



UK Research
and Innovation

DDN

Datacentre Efficiency

Gareth Wilson

DiRAC

High Performance
Computing Facility

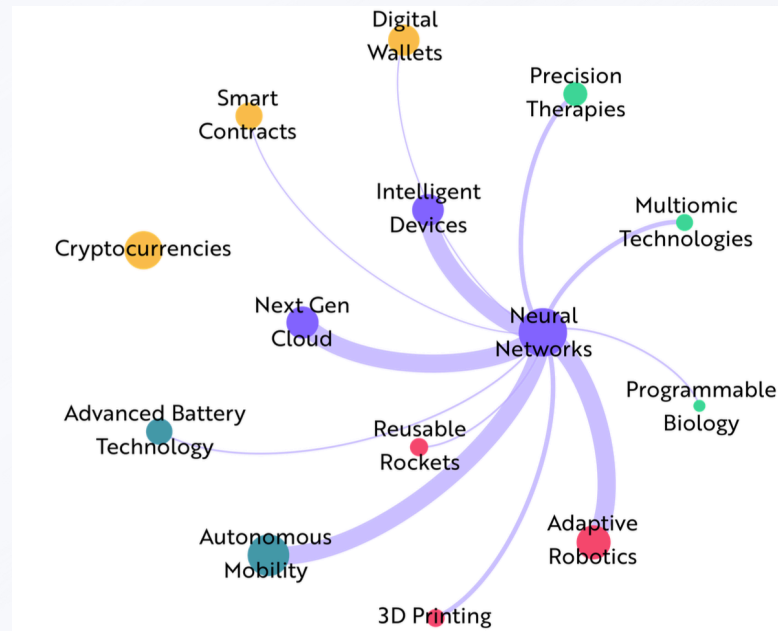


ddn

The AI Data Company

Our Era is Being Reshaped by Accelerated Computing

- Revolutionizes chatbots
- Drives value in financial services
- Powers autonomous driving
- Enhances personalized health
- Creates dynamic experiences in gaming





“

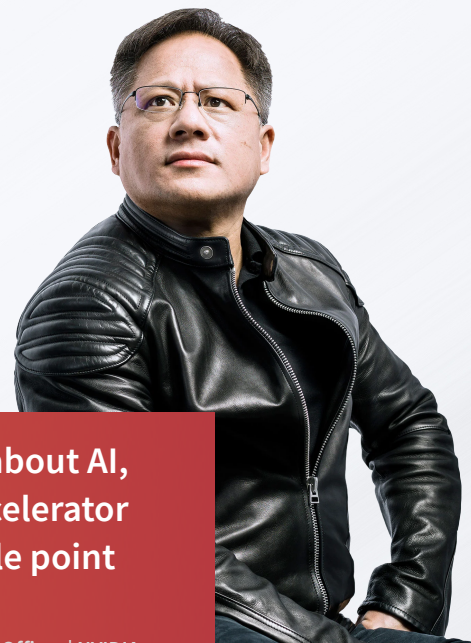
**I have a very simple statement
for you; NVIDIA uses DDN.**

~ Manuvir Das | Head of Enterprise Computing | NVIDIA

”

Accelerated Computing Requires a Full Stack

- Optimization of AI Models, GPUs, Network, Data Centre
- Data Management, Data Movement and Storage at Scale
- Power Efficiency, Footprint, Workflow Acceleration



So, nearly everybody who thinks about AI, they think about that chip, the accelerator chip and in fact, it misses the whole point nearly completely

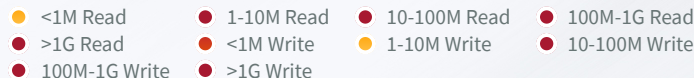
- Jensen Huang
Chief Executive Officer | NVIDIA

Machine Learning is Write and Read Intensive

- 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...**”

