

---

# VMS VAXcluster Manual

Order Number: AA-LA27C-TE

**November 1991**

This manual describes procedures for configuring and managing VAXcluster systems.

**Revision/Update Information:** This manual supersedes the *VMS VAXcluster Manual*, Version 5.4.

**Software Version:** VMS Version 5.5

---

**November 1991**

The information in this document is subject to change without notice and should not be construed as a commitment by Digital Equipment Corporation. Digital Equipment Corporation assumes no responsibility for any errors that may appear in this document.

The software described in this document is furnished under a license and may be used or copied only in accordance with the terms of such license.

No responsibility is assumed for the use or reliability of software on equipment that is not supplied by Digital Equipment Corporation or its affiliated companies.

© Digital Equipment Corporation 1991.

All Rights Reserved.

The postpaid Reader's Comments forms at the end of this document request your critical evaluation to assist in preparing future documentation.

The following are trademarks of Digital Equipment Corporation: CDA, DEC, DECdtm, DECnet, DECrouter, DECserver, DECUS, DECwindows, DECwrite, DEQNA, Digital, GIGI, HSC, LAT, LiveLink, LN03, MASSBUS, MicroVAX, PrintServer, Q-bus, RA, ReGIS, RSX, ThinWire, TK, TU, ULTRIX, UNIBUS, VAX, VAXcluster, VAX RMS, VAXserver, VAXstation, VMS, XUI, and the DIGITAL logo.

PostScript is a registered trademark of Adobe Systems Incorporated.

ZK4477

This document was prepared using VAX DOCUMENT, Version 1.2

---

# Contents

<b>Preface</b> .....	xi
<b>1 Introduction to VAXcluster Systems</b>	
1.1 Shared Resources .....	1-1
1.1.1 Disk and Tape Storage .....	1-2
1.1.2 Batch and Print Job Processing .....	1-2
1.2 Interconnect Devices .....	1-3
1.3 Software Components .....	1-4
1.4 Configuration Planning .....	1-5
<b>2 VAXcluster Interconnect Configurations</b>	
2.1 Configuration Types .....	2-1
2.2 CI-Based VAXcluster Systems .....	2-1
2.3 Local Area VAXcluster Systems .....	2-2
2.4 Mixed-Interconnect VAXcluster Systems .....	2-4
2.5 DSSI-Based VAXcluster Configurations .....	2-5
2.6 FDDI VAXcluster Systems .....	2-6
2.6.1 Multilobe FDDI VAXcluster System .....	2-7
2.6.2 Network Configurations Using FDDI as VAXcluster Interconnect .....	2-8
2.7 Configuring Multiple CI Adapters .....	2-9
2.8 Configuring Multiple DSSI Adapters .....	2-10
2.9 Configuring Multiple LAN Adapters .....	2-10
2.9.1 VAXcluster with Two LAN Segments .....	2-11
2.9.2 VAXcluster with Three LAN Segments .....	2-12
2.9.3 Guidelines to Allow for LAN Bridge Failover .....	2-13
2.9.4 Adjusting LRPSIZE for FDDI Configurations .....	2-14
2.10 Security for Local Area and Mixed-Interconnect VAXcluster Systems .....	2-15
<b>3 VAXcluster Integrity and Availability Features</b>	
3.1 Connection Management .....	3-1
3.1.1 The Quorum Scheme .....	3-1
3.1.2 Quorum Disk .....	3-2
3.1.3 State Transitions .....	3-3
3.2 VMS Volume Shadowing .....	3-5
3.2.1 Volume Shadowing Concepts .....	3-5
3.2.2 Distributing Shadow Sets Across VMS Configurations .....	3-7

## 4 Preparing the Cluster Operating Environment

4.1	Directory Structure on a Common System Disk . . . . .	4-2
4.2	Installing the VMS Operating System in the VAXcluster Environment . . .	4-4
4.3	Configuring and Starting the DECnet-VAX Network . . . . .	4-6
4.3.1	Copying Remote Node Databases . . . . .	4-8
4.3.2	Enabling VAXcluster Alias Operations . . . . .	4-9
4.4	Coordinating Startup Command Procedures . . . . .	4-9
4.4.1	Building Startup Procedures for a Common-Environment Cluster . . . .	4-10
4.4.1.1	Procedures for Existing Computers . . . . .	4-10
4.4.1.2	Procedures for Newly Installed Computers . . . . .	4-10
4.4.2	Building Startup Procedures for a Multiple-Environment Cluster . . . .	4-11
4.5	Coordinating System Files for a Common-Environment Cluster . . . . .	4-11
4.5.1	Coordinating User Accounts . . . . .	4-12
4.5.2	Preparing the Rights Database . . . . .	4-13
4.5.3	Preparing the MAIL Database . . . . .	4-14
4.5.4	Coordinating Shared System Files in Clusters with Multiple Common System Disks . . . . .	4-14

## 5 Setting Up and Managing Cluster Disks and Tapes

5.1	Cluster-Accessible Disk and Tape Devices . . . . .	5-1
5.1.1	HSC Disks and Tapes . . . . .	5-2
5.1.2	Served Disks and Tapes . . . . .	5-2
5.1.2.1	MSCP and TMSCP Server Functions . . . . .	5-3
5.1.2.2	MSCP Load Sharing . . . . .	5-4
5.1.3	Dual-Pathed Disks and Tapes . . . . .	5-4
5.1.3.1	Dual-Ported HSC Disks and Tapes . . . . .	5-4
5.1.3.2	Dual-Pathed DSA Disks and Tapes on Local UDA/KDA/KDB Controllers . . . . .	5-5
5.1.3.3	Preferred Path Specification . . . . .	5-5
5.1.3.4	DSSI-Connected ISEs . . . . .	5-6
5.1.3.5	Dual-Ported MASSBUS Disks . . . . .	5-6
5.2	Cluster Device-Naming Conventions . . . . .	5-7
5.2.1	Rules for Specifying Allocation Class Values . . . . .	5-7
5.2.2	Sample Configurations with Named Devices . . . . .	5-9
5.3	Shared Disks . . . . .	5-12
5.4	Configuring Cluster Disks . . . . .	5-12
5.5	Rebuilding Cluster Disks . . . . .	5-13

## 6 Setting Up and Managing Cluster Queues

6.1	Clusterwide Queues . . . . .	6-1
6.2	Cluster Printer Queues . . . . .	6-3
6.2.1	Setting Up Printer Queues . . . . .	6-4
6.2.2	Setting Up Clusterwide Generic Printer Queues . . . . .	6-4
6.3	Cluster Batch Queues . . . . .	6-5
6.3.1	Setting Up Execution Batch Queues . . . . .	6-7
6.3.2	Setting Up Clusterwide Generic Batch Queues . . . . .	6-8
6.4	Using a Common Command Procedure to Set Up Cluster Queues . . . . .	6-8
6.5	Starting Local Batch Queues . . . . .	6-13

## 7 Building and Maintaining the Cluster

7.1	CLUSTER_CONFIG.COM Functions	7-1
7.2	Determining Locations and Sizes for Satellite Page and Swap Files	7-3
7.3	Selecting MOP and Disk Servers	7-3
7.4	Determining Allocation Class Values in Mixed-Interconnect Clusters	7-4
7.5	Configuring the Cluster	7-5
7.5.1	Adding a Computer to the Cluster	7-6
7.5.1.1	Updating Network Data After Adding a Satellite	7-11
7.5.1.2	Restoring a Satellite's Network Data	7-11
7.5.1.3	Controlling Clusterwide Broadcast Messages on Satellites and Boot Servers	7-12
7.5.2	Removing a Computer from the Cluster	7-12
7.5.3	Changing a Computer's Characteristics	7-13
7.5.4	Changing the Cluster Configuration Type	7-18
7.5.4.1	Changing an Existing CI-Only Cluster to a Mixed-Interconnect Configuration	7-18
7.5.4.2	Changing an Existing Local Area Cluster to a Mixed-Interconnect Configuration	7-19
7.5.5	Converting a Standalone Computer to a VAXcluster Computer	7-19
7.5.6	Creating a Duplicate System Disk	7-20
7.6	Reconfiguring the Cluster After a Major Change	7-22
7.6.1	Updating MODPARAMS.DAT Files to Adjust Cluster Quorum	7-22
7.6.2	Shutting Down the Cluster	7-22
7.6.3	Changing Allocation Class Values on HSC Subsystems	7-23
7.6.4	Changing Allocation Class Values on DSSI Subsystems	7-23
7.6.5	Rebooting the Cluster	7-23
7.7	Maintaining the Cluster	7-24
7.7.1	Running AUTOGEN with the FEEDBACK Option	7-24
7.7.2	Recording Configuration Data	7-25
7.7.3	Monitoring LAN Activity in Local Area and Mixed-Interconnect Clusters	7-26
7.7.4	Performing VAXcluster Network Failure Analysis	7-26
7.7.5	Restoring Cluster Quorum After an Unexpected Computer Failure	7-26
7.7.6	Selecting Cluster Shutdown Options	7-28
7.7.6.1	The REMOVE_NODE Option	7-28
7.7.6.2	The CLUSTER_SHUTDOWN Option	7-28
7.7.6.3	The REBOOT_CHECK Option	7-28
7.7.6.4	The SAVE_FEEDBACK Option	7-29
7.7.7	Rebooting a Satellite with an Operating System on a Local Disk	7-29
7.7.8	Performing Security Functions in Local Area and Mixed-Interconnect Clusters	7-29
7.7.8.1	Maintaining Cluster Security Data	7-30
7.7.8.2	Controlling Conversational Bootstrap Operations for Satellites	7-31
7.8	Guidelines for Configuring Large Clusters	7-32
7.8.1	Local Area VAXcluster Satellite Booting	7-32
7.8.1.1	Alternate Adapter Booting	7-32
7.8.1.2	Changing the LAN Address in the DECnet Database to Allow a Cluster Satellite to Boot with Any Adapter	7-33
7.8.1.3	Displaying Connection Messages During Cluster Satellite Booting	7-33
7.8.1.4	Configuring MOP Service on DECnet Routers	7-34
7.8.2	Configuring Disk Server LAN Adapters and Memory	7-34

7.8.3	Configuring System Disks . . . . .	7-34
7.8.3.1	Concurrent User Activity . . . . .	7-35
7.8.3.2	Concurrent Booting Activity . . . . .	7-35
7.8.3.3	Boot Time Costs . . . . .	7-37
7.8.3.4	Moving High-Activity Files off System Disks . . . . .	7-37
7.8.3.5	Controlling Dump File Size and Creation . . . . .	7-37
7.8.3.6	Sharing Dump Files . . . . .	7-38
7.8.4	Adding Computers to an Existing Cluster . . . . .	7-39
7.8.4.1	Running AUTOGEN with FEEDBACK for Initial Configuration . . . . .	7-39
7.8.4.2	Creating a Command File to Run AUTOGEN with FEEDBACK . . . . .	7-40
7.8.5	Setting Up a New Large VAXcluster System . . . . .	7-41
7.8.6	Defining the VAXcluster Alias . . . . .	7-42

## A Cluster SYSGEN Parameters

## B Building a Common SYSUAF.DAT File

## C Cluster Troubleshooting

C.1	Diagnosing Failures of Computers to Boot or Join the Cluster . . . . .	C-1
C.1.1	Events for Computers Booting and Joining the Cluster . . . . .	C-1
C.1.2	CI-Connected Computer Fails to Boot . . . . .	C-3
C.1.3	Satellite Fails to Boot . . . . .	C-4
C.1.4	Computer Fails to Join the Cluster . . . . .	C-6
C.1.5	Startup Procedures Fail to Complete . . . . .	C-6
C.1.6	Diagnosing LAN Component Failures . . . . .	C-7
C.2	Diagnosing Cluster Hangs . . . . .	C-7
C.2.1	Cluster Quorum Is Lost . . . . .	C-7
C.2.2	A Shared Cluster Resource Is Inaccessible . . . . .	C-7
C.3	Diagnosing CLUEXIT Bugchecks . . . . .	C-8
C.4	Diagnosing VAXport Device Problems . . . . .	C-9
C.4.1	VAXport Communication Mechanisms . . . . .	C-9
C.4.2	Port Failures . . . . .	C-10
C.4.2.1	Verifying CI Port Functions . . . . .	C-11
C.4.2.2	Verifying CI Cable Connections . . . . .	C-12
C.4.2.3	Repairing CI Cables . . . . .	C-14
C.4.2.4	Verifying LAN Connections . . . . .	C-15
C.4.3	Analyzing Error Log Entries for VAXport Devices . . . . .	C-15
C.4.3.1	Error Log Entry Formats . . . . .	C-16
C.4.3.2	Device-Attention Entries . . . . .	C-16
C.4.3.3	Logged-Message Entries . . . . .	C-19
C.4.3.4	Error Log Entry Descriptions . . . . .	C-21
C.4.4	OPA0 Error Messages . . . . .	C-28

## D Local Area VAXcluster Network Connections

## E Local Area VAXcluster Sample Programs

E.1	Sample Programs for Local Area VAXclusters . . . . .	E-1
E.1.1	Starting the Local Area VAXcluster Protocol on a LAN Adapter . . . . .	E-1
E.1.2	Stopping the Local Area VAXcluster Protocol on a LAN Adapter . . . . .	E-2
E.1.3	Enabling VAXcluster Network Failure Analysis . . . . .	E-3
E.2	Using the Local Area VAXcluster Network Failure Analysis Program . . . . .	E-3
E.2.1	Collecting Information for the Network Failure Analysis Program . . . . .	E-3
E.2.2	Editing the Network Failure Analysis Program . . . . .	E-5
E.2.3	Assembling and Linking the Failure Analysis Program . . . . .	E-8
E.2.4	Executing the Network Failure Analysis Program . . . . .	E-8
E.2.5	Testing the Network Failure Analysis Subsystem . . . . .	E-9
E.2.6	PEDRIVER Suspect Network Component Display . . . . .	E-9

## F Local Area VAXcluster Subroutine Package

F.1	Subroutine Package to Start the Protocol on a LAN Adapter . . . . .	F-1
F.2	Subroutine Package to Stop the Protocol on a LAN Adapter . . . . .	F-3
F.3	Subroutine Package to Create a Network Component . . . . .	F-4
F.4	Subroutine Package to Create a Network Component List . . . . .	F-5
F.5	Subroutine Package to Start the Network Component Failure Analysis . . . . .	F-7
F.6	Subroutine Package to Stop the Network Component Failure Analysis . . . . .	F-7

## G Transmit Channel Selection

### Index

### Examples

4-1	Sample Interactive Network Configuration Session . . . . .	4-7
6-1	Sample Commands for Creating VAXcluster Queues . . . . .	6-10
6-2	Common Procedure to Start VAXcluster Queues . . . . .	6-11
7-1	Sample Interactive CLUSTER_CONFIG.COM Session to Add a CI-Connected Computer as a Boot Server . . . . .	7-7
7-2	Sample Interactive CLUSTER_CONFIG.COM Session to Add a Satellite with Local Page and Swap Files . . . . .	7-9
7-3	Sample NETNODE_UPDATE.COM File . . . . .	7-12
7-4	Sample Interactive CLUSTER_CONFIG.COM Session to Remove a Satellite with Local Page and Swap Files . . . . .	7-13
7-5	Sample Interactive CLUSTER_CONFIG.COM Session to Enable the Local Computer as a Disk Server . . . . .	7-15
7-6	Sample Interactive CLUSTER_CONFIG.COM Session to Change the Local Computer's ALLOCLASS Value . . . . .	7-16
7-7	Sample Interactive CLUSTER_CONFIG.COM Session to Enable the Local Computer as a Boot Server . . . . .	7-17
7-8	Sample Interactive CLUSTER_CONFIG.COM Session to Change a Satellite's Hardware Address . . . . .	7-18

7-9	Sample Interactive CLUSTER_CONFIG.COM Session to Convert a Standalone Computer to a Cluster Boot Server . . . . .	7-20
7-10	Sample Interactive CLUSTER_CONFIG.COM CREATE Session . . . . .	7-21
7-11	Sample SYSMAN Session to Change the Cluster Password . . . . .	7-31
C-1	CI Device-Attention Entry . . . . .	C-16
C-2	LAN Device-Attention Entry . . . . .	C-18
C-3	CI Logged-Message Entry . . . . .	C-19
E-1	Portion of LAVC\$FAILURE_ANALYSIS.MAR to Edit . . . . .	E-5

## Figures

2-1	Typical CI-Based VAXcluster Configuration . . . . .	2-2
2-2	Local Area VAXcluster System with Single Boot Server . . . . .	2-3
2-3	VAXcluster Configuration with Redundant Boot Servers and DSA Disks . . . . .	2-4
2-4	Typical Mixed-Interconnect VAXcluster System . . . . .	2-5
2-5	DSSI VAXcluster Configuration . . . . .	2-6
2-6	FDDI VAXcluster System: Example Configuration . . . . .	2-7
2-7	FDDI in Conjunction with Ethernet in a VAXcluster System . . . . .	2-8
2-8	Multi-Datacenter VAXcluster System . . . . .	2-9
2-9	Sample Two-LAN Segment VAXcluster Configuration . . . . .	2-12
2-10	Sample Three-LAN Segment VAXcluster Configuration . . . . .	2-13
3-1	Elements of a Shadow Set . . . . .	3-6
3-2	Shadow Set Members Accessed Through the VMS MSCP Server . . . . .	3-8
4-1	Directory Structure on Common System Disk . . . . .	4-2
4-2	File Search Order on Common System Disk . . . . .	4-3
5-1	CI-Based Configuration with Shared Disks and Tapes . . . . .	5-2
5-2	VAXcluster with Dual-Pathed HSC Disk and Tape . . . . .	5-9
5-3	VAXcluster with Dual-Pathed DSA Disk and Tape . . . . .	5-10
5-4	Device Names in a Mixed-Interconnect Cluster . . . . .	5-11
6-1	Sample Printer Configuration . . . . .	6-3
6-2	Printer Queue Configuration . . . . .	6-5
6-3	Clusterwide Generic Printer Queue Configuration . . . . .	6-6
6-4	Sample Batch Queue Configuration . . . . .	6-7
6-5	Clusterwide Generic Batch Queue Configuration . . . . .	6-9
C-1	A Correctly Connected Two-Computer CI Cluster . . . . .	C-12
C-2	Crossed CI Cable Pair . . . . .	C-13
D-1	Sample Transceiver/Thick Wire Ethernet Connection . . . . .	D-1
D-2	Sample DELNI Connection . . . . .	D-2
D-3	Sample DESTA Connection . . . . .	D-2
D-4	Sample Connection to an Existing Thin Wire Segment . . . . .	D-3
D-5	Sample New Connection to an Existing DEMPR . . . . .	D-3



## Tables

4-1	Information Requested for CI-Based Configurations .....	4-5
4-2	Information Requested for Local Area and Mixed-Interconnect Configurations .....	4-5
5-1	Specifying Values for MSCP_LOAD, MSCP_SERVE_ALL, and TMSCP_LOAD Parameters .....	5-3
7-1	Summary of CLUSTER_CONFIG.COM Functions .....	7-2
7-2	Data Requested by CLUSTER_CONFIG.COM .....	7-5
7-3	CLUSTER_CONFIG.COM CHANGE Options .....	7-14
7-4	Summary of SYSMAN CONFIGURATION Commands for Cluster Authorization .....	7-30
7-5	System Disk I/O Activity and Boot Time for Single Satellite .....	7-36
7-6	System Disk I/O Activity and Boot Times for Multiple Satellites .....	7-36
7-7	AUTOGEN Dump File Symbols .....	7-38
A-1	Cluster SYSGEN Parameters .....	A-1



---

# Preface

## Intended Audience

This document addresses persons responsible for setting up and managing VAXcluster systems. To use the document as a guide to cluster management, you must have a thorough understanding of VMS system management concepts and procedures, as described in the *Introduction to VMS System Management*, the *Guide to Setting Up a VMS System*, and the *Guide to Maintaining a VMS System*.

## Document Structure

The *VMS VAXcluster Manual* contains seven chapters and seven appendixes.

Chapter 1 is an introduction to VAXcluster systems.

Chapter 2 describes the ways to interconnect VAXcluster CPUs.

Chapter 3 presents the software concepts that maintain VAXcluster integrity, and describes Digital software that provides for data availability.

Chapter 4 explains how to prepare the cluster operating environment.

Chapter 5 discusses disk and tape management concepts and procedures.

Chapter 6 discusses queue management concepts and procedures.

Chapter 7 explains how to build a VAXcluster system once the necessary preparations are made, and how to reconfigure and maintain the cluster.

Appendix A lists and defines cluster SYSGEN parameters.

Appendix B provides guidelines for building a cluster common user authorization file.

Appendix C provides VAXcluster troubleshooting information.

Appendix D provides information on local area VAXcluster network connections.

Appendix E describes three sample programs for local area VAXclusters and explains how to use the Local Area Network Failure Analysis Program.

Appendix F describes the subroutine package used with local area VAXcluster sample programs.

Appendix G discusses transmit channel selection by PEDRIVER.

## Associated Documents

This document is not a one-volume reference manual. The VMS utilities and commands discussed are described in detail in separate VMS Utility Reference Manuals and in the *VMS DCL Dictionary*.

For additional information on the topics covered in this manual, refer to the following VMS documents:

- *Introduction to VMS System Management*
- *Guide to Setting Up a VMS System*
- *Guide to Maintaining a VMS System*
- *Guide to VMS File Applications*
- *VMS I/O User's Reference Manual: Part I*
- *VMS System Services Volume*
- *VMS System Generation Utility Manual*
- *VMS Networking Manual*
- *VMS Network Control Program Manual*
- VMS Utility Reference Manuals
- *VMS License Management Utility Manual*
- *Building Dependable Systems: The VMS Approach*
- *VMS Volume Shadowing Manual* (optional manual)

Another related document is the *Guidelines to VAXcluster System Configurations*.

## Conventions

The following conventions are used in this manual:

Ctrl/x

A sequence such as Ctrl/x indicates that you must hold down the key labeled Ctrl while you press another key or a pointing device button.

Return

In examples, a key name is shown enclosed in a box to indicate that you press a key on the keyboard. (In text, a key name is not enclosed in a box.)

...

In examples, a horizontal ellipsis indicates one of the following possibilities:

- Additional optional arguments in a statement have been omitted.
- The preceding item or items can be repeated one or more times.
- Additional parameters, values, or other information can be entered.

.  
. .  
. . .

A vertical ellipsis indicates the omission of items from a code example or command format; the items are omitted because they are not important to the topic being discussed.

()	In format descriptions, parentheses indicate that, if you choose more than one option, you must enclose the choices in parentheses.
[]	In format descriptions, brackets indicate that whatever is enclosed within the brackets is optional; you can select none, one, or all of the choices. (Brackets are not, however, optional in the syntax of a directory name in a file specification or in the syntax of a substring specification in an assignment statement.)
red ink	Red ink indicates information that you must enter from the keyboard or a screen object that you must choose or click on.  For online versions of the book, user input is shown in <b>bold</b> .
<b>boldface text</b>	Boldface text represents the introduction of a new term or the name of an argument, an attribute, or a reason.  Boldface text is also used to show user input in online versions of the book.
<i>italic text</i>	Italic text represents information that can vary in system messages (for example, Internal error <i>number</i> ).
UPPERCASE TEXT	Uppercase letters indicate that you must enter a command (for example, enter OPEN/READ), or they indicate the name of a routine, the name of a file, the name of a file protection code, or the abbreviation for a system privilege.
-	Hyphens in coding examples indicate that additional arguments to the request are provided on the line that follows.
numbers	Unless otherwise noted, all numbers in the text are assumed to be decimal. Nondecimal radices—binary, octal, or hexadecimal—are explicitly indicated.



---

# Introduction to VAXcluster Systems

A VAXcluster system is a highly integrated organization of VAX computers. As members of a VAXcluster system, these computers can share processing resources, data storage, and queues under a single VMS security and management domain, and they can boot or fail independently.

Using procedures described in Chapter 4, system managers can tailor the cluster operating environment to create a **common-environment** or a **multiple-environment** VAXcluster system.

- In a common-environment VAXcluster system, the same resources are available on all computers. User accounts are identical, the same known images are installed, the same logical names are defined, and mass storage devices and queues are shared.
- In a multiple-environment VAXcluster system, a group of computers shares one set of resources, while another group shares a different set. Alternatively, an individual computer can perform a specialized function using restricted resources, while other computers perform general timesharing work.

Although most cluster resources can be shared, user processes and memory are computer specific. When a process is created on a VAXcluster computer, the process must complete on that computer, using local memory. If the computer fails before the process completes, the process is terminated. However, users can recover from such a failure more quickly than on a standalone computer, because they need not wait until the computer is rebooted. Typically, they can log in on another VAXcluster computer to create a new process and continue working, provided that the resources required by the process (such as images and global sections) are available on that computer.

This chapter describes VAXcluster operating features and components, including the following:

- Shared resources
- Interconnect devices
- Software components
- Configuration planning

Be sure you understand these topics before you attempt to perform any cluster setup operations.

## 1.1 Shared Resources

In any VAXcluster system, users can share computing, disk storage, and batch and print job processing resources. The ability to share resources facilitates workload balancing, because work can be distributed across the cluster. To keep pace with user demand, resources can be added without disrupting normal cluster operation.

# Introduction to VAXcluster Systems

## 1.1 Shared Resources

### 1.1.1 Disk and Tape Storage

A major advantage of VAXcluster systems is the ability to make disk and tape storage resources accessible to all VAXcluster computers. Storage devices such as DIGITAL Storage Architecture (DSA) disks and tapes, RF-series **Integrated Storage Elements (ISEs)**, and ESE20 solid state disks can be configured for local or clusterwide access. A **cluster-accessible** storage device can be used directly by multiple computers in the cluster. By means of VMS MSCP and TMSCP server software, disks and tapes can be made accessible to nodes that are not directly connected.

Cluster-accessible disks offer the following advantages:

- More efficient use of mass storage, because more than one computer can use the same disk.
- Access by users to their default work disks when logging in to any computer on which the disks are accessible.
- Clusterwide file sharing. Because computers can share common versions of files, updates to a file are made only once to a single copy of the file.
- Implementation of clusterwide job-controller queues. Batch and print jobs can be processed on any computer that has access to the disks.

Some VAXcluster systems include hierarchical storage controller (HSC) subsystems. These are self-contained, intelligent mass storage subsystems that enable VAXcluster computers to share DSA disks and, in the VAXcluster configurations described in Section 2.2, DSA tapes. Because the HSC subsystem is an intelligent controller, it optimizes physical disk and tape operations and supports many combinations of standard disk interfaces (SDIs) and standard tape interfaces (STIs), which connect disks and tapes. HSC disk configurations provide flexibility, expansion potential, maintenance and backup capability.

Disk data can be replicated between VAXcluster systems for better read performance and higher availability using VMS volume shadowing, as described in Section 3.2. For detailed information on volume shadowing, see the *VMS Volume Shadowing Manual*.

Procedures for setting up and managing cluster disks and tapes are described in Chapter 5.

### 1.1.2 Batch and Print Job Processing

System managers control how jobs share batch processing and printer resources by setting up and maintaining clusterwide generic queues. The strategy for setting up and managing these queues determines how well workloads are matched to available resources.

With VMS Version 5.5, queues are controlled by a clusterwide queue manager process that accesses the clusterwide queue database for all processes in a cluster. Job controllers, user processes, and symbionts all communicate directly with the centralized queue manager through a shared interprocess communication (IPC) link.

Procedures for setting up and managing cluster queues are described in Chapter 6.

(Note that a generic queue holds a job that will execute on an execution queue when it is available to process jobs. When the job is sent to an execution queue on a specific node, it is executed on that queue.)



## 1.2 Interconnect Devices

Interconnect devices used to configure a VAXcluster system include the following:

- **Computer Interconnect (CI).** This high-speed, dual-path interface links computers and HSC subsystems in a computer room environment. A Computer Interconnect consists of several components such as CI port controllers (adapters), the star coupler, star coupler expander (CISCE), and the high-bandwidth CI cables themselves.
- **CI port controllers.** Port controllers like the CI780, CIBCA, CIBCI, and CIXCD (CI to XMI adapter) are microcoded, intelligent adapters that connect computers to CI cables. Each interface connects to one pair of transmitter and one pair of receiver cables.

Under normal operating conditions, both pairs of cables are available to meet traffic demands. If one path fails, all traffic uses the remaining path. The VMS operating system periodically tests a failed path. As soon as a failed path is restored, it is automatically used for normal traffic.

- **Star couplers and star coupler expanders (CISCE).** Star couplers and star coupler expanders provide common connection points for CI-connected computers and HSC subsystems. Both coupler devices connect CI cables from computers and HSC subsystems, creating a radial or “star” arrangement that has a maximum radius of 45 meters (147 feet). These devices support the physical connection and disconnection of individual computers and HSC subsystems without affecting other computers or HSC subsystems.

The star coupler and CISCE are dual-pathed devices that contain separate components for each path. The star coupler is a passive device; the CISCE consists of redundant amplifiers. Both devices are designed so that all CI cables are transformer coupled and independent of earth ground reference. These attributes help to ensure signal integrity.

- **DIGITAL Storage System Interconnect (DSSI).** The DSSI bus permits multiple computers to communicate directly with storage devices. The DSSI bus connects as many as eight nodes that can be ISEs or host CPU interfaces.
- **DSSI port controllers.** Port controllers like the KFMSA, KFQSA, and EDA-type adapters are intelligent adapters that connect computers to DSSI busses.
- **Local area network (LAN) adapters.** LAN adapters include:
  - **Ethernet.** The Ethernet is a bus that uses digital baseband signaling.
  - **Fiber Distributed Data Interface (FDDI).** FDDI is an ANSI-standard LAN interconnect based on a fiber optic transmission medium.

The LAN is used both for DECnet–VAX transmissions and, in some VAXcluster systems, for cluster communications. The LAN must be configured according to requirements specified in the VAXcluster Software Product Description (SPD).

Depending on how a VAXcluster system is configured, VAXcluster software can use multiple interconnects for cluster communications.

## 1.3 Software Components

The software components used to implement VAXcluster communication and resource-sharing functions always run on each computer in the cluster. Thus, if one computer fails, the VAXcluster system continues operating, because the components still run on the remaining computers. These software components are as follows:

- **System Communications Services (SCS)** software implements intercomputer communication, according to the DIGITAL System Communications Architecture (SCA).
- **VAXport drivers** control the communication paths between local and remote ports. (Examples are PADRIVER for the CI, PEDRIVER for the LAN, and PIDRIVER for the DSSI.)
- The **connection manager** dynamically defines the VAXcluster system and coordinates participation of computers in the cluster. The connection manager uses SCS to provide an acknowledged message delivery service for higher VMS software layers. The connection manager also maintains cluster integrity when computers join or leave the cluster—that is, when cluster **state transitions** occur. (State transitions are discussed in Section 3.1.3.)
- The VAXcluster **distributed file system** allows all computers to share mass storage, whether the storage device is connected to an HSC subsystem or to a computer. A local disk can be made available to the entire cluster. All cluster-accessible disks appear as if they are local to every computer.

The distributed file system and VMS Record Management Services (VMS RMS) provide the same access to disks and files across the cluster that is provided on a standalone computer. VMS RMS files may be shared to the record level.

- The **distributed lock manager** is used for synchronization functions by the distributed file system, job controller, device allocation, and other cluster facilities. It is available to users for developing cluster applications. The distributed lock manager implements the \$ENQ and \$DEQ system services to provide clusterwide synchronization of access to resources by allowing the locking and unlocking of resource names. (For detailed information on system services, refer to the *VMS System Services Volume*.) The distributed lock manager also provides a queuing mechanism so that processes can be put into a wait state until a particular resource is available. As a result, cooperating processes can synchronize their access to shared objects such as files and records.

If a VAXcluster computer fails, all locks that it holds are released. This mechanism allows processing to continue on the remaining computers. The distributed lock manager also supports clusterwide deadlock detection.

- The **distributed job controller** makes queues available across the cluster. VAXcluster computers can share batch and print queues. Users can submit jobs to any queue in the cluster, provided that the necessary mass storage volumes and peripheral devices are accessible to the computer on which the job executes. System managers can also set up generic batch queues that distribute batch processing workloads among computers. For detailed information on VAXcluster queues, see Chapter 6.

- The **mass storage control protocol (MSCP) server** implements the MSCP protocol, which is used to communicate with a controller for DSA disks, such as RA-series disks. In conjunction with one or both of the disk class drivers (DUDRIVER, DSDRIVER), the MSCP server implements this protocol on a computer, allowing the computer to function as a storage controller. The computer submits I/O requests to locally accessed disks, such as locally connected RA disks or Small Computer System Interface (SCSI) disks, and accepts the I/O requests from any computer in the cluster. In this way, the MSCP server makes locally connected disks available across the cluster. In the local area and mixed-interconnect VAXcluster systems described in Section 2.3 through Section 2.6, the MSCP server can also make HSC disks and DSSI ISE disks accessible over the LAN.
- The **tape mass storage control protocol (TMSCP) server** implements the TMSCP protocol, which is used to communicate with a controller for local MSCP tapes, such as TU series tapes. In conjunction with the tape class driver (TUDRIVER), the TMSCP Server implements this protocol on a processor, allowing the processor to function as a storage controller. The processor submits I/O requests to locally accessed tapes, such as Q-bus tapes, and accepts the I/O requests from any node in the cluster. In this way, the TMSCP Server makes locally connected tapes available to all nodes in the cluster. The TMSCP Server can also make HSC tapes and DSSI ISE tapes accessible to VAXcluster satellites.

In addition to these components, all VAXcluster systems require **DECnet-VAX software**, which ensures that system managers can access all VAXcluster computers from a single terminal, even if terminal switching facilities are unavailable.

In local area and mixed-interconnect VAXcluster systems, DECnet-VAX software is required both for system management functions and remote booting operations.

In these systems, DECnet and SCS software coexist on the same extended local area network (LAN). They share the same data link and physical link protocols, which are implemented by the LAN data link drivers, the LAN adapters, and the LAN itself.

## 1.4 Configuration Planning

The process of setting up a VAXcluster system requires careful preparation. In planning your configuration, you must determine the following:

- **Interconnect type.** If you want to include VAX 3000-class computers or workstations in your VAXcluster system, you must set up a local area or mixed-interconnect configuration. These configurations are described in Chapter 2.
- **Integrity and availability.** Chapter 3 describes the rules used to maintain VAXcluster integrity, and the VMS Volume Shadowing product that provides for data availability.
- **Operating environment (common or multiple).** These environments are described at the beginning of this chapter and in Chapter 4, which provides information on configuring the DECnet-VAX network and on preparing the startup, user authorization, and other files that define the operating environment.

# Introduction to VAXcluster Systems

## 1.4 Configuration Planning

- **Disk and tape storage configuration.** Chapter 5 provides information about disk and tape storage configurations, including descriptions of disk types, rules for specifying disk and tape names in VAXcluster systems, and sample disk and tape configurations.
- **Queue configuration.** Chapter 6 provides information about VAXcluster queues and includes a sample queue setup command procedure.
- **Computer configuration.** Procedures for configuring VAXcluster computers are described in Chapter 7. That chapter also includes a detailed discussion of the cluster configuration command procedure, `SYS$MANAGER:CLUSTER_CONFIG.COM`.

Once you have planned your configuration, installed the necessary hardware, and checked hardware devices for proper operation, you can set up the cluster using various system software facilities. Setup procedures are typically as follows:

- Installing or upgrading the VMS operating system on the first VAXcluster computer. Follow instructions in the installation and operations guide for your computer.
- Installing required software licenses. Follow instructions in the *VMS License Management Utility Manual*.
- Configuring and starting the DECnet-VAX network. Follow instructions in Chapter 4. For more detailed information on network operations, refer to the *VMS Networking Manual*.
- Preparing files that define the cluster operating environment and that control disk and queue operations. Follow instructions in Chapters 4, 5, and 6.
- Adding computers to the cluster. Follow instructions in Chapter 7.

Depending on various factors, the order in which these operations are performed can vary from site to site, as well as from cluster to cluster at the same site.

---

# VAXcluster Interconnect Configurations

This chapter discusses the various types of VAXcluster configurations and the ways they are interconnected.

## 2.1 Configuration Types

While processing needs and available hardware resources must determine how individual VAXcluster systems are configured, sites can choose from the following configuration types:

- CI-based VAXcluster systems
- Local area VAXcluster systems
- Mixed-interconnect VAXcluster systems
- DSSI-based VAXcluster systems

These configuration types are described in Section 2.2, Section 2.3, Section 2.4, and Section 2.5, respectively. Configurations that include FDDI connections are described in Section 2.6. For complete information on currently supported configurations, refer to the VAXcluster SPD.

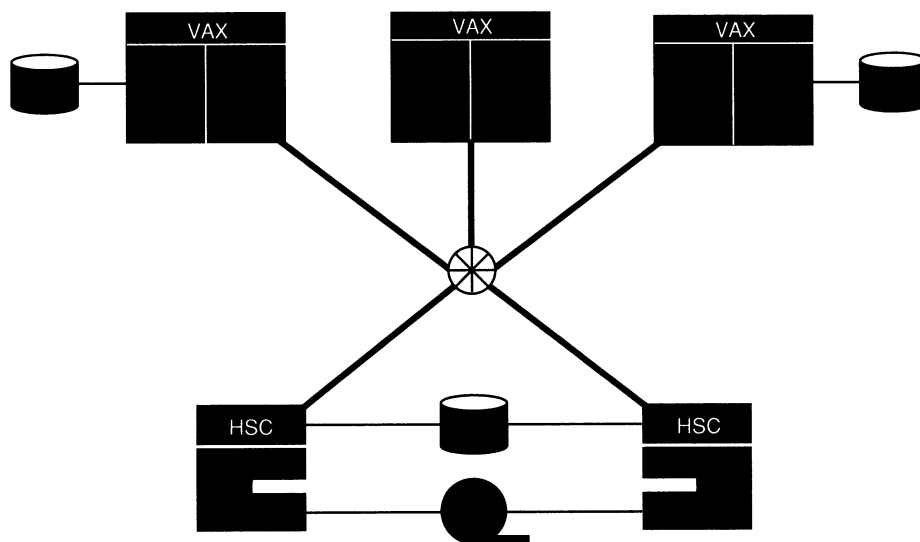
## 2.2 CI-Based VAXcluster Systems

A CI-based VAXcluster system uses CI components for cluster communications, with star couplers or CISCEs as common connection points for computers and HSC subsystems. Figure 2-1 shows how the components are typically configured. Note that if you want to add workstations to a CI-based VAXcluster system, you must convert it to a mixed-interconnect system. Refer to Section 2.4 for a description of mixed-interconnect systems and to Section 7.5.4 for conversion instructions.

## VAXcluster Interconnect Configurations

### 2.2 CI-Based VAXcluster Systems

Figure 2-1 Typical CI-Based VAXcluster Configuration



ZK-1640-GE

### 2.3 Local Area VAXcluster Systems

In a local area VAXcluster system, cluster communications are carried out over the LAN by a VAXport driver that emulates certain CI port functions. Because HSC subsystems require CI connections, local area VAXcluster systems do not include HSC subsystems.

A single LAN can support multiple local area VAXcluster systems, each system identified and secured by a unique **group number** and a **cluster password**. (For information on cluster security, see Section 2.10.)

Computers in a local area cluster are generally configured either as **servers** (MOP servers, disk servers, and tape servers) or as **satellites** (computers without a local VMS system disk). Using MSCP and TMSCP server software, the servers make their locally connected disks and tapes available to satellites over the LAN.

**MOP servers** downline load the VMS operating system to satellites by means of the DECnet Maintenance Operation Protocol (MOP). When a satellite requests an operating system load, a MOP server sends an image to the satellite that allows the satellite to load the VMS operating system from a disk server and join the cluster. VAXcluster satellites can be configured as additional MOP servers.

**Boot servers** are a combination of a MOP server and a disk server that serves the system disk. Boot servers must always run the MSCP server software.

Typically, a boot server is both a management center for the cluster and a major resource provider. Its system disk contains the cluster common files for startup, authorization, and queue setup, as well as the root directories from which the satellites are booted and in which their specific system files reside. These root directories, one for each satellite, are created when system managers add satellites to the cluster using the CLUSTER\_CONFIG.COM command procedure described in Chapter 7.

## VAXcluster Interconnect Configurations

### 2.3 Local Area VAXcluster Systems

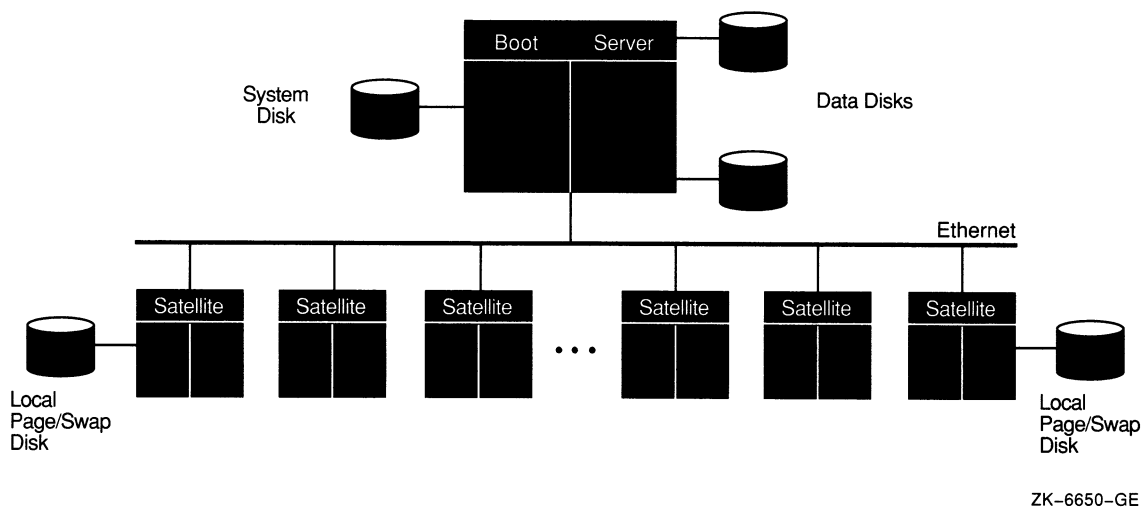
Boot and disk servers make user and application data disks available across the cluster. These servers should be the most powerful computers in the cluster and should use the highest bandwidth LAN adapters in the cluster.

Satellites are booted remotely from a boot server or a MOP server and a disk server serving the system disk. Generally, satellites are consumers of cluster resources, though they can also sometimes provide facilities for disk serving, MOP serving, tape serving, and batch processing. If satellites are equipped with local disks, they can enhance performance by using such local disks for paging and swapping.

Any local area VAXcluster system can be converted to a mixed-interconnect system. Refer to Section 7.5.4 for instructions.

Figure 2-2 shows a local area VAXcluster system with a single boot server. Note that because all computers in this configuration rely on the boot server's system disk, the boot server (or its system disk) is a single point of failure.

**Figure 2-2 Local Area VAXcluster System with Single Boot Server**



You can increase the availability of the VAXcluster system by using two boot servers with locally connected system disks. A satellite system disk can be dual ported between the servers.

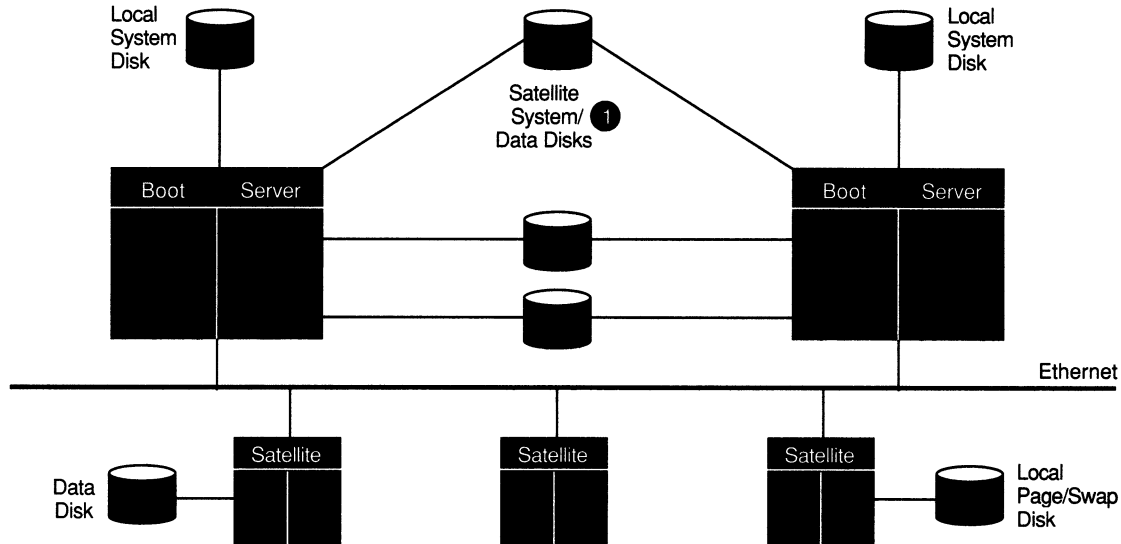
Figure 2-3 shows such a configuration using DSA disks. The boot servers boot from their locally connected system disks, and the satellites boot from the satellite system disk, which contains their root directories. Thus, if one boot server fails, satellites can access their system disk through the other server. If the satellite system disk fails, it can be restored from the remaining boot server, and the satellites can then be rebooted.

To increase availability ever further, use a DSSI configuration (see Section 2.5).

## VAXcluster Interconnect Configurations

### 2.3 Local Area VAXcluster Systems

Figure 2-3 VAXcluster Configuration with Redundant Boot Servers and DSA Disks



1 System Disk for Booting Satellites

ZK-1877A-GE

## 2.4 Mixed-Interconnect VAXcluster Systems

A mixed-interconnect VAXcluster system is any VAXcluster system that utilizes more than one interconnect for SCA traffic. Examples of mixed interconnect VAXcluster configurations are:

- A CI-based VAXcluster with satellite workstations on the LAN
- A DSSI-based VAXcluster with satellite workstations on the LAN
- A CI-based VAXcluster where one of the VAX nodes connects to another VAX node via the DSSI
- Any distributed LAN-based VAXcluster

(DSSI-based configurations are described in Section 2.5 and configurations that use FDDI are discussed in Section 2.6.)

CI-connected computers can serve HSC disks to satellites by means of MSCP server software and drivers; therefore, the satellites can access the large amount of storage that is available through HSC subsystems.

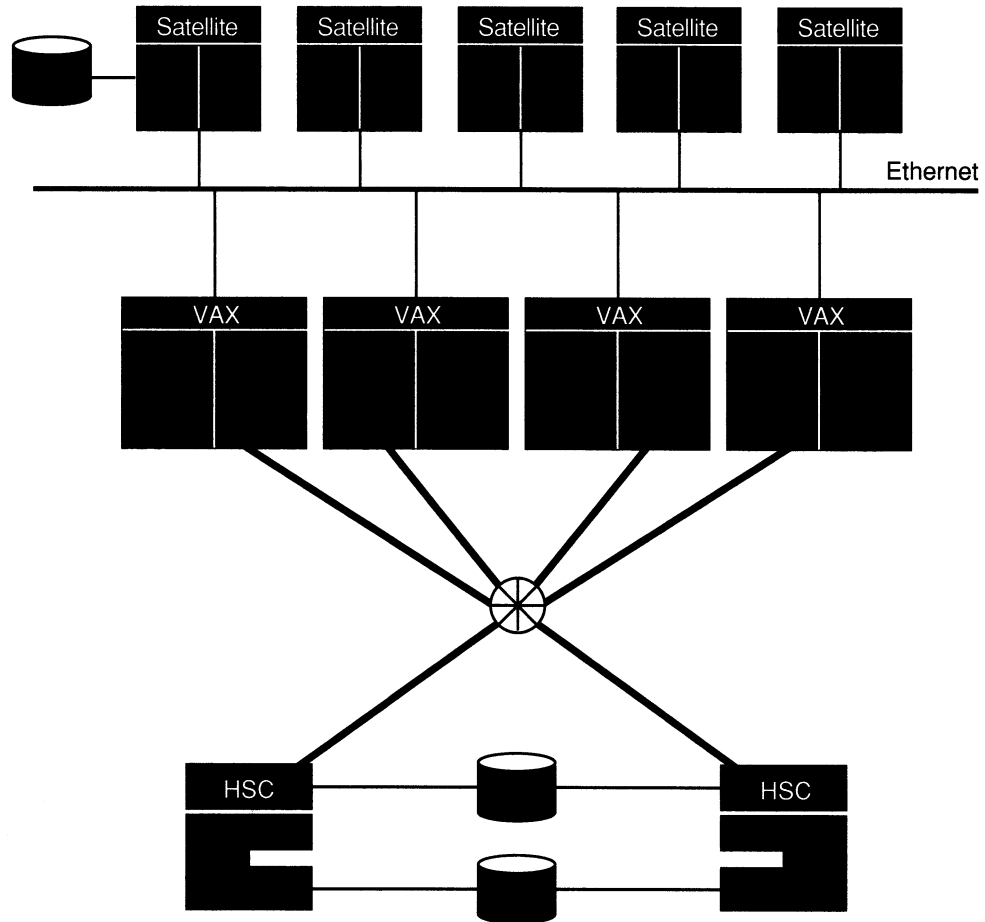
Mixed-interconnect VAXcluster systems provide maximum flexibility in combining CPUs, storage, and workstations into highly available configurations.

Figure 2-4 shows a typical mixed-interconnect VAXcluster system.



## VAXcluster Interconnect Configurations 2.5 DSSI-Based VAXcluster Configurations

Figure 2-4 Typical Mixed-Interconnect VAXcluster System



ZK-6659-GE

### 2.5 DSSI-Based VAXcluster Configurations

The DSSI is an interconnect that VAX systems can use to access disk and tape peripherals. Each peripheral is an ISE that contains its own controller and its own MSCP server that works in parallel with the other ISEs on the DSSI.

Because VAX nodes have direct access to the system disks, DSSI VAXcluster systems can support the VMS rolling upgrade feature. This capability allows the version of VMS to be upgraded on one node and that node to be rebooted while the other node is still providing service.

Generic guidelines for DSSI VAXcluster systems are as follows:

- Currently, three VAX nodes can be connected to a common DSSI.
- Multiple DSSI buses can operate in a VAXcluster configuration. The maximum number of VAX nodes in a VAXcluster configuration is not increased by having multiple DSSI buses. However, the maximum number of ISE controllers is dramatically increased, allowing much more storage to be configured into the system.

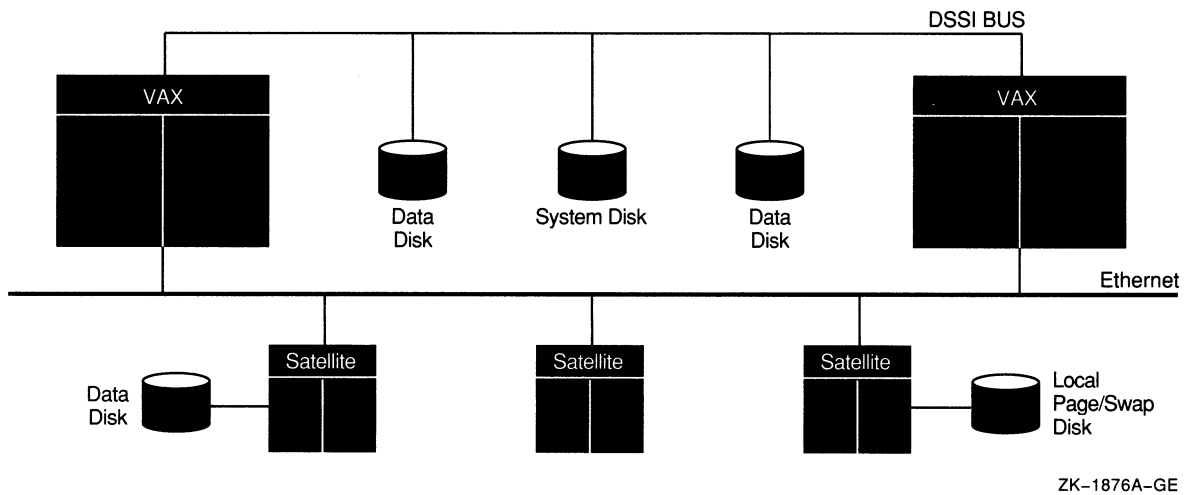
## VAXcluster Interconnect Configurations

### 2.5 DSSI-Based VAXcluster Configurations

- Some restrictions apply to the type of CPUs and DSSI I/O adapters that can reside on the same DSSI bus. See the appropriate VAXcluster configuration manual and the DSSI installation manual for the specific CPU or adapter type.
- Consult your Digital Services support person for more complete and timely configuration details on DSSI VAXcluster systems.

In the DSSI configuration shown in Figure 2-5, the two boot servers and all satellites boot from a common system disk.

Figure 2-5 DSSI VAXcluster Configuration



In addition to the two boot servers, a DSSI VAXcluster configuration typically includes several DSSI-connected ISEs and either a TK70 tape or the new DSSI TF85 series tape serving the configuration. Using additional DSSI cables, you can connect either one or two **storage expansion boxes** containing additional ISEs. If you decide to use a storage expansion box, it is a good idea to place a common system disk and critical data disks in the expansion box, which has a dedicated power supply. Thus, if one boot server fails, the other server and satellites can still access the disks.

Note that there are specific rules for various DSSI configurations; that is, the setup for a DSSI MicroVAX 3800 system is different from that for a DSSI VAX 4000 system. These rules are provided in the appropriate installation and user manuals.

### 2.6 FDDI VAXcluster Systems

The Fiber Distributed Data Interface (FDDI) is an interconnect that utilizes fiber optic cables for very high bandwidth over long distances. FDDI uses a token-ring topology with two counter-rotating rings. The available bandwidth depends on the target token rotation time (TTRT) and the time required to send the token around an idle ring (ring latency). The ring latency depends on the FDDI configuration. Providing alternative cluster communications paths will help cluster performance because the performance is impacted by increases in ring latency.

# VAXcluster Interconnect Configurations

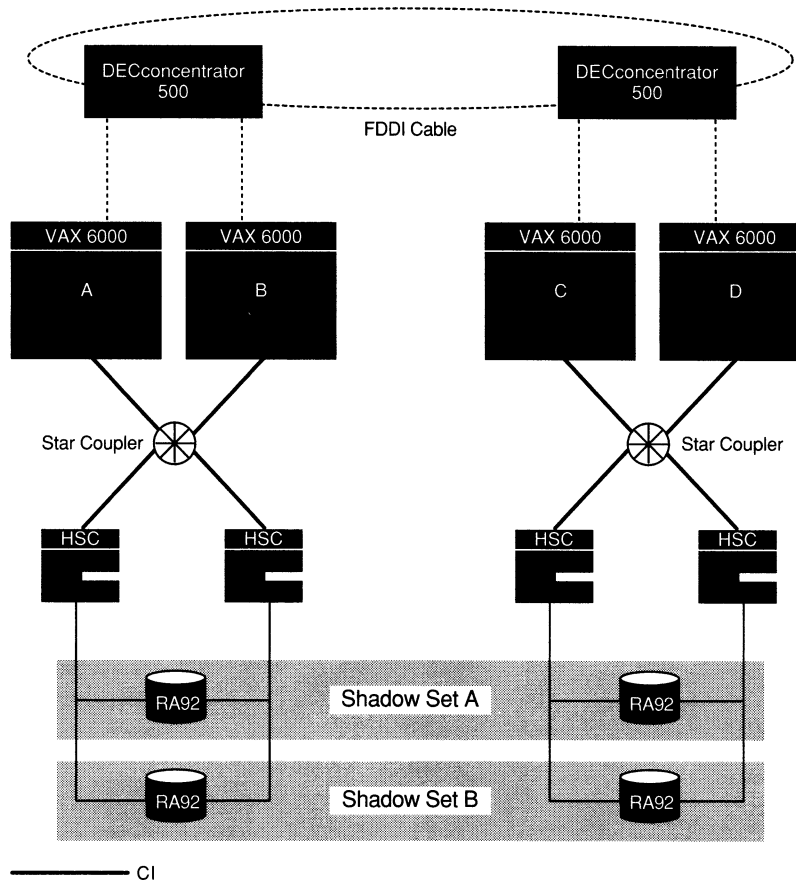
## 2.6 FDDI VAXcluster Systems

A **lobe** is a collection of VAX nodes connected by a CI and star coupler or connected by a DSSI. Such a collection could exist on its own as a VAXcluster system, but when two or more lobes are connected by the FDDI, they become one large, multilobe VAXcluster system.

### 2.6.1 Multilobe FDDI VAXcluster System

An example of a multilobe FDDI VAXcluster system is shown in Figure 2-6.

**Figure 2-6 FDDI VAXcluster System: Example Configuration**



ZK-3967A-GE

The elliptical line connecting the two DECconcentrator 500 units represents the dual-direction FDDI cables. Nodes VAX A and VAX B make up one lobe of the VAXcluster. Nodes VAX C and VAX D make up a second lobe (perhaps located at a distance from the first lobe). The FDDI fiber cable connects both lobes into a single VAXcluster system.

Note that VMS Volume Shadowing phase II is used to maintain key data storage devices in identical states (Shadow Sets A and B). Any sectors (on the shadowed disks) written at one site also will be written (virtually simultaneously) at the other site, and vice versa. Disks must be of the same type (for example, RA92).

# VAXcluster Interconnect Configurations

## 2.6 FDDI VAXcluster Systems

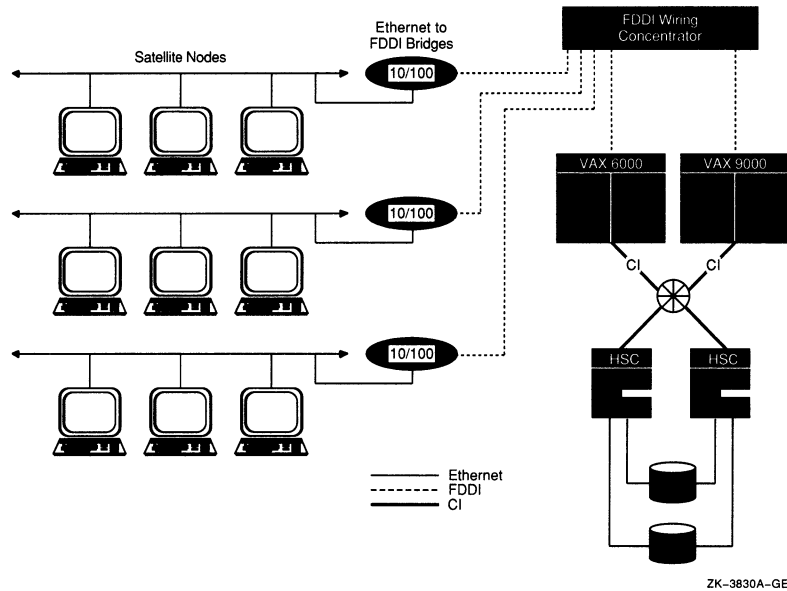
### 2.6.2 Network Configurations Using FDDI as VAXcluster Interconnect

Fiber Distributed Data Interface (FDDI) fiber optic cable can provide distinct advantages for networked client/server applications. Desktop environments can be integrated into VAXcluster systems by using FDDI's higher bandwidth (ten-fold speed improvement over Ethernet) and FDDI's ability to connect computing resources that are located (up to) 40 kilometers (25 miles) away.

Before FDDI, VAXcluster systems that were used for desktop integration typically had several large VAX processors acting as MOP and disk servers, plus many satellite nodes (clients) taking advantage of the served resources. Ethernet is used in this configuration to connect the satellites to the servers. Typically, satellites (workstations) on multiple Ethernet segments are bridged (using multiple LAN Bridge 200 devices, or DEBAMs) into a DELNI Ethernet wiring concentrator. From the DELNI, connections are made to the boot and disk server nodes of the CI VAXcluster.

VAXcluster systems that have heavily utilized Ethernet segments can replace the Ethernet backbone with FDDI to alleviate the Ethernet as a performance bottleneck. As shown in Figure 2-7, FDDI can replace the Ethernet from the bridges to the server CPU nodes. This configuration can increase overall throughput.

Figure 2-7 FDDI in Conjunction with Ethernet in a VAXcluster System



The longer distances provided by FDDI might permit you to create new VAXcluster systems in your computing environment. For example, another way of looking at Figure 2-7 is that this VAXcluster system is new. The large VAX nodes on the right could have replaced server nodes that previously existed on the individual Ethernet segments.

Currently there are no storage controllers that connect directly to FDDI. CPU nodes connected to FDDI must have local storage or access to storage over another interconnect. In Figure 2-7, the VAX 6000 and VAX 9000 computers have CI connections for storage. If a VAXcluster system has more than one FDDI-connected node, then those CPU nodes will probably use CI or DSSI

## VAXcluster Interconnect Configurations 2.6 FDDI VAXcluster Systems

connections for storage. The CI-connected VAX 6000 and VAX 9000 in Figure 2-7 are considered a **lobe** of the VAXcluster system.

For data center consolidation, FDDI can expand your data center to include additional resources while retaining a single system management domain. Computing resources and their associated storage that are today physically located outside your data center can now be tied in with FDDI; these resources could include standalone systems or another VAXcluster system. When connecting two or more existing VAXcluster systems with FDDI, you must create a single new VAXcluster system. The previously independent VAXcluster systems can remain physically separated within a building or in several buildings. But the systems now logically form a single operating environment, as depicted in Figure 2-8.

Figure 2-8 Multi-Datcenter VAXcluster System

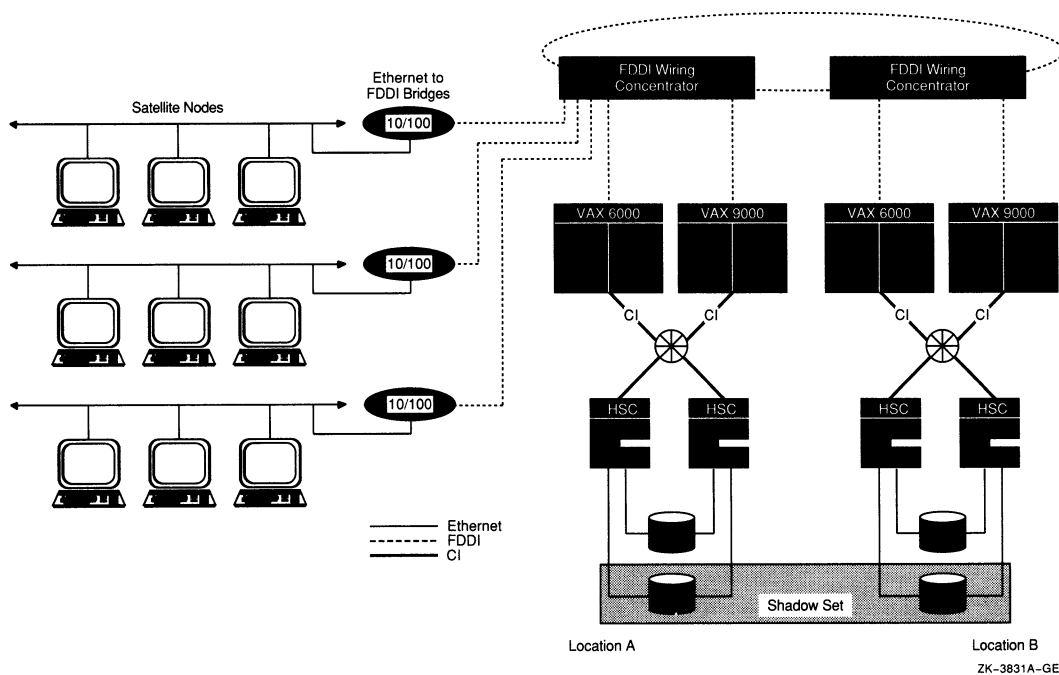


Figure 2-8 shows a multilobe VAXcluster configuration. It builds on Figure 2-7 where one CI based VAXcluster system was serving a large number of satellites. The VAX 6000 and VAX 9000 pair at *location A* and their CI-connected storage have been identified as one lobe. If more compute power or storage resources are required, additional systems can be connected in through FDDI.

### 2.7 Configuring Multiple CI Adapters

Extensions to the CI architecture allow the application of multiple CI interfaces per CPU and multiple star couplers per VAXcluster system. These extensions make possible VAXcluster systems with many times the data throughput capacity of current VAXcluster systems with a single star coupler.

## VAXcluster Interconnect Configurations

### 2.8 Configuring Multiple DSSI Adapters

## 2.8 Configuring Multiple DSSI Adapters

The DSSI architecture allows the application of multiple DSSI interfaces per CPU, making possible VAXcluster systems with many times the data throughput capacity of current VAXcluster systems with a single DSSI adapter.

## 2.9 Configuring Multiple LAN Adapters

VMS now supports multiple local area network (LAN) adapters on each local area VAXcluster node. Local area VAXcluster support for multiple adapters allows PEDRIVER (the VAXport driver for the LAN) to establish more than one channel between the local and remote cluster nodes. A channel is a unique network path between two nodes that is represented by a pair of LAN addresses or LAN adapters.

Local area VAXcluster systems with multiadapter support have the following characteristics:

- At boot time, *all* Ethernet and FDDI adapters are automatically configured for local area VAXcluster use.
- PEDRIVER automatically detects and creates a new channel for each unique pair of LAN adapters between the local node and each remote cluster node.
- Channel viability is continuously monitored at least every 3 seconds.
- Channel failure does not interfere with node-to-node (virtual circuit) communications as long as there is at least one remaining functioning channel between the nodes.

Configurations for multiadapter local area VAXcluster systems must meet the following requirements:

- The MOP server and the system disk server for a given satellite must be connected to the same LAN segment.
- All nodes must have a direct path to all other nodes. A direct path can be a bridged or unbridged LAN segment.

For each node, DECnet and MOP serving can be performed by only one adapter per extended LAN to prevent LAN address duplication.

The following guidelines are for configuring local area VAXcluster systems with multiple LAN adapters. If you configure these systems according to the guidelines, server nodes (nodes serving disks, tape, and lock traffic) can typically use some of the additional bandwidth provided by the added LAN adapters and increase the overall performance of the cluster. However, the performance increase depends on the configuration of your cluster (see Appendix G) and the applications it supports.

Configurations for multiadapter local area VAXclusters should follow these guidelines:

- All cluster nodes should run VMS Version 5.4-3 or later.
- Connect each LAN adapter to a separate LAN segment. A LAN segment can be bridged or unbridged. Doing this can help provide higher performance and availability in the cluster. The LAN segments can be either Ethernet segments or FDDI rings.
- Distribute satellites equally among the LAN segments. Doing this can help to distribute the cluster load more equally across all of the LAN segments.

## VAXcluster Interconnect Configurations

### 2.9 Configuring Multiple LAN Adapters

- LAN adapters providing MOP service should be distributed among the LAN segments to ensure that LAN failures do not prevent satellite booting.
- For the number of LAN adapters supported per node, refer to the VAXcluster SPD.

The following guidelines are for configuring the local area VAXcluster to be highly available:

- Bridge cluster LAN segments together to form a single extended LAN.
- Provide redundant LAN segment bridges for failover support.
- Configure LAN bridges to pass the local area VAXcluster and Maintenance Operations Protocol (MOP) multicast messages. You can use any of the following to build the configuration:
  - LAN bridge configuration documentation
  - Remote Bridge Management Software (RBMS)
  - Digital Extended LAN Management Software (DECelms)

Refer to the documentation for your LAN bridge or for the RBMS to obtain more information on configuring LAN bridges to pass these multicast messages.

- Use the Local Area VAXcluster Network Failure Analysis program to monitor and maintain network availability (see Section E.1.3).

