

# In Support of BMP and Economic Feasibility

---

BRAD BOOTH, MICROSOFT

SEPTEMBER 2014

# Supporters

---

Your name could be here...

Or here...

# Broad Market Potential (BMP)

---

Required to address the following:

Each proposed IEEE 802 LMSC standard shall have broad market potential. At a minimum, address the following areas:

- a) Broad sets of applicability.
- b) Multiple vendors and numerous users.

# Economic Feasibility (EF)

---

Required to address the following:

Each proposed IEEE 802 LMSC standard shall provide evidence of economic feasibility. Demonstrate, as far as can reasonably be estimated, the economic feasibility of the proposed project for its intended applications. Among the areas that may be addressed in the cost for performance analysis are the following:

- a) Balanced costs (infrastructure versus attached stations).
- b) Known cost factors.
- c) Consideration of installation costs.
- d) Consideration of operational costs (e.g., energy consumption).
- e) Other areas, as appropriate.

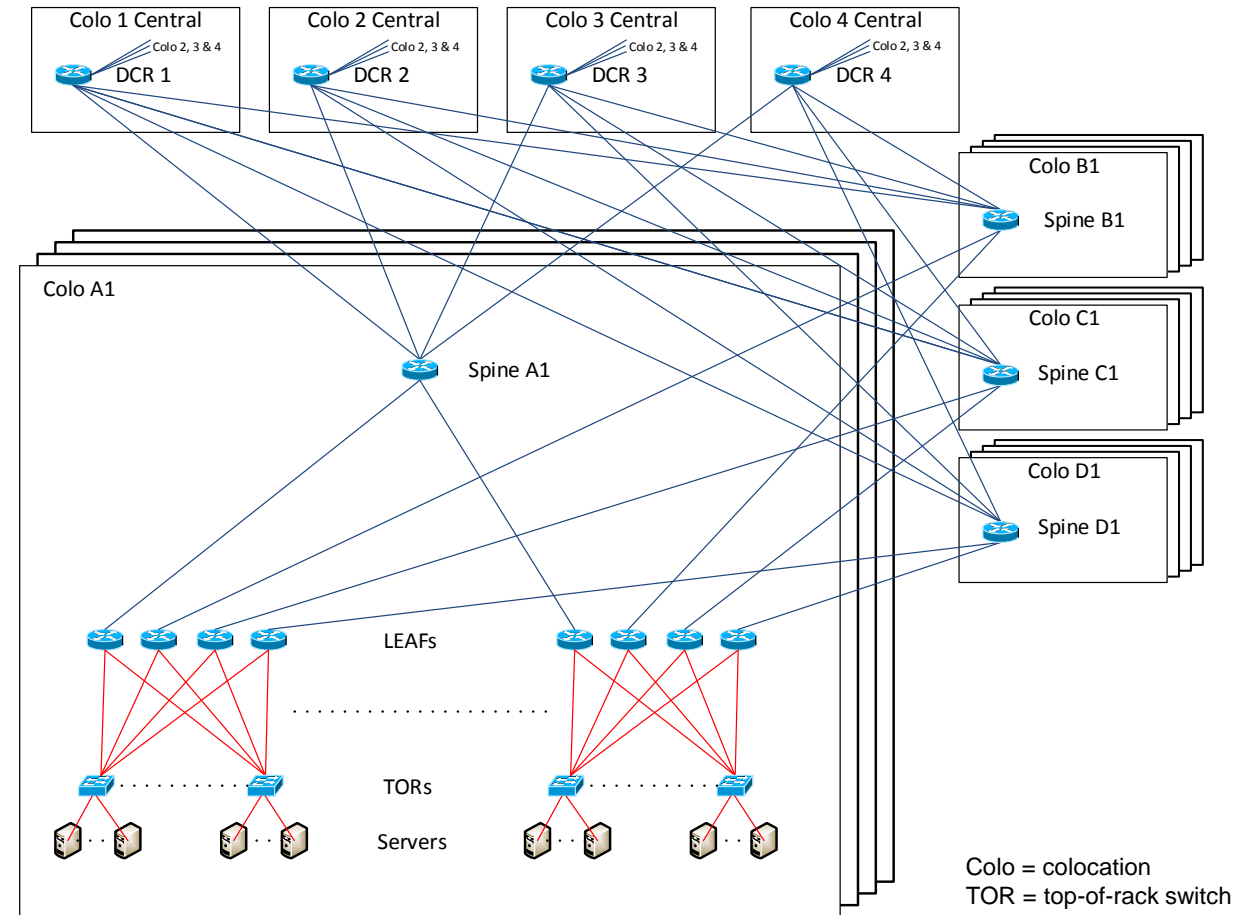
# Server to TOR Interconnect in Cloud Data Centers

