THE POSSIBILITIES ARE INFINITE

FUJITSU

Reducing Power and Costs of 10G Solutions In Data Centers

Fueled by data-center performance demands and declining price per port, the 10GbE switch market is poised for rapid growth in the next few years.

Gennum and Fujitsu have combined forces to develop 10GbE technology that helps reduce both cost and power. Using SFP+ modules based on Gennum's 10G CDR technology in combination with Fujitsu's 10GbE switch silicon results in significant power savings and cost reductions in the form of lower component count, lower cost manufacturing and a "pay-as-you populate" capability.

This case study discusses in detail how networking systems designers are under constant pressure to reduce cost and power while improving performance and functionality. The latest SFP+ architecture described in this case study represents a material improvement in all these areas, including cost (a $40 savings per port), power consumption (1W savings per port) and performance (guaranteed IEEE compliance) with the margin observed in historical form factors.

gennum

CASE STUDY

Introduction

Several trends in network demands are putting strong pressure on data centers to increase port density and performance while reducing cost per port and power consumption — goals that can only be met by using more highly integrated network switch chips and modules.

To fully support the benefits that 10GbE can provide, it is important to have the ability to maintain the signal path at 10Gb/s serial. Historically, an XAUI interface was used, which required that the 10Gb/s signal be converted into four parallel lanes of 3.125Gb/s. This conversion uses extra power and requires using board space to route four signals for each direction of every 10GbE port. The continued improvement in process technology has enabled Fujitsu to develop 10Gb/s serial ports for the next-generation 10GbE switch chip. The Fujitsu MB86C69RBC switch has 26 ports, each capable of serially transferring data at 10Gb/s, while also supporting XAUI if required, as shown in Figure 1. The device is capable of interfacing directly with optical modules.

Although the integration of the 10Gb/s serial I/Os into the switch silicon can save both power and space, that integration still poses challenges in terms of signal integrity. The 10GbE serial data has a bit period of only 97ps. Furthermore as the signal traverses the PCB, it encounters connectors and vias that degrade the signal integrity. As a result, noise on the signal can quickly accumulate to a level that can cause excessive bit errors. One solution is to place signal-conditioning devices on the PCB, this helps in simplifying layout, enhancing signal integrity and thereby saving time. These devices can receive noisy inputs and transmit a clean output, reducing the burden on the switch silicon I/Os. However, this will consume more board area, cost and power.

The challenge is to increase the data-center's bandwidth, while minimizing cost and power.
To minimize the high-speed layout challenge for the designers, the ultimate solution is to use a 10GbE switch with 10Gbps serial-capable I/Os that can directly connect to the 10GbE optical modules with integrated signal-conditioning devices. This removes the extra power requirement and cost that comes from using extra components in the path.

This can be achieved using SFP+ modules based on Gennum's 10G CDR technology in combination with Fujitsu's 10GbE switch silicon. Using these technologies together will result in significant power savings and cost reductions in the form of lower component count, reduced cost manufacturing and a "pay-as-you-populate" capability.

