IDF16
INTEL DEVELOPER FORUM

WICKED FAST STORAGE AND BEYOND

Dr. Frank T. Hady Intel Fellow
Chief Architect of Intel® Optane™ SSDs
Non-Volatile Memory Solutions Group
Intel Corporation

Wilson Fang Intel Optane SSD Product Team Lead
Non-Volatile Memory Solutions Group
Intel Corporation

intel
experience what's inside™

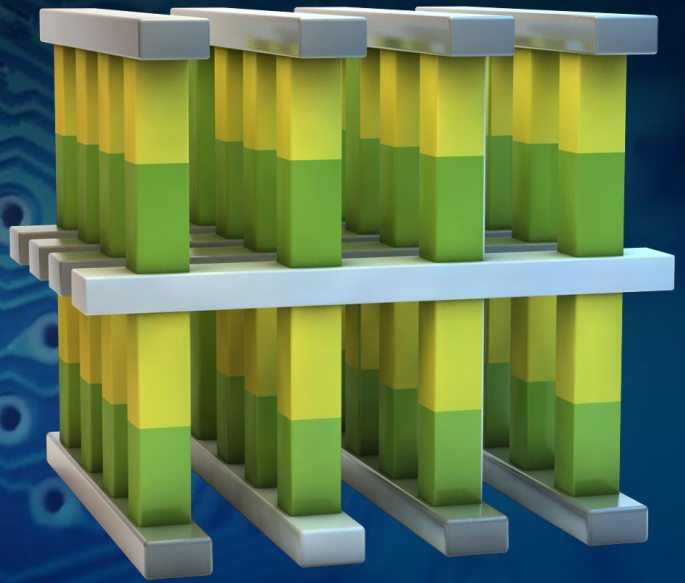
Agenda

- *Wicked Fast Storage*: Intel® Optane™ SSDs Based on 3D XPoint™ Technology
- Client/Workstation Impact (and why)
- Data Center Impact (and why)
- *And Beyond*: Intel® DIMMs Based on 3D XPoint™ Technology

3D XPoint™ Technology

IDF16
INTEL DEVELOPER FORUM

Cross Point Structure
Selectors allow dense packing and individual access to bits



Breakthrough Material Advances
Compatible switch and memory cell materials

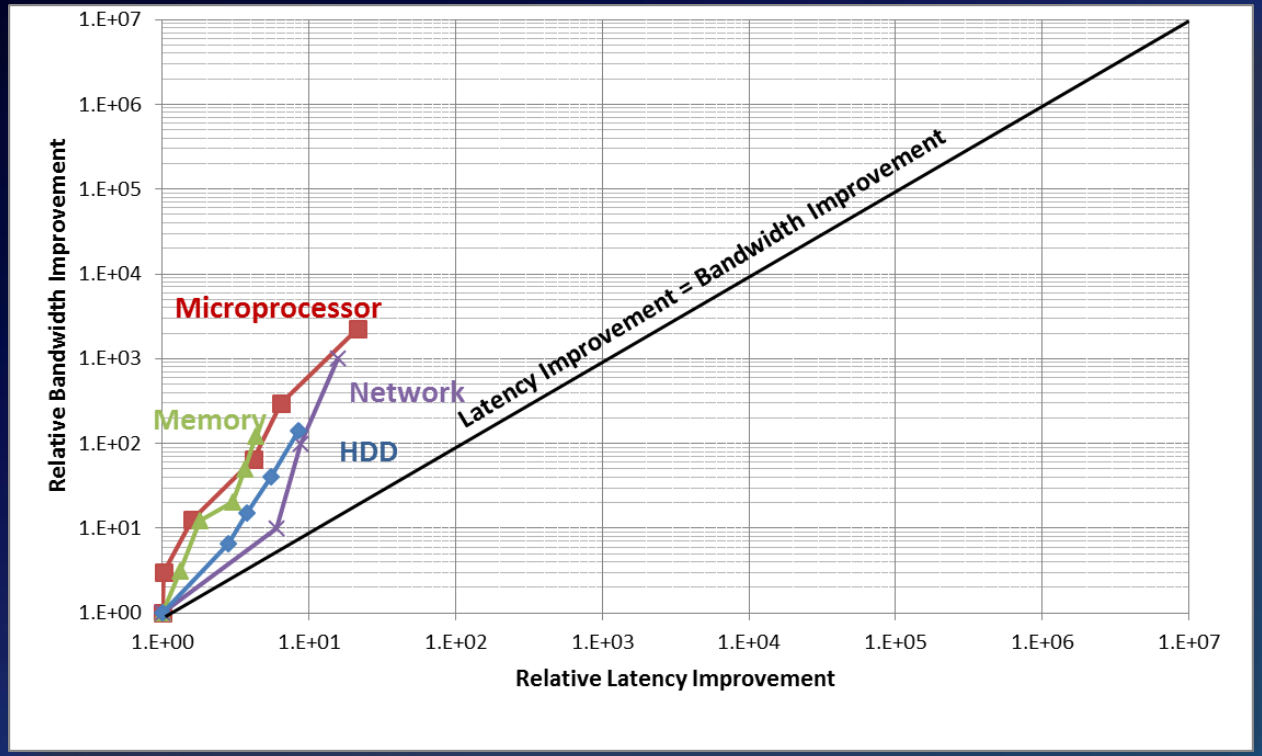
Scalable
Memory layers can be stacked in a 3D manner

High Performance
Cell and array architecture that can switch states 1000x[1] faster than NAND

intel
experience what's inside™

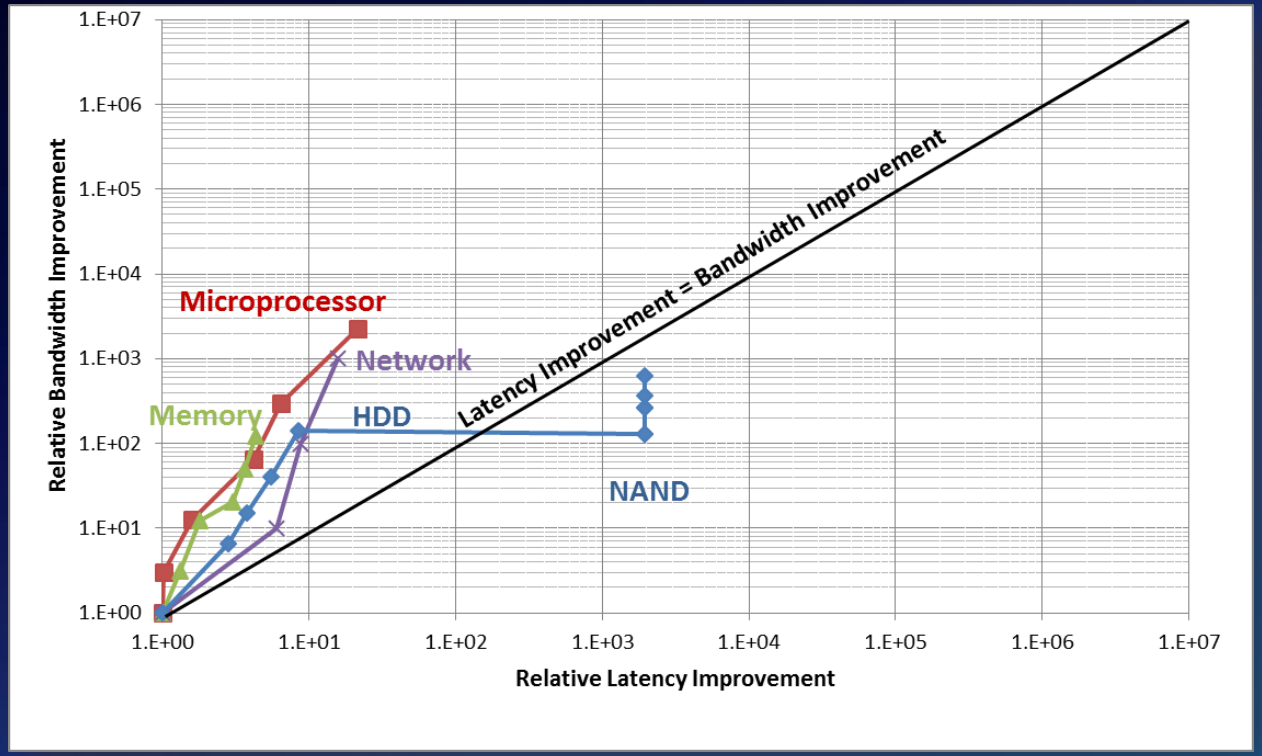
[1]Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.

Historically Significant Storage Advances



Source: "Latency lags Bandwidth" – David Patterson Comms. of the ACM, Oct 2004 Vol 47, No 10

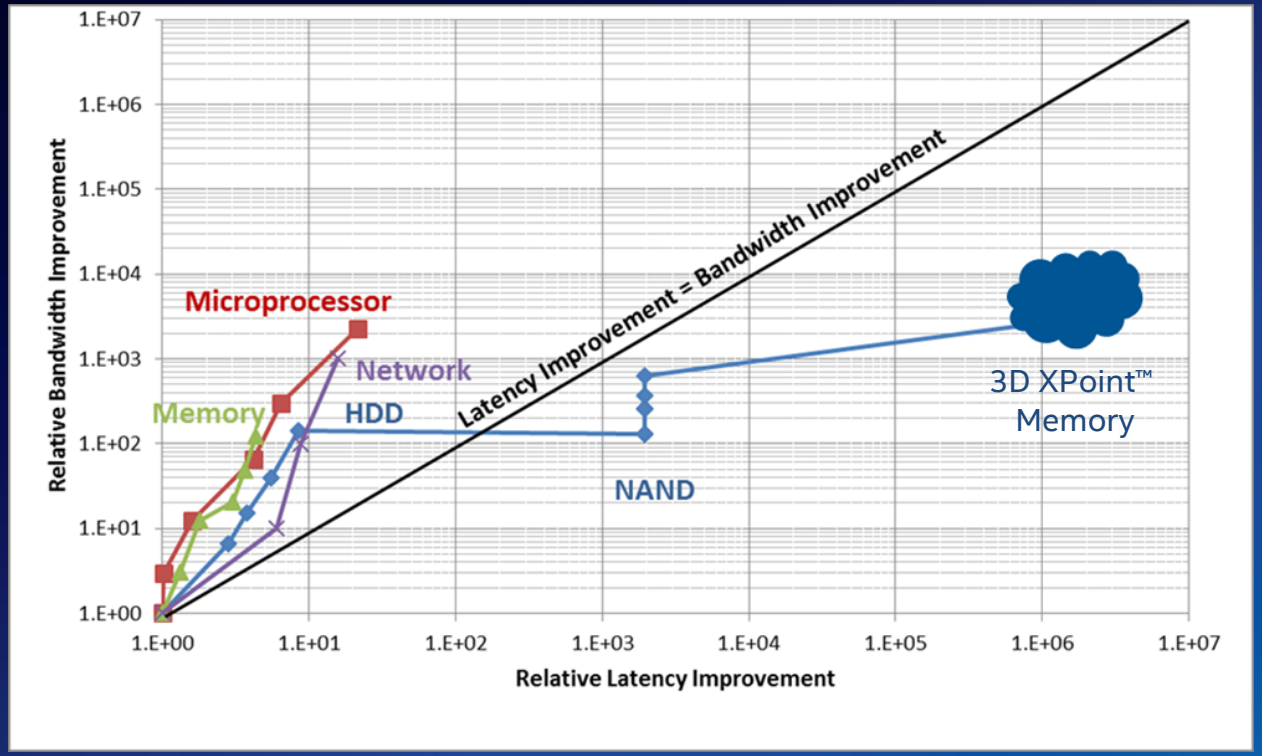
Historically Significant Storage Advances



