

# **DesignCon East 2005**

## Performance Model for Inter-Chip Busses Considering Bandwidth and Cost

Brock J. LaMeres, Agilent Technologies Inc.

Sunil P. Khatri, Texas A&M University

## **Abstract**

We present an analytical method to perform the design of the I/O subsystem of an IC given its throughput requirements. Our method can be used to select the IC package, along with the bus size and speed so as to minimize I/O cost. We have validated our model by conducting simulations on three industry-standard packages while varying the bus width, slew rate, and signal-to-power/ground ratio. Our experimental results track closely with the analytical model. We demonstrate for the packages considered that it is more cost effective to use faster, narrower busses rather than slower wider busses to achieve a desired system throughput.

## **Author(s) Biography**

**Brock J. LaMeres** received his BSEE from Montana State University in 1998 and his MSEE from the University of Colorado in 2001. He is currently a Ph.D. candidate at the University of Colorado where his research focus is VLSI Circuit Design and High-Speed I/O for next generation IC's. For the past 6 years he has worked as a hardware design engineer for Agilent Technologies in Colorado Springs where he designs logic analyzer probes and acquisition boards. LaMeres has published 25 technical articles in the area of signal integrity and has a patent in the field of logic analyzer probing. LaMeres is a registered Professional Engineer in the State of Colorado.

**Sunil P. Khatri** is an Assistant Professor in the Department of Electrical Engineering at Texas A&M University. He is affiliated with the VLSI CAD group. He completed his Ph.D. from the University of California, Berkeley in 1999. Before this, he worked with Motorola, Inc on the designs of the MC88110 and PowerPC 603 RISC Microprocessors. Khatri obtained his M.S from the University of Texas at Austin, which followed his B.Tech. from the Indian Institute of Technology, Kanpur. His research is in the areas of VLSI Design and VLSI CAD. Some recent areas of interest are design automation for datapath circuits, cross-talk avoidance in on-chip buses, leakage-power reduction, extreme low power circuit design, asynchronous circuit design methodologies, timing estimation, efficient test generation, fast logic simulation and cross-talk immune VLSI design.

## I. Introduction

Advances in CMOS technology have led to a dramatic increase in the on-chip performance of ICs. While the computational power of on-chip circuitry continues to grow, the inter-chip interconnect significantly limits the performance of digital systems [1,2]. The core speed for today's ICs is many times faster than the speed of inter-chip busses. As a consequence, inter-chip bus design is becoming a very important challenge in digital system design. Simply widening I/O buses to increase the total bus throughput is not practical due to the high cost of each I/O pin. In addition, the electrical parasitics of standard packaging limits not only the per channel bandwidth, but also the total number of signals that can switch simultaneously. Due to all of these factors, inter-chip bus design requires a careful analysis of the cost versus performance tradeoff.

Traditionally, inter-chip communication is performed using wide parallel busses. The standard approach to achieving the desired system bandwidth is to increase the number of pins on the package until the desired throughput is attained. There are three main problems with this approach.

- Cost of packaging. Package cost scales faster than linearly with the number of I/O pins that are needed and accounts for a large contribution to the overall chip price [3].
- Performance. Wide parallel busses experience a host of signal integrity issues associated with simultaneous switching of digital signals [1,4,5]. Problems such as ground bounce and power supply droop occur when the large dynamic currents of the CMOS output drivers induce a voltage across the inductance of the package [6,7,8]. Solutions to this problem include increasing the number of power and ground pins to reduce the inductance in the power supply current path. However, this increases the cost of the package because the number of I/O pins increases. Another solution to the package parasitic problem is to move toward advanced packaging technologies such as flip-chip packaging, to reduce the inductance in the power and ground leads. This reduces the voltage induced across the pins when large AC currents are present. However, advanced package technologies also increase the price of the IC.
- The increases in package bandwidth do not scale at the same rate as on-chip core frequencies [1]. The traditional approach of widening parallel busses to match the inner core's data rate is impractical not only from a cost viewpoint, but also because the signal integrity problems mentioned above limit how wide busses can be. The paradox of the wide parallel bus is that adding I/O should produce a linear increase in system throughput but in reality suffers an asymptotic limit due to signal integrity issues, as our experiments demonstrate. Parallel busses have to be ran at lower speeds as their width increases, which inherently limits their utility.

Recently, we have seen the emergence of narrower busses that run at higher per-pin data rates [1,2]. These new busses include Rapid I/O [9], PCI Express [10], and Hyper Transport [2]. All of these busses take advantage of the fact that the per-pin bandwidth of modern packages is much higher than that of parallel busses. Instead of widening the busses to achieve the system throughput, these new bus standards operate at much faster data rates, using fewer I/O pins, and therefore achieve the same or greater system throughput [11]. These busses are narrow enough to avoid the mutual inductance problems of modern packages. Further, since these busses are narrow they achieve a cost reduction in I/O. These factors enable the data rates to be equal to or near the theoretical maximum bandwidth of the I/O structure under ideal conditions.

