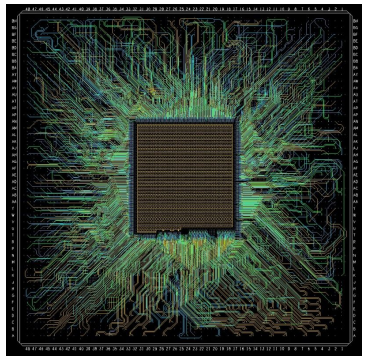
An update on Networking Standards

(The great thing about standards is there are so many to choose from)

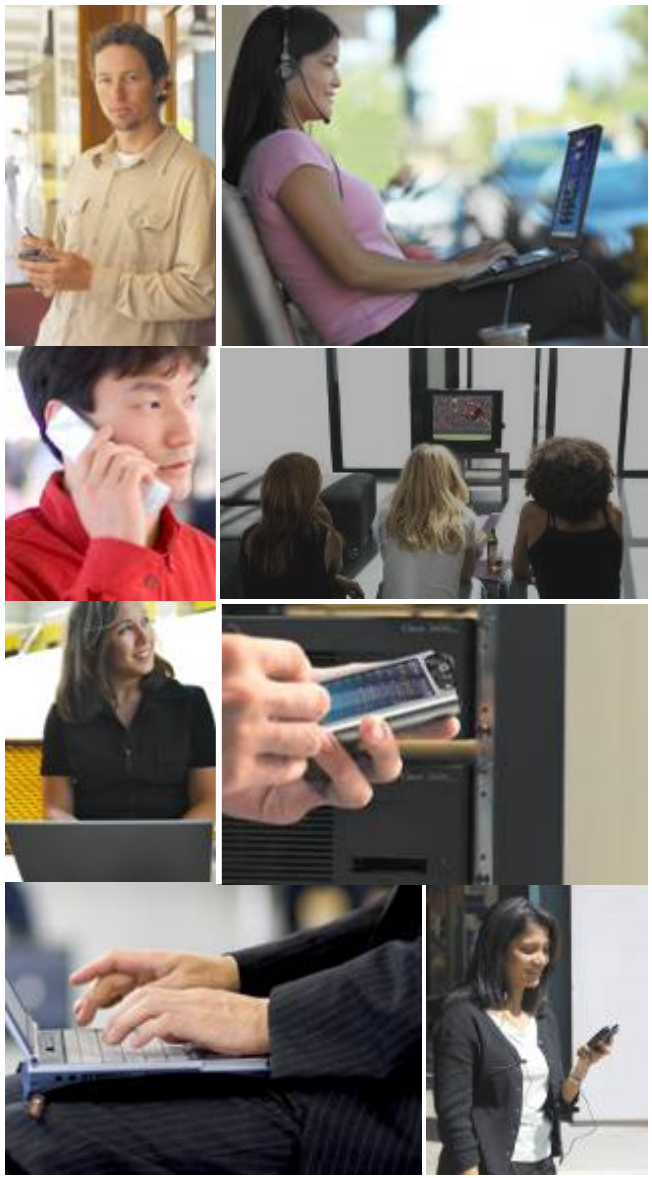
Lincoln Dale ltd@cisco.com
Distinguished Engineer
Data Center Switching Technology Group

Cisco Systems Inc.



Agenda

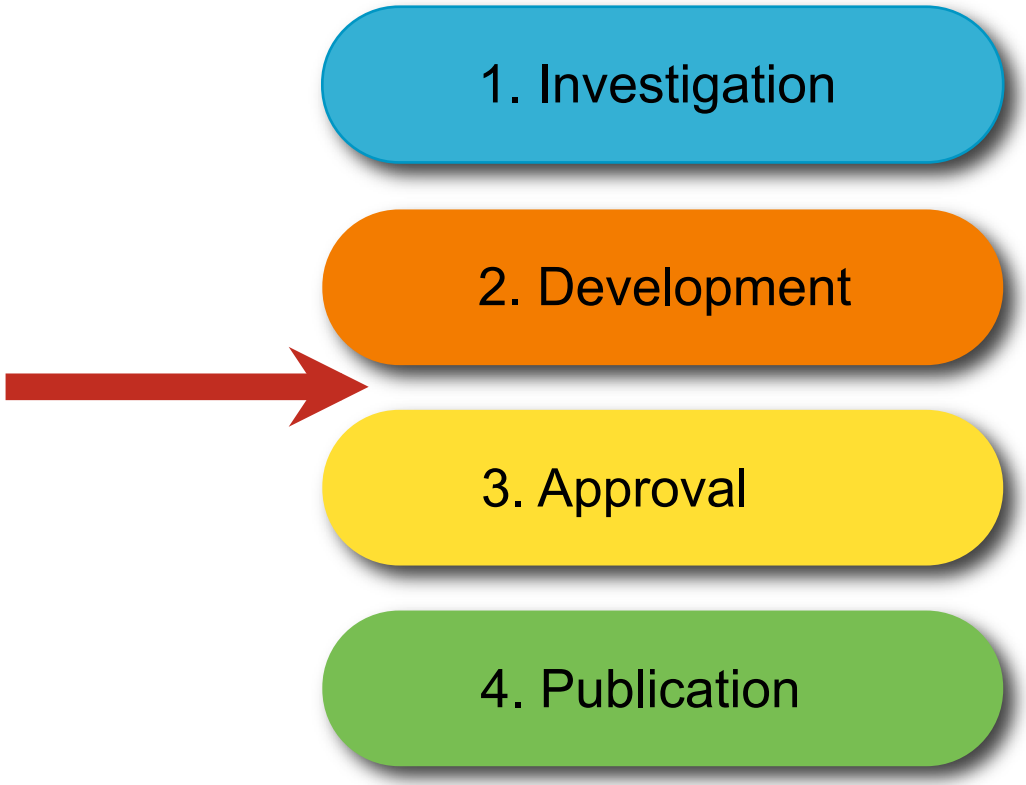
- **Introduction**
  - How standards processes work
  - State of network feeds and speeds
- **Evolving to 10G**
  - IEEE 802.3an 10GBaseT
  - Energy Consumption @ 10G
- **IEEE 802.3ba 40 GbE / 100 GbE**
- **Virtual machine awareness**
  - IEEE 802.1Qbg VEPA / Reflective Relay
  - IEEE 802.1Qbh vnTag / Port Extension
- **Merging SANs and LANs: FCoE**
- **IETF TRILL / Rbridges**



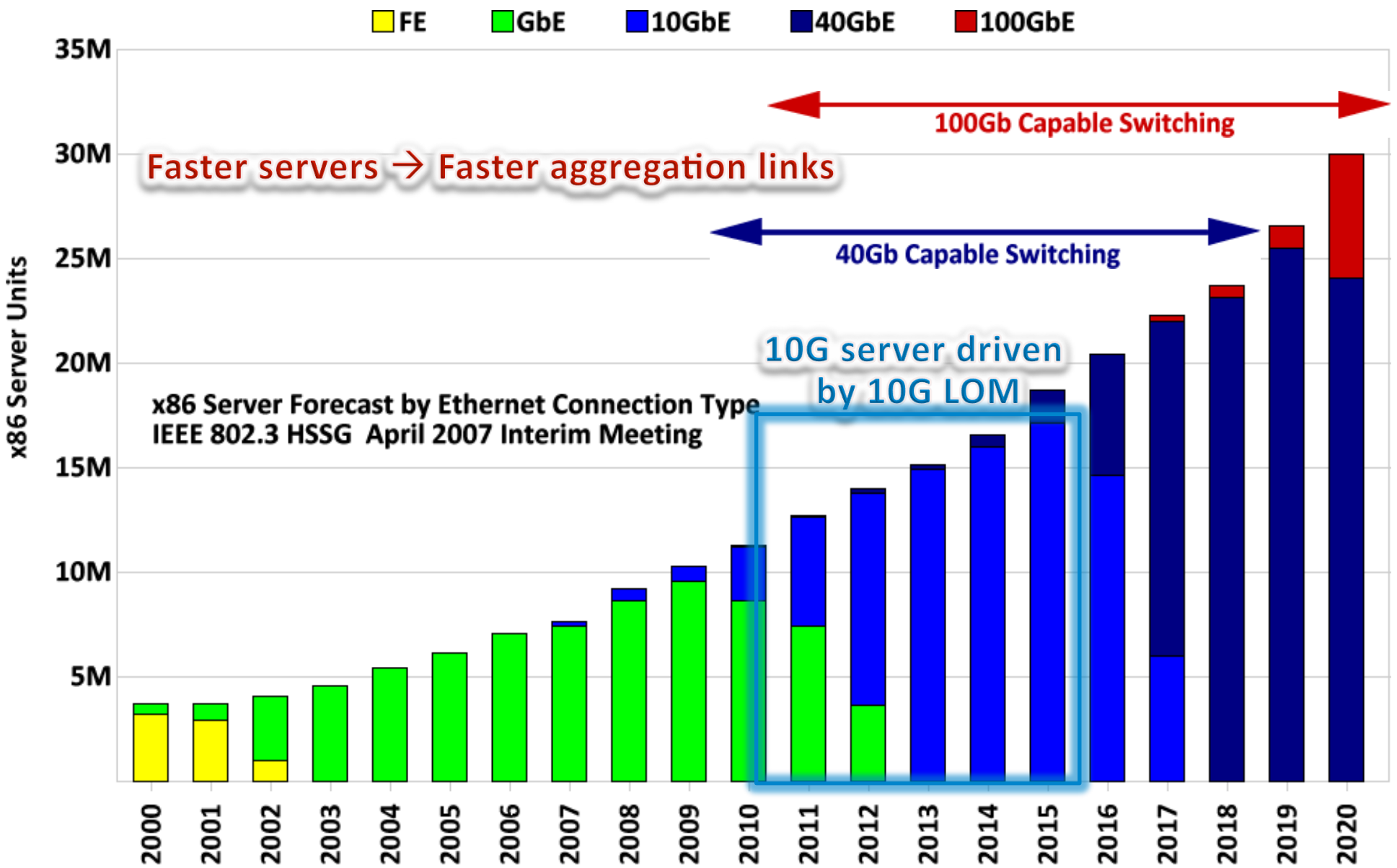
When Are Standards “Done?”

4 Phases of Standards Development

Standard is technically stable, a.k.a "Done," when it moves from Development to Approval phase

