THE STORAGE LAYER IN THE NVIDIA DGX SUPERPOD

HPCXXL 2019 Summer Meeting

Spectrum Scale/GPFS User Group

9/23/19

NVIDIA

IBM Spectrum Storage

Jacci Cenci, Technical Marketing Engineer, NVIDIA Systems Engineering

“As the pioneer in both AI and autonomous control, NVIDIA has continued to push the innovation boundaries for both the car and development platforms. NVIDIA just took that innovation to a new level with a turnkey data center called the DGX SuperPOD that ranks as number 22 on the list of global supercomputers.”

– Jim McGregor, Forbes

“In a clear demonstration of why artificial intelligence leadership demands the best compute capabilities, NVIDIA has unveiled ‘the world’s 22nd fastest supercomputer’ – DGX SuperPOD – which provides AI infrastructure that meets the massive demands of the company’s autonomous-vehicle deployment program.”

– Mai Tao, Robotics & Automation

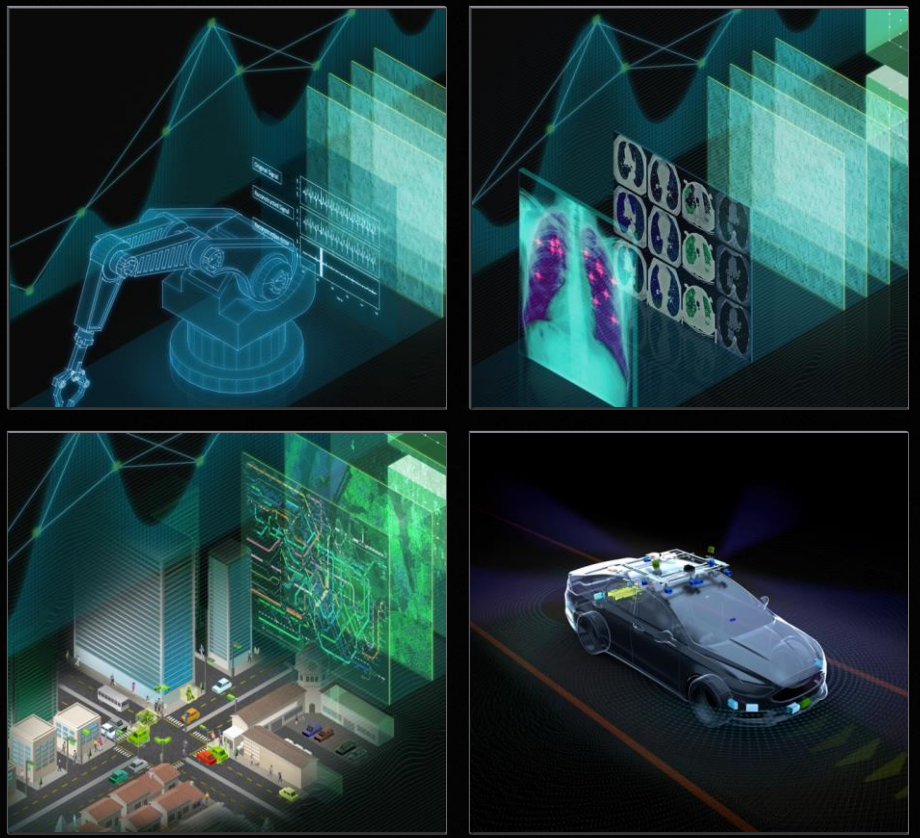
AI LEADERSHIP STARTS WITH AI COMPUTING LEADERSHIP

Researchers racing to advance AI for the world's largest industries – auto, healthcare, manufacturing

Increasingly complex AI models and larger data size demand powerful computers

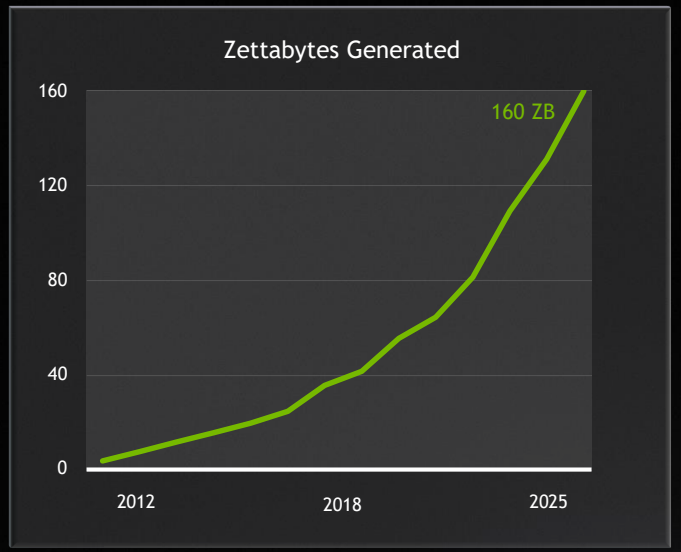
Iteration speed and time-to-train fuels innovation

NVIDIA created DGX SuperPOD to serve as the essential instrument of AI research

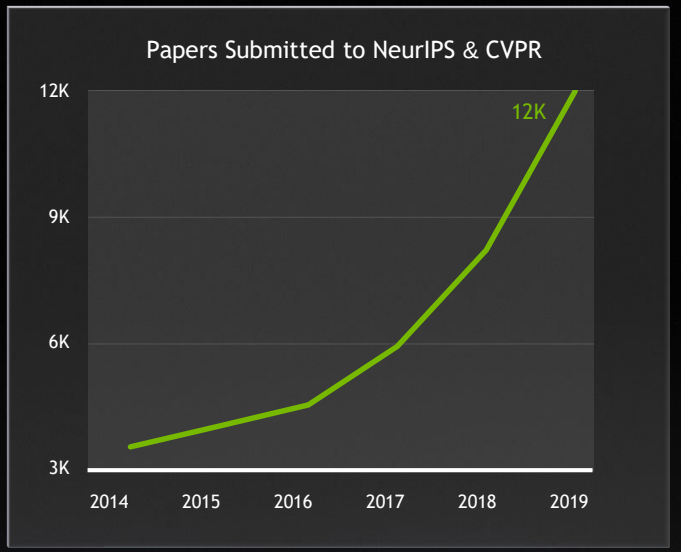


EXPONENTIAL GROWTH IN COMPUTING DEMAND

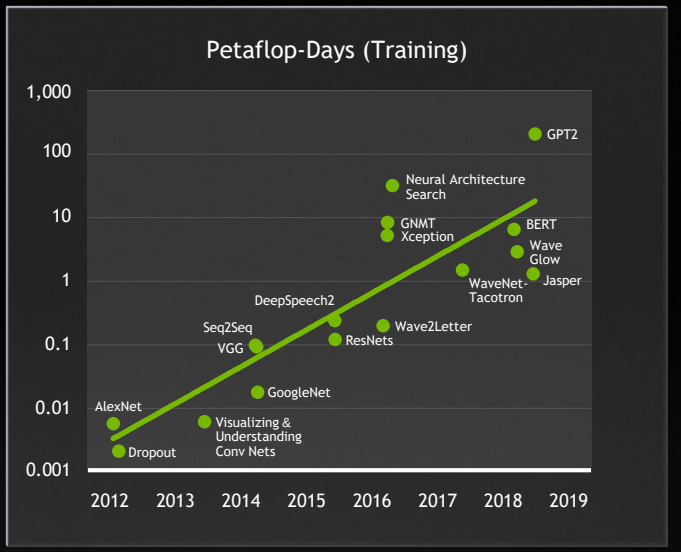
DATA SIZE GROWING



AI RESEARCH GROWING



AI MODEL COMPLEXITY GROWING



Source: IDC, GitHub, and OpenAI / NVIDIA

NVIDIA'S MOTIVATION

- AI leadership demands AI computing infrastructure leadership
- Attract world leading researchers to do groundbreaking work
- Assert industry leadership through groundbreaking performance - win MLPerf
- Explore and innovate infrastructure advancements in networking, storage, cooling, power delivery
- Platform to optimize large scale distributed computing stack

