10 Gbps Cabling
DC Facilities Top of Mind

Complexity, Cost, Power, Cooling Standards Compliance Reliability, Availability Management, Security Future Proofing

Increased Efficiency, Simpler Operations Scalability, Flexibility Technology Modularity, Mobility
10GE Server Evolution in the DC access

- The 10GbE LOM market will start in 2011 (Intel server arch. release cycle)
- 10G LOMs will show up first on premium servers.
- From 2011 it might take about 2 years to see the majority of servers 10GBASE-T LOM capable.
- Fiber & CX1 today (Multiple optics choices)
- Fiber, CX1, 10GBASE-T tomorrow
10 Gigabit Transmissions

Different Standards
- 10GBase-T (IEEE 802.3an)
- 10GBase-CX4 (IEEE 802.3ak)
- 10GBase-*X (IEEE 802.3ae)
- SFF 8431 (SFP+ Fiber & cu)

Applications
- Server Interconnects
- Aggregation of Network Links
- Switch to Switch Links
- Storage Area Networks (SAN)
# 10 Gbps Transciever Options

## 10G Sever Connectivity Options—UTP/F-UTP, MMF, SMF, TwinAx, CX4

<table>
<thead>
<tr>
<th>Connector (Media)</th>
<th>Cable</th>
<th>Distance</th>
<th>Power (each side)</th>
<th>Transceiver Latency (link)</th>
<th>Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>SFP+ CU* copper</td>
<td>Twinax</td>
<td>&lt;10m</td>
<td>~1.5W</td>
<td>~0.1μs</td>
<td>SFF 8431**</td>
</tr>
<tr>
<td>X2 CX4 copper</td>
<td>Twinax</td>
<td>15m</td>
<td>4W</td>
<td>~0.1μs</td>
<td>IEEE 802.3ak</td>
</tr>
<tr>
<td>SFP+ USR MMF, ultra short reach</td>
<td>MM OM2</td>
<td>10m</td>
<td>1W</td>
<td>~0</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>MM OM3</td>
<td>100m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SFP+ SR MMF, short reach</td>
<td>MM OM2</td>
<td>82m</td>
<td>1W</td>
<td>~0</td>
<td>IEEE 802.3ae</td>
</tr>
<tr>
<td></td>
<td>MM OM3</td>
<td>300m</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RJ45 10GBASE-T copper</td>
<td>Cat6</td>
<td>55m</td>
<td>~6W***</td>
<td>2.5μs</td>
<td>IEEE 802.3an</td>
</tr>
<tr>
<td></td>
<td>Cat6a/7</td>
<td>100m</td>
<td>~6W***</td>
<td>2.5μs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cat6a/7</td>
<td>30m</td>
<td>~4W***</td>
<td>1.5μs</td>
<td></td>
</tr>
</tbody>
</table>

* Terminated cable  ** Draft 3.0, not final  *** As of 2008; expected to decrease over time
10GE SFP+ Optical

- Smallest 10GE form factor
- Low Power
- Low Latency
- Hot swappable
- High density
- Optical SFP+ interoperates with other 10GE modules
  - XFP
  - XENPAK
  - X2

SFP+ Optical Module
Optics Positioning for Data Center Applications

- **1G Optics Type**
  - 1000 BASE-LX
  - 1000 BASE-SK
  - 1000 BASE-T

- **10G Optics Type**
  - 10GBASE-LR
  - 10GBASE-LRM
  - 10GBASE-SR
  - 10GBASE-T
  - 10GBASE-USR
  - 10GBASE-CX4
  - 10GBASE-CX1

- **Max PMD Distance (m)**
  - 1000 BASE-LX: ~10000 m
  - 1000 BASE-SK: ~820 m
  - 1000 BASE-T: ~220 m
  - 10GBASE-LR: ~10000 m
  - 10GBASE-LRM: 30M/100M
  - 10GBASE-SR: ~820 m
  - 10GBASE-USR: OM3 MMF Only
  - 10GBASE-T: ~220 m
  - 10GBASE-CX4: OM3 MMF Only
  - 10GBASE-CX1: OM3 MMF Only

- **In Rack X-rack**
  - <10M

- **Mid to End of Rack**
  - <100 M

- **Across Aisles**
  - <300 M

- **Across Sites**
  - <10 KM
A look at the 100M > 1G transition (Desktop and Server driven)

- 1G/10G server transition expected to be within ~2-years once LOM is available (slower than at 1G – source Intel)
- Switch edge port adoption of high speed technology historically took longer than server transition (hence the need for 100/1000/10000 technology on LOM)
10GBASE-T PHY Evolution

