Hammerspace
Global Data Platform

David Flynn – Chief Executive Officer

April 3, 2024

HAMMERSPACE

the **Future** of **Memory** and **Storage**
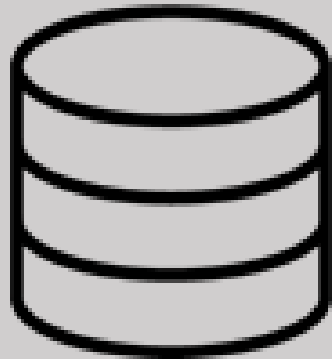
Hammerspace Is Like Magic

"Hammerspace" is an extra-dimensional space
that is instantly accessible and infinite in size

FMS
the Future of Memory and Storage

Goal to Radically Improve How Data Is Used

AI is forcing a long-overdue industry reckoning to implement radical changes in how data is used and preserved
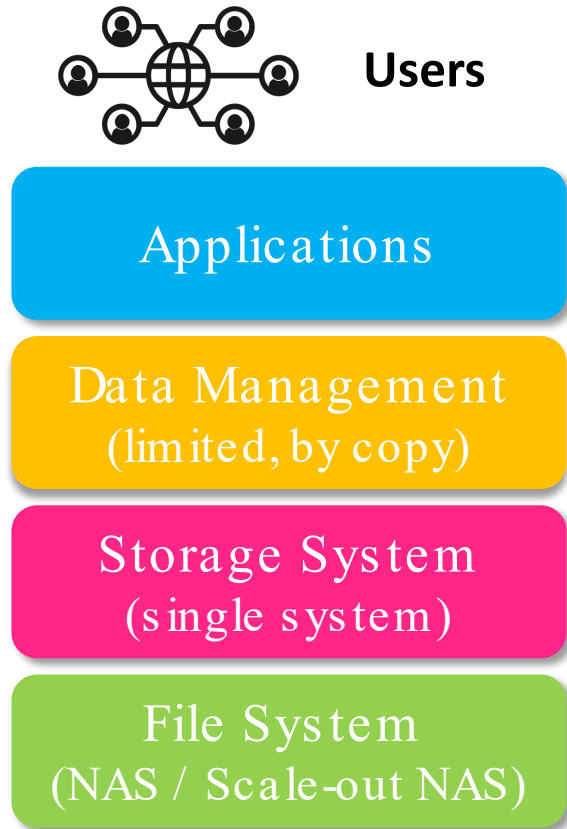
FROM: Data Placed to Rest in Storage

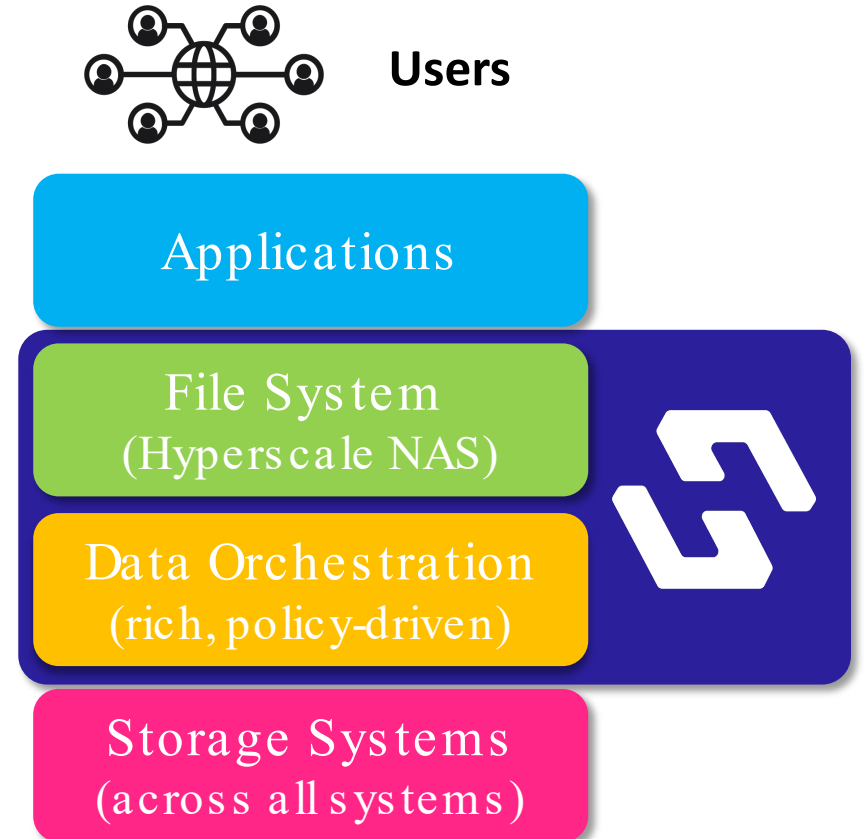TO: Data in Motion Across a Global Data Environment

the Future of Memory and Storage

A Fundamentally Different, Data-Centric Architecture

Storage-Centric Approach

Users

Applications
Data Management (limited, by copy)
Storage System (single system)
File System (NAS / Scale-out NAS)

❌ Data is Trapped in Silos

Hammerspace Data-Centric Approach

Users

Applications
File System (Hyperscale NAS)
Data Orchestration (rich, policy-driven)
Storage Systems (across all systems)

