Flash Memory SUMMIT

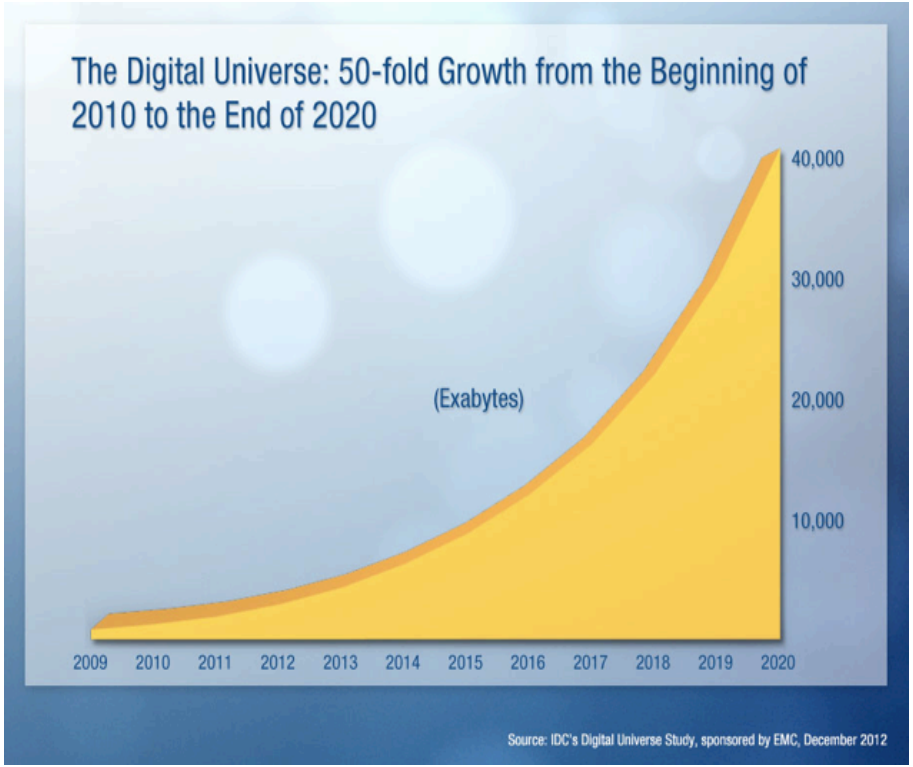
Higher Performance - Lower Cost

How higher storage performance can enable drastically lower data center TCO

Cameron Crandall – Kingston Technology

Sumit Puri – Liqid Inc.

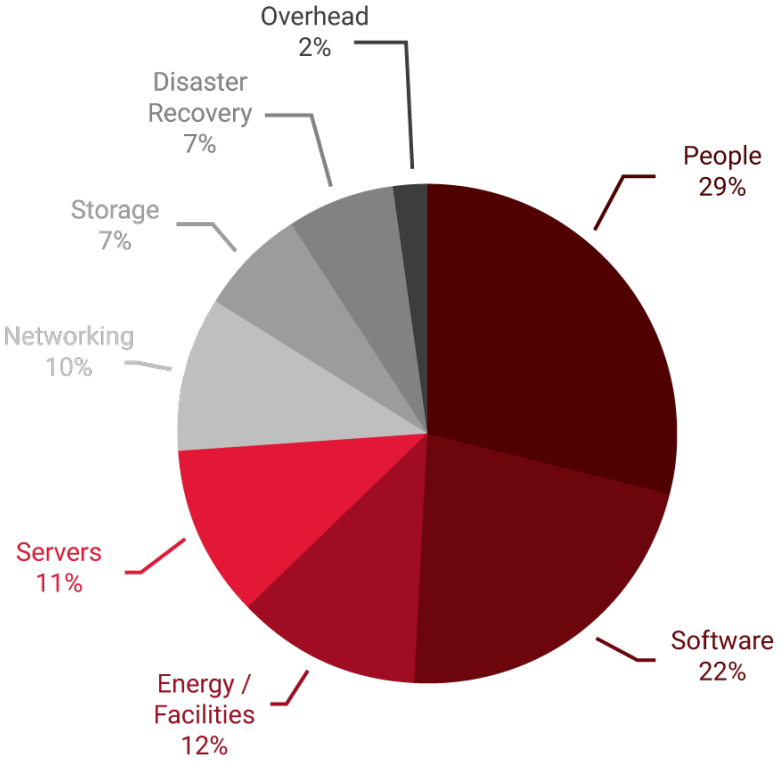
HYPERSCALE GROWTH



Next 60 Seconds on the Internet:

