STORAGE INDUSTRY SUMMIT

Convergence of Storage and Memory Developing the Needed Ecosystem

JANUARY 20, 2016, SAN JOSE, CA

Edward Sharp
PMC-Sierra
Chief Strategy & Technology Officer

NVM Changes the Computing Stack / Who Cares about NVM?

SNIA SSSI | SOLID STATE STORAGE

The Opportunity ahead

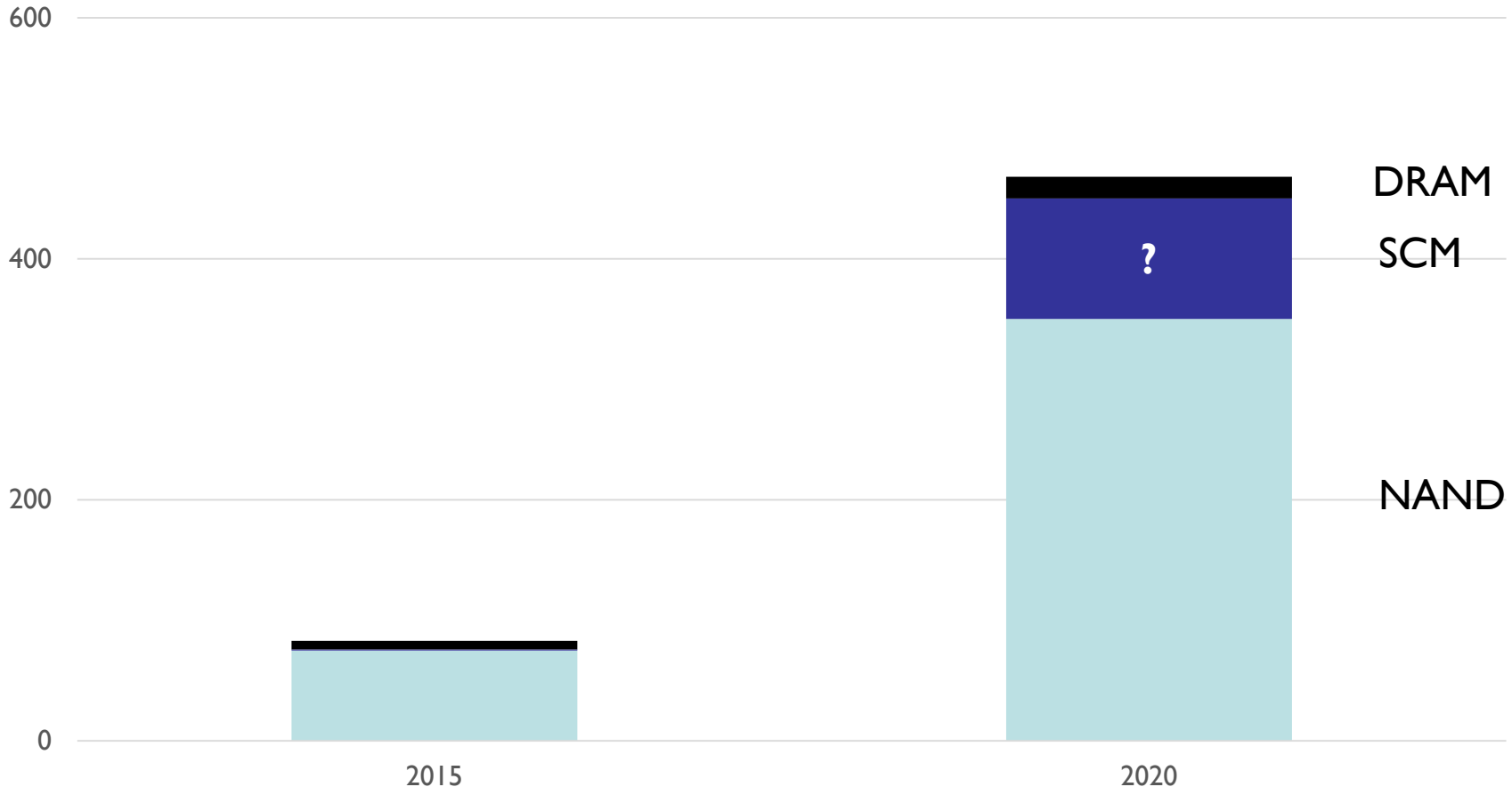
SNIA™
Solid State Storage Initiative

Categories of NVM

SNIA™ Solid State Storage Initiative

Category	Characteristics
DRAM "Drop-in"	Fast, but expensive Byte addressable Memory interface SuperCap replacement
Storage Class Memory (SCM)	Cheaper, but slower, than DRAM Byte addressable Block and Memory interfaces
NAND Flash	Cheap and relatively slow Block interface

EB of Solid State Memory / Storage shipped



Source: PMC estimates; Micron & HGST

We are gathered here because we believe in the value of low latency

SNIA™
Solid State Storage Initiative

msecs

<100us

<10us

Overhead of driver, OS and applications becomes more noticeable as we move right!