- *At NVIDIA, we built DGX SuperPOD to address these motivations*

DL TRAINING RAPID PROGRESS

36000 Mins (25 Days)
1xK80 | 2015
CUDA

2015

1200 Mins (20 Hours)
DGX-1P | 2016
NVLink

2016

480 Mins (8 Hours)
DGX-1V | 2017
Tensor Core

2017

70 Minutes on MLPerf
DGX-2H | 2018
NVSwitch

6.3 Minutes on MLPerf
At Scale | 2018
DGX-1V Saturn

2018

52.7 Minutes on MLPerf
DGX-2H | 2019
NVSwitch

1.33 Minutes on MLPerf
At Scale | 2019
DGX SuperPOD

2019

ResNet50 v1.5 training

ANNOUNCING NVIDIA DGX SUPERPOD

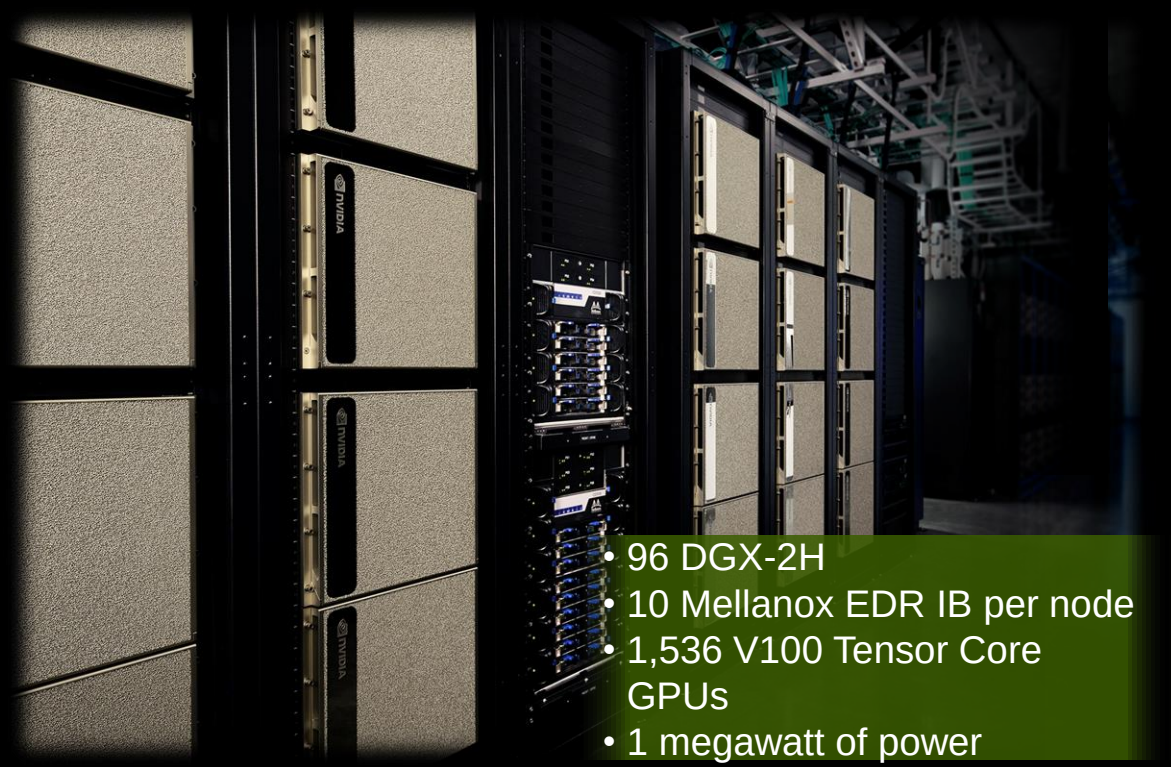
AI LEADERSHIP REQUIRES AI INFRASTRUCTURE LEADERSHIP

Autonomous Vehicles | Speech AI | Healthcare | Graphics | HPC

Test Bed for Highest Performance Scale-Up Systems
- 9.4 PF on HPL | ~200 AI PF | #22 on Top500 list
- <2 mins To Train RN-50

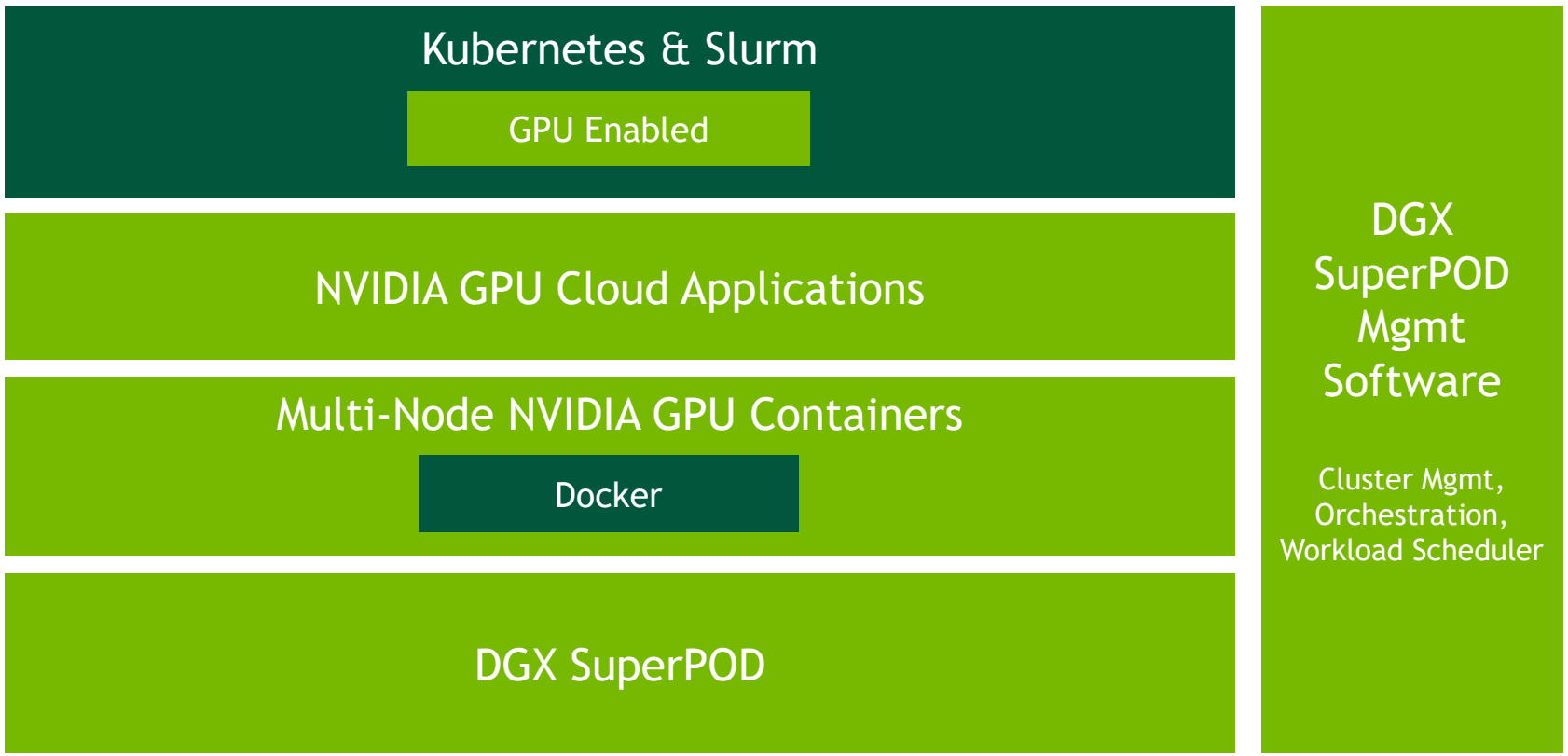
Modular & Scalable GPU SuperPOD Architecture
- Built in 3 Weeks
- Optimized For Compute, Networking, Storage & Software

Integrates Fully Optimized Software Stacks
- Freely Available Through NGC



- 96 DGX-2H
- 10 Mellanox EDR IB per node
- 1,536 V100 Tensor Core GPUs
- 1 megawatt of power

