




Enabling a Converged World™

# **An overview of Next-Generation 100 and 40 Gigabit Ethernet Technologies**



# **An Overview of** Next-Generation 100 and 40 Gigabit Ethernet Technologies





# Contents

- Abstract ..... 2**
- The Need for Speed ..... 2**
  - Different Uses – Different Speeds ..... 4
  - The Costly Alternative ..... 4
- Recent Developments and Events..... 5**
- Higher Speed Technology ..... 6**
  - Use of Existing Fiber ..... 6
  - Not reinventing the wheel ..... 7
- Efforts to bring 100 GbE to market in the near-term ..... 8**



**Backbone Internet bandwidth maintained by the major Internet carriers has been expanding at 75 to 125% per year, driven by expansion in the number of broadband Internet users and growing use of bandwidth-hungry applications.**

## Abstract

Advances in data center technology and Internet usage have pushed enterprises to 10 Gigabit Ethernet (Gbps) links and search engines, carriers, ISPs to multiple 10 Gbps links. Projected growth mandates the need for higher speed Internet connections. An IEEE Task Force has been established with the objective of standardizing 40 Gbps within the data center and 100 Gbps between major Internet nodes. If higher speed Ethernet is to be a success it must take advantage of existing data center copper and longer-reach fiber cables. The first implementation of the adopted PCS layer for PCS lane distribution was demonstrated by Ixia in mid 2008.

## The Need for Speed

The amount of backbone Internet bandwidth maintained by the major Internet carriers has been expanding at 75 to 125% per year<sup>1</sup>, driven by the explosion of broadband Internet users and growing use of bandwidth-hungry applications. Video-centric applications are the main contributors, including high-definition IPTV, video distributors such as YouTube™, and video conferencing services. Other significant bandwidth gobblers include peer-to-peer sharing, virtual private networks, wireless backhaul, gaming, and various forms of online collaboration. Figure 1<sup>2</sup> illustrates the changing nature of Internet usage from 2006 to 2012 (estimated).

