Metastability Report

Introduction

The dictionary definition of metastability is “a situation that is characterized by a slight margin of stability.” When applied to bi-stable (digital) logic, the term refers to an undesirable, marginally stable output state between V_{IL} max and V_{IH} min.

Metastability can occur in bi-stable storage elements (registers, latches, memories, etc.) when setup and/or hold times are violated. Since setup and hold times vary with temperature and operating voltage, among other factors, the times referred to here are not the min/max numbers printed in data sheets, but rather the actual times for the given set of operating conditions. Typical applications where such times are likely to be violated include bus and memory arbiters, interfaces, synchronizers, and other state machines employing asynchronous inputs or asynchronous clocks.

Metastability manifests itself in a number of different ways. Common responses are (shown as they might be captured on a digital oscilloscope in Figure 1): runt pulse (1a), decreased output slew rate (1b), output oscillation (1c), and increased clock-to-output time (1d). By definition, the phenomenon of metastability is statistical in nature. Not only is entry into the metastable state uncertain, but the time spent there can also vary.

Because PLDs are commonplace in today’s designs, a thorough understanding of their metastable behavior is crucial. In some applications, output anomalies shorter than one clock cycle may be acceptable, but in applications where the register output is used as a control signal (clock, bus grant, chip select, etc.) for other circuitry, faults such as runt pulses and oscillation cannot be tolerated.

This report will not study the causes or characteristics of metastability in great detail; excellent material has already been prepared on this subject [1-5]. Rather, this report will introduce a mathematical model for the metastable phenomenon, discuss potential test methodologies, present and compare test results from various bipolar and CMOS PLDs, and discuss how to interpret the data. This report will close with suggestions on how to design metastable tolerant systems.

Derivation of Constants

The basic premise of all metastability models is that a device’s output is more likely to have settled to a valid state in time(t) than in time(t-n). In fact, the failure probability distribution follows an exponential curve. Figure 2 shows a typical failure frequency plot.

It is accepted [1] that metastable failures can be accurately modeled by the equation:

$$\log \text{Failure} = \log \text{MAX} - b(\Delta - \Delta o) \qquad (1)$$

In this equation, MAX represents the maximum failure rate for a particular environment, Δ is the time delayed before sampling the DUT (Device Under Test) output, and Δo is the time at which the number of failures starts to decrease. On a failure frequency plot (such as the one in Figure 2), Δo represents the knee of the curve. The constant b is the rate at which the frequency of failures decreases after the knee is reached.

Recall that:

$$\log X = a \ln (X), \text{ where } a = \log (e)$$

Substituting this into (1):

$$a \bullet \ln \text{Failure} = a \bullet \ln \text{MAX} - b(\Delta - \Delta o) \qquad (2)$$

MAX is related to the clock frequency (fCLOCK) and data frequency (fDATA). That is,

$$\text{MAX} = (k1 \bullet \text{fCLOCK} \bullet \text{fDATA}) \qquad (3)$$

Substituting (3) into (2) and applying some algebra:

$$a \bullet \ln \text{Failure} = a \bullet \ln (k1 \bullet \text{fCLOCK} \bullet \text{fDATA}) - b(\Delta - \Delta o)$$

$$\ln \text{Failure} - \ln (k1 \bullet \text{fCLOCK} \bullet \text{fDATA}) = -b/a(\Delta - \Delta o)$$

Setting k2 = b/a and rearranging the equation yields:

$$\text{Failure} = (k1 \bullet \text{fCLOCK} \bullet \text{fDATA})e^{-k2\,(\Delta - \Delta o)} \qquad (4)$$

When used with equation (4), the constants k1, k2, and Δo, completely describe a particular device’s metastable characteristics; they indicate how quickly a device can resolve the metastable condition. Devices which transition out of the metastable region quickly are characterized by a small Δo and a large k2.

The constant k1 is peculiar to the test apparatus (it can be thought of as a “scaling factor”). The maximum metastable failure rate (MAX) is limited by fCLOCK; a failure cannot occur if the device isn’t clocked. Likewise, it is true that a metastable failure cannot occur unless data has changed. So, if fDATA < fCLOCK, then MAX

= fDATA. This was the case in the test fixture Lattice used (fCLOCK=10MHZ, fDATA=2.5MHz). Substituting MAX = fDATA back into equation (3) yields: k1 = 1/fCLOCK, so k1 = 100ns for our tests.

Test Fixture

The goal of testing a particular device's metastable characteristics is to generate real numbers for the constants k2 and Δo. To do this, the device must first be forced into the metastable state. This is done by intentionally violating setup and/or hold times. Once metastable, the output can be observed on an oscilloscope or used to increment an event counter.

Traditional Approach

One approach to characterizing a device's metastable behavior employs a test fixture similar to that shown in Figure 3a. In such a fixture, data to the device includes a "jitter band" so that the device sees changing data as it is clocked. The DUT output is fed to a window comparator to determine when it is in the metastable region (between V_{IL} max and V_{IH} min). The comparator output can be sampled periodically and used to increment an event counter.

This method of testing, though it directly yields MTBF numbers, has some drawbacks. The first is that it does not distinguish between the different types of metastable behavior (runt pulse, oscillation, slow rise/fall time, delayed transition), and it may have difficulty

