Toward Higher-Speed Modular Interconnection Compatibility

Contributors:

Greg McSorley, Amphenol

John D’Ambrosia, Dell

Executive Summary

Driven by the ever-increasing server speeds and storage explosion, enterprises today are migrating infrastructures from 10 Gb/s to 40 Gb/s and establishing a framework for 100 Gb/s systems. A basic building block of this framework is the development of a high-speed modular (HSM) port approach, which will enable data centers to pick and choose the physical-layer specifications that will deliver the performance needed at the right cost point. The HSM concept has been implemented for both 40 Gigabit Ethernet (GbE) and 100GbE, allowing networks to continue to scale with rate without causing major architectural changes.

This technical brief will provide an overview of the HSM approach, helping users take maximum advantage of the 40GbE and 100GbE architecture to come up with the appropriate interconnects and cost-effectively select from multiple vendors where interoperability can be ensured.

Introduction

IEEE 802.3™-2012 "Standard for Ethernet" defines a number of physical-layer specifications, including 40GBASE-CR4 and 100GBASE-CR10, which target twin-axial cabling, as well as optional electrical interfaces. The first of those optional interfaces, an Attachment Unit Interface (AUI), is intended for retimed applications, while the Parallel Physical Interface (PPI) is designed for non-retimed applications.[1]

HSM is envisioned as a kind of universal port type for the globally expanding Ethernet ecosystem. This powerful concept is being leveraged by system vendors to create ports capable of supporting a single form factor that can be populated with different modules that are each capable of supporting a specific physical-layer specification, such as multi-mode fiber, single-mode fiber, twin-axial copper cable or even an active cable solution. Care must be taken with such a flexible system, however, to minimize possibilities of mismatching system port types and devices. Incompatibility among system port capabilities, module families (CXP, QSFP or CFP), physical-layer specifications or differing cable types (active, passive or active optical cable) can lead to interoperability issues.

[1] For more details, the reader is encouraged to read the Ethernet Alliance white paper "40 Gigabit Ethernet and 100 Gigabit Ethernet Technology Overview" at http://www.ethernetalliance.org/wp-content/uploads/2011/10/document_files_40G_100G_Tech_overview.pdf.

Fortunately, the IEEE 802.3 architecture is elegant in its simplicity and flexibility, and a basic understanding will greatly help in avoiding any such occurrences.

Understanding the Problem

Compatibility among the various aspects of an HSM implementation is key to its successful deployment. Understanding these aspects will help reduce the potential for interoperability issues. As noted, HSM is envisioned as a kind of universal port, but it is the actual connection to that universal port that is often at the heart of potential interoperability issues. It might be argued that an "identity crisis" is truly the villain in any interoperability issues.

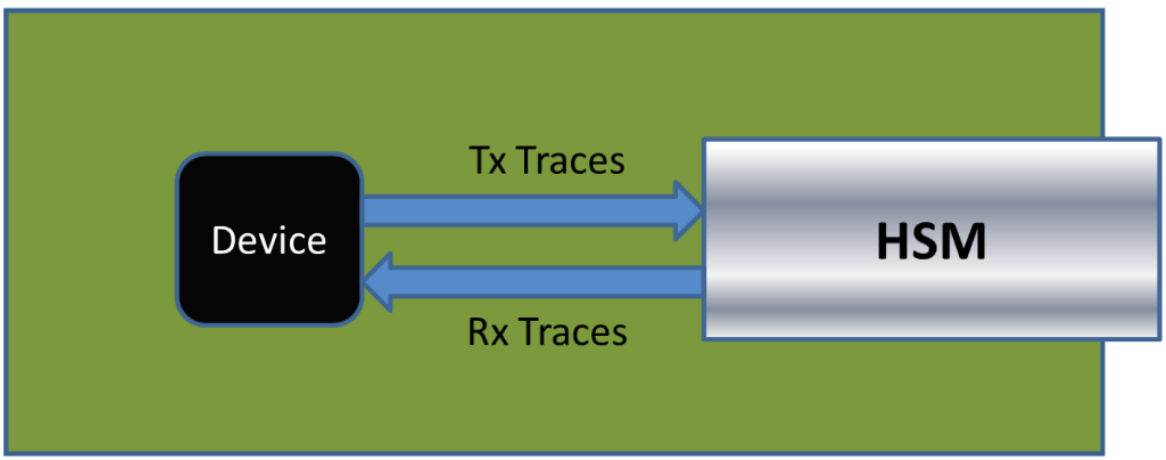

Figure 1: Generic HSM Configuration

As shown in Figure 1, there is a device driving traces that connect to the HSM connector interface. The ultimate personality of this configuration will be dependent upon the capabilities of the device, as well as the HSM being plugged in.

