

**This page is not part of document, please discard if its
inadvertently printed (The Claude Clause).**

Current Date: September 25, 1996 Modification Date: August 19, 1996 5:22 pm
Directory: /project/gbn_hw/switch/docs/tech_rpts/se/APP-NOTE-Timing-deskewers.fm

June 27th, 7:20 Started document, base for doc is SE report

July 4: cc_tap is not mentioned, what does it do? Is it useful with present chip design
answer from Tom, included to allow forcing integer cell_clk time per cell in the
SE but not needed. Can also provide compensation for clock skew.

ARL Working Note ARL-96-04

Timing Parameters for GigaBit Switch Chips (SE, IPP, and OPP)

Thomas J. Chaney

Fred U. Rosenberger

September 25, 1996 (Version 0.1)

This document describes timing parameters and usage conditions for the SE, IPP, and OPP of the Washington University Gigabit Switch described in "System Architecture Document for GIGABIT SWITCHING TECHNOLOGY," Version 3.1, Working Note ARL-94-11, by Jonathan S. Turner and staff.

Introduction

This document describes timing parameters and guidelines for the GigaBit Switch chips, including conditions which must be satisfied for operation of the deskewer circuits incorporated at some of the data inputs to these chips. These deskewer circuits eliminate the need to satisfy setup and hold times at the input to a chip with respect to a particular chip clock transition. Thus the clock distribution need not have balanced delays to each chip, any phase relation between chip clocks is acceptable. This is accomplished by sampling the inputs three times per clock period, and then choosing one of those samples based on when a control signal changes level with respect to the three sampling times. For a detailed description of the deskewer operation see: "A Digital Adjustment Circuit for ATM and ATM Like Data Formats", by Thomas J. Chaney, Working Note ARL-96-08. Specific numbers and times given in this document apply to the initial design of the SE, IPP, and OPP using the ES2/Atmel ECLP07 process with Vdd between 3.3 and 3.6 V and chip temperature between 15 and 80 C. The following sections give a brief description of the timing parameters, and the values are listed in Table 1. Portions of this document, including some parameter values, are not yet completely determined and are indicated by "??".

System Clock

Each chip has a 120 Mhz input clock (CLK) which must have duty factor between 40% and 60% and rise and fall times < 2ns. All CLK signals must be derived from a single source in order to avoid phase shifts between clocks with slightly different frequencies. Any constant phase shift between the CLK inputs to different chips is acceptable but changes in the phase shift during operation may cause the deskewers to fail. Slow changes in phase shift, such as those caused by 60 or 120Hz power supply ripple, or caused by IC circuit heating, will not cause the deskewers to fail.

Input Signal Timing

While the initialization and grant input signals have setup and hold times with respect to CLK, their values are sampled on only one out of every 16 positive clock transitions (equivalent to approximately a 7.5 Mhz clock) so the setup and hold times are easy to satisfy. They would be difficult or impossible to satisfy if the inputs were sampled every CLK period. The setup and hold time values for the initialization and grant input signals are $t_s = -3.??\text{ns}$ (negative value means the input value must be stable by 3.?? ns *after* the clock transition) and $t_h = 11.??\text{ns}$.

The cell clock input signal (CELL_CLK) which indicates the beginning of an ATM cell and which occurs once every 16 clock periods has setup (t_{s_cclk}) and hold (t_{h_cclk}) times with respect to the clock of: $t_{s_cclk} = 1.75\text{ns}$ and $t_{h_cclk} = 1.0\text{ns}$. This is the only signal which must meet a setup and hold time with respect to every positive clock transition.

Output Signal Timing

All output signals, with two exceptions, are driven by the Atmel ECLP07 Standard Cell family OPS1T pad drivers. Normal use in the GigaBit Switch is with a 61 ohm, plus or minus 7 ohms, transmission-line load using the pad-driver output resistance for series termination. There are specifications in Table 1 that apply for this case.

The two output signal exceptions are: (1) The two bi-directional signal pins in the IPP (RESET_REQ and CLR_ERR); and (2) The OPP to IPP signals (SOC_OPP, D_OPP,31 . .0>, and PARI_OPP). The RESET_REQ and CLR_ERR signals are driven by open-drain outputs with an external pullup resistor (typically a 1k ohm resistor to +3.3v), and have very relaxed timing requirements with acceptable rise and

fall times of microseconds or more. It is expected that the OPP to IPP data paths will always be short, one to two inches at most. Thus these signals are driven (from the OPP) using the weaker OPS0T pad drivers.

