CHEOPS Workshop

Form HPC to AI a data journey

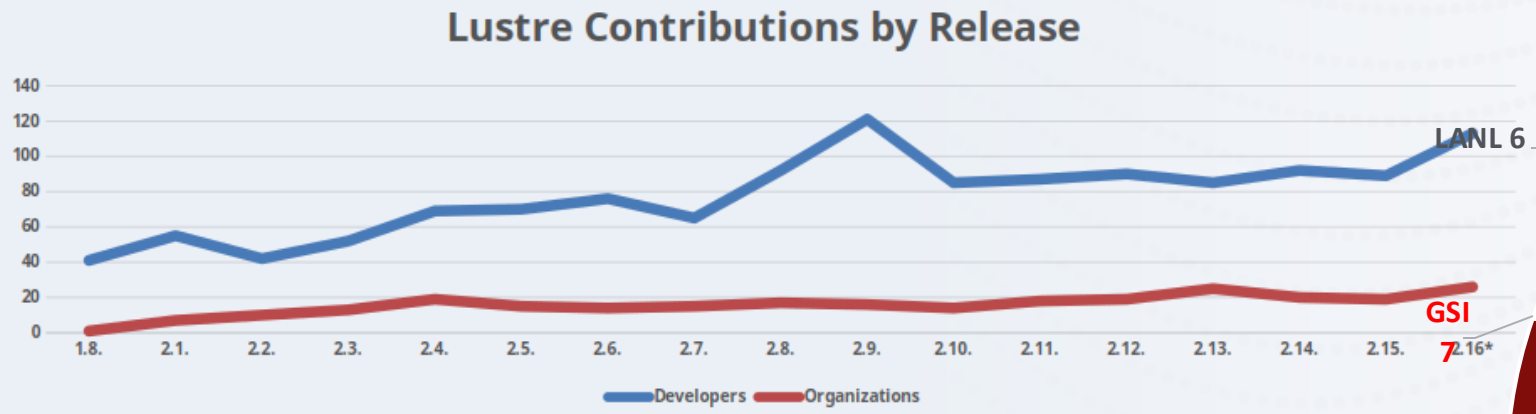
March 31st, Rotterdam

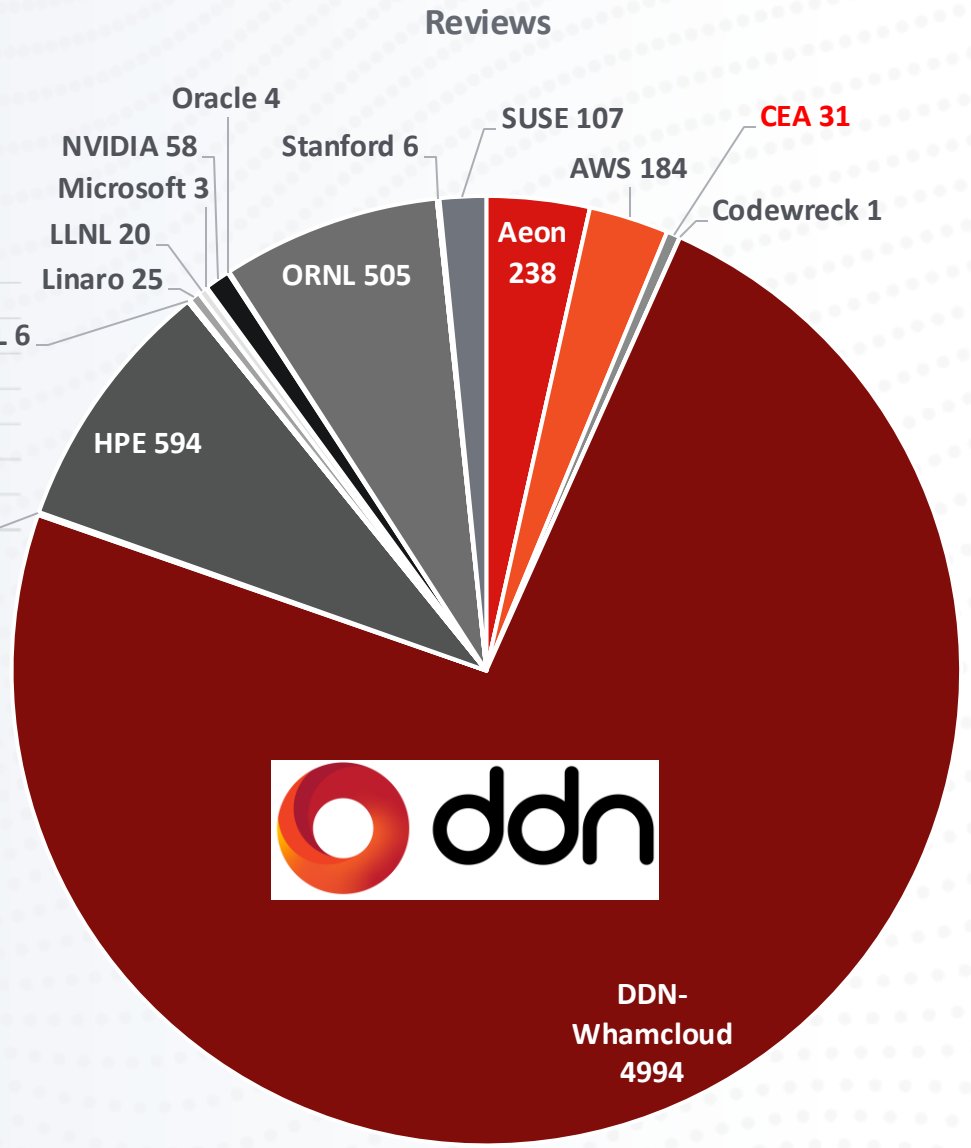
Jean-Thomas Acquaviva, jtacquaviva@ddn.com
DDN Senior Researcher

[Any Scale] [Any Data Center] [Any cloud]

ddn

An Open-Source Driven Company





Lustre Open-Source Parallel File system (OpenSFS)

- Designed for HPC: data extension of the compute platform
- OpenSFS provides overall directions and a forum for discussion among users
- DDN is the lead contributor to Lustre
- User meetings in Europe / ASIA co-organized by DDN

Data courtesy of Dustin Leverman (ORNL)

Source: https://git.whamcloud.com 2.15.50 to 2.15.91

A Presence at Every Stages of the Industrial Cycle

In Europe DDN is active at every stages of the Industrial cycle

- **Knowledge generation**
  - Mentorship, training, teaching and funding Ph.Ds
- **Research collaboration**
  - Participation to EU/US/JP research programs
  - Seed funding for partners research programs
- **Product definition**
  - Open-Source development + Industrial partnership
- **Customer feed-back**
  - Joint lab with strategic partners

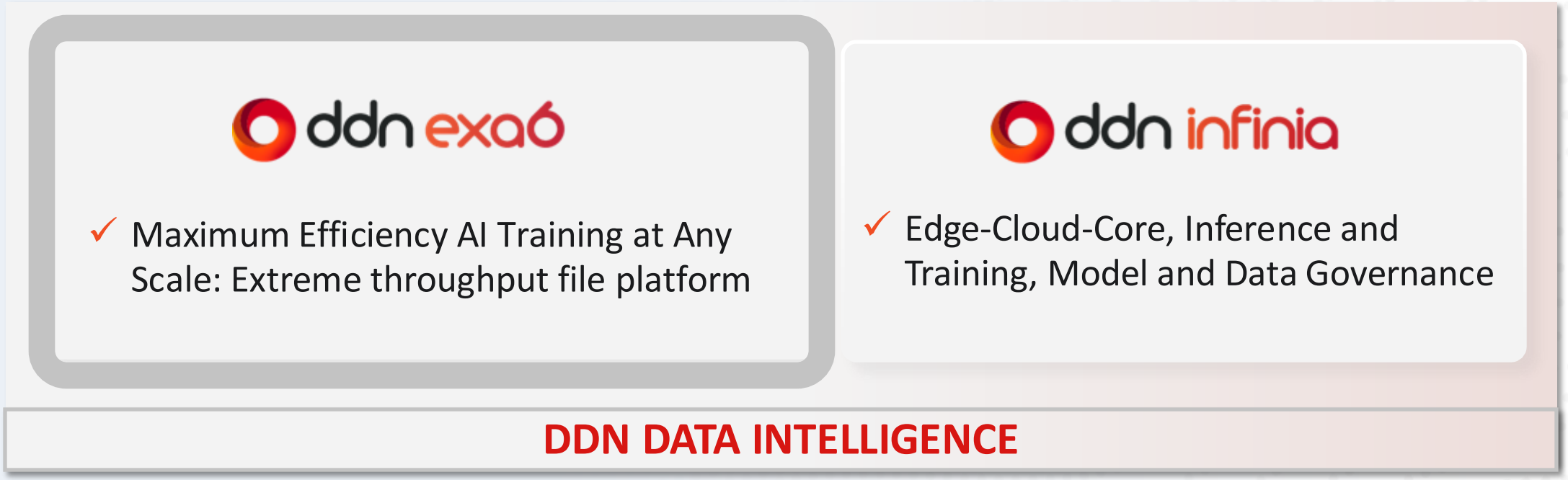
Jensen Huang
CEO and co-founder NVIDIA
Alex Bouzari,
CEO and co-founder
DDN

"NVIDIA is powered by DDN … Without DDN, NVIDIA supercomputers wouldn't be possible."

After Feb. 20, 2025

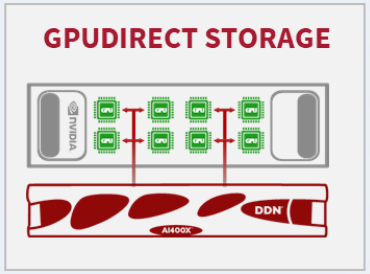
DATA PLATFORMS VISION

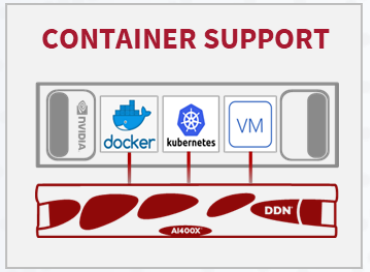
DDN's Data Intelligence Platform is Powering the World's AI Transformation



EXAScaler: Industrialization of Lustre with Specific Features

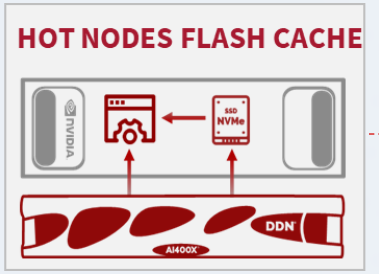
GPU Direct Storage (GDS) enables a direct data path for direct memory access (DMA) transfers between GPU memory and external storage bypassing CPU. GDS Increases throughput, reduces latency, eliminates memory copies, bypass internal architecture bottlenecks.

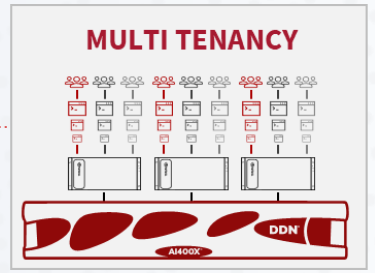




EXAScaler supports all popular container frameworks including Docker, Kubernetes with Container Storage Interface (CSI) with automated mounting of filesystem without needing roto access.

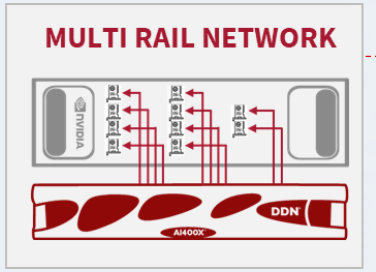
Hot Nodes converts local SSDs in application servers into read-only cache to improve re-read performance. AI applications benefit significantly by avoiding network hops to fetch data. Cached files persist through client node downtime and are available again once the client node reboots. Files get detached from cache automatically if modified.

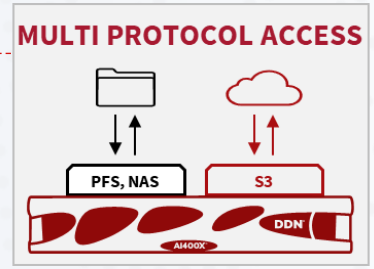




EXAScaler supports tenant isolation via subdirectory mount & node-maps to identify client nodes/VM/container used by a tenant. It also supports VLAN based isolation with LNET Routers.

Multi-rail network enable aggregation of bandwidth of multiple network interfaces into one single fat pipe. It simplifies network configuration, increases bandwidth and increases resilience, by allowing alternative transmission paths.

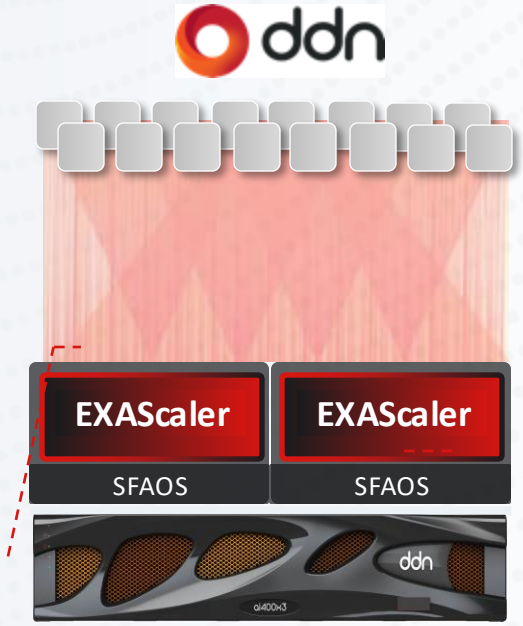




Along with default access via Lustre native clients, EXAScaler also supports NFS, SMB, S3 and CSI interfaces for data sharing. Ingest via one protocol and egest via another is supported. Commodity servers are required to run the gateway protocol.

