
WEKA
G2M
RESEARCH
Excelero
AIC
G2M
RESEARCH
AI/ML Storage: Distributed vs Centralized Architectures
Multi-Vendor Webinar
Tuesday August 17, 2021

G2M Research Introduction and Ground Rules

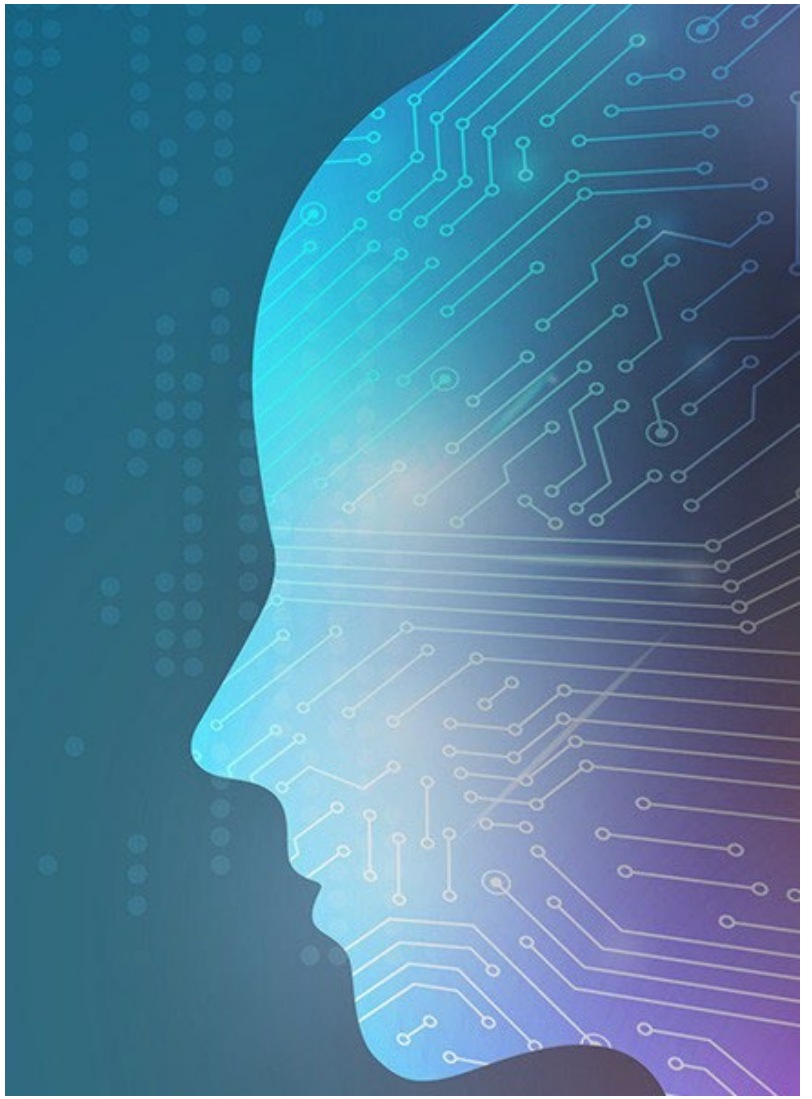
Mike Heumann
Managing Partner, G2M Research

Webinar Agenda

9:00-9:05	Ground Rules and Webinar Topic Introduction (G2M Research)
9:06-9:30	Sponsoring Vendor presentations on topic (10 minute each)
9:31-9:36	Panel Discussion Question #1
9:37-9:37	Audience Survey #1
9:38-9:43	Panel Discussion Question #2
9:44-9:44	Audience Survey #2
9:45-9:50	Panel Discussion Question #3
9:51-9:58	Audience Q&A (8 minutes)
9:59-10:00	Wrap-Up

Artificial Intelligence and Storage

- Artificial Intelligence (AI) and Machine Learning (ML) have unique storage needs
  - Extremely large datasets for learning (up to or over PB in size)
  - Large datasets for model validation
- Training and validation datasets must be cataloged and archived for future analysis
  - Exploring the impact of model changes on previous datasets
- The “live” data lake must have an extremely fast connection to compute resources



AI/ML Storage Architectures

- Most storage vendors are now optimizing their architectures for AI workloads
- Both cloud and on-prem solutions are now available for AI workloads
- These systems must provide large, scalable storage with high performance
- These storage architectures must also be able to provide data management to store training data sets and training results