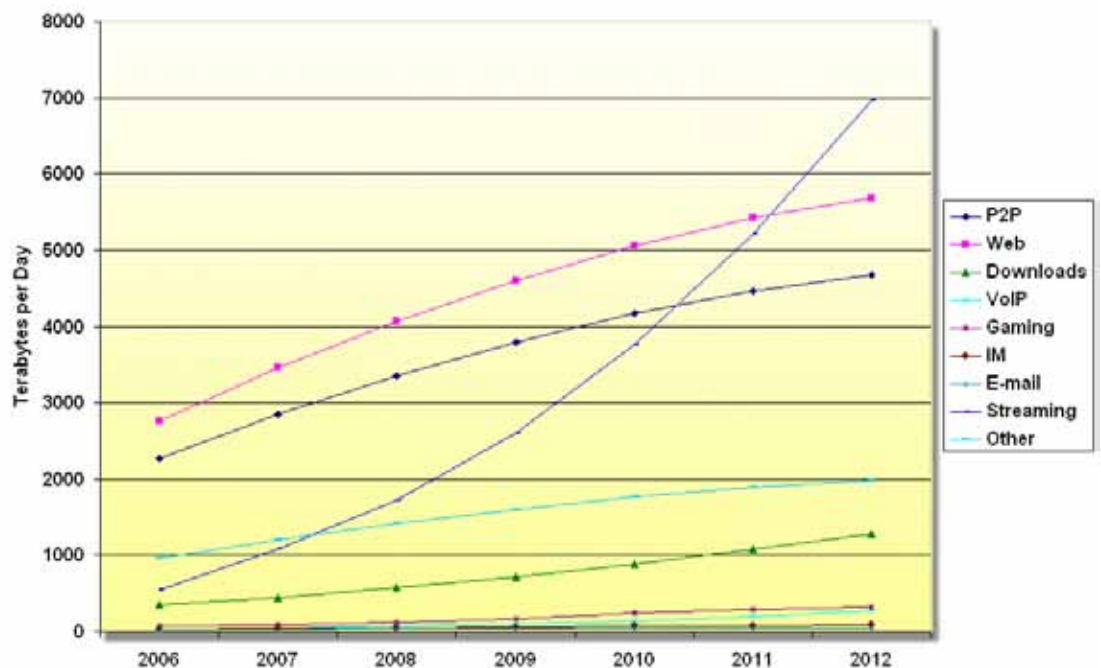


Figure 1. Changing Internet Usage

<sup>1</sup> G. Coffman and A. M. Odlysko, "Growth of the Internet," Optical Fiber Telecommunications IV B: Systems and Impairments, I. P. Kaminow and T. Li, eds. Academic Press, 2002, pp. 17-56.

<sup>2</sup> IDC Report on U.S. Consumer Internet 2007-2011 Forecast, June 2007

Advances in computer and networking technologies have likewise led to an increased demand for bandwidth within the data center. Multicore processing and virtualization have expanded the capacity and utility of individual servers.

Network attached storage (NAS) is rapidly replacing local disks, allowing access to larger amounts of data, centralized maintenance, and improved disk usage. Reduced costs are realized by consolidating servers and by employing NAS, as well as through the reduction of the number of network interfaces required both on servers and on network switches. With this consolidation, however, comes higher network data volume on a per-server basis. Most data centers have switched from Gigabit to 10 Gbps technologies to meet their bandwidth needs.

A growing number of the largest Internet services – Google, Yahoo, MySpace, Facebook, AOL, for example – have a near-term expected need for 40 Gbps in the data center and 100 Gbps between data centers, as illustrated in Figure 2 (courtesy of Google).

