



## The Real Story on Flash Storage Performance

Session **TEST-101B-1**

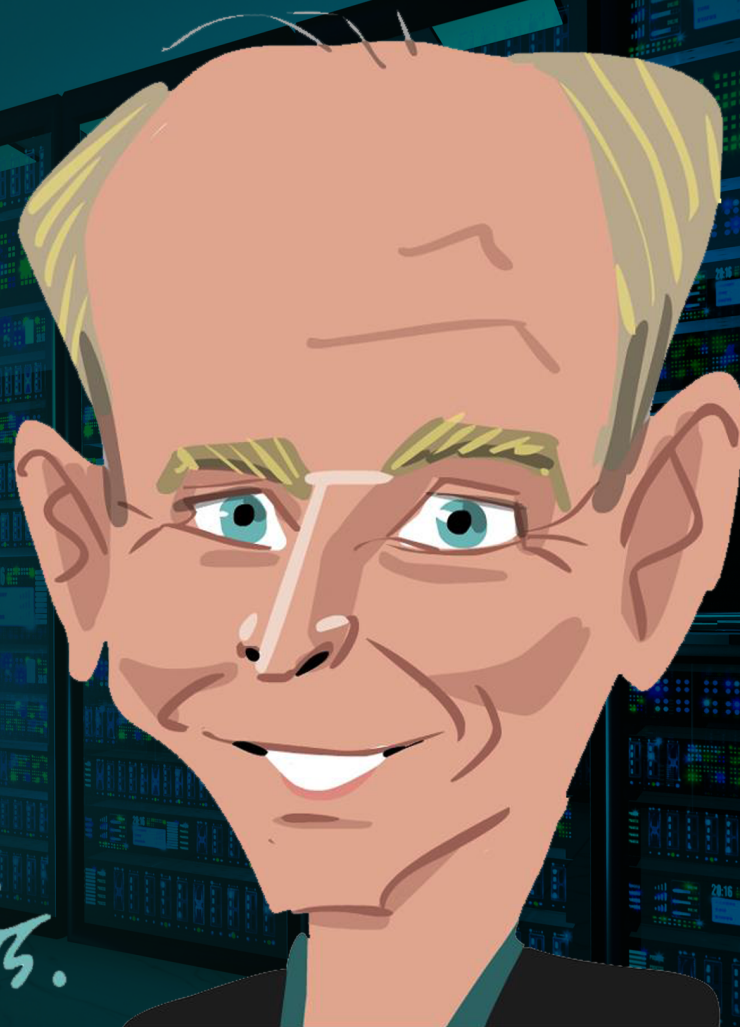
9:45 a.m. - 10:50 a.m. PDT,

**Tuesday, August 7, 2018**

Ballroom G




Flash Memory Summit






Flash Memory Summit

# This Presentation

 **Demartek**<sup>®</sup>

The Real Story on  
Flash Storage Performance

Session **TEST-101B-1**  
9:45 a.m. - 10:50 a.m. PDT,  
**Tuesday, August 7, 2018**  
Ballroom G

  
Flash Memory Summit

*J.S.*

<https://www.demartek.com/FMS2018/>

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek



Flash Memory Summit

## Agenda

- ◆ About Demartek
- ◆ Synthetic vs. Real-world workloads
- ◆ Performance Results – Various Flash Solutions  
*(new since last year's Flash Memory Summit presentation)*
- ◆ Industry Trends & Future Directions

Some of the images in this presentation are clickable links to web pages or videos → 



Flash Memory Summit

## About Demartek



Click to view this one minute video

[https://www.demartek.com/Demartek\\_Video\\_Library.html](https://www.demartek.com/Demartek_Video_Library.html)



Flash Memory Summit

## About Demartek



- ◆ Industry Analysis and ISO 17025 accredited test lab
- ◆ Lab includes enterprise servers, networking & storage: DAS, NAS, SAN, 10/25/40/100 GbE, 16/32 GFC, NVMe, NVMe over Fabrics
- ◆ We prefer to run real-world applications to test servers, storage and HCI solutions (databases, VMware, IoT, etc.)
- ◆ Demartek is an EPA-recognized test lab for **ENERGY STAR Data Center Storage** testing
- ◆ Website: <https://www.demartek.com/TestLab/>



**SNIA Emerald™**  
RECOGNIZED TESTER



Flash Memory Summit

## Demartek – Independent Test Lab

- ◆ We are not a product manufacturer
- ◆ We work with most product manufacturers
- ◆ We use almost every interface, device type, etc.
- ◆ We run system-level tests with real operating systems and applications – just like end-users
- ◆ We test current and new technologies



Flash Memory Summit

# Synthetic vs. Real-world Workloads

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek



Flash Memory Summit

## Synthetic Workloads

- ◆ Synthetic workload generators allow precise control of I/O requests with respect to:
  - ◆ Read/write mix, block size, random vs. sequential & queue depth
- ◆ These tools are used to generate the *“hero numbers”*
  - ◆ 4KB 100% random read, 4KB 100% random write, etc.
  - ◆ 256KB 100% sequential read, 256KB 100% sequential write, etc.
- ◆ Manufacturers advertise the hero numbers to show the top-end performance in the corner cases
  - ◆ Demartek also sometimes runs these tests





Flash Memory Summit

## Real-world Workloads

- ◆ Use variable levels of compute, memory and I/O resources as the work progresses
  - ◆ May use different and multiple I/O characteristics simultaneously for I/O requests (block sizes, queue depths, read/write mix and random/sequential mix)
- ◆ Many applications capture their own metrics such as database transactions per second, etc.
- ◆ Operating systems can track physical and logical I/O metrics
- ◆ *End-user customers have these applications*