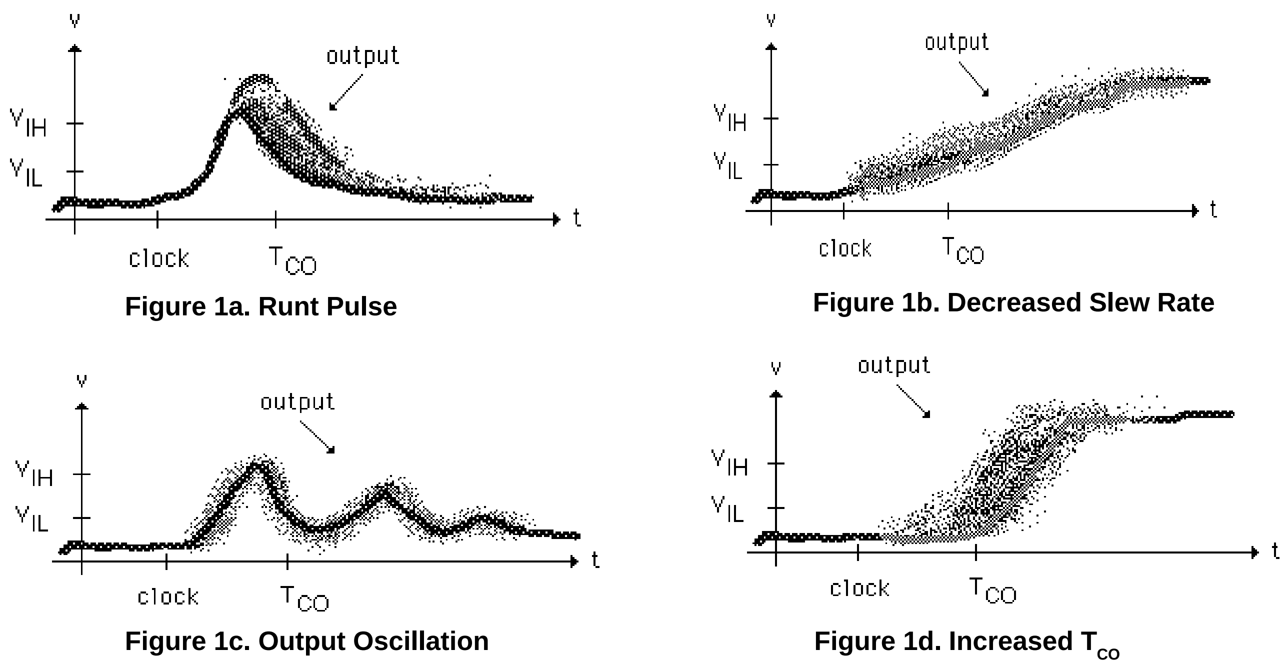


Figure 1a. Runt Pulse

Figure 1b. Decreased Slew Rate

Figure 1c. Output Oscillation

Figure 1d. Increased T_{co}

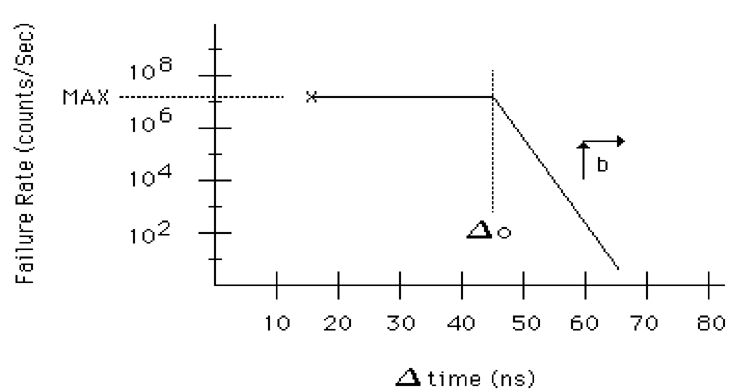


Figure 2. Typical Failure Frequency Plot

detecting every type. Also, the registers used in the detector circuit itself may become metastable, which would adversely affect the results.

A New Approach

The test method used to gather data for this report used the circuit shown in Figure 3b. The tester employed an “infinite precision” variable delay circuit to control clock placement with respect to data. This arrangement allowed exact worst case placement of the clock, so as to induce metastability with nearly every clock pulse.

Using a digital oscilloscope (Tektronix 11403A) in point accumulate mode, metastable failures were recorded over a lengthy period of time. A hardcopy was then made and the constants empirically obtained (details below).

The oscilloscope approach, being visual in nature, enables the designer to make educated decisions regarding maximum clock and data rates, as well as the suitability of using the output to drive other circuitry. The five minute sample period used in our tests contained approximately 750 million failures. Much longer sample periods were evaluated, but they provided no perceptible gain in usable information.

A slight disadvantage of this approach is that extracting k2 and Δo values from the hardcopies is not straightforward. Because each point on the hardcopy can represent any number of actual samples (between one and 1.5 million), one cannot simply count the points at time(t) for the MTBF at that time (although, in the case of the scattered points, the probability is low that a single isolated point represents more than one sample).

To generate values for k2 and Δo, it was necessary to refer to previous metastability studies [1]. By studying the output plots of devices with known constants, certain relationships were established. For example, it was determined that Δo represents the time from the leading

