The Composable Platform: Enabling Innovation Through Flexible Infrastructure

Andrew Dieckmann, VP Marketing & Applications
Data Center Solutions Division
one size fits
NONE
Challenge:
- Data centers have the formidable task of improving operating efficiency and maximizing their IT investments in hardware infrastructure in the face of evolving and varied application requirements.
- Data Center WW spent over $100B last year on server and storage infrastructure.
- What's in all that server and storage infrastructure.. A lot of compute.. A lot of memory.. A lot of storage.
- Accelerators and GPUs; The data center accelerator market in 2018 was $2.8B; expected to grow at a CAGR of >50% over next 5 years to eclipse over $20M/y 2023.
- A LOT of money is going into accelerators.. FPGAs, GPUs and the like..
- One of the challenges with accelerators is that the amount each workload can have a different optimum accelerator to compute and storage ratio.

Memory Spend
- $10's of Billion in DRAM is spent every year in the data center.
- By some estimates average DRAM utilization sits around 50%..
- How many Billions of Dollars of stranded DRAM do we have sitting in data centers around the world?

Storage Spend
- We all love our SSDs here at FMS.
- NVMe SSD revenue hit $XB in 2018.

- We are spending are spending $10's of Billions of dollars on underutilized: storage, memory and compute.

2018: >$140B
Accelerator Spending is Exploding

50% CAGR end-market growth; >$20B market in 2023
Accelerator Efficiency Challenges

- IO Bandwidth Constrained
- Workload-dependent optimizations for resources
Accelerator Efficiency Challenges

- IO Bandwidth Constrained
- Workload-dependent optimizations for resources
The Cost of Memory

• > $20B in DRAM purchased by data centers in 2019…

• Much of this DRAM is.. **not being used**… stranded in machines with workloads not requiring it

• Not only a capex problem; DRAM consumes 15-20% of the data center power!
The Cost of Memory

• > $20B in DRAM purchased by data centers in 2019…

• Much of this DRAM is.. **not being used**… stranded in machines with workloads not requiring it

• Not only a capex problem; DRAM consumes 15-20% of the data center power!
Memory Bandwidth & Latency Bottlenecks

CPU DRAM buses are limited in quantity and performance. Multi-core processor memory latency is increasing.

https://blog.dell EMC.com/en-us/memory-centric-architecture-vision/
Many Types of Storage Media

- PCIe Cards
- U.2 & U.3
- Enterprise 3.5” HDD
- EDSFF
- M.2
Many Types of Storage Media

- PCIe Cards
- U.2 & U.3
- EDSFF
- M.2
- Enterprise 3.5” HDD
Many Types of Storage Media

- PCIe Cards
- U.2 & U.3
- EDSFF
- M.2
- Enterprise 3.5” HDD
Many Types of Storage Media

- PCIe Cards
- U.2 & U.3
- EDSFF
- M.2
- Enterprise 3.5” HDD
Drive Capacities are Growing

32TB+ in the palm of your hand

But where do you put it?
1. We are wasting Billions of dollars on under-utilized storage, memory and compute resources

2. Rapid innovation in the data center requires flexible infrastructure
1. We are wasting Billions of dollars on under-utilized storage, memory and compute resources.

2. Rapid innovation in the data center requires flexible infrastructure.
1. We are wasting Billons of dollars on under-utilized storage, memory and compute resources

2. Rapid innovation in the data center requires flexible infrastructure
1. We are wasting Billons of dollars on **under-utilized** storage, memory and compute resources

2. Rapid innovation in the data center requires **flexible infrastructure**
Stop the Madness!!
There must be a better way!
Agile Infrastructure

Composable ↔ Flexible
Agile Infrastructure

Composable ↔ Flexible
Innovation Required for Agile Infrastructure

- **Composable** storage accelerators & memory
- Optimized resources by workload
- No resource stranding
- Remove BW bottlenecks
Innovation Required for Agile Infrastructure

• **Composable** storage accelerators & memory
• Optimized resources by workload
• No resource stranding
• Remove BW bottlenecks

• **Adaptable** memory and storage technology
• Memory bandwidth scaling
Increased GPU Utilization

Switche\textsuperscript{TM} PAX Advanced Fabric Switches enable Composable Heterogeneous Computing

- Scalable non-hierarchical fabric
- Dynamic end-point allocation
- Low latency data transfers
- ...using standard drivers
Switchtec™ PCIe Fabrics

The Switchtec PCIe fabric is built on the concept of virtual domains
Switchtec™ PCIe Fabrics

The Switchtec PCIe fabric is built on the concept of virtual domains.
Switchtec™ PCIe Fabrics

The Switchtec PCIe fabric is built on the concept of virtual domains.
Increased GPU Utilization with PCIe Fabrics

