

G2M Research Multi-Vendor Webinar Scale-Out Flash Storage: Breaking Old-School Storage Rules



February 25, 2020



▶ Webinar Agenda

- 9:00-9:05** Ground Rules and Webinar Topic Introduction (G2M Research)
- 9:06-9:32** Sponsoring Vendor presentations on topic (9 minute each)
- 9:33-9:40** Key Question 1 (2-minute question; 2 minutes response per vendor)
- 9:41-9:42** Audience Survey 1 (2 minutes)
- 9:43-9:50** Key Question 2 (2-minute question; 2 minutes response per vendor)
- 9:51-9:52** Audience Survey 2 (2 minutes)
- 9:53-10:00** Key Question 3 (2-minute question; 2 minutes response per vendor)
- 10:01-10:13** Audience Q&A (13 minutes)
- 10:14-10:15** Wrap-Up

G2M Research Introduction and Ground Rules

- ▶ Mike Heumann
Managing Partner, G2M Research

Panelists



Shailesh Manjrekar
Head of AI, Strategic Alliances
www.weka.io



Josh Goldenhar
VP Outbound Marketing
www.lightbitslabs.com



Tom Leyden
VP Corporate Marketing
www.excelero.com



Host/Emcee:
Mike Heumann
Managing Partner
www.g2minc.com

What is Scale-Out Flash Storage (SOFS)?

- ▶ Scale-Out Flash Storage (SOFS) solutions consist of the combination of an SOFS software package and servers (typically) containing NVMe™ storage devices.
- ▶ SOFS evolved from Software-Defined Storage (SDS), but scales out to hundreds or even thousands of nodes.
- ▶ SOFS also borrows heavily from the DAS-centric storage architectures utilized by hyperscale companies (AWS, Google, Azure, Facebook, and the BATs).
- ▶ Utilizing NVMe over Fabrics as a backbone, SOFS solutions can provide remote file system access across large number of servers and storage appliances with SSDs, but with local-like performance and latency.



Why Is SOFS Successful When SDS Struggled for Acceptance?

- ▶ SDS was built around a handful of storage appliances connected to compute nodes over Ethernet, usually NFS or iSCSI
 - Storage appliances connected to back-end storage (JBODs) via SAS or SATA with limited flexibility due to SAS/SATA networking limitations
 - To users, SDS solutions looked just like storage array-based SAN or NAS deployments, but (typically) at a lower cost and with lower performance
 - ***SDS was essentially a “poor-man’s SAN”***
- ▶ SOFS uses NVMe Over Fabric™ (NVMe-oF™) to connect storage, compute nodes
 - Nodes can be compute nodes, storage nodes, or both – no differentiation
 - Completely flexible/composable network topology
 - Remote storage performance approaches local storage performance
- ▶ This allows SOFS solutions to easily scale to ***thousands*** of nodes
- ▶ Also eliminates the “how much storage where” issue





Weka

- ▶ Shailesh Manjrekar
Head of AI and Strategic Alliances
www.weka.io

Agenda

- New Workloads - AI/ML/DL apps are inherently different
- New Architecture - Edge to Core to Multi-Cloud
- New Approach - Accelerated DataOps

**“There is no AI without IA
– Information Architecture”**



Company Overview

The Company



Founded in 2013
R&D in Tel Aviv, Israel
US HQ in San Jose, CA

The Investors



The Technology

10X



Faster than All Flash Filers

First True "Fits All Storage"

Strong patent protection

54 patents filed

14 issued

The Board

- Liran Zvibel, Co-founder
- Maor Ben Dayan, Co-founder
- Omri Palmon, Co-founder
- Jim Sullivan, Independent
- Dror Nahumi, NVP
- Lip-Bu Tan, WRV
- Menashe Ezra, Gemini

The Advisors

- Suresh Vasudevan, Nimble CEO
- Bill Moore, DSSD CEO
- Randy Seidl, HPE, Sun, EMC

Customer Successes

Genomics/Life Sciences



AI/Machine Learning



Financial Services



Traditional HPC



Media & Entertainment



Manufacturing/Engineering

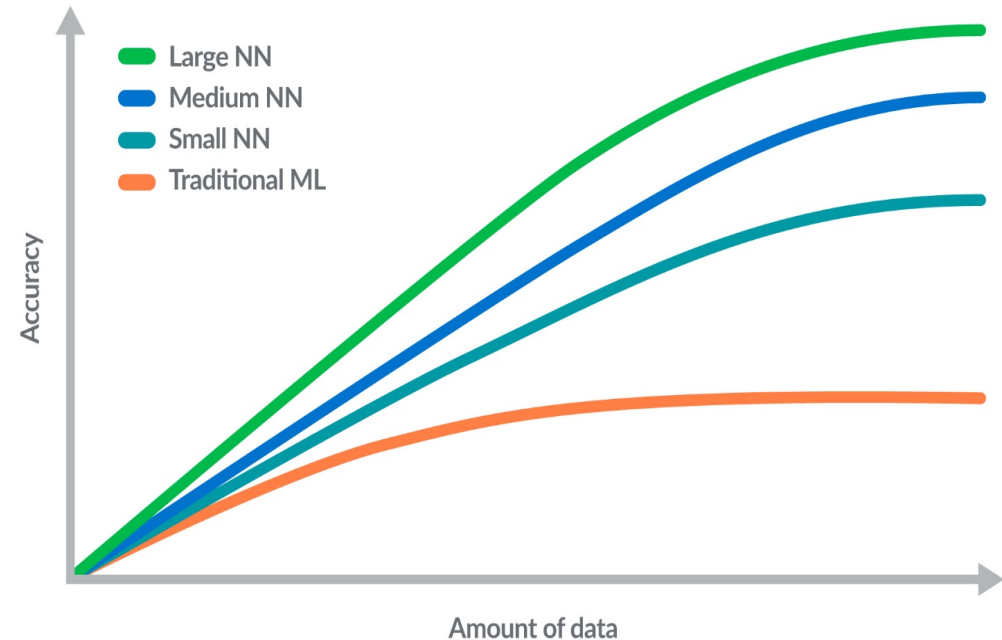


Deep Learning - Data Centric Use Cases

Advanced Driver Assist System (ADAS)



Level of Autonomy	Deep Neural Networks Needed	Survey Car	Data Storage	Data Processing
Level 2+	10	Equipped with 2MP Cameras	2 PB per Year per Car	100 Nvidia DGX-1 Servers per Car