#### 2.9.1 VAXcluster with Two LAN Segments

Figure 2–9 shows a sample configuration for a local area VAXcluster system connected to two different LAN segments. The figure illustrates the following points:

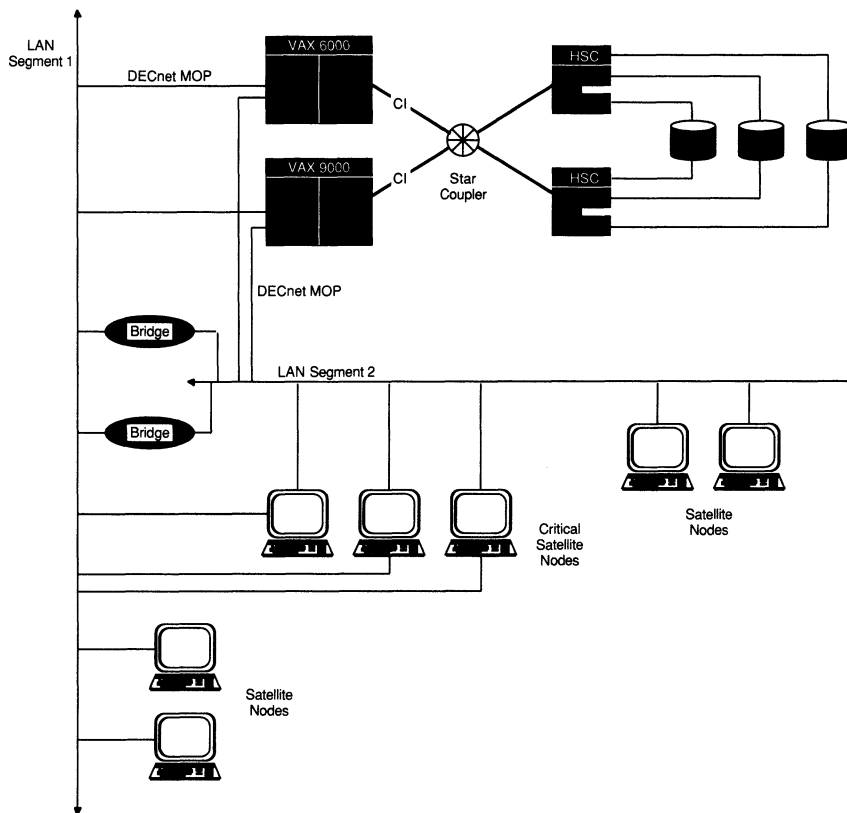
- Connecting critical nodes to multiple LAN segments provides increased availability in the event of segment or adapter failure. Disk and tape servers can use some of the network bandwidth provided by the additional network connection. Critical satellites can be booted using the other LAN adapter if one LAN adapter fails.
- Connecting noncritical satellites to only one LAN segment helps to balance the network load by equally distributing systems among the LAN segments. These systems communicate with satellites on the other LAN segment through one of the bridges.
- Only one LAN adapter per node can be used for DECnet and MOP service to prevent duplication of LAN addresses.
- LAN adapters providing MOP service should be distributed among the LAN segments to ensure that LAN failures do not prevent satellite booting.
- Using redundant LAN bridges prevents the bridge from being a single point of failure.

For descriptions of sample local area VAXcluster network connections, see Appendix D.

## VAXcluster Interconnect Configurations

### 2.9 Configuring Multiple LAN Adapters

Figure 2-9 Sample Two-LAN Segment VAXcluster Configuration



ZK-3828A-GE

#### 2.9.2 VAXcluster with Three LAN Segments

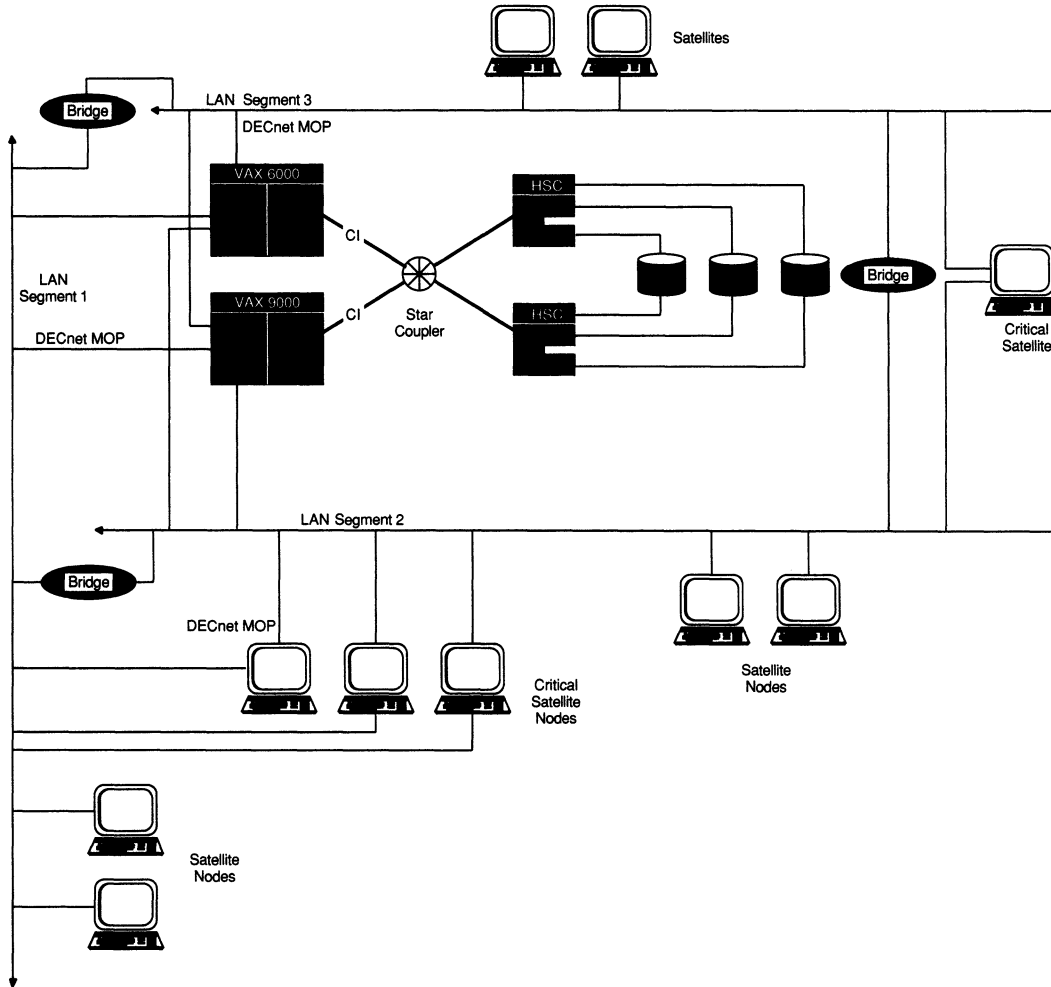
Figure 2-10 shows a sample configuration for a local area VAXcluster system connected to three different LAN segments. The figure illustrates the following points:

- Connecting disk and tape servers to two or three LAN segments can help provide higher availability and better I/O throughput.
- Connecting critical satellites to two or more LAN segments can also increase availability. If any of the network components fail, these satellites can use the other LAN adapters to boot and still have access to the critical disk servers.
- Distributing noncritical satellites equally among the LAN segments can help balance the network load.
- A MOP server is provided for each LAN segment.

For descriptions of sample local area VAXcluster network connections, see Appendix D.



Figure 2-10 Sample Three-LAN Segment VAXcluster Configuration



ZK-3829A-GE

### 2.9.3 Guidelines to Allow for LAN Bridge Failover

To achieve high availability, Digital recommends redundant bridges between LAN segments. If one bridge fails, another bridge can support message traffic between the LAN segments. To help ensure that there is little delay between the failure of one bridge and the continuing of message traffic support by another bridge, you should make sure the local area VAXcluster follows these guidelines:

- Bridge timers should be set to be faster than VAXcluster timers.
- Bridge self-test times should be less than the value for the SYSGEN parameter RECNXINTERVAL (described in Appendix A).

## VAXcluster Interconnect Configurations

### 2.9 Configuring Multiple LAN Adapters

The time required for one LAN bridge to fail over to another LAN bridge must adhere to the following parameters:

- Root bridge parameters:
  - LISTEN\_TIME (default equals 15 seconds)
  - FORWARDING\_DELAY (default equals 30 seconds)
- Bridge self-test time

If the VAXcluster is to survive a failover between LAN bridges, you must specify a value for the SYSGEN parameter RECNXINTERVAL that is greater than the following:

```
Bridge self-test time +  
LISTEN_TIME (of the root bridge) +  
FORWARDING_DELAY (of the root bridge)
```

The default value of RECNXINTERVAL is 20. For a value that is larger than the default failover time of a LAN bridge, use a value greater than 120. Note that the value of RECNXINTERVAL can be rounded up based on port timers.

A LAN bridge receives a HELLO message from the other bridges in an interval specified by the root bridge parameter HELLO\_INTERVAL. If the LAN bridge does not receive the HELLO message within the time specified by the bridge parameter LISTEN\_TIME, then the bridge changes the topology and possibly chooses a new root bridge. After the topology change, the bridge delays for the number of seconds specified by the value in the bridge parameter FORWARDING\_DELAY. This delay allows the LAN bridge to learn which LAN addresses are on either side of it. The bridge then starts forwarding packets.

Decreasing the LISTEN\_TIME value allows the bridge to detect topology changes more quickly. If you reduce the LISTEN\_TIME parameter value, you should also decrease the value for the HELLO\_INTERVAL bridge parameter. However, note that decreasing the value for the HELLO\_INTERVAL parameter causes the bridge to send the HELLO messages more frequently and thus increases LAN traffic.

Decreasing the FORWARDING\_DELAY value can cause the bridge to forward packets unnecessarily to the other LAN segment. Unnecessary forwarding can temporarily cause more traffic on both LAN segments until the bridge software determines which LAN address is on each side of the bridge.

If you change a parameter on one LAN bridge, you should change that parameter on all bridges to ensure that selection of a new root bridge does not change the value of the parameter. The actual parameter value the bridge uses is the value specified for the root bridge.

#### 2.9.4 Adjusting LRPSIZE for FDDI Configurations

FDDI supports transfers using large packets (up to 4468 bytes). PEDRIVER does not use large packets by default, but can take advantage of the larger packet sizes if you increase the SYSGEN parameter LRPSIZE to 1504 or higher. The SYSGEN parameter LRPSIZE (described in Appendix A) specifies the size of the large request packets.

PEDRIVER will use the full packet size if the LRPSIZE is set to 4474 or higher. However, only FDDI nodes connected to the same ring make use of large packets. Nodes connected to an Ethernet segment restrict packet size to that of an Ethernet packet (1498 bytes).

## VAXcluster Interconnect Configurations

### 2.9 Configuring Multiple LAN Adapters

If your local area VAXcluster configuration can take advantage of large FDDI packets, you should increase LRPSIZE to 4474 bytes on those nodes that have FDDI adapters on a common FDDI ring. Note that PEDRIVER does not combine multiple messages into a single packet. Therefore, increasing LRPSIZE to use large packets is beneficial only if the workload does large data transfers.

Configurations mixing Cluster Interconnect (CI) or Distributed System Services Interconnect (DSSI) with FDDI for cluster communication typically cannot benefit from using the large packets on the FDDI. Systems in these configurations choose the CI or DSSI path over the FDDI path. Increasing the SYSGEN parameter LRPSIZE for these mixed configurations increases the demand for physical memory (nonpaged pool) with no corresponding gain.

PEDRIVER detects that a message has traveled across an Ethernet segment by testing the value of the Priority Field in the Frame Control Byte of the FDDI header. Each bridge should set the Priority Field value to zero. (Digital's bridges set this value to zero.) PEDRIVER sends all FDDI messages with a nonzero Priority Field value. If a bridge does not properly set the Field Priority value to zero, PEDRIVER logs port errors on the console and the virtual circuit breaks connections to other FDDI nodes. If this condition occurs, reduce LRPSIZE to the default value of 1504 to prevent large packet use and reboot the FDDI nodes in the cluster.

If you decide to change the value of the LRPSIZE, edit the file `SYS$SPECIFIC:[SYSEXE]MODPARAMS.DAT` to permit AUTOGEN to factor the changed LRPSIZE into its calculations.

## 2.10 Security for Local Area and Mixed-Interconnect VAXcluster Systems

Local area and mixed-interconnect VAXcluster systems use a **group number** and a **cluster password** to allow multiple independent VAXcluster systems to coexist on the same extended LAN and to prevent accidental access to a cluster by unauthorized computers.

- The group number uniquely identifies each mixed-interconnect and local area VAXcluster system on a LAN. This number must be in the range from 1 to 4095 or from 61440 to 65535. Note that if you plan to have more than one of these clusters at your site, you must coordinate the assignment of group numbers among system managers.
- The cluster password serves as an additional check to ensure the integrity of individual clusters on the same LAN that accidentally use identical group numbers. (If each cluster's password is unique, the clusters will form independently.) The password also prevents an intruder who discovers the group number from joining the cluster. The password must be from 1 to 31 alphanumeric characters in length, including dollar signs (\$) and underscores (\_).

Security data for the cluster is maintained in the cluster authorization file, `SYS$COMMON:[SYSEXE]CLUSTER_AUTHORIZE.DAT`. This file is created during installation of the VMS operating system if you indicate that you want to set up a local area or mixed-interconnect cluster. The installation procedure then prompts you for the cluster group number and password. Cluster security functions are described in detail in Chapter 7. (If you convert a CI-based cluster to a mixed-interconnect configuration, the file is created when you execute the `CLUSTER_CONFIG.COM` command procedure, as described in Chapter 7.)



---

## VAXcluster Integrity and Availability Features

This chapter describes how the integrity of the VAXcluster configuration is provided for and explains how data availability in the cluster is maintained, including usage of the optional volume shadowing software.

### 3.1 Connection Management

The integrity of a VAXcluster system is controlled by a software component called the **connection manager**, which determines and coordinates membership in the cluster. The connection manager creates a cluster when the first active computers are booted and then reconfigures the cluster when computers join or leave it.

VAXcluster computers can share various data and system resources, such as disk volumes. To achieve the coordination that is necessary to maintain resource integrity, the computers must share a clear sense of cluster membership, which is maintained by the connection manager.

Within a single VAXcluster system, the VMS operating system guarantees the integrity of shared resources by carefully coordinating their use. However, because use of shared resources is not coordinated between computers in separate clusters—a condition known as **cluster partitioning**—the connection manager prevents this condition using a scheme called **quorum**.

#### 3.1.1 The Quorum Scheme

The quorum scheme is based on the arithmetic principle that the whole cannot be divided into multiple parts in such a way that more than one part is greater than half of the whole. (Integer arithmetic is used in this section.)

The quorum scheme functions as follows:

- In a VAXcluster system, each **voting member** (a computer with a nonzero value for the SYSGEN parameter VOTES) contributes a fixed number of votes toward quorum. On satellites, the VOTES value is always set to zero by default.
- Each active computer in the cluster (including satellites) indirectly specifies an initial quorum value using the SYSGEN parameter EXPECTED\_VOTES. This parameter is the sum of all votes held by potential cluster members. It is used to derive an estimate of the correct quorum value for the cluster, according to the following formula:  
$$\text{Estimated quorum} = (\text{EXPECTED\_VOTES} + 2) / 2$$
- During certain cluster state transitions, the system dynamically computes the cluster quorum to be the *maximum* of the following:
  - The current cluster quorum value

## VAXcluster Integrity and Availability Features

### 3.1 Connection Management

- The largest of the values calculated from the following formula, where EV is the EXPECTED\_VOTES value specified by each computer:

$$(EV+2) / 2$$

- The value calculated from the following formula, where V is the total of the SYSGEN parameter VOTES held by all cluster members:

$$(V+2) / 2$$

The cluster state transitions that cause cluster quorum to be recalculated occur when a computer joins the cluster and when the cluster recognizes a quorum disk. (The role of the quorum disk is discussed in Section 3.1.2.)

- If the current number of votes ever drops below quorum (because of computers leaving the cluster), the remaining cluster members suspend all process activity and all I/O operations to cluster-accessible disks and tapes until sufficient votes are added (computers joining the cluster) to bring the total number of votes to a value greater than or equal to quorum.
- As the cluster configuration changes, a computer only raises the cluster quorum value; it never lowers the value. (However, system managers can lower the value; for details, see Section 7.7.5.)

For example, consider a cluster consisting of three computers, each computer having its VOTES parameter set to 1 and its EXPECTED\_VOTES parameter set to 3. The connection manager dynamically computes the cluster quorum value to be 2. In this example, any two of the three computers constitute a quorum and can run in the absence of the third computer. No single computer can constitute a quorum by itself. Therefore, there is no way the three VAXcluster computers can be partitioned and run as two independent clusters.

#### 3.1.2 Quorum Disk

A quorum disk acts as a virtual computer, adding to the total cluster votes. By establishing a quorum disk in configurations with a small number of voting computers, you can increase the availability of the cluster. Such configurations can tolerate the failure either of the quorum disk or of a computer and continue operating.

To use a quorum disk, one or more computers must have a direct (non-MSCP-served) connection to the disk. Such computers are known as **quorum disk watchers**. Computers that cannot access the disk directly rely on the quorum disk watchers for information about the status of votes contributed by the quorum disk.

You should enable as quorum disk watchers any computers that have an active direct connection to the quorum disk or that have the potential for a direct connection. To enable a computer as a quorum disk watcher, use the CLUSTER\_CONFIG.COM CHANGE function described in Section 7.5.3. The procedure prompts for the name of the quorum disk and specifies that name as a value for the SYSGEN parameter DISK\_QUORUM in MODPARAMS.DAT. The procedure also sets an appropriate value for the QDSKVOTES parameter. The number of votes contributed by the quorum disk is equal to the smallest value of the SYSGEN parameter QDSKVOTES on any quorum disk watcher.

# VAXcluster Integrity and Availability Features

## 3.1 Connection Management

---

### Note

---

You can also enable the first installed cluster computer as a quorum disk watcher by answering YES when the VMS installation procedure asks whether the cluster will contain a quorum disk.

---

For the quorum disk's votes to be counted in the cluster votes total, the following conditions must be met:

- On one or more computers capable of becoming watchers, you must specify the same *physical* device name as a value for the SYSGEN parameter DISK\_QUORUM. The remaining computers (which must have a blank value for DISK\_QUORUM) recognize the name specified by the first quorum disk watcher with which they communicate. *A VAXcluster system can include only one quorum disk.*
- At least one quorum disk watcher must have a direct, active connection to the quorum disk. Thus, the quorum disk may be a dual-ported DSA disk, which has an active direct connection to only one computer at a time.
- The disk must contain a valid format file named QUORUM.DAT in the master file directory (MFD). The QUORUM.DAT file is created automatically after a system specifying a quorum disk has booted into the cluster. This file is used on subsequent reboots. Note that the file is not created if the SYSGEN parameter STARTUP\_P1 is set to *MIN*.
- To permit recovery from failure conditions, the quorum disk must be mounted by all disk watchers.

### 3.1.3 State Transitions

VAXcluster state transitions occur when a computer joins or leaves a VAXcluster system. The connection manager controls these events to ensure the preservation of data integrity throughout the cluster. A state transition's duration and effect on users (applications) are determined by the reason for the transition, the configuration, and the applications in use.

Every transition goes through one or more phases, depending on whether its cause is the addition of a new VAXcluster member or the failure of a current member. If the transition is caused by the addition of a new member, the phases are as follows:

- **New member detection**  
Early in its boot sequence, a computer seeking membership in a VAXcluster system sends messages to current members asking to join the cluster. The first cluster member that receives the membership request acts as the new computer's advocate and proposes reconfiguring the cluster to include the computer in the cluster. While the new computer is booting, no applications are affected.
- **Reconfiguration**  
All current VAXcluster members must establish communications with the new computer. Once communications are established, the new computer is admitted to the cluster. In some cases, the lock database is rebuilt.

# VAXcluster Integrity and Availability Features

## 3.1 Connection Management

If the transition is caused by the failure of a current VAXcluster member, the phases are as follows:

- Failure detection

The duration of this phase depends on the cause of the failure and on how the failure is detected.

During normal cluster operation, messages sent from one computer to another are acknowledged when received. If a message is not acknowledged within a period determined by VAXcluster communications software, the repair attempt phase begins.

If a cluster member is shut down or crashes, the VMS operating system causes datagrams to be sent from the computer shutting down to the other members. These datagrams state the computer's intention to sever communications and stop sharing resources. Because sending these datagrams is virtually the last activity of a dying computer, they are called "last gasp" datagrams. If any current cluster member receives a last gasp datagram, the "gasping" computer is removed from the cluster. The failure detection and repair attempt phases are bypassed, and the reconfiguration phase begins immediately.

- Repair attempt

If the communication path to a VAXcluster member is broken, attempts are made to repair the path. Repair attempts continue for an interval specified by the SYSGEN parameter RECNXINTERVAL. (System managers can adjust the value of this parameter to suit local conditions.) Thereafter, the path is considered irrevocably broken, and steps must be taken to reconfigure the VAXcluster system so that all computers can once again communicate with each other, and so that any computers that cannot communicate are removed from the cluster.

- Reconfiguration

When a VAXcluster member fails, the cluster must be reconfigured. One of the remaining computers acts as coordinator and exchanges messages with all other cluster members to determine the configuration of an **optimal subcluster** with the most members and the most votes. This phase, during which all user (application) activity is blocked, usually lasts less than 1 second.

- VAXcluster system recovery

Recovery includes the following stages, some of which can take place in parallel:

- I/O completion

When a computer is removed from the cluster, VAXcluster software ensures that all I/O operations that are started by the old configuration complete before I/O operations that are generated by the new configuration start. There is usually little or no effect on applications.

- Lock database rebuild

Because the lock database is distributed among all members, some portion of the database may need rebuilding. A rebuild is always performed when a computer leaves the cluster, but only in certain cases when a computer is added.



## VAXcluster Integrity and Availability Features

### 3.1 Connection Management

- Disk mount verification

This stage occurs only when the failure of a voting member causes quorum to be lost. To protect data integrity, all I/O activity is blocked until quorum is regained. Mount verification is the mechanism used for this purpose.

- Quorum votes validation

If, when a computer is removed, the remaining members can determine that it has shut down or crashed, the votes contributed by the quorum disk are included without delay in quorum calculations that are performed by the remaining members. However, if they cannot determine that the computer has shut down or crashed (for example, if a console halt, power failure, or communications failure has occurred), the votes are not included for a period equal to four times the value of QDSKINTERVAL seconds. This period is sufficient to determine that the failed computer is no longer using the quorum disk.

- Disk rebuild

If the transition is the result of a computer rebooting after a crash, the disks are marked as improperly dismounted, and they must be rebuilt before they can be remounted. The rebuild reclaims space that was cached by the failed computer but never returned, as with a normal dismount. A rebuild makes the disk briefly inaccessible to users unless the rebuild is deferred by using the /NOREBUILD qualifier to the MOUNT commands in the system startup files. (See Section 5.5 for more information on rebuilding disks.)

- Application recovery

When assessing the effect of a state transition on application users, consider that the application recovery phase includes activities such as replaying a journal file, cleaning up recovery units, and users logging in again because the terminal server has failed over to another computer.

VAXcluster troubleshooting information appears in Appendix C.

## 3.2 VMS Volume Shadowing

This section explains volume shadowing concepts and introduces VMS Volume Shadowing terms and requirements.

### 3.2.1 Volume Shadowing Concepts

Data availability is an essential requirement of most VAXcluster systems. Volume shadowing provides high data availability by protecting against data loss resulting from media deterioration, communication path failure, or through controller or device failure.

Volume shadowing is the process of maintaining multiple copies of the same data on two or more disk volumes. The duplication of data provides greater data availability and, in some cases, faster data accessibility. When the same data is recorded on more than one disk volume, you have access to that data even when one volume is unavailable. Disk input/output operations continue transparently with the remaining disk volumes.

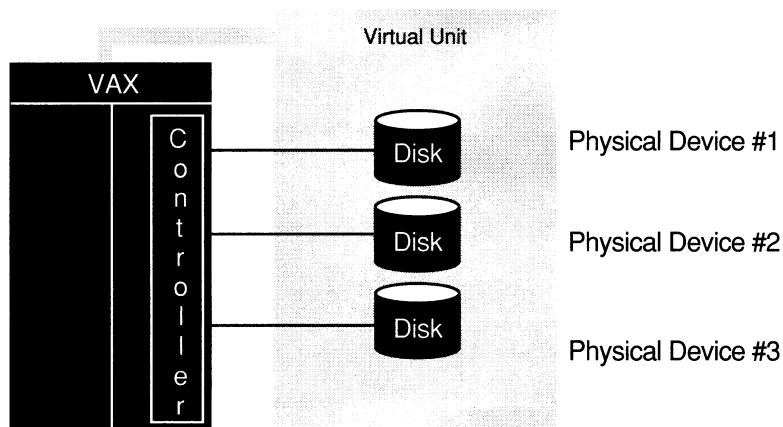
# VAXcluster Integrity and Availability Features

## 3.2 VMS Volume Shadowing

Volume shadowing allows you to unite individual disk volumes into a shadow set. A shadow set consists of one, two, or three compatible disk volumes referred to as shadow set members.<sup>1</sup> To be compatible, disks that “shadow” each other must be the same type of device with the same physical geometry, such that each disk in a given shadow set must have the same number of identical logical blocks (LBNs) as the other disks in the shadow set. For example, an RF72 disk can shadow another RF72, but an RF72 disk cannot shadow an RA81.

After you create a shadow set, applications and users can access the shadow set as a single virtual unit. A virtual unit is a software-created representation of a group of disks and disk drives that compose a shadow set. Figure 3–1 illustrates the basic components of a system configuration with a shadow set.

**Figure 3–1 Elements of a Shadow Set**



ZK-9657-GE

Although the primary objective of volume shadowing is to provide for high data availability, there are other benefits. Because a shadow set is made up of multiple disks containing the same data, the shadow set can use the additional disk read heads to respond to multiple read requests at the same time. Having multiple read heads can result in a higher effective search speed when an application or user reads data from the disk. An additional benefit of volume shadowing is its potential for repairing data. When normal media deterioration renders sections of a volume unreadable, systems with volume shadowing can read the duplicate data and copy it to the failing volume to replace the lost data on the other volume.

**Note**

VMS Volume Shadowing provides for duplicate data at the LBN level, and it should not be mistaken for a journaling product that provides application-level (data integrity) recovery.

<sup>1</sup> VMS Volume Shadowing supports a shadow set consisting of only one disk volume, but this configuration usually exists only for the purpose of continuing operation during repairs or failures of other disks in the shadow set. Note that because VMS Volume Shadowing duplicates data by writing it to identical logical blocks (LBNs) on multiple disks, the volume shadowing software does not duplicate data when only one disk is available.

## VAXcluster Integrity and Availability Features

### 3.2 VMS Volume Shadowing

The VMS operating system supports shadowing for users of Files-11 On-Disk Structure Level 2 (ODS2) data disks. You can mount shadow sets as private volumes or public volumes either on a standalone system or in a VAXcluster system.

Volume shadowing phase I provides for centralized shadowing on VMS systems using hierarchical storage controllers (HSCs) with compatible DIGITAL Storage Architecture (DSA) disks. Volume shadowing phase II provides distributed shadowing in a controller-independent manner. Phase II supports clusterwide shadowing of all VMS MSCP-compliant DSA disks and SCSI disks having the same physical geometry, on a single system or located anywhere in a VAXcluster system. Note, however, that it is not necessary to have VAXcluster software to run volume shadowing on a single, standalone node.

Volume shadowing phase II supports clusterwide shadowing of all DSA devices. Phase II provides volume shadowing capabilities for all DSA disks including those on both HSC and local adapters, all Digital Small Systems Interconnect (the DSSI RF-series) disk devices on any VAX computer, the Small Computer System Interface (SCSI) disk devices, all interfaces (CI, LAN, DSSI, mixed-interconnect), and across VMS MSCP servers.

A phase II shadow set can have a maximum of three shadow set members.

---

#### Note

---

Phase II volume shadowing places no explicit restrictions on the total number of shadow sets that can be configured. However, interconnect and adapter bandwidth capacities can force limits on the number of shadow sets that can be configured.

---

For detailed information on volume shadowing, see the *VMS Volume Shadowing Manual*. For configuration restrictions for phases I and II, refer to the Volume Shadowing SPD.

#### 3.2.2 Distributing Shadow Sets Across VMS Configurations

One of the primary advantages of volume shadowing phase II is its ability to work across a wide set of hardware configurations. Shadowed disk volumes can be located anywhere within the full range of VMS configurations—single nodes through large VAXcluster systems. By supporting clusterwide shadow sets, volume shadowing offers improved data availability and configurations.

Phase II manages shadow sets independently of the controllers, thus allowing both HSC and non-HSC configurations to utilize the benefits of shadowing. Shadow set member units can be on **local** controllers located on different VMS nodes. For example, Figure 3-2 shows that there is a disk volume local to each of the VAX nodes ATABOY and ATAGRL. Even though the disk volumes are local to different VAX nodes, the disks are members of the same shadow set. A member unit that is local to one node can be accessed by the remote node over the VMS mass storage control protocol (VMS MSCP) server. The VMS MSCP server is the software that makes disks that are connected to a remote system as accessible as if they were on the local system. (You must use the allocation class device name format to manipulate remote shadow set members. The shadowing software requires that you use allocation classes regardless of whether or not the device is local or remote.)

## VAXcluster Integrity and Availability Features

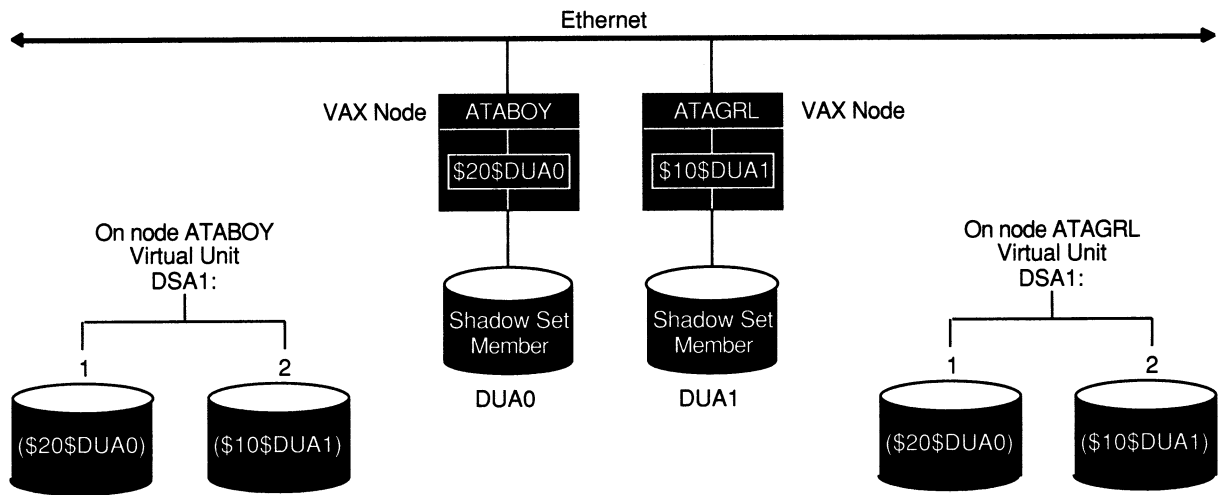
### 3.2 VMS Volume Shadowing

The VMS MSCP server accesses phase II shadow set members on a single system or on a VAXcluster system where interprocessor communication is carried out over the following types of network connections:

- Computer interconnect (CI)
- LAN (Ethernet or FDDI)
- Digital Small Systems Interconnect (DSSI)
- Mixed-interconnects (a combination of CI, LAN, and/or DSSI)

In Figure 3–2, the VMS MSCP server accesses the shadow set members over the LAN. These controller-independent, distributed configuration options make it possible to manage shadow sets regardless of their controller connection or location in the VAXcluster system, and help provide a high degree of data availability.

**Figure 3–2 Shadow Set Members Accessed Through the VMS MSCP Server**



ZK-3966A-GE

Phase II volume shadowing provides tolerance from the following types of faults: disk media errors, disk failures, controller communication failures, or controller errors across the full range of VAX computers and configurations.

The phase II shadowing software maintains virtual units in a distributed fashion on each node that mounts the shadow set in the VAXcluster system. Phase II volume shadowing can provide distributed shadowing because the virtual unit is maintained and managed by all nodes that have the shadow set mounted. Figure 3–2 illustrates a simple example of how all nodes that have a phase II shadow set mounted know the shadow set membership, can perform read and write I/O operations through the virtual unit to the shadow set members, and can use the VMS MSCP server to access shadow sets mounted on remote nodes.

## VAXcluster Integrity and Availability Features

### 3.2 VMS Volume Shadowing

---

#### Note

---

VMS Volume Shadowing does not support the shadowing of quorum disks. This is because volume shadowing makes use of the VMS distributed lock manager and the quorum disk must be accessed in order to allow locking.

---



---

## Preparing the Cluster Operating Environment

By setting up appropriate startup and other system files, you can prepare the VAXcluster operating environment on the first installed computer before adding other computers to the cluster. Depending on your processing needs, you can prepare either a **common-environment** or a **multiple-environment** cluster.

In a common-environment cluster, the operating environment is identical on each VAXcluster computer because the computers are run from common system files (which may include common resources and applications). The computers are set up with identical user accounts, the same known images are installed, the same logical names are defined, and mass storage devices and queues are shared. In effect, users in a common-environment cluster can log in to any computer and work in the same operating environment.

In a multiple-environment cluster, the environment (resources, applications) varies from computer to computer, and users can work in environments that are specific to the computer they are logged in to. A multiple-environment cluster is effective when you want to share data among computers but want certain computers to serve specialized needs. For example, you might want to set up a three-computer cluster, in which the timesharing environments on two computers are the same, while the third computer is set up exclusively for batch processing of large inventory jobs. In this case, the timesharing computers are set up with a common environment, sharing users, queues, and access to mass storage devices, while the third computer runs in its own restricted environment.

This chapter concentrates on the steps necessary to prepare a common-environment cluster. Approaches for preparing a multiple-environment cluster are also described, but are presented as general guidelines.

Topics include the following:

- Directory structure on a common system disk
- Installing the VMS operating system in the VAXcluster environment
- Configuring and starting the DECnet-VAX network
- Coordinating startup command procedures
- Coordinating system files for a common-environment cluster

Once you have prepared the cluster operating environment as described in this chapter and determined your disk and queue configurations using the information in Chapter 5 and Chapter 6, you can build the cluster following instructions in Chapter 7.

# Preparing the Cluster Operating Environment

## 4.1 Directory Structure on a Common System Disk

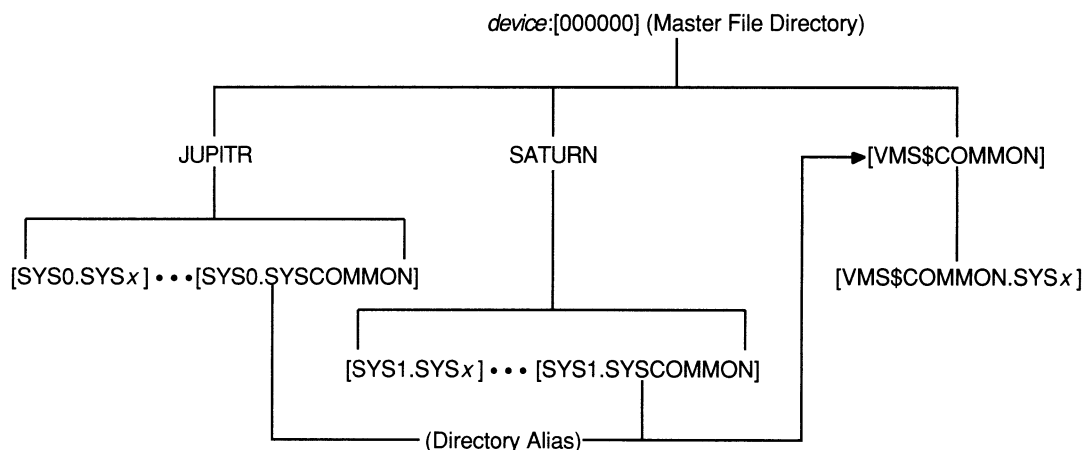
### 4.1 Directory Structure on a Common System Disk

The VMS installation or upgrade procedure generates a *common system disk*, on which most operating system and optional product files are stored in a common root directory. The entire directory structure—that is, the common root plus each computer's local root—is stored on the same disk. After the installation or upgrade completes, you use the CLUSTER\_CONFIG.COM command procedure described in Chapter 7 to create a local root for each new computer and boot it into the cluster.

Each local root contains, in addition to the usual system directories, a [SYSx.SYSCOMMON] directory that is an alias for [VMS\$COMMON], the cluster common root directory in which cluster common files actually reside. When you add a computer to the cluster, CLUSTER\_CONFIG.COM defines the alias.

Figure 4-1 illustrates the directory structure set up for computers JUPITR and SATURN, which are run from a common system disk. The disk's master file directory (MFD) contains the local roots (SYS0 for JUPITR, SYS1 for SATURN) and the cluster common root directory, [VMS\$COMMON].

Figure 4-1 Directory Structure on Common System Disk



SYS\$SPECIFIC = device:[SYSn.]

SYS\$COMMON = device:[SYS n.SYSCOMMON.]

SYS\$SYSROOT = device:[SYS n.], device:[SYS n.SYSCOMMON.]

Key: n = System Root

x = System Subdirectory

ZK-6658-GE

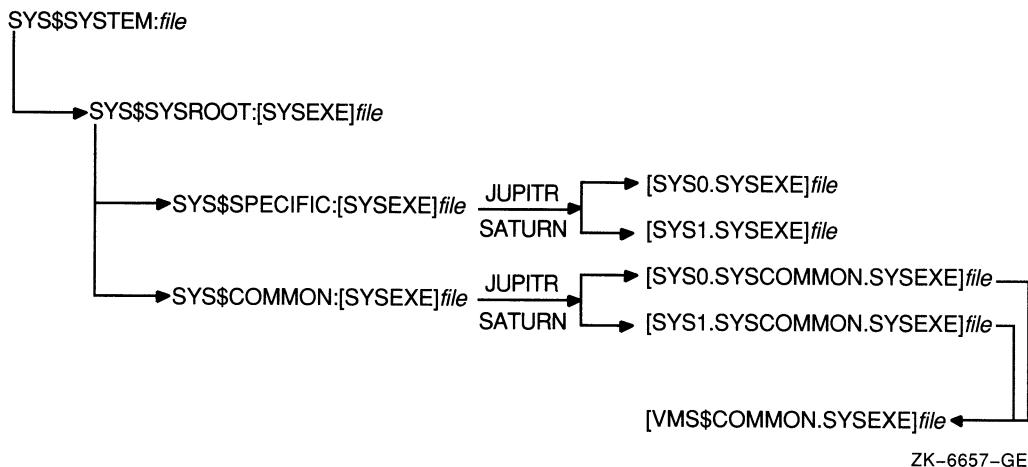
The logical name SYS\$SYSROOT is defined as a search list that points to a local root first (SYS\$SPECIFIC) and then to the common root (SYS\$COMMON). Thus, the logical names for the system directories (SYS\$SYSTEM, SYS\$LIBRARY, SYS\$MANAGER, and so forth) point to two directories: a local root (for example, SYS\$SPECIFIC:[SYSEXE]) and a common root (for example, SYS\$COMMON:[SYSEXE]). Figure 4-2 shows how directories on a common system disk are searched when the logical name SYS\$SYSTEM is used in file specifications.



## Preparing the Cluster Operating Environment

### 4.1 Directory Structure on a Common System Disk

**Figure 4-2 File Search Order on Common System Disk**



It is important to keep this search order in mind when manipulating system files on a common system disk. Computer-specific files must always reside and be updated in the appropriate computer's system subdirectory. For example, MODPARAMS.DAT must reside in SYS\$SPECIFIC:[SYSEXE], which is [SYS0.SYSEXE] on JUPITR, and [SYS1.SYSEXE] on SATURN. Thus, to create a new MODPARAMS.DAT file for JUPITR when logged in on JUPITR, you would enter the following command:

```
$ EDIT SYS$SPECIFIC:[SYSEXE]MODPARAMS.DAT
```

Once the file is created, you could use the following command to modify it:

```
$ EDIT SYS$SYSTEM:MODPARAMS.DAT
```

Note that if a MODPARAMS.DAT file does not exist in JUPITR's SYS\$SPECIFIC:[SYSEXE] directory when you enter this command, but there is a MODPARAMS.DAT file in the directory SYS\$COMMON:[SYSEXE], the command edits the MODPARAMS.DAT file in the common directory. If there is no MODPARAMS.DAT file in either directory, the command creates the file in JUPITR's SYS\$SPECIFIC:[SYSEXE] directory.

To modify JUPITR's MODPARAMS.DAT when logged in on any other computer that boots from the same common system disk, you enter the following command:

```
$ EDIT [SYS0.SYSEXE]MODPARAMS.DAT
```

If you want to modify records in the cluster common system authorization file in a cluster with a single cluster common system disk, you could enter the following commands on any computer:

```
$ SET DEFAULT SYS$COMMON:[SYSEXE]
$ RUN AUTHORIZE
```

But if, for example, you have set up a computer-specific system authorization file (SYSUAF.DAT) for computer JUPITR and you want to modify records in that file when logged in on another computer that boots from the same cluster common system disk, you must, before invoking AUTHORIZE, set your default directory to JUPITR's computer-specific [SYSEXE] directory. For example:

```
$ SET DEFAULT [SYS0.SYSEXE]
$ RUN AUTHORIZE
```

#### 4.2 Installing the VMS Operating System in the VAXcluster Environment

You must perform the installation or upgrade once for each system disk in the cluster. Because, however, several computers normally run from the same cluster common system disk, you need not perform the installation or upgrade on each computer. Refer to the Release Notes for the required version numbers for hardware and firmware. When mixing VMS versions, check the Release Notes for compatibility.

You may want to set up a cluster that has a combination of one or more common system disks and one or more individual system disks. Again, you must do the installation or upgrade once for each system disk. For example, if your cluster consists of 10 computers, 4 of which share one common system disk, 4 of which share a second common system disk, and each of the other 2 has its own system disk, you would do the installation or upgrade four times. Note that if your cluster includes multiple common system disks, *you must later coordinate system files to define the cluster operating environment, as described in Section 4.5.4.*

To perform the installation, follow instructions in the installation and operations guide for your computer. However, before you start the installation, be sure you have determined which *VAXcluster system configuration type* you want to create (CI-based, DSSI-based, local area, or mixed-interconnect), because the installation procedure requests configuration-specific information. (Configuration types are described in Section 2.1.)

Table 4-1 lists the information requested for CI-based configurations; this information is also required for DSSI-based configurations. Table 4-2 lists the information requested for local area and mixed-interconnect configurations. Typical responses are explained in the tables. Note that initial questions are the same for all configuration types. All references to the Ethernet are also applicable to the FDDI LAN.

If your system disk is on an HSC or ISE subsystem, you must obtain the HSC or ISE subsystem's *disk allocation class* value before starting the installation, because the installation procedure requests that information. (Allocation classes are discussed in detail in Section 5.2.) For example, to obtain the value, enter a command sequence like the following at the HSC or ISE console. The information displayed includes the allocation class value.

```
Ctrl/C
HSC> SHOW SYS
15-Jun-1990 14:31:43.41  Boot:  13-Jun-1990 11:31:11.41  Up:   51:00
.
.
.
DISK allocation class = 1          TAPE allocation class = 0
Start command file m Disabled
SETSHO - Program Exit
```

If you later want to change the allocation class value, follow the instructions in Section 7.6.

---

#### Note

---

While rebooting at the end of the installation procedure, the system displays messages warning that you must install VMS and VAXcluster software licenses. Be sure to install these licenses, as well as the DECnet-VAX license, as soon as the system is available. Procedures for

## Preparing the Cluster Operating Environment

### 4.2 Installing the VMS Operating System in the VAXcluster Environment

installing licenses are described in the release notes distributed with the software kit and in the *VMS License Management Utility Manual*.

**Table 4–1 Information Requested for CI-Based Configurations**

Item	Response
Will this node be a cluster member (Y/N)?	Enter Y.
What is the node's DECnet node name?	Enter DECnet node name—for example, JUPITR. The DECnet node name can be from 1 to 6 alphanumeric characters in length and cannot include dollar signs (\$) or underscores (_).
What is the node's DECnet node address?	Enter DECnet node address—for example, 2.2.
Will the Ethernet be used for cluster communications (Y/N)?	Enter N. The LAN is not used for cluster communications in CI-based VAXcluster systems.
Will JUPITR be a disk server (Y/N)?	Enter Y or N, depending on your configuration requirements. Refer to Section 2.4 and Chapter 5 for information on served cluster disks.
Enter a value for JUPITR's ALLOCLASS parameter:	If the system disk is connected to a dual-ported disk, enter a value from 1 to 255 that will be used on both sides. Otherwise, enter 0. (For detailed information on allocation classes see Section 5.2.)
Does this cluster contain a quorum disk [N]?	Enter Y or N, depending on your configuration. If you enter Y, the procedure prompts for the name of the quorum disk. Enter the device name of the quorum disk. (For detailed information on quorum disks, see Section 3.1.2.)

**Table 4–2 Information Requested for Local Area and Mixed-Interconnect Configurations**

Item	Response
Will this node be a cluster member (Y/N)?	Enter Y.
What is the node's DECnet node name?	Enter DECnet node name—for example, JUPITR. The DECnet node name may be from 1 to 6 alphanumeric characters in length and cannot include dollar signs (\$) or underscores (_).
What is the node's DECnet node address?	Enter DECnet node address—for example, 2.2.

(continued on next page)

## Preparing the Cluster Operating Environment

### 4.2 Installing the VMS Operating System in the VAXcluster Environment

Table 4–2 (Cont.) Information Requested for Local Area and Mixed-Interconnect Configurations

Item	Response
Will the Ethernet be used for cluster communications (Y/N)?	Enter Y. The LAN is required for cluster communications in local area and mixed-interconnect VAXcluster systems.
Enter this cluster's group number:	Enter a number in the range from 1 to 4095 or 61440 to 65535.
Enter this cluster's password:	Enter the cluster password. The password must be from 1 to 31 alphanumeric characters in length and can include dollar signs and underscores.
Reenter this cluster's password for verification:	Reenter the password.
Will JUPITR be a disk server (Y/N)?	Enter Y. In local area and mixed-interconnect configurations, the system disk is always served to the cluster. Refer to Section 2.4 and Chapter 5 for information on served cluster disks.
Will JUPITR serve HSC disks (Y/N)?	Enter a response appropriate for your configuration.
Enter a value for JUPITR's ALLOCLASS parameter:	If the system will serve HSC disks, enter the HSC's allocation class value. If the system disk is connected to a dual-ported disk, enter a value from 1 to 255 that will be used on both sides. If you are using multiple hosts and RF disks, assign a non-zero value to the allocation class. Otherwise, enter 0. (For detailed information on allocation classes see Section 5.2.)
Does this cluster contain a quorum disk [N]?	Enter Y or N, depending on your configuration. If you enter Y, the procedure prompts for the name of the quorum disk. Enter the device name of the quorum disk. (For detailed information on quorum disks, see Section 3.1.2.)

### 4.3 Configuring and Starting the DECnet–VAX Network

After you have installed the operating system and required licenses on the first VAXcluster computer, you configure, tailor, and start the DECnet–VAX network. If you locate certain network files in the `SYS$COMMON:[SYSEXE]` directory as described in step 3, other computers can share the data when they join the cluster. The process of configuring the network typically entails several operations:

- Executing the `SYS$MANAGER:NETCONFIG.COM` command procedure.
- Selecting a single LAN adapter when using multiple LAN adapters on a single extended LAN.
- Making remote node data available clusterwide.

## Preparing the Cluster Operating Environment

### 4.3 Configuring and Starting the DECnet-VAX Network

- Optionally defining a VAXcluster alias. You establish an alias using NCP commands like those shown in step 3 for alias SOLAR. (For more information on the VAXcluster alias, refer to the *VMS Networking Manual*.) Note that if you plan to define an alias, you must specify that one computer operate as a *router* node when you execute NETCONFIG.COM. Note further that you must later enable alias operations for other computers, as described in Section 4.3.2.
- Starting the network.

To perform these operations, proceed as follows:

1. Log in as system manager and execute NETCONFIG.COM. Enter information about your node when prompted and answer NO when the procedure asks whether you want network configuration commands to be executed. This will allow you to configure a VAXcluster alias and to select a LAN adapter when using multiple adapters connected to a single extended LAN.

Example 4-1 shows a typical NETCONFIG.COM session.

#### Example 4-1 Sample Interactive Network Configuration Session

```
$ @NETCONFIG.COM
```

```
DECnet-VAX network configuration procedure
```

This procedure will help you define the parameters needed to get DECnet running on this machine. You will be shown the changes before they are executed, in case you want to perform them manually.

```
What do you want your DECnet node name to be? [JUPITR]: 
What do you want your DECnet address to be? [2.2]: 
Do you want to operate as a router? [NO (nonrouting)]: YES
Do you want a default DECnet account? [NO]: 
```

```
Here are the commands necessary to set up your system.
```

```
·
·
·
```

```
Do you want these commands to be executed? [YES]: 
```

```
·
·
·
```

The changes have been made.

If you have not already registered the DECnet-VAX key, then do so now. After the key has been registered, you should invoke the procedure SYSSMANAGER:STARTNET.COM to start up DECnet-VAX with these changes.

```
(If the key is already registered) Do you want DECnet started? [YES] NO
$
```

2. NETCONFIG.COM configures all the LAN adapters for DECnet use. However, on a single extended LAN, this violates LAN configuration rules. A single LAN adapter must be chosen to connect to each extended LAN.

For example, to disable DECnet use of the LAN device XQB0, invoke the Network Control Program (NCP) utility and issue the following commands:

```
$ RUN SYSSSYSTEM:NCP
NCP>PURGE CIRCUIT QNA-1 ALL
NCP>PURGE LINE QNA-1 ALL
NCP>EXIT
```

## Preparing the Cluster Operating Environment

### 4.3 Configuring and Starting the DECnet-VAX Network

For further details, see the *VMS Networking Manual* and the *VMS Network Control Program Manual*.

3. NETCONFIG.COM creates, in the SYS\$SPECIFIC:[SYSEXE] directory, the permanent remote node database file NETNODE\_REMOTE.DAT, in which remote node data is maintained. To make this data available clusterwide, you must rename the file to the SYS\$COMMON:[SYSEXE] directory:

```
$ RENAME SYSSSPECIFIC:[SYSEXE]NETNODE_REMOTE.DAT -  
_ $ SYSSCOMMON:[SYSEXE]NETNODE_REMOTE.DAT
```

For information on sharing other network data, see the *VMS Networking Manual*.

4. If you want to define a VAXcluster alias, invoke the NCP Utility to do so. For example:

```
$ RUN SYSSSYSTEM:NCP  
NCP> DEFINE NODE 2.1 NAME SOLAR  
NCP> DEFINE EXECUTOR ALIAS NODE SOLAR  
NCP> EXIT  
$
```

The information you specify using these commands is entered in the DECnet-VAX permanent executor database and takes effect when you start the network.

5. Start the network:

```
$ @SYSSMANAGER:STARTNET.COM
```

6. To ensure that the network is started each time a VAXcluster computer boots, add the following line to the appropriate startup command file or files:

```
$ @SYSSMANAGER:STARTNET.COM
```

For information on preparing startup command files, see Section 4.4. For more detailed information on DECnet-VAX configuration issues and procedures, refer to the *VMS Networking Manual*.

#### 4.3.1 Copying Remote Node Databases

Some sites with large networks maintain remote node data in a central database file. If this is the case at your site, and if you want to make the data available clusterwide, you can, *after starting the network*, copy remote node database entries from that central file. For example, if the file resides on node SATURN, you could enter the following NCP commands to copy entries from the permanent database on SATURN to the permanent database on your system disk, and then to update your volatile database:

```
NCP> SET NODE 2.2 NAME SATURN  
NCP> COPY KNOWN NODES FROM SATURN USING PERMANENT TO PERMANENT  
NCP> SET KNOWN NODES ALL
```

Note that only node names and addresses are copied. See the *VMS Networking Manual* for more information on copying node databases.

## Preparing the Cluster Operating Environment

### 4.3 Configuring and Starting the DECnet-VAX Network

#### 4.3.2 Enabling VAXcluster Alias Operations

If you have defined a VAXcluster alias as described in Section 4.3, you can enable alias operations for other computers *after the computers are up and running in the cluster*. To enable such operations (that is, to allow a computer to accept incoming connect requests directed toward the alias), follow these steps:

1. Log in as system manager and invoke the SYSMAN Utility:

```
$ RUN SYS$SYSTEM:SYSMAN
```

2. At the SYSMAN> prompt, enter the following commands:

```
SYSMAN> SET ENVIRONMENT/CLUSTER
%SYSMAN-I-ENV, current command environment:
      Clusterwide on local cluster
      Username LAZARUS      will be used on nonlocal nodes
SYSMAN> SET PROFILE/PRIVILEGES=(OPER,SYSPRV)
SYSMAN> DO MCR NCP SET EXECUTOR STATE OFF
%SYSMAN-I-OUTPUT, command execution on node X...
.
.
.
SYSMAN> DO MCR NCP DEFINE EXECUTOR ALIAS INCOMING ENABLED
%SYSMAN-I-OUTPUT, command execution on node X...
.
.
.
SYSMAN> DO @SYS$MANAGER:STARTNET.COM
%SYSMAN-I-OUTPUT, command execution on node X...
.
.
.
```

#### 4.4 Coordinating Startup Command Procedures

You must coordinate your site-specific SYSTARTUP and SYLOGIN command procedures according to the type of cluster operating environment you want to prepare. For a common-environment cluster, these procedures should perform the same system startup and login functions for each computer. For a multiple-environment cluster, you may want some startup commands to remain specific to certain computers, as described in Section 4.4.2.

In a common-environment cluster, you can prepare SYSTARTUP\_V5.COM procedures using one of the following methods:

- In each computer's SYS\$SPECIFIC:[SYSMGR] directory, set up a SYSTARTUP\_V5.COM procedure that performs computer-specific startup functions and then invokes a common SYSTARTUP procedure, typically named SYSTARTUP\_COMMON.COM. This procedure is usually located in the SYS\$COMMON:[SYSMGR] directory on a common system disk but can reside on any disk, provided that the disk is cluster accessible and is mounted when the procedure is invoked.
- After setting up computer-specific SYSTARTUP\_V5.COM procedures, create a copy of SYSTARTUP\_COMMON.COM for each computer. However, if you use multiple SYSTARTUP\_COMMON.COM files, you must update all copies whenever you make changes.

## Preparing the Cluster Operating Environment

### 4.4 Coordinating Startup Command Procedures

To set up a common SYLOGIN procedure, define the logical name SYS\$SYLOGIN on each computer to be the full file specification of the procedure. If the common SYLOGIN file is on a cluster-accessible disk, include the command that defines SYS\$SYLOGIN in your common SYSTARTUP command file. If the computers use separate duplicate copies of SYLOGIN.COM, include the definition in each computer's specific startup procedure.

For example, the following command defines SYS\$SYLOGIN to be the common file [SYSMGR]SYLOGIN.COM on the cluster-accessible disk WORK5:

```
$ DEFINE/SYSTEM/EXEC SYSSYLOGIN WORK5:[SYSMGR]SYLOGIN
```

Certain startup functions, even in a common-environment cluster, are computer specific. Therefore, you must include commands in the computer-specific startup procedure on each computer to do the following:

- Set up dual-ported and local disks
- Load device drivers
- Set up local terminals and terminal server access
- Invoke the common SYSTARTUP command procedure

Section 4.4.1 and Section 4.4.2 present guidelines for using common and computer-specific command procedures to build a cluster environment.

#### 4.4.1 Building Startup Procedures for a Common-Environment Cluster

The first step in preparing a common-environment cluster is to build cluster common SYSTARTUP and SYLOGIN command procedures. In a common-environment cluster, each computer executes the common procedures at startup time to define the same operating environment. Because each computer is set up with the common procedure, users can work in the same operating environment on any VAXcluster computer.

##### 4.4.1.1 Procedures for Existing Computers

To build procedures for a cluster in which existing computers are to be combined in a VAXcluster system, you should compare both the computer-specific SYSTARTUP and SYLOGIN command procedures on each computer and make any adjustments required. For example, you can compare the procedures from each computer and include commands that define the same logical names in your common SYSTARTUP command file.

An easy method of comparing the existing procedures and creating common versions is to log in to each computer (in the single-computer environment) and print the existing SYSTARTUP and SYLOGIN command procedure files. You can then use the file listings to compare the procedures. After you have chosen which commands to make common, you can build the common procedures on one of the VAXcluster computers.

##### 4.4.1.2 Procedures for Newly Installed Computers

The strategy for clusters being formed from newly installed VMS systems is basically the same as that used for clusters that are to include previously installed systems: include common elements in a common command procedure file (for example, SYSTARTUP\_COMMON.COM). With newly installed systems, however, the SYSTARTUP and SYLOGIN command procedure files are empty. Therefore, you must build the common procedures from scratch.



## Preparing the Cluster Operating Environment

### 4.4 Coordinating Startup Command Procedures

For example, you could build a common startup command procedure named SYSTARTUP\_COMMON.COM and include the commands that you want to be common to all computers. You must decide which of the following elements you want to include in the common procedure:

- Commands that install images.
- Commands that define logical names; for example, the logical name that refers to the location of SYLOGIN.COM.
- Commands that set up queues. (See Chapter 6 for information on setting up cluster queues.)
- Commands that set up and mount physically accessible mass storage devices. (See Chapter 5 for information on setting up cluster disks.)
- Commands that perform any other common startup functions. (See the *Guide to Setting Up a VMS System* for more information on startup command procedures.)

To build a common SYLOGIN.COM command file, include in the file commands that define clusterwide logical names and symbols.

You can include commands that mount cluster-accessible storage devices and set up queues in the common SYSTARTUP procedure or in separate command files (such as MSCPMOUNT.COM and STARTQ.COM) that are invoked by the common procedure. However, because such commands are computer specific, they must be executed by the local computer. Therefore, you must use conditional logic to control their execution. Sample command files for mounting storage devices and setting up queues are described in Chapter 5 and Chapter 6, respectively.

#### 4.4.2 Building Startup Procedures for a Multiple-Environment Cluster

To build SYSTARTUP and SYLOGIN command files for a multiple-environment cluster, include in the files elements that you want to remain unique to a computer, such as commands to define computer-specific logical names and symbols. These files must be placed in the SYS\$SPECIFIC root on each computer.

For example, consider a three-member cluster consisting of computers JUPITR, SATURN, and PLUTO. The timesharing environments on JUPITR and SATURN are the same. However, PLUTO runs applications for a specific user group. In this cluster, you would create common SYSTARTUP and SYLOGIN command procedures for JUPITR and SATURN that define identical environments on these computers. But the command procedures for PLUTO would be different; they would include commands to define PLUTO's special application environment.

### 4.5 Coordinating System Files for a Common-Environment Cluster

To prepare a common VAXcluster user environment, you must coordinate the following system files:

- SYSUAF.DAT
- NETPROXY.DAT
- RIGHTSLIST.DAT
- VMSMAIL\_PROFILE.DATA
- NETNODE\_REMOTE.DAT
- QMAN\$MASTER.DAT

## Preparing the Cluster Operating Environment

### 4.5 Coordinating System Files for a Common-Environment Cluster

- SYS\$QUEUE\_MANAGER.QMAN\$QUEUES
- SYS\$QUEUE\_MANAGER.QMAN\$JOURNAL

These files, which are part of the VMS operating system, control such functions as user logins, proxy login access, mail, and access to files and job queues. By coordinating these files, you can define either a common-environment or a multiple-environment cluster.

In a common-environment cluster, you use a common version of each system file and place the files in the SYS\$COMMON:[SYSEXE] directory on a common system disk or on a disk that is mounted by all cluster nodes (see Section 4.5.4). In contrast, in a multiple-environment cluster, you would use computer-specific versions of the files and place the files in each computer's SYS\$SPECIFIC:[SYSEXE] directory.

Section 4.5.1 describes procedures for coordinating user accounts in common SYSUAF.DAT and NETPROXY.DAT files. Section 4.5.2 and Section 4.5.3 describe procedures for preparing the cluster RIGHTSLIST and VMSMAIL\_PROFILE database files, respectively. For detailed information on queue management, refer to Chapter 6. The NETNODE\_REMOTE.DAT file is described in Section 4.3.

---

#### Note

---

If you want to set up a common-environment cluster with more than one common system disk, you must coordinate files on each disk and ensure that the disks are mounted with each cluster reboot. Refer to Section 4.5.4 for instructions.

---

#### 4.5.1 Coordinating User Accounts

In a common-environment cluster, you must coordinate the user accounts from each computer and build common versions of the following files:

- SYSUAF.DAT
- NETPROXY.DAT

If you are setting up a common-environment cluster that consists of newly installed systems, you can follow the instructions in the *Guide to Setting Up a VMS System* to build these files. Because the SYSUAF.DAT file on new VMS systems is empty except for the Digital-supplied accounts, very little coordination is necessary.

However, if the cluster will include one or more computers that have been running with computer-specific SYSUAF.DAT and NETPROXY.DAT files, you must create common versions of the files. Procedures for creating a common SYSUAF.DAT file from computer-specific files are described in Appendix B.

Procedures for creating a common NETPROXY.DAT file are basically the same as those for creating a common SYSUAF.DAT file, except that less coordination is needed when you merge the individual NETPROXY.DAT files. For example, user identification codes (UICs) are not used in the NETPROXY records and therefore need not be coordinated. You should decide which existing proxy login records you want to keep and include these records in the common NETPROXY.DAT file.

## Preparing the Cluster Operating Environment

### 4.5 Coordinating System Files for a Common-Environment Cluster

Once you have prepared SYSUAF.DAT and NETPROXY.DAT files, you can set up each of them either as a common file on a cluster-accessible disk or as separate duplicate files. Note, however, that if you choose to use duplicate files, you must update all copies whenever you make changes.

If your cluster is running from one common system disk, make sure that SYSUAF.DAT and NETPROXY.DAT are located in the directory SYS\$COMMON:[SYSEXE].

If your cluster is running from any other system disk configuration, you must decide where to locate SYSUAF.DAT and NETPROXY.DAT. Once you have placed these two files in a directory, you must define clusterwide logical names to point to them.

Assume that disk WORK5 is shared by all computers in the cluster and that it contains cluster common SYSUAF.DAT and NETPROXY.DAT files. The following commands define system logical names that point to the location of the common files:

```
$ DEFINE/SYSTEM/EXEC SYSUAF WORK5:[SYSEXE]SYSUAF
$ DEFINE/SYSTEM/EXEC NETPROXY WORK5:[SYSEXE]NETPROXY
```

You must add the DEFINE commands to your common SYSTARTUP command file. After you have copied the files to the appropriate directory on the cluster-accessible disk, you should delete these files from the system disk.

#### 4.5.2 Preparing the Rights Database

The rights database file, RIGHTSLIST.DAT, associates users of the system or cluster with special names called **identifiers**. This file is the basis of a VMS protection scheme that uses access control lists (ACLs). For a detailed description of this scheme, see the *Guide to VMS System Security*. For information about how the rights database is created, refer to the *VMS Authorize Utility Manual*.

The cluster manager or security manager maintains the rights database, adding and removing identifiers as needed. By allowing groups of users to hold identifiers, the manager can create a different kind of group designation than the one based on UICs. This alternative grouping allows the holders of the identifier to make more efficient use of resources. It also permits each user to be a member of multiple overlapping groups.

If your cluster is running from one common system disk, the installation or upgrade procedure places the RIGHTSLIST.DAT file in the directory SYS\$COMMON:[SYSEXE]. No further action is required on your part.

If your cluster is running from any other system disk configuration, copy SYS\$SYSTEM:RIGHTSLIST.DAT to the directory in which you placed the SYSUAF.DAT and NETPROXY.DAT files. Then define a clusterwide logical name for the RIGHTSLIST.DAT file. For example:

```
$ DEFINE/SYSTEM/EXEC RIGHTSLIST -
    WORK5:[SYSEXE]RIGHTSLIST
```

You must also add this DEFINE command to your common SYSTARTUP command file.

## Preparing the Cluster Operating Environment

### 4.5 Coordinating System Files for a Common-Environment Cluster

#### 4.5.3 Preparing the MAIL Database

In a common-environment cluster, you may want to prepare a common mail database to allow users to use the Mail Utility (MAIL) to send and read their MAIL messages from any computer in the cluster.

Each time MAIL executes in a single-system environment, it accesses a database file named SYS\$SYSTEM:VMSMAIL\_PROFILE.DATA. To set up VMSMAIL\_PROFILE.DATA as a common file, define the logical name VMSMAIL\_PROFILE to be the complete file specification of the common file by specifying the DEFINE command in the following format:

```
$ DEFINE/SYSTEM/EXEC VMSMAIL_PROFILE file-spec
```

You must make sure that you define the logical name before you invoke MAIL for the first time. When invoked for the first time, MAIL creates the database file, VMSMAIL\_PROFILE.DATA, in SYS\$SYSTEM by default. By defining VMSMAIL\_PROFILE to be the location of a common file on a cluster-accessible disk, you cause MAIL to create and use that file.

If your cluster is running from one common system disk, define VMSMAIL\_PROFILE to be SYS\$COMMON:[SYSEXE]VMSMAIL\_PROFILE and invoke the Mail Utility, by entering the following two commands:

```
$ DEFINE/SYSTEM/EXEC VMSMAIL_PROFILE -  
_ $ SYS$COMMON:[SYSEXE]VMSMAIL_PROFILE  
$ MAIL
```

VMSMAIL\_PROFILE.DATA is created in the common system directory. You no longer need to use the logical name or make changes to your common SYSTARTUP command file.

If your cluster is running from any other system disk configuration, you must decide where to locate the common VMSMAIL\_PROFILE.DATA file. (Typically, you would place this file in the same directory in which SYSUAF.DAT and NETPROXY.DAT reside—for example, WORK5:[SYSEXE].) You then define a logical name for the file and invoke the Mail Utility:

```
$ DEFINE/SYSTEM/EXEC VMSMAIL_PROFILE -  
_ WORK5:[SYSEXE]VMSMAIL_PROFILE  
$ MAIL
```

The DEFINE command defines VMSMAIL\_PROFILE.DATA to be a file located in [SYSEXE] on the cluster-accessible disk volume WORK5. The first time MAIL is invoked, VMSMAIL\_PROFILE.DATA is created in WORK5:[SYSEXE]. Subsequently, MAIL uses this file as the database. You must also add the DEFINE command to your common SYSTARTUP command file.

#### 4.5.4 Coordinating Shared System Files in Clusters with Multiple Common System Disks

To prepare a common user environment for a VAXcluster system that includes more than one common system disk, you must coordinate on those disks the system files listed in Section 4.5. In local area and mixed-interconnect clusters, you must also coordinate the file SYS\$MANAGER:NETNODE\_UPDATE.COM, which is described in Section 7.5.1.2.

## Preparing the Cluster Operating Environment

### 4.5 Coordinating System Files for a Common-Environment Cluster

Proceed as follows:

1. Edit the file [VMS\$COMMON.SYSMGR]SYLOGICALS.COM on each system disk and define logical names that specify the location of the cluster common files. For example, if the files will be located on \$1\$DJA16, you could define logical names like the following:

```
$ DEFINE/SYSTEM/EXEC SYSUAF -
    $1$DJA16:[VMS$COMMON.SYSEXE]SYSUAF.DAT
$ DEFINE/SYSTEM/EXEC NETPROXY -
    $1$DJA16:[VMS$COMMON.SYSEXE]NETPROXY.DAT
$ DEFINE/SYSTEM/EXEC RIGHTSLIST -
    $1$DJA16:[VMS$COMMON.SYSEXE]RIGHTSLIST.DAT
$ DEFINE/SYSTEM/EXEC VMSMAIL_PROFILE -
    $1$DJA16:[VMS$COMMON.SYSEXE]VMSMAIL_PROFILE.DATA
$ DEFINE/SYSTEM/EXEC NETNODE_REMOTE -
    $1$DJA16:[VMS$COMMON.SYSEXE]NETNODE_REMOTE.DAT
$ DEFINE/SYSTEM/EXEC NETNODE_UPDATE -
    $1$DJA16:[VMS$COMMON.SYSMGR]NETNODE_UPDATE.COM
$ DEFINE/SYSTEM/EXEC QMAN$MASTER -
    $1$DJA16:[VMS$COMMON.SYSEXE]QMAN$MASTER.DAT
```

2. To ensure that the system disks are correctly mounted with each reboot, follow these steps:
  - a. Copy the file SYS\$EXAMPLES:CLU\_MOUNT\_DISK.COM to the directory [VMS\$COMMON.SYSMGR].
  - b. Edit SYLOGICALS.COM and include commands to mount, with appropriate volume label, the system disk containing the shared files. For example, if the system disk is \$1\$DJA16, you would include a command like the following:

```
$ @SYS$SYSDEVICE:[VMS$COMMON.SYSMGR]CLU_MOUNT_DISK.COM -
    $1$DJA16: volume-label
```

3. When you are ready to start the queuing system, be sure you have moved the queue and journal files to a cluster-available disk. The cluster common system disk is a good choice if the disk has sufficient space. Execute the following command:

```
$ START/QUEUE/MANAGER -
    $1$DJA16:[VMS$COMMON.SYSEXE]
```

When you execute CLUSTER\_CONFIG.COM to add computers to a cluster with more than one common system disk, a different device name must be used for each system disk on which computers are added. For this reason, CLUSTER\_CONFIG.COM supplies as a default device name the logical volume name (for example, DISK\$MARS\_SYS1) of SYS\$SYSDEVICE: on the local system.