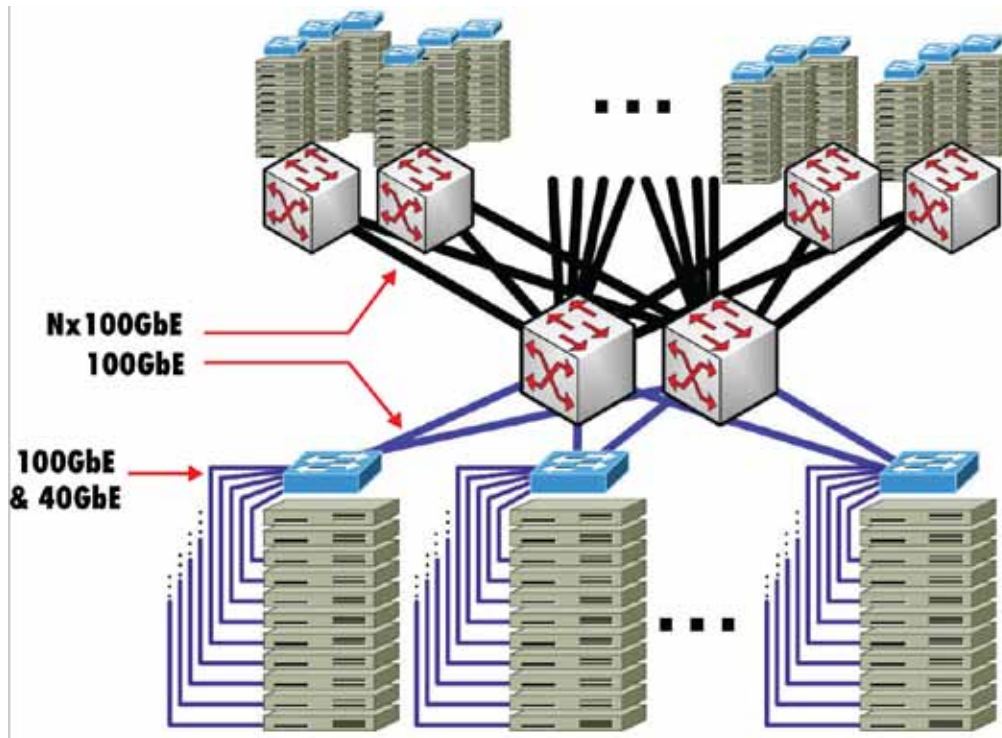


Figure 2. Data Center Network Aggregation

With the enterprise migration from frame relay and other private networks to IPSec and layer 2-3 VPNs, ISPs have found it necessary to offer 10 Gbps customer connections. ISPs traditionally require 4 to 10 times the bandwidth of their largest customer connection in order to provide the proper levels of service.

**ISPs traditionally require 4 to 10 times the bandwidth of their largest customer connection in order to provide the proper levels of service.**

## Different Uses – Different Speeds

Data center network usage is driven by Moore’s law – doubling every 24 months. While 10 Gbps appears to be sufficient for most data centers today, 40 Gbps is expected to meet data center needs through 2014.

Aggregated Internet traffic, however, is related to general Internet and telecom growth – doubling every 12 to 18 months. A higher bandwidth solution is required. According to the Ethernet Alliance<sup>3</sup>:

“The increasing trend of Ethernet being an end-to-end network solution and the understanding that a 40 Gbps interface will not be sufficient to address the aggregation demand has resulted in the 100 Gbps rate being identified as the next high-speed network operator interface. The core network needs to evolve to a higher rate, and operators are looking to Ethernet as the new ‘fat pipe’ for the core of their backbone; positioning 100 Gbps to complete the transition to an end-to-end all-Ethernet infrastructure.”

Based on the current rate of backbone bandwidth expansion, 100 Gbps should meet the needs of carriers and ISPs through 2015.

## The Costly Alternative

Lacking alternatives to 10 Gbps connections, carriers and enterprises have resorted to using multiple 10 Gbps connections to satisfy their aggregated bandwidth requirements. Each new 10 Gbps bandwidth step comes with additional switch and/or computer interfaces – both of which are expensive affairs.

The complexity of such shear numbers can also be daunting, as shown in the metro network depicted in Figure 3. The benefits from a 100 Gbps network infrastructure are immediately visible.

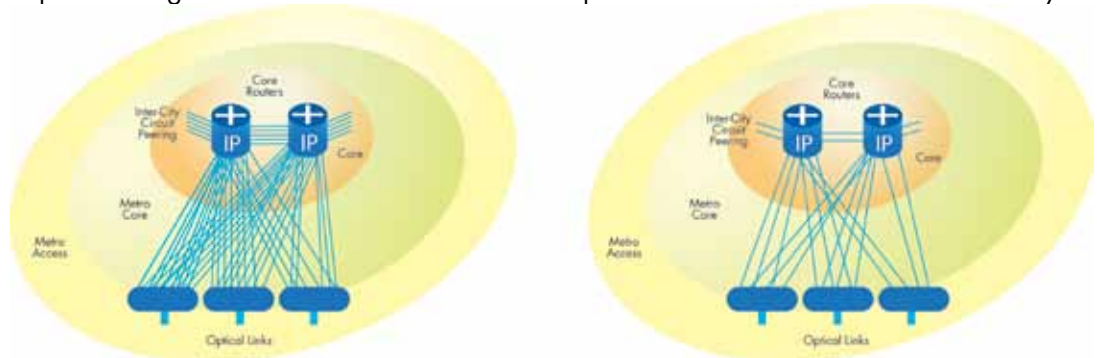



Figure 3. Network Complexity and Benefits of 100 Gbps Infrastructure

<sup>3</sup> Mark Nowell, Vijay Vusirikala, Robert Hays, “Overview of Requirements and Applications for 40 Gigabit and 100 Gigabit Ethernet,” Version 1.0, August 2007.