"There is an old network saying: Bandwidth problems can be cured with money. Latency problems are harder because the speed of light is fixed—you can't bribe God."
- Anonymous

Source: "Latency lags Bandwidth" – David Patterson Comms. of the ACM, Oct 2004 Vol 47, No 10
With NAND and 3D XPoint™ Memory added by Al Fazio

Historically Significant Storage Advances



"There is an old network saying: Bandwidth problems can be cured with money. Latency problems are harder because the speed of light is fixed—you can't bribe God."
- Anonymous

Source: "Latency lags Bandwidth" – David Patterson Comms. of the ACM, Oct 2004 Vol 47, No 10
With NAND and 3D XPoint™ Memory added by Al Fazio

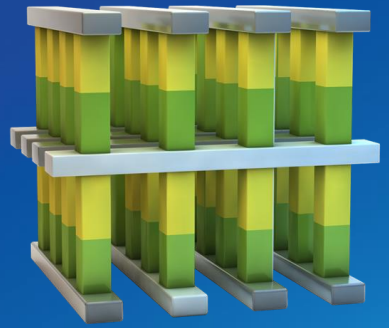
Memory and Storage Platform Connection

IDF16 INTEL DEVELOPER FORUM

Interfaces		Relative Latency (reads)
	Processor	
On Core CPU On Die	**SRAM Cache**	1x
Direct Attach	**DDR DRAM**	10x
PCI Express* NVM Express*	**SSD**	100,000x
SAS, SATA	**HDD**	10,000,000x

Performance

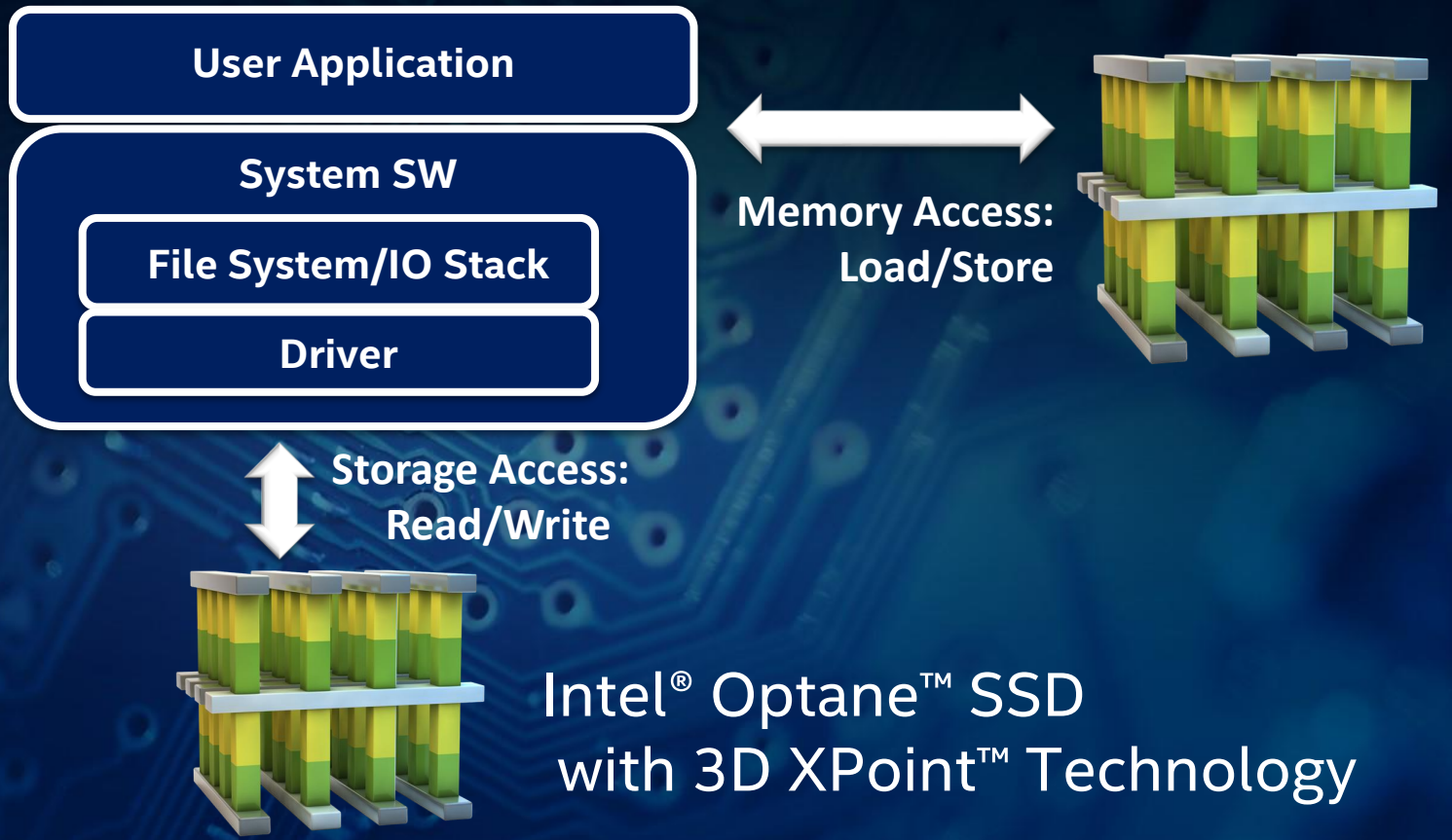
Cost



intel experience what's inside™

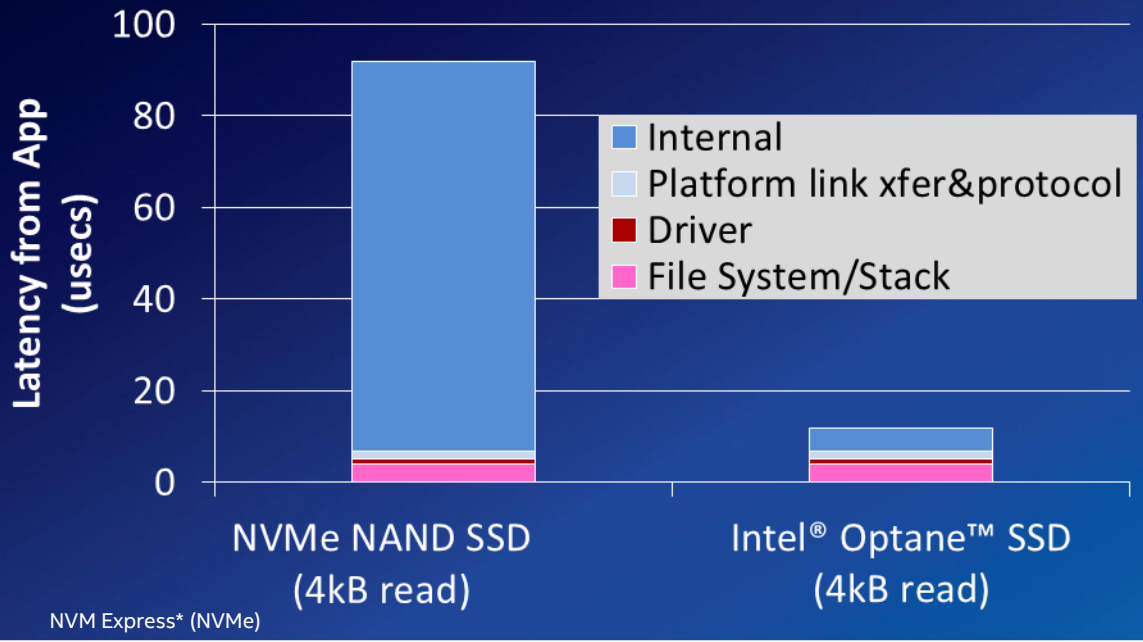
[1]Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.

Memory and Storage Platform Connection



Intel® Optane™ SSD
with 3D XPoint™ Technology

Storage System Interconnect



NVM Express* (NVMe)

Low Latency of Intel® Optane™ SSDs *is* accessible by User Applications on today's systems

Sources: "Storage as Fast as the rest of the system" 2016 IEEE 8th International Memory Workshop and measurement, Intel® Optane™ SSD measurements and Intel SSD P3700 Series measurements

IDF16
INTEL DEVELOPER FORUM

Intel® Optane™ SSD Prototype Performance Demonstration

intel
experience what's inside™

Intel® Optane™ SSD Latency

	Intel® SSD DC P3700 Series	Intel® Optane™ SSD Proto
RndRd 4K, QD 1	11,600	130,000
70/30 4k, QD 1	14,000	130,000

1/IOPs

Latency Math (Rnd Rd 4k, QD 1)		
Total = 1/IOPs (App level)	86 us	~8 us
Minus Systems HW/SW	~1 us	~1 us
Equals SSD HW Latency	~85 us	~7 us

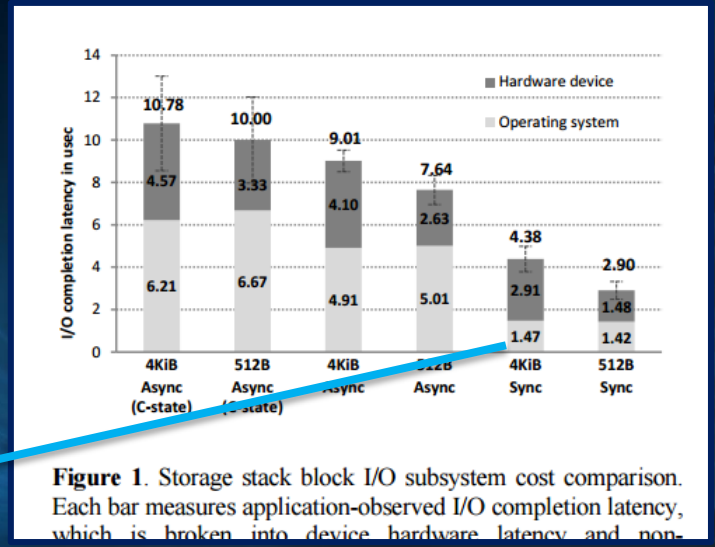


Figure 1. Storage stack block I/O subsystem cost comparison. Each bar measures application-observed I/O completion latency, which is broken into device hardware latency and non-

Source: "When Polls is Better than Interrupt" Jisoo Yang, FAST 2012 (Linux 2.6.33 Kernel)

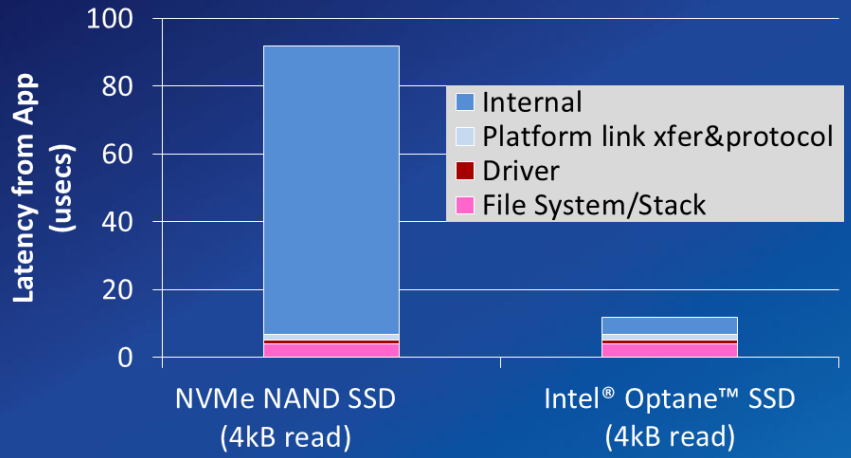
Intel® Optane™ SSD delivers 10x[1] advantage in SSD latency

Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, HT off, CPU C-state off, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11
[1]Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.

