Accelerate AI Even More with NVIDIA

Dr. Arts Yang, Sr. Product Architect DGX

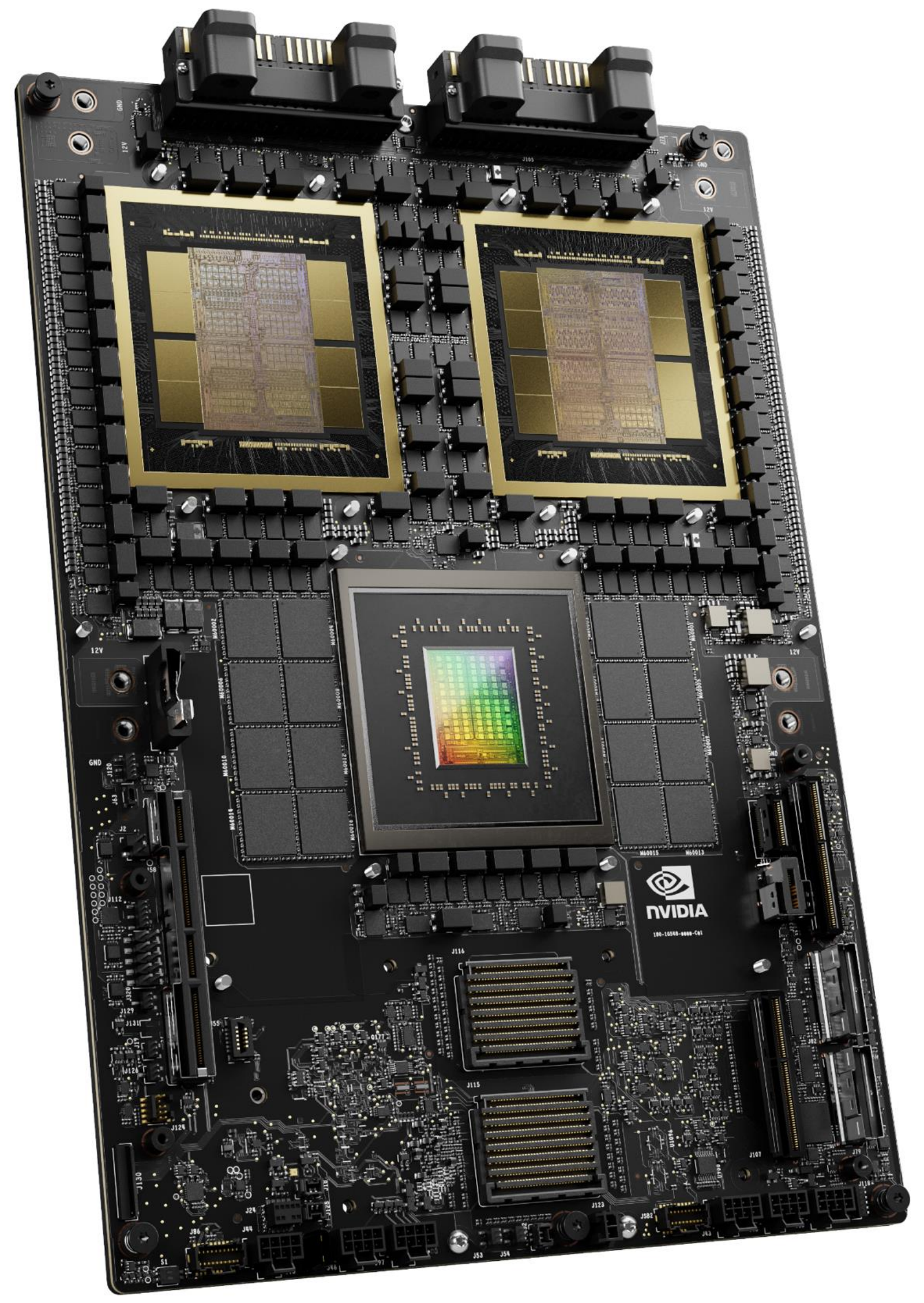
arts@nvidia.com

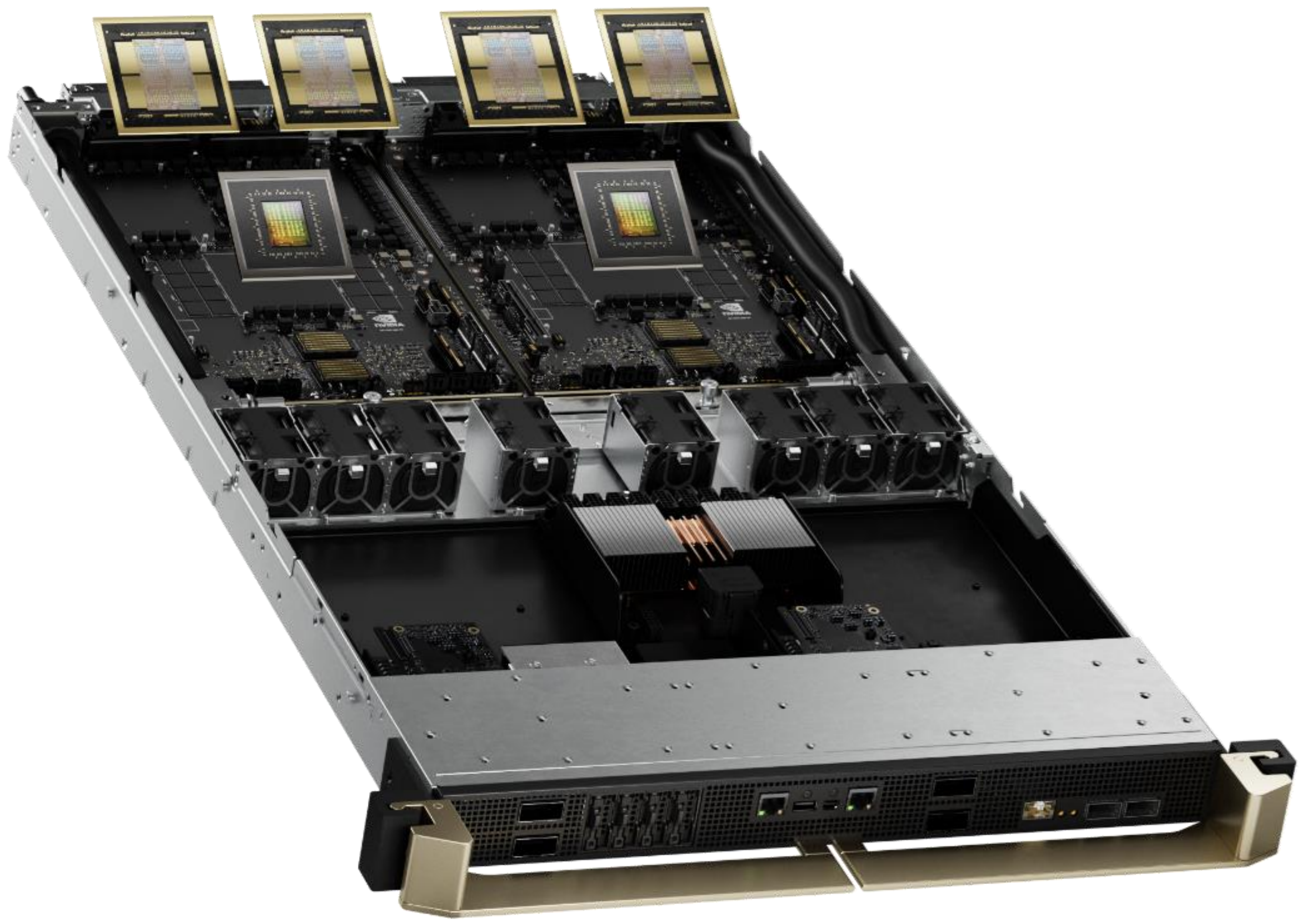
AI DAYS 2024 in Prague

23.10.2024

GB200 NVL72 Compute and Interconnect Nodes

Building Blocks for the GB200 NVL72 Rack