- **1st Gen silicon**: ~10W, 90nm
- **2nd Gen silicon**: ~6W, 90-65nm
- **3rd Gen silicon**: ~4W Peak, ~3W Avg w/ EEE, 65-40nm
- **4th Gen silicon**: ~3W Peak, <1W Avg w/ EEE, 40nm


*802.3az EEE
*30m mode will save about 1W

*Energy Efficient Ethernet*
10GBASE-T PHY Evolution – Switch View

- 24 ports (or more) in TOR
- 2nd Gen silicon
  - ~6W
  - 90-65nm
- 3rd Gen silicon
  - ~4W Peak
  - ~3W Avg w/ EEE
  - 65-40nm
- 4th Gen silicon
  - ~3W Peak
  - <1W Avg w/ EEE
  - 40nm

- Up to 32 port on modular
- 24+ ports**
- Same fundamental architecture
- NEW architecture
- 2008 . 2009
- 2010
- 2012+

- Energy Efficient Ethernet
- 802.3az
- EEE

- *30m mode will save about 1W
- **New MAC ASIC needed for EEE
10GBASE-T PHY Evolution – Server View

- 24 ports (or more) in TOR: Extreme/SMC/Fulcrum Micro
- Up to 32 port on modular: Nexus 7000 6500/4900M w/ 2nd gen

- Single port NIC: Shipping Platform
- Dual port NICs: Next Gen Platform (Q1C09)
- 1st Gen silicon: ~10W 90nm
- 2nd Gen silicon: ~6W 90-65nm
- 3rd Gen silicon: ~4W Peak ~3W Avg w/ EEE 65-40nm
- 4th Gen silicon: ~3W Peak <1W Avg w/ EEE 40nm

- Same fundamental architecture
- NEW architecture

- 2008 . 2009
- 2010
- 802.3az EEE

- 2012+
- *30m mode will save about 1W
- **New MAC ASIC needed for EEE
Twisted Pair Cabling For 10GBASE-T (IEEE 802.3an)

- **U/UTP** (Old designation UTP)
  - Outer Unshielded/Inner Pairs Unshielded

- **F/UTP** (Old designation FTP)
  - Outer Foil Shielded/Inner Pairs Unshielded

- **Cat 6a**:
  - *100m 10GBASE-T*
  - **largest diameter up to 0.354 in**

- **Cat 6**:
  - *55m 10GBASE-T*
  - **larger diameter than Cat5 (~0.3 in)**

- **Cat 6/6a**:
  - *100m 10GBASE-T*
  - **More flexible/easier to manage than Cat6a U/UTP**
  - **Equivalent diameter to Cat6**

- **S/FTP** (Old designation S/STP)
  - Outer Foil Shielded/Inner Pairs Foil shielded

- **Cat 7**:
  - *100m 10GBASE-T*
  - **Most expensive**
  - **Smaller diameter than Cat6a**
  - **Not popular in North America**
Shielded (F/UTP) vs. Unshielded (6a U/UTP)

**F/UTP**
- Alien XT reduction w/ shielding
- Diameter 0.3in/7.6mm
- Up to* 40% cable conduit fill-ratio higher than U/UTP
- Cabling choice in Cisco’s new Richardson Data Center

**U/UTP**
- Alien XT reduction w/ spacing
- Diameter 0.354in/9mm (worst case)
- Stiffer/less flexible cable than F/UTP

*Assumption of 40% conduit fill ratio and various conduit sizes (Source Tyco)
## Copper Cabling for Data Centers

<table>
<thead>
<tr>
<th>Frequency Bandwidth</th>
<th>TIA (Components)</th>
<th>TIA (Cabling)</th>
<th>ISO (Components)</th>
<th>ISO (Cabling)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 100 MHz</td>
<td>Category 5e</td>
<td>Category 5e</td>
<td>Category 5e</td>
<td>Class D</td>
</tr>
<tr>
<td>1 - 250 MHz</td>
<td>Category 6</td>
<td>Category 6</td>
<td>Category 6</td>
<td>Class E</td>
</tr>
<tr>
<td>1 - 500 MHz</td>
<td>Category 6A</td>
<td>Category 6A</td>
<td>Category 6A</td>
<td>Class E&lt;sub&gt;A&lt;/sub&gt;</td>
</tr>
<tr>
<td>1 - 600 MHz</td>
<td>n/s</td>
<td>n/s</td>
<td>Category 7</td>
<td>Class F</td>
</tr>
<tr>
<td>1 - 1,000 MHz</td>
<td>n/s</td>
<td>n/s</td>
<td>Category 7&lt;sub&gt;A&lt;/sub&gt;</td>
<td>Class F&lt;sub&gt;A&lt;/sub&gt;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application</th>
<th>Category 5e Class D</th>
<th>Category 6 Class E</th>
<th>Category 6A Class E&lt;sub&gt;A&lt;/sub&gt;</th>
<th>Class F</th>
<th>Class F&lt;sub&gt;A&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/16 MBPS Token Ring</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>10BASE-T</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>100BASE-T4</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>155 MBPS ATM</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>1000BASE-T</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>TIA/EIA-854</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>10GBASE-T</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Broadband CATV</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Charts extracted from www.siemon.com
10G SPF+ Cu

