ML Resources

Yihui (Ray) Ren, Shinjae Yoo, Meifeng Lin

yren@bnl.gov

May 07 2021

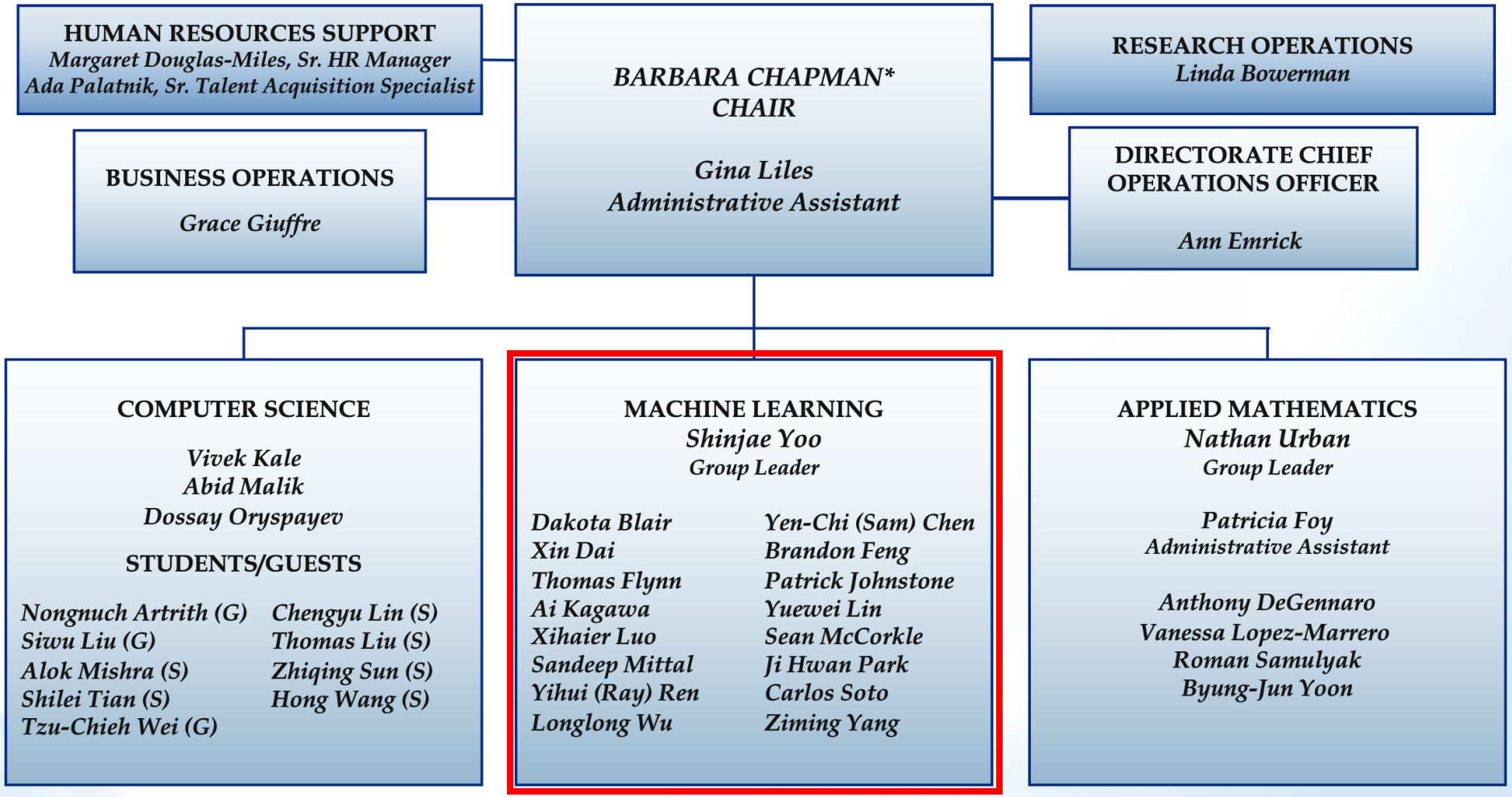
Outline

- CSI Organization Overview
- ML group (Human resources?)
- Computation resources
- Learning materials
- Funding opportunities

