Expanding CERN's EOS Storage with Seagate Kinetic Disk Technology
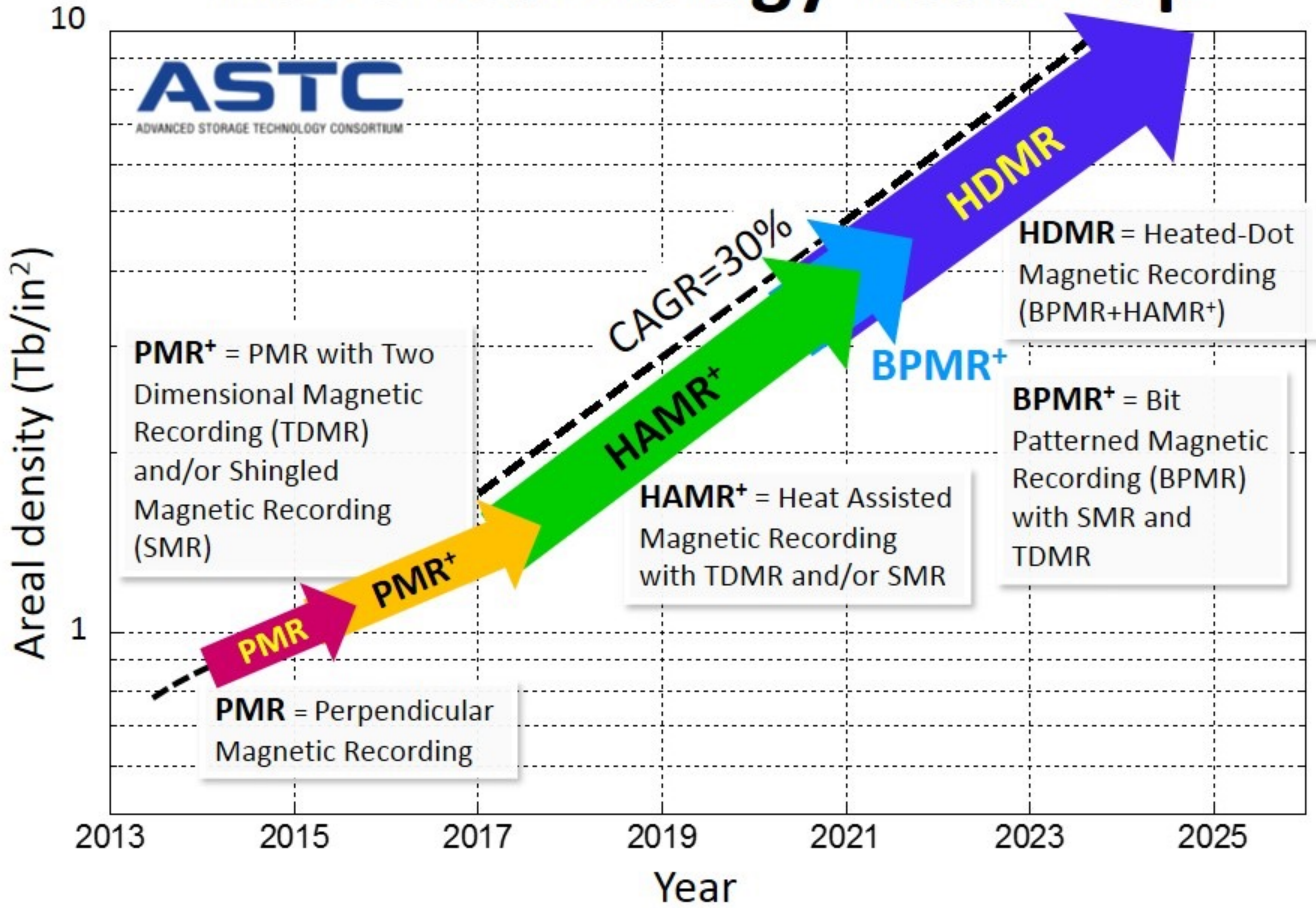