



Flash Memory Summit

Computational Storage Workloads - Implications for Datacenter Architectures

Jamon Bowen

Storage Segment Marketing Director
Xilinx



Flash Memory Summit

Agenda

- What workloads map well to storage acceleration?
- How is computational storage delivered?
- Future architectures?
- Call to Action



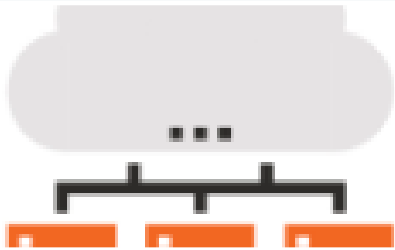
Computational Storage Motivation and Vision

- Data is
 - Big
 - Growing
 - Valuable
- Moving Data to Compute is
 - Expensive
 - Power Hungry
 - Best Minimized
- Let's move compute to storage where data resides instead!
 - Bulk of data crunching happens in storage
 - Results passed up to the CPU/network





What workloads map well to computational storage?



Storage
Services



Format
Conversion

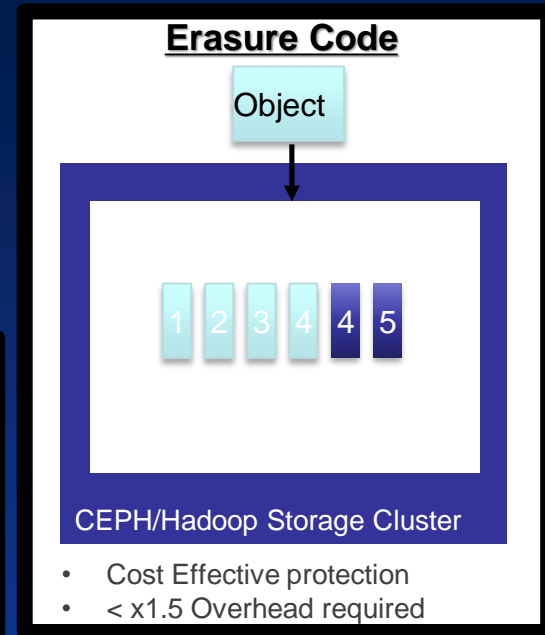
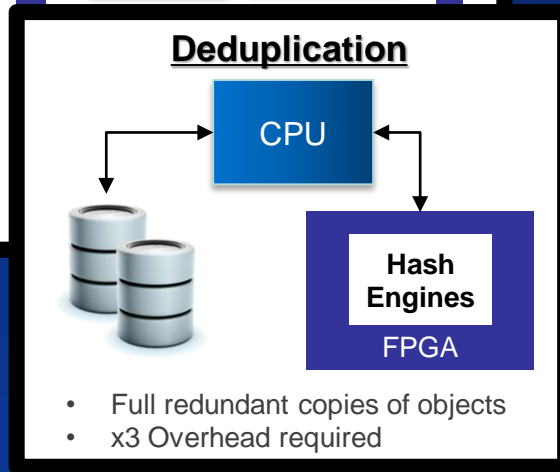
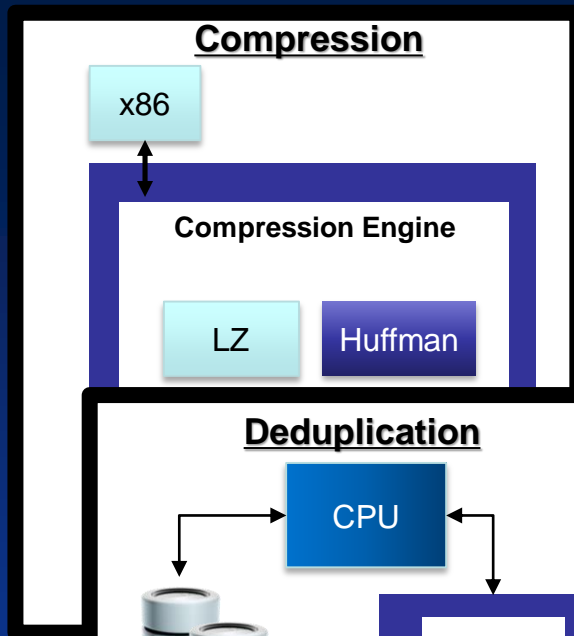
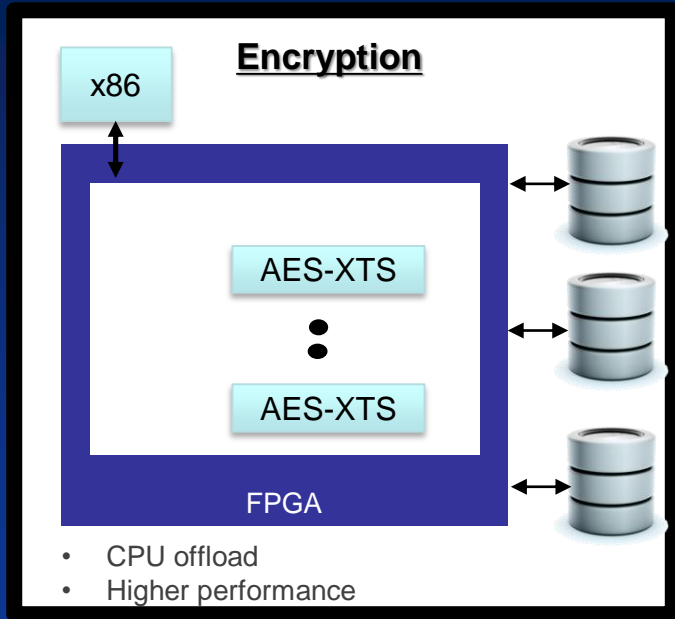


