Performance Assessment of an All-RRAM Solid State Drive Through a Cloud-Based Simulation Framework

Lorenzo Zuolo*

Michele Cirella, Cristian Zambelli, Rino Micheloni*, and Piero Olivo

Lorenzo Zuolo, Michele Cirella, Cristian Zambelli and Piero Olivo are with the with the Dipartimento di Ingegneria, Università degli Studi di Ferrara, via G. Saragat,1 - 44122 Ferrara (Italy)

Lorenzo Zuolo, is also with SSDVision S.R.L., Via Dosso Dossi 11, 44122 Ferrara (Italy)

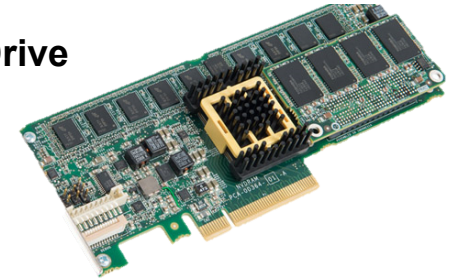
Rino Micheloni is with Microsemi Corporation, Via Torri Bianche 1, 20871 Vimercate (Italy)

*Work done at Università degli Studi di Ferrara

It is a fact: SSDs are now ubiquitous!

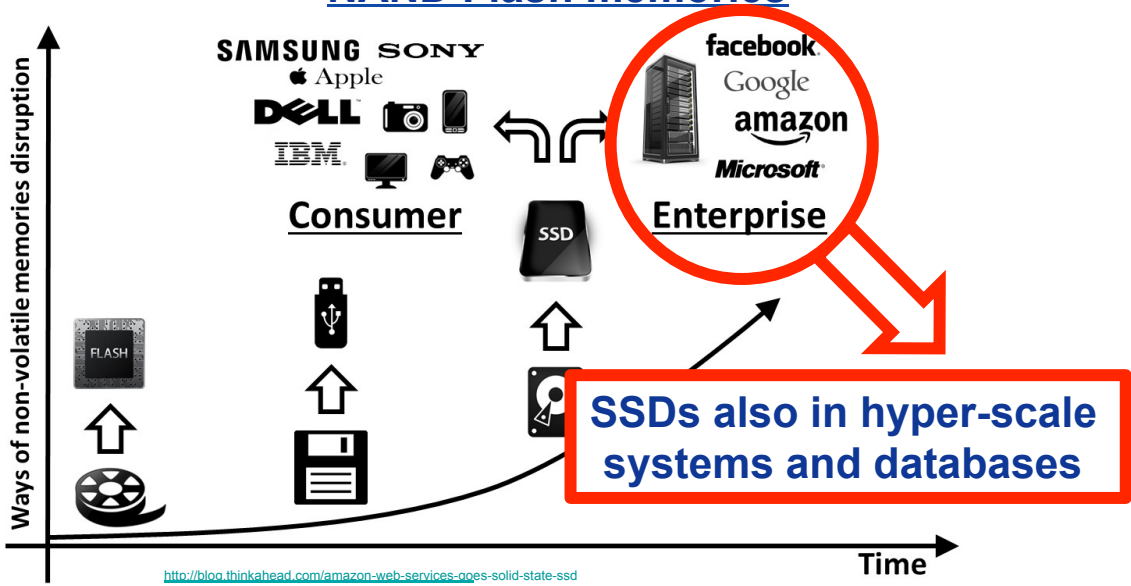
- **The Flash Memory Summit take away is:**
  - NAND-Flash based SSDs are ubiquitous and now they are also the most effective solution for high-performing mass-storage applications

FlashTec NVRAM Drive



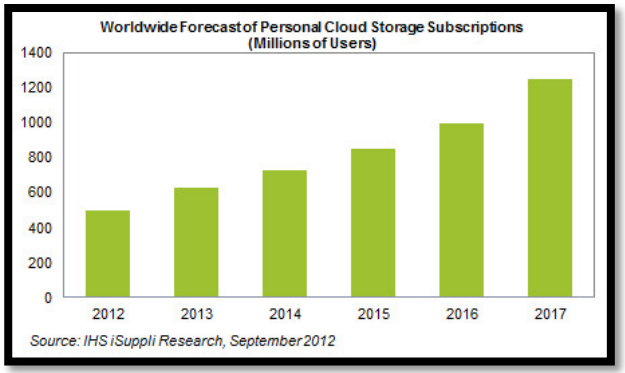
- **High speed**
  - Millions of IOPS
- **High robustness**
  - No mechanical parts
- **Low power consumption**
  - Power wall at 25W
- **Good reliability**
  - Two years with ten Disk Fills Per Day