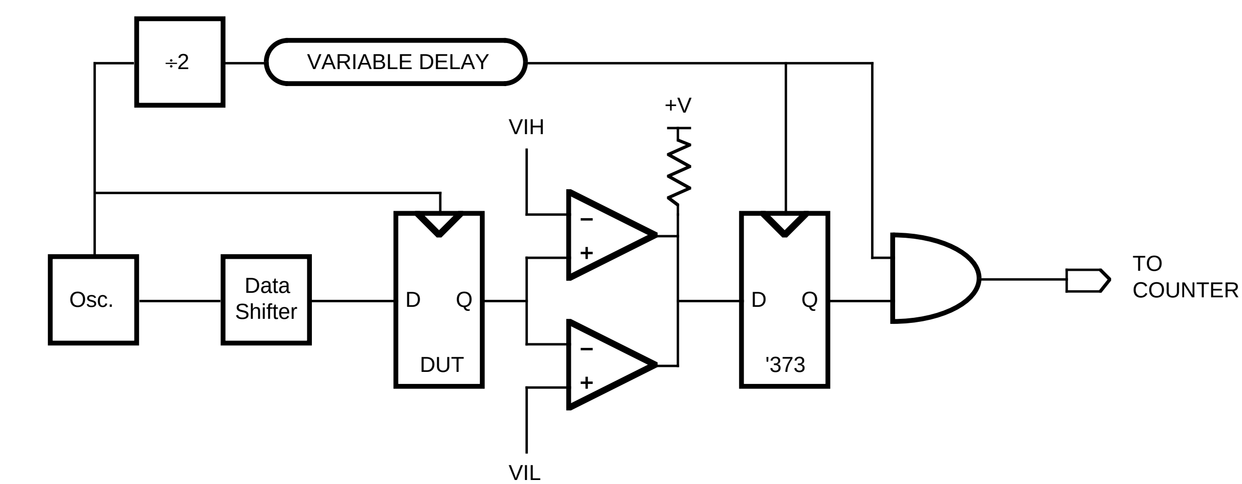


Figure 3a. Traditional Metastability Test Circuit

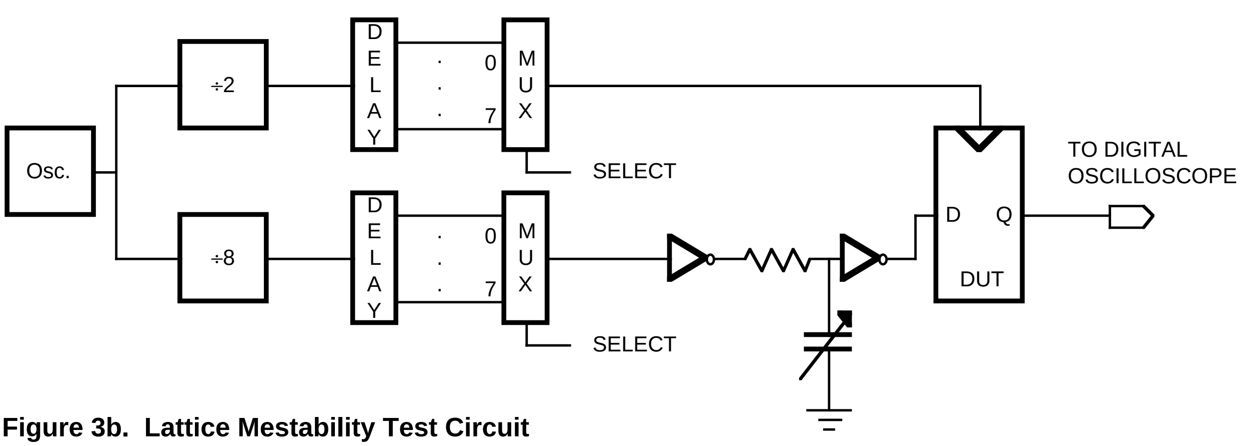


Figure 3b. Lattice Mestability Test Circuit

edge of the output until the “dot density” starts to decrease measurably. It should be noted that Δo in previous studies included device propagation delays, whereas in our test it does not.

The time from Δo until the dot density equals zero was defined to be the “time to metastable release” or simply time(r). The relationship between k2 and time(r) is given below in (5), and shown graphically in Figure 4. Recall that MAX=2.5x10^{6} and a=log(e).

$$k2 = \log(MAX) / (time(r) \cdot a) = 14.73/time(r) \qquad (5)$$

Interpreting the Results

In addition to examining E^{2}CMOS GAL devices, this study also tested several bipolar PAL devices as well as other CMOS PLDs. To insure that the results of this study would be relevant, all necessary precautions were observed: the devices were of recent vintage and were acquired blindly through distributors; multiple samples of each device were tested and the results combined; all devices had either fixed 16R8 architectures or were configured to emulate the 16R8 architecture; the devices were programmed from the same JEDEC fuse map file (the source equations and the JEDEC fuse map file are presented in Listing 1).

Plots 1 through 11 on the following pages are some of the oscilloscope plots generated for this study. The top waveform in each plot is the clock signal, the middle trace is the metastable data output and the bottom trace is the histogram of the accumulated samples between 1V and 2V of the output signal. The horizontal scale is 2ns per division, so the exact clock to output time of the metastable output condition can be read directly. The vertical scale is 2V per division for the top trace, and 1V per division for the middle trace.

The middle waveform in each plot is the metastable device output which is the only signal captured in point accumulate mode. In every case, the output signal plot shows two stable levels after the transition. This is a direct result of the “indecision” caused by metastability; on some cycles the output settled to a high level, while on others it settled to a low level.

Plot 9 shows the response of a bipolar PAL16R8-7. Notice the very well defined runt pulse (this correlates with previous data gathered on similar devices by the manufacturer [1]). The absence of a secondary trace along ground indicates that the output always starts to transition to a high level, even when it finally settles to a low level. This characteristic makes the device unsuitable for use in control path applications (when metastability is possible). All of the bipolar parts examined showed similar results.

Plots 4 through 8 are from GAL16V8C-5, ispLSI 1016-80, GAL16V8B-7, GAL22V10B-10 and GAL6002B-15, respectively. Aside from the fact that setup time violations may cause t_{CO} to increase by a small (but random) amount, the outputs are very clean and well behaved. The fact that there are no runt pulses or other anomalies is extremely significant, as the GAL6002B not only allows asynchronous clocking, but encourages that activity. Although GAL6002B is a much slower device as compared to GAL16V8 and GAL22V10, the similar metastable characteristics of the GAL6002B to the

