

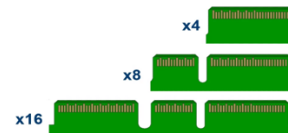
Challenges and Opportunities of EDSFF based Storage Solution

Agenda

- EDSFF Benefits
- Poseidon V1 - E1.S Reference System
- Poseidon V2 - E3.X Reference System
- V1 Validation Result

EDSFF – Benefits

- Designed to overcome conventional device limitations
- Capacity & Performance scaling
 - ✓ Improves thermals, power, and scalability
- Flexibility
 - ✓ Various power (12W, 16W, 20W, 25W, 40W, 70W) and connector (x4, x8, x16) options
 - ✓ Various size: E1.S (5.9/8.01/9.5/15/25mm), E1.L (9.5/18mm), E3.S (1T/2T), E3.L (1T/2T)



E1.L			E3				
Size	38.4 x 318.75 x 9.5mm			76 x 112.75 x 7.5mm	76 x 142.2 x 7.5mm	76 x 112.75 x 16.8mm	76 x 142.2 x 16.8mm
Recommended Power(W)	25W			25W	40W	40W	70W

E1.S					
Size	31.5 x 111.49 x 5.9mm	31.5 x 111.49 x 8.01mm	33.75 x 118.75 x 9.5mm	33.75 x 118.75 x 15mm	33.75 x 118.75 x 25mm
Recommended Power(W)	12W	16W	20W	20W	25W

EDSFF Reference System

- E1 and E3 based system can increase the performance and density
- Have more flexibilities than a traditional system



U.2 SSD x 10ea

VS.



E1.S SSD x 32ea

1RU



U.2 SSD x 24ea

VS.



E3.S 1T SSD x 40ea

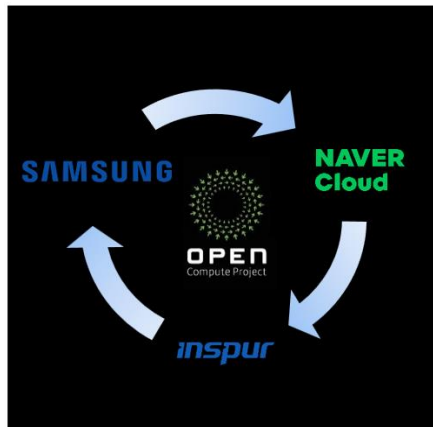
2RU



E3.S 2T SSD

Poseidon Project

- OCP based industrial collaboration b/w “Component Vendor ↔ System Vendor ↔ Data Center”
- Open-source HW & SW project to expand NVMe eco-system



Poseidon SW
Storage OS for NVMe-oF



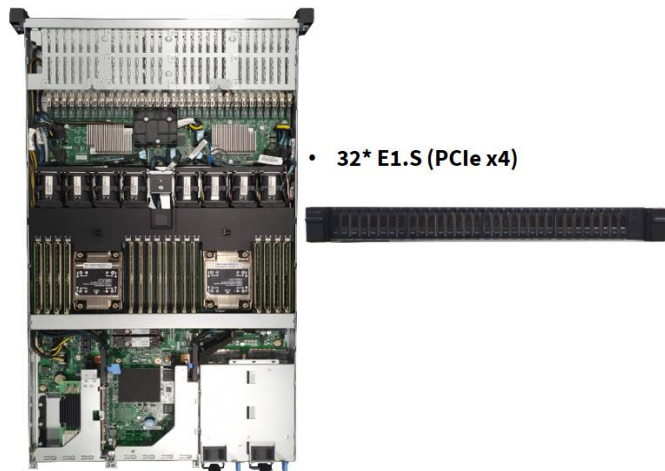
Poseidon HW
E1.S/E3.S Ref. System



EDSFF Reference System

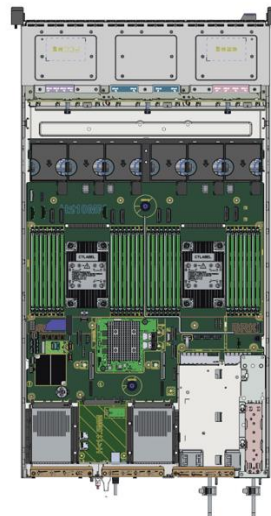
- PSD V1 announced in 2020 OCP virtual summit
- V2 Reference Server announced in 2021 OCP global summit