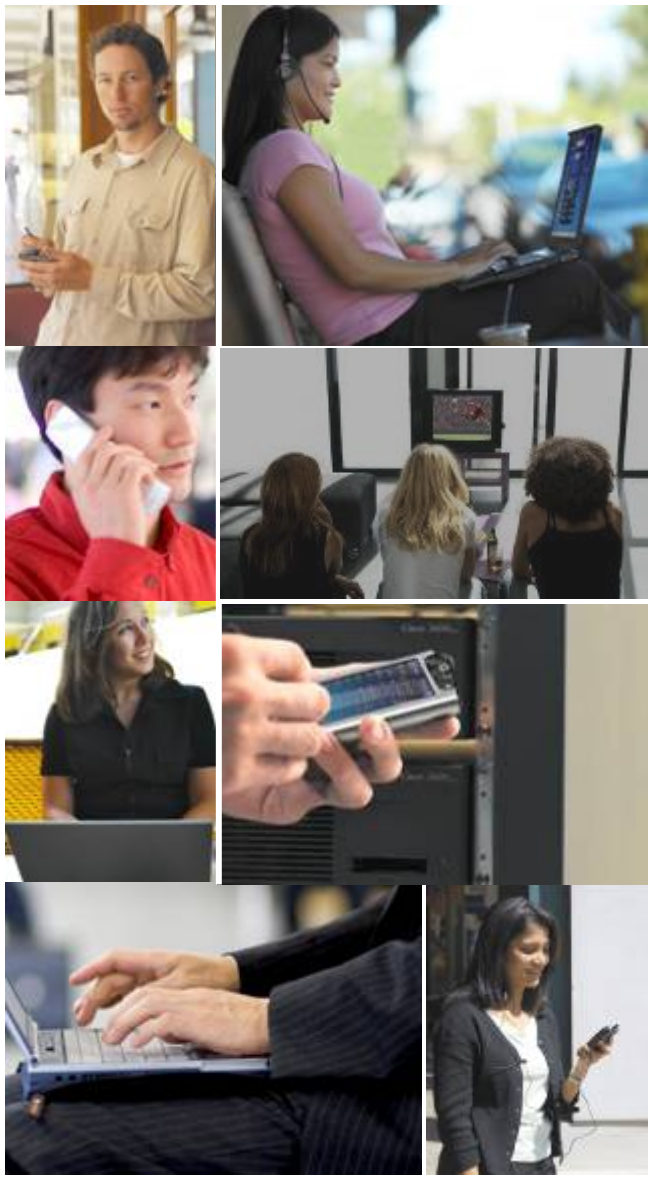


High-Speed Ethernet Server Adoption



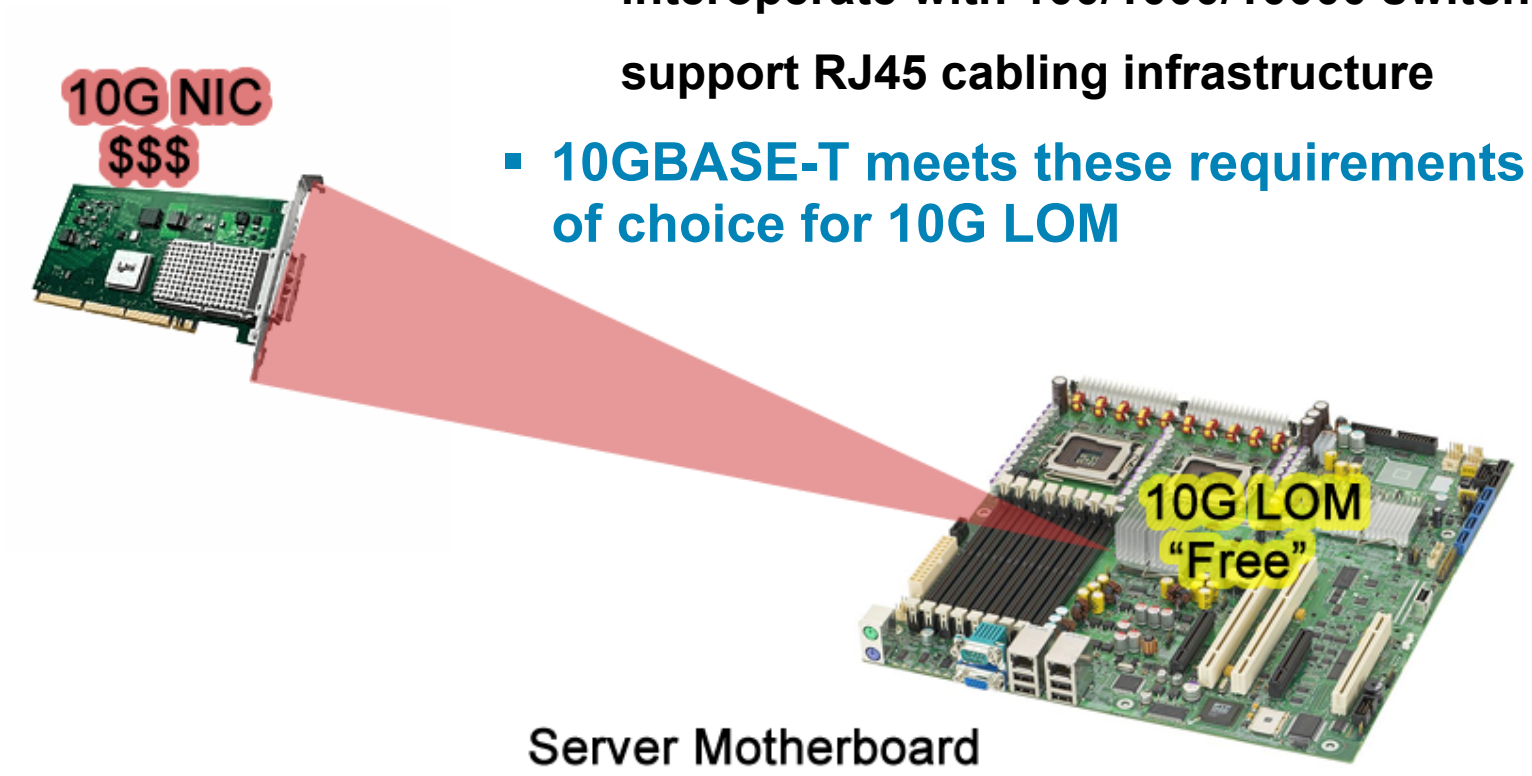
Agenda

- Introduction
  - How standards processes work
  - State of network feeds and speeds
- Evolving to 10G
  - IEEE 802.3an 10GBaseT
  - Energy Consumption @ 10G
- IEEE 802.3ba 40 GbE / 100 GbE
- Virtual machine awareness
  - IEEE 802.1Qbg VEPA / Reflective Relay
  - IEEE 802.1Qbh vnTag / Port Extension
- Merging SANs and LANs: FCoE
- IETF TRILL / Rbridges



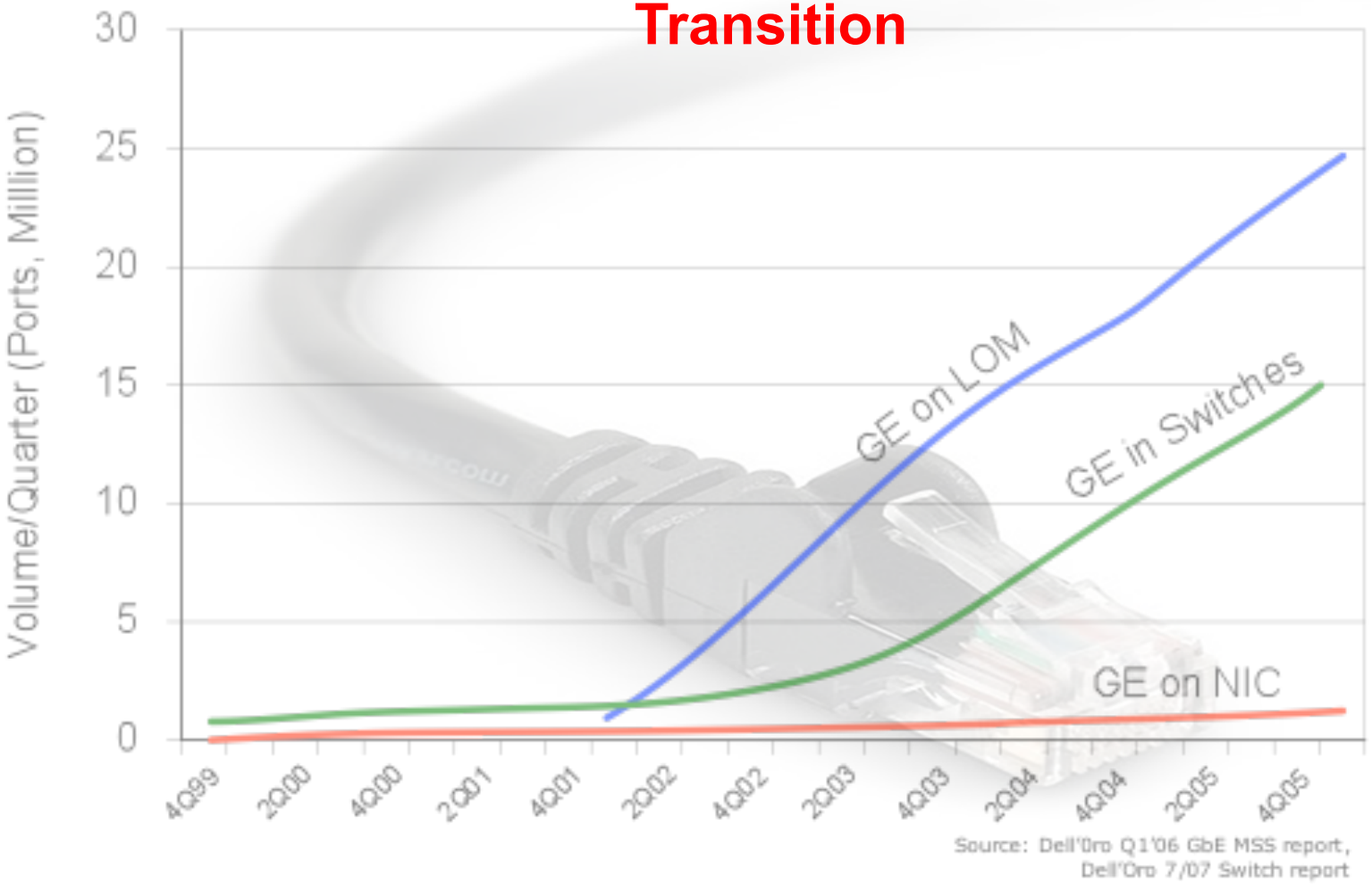
10G server adoption: from NIC to LAN-On-Motherboard (LOM)

- **LOM removes the cost barrier to adopt 10G on servers.**
- **Server vendors require LOM to be backward compatible, hence LOMs should support:**
  - **interoperate with 100/1000/10000 switches**
  - **support RJ45 cabling infrastructure**
- **10GBASE-T meets these requirements and is the PHY of choice for 10G LOM**



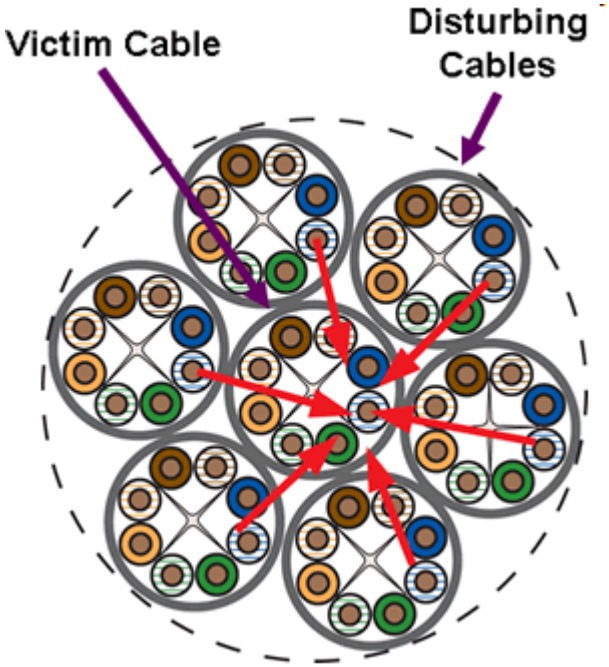
LOM Leads to 10G Server Volume Inflection Points

A look at the Gigabit Ethernet Transition



The ultimate challenge of 10GBASE-T: Alien X-Talk (AXT)

- **Undesired coupling of signal between adjacent cables**
- **Main electrical parameter limiting the performance of 10G**
- **Cannot be cancelled**
- **Can be prevented or mitigated by:**
  - **space (Cat6a solution)**
  - **shield (Cat6/Ca6a/Cat7 shielded solutions)**



Cat 5 and Cat 6 Twisted Pair Cabling At-a-Glance

Cat 5

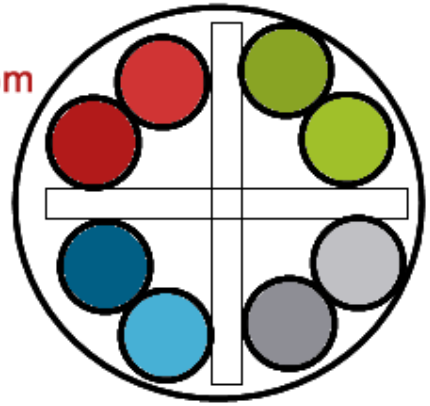
4 pairs, uncontrolled arrangement in cable

Not supported @ 10G

Cat 6

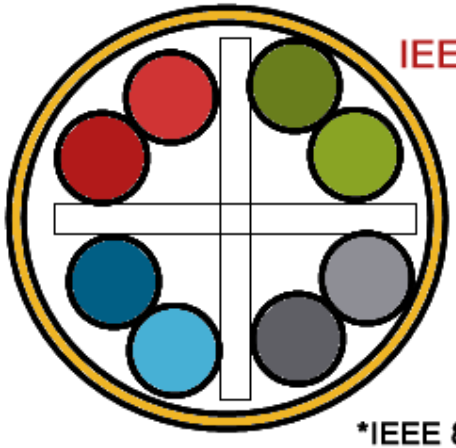
X-shaped spacer keeps pair-pair spacing controlled

IEEE: 55m



Cat 6 Shielded

Shield reduces alien noise



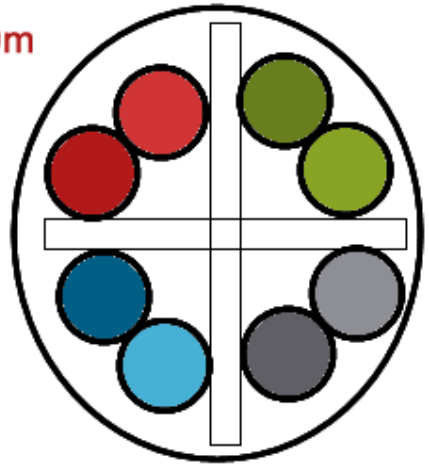
IEEE: 100m

*IEEE 802.3an Table 55-12
**Cat6a can also be shielded

Cat 6a**

Asymmetric spacer (egg-shape) prevents close packing of cables, reduces alien noise

IEEE: 100m



Twisted Pair Cabling For 10GBASE-T (IEEE 802.3an)

U/UTP (Old designation UTP)
Outer Unshielded/Inner Pairs Unschielded

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

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

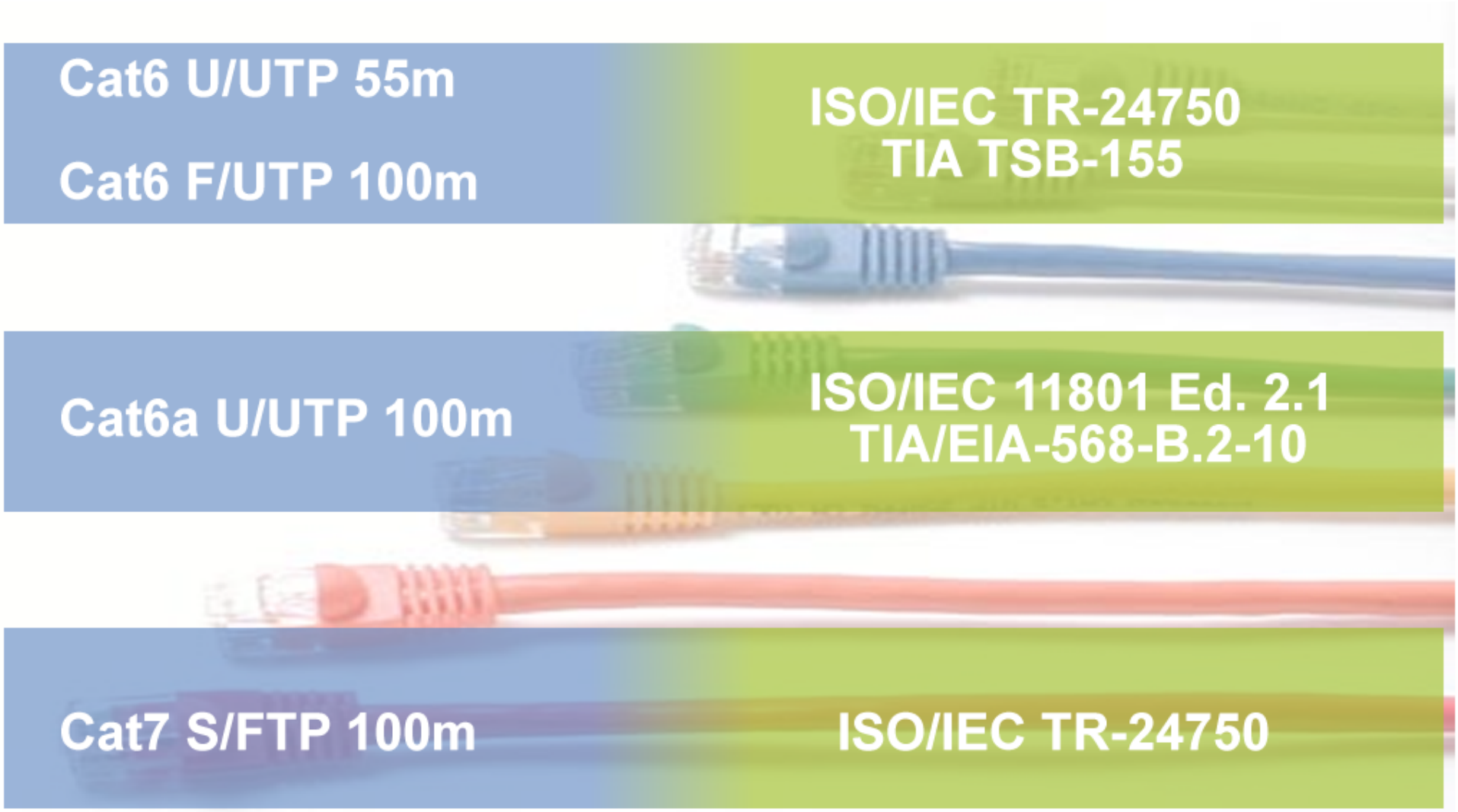
Cat 6a:
*100m 10GBASE-T
**largest diameter up to 0.354 in

