10GE Pluggable Transceiver Technology

Richard A Steenbergen <ras@nlayer.net> nLayer Communications, Inc.

A Quick Review: What are Pluggables?

- Transceivers which you can plug into routers, switches, transport gear, or pretty much any network device which will transmit and receive a signal.
- Hot swappable while the device is operating.
- Capable of operating over many different physical mediums and at different distances.
  - Copper, MMF, SMF, 10km, 40km, 80km, etc, etc.
- Standardized to be interchangeable among vendors.
  - Well, between pluggable vendors at any rate.
  - But more on this one later.

The Benefits of Using Pluggables

- Technical Benefits
  - Accommodates various media type and reach needs
  - Easy replacement in the event of component failures

- Financial Benefits
  - “Pay as you Populate” model lowers initial costs.
  - Pluggables are reusable in new cards or new systems.
  - Cards are reusable as the optical technology evolves.
  - Standardization and increased competition lowers costs.

How Pluggables are Standardized

- Interchangeable Components Require Standards
  - Allows for interoperability and mass production
- Standardization is achieved through a vendor MSA
  - Multi Source Agreement – A group of vendors who get together to develop a specification for a standardized hardware component
  - Specifications published under SFF - defines a strict standard for the physical, electrical, mechanical, and management interfaces.

First Generation Pluggable Technology

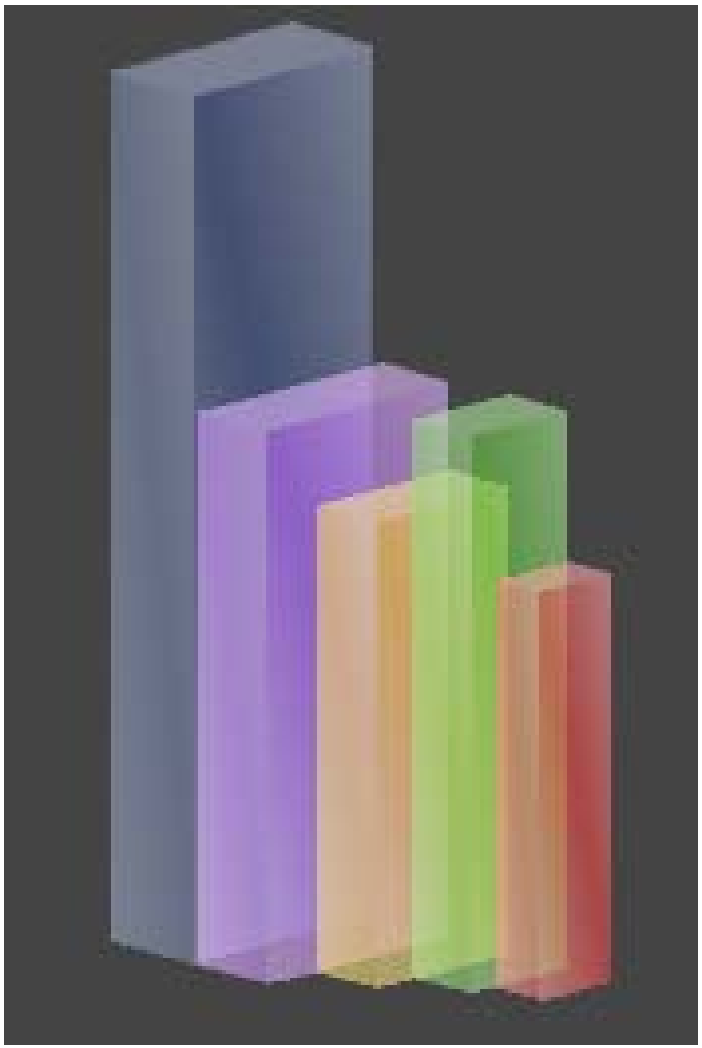
- GBIC (GigaBit Interface Converter)
  - Originally designed for 1G Fibre Channel (FC)
  - Quickly adopted for use in Gigabit Ethernet applications
  - Common on new gear from 1998 ~ 2002, still used today
- SFP (Small Form-factor Pluggable)
  - First published in 2002, extended in 2004 and 2007
  - Introduced Digital Optical Monitoring (DOM) in 2004
  - Multi-Rate SFPs are widely deployed today
    - 1G/2G FC / 1.25G (GigE) / 2.5G (OC-48) support is common
    - 4G FC SFPs are readily available as well
- Current standard for 1GE fiber – 48xSFP cards