GB200 SUPERCHIP

40 PETAFLOPS FP4 AI INFERENCE
20 PETAFLOPS FP8 AI TRAINING
864GB FAST MEMORY

GB200 SUPERCHIP COMPUTE TRAY

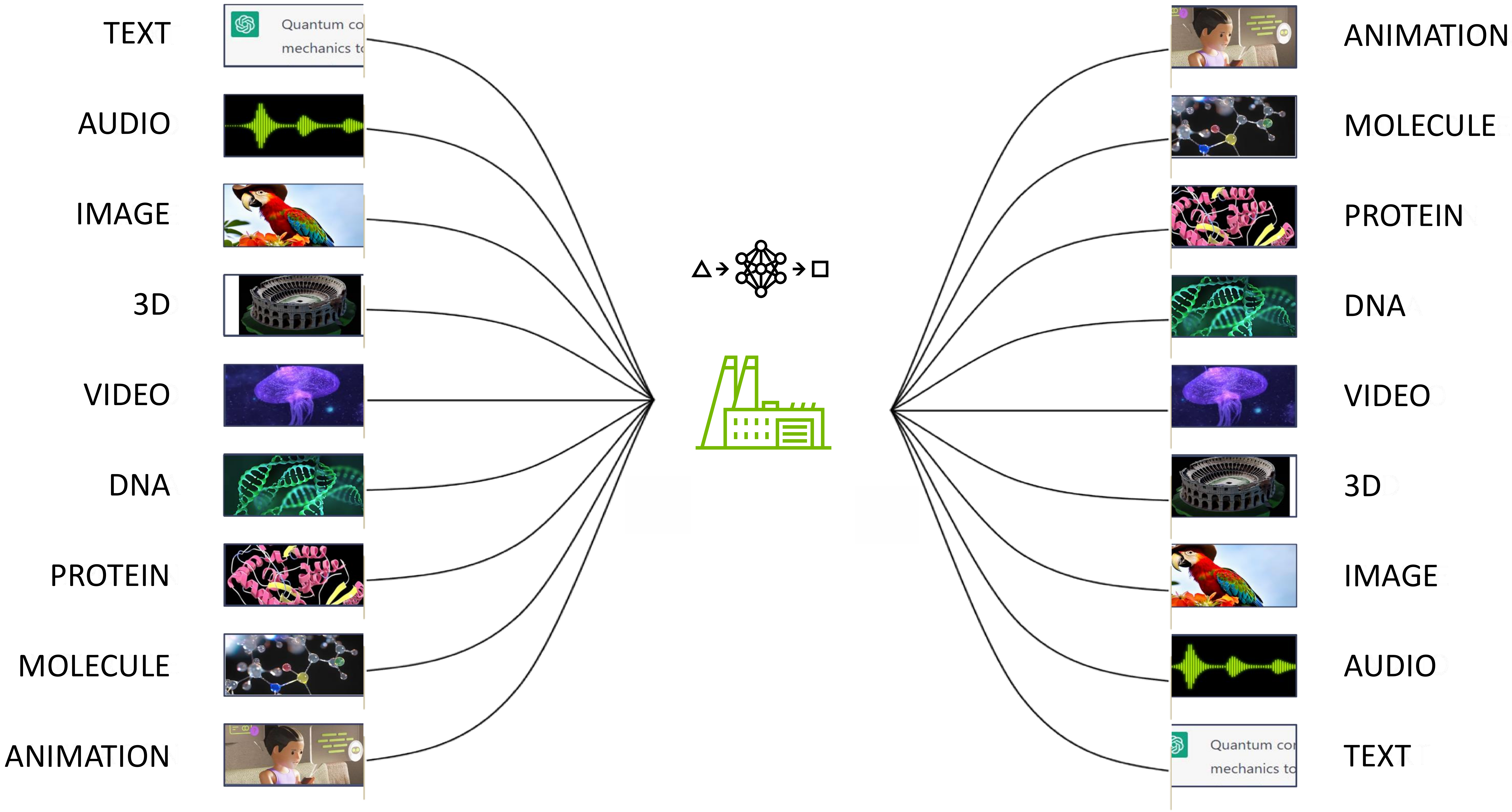
2x GB200
80 PETAFLOPS FP4 AI INFERENCE
40 PETAFLOPS FP8 AI TRAINING
1728 GB FAST MEMORY
1U Liquid Cooled
18 Per Rack

