DDN STORAGE

Parallel File System for AI and Deep Learning

Managing exponential data growth with Lustre File System

Carlos Thomaz – Technical Product Manager

SCAsia 2018 Supercomputing


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

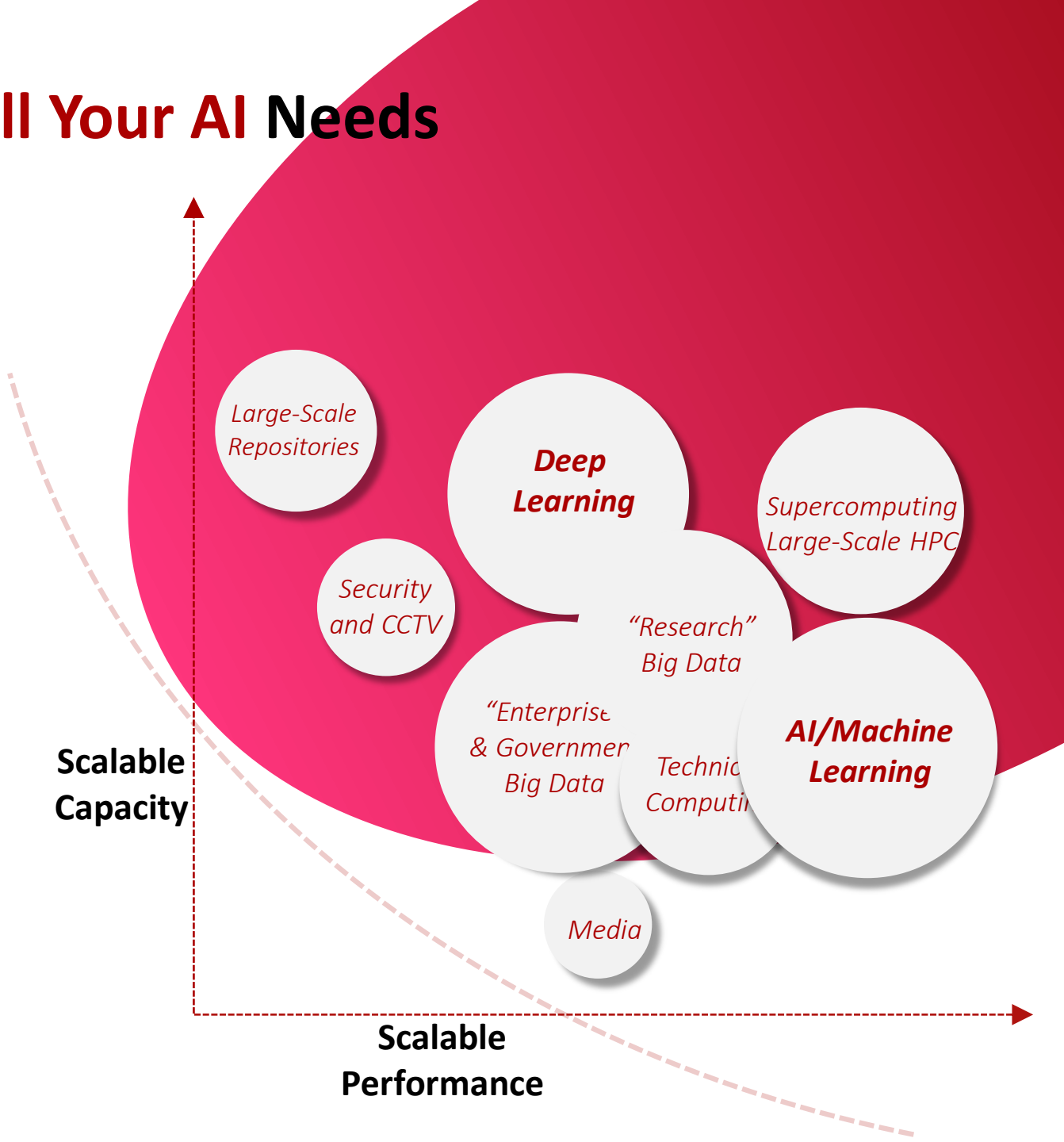
DDN | Storage Optimized for All Your AI Needs

We Deliver the Most Efficient Storage Solutions to Transform Your Data Science Into Real Business Benefits in AI and Deep Learning

Leading Storage, AI, Deep Learning Expertise

Highest Performing Analytics > 12X Faster

Scale IOPS, Bandwidth, Capacity



DDN: Leading Deployments Across All AI & Deep Learning

Driverless Cars	Data Security Fraud Detection	Augmented Reality	Healthcare Diagnostics	Personalized Marketing	Natural Language
Massive Scale Out. Real Time Ingest & Process	Fast Flash High IO/Client Security	Globally Distributed Optimized IO	Sensor Data Live Diagnostics Security	Distributed Analytics Resilient	Massive Ingest Flash/Disk Large Datasets

Data-at-scale Services

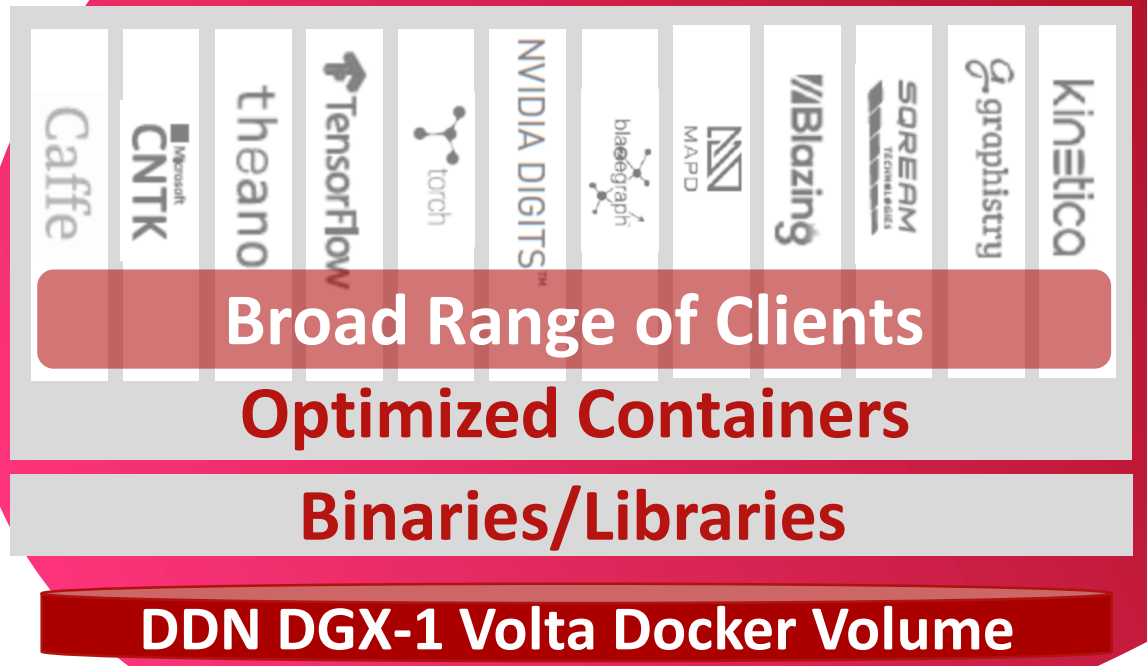
Delivered Real-time
Resource Optimized
Flash/GPU/CPU/Network

