HOW END TO END NVMe WILL IMPACT APPLICATIONS

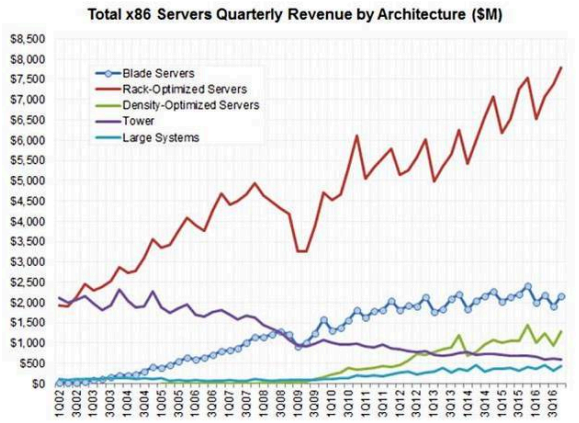
Ivan Iannaccone

Director of Product Management

@ivaniannaccone

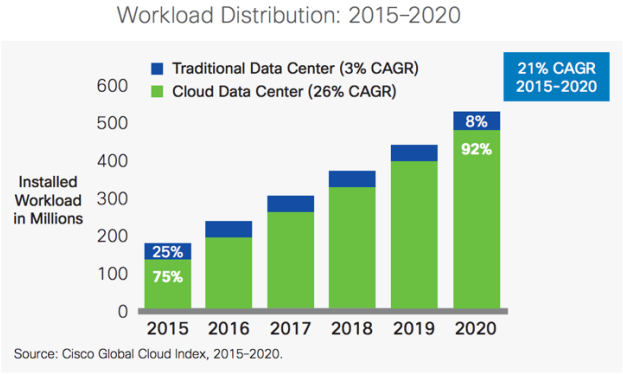
CLOUD WORKLOADS

SHAPING THE DATA CENTERS ARCHITECTURE

