

Standards-based Data Platforms for HPC and AI

Floyd Christofferson

Hammerspace

fc@Hammerspace.com

Oct 25, 2024



Major Industry Trends Driving New Storage Requirements

**LLM Training
Load and Iterate in GenAI**



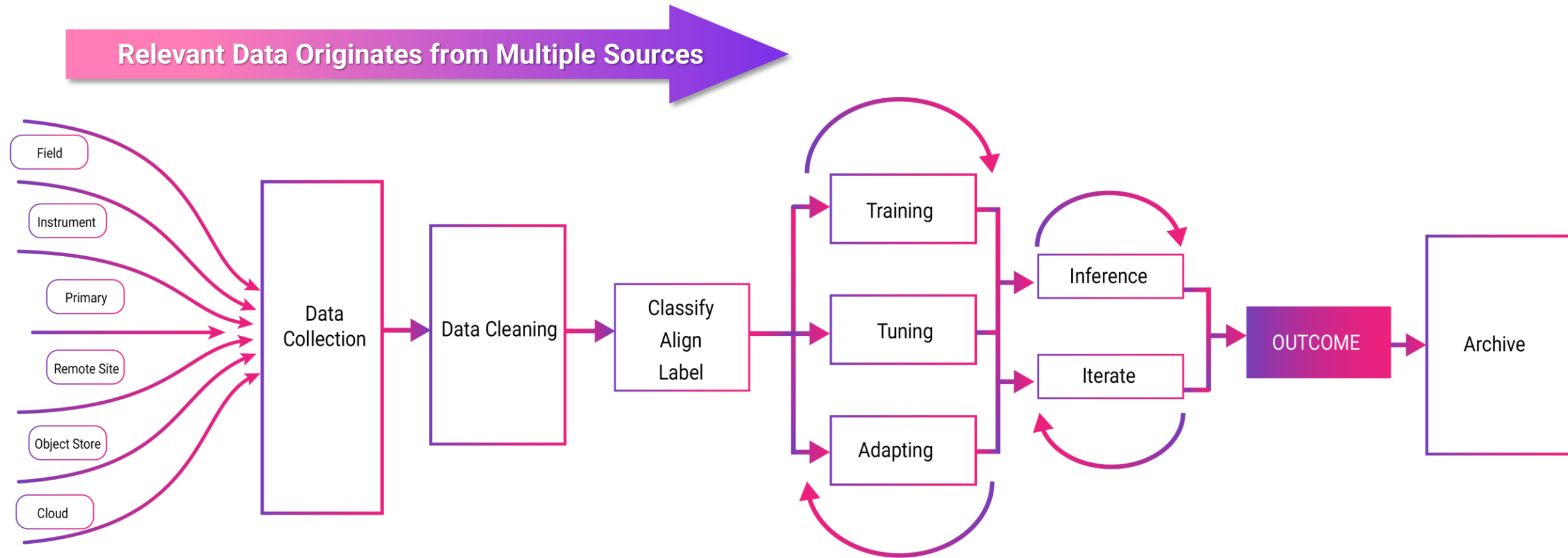
**Large, Decentralized
Data Sets**



**Multi-Site
Multi-Cloud
Remote AI Researchers**



AI/DL Pipelines Need Data From Across Silos



Why? Data Silos Inhibit Progress, Reduce Data Quality



Valuable data
trapped in silos

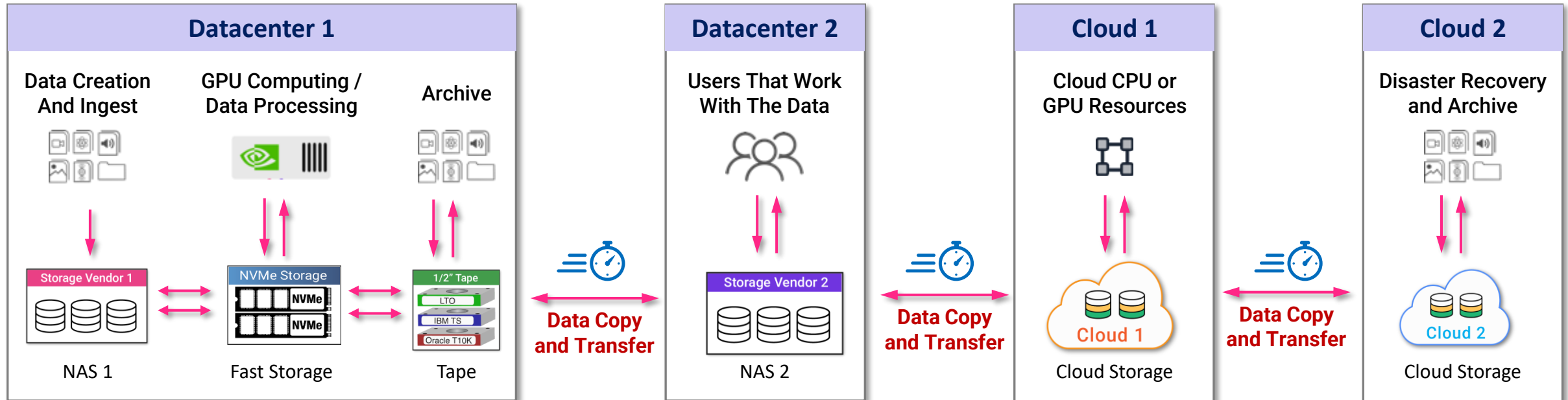
Data copy sprawl **impacts cost, governance and security**