Link data rate is typically a fractional rate of the data rate in the Core (TOR to Leaf, Leaf to Spine, etc.)

- Example, 10G server to TOR data rate with Core at 40G

Volume is an order of magnitude greater than any other interconnect in the data center

Most cost sensitive portion of the network



# Single-lane Interfaces in 10GbE Server

10GbE volume ramp in servers coincided with the availability of single-lane interfaces

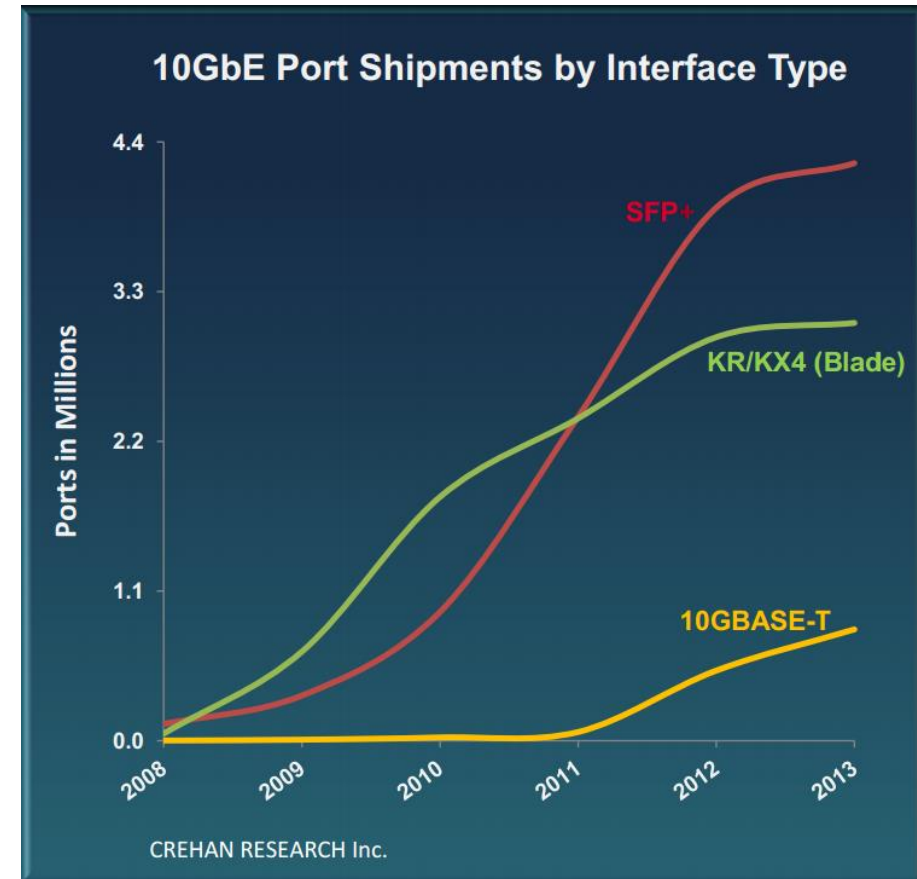
Early adopters (2004-2008) used

- XAUI-based optics
- 10GBASE-CX4
- 10GBASE-KX4

Single-lane backplane and twinax solutions eclipsed the early-adopter volume starting in 2009