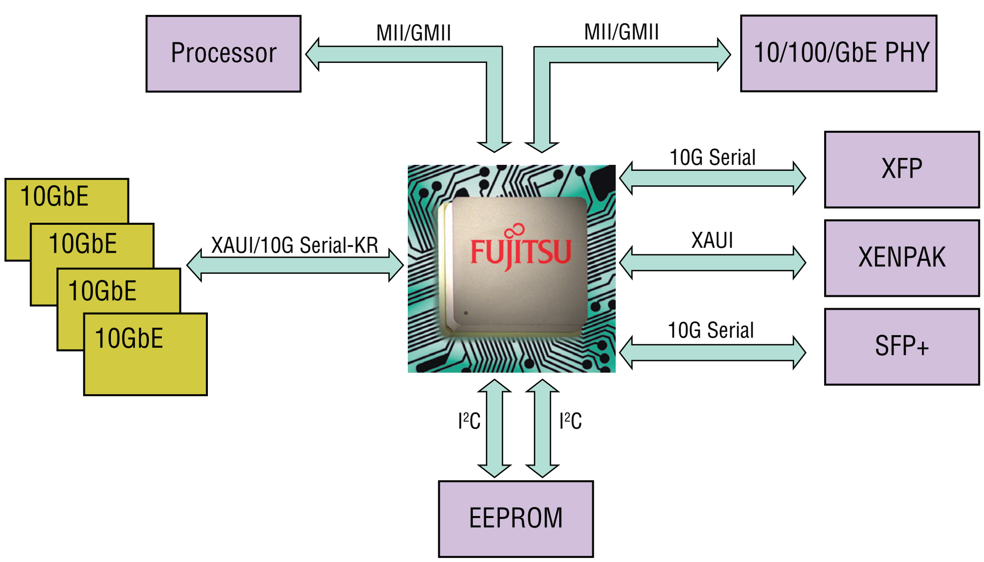

Figure 1 – The Fujitsu MB86C69RBC

The Transition to 10G Serial I/O and SFP+

Many older 10GbE systems (Figure 2) are based on X2/XPAK/XENPAK 10GbE modules connecting to the switch silicon with XAUI ports. In this case, each module contains a Serializer/Deserializer (SerDes) device and has a XAUI electrical interface. A XAUI interface uses four lanes, each running at 3.125Gb/s, instead of a single lane running at 10.3125Gb/s. Since each lane runs at a relatively low data rate of 3.125Gb/s, it is possible to route the signal directly to the switch silicon without additional signal conditioning. However, four lanes are required in each direction, instead of just the one that would be required for a serial connection, increasing the necessary board space. With the SerDes in the module, the power consumption is high, up to approximately 4W to 8W per module. Also, the module is large, which doesn't allow for high port-density configurations. Thus, this solution cannot effectively address the power and port density requirements of new data centers.

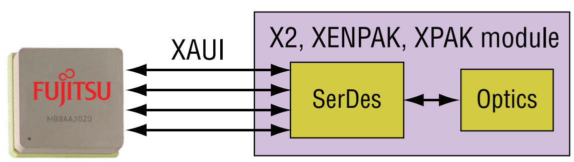

Figure 2 – XAUI Connection from X2/XENPAK/XPAK Module

Improvements can be made by upgrading the optical module type to SFP+ or XFP. The block diagrams for these module types are shown in Figure 3. The SerDes is no longer in the module, reducing its size and power. However, if the switch silicon uses a XAUI interface, a SerDes device is still required on the PCB to convert XAUI to 10Gbps serial. This device will consume power, increase the component count and require four lanes to be routed for each 10Gb/s signal.

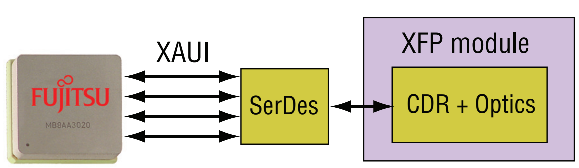

Figure 3 – XFP Module with SerDes connection to Switch Silicon

The next improvement can be made by removing the SerDes device altogether, through integration of the SerDes into the ASIC. This requires that the switch silicon be able to handle 10Gb/s serial input and outputs. The Fujitsu MB86C69RBC incorporates a new multi-rate PHY that supports the recently standardized 10GbE backplane (802.3ap). The PHY can drive 1m of single-channel backplane (also known as KR), XFI (the electrical interface for XFP modules) and SFP+'s that contain Gennum's CDR technology. The device supports KR on all 26 ports, and users can configure any of the ports to run in CX4, 1000BaseX, SGMII or KR mode. Therefore, the Fujitsu switch can connect directly to an XFP module, as shown in Figure 4.