Database
Acceleration



Storage Services

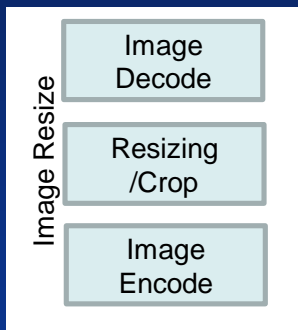
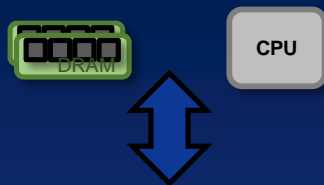
Flash Memory Summit



Functions every byte of data goes through on every access.



Format Conversion

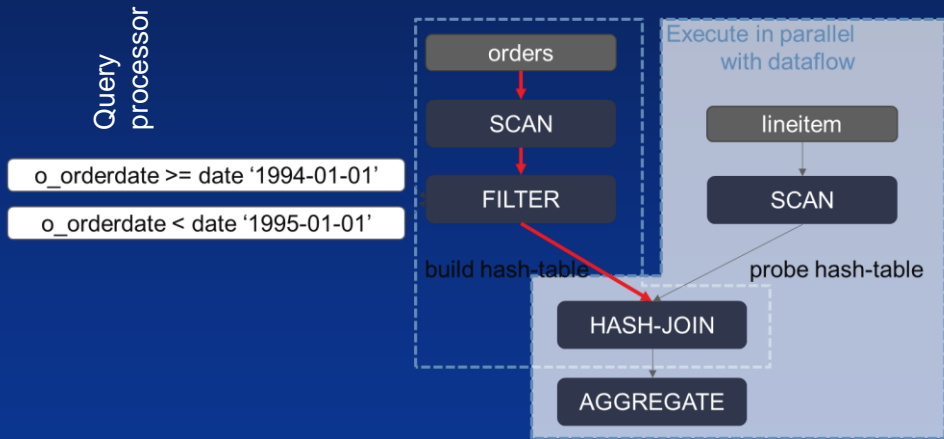


- Problem: CPU based format conversion requires both significant cpu cycles, data movement and latency.
- Example Jpeg: Image processing
 - Image Transcoding (JPEG2JPEG, JPEG2WebP, etc.)
 - Pixel Processing (Resize, Crop, etc)
 - Thumbnail Generation
 - Intelligent Analysis (Classifications)
- Solution: Push down conversion to the storage.
- Why?
 - Higher Throughput
 - Lower Latency
 - CPU offload / Space savings
 - Lower TCO



Database acceleration

```
SELECT sum(l_extendedprice * (1 - l_discount))
FROM   orders, lineitem
WHERE  l_orderkey = o_orderkey
       and o_orderdate >= date '1994-01-01' and o_orderdate < date '1995-01-01';
```



> **Problem:** Need to parse through large amount of data to find the records of interest.