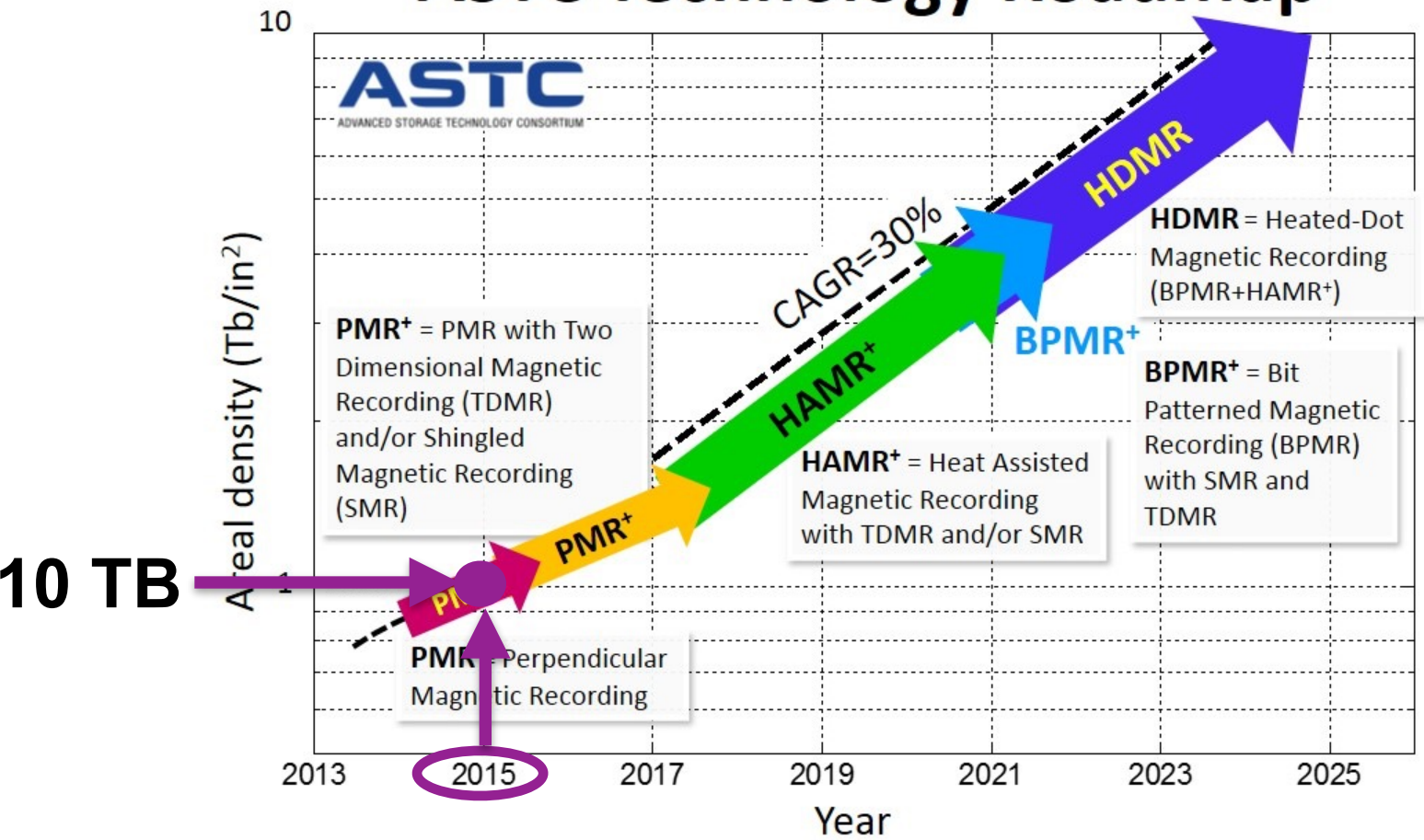
James Hughes
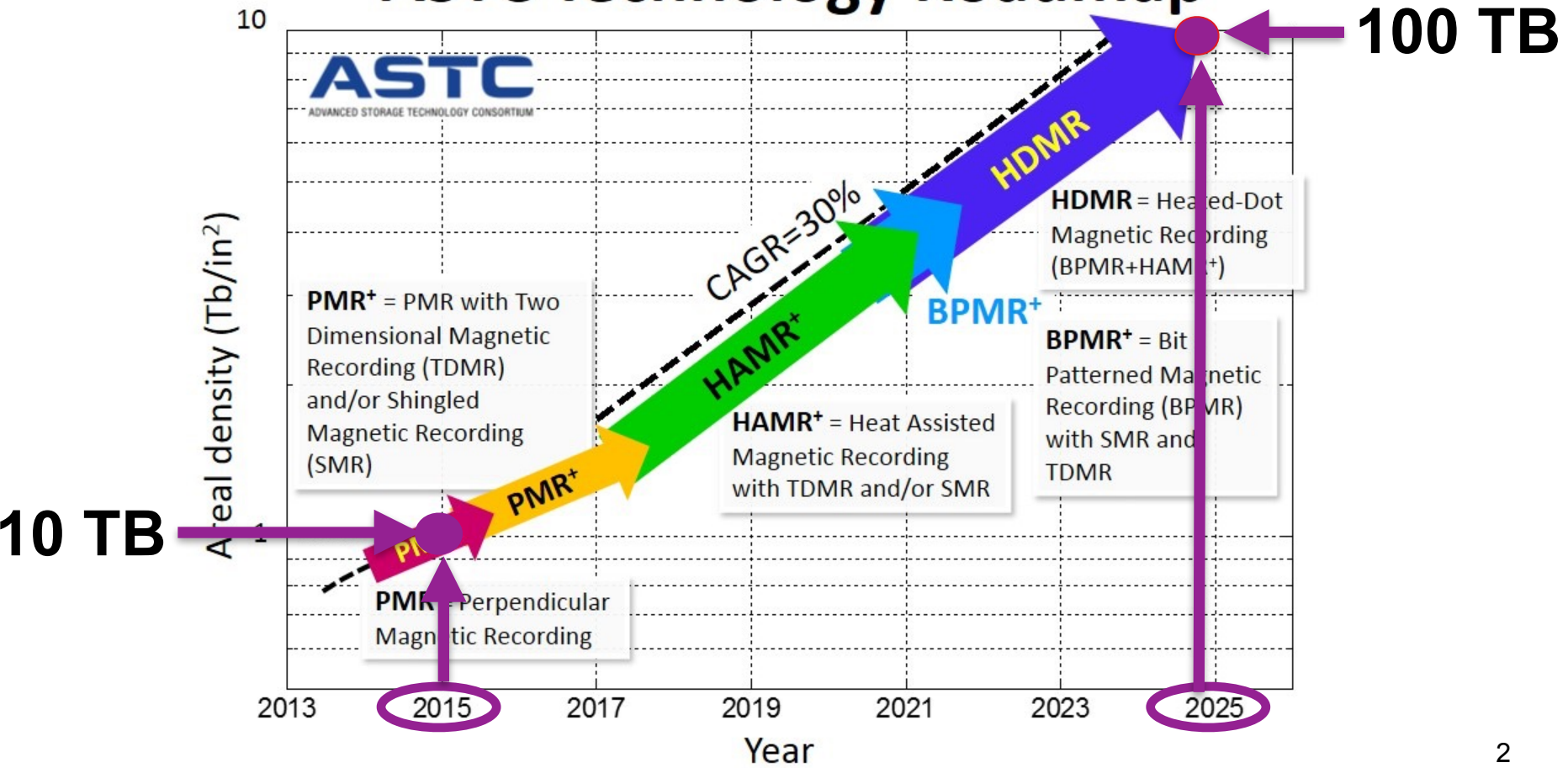**Paul Lensing**

