# IEEE 802.3
# 2.5G and 5G Ethernet Backplane and Short Reach Copper Cable
# CALL FOR INTEREST

## A Presentation to help form consensus

Contact:

Yong Kim (ybkim at broadcom com)
Thomas Skaar (thomas.j.skaar at seagate com)

# CFI request for "2.5G and 5G Ethernet Backplane and Short Reach Copper Cable"

This is a call for interest to initiate a Study Group to explore the need for 2.5 Gb/s and 5 Gb/s Ethernet on Backplane and short reach copper (e.g. Twin-axial) cabling. We believe there is a market need, driven by enterprise rotational and solid state storage devices that currently use 1 Gb/s Ethernet to migrate to higher speed applications.

Devices currently using 1 Gb/s in backplanes, as well as other 1 Gb/s applications for Ethernet, will also benefit from the new data rates provided by this work.

# CFI Objectives

- To gauge the interest in starting a study group to **investigate a "2.5Gb/s and 5Gb/s Ethernet Backplane and Short Reach Copper Cable" project.**

- This Meeting will **NOT**:
  - Fully explore the problem
  - Choose any one solution
  - Debate strengths and weaknesses of solutions
  - Create a PAR or CSD (including 5 Criteria)
  - Create a standard or specification

- Anyone in the room may speak / vote
- **RESPECT**… **give** it, **get** it.

# Agenda

- Overview — Yong Kim, Broadcom
- Presentations
  - Market Drivers — Thomas Skaar, Seagate
  - Technical Feasibility — William Lo, Marvell
  - Why Now? — Yong Kim, Broadcom
- Q&A – Presenters and Panelists
- Supporters
- Straw Polls

# Overview: 2.5Gb/s & 5Gb/s Ethernet Backplane and TwinAx

- Provide speed upgrade to 1Gb/s connections via support of 2.5Gb/s and 5Gb/s over backplane and short reach copper connections, e.g. TwinAx.
- Provide a speed upgrade to installed base of 1Gb/s Ethernet.
  - 1Gb/s Ethernet is massively successful, and used as the mainstream and the preferred Ethernet connections in modern networks.
  - End devices evolved and need higher speed than 1Gb/s but much lower than 10Gb/s. (e.g. >1Gb/s needing the next higher speed)
  - Modern Ethernet systems could offer 2.5Gb/s speed at relatively similar cost points as 1Gb/s. Similarly, 5Gb/s at a lower cost channel than 10Gb/s.
  - The key market driver is the use of native Ethernet interfaces to the "object" spinning and solid state storage devices, currently at 1Gb/s speed.
- Leverage 802.3bz 2.5Gb/s and 5Gb/s BASE-T project on MAC.
- Leverage 1Gb/s and 10Gb/s SerDes work of 802.3 on PCS, channel, etc.

# Overview: 2.5Gb/s & 5Gb/s Ethernet Backplane and TwinAx

- Support needs of storage transitions to "Object" access and Ethernet.
  - Annual storage unit volume (all types) is huge (500 Million+ /year, total market that includes SATA, SAS, FC, USB, eSATA, PCIe, etc, and HDD and SSD types.)
  - "Object" Storage transition over to Ethernet is expected to be 5~8% of the total storage drives.
  - Allows for rapid deployment and expansion of growing cloud storage infrastructure, especially "cold" or bulk storage use cases in the cloud.
  - Ethernet connected HDD/SSD allows for scalable storage systems while maintaining single network infrastructure – it's just Ethernet!
  - First generation Ethernet "Object" drives use 1GE, and uses native Ethernet connections in drives.
  - Second generation Ethernet drives will use 2.5Gb/s and 5Gb/s over backplanes. Short reach copper solution completes this use case.
  - 2.5Gb/s will be needed within the next 12 months. 5Gb/s will be needed in 24-36 months.

# MARKET DRIVERS

Thomas Skaar, Seagate

# Object Storage - Meaning

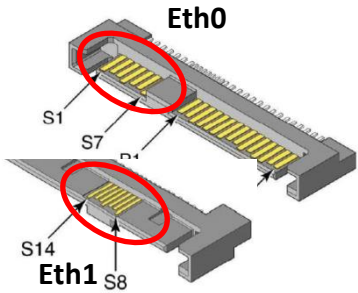
- Traditional storage access has been done through “block” access schemes.
  - Legacy CHS (cylinder-head-sector) physical mapping/addressing proven to be unfriendly and not scaleable. Thus a logical abstraction was defined -- LBA (Logical Block Address) to offer easier access.
  - LBA scheme assumes host access of sequential logical blocks on a given storage device. In SCSI (or SAS), 512 byte – block/sector size (“chunks”) – directly mapped to SCSI commands for a given drive.
  - Object Storage takes the next step in the storage abstraction -- LBA oriented access to “object” access.
- Benefits of Object Storage
  - Serves the needs of cloud storage requirements – “low-cost” and “deep” cold storage that is easy to scale, add, move, and change.
  - Fault tolerance, redundancy, remote storage, etc., just comes from Ethernet system access.