NVM delivers low latency and high throughput

SNIA™ Solid State Storage Initiative

Category	Latency (usec)	IOPS	Bandwidth (MB/s)
DRAM "Drop-in"	10^{-1}	10^{9}	10^{4}
Storage Class Memory (SCM)	10	10^{7}	10^{4}
NAND Flash	10^{2}	10^{6}	10^{3}
HDDs	10^{4}	10^{2}	10^{2}

To make this happen we've all been working on multiple fronts

SNIA™
Solid State Storage Initiative

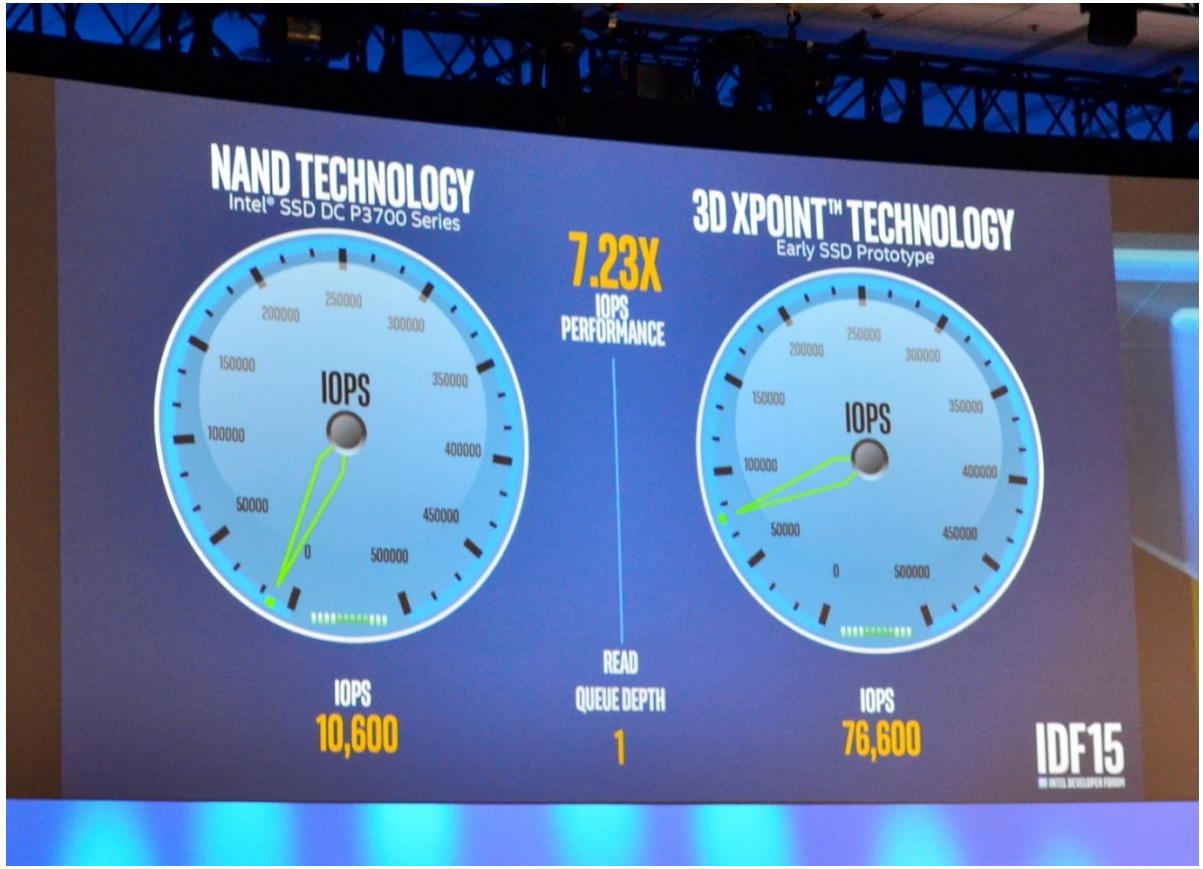
- Media
- Controllers
- Networks
- CPU architectures
- OS

Congratulations – New Media

Product	Vendors	Medium	Comments
NVDIMM	Micron Agiga Netlist	DRAM (with Flash)	DRAM performance and price
UltraDIMM	SanDisk/ Smart/ Diablo	NAND	Poor adoption NAND too slow for good performance
Optane DIMM	Intel	3DX	SkyLake Xeon (2017)

