

---

# *Throughput Expansion with FET Based Crossbar Switching*

*by Refugio Jones*

## **Introduction**

Today's embedded computing applications, particularly DSP and array processing, have ever-escalating requirements for bandwidth. Radar, medical imaging, geological...you name it, require massively parallel array processing to complete tasks. That means having many processors, the more powerful the better, working together on a problem to solve it and display results quickly. The weak link in a massively parallel system is not the processor however. Processor manufacturers are introducing faster and faster chips at breakneck speed. Processors today provide more than enough horsepower for today's applications, and newer, faster processors are coming on the scene all the time.

No, the weak link is the bus that connects the processors together so that they can share data and memory. Today's buses simply can't keep pace with processors. While the fastest processors today are exceeding 2GHz, the fastest buses out there are lucky to break 133 MHz. Clearly, the bus is the bottleneck in any multiprocessing system. The delay effect of buses is magnified exponentially when you have a system with hundreds of processors, typical of today's high-end DSP and array processing applications.

Bandwidth is affected by three factors: 1) width of the bus, measured in bits (that's how many bits per clock cycle can be transferred); 2) clock speed, measured in MHz (that's how fast those bits can be moved, or, alternatively, how quick a clock cycle is); and 3) the number of concurrent bus transactions, or "conversations" that can occur on a bus at once (that's how many data transfers between two processors, or a processor and memory, that can occur simultaneously).

Established bus standards, such as PCI have set bit width and clock speeds. What you see is what you get. If your bandwidth needs to exceed the bandwidth of the bus, you're out of luck.

## ***Enter the crossbar***

The crossbar does an amazing thing: it multiplies the number of conversations that can take place at once. Think of a crossbar as a central hub with several pathways, as opposed to a bus' one and only gate. Through this hub's pathways pass multiple conversa-

tions, all happening at the same time. Suppose a hub has three pathways with a port on each end of the pathway (six total), like RACEway<sup>™</sup>. Each pathway allows a point-to-point (processor-to-processor or processor-to-memory) conversation. Using three pathways results in three concurrent conversations. So, if the bandwidth of the bus is 132 MB/second, with three pathways, you get three times that bandwidth, or 396 MB/second.

The more crossbars you have in a system, the more conversations you can have at once. A typical topology of crossbars is the "tree" structure, with levels of crossbars. The bigger the tree (and hence the more crossbars you have) the higher the number of processors that can be linked to other processors. So, massively parallel systems not only have lots of processors, but lots of crossbars as well.

Obviously, crossbars are the way to go if you need expanded throughput or bandwidth. Several companies now offer crossbar "switching fabrics" allowing you to set up multi-level trees of crossbars and up to a thousand processors working together. The competing architectures are all proprietary to the company that invented them. As a result, these companies can and do charge top dollar for the hardware based on their crossbar. They further differentiate their offerings by allowing their systems to work only with their own proprietary software. Therefore, once a customer decides to go with a proprietary architecture, he/she is locked into that architecture and is at the mercy of the vendor's pricing, not to mention the exigencies of using a proprietary operating system. Current examples of such architectures are RACEway<sup>™</sup> by Mercury Computer Systems, SKY Channel<sup>™</sup> by Sky, and Myrinet<sup>™</sup> by CSPI.

Newer dedicated DSP boards are incorporating crossbars right on the board. They allow multiprocessors on the same board to talk to each other and a shared global memory at higher speeds than possible before. However, when these boards must talk to other boards, they are limited to a single conversation on the bus connecting multiple boards.

A further limitation of today's crossbars is the way switching is handled. The common methods, MUX (multiplexing) and "Sea of Gates", both described in detail later in this paper, are inefficient, resulting in switching delays. There is room for improvement in this area.

Pericom has taken a long, close look at the state of current crossbar technology. After much research and development, a superior crossbar technology has emerged. Pericom's PI5X1018 crossbar improves every variable affecting bandwidth: number of concurrent conversations, bit width, and clock speed. The PI5X1018 features five pathways (with two ports each), allowing up to five conversations at one time. Each port is an ample 18 bits wide, and can run at an unprecedented 133 MHz.

