# IBM Enterprise System/9000 clock system: A technology and system perspective

by K. Chiu
J. J. DeFazio
T. G. McNamara

**The minimization of system clock skew is critical to the overall performance of high-speed computer systems. This paper discusses the statistical clock skew calculation methodology employed in the analysis of the IBM Enterprise System/9000™ (ES/9000™) computer systems. Comparisons made to a worst-case design approach show the advantages of a statistical clock skew calculation and its use in the timing analysis of ES/9000 systems. Design techniques that aid in the minimization of system clock skew are discussed throughout this paper. While many details concerning these design techniques and the statistical clock skew calculation methodology are tutorial in nature and have been used in the design of past IBM high-end machines, it is hoped that this paper will give the reader a useful understanding of the major considerations affecting the clock system and its application to an ES/9000 system.**

## Introduction

The IBM Enterprise System/9000™ (ES/9000™) systems have advanced the state of the art in both system design and technology. From a system perspective, an advanced central processor design, a unique system control element implementation utilizing a common second-level cache, and a flexible interconnect communication element combine to provide impressive processor performance and function. From the technology side, some of the most advanced circuit packages and silicon chips available have enabled the ES/9000 systems to become the first IBM mainframe to operate at a clock rate greater than 100 Mhz.

The clocking system, or "heartbeat," of the ES/9000 processors overlaps at the boundary between technology and system designs. As with most high-end systems, the ES/9000 processors are synchronous machines which require a clock signal to control the execution of sequential functions. The clock system distributes and generates clock signals to all storage elements and ensures that proper pulse widths and timing relationships meet their specifications. The accuracy of these timing relationships is extremely critical to the overall system performance, and a

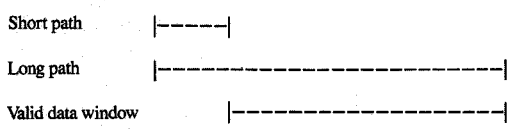


Figure 1
Valid data window.

clock performance parameter called *clock skew* is used as a measurement of this accuracy. A clock design must reduce the impact of clock skew on the overall system performance.

To minimize clock skew, clock design is treated as a customized product, with its own wiring rules, latch design, clock generation circuits, etc., which is integrated into the system design. The ES/9000 clock system is a multiphase clock system which supports a number of different clock arrival times within the machine cycle. Local customization (even at a chip level) of clock signal arrival times enables personalized, flexible designs in which system design trade-offs occur. This paper, which is tutorial in nature, describes the statistical calculation of clock performance parameters, examines the application of these parameters within the system timing environment, contrasts the statistical analysis to a worst-case approach [1], and discusses how customized design has reduced clock skew in the ES/9000 system.

A comparison between a statistical and worst-case analysis can best be illustrated by a simple example. For the statistical case we take the sum of two variables with normal distributions, each with a mean of $\mu_1$ and $\mu_2$, and a standard deviation of $\sigma_1$ and $\sigma_2$:

$$\mu_1, \sigma_1 + \mu_2, \sigma_2 = \mu_s, \sigma_s . \quad (1)$$

This produces a normal distribution with the resultant sum, $\mu_s$, equal to the sum of the two means,

$$\mu_s = \mu_1 + \mu_2 . \quad (2)$$

The resultant standard deviation of the sum $\sigma_s$ is a function of the individual variable standard deviations and a correlation coefficient, $\rho$, between the two variables,

$$\sigma_s = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2} . \quad (3)$$

In a worst-case treatment of the sum of two variables, the overall path tolerance $\sigma_s$ equals the linear addition of the two individual standard deviations shown in Equation (4):

$$\sigma_s = \sigma_1 + \sigma_2 . \quad (4)$$

If, as an example, $\sigma_1$ = data path tolerance and $\sigma_2$ = clock skew, and both these values equal 1.0, the standard deviation of the sum of these two variables for the worst case and the statistical case yields ($\rho = 0$ for the statistical case) 2.0 and 1.414, respectively. The statistical effect of a clock skew parameter equal to 1.0 on a data path tolerance also equal to 1.0 is only 0.414, or 41.4% of the effect. In the worst-case calculation, the full effect of the clock skew parameter is felt on the data path. Using timing relationships known as long and short paths, this example is expanded into a more detailed analysis between clock and data to determine critical timing edges and the allowable logic design space.

## Definition of terms

Terms that are used throughout this paper are defined in this section.

*Shift register latch*

The shift register latch (SRL) consists of an L1 latch portion (or "latch") that is controlled by a latch clock, C1. The latch captures data and provides a stable data input level to the next section of the register. An L2 latch portion (or "trigger") of the SRL is controlled by a trigger clock, C2. The trigger transfers data from the register into the logic network.