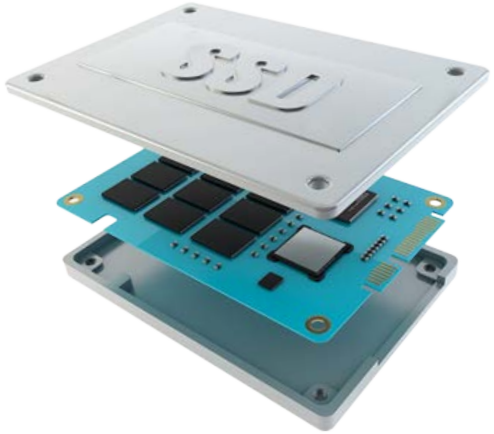
Congratulations – Media

- Intel Optane – NVM Express and DIMM form-factors



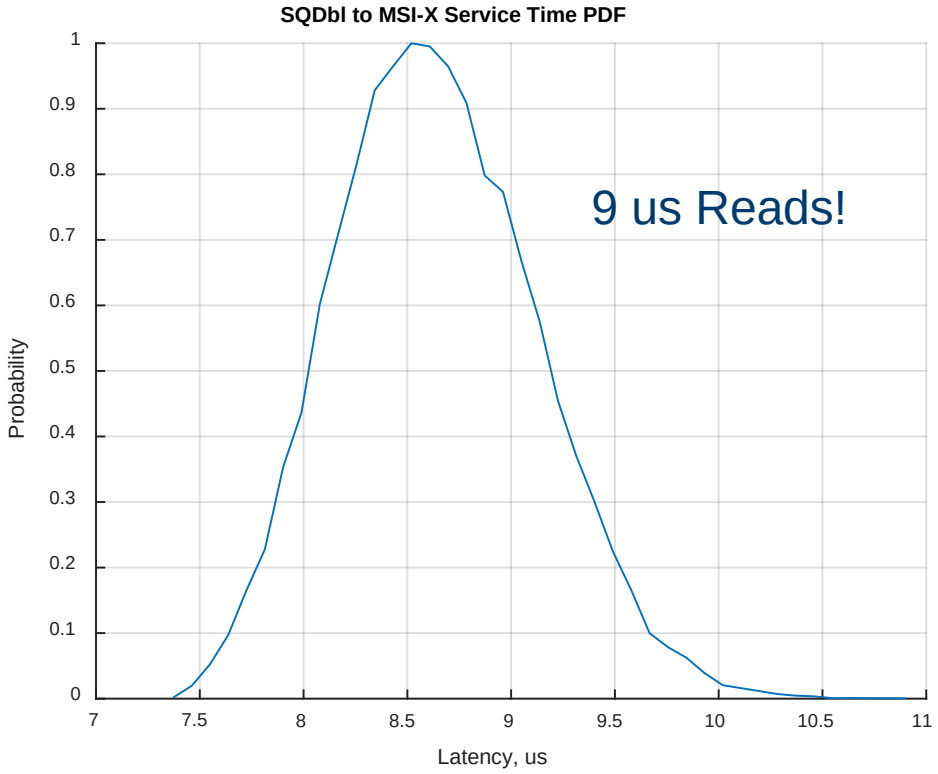
- **Faster (than NAND)**
- **Denser (than DRAM)**
- **Cheaper (than what)??**

Congratulations – New Controllers



- 1M IOPS
- >20TB capacity
- Next generation ECC – LDPC
- SW-Defined-Flash (SDF) interface

SCM can improve latency an order of magnitude over NAND



A proxy for a NG-NVM based SSD!

Low latency and great QoS!