Flash Memory Summit

# Performance Results

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek

10

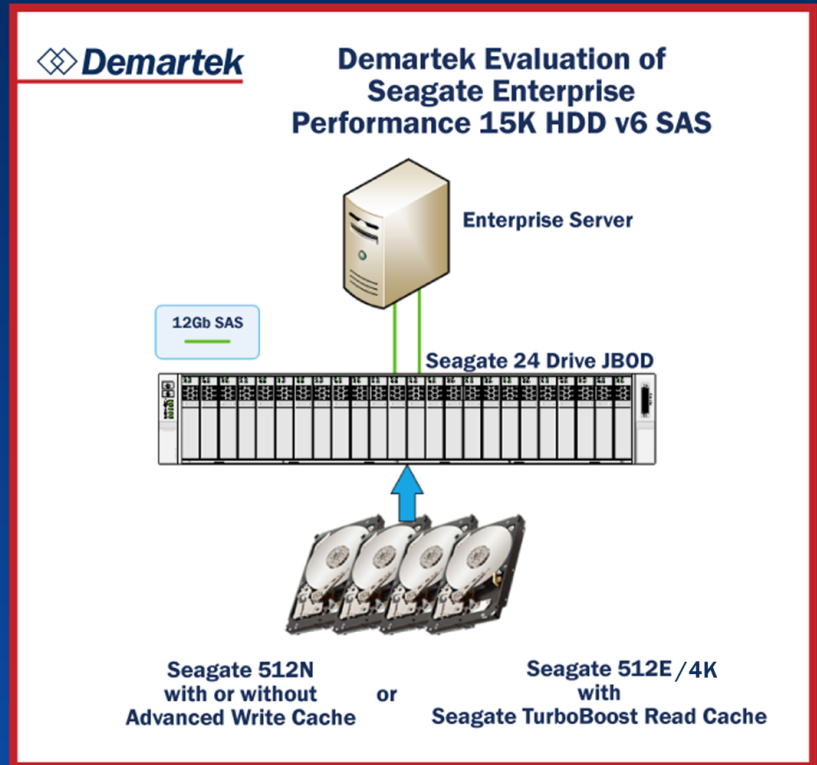


Flash Memory Summit

## Adding NAND Flash to HDDs

- ◆ 24x Seagate TurboBoost HDDs with flash cache in each drive
- ◆ Multiple synthetic & real-world workloads

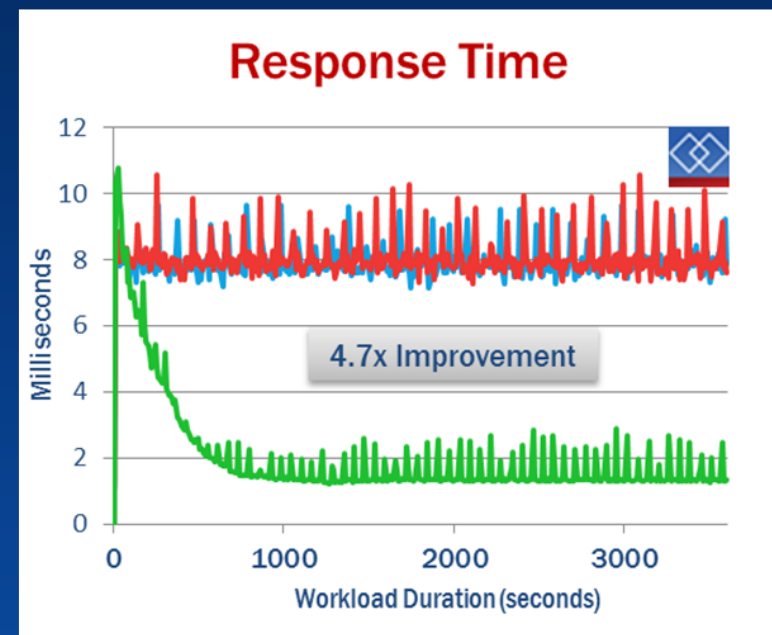
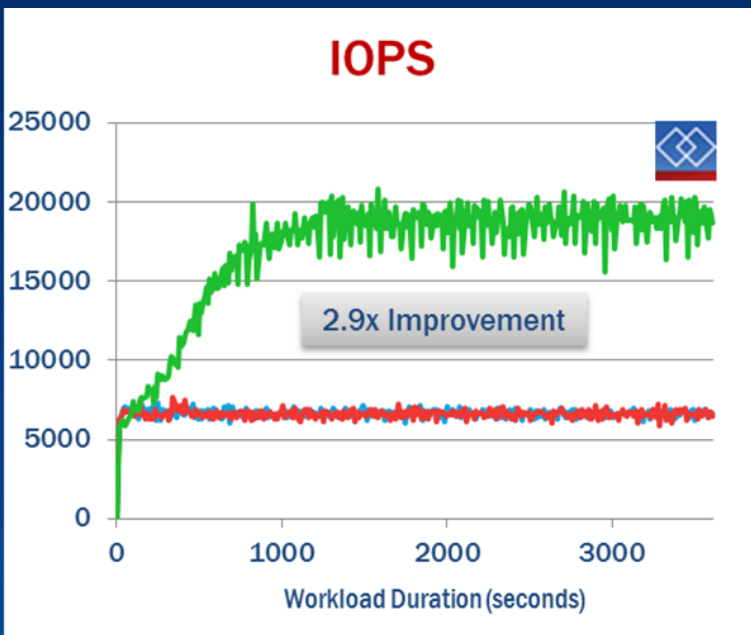
<https://www.demartek.com/SeagateEnhancedCache/>





Flash Memory Summit

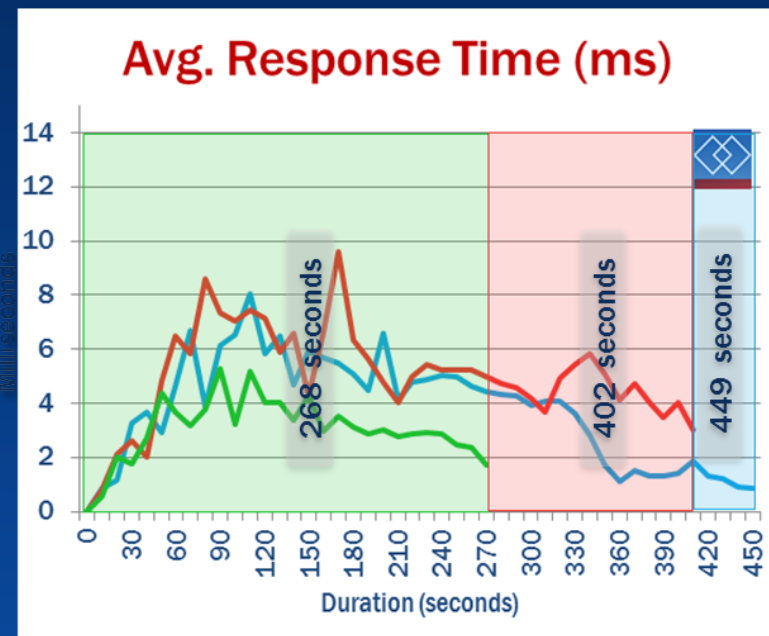
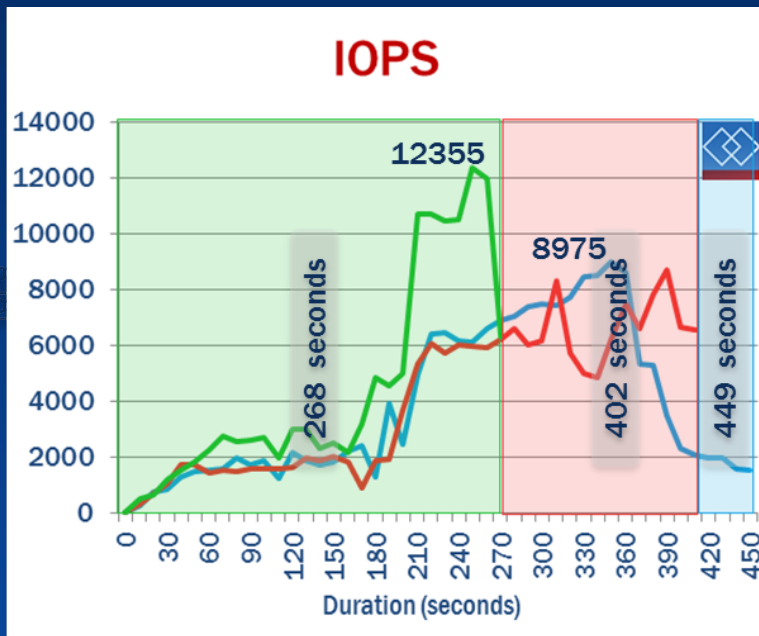
# Adding NAND Flash to HDDs



