

Looking to the Future: What Hyperscale Cares About

Ross Stenfort, Meta

Agenda

- ❖ Background
- ❖ Hyperscale Features Perspective
- ❖ Future of Flash
 - Form Factors
 - Open Compute Project
 - Datacenter SSD
 - Hyperscale Boot SSD
 - Flexible Data Placement Mode



**Meta's mission is to give people the power to build
community and bring the world closer together.**



Meta @ Scale

Hardware Optimized For User
Experience

Hyperscale Feature Perspective

Important

- Scalable & Flexible
- High volume & Low cost
- Power & Thermal Efficiency
- Hot-swappable & Serviceable
- Performance per TB & Quality of Service
- Security

Less Important

- Backwards compatible
- Maximum density
- Peak Performance (Peak IOPs/BW)

Form Factors

M.2 Challenges

- Power Limited to 8.25W
- Performance
 - Need connector designed for PCIe® 5.0 and beyond
- Capacity
 - Need more NAND Placements
- Serviceability
- Security

Market Needs

- Connector designed for PCIe 5.0 and beyond
- Scaling
 - Performance
 - Power
 - Thermal
- Density in 1 OU
- Serviceability
- Security

Market Solution: E1.S

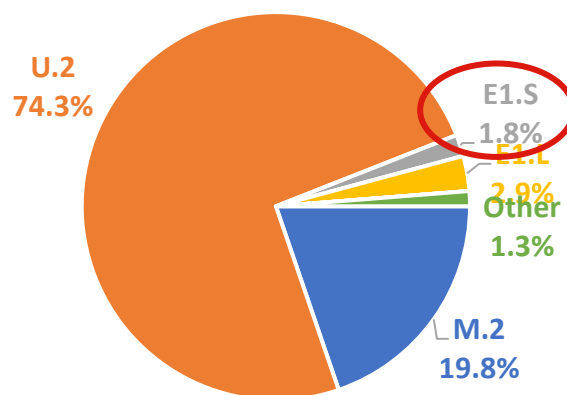
- Connector designed for PCIe 5.0 and beyond
- Enabling Scaling of:
 - Performance
 - Power
 - Thermal
- Density in 1 OU
- Serviceability
- Security

PCIe® Enterprise/ Datacenter Form Factor Trends

TRENDFOCUS

2021

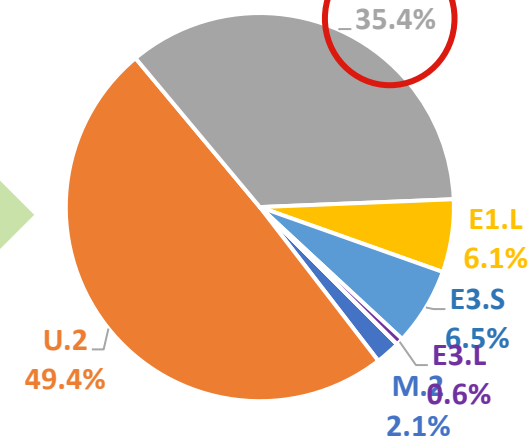
78 Exabytes



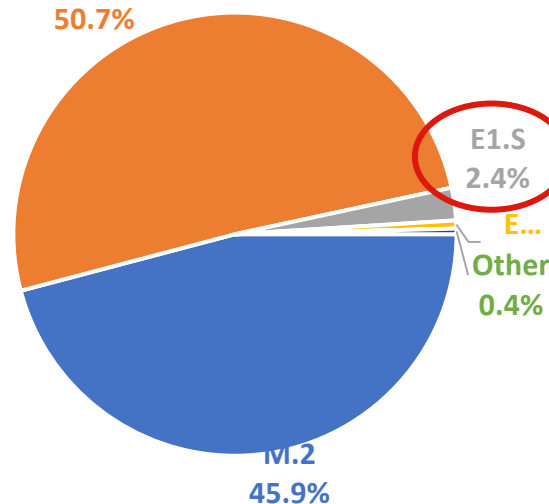
By 2026 more Exabytes of E1.S are predicted to ship than total ePCIe Exabytes in 2021

