



Flash Memory Summit

# The Straight Truth: How Today's Storage Performs on Real Workloads

TEST-101B-1

Dennis Martin, Principled Technologies



Flash Memory Summit

# Agenda

- Principled Technologies: Who we are and what we do
- Interface vs. device speeds
- Synthetic vs. real-world workloads
- Performance results
- Industry trends & future directions

(If you'd like a copy of this presentation, I provide my contact information at the end)



Flash Memory Summit

# Demartek → Principled Technologies

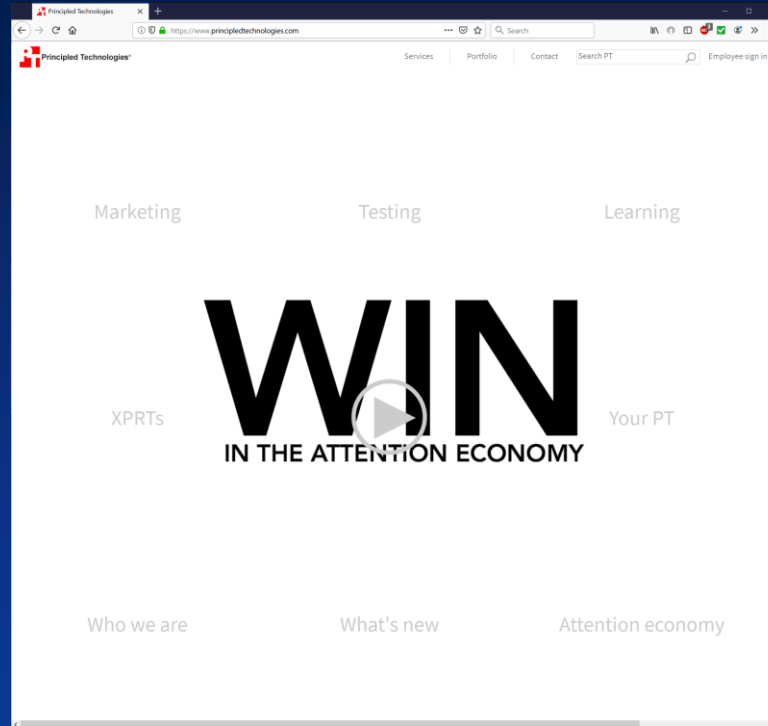
- I've spoken here at FMS for several years
- Many of you know me as the President and Founder of Demartek
- Principled Technologies (PT) acquired Demartek in September 2018
- Introducing PT...





Flash Memory Summit

# About Principled Technologies



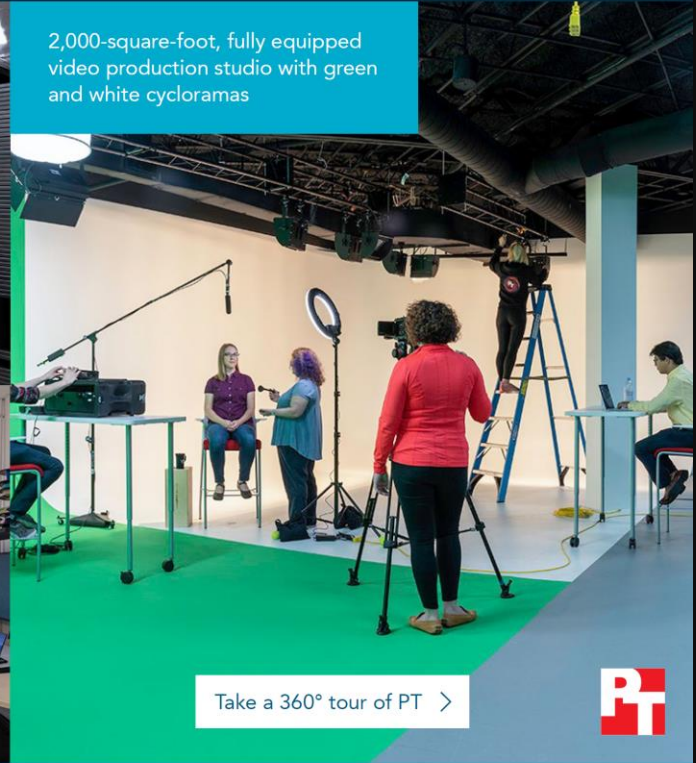
<https://facts.pt>

# Get the resources your project needs with top-of-the-line testing and creative facilities.

Modern, on-premises data centers



2,000-square-foot, fully equipped video production studio with green and white cycloramas