30x Faster and More Reliable Training and Inference Across GPU's

- **Full GPU Efficiency Across Generation (Blackwell and Beyond):** 10x Reduction in wait time for data access & idle time for checkpointing
- **Massive Data Set Performance and Security:** Store, access & process exabytes of data
- **Smallest Data Center Footprint and Power:** 10x less power, cooling & rack space.
- **Proven at Any Scale: Exabyte Level with 100,000 GPUs in single cluster**



Intelligent Clients

Fastest Network

Fastest Servers

Fastest I/O

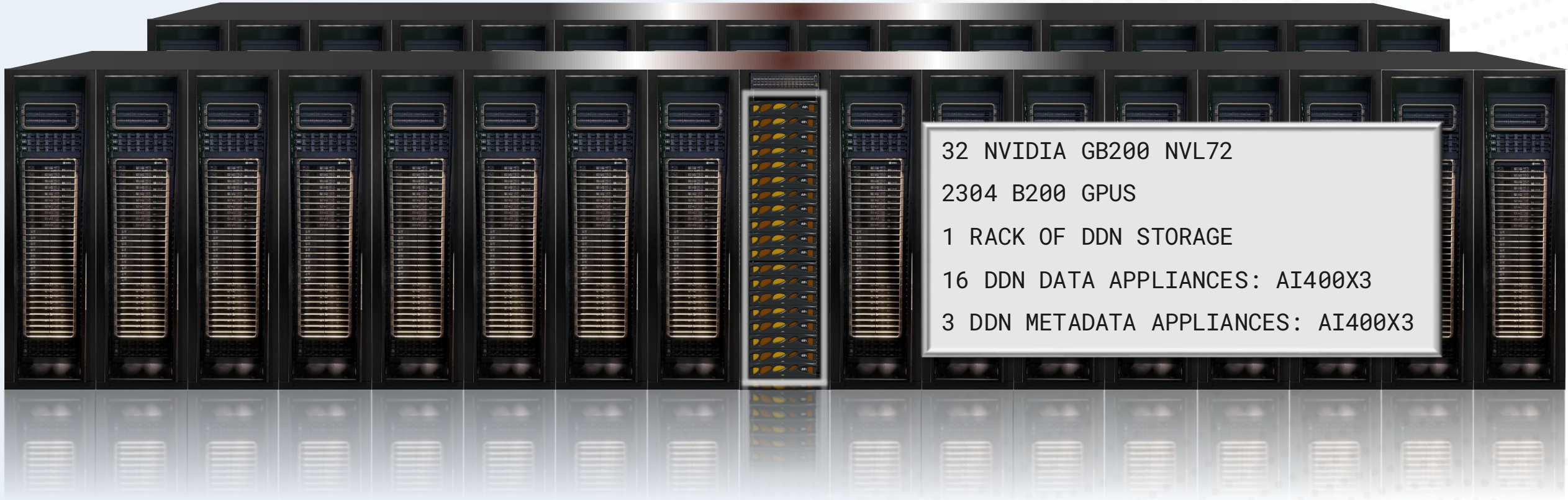
Resilient, High-Density Enclosures

Massive Expansion NVMeoF or SAS

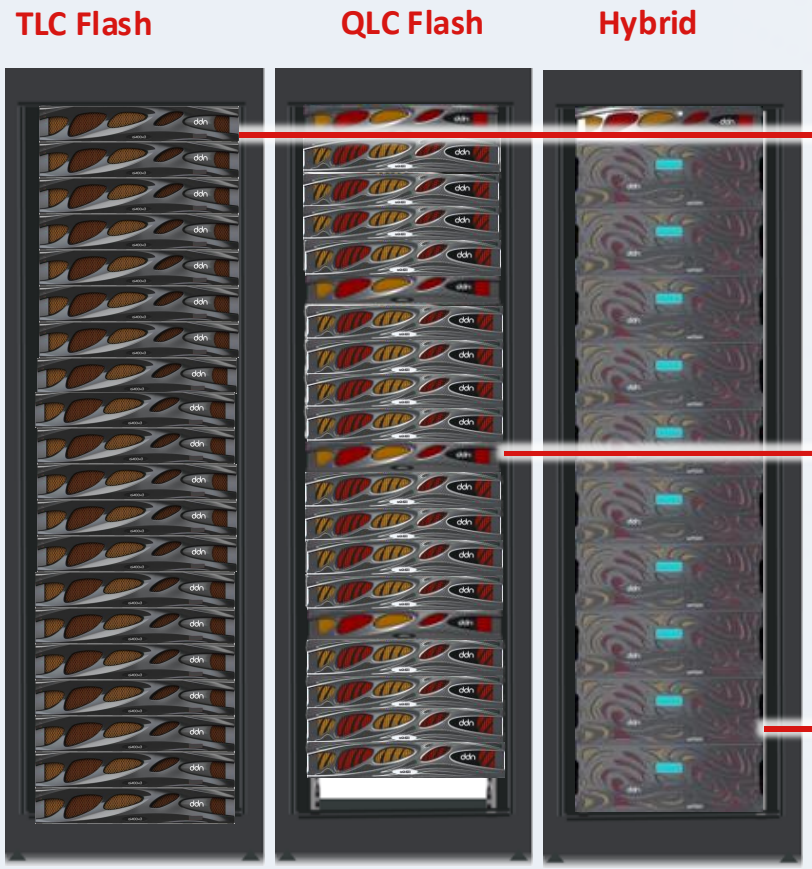
Fastest and Safest

Across QLC, TLC SSD and HDD

NVIDIA Reference Supercomputer: single tier full flash EXAScaler



AI is Flash HPC and Scientific AI is hybrid



Best IOPS & Throughput per rack

Up to 80M IOPs in a single rack
3.0 TB/s of Read throughput
1.9 TB/s of Write throughput
20PB of Flash

Best Price per Flash TB

Up to 14M IOPs in a single rack
360 GB/s of read throughput
240 GB/s or write throughput
Up to 26PB of Flash

Best Price per TB

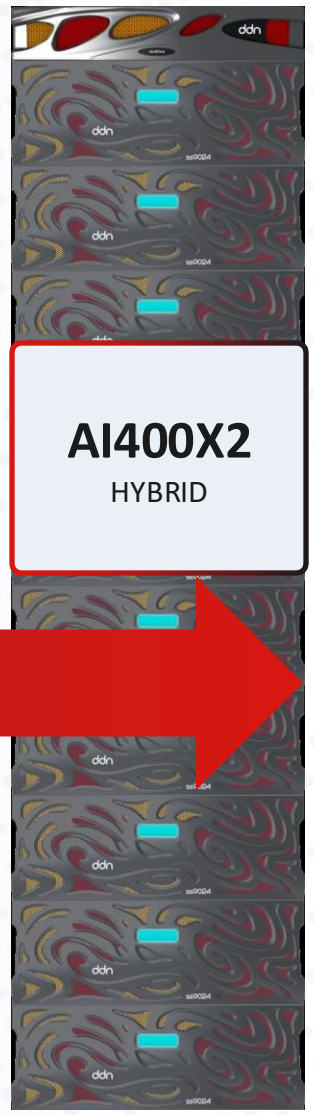
20PB per Rack
90 GB/s of read throughput
65 GB/s of write throughput

MASSIVE AI DATACENTER ACCELERATION

DDN HPC Storage tailored Data Solution

Q2 2025

Q1 2025



AI400X3I
Metadata and IOPS

AI400X3
Next Gen A3I

AI400X2T
A3I

AI400X2
A3I

AI400X2
QLC

AI200X2
HYBRID

AI400X2
HYBRID

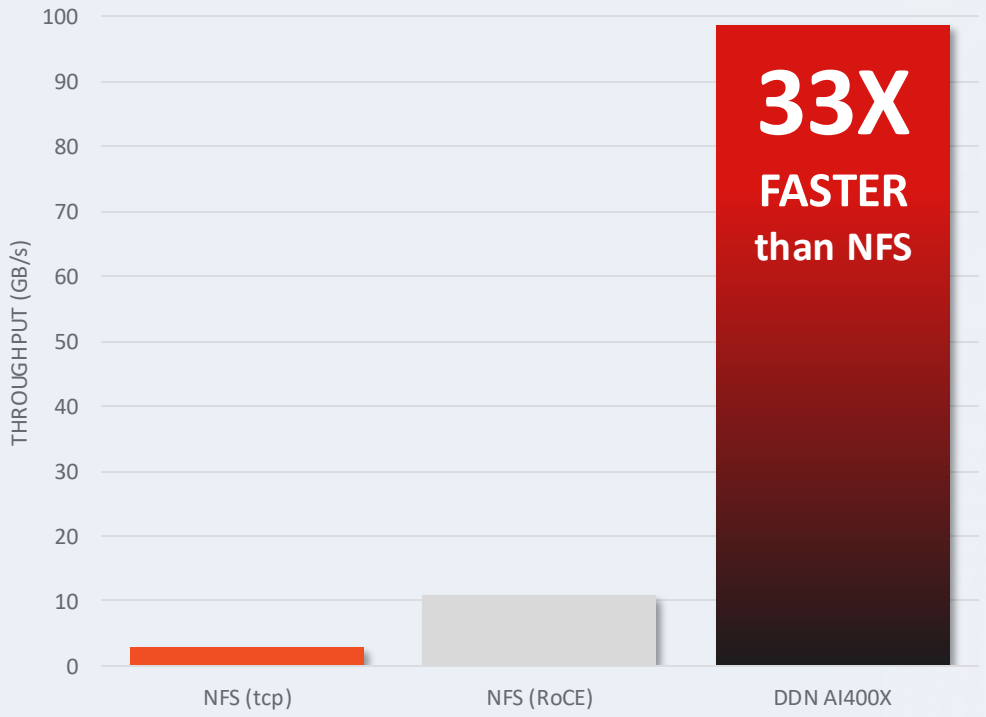
DDN DELIVERS Best $ per GB/s

DDN Delivers Best $/TB

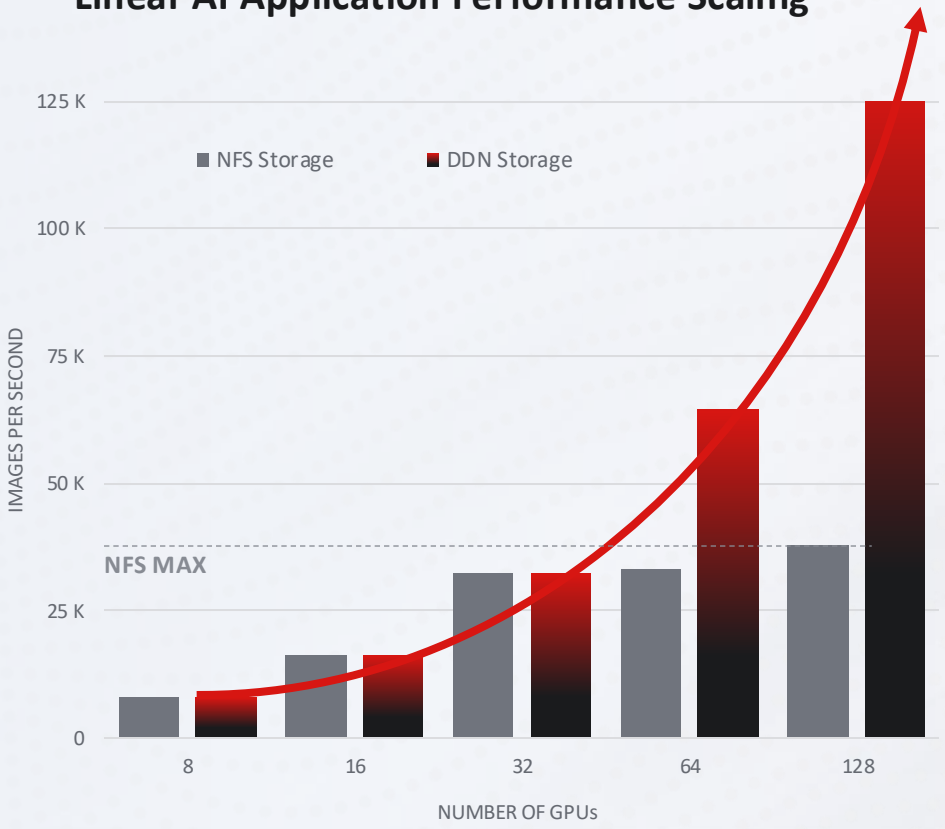
Performance with scalability is meaningless

DDN AI Solutions deliver full multi-dimensional performance and scaling with GPUs

