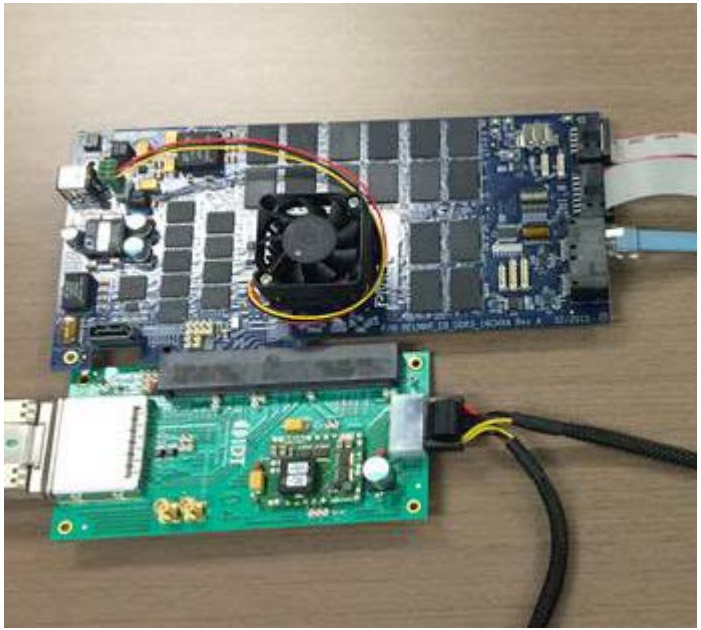
Modelling a High-Performance NVMe SSD constructed from ReRAM

Stephen Bates, PMC
Mehdi Asnaashari, Crossbar Inc
Lorenzo Zuolo, PMC

PMC Flashtec SSD Controller

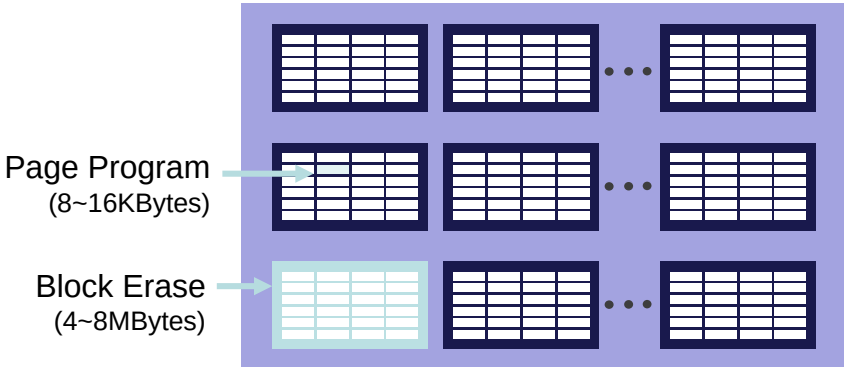
- High performance NVM Express SSD controller
- Up to 32 ONFI/Toggle flash channels
- DDR3/4 DRAM interface
- Flexible LDPC ECC
- Data Integrity, encryption etc.



The PMC Flashtec evaluation card with Micron NAND and DDR. No ReRAM attached (yet) unfortunately ;-).