*Edge tolerance*

Because of variations in temperature, power supplies, and processing of chips and packages, the edge of a clock signal may arrive faster or slower than a predicted nominal time. An edge tolerance provides an acceptable window for an early or late arrival of a signal. The edge tolerance is usually defined as a Gaussian distribution in which the early edge is $-3\sigma$ and the late edge is $+3\sigma$.

*Skew*

Skew is a measure of the timing relationship between any two clock edges. It is a statistical parameter that depends on the edge tolerances as well as the correlation or "tracking" between the two clock edges. For instance, if one clock signal is faster than predicted by 50 ps and the other signal is faster by the same time, the skew between these two signals is zero.

*Short path and long path*

Timing paths are the times required to transmit data between two different storage elements. A storage element is a latch or an array (memory) chip. A long path is the maximum time available for data to travel between storage elements, and a short path is the minimum time required. If the data arrival time is longer than the long-path time, the data are captured in the following cycle; if it is shorter than the short path, the data are captured one cycle too

soon. A valid data window, illustrated in **Figure 1**, is required, in which an actual data transmission delay is longer than the short-path limit and shorter than the long-path limit. If either limit is violated, there is a data integrity problem that could produce incorrect decisions within the system logic. In general, clock skew increases the minimum time required for a short path and decreases the maximum time available for a long path. This, of course, reduces the valid data window and imposes much tighter timing constraints on the system designer. Reduction of clock skew is therefore critical to improved system performance.

## Distribution and generation of clock signals

In general, there are two methodologies used to design the clock signal distribution and generation network. In one method, clock pulse widths are generated in the network and these signals are distributed to the various chips. The IBM 3090™ systems were designed with this methodology. An alternative method is to distribute a copy of the oscillator signal to each logic and array chip and then generate the appropriate pulse width within each chip. The ES/9000 computers are the first high-end IBM systems to fully utilize this alternative method. A "distribute-then-generate" scheme was chosen because some chips require pulse widths which are too small to propagate along the package.

- *Clock signal distribution*

The clock distribution system is one of the most pervasive functions in the ES/9000 system. Starting from a single-crystal oscillator source, the clock system distributes a copy of this signal to all of the clock storage elements in the ES/9000 system. Clock signals are distributed through all package levels via cables and the various wires used to interconnect components, i.e., card, board, and module wires. Transmission line delays cause more than 60% of the total on-board clock distribution delay. Thus, minimizing the delay tolerance of these nets is critical to reducing the overall clock skew.

Since a clock generates the highest-frequency signal in a machine, a distribution of clock signals often tests the frequency limits of the technologies used in the computer. When an oscillator frequency is 100 Mhz, the data paths switch at a 10-ns rate, but the clock signal will switch twice as fast, every 5 ns, given a 50% duty cycle oscillator signal. To guarantee proper ac functionality of clock nets and limit the delay tolerance to an acceptable level, customized clock wiring rules are developed for the distribution of clock signals. All clock nets in the ES/9000 systems are point-to-point nets resistively terminated to match the characteristic impedance of the package, which is modeled as a transmission line. Clock nets are only wired on a limited number of planes in boards and thermal conduction modules (TCMs), and TCM and chip I/Os are reserved for clock signals. These design techniques minimize and control reflections that are generated by capacitive and impedance discontinuities. Since a copy of the actual oscillator signal, rather than signals with different pulse widths, is distributed at all packaging levels, line lengths can be chosen to position reflection away from the critical edge transition. The methodology for developing and analyzing wiring rules that control these reflections is described in [2].

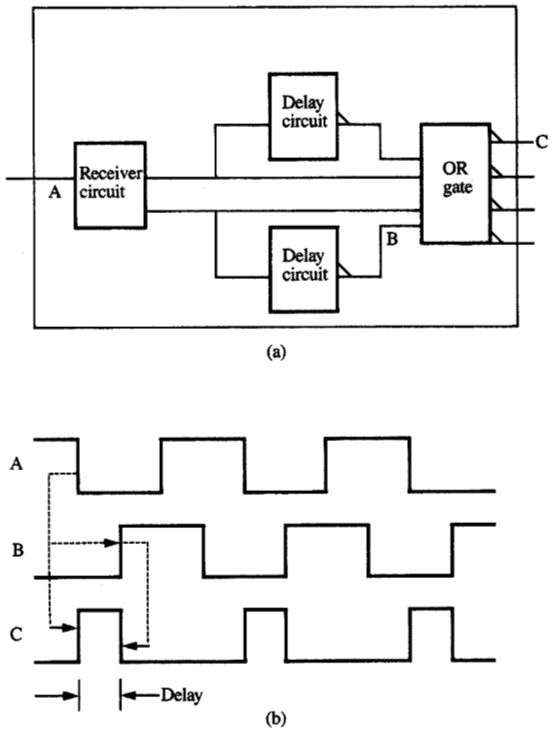


**Figure 2**

Clock-shaping circuit to minimize leading-edge tolerance: (a) Block diagram of circuit. (b) Timing diagrams at points A, B, C.

