Unleashing Data Driven Applications with Computational Storage

Thad Omura, EVP Marketing

Computational Storage Application Acceleration

Terabytes of Low Latency Flash

Compute Engines

Data-Driven Applications

Database   Big Data   Content Delivery   AI/ML

CSS 1000 Series
Computational Storage Subsystem™

ScaleFlux

PCI EXPRESS®

CSS 1000 Series
Computational Storage Subsystem™

ScaleFlux

Unlock Storage I/O AND Compute Bottlenecks

Responsive Performance

Affordable Scaling

Agile Infrastructure

- "Computational Storage"…with no Storage

- Complex SW interface to HW → hard to use / optimize

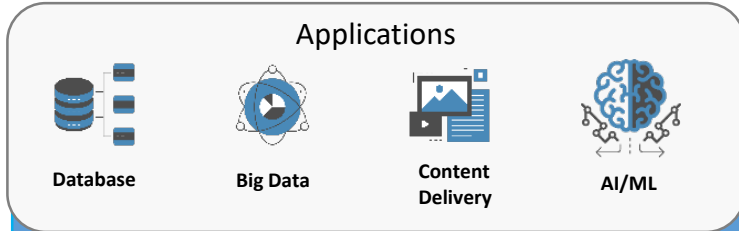- Highly specific workloads → lacking broad adoption

Requires highest expertise in Flash storage AND compute (applications)

Lay the Law Down (ScaleFlux POV)

- More storage, more compute (parallelize workloads)

- Compute MUCH faster than x86 host (5 to 100x)

- State-of-the-art storage performance / latency (Cloud)

- Tune storage I/O and compute cohesively

- Ability to quickly tune and customize (software & hardware)
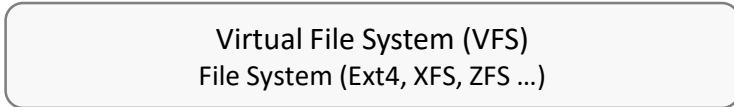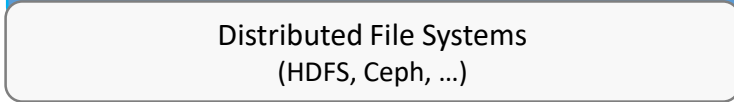
- Easy-to-use turnkey solutions

Solution Agility Across Whole Stack

Applications

Database   Big Data   Content Delivery   AI/ML

Compute Libraries / APIs

Distributed File Systems
(HDFS, Ceph, …)

Virtual File System (VFS)
File System (Ext4, XFS, ZFS …)

**Open Channel Flash Management
Computational Storage Driver**

PCIe

AIC or U.2

**FPGA
Re-programmable
Compute Engines**

Simple Interface

Easy Integration

Controllable Data Placement

Performance/QoS (latency)

Programmable HW engines

Identify the Right Workloads

INFRASTRUCTURE

STORAGE

Compression (GZIP)
Erasure Coding (RS)
Security (AES)
Authentication (SHA)
Error Checking (CRC)

PLATFORM

DATABASE, ANALYTICS

KV-Store
Transactional-Analytical
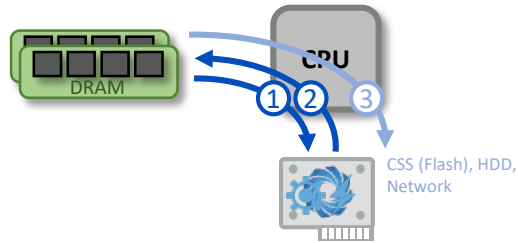SQL Processing
Big Data Analytics

APPLICATION

AI, Genomics, CDN, Search

Media Scaling & Transcoding
Neural Networks
Fuzzy Search
Filtering, Matching

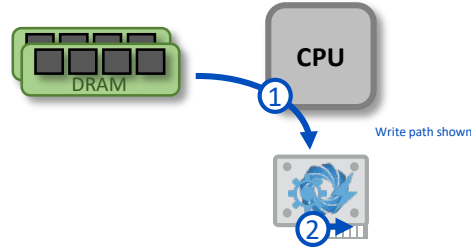# CSS Usage Models Based on Application Demand

