A Light Reading Webinar

Putting 40/100G To The Test

Wednesday, October 28, 2009

Hosted by

Andy Bray
Independent Analyst

Sponsored by:

SPIRENT

LIGHT READING

About Spirent Communications

Spirent Communication is a….

- Global test vendor with solutions for the lab and into live networks
- Next Generation Technology coverage from Data Centers, Virtualization, 3G/4G Wireless Technologies, Mobile Backhaul, Satellite/ Positioning and 40/100G Ethernet

Michael Lynge

- Ethernet Testing Evangelist
- Data Communications Test Industry for over 15 years
- Over the last 9 years at Spirent, has been the Product Manager of SmartBits, AX/4000, Avalanche and Spirent TestCenter products families

http://www.spirent.com

About H3C

http://www.h3cnetworks.com

- **One company (3Com) - Three Brands**
  - $1.3B networking specialist
  - H3C high-end networking brand launched in May 2009
  - #2 in global enterprise switch (ports) and routers (units) market share
  - #1 in China, securing over 30% of the Fortune 1000

- **David Law**
  - Engineering Consultant at 3Com / H3C
  - Specification, development of Ethernet products since 1989
  - IEEE 802.3 Ethernet Working Group
  - Editor for the IEEE 802.3u 100BASE-T Repeater Sub Task Force,
  - Chair of the IEEE 802.3z 1Gb/s Ethernet Mgt Sub Task Force
  - Editor of the IEEE 802.3ae 10Gb/s Ethernet Mgt Sub Task Force
  - Vice-Chair of the IEEE 802.3 Ethernet Working Group from 1996 until 2008
  - Presently the Chair of the IEEE 802.3 Ethernet Working Group

Agenda

- Market
- 40/100G Ethernet Basics
- How It Works
- Validating That It Works
- Summary
- Q&A

SPIRENT

LIGHT READING

Market

40/100G Drivers

- Computing
  - Driven by key technologies
    - Internal bus and memory performance
  - System throughput doubles approximately every 2 yrs
- Core networking
  - Driven by cumulative effect of more users, bandwidth, applications
  - Wide area, ISPs, IXs
    - Increase number of users
    - Increase in bandwidth available
      - xDSL, xPON, Cable, 3G
    - Increase number of applications
      - YouTube, Facebook, Netflix
  - Enterprise
    - Aggregation of multiple computing systems
  - Core throughput doubles approximately every 18 months

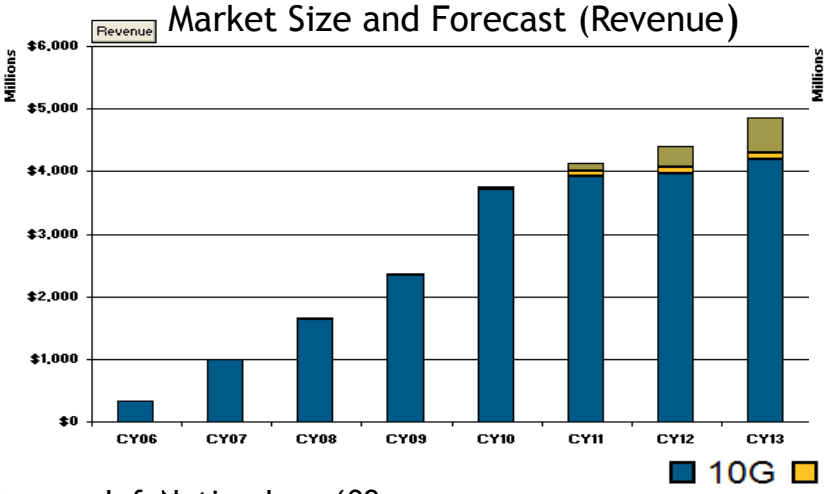
SPIRENT

LIGHT READING

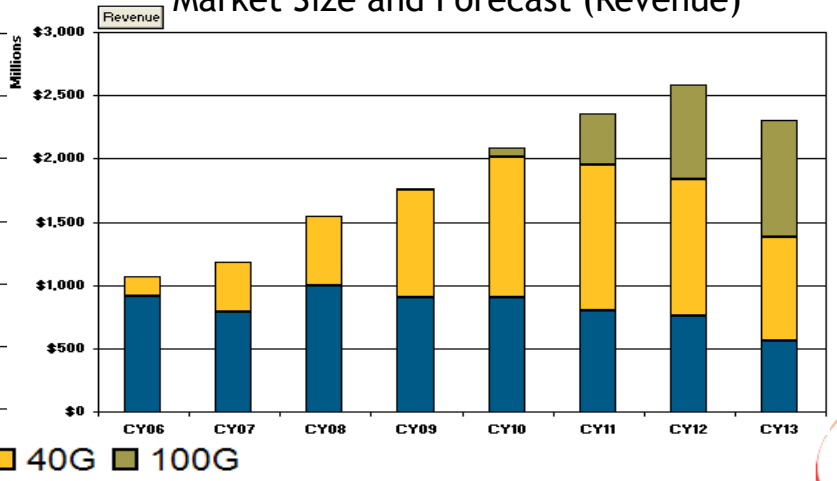
High Speed Ethernet at Service Provider

- Largest 40G application is router-router interconnects (OC-268/STM-256) with Comcast and AT&T as largest commercial deployments
  - 40G 42%+ CAGR for IP Core routers till 2013
  - 40GbE Ethernet viable alternative to 40G POS for short connection to OTN
- 100GbE SP Routing market takes of in 2010

10G/40G/100G **IP Edge Routers**
Market Size and Forecast (Revenue)



10G/40G/100G **IP Core Routers**
Market Size and Forecast (Revenue)