Peak per client performance

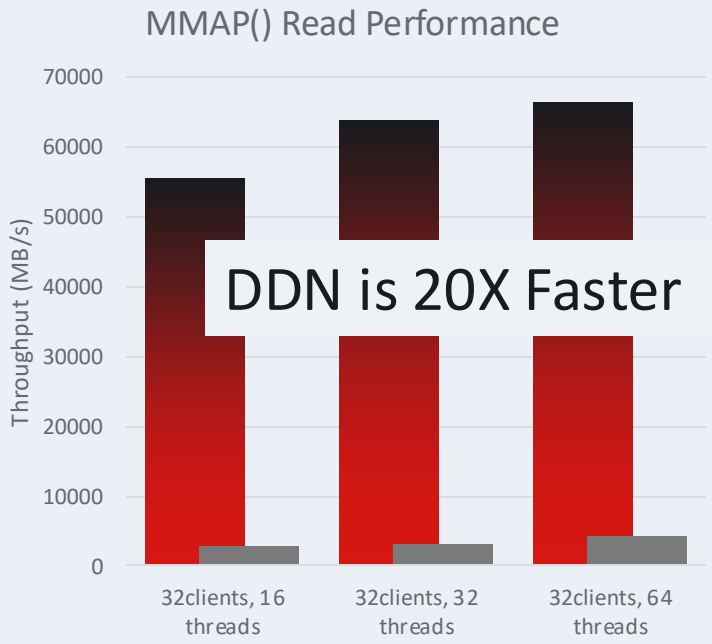


Linear AI Application Performance Scaling

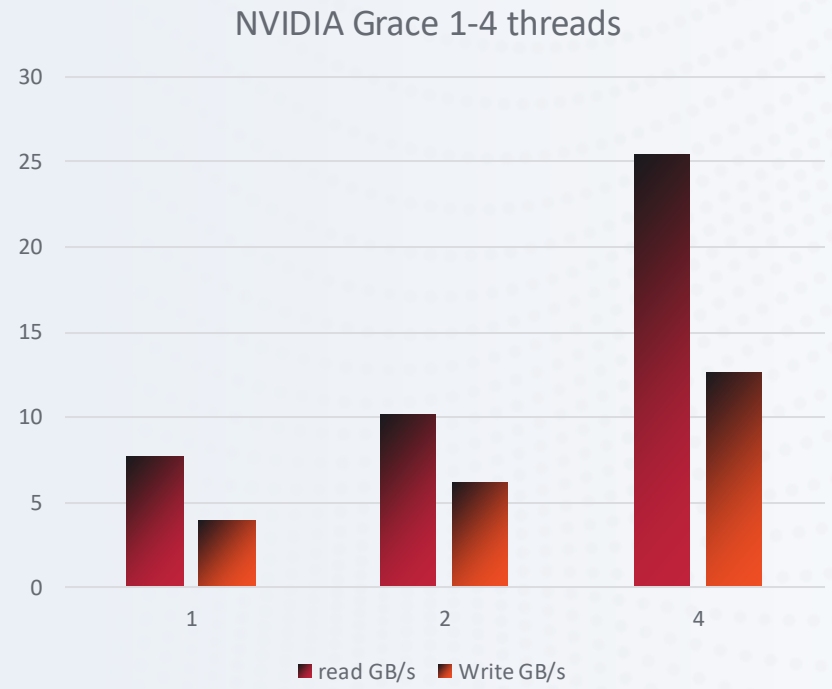


Performance is more complex than Bandwidth

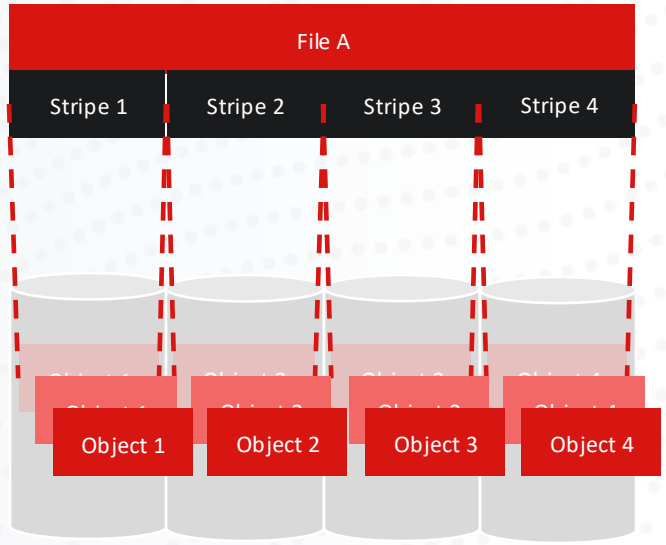
1 Optimize for the IO Call



2 Optimize for 1 thread



3 Scale many threads to 1 file



Complexity of AI IO workloads

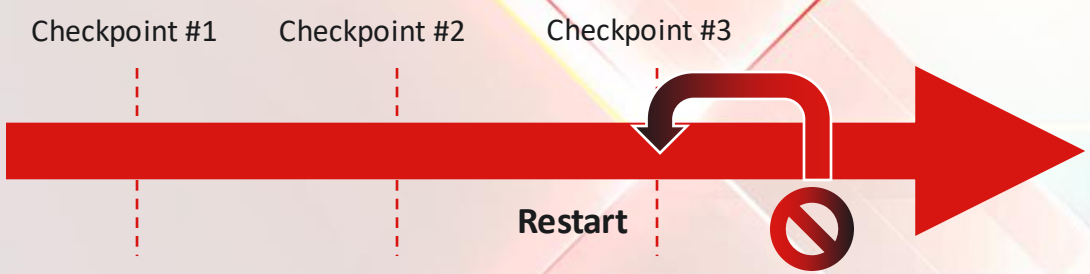
- Analysis of over 23,000 Machine Learning Jobs
- “Most ML jobs are perceived to be read-intensive with a lot of small reads while a few ML jobs also perform small writes.”
- “Our study showed that ML workloads generate a **large number of small file reads and writes**…”

~50% W/R ALSO VALIDATED BY NVIDIA DURING EOS PRESENTATION AT SC23. AVAILABLE ON DDN YOUTUBE.

Characterizing Machine Learning I/O Workloads on Leadership Scale HPC Systems
https://arnabkrpaul.github.io/publications/mascots21.pdf

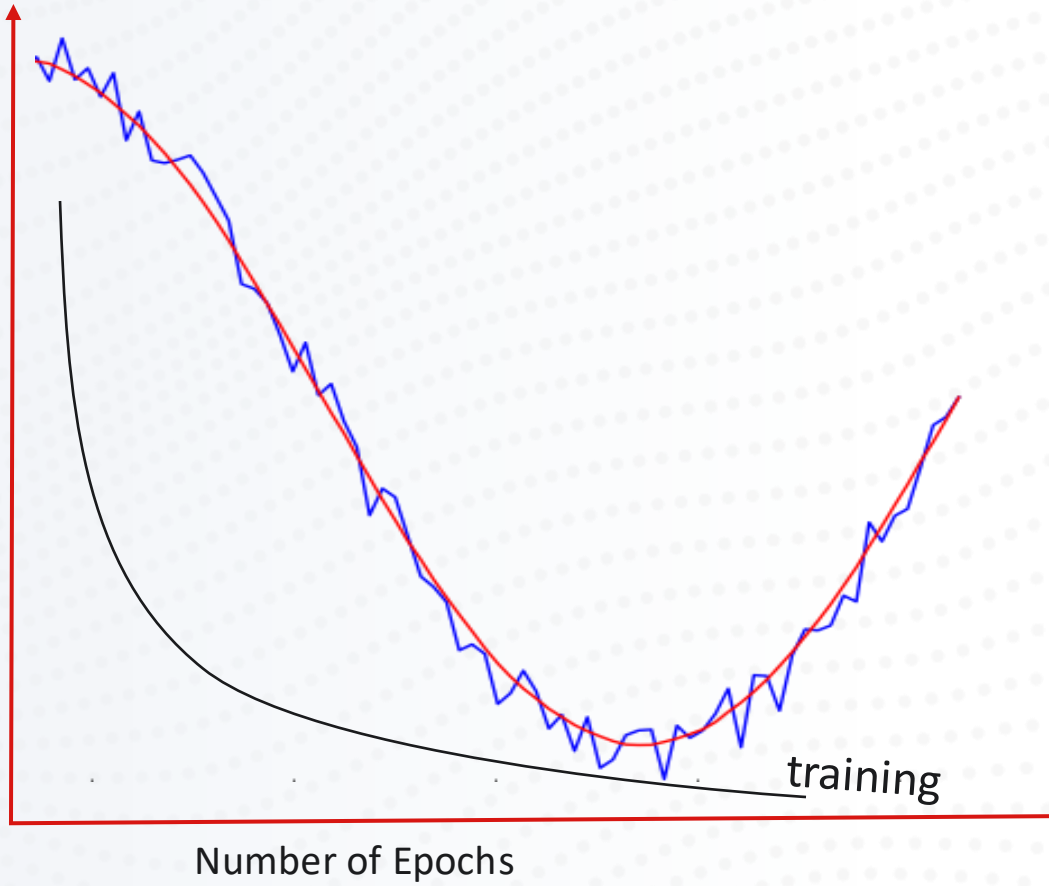
DDN improving AI pretraining time

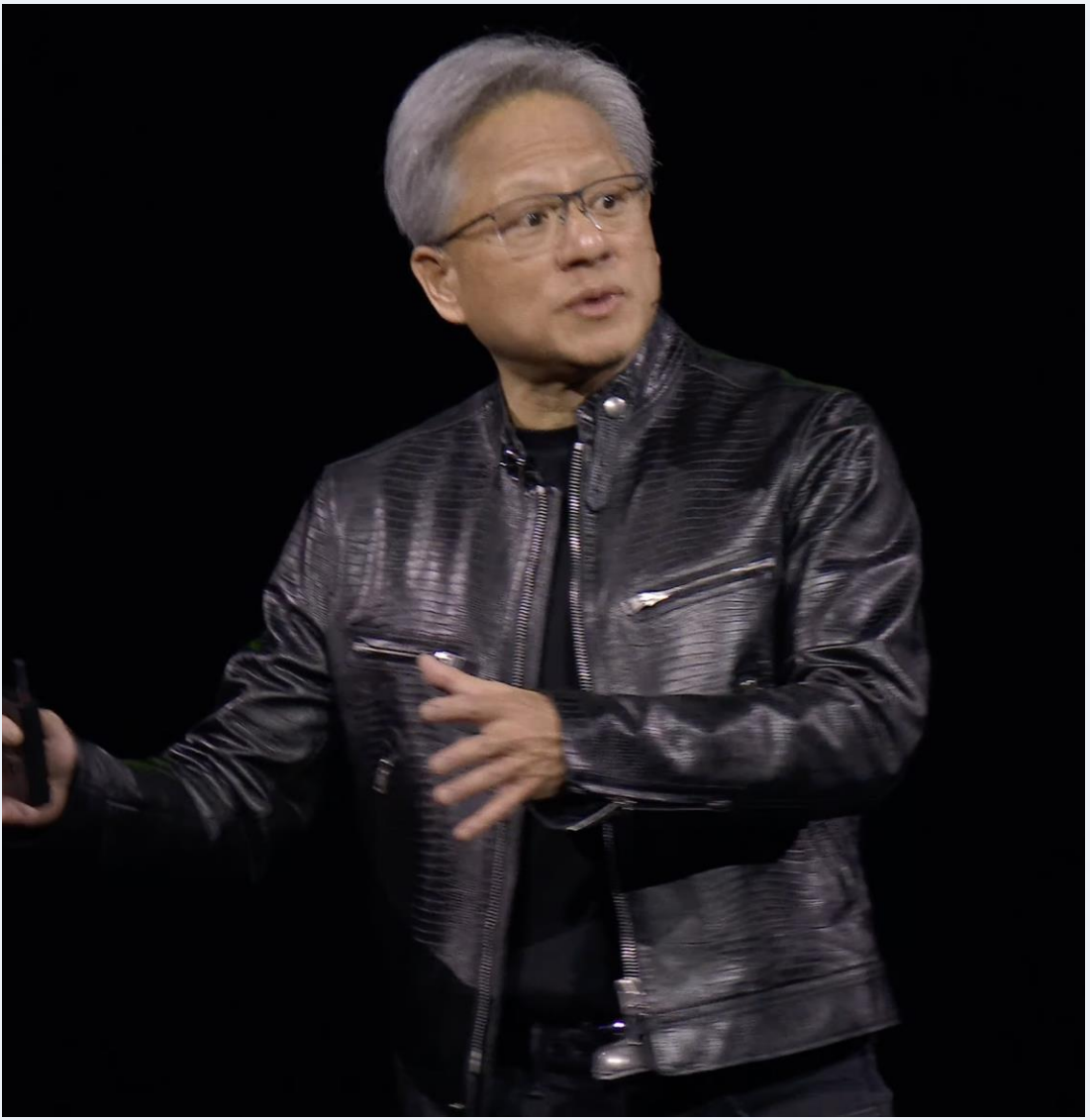
- Checkpoints are an intermediate dump or a snapshot of a model's entire internal state, including.
- Weights
- Learning rate
- Number of epochs executed, etc
- Checkpoints are a jumping-off point so that the framework can pick up on its training from here whenever needed.



EXAScaler: Extreme Performance Future proof architecture

- Checkpoint 176B parameters LLM: 3 sec
  - Checkpointing the model every 1h: 0.1% Overhead
  - Checkpointing a 530B parameters every 20 minutes: less than 1% overhead
- Frugal AI up to 13 GB/s per client node
  - Optimal usage of I/O bandwidth with a minimal number of GPUs
  - Tenant get storage performance without allocation of large chunks of the GPU space





The Challenges for Storage are Growing Rapidly

"…we checkpoint and we restart as often as we can"

"and that's our goal. Our goal is to **continuously drive down the cost and energy associated with the computing so that we can continue to expand and scale up the computation that we have to do to train the next generation of models**"

Jensen Huang - President, NVIDIA, GTC 2024

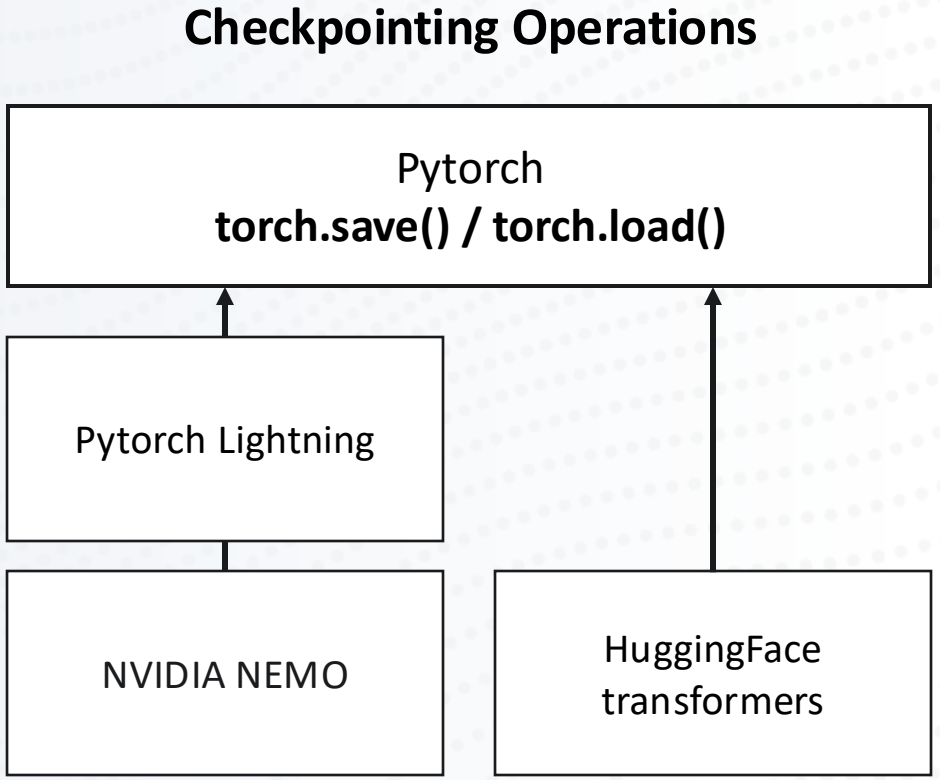
DDN Can Dramatically Increase AI Productivity

3x Faster Data Loads and 15x Faster Checkpoints

How checkpoints are performed?