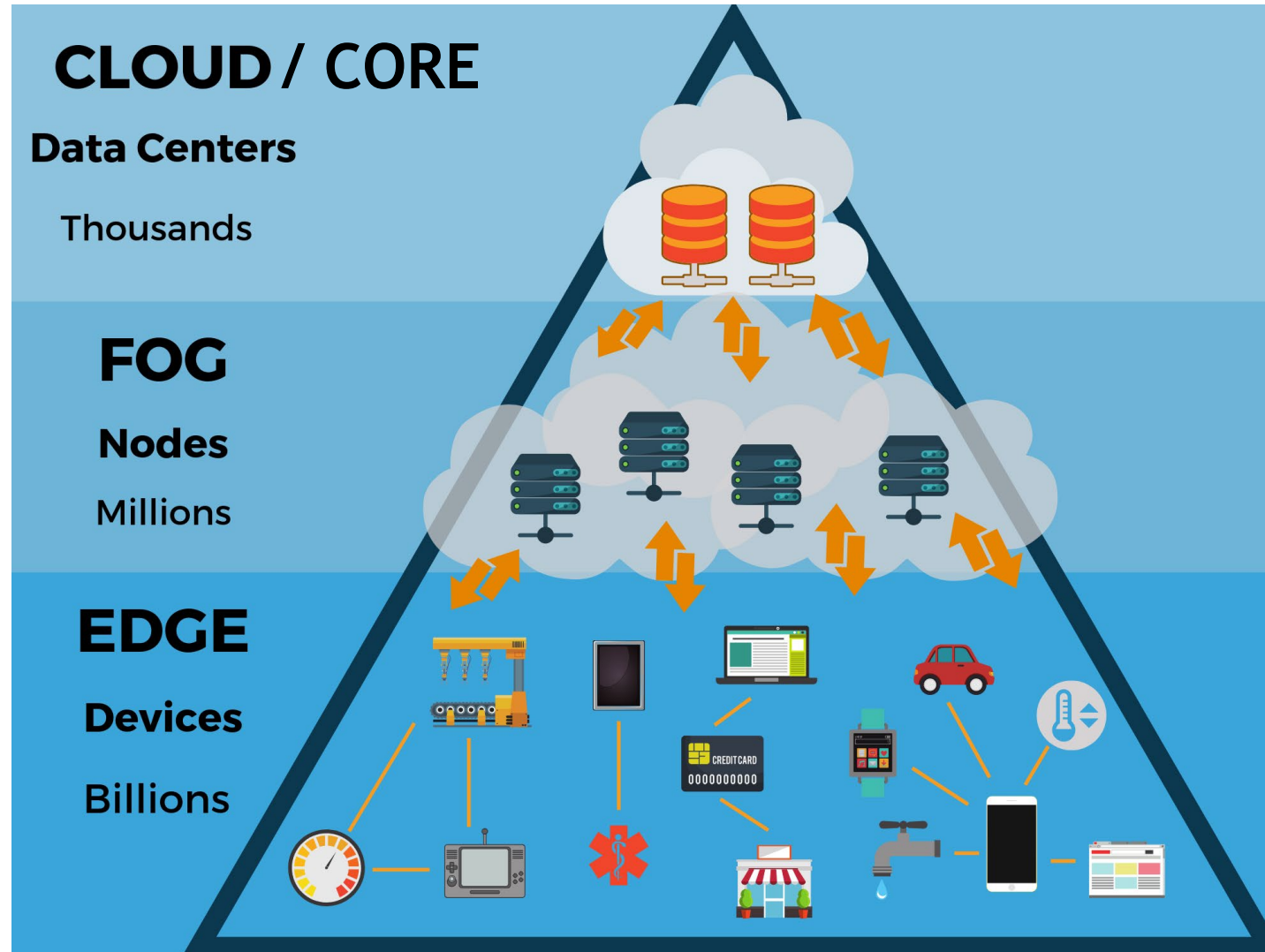


Intelligent 5G Edge – Bigger Than The Cloud !

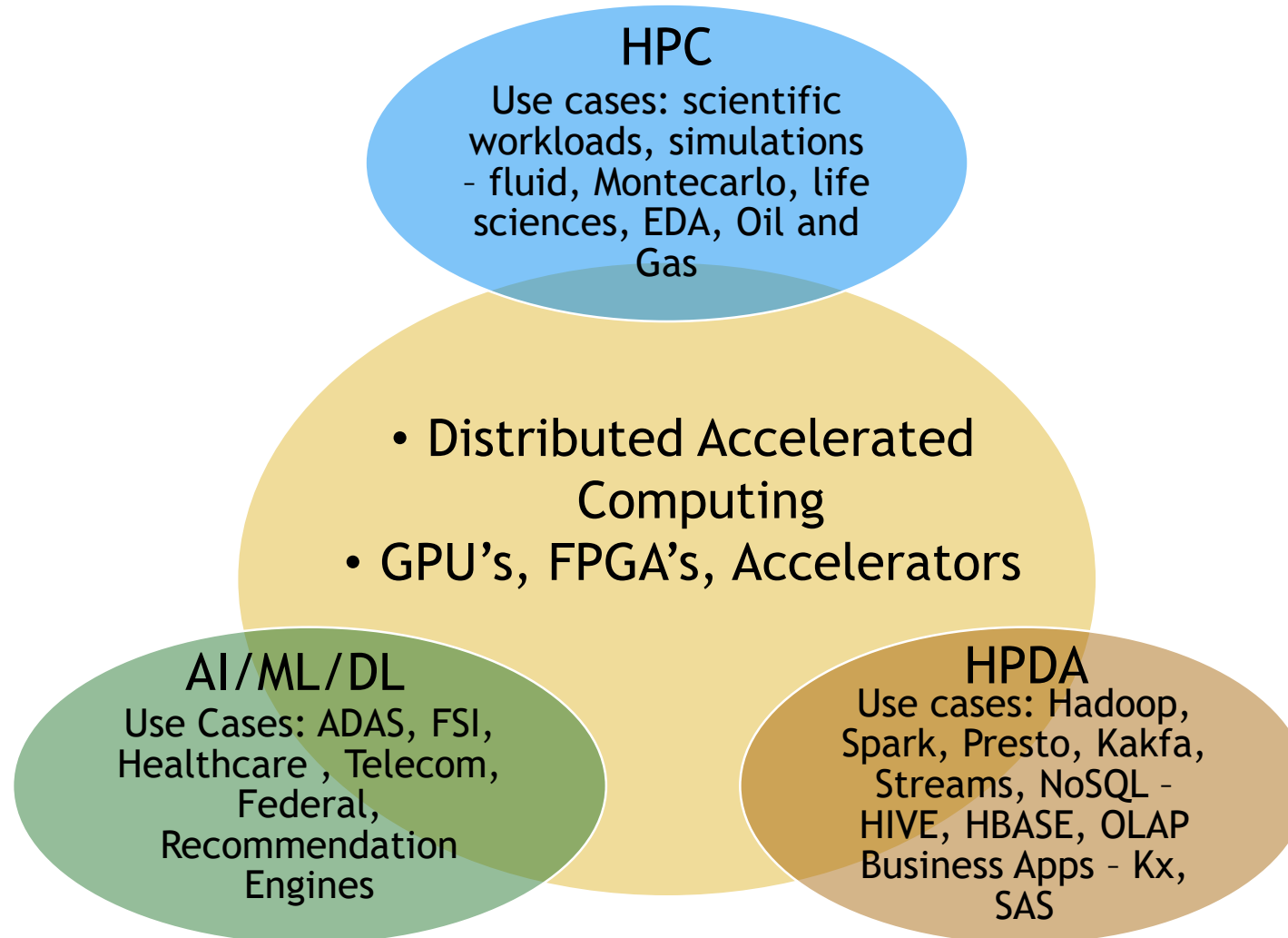
- Training / Inferencing testing
- Application specific processing
- High cost