DGX SUPERPOD AI SOFTWARE STACK

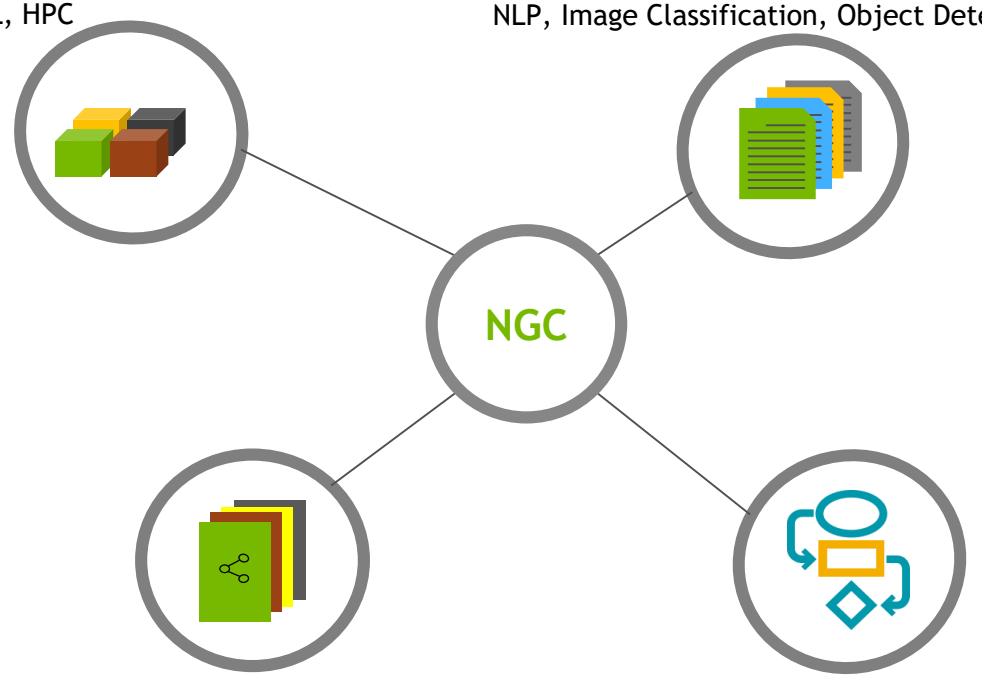


NGC: GPU-OPTIMIZED SOFTWARE HUB

Simplifying DL, ML and HPC Workflows

50+ Containers
DL, ML, HPC

15+ Model Training Scripts
NLP, Image Classification, Object Detection & more

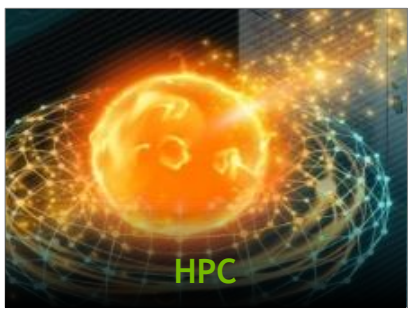


60 Pre-trained Models
NLP, Image Classification, Object Detection & more

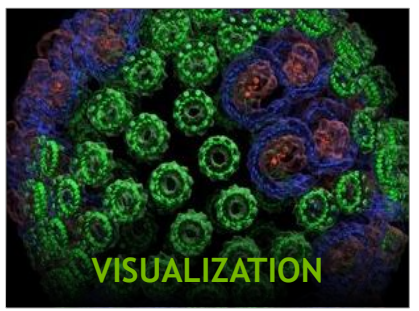
Industry Workflows
Medical Imaging, Intelligent Video Analytics

TensorFlow | PyTorch | more

RAPIDS | H2O | more



NAMD | GROMACS | more



ParaView | IndeX | more

From Concept to World-Record Setting Supercomputer in Three Weeks

The CIRCE Supercomputer and The MLPerf Benchmark Suite

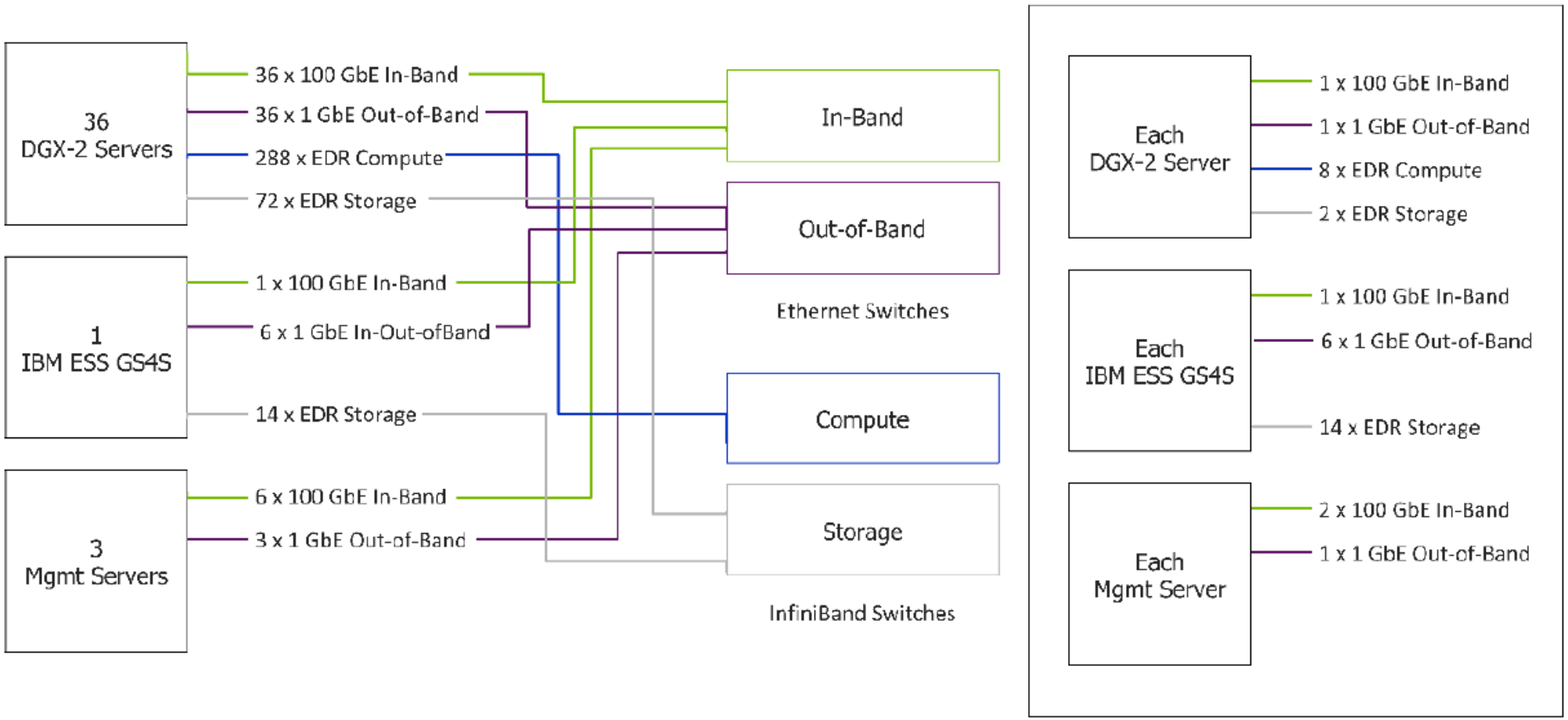
- The MLPerf benchmark suite can be used as a proxy to characterize the performance of GPU-based computer systems.
- Developed with support from over 40 commercial and academic organizations including NVIDIA, Google, Microsoft, Facebook, Intel, AMD, Stanford, and Harvard.
- MLPerf is the first benchmark covering such a broad range of AI algorithms and is representative of many of the workloads in use across industry, research and government.

NVIDIA DGX 2 – CIRCE - 2019-02-15

AI Compute	• 36 NVIDIA DGX2 Nodes • 576 Tesla V100 SXM3 GPUs • 18.4 TB of HBM2 memory • 72 petaFLOPS via Tensor Cores
Networking	• 324 EDR/100 Gbps ports • Eight connections per node • 216 EDR/100 Gbps ports • Two connections per node • Each DGX-2 server has ethernet connections to both switches
Storage	• IBM Elastic Storage Server (ESS) GS4S • 276 TB raw storage • > 40 GB/s read performance

