How Would it Impact Your Business if You Could Seamlessly Extract a Lot More Answers per Hour From Your Data at Any Scale?

Ran Pergamin, Storage Solution Architect
rpergamin@ddn.com - 052-722-0080
DDN | Scalable Storage Solutions for Your Specific Applications

- **Performance** Storage Platforms designed for scale
- All Aspects of **Scale**: IOPs, Throughput, Latency, Capacity, Distribution
- **Flexible** Architecture Design
- Global Services and Support from **Storage Specialists**

DDN Storage | ©2018 DataDirect Networks, Inc.
DDN Complete AI Data Lifecycle Solutions

Your AI Workflow Concurrently, Continuously, In-Place

DDN Parallel AI solutions are all-round performers able to saturate GPUs for training at the same time as inference and data movements and preparation.
Match your business growth with flexible all flash or hybrid scale-out

With data volumes continuing to rise, AI is getting smarter and learning faster, enabling organizations such as yours to leverage more-advanced Machine Learning (ML) and Deep Learning (DL) solutions.
ACCELERATED ANY-SCALE AI SOLUTIONS FROM DDN

Optimized and integrated data solutions for your AI frameworks

DDN brings you to production faster by creating a fully optimized and integrated data platform for AI frameworks, managing data efficiently whatever your choice of AI framework.
BUILD YOUR AI DATACENTER WITH DDN’S DATA PLATFORMS FOR AI

AI200 - 20GB/s
2RU, 24 NVME DRIVES

Capacities: 20TB to 368TB in 2-RU
Start with 1 system, Scale out to 100+
InfiniBand - EDR or Ethernet 100Gb

©2018 DataDirect Networks, Inc.
BUILD YOUR AI DATACENTER WITH DDN’S DATA PLATFORMS FOR AI

**AI7900 - 20GB/s**

4RU

90 DRIVES IN BASE + 360 MORE

6.3PB RAW IN 20-Us

---

Hybrid Configurations: SSDs and NL-SAS

Start with 1 system, Scale to 100+

InfiniBand - EDR or Ethernet 100Gb
DDN A³I Solutions: Turnkey, integrated and optimized for NVIDIA DGX-1

Simple, Fully-Packaged Solution for DGX

Start with a single AI200 storage appliance and a single DGX-1 server

Pre-configured solutions that are easy to deploy and manage, fully-integrated and optimized for maximum DGX server productivity

support: EDR IB and 100 GbE networking
**DDN A³I Solutions: Massive scale with multiple NVIDIA DGX-1**

**Scalable, End-to-End Solution for DGX**

Flexible architecture scales seamlessly in capacity, performance and capacity to match evolving workflow requirements.

Additional DGX-1 servers and AI200 storage appliances can be integrated rapidly, and immediately provisioned to applications.

Architecture performance scales linearly and delivers optimized, extremely cost-effective solution with multi-DGX deployments.
**DDN A³I Solutions**: Fully integrated and optimized for NVIDIA DGX-1

**Full NVMe Performance, In-Container**

Extreme amounts of performance and capacity through a tight integration between DDN and NVIDIA platforms

Every layer of hardware and software engaged is optimized for fast, responsive and reliable file access

Highly optimized parallel data paths run from containerized applications, all the way to disk drives
**DDN A³I Solutions: Performance to a single container**

Over 100K IOPs and 10GB/s to 1 Container

EXAScaler client demonstrates over 10GB per second to a single container and over 100K IOPs.

Typical Deep Learning Codes will perform IOs around 128K IOs – Scale up to 38GB/s with 4 Containers in a single DGX-1 System.
60% Faster - Consistent, Linear AI Performance

AI benchmarking of AI Frameworks with popular test sets demonstrates images per second rates 60% higher than competing systems.
DDN PARALLEL ARCHITECTURE AND PROTOCOL vs NFS

FULL GPU SATURATION
MAXIMUM AI PRODUCTIVITY

DDN AI200

GPU STARVATION
CRIPPLED AI PERFORMANCE

NFS
DDN PARALLEL ARCHITECTURE AND PROTOCOL vs NFS

DDN FULL GPU SATURATION

- Consistent DL performance
- Achieves full GPU utilization
- Efficient loading of data set

2X FASTER

NFS GPU STARVATION
DDN PARALLEL ARCHITECTURE AND PROTOCOL vs NFS

**CAFFE GOOGLENET TRAINING THROUGHPUT**

7.5X MORE IMAGES WITH DDN STORAGE

--

**CAFFE GOOGLENET TRAINING TIME**

5.5X FASTER WITH DDN STORAGE
AI Benchmarking: Optimized TensorFlow Data Sets

Optimized Data Sets, Accelerated Training

Convert discrete data and metadata asset collections into series of streamlined binary files for TensorFlow

Maximize DL training and validation efficiency while minimizing data management and preparation overhead

Achieve more accurate results by eliminating manual data shuffle bias
DDN – Large Autonomous car company deployment
Maximize efficiency
Profit from greatest technical and economic benefits

Ensure data availability
Trust in robust data protection, integrity and access

Accelerate AI performance
Achieve full GPU saturation for greatest AI productivity

Streamline DL workflows
Engage all workloads concurrently and continuously

Scale seamlessly
Increase performance, capacity and capability at any scale

Deploy and manage easily
Fully-integrated and optimized solutions from DDN
THANK YOU

CONTACTS:
NIR TALMOR, REGIONAL SALES DIRECTOR, NTALMOR@DDN.COM – 052-722-0070
RAN PEGAMIN, SOLUTION ARCHITECT, RPEGAMIN@DDN.COM - – 052-722-0080