Regardless of whether the inter-chip communication uses a slower, parallel bus or a faster and narrow bus, the objective is the same – the inter-chip bus must deliver the highest throughput in the most cost-effective manner. This is a challenging problem due to the faster than linear increase in the cost of adding I/O pins that must be balanced with the asymptotic limit to how much bandwidth can be attained by widening the bus. The most cost-effective solution to this problem occurs at the inflection point of where adding I/O pins increases the throughput of the bus at such a small rate that the cost increase negates adding I/O pins. To assist in the design of cost-effective inter-chip busses, this paper will present an analytical model for selecting the width and speed of the bus. Our approach considers the maximum data rate that a package can accommodate as the number of channels is increased. In addition, the cost of adding I/O pins is considered for three different Signal/Power/Ground (SPG) ratios – 8:1:1, 4:1:1, and 2:1:1. SPICE simulations are performed on three industry standard packages to validate the analytical model. The three packages considered are a Quad Flat Pack (QFP) with wire-bonding, a Ball Grid Array (BGA) with wire-bonding, and a BGA using flip-chip. This paper presents a selection methodology for inter-chip bus designers that will aid in selecting the package, bus width, and signal speed which results in minimal cost given a desired system throughput. It is shown that the most cost-effective bus design is obtained at the inflection point of the curve of bandwidth per unit cost versus the number of I/O pins. At this point, the cost of adding IC pins negates the small addition in throughput. This point depends on the package parasitics and SPG ratio that are used. This paper provides an analytical method to find this optimal point. Experimental results matched closely with our analytical predictions.

The rest of this paper is organized as follows. Section II describes the methodology used in constructing the analytical model including the variables considered and the failure mechanism. Section III presents the analytical model. Section IV presents the experimental results and conclusions are drawn in Sections V.

## II. Methodology

In order to develop the analytical model, a typical CMOS driver/receiver circuit topology was used. This circuit topology was also used in the SPICE simulations to validate the model. In this topology, the following parameters were varied:

1. Number of Channels
2. Slew Rate
3. SPG Ratio
4. Package

### A. Test Circuit

The circuit used to formulate the model and for simulations is shown in the Figure 1. We used the BPTM 0.1 $\mu$ m [12] technology using BSIM3 model cards [13]. All simulations were done using SPICE [14]. A CMOS inverter was used to model the driver and the receiver load. The driver was designed to drive a 75 $\Omega$  PCB trace which was 2" long with a drive strength of 25mA. The CMOS inverter had  $V_{DD}=1.5$ v and  $V_{SS}=0$ v. A series termination resistor was placed on the PCB at the output pins of the driving IC. The resistor value was chosen so that the cumulative output impedance of the resistor in series with the  $R_{ON}$  of the inverter is 75 $\Omega$ . The optimal size of the inverter that can drive 25mA into a 75 $\Omega$ , 2" long PCB trace that was series terminated with an equivalent output impedance of 75 $\Omega$  is  $W_N=80\mu$ m and  $W_P=260\mu$ m. The inverter is sized to have an equal drive strength on both the PMOS and NMOS transistors by using  $(W_P/W_N) = (u_n/u_p) = 3.25$  [4].

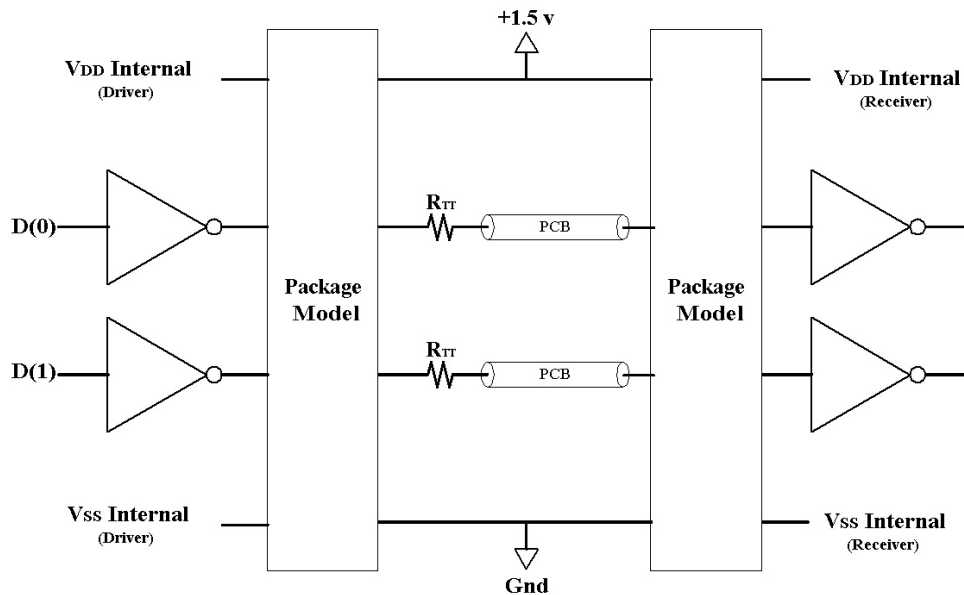


Figure 1. Test Circuit Used to Analyze Bus Configurations (showing a 2:1:1 SPG Ratio)

The package model included the self RLC of the leads and wire bonds (if used). The model included coupling capacitance out to the nearest two adjacent signals for both the package leads and the wire bonds. The mutual inductance of the leads and wire bonds was considered out to the nearest 5 signals. Coupling was not considered on the PCB since the geometries on the PCB are such that coupling can be and often is eliminated with trace spacing [5,15].

## B. Failure Condition

In our model, a failure was defined as ground bounce (or VDD droop) that had a magnitude greater than 5% of the VDD supply. The magnitude of the ground bounce was measured on the die of the driver (VSS-Internal-Driver). The worst case ground bounce was present when all of the CMOS inverters switched their outputs from a logic 1 to a logic 0 at the same time. This failure mechanism only accounts for the magnitude of the ground bounce. Other limitations such as delay and signal shape were not considered.

It was found that the power supply droop during a low-to-high transition was of the same order of magnitude as the ground bounce during a high-to-low transition. This was due to the fact that the CMOS inverter was sized so that the PMOS and NMOS transistors had the same AC characteristics and the number of VSS and VDD pins were matched. Based on this observation, only the ground bounce was monitored for failure conditions. A similar failure condition could have been created that monitored only power supply droop, but the results would have been identical.