- **Checkpointing with PyTorch**
- For pytorch based libraries, such as NVIDIA NEMO, Pytorch Lightning, or even HuggingFace, the checkpointing operations are handled by **torch.save()** and **torch.load()**

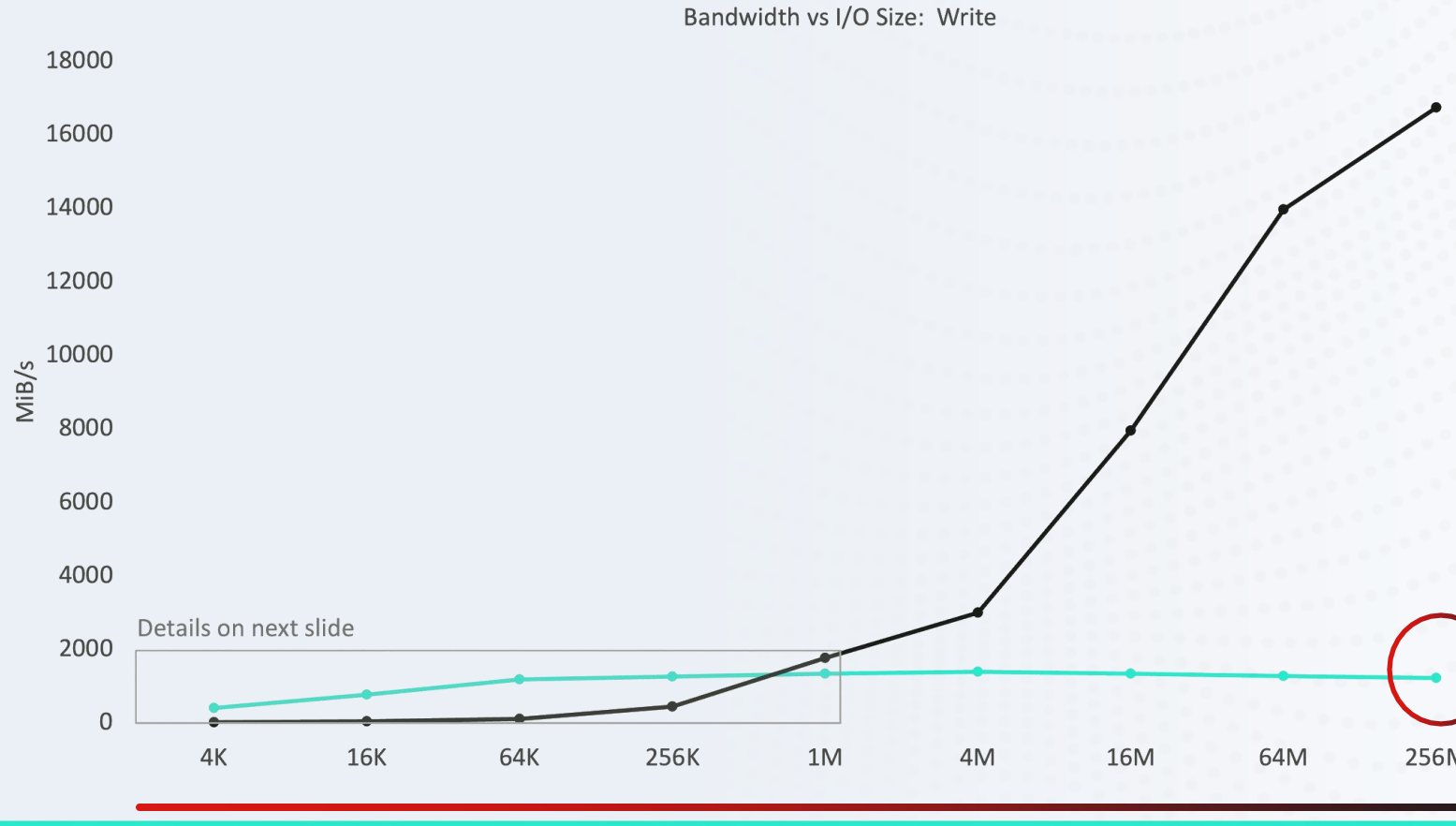
- **These two operations perform large buffered I/O operations**
- *write(60GB), write(58GB), write(56GB),…*



Torch.save/load

- **Performance issue**

Buffered vs Direct: Performance with I/O Size



Whamcloud

Buffered IOs

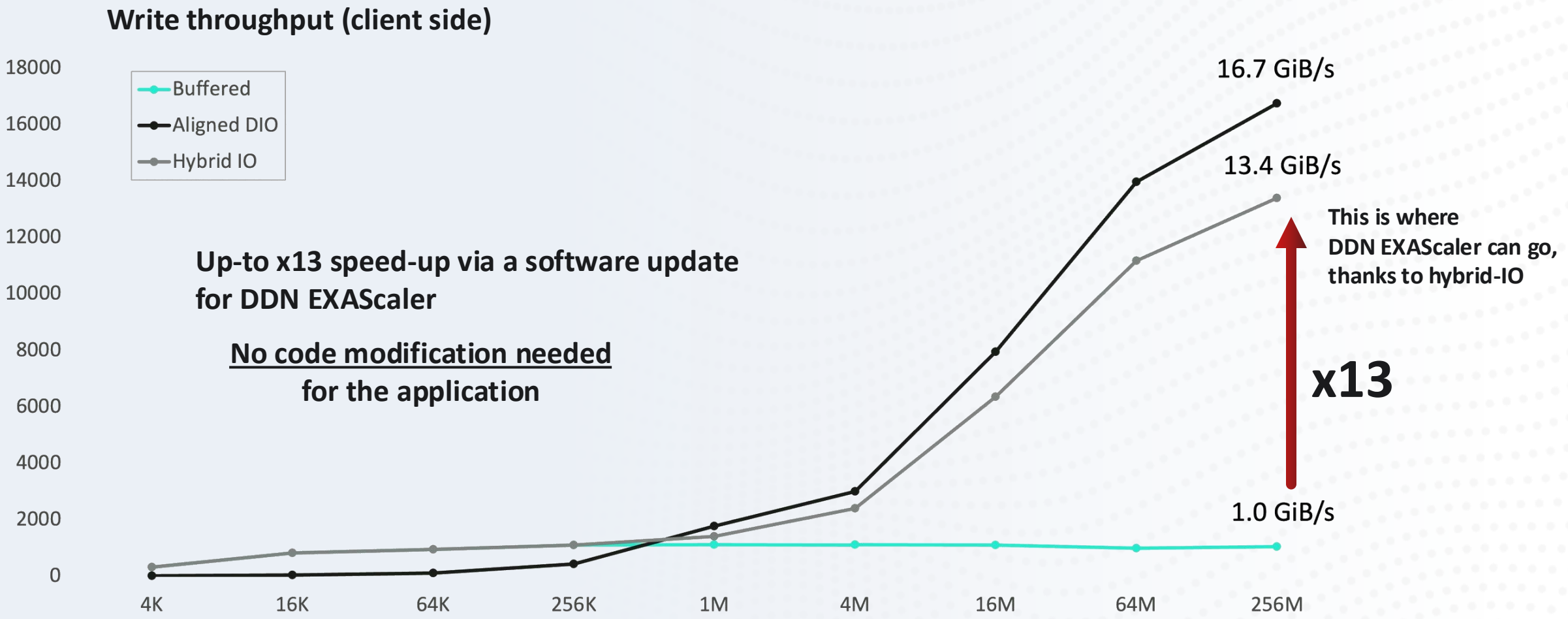
«very large buffered IOs »
Torch.save/Torch.load

Very large

dən

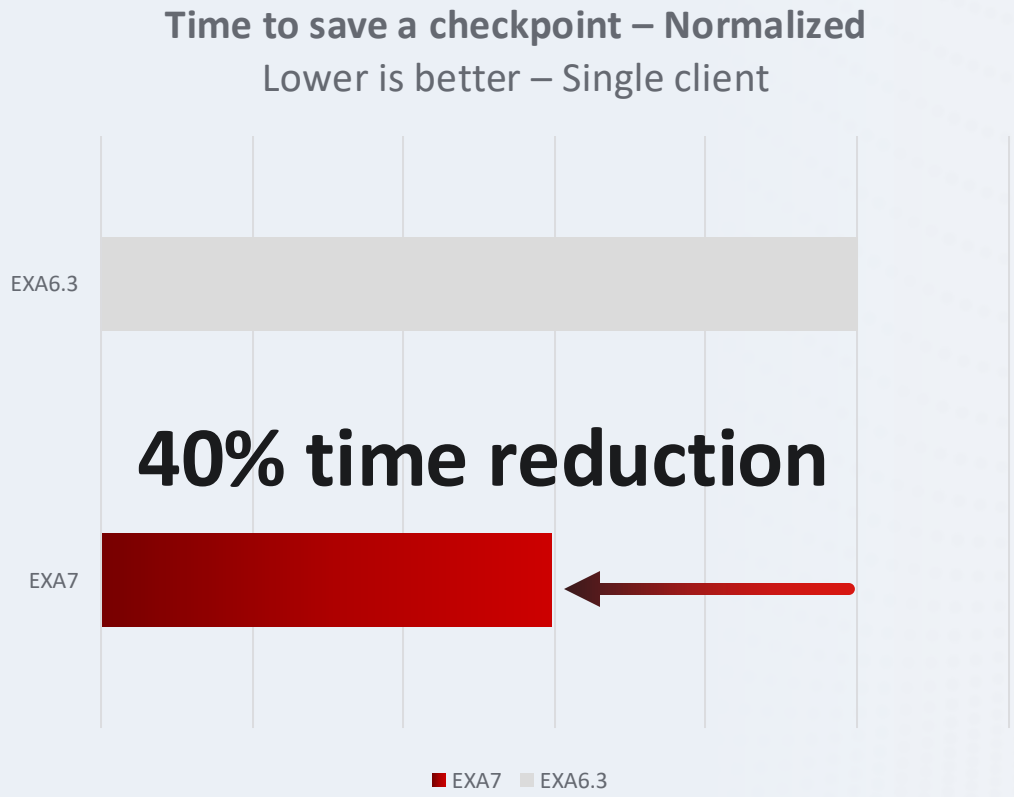
Hybrid-IO – From EXA7 software update

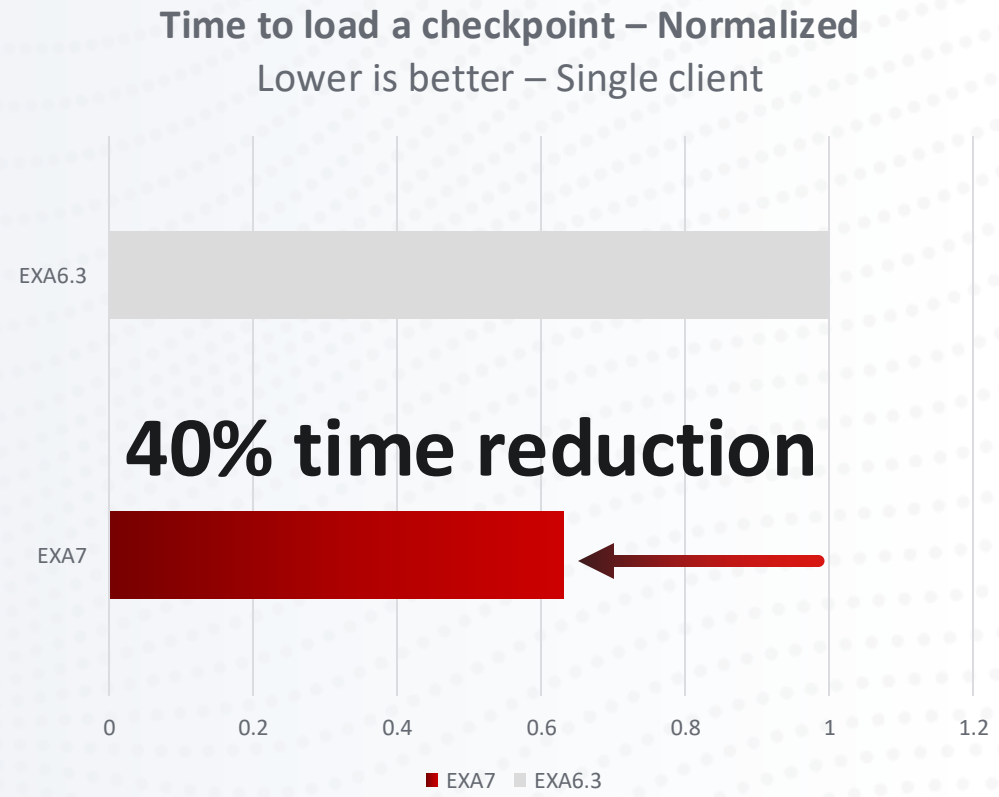
- **The best of both worlds**



Checkpointing kernel performance

- **EXA6.3 vs EXA7**





HPC workloads are well characterized

IO500

- Performance comparison across the industry
- Evolution over time
- https://io500.org/

Darshan

- De facto standard tool: https://github.com/darshan-hpc/darshan
- EuroHPC I/O trace analysis: https://hpcioanalysis.zdv.uni-mainz.de/involved

AI characterization is still elusive

MLPerf Storage

- https://mlcommons.org/benchmarks/storage/Evolution over time

Base II for Scientific AI

- https://excalibur.ac.uk/resources/base-ii-blueprinting-ai-for-science-at-exascle/