✔ Data Becomes a Global Resource

the Future of Memory and Storage

A New Solution Was Inevitable

❌ Valuable data trapped in silos

❌ Getting data to global users takes too long

❌ Infrastructure is not ready for AI

❌ Data copy sprawl impacts governance and security

❌ Lack performance to keep GPUs utilized

❌ Lack agility to use elastic cloud resources
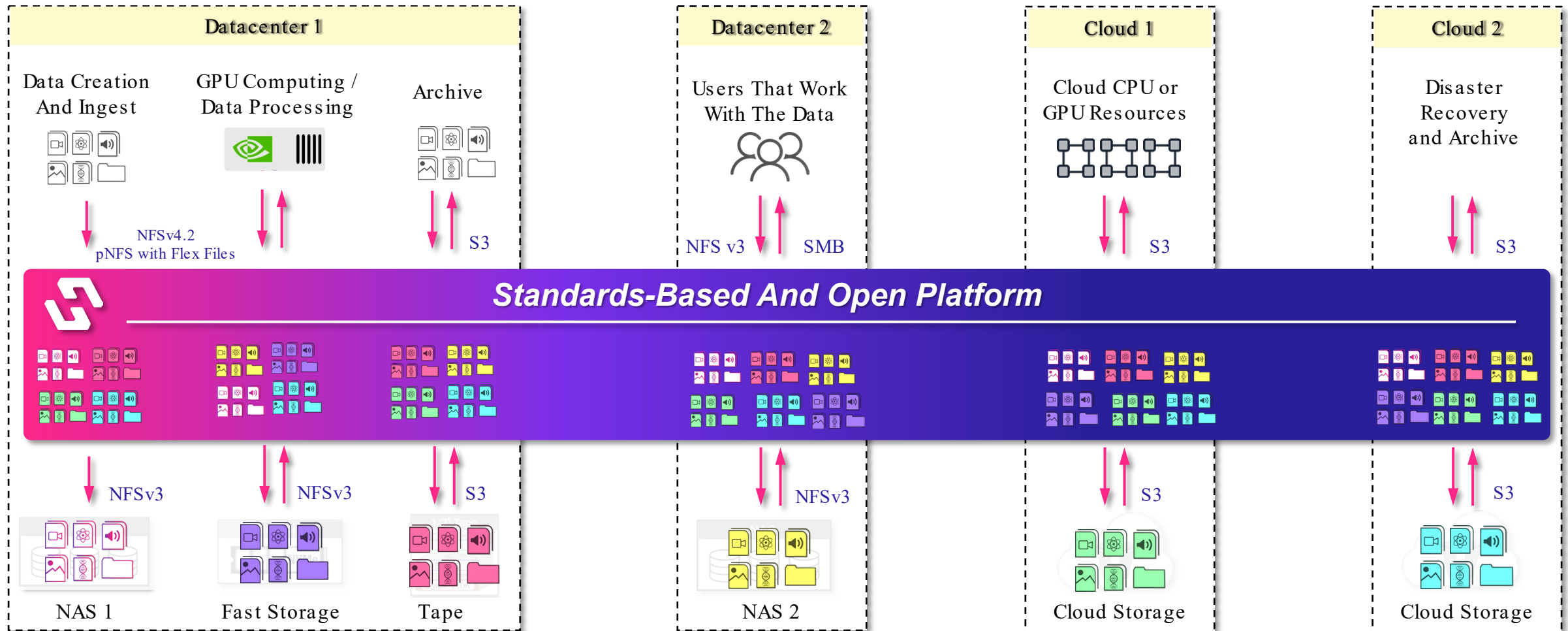
the **Future** of Memory and Storage
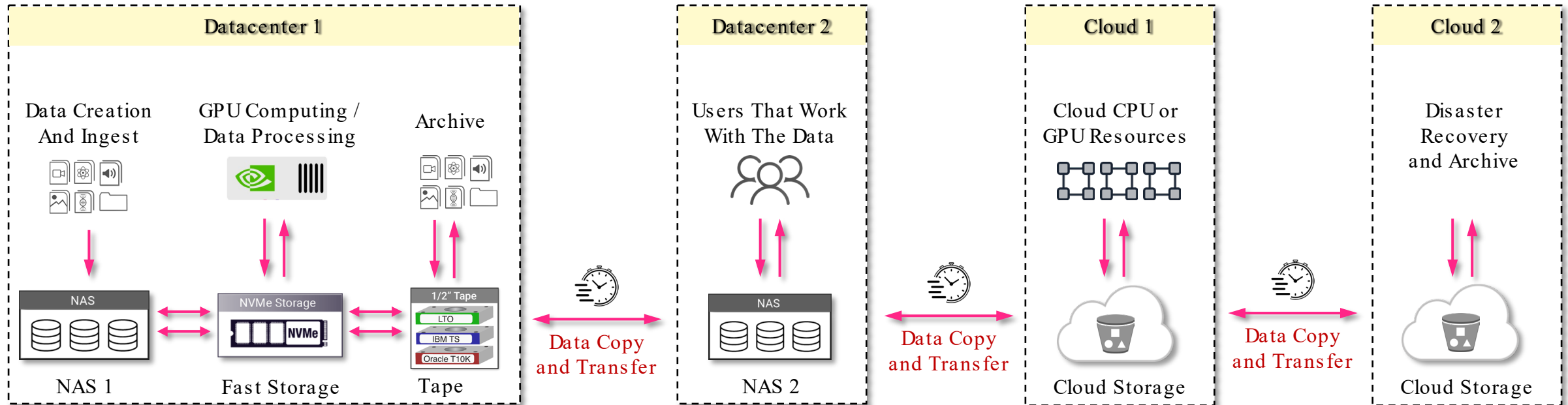
Hammerspace Changes the Paradigm

Hammerspace virtualizes the underlying storage infrastructure,
so all authorized users and applications can access the same data simultaneously.

Datacenter 1

Data Creation
And Ingest

GPU Computing /
Data Processing

Archive

Datacenter 2

Users That Work
With The Data

Cloud 1

Cloud CPU or
GPU Resources

Cloud 2

Disaster
Recovery
and Archive

Single Global File System That Spans Multiple Locations

NAS 1

Fast Storage

Tape

NAS 2

Cloud Storage

Cloud Storage

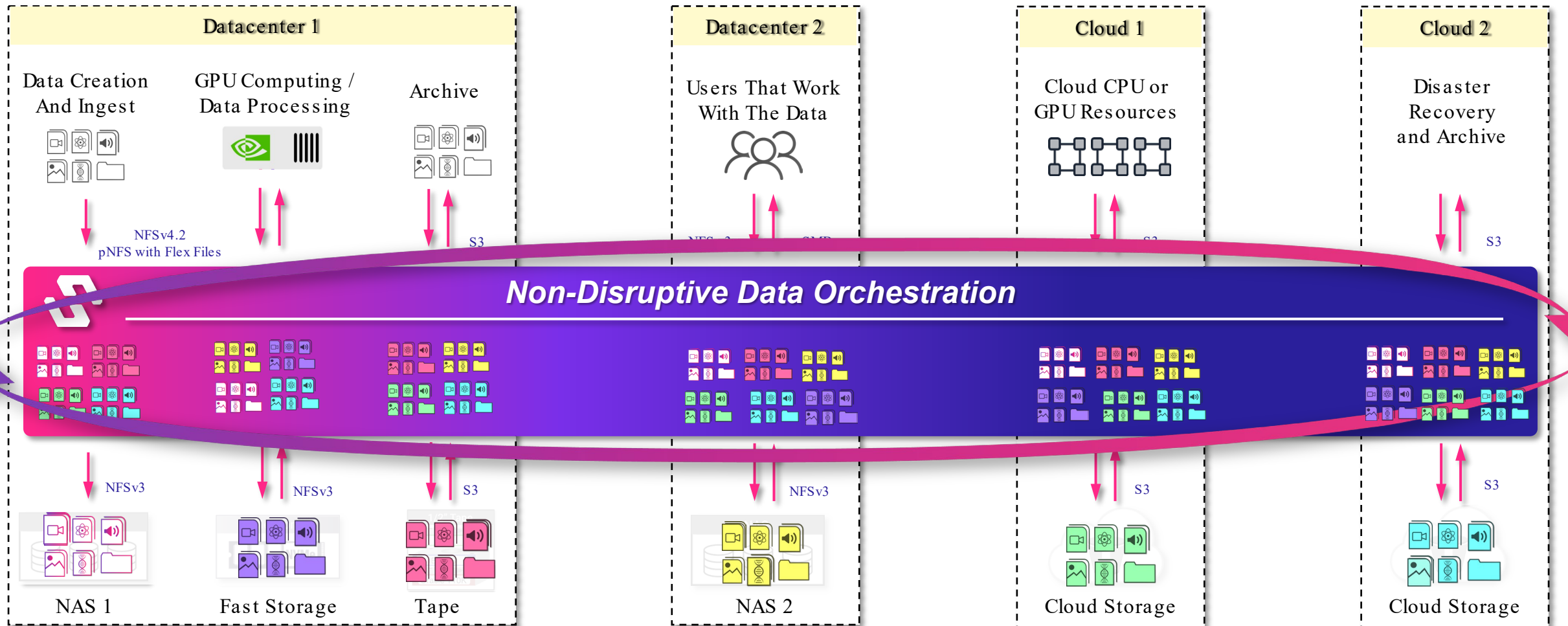the **Future** of Memory *and* Storage
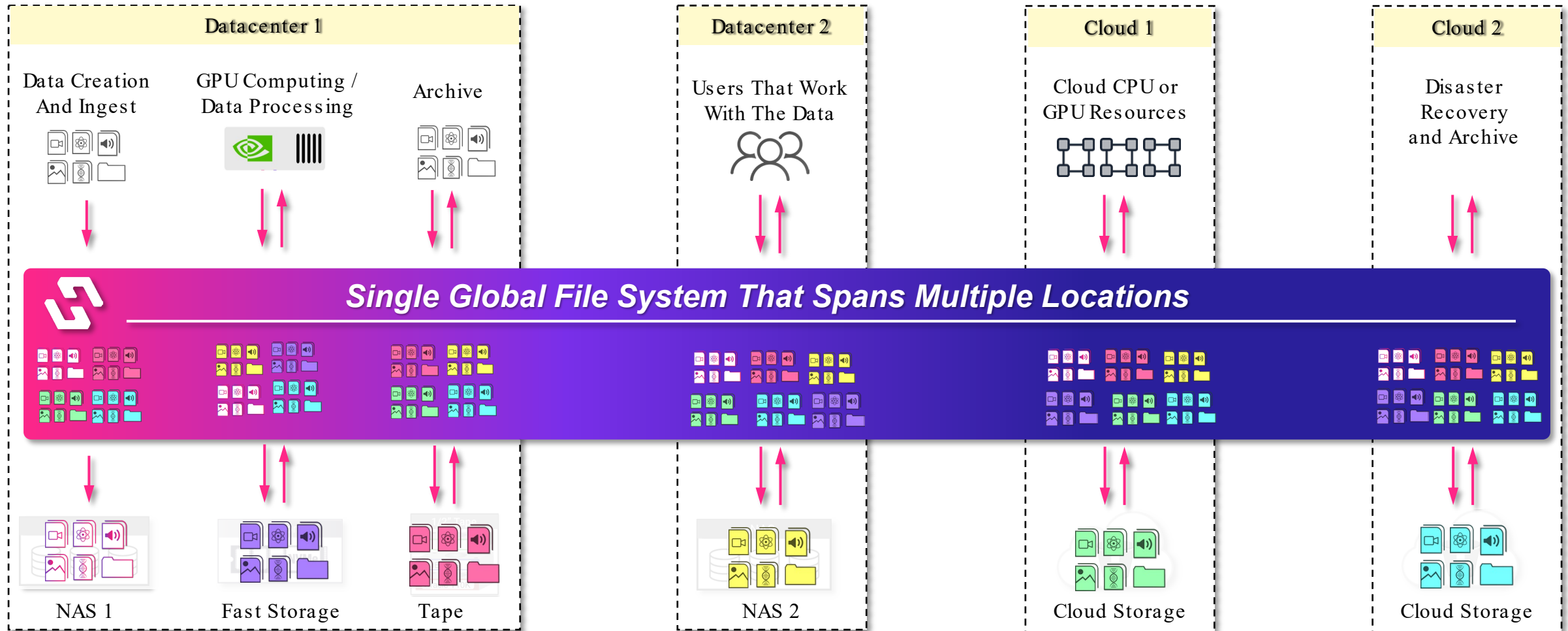
Hammerspace Changes the Paradigm

