POWER4

Synchronous Wave-Pipelined Interface

IBM

Authors: Frank Ferraiolo

Edgar Cordero

Daniel Dreps

Michael Floyd

Kevin Gower

Bradley McCredie

Worse Case Timing for a Typical Interface

Timing Parameters	pS
Clock Skew on Transmit Side	100
Silicon on Transmit Side	1,000
*Wiring - drv to chip pad, drv module, card(40cm), rcv. module, chip pad to rcv.	3,000
Silicon on Receive Side	1,000
Clock Skew on Receive Side	100
Coupling, Jitter, reflections	700
PLL - long term jitter	250
TOTAL	6,150

Bus Cycle Time = W.C. Path + the time to next bus clock = < 150 Mhz

Timing Variations for a Typical Interface

Timing Parameter Variation	pS
Clock Skew on Transmit Side	100
Silicon process on Transmit Side	600
Wiring - drv to chip pad, drv module,card,rcv. module, chip pad to rcv.	500
Silicon process on Receive Side	600
Clock Skew on Receive Side	100
Coupling, Jitter, reflections	700
PLL - long term jitter	250
TOTAL	2,850

Bandwidth & Cycle Limited ~ 200Mhz where the tgt cycle = 1.5 or 2

Conventional Interface Timing

1a) The period of the bus clock is slightly greater than the worse case latency of the interface.

1b) The minimum latency of the interface must be greater than the skew between the sending and receiving latch pairs. (trivial)

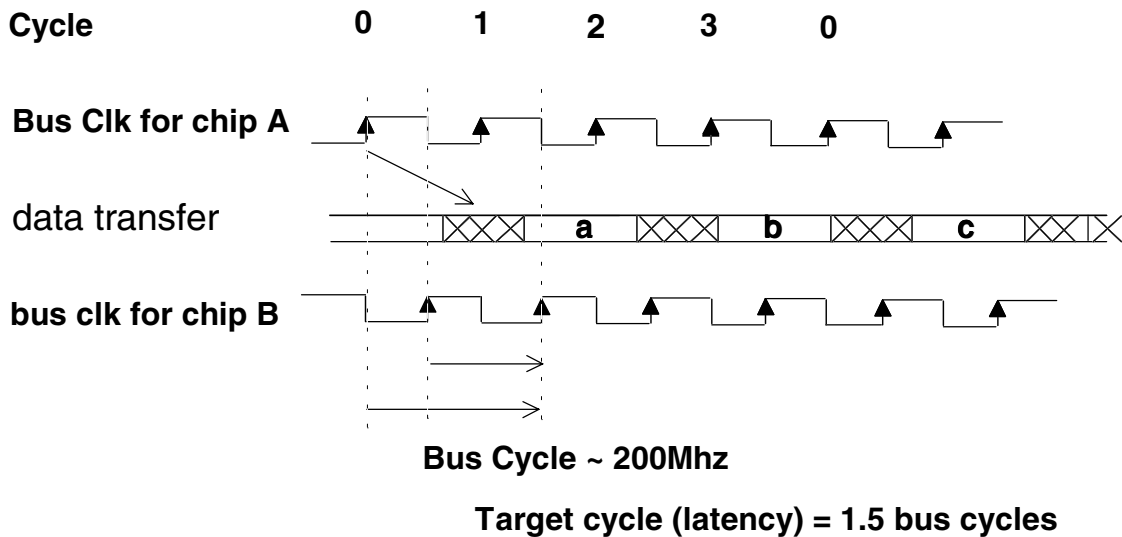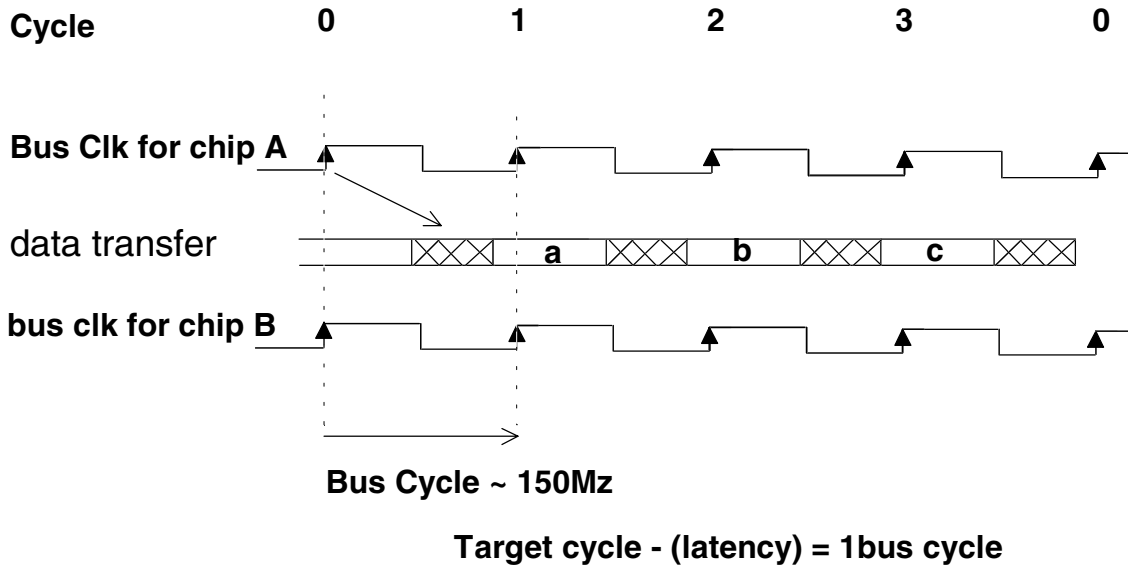Note the data is always captured (target cycle) in the first bus cycle.

Very Limited Use of Pipelining, better described as Bus "Tuning"

2a) The data's target cycle (capture cycle) was the second cycle and the card wire or silicon path was "tuned" to ensure the proper timing.

2b) The best case latency > (target cycle -1). The fast and slow case are equally important.

Note the data had to arrive in the same bus cycle.

Conventional Interface Timing

Cycle   0   1   2   3   0

Bus Clk for chip A

data transfer   a   b   c

bus clk for chip B

Bus Cycle ~ 150Mz

Target cycle - (latency) = 1bus cycle

Cycle   0   1   2   3   0

Bus Clk for chip A

data transfer   a   b   c

bus clk for chip B

Bus Cycle ~ 200Mhz

Target cycle (latency) = 1.5 bus cycles

Bus speed f {interconnect latency,timing variations, processor speed}

Power4 Design Goals

Generic High Performance Interface

Current Design > 500 MHz

Next generation PowerPC processor - 1Ghz I/O

Minimize latency - Synchronous transfer

In order to allow multiple system configurations the latency between chips must be a variable.