# SAS Block Storage vs Object Storage

## SAS Block Storage Drive

- Standard form factor
  - 2 SAS ports
- SCSI command set
  - data = read (LBA, count)
  - write (LBA, count, data)
  - LBA :: [0, max]
  - data :: count * 512 bytes

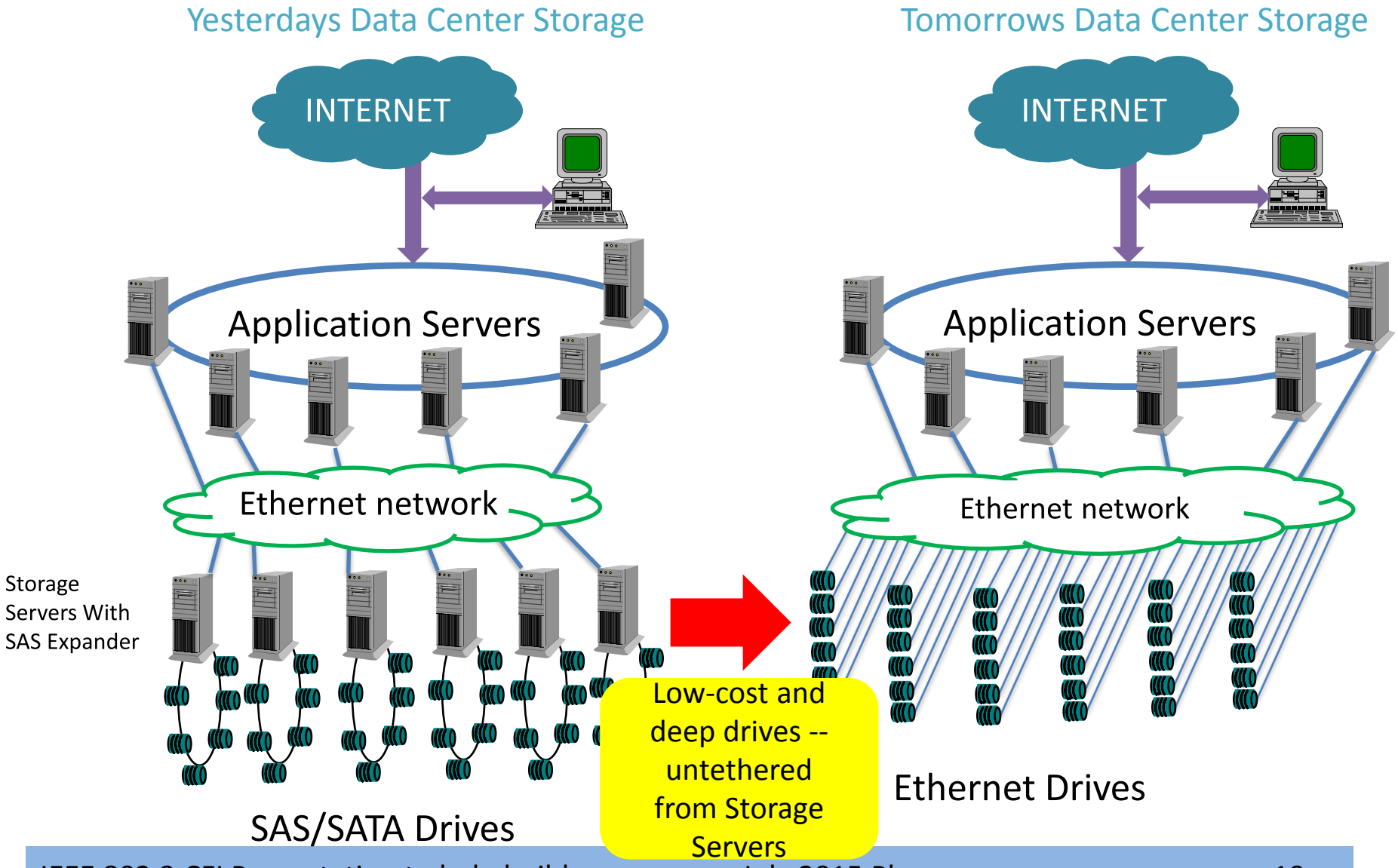
## Ethernet Object Storage Drive



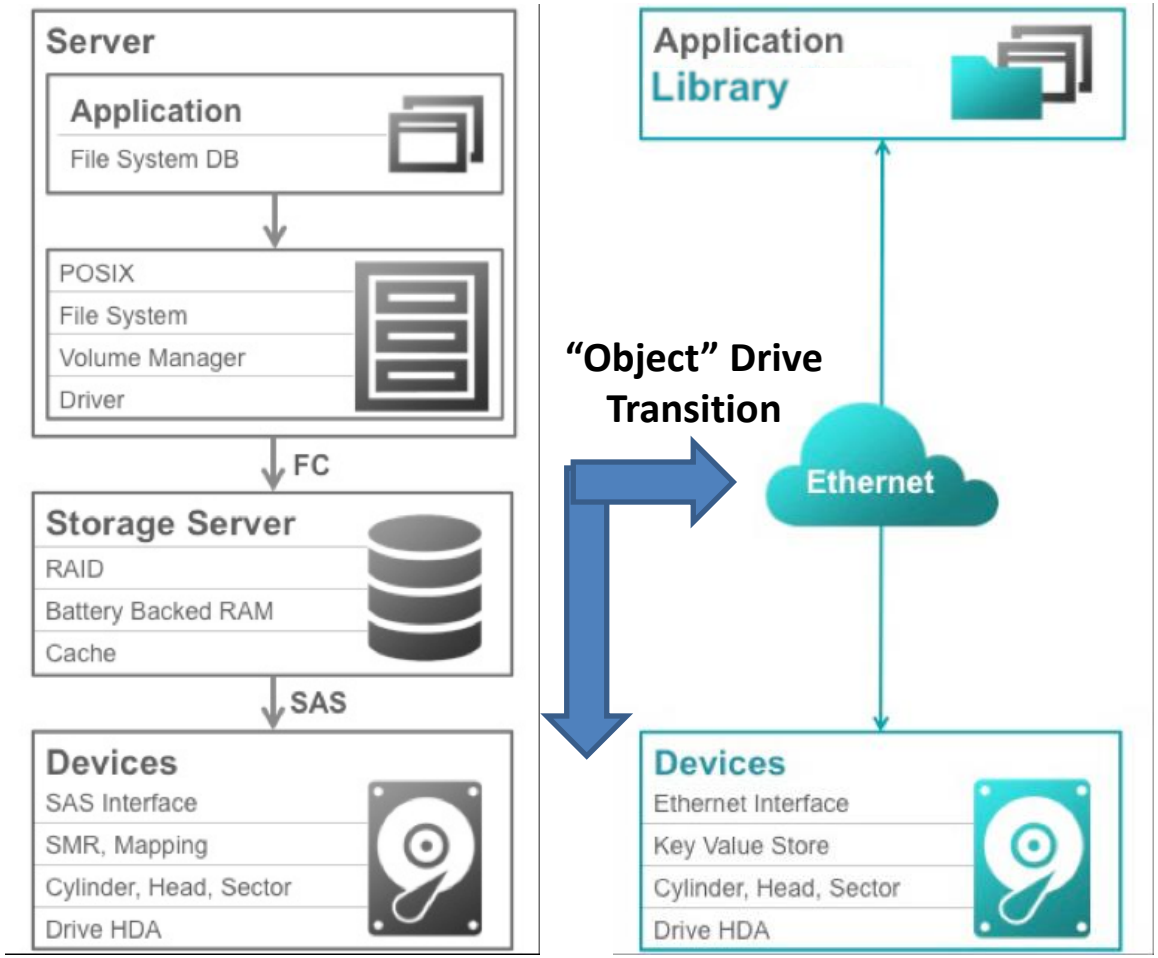

- Standard form factor
  - 2 Ethernet ports (re-use SAS connector)
- Object key/value API