Panelists

WEKA

Shimon Ben David
Chief Technical Officer
www.weka.io

Excelero

Kirill Shoikhet
Chief Technical Officer
www.excelero.com

AIC

Joe Kimpler
Office of the CTO
www.aicipc.com

G2M RESEARCH

Mike Heumann
Principal Analyst
www.g2minc.com

WEKA

Shimon Ben David

Chief Technology Officer

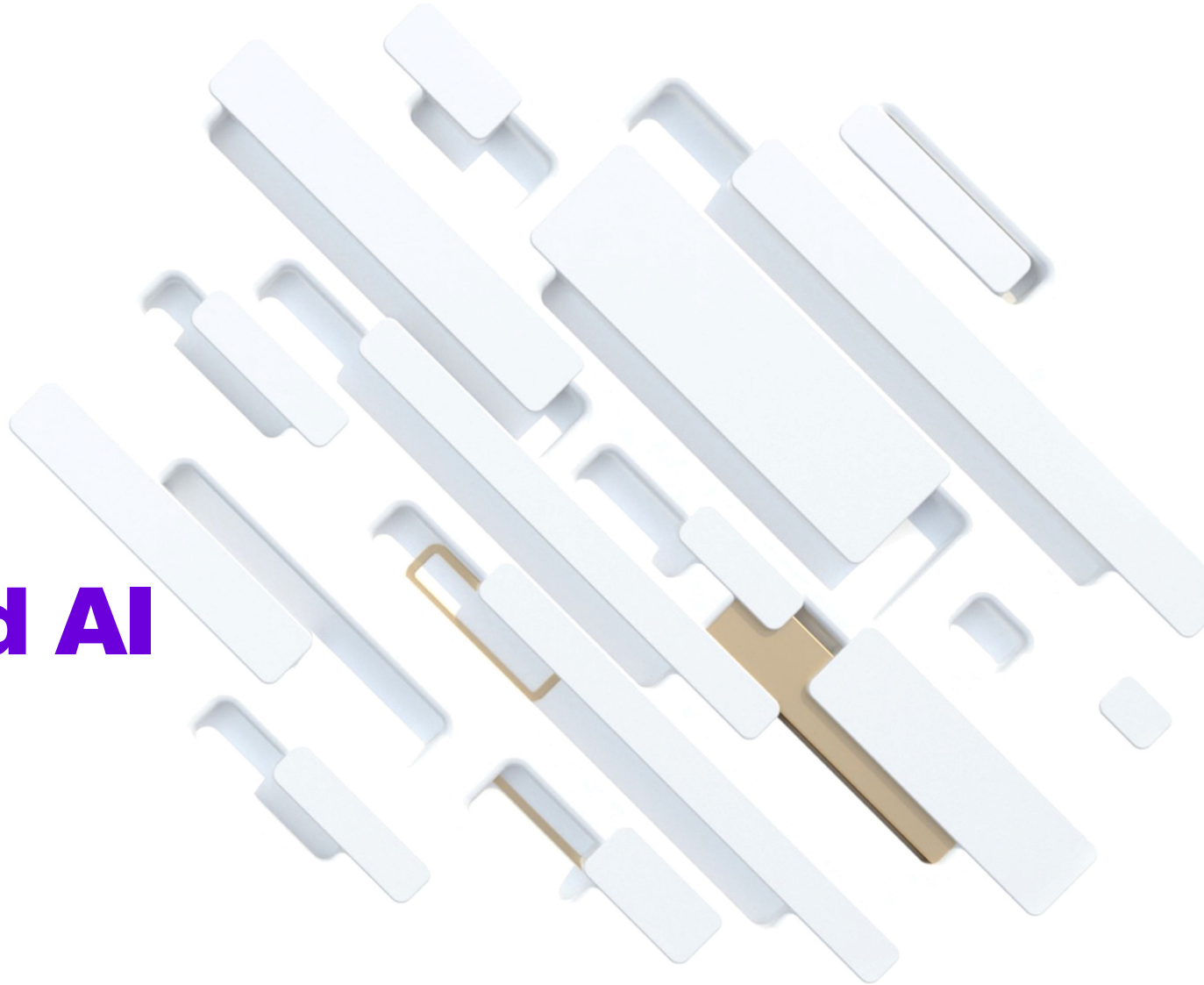
www.Weka.io

WEKA

Weka for Distributed AI Storage

Shimon Ben David

Chief Technology Officer



WELCOME TO WEKA

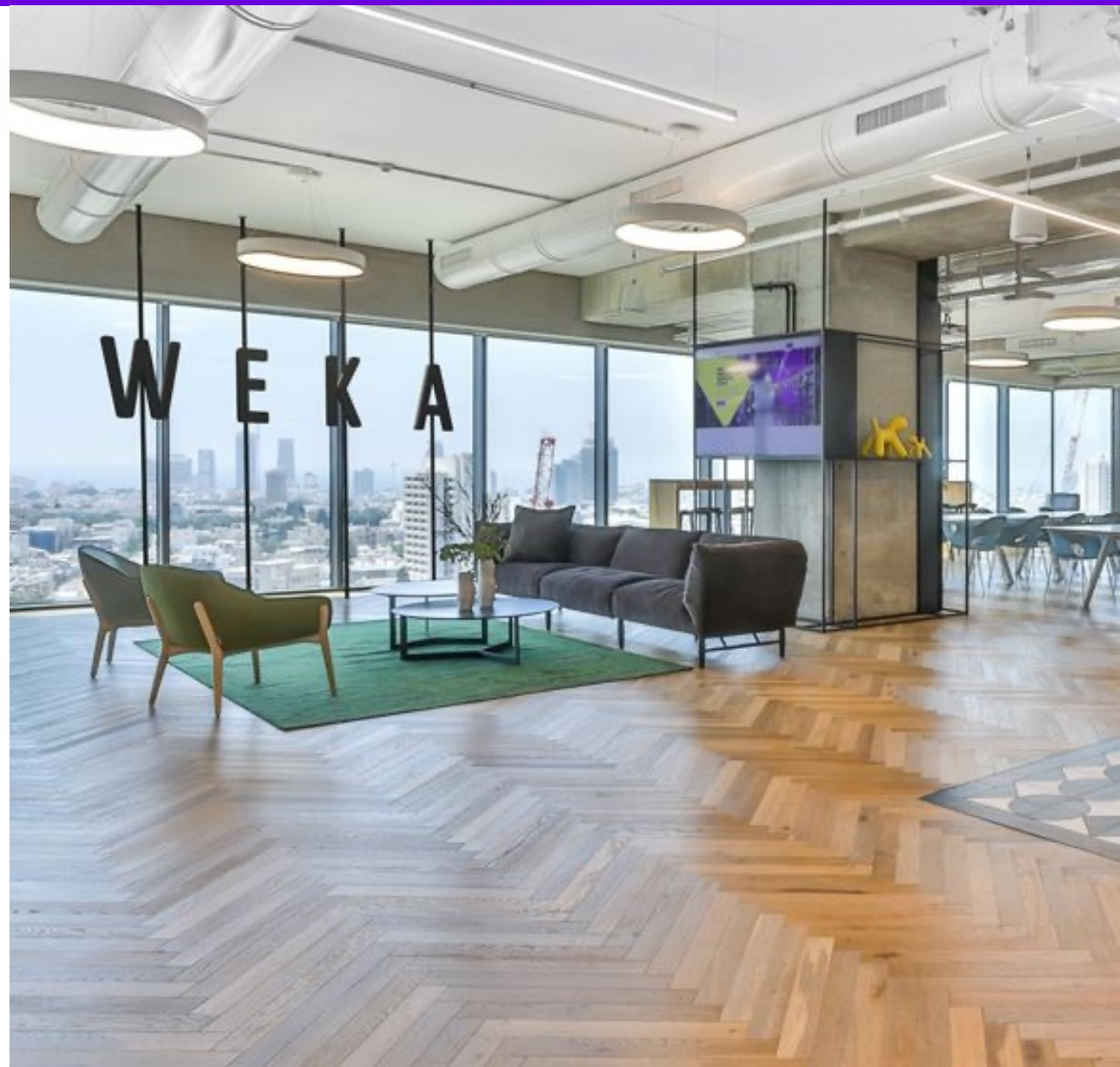
Our Mission

Make storage a utility by delivering simplicity, speed, scale, and better economics