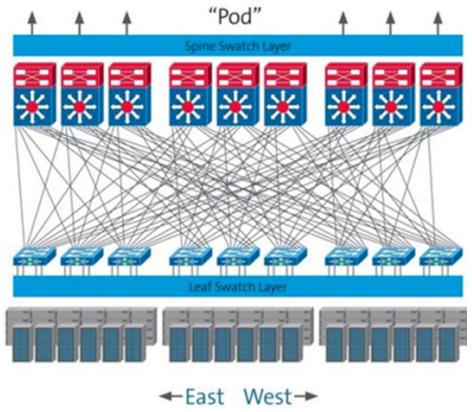


Source:TheRegister*

Rack-Optimized servers on the raise



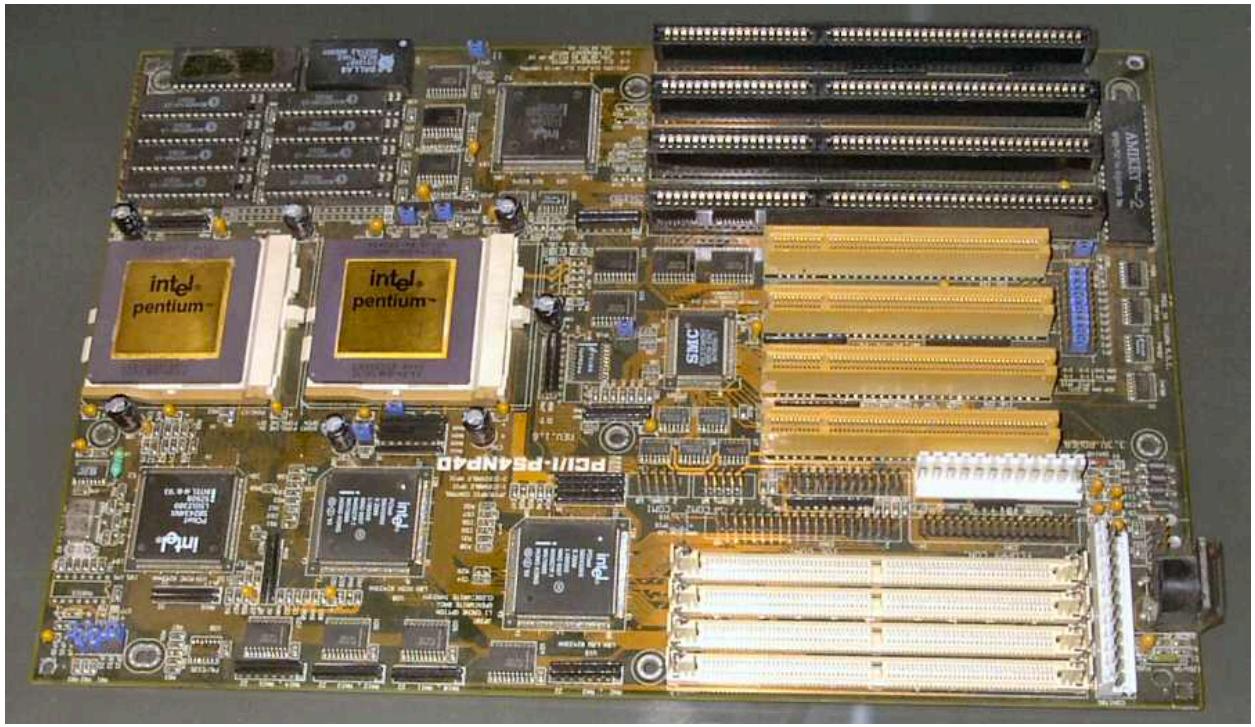
Cloud Workloads prevail



Applications shaping new network architectures

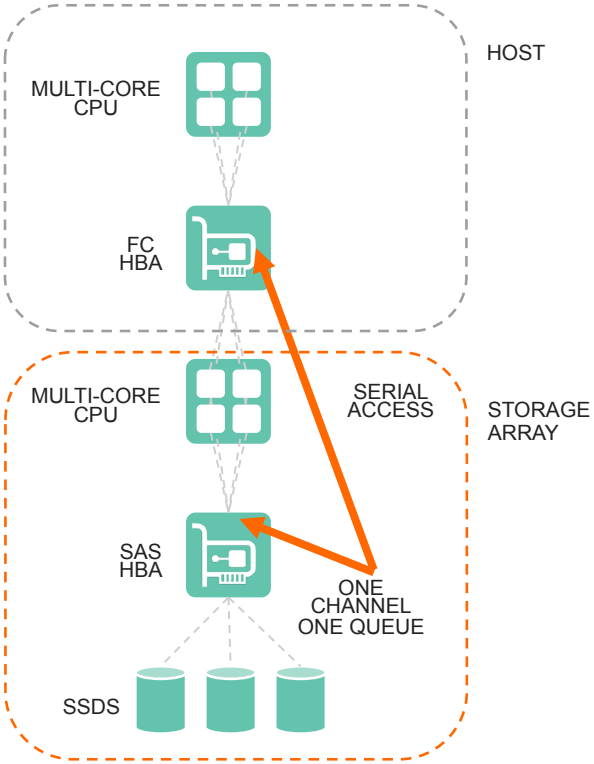
PURESTORAGE®

REMEMBER THIS?