The latency of the interface can vary over a wide range (multiple bus cycles) while maintaining synchronous operation.

--> Transfer rate (bus cycle time) is made nearly independent from the latency.

--> **Increased Performance / Bandwidth** - Reduce the amount of timing variations without imposing stringent and costly manufacturing or design constraints.

Power4 I/O Design Goals

Point to point, uni & bi-directional bus types

Wide Bus Widths

Multiple System Configurations

VLSI Compatible - All Digital Design

Low Power - Source Terminated Drivers

   - Active Clamps on Receivers

Easily Mapped between Technologies

Easily Customized to application
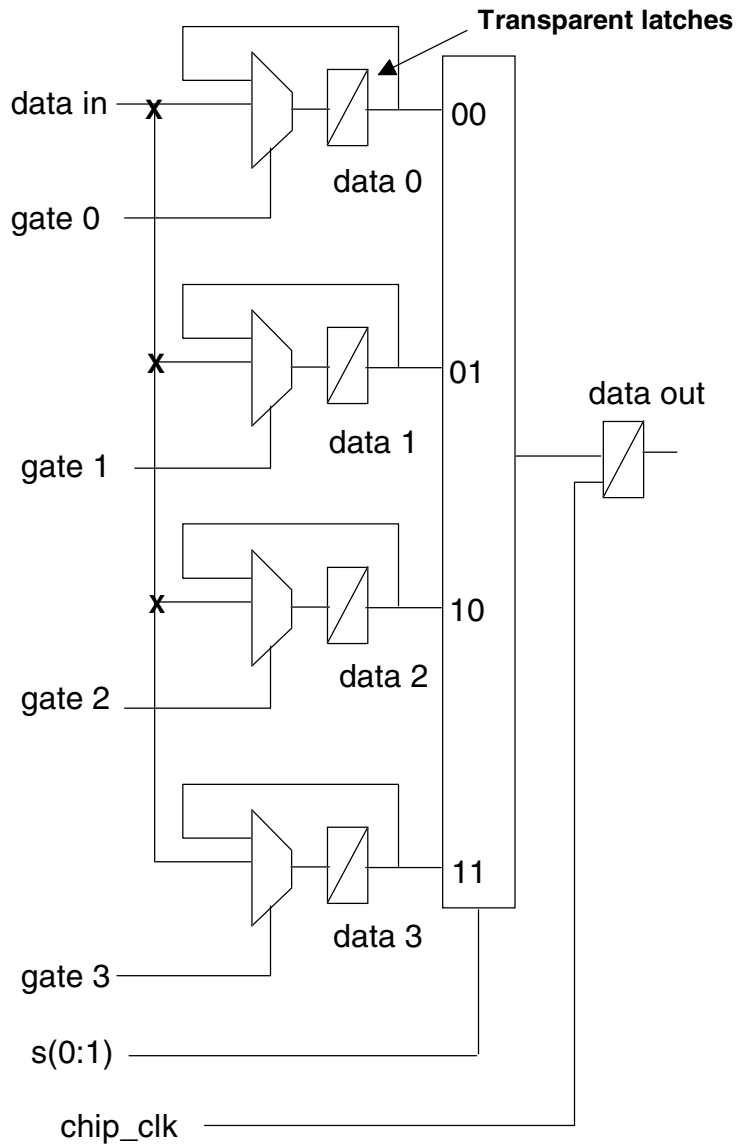
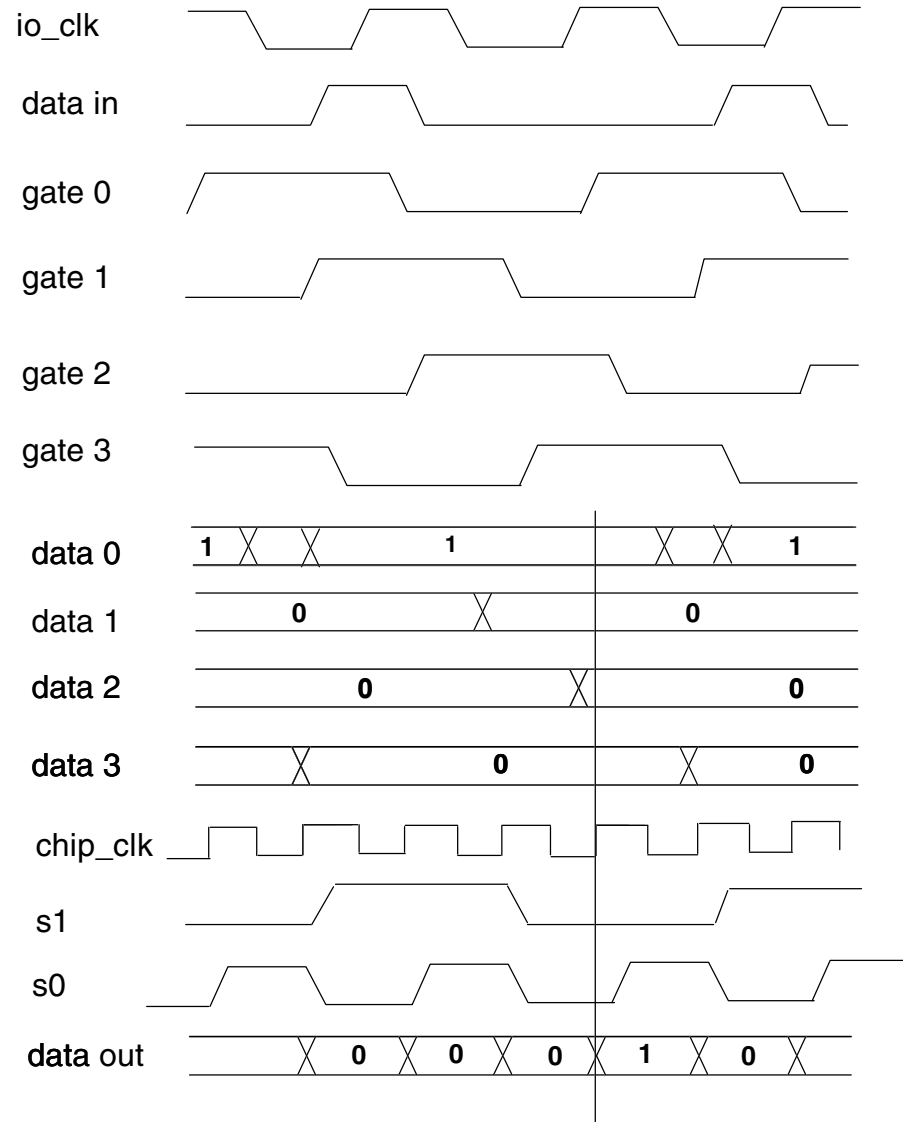 - Bus width, Speed, Distance, skew

Power4 Design Point

Latency:

FIFO on Receive Side

- Initialized at Power On

- Provides multiple bit times of data valid (Elasticity) allowing for a wide range of arrival times.