- SFF 8431
- Supports 10GE passive direct attached up to 10 meters
- Active cable options to be available
- Twinax with direct attached SFP+
- Primarily for in rack and rack-to-rack links
- Low Latency, low cost, low power
10G Copper Infiniband - 10GBase-CX4

- IEEE 802.3ak
- Supports 10G up to 15 meters
- Quad 100 ohm twinax, Infiniband cable and connector
- Primarily for rack-to-rack links
- Low Latency
HSSG: Higher Speed Study Group

HSSG Objectives

- Support full-duplex operation only
- Preserve the 802.3 / Ethernet frame format utilizing the 802.3 MAC
- Preserve minimum and maximum FrameSize of current 802.3 standard
- Support a BER better than or equal to 10^{-12} at the MAC/PLS service interface
- Provide appropriate support for OTN

- Support a MAC data rate of 40 Gb/s
  - Provide Physical Layer specifications which support 40 Gb/s operation over:
    - at least 100m on OM3 MMF
    - at least 10m over a copper cable assembly
    - at least 1m over a backplane

- Support a MAC data rate of 100 Gb/s
  - Provide Physical Layer specifications which support 100 Gb/s operation over:
    - at least 40km on SMF
    - at least 10km on SMF
    - at least 100m on OM3 MMF
    - at least 10m over a copper cable assembly
Access Layer Network Model
End of Row, Top of Rack & Blade Switches

GE Access / 10GE

What it used to be…
End of Row / MoR
Top of Rack
Active Zone Above Row/Rack
Blade Switches

What else is coming…
End of Row w/ Blade Enclosures
ToR w/ Blade Enclosures
End of Row w/ FEX
EoR FEX

What influences the physical layout…
Primarily:
- Power
- Cooling
- Cabling
Secondarily:
- Access Model
- Port Density
Pod Concept - Sizing

Pod Concept

Network Zones and Pods

Sizing

Zone: Typically mapped area in the DC
Pod: Typically mapped to agg pair with ToR switches in EDA/access cabinets
Size: determined by distance and density
Cabling distance from server racks to network racks
Fiber
200-500m Fiber
Cabling density: # of servers per rack and I/Os per server
Rack
Server: 6-30 Servers per rack — limited by power
Network: Depends on access model: Modular, ToR or Blade
Storage: special Cabinets
Physical Infrastructure and Network Topology
Modular Cabling Architecture Methodology
TIA-942 Logical Layout
EIA/TIA 568 Copper & Fiber Cabling

Access Providers

Office & Operations Center LAN Switches

Main Dist Area
(Routers/Backbone LAN/SAN Switches, PBX, M13 Muxes)

Horiz Dist Area
(LAN/SAN/KVM Switches)

Access Providers

Telecom Room
(Office & Operations Center LAN Switches)

Horiz Dist Area
(LAN/SAN/KVM Switches)

Horiz Dist Area
(LAN/SAN/KVM Switches)

Horiz Dist Area
(LAN/SAN/KVM Switches)

Computer Room

Front Desk Area
(Rack/Cabinet)

Backbone Cabling

Core / EoR Agg

EoR Access / ToR Agg

ToR

Horizontal Cabling

Backbone Cabling

Equipment Dist Area
(Rack/Cabinet)

Equipment Dist Area
(Rack/Cabinet)

Equipment Dist Area
(Rack/Cabinet)

Equipment Dist Area
(Rack/Cabinet)
Dual-Homed 1-G Attached Servers

10G SFP+Fibre cables connecting FEX to Nexus 5010
10G SFP+Fibre cables to Aggregation Layer
ToR Configuration

- Direct Attach CXI Copper cables connecting FEX to Nexus 5010
- 10G SFP+Fibre cables to Aggregation Layer
Mixed 1-G and 10-G Attached Servers
Review of the Nexus 2000 Fabric Extender
Nexus 5000 Family Overview