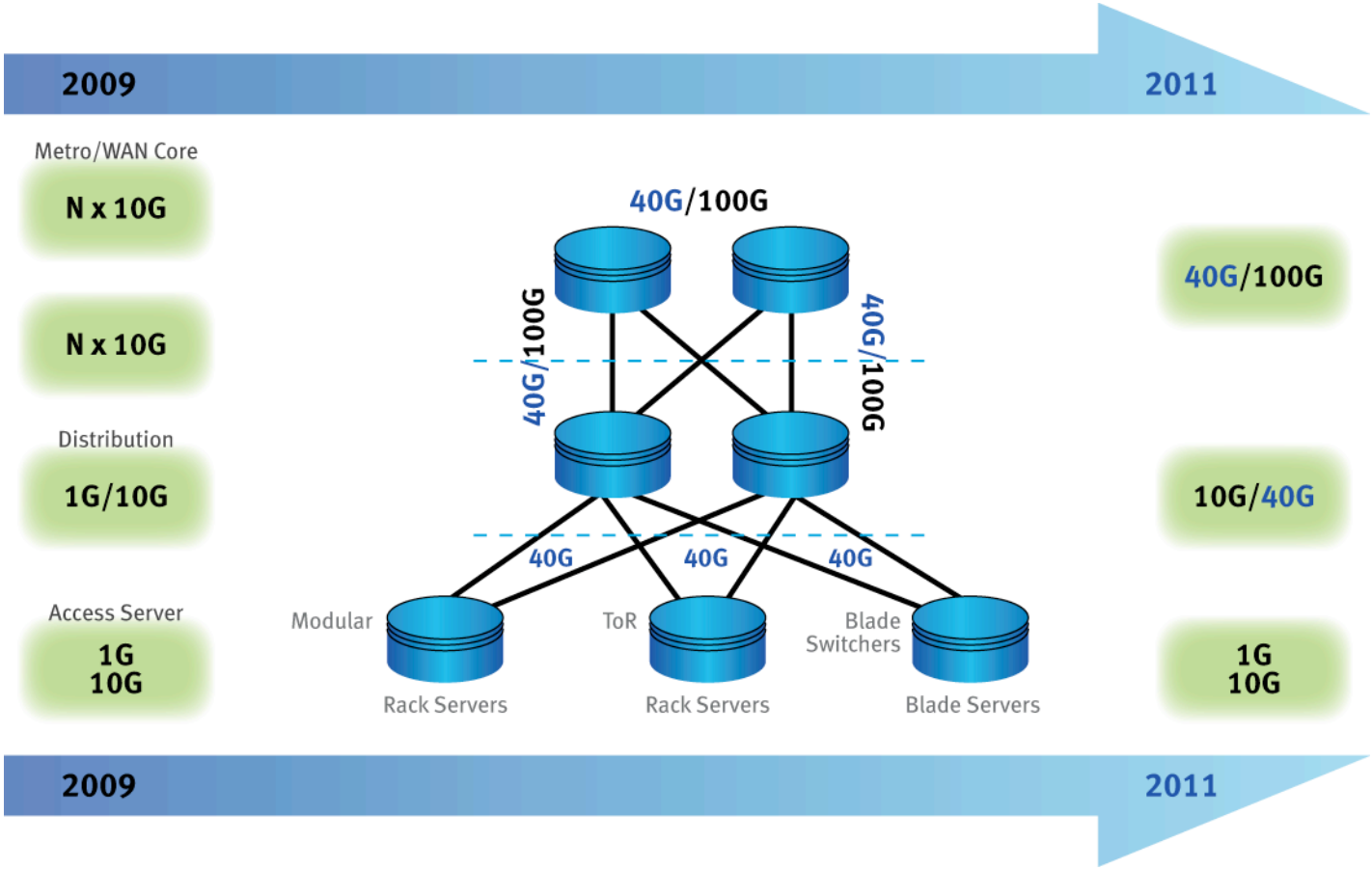


10G 40G 100G

Source: InfoNetics June '09

LIGHT READING

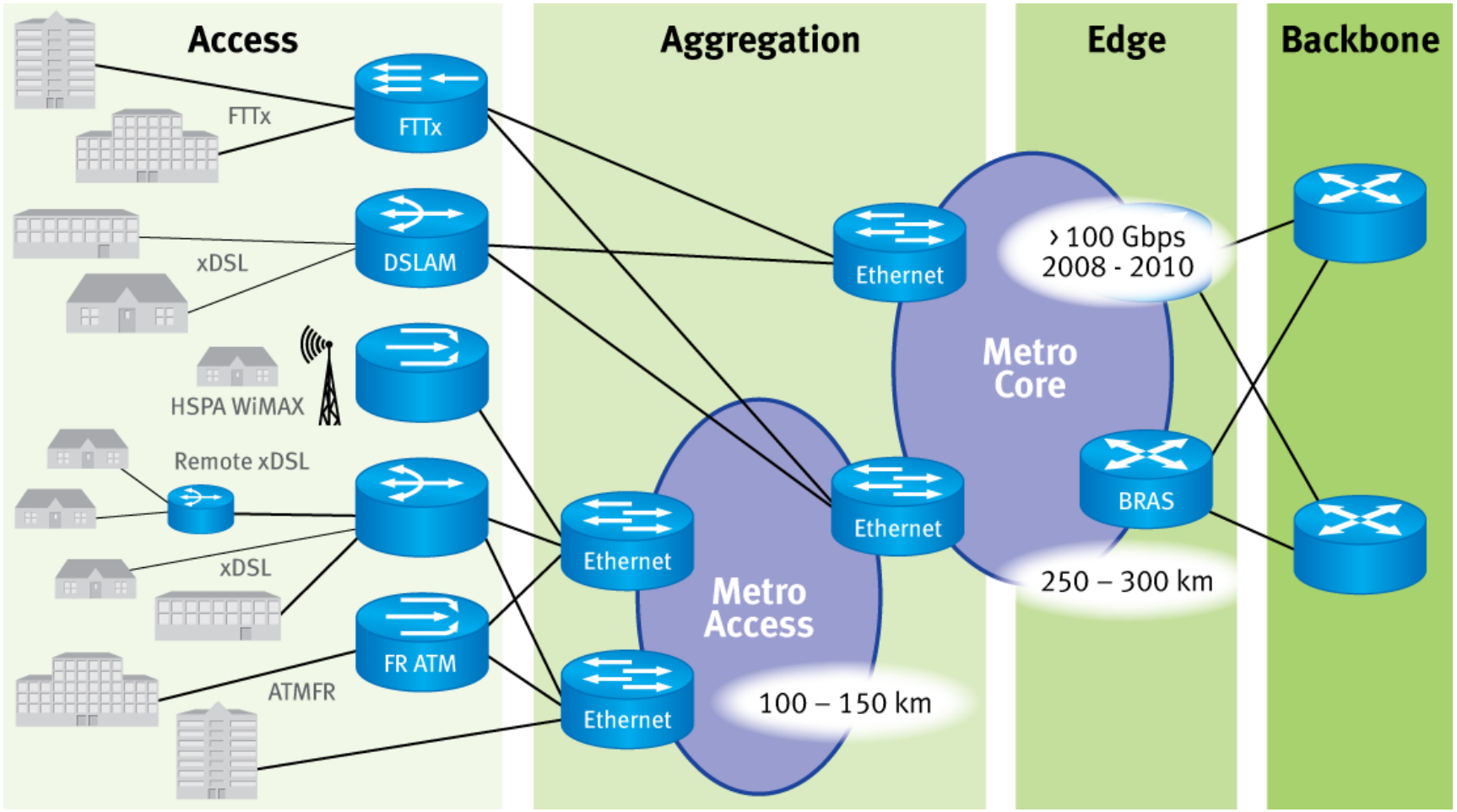
Data Center Network



SPIRENT

LIGHT READING

Service Provider Network



SPIRENT

LIGHT READING

Validation

- Each element has to be characterized and then the system as whole
- Different Types of Test
  - Conformance
  - Functional
  - Performance
- Different Levels of Focus

Device	System
Data plane	End-to-end capability
Control plane	Network Impairment
Backplane	Convergence Times
	Prioritization
	Resiliency

But first we need to understand the technology

SPIRENT

LIGHT READING

IEEE P802.3ba
40/100G Ethernet Basics

IEEE P802.3ba 40/100G Ethernet PHY Types

PHY Type	Data rate	Distance	Media	Technology
100GBASE-ER4	100Gb/s	40km	Single Mode Fibre	4 x 25Gb/s (28.78125GBaud) 1310nm DWDM (5nm), SOA
100GBASE-LR4	100Gb/s	10km		4 x 25Gb/s (28.78125GBaud) 1310nm DWDM (5nm)
40GBASE-LR4	40Gb/s			4 x 10Gb/s (10.3125GBaud) 1310nm CWDM (20nm)
100GBASE-SR10	100Gb/s	100m	OM3 multimode fibre	10 x 10Gb/s (10.3125GBaud) 850nm, 10 pairs of fibres
40GBASE-SR4	40Gb/s			4 x 10Gb/s (10.3125GBaud) 850nm, 4 pairs of fibres
100GBASE-CR10	100Gb/s	7m	Copper cable assembly	10 x 10Gb/s (10.3125GBaud) 10 differential pairs
40GBASE-CR4	40Gb/s			4 x 10Gb/s (10.3125GBaud) 4 differential pairs
40GBASE-KR4	40Gb/s	1m	Backplane	4 x 10Gb/s (10.3125GBaud) 4 10GBASE-KR channels

SPIRENT

LIGHT READING

IEEE P802.3ba 40/100G Ethernet Layers & Interfaces

XLGMII / CGMII

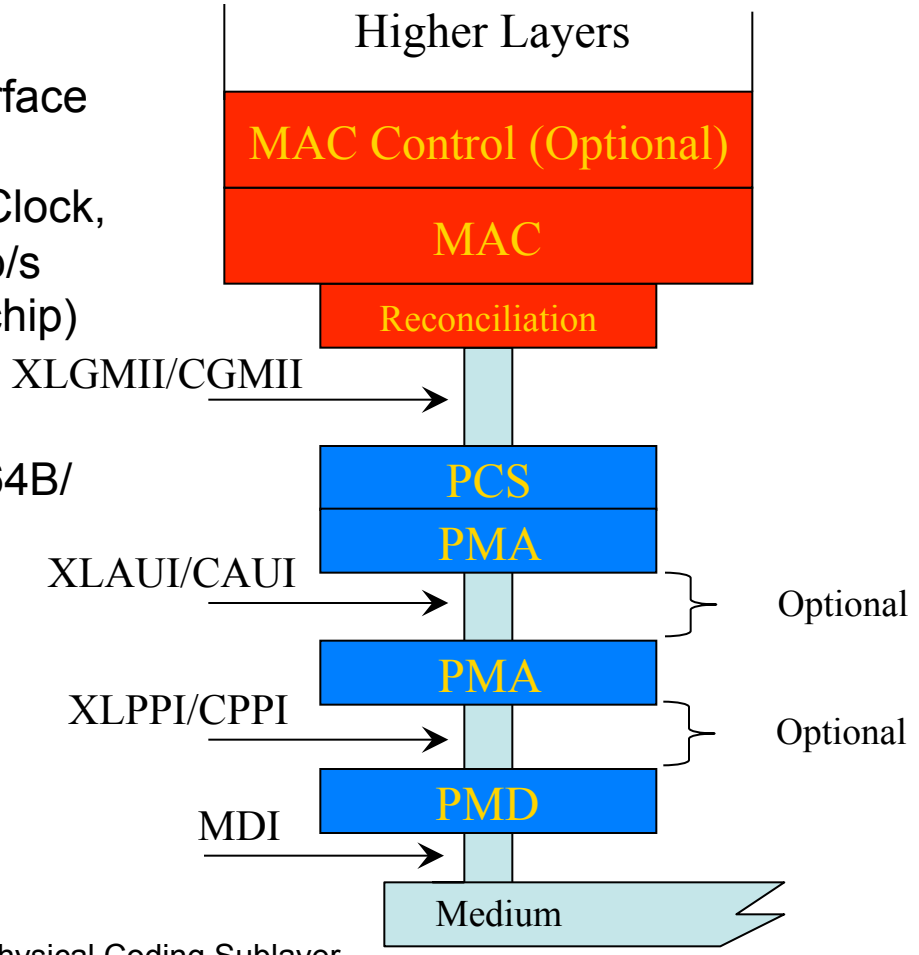
40 / 100 Gigabit **M**edia **I**ndependent **I**nterface
TX and RX data paths
64 data (8 'lanes' of 8 bits), 8 Control, 1 Clock,
625Mhz @ 40Gb/s,1.5625GHz @ 100Gb/s
Logical interface (supports system on a chip)