Accelerator + Flash
Workload Starts in Memory

CSS (Flash), HDD, Network

➤ Add CSS Library for HW acceleration
➤ Examples: GZIP, EC (RS), AES, SHA, transcoding…
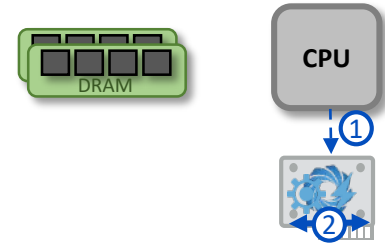
Data Path Processing
Workload Starts in Memory (Write) or CSS (Read)

Write path shown

➤ Compute while data read/written to CSS
➤ Save data movement
➤ Examples: in-line GZIP, AES, SHA, transcoding…

In-Storage Processing
Workload in CSS

➤ Compute locally on each CSS
➤ Optimize data movement
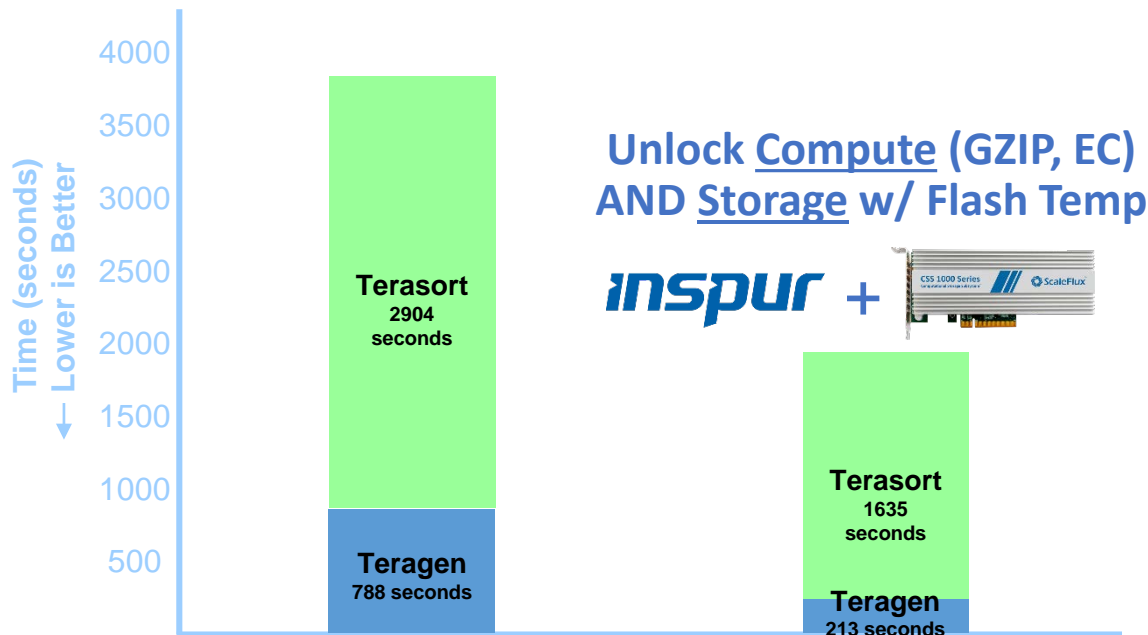➤ Examples: SQL queries, search…

- PCIe slot consolidation (Accelerator + Flash)
- Parallelize computation across multiple CSS
- In system programmability of compute engines
- Same hardware

Common CSS benefits

200% Job Throughput!