Dedicated client test beds



Take a 360° tour of PT >



PROVE BENEFITS

WIN ATTENTION

SELL MORE

CHOOSE PT

CONTACT US

<https://facts.pt/FBM-outcomes>



Flash Memory Summit

# BenchmarkXPRT

## BenchmarkXPRT

AIXPRT WebXPRT MobileXPRT BatteryXPRT TouchXPRT CrXPRT HDXPRT

The BenchmarkXPRT Development Community provides registered members with the opportunity to contribute to the process of creating and improving the XPRTs – **benchmark apps** that measure how well **PCs, tablets, and smartphones** handle everyday tasks. The XPRTs empower people all over the world to know how well their gear handles workloads like editing photos, playing movies, and browsing the Web.

<http://benchmarkxprt.com>



# BenchmarkXPRT selector

principlotechnologies.com/... X

https://www.principlotechnologies.com/benchmarkxprt/the-xprt-selector/

## BenchmarkXPRT

Whether you're using, making, or writing about tech products, you need to know how well they'll do what you want them to do.

Adjust the dial to find the testing app for your device

<http://facts.pt/the-xprt-selector>



Flash Memory Summit

# Interface vs. device speeds





# Interface vs. device speeds

	Interface speeds	Device speeds
Application	Bluetooth, Ethernet, Fibre Channel, InfiniBand, Mobile phones (3G / 4G / 5G), SAS, SATA, Thunderbolt, USB, Wi-Fi	HDDs, SSDs, NVMe, PCIe, consumer device storage cards (SD, CompactFlash, etc.), Memory (DRAM, NVDIMM, etc.)
Throughput	Bits per second	Bytes per second
Abbreviation	Mbps, Gbps, or Mb/s, Gb/s	MBps, GBps, or MB/s, GB/s
Case	Lowercase "b"	Uppercase "B"

*Converting from bits per second to bytes per second is **not** a simple divide by 8. You must know the **encoding scheme** for that interface.*



# What about memory?

Memory	
Addressing	Byte-addressable Applies to volatile and non-volatile types
Speed	Clock rate, such as 2133 MHz
Throughput	Gigabytes per second (GBps or GB/s)



Flash Memory Summit

# Synthetic vs. real-world workloads



# Synthetic vs. real-world workloads

<i>Storage testing</i>	<b>Synthetic</b>	<b>Real-world</b>
<b>I/O profile</b>	Fixed, controlled, specific block size, read/write mix, length of time, random vs. sequential, etc.	Variable block sizes, read/write mix, length of time, random, sequential, etc. May use multiple different parameters simultaneously. May issue few or no I/O's for some periods of time.
<b>CPU usage</b>	Low to medium, but generally steady, focused on I/O tasks	Variable: low to high (performs other tasks in addition to I/O)
<b>Memory usage</b>	Generally fixed and small	Variable: small to large
<b>Results</b>	"Hero numbers"	Application-specific, such as database transactions per second, orders per minute, etc.
<b>Who</b>	Storage product vendors	End-users