XLAUI / CAUI

40 / 100 Gb/s **A**ttachment **U**nit **I**nterface
4 (XLAUI) / 10 (CAUI) lanes of 10Gb/s, 64B/
66B encoded, 10.3125GBaud/s
To support 25cm FR- 4 PCB traces
Total: XLAUI 16 pins; CAUI 40 pins

XLPPI / CPPI

40 / 100 Gb/s **P**arallel **P**hysical **I**nterface
40GBASE-SR4 or 100GBASE-SR10
Encoding the same as XLAUI/CAUI
No retiming (short distance)

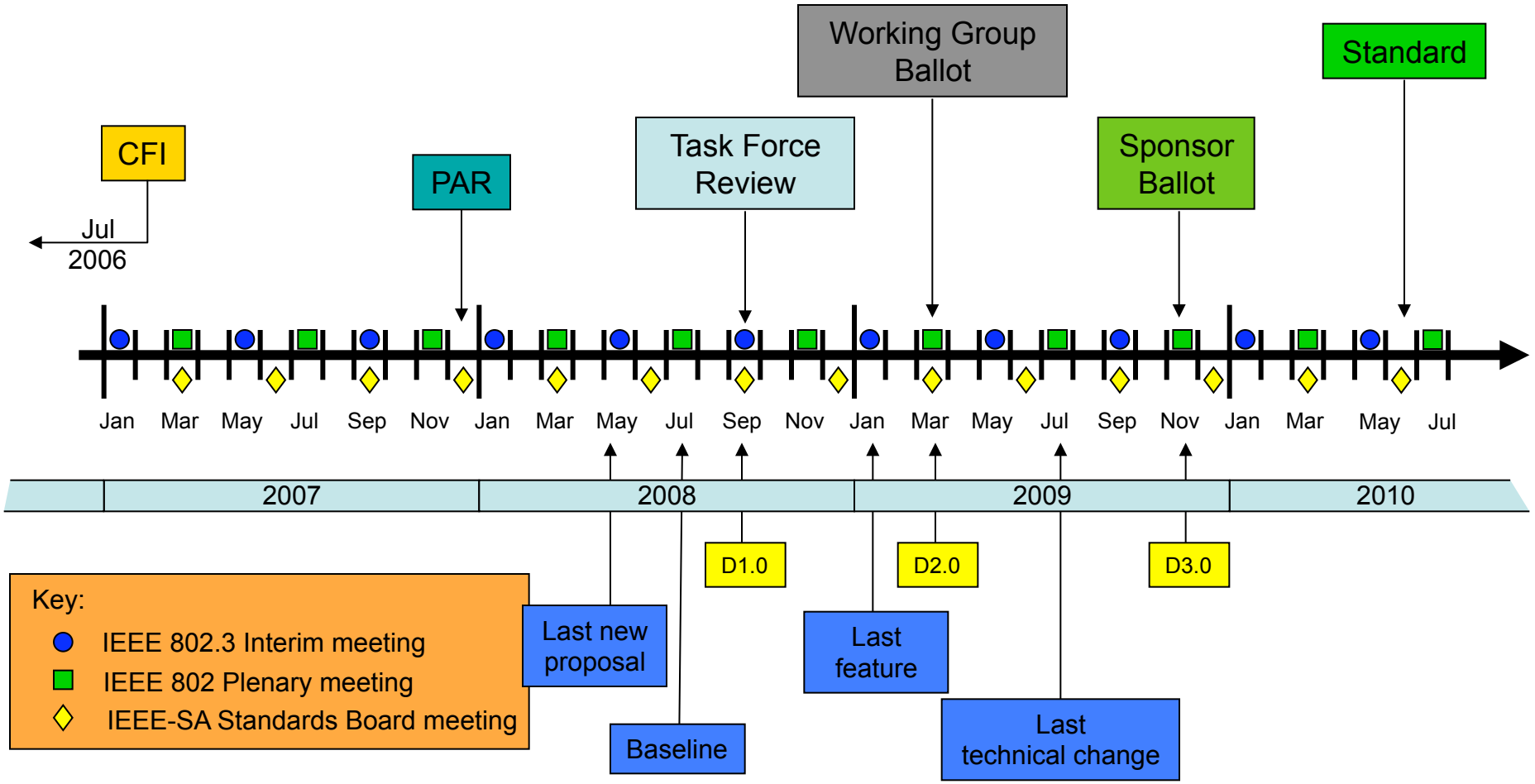


PCS = Physical Coding Sublayer
PMA = Physical Medium Attachment
PMD = Physical Medium Dependent

SPIRENT

LIGHT READING

IEEE P802.3ba 40/100G Ethernet Standards/Timeline



SPIRENT

LIGHT READING

Poll Question #1

What timeframe do you think 40/100G Ethernet will start to be adopted or deployed?

- Pre standard (2009-mid 2010)
- Late 2010
- Early 2011
- After mid to late 2011
- Even later still

SPIRENT

LIGHT READING

How It Works

IEEE P802.3ba 40/100G Ethernet Layers & Interfaces

XLGMII / CGMII

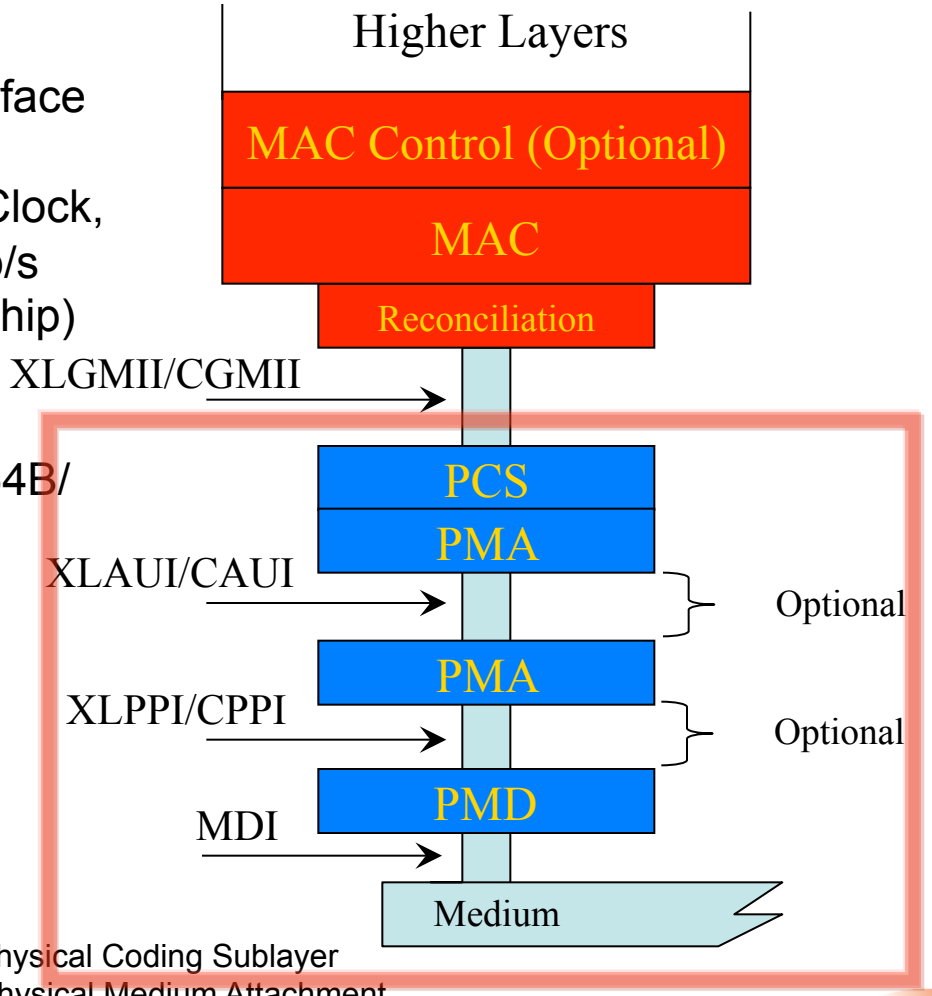
40 / 100 Gigabit **M**edia **I**ndependent **I**nterface
TX and RX data paths
64 data (8 'lanes' of 8 bits), 8 Control, 1 Clock,
625Mhz @ 40Gb/s,1.5625GHz @ 100Gb/s
Logical interface (supports system on a chip)