In addition, Pericom used an entirely different switching method, called FET. FET allows much faster switching than either the Sea of Gates approach or the MUX approach. So not only are conversations faster before, but switching between completed conversations to new conversations is faster as well.

Pericom also went one very important step further. The PI5X1018 is not constrained to a proprietary bus, nor to just one bus. The crossbar is designed to work in any bus architecture. Now an engineer can select any industry standard bus, such as the ubiquitous PCI for instance, and use commonly available parts and operating systems to put his/her system together.

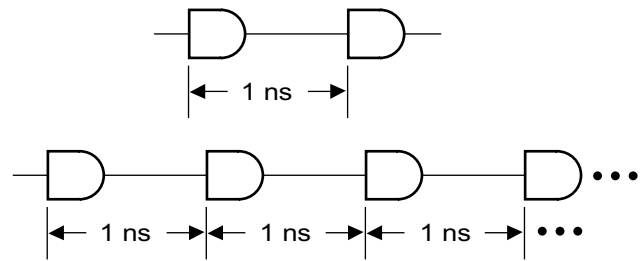
With the right bus (PCI is a good one) a system with the PI5X1018 will enjoy significantly higher bandwidth than systems using the proprietary crossbar fabrics that are currently available. And as higher-bandwidth buses become adopted by the industry, the PI5X1018 will support them as well. No other crossbar technology gives you such a wide and free upgrade path.

With the PI5X1018, the future bandwidth potential of massively parallel systems has been greatly expanded.

The following sections will discuss in more technical detail a comparison of crossbar technologies, the superior features and corresponding benefits of the PI5X1018, and applications that will benefit from the PI5X1018, including a real-life example of a company already using the PI5X1018.

### Discussion of Available Technologies

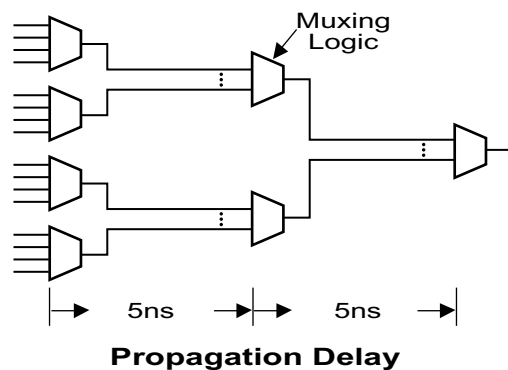
Digital FET technology offers advantages over using standard logic gates in terms of propagation delay and low quiescent power. Pericom's BusSwitch technology typically runs at 250ps or less, which contributes no major delays. In contrast, gates switch at 10ns or slower which can burden tight timing budgets. In the case of say an ASIC solution, gate switching is fast (<1ns), but each additional layer in a 'Sea of Gates' results in additional delays. Figure 1 basically demonstrates the additive nature of delays through a dense ASIC solution.



**Propagation Delay Increases w/ Each Layer of Gates**

**Figure 1. ASIC Sea of Gates Model for Crossbar Switching**

Yet another means of implementing a crossbar consists of the use of multiplexing (MUX) logic. The clear advantage of such a model is the configurable nature of the switch matrix. Such a topology bridges the gap between ASIC and individual bus switching. On the other hand, each individual MUX layer can effectively switch at rates of roughly 5ns. Thus each subsequent layer adds further propagation delay to the switching of a signal. Figure 2. demonstrates the slow switching nature inherent in MUX switching architectures.