- Synchronous

- Programmable "target cycle"

- Minimal Added latency

Receive Side FIFO:

Transparent latches

data in

data 0
gate 0

data 1
gate 1

data out

data 2
gate 2

data 3
gate 3

s(0:1)

chip_clk

00
01
10
11

Signals after alignment

io_clk

data in

gate 0

gate 1

gate 2

gate 3

| data 0 | 1 | | | 1 | | | 1 |

| data 1 | | 0 | | | | 0 | |

| data 2 | | 0 | | | | | 0 |

| data 3 | | | 0 | | | | 0 |

chip_clk

s1

s0

| data out | | 0 | 0 | 0 | 1 | 0 | |

Note: gates are shifted one bit time at a time until a 1 is detected in data 0 and a 0 is detected in data 3. Thus, after alignment, the '1' in the IAP pattern (time 0) is captured by latch 0.

* target cycle

Elasticity Facilitates

NOT PRESENTED

o Synchronization and Timing

- Synchronization is maintained over a much wide range of latencies. The bandwidth of the interconnect determines the speed of the bus.
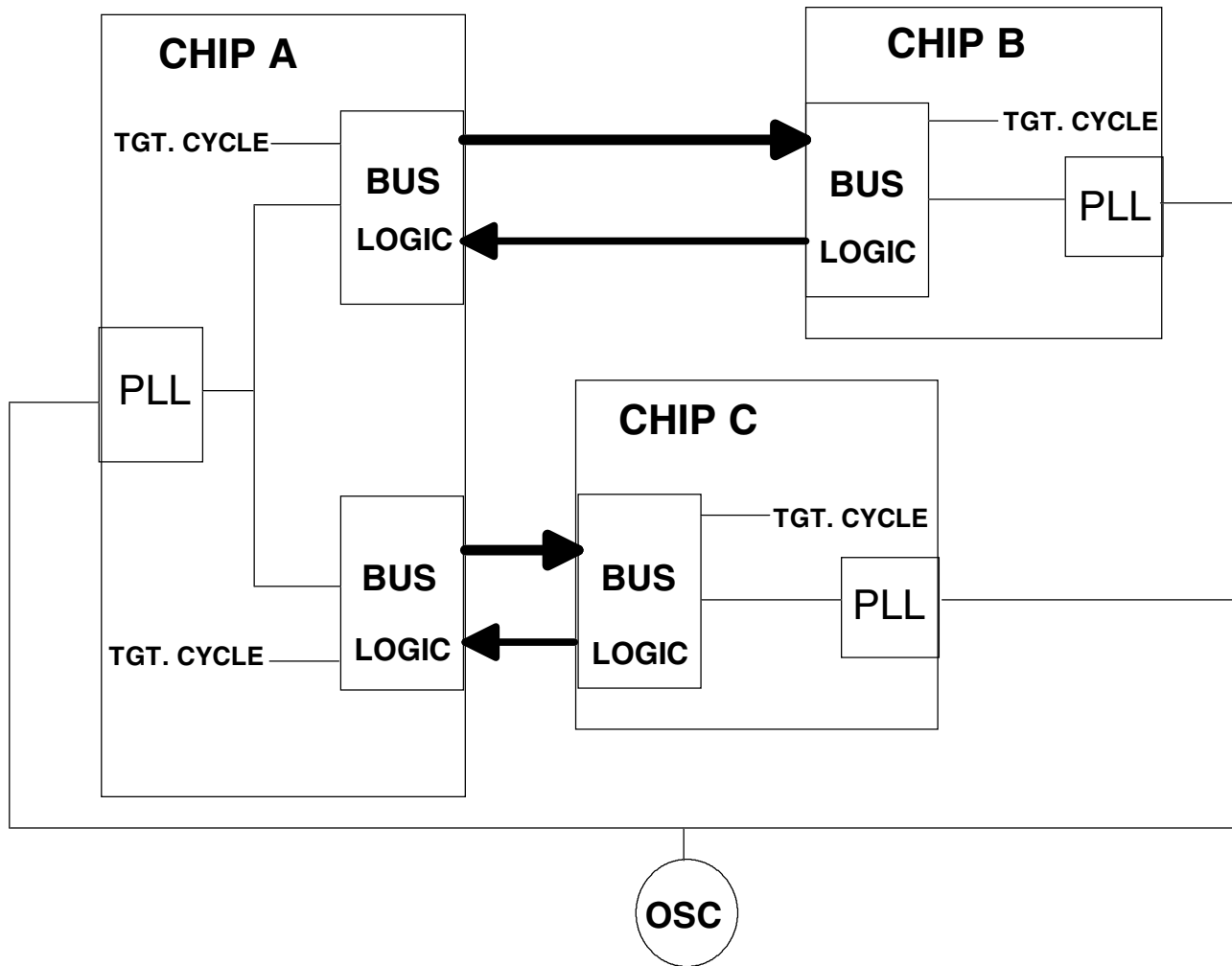
- Target time (capture cycle) is programmable (MOD4).

o Chip, Module & Board Wiring - Elasticity is used to synchronize chips which are at different electrical lengths. No need to "tune" the timing to a particular chip or balance the wiring between chips.

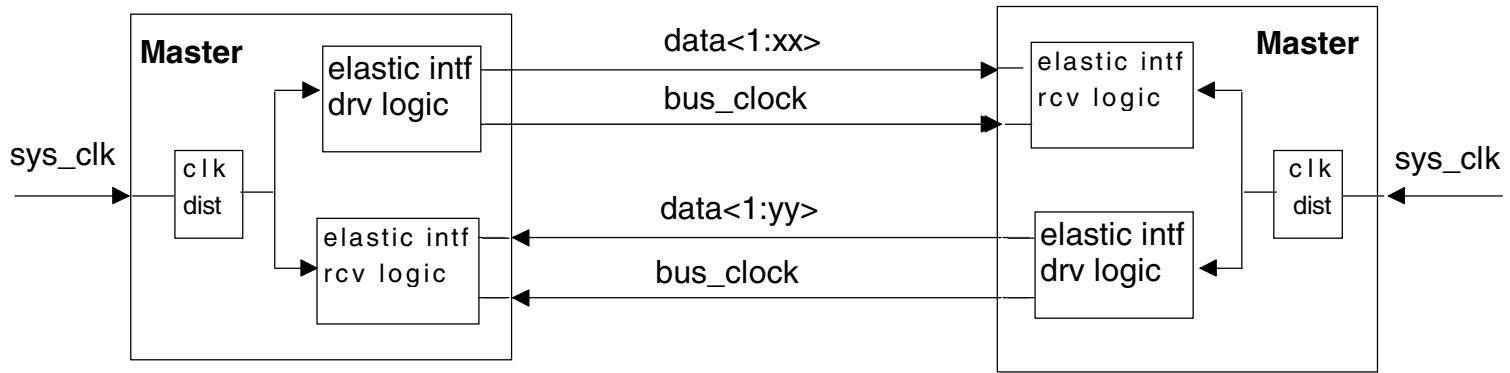o Can be used with receive chip with or without a PLL - i.e. SRAM (Slave) Chip Timing

o Multiple System Configurations can be easily supported and optimized without significant redesign and timing analysis.

