SIMPLIFYING NETWORK CONNECTIVITY WITH REMOTEPHY TECHNOLOGY IN DELL BLADE SERVERS

By David Christensen

The new Broadcom® RemotePHY™ technology offered in Dell™ PowerEdge™ M600 and PowerEdge M605 server blades helps simplify the integration of blade server enclosures into data centers wired with 10/100 Mbps copper Ethernet.

Related Categories:

Broadcom

Blade servers

Dell PowerEdge blade servers

Ethernet

Visit DELL.COM/PowerSolutions for the complete category index.

The introduction of blade servers, combined with the amplifying effects of server virtualization, has allowed enterprise data centers to achieve extremely high-density computing. Even as computing power has increased, system cabling has been simplified, and the number of failure points within the system has decreased, the design of blade servers has still required some trade-offs that affect how enclosures are deployed within existing network infrastructures, especially when legacy networking equipment is involved. This article discusses why such situations exist today and how enterprises can use Broadcom RemotePHY technology to integrate the Dell PowerEdge M1000e modular blade server enclosure and PowerEdge M600 and PowerEdge M605 server blades into their data centers to help achieve optimal efficency.[1]

Among the many benefits of blade computing, one strong motivation for moving to a Dell modular blade enclosure is the ability to reduce cable sprawl. Common computer system connections including power, KVM (keyboard, video, mouse), and networking are routed over the backplane in the enclosure, connecting each blade to the outside world. Networking is routed over the backplane using the IEEE 802.3-2005 1000Base-X protocol, which is typically found in fiber-optic-based network controllers but is also well suited to backplane implementations without using the optical interconnects. Each server blade has two or more such network connections to a modular I/O bay, which typically contains an Ethernet switch that allows the blades to communicate with each other and with the rest of the enterprise network.

UNDERSTANDING PASS-THROUGH MODULES

While a network connection from the server blade to an Ethernet switch is suitable in many situations, some organizations must directly connect a particular server blade to an external network device such as a router that provides Internet access, or to a network security appliance. In these situations, an Ethernet pass-through module is used to perform a media-conversion function from the 1000Base-X signaling used over the backplane to a 1000Base-T physical layer (PHY) transceiver built into the pass-through module. As far as the blade server is concerned, it is directly connected to the external device at a 1,000 Mbps link speed.

[1] For more information about PowerEdge M-Series blade servers, see "The Next-Generation Dell PowerEdge M1000e Modular Blade Enclosure," by Chad Fenner, in *Dell Power Solutions*, February 2008, DELL.COM/Downloads/Global/Power/ps1q08-20080206-Fenner.pdf; and "Exploring the Dell PowerEdge M1000e Network Fabric Architecture," by John Loffink, in *Dell Power Solutions*, February 2008, DELL.COM/Downloads/Global/Power/ps1q08-20070500-Loffink.pdf.

Figure 1 shows a simplified representation of the Dell PowerEdge M1000e modular blade enclosure with a Dell Ethernet pass-through module installed. Because two different link connections are involved with a pass-through module, the following terminology is used in the remainder of this article:

- **Media Access Control (MAC):** The Broadcom BCM5708S Gigabit Ethernet controller used on the blade server
- **Serializer/deserializer (SerDes) link:** The internal 1000Base-X link between the blade and the pass-through module; typical Gigabit Ethernet adapters use a SerDes to connect the MAC with the fiber-optic module, but when using a blade architecture, the fiber-optic module on both sides of the Ethernet link can be eliminated, allowing the SerDes interface for both Ethernet adapters to be directly connected
- **Copper link:** The external 1000Base-T link between the pass-through module and the external network infrastructure
- **Copper PHY:** The Broadcom BCM54981 Gigabit Ethernet transceiver built into the pass-through module

Previously, when an Ethernet pass-through module was used in a blade enclosure, the link speed available through the external copper port exactly matched the link speed of the enclosure backplane. Because the MAC operated at 1,000 Mbps, the copper PHY was forced to operate at 1,000 Mbps as well. This fixed link speed increased the difficulty of integrating blade servers into existing data centers that used 10 Mbps or 100 Mbps network infrastructures. External Ethernet switches could be used to connect the 1,000 Mbps copper PHY to the lower-speed network, but this work-around typically added expense to the overall solution and increased cable sprawl rather than reducing it.

"Broadcom has developed RemotePHY technology to help avoid the need for additional switches and allow individual server blades to connect to external network devices at speeds other than 1,000 Mbps."

OVERCOMING HISTORICAL LIMITATIONS WITH BROADCOM REMOTEPHY TECHNOLOGY

Broadcom has developed RemotePHY technology to help avoid the need for additional switches and allow individual server blades to connect to external network devices at speeds other than 1,000 Mbps. The RemotePHY feature is currently supported on the Broadcom BCM5708S Gigabit Ethernet controller (available in Dell PowerEdge M600 and PowerEdge M605 server blades) and on the Broadcom BCM54981 Gigabit Ethernet transceiver (available in the Dell Ethernet pass-through module). The Broadcom networking drivers for Linux® and Microsoft® Windows Server® 2003 operating systems used on these Dell server blades are RemotePHY aware and can automatically detect and enable support for RemotePHY when the Dell Ethernet pass-through module is detected.

RemotePHY functionality allows the blade server to communicate directly with the copper PHY in the pass-through module using an in-band signaling mechanism supported by the IEEE 802.3-2005 Ethernet specification. This communications channel allows blades to configure the copper PHY as directed by the administrator, including setting auto-negotiation parameters or forcing the link speed to a specific value—the same functionality typically provided on a traditional rack or tower server. Once a copper link has been established to an external network device, the RemotePHY feature also provides a mechanism for reducing the fixed 1,000 Mbps SerDes signaling rate to a 10 Mbps or 100 Mbps data rate on the copper link.

