Beyond NVDIMM: Future Interfaces for Persistent Memory

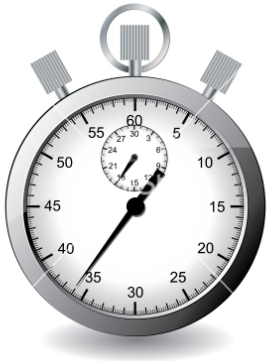
Stephen Bates, Microsemi

Focused Markets and Applications
Solving the Difficult Problems

Aerospace and Defense *26% of Revenue*	Communications *38% of Revenue*	Data Center *21% of Revenue*	Industrial *15% of Revenue*
Military Communication	Wireless Backhaul	Hyperscale and Enterprise	Industrial Control
Information Assurance	Base Station	Servers	Medical Devices
Engine Control	Routing and Switching	Storage Systems	Energy
Avionics	Networking	Rack Disaggregation	Automotive
Electronic Actuation	Access and CPE	NVM Solutions	SEMI Cap Equipment
Bus and Payload Electronics	Wireless Terminal	Security	Machine to Machine
Launch Systems	Fiber Optic Backhaul		

Based on FQ4 2016 revenue

Persistent Memory (PM)



Low Latency

Memory Semantics

Storage Features

Throughput easy; latency hard

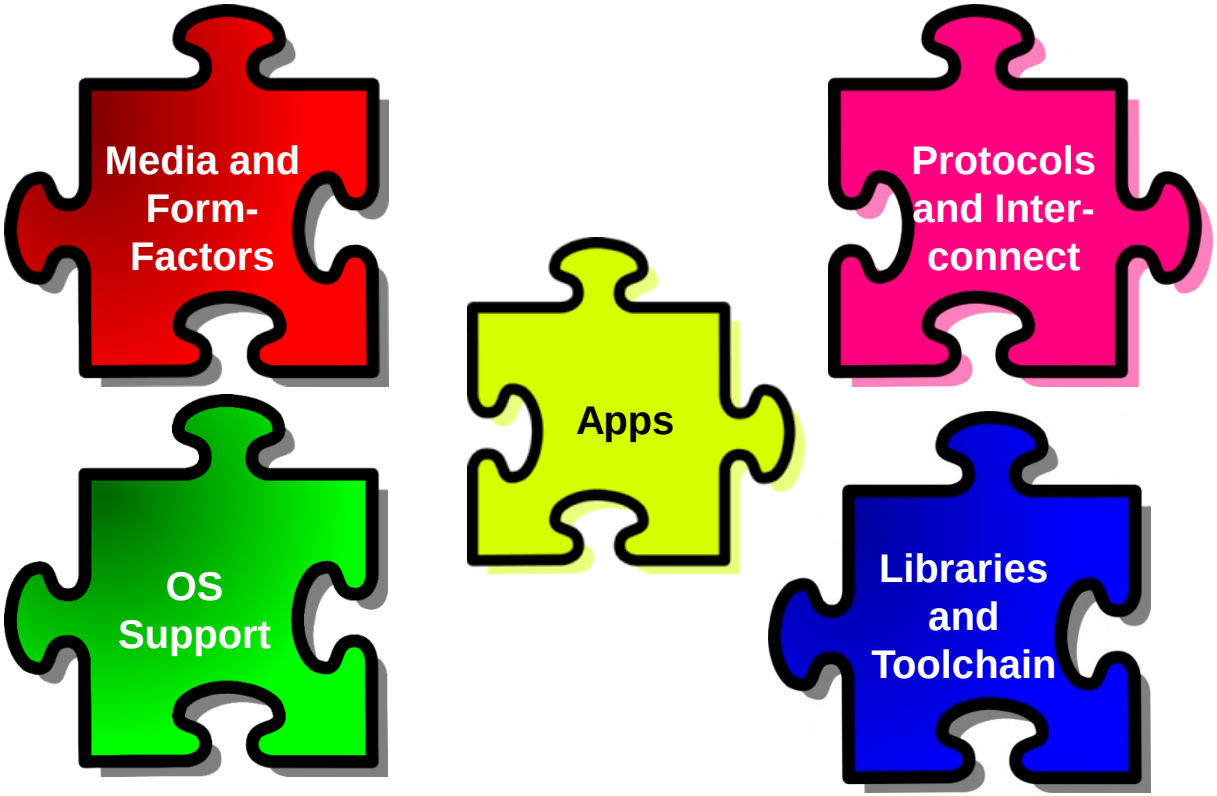
Throughput is easy

Latency is hard

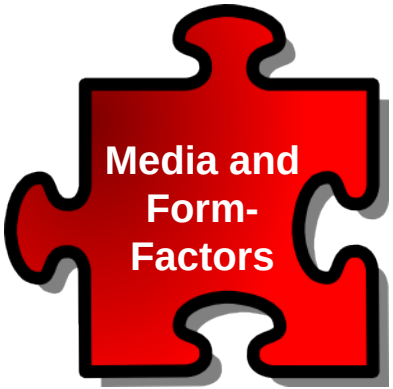
Throughput is an engineering problem; latency is a physics problem!

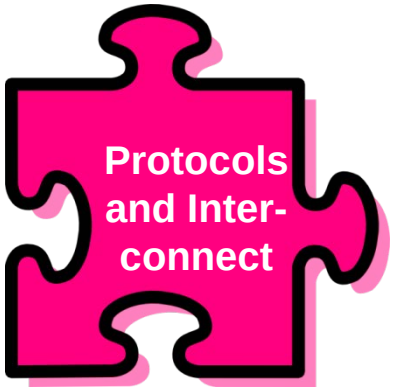
Where Are We?

What is Needed?



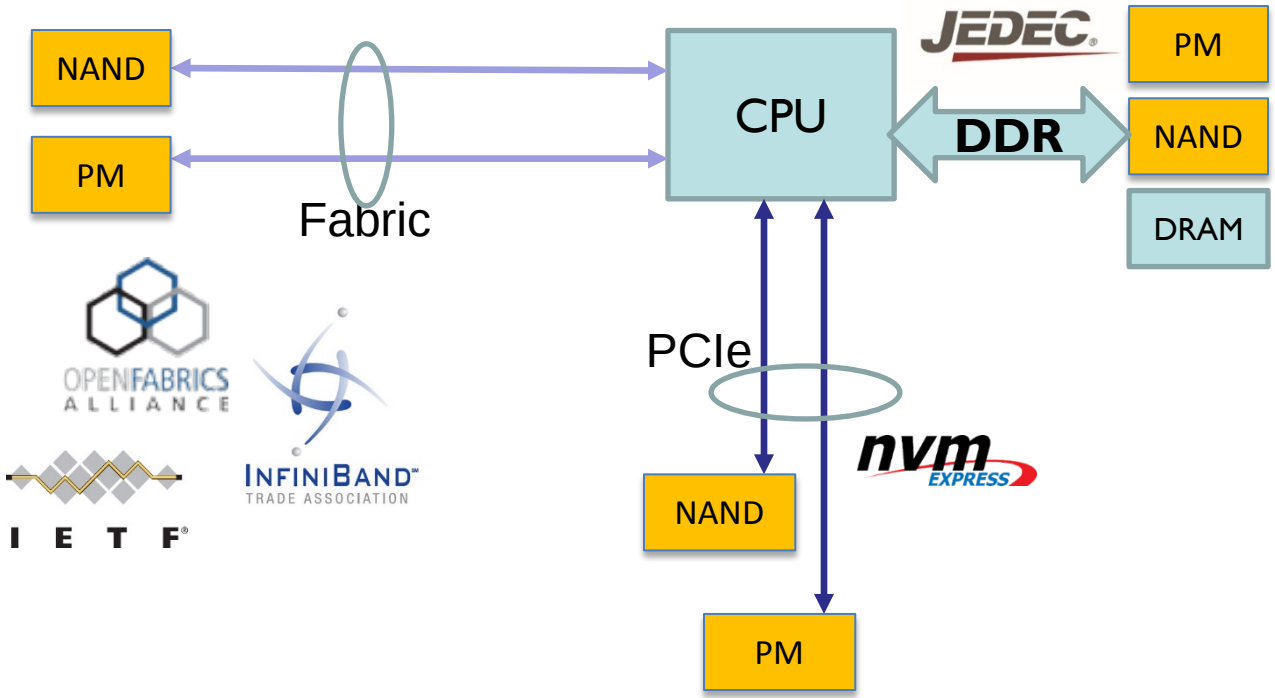
What is Needed?