Getting data to global users **takes too long**

Lack performance to keep GPUs utilized

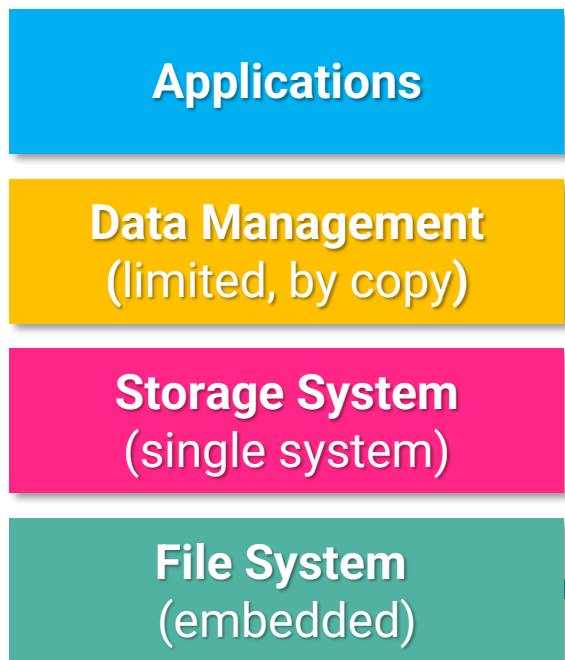
Infrastructure is **not ready for AI**

Lack agility to use elastic cloud resources



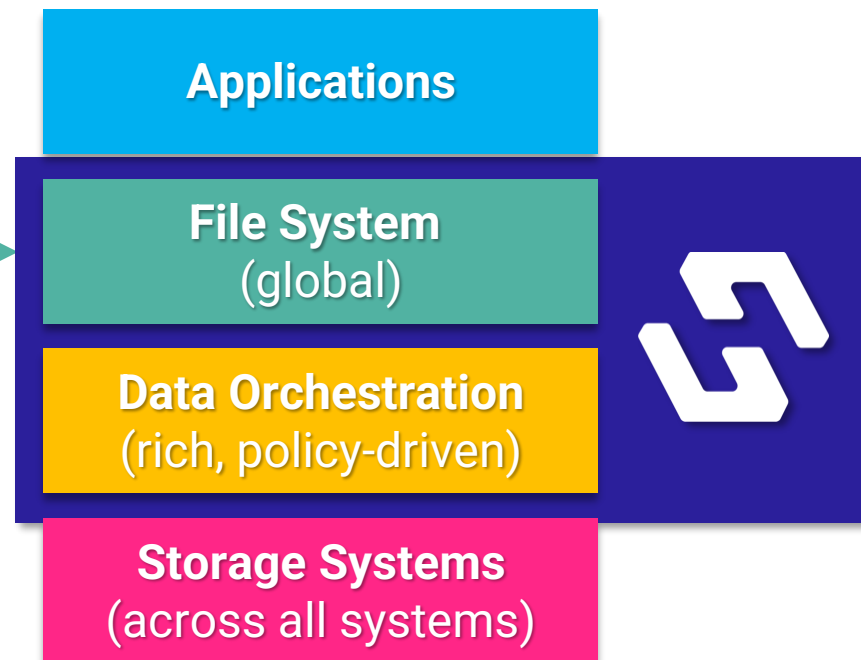
Distributed Workloads Require a New Architecture