1

ASTC Technology Roadmap

PMR+ = PMR with Two Dimensional Magnetic Recording (TDMR) and/or Shingled Magnetic Recording (SMR)

PMR = Perpendicular Magnetic Recording

HAMR+ = Heat Assisted Magnetic Recording with TDMR and/or SMR

BPMR+ = Bit Patterned Magnetic Recording (BPMR) with SMR and TDMR

HDMR = Heated-Dot Magnetic Recording (BPMR+HAMR+)

CAGR=30%

PMR
PMR+
HAMR+
BPMR+
HDMR

Areal density (Tb/in^2)

Year

2

ASTC Technology Roadmap

10

Areal density (Tb/in²)

10 TB

PMR⁺ = PMR with Two Dimensional Magnetic Recording (TDMR) and/or Shingled Magnetic Recording (SMR)

PMR = Perpendicular Magnetic Recording

PMR⁺

HAMR⁺

HAMR⁺ = Heat Assisted Magnetic Recording with TDMR and/or SMR

CAGR=30%

BPMR⁺

BPMR⁺ = Bit Patterned Magnetic Recording (BPMR) with SMR and TDMR

HDMR

HDMR = Heated-Dot Magnetic Recording (BPMR+HAMR⁺)

2013   2015   2017   2019   2021   2023   2025

Year

2

ASTC Technology Roadmap

Data Center Stack

Conventional vs Kinetic

Storage Application
File System DB

POSIX
File System
Volume Manager
Driver

FC

Storage Server
RAID
Battery Backed RAM
CACHE

SAS

Devices
SAS Interface
SMR, Mapping
Cylinder, Head, Sector, Drive HDA

Application
Kinetic Library

Ethernet

Devices
Ethernet Interface
Key Value Store
Cylinder, Head, Sector
Drive HDA

3
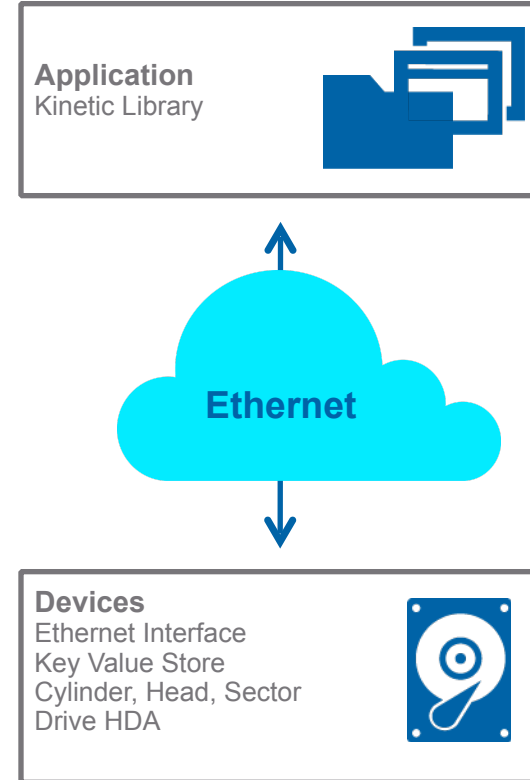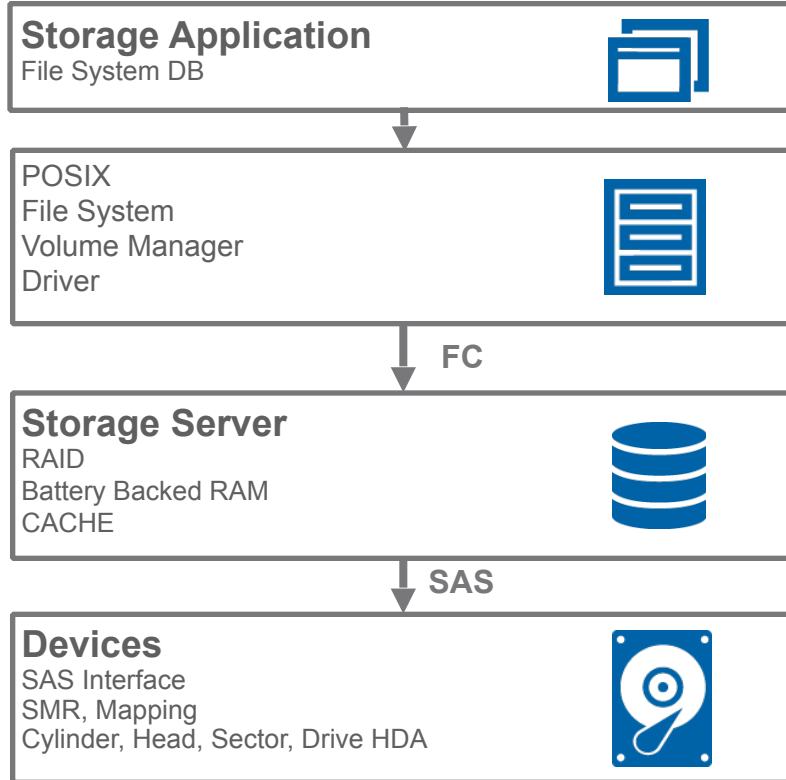
Green is good