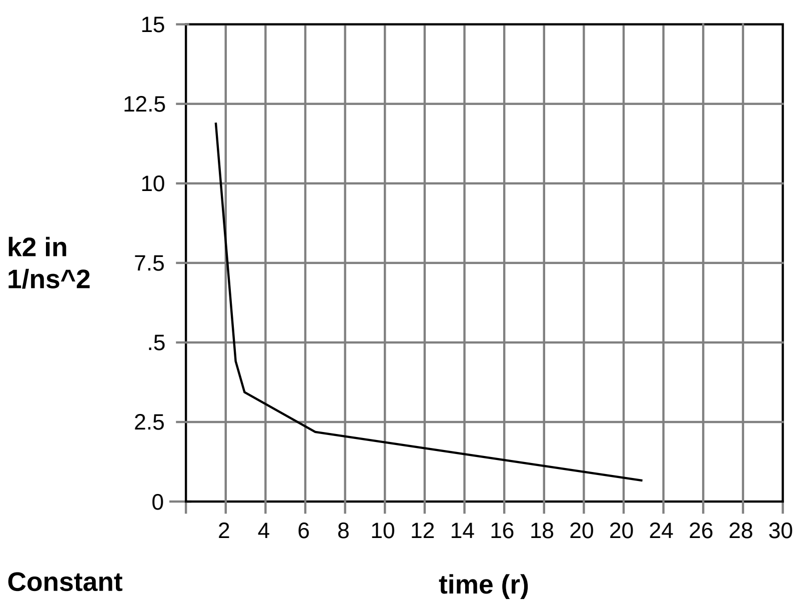


Figure 4. K2 Constant

much faster GAL devices indicate that the inherent metastable characteristics of all the GAL devices have consistently desirable characteristics across all speed grades. Comparing Plots 4 through 8 with Plots 9 and 10 shows that characteristics of the GAL devices are superior to those of bipolar PLDs. Plot 11 illustrates metastable characteristics of the TTL flip-flop (TISN74AS74).

For reference purposes, Plots 12 through 14 are included. Plot 12 shows a normal (i.e. non-metastable) GAL16V8B-7 transition, and Plot 13 a normal PAL16R8-7 transition. Plot 14 is the normal transition of the TTL flip-flop (TI SN74AS74). For consistency, only rising edges have been shown. Our tests also covered falling edges which, in general, were interesting but did not provide any additional information.

For a more quantitative look at the phenomenon of metastability, refer to the table beneath each plot. These tables list the measured values of the constants Δo and k2 for the device whose plot is shown, and for similar devices. Recall that large k2 and small Δo values are desirable. The numbers in the tables correlate closely with the results of earlier tests [1,5], confirming the validity of our test method.

Since all current GAL devices possess very similar register and output buffer circuitry, and all are fabricated using the same basic process, the data shown in Table 1 for the GAL16V8 is considered applicable to all devices and speed grades in the GAL family.

Using the Results

If a register enters the metastable state in a system, then data was obviously unstable as the register was being clocked. The argument over which data should have been captured (old or new) is academic as the register will randomly pick one or the other. Signals in most asynchronous systems are active for more than one clock cycle, so if they are missed initially, they could be captured on a subsequent clock cycle.

It is the task of the state machine designer to take adequate precautions against metastability causing illegal states to be entered. One way to do this is by using "gray codes" when ordering states. Gray code state equations allow only one state bit to change during a state transition. Thus, the worst metastability could do would be to delay a state transition by one clock cycle. If more than one bit were allowed to change, the outcome would be purely random, and probably illegal. Figure 5 shows examples of both cases.

Other solutions are to externally (or internally) synchronize the asynchronous signals, or to increase cycle times to allow time for metastable outputs to settle. An example of the latter solution is given below.

It is worth noting at this point that state machines (synchronous or asynchronous) can fail for reasons other than metastability. A not insignificant component of a PLD's specified setup time is directly attributable to internal data skewing [2]. Data skewing is the inevitable result of differing signal path lengths, loading conditions, and gate delays. Stated another way, each input to output path has its own set of actual AC specifications. If insufficient setup time has passed, different "versions" of the same data may be present at the inputs of different registers as they are clocked. A good example of this is:

```
Output_Pin19 := Input_Pin2;
Output_Pin15 := !Input_Pin2;
```

If clocked at precisely the right moment after an input transition, one register will capture old data while the other captures new data, resulting in a system failure. This condition, though also the result of a setup time violation, should not be confused with metastability (the "incorrect" data that is captured has normal output characteristics); it is, pure and simply, the result of a violation of specifications.

Example

To determine the maximum clock rate (given an acceptable error rate) that a particular device will allow in an asynchronous environment, equation (4) is used. For example, the system shown in Figure 6 utilizes a 9600 baud (bits/sec) asynchronous data stream. The system clock period is tCO+tPD+tSU+Δ. For one failure per year:

$$3.2\text{x}10\text{-}8 = [(1\text{x}10\text{-}7)(1/(\Delta+22))(9600)]e^{-[4(\Delta-.44)]}$$

Solving for Δ yields Δ=2.22ns, or about 2ns, for a cycle time of 24ns. Referring back to Plot 1, the additional delay of 2ns intuitively makes sense. Remember, in terms of setup and hold time violations, the oscilloscope plots were made under worst case failure conditions; the scattered dots could represent MTBFs of days, years, or even millenniums in a typical asynchronous environment.

Due to the extremely quick metastable settling times of GAL devices, a relatively small increase in the cycle time will produce a dramatic improvement in reliability.

Bibliography

1. D.M.Tavana (MMI), "Metastability - A study of the Anomalous Behavior of Synchronizer Circuits," in: Programmable Array Logic Handbook, Monolithic Memories Inc., 1986, pp 11-13 - 11-16.
2. K.Rubin (Force Computers), "Metastability Testing in PALs," Wescon/87 Conference Record (San Francisco, November 17-19, 1987). Los Angeles: Electronics Conventions Management, Inc, 1987, pp 16/1 1-10.
3. K.Nootbaar (Applied Microcircuits Corp.), "Design, Testing, and Application of a Metastable Hardened Flip-Flop," ibid., pp 16/2 1-9.
4. J.Birkner (MMI), "Understanding Metastability," ibid., pp 16/3 1-3.
5. R.K.Breuninger, K.Frank, "Metastable Characteristics of Texas Instruments Advanced Bipolar Logic Families," application note SDAA004, Texas Instruments, 1985.