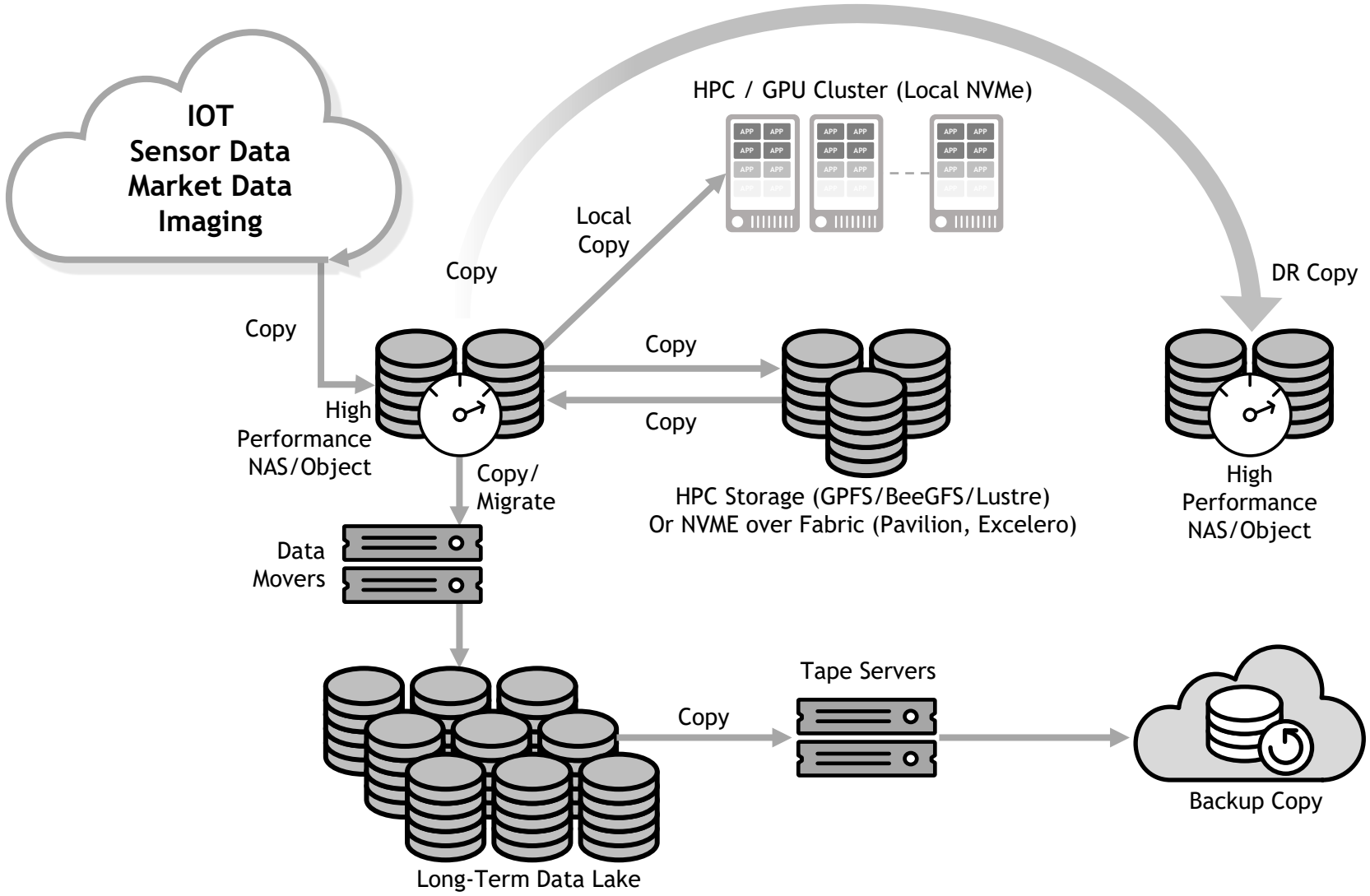
8 of the Fortune 50 are customers

Backed By Industry Leaders

cisco · Hewlett Packard Enterprise · NVIDIA · Mellanox TECHNOLOGIES · Qualcomm · Micron · Western Digital · SEAGATE



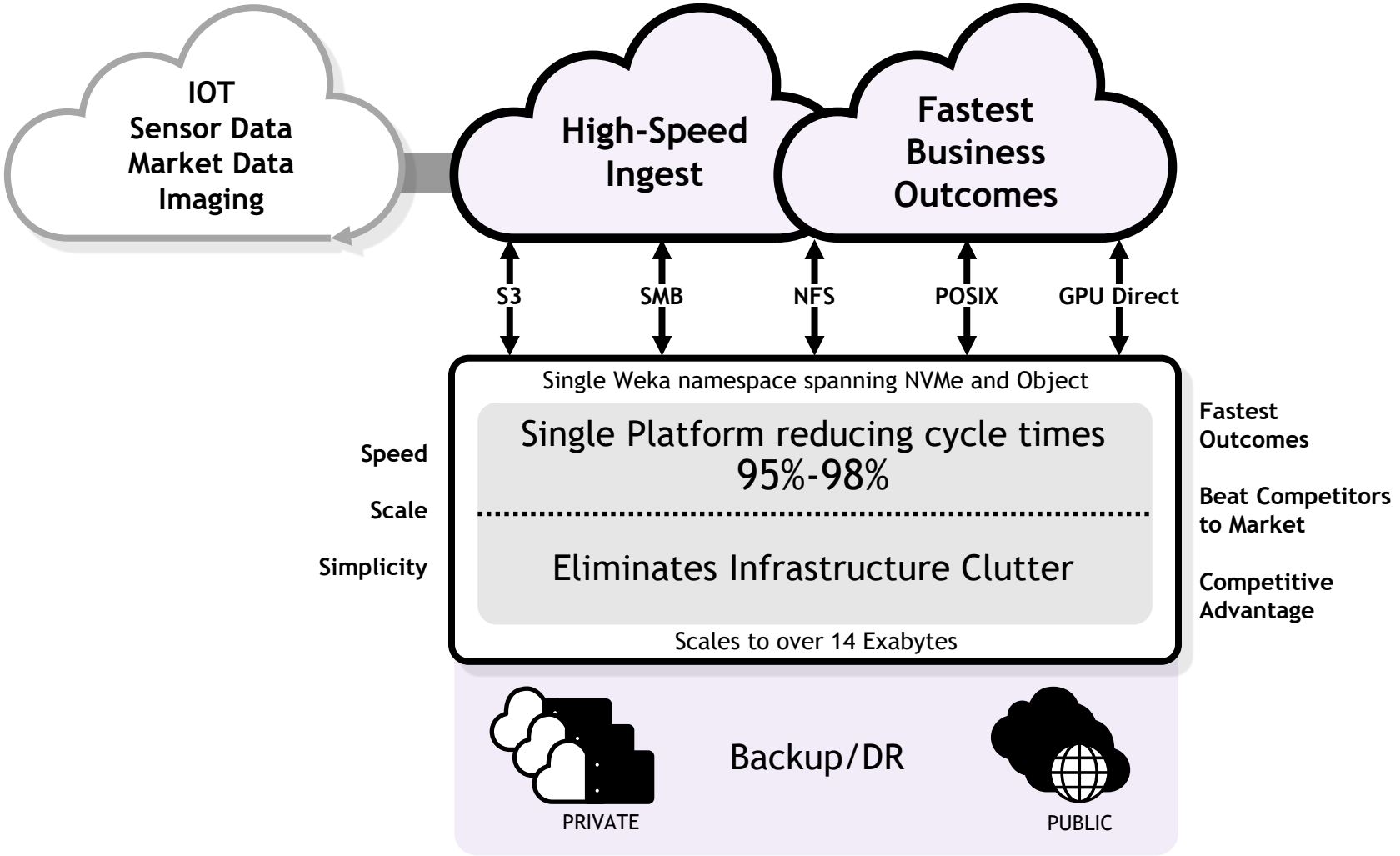
Existing Problem: Multi-Copy Architecture