XLAUI / CAUI

40 / 100 Gb/s **A**ttachment **U**nit **I**nterface
4 (XLAUI) / 10 (CAUI) lanes of 10Gb/s, 64B/66B encoded, 10.3125GBaud/s
To support 25cm FR- 4 PCB traces
Total: XLAUI 16 pins; CAUI 40 pins

XLPPI / CPPI

40 / 100 Gb/s **P**arallel **P**hysical **I**nterface
40GBASE-SR4 or 100GBASE-SR10
Encoding the same as XLAUI/CAUI
No retiming (short distance)



PCS = Physical Coding Sublayer
PMA = Physical Medium Attachment
PMD = Physical Medium Dependent

SPIRENT

LIGHT READING

IEEE P802.3ba 40/100G Ethernet Physical Coding Sublayer (PCS)

- Serial 40Gb/s Ethernet support some way off
  - Not included in initial 40Gb/s PHY set
  - Work on project proposal starts at November 2009 meeting
- Serial 100Gb/s Ethernet operation further in the future
  - Technology demonstrations just taking place
- IEEE P802.3ba PHYs therefore lane based
  - Encoding termed Multi-Lane Distribution (MLD)
  - Logical lanes called PCS lanes (PCSL)
- IEEE P802.3ba PCS designed to support future PHYs
  - Don't want a new PHY to require a new PCS
  - This is what happened for some of the 10Gb/s PHYs

SPIRENT

LIGHT READING

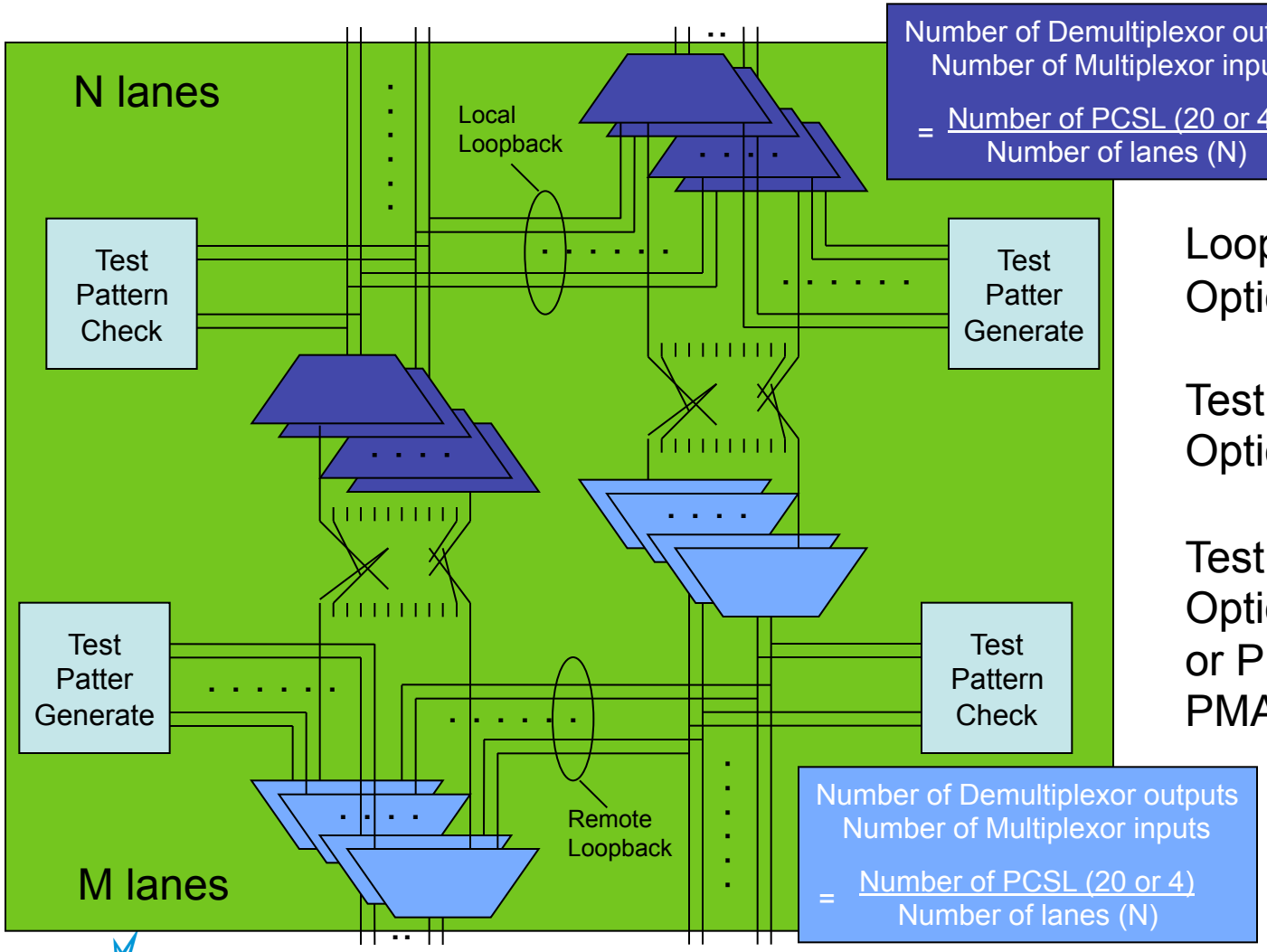
Multi-Lane Distribution (MLD)

- Based on 10GBASE-R 64B/66B PCS
  - Data striped round robin across lanes 66 bit blocks at a time
  - Periodic alignment blocks are added to allow deskew
    - Allows random ordering of lanes across link
    - Rate of marker insertion low, every 16383 blocks
      - Accommodated in normal IPG
- Support multiple PCSL widths with the same PCS layer
  - PMA maps n lane interface to m lane interface
  - PMA is simple bit level muxing
    - Does not know or care about PCS coding
  - Alignment performed at RX PCS
  - Compensation for 'fixed' skew performing at RX PCS
    - Compensation for 'dynamic' skew performed at each sublayer

SPIRENT

LIGHT READING

PMA Logical Paths



Loopback paths
Optional

Test Pattern, upper
Optional if XLAUI/CAUI

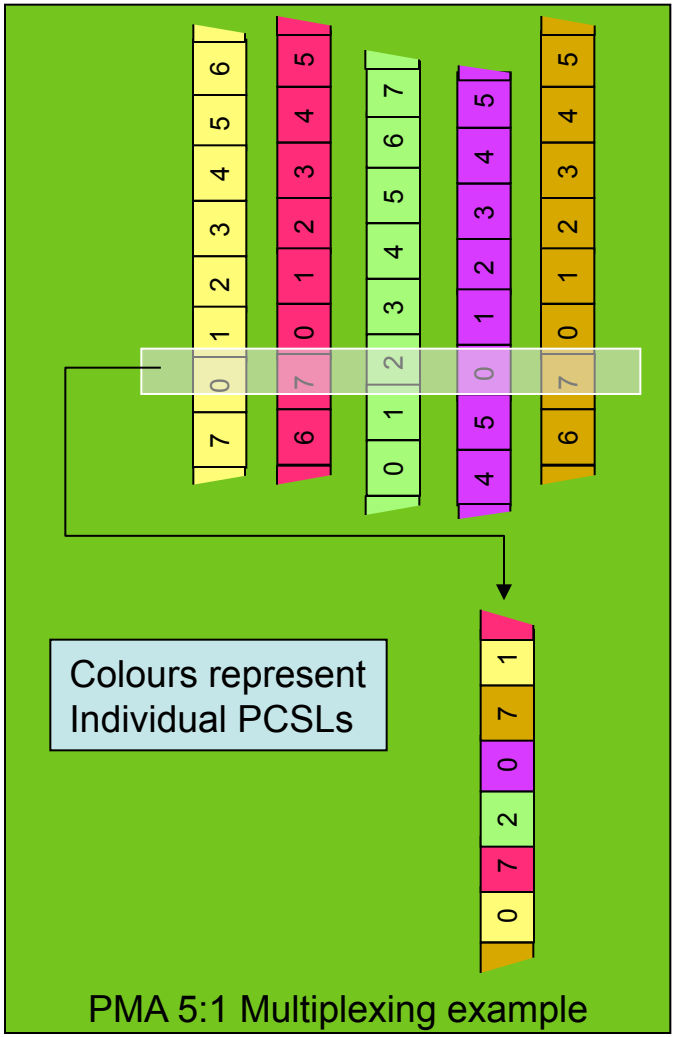
Test pattern, lower
Optional if XLAUI/CAUI
or PMD below if no other
PMA below

SPIRENT

LIGHT READING

PMA Multiplexing and Demultiplexing

- Chosen order not important
  - In this example must always be Yellow, Red, Green, Purple and Brown, Red bit 7 is followed by Green 2 then Purple 0
- Once chosen order must be maintained
- Order may change after each reset
  - PMA operates at bit level, has no knowledge of lane markers
- Skew variation compensated
  - If the skew varies due to changes in channel the PMA must track to maintain lane order
  - Fixed Skew not removed



PMA 5:1 Multiplexing example

SPIRENT

LIGHT READING

Poll Question #2

Are you planning on implementing 40G or 100G Ethernet?