Computational Science Initiative (CSI)

- Computer Science and Applied Mathematics [Math, **ML**, CS]
- Computational Science Laboratory [Quantum, HPC]
- Computing for National Security [ACL]
- Scientific Data and Computing Center (SDCC)
- Computation and Data-Driven Discovery (C3D)

COMPUTER SCIENCE AND APPLIED MATHEMATICS



*Reports to Kerstin Kleese van Dam, Director of Computational Science Initiative
ESH Coordinator – Bob Colichio, Pat Carr

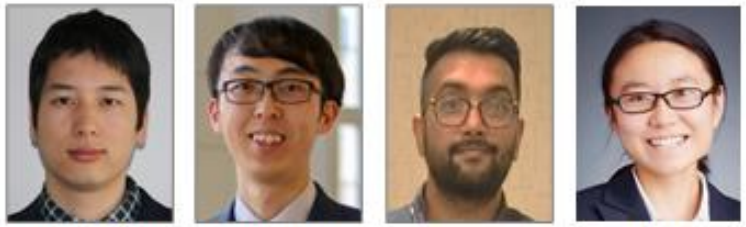
Approved: Barbara Chapman **Date: February 01, 2021**

https://www.bnl.gov/orgcharts/CCSM.pdf

Machine Learning Group

- Group Lead: **Dr. Shinjae Yoo**
- 8 staff scientists
- 6 postdocs (+2 will join in May)
- 3 full time software engineers
- Many collaborative projects with other departments and other labs.

New Members



Yumin Liu

Spatio-temporal Modeling, UQ, Multi-task Learning, Computer Vision

Xihaier Luo

Bayesian Optimization, UQ, Stochastic Simulation, Surrogate Modeling, PDEs

Sandeep Mittal

Real-time System Optimization, Edge Inference, Autonomous Control, Computer

Ziming Yang

ML Software Stack Development, Continuous Integration, Containerization

Chuntian Cao

Multimodal Analysis, Inverse Problem Solving, Energy Storage, Electrochemistry

Matthew Carbone*

Microscopic Theory, Methods for Quasi-particle Systems, ML for Chem-informatics, Computational

Xin Dai

Longitudinal Modeling, Quantum Mechanics, Multi-modal Analytics

Yi Huang

Spatio-temporal Modeling, Medical Data Analytics, Graph/ Network Analysis

Patrick Johnstone

Distributed Algorithms, Large-scale Training, Stochastic Optimization

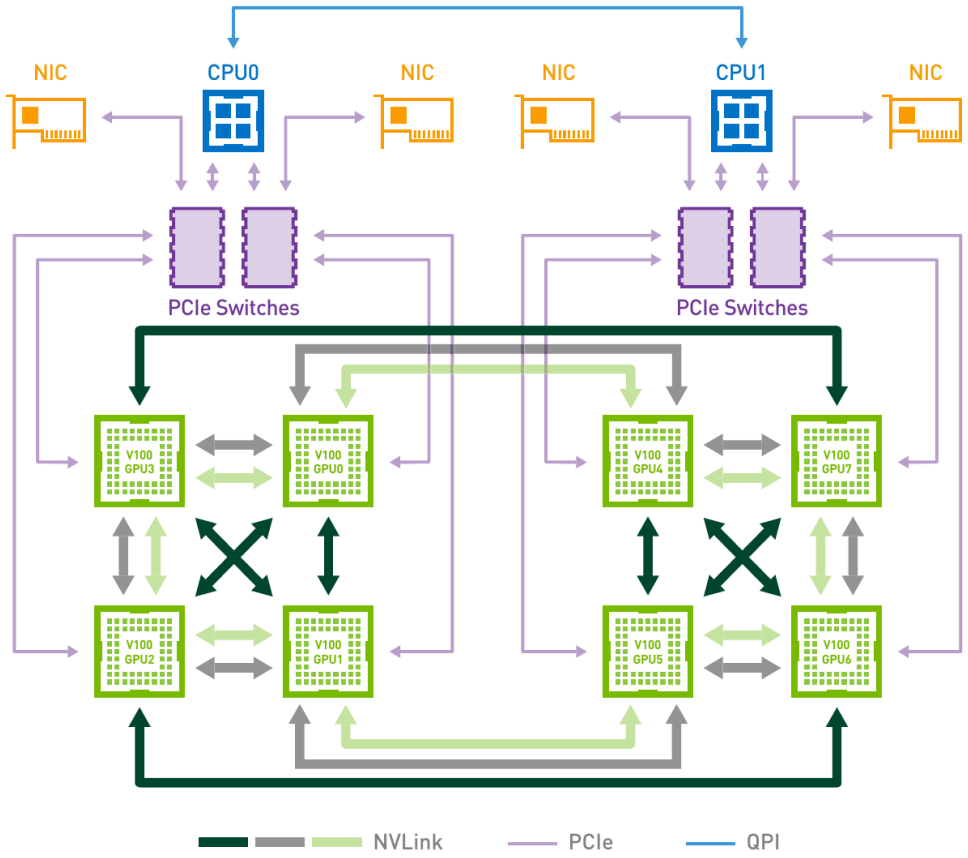
Computation Resources (SDCC)