Storage-Centric Approach



 Data is Trapped in Silos

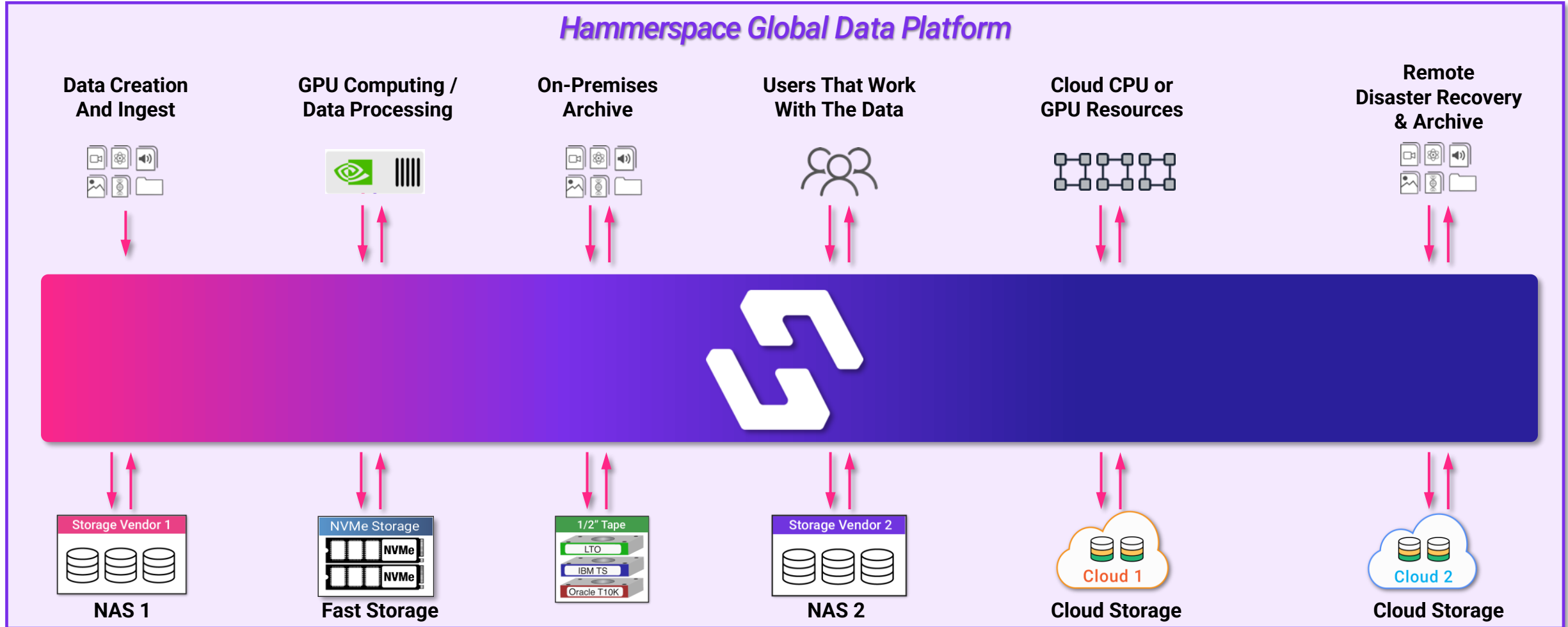
Data-Centric Approach



 Data Becomes a Global Resource

Standards Are A Must for Global Data Use

Unify and automate unstructured data across any data center, any cloud, anywhere



Goal: Consolidate File Systems and Data Sources