> **Example:**

>> Analytics – Need the records for a time range for just one of many products included in the database.

> **Solution:** Push down Scan, Filter, Aggregate to storage.

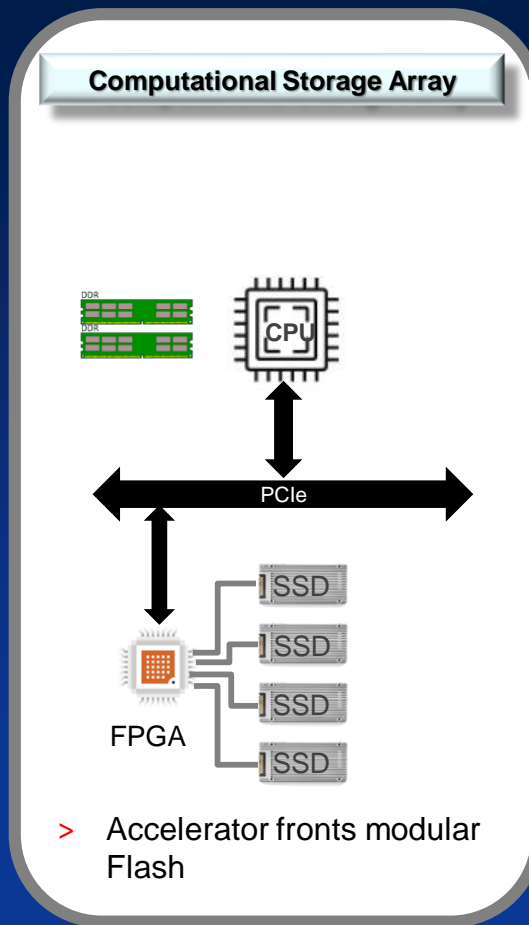
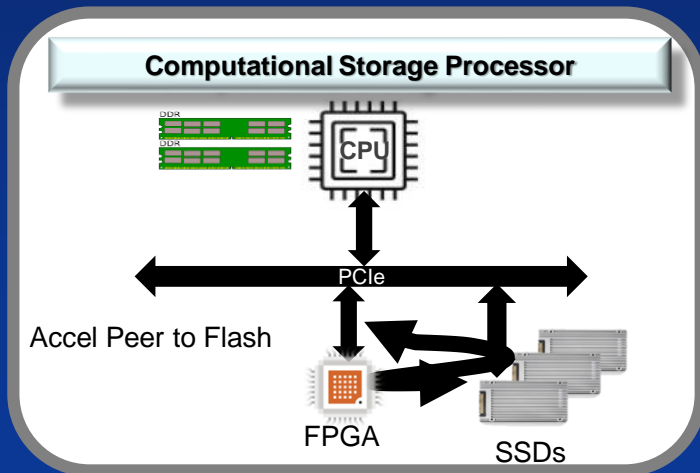
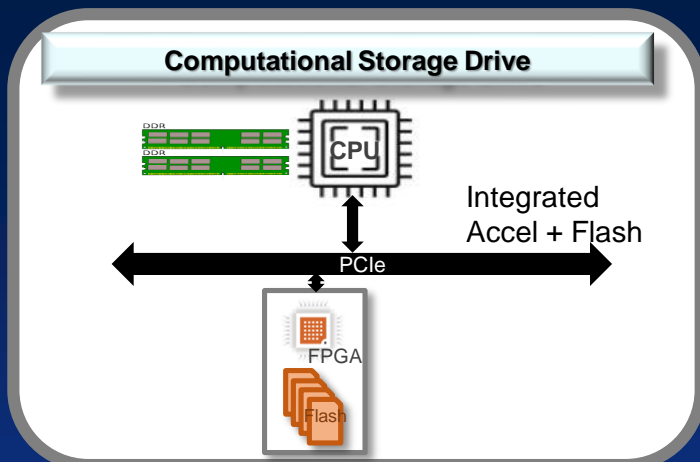
> **Why?**