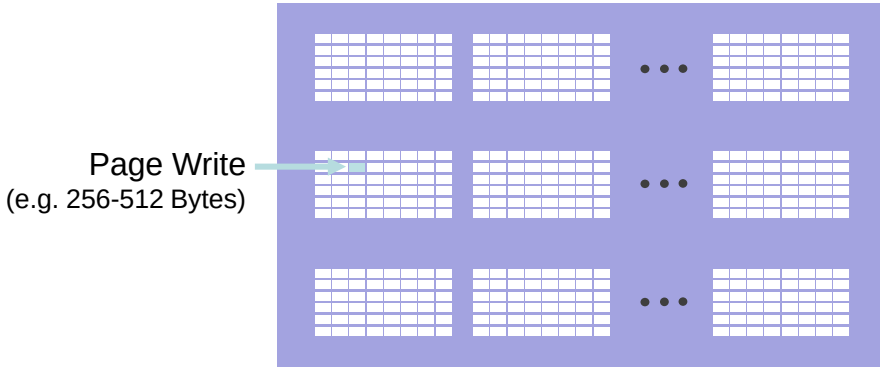
Crossbar RRAM Memory

NAND Flash



No possibility to update a page
Requires **Erase** Operation
10-30% **Over-Provisioning**
Write Amplification of 2.5-3

Crossbar RRAM



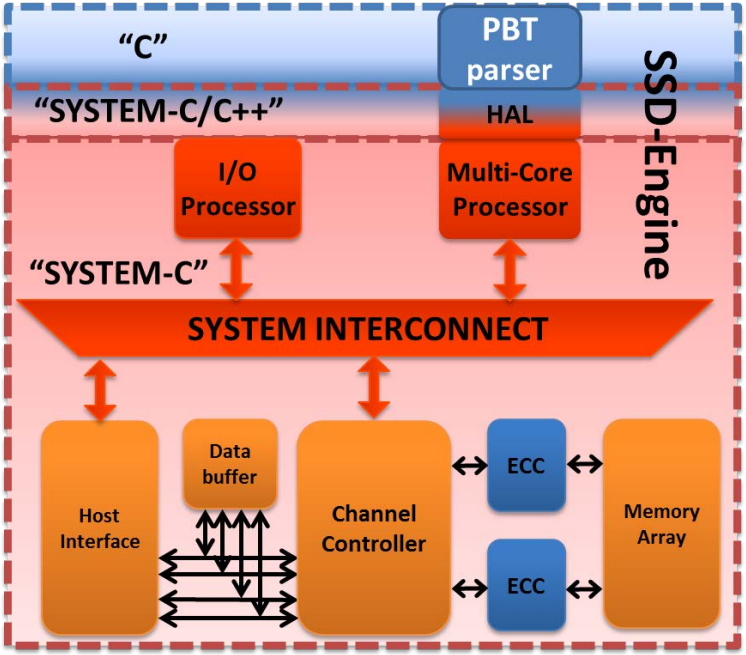
Every page can be updated
No Erase Operation is Required
No Over-Provisioning
Write Amplification of 1

Crossbar RRAM Memory

	NAND	ReRAM	Comments
Page write	1,000 us	2 us	
Block Erase	10,000 us	0 us	Not required
Read latency	100 us	1 us	
Write cycles	1,500	100,000	
Over-writes	N/A Resulting in Write Amplification	Allowed No Write Amp.	
Page size	16-32 KB	0.2-0.5 KB	Small is better
Retention	3 yrs @ 25C	10 yrs @ 85C, after 10K cycles	
Scaling	Limit ~20nm	<10nm	

