Computational Storage

Using Computational Storage to Handle Big Data

Andy Walls

IBM Fellow, CTO and Chief Architect FlashSystem
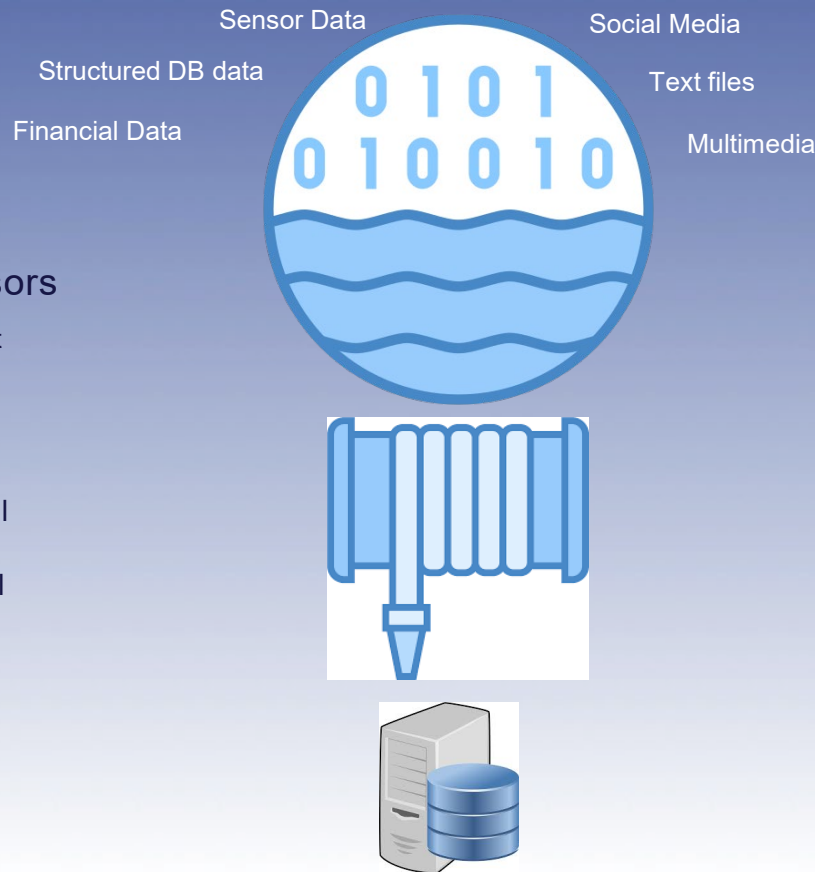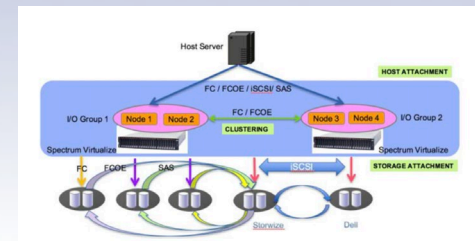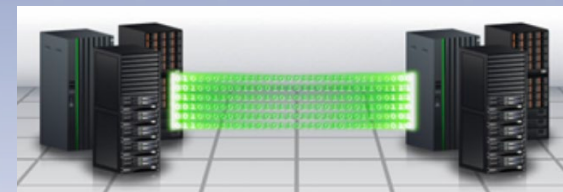
Flash Memory Summit

The ~Problem Today

- CPUs tend to do everything?
  - Not quite true – GPUs are there too.