As noted, 40GbE and 100GbE each have two optional interfaces designed for chip-to-chip and chip-to-module applications: AUI for retimed applications and PPI for non-retimed applications. The interfaces for 40GbE are XLAUI and XLPPI, and 100GbE has CAUI-10 and CPPI-10. The "XL" and "C" represent 40 and 100 in Roman numerals, respectively. The "-10" is currently being introduced into the IEEE 802.3 Ethernet specification to indicate that the electrical interface is 10 lanes wide in each direction. Users are advised to request clarification of the type of interface if only the term "CAUI" is defined. A four-lane retimed interface that supports 100GbE, based on 25 Gb/s electrical signaling, is being developed at the time of writing this technical brief and will be referred to as CAUI-4. Furthermore, this width is the same for the 40GbE XLAUI interface. Thus, a module that is four lanes wide may support either 40GbE or 100Gbe, introducing another source of potential deployment confusion.

Table 1 - Summary of Electrical Interfaces for HSM Modules

Name	Ethernet Rate	Description
XLPPI	40GbE	A non-retimed interface, based on four lanes of 10 Gb/s in each direction.
XLAUI	40GbE	A retimed interface, based on four lanes of 10 Gb/s in each direction.
CPPI	100GbE	A non-retimed interface, based on 10 lanes of 10 Gb/s in each direction.
CAUI-10	100GbE	A retimed interface, based on 10 lanes of 10 Gb/s in each direction.
CAUI-4	100GbE	A retimed interface, based on four lanes of 25 Gb/s in each direction.

These electrical interfaces may be utilized inside the system or in the module that gets plugged into the HSM port. Therefore, a fundamental key to ensuring interoperability is to make certain that the system port and the module itself are capable of supporting the same interface and are configured to be operating with the desired interface.

This same port may also be used to support physical-layer specifications for passive copper cables, such as 40GBASE-CR4, 100GBASE-CR10 or 100GBASE-CR4. In this instance, the traces that connect to the HSM port are actually considered to be part of the channel that these physical-layer specifications have been designed to solve. Therefore, they employ more advanced signaling techniques, given the longer inherent channel being targeted. So, another potential interoperability issue could occur if a system is configured to be driving an AUI or PPI interface, intended for chip-to-module applications, but actually has a passive copper cable plugged into the system.

Thus, while the HSM port is extremely powerful in its flexibility, care must be taken during deployment and configuration to match its various capabilities with the intended I/O module or cable that is ultimately plugged in.

Specifications & Form Factors

While the IEEE 802.3 Ethernet standard does not specify implementations of optical modules, it does define the connector interface, or Media Dependent Interface (MDI), for the passive copper cable physical specifications 40GBASE-CR4, 100GBASE-CR10 and 100GBASE-CR4.

Table 2 - IEEE 802.3 Copper Cable MDI Specifications

Specification	IEC 61076-3-113*	QSFP**	CXP	CFP4
40GBASE-CR4	√	√		
100GBASE-CR10			√	
100GBASE-CR4		√		√

* Note: This implementation is compatible with 10GBASE-CX4.
** Note: The actual implementation is QSFP+ for 40GBASE-CR4 and QSFP+28 for 100GBASE-CR4. The two share a common connector interface and are compatible.

The QSFP, CXP and CFP4 MDIs are also used as connector interfaces to optical module form factors that support one or more optical specifications.

Table 3 - IEEE 802.3 Ethernet Optical Specifications

Port Type	Reach	40 GbE	100 GbE
40GBASE-SR4	At least 100m OM3 MMF/150m OM4 MMF	√	
40GBASE-FR	At least 2km SMF	√	
40GBASE-LR4	At least 10km SMF	√	
40GBASE-ER4	At least 40km SMF	√	
100GBASE-SR4	At least 100m MMF		√
100GBASE-SR10	At least 100m OM3 MMF/150m OM4 MMF		√
100GBASE-LR4	At least 10km SMF		√
100GBASE-ER4	At least 40km SMF		√

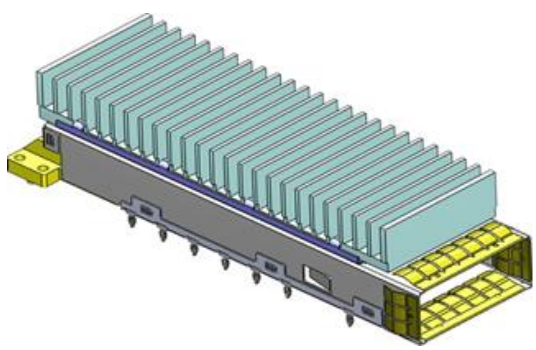

Figure 2: QSFP Form Factor

The acceptance of the QSFP form factor, Figure 2, has led to its adoption in data center applications for 40GbE and has been a guiding force as 100GbE, based on a 4x25 Gb/s configuration, is being developed. In addition to the development of 25Gb/s electrical signaling, there has been additional industry effort underway since 100GbE was defined, in order to increase the thermal capacity of the QSFP form factor.