SSD Performance Metrics that Matter

IDF16 INTEL DEVELOPER FORUM

- Latency
- Throughput (IOPs)
- Quality of Service -
  - variation in latency and IOPs

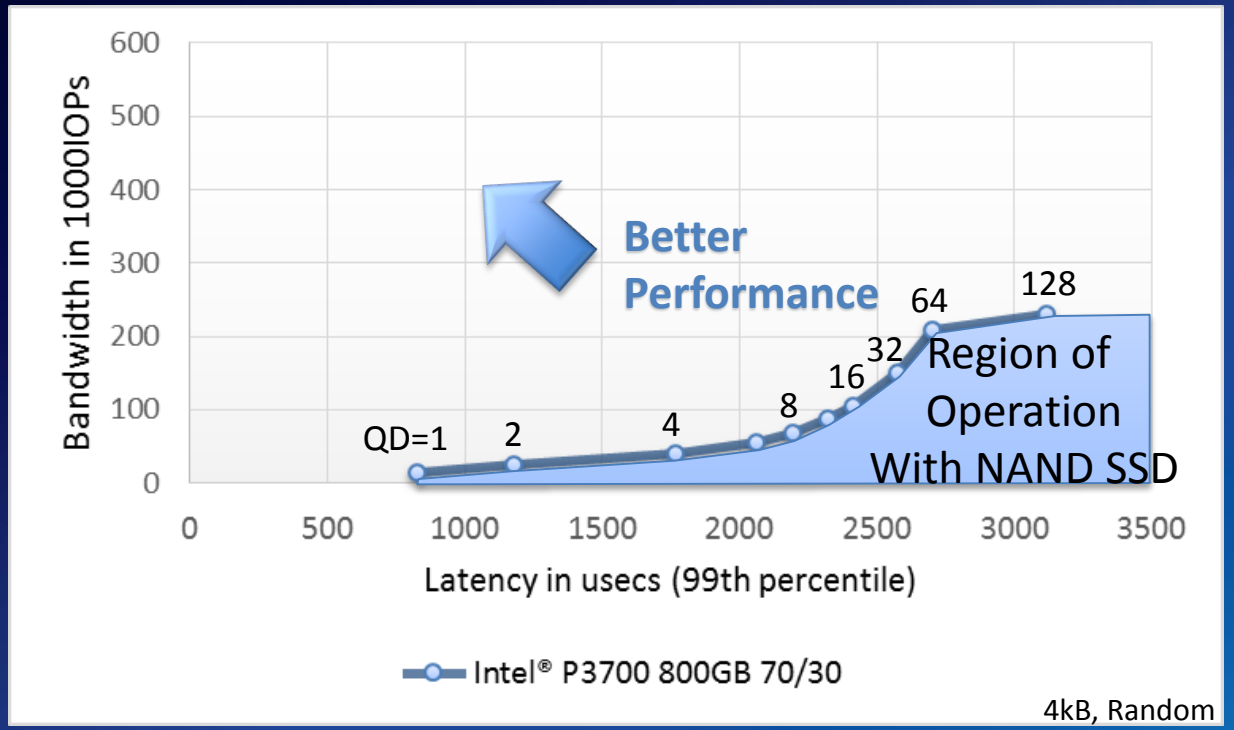


- Not Queue Depth -
  - function of interaction between application, OS, platform AND SSD

Sources: "Storage as Fast as the rest of the system" 2016 IEEE 8th International Memory Workshop and measurement, Intel® Optane™ SSD measurements and Intel SSD P3700 Series measurements

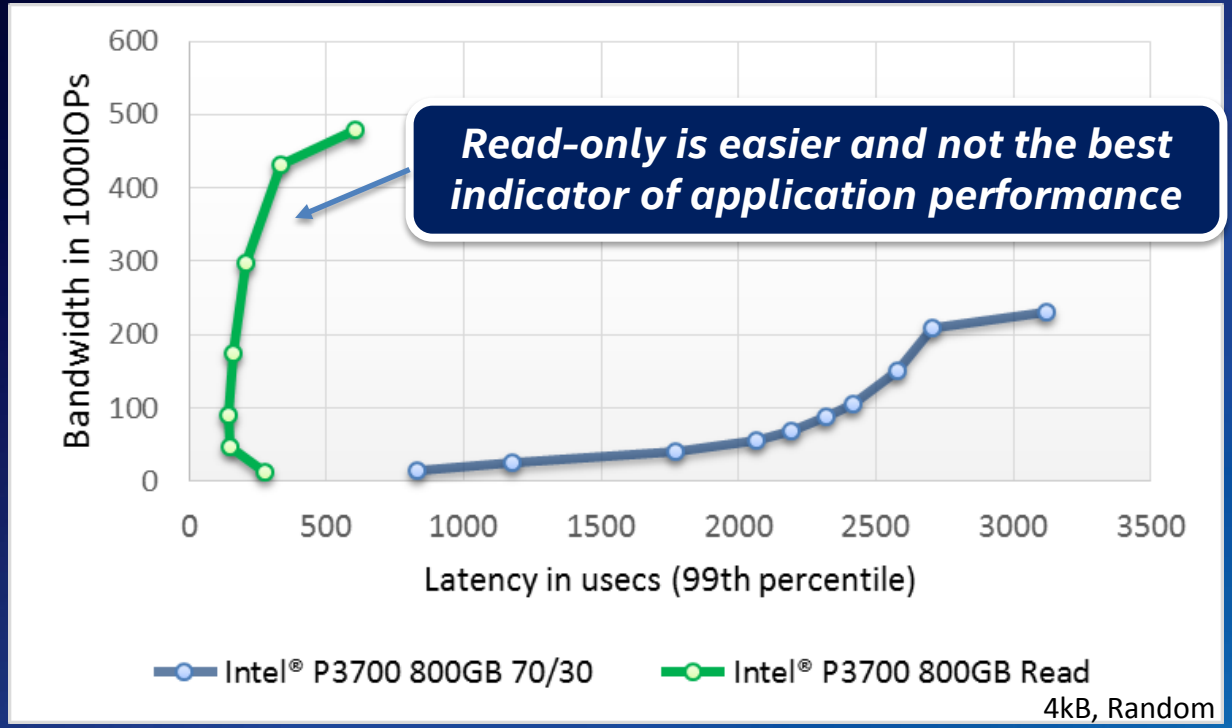
intel experience what's inside™

SSD Performance



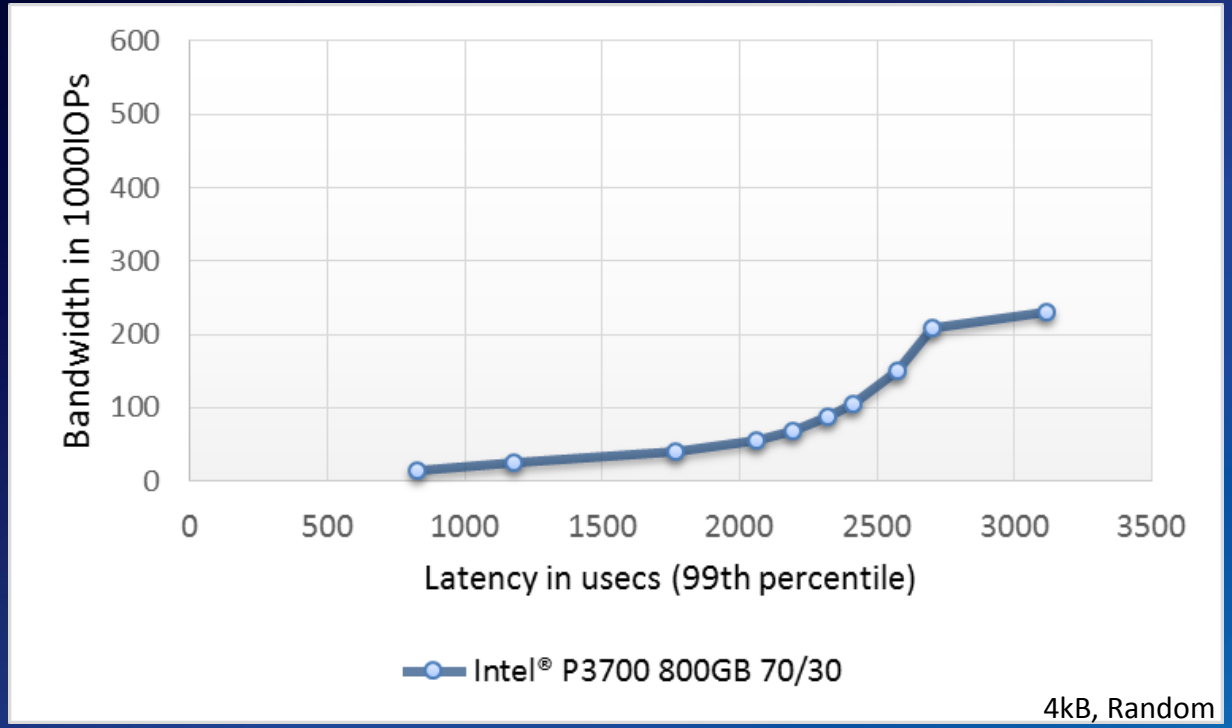
Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