PSD V1 server



- 32* E1.S (PCIe x4)

V2 reference server



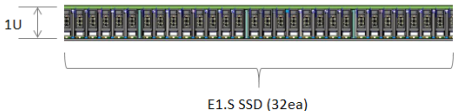
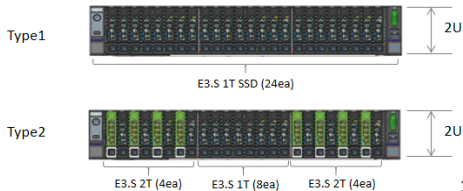
- Type1: 24* E3.S (PCIe x4)

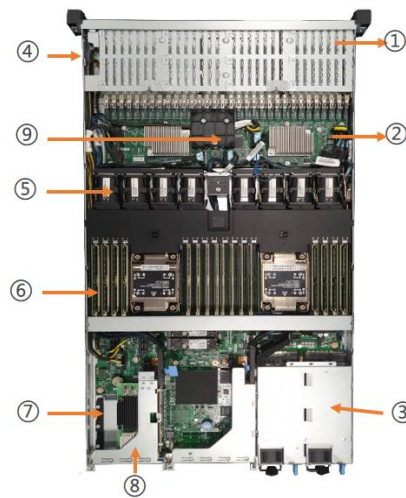


- Type2: 8*E3.S 2T (PCIe x8) + 8* E3.S (PCIe x4)



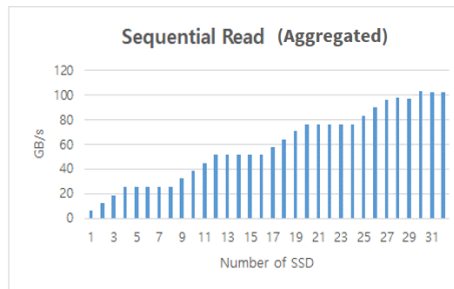
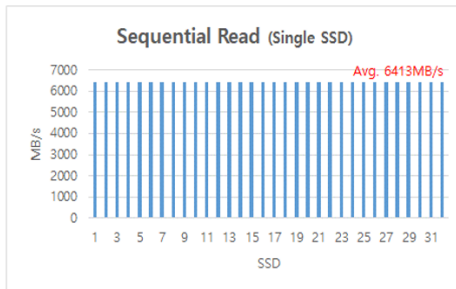
EDSFF Reference System

	E1.S Ref. System (iBoF V1)	E3.S Ref. System (iBoF V2)
System Design	 1U E1.S SSD (32ea) 1.1kW	 Type1 E3.S 1T SSD (24ea) 2U Type2 E3.S 2T (4ea) E3.S 1T (8ea) E3.S 2T (4ea) 2U 1.5kW
Reference Target	ICX / PCIe Gen4 / DDR4 3200 / E1.S * 32ea	SPR / PCIe Gen5 / DDR5 4800 / E3.S 1T * 24ea or 2T * 8ea
Application	NVMe-oF (POS), DB	NVMe-oF (POS), DB, VM, Video Transcoding
Relative Characteristic	High Performance / Density (Max. 8TB * 32ea)	High Performance / Density / Flexible
System Structure	1RU / E1.S * 32ea / 2S (Ice Lake)	2RU / E3.S * 16ea / 2S (Next gen CPU)
Note	차세대 SSD F/F인 E1.S의 Reference 시스템으로 PCIe Gen4 선행 검증 환경 제공 및 당사 E1.S 제품 생태계 확장 목표	- E3.S Reference 시스템으로 PCIe Gen5 기반 제품 생태계 확장 및 선행 검증 환경 제공 - SSD 뿐만 아니라 다양한 형태의 Device를 제공함으로써 고객의 Application 맞는 유연한 구성 가능



IO Performance – Single vs Aggregated

- Each 32 SSDs shows stable IO performance around 6.4GB/s
- Aggregated IO performance achieved > 102GB/s
- Lanes b/w CPU↔PCIe Switch (x16) are saturated with 4 SSDs