NVLINK SWITCH TRAY

2x NVLINK SWITCH CHIP
14.4 TB/s Total Bandwidth
SHARPv4 FP64/32/16/8
1U Liquid Cooled
9 Per Rack

The Next Era of Generative AI

AI factories unlock $100T industries



TEXT
AUDIO
IMAGE
3D
VIDEO
DNA
PROTEIN
MOLECULE
ANIMATION

ANIMATION
MOLECULE
PROTEIN
DNA
VIDEO
3D
IMAGE
AUDIO
TEXT

Building Generative AI Applications for the Enterprise

Build, customize and deploy generative AI models with NVIDIA NeMo

Data Curation
Distributed Training
Model Customization
Accelerated Inference
Retrieval Augmented Generation
Guardrails
NeMo Curator
Megatron Core
NeMo Aligner
In-domain, secure, cited responses
In-domain queries
Triton & TensorRT-LLM
NeMo Retriever
NeMo Guardrails
Model Development
Enterprise Application Deployment
NVIDIA NeMo
NVIDIA AI Enterprise

Microsoft Azure aws ORACLE NVIDIA DGX Cloud DELL Technologies Hewlett Packard Enterprise Lenovo SUPERMICRO

NVIDIA

Generative AI Trends

The iPhone Moment of AI

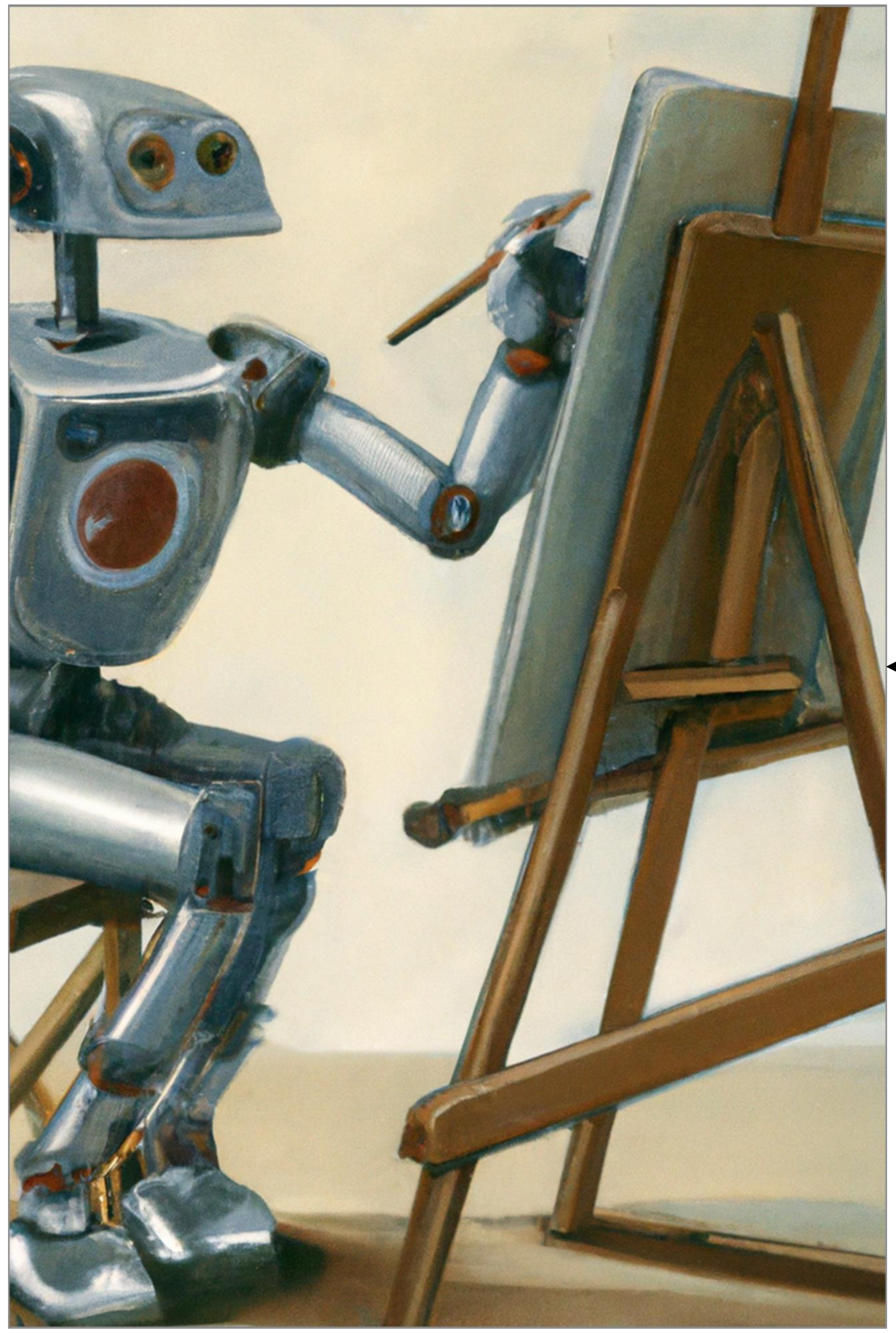
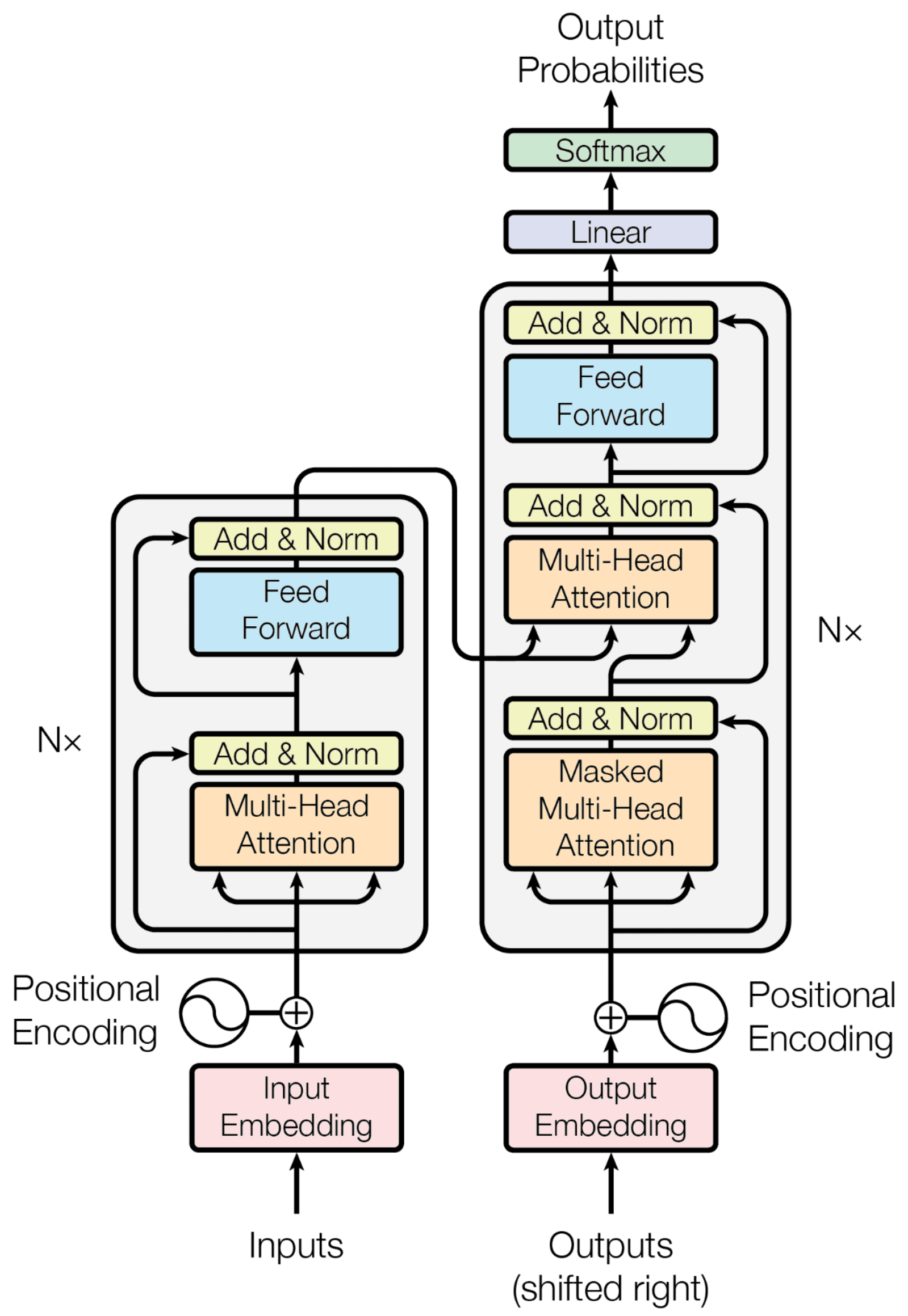