Deskewing

At the high clock rates used in the GIGABIT SWITCH (120MHz, approximately an 8.3ns clock period), very careful control of the PC board trace lengths and propagation delays would be required to transmit data between chips without the deskewer circuits. While the deskewer circuits dramatically relax the over all system time matching of clock and data paths, in order to successfully use the deskewer circuits, some constraints must be satisfied in the timing of signal "groups". Each group has one control signal. This signal sometimes also conveys one bit of data, depending on the specific group. At power-up or reset time, the control signal must have synchronization¹ transitions for initialization before any data values are used, and synchronization transitions at some minimum rate if long term clock shift (jitter) is to be compensated for. These transitions are used to determine when to sample the value for all data signals in the group. Thus the skew between the control signal and any other data signals in the group must be kept small enough that the correct values will be sampled. Sampling is controlled by a three phase clock generated from the receiver-chip input clock by internal-logic circuit delays, thus the phase to phase delay is not very well controlled. The minimum phase to phase delay under the extremes of operating conditions and process variations is 1.3ns.² To insure that the correct data values are sampled, the maximum skew between each control signal transition and corresponding data value transition must be less than 1.3ns, the minimum time between samples. The specific signal used as the control signal in each group is identified in the chip specific signal description documents.

Sources of Skew

Skew at the receiving-chip inputs is the sum of the driving-chip output skew (maximum difference in transition time between a control signal positive transition and the either direction of data signal transition), the skew introduced by differences in propagation delays on paths between the driving chip and receiving chip flip-flops, and the skew between the receiving-chip and driving-chip clock signals. The source for skew at chip outputs is dominated by the difference between the H-L and L-H propagation delays in the chip final flip-flop and output pad circuits since all output signals in a group are driven from flip-flops clocked on the same internal clock signal and the on-chip data path delays from these final flip-flops to the chip pads are balanced.

The control signal will not have a synchronization transition on every cycle, so the previous sampling time is used when the control signal does not have a transition. This means there must be an accounting for the variations in the clock at the driving chip flip-flops with respect to the clock at the receiving chip flip-flops. The deviation of the clock transition time from its ideal time is called jitter; we will take the peak-to-peak deviation as the jitter. Jitter may be caused by small variations in the clock oscillator, by coupling of data signals to the clock signal, by variations in clock buffer Vdd caused by switching transients, and other effects. Worst case occurs if the transmit clock is early and the receive clock late when a control signal positive transition is sensed to determine the sample phase, and the transmit clock is late and the receive clock early on a following cycle that uses the sample phase determined previously. Exchanging early and late in the above case is another worst case. This has the effect of increasing the skew by the sum of the receiver and transmitter clock jitter. (decreasing the margin).

1. The internal cell format used in the Gigabit Switch includes an "idle" cell format that acts as a clocking synchronization cell. Thus all cell periods during which "real" cells are not being passed on a particular data link are used as cell timing synchronization periods.

2. With ideal delay elements, the phase to phase delay could be up to one third of the clock period, or 2.7ns.

Setup and Hold for Deskewed Inputs

The deskewing circuit greatly relaxes the setup and hold time requirements for data signals. There are three variations of the deskewing circuit used in the Gigabit Switch chip set. The deskewer circuit used between the IPP and OPP has no setup and hold restrictions. The data values, and time position of the first word of a cell, received by the IPP can have any timing relationship with the IPP clock and the IPP cell clock signals. The deskewer circuit used in the upper 16 bits of the IPP Link data interface allows a data skew of more than two link clock periods. The deskewer circuit used for data flow from the IPP to SE, SE to SE, and SE to OPP, allows an interval of approximately 2.5 clock periods, or about 21ns at 120Mhz, to the driving-chip to the receiving-chip clock skew and the data skew. Because of the nature of the deskewer input signals, their timing is not specified as standard setup and hold times, but rather by the limits over which the data signal is allowed to change from one value to the succeeding value. The two values used in the timing diagrams for the IPP-to-SE(s)-to-OPP timing of the individual chips are **Tmin** and **Tmax**¹ as shown in Figure 3 of the SE Signal Description Working Note ARL-96-05 for example. Note these two time intervals are related to different clocks.