## C. Ground Bounce

There are two factors that contribute to ground bounce. The first component is due to the voltage induced across the self inductance of the VSS pin of the driver. This voltage follows the relationship:

$$V_1 = L_{11} \frac{di_1}{dt} \quad (1)$$

The AC current ( $i_1$ ) in this expression is the cumulative drain current of the CMOS inverters as they transition in the same direction. This current is directly proportional to the number of inverters that are switching. The subscripts on  $V_1$  and  $i_1$  represent the fact that voltage on the ground pin is caused by the current through the same ground pin. This current induces a voltage across the self inductance of the pin ( $L_{11}$ ).

The second component is due to the mutual inductance from neighboring signal pins. This contribution follows the relationship:

$$V_1 = M_{1k} \frac{di_k}{dt} \quad (2)$$

For this type of contribution to ground bounce, the voltage ( $V_1$ ) induced on the ground pin is caused by the mutual inductive coupling from adjacent signal pins that are transitioning. The subscript  $k$  represents an arbitrary neighboring pin that is  $k$  pins away from the VSS pin. The current  $i_k$  in this  $k^{th}$  neighboring pin induces a voltage across the mutual inductance  $M_{1k}$  of the ground pin and the  $k^{th}$  neighbor.

It is clear to see that increasing the number of signals that are switching on a package will increase the amount of ground bounce that will occur. As mentioned earlier, a common way to combat ground bounce is to increase the number of ground and power pins on a package. This has the effect of reducing the equivalent self and mutual inductance in the ground return path. This decreases the ground bounce contribution in both equations 1 and 2. Moving toward advanced packaging also has the effect of reducing both the self and mutual inductance of the package.

#### D. Slew Rate

$di/dt$  is proportional to the slew rate of the bus signals. As the slew rate increases, the amount of time it takes for the charging and discharging of the load decreases which increases the data rate at which the bus can operate. This also means that as the slew rate gets faster, the more ground bounce will be present and thus limit the maximum data rate that the bus can run at.

The slew rate  $dv/dt$  can be found as follows:

$$slewrate = \frac{dv}{dt} = \frac{di}{dt} \cdot Z_{load} \quad (3)$$

The rise time of the signal is defined as the time it takes to switch from 10% to 90% of the DC output value (80% of  $V_{DD}$ ).

$$t_{rise} = \frac{0.8 \cdot V_{DD}}{slewrate} \quad (4)$$

The rise time can then be used to define the minimum Unit Interval (UI) that can be used in a robust digital system [2,5,15]:

$$UI_{min} = (1.5) \cdot (t_{rise}) \quad (5)$$

The  $UI_{min}$  defines the minimum time that the data valid window must be present in order to transmit a logic symbol successfully. This corresponds to the maximum data rate of a signal as follows [2,5,15]:

$$DR_{max} = \frac{1}{UI_{min}} \quad (6)$$

## E. Packaging

The package selection dictates the magnitudes of the electrical parasitics present in the inter-chip bus. Packages traditionally add a large inductive component to the I/O system. This inductiveness results in ground (and supply) bounce (equations 1 and 2). As package technology advances, the electrical parasitics are reduced [1]. However, these advanced packages add to the overall cost of the IC [3]. In this paper, we study three industry packages, the QFP wire bond, BGA wire bond, and the BGA flip chip [3,16,17,18].

### i. QFP, Wire Bond Package

One of the most widely used packages over the past 10 years has been the QFP (Quad Flat Pack) with wire bonding. This package is attractive due to its relatively simple assembly in addition to its ability to easily be loaded onto a PCB. Wire Bonding from the die to the lead frame has been refined over the years to yield a robust and efficient assembly process. This has driven the cost out of this package.

The drawback of this package is that as rise times decrease into the multiple nanosecond range, the electrical parasitics cause significant noise [19]. The lead frame itself contains a large amount of inductive and capacitive coupling between signals. In addition, the dense wire bonding pattern from the die to the lead frame has high mutual inductive coupling that can induce a voltage on neighboring wire bonds many signals away. The mutual inductive coupling causes severe ground bounce and can cause the package to resonate. Figure 2 shows the cross-section of a typical QFP package using wire bonding technology.

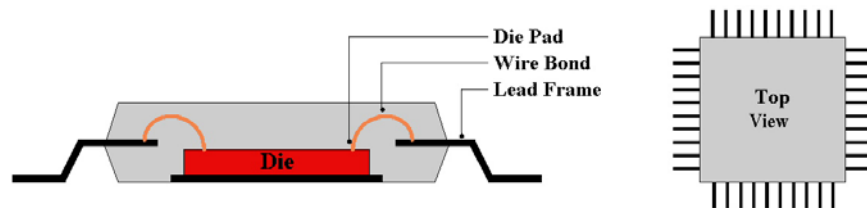


Figure 2. Cross-Section of QFP, Wire Bond Package

### ii. BGA, Wire Bond Package

BGA (Ball Grid Array) packaging emerged in the late 1990's as a way to increase the density of IC packages. This package reduces the coupling within the lead frame that is present in the QFP package. However, the same coupling issues remain within the wire bonds that connect the die to the PCB. The technology to implement the BGA connection is slightly more expensive than the QFP lead frame processing [3]. Figure 3 shows the cross-section of a typical BGA package using wire bonding technology.