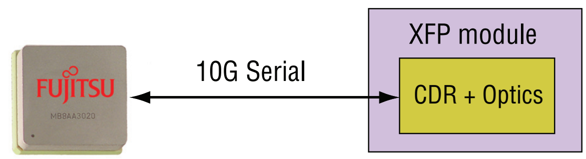

Figure 4 – The Fujitsu Switch Connected Directly to an XFP Module

The connection can be made directly to the XFP module because the signal-conditioning devices are integrated into the module. According to the XFP MSA, the input jitter at B', as shown in Figure 5, can be as high as 0.61UI and the output jitter at C' is guaranteed to be a maximum of 0.34UI. Therefore, the ASIC device can connect directly to the XFI module without any signal conditioning in between. This means that using the Fujitsu MB86C69RBC with XFP modules will result in an extremely low-component-count, cost-efficient solution. No SerDes devices are required, either in the module or on the PCB. The Fujitsu MB86C69RBC consumes only 22W for 26 ports or about 850mW/port, with the XFP module consuming ~1.5W/port for a combined solution of 2.3W per 10GbE port.

Although the XFP module interfacing directly to the Fujitsu MB86C69RBC is a cost-efficient, low-power solution, further reductions are possible with existing technology. SFP+ is the latest pluggable optical-module form factor for use in 10Gb/s Ethernet and 8.5Gb/s Fibre Channel systems. The objectives of this new form

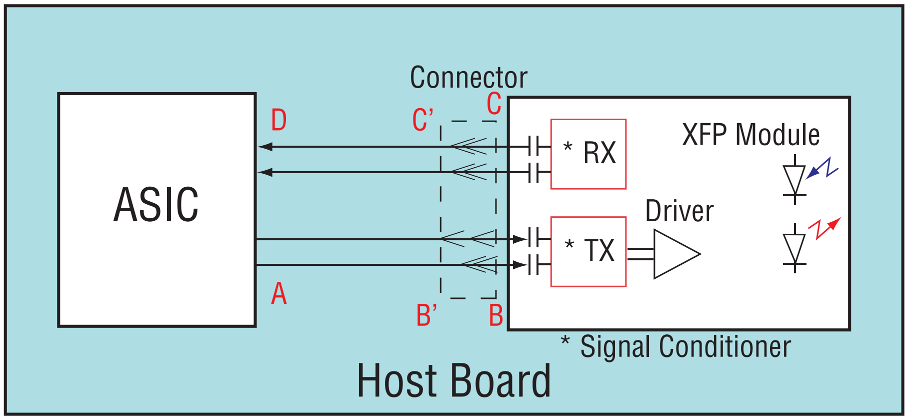

Figure 5 – XFP Module Connected to ASIC

factor are to increase port density through its smaller footprint, to provide a low-cost module to drive per-port costs down and to encourage increased adoption of the 10GbE technology. Semiconductor companies and module vendors have begun offering products for this space, which can be used in a number of potential system architectures.

A block diagram of an SFP+ module is shown below in Figure 6. Like an XFP module, the electrical interface of an SFP+ module runs at 10Gb/s. The main difference between an XFP module and SFP+ module is that signal-conditioning devices are optional in SFP+.

If a signal-conditioning device is not within the module, the jitter specifications for an SFP+ module are quite different from that of an XFP. The requirements are compared in Table 1 on the next page.

The decrease in the input jitter into the module and the increase in the output jitter from the module (plus additional jitter due to routing through various combinations of connectors, proximity to other signals, and the vias the signal needs to traverse) have a large impact on the ability to directly connect the module to the ASIC. Even with an ASIC that has advanced I/Os with powerful transmit and receive equalization (like Fujitsu's switch, which can function effectively when directly attached to SFP+ module), the jitter budget is tight.

The most common way to solve this signal-integrity issue is to use an Electronic Dispersion Compensation (EDC) device between the SFP+ module and the ASIC. EDC devices are available with either XAUI interfaces or XFI interfaces on the host side. EDC devices are sometimes required in the path for legacy LRM applications. Such applications and the requirement for EDC will be discussed in more detail next.