NVIDIA NSight

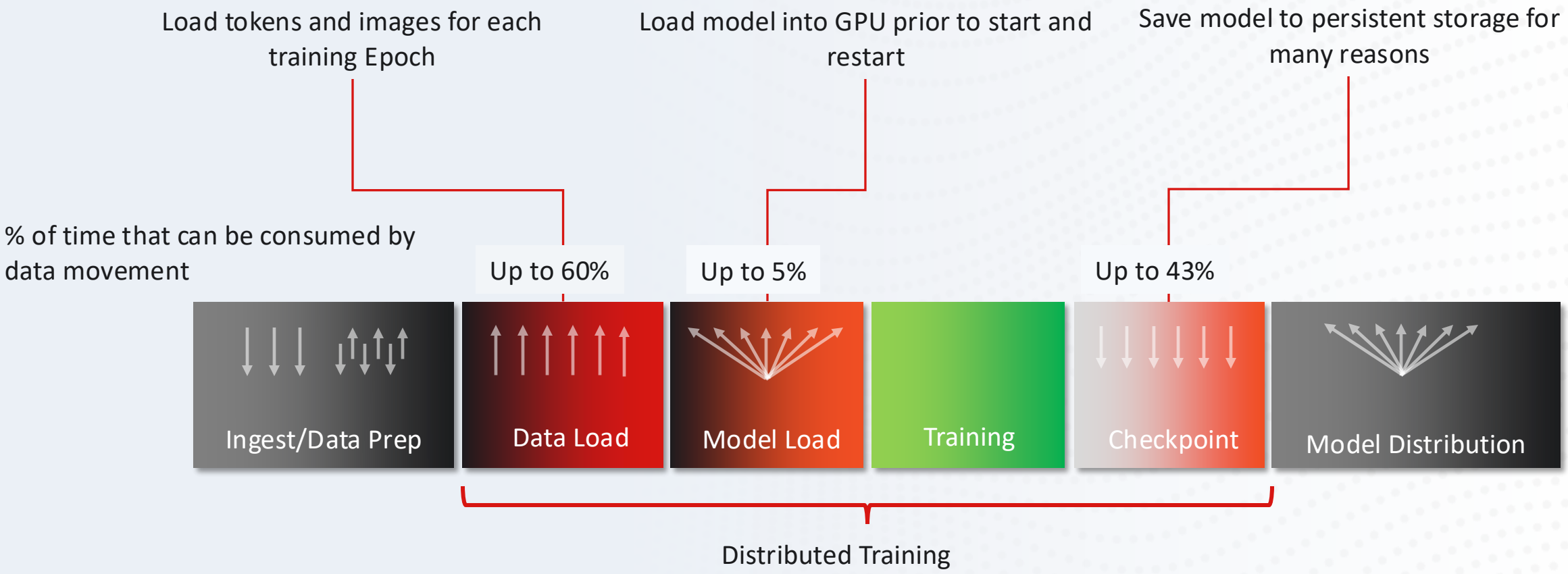
- Working with NVIDIA to extend NSight profiling tools for IO
- https://developer.nvidia.com/nsight-systems

Up to 50X Data Reduction

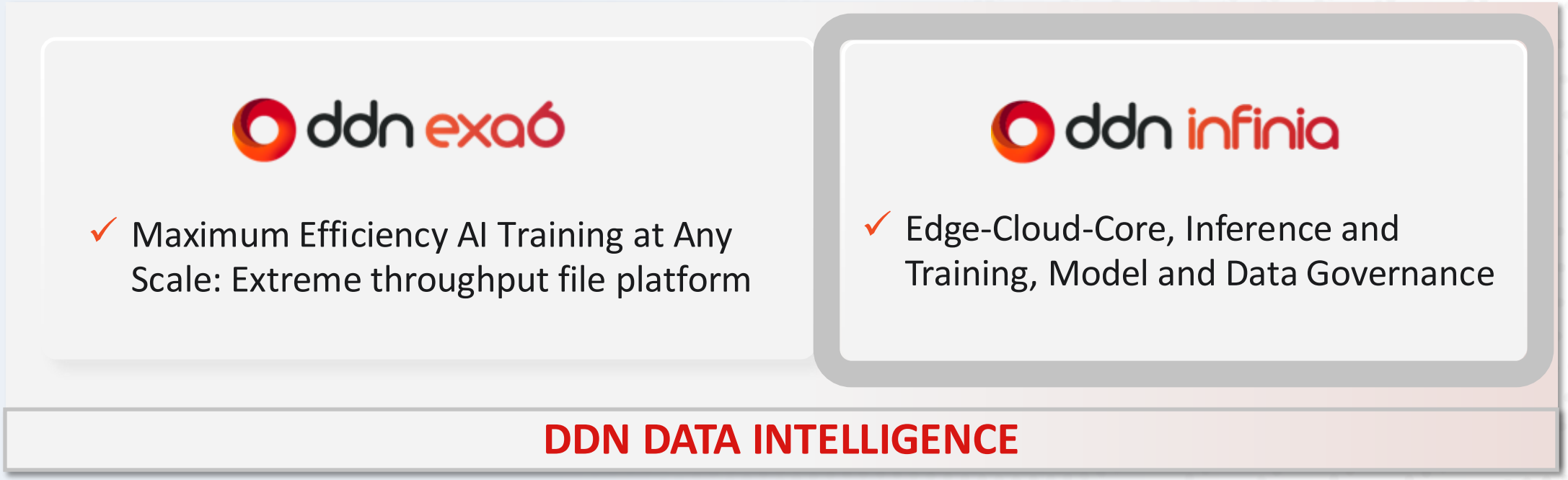
	FAST Compression ZSTD:1 (1MiB Chunk)	**Deep Compression** ZSTD:11 (1MiB Chunk)
Life Sciences		
Genomic Data: Sequence files (e.g., FASTQ, BAM)	**3.33**	**3.66**
Bioinformatics Data: Processed genomic data and analysis results (VCF)	**5.85**	**8.68**
Artificial Intelligence (AI) and Machine Learning (ML)		
Training Data: Large datasets including images, text, audio, video		
Training Data: Fineweb - 15T token web dataset for LLM Training in parquet format	**2.37**	**2.89**
Training Data: Audio (google-fleurs speect international training set, in FLAC format)	**1.1**	**1.03**
Training Data: Images (RAISE - RAW image dataset)	**1.09**	**1.17**
Model Data: Parameters and weights of trained models	**2.08**	**2.35**
Financial Services		
Transactional Data: High-volume transactional records and logs	**6.07**	**7.53**
Operational Data		
Operational Data: Logs, configuration files, monitoring data (SPLUNK)	**2.33**	**2.43**
Operational Data: Raw Text logs and configuration files (DataFlow and system sos log bundles)	**41.07**	**49.43**

[TRAINING IS ONLY A LINK OF A VALUE CHAIN]

AI data life cycle goes beyond training



From Raw Performances to Business Outcomes



DDN Infinia: A unified, cloud-optimized platform for high-performance data and AI-driven insights at scale.

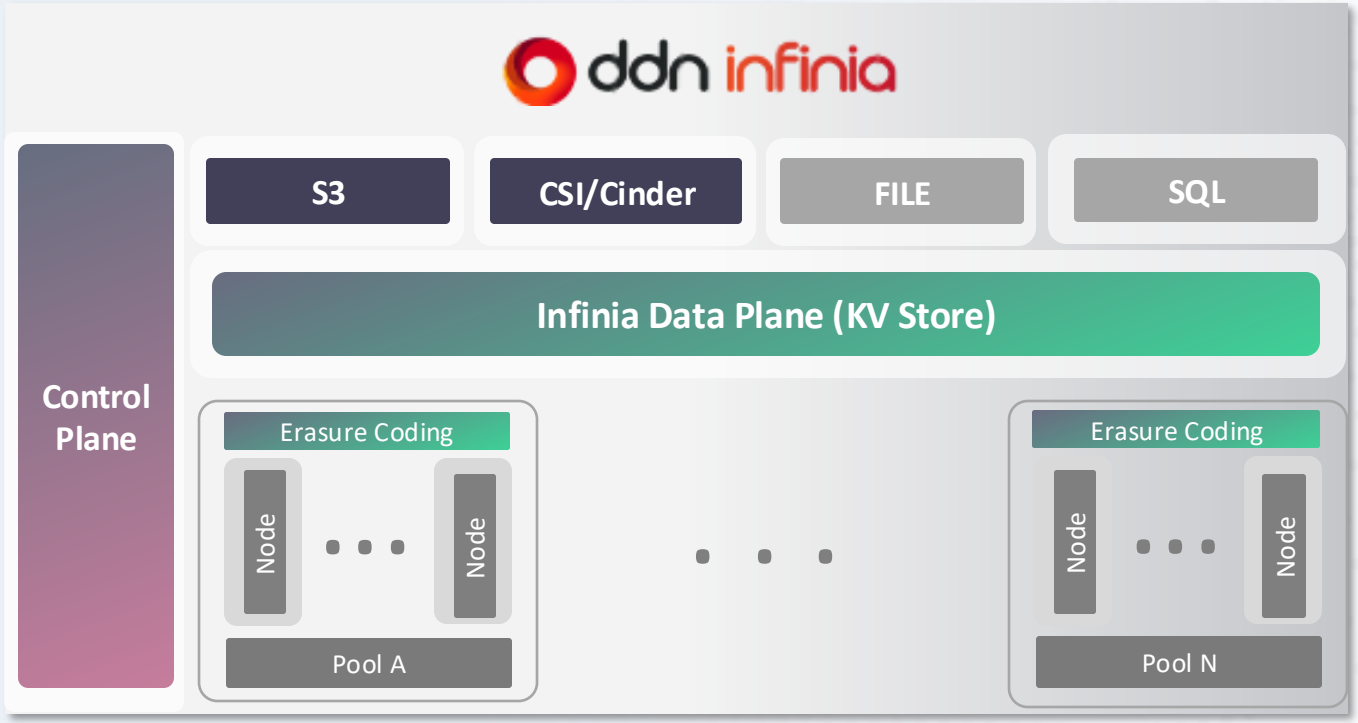
Unlock the power of structured, semi-structured and unstructured data with seamless access, management, and analytics across edge, core and cloud.

- **Structured AND Unstructured data**
- **Storage Delivered as a Service**
- **Hyper Simple Operation**

DDN Infinia's Software Architecture: Versatile and Powerful

Control Plane

- Built for Integration
- API First
- GUI & CLI



Data Services

- Lightweight & Containerized

Enterprise Data Plane

- Scale Out
- Elastic IOPS/Bulk
- Always On
- Data Reduction
- Encryption
- Snapshot
- Multi-tenancy

100% Hardware Agnostic

- Any Intel, Arm or AMD CPU
- Any FLASH (TLC/QLC/PLC)

Fully Clustered

- Fault Domain Aware
- Network Erasure Coding

[S3 AS A DEFACTO STANDARD]

S3 is the critical data service for AI

S3 Compatibility details ~80% passed

testsuite	s3tests_boto3		
category	(Multiple Items)		
Count of result_0711	**Column Labels**		
Row Labels	**FAILED**	**PASSED**	**Grand Total**
acl	15	5	20
bucket_ops	3	28	31
bucket_policy	17		17
bucket_policy_status	7	2	9
ceph_utils		1	1
copy_ops		4	4
general	14	32	46
getput	2	26	28
headers	12	36	48
list bucket	28	55	83
mpu	10	14	24
objectcopy	2	15	17
post	19	16	35
presigned	7	12	19
versioning	6	16	22
Grand Total	**142**	**262**	**404**

7/10/2024

testsuite	s3tests_boto3		
category	(Multiple Items)		
Count of testcase	**Column Labels**		
Row Labels	**FAILED**	**PASSED**	**Grand Total**
acl	15	5	20
bucket_ops	3	28	31
bucket_policy	4	13	17
bucket_policy_status	7	2	9
ceph_utils		1	1
copy_ops		4	4
general	11	36	47
getput	3	25	28
headers	12	35	47
list bucket	7	76	83
mpu	8	16	24
objectcopy	2	15	17
post	1	34	35
presigned	6	13	19
versioning	4	18	22
Grand Total	**83**	**321**	**404**

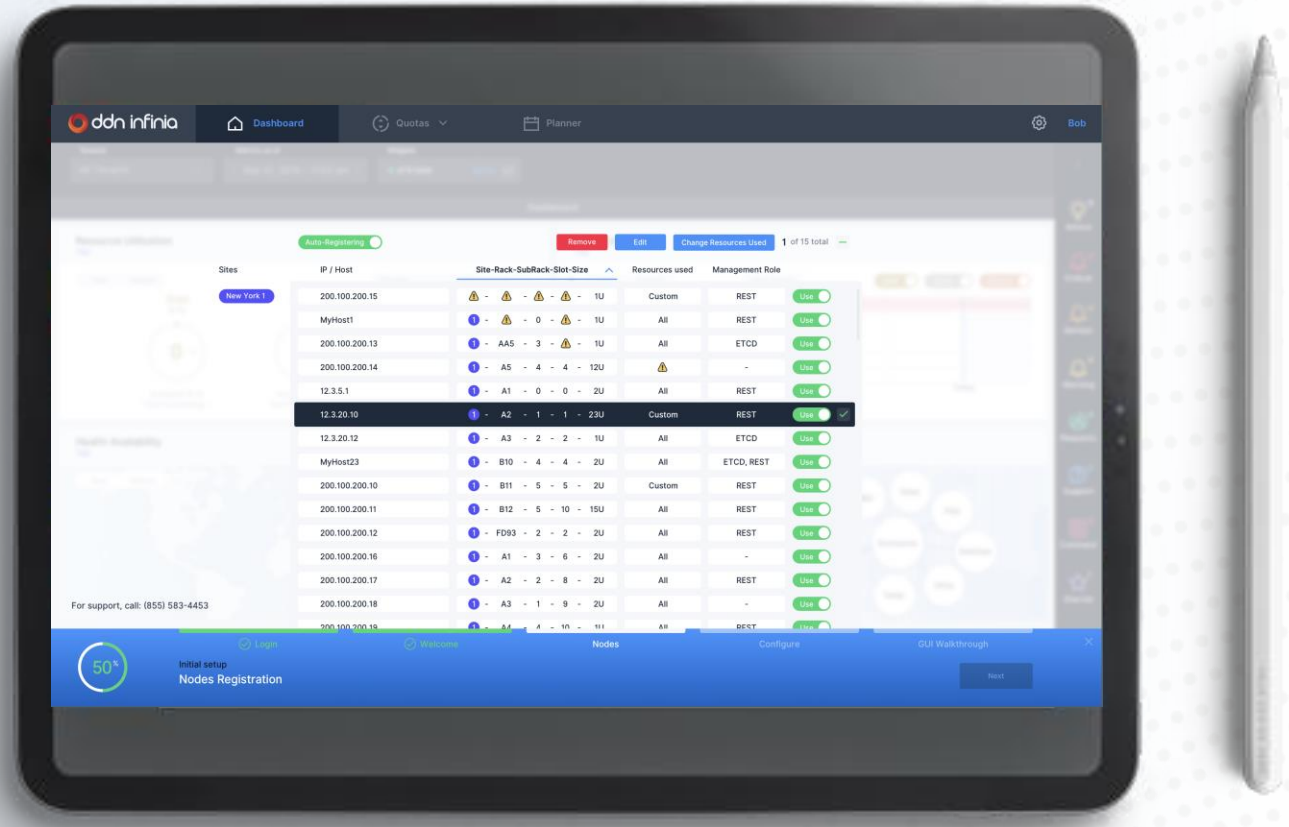
8/8/2024

Pre-GA numbers illustrative purpose only

DDN INFINIA: Flash only consumer grade QLC

Auto-Installation, **10 minutes to Deploy**

Deploy new services and tenants in just 4 minutes



No kernel module for AI workloads

User space client only

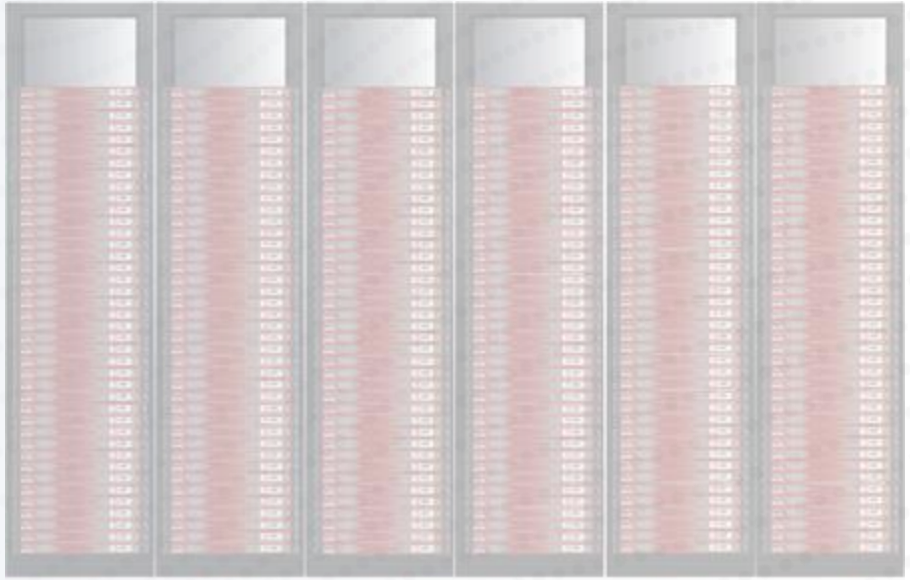
- Deployment constraints, abstraction level
- Operating systems understand files and blocks, not objects