- 40G
- 100G
- 40G and 100G
- No plans

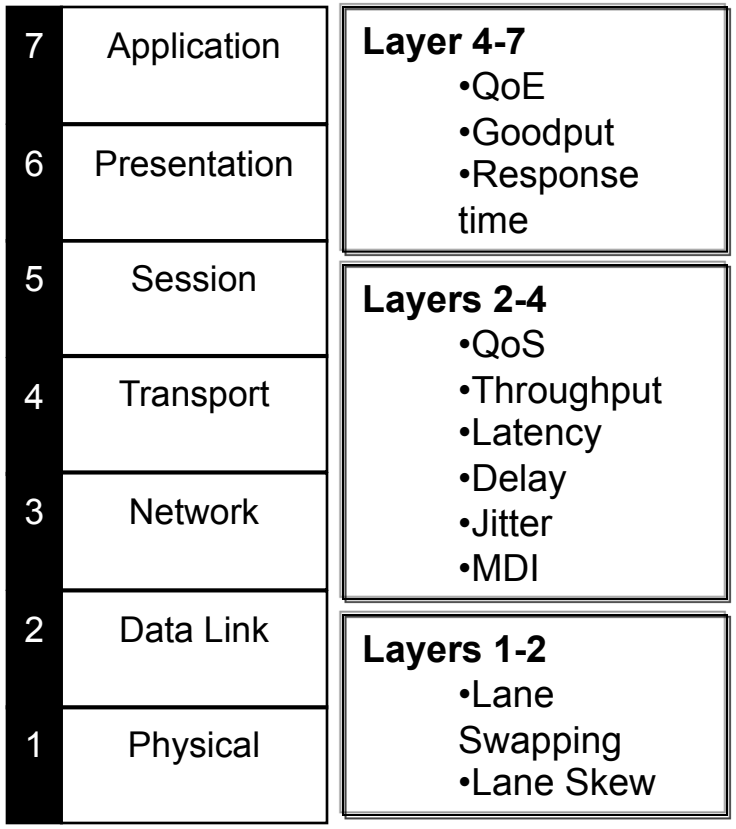
SPIRENT

LIGHT READING

Validating That It Works

Test Methodologies

- Many RFC methodologies available (IETF)
  - RFC 2544 – Network Interconnect Devices
  - RFC 2889 – LAN Switching
  - RFC 3511 – Firewall
  - RFC 3918 – IP Multicast
  - RFC 4445 – MDI
- Other methodologies/standards exist
  - TIA -921 / ITU-T G.1050 – Network Model
- Provides means to understanding each Layer and the interaction between Layers



For 40/100G, initial focus will be at the lower layers

SPIRENT

LIGHT READING

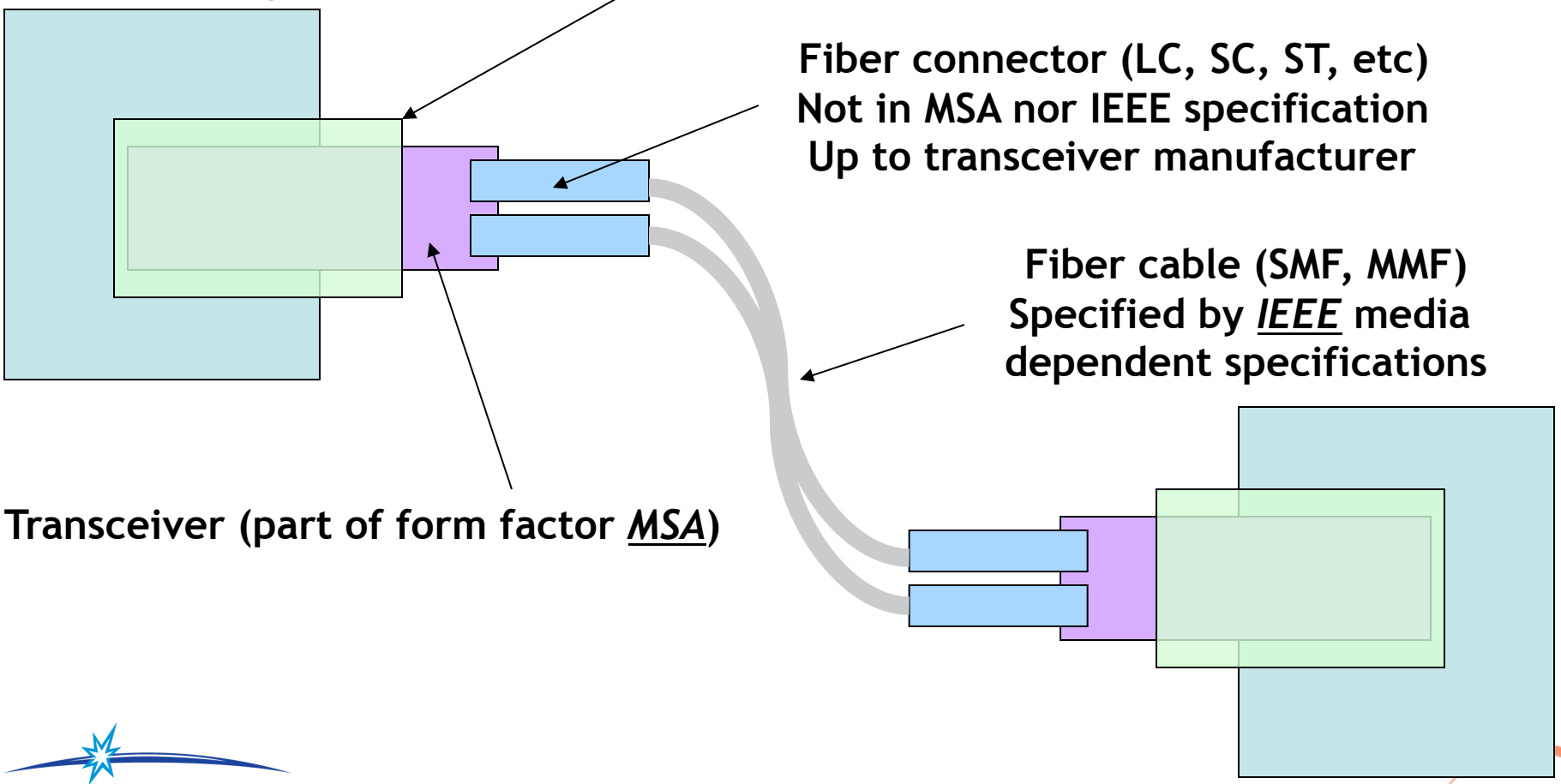
Testing The Link Components

PHY Components – PCS, PMA, etc (specified by IEEE)

Cage (part of form factor *MSA*)

Fiber connector (LC, SC, ST, etc)
Not in MSA nor IEEE specification
Up to transceiver manufacturer