Before Weka

1. High cost of infrastructure (Need of multiple systems to store and run data)
2. Slow time for data to become actionable
3. Limited scalability (cold data needs to be moved to different systems – additional software required)
4. High management overhead (multiple systems need to be purchased, managed, powered, cooled & housed in Data Centers)

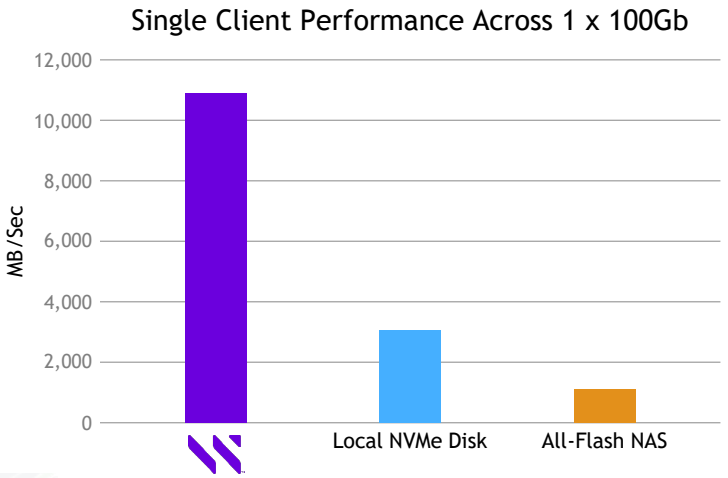
Solution: Zero-Copy Architecture



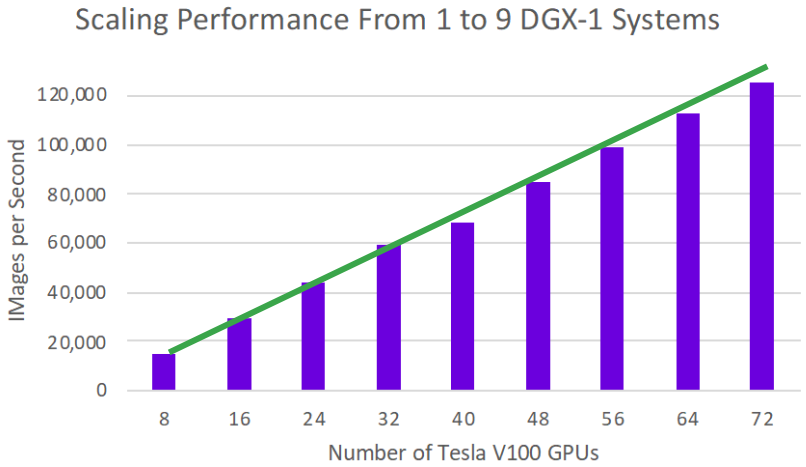
After Weka

1. Collapse into single system, reducing infrastructure costs by up to 75%
2. 95% - 98% reduction in time for data to be actionable (first to market advantage)
3. Immense scalability up to 14EB's in single namespace (no additional software needed)
4. 20x reduction in management overhead

WekaFS PERFORMANCE



162GB/Sec & 1M IOPs Performance to a Single GPU client



HIGH-SPEED NETWORK SATURATION

MASSIVE SINGLE CLIENT PERFORMANCE

Benchmark	DAS (Optane SSD Server)	SAN (NVMe-oF)	NAS (All-Flash)	WekaFS (HPE NVMe Servers)
100T.YR1VWAB-12D-HO	15633	1886	4183	1028
100T.YR2VWAB-12D-HO	18114	1418	3294	892
100T.YR3VWAB-12D-HO	20730	1910	4773	1141
100T.YR4VWAB-12D-HO	24741	3317	7037	1550
100T.YR5VWAB-12D-HO	36888	22389	11376	4808
10T.YR2-MKTSNAP	176	355	6898	655
10T.YR3-MKTSNAP	176	358	7855	675
10T.YR4-MKTSNAP	149	375	8531	711
10T.YR5-MKTSNAP	155	393	8684	726
1T.2YRHIBID	645	374	1419	309
1T.3YRHIBID	1129	630	2737	480
1T.4YRHIBID	1957	1082	4881	804
1T.5YRHIBID	3234	1804	8589	1234
1T.OLDYRHIBID	61	46	129	48
1T.YR1VWAB-12D-HO	334	226	545	294
1T.YR2VWAB-12D-HO	394	268	632	355
1T.YR3VWAB-12D-HO	462	347	750	430
1T.YR4VWAB-12D-HO	553	517	928	547
1T.YR5VWAB-12D-HO	841	769	1298	732
50T.YR1VWAB-12D-HO	1089	1748	4302	2300
50T.YR2VWAB-12D-HO	1988	1774	4798	1971
50T.YR3VWAB-12D-HO	2865	2278	6253	2409
50T.YR4VWAB-12D-HO	4195	3118	8840	3077
50T.YR5VWAB-12D-HO	6731	4625	13597	4111
Average Result (lower is better)	**5968.33**	**2166.958**	**5097.041**	**1303.625**