SSD Explorer

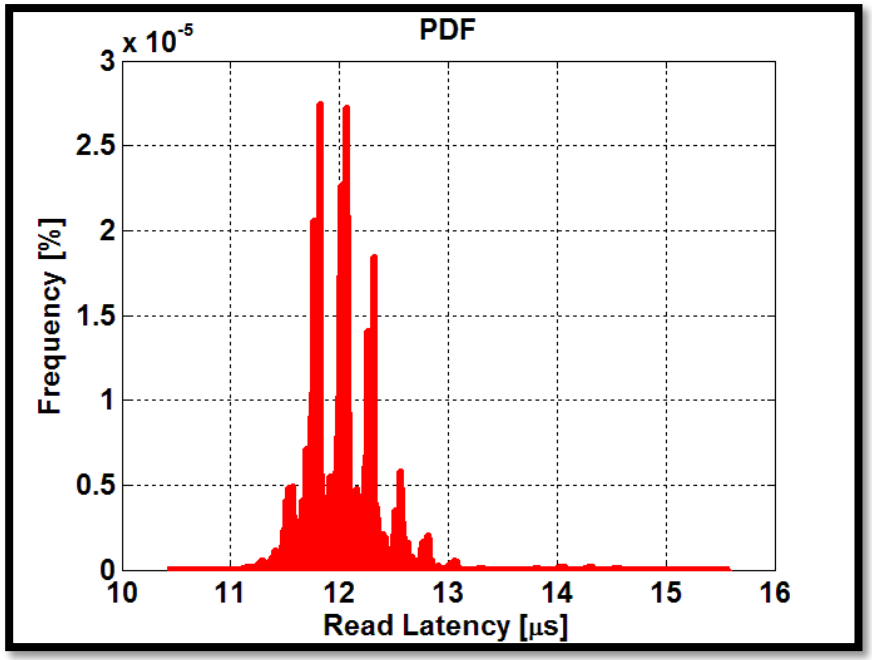
- C simulation of SSD controller, FW, NAND and DRAM
- Inputs include NAND topology, FTL algorithms, ECC, RBER, DDR timings
- Outputs include bandwidth, IOPS, latency
- Calibrated against real SSD data