## Microsoft SQL Server OLTP workload



# Adding NAND Flash to HDDs



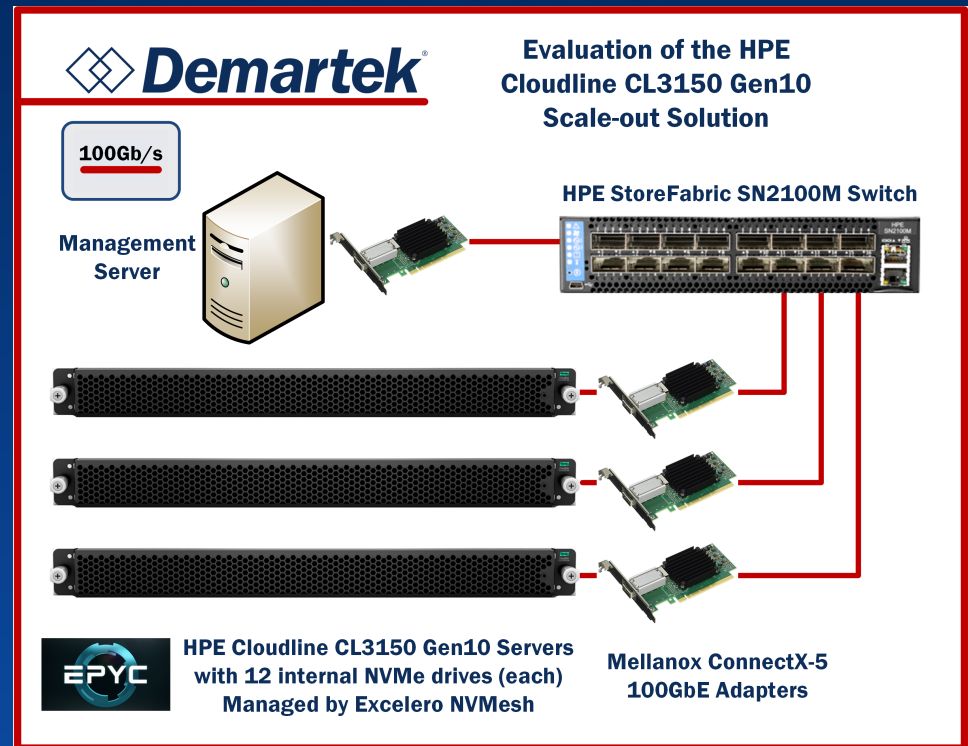
VMware ESXi Bootstorm: Fixed amount of work, 60 Win10 VMs



Flash Memory Summit

# 12 NVMe Drives in Cloud Server

- ◆ HPE AMD EPYC cloud server cluster
- ◆ 100 GbE network
- ◆ Excelero NVMeMesh
- ◆ Yahoo Cloud Serving Benchmark (YCSB)



<https://www.demartek.com/HPE-Cloudline-CL3150-Benchmark/>

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek



# Yahoo Cloud Serving Benchmark (YCSB)

- ◆ Common cloud datacenter workloads
  - ◆ **Workload A:** Update heavy (50% read, 50% write)
  - ◆ **Workload B:** Read mostly (95% read, 5% write)
  - ◆ **Workload C:** Read only (100% read)
  - ◆ **Workload D:** Read latest (new records inserted and then read)
  - ◆ **Workload E:** Short ranges (ranges of reads, such as email threads)
  - ◆ **Workload F:** Read-modify-write
- ◆ Uses NoSQL database (MongoDB, Cassandra, etc.)



# Yahoo Cloud Serving Benchmark (YCSB)

- ◆ Common cloud datacenter workloads



- ◆ **Workload A:** Update heavy (50% read, 50% write)



- ◆ **Workload B:** Read mostly (95% read, 5% write)

- ◆ **Workload C:** Read only (100% read)

- ◆ **Workload D:** Read latest (new records inserted and then read)

- ◆ **Workload E:** Short ranges (ranges of reads, such as email threads)



- ◆ **Workload F:** Read-modify-write

- ◆ Uses NoSQL database (MongoDB, Cassandra, etc.)





Flash Memory Summit

## Cloud compute / storage nodes

- ◆ Each server was configured identically
  - ◆ One node was designated the compute node
  - ◆ Two nodes were designated the storage nodes (where the application database resided)
- ◆ All the data had to traverse the network
- ◆ In the event of a compute node failure, it can be replaced without moving any data



Flash Memory Summit

## YCSB Database Record Counts

- ◆ 700,000 records (700K)
- ◆ 200,000,000 records (200M)
- ◆ 500,000,000 records (500M)
  
- ◆ Fixed amount of work to be processed



Flash Memory Summit

## Bottleneck

- ◆ With 12 NVMe drives in each server, we found that the bottleneck was the 100GbE network
- ◆ See my NVMe over Fabrics Rules of Thumb later in this presentation