The primary function of the CC_TAP input is to allow Gigabit Switch chips which do not have input pad to output pad delays of an integer number of clock periods. The CC_TAP allowed the switch designer the ability to adjust the cell level timing of each rank of switch chips. The value of CC_TAP determines the gross relative position of input and output data with respect to CELL_CLK. CC_TAP could be used to compensate for data distribution delays to individual chips. CC_TAP would be decreased by 1 (mod 16) for each 8.333ns additional delay in the data distribution to a chip. However, the timing margin for the Grant signals will be reduced to the point that such adjustments of the CC_TAP value will be unworkable.

Notes: (left over remarks to assist the authors, will be deleted when we are sure all these issues have been covered.)

1. All chip CLK signals should be derived from the same master clock.
2. Clock jitter would not be a problem if a control signal positive transition were associated with every data value.
3. Correlation in the jitter of the transmitter CLK and the receiver CLK does not help because the receiving flip-flop strobe may be generated from a different edge of the master clock than the transmitter flip-flop strobe.
4. Deliberately introduced PC board skew may be used to advantage if the on-chip data skew is asymmetrical with respect to the control signal.

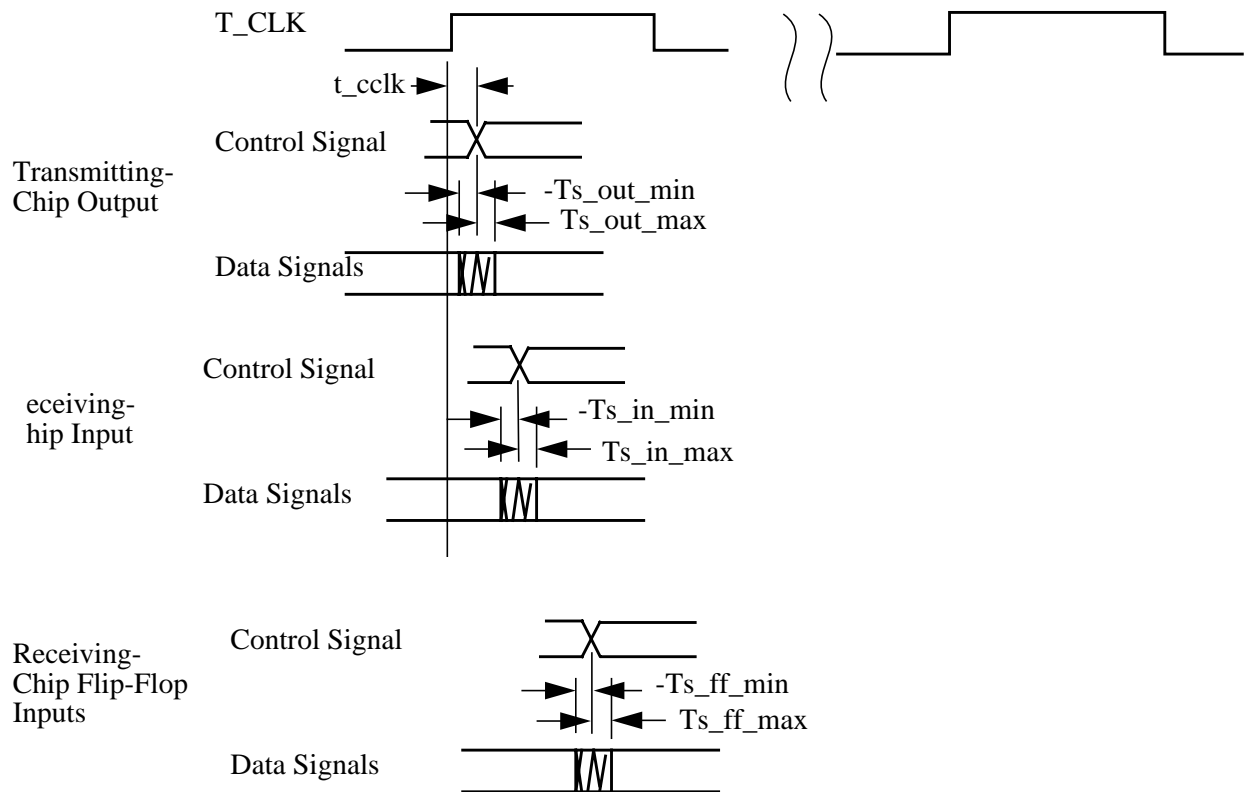
Parameters:

Figure 1 shows timing diagrams for transmission of data from a driver to a receiver with the sources of skew, the relations between signals, and the constraints necessary for the deskewing circuit operation. Table 1 gives the GigaBit Switch timing parameter values. Table 2 gives a representative skew budget for a deskewing connection.