Using different device names ensures that each computer added has a unique root directory specification, even if the system disks contain roots with the same name—for example, DISK\$MARS\_SYS1:[SYS10] and DISK\$MARS\_SYS2:[SYS10].



---

# Setting Up and Managing Cluster Disks and Tapes

A VAXcluster system can include two types of disk and tape devices:

- Restricted-access devices, which are accessible only by the local computer or computers to which they are directly connected
- Cluster-accessible devices, which are accessible by any computer in the cluster

As system manager, you are responsible for planning, organizing, and setting up the proper cluster device configuration for your site. You must decide which disk and tape devices should have access restricted to the local computer and which should be accessible to the cluster. For example, you may want to restrict access to a particular device to the users on the computer that is directly connected to the device. Alternatively, you may decide to set up a disk as a cluster-accessible device, so that any user on any computer can allocate and use it.

You can use the information in this chapter to plan and set up your disk and tape configuration. Topics include the following:

- Cluster-accessible disk and tape devices
- Cluster device-naming conventions
- Shared disks
- Configuring cluster disk and tape devices
- Rebuilding cluster disks

## 5.1 Cluster-Accessible Disk and Tape Devices

A cluster-accessible device is a disk or tape that multiple computers in the cluster can recognize and access. The following types of devices are cluster accessible:

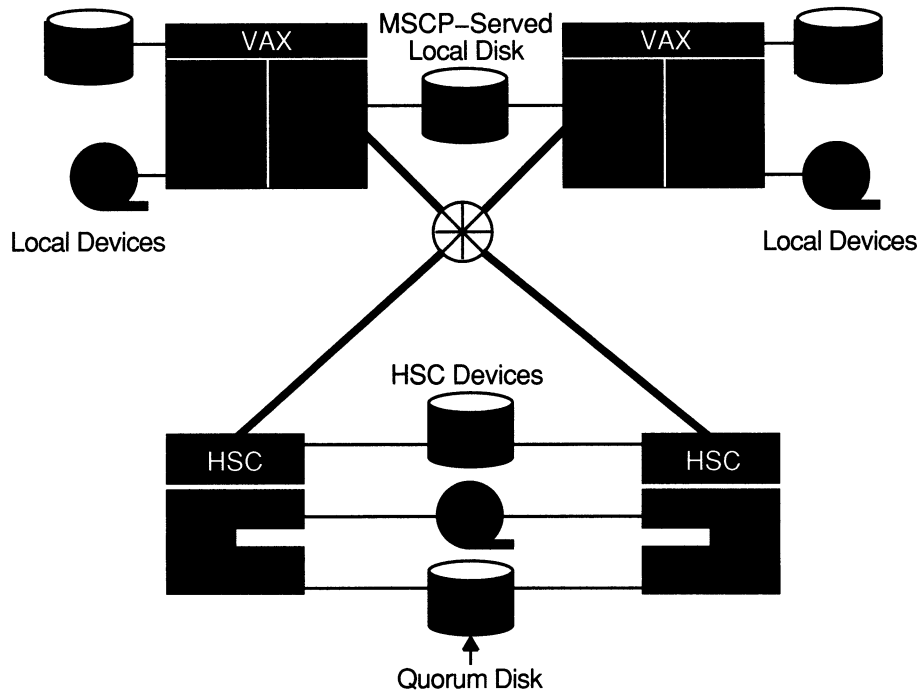
- HSC disks and tapes
- Served disks and tapes
- Dual-pathed disks and tapes
- DSSI disks and tapes

Figure 5–1 illustrates how disks and tapes might be configured in a typical CI-based cluster. The HSC disks and tapes and the dual-ported MSCP-served local disk are considered cluster accessible.

# Setting Up and Managing Cluster Disks and Tapes

## 5.1 Cluster-Accessible Disk and Tape Devices

Figure 5-1 CI-Based Configuration with Shared Disks and Tapes



ZK-1637-GE

### 5.1.1 HSC Disks and Tapes

An HSC disk or tape is a DIGITAL Storage Architecture (DSA) device that is connected to an HSC subsystem. If an HSC subsystem is connected in a cluster, its disks and tapes are accessible by all VAXcluster nodes connected to the same star coupler. These devices can be served to allow access by satellite nodes or by other large VAX computers not connected to the same star coupler. You can set up HSC storage devices (including a quorum disk) to be dual pathed between two HSC subsystems, as shown in Figure 5-1. Dual-pathed disks and tapes are described in Section 5.1.3.

### 5.1.2 Served Disks and Tapes

The mass storage control protocol (MSCP) server and tape mass storage control protocol (TMSCP) server are used to communicate between a computer and a DSA controller. The servers enable a computer to make locally connected devices, such as MASSBUS, UNIBUS, and BI devices, available to all other cluster members.

Locally connected disks and tapes are not automatically cluster accessible. Access to these devices is restricted to the local computer unless you explicitly set them up as cluster accessible using the MSCP server for disks or the TMSCP server for tapes.



## Setting Up and Managing Cluster Disks and Tapes

### 5.1 Cluster-Accessible Disk and Tape Devices

#### 5.1.2.1 MSCP and TMSCP Server Functions

To make a disk or tape accessible to all VAXcluster computers, the MSCP or TMSCP server must be loaded on the local computer, and it must be instructed to make the device available across the cluster. MSCP server functions are enabled with the SYSGEN parameters `MSCP_LOAD` and `MSCP_SERVE_ALL` and TMSCP server functions are enabled with the `TMSCP_LOAD` parameter (see Table 5–1). By specifying appropriate values for these parameters in a computer's `MODPARAMS.DAT` file and then running `AUTOGEN` to reboot the computer, you enable the computer to serve all suitable devices to the cluster early in the boot sequence. (You can also use the `CLUSTER_CONFIG.COM CHANGE` function to perform these operations for disks.) The served devices thus become accessible with minimal interruption whenever the serving computer reboots. Further, the servers automatically serve any suitable device that is added to the system later. For example, if new drives are attached to an HSC subsystem, the devices become available within seconds after the cables are connected.

Table 5–1 summarizes the SYSGEN parameter values you can specify to configure the MSCP and TMSCP servers. Initial values are determined by your responses when you execute the VMS installation or upgrade procedure, or when you execute the `CLUSTER_CONFIG.COM` command procedure described in Chapter 7 to set up your configuration. Note that if you later change the values, you must reboot the computer on which you changed the values before the new values can take effect (see Section 7.5.3).

**Table 5–1 Specifying Values for `MSCP_LOAD`, `MSCP_SERVE_ALL`, and `TMSCP_LOAD` Parameters**

Parameter	Value	Function
<code>MSCP_LOAD</code>	0	Do not load the MSCP server (default value).
	1	Load the MSCP server with attributes specified by <code>MSCP_SERVE_ALL</code> parameter, using the default CPU load capacity.
	>1	Load the MSCP server with attributes specified by the <code>MSCP_SERVE_ALL</code> parameter. Use this value as the CPU load capacity.
<code>MSCP_SERVE_ALL</code>	0	Do not serve any disks (default value).
	1	Serve all available disks.
	2	Serve only locally connected (non-HSC) disks.
<code>TMSCP_LOAD</code>	0	Do not load the TMSCP server and do not serve any tapes (default value).
	1	Load the TMSCP server and serve all available tapes, including all local tapes and all multihost tapes with matching <code>TAPE_ALLOCLASS</code> value.

For a sample use of these parameters, see the discussion of Figure 5–4 in Section 5.2.2.

## Setting Up and Managing Cluster Disks and Tapes

### 5.1 Cluster-Accessible Disk and Tape Devices

#### 5.1.2.2 MSCP Load Sharing

MSCP servers monitor their I/O traffic and periodically calculate a load available rating to indicate available capacity for I/O requests. (TMSCP servers do not perform this monitoring function.)

Load available is calculated by counting the read and write requests sent to the server and periodically converting this value to requests per second and subtracting this calculated value from the server's load capacity (also specified in requests per second).

This information is communicated to the MSCP class driver (DUDRIVER and DSDRIVER). When a disk is mounted or a failover occurs, the class driver selects the server with the highest load available rating to access the disk.

Load balancing is enabled and controlled by the SYSGEN parameters MSCP\_LOAD and MSCP\_SERVE\_ALL. In most cases, the values established by CLUSTER\_CONFIG.COM are appropriate.

The MSCP\_SERVE\_ALL parameter determines whether the server participates in load sharing. If the parameter is set to 2 (serve only local disks), the server does not monitor its I/O traffic and does not participate in load balancing. Other valid settings for MSCP\_SERVE\_ALL (0 or 1) result in the server monitoring I/O traffic and communicating load available information to the class drivers.

The MSCP\_LOAD parameter is used to communicate load capacity to the server, in addition to its existing function of controlling the loading of the MSCP server. If the parameter is set to 1, the MSCP server is loaded and its load capacity is set to a default value based on CPU type. If MSCP\_LOAD is set to a value greater than 1, the server is loaded and its load capacity set to that value. Setting MSCP\_LOAD to zero disables loading of the MSCP server.

#### 5.1.3 Dual-Pathed Disks and Tapes

A **dual-pathed device** is a disk or tape that is accessible to all the computers in the cluster, not just to the computers that are physically connected to the device. The term dual-pathed refers to the two paths through which computers can access a device to which they are not directly connected. If one path fails, the device is accessed over the other path. (Note that with a dual-ported MASSBUS device, a computer directly connected to the device always accesses it locally.)

Dual-pathed devices can be any of the following:

- Dual-ported HSC disks or tapes
- Dual-pathed DSA disks or tapes on local UDA/KDA/KDB controllers
- DSSI-connected integrated storage elements (ISEs)
- Dual-ported MASSBUS disks

##### 5.1.3.1 Dual-Ported HSC Disks and Tapes

By design, HSC disks and tapes are accessible by all VAXcluster nodes connected to the same star coupler. Therefore, if they are dual ported, they are automatically dual pathed. CI-connected computers can access a dual-pathed HSC device by way of a path through either HSC subsystem connected to the device. If one HSC subsystem fails, access fails over to the other subsystem.

For each dual-ported HSC device, you can control failover to a specific port using the port select buttons on the front of each drive. By pressing either port select button (A or B) on a particular drive, you can cause the device to fail over to the specified port.

# Setting Up and Managing Cluster Disks and Tapes

## 5.1 Cluster-Accessible Disk and Tape Devices

With the port select buttons, you can select alternate ports to balance the device controller workload between two HSC subsystems. For example, you can set half of your disks to use port A and set the other half to use port B. The port select buttons also enable you to fail over all the devices to an alternate port manually when you anticipate the shutdown of one of the HSC subsystems.

Additionally, you can use software controllable port selection, as described in Section 5.1.3.3.

### 5.1.3.2 Dual-Pathed DSA Disks and Tapes on Local UDA/KDA/KDB Controllers

A dual-ported DSA disk or tape can be failed over between the two computers that serve it to the cluster, provided that:

- The same device controller letter is generated and the same allocation class is specified on each computer, with the result that the device has the same name on both systems, and
- Both computers are running the MSCP and/or TMSCP server.

---

#### Caution

---

Failure to observe these requirements can endanger data integrity.

---

However, because a DSA device can be on line to only one controller at a time, only one of the computers can use its local connection to the device. The second computer accesses the device through the MSCP or TMSCP server. If the computer that is currently serving the device fails, the other computer detects the failure and fails the device over to its local connection. The device is thereby made available to the cluster once more.

---

#### Note

---

A dual-ported DSA disk cannot be used as a system disk for a directly connected CPU.

---

Software controllable port selection is described in Section 5.1.3.3.

### 5.1.3.3 Preferred Path Specification

The VMS operating system supports specifying a preferred path for DSA disks, including RA-series disks and disks that are accessed through the MSCP server. (This function is not available for tapes.) If a preferred path is specified for a disk, the MSCP disk class drivers (DUDRIVER and DSDRIVER) use that path as their first attempt to locate the disk and bring it on line with a DCL MOUNT command or failover of an already mounted disk.

In addition, you can initiate failover of a mounted disk to force the disk to the preferred path or to use loadsharing information for disks accessed by MSCP servers.

The preferred path is specified by a \$QIO function (IO\$\_SETPRFPATH), with the P1 parameter containing the address of a counted ASCII string (.ASCIC). This string is the node name of the HSC or VMS system that is to be the preferred path. The node name must match an existing node that is known to the local node, and, if it is a VMS system, it must be running the MSCP server. This function does not move the disk to the preferred path. For more information on

## Setting Up and Managing Cluster Disks and Tapes

### 5.1 Cluster-Accessible Disk and Tape Devices

the use of the `IO$_SETPRFPATH` function, refer to the *VMS I/O User's Reference Manual: Part I*. See also the `PREFER.MAR` program in `SYS$EXAMPLES`.

#### 5.1.3.4 DSSI-Connected ISEs

In a DSSI VAXcluster system, a DSSI bus can connect as many as eight nodes that can be ISEs or host CPU interfaces. In a dual-host configuration with two CPUs sharing a DSSI bus, as many as six ISEs can be connected between DSSI controllers (see Section 2.5). Simultaneously accessible to both servers, these storage elements can be served to satellites. If one boot server fails, access fails over to the other server and applications continue to run. Any DSSI-connected ISE can be used as a system disk. Note that, because most failures occur in system enclosures, you should try to locate the system disk in a storage expansion box, which has a dedicated power supply.

Like configurations based on HSC subsystems, DSSI configurations provide multihost access, flexibility, and expansion potential.

#### 5.1.3.5 Dual-Ported MASSBUS Disks

A dual-ported MASSBUS disk can be connected between two computers, if it has the same controller letter and allocation class on both.

Before mounting the disk, enter the DCL command `SET DEVICE` in the following format on both computers:

```
SET DEVICE/DUAL_PORT device-name
```

---

#### Note

---

A MASSBUS disk can be used either as a dual-ported disk or as a system disk, but not as both.

---

In clusters with more than two computers, you can set up a dual-ported MASSBUS disk to be cluster accessible through the MSCP server on either or both computers to which the disk is connected. Be sure, however, *not* to use the `SYSGEN` commands `AUTOCONFIGURE` or `CONFIGURE` to configure a dual-ported MASSBUS disk that is already available on the computer through the MSCP server. Establishing a local connection to the disk when a remote path is already known creates two uncoordinated paths to the same disk. Use of these two paths can endanger data integrity on any disk that is mounted on the drive.

If the local path to the disk is not found during the system bootstrap procedure, the MSCP server path from the remote computer is the only available access to the drive. The local path is not found during a boot if any of the following conditions exist:

- The port select switch for the drive is not enabled for the local computer.
- The disk, cable, or adapter hardware for the local path is broken.
- There is sufficient activity on the other port to “mask” the existence of the port.
- The computer is booted in such a way that the `SYSGEN` command `AUTOCONFIGURE ALL` in the site-independent startup procedure (`SYS$SYSTEM:STARTUP.COM`) was not executed.

# Setting Up and Managing Cluster Disks and Tapes

## 5.1 Cluster-Accessible Disk and Tape Devices

Use of the disk is still possible through the MSCP server path.

---

### Caution

---

Under these conditions, do not attempt to add the local path back into the system I/O database using the SYSGEN commands AUTOCONFIGURE or CONFIGURE. SYSGEN is currently unable to detect the presence of the disk's MSCP path and would incorrectly build a second set of data structures to describe it. Subsequent events could lead to incompatible and uncoordinated file operations, which might endanger data integrity.

---

To recover the local path to the disk, you must reboot the computer connected to that local path.

Note that if the disk is not dual ported or is never MSCP served on the remote computer, this restriction does not apply.

## 5.2 Cluster Device-Naming Conventions

To manage cluster devices properly, you must understand the conventions used to identify them. Every cluster device is identified by a unique name, which provides a reliable way to access it in the cluster.

Disk and tape devices that are local to a VAXcluster computer can be accessed by that computer through the traditional device name (for example, DUA1 or MUB6) or through a cluster device name in the format *node-name\$device-name* (for example, JUPITR\$DUA1).

However, a device that is dual pathed between two computers must be identified by a unique, path-independent name that includes an **allocation class**. The allocation class is a numeric value from 0 to 255 that is used to create a device name in one of the following formats:

```
$allocation-class$device-name  
$tape-allocation-class$device-name
```

For example, the allocation class device name \$1\$DJA17 identifies a disk that is dual ported between two computers or HSC subsystems that both have an allocation class value of 1. Similarly, the allocation class device name \$1\$MUA12 identifies a dual-ported tape.

Each time a computer that is not directly connected to such a device tries to access it, the choice of which path to take is made arbitrarily; no one path is ever guaranteed. Because the access path is chosen without regard to the names of the computers or HSC subsystems serving the device, an allocation class device name is required to identify the device uniquely.

### 5.2.1 Rules for Specifying Allocation Class Values

Allocation classes play an important role in determining strategies for configuring and naming disks and tapes. In fact, the VMS operating system uses allocation class values above all other available information when determining the configuration of cluster devices.

The following rules apply for specifying allocation class values:

- Computers or HSC subsystems to which a dual-pathed device is connected must have the same nonzero allocation class value.

## Setting Up and Managing Cluster Disks and Tapes

### 5.2 Cluster Device-Naming Conventions

- All cluster-accessible devices on computers with a nonzero allocation class value must have unique names. For example, if two computers have the same allocation class value, it is invalid for both computers to have a disk named DJA0 or a tape named MUA0. This restriction also applies to HSC subsystems.
- Single-ported devices with an allocation class value of zero can have the same unit number on different computers.

Zero is the default allocation class value. Any computer in a cluster that is serving either multihost (HSC or DSSI) or dual-path disks or tapes must be assigned a nonzero allocation class value. A zero allocation class value may only be appropriate when serving a local single-path disk. In a mixed-interconnect cluster, all of the following must have a nonzero allocation class value:

- HSC subsystems
- Computers that serve HSC disks and tapes
- DSSI ISEs
- Computers connected to dual-path disks and tapes
- Members of shadow sets

---

#### Caution

---

Failure to set allocation class values correctly can endanger data integrity and cause locking conflicts that suspend normal cluster operations.

---

The disk allocation class and the tape allocation class are not required to have the same value on a node in a VAXcluster. It is recommended that they have the same value to reduce confusion.

To assign an allocation class value to a VAXcluster computer that supports dual-path devices, specify the value with the SYSGEN parameters ALLOCLASS and TAPE\_ALLOCLASS. To assign disk allocation class for an HSC subsystem, specify the value using the HSC console to enter a command in the following format, where *n* is the allocation class value:

```
SET ALLOCATE DISK n
```

To assign a tape allocation class, enter a command in the following format:

```
SET ALLOCATE TAPE n
```

For complete information on HSC console commands, refer to the HSC hardware documentation.

For more information on setting the HSC allocation class, see Section 7.6.3. For additional information about changing a DSSI subsystem allocation class, see Section 7.6.4.

Note that multihost disks and tapes must have matching allocation class values on both the CPU and the controller to be served.

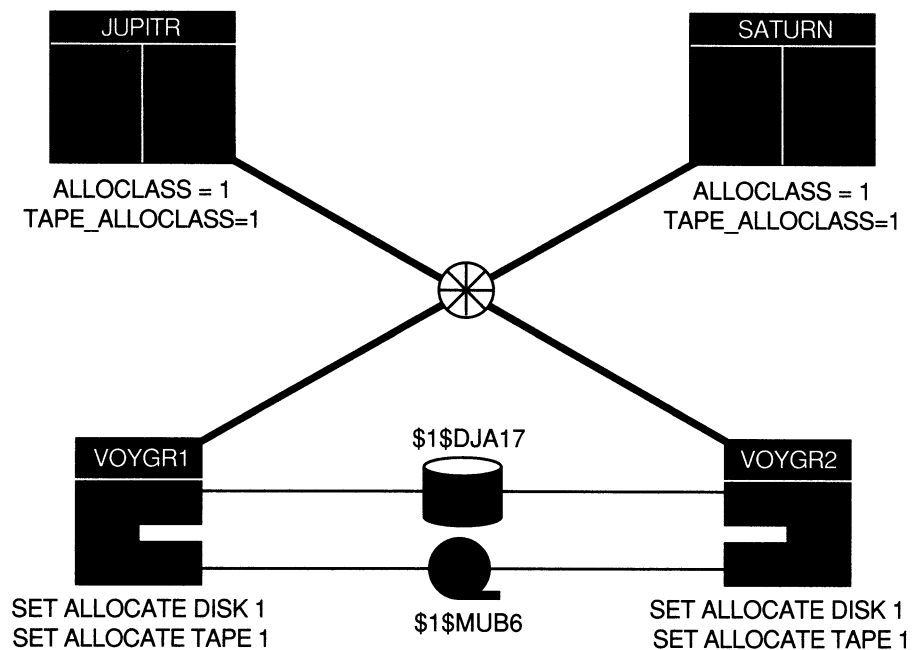
### 5.2.2 Sample Configurations with Named Devices

Figure 5–2 and Figure 5–3 show how cluster device names are specified for the following:

- Dual-pathed HSC devices
- Dual-pathed DSA devices

Figure 5–2 shows a VAXcluster with a dual-pathed HSC disk. Note that the allocation class value (1) is the same on VAX computers JUPITR and SATURN, and that the disk's device name (\$1\$DJA17) is constructed using that value. JUPITR and SATURN can access the disk through either of the HSC subsystems VOYGR1 or VOYGR2. If this configuration includes a dual-pathed HSC tape, the tape's device name is constructed in the same way (for example, \$1\$MUB6).

Figure 5–2 VAXcluster with Dual-Pathed HSC Disk and Tape



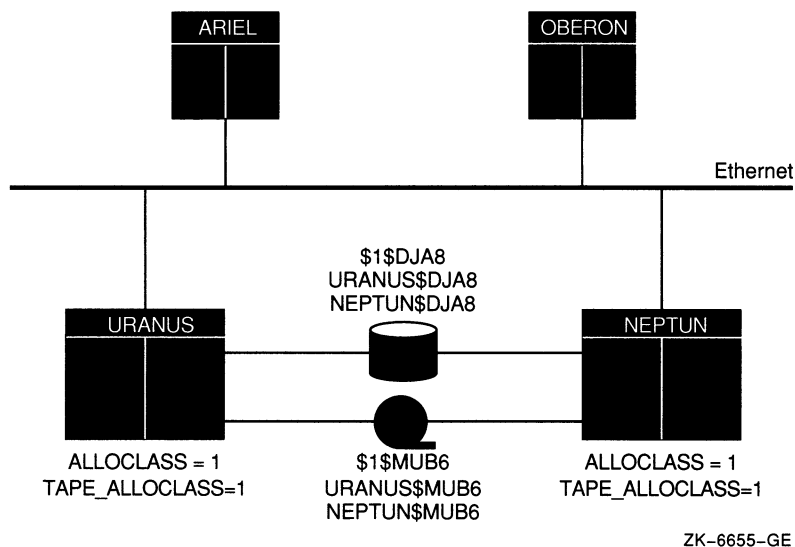
ZK-6656-GE

## Setting Up and Managing Cluster Disks and Tapes

### 5.2 Cluster Device-Naming Conventions

Figure 5-3 shows a VAXcluster with a dual-pathed DSA disk.

**Figure 5-3 VAXcluster with Dual-Pathed DSA Disk and Tape**



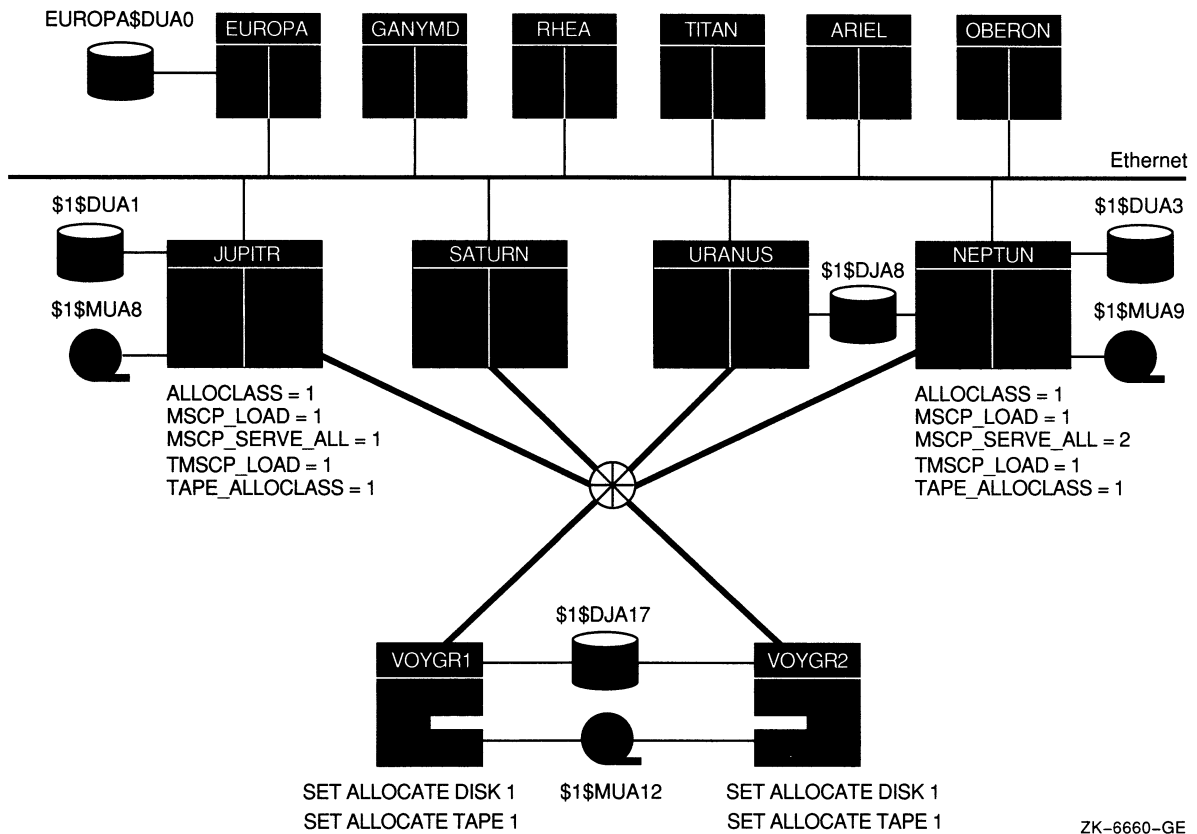
URANUS and NEPTUN can access the disk either locally or through the other computer's MSCP server. However, when satellites ARIEL and OBERON access the disk, they arbitrarily choose a path through either URANUS or NEPTUN. If the satellites try to access the disk by using the computer-specific device name URANUS\$DJA8, and the disk is not currently accessible through URANUS, access will fail. But if the satellites use the allocation class device name \$1\$DJA8, they can access the disk through NEPTUN. Similarly, if the configuration includes a dual-pathed DSA tape with the tape allocation class name \$1\$MUB6, satellites can access the tape through NEPTUN. *As a general rule, you should always use a path-independent allocation class (or tape allocation class) device name to identify dual-pathed cluster devices.*

Figure 5-4 shows how device names are typically specified in a mixed-interconnect cluster. This figure also shows how relevant SYSGEN parameter values are set in each CI-connected computer's MODPARAMS.DAT file. Note that the values shown for JUPITR are the same for SATURN and URANUS, but that NEPTUN has a different value for MSCP\_SERVE\_ALL.



## Setting Up and Managing Cluster Disks and Tapes 5.2 Cluster Device-Naming Conventions

**Figure 5-4 Device Names in a Mixed-Interconnect Cluster**



In this configuration, a disk and a tape are dual-pathed to the HSC subsystems named VOYGR1 and VOYGR2; these subsystems are connected to JUPITR, SATURN, URANUS, and NEPTUN through the star coupler. The MSCP and TMSCP servers are loaded on all four computers (MSCP\_LOAD = 1, TMSCP\_LOAD = 1) and the ALLOCLASS and TAPE\_ALLOCLASS parameters are set to the same value (1) on these computers and on both HSC subsystems. But MSCP\_SERVE\_ALL is set to 1 only on JUPITR, SATURN, and URANUS. Therefore, only these three computers can serve the devices on VOYGR1 and VOYGR2 to satellites. Because MSCP\_SERVE\_ALL is set to 2 on NEPTUN, NEPTUN can serve only its local disks.

The HSC-connected disk and tape have allocation class names in the form \$1\$ddcu. For example, the disk DUA17 is named \$1\$DUA17. On CI-connected computers, VMS software would also recognize the disk as JUPITR\$DUA17 and as either VOYGR1\$DUA17 or VOYGR2\$DUA17. On satellites, it would recognize the disk as JUPITR\$DUA17 or as \$1\$DUA17. This example shows why you should always use an allocation class name like \$1\$DUA17 when you configure cluster devices: the allocation class name is the only name that all computers recognize at all times.

Note that, for optimal availability, two or more CI-connected computers should serve HSC disks and tapes to the cluster.

### 5.3 Shared Disks

A **shared disk** is a disk that is mounted on a cluster-accessible device by one or more VAXcluster computers. Shared disks play a key role in common-environment clusters, because when you place system files or command procedures on a shared disk, computers can share a single copy of each common file (see Chapter 4). Note, however, that a shared disk is a single point of failure for data access by the computers sharing the disk.

To mount cluster-accessible disks that are to be shared among all computers, specify the same MOUNT command on each computer or specify the MOUNT command with the /CLUSTER qualifier on one or more computers. (Typically, you specify MOUNT/CLUSTER on all computers.) When you execute MOUNT/CLUSTER on one computer, the disk is mounted on every computer that is active in the cluster at the time the command executes. Note that only system or group disks can be mounted across the cluster. Thus, if you specify MOUNT/CLUSTER without the /SYSTEM or /GROUP qualifier, /SYSTEM is assumed. Also note that each cluster disk mounted with the /SYSTEM, /GROUP, or /SHARED qualifiers must have a unique volume label.

If you want to mount a shared disk on some but not all VAXcluster computers, execute the same MOUNT command (without the /CLUSTER qualifier) on each computer that shares the disk.

For example, suppose you want all the computers in a three-member cluster to share a disk named COMPANYDOCS. To share the disk, each of the three computers can execute identical MOUNT commands, or one of the three computers can mount COMPANYDOCS using the MOUNT/CLUSTER command, as follows:

```
$ MOUNT/SYSTEM/CLUSTER/NOASSIST $1SDUA4: COMPANYDOCS
```

If you want just two of the three computers to share the disk, those two computers must both mount the disk with the same MOUNT command. For example:

```
$ MOUNT/SYSTEM/NOASSIST $1SDUA4: COMPANYDOCS
```

To mount the disk at startup time, include the mount command either in a common command procedure that is invoked at startup time or in the computer-specific startup command file.

### 5.4 Configuring Cluster Disks

To configure cluster disks, you can create command procedures to set up and mount them. You may want to include commands that set up and mount cluster disks in a separate command procedure file that is invoked by a site-specific SYSTARTUP procedure. Depending on your cluster environment, you can set up your command procedure in either of the following ways:

- As a separate file specific to each computer in the cluster
- As a common computer-independent file

You can set up the common procedure as a shared file on a shared disk, or you can make copies of the common procedure and store them as separate files. With either method, each computer can invoke the common procedure from the site-specific SYSTARTUP procedure.

## Setting Up and Managing Cluster Disks and Tapes

### 5.4 Configuring Cluster Disks

The MSCPMOUNT.COM file in the SYS\$EXAMPLES directory on your system is a sample common command procedure that contains commands typically used to mount cluster disks. The example includes comments explaining each phase of the procedure.

## 5.5 Rebuilding Cluster Disks

To minimize disk I/O operations (and thus improve performance) when files are created or extended, the VMS file system maintains a cache of preallocated file headers and disk blocks.

If a disk is improperly dismounted—for example, if a system crashes or is removed from a cluster without running SYS\$SYSTEM:SHUTDOWN.COM—this preallocated space becomes temporarily unavailable. When the disk is remounted, MOUNT scans the disk to recover the space as it rebuilds the disk.

On a nonclustered computer, the scan operation merely prolongs the boot process. In a VAXcluster system, however, this operation can degrade response time for all user processes in the cluster. While the scan is in progress on a particular disk, most activity on that disk is blocked. User processes that attempt to read or write to files on the disk can experience delays of several minutes or longer, especially if the disk contains a large number of files or has many users.

Because the rebuild operation can delay access to disks during the startup of any VAXcluster computer, Digital recommends that procedures for mounting cluster disks use the /NOREBUILD qualifier. When MOUNT/NOREBUILD is specified, disks are not scanned to recover lost space, and users experience minimal delays while computers are mounting disks.

System disks are especially critical in this regard, because most system activity requires access to a system disk. When a system disk rebuild is in progress, very little activity is possible on any computer that uses that disk. Unlike other disks, the system disk is automatically mounted early in the boot sequence. If a rebuild is necessary, and if the value of the SYSGEN parameter ACP\_REBLDSYSD is 1, the system disk is rebuilt during the boot sequence. (The default setting of 1 for the SYSGEN parameter ACP\_REBLDSYSD specifies that the system disk should be rebuilt.)

In local area and mixed-interconnect clusters, however, the ACP\_REBLDSYSD parameter should normally be set to zero on all satellites. This setting prevents them from rebuilding a system disk when it is mounted early in the boot sequence and eliminates delays caused by such a rebuild when satellites join the cluster.

In large clusters, a substantial amount of system disk space (some for each computer) might be preallocated to caches, and, if many computers abruptly leave the cluster (for example, during a power failure), this space can become temporarily unavailable. Thus, ACP\_REBLDSYSD on boot servers in local area and mixed-interconnect clusters with many computers should be set to the default value of 1, and procedures that mount disks on the boot servers should use the /REBUILD qualifier. While these measures can make boot server rebooting more noticeable, they ensure that system disk space is available after an unexpected shutdown.

Once the cluster is up and running, system managers can submit one or more batch procedures that execute SET VOLUME/REBUILD commands to recover lost disk space. Such procedures can run at a time when users would not be inconvenienced by the blocked access to disks (for example, between midnight and 6 a.m. each day). Because the SET VOLUME/REBUILD command determines

## Setting Up and Managing Cluster Disks and Tapes

### 5.5 Rebuilding Cluster Disks

whether a rebuild is needed, the procedures can execute the command for each disk that is usually mounted. Note that the procedures run more quickly and cause less delay in disk access if they are executed on powerful computers. Moreover, several such procedures, each of which rebuilds a different set of disks, can be executed simultaneously.

---

#### Caution

---

If either or both MOUNT/NOREBUILD and ACP\_REBLDSYSD = 0 are specified when mounting disks, it is essential to run a procedure with SET VOLUME/REBUILD commands on a regular basis to rebuild the disks. Failure to rebuild disk volumes can result in a loss of free space and in subsequent failures of applications to create or extend files.

---

---

## Setting Up and Managing Cluster Queues

On a standalone computer, print and batch job processing is limited to local devices. In VAXcluster systems, however, computers can share device and processing resources. This ability to share resources allows for better workload balancing because batch and print jobs can be distributed across the cluster.

You control how jobs share device and processing resources in a cluster by setting up and maintaining cluster queues. The strategy you use to set up and manage these queues determines how well workloads are matched to your cluster's resources. For high availability, you can set up autostart queues.

This chapter explains how to set up and manage VAXcluster queues. Topics include the following:

- Clusterwide queues
- Cluster printer queues
- Cluster batch queues
- Using a common command procedure to set up cluster queues

Because queues in a VAXcluster system are established and controlled with the same commands used to manage queues on a standalone computer, the discussions in this chapter assume some knowledge of queue management on a standalone system, as described in the *Guide to Setting Up a VMS System*.

### 6.1 Clusterwide Queues

VMS Version 5.5 supports a new queue manager. For Version 5.5, queues are controlled by a clusterwide queue manager process that accesses the clusterwide queue database for all processes in a cluster. Job controllers, user processes and symbionts all communicate directly with the centralized queue manager through a shared IPC link.

The `START/QUEUE/MANAGER` command starts up the clusterwide queue manager process. By default, any node in the cluster is eligible to run the queue manager process. However, the `/ON` qualifier lets you to specify an ordered list of nodes that can claim the queue manager if the node running the queue manager should exit the cluster.

```
$ START/QUEUE/MANAGER/ON=(node-list)
```

In this command, *node-list* is a list of nodes specified in the order the nodes will claim the queue manager process.

If the node on which the queue manager is running leaves the cluster, the queue manager will fail over to the next available node in the node list specified in the `START/QUEUE/MANAGER` command.

## Setting Up and Managing Cluster Queues

### 6.1 Clusterwide Queues

Specify the asterisk wildcard (\*) as the last node in the node list to indicate that any remaining, unlisted node can start the queue manager, with no preferred order. If you want to exclude certain nodes from being eligible to run the queue manager process, you cannot use the asterisk wildcard character. You cannot specify the asterisk wildcard character as part of a node name.

In the following example, the queue manager process will be started on node JUPITR (if JUPITR is available):

```
$ START/QUEUE/MANAGER/ON=(JUPITR,SATURN,URANUS)
```

If node JUPITR exits the cluster, node SATURN will start up the process (if SATURN is available). During the transition from JUPITR to SATURN, queues on SATURN and URANUS are not stopped. All requests to the queuing system, for example, PRINT, SUBMIT and SHOW ENTRY requests, will complete as expected.

For more information on the queue manager, see the *Guide to Maintaining a VMS System*.

To control queues, the queue manager communicates with a clusterwide queue database. This database makes queues available across the cluster and allows jobs to execute on any queue from any computer, provided that the necessary mass storage volumes can be accessed by the computer on which the job executes.

For clusterwide queues, only one queue database can be used, and it must be located on a disk that is mounted clusterwide or is accessible to the computers participating in the clusterwide queue scheme. Note that performance can be enhanced by locating the file on a shared disk that has a low level of activity.

By default, the following three queue database files are located in SYS\$COMMON:[SYSEXE]:

- QMAN\$MASTER.DAT
- SYS\$QUEUE\_MANAGER.QMAN\$QUEUES
- SYS\$QUEUE\_MANAGER.QMAN\$JOURNAL

However, you can move the files to another location to control which computers share clusterwide queues. You specify the location of the queue and journal files with a command in the following format:

```
$ START/QUEUE/MANAGER device:[directory]
```

You can specify an alternate location for the master file by defining the logical name QMAN\$MASTER with a command in the following format:

```
$ DEFINE/SYSTEM/EXECUTIVE_MODE QMAN$MASTER device:[directory]
```

If the master file is moved, the new device must be mounted in SYLOGICALS.COM. Section 4.5.4 explains this procedure.

The same logical name must be defined on all cluster nodes. You must include the definition in SYLOGICALS.COM. A sample common procedure for setting up a VAXcluster batch and print system is provided in Section 6.4.

For information on storing the queue database files in another location, see the *Guide to Maintaining a VMS System*.

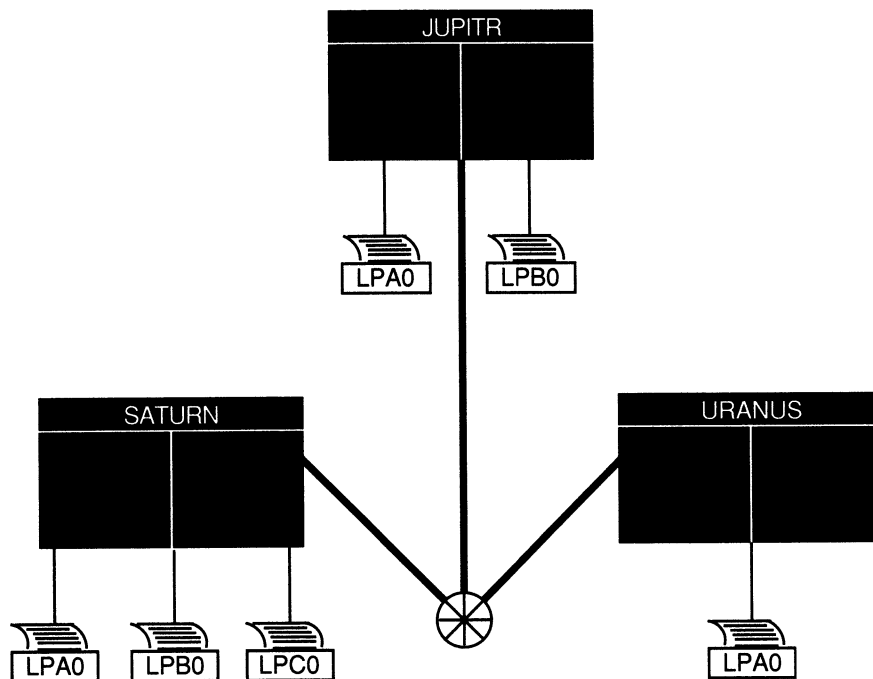
## 6.2 Cluster Printer Queues

To establish printer queues, you must determine the type of queue configuration that best suits your VAXcluster system. You have several alternatives that depend on the number and type of print devices you have on each computer and on how you want print jobs to be processed. For example, make these decisions:

- Which printer queues you want to establish on each computer
- Whether to set up any clusterwide generic queues to distribute print job processing across the cluster
- Whether to set up autostart queues for availability or improved startup time

Once you determine the appropriate strategy for your cluster, you can create your queues. Figure 6-1 shows the printer configuration for a cluster consisting of the active computers JUPITR, SATURN, and URANUS. Section 6.2.1 and Section 6.2.2 describe various methods for establishing and naming the cluster printer queues shown in this configuration.

Figure 6-1 Sample Printer Configuration



ZK-1631-GE

## Setting Up and Managing Cluster Queues

### 6.2 Cluster Printer Queues

#### 6.2.1 Setting Up Printer Queues

You set up VAXcluster printer queues using the same method that you would use for a standalone computer (see the *Guide to Setting Up a VMS System*). However, in a VAXcluster system, you must provide a unique name for each queue you create.

You create a printer queue by specifying the DCL command INITIALIZE/QUEUE in the following format:

```
$ INITIALIZE/QUEUE/ON=node-name::device[/START] queue-name
```

The /ON qualifier specifies the computer and printer to which the queue is assigned. If you specify the /START qualifier, the queue is started.

You can also create an autostart queue and with a list of nodes on which it can run by specifying the DCL command INITIALIZE/QUEUE in the following format:

```
$ INITIALIZE/QUEUE/AUTOSTART_ON=(node-name::device:) queue-name
```

Generic queues cannot be autostart queues. Note that the /ON and /AUTOSTART\_ON qualifiers are mutually exclusive.

The following commands make the local printer queue assignments for JUPITR shown in Figure 6-2 and start the queues:

```
$ INITIALIZE/QUEUE/ON=JUPITR::LPA0/START JUPITR_LPA0
$ INITIALIZE/QUEUE/ON=JUPITR::LPB0/START JUPITR_LPBO
```

#### 6.2.2 Setting Up Clusterwide Generic Printer Queues

The clusterwide queue database enables you to establish generic queues that function throughout the cluster. Jobs queued to clusterwide generic queues are placed in any assigned printer queue that is available, regardless of its location in the cluster. However, the file queued for printing must be accessible to the computer to which the printer is connected.

Figure 6-3 illustrates a clusterwide generic printer queue in which the queues for all LPA0 printers in the cluster are assigned to a clusterwide generic queue named SYS\$PRINT.

The following command initializes and starts the clusterwide generic queue SYS\$PRINT:

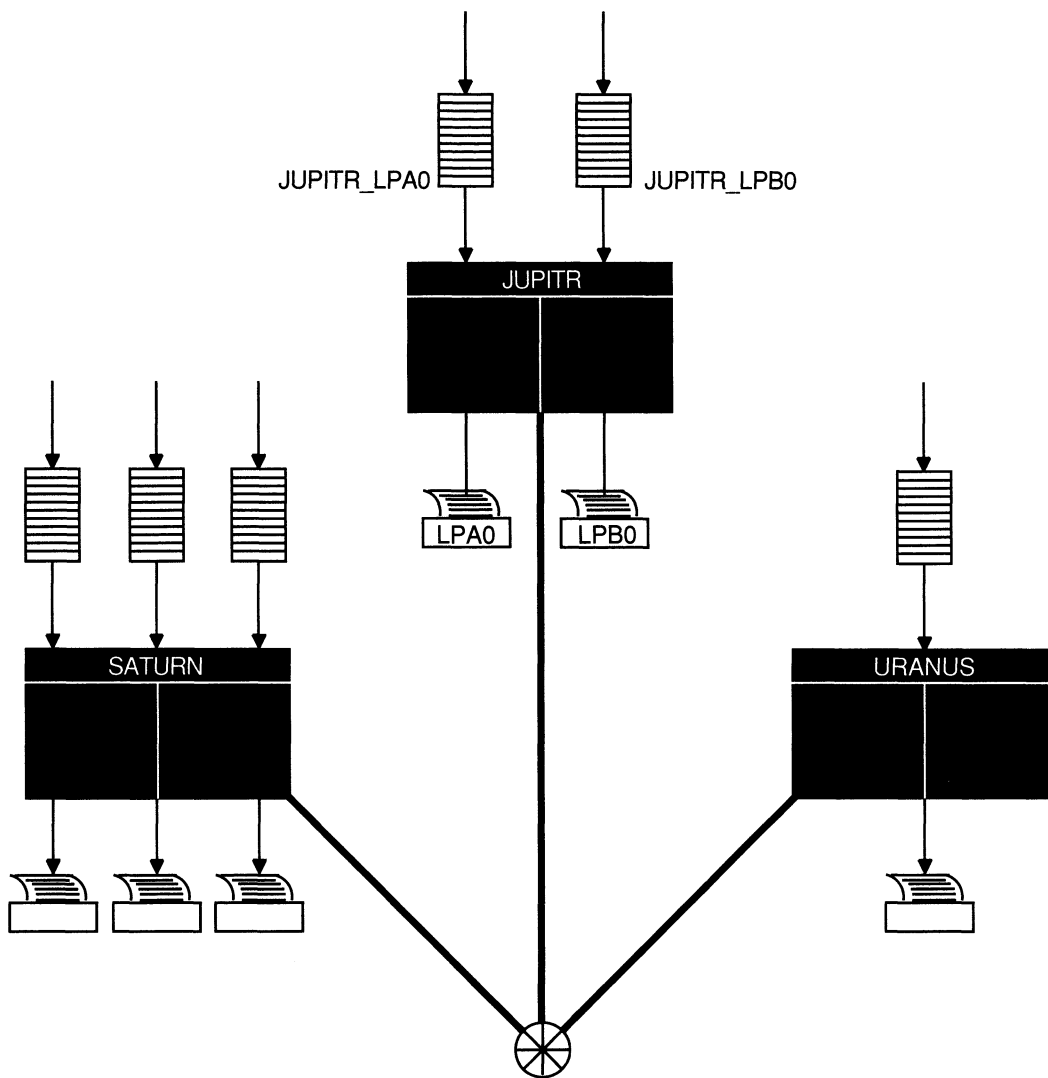
```
$ INITIALIZE/QUEUE/GENERIC=(JUPITR_LPA0,SATURN_LPA0,-
  URANUS_LPA0)/START SYS$PRINT
```

Jobs queued to SYS\$PRINT are placed in whichever assigned printer queue is available. Thus, in this example, a print job from JUPITR that is queued to SYS\$PRINT can be queued to JUPITR\_LPA0, SATURN\_LPA0, or URANUS\_LPA0.

A clusterwide generic printer queue needs to be initialized and started only once. The most efficient way to start your queues is to create a common command procedure that is executed by each VAXcluster computer (see Example 6-2).



Figure 6–2 Printer Queue Configuration



ZK-1632-GE

## 6.3 Cluster Batch Queues

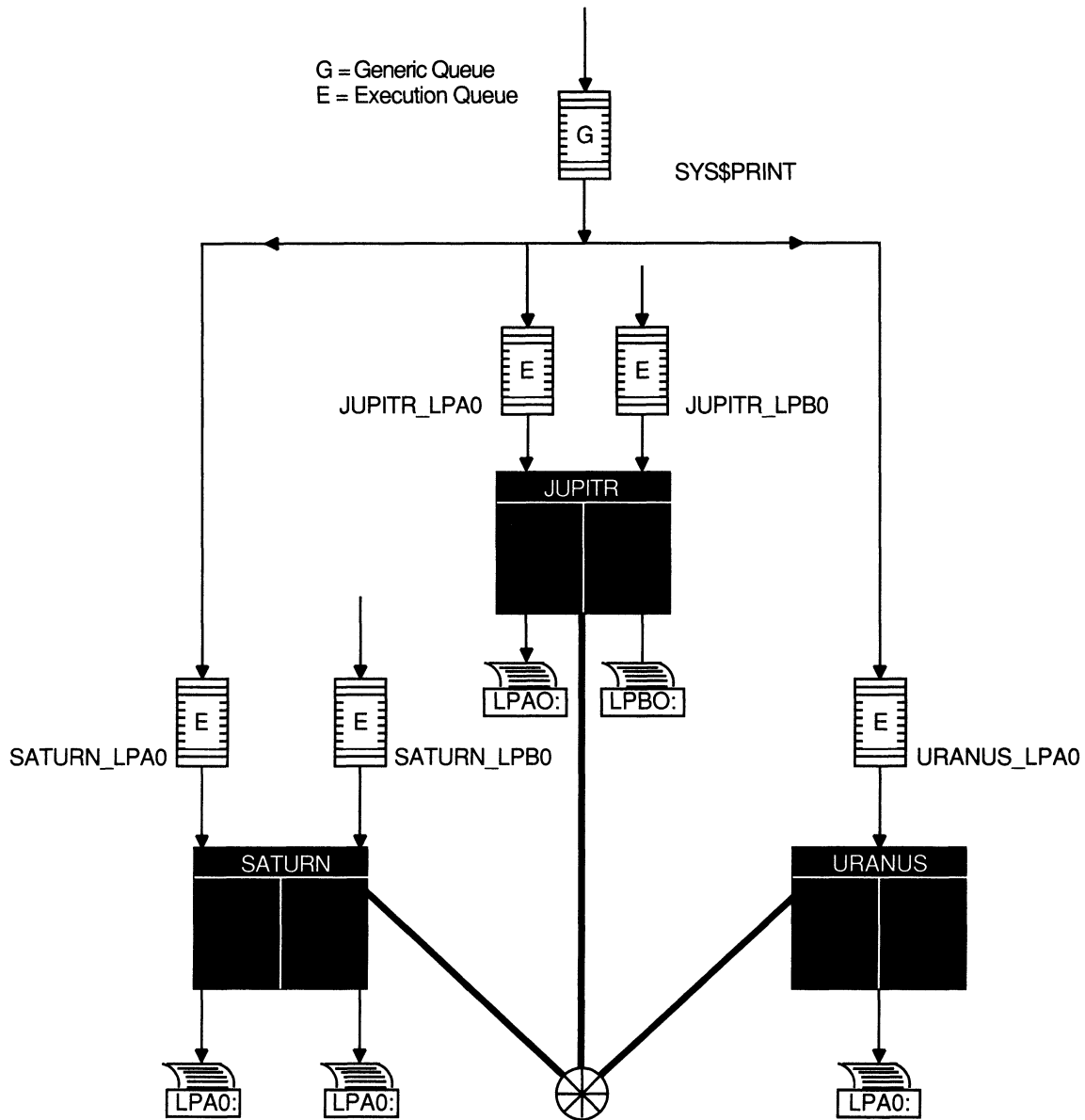
Before you establish batch queues, you should decide on the type of queue configuration that best suits your cluster. As system manager, you are responsible for setting up batch queues to maintain efficient batch job processing on the cluster. For example, you should do the following:

- Determine what type of processing will be performed on each computer
- Set up local batch queues that conform to these processing needs
- Decide whether to set up any clusterwide generic queues that will distribute batch job processing across the cluster
- Decide whether to use autostart queues for high availability.

# Setting Up and Managing Cluster Queues

## 6.3 Cluster Batch Queues

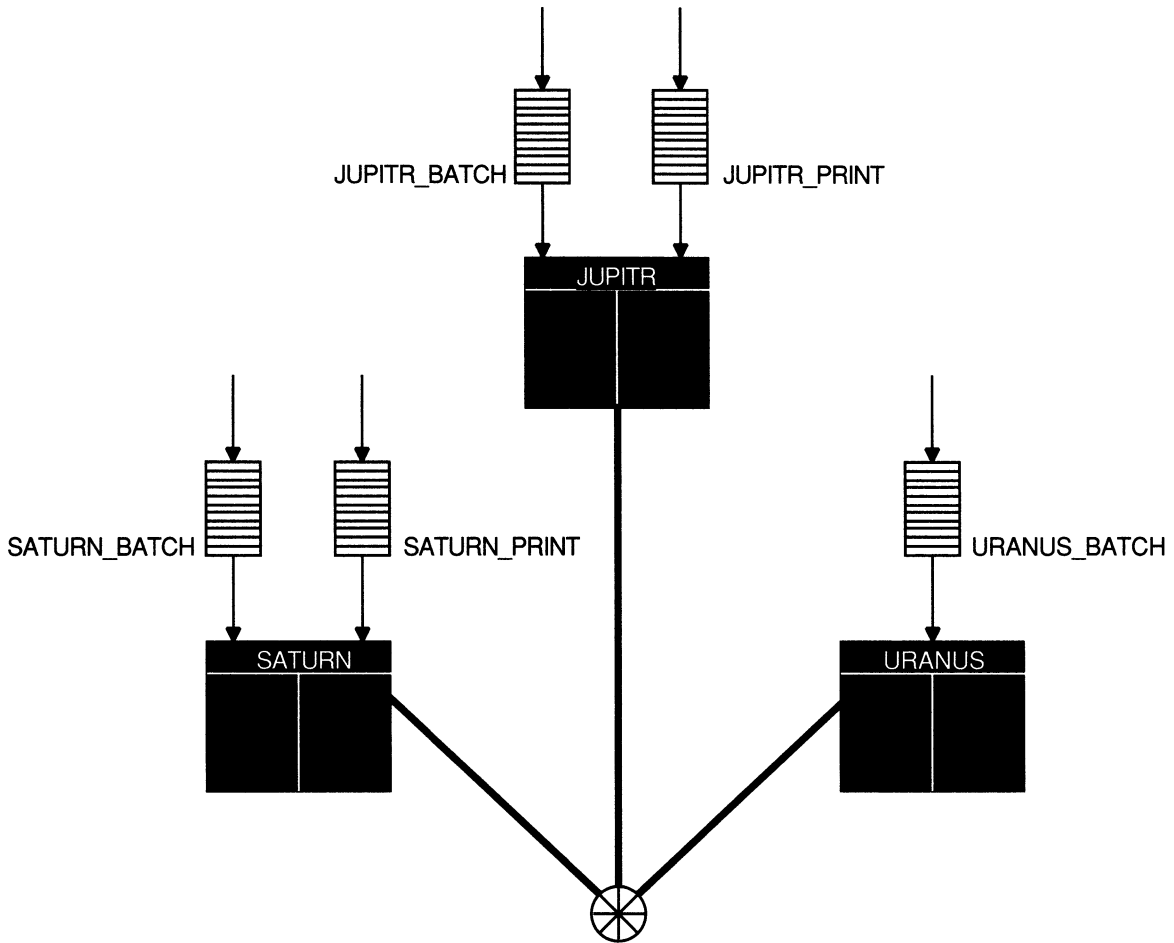
Figure 6-3 Clusterwide Generic Printer Queue Configuration



ZK-1634-GE

Once you determine the strategy that best suits your needs, you can create a command procedure to set up your queues (see Example 6-2). Figure 6-4 shows a batch queue configuration for a cluster consisting of computers JUPITR, SATURN, and URANUS.

Figure 6-4 Sample Batch Queue Configuration



ZK-1635-GE

### 6.3.1 Setting Up Execution Batch Queues

Generally, you set up execution batch queues on each VAXcluster computer using the same procedures you use for a standalone computer. For more detailed information on how this is done, see the *Guide to Maintaining a VMS System*.

You create a batch queue with a unique name by specifying the DCL command INITIALIZE/QUEUE/BATCH in the following format:

```
$ INITIALIZE/QUEUE/BATCH/ON=node::[/START] queue-name
```

The /ON qualifier specifies the computer on which the batch queue runs. If you specify the /START qualifier, the queue is started.

You can create an autostart batch queue by specifying the DCL command INITIALIZE/QUEUE/BATCH in the following format:

```
$ INITIALIZE/QUEUE/BATCH/AUTOSTART_ON=node:... queue-name
```

Note that you cannot specify both /ON and /AUTOSTART\_ON.

## Setting Up and Managing Cluster Queues

### 6.3 Cluster Batch Queues

The following commands make the local batch queue assignments for JUPITR, SATURN, and URANUS shown in Figure 6-4:

```
$ INITIALIZE/QUEUE/BATCH/ON=JUPITR::/START JUPITR_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=SATURN::/START SATURN_BATCH
$ INITIALIZE/QUEUE/BATCH/ON=URANUS::/START URANUS_BATCH
```

Because batch jobs on each VAXcluster computer are queued to SYS\$BATCH by default, you should consider defining a logical name to establish this queue as a clusterwide generic batch queue that distributes batch job processing throughout the cluster (see Example 6-2). Note, however, that you should do this only if you have a common-environment cluster. Guidelines for establishing clusterwide generic batch queues are presented in Section 6.3.2.

#### 6.3.2 Setting Up Clusterwide Generic Batch Queues

In a VAXcluster system, you can distribute batch processing among computers to balance the use of processing resources. You can achieve this workload distribution by assigning local batch queues to one or more clusterwide generic batch queues. These generic batch queues control batch processing across the cluster by placing batch jobs in assigned batch queues that are available. You can create a clusterwide generic batch queue as shown in Example 6-2.

In Figure 6-5, batch queues from each VAXcluster computer are assigned to a clusterwide generic batch queue named SYS\$BATCH. Users can submit a job to a specific queue (for example, JUPITR\_BATCH or SATURN\_BATCH) or, if they have no special preference, they can submit it by default to the clusterwide generic queue SYS\$BATCH. The generic queue in turn places the job in an available assigned queue in the cluster.

If more than one assigned queue is available, the VMS operating system selects the queue that minimizes the ratio (executing jobs/job limit) for all assigned queues.

A clusterwide generic batch queue needs to be initialized and started only once. The most efficient way to perform these operations is to create a common command procedure that is executed by each VAXcluster computer (see Example 6-2).

## 6.4 Using a Common Command Procedure to Set Up Cluster Queues

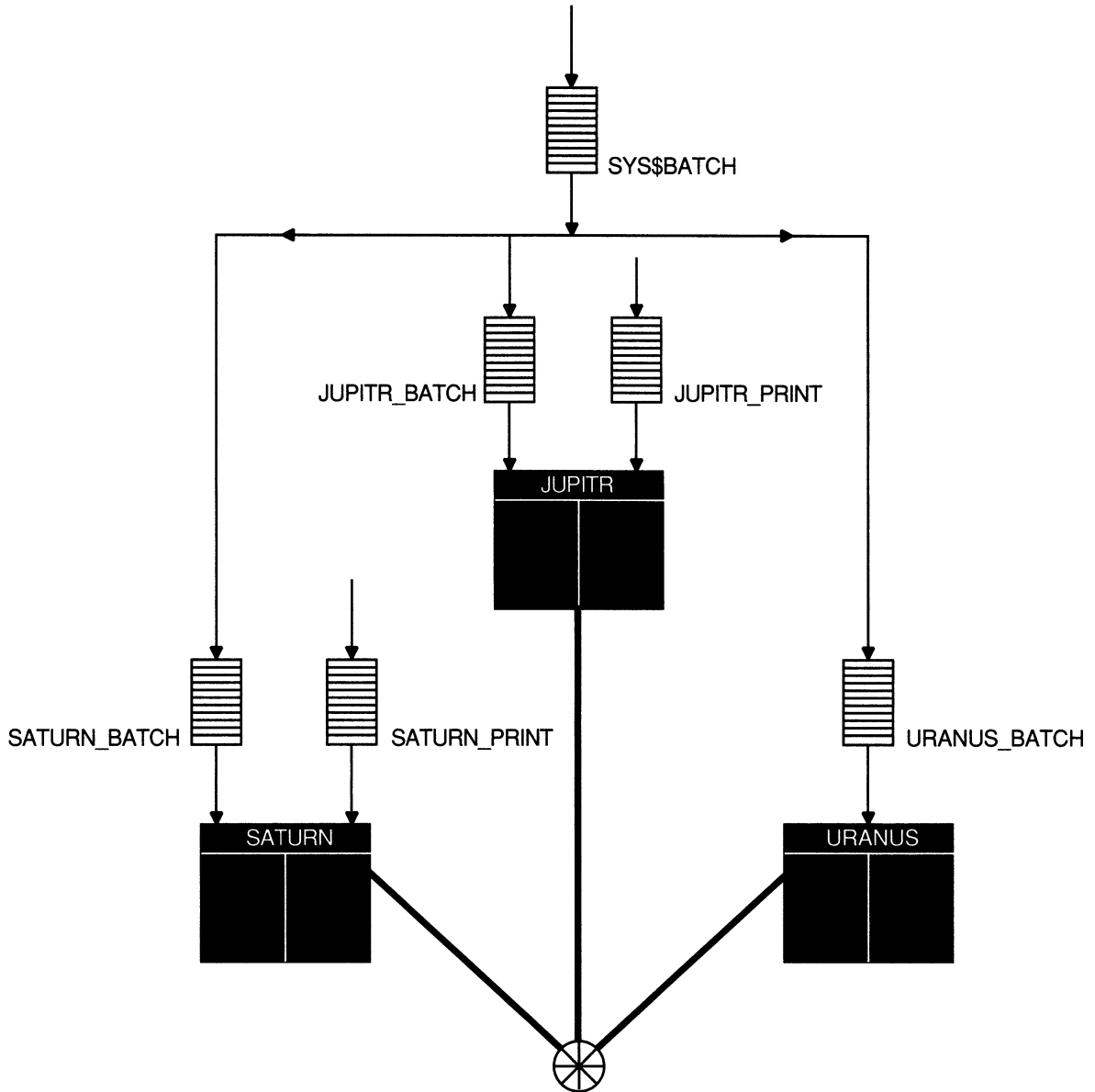
Once you have created queues, you must start them to begin processing batch and print jobs. In addition, you must make sure the queues are started each time the system reboots, by enabling autostart for autostart queues, or by entering START/QUEUE commands for nonautostart queues. To do so, create a command procedure containing the necessary commands.

You can create a common command procedure named, for example, QSTARTUP.COM, and store it on a shared disk. With this method, each node can share the same copy of the common QSTARTUP.COM procedure. Each node invokes the common QSTARTUP.COM procedure from the common version of SYSTARTUP. You can also include the commands to start queues in the common SYSTARTUP file instead of in a separate QSTARTUP.COM file.

# Setting Up and Managing Cluster Queues

## 6.4 Using a Common Command Procedure to Set Up Cluster Queues

Figure 6-5 Clusterwide Generic Batch Queue Configuration



ZK-1636-GE

Example 6-1 shows commands used to create VAXcluster queues; Example 6-2 illustrates the use of a common QSTARTUP command procedure on a shared disk.

## Setting Up and Managing Cluster Queues

### 6.4 Using a Common Command Procedure to Set Up Cluster Queues

#### Example 6-1 Sample Commands for Creating VAXcluster Queues

```
$
①
$ DEFINE/FORM LN FORM 10 /WIDTH=80 /STOCK=DEFAULT /TRUNCATE
$ DEFINE/CHARACTERISTIC 2ND_FLOOR 2
.
.
②
$ INITIALIZE/QUEUE/AUTOSTART_ON=(JUPITR::LPA0:)/START JUPITR_PRINT
$ INITIALIZE/QUEUE/AUTOSTART_ON=(SATURN::LPA0:)/START SATURN_PRINT
$ INITIALIZE/QUEUE/AUTOSTART_ON=(URANUS::LPA0:)/START URANUS_PRINT
.
.
③
$ INITIALIZE/QUEUE/BATCH/START/ON=JUPITR:: JUPITR_BATCH
$ INITIALIZE/QUEUE/BATCH/START/ON=SATURN:: SATURN_BATCH
$ INITIALIZE/QUEUE/BATCH/START/ON=URANUS:: URANUS_BATCH
.
.
④
$ INITIALIZE/QUEUE/START -
  $ /AUTOSTART_ON=(JUPITR::LTA1:,SATURN::LTA1,URANUS::LTA1) -
  $ /PROCESSOR=LATSYM /FORM_MOUNTED=LN FORM -
  $ /RETAIN=ERROR /DEFAULT=(NOBURST,FLAG=ONE,NOTRAILER) -
  $ /RECORD_BLOCKING LN03$PRINT
$
$ INITIALIZE/QUEUE/START -
  $ /AUTOSTART_ON=(JUPITR::LTA2:,SATURN::LTA2,URANUS::LTA2) -
  $ /PROCESSOR=LATSYM /RETAIN=ERROR -
  $ /DEFAULT=(NOBURST,FLAG=ONE,NOTRAILER) /RECORD_BLOCKING -
  $ /CHARACTERISTIC=2ND_FLOOR LA210$PRINT
$
⑤
$ ENABLE AUTOSTART/QUEUES/ON=SATURN
$ ENABLE AUTOSTART/QUEUES/ON=JUPITR
$ ENABLE AUTOSTART/QUEUES/ON=URANUS
⑥
$ INITIALIZE/QUEUE/START SYS$PRINT -
  $ /GENERIC=(JUPITR_PRINT,SATURN_PRINT,URANUS_PRINT)
$
⑦
$ INITIALIZE/QUEUE/BATCH/START SYS$BATCH -
  $ /GENERIC=(JUPITR_BATCH,SATURN_BATCH,URANUS_BATCH)
$
```

The following are descriptions of each command or group of commands in Example 6-1:

- ① Define all printer forms and characteristics.
- ② Initialize local printer queues. In the example, these queues are autostart queues and will be automatically started when the node executes the `ENABLE AUTOSTART/QUEUES` command. Although the `/START` qualifier is specified to activate the autostart queues, they will not begin processing jobs until autostart is enabled.

To enable autostart each time the system reboots, add the `ENABLE AUTOSTART/QUEUES` command to your queue startup command procedure as described in Example 6-2.

## Setting Up and Managing Cluster Queues

### 6.4 Using a Common Command Procedure to Set Up Cluster Queues

- ③ Initialize and start local batch queues on all nodes, including satellite nodes. In this example, the local batch queues are not autostart queues. For more information on local batch queues, see Section 6.5.
- ④ Initialize queues for remote LAT printers. In the example, these queues are autostart queues and are set up to run on one of three nodes. The queues will be started on the first of those three nodes to execute the ENABLE AUTOSTART command.

You must establish the logical devices LTA1 and LTA2 in the LAT startup command procedure LTLOAD.COM on each node on which the autostart queue can run. For more information see the description of editing LTLOAD.COM in the *Guide to Setting Up a VMS System*.

Although the /START qualifier is specified to activate these autostart queues, they will not begin processing jobs until autostart is enabled.

- ⑤ Enable autostart to automatically start the autostart queues. In the example, autostart is enabled on node SATURN first, so SATURN will start those autostart queues which are set up to run one of several nodes.
- ⑥ Initialize and start the generic output queue SYS\$PRINT. This is a nonautostart queue (generic queues cannot be autostart queues). However, generic queues are not automatically stopped when a system is shut down, so you do not need to restart the queue each time a node reboots.
- ⑦ Initialize and start the generic batch queue SYS\$BATCH. Because this is a generic queue, it is not stopped when the node shuts down. Therefore, you do not need to restart the queue each time a node reboots.