- *Clock signal generation*

Once an oscillator signal arrives at a logic or array chip, a special clock-shaping circuit generates the appropriate pulse width. The clock-shaping circuits are unique because the different array and logic chips in the ES/9000 have different pulse width requirements. Since certain edges of the clock signal are more critical, the shaping circuits minimize the delay tolerance of these critical edges. Special clock-shaping circuits [Figures 2(a) and 3(a)] were

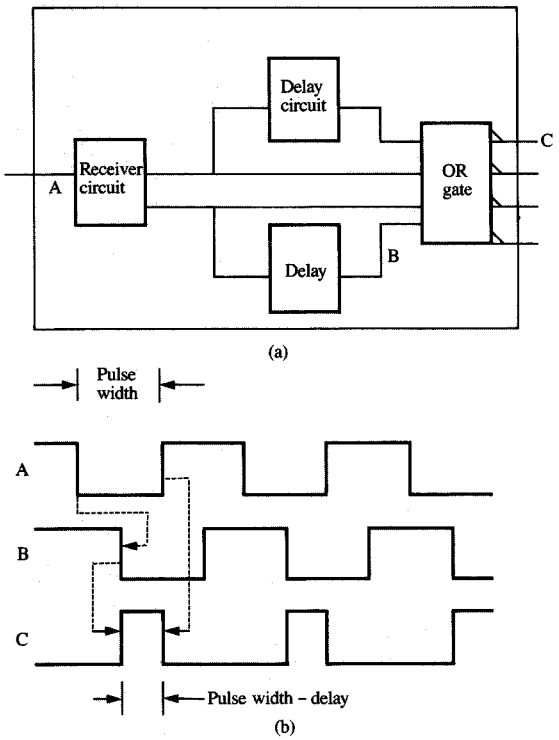


Figure 3

Clock-shaping circuit to minimize trailing-edge tolerance: (a) Block diagram of circuit. (b) Timing diagrams at points A, B, C.

designed for high-performance 11 000-circuit gate arrays to minimize the delay tolerance on these "critical edges."

**Figure 2(a)** is a schematic diagram of a single-edge clock-chopper circuit. The oscillator signal (50% duty cycle) is sent to a receiver circuit that drives an out-of-phase delay circuit and an OR gate. The operation of the single-edge clock-chopper circuit is illustrated in the timing diagrams shown in **Figure 2(b)**. A single-edge transition at point A transmitted in the negative direction produces a clock pulse at point C, whose pulse width equals the delay of the delay circuit. This circuit minimizes the tolerance on the leading edge of the generated clock pulse at point C. Only the delay tolerance of the receiver circuit and the OR gate affects the leading edge of a pulse. The trailing edge of the clock pulse at point C is a function of the tolerance of the delay circuit as well as the receiver circuit and the OR gate tolerance.

**Figure 3(a)** is a schematic diagram of a two-edge clock-chopper circuit. This circuit is used to minimize the tolerance on the trailing edge of a clock pulse. The circuits in Figures 2(a) and 3(a) are similar except that the delay circuit in Figure 3(a) is in phase. The timing diagrams in **Figure 3(b)** show that the clock pulse produced at point C depends on the input pulse width of the oscillator at point A minus the delay circuit delay. The trailing-edge tolerance of the clock pulse at point C is a function of the tolerance of the receiver circuit and the OR gate, while the leading-edge tolerance is a function of the tolerance of the delay circuit as well as the tolerance of the receiver circuit and the OR gate. Notice that the clock-shaping circuitry is duplicated in Figures 2(a) and 3(a). The necessity for this redundancy is discussed below.

## Calculation of clock skew

Clock skew, which is the delay between any two clock signals, can be computed using a worst-case linear or a statistical approach. However, Shelly and Tryon [3] have shown that statistical techniques predict system performance more accurately than worst-case analysis, and that no physical mechanisms justify a worst-case calculation. Properly specified statistical input parameters (correlation coefficients and delay tolerances) lead to timing analysis programs that yield better predictions of the performance of an actual system [4].

The clock skew calculations for the ES/9000 system use the statistical approach. A number of factors contribute to the calculation of clock skew, but the three most important are the component delay tolerances that comprise the clock distribution system, the delay correlation or "tracking" between these components, and the clock distribution topology.

