

On Drive Linux – **Computational Storage** – World's First Computing SSD

Scott Shadley, VP, [NGD Systems](#)

SNIA, Board of Directors, [Computational Storage TWG](#) Co-Chair

The Need for Computational Storage with On-Drive Linux

Topics covered today will include:

- The Market Evolution and need for Local Compute
- The Paths to Computing
 - Smart 'Things' and where they reside
- The difference between 'Smart' and 'Intelligent' Solutions
- How On-Drive Linux is Accomplished
 - Why Linux? Why Now?
 - K.I.S.S. - Discussion on the Deployment Model
- The Use Cases of On-Drive Linux to Accelerate AI, ML, and Database

The Market Evolution and Need for Local Computing

Our Friends at Gartner Say it best...

Structured Data is great for current infrastructure

Allows for ease of data movement, location, access, compute

Only a small subset of the real data Iceberg

Unstructured Data is the greatest threat to results

As more and more data is generated, it is more random

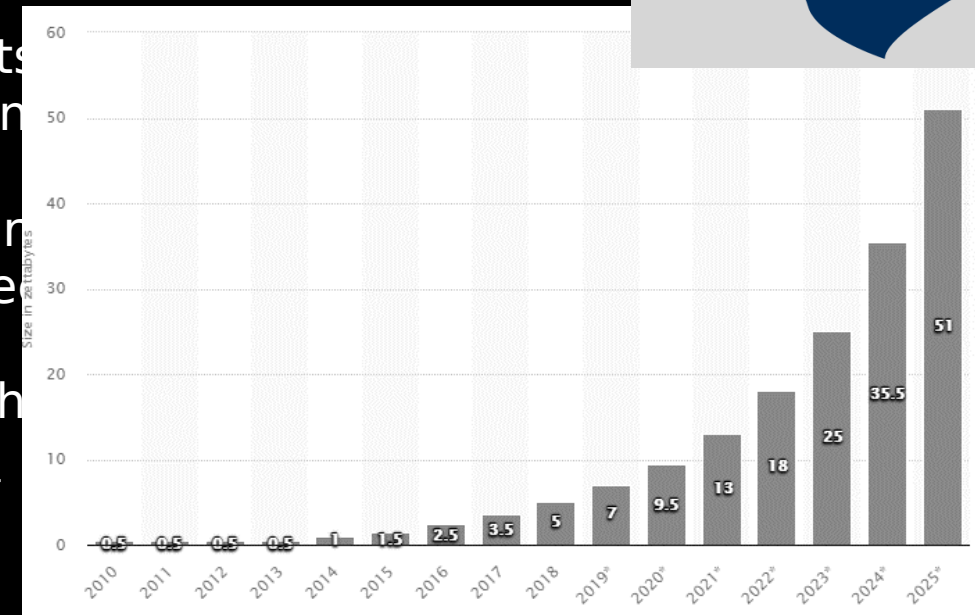
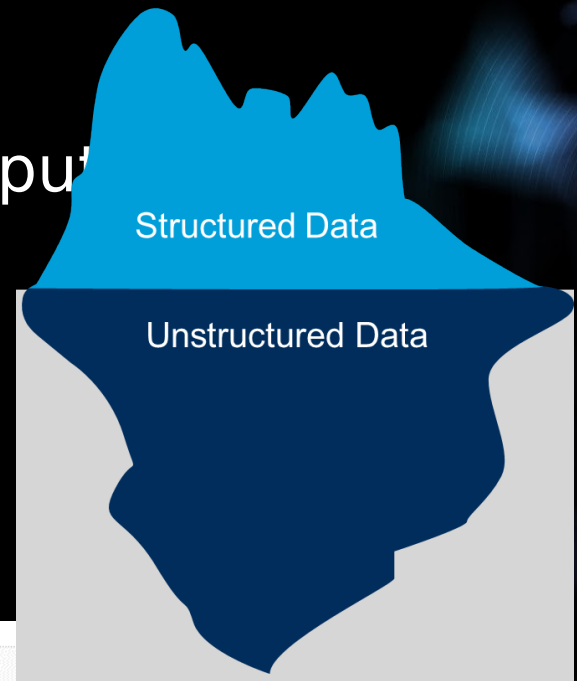
Needs to manage this data locally are key

Edge Computing is not able to scale at data growth

A new way to compute on random, local data is needed

The Global DataSphere (Statista.com) shows how the data growth is overshadowing the compute growth

CHANGE IS NEEDED



The Path to Compute Solutions is Paved with Smart Intentions

Finding paths to compute is easy... But one thing is very lacking in these 'Smart Things'