Scalable storage Platforms

Line Speed Performance
Tiered Storage At Scale
Extended Flash Life

DDN | Your Data Was Never as Close to Your GPU's

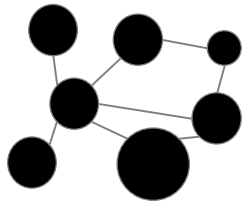
DDN Storage Accelerates Your Processors and GPU's and Optimized Containers So That You Achieve the Full Business Value of Your AI System



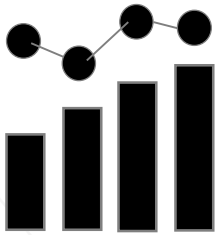
RDMA Accelerated I/O Delivered Directly to GPU's

Parallel File System Which is So Much Faster Than NFS!

Leverages a Wide Range of GPU Optimized Applications



Machine and Deep Learning



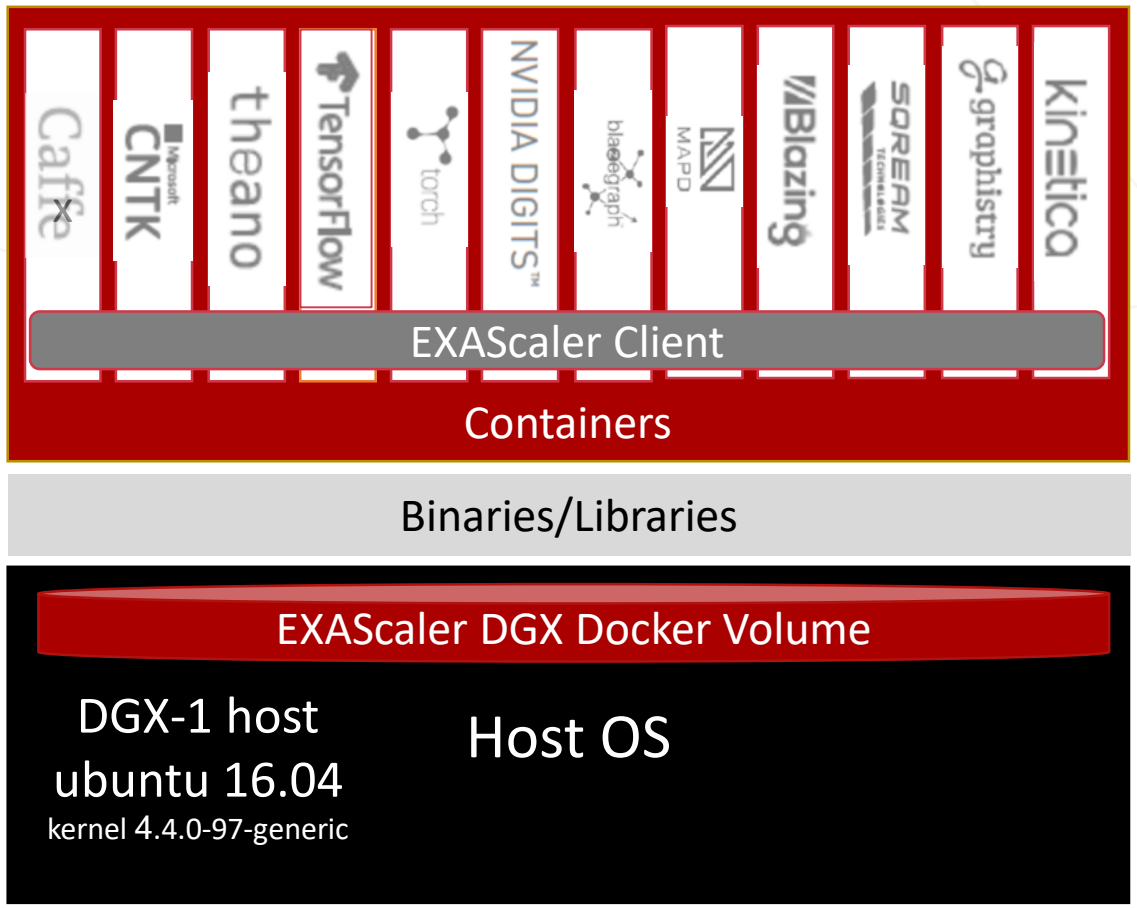
NoSQL Analytics

Big Data

IO Characteristics: Read, Random, High Throughput per Client, File and IO Sizes between a few kb and a few MB

Training Sets typically larger than local cache

EXASCALER FOR AI AND DEEP LEARNING ENVIRONMENTS (DGX-1 STACK)



- Integrated Flash Parallel File System Access via TCP or IB
- Extreme Data Access Rates for concurrent DGX Containers

EXAScaler DGX Docker Volume

- Lustre ES3.2 kernel modules compiled for Ubuntu kernel and host's OFED
- Lustre userspace tools
- scripts for Lustre mount/umount

WHY LUSTRE FILE SYSTEM – DDN EXASCALER

Feature	Importance for AI	Others	Lustre
Unique Metadata Operations	**Medium** - depends on Installation Size and Application Workflow	✓ Highly scalable	✓ Highly scalable with DNE I and II
Shared Metadata Operations	**High** - training data are usually curated into a single directory	✗ Lower than 10K	✓ Up to 200K
Support for high-performance mmap() I/O Calls	**High** - many AI applications use mmap() calls	✗ Extremely poor	✓ Strong
Container Support	**High** - most AI applications are containerized	✗ Poor (network complexity & root issues)	✓ Available
Data Isolation for Containers	**Medium/High** – important for shared environments	✗ Not available or very limited today	✓ Available
Data-on-Metadata (small file support)	**Medium/High** – depends on data set	✗ DOM limited to tiny files	✓ DOM is highly tunable

