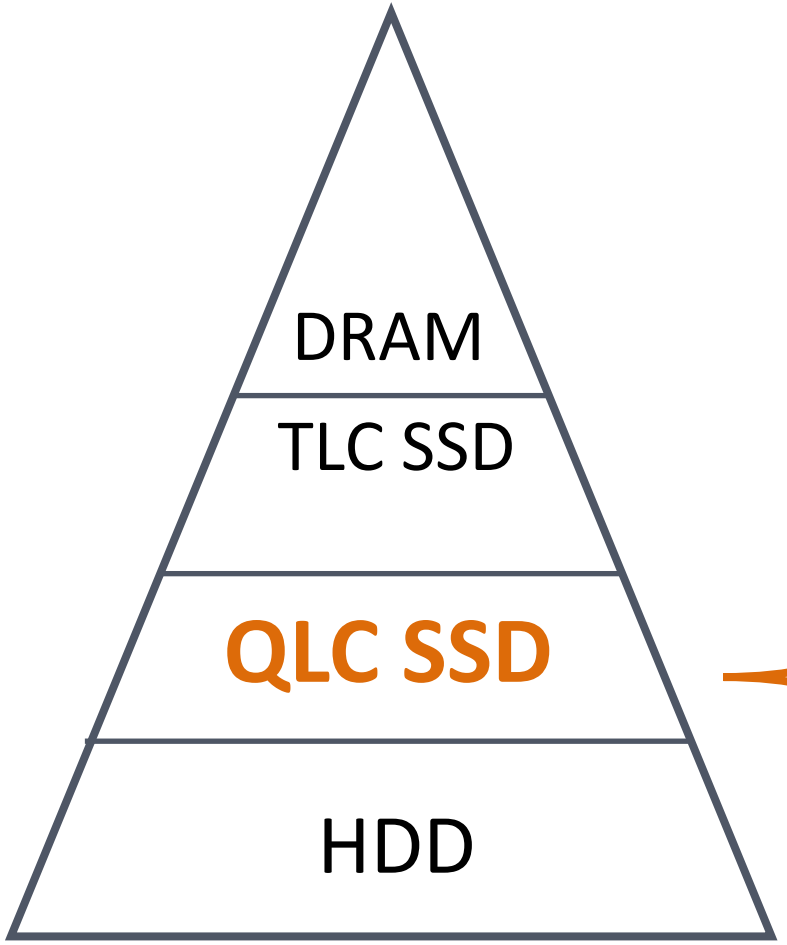
Storage Industry Update

Ross Stenfort, Meta

Agenda

- High Capacity QLC Storage
- OCP in the real world

WHY QLC?



- Increases Storage Chassis Density
- Increases Device TB / W
- Improves Performance / TB Scaling
  - Power Based Performance Scaling

QLC Creates Storage Tier Between TLC SSD and HDD

QLC Challenges

- What should the performance/power be?
- What should the form factor be?

What should QLC Performance & Power be?

Performance and Power Guidance

Performance

Read Bandwidth (MB/s)/ Useable TB	Write Bandwidth (MB/s) / Useable TB
32	0
28.8	0.8
25.6	1.6
22.4	2.4
19.2	3.2
16	4.0
12.8	4.8
9.6	5.6
6.4	6.4
3.2	7.2
0	8

Read:
32 (MB/s) / Useable TB

Write:
8 (MB/s) / Usable TB

Power

- 128 TB: 20W
- 256 TB: 30W

Background: High Capacity Form Factor

- Issue
  - Current E1 and E3 form factors are unable to effectively scale to higher capacities

- E2: Next Gen QLC Form Factor
  - Objectives:
    - Minimum 64 Nand Package Sites
    - Leverage EDSFF E1 and E3 form factors and learnings