Compute is Needed, **DATA is Mandatory!**

CPU - The Brain of the operations, starved for data, overwhelmed with requests

GPU - The Parallel processing Master, Nothing Persistent about it

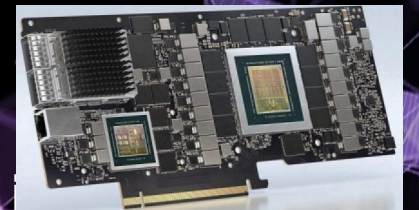
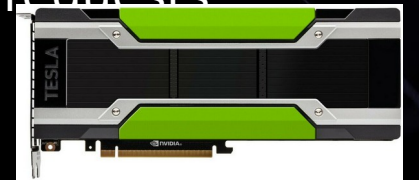
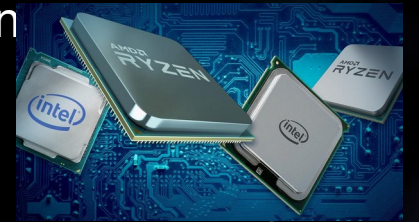
NIC - The great Mover, not so great at processing

Smart NIC - The intelligent mover, but still doesn't know what it is moving

DPU - The Processor closer to data, but still not persistent, still **Volatile**

All these pieces are needed parts of the new ecosystem. But NONE of them solve the

Real Issue... The Data, where it is, where it comes from, and **how to Store**



Being Smart is Good, Being Intelligent is KEY

Data is Stored. Data is Moved. Data is Managed. Data locality is largely ignored.

Processing Data Requires the Right Recipe!!

Raw Ingredients are useless with out the right way to get to a finished product

SMART things are great, but they need real intelligence

Solutions exist in many forms, but all require some sort of 'Host Guidance'

SmartNIC - CPU must provide routing solutions

GPU - CPU must provide instructions

CSP/CSD - Computational Storage Processors and Drives - Based on FPGA today

Linux-Based Computational Storage Drives (CSDs) are THE path to the true Intelligent Data Processing

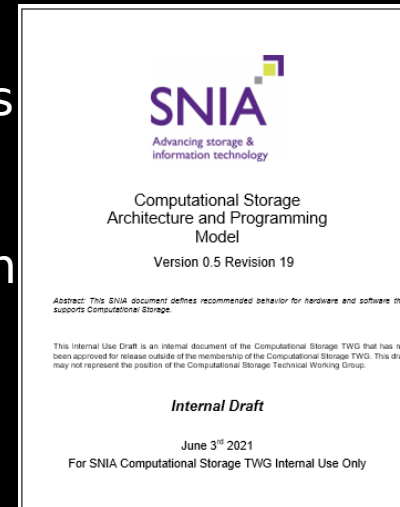
Especially when the storage devices have a local Compute Infrastructure

Computational Storage Exists To Solve Data Movement and Scale

SNIA is driving for an Architectural Solutions

NVM Express is working on an Initial Implem

Prototyping and Deploying Now



The charter of Computational Storage Task Group is to develop features associated with the concept of **Computational Storage on NVMe Express devices**.

The target audience consists of the vendors and customers of **NVMe Storage Devices** that support computational features.