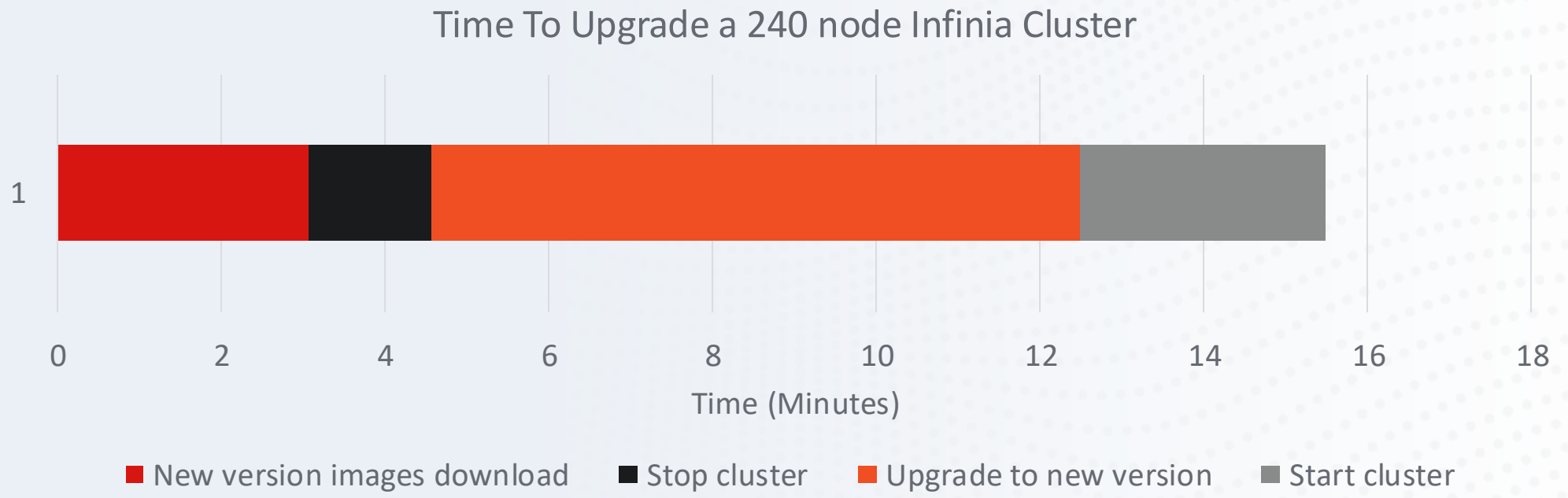
DDN Libfuse maintainer since Q1 2024

- FUSE-over-IO-URING Goal:
- Performance
  - Reduction of kernel/user-space transitions
  - NUMA awareness and core affinity

[OPERATIONAL EFFICIENCY]

Fast Installation, Fast Upgrades

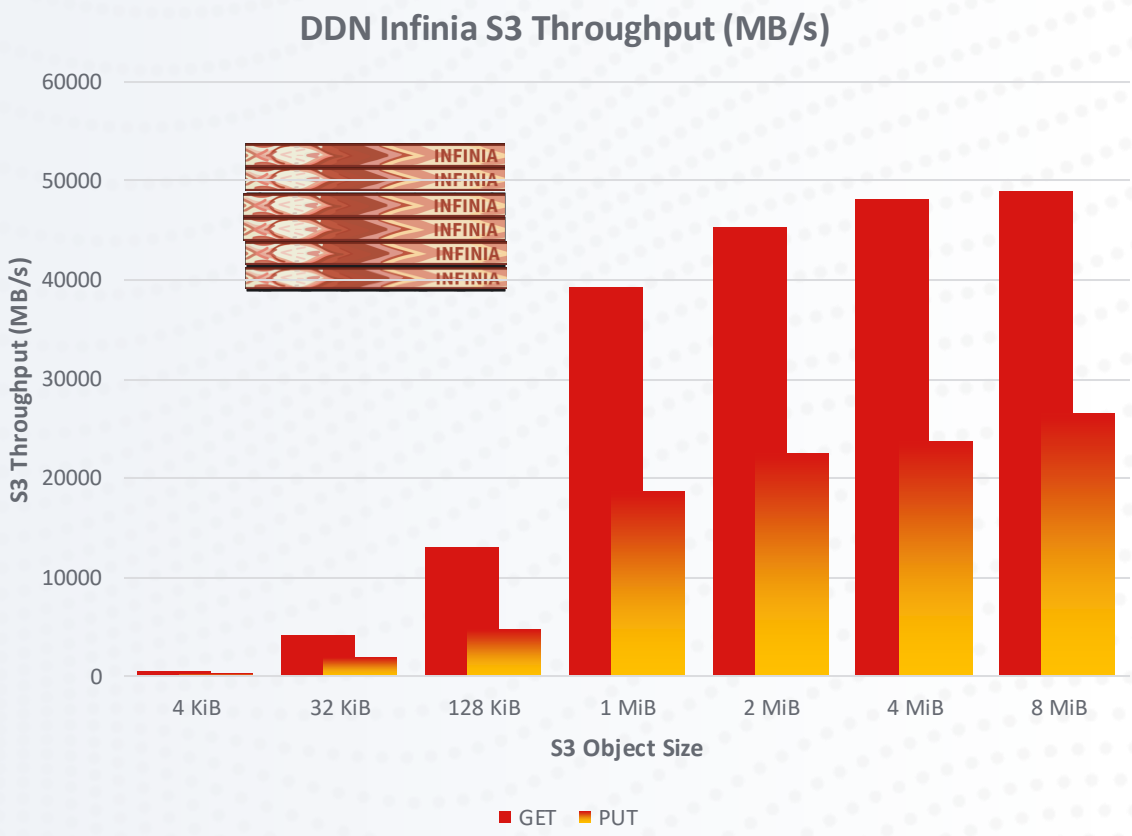




Infinia is Fast - Accelerates S3 workloads by 10-100X

- DDN Infinia exceeds the claimed performance of all other object stores in Objects/sec and Throughput
- Accelerate Enterprise Analytics workloads like Apache Spark, Starburst Presto/Trino, Clickhouse,

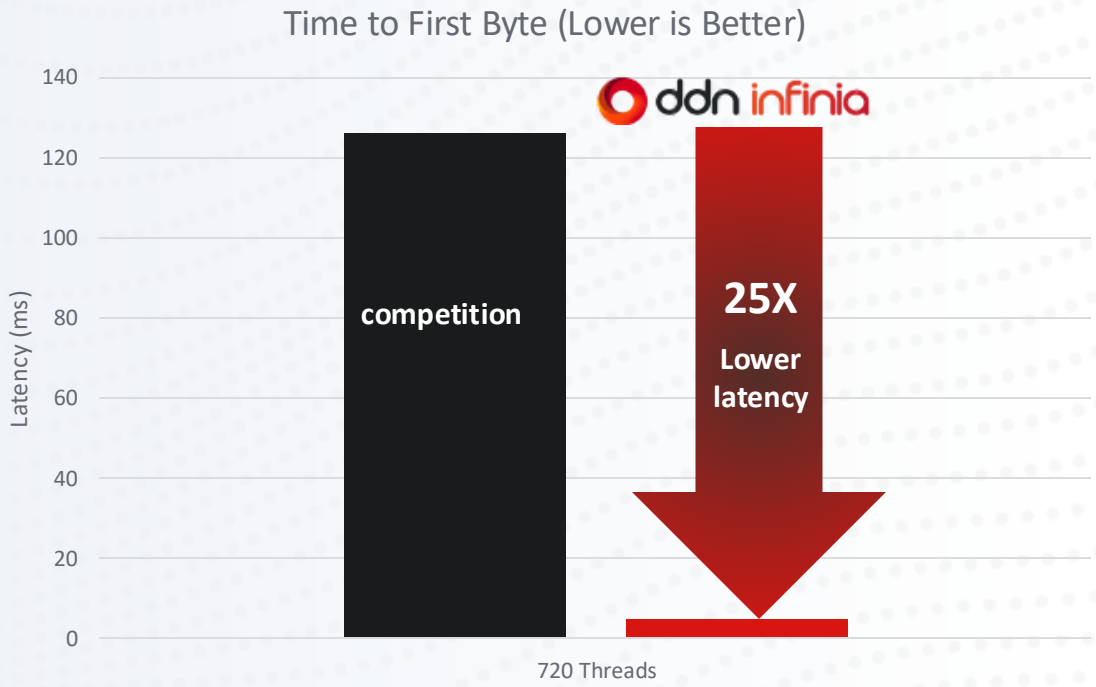
Object Size	GET	PUT
4K	**346 K**	**69 K**
4M	**49 GB/s**	**34 GB/s**



Importance of Time to First Byte Latency

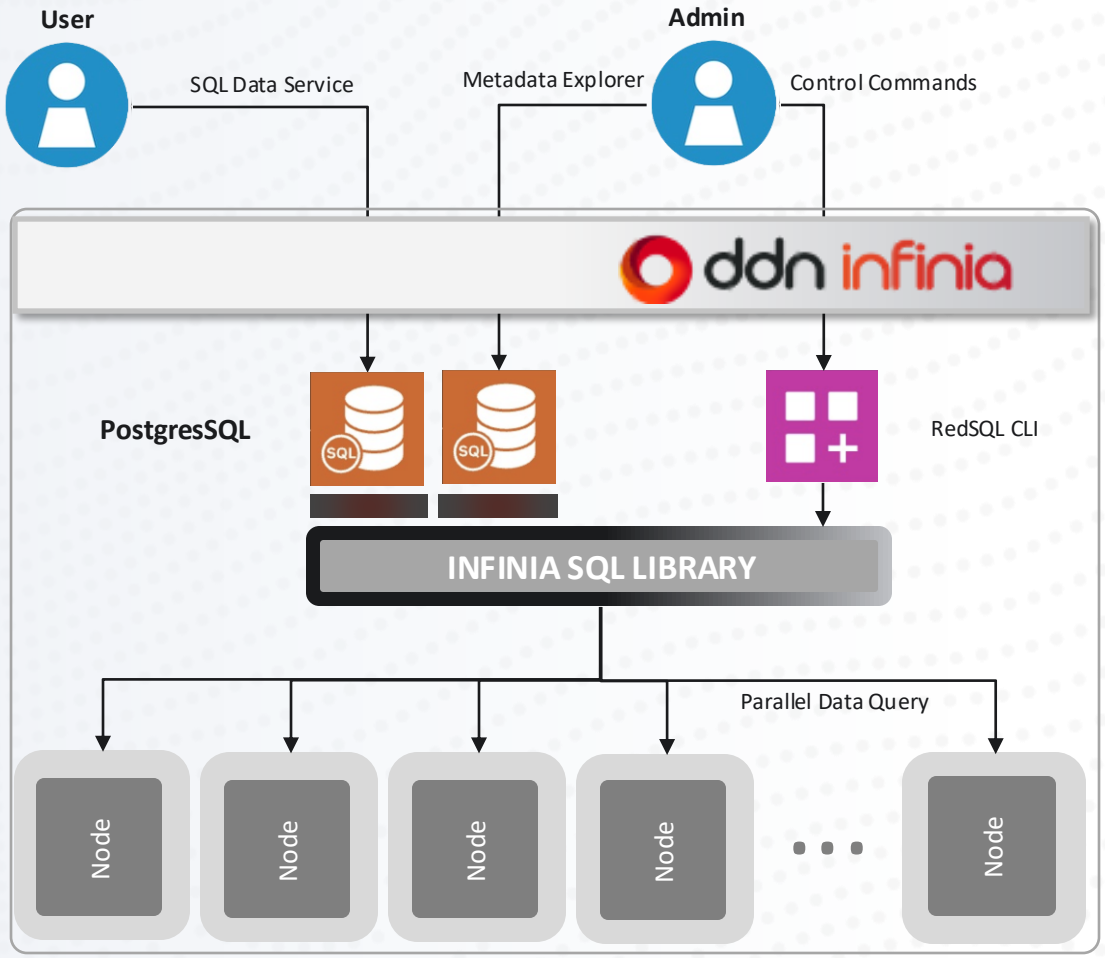
- DDN Infinia exhibits the lowest time to first byte of any object store.