- 1200w
- $7,800

Conventional

- 400w
- $4,200

Kinetic

Green is good

- 1200w
- $7,800

Conventional

- 400w
- $4,200

Kinetic

4

Green is good

- 1200w
- $7,800

Conventional

- 400w
- $4,200

Kinetic

Conventional EOS
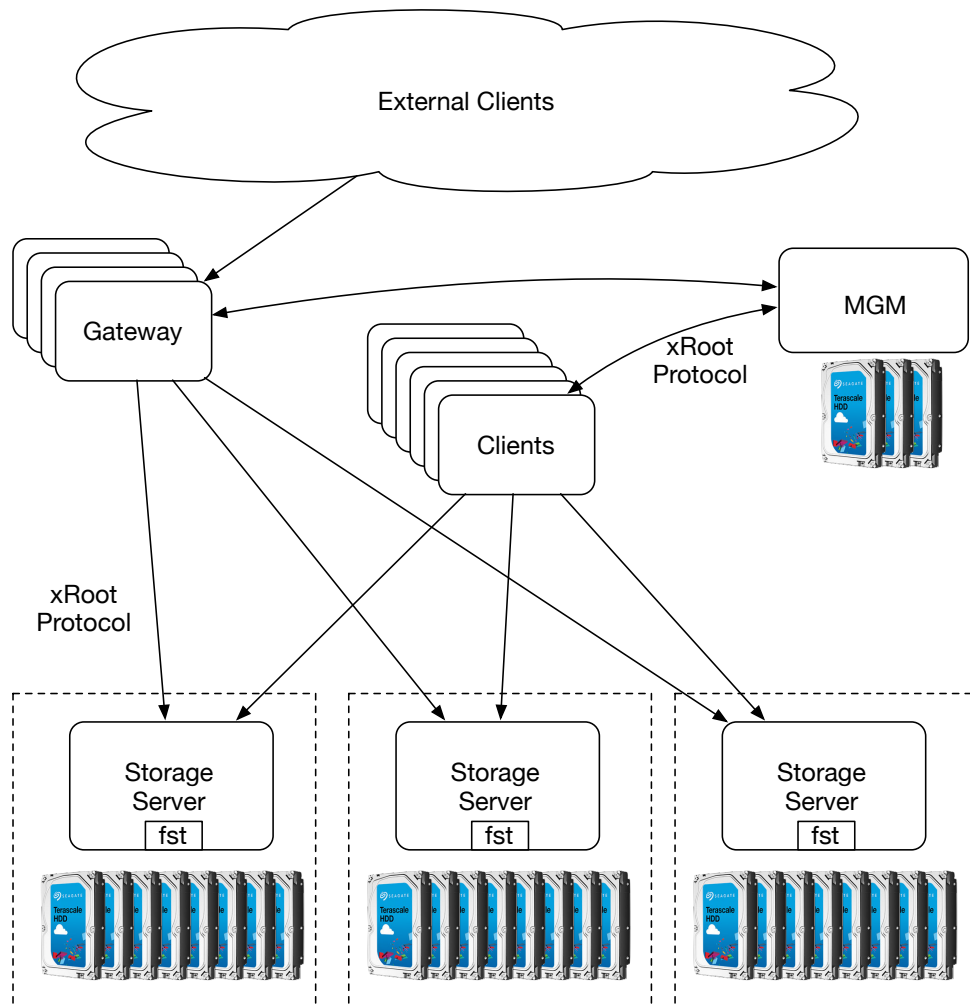
Gateways

Clients

- External Clients Through Gateway
- Internal Clients Connecting Directly

Metadata Server

- MGM (EOS Management Module)
- Namespace, Pool Configuration, Placement & Access Scheduler

Storage Servers

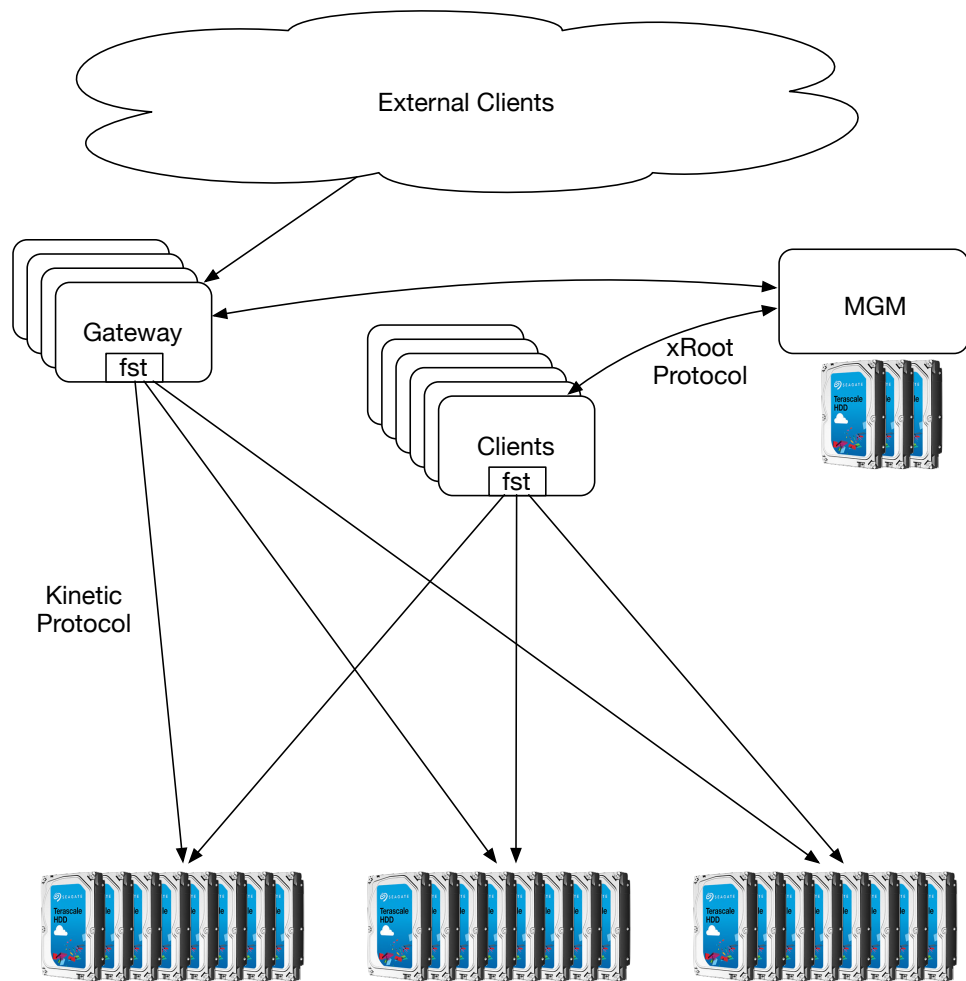- FST (EOS File Storage Module)
- Expose Local Disks to EOS