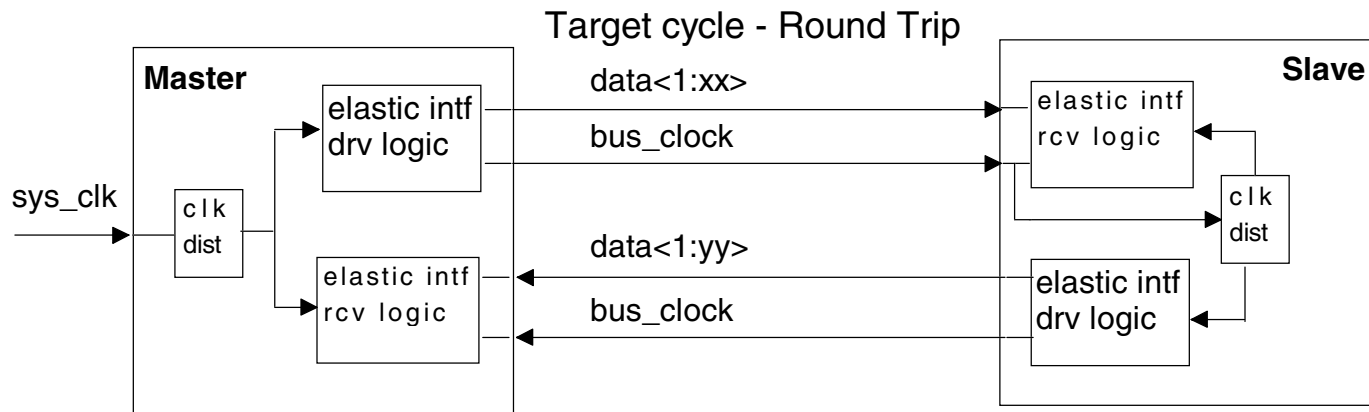o The bus speed more easily scales with the speed of the processor.

System Configuration

CHIP A

TGT. CYCLE

BUS

LOGIC

PLL

TGT. CYCLE

BUS

LOGIC

CHIP B

TGT. CYCLE

BUS

LOGIC

PLL

CHIP C

TGT. CYCLE

BUS

LOGIC

PLL

OSC

Master to master - both chips are synchronous, common time reference

sys_clk

Master

clk dist

elastic intf drv logic

data<1:xx>

bus_clock

elastic intf rcv logic

Master

clk dist

sys_clk

elastic intf rcv logic

data<1:yy>

bus_clock

elastic intf drv logic

Master to slave to master - Slave's time reference is initialized @ power on

Target cycle - Round Trip

sys_clk

Master

clk dist

elastic intf drv logic

data<1:xx>

bus_clock

elastic intf rcv logic

Slave

clk dist

elastic intf rcv logic

data<1:yy>

bus_clock

elastic intf drv logic

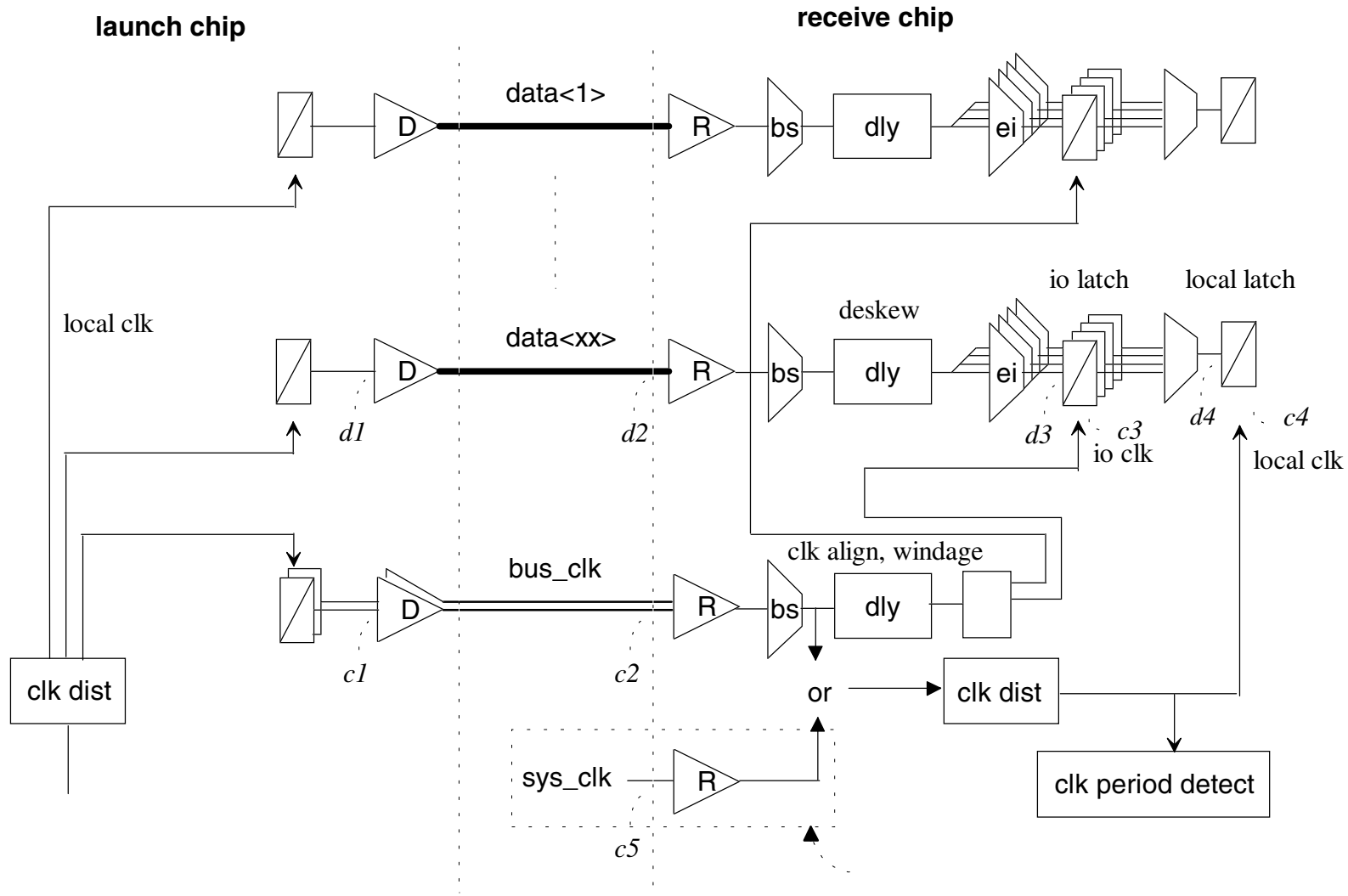**slave's chip clock derived from interface**