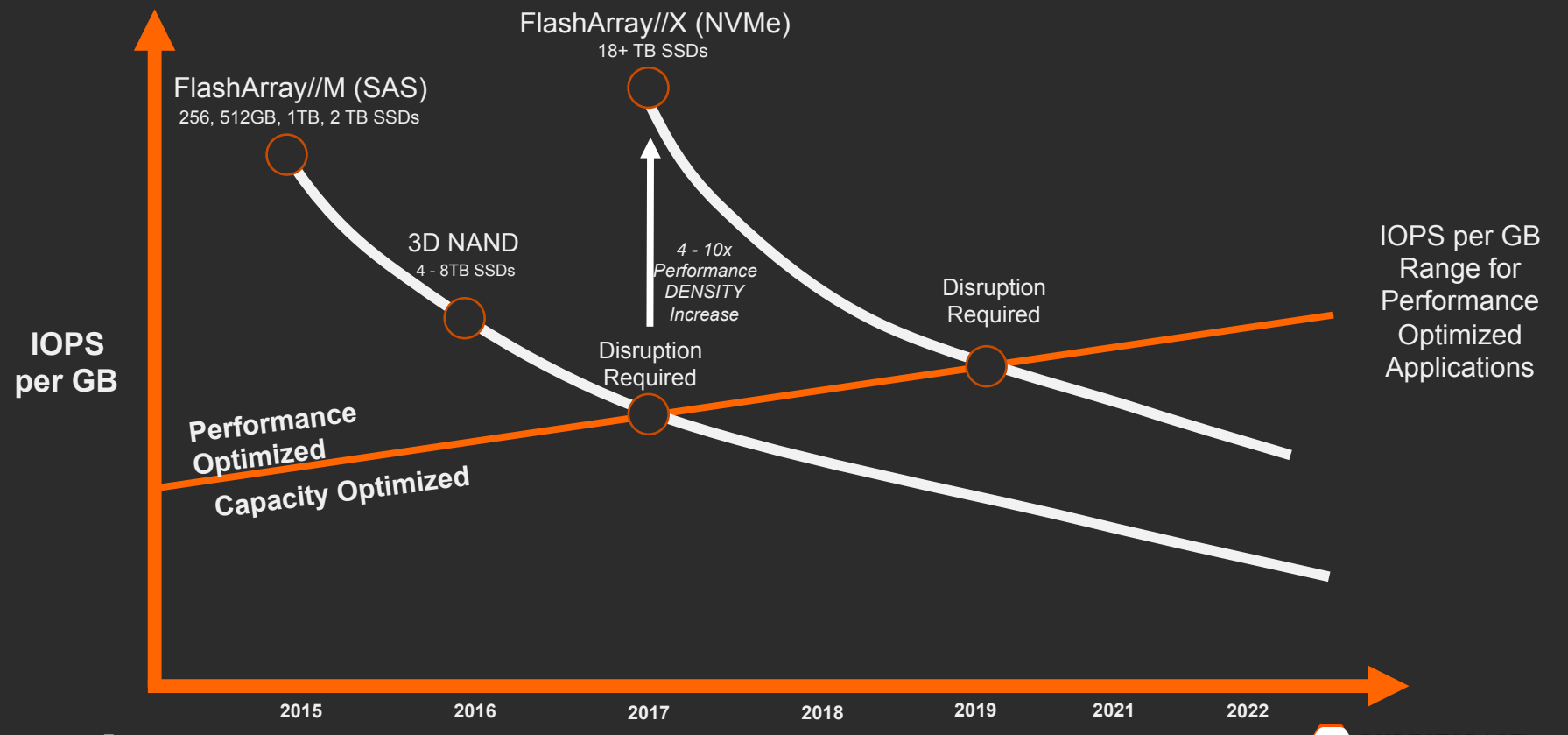
SCSI BOTTLENECK

SCSI
- HBA bottleneck
- Adequate for HDDs



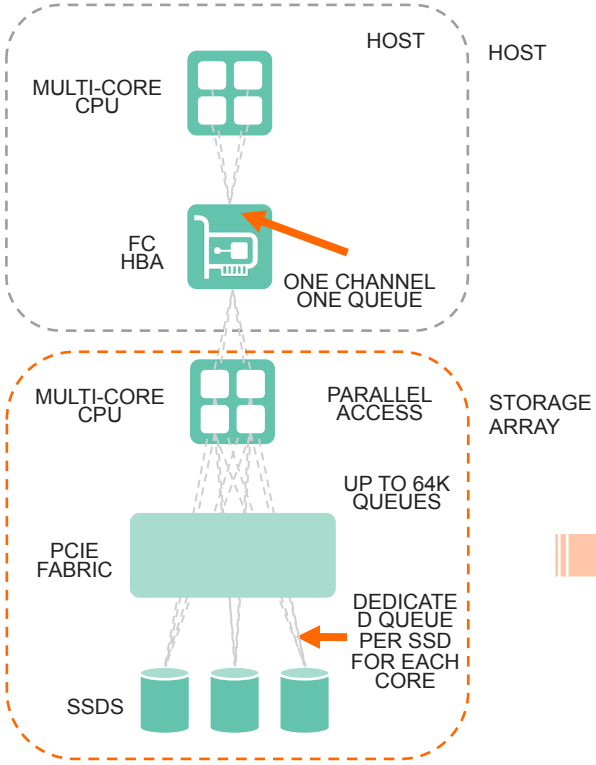
Flash
- Inherently parallel
- SCSI limits Flash performance