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

Cat 7:
*100m 10GBASE-T
**Most expensive
***Smaller diamter than Cat6a
****Not popular in North America

Standard Cabling Referenced by IEEE 802.3an



10GbaseT IEEE 802.3an power consumption

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

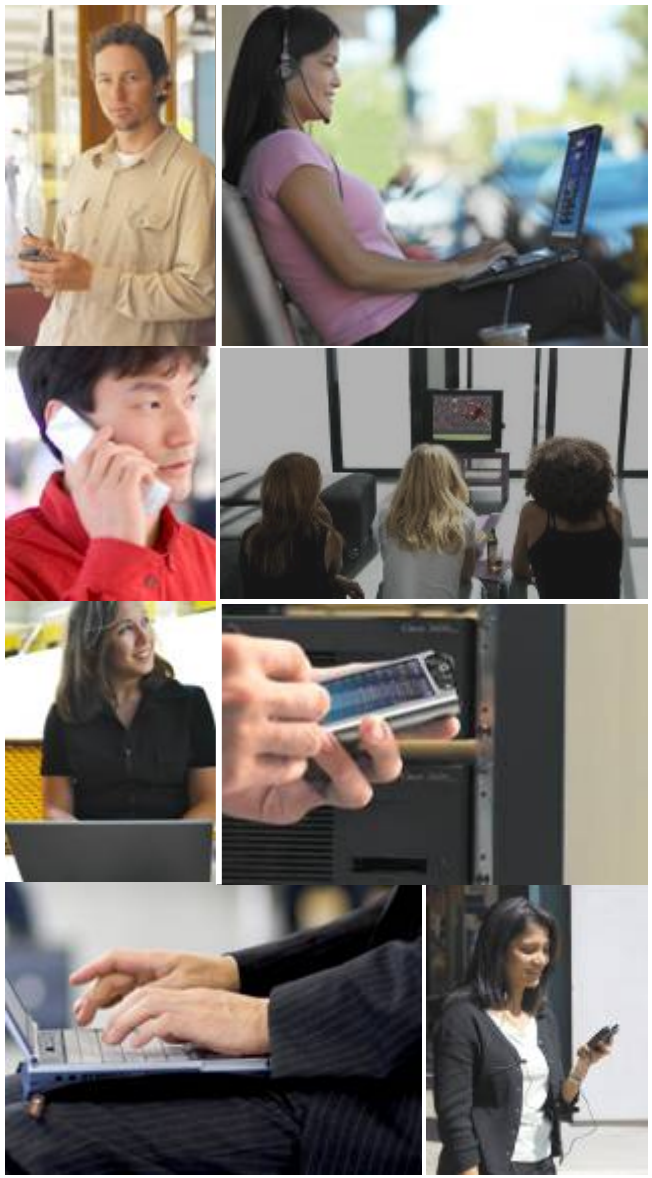
* Terminated cable ** Draft 3.0, Not Final *** As of 2008; expected to decrease over time 1

Introducing Energy Efficient Ethernet IEEE 802.3az

- EEE is a method to facilitate transition to and from lower power consumption in response to changes in network demand.
- In the process of being defined for these copper PHYs
  - 100BASE-TX (Full Duplex)
  - 1000BASE-T (Full Duplex)
  - 10GBASE-T
  - 10GBASE-KR
  - 10GBASE-KX4
- Uses Low Power Idle (LPI) to save energy
- EEE is an Energy Star requirement for PCs in 2010
- Likely 'network' will follow EEE requirements moving forward

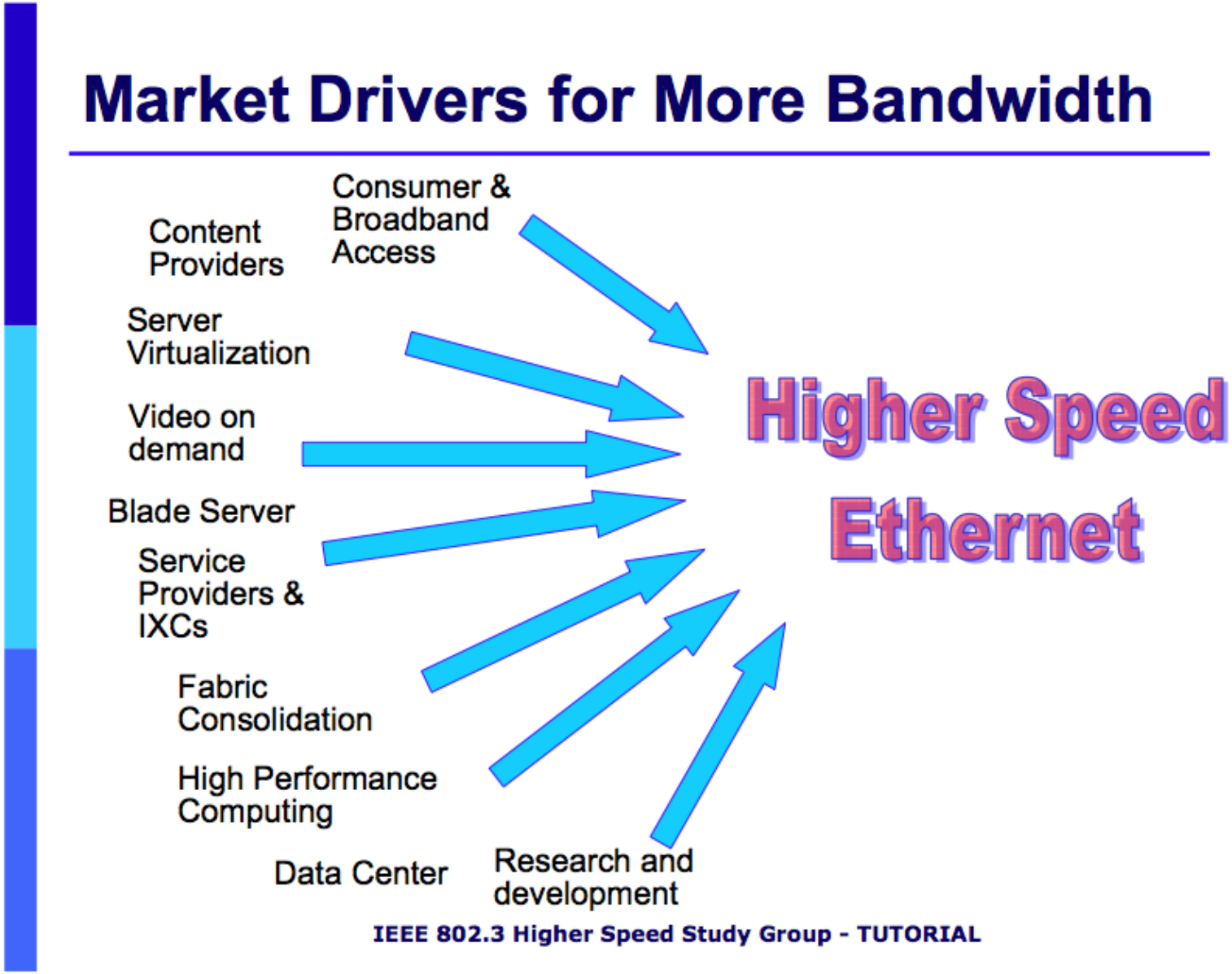
Agenda

- Introduction
  - How standards processes work
  - State of network feeds and speeds
- Evolving to 10G
  - IEEE 802.3an 10GBaseT
  - Energy Consumption @ 10G
- IEEE 802.3ba 40 GbE / 100 GbE
- Virtual machine awareness
  - IEEE 802.1Qbg VEPA / Reflective Relay
  - IEEE 802.1Qbh vnTag / Port Extension
- Merging SANs and LANs: FCoE
- IETF TRILL / Rbridges