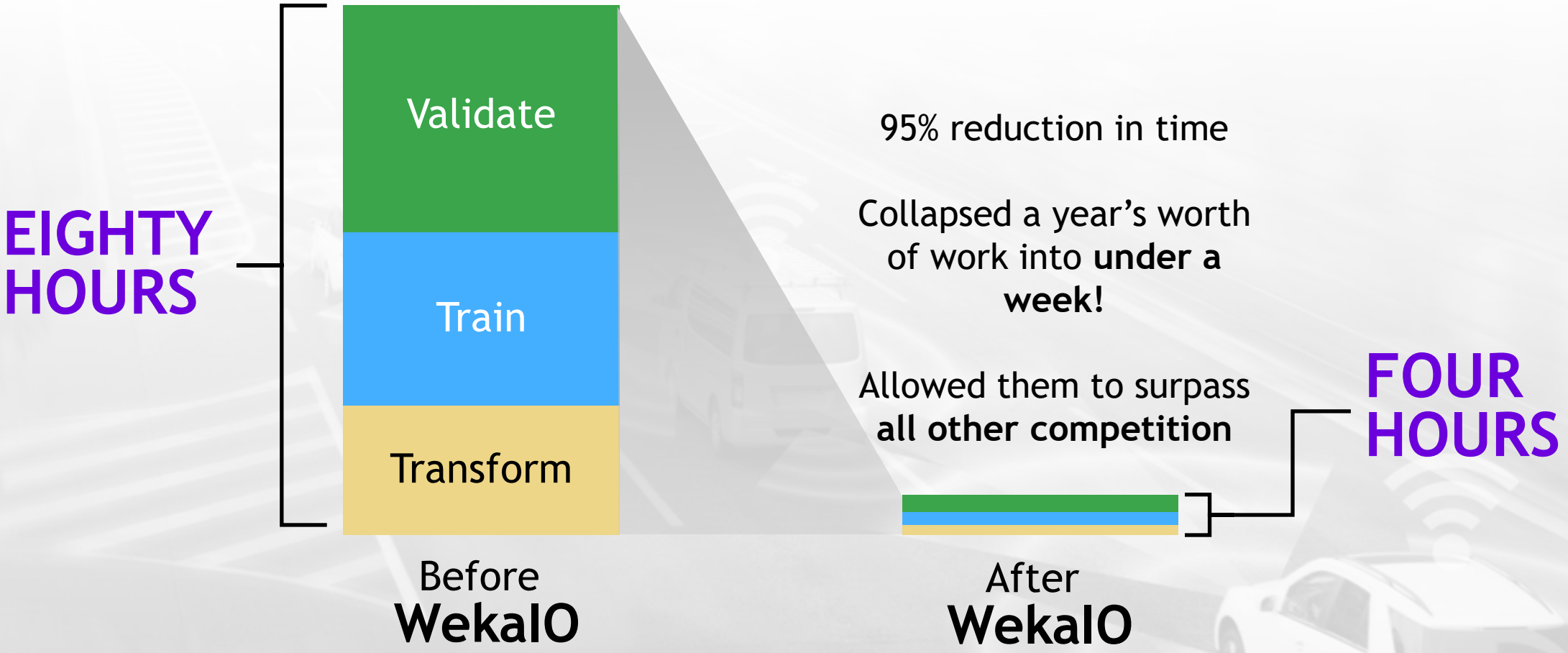
PERFORMANCE SCALES LINEARLY TO MULTIPLE CLIENTS

FINANCE, TICK DATA ANALYTICS APPLICATION PERFORMANCE

- 4X FASTER THAN ALL-FLASH NAS
- 4.5X FASTER THAN DAS WITH OPTANE
- 1.6X FASTER THAN ALL-FLASH ARRAY

ACCELERATING AUTONOMOUS DRIVING REVOLUTION

Large Autonomous Vehicle Company AI Project




WEKA

Excelero

Kirill Shoikhet

Chief Technology Officer

www.excelero.com

Many Facets of Storage for AI

Proper storage solution for AI is a far from static

- Research to Production
- On-premises to Cloud to Hybrid
- Containers to Large-scale Kubernetes
- Research datasets to Data Lakes and Data Governance

But storage characteristics should remain stable so there is a delicate balance between

Predictability  Flexibility

Jack of All Trades, Master of **Some**

G2M RESEARCH

Some real-world AI workload are OK with large IOs and high BW

But there are real-world AI workloads that require low latency

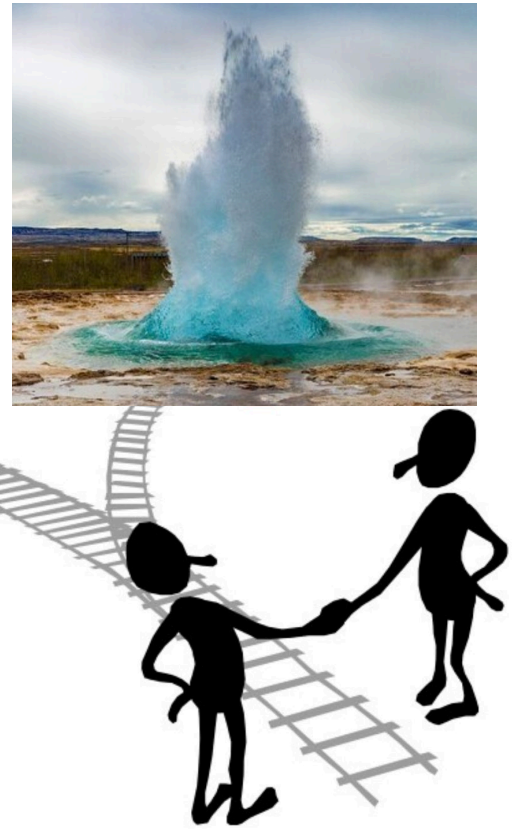
- Efficiently working with small/medium files
- Example: Deep Learning training on images or video fragments

50 Ways to Hide Your Latency:

- High bandwidth numbers using 1MB reads
- Scaling capacity to 10s of PBs to get high IOPS
- Performance benchmarking by obscurity
- …

Divide and Conquer

- Cold object storage for data lake
  - Cloud-based solutions bring highest reliability levels
- Fast and efficient storage for the immediate tasks
  - Higher GPU utilization
  - Faster turnaround times
  - Faster convergence
- Intelligent data management
  - There are no shortcuts without tradeoffs



Storage as an Agile Companion

	Centralized	Disaggregated
Scales easily	Scale-out solutions do scale but the scale is often limited	Software-defined disaggregated solutions are natively scalable
Cloud/Hybrid support	Only in co-located datacenters. Hybrid deployments are challenging	Allows running on top of public cloud IAAS as well as on-premises. Hybrid deployments are feasible
Kubernetes support	Container-ready CSI plugin – storage is deployed alongside Kubernetes clusters. Extremely limited mobility	Container-native deployments. Can run on-premises and on managed Kubernetes offerings on public clouds
	Limiting factor	Enabling factor

Common Myths about Disaggregated Storage

G2M RESEARCH

Frequently cited concerns about disaggregated solutions

Excelero NVMesh

High/Unpredictable latency

Direct [Application Stack] to [NVMe drives] data path and
Client-side Architecture with no East-West traffic ensure

→ low and predictable latency across platforms

Lack of proper data protection

Multi-node, multi-rack rack distributed RAID10 and Erasure Coding (RAID6) tested on 1000-node clusters