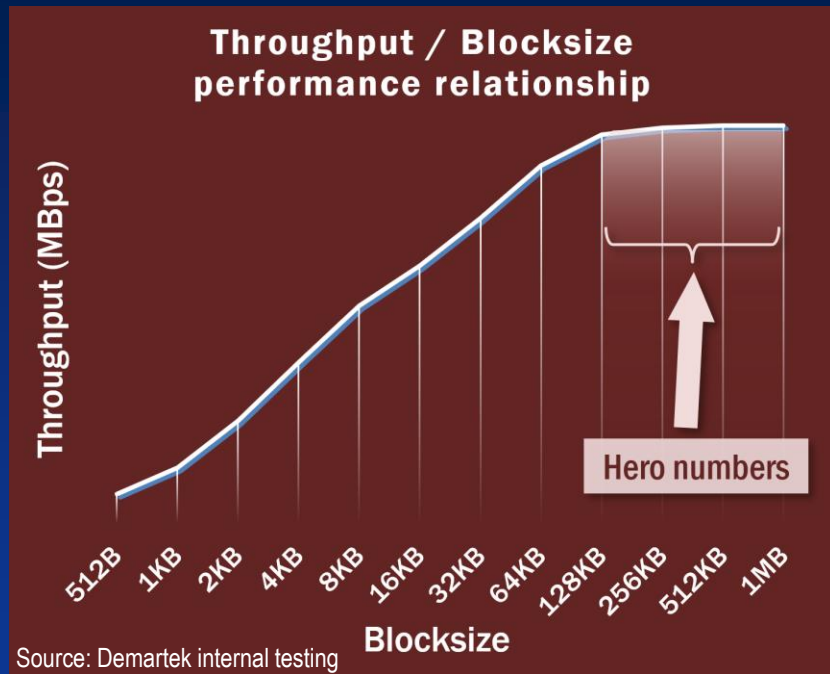
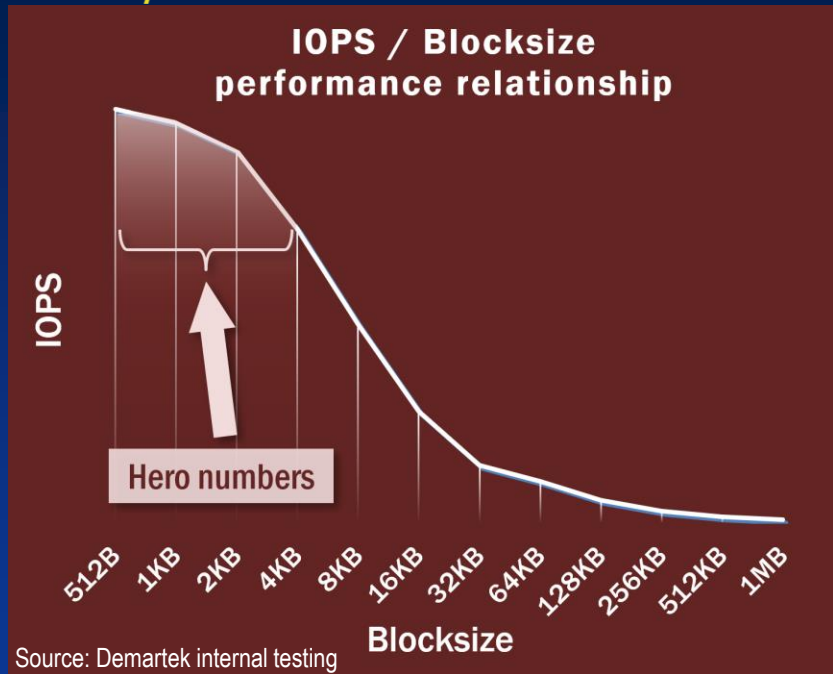
# Real-world workload types

	Transactional	Streaming
I/O pattern	Mostly random	Mostly sequential
Emphasis	I/O's per second (IOPS)	Throughput (MBps or GBps)
Latency	Important	Not very important

*We observe different latency result patterns for different workload types.*



# Blocksize performance relationships (synthetic benchmarks)



These generic performance curves apply to network and storage performance.  
For network performance, replace *IOPS* with *Frames per Second (FPS)*.