Power4 Design Point

Increase Bandwidth

- Reduced clock skew on Send & Receive Side

- Reduced Silicon Variation on both sides

- Reduced Wiring Variations - chips, module, card

- Eliminate Long Term PLL Jitter

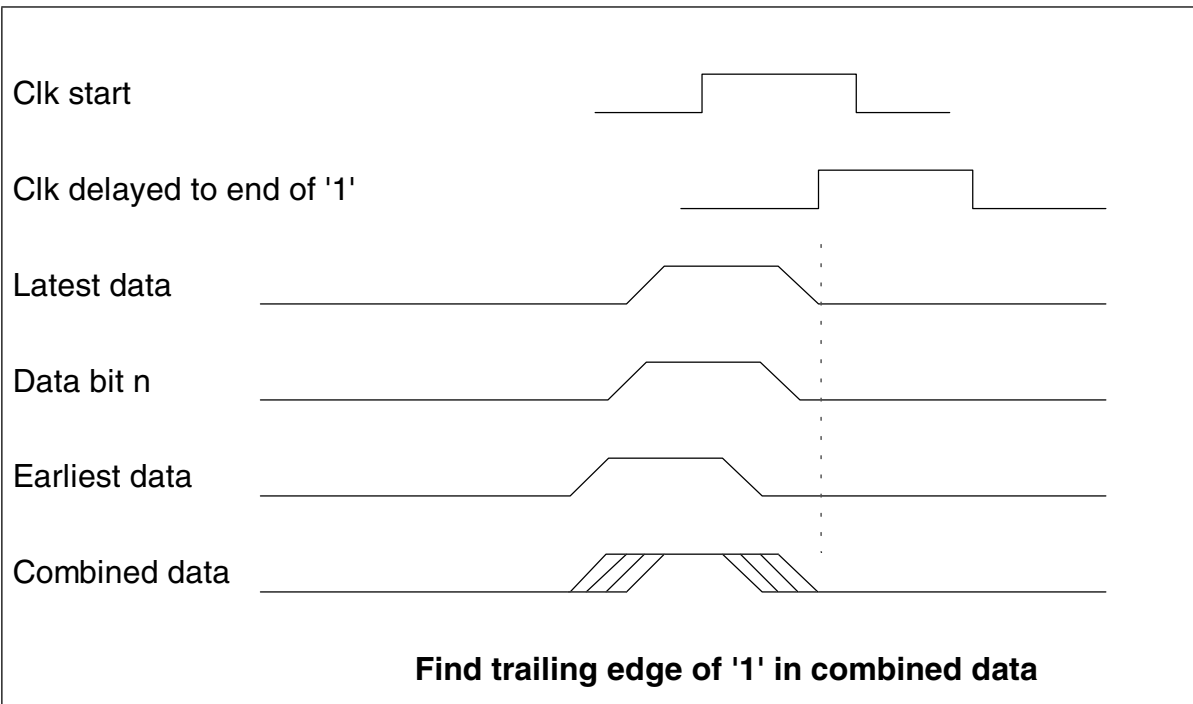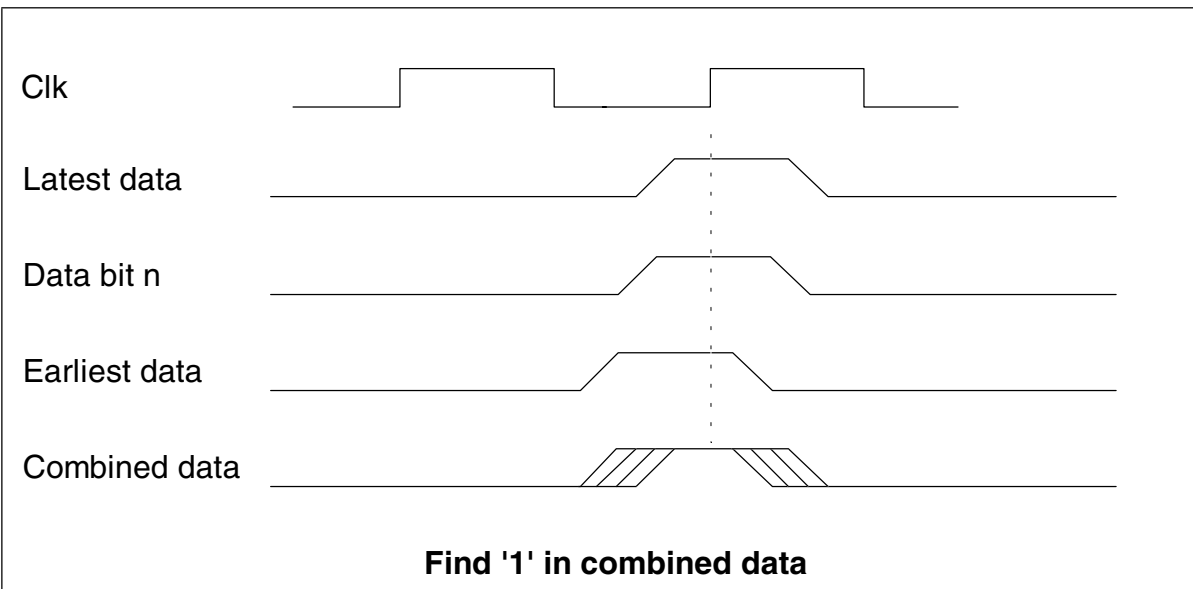## Achieved By