Congratulations – NVMe delivers ~10 times throughput

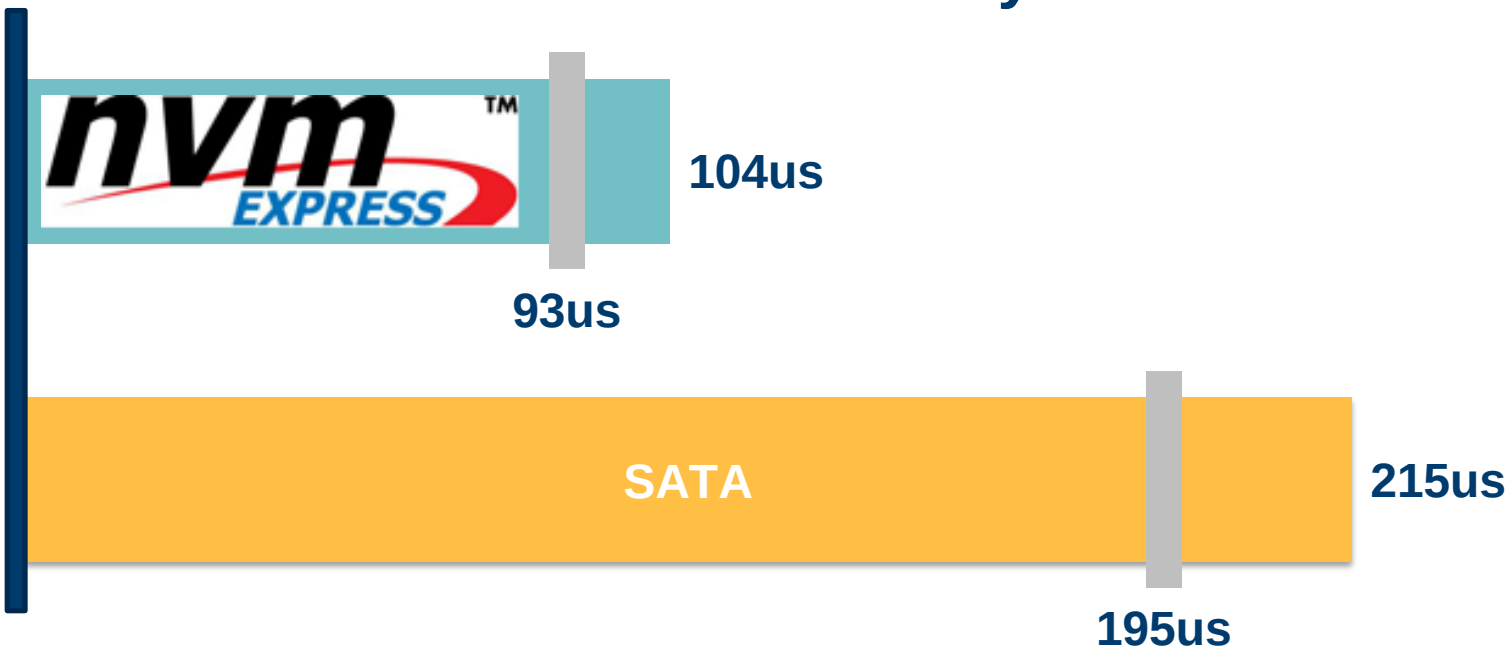
SNIA™ Solid State Storage Initiative

IOPS per Hardware Thread

Congratulations – NVMe Latency half that of SATA

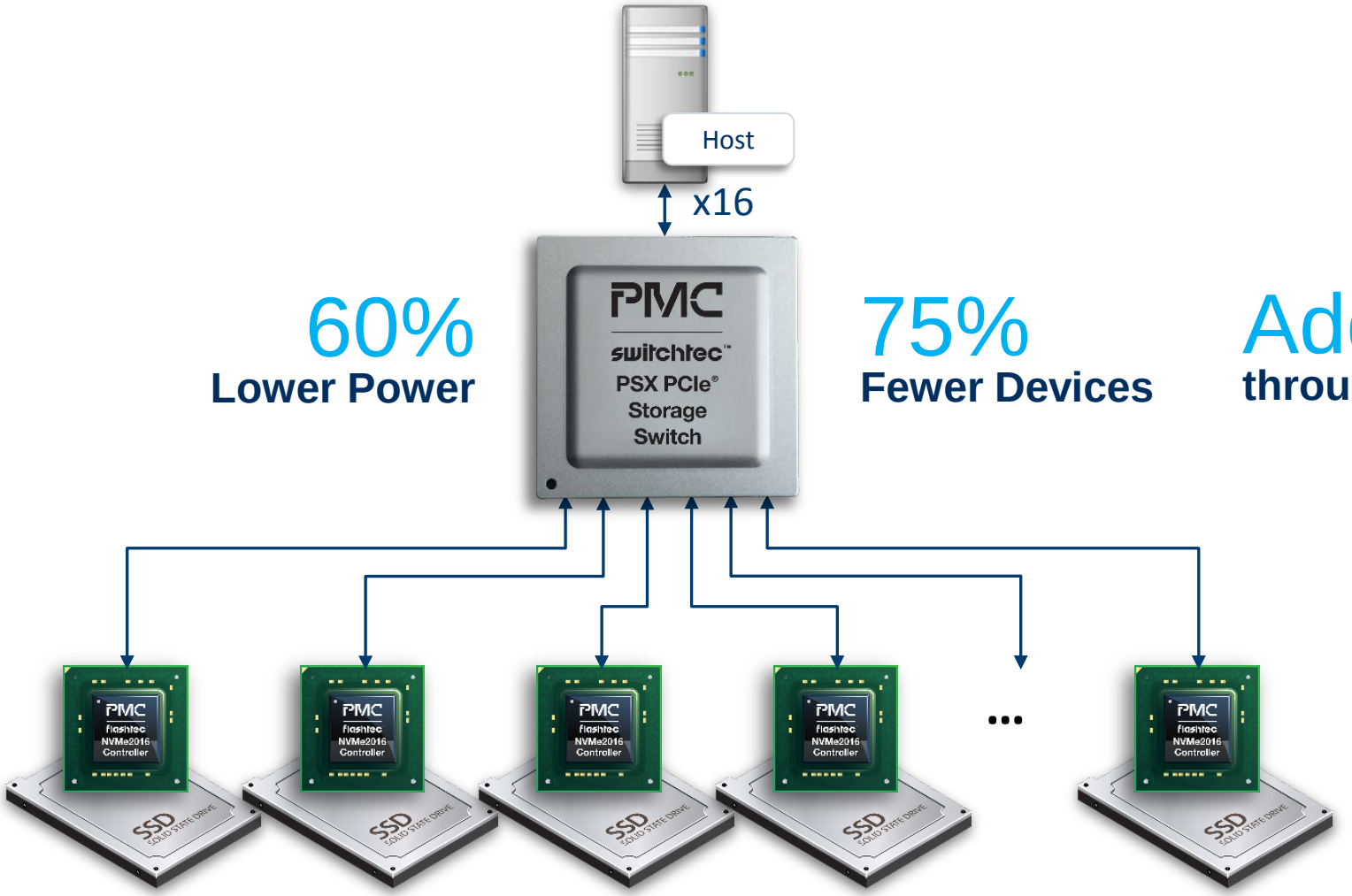
SNIA™ Solid State Storage Initiative

Average and 99.9% Random Read Latency



Congratulations – PCIe Storage Switch

SNIA™ Solid State Storage Initiative

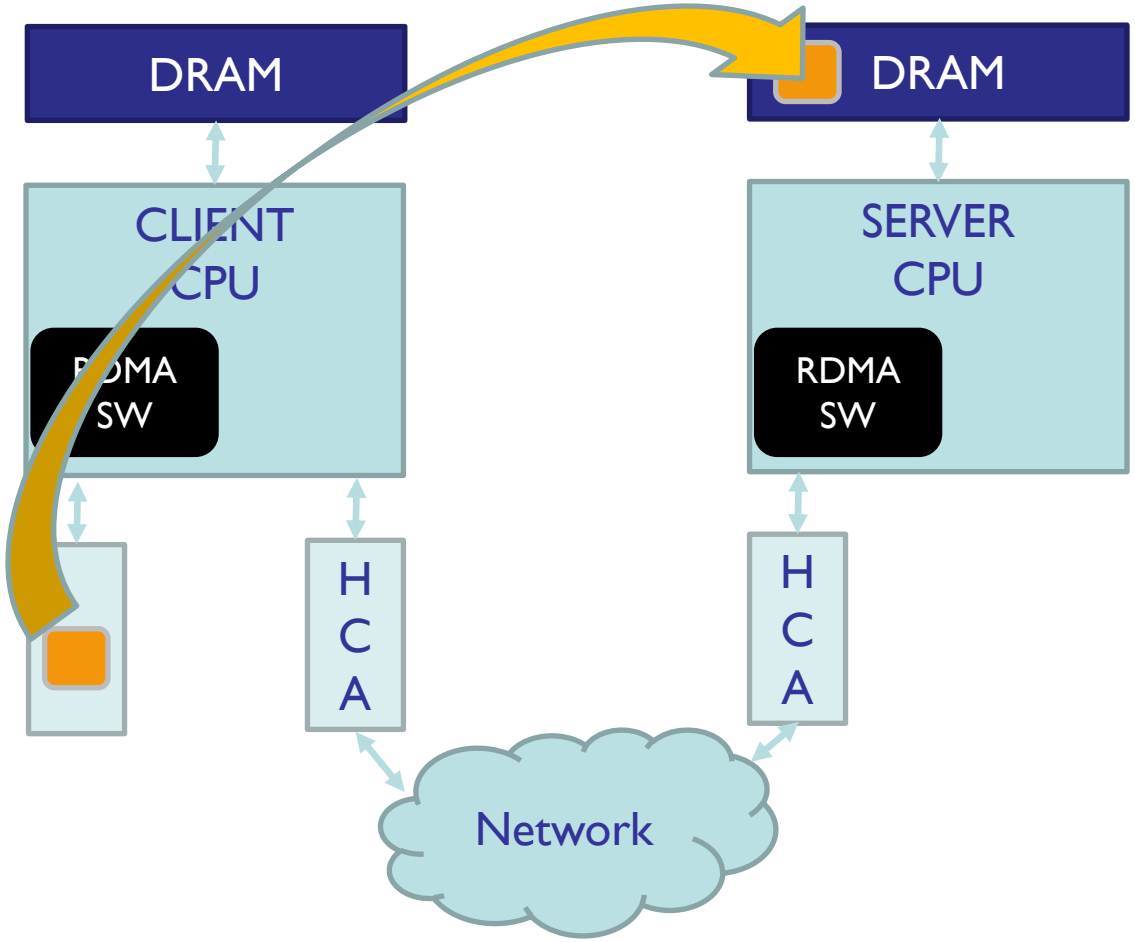