This storage revolution is made possible thanks to NAND Flash memories

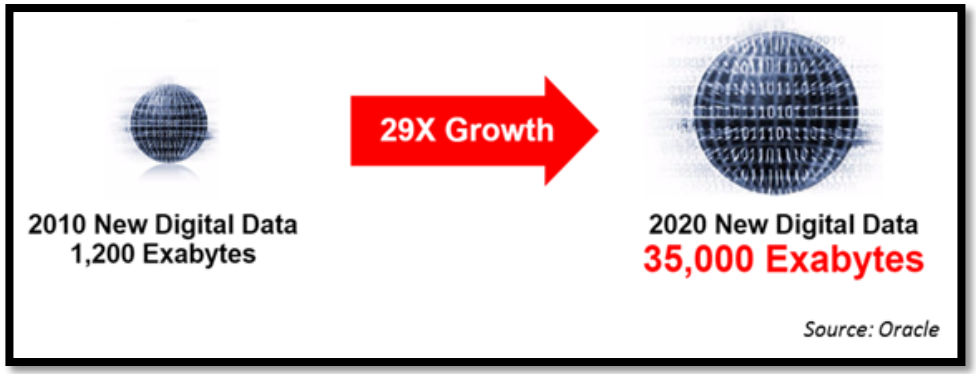


http://blog.thinkahead.com/amazon-web-services-goes-solid-state-ssd
http://www.wired.com/2012/06/flash-data-centers/
http://azure.microsoft.com/blog/2014/09/22/new-d-series-virtual-machine-sizes/
http://www.theregister.co.uk/2014/06/16/google_load_balancer_persistent_disks/

…Numbers in the hyper-scale market: the need for storage capacity...



IHS Technology said:
"In 2017 more than 1Billion users will subscribe a cloud storage service"



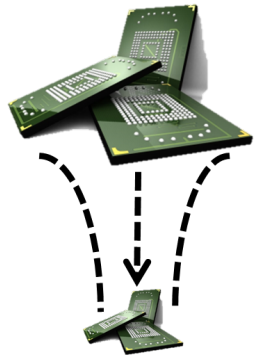
Oracle said:
"In 2020 the amount of new digital data produced will be around 35000 Exabyte"

…The challenge: how to cover these numbers with NAND flash-based SSDs?

…Challenge accepted: memory vendors improved the bit density!

- New storage paradigms: **SLC→MLC→TLC**
- An aggressive technology scaling: **3X→2X→Mid-1X**

} **1Tbit chips**



http://www.isuppli.com/mobile-and-wireless-communications/news/pages/subscriptions-to-cloud-storage-services-to-reach-half-billion-level-this-year.aspx
http://www.oracle.com/us/products/servers-storage/storage/oraclestoragestrategyupdate-424276.pdf

The side-effects of the improved bit density in NAND flash memories

1. **Memory vendors partially solved the need for storage capacity… but…**

2. Changing the storage paradigm **worsened the Read, Program, and Erase latencies**

3. **the aggressive technology scaling shifted the RBER from $1x10^{-6}$ (of 3X SLC) to $1x10^{-2}$ (of mid-1X TLC)**
   - LDPC ECCs are needed which take too long to correct read data

Storage paradigm	Avg. read latency	Avg. program latency	Avg. Erase latency
SLC	25 µs	200 µs	1000 µs
MLC	45 µs	800 µs	2000 µs
TLC	**85 µs**	**2200 µs**	**5000 µs**

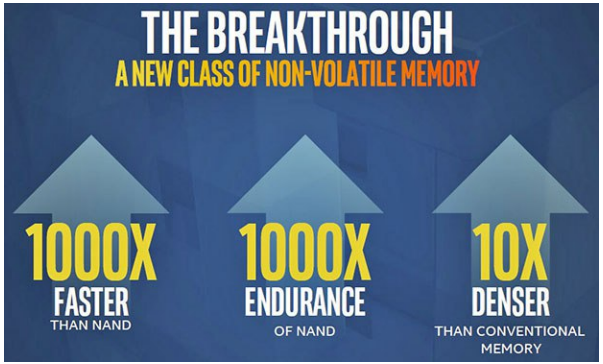
Technology node	Avg. RBER (BOL)	Avg. RBER (EOL)	Needed ECC
3X	$1x10^{-6}$	$1x10^{-3}$	BCH
2X	$1x10^{-5}$	$5x10^{-3}$	BCH
Mid-1X	**$1x10^{-4}$**	**$1x10^{-2}$**	**LDPC**

- **Final take away: the density/performance tradeoff**
  - **SSDs become no longer "high performing storage systems" when you need a high storage density!**

The density/performance tradeoff: Is there any countermeasure?...

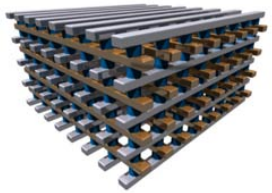
- **A step back… What is the "holy grail" of storage?:**
  - DRAM-like performance (latency and bandwidth)
  - DRAM-like reliability
  - NAND flash-like storage density
  - NAND flash-like data non-volatility

- **Storage-Class Memories (SCMs) seem to answer the question**
  - **Micron's 3D-Xpoint**
  - **Crossbar's RRAM**
  - **Everspin's MRAM**
  - **Etc…**



CROSSBAR MEMORY
A COMPLETELY NEW CLASS OF RRAM

- Up to 1 Terabyte per chip
- 20X faster write than NAND
- 10X better endurance



http://www.techshout.com/hardware/2013/06/crossbar-unveils-rram-with-whopping-1tb-memory/
http://www.innovationtoronto.com/2015/07/storage-technology-thats-1000-times-faster-than-current-ssds/