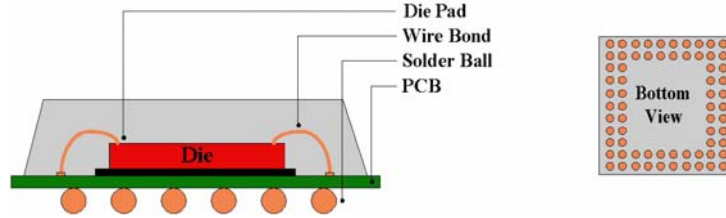


Figure 3. Cross Section of a BGA, Wire Bond Package

iii. BGA, Flip-Chip Package

The most recent package to emerge is the BGA package using flip-chip technology to connect the die to the package PCB. In this style of packaging, the die has an array of pads on its outer most metal layer. The die is flipped upside down and mounted to a complementary array on the package PCB. The process technology used to connect the die to the package PCB is similar to the BGA connection to the target PCB. This involves solder bumps that are reflowed to form the connection. This style of packaging has all of the benefits of a standard BGA package in that it reduces the coupling associated with a lead frame and greatly increases the pin density per area. Its most attractive characteristic is that it alleviates the problem associated with mutual inductive coupling present in wire bond technology. The one disadvantage is that the process time for the solder reflow and under fill diffusion takes longer than the industry standard wire bonding. This causes this package to be more expensive. However, its electrical performance outweighs its cost when designing high-speed inter-chip busses. Figure 4 shows the cross-section of this package.

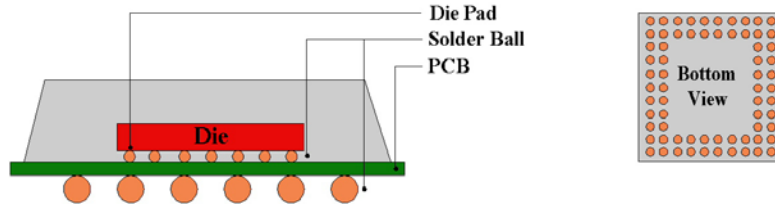


Figure 4. Cross Section of a BGA, Flip Chip Package

Table I shows the electrical parameters and averaged per-pin cost for the three packages studied in this paper.

Package	$L_{11}$	$K_{12}$	$K_{13}$	$K_{14}$	$K_{15}$	$K_{16}$	Package	Cost Per-Pin
QFP-wb	4.550n	0.744	0.477	0.352	0.283	0.263	QFP-wb	\$0.22
BGA-wb	3.766n	0.537	0.169	0.123	0.097	0.078	BGA-wb	\$0.34
BGA-fc	1.244n	0.630	0.287	0.230	0.200	0.175	BGA-fc	\$0.63

Table I. Electrical and Cost Characteristics for Packages Studied

### III. Analytical Model

#### A. Performance of the Bus

This section presents an analytical model that describes the maximum data rate for an inter-chip bus considering the magnitude of ground bounce on the IC as the failure condition.

Using equations 1 and 2, the net ground bounce of a bus can be expressed as:

$$V_{gnd-bnc} = \left( \frac{W_{bus} \cdot L_{11}}{N_g} \right) \left( \frac{di}{dt} \right) + \sum_{k=2}^{W_{bus}} \left( M_{1k} \frac{di}{dt} \right) \quad (7)$$

In this expression,  $W_{bus}$  is the number of signals in the bus. For this model, it is assumed that all of the signal in the bus are transitioning in the same direction to represent the worst case ground bounce situation.  $N_g$  is the number of ground pins in the bus and is dictated by the SPG ratio that is selected. Increasing the number of grounds will have the effect of reducing the inductance of the ground path.  $i$  is the current in any pin.

$V_{gnd-bnc}$  is set to an acceptable magnitude ( $p \cdot V_{DD}$ , where  $p < 1$ ) depending on the desired noise margin for the bus. Therefore the maximum slew rate achievable for an inter-chip bus is:

$$\left( \frac{dv}{dt} \right)_{\max} = \frac{p \cdot V_{DD} \cdot Z_{load}}{\left( \frac{W_{bus} \cdot L_{11}}{N_g} \right) + \sum_{k=2}^{W_{bus}} M_{1k}} \quad (8)$$

From equation 4, we get the minimum tolerable rise time as:

$$t_{rise-min} = \frac{(0.8) \cdot \left[ \left( \frac{W_{bus} \cdot L_{11}}{N_g} \right) + \sum_{k=2}^{W_{bus}} (M_{1k}) \right]}{p \cdot Z_{load}} \quad (9)$$

The minimum Unit Interval can be computed by combining equations 5 and 9. We can then use equation 6 to get an expression for the maximum per-pin data rate  $DR_{max}$  that can be achieved:

$$DR_{\max} = \frac{p \cdot Z_{load}}{(1.5) \cdot (0.8) \cdot \left[ \left( \frac{W_{bus} \cdot L_{11}}{N_g} \right) + \sum_{k=2}^{W_{bus}} M_{1k} \right]} \quad (10)$$

The total system throughput  $TP$  of the bus can now be expressed as  $TP = (DR_{max} \cdot W_{bus})$ .

## B. Cost-Effectiveness of Bus

We now formulate a method to analyze the cost-effectiveness of the bus design. The following expression defines the number of I/O pins needed to implement an inter-chip bus of width  $W_{bus}$  with an equal number of VSS and VDD pins set by the SPG ratio:

$$N_{I/O} = W_{bus} + 2 \cdot \left\lceil \left( \frac{W_{bus}}{SPR} \right) \right\rceil \quad (11)$$

In this expression,  $SPR$  refers to the Signal/Power/Ground ratio. For example, if  $SPG = k:1:1$ , then  $SPR = k$ . The cost of the inter-chip bus is given by:

$$Cost_{bus} = (N_{I/O}) \cdot (Cost_{per-pin}) \quad (12)$$

where the  $Cost_{per-pin}$  will vary depending on which package is selected.