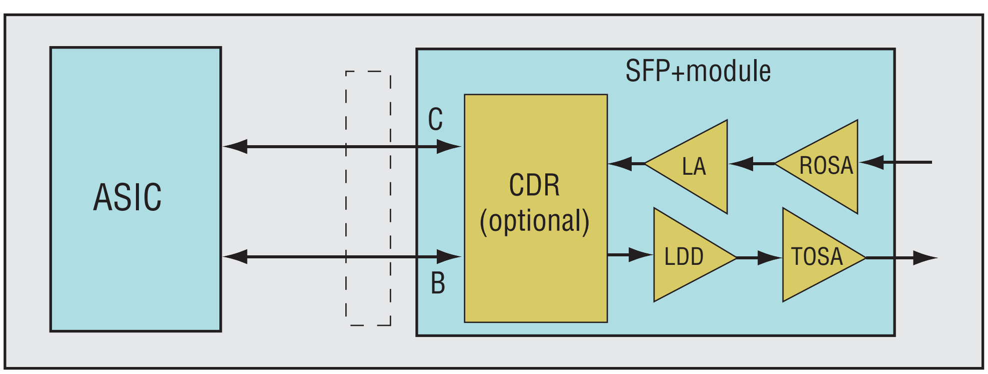

Figure 6 – SFP+ Module Block Diagram

	Input Jitter at B (UIpp)	Data Dependent Pulse Width Shrinkage at B (ps)	Output Jitter at C (UIpp)
XFP	0.61	N/A	0.34
SFP+	0.28	5.3	0.70

Table 1 – Jitter Specifications

When the dispersion-compensation functionality is not required, the EDC part is still used as a signal conditioner between the SFP+ module and the ASIC. However, this comes at a cost. A typical EDC device will consume more than 1W of power per port, resulting in a total power of at least 2.8W/port (850mW for switch silicon, 1W for EDC, and 1W for the SFP+ module). Using the EDC device also adds significantly to the BOM cost since it is an extra component on the host board and can require exotic PCB material. By using SFP+ modules and EDC devices on the line card, designers can take advantage of the high port density and low module power benefits of SFP+ modules. Designers can also use the 10G serial I/O of the Ethernet switch chip to reduce the number of signals that would be required for a XAUI-based solution. However, the component count is still high because an EDC device is required on every port; this can add substantially to the cost and power requirements.

The ideal solution is to be able to connect an SFP+ module directly to the switch ASIC. For this to be a robust solution, a switch ASIC with high-performance 10Gb/s I/Os and an SFP+ module with the signal-conditioning capability of an XFP module are necessary. Fujitsu has made the ASIC available. The SFP+ module with XFP-like signal-conditioning capabilities is also available by using an SFP+ module with integrated Clock and Data Recovery (CDR) devices, which guarantees physical layer performance. This can be achieved while still meeting the 1W power requirement of a class 1 SFP+ module.

The signal conditioning in an XFP module is achieved through the use of CDRs in both the transmit and receive directions. The purpose of a CDR is to accept a signal with a large amount of jitter and to regenerate this signal with much lower jitter. Each point in the system that uses a CDR resets the jitter budget.

Within the module, in the transmit direction, the CDR has high-input jitter tolerance, allowing the module to have high-input jitter tolerance. This means that the switch ASIC can output a signal that can accumulate jitter as it travels to the module, and the module will still be able to handle this signal since it is regenerated within the module. This approach eliminates the need for signal-conditioning components along the path. In the receive direction, without a CDR, a high jitter signal (i.e., 0.70UI or more) is transmitted from the module. This will accumulate jitter on the way to the ASIC, resulting in a signal that few, if any, ASICs can handle reliably, particularly after transmission over additional trace. With a CDR in the module, the ASIC can handle the high input jitter and output a clean signal that the ASIC can handle.