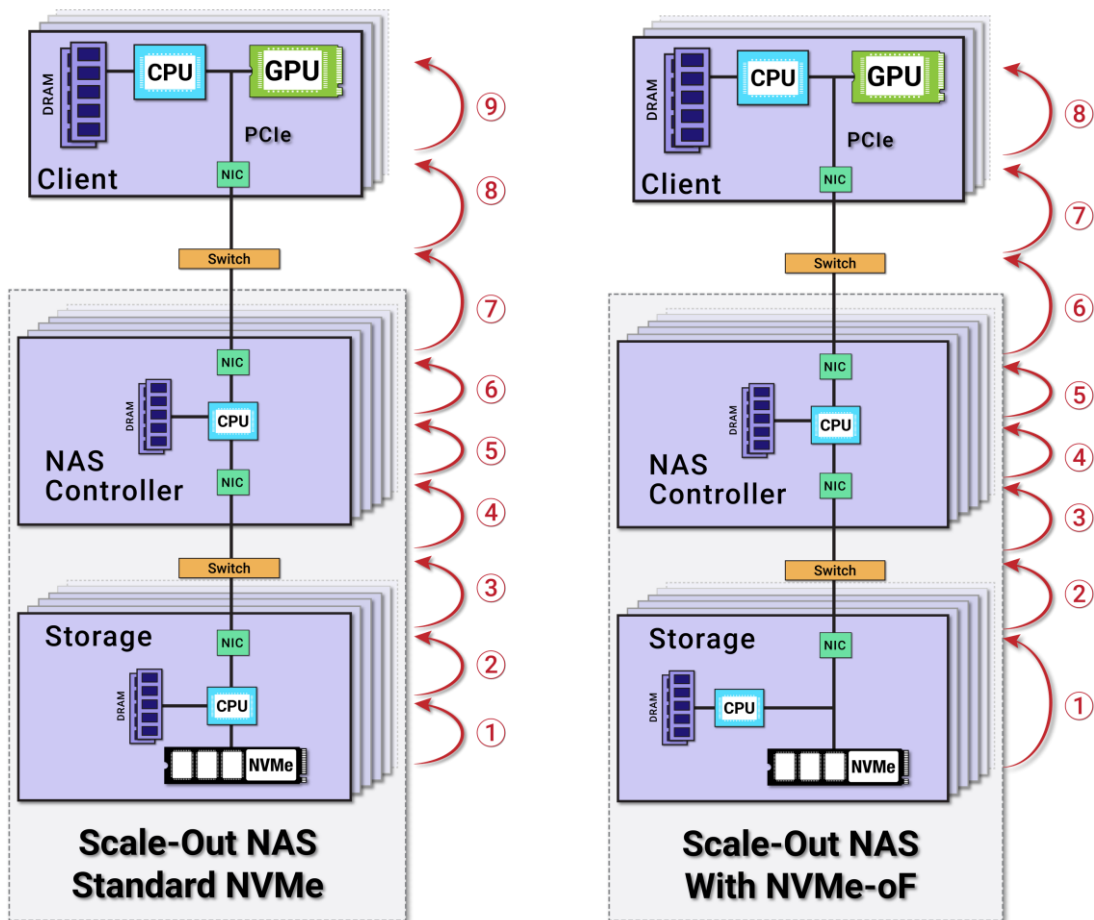


	Traditional HPC Parallel File Systems	Scale-out NAS w/ NVMeoF	Global Data Platform <i>(parallel file system + data orchestration)</i>
Large block, large file read/write	●	●	●
Read intensive	●	●	●
AI random read / write	●	●	●
Data distribution	●	●	●
Archive	●	●	●