Fiber cable (SMF, MMF)
Specified by *IEEE* media dependent specifications



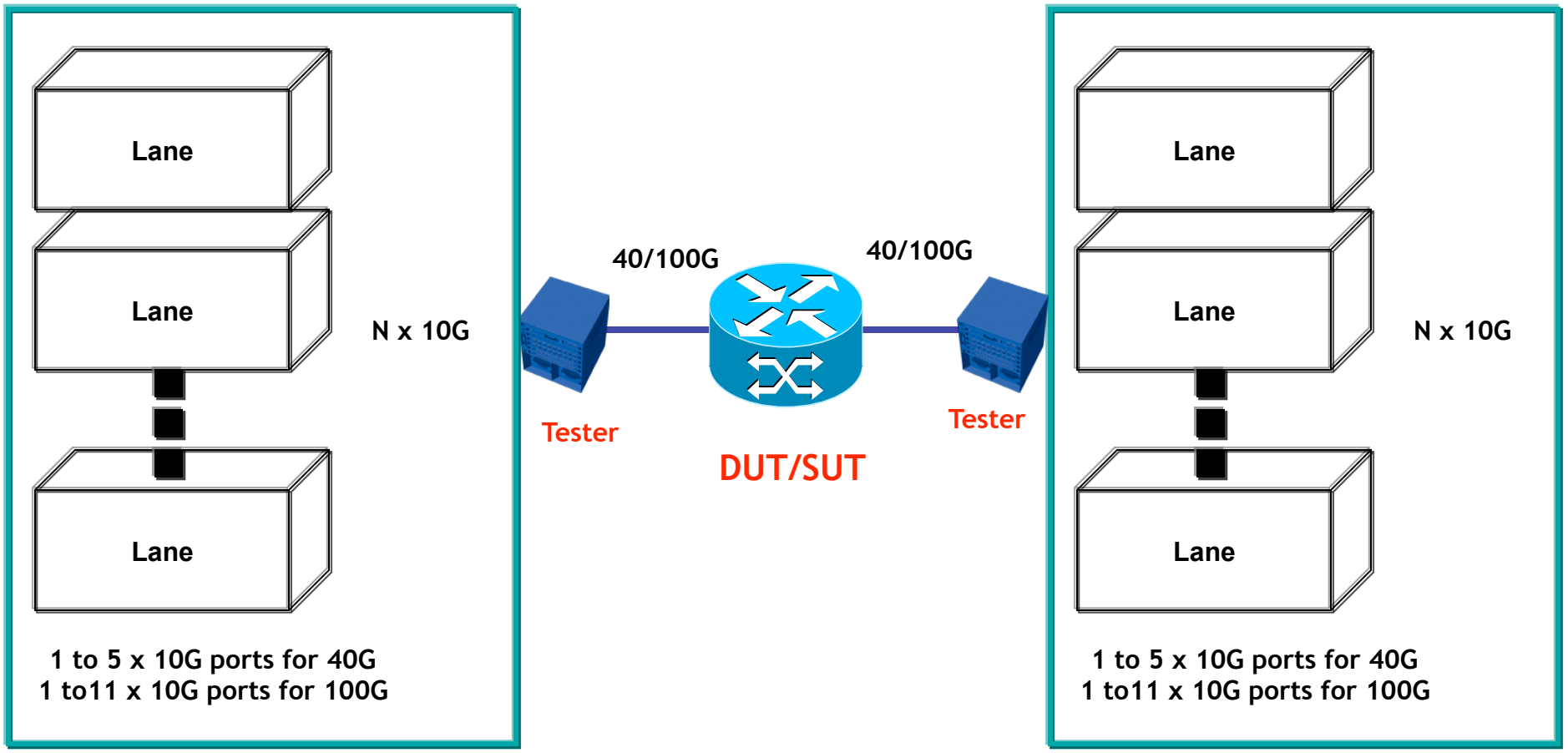
Transceiver (part of form factor *MSA*)