EOS with Kinetic

Kinetic Drives

- Ethernet Storage
- Not Locked to Single Server
- Accessible from Multiple FST

Internal EOS Clients

- Run FST Module for Direct Access

External EOS Clients

- Continue Using Gateway Servers

1PB test system

Twenty One Chassis

- 1u12 from Supermicro
  – 2x10Gbe

252 Kinetic drives

- 504 IP addresses
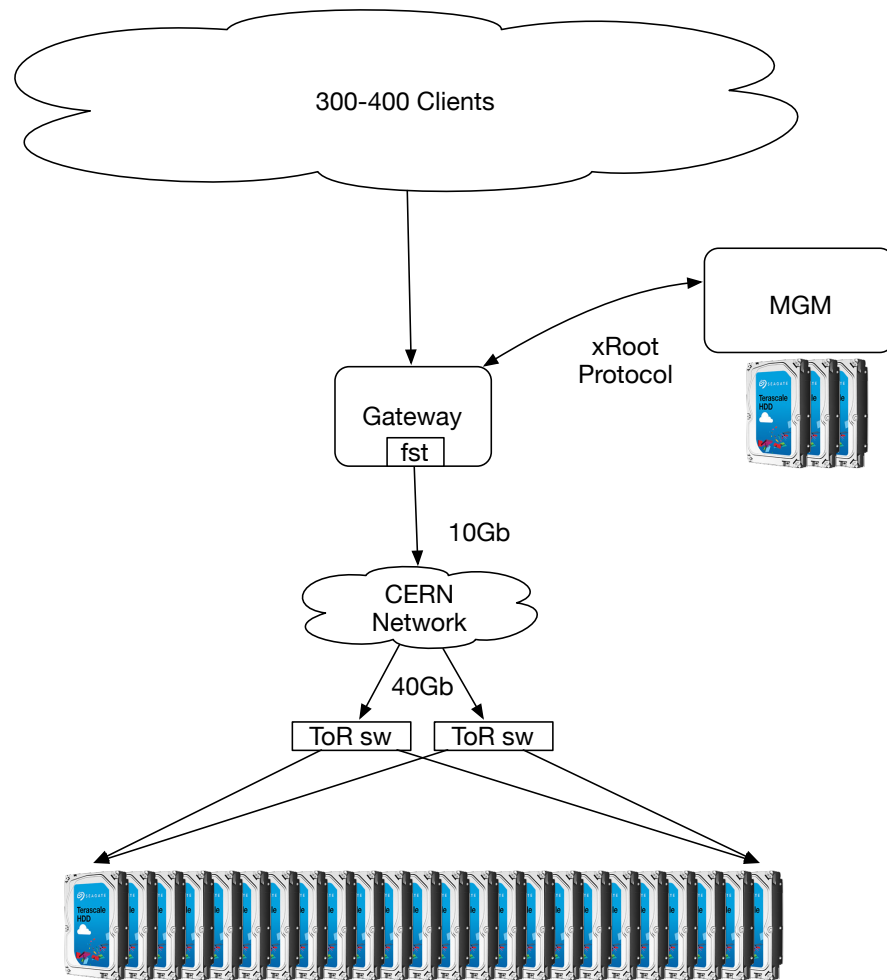- 126GBytes of RAM cache