- 2,000,000 Google searches
- 27,800 Pictures uploaded to Instagram
- 514,000 Likes on Instagram
- 278,000 Tweets on Twitter
- 1,750,000 Likes on Facebook
- 208,700 Pictures uploaded to Facebook
- 204,000,000 Emails sent WW
- $83,000 Worth of product sold on Amazon
- 100 Hours of video uploaded to YouTube

EXPLODING INFRASTRUCTURE COST



~75% of Data Center Spending is OpEx

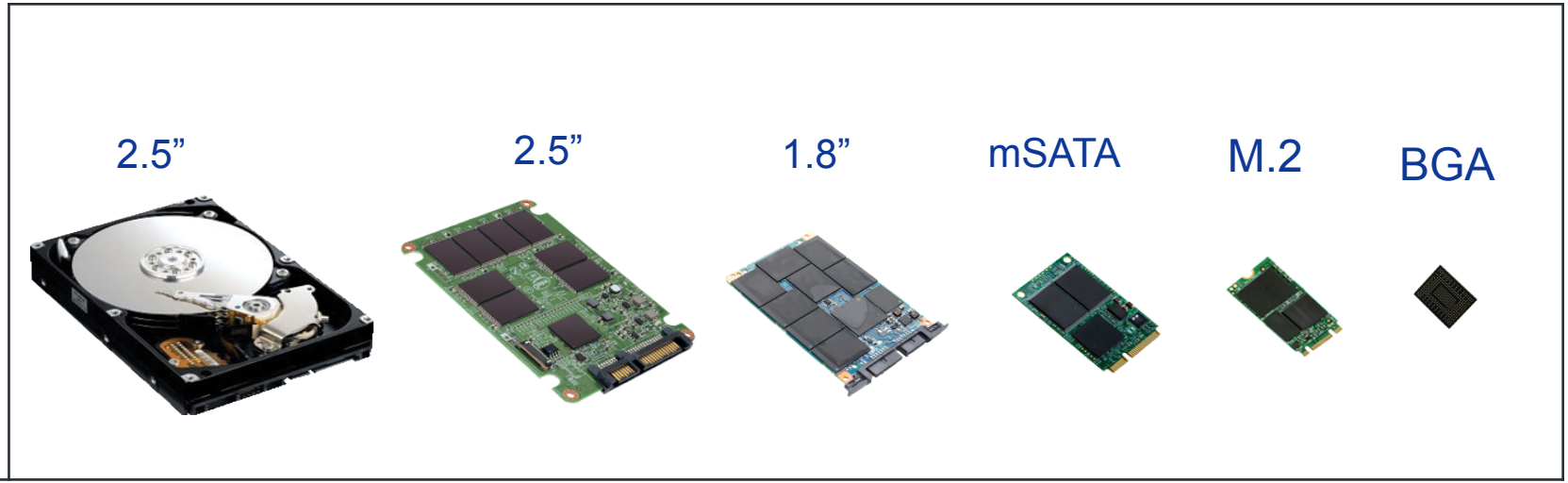
Power, Cooling, Floor Space, People and Overhead consume majority of the IT budget

“Do More – With Less”