#open_source #no_hidden_fees #fully_supported

DDN EXASCALER FEATURE HIGHLIGHTS

Lustre Native Container Support

#subdirectory_mounts

#docker_ready_containers

#docker_optimizations

#nvidia_DGX

Lustre Persistent Client Cache

#local_lustre_cache

#lustre_aware

#designed_for_ai

#machine_learning_oriented

#containers #read_write

Lustre Enhanced security

#mls

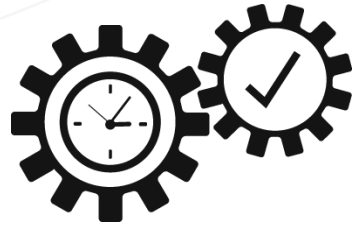
#changelog_audit

#isolation

#kerberos

#encryption_at_rest

Integrated Policy Engine



#fast_metadata_scan

#lightweight

#no_sql

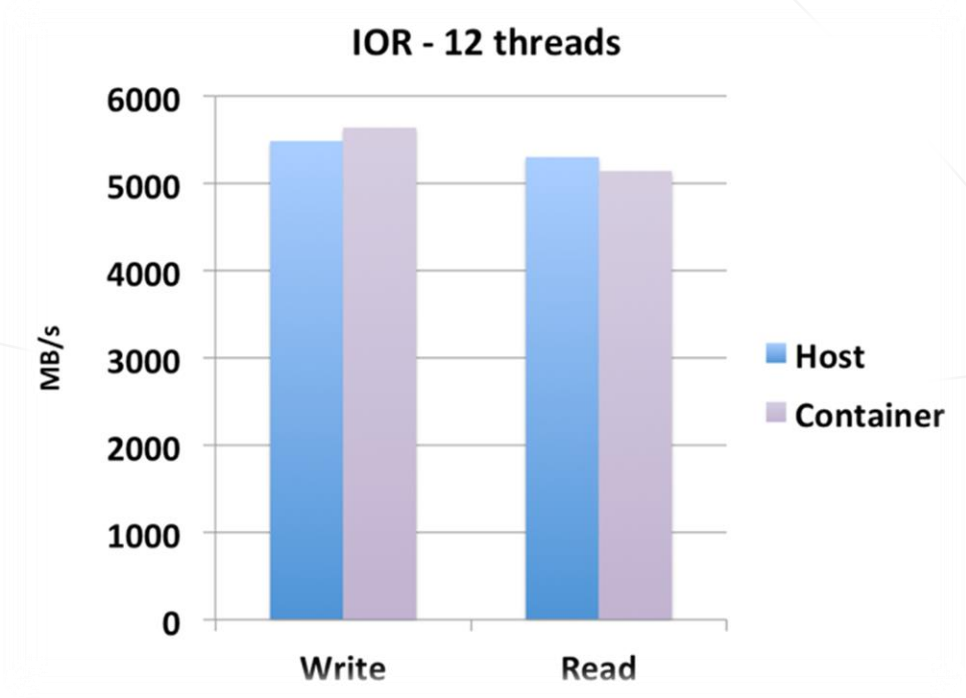
#scalable

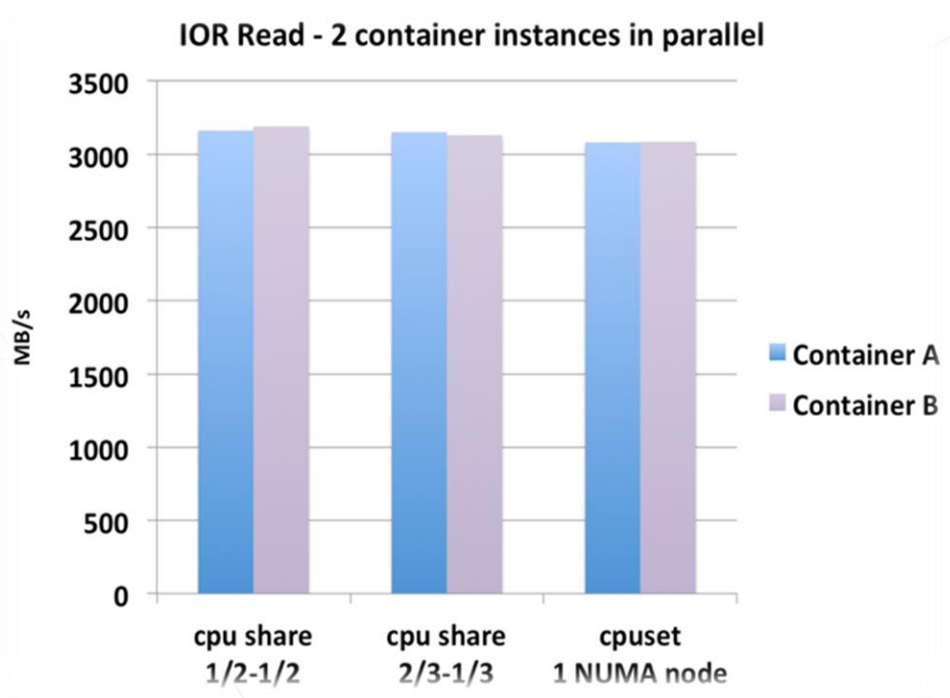
#mds_integrated

#rh_alternative

Features under development

DDN EXASCALER LUSTRE CONTAINER
MINIMAL OVERHEAD

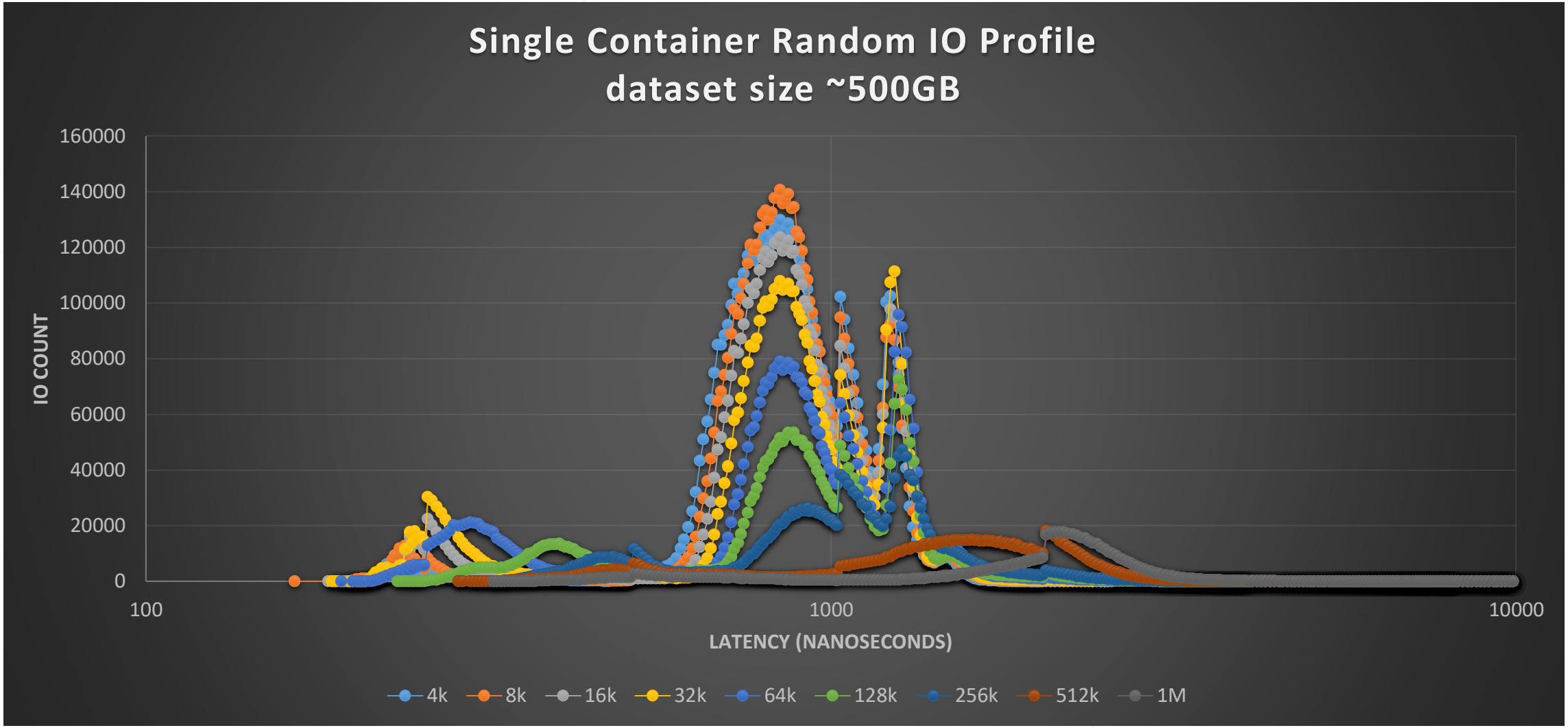




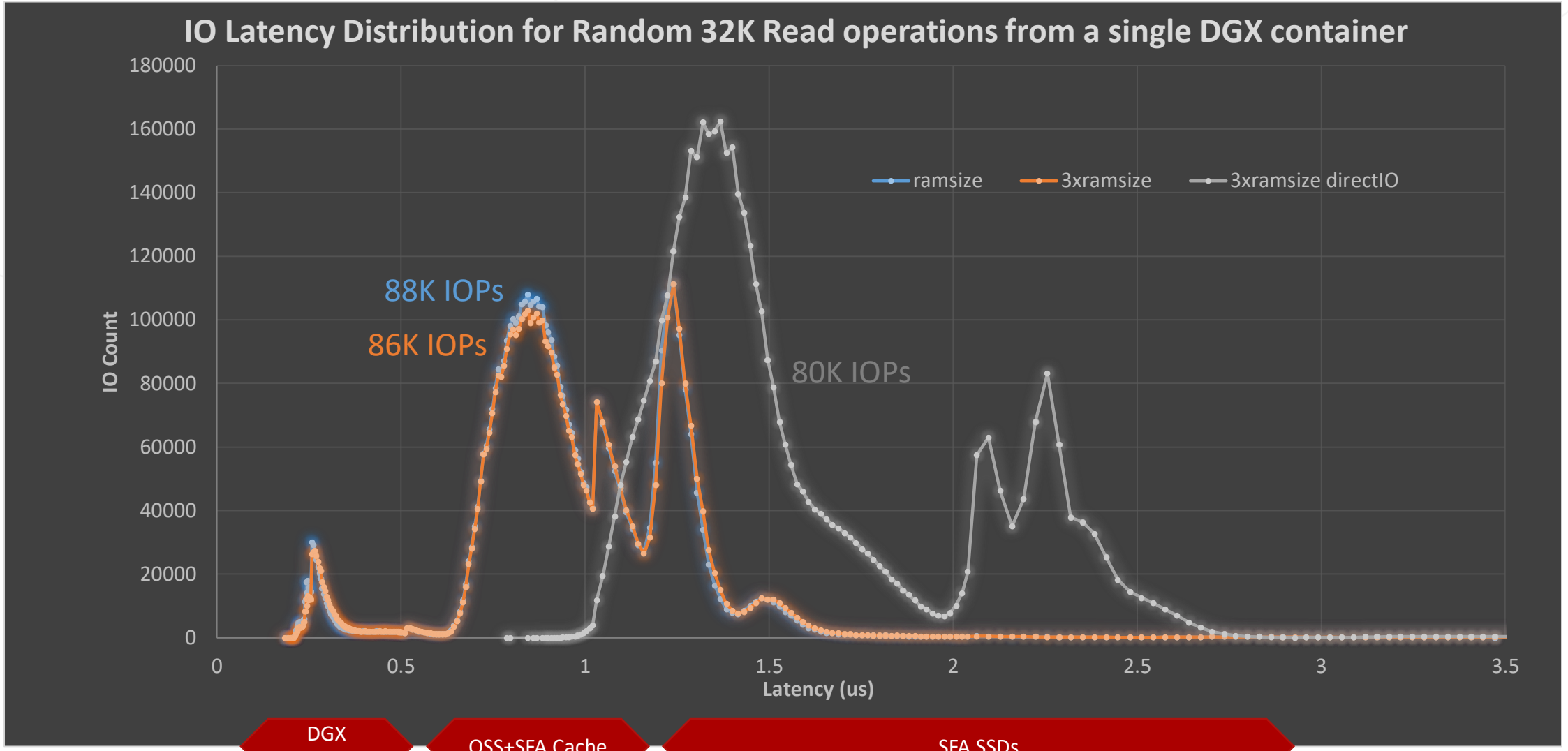
- SFA7700 delivering 7BG/s+
- Client node: 16 cores, 128 GB RAM, IB 4X FDR

Software: CentOS 7 (3.10 kernel), Lustre pre-IEEL3.0 (2.7 + patches), OFED 3.18-1, Docker 1.8.2