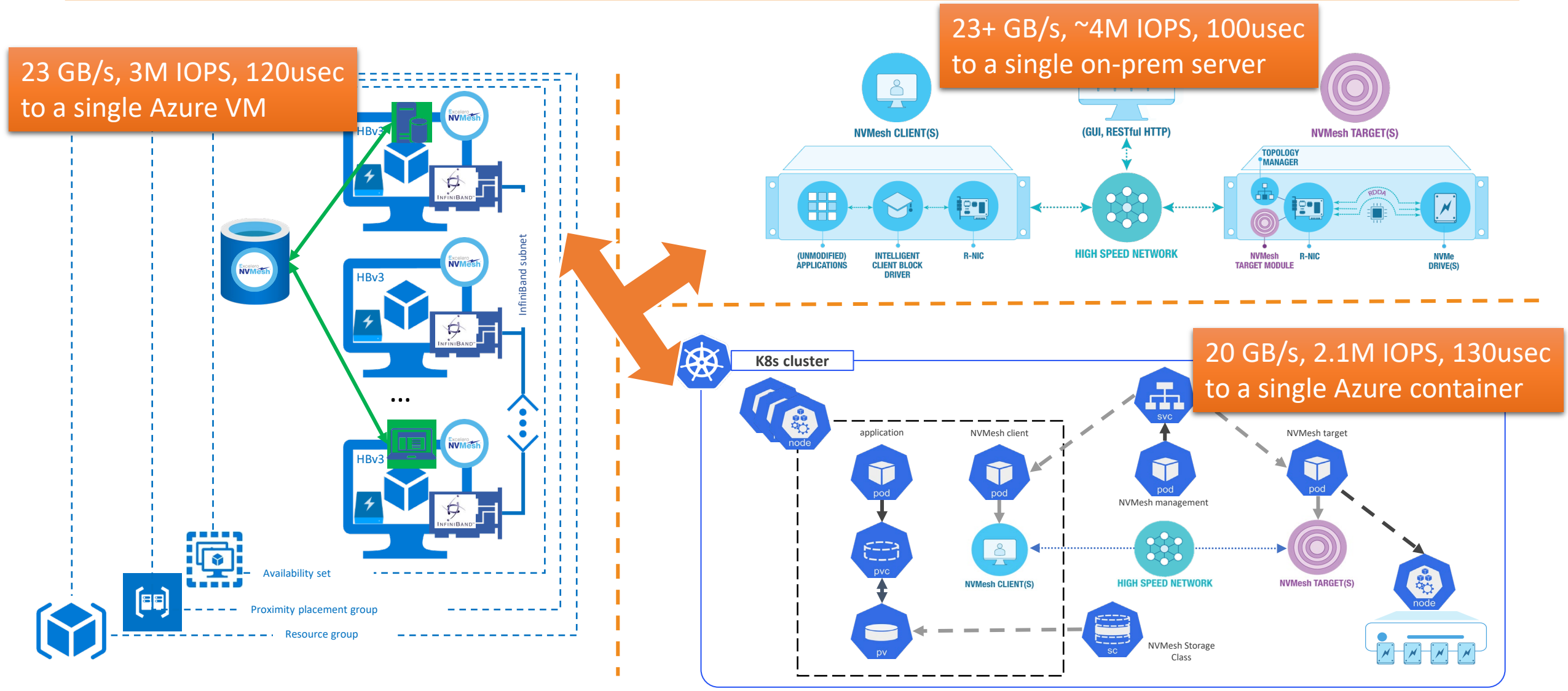
Difficult to integrate w/ file systems

Over 9M file IOPS with IBM Spectrum Scale in SAN mode in just 2 rack units
Full line rate bandwidth with 200Gbps and 400Gbps networking

Setup complexity

Deploy in minutes via standard package managers on-premises;
Kubernetes Operator on K8s clusters;
Native deployment mechanisms on public clouds (ARM, CloudFormation, …)

Predictability Across Environments

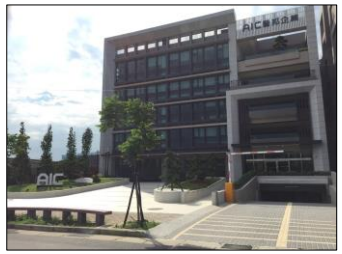


AIC

Joe Kimpler
Office of the CTO
www.aicipc.com

AIC – Who We Are

- AIC is a leading provider of OEM/ODM and COTS, server and storage solutions
- In-house mechanical, electronic, system-level engineering design teams, thermal & system level validation capabilities and world-class manufacturing
- Products are highly flexible and configurable to any form factor, standard or custom
- Founded in 1996, the Global Headquarter is in Taiwan, with offices and operations throughout the United States, Asia and Europe