  - value = get (key)
  - put (key, value)
    delete (key)
  - key :: 1 byte to 4 KB
  - value :: 0 bytes to 1 MB

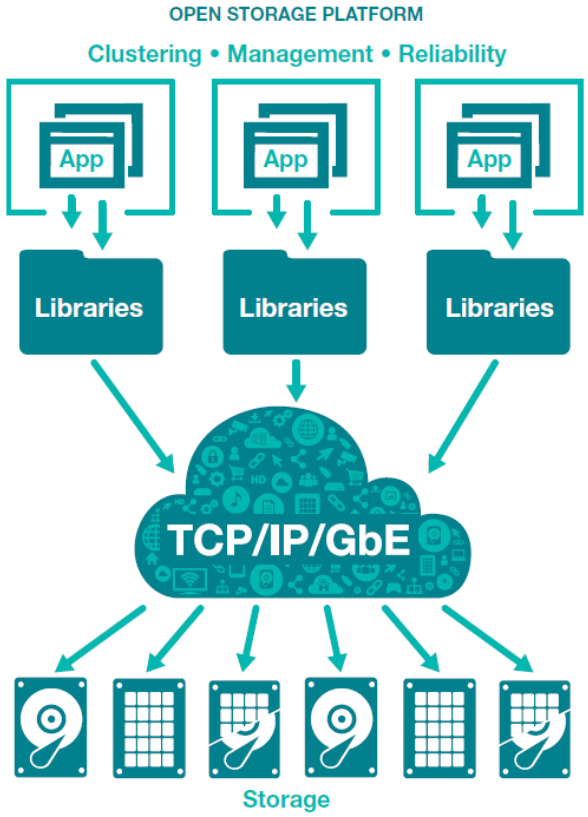
LBA: Logical Block Address (abstraction of CHS – cylinder-head-sector)

# Data Center Architectures



# Ethernet Drives Yesterday and Tomorrow



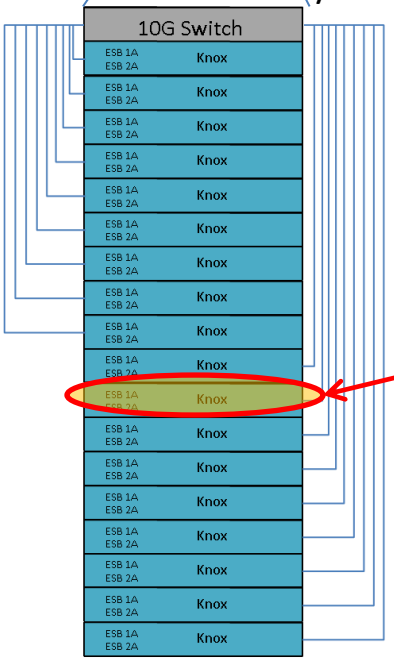



# Use Cases – Cloud Object Storage

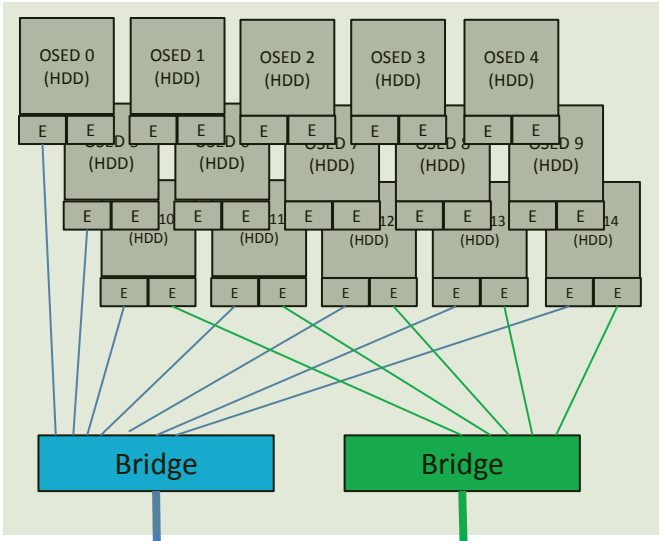
"Rack of Blade Drives w/ ToR Switch"