The end result is that the pass-through module operates as if the copper PHY were directly connected to the blade server, allowing the blade to connect directly not only to Gigabit Ethernet devices, but also to legacy 10 Mbps or 100 Mbps network equipment as well—thereby helping eliminate the need for an external Ethernet switch as intermediary.

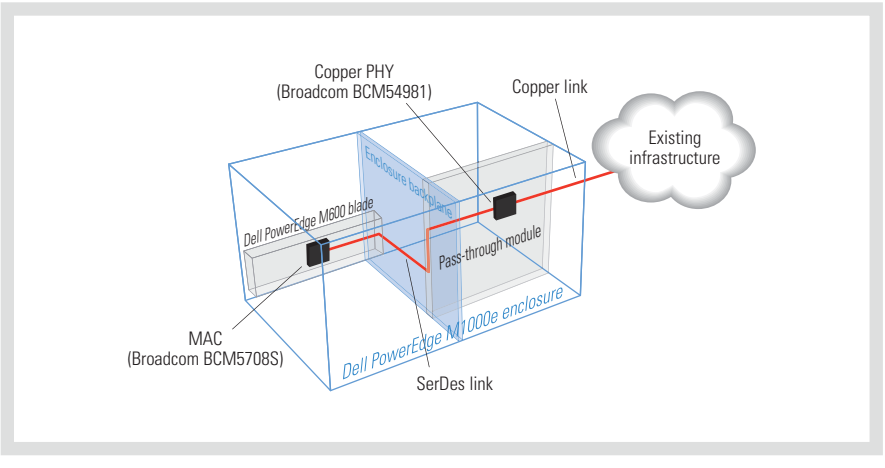

Figure 1. *Dell PowerEdge M-Series blade server architecture*

> **"RemotePHY functionality allows the blade server to communicate directly with the copper PHY in the pass-through module using an in-band signaling mechanism supported by the IEEE 802.3-2005 Ethernet specification."**

EXPLORING HOW REMOTEPHY TECHNOLOGY WORKS

RemotePHY technology is built on two standard components of the IEEE 802.3-2005 1000Base-X specification: an electrical signaling specification and a standards-compliant extension to the Ethernet auto-negotiation protocol. The first allows a 10 Mbps or 100 Mbps data rate to be transmitted over a fixed 1,000 Mbps connection, while the second provides a message-passing mechanism between the Ethernet MAC and the remote copper PHY over the blade enclosure backplane.

Electrical signaling

RemotePHY technology continues to use the IEEE 802.3-2005 1000Base-X specification, which operates at 1,000 Mbps over the blade enclosure backplane. To accommodate the 10 Mbps or 100 Mbps connections that may occur on the copper link, the data transferred over the backplane is expanded or elongated by a factor of 100 or 10, respectively, allowing a slower data rate to be transmitted over a faster data path.

Auto-negotiation

The extensions to the IEEE 802.3-2005 auto-negotiation protocol are more complicated than the use of electrical signaling. For 1000Base-X devices, auto-negotiation is mandatory. Each side of the link sends its link partner a base page (BP) that is encoded as a 16-bit word that indicates the features supported by the device: half- or full-duplex operation, symmetric or asymmetric pause frame generation, remote fault indication, and an acknowledgment bit. The most significant bit of the BP is the next page (NP) bit, which indicates whether additional 16-bit pages are available.

Note: Although the current 1000Base-X auto-negotiation process does not negotiate the link speed, Broadcom has released an extension known as the Multirate Backplane Ethernet (MRBE) specification, which allows link speed to be negotiated between supported devices such as the BCM5708S. The RemotePHY feature is designed to be fully compatible with MRBE.

When two RemotePHY-compatible devices are engaged in auto-negotiation, the NP bit of the BP is set, allowing additional information to be exchanged between the two PHYs in the form of message pages (MPs) and unformatted pages (UPs). The RemotePHY specification defines the format for these MP and UP exchanges, which start with an exchange of message page 5 (MP-5), which is defined in the 802.3 specification as an organizationally unique identifier (OUI) tag code. The presence of the Broadcom OUI, along with additional mode bits that are sent in subsequent UPs, allows the link partners to negotiate which of these RemotePHY services are supported:

- The MAC can control the copper PHY auto-negotiation settings for the copper link.
- The copper PHY can report the results of the copper link auto-negotiation to the MAC.
- The MAC can send a register write request to the copper PHY.
- The MAC can request a register read from the copper PHY.
- The copper PHY can send the results of a register read to the MAC.

Through a combination of these services, the MAC can configure the copper PHY as if it were locally connected to the MAC, including setting auto-negotiation parameters, forcing the link speed, and setting half- or full-duplex operation.

STANDARDIZING SIMPLIFIED NETWORK CONNECTIVITY

To promote adoption of RemotePHY technology, Broadcom has published the combined MRBE and RemotePHY specification on its Web site at www.broadcom.com/MRBE. Broadcom's goal is to simplify IT operation through application of industry standards to blade server environments.

David Christensen is an applications engineering manager for the Broadcom Ethernet controller business unit. David has a B.S. degree in Computer Science from California State University, Fullerton, and has worked at a variety of software engineering and product support roles at AST Computer, Phoenix Technologies, and Broadcom over his 17-year career in the PC industry.

MORE ONLINE
DELL.COM/PowerSolutions

QUICK LINKS

Broadcom Multirate Backplane Ethernet and RemotePHY specification:
www.broadcom.com/MRBE

Dell PowerEdge blade servers:
DELL.COM/Blades