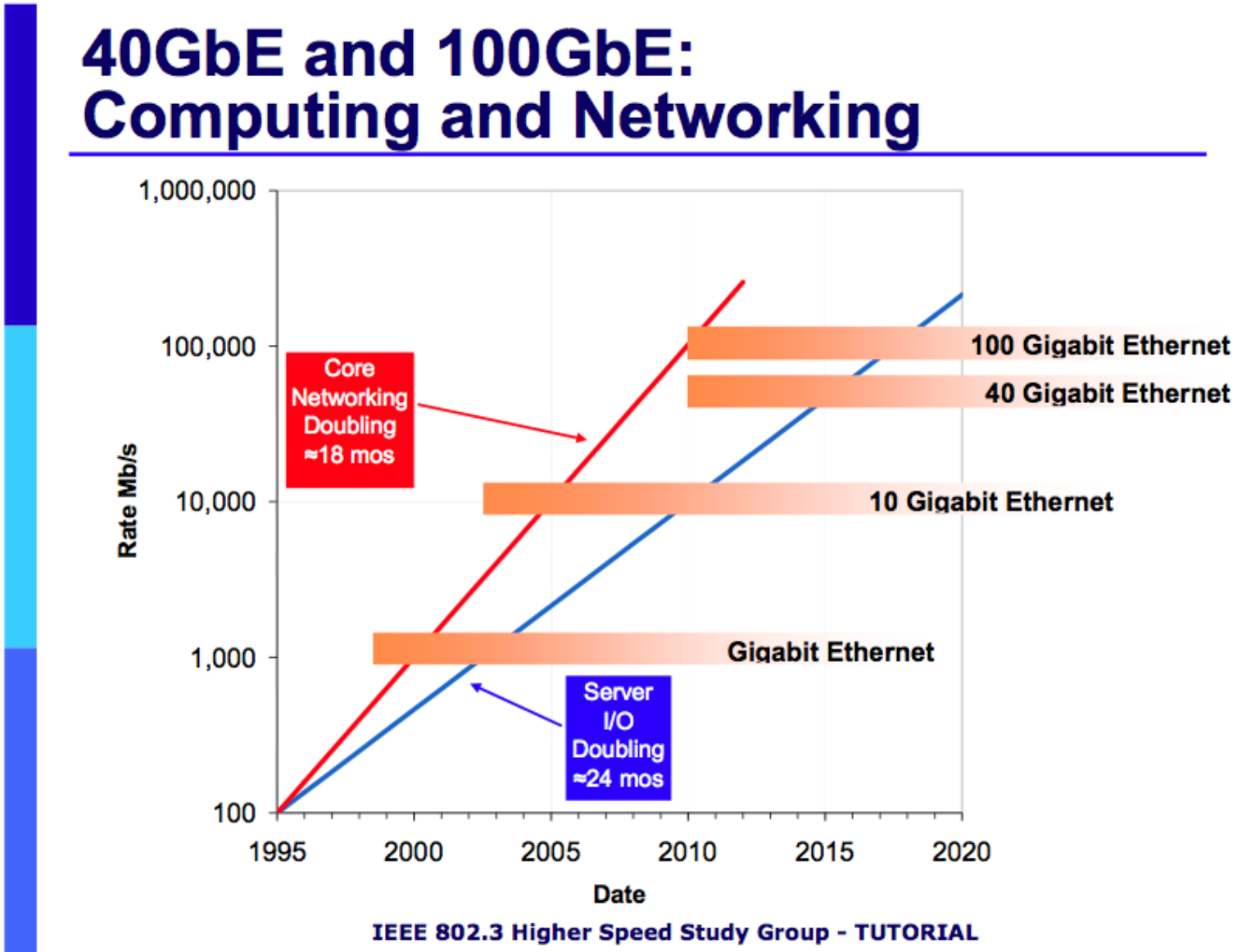


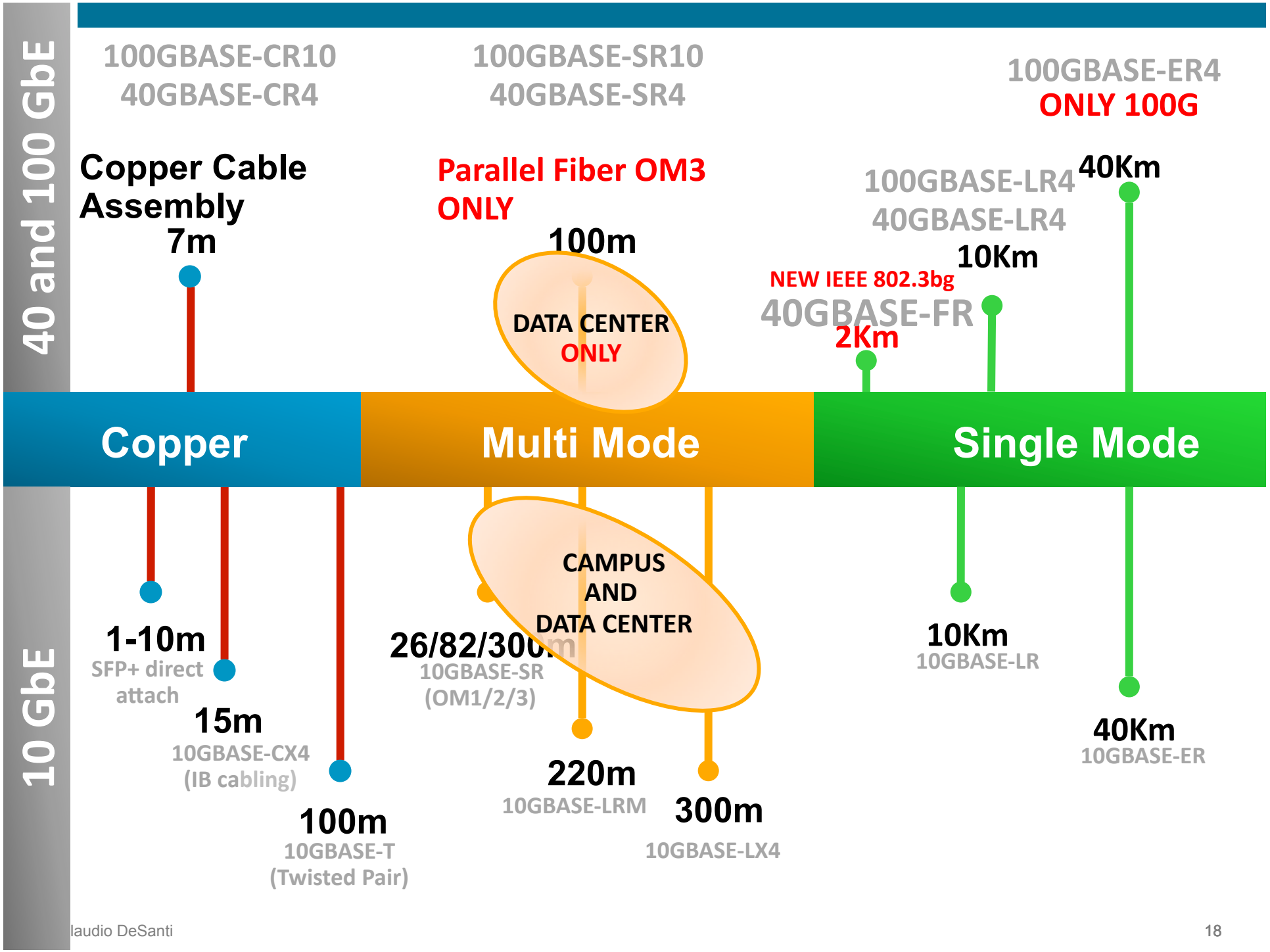

June 2010
IEEE 802.3ba
STANDARD COMPLETED
JUNE 2010
SUN MON TUES WED THURS FRI SAT
1 2 3 4 5
6 7 8 9 10 11 12
13 14 15 16 17 18 19
20 21 22 23 24 25 26
27 28 29 30
Note:

High-Speed Ethernet Drivers (Source IEEE)



High-Speed Ethernet Trends (Source IEEE)




40 and 100 GbE
100GBASE-CR10
40GBASE-CR4
100GBASE-SR10
40GBASE-SR4
100GBASE-ER4
ONLY 100G
Copper Cable Assembly
7m
Parallel Fiber OM3 ONLY
100m
DATA CENTER ONLY
100GBASE-LR4
40GBASE-LR4
10Km
40Km
NEW IEEE 802.3bg
40GBASE-FR
2Km
Copper
Multi Mode
Single Mode
10 GbE
1-10m
SFP+ direct attach
15m
10GBASE-CX4
(IB cabling)
100m
10GBASE-T
(Twisted Pair)
CAMPUS AND DATA CENTER
26/82/300m
10GBASE-SR
(OM1/2/3)
220m
10GBASE-LRM
300m
10GBASE-LX4
10Km
10GBASE-LR
40Km
10GBASE-ER

What 40G / 100G types will NOT be Available

- No Duplex over Multimode fiber
  - No “40G-SR” equivalent to 10G-SR
    (BUT MMF can be used – think ‘parallel fibers’)
  - No support for OM1 or OM2 with 40G-SR4
- No Cat6A or equivalent twisted pair solutions
- 40G beyond 10km reaches even on SM
- 100G beyond 40km – nothing agreed for >40km

These are all areas where vendors can/will develop MSA options:

- Market forces frequently drive more options (eg LRM, LX4)

Notes on 100G

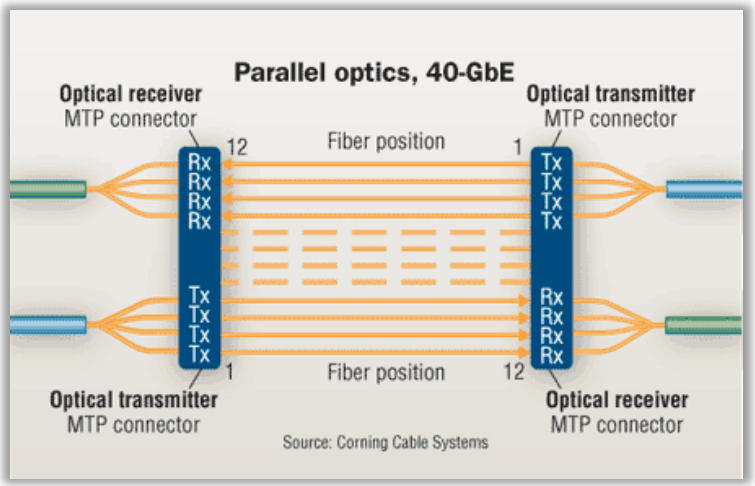
- 100G-SR10 is 10x 10G SR – based on 10G optics
  - Estimated pricing around 10x existing 10G-SR
- 100G-LR4 is 4x 25G lamda – based on 25G wavelength optics (10km reaches)
- 100G-ER4 is 4x 25G with long haul 25G wavelength optics
- 100G-CR10 is limited to 7m and is not intended for use outside of systems (compute or network)
- No standards for 2x50G or 10x10 DWDM packages
  - Can be considered as technology evolutions

Parallel Optics – 100G and 40G over Multimode OM3 fiber

40-GbE

4 Fiber Pairs

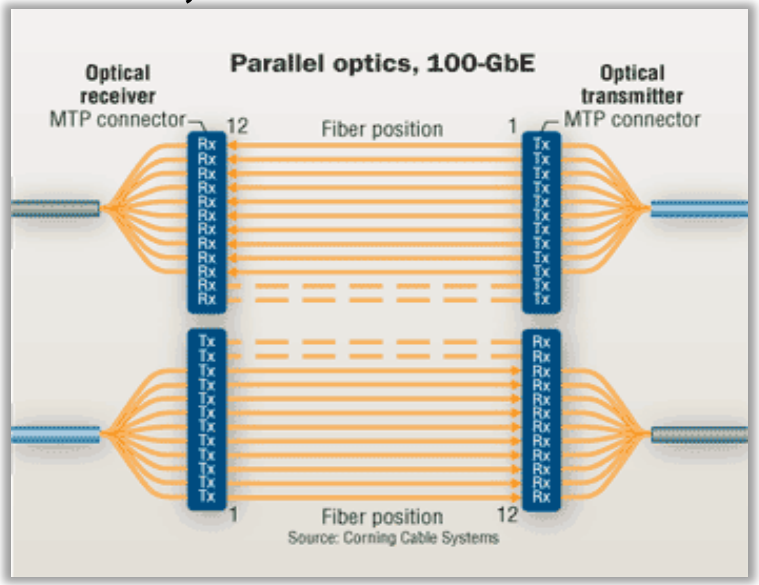
- MTP Connector with 12 pins
- 4 TX, 4 RX allows for 40G



100-GbE

10 Fiber Pairs

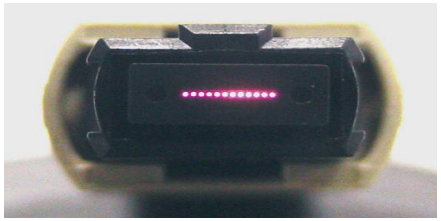
- MTP Connector with 2x12 pins
- 10 TX, 10 RX allows for 100G



40G/100G Multimode OM3 Fiber Array Cables

- MPO cables used for optical cabling trunks today
  - Custom-length cabling delivered with factory-installed connectors on both ends
    - MPO Structured Cabling trunks and LC patch leads
  - MPO Cables are plugged into the back of patch panels
- At 40GbE/100GbE “MPO” will plug directly into transceiver

MPO Plugs for 40G (12-fiber array connectors)



MPO Plugs for 100G (24-fiber array connectors)

Common MTP Fiber Patch Cables

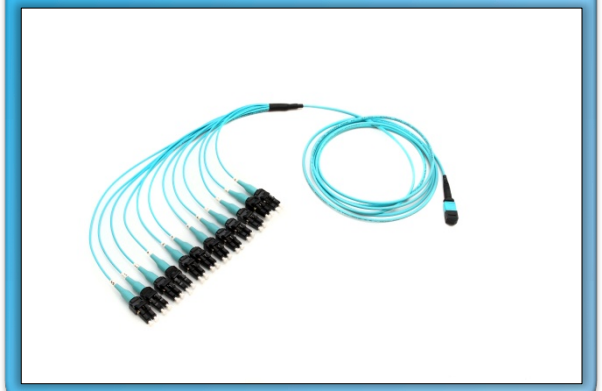
MTP-24 – 12x LC

Common Uses:

12 x 10GE – MTP-24

100G-SR10 MPO – LC fiber pairs

Allows connecting to MTP fiber from 10G LC ports today



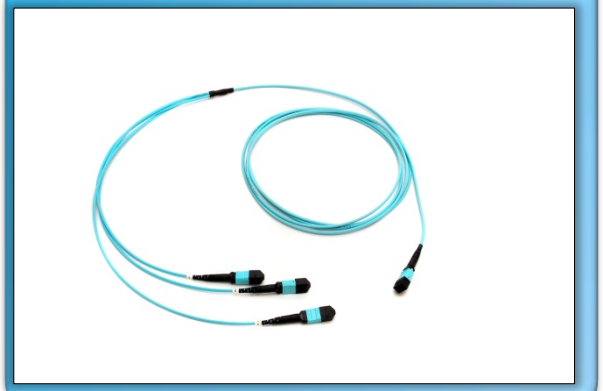
MTP-24 – 3xMTP-12

Common Uses:

Connect 3 x 40G to 1 MTP24

Trunk 1 100G into 3 x MTP12 where MTP24 does not exist

Harness MTP12 to MTP24 paths



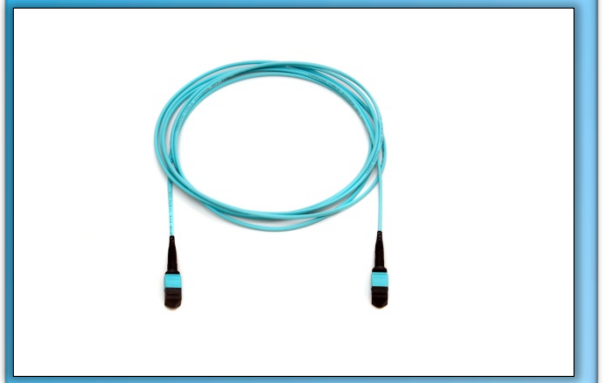
MTP-24 – MTP-24

MTP-12 – MTP-12

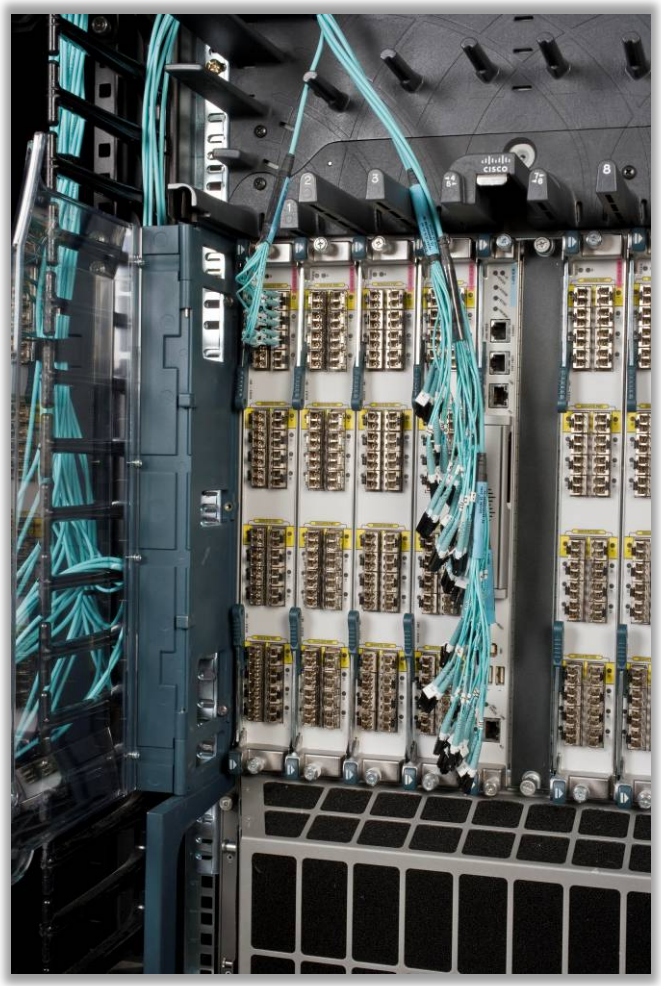
Common Uses:

100G to 100G (24)

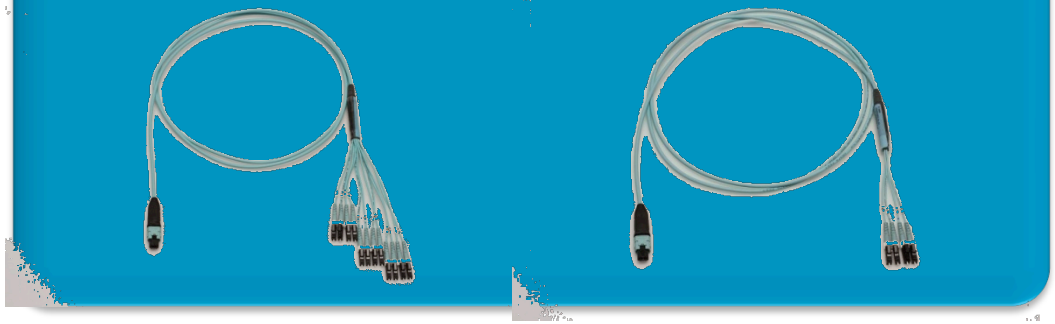
40G to 40G (12)