Hammerspace provides fast data access using industry-standard protocols, including a direct data path between Linux clients and storage for high-speed data processing.
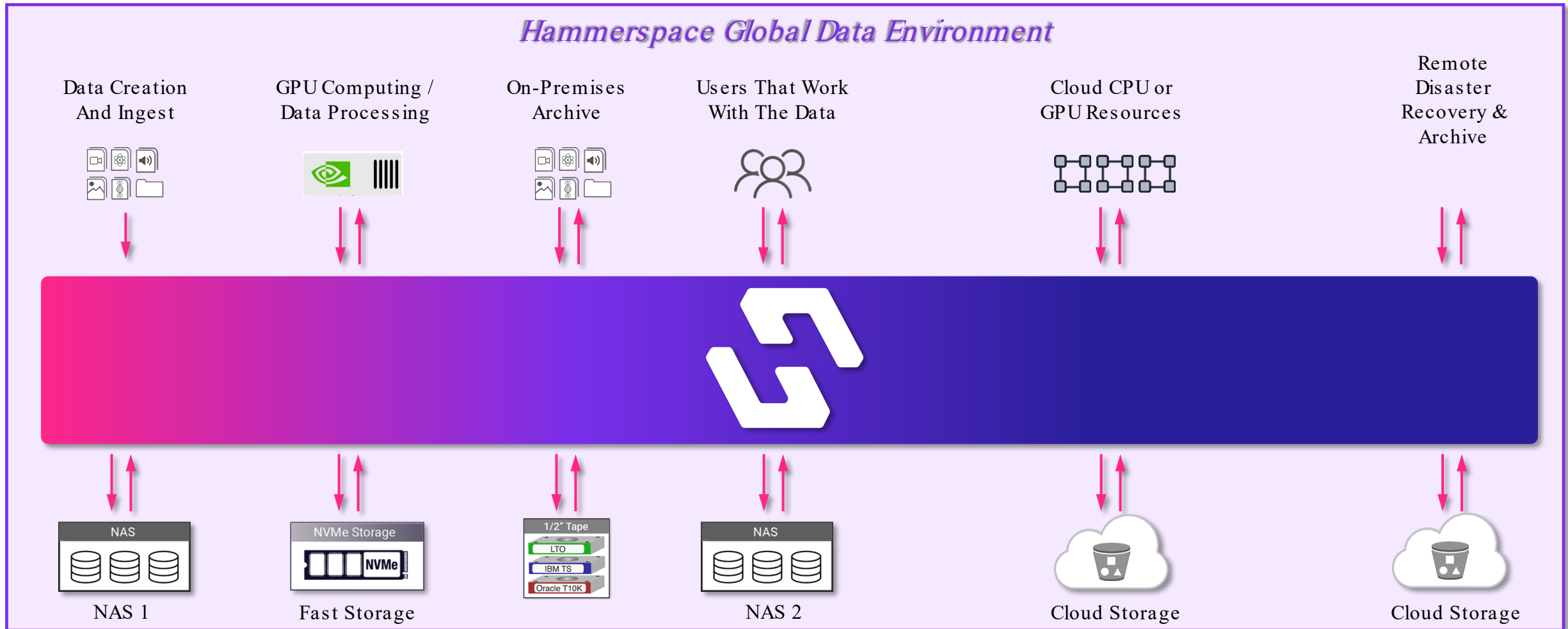

Datacenter 1

Data Creation And Ingest

GPU Computing / Data Processing

Archive

NFSv4.2 pNFS with Flex Files

S3

Datacenter 2

Users That Work With The Data

NFS v3   SMB

Cloud 1

Cloud CPU or GPU Resources

S3

Cloud 2

Disaster Recovery and Archive

S3

Standards-Based And Open Platform

NFSv3   NFSv3   S3

NAS 1   Fast Storage   Tape

NFSv3

NAS 2

S3

Cloud Storage

S3

Cloud Storage

FMS — the Future of Memory and Storage

Hammerspace Changes the Paradigm

Hammerspace automates data orchestration behind the scenes, transparent to users even on live data.

Datacenter 1

Data Creation And Ingest

GPU Computing / Data Processing

Archive

NFSv4.2 pNFS with Flex Files

S3

Datacenter 2

Users That Work With The Data

NFSv3 — SMB

Cloud 1

Cloud CPU or GPU Resources

S3

Cloud 2

Disaster Recovery and Archive

S3

Non-Disruptive Data Orchestration

NFSv3

NAS 1

NFSv3

Fast Storage

S3

Tape

NFSv3

NAS 2

S3

Cloud Storage

S3

Cloud Storage

the Future of Memory and Storage

Hammerspace Global Data Environment

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

Hammerspace Global Data Environment

| Data Creation And Ingest | GPU Computing / Data Processing | On-Premises Archive | Users That Work With The Data | Cloud CPU or GPU Resources | Remote Disaster Recovery & Archive |

NAS 1 · Fast Storage · 1/2" Tape (LTO, IBM TS, Oracle T10K) · NAS 2 · Cloud Storage · Cloud Storage

Standards-Based, Efficient Architectures

Hyperscale NAS for Data Intensive Workloads

Hyperscale NAS is a fundamentally different NAS architecture urgently needed to power AI initiatives and the GPU computing boom.

HYPERSCALE NAS

- **Parallel file system + Enterprise NAS**
- **Standards-based**
    - **- pNFS client in Linux**
    - **- Networking**
    - **- Storage Interfaces**
- **Linear scale to thousands of nodes**
- **Software-defined**
- **Storage-agnostic**

Scale-Up NAS

- **Introduced 1990s**
- **Single storage controller**
- **For Corporate data**

Scale-Out NAS

- **Introduced early 2000s**
- **Multiple storage controllers**
- **For Unstructured Data**
- **Good for scaling capacity**

SCALING NAS CAPACITY

SCALING NAS PERFORMANCE

MOSTLY HUMAN-GENERATED DATA

MOSTLY MACHINE-GENERATED DATA

the **Future** of **Memory** and **Storage**

Hyperscale NAS Delivers Linear Performance Scalability

Performance (TB/sec) →

Hyperscale NAS architecture

Scale-out NAS architecture

Number of Storage Nodes →

Hyperscale NAS increases performance linearly as node count increases.

So far proven at over 1,000+ storage nodes.