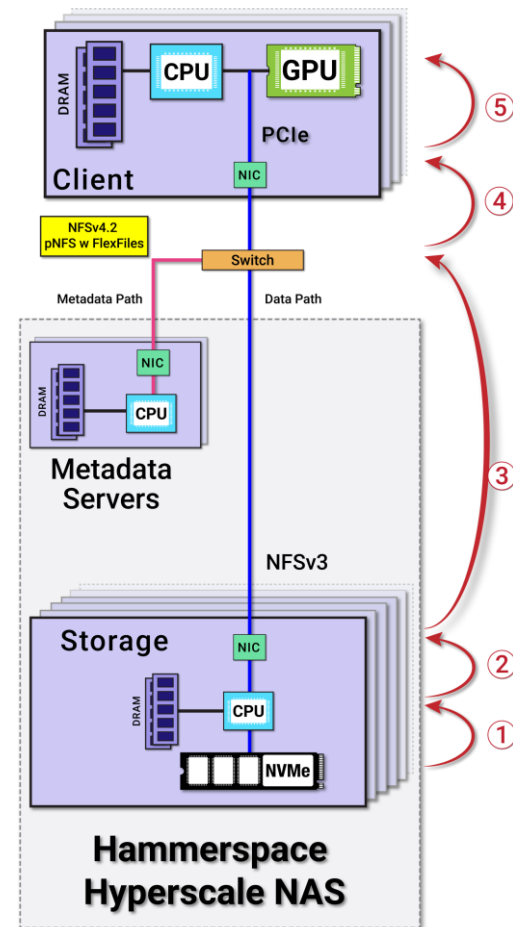


NFSv4.2 Enables Storage Efficiency & Performance

Scale-Out NAS Architectures



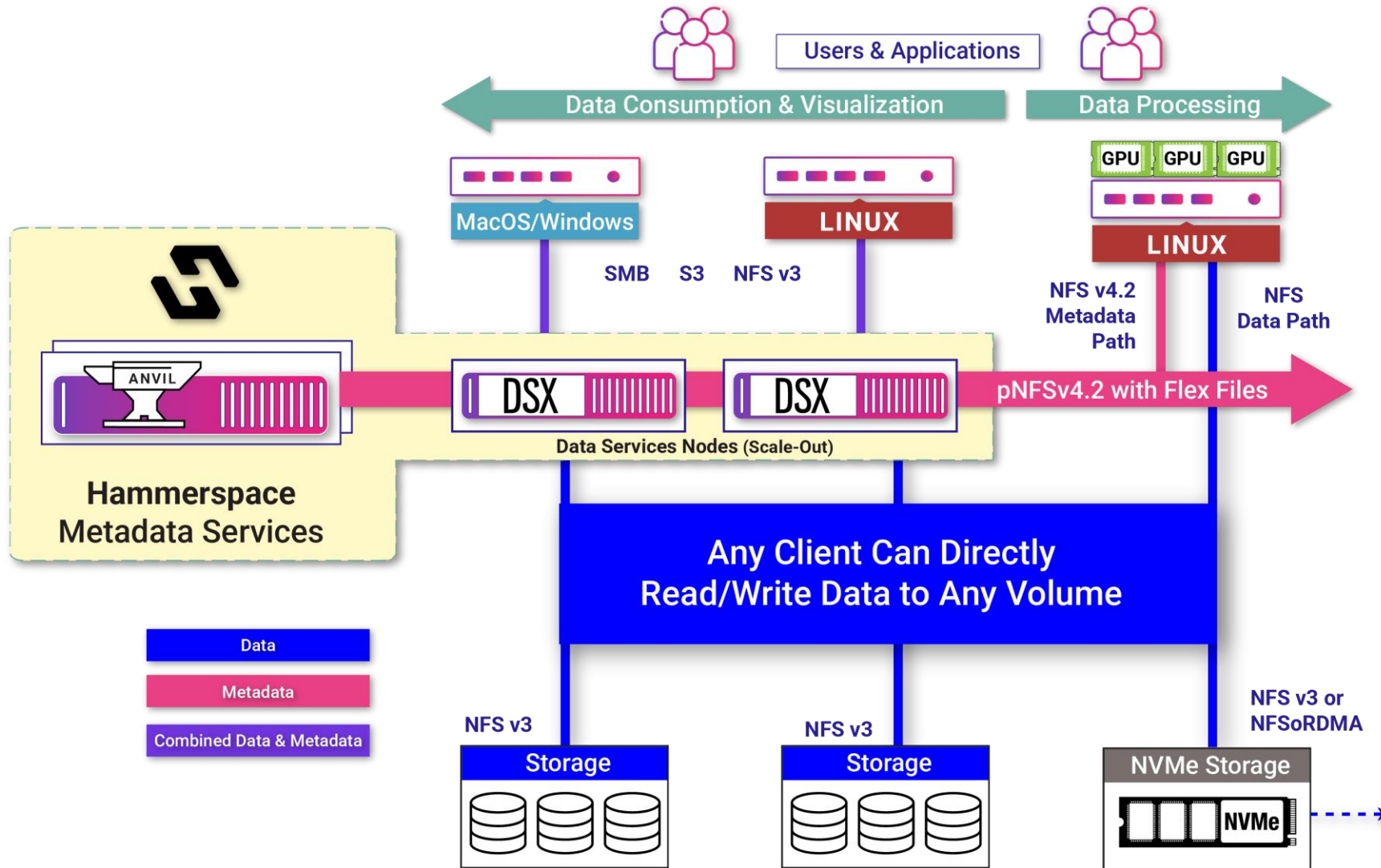
Hyperscale NAS Architectures



2x
Reduction in
Servers,
Networking,
Latency,
Watts and Rack U



Architecture Extends to Address Multiple Use Cases



- Read/write data using standard protocols

- Scale access and performance by adding more DSX Nodes.

- Use any vendor storage including NAS, block, object, cloud, and even tape.

Customer Story: LLM Training

"What Hammerspace does is pure magic."