MTP to Duplex LC Harness



MTP to LC Duplex Harnesses



12-fiber
MTP to 6x Duplex LC

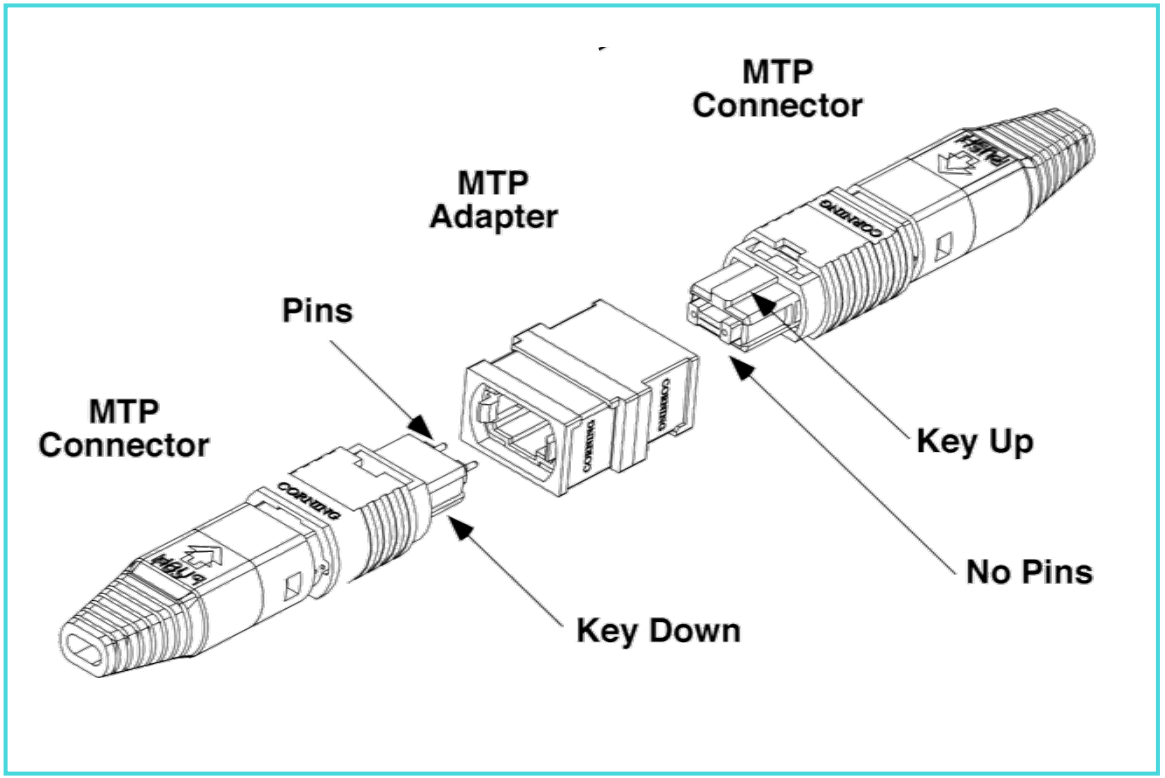
4-fiber
MTP to 2x Duplex LC

MTP to MTP Trunk cables

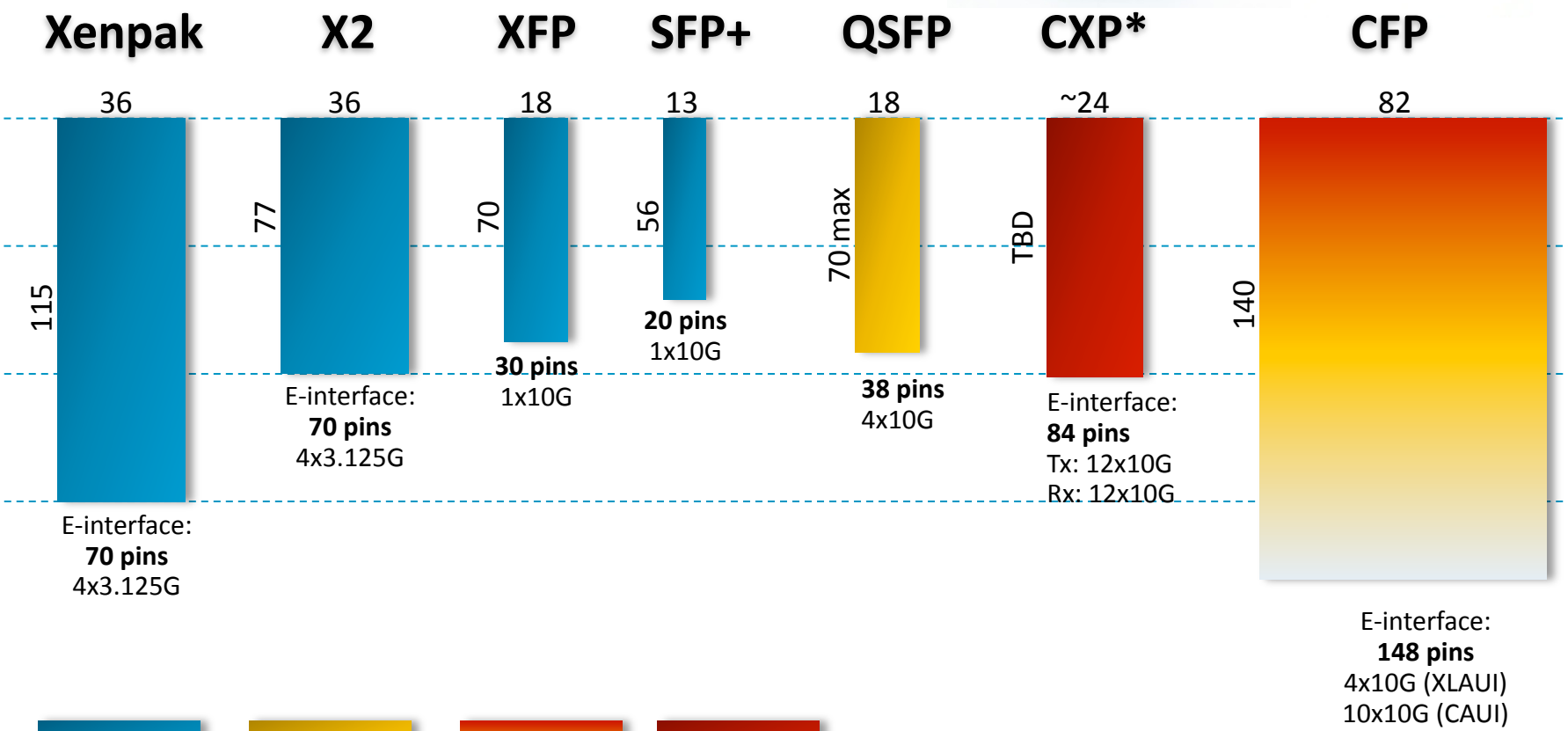
- Reduced cable bulk
- Fewer connectors
- Improved airflow
- Support for 40/100G
- Modular Solutions

MTP Plug & Play

- High-density 12-fiber connector
- Push-pull style latching
- Pinned to pinless mating



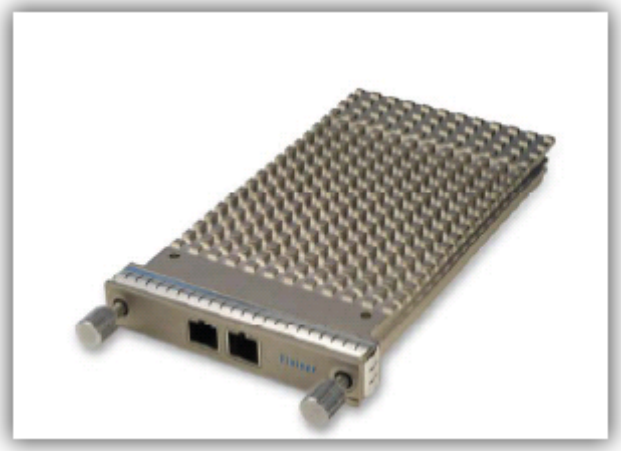
High-Speed Transceivers Form Factors



All units are in millimeters and round numbers

1st Gen 40GbE Transceivers

CFP



Applications:
Single Mode Fiber 10Km
Multi Mode OM-3 100m
Twinax Copper
“FourX” converter for 4x10GbE (SFP+)

Power Consumption:
Up to 8W @ 40GbE

QSFP

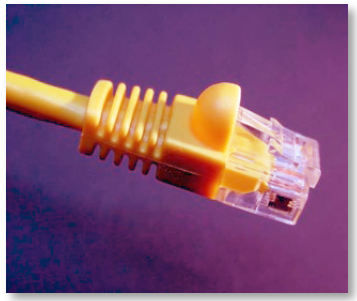
Applications:
Multimode Parallel Fiber
Twinax Copper
10 KM Single Mode (Future)

Power Consumption:
Up to 3.5W

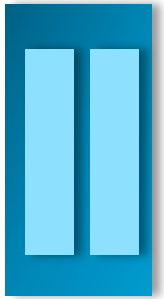
CFP to ease the transition to 40GbE

10/100/1000/10G 1 -> 10

RJ45



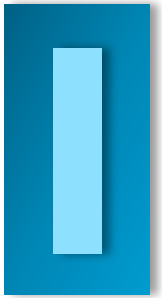
X2



SFP

10 -> 10

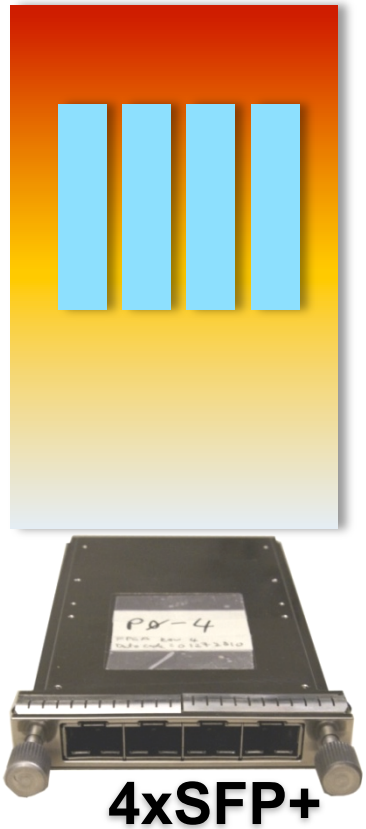
X2



SFP+

10 -> 40

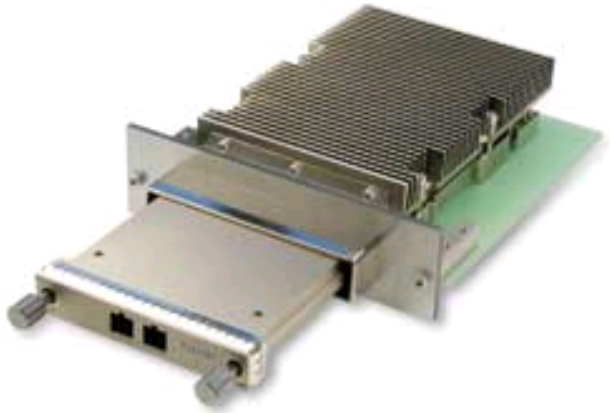
CFP (40G/100G)



4xSFP+

1st Gen 100GbE Transceivers

100GbE CFP requires "Riding HeatSink" SMF optimized



CFP features a new concept known as the riding heat sink, in which the heat sink is attached to rails on the host card and "rides" on top of the CFP, which is flat topped.

Applications:
Single Mode Fiber 10Km and 40Km
Multi Mode Fiber OM-3 100m
Power Consumption:
Up to 25W

CXP
MMF/Twinax optimized

CXP was created to satisfy the high-density requirements of the data center, targeting parallel interconnections for 12x QDR InfiniBand (120 Gbps), 100 GbE, and proprietary links between systems collocated in the same facility. The InfiniBand Trade Association is currently standardizing the CXP.

Applications:
Multi Mode Fiber OM-3 100m
Twinax copper assembly 7m
Power Consumption:
Up to 3.5W

1st Generation Std. 40G/100G Transceivers

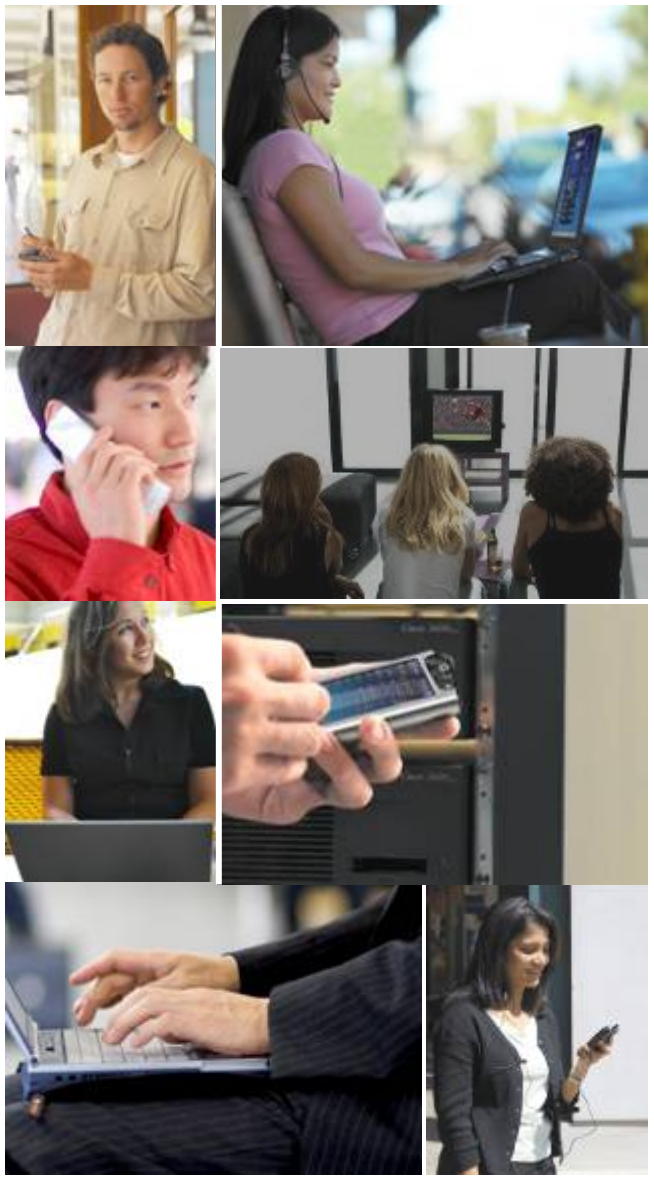
- 1st generation optics will be in multiple packages:
  - Single Mode requires more power which increases size / restricts density
  - Multi Mode optics use less power and so allow higher port density
- 2nd generation may rationalize these to fewer choices or increase overlap
- Do expect additional non-standard interconnect options as well in 2012+