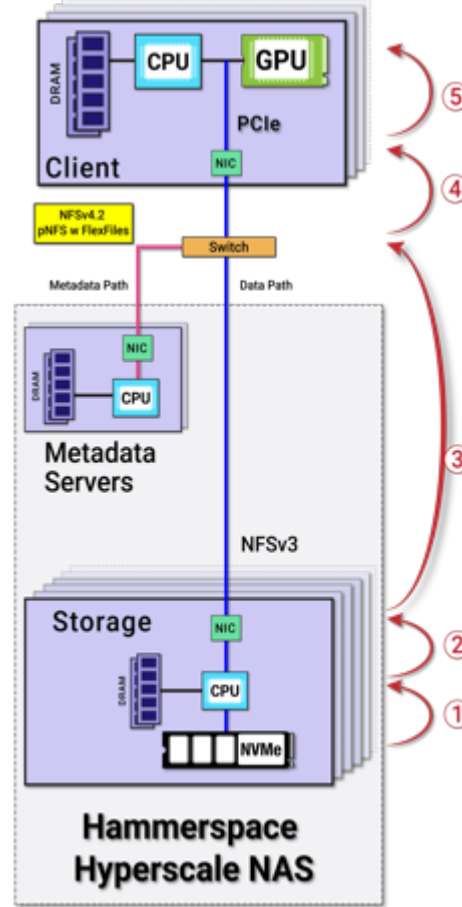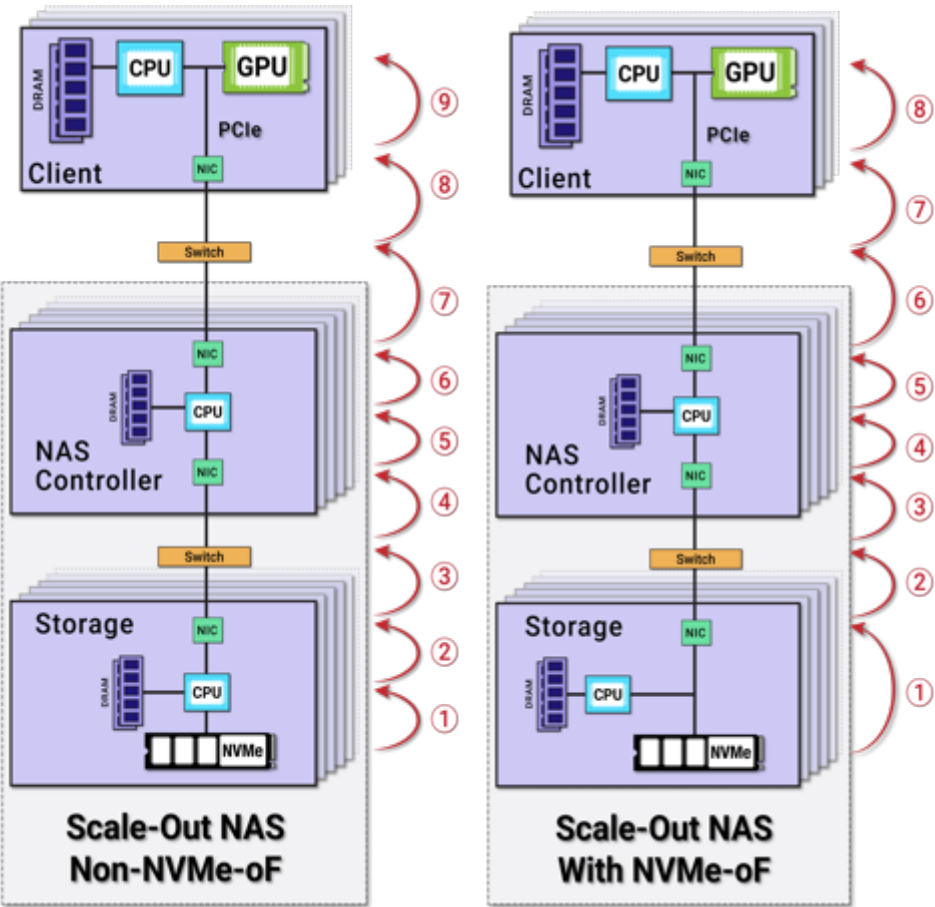Scale-Out NAS performance plateaus as storage nodes increase

the **Future of Memory** *and* **Storage**

File Storage Architecture Comparison

File Storage Architecture Comparison

Scale-Out NAS Non-NVMe-oF

Scale-Out NAS With NVMe-oF

Hammerspace Hyperscale NAS

Hyperscale NAS with HCI Clients

2x	**2x**	**2x**	**2x**	**2x**
Reduction in Servers	Reduction in Networking	Reduction in Latency	Reduction in Power	Reduction in Rack Space