2026

246 Exabytes

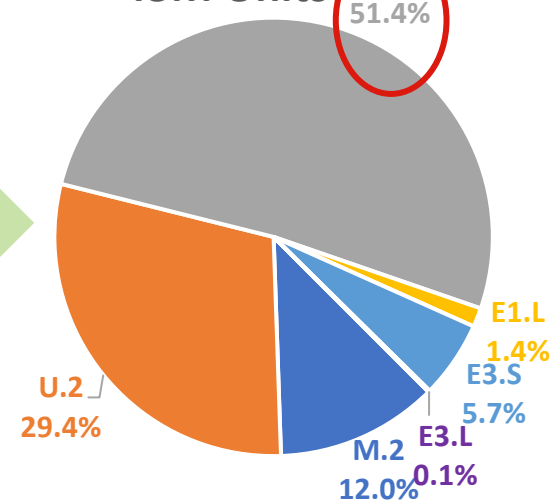


26M Units



E1.S is predicted to have largest number of ePCIe units by 2026

43M Units



E1.S market growth is very significant.

Note: Data excludes SSD consumption where companies buy NAND and build SSDs for internal use.

**2019 Flash
Industry
Perspective:
Data Drive
Perspective**

Challenge

- Customer requirements are confidential
 - Standards have many optional features
 - What do customers really require?
 - Limited competition
 - Access to specifications are limited to customer/supplier relationships
- SSD industry highly fragmented with lots of SKUs
 - Many customers ask for similar but different features
- SSD Suppliers have finite resources
- 3rd party test providers don't know what customers require



Result

- Delayed schedules
- Quality suffers
- Difficult product and feature decisions

Customers and Suppliers struggle.

Solution: OCP NVMe™ Datacenter SSD Specification

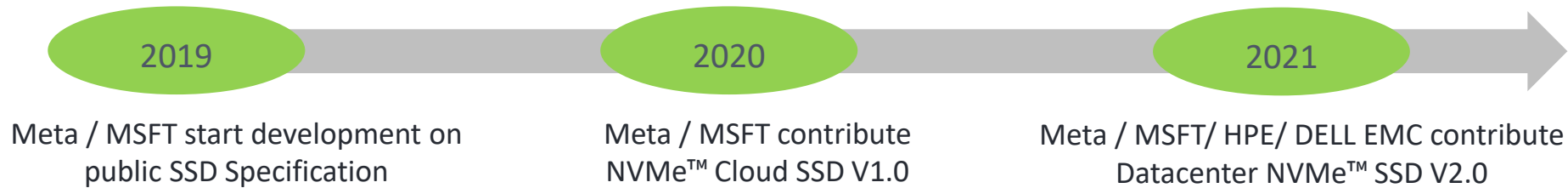
Datacenter NVMe SSD Specification Goals:

- Aligns SSD needs and requirements between Hyperscale/OEMs and SSD makers
- Share flash learnings based on Hyperscale and Enterprise deployments
 - Example: Latency Monitoring
- Provide everything needed to build a SSD for Hyperscale/OEM

Datacenter NVMe SSD Specification Coverage:

- NVM Express®
- PCI Express®
- Reliability
- Thermal
- Security
- Form Factor
- SMART Logs
- Power
- SMBUS

Timeline:

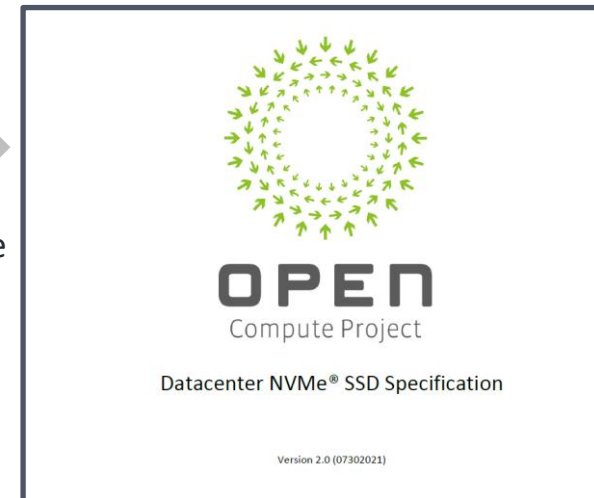


Open Compute Project Contribution links:

- <https://www.opencompute.org/documents/nvme-cloud-ssd-specification-v1-0-3-pdf>
- <https://www.opencompute.org/documents/datacenter-nvme-ssd-specification-v2-0r21-pdf>

Open-Source Tooling (NVMe-CLI/ plugins / OCP):

- <https://github.com/linux-nvme/nvme-cli>



OCP Datacenter NVMe SSD Specification is an industry collaboration win.