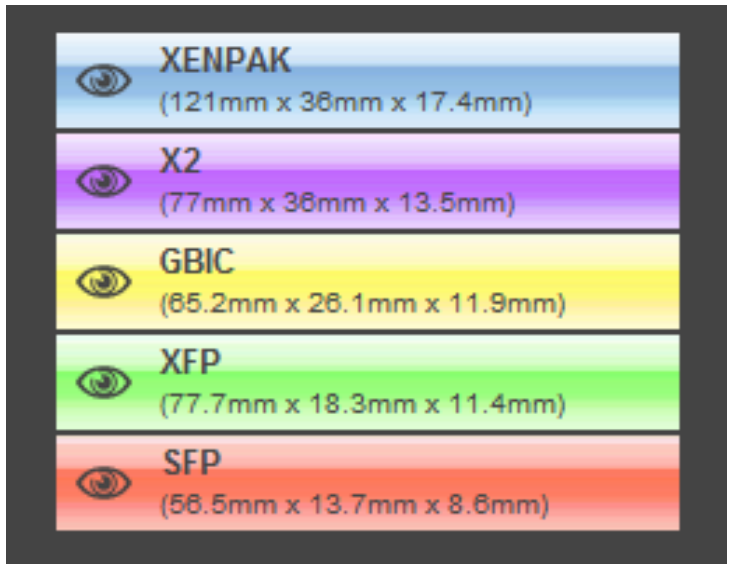
Pick Your Pluggable Technology

- With 1G – 2.5G the choices were few, and simple
- 10GE offers a much wider range of choices
  - 300-pin
  - XENPAK
  - XPAK
  - X2
  - XFP
  - XFP-E
  - SFP+
- With many different designs and implementations
- And each with their own pros and cons

Maybe a Few Too Many Choices?

Pluggable Physical Size Comparison





300-Pin MSA

- First generation 10G
- Interface: 16 x 622Mbps
- Power: Max 8-14W
- Notes: Comes in various sizes
  - Not really a pluggable, uses a “snap-on” connector.
  - Naturally suited to a first generation technology
    - Large size footprint and many low-speed lanes.
  - Similar format being used today for 40G optics.
- If you’re still using these for 10G, I’m sorry.

XENPAK MSA

- MSA Founded: March 2001
- Interface: XAUI (4x3.125G)
- Power: Max use of 6-10W
- PHY Framer: Onboard
- Deployment: Common
- Exotic Optics: Largest collection
  - First platform for 80km+ ZR optics, DWDM tuned optics, etc
- Notes: Popular among enterprises (CX4/LX4 support)

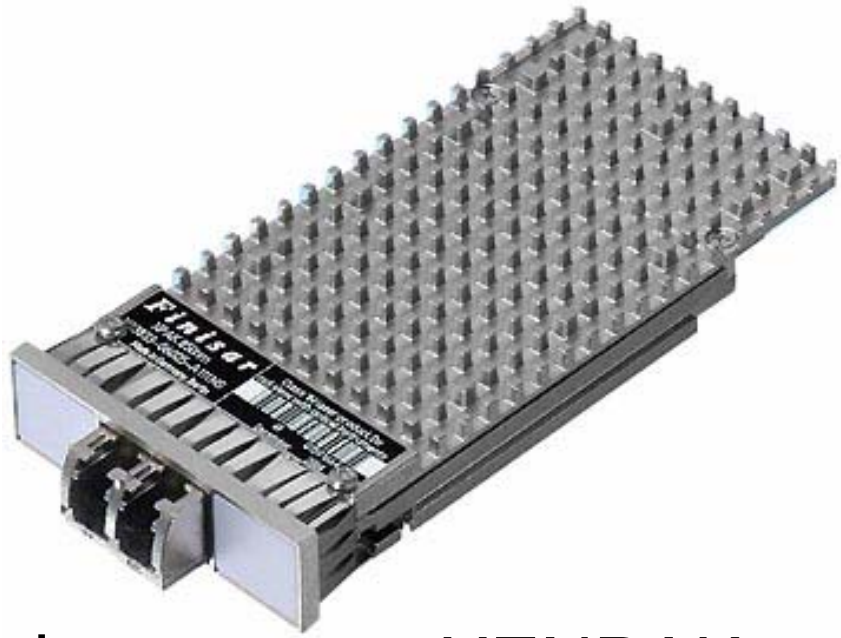
X2 MSA

- MSA founded: July 2002
- Interface: XAUI (4x3.125G)
- Power: Max use of 4-5W
- PHY Framer: Onboard
- Deployment: Limited
- Exotic Optics: Very limited