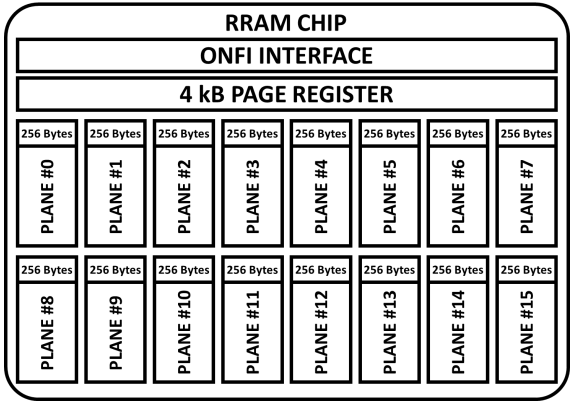
SCMs: the system-level perspective (1)

1. **If we follow the memory vendors claims' it seems that SCMs will become one the main storage layers in:**
   - Datacenters
   - hyper-scale systems
   - high-performance computing (HPC)
   - Etc. …

2. **To do that… SCMs have to be fully compatible with traditional SSD architectures:**
   - ONFI, Toggle, or DRAM-like interface
   - Plug 'n play with NAND flash memories (command set, timings, etc.)
   - Plug 'n play with SSD controller designs (internal organization, bus frequency, etc.)

3. **Let's take an example of SCMs: RRAMs**

Question is: can I build a All-RRAM SSD?



Chip Parameters	Configuration
IO-Bus interface	ONFI
IO-Bus speed	800 MT/s
Native page size	1T-nR 256 Bytes
Emulated page size	512-1024-4096 Bytes
T_{READ} per Page	1 µs

Data from real chip datasheet

…Let's do it: the "All-RRAM" SSD

- **Unfortunately we do not have real RRAM chips…**
  - **We do have the datasheet and all the specs, but no real silicon**

- **… And we cannot spend months in developing an FPGA-based design for:**
  1. **Emulate the controller behavior**
  2. **Emulate the RRAMs behavior and timings (… this point could be very tricky…)**

- **We need numbers, and we need them now…**

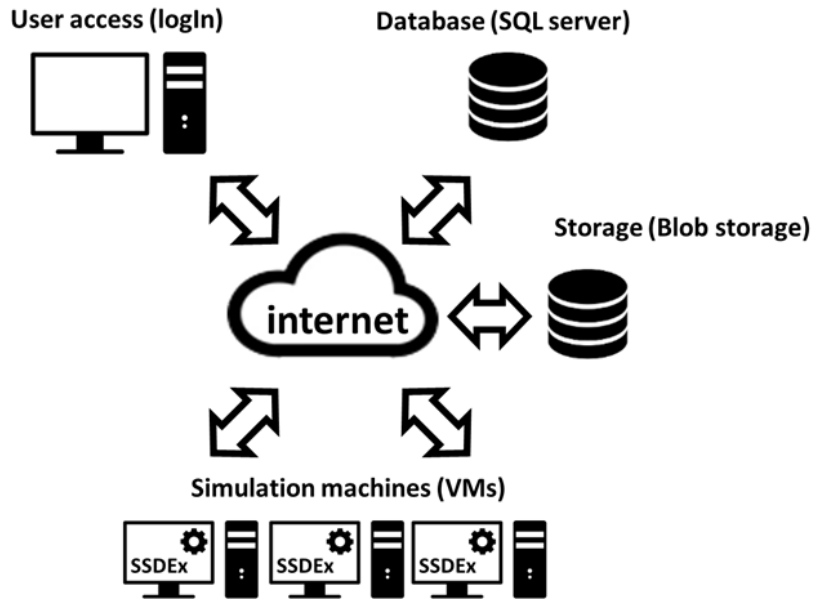
- **Ok, no problem, we can simulate it but…**
  1. Simulations have to be **accurate**
  2. We **don't want to spend too much time in simulations**
  3. We want to **test the SSD in several different configurations**
  4. We want to **understand what are the "corner-case" working conditions**

SSDExplorer: a cloud-based service for SSD simulation

SSDExplorer By SSDVision

- **The world first cloud-based tool for SSD design-space exploration: www.ssdvision.com**

- **Cycle-Accurate** (tuned on real HW)
- **Reconfigurable**
- **Fast** (up to 10 Million-transactions per day)



The strength: the simulation environment is in the cloud

All-RRAM SSD testing conditions: baseline (1)

- **The question to answer: will an All-RRAM SSD outperform a NAND flash-based SSD?**
- **Remember: the main feature of SCMs is that they are fully compatible with traditional SSD architectures**

Working assumptions are:

1. **Use the same SSD controller configuration used in traditional NAND-flash based SSDs**
2. **Change only the storage paradigm**
   - 3X-SLC NAND Flash
   - 1X-MLC NAND Flash
   - 1T-nR RRAMs
3. **Host queue depth is 32 (real application)**
4. **RRAMs use the 4 KBytes emulated page size**

Parameters	Configuration
Host Interface	PCIe Gen 2x8
Host protocol	NVMe 1.2
Host workload	4 kBytes 100% Random Read
SSD Channels	16
SSD Targets	8
SSD size	2 TBytes

All-RRAM SSD testing conditions: baseline (2)