Where does PM sit?

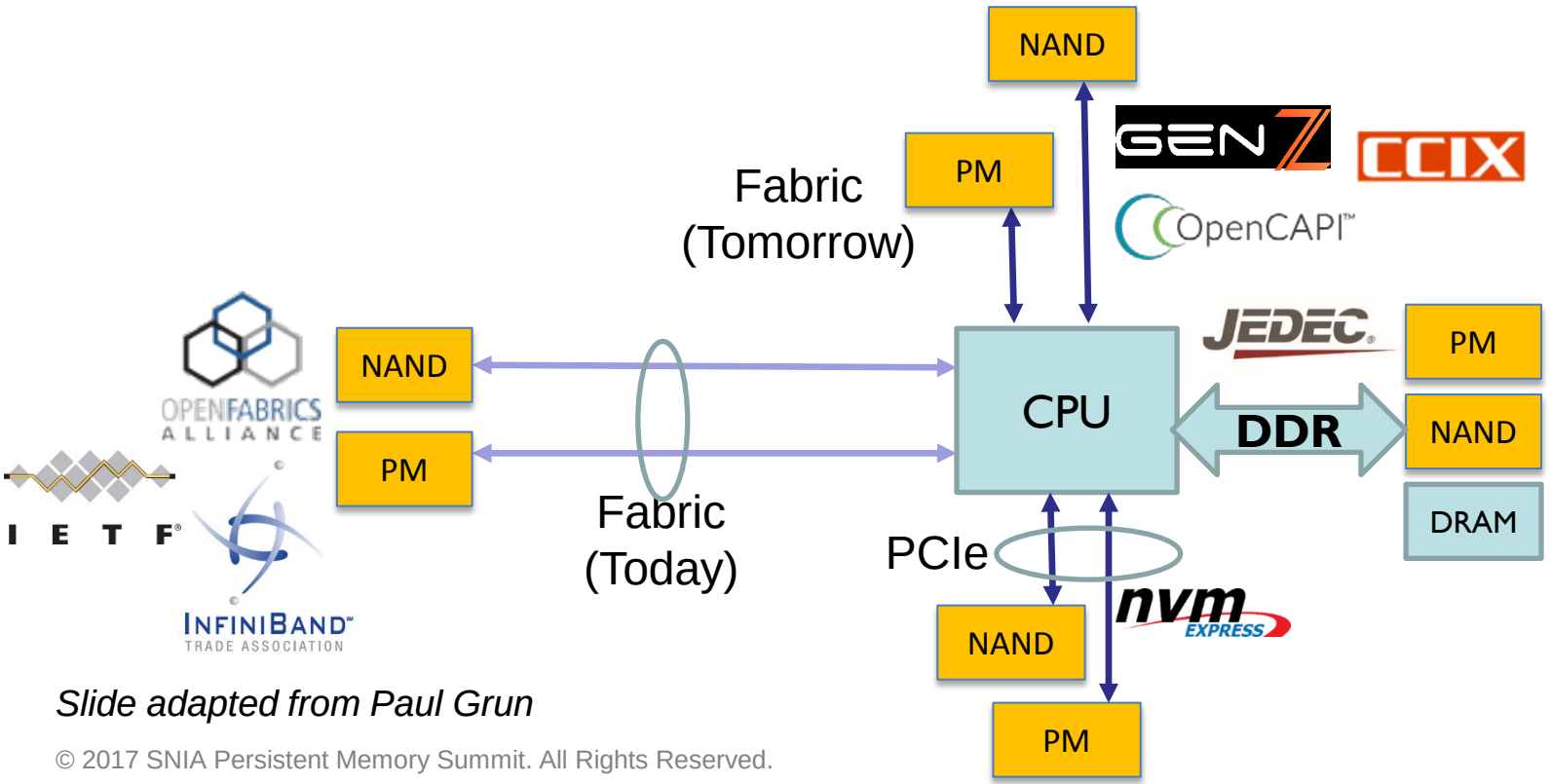
(Answer – anywhere it wants to)