- Notes: Electrical interface is the same as XENPAK
  - Easy for existing XENPAK boards to switch to X2 with very little board redesign and no ASIC changes.

XPAK MSA

- MSA founded: ???
- Interface: XAUI (4x3.125G)
- PHY Framer: Onboard
- Deployment: VERY Limited
- Exotic Optics: None



- Notes: Electrical interface is the same as XENPAK
  - VERY similar to X2, but optimized for use on PCI cards
    - A lot of talk about merging with X2, but at this point nobody cares
  - All but unheard of in the networking world.
  - And can be safely ignored for the rest of this talk.

XFP MSA

- MSA founded: March 2002
- Interface: XFI (9.95-11.1G)
- Power: Max use 1.5-3.5W
- PHY Framer: Offloaded
- Deployment: Very common

- Exotic Optics: Full ZR/DWDM, limited CX4, no LX4
- Notes: Extremely popular on new 10G equipment
  - Eliminating SerDes for 10GBASE-R/W is a big power saver
  - Some CX4 support recently added, but not optimized for it.

SFP+ MSA

- MSA: Latest draft December 2007
- Interface: SFI (8.5 – 11.1G)
- PHY Framer: Offloaded
- Deployment: Barely shipping
- Exotic Optics: None today
- Notes: Physically compatible with original SFPs
  - Around 30% smaller than XFP (offloads CDR function)
  - SFI similar to XFI, adds support for 8G FC speeds
  - Extremely limited power use, no long reach optics at all
  - Target market: FC, Enterprises, high-density 10GE LAN

Summary of Pluggable Characteristics

	XENPAK	X2	XFP	SFP+
Interface Type	XAUI	XAUI	XFI	SFI
Interface Speed	4x3.125Gb	4x3.125Gb	9.95-11.1Gb	8.5-11.1Gb
PHY/Framer	Pluggable	Pluggable	Host	Host
SerDes	Pluggable	Pluggable	Host (Optional)	Host (Optional)
CDR	Pluggable	Pluggable	Pluggable	Host
Max Power Use	6.0-10.0W	4.0-5.0W	1.5-3.5W	1.0-1.5W
Max Ports/Blade	4	8	16	48
Protocol Agnostic	No	No	Yes	Yes

10GE Pluggable Component Terminology

- PHY – Physical layer component
  - PCS – Physical Coding Sublayer
  - PMA – Physical Medium Attachment Sublayer
  - PMD – Physical Medium Dependant Sublayer
- SerDes – Serializer / Deserializer
  - Converts between serial and parallel signals
- CDR – Clock and Data Recovery
  - Provides retiming and signal conditioning

10GE PHY PCS – Serial or Parallel

- The 10GE PHY/PCS comes in 3 basic flavors
  - 10GBASE-R – LAN PHY serial 10G Signal
  - 10GBASE-W – WAN PHY serial 10G Signal
    - Similar to –R but wrapped in a OC192 SONET compatible frame
  - 10GBASE-X – LAN PHY 4x2.5G parallel Signal
- Low speed signals are easier to TX/RX cleanly
- But multiplexing adds overhead (extra chips, power)
- And parallel lanes requires separate paths
  - Optical: LX4 – 4 wavelengths of light + CWDM multiplexer
  - Copper: CX4 – 4 parallel paths over copper

Pluggable Interconnection Technology

- Ironically, 10G pluggable components have similar options for talking to each other in serial or parallel
  - XAUI – uses 4 x 3.125G parallel lanes
    - 2.5G signal + 8B/10B encoding overhead = 3.125G signal
  - XFI/SFI – uses single variable speed ~ 10G lanes
    - 10.0G signal + 64B/66B encoding overhead = 10.3125G signal
- SerDes ASIC translates serial and parallel streams
  - But adds overhead every time you do the conversion
    - Every SerDes adds to the costs to the component
    - Consumes board space, limiting physical density
    - Consumes power, limiting thermal density