- Edge Aggregation
- Tagging / High Ingest

- Inferencing
- Time sensitive
- Task specific processing
- Low cost

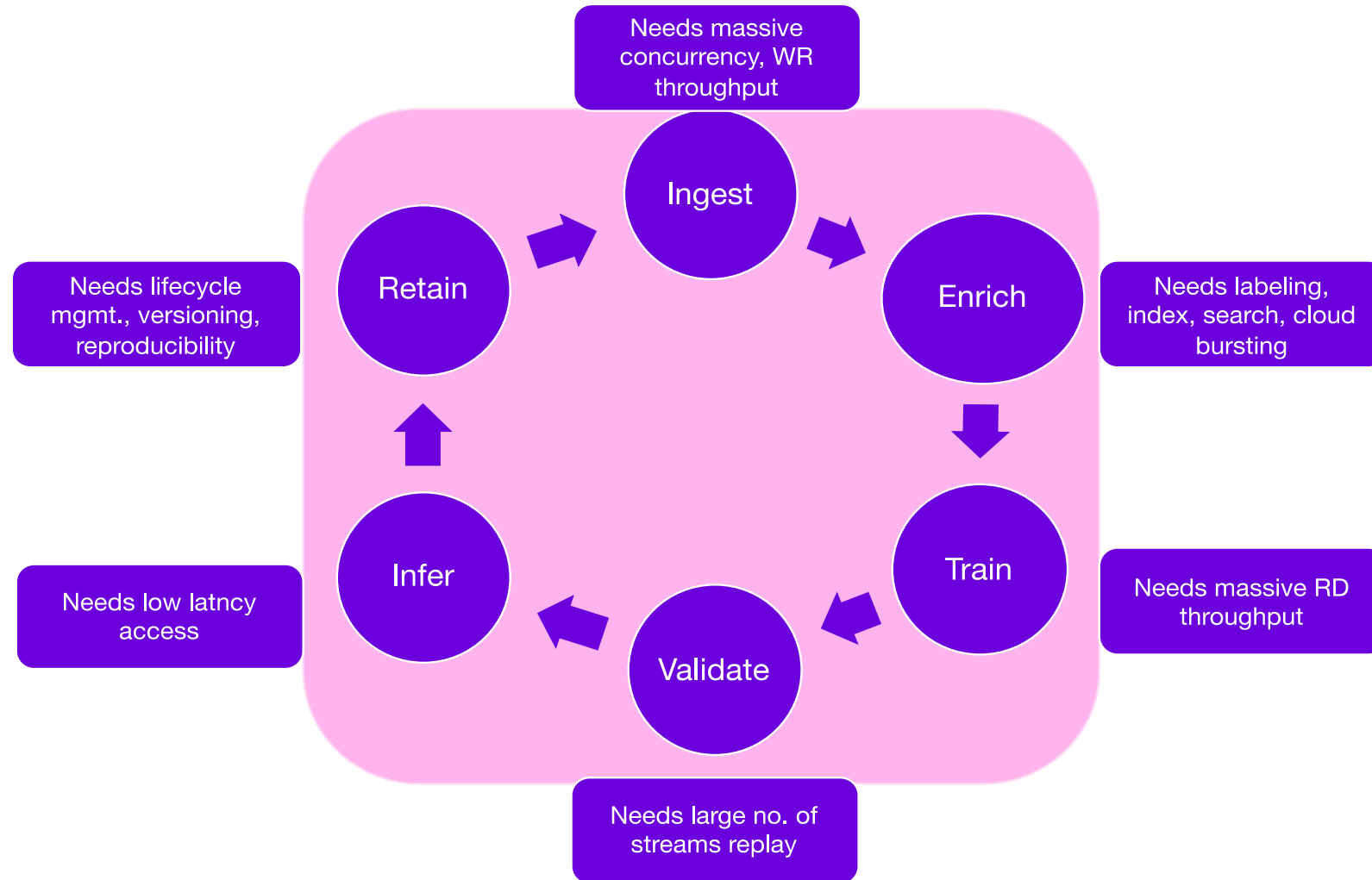


AI 2.0 Market-scape – Use cases

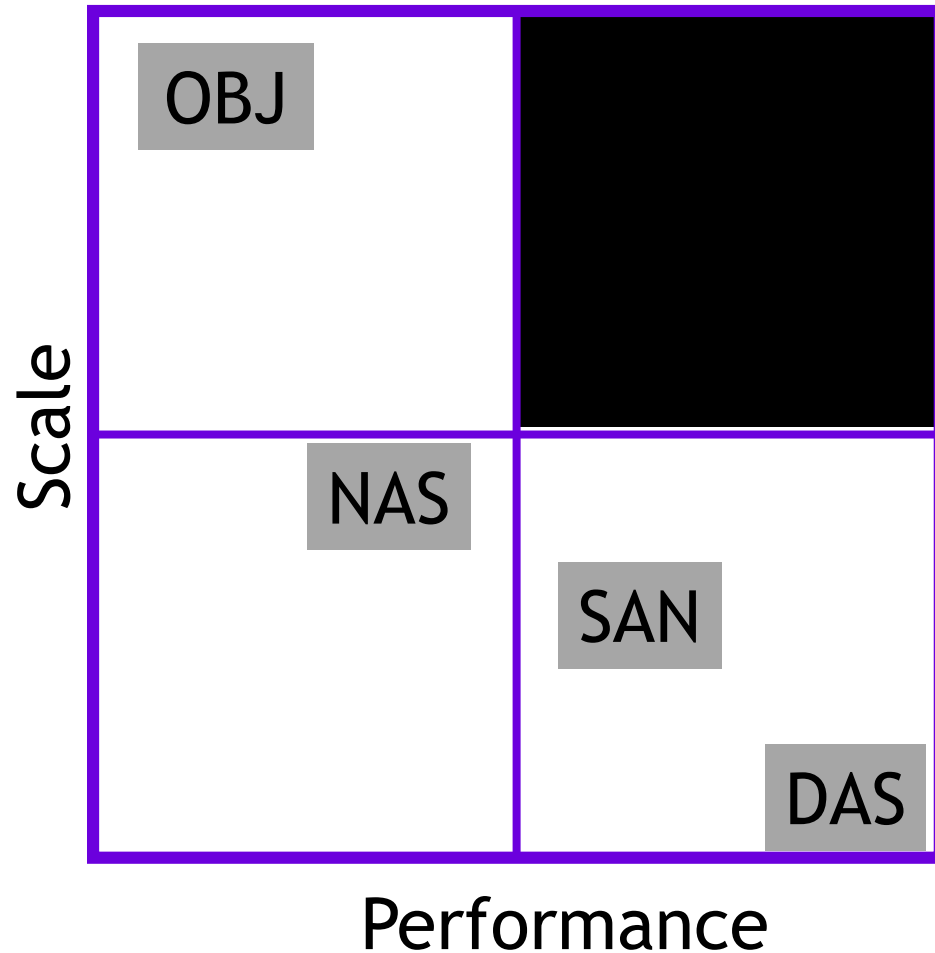


AI 2.0 Workflow last mile problem – Storage I/O

Results in storage silos and delayed time-to-value



Storage is Broken Into “Islands Of Compromise”