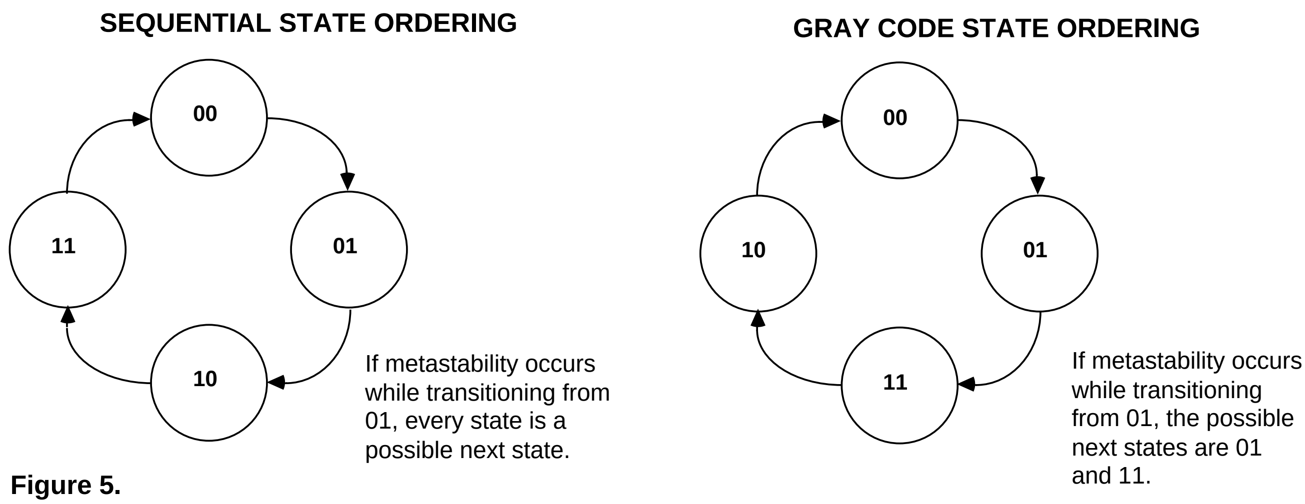


Figure 5.

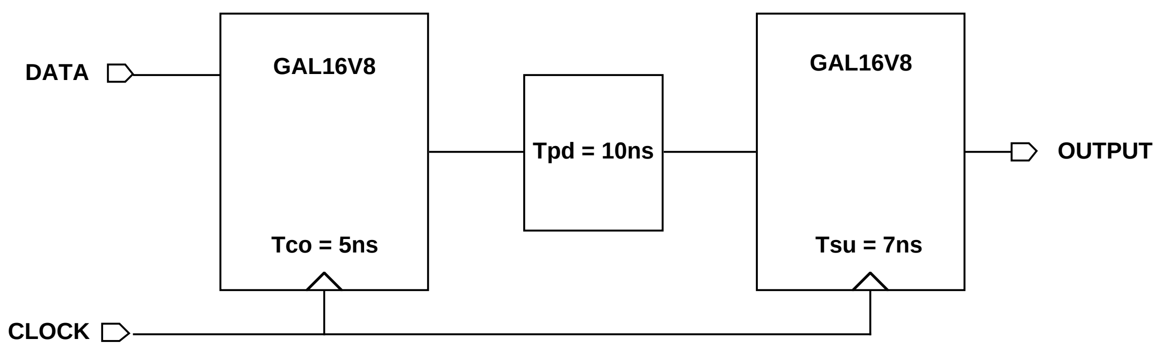


Figure 6.

```
MODULE metastable
TITLE 'Metastable Test
Pattern'
uOO Device 'P16R8';
d       PIN   2;
q1,q2   PIN   12,19;
        EQUATIONS
q1 := d;
q2 := d;
        End metastable
```

Listing 1a. Source equations

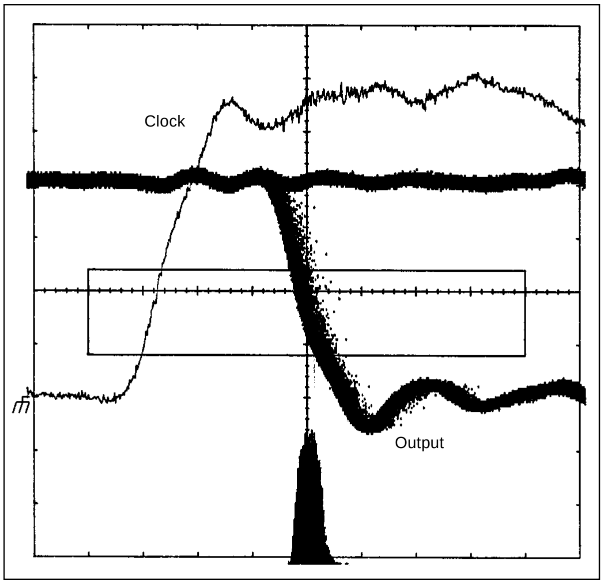
```
JEDEC file for: P16R8
Metastability Test Pattern*
QP20* QF2048* F0*
L0000 10111111111111111111111111111111*
L1792 10111111111111111111111111111111*
C07F4*
```

Listing 1b. JEDEC file

Listing 2. ispLSI 1016 Metastability Test Source Equation from Lattice pDS Software

```
//
// metastbl.ldf generated using Lattice pDS Version 2.20

LDF 1.00.00 DESIGNLDF;
DESIGN metastbl;
REVISION 00;
AUTHOR ;
PROJECTNAME METASTABILITY STUDY;
PART pLSI1016-80LJ;
OPTION Y1_AS_RESET ON;
DECLARE
END;  //DECLARE
SYM GLB  B7  1  ;
SIGTYPE IMOUT REG OUT;
SIGTYPE IROUT REG OUT;
EQUATIONS
   IMOUT.CLK=ICLK;
   IROUT.CLK=ICLK;
   IMOUT.D = IMIN;
   IROUT.D = IRIN;
END;
END;

SYM IOC  IO31  1  ;
XPIN IO MOUT LOCK 10;
OB11 (MOUT,IMOUT);
END;

SYM IOC  IO30  1  ;
XPIN IO ROUT LOCK 9;
OB11 (ROUT,IROUT);
END;

SYM IOC  IO29  1  ;
XPIN IO MIN LOCK 8;
IB11 (IMIN,MIN);
END;

SYM IOC  IO28  1  ;
XPIN IO RIN LOCK 7;
ID11 (IRIN,RIN,ICLK);
END;

SYM IOC  Y2  1  ;
XPIN CLK XCLK LOCK 33;
IB11 (ICLK,XCLK);
END;
END;  //LDF DESIGNLDF
```