Serial or Parallel: 10G vs 4x3.125G

10G PHY/PCS	XAUI (XENPAK/X2)	XFI/SFI (XFP/SFP+)
10GBASE-R/W	Requires SerDes	Native
10GBASE-X	Native	Requires SerDes

- Optimal configuration is a native pairing
  - XAUI talking to 10GBASE-X
  - XFI/SFI talking to 10GBASE-R/W
- Sub-optimal configurations require a conversion
  - XAUI talking to 10GBASE-R/W
    - Plenty of room for SerDes on a XENPAK board at least
  - XFI/SFI talking to 10GBASE-X
    - Fitting a SerDes in a XFP/SFP+ is not quite so easy

The Future of Parallel PHY/PCS

- 10GBASE-X exists to reduce costs, extend distances
  - 10GBASE-LX4 can run 300m over FDDI grade multimode
    - And 10km over SMF, making it a “10G Swiss Army Knife”.
  - 10GBASE-CX4 can run 15m over Infiniband style cables
    - But VERY cheaply, perfect for short distance deployments
- But 10G technologies are being developed to replace them
  - 10GBASE-LRM – Long Reach Multimode to replace LX4
    - And offer extended reach over SR’s “26 meters on a good day”.
  - 10GBASE-T – 10G copper to replace CX4
    - Longer reach, uses standard Cat6/Cat7, good old 8P8C (RJ45).
- In the long term, 10GBASE-X has a limited lifespan