Finally, we define a cost effectiveness metric for any bus configuration called *Bandwidth-per-Cost (BPC)*. This metric has units  $Mb/\$$  and takes into account the total bus throughput for a given inductive noise margin as well as the I/O cost including the number of the power and ground pins.

$$BPC = \left( \frac{TP}{Cost_{bus} \cdot 1e^6} \right) \quad (13)$$

## IV. Experimental Results

### A. QFP Wire Bond Results

Using the methodology outlined in section II, we simulated the test circuit and compared the results with the analytical model. Figure 5 shows the maximum data rate per-pin ( $DR_{max}$ ) for a QFP package with wire bonding as a function of the number of channels that are simultaneously switching. Both the simulation and analytical model data are displayed. These results illustrate that as the number of simultaneously switching channels is increased, the per-pin data rate is decreased. In addition, it shows that the effect of adding more grounds can increase the per-pin data rate by reducing the self and mutual inductance in the ground path.

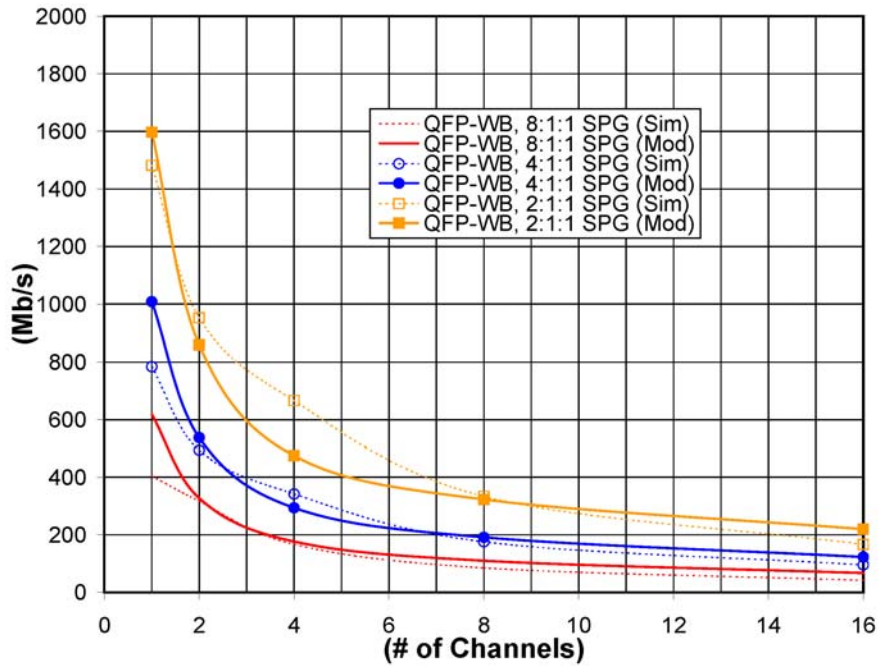


Figure 5. Maximum Data Rate Per-Pin for a QFP Wire Bonded Package

Figure 6 shows the total throughput ( $TP$ ) of the bus for the same package. This figure shows that the system throughput actually approaches an asymptotic limit as more channels are added to the bus. This is due to the fact that adding more channels to the bus actually degrades the speed at which each individual channel can switch. *The linear increase expected by adding additional I/O is negated to the dramatic decrease in per-pin performance due to the package parasitics.*

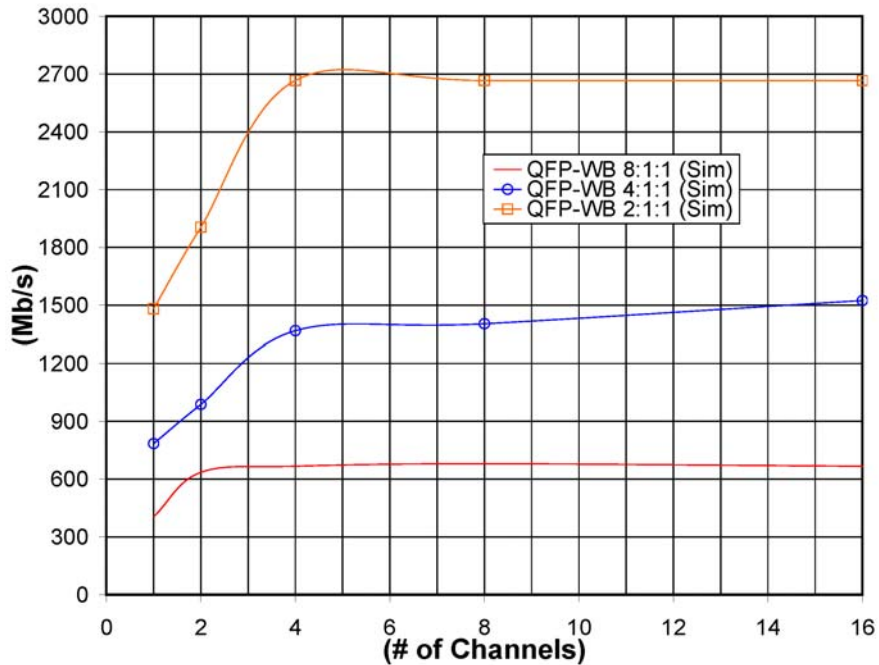


Figure 6. Total System Throughput for a QFP Wire Bonded Package

## B. BGA Wire Bond Results

We performed the same experiment on a BGA Wire Bond Package. Figures 7 and 8 respectively show the maximum data rate per-pin  $DR_{max}$  and the maximum throughput  $TP$  as a function of channels switching. Again, the simulation results and analytical model data are presented to verify the analytical model's accuracy.