SSD Performance



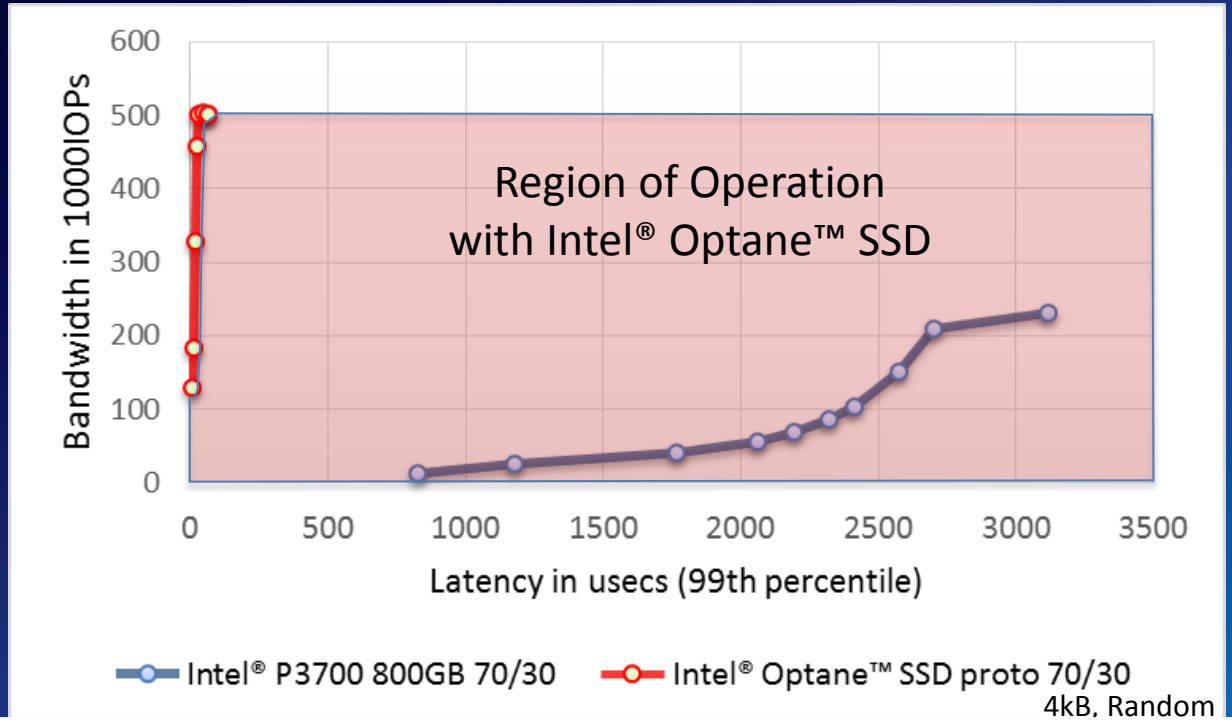
Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

SSD Performance



Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

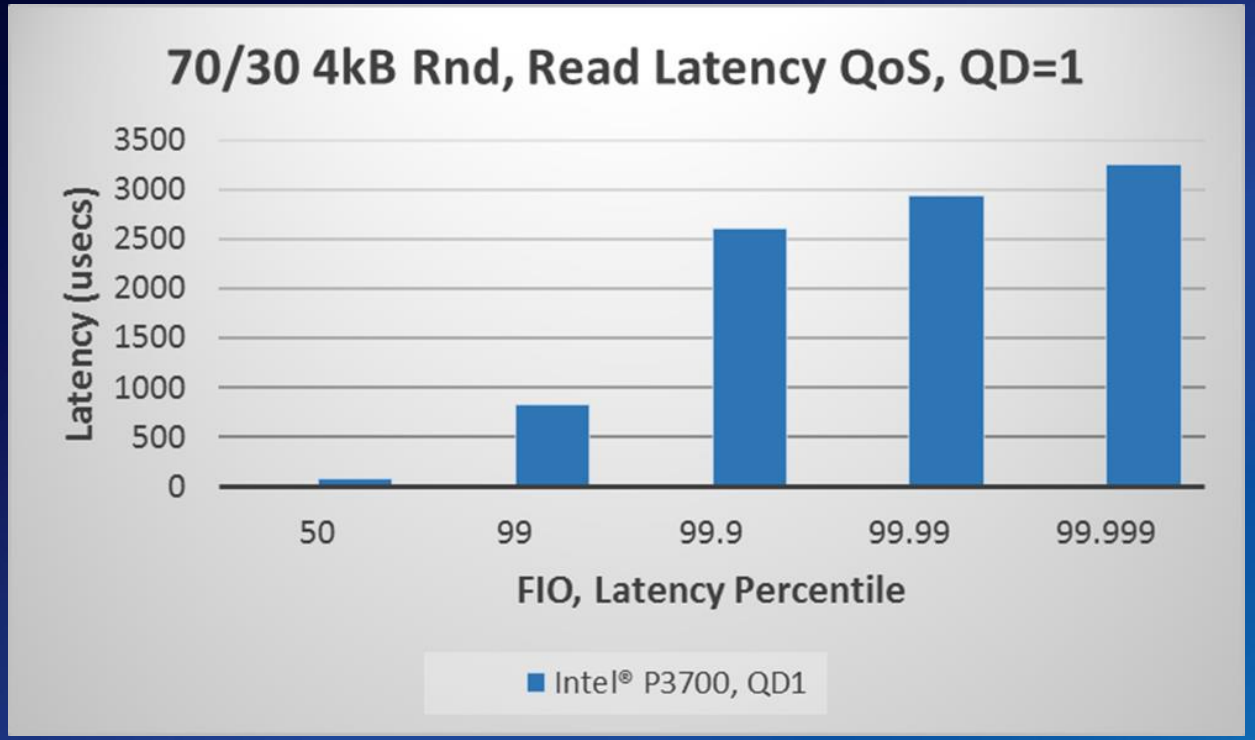
SSD Performance



Intel® Optane™ SSDs change the game, operating at previously unreachable IOPs/Latency combinations

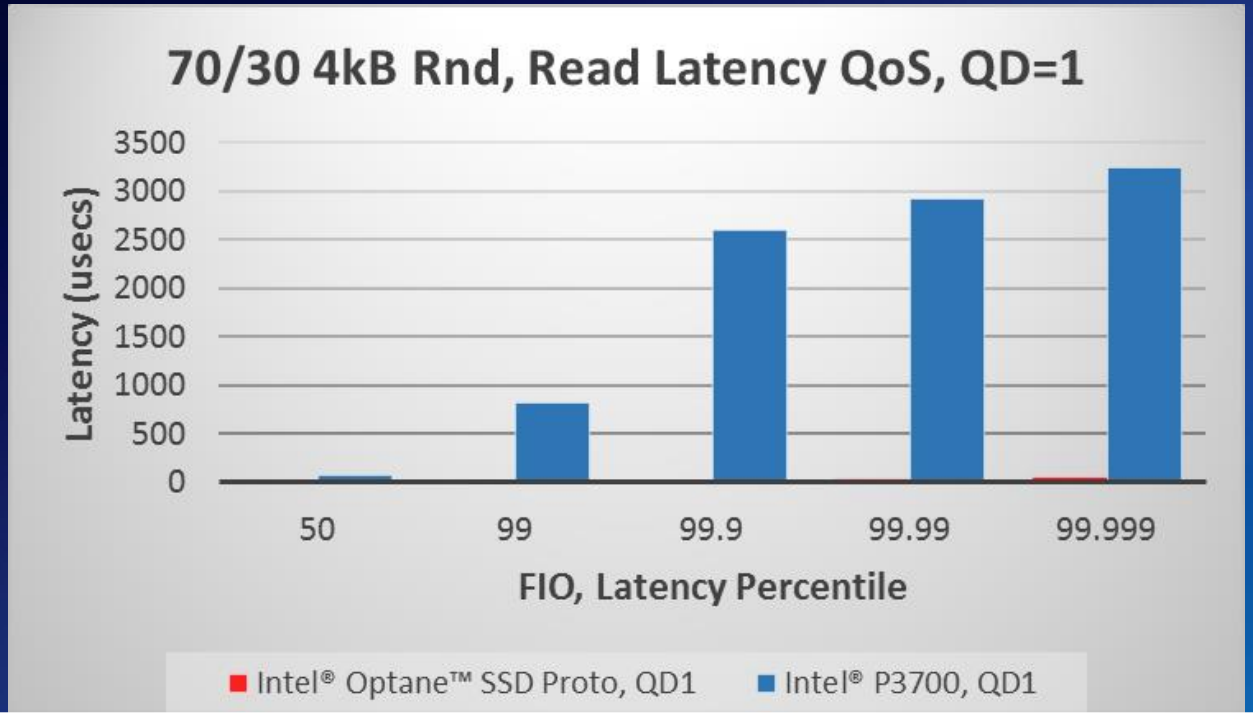
Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

SSD Quality of Service



Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Intel SSD P3700 Series 800GB, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

SSD Quality of Service



Intel® Optane™ SSDs change the game, operating at previously unreachable response latency consistency

Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus 4x4GB DDR-2133, Hyperthreading disabled CPU C-state disabled, Intel P3700 SSD 800GB, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

Agenda

- *Wicked Fast Storage*: Intel® Optane™ SSDs Based on 3D XPoint™ Technology
- **Client/Workstation Impact** (and why)
- Data Center Impact (and why)
- *And Beyond*: Intel® DIMMs Based on 3D XPoint™ Technology

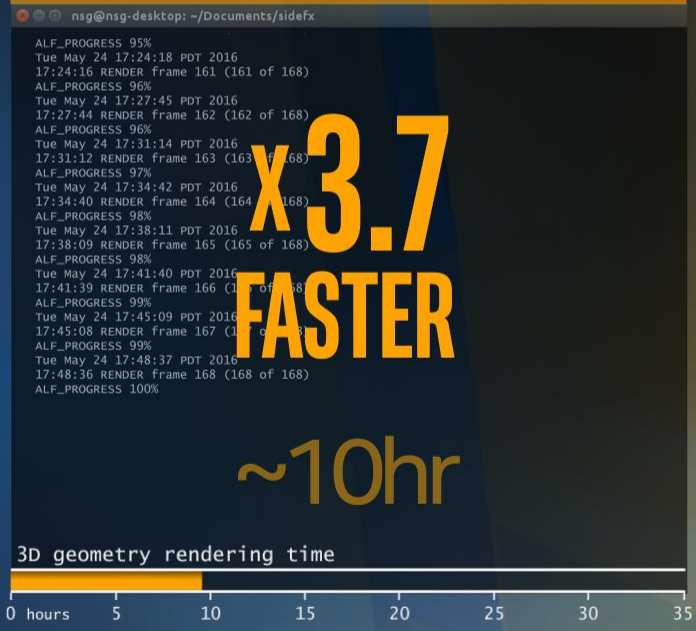
SideFX* Houdini* Use Case


[FLUID DYNAMICS: 7sec; 168 frames; 1.1B particles]