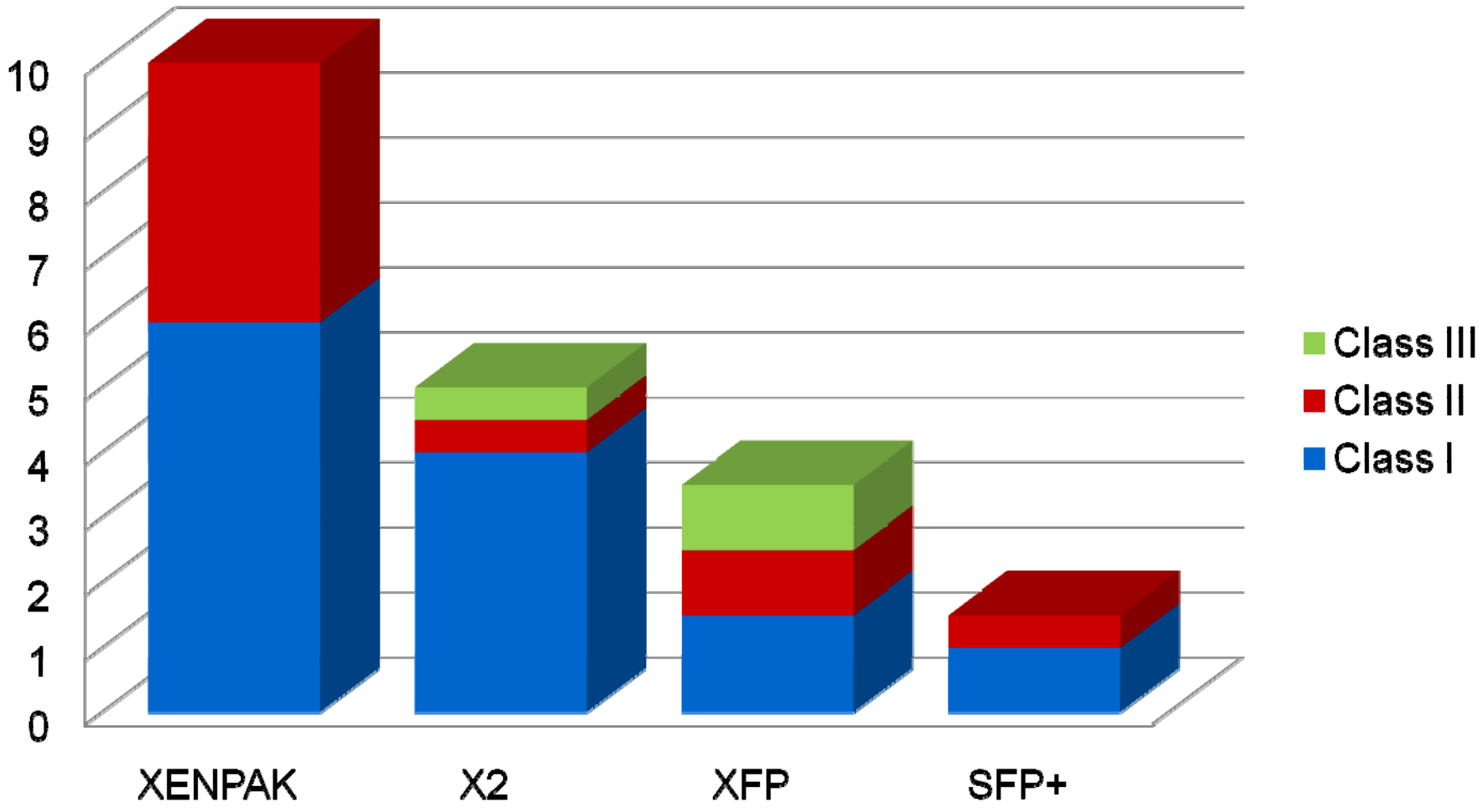
Advantages of Offloading the PHY/PMA

- Makes the pluggable “protocol agnostic”
  - Can use the same pluggable for 10GE LAN/WAN, OC-192 SONET, 10G FC, G.709 Forward Error Correction, OTN, etc.
  - Allows component reuse, lowers costs between industries.
- 10GE WAN PHY **much** better when done on the host
  - A host which implements WAN PHY can now use any PMD
    - SR LR ER ZR DWDM etc, without having to buy special pluggables
    - Improves sparing, lowers costs (significantly), expands PMD options
  - Improves WAN layer control signaling and alarms
    - Provides access to SONET alarms, path trace, etc.
    - Vastly improves troubleshooting when working with a OC192 carrier

Advantages of Offloading the CDR

- Makes the pluggable smaller and use less power
- May be outweighed by disadvantages though
  - Mostly a zero sum game
    - Doesn’t actually eliminate components, like SerDes offloading
    - Doesn’t provide technical advantages like Framer offloading
    - Basically just moving the component from one place to another
  - Not all CDRs are created equally
    - The CDR in SR optics may not need to be as good in a ZRD unit
    - Not possible to upgrade optics to keep up with advancements in EDC (Electronic Dispersion Compensation) technology.
  - One more reason why SFP+ may never be suitable for use in medium/long reach or DWDM applications.

Evolution of Pluggable Power Use (Watts)



Comparison of Optics by Pluggable

	XENPAK	X2	XFP	SFP+
SR (26m)	✓	✓	✓	✓
LR (10km)	✓	✓	✓	✓
ER (40km)	✓	✓	✓	⊘
ZR (80km)	✓	Not Today	✓	⊘
LX4/CX4	✓	✓	⊘	⊘
DWDM	✓	Not Today	✓	⊘
LW (10km)	✓	✓	✓	✓
EW (40km)	Not Today	Not Today	✓	⊘
ZW (80km)	Not Today	Not Today	✓	⊘
LRM (300m)	✓	✓	✓	✓

The Software Side of Pluggables

- Pluggables talk to host via a low speed control bus
  - Control commands such as “power up”, “power down”, etc.
  - Diagnostic information such “loss of signal”, etc.
  - Another recent addition is Digital Optical Monitoring (DOM)
    - Essentially a built-in light meter in your optics
      - Or Time-Domain Reflectometer (TDR) for copper
    - Absurdly helpful in troubleshooting layer 1 issues
  - But hosts are also capable of reading EEPROM data
    - Optic type, media type, reach, connector type, etc
    - But also: vendor name, part number, serial number, etc
    - The vendor information turns out to be very important

Money, Cash, Cisco's…

- First, a couple of key facts:
  - **No** major networking vendor makes their own pluggables.
  - Pluggables are manufactured by OEMs like Finisar, Intel, Emcore, Agilent, Opnext, Hitachi, JDS/Uniphase, etc.
  - Pluggable EEPROM Vendor IDs can be easily programmed
    - Making it possible for every network vendor to have a "store brand".
  - There is a significant market in reselling pluggables
    - Typical markup during resale by the router vendors is 10-25x.
    - Accounted for 25% of Cisco's FY06 profit: **$1.4 BILLION dollars**.
- This creates a strong incentive for network vendors to keep customers using only "their" brand of optics.