NVIDIA GPUDirect

File Storage Architecture Comparison

2x
Reduction in
Servers

2x
Reduction in
Networking

2x
Reduction in
Latency

2x
Reduction in
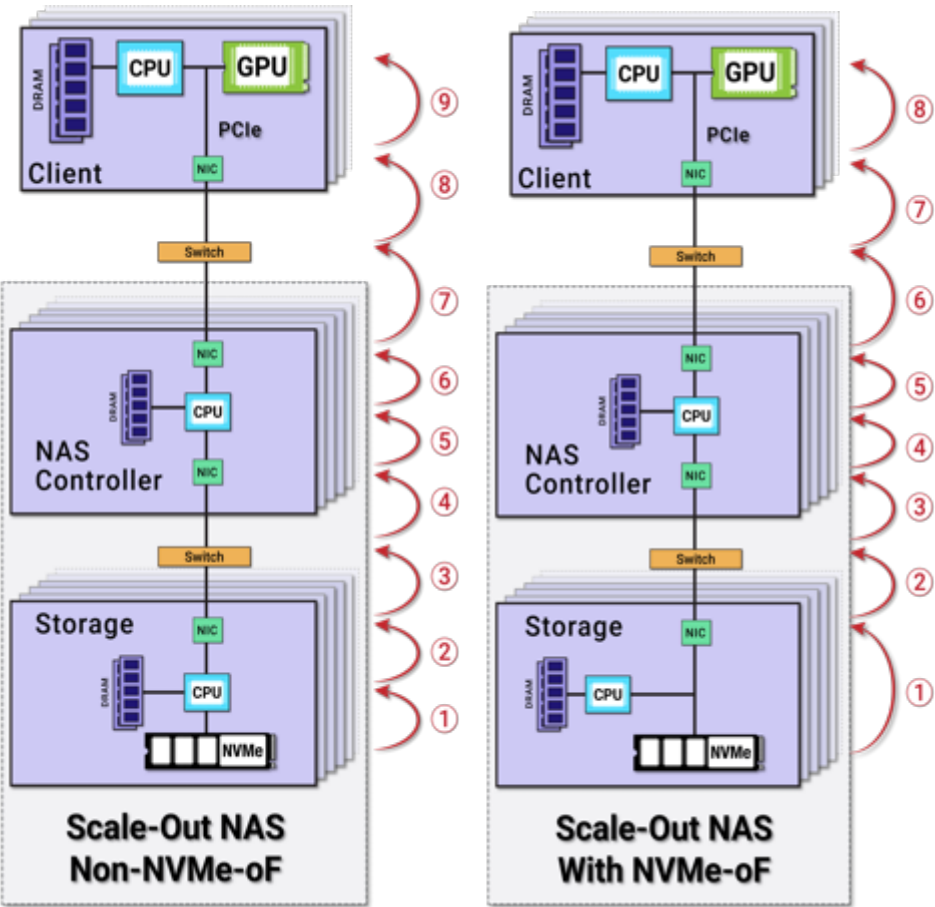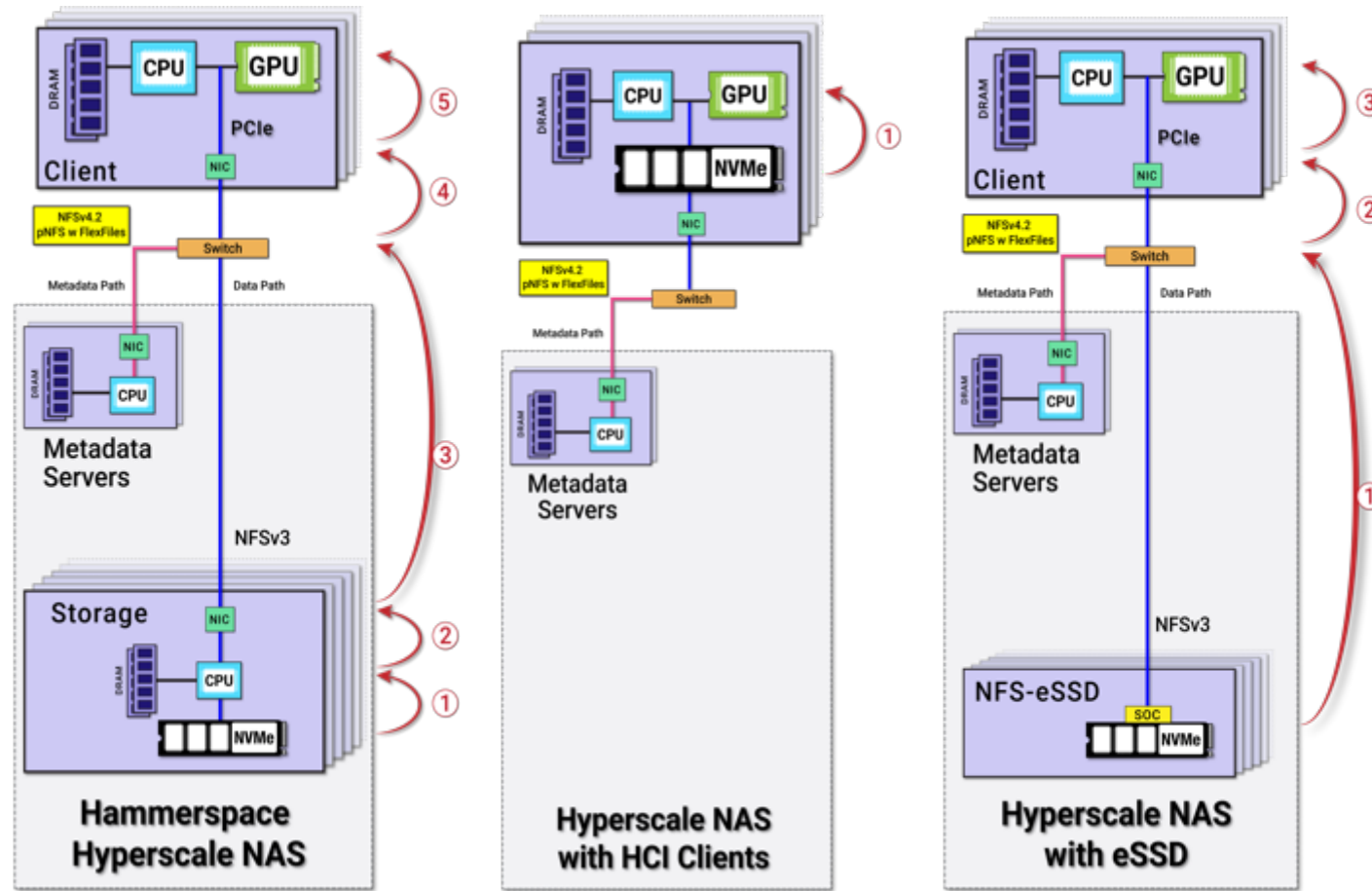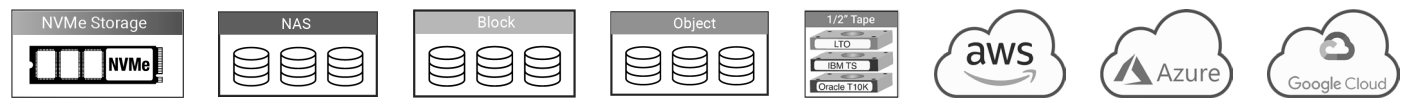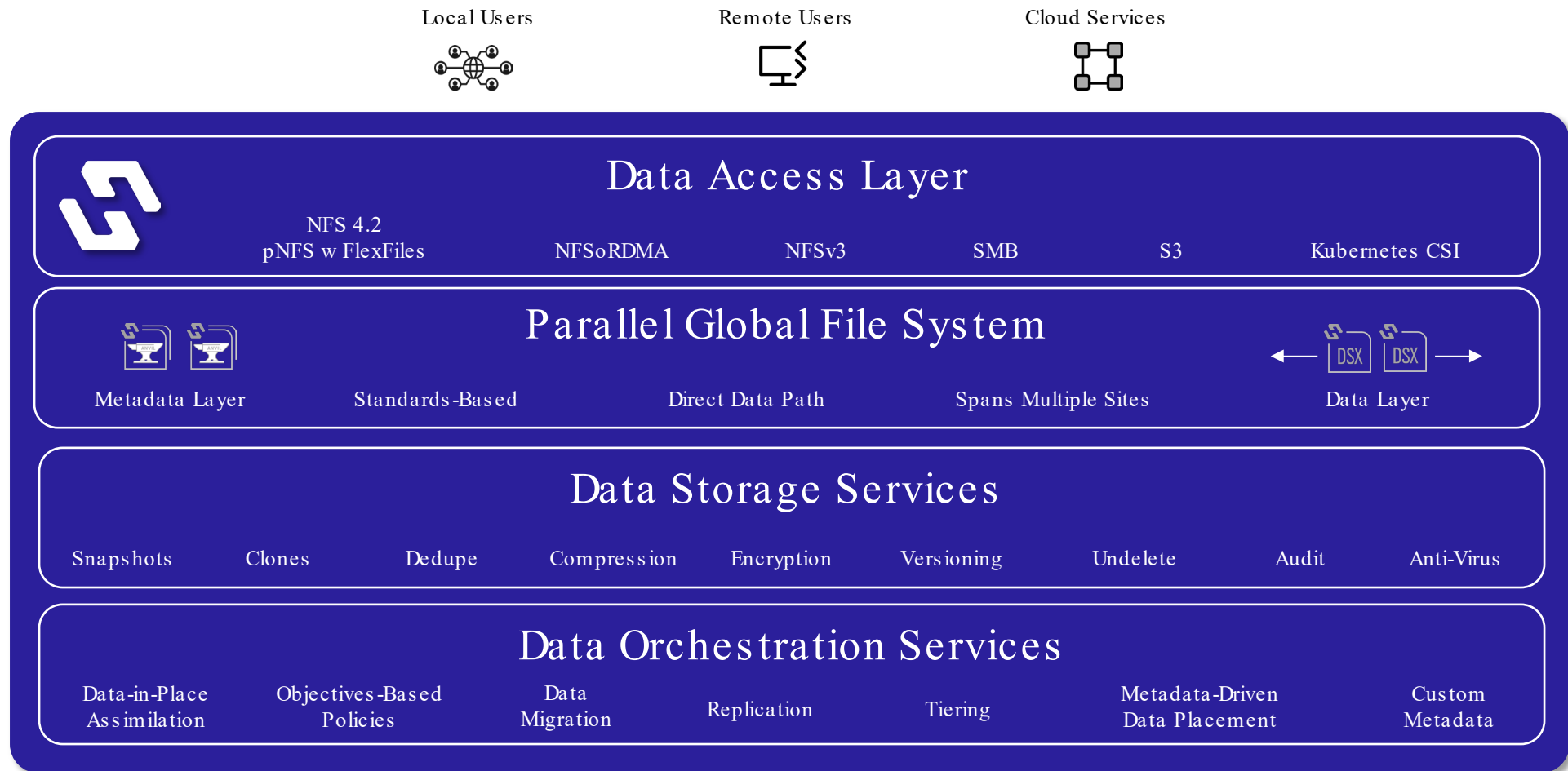Power

2x
Reduction in
Rack Space

NVIDIA GPUDirect

the **Future** of Memory and Storage

Architecture

Software Architecture

Local Users

Remote Users

Cloud Services

Data Access Layer

NFS 4.2
pNFS w FlexFiles   NFSoRDMA   NFSv3   SMB   S3   Kubernetes CSI

Parallel Global File System

Metadata Layer   Standards-Based   Direct Data Path   Spans Multiple Sites   Data Layer