- Source Synchronous Interface - Clock is edge aligned w/ data

- Optimizing the phase of the clock at the sampling latch

- Per Bit Deskew

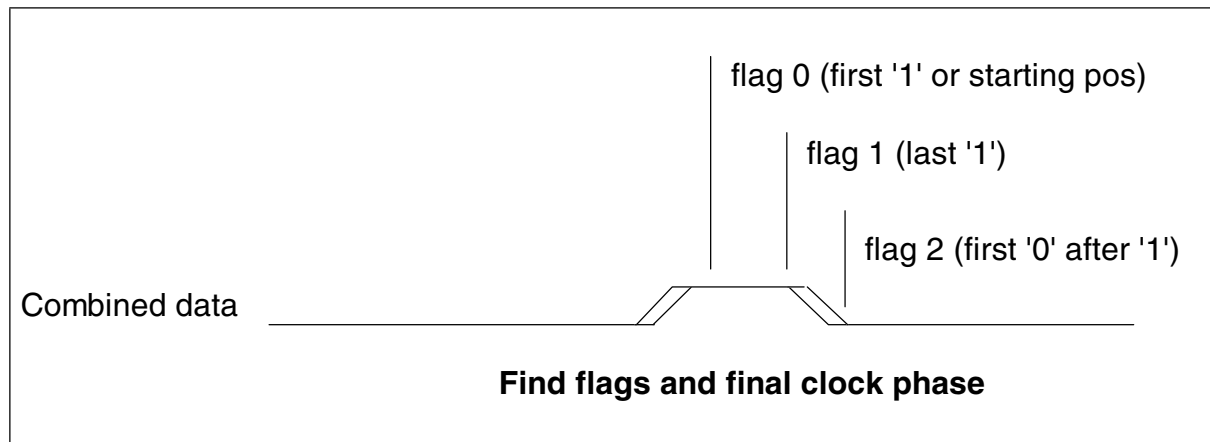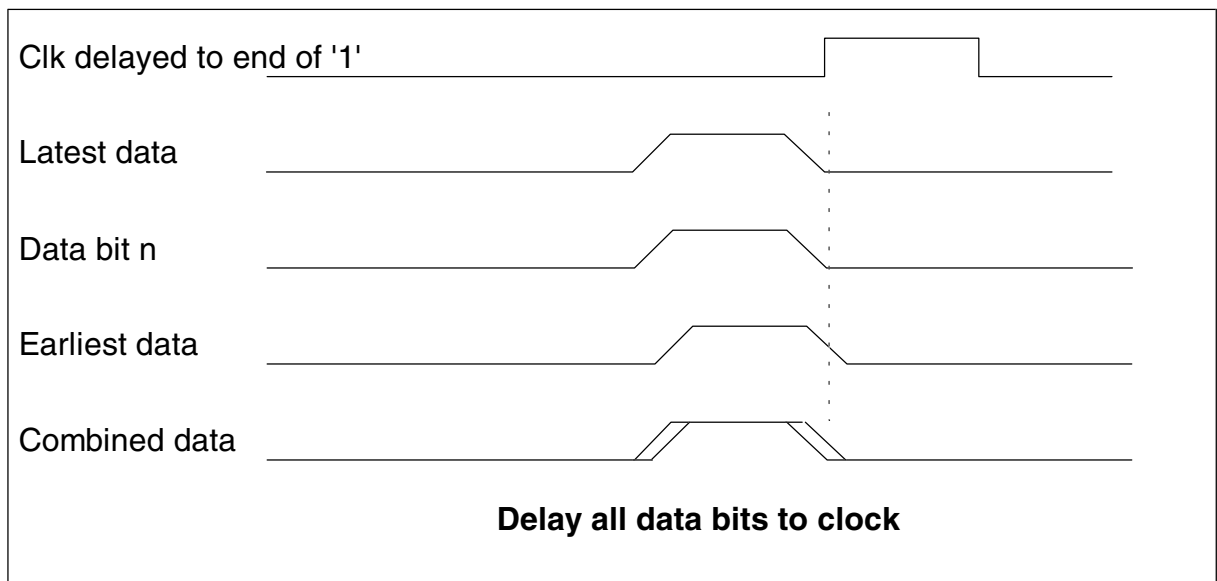# POWER4 Wave-Pipelined Receive Macro

ALGORITHM

Find '1' in combined data

Clk
Latest data
Data bit n
Earliest data
Combined data

Find trailing edge of '1' in combined data

Clk start
Clk delayed to end of '1'
Latest data
Data bit n
Earliest data
Combined data

ALGORITHM Con't

Clk delayed to end of '1'

Latest data

Data bit n

Earliest data

Combined data

Delay all data bits to clock

flag 0 (first '1' or starting pos)

flag 1 (last '1')

flag 2 (first '0' after '1')

Combined data

Find flags and final clock phase

High Speed Operation

Clock Sample = (flag1 + flag2 - Tc)/2 where Tc = bit time

Low Speed Operation

Clock Sample = (flag0 + flag1)/2
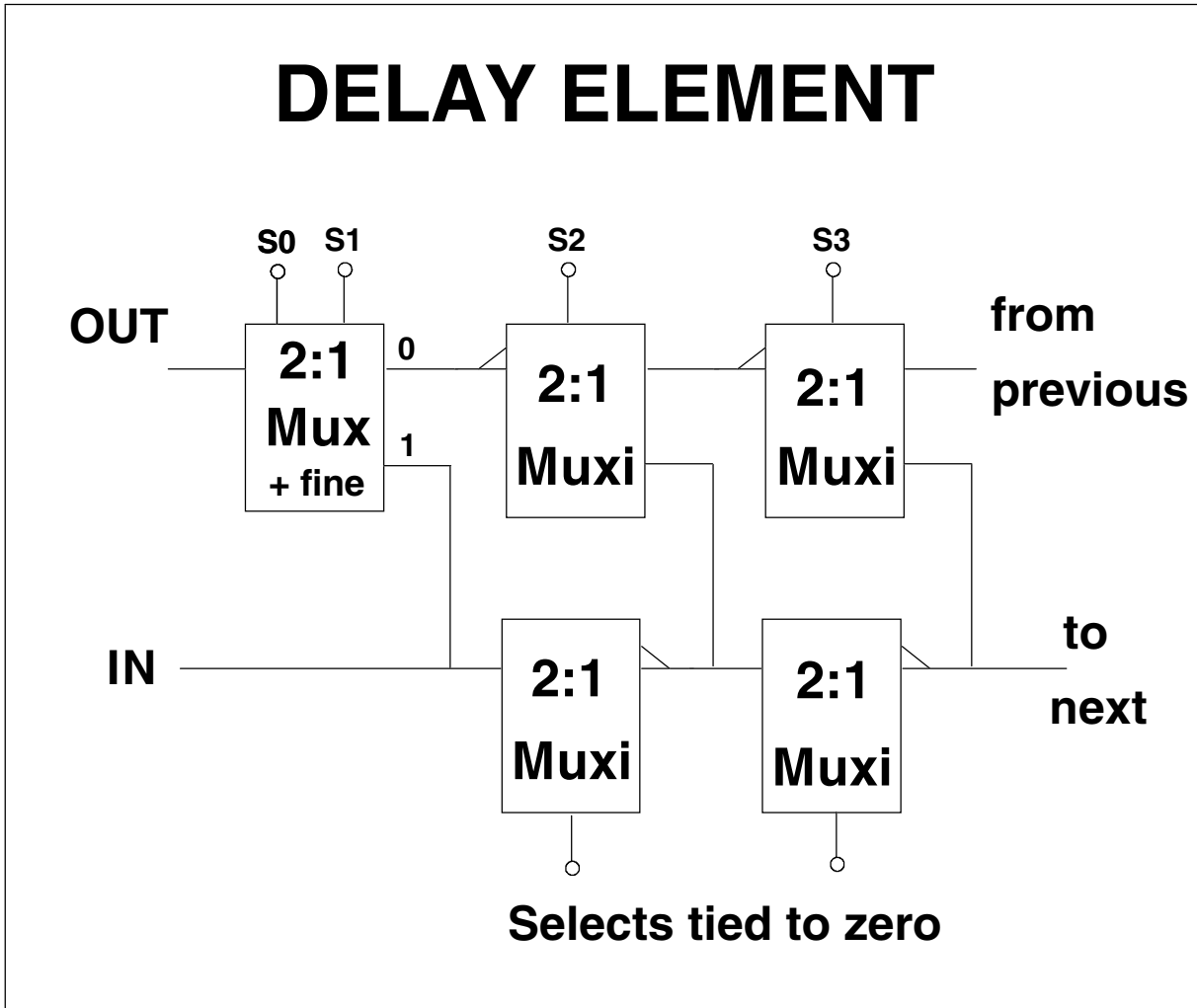
DELAY ELEMENT