SPIRENT

Spirent TestCenter Port

LIGHT READING

Test Stage 1 : Lower Layers



Emulated Devices

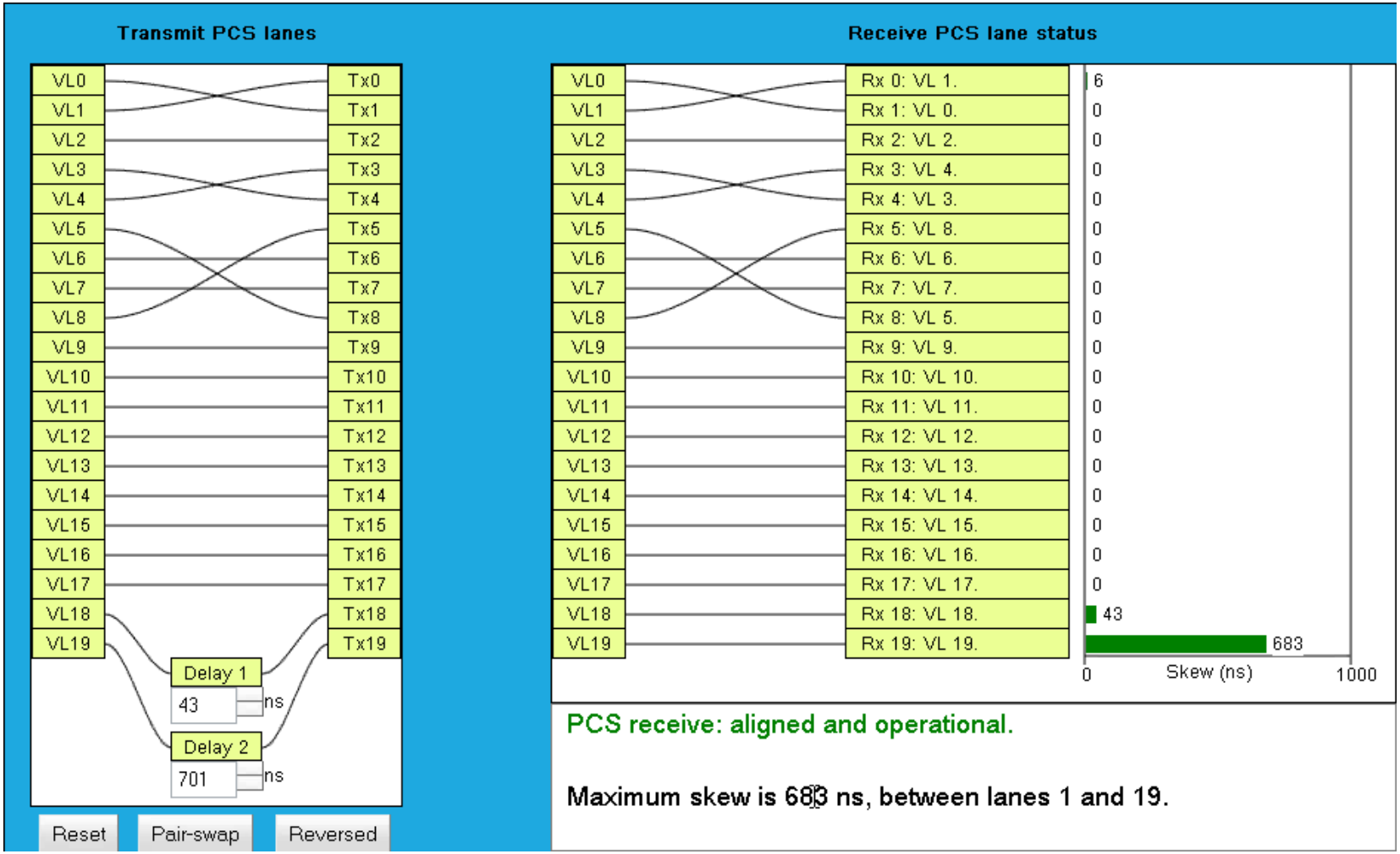
Emulated Devices

- Lane Swapping
- Lane Skew

SPIRENT

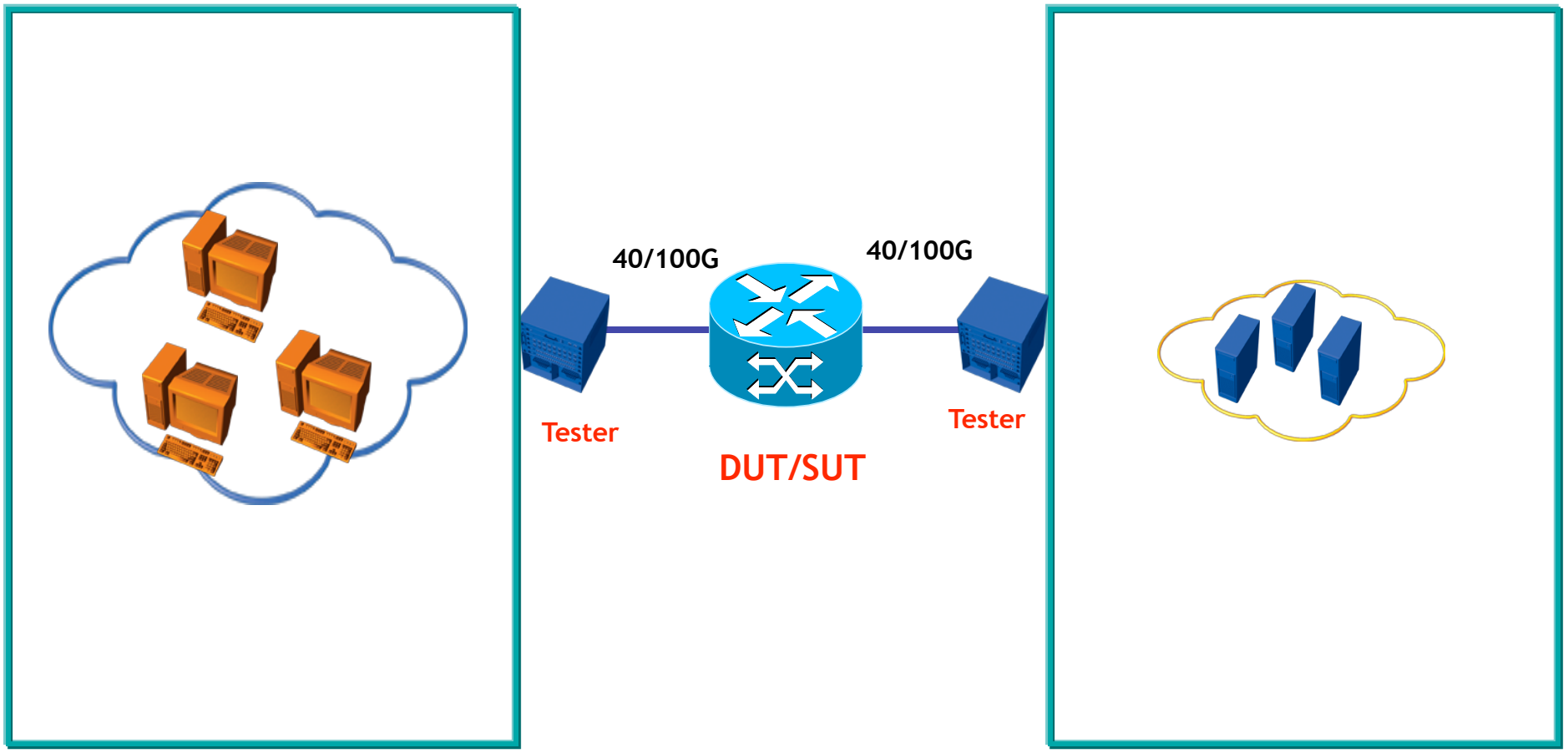
LIGHT READING

Example of PCS Skew



Testing at the PMA and MAC layer should also be verified

Test Stage 2: Data plane

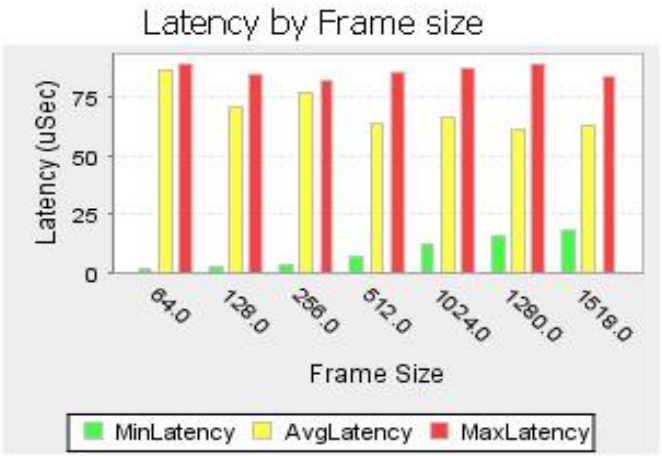


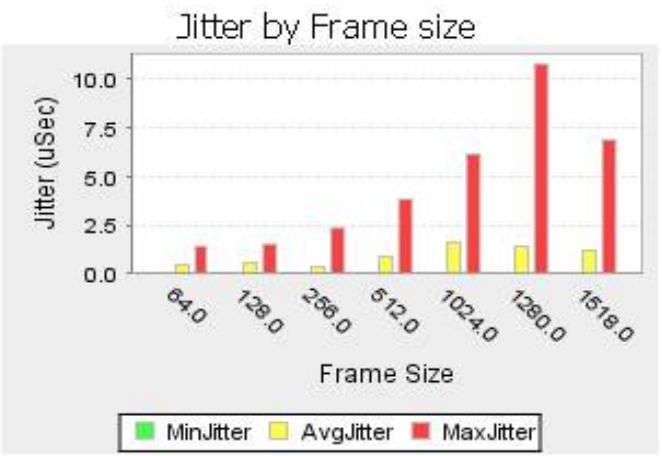
- IP

SPIRENT

LIGHT READING

Example of Data Performance





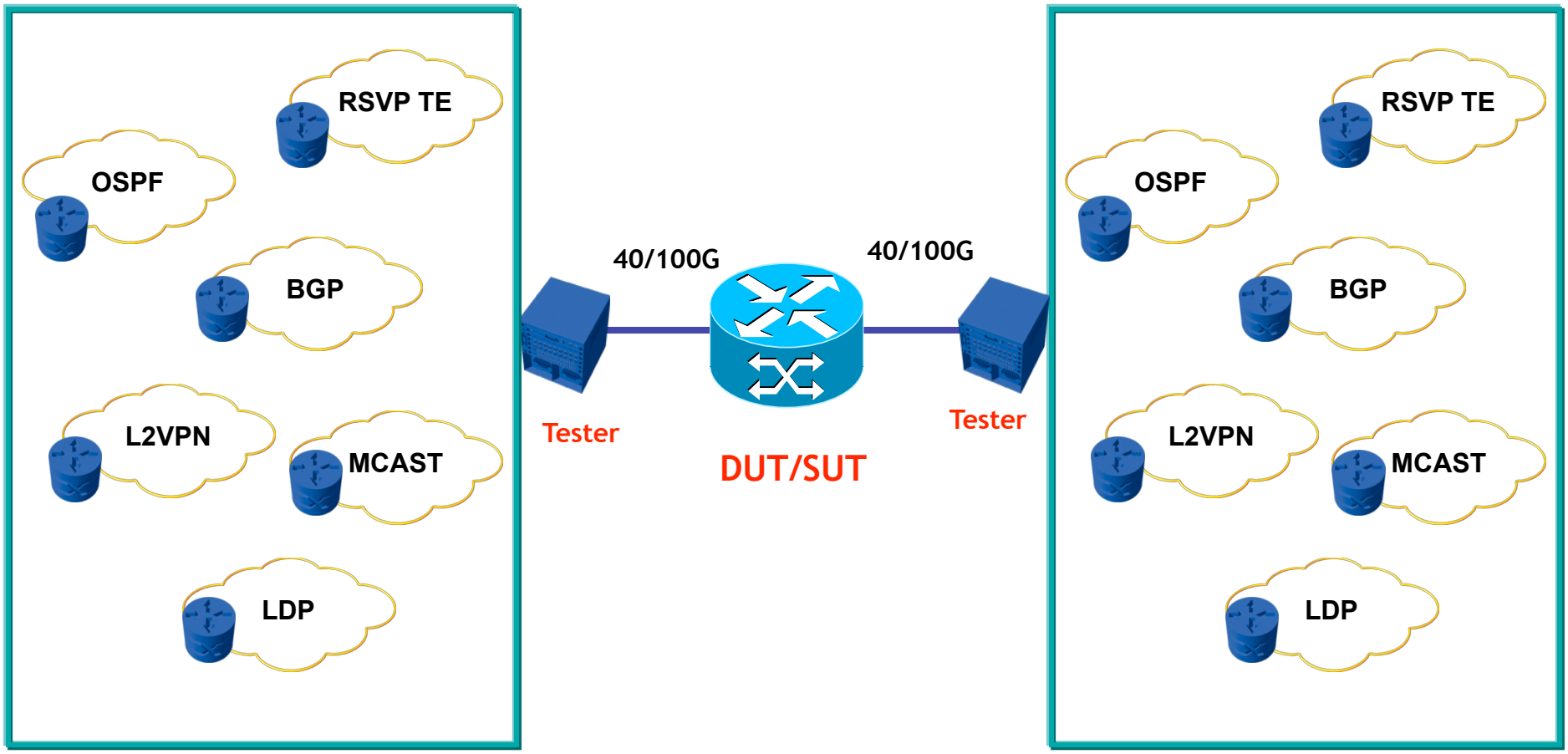
Frame Size(bytes)	Load (%)	Latency Min (uSec)	Latency Avg (uSec)	Latency Max (uSec)	Jitter Min (uSec)	Jitter Avg (uSec)	Jitter Max (uSec)
64	10	1.97	2.065	2.16	0.03	0.058	0.13
64	32.5	1.97	2.066	2.16	0.05	0.076	0.1
64	43.75	1.97	2.066	2.16	0	0.005	0.17
64	49.375	1.97	2.066	2.16	0	0.037	0.15
64	50.078	1.97	168.24	173.03	0	0.648	2.14
64	50.781	1.97	170.861	177.3	0	0.921	2.83
64	52.187	1.97	171.532	178.14	0.04	0.883	2.87
64	55	1.97	171.923	175.88	0	0.827	2.53
128	10	2.69	2.783	2.88	0	0.023	0.15