Image Generation

Stable Diffusion

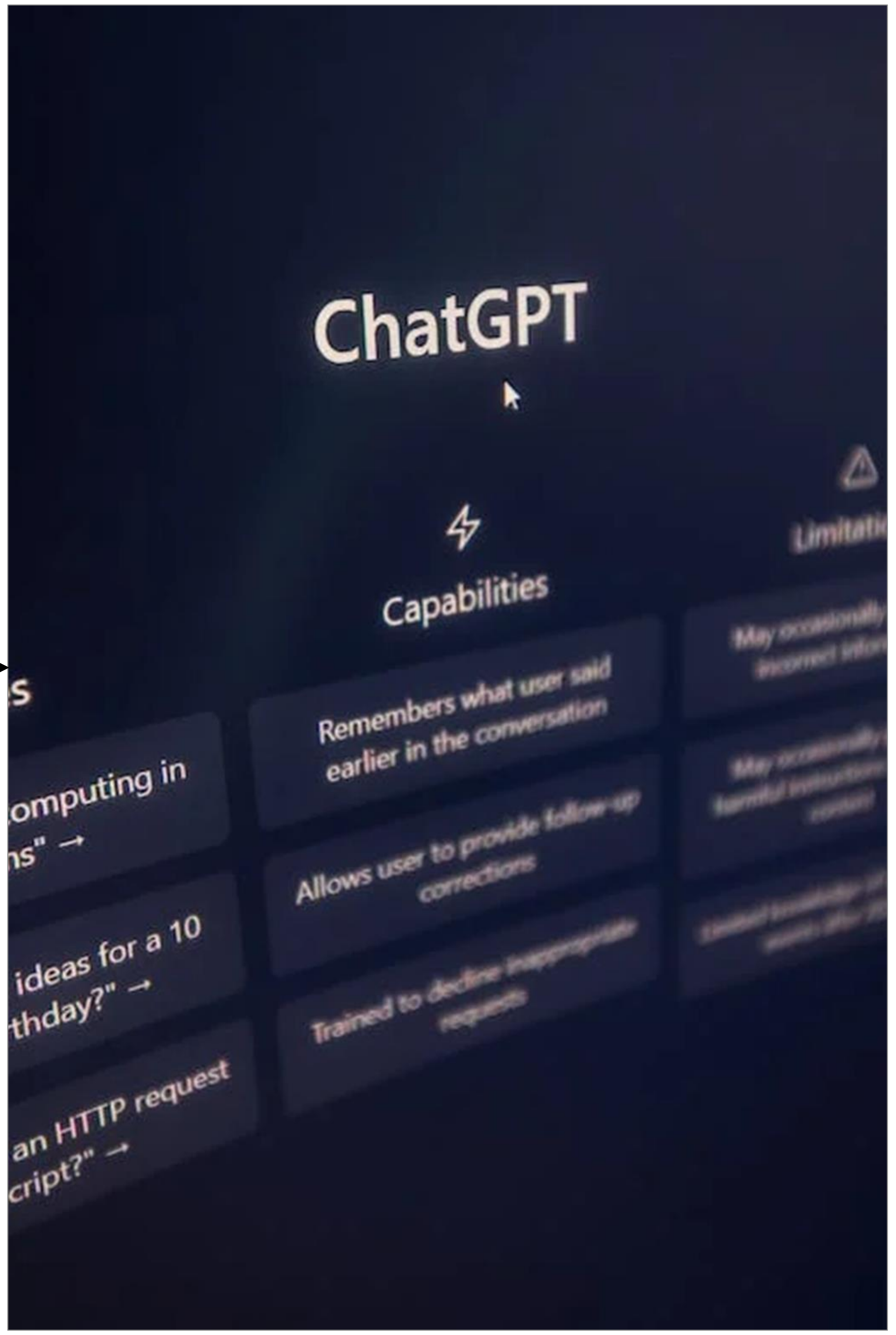
Single GPU

Latent Diffusion model with NLP on text prompts via Transformers



Transformer Architecture

Unlabeled Datasets



ChatGPT

Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests

omputing in
ns" →

ideas for a 10
thday?" →

an HTTP request
cript?" →

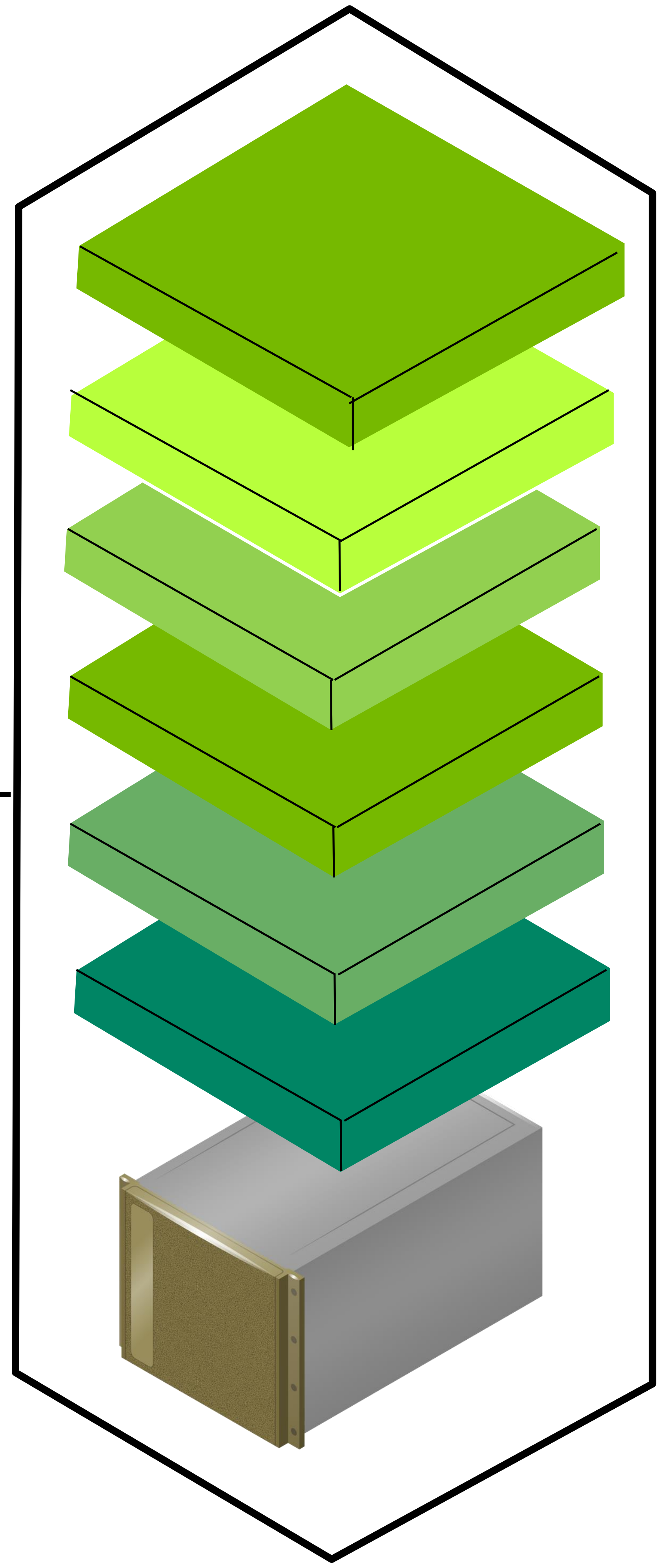
LLM

GPT-3/GPT-4

10,000 GPU Clusters

DGX Systems: Enterprise-class, Full-stack Validated Solutions

Backed by direct access to full-stack AI expertise found nowhere else



The Best of NVIDIA Software – Included with DGX
Pre-trained models, optimized frameworks, data science tools, infrastructure management tools and more, with full support from NVIDIA

NVIDIA Backed
Even ”supported systems” might not be backed by people who know AI intimately – ours do

Single Point of Contact
From framework to libraries to drivers to network, storage and compute – we’re one stop for answers and uptime

Fastest Path to Resolution
Unlike commodity servers offering commodity support - AI problems resolve quicker with direct access to NVIDIA Enterprise Support

Direct Relationship with DGXperts
NVIDIA Enterprise Support gets you a DGXpert who can offer valuable advice now, not later

Insider Access
Get orientation and on-boarding classes + DGX customer exclusive sessions taught by experts in the field

DGX-Ready Software

Operationalize AI development at scale

- For organizations that need to:
- Improve data science productivity and speed workflow with MLOps
- Simplify deployment and maximize utilization of DGX infrastructure
- Scale projects easily
- Accelerate the ROI of AI

ALTAIR
backend AI
Canonical
CLEAR ML
D2 IQ
data iku