*Theoretical B/W limit: 128GB/s

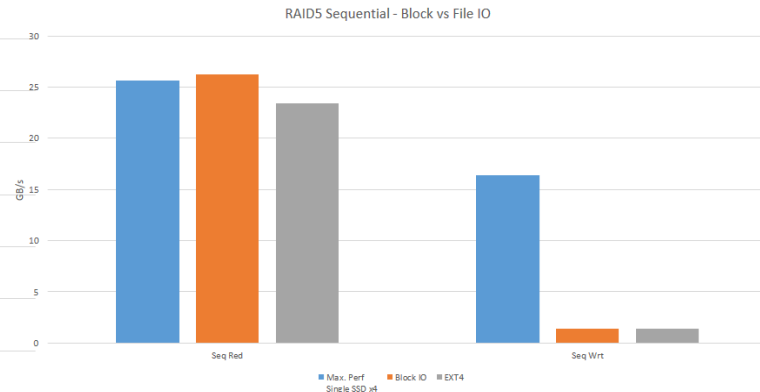
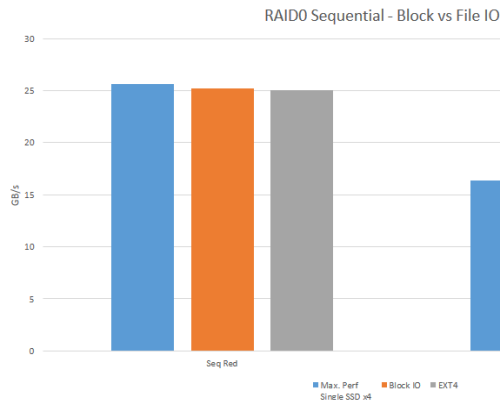
• Samsung PM9A3 Specification

Form factor	E1.S
Capacity	960 GB, 1.92 TB, 3.82 TB, 7.68 TB
Sequential read	Up to 6,500 MB/s
Sequential write	Up to 3,200 MB/s
Random read	Up to 900,000 IOPS
Random write	Up to 150,000 IOPS
Physical Dimensions	31.5 x 111.49 x 5.9 mm
Power consumption	Read: <= 9.7W, Write: <= 11.7W
Host interface	PCIe Gen 4 x4



IO Performance – Raid 0&5

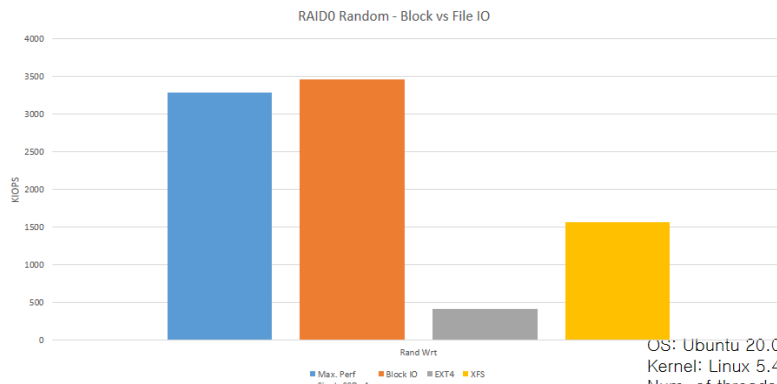
- Compared Raid 0 & 5 performance of 4 x NVMe SSDs
- Observed significant performance degradation on Raid 5 write operation
- Further study would be needed for S/W Raid on NVMe storage system



OS: Ubuntu 20.04.2 LTS
Kernel: Linux 5.4.0-65-generic
Num. of threads : 112 (56 * 2)
Model name : Intel Xeon Gold 6330N CPU @ 2.20GHz
fio: fio-3.16

IO Performance – Raid 0&5

- Compared Raid 0 random performance of 4 x NVMe SSDs
- Observed random write performance increase in XFS compared to EXT4
- Further study would be needed for S/W Raid on NVMe storage system



OS: Ubuntu 20.04.2 LTS
Kernel: Linux 5.4.0-65-generic
Num. of threads : 112 (56 * 2)
Model name : Intel Xeon Gold 6330N CPU @ 2.20GHz
fio: fio-3.16

Thank You!