Memories characteristics	NAND Flash 3X-SLC	NAND Flash 1X-MLC	1T-nR RRAM
T_{READ}	25 µs	40 µs	1 µs
IO-Bus Interface	ONFI	ONFI	ONFI
IO-Bus Speed	800 MT/s	800 MT/s	800 MT/s
Page size	4 kB	4 kB	4 kB

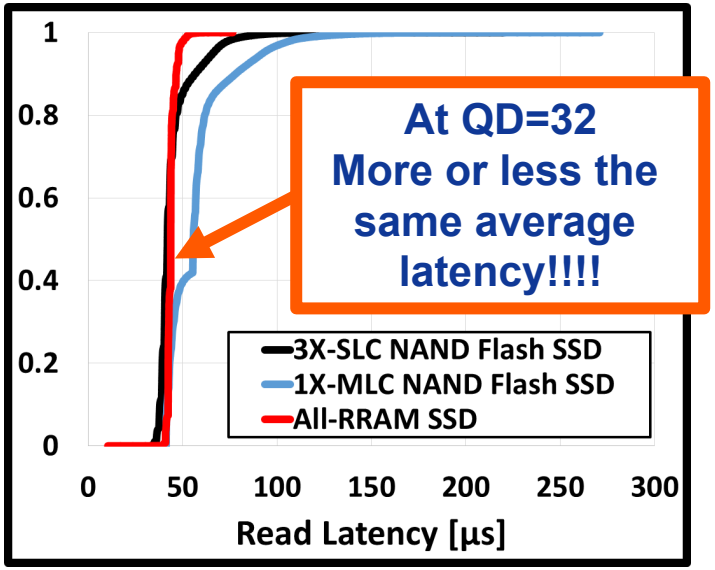
The answer is: NO the All-RRAM SSD works like a NAND-based SSD

…

Any available optimization?

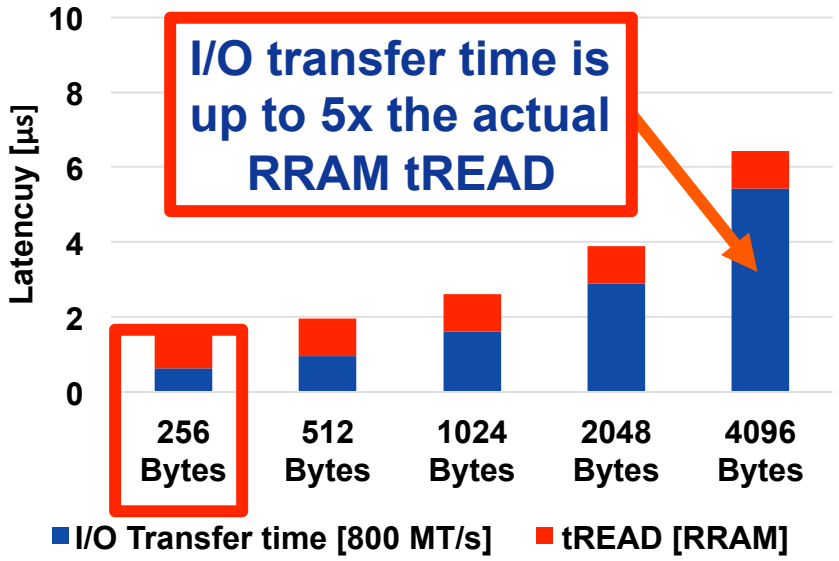
Well the expected conclusion is:

- **The All-RRAM SSD should be faster than any NAND-flash based SSD**



Optimized All-RRAM SSD: v1 (1)

- Let's do a step back…
- RRAMs are designed to work with a page size of 256 Bytes...
- But filesystems work with a 4 KBytes sector size
- RRAMs can emulate the 4 KBytes page but it is not a efficient solution
- The I/O transfer time takes a lot of time

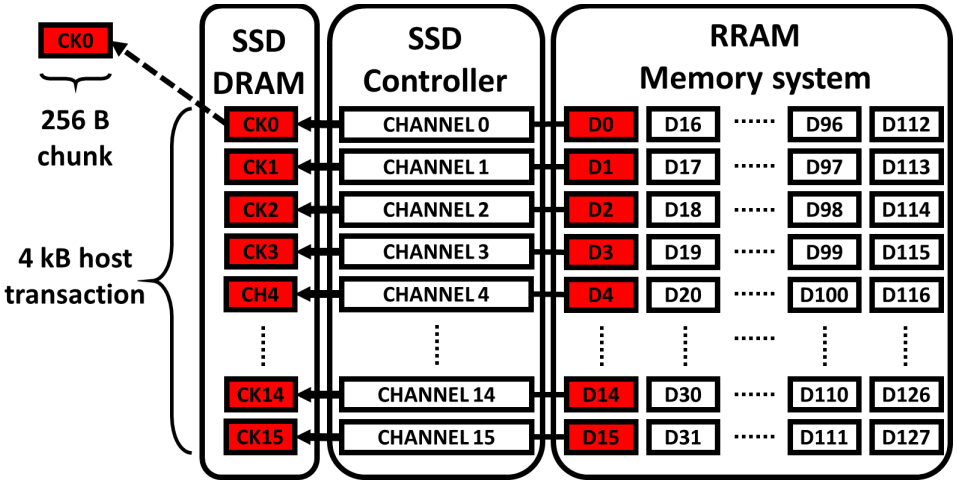


What about reducing the data transfer time using RRAM chips with the native 256 Bytes page size?