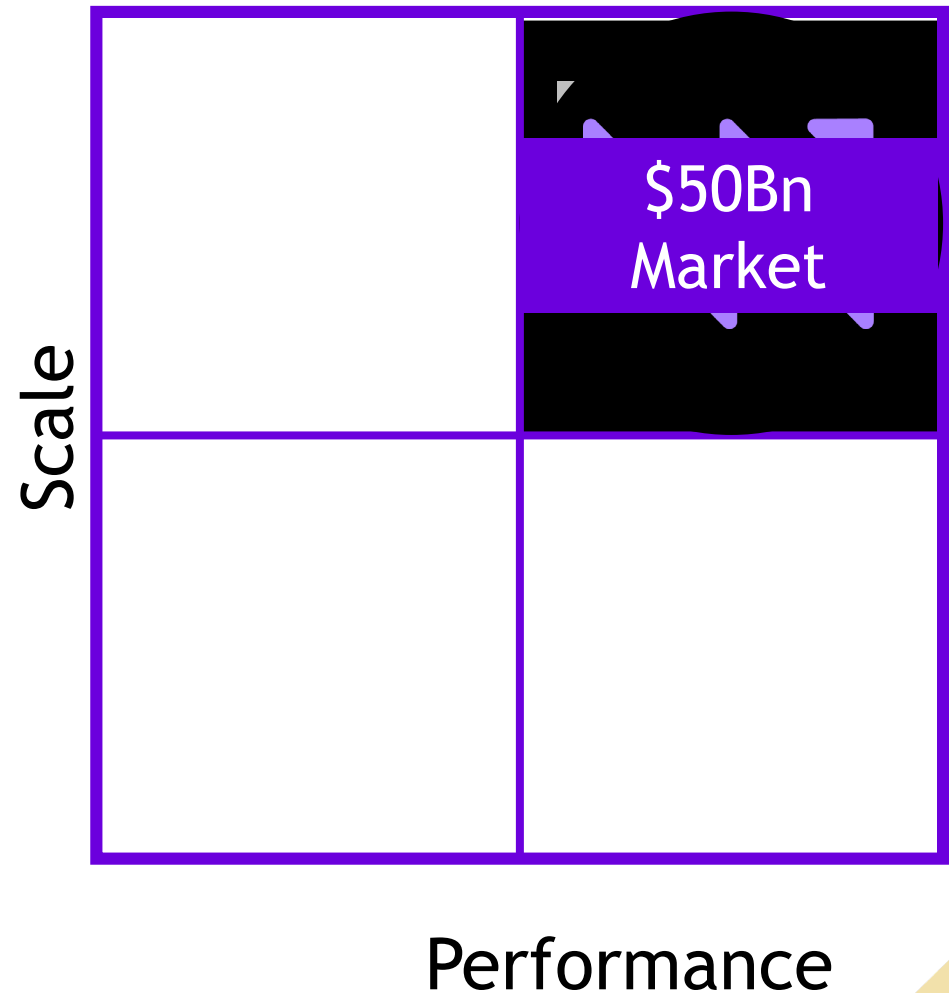


DAS and SAN - Cannot share data
NAS and Object - Performance sucks

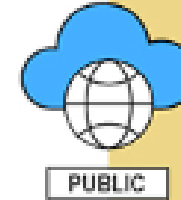
WekaIO – Creating a New Storage Category

Accelerated DataOps

- Fastest time to market (performance)
- Actionable Intelligence (Analytics, AI)
- Collaboration (shareability)
- Mixed workload support (small, large I/O)
- Cloud native, hybrid cloud
- Enterprise Ready



Weka: Full-Featured and Flexible



FILE SYSTEM

DATA SERVICES

CLUSTER RESOURCE MANAGER

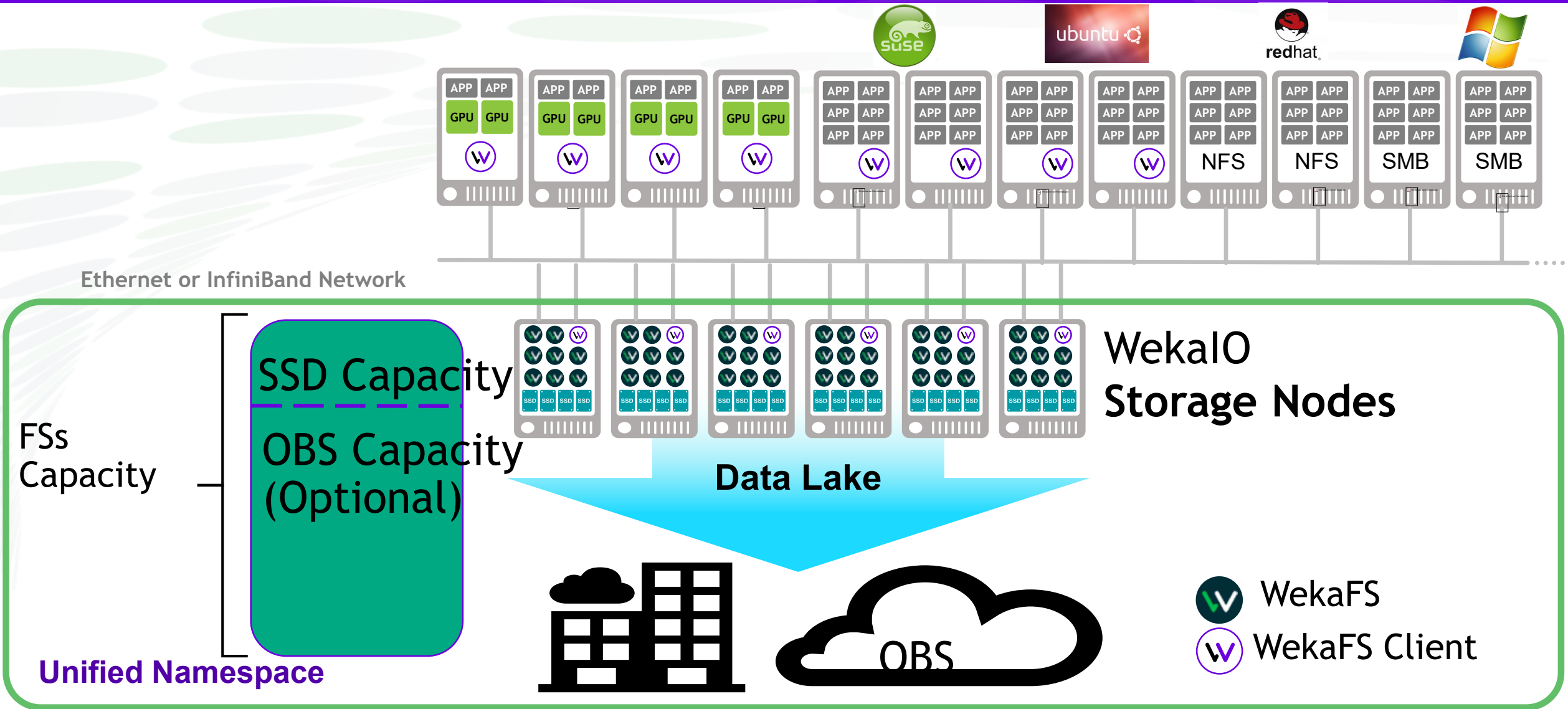
NETWORKING STACK

HARDWARE
RESOURCES

PHYSICAL HARDWARE

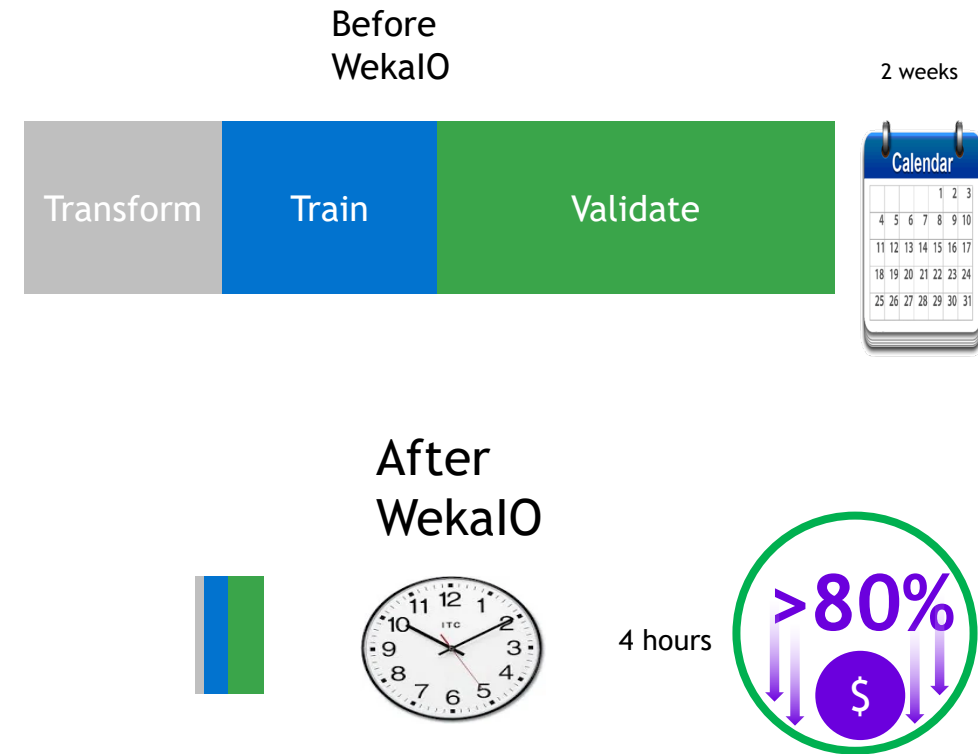
Management

Extend File namespace over object tier



Accelerated DataOps for Data Scientists, CDO's, CAO's

- Improve productivity and faster time to market and value
 - accelerate large scale data pipelines with reduced epoch times, fastest inferencing and highest images / secs benchmarks
 - run entire pipeline on the same storage backend
 - Faster than local storage