Slide adapted from Paul Grun

Where does PM sit?

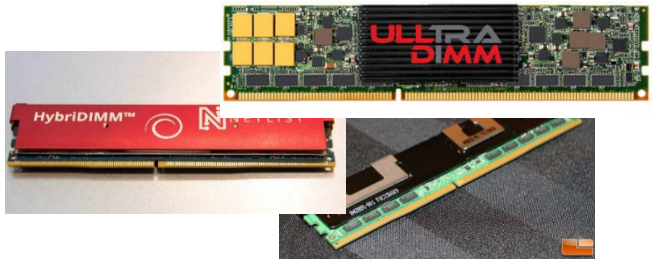
(Answer – anywhere it wants to)



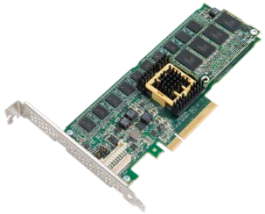
Slide adapted from Paul Grun

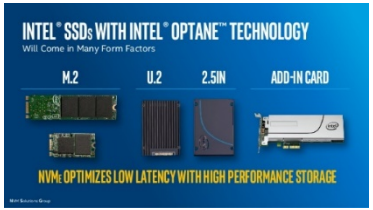
PM Form Factors

NVDIMM-N



NVDIMM-P

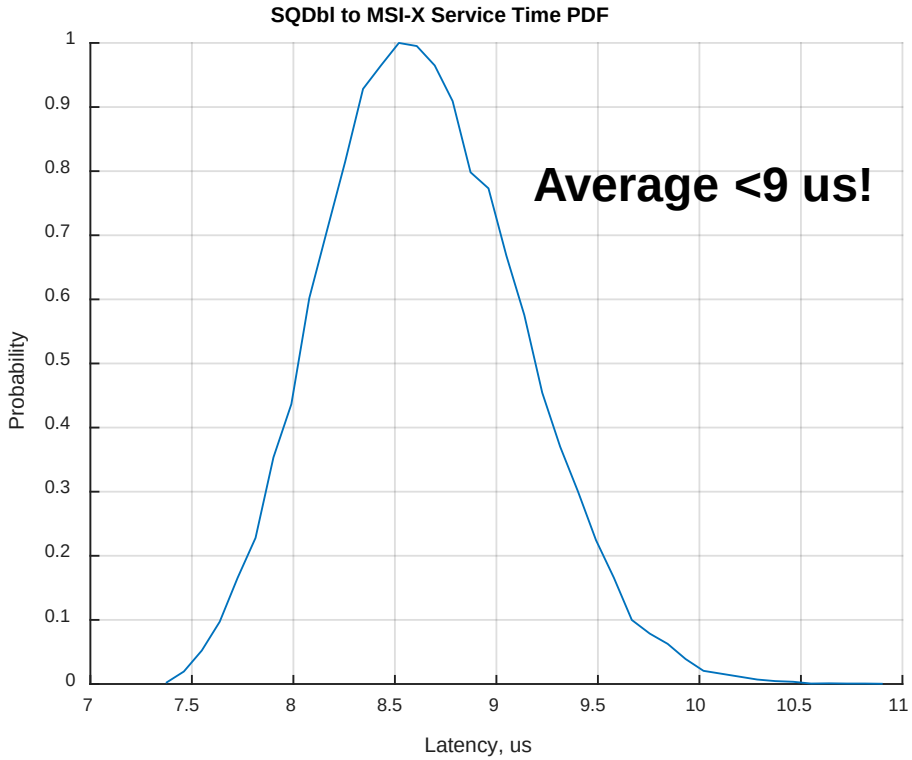




Not-NAND NVMe

NAND NVMe

NVMe Latency



NVMe is fast but not PM fast (nor byte addressable, nor coherent).