Since the XFP module form-factor is mature, there is no opportunity to reduce its size. On the other hand, with SFP+ there is an opportunity to use newer CDR technology that can fit into the SFP+ module and provide XFP-like performance.

No doubt that the SFP+ module is physically smaller than an XFP module, but there is a perception that CDRs cannot be made to fit in the module. Actually, CDRs can comfortably fit within the SFP+ when integrated with existing SFP+ module components, as shown in Gennum's SFP+ optical reference design, the RDK-10G-SFP+850. This module is based on Gennum's 10G CDR technology that is used in most of the XFP modules in operation today. This reference design also proves that CDR functionality can be used in both the transmit and receive direction with total power less than 1W, particularly when integrated with other module components such as the laser driver and limiting amplifier. An 8.5G version of the module for 8G Fibre Channel applications is also available.

A combination of the Fujitsu MB86C69RBC and SFP+ modules based on Gennum's reference design can achieve the ultimate data-center switch, with 10Gb/s SFP+ ports connected directly to the 10G switch silicon I/Os. It is low power, at <1.8W/port for both switch and module. It is cost effective, using the low-cost SFP+ form factor and no extra components on the PCB. It is robust, using proven 10G technology to maintain low BER performance. A comparison of the different solutions based on module and port type is shown in Table 2.

When analyzing the total component cost per port of each of the above solutions, it can be seen that the CDR-based SFP+, in conjunction with 10Gb/s serial ports, is both the lowest power solution and the lowest cost solution. By using low-cost SFP+ modules and removing the need for signal conditioning on the line card, a savings of $40 per port can be realized compared to the next most cost-efficient solution.

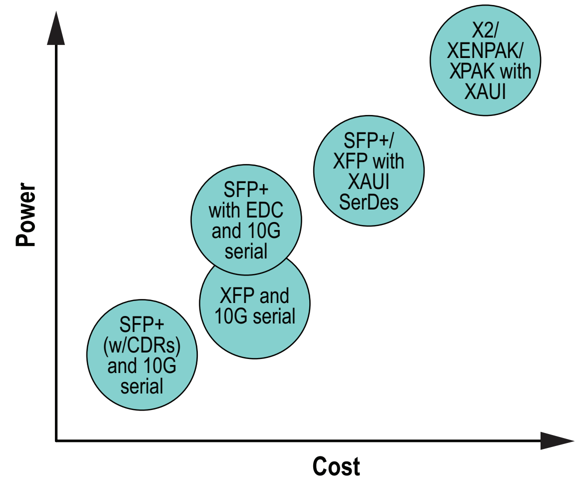

Figure 7 – Relative Cost and Power of 10GbE Switch Solutions

Supported Link Types

Two variants of SFP+ are currently being specified. The limiting variant, which closely resembles current SFP modules, includes a laser driver, a TOSA, a ROSA, and a limiting amplifier. The second variant, which is primarily intended for 10Gb/s Ethernet long-reach legacy multimode (LRM) applications, requires a linear optical receive path and is therefore referred to as the linear variant.

Optical Module	10GbE Switch Silicon Port Type	Line Card Silicon	Pros	Cons	Power/Port
X2/XENPAK/XPAK	XAUI	None	• Signal conditioning not required	• High power • Many signals to route	5W to 9W
SFP+/XFP	XAUI	SerDes	• Higher port density	• SerDes required on PCB	3W to 3.5W
XFP	10G	None	• Signal conditioning not required	• Port density not optimal	2.3W
SFP+ (non CDR)	10G	EDC	• High port density	• Signal conditioning required	2.8W
SFP+ (CDR)	10G	None	• High port density • Signal conditioning not required		1.8W

Table 2 – Data Switch Solution Options

The limiting variants are comprised of 10Gbase-SR and 10Gbase-LR links. 10Gbase-SR supports 300m over OM3 fiber and 10Gbase-LR supports 10km over single mode fiber. A CDR-based SFP+ design would classify as a superior performing limiting module.