**Each new 10 Gbps bandwidth step comes with additional switch and/or computer interfaces – both of which are expensive affairs.**



There is a further complexity associated with splitting up large aggregated flows that result from server consolidation into 10 Gbps links. Bundling multiple links is usually handled through the use of link aggregation groups (LAGs), as specified in IEEE standard 802.3ad. Shorter data flows, such as VoIP packets, lend themselves to easy distribution across multiple links and result in efficient packing due to their small size and even flow rate. The single data flows associated with VPNs, streaming, and long-term data storage protocols are bursty in nature and cannot be split across multiple connections. Such data flows can generate bursts of up to 5 Gbps in traffic rate, resulting in high latency. ISPs often reserve bandwidth to handle maximum traffic rates without SLA penalties – resulting in low bandwidth usage. Misallocation of bandwidth can lead to dropped packets and low quality of real-time applications such as video transmissions.

Larger 40 Gbps and 100 Gbps pipes will serve to average out multiple bursty flows, resulting in higher bandwidth efficiency and lower per-Gbps costs.

## Recent Developments and Events


In July of 2006, the IEEE established the high speed study group (HSSG) to look into the three- to seven-year requirements for Ethernet speeds beyond 10 Gbps. The HSSG reached the conclusion that both 40 and 100 Gbps solutions should be pursued, and forwarded a project authorization request (PAR) for a higher speed Ethernet amendment to the IEEE 802.3 standard.

With approved objectives and PAR, the IEEE 802.3ba Task Force was formed in December of 2007. The task force has been meeting regularly to discuss technology proposals that are nearing finalization (expected in early 2010).

The approved task force objectives were:

- Exclusive support full-duplex operation
- Preserve the 802.3 Ethernet frame format using the 802.3 MAC
- Preserve the current 802.3 standard minimum and maximum frame sizes
- Support bit error rates better than or equal to 10<sup>-12</sup> at the MAC/PLS service interface
- Provide appropriate support for optical transport networks (OTN)
- Support a MAC data rate of 100 Gbps, with physical layer specifications that support operation over:
  - ✧ At least 40 km on SMF fiber
  - ✧ At least 10 km on SMF fiber
  - ✧ At least 100 m on OM3 MMF
  - ✧ At least 10 m over a copper cable assemblies

**The High Speed Study Group reached the conclusions that both 100 GbE and 40 GbE solutions should be pursued.**



## In May 2009, Ixia demonstrated the world's first 100 Gb/s test module with a CFP interface.

- Support a MAC data rate of 40 Gbps, with physical layer specifications that support operation over:
  - ✧ At least 10 km on single-mode fiber (SMF) fiber
  - ✧ At least 100 m on OM3 multi-mode (MMF) fiber
  - ✧ At least 10 m over a copper cable assemblies
  - ✧ At least 1 m over a backplane

In June 2008, at the NXTcomm'08 trade show, Ixia demonstrated the operation and testing of a 100 Gbps link based on the current state of their standardization efforts. In October 2008, Ixia unveiled a 100 Gbps Development Accelerator System for sale and demonstrated a 40 Gbps proof-of-concept system. In May 2009, Ixia demonstrated the world's first 100 Gbps test module with a CFP interface.

## Higher Speed Technology

### Use of Existing Fiber

One of the task force's recommendations was to use existing fiber connectors. In particular, the most common fiber cable types should be accommodated:

- 100 m multi-mode fiber (MMF) using a 850 nm wavelength,
- 10 km single-mode fiber (SMF) using 1310 and 1550 wavelengths, and
- 40 km single-mode fiber (SMF) using 1310 and 1550 wavelengths