Minimum Insertion Delay - one fine delay element

High Resolution achieved with scaling of device channel lengths

Monatomic Delay

High Bandwidth / Low Pulse Distortion

Scales with Technology

Trade off - Resolution vs. Deskew Range Vs. Area

Timing Variations for a Typical Interface using the POWER4 I/O Design

Timing Parameters	pS
Clock Skew on Transmit Side	25
Silicon process on Transmit Side	0
Wiring - drv to chip pad, drv module,card,rcv. module, chip pad to rcv.	0
Silicon process on Receive Side	0
Clock Skew on Receive Side	25
** Coupling, Jitter, reflections	700
PLL - long term jitter	0
TOTAL	750

STATIC SKEW COMPENSATION IS LIMITED TO ~ 800pS

** Note this term highly dependent on the distance of the interface. It can be significantly reduced with far end termination

Example Data Valid for a POWER4 Interface

Timing Parameters	pS
**Temp. & Power Supply Drift	+/-200
Clock Duty Cycle Distortion (+/-5%)	+/-100
*Clock Sampling Pt. Resolution	+/-135
*Per Bit Deskew Resolution	90
Latch Set Up Time	100
TOTAL	1,060

* This assumes a 90pS delay step under worse case conditions

** Future Design points will track Temp. & Power supply changes - The clock calibration circuit continuously operates. An update occurs upon issue of a reload command. A reload is accomplished in 24 bit times. With a 10uS interval this would impact the bandwidth by ~ .5%

Enhanced Diagnostics

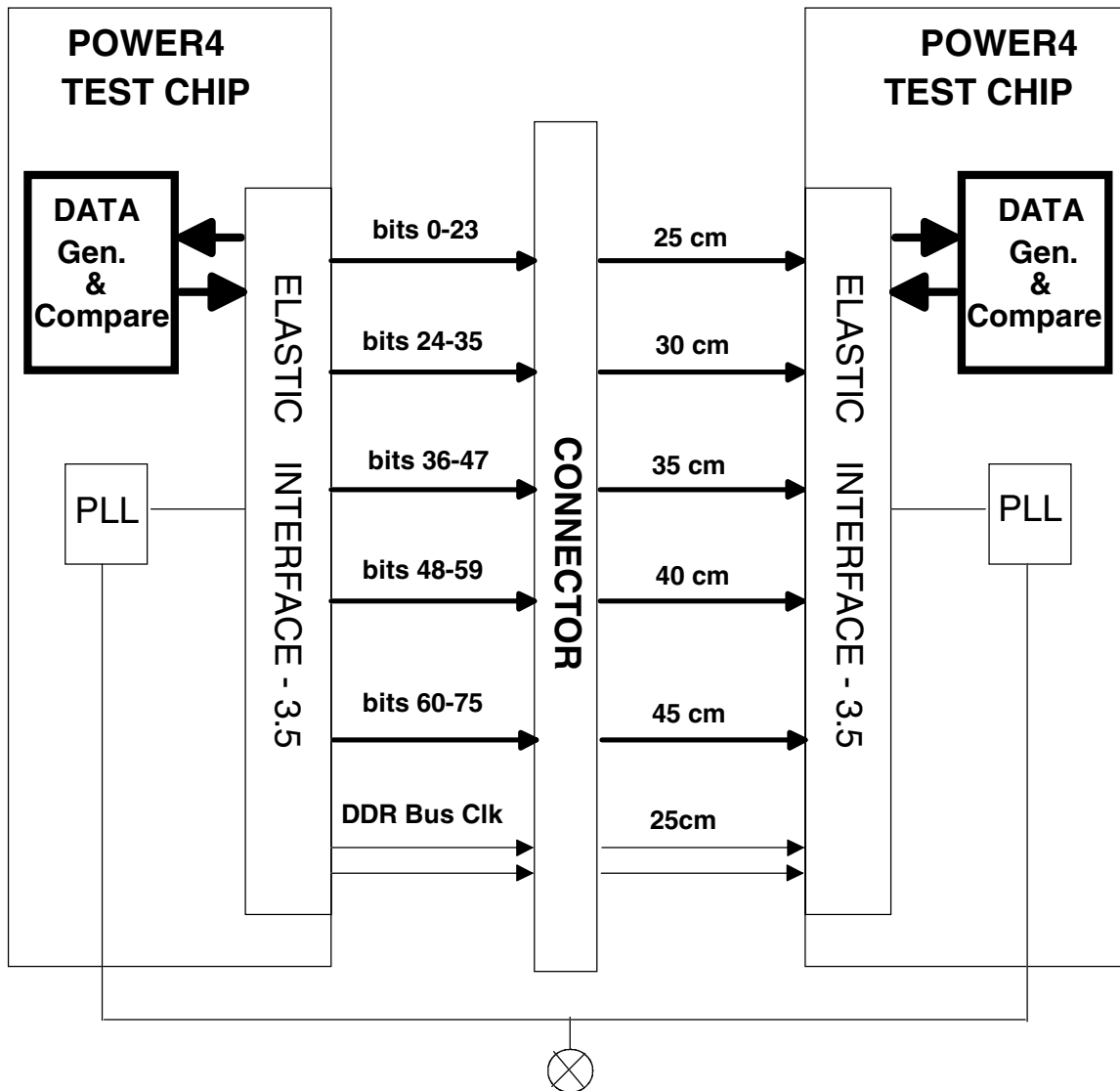**Reading out the register settings after initialization**