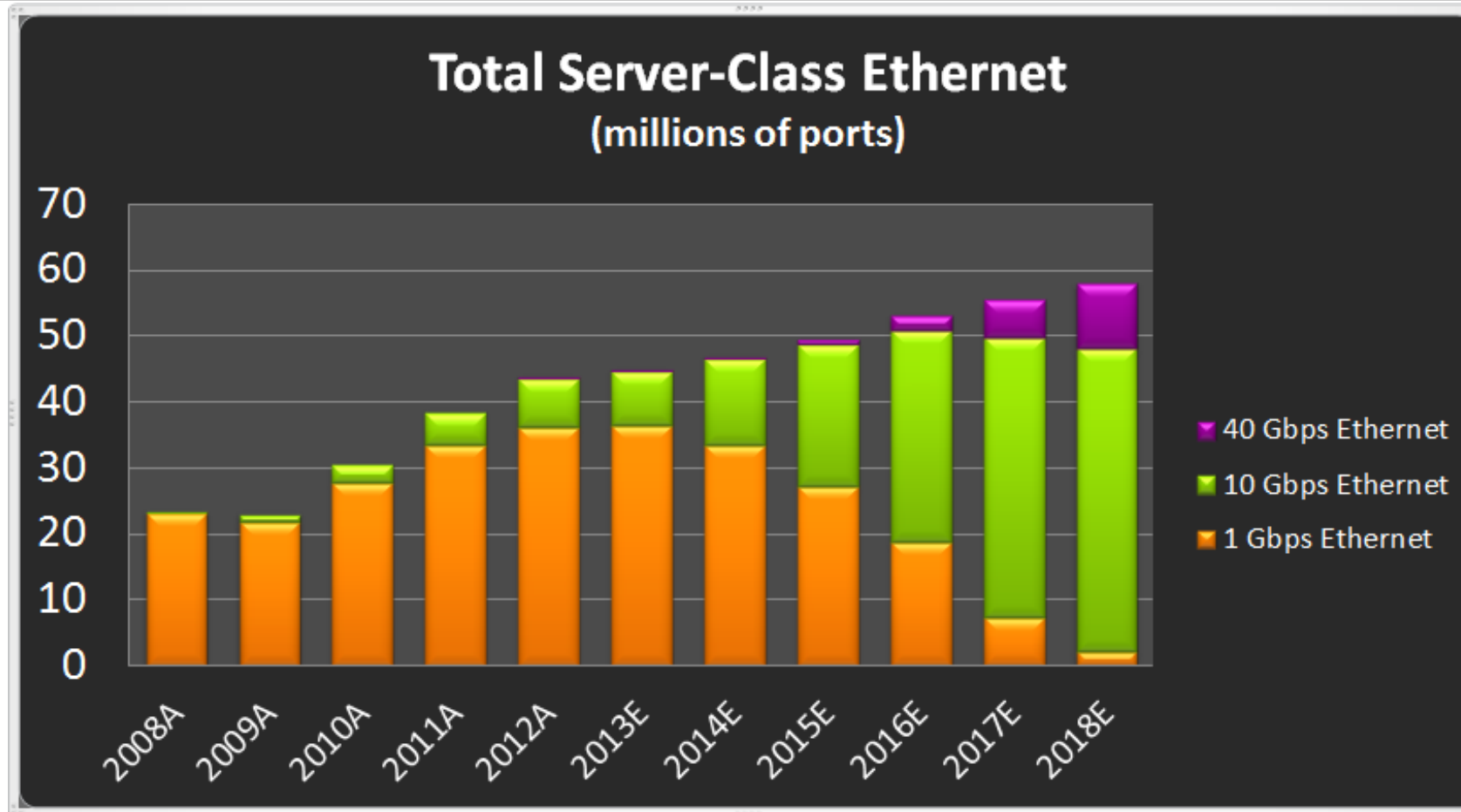
Notes:

- SFP+ majority use is twinax, then SR; accurate share data unavailable
- Blade server is mostly KR based upon system configuration. KX4 vs. KR split data unavailable.