NEW STORAGE ARCHITECTURES

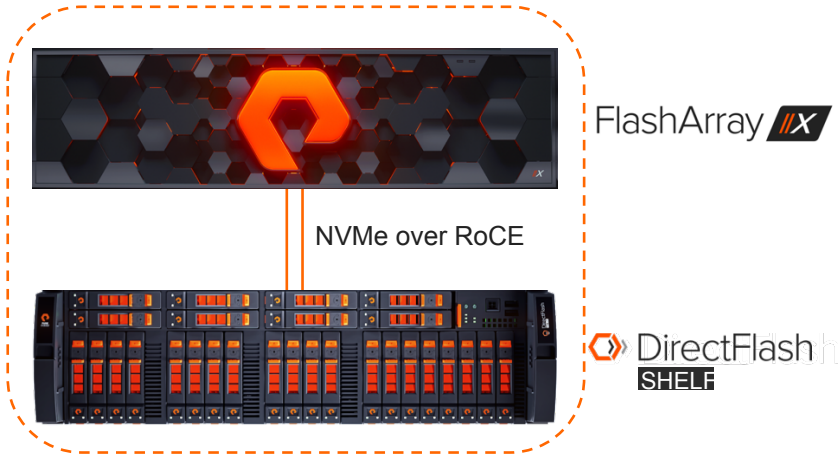


PURESTORAGE®

NVMe ARRAY

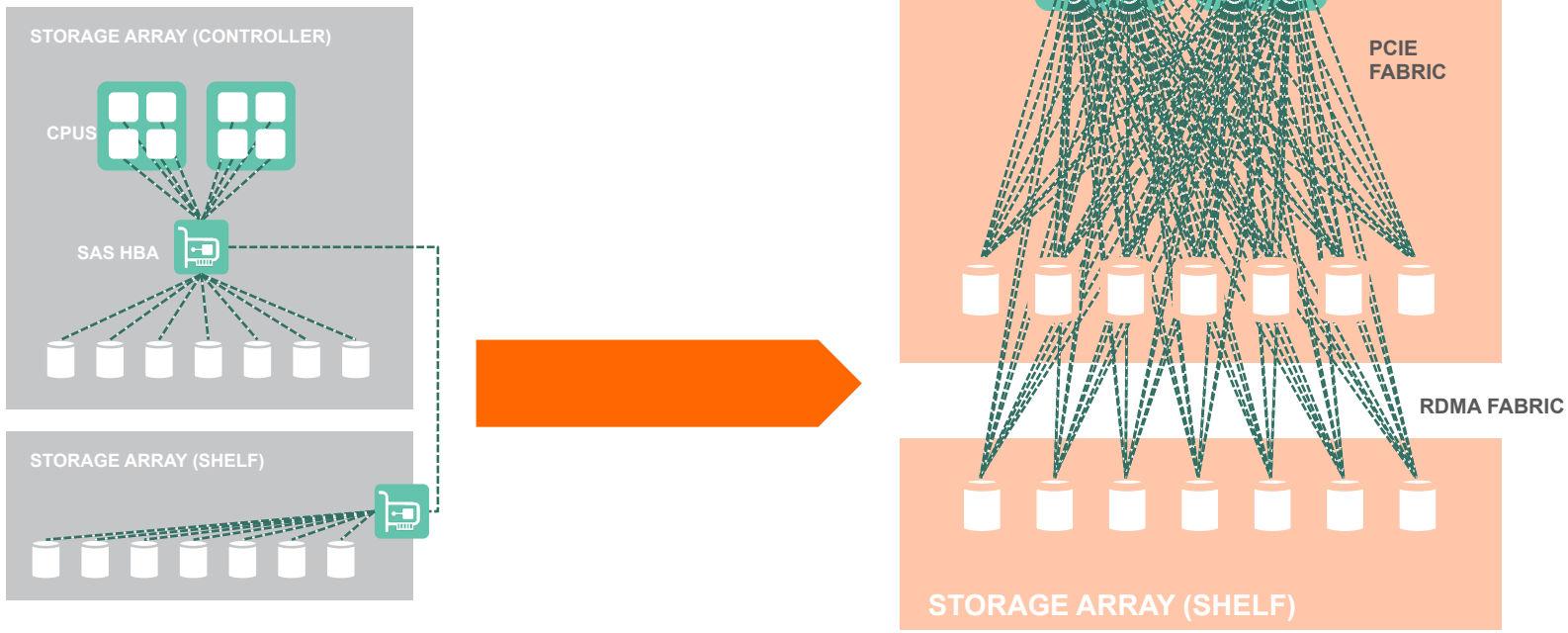


- SCSI Host
- NVMe Array



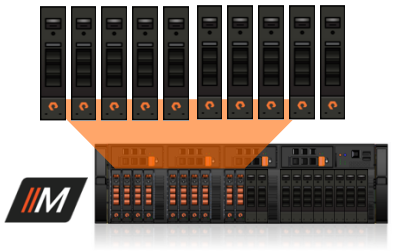
PURESTORAGE®

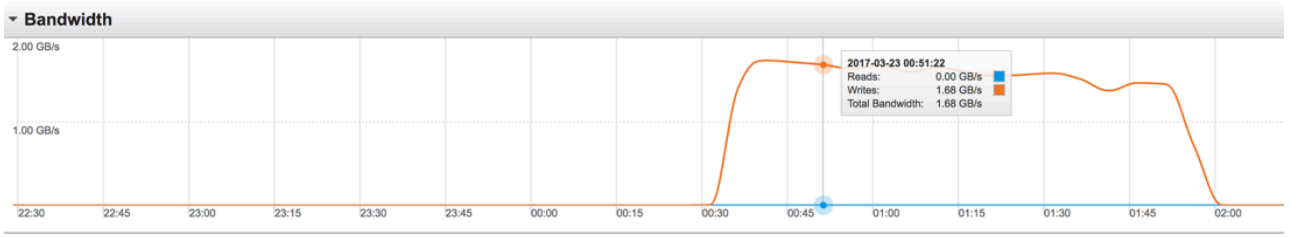
NVMe SHELF EXPANSION



PROVIDING BETTER HEALTHCARE

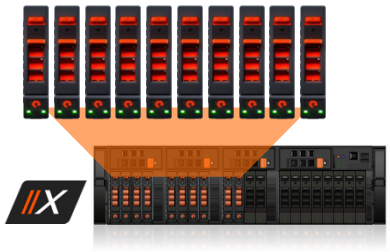
10 SAS FLASH MODULES (20XSSDs)



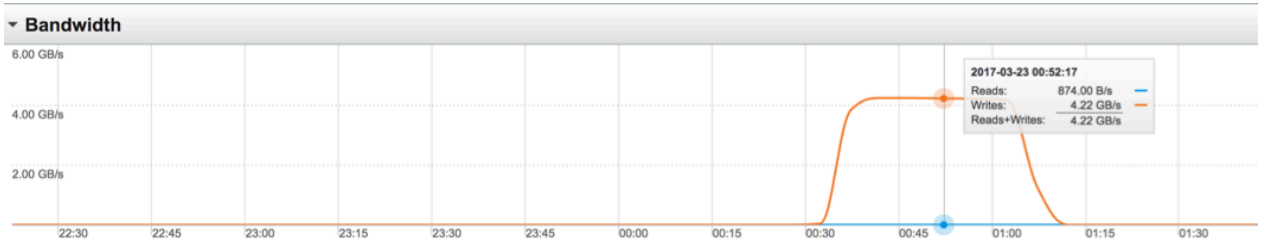


Test Started 0:33:12 and finished at 2:03:12 = 90 minutes - Running at **1.5GB/s Write Average**