- *Component tolerances*

The clock distribution system consists of bipolar chips, the electronic packages (cards, TCMs, and boards) on which the chips reside, and the interconnections (card, TCM, and board wires, and cables) between the chips. To limit variations in component delay when designing the ES/9000 clock system, the number of components used in the distribution of the clock signals was minimized, and components were selected with the smallest possible delay tolerance.

As a result of increased circuit densities and I/Os on chips, the clock system overhead of the TCM board distribution for the ES/9000 system was reduced by 33% in comparison to the 3090 system (two chips were required for the ES/9000 distribution compared to three chips on the 3090). Furthermore, only two copies of the clock-powering chip were required on each TCM in the ES/9000 system, while four copies of the clock-powering chip were needed on the TCMs in the 3090 machine (a 50% reduction in the TCM clock overhead).

- *Correlation coefficients*

The importance of correlation coefficients $\rho$, which describe how one variable tracks another variable, cannot

be understated when designing and analyzing a statistical clock system. Individual components are combined using statistical correlation coefficients that exist between any two components to accurately predict the overall clock skew. Two normally distributed variables with a mean of $\mu_1$ and $\mu_2$ and a standard deviation of $\sigma_1$ and $\sigma_2$ can be added or subtracted to produce a third normally distributed variable.

The sum of the two variables is

$$\mu_3 = \mu_1 + \mu_2, \tag{5}$$

$$\sigma_3 = \sqrt{\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2}. \tag{6}$$

The difference between the two variables is

$$\mu_3 = \mu_1 - \mu_2, \tag{7}$$

$$\sigma_3 = \sqrt{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}. \tag{8}$$

If two similar circuits track perfectly, then $\rho = 1$. For the case where both circuits have the same distribution and a correlation of 1, their skew, or difference in their delay, will be zero. The clock system must utilize components that have a high degree of tracking as well as design techniques which improve the tracking between components. One technique that is used to improve the tracking between clock interconnections on a TCM is to wire all clock nets on the same plane. The variation in the dielectric constant within a plane is smaller than the variation between planes. Since the dielectric constant has a direct effect on the delay of the signal as it propagates in the TCM, a smaller variation translates into less skew between different clock interconnections.

### • *Clock distribution topology*

The calculation of clock skew is critically dependent on the clock distribution topology. The ES/9000 clock system derives all critical clock signals from one common crystal oscillator. This oscillator signal is distributed to each system board, and each board, in turn, generates the appropriate latch, trigger, and array clocks and distributes them to chips that reside on the TCMs. To determine the clock skew between two clocks, it is necessary to know the clock locations. The skew between a latch clock on one board and a trigger clock on a different board is very different from the skew between a latch clock and a trigger clock on the same board. Two clocks on the same board have a greater similarity in signal distribution than clocks on different boards. From a topological point of view, clock skew is reduced for signals that traverse a common distribution. Choosing the best clock distribution topology for a system depends on a number of parameters, including technology, testability, diagnostics, and location of critical paths.

### • *Skew calculation*

The clock distribution system topology shown in **Figure 4** illustrates the statistical calculation of the clock skew. Clock skew is typically calculated at the clock input of the latch and/or array. However, for illustrative purposes, skew is calculated at the end of the two nets shown in Figure 4.

Figure 4 shows a single oscillator edge arriving at the input to the clock control chip (C1). This edge is common to both the trigger and latch clock and is the starting point for calculating skew. The oscillator edge goes through two different paths on C1: one path for the latch clock (C1L) and one path for the trigger clock (C1T). The latch clock is powered up via net 1 (N1) to its clock-powering chip (C2), while the trigger clock is sent to its clock-powering chip (C3) via net 2 (N2). The two clock-powering chips, C2 and C3, drive a corresponding clock signal through nets N3 and N4, respectively. The skew between the latch clock at the end of N3 and the trigger clock at the end of N4 is calculated by combining chip and net tolerances with their corresponding correlation coefficients in an expansion of the sum and difference equations presented earlier. Edge tolerances are summed along a path and subtracted between paths:

$$\begin{aligned}(\textit{Clock skew})^2 = {} & \sigma_{C1L}^2 + \sigma_{C1T}^2 + \sigma_{C2}^2 + \sigma_{C3}^2 + \sigma_{N1}^2 + \sigma_{N2}^2 + \sigma_{N3}^2 + \sigma_{N4}^2 \\ & + 2\rho_{cc}\sigma_{C1L}\sigma_{C2} + 2\rho_{cc}\sigma_{C1T}\sigma_{C3} + 2\rho_{cn}\sigma_{C1L}\sigma_{N1} + 2\rho_{cn}\sigma_{C1L}\sigma_{N3} \\ & + 2\rho_{cn}\sigma_{C1T}\sigma_{N2} + 2\rho_{cn}\sigma_{C1T}\sigma_{N4} + 2\rho_{cn}\sigma_{C2}\sigma_{N3} + 2\rho_{cn}\sigma_{C3}\sigma_{N4} \\ & + 2\rho_{cn}\sigma_{C2}\sigma_{N1} + 2\rho_{cn}\sigma_{C3}\sigma_{N2} + 2\rho_{nn}\sigma_{N1}\sigma_{N3} + 2\rho_{nn}\sigma_{N2}\sigma_{N4} \\ & - 2\rho_{cc}\sigma_{C1L}\sigma_{C1T} - 2\rho_{cc}\sigma_{C1L}\sigma_{C3} - 2\rho_{cc}\sigma_{C1T}\sigma_{C2} - 2\rho_{cc}\sigma_{C2}\sigma_{C3} \\ & - 2\rho_{nn}\sigma_{N1}\sigma_{N2} - 2\rho_{nn}\sigma_{N1}\sigma_{N4} - 2\rho_{cc}\sigma_{N3}\sigma_{N2} - 2\rho_{nn}\sigma_{N3}\sigma_{N4} \\ & - 2\rho_{cn}\sigma_{C1L}\sigma_{N2} - 2\rho_{cn}\sigma_{C1L}\sigma_{N4} - 2\rho_{cn}\sigma_{C2}\sigma_{N2} - 2\rho_{cn}\sigma_{C2}\sigma_{N4} \\ & - 2\rho_{cn}\sigma_{C1T}\sigma_{N1} - 2\rho_{cn}\sigma_{C1T}\sigma_{N3} - 2\rho_{cn}\sigma_{C3}\sigma_{N1} - 2\rho_{cn}\sigma_{C3}\sigma_{N3}, \end{aligned} \tag{9}$$

where $\rho_{cc}$ = chip-to-chip correlation, $\rho_{cn}$ = chip-to-net correlation, and $\rho_{nn}$ = net-to-net correlation.

In this equation, each path on a chip (C1L, C1T, C2, and C3) and the nets connecting these chips (N1, N2, N3, and N4) are treated as statistical variables. The correlation coefficients ($\rho_{cc}$, $\rho_{cn}$, and $\rho_{nn}$) can each be different. For example, two identical paths on the same chip will have a higher path correlation than two identical paths on different chips, since similar circuits on the same chips generally track very well. The clock skew equation illustrates that the sum equation tends to increase the clock edge tolerance, while the difference equation tends to reduce the clock skew. Since the ultimate goal for the clock designer is to minimize the clock skew, it is critical to design a clock distribution system that maximizes the correlations that are used in the difference equation. This is usually achieved by a judicious selection of circuits.

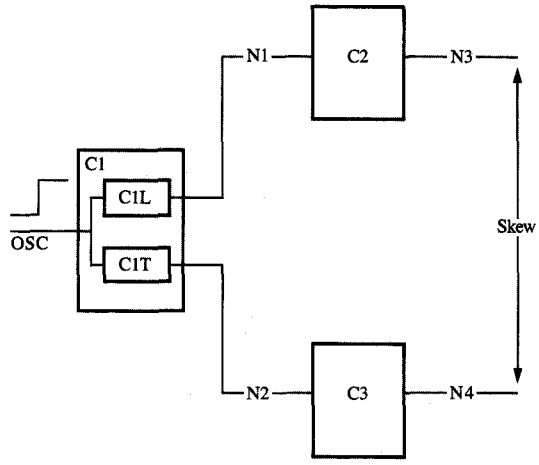


Figure 4

Clock distribution topology.

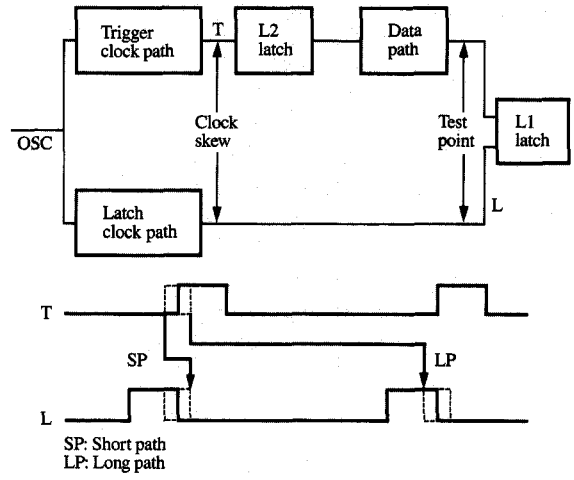


Figure 5

Data path timing: (a) Block diagram of circuit. (b) Timing diagrams at trigger clock (T) and latch clock (L).