NVMe QoS is pretty good in the system we tested.

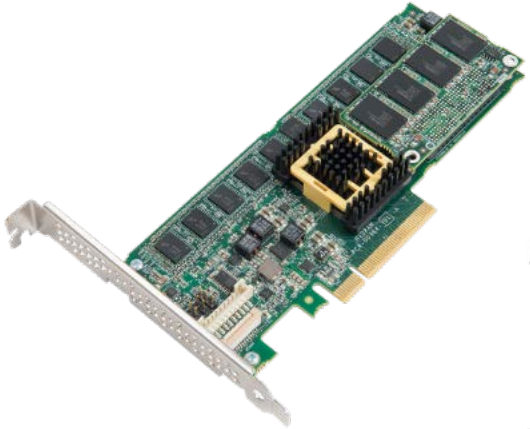
Device	Average	P99
/dev/nullb0	3.9us	5.3us
/dev/pmem0	3.31us	6.2us
/dev/nvme0n1	12us	18.5us

PM Form Factors

Form-Factor	Media	Latency	Memory Semantics	Storage Features
NVDIMM-N	DRAM/ MRAM			
NVDIMM-P	NAND/ PM			
Non-NAND NVMe	DRAM/ PM			
NAND NVMe	NAND			

Form factors impact features
(No DMA engines on a DIMM!)

Blucky[1] or Awesome



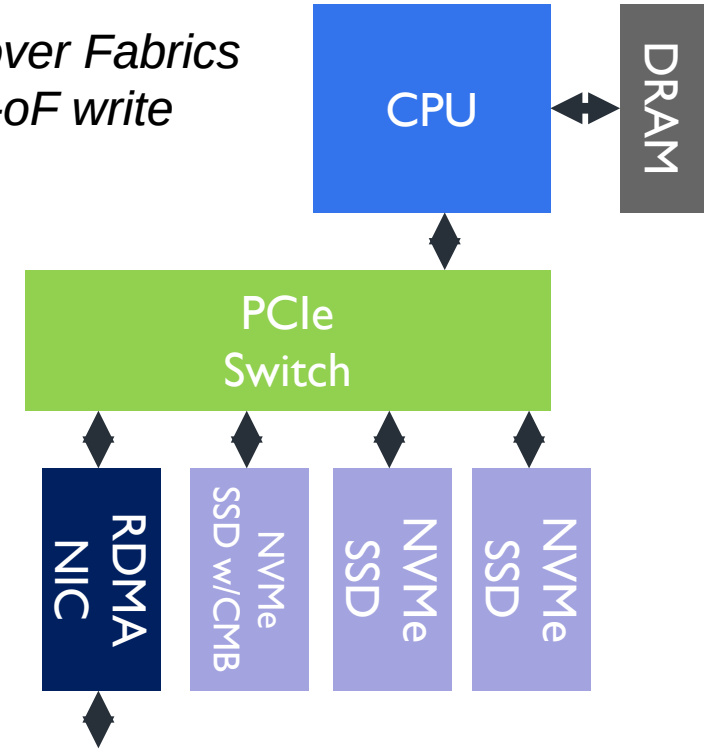
Microsemi

EVERSPIN TECHNOLOGIES
The MRAM Company™

[1] *Blucky = Blah+Yucky – coined by Oisin Bates*

NVDIMM-PCIe aka IOPMEM and P2P

e.g. PM over Fabrics or NVMe-oF write cache!



- RDMA NIC can push data direct to one NVMe w/CMBs. This SSD works as a write-back cache.
- Data is then lazily copied out of the NVMe SSD w/CMB onto standard NVMe SSDs.
- Avoids the need for all SSDs to be CMB enabled (cost reduction).

See SDC2016 Paper for details!