#### Example 6-2 Common Procedure to Start VAXcluster Queues

```
$!  
$! QSTARTUP.COM -- Common procedure to set up cluster queues  
$!  
$!  
①  
$ NODE = F$GETSYI("NODENAME")  
$!  
$! Determine the node-specific subroutine  
$!  
$ IF (NODE .NES. "JUPITR") .AND. (NODE .NES. "SATURN") .AND. (NODE .NES. "URANUS")  
$ THEN  
$   GOSUB SATELLITE_STARTUP  
$ ELSE  
②  
$!  
$! Configure remote LAT devices.  
$!  
$   SET TERMINAL LTA1: /PERM /DEVICE=LN03 /WIDTH=255 /PAGE=60 -  
$     /LOWERCASE /NOBROAD  
$   SET TERMINAL LTA2: /PERM /DEVICE=LA210 /WIDTH=255 /PAGE=66 -  
$     /NOBROAD  
$   SET DEVICE LTA1: /SPOOLED=(LN03$PRINT, SYS$SYSDEVICE:)  
$   SET DEVICE LTA2: /SPOOLED=(LA210$PRINT, SYS$SYSDEVICE:)  
③  
$   START/QUEUE/BATCH 'NODE' _BATCH  
$   GOSUB 'NODE' _STARTUP  
$   ENDF  
$ GOTO ENDING
```

(continued on next page)

## Setting Up and Managing Cluster Queues

### 6.4 Using a Common Command Procedure to Set Up Cluster Queues

#### Example 6–2 (Cont.) Common Procedure to Start VAXcluster Queues

```
$!  
$! Node-specific subroutines start here  
$!  
④  
$ SATELLITE_STARTUP:  
$!  
$! Start a batch queue for satellites.  
$!  
$ START/QUEUE/BATCH 'NODE'_BATCH  
$ RETURN  
$!  
⑤  
$JUPITR_STARTUP:  
$!  
$! Node-specific startup for JUPITR::  
$! Setup local devices and start nonautostart queues here  
$!  
$ SET PRINTER/PAGE=66 LPA0:  
$ RETURN  
$!  
$SATURN_STARTUP:  
$!  
$! Node-specific startup for SATURN::  
$! Setup local devices and start nonautostart queues here  
$!  
.  
.  
.  
$ RETURN  
$!  
$URANUS_STARTUP:  
$!  
$! Node-specific startup for URANUS::  
$! Setup local devices and start nonautostart queues here  
$!  
.  
.  
.  
$ RETURN  
$!  
$ENDING:  
⑥  
$! Enable autostart to start all autostart queues  
$!  
$ ENABLE AUTOSTART/QUEUES  
$ EXIT
```

Following are brief descriptions of each phase of the common QSTARTUP.COM command procedure:

- ① Determine the name of the node executing the procedure.
- ② On all large nodes, set up remote devices connected via LAT. The queues for these devices are autostart queues and will automatically be started when the ENABLE AUTOSTART/QUEUES command is executed at the end of this procedure.

In the example, these autostart queues were set up to run on one of three nodes. The queues will start when the first of those nodes executes the ENABLE AUTOSTART/QUEUES command. The queue will remain running as long as one of those nodes is running and has autostart enabled.



### 6.4 Using a Common Command Procedure to Set Up Cluster Queues

- ③ On large nodes, start the local batch queue. In the example, the local batch queues are nonautostart queues, and must be explicitly started with `START /QUEUE` commands.
- ④ On satellite nodes, start the local batch queue.
- ⑤ Each node executes its own subroutine. On node JUPITR, set up the line printer device LPA0:. The queue for this device is an autostart queue, and will be started automatically when the `ENABLE AUTOSTART/QUEUES` command is executed.
- ⑥ Enable autostart to start all autostart queues.

### 6.5 Starting Local Batch Queues

Normally, you use local batch execution queues during startup, to run batch jobs to start layered products. For this reason, these queues must be available before printers are configured and the `ENABLE AUTOSTART` command is executed in `QSTARTUP.COM`.

Start the local batch execution queue in each node's startup command procedure `SYSTARTUP_V5.COM`. If you use a common startup command procedure, add commands similar to the following to your procedure:

```
$ SUBMIT/PRIORITY=255/NOIDENT/NOLOG/QUEUE='NODE'_BATCH LAYERED_PRODUCT.COM
$ START/QUEUE 'NODE'_BATCH
$ DEFINE/SYSTEM/EXECUTIVE SYSSBATCH 'NODE'_BATCH
```

Submitting the startup command procedure `LAYERED_PRODUCT.COM` as a high-priority batch job before the queue starts ensures that the job is executed immediately, regardless of the job limit on the queue. If the queue were started before the command procedure was submitted, the queue might reach its job limit by scheduling user batch jobs, and the startup job would have to wait.



---

## Building and Maintaining the Cluster

Before you attempt to build your cluster, be sure you have read the previous chapters and made the following preparations:

- Determined the VAXcluster configuration type (CI-only, DSSI, local area, or mixed-interconnect)
- Determined whether you want a common-environment or multiple-environment VAXcluster system
- Determined how you will set up and distribute the startup and system files that define the operating environment
- Planned your disk, tape, and queue configurations
- Installed or upgraded the VMS operating system on the first VAXcluster computer and installed any required licenses
- Configured and started the DECnet-VAX network

Once you have made these preparations, you can use the information in this chapter to build and maintain your cluster. Topics include the following:

- CLUSTER\_CONFIG.COM functions
- Determining locations and sizes for satellite page and swap files
- Specifying allocation classes in mixed-interconnect clusters
- Configuring the cluster
- Reconfiguring the cluster after a major change
- Maintaining the cluster

### 7.1 CLUSTER\_CONFIG.COM Functions

When you invoke CLUSTER\_CONFIG.COM, the procedure displays a menu of configuration options. By selecting the appropriate option, you can configure the cluster easily and reliably, without invoking VMS utilities directly. You use CLUSTER\_CONFIG.COM to perform these functions:

- Add a computer to the cluster
- Remove a computer from the cluster
- Change a computer's characteristics
- Create a duplicate system disk

# Building and Maintaining the Cluster

## 7.1 CLUSTER\_CONFIG.COM Functions

Table 7–1 summarizes the operations that CLUSTER\_CONFIG.COM performs for each configuration function.

**Table 7–1 Summary of CLUSTER\_CONFIG.COM Functions**

Function	Operations Performed
ADD	<p>Establish the new computer's root directory on a cluster common system disk and generate the computer's system parameter files (VAXVMSSYS.PAR and MODPARAMS.DAT) in its SYS\$SPECIFIC:[SYSEXE] directory.</p> <p>Update the permanent and volatile remote node network databases for the computer on which CLUSTER_CONFIG.COM is executed to add the new computer. If the new computer is a satellite, update SYS\$MANAGER:NETNODE_UPDATE.COM on the local computer (see Section 7.5.1.2).</p> <p>Generate the new computer's page and swap files (PAGEFILE.SYS and SWAPFILE.SYS).</p> <p>Optionally set up a cluster quorum disk.</p> <p>Set allocation class (ALLOCLASS) value for the new computer, if the computer is being added as a disk server.</p> <p>Generate an initial (temporary) startup procedure for the new computer. This initial procedure runs NETCONFIG.COM to configure the network, runs AUTOGEN to set appropriate SYSGEN parameter values for the computer, and reboots the computer with normal startup procedures.</p>
REMOVE	<p>Delete another computer's root directory and its contents from the local computer's system disk. If the computer being removed is a satellite, update SYS\$MANAGER:NETNODE_UPDATE.COM on the local computer.</p> <p>Update the permanent and volatile remote node network databases on the local computer.</p>
CHANGE	<p>Enable or disable the local computer as a disk server; enable or disable the local computer as a boot server; enable or disable the Ethernet or FDDI LAN for cluster communications on the local computer; enable or disable a quorum disk on the local computer; change the local computer's ALLOCLASS value; change a satellite's LAN hardware address. Procedure displays CHANGE menu and prompts for appropriate information.</p>
CREATE	<p>Duplicate the local computer's system disk and remove all system roots from the new disk.</p>

If you intend to set up a local area or mixed-interconnect cluster, you must, before executing CLUSTER\_CONFIG.COM, do the following:

- Determine locations and sizes for satellite page and swap files
- Select cluster boot servers and disk servers
- Determine allocation classes for computers and disks (also applicable for CI-based configurations)

Guidelines are provided in Section 7.2, Section 7.3, and Section 7.4, respectively.

Note that some configuration functions, such as adding or removing a **voting member** (a computer with a nonzero value for the SYSGEN parameter VOTES), or enabling or disabling a quorum disk, require one or more additional operations. Refer to Section 7.6 for instructions.

## 7.2 Determining Locations and Sizes for Satellite Page and Swap Files

When you add a computer to the cluster, CLUSTER\_CONFIG.COM prompts for the sizes and location of the computer's page and swap files. (The default sizes supplied by the procedure are minimums.) Depending on the configuration of your VAXcluster system disk and your network, you may realize a performance improvement in local area and mixed-interconnect configurations by locating page and swap files for satellites on a satellite's local disk, if such a disk is available.

To set up page and swap files on a satellite's local disk, CLUSTER\_CONFIG.COM creates (in the satellite's [SYSx.SYSEXx] directory on the boot server's system disk) the command procedure SATELLITE\_PAGE.COM. This procedure executes when AUTOGEN reboots the satellite at the end of CLUSTER\_CONFIG.COM. SATELLITE\_PAGE.COM performs the following functions:

- Mounts the satellite's local disk with a volume label that is unique in the cluster in the format *node-name\_SCSSYSTEMID*
- Installs the page and swap files on the satellite's local disk

If you want to alter the volume label, follow these steps after the satellite has been added to the cluster:

1. Log in as system manager and enter a DCL command in the following format:  
SET VOLUME/LABEL=*volume-label device-spec*[:]

Note that the SET VOLUME command requires write access (W) to the index file on the volume. If you are not the volume's owner, you must have either a system UIC or the SYSPRV privilege.

2. Update SATELLITE\_PAGE.COM to reflect the new label.

To relocate the satellite's page and swap files (for example, from the satellite's local disk to the boot server's system disk, or the reverse), or to change file sizes, the easiest way is to remove the satellite from the cluster and then add it again, using CLUSTER\_CONFIG.COM.

## 7.3 Selecting MOP and Disk Servers

While every local area and mixed-interconnect cluster must include at least one MOP and disk server, multiple MOP and disk servers offer the following advantages:

- Higher availability—Satellites can access served disks and boot, even if one of the servers is temporarily unavailable.
- Better workload balancing—The task of serving disks to satellites can place a significant load on a server. With multiple servers, this workload is distributed across more computers and LAN adapters.

As a general rule, use as MOP and disk servers the most powerful computers in the cluster. Low-powered computers can become overloaded when serving many busy satellites, or when many satellites boot simultaneously. Note, however, that two or more moderately powered servers can provide better performance than a single high-powered server. Multiple servers give better availability, and they distribute the workload across more LAN adapters. If you have several computers of roughly comparable power, it is reasonable to use them all as boot servers. This arrangement gives optimal load balancing. In addition, if one computer fails or is shut down, others remain available to serve satellites.

## Building and Maintaining the Cluster

### 7.3 Selecting MOP and Disk Servers

After compute power, the most important factor in selecting a server is the speed of its LAN adapter. Servers should be equipped with the highest-bandwidth LAN adapters in the cluster.

## 7.4 Determining Allocation Class Values in Mixed-Interconnect Clusters

The allocation class value coupled with unit numbers and, for local controllers, with the controller letter, must form a clusterwide unique name for the device.

Before setting up any mixed-interconnect cluster, you must determine allocation class values for disk and tape servers and HSC subsystems. It is easiest to use the same value for all servers and HSC subsystems; you can arbitrarily choose a number between 1 and 255. Note, however, that to change the allocation class value on any CI-connected computer or HSC subsystem, you must shut down and reboot the entire cluster. (See Section 7.6.)

As explained in Section 5.2, every device allocation class name (in the form \$1\$ddcu) must be the same for all servers and HSC subsystems that share the devices. For RA-series disks, make sure that all the removable unit plugs on all disks of that allocation class are unique. As long as you have no more than 256 such disks, this is easy to accomplish.

Assume, for instance, that 10 disks are dual pathed between the HSC subsystems VOYGR1 and VOYGR2, and assume that 10 others are dual pathed between the HSC subsystems VIKNG1 and VIKNG2. Provided that all 20 disks have unique unit numbers, you can assign the same allocation class value to all four HSC subsystems.

However, if you run out of unique disk unit numbers, you must define unique disk names by using two or more allocation class values for the HSC subsystems. You must also configure one or more computers to serve HSC disks and assign allocation class values accordingly. To perform those operations, you can execute the CLUSTER\_CONFIG.COM CHANGE function, which is described in Section 7.5.3.

Additionally, you must ensure that all locally connected disks have unique device names. For example, if SATURN and URANUS each have a single-pathed RA81 disk connected to a local BDA controller with unit plug 0, and if both computers have an allocation class value of 1, both RA81 disks receive the same device name (\$1\$DUA0). Because both disks have the same device name, they appear to VMS software to be the same disk. This condition can endanger data integrity. You can avoid potential problems by selecting a different unit number for one of the disks or by using a different allocation class on one of these nodes.

Note that because fewer unit numbers are available for MASSBUS or UNIBUS disks, fewer unique device names can be defined. To ensure that device names remain unique in your cluster, you may have to relocate such disks or disqualify a computer as a disk server.

## 7.5 Configuring the Cluster

To perform configuration functions, you execute CLUSTER\_CONFIG.COM. Before invoking the procedure, be sure to verify the following:

- You are logged in to the system manager's account on an appropriate computer. If you are building a new local area or mixed-interconnect cluster, you must be logged in on a computer that you want to set up as a boot server. If you are adding a satellite, you must be logged in on a boot server. Note that the process privileges SYSPRV, OPER, CMKRNL, BYPASS, and NETMBX are required, because the procedure performs sensitive system operations.
- The DECnet-VAX network is up and running and all computers are connected to the LAN.
- You have at hand the data listed in Table 7-2. Note that some items are configuration specific.
- If your configuration has two or more system disks, you have coordinated cluster common files, as described in Section 4.5.4.

Sections 7.5.1 through 7.5.6 provide examples of typical interactive CLUSTER\_CONFIG.COM sessions. Section 7.6 describes tasks you must perform after executing CLUSTER\_CONFIG.COM to make major configuration changes.

**Caution**

You may not initiate concurrent CLUSTER\_CONFIG.COM sessions.

**Table 7-2 Data Requested by CLUSTER\_CONFIG.COM**

Item	How to Specify or Obtain
Device name of cluster system disk on which root directories will be created.	System manager specifies. Default is logical volume name of SYS\$SYSDEVICE: (for example, DISK\$VAXVMSRL5:).
Computer's root directory name on cluster system disk.	System manager specifies. Name must be of the form SYSx. For CI-connected computers, x is a hexadecimal digit in the range 1 through 9 or A through D (for example, SYS1 or SYSA). For satellites, x must be in the range from 10 through FFFF. Procedure supplies valid default.
Computer's DECnet node name.	Network manager supplies. Name must be from 1 to 6 alphanumeric characters and <i>may not</i> include dollar signs (\$) or underscores (_).
Computer's DECnet node address.	Network manager supplies.
Cluster group number and password if CHANGE is run to enable cluster communications over the LAN.	System manager specifies.

(continued on next page)

# Building and Maintaining the Cluster

## 7.5 Configuring the Cluster

**Table 7-2 (Cont.) Data Requested by CLUSTER\_CONFIG.COM**

Item	How to Specify or Obtain
<p>If computer is a satellite on a LAN, satellite's LAN hardware address. Address has the form xx-xx-xx-xx-xx-xx. Note that you must include the dashes when you specify a hardware address.</p>	<p>When DECnet-VAX network is running on boot server, proceed as follows:</p> <ul style="list-style-type: none"> <li>For MicroVAX II and VAXstation II satellites, enter the following commands at satellite's console: <pre data-bbox="949 472 1204 535">&gt;&gt;&gt; B/100 XQAO Bootfile: READ_ADDR</pre> </li> <li>For MicroVAX 2000 and VAXstation 2000 satellites, enter the following commands at successive console-mode prompts: <pre data-bbox="949 651 1204 756">&gt;&gt;&gt; T 53 2 ?&gt;&gt;&gt; 3 &gt;&gt;&gt; B/100 ESAO Bootfile: READ_ADDR</pre> <p>If the second prompt appears as 3 ?&gt;&gt;&gt;, press the Return key.</p> </li> <li>For MicroVAX 3xxx and 4xxx series satellites, enter the following command at satellite's console: <pre data-bbox="949 934 1173 966">&gt;&gt;&gt; SHOW ETHERNET</pre> </li> </ul>
<p>Workstation windowing system.</p>	<p>System manager specifies. Workstation software must be installed before workstation satellites are added. If it is not, the procedure indicates that fact.</p>
<p>Location and sizes of page and swap files.</p>	<p>System manager specifies.</p>
<p>Value for local computer's allocation class (ALLOCLASS) parameter.</p>	<p>System manager specifies.</p>
<p>Physical device name of quorum disk.</p>	<p>System manager specifies.</p>

### 7.5.1 Adding a Computer to the Cluster

Once you have made the necessary preparations, you can execute CLUSTER\_CONFIG.COM to add a new computer to the cluster.

- If you are setting up a CI-only cluster or a DSSI cluster, invoke CLUSTER\_CONFIG.COM on an active VAXcluster computer and select the ADD function.
- If you are setting up a new local area or mixed-interconnect cluster, follow these steps:
  1. Invoke CLUSTER\_CONFIG.COM and execute the CHANGE function described in Section 7.5.3 to enable the local computer as a boot server.
  2. After the CHANGE function completes, execute the ADD function to add either CI-connected computers or satellites to the cluster. To add satellites, you must be logged in on a cluster boot server.

While adding computers, you may want to disable broadcast messages to your terminal—the ADD function generates many such messages. To disable the messages, you can enter the DCL command REPLY/DISABLE=(NETWORK, CLUSTER).



Whenever you add a voting member to the cluster, you must, after the ADD function completes, reconfigure the cluster, following instructions in Section 7.6. In addition, if you add a CI-connected computer that boots from a cluster common system disk, you must create a new default bootstrap command procedure for the computer before booting it into the cluster. For instructions, refer to your computer-specific installation and operations guide.

Example 7-1 and Example 7-2 illustrate the use of CLUSTER\_CONFIG.COM on JUPITR to add, respectively, CI-connected computer SATURN and satellite computer EUROPA to the cluster.

---

**Caution**

---

If either the local or the new computer fails before the ADD function completes, you must, after normal conditions are restored, perform the REMOVE function to erase any invalid data and then restart the ADD function.

---

**Example 7-1 Sample Interactive CLUSTER\_CONFIG.COM Session to Add a CI-Connected Computer as a Boot Server**

```
$ @CLUSTER_CONFIG.COM
      Cluster Configuration Procedure
Use CLUSTER_CONFIG.COM to set up or change a VAXcluster configuration.
To ensure that you have the required privileges, invoke this procedure
from the system manager's account.
Enter ? for help at any prompt.
      1. ADD a node to the cluster.
      2. REMOVE a node from the cluster.
      3. CHANGE a cluster node's characteristics.
      4. CREATE a second system disk for JUPITR.
Enter choice [1]: 
The ADD function adds a new node to the cluster.
If the node being added is a voting member, EXPECTED_VOTES in all
other cluster members' MODPARAMS.DAT must be adjusted, and the
cluster must be rebooted.
If the new node is a satellite, the network databases on JUPITR are
updated. The network databases on all other cluster members must be
updated.
For instructions, see the VMS VAXcluster Manual.
What is the node's DECnet node name? SATURN
What is the node's DECnet address? 2.3
Will SATURN be a satellite [Y]? N
Will SATURN be a boot server [Y]? 
```

(continued on next page)

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

#### Example 7-1 (Cont.) Sample Interactive CLUSTER\_CONFIG.COM Session to Add a CI-Connected Computer as a Boot Server

This procedure will now ask you for the device name of SATURN's system root. The default device name (DISK\$VAXVMSRL5:) is the logical volume name of SYS\$SYSDEVICE:.

What is the device name for SATURN's system root [DISK\$VAXVMSRL5:]?

What is the name of the new system root [SYSA]?

Creating directory tree SYSA...

%CREATE-I-CREATED, \$1\$DJA11:<SYSA> created

%CREATE-I-CREATED, \$1\$DJA11:<SYSA.SYSEXE> created

.

.

System root SYSA created.

Enter a value for SATURN's ALLOCLASS parameter: 1

Does this cluster contain a quorum disk [N]? Y

What is the device name of the quorum disk? \$1\$DJA12

Updating network database...

Size of page file for SATURN [10000 blocks]? 50000

Size of swap file for SATURN [8000 blocks]? 20000

Will a local (non-HSC) disk on SATURN be used for paging and swapping? N

If you specify a device other than DISK\$VAXVMSRL5: for SATURN's page and swap files, this procedure will create PAGEFILE\_SATURN.SYS and SWAPFILE\_SATURN.SYS in the <SYSEXE> directory on the device you specify.

What is the device name for the page and swap files [DISK\$VAXVMSRL5:]?

%SYSGEN-I-CREATED, \$1\$DJA11:<SYSA.SYSEXE>PAGEFILE.SYS;1 created

%SYSGEN-I-CREATED, \$1\$DJA11:<SYSA.SYSEXE>SWAPFILE.SYS;1 created

The configuration procedure has completed successfully.

SATURN has been configured to join the cluster.

Before booting SATURN, you must create a new default bootstrap command procedure for SATURN. See your processor-specific installation and operations guide for instructions.

The first time SATURN boots, NETCONFIG.COM and AUTOGEN.COM will run automatically.

The following parameters have been set for SATURN:

VOTES = 1

QDSKVOTES = 1

After SATURN has booted into the cluster, you must increment the value for EXPECTED VOTES in every cluster member's MODPARAMS.DAT. You must then reconfigure the cluster, using the procedure described in the VMS VAXcluster Manual.

## Building and Maintaining the Cluster 7.5 Configuring the Cluster

### Example 7-2 Sample Interactive CLUSTER\_CONFIG.COM Session to Add a Satellite with Local Page and Swap Files

```
$ @CLUSTER_CONFIG.COM
```

```
Cluster Configuration Procedure
```

```
Use CLUSTER_CONFIG.COM to set up or change a VAXcluster configuration.
To ensure that you have the required privileges, invoke this procedure
from the system manager's account.
```

```
Enter ? for help at any prompt.
```

1. ADD a node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster node's characteristics.
4. CREATE a second system disk for JUPITR.

```
Enter choice [1]: 
```

```
The ADD function adds a new node to the cluster.
```

```
If the node being added is a voting member, EXPECTED_VOTES in all
other cluster members' MODPARAMS.DAT must be adjusted, and the
cluster must be rebooted.
```

```
If the new node is a satellite, the network databases on JUPITR are
updated. The network databases on all other cluster members must be
updated.
```

```
For instructions, see the VMS VAXcluster Manual.
```

```
What is the node's DECnet node name? EUROPA
```

```
What is the node's DECnet address? 2.21
```

```
Will EUROPA be a satellite [Y]? 
```

```
Verifying circuits in network database...
```

```
This procedure will now ask you for the device name of EUROPA's system root.
The default device name (DISK$VAXVMSRL5:) is the logical volume name of
SYS$SYSDEVICE:.
```

```
What is the device name for EUROPA'S system root [DISK$VAXVMSRL5:]? 
```

```
What is the name of the new system root [SYS10]? 
```

```
Allow conversational bootstraps on EUROPA [NO]? 
```

```
The following workstation windowing options are available:
```

1. No workstation software
2. VWS Workstation Software
3. DECwindows Workstation Software

```
Enter choice [1]: 3
```

```
Creating directory tree SYS10...
```

```
%CREATE-I-CREATED, $1$DJA11:<SYS10> created
```

```
%CREATE-I-CREATED, $1$DJA11:<SYS10.SYSEXEXE> created
```

```
.
```

```
.
```

```
.
```

```
System root SYS10 created.
```

```
Will EUROPA be a disk server [N]? 
```

```
What is EUROPA's Ethernet hardware address? 08-00-2B-03-51-75
```

```
Updating network database...
```

```
Size of pagefile for EUROPA [10000 blocks]? 20000
```

```
Size of swap file for EUROPA [8000 blocks]? 12000
```

```
Will a local disk on EUROPA be used for paging and swapping? YES
```

```
Creating temporary page file in order to boot EUROPA for the first time...
```

```
%SYSGEN-I-CREATED, $1$DJA11:<SYS10.SYSEXEXE>PAGEFILE.SYS;1 created
```

```
This procedure will now wait until EUROPA joins the cluster.
```

(continued on next page)

# Building and Maintaining the Cluster

## 7.5 Configuring the Cluster

### Example 7-2 (Cont.) Sample Interactive CLUSTER\_CONFIG.COM Session to Add a Satellite with Local Page and Swap Files

Once EUROPA joins the cluster, this procedure will ask you to specify a local disk on EUROPA for paging and swapping.

Please boot EUROPA now.

Waiting for EUROPA to boot...

.

.

(User enters boot command at satellite's console-mode prompt (>>>)). For MicroVAX II, VAXstation II, and MicroVAX 3xxx series satellites, user enters B XQ. For MicroVAX 2000 and VAXstation 2000 satellites, user enters B ES.)

.

.

The local disks on EUROPA are:

Device Name	Device Status	Error Count	Volume Label	Free Blocks	Trans Count	Mnt Cnt
EUROPA\$DUA0:	Online	0				
EUROPA\$DUA1:	Online	0				

Which disk can be used for paging and swapping? EUROPA\$ DUA0:  
May this procedure INITIALIZE EUROPA\$DUA0: [YES]? NO  
Mounting EUROPA\$DUA0:..  
PAGEFILE.SYS already exists on EUROPA\$DUA0:  
\*\*\*\*\*  
Directory EUROPA\$DUA0:[SYS0.SYSEXE]  
PAGEFILE.SYS;1        23600/23600  
Total of 1 file, 23600/23600 blocks.  
\*\*\*\*\*  
What is the file specification for the page file on EUROPA\$DUA0: [ <SYS0.SYSEXE>PAGEFILE.SYS ]?   
%CREATE-I-EXISTS, EUROPA\$DUA0:<SYS0.SYSEXE> already exists  
This procedure will use the existing pagefile,  
EUROPA\$DUA0:<SYS0.SYSEXE>PAGEFILE.SYS;.  
SWAPFILE.SYS already exists on EUROPA\$DUA0:  
\*\*\*\*\*  
Directory EUROPA\$DUA0:[SYS0.SYSEXE]  
SWAPFILE.SYS;1        12000/12000  
Total of 1 file, 12000/12000 blocks.  
\*\*\*\*\*

(continued on next page)

**Example 7-2 (Cont.) Sample Interactive CLUSTER\_CONFIG.COM Session to Add a Satellite with Local Page and Swap Files**

What is the file specification for the swap file on  
EUROPA\$DUA0: [ <SYS0.SYSEXE>SWAPFILE.SYS ]?   
This procedure will use the existing swapfile,  
EUROPA\$DUA0:<SYS0.SYSEXE>SWAPFILE.SYS;.

AUTOGEN will now reconfigure and reboot EUROPA automatically.  
These operations will complete in a few minutes, and a  
completion message will be displayed at your terminal.

The configuration procedure has completed successfully.

**7.5.1.1 Updating Network Data After Adding a Satellite**

Whenever you add a satellite, CLUSTER\_CONFIG.COM updates both the permanent and volatile remote node network databases on the boot server. However, the volatile databases on other cluster members are not automatically updated. To share the new data throughout the cluster, you must update the volatile databases on all other cluster members. Log in as system manager, invoke the SYSMAN Utility, and enter the following commands at the SYSMAN> prompt:

```
SYSMAN> SET ENVIRONMENT/CLUSTER
%SYSMAN-I-ENV, current command environment:
Clusterwide on local cluster
Username LAZARUS will be used on nonlocal nodes
SYSMAN> SET PROFILE/PRIVILEGES=(OPER, SYSPRV)
SYSMAN> DO MCR NCP SET KNOWN NODES ALL
%SYSMAN-I-OUTPUT, command execution on node X...
.
.
SYSMAN> EXIT
$
```

Note that the file NETNODE\_REMOTE.DAT must be located in the directory SYS\$COMMON:[SYSEXE].

**7.5.1.2 Restoring a Satellite's Network Data**

The first time you execute CLUSTER\_CONFIG.COM to add a satellite, the procedure creates the file NETNODE\_UPDATE.COM in the boot server's SYS\$SPECIFIC:[SYSMGR] directory. (For a common-environment cluster, you must rename this file to the SYS\$COMMON:[SYSMGR] directory, as described in Section 4.5.4.) This file, which is updated each time you add or remove a satellite, or change its Ethernet or FDDI hardware address, contains all essential network configuration data for the satellite. If an unexpected condition at your site should cause configuration data to be lost, you can use NETNODE\_UPDATE.COM to restore it. You can also read the file when you need to obtain data about individual satellites. Note that you may want to edit the file occasionally to remove obsolete entries.

Example 7-3 shows the contents of the file after satellites EUROPA and GANYMD have been added to the cluster.

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

#### Example 7–3 Sample NETNODE\_UPDATE.COM File

```
$ run sys$system:ncp
  define node EUROPA address 2.21
  define node EUROPA hardware address 08-00-2B-03-51-75
  define node EUROPA load assist agent sys$share:niscs_laa.exe
  define node EUROPA load assist parameter $1$DJAI1:<SYS10.>
  define node EUROPA tertiary loader sys$system:tertiary_vmb.exe
  define node GANYMD address 2.22
  define node GANYMD hardware address 08-00-2B-03-58-14
  define node GANYMD load assist agent sys$share:niscs_laa.exe
  define node GANYMD load assist parameter $1$DJAI1:<SYS11.>
  define node GANYMD tertiary loader sys$system:tertiary_vmb.exe
```

#### 7.5.1.3 Controlling Clusterwide Broadcast Messages on Satellites and Boot Servers

When a satellite joins the cluster, broadcasts for all message classes are initially enabled for the satellite by default. Users can disable such broadcasts selectively by including a form of the DCL command SET BROADCAST in their LOGIN.COM files. For example, the following command would disable OPCOM and shutdown messages:

```
$ SET BROADCAST=(NOOPCOM,NOSHUTDOWN)
```

Note that broadcasts to the operator console terminal (OPA0:) on satellite workstation computers are disabled by default and should remain disabled at all times. Users who want to receive broadcast messages can create a terminal window and then enter the DCL command REPLY/ENABLE. (This command requires OPER privilege.) For more detailed information on workstation operations, refer to the documentation supplied with the workstation software.

In large clusters, state transitions (computers joining or leaving the cluster) generate many multiline OPCOM messages on a boot server's console device. You can abbreviate such messages by including the DCL command REPLY /DISABLE=CLUSTER in the appropriate site-specific startup command file or by entering the command interactively from the system manager's account.

#### 7.5.2 Removing a Computer from the Cluster

You must shut down a computer before removing it from the cluster. If possible, use the command procedure SYS\$SYSTEM:SHUTDOWN.COM to perform an orderly shutdown. Otherwise, halt the computer.

Note that because the REMOVE function deletes the computer's entire root directory tree, it generates VMS RMS error messages while deleting directory files. You can ignore these messages.

Whenever you remove a voting member, you must, after the REMOVE function completes, reconfigure the cluster, following instructions in Section 7.6.

Example 7–4 illustrates the use of CLUSTER\_CONFIG.COM on JUPITR to remove satellite EUROPA from the cluster.

---

#### Note

---

If the page and swap files for the computer being removed do not reside on the same disk as the computer's root directory tree, the REMOVE function does not delete these files. It displays a message warning that the files will not be deleted, as in Example 7–4. If you want to delete the files, you must do so after the REMOVE function completes.

---

**Example 7-4 Sample Interactive CLUSTER\_CONFIG.COM Session to Remove a Satellite with Local Page and Swap Files**

```
$ @CLUSTER_CONFIG.COM

Cluster Configuration Procedure

Use CLUSTER_CONFIG.COM to set up or change a VAXcluster configuration.
To ensure that you have the required privileges, invoke this procedure
from the system manager's account.

Enter ? for help at any prompt.

    1. ADD a node to the cluster.
    2. REMOVE a node from the cluster.
    3. CHANGE a cluster node's characteristics.
    4. CREATE a second system disk for JUPITR.

Enter choice [1]: 2

The REMOVE function disables a node as a cluster member.

    o It deletes the node's root directory tree.
    o It removes the node's network information
      from the network database.

If the node being removed is a voting member, you must adjust
EXPECTED_VOTES in each remaining cluster member's MODPARAMS.DAT.
You must then reconfigure the cluster, using the procedure described
in the VMS VAXcluster Manual.

What is the node's DECnet node name? EUROPA
Verifying network database...
Verifying that SYS10 is EUROPA's root...

    WARNING - EUROPA's page and swap files will not be deleted.
              They do not reside on $1$DJ11:.

Deleting directory tree SYS10...
%DELETE-I-FILDEL, $1$DJ11:<SYS10>SYSCBI.DIR;1 deleted (1 block)
%DELETE-I-FILDEL, $1$DJ11:<SYS10>SYSERR.DIR;1 deleted (1 block)
.
.
.
System root SYS10 deleted.
Updating network database...
The configuration procedure has completed successfully.
```

### 7.5.3 Changing a Computer's Characteristics

You select the **CHANGE** function when you want to accomplish any of the operations described in Table 7-3. When you select this function, **CLUSTER\_CONFIG.COM** displays a menu of **CHANGE** options. Note that all operations except changing a satellite's LAN (Ethernet or FDDI) hardware address must be executed on the computer whose characteristics you want to change.

Before adding computers in a new local area or mixed-interconnect cluster, you must execute the **CHANGE** function to enable the first installed computer as a boot server (see Example 7-7).

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

#### Caution

Whenever you enable or disable disk-serving functions, you must run AUTOGEN with the REBOOT option to reboot the local computer. For all other CHANGE operations (except changing a satellite's hardware address), you must reconfigure the cluster, following instructions in Section 7.6.

**Table 7-3 CLUSTER\_CONFIG.COM CHANGE Options**

Option	Operation Performed
Enable the local computer as a disk server.	Load the MSCP server by setting, in MODPARAMS.DAT, the value of the MSCP_LOAD parameter to 1, and setting an appropriate value for the MSCP_SERVE_ALL parameter.
Disable the local computer as a disk server.	Set MSCP_LOAD to 0.
Enable the local computer as a boot server.	If you are setting up a local area or mixed-interconnect cluster, you must execute this operation once before you attempt to add computers to the cluster. You thereby enable DECnet MOP service for the LAN adapter circuit that the computer uses to service operating system load requests from satellites. When you enable the computer as a boot server, it automatically becomes a disk server (if it is not one already), because it must serve its system disk to satellites.
Disable the local computer as a boot server.	Disable DECnet MOP service for the computer's adapter circuit.
Enable the LAN for cluster communications on the local computer.	Load the VAXport driver PEDRIVER by setting the value of the NISCS_LOAD_PEA0 parameter to 1 in MODPARAMS.DAT. Create the cluster security database file, SYS\$SYSTEM:[SYSEXE]CLUSTER_AUTHORIZE.DAT, on the local computer's system disk.
Disable the LAN for cluster communications on the local computer.	Set NISCS_LOAD_PEA0 to 0.
Enable a quorum disk on the local computer.	In MODPARAMS.DAT, set an appropriate value for the SYSGEN parameter DISK_QUORUM; set the value of QDSKVOTES to 1 (default value).
Disable a quorum disk on the local computer.	In MODPARAMS.DAT, set a blank value for the SYSGEN parameter DISK_QUORUM; set the value of QDSKVOTES to 1.

(continued on next page)



Table 7-3 (Cont.) CLUSTER\_CONFIG.COM CHANGE Options

Option	Operation Performed
Change the local computer's allocation class value.	Set a value for the computer's ALLOCLASS parameter in MODPARAMS.DAT.
Change a satellite's LAN hardware address.	Change a satellite's hardware address if its LAN device needs replacement. Both the permanent and volatile network databases, and NETNODE_UPDATE.COM, are updated on the local computer. <i>You must execute this operation on a computer enabled as a boot server for the satellite.</i>

**Note**

When CLUSTER\_CONFIG.COM sets or changes values in MODPARAMS.DAT, the new values are always appended at the end of the file, so that they override earlier values. You may want to edit the file occasionally and delete lines that specify earlier values.

Example 7-5, Example 7-6, Example 7-7, and Example 7-8, respectively, show the use of CLUSTER\_CONFIG.COM to perform the following operations:

- Enable node URANUS as a disk server
- Change node URANUS's ALLOCLASS value
- Enable node URANUS as a boot server
- Specify a new hardware address for satellite node ARIEL, which boots from URANUS's system disk.

**Example 7-5 Sample Interactive CLUSTER\_CONFIG.COM Session to Enable the Local Computer as a Disk Server**

```
$ @CLUSTER_CONFIG.COM
```

Cluster Configuration Procedure

Use CLUSTER\_CONFIG.COM to set up or change a VAXcluster configuration. To ensure that you have the required privileges, invoke this procedure from the system manager's account.

Enter ? for help at any prompt.

1. ADD a node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster node's characteristics.
4. CREATE a second system disk for URANUS.

Enter choice [1]: 3

(continued on next page)

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

#### Example 7-5 (Cont.) Sample Interactive CLUSTER\_CONFIG.COM Session to Enable the Local Computer as a Disk Server

CHANGE Menu

1. Enable URANUS as a disk server.
2. Disable URANUS as a disk server.
3. Enable URANUS as a boot server.
4. Disable URANUS as a boot server.
5. Enable Ethernet for cluster communications on URANUS.
6. Disable Ethernet for cluster communications on URANUS.
7. Enable a quorum disk on URANUS.
8. Disable a quorum disk on URANUS.
9. Change URANUS's ALLOCLASS value.
10. Change a satellite's hardware address.

Enter choice [1]:

Will URANUS serve HSC disks [Y]?

Enter a value for URANUS's ALLOCLASS parameter: 2

The configuration procedure has completed successfully.

URANUS has been enabled as a disk server. MSCP\_LOAD has been set to 1 in MODPARAMS.DAT. Please run AUTOGEN to reboot URANUS:

```
$ @SYS$UPDATE:AUTOGEN GETDATA REBOOT
```

If you have changed URANUS's ALLOCLASS value, you must reconfigure the cluster, using the procedure described in the VMS VAXcluster Manual.

#### Example 7-6 Sample Interactive CLUSTER\_CONFIG.COM Session to Change the Local Computer's ALLOCLASS Value

```
$ @CLUSTER_CONFIG.COM
```

Cluster Configuration Procedure

Use CLUSTER\_CONFIG.COM to set up or change a VAXcluster configuration. To ensure that you have the required privileges, invoke this procedure from the system manager's account.

Enter ? for help at any prompt.

1. ADD a node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster node's characteristics.
4. CREATE a second system disk for URANUS.

Enter choice [1]: 3

CHANGE Menu

1. Enable URANUS as a disk server.
2. Disable URANUS as a disk server.
3. Enable URANUS as a boot server.
4. Disable URANUS as a boot server.
5. Enable Ethernet for cluster communications on URANUS.
6. Disable Ethernet for cluster communications on URANUS.
7. Enable a quorum disk on URANUS.
8. Disable a quorum disk on URANUS.
9. Change URANUS's ALLOCLASS value.
10. Change a satellite's hardware address.

Enter choice [1]: 9

(continued on next page)

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

#### Example 7-6 (Cont.) Sample Interactive CLUSTER\_CONFIG.COM Session to Change the Local Computer's ALLOCLASS Value

Enter a value for URANUS's ALLOCLASS parameter [2]: 1  
The configuration procedure has completed successfully

If you have changed URANUS's ALLOCLASS value, you must reconfigure the cluster, using the procedure described in the VMS VAXcluster Manual.

#### Example 7-7 Sample Interactive CLUSTER\_CONFIG.COM Session to Enable the Local Computer as a Boot Server

\$ @CLUSTER\_CONFIG.COM

##### Cluster Configuration Procedure

Use CLUSTER\_CONFIG.COM to set up or change a VAXcluster configuration. To ensure that you have the required privileges, invoke this procedure from the system manager's account.

Enter ? for help at any prompt.

1. ADD a node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster node's characteristics.
4. CREATE a second system disk for URANUS.

Enter choice [1]: 3

##### CHANGE Menu

1. Enable URANUS as a disk server.
2. Disable URANUS as a disk server.
3. Enable URANUS as a boot server.
4. Disable URANUS as a boot server.
5. Enable Ethernet for cluster communications on URANUS.
6. Disable Ethernet for cluster communications on URANUS.
7. Enable a quorum disk on URANUS.
8. Disable a quorum disk on URANUS.
9. Change URANUS's ALLOCLASS value.
10. Change a satellite's hardware address.

Enter choice [1]: 3

Verifying circuits in network database...

Updating permanent network database...

In order to enable or disable DECnet MOP service in the volatile network database, DECnet traffic must be interrupted temporarily.

Do you want to proceed [Y]?

Enter a value for URANUS's ALLOCLASS parameter [1]:   
The configuration procedure has completed successfully.

URANUS has been enabled as a boot server. Disk serving and Ethernet capabilities are enabled automatically. If URANUS was not previously set up as a disk server, please run AUTOGEN to reboot URANUS:

\$ @SYSSUPDATE:AUTOGEN GETDATA REBOOT

If you have changed URANUS's ALLOCLASS value, you must reconfigure the cluster, using the procedure described in the VMS VAXcluster Manual.

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

#### Example 7-8 Sample Interactive CLUSTER\_CONFIG.COM Session to Change a Satellite's Hardware Address

```
$ @CLUSTER_CONFIG.COM
```

```
Cluster Configuration Procedure
```

```
Use CLUSTER_CONFIG.COM to set up or change a VAXcluster configuration.
To ensure that you have the required privileges, invoke this procedure
from the system manager's account.
```

```
Enter ? for help at any prompt.
```

1. ADD a node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster node's characteristics.
4. CREATE a second system disk for URANUS.

```
Enter choice [1]: 3
```

```
CHANGE Menu
```

1. Enable URANUS as a disk server.
2. Disable URANUS as a disk server.
3. Enable URANUS as a boot server.
4. Disable URANUS as a boot server.
5. Enable Ethernet for cluster communications on URANUS.
6. Disable Ethernet for cluster communications on URANUS.
7. Enable a quorum disk on URANUS.
8. Disable a quorum disk on URANUS.
9. Change URANUS's ALLOCLASS value.
10. Change a satellite's hardware address.

```
Enter choice [1]: 10
```

```
What is the node's DECnet node name? ARIEL
```

```
What is the new hardware address [XX-XX-XX-XX-XX-XX]? 08-00-3B-05-37-78
```

```
Updating network database...
```

```
The configuration procedure has completed successfully.
```

#### 7.5.4 Changing the Cluster Configuration Type

As your processing needs change, you may want to add satellites to an existing CI-based cluster, or you may want to add CI-connected computers or HSC subsystems to an existing local area cluster. In either case, you can use CLUSTER\_CONFIG.COM to convert your existing cluster to a mixed-interconnect configuration.

##### 7.5.4.1 Changing an Existing CI-Only Cluster to a Mixed-Interconnect Configuration

If you want to convert an existing CI-only cluster to a mixed-interconnect configuration, you must enable cluster communications over the LAN on all computers, and you must enable one or more computers as boot servers. Proceed as follows:

1. Log in as system manager on each computer, invoke CLUSTER\_CONFIG.COM, and execute the CHANGE function to enable the LAN for cluster communications. *You must perform this operation on all computers.*
2. Execute the CHANGE function to enable one or more computers as boot servers.
3. Shut down and reboot the cluster, following instructions in Section 7.6.

#### 7.5.4.2 Changing an Existing Local Area Cluster to a Mixed-Interconnect Configuration

Before performing the operations described in this section, be sure that the computers and HSC subsystems you intend to include in your new mixed-interconnect configuration are correctly installed and checked for proper operation.

The method you use to convert an existing local area cluster to a mixed-interconnect configuration depends on whether your current boot server is a CI-capable computer. Note that the following procedures assume that the system disk containing satellite roots will reside on an HSC disk.

If the boot server is a CI-capable computer, proceed as follows:

1. Log in as system manager on the boot server and perform an image backup operation to back up the current system disk to a disk on an HSC subsystem. (For complete information on backup operations, refer to the *VMS Backup Utility Manual*.)
2. Modify the computer's default bootstrap command procedure to boot the computer from the HSC disk, following instructions in the appropriate computer-specific installation and operations guide.
3. Shut down the cluster. Shut down the satellites first, and then shut down the boot server.
4. Boot the boot server from the newly created system disk on the HSC subsystem.
5. Reboot the satellites.

If your current boot server is not a CI-capable computer, proceed as follows:

1. Shut down the old local area cluster. Shut down the satellites first, and then shut down the boot server.
2. Install the VMS operating system on the new CI-connected computer's HSC system disk. When the installation procedure asks if you want to enable the LAN for cluster communications, answer YES.
3. When the installation completes, log in as system manager and configure and start the DECnet-VAX network, as described in Chapter 4.
4. Execute the CLUSTER\_CONFIG.COM CHANGE function to enable the computer as a boot server.
5. Log in as system manager on the newly added CI-connected computer and execute CLUSTER\_CONFIG.COM's ADD function to add the former local area cluster members (including the former boot server) as satellites on the new HSC system disk.

#### 7.5.5 Converting a Standalone Computer to a VAXcluster Computer

You execute CLUSTER\_CONFIG.COM on a standalone computer to perform the following operations:

- Add the standalone computer to an existing cluster.
- Set up the standalone computer to form a new cluster, if the computer was not set up as a cluster computer during installation of the VMS operating system.

## Building and Maintaining the Cluster

### 7.5 Configuring the Cluster

Example 7-9 illustrates the use of CLUSTER\_CONFIG.COM on standalone computer PLUTO to convert PLUTO to a cluster boot server.

#### Example 7-9 Sample Interactive CLUSTER\_CONFIG.COM Session to Convert a Standalone Computer to a Cluster Boot Server

```
$ @CLUSTER_CONFIG.COM

Cluster Configuration Procedure

This procedure sets up this standalone node to join an existing
cluster or to form a new cluster.

What is the node's DECnet node name? PLUTO
What is the node's DECnet address? 2.5
Will the Ethernet be used for cluster communications (Y/N)? Y
Enter this cluster's group number: 3378
Enter this cluster's password:
Re-enter this cluster's password for verification:
Will PLUTO be a boot server [Y]? 
Verifying circuits in network database...
Enter a value for PLUTO's ALLOCLASS parameter: 1
Does this cluster contain a quorum disk [N]? 

AUTOGEN computes the SYSGEN parameters for your configuration
and then reboots the system with the new parameters.
```

#### 7.5.6 Creating a Duplicate System Disk

As you continue to add computers to a VAXcluster system with a single common system disk, you may eventually reach the disk's storage or I/O capacity. In that case, you may want to add one or more common system disks to handle the increased load. You can use CLUSTER\_CONFIG.COM to set up these disks. Proceed as follows, *after* you have coordinated cluster common files, as described in Section 4.5.4.

1. Log in as system manager.
2. Place a blank disk in an appropriate drive and spin up the disk.
3. Invoke CLUSTER\_CONFIG.COM and select the CREATE function. The procedure prompts for the device names of the current and new system disks, as shown in Example 7-10. It then backs up the current system disk to the new one, deletes all directory roots from the new disk, and mounts that disk clusterwide. Note that VMS RMS error messages are displayed while the procedure deletes directory files. You can ignore these messages.

## Building and Maintaining the Cluster 7.5 Configuring the Cluster

### Example 7-10 Sample Interactive CLUSTER\_CONFIG.COM CREATE Session

```
$ @CLUSTER_CONFIG.COM
```

#### Cluster Configuration Procedure

Use CLUSTER\_CONFIG.COM to set up or change a VAXcluster configuration. To ensure that you have the required privileges, invoke this procedure from the system manager's account.

Enter ? for help at any prompt.

1. ADD a node to the cluster.
2. REMOVE a node from the cluster.
3. CHANGE a cluster node's characteristics.
4. CREATE a second system disk for JUPITR.

Enter choice [1]: 4

The CREATE function generates a duplicate system disk.

- o It backs up the current system disk to the new system disk.
- o It then removes from the new system disk all system roots.

WARNING - Do not proceed unless you have defined appropriate logical names for cluster common files in your site-specific startup procedures. For instructions, see the VMS VAXcluster Manual.

Do you want to continue [N]? YES

This procedure will now ask you for the device name of JUPITR's system root. The default device name (DISK\$VAXVMSRL5:) is the logical volume name of SYS\$SYSDEVICE:.

What is the device name of the current system disk [DISK\$VAXVMSRL5:]?

What is the device name for the new system disk? \$1\$DJA16:

%DCL-I-ALLOC, \$1\$DJA16: allocated

%MOUNT-I-MOUNTED, SCRATCH mounted on \_\$1\$DJA16:

What is the unique label for the new system disk [JUPITR\_SYS2]?

Backing up the current system disk to the new system disk...

Deleting all system roots...

Deleting directory tree SYS1...

%DELETE-I-FILDEL, \$1\$DJA16:<SYS0>DECNET.DIR;1 deleted (2 blocks)

.  
.  
.

System root SYS1 deleted.

Deleting directory tree SYS2...

%DELETE-I-FILDEL, \$1\$DJA16:<SYS1>DECNET.DIR;1 deleted (2 blocks)

.  
.  
.

System root SYS2 deleted.

All the roots have been deleted.

%MOUNT-I-MOUNTED, JUPITR\_SYS2 mounted on \_\$1\$DJA16:

The second system disk has been created and mounted clusterwide. Satellites can now be added.

## Building and Maintaining the Cluster

### 7.6 Reconfiguring the Cluster After a Major Change

## 7.6 Reconfiguring the Cluster After a Major Change

Because the following operations affect the integrity of the entire cluster, you must reconfigure the cluster after executing any of them:

- Adding or removing a voting member
- Enabling or disabling the LAN for cluster communications
- Enabling or disabling a quorum disk
- Changing allocation class values
- Changing the cluster group number or password (see Section 7.7.8)

In all cases, you must shut down and reboot the entire cluster. Note that if you add or remove a voting member, or if you enable or disable a quorum disk, you must update MODPARAMS.DAT files before shutting down the cluster. To perform reconfiguration tasks, follow instructions in Section 7.6.1, Section 7.6.2, Section 7.6.3, and Section 7.6.5.

### 7.6.1 Updating MODPARAMS.DAT Files to Adjust Cluster Quorum

Whenever you add or remove a voting cluster member, or whenever you enable or disable a quorum disk, you must edit MODPARAMS.DAT in all other cluster members' [SYSx.SYSEX] directories and adjust the value for the SYSGEN parameter EXPECTED\_VOTES appropriately. For example, if you add a voting member, or if you enable a quorum disk, you must increment the value by the number of votes assigned to the new member (usually 1). If you add a voting member with 1 vote and enable a quorum disk with 1 vote on that computer, you must increment the value by 2.

You must then prepare to shut down and reboot the entire cluster. To ensure that the new values take effect when you reboot, log in on each computer as system manager and run AUTOGEN to propagate the values to the computer's VAXVMSSYS.PAR file. Enter the following command:

```
$ @SYSSUPDATE:AUTOGEN GETDATA SETPARAMS
```

Be sure *not* to specify the SHUTDOWN or REBOOT options.

---

#### Caution

---

Do not perform this operation until you are ready to shut down and reboot the entire cluster. If a computer fails and then reboots with the new parameters, normal cluster operations can be seriously compromised.

---

### 7.6.2 Shutting Down the Cluster

After you have run AUTOGEN to set parameter values correctly, you must shut down the entire cluster. Shut down nonvoting members (such as satellites) before shutting down voting members. Log in as system manager on each computer *locally* and enter the following command to perform an orderly shutdown:

```
$ @SYSSSYSTEM:SHUTDOWN
```

When you are prompted for shutdown options, specify CLUSTER\_SHUTDOWN. Note that you must run the shutdown procedure and specify this option on each computer. When all computers have reached a point in the procedure where activity is suspended, you must halt each computer at its console. You cannot



## Building and Maintaining the Cluster

### 7.6 Reconfiguring the Cluster After a Major Change

shut down the entire cluster from one computer. (For more information on the CLUSTER\_SHUTDOWN option see Section 7.7.6.2.)

#### 7.6.3 Changing Allocation Class Values on HSC Subsystems

If it is necessary to change allocation class values on any HSC subsystem, you must do so while the entire cluster is shut down. For example, to change the allocation class value to 1, set the HSC internal door switch to the Enable position, and enter a command sequence like the following at the appropriate HSC consoles:

```
Ctrl/C
HSC> RUN SETSHO
SETSHO> SET ALLOCATE DISK 1
SETSHO> EXIT
SETSHO-Q Rebooting HSC; Y to continue, Ctrl/Y to abort:? Y
```

Restore the HSC internal door switch setting.

#### 7.6.4 Changing Allocation Class Values on DSSI Subsystems

If it is necessary to change allocation class values on any DSSI subsystem, you must enter a command sequence. For example, to change the allocation class value to 1 for a DSSI node R2RCBA from the VAX node BARODA, enter a command sequence like the following at the \$ prompt from the SYSTEM account of node BARODA:

```
$ MCR SYGEN CONN FYA0:/NOADAP/DRIVER=FYDRIVER
$ SH DEV FY

Device Device Error
Name Status Count
FYA0: offline 0

$ SET HOST/DUP/SERVER=MSCP$DUP/TASK=PARAMS R2RCBA
params> set allclass 1
params> write
Changes require controller initialization, ok?[Y/N] Y
Initializing...
%HSCPAD-S-REMPGMEND, Remote program terminated - message number 3.
%PAxx, Port has closed virtual circuit - remote node R2RCBA
%HSCPAD-S-END, control returned to node BARODA
$
```

On a directly connected CPU, after the value is changed, the computer must be rebooted.

#### 7.6.5 Rebooting the Cluster

For VAXclusters with HSC subsystems, after all HSC subsystems have been set and rebooted, reboot each computer. Watch the console listings for unusual messages or warnings.

---

#### Caution

---

In local area and mixed-interconnect clusters, you must reboot boot servers before rebooting satellites.

---

## Building and Maintaining the Cluster

### 7.6 Reconfiguring the Cluster After a Major Change

Note that several new messages may appear. For example, if you have used the `CLUSTER_CONFIG.COM CHANGE` function to enable cluster communications over the LAN, one message reports that the local area VAXcluster security database is being loaded. Then, for every disk-serving computer, another message reports that the MSCP server is being loaded. This message is followed by a list of all the disks being served by that computer. You should verify that all disks are being served in the manner that you specified when you designed the configuration.

For DSSI VAXclusters, after all the DSSI subsystems have been set, reboot the system. Watch the console listings for unusual messages or warnings. Note that several new messages may appear. For every disk-serving computer, a message reports that the MSCP server is being loaded. Note that, for DSSI VAXclusters, there is no message containing the list of all the disks being served by the computer. To verify that all disks are being served in the manner in which you designed the configuration, at the `$` prompt of the node serving the disks, type:

```
$ SHOW DEVICE/SERVED
```

Device:	Queue	Status	Requests	Total Size	Current	Max	Hosts
\$1\$DIA0	Avail			1954050	0	0	0
\$1\$DIA2	Avail			1800020	0	0	0

## 7.7 Maintaining the Cluster

Once your cluster is up and running, you can implement routine site-specific maintenance operations—for example, backing up disks or adding user accounts. You should plan to run `AUTOGEN` with the `FEEDBACK` option on a regular basis, as described in Section 7.7.1.

You should also maintain records of current configuration data, especially any changes to hardware or software components. Section 7.7.2 lists items that should be included in your records.

If you are managing a local area or mixed-interconnect cluster, it is important to monitor LAN activity. Section 7.7.3 provides information to help you set up a monitoring procedure.

From time to time conditions may occur that require the following special maintenance operations:

- Restoring cluster quorum after an unexpected computer failure
- Executing conditional shutdown operations
- Performing security functions in local area and mixed-interconnect clusters

These operations are discussed in Section 7.7.5, Section 7.7.6, and Section 7.7.8, respectively.

### 7.7.1 Running `AUTOGEN` with the `FEEDBACK` Option

`AUTOGEN` includes a mechanism called **feedback**. This mechanism examines data collected during normal system operations, and it adjusts system parameters on the basis of the collected data whenever you run `AUTOGEN` with the `FEEDBACK` option. For example, the system records each instance of a disk server waiting for buffer space to process a disk request. Based on this information, `AUTOGEN` can automatically size the disk server's buffer pool to ensure that sufficient space is allocated.

Digital strongly recommends that you use the FEEDBACK option. Without FEEDBACK, it is difficult for AUTOGEN to anticipate patterns of resource usage, particularly in complex configurations. Factors such as the number of computers and disks in the cluster and the types of applications being run, require adjustment of system parameters for optimal performance.

You should therefore run AUTOGEN with FEEDBACK frequently. As a cluster grows, settings for many parameters must be adjusted. The settings AUTOGEN chooses for a cluster with three CI-connected computers and five satellites will no longer be appropriate when you add more computers or satellites. In summary, you should run AUTOGEN on a regular basis to compensate for changes in user workloads and whenever you make significant changes in your configuration. For detailed information on AUTOGEN, refer to the *Guide to Setting Up a VMS System*.

### 7.7.2 Recording Configuration Data

To maintain a VAXcluster system effectively, you must keep accurate records on the current status of all hardware and software components and on any changes made to those components. Changes to cluster components can have a significant effect on the operation of the entire cluster. If a failure occurs, you will need to consult your records when diagnosing problems.

At a minimum, your configuration records should include the following:

- SCSNODE and SYSSYSTEMID parameter values for all computers.
- DECnet names and addresses for all computers.
- Current values for cluster-related SYSGEN parameters, especially ALLOCLASS and TAPE\_ALLOCLASS values for HSC subsystems and computers. (Cluster SYSGEN parameters are described in Appendix A.)
- Names and locations of default bootstrap command procedures for all CI-connected computers.
- Names of cluster disk and tape devices.
- In local area and mixed-interconnect clusters, LAN hardware addresses for satellites.
- Names of LAN adapters.
- Names of LAN segments or rings.
- Names of LAN bridges.
- Names of wiring concentrators, DELNIs, or DEMPRs
- Serial numbers of all hardware components.
- Changes to any hardware or software components (including site-specific command procedures) along with dates and times when changes were made.

Maintaining current records for your configuration is necessary both for routine operations and for eventual troubleshooting activities. (Section E.2.1 describes collecting information for VAXcluster network failure analysis.)

## Building and Maintaining the Cluster

### 7.7 Maintaining the Cluster

#### 7.7.3 Monitoring LAN Activity in Local Area and Mixed-Interconnect Clusters

In local area and mixed-interconnect clusters, it is important that you monitor LAN (Ethernet or FDDI) activity on a regular basis. Using NCP commands like the following (where BNA-0 is the line-id of the Ethernet line), you can set up a convenient monitoring procedure to report activity for each 12-hour period. Note that DECnet event logging for event 0.2 (automatic line counters) must be enabled. (For detailed information on DECnet-VAX event logging, refer to the *VMS Network Control Program Manual*.)

```
NCP> DEFINE LINE BNA-0 COUNTER TIMER 43200
NCP> SET LINE BNA-0 COUNTER TIMER 43200
```

Every timer interval (in this case 12 hours) DECnet will create an event that sends counter data to the DECnet event log. If you experience a performance degradation in your cluster, check the event log for increases in counter values that exceed normal variations for your cluster. If all computers show the same increase, there may be a general problem with your Ethernet configuration. If, on the other hand, only one computer shows a deviation from usual values, there is probably a problem with that computer or its Ethernet interface device.

The following Digital layered products can be used in conjunction with one of Digital's LAN bridges to monitor the LAN traffic levels:

- Remote Bridge Management Software (RBMS)
- DIGITAL Extended LAN Management Software (DECelms)
- LAN Traffic Monitor (LTM)

#### 7.7.4 Performing VAXcluster Network Failure Analysis

VMS provides in SYS\$EXAMPLES a sample program to enable local area VAXcluster network failure analysis. You can edit and use the program to detect and isolate failed network components. Using the network failure analysis program can help reduce the time required to detect and isolate a failed network component, therefore providing a significant increase in cluster availability. For a description of the network failure analysis program, refer to Appendix E.

#### 7.7.5 Restoring Cluster Quorum After an Unexpected Computer Failure

During the life of a VAXcluster system, computers join and leave the cluster. For example, you may need to add more computers to the cluster to extend the cluster's processing capabilities, or a computer may shut down unexpectedly because of a hardware or fatal software error. The connection management software coordinates these cluster transitions and controls cluster operation.

When a computer shuts down unexpectedly, the remaining computers, with the help of the connection manager, reconfigure the cluster, excluding the computer that shut down. The cluster can survive the failure of the computer and continue process operations, as long as the cluster votes total is greater than the cluster quorum value. If the cluster votes total falls below the cluster quorum value, the cluster suspends the execution of all processes.

For process execution to resume, the cluster votes total must be restored to a value greater than or equal to the cluster quorum value. Often, the required votes are added as computers join or rejoin the cluster. However, waiting for a computer to join the cluster and raising the votes value is not always a simple or convenient remedy. An alternative solution, for example, might be to shut down and reboot all the computers with a lower quorum value.

## Building and Maintaining the Cluster

### 7.7 Maintaining the Cluster

Following the failure of a computer, you may want to run the Show Cluster Utility and examine values for the VOTES, EXPECTED\_VOTES, CL\_VOTES, and CL\_QUORUM fields. (See the *VMS Show Cluster Utility Manual* for a complete description of these fields.) The VOTES and EXPECTED\_VOTES fields show the settings for each cluster member; the CL\_VOTES and CL\_QUORUM fields show the cluster votes total and the current cluster quorum value.

To examine these values, enter the following commands:

```
$ SHOW CLUSTER/CONTINUOUS
COMMAND> ADD VOTES,EXPECTED_VOTES,CL_VOTES,CL_QUORUM
```

---

#### Note

---

If you want to enter SHOW CLUSTER commands interactively, you must specify the /CONTINUOUS qualifier as part of the SHOW CLUSTER command string. If you do not specify this qualifier, SHOW CLUSTER displays cluster status information returned by the DCL command SHOW CLUSTER and returns you to the DCL command level.

---

If the display from the Show Cluster Utility shows the CL\_VOTES value equal to the CL\_QUORUM value, the cluster cannot survive the failure of any remaining voting member. If one of these computers shuts down, all process activity in the cluster stops.

To prevent the disruption of cluster process activity, you can lower the cluster quorum value. You can use the DCL command SET CLUSTER/EXPECTED\_VOTES to adjust the cluster quorum to a value you specify. If you do not specify a value, the operating system calculates an appropriate value for you. You need enter the command on only one computer to propagate the new value throughout the cluster. When you enter the command, the operating system reports the new value.

Normally, you use the SET CLUSTER/EXPECTED\_VOTES command only when a computer is leaving the cluster for an extended period. (For more information on this command, see the *VMS DCL Dictionary*.)

If, for example, you want to change expected votes to set the cluster quorum to 2, enter the following command:

```
$ SET CLUSTER/EXPECTED_VOTES=3
```

The resulting value is  $(3 + 2) / 2 = 2$ .

Note that no matter what value you specify for the SET CLUSTER/EXPECTED\_VOTES command, you cannot increase quorum to a value that is greater than the number of the votes present, nor can you reduce quorum to a value that is half or fewer of the votes present.

To make the new value active clusterwide, you must adjust the SYSGEN parameter EXPECTED\_VOTES in MODPARAMS.DAT files on each VAXcluster computer, and then reconfigure the cluster, following instructions in Section 7.6.

When a computer that was previously a cluster member is ready to rejoin, you must reset the SYSGEN parameter EXPECTED\_VOTES to its original value in MODPARAMS.DAT on all computers and then reconfigure the cluster, following instructions in Section 7.6. You do not need to use the SET CLUSTER/EXPECTED\_VOTES command to increase cluster quorum, because the quorum value is increased automatically when the computer rejoins the cluster.

## Building and Maintaining the Cluster

### 7.7 Maintaining the Cluster

You can also reduce cluster quorum by selecting one of the cluster-related shutdown options described in Section 7.7.6.

#### 7.7.6 Selecting Cluster Shutdown Options

In addition to the default shutdown option NONE, the VMS operating system provides four options for shutting down VAXcluster computers:

- REMOVE\_NODE
- CLUSTER\_SHUTDOWN
- REBOOT\_CHECK
- SAVE\_FEEDBACK

These options are described in Section 7.7.6.1, Section 7.7.6.2, Section 7.7.6.3, and Section 7.7.6.4, respectively.

If you do not select any of these options (that is, if you select the default SHUTDOWN option NONE), the SHUTDOWN procedure performs the normal operations for shutting down a standalone computer. If you want to shut down a computer that you expect will rejoin the cluster shortly, you can specify the default option NONE. In that case, cluster quorum is not adjusted because the VMS operating system assumes that the computer will soon rejoin the cluster.

##### 7.7.6.1 The REMOVE\_NODE Option

If you want to shut down a computer that you expect will not rejoin the cluster for an extended period, select the REMOVE\_NODE option. For example, a computer may be waiting for new hardware, or you may decide that you want to use a computer for standalone operation indefinitely.

When you use the REMOVE\_NODE option, the active quorum in the remainder of the cluster is adjusted downward to reflect the fact that the removed computer's votes no longer contribute to the quorum value. The SHUTDOWN procedure readjusts the quorum by issuing the SET CLUSTER/EXPECTED\_VOTES command, which is subject to the usual constraints described in Section 7.7.5.

Note that it is still the responsibility of the system manager to change the SYSGEN parameter EXPECTED\_VOTES on the remaining VAXcluster computers to reflect the new configuration.

##### 7.7.6.2 The CLUSTER\_SHUTDOWN Option

You must run the SHUTDOWN procedure and specify this option on every VAXcluster computer. Each computer suspends activity, just short of shutting down completely, until all other computers in the cluster have reached the same point in the SHUTDOWN procedure. You must then shut down every computer individually by halting each computer at its console. If any one computer is not completely shut down, clusterwide shutdown cannot occur. Instead, operations on all other computers are suspended. Be sure to shut down nonvoting members (such as satellites) before shutting down other computers.

##### 7.7.6.3 The REBOOT\_CHECK Option

When you select the REBOOT\_CHECK option, the SHUTDOWN procedure checks for the existence of basic system files that are needed to reboot the computer successfully and notifies you if any files are missing. You should replace such files before proceeding. If all files are present, the following informational message appears:

```
%SHUTDOWN-I-CHECKOK, Basic reboot consistency check completed.
```

Note that you can select the REBOOT\_CHECK option separately or in conjunction with either the REMOVE\_NODE or CLUSTER\_SHUTDOWN option. If you select REBOOT\_CHECK with one of the other options, you must specify options as a comma-separated list.

#### 7.7.6.4 The SAVE\_FEEDBACK Option

You select the SAVE\_FEEDBACK option to enable AUTOGEN feedback operation. Note that you should select this option only when a computer has been running long enough to reflect your typical workload. For detailed information on AUTOGEN feedback, see the *Guide to Setting Up a VMS System*.

#### 7.7.7 Rebooting a Satellite with an Operating System on a Local Disk

In some circumstances, cluster software reboots satellites automatically. Before booting a satellite, the boot procedures check for the presence of an operating system on the satellite's local disk. If an operating system is found, that "local" operating system—not the VAXcluster operating system—is booted.

If an operating system is installed on a satellite's local disk, you should take one of the following measures before performing any operation that causes an automatic reboot—for example, executing SYS\$SYSTEM.SHUTDOWN.COM with the REBOOT option or using CLUSTER\_CONFIG.COM to add that satellite to the cluster:

- Rename the directory file ddcu:[000000]SYS0.DIR on the local disk to ddcu:[000000]SYSx.DIR (where SYSx is a root other than SYS0, SYSE, or SYSF). Then enter the DCL command SET FILE/REMOVE to remove the old directory entry for the boot image SYSBOOT.EXE:

```
$ RENAME DUA0:[000000]SYS0.DIR DUA0:[000000]SYS1.DIR
$ SET FILE/REMOVE DUA0:[SYSEX]SYSBOOT.EXE
```

For subsequent reboots of the computer from the local disk, enter a command in the format B/x0000000 at the console-mode prompt (>>>). For example:

```
>>> B/10000000
```

- Disable the local disk. For instructions, refer to your computer-specific installation and operations guide. Note that this option is not available if the satellite's local disk is being used for paging and swapping.