- >> Higher Throughput
- >> Lower Latency
- >> CPU offload
- >> Lower TCO



Flash Memory Summit

Computational Storage Accelerator Types

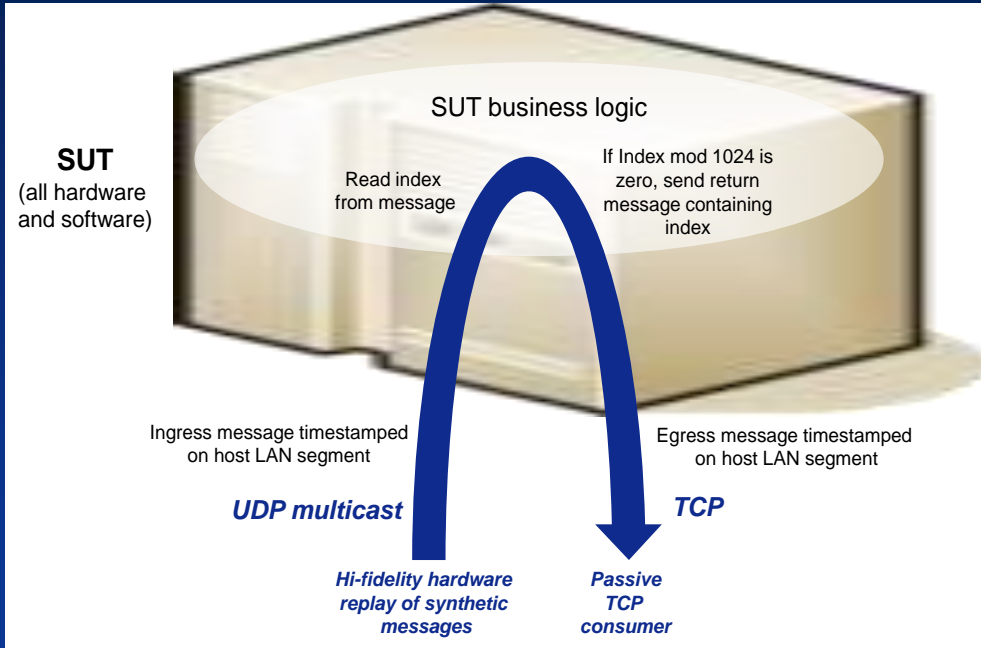




Other domains - Finance

Flash Memory Summit

Just how much latency does a network connection have to add?



STAC-T0™ Benchmarks

STAC Benchmark™ specifications for assessing tick-to-trade network I/O



Flash Memory Summit

98 ns!

The Ultimate Trading Machine

World Record 98ns Tick-to-Trade Latency Based on STAC T0 Benchmark



The new STAC T0 benchmark simulates the time for the Chicago Mercantile Exchange to emit a UDP trade data packet and for the CME to receive a TCP market order packet. That 1/2 round trip represents the minimum amount of time possible to execute a trade.

Intel® Xeon® Scalable (Skylake) Processors

Intel Xeon Gold processors offer monumental leaps in I/O, memory, storage, and network technologies.



LDA Technologies LightSpeed TCP™ Cores

An ultra-light, ultra-high-speed, and ultra-low-latency FPGA-based distributed TCP offload with processing latencies under 20ns and thousands of TCP connections.



Xilinx Kintex® UltraScale™ FGAs

Kintex UltraScale FGAs are optimized for best-in-class performance per watt fabric in 10G to 100G networking applications.

Improving tick-to-trade latency is critically important in electronic trading because it means improving the queue position of trades, which increases the probability that trades will be executed. The Ultimate Trading Machine server configured by Penguin Computing, Solarflare, LDA and XILINX achieved 98 nanosecond tick-to-trade latency in STAC T0 benchmark testing, 18% faster than the previous world record.