STORAGE EVOLUTION

Flash Memory SUMMIT

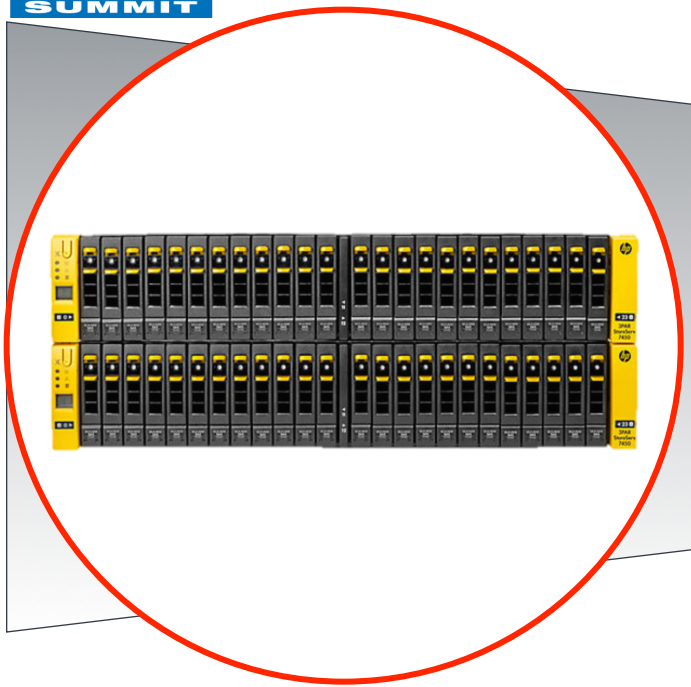
150MB/s → 4000MB/s



1989 2000 2010 2011 2014 2016

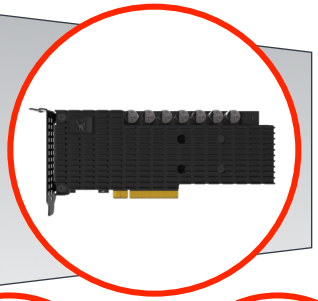
- 2.5-inch HDD
- 2.5-inch SSD
- 1.8-inch SSD
- mSATA SSD
- M.2 SSD
- BGA SSD

PERFORMANCE CONVERGENCE



Flash Storage Array

Convergence of Storage Performance to a PCIe Slot



10X IMPROVED TCO

10X LOWER LATENCY

High Performance PCIe SSD

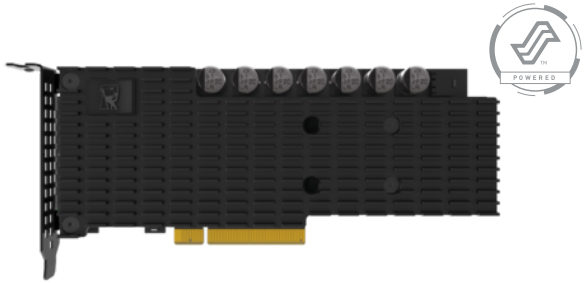
High performance PCIe storage enables FSA performance from single device

ACHIEVING 1M IOPS AND 5GBS

Storage Performance: 1 Million IOPS and 5GB/s

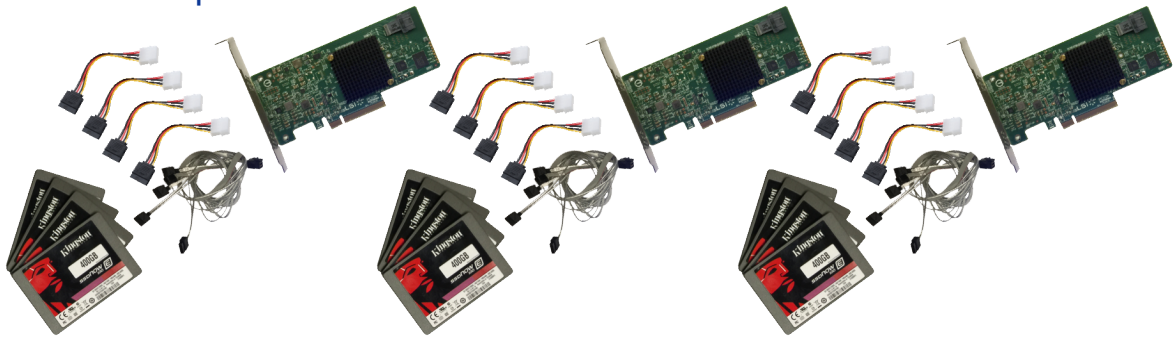
The primary goal of this white paper is to study different methods of achieving 1M IOPS and 5GB/s of storage performance with SATA vs. NVMe SSDs.