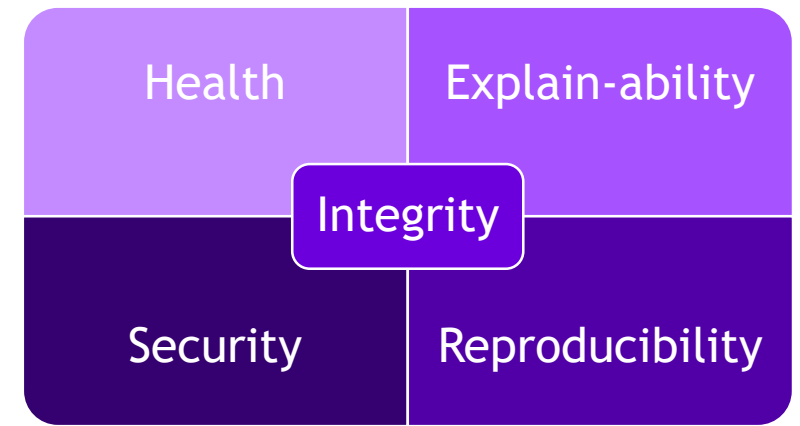


30% better utilization results into \$1.13M in savings for 10 node GPU cluster with 3 Data scientists, over 3 years

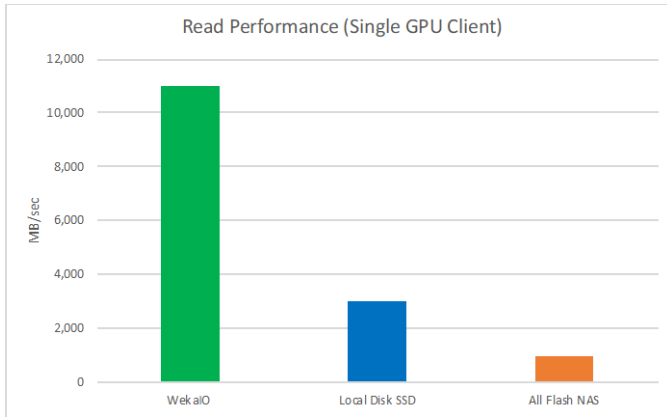
Accelerated DataOps for Data Scientists

- Data compliance and security
 - in-line, no-impact encryption support enables compliance
- Manage data pipelines at the Edge Aggregation, Core Datacenter and Cloud
 - SaaS offering in public cloud for cloud bursting and test dev
- Explain-ability and Reproducibility for experiments
 - Snap2object retains versions for reproducibility and explain-ability
 - instant snapshots make it easy to maintain versions
 - Data mobility by moving snap2object to public cloud and rehydrate

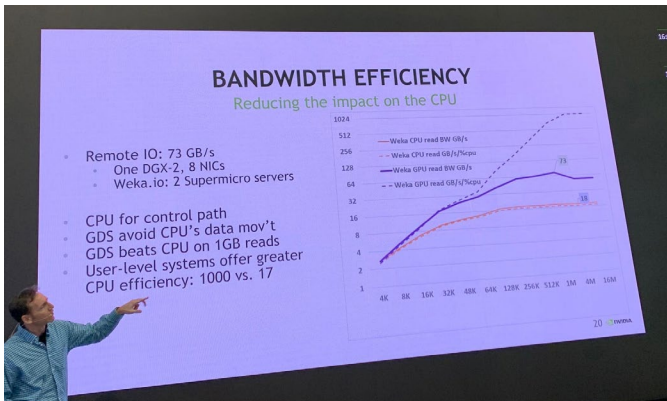
Pillars of AI Trust



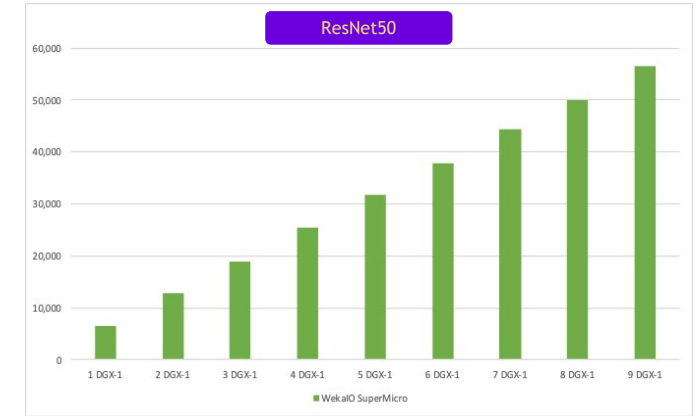
Highest Performance to GPU Servers



- NVIDIA demonstrated Weka performance over 73GB/second to a single DGX-2



- Fully saturate 100Gbit Network link
- 3x faster than local drive Storage
- 10x faster than all flash NAS
- Perfect linear scaling as cluster expands
- NVIDIA validated reference architecture



IO500

This is the official list from [Supercomputing 2019](#). The list shows the best result for a given combination of system/institution/filesystem.

Please see also [the 10 node challenge ranked list](#).



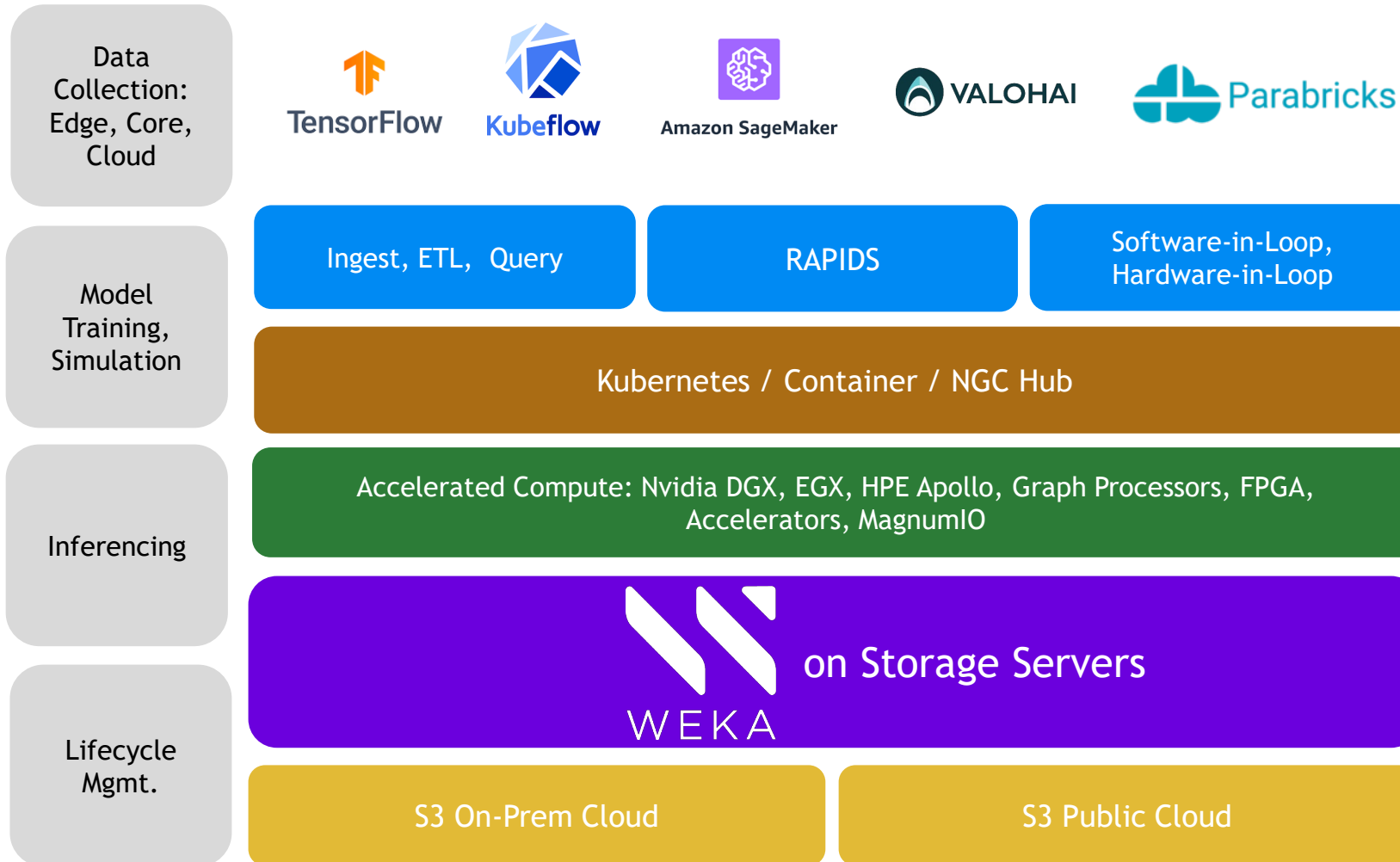
#	information								io500		
	list id	institution	system	storage vendor	filesystem type	client nodes	client total procs	data	score	bw	md
										GiB/s	kIOP/s
1	sc19	WekaIO	WekaIO on AWS	WekaIO	WekaIO Matrix	345	8625	zip	938.95	174.74	5045.33
2	sc19	Intel	Wolf	Intel	DAOS	26	728	zip	933.64	183.36	4753.79
3	sc19	National Supercomputing Center in Changsha	Tianhe-2E	National University of Defense Technology	Lustre	480	5280	zip	453.68	209.43	982.78
4	sc19	NVIDIA	DGX-2H SuperPOD	DDN	Lustre	10	400	zip	249.50	86.97	715.76
5	sc19	University of Cambridge	Data Accelerator	Dell EMC	Lustre	128	2048	zip	229.45	131.25	401.13