2021 Flash Industry Perspective: Boot SSD

❖ Challenges

- Hyperscale and Client both use 2280 M.2s for boot SSDs
- Hyperscale features/needs are different than client boot SSDs
- Hyperscale Boot SSD requirements are confidential

Metric	Client	Hyperscale
Device Idle Time	More	Less
Power Savings Features	Required	Not Required
Performance Metric	Fresh Out-Of-Box	Sustained
Monitoring Capabilities	Not Important	Important
Endurance Requirements	Low	High

❖ Result

- Hyperscale boot SSD needs are misaligned to client SSD market
- Products struggle to align to hyperscale needs

Customers and Suppliers struggle.

Solution: OCP Hyperscale NVMe™ Boot SSD Specification

OCP Hyperscale NVMe Boot SSD Specification Goals:

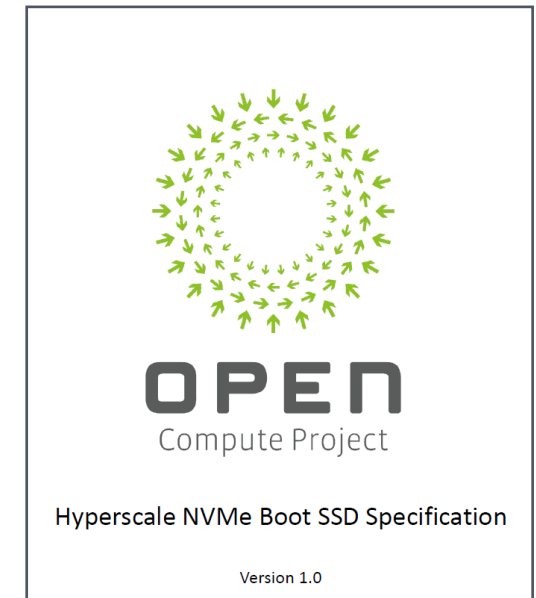
- Aligns NVMe Boot SSD needs and requirements between Hyperscale and SSD makers
- Share flash learnings based on Hyperscale deployments
- Provide everything needed to build a boot SSD for hyperscale deployments

OCP Hyperscale NVMe Boot SSD Specification Coverage :

- NVM Express®
- Reliability
- Security
- Sustainability
- PCI Express®
- Thermal
- Form Factor
- SMART Logs
- Power
- SMBUS

Open Compute Project Contribution link:

<https://www.opencompute.org/documents/hyperscale-nvme-boot-ssd-specification-v1-0-pdf>



OCP Hyperscale NVMe Boot Specification is an industry collaboration win.

Problem: Write Amplification (WA)

- ❖ Write Amplification Negative Impacts:
 - Read/Write performance and QOS
 - Primary contributor to sequential and random write performance being different
 - Device wear
 - Useable Capacity
- ❖ Current Solutions:
 - Over provisioning (~1991)
 - Increases cost as there is capacity that is not usable.
 - De-allocate/TRIM (~2007/2008)
 - Data is poorly placed on media causing data on the media to be re-located.

**Hyperscale Driven Innovation and Collaboration
is enabling a significant step forward in reducing WA.**

The Solution: Flexible Data Placement (FDP – TP4146)

❖ Enables the following:

- Device to advertise virtual media information
- Host to provide media placement hints
 - Optimizes data placement on media

❖ Benefits

- Improved Performance/ QOS
- Reduced Device Wear/ Life
- Increased Useable Capacity



For more information come hear Google/Meta NVM Express Talk:
Hyperscale Innovation with Flexible Data Placement

Summary

Important Hyperscale Features

- Scalable and Flexible
- High volume and Low Cost
- Power and Thermal Efficiency
- Serviceability
- Performance per TB & Quality of Service
- Security

Hyperscale Solutions

- E1.S Form Factor
- OCP Datacenter NVMe™ SSD Specification
- OCP Hyperscale NVMe™ Boot SSD Specification
- Flexible Data Placement Mode

**Innovation + Collaboration + Scalability =
Solutions for the next generation of Flash**



A perspective view of a long, narrow server room aisle. The aisle is flanked by rows of server racks on both sides. The racks are illuminated with a strong blue light, creating a cool, technological atmosphere. The floor is dark and reflective. In the far distance, at the end of the aisle, a person is visible, standing near a bright white light source that illuminates the end of the hallway. The overall scene conveys a sense of depth and scale within a data center.

Thank You