Data Storage Services

Snapshots   Clones   Dedupe   Compression   Encryption   Versioning   Undelete   Audit   Anti-Virus

Data Orchestration Services

Data-in-Place
Assimilation   Objectives-Based
Policies   Data
Migration   Replication   Tiering   Metadata-Driven
Data Placement   Custom
Metadata

NVMe Storage   NAS   Block   Object   1/2" Tape
NVMe   LTO
IBM TS
Oracle T10K

aws   Azure   Google Cloud

Runs on New or Existing Storage from All Leading Vendors

the Future of Memory and Storage

About Parallel NFS with Flex Files

- Parallel NFS (pNFS) introduced as optional feature of NFS in 2010, enhanced in later RFCs

- Flex Files layout type engineered by Hammerspace and added to NFSv4.2 spec in 2018

- Defines a standards-based parallel file system architecture using NFS

- Architecture requires NFSv4.2 client which is part of the Linux kernel

- Hammerspace Global Parallel File System is the first to leverage this architecture

Linux Server

NFSv4.2 Client

Metadata Path: pNFSv4.2 w/ Flex Files
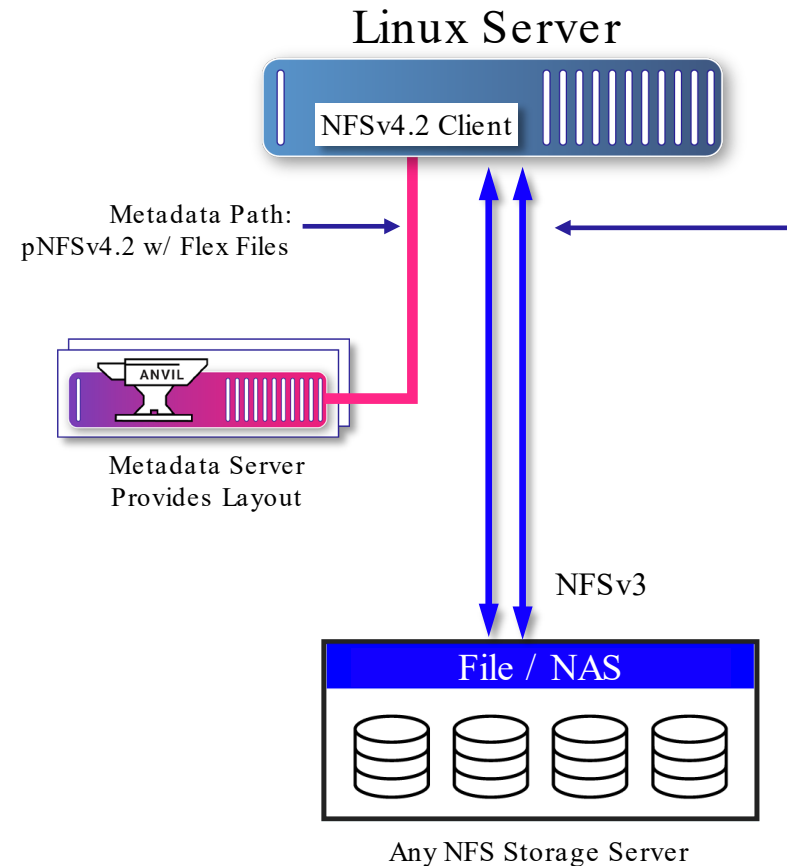
Direct data path between Linux client and storage volumes

ANVIL

Metadata Server Provides Layout
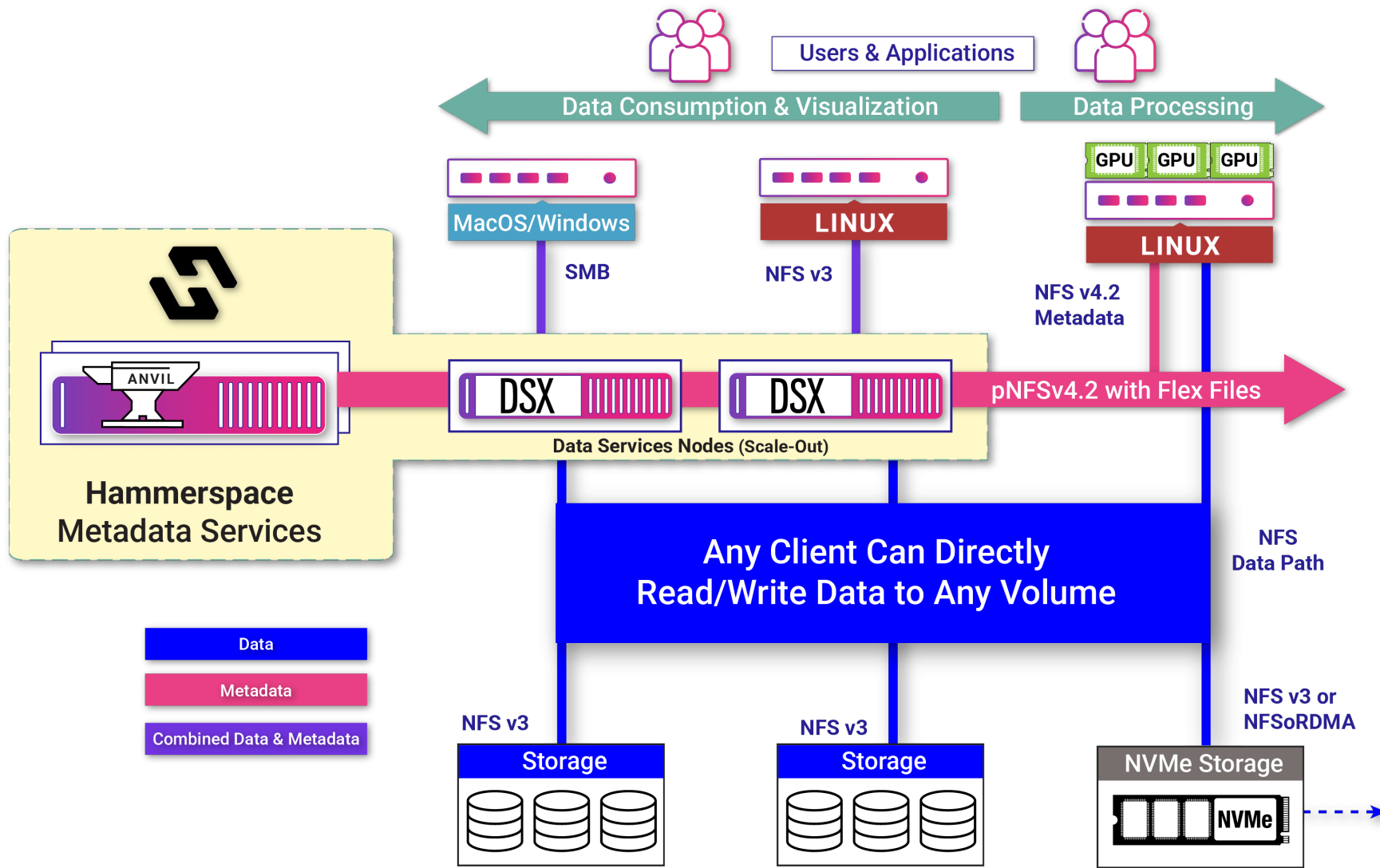
Also provides for:

- Multiple parallel network connections between client and server
- Ability to write to multiple storage nodes synchronously
- Ability to move data while it is being accessed without interruption
- Eliminates protocol chatter
- File granular access and telemetry
- TCP or RDMA Supported