Accelerated DataOps for Data Engineers

- Best TCO, with performance at scale
 - leverage NVMe flash for performance and object store for capacity
 - built in data protection eliminates need for another solution



WekaAI for Accelerated DataOps – Small / Medium / Large Bundles





lightbits

Lightbits Labs

- ▶ Josh Goldenhar
VP, Outbound Marketing
www.lightbitslabs.com

Over a Decade of Innovation

Lightbits Team accomplishments and contributions in the NVMe Space

NVMe

Direct-attached
High-performance PCIe SSDs

2009 - 2013

- **First NVMe SSD controller**
- Adopted by top Hyperscalers and all-flash-arrays
- First Linux & VMware drivers

NVMe-oF

Rack-scale
Remote NVMe SSDs over
RDMA or FC fabrics.

2014 - 2016

- **Defined NVMeoF**
- First NVMe-rack-scale storage solution

NVMe/TCP

Cloud & Hyper-scale
NVMe across the data-center
TCP/IP fabric

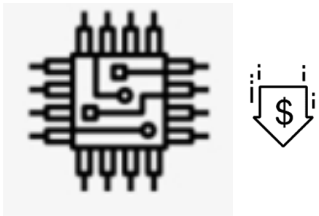
2017 - 2018

- **Pioneered NVMe/TCP** in collaboration with industry leaders
- Contributed initiator code to upstream kernel
- First NVMe/TCP product

Data Center and Market Trends

Co-dependence driving rapid change

Flash Storage



- Flash cost dropping
- ...but QLC needs special treatment

Networking Technology



- Bandwidth is cheap!
- Low latency the norm
- Speeds rapidly increasing

Applications



- Storage growing faster than compute
- Move to Cloud native applications

Private Cloud



- Operational efficiency
- Hyperscale innovation is tomorrow's enterprise practice

Need: simple & efficient storage scaling independently of compute

Transition to Cloud-Native Applications

NoSQL, In-memory, Distributed



PostgreSQL

They All Share:

- Need for low latency and high bandwidth
- Need for consistent response time
- Deployed on local flash (NVMe)
- Usually bare-metal

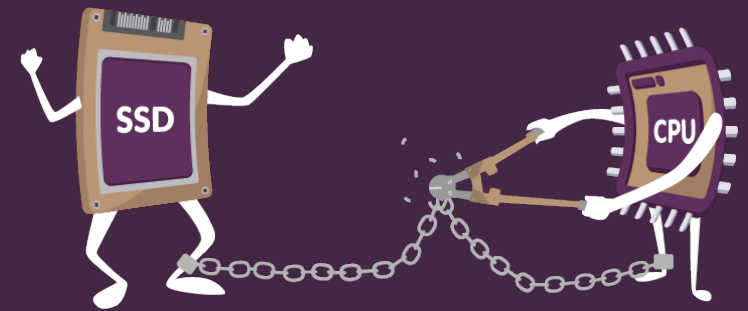
They All Suffer From:

- Poor local flash utilization
- Long recoveries detrimental to network
- Physical ties to application servers
- Flash endurance issues

LightOS: The Modern Data Center Storage Solution

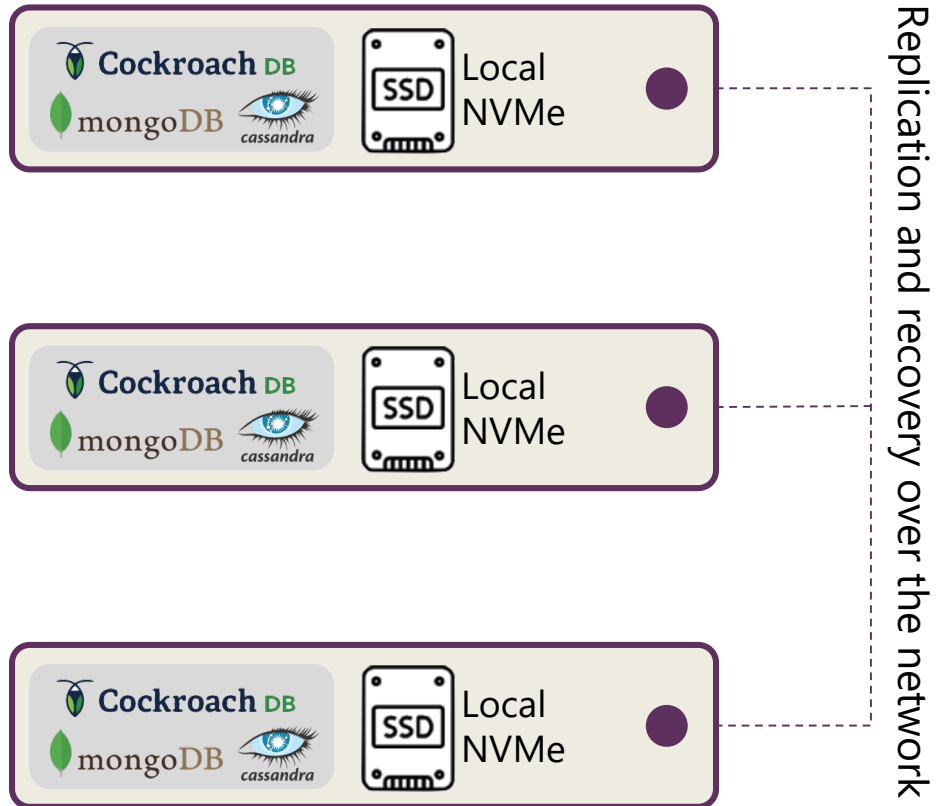
Disaggregated, virtualized NVMe over TCP that performs like local flash

- High performance **software-defined storage**
- **Scale** storage & compute independently
- **Standard infrastructure:** Client/Servers, networking, protocol
- **Reduce cost:**
 - Maximize utilization
 - Increase Flash endurance
 - Improve Operational efficiency
- **Rich data services in software or hardware** (optional)