#### 7.7.8 Performing Security Functions in Local Area and Mixed-Interconnect Clusters

Because multiple local area and mixed-interconnect clusters coexist on a single extended LAN, the VMS operating system provides mechanisms to ensure the integrity of individual clusters and to prevent accidental access to a cluster by an unauthorized computer.

Cluster security mechanisms prevent problems that could otherwise occur under circumstances like the following:

- When setting up a new cluster, the system manager specifies a group number identical to that of an existing cluster on the same Ethernet. (This condition is not as unlikely as it may at first appear, because system managers probably do not assign group numbers randomly.) However, if each cluster's password is unique, the new cluster can form independently.
- A satellite user with access to a local system disk tries to join a cluster by executing a conversational SYSBOOT operation at the satellite's console.

## Building and Maintaining the Cluster

### 7.7 Maintaining the Cluster

The following mechanisms are designed to help system managers perform security functions:

- A cluster authorization file (SYS\$COMMON:[SYSEXE]CLUSTER\_AUTHORIZE.DAT), initialized during installation of the VMS operating system or during execution of the CLUSTER\_CONFIG.COM CHANGE function. The file is maintained with the SYSMAN Utility.
- Control of conversational bootstrap operations on satellites.

These mechanisms are discussed in Section 7.7.8.1 and Section 7.7.8.2, respectively.

#### 7.7.8.1 Maintaining Cluster Security Data

Security data is maintained in the cluster authorization file, SYS\$COMMON:[SYSEXE]CLUSTER\_AUTHORIZE.DAT, which contains the cluster group number and (in encrypted form) the cluster password. The file is accessible only to users with the SYSPRV privilege.

Under normal conditions, you need not alter records in the CLUSTER\_AUTHORIZE.DAT file interactively. However, if you suspect a security breach, you may want to change the cluster password. In that case, you use the SYSMAN Utility to make the change.

Note that if your configuration has multiple system disks, each disk must have a copy of CLUSTER\_AUTHORIZE.DAT. You must run the utility to update all copies.

---

#### Caution

---

If you change either the group number or password, you must reboot the entire cluster. For instructions, see Section 7.6.

---

To invoke the SYSMAN Utility, log in as system manager on a boot server and enter the following command:

```
$ RUN SYS$SYSTEM:SYSMAN
```

When the utility responds with the SYSMAN> prompt, you can enter any of the CONFIGURATION commands listed in Table 7-4.

**Table 7-4 Summary of SYSMAN CONFIGURATION Commands for Cluster Authorization**

Command	Qualifiers	Function
HELP CONFIGURATION SET - CLUSTER_AUTHORIZATION	None	Explains the command's functions.
CONFIGURATION SET - CLUSTER_AUTHORIZATION		Updates the cluster authorization file, CLUSTER_AUTHORIZE.DAT, in the directory SYS\$COMMON:[SYSEXE]. (The SET command creates this file if it does not already exist.)
	/GROUP_NUMBER	Specifies a cluster group number. Group number must be in the range from 1 to 4095 or 61440 to 65535.

(continued on next page)



**Table 7–4 (Cont.) Summary of SYSMAN CONFIGURATION Commands for Cluster Authorization**

Command	Qualifiers	Function
	/PASSWORD	Specifies a cluster password. Password may be from 1 to 31 characters in length and may include alphanumeric characters, dollar signs (\$), and underscores (_).
CONFIGURATION SHOW - CLUSTER_AUTHORIZATION	None	Displays the cluster group number.

Example 7–11 illustrates the use of the SYSMAN Utility to change the cluster password.

#### Example 7–11 Sample SYSMAN Session to Change the Cluster Password

```

$ RUN SYSSYSTEM:SYSMAN
SYSMAN> SET ENVIRONMENT/CLUSTER
%SYSMAN-I-ENV, current command environment:
    Clusterwide on local cluster
    Username LAZARUS      will be used on nonlocal nodes
SYSMAN> SET PROFILE/PRIVILEGES=SYSPRV
SYSMAN> CONFIGURATION SET CLUSTER_AUTHORIZATION/PASSWORD=newpassword
%SYSMAN-I-CAFOLDGROUP, existing group will not be changed
%SYSMAN-I-CAFREBOOT, cluster authorization file updated
    The entire cluster should be rebooted.
SYSMAN> EXIT
$

```

#### 7.7.8.2 Controlling Conversational Bootstrap Operations for Satellites

When you add a satellite to the cluster using CLUSTER\_CONFIG.COM, the procedure asks whether you want to allow conversational bootstrap operations for the satellite (default is NO). If you press the Return key, SYSGEN parameter NISCS\_CONV\_BOOT in the satellite's SYSGEN parameter file remains set to 0 to disable such operations. The parameter file, VAXVMSSYS.PAR, resides in the satellite's root directory on a boot server's system disk (*device:[SYSx.SYSEXE]*). You can later enable conversational bootstrap operations for a given satellite at any time by setting this parameter to 1.

For example, to enable such operations for a satellite booted from root 10 on device \$1\$DJA11, you would proceed as follows:

1. Log in as system manager on the boot server.
2. Invoke the System Generation Utility (SYSGEN) and enter the following commands:

```

$ RUN SYSSYSTEM:SYSGEN
SYSGEN> USE $1$DJA11:[SYS10.SYSEXE]VAXVMSSYS.PAR
SYSGEN> SET NISCS_CONV_BOOT 1
SYSGEN> WRITE $1$DJA11:[SYS10.SYSEXE]VAXVMSSYS.PAR
SYSGEN> EXIT
$

```

The change remains in effect until the next time AUTOGEN runs on the satellite or until you reset NISCS\_CONV\_BOOT back to 0.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

## 7.8 Guidelines for Configuring Large Clusters

This section provides guidelines for configuring VAXcluster systems that include many computers—approximately 20 or more—and describes procedures that you might find helpful. Typically, such VAXcluster systems are local area or mixed-interconnect configurations with a large number of satellites. Topics include the following:

- Configuring disk server LAN adapters and memory
- Configuring system disks
- Adding computers to an existing cluster
- Setting up a new large cluster
- Defining the VAXcluster alias

Note that the recommendations in Section 7.8.1, Section 7.8.2, and Section 7.8.3 can prove beneficial in some clusters with fewer than 20 computers.

### 7.8.1 Local Area VAXcluster Satellite Booting

VAXcluster satellites use a single LAN adapter for the initial stages of booting. This section describes selecting this adapter and configuring MOP servers, as well as some troubleshooting support for the early stages of booting.

#### 7.8.1.1 Alternate Adapter Booting

VMS supports booting from any LAN adapter on a local area VAXcluster satellite with multiple LAN adapters. You can use alternate adapter booting to work around broken adapters or network problems. You can also use this feature to boot into different clusters, depending on the adapter you use to boot the system.

To boot a cluster system from an alternate adapter, specify the full device name in the boot command, as in the following example:

```
>>> B XQB0
```

To use alternate adapter booting, you need the physical address of the alternate LAN adapter. You use the address to update the satellite's node definition in the DECnet database on the boot servers, so they recognize the satellite. You can find the LAN address of the additional adapters by using one of the following methods:

- Use the console command `SHOW ETHERNET`.
- Boot the `READ_ADDR` program using the following commands:

```
>>> B/100 XQB0  
Bootfile: READ_ADDR
```

See Section 7.8.1.2 for information on changing the LAN address in the DECnet database.

Until the run-time driver is loaded, the system uses only the LAN adapter that it used for booting. The satellite must have a path to the system disk using this adapter. After the node joins the cluster, `PEDRIVER` starts the local area VAXcluster protocol on the remaining LAN adapters.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

#### 7.8.1.2 Changing the LAN Address in the DECnet Database to Allow a Cluster Satellite to Boot with Any Adapter

DECnet Phase IV supports one LAN hardware address per node definition. To allow a cluster satellite with multiple LAN adapters to use any LAN adapter to boot into the cluster, use one of the following methods:

- Define a synonym node with a different DECnet address. Have the address point to the same cluster satellite root as the existing node definition. You can display the existing node definition with the NCP command `SHOW NODE`. Then you can use the NCP commands `DEFINE NODE` or `SET NODE` to create a synonym. For the command syntax, refer to the *VMS Network Control Program Manual*.
- Create and maintain different DECnet databases on the different boot nodes within the cluster. In each database, list a different LAN address for the same node definition. A system booting from one LAN adapter receives responses from a subset of the boot nodes. The same system booting from a different LAN adapter receives responses from a different subset of the boot nodes.

Once the satellite receives the Maintenance Operations Protocol (MOP) downline load from the MOP server, it uses the booting LAN adapter to connect to any node serving the system disk. The satellite continues to use the booting LAN adapter exclusively until after the run-time drivers are loaded. The satellite then switches over to using the run-time drivers and starts the local area VAXcluster protocol on all of the LAN adapters.

#### 7.8.1.3 Displaying Connection Messages During Cluster Satellite Booting

In previous versions of the VMS operating system, the system did not display connection messages when it accessed the system disk during local area VAXcluster satellite booting. In VMS Version 5.4-3 and later, a system displays these connection messages during a conversational boot. To enable the display, perform the following steps:

1. Enable conversational booting by updating the satellite's `SYSGEN` parameters in the file `VAXVMSSYS.PAR` in the system root on the disk server.
2. Make sure that the `SYSGEN` parameter `NISCS_CONV_BOOT` is set to 1.
3. Set bit zero (0) in register R5.

Displaying connection messages during a satellite boot allows you to determine which system in a large cluster is serving the system disk to a cluster satellite during the boot process. If booting problems occur, you can use this display to help isolate the problem with the system that is currently serving the system disk. Then, if your system has multiple LAN adapters, you can perform LAN adapter isolation.

Perform LAN adapter isolation by disconnecting all but one of the LAN adapters on the system and then rebooting the satellite. If the satellite boots when it is connected to the system disk server, then a different LAN adapter is at fault. Use a different LAN adapter on the system to reattempt the satellite boot until the satellite does not boot. When the system does not boot, you have located the bad adapter.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

#### 7.8.1.4 Configuring MOP Service on DECnet Routers

On a boot node, CLUSTER\_CONFIG.COM enables the DECnet MOP downline load service on the first circuit that is found in the DECnet database. The circuit state and the service (MOP downline load service) state can be displayed using the following command:

```
$ MCR NCP SHOW CHAR KNOWN CIRCUITS
.
.
.
Circuit = SVA-0
State           = on
Service         = enabled
.
.
.
```

The above example shows that circuit SVA-0 is in the ON state with the MOP downline service enabled. This is the correct state to support MOP downline loading for local area VAXcluster satellites.

Some boot nodes may perform routing between multiple extended LANs. In these configurations, you may want to enable the MOP downline load service on the other DECnet circuits. This allows VAXcluster satellites to be connected to both extended LANs, and possibly boot from either LAN. NCP can be used to enable the MOP downline service for each of the circuits. Note that the circuit must be turned off prior to enabling the MOP downline load service.

The following example enables service for the circuit QNA-1.

```
$ MCR NCP SET CIRCUIT qna-1 STATE OFF
$ MCR NCP SET CIRCUIT qna-1 SERVICE ENABLE STATE ON
$ MCR NCP DEFINE CIRCUIT qna-1 SERVICE ENABLE
```

For further details, refer to the *VMS Network Control Program Manual*.

#### 7.8.2 Configuring Disk Server LAN Adapters and Memory

Because disk-serving activity in a large local area or mixed-interconnect VAXcluster system can generate a substantial amount of I/O traffic on the LAN, boot and disk servers should use the highest-bandwidth LAN adapters in the cluster, and can also use multiple LAN adapters in a single system to distribute the load across the LAN adapters. In addition, a large local area or mixed-interconnect cluster should include multiple boot and disk servers to enhance availability and to distribute I/O traffic over several cluster nodes.

Relatively little memory is required to serve disks. Even busy boot and disk servers probably require no more than 0.25 to 0.5 megabyte of physical memory for disk serving activity. However, if boot and disk servers must also support timesharing users or run batch queues for the cluster, the servers should be configured with memory appropriate for those additional tasks.

#### 7.8.3 Configuring System Disks

Depending on the number of computers to be included in a large cluster, you must evaluate the tradeoffs involved in configuring a single system disk or multiple system disks.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

While a single system disk is easier to manage, a large cluster might require more system disk I/O capacity than a single system disk can provide. (Consider using the optional VMS Volume Shadowing software product to increase disk I/O capacity. For more information on VMS Volume Shadowing software, see the *VMS Volume Shadowing Manual*.) To achieve satisfactory performance, multiple system disks might be needed. However, you should recognize the increased system management efforts involved in maintaining multiple system disks.

#### 7.8.3.1 Concurrent User Activity

In clusters with many workstation satellites, the amount and type of user activity on those satellites (for example, any active batch job or other task created on the workstation by or for the user) influence system disk load and therefore the number of satellites that can be supported by a single system disk. For example, if many users are active or run multiple applications simultaneously, the load on the system disk can be significant. Conversely, in an environment where few users are simultaneously active, or where most users run a single application for extended periods, a single system disk might support a large number of satellites. Note, however, that in these environments significant numbers of I/O requests can be directed to application data disks.

This situation is similar to the traditional timesharing model, because the probability is low that most users are simultaneously active at any given time. Thus, while a VAXcluster system can be configured assuming that all users are constantly active, a smaller and less expensive one can be configured for more typical working conditions. The tradeoff is between a more expensive VAXcluster system that handles rare peak loads without performance degradation and a less expensive one that handles most normal activity as well as the more expensive one, but suffers some performance degradation during peak load periods.

Note one difference from the traditional timesharing model. In a timesharing system, the most important shared resource is the processing power of a shared computer. But because each workstation user in a VAXcluster system has a dedicated computer, a user who runs large compute-bound jobs on that dedicated computer does not significantly affect users of other computers in the VAXcluster system.

For clustered workstations, the critical shared resource is a disk server. Thus, if a workstation user runs even a small I/O-intensive job, its effect on other workstations sharing the same disk server might be noticeable.

#### 7.8.3.2 Concurrent Booting Activity

One of the rare times when all VAXcluster computers are simultaneously active is during a cluster reboot—for example, after a power failure. All satellites are waiting to reload the VMS operating system, and as soon as a boot server is available, they begin to boot in parallel. This booting activity places a significant I/O load on the system disk or disks.

For example, Table 7-5 shows system disk I/O activity and elapsed time until login for a single satellite with minimal startup procedures when the satellite is the only one booting. Table 7-6 shows system disk I/O activity and times elapsed between boot server response and login for various numbers of satellites booting from a single system disk. The disk in these examples has a capacity of 40 I/O operations per second.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

Note that the numbers in the tables are fabricated and are meant to provide only a generalized picture of booting activity. Elapsed times until login on satellites in any particular cluster depend on the complexity of the site-specific system startup procedures. Computers in clusters with many layered products or site-specific applications require more system disk I/O operations to complete booting operations.

**Table 7-5 System Disk I/O Activity and Boot Time for Single Satellite**

Total I/O Requests to System Disk	Average System Disk I/O Operations per Second	Elapsed Time Until Login (Minutes)
4200	6	12

**Table 7-6 System Disk I/O Activity and Boot Times for Multiple Satellites**

Number of Satellites	I/Os per Second Requested	I/Os per Second Serviced	Elapsed Time Until Login (Minutes)
1	6	6	12
2	12	12	12
4	24	24	12
6	36	36	12
8	48	40	14
12	72	40	21
16	96	40	28
24	144	40	42
32	192	40	56
48	288	40	84
64	384	40	112
96	576	40	168

While the elapsed times shown in Table 7-6 do not include the time required for the boot server itself to reload, they illustrate that the I/O capacity of a single system disk can be the limiting factor for cluster reboot time.

Note that you can reduce overall cluster boot time by configuring multiple system disks and distributing system roots for computers evenly across those disks. This technique has the advantage of increasing overall system disk I/O capacity but the disadvantage of requiring additional system management effort. For example, layered product installation or VMS operating system upgrades must be repeated once for each system disk.

In a mixed-interconnect VAXcluster system, VMS Volume Shadowing software can be used to increase the I/O capacity of a single system disk. Installations or updates need only be applied once to a volume-shadowed system disk. For clusters with substantial system disk I/O requirements, you can use multiple system disks, each configured as a shadow set.

### 7.8.3.3 Boot Time Costs

When configuring a VAXcluster system for minimum boot times, consider the following:

- Costs of workstations being unavailable during a cluster reboot
- Hardware costs of additional disk drives
- Cost of VMS Volume Shadowing software, if needed
- System management effort required to maintain multiple system disks
- Probability of power interruptions

---

**Note**

---

Sites with stringent demands for uptime should investigate power conditioning options to minimize power interruption problems.

---

### 7.8.3.4 Moving High-Activity Files off System Disks

To reduce I/O activity on system disks, you can move page and swap files for computers off system disks, and you can set up page and swap files for satellites on the satellites' local disks, if such disks are available. You specify the sizes and locations of page and swap files when you run CLUSTER\_CONFIG.COM to add computers.

You should also move off the system disk such high-activity files as the following:

SYSUAF.DAT	NETPROXY.DAT	RIGHTSLIST.DAT
ACCOUNTNG.DAT	VMSMAIL_PROFILE.DATA	QMAN\$MASTER.DAT

To specify the location of the files, follow instructions in Chapter 4.

### 7.8.3.5 Controlling Dump File Size and Creation

Whether your VAXcluster system uses a single common system disk or multiple system disks, you should plan a strategy to manage dump files. Dump file management is especially important for large clusters with a single system disk.

In the event of a software-detected system failure, each computer normally writes the contents of memory to a full dump file on its system disk for analysis. By default, this full dump file is the size of physical memory plus a small number of pages. If system disk space is limited (as is probably the case if a single system disk is used for a large cluster), you may want to specify that no dump file be created for satellites, or that AUTOGEN create a selective dump file. The selective dump file is typically 30 to 60 percent of the size of a full dump file.

You can control dump file size and creation for each computer by specifying appropriate values for the AUTOGEN symbols DUMPSTYLE and DUMPFIL in the computer's MODPARAMS.DAT file. Dump files are specified as shown in Table 7-7.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

**Table 7-7 AUTOGEN Dump File Symbols**

Value Specified	Effect
DUMPSTYLE = 0	Full dump file created (default)
DUMPSTYLE = 1	Selective dump file created
DUMPFIL = 0	No dump file created

#### Caution

While it is possible to configure computers without dump files, the lack of a dump file can make it difficult or impossible to determine the cause of a system failure.

#### 7.8.3.6 Sharing Dump Files

Another option for saving dump file space is to share a single dump file among multiple computers. This technique makes it possible to analyze isolated computer failures. But dumps are lost if multiple computers fail at the same time or if a second computer fails before you can analyze the first failure. Because boot server failures have a greater impact on cluster operation than failures of other computers, you should configure full dump files on boot servers to help ensure speedy problem analysis.

The VMS operating system attempts to ensure that dump files are not unintentionally shared. However, if you want to share dump files, you can follow these steps:

1. Decide whether to use full or selective dump files.
2. Determine the size of the largest dump file needed by any satellite.
3. Select a satellite whose memory configuration is the largest of any in the cluster and do the following:
  - a. Set DUMPSTYLE = 0 (or DUMPSTYLE = 1) in that satellite's MODPARAMS.DAT file.
  - b. Remove any DUMPFIL symbol from the satellite's MODPARAMS.DAT file.
  - c. Run AUTOGEN on that satellite to create a dump file.
4. Rename the dump file to  
SYS\$COMMON:[SYSEXE]SYSDUMP-COMMON.DMP,  
or create a new dump file named SYSDUMP-COMMON.DMP  
in SYS\$COMMON:[SYSEXE].
5. For each satellite that is to share the dump file, do the following:
  - a. Create a file synonym entry for the dump file in the system-specific root. For example, to create a synonym for the satellite using root SYS1E, you could enter a command like the following:

```
$ SET FILE SYS$COMMON:[SYSEXE]SYSDUMP-COMMON.DMP -  
_ $ /ENTER=SYSS$SYSDEVICE:[SYS1E.SYSEXE]SYSDUMP.DMP
```



## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

- b. Add the following lines to the satellite's MODPARAMS.DAT file:

```
DUMPFIL = 0  
DUMPSTY = 0 (or DUMPSTY = 1)
```

6. After a satellite has rebooted, you can delete any SYSDUMP.DMP file in its SYS\$SPECIFIC directory. Note that if the old dump file is deleted before the satellite reboots, the disk space is lost. You can recover the space by entering the DCL command ANALYZE/DISK/REPAIR.

#### 7.8.4 Adding Computers to an Existing Cluster

When a computer is first added to a cluster, SYSGEN parameters that control the computer's system resources are normally adjusted in several steps, as follows:

1. CLUSTER\_CONFIG.COM sets initial parameters that are adequate to boot the computer in a minimum environment.
2. When the computer boots, AUTOGEN runs automatically to size the static operating system (without using any dynamic FEEDBACK data), and the computer reboots into the production environment.
3. After the newly added computer has been subjected to typical use for a day or more, you should manually run AUTOGEN with FEEDBACK to adjust parameters for the production environment.
4. At regular intervals, and whenever a major change occurs in the cluster configuration or production environment, you should manually run AUTOGEN with FEEDBACK to readjust parameters for the changes.

Because, however, the first AUTOGEN run (initiated by CLUSTER\_CONFIG.COM) is performed both in the minimum environment and without FEEDBACK, a newly added computer may be inadequately configured to run in the production environment of some large clusters. For this reason, you might want to implement additional configuration measures like those described in Section 7.8.4.1 and Section 7.8.4.2.

##### 7.8.4.1 Running AUTOGEN with FEEDBACK for Initial Configuration

To ensure that computers are adequately configured for production use when they first join the cluster, you can run AUTOGEN with FEEDBACK automatically as part of the initial boot sequence. While this step adds an additional reboot before the computer can be used for production work, the computer's performance can be substantially improved for the first few days of use.

When a computer first boots into a large cluster, much of the computer's resource utilization is determined by the current cluster configuration. Factors such as the number of computers, the number of disk servers, and the number of disks available or mounted contribute to a fixed minimum resource requirement. Because this minimum does not change with continued use of the computer, FEEDBACK information on the required resources is immediately valid.

Other FEEDBACK information, however, such as that influenced by normal user activity, is not immediately available, because the only "user" has been the system startup process. If AUTOGEN were run with FEEDBACK at this point, some system values might be set too low.

By running a simulated user load at the end of the first production boot, you can ensure that AUTOGEN has reasonable FEEDBACK information. The User Environment Test Package (UETP) supplied with the VMS operating system contains a test that simulates such a load. You can run this test (the UETP

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

LOAD phase) as part of the initial production boot and then run AUTOGEN with FEEDBACK before a user is allowed to log in.

To implement this technique, you can create a command file like that in step 1 of the procedure in Section 7.8.4.2 and submit the file to the computer's local batch queue from the cluster common SYSTARTUP procedure. Your command file conditionally runs the UETP LOAD phase and then reboots the computer with AUTOGEN FEEDBACK.

#### 7.8.4.2 Creating a Command File to Run AUTOGEN with FEEDBACK

As shown in the sample file, UETP lets you specify a typical user load to be run on the computer when it first joins the cluster. The UETP run generates data that AUTOGEN uses to set appropriate SYSGEN values for the computer when rebooting it with FEEDBACK. Note, however, that the default setting for the UETP user load assumes that the computer is used as a timesharing system. This calculation can produce SYSGEN values that might be excessive for a single-user workstation, especially if the workstation has large memory resources. Therefore, you might want to modify the default user load setting, as shown in the sample file.

Follow these steps:

1. Create a command file like the following:

```
$!  
$! ***** SYSS$COMMON:[SYSMGR]UETP_AUTOGEN.COM *****  
$!  
$! For initial boot only, run UETP LOAD phase and  
$! reboot with AUTOGEN FEEDBACK.  
$!  
$ SET NOON  
$ SET PROCESS/PRIVILEGES=ALL  
$!  
$! Run UETP to simulate a user load for a satellite  
$! with 8 simultaneously active user processes. For a  
$! CI-connected computer, allow UETP to calculate the load.  
$!  
$ LOADS = "8"  
$ IF F$GETDVI("PAA0:", "EXISTS") THEN LOADS = ""  
$ @UETP LOAD 1 'loads'  
$!  
$! Create a marker file to prevent resubmission of  
$! UETP_AUTOGEN.COM at subsequent reboots.  
$!  
$ CREATE SYSS$SPECIFIC:[SYSMGR]UETP_AUTOGEN.DONE  
$!  
$! Reboot with AUTOGEN to set SYSGEN values.  
$!  
$ @SYSS$UPDATE:AUTOGEN SAVPARAMS REBOOT FEEDBACK  
$!  
$ EXIT
```

2. Edit the cluster common SYSTARTUP file and add commands like the following at the end of the file. Assume that queues have been started and that a batch queue is running on the newly added computer. Submit UETP\_AUTOGEN.COM to the computer's local batch queue:

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

```
$!  
$ NODE = F$GETSYI("NODE")  
$ IF F$SEARCH ("SYS$SPECIFIC:[SYSMGR]UETP_AUTOGEN.DONE") .EQS. ""  
$ THEN  
$   SUBMIT /NOPRINT /NOTIFY /USERNAME=SYSTEST -  
$     /QUEUE='NODE'_BATCH SYS$MANAGER:UETP_AUTOGEN  
  
$ WAIT FOR UETP:  
$   WRITE SY$OUTPUT "Waiting for UETP and AUTOGEN... 'F$TIME()'"  
$   WAIT 00:05:00.00           ! Wait 5 minutes  
$   GOTO WAIT_FOR_UETP  
$ ENDIF  
$!
```

Note that UETP must be run under the username SYSTEST.

3. Execute CLUSTER\_CONFIG.COM to add the computer.

When you boot the computer, it runs UETP\_AUTOGEN.COM to simulate the user load you have specified and then reboots with AUTOGEN FEEDBACK to set appropriate SYSGEN values.

#### 7.8.5 Setting Up a New Large VAXcluster System

When building a new large cluster, you must be prepared to run AUTOGEN and reboot the cluster several times during the installation. The parameters that AUTOGEN sets for the first computers added to the cluster will probably be inadequate when additional computers are added. Readjustment of parameters is especially critical for boot and disk servers.

One solution to this potential problem is to run UETP\_AUTOGEN.COM to reboot computers at regular intervals as new computers are added. You should run the procedure according to the percentage of growth. For example, each time there is a significant percentage increase in the number of computers (from 5 to 10, from 10 to 20, and so forth), you should run UETP\_AUTOGEN.COM. For best results, the cluster environment should be as close as possible to the final production environment when you run the procedure.

To set up the cluster, you can follow these steps:

1. Configure boot and disk servers, using the CLUSTER\_CONFIG.COM command procedure.
2. Install all layered products and site-specific applications required for the cluster production environment, or as many as possible.
3. Prepare the cluster startup procedures so that they are as close as possible to those to be used in the final production environment.
4. Add a small number of satellites (perhaps 2 or 3), using CLUSTER\_CONFIG.COM.
5. Reboot the cluster to verify that the startup procedures work as expected.
6. After you have verified that startup procedures work, run UETP\_AUTOGEN.COM on every computer's local batch queue to reboot the cluster again and to set initial production environment values. When the cluster has rebooted, all computers should have reasonable parameter settings. However, check the settings to be sure.
7. Add additional satellites to double their number, and then rerun UETP\_AUTOGEN on each computer's local batch queue to reboot the cluster and set values appropriately to accommodate the newly added satellites.

## Building and Maintaining the Cluster

### 7.8 Guidelines for Configuring Large Clusters

8. Repeat the previous step until all satellites have been added.
9. When all satellites have been added, run UETP\_AUTOGEN a final time on each computer's local batch queue to reboot the cluster and to set new values for the production environment.

Note that for best performance, you might not want to run UETP\_AUTOGEN on every computer simultaneously, because the procedure simulates a user load that is probably more demanding than that for the final production environment. A better method is to run UETP\_AUTOGEN on several satellites (those with the least recently adjusted parameters) while adding new computers. This technique increases efficiency, because little is gained when a satellite reruns AUTOGEN shortly after joining the cluster. For example, if the entire cluster is rebooted after 30 satellites have been added, few adjustments are made to system parameter values for the 28th satellite added—only two satellites have joined the cluster since that satellite ran UETP\_AUTOGEN as part of its initial configuration.

#### 7.8.6 Defining the VAXcluster Alias

The VAXcluster alias acts as a single network node identifier for a VAXcluster system. Computers in the cluster can use the alias for communications with other computers in a DECnet-VAX network. A maximum of 64 VAXcluster computers can participate in a VAXcluster alias. If your cluster includes more than 64 computers, you must determine which 64 should participate in the alias and then define the alias on those computers. For detailed information on the VAXcluster alias, refer to the *VMS Networking Manual*.

---

## Cluster SYSGEN Parameters

For systems to boot properly into a cluster, certain system parameters must be set on each cluster computer. Table A-1 lists SYSGEN parameters used in cluster configurations.

**Table A-1 Cluster SYSGEN Parameters**

Parameter	Description
ALLOCLASS	Specifies a numeric value from 0 to 255 to be assigned as the disk allocation class for the computer. The default value is 0.
DISK_QUORUM	The physical device name, in ASCII, of an optional quorum disk. ASCII spaces indicate that no quorum disk is being used. DISK_QUORUM must be defined on one or more cluster computers capable of having a direct (non-MSCP served) connection to the disk. These computers are called <b>quorum disk watchers</b> . The remaining computers (computers with a blank value for DISK_QUORUM) recognize the name defined by the first watcher computer which which they communicate.
EXPECTED_VOTES	Specifies a setting that is used to derive the initial quorum value. This setting is the sum of all VOTES held by potential cluster members.  By default, the value is 1. The connection manager sets a quorum value to a number that will prevent cluster partitioning (see Section 3.1). To calculate quorum, the system uses the following formula:  $\text{estimated quorum} = (\text{EXPECTED\_VOTES} + 2) / 2$
LRPSIZE	Specifies the size (in bytes) of the large request packets. The actual physical memory consumed by a large request packet is LRPSIZE plus overhead for buffer management.  Normally the default value is adequate. The value of LRPSIZE does affect the transfer size used by VAXcluster nodes on an FDDI ring (see Section 2.9.4).
MSCP_LOAD	Controls whether the MSCP server is loaded. Specify 1 to load the server and use the default CPU load rating. A value greater than 1 loads the server and uses this value as a constant load rating. By default, the value is set to zero, and the server is not loaded.
MSCP_SERVE_ALL	Specifies MSCP disk-serving functions when the MSCP server is loaded. The default value of zero specifies that no disks are served. A value of 1 specifies that all available disks are served. A value of 2 specifies that only locally connected (non-HSC) disks are served.
NISCS_CONV_BOOT	During booting as a VAXcluster satellite, specifies whether conversational bootstraps are enabled on the computer. The default value of zero specifies that conversational bootstraps are disabled. A value of 1 enables conversational bootstraps.

(continued on next page)

## Cluster SYSGEN Parameters

**Table A-1 (Cont.) Cluster SYSGEN Parameters**

Parameter	Description
NISCS_LOAD_PEA0	Specifies whether the VAXport driver PEDRIVER is to be loaded to enable cluster communications over the Local Area Network (LAN). The default value of zero specifies that the driver is not loaded. A value of 1 specifies that that driver is loaded.
NISCS_PORT_SERV	Specifies whether data checking is enabled for the computer. The default value of zero specifies that data checking is disabled. The bit setting for the parameter value is: <ul style="list-style-type: none"> <li>• Bit 0—Set, enables data checking on transmit.</li> <li>• Bit 1—Set, enables data checking on receive.</li> <li>• Bits 2-31—Reserved; must be zero.</li> </ul>
QDSKVOTES	Specifies the number of votes contributed to the cluster votes total by a quorum disk. The maximum is 127, the minimum is 0, and the default is 1. This parameter is used only when DISK_QUORUM is defined.
QDSKINTERVAL	Specifies the disk quorum polling interval, in seconds. The maximum value is 32767, the minimum value is 1, and the default is 10. Lower values trade increased overhead cost for greater responsiveness. Digital recommends that this parameter be set to the same value on each cluster computer.
RECNXINTERVAL	Specifies, in seconds, the interval during which the connection manager attempts to reconnect a broken connection to another computer. If a new connection cannot be established during this period, the connection is declared irrevocably broken, and either this computer or the other must leave the cluster. This parameter trades faster response to certain types of system failures against the ability to survive transient faults of increasing duration. Digital recommends that this parameter be set to the same value on each cluster computer. This parameter also affects the tolerance of the VAXcluster system for LAN bridge failures. (See Section 2.9.3.)
TAPE_ALLOCLASS	Specifies a numeric value from 0 to 255 to be assigned as the tape allocation class for tape devices connected to the computer. The default value is 0.
TIMVCFAIL	Specifies the time required for an adapter or virtual circuit failure to be detected. Digital recommends that the default value be used. Digital also recommends that this value be lowered only in VAXclusters of three CPUs or less, that the same value be used on each computer in the cluster, and that dedicated LAN segments be used for cluster I/O.
TMSCP_LOAD	Controls whether the TMSCP server is loaded. Specify 1 to load the server and set all available TMSCP tapes served. By default, the value is set to zero, and the server is not loaded.

(continued on next page)

Table A-1 (Cont.) Cluster SYSGEN Parameters

Parameter	Description
VAXCLUSTER	<p>Controls whether the computer should join or form a cluster. This parameter accepts the following three values:</p> <ul style="list-style-type: none"> <li>• 0—Specifies that the computer will not participate in a cluster.</li> <li>• 1—Specifies that the computer should participate in a cluster if hardware supporting SCS is present (CI, UDA, HSC50).</li> <li>• 2—Specifies that the computer should participate in a cluster.</li> </ul> <p>You should always set this parameter to 2 on computers intended to run in a cluster, 0 on computers that boot from a UDA and are not intended to be part of a cluster, and 1 (the default) otherwise.</p>
VOTES	<p>Specifies the number of votes towards a quorum to be contributed by the computer. By default, the value is 1.</p>
<b>SCS Parameters</b>	
PANUMPOLL	<p>Specifies the number of ports to poll at each interval. Digital recommends that this parameter be set to the same value on each cluster computer.</p>
PASTIMOUT	<p>Specifies the interval at which the CI port driver performs time-based bookkeeping operations. This interval is also the period after which a start handshake datagram is assumed to have timed out.</p> <p>Normally the default value is adequate. Digital recommends that this parameter be set to the same value on each cluster computer.</p>
PASTDGBUF	<p>Specifies the number of datagram receive buffers to queue for the CI port driver's configuration poller, that is, the maximum number of start handshakes that can be in progress simultaneously.</p> <p>Normally the default value is adequate. Digital recommends that this parameter be set to the same value on each cluster computer.</p>
PAMAXPORT	<p>Specifies the maximum number of CI ports the CI port driver polls for a broken port-to-port virtual circuit, or a failed remote computer.</p> <p>You can decrease this parameter in order to reduce polling activity if the hardware configuration has fewer than 16 ports. For example, if the configuration has a total of five ports assigned port numbers 0-4, then you should set PAMAXPORT to 4.</p> <p>The default for this parameter is 15 (poll for all possible ports 0 through 15). Digital recommends that this parameter be set to the same value on each cluster computer.</p>
PANOPOLL	<p>Disables CI polling for ports if set to 1. (The default is 0.) When PANOPOLL is set, a computer will not discover that another computer has shut down or powered down promptly and will not discover a new computer that has booted. This parameter is useful when you want to bring up a computer detached from the rest of the cluster for checkout purposes. It is roughly equivalent to uncabing the computer from the star coupler.</p> <p>PANOPOLL = 0 is the normal setting and is required if you are booting from an HSC.</p>

(continued on next page)

## Cluster SYSGEN Parameters

Table A-1 (Cont.) Cluster SYSGEN Parameters

Parameter	Description
PAPOLLINTERVAL	<p>Specifies in seconds, the polling interval the computer interconnect (CI) port driver uses to poll for a newly booted computer, a broken port-to-port virtual circuit, or a failed remote computer.</p> <p>This parameter trades polling overhead against quick response to virtual circuit failures. Digital recommends that you use default value for this parameter.</p> <p>Digital recommends that this parameter be set to the same value on each cluster computer.</p>
PAPOOLINTERVAL	<p>Specifies in seconds, the interval at which the PA port driver checks for available nonpaged pool after a failure to allocate.</p> <p>Normally the default value is adequate.</p>
PASANITY	<p>Controls whether the port sanity timer is enabled to permit remote computers to detect a computer that has been halted or retained at IPL 7 for a prolonged period. This parameter is normally set to 1 and should only be set to 0 when debugging with XDELTA.</p> <p>PASANITY is a dynamic parameter (altered the next time the port is initialized) and has a default value of 1.</p>
PRCPOLINTERVAL	<p>Specifies, in seconds, the polling interval used to look for SCS applications, such as the connection manager and MSCP disks, on other computers. Each computer is polled, at most, once each interval.</p> <p>This parameter trades polling overhead against quick recognition of new computers or servers as they appear. Digital recommends that you set this parameter to 15, which is the default.</p>
SCSBUFFCNT	<p>Specifies the number of computer interconnect (CI) buffer descriptors configured for all CI ports on the computer.</p>
SCSCONNCNT	<p>Specifies the total number of SCS connections that are configured for use by all system applications.</p> <p>Normally, the default value is adequate.</p>
SCSMAXMSG	<p>Specifies the SCS maximum sequenced message size.</p> <p>Normally, the default value is adequate.</p>
SCSMAXDG	<p>Specifies the maximum number of bytes of application data in one datagram.</p> <p>Normally the default value is adequate.</p>
SCSFLOWCUSH	<p>Specifies the lower limit for receive buffers at which point SCS starts to notify the remote SCS of new receive buffers. For each connection, SCS tracks the number of receive buffers available. SCS communicates this number to the SCS at the remote end of the connection. However, SCS does not need to do this for each new receive buffer added. Instead, SCS notifies the remote SCS of new receive buffers if the number of receive buffers falls as low as the SCSFLOWCUSH value.</p> <p>Normally the default value is adequate.</p>
SCSSYSTEMID	<p>Specifies the lower-order 32 bits of the 48-bit system identification number. This parameter is not dynamic and must be the same as the DECnet computer number (1024 * &lt;DECnet area&gt; + DECnet computer number).</p>

(continued on next page)



**Table A-1 (Cont.) Cluster SYSGEN Parameters**

Parameter	Description
SCSSYSTEMIDH	Specifies the high-order 16 bits of the 48 bit system identification number. This parameter must be set to 0. It is reserved by Digital for future use.
SCSNODE	Specifies the SCS system name. This parameter is not dynamic. You should use a name that is the same as the DECnet computer name (limited to six characters) since the name must be unique among all computers in the cluster.  Note that once a computer has been recognized by another computer in the cluster, you cannot change the SCSSYSTEMID or SCSNODE parameter without changing both.
SCSRESPCNT	Specifies the total number of response descriptor table entries configured for use by all system applications.

VMS SYSGEN parameters, including cluster SYSGEN parameters, are specified in the *VMS System Generation Utility Manual*. Volume shadowing SYSGEN parameters are described in the *VMS Volume Shadowing Manual*.



---

## Building a Common SYSUAF.DAT File

This appendix provides guidelines for building a common user authorization file from computer-specific files. For more detailed information on how to set up a computer-specific authorization file, see the descriptions in the *VMS Authorize Utility Manual* and in the *Guide to Setting Up a VMS System*.

To build a common SYSUAF.DAT file, follows these steps:

1. Print a listing of SYSUAF.DAT on each computer. To print this listing, invoke `AUTHORIZE` and specify the `AUTHORIZE` command `LIST` as follows:

```
$ SET DEF SYSSYSTEM
$ RUN AUTHORIZE
UAF> LIST/FULL [*,*]
```

2. Use the listings to compare the accounts from each computer. On the listings, mark down any necessary changes.

One such change is to delete any accounts that you no longer need. You should also make sure that each user account in the cluster has a unique UIC.

For example, VAXcluster member VENUS may have a user account JONES that has the same UIC as user account SMITH on computer MARS. When computers VENUS and MARS are joined to form a cluster, accounts JONES and SMITH will exist in the cluster environment with the same UIC. If the UICs of these accounts are not differentiated, each user will have the same access rights to various objects in the cluster. In this case you should assign each account a unique UIC.

Make sure that accounts that perform the same type of work have the same group UIC. Accounts in a single-computer environment probably follow this convention. However, there may be groups of users on each computer that will perform the same work in the cluster but have group UICs unique to their local computer. As a rule, the group UIC for any given work category should be the same on each computer in the cluster. For example, data entry accounts on VENUS should have the same group UIC as data entry accounts on MARS.

Note that if you change the UIC for a particular user, you should also change the owner UICs for that user's existing files and directories. You can use the DCL commands `SET FILE` and `SET DIRECTORY` to make these changes. These commands are described in detail in the *VMS DCL Dictionary*.

3. Choose the SYSUAF.DAT file from one of the computers to be a master SYSUAF.DAT.

## Building a Common SYSUAF.DAT File

4. Merge the SYSUAF.DAT files from the other computers to the master SYSUAF.DAT by running the Convert Utility (CONVERT) on the computer that owns the master SYSUAF.DAT. (See the *VMS Convert and Convert /Reclaim Utility Manual* for a description of CONVERT.) To use CONVERT to merge the files, each SYSUAF.DAT file must be accessible to the computer that is running CONVERT.

To merge the UAFs into the master SYSUAF.DAT file, specify the CONVERT command in the following format:

```
CONVERT SYSUAF1,SYSUAF2,...SYSUAFn MASTER_SYSUAF
```

Note that if a given user name appears in more than one source file, only the first occurrence of that name will appear in the merged file.

The command sequence in the following example adds the SYSUAF.DAT file from two VAXcluster computers to the master SYSUAF.DAT in the current default directory:

```
$ SET DEFAULT SYSS$SYSTEM
$ CONVERT [SYS1.SYSEXE]SYSUAF.DAT, -
_ $ [SYS2.SYSEXE]SYSUAF.DAT SYSUAF.DAT
```

The CONVERT command in this example adds the records from the files [SYS1.SYSEXE]SYSUAF.DAT and [SYS2.SYSEXE]SYSUAF.DAT to the file SYSUAF.DAT on the local computer.

After you run CONVERT, you have a master SYSUAF.DAT that contains records from the other SYSUAF.DAT files.

5. Use AUTHORIZE to modify the accounts in the master SYSUAF.DAT according to the changes you marked on the initial listings of the SYSUAF.DAT files from each computer.
6. Place the master SYSUAF.DAT file in SYS\$COMMON:[SYSEXE].
7. Remove all node-specific SYSUAF.DAT files.

### Merging RIGHTSLIST.DAT Files

If you need to merge RIGHTSLIST.DAT files, you can use a command sequence like the following:

```
$ SET DEFAULT SYSS$SYSTEM
$ ANALYZE/RMS/FDL RIGHTSLIST.DAT
$ CONVERT/STATISTICS/FDL=RIGHTSLIST -
_ $ [SYS1.SYSEXE]RIGHTSLIST.DAT, [SYS2.SYSEXE]RIGHTSLIST.DAT -
_ $ RIGHTSLIST.DAT
```

The commands in this example add the RIGHTSLIST.DAT files from two VAXcluster computers to the master RIGHTSLIST.DAT file in the current default directory. For detailed information of creating and maintaining RIGHTSLIST.DAT files, refer to the *Guide to VMS System Security*.

---

# Cluster Troubleshooting

This appendix contains information to help you perform troubleshooting operations for the following:

- Failures of computers to boot or to join the cluster
- Cluster hangs
- CLUEXIT bugchecks
- VAXport device problems

## C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

Before you initiate diagnostic procedures, be sure to verify that these conditions are met:

- All cluster hardware components are correctly connected and checked for proper operation.
- VAXcluster computers and mass storage devices are configured according to requirements specified in the VAXcluster *Software Product Description* (SPD).

When attempting to add a new or recently repaired CI-connected computer to the cluster, you must verify that the CI cables are correctly connected, as described in Section C.4.2.2.

When attempting to add a satellite to a local area or mixed-interconnect cluster, you must verify that the LAN is configured according to requirements specified in the VAXcluster SPD, and that the machine's memory resources and LAN adapters meet the requirements specified in that document. You must also verify that you have correctly configured and started the DECnet-VAX network, following the procedures described in Section 4.3.

If after performing preliminary checks and taking appropriate corrective action, you find that a computer still fails to boot or to join the cluster, you can follow the procedures in Sections C.1.2 through C.1.4 to attempt recovery.

### C.1.1 Events for Computers Booting and Joining the Cluster

To perform diagnostic and recovery procedures effectively, you must understand the events that occur when a computer boots and attempts to join the cluster. This section outlines those events and shows typical messages displayed at the console.

Note that events vary, depending on whether a computer is the first to boot in a new cluster or whether it is booting in an active cluster. Note further that some events (such as loading the cluster security database) occur only in local area and mixed-interconnect clusters.

# Cluster Troubleshooting

## C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

The normal sequence of events is as follows:

1. The computer boots. If the computer is a satellite, a message like the following shows the name and LAN address of the MOP server that has downline loaded the satellite. At this point, the satellite has completed communication with the MOP server and further communication continues with the system disk server, using VAXcluster communications.

```
%VAXcluster-I-SYSLOAD, system loaded from Node X... (XX-XX-XX-XX-XX-XX)
```

For any booting computer, the VMS "banner message" is displayed in the following format:

```
VAX/VMS Version n.n DD-MMM-YYYY hh:mm:ss
```

2. The computer attempts to form or join the cluster, and the following message appears:

```
waiting to form or join a VAXcluster system
```

If the computer is a member of a local area or mixed-interconnect cluster, the cluster security database is loaded. Optionally, the MSCP server and TMSCP server may be loaded:

```
%VAXcluster-I-LOADSECDB, loading the cluster security database
%MSCPLOAD-I-LOADMSCP, loading the MSCP disk server
%TMSCPLOAD-I-LOADTMSCP, loading the TMSCP tape server
```

3. If the computer discovers a cluster, the computer attempts to join. If a cluster is found, the connection manager displays one or more messages in the following format:

```
%CNXMAN, Sending VAXcluster membership request to system X...
```

Otherwise, the connection manager forms the cluster when it has enough votes to establish quorum (that is, when enough voting computers have booted).

4. As the booting computer joins the cluster, the connection manager displays a message in the following format:

```
%CNXMAN, now a VAXcluster member -- system X...
```

Note that if quorum is lost while the computer is booting, or if a computer is unable to join the cluster within 2 minutes of booting, the connection manager displays messages like the following:

```
%CNXMAN, Discovered system X...
%CNXMAN, Deleting CSB for system X...
%CNXMAN, Established "connection" to quorum disk
%CNXMAN, Have connection to system X...
%CNXMAN, Have "connection" to quorum disk
```

The last two messages show any connections that have already been formed.

If the cluster includes a quorum disk, you may also see messages like the following:

```
%CNXMAN, Using remote access method for quorum disk
%CNXMAN, Using local access method for quorum disk
```

The first message indicates that the connection manager is unable to access the quorum disk directly, either because the disk is unavailable or because it is accessed through the MSCP server. Another computer in the cluster that can access the disk directly must verify that a reliable connection to the disk exists.

## Cluster Troubleshooting

### C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

The second message indicates that the connection manager can access the quorum disk directly and can supply information about the status of the disk to computers that cannot access the disk directly.

Note that the connection manager may not see the quorum disk initially, because the disk may not yet be configured. In that case, the connection manager first uses remote access, then switches to local access.

5. Once the computer has joined the cluster, normal startup procedures execute. One of the first functions is to start the OPCOM process:

```
%%%%%%%%%% OPCOM 15-APR-1990 16:33:55.33 %%%%%%%%%%%
Logfile has been initialized by operator_X...$OFA0:
Logfile is SYS$SYSROOT:[SYSMGR]OPERATOR.LOG;17

%%%%%%%%%% OPCOM 15-APR-1990 16:33:56.43 %%%%%%%%%%%
16:32:32.93 Node X... (csid 0002000E) is now a VAXcluster member
```

When other computers join the cluster, OPCOM displays messages like the following:

```
%%%%%%%%%% OPCOM 15-APR-1990 16:34:25.23 %%%%%%%%%%% (from node X... at 16:34:25.23)
16:34:24.42 Node X... (csid 000100F3) received VAXcluster membership request from node X...
```

As startup procedures continue, various messages report startup events.

---

#### Note

---

For troubleshooting purposes, you may want to include in your site-specific startup procedures messages announcing each phase of the startup process—for example, mounting disks or starting queues.

---

### C.1.2 CI-Connected Computer Fails to Boot

If a CI-connected computer fails to boot, perform the following checks:

- Verify that the computer's SCSNODE and SYSSYSTEMID parameters are unique in the cluster. If they are not, you must either alter *both* values or reboot all other computers.
- Verify that you are using the correct bootstrap command file. This file must specify the internal bus computer number (if applicable), the HSC node number, and the HSC disk from which the computer is to boot. Refer to your processor-specific installation and operations guide for information on setting values in default bootstrap command procedures.
- Verify that the SYSGEN parameter PAMAXPORT is set to a value greater than or equal to the largest CI port number.
- Verify that the HSC subsystem is ONLINE. The ONLINE switch on the HSC operator control panel should be pressed in.
- Verify that the disk is available. The correct port switches on the disk's operator control panel should be pressed in.
- Verify that the computer has access to the HSC subsystem. The SHOW HOSTS command of the HSC SETSHO Utility displays status for all VAX computers (hosts) in the cluster. (For complete information on the SETSHO Utility, consult the HSC hardware documentation.) If the computer in question appears in the display as DISABLED, use the SETSHO Utility to set the computer to the ENABLED state.

## Cluster Troubleshooting

### C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

- Verify that the HSC subsystem allows access to the boot disk. Invoke the SETSHO Utility to ensure that the boot disk is available to the HSC subsystem. The utility's SHOW DISKS command displays the current state of all disks visible to the HSC subsystem and displays all disks in the no-host-access table. If the boot disk appears in the no-host-access table, use the SETSHO Utility to set the boot disk to host-access. If the boot disk is AVAILABLE or MOUNTED and host-access ENABLED, but does not appear in the no-host-access table, contact your Customer Services representative and explain both the problem and the steps you have taken.

#### C.1.3 Satellite Fails to Boot

To boot successfully, a satellite must communicate with a MOP server over the LAN. You can use DECnet event logging to verify this communication. Proceed as follows:

1. Log in as system manager on the MOP server.
2. If event logging for management layer events is not already enabled, enter the following NCP commands to enable it:

```
NCP> SET LOGGING MONITOR EVENT 0.*  
NCP> SET LOGGING MONITOR STATE ON
```

3. Enter the following DCL command:

```
$ REPLY/ENABLE=NETWORK
```

This command enables the terminal to receive DECnet messages reporting downline load events.

4. Boot the satellite. If the satellite and the MOP server can communicate and all boot parameters are correctly set, messages like the following are displayed at the MOP server's terminal:

```
DECnet event 0.3, automatic line service  
From node 2.4 (URANUS), 15-APR-1990 09:42:15.12  
Circuit QNA-0, Load, Requested, Node = 2.42 (OBERON)  
File = SYSSYSDEVICE:<SYS10.>, Operating system  
Ethernet address = 08-00-2B-07-AC-03
```

```
DECnet event 0.3, automatic line service  
From node 2.4 (URANUS), 15-APR-1990 09:42:16.76  
Circuit QNA-0, Load, Successful, Node = 2.42 (ARIEL)  
File = SYSSYSDEVICE:<SYS11.>, Operating system  
Ethernet address = 08-00-2B-07-AC-13
```

If the satellite cannot communicate with the MOP server, no message for that satellite appears. There may be a problem with a LAN cable connection or adapter service.

If the satellite's data in the DECnet database is incorrectly specified (for example, if the hardware address is incorrect), a message like the following displays the correct address and indicates that a load was requested:

```
DECnet event 0.7, aborted service request  
From node 2.4 (URANUS), 15-APR-1990 09:42:09.67  
Circuit QNA-0, Line open error, Ethernet address = 08-00-2B-03-29-99
```

Note the absence of the node name, node address, and system root.

If a satellite fails to boot, perform the following checks:

- Verify that the boot device is available. This check is particularly important for local area and mixed-interconnect clusters in which satellites boot from multiple system disks.



## C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

- Verify that the satellite's SCSSNODE and SCSSYSTEMID values and its DECnet node name and address are unique in the cluster.
- Verify that the DECnet-VAX network is up and running.
- Verify that circuit service is enabled for the MOP server's LAN adapter device. Invoke the NCP Utility and enter an NCP command in the following format, where *circuit-id* is the name of the LAN adapter circuit that the boot server uses to service downline load requests from satellites:

```
NCP> SHOW CIRCUIT circuit-id
```

If service is not enabled, you can enter NCP commands like the following to enable it:

```
NCP> SET CIRCUIT circuit-id STATE OFF
NCP> DEFINE CIRCUIT circuit-id SERVICE ENABLED
NCP> SET CIRCUIT circuit-id SERVICE ENABLED STATE ON
```

The DEFINE command updates the permanent database and ensures that service is enabled the next time you start the network. Note that DECnet traffic will be interrupted while the circuit is off.

- Verify that you have specified the correct LAN hardware address for the satellite. Proceed as follows:

1. Enter an NCP command in the following format on the MOP server, specifying the satellite's node name:

```
NCP> SHOW NODE X... CHARACTERISTICS
```

The system displays data like the following:

```
Node Volatile Characteristics as of 15-APR-1990 13:15:28
Remote node =      2.41 (ARIEL)
Hardware address      = 08-00-2B-03-27-95
Tertiary loader      = SYSS$SYSTEM:TERTIARY VMB.EXE
Load Assist Agent    = SYSS$SHARE:NISCS_LAA.EXE
Load Assist Parameter = DISK$VAXVMSRL5:<SYS12.>
```

2. At the satellite's console prompt (>>>), enter the commands shown in Table 7-2 to display the satellite's current LAN hardware address.
  3. Compare the hardware address values displayed by NCP and at the satellite's console. The values should be identical and should also match the value shown in the file SYS\$MANAGER:NETNODE\_UPDATE.COM. If the values do not match, you must make appropriate adjustments. For example, if you have recently replaced the satellite's LAN adapter device, you must execute CLUSTER\_CONFIG's CHANGE function to update the network database and NETNODE\_UPDATE.COM on the appropriate MOP server.
- Verify that the satellite's load assist parameter specifies the correct device and root directory name and that the satellite's root is unique in the cluster. If changes are needed, you can use CLUSTER\_CONFIG.COM to remove the satellite and then add it again with correct values.

## Cluster Troubleshooting

### C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

#### C.1.4 Computer Fails to Join the Cluster

If a computer boots but fails to join the cluster, proceed as follows:

- Verify that VAXcluster software has been loaded. Look for connection manager (%CNXMAN) messages like those shown in Section C.1.1. If no such messages are displayed, it is likely that VAXcluster software was not loaded at boot time. Reboot the computer in conversational mode. At the SYSBOOT> prompt, set the VAXCLUSTER parameter to 2. (In local area or mixed-interconnect clusters, you must also set NISCS\_LOAD\_PEA0 to 1.) Note that these parameters should also be set in the computer's MODPARAMS.DAT file. For more information on booting a computer in conversational mode, consult your processor-specific installation and operations guide.

In local area and mixed-interconnect clusters, verify that the cluster security database file (SYS\$COMMON:CLUSTER\_AUTHORIZE.DAT) exists and that you have specified the correct group number for this cluster. (See Section 7.7.8.1.)

- Verify that the computer has booted from the correct disk and system root. If %CNXMAN messages are displayed, and if after the conversational reboot the computer still does not join the cluster, check the console output on all active computers and look for messages indicating that one or more computers found a remote computer that conflicted with a known or local computer. Such messages suggest that two computers have booted from the same system root.

Review the boot command files for all CI-connected computers and ensure that all are booting from the correct disks and from unique system roots. If you find it necessary to modify the computer's bootstrap command procedure (console media), you may be able to do so on another processor that is already running in the cluster. Replace the running processor's console media with the media to be modified, and use the Exchange Utility and a text editor to make the required changes. Consult the appropriate processor-specific installation and operations guide for information on examining and editing boot command files.

- Verify that the computer's SCSNODE and SCSSYSTEMID parameters are unique in the cluster. To be eligible to join a cluster, a computer must have unique SCSNODE and SYSSYSTEMID parameter values. Check that the current values do not duplicate any values set for existing VAXcluster computers. Note that if you discover that one or the other value is not unique, you must alter *both* values or reboot all other computers. To check or modify values, you can perform a conversational bootstrap operation. However, for reliable future bootstrap operations, you must specify appropriate values for these parameters in the computer's MODPARAMS.DAT file.

#### C.1.5 Startup Procedures Fail to Complete

If a computer boots and joins the cluster but appears to hang before startup procedures complete—that is, before you are able to log in to the system, be sure that you have allowed sufficient time for the startup procedures to execute.

If the startup procedures fail to complete after a period that is normal for your site, try to access the procedures from another VAXcluster computer and make appropriate adjustments. For example, verify that all required devices are configured and available.

### C.1 Diagnosing Failures of Computers to Boot or Join the Cluster

One potential cause of such a failure is the lack of some system resource such as NPAGEDYN or page file space. If you suspect that the value for the NPAGEDYN parameter is set too low, you can perform a conversational bootstrap operation to increase it. Use SYSBOOT to check the current value, and then double the value. If this procedure is unsuccessful, double the value once more.

If you suspect a shortage of page file space, and if another VAXcluster computer is available, you can log in on that computer and use the System Generation Utility (SYSGEN) to provide adequate page file space for the problem computer. (Note that insufficient page file space on the booting computer may cause other computers to hang.) If the computer still cannot complete the startup procedures, contact your Customer Services representative.

#### C.1.6 Diagnosing LAN Component Failures

The local area VAXcluster Network Failure Analysis program described in Section E.2 can be used to trouble shoot LAN component failures. This analysis program requires an accurate description of the portion of the LAN used for VAXcluster traffic.

### C.2 Diagnosing Cluster Hangs

Conditions like the following can cause a VAXcluster computer to suspend process or system activity (that is, to hang):

- Cluster quorum is lost.
- A shared cluster resource is inaccessible.

Sections C.2.1 and C.2.2 discuss these conditions.

#### C.2.1 Cluster Quorum Is Lost

The VAXcluster quorum scheme coordinates activity among VAXcluster computers and ensures the integrity of shared cluster resources. (The quorum scheme is described fully in Section 3.1.1.) Quorum is checked after any change to the cluster configuration—for example, when a voting computer leaves or joins the cluster. If quorum is lost, process creation and I/O activity on all computers in the cluster are blocked.

Information about the loss of quorum and clusterwide events that cause loss of quorum are sent to the OPCOM process, which broadcasts messages to designated operator terminals. The information is also broadcast to each computer's operator console (OPA0), unless broadcast activity is explicitly disabled on that terminal. Because, however, quorum may be lost before OPCOM has been able to inform the operator terminals, the messages sent to OPA0 are the most reliable source of information about events that may cause loss of quorum.

If quorum is lost, you can follow instructions in Section 7.7.5 to recover.

#### C.2.2 A Shared Cluster Resource Is Inaccessible

Access to shared cluster resources is coordinated by the distributed lock manager. If a particular process is granted a lock on a resource (for example, a shared data file), other processes in the cluster that request incompatible locks on that resource must wait until the original lock is released. If the original process retains its lock for an extended period, other processes waiting for the lock to be released may appear to hang.

## Cluster Troubleshooting

### C.2 Diagnosing Cluster Hangs

Occasionally a system activity must acquire a restrictive lock on a resource for an extended period. For example, to perform a volume rebuild, system software takes out an exclusive lock on the volume being rebuilt. While this lock is held, no processes can allocate space on the disk volume. If they attempt to do so, they may appear to hang.

Access to files that contain data necessary for the operation of the system itself is coordinated by the distributed lock manager. For this reason, a process that acquires a lock on one of these resources and is then unable to proceed may cause the cluster to appear to hang.

For example, this condition may occur if a process locks a portion of the system authorization file (SYS\$SYSTEM:SYSUAF.DAT) for write access. Any activity that requires access to that portion of the file, such as logging into an account with the same or similar user name or sending mail to that username, will be blocked until the original lock is released. Normally this lock would be released quickly, and users would not notice the locking operation.

However, if the process holding the lock is unable to proceed, other processes could enter a wait state. Because the authorization file is used during login and for most process creation operations (for example, batch and network jobs), blocked processes could rapidly accumulate in the cluster. Because the distributed lock manager is functioning normally under these conditions, users are not notified by broadcast messages or other means that a problem has occurred.

### C.3 Diagnosing CLUEXIT Bugchecks

The VMS operating system performs **bugcheck** operations only when it detects conditions that could compromise normal system activity or endanger data integrity. A **CLUEXIT bugcheck** is a type of bugcheck initiated by the connection manager, the VAXcluster software component that manages the interaction of cooperating VAXcluster computers. Most such bugchecks are triggered by conditions resulting from hardware failures (particularly failures in communications paths), configuration errors, or system management errors.

The conditions that most commonly result in CLUEXIT bugchecks are as follows:

- The cluster connection between two computers is broken for longer than **RECNXINTERVAL** seconds. Thereafter, the connection is declared irrevocably broken. If the connection is later reestablished, either or both of the computers shut down with a CLUEXIT bugcheck.

This condition can occur upon power failure recovery with battery backup, after the repair of an SCS communication link, or after the computer was halted for a period longer than **RECNXINTERVAL** seconds, and was restarted with a **CONTINUE** command entered at the operator console. You must determine the cause of the interrupted connection and correct the problem. For example, if recovery from a power failure is longer than **RECNXINTERVAL** seconds, you may want to increase the value of the **RECNXINTERVAL** parameter on all computers.

- Cluster partitioning occurs. A member of a cluster discovers or establishes connection to a member of another cluster, or a foreign cluster is detected in the quorum file. In this case, you must review the setting of **EXPECTED\_VOTES** on all computers.

- The value specified for the SYSGEN parameter SCSMAXMSG on a computer is too small. Verify that the value of SCSMAXMSG on all VAXcluster computers is set to a value that is at the least the default value.

## **C.4 Diagnosing VAXport Device Problems**

The following sections present information on the CI and LAN VAXport devices. Information is also provided on entries in the system error log and on corrective actions to take when errors occur. Topics include the following:

- VAXport communication mechanisms
- Port failures
- VAXcluster error log entries
- OPA0 error messages

### **C.4.1 VAXport Communication Mechanisms**

This section describes CI and LAN port communication mechanisms and System Communications Services (SCS) connections.

#### **Port Polling**

Shortly after a CI-connected computer boots, the CI port driver (PADRIVER) begins configuration polling to discover other active ports on the CI. Normally the poller runs every 5 seconds (the default value of the SYSGEN parameter PAPOLLINTERVAL). In the first polling pass, all addresses are probed over cable path A; on the second pass all addresses are probed over path B; on the third pass path A is probed again, and so on.

The poller probes by sending request ID (REQID) packets to all possible port numbers, including itself. Active ports receiving the REQIDs return ID packets (IDREC) to the port issuing the REQID. A port may respond to a REQID even if the computer attached to the port is not running.

In any CI-only, local area, or mixed-interconnect cluster, the port drivers perform a start handshake when a pair of ports and port drivers has successfully exchanged ID packets. The port drivers exchange datagrams containing information about the computers, such as the type of computer and the operating system version. If this exchange is successful, each computer declares a virtual circuit open. An open virtual circuit is prerequisite to all other activity.

#### **LAN Communications**

In local area and mixed-interconnect clusters, a multicast scheme is used to locate computers on the LAN. Every 3 seconds the port emulator driver (PEDRIVER) sends a HELLO message via each LAN adapter to a cluster-specific multicast address that is derived from the cluster group number. The driver also enables the reception of these messages from other computers. When the driver receives a HELLO message from a computer with which it does not currently share an open virtual circuit, it attempts to create a circuit. HELLO messages received from a computer with a currently open virtual circuit indicate that the remote computer is operational.

A standard three-message exchange handshake is used to create a virtual circuit. The handshake messages contain information about the transmitting computer and its record of the cluster password. These parameters are verified at the receiving computer, which continues the handshake only if its verification is

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

successful. Thus, each computer authenticates the other. After the final message, the virtual circuit is opened for use by both computers.

#### System Communications Services (SCS) Connections

System services such as the disk class driver, connection manager, and the MSCP and TMSCP servers communicate between computers with a protocol called System Communications Services (SCS). Primarily, SCS is responsible for the formation and breaking of intersystem process connections and for flow control of message traffic over those connections. SCS is implemented in the VAXport driver (for example, PADRIVER, PBDRIVER, PEDRIVER, PIDRIVER), and in a loadable piece of the VMS operating system called SCSLOA.EXE (loaded automatically during system initialization).

When a virtual circuit has been opened, a computer periodically probes a remote computer for system services that the remote computer may be offering. The SCS directory service, which makes known services that a computer is offering, is always present both on computers and HSC subsystems. As system services discover their counterparts on other computers and HSC subsystems, they establish SCS connections to each other. These connections are full duplex and are associated with a particular virtual circuit. Multiple connections are typically associated with a virtual circuit.

#### C.4.2 Port Failures

Taken together, SCS, the VAXport drivers, and the port itself support a hierarchy of communications paths. Working up from the most fundamental level, these are as follows:

- The physical wires. The Ethernet is a single coaxial cable. FDDI has a pair of fiber optic cables for transmit and receive. The CI has two pairs of transmit and receive cables (path A transmit and receive and path B transmit and receive). For the CI, VMS software normally sends traffic in automatic path select mode. The port chooses the free path or, if both are free, an arbitrary path (implemented in the cables and star coupler, and managed by the port).
- The virtual circuit (implemented partly in the CI port or LAN port emulator driver (PEDRIVER) and partly in SCS software).
- The SCS connections (implemented in system software).

Failures can occur at each communications level and in each component. Failures at one level translate into failures at other levels as follows:

- **Wires.** If the LAN fails or is disconnected, LAN traffic stops or is interrupted, depending on the nature of the failure. For the CI, either path A or B can fail while the virtual circuit remains intact. All traffic is directed over the remaining good path. When the wire is repaired, the repair is detected automatically by port polling, and normal operations resume on all ports.
- **Virtual circuit.** If no path works between a pair of ports, the virtual circuit fails and is closed. A path failure is discovered as follows:
  - For the CI, when polling fails, or when attempts are made to send normal traffic, and the port reports that neither path yielded transmit success.
  - For the LAN, when no multicast HELLO message or incoming traffic is received from another computer.

When a virtual circuit fails, every SCS connection on it fails. The software automatically reestablishes connections when the virtual circuit is reestablished. Normally, reestablishing a virtual circuit takes several seconds after the problem is corrected.

- **CI port.** If a port fails, all virtual circuits to that port fail, and all SCS connections on those virtual circuits fail. If the port is successfully reinitialized, virtual circuits and connections are reestablished automatically. Normally, port reinitialization and reestablishment of connections take several seconds.
- **LAN adapter.** If a LAN adapter device fails, attempts are made to restart it. If repeated attempts fail, all channels using that adapter are broken. A channel is a pair of LAN addresses, one local and one remote. If the last open channel for a virtual circuit fails, the virtual circuit is closed and the connections are broken.
- **SCS connection.** When the software protocols fail or, in some instances, when the software detects a hardware malfunction, a connection is terminated. Other connections are normally unaffected, as is the virtual circuit. Breaking of connections is also used under certain conditions as an error recovery mechanism—most commonly when there is insufficient nonpaged pool available on the computer.
- **Computer.** If a computer fails because of operator shutdown, bugcheck, or halt and reboot, all other computers in the cluster record the failure as failures of their virtual circuits to the port on the failed computer.

#### C.4.2.1 Verifying CI Port Functions

Before you boot in a cluster a CI-connected computer that is new, just repaired, or suspected of having a problem, you should have Customer Services verify that the computer runs correctly on its own.