THE CIRCE SUPERCOMPUTER

Network Topology



DGX-2H STORAGE AND CACHING

Hierarchy

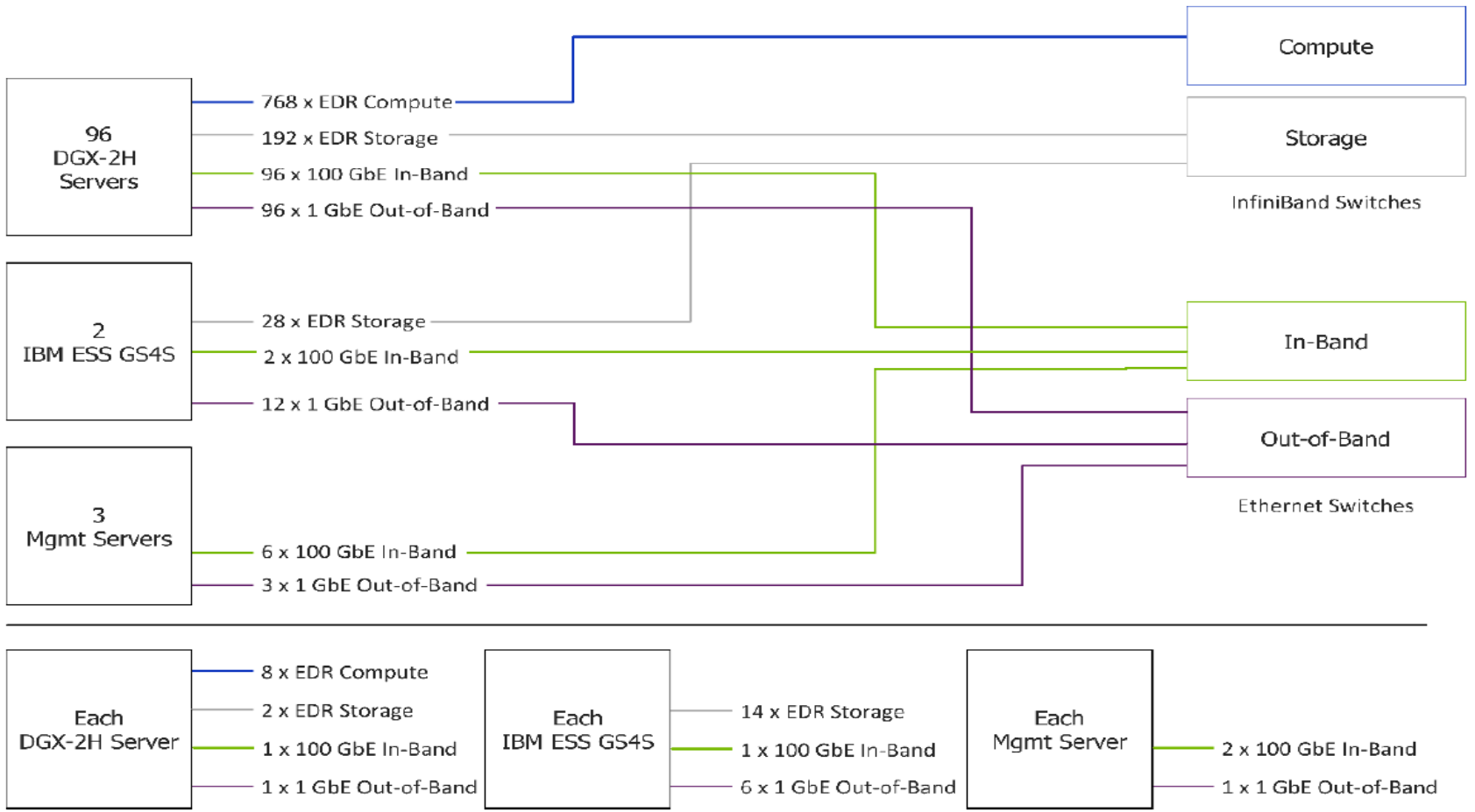
Storage Hierarchy Level	Technology	Total Capacity	Read Performance
RAM	DDR4	1.5 TB per node	> 100 GB/s
Internal Storage	NVMe	30 TB per node	> 15 GB/s
High-Speed Storage	IBM SPECTRUM SCALE using SSD	276 TB	> 40 GB/s aggregate > 15 GB/s per node
Long Term Storage	NFS using SSD	> 1 PB	10 GB/s aggregate 1 GB/s per node
Home	NFS using NVMe	30 TB	N/A

NVIDIA DGX SUPERPOD

AI Compute	• 96 NVIDIA DGX2 Nodes • 1,536 V100 GPUs • 192 PF Peak • 49 TB HBM2 memory
Networking	• 1 Terabit Data Bandwidth per Node • 10 Mellanox EDR InfiniBand per Node • Fully Connected EDR InfiniBand Switch
Storage	• IBM Spectrum Scale - 554 TB total flash storage • 8 storage disk shelves and two ESS nodes with 1 management server • 26 IB ports allocated to the management server and ESS nodes for client data and inter cluster traffic • 100 Gbps ports are allocated on the management server. • 6 - 1 Gbps or 10 Gbps ethernet ports are allocated for BMC/XCAT. Note: this is an allocation on the internal switch for the appliance.

THE DGX SUPERPOD

Network Topology



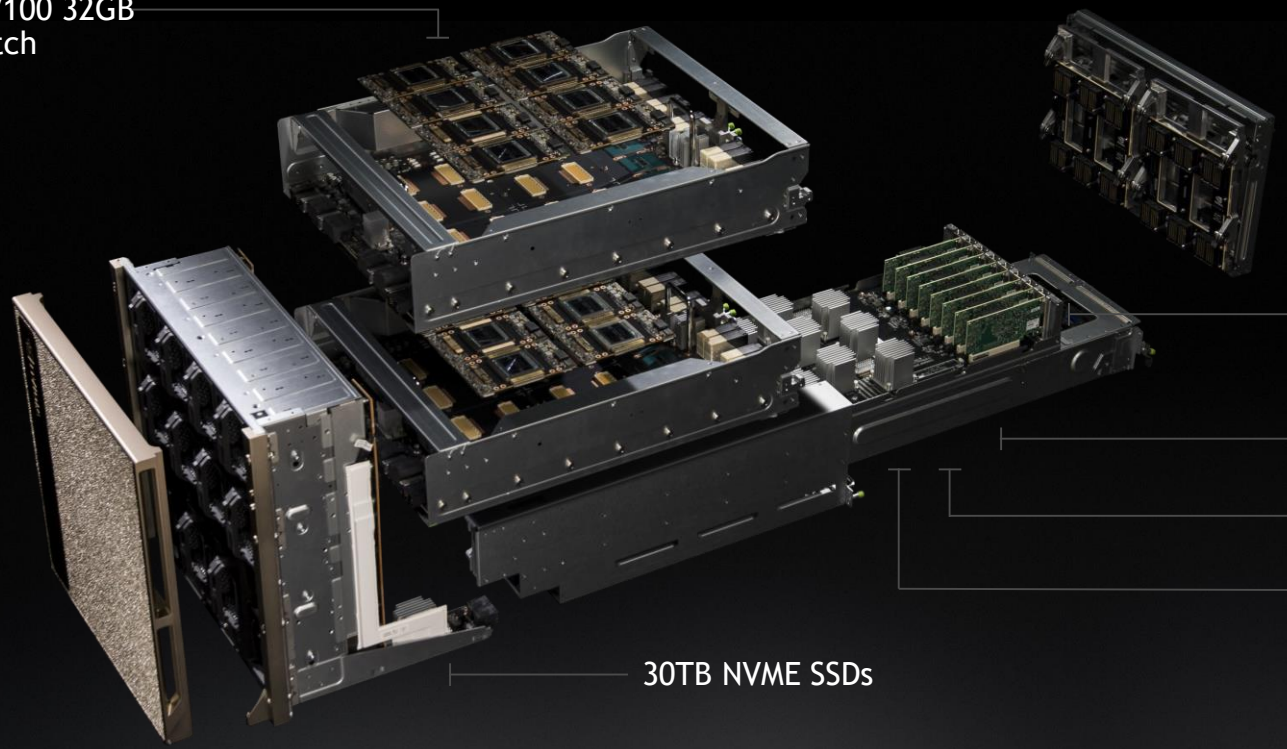
DGX-2H STORAGE AND CACHING

Hierarchy

Storage Hierarchy Level	Technology	Total Capacity	Read Performance
RAM	DDR4	1.5 TB per node	> 100 GB/s
Internal Storage	NVMe	30 TB per node	> 15 GB/s
High-Speed Storage	IBM SPECTRUM SCALE using SSD	552 TB	> 80 GB/s aggregate > 15 GB/s per node
Long Term Storage	NFS using SSD	> 1 PB	10 GB/s aggregate 1 GB/s per node
Home	NFS using NVMe	30 TB	N/A