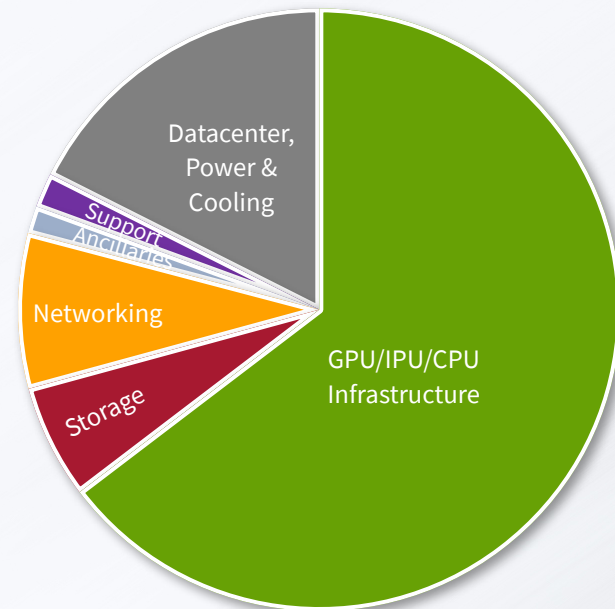
Average Number of Calls per Job



Spend is Large, Risks are High. How can Storage Help?

- A Storage System typically represents 5% of the 3 year Capex & Opex budget of an AI system for Deep Learning/LLM training
- **IO Wait and associated elements of the training process can consume up to 43%¹ of runtime**
- How can the efficiency architecture and consumption of storage resources impact overall productive output of this System?

Large AI Supercomputer Capex + Opex



DDN Increases AI Data Center Productivity Making the Storage Free

By reducing storage wait time for Load, hugely reducing wait time for checkpoints, and enabling more checkpointing to eliminate lost time of recomputation, DDN reduces data center runtime by 5-12%, delivering higher productivity to LLM's and Generative AI.

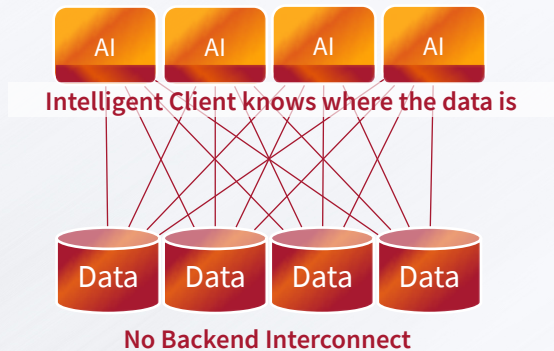
Spend \$1 on DDN, gain \$2 in Infrastructure productivity

Significantly Higher Efficiency and Faster AI Training with DDN

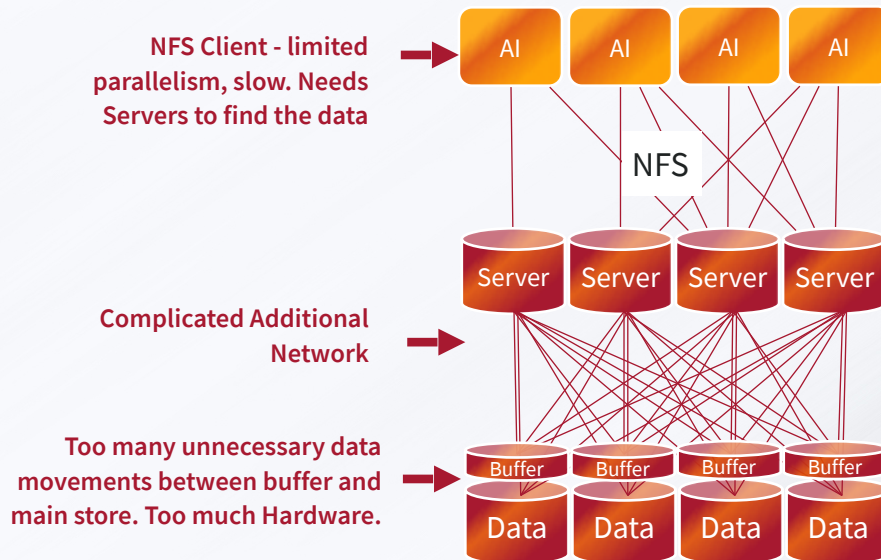


Parallel Filesystems Enable Scale Efficiency

An Efficient Architecture



You can never get efficiency with this architecture.