DELL EMC



arm



Microsoft



Booking.com

AIC

Lenovo

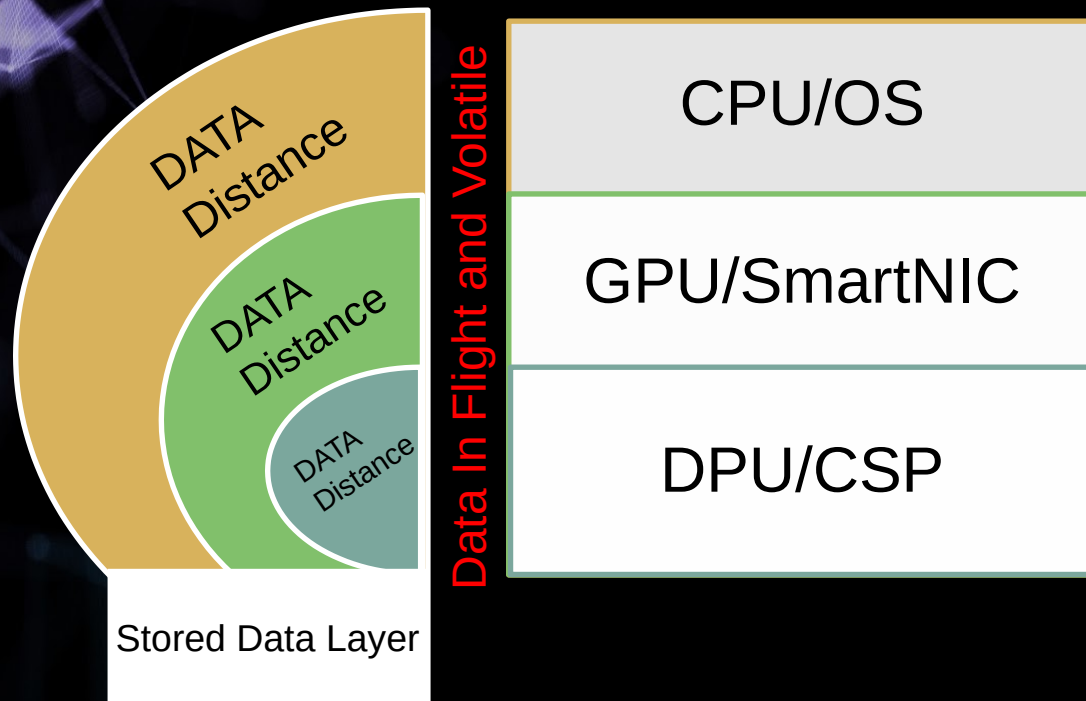


sodacon
— Global 2021 — July 13-14

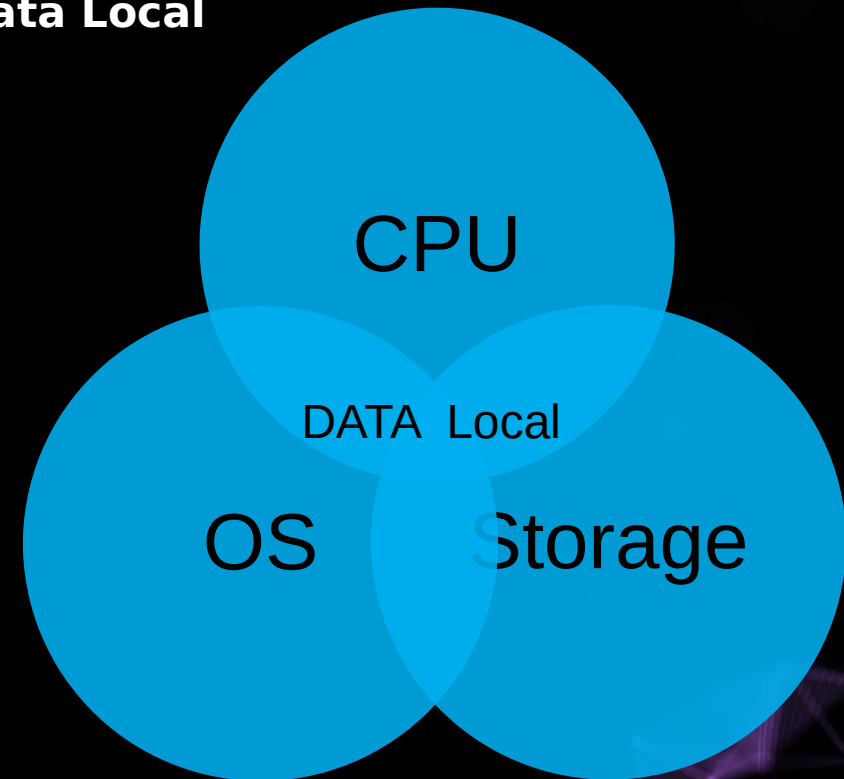
#sodacon2021
#ComputationalStora