Optimized All-RRAM SSD: v1 (2)

- **But how can we handle it?**
  - **4 Kbytes from the host**
  - **Page-size of 256 Bytes**
- **Hint #1: the SSD controller has 16 channels**
- **Hint #2: 16 channels x 256Bytes = 4096Bytes**
- **Hint #3: the SSD controller has a DRAM buffer**

- **Let's do it in the SSD controller!**
  1. **Striping the host LBA across the 16 SSD channels**
  2. **Rebuild the LBA in DRAM**



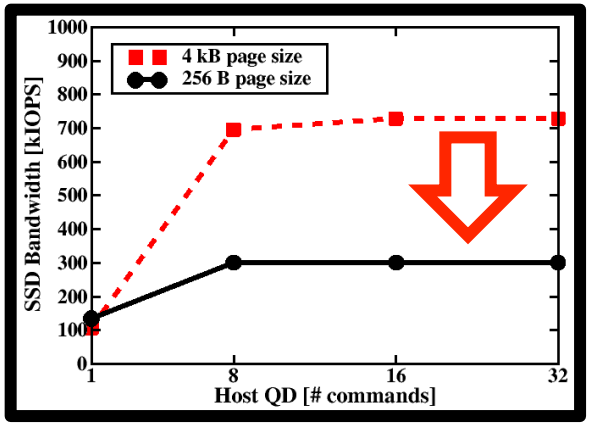
Optimized All-RRAM SSD: v1 (3)

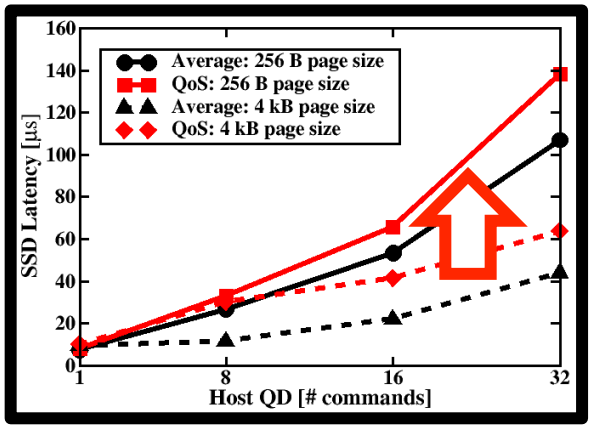
- **The quality metrics we assessed with SSDExplorer:**
  - **Bandwidth: kIOPS**
  - **Average latency: µs**
  - **Quality of Service (QoS): 99.99th percentile of the latency distribution**
  - **Host queue depth (QD): 1, 8, 16, 32**

- **Expected result: because of the 256 Bytes page size, the All-RRAM SSD should show high IOPs and low read latency**

- **Result: with respect to the traditional 4 kB page size mode (dashed lines), splitting the host transaction in 16 chunks of 256 Bytes (solid line) worsen the All-RRAM SSD performance**

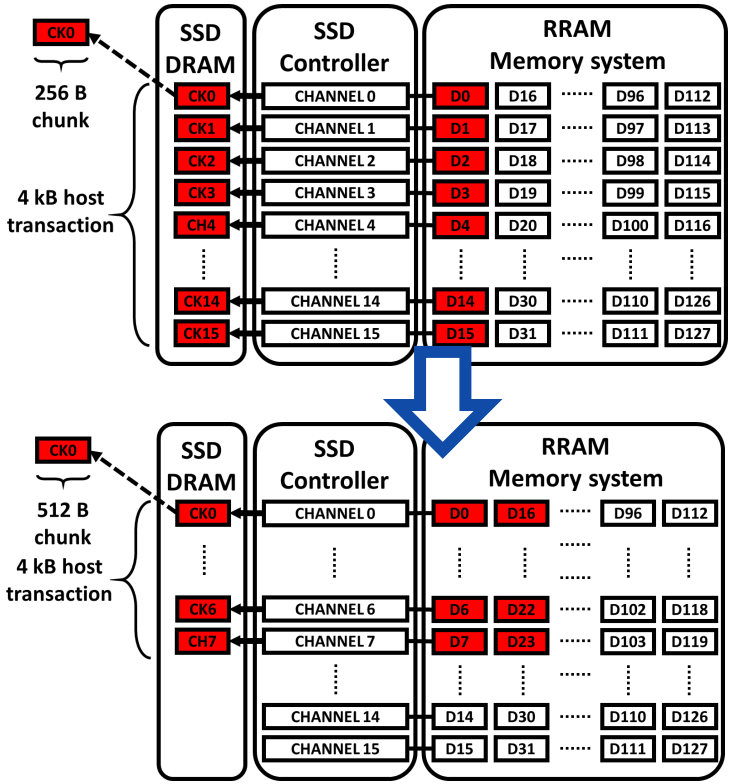
- **Why?:**
  - **we are multiplying by 16 the number of commands the SSD controller internally manages**