1V/div

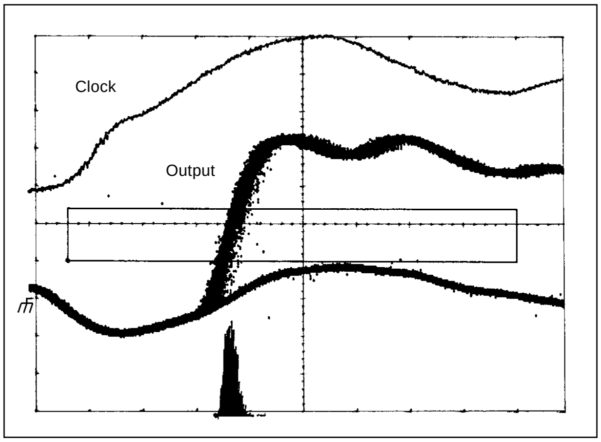


2ns/div

Plot 1. ispLSI 2032 Metastable Output

Part #	Manufacturer	∆ o (ns)	k2 (1/ns²)
ispLSI 2032	Lattice	.986	13.9

1V/div

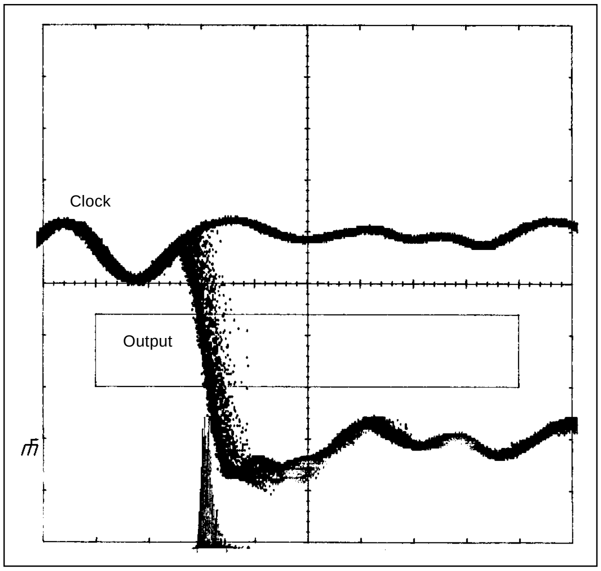


2ns/div

Plot 2. ispLSI 2032LV Metastable Output

Part #	Manufacturer	∆ o (ns)	k2 (1/ns²)
ispLSI 2032LV	Lattice	1.044	13.9

1V/div

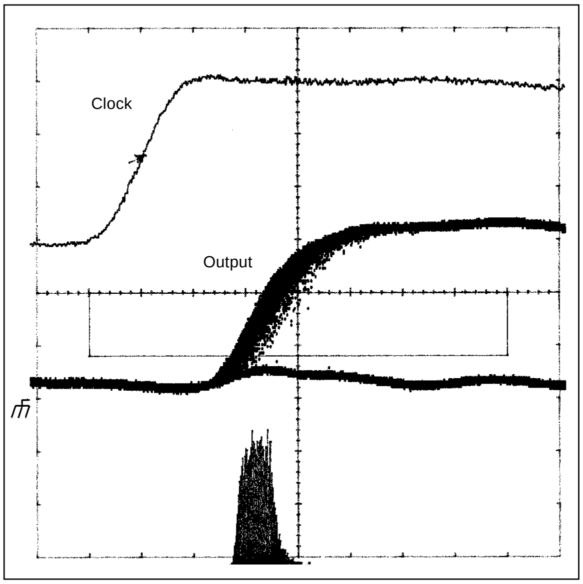


2ns/div

Plot 3. ispLSI 3192 Metastable Output

Part #	Manufacturer	Δ o (ns)	k2 (1/ns²)
ispLSI 3192	Lattice	.772	13.9

1V/div

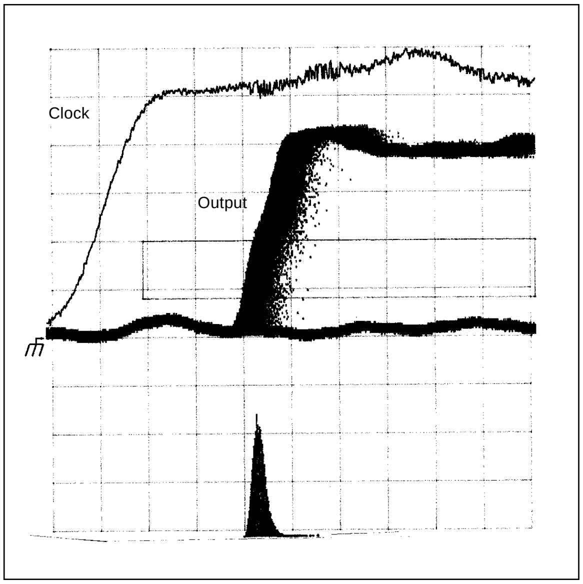


2ns/div

Plot 4. GAL16V8C-5 Metastable Output

Part #	Manufacturer	Δ o (ns)	k2 (1/ns²)
GAL16V8C-5	Lattice	1.4	9.82

1V/div



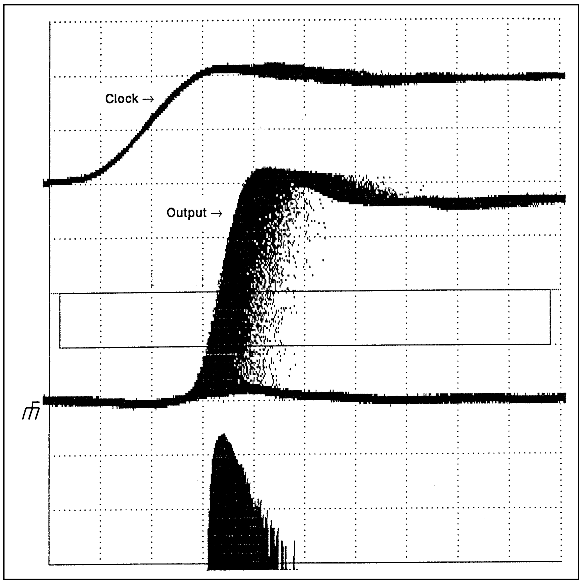
2ns/div