-Paul Saab, Principal Engineer Meta

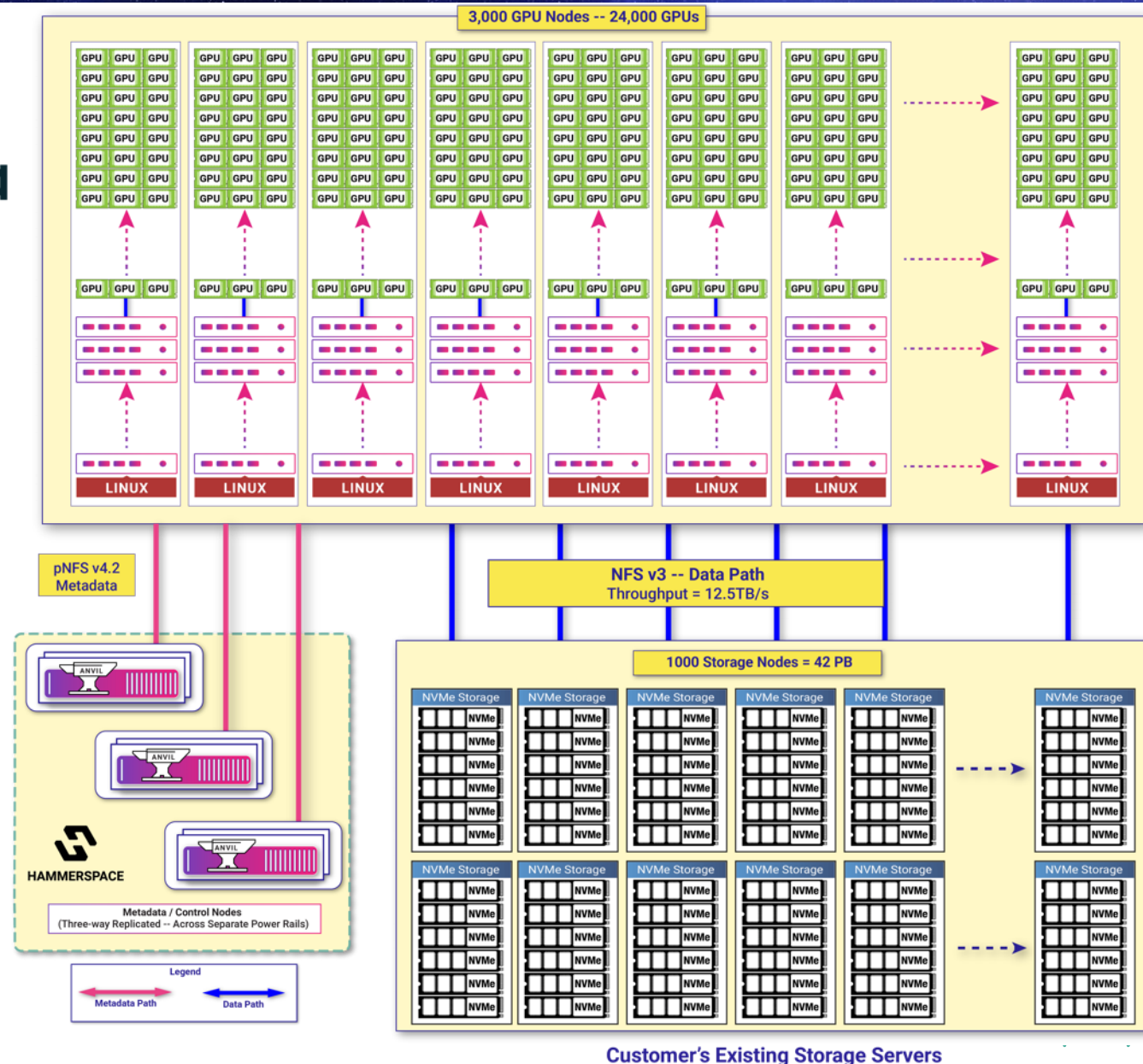
About the Customer

- Largest web property in the world
- Training LLMs and other Gen AI models
- Massive performance and scale demands
- Evaluated leading storage vendors