BREAKTHROUGH TECHNOLOGY

INTEL® 750 SERIES SSD

INTEL® OPTANE™ SSD

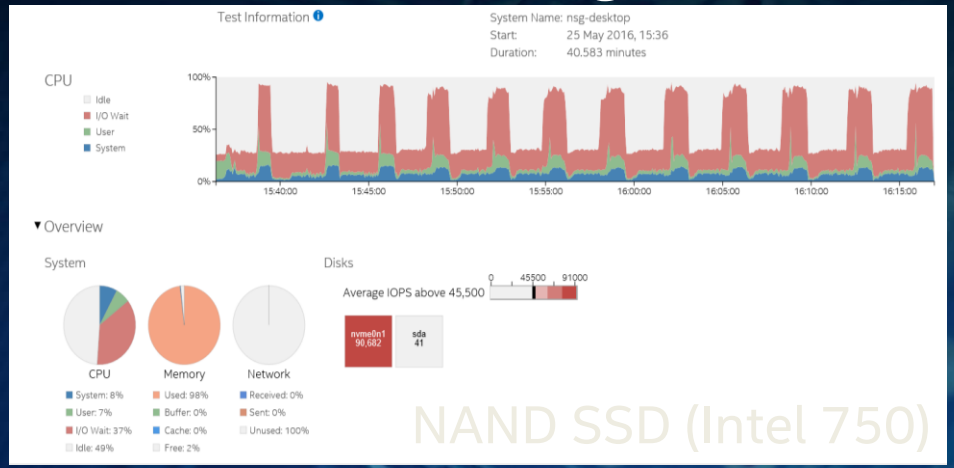


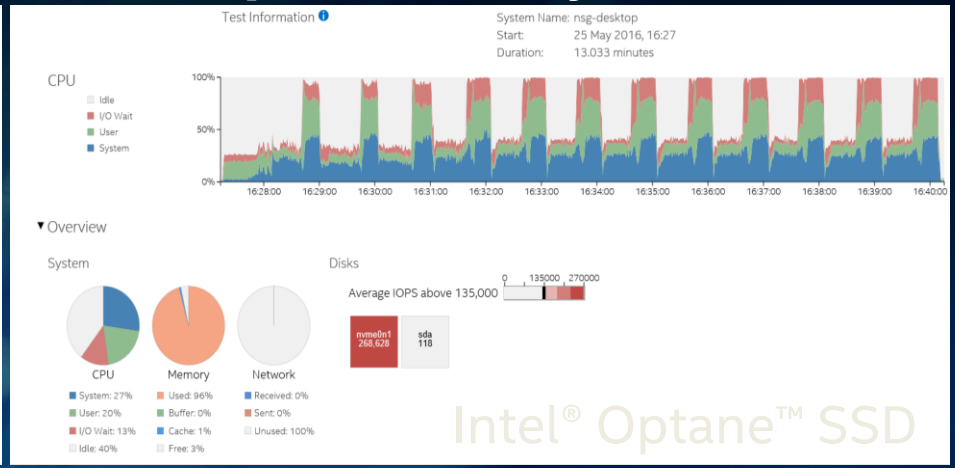
Config: Single socket Broadwell E CPU on the X99 platform. 64GB of DDR-4, single SATA OS drive with secondary 750 and Intel® Optane™ SSD prototype configured as scratch partitions.

Houdini* 3D Rendering

IDF16 INTEL DEVELOPER FORUM

Intel® Storage Performance Snapshot Analysis





- 40.6 minutes to render
- 70% of CPU active waiting on storage

- 13 minutes to render
- 20% of CPU active waiting on storage
- More time highly parallel

intel experience what's inside™

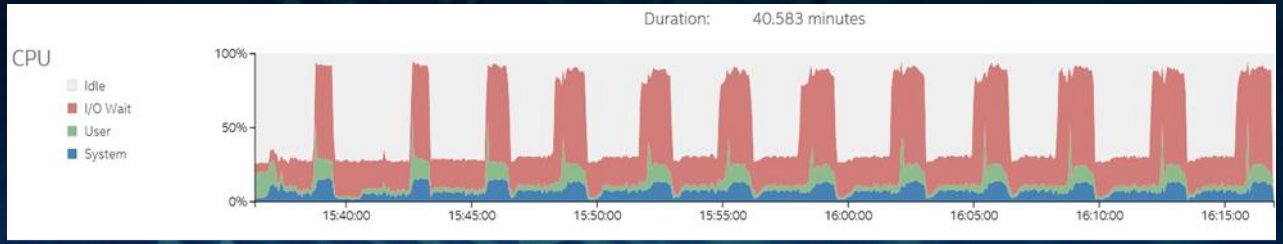
Render only 4 frames to enable shorter run
Enable only 4 cores (not HT) to get clearer trace stats

Config: Single socket Broadwell E CPU on the X99 platform. 64GB of DDR-4, single SATA OS drive with secondary 750 and Intel® Optane™ Prototype SSD configured as scratch partitions.

Houdini* 3D Rendering

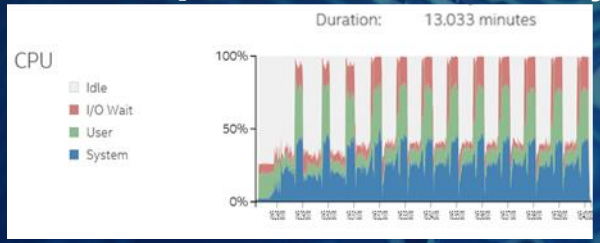
IDF16 INTEL DEVELOPER FORUM

Intel® 750 SSD



- 40.6 minutes to render
- 70% of CPU active waiting on storage

Intel® Optane™ SSD Prototype



- 13 minutes to render
- 20% of CPU active waiting on storage
- More time highly parallel

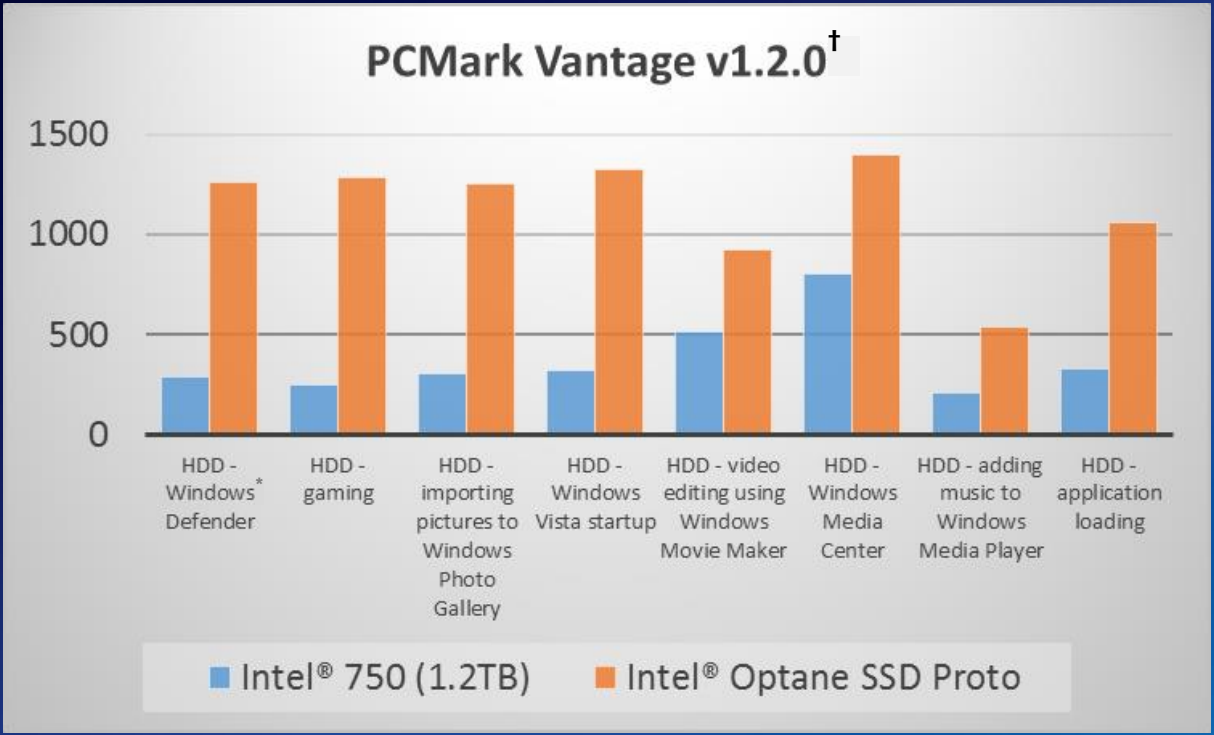
Lower Latency results in >3x Speedup

intel experience what's inside™

Render only 4 frames to enable shorter run
Enable only 4 cores (not HT) to get clearer trace stats

Config: Single socket Broadwell E CPU on the X99 platform. 64GB of DDR-4, single SATA OS drive with secondary 750 and Intel® Optane™ Prototype SSD configured as scratch partitions.

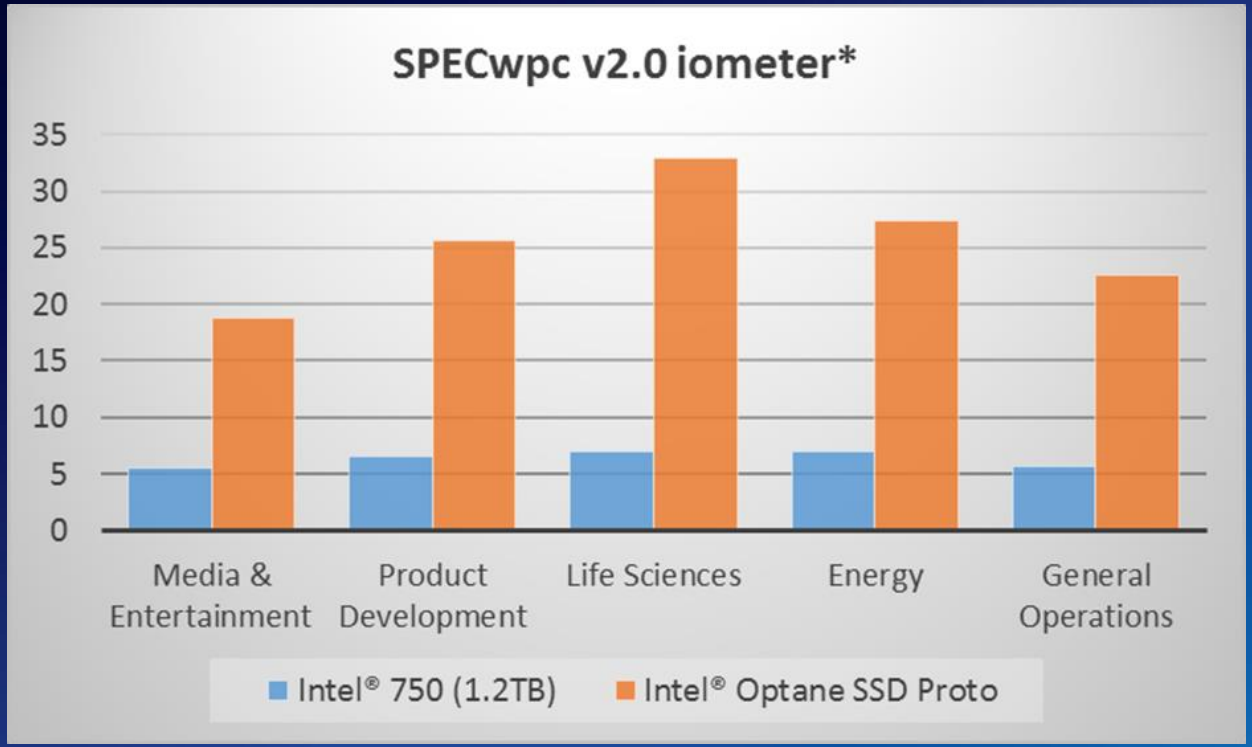
Client/Workstation Benchmarks



3x Client Application Storage Performance Advantage

†See Appendix for configs and trademarks (slide 47)