Planned for 1st Generation Not Planned for 1st Generation

Media	Reach	Speed	CFP	QSFP	CXP
Single Mode	10Km	100G			
		40G		Possible	
	40Km	100G			
		40G	No Std	No Std	No Std
Multi Mode	100m	100G			
		40G			
Copper	3-7m	100G			
		40G			

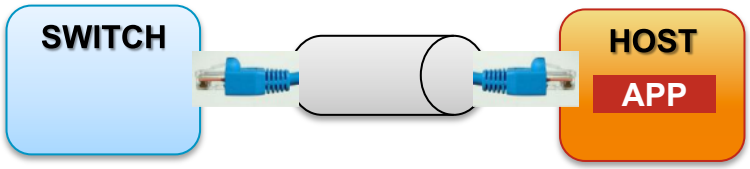
Agenda

- Introduction
  - How standards processes work
  - State of network feeds and speeds
- Evolving to 10G
  - IEEE 802.3an 10GBaseT
  - Energy Consumption @ 10G
- IEEE 802.3ba 40 GbE / 100 GbE
- Virtual machine awareness
  - IEEE 802.1Qbg VEPA / Reflective Relay
  - IEEE 802.1Qbh vnTag / Port Extension
- Merging SANs and LANs: FCoE
- IETF TRILL / Rbridges

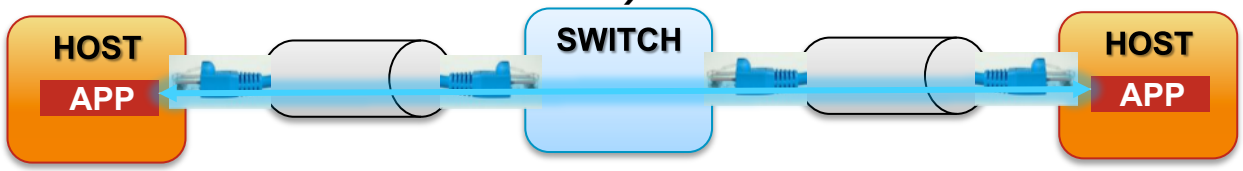


Background

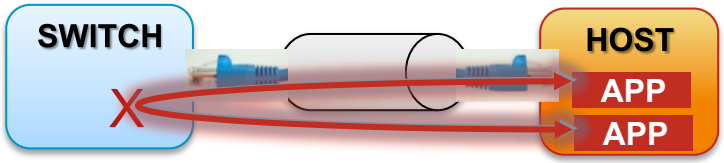
- **Before Server Virtualization… server typically ran one App…**



- **That App typically communicated with other Apps**
  (on other hosts on the same network)

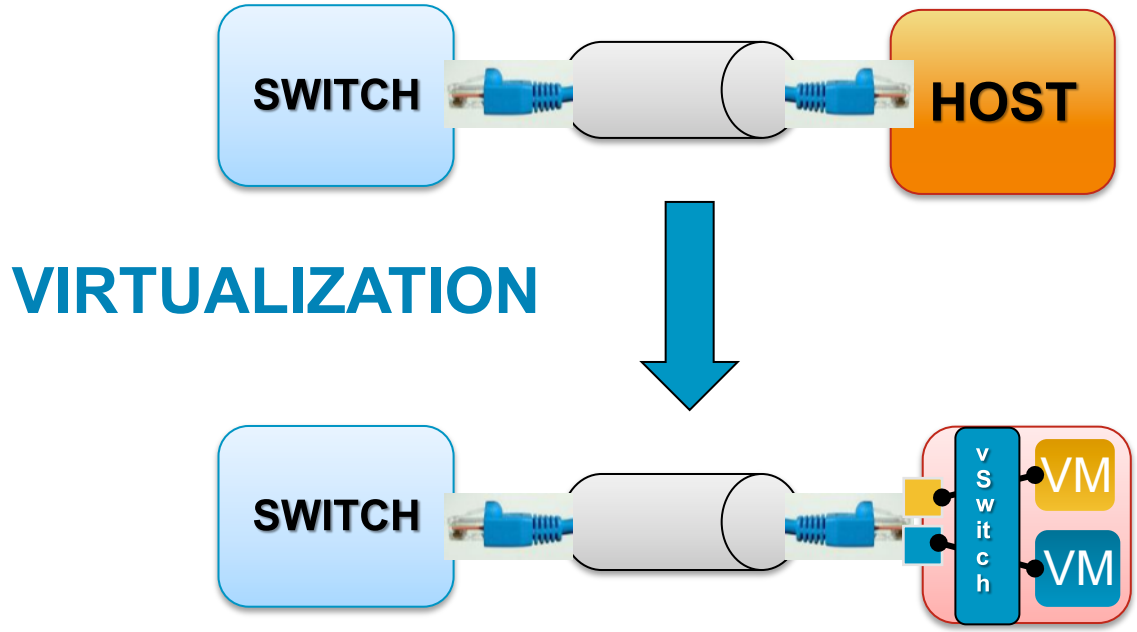


- **Server Virtualization came along**
  (which meant that you could run multiple Apps on the same server – cost savings, do more with less, etc.)



- **PROBLEM: Switches aren't allowed to forward a frame out the same port they received it on**
  (because this could result in a never-ending loop)

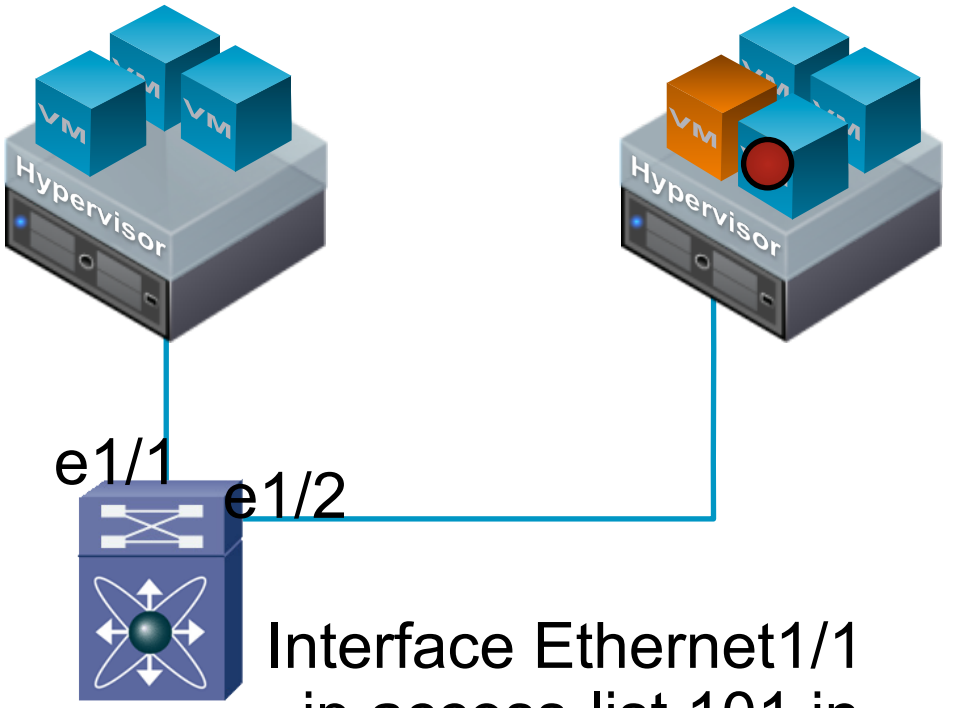
Datacenter Evolution



- **Difficult to correlate network back to virtual machines**
- **Scaling globally depends on maintaining transparency while also providing operational consistency**

Server Virtualization Issues

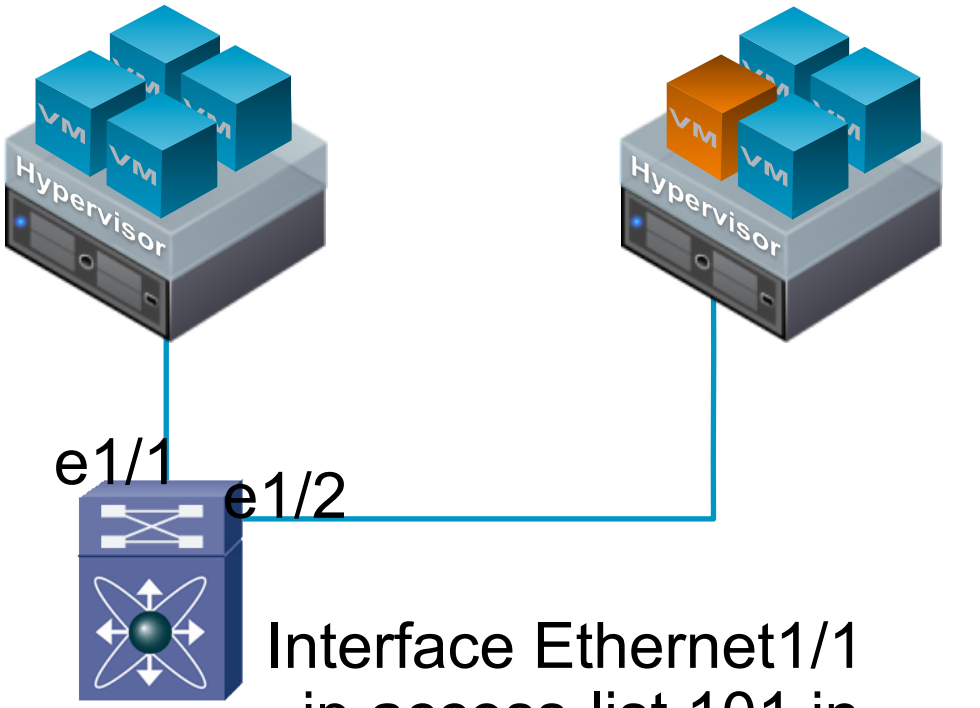
1 Impossible to View or Apply Network Policy to Locally Switched Traffic




```
Interface Ethernet1/1
  ip access-list 101 in
  ip flow monitor foo input
```

Server Virtualization Issues

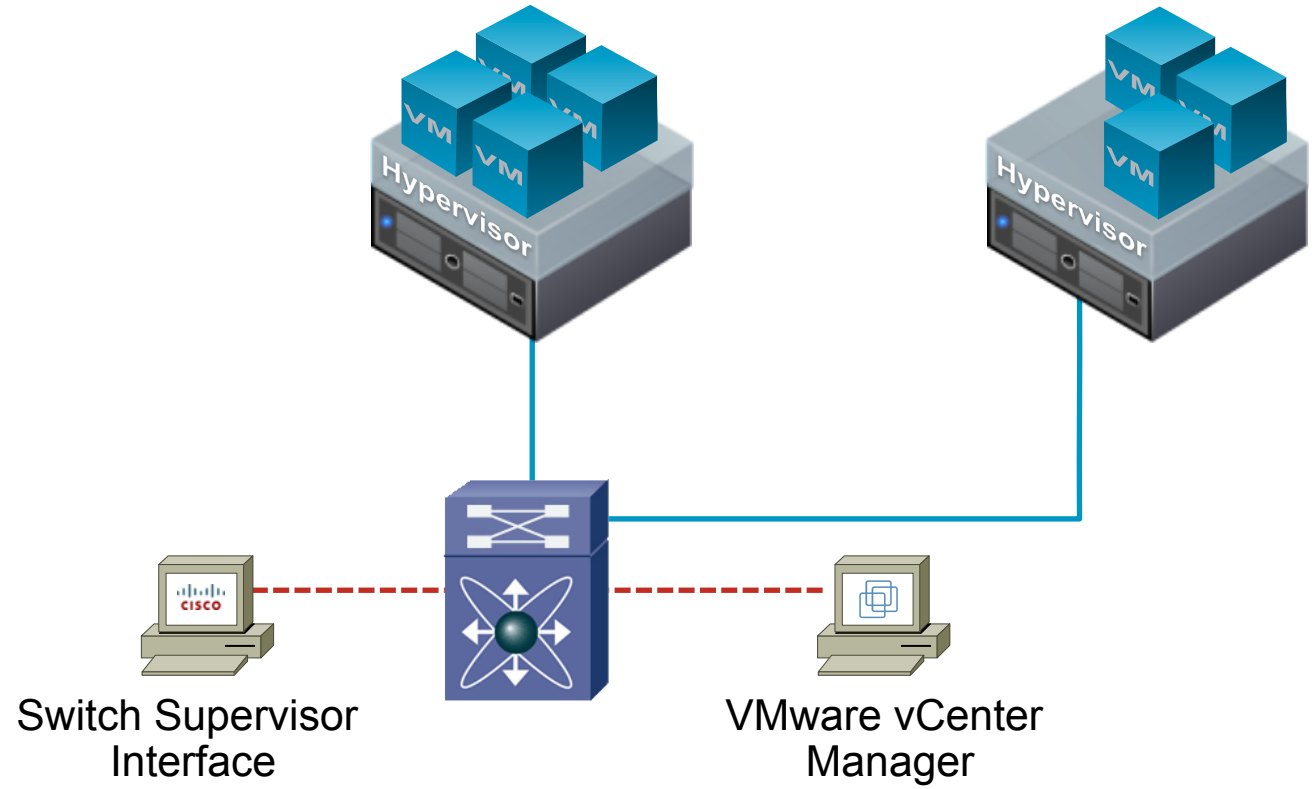
2 vMotion Moves VMs Across Physical Ports—the Network Policy Should Follow




```
Interface Ethernet1/1
  ip access-list 101 in
  ip flow monitor foo input
```

