DDN LEVERAGES YEARS OF HPC KNOWLEDGE TO FUSE INGREDIENTS OF AI THANKS TO A3I INTO TRADITIONAL HPC WORKLOADS
Extract a Lot More Answers per Hour From Your Data at Any Scale
### DDN LEADS ARTIFICIAL INTELLIGENCE AND DEEP LEARNING AT SCALE

<table>
<thead>
<tr>
<th>Autonomous Vehicles</th>
<th>Life Sciences Research</th>
<th>Precision Healthcare</th>
<th>Smart Manufacturing Robotics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial Services</td>
<td>Security &amp; Intelligence</td>
<td>Predictive Analytics</td>
<td>Supply Chain Logistics</td>
</tr>
<tr>
<td>Fraud Detection</td>
<td>Personalized Marketing</td>
<td>Consumer Retail</td>
<td>Language Processing</td>
</tr>
<tr>
<td>Augmented Reality</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

©2018 DataDirect Networks, Inc.
DDN A³I SOLUTIONS FOR END-TO-END AI ENABLEMENT

Accelerate AI performance
Achieve full GPU saturation for greatest AI productivity

Maximize efficiency
Profit from greatest technical and economic benefits

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

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

©2018 DataDirect Networks, Inc.
DDN COMPLETE AI DATA LIFECYCLE SOLUTIONS

PICK YOUR COMPUTE NODES, 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.
DDN ENABLES AND ACCELERATES DEEP LEARNING

Data is the new source code

► Collect and access large amounts of data from a variety of sources
► Easy data discovery
► Rapid iteration of model training and refreshing
► Quick transition of new and refreshed models to production
► At scale, concurrently, continuously
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
DDN PARALLEL FILE STORAGE APPLIANCES FOR AI AND DL

**AI200**

All NVME parallel file storage appliance
Optimized for the most intensive workloads
30TB, 60TB, 120TB in 2RU

**AI7990**

Hybrid parallel file storage appliance
Optimized for capacity, intermix with flash
1PB in 4RU

20GB/s, 350K IOPS • 100Gb Ethernet or EDR InfiniBand • Start with 1, Scale out and mix
Superior performance, scalability and reliability

GPU saturation is critical for achieving the full potential of AI and DL.

The DDN parallel architecture and protocol delivers high-throughput, low-latency, massive concurrency to applications.

For DL frameworks, provides increased performance and faster processing than NFS.
DDN AI SOLUTIONS FOR AUTONOMOUS VEHICLES

DDN enables the largest autonomous vehicle programs in the world

FLEET OPERATIONS
- Vehicle data generation
  - internal and external sensors
- Fleet data collection
  - and processing

TRAINING
- Data curation and indexing
- Model development, selection
- Discovery

SIMULATION
- Safety
- Resimulation
- Regression testing
Traditional HPC simulation leveraging AI/ML approached

FUSION
Ground sensor data collection
Segmenting, pattern labelling

TRAINING
Data curation and indexing
Model development, selection
Discovery

SIMULATION
HPC processing
Interpolation
Regression testing
DDN AI SOLUTIONS FOR LIFE SCIENCES AND HEALTHCARE

DDN accelerates life sciences and healthcare hundred fold with GPU-enabled genome and image analytics.

<table>
<thead>
<tr>
<th>CORE RESEARCH</th>
<th>DRUG DISCOVERY</th>
<th>PATIENT CARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomics and microscopy</td>
<td>Pharmacology and pharmacotherapy</td>
<td>Diagnostics and precision medicine</td>
</tr>
<tr>
<td>Biostatistics and bioinformatics</td>
<td>Clinical trials</td>
<td>Predictive health</td>
</tr>
<tr>
<td></td>
<td>Post-market drug safety</td>
<td>Insurance</td>
</tr>
</tbody>
</table>

DDN Storage | ©2018 DataDirect Networks, Inc.
DDN AI SOLUTIONS FOR SMART MANUFACTURING

ROBOTICS

Collaborative teams
Self-training and repair
Agility and adaptability

QUALITY CONTROL

Industrial inspection
Better anomaly detection and fewer false positives
Failure analysis and reporting

PLANT EFFICIENCY

Raw materials ordering just in time delivery
Operations scheduling
Predictive maintenance
DDN AI SOLUTIONS FOR FINANCIAL SERVICES AND ANALYTICS

ANALYTICS
- Market prediction
- Contract intelligence
- Organizational performance

BANKING
- Investment and asset management
- Creditworthiness
- Customer engagement

SECURITY
- Risk management
- Fraud detection
- Regulatory compliance
DDN AI SOLUTIONS FOR CONSUMER EXPERIENCES

DDN enables natural language processing for the top tier virtual assistant.

RETAIL
Cashier-free checkout
Personalized services
Inventory and logistics

ADVERTISING
Personalized marketing
Marketing forecasting

ENTERTAINMENT
Experience innovation: AR, VR
Recommendation engines
Content protection
HPC simulation accelerated

1. **Macro simulation**
   Using traditional methods

2. **Interpolation**
   Inference spatial interpolation

3. **Iteration loop**
   Ground truth confrontation and minimum error heuristic

→ Harness the ever growing computation needs
→ Allow faster time to insight
→ Remains still highly demanding on data access
→ Smart staging and closely connected hot data storage required

Courtesy from Michael Houston SC18 conference
NVIDIA: Exascale Deep Learning for Climate Analytics
Reaching scale remains a challenge

<table>
<thead>
<tr>
<th>Dataset Size</th>
<th>Required BW (27K GPUs)</th>
<th>GPFS/LUSTRE</th>
<th>BurstBuffer</th>
<th>NVMe or DRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 TB (~63K samples)</td>
<td>3.8 TB/s</td>
<td>~400 GB/s</td>
<td>~2 TB/s</td>
<td>~26 TB/s</td>
</tr>
</tbody>
</table>

- 250 training samples/GPU (~15 GB), sample w/ replacement
- Each file will be read at most once from FS
- Files shared between nodes via MPI (mpi4py)

→ Using local cache is a smart approach yet requires tradeoff on the ability to update and synchronize parameters
→ Current R&D focus on a smart LPCC Lustre Persistent Client Cache - Enabling client side cache

Courtesy from Michael Houston SC18 conference
NVIDIA: Exascale Deep Learning for Climate Analytics
Reaching scale remains a challenge

→ DDN approach to remove limits to scaling greatly improve performances

→ Higher compute density is a real opportunity for DDN and its ability to feed data locally

→ Staging the data you need to process is key (pseudo random brings efficiency)

→ The relevance of IO benchmark is extremely relevant

THANK YOU