Plot 5. ispLSI 1016-80 Metastable Output

Part #	Manufacturer	Δ o (ns)	k2 ($1/ns^2$)
ispLSI 1016-80	Lattice	.854	11.0

2V/div

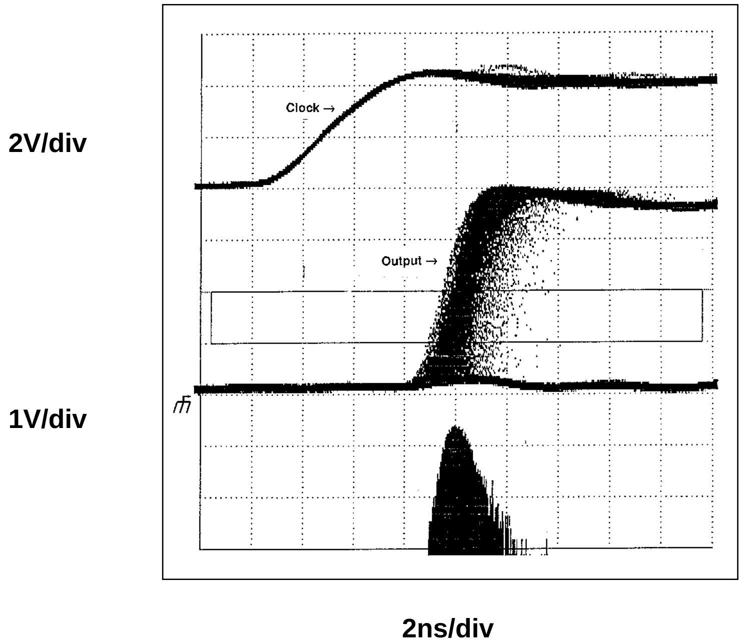
1V/div



2ns/div

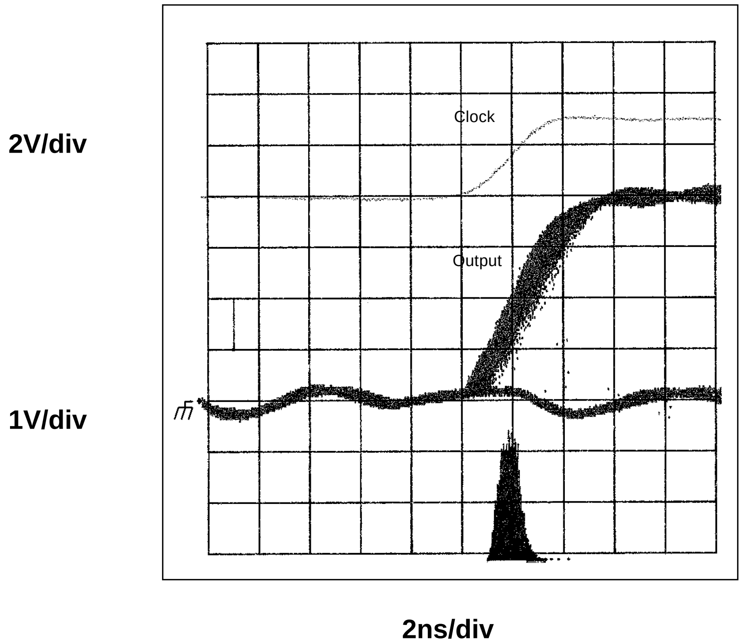
Plot 6. GAL16V8B-7 Metastable Output

Part #	Manufacturer	Δ o (ns)	k2 ($1/ns^2$)
GAL16V8B-7	Lattice	.44	5.0



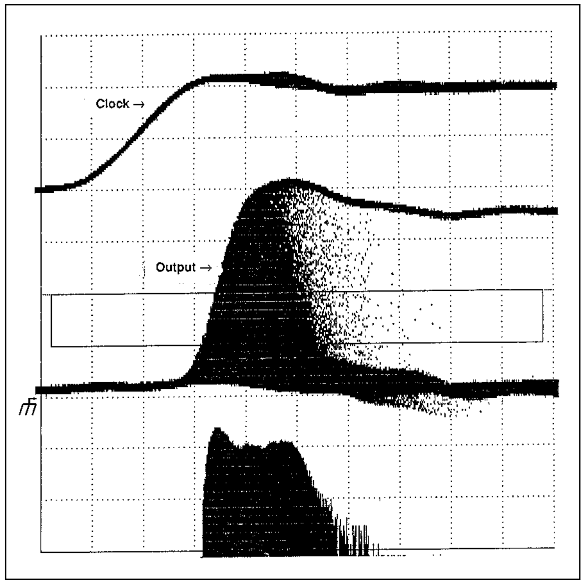
Plot 7. GAL22V10B-10 Metastable Output

Part #	Manufacturer	Δo (ns)	k2 ($1/ns^2$)
GAL22V10B-10	Lattice	.51	5.2



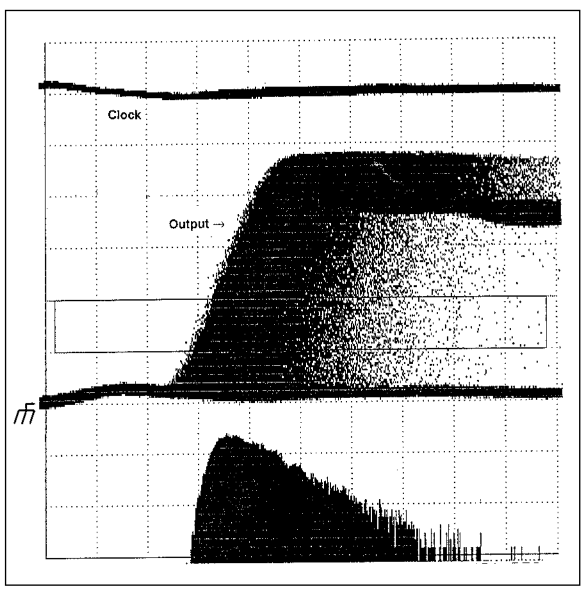
Plot 8. GAL6002B-15 Metastable Output

Part #	Manufacturer	Δo (ns)	k2 ($1/ns^2$)
GAL6002B-15	Lattice	1.1	6.52



Plot 9. PAL16R8-7 Metastable Output

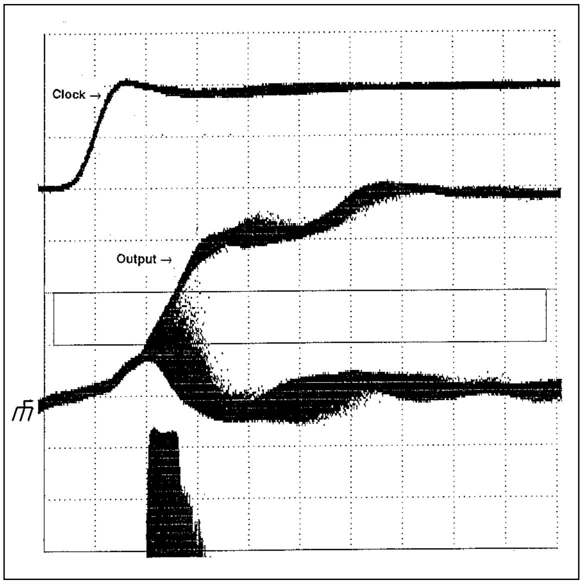
Part #	Manufacturer	Δ o (ns)	k2 (1/ns²)