The linear variant is necessary to support legacy multi-mode fiber. For short-reach applications a direct connection between the ASIC (such as the Fujitsu chip) and the SR version of the SFP can achieved. However, for longer reaches of the installed MMF fiber, the optical signal will experience a significant amount of dispersion. This dispersion limits the maximum length of fiber that can be used and is proportional to the data rate of the signal. Therefore, the transition from 1GbE to 10GbE would have meant that only short lengths of legacy fiber could carry a 10GbE signal. The solution is to use Electronic Dispersion Compensation (EDC), enabling transmission over 220m of legacy FDDI fiber. However, as new data centers continue to be deployed and infrastructure upgrades occur, the small percentage of long FDDI links will continue to diminish. This is an opportunity for system optimization, where network equipment users don't pay for EDC (in power and cost) for links that do not require it. This optimized case is depicted in Figure 8.

According to Lightcounting's Q1-2008 Transceiver Market Survey Result, almost 30 limiting modules were sold for every one linear module in 2007. Therefore, more than 95% of the time, the EDC device was not required. The EDC device, along with system design implications, adds significantly to the system cost ($40 per port) and power (greater than 1W). The optimized EDC-less architecture also enables a "pay-as-you-populate" model. For the unused links, no extra silicon is populated on the line card, so there is no unnecessary cost. When a module is required, one with signal-conditioning capabilities can interface with the switch ASIC to create a robust link, with no unneeded components in the path.