To diagnose communication problems, you can invoke the Show Cluster Utility and tailor the SHOW CLUSTER report by entering the SHOW CLUSTER command `ADD CIRCUIT CABLE_ST`. This command adds a class of information about all the virtual circuits as seen from the computer on which you are running SHOW CLUSTER. Primarily, you are checking whether there is a virtual circuit in the OPEN state to the failing computer. Common causes of failure to open a virtual circuit and keep it open are the following:

- Port errors on one side or the other
- Cabling errors
- A port set off line because of software problems
- Insufficient nonpaged pool available on both sides
- Failure to set correct values for the SYSGEN parameters `SCSNODE`, `SCSSYSTEMID`, `PAMAXPORT`, `PANOPOLL`, `PASTIMOUT`, and `PAPOLLINTERVAL`.

Run SHOW CLUSTER from each active computer in the cluster to verify whether each computer's view of the failing computer is consistent with every other computer's view. If all the active computers have a consistent view of the failing computer, the problem may be in the failing computer. If, on the other hand, only one of several active computers detects that the newcomer is failing, that particular computer may be experiencing a problem.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

If no virtual circuit is open to the failing computer, check the bottom of the SHOW CLUSTER display for information on circuits to the port of the failing computer. Virtual circuits in partially open states are shown at the bottom of the display. If the circuit is shown in a state other than OPEN, communications between the local and remote ports are taking place, and the failure is probably at a higher level than in port or cable hardware. Next, check that both paths A and B are good to the failing port. The loss of one path should not prevent a computer from participating in a cluster.

#### C.4.2.2 Verifying CI Cable Connections

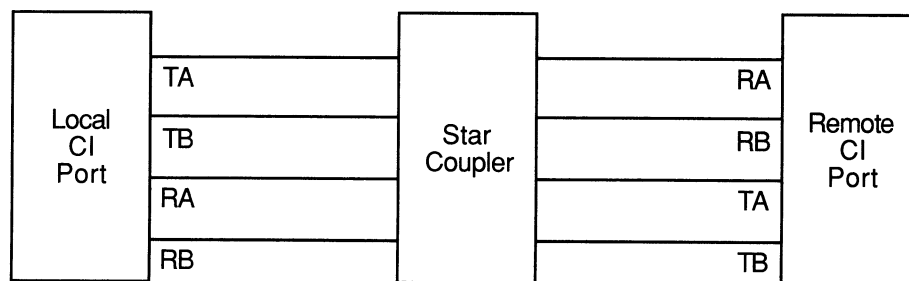
Whenever the configuration poller finds that no virtual circuits are open and that no handshake procedures are currently opening virtual circuits, the poller analyzes its environment. It does so by using the send-loopback-datagram facility of the CI port.

The send-loopback-datagram facility tests the connections between the CI port and the star coupler by routing messages across them. The messages are called loopback datagrams. (The port processes other self-directed messages without using the star coupler or external cables.)

The configuration poller makes entries in the error log whenever it detects a change in the state of a circuit. Note, however, that it is possible for two changed-to-failed-state messages to be entered in the log without an intervening changed-to-succeeded-state message. Such a series of entries means that the circuit state continues to be faulty.

The following paragraphs discuss various incorrect CI cabling configurations and the entries made in the error log when these configurations exist. Figure C-1 shows a two-computer configuration with all cables correctly connected. Figure C-2 shows a CI cluster with a pair of crossed cables.

**Figure C-1 A Correctly Connected Two-Computer CI Cluster**

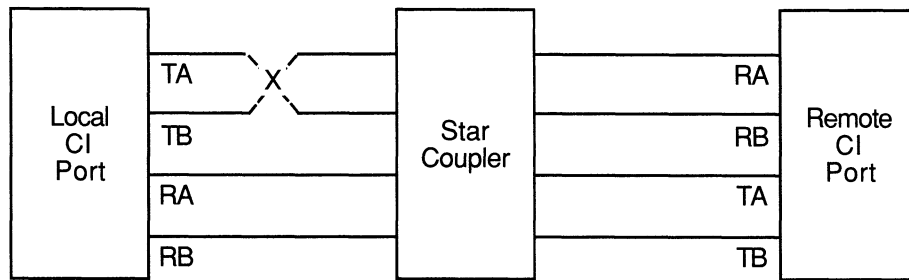


ZK-1924-GE



## Cluster Troubleshooting C.4 Diagnosing VAXport Device Problems

**Figure C-2 Crossed CI Cable Pair**



ZK-1925-GE

If a pair of transmitting cables or a pair of receiving cables is crossed, a message sent on TA is received on RB, and a message sent on TB is received on RA. This is a hardware error condition from which the port cannot recover. An entry is made in the error log to say that a single pair of crossed cables exists. The entry contains the following lines:

```
DATA CABLE(S) CHANGE OF STATE
PATH 1. LOOPBACK HAS GONE FROM GOOD TO BAD
```

If this situation exists, you can correct it by reconnecting the cables properly. The cables could be misconnected in several places. The coaxial cables that connect the port boards to the bulkhead cable connectors can be crossed, or the cables can be misconnected to the bulkhead or the star coupler.

The information in Figure C-2 can be represented more simply. Configuration 1 shows the cables positioned as in Figure C-2, but it does not show the star coupler or the computers. The letters LOC and REM indicate the pairs of transmitting (T) and receiving (R) cables on the local and remote computers, respectively.

### Configuration 1

```
T x   = R
R =   = T
LOC  REM
```

The pair of crossed cables causes loopback datagrams to fail on the local computer, but succeed on the remote computer. Crossed pairs of transmitting cables and crossed pairs of receiving cables cause the same behavior.

Note that only an odd number of crossed-cable pairs causes these problems. If an even number of cable pairs is crossed, communications succeed. An error log entry is made in some cases, however, and the contents of the entry depends on which pairs of cables are crossed.

Configuration 2 shows two-computer clusters with the combinations of two pairs of crossed-cable pairs. These crossed pairs cause the following entry to be made in the error log of the computer that has the cables crossed:

```
DATA CABLE(S) CHANGE OF STATE
CABLES HAVE GONE FROM UNCROSSED TO CROSSED
```

Loopback datagrams succeed on both computers, and communications are possible.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

#### Configuration 2

```
T x   = R           T =   x R
R x   = T           R =   x T
LOC   REM           LOC   REM
```

Configuration 3 shows the possible combinations of two pairs of crossed cables that cause loopback datagrams to fail on both computers in the cluster. Communications can still take place between the computers. An entry stating that cables are crossed is made in the error log of each computer.

#### Configuration 3

```
T x   = R           T =   x R
R =   x T           R x   = T
LOC   REM           LOC   REM
```

Configuration 4 shows the possible combinations of two pairs of crossed cables that cause loopback datagrams to fail on both computers in the cluster, but allow communications. No entry stating that cables are crossed is made in the error log of either computer.

#### Configuration 4

```
T x   x R           T =   = R
R =   = T           R x   x T
LOC   REM           LOC   REM
```

Configuration 5 shows the possible combinations of three pairs of crossed cables. In each case, loopback datagrams fail on the computer that has only one crossed pair of cables. Loopback datagrams succeed on the computer with both pairs crossed. No communications are possible.

#### Configuration 5

```
T x   x R           T x   = R           T =   x R           T x   x R
R x   = T           R x   x T           R x   x T           R =   x T
LOC   REM           LOC   REM           LOC   REM           LOC   REM
```

If all four cable pairs between two computers are crossed, communications succeed, loopback datagrams succeed, and no crossed-cable message entries are made in the error log. Such a condition might be detected by noting error log entries made by a third computer in the cluster, but only if the third computer has one of the crossed-cable cases described.

#### C.4.2.3 Repairing CI Cables

This section describes some ways in which Customer Services can make repairs on a running computer. This information is provided to aid system managers in scheduling repairs.

For cluster software to survive cable-checking activities or cable-replacement activities, you must be sure that either path A or path B is intact at all times between each port and between every other port in the cluster.

You can, for example, remove path A and path B in turn from a particular port to the star coupler. To make sure that the configuration poller finds a path that was previously faulty but is now operational, follow these steps:

1. Remove path B.
2. After the poller has discovered that path B is faulty, reconnect path B.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

3. Wait two poller intervals, and then enter the DCL command `SHOW CLUSTER` to make sure that the poller has reestablished path B. Or, enter the DCL command `SHOW CLUSTER/CONTINUOUS` followed by the `SHOW CLUSTER` command `ADD CIRCUITS, CABLE_ST`. Wait until `SHOW CLUSTER` tells you that path B has been reestablished.
4. Remove path A.
5. After the poller has discovered that path A is faulty, reconnect path A.
6. Wait two poller intervals to make sure that the poller has reestablished path A.

If both paths are lost at the same time, the virtual circuits are lost between the port with the broken cables and all other ports in the cluster. This condition will in turn result in loss of SCS connections over the broken virtual circuits. However, recovery from this situation is automatic after an interruption in service on the affected computer. The length of the interruption varies, but it is usually approximately two poller intervals (or 10 seconds) at the default `SYSGEN` parameter settings.

#### C.4.2.4 Verifying LAN Connections

The Local Area VAXcluster Network Failure Analysis program described in Section E.1.3 uses the `HELLO` messages to verify continuously the network paths (channels) used by `PEDRIVER`. This verification process, combined with physical description of the network, can isolate failing network components. Failing channels are grouped together and mapped onto the physical network description. The failure analysis calls out the common components related to the channel failures.

#### C.4.3 Analyzing Error Log Entries for VAXport Devices

To anticipate and avoid potential problems, you must monitor events recorded in the error log. From the total error count, displayed by a DCL command in the format `SHOW DEVICE device-name`, you can determine whether errors are increasing. If so, you should examine the error log.

The DCL command `ANALYZE/ERROR_LOG` invokes the Error Log Utility to report the contents of an error log file. (For more information on the Error Log Utility, see the *VMS Error Log Utility Manual*.)

Note that some error log entries are informational only, and require no action. For example, if you shut down a computer in the cluster, all other active computers that have open virtual circuits between themselves and the computer that has been shut down make entries in their error logs. Such computers record up to three errors for the event:

1. Path A received no response.
2. Path B received no response.
3. The virtual circuit is being closed.

These messages are normal and reflect the change of state in the circuits to the computer that has been shut down.

On the other hand, some error log entries are made for problems that degrade operation, or for nonfatal hardware problems. The VMS operating system might continue to run satisfactorily under these conditions. The purpose of detecting these problems early is to prevent nonfatal problems (such as loss of a single CI path) from becoming serious problems (such as loss of both paths).

# Cluster Troubleshooting

## C.4 Diagnosing VAXport Device Problems

### C.4.3.1 Error Log Entry Formats

Errors and other events on the CI or LAN cause VAXport drivers to enter information in the system error log. The two formats used for error log entries are the **device-attention** format and the **logged-message** format. Sections C.4.3.2 and C.4.3.3 describe those formats.

Device-attention entries for the CI record events that, in general, are indicated by the setting of a bit in a hardware register. For the LAN, device-attention entries typically record errors on an LAN adapter device. Logged-message entries record the receipt of a message packet that contains erroneous data or that signals an error condition.

### C.4.3.2 Device-Attention Entries

Example C-1 shows device-attention entries for the CI. The left column gives the name of a device register or a memory location. The center column gives the value contained in that register or location, and the right column gives an interpretation of that value.

#### Example C-1 CI Device-Attention Entry

```

***** ENTRY      83. ***** ①
ERROR SEQUENCE 10.          LOGGED ON:      SID 0150400A
DATE/TIME 15-APR-1990 11:45:27.61          SYS_TYPE 01010000 ②
DEVICE ATTENTION  KA780          ③
                SCS NODE: MARS

CI SUB-SYSTEM, MARS$PAA0: - PORT POWER DOWN ④
    CNFGR          00800038          ADAPTER IS CI
                                ADAPTER POWER-DOWN
    PMCSR          000000CE          MAINTENANCE TIMER DISABLE
                                MAINTENANCE INTERRUPT ENABLE
                                MAINTENANCE INTERRUPT FLAG
                                PROGRAMMABLE STARTING ADDRESS
                                UNINITIALIZED STATE
    PSR            80000001          RESPONSE QUEUE AVAILABLE
                                MAINTENANCE ERROR
    PFAR           00000000
    PESR           00000000
    PPR            03F80001
    UCB$B_ERTCNT   32          ⑤
                                50. RETRIES REMAINING
    UCB$B_ERTMAX   32          ⑥
                                50. RETRIES ALLOWABLE
    UCB$L_CHAR     0C450000
                                SHAREABLE
                                AVAILABLE
                                ERROR LOGGING
                                CAPABLE OF INPUT
                                CAPABLE OF OUTPUT
    UCB$W_STS      0010
                                ONLINE
    UCB$W_ERRCNT   000B          ⑦
                                11. ERRORS THIS UNIT

```

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

- ❶ The first two lines are the entry heading. These lines contain the number of the entry in this error log file, the sequence number of this error, and the identification number (SID) of this computer. Each entry in the log file contains such a heading.
- ❷ The next line contains the date and time, and the computer type.
- ❸ The next two lines contain the entry type, the processor type (KA780), and the computer's SCS node name.
- ❹ The line `CI SUB-SYSTEM, MARS$PAA0: - PORT POWER DOWN` contains the name of the subsystem and the device that caused the entry, and the reason for the entry. The CI subsystem's device PAA0 on MARS was powered down.

The next 15 lines contain the names of hardware registers in the port, their contents, and interpretations of those contents. See the appropriate CI hardware manual for a description of all the CI port registers.

The CI port can recover from many errors, but not all. When an error occurs from which the CI cannot recover, the port notifies the port driver. The port driver logs the error and attempts to reinitialize the port. If the port fails after 50 such initialization attempts, the driver takes it off line, unless the system disk is connected to the failing port or this computer is supposed to be a cluster member. If the CI port is required for system disk access or cluster participation and all 50 reinitialization attempts have been used, then the computer bugchecks with a CIPORT-type bugcheck. Once a CI port is off line, you can put the port back on line only by rebooting the computer.

- ❺ The `UCB$B_ERTCNT` field contains the number of reinitializations that the port driver can still attempt. The difference between this value and `UCB$B_ERTMAX` is the number of reinitializations already attempted.
- ❻ The `UCB$B_ERTMAX` field contains the maximum number of times the port can be reinitialized by the port driver.
- ❼ The `UCB$W_ERRCNT` field contains the total number of errors that have occurred on this port since it was booted. This total includes both errors that caused reinitialization of the port and errors that did not.

Example C-2 shows device-attention entries for the LAN. The left column gives the name of a device register or a memory location. The center column gives the value contained in that register or location, and the right column gives an interpretation of that value.

# Cluster Troubleshooting

## C.4 Diagnosing VAXport Device Problems

### Example C-2 LAN Device-Attention Entry

```

***** ENTRY      80. *****
ERROR SEQUENCE 26.          LOGGED ON:      SID 08000000
DATE/TIME 15-APR-1990 11:30:53.07         SYS_TYPE 01010000
DEVICE ATTENTION KA630
                                SCS NODE: PHOBOS
NI-SCS SUB-SYSTEM, PHOBOS$PEA0:
FATAL ERROR DETECTED BY DATALINK

STATUS1          0000002C
STATUS2          00000000
DATALINK UNIT    0001
DATALINK NAME    41515803
                                00000000
                                00000000
                                00000000
                                DATALINK NAME = XQA1:
REMOTE NODE      00000000
                                00000000
                                00000000
                                00000000
REMOTE ADDR      00000000
                                0000
LOCAL ADDR       000400AA
                                4C07
                                ETHERNET ADDR = AA-00-04-00-07-4C
ERROR CNT        0001
UCB$W_ERRCNT    0007
1. ERROR OCCURENCES THIS ENTRY
7. ERRORS THIS UNIT

```

- ① The first two lines are the entry heading. These lines contain the number of the entry in this error log file, the sequence number of this error, and the identification number (SID) of this computer. Each entry in the log file contains such a heading.
- ② The next line contains the date and time, and the computer type.
- ③ The next two lines contain the entry type, the processor type (KA630), and the computer's SCS node name.
- ④ This line shows the name of the subsystem and component that caused the entry.
- ⑤ This line shows the reason for the entry. The LAN driver has shut down the data link because of a fatal error. The data link will be restarted automatically, if possible.
- ⑥ STATUS 1 shows the I/O completion status returned by the LAN driver. STATUS 2 is the VCI event code delivered to PEDRIVER by the LAN driver. If a message transmit was involved, the status applies to that transmit.
- ⑦ DATALINK UNIT shows the unit number of the LAN device on which the error occurred.
- ⑧ DATALINK NAME is the name of the LAN device on which the error occurred.
- ⑨ REMOTE NODE is the name of the remote node to which the packet was being sent. If it is zero, no remote node was available or no packet was associated with the error.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

- ⑩ REMOTE ADDR is the LAN address of the remote node to which the packet was being sent. If zero, no packet was associated with the error.
- ⑪ LOCAL ADDR is the LAN address of the local node.
- ⑫ ERROR CNT. Because some errors can occur at extremely high rates, some error log entries represent more than one occurrence of an error. This field indicates how many. The errors counted occurred in the 3 seconds preceding the time stamp on the entry.

#### C.4.3.3 Logged-Message Entries

Logged-message entries are made when the CI or LAN port receives a response that contains either data that the port driver cannot interpret or an error code in the status field of the response.

Example C-3 shows a CI logged-message entry with an error code in the status field PPD\$B\_STATUS.

#### Example C-3 CI Logged-Message Entry

```

***** ENTRY      3. ***** ①
ERROR SEQUENCE 3.          LOGGED ON SID 01188542
ERL$LOGMESSAGE, 15-APR-1990 13:40:25.13 ②
      KA780 REV #3. SERIAL #1346.    MFG PLANT 15. ③
CI SUB-SYSTEM, MARS$PAA0: ④
DATA CABLE(S) STATE CHANGE - PATH #0. WENT FROM GOOD TO BAD ⑤
      LOCAL STATION ADDRESS, 000000000002 (HEX) ⑥
      LOCAL SYSTEM ID, 000000000001 (HEX) ⑦
      REMOTE STATION ADDRESS, 000000000004 (HEX) ⑧
      REMOTE SYSTEM ID, 0000000000A9 (HEX) ⑨
UCB$B_ERTCNT      32 ⑩
                                50. RETRIES REMAINING
UCB$B_ERTMAX      32
                                50. RETRIES ALLOWABLE
UCB$W_ERRCNT      0001
PPD$B_PORT        04 ⑪
                                REMOTE NODE #4.
PPD$B_STATUS      A5 ⑫
                                FAIL
                                PATH #0., NO RESPONSE
                                PATH #1., "ACK" OR NOT USED
                                NO PATH
PPD$B_OPC         05 ⑬
                                IDREQ
PPD$B_FLAGS       03 ⑭
                                RESPONSE QUEUE BIT
                                SELECT PATH #0.

```

(continued on next page)

# Cluster Troubleshooting

## C.4 Diagnosing VAXport Device Problems

### Example C-3 (Cont.) CI Logged-Message Entry

"CI" MESSAGE

15

```
00000000
00000000
80000004
0000FE15
4F503000
00000507
00000000
00000000
00000000
00000000
00000000
00000000
00000000
00000000
00000000
00000000
00000000
```

- ① The first two lines are the entry heading. These lines contain the number of the entry in this error log file, the sequence number of the error, and the identification number (SID) of the computer. Each entry in the log file contains a heading.
- ② The next line contains the entry type, the date, and time.
- ③ The next line contains the processor type (KA780), the hardware revision number of the computer (REV #3), the serial number of the computer (SERIAL #1346), and the plant number (15).
- ④ The line CI SUB-SYSTEM, MARS\$PAA0: contains the name of the subsystem and the device that caused the entry.
- ⑤ The next line gives the reason for the entry (one or more data cables have changed state), and a more detailed reason for the entry. Path 0, which the port used successfully before, cannot be used now.

---

#### Note

---

ANALYZE/ERROR\_LOG uses the notation path 0 and path 1; cable labels use the notation path A (=0) and path B (=1).

---

- ⑥ The local (⑥) and remote (⑧) station addresses are the port numbers (range 0–15) of the local and remote ports. The port numbers are set in hardware switches by Customer Services. The local (⑦) and remote (⑨) system IDs are the SCS system IDs set by the SYSGEN parameter SCSSYSTEMID for the local and remote systems. For HSC subsystems, the system ID is set with the HSC console.
- ⑦ The rest of the entry, which consists of the entry fields that begin with UCB\$, gives information on the contents of the unit control block (UCB) for this CI device.

The following fields, which begin with PPD\$, are fields in the message packet that the local port has received.



- ⑧ PPD\$B\_PORT contains the station address of the remote port. In a loopback datagram, however, this field contains the local station address.
- ⑨ The PPD\$B\_STATUS field contains information on the nature of the failure that occurred during the current operation. When the operation completes without error, ERF prints the word NORMAL beside this field; otherwise, ERF decodes the error information contained in PPD\$B\_STATUS. Here a NO PATH error occurred because of a lack of response on path 0, the selected path.
- ⑩ The PPD\$B\_OPC field contains the code for the operation that the port was attempting when the error occurred. The port was trying to send a request-for-ID message.
- ⑪ The PPD\$B\_FLAGS field contains bits that indicate, among other things, the path that was selected for the operation.
- ⑫ The "CI" MESSAGE is a hexadecimal listing of bytes 16 through 83 (decimal) of the response (message or datagram). Since responses are of variable length depending upon the port opcode, bytes 16 through 83 may contain either more or fewer bytes than actually belong to the message. Here the request-for-id contains no information in bytes 16 through 83.

#### C.4.3.4 Error Log Entry Descriptions

This section describes error log entries for the CI and LAN ports. Each entry shown is followed by a brief description of what the associated VAXport driver (PADRIVER, PBDRIVER, PEDRIVER) does, and the suggested action a system manager should take. In cases where Software Performance Reports (SPRs) with crash dumps are requested, it is important to capture the crash dumps as soon as possible after the error. For CI entries, note that path A and path 0 are the same path, and that path B and path 1 are the same path.

##### BIIC FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services.

##### CI PORT TIMEOUT

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device offline.

**User Action:** First, increase the SYSGEN parameter PAPOLLINTERVAL. If the problem disappears and you are not running privileged user-written software, submit an SPR. Otherwise, call Customer Services.

##### 11/750 CPU MICROCODE NOT ADEQUATE FOR PORT

**Explanation:** The VAXport driver sets the port off line with no retries attempted. In addition, if this port is needed because the computer is booted from an HSC subsystem or is participating in a cluster, the computer bugchecks with a UCODEREV code bugcheck.

**User Action:** Read the appropriate section in the current VAXcluster SPD for information on required computer microcode revisions. Call Customer Services if necessary.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

#### PORT MICROCODE REV NOT CURRENT, BUT SUPPORTED

**Explanation:** The VAXport driver detected that the microcode is not at the current level, but will continue normally. This error is logged as a warning only.

**User Action:** Contact Customer Services when it is convenient to have the microcode updated.

#### PORT MICROCODE REV NOT SUPPORTED

**Explanation:** The VAXport driver sets the port off line without attempting any retries.

**User Action:** Read the VAXcluster SPD for information on the required CI port microcode revisions. Contact Customer Services if necessary.

#### DATA CABLE(S) STATE CHANGE CABLES HAVE GONE FROM CROSSED TO UNCROSSED

**Explanation:** The VAXport driver logs this event.

**User Action:** No action needed.

#### DATA CABLE(S) STATE CHANGE CABLES HAVE GONE FROM UNCROSSED TO CROSSED

**Explanation:** The VAXport driver logs this event.

**User Action:** Check for crossed-cable pairs. See Section C.4.2.2.

#### DATA CABLE(S) STATE CHANGE PATH 0. WENT FROM BAD TO GOOD

**Explanation:** The VAXport driver logs this event.

**User Action:** No action needed.

#### DATA CABLE(S) STATE CHANGE PATH 0. WENT FROM GOOD TO BAD

**Explanation:** The VAXport driver logs this event.

**User Action:** Check path A cables to see that they are not broken or improperly connected.

#### DATA CABLE(S) STATE CHANGE PATH 0. LOOPBACK IS NOW GOOD, UNCROSSED

**Explanation:** The VAXport driver logs this event.

**User Action:** No action needed.

#### DATA CABLE(S) STATE CHANGE PATH 0. LOOPBACK WENT FROM GOOD TO BAD

**Explanation:** The VAXport driver logs this event.

**User Action:** Check for crossed-cable pairs or faulty CI hardware. See Sections C.4.2.1 and C.4.2.2.

#### DATA CABLE(S) STATE CHANGE PATH 1. WENT FROM BAD TO GOOD

**Explanation:** The VAXport driver logs this event.

**User Action:** No action needed.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

DATA CABLE(S) STATE CHANGE  
PATH 1. WENT FROM GOOD TO BAD

**Explanation:** The VAXport driver logs this event.

**User Action:** Check path B cables to see that they are not broken or improperly connected.

DATA CABLE(S) STATE CHANGE  
PATH 1. LOOPBACK IS NOW GOOD, UNCROSSED

**Explanation:** The VAXport driver logs this event.

**User Action:** No action needed.

DATA CABLE(S) STATE CHANGE  
PATH 1. LOOPBACK WENT FROM GOOD TO BAD

**Explanation:** The VAXport driver logs this event.

**User Action:** Check for crossed-cable pairs or faulty CI hardware. See Sections C.4.2.1 and C.4.2.2.

DATAGRAM FREE QUEUE INSERT FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, and 8800) contention.

DATAGRAM FREE QUEUE REMOVE FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures, or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, and 8800) contention.

FAILED TO LOCATE PORT MICRO-CODE IMAGE

**Explanation:** The VAXport driver marks device off line and makes no retries.

**User Action:** Make sure console volume contains the microcode file CI780.BIN (for the CI780, CI750, or CIBCI) or the microcode file CIBCA.BIN for the CIBCA-AA, and then reboot the computer.

HIGH PRIORITY COMMAND QUEUE INSERT FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, 8800) contention.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

#### MSCP ERROR LOGGING DATAGRAM RECEIVED

**Explanation:** On receipt of an error message from the HSC subsystem, the VAXport driver logs the error and takes no other action. It is recommended that you disable the sending of HSC informational error log datagrams with the appropriate HSC console command. Informational error log datagrams take considerable space in the error log data file.

**User Action:** They are useful to read only if they are not captured on the HSC console for some reason (for example, the HSC console ran out of paper.) This logged information is a duplicate of the messages logged on the HSC console.

#### INAPPROPRIATE SCA CONTROL MESSAGE

**Explanation:** The VAXport driver closes the port-to-port virtual circuit to the remote port.

**User Action:** Submit a Software Performance Report to Digital including the error logs and the crash dumps from the local and remote computers.

#### INSUFFICIENT NON-PAGED POOL FOR INITIALIZATION

**Explanation:** The VAXport driver marks device off line and makes no retries.

**User Action:** Reboot the computer with a larger value for NPAGEDYN or NPAGEVIR.

#### LOW PRIORITY CMD QUEUE INSERT FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, and 8800) contention.

#### MESSAGE FREE QUEUE INSERT FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, and 8800) contention.

#### MESSAGE FREE QUEUE REMOVE FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, and 8800) contention.

#### MICRO-CODE VERIFICATION ERROR

**Explanation:** The VAXport driver detected an error while reading the microcode that it just loaded into the port. The driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

#### NO PATH-BLOCK DURING VIRTUAL CIRCUIT CLOSE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Submit a Software Performance Report to Digital including the error log and a crash dump from the local computer.

#### NO TRANSITION FROM UNINITIALIZED TO DISABLED

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services.

#### PORT ERROR BIT(S) SET

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** A **maintenance timer expiration** bit may mean that the PASTIMOUT SYSGEN parameter is set too low and should be increased, especially if the local computer is running privileged user-written software. For all other bits, call Customer Services.

#### PORT HAS CLOSED VIRTUAL CIRCUIT

**Explanation:** The VAXport driver closes the virtual circuit that the local CI port opened to the remote port.

**User Action:** Check the PPD\$B\_STATUS field of the error log entry for the reason the virtual circuit was closed. This error is normal if the remote computer crashed or was shut down. For PEDRIVER, ignore the PPD\$B\_OPC field value; it is an unknown opcode.

#### PORT POWER DOWN

**Explanation:** The VAXport driver halts port operations, and then waits for power to return to the port hardware.

**User Action:** Restore power to the port hardware.

#### PORT POWER UP

**Explanation:** The VAXport driver reinitializes the port and restarts port operations.

**User Action:** No action needed.

#### RECEIVED CONNECT WITHOUT PATH-BLOCK

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Submit a Software Performance Report to Digital including the error log and a crash dump from the local computer.

#### REMOTE SYSTEM CONFLICTS WITH KNOWN SYSTEM

**Explanation:** The configuration poller discovered a remote computer with SCSYSTEMID and/or SCSNODE equal to that of another computer to which a virtual circuit is already open.

**User Action:** Shut the new computer down as soon as possible. Reboot it with a unique SCSYSTEMID and SCSNODE. Do not leave the new computer up any longer than necessary. If you are running a cluster and two computers with conflicting identity are polling when any other virtual circuit failure

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

takes place in the cluster, then computers in the cluster may crash with a CLUEXIT bugcheck.

#### RESPONSE QUEUE REMOVE FAILURE

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services. This error is caused by a failure to obtain access to an interlocked queue. Possible sources of the problem are CI hardware failures or memory, SBI (11/780), CMI (11/750), or BI (8200, 8300, and 8800) contention.

#### SCSSYSTEMID MUST BE SET TO NON-ZERO VALUE

**Explanation:** The VAXport driver sets the port off line without attempting any retries.

**User Action:** Reboot the computer with a conversational boot and set the SCSSYSTEMID to the correct value. At the same time, check that SCSNODE has been set to the correct nonblank value.

#### SOFTWARE IS CLOSING VIRTUAL CIRCUIT

**Explanation:** The VAXport driver closes the virtual circuit to the remote port.

**User Action:** Check error log entries for the cause of the virtual circuit closure. Faulty transmission or reception on both paths, for example, causes this error and may be detected from the one or two previous error log entries noting bad paths to this remote computer.

#### SOFTWARE SHUTTING DOWN PORT

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Check other error log entries for the possible cause of the port reinitialization failure.

#### UNEXPECTED INTERRUPT

**Explanation:** The VAXport driver attempts to reinitialize the port; after 50 failed attempts, it marks the device off line.

**User Action:** Call Customer Services.

#### UNRECOGNIZED SCA PACKET

**Explanation:** The VAXport driver closes the virtual circuit to the remote port. If the virtual circuit is already closed, the port driver inhibits datagram reception from the remote port.

**User Action:** Submit a Software Performance Report to Digital, including the error log file that contains this entry and the crash dumps from both the local and remote computers.

#### VIRTUAL CIRCUIT TIMEOUT

**Explanation:** The VAXport driver closes the virtual circuit that the local CI port opened to the remote port. This closure occurs if the remote computer is running CI microcode version 7 or later, and if the remote computer has failed to respond to any messages sent by the local computer.

**User Action:** This error is normal if the remote computer has halted, crashed, or was shut down. This error may mean that the local computer's TIMVCFAIL SYSGEN parameter is set too low, especially if the remote computer is running privileged user-written software.

#### INSUFFICIENT NON-PAGED POOL FOR VIRTUAL CIRCUITS

**Explanation:** The VAXport driver closes virtual circuits because of insufficient pool.

**User Action:** Enter the DCL command SHOW MEMORY to determine pool requirements, and then adjust the appropriate SYSGEN requirements.

---

**Note**

---

The following descriptions apply only for LAN devices.

---

#### FATAL ERROR DETECTED BY DATALINK

**Completion status:** SS\$\_ABORT (0000002C)

**Explanation:** The LAN driver has shut down the device because of a fatal error and is returning all outstanding transmits with SS\$\_OPINCOMPL. The LAN device is automatically restarted, and all the aborted transmits are logged in the error log.

**User Action:** Infrequent occurrences of this error are probably not a problem. If the error occurs frequently or is accompanied by loss or reestablishment of connections to remote computers, there is probably a hardware problem. Check for the proper LAN adapter revision level or call Customer Services.

#### TRANSMIT ERROR FROM DATALINK

**Completion status:** SS\$\_OPINCOMPL (000002D4)

**Explanation:** The LAN driver is in the process of restarting the data link because an error forced the driver to shut down the controller and all users (see FATAL ERROR DETECTED BY DATALINK).

**Completion status:** SS\$\_DEVREQERR (00000334)

**Explanation:** The LAN controller tried to transmit the packet 16 times and failed because of defers and collisions. This condition indicates that LAN traffic is heavy.

**Completion status:** SS\$\_DISCONNECT (0000204C)

**Explanation:** There was a loss of carrier during or after the transmit.

**User Action:** The port emulator automatically recovers from any of these errors, but many such errors indicate either that the LAN controller is faulty or that the LAN is overloaded. If you suspect either of these conditions, contact Customer Services.

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

#### INVALID CLUSTER PASSWORD RECEIVED

**Explanation:** A computer is trying to join the cluster using the correct cluster group number for this cluster but an invalid password. The Port Emulator discards the message. The probable cause is that another cluster on the LAN is using the same cluster group number.

**User Action:** Provide all clusters on the same LAN with unique cluster group numbers.

#### NISCS PROTOCOL VERSION MISMATCH RECEIVED

**Explanation:** A computer is trying to join the cluster using a version of the cluster LAN protocol that is incompatible with the one in use on this cluster.

**User Action:** Install a version of the VMS operating system that uses a compatible protocol, or change the cluster group number so that the computer joins a different cluster.

### C.4.4 OPA0 Error Messages

VAXport drivers detect certain error conditions and attempt to log them. Under some circumstances, attempts to log errors to the error logging device can fail. Such failures can occur because the error logging device is not accessible when attempts are made to log the error condition. Because of the central role that the VAXport device plays in clusters, the loss of error-logged information in such cases makes it difficult to diagnose and fix problems.

A second, redundant method of error logging captures at least some of the information about VAXport device error conditions that would otherwise be lost. This second method consists of broadcasting selected information about the error condition to OPA0, in addition to the port driver's attempt to log the error condition to the error logging device. The VAXport driver attempts both OPA0 error broadcasting and standard error logging under any of the following circumstances:

- The system disk has not yet been mounted.
- The system disk is undergoing mount verification.
- During mount verification, the system disk drive contains the wrong volume.
- Mount verification for the system disk has timed out.
- The local computer is participating in a cluster, and quorum has been lost.

Note the implicit assumption that the system and error logging devices are one and the same.

This second method of reporting errors is also not entirely reliable. Because of the way OPA0 error broadcasting is performed, some error conditions may not be reported. This situation occurs whenever a second error condition is detected before the VAXport driver has been able to broadcast the first error condition to OPA0. In such a case, only the first error condition is reported to OPA0, because that condition is deemed to be the more important one.

Certain error conditions are always broadcast to OPA0, regardless of whether the error logging device is accessible. In general, these are errors that cause the port to shut down either permanently or temporarily.



## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

One OPA0 error message for each error condition is always logged. The text of each error message is similar to the text in the summary displayed by formatting the corresponding standard error log entry using the Error Log Utility. (See Section C.4.3.4 for a list of Error Log Utility summary messages and their explanations.)

Many of the OPA0 error messages contain some optional information such as the remote port number, CI packet information (flags, port operation code, response status, and port number fields), or specific CI port registers.

Following is a list of OPA0 error messages, subdivided by error type. See the CI hardware documentation for a detailed description of the CI port registers (CNF = Configuration Register; PMC = Port Maintenance and Control Register; PSR = Port Status Register), which are optionally displayed for certain of the error conditions. The codes, always file accessible, specify whether the message is *always* logged on OPA0 or is logged only when the system device is *inaccessible*.

#### **Software Errors During Initialization (Always Logged on OPA0)**

%Pxxn, Insufficient Non-Paged Pool for Initialization

%Pxxn, Failed to Locate Port Micro-code Image

%Pxxn, SCSSYSTEMID has NOT been set to a Non-Zero Value

#### **Hardware Errors (Always Logged on OPA0)**

%Pxxn, BIIC failure - BICSR/BER/CNF xxxxxx/xxxxxx/xxxxxx

%Pxxn, Micro-code Verification Error

%Pxxn, Port Transition Failure - CNF/PMC/PSR xxxxxx/xxxxxx/xxxxxx

%Pxxn, Port Error Bit(s) Set - CNF/PMC/PSR xxxxxx/xxxxxx/xxxxxx

%Pxxn, Port Power Down

%Pxxn, Port Power Up

%Pxxn, Unexpected Interrupt - CNF/PMC/PSR xxxxxx/xxxxxx/xxxxxx

%Pxxn, CI Port Timeout

%Pxxn, CI port ucode not at required rev level. RAM/PROM rev is xxxx/xxxx

%Pxxn, CI port ucode not at current rev level. RAM/PROM rev is xxxx/xxxx

%Pxxn, CPU ucode not at required rev level for CI activity

#### **Queue Interlock Failures (Always Logged on OPA0)**

%Pxxn, Message Free Queue Remove Failure

%Pxxn, Datagram Free Queue Remove Failure

%Pxxn, Response Queue Remove Failure

%Pxxn, High Priority Command Queue Insert Failure

%Pxxn, Low Priority Command Queue Insert Failure

%Pxxn, Message Free Queue Insert Failure

%Pxxn, Datagram Free Queue Insert Failure

# Cluster Troubleshooting

## C.4 Diagnosing VAXport Device Problems

### Errors Signaled with a CI Packet

- %Pxxn, Unrecognized SCA Packet - FLAGS/OPC/STATUS/PORT xx/xx/xx/xx  
(ALWAYS)
- %Pxxn, Port has Closed Virtual Circuit - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Software Shutting Down Port  
(ALWAYS)
- %Pxxn, Software is Closing Virtual Circuit - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Received Connect Without Path-Block - FLAGS/OPC/STATUS/PORT xx/xx/xx/xx  
(ALWAYS)
- %Pxxn, Inappropriate SCA Control Message - FLAGS/OPC/STATUS/PORT xx/xx/xx/xx  
(ALWAYS)
- %Pxxn, No Path-Block During Virtual Circuit Close - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, HSC Error Logging Datagram Received - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Remote System Conflicts with Known System - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Virtual Circuit Timeout - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Parallel Path is Closing Virtual Circuit - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Insufficient Non-paged Pool for Virtual Circuits  
(ALWAYS)

### Cable Change-of-State Notification

- %Pxxn, Path #0. Has gone from GOOD to BAD - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Path #1. Has gone from GOOD to BAD - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Path #0. Has gone from BAD to GOOD - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Path #1. Has gone from BAD to GOOD - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Cables have gone from UNCROSSED to CROSSED - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Cables have gone from CROSSED to UNCROSSED - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Path #0. Loopback has gone from GOOD to BAD - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Path #1. Loopback has gone from GOOD to BAD - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Path #0. Loopback has gone from BAD to GOOD - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Path #1. Loopback has gone from BAD to GOOD - REMOTE PORT xxx  
(ALWAYS)
- %Pxxn, Path #0. Has become working but CROSSED to Path #1. - REMOTE PORT xxx  
(INACCESSIBLE)
- %Pxxn, Path #1. Has become working but CROSSED to Path #0. - REMOTE PORT xxx  
(INACCESSIBLE)

## Cluster Troubleshooting

### C.4 Diagnosing VAXport Device Problems

Note that if the port driver can identify the remote SCS node name of the affected computer, the driver replaces the “REMOTE PORT xxx” text with “REMOTE SYSTEM X...”, where X... is the value of the SYSGEN parameter SCSNODE on the remote computer. If the remote SCS node name is not available, the port driver uses the existing message format.

Two other messages concerning the CI port appear on OPA0. They are as follows:

```
%Pxxn, CI port is reinitializing (xxx retries left.)
```

```
%Pxxn, CI port is going off line.
```

The first message indicates that a previous error requiring the port to shut down is recoverable, and that the port will be reinitialized. The “xxx retries left” specifies how many more reinitializations are allowed before the port must be left permanently off line. Each reinitialization of the port (for reasons other than power fail recovery) causes approximately 2kb of nonpaged pool to be lost.

The second message indicates that a previous error is not recoverable, and that the port will be left off line. In this case, the only way to recover the port is to reboot the computer.



---

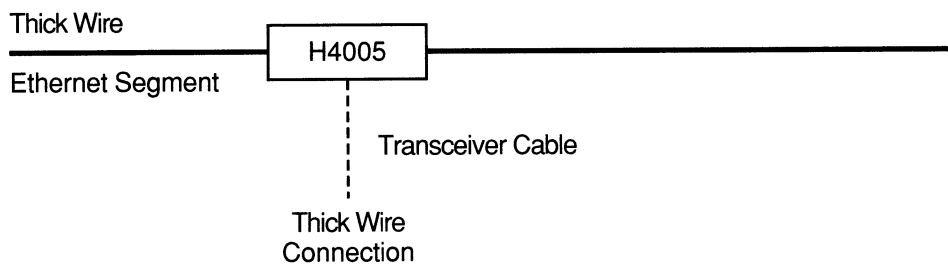
## Local Area VAXcluster Network Connections

This appendix describes the sample local area network connections in a local area VAXcluster.

Each network connection requires a LAN adapter in the host system. Thick wire connections require a transceiver cable connected to one of the following network components:

- Transceiver, H4000, H4005, or equivalent. (See Figure D-1.)
- Digital Ethernet Local Network Interconnect (DELNI), or equivalent. (See Figure D-2.)
- Digital Ethernet Station Adapter (DESTA), or equivalent, plus the components required to make a thinwire connection. (See Figure D-3.)

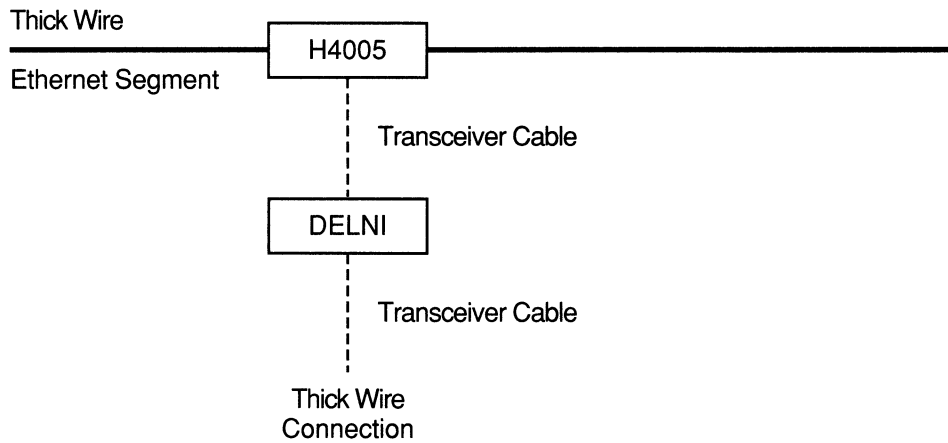
**Figure D-1 Sample Transceiver/Thick Wire Ethernet Connection**



ZK-3737A-GE

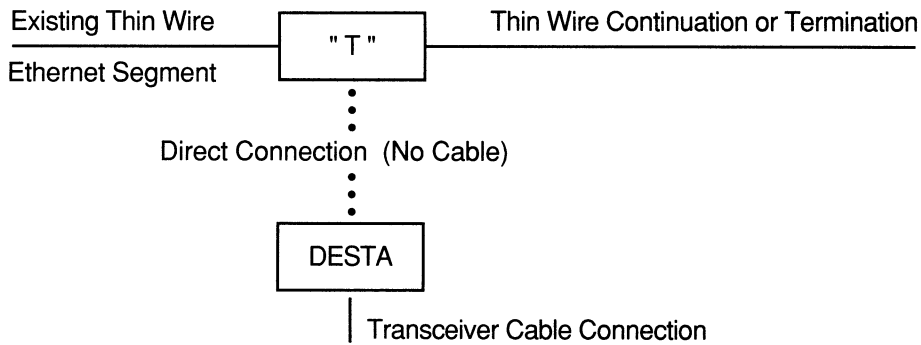
# Local Area VAXcluster Network Connections

**Figure D-2 Sample DELNI Connection**



ZK-3738A-GE

**Figure D-3 Sample DESTA Connection**



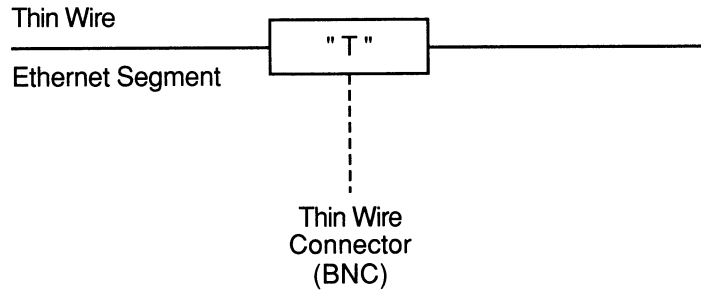
ZK-3739A-GE

Thinwire connections are typically created to:

- Extend an existing thinwire segment. This connection requires a segment of thinwire and a T-connector. Figure D-4 illustrates this type of connection.
- Making a new connection to an existing Digital to Ethernet Multiport Repeater (DEMPR). This connection requires a segment of thinwire, a T-connector, and a terminator. Figure D-5 illustrates this type of connection.

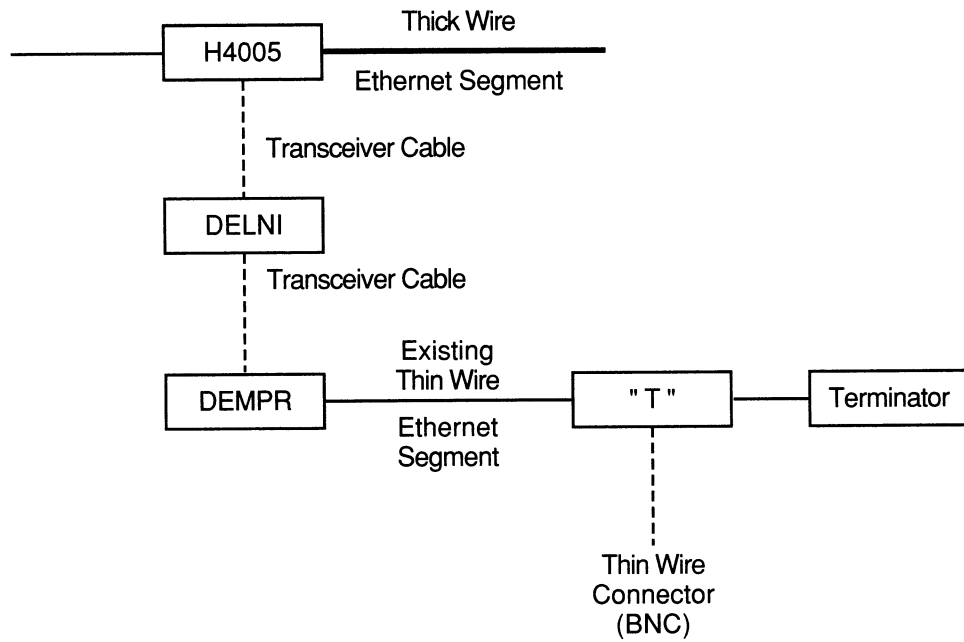
# Local Area VAXcluster Network Connections

Figure D-4 Sample Connection to an Existing Thin Wire Segment



ZK-3740A-GE

Figure D-5 Sample New Connection to an Existing DEMPR



ZK-3741A-GE





---

## Local Area VAXcluster Sample Programs

Sample programs are provided in SYS\$EXAMPLES to start and stop the local area VAXcluster protocol on a LAN adapter, and to enable local area VAXcluster network failure analysis.

For descriptions of the programs, see Section E.1. For information on editing and using the Network Failure Analysis program, refer to Section E.2.

### E.1 Sample Programs for Local Area VAXclusters

VMS provides three sample programs for local area VAXclusters. You can find these sample programs in SYS\$EXAMPLES:

- LAVC\$START\_BUS.MAR—Program to start the local area VAXcluster protocol on a specified LAN adapter
- LAVC\$STOP\_BUS.MAR—Program to stop the local area VAXcluster protocol on a specified LAN adapter
- LAVC\$FAILURE\_ANALYSIS.MAR—Template program to enable the Local Area VAXcluster Network Failure Analysis

PEDRIVER starts the protocol on all of the LAN adapters in the cluster. LAVC\$START\_BUS.MAR and LAVC\$STOP\_BUS.MAR are provided for cluster managers who want to split the network load according to protocol type and therefore do not want the local area VAXcluster protocol running on all of the LAN adapters.

VMS also provides a command file SYS\$EXAMPLES:LAVC\$BUILD.COM to assemble and link these sample programs.

The following sections describe these programs in more detail.

#### E.1.1 Starting the Local Area VAXcluster Protocol on a LAN Adapter

The sample program LAVC\$START\_BUS.MAR, provided in SYS\$EXAMPLES, starts the local area VAXcluster protocol on a specified LAN adapter.

To build this program, first copy the files LAVC\$START\_BUS.MAR and LAVC\$BUILD.COM from SYS\$EXAMPLES to your local directory. Then assemble and link the sample program using the following command:

```
$ @LAVC$BUILD.COM LAVC$START_BUS.MAR
```

## Local Area VAXcluster Sample Programs

### E.1 Sample Programs for Local Area VAXclusters

To start the local area VAXcluster protocol on a LAN adapter, perform the following steps:

- Define the foreign command (DCL symbol).
- Execute the foreign command (LAVC\$START\_BUS.EXE), followed by the name of the LAN adapter on which you want to start the protocol.

---

#### Note

---

In order to execute, LAVC\$START\_BUS.EXE requires the PHY\_IO privilege.

---

The following example starts the cluster protocol on LAN adapter ETA0:

```
$ START_BUS:==$SYSSDISK:[ ]LAVC$START_BUS.EXE
$ START_BUS ETA
```

#### E.1.2 Stopping the Local Area VAXcluster Protocol on a LAN Adapter

The sample program LAVC\$STOP\_BUS.MAR, provided in SYS\$EXAMPLES, stops the local area VAXcluster protocol on a specified LAN adapter.

---

#### Caution

---

Stopping the local area VAXcluster protocol on all LAN adapters will hang the local area VAXcluster satellites and could cause cluster systems to crash with a CLUEXIT bugcheck on mixed interconnect nodes.

---

To build this program, first copy the files LAVC\$STOP\_BUS.MAR and LAVC\$BUILD.COM from SYS\$EXAMPLES to your local directory. Then assemble and link the sample program using the following command:

```
$ @LAVC$BUILD.COM LAVC$STOP_BUS.MAR
```

To stop the local area VAXcluster protocol on a LAN adapter, perform the following steps:

- Define the foreign command (DCL symbol).
- Execute the foreign command (LAVC\$STOP\_BUS.EXE), followed by the name of the LAN adapter on which you want to stop the protocol.

---

#### Note

---

In order to execute, LAVC\$STOP\_BUS.EXE requires the PHY\_IO privilege.

---

The following example stops the cluster protocol on LAN adapter ETA0:

```
$ STOP_BUS:==$SYSSDISK:[ ]LAVC$STOP_BUS.EXE
$ STOP_BUS ETA
```

## Local Area VAXcluster Sample Programs

### E.1 Sample Programs for Local Area VAXclusters

#### E.1.3 Enabling VAXcluster Network Failure Analysis

VMS provides LAVC\$FAILURE\_ANALYSIS.MAR, a sample program that you can edit and use to help detect and isolate a failed network component. LAVC\$FAILURE\_ANALYSIS.MAR is located in SYS\$EXAMPLES. Edit this sample program to include in it a physical description of the communications network for your cluster. Then assemble and link the program and execute it on one (or more) cluster system(s) that will perform the failure analysis. When the program is executed, it provides the physical description of your cluster communications network to the set of routines that perform the failure analysis.

Using the network failure analysis program can help reduce the time necessary for detection and isolation of a failing network component and, therefore, significantly increase cluster availability. The program groups together channels that fail and compares them with the physical description of the cluster network. The program then develops a list of nonworking network components related to the failed channels and uses OPCOM messages to display the names of components with a probability of causing one or more channel failures. If the network failure analysis cannot verify that a portion of a path (containing multiple components) works, the program calls out the first component in the path as the primary suspect (%LAVC-W-PSUSPECT). The other components are listed as secondary or additional suspects (%LAVC-I-ASUSPECT). When the component works again, OPCOM displays the message %LAVC-S-WORKING.

For information on how to edit and use LAVC\$FAILURE\_ANALYSIS.MAR, refer to Section E.2.

#### E.2 Using the Local Area VAXcluster Network Failure Analysis Program

LAVC\$FAILURE\_ANALYSIS.MAR is a sample MACRO program, which you must edit, assemble, link, and execute in order to use. This sample program is located in SYS\$EXAMPLES.

To use the Network Failure Analysis Program, perform these steps:

1. Collect information specific to your cluster communications network. See Section E.2.1 for more information on this step.
2. Edit a copy of LAVC\$FAILURE\_ANALYSIS.MAR to include the information you collected. See Section E.2.2.
3. Assemble, link, and debug the program. See Section E.2.3.
4. Execute the program on one or more of the nodes where you plan to perform the network failure analysis. See Section E.2.4.
5. Test the Local Area VAXcluster Network Failure Analysis subsystem. See Section E.2.5.

##### E.2.1 Collecting Information for the Network Failure Analysis Program

To collect the information you will need to edit the Network Failure Analysis program, perform the following steps:

1. Draw a diagram of your VAXcluster communications network. When you edit LAVC\$FAILURE\_ANALYSIS.MAR, you include this drawing (in electronic form) in the program. Your drawing should show the physical layout of the cluster and include the following components:

## Local Area VAXcluster Sample Programs

### E.2 Using the Local Area VAXcluster Network Failure Analysis Program

- LAN segments or rings
- LAN bridges
- Wiring concentrators, DELNIs, or DEMPRs
- LAN adapters
- VAX systems

For large clusters, it may be necessary to verify the configuration by tracing the cables.

2. Give each component on the drawing a unique label. If your VAXcluster contains a large number of nodes, you may want to replace each node name with a shorter abbreviation. Abbreviating node names can help save space in the electronic form of the drawing when you include it in LAVC\$FAILURE\_ANALYSIS.MAR. For example, you can replace the node name ASTRA with A and call node ASTRA's two LAN adapters A1 and A2.
3. List the following information for each component:
  - Unique label.
  - Type [SYSTEM, LAN\_ADP, DELNI].
  - Location (the physical location of the component).
  - LAN address or addresses (if applicable). (Devices such as DELNIs, DEMPRs, and cables do not have LAN addresses.)
4. Classify each component into one of the following categories:
  - Node: VAX system in the VAXcluster configuration.
  - Adapter: LAN adapter on the VAX system that is normally used for VAXcluster communications.
  - COMPONENT: Generic component in the network. Components in this category can usually be shown to be working if at least one path through them is working. Wiring concentrators, DELNIs, DEMPRs, LAN bridges, and LAN segments and rings typically fall into this category.
  - CLOUD: Generic component in the network. Components in this category cannot be shown to be working even if one or more paths are shown to be working. This type of component is only necessary when multiple paths exist between two points within the network. An example is redundant bridging between LAN segments. At a high level, multiple paths can exist; however, during operation, this bridge configuration allows only one path to exist at one time. In general, this bridge example is probably better handled by representing the active bridge in the description as a COMPONENT and ignoring the standby bridge. (You can identify the active bridge with such network monitoring software as Remote Bridge Management System or Digital Extended LAN Management Software.) With the default bridge parameters, failure of the active bridge will be called out.
5. Use the component labels from step 3 to describe each of the connections in the VAXcluster communications network.

## Local Area VAXcluster Sample Programs

### E.2 Using the Local Area VAXcluster Network Failure Analysis Program

6. Choose a node, or group of nodes, to run the network failure analysis program. You should run the network failure analysis program only on a node that you included in the physical description when you edited LAVC\$FAILURE\_ANALYSIS.MAR. The network failure analysis program on one node operates independently from other systems in the VAXcluster. So, for executing the network failure analysis program, you should choose systems that are not normally shut down. Other good candidates for running the program are systems with the following characteristics:

- Faster CPU speed
- Larger amounts of memory
- More LAN adapters (running local area VAXcluster protocol)

---

#### Note

---

The physical description will be loaded into nonpaged pool and all processing will be performed at IPL 8. CPU use increases as the average number of network components in the network path increases. CPU use also increases as the total number of network paths increases.

---

#### E.2.2 Editing the Network Failure Analysis Program

To edit the Network Failure Analysis program, first copy the files LAVC\$FAILURE\_ANALYSIS.MAR and LAVC\$BUILD.COM from SYS\$EXAMPLES to your local directory. Then use the VAXcluster network map and the other information you collected to edit the copy of the LAVC\$FAILURE\_ANALYSIS.MAR.

Example E-1 shows the part of LAVC\$FAILURE\_ANALYSIS.MAR that you edit.

#### Example E-1 Portion of LAVC\$FAILURE\_ANALYSIS.MAR to Edit

```
;      *** Start edits here ***
;      Edit 1. ❶
;
;      Define the hardware components needed to describe
;      the physical configuration.
;

NEW_COMPONENT  SYSTEM          NODE
NEW_COMPONENT  LAN_ADP         ADAPTER
NEW_COMPONENT  DEMPR           COMPONENT
NEW_COMPONENT  DELNI           COMPONENT
NEW_COMPONENT  SEGMENT         COMPONENT
NEW_COMPONENT  NET_CLOUD       CLOUD
```

(continued on next page)



## Local Area VAXcluster Sample Programs

### E.2 Using the Local Area VAXcluster Network Failure Analysis Program

#### Example E-1 (Cont.) Portion of LAVC\$FAILURE\_ANALYSIS.MAR to Edit

```
CONNECTION      LNI_A, A2
CONNECTION      LNI_A, B2

CONNECTION      Sa, BRIDGES
CONNECTION      Sa, MPR_A
CONNECTION      Sa, D1

CONNECTION      Sb, BRIDGES
CONNECTION      Sb, LNI_A
CONNECTION      Sb, D2

.PAGE
.SBTTL Define the Local Data.

;      *** End of edits ***
```

In the program, EDIT [number] identifies a place where you must edit the program to incorporate information about your network. Make the following edits to the program:

- ① At EDIT 1 in the template, define a category for each component in the configuration. Use the information from step 5 in Section E.2.1. Use the format: NEW\_COMPONENT Component\_Type Category  
In the following example, a DEMPR is defined as part of the Component category: NEW\_COMPONENT DEMPR COMPONENT
- ② At EDIT 2, draw the network map you drew for step 1 of Section E.2.1. Including the map here in LAVC\$FAILURE\_ANALYSIS.MAR gives you an electronic record of the map you can locate and update more easily than a drawing on paper.
- ③ At EDIT 3, list each VAXcluster node and its LAN adapters. Use one line for each node. Each line should include the following information. Separate the items of information with commas to create a table of the information.
  - Component type, followed by a comma.
  - Label from the network map, follow by a comma.
  - Node name (for SYSTEM components only). If there is no node name, type a comma.
  - Descriptive text that the network failure analysis program displays if it detects a failure with this component. Put this text within angle brackets (<>). This text should include the component's physical location.
  - LAN hardware address (for LAN adapters).
  - DECnet LAN address for the LAN adapter that DECnet uses.
- ④ At EDIT 4, list each of the other network components. Use one line for each component. Each line should include the following information:
  - Component name and category you defined with NEW\_COMPONENT.
  - Label from the network map.
  - Descriptive text that the Network Failure Analysis program displays if it detects a failure with this component. Include a description of the physical location of the component.

## Local Area VAXcluster Sample Programs

### E.2 Using the Local Area VAXcluster Network Failure Analysis Program

- LAN hardware address (optional).
  - Alternate LAN address (optional).
- ⑤ At EDIT 5, define the connections between the network components. Use the CONNECTION macro and the labels for the two components that are connected. Include the following information:
- CONNECTION macro name
  - First component label
  - Second component label

#### E.2.3 Assembling and Linking the Failure Analysis Program

Use the following command procedure to assemble and link the program:

```
$ @LAVC$BUILD.COM LAVCSFAILURE_ANALYSIS.MAR
```

Make the edits necessary to fix the assembly or link errors, such as errors caused by mistyping component labels in the path description. Assemble the program again.

#### E.2.4 Executing the Network Failure Analysis Program

Before you execute LAVC\$FAILURE\_ANALYSIS.EXE, modify the startup files in SYS\$COMMON:[SYSMGR] to add a conditional statement to run LAVC\$FAILURE\_ANALYSIS.EXE only on the node for which you supplied data. The following is an example of such a conditional statement:

```
$ If F$GETSYI ("nodename").EQS."OMEGA"  
$ THEN  
$ RUN SYS$MANAGER:LAVC$FAILURE_ANALYSIS.EXE  
$ ENDF
```

---

#### Note

---

In order to execute, LAVC\$FAILURE\_ANALYSIS.EXE requires the PHY\_IO privilege.

---

Execute the linked program on one of the nodes that will perform the network failure analysis. Use an account that has the PHY\_IO privilege. Then execute the program on each of the nodes that will perform the network failure analysis.

After it executes, the program displays the approximate amount of nonpaged pool required for the network description. The display is similar to the following display:

```
Non-paged Pool Usage: ~ 10004 bytes
```

On each system running the Network Failure Analysis, modify the file SYS\$SPECIFIC:[SYSEXE]MODPARAMS.DAT to include the following lines, where *value* is the value displayed for non-paged pool usage.

```
ADD_NPAGEDYN = value  
ADD_NPAGEVIR = value
```

Run AUTOGEN on each system for which you modified MODPARAMS.DAT.



# Local Area VAXcluster Sample Programs

## E.2 Using the Local Area VAXcluster Network Failure Analysis Program

### E.2.5 Testing the Network Failure Analysis Subsystem

Test the program by causing a failure. For example, disconnect a transceiver cable or thinwire segment or power fail a bridge, a DELNI, or a DEMPR. Then check the OPCOM messages to see if LAVC\$FAILURE\_ANALYSIS reports the failed component correctly. If the failure has not been reported, check your edits to the network failure analysis program.

### E.2.6 PEDRIVER Suspect Network Component Display

When a VAXcluster network component failure occurs, OPCOM displays a list of suspected components. Displaying the list through OPCOM allows the system manager to enable and disable selectively the display of these messages.

The following are sample displays:

```
%%%%%%%%%% OPCOM 1-AUG-1991 14:16:13.30 %%%%%%%%%% (from node BETA at
1-AUG -1991 14:15:55.38) Message from user SYSTEM on BETA
LAVC-W-PSUSPECT, component_name

%%%%%%%%%% OPCOM 1-AUG-1991 14:16:13.41 %%%%%%%%%% (from node BETA at
1-AUG-1991 14:15:55.49) Message from user SYSTEM on BETA
%LAVC-W-PSUSPECT, component_name

%%%%%%%%%% OPCOM 1-AUG-1991 14:16:13.50 %%%%%%%%%% (from node BETA at
1-AUG-1991 14:15:55.58) Message from user SYSTEM on BETA
%LAVC-I-ASUSPECT, component_name
```

The OPCOM display of suspected failures uses the following prefixes to list suspected failures:

- %LAVC-W-PSUSPECT—Primary suspects
- %LAVC-I-ASUSPECT—Secondary or additional suspects
- %LAVC-S-WORKING—Suspect component is now working

The text following the message prefixes is the network component description you supply when you edit LAVC\$FAILURE\_ANALYSIS.MAR.



---

## Local Area VAXcluster Subroutine Package

VMS provides a subroutine package to control the local area VAXcluster features described in Appendix E. The sample programs LAVC\$FAILURE\_ANALYSIS.MAR, LAVC\$START\_BUS.MAR, and LAVC\$STOP\_BUS.MAR use this subroutine package, and these programs may be sufficient for your needs. The subroutine package provides a supported way to extend the sample programs as your needs require.

This subroutine package offers the following function calls to manage LAN adapters:

- SYS\$LAVC\_START\_BUS—Directs PEDRIVER to start the local area VAXcluster protocol on a specified LAN adapter
- SYS\$LAVC\_STOP\_BUS—Directs PEDRIVER to stop the local area VAXcluster protocol on a specified LAN adapter

The subroutine package also offers the following function calls to control the Network Failure Analysis subsystem:

- SYS\$LAVC\_DEFINE\_NET\_COMPONENT—Creates a representation of a physical network component
- SYS\$LAVC\_DEFINE\_NET\_PATH—Creates a directed list of network components between two network nodes
- SYS\$LAVC\_ENABLE\_ANALYSIS—Enables the Network Failure Analysis, which makes it possible to analyze future channel failures
- SYS\$LAVC\_DISABLE\_ANALYSIS—Stops the Network Failure Analysis and deallocates the memory used for the physical network description

### F.1 Subroutine Package to Start the Protocol on a LAN Adapter

The SYS\$LAVC\_START\_BUS routine starts the local area VAXcluster protocol on a specified LAN adapter.

To use the routine SYS\$LAVC\_START\_BUS, specify the following parameter:

- BUS\_NAME: String descriptor representing the LAN adapter name buffer, passed by reference. The LAN adapter name must be 15 characters or less.

The following FORTRAN sample program uses the SYS\$LAVC\_START\_BUS call to start the local area VAXcluster protocol on the LAN adapter XQA:

## Local Area VAXcluster Subroutine Package

### F.1 Subroutine Package to Start the Protocol on a LAN Adapter

```
PROGRAM START_BUS
EXTERNAL SYS$LAVC_START_BUS
INTEGER*4 SYS$LAVC_START_BUS
INTEGER*4 STATUS
STATUS = SYS$LAVC_START_BUS ( 'XQA0:' )
CALL SYS$EXIT ( %VAL ( STATUS ))
END
```

The `SYS$LAVC_START_BUS` call returns a VMS status value in register R0. A success status indicates that PEDRIVER is attempting to start the local area VAXcluster protocol on the specified adapter. A failure status indicates that PEDRIVER cannot start the protocol on the specified LAN adapter.

`SYS$LAVC_START_BUS` can return the following errors:

- `SS$_ACCVIO`: This status is returned for the following conditions:
  - No access to the argument list
  - No access to the LAN adapter name buffer descriptor
  - No access to the LAN adapter name buffer
- `SS$_DEVACTION`: BUS already exists. PEDRIVER is already trying to use this LAN adapter for the local area VAXcluster protocol.
- `SS$_INSFARG`: Not enough arguments supplied
- `SS$_INSFMEM`: Insufficient nonpaged pool to create the BUS data structure
- `SS$_INVBUSNAM`: Invalid BUS name specified. The device specified does not represent a LAN adapter that can be used for the protocol.
- `SS$_IVBUFLN`: This status value is returned under the following conditions:
  - Length of the LAN adapter name is 0.
  - Length of the LAN adapter name is greater than 15.
- `SS$_NOSUCHDEV`: This status value is returned under the following conditions:
  - The LAN adapter name specified does not correspond to a LAN device available to PEDRIVER on this system.
  - No LAN drivers are loaded in this system; `NET$AR_LAN_VECTOR` is zero.
  - PEDRIVER is not initialized; PEDRIVER's PORT structure is not available.
- `SS$_NOTNETDEV`: PEDRIVER does not support the specified LAN device.
- `SS$_SYSVERDIF`: The specified LAN device's driver does not support the VCI interface version required by PEDRIVER.
- PEDRIVER can return additional errors that indicate it has failed to create the connection to the specified LAN adapter.

## **F.2 Subroutine Package to Stop the Protocol on a LAN Adapter**

The SYS\$LAVC\_STOP\_BUS routine stops the local area VAXcluster protocol on a specified LAN adapter. To use this routine, specify the following parameter:

- **BUS\_NAME:** String descriptor representing the LAN adapter name buffer, passed by reference. The LAN adapter name must be 15 characters or less.

The following FORTRAN sample program shows the SYS\$LAVC\_STOP\_BUS call used to stop the local area VAXcluster protocol on the LAN adapter XQB:

```
PROGRAM STOP_BUS
EXTERNAL SYS$LAVC_STOP_BUS
INTEGER*4 SYS$LAVC_STOP_BUS
INTEGER*4 STATUS

STATUS = SYS$LAVC_STOP_BUS ( 'XQB' )
CALL SYS$EXIT ( %VAL ( STATUS ) )
END
```

The SYS\$LAVC\_STOP\_BUS call returns a VMS status value in register R0. A success status indicates that PEDRIVER is attempting to shut down the local area VAXcluster protocol on the specified adapter. A failure status indicates that PEDRIVER cannot shut down the protocol on the specified LAN adapter. However, PEDRIVER performs the shutdown asynchronously and there could be other reasons why PEDRIVER is unable to complete the shutdown.