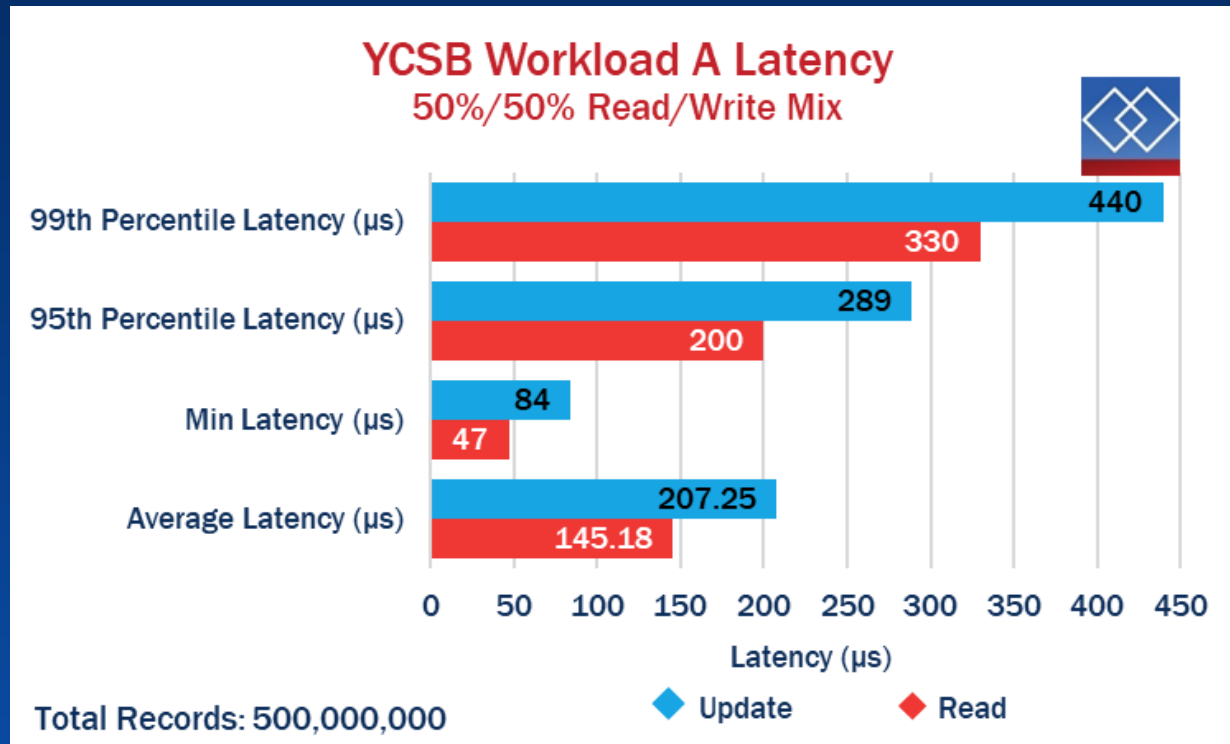
## Run time

- ◆ Workload F (the longest of the three we chose)

Workload F	Milliseconds	Seconds	Minutes
700K records	43120	43	0.7
200M records	2777729	2778	46.3
500M records	5230121	5230	87.2