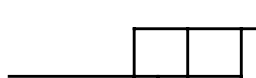
1. These two values are functions of chip internal delays as shown in the following equations. The internal delays are not defined in this document, but are here to assist the authors in the correct calculation of the total values.

$$\mathbf{Tmin} = (\text{Deskewer T1 to T2 max}) + (\text{FF max hold}) + (\text{max L0 to L6 clock tree delay}) = 3.6?? + .35???? + 4.5?? = 7.75??\text{ns}$$

$$\mathbf{Tmax} = (\text{Deskewer T1 to T2 min}) - (\text{FF min setup}) + (\text{min L0 to L6 clock tree delay}) = 1.35 - 1.5???? + 1.5???? = 1.35??\text{ns}$$



Receiving Chip Flip-Flop Clocks
(Three Sample Times Derived from R_CLK. Sample Used is Based on Arrival Time of Control Signal)



$$Ts_out_min = -0.45ns$$

$$Ts_in_min = Ts_out_min + Ts_PC_bd_min$$

$$Ts_ff_min = Ts_in_min + Ts_Receiving_Chip_min$$

$$Ts_out_max = 0.45ns$$

$$Ts_in_max = Ts_out_max + Ts_PC_bd_max$$

$$Ts_ff_max = Ts_in_max + Ts_Receiving_Chip_max$$

- Skew: Time of data signal transition with respect to control signal (may be positive or negative)
- Ts_out: Skew at output of transmitting package
- Ts_PC-bd: Skew due to pc board
- Ts_in: Skew at input to receiver package
- Ts_Receiving_Chip: Skew due to receiving package and chip
- Ts_ff: Skew at input to receiver flip-flops
- R_CLK_jitter: Receiving Clock jitter
- T_CLK_jitter: Transmitting Clock jitter

Based on minimum time between sample clock phases in the receiver of 1.3ns:
 $Ts_ff_max + R_CLK_jitter + T_CLK_jitter \leq 1.3ns$
 $Ts_ff_min - R_CLK_jitter - T_CLK_jitter \geq -1.3ns$

Figure 1. Timing Relation Between Control and Data Signals in a Transmission Group with Deskewing Circuit.

Table 1: Timing Parameter Values (Values in ns)

Parameter	Min	Max	Definition
Clock Frequency	-	120 Mhz	
Clock Duty Cycle	40%	60%	
Clock Rise/ Fall Time	0.3ns	2ns	
tpd	4.??+ 0.0137*C _L (pf)	12.4+ 0.0635*C _L (pf)	Clock to data out (ns)
tpd	4.??ns	12.4ns	Clock to Data out (With 61-Ohm transmission line and fan-out of 1)
ts	-	-3.??ns	Input setup time
th	-	11.??ns	Input hold time
ts_cclk	-	1.75ns	cell clock setup time
th_cclk	-	1.0ns	cell clock hold time
Tmin	9.6??ns		
Tmax		2.5??ns	
Ts_out	0.45??ns	0.45??ns	Output Skew (With 61-Ohm transmission line and fan-out of 1)
Ts_Receiving_ Chip			Receiver On-Chip Skew
T_phase	1.30ns		Time between input deskewer sampling intervals

**Table 2: Skew Budget for the GigaBit Switch Data
Group Transfers**

Skew Source	min	max
Ts_out	0.45??ns	0.45??ns
Ts_PC_Board	-0.2??ns	0.2??ns
Ts_Receiving_Chip	-0.2??ns	0.2??ns
R_CLK_jitter	-0.2??ns	0.2??ns
T_CLK_jitter	-0.2??ns	0.2??ns
Margin	-0.05ns	0.05??ns
Total	-1.30ns	1.3ns

Setup Time Requirement for Data Transfer from Chip to Chip

Figure 2 highlights the paths involved for data transfer from Chip B to Chip C. Data must change at the chip C input before it is needed by chip C. Figure 3 of the SE document shows the timing diagram for the case where chips B and C are IPP and SE, respectively. Times are referenced to the clock cycle for which CELL_CLK is asserted. T is included in the data path delay since Chip B data output flip-flops are clocked on the cycle following the cycle on which CELL_CLK is asserted, and 3T is added to the time at which data is required at Chip C since Chip C data input flip-flops are clocked three cycles after the cycle for which CELL_CLK is asserted. T is one clock period (normally 8.333ns).