**Figure 2. MUX Crossbar Matrix Topology**

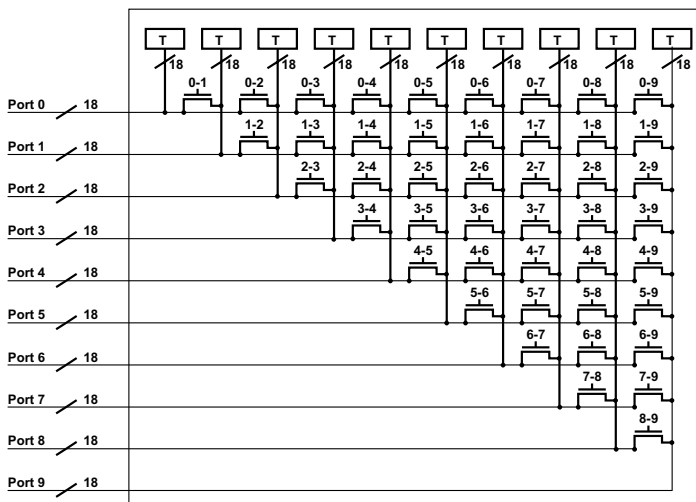
### FET Crossbar Solution

With very low propagation delays (2ns typical), the FET switch functions in application environments such as the 66 MHz PCI bus or it can fit into almost any environment requiring greater bandwidth. Pericom's FET 10 port x 18 bit Crossbar switch consumes as little as 1mA of current. The low current draw of the device makes it an attractive fit for low power consumption designs.

The Pericom solution is comprised of a bare switch matrix without glue logic. Such a device allows for immediate implementation in general applications where routing logic can be implemented at the board level. The switch itself represents a building block where custom circuitry is built around the matrix so as to utilize its functionality without having additional built-in overhead. This topology clearly maximizes the advantage of the fast FET switching matrix. No delays are introduced into the system through the switching of signals.

The bare matrix approach allows for greater flexibility in designs where greater throughput is required in combination with custom routing logic. Implementation is simplified to the use of external programmable logic to intelligently route signals through the device.

The matrix block itself is comprised of several groups of individual matrices corresponding to each of the 18 bits. A general representation of the matrix block is shown in Figure 3. below.



**Figure 3. PI5X1018 Crossbar Switch Block Diagram**

The device is comprised of ten, 18 bit wide ports. It has 45 pin level switch controls, which control the individual switch matrices found in each of the 10 ports. Each matrix is composed of a group of 45 stacked FET switches, which are thus controlled by the 45 control pins.

**Pull-up Resistors**

As an additional feature, each input/output pin of the PI5X1018 has an internal weak pull-up resistor that can be adjusted through biasing via the device’s VBIAS pin from 5KΩ to 100KΩ. These pull-ups improve the performance of digital signals by effectively decreasing the rise times of the signal edges. Furthermore, due to the nature of the FET switch, the pull-up feature helps guarantee consistent voltage levels on the output of the device. Finally, for the control signals on pins S<sub>N-M</sub>, the internal pull-up resistors ensure that the unused signals are not left floating and thereby help in robustness.

**Throughput Expansion with the PI5X1018**

Crossbar solutions offer greater throughput switching. Pericom’s Crossbar is configured as a 5 x 5 device (5 through connections) and offers a standard per port throughput of 148 MBytes/sec at 66 MHz and 300 MBytes/sec at 133 MHz. This is calculated using the formula given below.

$$\text{Port to Port Throughput} = \frac{\text{Frequency (MHz)} \times \text{Bus Width (bits)}}{\text{Word Length (bits/Byte)}}$$

The aggregate throughput of a Crossbar device takes into consideration the total number of ports available. This results in the maximum number of possible connections that the device can accommodate at any given time. This can be calculated by expanding upon the formula used for port to port throughput to include the configuration as shown below.

$$\text{Aggregate Throughput} = \frac{\text{Configuration} \times \text{Frequency (MHz)} \times \text{Bus Width (bits)}}{\text{Word Length (bits/Byte)}}$$

$$\begin{aligned} \text{Aggregate Throughput of One PI5X1018} &= \frac{5 \times 66 \text{ MHz} \times 18 \text{ bits}}{8 \text{ bits/Byte}} \\ &= 742 \text{ MBytes/sec} \end{aligned}$$

Furthermore, one PI5X1018 has a maximum aggregate throughput (for 8-bits/Byte) of 1.5 GBytes/sec at 133 MHz.