"Blade Drives"

Ethernet Drives

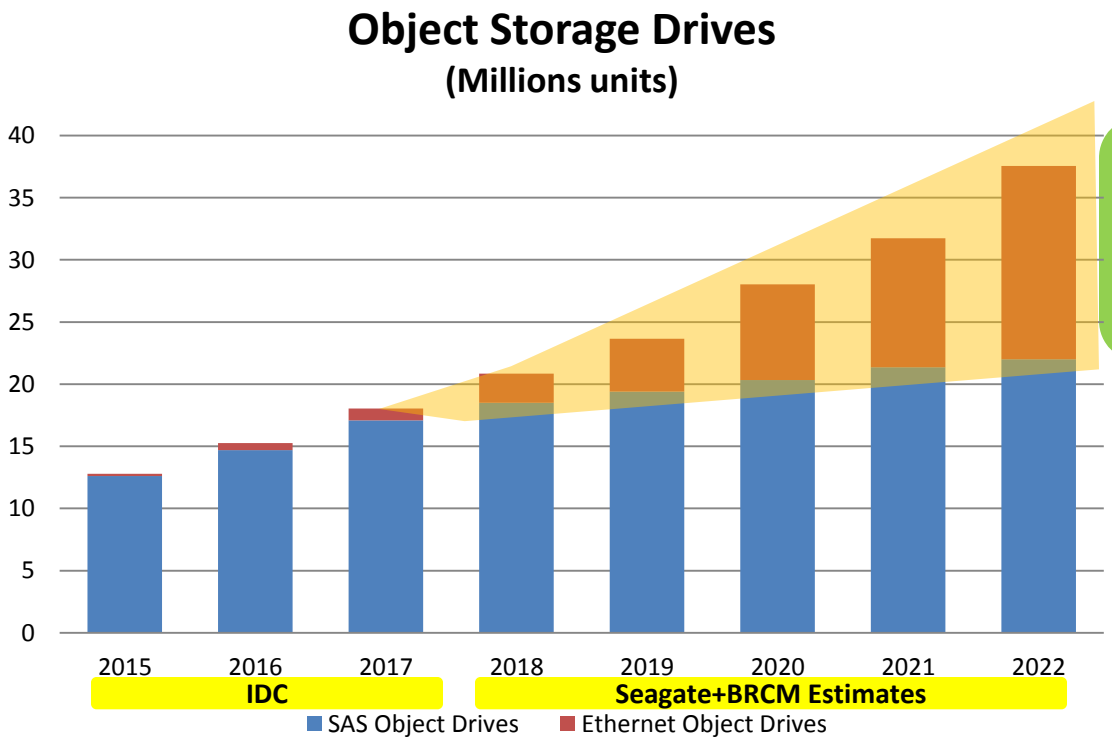



2 X Ethernet cables




2 x Ethernet to ToR Switch

# The Ethernet Object Storage Market

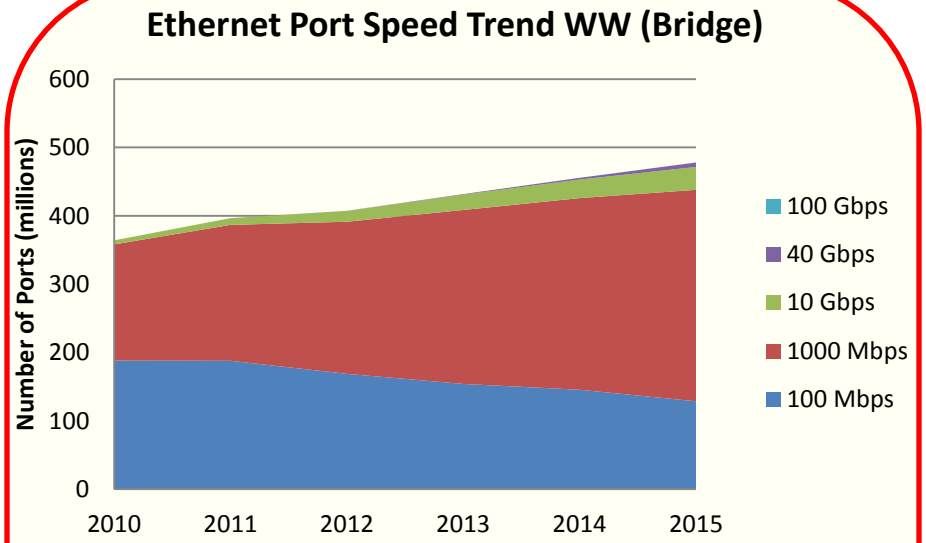


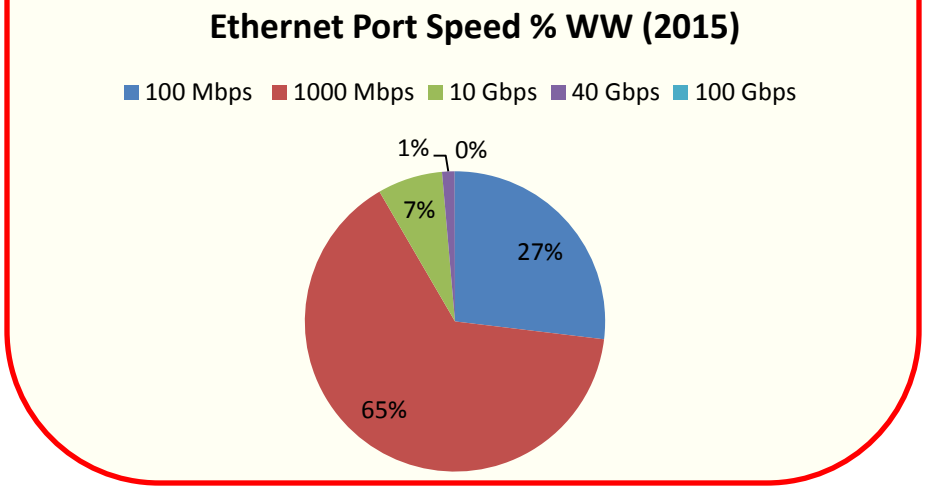

Expecting the object storage growth to be on Ethernet Drives

***Notes:***

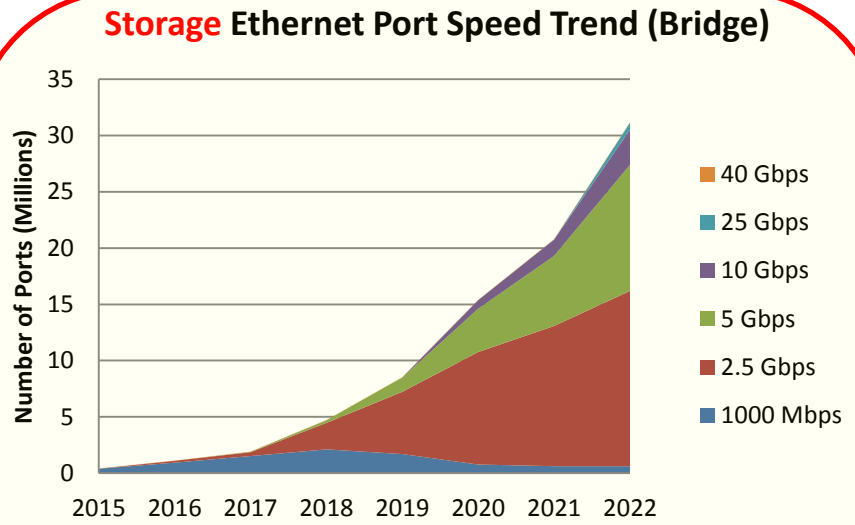
- ***Source: IDC Object Storage market (~2017) plus market estimates by Seagate and Broadcom (2018~2022)***