Infinia is 25x lower latency at high concurrency



Data Lakehouse: Infinia SQL Service

- PostgreSQL Front-End
  - Comprehensive SQL support, widely adopted, easy integration
- Infinia SQL
  - Distributed SQL data service built on Infinia KV
  - table operations: CREATE/DROP
  - row operations: INSERT/UPDATE/DELETE/SELECT
- Infinia SQL metadata catalog
  - System metadata explorer with SQL interface
  - Efficient! no additional snapshot/copy/space overhead, consistent catalog

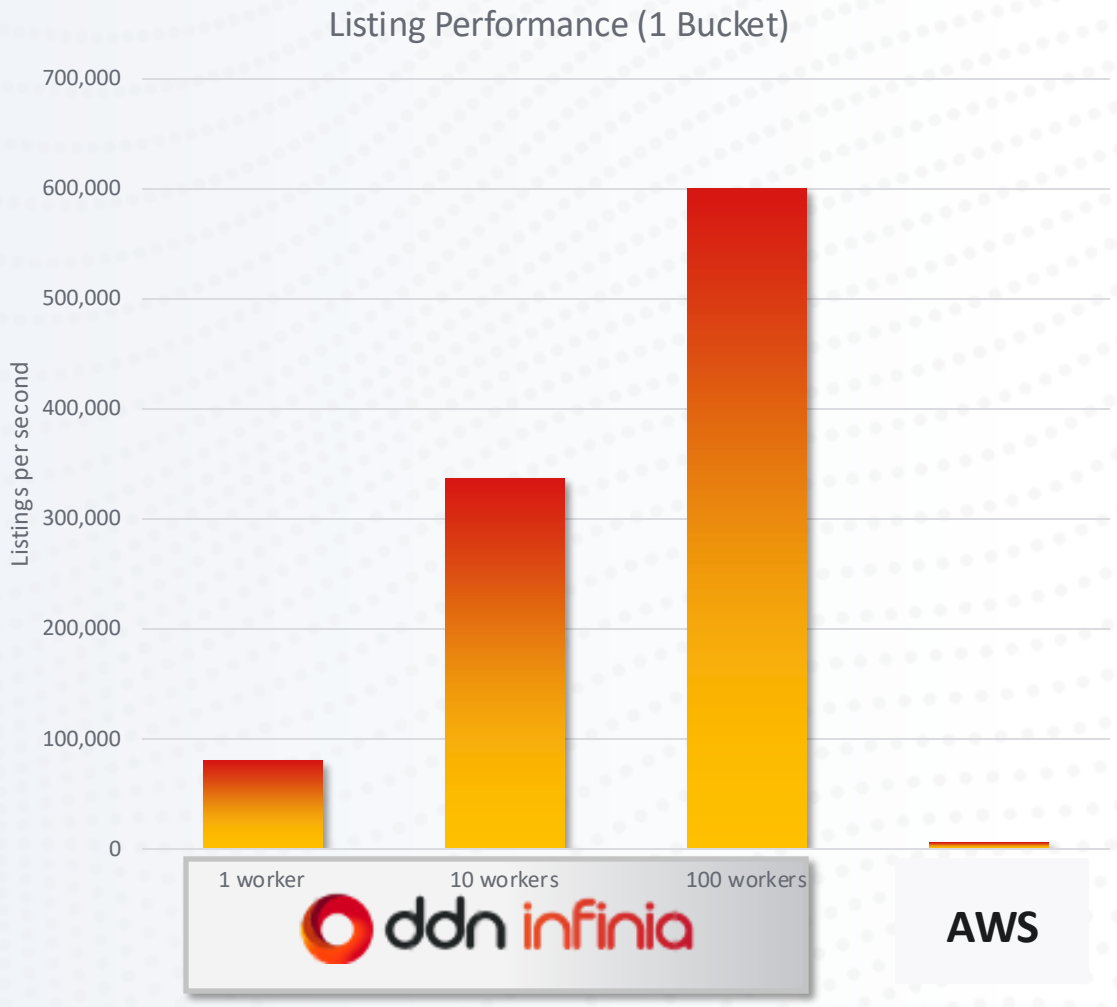


Fast Object Listing + SQL = Dataset selection

```
-- Satellite images by time range
Sql> SELECT * FROM Copernicus WHERE date >= '2025-01-01' AND
date < '2025-02-02';

-- Satellite images for a given area and time range
Sql> SELECT * FROM Copernicus WHERE date >= '2025-01-01' AND
date < '2025-02-02' AND latitude BETWEEN min_latitude AND
max_latitude AND longitude BETWEEN min_longitude AND
max_longitude;

-- Satellite images for a given area and time range and
a specific feature
Sql> SELECT * FROM Copernicus WHERE date >= '2025-01-01' AND
date < '2025-02-02' AND latitude BETWEEN min_latitude AND
max_latitude AND longitude BETWEEN min_longitude AND
max_longitude AND water_body_presence > 0.85;
```



Comparison with Other Catalog

	Other Catalog	Infinia SQL
Catalog mechanism	Periodic snapshot (e.g., 15 seconds)	Realtime mapping
Consistency	Always behind the file system	Consistent
Tables	Single table	Multiple tables (extensible)
Scope	Filesystem	All system metadata including filesystem
Integration	Trino, Python-based SDK	PostgreSQL FDW
CLI	big-catalog (UNIX find like CLI)	redsql, psql

[CLOUD CONSUMPTION MODEL]

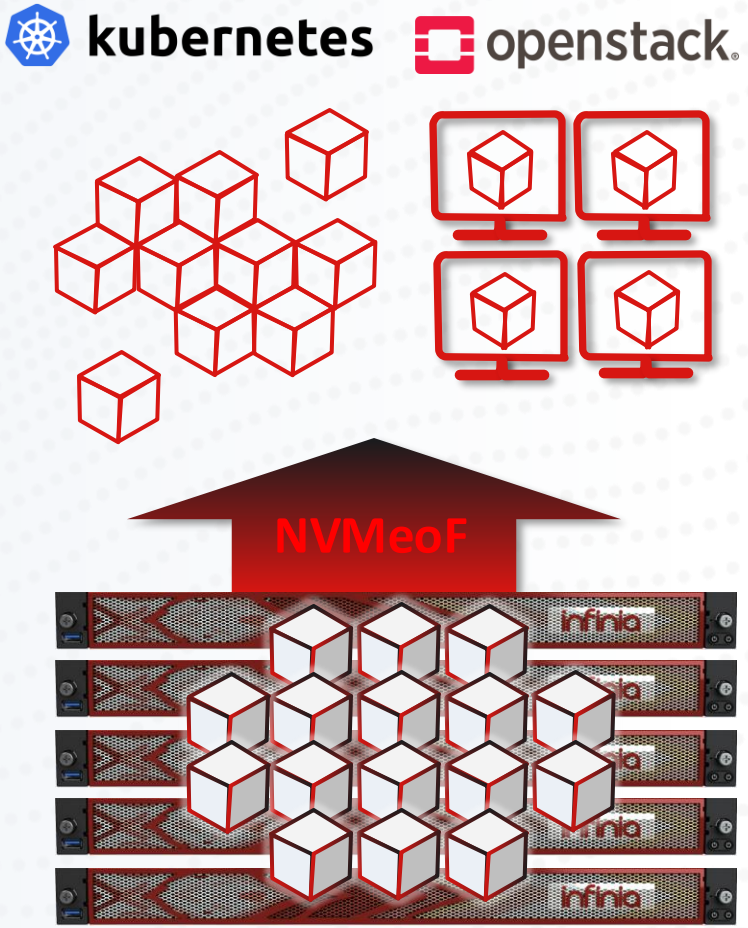
Multi-Cloud Ready Kubernetes Storage

The complete multi-cloud ready Kubernetes storage platform with elastic scalability, unmatched availability, and self-service access to any storage.

- Scalable persistent storage
- Multi-cloud Data Mobility
- Zero RPO Disaster Recovery

DDN Infinia provides Low Latency, High Bandwidth Block Exports via NVMeoF TCP(1.0) / RoCE(1.1) for Kubernetes and Openstack

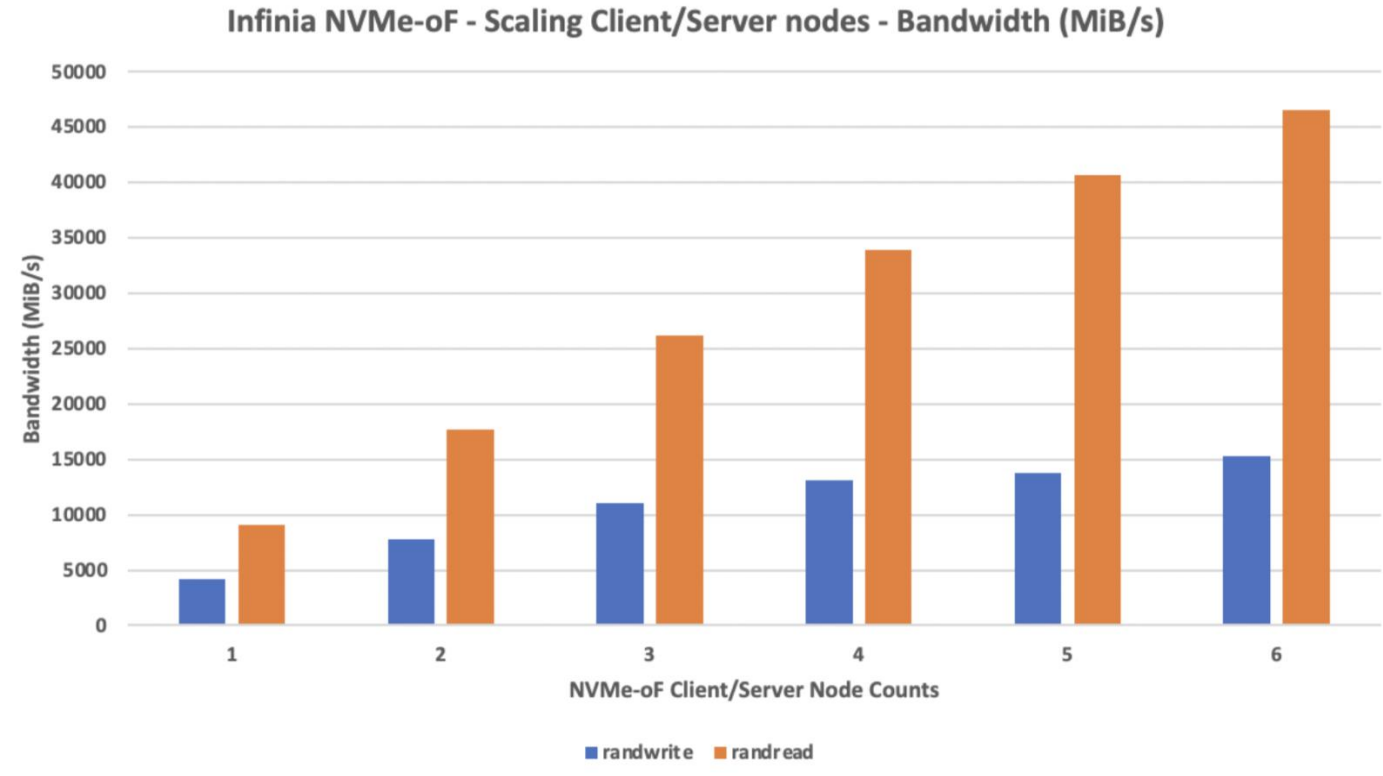
- CSI Drivers 1.0
- Openstack Cinder Drivers 1.1
- Single NVMe-oF Client Read over 15 GB/s



Infinia NVMeoF Performance

- The scaling Infinia NVMe-oF (NVMF) server/node performance of the latest Infinia Build (Feb 9th) is assessed using the fio benchmark.

- The setup comprises of 6 x NVMF clients connected to 6 x Infinia NVMF servers over 2 x 100GbE.