The accuracy of input parameters, such as correlation coefficients and tolerances, is critical to the statistical calculation just described. Since the determination of these parameters is generally difficult, analyses are first performed to determine which parameters are most critical to the clock skew analysis. Efforts should then be made to predict these parameters accurately.

The clock skew equation above is simplified to illustrate the calculation of ES/9000 clock skew. A timing analysis problem [4, 5] uses circuit-to-circuit correlation coefficients in the expansion of the sum and difference equations. Since each path on a chip consists of a number of circuits, the path-to-path or chip-to-chip correlation coefficients ($\rho_{cc}$) are really a combination of circuit-to-circuit correlation coefficients. A clock distribution path consisting of three chips with three levels of circuits on each chip has a much larger number of terms in the full expansion of the sum and difference equations.

## Timing analysis with clock skew

Clock skew is not sufficient to determine the impact of clock performance on the overall ES/9000 system performance. In this section, the relationship between clock skew and the data path is discussed, as well as the statistical effect clock skew has on system performance.

Since ES/9000 systems are synchronous, all data transfers are initiated and terminated by the system clocks. As shown in **Figure 5**, a common oscillator signal propagates through the trigger and latch clock paths and produces the respective trigger (T) and latch (L) clock timing relationships. The trigger and latch clock paths generally consist of a number of chips and the wiring interconnections between them. The leading edge of the trigger (T) clock transmits new data from the L2 latch into the data path. The data are then transmitted to the L1 latch. As indicated in Figure 5, the earliest (short-path test) and latest (long-path test) arrival times of the data are compared to the latch (L) clock at a "test point" which tests for short and long paths.

For a short-path (SP) test, the data transmitted by the earliest trigger leading edge in cycle $N$ are compared to the arrival of the latest latch trailing edge in cycle $N$. The data must arrive after the latch clock trailing edge in order to "pass" the SP test and be captured in cycle $N + 1$ and not in cycle $N$. For a long-path (LP) test, the data transmitted by the latest trigger leading edge in cycle $N$ are compared to the arrival time of the earliest latch trailing edge in cycle $N + 1$. The data must arrive before the latch clock trailing edge turns "off" in order to be captured in cycle $N + 1$. These tests determine the minimum and maximum data path delay required for proper machine operation, i.e., the valid data window.

If the trigger clock (T), the latch clock (L), and the data path (D) are treated as statistical variables (assuming that the L1 and L2 latch tolerances and their appropriate set-up and hold times are included in the data path tolerance), the standard deviation at the test point ($\sigma_{TP}$) is calculated by expanding the sum and difference equations:

$$\sigma_{TP} = \sqrt{\sigma_T^2 + \sigma_L^2 + \sigma_D^2 + 2\rho_{TD}\sigma_T\sigma_D - 2\rho_{TL}\sigma_T\sigma_L - 2\rho_{LD}\sigma_L\sigma_D}, \quad (10)$$

where $\rho_{TD}$ = trigger-to-data correlation, $\rho_{TL}$ = trigger-to-latch correlation, and $\rho_{LD}$ = latch-to-data correlation. Note that the term $\sigma_T^2 + \sigma_L^2 - 2\rho_{TL}\sigma_T\sigma_L$ is equal to $(clock\ skew)^2$ and can be rewritten as

$$\sigma_{TP} = \sqrt{(clock\ skew)^2 + \sigma_D^2 + 2\rho_{TD}\sigma_T\sigma_D - 2\rho_{LD}\sigma_L\sigma_D}. \quad (11)$$

If the trigger and latch clock path distributions have the same type and number of circuits in their paths, the last two terms in the above equation are nearly equal, and a first-order representation of $\sigma_{TP}$ is the RSS (root sum squared) of the clock skew and the data path tolerance:

$$\sigma_{TP} = \sqrt{(clock\ skew)^2 + \sigma_D^2}. \quad (12)$$

The statistical combination of clock skew and data path tolerance reduces the impact of clock performance on system performance more than a linear application of clock skew. The analysis described was simplified to illustrate the effect of clock skew on system performance. Timing analysis programs calculate the full sum and difference equations with the appropriate circuit-to-circuit correlation coefficients.