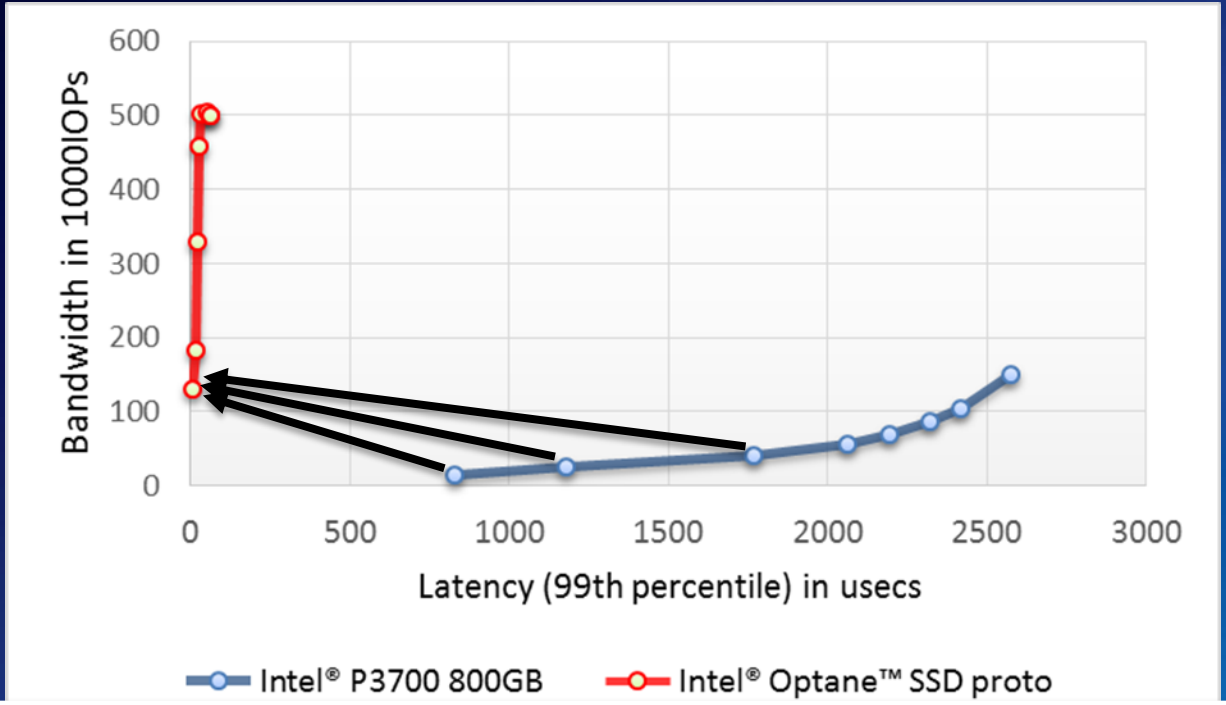
Client/Workstation Benchmarks



4x Workstation Application Storage Performance Advantage

*See Appendix for configs and trademarks (slide 47)

Speeding Client/Workstation Applications



Intel® Optane™ SSDs enable a higher performance operating point for client/workstation applications

PMBench 0.71 2GB file w/random access, 1 thread, Ubuntu* server 4.4.0-31, third-party platform, 1.5GHz, 32GB DDR4 – 200 limited to 1GB with GRUB, Neon Coty MB

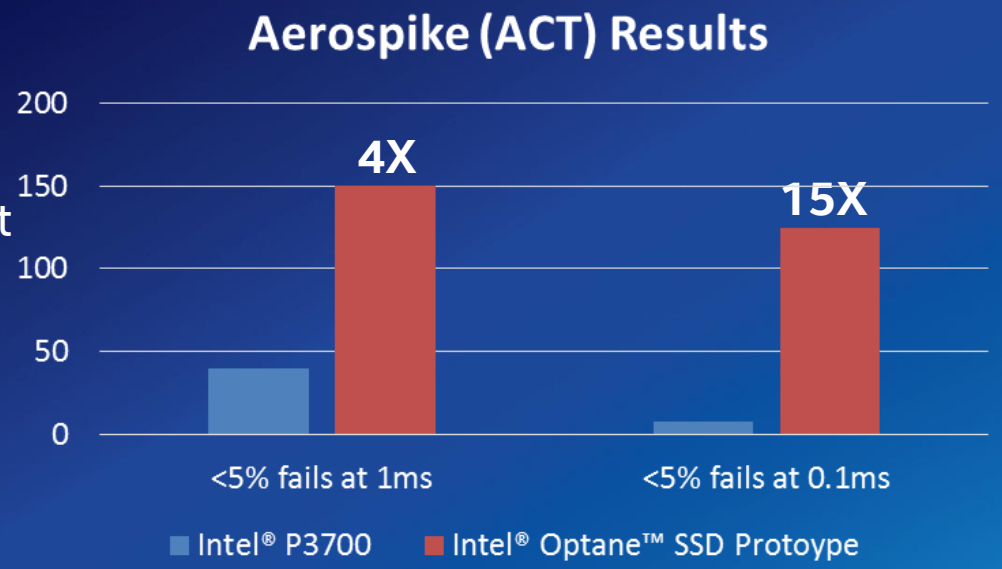
Agenda

- *Wicked Fast Storage*: Intel® Optane™ SSDs Based on 3D XPoint™ Technology
- Client/Workstation Impact (and why)
- **Data Center Impact (and why)**
- *And Beyond*: Intel® DIMMs Based on 3D XPoint Technology

Data Center: ACT Benchmark

IDF16 INTEL DEVELOPER FORUM

- **Aerospike* Certification Tool emulates the I/O pattern of a real-time database:**
  - 1.5kB random reads that meet Service Level Agreement
  - 128kB background writes
  - Measure multiplier while maintaining SLA

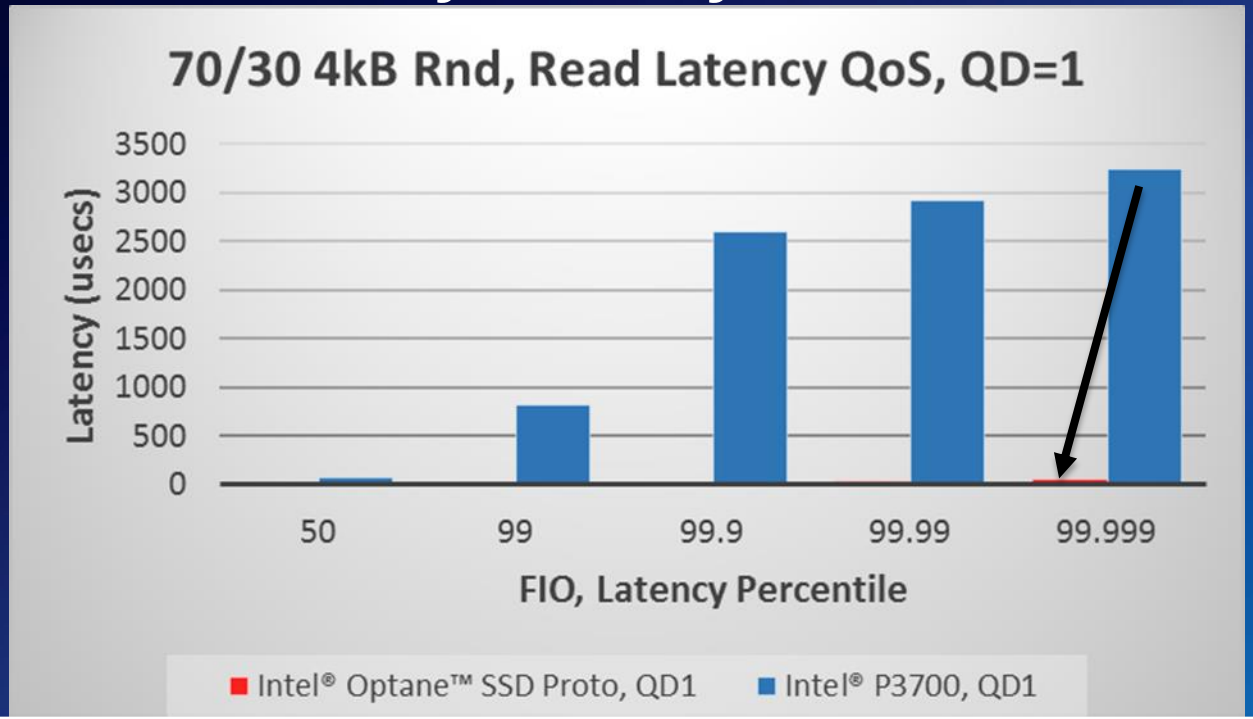


Better SSD QoS results in higher real-time database throughput

intel experience what's inside™

CentOS* Linux* release 7.1.1503, Intel® Core™ i7 4770, ASUSTeK COMPUTER INC., H87I-PLUS, Boot Drive: Intel SSD DC S3500 160GB SATA SSD, 4GB DDR3 Single Channel

Superior IO Latency Quality of Service



Intel® Optane™ SSDs operate at a higher performance level for a given Service Level Agreement

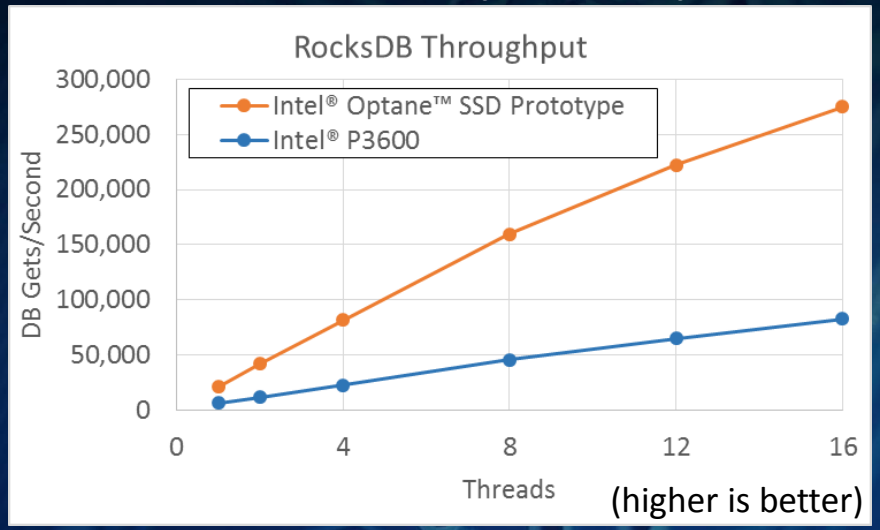
[30] Config: I7-6700K Turbo to 4.3GHz, ASUS* Z170m-plus, 4x4GB DDR-2133, Hyperthreading disabled, CPU C-state disabled, Ubuntu* 14.04 LTS 64 bit server, kernel 4.4 (polling enabled), FIO 2.1.11