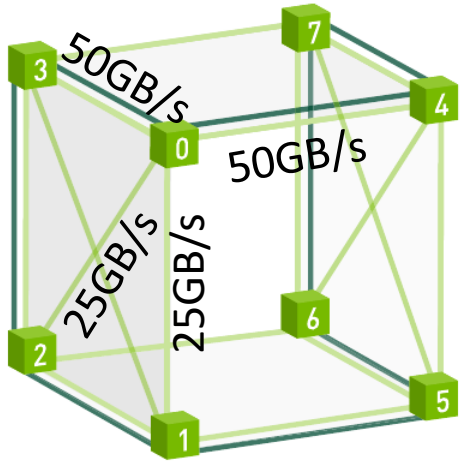
- 108 P100-based nodes with 2 P100s per node
- 5 V100-based nodes with 8 V100s per node
- SDCC is in the process of procuring the next-gen IC.
- If you have any feedback or suggestions on SDCC resources, feel free to contact Christopher Hollowell hollowec@bnl.gov .

Computation Resources (CSI)

- Internal (CSI):
  - DGX-2, 1 testbed. “minerva”
    - 16x V100 (32GB) per node
  - DGX A100 + DDN AI400X, 1 testbed, “athena”
    - 8x A100 (40GB) per node
  - RTX, 1 devbox “mlgpu01”
    - 8x RTX 2080Ti
  - Desktops and laptops
  - External HPC (Summit, NERSC, etc…)

IC Volta (DGX-1)