- MAIN Point: All data goes from storage to processors
  - Network and memory speeds and even processor speeds limit analysis

- Always better to process data closer to where it resides:
  - Internal fabric of the SSD may be faster than the external fabric to the server/AFA
  - Internal fabric of the AFA may be faster than the external fabric to the server

- Analytics, AI, Machine Learning, Deep Learning require huge amounts of data
  - Even 100Gb Ethernet and DDR5 and Gen 4/PCIE5 pose limits

Sensor Data

Social Media

Structured DB data

Text files

Financial Data

Multimedia

0101
010010

Offload is not new

- Data Reduction is ubiquitous in AFAs – (and in IBM FCM SSDs)

- Replication and Hyperswap offload

- Storage Virtualization

- RAID

The New Motivation

Storage

- The Data Lake is simply enormous and growing daily
- The more data analyzed the more value and accuracy
- Network speeds really can not grow fast enough
- Server compute has some limitations as well with physics and Moore's Law ending.
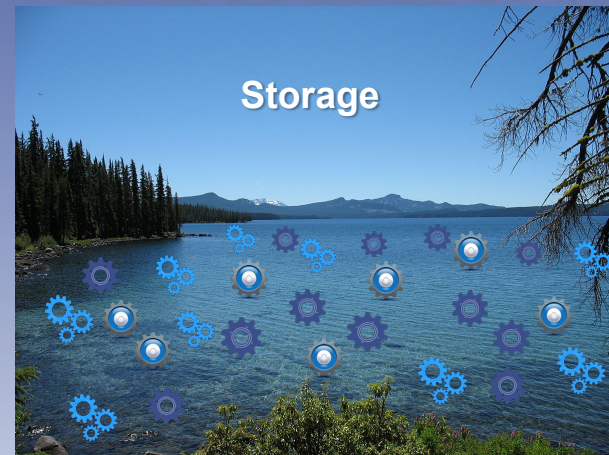
Move some of the computation closer to the lake or even in the lake!
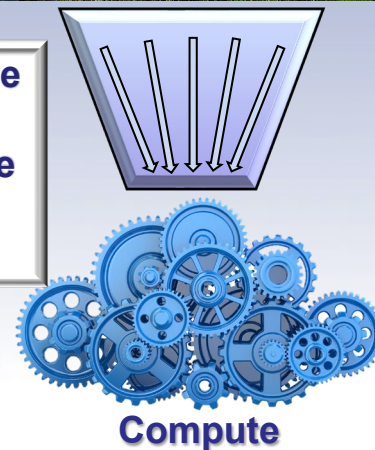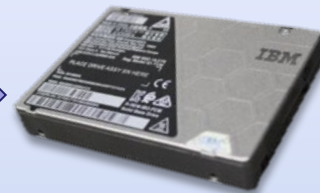
Compute

The SSD as a Computational Storage Platform

- **Example: IBM FCM**
- **FPGA Controller**
- **4 arm cores**
  - **2 used for FTL and NVMe**
- **Has an efficient Compressor INLINE**
- **Encryption inline**
- **No Change to Storage Stack data path**
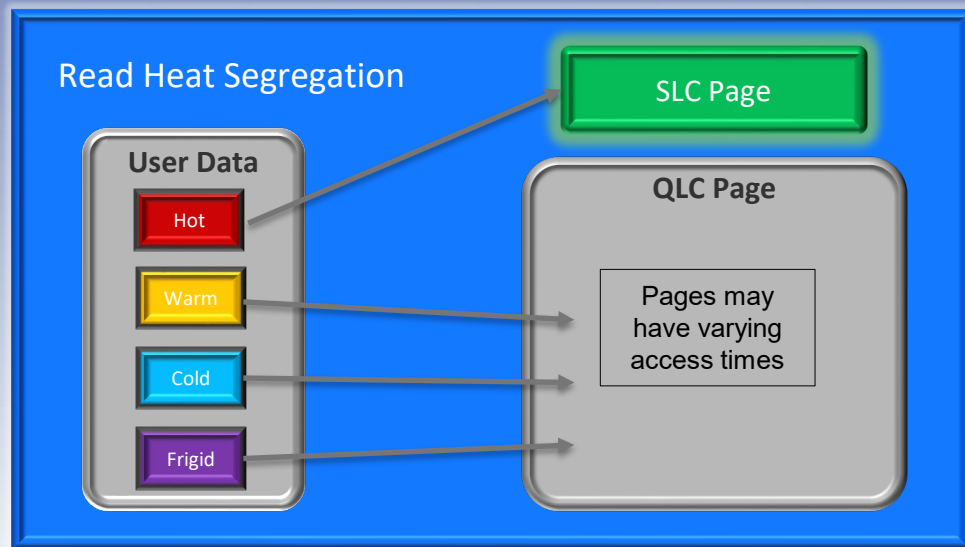- **Hinting Architecture**