DOMINO
Determined AI
IBM
iguazio
Pachyderm
Paperspace
Red Hat
run: ai
slurm SchedMD
SHAKUDO
RANCHER BY SUSE
Weights & Biases

nvidia.com/dgx-mlops

DGX SuperPOD Storage Vendor Support

NVIDIA-validated to ensure maximum application performance

- Validation against a variety of workloads and their corresponding typical datasets
- Validation of storage and access of model data for training of deep learning models
- Validation of storage architecture and design to ensure a balance performance, capacity, and cost
- Tier-1 support provided by NVIDIA Enterprise Support (NVES)

ddn

DELL Technologies

IBM®

NetApp®

VAST

WEKA

Announcing GB200 NVL72

Delivers New Unit of Compute

GB200 NVL72

- 36 GRACE CPUs
- 72 BLACKWELL GPUs
- Fully Connected NVLink Switch Rack

Training FP8	720 PFLOPs
Inference FP4	1,440 PFLOPs
NVL Model Size	27T params
Multi-Node All-to-All	130 TB/s
Multi-Node All-Reduce	260 TB/s

ARM Support Across Platform

Current Status

Simulation

Data Analytics

AI

HPC SDK

COMPILERS

- NVC++
- NVC
- NVFORTRAN
- NVCC

CORE C++ LIBRARIES

- CUB
- Thrust
- libcu++

PROFILING AND DIAGNOSTICS

- CUDA-gdb
- Nsight Compute
- Nsight Systems
- CUPTI

COMMUNICATION LIBRARIES

- MPI
- **HPC-X**
- OpenSHMEM
- UCX
- SHARP
- HCOLL
- NCCL
- NVSHMEM

RAPIDS

- cuDF
- cuxfilter
- cuSpatial
- cuSignal
- cuGRAPH
- cuML

DL INFERENCE

- Triton Inference Server
- TensorRT

TRAINING (DLFW)

- Pytorch
- TensorFlow

DL LIBRARIES AND SOFTWARE

- DALI
- cuDNN
- CUTLASS

MATH LIBRARIES

- cuBLAS
- cuTENSOR
- cuSOLVER
- cuSPARSE
- cuFFT
- cuRAND
- MATH API
- (OpenBLAS)

CUDA Runtime

Platform Software

- Fabric Manager
- GPUDirect
- NVML
- NVIDIA Linux GPU Driver for ARM
- DCGM
- (Slurm)
- (Docker)
- (Singularity)

