FPGAs: The Key to Accelerating High-Speed Storage Systems

Salil Raje
Executive Vice President & GM
Xilinx Data Center Business
SSDs Have Been a Game Changer for Storage

![Graph showing latency comparison between HDD, SSD, and SCM. HDD has the highest latency, SSD has lower latency, and SCM has the lowest latency.]
Explosion of Unstructured Data
Continuously Evolving Standards

Data Filtering
- Hadoop
- Spark
- Aerospike
- RocksDB
- Cassandra
- Foundation DB

Compression
- GZip
- zSTD
- Huffman
- LZ
- Zipline
- Brotli

Decompression
- LZ
- Brotli
- Zipline

Encryption
- DES
- AES-XST
- SHA1-256
- Block chain
Bottlenecks Remain for Data Intensive Applications

Processor-centric architecture

- CPU not optimized for these tasks
- Excessive data transfers
- High latency
- Limited BW

- Controller
- Flash
- PCIe
- DRAM
- Compute
- CPU
Emergence of Computational Storage as the Solution

Computational storage architecture

- More available CPU cycles
- Reduces required bandwidth
- Reduces latency
- Compute acceleration close to storage

Controller

CPU

DRAM

Flash

PCle
Growing Industry Momentum for Computational Storage

Participating Companies

- EIDETICOM
- NGD systems
- ScaleFlux™
- arm
- CALYPSO Systems
- GIGAIO
- inspur
- KALRAY
- Lenovo
- MARVELL®
- Micron
- NetApp®
- NETINT
- NYRIAD®
- SK hynix
- SAMSUNG
- TOSHIBA
- WD Western Digital
- XILINX®
How FPGAs Address the Computational Storage Problem
FPGAs in Storage Today

- Flash controllers
- Storage Systems
  - Cache-offload
  - Storage System & Switching connectivity
  - Data Reduction
FPGA Advantages for Computational Storage

- Flexible, fully customizable architecture adapts to specific applications
  - Massive parallelism, I/O and customizable data path
- Performance, power and latency of dedicated HW + reconfigurability of SW
- More economical than ASIC/ASSP for many applications
FPGA Advantages for Changing Standards

Architecture easily adapts to latest compression algorithms

Gzip Accelerator

Brotli Accelerator

Zipline Accelerator
Example of Analytics Acceleration

Q1: “Which cities originate the most flights with >10min delays?
Q2: “Which airport in the Bay Area has the worst record?

Airline traffic in the USA from 1970 to Present
Flight Data — 1.2B Entries
Airport Data — 500M Entries
Planes Data — 700M Entries

Realtime processing on Data Lake

FPGA
SSD Ctrl

Scan, filter, Hash-Agg

Databases
Spark

Computational Storage Drives

QUERY PERFORMANCE

Relative Performance

<table>
<thead>
<tr>
<th># FPGA Accelerators</th>
<th>Relative Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>1x</td>
</tr>
<tr>
<td>1</td>
<td>4x</td>
</tr>
<tr>
<td>2</td>
<td>7x</td>
</tr>
<tr>
<td>4</td>
<td>13x</td>
</tr>
</tbody>
</table>

XILINX
Example of Line Rate Hadoop Compression Acceleration

The challenge: Ingest real-time retail sales data during peak shopping season

CPU can’t keep up with line-rate data ingestion making compression impractical

<table>
<thead>
<tr>
<th>FPGA</th>
<th>CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>20x</td>
<td>1x</td>
</tr>
</tbody>
</table>

Intel Skylake-SP 6152 @2.10GHz CPU (Ubuntu 16.04), 6GB/s compression per CPU core = .0229. Alveo U50 = 10GB/s
FPGA-based Data Compression Enables Server Consolidation

<table>
<thead>
<tr>
<th>Without Compression Acceleration</th>
<th>With FPGA Compression Acceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td>2x Dual CPU Servers</td>
<td>Single Socket Server</td>
</tr>
<tr>
<td>With 192TB (uncompressed)</td>
<td>2x Accelerators, 96 TB (compressed)</td>
</tr>
</tbody>
</table>

- 50% Reduction in Nodes
- 40% Lower Cost

Intel Skylake-SP 6152 @2.10GHz CPU (Ubuntu 16.04), GB/s compression per CPU core = .0229. Alveo U50 = 10GB/s, Assume 2:1 compression.
Computational Storage Deployment Options
Computational Storage Drive (CSD)

> Integrated Accelerator and Flash

> Benefits:
  > Easy to implement - plug & play
  > Adding capacity adds accelerators + performance
  > Ability to optimize BW between accelerator and flash
  > Ability to customize FTL for specific workloads

> Vendors at FMS:
  > Samsung
  > Scaleflux
Computational Storage Processor (CSP)

- **Accelerator and Storage on same PCIe subsystem**

- **Benefits:**
  - SSD vendor independence
  - Plugs into standard slot
  - PCIe peer-to-peer transfers for high bandwidth and low latency

- **Vendors at FMS:**
  - Bittware
  - Eideticom
  - Xilinx
Computational Storage Array (CSA)

- Accelerator in-line with storage
- Benefits:
  - SSD vendor independence
  - Independently scale accelerators and SSDs
  - Ability to optimize BW between accelerator and SSDs
- Vendors at FMS:
  - Bittware
Future Directions
Current Data Center Architecture:
Fixed Resources, Sub-optimal Utilization
Future Data Center: Disaggregated and Composable

Challenge: Reduced Bandwidth and Increased Latency

Ethernet

Workload 1

Workload 2

Workload 3
Introducing Composable Storage Acceleration

> Enables composability without significant performance penalty

> Benefits
  >> Performance and latency benefits of computational storage
  >> Scale compute / storage independently
  >> Higher density per rack
  >> Lowest TCO

> Vendors at FMS:
  >> Xilinx
Future DC: Composable + Adaptable Computational Storage

- Moves some compute next to the data
- Network traffic reduced
- Latency improved
- Higher utilization with composable infrastructure
Future DC: Composable + Adaptable Network Acceleration

- Enables low latency high bandwidths acceleration of network interface workloads
- Enables significantly higher packets per second
- Offloads network functions from the CPU
Future DC: Composable + Adaptive Compute Acceleration

> Customizable acceleration up to 100x faster than CPUs for:
  > Video transcoding
  > ML inferencing
  > Financial modeling
  > …
Future DC: Composable + Distributed Adaptive Acceleration

- Composable accelerated storage, networking and compute
- Optimized for each workload
- Optimal infrastructure utilization
FPGAs are Key to Accelerating High-Speed Storage Systems

Computational storage addresses a broad range of application bottlenecks

Offers data center operators >5x performance boost and up to 2x reduction of TCO

Xilinx is leading the way in distributed adaptive acceleration
Computational Storage in Action

> Visit Xilinx in booth 313

> Visit our partners

   > Alpha Data, Bittware, Burlywood, Codelucida, GigaIO, Echo Streams, Eideticom, Everspin Technologies, IP-Maker, Mobiveil, Pliops, PLDA, Scaleflux, Smart IOPS, Samsung, SMART Modular, Toshiba Memory America, Western Digital

> Visit our Computational Storage microsite

   www.xilinx.com/computational-storage

> Join SNIA working group for Computational Storage
Adaptable.
Intelligent.