OFFLOAD

Leaves processors and GPUs for other critical work

Simple Hinting Architecture

- **Flash can have asymmetrical accesses**
- **SLC Mode is fastest**
- **App passes hints about the type of data and where to store.**

Read Heat Segregation

SLC Page

User Data

Hot

Warm

Cold

Frigid

QLC Page

Pages may have varying access times

Future of Computational Storage Devices

Especially for navigating data lakes

- Inline filtering or reduction
  - Single PCIE Device
- Reducing the amount of data to read
- Reducing the amount of data to store
- "Self Describing Data"
- Offloading tasks that can run on the cores in CSDs that do background searches or filters
- Report on access temperature, age, duplication, other

The AFA as a Computational Storage Platform

- Disaggregated Storage

- External Storage

- HCI

- Takes significant compute to run a storage stack

AFA as a Computational Storage Platform

- **Accelerators and offload functionality could be built into the storage stack**

- **Or – Storage stack can be virtualized or containerized and allow for guest VMs or Containers**

- **Smart NICs can offload functionality**

- **Accelerators and FPGAs could be connected off of CPU**

  - **PCIE today**

  - **CXL tomorrow**
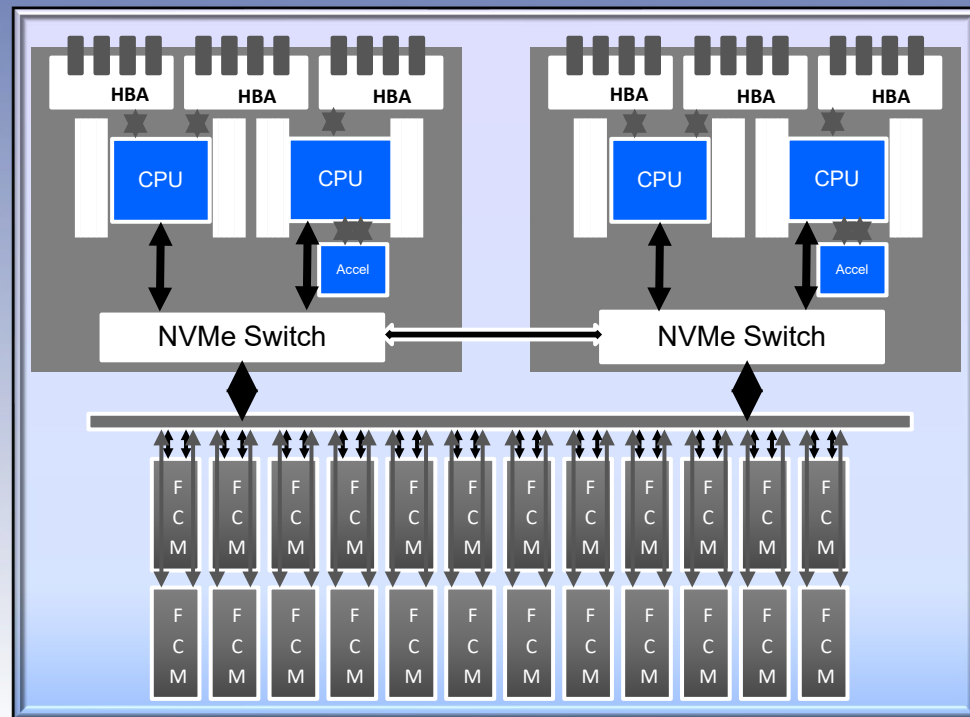
Thank You!

Flash Memory Summit

Everything You Need To Know

For Success