Optimized All-RRAM SSD: v2 (1)

- **Some considerations:**
  1. The plug 'n play approach, (NANDs to RRAM + same SSD controller) was not a good solution
  2. RRAM with 256 Bytes native page size + striping the LBAs across the 16 channels was not a good solution
- **New idea: what about co-design?**
  - Co-design the RRAM page size with the SSD controller considering the whole memory system architecture
- **How?**
  - Still striping the page across channels
  - But using different RRAM page sizes
- **E.g.:**
  - 16 Channels → 256 Bytes page size: TESTED
  - 1 Channel → 4096 Bytes page size: TESTED (Baseline)
  - 8 Channels → 512 Bytes page size
  - 4 Channels → 1024 Bytes page size
  - 2 Channels → 2048 Bytes page size



Optimized All-RRAM SSD: v2 (2)

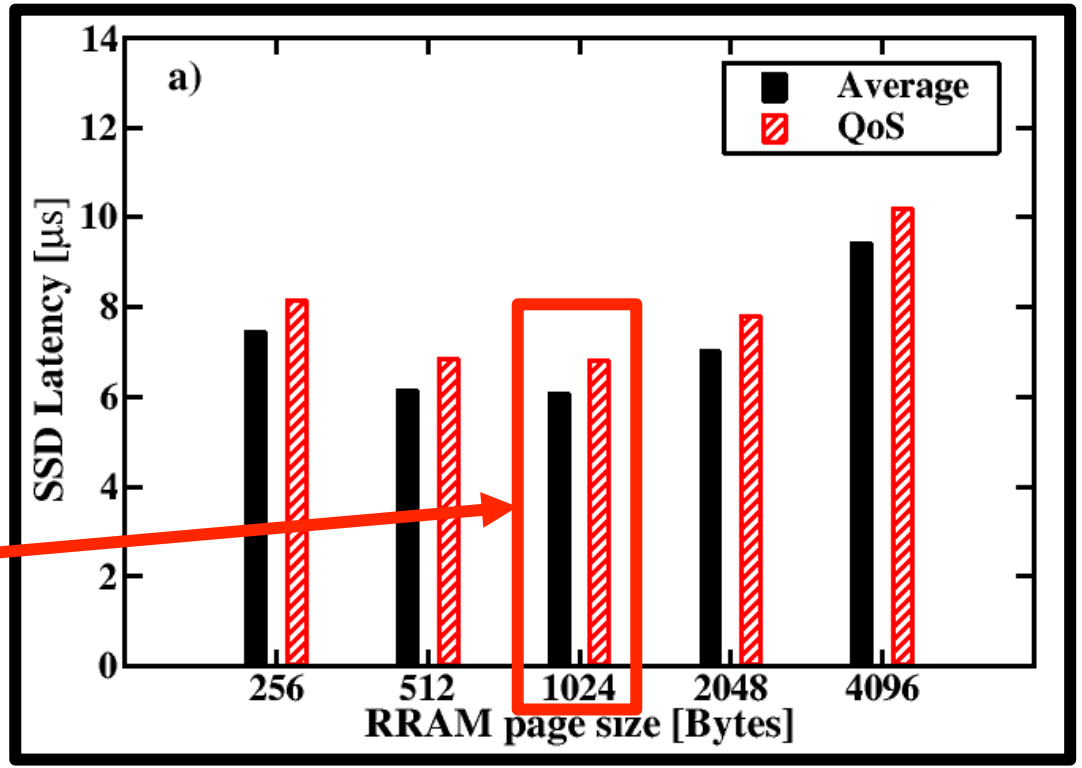
QD = 1 is considered

Take away #1:

- the optimum disk latency is achieved neither with the standard 256 Byte page size nor with the 4 KByte NAND-like mode.

Take away #2:

- There is a minimum in both the average latency and the QoS at a page size of 1 Kbytes
- 4 Channels striping → 1024 Bytes page size



Optimized All-RRAM SSD: v2 (3)

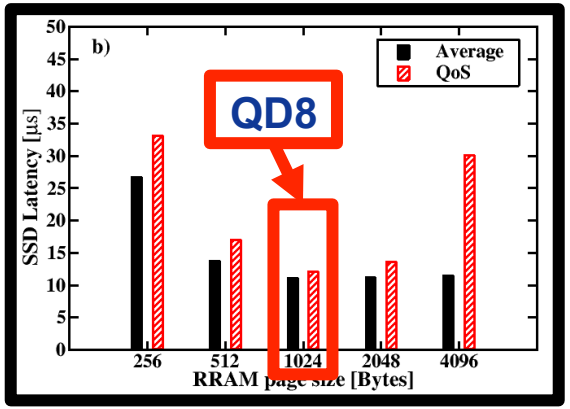
QD = 8, 16, 32 are considered

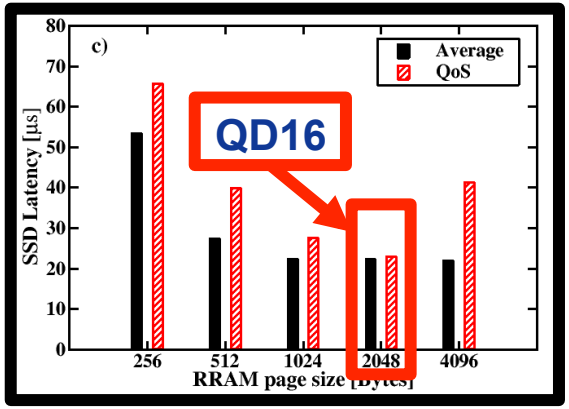
Take away #1:

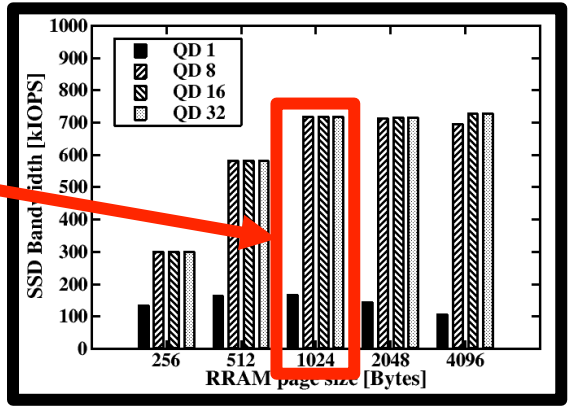
- **Also at different QD, depending on what it is necessary to optimize either the average latency or the QoS, a page size of 1 Kbytes or 2 Kbytes seems to be the better choice**

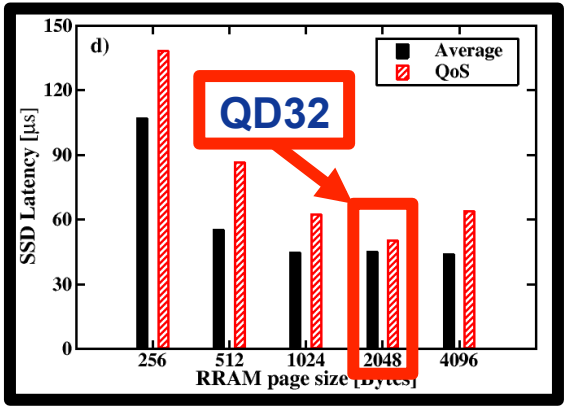
Take away #2:

- **These considerations hold on also for bandwidth**









Conclusions & open-questions

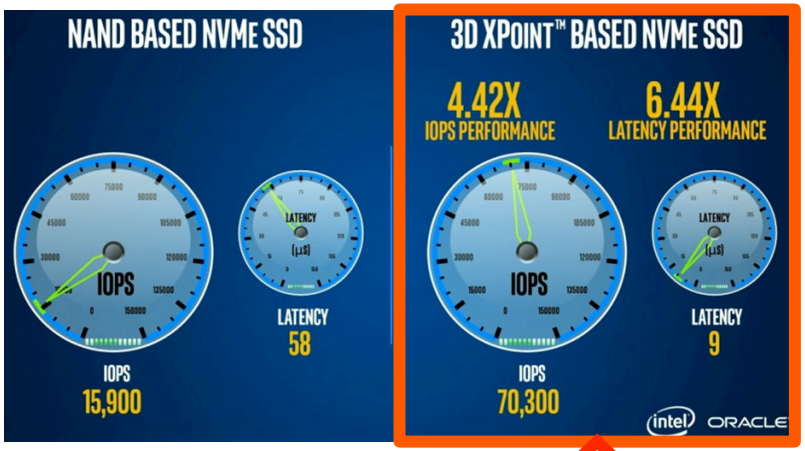
1. **In this work we learned that:**
   - **RRAMs are good SCMs but not gold when used in SSDs**
   - **to build a high-performing SSD with RRAMs it is mandatory to co-design the memory architecture with the SSD controller**

2. **SSDExplorer helped us to identify the architectural bottlenecks and to design a more efficient solution**
   - **More than 100 different simulations were needed to find the optimal design point**

3. **…These considerations can be applied to any SCM with the same characteristics of the tested RRAMs…**

… So… What about Micron's 3D-Xpoint and the Intel Optane? (1)

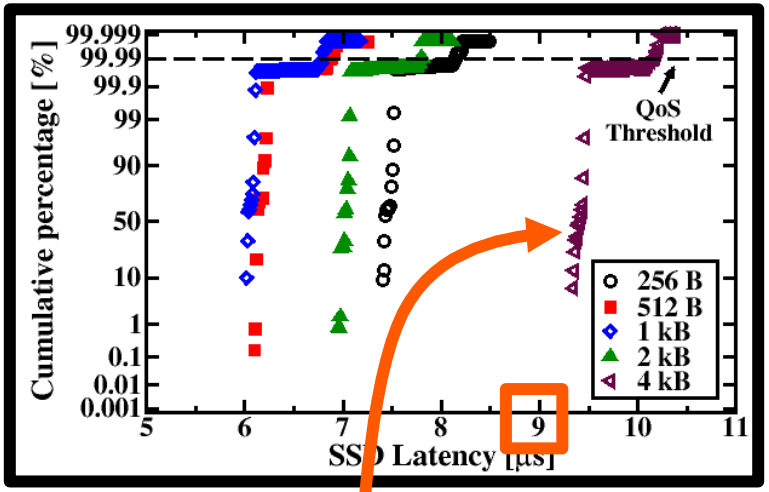
- **Intel Optane, is the world first "All-SCM" 3D-Xpoint-based SSD**
  - **(Still not in mass production)**
- **3D-Xpoint are used as NAND flash replacement**
- **…Sounds familiar… The same approach we used in All-RRAM SSD**
- **At IDF Intel showed some performance numbers…**