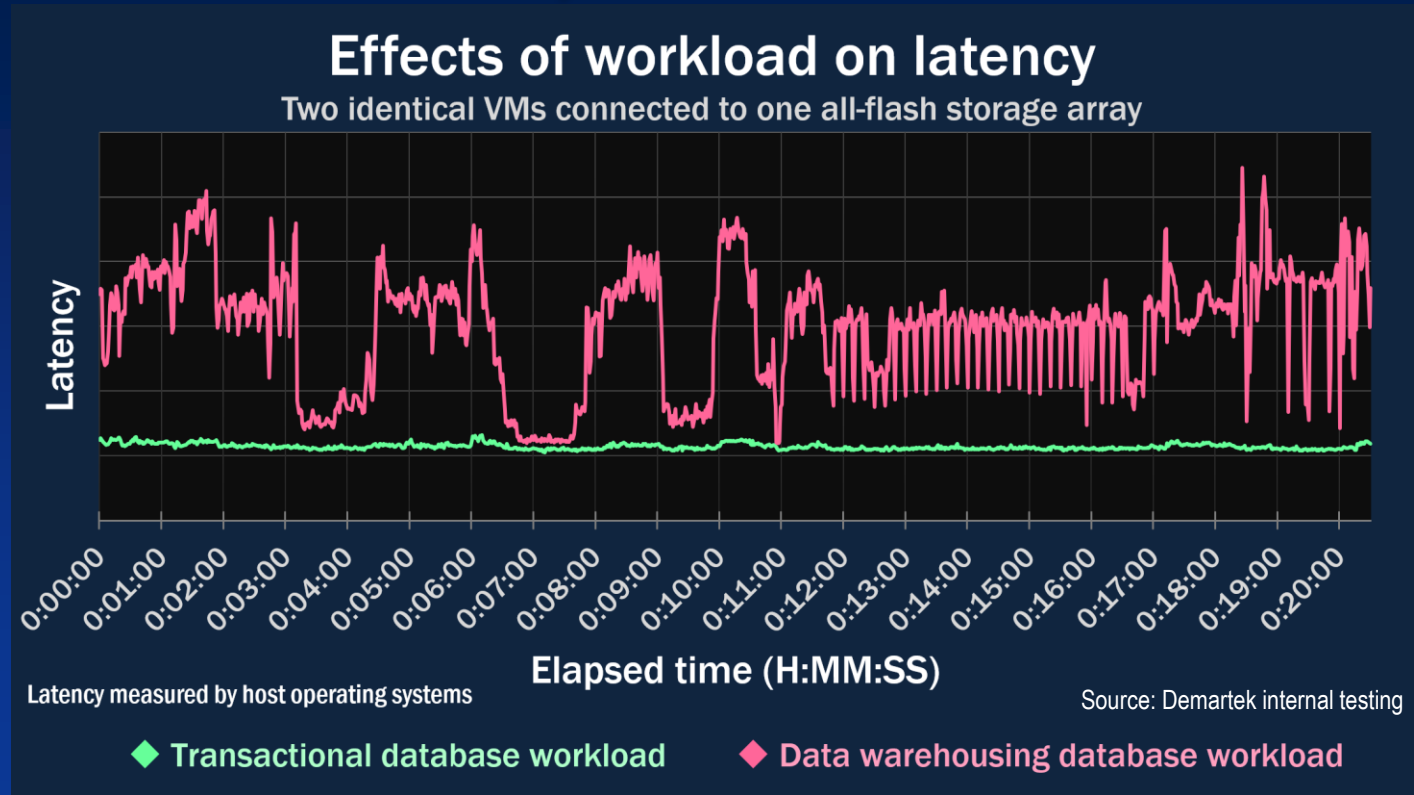


# Generic latency results (real-world workloads)

One all-flash array.  
Two different  
workloads running  
simultaneously.

The nature of each  
workload has a large  
impact on latency.

At 06:00 & 10:00 the  
pink workload  
affected the latency of  
the green workload.

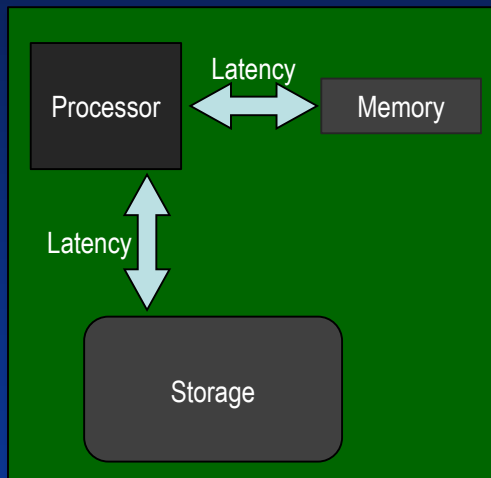




# Real-world latency measurements

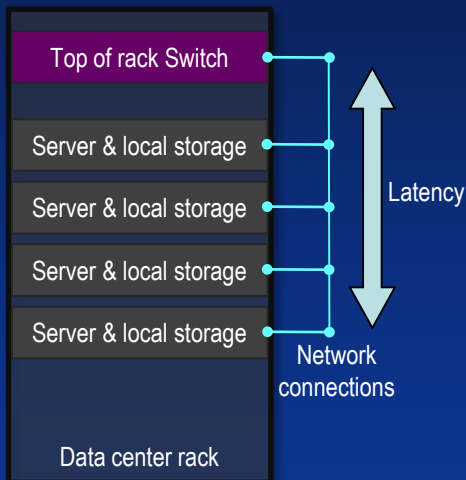
*We measure end-to-end latency from the source to the target and back to the source, through all the hardware and software layers*