Time Data Arrives at C < Time Data Required at C

$$T_{c1} + t_{pd} + T_{pcbd_dat1} + T < T_{c2} + T_{max} + 3T$$

$$T_{c1} - T_{c2} < 6.766ns - T_{pcbd_dat1}$$

Hold Time Requirement for Data Transfer from Chip to Chip

Figure 2 highlights the paths involved in data transfer from Chip B to Chip C. Data must not change at the chip C input until after it is no longer needed by Chip C. Figure 3 of the SE document shows the timing diagram for the case where chips B and C are IPP and SE, respectively. Times are referenced to the clock cycle for which CELL_CLK is asserted. T is included in the data path delay since Chip B data output flip-flops are clocked on the cycle following the cycle on which CELL_CLK is asserted.

Time Data Changes at C > Time Data Required to Remain at C

$$T_{c1} + t_{pd} + T_{pcbd_dat1} + T > T_{c2} + T_{min}$$

$$T_{c1} - T_{c2} > -2.73ns \text{ (taking } T_{pcbd_dat1} \text{ as zero)}$$

Setup Time Requirement for Grant from Chip to Chip (Hold time is easily satisfied since the Grant signals change only once per CELL_CLK period)

Figure 3 highlights the paths involved in a grant from Chip C to Chip B. The grant signal from chip C must reach chip B before it is required.

Time Grant at Chip B Input < Time Grant Required at Chip B Input

$$T_{c2} + 4T + t_{pd} + T_{pcbd_gnt} < T_{c1} + 8T - t_s$$

$$T_{c2} - T_{c1} < 23.9ns - T_{pcbd_gnt}$$

Timing Requirements for CC_TAP Not Equal 0.

The equations from above are modified here to show the effect of CC_TAP on transfers between chips.

$$\text{Setup Time: } T_{c1} - T_{c2} < 6.766ns - T_{pcbd_dat1} + T(CC_TAP_C - CC_TAP_B)$$

$$\text{Hold Time: } T_{c1} - T_{c2} > -2.73ns - T_{pcbd_dat1} + T(CC_TAP_C - CC_TAP_B)$$

$$\text{Grant Setup: } T_{c2} - T_{c1} < 23.9ns - T_{pcbd_gnt} + T(CC_TAP_B - CC_TAP_C)$$