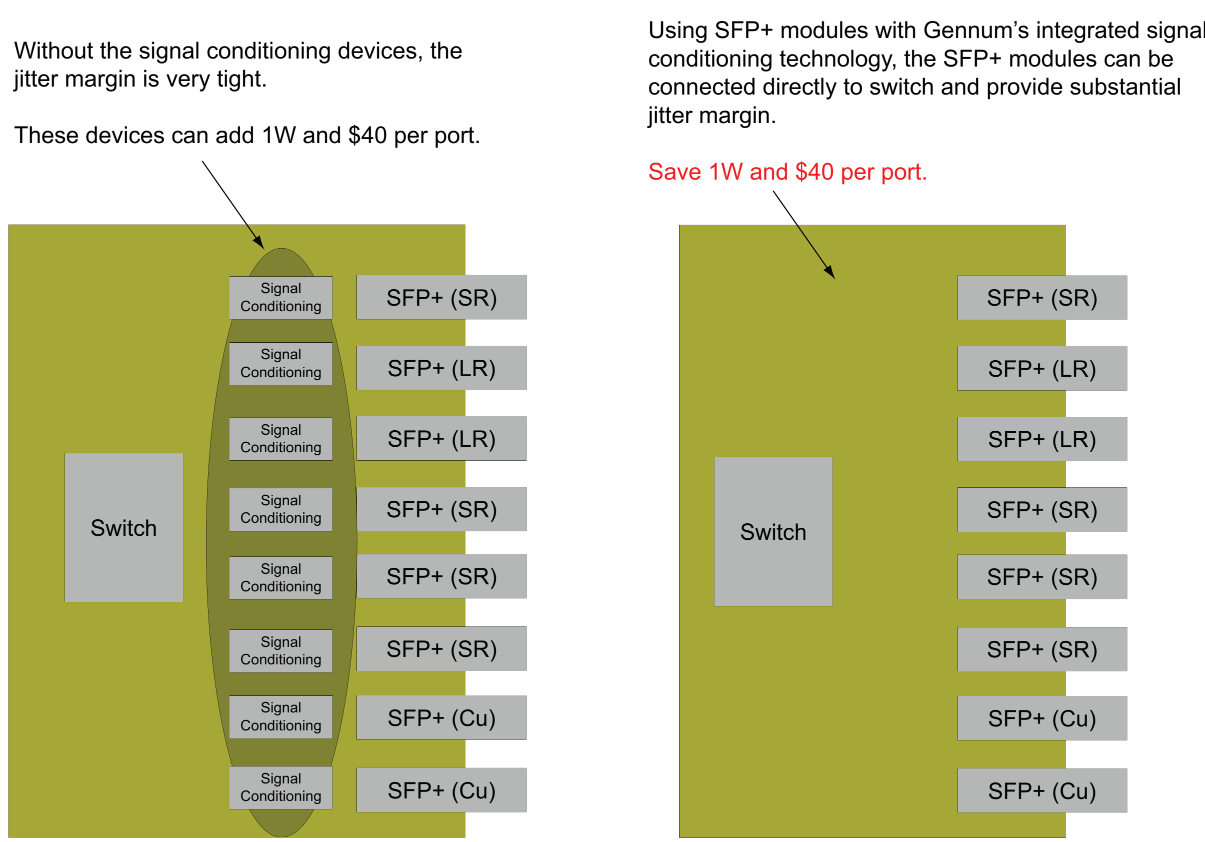

Figure 8 – Typical SFP+ Architecture with EDC on Every Port

Copper

In addition to optical links, copper links are common in the data center. For connections of less than 15m, it may be more cost-effective to drive the signal electrically over a twinax cable than optically over fiber. The "pay-as-you-populate" model applies to these copper links also. An SFP+ active copper cable assembly that uses CDRs as signal conditioners within the SFP+ housing can be used. A CDR-based reference design kit, the RDK-SFP+CU, is available from Gennum. This is shown in Figure 9.

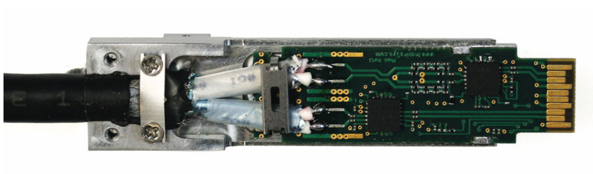

Figure 9 – SFP+ Active Copper Cable Assembly using CDRs

As with the SFP+ optical module using CDRs, this type of copper cable does not require any components on the line card to connect directly to the 10Gb/s I/Os on the Fujitsu switch ASIC. Again, this combination results in the lowest cost, lowest power solution, and enables longer reaches over the cable than otherwise possible. An SFP+ active copper cable assembly reference design is currently available from Gennum Corp.

For more information on Gennum's CDR-based SFP+ reference design kits, visit www.Gennum.com/CDR-RDK.

Summary

Fueled by data-center performance demands and declining price per port, the 10GbE switch market is poised for rapid growth in the next few years. During 2006 and 2007, developers of integrated silicon solutions made tremendous progress in refining 10GbE technology and its infrastructure. By reducing the cost per port, 10GbE based on SFP+ optical transceivers are enabling many new applications using 10GbE NICs and switches for low-cost interconnects.

On top of technology improvements and cost reductions, rising energy costs and an increased global focus on environmental impact mean that power reduction is crucial for all data centers.

Gennum and Fujitsu have combined to develop 10GbE technology that is superior to previous solutions in terms of both cost and power, and lowers the per-port cost of the solution. Fujitsu's 26-port, 10GbE switch IC, the MB86C69RB, consumes only 22W with all 26 ports fully loaded. It has 10Gb/s serial I/Os with transmit-and-receive equalization. When combined with SFP+ modules based on Gennum's 10G CDR technology, the switch IC allows for direct connection to the module with enhanced SFP+ performance. This reduces the component count and saves 1/3 the power and $40 per port, resulting in the ultimate data-center solution.

FUJITSU MICROELECTRONICS AMERICA, INC.
Corporate Headquarters
1250 East Arques Avenue, M/S 333, Sunnyvale, California 94085-5401
Tel: (800) 866-8608 Fax: (408) 737-5999
E-mail: inquiry@fma.fujitsu.com Web Site: http://us.fujitsu.com/micro

Printed in U.S.A. 10GE-CS-21336-10/2008