Scale-out, Distributed Databases

One example, many different kinds



Pros:

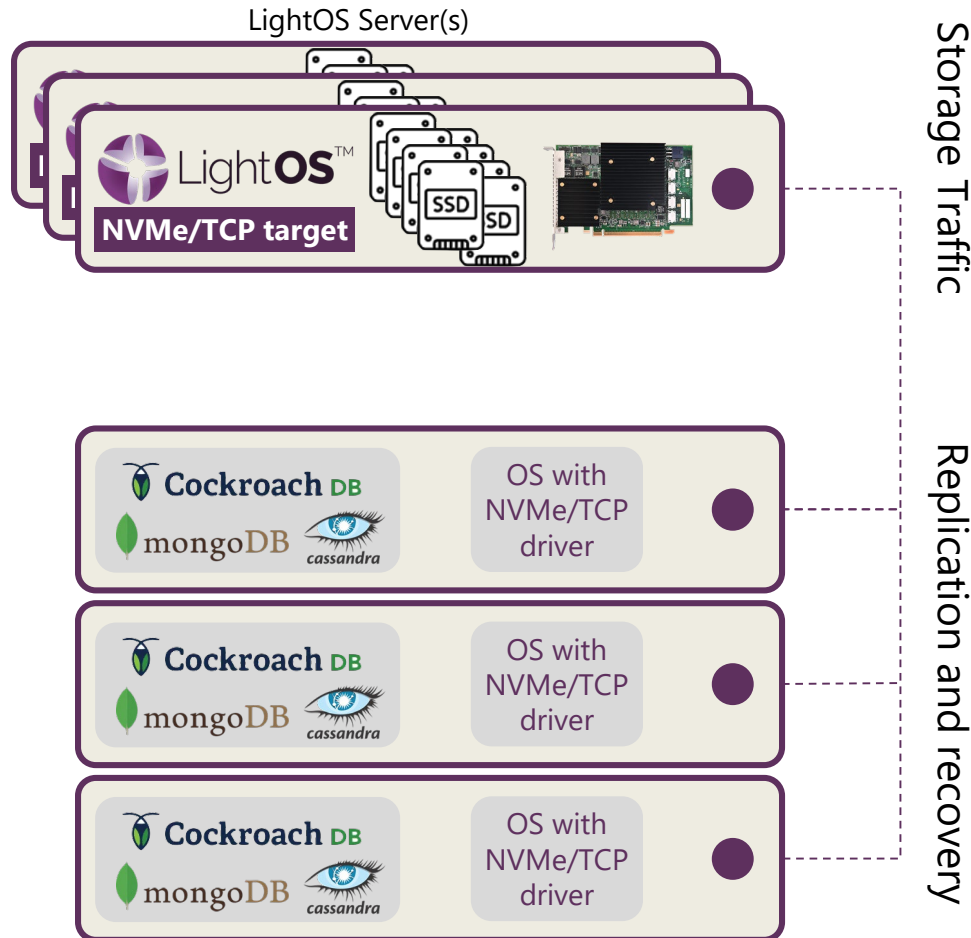
- Local NVMe is very fast:

Cons:

- Limited to one drive
- Tendency to choose a larger drive
- When a drive fails:
 - Severe network impact during rebuild
 - Degraded service for potentially many hours
 - Difficulty meeting SLAs
- If a server fails, full drive rebuild required

Scale-out, Distributed Databases with NVMe/TCP

More uptime, faster recovery, higher utilization = lower TCO



Benefits:

- Local NVMe drive performance
- Drive failures have no effect on applications
- Volumes can be any size and thin provisioned
- On server failure, a replacement instance can be launched anywhere: recovery takes seconds to minutes
- Lessens network impact
- Higher uptime, increased SLAs
- Reduced SKU count

Summary - Lightbits Differentiation

Utilize disaggregated NVMe SSDs as a remote low-latency pool at local flash speeds



Bring Your Own Hardware

- **Software-defined solution**
- **Runs on commodity storage server HW**
- **Standard NVMe SSDs**



Global Flash Management

- **Thin provisioning**
- **Compression**
- **Data protection**
- **QoS services**
- **Higher Endurance**
- **QLC optimized**
- **Optional HW accelerator**



No change in network (NVMe/TCP)

- **TCP/IP network infrastructure - ubiquitous, simple & efficient**
- **Run on standard Ethernet NICs**



Don't touch clients (Target side solution)

- **No proprietary client software**
- **High availability targets with ANA**
- **Standard NVMe/TCP client driver**