- ***Each Storage Drives are dual homed (2 ports of Ethernet and companion Switch ports)***
- ***With Standardization, more Ethernet based drives are likely to be adopted (orange colored region) -- reduced complexity and system cost reduction (compared to SAS based object storage drives)***

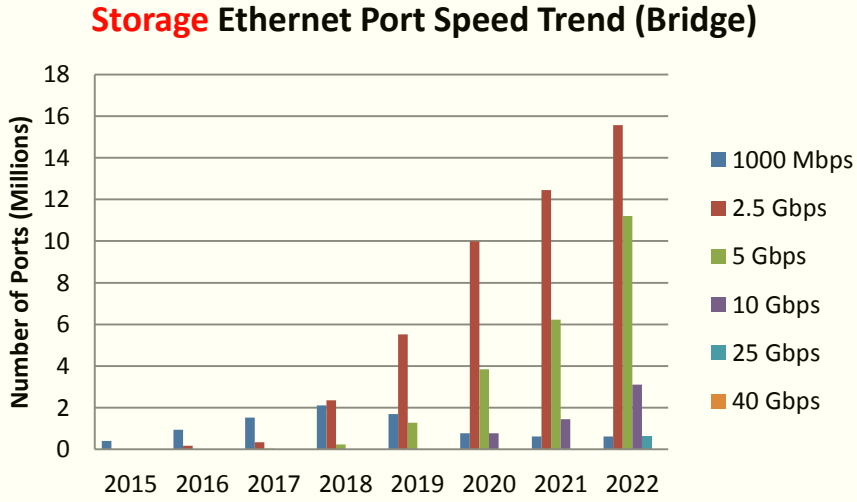
# Ethernet Drive Interface Market Estimate