60%
Lower Power

75%
Fewer Devices

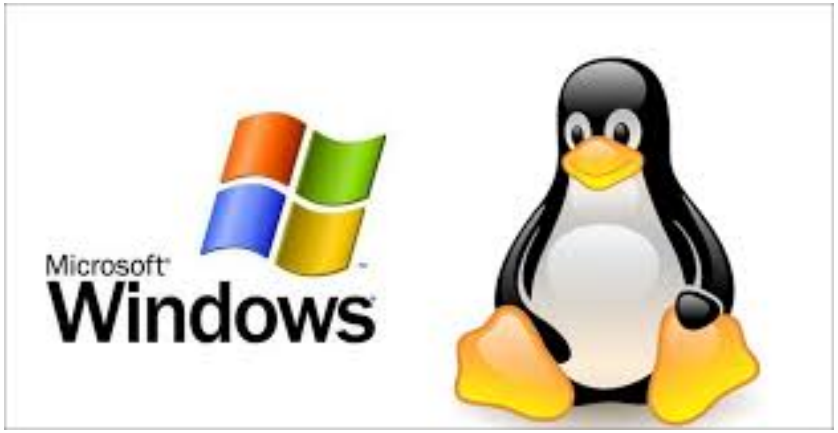
Add Value
through Firmware

Congratulations – RDMA scales latency benefit



- RDMA doesn't have to be DRAM to DRAM
- <<10us kernel bypass
- CPU offloaded

Congratulations – OS



- NVM Express support
- NVMe over Fabrics support
- Persistent MEMory (PMEM) support
- mmap()-able NVM either in NVDIMM or PCIe form-factor
- Heterogeneous Memory Types