Data source: Crehan Research, Inc., Q1'2014

# Server Ethernet Port Speed Forecast Based Upon 10 Gbps SerDes



Data source:  
Crehan Research, Inc., Q1'2014

# Port Speed and Media Observations

---

Market is wide and varied

- Compute server to switch links may operate at 10G
- Storage server to switch links may operate at 40G
- Core network may operate at 40G
- Management network typically operates at 1G

No single answer to the bandwidth need question... driven by the application need vs. CPU compute power

- Volume adoption is cost sensitive
- Servers are the “big ticket” items in data centers

25G creates a window where latest SerDes technology can provide the higher port speed at lower cost



# Transition of Technology

---

10 Gbps SerDes has been the primary interface technology for Ethernet for over 10 years

- Started with IEEE Std. 802.3ae (10 Gigabit Ethernet)
- Continued through IEEE Std. 802.3ba (40 and 100 Gigabit Ethernet)

25 Gbps SerDes technology is posed to be the next primary interface technology

- Work completed in the OIF for CEI-28G
- Adopted into specifications such as IEEE Std. 802.3bj and IEEE P802.3bm
- Even under consideration by IEEE P802.3bs
- Lifetime of the technology could be extended with modulation schemes

Ecosystem is developing that supports 25 Gbps technology

- Next generation silicon for switches, NICs and components
- Cabling, connectors, etc.

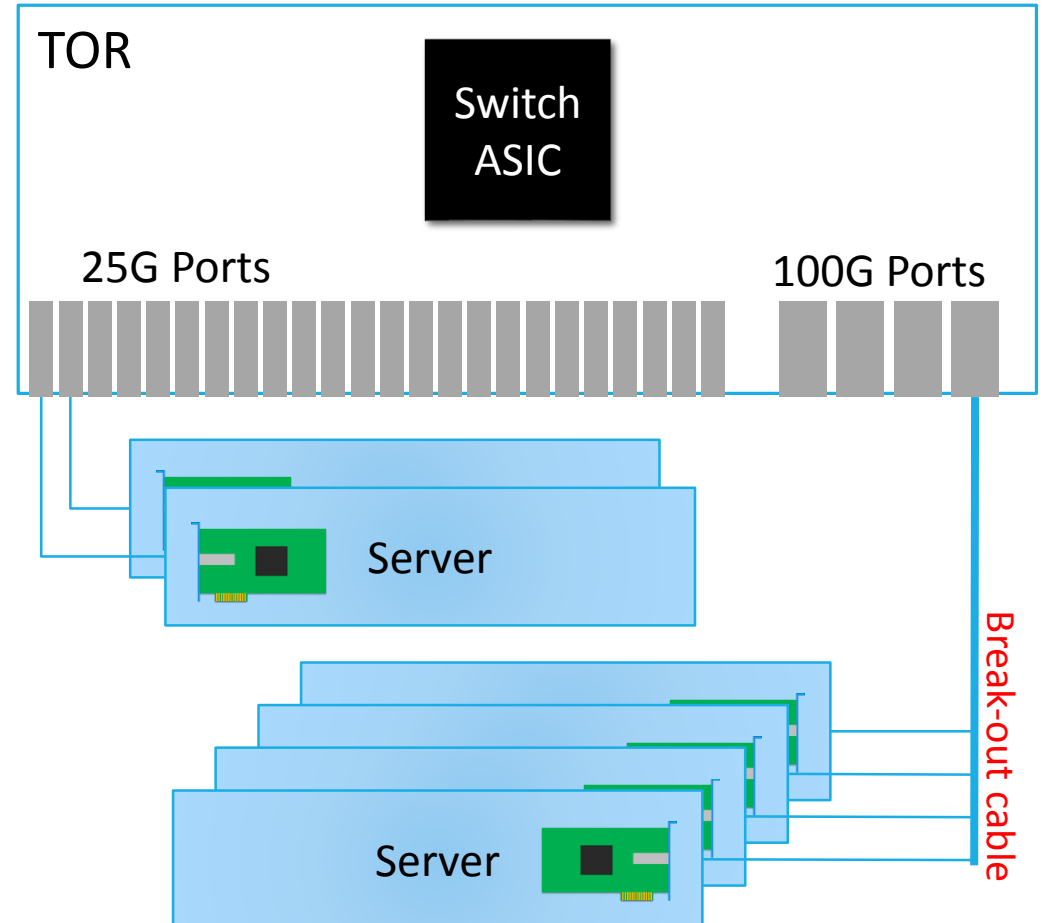
# 25 Gb/s Ethernet Connectivity

## Network topology re-use

- Single 25Gb/s SFP28 port implementation
- Quad 25Gb/s QSFP28 breakout implementation possible