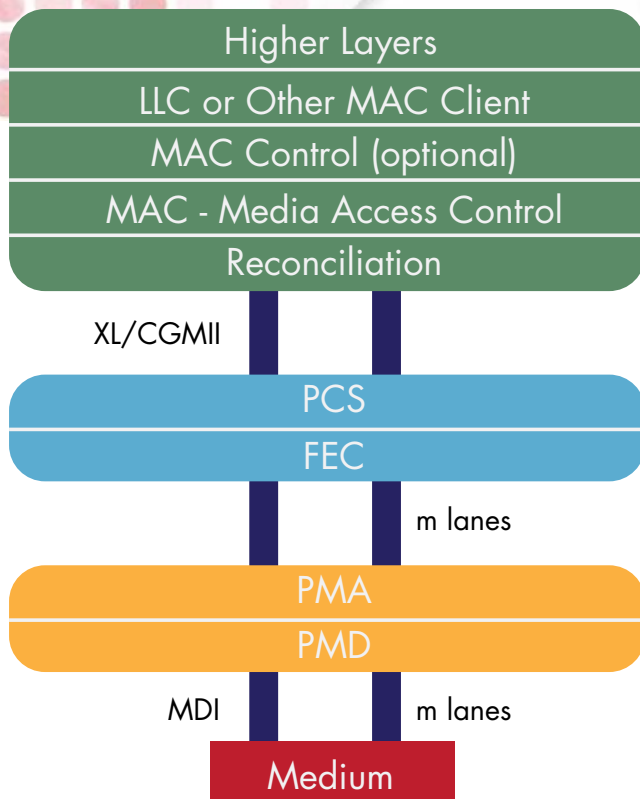
Advances in wavelength-division multiplexing (WDM) and dense-WDM (DWDM) enabled the use of multiple simultaneous wavelengths on individual fibers, each capable of 10 Gbps traffic. Today's optical transport networks uses DWDM to implement ITU-T recommendation G.709, which defines ODU4 for use with 100 Gbps and ODU3 with 40 Gbps.

Work is progressing toward higher bandwidth per wavelength and fiber. Industry pundits see near-term bundles of various combinations that will handle 100 Gbps information flows, including:

- 10 wavelengths of 10 Gbps
- 4 wavelengths of 25 Gbps
- 5 wavelengths of 20 Gbps

### Not reinventing the wheel

The 802.3ba task force's guidelines include leaving the majority of 802.3 standards in place. Figure 4 is a typical Ethernet architecture as it applies to 40 and 100 Gbps.



**The 802.3ba task force's guidelines include leaving the majority of 802.3 standards in place.**

Figure 4. 40/100 Gbps Ethernet Architecture

No changes will be made to the medium access control (MAC) layer beyond increasing the data rates to 40 and 100 Gbps. This means that other characteristics, including frame format, minimum/maximum frame sizes, and full-duplex only operation remain the same.

The reconciliation sublayer (RS) maps the MAC's serial bitstream into its respective 40 and 100 Gbps media independent interfaces. These interfaces are being developed based on the earlier XGMII (10 Gbps) interface: CGMII (100 Gbps) and XLGMII (40 Gbps). These interfaces in the architecture along with the components that use these interfaces are being standardized through multiple multi-source agreements (MSAs) between industry component vendors, such as the CFP MSA (see the transceiver section below).

The remaining sublayers serve to serialize the data received through the MII and present it to the copper or optical medium. The physical coding sublayer (PCS) provides coding, including scrambling and control block encoding. 64/66b coding, as used in 10GBASE-R, is the choice for these higher speeds. The optional FEC sublayer performs encoding for forward error correction. The physical medium attachment (PMA) sublayer serializes the coded data as required by the medium. The physical medium dependent (PMD) sublayer is responsible for signaling of the serial stream to the medium through the media dependent interface (MDI).



## Transceiver Options

As the IEEE moves forward in finalizing the 40 Gbps and 100 Gbps standard P802.3ba, many wonder which transceiver type will win out as the final industry higher speed Ethernet connector. With the recent formation of a new multi-source agreement (MSA) between several of the top transceiver manufacturers (Avago, Finisar, Opnext, and Sumitomo), it appears that hot-pluggable optical CFP transceivers are the clear front runner.