1TB Teragen + 1TB Terasort

hadoop
3.0 w/ EC (6+3)

**Unlock Compute (GZIP, EC)
AND Storage w/ Flash Temp**

inspur + CS5 1000 Series ScaleFlux

Time (seconds)
← Lower is Better

- 4000
- 3500
- 3000
- 2500
- 2000
- 1500
- 1000
- 500

**Terasort
2904
seconds**

**Teragen
788 seconds**

**Terasort
1635
seconds**

**Teragen
213 seconds**

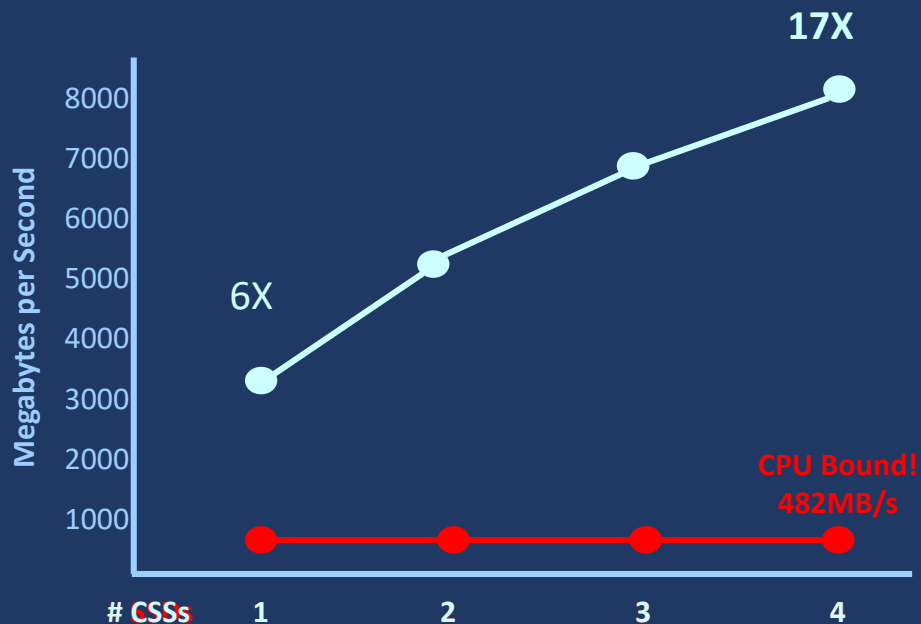CPU GZIP, EC
HDD Storage

CSS HW Accelerated GZIP, EC
HDD Storage
CSS Flash Temp

Parallelizing Computational Storage

GZIP Compression
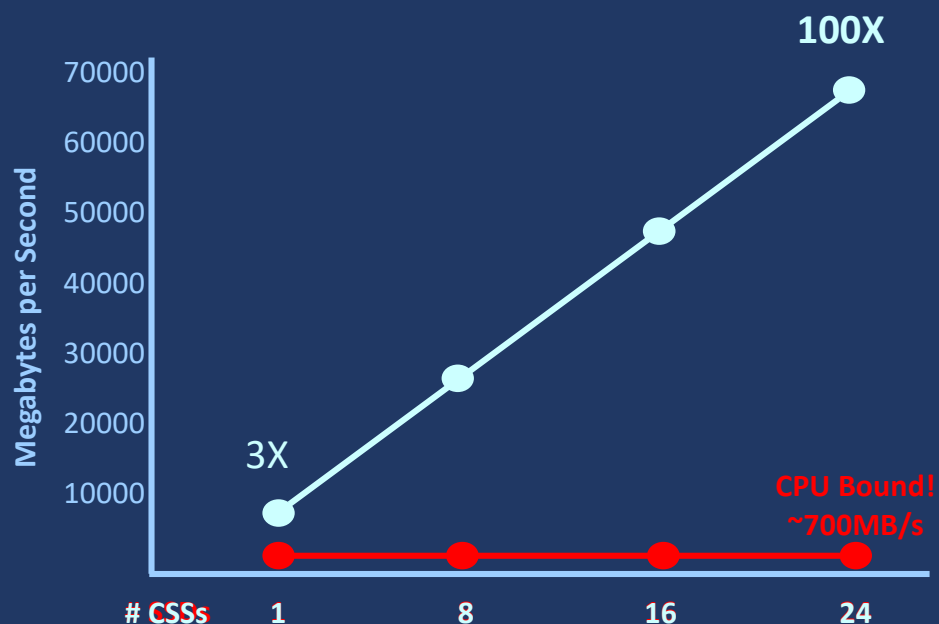(CPU zlib vs. ScaleFlux css_zlib, corpus.cantebury E5-2667v4)