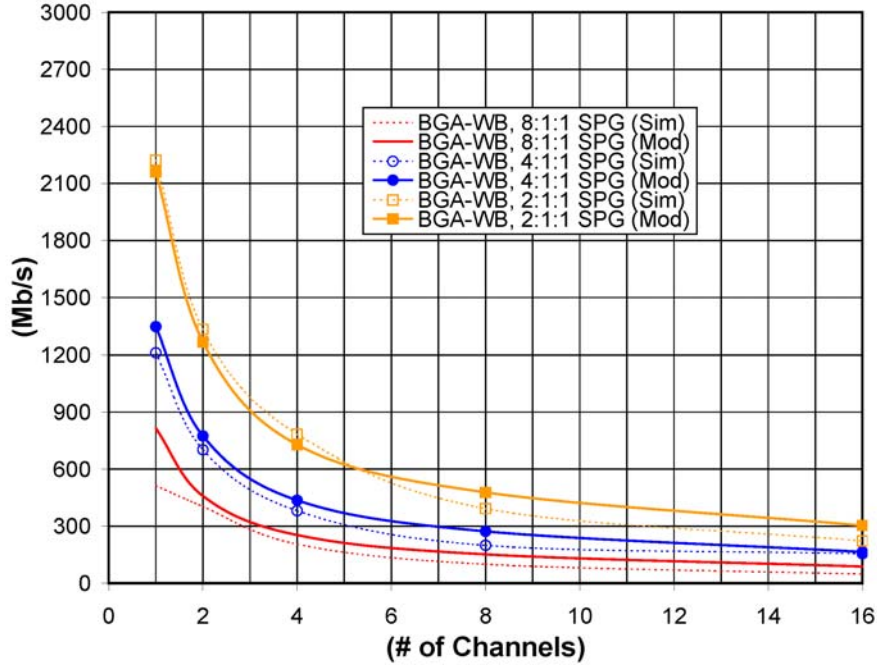


Figure 7. Maximum Data Rate Per-Pin for a BGA Wire Bonded Package

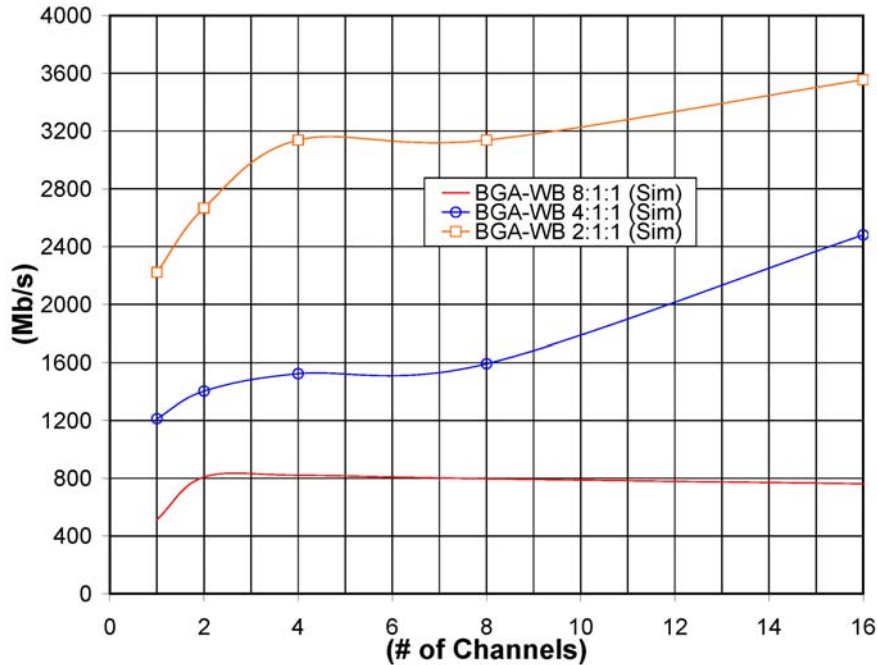


Figure 8. Total System Throughput for a BGA Wire Bonded Package

The BGA Wire Bond package has slightly better electrical performance over the QFP wire bonded package. The main advantages arise due to the elimination of the coupling and self inductance within the lead frame.

### C. BGA Flip-Chip Results

Figures 9 and 10 respectively show the maximum data rate per-pin  $DR_{max}$  and the maximum system throughput  $TP$  as a function of the number of simultaneously switching channels for a BGA Flip Chip package. The electrical advantages of the flip-chip package are evident in Figure 9. The single channel data rate that can be achieved with sufficient grounding is over twice the frequency of the BGA with wire bonding and over three times that of the QFP with wire bonding. Figure 10 shows that flip-chip technology is still vulnerable to the simultaneous switching problems that the other packages have, albeit at a higher frequency. The maximum throughput still approaches an asymptotic limit as the number of channels is increased.