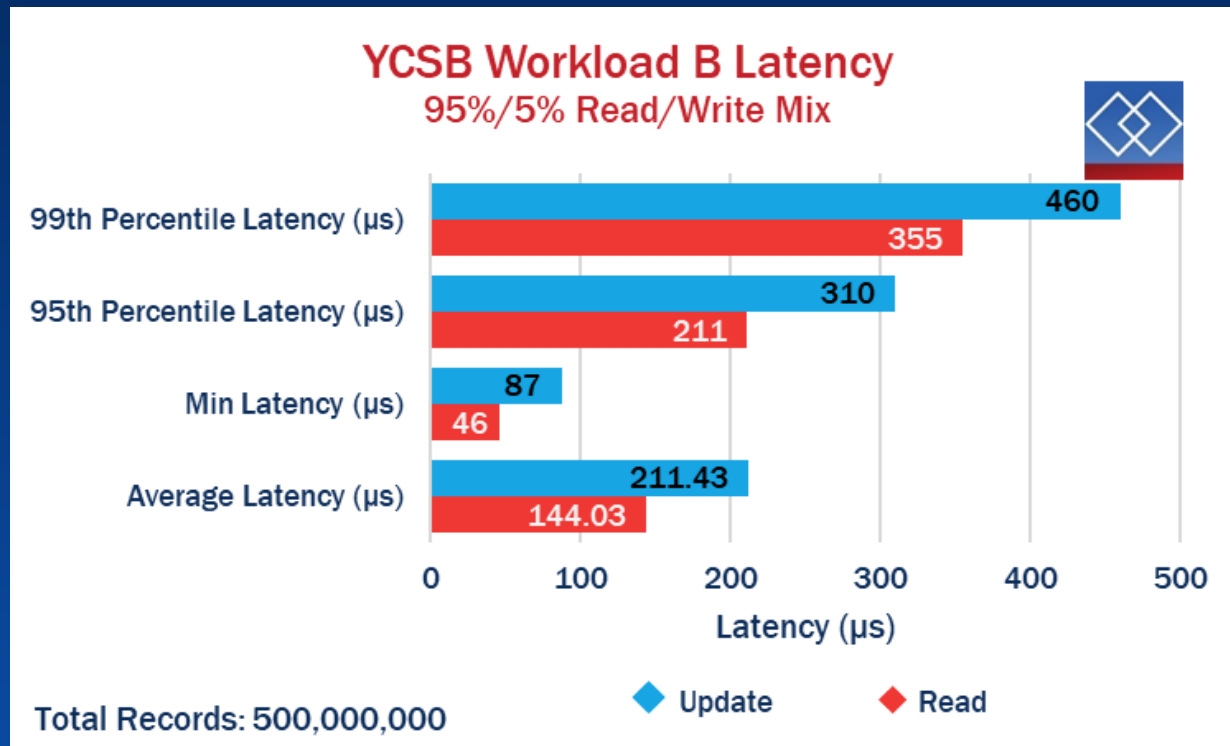


# Results: Workload A



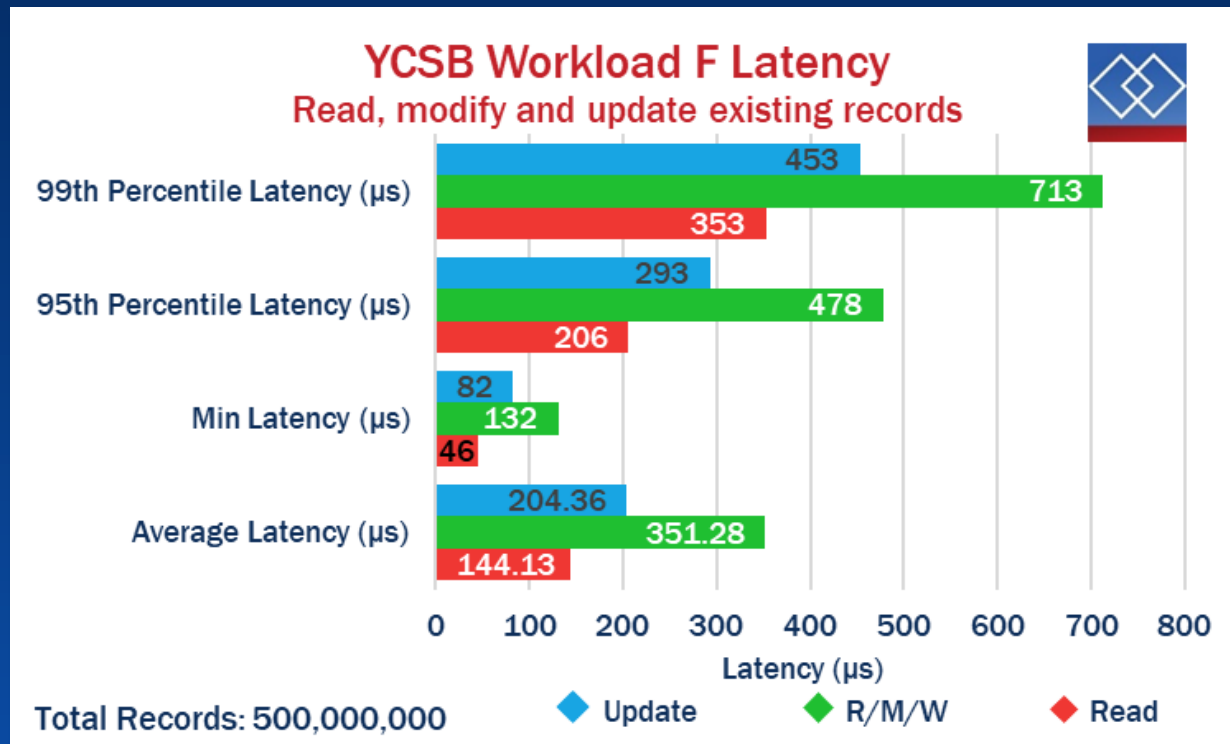


# Results: Workload B





# Results: Workload F

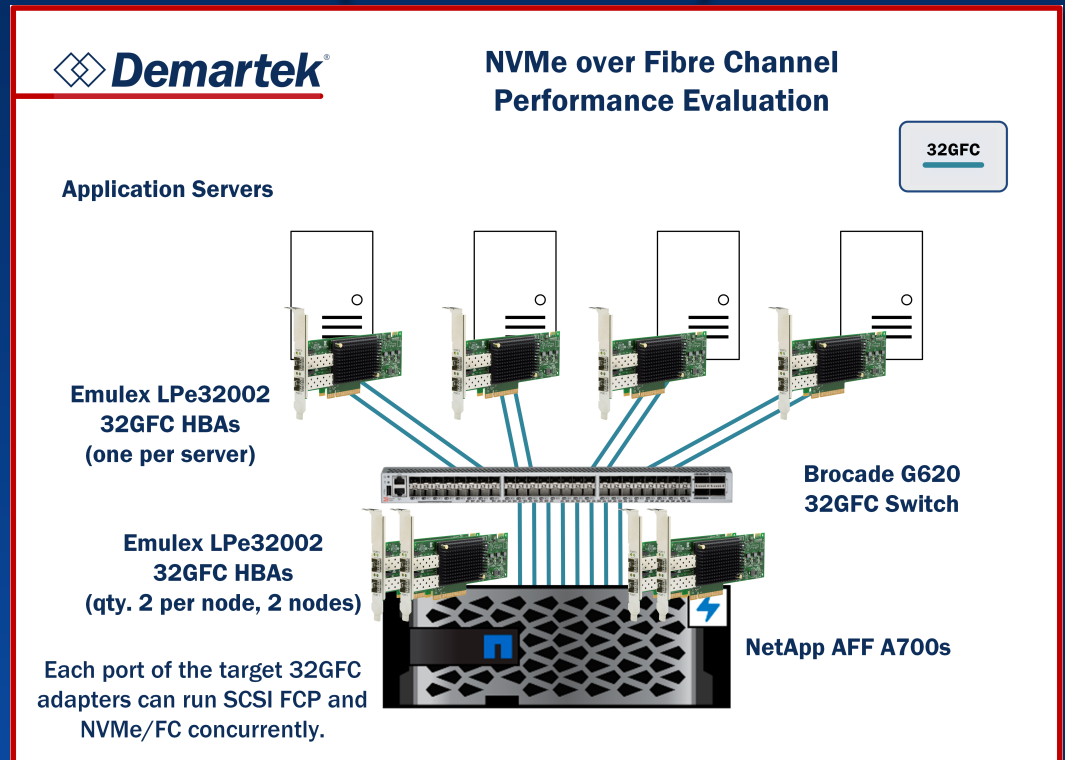




Flash Memory Summit

# NVMe over Fabrics (FC-NVMe)

- ◆ Comparison of FC-SCSI to FC-NVMe
- ◆ Same hardware, different protocol

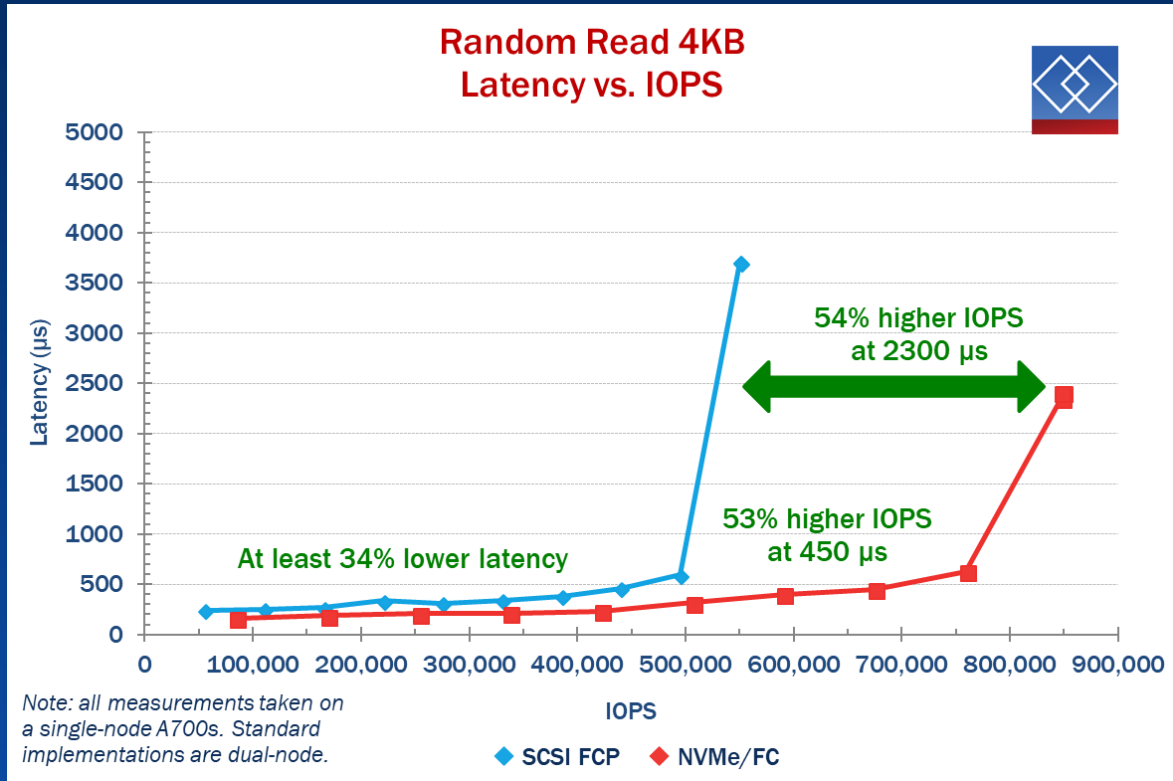


<https://www.demartek.com/ModernSAN/>



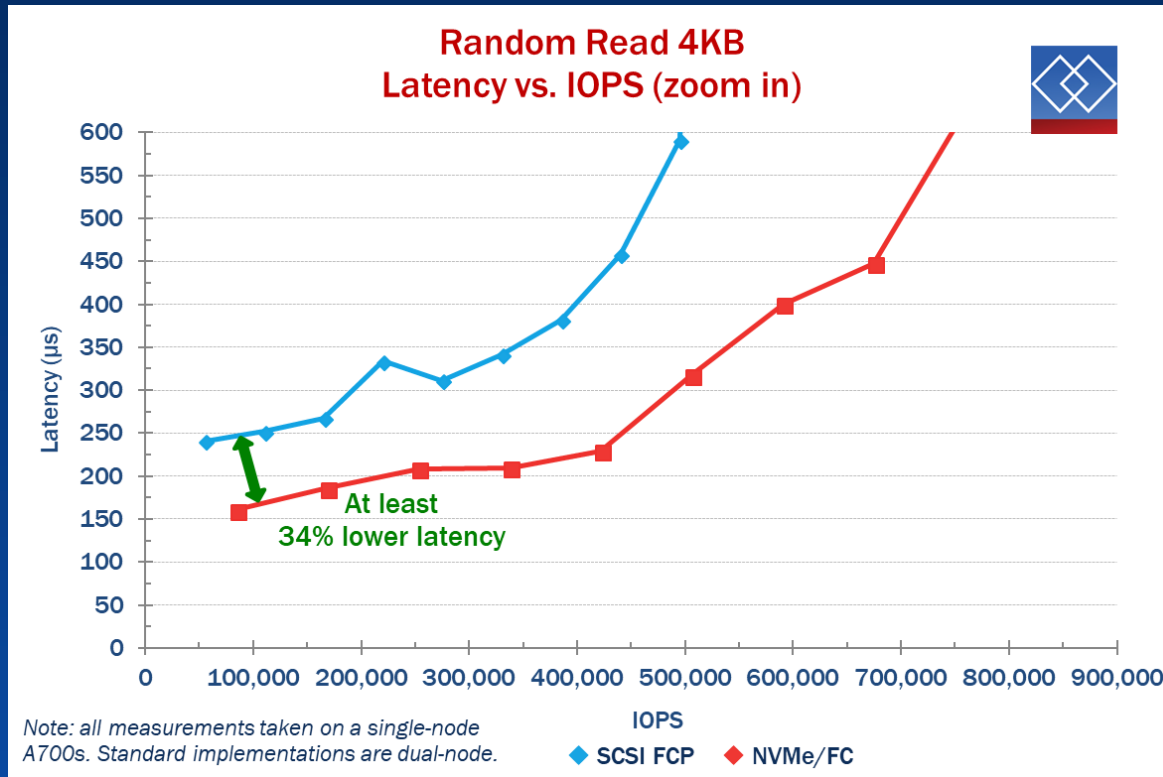


