Beyond NVDIMM: Future Interfaces for Persistent Memory

Stephen Bates, Microsemi
Focused Markets and Applications
Solving the Difficult Problems

Aerospace and Defense
26% of Revenue
Military Communication
Information Assurance
Engine Control
Avionics
Electronic Actuation
Bus and Payload
Electronics
Launch Systems

Communications
38% of Revenue
Wireless Backhaul
Base Station
Routing and Switching
Networking
Access and CPE
Wireless Terminal
Fiber Optic Backhaul

Data Center
21% of Revenue
Hyperscale and Enterprise
Servers
Storage Systems
Rack Disaggregation
NVM Solutions
Security

Industrial
15% of Revenue
Industrial Control
Medical Devices
Energy
Automotive
Machine to Machine

Based on FQ4 2016 revenue
© 2016 Microsemi Corporation.
Persistent Memory (PM)

- Low Latency
- Memory Semantics
- Storage Features
Throughput is easy; latency hard

Throughput is an engineering problem; latency is a physics problem!
Where Are We?
What is Needed?

- Media and Form-Factors
- Protocols and Interconnect
- OS Support
- Libraries and Toolchain
- Apps
What is Needed?

- Media and Form-Factors
- Protocols and Interconnect
Where does PM sit?
(Answer – anywhere it wants to)

Slide adapted from Paul Grun

© 2017 SNIA Persistent Memory Summit. All Rights Reserved.
Where does PM sit?
(Answer – anywhere it wants to)

Slide adapted from Paul Grun

© 2017 SNIA Persistent Memory Summit. All Rights Reserved.
PM Form Factors

NVDIMM-N

NVDIMM-P

Not-NAND NVMe

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

NVMe QoS is pretty good in the system we tested.

<table>
<thead>
<tr>
<th>Device</th>
<th>Average</th>
<th>P99</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dev/nullb0</td>
<td>3.9us</td>
<td>5.3us</td>
</tr>
<tr>
<td>/dev/pmem0</td>
<td>3.31us</td>
<td>6.2us</td>
</tr>
<tr>
<td>/dev/nvme0n1</td>
<td>12us</td>
<td>18.5us</td>
</tr>
</tbody>
</table>
## PM Form Factors

<table>
<thead>
<tr>
<th>Form-Factor</th>
<th>Media</th>
<th>Latency</th>
<th>Memory Semantics</th>
<th>Storage Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVDIMM-N</td>
<td>DRAM/ MRAM</td>
<td>😊</td>
<td>😊</td>
<td>😕</td>
</tr>
<tr>
<td>NVDIMM-P</td>
<td>NAND/ PM</td>
<td>😒</td>
<td>😒</td>
<td>😒</td>
</tr>
<tr>
<td>Non-NAND NVMe</td>
<td>DRAM/ PM</td>
<td>😒</td>
<td>😒</td>
<td>😊</td>
</tr>
<tr>
<td>NAND NVMe</td>
<td>NAND</td>
<td>😒</td>
<td>😒</td>
<td>😊</td>
</tr>
</tbody>
</table>

Form factors impact features
(No DMA engines on a DIMM!)
Blucky\(^1\) or Awesome

\(^1\) Blucky = Blah+Yucky – coined by Oisin Bates
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

What’s it good for?
Absolutely nothing.
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!

OpenCAPI
OmniPath
NVLink
CCIX
DDR
QPI

Not Coherent!

RDMA
PCIe
Ethernet
NVMe
GenZ?
Lots to do – Sisyphean?

Very Excited!
A Final Thought

64 bits ~ 18 EB

180ZB\textsuperscript{1} ~ 73 bits

\textsuperscript{1} IDC estimate of new data in 2025