NFSv3

File / NAS

Any NFS Storage Server

the Future of Memory and Storage

About the Hyperscale NAS Architecture
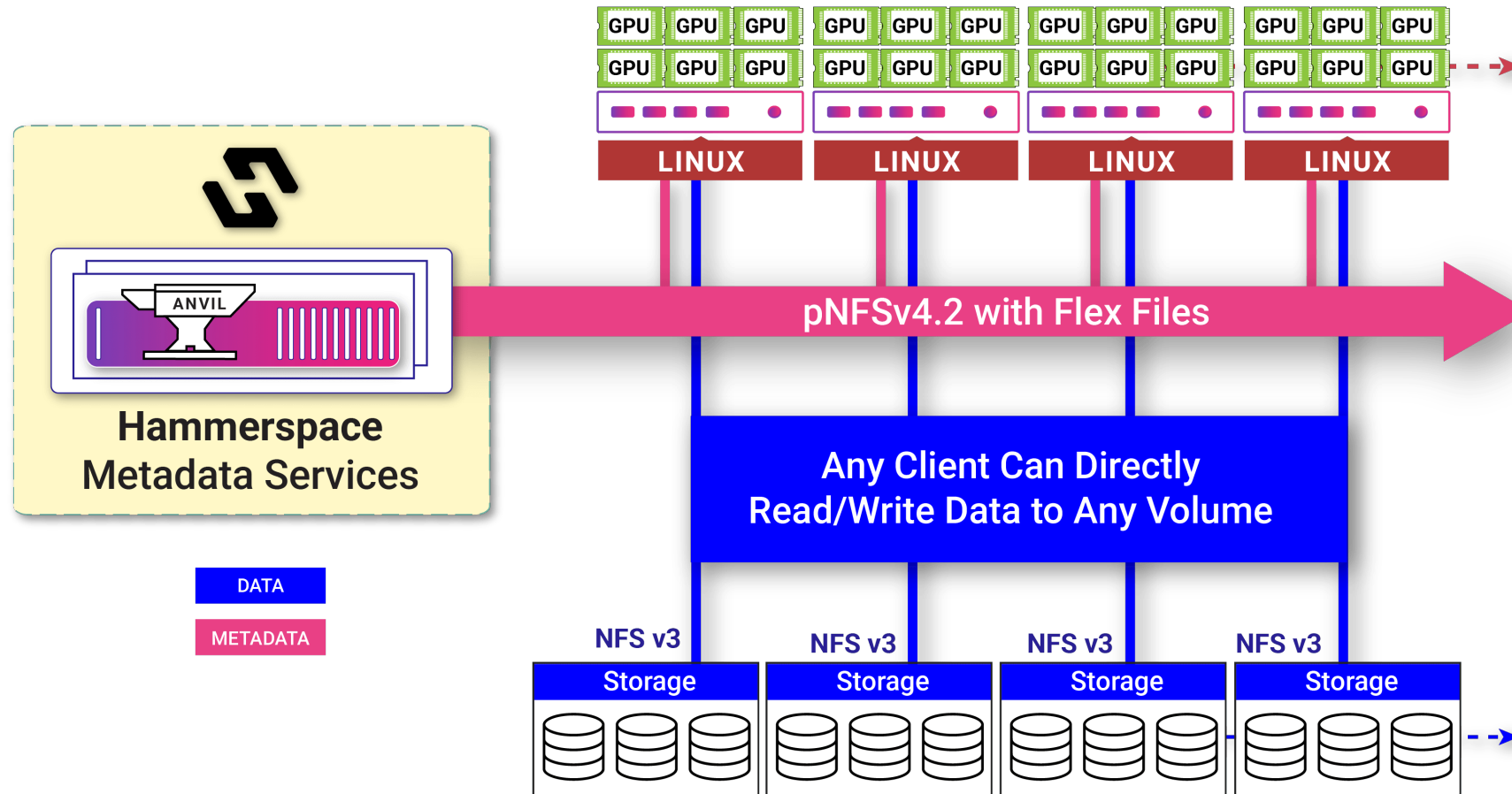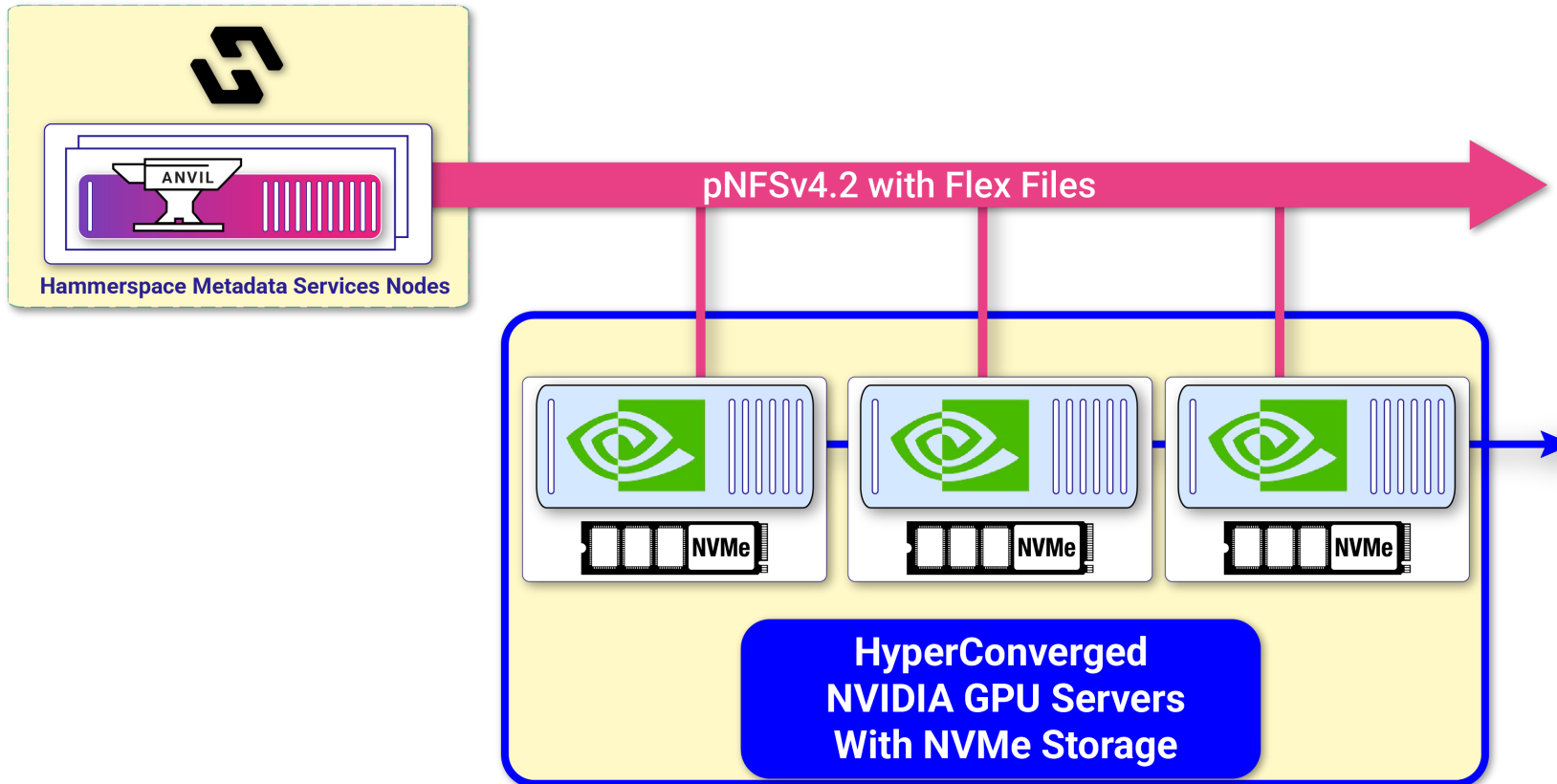

- Extreme Parallel Performance
- Linear scale to >1000s of nodes
- Standards-Based
- Choice of networking
- Software-defined
- Storage-agnostic

Hammerspace Metadata Services

pNFSv4.2 with Flex Files

Any Client Can Directly Read/Write Data to Any Volume

DATA

METADATA

LINUX LINUX LINUX LINUX

NFS v3 NFS v3 NFS v3 NFS v3

Storage Storage Storage Storage

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.