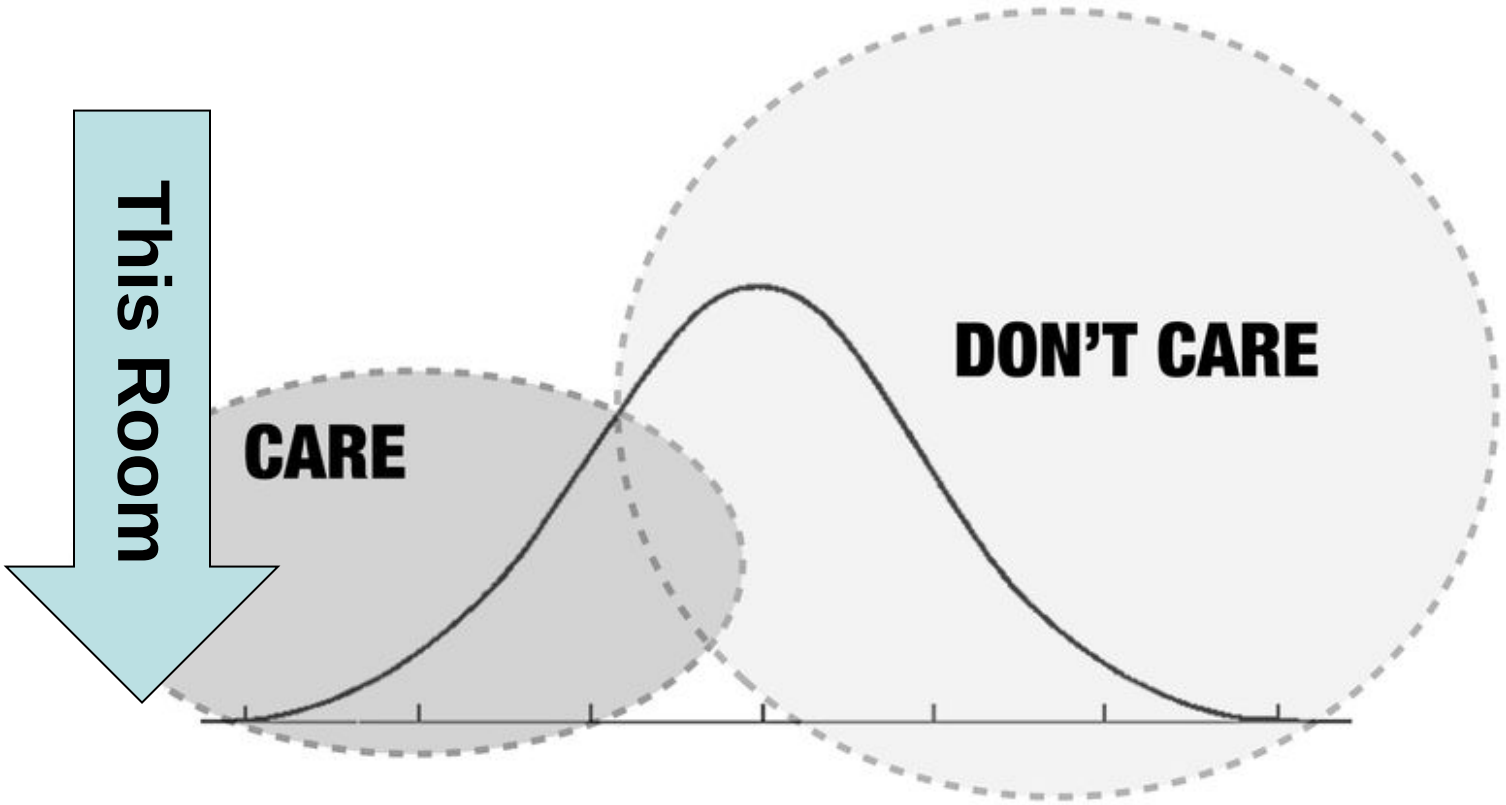
To make this happen we've all been working on multiple fronts

- Media – NAND, PCM, 3DX, ST-MRAM
- Controllers – PCIe, NVM Express; LDPC
- Networks – 25/50/100Gbe Ethernet, RDMA, NVM Express over Fabrics
- CPU architectures – Optane, NUMA
- OS – blk-mq, polling, NVM Express

What Next?
SCM

- Media
  - SCM
- Controllers
  - SCM, Fabric support, Near data processing
- Networks
  - RDMA, NVMe over Fabrics
- CPU architectures
  - Enable Controller Memory Buffers (CMBs) for NVMe
  - Standardize PMEM access over PCIe
  - RDMA into SCM support
- OS
  - NVMe over Fabrics & RDMA into SCM support
  - Make OS aware of heterogeneous memory types
  - Management

But who cares?



Most people don't care

Different priorities, too much trouble, fear, resources

Good enough

But who cares?
The Application

SNIA™
Solid State Storage Initiative

Computing Stack
Application
Compute
Memory
IO
Storage

We've made all the building blocks fast – but do the applications care?

SNIA™ Solid State Storage Initiative

Layer	Customer needs	
Application	Serve business goals	Simple, Efficient, Fast (enough)
Compute	Serve application / workload	Efficient, Fast
Memory	Serve compute	Fast, BW
IO	Connect to data	Latency, BW, Efficient
Storage	Available for application	Efficient, Fast, Scale, Reliable