10x DirectFlash MODULE



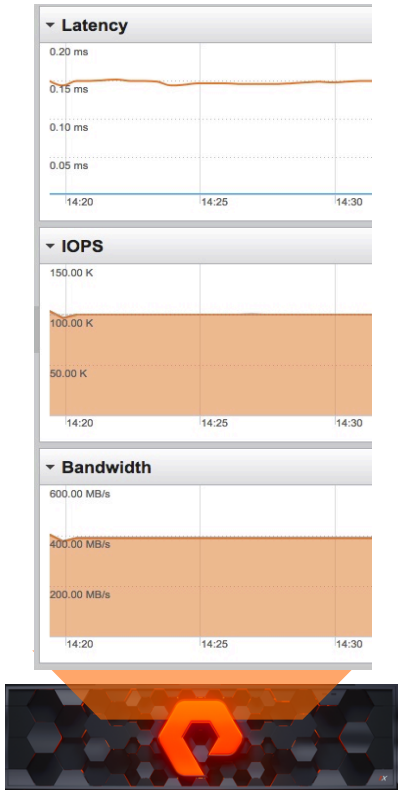
> 2.5x FASTER WRITES IN EPIC ENVIRONMENTS



Test Started 0:33:12 and finished at 1:12:16 = 39 minutes - Running at **4.2GB/s Write Average**

PURESTORAGE®

THE NEXT BOTTLENECK: LATENCY IN THE HOST



Source:blog.koehntopp.info*

Customer running 100k random writes of 4K per second (benchmark) over iSCSI, zero jitter at **350µs** end-to-end write latency across six network switches.