NVIDIA DGX-2

The World's Most Powerful AI Computer



2 PFLOPS | 512GB HBM2 | 10kW | 350 lbs

NVIDIA DGX SUPERPOD

Terabit-Speed InfiniBand Networking per Node

Mellanox EDR 100G InfiniBand Network

Mellanox Smart Director Switches

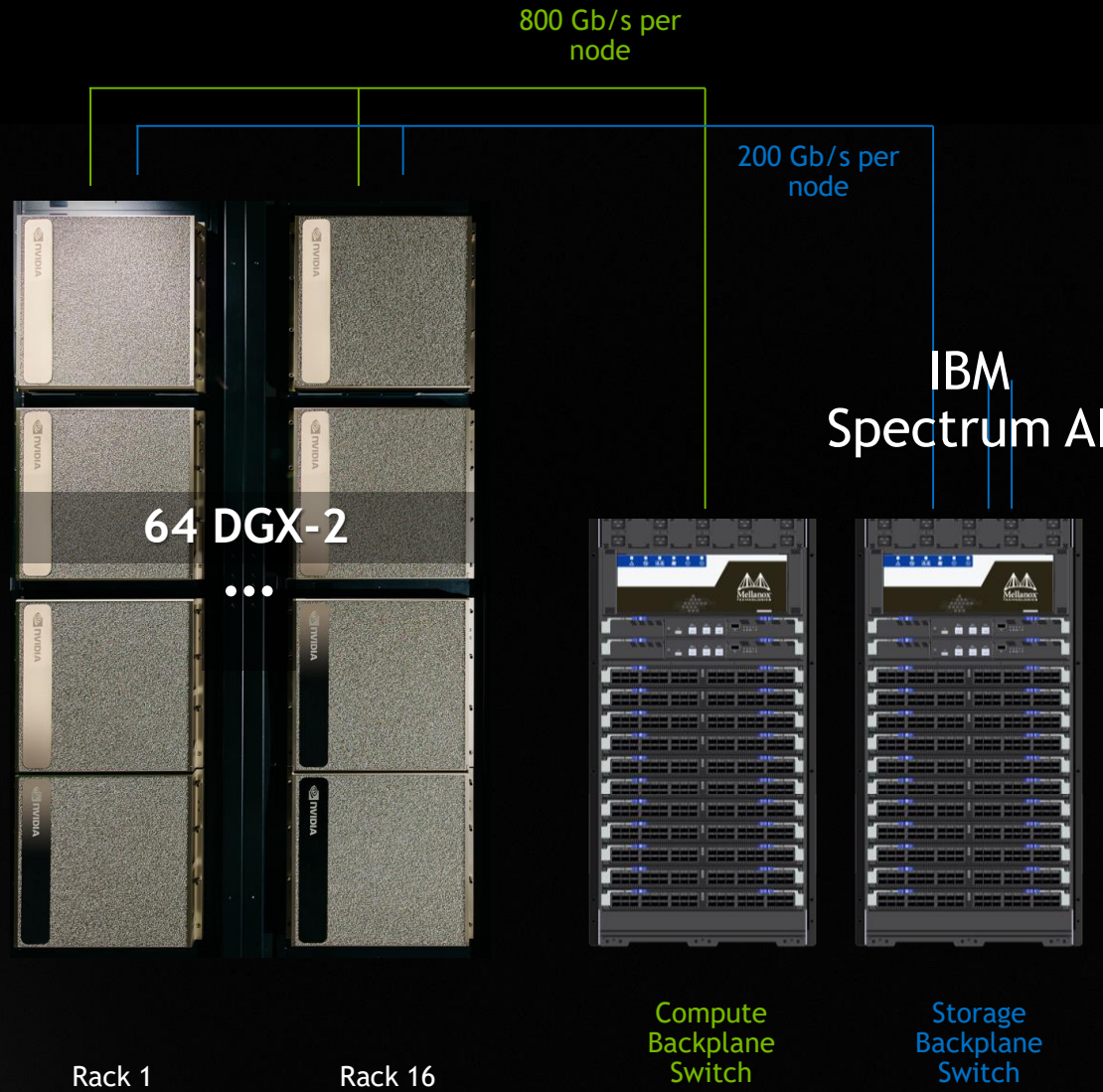
In-Network Computing Acceleration Engines

Fast and Efficient Storage Access with RDMA

Up to 130Tb/s Switching Capacity per Switch

Ultra-Low Latency of 300ns

Integrated Network Manager



White paper:
https://nvidia.highspot.com/items/5d073ad681171721086b2788

INSTANT AI INFRASTRUCTURE

Tackling the Most Complex AI Problems with a Proven Solution that’s Easy to Deploy

Installed and operational in 2 weeks

Record-breaking performance “out of the box”

Fully integrated and optimized AI HPC stack for max performance and productivity

Fast today, faster tomorrow as NVIDIA continues to create optimized algorithms

NVIDIA DGX SUPERPOD BREAKS AI RECORDS

MLPERF 2019

Record Type	Benchmark	Record
Max Scale (Minutes to Train)	Object Detection (Heavy Weight) Mask R-CNN	18.47 Mins
	Translation (Recurrent) GNMT	1.8 Mins
	Reinforcement Learning (MiniGo)	13.57 Mins
Per Accelerator (Hours to Train)	Object Detection (Heavy Weight) Mask R-CNN	25.39 Hrs
	Object Detection (Light Weight) SSD	3.04 Hrs
	Translation (Recurrent) GNMT	2.63 Hrs
	Translation (Non-recurrent) Transformer	2.61 Hrs
	Reinforcement Learning (MiniGo)	3.65 Hrs

Per Accelerator comparison using reported performance for MLPerf 0.6 NVIDIA DGX-2H (16 V100s) compared to other submissions at same scale except for MiniGo where NVIDIA DGX-1 (8 V100s) submission was used| MLPerf ID Max Scale: Mask R-CNN: 0.6-23, GNMT: 0.6-26, MiniGo: 0.6-11 | MLPerf ID Per Accelerator: Mask R-CNN, SSD, GNMT, Transformer: all use 0.6-20, MiniGo: 0.6-10

NVIDIA DGX-READY DATA CENTERS

Validated network of colocation service providers

DGX-1 and DGX-2 optimized, next-gen data centers

Cost-effective OpEx model for infrastructure hosting

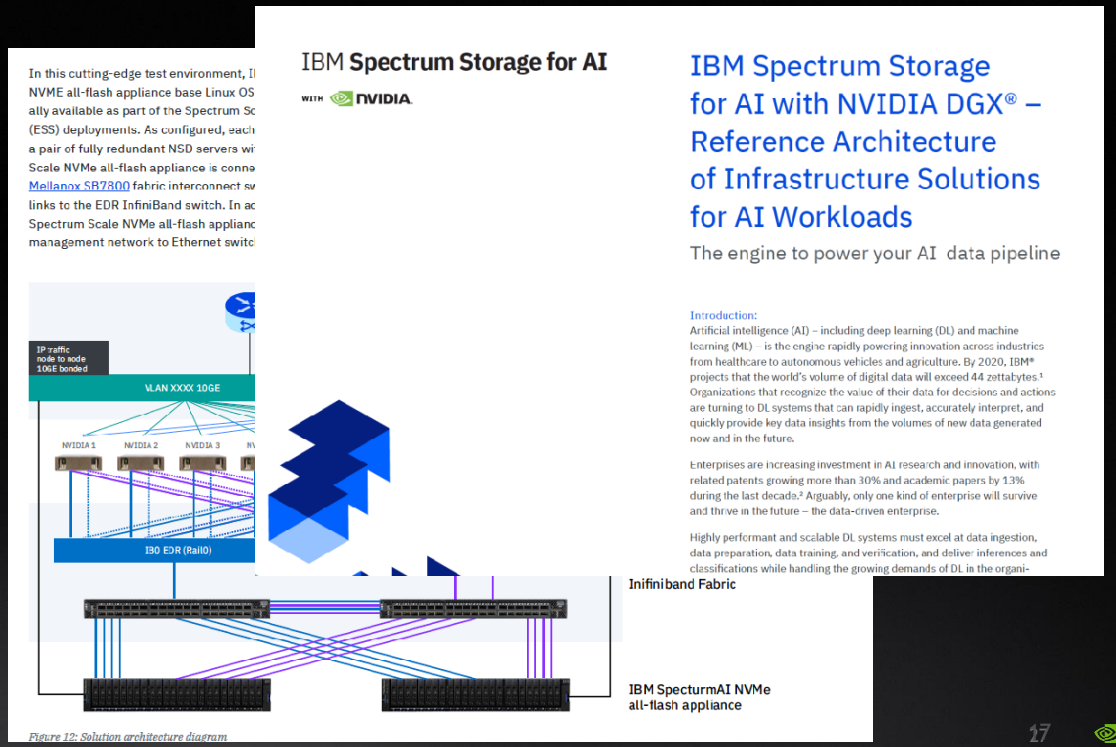
NVIDIA Partners Ready to Host Your DGX SuperPODs

IBM & NVIDIA REFERENCE ARCHITECTURE

VALIDATED DESIGN FOR DEPLOYING DGX AT-SCALE WITH IBM STORAGE

Download at
https://bit.ly/2GcYbgO

Learn more about
DGX RA Solutions at:
https://bit.ly/2OpXYeC



THANK YOU!