Applications historically designed for HDD performance

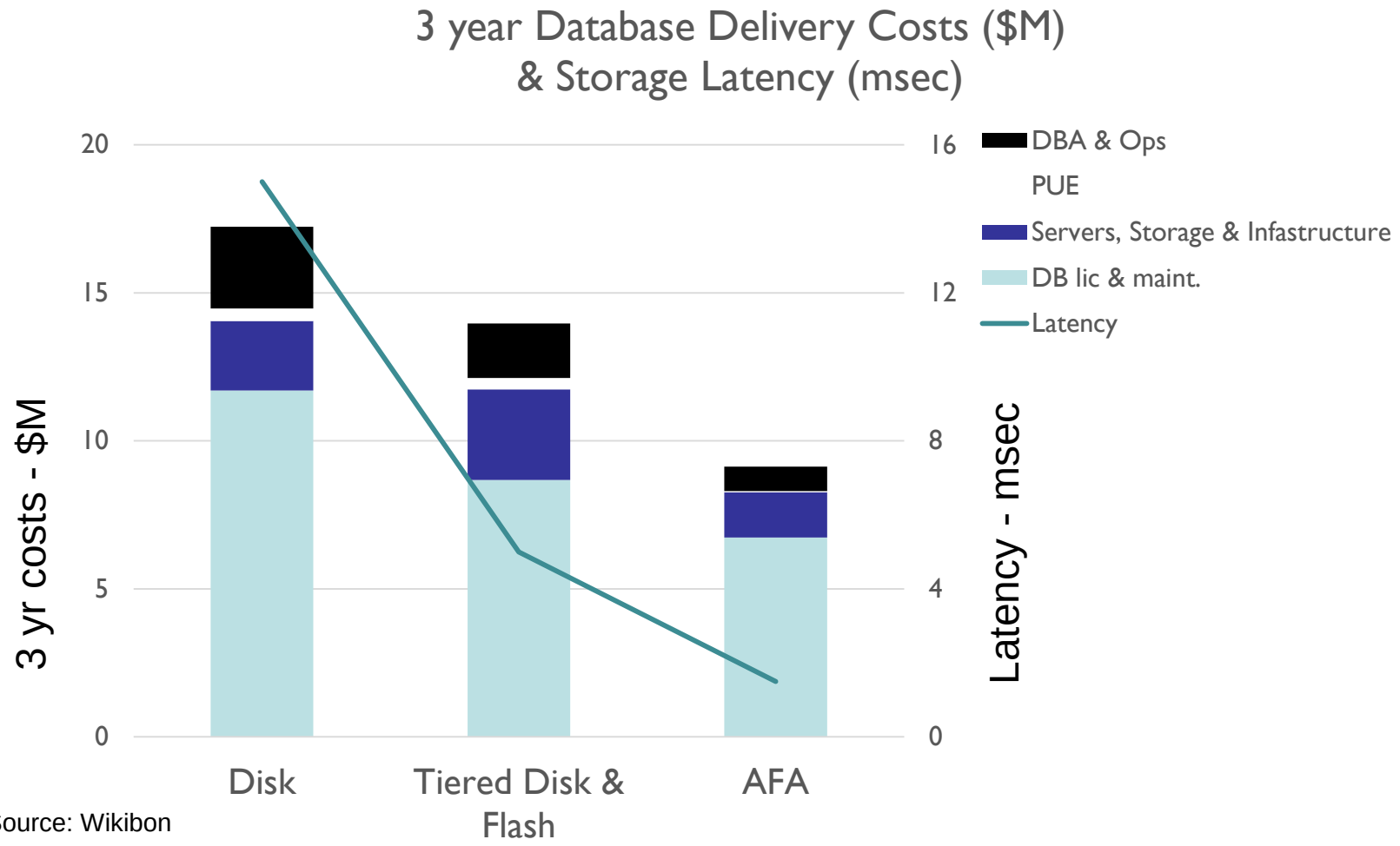
OTLP
Virtualized Apps
Backup & Archive
Media
Big Data / Machine Learning

NVM excelled where no SW changes are needed, removing HDD latency

NVM as Memory	NVM as Cache	NVM as Capacity
Software Changes	***No Software Changes***	***No Software Changes***
Initially workload specific	Huge performance gains from little Flash (10%)	TCO Battleground with HDD
Niche *Mainstream 2020?*	Mainstream *Took about 10Yrs*	Niche *Mainstream 2020?*

Database Example

Savings from fewer licenses – NAND as Disk



Source: Wikibon

Latency drives TCO

What does more performance (than NAND) deliver to legacy apps?

NAND Flash (and HDDs) deliver low cost

SNIA™ Solid State Storage Initiative

Category	Latency (usec)	$/GB
DRAM "Drop-in"	10^{-1}	10^{1}
Storage Class Memory (SCM)	10	10^{0}
NAND Flash	10^{2}	10^{-1}
SMR	10^{4}	10^{-2}