SPIRENT

Testing with Convergence should also be verified

LIGHT READING

Test Stage 3: Control plane



Emulated Devices

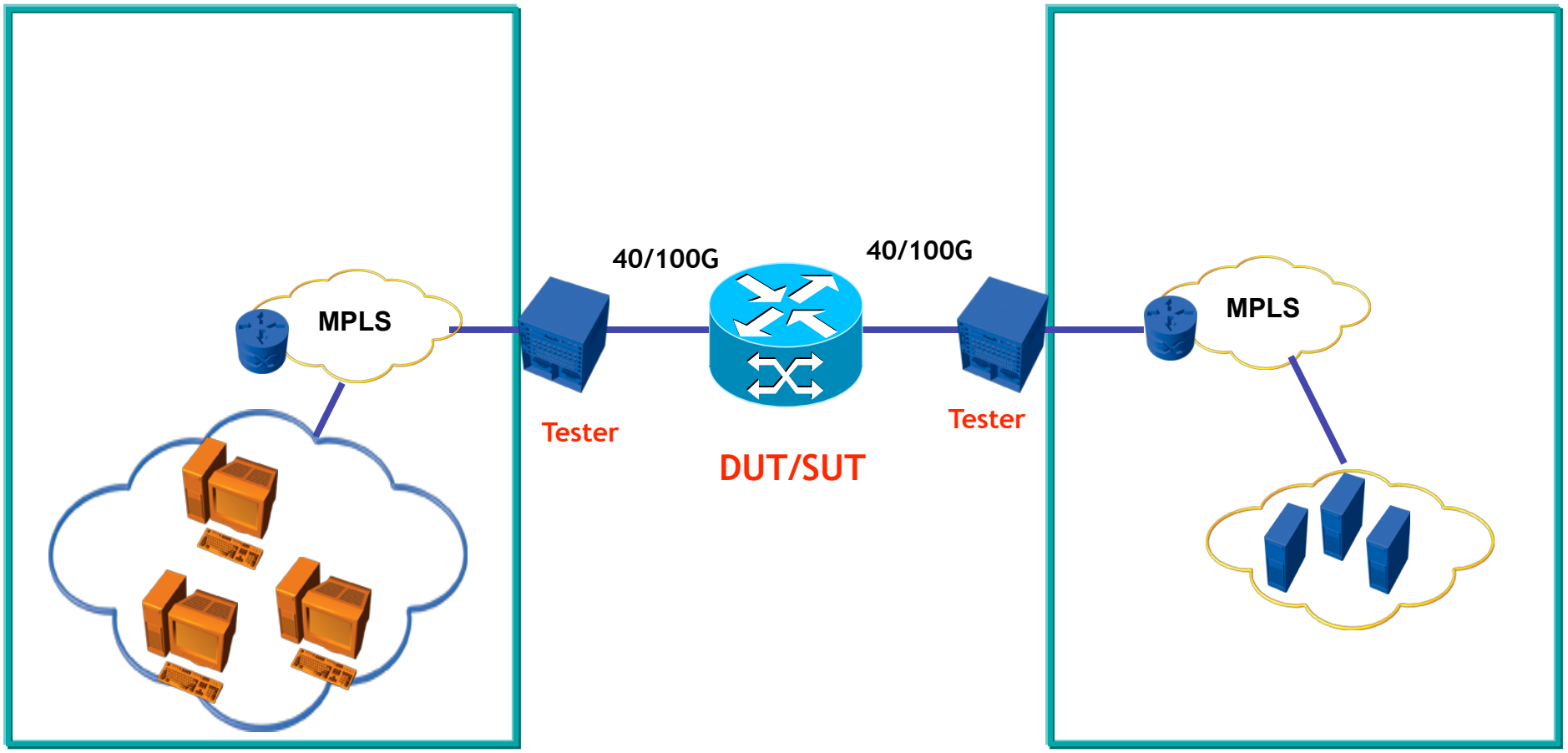
- **Routing**

Emulated Devices

SPIRENT

LIGHT READING

Test Stage 4: Data / Control plane



Emulated Devices Behind Emulated Devices

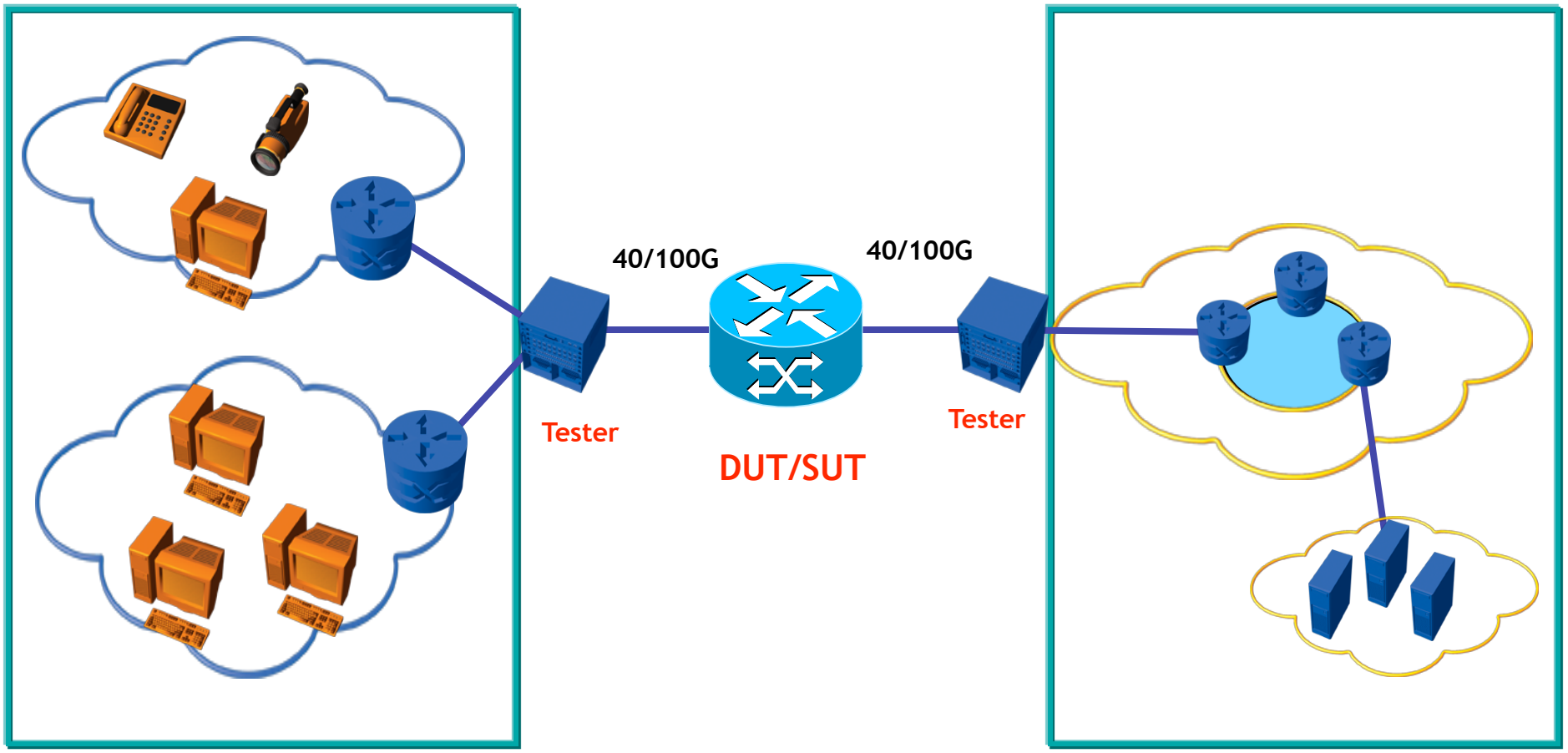
- **Routing and IP**

Emulated Devices Behind Emulated Devices

SPIRENT

LIGHT READING

Test Stage 5: End Users



Emulated Devices Behind Emulated Devices

- MOS
- MOS-V
- PSEQ

Emulated Devices Behind Emulated Devices

SPIRENT

LIGHT READING

What The User Will See

0.0% Packet Loss

Small amounts of Packet Loss can cause high levels of distortion

0.5 % Packet Loss

5.0 % Packet Loss

SPIRENT

LIGHT READING

Poll Question #3

What equipment in the network will your organization probably invest in first for 40/100G support?

- Top-of-rack
- Edge Routers
- Core Routers
- Other

SPIRENT

LIGHT READING

Summary

- 40/100G is the next major advancement for Ethernet offering large amount of bandwidth
- Creates new challenges on multiple levels
  - Lower layers are the first hurtle
  - Upper layer will be next
- End goal is improving services such as Video
- Validate that the system is ready

SPIRENT

LIGHT READING

Q&A

www.spirent.com

www.h3cnetworks.com