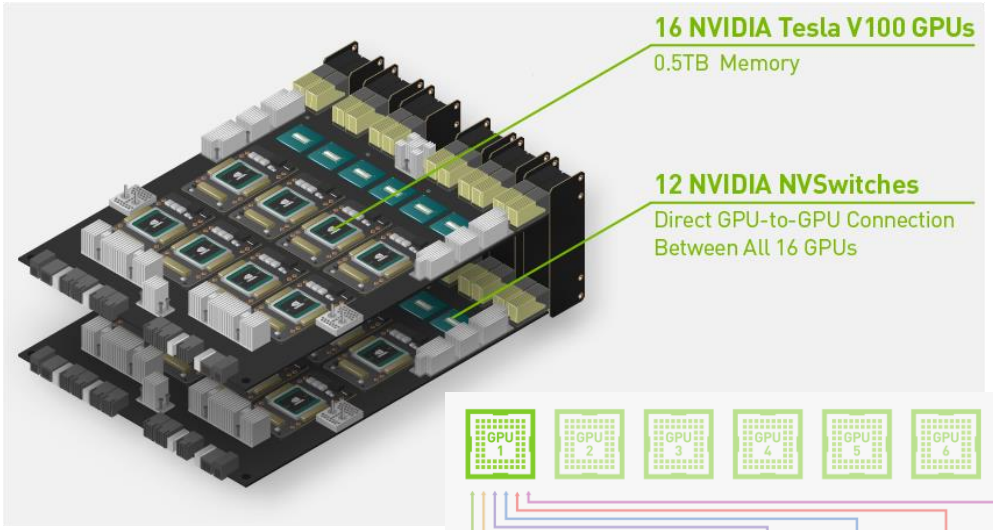


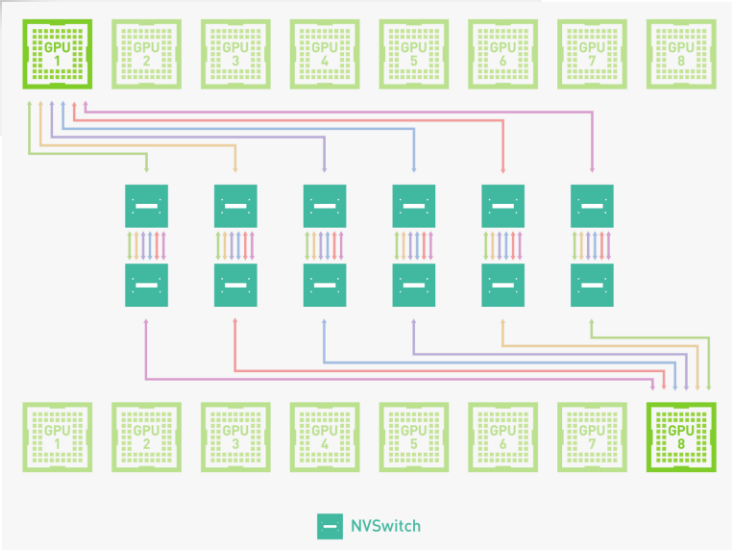


- 8x V100 (32GB) GPUs
- Hybrid cube-mesh topology
- 2x 24-core Xeon 8175M (96 logic cores in total)
- 768 GB system memory
- 2 TB NVMe SSD
- 2x AWS P3 (16 GPUs in total)
- Connected through 1.25GB/s Ethernet.
- (AWS P3dn.24xlarge)

Image Ref:
https://images.nvidia.com/content/pdf/dgx1-v100-system-architecture-whitepaper.pdf