Native Multi-Tenancy: Building a Cloud Infrastructure

Realm Admin

Manage HW

Manage Tenants

realm_admin_001

Tenant Admin

Manage users/groups/identity

Manage dataservice and subTenants

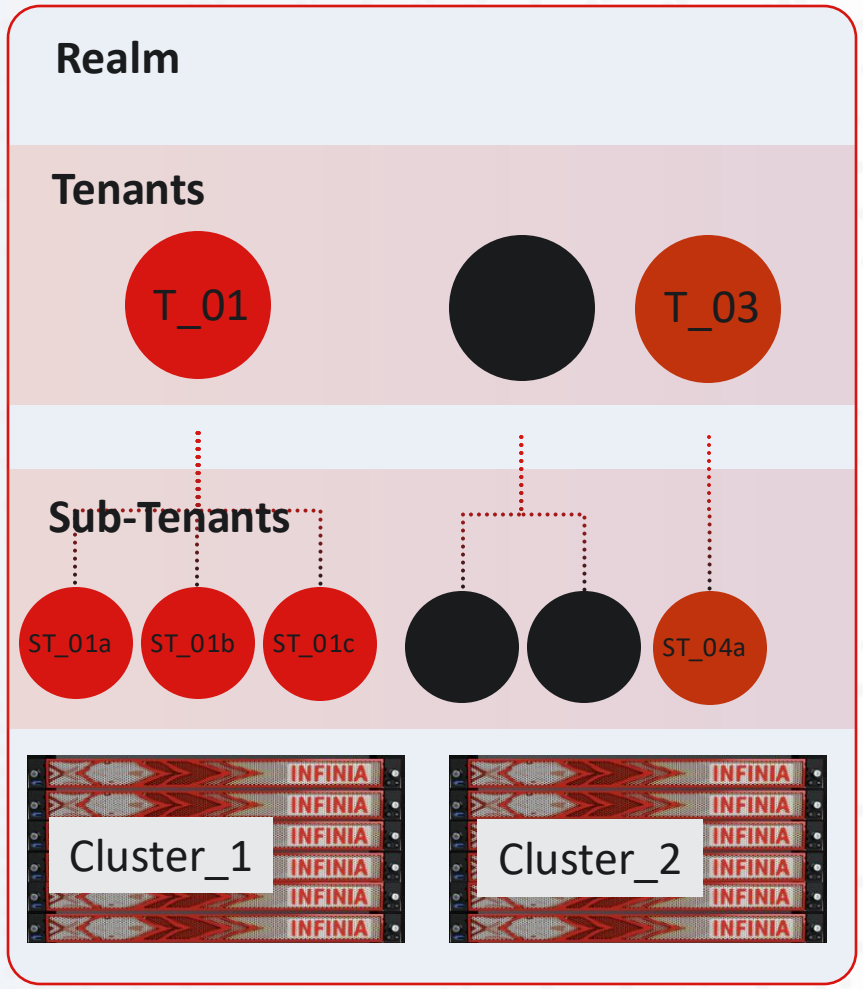
tenant_admin_001

SubTenant Admin

Manage Dataserice

Manage Dataset Sizes

redadmin_01



Protect Tenant service levels with no compromise on efficiency

Native Multi-Tenancy

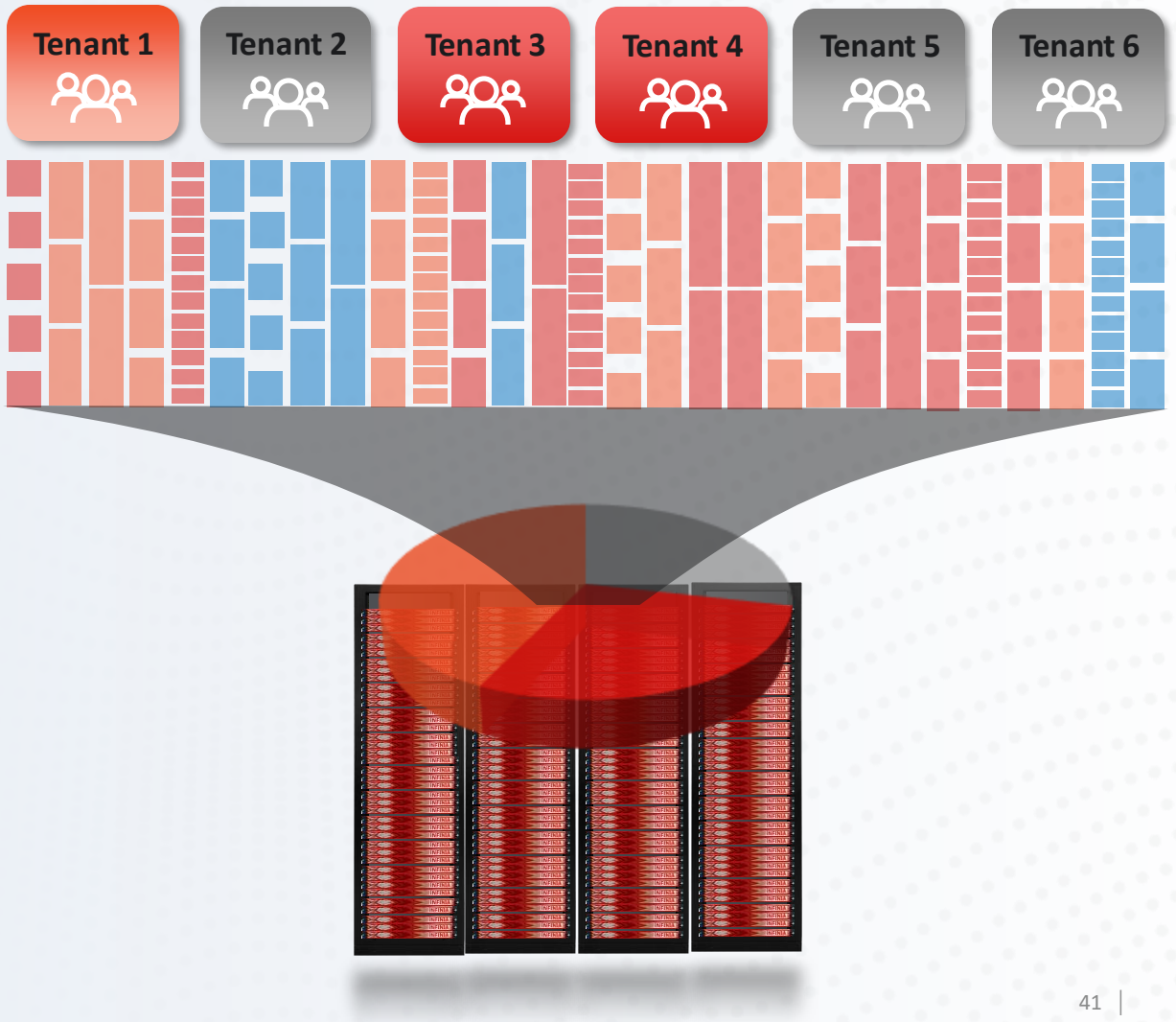
Easy Simple Creation of Tenants and Subtenants

Real-Time QoS

Real-Time Performance Allocator delivers to policy

Maximum Utilization

DDN Fair Share gives full utilization regardless of tenant setup

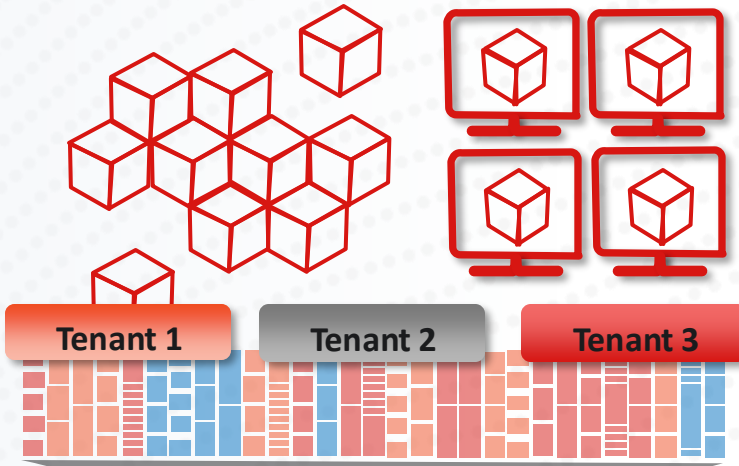


VERSATILE MULTI CLOUD MULTI-TENANT

CLOUD CONSUMPTION ORIENTED

Multi-Site / Multi-usage AI Factory

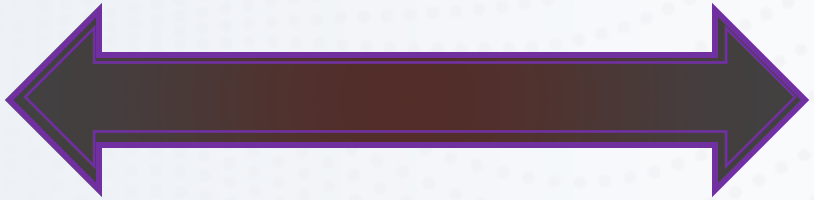
kubernetes openstack



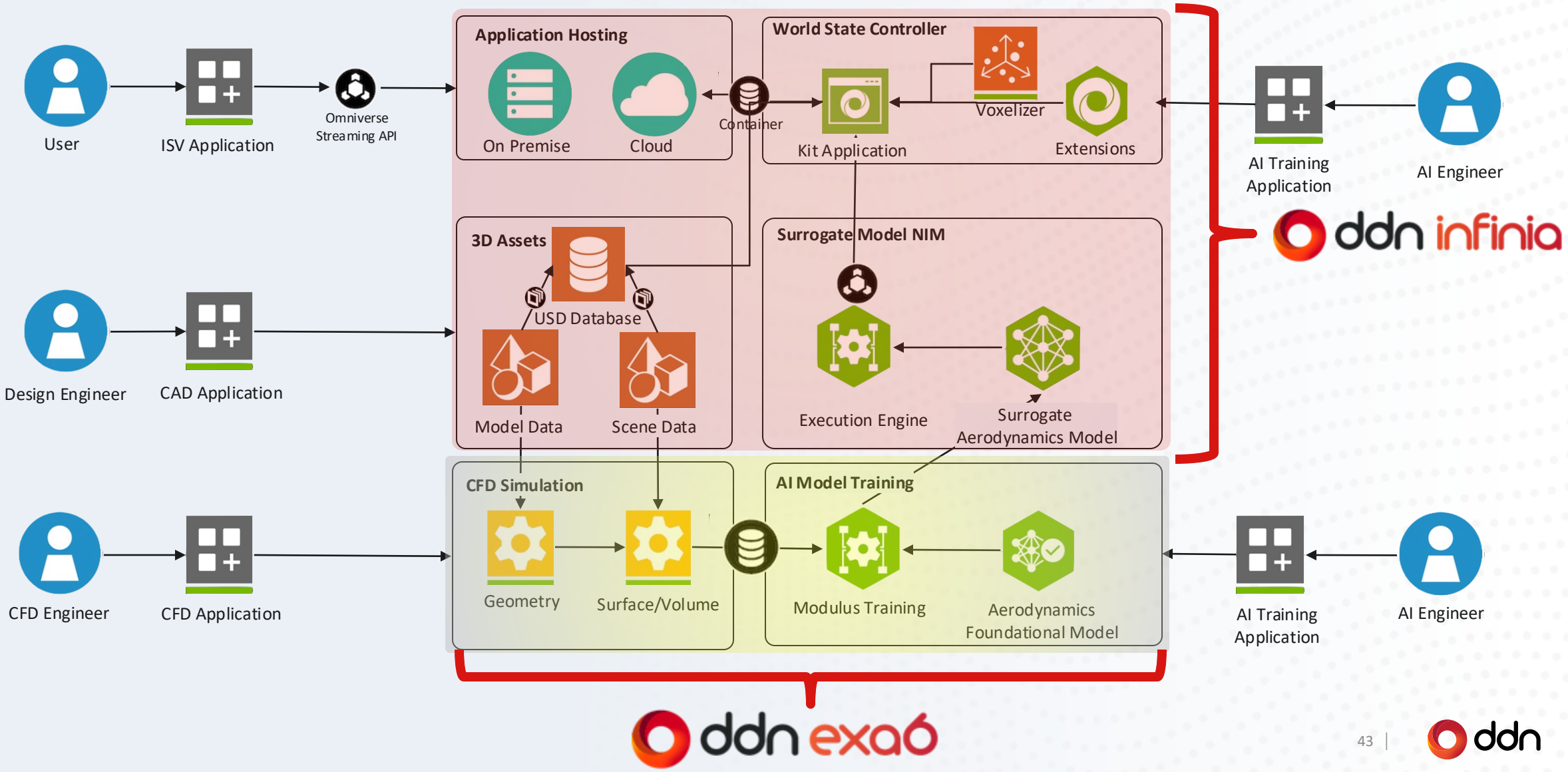
FILE / CSI

MULTI-TENANT

EXTREME PERFORMANCE

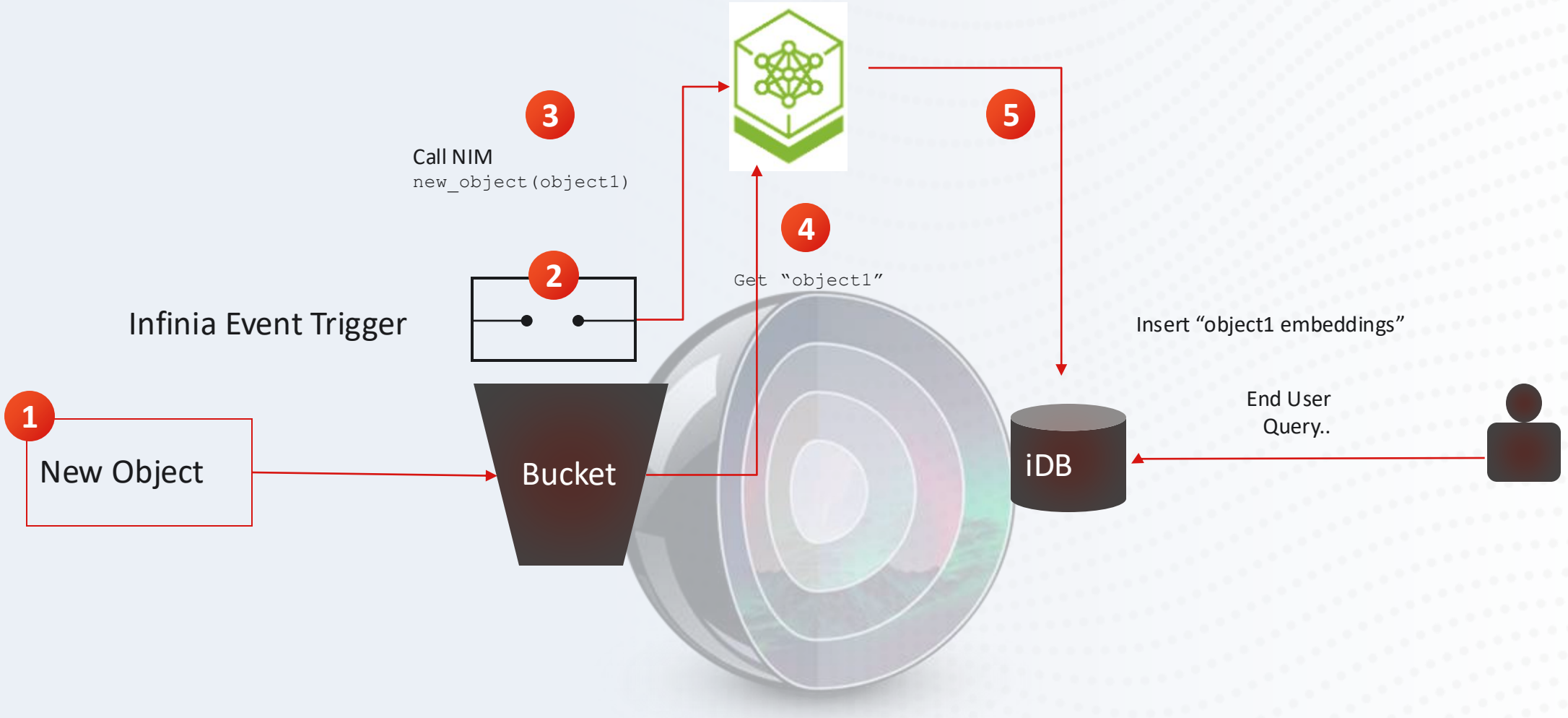


Industrial use-case: Manufacturing end-to-end CFD



INDUSTRIAL USE CASE

Automated Data ingestion and inference



AI Factories: NVIDIA + DDN: End-to-End AI Solution

TRAINING

- ✓ Maximum GPU Productivity
- ✓ NVIDIA DGX Platforms

DIGITAL TWIN

- ✓ Omniverse with Cosmos
- ✓ Accelerate NIMs and Nemo

EDGE

- ✓ Edge-Cloud-Core
- ✓ Collection, Curation, Transfer

THANKS!