A Comparison at Compute Infrastructure – Why CSDs are Needed

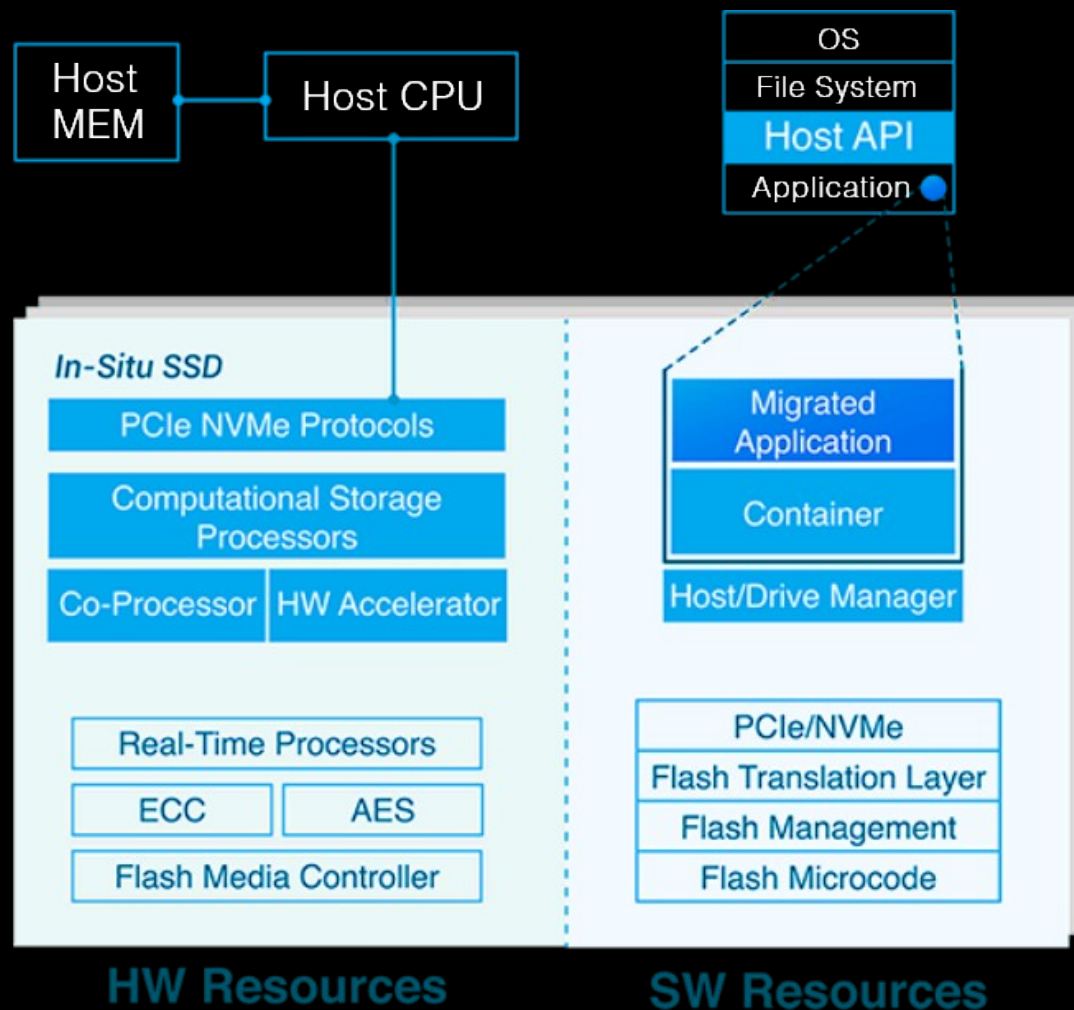
Today's Standard Infrastructure – **Data Distant**



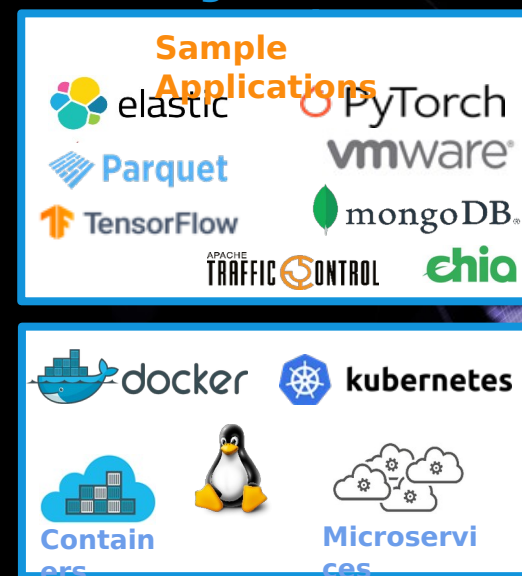
Linux-Based Computational Storage Drive – **Data Local**