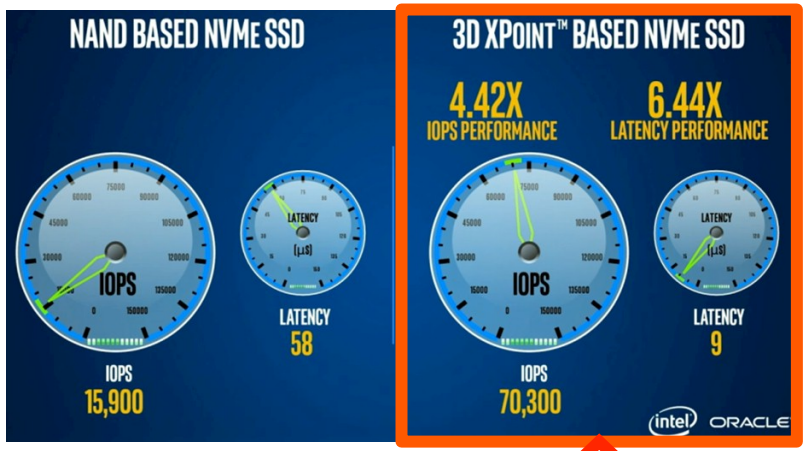


These numbers are achieved at Queue Depth (QD) = 1

http://wccftech.com/intels-3d-xpoint-memory-featured-optane-ssds-optane-dimms-8x-performance-increase-conventional-ssds/

… So… What about Micron's 3D-Xpoint and the Intel Optane? (2)





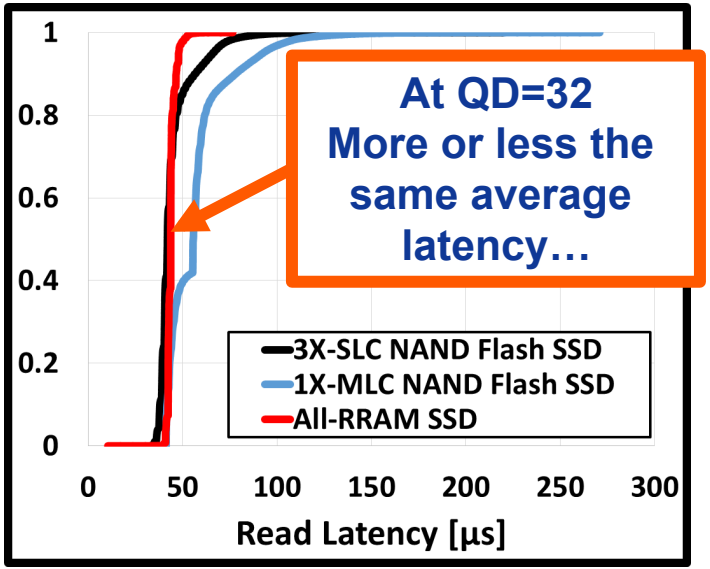
All-RRAM SSD simulated in this work

- **QD = 1 is used**
- **With a 4 Kbytes page size we have roughly the same latency of Optane…**

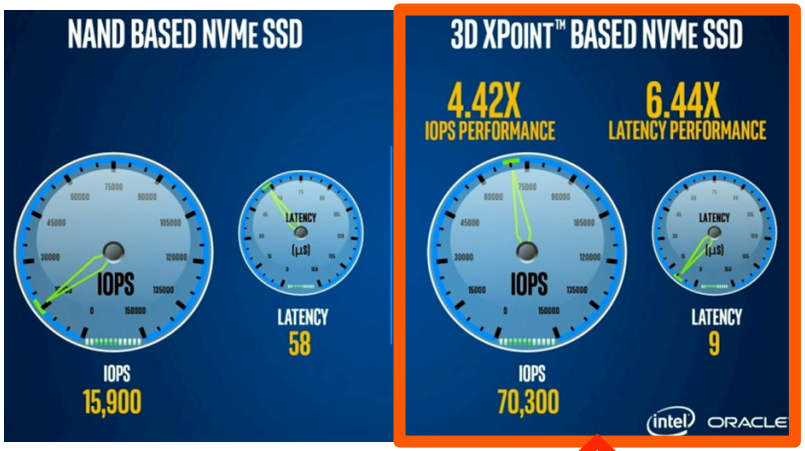
These numbers are achieved at Queue Depth (QD) = 1

http://wccftech.com/intels-3d-xpoint-memory-featured-optane-ssds-optane-dimms-8x-performance-increase-conventional-ssds/

… So… What about Micron's 3D-Xpoint and the Intel Optane? (3)



At QD=32 More or less the same average latency…



These numbers are achieved at Queue Depth (QD) = 1

All-RRAM SSD simulated in this work

- **Compared to SLC-MLC NAND flash**
- **QD = 32 is used**

Thanks

Q&A