AIC – What We Do

All-Flash NVMe
High Performance Computing

GPU Server
AI, Deep Learning

Storage Server
High Density,
Performance optimized

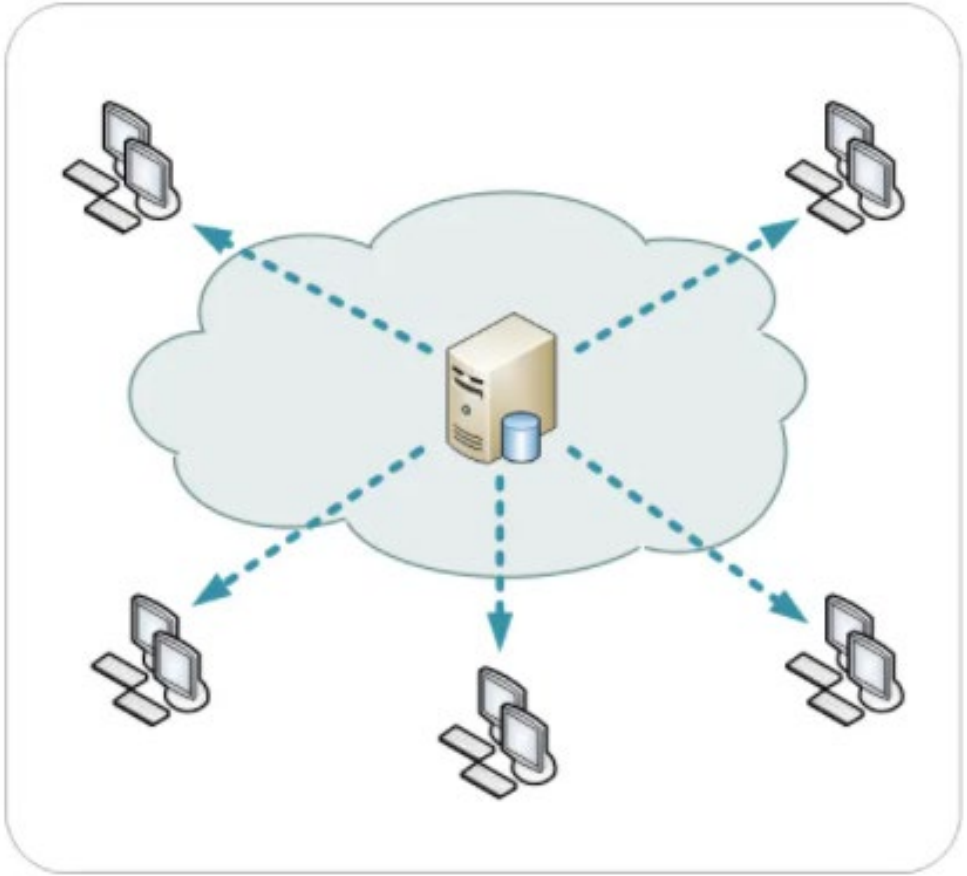
Network Security
FIPS, NEBS Compliant

Multi-node Server
Hyper Converged
Infrastructure System

Rack
Cloud Datacenter Total Solution

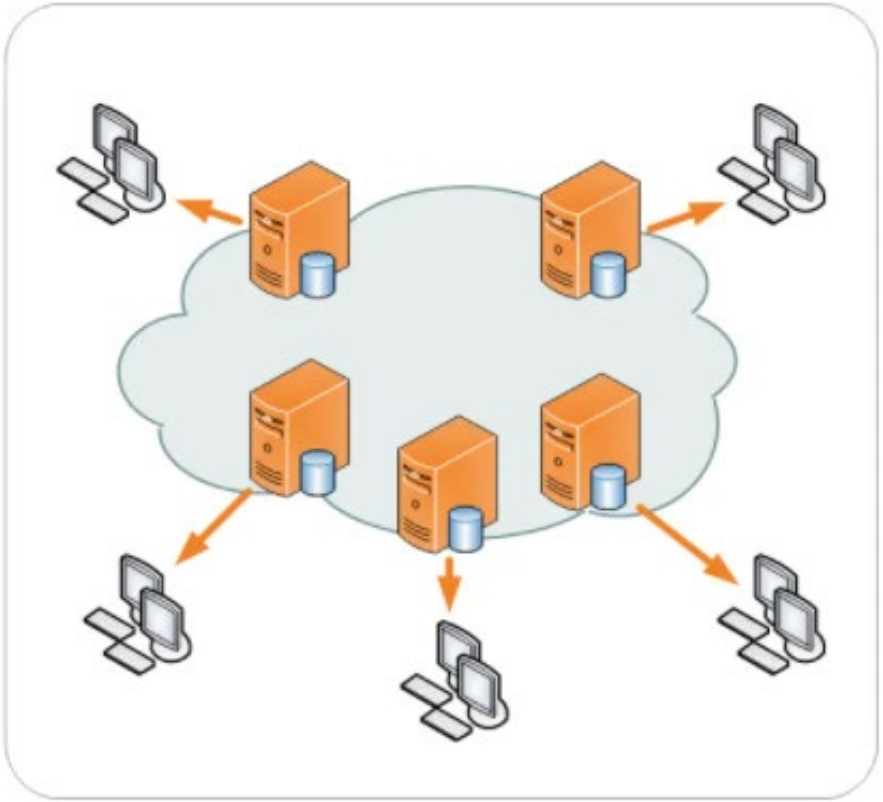
Centralized Storage Characteristics

- **The characteristics of centralized data storage include:**
  - One central system, enables universal communications
  - Presence of a global clock, synchronizing information
- **Advantages of a centralized data system include:**
  - Consistent system, increases data integrity
  - Single location security, including system, data and physical
  - Central location facilitates data access & coordination
  - Can be less expensive to maintain
  - Reduction of distribution, reduces data redundancy
  - The data is easily portable
  - Single location reduces data retrieval time
- **Centralized data system has disadvantages including:**
  - Catastrophic failure can destroy system and data set
  - System down time affects all users
  - Issues can occur when all users want to access the data at the same time

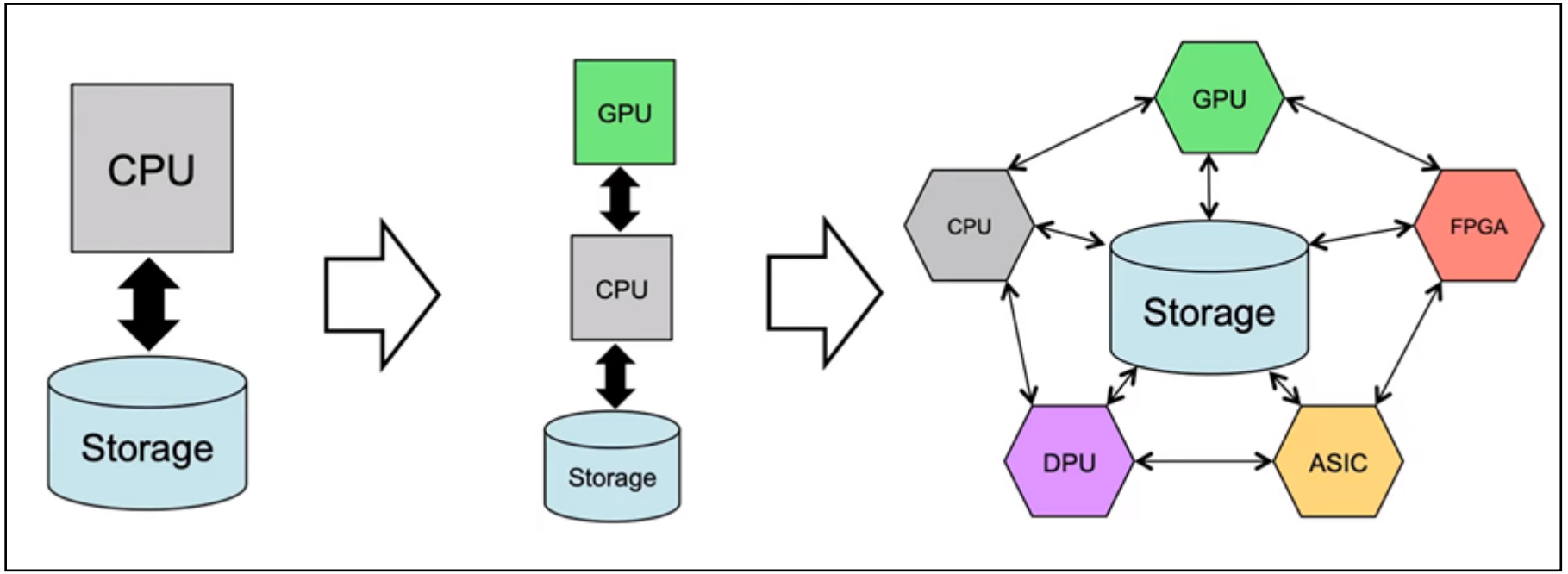


Distributed Storage Characteristics

- **Distributed data storage is characterized by:**
  - Different systems, each running individual clocks
  - Multiple servers or central units
- **Advantages of distributed data storage include:**
  - Since data is spread across physical locations, it can be easily expanded
  - It can be accessed from many networks
  - Multiple locations provide resiliency
  - It has a high-performance capability as the system load is spread over the multiple nodes
- **Disadvantages include:**
  - Due to its complexity, it is difficult to maintain
  - It can increase management costs
  - It may difficult to detect which node has failed hence may take longer to be rectified