Tightly Coupled in a Hyperconverged Architecture

pNFSv4.2 with Flex Files

Hammerspace Metadata Services Nodes

NVMe NVMe NVMe

HyperConverged NVIDIA GPU Servers With NVMe Storage

- Supports Hyperconverged GPU servers with NVMe storage for greater performance & network efficiency

the **Future** of **Memory** and **Storage**
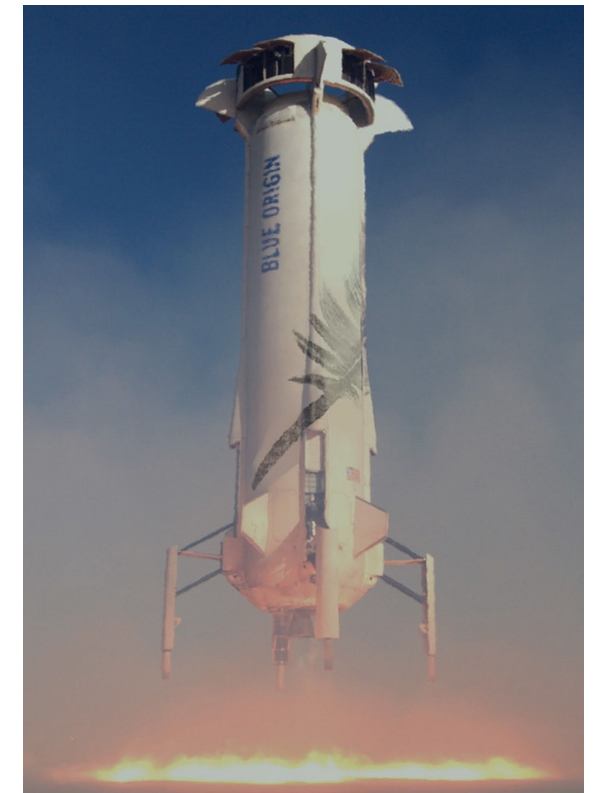
Customer Stories

Customer Story: Aerospace Company

About the Customer, Business Goals, Requirements

- Aerospace company building and testing reusable rockets and engines

- Remote sites producing large amounts of data

- NAS Storage at each site filling up

- Engineers in various offices and work from home

- 11,000 users across 9 locations

- Moving data to be "close to the engineer" is a challenge

Solution and Why Hammerspace

- Hammerspace Global Data Environment spans locations

- Engine test data now proactively made available at multiple sites

- Local data stays on fast storage, then moved to AWS S3

- Data always available

- Transition from legacy storage to Hammerspace was transparent to users

the **Future of Memory** and **Storage**

Customer Story: VFX Studio

About the Customer, Business Goals, Requirements

- Leading global Visual FX and Animation Studio

- 300 artists in two UK locations, remote work around the world

- Virtual since 2019, storage in multiple locations and clouds

- Sharing large files globally to keep artists productive was a "real issue"

Solution and Why Hammerspace

- Hammerspace orchestrates data movement globally

- Both on-premises and Azure cloud environments

- Integrated with Jellyfish's production scheduler: Autodesk Shotgrid

- 100M+ files globally accessible for read/write access in minutes

- Saved 30% in initial render costs by moving files to lower cost clouds

- Enabled access to global talent pool of artists

JELLYFISH PICTURES

"Hammerspace is a core part of the Jellyfish strategic vision, helping us expand our global workforce in a highly competitive industry, increasing our productivity to meet and exceed our clients' expectations while greatly reduce costs on multiple levels."
Jeremy Smith, CTO, Jellyfish Pictures

the **Future** of Memory and Storage

Customer Story: LLM Training at Hyperscale

"What Hammerspace does is pure magic."
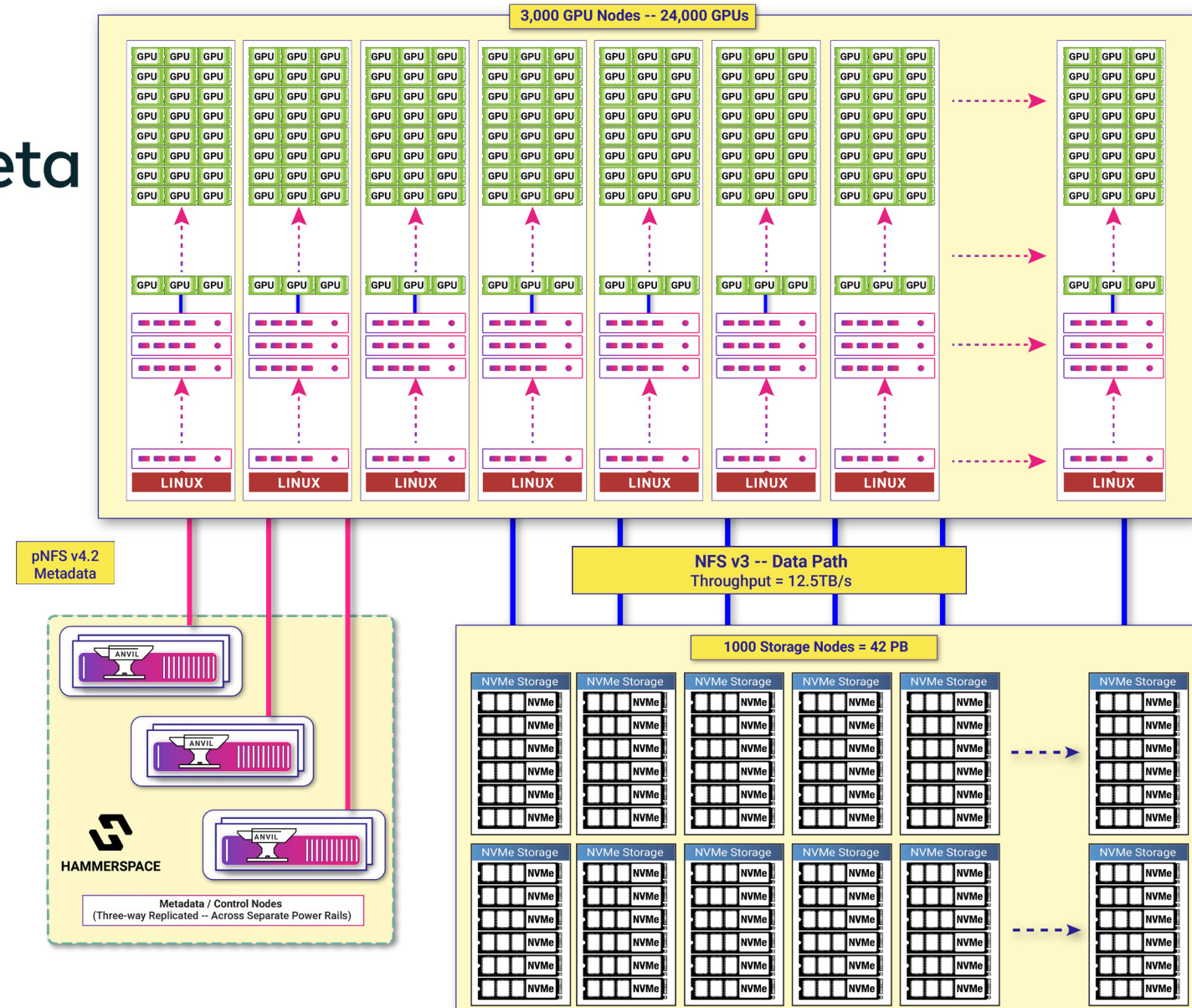-Paul Saab, Principal Engineer Meta

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

3,000 GPU Nodes -- 24,000 GPUs

pNFS v4.2 Metadata

NFS v3 -- Data Path
Throughput = 12.5TB/s

Metadata / Control Nodes
(Three-way Replicated -- Across Separate Power Rails)

1000 Storage Nodes = 42 PB

the Future of Memory and Storage

Addressing Requirements of AI and GPU Computing

Deep Learning   LLM Training   Generative AI   Scientific Computing   Video and Image Rendering   Data Analytics

Extreme Parallel Performance to Feed GPUs	Hyperscale NAS Architecture
Gather and Assemble Existing Data	Data-in-Place Assimilation
Efficiently Move Data to Compute Resources	Data Orchestration
Multi-Site, Hybrid-Cloud, Multi-Cloud	Global File System
Metadata-Driven Workloads and Pipelines	Programmable Metadata
Data Security and Governance	Advanced Data Services

the **Future** of Memory *and* **Storage**

Thank you

HAMMERSPACE