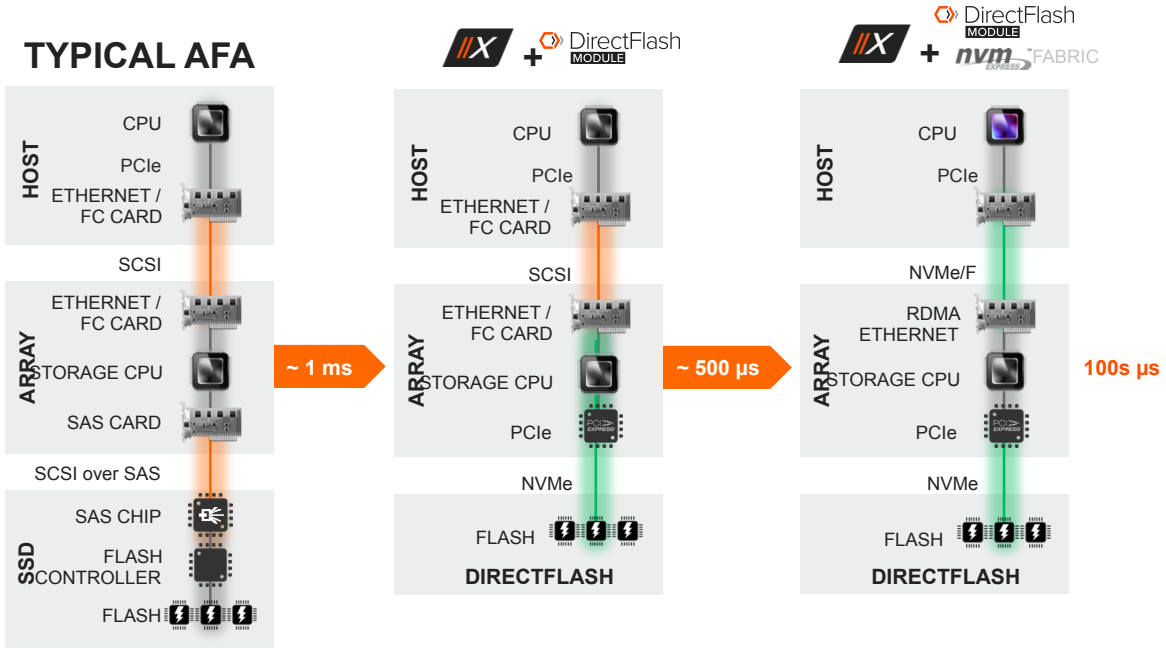
Jul 26, 2017	interval	i/o rate	MB/sec 1024**2	bytes i/o	read pct	resp time	read resp	write resp
14:19:40.047	4051	99736.00	389.59	4096	0.00	0.353	0.000	0.353
14:19:41.044	4052	99864.00	390.09	4096	0.00	0.357	0.000	0.357
14:19:42.044	4053	99786.00	389.79	4096	0.00	0.355	0.000	0.355
14:19:43.044	4054	99799.00	389.84	4096	0.00	0.353	0.000	0.353
14:19:44.046	4055	100000.00	390.63	4096	0.00	0.355	0.000	0.355
14:19:45.044	4056	100081.00	390.94	4096	0.00	0.359	0.000	0.359
14:19:46.044	4057	100000.00	390.63	4096	0.00	0.354	0.000	0.354
14:19:47.045	4058	100557.00	392.80	4096	0.00	0.356	0.000	0.356
14:19:48.046	4059	100112.00	391.06	4096	0.00	0.356	0.000	0.356
14:19:49.047	4060	99587.00	389.01	4096	0.00	0.352	0.000	0.352

FlashArray//X response time is **150µs** latency; with the host experiencing around 350µs. Switches are accounting for around 50µs and 150µs or more from the Linux storage stack.

How do you further reduce latency by improving host storage stack?