The industry needed a solution for 100GbE that would be based on a 10x10 Gb/s electrical interface. In addition, the power/thermal cooling capability required for the 100GbE 10km optics could not be met by the QSFP form factor. The CFP MSA emerged and quickly developed an optical module roadmap for three generations of modules: CFP, CFP2 and CFP4, shown in Figure 3. The development of this family was driven by the electrical, optical and packaging technology directions being pursued in industry. It should be noted that the CFP MSA is a generic form factor that supports Ethernet and other specifications. The reader is directed to the MSA website, http://www.cfp-msa.org/, for more information.

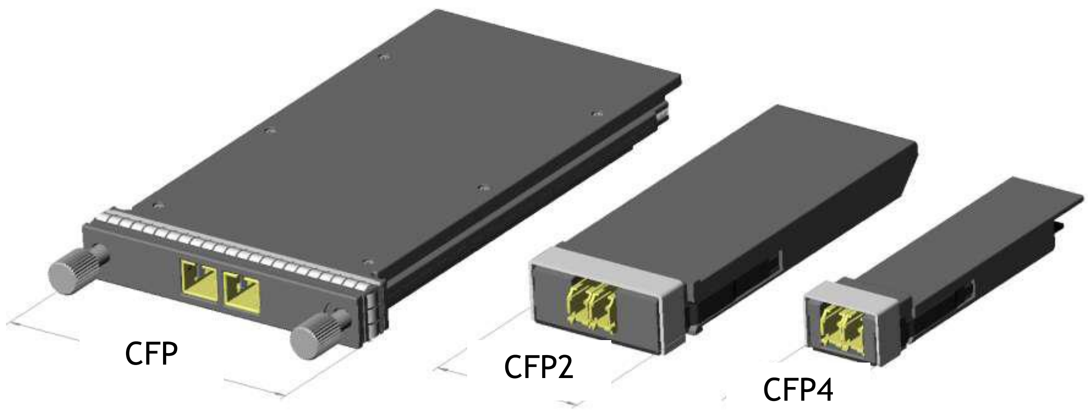

Figure 3: CFP Family of Modules[2]

Table 4: CFP Ethernet Roadmap

Module	Electrical Interface Supported	Module Width
CFP	4x10 Gb/s, 10x10Gb/s	82.00 mm
CFP2	4x10 Gb/s, 10x10Gb/s, 4x25Gb/s	41.50 mm
CFP4*	4x25Gb/s	21.70 mm

* Note: Per baseline concept.

The last form factor, CXP, is typically used for support of 100GBASE-CR10 and 100GBASE-SR10.

As noted, there are several high-speed interfaces available. The user should take care to make sure that the electrical interface of the specific module being used matches the capabilities of the HSM port on the module.

Also, the deployment of HSM ports means a port may be able to support many physical-layer specifications through the population of the desired module type. Using on-board storage and defined registers, the personality (type of module or cable itself) is indicated to a system, allowing the system to configure itself properly and drive modules appropriately. In some cases, the ports on a given board can exhibit

[2] Courtesy of CFP MSA, http://www.cfp-msa.org/Documents/cole_EAb_OFC12.pdf.

different personalities, in terms of being a retimed or a non-retimed module. Electrical interfaces can and do vary. There are multiple variants of CFP modules, for example. And the industry develops new modules dynamically driven by new application needs, so new opportunities for incompatibility can crop up.

Proprietary Implementations

The definition of chip-to-module interfaces has enabled the development of implementations where the stated electrical requirements via the module interface are known. This has enabled modules to be developed that may be multi-rate and may support retimed and non-retimed applications. The presence of an IEEE-specified electrical interface has actually empowered a number of proprietary solutions. A range of implementations may be in play at a given port, such as passive direct attach copper cables, active copper cables, active optical cables (AOC), non-retimed optical modules and retimed optical modules, based on duplex or multiple parallel fibers. In addition, the Optical Internetworking Forum (OIF) has developed a series of specifications, known as Multi-Link Gearbox (MLG), that will enable 10GbE or 40GbE to be driven over the 4x25 Gb/s interface. This last feature will play a key role in enabling even higher densities, as less face-plate space may be used for modules empowered with MLG. This will also improve the number of lower-speed ports that can get into application-specific integrated circuits (ASICs) that are limited by the number of pins at the ASIC.

Summary

The proliferation of HSM interconnect devices is giving operators of data centers the ability to pick and choose the interconnects that give them the performance they need at the right price point. To ensure that equipment can be sourced flexibly and cost-effectively from multiple vendors and real-world interoperability at the physical level of the network can be achieved, there are critical questions to be asked of suppliers beyond checking for a given module's mere compliance with the IEEE 802.3 Standard for Ethernet.

Understanding such nuances of higher-speed modular I/Os fosters more intelligent connectivity and cost-effective decision making with regard to different applications. Performing the due diligence to ensure compatibility across variables such as form factor and module, interface and cable types will help to ensure interoperability and unlock the true potential envisioned by the HSM module concept.