Excelero

- ▶ Tom Leyden
VP, Corporate Marketing
www.excelero.com

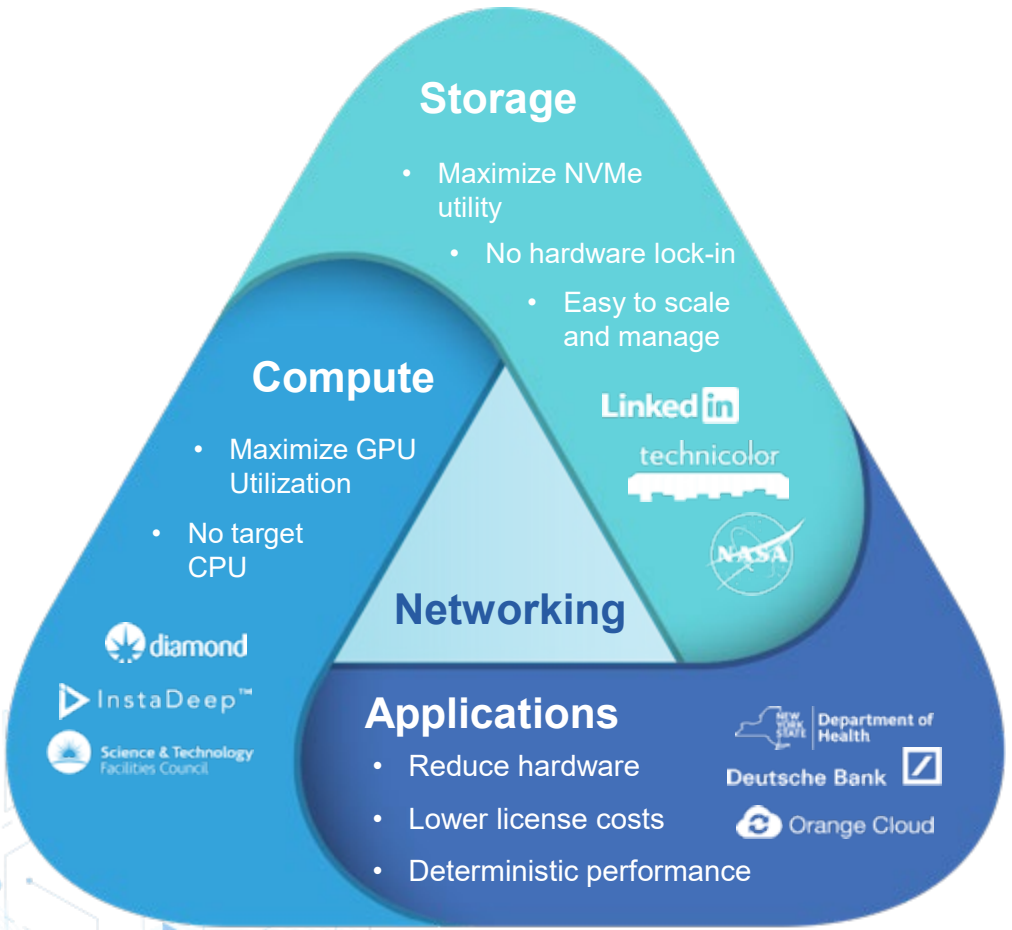


Excelero Elastic NVMe



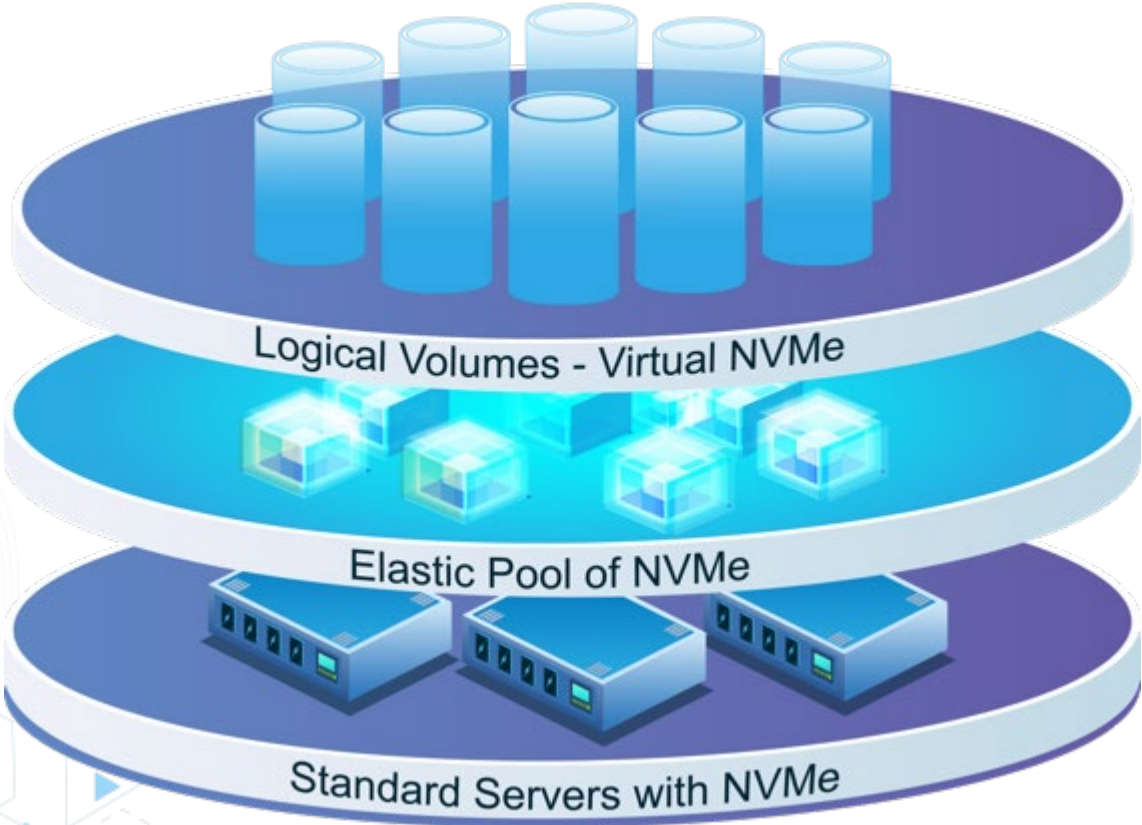
Modern data centers run increasingly data-intensive workloads

Cloud, HPC and AI are converging into the High-performance Datacenter



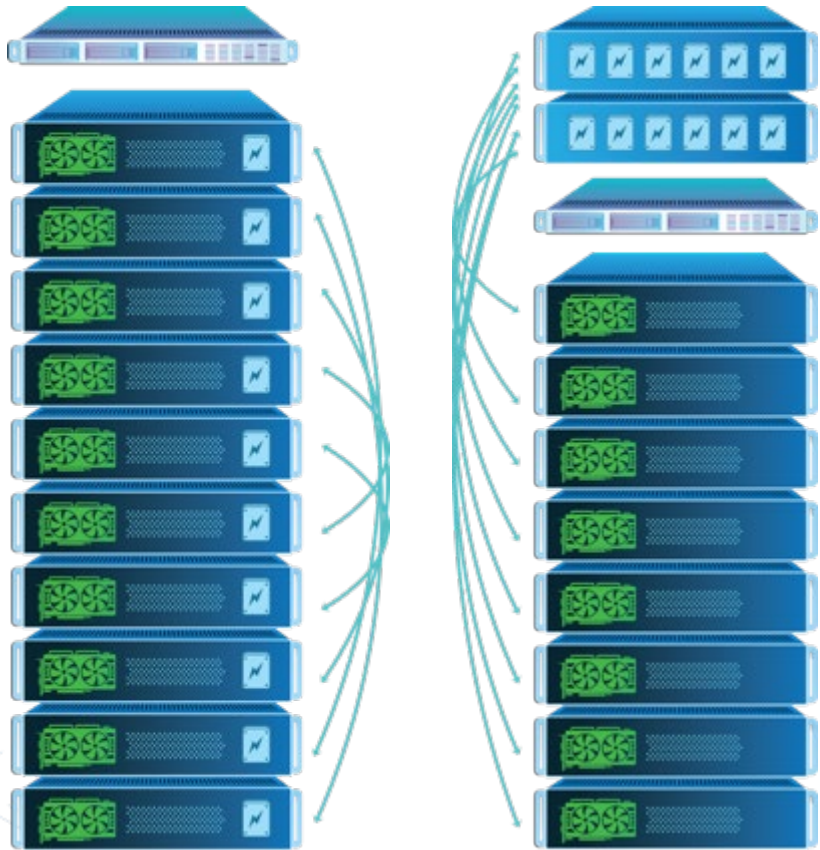
Modern Data Centers – Cloud and Edge
Are adopting new AI, ML & HPC workloads
Double capacity & processing every 18 months
Need to maximize data center ROI: compute, storage, networking, applications

NVMeSH is SOFS



- Elastic NVMe through NVMe virtualization
- Distributed architecture: no controller bottleneck
- Local NVMe performance across the network
- Works with any local or distributed file system
- Leverage RDMA, TCP/IP or NVMe-of

NVMeSH is SDS



- 100% Software-only – leverage any hardware
- Lowest overhead – maximize datacenter ROI
- Separates the data plane, control plane and management plane
- Converged or disaggregated architecture
- Choose any data protection

Successful with partners

Tens of production customers!





Thank you!

 excelero

▶ Panel Discussion

▶ Panel Question #1

- ▶ Each of you has deployed a large number of SOFS solutions. What do you see as the leading use cases for SOFS?
 - Weka.IO
 - Lightbits Labs
 - Excelero

Audience Survey Question #1

- ▶ Does SOFS look like a solution that makes sense for your organization?
(check one):
- Definitely; we have several potential SOFS use cases: %
 - Probably; we have one/a few SOFS use cases: %
 - Maybe; would likely evaluate SOFS as a solution: %
 - Probably not; we would likely continue to acquire array-based storage solutions: %
 - Don't know/haven't studied SOFS as a solution: %

▶ Panel Question #2

- ▶ What best practices would you suggest to a potential customer using a SAN or NAS storage solution today who is looking to migrate to an SOFS solution?
 - Lightbits Labs
 - Excelero
 - Weka.IO

Audience Survey Question #2

▶ How do you see SOFS fitting into the overall storage solutions environment? (select one):

- It is a game-changer – we see it as the leading way to deploy storage in the near-future: %
- It is very important – we see it eventually replacing classical storage arrays over time: %
- It is a useful “tool in the toolbelt” that gives our organization new storage options: %
- It is interesting, but is likely a “niche” technology: %
- We don’t see it as a relevant technology: %

▶ Panel Question #3

- ▶ Hybrid cloud and private cloud architectures are becoming commonplace in enterprises today. What advantages do SOFS solutions have over array-based storage solutions in this regard?
 - Excelero
 - Weka.IO
 - Lightbits Labs



Thank You For Attending

A small orange triangle pointing to the right, located on the left side of the slide, below the dotted line.

G2M
RESEARCH