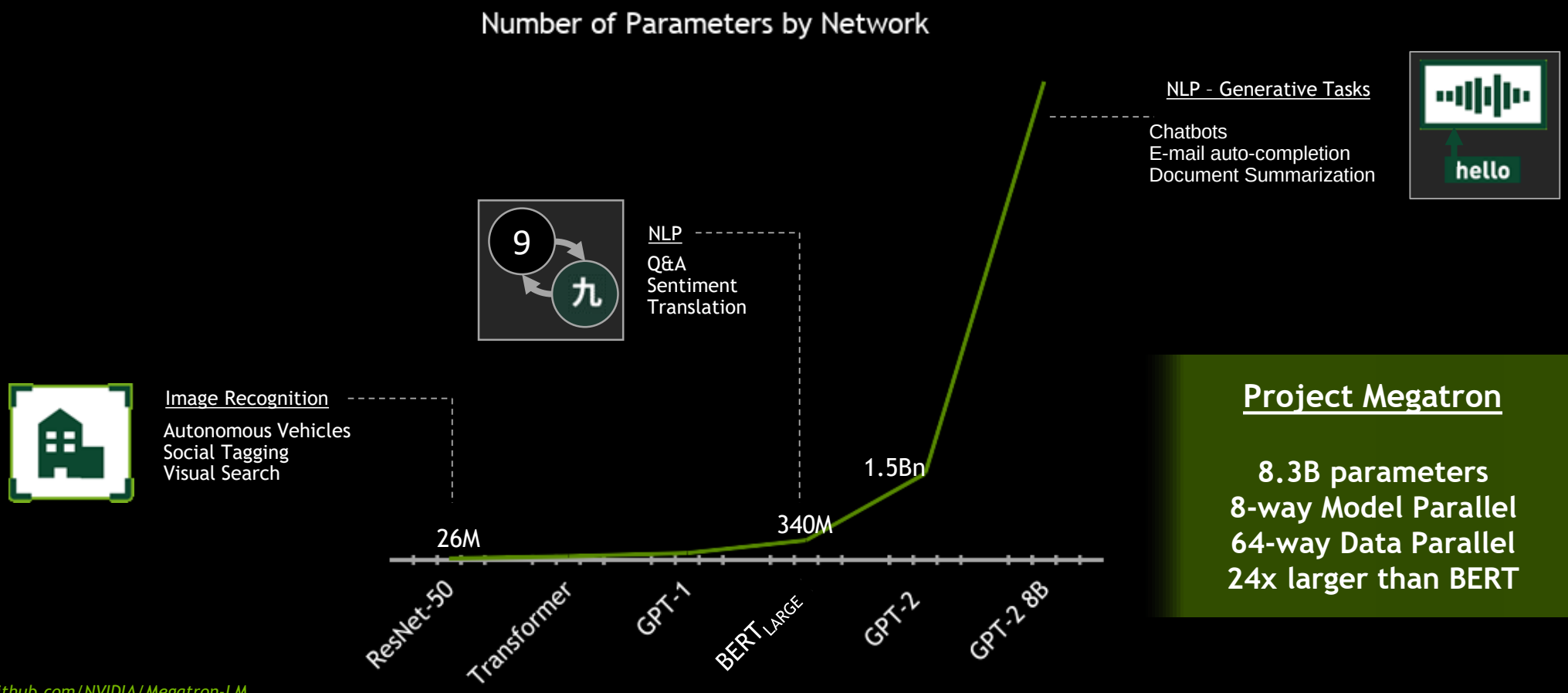
DGX SUPERPOD AND IBM SPECTRUM SCALE AI

Instant AI HPC Infrastructure

BACKUP SLIDES

Deep Learning Models Increasing in Complexity

Next-Level Use-Cases Require Gigantic Models



Project Megatron

8.3B parameters
8-way Model Parallel
64-way Data Parallel
24x larger than BERT

https://github.com/NVIDIA/Megatron-LM

DEEP LEARNING DATASETS INCREASING IN SIZE

- Advances in unsupervised learning allow for training DL models on large amounts of non-annotated data
  - Language model data: BooksCorpus (800 million words), English Wikipedia (2.5 billion words), multi-language Wikipedia, WebText (OpenAI, 8M documents, 40 GB of text)
  - Data for GAN training: unlabeled images and videos

- Labeled data also increasing:
  - From ImageNet (2012) - 1.3M Images on 1000 categories to Open Images (2019) - 9M Images on 6000 categories
  - Semi-autonomous vehicles generate 20~40MB of image data per second, or 0.5~1.1TB of data for every 8h of driving.

WHY MULTI-NODE TRAINING?

ABCi
Supercomputer
Japan AIST
~550 Petaflop/s

aws

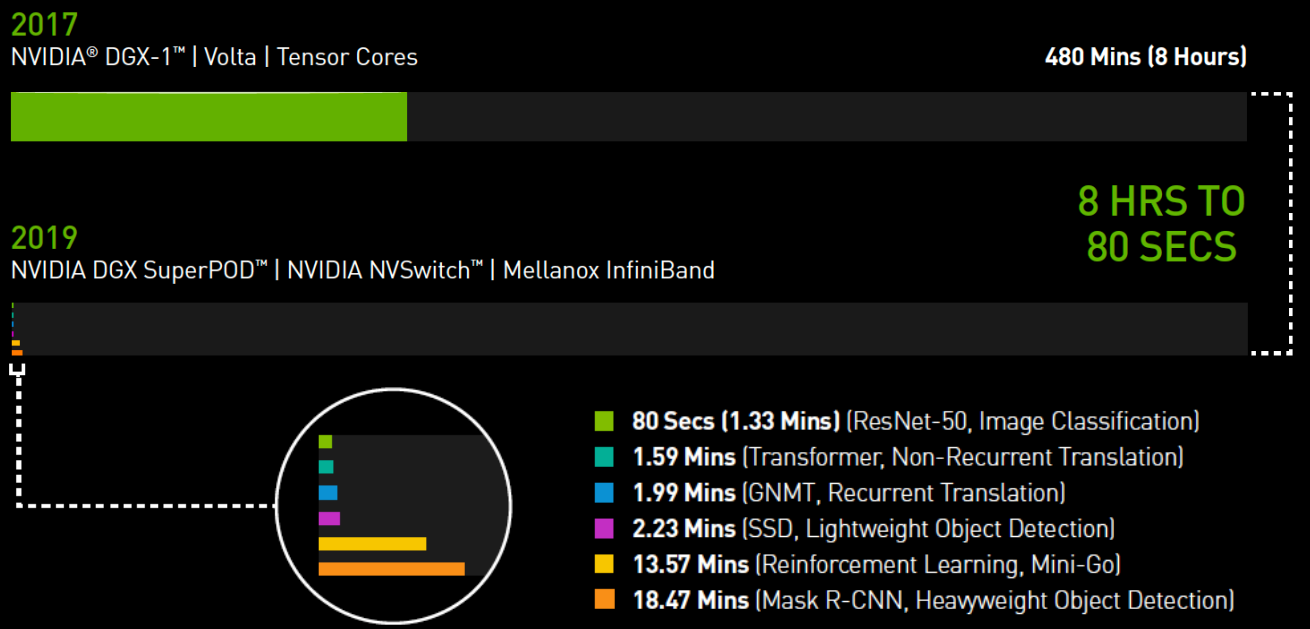
p3dn.24xlarge
cloud instances
~64 Petaflop/s

NVIDIA DGX
SaturnV
1.8 Exaflop/s

Would roughly take 1-30 days to train AlphaGo Zero

AT SCALE MLPERF TRAINING

Smashing Time to Train From 8 Hours to 80 Seconds On V100



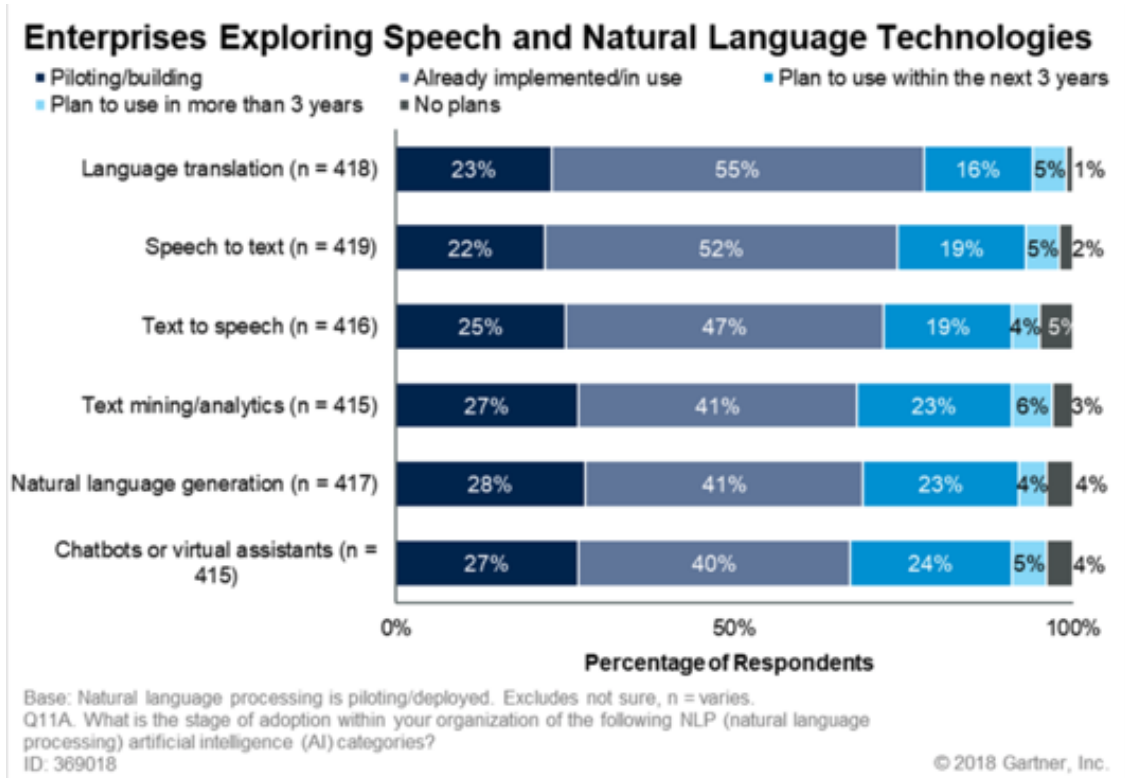
2019 MLPerf ID (in order from top to bottom of chart): ResNet-50: 0.6-30 | Transformer: 0.6-28 | GNMT: 0.6-14 | SSD: 0.6-27 | Mini-Go: 0.6-11 | Mask R-CNN: 0.6-23

Enterprise NLP Trend

Unstructured content represents as much as **80%** of enterprise information resources.