# Results: Random Read 4KB



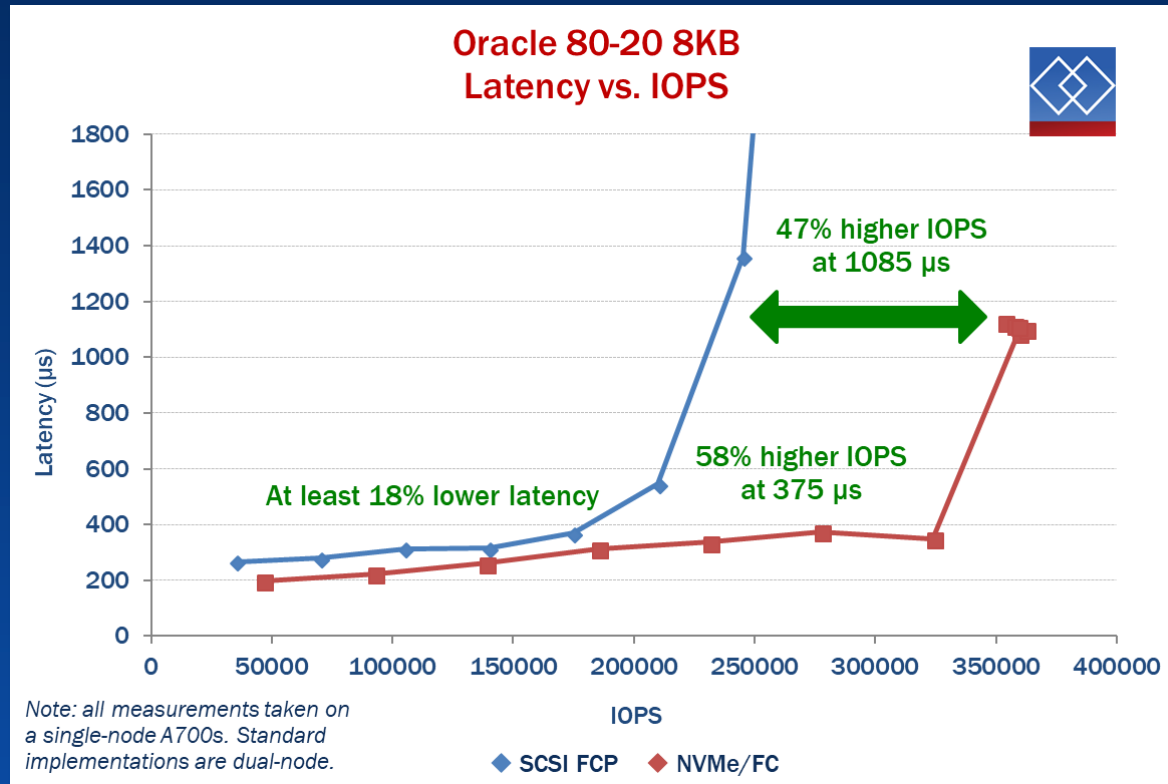


# Results: Random Read 4KB (zoom-in)





# Results: Oracle 80-20 8KB





Flash Memory Summit

## NVDIMM comments

- ◆ Faster technology can have some interesting effects.
- ◆ We installed some NVDIMMs in a server running Microsoft SQL Server. Because of the speed of the NVDIMMs, we had to adjust the SQL Server recovery interval setting. The default setting was slowing things down.