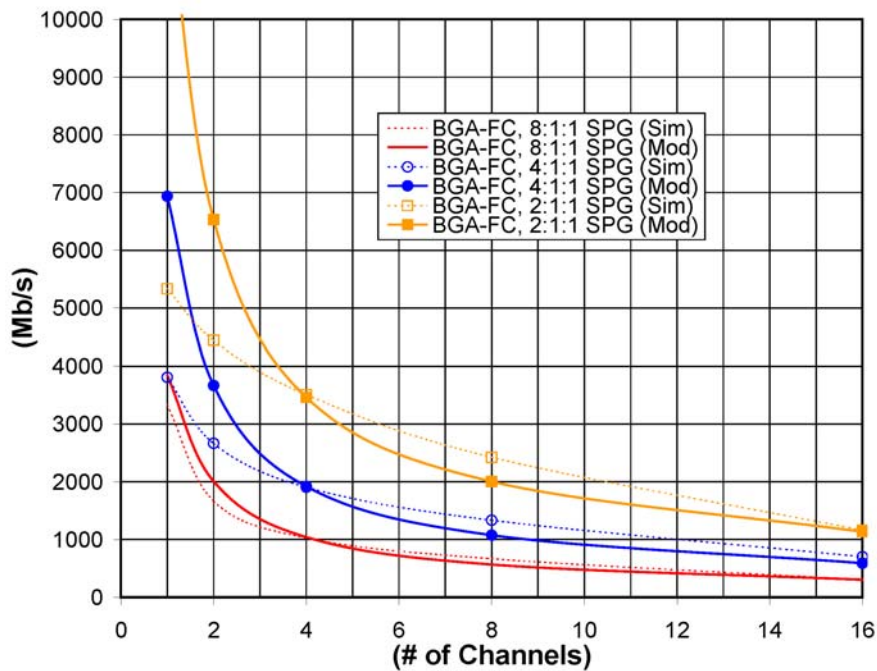


Figure 9. Maximum Data Rate Per-Pin for a BGA Flip-Chip Package

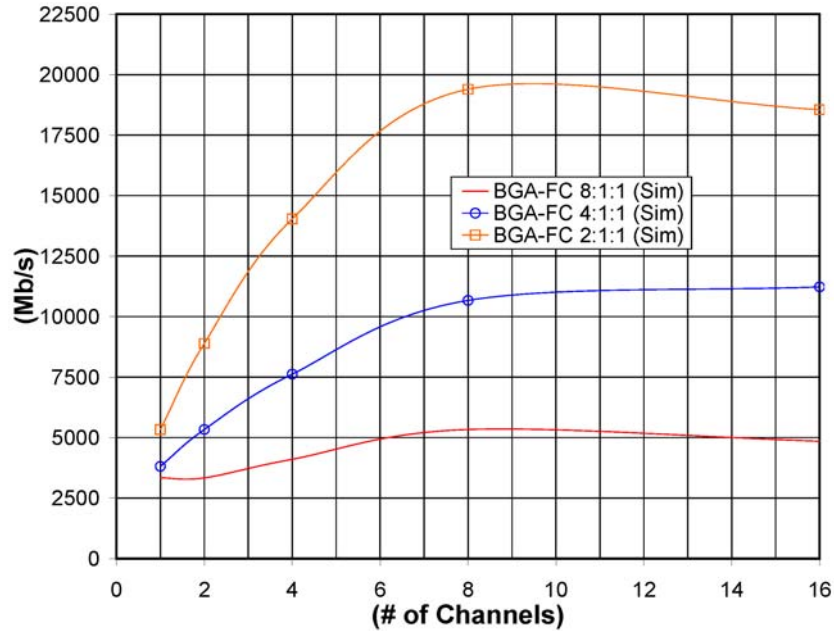


Figure 10. Total System Throughput for a BGA Flip-Chip Package

#### D. Cost Analysis

All of the packages that were analyzed reached an asymptotic limit in total throughput as the width of the bus was increased. In all cases this was due to the ground bounce failure mechanism decreasing the maximum data rate per-pin at a rate that was similar to the increase in the throughput achieved by adding channels. This indicates that after the failure mechanism begins to dominate the per-pin performance, simply adding I/O to the bus does not increase system throughput. A more thorough analysis of this should include the cost of the bus. This section performs such a cost analysis for the three packages by considering the maximum throughput as well as I/O cost as channels are added.

The metric introduced in Equation 13 represents the cost-effectiveness of an inter-chip bus. This metric considers the SPR in the cost of the I/O, providing insight into the most cost-effective bus configuration.

Table II shows the number of I/O pins needed to implement the various bus configurations considered in this paper. This table accounts for the number of VDD and VSS pins as different SPR's selected.

Bus Configuration	Number of Channels				
	1	2	4	8	16
QFP-WB 8:1:1	3	4	6	10	20
QFP-WB 4:1:1	3	4	6	12	24
QFP-WB 2:1:1	3	4	8	16	32
BGA-WB 8:1:1	3	4	6	10	20
BGA-WB 4:1:1	3	4	6	12	24
BGA-WB 2:1:1	3	4	8	16	32
BGA-FC 8:1:1	3	4	6	10	20
BGA-FC 4:1:1	3	4	6	12	24
BGA-FC 2:1:1	3	4	8	16	32

Table II. Number of I/O Pins Needed Per Bus Configuration

Table III shows the cost of the various bus configurations. The effect of a better grounding scheme (i.e.,  $SPG=2:1:1$ ) is that the cost increases at a faster rate as channels are added. This table shows the relative expense between the packages using per-pin cost data from Table I in Section II. The BGA wire bond package is only slightly more expensive than the QFP wire bond. Moving toward the more advanced packaging such as BGA and flip chip assembly will increase the cost of I/O.

Bus Configuration	Number of Channels				
	1	2	4	8	16
QFP-WB 8:1:1	0.66	0.88	1.32	2.20	4.40
QFP-WB 4:1:1	0.66	0.88	1.32	2.62	5.28
QFP-WB 2:1:1	0.66	0.88	1.76	3.52	7.04
BGA-WB 8:1:1	1.02	1.36	2.04	3.40	6.80
BGA-WB 4:1:1	1.02	1.36	2.04	4.08	8.16
BGA-WB 2:1:1	1.02	1.36	2.72	5.44	10.88
BGA-FC 8:1:1	1.89	2.52	3.78	6.30	12.60
BGA-FC 4:1:1	1.89	2.52	3.78	7.56	15.12
BGA-FC 2:1:1	1.89	2.52	5.04	10.08	20.16

Table III. Cost of I/O Per Bus Configuration (\$)

Table IV shows the  $BPC$  for the three different packages. This table illustrates that it is more cost-effective to use busses that are narrower and faster rather than expanding the bus which actually decreases the data rate per-pin.