Nexus 5010
- 20 Fixed Ports 10G/FCoE/Data Center Ethernet
- Line Rate, Non-Blocking 10G
- 1 Expansion Module slot

Nexus 2000 Fabric Extender (Q1CY09)
- 48 Fixed Ports 1G Ethernet
- 4 Fixed Ports 10G Uplink
- Distributed Virtual Line Card
- Hardware VN-Link Technology

Nexus 5020
- 40 Fixed Ports 10G/FCoE/Data Center Ethernet
- Line Rate, Non-Blocking 10G
- 2 Expansion Module slots

Ethernet
- 6-Port 10G/FCoE/Data Center Ethernet

Ethernet + Fibre Channel
- 4-Port 10G/FCoE/Data Center Ethernet
- 4-Port 1/2/4G Fibre Channel

Fibre Channel
- 8-Port 1/2/4/8G Fibre Channel

Fibre Channel (Q3CY09)
- 2/4/8G Fibre Channel

NX-OS, DCNM, CiscoWorks and Fabric Manager
Nexus 5020 Overview

- Ethernet Out-of-Band Management
- Wire-Speed 10GE/FCoE/DCE
- Front-to-Back Airflow
- 2 Expansion Modules
- Redundant Power Supplies
- Redundant Fans
Nexus 5010 Overview

- Ethernet Out-of-Band Management
- Wire-Speed 10GE/FCoE/DCE
- Front-to-Back Airflow
- 1 Expansion Module
- Redundant Power Supplies
- Redundant Fans
Nexus 2000 Fabric Extender (FEX)

The Nexus 5000 Fabric Extender (FEX) acts as a remote line card (module) for the Nexus 5000, retaining all centralized management and configuration on the Nexus 5000, transforming it to a Virtualized Chassis.
Nexus 2000 Fabric Extender
1GE Connectivity

- 48 x 1 GigabitEthernet Interfaces
- 4 x 10 GigabitEthernet Interfaces
- Beacon & Status LEDs
- Redundant, Hot-Swappable Power Supplies
- Hot-Swappable Fan Tray
From a management perspective, fabric extender is modeled as a NX5K linecard.

Nexus 5000 and Fabric extender communicate via inband connectivity.

There is no configuration stored on the Fabric Extender.

Fabric extender is not an independent manageable entity.
Nexus 2000 Fabric Extender
Network Topology – Physical vs. Logical

**Physical Topology**

- Central Point of Management
- 4x10G uplinks from each rack
- Nexus 5000 and Fabric Extender

**Logical Topology**

- Central Point of Management
- Nexus 5020
- Nexus 2000 Fabric Extender

*Nexus 5000 and Fabric Extender are one logical access switch in the layer 2 topology ToR flexibility *without* a larger STP topology*
Fabric Extender
Design Option 1 – Dual NIC

- Dual NIC (redundant NIC) designs can leverage server side redundancy
- A common configuration for VMWare servers
- Two vSwitches configured with redundant NIC’s
  - vSwitch 1 - Virtual Machine port groups
  - vSwitch 2 – VMkernel and service console port groups
- Common practice is to assign specific NIC’s to specific traffic types
Fabric Extender
Design Option 2 – Single NIC

- Single NIC (non-redundant NIC) designs usually are used:
  - For non critical systems
  - As compute nodes in a larger application architecture leveraging ACE and application level redundancy
- All upstream traffic is multiplexed over a single NIC
- Note: Dual NIC servers often leverage ACE and application redundancy in addition to the physical NIC redundancy
Fabric Extender
MCEC and FEX uplink

- Design Option 2 - Nexus 5000 will support a vPC uplink configuration from the Fabric Extender
- Multiple uplinks from FEX to Nexus 5000 provide data plane redundancy
- MCEC provides redundancy for the control plane
- Two elements of the MCEC configuration
  - vPC which provides the multi-chassis capabilities
  - Physical Port Channel on each of the individual Nexus 5000

In this example FEX is configured with two 10GE vPC uplinks
Fabric Extender MCEC and FEX uplink

- The Fabric Extender MCEC vPC configuration can either be built from
  - Two individual 10GE links
  - Two physical port channels
- NX-OS can theoretically support up to 768 virtual port channels (vPC’s) per pair (what will be supported at FCS still to be determined)
- Nexus 5000 will support up to 12 Ethernet physical port channels per switch

In this example FEX is configured with two 20GE ‘Port Channel’ vPC uplinks