A Look at the Hardware and Software of a Linux-Based CSD

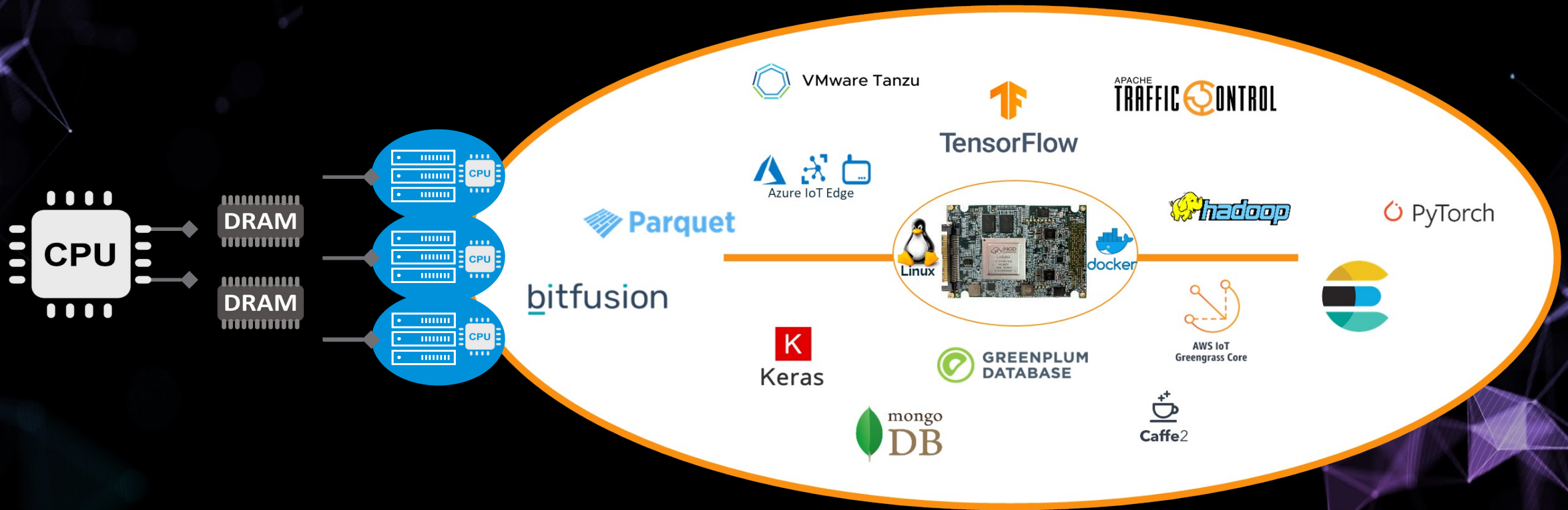


Computational Storage Software



Some Examples of Linux Deployments with K.I.S.S.

- **Keep It Simple & Seamless** –
 - The best way to move technology forward is to leverage architectures already in use



On Drive Linux Interacting Directly with AZURE Cloud



HOST System Linux

ON DRIVE Linux

Used for simply saving data
ALL

Used for
Azure Interaction

**Computer Vision Application executed
using only the NGD Systems Newport
Drive OS,**

**No Host Interaction with Cloud
Required**



<https://www.ngdsystems.com/media/NGDs-Computational-Storage-device-is-powered-by-Azure-IoT-Edge>