Distributed Processing & Storage



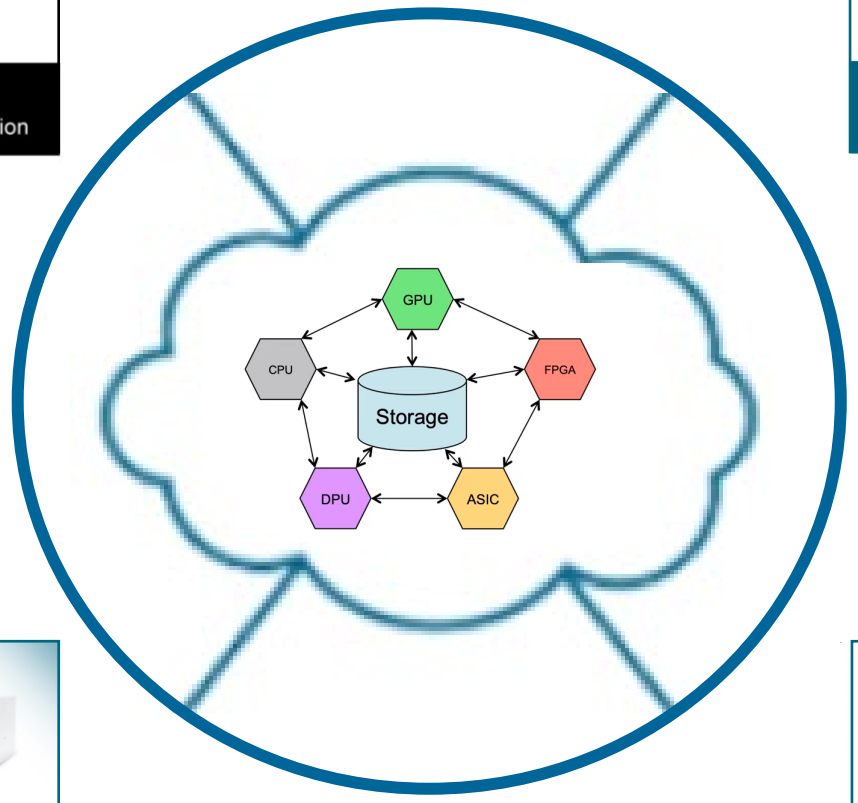
- Storage supplies data to the CPU for processing
- Storage provides data to the CPU, then shipped to the GPU
- Storage provides data to the most efficient compute

AIC Server, Storage & Security Solutions

- **420U OCP Rack**
- Ultra-low Power, 6 Zone-48 Smart Fan Control
- Tool-less
- Redfish Support

- **Computational Storage**
- 36 Gen4 E1.S NVMe SSD
- AMD EYPC or Intel Xeon
- Multiple PCIe Gen4 & OPC Slots

- **1U FIPS/NEBS Compliant Network** Appliance
- FIPS 140-3
- Xeon E3 Scalable Processor



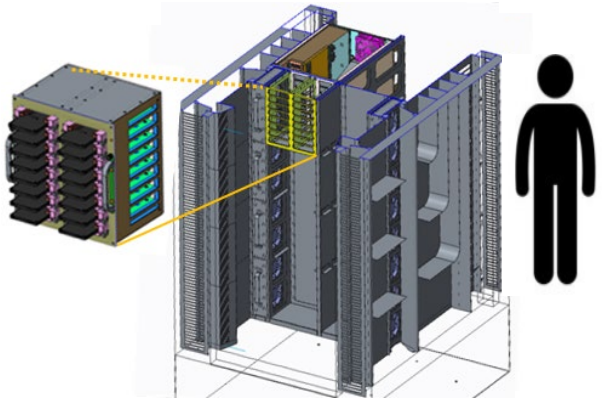
- **High Density 1U4bay FPGA/GPU**
- 6 PCIe Gen 4 & 1 OCP Slots
- 2 Intel Gen 3 Scalable
- 32x DDR4 RDIMM

- **Ultra-density Storage Servers**
- 2U 24-Bay to 4U 102-Bay
- Front end enabled-JBOF's & JBOD's

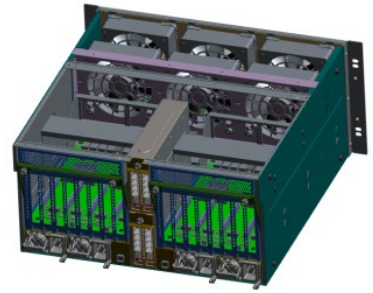
- **HCI Optimized - 2U 4-Node**
- AMD EPYC or Intel Xeon
- 12x 3.5" SAS/SATA + 8x 2.5" SATA Bays
- 3 PCIe X16 Add-on Cards

Why AIC?

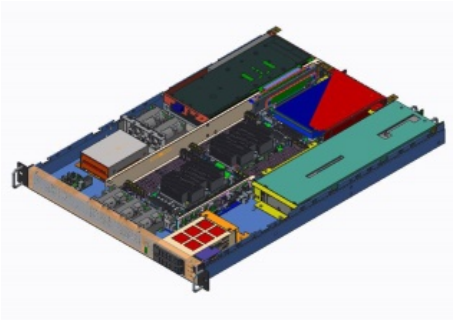
- **Design Standard Solutions to Complex Systems**
  - Experienced Mechanical, Electrical, Software, System Level and Thermal engineering teams design from PCBA to enormous test chambers
  - All manufacturing is in Taiwan, not China



- **Two Platforms** – Intel and AMD (PCIe Gen 4)
  - AIC has products released or in design with both Intel (Cascade Lake Refresh & Ice Lake) and AMD (Naples/Rome & Milan) processors
- **Complete Systems** – Unique ability to customize
  - PCIe Gen 4 & Gen 5 Support – Riser cards can be redesigned for frequency modifications
  - BIOS Customization – Ability to access the BIOS source code for customization
  - Configurability & Routing of PCIe Lanes – AIC's MAX I/O technology can manage the PCIe lanes individually or as a group to meet the design requirements
  - Design excels at accommodating unique requirement for additional cards like SmartNICs, GPUs, Frame Grabbers, etc.
  - Short/small footprint multi-node high density systems
  - FIPS/NEBs compliant designs



- **Extended Life Cycle** – 5-7 years, with extended support, some products having 10+ years
- **Demand Creation** – Partner with customers to create customer awareness & demand



Engineering, Design and Manufacturing

jkimpler@aicipc.com

Panel Questions and Audience Surveys

Panel Question #1

What is the right approach to determining how much storage capacity goes into a particular storage node?

- Shimon Ben David - Weka
- Kirill Shoikhet - Excelero
- Joe Kimpler - AIC

Audience Survey Question #1

How big is your largest AI/ML storage pool today? (check one):

- 1PB or greater: 12%
- Between 250TB and 1PB: 9%
- Between 100TB and 250TB: 6%
- Between 25TB and 100TB: 6%
- Less than 25TB: 21%
- Don’t know/no opinion: 47%

Panel Question #2

How do you balance the needs of long-term storage for archived datasets vs datalake storage for training?

- Kirill Shoikhet - Excelero
- Joe Kimpler - AIC
- Shimon Ben David - Weka

Audience Survey Question #2

How much cloud storage capacity do you use for AI/ML today? (check one):

- Greater than 5PB of cloud storage: 8%
- 1PB to 5PB of cloud storage: 0%
- 100TB to 1PB of cloud storage: 13%
- Less than 100TB of cloud storage: 13%
- We don’t use cloud storage for AI/ML: 29%
- Don’t know 38%

Panel Question #3

NVMe is clearly the fastest flash storage technology today. Where does it fit in the ecosystem of AI/ML storage solutions?

- Joe Kimpler - AIC
- Shimon Ben David - Weka
- Kirill Shoikhet - Excelero

Audience Q&A


Thank You For Attending!