Source: Dell O'ro Ethernet Switch Report 1Q15



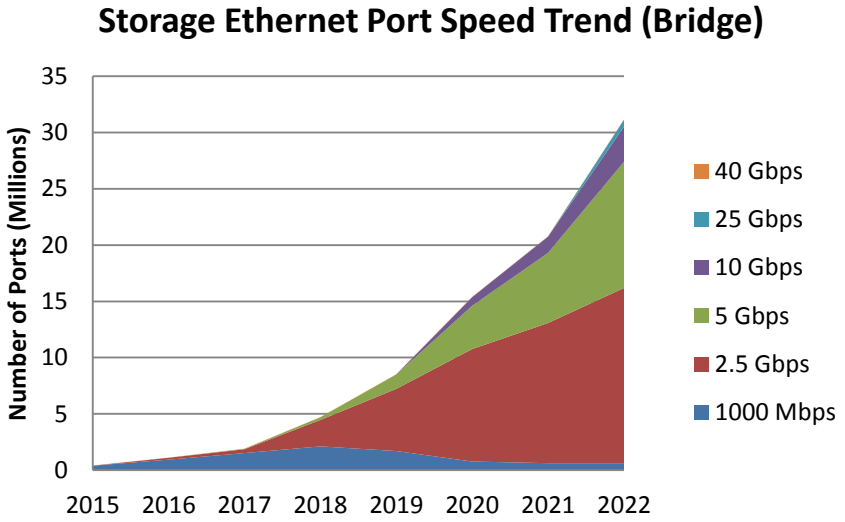

- *Notes: # of Switch ports reflects x2 of drive unit*
- *Market estimates by Seagate and Broadcom (2018~2022)*

# Ethernet Drive Interface Market Estimate

Total # of Switch ports per year is around 460 million ports per year.

Stable single digit CAGR in # of ports.

A good indication of actual # of Ethernet ports that gets used.




2.5G/b and 5Gb/s rates serve the mainstream HDD Object storage market for the long term.

- *Notes: # of Switch ports reflects x2 of drive unit*
- *Market estimates by Seagate and Broadcom (2018~2022)*

**Source: Dell O'ro Ethernet Switch Report 1Q15**

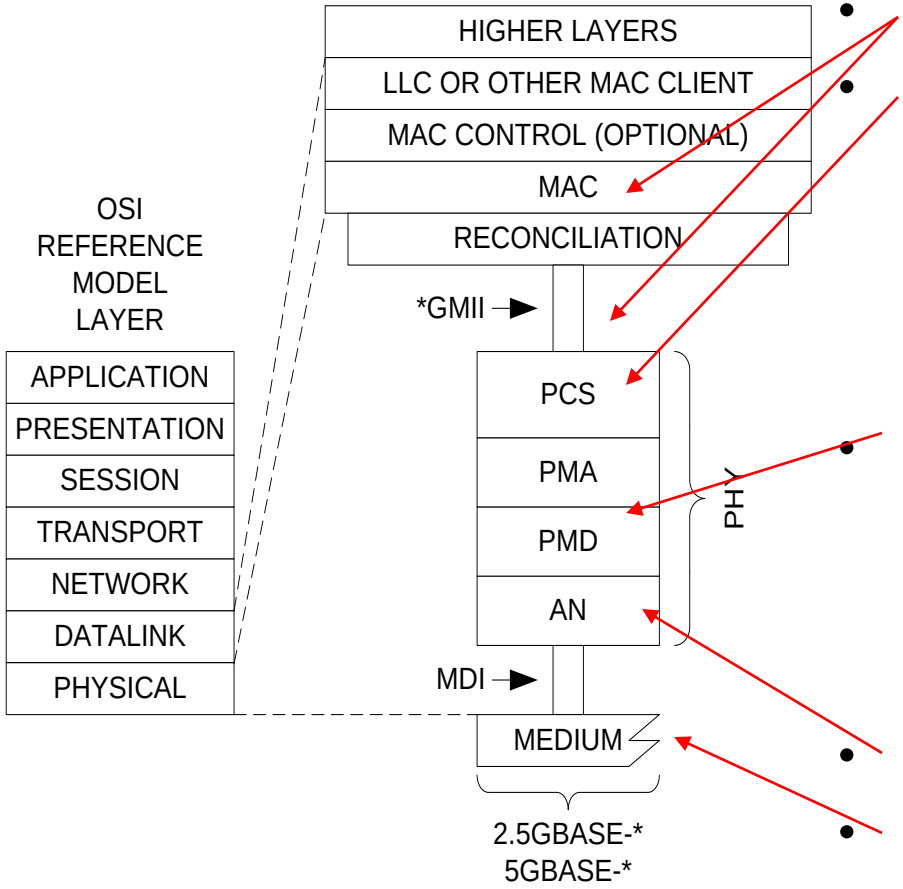
# Market Driver Summary

- “Object” storage has emerged
- “Low-cost and Deep” Cloud cold storage will adopt Object HDD storage on Ethernet.
- Bandwidth need for HDD (regardless of SAS/SATA I/O speed) is ~2x the sustained bandwidth, now at round ~175 MBytes/s.
- 2.5G and 5G Ethernet speeds will serve the mainstream HDD needs for the long term.
- Native Ethernet I/O on Object storage to be 5~8% or more of overall Ethernet switch connections over time.