PAL16R8-7	AMD	1.2	2.5



Plot 10. TIBPAL16R6-7 Metastable Output

Part #	Manufacturer	Δ o (ns)	k2 (1/ns²)
TIBPAL16R6-7	TI	1.5	1.5

2V/div



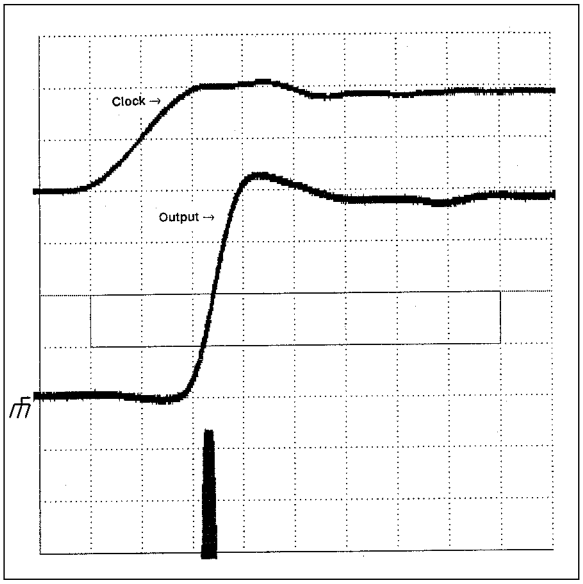
1V/div

2ns/div

Plot 11. SN74AS74 Metastable Output

Part #	Manufacturer	Δo (ns)	k2 (1/ns²)
SN74AS74	TI	.91	3.5

2V/div



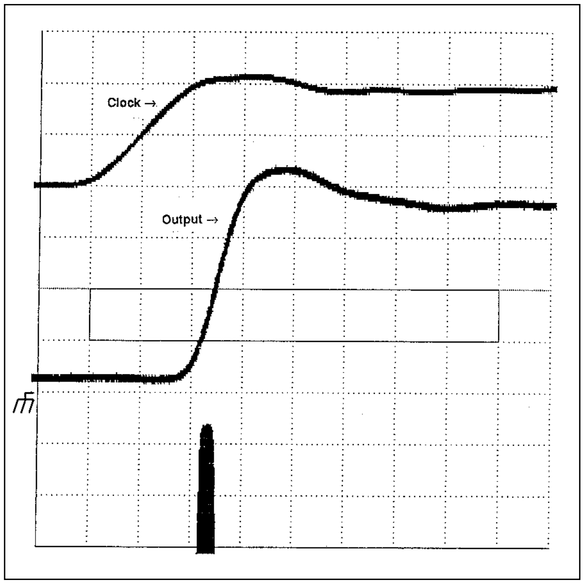
1V/div

2ns/div

Plot 12. Normal GAL16V8B-7 Transition

2V/div

1V/div

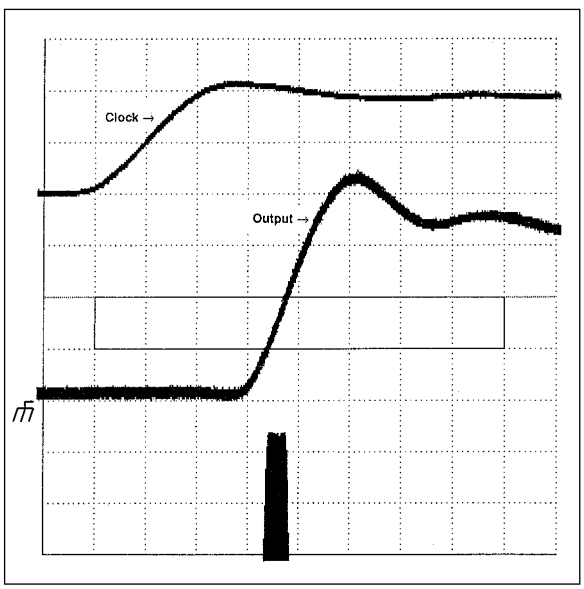


2ns/div

Plot 13. Normal PAL16R8-7 Transition

2V/div

1V/div



2ns/div

Plot 14. Normal SN74AS74 Transition

LATTICE SEMICONDUCTOR CORPORATION
5555 Northeast Moore Court
Hillsboro, Oregon 97124 U.S.A.
Tel.: (503) 681-0118
FAX: (503) 681-3037
http://www.latticesemi.com

November 1996