How Vendors Keep Customers Locked In

- Psychology: Fear / Uncertainty / Doubt (FUD)
  - “If you don’t use us you’re getting inferior knockoffs”
  - “This might void your warranty and fry your router”
  - “We can’t provide support if you don’t use our optics”
- Finance: Market Oligopsony
  - AKA Many sellers, few buyers.
  - Cisco purchases 70% of all pluggables sold by OEMs
    - OEMs who don’t play by Cisco’s rules risk losing all business.
- When all else fails, implement vendor locking
  - If the EEPROM doesn’t say “our brand”, disable the port.

The Latest Trends in Vendor Locking

- Vendor locking has been going on for some time
  - But customer outrage has kept it somewhat restrained.
  - Most vendors at least offer a (hidden) disable command.
- Professional counterfeiters are not being deterred
  - Vendor locking actually seems to encourage counterfeits
    - Locking prevents vendors from selling to consumers legitimately
      - So their only option is to produce and sell counterfeits
    - Cloned optics are finding their way into VAR supply chains.
- Latest strategy is “Feature Impairment”
  - “If we can’t disable it, we can just not support all features”
  - One targeted “extra” seems to be DOM support.

Which Format is Right for You? - XENPAK

- Advantages
  - Large established base, stocked by every vendor.
  - Currently has the best selection of long reach / DWDM
  - Full support for LX4/CX4 for 10G over MMF/Copper
- Disadvantages
  - Large format, draws a lot of power, not high density
  - Not protocol agnostic, not friendly with WAN PHY
    - WAN PHY may cost 3x more than equivalent LAN XENPAK
  - Not the lowest cost solution, either by volume or COGS
- Still useful in many respects, but maybe not the best choice for new deployments or high density.

Which Format is Right for You? – X2

- Advantages
  - Smaller than XENPAK, lower power allows higher density
  - Easy for XENPAK users to adopt with little cost/effort.
  - Full support for 10GBASE-X LX4/CX4 technologies.
  - Being pushed by Cisco in new stackables, 6708, 6716, etc
- Disadvantages
  - Very little deployed base, few vendors supporting this.
  - Another thing to spare, exotic optics not currently available.
  - Not protocol agnostic, same limitations as XENPAK.
- “The worst of both worlds” between XENPAK and XFP

Which Format is Right For You? – XFP

- Advantages
  - Large established deployed base, cheap/easy to buy.
  - LRM and 10GBASE-T will offer alternatives to LX4 / CX4
  - Power and density still very reasonable for most users
  - Protocol agnostic, easier to spare, better for WAN PHY
  - The cheapest option for long reach / DWDM optics today
- Disadvantages
  - Still no LX4, CX4 available but not optimal.
  - Will never be able to achieve 48-port per blade densities
- Still the best all-around choice for today and the foreseeable future.

Which Format is Right for You? – SFP+

- Advantages
  - Extremely high densities are possible (48-ports per blade)
  - Physically the same as SFP, possible compatibilities
    - Goal is 1G/10G Ethernet or 1G/2G/4G/8G/10G FC ports.
  - May allow for extremely low cost devices (“GoogleSwitch”)
- Disadvantages
  - Limited power budget may never support long reach optics
  - No adoption yet, barely even announced let alone shipping
  - No interest outside of Enterprise or converged Ethernet/FC
- Not inherently bad, but not a complete replacement for XFP and doubtful that it ever will be.

Further Reading

MSA	URL
XPAK	ftp://ftp.seagate.com/sff/INF-8475.PDF
XENPAK	http://www.xenpak.org/MSA/XENPAK_MSA_R3.0.pdf
X2	http://www.x2msa.org/X2_MSA_Rev2.0b.pdf
XFP	http://www.xfpmsa.org/XFP_Rev4_5_SFF_INF_8077i.pdf
SFP+	ftp://ftp.seagate.com/sff/SFF-8431.PDF

Send questions, comments, complaints to:

Richard A Steenbergen <ras@nlayer.net>