NVMe Setup:



* NVMe setup based on direct connect Kingston EP1000 Gen3x8 PCIe NVMe SSD

SATA Setup:



* SATA setup based on Kingston SATA SSD and LSI HBA

Best Method to Achieve 1M IOPS and 5GB/s – SATA vs. PCIe

TCO BENEFIT IN HIGHER PERFORMANCE

	SATA SSD	NVMe SSD
# of SSD	12	1
# of HBA	3	-
IOPS	~850K IOPS	~1.1M IOPS
Seq. Read	~6.7 GB/s	~6.4 GB/s
Latency	~175 us	~60 us
Total Power	~90W	~27W
Total Part Count	39	1

90% Smaller Foot Print

80% Reduced Part Count

70% Lower Power

60% Lower Latency

Lower cost and complexity enabled by higher performance PCIe

WHY PERFORMANCE MATTERS

Use Case Examples:

- *Virtualization*
- *Video Streaming*
- *QoS Software Layer*

HIGH PERFORMANCE VIRTUALIZATION

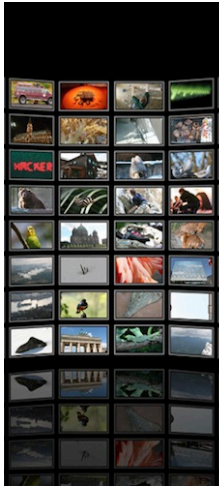
Vendor A = 450k IOPS
= 8 VMs

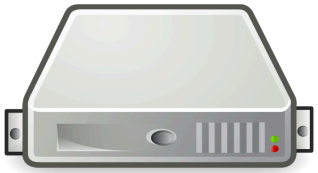
Vendor B = 1.25M IOPS
= 25 VMs

* Assumption: High-performance VM requires 50K IOPS

3x more infrastructure required to address same number of VMs

VIDEO STREAMING ENVIRONMENTS

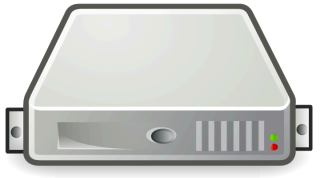




Vendor A = 2.8 GB/s
= ~9 HD Stream

2x more infrastructure required
to address same number of streams

HD Stream 24-bit @ 1080p @ 60 fps:
24× 1920x1080*60=2.99 Gbit/s



Vendor B = 7.2 GB/s
= ~20 HD Stream

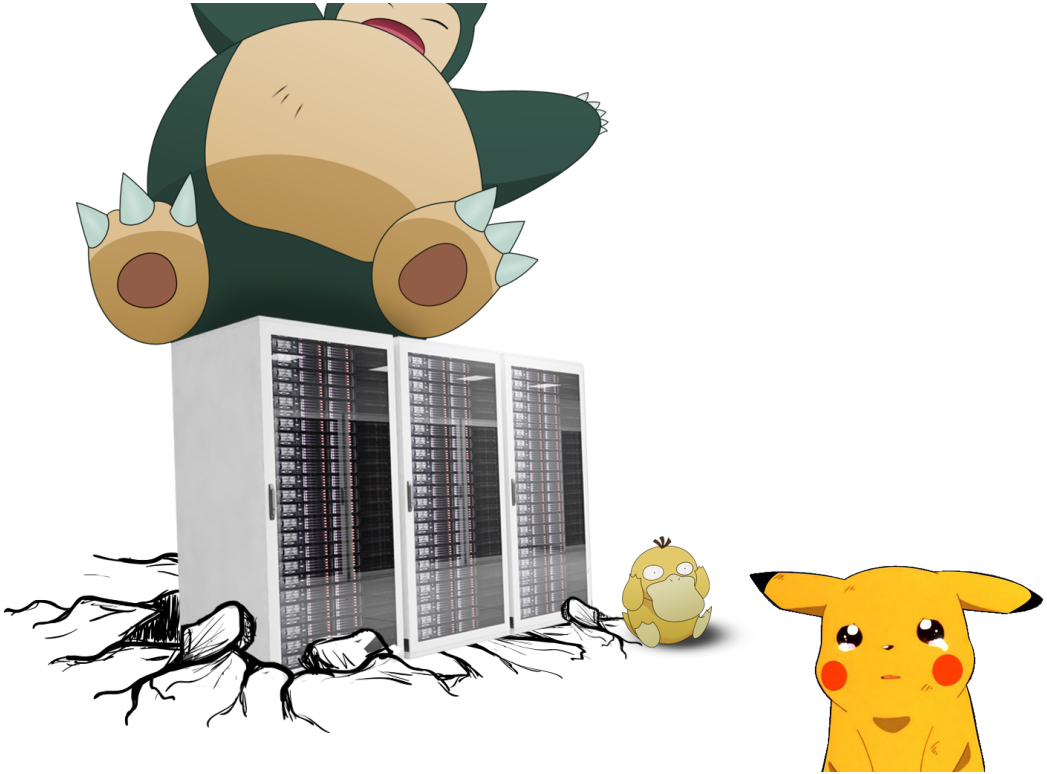
QUALITY OF SERVICE

QoS = Performance Rationing

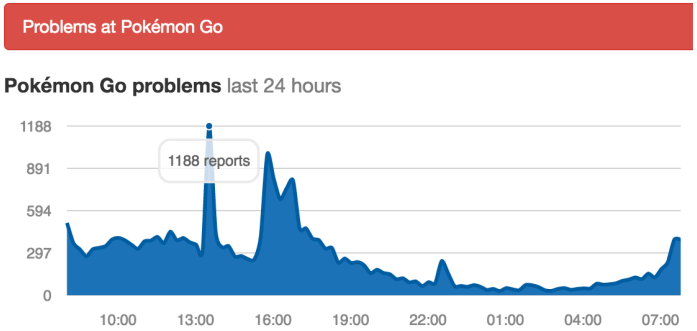
- QoS methods are required when not enough performance is available from underlined HW