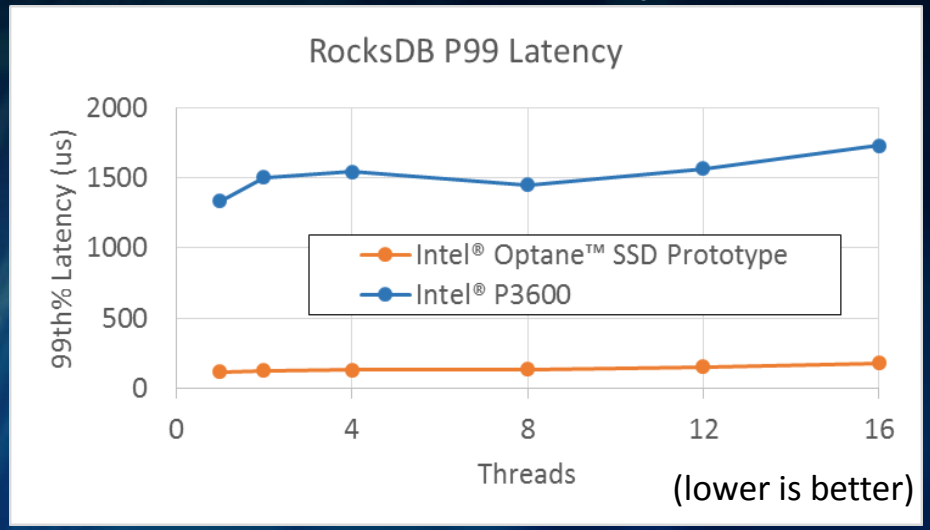
Intel® Optane™ SSD Prototype RocksDB Demonstration

Data Center: RocksDB Perf on Test5 (from rocksdb.org)

IDF16 INTEL DEVELOPER FORUM

- Open source persistent key-value store
- All threads randomly reads keys, one writer thread updates up to ~80K keys/second





~3x Throughput advantage

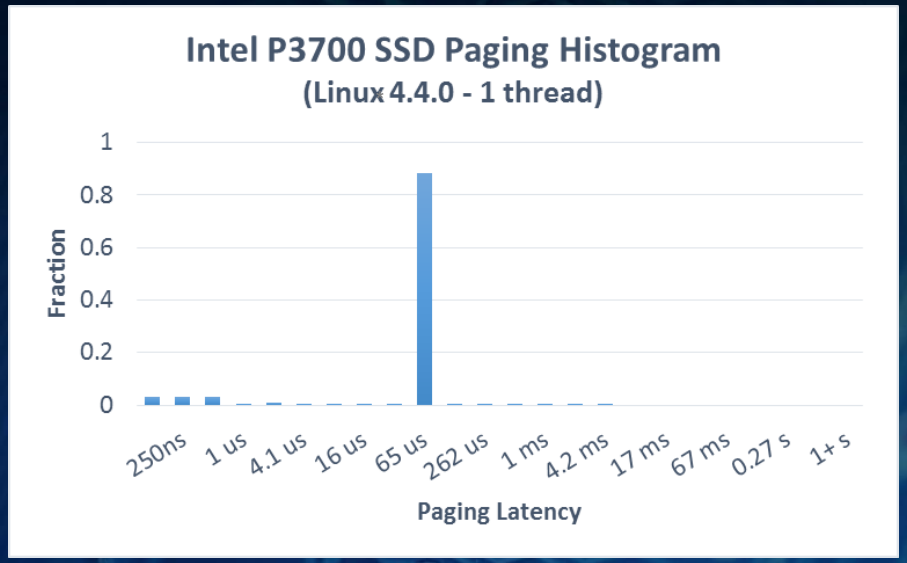
~10x Latency advantage (99th percentile)

Increased persistent key-value store throughput with better QoS

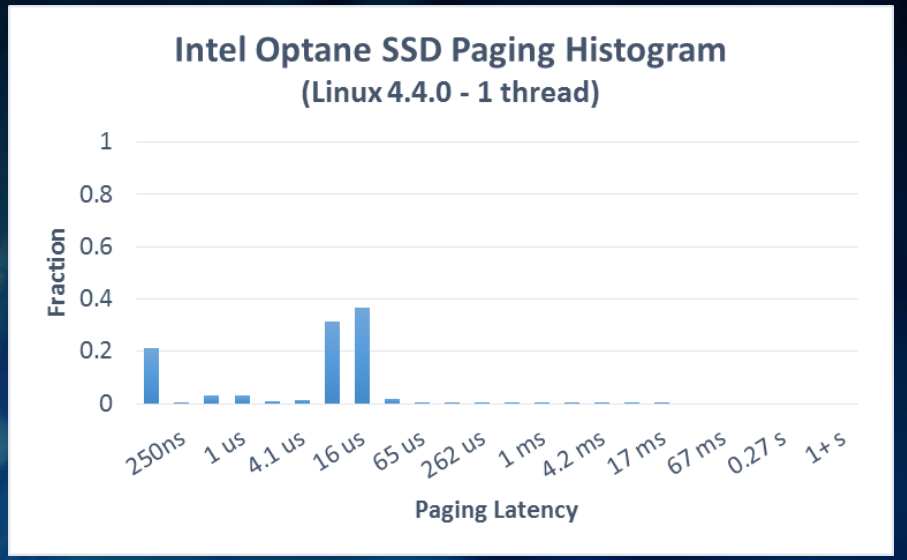
intel experience what's inside™

RocksdB setup based on published tests at rocksdb.org: 1B Key Database used, 8 "Shards" of 25M Key/Values each, 20 byte keys, 800 byte values, 50% compression, ~100 GB on-disk. Read: All threads randomly read all keys. Read/Write: All threads randomly reads keys 1 writer thread updates up to ~80K keys/second. Quanta Leopard base board, 2x Intel CPUs (2.5 GHz, 12 core, HT Enabled, 8 DDR4 DIMMs, 256GB, 32GB Used, CentOS* 7.2, no OS changes XFS FS with FB build/mount opts, TRIM enabled, P3700 (50% capacity used) and Intel Optane Based Prototype (75% capacity used).

Paging Performance



Average paging time = 88 usecs



Average paging time = 15 usecs

Paging to extend system memory now a viable strategy

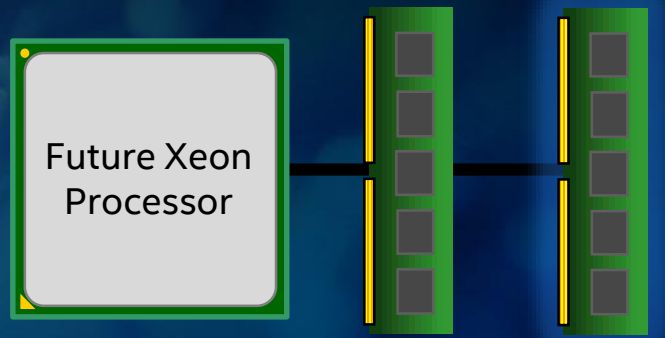
PMBench 0.71 2GB file w/random access, 1 thread, Ubuntu* server 4.4.0-31,
third-party platform, 1.5GHz, 32GB DDR4 – 200 limited to 1GB with GRUB, Neon Coty MB

Agenda

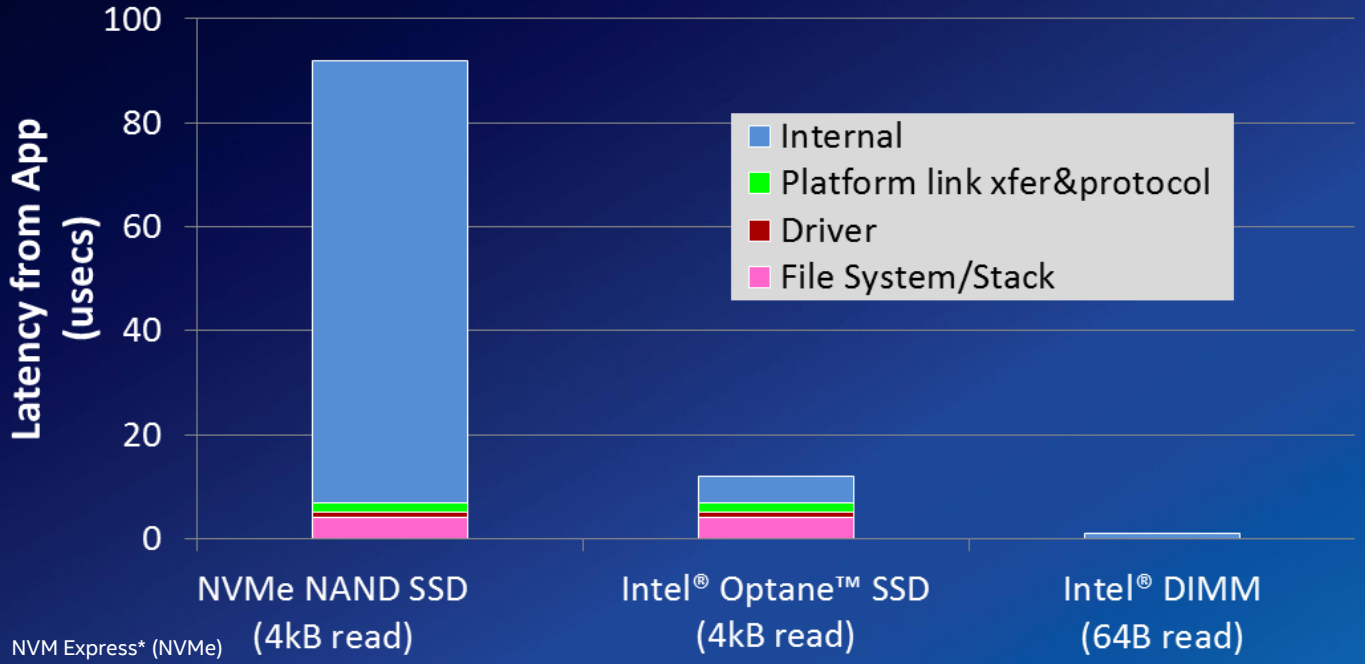
- *Wicked Fast Storage*: Intel® Optane™ SSDs Based on 3D XPoint™ Technology
- Client/Workstation Impact (and why)
- Data Center Impact (and why)
- *And Beyond*: Intel® DIMMs Based on 3D XPoint Technology

Intel® DIMMs Based on 3D XPoint™ Technology

- DDR4 electrical & physical compatible
- Supported on next generation Intel® Xeon® platform
- Up to 2X system memory capacity, at significantly lower cost than DRAM
- Can deliver big memory benefits without modifications to OS or applications



