Future of Persistent Memory, DRAM and SSD Form Factors Aligned with New System Architectures

Arthur Sainio
Director, Product Marketing, SMART Modular Technologies
Large Datasets Need Memory that Scales

Processor memory and I/O technologies …

... are being stretched to their limits

More than 2X digital data will be created over the next five years compared to the combined amount since the advent of digital storage
(Source: IDC, Mar 2020)
Memory and Accelerator Alignment with Fabrics

Memory Controller
- L1
- L2
- LLC

Cache Coherence
- OpenCAPI
- DDR
- RDIMM
- NVDIMM
- NVDIMM-P
- DRAM

Processor
- Gen-Z
- DDR
- PCIe
- NVMe
- CPU
- NAND Flash
- PM/NVDIMM
- MRAM

Storage Controller
- NAND Flash
- NVDIMM-P
- NVDIMM-N

Pooled Memory
- 64GB/s (PCIe 4.0) – 128GB/s (PCIe 5.0)
- 25GB/s (DDR4) – 50GB/s (DDR5)

Pooled Accelerators
- 200GB/s

- More Memory Bandwidth
  - 1 Gen-Z port = DDR5 memory channels
- More I/O Bandwidth
  - 1 Gen-Z port = 3 PCs Gen4 ports

© 2021 SNIA Persistent Memory + Computational Storage Summit. All Rights Reserved.
Form Factor Migration


U.2 → Memory Expansion → E3 (SFF-TA-1009)

Source: Modified from J. Adiran, Microsoft SNIA, FMS, 2020
## Performance

<table>
<thead>
<tr>
<th></th>
<th>DDR DIMM</th>
<th>E1.S 1C (x4)</th>
<th>E1.S 2C (x8)</th>
<th>E3.S</th>
<th>AIC (x16)</th>
<th>E3 with x8 (2C)</th>
<th>Across network</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current Generation</strong></td>
<td>DDR4@3200 25.6GB/s</td>
<td>PCIe-Gen4-x4 7.8GB/s</td>
<td>PCIe-Gen4-x8 15.7GB/s</td>
<td>PCIe-Gen4-x16 31.5GB/s</td>
<td>OMI 25.6 GB/s  8 lanes</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Future Generation</strong></td>
<td>DDR5@4800 63.0GB/s</td>
<td>PCIe-Gen5-x4 15.7GB/s</td>
<td>PCIe-Gen5-x8 31.5GB/s</td>
<td>PCIe-Gen5-x16 63.0GB/s</td>
<td>DDR5 DDIMM (TBD) GB/s</td>
<td></td>
<td>Gen-Z, NVMe-oF (TBD)</td>
</tr>
</tbody>
</table>

* Source(s): [https://en.wikipedia.org/wiki/PCI_Express#History_and_revisions](https://en.wikipedia.org/wiki/PCI_Express#History_and_revisions)  

Latency

Direct attached (Parallel Bus) 100’s of GB
- DIMM
- NVDIMM-N, NVDIMM-P

Serial attached and PCIe attached 100’s of GB to TB’s
- DDIMM (40ns latency)
- Cache Coherent AIC
- XMM, NV-XMM
- OpenCAPI
- CCIX

Network Attached TB’s to PB’s
- E1.S
- ZMM
- GENZ

<100ns  <350ns  <800ns

© 2021 SNIA Persistent Memory + Computational Storage Summit. All Rights Reserved.
Form Factors

Direct attached (Parallel Bus) 100’s of GB

- DDR4 DIMM
  - E1.S (x4)
  - E1.S (x8)
  - E3.S (x16)

- NVDIMM-N, NVDIMM-P

Serial attached and PCIe attached 100’s of GB to TB’s

Network Attached TB’s to PB’s

<table>
<thead>
<tr>
<th>Form Factors</th>
<th>E1.S with x4 (1C)</th>
<th>E1.S with x8 (2C)</th>
<th>E3.S with x16 (4C)</th>
<th>E3 with x8 (2C)</th>
<th>Network Card</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pins</strong></td>
<td>288 pins (64 data, 87 sideband, rest power)</td>
<td>56 pins (16 diff-data, 16 sideband, 24 power)</td>
<td>84 pins (32 diff-data, 18 sideband, 34 power)</td>
<td>140 pins (64 diff-data, 24 sideband, 52 power)</td>
<td>84 pins (32 diff-data, 18 sideband, 34 power)</td>
</tr>
<tr>
<td><strong>Connect or (LxW)</strong></td>
<td>142.0mm x 6.5mm</td>
<td>23.8mm x 6.0mm</td>
<td>35.6mm x 6.0mm</td>
<td>57.0mm x 6.0mm</td>
<td>35.6mm x 6.0mm</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td>Input voltage=1.2V VPP 2.5</td>
<td>Input voltage 12V Vaux 3.3 (optional)</td>
<td>Input voltage 12V Vaux 3.3 (optional)</td>
<td>Input voltage 12V Vaux=3.3 (optional)</td>
<td>Input voltage 12V</td>
</tr>
</tbody>
</table>

Input voltage 12V
Vaux 3.3 (optional)

Vendor specific
Bandwidth

Current Generation configuration
Dual socket server with 64 core CPU and 12 x DDR4 channels populated with 128GB DDR4 DIMMs in 2 DIMMs/Channel

- Total Memory = 3TB per server
- Theoretical maximum bandwidth of 614GB/s
- Bandwidth per core = 4.79GB/s