Two ToR Switches

- 40Gbe to CERN

Initial Test Setup
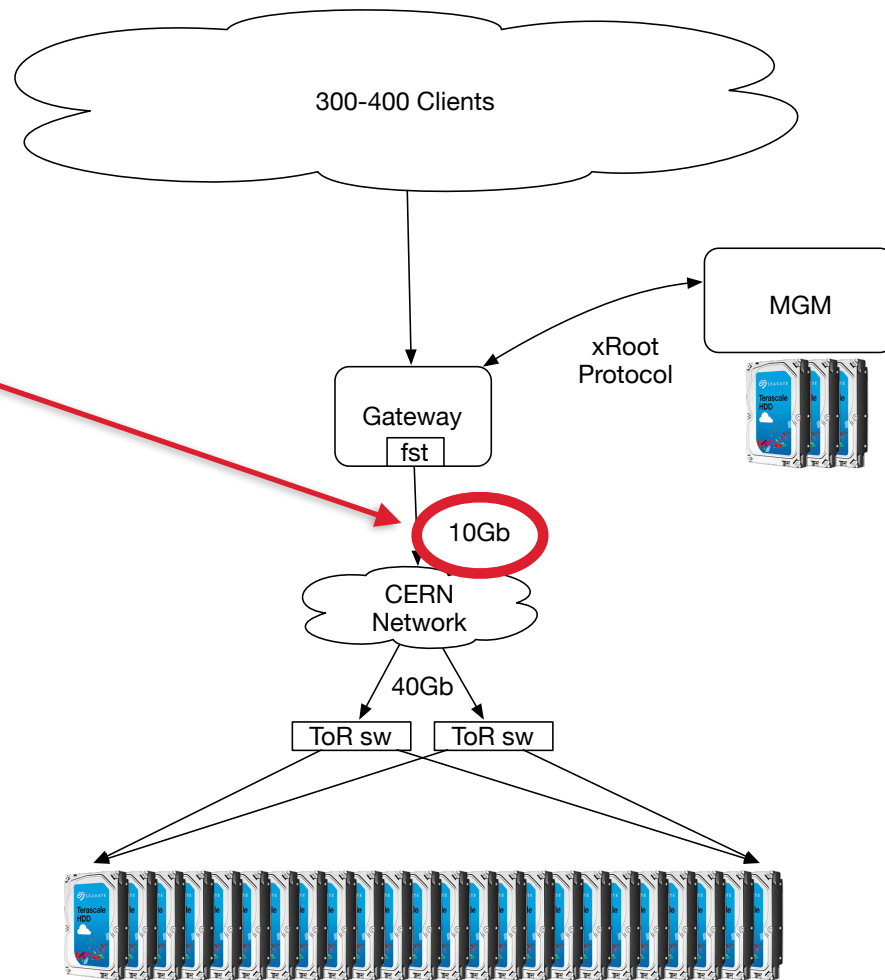
Single Gateway Test

- 10Gb/s Over a 10Gb/s Network

Initial Test Setup

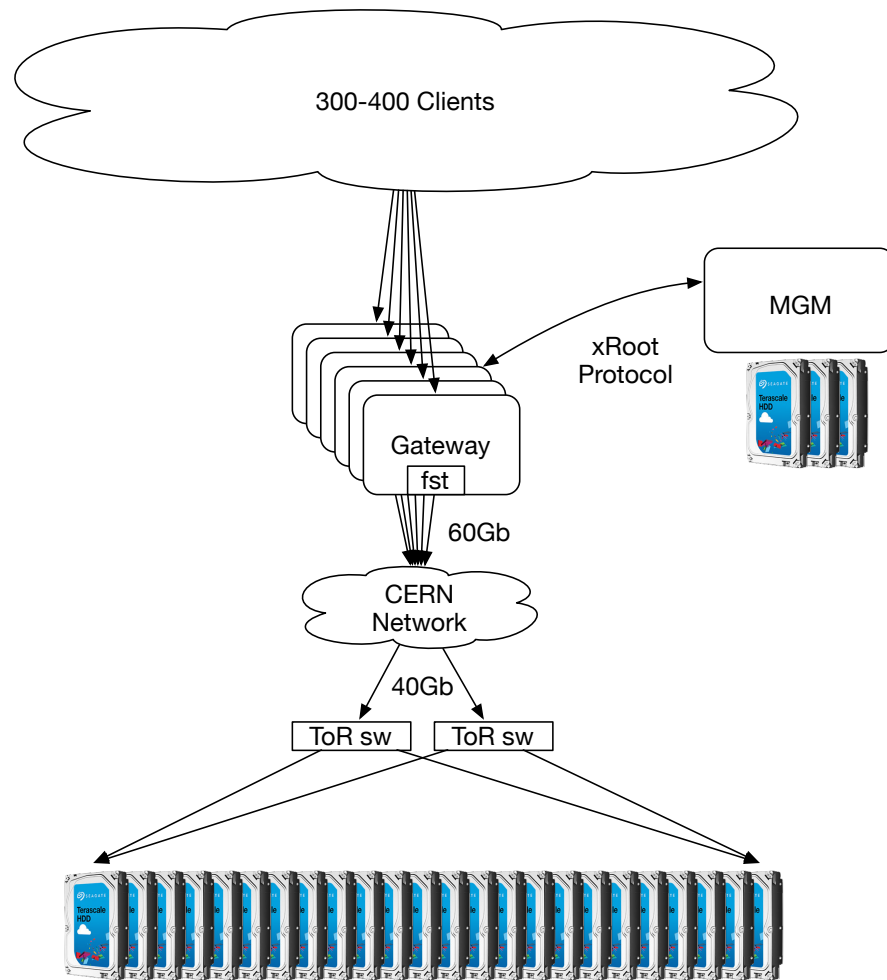## Single Gateway Test