IO LATENCY PROFILE FOR RANDOM READ DGX
LUSTRE FULL FLASH

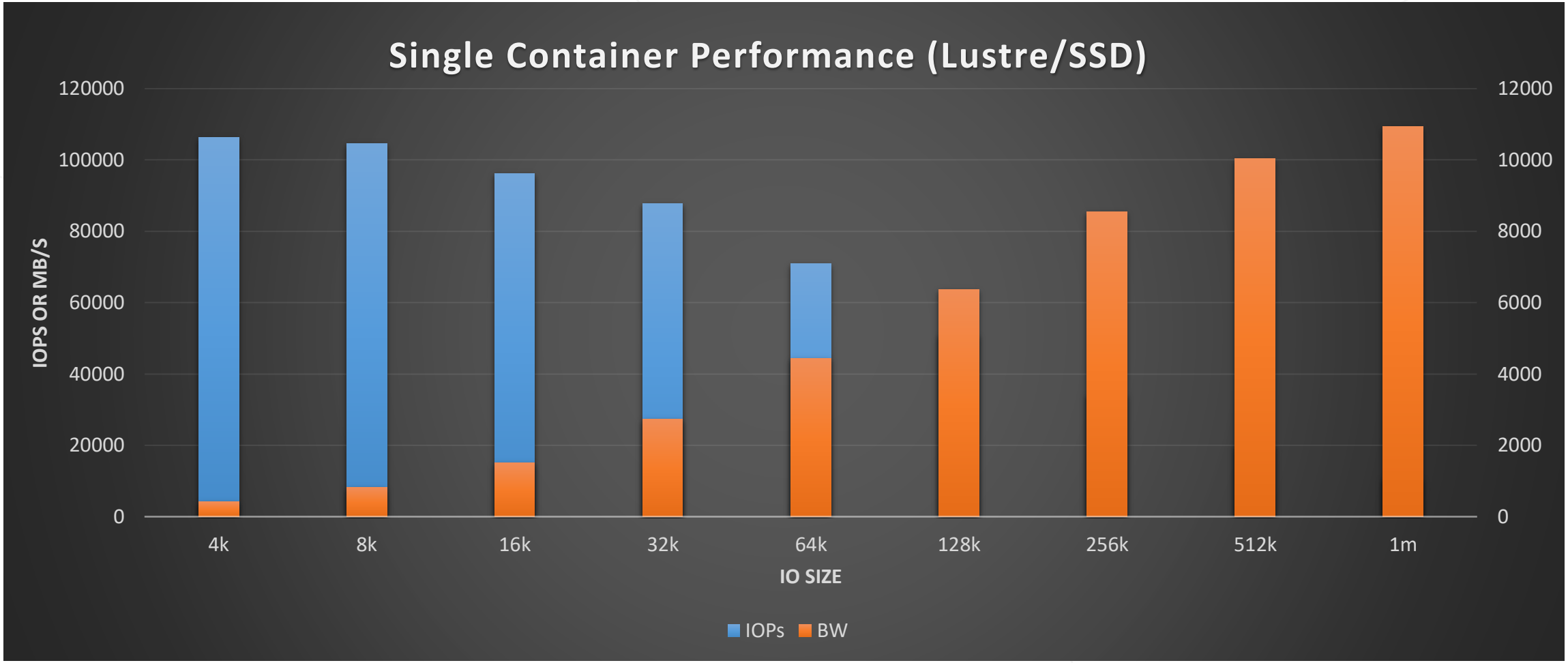


IO LATENCY PROFILE FOR RANDOM READ DGX
LUSTRE FULL FLASH