For a worst-case calculation of $\sigma_{TP}$, clock skew and the combination of clock skew and the data path would each be calculated linearly, thereby increasing the system cycle time:

$$clock\ skew = \sigma_T + \sigma_L, \quad (13)$$

$$\sigma_{TP} = \sigma_D + \sigma_T + \sigma_L. \quad (14)$$

By using typical values for the clock edge and the data path tolerances, a worst-case linear application is compared with a statistical application for the short- and long-path calculations. All correlation coefficients are set to zero ($\rho = 0$) in these calculations. For

C2 clock tolerance = 500 ps,
C1 clock tolerance = 500 ps,
Long-path data tolerance = 1000 ps,
Short-path data tolerance = 250 ps,

long-path calculations are

$$\sigma_{TP}(statistical) = \sqrt{500^2 + 500^2 + 1000^2} = 1225 \text{ ps};$$

$$\sigma_{TP}(linear) = 500 + 500 + 1000 = 2000 \text{ ps}.$$

Short-path calculations are

$$\sigma_{TP}(statistical) = \sqrt{500^2 + 500^2 + 250^2} = 750 \text{ ps};$$

$$\sigma_{TP}(linear) = 500 + 500 + 250 = 1250 \text{ ps}.$$

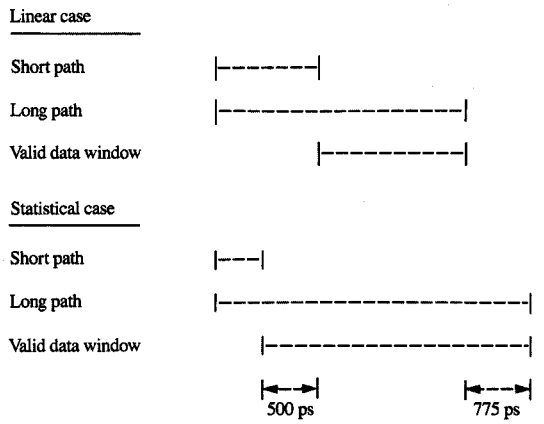


**Figure 6**

Comparison of valid data window times computed from linear (or worst-case) and statistical calculations.

The statistical calculation increases the valid data window by 1275 ps (775 ps + 500 ps for the long and short paths). As noted earlier, Shelly and Tryon [3] have shown that statistical techniques provide a more accurate prediction of system performance than a worst-case (linear) analysis. On the basis of the statistical calculations, system performance is improved by increasing the logic per cycle or by reducing the cycle time. The minimum path delay required to meet the short-path test is reduced in the statistical case. This reduction increases the valid data window, as illustrated in **Figure 6.**

## CSEF and DCS on-chip clock distribution

The multipurpose high-performance gate array [6] used in the ES/9000 systems supports two circuit families: an emitter-coupled logic family called current switch emitter follower (CSEF), and a logic family called differential current switch (DCS). Latches in both circuit families permit logic designers full flexibility to trade off speed, power, and functionality in their design. Since CSEF and DCS require different signal swings to ensure proper operation, special clock interface circuits are needed to support latches in both families. These circuits are used in an on-chip clock distribution tree.

A clock-shaping circuit in the on-chip clock distribution tree, which is similar to circuits in Figures 2 and 3, receives an oscillator signal and generates a "shaped" pulse width. The clock shaper then distributes this "shaped" signal to a special clock-driver circuit that transmits the clock signal to a latch. A unique driver

circuit interfaces with latches in both logic families. Similar delays in the clock-driver circuits guarantee that the clock arrival times at a latch input are similar regardless of the latch type used. One clock distribution tree supports both types of latches with the same timing relationship requirement.

"Loading block" circuits are designed and used in clock distribution networks by automated chip placement and wiring programs to equalize loading and wire capacitance. Minimum and maximum capacitance limits placed on the on-chip clock nets reduce the overall clock skew.

To guarantee product quality levels, all bipolar chips in the system must meet a specified test coverage for "stuck-at" faults. The clock-shaping circuits have some untested stuck-at faults due to their inherent reconvergent fanout. As shown in Figures 2 and 3, the clock-shaping circuits were designed with redundancy to eliminate these faults.

## Clock system diagnostics: Path stressing

Path stressing is a technique that measures system performance and the operating margin of the long and short paths during system testing, and verifies the valid data window.

Path stressing can induce long or short paths to fail by changing the timing relationships of a latch, trigger, and/or array clock [7]. The nominal arrival time of a clock edge is stressed by adding internal circuit delay. The ES/9000 system is designed with small, medium, and large stress levels which add delays of approximately 0.25, 0.5, and 0.75 nanoseconds, respectively. The number of logic levels needed for each stressed level depends on the technology offered for the internal circuits. The actual added delay of the same stressed level varies depending on the wire capacitance of the path and on the critical edge (rising or falling) used.

The ES/9000 path-stressing design allows this capability on an individual clock tree at the lowest point in the clock distribution. This results principally from the 5× increase in logic cells compared to the 3090 system, in which path stressing was done on a board basis. This increased capability allows the system designer to deal more precisely with timing problems, sometimes isolating timing problems to an individual chip. This has reduced system debugging times on the test floor as well as determining the long-path and short-path timing margins.