But who cares?
New Applications

SNIA™ Solid State Storage Initiative

ceph

SwiftStack

Red Hat
Canonical
SuSe

openstack

SDS
Broad Market

Windows Server 2012 Storage Spaces

VMware Virtual SAN

IaaS Bare Metal SDS

Hyper-Converged SDS

Big Data Frameworks

HDFS

Flash as Memory
Flash as Cache

cloudera

Hortonworks

MAPR

NoSQL

Flash as Cache

mongoDB

Cassandra

CouchDB relax

IBM Cloudant

In-Memory

Flash as Memory
Flash as Cache

Spark

redis

AEROSPIKE

Transactional

Flash as Cache

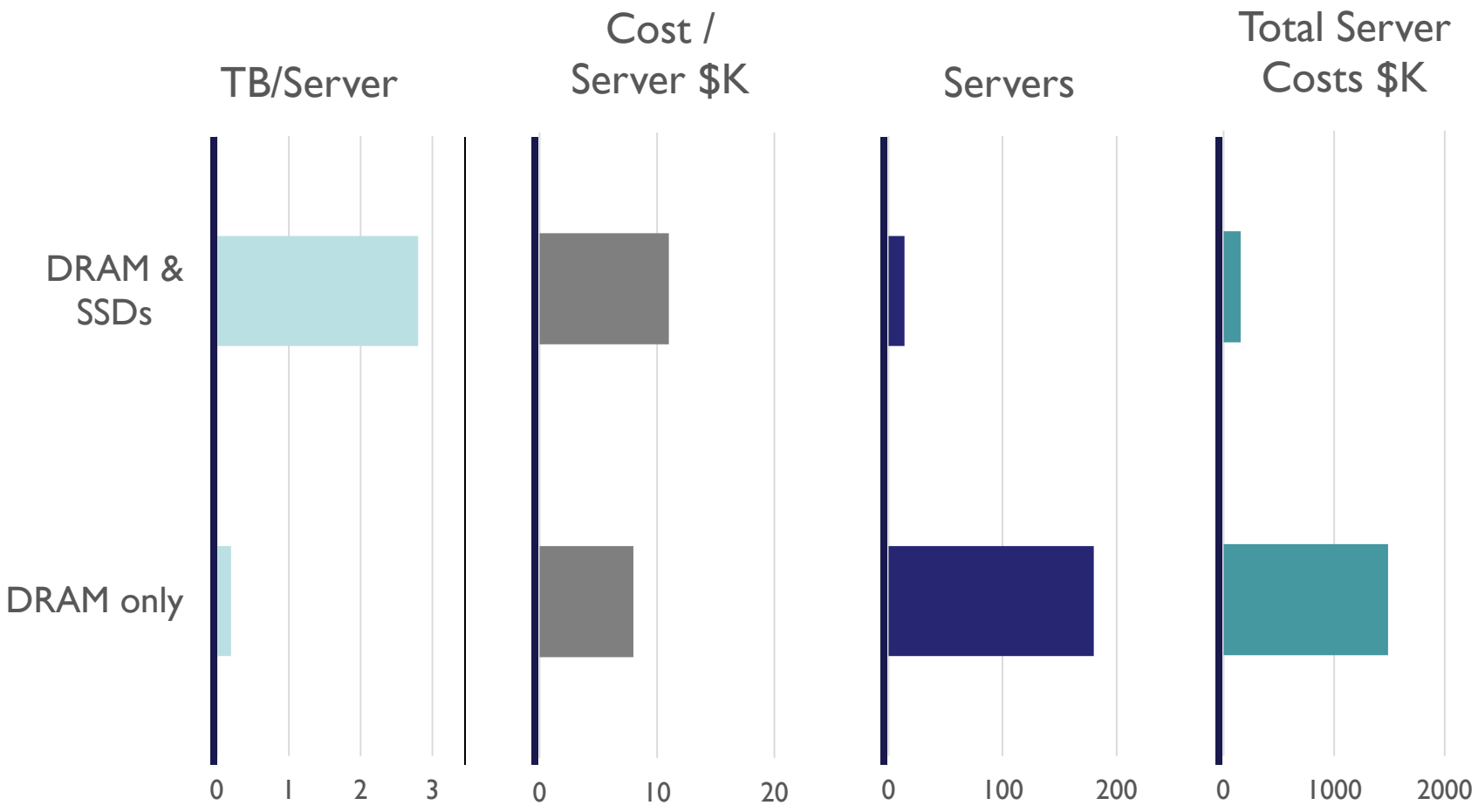
ceph

ScaleIO

nexenta Global Leader in Software-Defined Storage.

Big Data Example

Aerospike NoSQL DB – NAND as Memory



Servers, with storage, needed for 0.5M TPS on Aerospike Database with SSDs vs. NoSQL with DRAM only

Get Real

Real-Time creates an arms race for low and deterministic latency

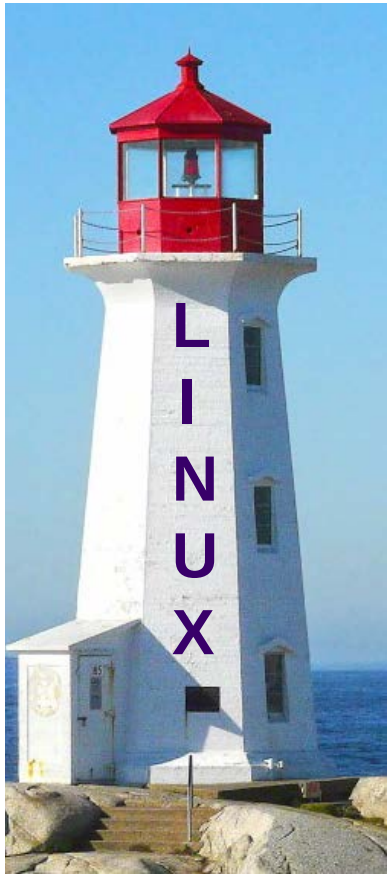
Engage with SW Ecosystem

Call to Action

Engage with application ecosystem

- Promote performance apps: Real-time
- Solve **Price-**performance challenge with NVM
- Help applications take advantage of performance
- SW defined NVM: Flexibility
- Learn what is coming

Continue working the SW



Linux Shows us the Way

- OSes not accustomed to persistence on memory channel
- Changes to OS tell us what is coming

Applications are key

Computing Stack
Application
Compute
Memory
IO
Storage