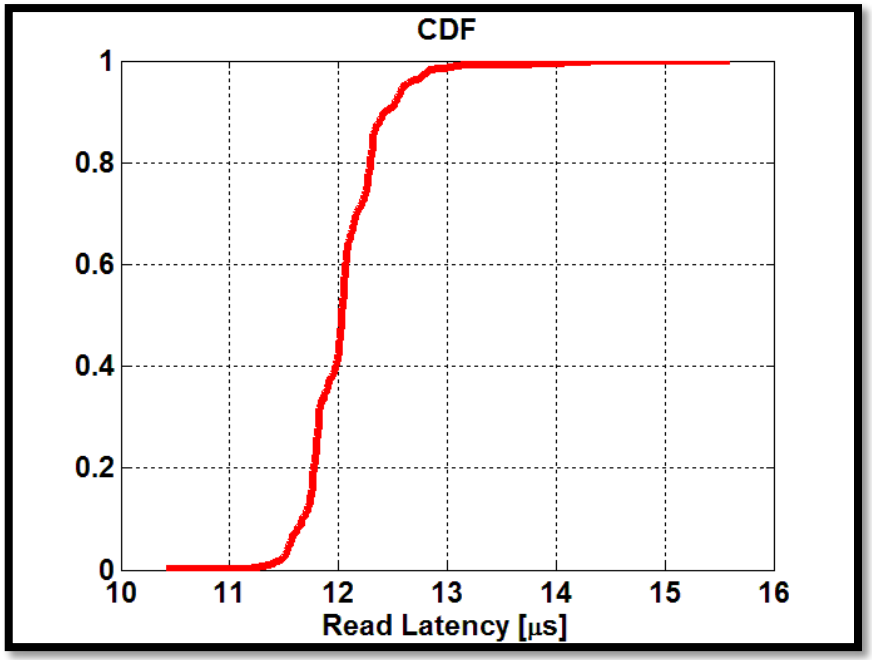


QD=1 Read Latency

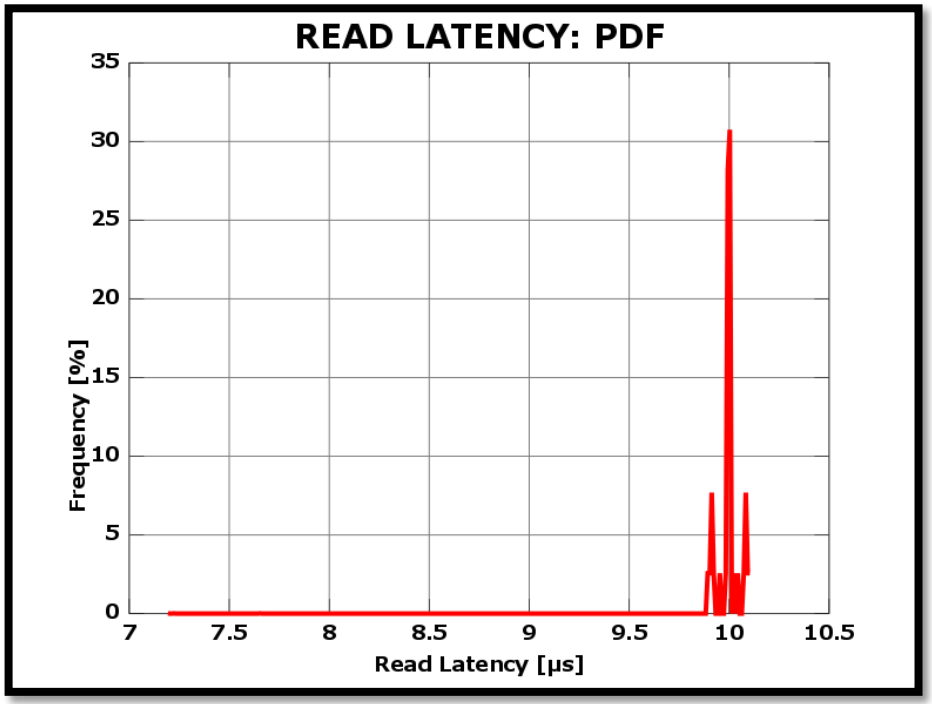
- Queue Depth = 1
- 256B LBA
- Random Reads
- NVMe Protocol
- MSI-X is not simulated (i.e. polling mode)