Future Generation configuration – 2022
Dual socket server with 96 core CPU and 12 x DDR5 channels populated with 128GB DDR5 DIMMs in 2 DIMMs/Channel

- Total Memory = 3TB (DDR5) + 2TB (CXL) per server
- Theoretical maximum bandwidth of 768GB/s (DDR5)+ 252GB/s (CXL)
- Bandwidth per core = 5.3GB/s (higher even with more CPU cores)

Source: https://www.tomshardware.com/news/what-we-know-ddr5-ram,39079.html
E1.S and E1.L for Memory Acceleration and Expansion

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
</table>
| Host Interface           | - Data: PCIe x4,x8  
                          - Sideband: SMBus (I2C)  
                          - Wake-up, Low-power (PWRDIS), ... |
| Memory                   | 64-128GB with DDR4 or DDR5 |
| Protocols                | NVMe, CXL, CCIX, Gen-Z |
| Power                    | - Multiple profiles from 12,16, 20, 25W  
                          - Completely bus powered: 12V (main), 3.3V Aux  
                          - Supports low power modes (CLKREQ#, PWRDIS signaling) |
| Targeted Use Cases       | Targeted for 1U Servers  
                          - 16 – 32 Slots per 1U Server  
                          - Improves performance by offloading fixed functions like encryption, compression or Key-Value semantics to Memory modules |

Source: https://www.anandtech.com/show/16248/edsff-form-factor-updates
© 2021 SNIA Persistent Memory + Computational Storage Summit. All Rights Reserved.
E3.S and E3.L for Memory Expansion

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
</table>
| Host Interface   | • Data: PCIe x16  
                  • Sideband: SMBus (I2C)  
                  • Wake-up, Low-power (PWRDIS), ...                                                                                                       |
| Memory           | Up to 256GB with DDR4 or DDR5  
                  * Non-volatile persistent memory feature could be support on this form-factor using back-up and restore functionality like in NVDIMM-N. |
| Protocols        | NVMe, CXL, CCIX, Gen-Z                                                                                                                         |
| Power            | • 2 profiles 25W (thin) and 40W (thick)  
                  • Bus powered: 12V (main), 3.3V Aux  
                  • Supports low power modes (CLKREQ#, PWRDIS signaling)                                                                                   |
| Targeted Use Cases | Enabled for 2U Server  
                   Enables 4TB – 8 TB of Memory expansion with 16 E3.S modules in single 2U server, achieving better throughput than direct attached DDR4 DIMM. |
OpenCAPI High Bandwidth Memory - DDIMM

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host Interface</td>
<td>• OpenCAPI</td>
</tr>
<tr>
<td>Memory</td>
<td>• Up to 256GB</td>
</tr>
<tr>
<td>Protocols</td>
<td>• OMI – Open Memory Interface</td>
</tr>
<tr>
<td></td>
<td>• The memory bus is defined with one read port and one write port per channel, each having eight unidirectional differential lanes</td>
</tr>
<tr>
<td>Performance</td>
<td>• DDR4-3200</td>
</tr>
<tr>
<td></td>
<td>• Latency 40ns</td>
</tr>
<tr>
<td></td>
<td>• Data throughput rate of 25.6GB/s with 8 lanes</td>
</tr>
<tr>
<td></td>
<td>• The DDIMM/OMI approach delivers up to 4TB of memory on a server at about 320GB/second or 512GB at up to 650GB/s sustained rates.</td>
</tr>
<tr>
<td>Targeted Use Cases</td>
<td>• Targeted for servers</td>
</tr>
<tr>
<td></td>
<td>• High bandwidth, low latency serial connection for memory, accelerators, network, storage, and other devices like ASICs</td>
</tr>
</tbody>
</table>

© 2021 SNIA Persistent Memory + Computational Storage Summit. All Rights Reserved.
NVDIMM for Persistent Memory

Key Features of DDR4 and DDR5 NVDIMM-N
• Operation like DRAM
• Fast recovery from system power loss
• Software overhead can be eliminated

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host Interface</td>
<td>• DDR</td>
</tr>
<tr>
<td>Memory</td>
<td>• DDR4 16GB, 32GB</td>
</tr>
<tr>
<td></td>
<td>• DDR5 32GB, 64GB</td>
</tr>
<tr>
<td>Protocol</td>
<td>• JEDEC Compliant DDR4 / DDR5</td>
</tr>
<tr>
<td>Features</td>
<td>• Throughput of 25.6GB/s (DDR4)</td>
</tr>
<tr>
<td></td>
<td>• Latency ~20ns</td>
</tr>
<tr>
<td></td>
<td>• AES 256 bit Encryption</td>
</tr>
<tr>
<td>Targeted Use Cases</td>
<td>• All Flash Arrays, Storage Servers, HPC, AI Training Servers</td>
</tr>
<tr>
<td></td>
<td>• Needed for very low latency tiering, caching, write buffering, metadata storage, checkpointing</td>
</tr>
<tr>
<td></td>
<td>• Needed for AI/ML algorithm processing</td>
</tr>
</tbody>
</table>
CXL-based NVDIMM (NV-XMM)

High-speed DRAM with built-in back-up power to back-up data to on-module Flash during power loss.

Conclusions
Future of Persistent Memory, DRAM and SSD Form Factors

NVDIMM-N, NVDIMM-P (Persistent)
DIMM

PCIe AIC (x16)
OpenCAPI DDIMM


Scalability, Serviceability, Savings
Modularize the Technology

Dependence on CPU

2020 2022 2025
Thank you

Please visit snia.org/pm-summit for presentations