## Single-server solution



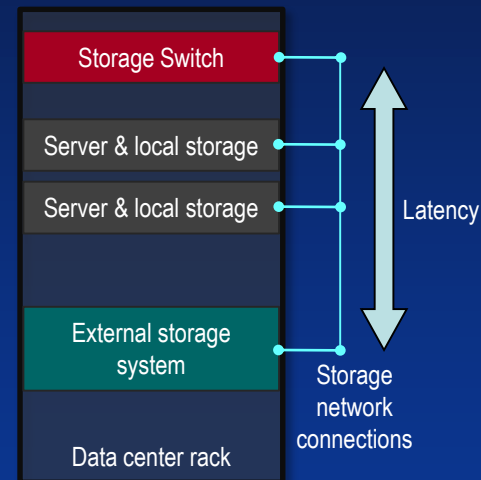
- Between processor and memory
- Between processor and storage

## Hyper-converged solution



- Between server and local storage
- Between server and remote storage in a different node

## External storage solution



- Between server and external storage system





Flash Memory Summit

# Performance results



## Flash Memory Summit

# Desktop with Intel Optane memory

Two desktop computers:

- Intel Core i5 8600

Machine 1:

- 32GB RAM
- \$2012

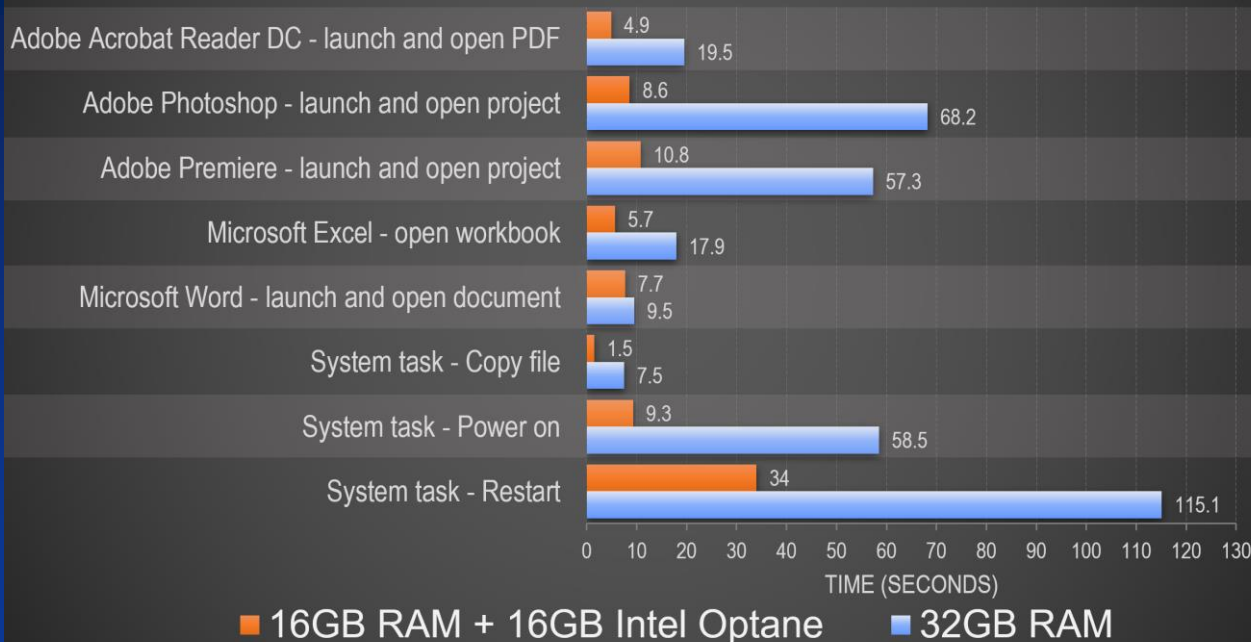
Machine 2:

- 16GB RAM
- 16GB Intel Optane mem.
- \$1706

<http://facts.pt/pqv8ua2>

Intel Optane in memory  
mode as a cache

## Time to complete tasks (Lower is better)





## Flash Memory Summit

# Intel Optane SSDs in VMware vSAN

## 4-node VMware vSAN cluster

### Each server:

- 2x Intel Xeon® Gold 6154 CPU @ 3.00GHz
- 384 GB RAM

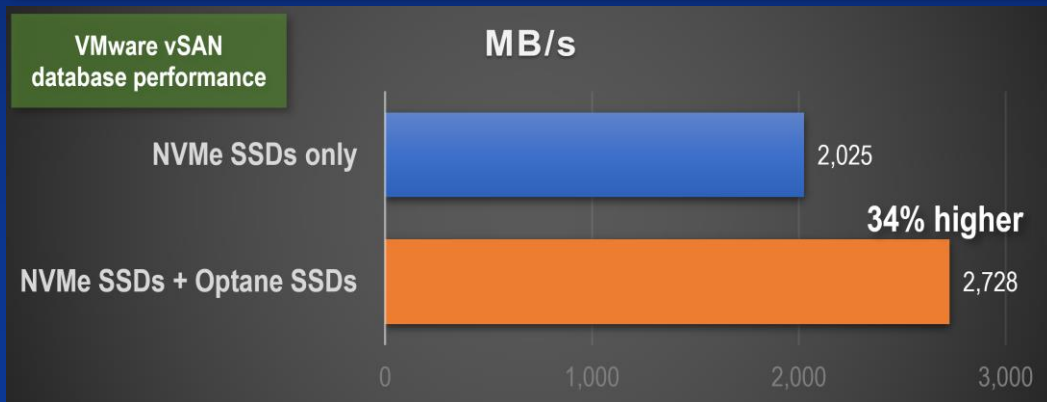
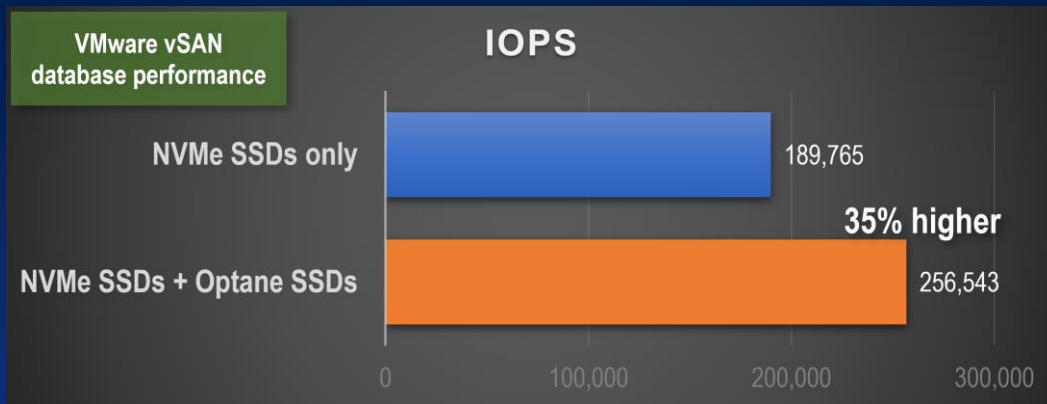
### Config 1

- 6x 2 TB NVMe SSDs (capacity)
- 2x 2 TB NVMe SSDs (cache)

### Config 2

- 6x 2 TB NVMe SSDs (capacity)
- 2x 375 GB Optane SSDs (cache)