# TECHNICAL FEASIBILITY

William Lo, Marvell

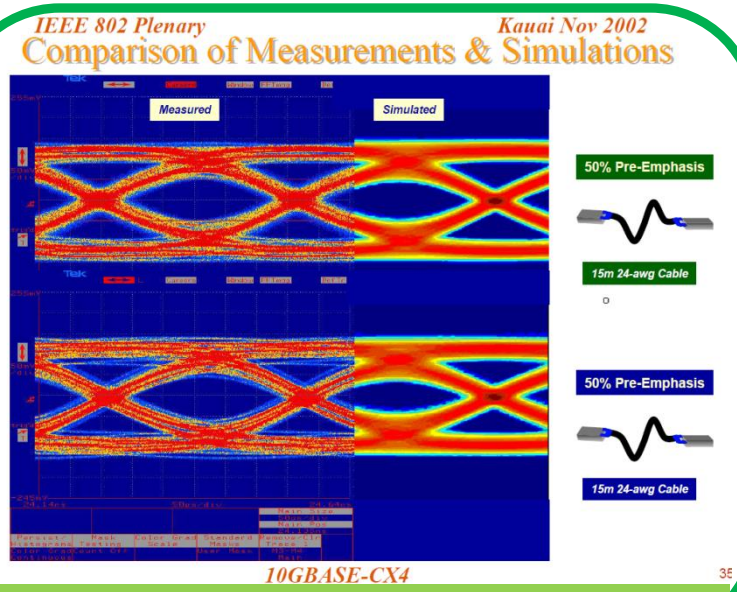
# Wealth of Existing Proven Technology



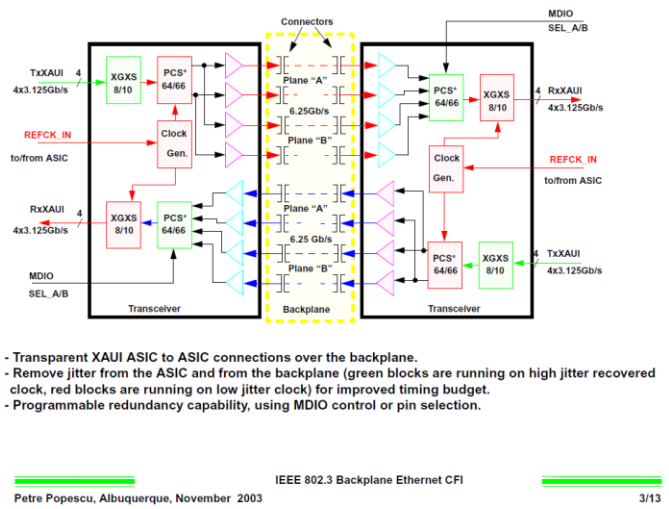

- P802.3bz defining 2.5G/5G MAC & *GMII
- PCS is feasible in existing technology:
  - Possible options:
    - Speed up 1000BASE-X
    - Single lane XAUI
    - Slow down 10GBASE-R
    - Something else…
- PMA/PMD feasible in existing technology:
  - Possible options:
    - Single lane BASE-KX4
    - Reduced speed BASE-KR
    - Something else…
- Auto-Negotiation – Clause 73
- Significant knowledge base of backplane and short reach copper channel and component characteristics

- Energy efficient Ethernet approaches are well known for 1GE and 10GE

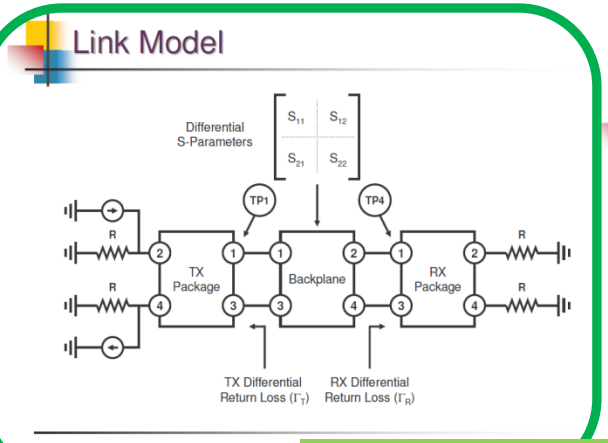
# Some Past Backplane Contributions in IEEE 802.3

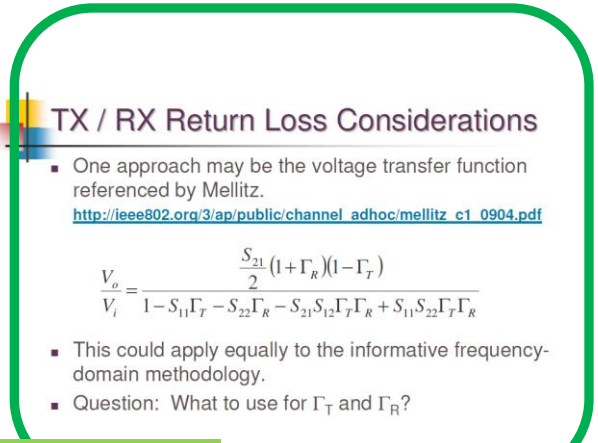



**CFI for P802.3ak 10GBASE-CX4, w/ XAUI, 2002**

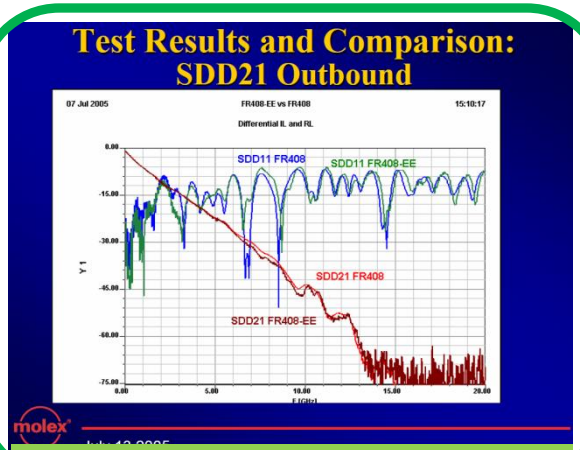



**CFI for P802.3ap Backplane Ethernet w/ R XAUI, 2003**




**P802.3ap Channel Ad hoc, 2005**




**P802.3ap FR408 Backplane, 2005**

# 2.5G/5G Technology for installed base well understood

- 1G backplane leverage
- XAUI and 10G backplane & TwinAx leverage
  - 10GBASE-KR
  - 10GBASE-KX4
  - 10GBASE-CX4
- Fibre Channel Rates: 1G, 2G, 4G, 8G, 16G, 32G,
- SATA, SAS Rates: 3G, 6G, 12G, SATA Express
- Ethernet Rates: 1G, 10G, 25G, 40G, 100G
- Conclusions:
  - **Technical feasibility demonstrated in multiple platforms**
  - **Leverage prior work to help specify PCS, PMA, Channel**
  - **Extend auto-negotiation to include new speeds**