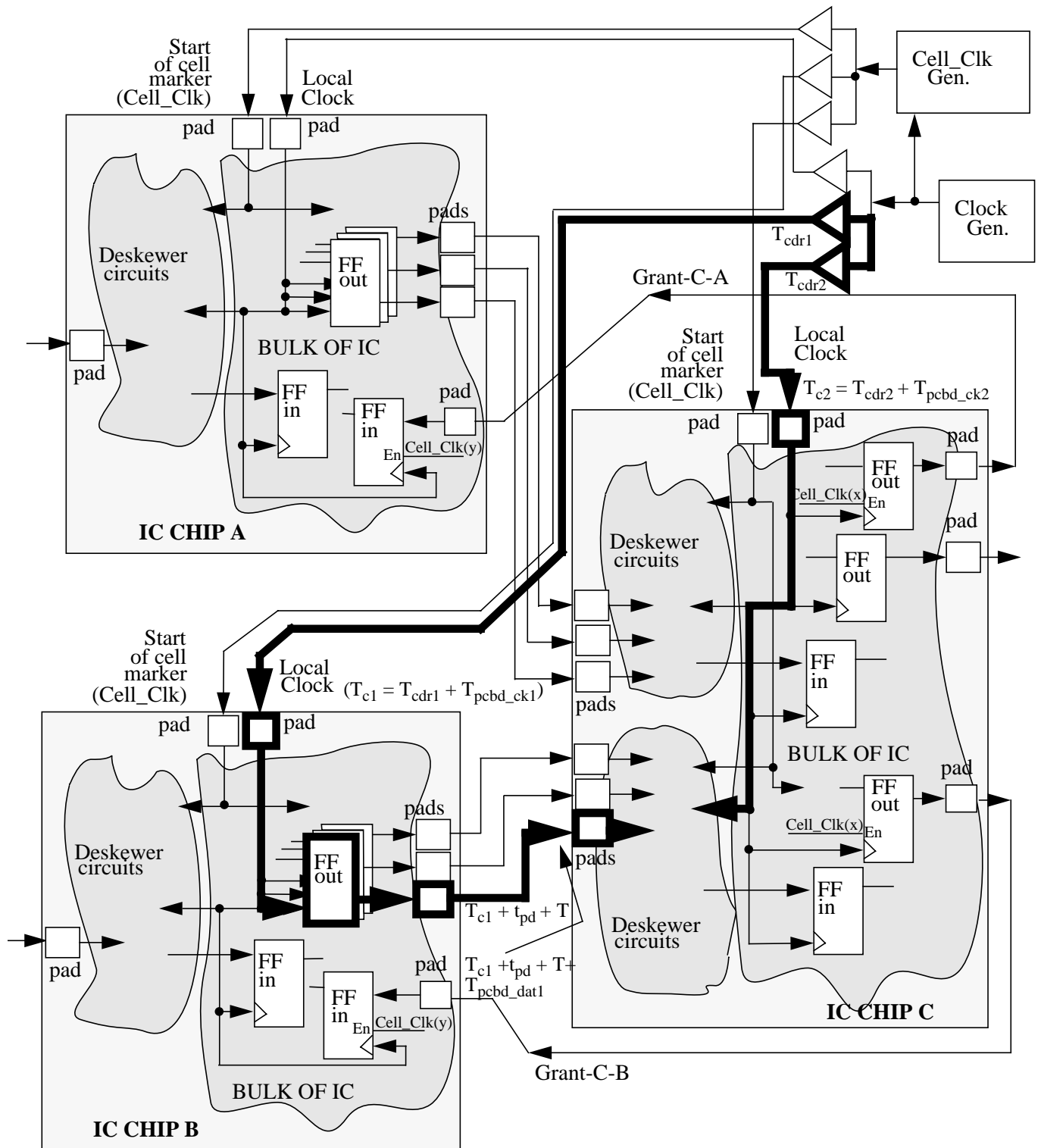


Figure 2

Setup and Hold Time Paths for Data from Chip B to Chip C

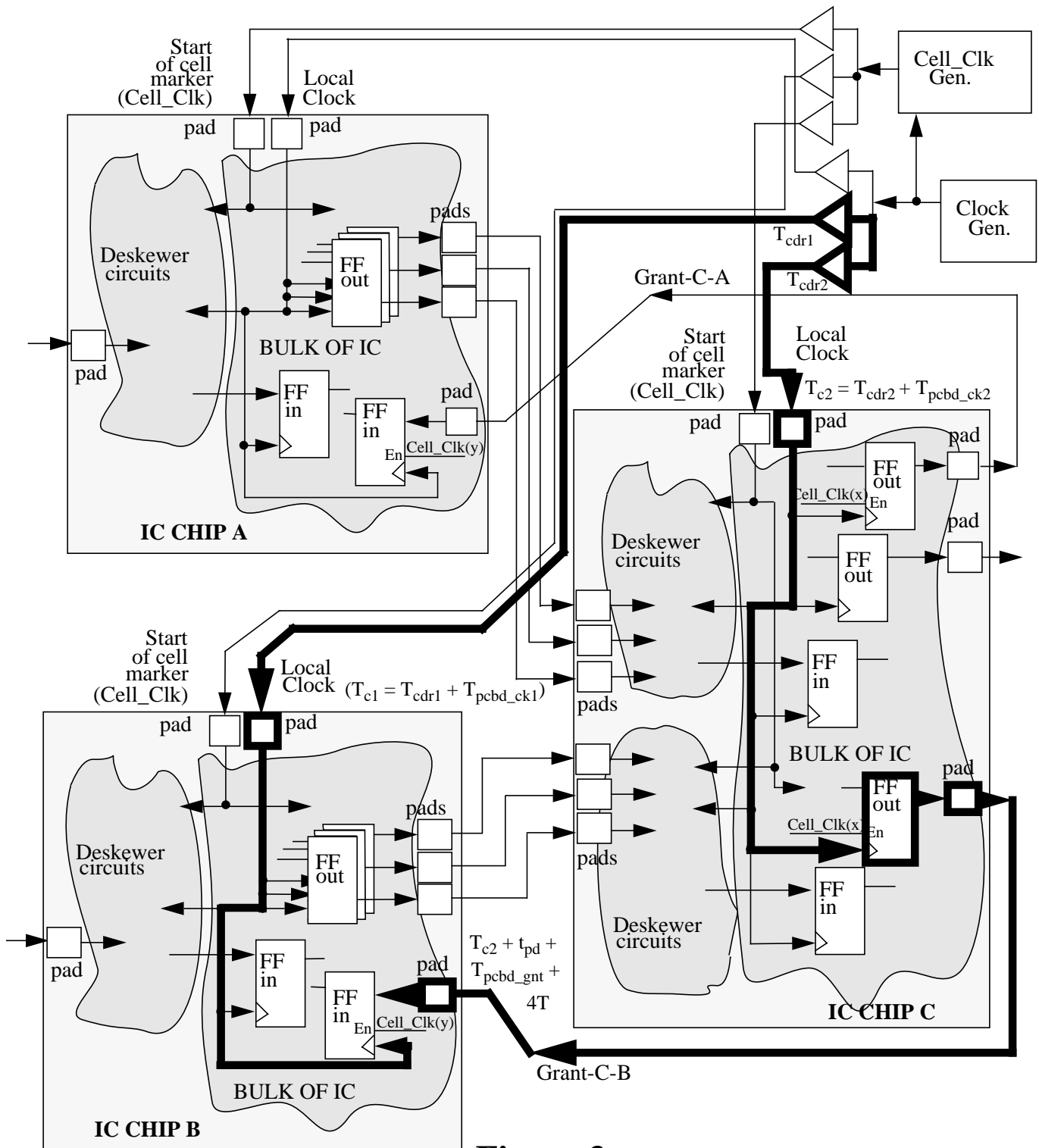


Figure 3

Setup Time Paths for Grant from Chip B to Chip C

An Example of Allocating Timing and Skew Between SE, IPP, and OPP Chips

Allowed Clock Skew for Minimum Data Path (Data Hold Time)

As an example, let us assume that the minimum data path can be effectively zero (an inch or less). From the data hold time equations we can have up to 2.73ns of clock skew between chip inputs. If we allow 0.5ns¹ skew for the clock line drivers, then the printed circuit board path delays for the clock signals can differ by up to 2.2ns, or approximately 13 inches of PC board path length (PC board strip-line has propagation delay of about 2ns per foot). If the clock driving circuits are located in the center of the switch, then IC chips can be located up to 26 inches apart.

Allowed Clock Skew for Maximum Data Path (Data Setup Time)

From the data setup time equations, clock skew minus PC board propagation delay must be less than 6.766ns. If we allow 2.7ns of clock skew as determined above for hold time, then the data path delay can be as much as 4ns or 24 inches. Very good control of clock skew can increase this by about 12 inches.

Max Grant Signal Delay

If we allow the clock path skew to again be 2.7ns, the maximum delay for the Grant signals is 23.9ns - 2.7ns or 21.2ns. Thus the grant signal PC board trace length can be over 10 feet, easily within the 26 inch diameter dictated by our clock signal distribution required for hold time.

Increasing Switch Size

To increase the size of a switch that can be built from the SE, IPP and OPP chips, we can adjust our allotments of the timing skew. With a little care, the clock-signal-distribution skew could be held to the 2.7ns we allotted above. IC chips close to the clock generation circuit would be fed from clock signals that are deliberately delayed.