<https://docs.microsoft.com/en-us/sql/database-engine/configure-windows/configure-the-recovery-interval-server-configuration-option?view=sql-server-2017#SSMSProcedure>



Flash Memory Summit

# Industry Trends & Future Directions

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek

29



Flash Memory Summit

# NVMe over Fabrics Rules of Thumb



<https://www.demartek.com/NVMeoF-Rules/>

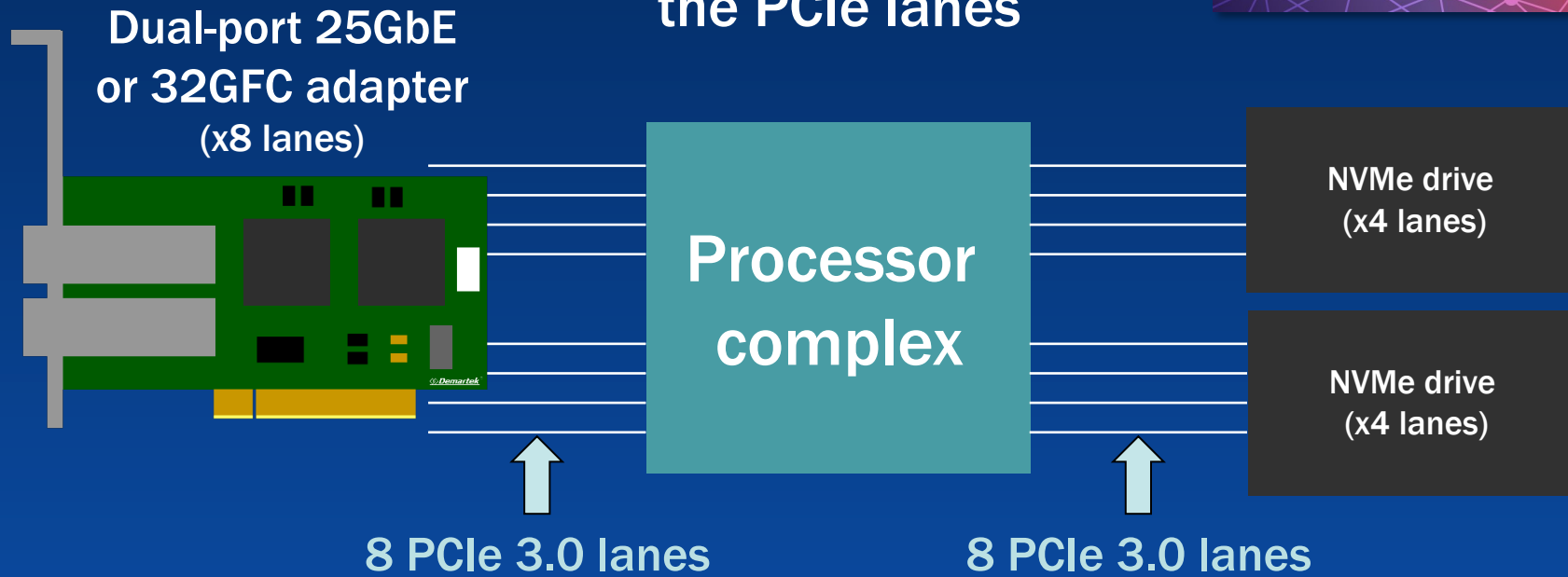
Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek



Flash Memory Summit

# Balanced Configuration without oversubscribing the PCIe lanes



<https://www.demartek.com/NVMeoF-rules/>



Flash Memory Summit

## Demartek 25GbE Deployment Tips

A graphic for '25 GbE' with the text in large white font. The background is a dark blue rectangle with glowing, swirling light trails in shades of green and yellow.