Coherency

Coherency

~~War~~ what's it good for?

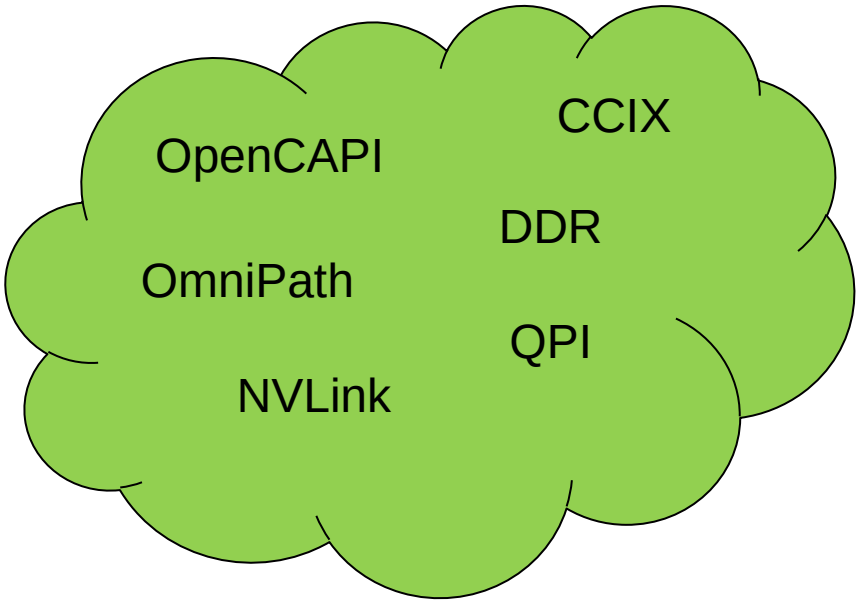
Absolutely ~~nothing~~. **Loads[1]**.

Say it again.

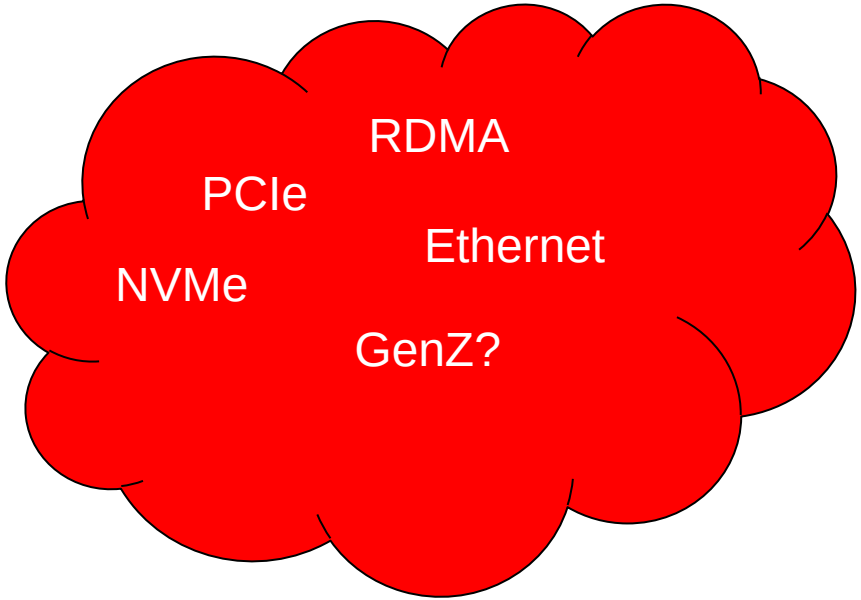
~ Borrowed from Edwin Starr ~

[1].Virtual addressing, simple (no) driver,
shared memory across CPUs and IO devices,
no DMA setup, no get_user_pages()mappings etc…

Coherency



Coherent!



Not Coherent!

Call to Arms

Lots to do – Sisyphean?

Very Excited!

A Final Thought

64 bits ~ 18 EB

180ZB[1] ~ 73 bits

[1] IDC estimate of new data in 2025