Today's QLC Scale Out NAS Are Highly Inefficient

Low IOPS and Throughput per Server

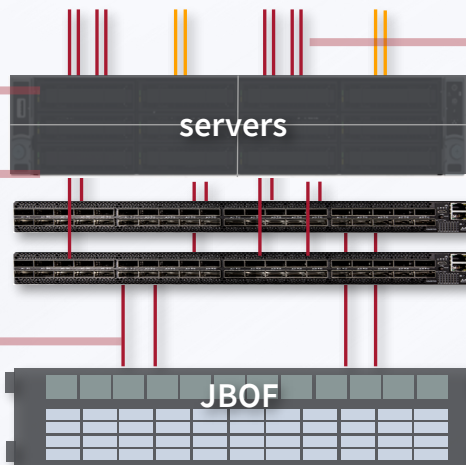
10x lower IOPS than DDN in 3x the footprint

Very Low Write Performance

12x less write performance compared to DDN

10x More Cabling

Expensive and complicated cabling at the back end. More to Fail and more to manage



Low Per-Port Performance

Customer needs 10x more switch ports for performance

Backend Switching

Adds complexity and Cabling. Makes it difficult to expand

Everything is Tiered

- Poor Write Performance
- Lowers Efficiency with many data movements per write
- High Latencies
- Poor Metadata Performance

In 2RU and 2KW **DDN Delivers:**

90GB/s
READS

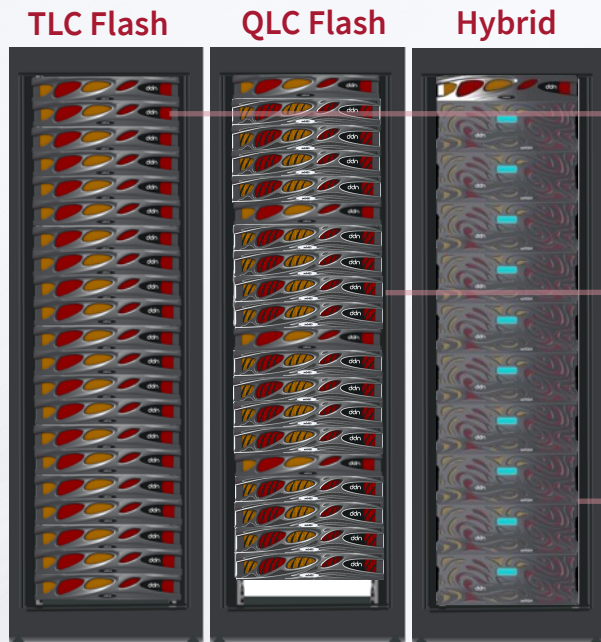
3M
IOPS

65GB/s
WRITES

Up to **720TB**
NVMe TLC CAPACITY



World's Most Performing and Flexible Storage At Any Scale



Best IOPS & Throughput per rack

Up to 70M IOPs in a single rack
 1.8 TB/s of read throughput
 1.4 TB/s of write throughput
 16PB of Flash

Best Price per Flash TB

Up to 14M IOPs in a single rack
 360 GB/s of read throughput
 240 GB/s of 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