**NGD Systems Computational Storage
Drive (CSD)**



Image
Analyz
Edge
Edge Agent

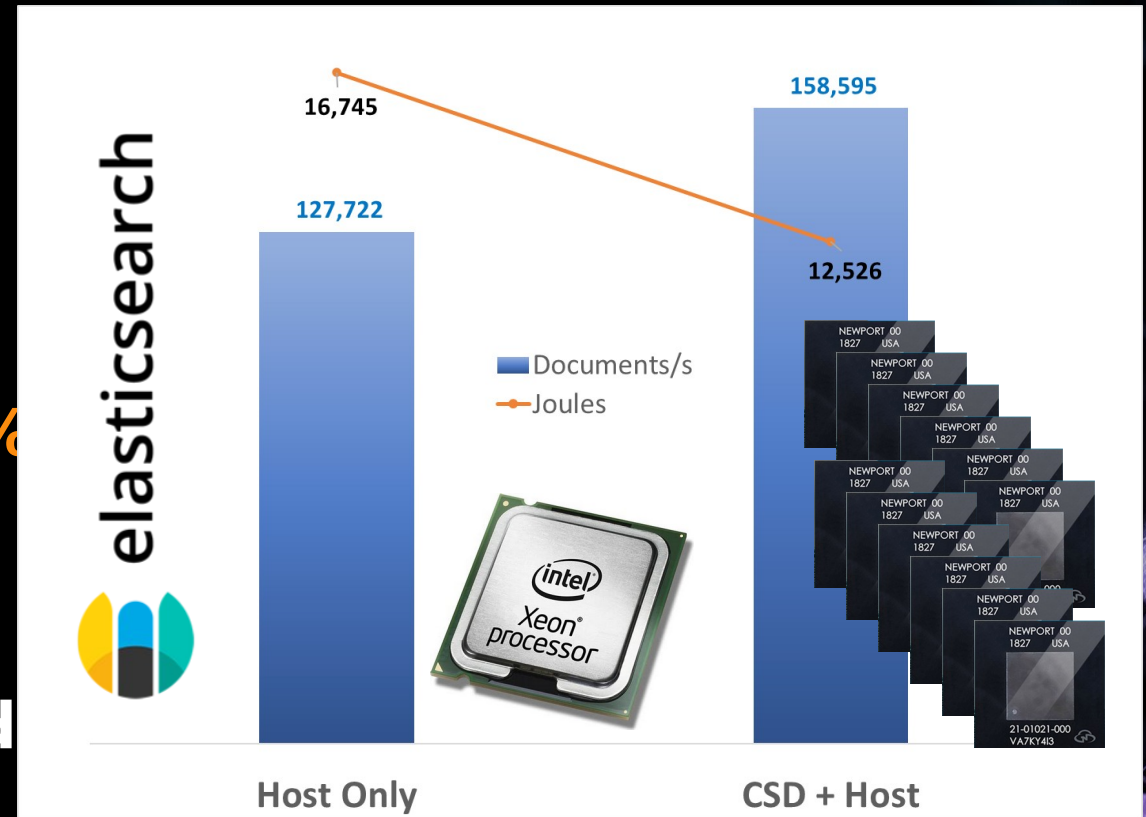
#sodacon2021
#ComputationalStora

DATA CENTER



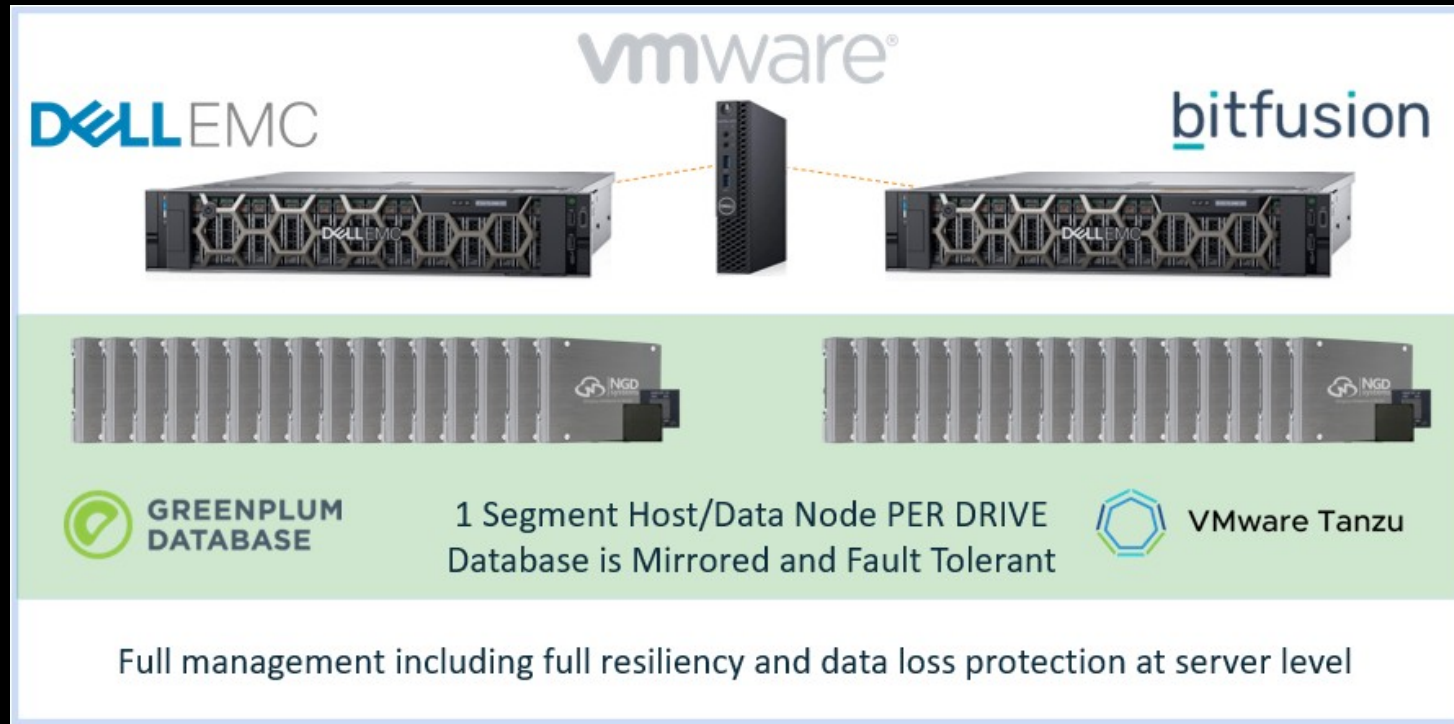
System with 2 Host CPUs, & 144 added On Drive Linux Cores