DGX-2 and NVSwitch



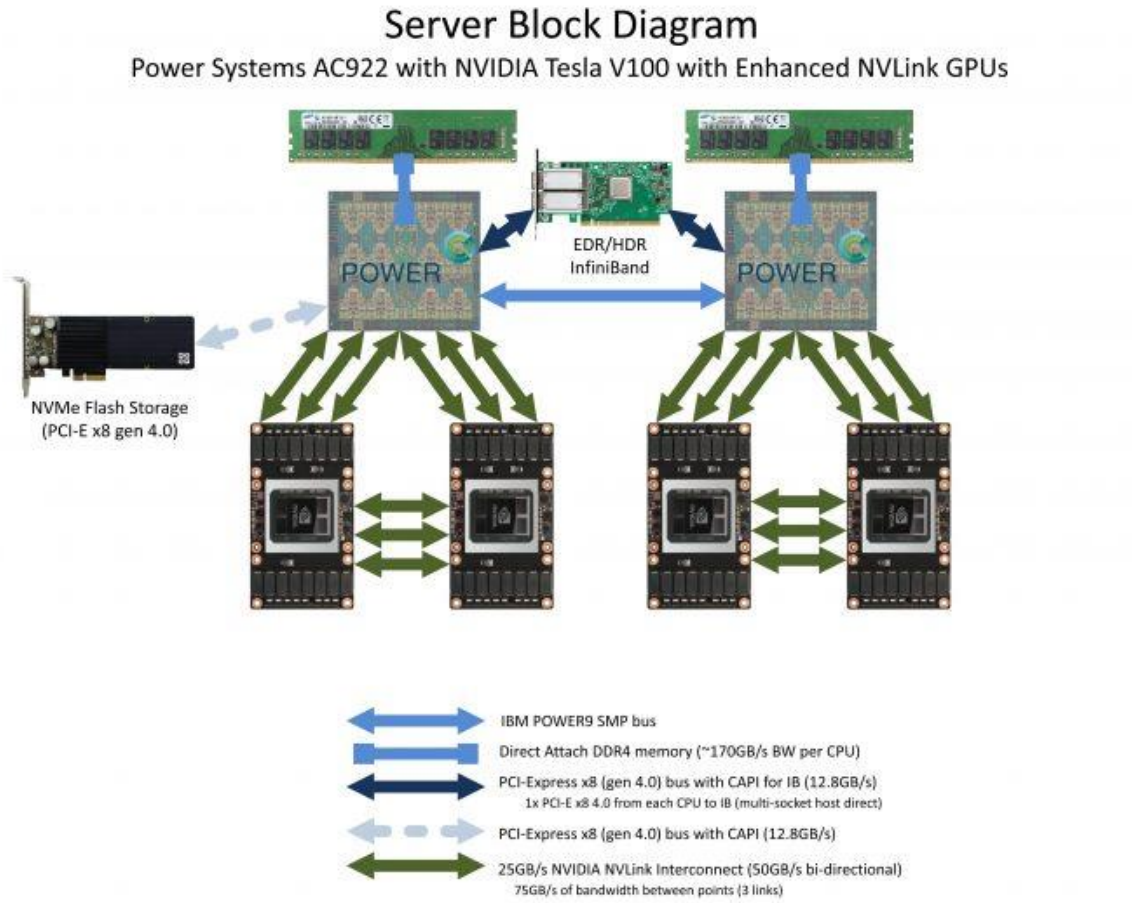


- 16x V100 (32GB) GPUs
- 12x on-node NVSwitches
- Each NVSwitch has 18 NVLink ports (16 in use).
- 2x 24-core Xeon 8186 (96 logic cores in total)
- 1.5 TB system memory
- 30 TB NVMe SSD in 8-way RAID0

Image Ref:

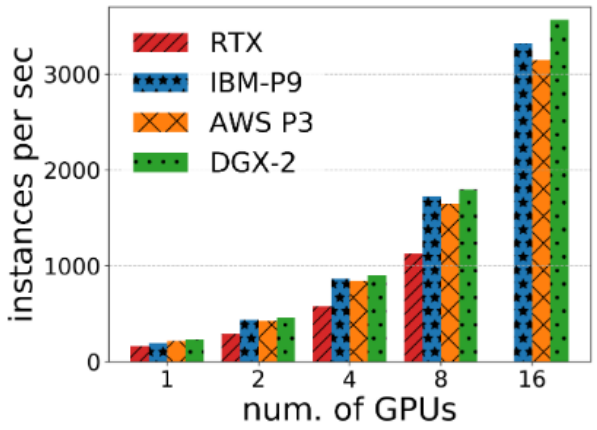
https://images.nvidia.com/content/pdf/nvswitch-technical-overview.pdf
https://www.nvidia.com/en-us/data-center/hgx/

IBM Power System AC922 (8335-GTH)

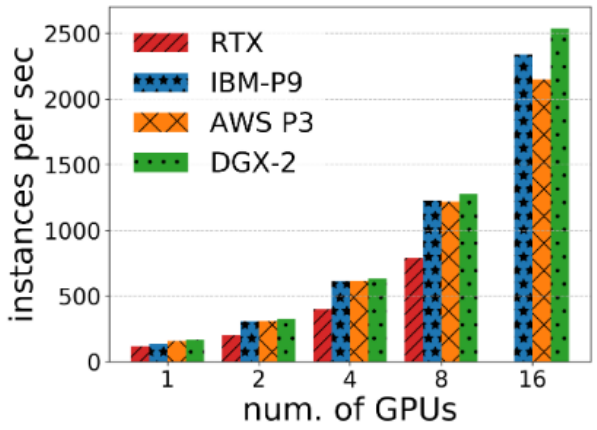


- 4x V100 (32GB) GPUs
- 2x IBM 20-core Power9 CPU (160 logic cores in total)
- Each IBM Power9 CPU has 6 NVLinks.
- Two CPUs are connected by a SMP bus (32GB/s).
- 4x IBM P9 systems (16 GPUs in total)
- Connected through InfiniBand (24 GB/s).
- The tested system uses GPFS (remote filesystem) with block size of 16 MB and bandwidth ~18 GB/s.