For even greater bandwidth expansion, Crossbar devices can be laid out in a variety of ways including in parallel as in Figure 4. Such a layout, supports a wider bus without decreasing the number of usable ports. In such cases, the throughput increases by increasing the Bus Width in either of the formulas listed above. For a 64-bit bus, four PI5X1018N devices (72 bits wide) are laid out in parallel and the resulting aggregate throughput will be ~3 GBytes/sec at 66 MHz.

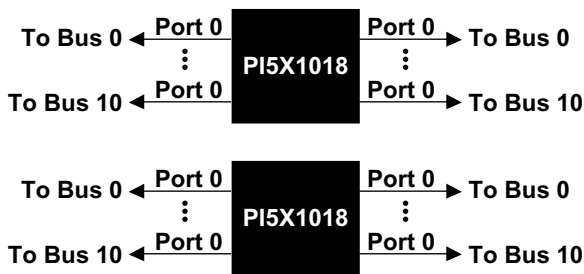


Figure 4.

Parallel layout of crossbar switches for wider bus widths

## Applications of Pericom's Crossbar Switch

### 1. Industrial and DSP Computing

PCI is quickly becoming the architecture of choice in industrial computing under the Compact PCI specification created by the PCI Industrial Computing Manufacturers Group (PICMG). Back plane architectures are typically used in industrial or medical systems where boards can be quickly swapped for upgrades. Industrial computer manufacturers use this architecture with a common bus such as VME or Compact PCI for system operation. But, they have speed and capacitance issues that limit the total bandwidth and number of usable slots. For example the VME bus has a maximum bandwidth of 80 Mbytes/sec while 64-bit Compact PCI can run between 512 Mbytes/sec and 1Gbytes/sec maximum. To expand the bandwidth and control data traffic on the bus manufacturers have been using discrete fast high-density bus switches. The switches add to the cost of their systems and take up large amounts of board space, but are far less expensive than the proprietary RACEWAY™, as an alternative, the PI5X1018 offers a higher integration solution to address these issues.

In a DSP board application, a Pericom Crossbar is used to increase the bandwidth to over 2 GBytes/sec. The switch enables the DSPs to eliminate I/O bottlenecks. It increases bandwidth by expanding the PCI bus so that several DSP's can carry on non-blocked transfers at standard 66 MHz speeds. Figure 5 illustrates a method of implementing the PI5X1018 for DSP switching.

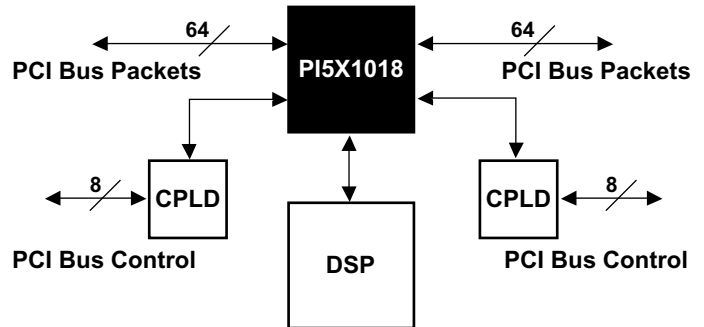


Figure 5. Eliminating DSP Switching Bottlenecks

### 2. Multiprocessor Switching and Memory Sharing

Multi processor applications require shared access to system memory without sacrificing on quick access time. Typically, layering the L2 or L3 cache between the processors and the memory itself solves this problem. This makes the system into a hierarchical memory bus where the first or most important functions have first access to memory. As a result secondary processors must wait for access to memory.

A more elegant solution is to use a crossbar, which allows for the fast switching of the memory bus itself without adding excessive delays into the system. Figure 6 below illustrates one example of a memory sharing application.