Bus Configuration	Number of Channels				
	1	2	4	8	16
QFP-WB 8:1:1	612	722	505	309	152
QFP-WB 4:1:1	1188	1122	1036	532	289
QFP-WB 2:1:1	2245	2165	1515	758	379
BGA-WB 8:1:1	503	594	402	234	112
BGA-WB 4:1:1	1188	1032	747	390	304
BGA-WB 2:1:1	2179	1961	1153	577	327
BGA-FC 8:1:1	1764	1323	1085	847	385
BGA-FC 4:1:1	2016	2116	2016	1411	743
BGA-FC 2:1:1	2822	3527	2785	1924	920

Table IV. BPC of Different Bus Configurations

## V. Conclusion

In this paper we presented an analytical performance model for inter-chip bus design. Our model considered performance and cost as the number of channels, grounding scheme, and various packaging options were explored. We demonstrated that the maximum data rate per-pin decreased significantly as the number of simultaneously switching channels was increased. This shows that *simple expansion of an inter-chip bus does not yield a linear increase in the throughput* of the system as one would expect. It was also shown that the total system throughput reached an asymptotic limit as the number of channels was increased. This means that the same throughput can be achieved by using faster narrower busses rather than a traditional wider and slower bus design.

A cost analysis was also performed which considered various packaging and grounding schemes. A new Bandwidth per Unit Cost (*BPC*) metric was defined as a means to evaluate the most cost-effective bus configuration. It was found that the most cost-effective bus was faster and narrower rather than slower and wider. By running the individual channel near its theoretical maximum data rate (i.e., with no mutual inductive coupling), a cost advantage is achieved because additional I/O are not needed to obtain the desired system throughput. The BGA Flip-Chip package was found to be the most cost effective. Even though the cost per channel is higher for this advanced style of package, the increased bandwidth far outweighs the cost increase when considering *BPC*. For all packages studied, the optimal bus configuration occurred at the inflection point of where adding I/O pins increased the throughput of the bus at such a small rate that the cost increase negates adding I/O pins. The technique presented in this paper to analyze the cost-effectiveness of a bus configuration (considering cost, package, and grounding schemes) can be applied to any style of packaging.

## References

- [1] M. Horowitz, C. Yang, and S. Sidiropoulos, "High-Speed Electrical Signaling: Overview and Limitations", *IEEE Micro.*, vol.18, pp. 12-24, Jan, 1998.
- [2] C. Kinnaird, "Standards are key to optimizing high-speed data bus communications," *Planet Analog* (planetanalog.com), Oct 07, 2002.
- [3] Agilent Technologies, Inc. Packaging Group, Ft. Collins, CO, Personal Communication, 2004.
- [4] W. Dally and J. Poulton, *Digital Systems Engineering*, Cambridge University Press, Cambridge, U.K., 1998.
- [5] H. Johnson and M. Graham, *High-Speed Digital Design*, Prentice Hall PTR, 2003.
- [6] N. Hirano, M. Miura, Y. Hiruta, T. Sudo, "Characterization and Reduction of Simultaneous Switching Noise for a Multilayer Package", *Proceedings of 44th Electronic Components and Technology Conference*, pp. 494-956, May, 1994.
- [7] E. Mejia-Motta, F. Sandoval-Ibarra, J. Santana, "Design of CMOS buffers using the settling time of the ground bounce voltage as a key parameter", *Proceedings of 43rd IEEE Midwest Symposium on Circuits and Systems*, vol. 2, pp. 718, Aug, 2000.
- [8] B. Young, "Return path inductance in measurements of package inductance matrixes", *IEEE Transactions on Components, Packaging, and Manufacturing Technology*, vol. 20, pp. 50-55, Feb, 1997.
- [9] "Rapid IO Trade Association", [www.rapidio.org/](http://www.rapidio.org/).
- [10] "PCI-SIG Trade Organization", [www.pcisig.com/](http://www.pcisig.com/).
- [11] W.H. Dally, J. Poulton "Transmitter Equalization for 4-Gbps Signaling", *IEEE Micro*, vol. 1, ver. 1, pp. 48-56, Feb, 1997.
- [12] "BPTM Homepage", [www-device.eecs.berkeley.edu/~ptm/](http://www-device.eecs.berkeley.edu/~ptm/).
- [13] "BSIM3 Homepage", [www-device.eecs.berkeley.edu/~bsim3/](http://www-device.eecs.berkeley.edu/~bsim3/).
- [14] L. Nagel, "SPICE: A Computer Program to Simulate Computer Circuits", University of California, Berkeley UCB/ERL Memo M520, May, 1995.
- [15] H. Johnson and M. Graham, *High-Speed Signal Propagation*, Prentice Hall PTR, 2003.
- [16] ASAT Inc., "Application Note: Peak Performance Enhanced Lead Packages", [www.asat.com/](http://www.asat.com/).
- [17] ASAT Inc., "Application Note: Peak Performance Array Standard, HS fpBGA Packages", [www.asat.com/](http://www.asat.com/).
- [18] ASAT Inc., "Application Note: Peak Performance Flip Chip Packages", [www.asat.com/](http://www.asat.com/).
- [19] M. Miura, N. Hirano, Y. Hiruta, T. Sudo, "Electrical characterization and modeling of simultaneous switching noise for leadframe packages", *Proceedings of 45th Electronic Components and Technology Conference*, pp. 857-846, May, 1995.