SYS\$LAVC\_STOP\_BUS can return the following errors:

- **SS\$\_ACCVIO:** This status is returned for the following conditions:
  - No access to the argument list
  - No access to the LAN adapter name buffer descriptor
  - No access to the LAN adapter name buffer
- **SS\$\_INVBUSNAM:** Invalid BUS name specified. The device specified does not represent a LAN adapter that can be used for the local area VAXcluster protocol.
- **SS\$\_IVBUFLEN:** This status value is returned under the following conditions:
  - Length of the LAN adapter name is 0.
  - Length of the LAN adapter name is greater than 15.
- **SS\$\_NOSUCHDEV:** This status value is returned under the following conditions:
  - The LAN adapter name specified does not correspond to a LAN device that is available to PEDRIVER on this system.
  - No LAN drivers are loaded in this system. NET\$AR\_LAN\_VECTOR is zero.
  - PEDRIVER is not initialized. PEDRIVER's PORT structure is not available.

### F.3 Subroutine Package to Create a Network Component

The SYS\$LAVC\_DEFINE\_NET\_COMPONENT subroutine call creates a representation for a physical network component.

Specify the following parameters for SYS\$LAVC\_DEFINE\_NET\_COMPONENT:

- **Component\_description:** Address of a string descriptor representing network component name buffer. The network component name must be less than or equal to COMP\$C\_MAX\_NAME\_LEN characters.
- **Nodename\_length:** Address of the length of the node name. This address is located at the beginning of the network component name buffer for COMP\$C\_NODE types. You should use zero for other component types.
- **Component\_type:** Address of the component type. These values are defined by \$PEMCOMPDEF found in SYS\$LIBRARY:LIB.MLB.
- **Lan\_hardware\_addr:** String descriptor of a buffer containing the component's LAN hardware address (6 bytes). You must specify this value for COMP\$C\_ADAPTER types. For other component types, this value is optional.
- **Lan\_decnet\_addr:** String descriptor of a buffer containing the component's LAN DECnet address (6 bytes). This is an optional parameter for all component types.
- **Component\_id\_value:** Address of a longword that is written with the component ID value.

Use the following format to specify the parameters:

```
STATUS = SYS$LAVC_DEFINE_NET_COMPONENT (
    component_description,
    nodename_length,
    component_type,
    lan_hardware_addr,
    lan_decnet_addr,
    component_id_value )
```

This subroutine call defines a network component. If successful, this call creates a COMP data structure and returns its ID value. This call copies user-specified parameters into the data structure and sets the reference count to zero.

This call also returns a status, as well as the component ID value. The component ID value is a 32-bit value that has a one-to-one association with a network component. Lists of these component IDs are passed to SYS\$LAVC\_DEFINE\_NET\_PATH to specify the components used when a packet travels from one node to another.

SYS\$LAVC\_DEFINE\_NET\_COMPONENT can return the following errors:

- **SS\$\_ACCVIO**—This status is returned under the following conditions:
  - No access to the network component name buffer descriptor
  - No access to the network component name buffer
  - No access to the component's LAN hardware address if a nonzero value was specified
  - No access to the component's LAN DECnet address if a nonzero value was specified
  - No access to the lan\_hardware\_addr string descriptor

## Local Area VAXcluster Subroutine Package

### F.3 Subroutine Package to Create a Network Component

- No access to the lan\_decnet\_addr string\_descriptor
- No write access to the component\_id\_value
- No access to the component\_type
- No access to the nodename\_length
- No access to the argument list
- SS\$\_DEVACTIVE: Analysis program already running. You must stop the analysis by calling the SYS\$LAVC\_DISABLE\_ANALYSIS before you define the network components and the network component lists.
- SS\$\_INSFARG: Not enough arguments supplied.
- SS\$\_INVCOMPTYPE: The component type is 0 or greater than or equal to COMP\$C\_INVALID.
- SS\$\_IVBUFLen: This status value is returned under the following conditions:
  - Length of the component name is 0.
  - Length of the component name is greater than COMP\$C\_MAX\_NAME\_LEN.
  - Length of the node name is 0 and the component type is COMP\$C\_NODE.
  - Length of the node name is greater than 8 and the component type is COMP\$C\_NODE.
  - Length of lan\_hardware\_addr is less than 6.
  - Length of lan\_decnet\_addr is less than 6.

### F.4 Subroutine Package to Create a Network Component List

The SYS\$LAVC\_DEFINE\_NET\_PATH subroutine call creates a directed list of network components between two network nodes. A directed list is a list of all the components through which a packet passes as it travels from the failure analysis node to other nodes in the cluster network.

Specify the following parameters for SYS\$LAVC\_DEFINE\_NET\_PATH:

- Network\_component\_list: Address of a string descriptor for a buffer containing the component ID values for each of the components in the path. List the component ID values in the order in which a network message travels through them. Specify components in the following order:
  1. The local node
  2. The local LAN adapter
  3. Intermediate network components
  4. The remote network LAN adapter
  5. The remote nodeYou must list two nodes and two LAN adapters in the network path. The buffer length must be greater than 15 bytes and less than 509 bytes.
- Used\_for\_analysis\_status: Address of a longword status value that is written. This status indicates if this network path has any value for the network failure analysis.

## Local Area VAXcluster Subroutine Package

### F.4 Subroutine Package to Create a Network Component List

- `Bad_component_id`: Address of a longword value that contains the component ID that is in error if an error is detected while processing the component list.

Use the following format to specify the parameters:

```
STATUS = SYS$LAVC_DEFINE_NET_PATH (
    network_component_list,
    used_for_analysis_status,
    bad_component_id )
```

This subroutine call creates a directed list of network components that describe a specific network path. This call, if successful, creates a CLST data structure. If one node is the local node, then this data structure is associated with a PEDRIVER channel. In addition, the reference count for each network component in the list is incremented. If neither node is the local node, then the `used_for_analysis_status` contains an error status.

This call returns a status in register R0 indicating that the network component list has the correct construction. If this value is successful, the `used_for_analysis_status` value will indicate whether network path is useful for network analysis performed on the local node.

If a failure status returned in R0 is `SS$_INVCOMPID`, the `bad_component_id` will contain the value of the `bad_component_id` found in the buffer.

The following errors can be returned:

- `SS$_ACCVIO`: This status value can be returned under the following conditions:
  - No access to the descriptor or the network component ID value buffer
  - No access to the argument list
  - No write access to `used_for_analysis_status`
  - No write access to `bad_component_id`
- `SS$_DEVACTIVE`: Analysis already running. You must stop the analysis by calling the `SYS$LAVC_DISABLE_ANALYSIS` function before defining the network components and the network component lists.
- `SS$_INSFARG`: Not enough arguments supplied
- `SS$_INVCOMPID`: Invalid network component ID specified in the buffer. `Bad_component_id` contains the failed component ID.
- `SS$_INVCOMPLIST`: This status value can be returned under the following conditions:
  - Fewer than two nodes were specified in the node list.
  - Greater than two nodes were specified in the list.
  - The first network component ID was not a `COMP$C_NODE` type.
  - The last network component ID was not a `COMP$C_NODE` type.
  - Fewer than two adapters were specified in the list.
  - Greater than two adapters were specified in the list.
- `SS$_IVBUFLN`: Length of the network component ID buffer is less than 16, is not a multiple of 4, or is greater than 508.



## Local Area VAXcluster Subroutine Package

### F.4 Subroutine Package to Create a Network Component List

- **SS\$\_RMTPATH:** Network path is not associated with the local node. This status is only returned to indicate whether this path was needed for network failure analysis on the local node.

### F.5 Subroutine Package to Start the Network Component Failure Analysis

The `SYS$LAVC_ENABLE_ANALYSIS` subroutine call starts the Network Component Failure Analysis. The following is an example of using the call:

```
STATUS = SYS$LAVC_ENABLE_ANALYSIS ( )
```

This subroutine call attempts to enable the network component failure analysis code. The attempt will succeed if at least one component list is defined.

This call returns a status in register R0.

The following errors can be returned:

- **SS\$\_DEVOFFLINE:** PEDRIVER is not properly initialized. ROOT or PORT block is not available.
- **SS\$\_NOCOMPLSTS:** No network connection lists exist. Network analysis is not possible.
- **SS\$\_WASSET:** Network component analysis is already running.

### F.6 Subroutine Package to Stop the Network Component Failure Analysis

The `SYS$LAVC_DISABLE_ANALYSIS` subroutine call stops the Network Component Failure Analysis. The following is an example of using the call:

```
STATUS = SYS$LAVC_DISABLE_ANALYSIS ( )
```

This subroutine call disables the network component failure analysis code and, if analysis was enabled, deletes all the network component definitions and network component list data structures from nonpaged pool.

`SYS$LAVC_DISABLE_ANALYSIS` returns a status in register R0.

The following errors can be returned:

- **SS\$\_DEVOFFLINE:** PEDRIVER is not properly initialized. ROOT or PORT block is not available.
- **SS\$\_WASCLR:** Network component analysis already stopped.



---

## Transmit Channel Selection

PEDRIVER uses a single channel to transmit and all channels to receive. The driver software chooses a transmit channel on a per-remote-node basis. A transmit channel selection algorithm ensures that messages are sent in the order they are expected to be received. Sending the messages in this way also maintains compatibility with previous versions of VMS. The selected transmit channel is called the *preferred channel*.

PEDRIVER continuously computes a network delay value for each channel. PEDRIVER switches the preferred channel based on observed network delays or network component failures. Switching to a new transmit channel sometimes causes messages to be received out of the desired order. PEDRIVER now uses a receive cache to reorder these messages, instead of discarding them. PEDRIVER's use of the receive cache prevents the remote node from retransmitting the messages after a 2-second delay.

Some restrictions apply to remote nodes running a version of VMS prior to 5.4-3. Messages received out of order on remote nodes are discarded since the receive cache does not exist. For remote nodes, PEDRIVER cannot switch channels based on the observed network delays. In this case, PEDRIVER chooses a single transmit channel and uses it until the channel fails. Only at that time will PEDRIVER switch to another channel.

With these algorithms, PEDRIVER has a higher chance of utilizing adapters on a server node that continuously communicates with a number of clients. In a two-node cluster, PEDRIVER will actively use, at most, two LAN adapters: one to transmit and one to receive. Additional adapters provide higher availability and alternate paths that can be used to avoid network congestion. As more nodes are added to the cluster, PEDRIVER has a higher chance of using the additional adapters.



## A

---

- ACP\_REBLDSYSD parameter, 5-13
- Adding a computer, 7-7, 7-22, 7-39
  - adjusting EXPECTED\_VOTES, 7-22
- Allocation class, 5-7
  - assigning value to computers, 5-8
  - assigning value to HSC subsystems, 5-8
  - determining for mixed-interconnect cluster, 7-4
  - rules for specifying, 5-7
  - sample configurations, 5-9
  - using in a distributed environment, 3-7
- ALLOCLASS, 5-8
- Asterisk wildcard
  - in START/QUEUE/MANAGER command, 6-1
- Authorize Utility (AUTHORIZE), B-1
- AUTOGEN.COM command procedure
  - enabling or disabling disk server, 7-13
  - executed by CLUSTER\_CONFIG.COM, 7-2
  - running with feedback option, 7-24, 7-39
  - specifying dump file, 7-37
- Availability
  - of data, 3-1, 3-5
  - of data center, 2-9
  - of network, E-3
  - of queue manager, 6-1
  - surviving LAN component failures, 2-11, E-3

## B

---

- Batch queue, 6-5
  - assigning unique name to, 6-7
  - clusterwide generic, 6-8
  - initializing, 6-8
  - sample configuration, 6-5
  - setting up, 6-7
  - starting, 6-8
  - SYS\$BATCH, 6-8
- Boot events, C-1
- Boot node
  - See Boot server
- Boot server
  - functions, 2-2
- Broadcast messages, 7-12

## C

---

- CI (Computer Interconnect), 1-3, C-1
  - analyzing error log entry, C-15
  - cable repair, C-14
  - communication path, C-10
  - configuring multiple adapters, 2-9
  - connected computer
    - failure to boot, C-3
    - failure to join the cluster, C-6
  - device-attention entry, C-16
  - error log entry, C-21
    - formats, C-16
  - logged message entry, C-19
  - MSCP server accesses shadow sets, 3-8
  - port
    - controller, 1-3
    - loopback datagram facility, C-12
    - polling, C-9
    - verifying function, C-11
- CI-connected computer
  - adding, 7-7
- CLUEXIT bugcheck
  - diagnosing, C-8
- Cluster
  - boot events, C-1
  - CLUEXIT bugcheck, C-8
  - common-environment, 4-1
  - communication mechanisms, 1-5
  - configuration planning, 1-5
  - configuration type, 2-1
    - changing to mixed-interconnect, 7-18
  - configuring, 7-5, 7-32
  - connection manager, 1-4
  - disk class driver, 1-4
  - disk devices, 5-1
  - distributed file system, 1-4
  - distributed job controller, 1-4
  - distributed lock manager, 1-4
  - distributed processing, 1-1
  - error log entries for VAXport device, C-15
  - failure of computer to boot, C-1
  - failure of computer to join the cluster, C-1, C-6
    - group number, 7-30
    - hang condition, C-7

## Cluster (Cont.)

- hardware components, 1-3
  - installing license, 4-4
  - interconnect devices, 1-3
  - maintaining, 7-24
  - managing queues, 6-1
  - MSCP server, 1-4
  - multiple-environment, 4-1
  - password, 7-30
  - preparing operating environment, 4-1
  - queues, 1-2
  - quorum disk, 3-2
  - quorum disk watcher, 3-2
  - quorum file (QUORUM.DAT), 3-3
  - rebooting after configuration change, 7-22
  - reconfiguring, 7-22
  - recording configuration data, 7-25
  - recovering from startup procedure failure, C-6
  - resource access, 1-4
  - resource locking, 1-4
  - restoring quorum, 7-26
  - security management, 2-15, 7-29
  - shared disk resources, 1-2
  - shared processing and printer resources, 1-1
  - summary of configuration procedures, 1-5
  - System Communications Services (SCS), 1-4, C-9
  - tape class driver, 1-5
  - tape devices, 5-1
  - TMSCP server, 1-5
  - troubleshooting, C-1
  - types of operating environments, 4-1
  - updating MODPARAMS.DAT files, 7-22
  - VAXport device error log entries, C-15
  - VAXport driver, 1-4, C-9
  - voting member, 3-1
    - adding, 7-2, 7-7, 7-22
    - removing, 7-2, 7-12, 7-22
  - workload balancing, 1-1
- Cluster-accessible disks, 1-2, 5-1
- and MSCP server, 5-2
- Cluster-accessible tapes, 1-2, 5-1
- and TMSCP server, 5-2
- Cluster authorization file (CLUSTER\_AUTHORIZE.DAT), 2-15, 7-30
- Cluster SYSGEN parameters, A-1 to A-3
- Clusterwide queue manager, 6-1
- CLUSTER\_CONFIG.COM command procedure
- adding computers, 7-6
  - change options, 7-14
  - converting standalone computer to cluster computer, 7-19
  - creating a duplicate system disk, 7-20
  - enabling disk server, 5-3, 7-15
  - functions, 7-1
  - modifying satellite LAN hardware address, 7-13
  - preparing to execute, 7-5

## CLUSTER\_CONFIG.COM command procedure (Cont.)

- removing computers, 7-12
  - required information, 7-5
  - system files created for satellite, 7-2
- Common command procedure
- coordinating, 4-9
  - creating, 4-10
  - executing, 4-10
  - on cluster-accessible disks, 4-9
  - SYLOGIN.COM, 4-11
- Common-environment cluster, 4-1, 4-10
- Common file
- coordinating for multiple boot servers, 4-14
  - coordinating for multiple system disks, 4-14
  - mail database, 4-14
  - moving off system disk, 7-37
  - NETPROXY.DAT, 4-12
  - QMAN\$MASTER.DAT, 4-15, 6-2
  - RIGHTSLIST.DAT, 4-13
  - system, 4-11
  - SYSUAF.DAT, 4-12
  - VMSMAIL\_PROFILE.DATA, 4-14
- Common MAIL database, 4-14
- Common rights database, 4-13
- Common system disk
- directory structure, 4-2
- Computer Interconnect (CI)
- See CI (Computer Interconnect)
- Computer-specific startup functions, 4-11
- Configuration
- changing type from CI-only to mixed-interconnect, 7-18
  - changing type from local area to mixed-interconnect, 7-19
  - multiple adapters on local area VAXcluster, 2-10
  - recording data, 7-25
  - three LAN segments in a local area VAXcluster, 2-12
  - two LAN segments in a local area VAXcluster, 2-11
  - type, 2-1
- Connection manager, 1-4, 3-1
- Conversational bootstrap
- controlling, 7-31
- Convert Utility (CONVERT), B-2
- using to merge SYSUAF.DAT files, B-1
- Crossed cable, C-12

## D

- Data availability, 3-5
- DECnet
  - See DECnet-VAX network
- DECnet Maintenance Operation Protocol (MOP)
  - See MOP
- DECnet-VAX network

## DECnet-VAX network (Cont.)

- cluster functions, 1-5
- cluster satellite synonym, 7-33
- configuring, 4-6
- copying remote node databases in VAXcluster environments, 4-8
- disabling LAN device, 4-7
- downline loading, 7-34
- enabling circuit service for cluster MOP server, 4-6
- installing license, 4-4
- making databases available clusterwide, 4-8
- making remote node data available clusterwide, 4-6
- maximum address value, defining for cluster boot server, 4-6
- modifying satellite LAN hardware address, 7-13
- monitoring LAN activity, 7-26
- NETCONFIG.COM command procedure, 4-7
- NETNODE\_REMOTE.DAT file, renaming to SYS\$COMMON directory, 4-7
- Network Control Program (NCP), 4-8
- restoring satellite configuration data, 7-11
- starting, 4-8
- tailoring, 4-6
- troubleshooting, E-3
- VAXcluster alias, 4-6, 4-9, 7-42

## Device driver

- loading, 4-10

## DIGITAL Extended LAN Management Software (DECelms), 7-26

## DIGITAL Storage Architecture (DSA)

- See DSA

## DIGITAL Storage System Interconnect (DSSI)

- See DSSI

## Directory structure

- on common system disk, 4-2

## Disk

- See also Dual-pathed disk, Dual-ported disk
- allocation class, 5-8
- cluster-accessible, 5-1
  - storing common procedures on, 4-9
- command procedures for setting up, 4-11
- configuring, 5-12
- directory structure on common system disk, 4-2
- DSA, 5-5
- DSA controller, 5-2
- DSSI, 5-1, 5-6
- dual-pathed, 5-1
- HSC, 5-1, 5-2, 5-9
- local, 5-2
- managing, 5-1
- MASSBUS, 5-6
- mounting, 5-12
- MOUNT/NOREBUILD, 5-13
- MSCP-served, 5-1

## Disk (Cont.)

- naming conventions, 5-7
  - paths, 5-7
  - quorum, 3-2
  - rebuilding, 5-13
  - restricted access, 5-1
  - selecting server, 7-3
  - setting up, 4-11
- ## Disk class driver, 1-4
- ## Disk controller, 1-2
- ## Disk server
- configuring LAN adapter, 7-34
  - configuring memory, 7-34
  - functions, 2-2
  - selecting, 7-3
- ## DISK\_QUORUM parameter, 3-2, A-1
- ## Distributed file system, 1-4
- ## Distributed job controller, 1-4
- ## Distributed lock manager, 1-4
- ## Distributed processing, 1-1, 6-1
- ## Distributing members of shadow sets, 3-7
- ## Distributing shadow sets, 3-7
- ## DSA, 5-2
- disks and tapes, 1-2, 5-5
  - support for compliant hardware, 3-7
- ## DSSI
- changing allocation class values on DSSI subsystems, 7-23
  - configurations, 2-5, 2-10
  - configuring multiple adapters, 2-10
  - disk, 5-6
  - ISE, 2-5, 5-4, 5-6
  - MSCP server accesses shadow sets, 3-8
  - port controller, 1-3
- ## Dual-pathed disk, 5-4
- DSA, 5-5
  - DSSI, 5-6
  - HSC, 5-4, 5-9
  - MASSBUS, 5-6
- ## Dual-pathed tape
- DSA, 5-5
  - HSC, 5-4
- ## Dual-ported disk
- DSA, 5-5
  - MASSBUS, 5-6
  - setting up, 4-10
- ## Dual-ported tape
- DSA, 5-5
- ## Dump file
- controlling size, 7-37
  - managing, 7-37
  - sharing, 7-38
- ## DUMPFIL AUTOGEN symbol, 7-37
- ## DUMPSYLE AUTOGEN symbol, 7-37
- ## Duplicate system disk
- creating, 7-20

## E

---

### Ethernet

- adapter, 1-3
- configurations, 2-2, 2-10
- configuring adapter, 7-34
- error log entry, C-21
- hardware address, 7-5
- monitoring activity, 7-26
- multiple adapters, 2-10
- port, C-9
- EXPECTED\_VOTES parameter, 3-1, 7-22, 7-27, A-1

## F

---

### Failover

- dual-ported DSA disk, 5-5
- dual-ported DSA tape, 5-5
- of queue manager, 6-1

### Fault tolerance

- types of faults, 3-8

### FDDI

- adapter, 1-3
- configurations, 2-2, 2-6, 2-10
- configuring adapter, 7-34
- error log entry, C-21
- hardware address, 7-5
- influence of LRPSIZE on, 2-14
- large packet support, 2-14
- monitoring activity, 7-26
- multiple adapters, 2-10
- port, C-9
- use of priority field, 2-14

### Fiber Distributed Data Interface (FDDI)

- See FDDI

### File access

- controlling, 4-12

### Files-11 volume structure

- shadowing support, 3-7

### File system

- coordinating, 4-11

### FORWARDING\_DELAY parameter, 2-13

## G

---

### Generic queue

- clusterwide batch, 6-8
- clusterwide printer, 6-4
- establishing, 6-4

### Group number

- See Security management

## H

---

### Hang condition

- diagnosing, C-7

### HELLO\_INTERVAL parameter, 2-13

### Hierarchical storage controller (HSC) subsystem

- See HSC subsystem

### HSC disk, 1-2, 2-4, 5-2

- dual-pathed, 5-4, 5-9

### HSC subsystem, 1-2

- changing allocation class values, 7-23

### HSC tape, 5-2

- dual-pathed, 5-4

## I

---

### Integrated Storage Elements (ISE)

- See ISE

### Integrity

- of VAXcluster, 3-1

### ISE, 2-5

- DSSI connection, 5-4, 5-6

- in DSSI-based cluster, 1-3

- use in clusters, 1-2

## J

---

### Job controller, 1-4

- separation from queue manager, 6-1

### Job-controller queue database file

- specifying location of, 6-2

## K

---

### Known images

- installing, 4-11

## L

---

### LAN

- adapter, 1-3

- configurations, 2-2, 2-6, 2-10

- configuring adapter, 7-34

- error log entry, C-21

- hardware address, 7-5

- monitoring activity, 7-26

- MSCP server accesses shadow sets, 3-8

- multiple adapters, 2-10

- port, C-9

### LAN bridge

- failover, 2-13

- FORWARDING\_DELAY parameter, 2-13

- HELLO\_INTERVAL parameter, 2-13

- LISTEN\_TIME parameter, 2-13

- use of FDDI priority field, 2-14

### LAN segment



LAN segment (Cont.)  
 three in local area VAXcluster configuration, 2-12  
 two in local area VAXcluster configuration, 2-11

LAN Traffic Monitor (LTM), 7-26

LAVC\$FAILURE\_ANALYSIS.MAR sample program, E-3

LAVC\$START\_BUS.MAR sample program, E-1

LAVC\$STOP\_BUS.MAR sample program, E-2

%LAVC-I-ASUSPECT OPCOM message, E-9

%LAVC-S-WORKING OPCOM message, E-9

%LAVC-W-PSUSPECT OPCOM message, E-9

LBN (logical block number)  
 consistency across a shadow set, 3-6

LISTEN\_TIME parameter, 2-13

Lobe  
 VAXcluster, 2-7

Local area network (LAN)  
 See LAN

Local area VAXclusters, 2-2  
 alternate adapter booting, 7-32  
 configuration, 2-10, 2-11, 2-12  
 creating a network component list, F-5  
 creating a network component representation, F-4  
 creating cluster security database, 2-15  
 debugging satellite booting, 7-33, C-1, C-4  
 downline loading, 7-33, 7-34  
 LAN address for satellite, 7-33  
 large packet support for FDDI, 2-14  
 monitoring LAN activity, 7-26  
 network connections, D-1  
 network failure analysis, C-7, C-15, E-3  
 OPCOM messages, E-9  
 sample programs, E-1, E-2, E-3  
 satellite booting, 7-32, 7-33  
 starting network failure analysis, F-7  
 starting protocol on LAN adapter, E-1, F-1  
 stopping network failure analysis, F-7  
 stopping protocol on LAN adapter, E-2, F-3  
 subroutine package, F-1, F-3, F-4, F-5, F-7  
 surviving LAN bridge failover, 2-13  
 SYSGEN parameter LRPSIZE, 2-14  
 SYSGEN parameter NISCS\_CONV\_BOOT, 7-33

Local disk  
 setting up, 4-10

Logical block number  
 See LBN

Logical name  
 defining, 4-11  
 defining for NETPROXY.DAT, 4-13  
 defining for QMAN\$MASTER.DAT, 4-15, 6-2  
 defining for RIGHTSLIST.DAT, 4-13  
 defining for SYLOGIN.COM, 4-10  
 defining for SYSUAF.DAT, 4-13  
 defining for VMSMAIL\_PROFILE.DATA, 4-14

Logical name (Cont.)  
 system, 4-2

Login  
 controlling, 4-12

LRPSIZE, 2-14  
 parameter, A-1

## M

MAIL database  
 preparing common file, 4-14

Mail Utility (MAIL)  
 controlling, 4-12  
 preparing common database, 4-14

Manager, queue  
 see Queue manager

MASSBUS disk  
 dual-ported, 5-6

Messages  
 OPCOM messages, E-9

Mixed-interconnect cluster  
 See also local area VAXcluster  
 changing allocation class values on HSC subsystems, 7-23  
 creating cluster security database, 2-15  
 determining allocation class values, 7-4  
 monitoring LAN activity, 7-26  
 MSCP-served HSC disk, 2-4  
 MSCP server accesses shadow sets, 3-8

MODPARAMS.DAT file  
 created by CLUSTER\_CONFIG.COM, 7-2  
 specifying dump file, 7-37  
 specifying MSCP disk-serving parameters, 5-3  
 specifying TMSCP tape-serving parameter, 5-3  
 updating, 7-22

MOP server  
 enabling service, 7-34  
 functions, 2-2  
 selecting, 7-3

Mounting disks, 5-12

MSCP server, 1-4  
 and cluster-accessible disks, 5-2  
 initializing, 5-3  
 loading, 5-3  
 load sharing, 5-4

MSCP\_LOAD parameter, 5-3, A-1

MSCP\_SERVE\_ALL parameter, 5-3, A-1

Multiple-environment cluster, 4-1, 4-11

## N

NCP  
 copying node databases, 4-8  
 defining VAXcluster alias, 4-8  
 disabling LAN adapter, 4-7  
 enabling MOP service, 7-34, C-5  
 logging events, C-4  
 logging line counters, 7-26

NETCONFIG.COM command procedure  
 See DECnet-VAX network  
 NETNODE\_REMOTE.DAT  
 renaming to SYS\$COMMON directory, 4-7  
 NETNODE\_REMOTE.DAT file  
 sharing, 4-12  
 NETNODE\_UPDATE.COM command procedure,  
 7-11  
 NETPROXY.DAT file  
 creating common version, 4-12  
 defining logical name for, 4-13  
 setting up, 4-12  
 sharing, 4-12  
 Network  
 See DECnet-VAX network  
 Network Connections, D-1  
 Network Control Program (NCP)  
 See DECnet-VAX network  
 Network troubleshooting  
 See LAVC\$FAILURE\_ANALYSIS.MAR  
 NISCS\_CONV\_BOOT, 7-31, 7-33, A-1  
 NISCS\_LOAD\_PEA0, A-1, C-6  
 NISCS\_PORT\_SERV parameter, A-2

## O

OPA0: workstation operator console terminal,  
 7-12  
 OPCOM messages, 7-12, E-9  
 Operating system  
 coordinating files, 4-11  
 installing, 4-4  
 upgrading, 4-4

## P

Page file (PAGEFILE.SYS)  
 created by CLUSTER\_CONFIG.COM, 7-2, 7-3  
 Partitioning of cluster, 3-1, C-8  
 Password  
 See Security management  
 PEDRIVER  
 preferred channel, G-1  
 transmit channel selection, G-1  
 Port  
 select button, 5-4  
 software controllable selection, 5-5  
 Preferred channel, G-1  
 Preferred path  
 specification, 5-5  
 Printer queue, 6-3  
 assigning unique name to, 6-4  
 initializing, 6-4  
 sample configuration, 6-3  
 setting up, 6-3  
 starting, 6-4  
 Proxy login

Proxy login (Cont.)  
 controlling, 4-12  
 records, 4-12

## Q

QDSKINTERVAL parameter, A-2  
 QDSKVOTES parameter, 3-2, A-2  
 QMAN\$MASTER.DAT, 4-12, 6-2  
 Queue  
 batch  
 See Batch queue  
 command procedures, 4-11  
 common command procedure, 6-8  
 controlling, 1-2, 6-2  
 printer  
 See Printer queue  
 setting up, 4-11  
 single-computer and cluster, 6-1  
 Queue manager  
 availability, 6-1  
 clusterwide, 6-1  
 failover, 6-1  
 function, 6-1  
 nodes eligible to run, 6-1, 6-2  
 separation from job controller, 6-1  
 starting, 6-1  
 Quorum  
 adjusting EXPECTED\_VOTES, 7-22  
 equation, 3-1  
 EXPECTED\_VOTES parameter, 3-1, 7-22,  
 7-27  
 loss causes cluster hang condition, C-7  
 lowering value, 7-27  
 reasons for loss, C-7  
 restoring after unexpected computer failure,  
 7-26  
 VOTES parameter, 3-1  
 voting member, 3-1  
 adding, 7-2, 7-7, 7-22  
 removing, 7-2, 7-12, 7-22  
 QUORUM.DAT file, 3-3  
 Quorum disk, 3-2, 3-9  
 adding, 7-22  
 adjusting EXPECTED\_VOTES, 7-22  
 disabling, 7-2  
 enabling, 7-2  
 mounting, 3-3  
 removing, 7-22  
 Quorum disk watcher, 3-2  
 Quorum file (QUORUM.DAT), 3-3  
 Quorum scheme, 3-1

## R

- Rebooting a satellite with operating system
  - installed on local disk, 7-29
- RECNXINTERVAL, 2-13
  - parameter, A-2
- Reconfiguring the cluster, 7-22
- Recovery
  - repairing data, 3-6
- Remote Bridge Management Software (RBMS), 7-26
- Remote network node data
  - controlling, 4-12
- Remote node databases
  - copying, 4-8
- Removing a computer, 7-12, 7-22
  - adjusting EXPECTED\_VOTES, 7-22
  - shutting down before removing from cluster, 7-12
- Removing a satellite, 7-12
- Resource sharing, 3-1
- Restoring quorum, 7-26
- Restoring satellite configuration data, 7-11
- Restricted access disk, 5-1
- RIGHTSLIST.DAT file
  - defining logical name for, 4-13
  - merging, B-2
  - preparing common version of, 4-13
  - sharing, 4-12

## S

- Satellite
  - adding, 7-9
  - booting, 7-32
  - disabling conversational bootstrap operations, 7-31
  - failure to boot, 7-33, C-4
  - failure to join the cluster, C-6
  - functions, 2-3
  - local disk used for paging and swapping, 2-3
  - maintaining network configuration data, 7-11
  - modifying LAN hardware address, 7-13
  - obtaining LAN hardware address, 7-5
  - rebooting if operating system installed on local disk, 7-29
  - removing, 7-12
  - restoring network configuration data, 7-11
  - system files created by CLUSTER\_CONFIG.COM, 7-2
- SCSI
  - disks, 1-4, 3-7
- SCS SYSGEN parameters, A-3 to A-5
- Search list, 4-2
- Security management
  - controlling conversational bootstrap operations, 7-31
- Security management (Cont.)
  - modifying cluster group number, 7-30
  - modifying cluster password, 7-30
  - overview, 7-29
- Separation of job controller and queue manager, 6-1
- SET CLUSTER/EXPECTED\_VOTES command, 7-27
- Shadow set
  - components, 3-6
  - definition, 3-6
  - distributing, 3-7
  - in a VAXcluster system, 3-7
  - overview, 3-5
  - quorum disk, 3-9
- Shared command procedure files, 4-9
- Shared disk volume, 5-12
  - mounting, 5-12
- Shared file, 4-12
  - NETPROXY.DAT, 4-12
  - SYSUAF.DAT, 4-12
- Shared queues, 6-1
- Show Cluster Utility (SHOW CLUSTER), 7-26
  - CL\_QUORUM, 7-27
  - CL\_VOTES, 7-27
- Shutting down the cluster, 7-27
- Site-specific startup command file
  - elements, 4-11
- Standalone computer
  - converting to cluster computer, 7-19
- Star coupler, 1-3
- Star coupler expander (CISCE), 1-3
- Starting the queue manager, 6-1
- START/QUEUE/MANAGER command, 6-2
  - /ON qualifier, 6-1
- Startup
  - computer-specific function, 4-11
- Startup command file
  - coordinating, 4-9
  - creating common version, 4-10
  - site-specific elements, 4-11
- Startup procedure
  - failure to complete, C-6
- Swap file (SWAPFILE.SYS)
  - created by CLUSTER\_CONFIG.COM, 7-2, 7-3
- SYLOGIN.COM file
  - coordinating, 4-10
  - creating common version, 4-10, 4-11
  - defining logical name for, 4-10
- SYS\$LAVC\_DEFINE\_NET\_COMPONENT, F-4
- SYS\$LAVC\_DEFINE\_NET\_PATH, F-5
- SYS\$LAVC\_DISABLE\_ANALYSIS, F-7
- SYS\$LAVC\_ENABLE\_ANALYSIS, F-7
- SYS\$LAVC\_START\_BUS.MAR, F-1
- SYS\$LAVC\_STOP\_BUS, F-3
- SYS\$QUEUE\_MANAGER.QMAN\$JOURNAL, 4-12, 6-2

SYS\$QUEUE\_MANAGER.QMAN\$QUEUES,  
     4-12, 6-2  
 SYSBOOT.EXE image  
     renaming before rebooting satellite, 7-29  
 SYSGEN  
     cluster parameters, A-1 to A-3  
     EXPECTED\_VOTES, 3-1, 7-2, 7-7, 7-22  
     LRPSIZE parameter, 2-14  
     NISCS\_CONV\_BOOT parameter, 7-33  
     RECNXINTERVAL parameter, 2-13  
     SCS parameters, A-3 to A-5  
 SYSMAN Utility  
     enabling VAXcluster alias operations, 4-9  
     modifying cluster security data, 7-30  
 System command procedures  
     coordinating, 4-9  
 System Communications Services (SCS), 1-4,  
     C-9, C-10  
 System directory, 4-2  
 System disk  
     configuring in large cluster, 7-34, 7-37  
     creating duplicate, 7-20  
     directory structure, 4-2  
     moving high-activity files, 7-37  
     rebuilding, 5-13  
 System files, 4-11  
 SYSUAF.DAT file  
     creating common version, 4-12  
     defining logical name for, 4-13  
     merging, B-1  
     printing listing of, B-1  
     setting up, 4-12  
     sharing, 4-12

## T

---

Tape  
     See also Dual-pathed tape, Dual-ported tape  
     allocation class, 5-8  
     cluster-accessible, 5-1  
     DSA, 5-5  
     DSA controller, 5-2  
     DSSI, 5-1  
     dual-pathed, 5-1  
     HSC, 5-1, 5-2  
     local, 5-2  
     managing, 5-1  
     restricted access, 5-1  
     serving, 5-1  
     TMSCP-served, 5-1  
 Tape class driver, 1-5  
 TAPE\_ALLOCLASS, 5-8, A-2  
 TIMVCFail parameter, A-2  
 TMSCP server, 1-5  
     and cluster-accessible tapes, 5-2  
     initializing, 5-3  
     loading, 5-3  
 TMSCP\_LOAD parameter, 5-3, A-2

Transmit channel  
     selection, G-1  
 Troubleshooting  
     See LAVC\$FAILURE\_ANALYSIS.MAR  
     analyzing VAXport error log entries, C-15  
     CLUEXIT bugcheck, C-8  
     error log entries for CI and LAN ports, C-21  
     failure of computer to boot, 7-33, C-1  
     failure of computer to join the cluster, C-1,  
         C-6  
     failure of startup procedure to complete, C-6  
     hang condition, C-7  
     local area VAXcluster network failure analysis,  
         C-7, E-3  
     loss of quorum, C-7  
     network components, E-3  
     OPA0 error messages, C-28  
     repairing CI cables, C-14  
     shared resource is inaccessible, C-7  
     VAXport device problem, C-9  
     verifying CI cable connections, C-12  
     verifying CI port, C-11

## U

---

Unknown opcode error, C-25  
 Upgraded systems, 4-4  
 User accounts  
     comparing, B-1  
     coordinating, 4-12, B-1  
     group UIC, B-1  
 User environment  
     common-environment cluster, 4-1  
     creating common-environment cluster, 4-10  
     defining, 4-12  
     multiple-environment cluster, 4-1  
 User Environment Test Package (UETP)  
     creating command procedure to run, 7-40  
     running in large cluster, 7-39  
     specifying values for LOAD phase, 7-40  
 User identification code (UIC), B-1

## V

---

VAXcluster alias  
     defining, 4-6, 7-42  
     enabling operations, 4-9  
 VAXcluster environment  
     queue manager in, 6-1  
 VAXcluster sample program, E-1  
     LAVC\$FAILURE\_ANALYSIS.MAR, E-3  
     LAVC\$START\_BUS.MAR, E-1  
     LAVC\$STOP\_BUS.MAR, E-2  
 VAXport communication, C-9  
 VAXport driver, 1-4, C-9  
 VAXVMSSYS.PAR file  
     created by CLUSTER\_CONFIG.COM, 7-2  
 Virtual circuit, C-9

VMSMAIL\_PROFILE.DATA  
  defining logical name for, 4-14  
VMSMAIL\_PROFILE.DATA file  
  preparing common version of, 4-14  
  sharing, 4-12  
VMS operating system  
  installing license, 4-4  
VMS RMS distributed file system, 1-4  
Volume label  
  modifying for satellite's local disk, 7-3  
Volume shadowing  
  across clusters, 3-7  
  concepts, 3-5  
  distributing shadow sets, 3-7  
  for system disks, 7-35

  in mixed-interconnect cluster, 7-36  
  interprocessor communication, 3-8  
  overview, 3-5  
  phase I, 3-7  
  phase II, 3-7  
  shadow sets, 2-7  
  virtual units, 3-8  
VOTES parameter, 3-1, A-3  
Voting member, 3-1  
  adding, 7-2, 7-7, 7-22  
  removing, 7-2, 7-12, 7-22

## W

---

Workload balancing, 1-1, 6-1





## How to Order Additional Documentation

---

### Technical Support

If you need help deciding which documentation best meets your needs, call 800-343-4040 before placing your electronic, telephone, or direct mail order.

### Electronic Orders

To place an order at the Electronic Store, dial 800-DEC-DEMO (800-332-3366) using a 1200- or 2400-baud modem. If you need assistance using the Electronic Store, call 800-DIGITAL (800-344-4825).





### Telephone and Direct Mail Orders

<b>Your Location</b>	<b>Call</b>	<b>Contact</b>
Continental USA, Alaska, or Hawaii	800-DIGITAL	Digital Equipment Corporation P.O. Box CS2008 Nashua, New Hampshire 03061
Puerto Rico	809-754-7575	Local Digital subsidiary
Canada	800-267-6215	Digital Equipment of Canada Attn: DECdirect Operations KAO2/2 P.O. Box 13000 100 Herzberg Road Kanata, Ontario, Canada K2K 2A6
International	_____	Local Digital subsidiary or approved distributor
Internal <sup>1</sup>	_____	USASSB Order Processing - WMO/E15 <i>or</i> U.S. Area Software Supply Business Digital Equipment Corporation Westminster, Massachusetts 01473

---

<sup>1</sup>For internal orders, you must submit an Internal Software Order Form (EN-01740-07).







# Reader's Comments

VMS VAXcluster Manual

AA-LA27C-TE

Please use this postage-paid form to comment on this manual. If you require a written reply to a software problem and are eligible to receive one under Software Performance Report (SPR) service, submit your comments on an SPR form.

Thank you for your assistance.

<b>I rate this manual's:</b>	<b>Excellent</b>	<b>Good</b>	<b>Fair</b>	<b>Poor</b>
Accuracy (software works as manual says)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Completeness (enough information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clarity (easy to understand)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Organization (structure of subject matter)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Figures (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Examples (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Index (ability to find topic)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Page layout (easy to find information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

I would like to see more/less \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

What I like best about this manual is \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

What I like least about this manual is \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

I found the following errors in this manual:

Page	Description
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

Additional comments or suggestions to improve this manual:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

I am using **Version** \_\_\_\_\_ of the software this manual describes.

Name/Title \_\_\_\_\_ Dept. \_\_\_\_\_  
Company \_\_\_\_\_ Date \_\_\_\_\_  
Mailing Address \_\_\_\_\_  
\_\_\_\_\_ Phone \_\_\_\_\_

--- Do Not Tear - Fold Here and Tape ---

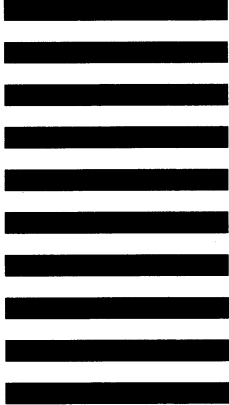


No Postage  
Necessary  
if Mailed  
in the  
United States

**BUSINESS REPLY MAIL**  
FIRST CLASS PERMIT NO. 33 MAYNARD MASS.

POSTAGE WILL BE PAID BY ADDRESSEE

DIGITAL EQUIPMENT CORPORATION  
Corporate User Information Products  
ZKO1-3/J35  
110 SPIT BROOK RD  
NASHUA, NH 03062-9987



--- Do Not Tear - Fold Here ---



# Reader's Comments

VMS VAXcluster Manual

AA-LA27C-TE

Please use this postage-paid form to comment on this manual. If you require a written reply to a software problem and are eligible to receive one under Software Performance Report (SPR) service, submit your comments on an SPR form.

Thank you for your assistance.

**I rate this manual's:**

	Excellent	Good	Fair	Poor
Accuracy (software works as manual says)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Completeness (enough information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clarity (easy to understand)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Organization (structure of subject matter)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Figures (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Examples (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Index (ability to find topic)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Page layout (easy to find information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

I would like to see more/less \_\_\_\_\_

What I like best about this manual is \_\_\_\_\_

What I like least about this manual is \_\_\_\_\_

I found the following errors in this manual:

Page	Description
_____	_____
_____	_____
_____	_____
_____	_____

Additional comments or suggestions to improve this manual:

I am using **Version** \_\_\_\_\_ of the software this manual describes.

Name/Title \_\_\_\_\_ Dept. \_\_\_\_\_

Company \_\_\_\_\_ Date \_\_\_\_\_

Mailing Address \_\_\_\_\_

Phone \_\_\_\_\_

--- Do Not Tear - Fold Here and Tape ---

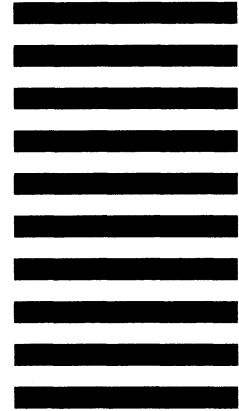


No Postage  
Necessary  
if Mailed  
in the  
United States

**BUSINESS REPLY MAIL**  
FIRST CLASS PERMIT NO. 33 MAYNARD MASS.

POSTAGE WILL BE PAID BY ADDRESSEE

DIGITAL EQUIPMENT CORPORATION  
Corporate User Information Products  
ZKO1-3/J35  
110 SPIT BROOK RD  
NASHUA, NH 03062-9987



--- Do Not Tear - Fold Here ---



# VMS Exchange Utility Manual

Order Number: AA-LA28A-TE

**April 1988**

This document describes the functions of the Exchange Utility for use on VAX processors.

**Revision/Update Information:** This document supersedes the *VAX/VMS Exchange Utility Reference Manual, Version 4.2.*

**Software Version:** VMS Version 5.0

**digital equipment corporation  
maynard, massachusetts**

---

**April 1988**

The information in this document is subject to change without notice and should not be construed as a commitment by Digital Equipment Corporation. Digital Equipment Corporation assumes no responsibility for any errors that may appear in this document.

The software described in this document is furnished under a license and may be used or copied only in accordance with the terms of such license.

No responsibility is assumed for the use or reliability of software on equipment that is not supplied by Digital Equipment Corporation or its affiliated companies.

---

Copyright ©1988 by Digital Equipment Corporation

All Rights Reserved.  
Printed in U.S.A.

---

The postpaid READER'S COMMENTS form on the last page of this document requests the user's critical evaluation to assist in preparing future documentation.

The following are trademarks of Digital Equipment Corporation:

DEC	DIBOL	UNIBUS
DEC/CMS	EduSystem	VAX
DEC/MMS	IAS	VAXcluster
DECnet	MASSBUS	VMS
DECsystem-10	PDP	VT
DECSYSTEM-20	PDT	
DECUS	RSTS	
DECwriter	RSX	<b>digital</b> ™

ZK4543

---

**HOW TO ORDER ADDITIONAL DOCUMENTATION  
DIRECT MAIL ORDERS**

**USA & PUERTO RICO\***

Digital Equipment Corporation  
P.O. Box CS2008  
Nashua, New Hampshire  
03061

**CANADA**

Digital Equipment  
of Canada Ltd.  
100 Herzberg Road  
Kanata, Ontario K2K 2A6  
Attn: Direct Order Desk

**INTERNATIONAL**

Digital Equipment Corporation  
PSG Business Manager  
c/o Digital's local subsidiary  
or approved distributor

In Continental USA and Puerto Rico call 800-258-1710.

In New Hampshire, Alaska, and Hawaii call 603-884-6660.

In Canada call 800-267-6215.

\* Any prepaid order from Puerto Rico must be placed with the local Digital subsidiary (809-754-7575).

International orders should be placed through the Software Distribution Center (SDC), Digital Equipment Corporation, Westminister, Massachusetts 01473.

---

---

## Production Note

This book was produced with the VAX DOCUMENT electronic publishing system, a software tool developed and sold by DIGITAL. In this system, writers use an ASCII text editor to create source files containing text and English-like code; this code labels the structural elements of the document, such as chapters, paragraphs, and tables. The VAX DOCUMENT software, which runs on the VMS operating system, interprets the code to format the text, generate a table of contents and index, and paginate the entire document. Writers can print the document on the terminal or line printer, or they can use DIGITAL-supported devices, such as the LN03 laser printer and PostScript<sup>®</sup> printers (PrintServer 40 or LN03R ScriptPrinter), to produce a typeset-quality copy containing integrated graphics.





---

# Contents

---

<b>PREFACE</b>	<b>vii</b>
<b>EXCHANGE Description</b>	<b>EXCH-1</b>
<b>1 THE EXCHANGE MEDIA</b>	<b>EXCH-1</b>
<b>2 THE EXCHANGE COMMAND STRING</b>	<b>EXCH-2</b>
<b>3 THE /[NO]MESSAGE QUALIFIER</b>	<b>EXCH-2</b>
<b>4 VOLUME FORMAT SPECIFICATION</b>	<b>EXCH-3</b>
<b>4.1 Files-11 Volume Format</b>	<b>EXCH-3</b>
4.1.1 Files-11 Device Specifications • EXCH-4	
4.1.2 Files-11 File Specifications • EXCH-4	
<b>4.2 RT-11 Volume Format</b>	<b>EXCH-4</b>
4.2.1 RT-11 Device Specifications • EXCH-4	
4.2.2 RT-11 File Specifications • EXCH-4	
<b>4.3 DOS-11 Volume Format</b>	<b>EXCH-5</b>
4.3.1 DOS-11 Device Specifications • EXCH-5	
4.3.2 DOS-11 File Specifications • EXCH-5	
4.3.3 DOS-11 Performance Optimization • EXCH-6	
<b>4.4 Volume Format Defaults</b>	<b>EXCH-6</b>
<b>5 RECORD FORMAT AND ATTRIBUTE SPECIFICATION</b>	<b>EXCH-6</b>
<b>5.1 Record Format and Attribute Options</b>	<b>EXCH-7</b>
5.1.1 BINARY Option • EXCH-7	
5.1.2 FIXED Option • EXCH-8	
5.1.3 STREAM Option • EXCH-8	
5.1.4 VARIABLE Option • EXCH-8	
5.1.5 PAD Option • EXCH-9	
<b>5.2 Record Format Defaults</b>	<b>EXCH-9</b>
<b>5.3 Carriage Control Options</b>	<b>EXCH-10</b>
<b>6 VIRTUAL DEVICES</b>	<b>EXCH-11</b>

---

# Contents

---

**EXCHANGE Usage Summary** **EXCH-12**

---

**EXCHANGE Commands** **EXCH-14**

<b>COPY</b>	<b>EXCH-15</b>
<b>DELETE</b>	<b>EXCH-23</b>
<b>DIRECTORY</b>	<b>EXCH-25</b>
<b>DISMOUNT</b>	<b>EXCH-28</b>
<b>EXIT</b>	<b>EXCH-29</b>
<b>HELP</b>	<b>EXCH-30</b>
<b>INITIALIZE</b>	<b>EXCH-31</b>
<b>MOUNT</b>	<b>EXCH-35</b>
<b>RENAME</b>	<b>EXCH-38</b>
<b>SHOW</b>	<b>EXCH-40</b>
<b>TYPE</b>	<b>EXCH-41</b>

---

## INDEX

---

## TABLES

<b>EXCH-1</b>	<b>EXCHANGE Commands</b>	<b>EXCH-2</b>
<b>EXCH-2</b>	<b>Record Format and Attribute Options</b>	<b>EXCH-7</b>
<b>EXCH-3</b>	<b>Pad Characters</b>	<b>EXCH-9</b>
<b>EXCH-4</b>	<b>Implied Output Record Formats</b>	<b>EXCH-10</b>

---

# Preface

---

## Intended Audience

This document is intended for users who transfer data between VMS volumes and foreign volumes.

---

## Document Structure

This document is composed of the following three sections:

- Description—Provides a full description of the Exchange Utility (EXCHANGE). This section shows the DCL command that invokes EXCHANGE, describes how to direct output from EXCHANGE, and details any applicable restrictions on EXCHANGE use.
- Usage Summary—Outlines the following EXCHANGE information:
  - Invoking the utility
  - Using the utility interactively and from the DCL command line
  - Exiting the utility
  - Directing output
  - Restrictions or privileges required
- Commands—Describes EXCHANGE commands, including format, parameters, qualifiers, and examples.

---

## Associated Documents

The *Guide to VMS Files and Devices* provides task-oriented instructions for performing EXCHANGE operations on private volumes.

The *Guide to Maintaining a VMS System* includes system management tasks that use the Exchange Utility.

The *VAXcluster Operations Manual* includes examples that show the use of EXCHANGE in a VAXcluster environment.

The *VMS System Messages and Recovery Procedures Reference Manual* explains error messages and, in most cases, suggests recovery procedures.

---

## Conventions

Convention	Meaning
<code>RET</code>	In examples, a key name (usually abbreviated) shown within a box indicates that you press a key on the keyboard; in text, a key name is not enclosed in a box. In this example, the key is the RETURN key. (Note that the RETURN key is not usually shown in syntax statements or in all examples; however, assume that you must press the RETURN key after entering a command or responding to a prompt.)
<code>CTRL/C</code>	A key combination, shown in uppercase with a slash separating two key names, indicates that you hold down the first key while you press the second key. For example, the key combination CTRL/C indicates that you hold down the key labeled CTRL while you press the key labeled C. In examples, a key combination is enclosed in a box.
<code>\$ SHOW TIME</code> <code>05-JUN-1988 11:55:22</code>	In examples, system output (what the system displays) is shown in black. User input (what you enter) is shown in red.
<code>\$ TYPE MYFILE.DAT</code> . . .	In examples, a vertical series of periods, or ellipsis, means either that not all the data that the system would display in response to a command is shown or that not all the data a user would enter is shown.
<code>input-file, . . .</code>	In examples, a horizontal ellipsis indicates that additional parameters, values, or other information can be entered, that preceding items can be repeated one or more times, or that optional arguments in a statement have been omitted.
<code>[logical-name]</code>	Brackets indicate that the enclosed item is optional. (Brackets are not, however, optional in the syntax of a directory name in a file specification or in the syntax of a substring specification in an assignment statement.)
quotation marks apostrophes	The term quotation marks is used to refer to double quotation marks ("). The term apostrophe (') is used to refer to a single quotation mark.

---

---

## EXCHANGE Description

The Exchange Utility (EXCHANGE) enables you to transfer files between some foreign format volumes and VMS native volumes. It converts the format of the files when transferring files between volumes of different structures. EXCHANGE recognizes all Files-11 volumes on VMS devices, as well as DOS-11 magnetic tape volumes and RT-11 block-addressable volumes.

You can use EXCHANGE commands to perform the following tasks:

- To locate bad blocks on volumes
- To list directories of volumes
- To transfer files to and from volumes
- To delete and rename files for block-addressable devices, such as RT-11 disks
- To create foreign volumes
- To manipulate virtual devices, Files-11 files that are images of foreign volumes

EXCHANGE employs defaults to ensure that the volume formats and file structures are compatible with the type of operation you want to perform. However, you can override these defaults by using volume format and record format qualifiers.

The following sections describe EXCHANGE media, commands, qualifiers, and defaults.

1

---

### The EXCHANGE Media

Use EXCHANGE to perform file transfers and format conversions on the following devices:

- DOS-11 magnetic tape volumes
- Files-11 volumes
- RT-11 block-addressable volumes

EXCHANGE recognizes RT-11 volumes on any VMS block-addressable device. However, RT-11 supports only some of the devices that are recognized by EXCHANGE. Also, RT-11 imposes a size limit of 65,535 blocks for a volume. If an RT-11 volume is created on a device that exceeds this limit, the size of the RT-11 volume is set to the maximum. Therefore, blocks 65,536 to the end of the volume cannot be used.

You can use EXCHANGE to manipulate Files-11 files that are images of foreign volumes; these files are called *virtual devices*. A virtual device volume is contained in a VAX Record Management Services (RMS) file on a VMS disk. EXCHANGE contains commands to copy a block-addressable volume both to and from the RMS virtual device file. The virtual device file can then be manipulated as if it were a block-addressable device.

# EXCHANGE Description

Files-11 Structure Level 1 or Structure Level 2 volumes are the default for disk volumes initialized by the DCL command INITIALIZE. DOS-11 and RT-11 volumes are initialized using the EXCHANGE command INITIALIZE. Since the formats of these volumes are not recognized by VMS, the volumes must be mounted as foreign volumes. To mount a foreign volume, use the DCL command MOUNT with the /FOREIGN or /NOLABEL qualifier. Refer to the *VMS Mount Utility Manual* for more information on the Mount Utility; refer to the *VMS DCL Dictionary* for more information on the DCL command INITIALIZE.

## 2 The EXCHANGE Command String

Most EXCHANGE command strings are similar to DCL commands for native volumes. The form for the EXCHANGE command string is as follows:

```
$ EXCHANGE command file-spec[, ... ] [file-spec]
EXCHANGE> command file-spec[, ... ] [file-spec]
```

You can enter only one EXCHANGE command at the DCL prompt. Use the second file-spec parameter only for EXCHANGE commands specifying data transfers or requiring additional parameters.

The command indicates the action you want EXCHANGE to perform. Each EXCHANGE command qualifier is the same as the DCL command qualifier that performs a similar function on native volumes. If no corresponding DCL command qualifier exists, the EXCHANGE command qualifier is taken from the command on the foreign operating system. Table EXCH-1 lists the EXCHANGE commands and their functions.

**Table EXCH-1 EXCHANGE Commands**

Command	Function
COPY	Copy files from input to output spec
DELETE	Delete files from block-addressable volumes
DIRECTORY	List names of files on volume
DISMOUNT	Dismount a volume
EXIT	Return to the DCL command level
HELP	Obtain help on EXCHANGE commands and features
INITIALIZE	Erase contents, if any, and create an empty volume
MOUNT	Mount a volume on a device (foreign only)
RENAME	Rename files on block-addressable volumes
SHOW	Show volumes known to EXCHANGE
TYPE	Display files on the default output device

## 3 The /[NO]MESSAGE Qualifier

Use the EXCHANGE command line qualifier /[NO]MESSAGE to control whether or not EXCHANGE displays messages related to the MOUNT, DISMOUNT, and INITIALIZE commands. The default is /MESSAGE, which

remains in effect throughout the EXCHANGE session unless you explicitly change it. You can change the default by specifying /NOMESSAGE, as follows:

```
$ EXCHANGE/NOMESSAGE  
EXCHANGE>
```

---

## 4 Volume Format Specification

EXCHANGE allows you to define the physical format of the volumes to be processed. Specify the volume format by using the qualifier /VOLUME\_FORMAT=option. If you do not define a volume format, EXCHANGE uses defaults.

The qualifier /VOLUME\_FORMAT=option determines which operations are allowed on the volume, and the option you specify often implies certain defaults. In addition, the volume format qualifier determines the format of the file name and directory specifications.

The /VOLUME\_FORMAT qualifier has three options. *FILES11* identifies a Files-11 formatted native volume. *RT11* identifies an RT-11 formatted block-addressable volume. *DOS11* identifies a DOS-11 formatted magnetic tape volume.

To use the volume format qualifiers, you must attach them to one or both of the file specification parameters; you cannot attach them directly to the command.

You can specify multiple volumes of different volume types, as in the following example:

```
EXCHANGE> COPY DMA1:*.FOR/VOLUME=RT11,MTA0:[11,132]*.FOR/VOLUME=DOS11 *
```

The following sections describe volume format options and explain EXCHANGE device and file specifications.

---

### 4.1 Files-11 Volume Format

Files-11 is the disk volume format maintained by the VMS operating system. Although the VMS operating system fully supports the volume formats of ANSI magnetic tape and Files-11 Structure Levels 1 and 2, EXCHANGE uses VAX Record Management Services (RMS) for operations on these file structures.

To specify a Files-11 volume, or a file on a Files-11 volume, use a standard VMS file specification. For a description of VMS file specifications, refer to the *VMS DCL Dictionary*. The following two sections provide additional information on the use of EXCHANGE with Files-11 volumes and files.

# EXCHANGE Description

---

## 4.1.1 Files-11 Device Specifications

Use the FILES11 option when referencing any VMS device, even though the Files-11 structure specifically refers to directory-structured devices such as disks, diskettes, and TU58 tape cassettes. This usage is consistent with VAX RMS, which allows unit-record devices (such as terminals) and magnetic tapes to be accessed in a device-independent manner.

Files-11 is the standard disk format for the VMS operating system. You only need to specify the /FOREIGN qualifier when mounting volumes that have a format other than Files-11.

If you specify FILES11 when referencing a magnetic tape device, an ANSI-formatted magnetic tape device is implied. The tape must have been mounted as an ANSI volume. For more information on mounting tape devices, refer to the *VMS Mount Utility Manual*.

---

## 4.1.2 Files-11 File Specifications

There are no restrictions on the use of logical names in Files-11 file specifications. If you omit device or directory specifications from a Files-11 specification, EXCHANGE uses the current defaults for your process.

You can use wildcard characters for both input and output file specifications. You should not, however, specify version numbers as wildcards on foreign volumes that do not support version numbers (such as RT-11 or DOS-11 volumes). If you use wildcards in any field of the output file specification, EXCHANGE takes the corresponding name from the input file.

---

## 4.2 RT-11 Volume Format

The RT11 option allows you to read and write block-addressable volumes in the format used by DIGITAL's RT-11 operating system. EXCHANGE permits you to use devices other than those supported by RT-11.

The following two sections describe how you can use EXCHANGE to specify RT-11 volumes and files.

---

### 4.2.1 RT-11 Device Specifications

The device name you specify for the RT-11 volume can be either a standard VMS device name in the form **ddcu:** or a logical name that translates to a VMS device name. This device must have been mounted with the /FOREIGN qualifier.

---

### 4.2.2 RT-11 File Specifications

You can specify an RT-11 file name that consists of any uppercase letters from A through Z and any digits from 0 through 9 (RADIX-50 characters). The file name field for RT-11 is six characters; the file type field is three characters. The RT-11 file specification has no directory field because RT-11 volumes do not have user directories.

If you copy a file with a name that is not valid under RT-11 to an RT-11 volume, the RT-11 file name is produced from the first six valid characters of the name. The RT-11 file type is extracted from the first three valid characters of the input file type (for input names that are similar to Files-11 names). Lowercase letters are converted to uppercase before EXCHANGE checks for a valid file name.



For example:

```
EXCHANGE> COPY/LOG SYS_CREATE_PROCESS.MACRO_32 DMAO:/VOLUME=RT11
           %EXCHANGE-S-COPNEWNAME, DEVDS:[SYS.SRC]SYS_CREATE_PROCESS.MACRO_32;34
           copied to DMAO:SYSCRE.MAC, 937 records
```

```
EXCHANGE> COPY/LOG MTAO:"ANSI (A) TAPENAME" DMAO:/VOLUME=RT11
           %EXCHANGE-S-COPNEWNAME, MTAO:[]"ANSI (A) TAPENAME".;1
           copied to DMAO:ANSIAT., 321 records
```

EXCHANGE allows you to rename a file on a copy operation, so you can specify any valid file name for the output file.

## 4.3 DOS-11 Volume Format

Use the DOS11 option with the /VOLUME\_FORMAT qualifier to specify that a magnetic tape is in the format defined by DIGITAL's early PDP-11 Disk Operating System, DOS-11.

The following two sections describe how you can specify DOS-11 devices and files using EXCHANGE.

### 4.3.1 DOS-11 Device Specifications

You can use the DOS-11 format only on magnetic tape devices. Therefore, a VMS device specification for a 9-track magnetic tape drive must be present in order for you to use the DOS11 option.

### 4.3.2 DOS-11 File Specifications

DOS-11 volumes accept directories only in the user identification code (UIC) format, such as [310,22]. The two numbers are octal and must be in the range 0 through 377. When you specify the directory in an input file specification, you can use a wildcard character to indicate either number.

The EXCHANGE command DIRECTORY/OWNER uses a default directory of [\*,\*], so that all the files on the tape are listed regardless of the DOS-11 directory. For all other commands, EXCHANGE uses your current default directory, as long as it is in UIC format. Otherwise, your process UIC becomes a directory. If a file that appears in an EXCHANGE command DIRECTORY is not found by another EXCHANGE command, enter the DIRECTORY/OWNER command for that file. Then enter the other EXCHANGE command again, specifying the same directory displayed for the file by the DIRECTORY/OWNER command.

The DOS-11 character set for file names is the same as the character set for RT-11: the uppercase letters from A through Z and the digits 0 through 9. The DOS-11 file name field is nine characters long; the file type field is three characters.

If you copy a file with a name that is invalid under DOS-11 to a DOS-11 magnetic tape, the file name on the magnetic tape is formed from the first nine valid characters of the input file name. The DOS-11 file type is extracted from the first three valid characters of the input file type (for input names that are similar to Files-11 names). Lowercase letters are converted to uppercase before EXCHANGE checks for a valid file name.

# EXCHANGE Description

For example:

```
EXCHANGE> COPY/LOG SYS_CREATE_PROCESS.MACRO_32 MFAO:/VOLUME=DOS11
           %EXCHANGE-S-COPNEWNAME, DEVDS:[SYS.SRC]SYS_CREATE_PROCESS.MACRO_32;34
           copied to MFAO:SYSCREATE.MAC, 937 records
```

```
EXCHANGE> COPY/LOG MTAO:"ANSI (A) TAPENAME" MFAO:/VOLUME=DOS11
           %EXCHANGE-S-COPNEWNAME, MTAO:["ANSI (A) TAPENAME".;1
           copied to MFAO:ANSIATAPE., 321 records
```

---

### 4.3.3 DOS-11 Performance Optimization

When you specify the COPY or TYPE command with the /REWIND qualifier, a DOS-11 tape device does not rewind to the beginning of tape.

DOS-11 maintains an in-memory list of the files it has read on the tape. The /REWIND qualifier directs DOS-11 to read the list of files from memory and to rewind only to the beginning of the file specified in the COPY or TYPE command. This is called a logical rewind operation.

---

## 4.4 Volume Format Defaults

EXCHANGE uses the following default volume formats if you do not specify the /VOLUME\_FORMAT qualifier:

- If a volume is mounted as a native volume (without the /FOREIGN or /NOLABEL qualifier), the default is /VOLUME\_FORMAT=FILES11. This default applies to both block-addressable and magnetic tape devices.
- If the volume is block-addressable (disks, diskettes, and TU58 tape cassettes) and mounted with the /FOREIGN qualifier, EXCHANGE defaults to /VOLUME\_FORMAT=RT11.
- If the volume is a 9-track magnetic tape mounted with the /FOREIGN qualifier, the default is /VOLUME\_FORMAT=DOS11.
- If no device is given, EXCHANGE defaults to the current default device. For example:

```
$ MOUNT/FOREIGN DMA0:
$ SET DEFAULT DMA0:
```

Since the DCL command SET DEFAULT assigns the RK07 disk named DMA0 as the default device, any subsequent EXCHANGE command defaults to DMA0 for a file reference that omits a device specification. EXCHANGE assigns DMA0 the default format option RT11 because the volume is mounted with the /FOREIGN qualifier.

---

## 5 Record Format and Attribute Specification

EXCHANGE allows you to specify the internal record structure and related attributes of a file. Define the record structure of a file by using the qualifier /RECORD\_FORMAT=option. In addition to the record structure options, the /RECORD\_FORMAT qualifier includes options you can use to define other related attributes of the records. The qualifier /CARRIAGE\_CONTROL=option allows you to define the carriage control attributes.

The following sections describe how you can use the /RECORD\_FORMAT and /CARRIAGE\_CONTROL qualifiers.

## 5.1 Record Format and Attribute Options

When you specify an input file, you can use the /RECORD\_FORMAT qualifier to give EXCHANGE additional information to use in reading the file. You need to specify this information when a foreign file structure is not entirely self-descriptive, or when a special interpretation of the file structure is needed.

To define the internal record structure of a file, the /RECORD\_FORMAT qualifier has four options: BINARY, FIXED, STREAM, and VARIABLE. In addition, there are two options to define related file attributes: FIXED=n and PAD=x. Table EXCH-2 summarizes the record format and attribute options.

**Table EXCH-2 Record Format and Attribute Options**

Option	Function
BINARY	Identifies PDP-11 formatted binary record format
FIXED	Identifies fixed-length record format; sets fixed-length record size
STREAM	Identifies ASCII stream record format
VARIABLE	Identifies variable-length record format
PAD	Sets the character to pad short records on fixed-length output, valid only with the FIXED option

You can specify only one of the four record format options at a time. The FIXED=n and PAD=n file attribute options are valid only with the record format FIXED.

You are not required to include input record format qualifiers on Files-11 input, since Files-11 stores file attributes that allow EXCHANGE to read the file correctly. If a Files-11 file is not in a suitable format, you can use the Convert Utility (CONVERT) to change the file to the desired format. For a detailed description of CONVERT, refer to the *VMS Convert and Convert/Reclaim Utility Manual*.

When you specify an output file, use the /RECORD\_FORMAT qualifier to specify the characteristics the file should have after it is copied. Specific options are available with the /RECORD\_FORMAT and /CARRIAGE\_CONTROL qualifiers to provide commonly needed attributes.

The following sections describe the record format and attribute options.

### 5.1.1 BINARY Option

The BINARY record format is valid with DOS-11 or RT-11 volumes. BINARY files (also known as formatted binary) have variable-length records with headers and checksums added to each record.

The BINARY option is not valid for Files-11 output. On conversions from DOS-11 or RT-11 BINARY input to Files-11 output, EXCHANGE normally generates variable-length records. The output record size equals the input record size.

# EXCHANGE Description

---

## 5.1.2 **FIXED Option**

The **FIXED** option directs **EXCHANGE** to read or write fixed-length records only. The file contains no control information; since the records are all the same length, they are easily located.