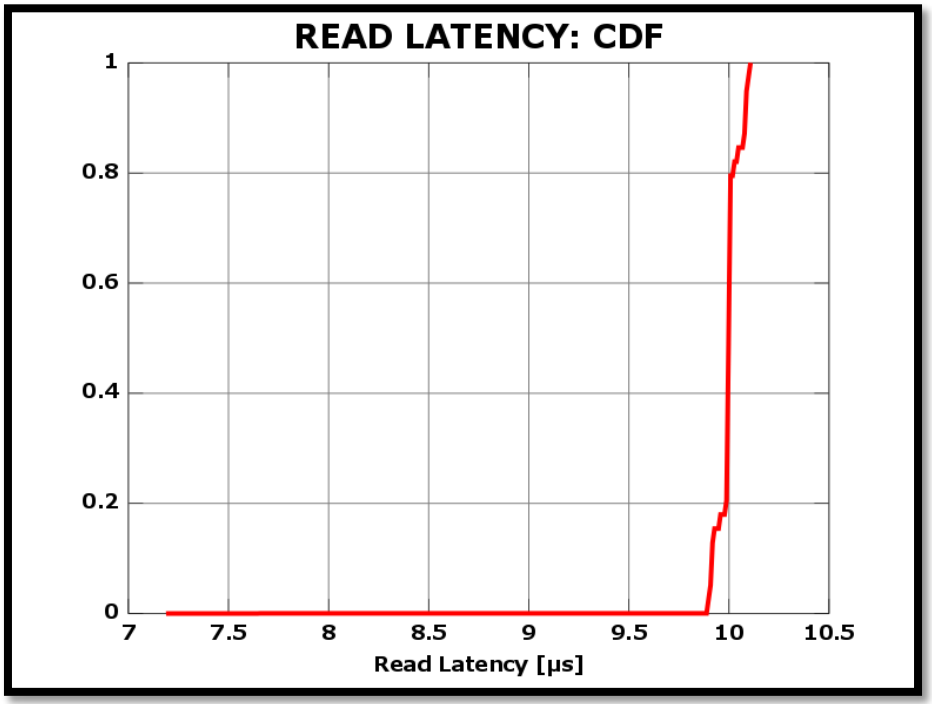
QD=1 Read Latency





QD=1 Read Latency





QD=1 Read Latency Results

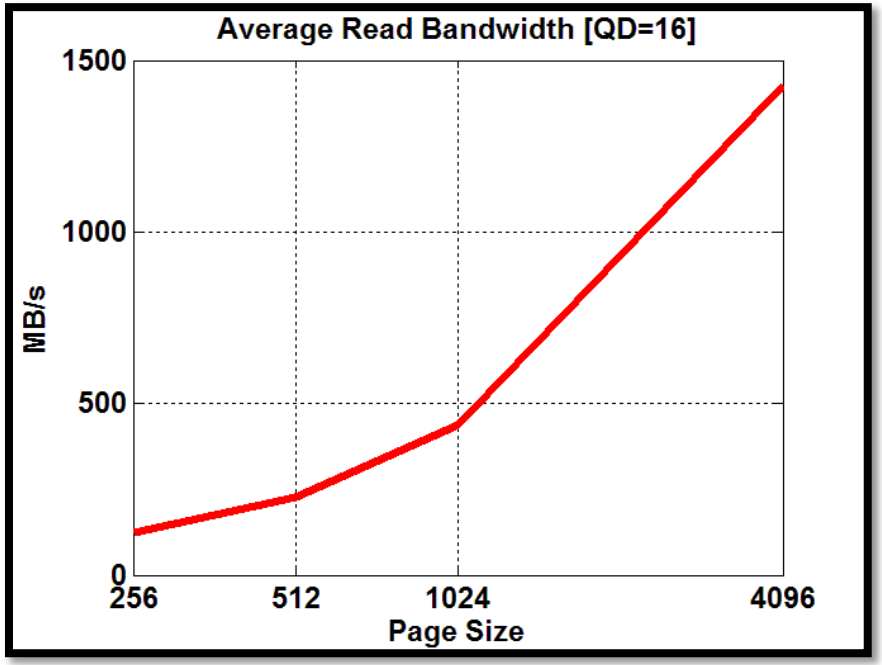
- Minimum Latency = 10.4 us.
- Average Latency = 12.05 us.
- 99.99% Latency = 15.6 us.
- 83000 IOPs ~= 22MB/s

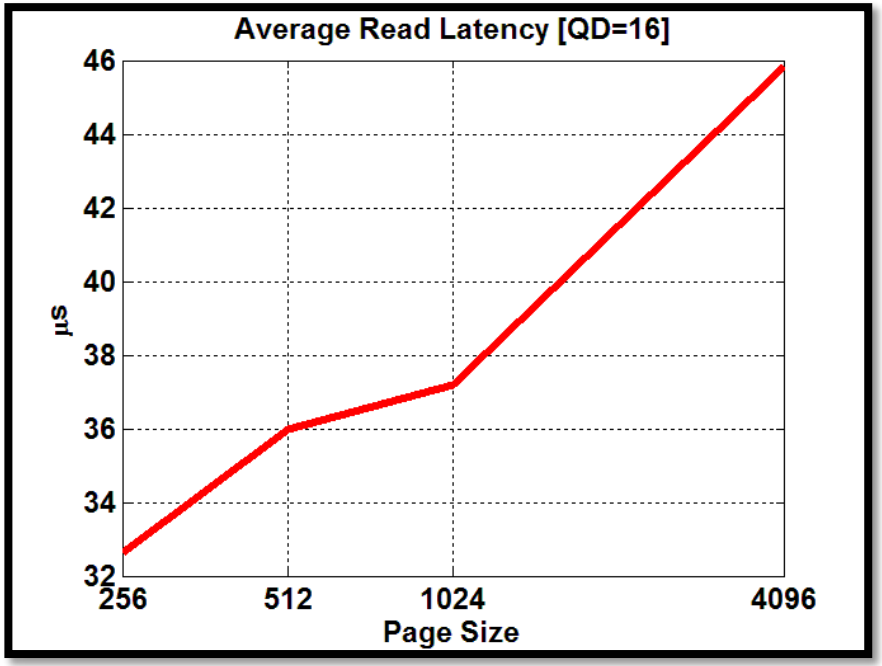
~12us Read Latency at QD=1!

