Use cases of Computational Storage Drives (CSD)

Eli Tiomkin, VP Business Development
The simplicity of scalable, ASIC-based computational storage

An enterprise class device capable of processing workloads in storage at the source

Needed Key Attributes:

- Use standard protocols (NVMe)
- Minimize data movement (Faster Response, Lower W/TB)
- Improve (TB/in³) with maximize (Customer TCO)
Platforms are growing to fill the need

- The highest density 1U solutions on the market with added compute!

- Up to 288TB of storage in 1U

- Adding 144 Processing Cores
  - Based on Arm® cores with 64-bit Linux

- The Market is consuming M.2 faster than U.2 driven by the cloud providers

Source: Market Analysts, NGD Analysis
Edge Computing
Processing the data at the edge

Azure IoT Edge running on NGD’s CSD
- Customer run Azure workloads at the edge and enables a seamless workflow to Azure Cloud

ML training at the edge
- MobileNet is a Tensorflow application that was ported into the drive
- Object identification and ML running on CSD
- Object Identification is a growing field and needed at the Edge more than anywhere
- Drive level ML and AI provides Scale in any platform where multiple storage devices are needed
Using Computational Storage Drives for ML.

Keras + TensorFlow

MobileNetV2.
Using Computational Storage Drives for ML.

Constant Updates to Training Model
Edge Computing
Processing the data at the edge

Image Classification at the edge
- This showcases the ability to run multilayer applications within the drives
- Drive level ML and AI provides Scale in any platform where multiple storage devices are needed
- PoseNet AI is a TensorFlow.js application that captures live streams and matches human position
- Human Pose mapping is valuable in Medical and Security field work. With the ability to scale by drives in a system, the accuracy and ability to manage the data is increased
Using Computational Storage Drives for ML.

**PoseNet AI Human Pose Identification**

TensorFlow & Node.js

Limited Host Interaction Required
Big Data Analytics

- Big Data Analytics
  - Collect 4TB of data from 50K sensors on SATA SSDs
  - Move Data to NVMe Storage system for analytics with 6PB of existing data
  - Optimize the applications to take advantage of computational storage
MongoDB on CSD

- **MongoDB** Execution of DB application and datasets within the drive offers unique opportunities to scale and deploy DB applications with limited external architectures.
- Retail website based on MongoDB to run within the drive. There are no host resources required to execute the shopping experience.
- This allows for scale of deployments at the Edge by reducing in store IT HW footprints.
- Scaling multiple DB instances that can be read/updated/stored with no Host resources has an numerous TCO advantage.
- As storage grows, adding CSD matches the compute capabilities without changing the storage infrastructure.
Using MongoDB within Computational Storage.
Hadoop on CSD

- **Hadoop** – Running Hadoop via Containers within the drive further illustrates the ability for Computational Storage to add even more value to Scale Out architectures

- Running Hadoop today users get ‘baseline’ performance regardless of the amount of storage behind the platform. However, acceleration is really needed to make it more valuable

- By reducing Host core usage from 16 Host cores active to 8 Host cores active, you save 50% of the host resources

- With added drives via NGD Computational Storage you can save over 40% of the time needed to run Hadoop workloads and save host and system level power

- Since adding drives adds capacity and compute, the scalability of platforms is impressive
Amplifying TCO for Hadoop

Host Platform

- DRAM
- Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz
- 32GB RAM

Datanode Config:
- Single E5-2620v4, 32GB DRAM, 12*8 TB SAS HDD
  - 18U Total Density in 18U = 864TB
  - 9 Cores for Data Processing

Saves Power!
Saves Space!
Saves Time!

Datanode Config:
- Single E5-2620v4, 32GB DRAM, 36*8TB NVMe
  - 3U Total Density in 3U = 864TB
  - 432 Cores for Data Processing

@ Scale

TERASORT
WORD COUNT

16 Cores Standard Host Use Case
8 Cores with NGD M.2 Computational SSDs
Thank You.