A recent Gartner Research Circle survey on data and analytics trends shows that organizations are **actively developing text analytics** as part of their data and analytics strategies.

80% of survey respondents either have text analytics in use or plan to use it within the next **two years**.



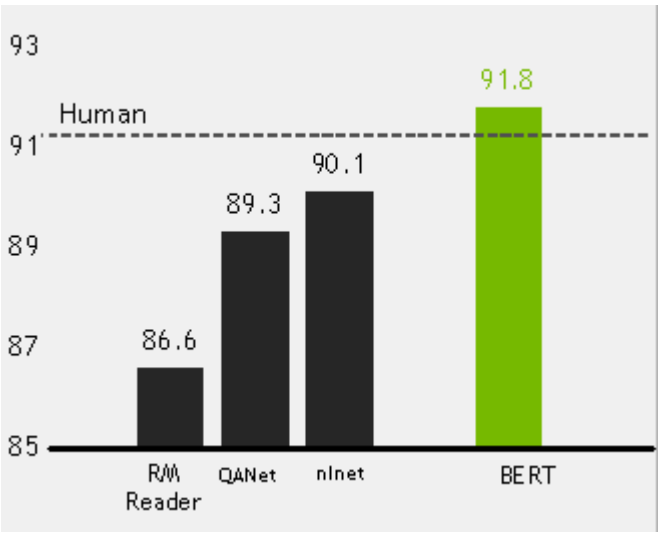
BERT: Flexibility + Accuracy for NLP Tasks

9th October, Google submitted GLUE benchmark

Sentence Pair Classification	MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG
Single Sentence Classificitaion	SST-2, CoLA
Question Answering	SQuAD
Single Sentence Tagging	CoNLL-2003 NER

"BERT is a method of pre-training language understanding" model on a large text corpus care about (like question answering).

BERT outperforms previous methods because it is the first *unsupervised*, *deeply bidirectional* system for pre-training NLP."



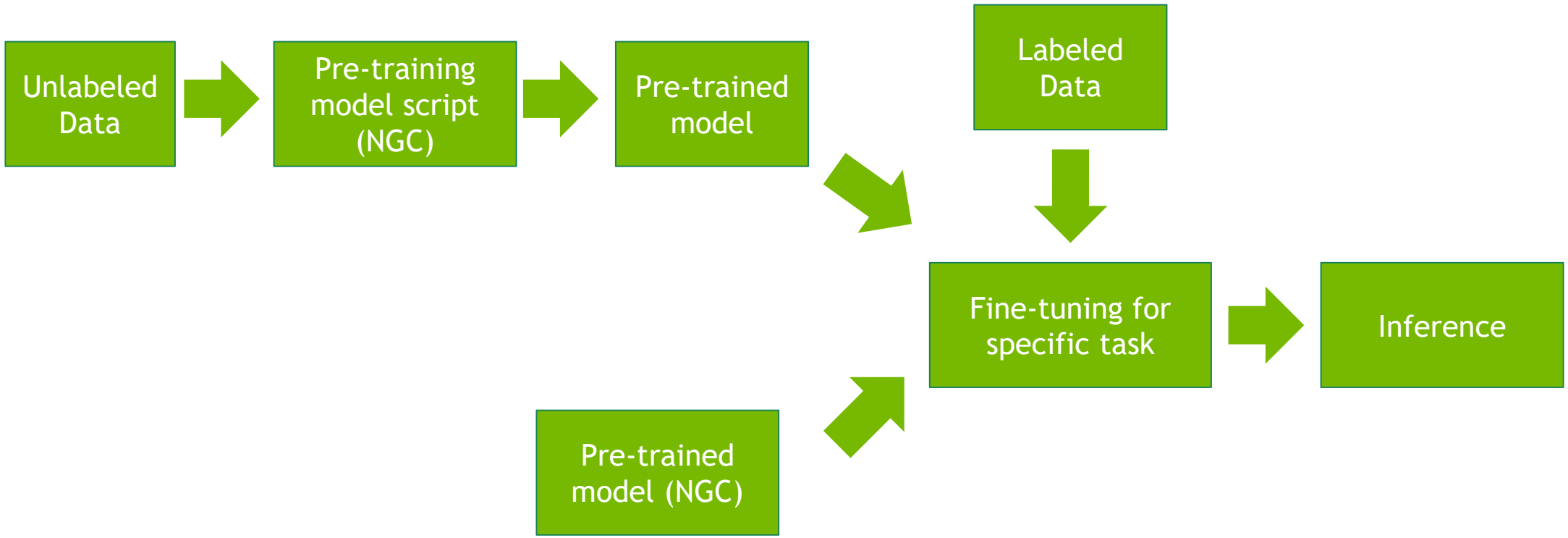
Super Human Question & Answering

9th October, Google submitted GLUE benchmark

- **Sentence Pair Classification:** MNLI, QQP, QNLI, STS-B, MRPC, RTE, SWAG
- **Single Sentence Classification:** SST-2, CoLA
- **Question Answering:** SQuAD
- **Single Sentence Tagging:** CoNLL - 2003 NER

System	MNLI-(m/mm) 392k	QQP 363k	QNLI 108k	SST-2 67k	CoLA 8.5k	STS-B 5.7k	MRPC 3.5k	RTE 2.5k	**Average** -
Pre-OpenAI SOTA	80.6/80.1	66.1	82.3	93.2	35.0	81.0	86.0	61.7	74.0
BiLSTM+ELMo+Attn	76.4/76.1	64.8	79.8	90.4	36.0	73.3	84.9	56.8	71.0
OpenAI GPT	82.1/81.4	70.3	87.4	91.3	45.4	80.0	82.3	56.0	75.1
$BERT_{BASE}$	84.6/83.4	71.2	90.5	93.5	52.1	85.8	88.9	66.4	79.6
$BERT_{LARGE}$	**86.7/85.9**	**72.1**	**92.7**	**94.9**	**60.5**	**86.5**	**89.3**	**70.1**	**82.1**

Two paths to use BERT



Resources:

- NGC Model Scripts: Single Node & Multi node (4 x DGX, 16 x DGX and 64 x DGX) for Pre-training and Fine-tuning
- Deep Learning Examples GitHub

BERT Multi-node Performance Metric Time to Train

DGX-1 (16 GB)	GPUs	Time to train (Hrs)
1	8	153.6 (6.3 days)
4	32	39.3
16	128	10.4

DGX-2H (32 GB)	GPUs	Time to train (Hrs)
1	16	58.4 (2.4 days)
4	64	15.4
16	256	3.9
64	1024	1.2
92	1472	0.9 (53 min)***

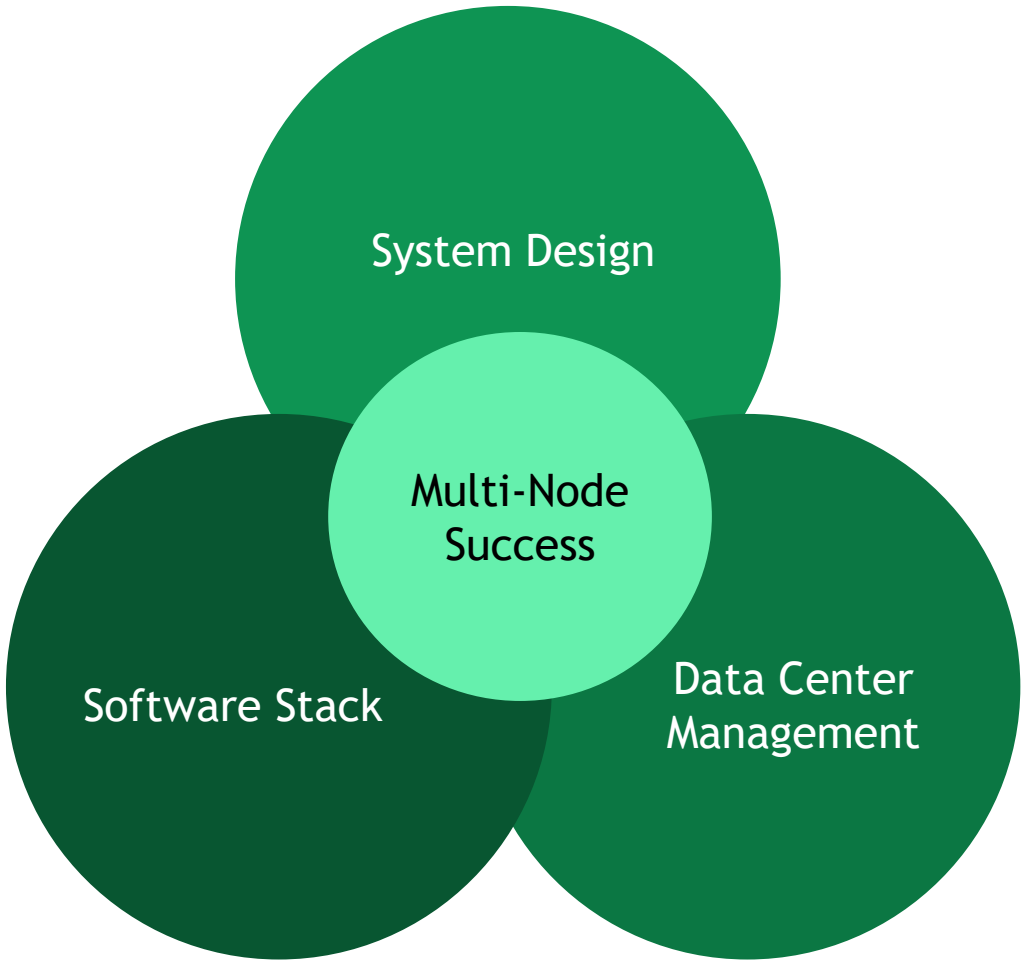
Source: https://github.com/NVIDIA/DeepLearningExamples/tree/master/PyTorch/LanguageModeling/BERT#pre-training-loss-results