DDN Delivers Highest Performance AND Highest Endurance in QLC

The Fastest QLC Appliance

Read: 90GB/s Write: 70GB/s, IOPs: 3M

Fully Flexible and Scalable Expansion

Add 2, 4 or 5 SE 2420 QLC Expansion Enclosures

Highest Density

Between 1 and 6 PB QLC Capacity

Simplest Redundant Infrastructure

DDN embedded NVMeoF network
- No External Switches



QLC Optimized for Performance or Capacity

Optimized for Performance/\$

**630GB/s**

READ Performance

490GB/s

WRITE Performance

14PB

Capacity

Optimized for Capacity/\$

**270GB/s**

READ Performance

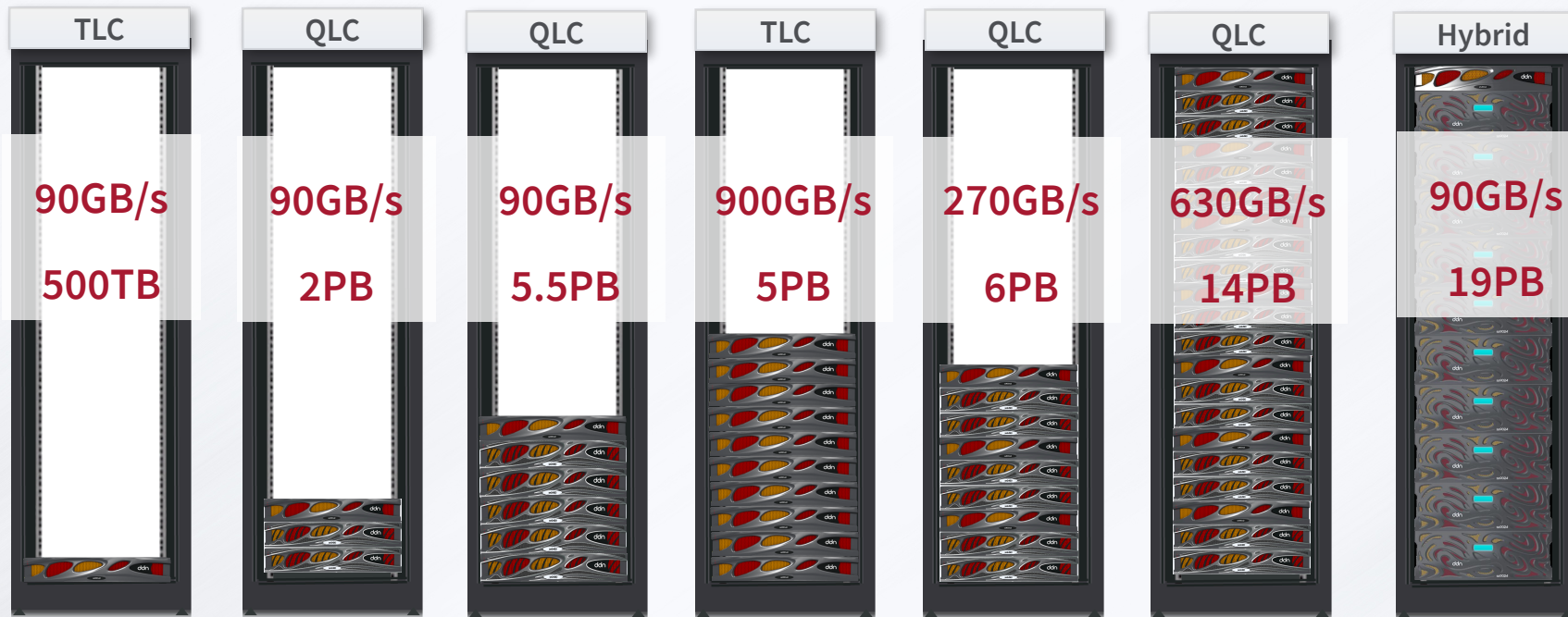
210GB/s

WRITE Performance

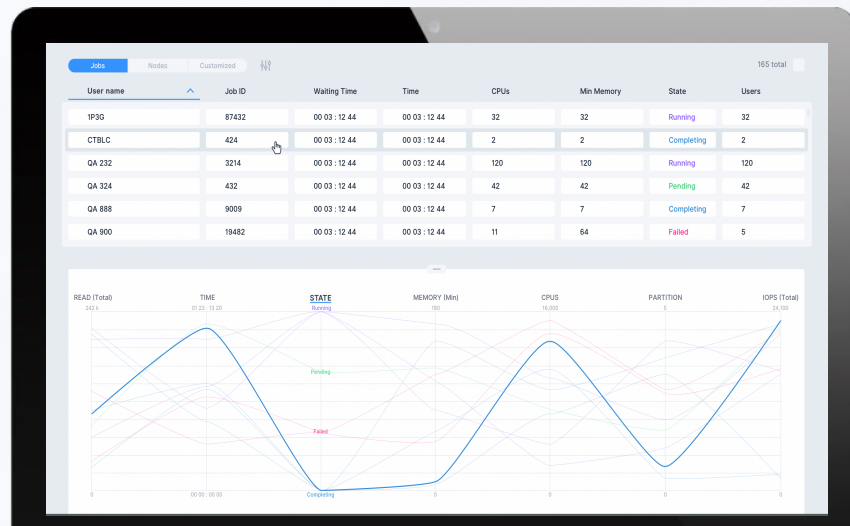
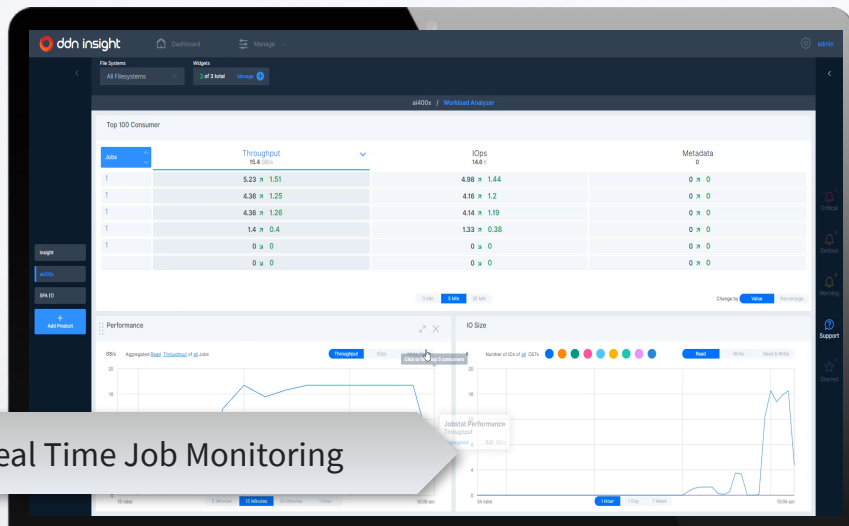
24PB