Both the OIF and the ITU-T are working on standardizing SDH/OTN telecom interfaces for long-haul transmission of 100 Gbps. The CFP MSA defines a hot-pluggable optical transceiver form factor to enable 40 and 100 Gbps applications. Pluggable CFP transceivers will support the ultra-high bandwidth requirements of data communications and telecommunication networks that form the backbone of the internet.

Pluggable transceiver modules compliant to the CFP MSA are an option for 40 and 100 Gbps interfaces. The CFP MSA is defining the specifications required to support multiple applications using the same form factor. These applications include various protocols (such as 40 Gbps, 100 Gbps, OC-768/STM-256, OTU3), media types (multimode and single mode fiber optics), and link distances. The CFP MSA uses numerous innovative features such as advanced thermal management, EMI management, and 10 Gbps signal integrity design to define the transceiver mechanical form factor, the optical connector, the 10x10Gbps electrical connector with its pin assignments, the MDIO-based transceiver management interface, and the hardware required on the system host board.

Another niche option being developed for shorter-length connections is quad small form-factor pluggable (QSFP) transceivers. The QSFP MSA defines the connector as a 4-channel QSFP bi-directional electrical-to-electrical connectivity over distances up to 100 meters. Each independent channel is capable of 5, 8 or 10 Gbps, for an aggregate bi-directional bandwidth of up to 40 Gbps.

QSPF transceivers are multi-mode fiber, short-reach only transceivers designed specifically for 40 Gbps devices. Their application is mainly for links within the data center, and are significantly less expensive than their CFP counterparts.

## Efforts to bring 40 and 100 Gbps to market in the near-term

While the 802.3ba task force hones in on the final standard, carriers and large information service organizations are expressing a dire need to begin testing deployments of 40 and 100 Gbps services. A number of proposals were offered, but PCS lane distribution (first described by Mark

**Pluggable transceiver modules compliant to the CFP MSA are an option for 40 and 100G Gbps interfaces.**



Gustlin at Cisco) was approved by the 802.3ba Task Force as the baseline PCS proposal on May 16, 2008.

Optical fiber module (PHY) vendors now have the standard of 20 PCS lanes for 100 Gbps and 4 PCS lanes for 40 Gbps. With this standard, development on 40 and 100 Gbps networking systems can begin in earnest. Such development will require intense and accurate testing solutions.

In the market's first public demonstration, Ixia tested a 100 Gbps link using PCS lanes and full line-rate traffic. Ixia, in conjunction with Hitachi, also demonstrated a real-world testing scenario in August of 2009 based on Hitachi's 100 Gbps frame transmission system. This system was partially supported by the Lambda Access Project at the National Institute of Information and Communications Technology (NICT) in Japan.

PCS lanes are a complicated and interesting subject, and are the subject of another Ixia-sponsored white paper, which can be found at <http://www.ixiacom.com/ixialabs>.

**Pluggable transceiver modules compliant to the CFP MSA are an option for 40 and 100G Gbps interfaces.**



**Ixia Worldwide  
Headquarters**

26601 Agoura Rd.  
Calabasas, CA 91302  
(Toll Free North America)  
1.877.367.4942  
(Outside North America)  
+1.818.871.1800  
(Fax) 818.871.1805  
[www.ixiacom.com](http://www.ixiacom.com)

**Other Ixia Contacts**

Info: [info@ixiacom.com](mailto:info@ixiacom.com)  
Investors: [ir@ixiacom.com](mailto:ir@ixiacom.com)  
Public Relations: [pr@ixiacom.com](mailto:pr@ixiacom.com)  
Renewals: [renewals@ixiacom.com](mailto:renewals@ixiacom.com)  
Sales: [sales@ixiacom.com](mailto:sales@ixiacom.com)  
Support: [support@ixiacom.com](mailto:support@ixiacom.com)  
Training: [training@ixiacom.com](mailto:training@ixiacom.com)