17X

6X

Megabytes per Second

8000
7000
6000
5000
4000
3000
2000
1000

CPU Bound!
482MB/s

CSSs   1   2   3   4

Fuzzy Search
(POC Unindexed Text Data, Edit Distance = 8, E5-2637v3)

100X

3X

Megabytes per Second

70000
60000
50000
40000
30000
20000
10000

CPU Bound!
~700MB/s

CSSs   1   8   16   24

Proven Value for Data-Driven Applications

System-Level Application Advantage
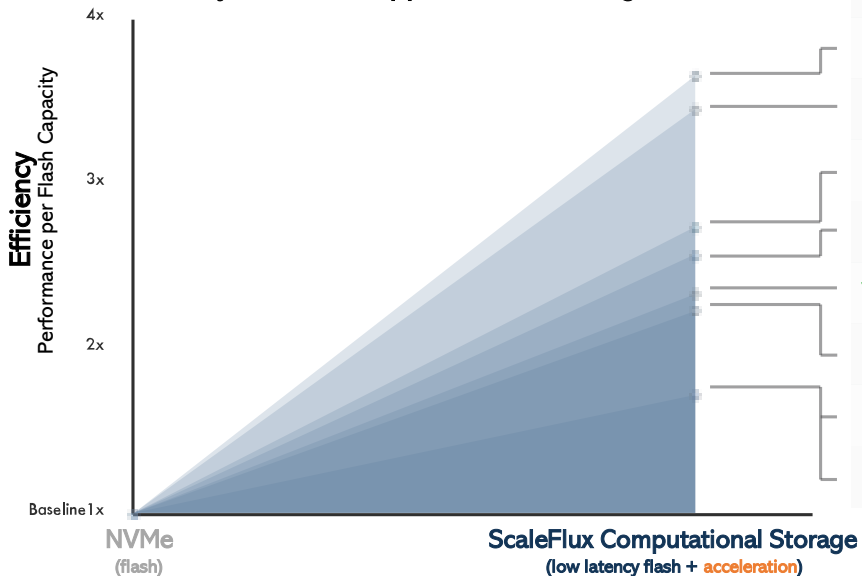
Application	Benchmark	NVMe (flash)		ScaleFlux Computational Storage (low latency flash + acceleration)			
		Performance	Capacity	Performance	Capacity	Efficiency Advantage	Acceleration
MySQL	SysBench	514 TPS	75GB	1872 TPS	76GB	3.6x	CSS GZIP
PostgreSQL OpenZFS	pgbench (50% R/W)	722 TPS	33GB	2404 TPS	31GB	3.5x	CSS GZIP
AEROSPIKE	ACT	109,500 TPS	1.6TB	300,000 TPS	1.6TB	2.7x	FTL
APACHE HBASE	YCSB Load	39k OPS	45GB	110k OPS	49GB	2.6x	CSS GZIP
VITESSE DATA DEEPGREEN DB	TPC-H (avg query time)	154 sec (lower is better)	360GB	75 sec (lower is better)	317GB	2.3x	CSS GZIP
RocksDB	RDB Random Write	8726 sec (lower is better)	1.6TB	3956 sec (lower is better)	1.6TB	2.2x	KV Store
hadoop MapReduce	Teragen & Terasort	4655 sec (lower is better)	72TB HDD	2801 sec (lower is better)	72TB HDD (+1.6TB CSS)	1.7x	CSS GZIP
Spark	Teragen & Terasort	2663 sec (lower is better)	1.6TB	1595 sec (lower is better)	1.6TB	1.7x	CSS GZIP

NVMe (flash)

ScaleFlux Computational Storage (low latency flash + acceleration)

- All benchmarks use a single NVMe/CSS card per server
- HBase, MapReduce and Spark are run on multiple servers
- MapReduce benchmarks adds NVMe/CSS card to HDD storage
- Detailed benchmark data at http://www.scaleflux.com/applications.html

Get the Most Out of Flash Storage at Same Economic Price Points

What Our Customers are Saying…

"…delivering **FANTASTIC OPERATIONS PER SECOND** for our latest NoSQL database…"