**25  
GbE**

**PRACTICAL TIPS FOR  
DEPLOYING 25GBE  
TECHNOLOGY...**

**BECAUSE THERE ARE  
SOME THINGS YOU NEED  
TO KNOW THAT MIGHT  
NOT BE OBVIOUS.**

*LEARN MORE >*

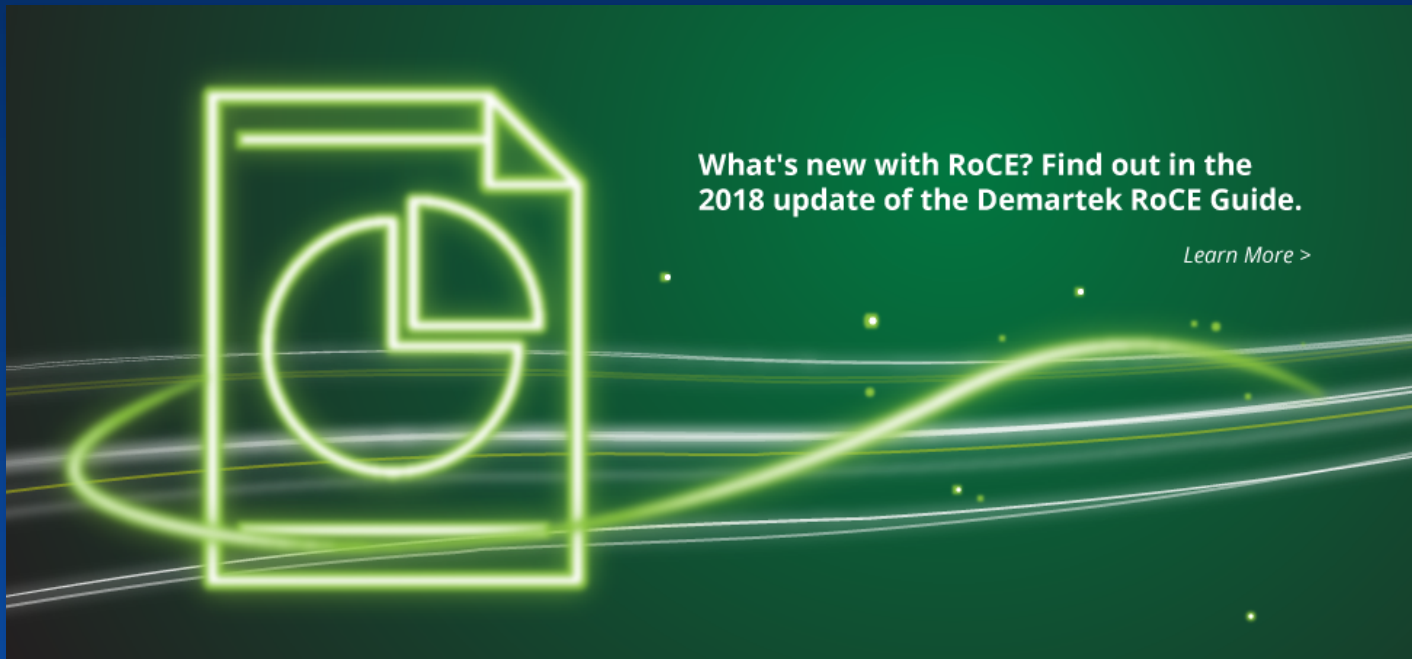
<https://www.demartek.com/25GbE-Tips/>





Flash Memory Summit

# Demartek RoCE Deployment Guide



<https://www.demartek.com/RoCE/>

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek



Flash Memory Summit

# Storage Interface Comparison

- ◆ Demartek Storage Interface Comparison reference page
  - ◆ Search engine: *Storage Interface Comparison*
  - ◆ Recent updates for PCIe 5.0, U.3, Fibre Channel, FC-NVMe & SATA



<https://www.demartek.com/Storage-Interface-Comparison/>

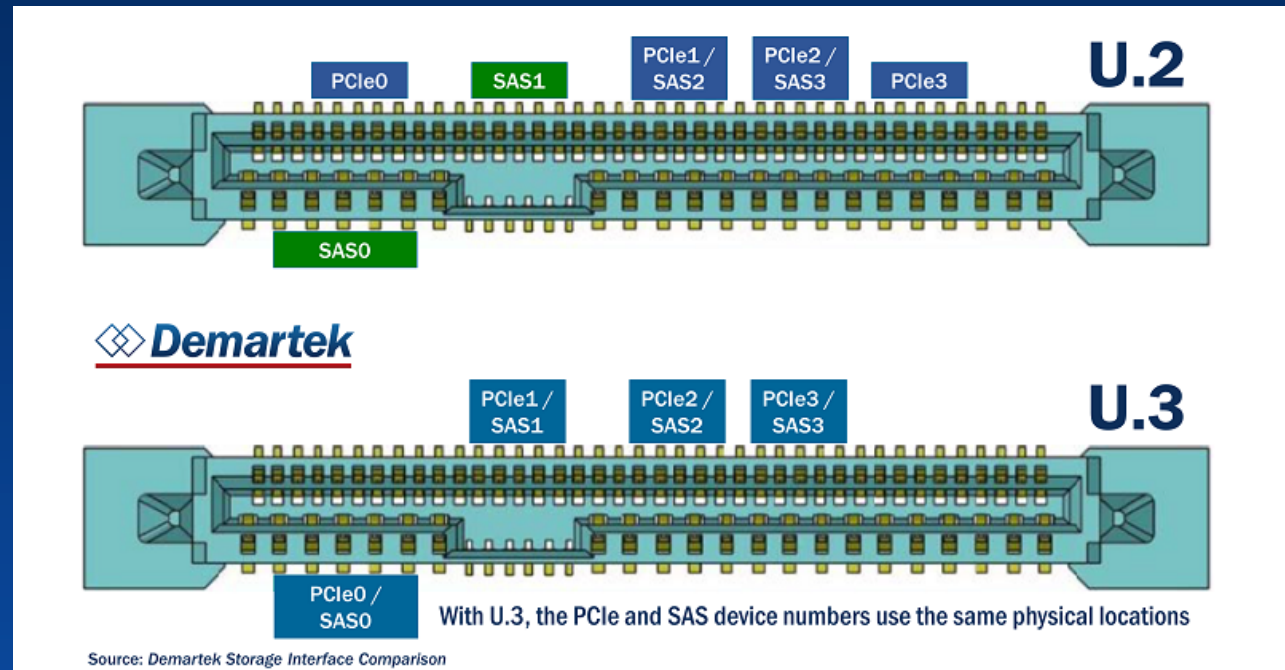


Flash Memory Summit

## U.2 and U.3 backplanes

U.2 – SFF8639

U.3 – SFF-TA-1001  
Rev. 1.0 was ratified  
in November 2017  
and Rev. 1.1 was  
ratified in May 2018.



<https://www.demartek.com/Storage-Interface-Comparison/>



Flash Memory Summit

## Roadmaps

- ◆ PCIe 4.0 – 1.0 spec. published October 2017
- ◆ PCIe 5.0 – revision 0.7 published May 2018
  - ◆ Target of Q1 2019 for spec. complete
- ◆ NVMe and NVMe over Fabrics (NVMe-oF) – next revision in 2019
- ◆ Ethernet & Fibre Channel – some of the same technology will drive single-lane 50GbE and 64GFC.

<https://www.demartek.com/Storage-Interface-Comparison/>



Flash Memory Summit

## Demartek Free Resources


- ◆ Demartek FC Zone – [www.demartek.com/FC/](http://www.demartek.com/FC/)
- ◆ Demartek iSCSI Zone – [www.demartek.com/iS/](http://www.demartek.com/iS/)
- ◆ Demartek NVMe Zone – [www.demartek.com/N/](http://www.demartek.com/N/)
- ◆ Demartek SSD Zone – [www.demartek.com/SSD/](http://www.demartek.com/SSD/)
- ◆ Demartek commentary: “Horses, Buggies and SSDs”  
[www.demartek.com/Demartek\\_Horses\\_Buggies\\_SSDs\\_Commentary.html](http://www.demartek.com/Demartek_Horses_Buggies_SSDs_Commentary.html)
- ◆ Demartek Video Library -  
[www.demartek.com/Demartek\\_Video\\_Library.html](http://www.demartek.com/Demartek_Video_Library.html)

Performance reports,  
Deployment Guides and  
commentary available  
for free download.




Flash Memory Summit

## This Presentation

 **Demartek**<sup>®</sup>

The Real Story on  
Flash Storage Performance

Session **TEST-101B-1**  
9:45 a.m. - 10:50 a.m. PDT,  
**Tuesday, August 7, 2018**  
Ballroom G

  
Flash Memory Summit

*JS.*

<https://www.demartek.com/FMS2018/>

Santa Clara, CA  
August 2018

The Performance Story: An Independent Evaluation of Flash Storage – by Demartek



Flash Memory Summit

# Thank You!



Demartek public projects and materials are announced on a variety of social media outlets. Follow us on any of the above.



Sign-up for the Demartek monthly newsletter, *Demartek Lab Notes*.  
[www.demartek.com/newsletter](http://www.demartek.com/newsletter)