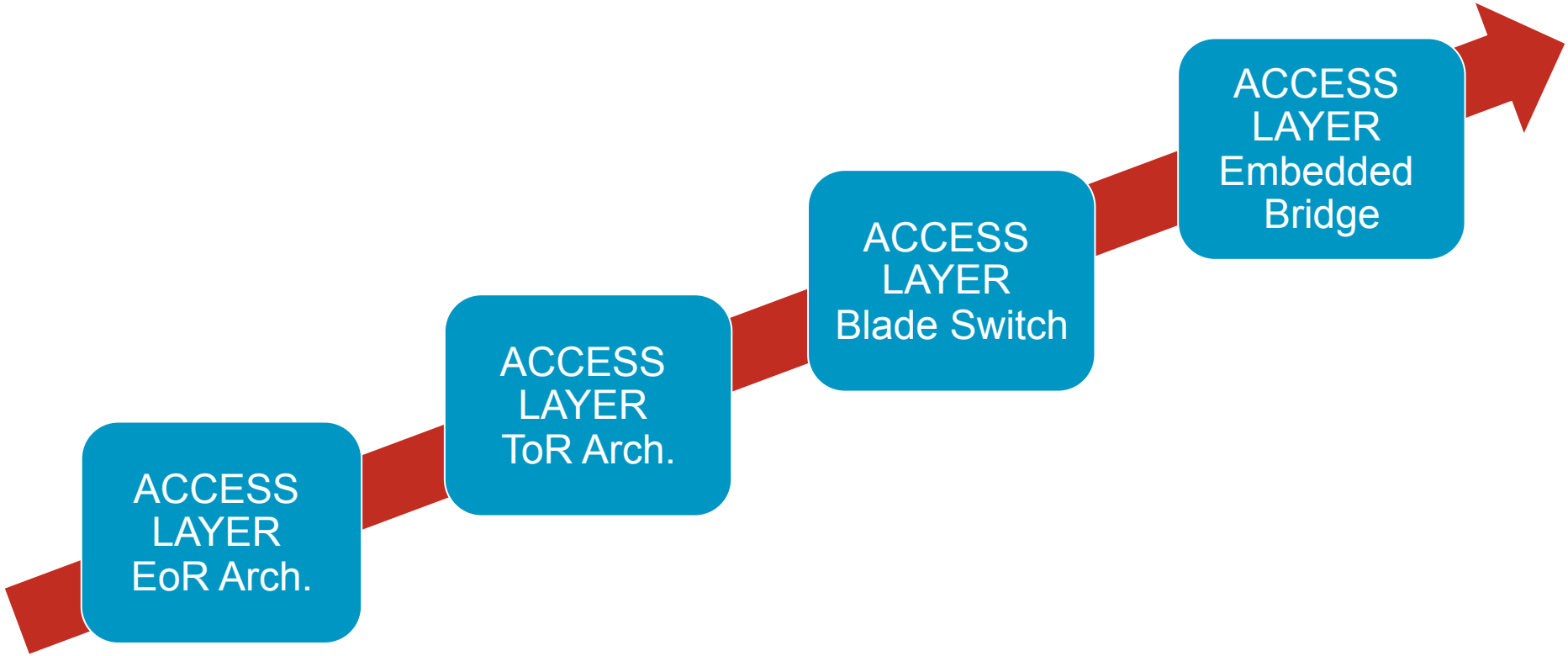
Server Virtualization Issues

3 Need Shared Nomenclature Between Network Admin and Server Admin

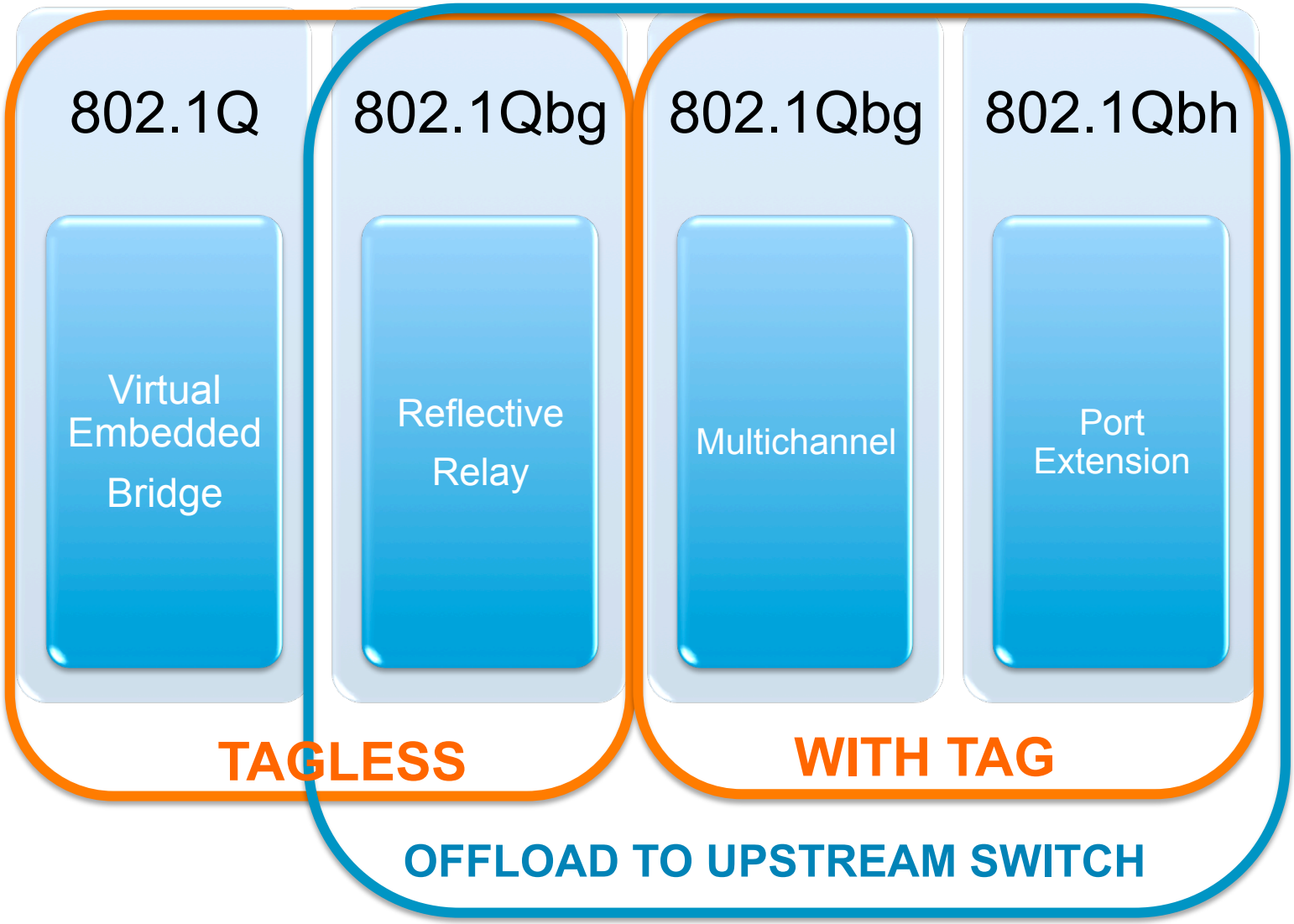


Server Virtualization Issues

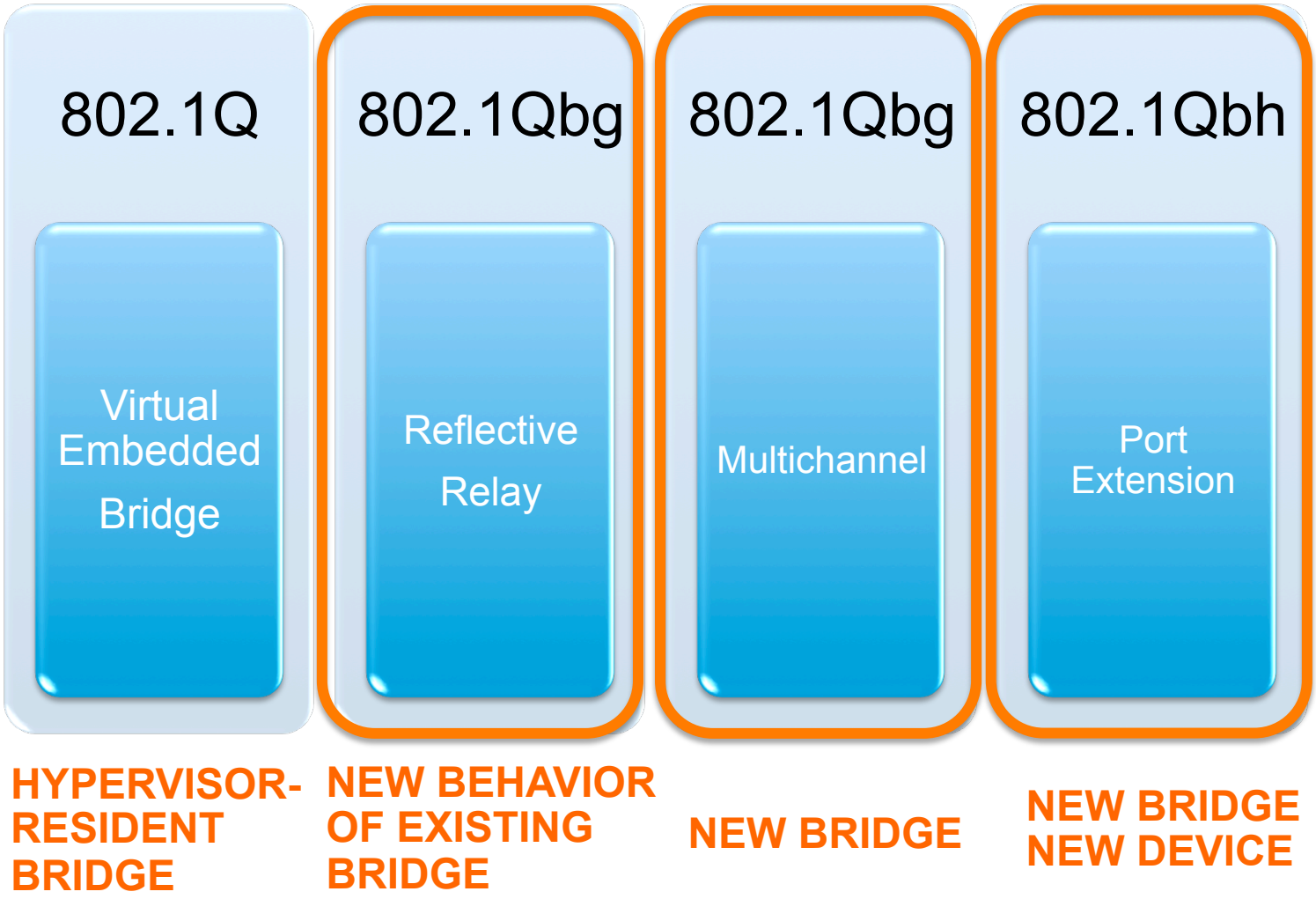
4 Proliferation of Management Points & new network devices