Capacity

Build to Specification: Price - Performance - Capacity



With DDN Analyze Live Workloads and Full History



Historical Job Analysis¹

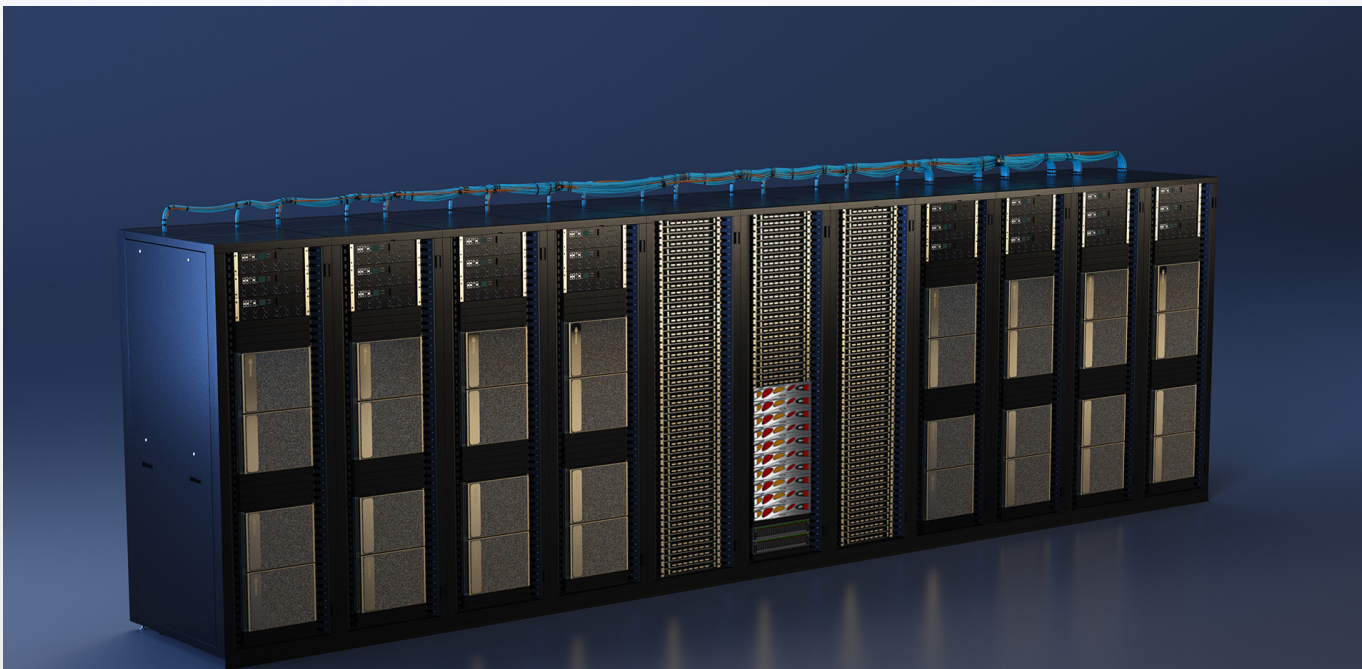
Prometheus Exporters Enable Open Integration



Real Time Job Monitoring



DDN Makes the Storage Which Drives Advances in AI, HPC, Life Sciences, Finance, Autonomous Cars at Any Scale



10X Faster. 10X More Capacity. 10X Less Power and Footprint



ddn