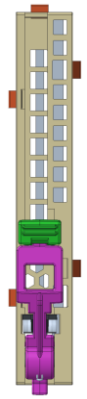
EDSFF Comparison: E2, E3, E1

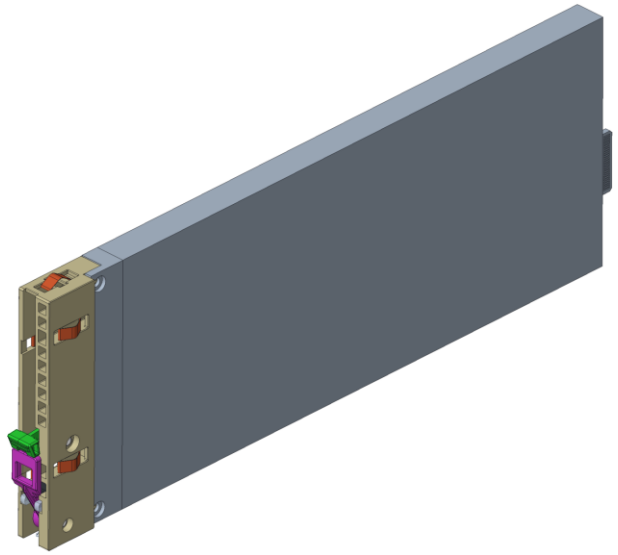
	E2	E3.L 1T	E1.L 9.5	E3.S 1T	E1.S 9.5
Protocol	NVMe	NVMe	NVMe	NVMe	NVMe
Transport	PCIe	PCIe	PCIe	PCIe	PCIe
Connector	SFF-TA-1002	SFF-TA-1002	SFF-TA-1002	SFF-TA-1002	SFF-TA-1002
Pinout/electricals	SFF-TA-1009	SFF-TA-1009	SFF-TA-1009	SFF-TA-1009	SFF-TA-1009
Number of packages	64+	32-48	32-48	16-32	8-16
Enclosure Length	200mm	142.2mm	318.75mm	112.75mm	118.75mm
Enclosure width	76mm	76mm	38.4mm	76mm	33.75mm
Enclosure thickness	9.5mm, symmetrical	7.5mm asymmetrical	9.5mm, symmetrical	7.5mm asymmetrical	9.5mm, symmetrical
Connector alignment	27.7mm from Datum	27.7mm from Datum	12.415mm from Datum	27.7mm from Datum	12.415mm from Datum
LEDs	Amber, Green opposite PCB side	Amber/Blue, Green Same PCB side	Amber, Green same PCB side	Amber/Blue, Green Same PCB side	Amber, Green opposite PCB side
Latch/Carrier mount	ledge, 2 thru holes	3 sides, 4 threaded holes	ledge, 2 thru holes	3 sides, 4 threaded holes	ledge, 2 thru holes
EMI/ESD contact point	Bottom Strike pad, latch area	Side contact pads, mounting holes	Bottom Strike pad, latch area	Side contact pads, mounting holes	Bottom Strike pad, latch area

E2 merges the best attributes and learnings from E3 and E1 to enable a scalable high capacity QLC form factor.

Industry Collaboration Results: E2

- E2 (SFF-TA-1042) V1.0 is ratified
- E2 enables efficient high-capacity flash storage
- Overview
  - Capacity Scaling:
    - Up to 1+ PB
  - EDSFF Connector Scaling:
    - x4 PCI Gen 6 and beyond
    - ~80W
  - Resource efficient simplicity:
    - Single PCB
  - Thermally optimized enclosure
- Enables
  - Dense Chassis:
    - 40 Devices
  - Thermal:
    - Cooling with low airflow above 25W
  - Performance
    - Scales with capacity
  - Serviceability





OCP Datacenter NVMe® SSD Specification Update

2020

- OCP NVMe Cloud SSD Spec v1.0
- Supporters:
  - **Meta**
  - **MSFT**

2021

- OCP Datacenter NVMe SSD V2.0
- Supporters:
  - Meta
  - MSFT
  - **Dell EMC**
  - **HPE**