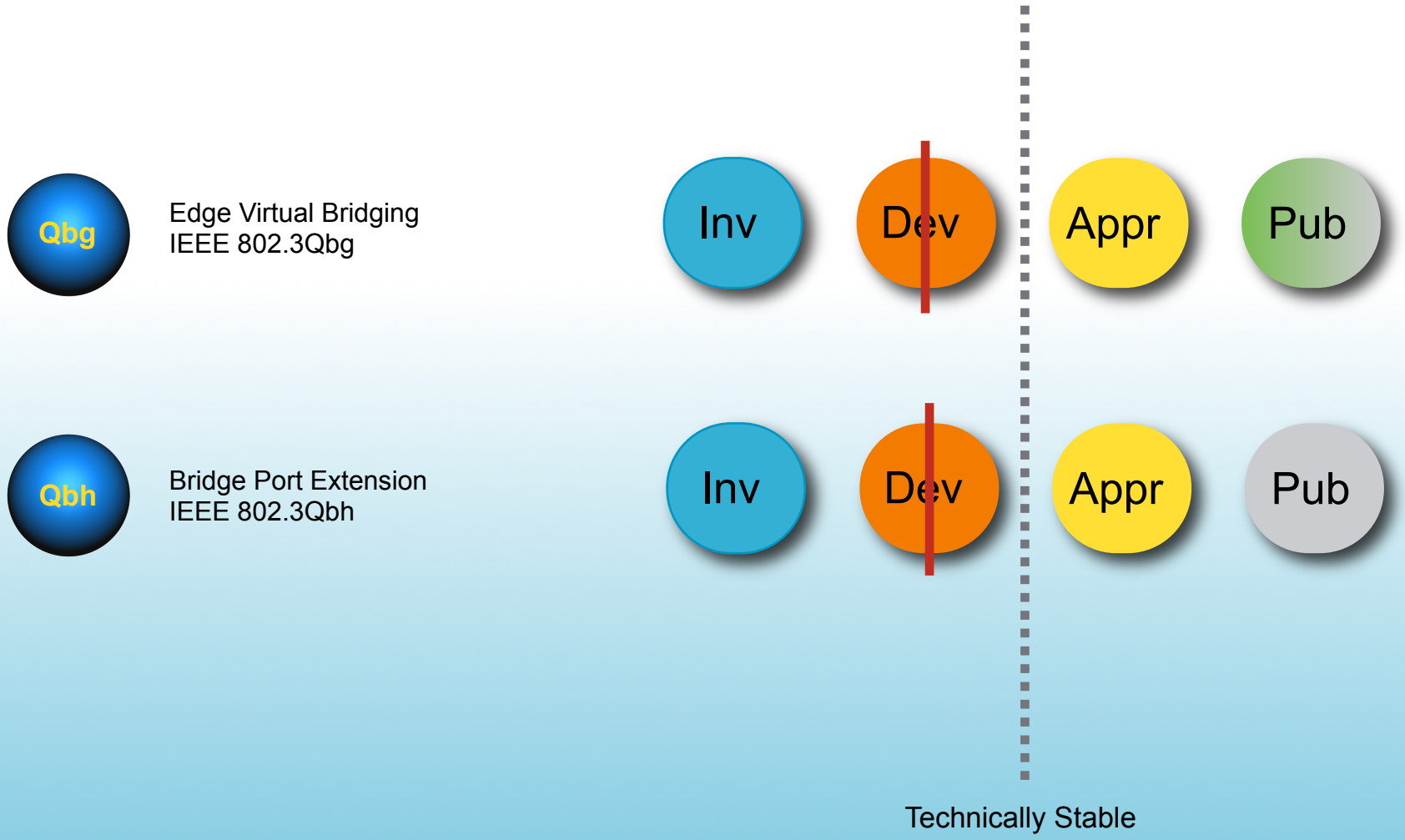
Virtual Networking Standards Components



Virtual Networking Standards Components

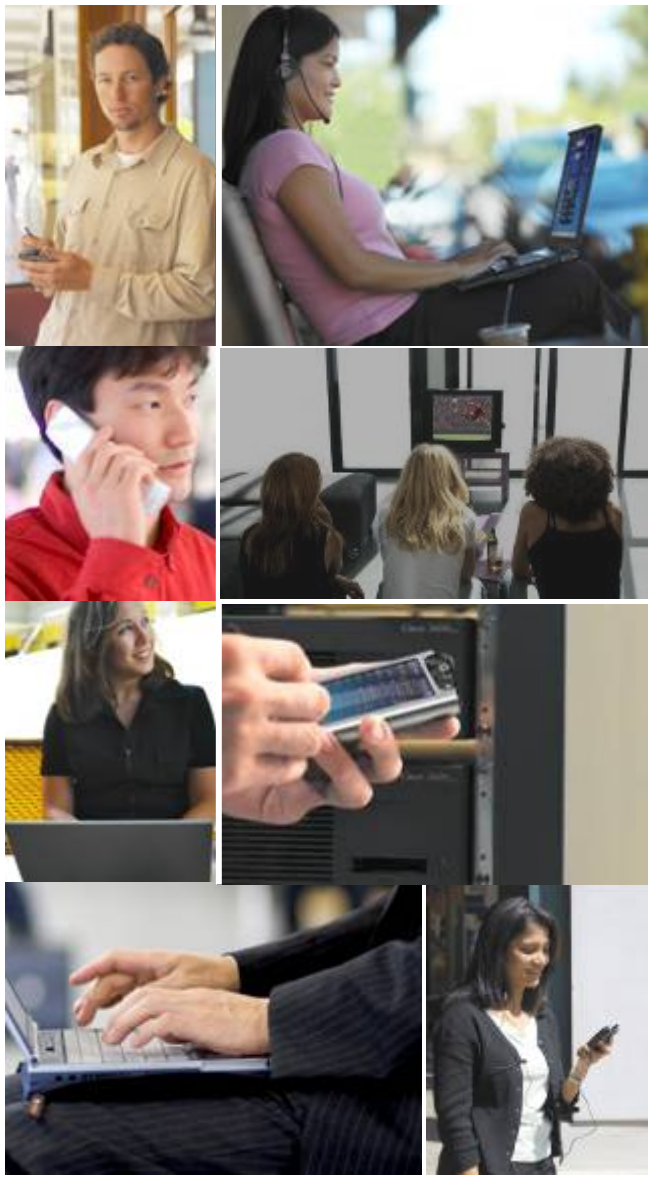


Virtual Bridging Status



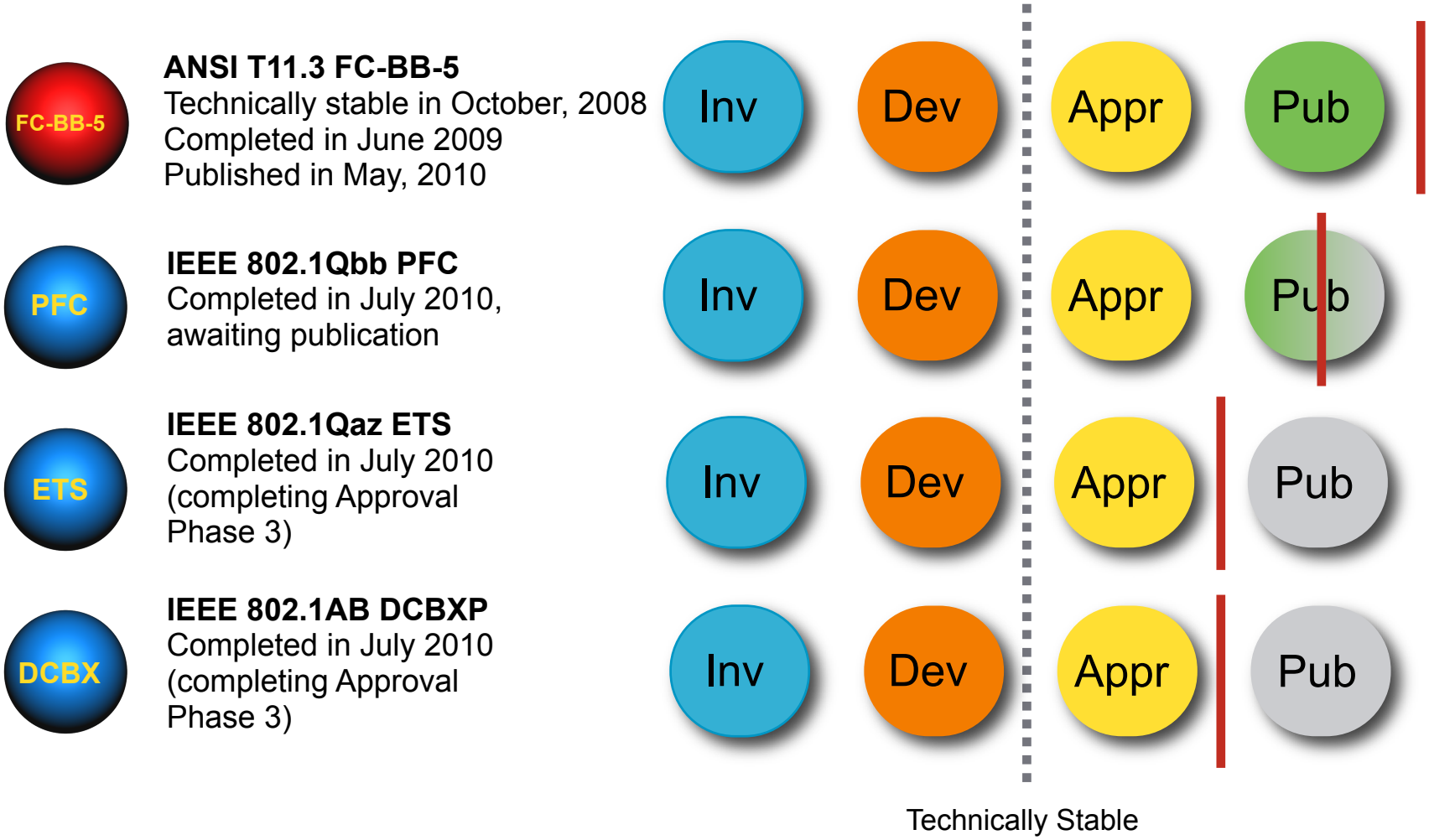
Agenda

- Introduction
  - How standards processes work
  - State of network feeds and speeds
- Evolving to 10G
  - IEEE 802.3an 10GBaseT
  - Energy Consumption @ 10G
- IEEE 802.3ba 40 GbE / 100 GbE
- Virtual machine awareness
  - IEEE 802.1Qbg VEPA / Reflective Relay
  - IEEE 802.1Qbh vnTag / Port Extension
- Merging SANs and LANs: FCoE
- IETF TRILL / Rbridges



Status of the Standards

All Standards for FCoE Are Technically Stable



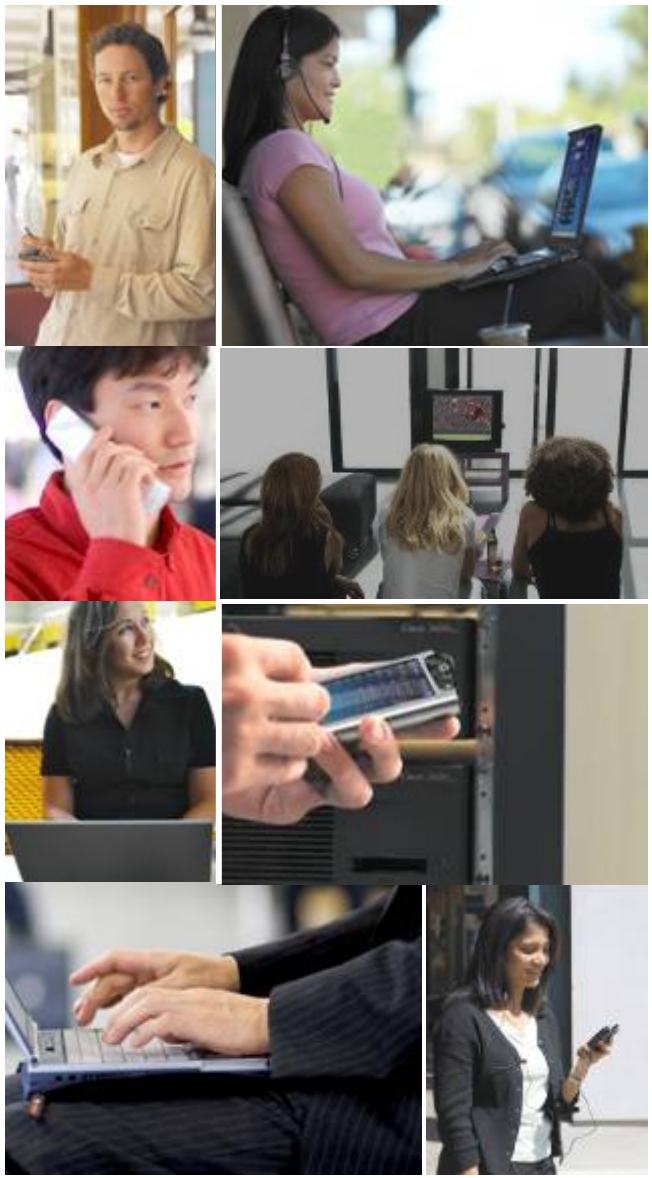
The FCoE Standard

- **T11 FC-BB-5**
  - **Supports the full Fibre Channel fabric functionality**
    - Including multi-hop FCoE
  - **Working Group established in June 2007**
  - **Technically stable in October 2008**
  - **Completed (forwarded for publication) in June 2009**
  - **Published by ANSI in May 2010**

- **T11 FC-BB-6**
  - **Developing additional FCoE functionality**
  - **Not needed for current deployments**
  - **Working Group established in August 2009**

Agenda

- Introduction
  - How standards processes work
  - State of network feeds and speeds
- Evolving to 10G
  - IEEE 802.3an 10GBaseT
  - Energy Consumption @ 10G
- IEEE 802.3ba 40 GbE / 100 GbE
- Virtual machine awareness
  - IEEE 802.1Qbg VEPA / Reflective Relay
  - IEEE 802.1Qbh vnTag / Port Extension
- Merging SANs and LANs: FCoE
- IETF TRILL / Rbridges



TRILL / Rbridges / "Cisco Fabricpath"

- **Network Address Scheme: Flat → *Hierarchical***

  Additional header is required to allow L2 "Routing" instead of "Bridging"

  Provide additional loop-prevention mechanism like TTL

- **Address Learning: Data Plane → *Control Plane***

  Eliminate the needs to program all MACs on every switches to avoid flooding

- **Control Plane: Distance-Vector → *Link-State***

  Improve scalability, minimize convergence time, and allow multipathing inherently

- **END RESULT: Faster, Smaller, Better, Cheaper**

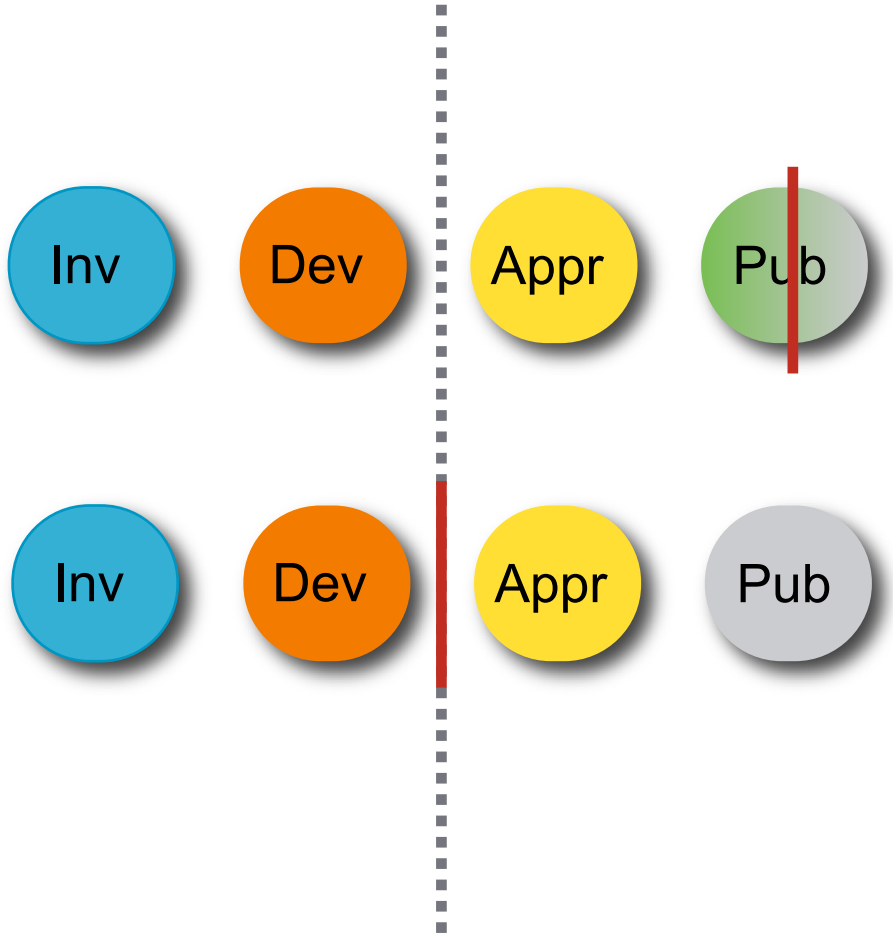
TRILL Status

Base proto

Completed in March 2010, awaiting publication. This is the part of the protocol affecting the hardware.

IS-IS Exts

Stable since long time, entering now the approval phase. This is control plane (i.e., software) only




CISCO