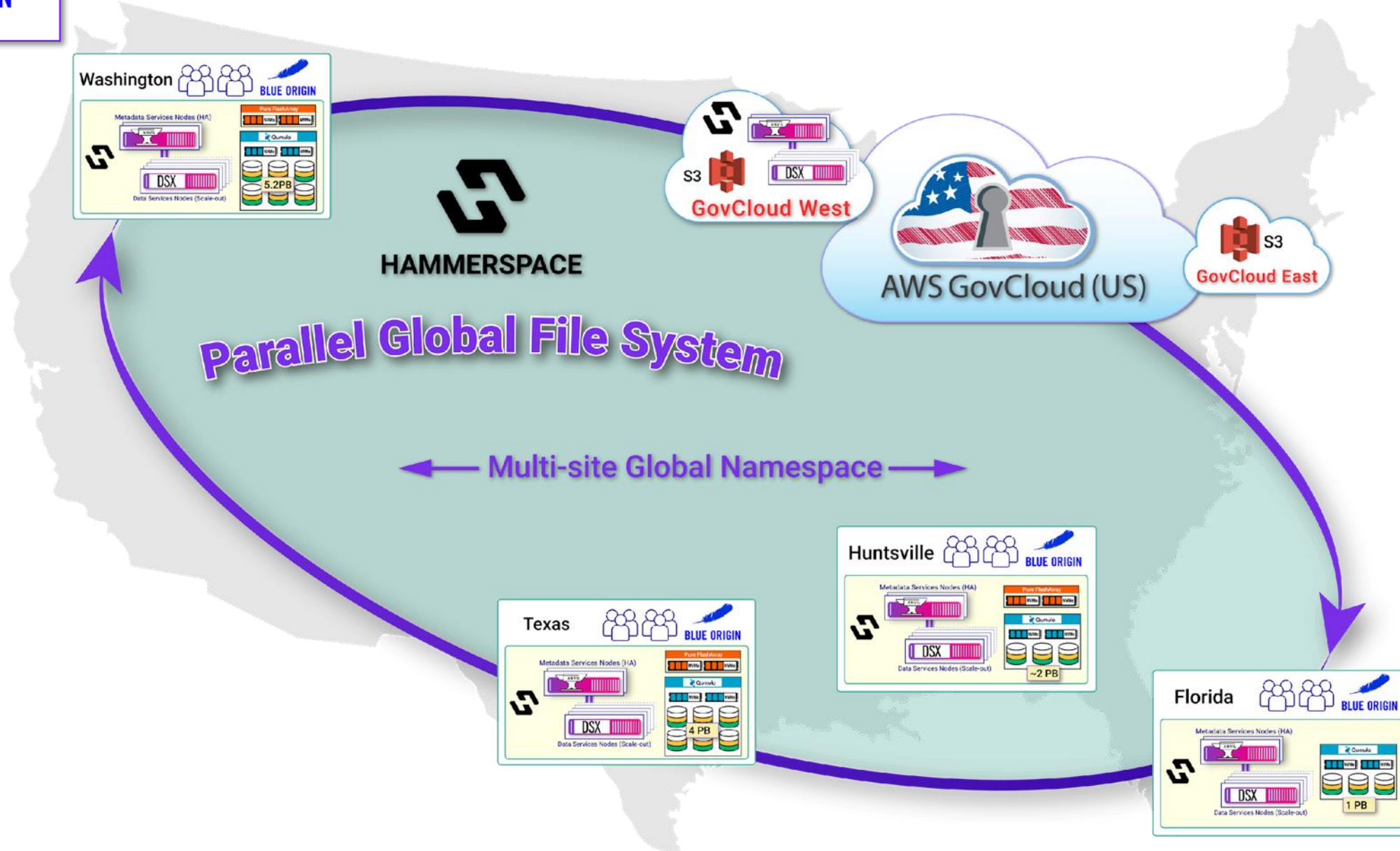


Hammerspace Solution

- No vendor even came close to Hammerspace's capabilities
- 1,000+ node Hammerspace storage cluster
- Feeding 24,000 GPUs, soon to be 350,000, then 1M
- Aggregate performance of 12.5TB/sec (100Tb/sec)
- Everything is standards-based and plug-n-play
- Customer was able to use existing OCP storage servers