<http://facts.pt/qtaj3ob>





# Upgrade to NVMe and Optane persistent memory

## Storage config 1 (HDD)

- **Data:** 4-drive RAID10 volume
- **Logs:** 2-drive RAID1 volume

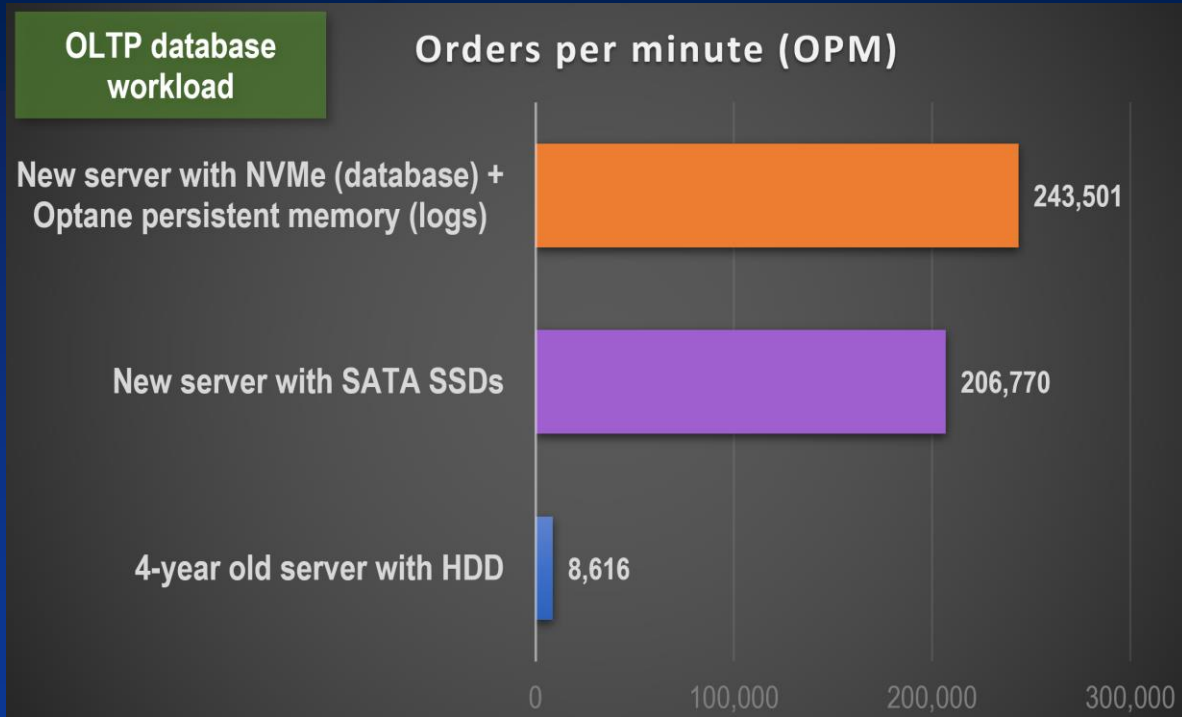
## Storage config 2 (SATA SSD)

- **Data:** 4-drive RAID10 volume
- **Logs:** 2-drive RAID1 volume

## Storage config 3 (NVMe + Optane)

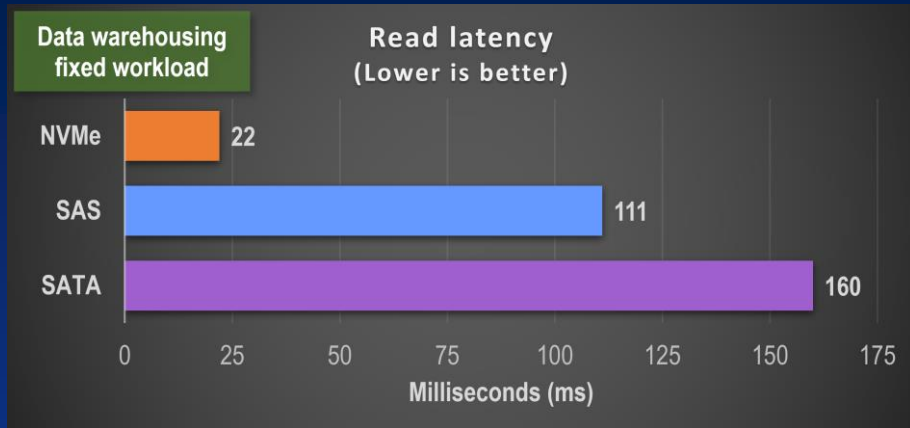
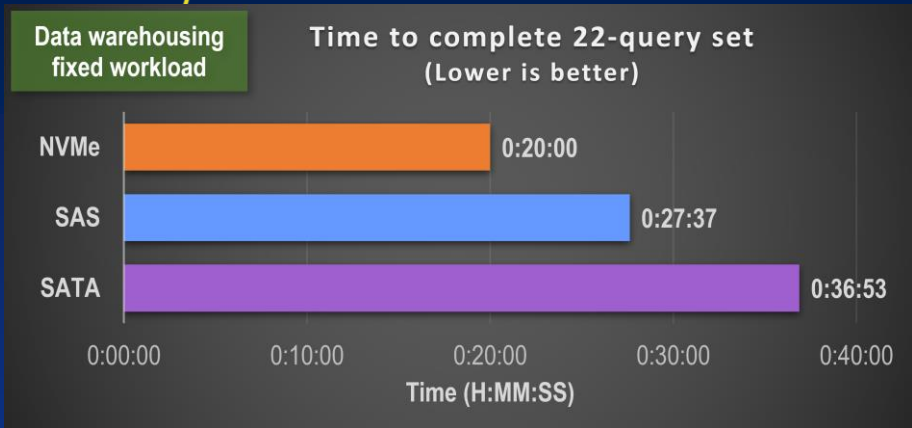
- **Data:** 1 Intel Optane NVMe SSD
- **Logs:** Intel Optane persistent memory (Non-Interleaved App Direct Mode)