- Embedded CPUs in each fabric element are the virtualizers
- Each virtual domain is a PCIe-compliant virtual switch
Increased GPU Utilization with PCIe Fabrics

- Data is routed directly by switch hardware
- **Peer-to-peer transfers** supported through the fabric
Removing IO Bottlenecks with Switchtec™

Peer-to-peer traffic through a fabric decreases congestion due to bypass of CPU to CPU interconnect

9.5 GB/s

26 GB/s

2.5x increase in bandwidth in 2 socket system
Composable Storage with PCIe Fabrics

• PCIe Fabric model is extensible to NVMe SSDs & other PCIe endpoints
  • End points are added to the fabric just like a spec-compliant GPU
  • Storage is now a flexible resource

Switchtec™ PAX enables
Composable Heterogenous Compute & Storage
No More Stranded Storage

SR-IOV and multi-host sharing provide for new granularity in composable heterogenous compute and storage

Switchtec™ + Flashtec™ provide end-to-end Multi-Host IO Virtualization with off-the-shelf drivers

- >8 NVMe SR-IOV vendors
- Leading GPU vendors support SR-IOV
- NVMe SR-IOV standardization complete
No More Stranded Storage

SR-IOV and multi-host sharing provide for new granularity in composable heterogeneous compute and storage.

Switchtec™ + Flashtec™ provide end-to-end Multi-Host IO Virtualization with off-the-shelf drivers

- >8 NVMe SR-IOV vendors
- Leading GPU vendors support SR-IOV
- NVMe SR-IOV standardization complete
Virtualization Will Be Everywhere

- Many use cases even beyond the Data Center
- Example: Autonomous Driving
Flexibility Building Block Platforms

Hyperscale:
- Compute Servers (NVMe)
- Balanced Servers (Mix)
- Cold Storage (SAS/SATA)

Enterprise:
- Universal Bay Servers (Mix)

Many form factors, many performance points:
- U.2, M.2, EDSFF, Custom
- Performance & Mainstream
- Up to 8GB/s+ bandwidth
- Up to 200TB+ capacity
Flexibility Building Block Platforms

Hyperscale:
- Compute Servers (NVMe)
- Balanced Servers (Mix)
- Cold Storage (SAS/SATA)

Enterprise:
- Universal Bay Servers (Mix)

Many form factors, many performance points:
- U.2, M.2, EDSFF, Custom
- Performance & Mainstream
- Up to 8GB/s+ bandwidth
- Up to 200TB+ capacity
Composable Memory Infrastructure

• Near memory innovation
  • Serialization of the memory

• Far memory innovation
  • Sharing pools of resources to reduce stranding

• New open load/store standards provide the low-latency connectivity required
High-Bandwidth Memory Solution – SMC 1000

Microchip Enters Memory Infrastructure Market with Serial Memory Controller for High-performance Data Center Computing

SMC 1000 8x25G enables high memory bandwidth required by next-generation CPUs and SoCs for AI and machine learning
SMC 1000 Smart Memory Controller

8x25G Open Memory Interface (OMI) Serial DDR4 Smart Memory Controller

**INCREASED MEMORY BANDWIDTH**
Enables 4x memory channels vs. x72 DDR4

**MEDIA INDEPENDENCE**
Single OMI interface provides for multiple media types

**LOWER SOLUTION COSTS**
Reduced silicon, IP and package costs for CPUs and SoCs
OMI-Enabled DDIMMs are Here!

The SMC 1000 8x25G available on standards-based DDIMMs in 1U and 2U:

1U DDIMM

Traditional RDIMM

Available from Micron, Samsung Electronics, and SMART Modular
Come visit us at Booth 601!
See a Live Demo of the Future of Memory

New Smart Memory Controller Breaks Through the Memory Bandwidth Bottleneck

CPU → Parallel Memory Bus → RDIMM → CPU → Serial Memory Bus → DDIMM
Live Demo: 24G SAS + Dynamic Channel Multiplexing

Unleash the Bandwidth of PCIe Gen 4 Infrastructure

16/32 GBps

>20 GBps

9.6 GBps → 19.2 GBps

x8 SASE-3 → x8 24G SAS

xN SATA

HDD

xN SAS

SSD
Boost GPU to Storage Transfer Rates with PAX

7x

Increase in GPU to Storage Transfer Rates
Machine Learning Benefits from Shared Storage

Machine Inference with Switchelec™ and SR-IOV enabled NVMe SSDs
Flashtec™ PCIe Gen 4 NVMe Controller

The world’s highest performance and most flexible PCIe Gen 4 NVMe SSD controller family
Microchip Data Center Solutions

Connecting, managing and securing the world’s information

Storage  Memory  Compute