** Above time to train is measured for Mixed precision, training loss 1.3 in PyTorch; with LAMB optimizer*

*** Gradient accumulation is applied to DGX-2H 1,4,16 node*

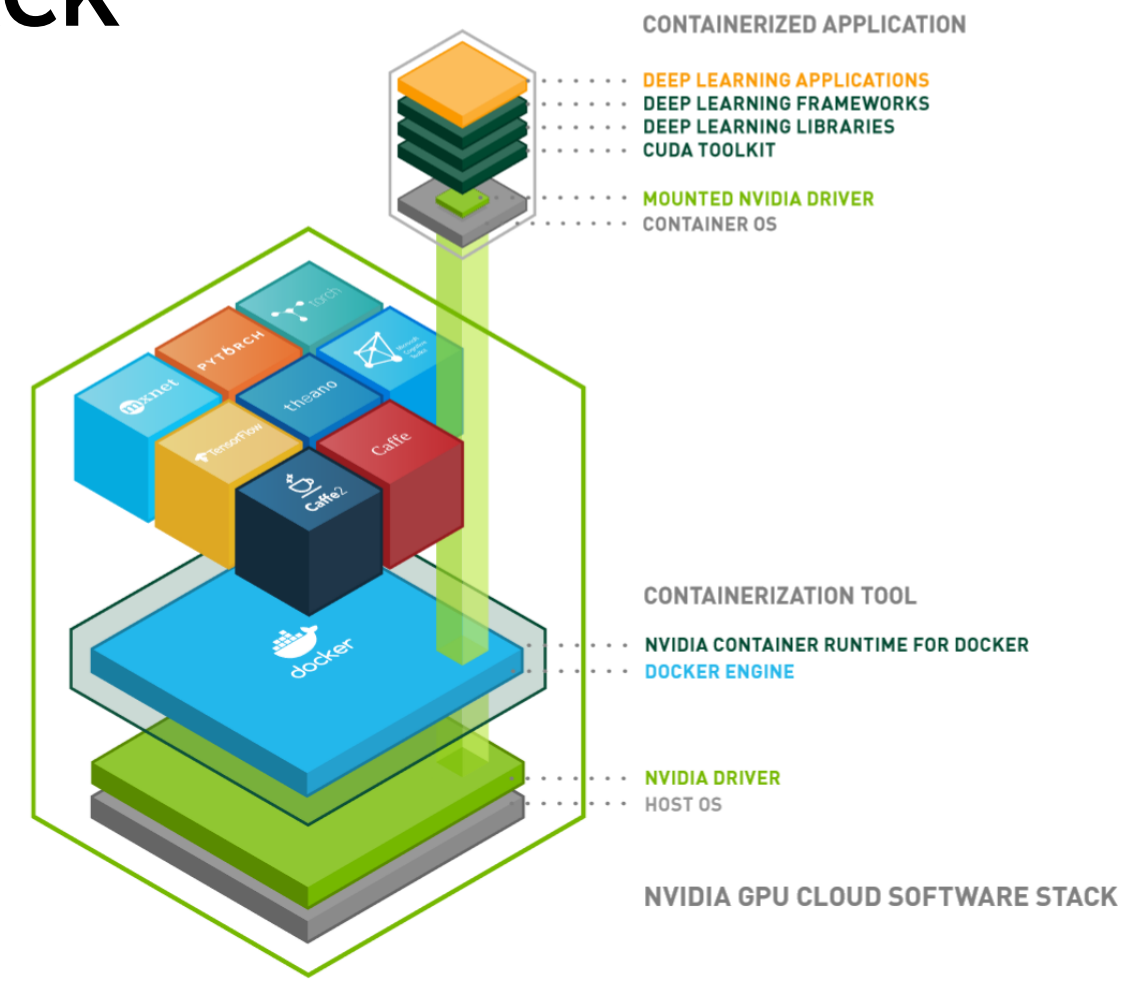
**** 92 DGX-2H results is not yet published to NGC/DeepLearningExamples. For details, please refer to this blog: Fastest training BERT*

Multi-node Requirements



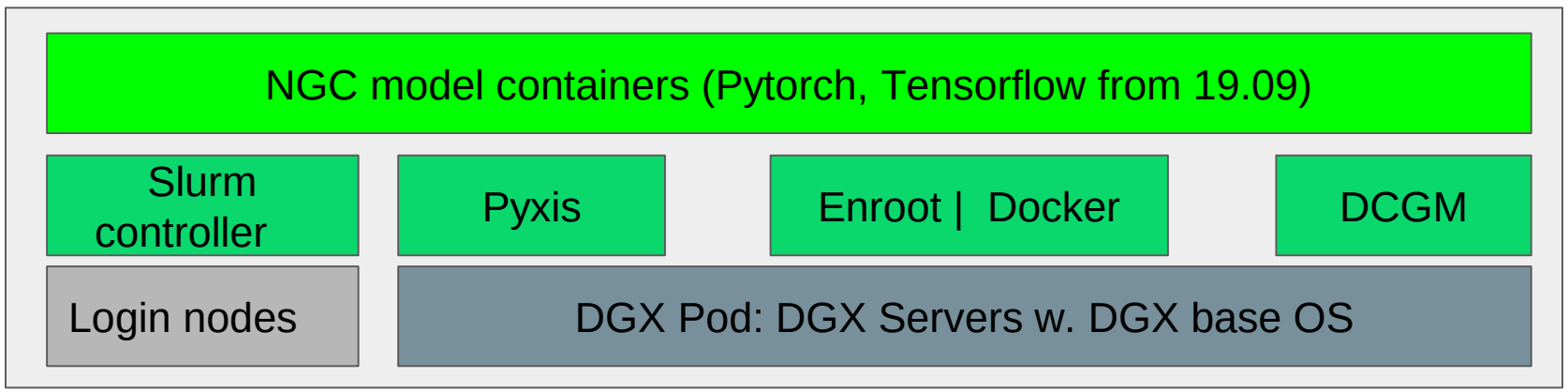
SINGLE NODE NGC DL FW CONTAINER SOFTWARE STACK

- Deep Learning Model:
  - Model hyperparameters tuned for multi-node scaling
  - Multi-node launcher scripts
- Deep Learning Container:
  - Deep learning framework, NVIDIA libraries and multi-node software
- Host:
  - Host OS (DGX OS), GPU drivers, IB driver, container runtime engine (docker, enroot)



ENABLE MULTI-NODE DEEP LEARNING

- **DGX POD Cluster Stack**
  - Slurm: Scheduling for user job management
  - Enroot: NVIDIA-developed tool to convert traditional container/OS images into unprivileged sandboxes
  - Pyxis: NVIDIA-developed plugin that integrates Enroot with Slurm



  - DeepOps: NVIDIA-developed toolbox for GPU cluster management w/Ansible playbooks

CUSTOMER CALL-TO-ACTION

Try Multi-Node BERT and give feedback

- Build your own GPU cluster following the DGX Pod and DGX SuperPOD reference architectures.
- Clone the DeepOps repo and follow the cluster setup guide. Open a GitHub issue if any problem.
- Clone the NVIDIA Deep Learning Example repo and train your own BERT model on multi-node OR download a pre-trained BERT model from NGC and fine-tune for your NLP task
- Provide feedback back to the Tesla product management teams on what's working well and any challenges