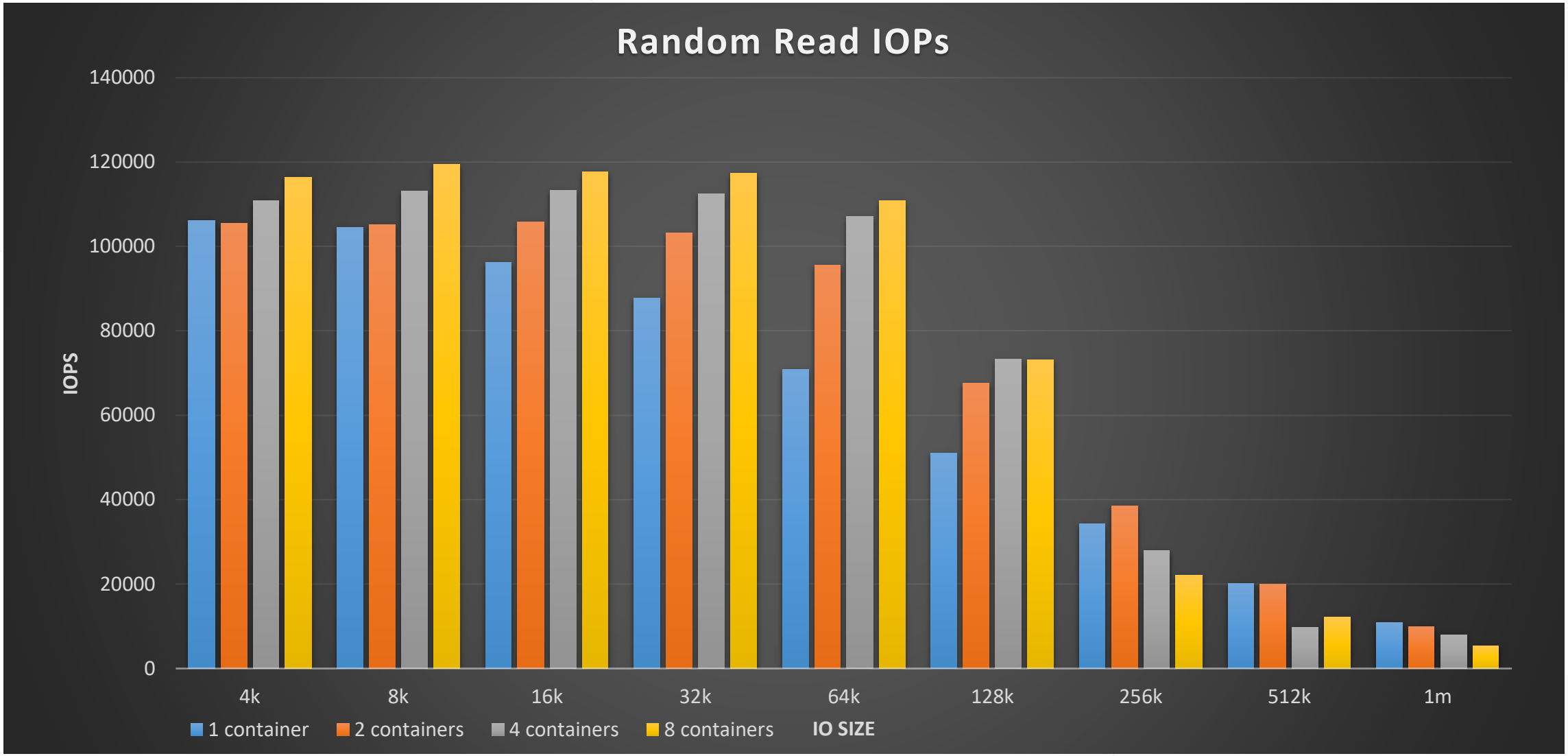


DGX DDN EXASCALER VOLUME PERFORMANCE

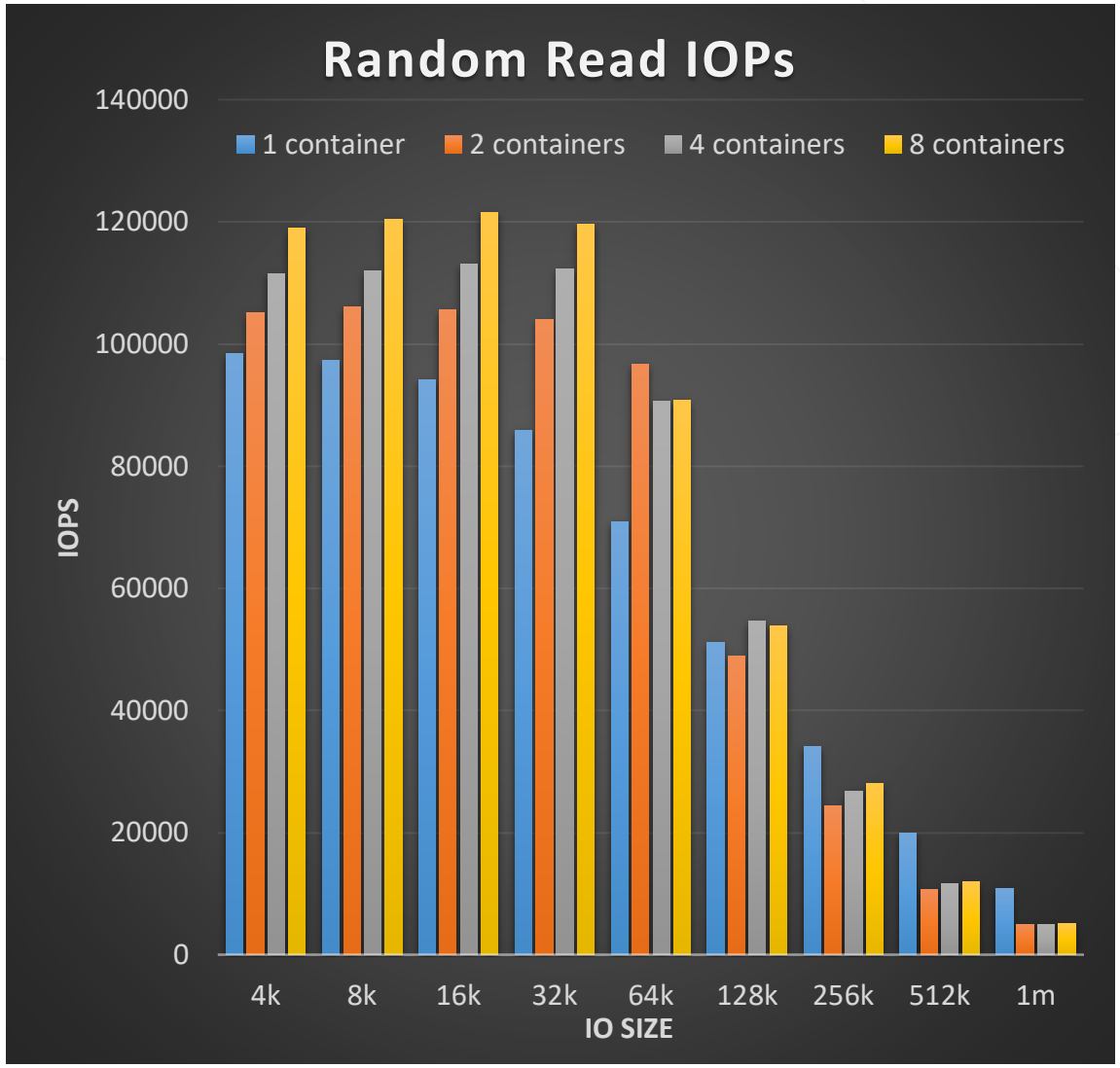
Over 100K IOPs and 11GB/s to a single container