The general structure of the switch architecture allows us to deliberately skew the clock (delay the clock at data receivers with respect to the transmitters) such that the allowed delay for grant signals is reduced in exchange for an increase in the allowed data path delay. The DIFFERENCE between max and min data path delays would have to held to 4ns, but the absolute data path delay could increase. If we balance the grant and data path delays, then we have the average of 21.2ns and 4ns, or about 12ns. Utilizing this would probably require padding the delay of short data paths. At 2ns/ft PC board delay, the Gigabit Switch could grow to a diameter of approximately 6ft using the current chip set. At 6ft, we are at least approaching a size switch where part of the signaling will be in cables that may have propagation delays closer to 8 inches per ns, thus 6ft is a conservative size.

With a modification to the SE chip so that the relationship between the grant and the returned data was increased one cell time, then the physical size of the Gigabit Switch (and thus the number of ports) could be increased to one half of [21.2ns + (16 clock periods)(8.3ns per period)], or about 77ns. At 6 inches per ns, the switch physical size could grow to about 40ft (over 11 meters). A switch of this size would not have interconnects that spanned the entire physical switch, thus the physical switch would be even larger! The

1. A 0.5ns clock-driver-skew budget is easy to meet for a single-level clock tree. For larger switches requiring multiple-level clock trees, a different allocation of clock skew between the drivers and the PC board paths may be more appropriate.

main point is that there is little reason to ever consider SE designs that change the grant to data cell timing relationship by more than an additional cell time. However, it would be necessary to also increase the difference between max and min data path delay to make this practical.