flag0 = clock position relative to latest arriving bit

flag2 - flag1 = data uncertainty region

max. - min. per bit deskew = skew between data bits

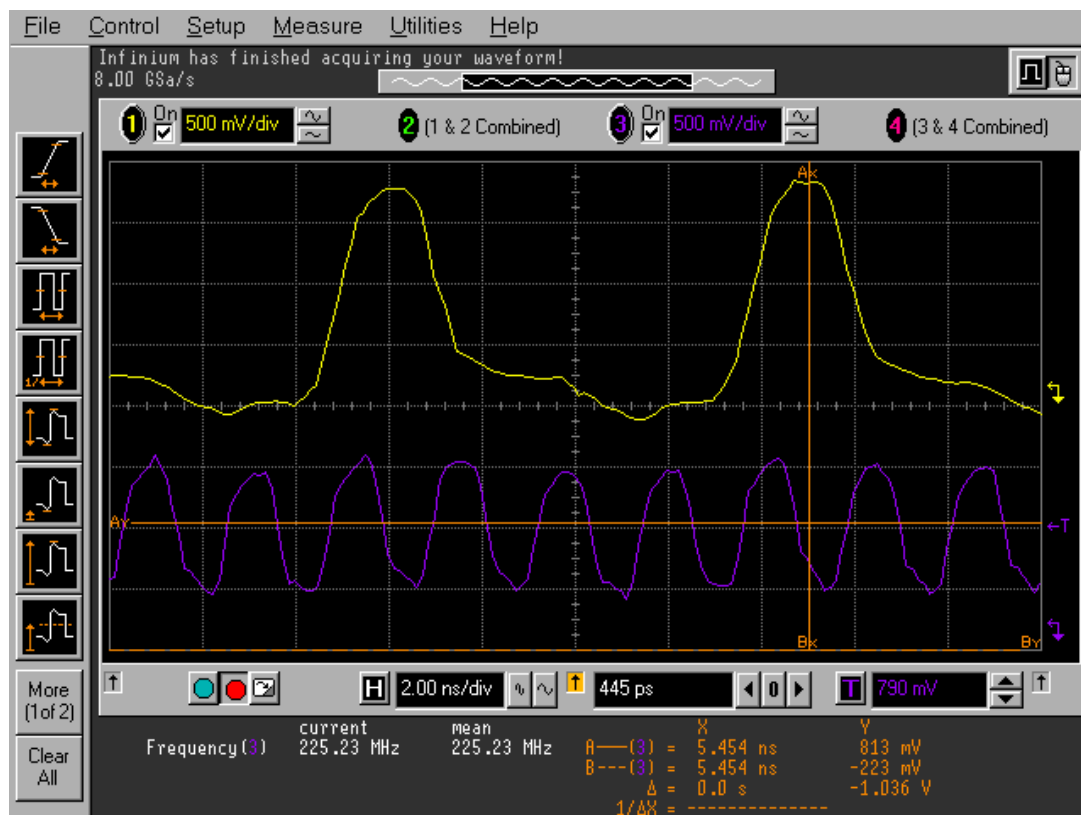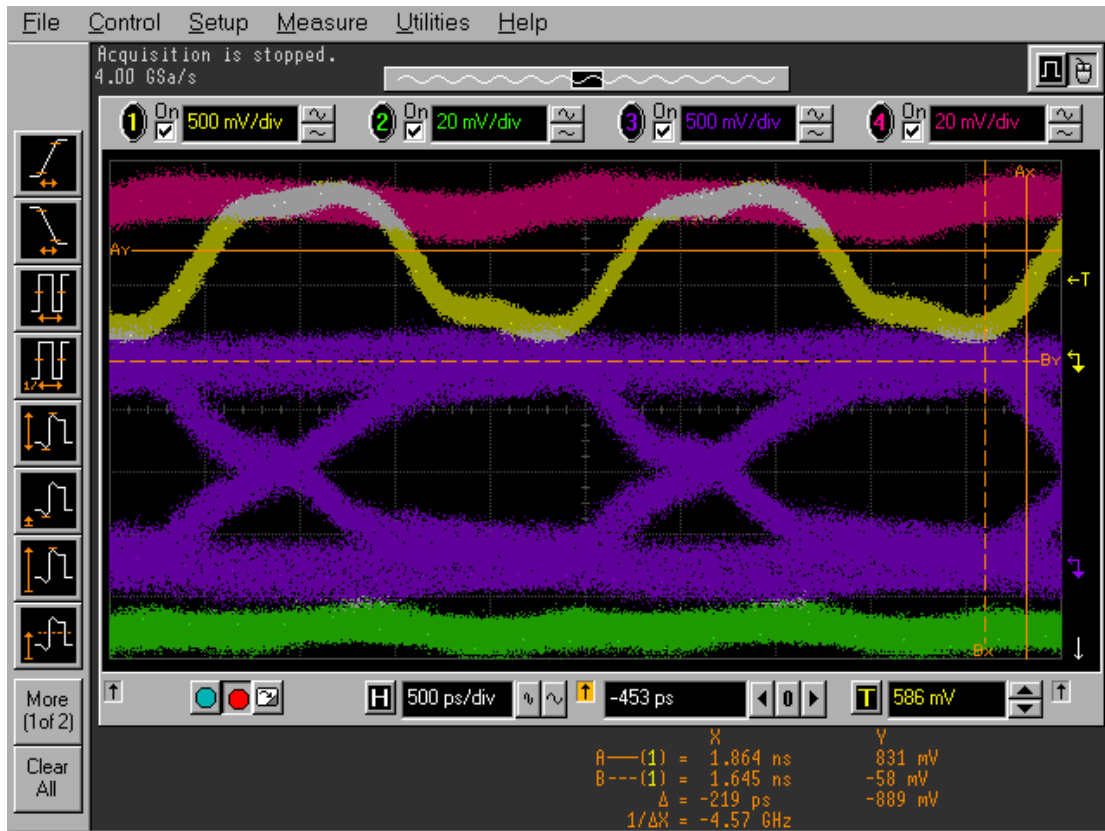**clock calibration = # delay elements in a clock period**

POWER4 TEST CHIP BOARD WIRING

WIRING SHOWN FOR ONE DIRECTION ONLY

Distance shown for total board wires with zero cable length

Below is a random data pattern with the associated bus clock at 30cm board trace, and the IAP pattern. Clamps are used on the receiver to provide soft termination.

POWER4 TEST CHIP RESULTS

2ns Operation

Skew on Data bits	Dly Element Cal.	Flag-0	Flag-1	Flag-2	Clock Delay pS	Data Valid pS	Bus Width	**Bus Length cm
2	39	2	28	41	15	1,337	24	25/20
15(max.)	39	15	41	55	28	1,286	48	35/20
15(max.)	38	24	32	63	28	357	73	45/20 *
2	39	10	27	49	18	787	24	25/70

* Initialization Only

** Bus length - first number is cm of card wire and second number is cm of Teflon cable