▪ 10Gb/s Over a 10Gb/s Network

Initial Test Setup

Single Gateway Test

- 10Gb/s Over a 10Gb/s Network

Six Gateway Test
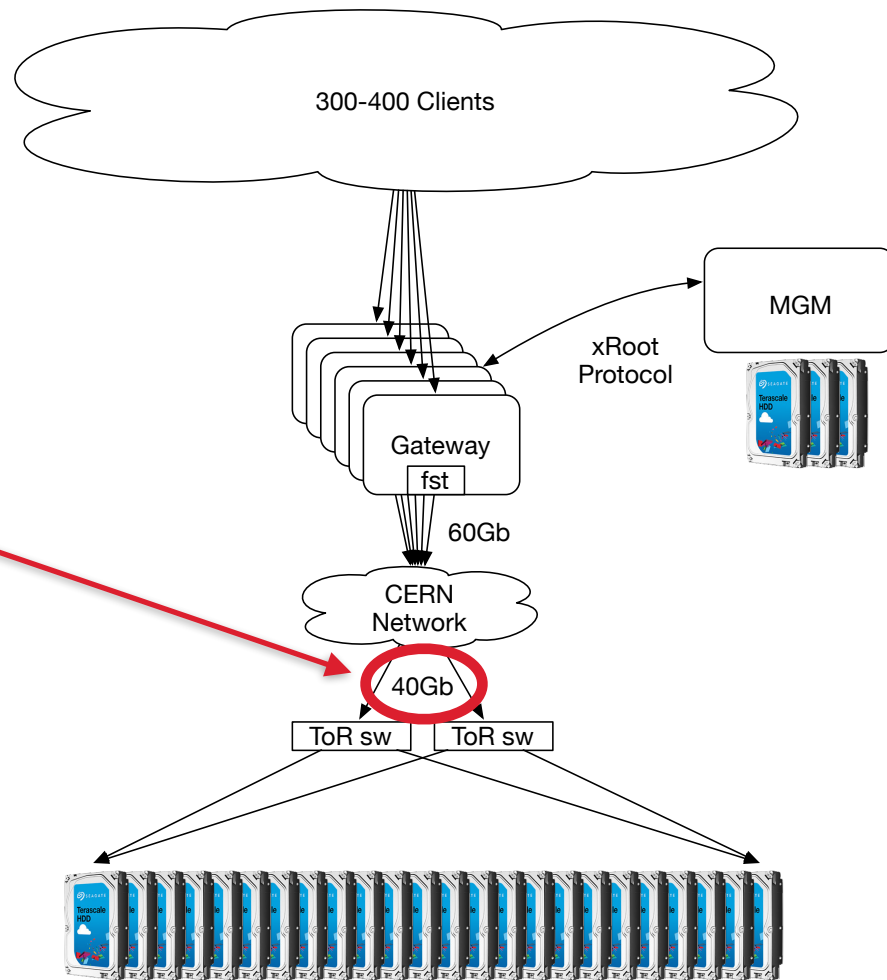
- 32Gb/s Over a 40Gb/s Network