## Reach requirement

- In a rack, placement of the TOR is typically in the center of the rack
  - Shorter cables tend to cost less
- In some situations, there is more than 1 TOR per rack
- Maximum reach within the rack is 3m
- Typical maximum reach from server to TOR is 2m



# Port Speed and Media Observations

---

Market is wide & varied – no single answer to the BW need question!

- Port speeds from 1Gb/s to 100Gb/s will co-exist
- Variety of CPU architectures, clock speeds, core counts, CPUs/system
- Mix of software applications with varied needs of I/O BW vs. CPU compute power

Leading edge drives the higher speed as soon as available

- Initial adoption: 10G ~2004; 40G ~2012; 100G ~2015

...but volume adoption is cost sensitive

1G→10G crossover forecast has repeatedly shifted right

- 2012→2014→2016
- In turn, transition to 40G slower than prior forecasts
- Creates a window where new technology can provide the higher port speed at lower cost
- Some portion of today's 10G KR & SFP+ users are likely to adopt 25Gb/s on the way to a higher speed

# Maintaining Ethernet Economics

---

Switch ASIC connectivity limited by SerDes that can fit on a die (aka pad limited)

25 Gb/s lane maximizes bandwidth/pin and switch fabric capability vs. older generation

Single-lane port maximizes server connectivity available in single ASIC

25 Gb/s port optimizes both port count and total bandwidth for server interconnect

Inefficiency with next generation TOR switch silicon using 25 Gb/s SerDes

Option	Oversubscription	Servers	100G Uplinks	Throughput (Tb/s)	Utilization (%)
10G (single-lane)	2.8:1	112	4	1.52	47.5
40G using 10Gb/s SERDES	2.8:1	28	4	1.52	47.5
25G (single-lane)	3:1	96	8	3.2	100

# Market Transition

---

## IDC Predictions 2014

- “Number of servers installed in large SP datacenters will reach parity with large internal datacenters in late 2014/early 2015...”

## Dell’Oro Group

- “Top 7 Cloud providers accounted for 20% of Ethernet switch revenue...”

## Deployments of servers and switches are growing in the Cloud provider market

- Ability to seamlessly connect and access data
- Transition from corporate data centers to cloud
- Storage of personal data and information

# BMP Responses

---

Ethernet is widely deployed for server to switch applications in data centers. An Ethernet data rate of 25 Gb/s enables a cost effective interconnect solution enabling 25 Gb/s rack server solutions and intersecting the networking 100Gb/s solutions based on 25 Gb/s SerDes technology.

There will be a significant market potential for 25 Gb/s Ethernet interfaces on servers that optimize the total cost of ownership while meeting the necessary IO bandwidth requirements in data centers.

148 participants attended the “25 Gb/s Ethernet over a single lane for server interconnect ” Call-For-Interest. 59 individuals representing at least 36 companies indicated that they would support the standardization process. It is anticipated that there will be sufficient participation to effectively complete the standardization process including representatives from end-users, equipment manufacturers and component suppliers.

# EF Responses

---

The cost factors for Ethernet components and systems are well known.

Prior experience in the development of 25 Gb/s technology for Ethernet establishes that the specifications developed by this project will entail a reasonable cost for the resulting performance.

In consideration of installation costs, the project is expected to use proven and familiar media.

Network design, installation and maintenance costs are minimized by preserving network architecture, management, and software.

A 25 Gb/s Ethernet interface will maintain a favorable cost balance for server to switch applications.

Energy Efficient Ethernet will reduce the operational costs and the environmental footprint.