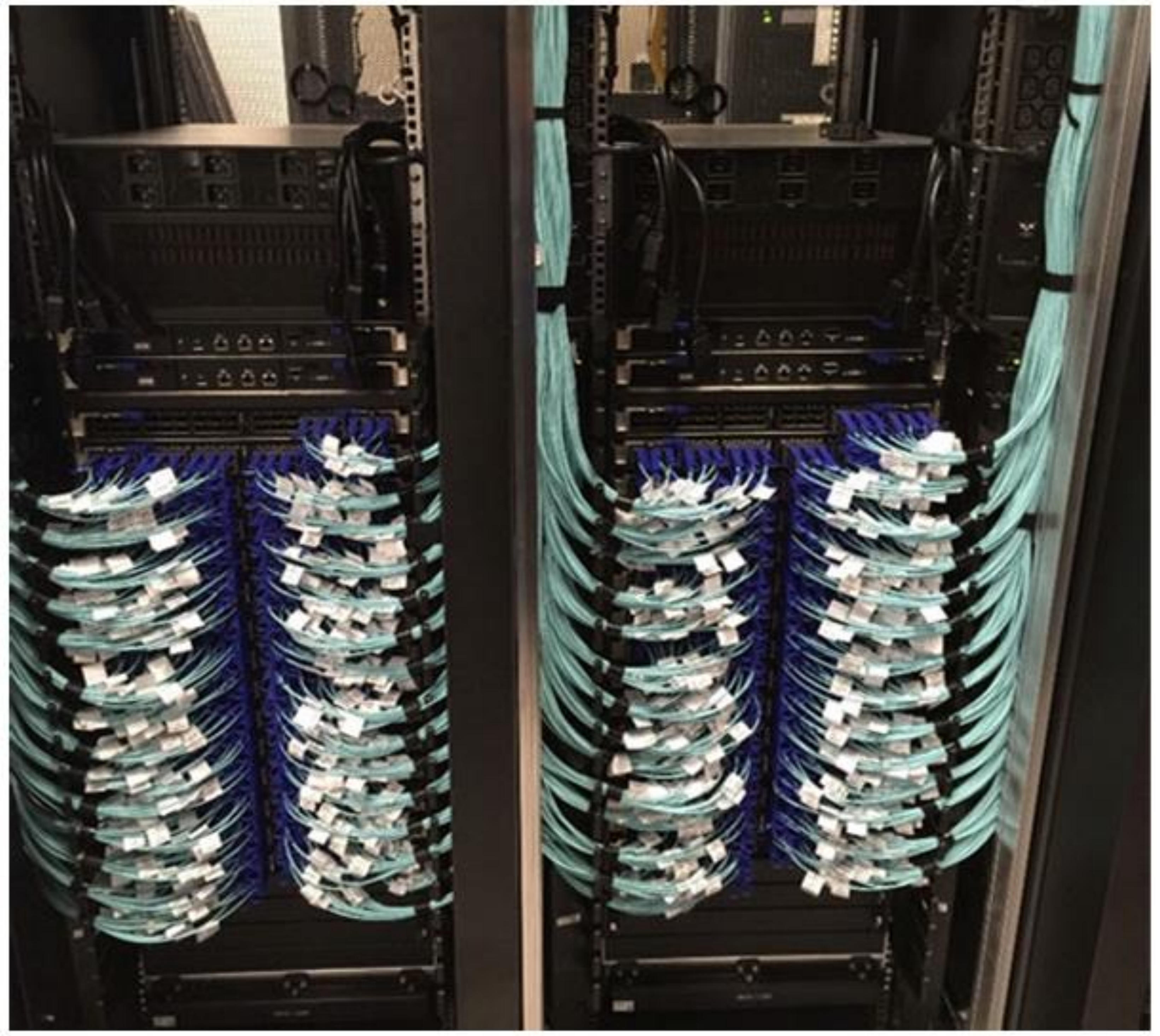
(Third party)