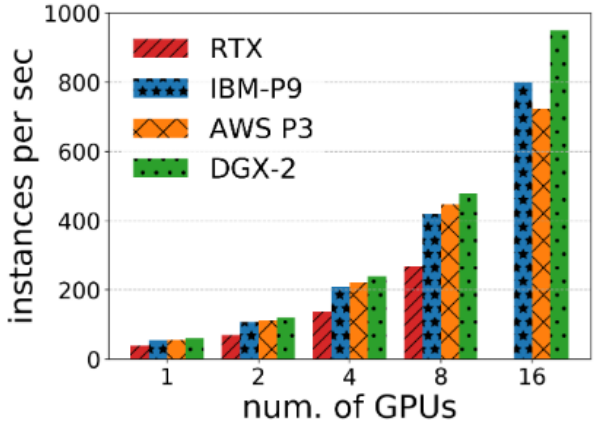
Image Ref:
https://www.microway.com/product/ibm-power-systems-ac922/



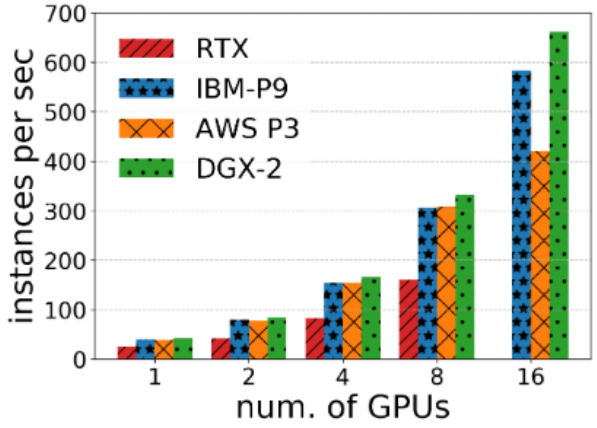
(a) ResNet101



(b) ResNet152



(c) BERT-SWAG



(d) BERT-SQuAD

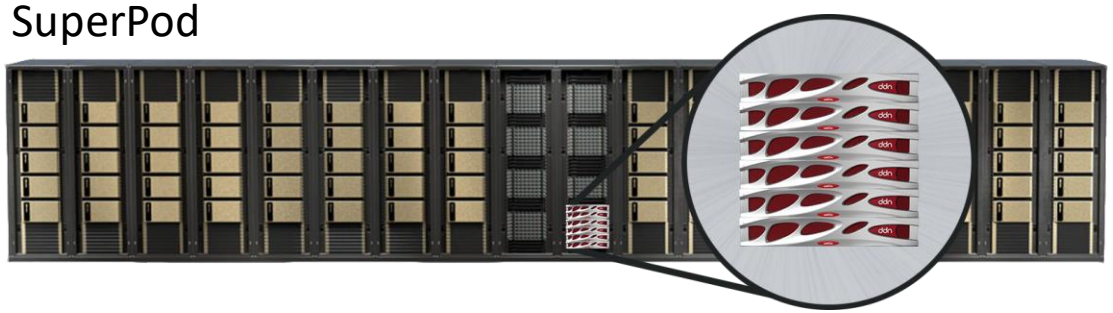
Benchmark Results

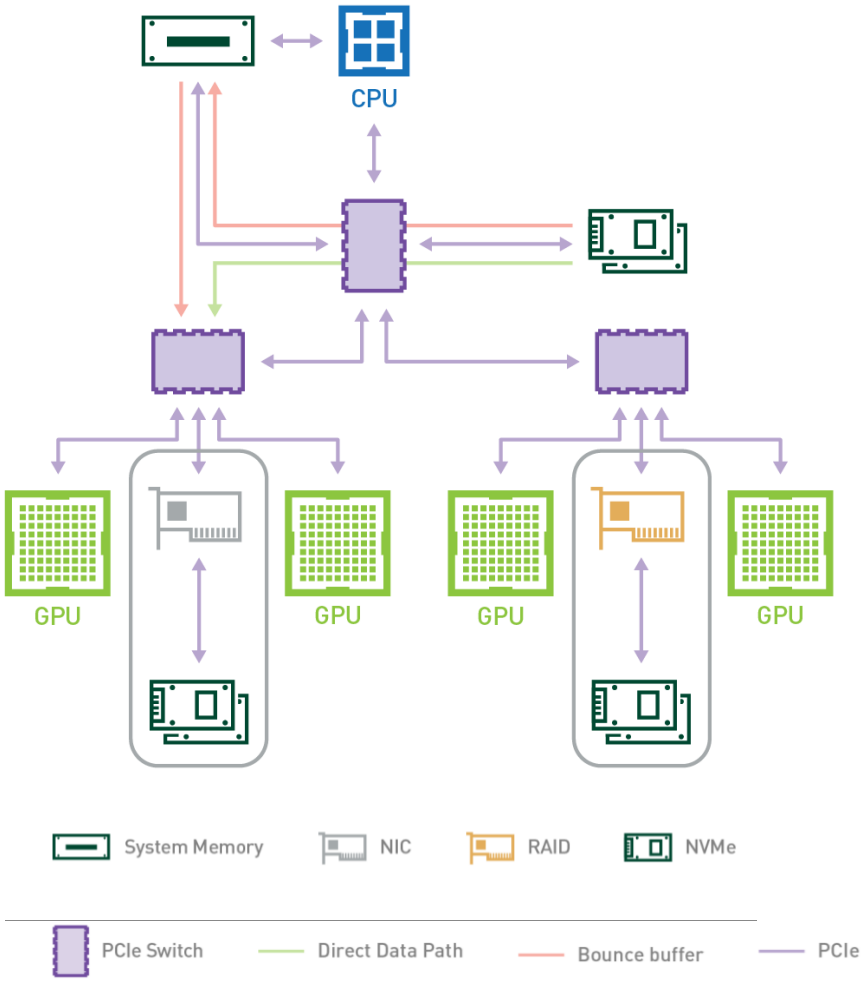
- ResNet152
  - CNN-based computer vision model
  - Model contains 60.34M parameters
  - Training on ImageNet dataset (150GB)
- BERT-SQuAD
  - The state-of-the-art natural language processing model
  - Model contains 109.5M parameters
  - Fine-tuning on Stanford Question Answering Dataset
- Similar linear scaling for both AWS-P3 and DGX2 up to 8 GPUs
- DGX2 has better linear weak scaling up 16 GPUs

https://arxiv.org/abs/1905.08764

DGX A100 + DDN AX400

SuperPod





- Direct Data I/O through PCIe bypassing host bounce buffer.
- Promised to reach 40GB/s between DGX and AX400.

2 x AMD EPYC 7742 64-Core Processor
8x A100 GPU (40GB)
2 TB memory
14 TB 4-way raid0 for data
2 TB raid1 for code
6 x IB ConnectX-6 VPI HDR/200GbE

Image credit:
https://www.ddn.com/
https://developer.nvidia.com/blog/gpudirect-storage/

Learning materials

- AI/ML working group (lead by Meifeng, under construction)[link]
  - In our plan, the purpose of the BNL AI/ML working group is to
  - Form a lab-wide AI/ML community to share common interests, challenges and collaboration opportunities
  - Provide timely and appropriate training events to the lab staff on fundamentals of scientific AI/ML
  - Share progress of scientific AI/ML to the lab community via seminars/workshops
  - Possible activities:
    - Bi-weekly journal clubs, ML/AI seminars
    - Tutorials and summer schools
    - GPU hackathons
- SDCC JupyterHub [link]

Funding Opportunities

- LDRD collaborations
  - Targeting scientific applications
  - Inspiring novel AI algorithms
  - Deploying AI solutions into products
- SciDAC Institutes (in future)
  - Currently, Shinjae co-leads the AI group in RAPIDS2 institute.
  - We expect future institutes targeting HEP and NP.
- ASCR proposal calls