You can use the **FIXED** option to indicate the desired number of bytes in the record. If you do not specify an explicit record length, **EXCHANGE** assumes a record length of 512 bytes.

If any input records are longer than the specified length, **EXCHANGE** truncates them. If input records are shorter than the output record, **EXCHANGE** extends them with the **PAD** character. The default **PAD** character is the null character. To specify an alternate **PAD** character, use the **PAD** option.

As an input option, **FIXED** is valid only for non-Files-11 **FIXED** transfers. It is not a valid option for **STREAM**, **VARIABLE**, or **BINARY** transfers because these operations determine the record length by delimiters or control bytes.

---

## 5.1.3 **STREAM Option**

Stream ASCII files, also known as formatted ASCII files, consist of ASCII data records terminated by carriage return/line feed (**CR-LF**), form feed (**FF**), or vertical tab (**VT**) characters.

In transfers from **DOS-11** or **RT-11** **STREAM** files to **Files-11** **VARIABLE** files, **CR-LF** pairs are removed from the ends of records. In transfers from **Files-11** **VARIABLE** files to **DOS-11** or **RT-11** **STREAM** files, **CR-LF** pairs are added to the end of each record that does not already end with **LF** or **FF**. **EXCHANGE** removes all null (**NUL**), delete (**DEL**), and vertical tab (**VT**) characters from input records in any of these transfers. ASCII data is transferred as 7-bit codes. Bit 7 (sometimes used as a parity bit) of each byte is set to 0 before the transfer.

If the **RT-11** or **DOS-11** file is input or output as an ASCII text file with fixed-length records, use the **FIXED** option rather than the **STREAM** option. An ASCII stream file with records of the same length is not the same as a file of fixed-length records. **STREAM** files have records delimited by control characters. **FIXED** files have no delimiters; **EXCHANGE** locates **FIXED** file records by counting bytes.

When you perform conversions from **DOS-11** or **RT-11** to **Files-11**, **EXCHANGE** normally generates variable-length records (**/RECORD\_FORMAT=VARIABLE**). The output record size equals the input record size.

---

## 5.1.4 **VARIABLE Option**

The variable-length record format is the most common type of **Files-11** format. The file uses internal control information to separate the records, so that any sequence of bytes (even **CR-LF**) can be contained in the data. By default, **EXCHANGE** generates **Files-11** files with **VARIABLE** records unless the input file is **FIXED**.

You cannot specify **VARIABLE** for **RT-11** and **DOS-11** output.

## 5.1.5 PAD Option

Use the PAD option to specify which character will be used to pad input records that are shorter than the fixed-length output record. By default, EXCHANGE pads with the NULL character (value 0). Table EXCH-3 shows which characters you can specify with the PAD option.

**Table EXCH-3 Pad Characters**

Option	Character
PAD	NULL (0 Decimal)
PAD=c	Where c is ASCII A through Z (uppercase), or 0 through 9
PAD="c"	Where c is any other ASCII character
PAD=%bx	Where x is a number representing an ASCII character and b represents the base of the number. Values for b: D = decimal O = octal X = hexadecimal

## 5.2 Record Format Defaults

EXCHANGE assumes defaults for the record formats if you do not specify an explicit /RECORD\_FORMAT option. The defaults are dependent on a number of factors and allow you to perform most transfer operations without using the /RECORD\_FORMAT qualifier.

Files-11 input files use a default /RECORD\_FORMAT option of FIXED, STREAM, or VARIABLE, depending on the actual record format of the file. If a Files-11 file is not in a suitable format, you can use the Convert Utility (CONVERT) to change the file format. For a detailed description of CONVERT, refer to the *VMS Convert and Convert/Reclaim Utility Manual*.

By default, EXCHANGE generates Files-11 output files with sequential organization. However, if the input file is in BINARY, STREAM, or VARIABLE format, the default Files-11 output record format is VARIABLE. If the input file has FIXED format, the default output format is FIXED.

For DOS-11 and RT-11, EXCHANGE assumes default record formats for a file based on the type field of the file name. If the file type is OBJ, STB, BIN, or LDA, the default record format is BINARY. If the file type is EXE, %LB (% is a wildcard character representing a single character), SAV, SML, SYS or TSK, the default record format is FIXED=512. For all other types of input files, the default record format is STREAM. The default for other types of output files depends on the record format of the input file.

EXCHANGE applies these defaults only for transfers between an RT-11 or DOS-11 volume and a Files-11 volume. When you perform a transfer between an RT-11 or DOS-11 volume and another RT-11 or DOS-11 volume, the file is copied block by block; no reformatting occurs unless you request a conversion with an explicit /RECORD\_FORMAT option.

In most cases, EXCHANGE applies a record format default to the input file and lets the resulting input record format imply the output record format. However, if any of the file types previously listed is included in an output specification, the default implied by the input record format is overridden.

# EXCHANGE Description

Table EXCH-4 shows how the output record formats are implied by different input record and volume formats. The input format can be defaulted, read from the file (for Files-11), implied by the type field of the input file, or explicitly stated. Note that the default output formats are overridden by record formats implied by the type field of the output files.

**Table EXCH-4 Implied Output Record Formats**

<b>Input Volume</b>	<b>Input Record</b>	<b>Output Volume</b>	<b>Output Record</b>
Files-11	Fixed	Files-11	Fixed
Files-11	Stream	Files-11	Variable
Files-11	Variable	Files-11	Variable
Files-11	Fixed	RT-11,DOS-11	Fixed
Files-11	Stream	RT-11,DOS-11	Stream
Files-11	Variable	RT-11,DOS-11	Stream
Files-11	Variable	RT-11,DOS-11	Binary
RT-11,DOS-11	Binary	Files-11	Variable
RT-11,DOS-11	Fixed	Files-11	Fixed
RT-11,DOS-11	Stream	Files-11	Variable

## 5.3 Carriage Control Options

For Files-11 files only, you can specify a file's carriage control attributes by using the qualifier `/CARRIAGE_CONTROL=option`. This qualifier has the following options: `CARRIAGE_RETURN`, `FORTTRAN`, and `NONE`.

The default option is `CARRIAGE_RETURN`. The `CARRIAGE_RETURN` option indicates that each record is to be preceded by a line feed and followed by a carriage return when the record is written to a carriage control device (such as a line printer or terminal).

The `FORTTRAN` option indicates that the first character of each record is to be interpreted as the carriage control specifier. This option does not alter any of the actual data. The `FORTTRAN` option is valid only with Files-11 output files and is incompatible with stream output to Files-11.

The `NONE` option indicates that carriage control is not implied for the file. If you specify `NONE`, EXCHANGE assumes that each record contains the control characters necessary for proper formatting.

## 6 Virtual Devices

A *virtual device* is an RMS file that contains an image of a foreign block-addressable device. EXCHANGE treats a virtual device as if it were a foreign volume.

You must use EXCHANGE interactively if you want to manipulate a virtual device. In other words, you cannot operate on a virtual device by entering the EXCHANGE commands at DCL level, because a virtual device must be mounted before it can be accessed. EXCHANGE executes only a single command in DCL mode, and it is not possible to both mount and operate on the virtual device in a single command.

To create a virtual device, use the command INITIALIZE/CREATE, as follows:

```
EXCHANGE> INITIALIZE/CREATE DRA1:[TEST]MYPHILE
%EXCHANGE-S-INITIALIZED, the RT-11 volume DRA1:[TEST]MYPHILE.DSK
has been initialized
```

This command creates a virtual device in the directory [TEST] on DRA1. EXCHANGE applies the default file type DSK.

When you use the EXCHANGE command MOUNT/VIRTUAL, include two parameters (as opposed to one parameter for MOUNT/FOREIGN). The first parameter supplies the name of the virtual device; the second parameter is the name of the RMS file containing the device image. Thereafter, EXCHANGE uses the virtual device name as a logical name to determine the physical location of the virtual device. For example:

```
EXCHANGE> INITIALIZE/CREATE DRA2:[DISKS]VIRTUAL
%EXCHANGE-S-INITIALIZED, the RT-11 volume DRA2:[DISKS]VIRTUAL.DSK
has been initialized
```

```
EXCHANGE> MOUNT/VIRTUAL DISK: VIRTUAL.DSK
%EXCHANGE-S-MOUNTVIR, the RT-11 virtual volume DISK: has been mounted
using the file DRA2:[DISKS]VIRTUAL.DSK;1
```

```
EXCHANGE> DIRECTORY DISK:
```

```
Directory DISK:      16-MAR-1985 08:30
using DRA2:[DISKS]VIRTUAL.DSK;1
```

```
ANSI.COM      1 16-Mar-1985    QARINV.TXT    3 16-Mar-1985
REASSI.COM    13 16-Mar-1985    SYSCRE.MAC    1 16-Mar-1985
```

```
Total of 4 files, 18 blocks. Free space 4944 blocks, largest 4474.
```

```
EXCHANGE> COPY /LOG SYS$LOGIN:WORKLIST.TXT DISK:
%EXCHANGE-S-COPIED, WRKD$: [JOYNER.LOGIN]WORKLIST.TXT;10 copied to
DISK:WORKLI.TXT, 73 records
```

```
EXCHANGE> DISMOUNT DISK:
%EXCHANGE-S-DISMOUNTED, the RT-11 volume DISK: has been dismounted
```

The name "DISK:" is used in the same manner as a device name. Therefore, when naming a virtual device, do not use a name that you have already defined as a logical name. EXCHANGE reads logical name translations first and will ignore a duplicate virtual device name, making that device inaccessible.

---

## EXCHANGE Usage Summary

The Exchange Utility (EXCHANGE) allows you to manipulate mass storage volumes written in formats other than those normally recognized by VMS.

---

**FORMAT**            **EXCHANGE** *command* [*file-spec*] [*file-spec*]  
**EXCHANGE**>        *command* [*file-spec*] [*file-spec*]

---

**PARAMETERS**    *command*

Defines the specific operation to be performed.

***file-spec***

Specifies the device name, directory, and file name for the EXCHANGE input or output device. It has the following general form:

device:[directory]filename.filetype;version

device:            The device name can be either a standard VMS device name of the form ddcu: or a logical name that translates to a VMS device name. If the device field is omitted for a reference, the current default device is assumed. When a virtual device is mounted, a name is created for the virtual device and is used as the device name in subsequent EXCHANGE commands.

[directory]        The syntax of the directory subfield is volume specific.

filename           The name field file specification for an input or output file. The exact format allowed for the filename is dependent on the volume format qualifier used.

filetype           The extension field of the file specification.

version            The version number of the file, if supported by the volume type.

---

**usage summary**

You can use EXCHANGE in two ways. You can work interactively (within the utility) by entering "EXCHANGE" at the DCL prompt. This invokes the utility, which responds with the EXCHANGE> prompt. You can then enter any EXCHANGE command. You must invoke the utility and use it interactively if you want to execute more than one EXCHANGE command. However, you can enter a single EXCHANGE command at DCL level. For example, the following command entered at DCL level lists the directory of a foreign volume:

```
$ EXCHANGE DIRECTORY DMA1:/VOLUME_FORMAT=RT11
```

When you use EXCHANGE at the DCL level, the utility returns you to the DCL prompt after it completes its task. If you are using EXCHANGE interactively, you can return to DCL at any time by typing EXIT or CTRL/Z.



## EXCHANGE Usage Summary

You can direct output from EXCHANGE operations in several ways. The command qualifier `/[NO]MESSAGE` allows you to control the default display of information from EXCHANGE MOUNT, INITIALIZE, and DISMOUNT operations. When you use the EXCHANGE commands COPY, DELETE, RENAME, or TYPE, include the `/LOG` qualifier to send information about the files being processed to `SYSS$OUTPUT`. When you use the EXCHANGE command DIRECTORY, use the `/OUTPUT[=file-spec]` qualifier to direct the output to a specified file. If you specify the `/OUTPUT` qualifier without a file specification, the output is directed to `SYSS$OUTPUT`. To send the output to a printer, use the `/PRINTER` qualifier with the DIRECTORY command.

# EXCHANGE

## EXCHANGE Commands

---

### EXCHANGE COMMANDS

The syntax of each of the EXCHANGE commands is similar to that of the corresponding DCL command. This section describes the functions and provides examples of the EXCHANGE commands.

---

## COPY

Transfers a file or files from an input volume to an output volume. You can use the COPY command to do any of the following:

- Copy a file from a foreign volume to a native volume
- Copy a file from a native volume to a foreign volume
- Copy a file from one foreign volume to another foreign volume
- Convert the format of the file during the transfer
- Copy groups of files from volume to volume
- Give the output file a different name from the input file

---

### FORMAT

**COPY** *input-file-spec[, . . . ] output-file-spec*

---

### PARAMETERS

***input-file-spec[, . . . ]***

Specifies the names of one or more input files to be copied. If you specify more than one input file, separate them with commas or plus signs. The syntax for input file names depends on the volume format option. Refer to Section 4 for a description of volume format options. You can specify standard VMS wildcards in both Files-11 and foreign file names. COPY supports wildcard directories for Files-11 and DOS-11 input.

#### ***output-file-spec***

Specifies the name of the output file, directory, or device to which the input files are to be copied. If the input is a single file, you can specify an explicit output name (which is equivalent to a rename on a copy operation). If the input is more than one file, the output specifier must be one of the following:

- Wildcards (\*, \*\*, or \*.\*\*) specifying current default device and directory
- An explicit device and/or directory for Files-11 output, such as BB:[EXCHANGE.TMP], with or without wildcards for the file name
- An explicit device for RT-11 as in DLA2:/VOLUME=RT11
- An explicit device or directory for DOS-11 output, such as TAPE:/VOLUME=DOS11 or TAPE:[11,132]/VOLUME=DOS11

The output file names are constructed according to rules implied by the input and output volume qualifiers. COPY does not concatenate multiple input files into a single output file. Wildcard directories are not permitted. The syntax for input file names depends on the volume format option. Refer to Section 4 for a description of volume format options.

You must specify at least one field in the output file specification; COPY replaces missing fields with the corresponding field of the related input file specification. If the input file has no corresponding field, COPY substitutes null text fields and maximizes version numbers.

# EXCHANGE

## COPY

The UIC of the output file is the UIC of the current process. For DOS-11 output in UIC format, EXCHANGE uses the current default directory; otherwise, it uses the current process UIC as a directory. You can specify an alternate directory for DOS-11 output in the command.

---

### QUALIFIERS

#### ***/BOOT[=nn]***

Copies bootstrap information from a monitor and the handler files to blocks 0 and 2 through 5 of an RT-11 volume, permitting you to use that volume as a system volume. The COPY/BOOT operation does not create any files on the volume; it is intended only to create bootable RT-11 systems.

The /BOOT qualifier implies /VOLUME\_FORMAT=RT11 for both input and output specifications. The output device can be omitted, as it is assumed to be identical to the input device. You cannot combine the /BOOT qualifier with qualifiers other than /LOG. The COPY/BOOT command requires that both the input and output devices be the same volume or virtual device. The file name of the desired monitor must be specified as the input specification.

RT-11 Version 1.0 through Version 3.0 monitors had the system device handler linked into the monitor image. For Version 4.0 of RT-11, the system device handler uses the standard device handler, and the COPY/BOOT command must dynamically link the handler into the bootstrap area. COPY/BOOT finds the default handler for the specific device type and merges the handler with the monitor as it is copied to the boot area.

You can use the two-letter argument *nn* to override the default system device handler. The most frequent use of this option occurs when a diskette is mounted in an RX02 drive, and you want to create a diskette bootable from an RX01 drive. (The diskette must be single density.) The default handler for the RX02 is DY.SYS, and the handler for the RX01 is DX.SYS; therefore, you would use the command COPY/BOOT=DX to create the bootable RX01 system diskette. Do not specify /BOOT=nn for Version 3.0 RT-11 and earlier systems; instead, choose the monitor file DYMNxx.SYS or DXMNxx.SYS as the source file.

#### ***/[NO]LOG***

Controls whether the EXCHANGE command COPY displays the file specifications of each file copied. If you specify /LOG, the system displays the following data for each copy operation: the file specifications of the input and output files, and the number of blocks or the number of records copied (depending on whether the file is copied on a block-by-block or record-by-record basis). The default is /NOLOG.

---

### FILE QUALIFIERS

#### ***/ALLOCATION=n***

Forces the initial allocation of the output file to the number of 512-byte blocks that you specified as *n*. The /ALLOCATION qualifier is valid only for Files-11 and RT-11 output files.

By default, COPY determines the initial allocation of the output file by the size of the input file. Typically, /ALLOCATION is needed only when you are creating a contiguous file on Files-11 (using /BEST\_TRY\_CONTIGUOUS or /CONTIGUOUS), when the input file is on magnetic tape, or when you want additional space at the end of the file.

If you specify /ALLOCATION, the file's allocated size does not change, unless you also specify /TRUNCATE. When you are unsure of the output size, you might want to specify both /ALLOCATION and /TRUNCATE.

## ***/[NO]BEST\_TRY\_CONTIGUOUS***

Indicates whether the Files-11 output file is to be allocated contiguously on a "best effort" basis; that is, whether EXCHANGE will attempt to place the file on consecutive physical disk blocks. If insufficient contiguous space is available, the file occupies the largest available contiguous space plus additional extents as necessary for the rest of the allocation. You can apply this qualifier only to a Files-11 output file.

The /BEST\_TRY\_CONTIGUOUS qualifier has no effect when you copy files to magnetic tape volumes. When you would like a file from a magnetic tape to be copied contiguously, use both the /ALLOCATION and the /BEST\_TRY\_CONTIGUOUS qualifiers, because the size of the file on magnetic tape cannot be determined until after it is copied to the disk. If you do not know the exact size of the file, you can overestimate the size and specify /TRUNCATE (along with /ALLOCATION and /BEST\_TRY\_CONTIGUOUS) to avoid wasted space.

The default is /NOBEST\_TRY\_CONTIGUOUS.

## ***/CARRIAGE\_CONTROL=option***

Defines the carriage control attributes of a file, as well as other attributes of the records. The carriage control options are: CARRIAGE\_RETURN, which implies carriage return/line-feed control; FORTRAN, which indicates that the first character of each record is to be interpreted as the carriage control specifier; and NONE, which indicates that carriage control is not implied. Refer to Section 5.3 for a detailed description of /CARRIAGE\_CONTROL options.

The default is /CARRIAGE\_CONTROL=CARRIAGE\_RETURN.

## ***/[NO]CONTIGUOUS***

Indicates whether the copied file is to be contiguous; that is, stored on consecutive physical blocks on an output disk volume. The /CONTIGUOUS qualifier is valid only for Files-11 output files.

The /CONTIGUOUS qualifier has no effect when you copy files to magnetic tape volumes. When you would like a file from a magnetic tape to be copied contiguously, use both the /ALLOCATION and /CONTIGUOUS qualifiers because the size of the file on magnetic tape cannot be determined until after it is copied to the disk. If you do not know the exact size of the file, you can overestimate the size and specify the /TRUNCATE qualifier (along with /ALLOCATION and /CONTIGUOUS) to avoid wasted space.

The default is /NOCONTIGUOUS.

## ***/[NO]DELETE***

Controls whether COPY deletes existing files of the same name during the copy operation. This qualifier is valid for RT-11 output only; it is equivalent to the RT-11 COPY command qualifier /REPLACE. In fact, you can use the EXCHANGE COPY command qualifier /REPLACE to control file deletion, although its function differs from that of /DELETE (see the description of the /REPLACE qualifier for details on its function).

# EXCHANGE COPY

If you want a message displayed when you delete a file, include the /LOG qualifier in your command. To prevent automatic file deletion, use /NODELETE.

The default is /DELETE. Files with the same name as the output file name are deleted *after* the new file has been copied.

## ***/EXTENSION=n***

Specifies the number of blocks to be added to the output file each time the file is extended. This qualifier is valid for Files-11 output files only.

EXCHANGE determines the default extension according to the following hierarchy:

- 1 An explicit value specified on the /EXTENSION qualifier
- 2 The current process default extension value set by the command SET RMS\_DEFAULT
- 3 The current system default extension value set at system generation or with the SET RMS\_DEFAULT/SYSTEM command

Use the /EXTENSION qualifier to set an extension quantity with magnetic tape input; EXCHANGE preallocates a file of the correct size when the input is on a directory-structured-device.

## ***/[NO]PROTECT***

Determines whether protection is set for an RT-11 output file. The owner UIC of the output file is the UIC of the current process. This qualifier is not valid for Files-11 or DOS-11 output files. Protection attributes for Files-11 output are taken from the current process default protection.

EXCHANGE does not attempt to transfer protection attributes from the input file to the output file, because protection mechanisms of various operating systems do not readily translate to one another.

The default is /NOPROTECT.

## ***/RECORD\_FORMAT=(option[, . . . ])***

Defines the internal record structure of a file, as well as other attributes of the records. Table EXCH-2 summarizes the format options.

Refer to Section 5.1 for a description of the /RECORD\_FORMAT qualifier and options.

## ***/[NO]REPLACE***

Requests that if an RT-11 output file already exists with the same file specification as that entered for the output file, the existing file is to be deleted *before* the copy proceeds. COPY allocates new space for the output file. The /REPLACE qualifier is valid for RT-11 output only; it is equivalent to the RT-11 COPY command qualifier /PREDELETE.

By default, COPY creates the new file first and then, after the copy operation is done, deletes the previous file. However, when you use /REPLACE, COPY deletes the previous file *before* it copies the new file. This can be a problem if the input file has been corrupted because the previous version of the file will have been deleted. Therefore, you should use /REPLACE only when there is insufficient room for two copies of the file.

## ***/[NO]REWIND***

Determines whether a DOS-11 input magnetic tape reel logically rewinds to the beginning-of-tape mark (BOT) before EXCHANGE searches for the file name specified in the input specifier. This qualifier is valid for DOS-11 magnetic tape only. The default is /NOREWIND.

Use the /REWIND qualifier when you want COPY to search for a file from the logical beginning of the magnetic tape, instead of from the current physical position of the tape.

## ***/START\_BLOCK=[n]***

For RT-11 volumes, specifies the logical block number where the file is to be placed. This qualifier is especially useful with TU58 tape cassettes, because performance can be significantly enhanced by careful placement of files.

## ***/[NO]SYSTEM***

Controls whether the COPY command copies files that have the file type SYS. Files with a file type of SYS are usually necessary for the operation of an RT-11 system. Only RT-11 volumes handle SYS files in this manner.

The default is /NOSYSTEM; the COPY command does not copy an RT-11 file with the type SYS, whether matched by a wildcard specification or explicitly named. EXCHANGE displays a message whenever it skips over a SYS file during a copy operation.

## ***/TRANSFER\_MODE=option***

Specifies the I/O method to be used in a transfer. This qualifier is useful for all volume formats.

<b>Option</b>	<b>Function</b>
AUTO	Select BLOCK transfer for efficiency if possible
BLOCK	Transfer block by block without looking at records
RECORD	Transfer record by record

The default is the AUTOMATIC transfer mode. In AUTOMATIC mode, EXCHANGE attempts to use a BLOCK transfer whenever possible. BLOCK transfers are possible between RT-11 volumes or between RT-11 and DOS-11 volumes, since the internal file structures are identical. AUTOMATIC does not use the BLOCK transfer if either file specification contains a /RECORD\_FORMAT qualifier.

A BLOCK transfer moves data between devices. Since no interpretation is done on the data, BLOCK transfers are more efficient than RECORD transfers. The block sizes on both devices must be identical. Both input and output must be in BLOCK format. Specifying BLOCK on one parameter implies BLOCK for the other file or device specification.

A BLOCK transfer produces an exact copy of the file. If the output device is Files-11, the file will be a sequential file with fixed-length 512-byte records. This feature is used primarily to avoid any interpretation of the data during the transfer. If the Files-11 file is a sequential file with 512-byte fixed-length records, there is no difference between a /TRANSFER\_MODE=BLOCK transfer and a /RECORD=FIXED=512 transfer.

# EXCHANGE

## COPY

A RECORD transfer moves the data record by record. A RECORD transfer requires more time than a BLOCK transfer, but it must be used if the input and output record structures differ.

When the /LOG qualifier is used in a COPY command, EXCHANGE displays the size of the file that was transferred. If BLOCK mode was used, the message gives the file size as the number of blocks transferred. If RECORD mode was used, the message displays the number of records.

### ***/[NO]TRUNCATE***

Controls whether COPY truncates an output file at the end-of-file when copying it. The default is /NOTRUNCATE; COPY uses the allocation of the input file to determine the size of the output file.

### ***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. The default format qualifier is dependent on the device type.

If used, volume format qualifiers must be attached to one or both of the file specification parameters; you cannot attach them directly to the command. A volume format qualifier determines the format of the file name and directory specifications, and often implies certain defaults. Refer to Section 4 for a description of the volume format options and defaults.

---

## DESCRIPTION

COPY transfers a file or files from an input volume to an output volume.

You can create multiple output files by specifying multiple input files. When multiple output files are created, the corresponding field from each input file is used in the output file name.

If you do not specify a version number for Files-11 output, COPY applies a version number as follows:

- The same version number as that of the input file, if the input volume structure supports version numbers and no file exists with the same name and type
- A version number that is one greater than the highest version number of an existing file with the same file name and file type
- Version 1 if neither of the above applies

If you use an asterisk (\*) wildcard character to specify the output file version number, COPY uses the version numbers of the associated input files (if any) as the version numbers of the output files.

Note that ANSI-formatted magnetic tapes do not handle version numbers in the same manner as disks. For a description of version numbers on magnetic tape, refer to the *Guide to VMS Files and Devices*.

EXCHANGE might reformat files during the copy operation. The defaults for reformatting are dependent on the record and volume format qualifiers that are attached to both the input and output file specifications, as well as the type fields of the file specifications. Refer to Section 5 for a description of the /RECORD\_FORMAT qualifier and record format defaults. Refer to Section 4.4 for a description of volume format defaults.



# EXCHANGE COPY

The COPY command does not copy a file with the SYS type unless you specify the /SYSTEM qualifier. EXCHANGE displays a message if it passes over one or more SYS files during a copy operation.

EXCHANGE does not copy files with the type BAD if the file specification contains wildcards. EXCHANGE does not display a message when it passes over one or more BAD files during a copy operation. Therefore, to copy a file with the type BAD, specify the file name explicitly instead of using wildcards.

---

## EXAMPLES

**1** EXCHANGE> COPY TEST.DAT DYAO:NEWTST.DAT/VOLUME=RT11

The command in this example copies the contents of the file TEST.DAT from the default disk and directory into a file named NEWTST.DAT on an RT-11 diskette (mounted on DYAO). If a file named NEWTST.DAT already exists, the COPY command replaces it. The record formats are variable length on the Files-11 input and ASCII stream on the RT-11 output.

**2** EXCHANGE> COPY/LOG TEST.DAT DYAO:NEWTST.DAT  
%EXCHANGE-S-DELETEPREV, previous copy of DYAO:NEWTST.DAT deleted  
%EXCHANGE-S-COPIED, WRKD\$: [FRED]TEST.DAT;5  
copied to DYAO:NEWTST.DAT, 93 records

The command in this example is the same as in the preceding example, except that the /LOG qualifier is included so that the actions of the command are displayed. The /VOLUME\_FORMAT qualifier is omitted; EXCHANGE defaults to RT-11 format for the foreign mounted diskette. Since a file named NEWTST.DAT already exists on the diskette (from the command in the previous example), the first copy of the file is deleted after the second copy is successfully transferred.

**3** EXCHANGE> COPY/LOG LARGE.DAT DYAO:/VOLUME=RT11  
%EXCHANGE-S-DELETEPREV, previous copy of \_DYAO:LARGE.DAT deleted  
%EXCHANGE-W-RTOUTEOF, end-of-file on output \_DYAO:LARGE.DAT,  
insufficient space on volume  
%EXCHANGE-I-PARTCOPIED, WRKD\$: [FRED]LARGE.DAT;9 partially copied to  
\_DYAO:LARGE.DAT, 1670 records

EXCHANGE> COPY/LOG/REPLACE LARGE.DAT DYAO:/VOLUME=RT11  
%EXCHANGE-S-DELETEPREV, previous copy of \_DYAO:LARGE.DAT deleted  
%EXCHANGE-S-COPIED, WRKD\$: [FRED]LARGE.DAT;9 copied to  
\_DYAO:LARGE.DAT, 3288 records

The first command in this example fails because there is insufficient space on DYAO for EXCHANGE to do a normal copy (that is, to copy the file into a temporary file on DYAO, delete the existing file of the same name, and then rename the temporary file).

The second command in this example includes the /REPLACE qualifier, which directs COPY to delete an existing version of the output file before copying the new file. The first message generated by this command indicates that EXCHANGE has deleted an existing file. The second message indicates that the copy operation has successfully completed.

# EXCHANGE

## COPY

4 EXCHANGE> COPY MTAO:[11,132]\*.COM/VOLUME=DOS11 [FRED.TEMP]

The command in this example copies all files with the file type COM owned by UIC [11,132] to the subdirectory [FRED.TEMP]. The default DOS-11 record format is STREAM. The default Files-11 record format is VARIABLE.

5 EXCHANGE> COPY/BOOT DYAO:RT11SJ.SYS DYAO:

The COPY/BOOT command in this example makes the diskette mounted on the RX02 drive DYAO a bootable RT-11 system. Bootstrap information is written to the volume using the RT-11 single job monitor RT11SJ.SYS and the system device handler DY.SYS. The diskette can be formatted in single or double density.

6 EXCHANGE> COPY/BOOT=DX DYAO:RT11SJ.SYS DYAO:

As in the previous example, the COPY/BOOT command in this example makes the diskette mounted on the RX02 drive DYAO a bootable RT-11 system. In this example, however, the system device handler is DX.SYS, the RX01 diskette handler. The diskette must be formatted in single density in order to boot on the RT-11 system.

7 EXCHANGE> COPY/BOOT CSA1:CONSOL.SYS

The command in this example writes bootstrap information on the console storage device on a VAX processor.

8 EXCHANGE> COPY DMAO:FROG.DAT/VOLUME=RT11/RECORD=STREAM -  
\_EXCHANGE> FROG4JUN.DAT/RECORD=(FIXED=80,PAD=" ")

The command in this example copies an ASCII stream file (the default) from an RT-11 volume to the current default device and directory. The output contains fixed-length records that are padded to 80 bytes with the space character.

---

## DELETE

Deletes one or more files from a foreign block-addressable mass storage volume. EXCHANGE does not delete files from Files-11 volumes; the DELETE command is supported only on RT-11 volumes.

---

**FORMAT**            **DELETE** *file-spec[, . . . ]*

---

**PARAMETERS**    ***file-spec[, . . . ]***

Specifies the names of one or more files to be deleted. You can specify wildcard characters in any of the file specification fields.

To delete more than one file, separate the file specifications with commas or plus signs.

The DELETE command does not delete a file with the SYS type unless you specify the /SYSTEM qualifier. EXCHANGE displays a message if it passes over one or more SYS files during a delete operation.

EXCHANGE does not delete files with the type BAD if the file specification contains wildcards. When this happens, you will not receive a warning. Therefore, to delete files with the type BAD, enter their file specifications explicitly.

---

**QUALIFIERS**        ***/[NO]LOG***

Controls whether the DELETE command displays the file specification of each deleted file. The default is /NOLOG.

***/[NO]SYSTEM***

Controls whether the DELETE command deletes files with the file type SYS. Files with the type SYS are usually necessary for the operation of an RT-11 system. Only RT-11 volumes handle .SYS files in this manner.

The default is /NOSYSTEM; the DELETE command does not delete an RT-11 file with the SYS type, whether matched by a wildcard specification or explicitly named. EXCHANGE displays a message whenever it skips a SYS file during a delete operation.

***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. RT-11 volumes are the only volumes on which DELETE is currently supported. Refer to Section 4 for a description of the volume format options and defaults.

---

## EXAMPLES

**1**    EXCHANGE> DELETE DMA0:COMMON.SUM/VOLUME=RT11

The command in this example deletes the file COMMON.SUM from the RT-11 device DMA0.

# EXCHANGE

## DELETE

2 EXCHANGE> DELETE DXAO:\*.OLD

The command in this example deletes all files with the file type OLD from the diskette.

3 EXCHANGE> MOUNT /VIRTUAL TEST: TEST.DSK/VOLUME=RT11  
EXCHANGE> DELETE TEST:ALPHA.TXT,BETA.TXT

The command in this example deletes the files ALPHA.TXT and BETA.TXT from the virtual device file TEST.DSK.

---

## DIRECTORY

Provides a list of files or information about a file or group of files. The files must reside on a foreign volume; EXCHANGE does not list directories of Files-11 volumes.

---

**FORMAT**            **DIRECTORY** *[file-spec[, . . . ]]*

---

**PARAMETERS**    *file-spec[, . . . ]*

Specifies one or more files to be listed. The /VOLUME\_FORMAT qualifier determines the syntax of a file specification.

To specify more than one file, separate the file specifications with either commas or plus signs. You can use wildcard characters in the directory specification, file name, file type, or version number fields of a file specification.

---

**QUALIFIERS**

***/[NO]ALL***

Lists all deleted or unused files on an RT-11 volume, in addition to other files selected by the command. For example, the following command lists all MACRO source files, in addition to deleted and unused files:

```
EXCHANGE> DIRECTORY DMAO:* .MAC/ALL
```

***/[NO]BADBLOCKS***

Scans the volume to find any blocks that return read errors. The data on the volume is not modified. If a bad block replacement table is present, the contents of the table are displayed. This is valid for RT-11 volumes only.

***/[NO]BLOCKS***

Lists the starting block number of the file. This qualifier is valid only for directories of RT-11 devices. The first block of the device is block number 0. The default is /NOBLOCKS.

***/[NO]BRIEF***

Includes only the file name of each file to be listed. Specifying the /BRIEF qualifier is equivalent to specifying /NODATE/NOSIZE. The default is /BRIEF.

***/COLUMNS=n***

Lists the files, using the specified number of columns on each line of the display. This qualifier is used in conjunction with the /BRIEF qualifier (either explicitly or by default). The default number of columns is dependent on the volume format and the information requested. The DIRECTORY command attempts to use as many columns as possible. If you request too many columns, DIRECTORY displays a message and reduces the number of columns to the number that fit on the listing.

***/[NO]DATE***

Includes the date for each file listed. If you omit this qualifier, the default is /DATE.

# EXCHANGE DIRECTORY

## ***/[NO]DELETED***

Lists a directory of files that have been deleted from an RT-11 device, but whose file name information has not been destroyed. The listing includes the file names, types, sizes, creation dates, and starting block numbers (in decimal, unless you also specify the */OCTAL* qualifier) of the deleted files. The */DELETED* qualifier is valid only with block-addressable volumes in RT-11 format. The default is */NODELETED*.

## ***/[NO]FREE***

Includes unused areas in the directory listing. The */FREE* qualifier is valid only with RT-11 formatted volumes.

## ***/FULL***

Lists all the available information for each file. The format of the listing depends on the format of the volume. The */FULL* qualifier overrides the default brief listing format.

## ***/[NO]OCTAL***

Controls whether numeric information is displayed in decimal or octal format. The default is */NOOCTAL*; numbers are displayed in decimal radix. Dates are always displayed in decimal format.

## ***/OUTPUT[=file-spec]***

Writes the DIRECTORY output to a specified file, rather than to the current SYS\$OUTPUT device. If you specify the */OUTPUT* qualifier without a file specification, the output is directed to SYS\$OUTPUT. If you omit the file type in the file specification, the default file type is LIS. If you specify a file type and omit the file name, the default file name is EXCHDIRE. No wildcard characters are allowed in the file specification.

## ***/OWNER***

Displays information about the owner of a volume and the files on the volume. For RT-11, the volume owner is shown. For DOS-11, the UIC of the file owner is shown.

## ***/PRINTER***

Queues the command output for printing under the name specified by the */OUTPUT* qualifier. If you specify */PRINTER* without the */OUTPUT* qualifier, the output is directed to a file named EXCHDIRE.LIS, which is spooled for printing and then deleted.

## ***/[NO]SIZE***

Displays the file size in blocks for each file listed. The default is */SIZE*.

## ***/[NO]SUMMARY***

Lists a summary of the usage of the directory segments for an RT-11 volume. If a bad block replacement table is present, the contents of the table are displayed.

## ***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. The default format is dependent on the device type. Refer to Section 4 for a description of the volume format options and defaults.

The EXCHANGE command DIRECTORY is not valid for Files-11 devices.

# EXCHANGE DIRECTORY

---

## DESCRIPTION

The output of the DIRECTORY command depends on the volume format and on certain formatting qualifiers and defaults. The following are the formatting qualifiers:

/ALL	/BLOCKS	/BRIEF
/COLUMNS	/DATE	/FULL
/OCTAL	/OWNER	/SIZE

The files that are listed always appear in the order in which they appear in the volume directory or the order in which they reside on a magnetic tape.

---

## EXAMPLES

**1** EXCHANGE> DIRECTORY DLA2:.OBJ/VOLUME=RT11/FULL

The command in this example lists all files with the type OBJ on the RT-11 volume mounted on DLA2. The /FULL qualifier causes the file sizes and dates to be listed along with the names.

**2** EXCHANGE> DIRECTORY MFA0:/VOLUME=DOS11

The command in this example lists all files on the DOS-11 magnetic tape mounted on MFA0. The magnetic tape is rewound before the files are listed.

# EXCHANGE

## DISMOUNT

---

### DISMOUNT

Releases a volume previously accessed by the EXCHANGE command MOUNT.

---

**FORMAT**            **DISMOUNT** *device-name[:]*

---

**PARAMETERS**    *device-name[:]*  
Specifies the name of the device to be dismounted. You can specify a physical device name or a logical name assigned to a physical device name. If you omit a controller designation or a unit number, the defaults are controller A and unit 0, respectively. You can also specify the name of a virtual device.

---

**QUALIFIERS**     */[NO]MESSAGE*  
Controls whether or not EXCHANGE displays a message that the volume was dismounted. The default is determined by the /MESSAGE qualifier on the EXCHANGE command when EXCHANGE was activated.

---

**DESCRIPTION**    The DISMOUNT command closes all connections that EXCHANGE maintains to the device. This command does not affect the state of the operating system mount; the device remains accessible to VMS. If you do not use the DISMOUNT command, an implicit DISMOUNT is automatically executed when you exit EXCHANGE.

The DISMOUNT command is valid only with foreign devices.

---

### EXAMPLE

```
EXCHANGE> MOUNT/FOREIGN MTA0:  
EXCHANGE> COPY MTA0:AVERAGE.FOR/VOLUME=DOS11 *  
EXCHANGE> DISMOUNT MTA0:
```

The first command in this example mounts the tape on the device MTA0. The second command in this example transfers a file from the magnetic tape to the current default directory. The last command releases EXCHANGE's access to the volume; however, the volume is still mounted on the operating system and is accessible to VMS.



---

**EXIT**

Terminates execution of EXCHANGE. Control is returned to the DCL command level. You can also use CTRL/Z to exit EXCHANGE.

---

**FORMAT**

**EXIT**

# EXCHANGE

## HELP

---

## HELP

Displays information about EXCHANGE commands and qualifiers.

---

### FORMAT

**HELP** [*command* [*qualifier* [*option* [*option*]]]]

---

### PARAMETERS

#### ***command***

Specifies the name of the EXCHANGE command that you want information about. If you omit the command, HELP displays general information listing all commands recognized by EXCHANGE.

#### ***qualifier***

Gives the name of the qualifier to be explained.

#### ***option***

Gives the name of the option to be explained.

---

### DESCRIPTION

For an overview of EXCHANGE and a listing of the EXCHANGE command names, enter the HELP command with no arguments.

If you enter HELP and the name of an EXCHANGE command, HELP displays a description of the command followed by a list of related qualifiers. For information on any of the related qualifiers, enter the qualifier name at the prompt.

You can also obtain information on any EXCHANGE command qualifier by entering HELP, the command, and the qualifier at the EXCHANGE prompt, as follows:

```
EXCHANGE> HELP COPY /CONTIGUOUS
```

For information on a qualifier with options, enter HELP, the command, the qualifier, and the option at the EXCHANGE prompt.

If you specify an asterisk (\*) in place of any keyword, the HELP command displays all information available at that level.

If you specify an ellipsis ( . . . ) after any keyword, the HELP command displays all information relating to that keyword.

You can specify percent signs and asterisks in the keyword as wildcard characters.

---

### EXAMPLE

```
EXCHANGE> HELP COPY/VOLUME...
```

The command in this example displays all the help that is available for the COPY qualifier /VOLUME\_FORMAT, including descriptions of each of the options.

---

## INITIALIZE

Formats and writes a label on a foreign mass storage volume. For directory-structured devices, the device directory is also initialized.

---

**FORMAT**            **INITIALIZE** *device-name* [*volume-label*]  
**INITIALIZE/CREATE** *file-name* [*volume-label*]

---

**PARAMETERS**    ***device-name***  
 Specifies the name of the device on which the volume to be initialized is physically mounted.

The device name can also refer to the name of a mounted virtual device to be reinitialized.

***file-name***  
 For INITIALIZE/CREATE, file-name refers to the name of a file to be created and initialized as a virtual device.

***volume-label***  
 Specifies the identification to be written onto the volume header for RT-11 volumes only. The volume label can contain up to a maximum of 12 alphanumeric characters. The default is *VMS Exchange*. Use quotation marks to specify a volume label with lowercase letters.

---

**QUALIFIERS**        ***/ALLOCATION=n***  
 Specifies the allocation of a new virtual device file in terms of 512-byte blocks. The allocation specified is the number you entered as n. If you do not specify the */ALLOCATION* qualifier when you create a new virtual device file, the default allocation is 494 blocks, the size of a single-density diskette. The maximum allocation is 65,536 blocks.

A virtual device file is usually the size of a standard device supported by both RT-11 and VMS. These sizes are as follows:

Device	Blocks
TU58	512
RX01	494
RX02	494 (single density)
RX50	800
RX02	988 (double density)
RX33	2400
RL02	20480
RK06	27126
RK07	53790

# EXCHANGE INITIALIZE

You can also use the `/ALLOCATION` qualifier to reduce the size of a physical device. For example, if you want to prepare an RL02 disk but have only an RK07 device available, you can initialize the RK07 to a volume of 20,480 blocks. When the RL02 is available, you can transfer the files to the RL02 knowing they will fit on the smaller device.

## **`/BADBLOCKS[=RETAIN]`**

Performs a bad block scan of the volume before initialization. A file named `FILE.BAD` is created on top of each bad block or group of bad blocks encountered on the device, preventing any future use of the bad areas.

If a bad block is found in either the boot block or the volume directory, the volume is not usable and EXCHANGE displays an error message. If the bad block is in a directory segment other than the first, you might be able to use the volume by reinitializing it with a smaller number of segments (see the `/SEGMENTS` qualifier description).

If you specify `/BADBLOCKS=RETAIN`, EXCHANGE uses the device's existing bad block information, instead of performing a bad block scan. Therefore, initializing takes less time. If you do not specify `RETAIN`, EXCHANGE writes a pattern on each block of the volume, then reads each block to verify that the block is usable. EXCHANGE prints a list of the bad blocks found on the device.

RK06, RK07, and RL02 disk volumes support bad block replacement. Therefore, DIGITAL recommends that you use the `/REPLACE=RETAIN` qualifier for these volumes. If you use the `/BADBLOCKS` qualifier with a volume initialized previously with the `/REPLACE` qualifier, EXCHANGE deletes the bad block replacement table and performs a new bad block scan. If you use `/BADBLOCKS=RETAIN` with such a volume, EXCHANGE uses the `FILE.BAD` files created during the volume initialization.

## **`/CREATE`**

Specifies that a virtual device is to be created and initialized. The specification is a file name; if a file type is not given, EXCHANGE applies the default type of `DSK`. Refer to Section 6 for a description of virtual devices.

## **`/DENSITY=density-value`**

Specifies, for magnetic tape volumes, the density in bytes per inch (bpi) at which the tape is to be written.

For magnetic tape volumes, the density value specified can be 800 or 1600, as long as the density is supported by the magnetic tape drive. If you do not specify a density value for a blank tape, the system uses a default of the lowest density supported by the tape drive.

For the RX02 dual-density diskette drive, use the DCL command `INITIALIZE/DENSITY=SINGLE` or `INITIALIZE/DENSITY=DOUBLE` to reformat the diskettes to a different density; then use the EXCHANGE command `INITIALIZE` to create the RT-11 directory structure.

# EXCHANGE INITIALIZE

**Note:** Diskettes formatted in double density cannot be read or written by the console block storage device (an RX01 drive) of a VAX-11/780 until they have been reformatted in single density.

## ***/EXTRA\_WORDS=n***

Specifies, for RT-11 volumes, the number of extra words to add to each directory entry, in addition to the required seven words. The ability to increase the length of directory entries is useful for some RT-11 applications. Increasing the size of the directory entries reduces the number of entries that fit in each directory segment.

## ***/[NO]MESSAGE***

Controls whether or not EXCHANGE displays a message that the volume was initialized. The default is determined by the /MESSAGE qualifier entered with the EXCHANGE command when EXCHANGE was activated.

## ***/REPLACE=RETAIN***

Retains, when an RT-11 volume is initialized, the bad block replacement table and any existing FILE.BAD files.

The RETAIN option is required; EXCHANGE cannot build a replacement table for a volume. The RT-11 system builds and uses the table based on specific hardware error conditions. The VMS I/O system is different, and cannot be relied upon to generate exactly the same error conditions. Therefore, it is not possible for EXCHANGE to generate the same replacement table that would be generated by RT-11.

If no replacement table is present, the qualifier /REPLACE=RETAIN is equivalent to /BADBLOCKS=RETAIN.

## ***/SEGMENTS=n***

Defines, for RT-11 volumes, the number of 2-block directory segments to allocate for the directory. The number of segments in the directory establishes the number of files that can be stored on a device. The system allows a maximum of 72 files per directory segment and 31 directory segments per device. The argument *n* represents the number of segments; the valid range for *n* is from 1 to 31 (decimal). The default values for *n* depend on the device type, as follows:

<b>Device</b>	<b>Segments</b>
TU58	1
RX01	1
RX02	1 (single density)
RX02	4 (double density)
RX50	4
RX33	16
RLO2	16
RK06	16
RK07	31

# EXCHANGE

## INITIALIZE

### ***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. Refer to Section 4 for a description of the volume format options and defaults.

The EXCHANGE command INITIALIZE is not valid for Files-11 devices.

---

## DESCRIPTION

The EXCHANGE command INITIALIZE erases all files from a volume. After initialization, the volume directory contains no files. DOS-11 magnetic tapes and RT-11 block-addressable devices can be initialized.

The device must be mounted with the /FOREIGN qualifier.

---

## EXAMPLES

**1** \$ MOUNT/FOREIGN DLA2:  
%MOUNT-I-MOUNTED, mounted on DLA2  
\$ EXCHANGE  
EXCHANGE> INITIALIZE DLA2:  
%EXCHANGE-S-INITIALIZED, the RT-11 volume \_DLA2: has been initialized

The command in this example initializes the volume mounted on the RL02 drive DLA2. Since DLA2 is a block-addressable device mounted with the /FOREIGN qualifier, RT-11 is the default format. EXCHANGE physically scans all blocks of the volume, builds a bad block replacement table, and displays a message indicating that it failed to turn up any bad blocks.

**2** EXCHANGE> INITIALIZE MTA0:/DENSITY=1600

The command in this example initializes the DOS-11 magnetic tape volume loaded on MTA0. The density is specified as 1600 bpi; the default would have been 800 bpi for an MT drive.

**3** EXCHANGE> INITIALIZE/CREATE/ALLOCATION=1000 VIRTUAL  
%EXCHANGE-S-INITIALIZED, the RT-11 volume DRB0:[LOGIN]VIRTUAL.DSK has been initialized

The command in this example creates a virtual device with an allocation of 1000 blocks in the directory [LOGIN] on DRB0. EXCHANGE applies the default file type DSK.

---

## MOUNT

Makes a foreign volume and the files or data it contains available for processing by EXCHANGE. The EXCHANGE command MOUNT enters the device into internal tables maintained by EXCHANGE.

---

**FORMAT**            **MOUNT** *device-name*  
**MOUNT/VIRTUAL** *device-name file-name*

---

**PARAMETERS**    ***device-name***  
Specifies the physical device name or logical name of the device on which the volume is to be mounted. For MOUNT/VIRTUAL, the device-name parameter supplies a name for the virtual device.

***file-name***

For MOUNT/VIRTUAL only, the file-name parameter gives the name of the file containing the image of the foreign volume.

---

**QUALIFIERS**    ***/[NO]DATA\_CHECK[=(READ,WRITE)]***

Determines whether EXCHANGE performs a second operation after every I/O operation to verify that the data was correctly transferred. If you specify /DATA\_CHECK=WRITE, after every write operation EXCHANGE rereads the data that was just written and compares it with the original data. If you specify /DATA\_CHECK=READ, EXCHANGE reads each block of data twice and verifies that both read operations received identical data.

It is usually more efficient to use the /DATA\_CHECK option on the DCL command MOUNT than to use the option on the EXCHANGE command MOUNT. If you mount a device with the DCL command MOUNT/FOREIGN/DATA\_CHECK, VMS can use features in the device hardware and device driver to perform the redundant I/O operations.

The RX01 and RX02 diskette drives do not contain the necessary features for the operating system to perform data checking. If you use the DCL command MOUNT/DATA\_CHECK with a diskette, the system is unable to perform data checking (no warning message is displayed). EXCHANGE is able to recognize, however, that a diskette was mounted with the data checking option; in this case, EXCHANGE performs the software data checking internally, even if you have not specified an explicit MOUNT/DATA\_CHECK command.

If you specify the /DATA\_CHECK qualifier without an option, the default is /DATA\_CHECK=WRITE.

***/FOREIGN***

Indicates that the volume is not in the standard format used by the VMS operating system; that is, a magnetic tape volume is not in the standard ANSI format, or a disk volume is not in Files-11 format. The EXCHANGE command MOUNT mounts only foreign volumes. The /FOREIGN qualifier is the default. You must use the DCL command MOUNT to mount VMS volumes.

# EXCHANGE MOUNT

The default protection applied to foreign volumes is RWLP (Read, Write, Logical I/O, Physical I/O) for the system and owner. If you mount a volume currently in Files-11 format with the /FOREIGN qualifier, you must have the user privilege to override volume protection (VOLPRO), or your UIC must match the UIC on the volume.

## ***/[NO]MESSAGE***

Controls whether EXCHANGE displays a message indicating that the volume was mounted. The default is determined by the /MESSAGE qualifier specified with the EXCHANGE command when EXCHANGE was invoked.

## ***/VIRTUAL***

Mounts a Files-11 file as a virtual device. When you specify /VIRTUAL, the MOUNT command requires two parameters. The first parameter is a device name assigned as the name of the virtual device. The second parameter is the name of the Files-11 file that is the image of a foreign volume.

## ***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. Refer to Section 4 for a description of the volume format options and defaults.

## ***/[NO]WRITE***

Controls whether the volume can be written. You can specify /NOWRITE to protect files by providing read-only access. Specifying /NOWRITE is equivalent to write-locking the device.

The default is /WRITE. If /WRITE is specified (either explicitly or by default) and the volume itself is write-locked, EXCHANGE displays a message to inform you that the volume is write-locked.

---

## **DESCRIPTION**

The EXCHANGE command MOUNT enters the description of the foreign volume in internal tables maintained by EXCHANGE. This command is different from the DCL command MOUNT, which enters the device in tables maintained by the VMS operating system.

A virtual volume must be explicitly mounted with the MOUNT/VIRTUAL command.

If an EXCHANGE command is given on an unmounted foreign volume, EXCHANGE attempts to execute an implied MOUNT/FOREIGN/WRITE/NODATACHECK on the device. This feature enables EXCHANGE to operate in the single-command DCL mode.

If a MOUNT/FOREIGN (either implied or explicit) command is given for a foreign device that has not been mounted on the VMS system, EXCHANGE issues the equivalent of the DCL command MOUNT/FOREIGN and attempts to make the volume known to the operating system. Any volume mounted in this way remains mounted after EXCHANGE exits.

When EXCHANGE issues the MOUNT/FOREIGN command, the system checks the following:

- That the device has not been allocated to another user
- That a volume is physically loaded on the specified device
- For magnetic tapes, the volume accessibility field of the VOL1 label



For a detailed description of the Mount Utility, see the *VMS Mount Utility Manual*.

---

## EXAMPLES

**1** EXCHANGE> MOUNT MT:  
%EXCHANGE-I-MOUNTED, MATH06 mounted on \_MTA0:

The command in this example requests that the magnetic tape loaded on the device MTA0 be mounted as a foreign volume. The tape label is displayed, since the tape has been previously initialized as an ANSI-labeled tape with the label MATH06. This tape cannot be accessed as a Files-11 tape; it should be reinitialized as a DOS-11 tape during the current EXCHANGE session.

**2** EXCHANGE> MOUNT DMA1:  
%EXCHANGE-I-WRITELOCK, volume is write-locked  
%EXCHANGE-S-MOUNTED, volume DMA1: mounted

The command in this example mounts the foreign volume that is loaded in the RK07 device DMA1, making the volume available for subsequent EXCHANGE commands. EXCHANGE recognizes that the volume itself is write-locked, and displays a message.

# EXCHANGE

## RENAME

---

## RENAME

Changes the file specification of an existing file on an RT-11 volume.

---

**FORMAT**            **RENAME** *input-file-spec output-file-spec*

---

**PARAMETERS**    ***input-file-spec***

Specifies the names of one or more files whose specifications are to be changed.

You can use wildcard characters in the file name and file type specification; if you do, all files that satisfy the specified fields are renamed.

***output-file-spec***

Provides the new file specification to be applied to the input file. The RENAME command uses the file name and file type of the input file specification to provide defaults for nonspecified fields in the output file.

You can specify an asterisk (\*) in place of the file name or file type of the output file; the RENAME command uses the corresponding field in the input file specification to name the output file. Specifying wildcard characters in corresponding fields of the input and output file specifications results in multiple rename operations.

You can omit the device name from the output specification. EXCHANGE uses the device name specified for the input, since it is not possible to rename a file from one device to another.

---

**QUALIFIERS**        ***[/NO]LOG***

Controls whether the RENAME command displays the file specification of each file that it renames. The default is /NOLOG.

***[/NO]PROTECT***

Determines whether protection is set for an RT-11 output file. The default is /NOPROTECT.

This qualifier is not valid for Files-11 or DOS-11 output files. Protection attributes for Files-11 output are taken from the current process default protection.

EXCHANGE does not attempt to transfer protection attributes from the input file to the output file. Protection mechanisms of various operating systems do not readily translate to one another.

The owner UIC of the output file is the UIC of the current process.

***[/NO]SYSTEM***

Controls whether the RENAME command renames files that have the file type SYS. These files are usually files necessary for the operation of an RT-11 system. Only RT-11 volumes handle SYS files in this manner.

# EXCHANGE

## RENAME

The default is /NOSYSTEM; the RENAME command does not rename an RT-11 file with the type SYS, whether it is matched by a wildcard specification or is named explicitly. EXCHANGE displays a message when it skips a SYS file during a rename operation.

EXCHANGE handles files with the file type BAD in a similar manner; that is, the rename operation skips BAD files. However, EXCHANGE does not warn that BAD files are being skipped, and the /SYSTEM qualifier has no effect on BAD files. To rename a file with the type BAD, specify the file explicitly instead of using wildcards.

### ***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. EXCHANGE supports the RENAME command on RT-11 volumes only. Refer to Section 4 for a description of volume format options and defaults.

---

## EXAMPLES

**1** EXCHANGE> RENAME DMA0:AVERAG.OBJ MEAN

The command in this example changes the file name of the file AVERAG.OBJ to MEAN.OBJ.

**2** EXCHANGE> RENAME DLA2:\*.TXT \*.OLD

The command in this example renames all files with the file type TXT to files with the file type OLD; the file names are not changed.

**3** EXCHANGE> RENAME/LOG DMA0:DATA.\* NEW  
%EXCHANGE-I-RENAMED, \_DMA0:DATA.AAA renamed to \_DMA0:NEW.AAA  
%EXCHANGE-I-RENAMED, \_DMA0:DATA.BBB renamed to \_DMA0:NEW.BBB  
%EXCHANGE-I-RENAMED, \_DMA0:DATA.CCC renamed to \_DMA0:NEW.CCC

The command in this example illustrates wildcard characters in the input file names. The device DMA0 contains three files with the file name DATA; the result is the renaming of all three files as displayed by the /LOG qualifier.

# EXCHANGE

## SHOW

---

## SHOW

Displays the devices currently mounted by EXCHANGE.

---

## FORMAT            SHOW

---

## EXAMPLE

```
EXCHANGE> MOUNT DBA0:
%EXCHANGE-I-VMSMOUNT, a "$ MOUNT /FOREIGN DBA0:" command was done by Exchange
%EXCHANGE-S-MOUNTED, the RT-11 volume _DBA0: has been mounted
EXCHANGE> MOUNT DLA2:
%EXCHANGE-I-VMSMOUNT, a "$ MOUNT /FOREIGN DLA2:" command was done by Exchange
%EXCHANGE-S-MOUNTED, the RT-11 volume _DLA2: has been mounted
EXCHANGE> INITIALIZE/CREATE WRKD:[USER]VIRT.DSK
%EXCHANGE-S-INITIALIZED, the RT-11 volume WRKD:[USER]VIRT.DSK;1 has been
initialized
EXCHANGE> MOUNT/VIRTUAL DISK: VIRT.DSK
%EXCHANGE-S-MOUNTVER, the RT-11 volume DISK: has been mounted
                  using the file WRKD:[USER]VIRT.DSK;1

EXCHANGE> SHOW

Mounted volumes:
    volume format:          RT-11
    volume class:          disk (virtual volume)
    virtual file name:      WRKD:[USER]VIRT.DSK;1
    volume size:            494 blocks

  _DLA2:
    volume format:          RT-11
    volume class:          disk
    physical device name:   _DLA2:
    volume size:            20480 blocks

  _DBA0:
    volume format:          RT-11
    volume class:          disk
    physical device name:   _DBA0:
    volume size:            65535 blocks
```

EXCHANGE>

The MOUNT commands in this example mount foreign devices on drives DBA0 and DLA2. The SHOW command displays all devices currently mounted by EXCHANGE.

---

## TYPE

Displays the contents of a file or group of files on the current output device.

---

**FORMAT**            **TYPE** *file-spec[, . . . ]*

---

**PARAMETERS**    *file-spec[, . . . ]*

Specifies the names of one or more input files to be copied. If you specify more than one input file, separate them with either commas or plus signs. You can specify standard VMS wildcards in file names, both Files-11 and foreign. You can use wildcard directories with Files-11 and DOS-11 input.

The syntax for the file names is dependent on the particular volume format option present or implied. Refer to Section 4 for a description of the volume format options and defaults.

---

**QUALIFIERS**      ***/[NO]LOG***

Controls whether TYPE displays the file specifications of each file displayed.

If you specify /LOG, the TYPE command displays the following for each copy operation:

- File specifications of the input and output files
- Number of blocks or the number of records copied (depending on whether the file is copied on a block-by-block or record-by-record basis)

***/RECORD\_FORMAT={option[, . . . ]}***

Defines the internal record structure of a file and other attributes of the records. Section 5.1 describes record format options.

***/[NO]REWIND***

Controls whether the DOS-11 input magnetic tape reel logically rewinds to the beginning-of-tape mark before EXCHANGE searches for the file name given in the input specifier.

Use this qualifier only for DOS-11 magnetic tape devices. The default is /NOREWIND; you should use /REWIND when you want TYPE to start searching for a file at the beginning of the magnetic tape rather than at the current position.

***/VOLUME\_FORMAT=option***

Defines the physical format of the volume to be processed. The default format qualifier is dependent on the device type. Refer to Section 4 for a description of the volume format options and defaults.

# EXCHANGE

## TYPE

---

### EXAMPLE

EXCHANGE> TYPE DYAO:BEAM.RAT/VOLUME=RT11/RECORD=STREAM

The command in this example copies the RT-11 file to the current SYS\$OUTPUT device. The two qualifiers are actually the defaults if DYAO was mounted as a foreign volume.

---

# Index

---

---

## B

---

Bad blocks  
  locating • EXCH-25, EXCH-26, EXCH-32  
  replacing • EXCH-32  
BINARY option  
  record format • EXCH-7

---

---

## C

---

Carriage control option  
  CARRIAGE\_RETURN • EXCH-10  
  FORTRAN • EXCH-10  
  NONE • EXCH-10  
COPY command • EXCH-15 to EXCH-22

---

---

## D

---

DCL prompt  
  entering EXCHANGE command at • EXCH-2,  
  EXCH-11  
DELETE command • EXCH-23 to EXCH-24  
Devices  
  displaying currently mounted • EXCH-40  
  supported by EXCHANGE • EXCH-1  
Device specification  
  DOS-11 volume format • EXCH-5  
  Files-11 volume format • EXCH-4  
  RT-11 volume format • EXCH-4  
Directory  
  listing contents of • EXCH-25  
DIRECTORY command • EXCH-25 to EXCH-27  
  default directory used • EXCH-5  
DISMOUNT command • EXCH-28  
  displaying messages from • EXCH-2  
DOS-11 volume format • EXCH-2, EXCH-5  
  devices supported • EXCH-5  
  specifying files • EXCH-5  
  using with COPY command • EXCH-6  
  using with TYPE command • EXCH-6

---

---

## E

---

Exchange Utility (EXCHANGE)  
  directing output from • EXCH-12  
  exiting • EXCH-12  
  invoking • EXCH-12  
  using at DCL level • EXCH-12  
EXIT command • EXCH-29

---

---

## F

---

File  
  deleting • EXCH-23  
  displaying contents of • EXCH-41  
  DOS-11 format • EXCH-5  
  Files-11 format • EXCH-4  
  format • EXCH-3  
  renaming • EXCH-38  
  RT-11 format • EXCH-4  
  transferring from one volume, to another •  
  EXCH-15  
File attribute options • EXCH-7  
File name  
  changing during copy operation • EXCH-5  
Files-11 volume format • EXCH-2, EXCH-3  
  devices supported • EXCH-4  
  specifying files • EXCH-4  
  using wildcards in file specifications • EXCH-4  
File specification  
  DOS-11 volume format • EXCH-5  
  Files-11 volume format • EXCH-4  
  RT-11 volume format • EXCH-4  
FIXED option  
  record format • EXCH-8  
Foreign volume  
  mounting • EXCH-2  
Format  
  See Volume format

---

---

## H

---

HELP command • EXCH-30

---

## Index

---

### I

---

INITIALIZE command • EXCH-31 to EXCH-34  
displaying messages from • EXCH-2

---

### L

---

Logical name  
using with Files-11 file specifications • EXCH-4  
with RT-11 file specifications • EXCH-4

---

### M

---

Media  
supported by EXCHANGE • EXCH-1  
MOUNT command • EXCH-35 to EXCH-37  
displaying messages from • EXCH-2  
foreign volume • EXCH-2

---

### P

---

PAD option  
specifying pad character • EXCH-9

---

### R

---

Record format  
BINARY option • EXCH-7  
FIXED option • EXCH-8  
PAD option • EXCH-9  
specifying • EXCH-6  
STREAM option • EXCH-8  
VARIABLE option • EXCH-8  
Record format defaults • EXCH-9 to EXCH-10  
Record format options • EXCH-7  
RENAME command • EXCH-38 to EXCH-39  
RT-11 volume format • EXCH-1, EXCH-4  
specifying devices • EXCH-4  
specifying files • EXCH-4

---

### S

---

SHOW command • EXCH-40  
Size limit  
for RT-11 volumes • EXCH-1  
STREAM option  
record format • EXCH-8

---

### T

---

TYPE command • EXCH-41 to EXCH-42

---

### V

---

VARIABLE option  
record format • EXCH-8  
Virtual device • EXCH-1  
allocating • EXCH-31  
creating • EXCH-11  
definition • EXCH-11  
dismounting • EXCH-11  
mounting • EXCH-11  
Volume  
dismounting • EXCH-28  
mounting • EXCH-35  
Volume format  
converting • EXCH-7  
defaults • EXCH-6  
disk default • EXCH-6  
DOS-11 • EXCH-2, EXCH-5  
Files-11 • EXCH-2, EXCH-3  
magnetic tape default • EXCH-6  
options • EXCH-3  
RT-11 • EXCH-1, EXCH-4  
specifying • EXCH-3



# Reader's Comments

VMS Exchange Utility  
Manual  
AA-LA28A-TE

Please use this postage-paid form to comment on this manual. If you require a written reply to a software problem and are eligible to receive one under Software Performance Report (SPR) service, submit your comments on an SPR form.

Thank you for your assistance.

### I rate this manual's:

	Excellent	Good	Fair	Poor
Accuracy (software works as manual says)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Completeness (enough information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clarity (easy to understand)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Organization (structure of subject matter)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Figures (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Examples (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Index (ability to find topic)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Page layout (easy to find information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

I would like to see more/less \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

What I like best about this manual is \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

What I like least about this manual is \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

I found the following errors in this manual:

Page	Description
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

Additional comments or suggestions to improve this manual:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

I am using **Version** \_\_\_\_\_ of the software this manual describes.

Name/Title \_\_\_\_\_ Dept. \_\_\_\_\_

Company \_\_\_\_\_ Date \_\_\_\_\_

Mailing Address \_\_\_\_\_

Phone \_\_\_\_\_

--- Do Not Tear - Fold Here and Tape ---

**digital**<sup>TM</sup>

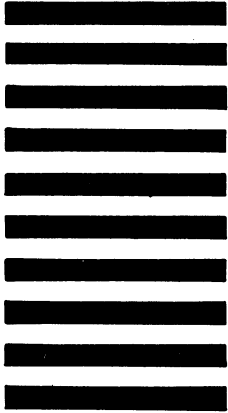


No Postage  
Necessary  
if Mailed  
in the  
United States

**BUSINESS REPLY MAIL**  
FIRST CLASS PERMIT NO. 33 MAYNARD MASS.

POSTAGE WILL BE PAID BY ADDRESSEE

DIGITAL EQUIPMENT CORPORATION  
Corporate User Publications—Spit Brook  
ZK01-3/J35 110 SPIT BROOK ROAD  
NASHUA, NH 03062-9987



--- Do Not Tear - Fold Here ---

Cut Along Dotted Line

# Reader's Comments

VMS Exchange Utility  
Manual  
AA-LA28A-TE

Please use this postage-paid form to comment on this manual. If you require a written reply to a software problem and are eligible to receive one under Software Performance Report (SPR) service, submit your comments on an SPR form.

Thank you for your assistance.

I rate this manual's:	Excellent	Good	Fair	Poor
Accuracy (software works as manual says)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Completeness (enough information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Clarity (easy to understand)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Organization (structure of subject matter)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Figures (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Examples (useful)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Index (ability to find topic)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Page layout (easy to find information)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

I would like to see more/less \_\_\_\_\_  
\_\_\_\_\_

What I like best about this manual is \_\_\_\_\_  
\_\_\_\_\_

What I like least about this manual is \_\_\_\_\_  
\_\_\_\_\_

I found the following errors in this manual:

Page	Description
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

Additional comments or suggestions to improve this manual:  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

I am using **Version** \_\_\_\_\_ of the software this manual describes.

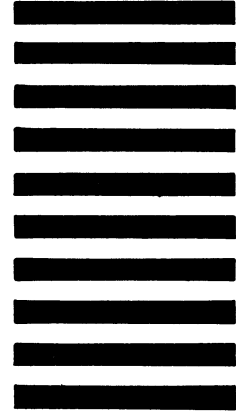
Name/Title \_\_\_\_\_ Dept. \_\_\_\_\_  
Company \_\_\_\_\_ Date \_\_\_\_\_  
Mailing Address \_\_\_\_\_  
Phone \_\_\_\_\_

--- Do Not Tear - Fold Here and Tape ---

**digital**<sup>TM</sup>



No Postage  
Necessary  
if Mailed  
in the  
United States



**BUSINESS REPLY MAIL**  
FIRST CLASS PERMIT NO. 33 MAYNARD MASS.

POSTAGE WILL BE PAID BY ADDRESSEE

DIGITAL EQUIPMENT CORPORATION  
Corporate User Publications—Spit Brook  
ZK01-3/J35 110 SPIT BROOK ROAD  
NASHUA, NH 03062-9987



--- Do Not Tear - Fold Here ---

Cut Along Dotted Line