Initial Test Setup

Single Gateway Test

- 10Gb/s Over a 10Gb/s Network

Six Gateway Test

- 32Gb/s Over a 40Gb/s Network

Initial Test Setup
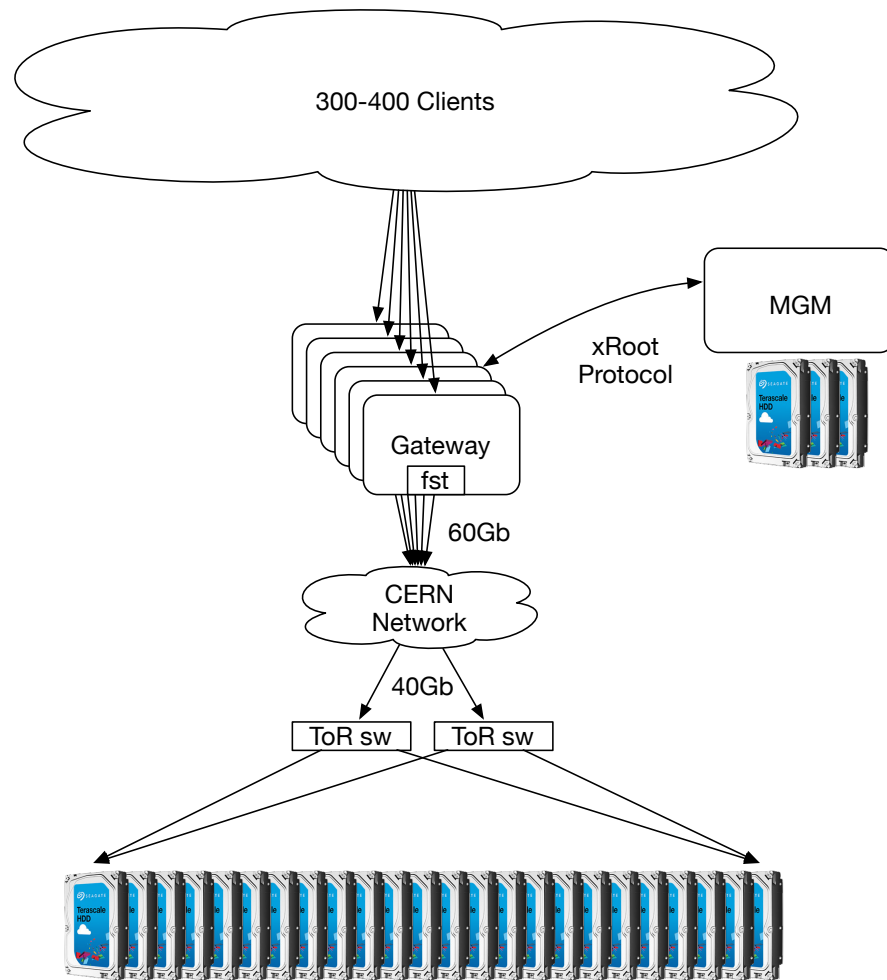
Single Gateway Test

- 10Gb/s Over a 10Gb/s Network

Six Gateway Test

- 32Gb/s Over a 40Gb/s Network

More work to do

- Deploy Kinetic/FST on Datacenter Clients
- Add Cluster Logic and Erasure Coding
- Begin Testing With Real Data
- Add Second Cluster in 2016
  – Next Generation Drives

Conclusion

On The Path to Prove TCO Gains of Kinetic at CERN

- No change to Existing EoS User Applications
- Looking Forward to Storing LHC Data

Thank You

HTTP://Developers.Seagate.Com