2023

- OCP Datacenter NVMe SSD V2.5
- Supporters:
  - Meta
  - MSFT
  - Dell EMC
  - HPE
  - **Google**

2024

- OCP Datacenter NVMe SSD V2.6
- Supporters:
  - Meta
  - MSFT
  - Dell EMC
  - HPE
  - Google

2025

- **Coming Soon** OCP Datacenter NVMe SSD V2.7
- Supporters:
  - Meta
  - MSFT
  - Dell EMC
  - HPE
  - Google

Improvements in V2.7

- QLC High Capacity SSD Improvements
  - Telemetry, E2, etc..
- Device Measured Power
- Continuous Improvements to Telemetry / Debug

OCP Datacenter NVMe SSD Enables:
More Features, Better Quality and Faster

Key Features for Managing and Deploying at Scale (1 of 2)

Improvements Based On Deployment @Scale

- **OCP Health Information Extended Log**
  - Telemetry Metrics based on deployments at scale
- **OCP Latency Monitoring Feature**
  - Isolates, monitors, debugs latency spikes at scale
- **OCP Formatted Telemetry for Human Readable Logs**
  - Customer usable telemetry with improved security
- **Open-Source OCP NVMe CLI**
  - Open-Source Tooling
- **DSSD Power State Support**
  - Simplifies power state control

Key Features for Managing and Deploying at Scale (2 of 2)

- **Hardware Component Log**
  - Hardware manufacturing information is available to customer
- **Device Self Test Improvements**
  - Failing Segment codes is universal rather than supplier/product dependent
- **Device Self-Reported Power**
  - Device Power measurements are simplified in qualification and at scale

Open-Source SSD Qualification Test Cases

- Meta OCP Framework
  https://github.com/opencomputeproject/ocp-diag-autoval
- Meta OCP Storage Tests
  https://github.com/opencomputeproject/ocp-diag-autoval-ssd

Improves Development & Qualification Timelines

OCP Storage Project Update

2023	NVMe HDD rev. 1.0	[spec]	Microsoft, Seagate, Western Digital
	Datacenter SAS-SATA Device rev. 1.0	[spec]	HPE, Meta, Microsoft
	Datacenter NVMe SSD rev. 2.5	[spec]	Dell, Google, HPE, Meta, Microsoft
	NVMe Telemetry Scripts	[scripts]	Samsung
	2nd Annual OCP Storage Tech Talk	[virtual]	May 2023
2024	Meta and Google Open-Source SSD Test Cases	Current	OCP Test Framework (Meta), OCP Storage Test Cases (Meta), ocp-diag-ssd-qual (Google)
	3rd Annual OCP Storage Tech Talk	[virtual]	May 2024
	EDSFF SSD Reference Design	[spec]	ScaleFlux PCIe Gen 5 EDSFF NVMe SSD Reference Design
	Datacenter NVMe SSD rev. 2.6	[spec]	Dell, Google, HPE, Meta, Microsoft
	L.O.C.K. rev. 0.5	[spec]	Google, Kioxia, Microsoft, Samsung, Solidigm
	HDD Dynamics Common Language and Tools Phase 2	[link]	Dell, IEIT Systems, Lenovo, Seagate, WD, ZT Systems
	HDD Dynamics Combined Acoustic & Vibration Surrogate	[white paper]	Dell, IEIT Systems, Lenovo, Seagate, WD, ZT Systems
2025	OCP Marketplace (Products/Storage)	[link]	20+ SSDs and 5 SSD Testing Services
	Datacenter SAS-SATA Device rev. 1.5	Coming	HPE, Dell, Meta, Microsoft
	Datacenter NVMe SSD rev. 2.7	Coming	Dell, Google, HPE, Meta, Microsoft
	L.O.C.K. rev. 0.8.1	[link]	Google, Kioxia, Microsoft, Samsung, Solidigm
	4th Annual OCP Storage Tech Talk	[link]	May 2025

- OCP Storage Project Link
- Meetings are 2nd Monday of the month

Thank You