- With sufficient performance from underline hardware – QoS can be removed/bypassed from architecture
- QoS adds significant cost & complexity – Software Cost, Deployment/Management Cost, Premium Product

Higher performance can eliminate/reduce cost of QoS

POKEMON DEMANDS PERFORMANCE



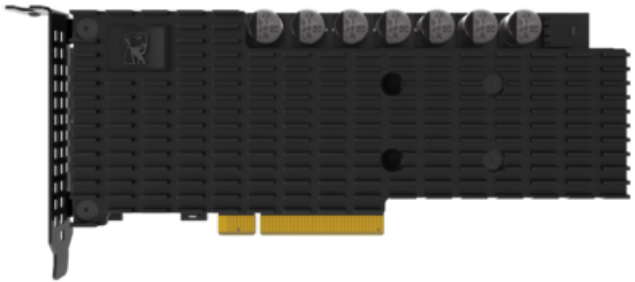
Pokémon Go's server issues have been driving people wild all day
- "Please come back later"



Infrastructure Unable to Meet Demand

KINGSTON'S NVME SSD

PCIe SSD
HHHL Card



POWERED

PCIe SSD
2.5" U.2

ULTRA PERFORMANCE SSD

ENTERPRISE HHHL NVME

	Vendor A	Vendor B	Vendor C	Kingston
Model	xxx	xxx	xxx	EP1000
Max Capacity	700 GB	2,000 GB	3,200 GB	7,800 GB
Interface	Gen 2.0 x8	Gen 3 .0 x4	Gen 3.0 x4	Gen 3.0 x8
Sequential Read 128KB	3,300	2,800	3,000	7,000
Sequential Write 128KB	630	2,000	1,600	6,200
Random Read 4KB	750,000	450,000	740,000	1,250,000
Random Write 4KB	95,000	90,000	140,000	275,000
Random Write 4KB (Peak)	150,000	400,000	450,000	900,000
P-fail	Yes	Yes	Yes	Yes

ENTERPRISE 2.5” NVME

	Vendor A	Vendor B	Vendor C	Kingston
Model	xxx	xxx	xxx	EP1000
Max Capacity	700 GB	2,000 GB	3,200 GB	3,920 GB
Sequential Read 128KB	1,800	2,800	3,000	3,600
Sequential Write 128KB	500	2,000	1,600	3,600
Random Read 4KB	430,000	450,000	740,000	900,000
Random Write 4KB	50,000	90,000	140,000	275,000
P-fail	Yes	Yes	Yes	Yes
Dual Port	No	No	No	Yes
SRIS	No	No	No	Yes

SUMMARY

Performance DOES matter

High Performance DOES drive lower TCO

Performance is fundamental to “do more – with less”

How can higher performance impact your business?