Solarflare XtremeScale™ Software Defined NIC

Leverages the Delegated Send™ capability of the Onload™ kernel bypass-enabled NIC--and Solarflare Application Nanosecond TCP Send (ANTS) technology--to maintain TCP connections that delivery blazingly fast network performance.



Penguin Computing Relion® Server

Optimal performance through carefully selected and vetted processors, memory, bus, storage, and other options, architected into a 1U 19" EIA traditional form-factor.

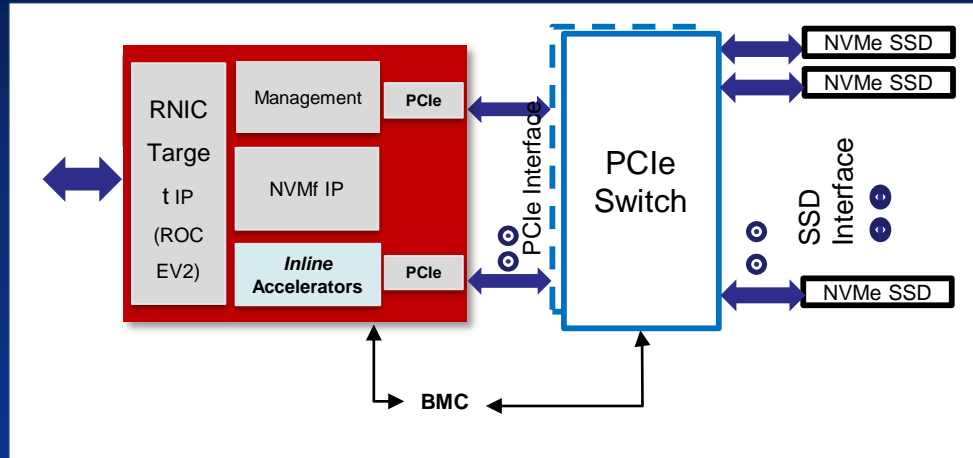


Source: Penguin Computing <https://www.penguincomputing.com/record-setting-high-frequency-trading-solution-unveiled-stac-summit/>

Benchmark Identifier <https://www.stacresearch.com/SFC170831>

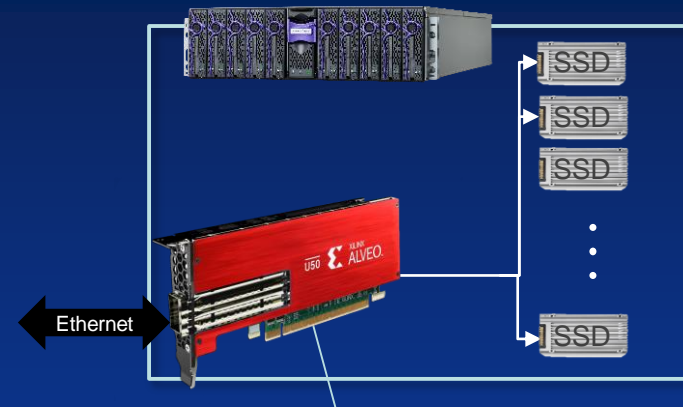


Computational Storage – Why not Fabric Attached?



- > Fabric connected accelerator fronts SSDs – brings the compute to the data.
- > NVMeoF target offloaded to U50 supporting 2.5 Million IOPS.
- > Storage accelerators Inline with NVMeoF hardware datapath.

Technology Preview: Fabric Attached Computational Storage Array



- Inline Accelerator Examples:**
- Storage services:
- (De)Compression
 - (De)Encryption
 - Data protection
- Database Acceleration:
- Scan
 - Filter
 - Aggregate



Flash Memory Summit

Conclusion

Computational storage improves performance by offloading bandwidth and reducing latency.

Visit our Computational Storage microsite:

www.xilinx.com/computational-storage

Join SNIA working group for Computational Storage!

www.snia.org/computational