QD=16 Read Latency

- Queue Depth = 16
- 256 to 4096 Byte LBA
- Random Reads
- NVMe Protocol
- MSI-X is not simulated (i.e. polling mode)

QD=16 Read Latency





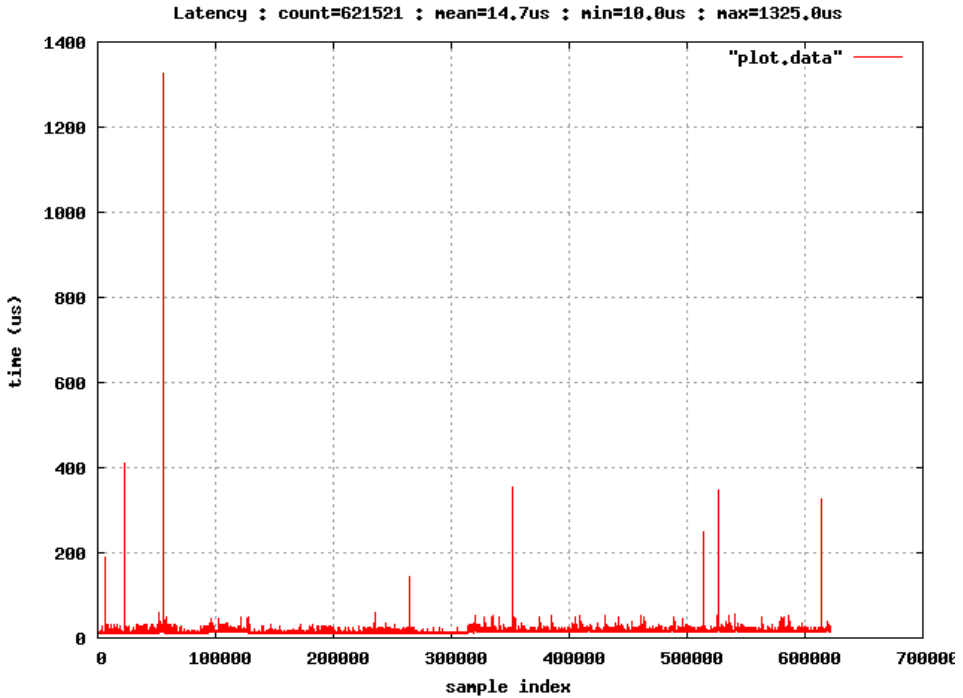
QD=16 Read Latency Results

- Average Latency = 33/46 us (256/4096B).
- Bandwidth = 100/1500 MB/s (256/4096B).

~33us Read Latecny at QD=16!

NVM Express

- Although the NVM Express driver is really good it does have issues as t_read drops.
- For example MSI-X service times are non-deterministic in Linux and can contribute to outliers.
- Polling modes may be needed to enhance NVMe for NG-NVM



Approximately 600,000 consecutive 4KB random reads from PMC Flashtec NVRAM card using Ubuntu and 3.13.0 in-box NVMe driver. FW time-stamping shows latency outliers are not from the drive but come from MSI-X handling.

Next Steps

- **Read/Write workloads**. The symmetric access times and lack of erase will help here.
- **Optimizations**. Currently just dropping in ReRAM. Need to retune SSD FW to increase utilization.
- **Real ReRAM!**

Conclusion

- NG-NVM is coming and will enable extremely low latency PCIe attached SSDs
- NVM Express is a great way to talk to NG-NVM
- Low consistent latency at QD=1
- OS and Driver latencies now dominate both average and outliers. Need to address!