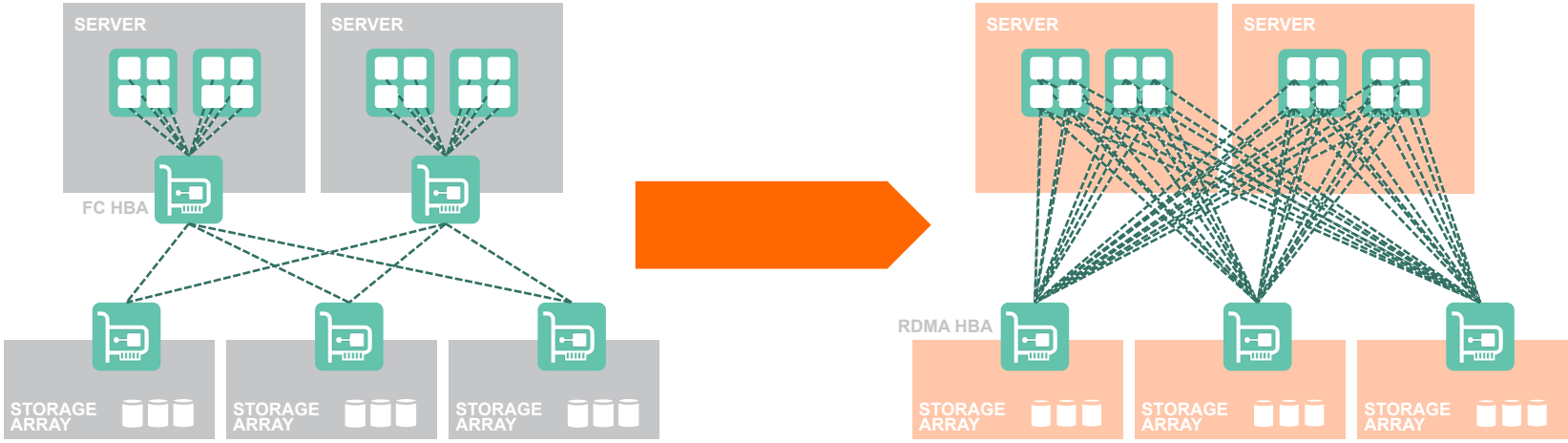
PURESTORAGE®

NVMe/F TO THE RESCUE



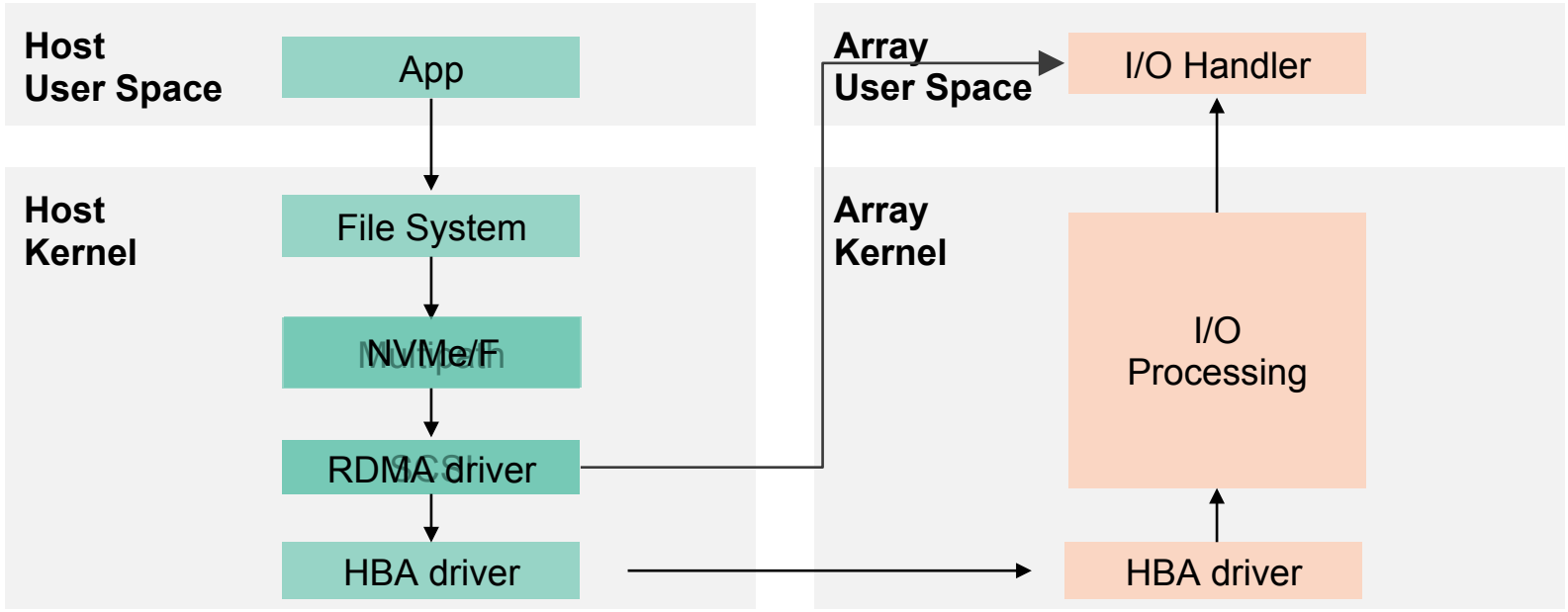
DEPLOYING NVMe/F

→ NVM$_{E/F}$ OVER RDMA

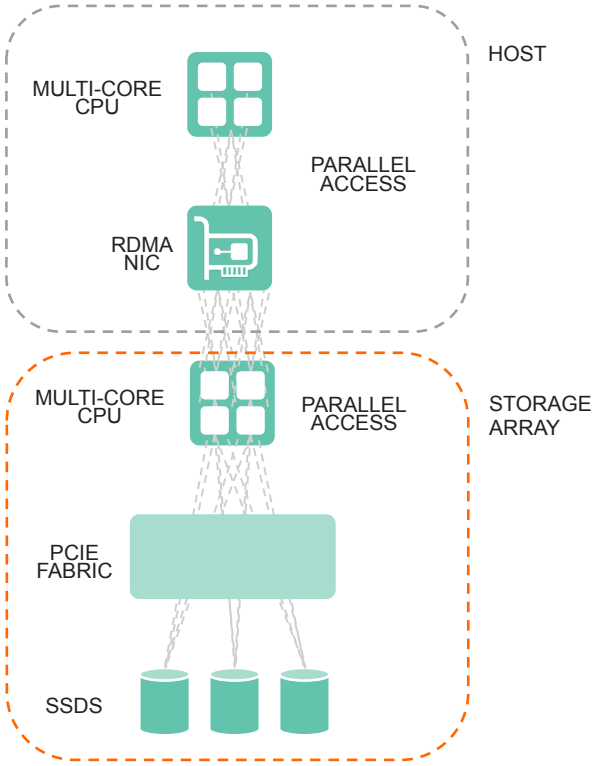


PURESTORAGE®

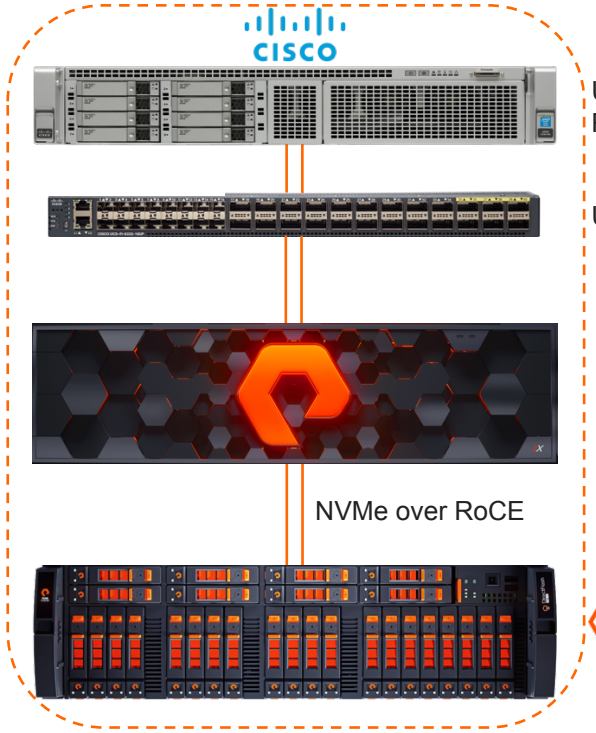
NVMe/F – I/O PATH



NVMe FABRIC



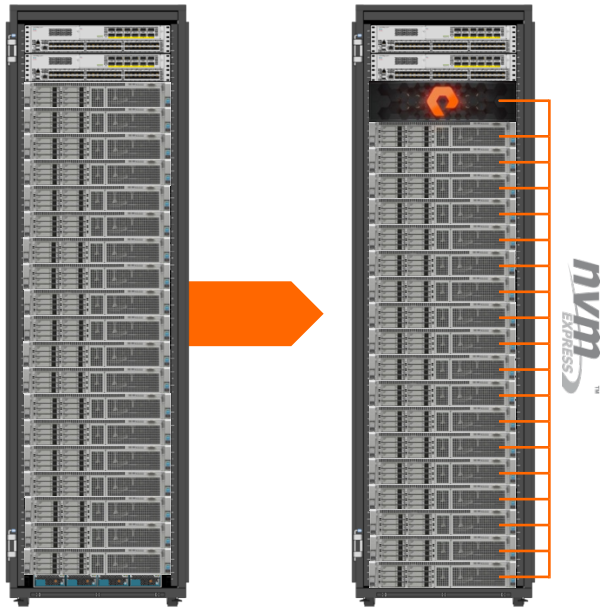
TECHNOLOGY PREVIEW



ALL NVMe
NVMe over Fabrics

PURESTORAGE®

NEW ARCHITECTURES FOR AFA



BRING STORAGE AND COMPUTE CLOSER

BRING ALL THE EFFICIENCY OF FLASHARRAY TO DAS APPLICATIONS

OFFLOAD STORAGE PROCESSING FROM SERVER CPUs

SUMMARY

1. NVMe access to media will enable customers to not have to overprovision flash capacity in order to get full performance for their applications.

2. Parallelism is key to future workloads :

    - High CPU core count
    - Parallel processing
    - Highly concurrent access

3. Parallelism should be at all layers of infrastructure
    - NVMe/F enables such parallelism

Tomorrow's Cloud Block

1,300 CORES

50 TB DRAM

2.6 PB FLASH

ELASTIC FILE, OBJECT, BLOCK

100 Gb/s Ethernet + NVMe/F

µs latency end-to-end

PURESTORAGE®