Data Orchestration Powering Space Flight



Global File Access & Automated Data Orchestration

Hammerspace Approach



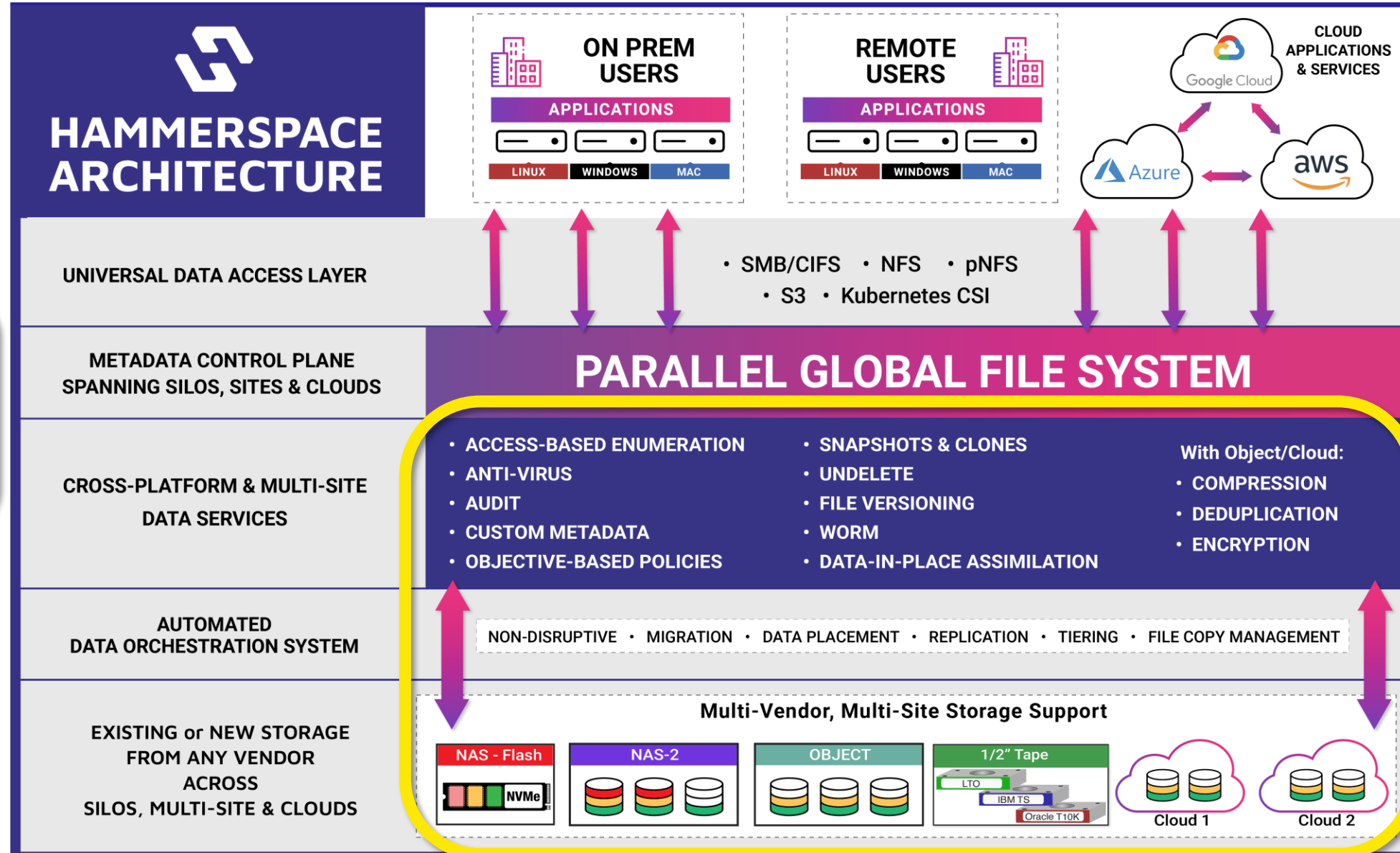
Applications

File System
(Global)

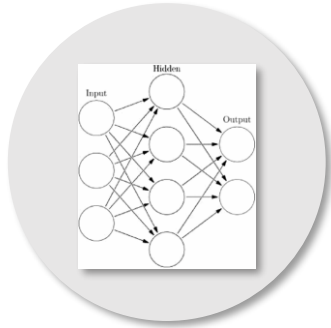
Data Orchestration
(By Objective)

Storage Systems
(Multiple Vendors)

HAMMERSPACE

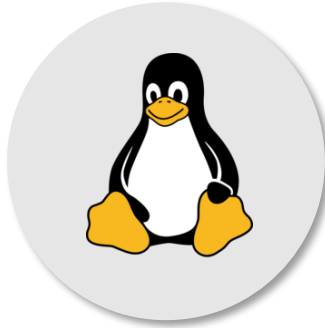


Standards-Based Global Data Platform



Why Needed Now

- AI and GPU Computing workloads are forcing HPC into the mainstream
- Performance and agility are needed now like never before



Why Possible Now

- Linux won the Unix wars.
- Now a much smarter NFS client needs only be done in Linux.



Why Hammerspace

- Unique talent: Technical CEO, CTO is Linux kernel maintainer, best-in-class team
- Unique technology: Addresses requirements of AI and GPU computing

Hammerspace Global Data Platform

Thank You

www.Hammerspace.com



HAMMERSPACE
DATA IN MOTION