Responsive Performance

"…**INSTANTLY** saw how this can help us **COST-EFFECTIVELY** scale our infrastructure …"

Affordable Scaling

"PhonePe is leading India's digital payments revolution through super reliable payment processing infrastructure that continues to grow our transactions threefold year on year. ScaleFlux is accelerating **MULTIPLE, BUSINESS CRITICAL APPLICATIONS** for us so we can **MOST EFFICIENTLY SCALE** our low-latency, Flash storage deployment."
-- Burzin Engineer, Chief Reliability Officer

Agile Infrastructure

Computational Storage: Alibaba Cloud

- HTAP: Hybrid Transactional/Analytical Processing

- No lag for analytics, low cost, unified storage

- ScaleFlux compute @ data enables:
  - HW accelerated real-time analytical processing from transactional data
  - Intelligent data management at the storage layer
  - Massive data movement, power, and latency reduction
  - Fast time-to-value with programmable HW

Alibaba Cloud   POLARDB

"By bringing compute to the data, ScaleFlux is **transforming** the way we are architecting our **Flash storage infrastructure**.

We're looking to fully utilize the values of Computational Storage in order to **cost-effectively** scale **real-time analytics** across exploding transactional data sets, all the while delivering the **most responsive, cloud-native** user experience."

- GM of Alibaba Cloud

10X Transactional-Analytical Processing, **Half** the Flash Capacity

First Computational Storage on Cloud!

- **Packet.net**
  - Leading bare metal cloud for developers
  - Hardware optimized solutions
  - Single tenant infrastructure

- **Easy, simple, cloud access**
  - m2.xlarge configuration
  - dual Xeon Gold 5120s, 384GB DRAM
  - Two ScaleFlux CSS 1000 U.2 3.2TB drives

- **Contact** info@scaleflux.com
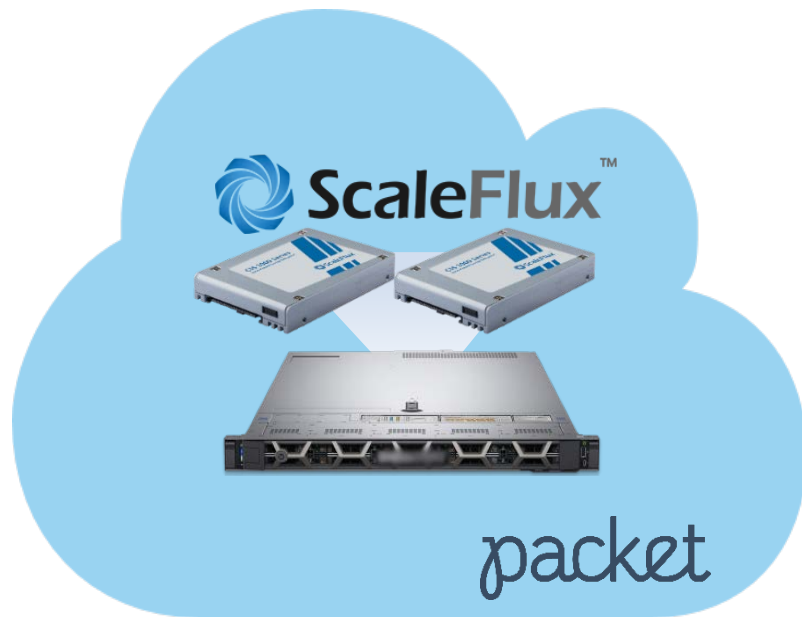
Computational Storage Summary

- Low-Latency, state-of-the-art Flash **AND** Compute Acceleration

- Proven turnkey apps **AND** customizable

- HW & SW agility to evolve quickly

- Hyperscale/Cloud, Webscale & Enterprise volume applicability

The pioneer in deploying Computational Storage at scale

Booth #113 – Check us out!

You can't go home without this!

Or without this…

Check us out in Booth #113

Thank You

97 East Brokaw Road, Suite 260
San Jose, CA 95112
www.scaleflux.com #compute2data

ScaleFlux™