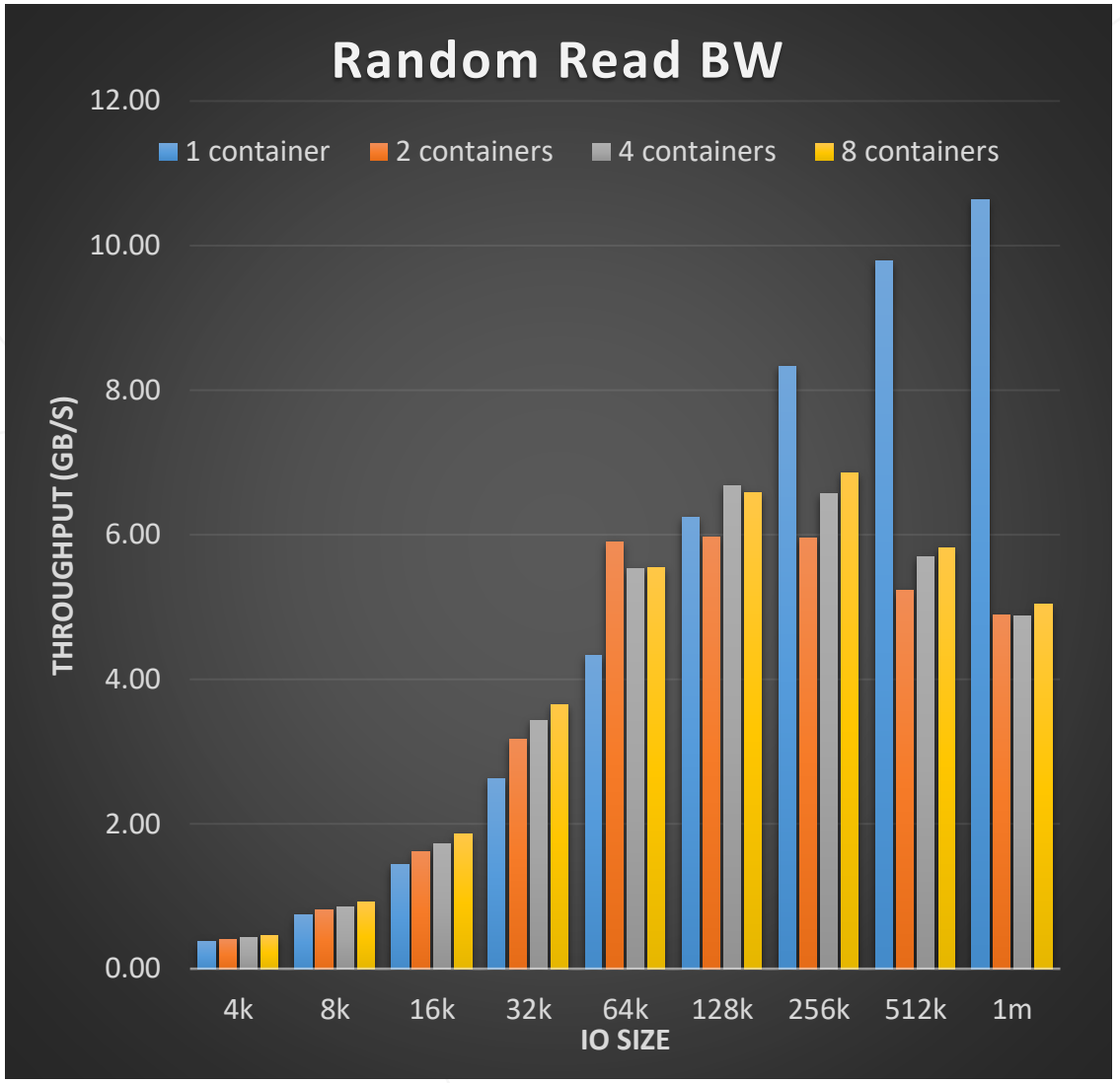


DGX DDN EXASCALER VOLUME PERFORMANCE



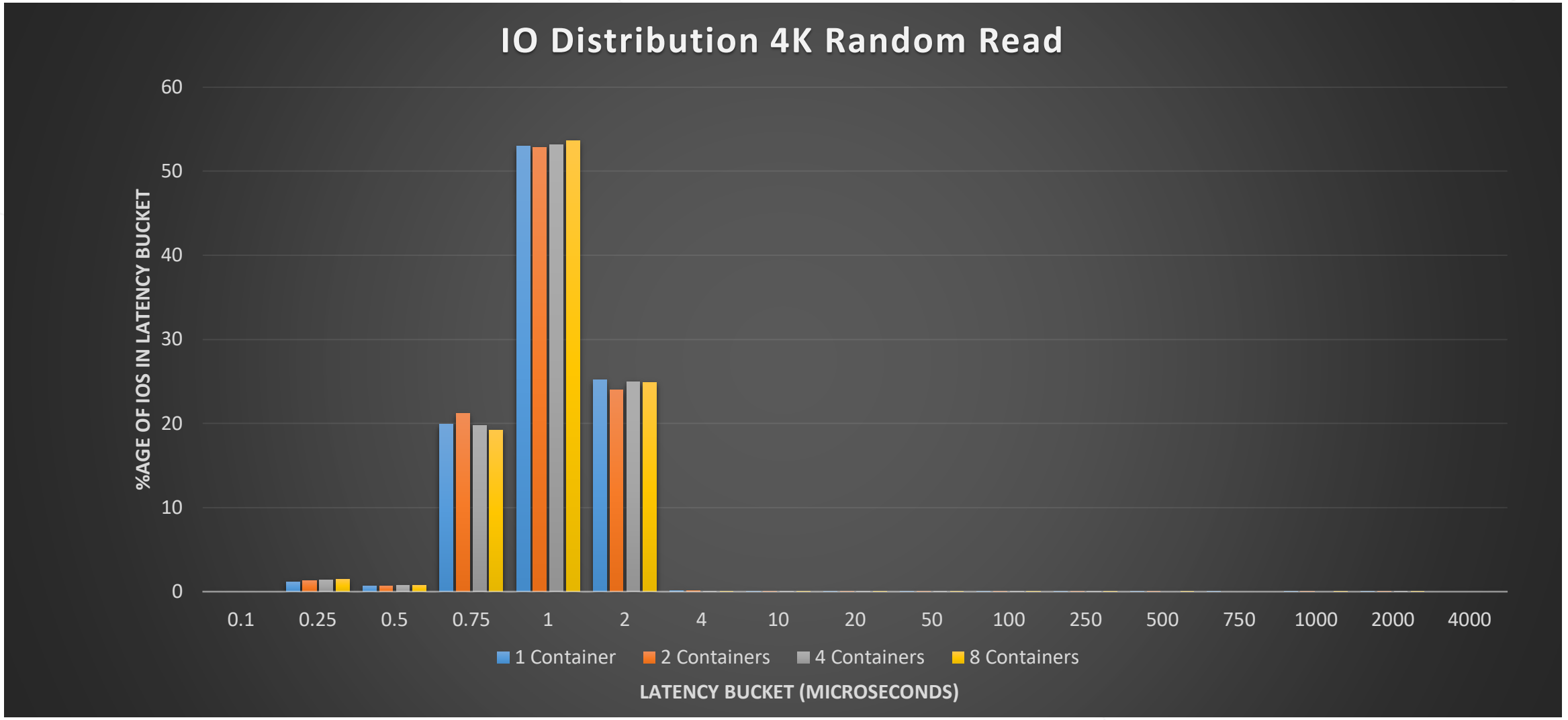
DATASET 3X DGX RAM CAPACITY RANDOM READ





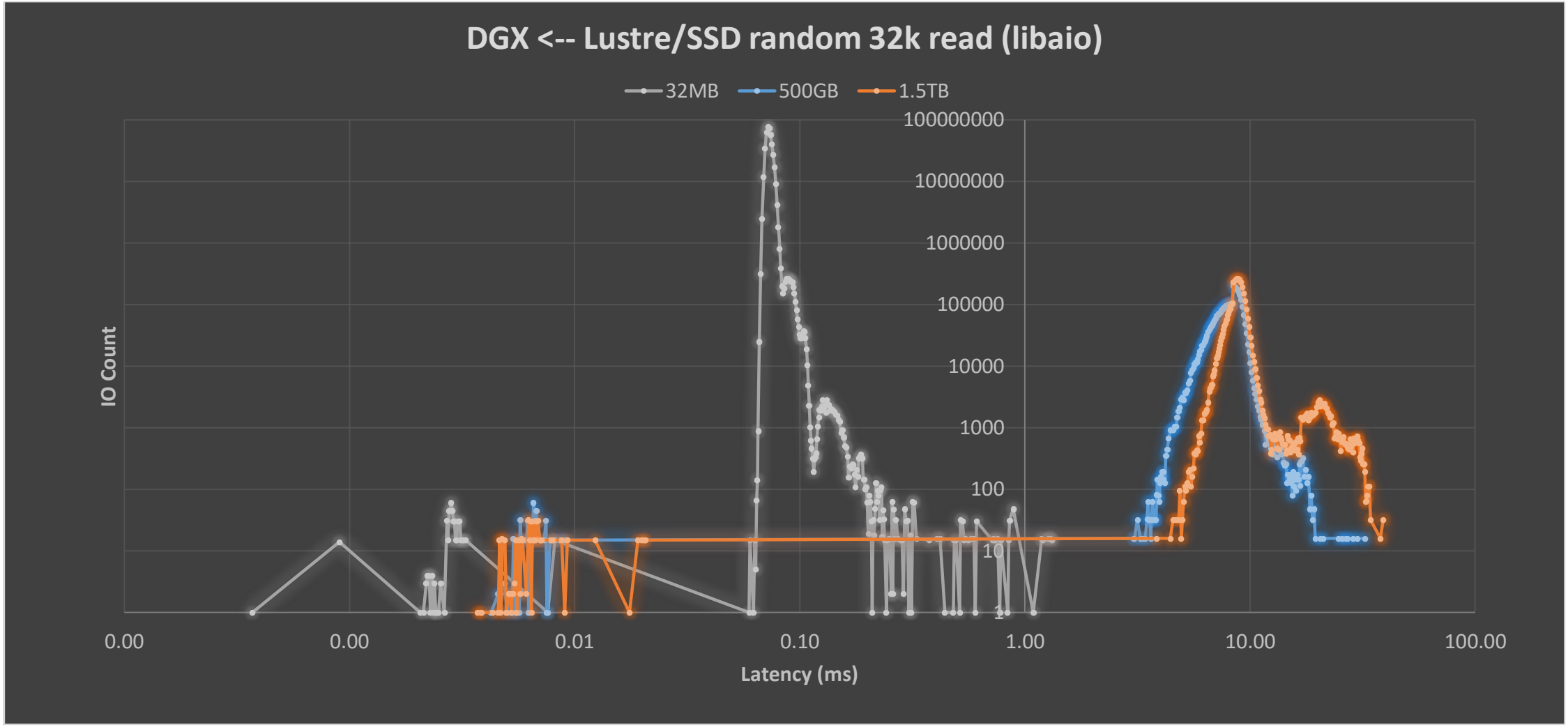
DATASET 3X DGX RAM CAPACITY RANDOM READ

Almost all reads in under 4 microseconds



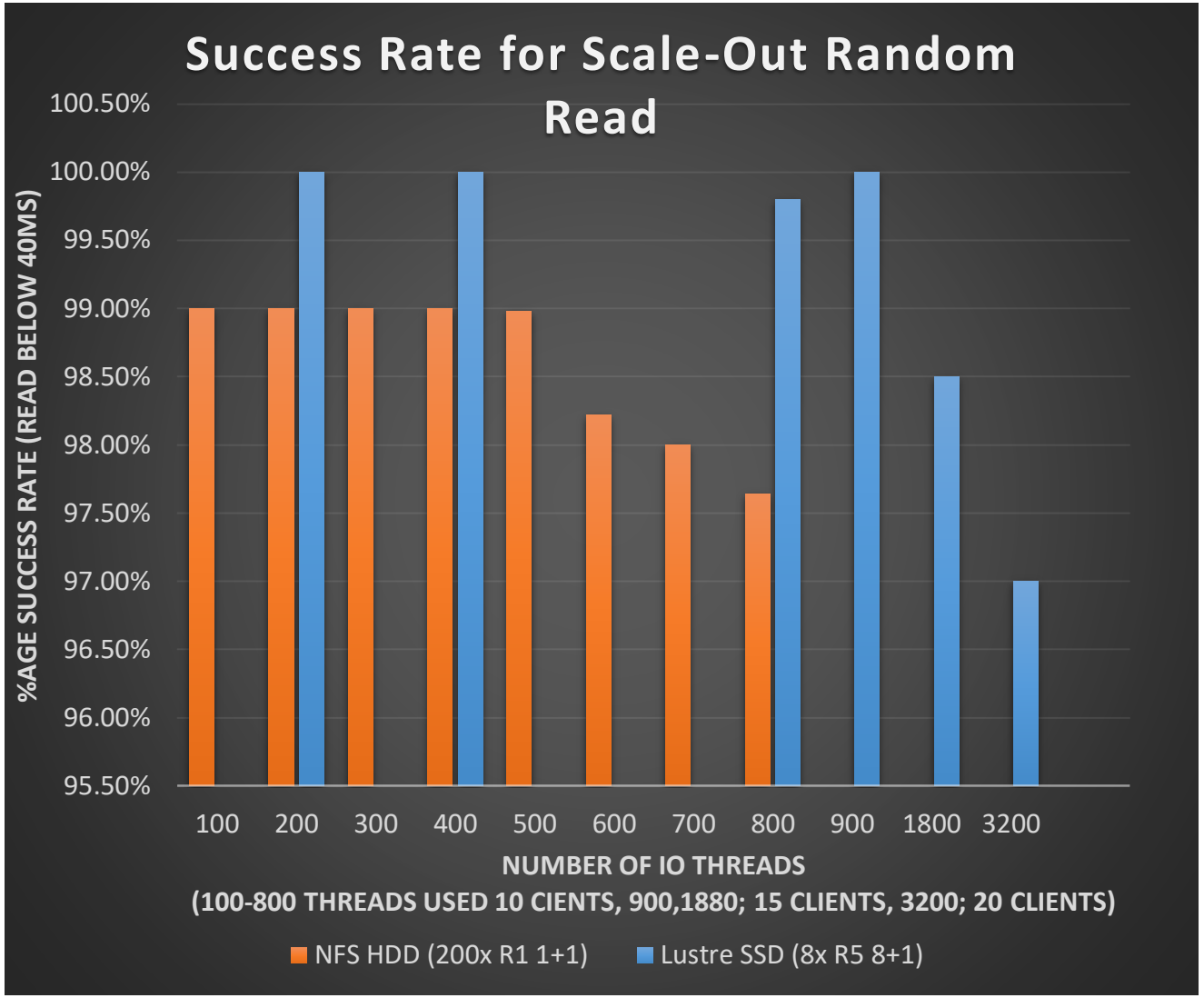
DGX Testing (libaio)

32k random read across different dataset sizes from a single DGX container



Autonomous Driving: Customer Benchmark

- Create 3M files, 100 per directory across 3000 directories
- Starting with 40 files per thread, kick off multiples of 10 threads across a number of clients
  - each thread randomly selects 40 files at a time to read sequentially
- Each thread iterates in batches of 40 files.
- A success rate is calculated corresponding to the percentage of files that are read within a set time window (40ms)
- Lustre DGX Docker Volume demonstrates 100% success rate with 900 IO threads



Thank You!

Keep in touch with us.

sales@ddn.com

@ddn_limitless

company/datadirect-networks

9351 Deering Avenue
Chatsworth, CA 91311

1.800.837.2298
1.818.700.4000