Long Term Clock Jitter

As detailed earlier, the total skew between control and data signals in a signal group includes transition clock-signal positive-transition jitter. The deskewing circuits can adjust for long term jitter, or timing “drift”. The deskewing circuit can pick a new sample phase each time an “idle” (or “synchronization”) cell is passed between two IC chips (every cell for the OPP to IPP path since there is a control signal transition for each cell). Thus between “idle” cells, the total signal jitter is composed of the cycle-to-cycle jitter, plus any long-term phase shift between the transmitter clock plus data path and the receiver clock. The phase shift might be caused by chip temperature changes and/or power supply voltage variations. The time between “idle” cells must be small enough to limit the long term phase shift to the budgeted amount. It is important to know the maximum length of a continuous run of “non-idle” cells. ATM switches can typically have average loads of no more than 90 to 95%¹ of the maximum bandwidth before the cell loss becomes a significant problem. Thus, on average, there will be an “idle” at least every 10 to 20 “real” cells. Another way to look at the problem is to realize that, for a particular path through the switch, there are less than a thousand cell storage locations along that path. Thus it is not reasonable to consider even bursts of “real” cells lasting more than a thousand cell times. For our example we will set the “period between idle cells” to 1000 cell times, or about 133us with a 120mhz clock, and 16 clock periods per cell time. For the example skew budget in Table 2, the signal timing can drift up to 0.1 ns every 133us. A 133us period is a 7.5kHz rate. If we restrict the long-term timing drift per period between “idle” cells to 2% of the total long-term timing drift, then the deskewers can follow drifts that change at a 150Hz rate, and have an amplitude of 5ns or less. Thus power supply power line “ripple” at 60 or 120Hz that could cause the total of the clock distribution, the transmitting IC chip, and the receiving IC chip timing drifts to change less than 5ns, peak to peak, can be handled by the deskewing circuits with small contribution to the skew budget. Timing drifts caused by IC chip heating will have much longer time constants, and can thus be also be easily handled, up to the deskewer range limits. Timing changes that result from on-chip voltage variations caused by data changes (simultaneous switching noise) will almost certainly be of high enough rate that the deskewer will not provide any timing adjustment help, and this jitter will need to be included in the skew budget. The deskewer timing drift adjustment feature is only useful for jitter effects that have a period of a few kilohertz or less. In practice, the biggest benefit of the deskewer tracking will be to adjust for temperature effects, with some small added benefit for power supply low frequency ripple, however, power supplies have very little 60 or 120Hz ripple.

1. For most forms of bursty traffic, the highest average switch loading, with low cell loss, will be even less.

Appendix A: Operating Conditions and Chip Timing Verification

In order to meet the stringent clock rate of 120Mhz, the operating conditions for the chips must be more tightly controlled than the normal commercial conditions. Atmel specifies MAX timing conditions at $V_{DD}=3.0V$ and $T=100C$. By restricting the operating conditions to $V_{DD}\geq 3.3V$ and $T\leq 80C$ the maximum delays are decreased and the maximum clock rate increased. From Figures 1.3 and 1.4 in the Atmel ES2 ECLP07 library data book we have:

$$\text{delay @ 3.0V} = 1.1(\text{delay @ 3.3V})$$

$$\text{delay @ 100C} = (1.26/1.195)(\text{delay @ 80C})$$

$$\text{delay @ 3.0V \& 100C} = 1.160(\text{delay @ 3.3V \& 80C})$$

$$\text{Clock Period @ 3.0V \& 100C} = 1.160(\text{Clock Period @ 3.3V \& 80C}) = 1.160 (1/120\text{Mhz}) = 9.7\text{ns}$$

Thus verification with Atmel worst case commercial parameters and clock period of 9.7ns verifies circuit operation for our more tightly controlled parameters and clock period of 8.333ns (120 Mhz). Timing verification and simulation are carried out with maximum allowed clock period of 9.7ns.

All times given in this document are for the Gigabit Switch operating conditions, $V_{DD}\geq 3.3V$ and $T\leq 80C$.