<http://facts.pt/1mgym04>





# SATA vs. SAS vs. NVMe storage



## Storage configurations

- **NVMe**: 8x 960GB NVMe SSD
- **SAS**: 8x 960GB SAS SSD
- **SATA**: 8x 960GB Enterprise SATA SSD

## Server configuration

- 2x AMD EPYC 7601, 32c/64t
- 128GB RAM

<http://facts.pt/2h8emuf>



Flash Memory Summit

# Industry trends & future directions



# PCI Express (PCIe)

	Specification announced	Specification completed	Gigatransfers per second (GT/s)	x1 bandwidth (half-duplex)	x16 bandwidth (half-duplex)
PCIe 4	November 2011	October 2017	16	2 GB/s	32 GB/s
PCIe 5	June 2017	May 2019	32	4 GB/s	64 GB/s
PCIe 6	June 2019	2021 (estimate)	64	8 GB/s	128 GB/s

*Half-duplex speeds indicate the maximum rate for storage reads or writes in one direction only. For transmissions in both directions simultaneously (full-duplex), double the rate.*

Source: PCI-SIG <https://pcisig.com/>



# Ethernet, Fibre Channel & InfiniBand roadmaps

	Single-lane speeds (Gbps)	Four-lane speeds (Gbps)	Eight-lane speeds (Gbps)	Twelve-lane speeds (Gbps)
Ethernet	1, 2.5, 5, 10, 25, <b>50*</b>	40, 100, <b>200*</b>	100, <b>400*</b>	-
Fibre Channel	8, 16, 32, <b>64*</b>	128, <b>256*</b>	-	-
InfiniBand	8, 14, 25, <b>50*</b>	32, 56, 100, <b>200*</b>	-	96, 168, 300, <b>600*</b>

*Ethernet, Fibre Channel and InfiniBand are all dependent on two underlying technologies and will achieve new speed increments in similar time frames. The two technologies are: **PCI Express** and **Transceiver technology**.*

*\* Speeds in **purple** generally require at least one PCIe 4.0 slot. Links to public roadmaps listed on the “References” page later in this presentation.*





# NVM Express (NVMe) news

“BEAVERTON, Ore.,—USA—July 23, 2019—NVM Express™, Inc. today announced the release of NVM Express™ (NVMe™) 1.4 Base Specification and that NVMe™ over Fabrics (NVMe-oF™) 1.1 specification has entered into final 45-day member review.”

<https://nvmexpress.org/new-nvm-express-inc-specifications-bolster-cloud-and-enterprise-advancements/>

- You will learn much more about **NVMe 1.4** and **NVMe-oF 1.1** at this conference in the NVMe track and in the expo hall.
- NVMe-oF 1.1 includes NVMe/TCP.



# References

- Ethernet Alliance 2019 public roadmap:  
<https://ethernetalliance.org/the-2019-ethernet-roadmap/>
- Fibre Channel public roadmap:  
<https://fibrechannel.org/roadmap/>
- InfiniBand public roadmap:  
<https://www.infinibandta.org/infiniband-roadmap/>
- PCI Express version 6 announcement:  
<http://pcisig.com/pioneering-interconnect-industry-pci-sig%C2%AE-announces-upcoming-pcie%C2%AE-60-specification>



# Questions?

- I'm happy to answer questions after this session
  - I'll be available during this conference
- Send me a message on LinkedIn
  - <https://www.linkedin.com/in/dennismartin>
  - To ask questions about this presentation
  - To get a copy of this presentation  
(not putting my email here to avoid spam)





Flash Memory Summit



Principled Technologies posts public projects and materials to [our website](#) and announces them on a variety of social media outlets. Find us on any of the above.