## Summary

To control the execution of sequential events in ES/9000 processors, a clock system generates and distributes signals to all storage elements and ensures that specified pulse widths and timing relationships are met. The timing relationships are characterized by a a clock-performance parameter called clock skew. A reduction in system clock skew plays a critical role in increasing the window for transmission of valid data and for improving the overall performance of ES/9000 processors. The calculation and application of clock skew within the system timing environment have been described.

To reduce clock skew, clock design has been customized for the IBM ES/9000 processors. The processors contain a multiphase clock system which supports a number of different clock arrival times within a machine cycle. A copy of the oscillator signal is distributed to each logic and array chip, and an appropriate pulse width is generated within the chips. On-chip clock-pulse-shaping circuitry is incorporated to minimize clock skew due to manufacturing tolerances. The number of components used to distribute clock signals, the placement of chips, and the data path topology have each been adjusted to enhance clock performance. ES/9000 clock diagnostics have been increased to facilitate the isolation of timing problems.

The ES/9000 systems are the first IBM high-end computer family to utilize clock-pulse-shaping circuitry fully on all logic and array chips to reduce clock skew. The new and customized clock design techniques distribute and generate clock signals at more than 100 MHz. The IBM ES/9000 family contains the first IBM processors to operate at this high clock frequency.


## Acknowledgments

The authors wish to thank their many colleagues at IBM for their help and contributions. We would specifically like to thank D. Chang, J. Collins, and J. Shelly for their ground-breaking work on clock system design and analysis.




## References


1. S. Unger and C. J. Tan, "Clocking Schemes for High-Speed Digital Systems," *IEEE Trans. Computers* **C-35**, No. 10, 880–895 (October 1986).
2. E. E. Davidson, "Electrical Design of a High Speed Computer Package," *IBM J. Res. Develop.* **26,** No. 3, 349–361 (May 1982).
3. J. Shelly and D. Tryon, "Statistical Techniques of Timing Verification," *Proceedings of the 20th Design Automation Conference*, 1983, pp. 396–402.
4. Robert B. Hitchcock, Sr., Gordon L. Smith, and David D. Cheng, "Timing Analysis of Computer Hardware," *IBM J. Res. Develop.* **26,** No. 1, 100–105 (January 1982).
5. R. B. Hitchcock, "Timing Verification and the Timing Analysis Program," *Proceedings of the 19th Design Automation Conference*, Las Vegas, 1982, pp. 594–603.
6. A. E. Barish, J. P. Eckhardt, M. D. Mayo, W. A. Svarczkopf, and S. P. Gaur, "Improved Performance of IBM Enterprise System/9000 Bipolar Logic Chips," *IBM J. Res. Develop.* **36,** No. 5, 829–834 (1992, this issue).
7. J. C. Collins and J. J. DeFazio, "System Path Stressing," U.S. Patent 4,564,943, 1986.

*Received September 3, 1991; accepted for publication March 3, 1992*

**Kin Chiu** *IBM Technology Products, East Fishkill facility, Route 52, Hopewell Junction, New York 12533 (resigned).* Mr. Chiu was an Associate Engineer in the Electrical Packaging and Integration Department at the East Fishkill facility. He received a B.E. in electrical engineering from the City College of New York in 1988 and joined IBM at the East Fishkill facility that same year.

**John J. DeFazio** *IBM Enterprise Systems, P.O. Box 950, Poughkeepsie, New York 12602 (DEFAZIO at PK705VMG).* Mr. DeFazio is a Senior Engineering Manager in the Design Systems Organization in Poughkeepsie. He received his B.S. in electrical engineering from the New Jersey Institute of Technology at Newark in 1980, and his M.S. in industrial administration from Union College, Schenectady, New York, in 1986. Mr. DeFazio joined IBM to work on high-performance clock systems in 1980, and in 1985 transferred to Poughkeepsie to work on large-system development. He currently manages a software tool design team in support of high-end systems. Mr. DeFazio has received an IBM First-Level Invention Achievement Award and an IBM Division Award; he is a member of Tau Beta Pi and Eta Kappa Nu.

**Timothy G. McNamara** *IBM Technology Products, East Fishkill facility, Route 52, Hopewell Junction, New York 12533 (MCNAMARA at FSHVMX).* Mr. McNamara is an Advisory Engineer in the Electrical Packaging Integration Department at the East Fishkill facility. He received his B.E. in electrical engineering from the State University of New York at Stony Brook in 1983 and his M.S. in computer engineering from Syracuse University in 1990. Since joining IBM in 1983, he has been working on high-performance clock system development. Mr. McNamara is a member of Tau Beta Pi and Eta Kappa Nu.