- **Total Performance Improves**
20% Better Results
- **Reduced Power Consumption**
30% LESS Power
- **DRAM Usage Reduced by >50%**
Host Only used **25GB**
Hybrid used **12GB**
- **CPU Usage Utilization Reduced**
Host Only used **24%**
Hybrid used **10%**



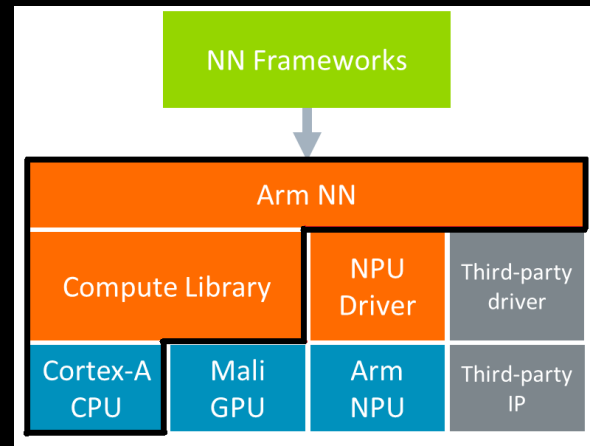
VMware - Tanzu - GreenplumDB - Bitfusion - On Drive!!

Computational Storage with On-Drive Linux allows it to be drive level.
Reducing footprint, server cost, while still offering full fault tolerance

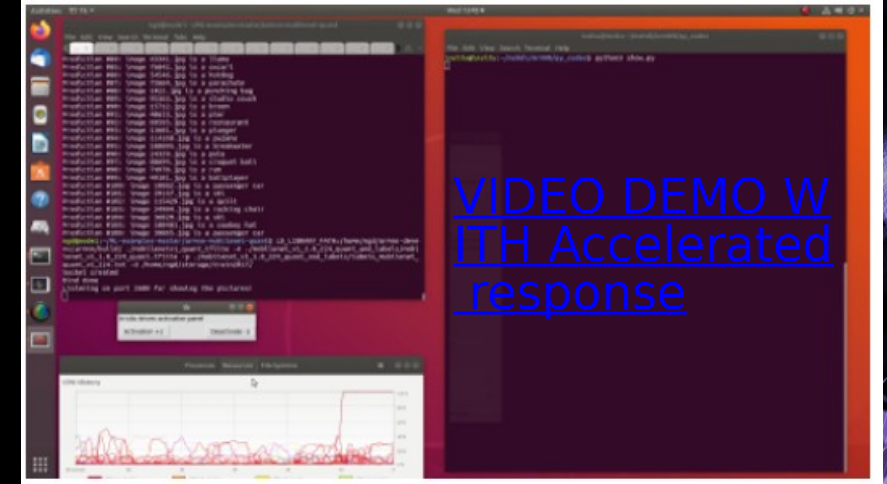
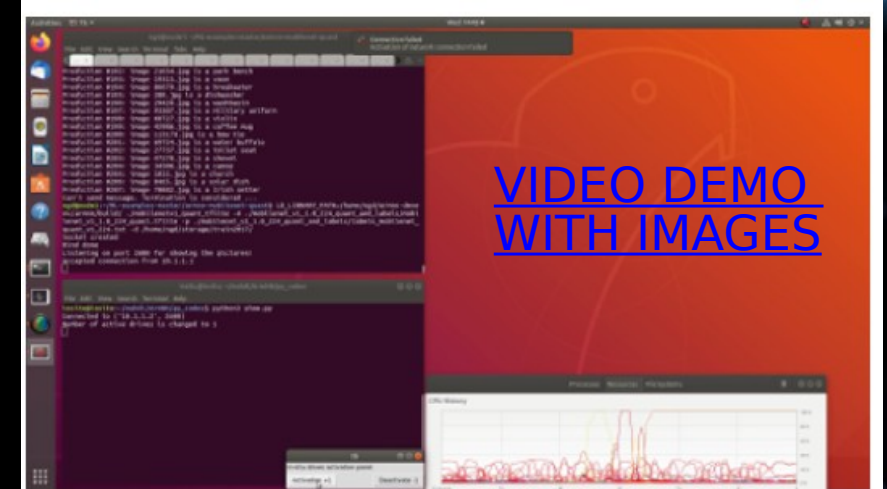