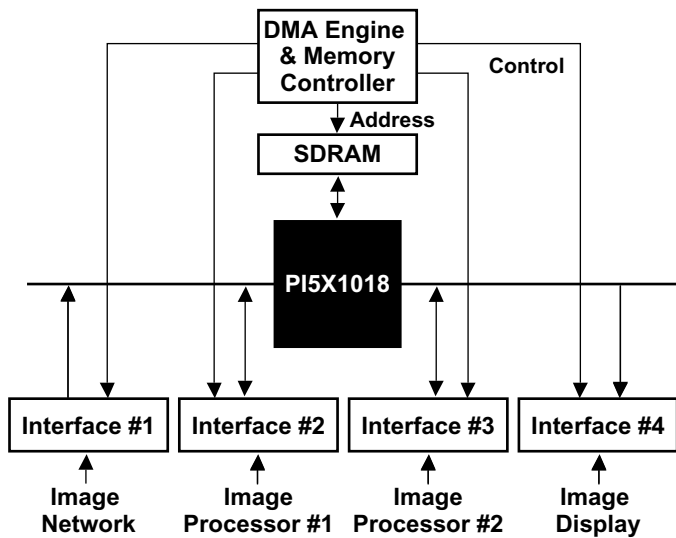


Figure 6. Memory Sharing Application

### 3. Data Communications Switching

The latest LAN and data network switches are designed to have several ports in the same box. In Ethernet the ports are composed of a Media Access Unit (MAC) and a physical layer device (PHY). A Routing Control Block is used to manage the data traffic. For a large number of ports it's common to have switched access to the bus for Routing Control so that the ports can access each other at any given time. Hence as the number of ports increases, so does the switching bandwidth. This increase in switching bandwidth can be done through using a number of discrete bus switch devices.

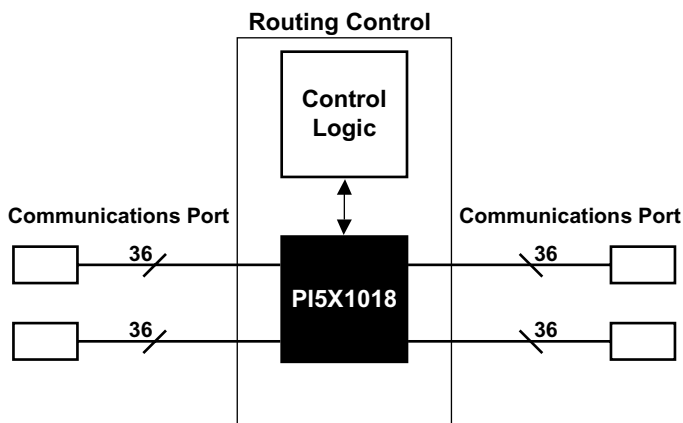


Figure 7. Crossbar in Data Communications Switching

A Crossbar Switch can effectively eliminate the need for several bus switches with a single chip solution. The device allows any port to send or receive data or ‘talk’ to any other device at any time. It can also be used in the one to many mode which would in fact broadcast data to all ports. In addition, the switch can again be used to concentrate data in a, many to one configuration. These two features and the fact that any port on the Crossbar switch can ‘talk’ to any other port make it a very flexible device for data communications applications. Some high-density network switches can require as many as 16 ports in a single box and hence multiple crossbar devices can be used to accomplish the entire switching between ports.

### Summary

FET Crossbar switching can significantly improve performance while reducing the dependence on costly proprietary solutions. Pericom’s crossbar switch technology can easily fit into any bus where low delay switching is needed. Furthermore, the crossbar can replace any multiple discrete switch solution without adding any significant overhead to the application while significantly increasing board space and performance.

Further advantages include its built-in ‘Any Bus’ feature where the PI5X1018 will easily function in any proprietary or generic bus such as VME, RACEWAY™, PCI, and Compact PCI. The fast FET switching matrix allows the Pericom’s crossbar to work at any bus speed up to 133 MHz.

Finally, with a propagation delay of 2ns (max) at 66 MHz, the FET based Crossbar can easily expand the useable bandwidth in applications such as back-plane switching, DSP, and Data Communications.