Optimized System Interconnect

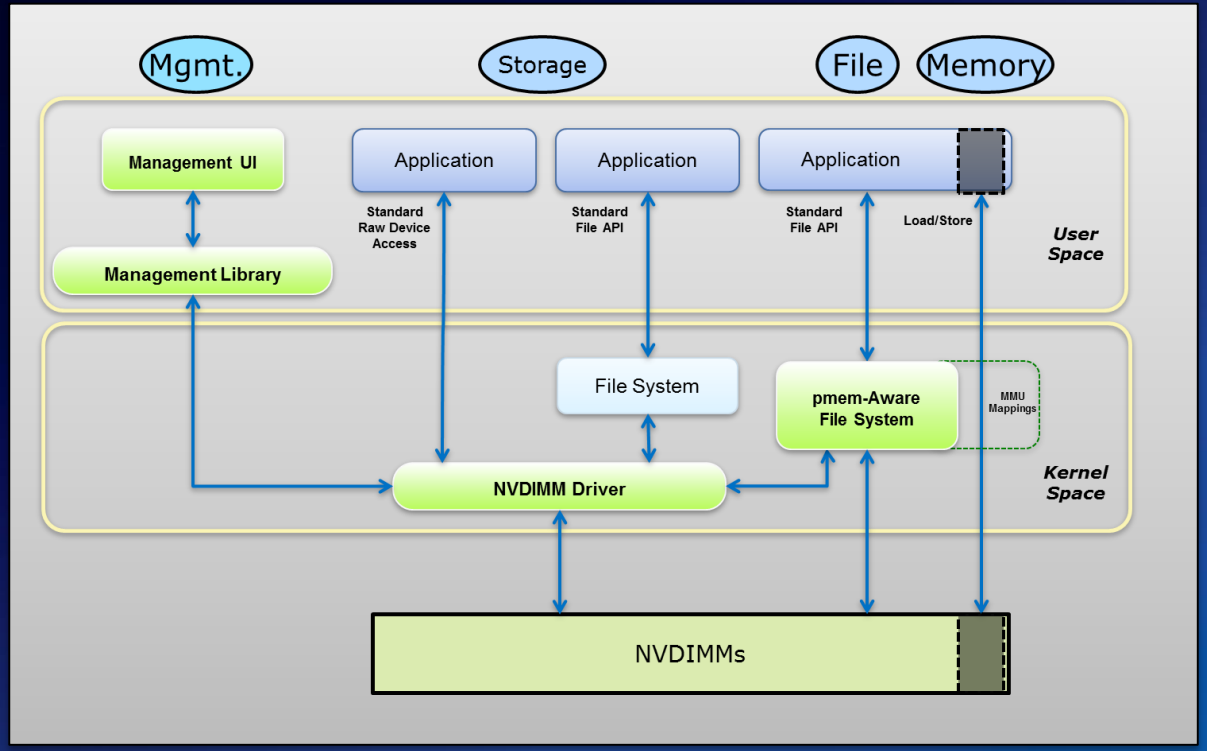


NVM Express* (NVMe)

Reach full potential of 3D XPoint™ Technology by connecting it as Memory

Sources: "Storage as Fast as the rest of the system" 2016 IEEE 8th International Memory Workshop and measurement, Intel® Optane™ SSD measurements and Intel P3700 measurements, and technology projections

Operating System Support



Linux*

- Direct Access (DAX) changes upstream
- ext4 is a pmem-aware file system (others may follow)
- Multi-Tiered Memory Management beginning development

Windows*

- Direct Access (DAX) changes announced publicly
  - Includes “SCM” stack
- NTFS is a pmem-aware file system (others may follow)

Application Level Advantages: Example
(Measurements use Persistent Memory (PM) Emulator)

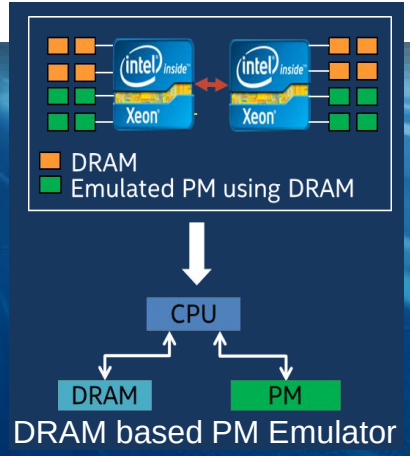
IDF16 INTEL DEVELOPER FORUM

Instant Recovery for Main-Memory Databases

Ismail Oukid*°, Wolfgang Lehner*, Thomas Kissinger*, Peter Bumbulis°, and Thomas Willhalm +

*TU Dresden °SAP SE + Intel GmbH

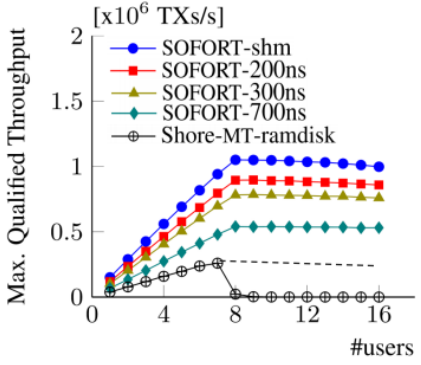
CIDR 2015, Asilomar, California, USA, January 5, 2015



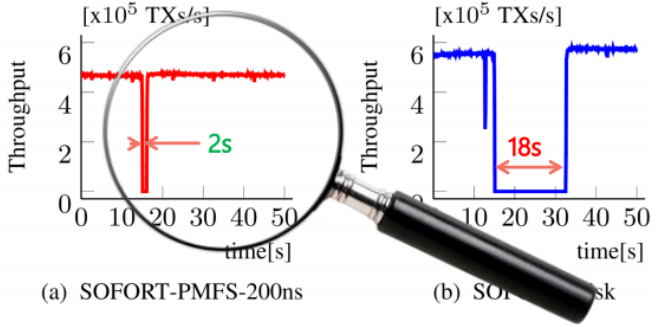
Performance Overview

Dresden Database Systems Group

THROUGHPUT



RESTART TIME



Competitive performance even in high latency environment

Fast restart time. No need to fetch data stored in SCM

Still not instant

TECHNISCHE UNIVERSITÄT DRESDEN

11

http://www.cidrdb.org/cidr2015/Slides/13_CIDR15_Slides_Paper13.pdf

intel experience what's inside™

Your Chance To Innovate early!

With the Intel® Optane™ Technology testbed

IDF16
INTEL DEVELOPER FORUM

Got an amazing app for Intel® Optane™ SSDs? Get ready!

Intel is working to provide access to the developer community:

- **Via a public cloud offering***
- **Free of charge**, before general availability**
- **Benchmarks performance hungry apps, build tomorrow's apps**

SSDTestBed@INTEL.COM

Sign up at: SSDTestBed@intel.com or scan the QR code

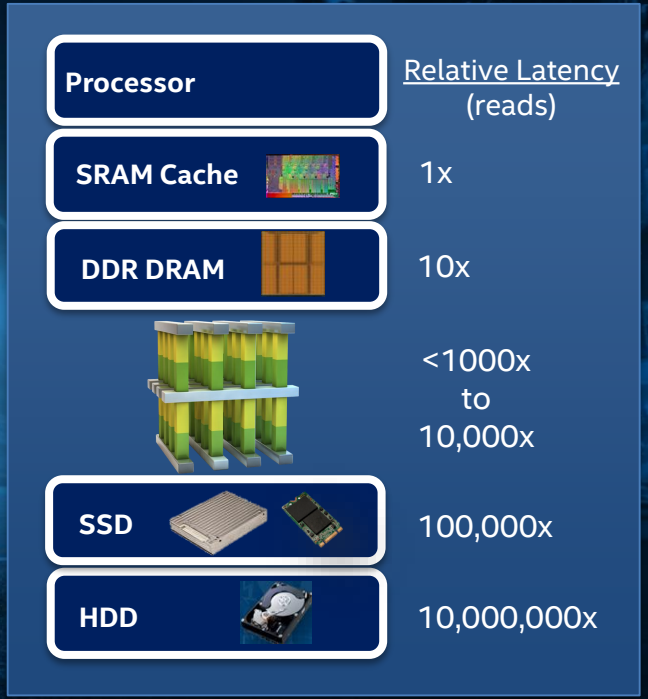
Looking for innovators/developers building innovative apps: blockchain, quantum compute, autonomous navigation, high frequency trading, real time analytics/databases, cloud gaming & more!

*More details in 30-45 days
**Limited time access

intel
experience what's inside™

Summary and Next Steps

- Historically low latency 3D XPoint™ Memory completes the hierarchy
- Intel® Optane™ SSDs deliver superior performance from client and data center.
  Get Ready:
  - *For a new operating point*
  - *To optimize applications*
  - *To use storage in new ways*
- Intel® DIMMs will deliver the full potential to the system



[1]Technology claims are based on comparisons of latency, density and write cycling metrics amongst memory technologies recorded on published specifications of in-market memory products against internal Intel specifications.

Additional Sources of Information

- A PDF of this presentation is available from our Technical Session Catalog: www.intel.com/idfsessionsSF. This URL is also printed on the top of Session Agenda Pages in the Pocket Guide.
- Please visit our Memory & Storage demonstrations in the showcase, booth numbers: 360, 361, 362, 364, 365, 366, 559, 561, 568, 586, 590, 593, 882, 882.
- Additional info in the Memory and NVM Express* Community
- More web based info: www.intel.com

IDF16
INTEL DEVELOPER FORUM

Legal Notices and Disclaimers

- Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at intel.com.
- Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.
- Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more complete information visit http://www.intel.com/performance.
- Cost reduction scenarios described are intended as examples of how a given Intel-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.
- This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.
- No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.
- Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.
- All products, computer systems, dates and figures specified are preliminary based on current expectations, and are subject to change without notice. The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.
- Intel does not control or audit third-party benchmark data or the web sites referenced in this document. You should visit the referenced web site and confirm whether referenced data are accurate.
-
- *Other names and brands may be claimed as the property of others.

IDF16
INTEL DEVELOPER FORUM

Appendix: Legal Notices and Disclaimers

For PCMark* and SPECwpc

System Configuration

Intel® Core™ i7-6700K Processor 4 GHz base up to 4.2 GHz, 8T/4C, 8 MB cache, 91 W TDP On **Motherboard:** Asus* Maximus VIII Hero Alpha Z170, **Graphics:** Intel® HD Graphics Driver 20.19.15.4380, **Memory:** 2 X 8GB G.Skill Ripjaws F4-2133C15Q-16GRB PC4-17000 (2133MHz), **OS Hard Disk:** Barracuda ST3100524AS 1TB 1000000, **Operating System:** Windows* 10 Pro (x64) TH2 **Intel® Rapid Storage Technology:** 14.6.0.1029

Workload Details

Drive under test is secondary drive and preconditioned to steady state.
0% Utilization defined as a formatted drive with no data added.
75% Utilization achieved by using a file generator to create a single, compressible file of size equal to 75% of user capacity and written to the drive under test.

PCMark Vantage

PCMark Vantage is a benchmark from Futuremark* that measures Windows everyday computing performance. PCMark Vantage is made up of several benchmarking suites: PCMark Suite (produces "PCMark" Score), Memories Suite, TV and Movies Suite, Gaming Suite, Music Suite, Communications Suite, Productivity Suite and HDD Suite. The HDD Suite contains an operating system start-up workload that is sensitive to HDD versus SSD boot devices. Operating Systems: Desktop Windows*.

SPECwpc* V2.0

SPECwpc* V2.0 is a benchmark from the SPEC* consortium that measures workstation performance based on a broad range of professional applications. SPECwpc runs professional application usage tests: Media and Entertainment, Financial Services, Product Development, Energy, Life Sciences and General Operations. Operating Systems: Desktop Windows

IDF16
INTEL DEVELOPER FORUM

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804