Showcased at Vmworld 2020 - Session ID [OCTO478] –
Computational Storage, Tanzu Greenplum, vSphere Bitfusion

TensorFlow, MongoDB, MobileNet - NO HOST USED



- More than **100k images per drive** are processed without sending the images to host.
- The output of the application is written directly to a MongoDB
- The database is implemented inside the **On-Drive Linux Environment** as well
- The user can access the output data using MongoDB APIs.





User Experience – Let them Show you how On-Drive Linux is
KISS

What's the Value Proposition
for Computational Storage in
eDiscovery?

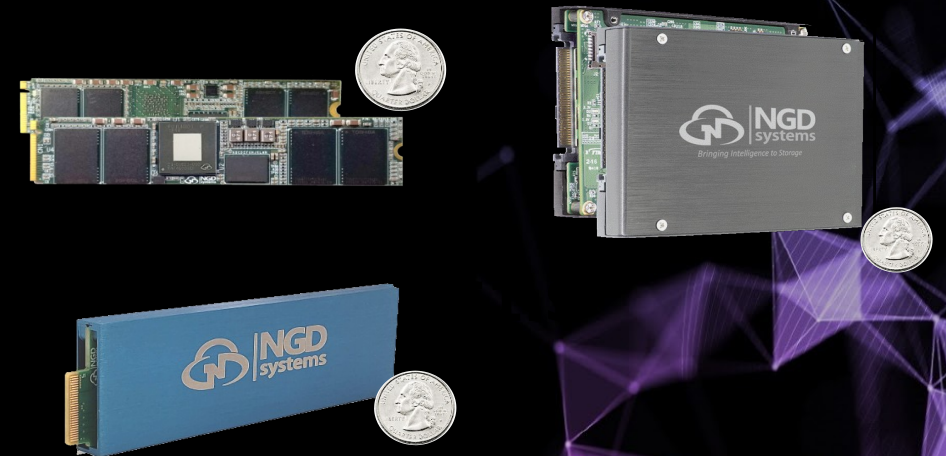
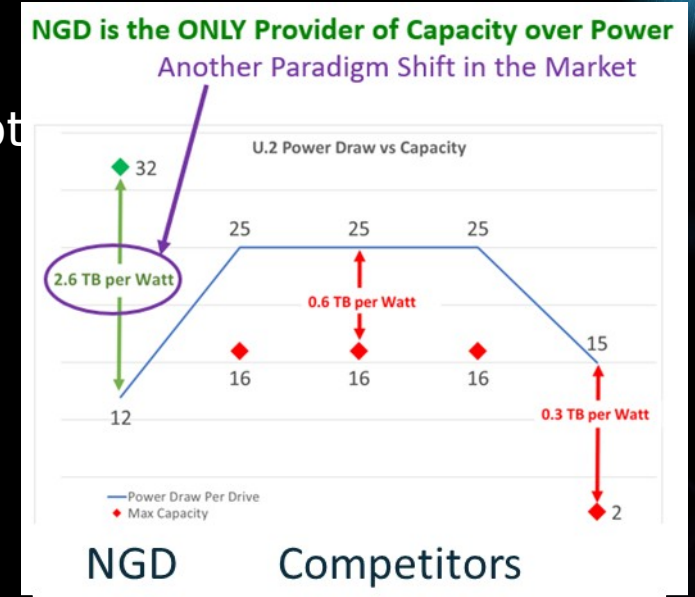


Dan Pollack
CEO, Data Storage Sciences

NVMe On-Drive Linux SSD Products at a Glance.

- Large breadth of **SSD** solutions and capacity options
- Leading **TB/W** Energy Efficiency
- Industry's **Largest capacity** NVMe SSDs
- Quad-Core **Computational Storage CPUs**

Form Factor	Availability	Raw Capacity TLC (TB)	MAX Power (W)
M.2 2280	CQ3'20	up to 4	8
M.2 22110	NOW	up to 8	8
U.2 15mm	NOW	up to 32	12
EDSFF E1.S	NOW	up to 12	12
EDSFF E3	Planned	up to 64	12





Thank You!
Happy Computing!

Scott.Shadley@NGDsystems.com

Info@NGDSystems.com

@SMShadley



@NGDSystems