# WHY NOW?

Yong Kim, Broadcom

# Why Now?

- The application is imminent!
  - 1G Ethernet connected Object HDD (and SSD in the future) are in use, and increasing.
  - The use of 2.5 Gb/s speed in the next 12 months, and needs 5 Gb/s speed in the next 24-36 months.
  - Cost and power sensitive -- to serve “Low-cost and Deep” Cold Cloud Storage
  - Balanced cost of overall system, including backplane, reach, compared to transition directly to 10G from 1G. Spinning drives sweet spot is 2.5G, migrating to 5G.
- Solutions are available and shipping at 1Gb/s in volume. ← Why do this now.
  - Many Switching & PHY silicon already has 2.5Gb/s. Some has 5Gb/s rates.
  - The market needs an interoperable, open standard -- avoid poor user experience from incompatible exposed interfaces.
  - One of the key benefit sought is adding Auto-negotiation of 2.5Gb/s, and 5Gb/s to existing rates of 1 Gb/s, 10 Gb/s, 25Gb/s (.3by), 40Gb/s, and beyond, and allowing for the parallel detect of legacy 1Gb/s in use.
- The Ethernet Ecosystem has been very successful
  - Open and common specifications that ensures Interoperability

# Why Now? Summary

- Proliferation and expansion of Ethernet family to serve the "Object" storage architecture over Ethernet is relevant to all Ethernet speeds and to both HDD and SSD.
  - Top end (10 Gb/s and up) and bottom end (1 Gb/s) of the Ethernet speeds are covered for object storage market.
  - The mainstream Ethernet speed (2.5Gb/s and 5Gb/s) for HDD is not covered by current 802.3 Stds. This project will fill the gap in the middle.
- Object storage using 1Gb/s Ethernet is being deployed now – and 1Gb/s already is a bottleneck.

# Possible topics for Study Group

- BER/Reliability requirements
  - SAS/SATA BER coordinated with IEEE 802.3 BER considerations.
- Clause 73 Auto-Negotiation
- Compatibility Considerations
  - Coexistence with other backplane(s) between 1G to 10Gb
- Energy Efficiency and EEE
- Market needs for the short reach copper (“TwinAx”)
- XGMII/GMII/PCS compatibility to 802.3bz

# QUESTIONS?

## Q&A Panelists

Systems, Storage, and Drive Experts

Thomas Skaar, Seagate

Yong Kim, Broadcom

Ethernet and PHY Experts

William Lo, Marvell

Adam Healey, Avago

Vasudevan Parthasarathy, Broadcom

# Supporters

**38 from 24 Entities**

Systems

Ghani Abbas, Ghani, Ericsson
Anthony Calbone, Seagate
Mark Carlson, Toshiba
Jackie Chang, HP
David Chen, Nokia
Alvin Cox , Seagate
John D'Ambrosia, Dell
Jon Lewis, Dell
Joel Goergen, Cisco
Philip Kufeldt, Toshiba
Larry McMillan, Western Digital
Thomas J Skaar, Seagate
Paul Suhler, HGST

Components

Rob Davis, Mellanox
Steve Gorshe, PMC-Sierra
Mark Gustlin, Xilinx
Adam Healey, Avago
Yong Kim, Broadcom
William Lo, Marvell
Andrew Lillie, Intel
Kent Lusted, Intel
Christopher Mash, Marvell
Greg McSorley, Amphenol TCS
Tom Palkert, Molex
Vasudevan Parthasarathy, Broadcom
Steve Sedio, FoxConn
Santanu Sinha, Broadcom
Scott Summers, Molex
Pat Thaler, Broadcom
Paul Wassenberg, Marvell
Greg Winn, Avago
Yaron Zimerman, Marvell

Ecosystem

Rita Horner, Synopsys
Arthur Marris, Cadence
Alon Regev, Ixia
Bogdan Tenea, Ixia

End Users & Others

Ali Ghiasi, Ghiasi Quantum LLC
Paul Nikolich, Self

# STRAW POLLS

# Call-for-Interest Consensus

- Should a study group be formed for “2.5G & 5G Ethernet on Backplane and TwinAxial Cables”?

- Yes:  53 = 34 + 19,
- No: 0,
- Abstain: 6 = 4 + 2,

- Room Count: 61 =41 + 20

# Participation

- I would participate in a “2.5G & 5G Ethernet on Backplane and TwinAxial Cables” study group in IEEE 802.3

  Tally: 20 = 15 + 5

- My company would support participation in a “2.5G & 5G Ethernet on Backplane and TwinAxial Cables” study group

  Tally: 15 = 10 + 5

# Future Work

- Ask 802.3 at Thursday's closing meeting to form a "2.5G & 5G Ethernet on Backplane and TwinAxial Cables" study group

- If approved:
  - 802 EC informed on Friday of formation of the study group.
  - First study group meeting would be during September 2015 IEEE 802.3 Interim meeting.

# THANK YOU!