NVIDIA supported

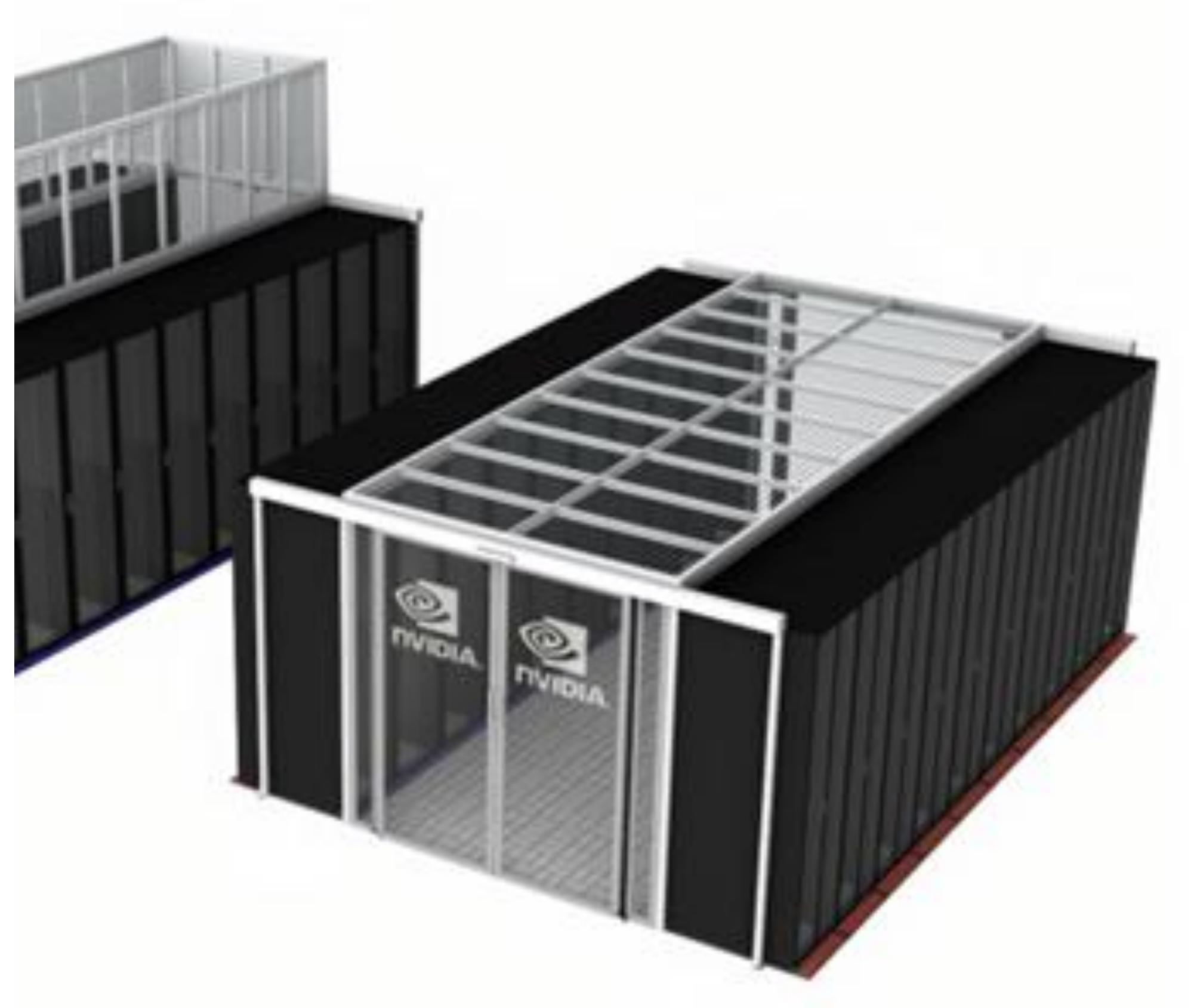
Work in Progress

Planning Power and Cooling per Rack

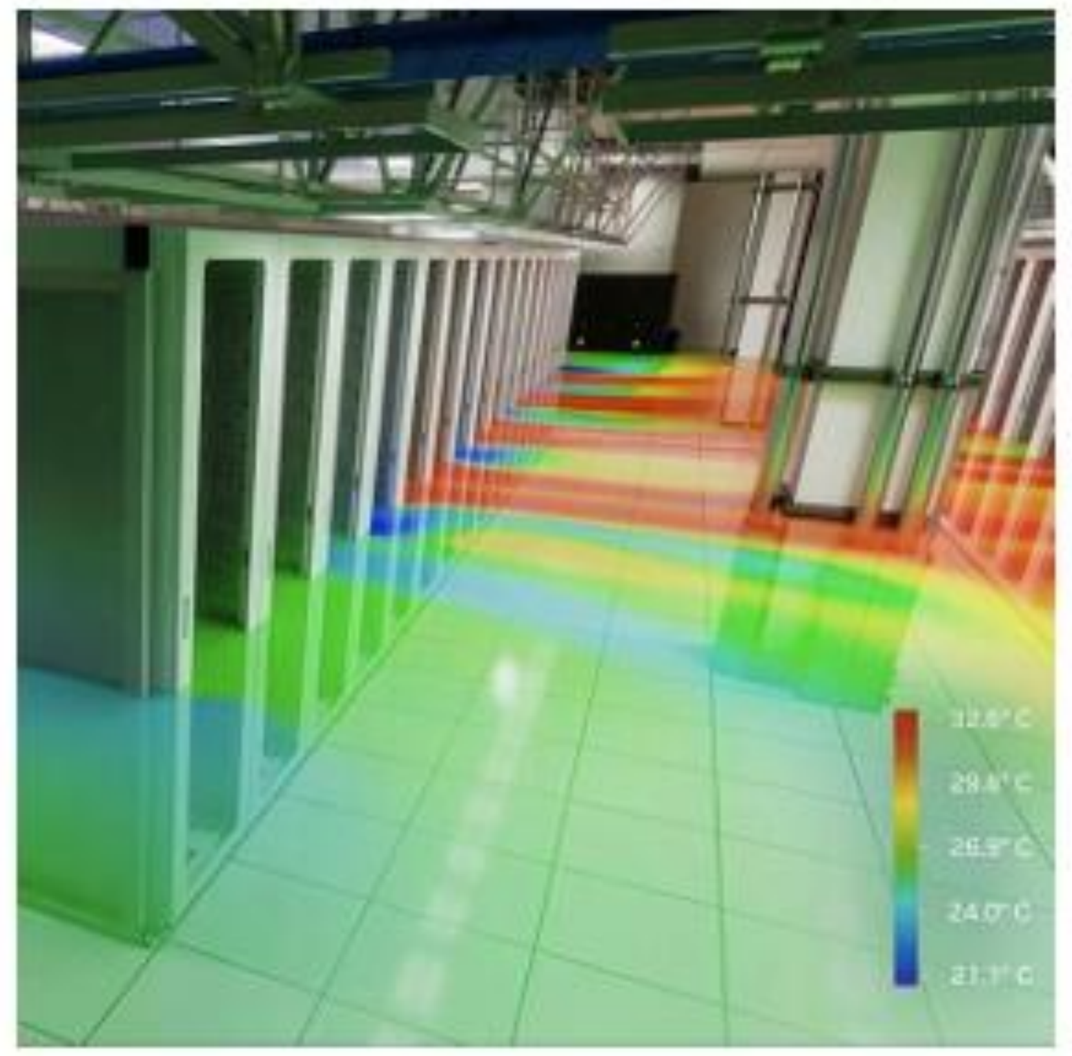
Increasing Complex Datacenter Design for Exponential Growing Computation Requirements



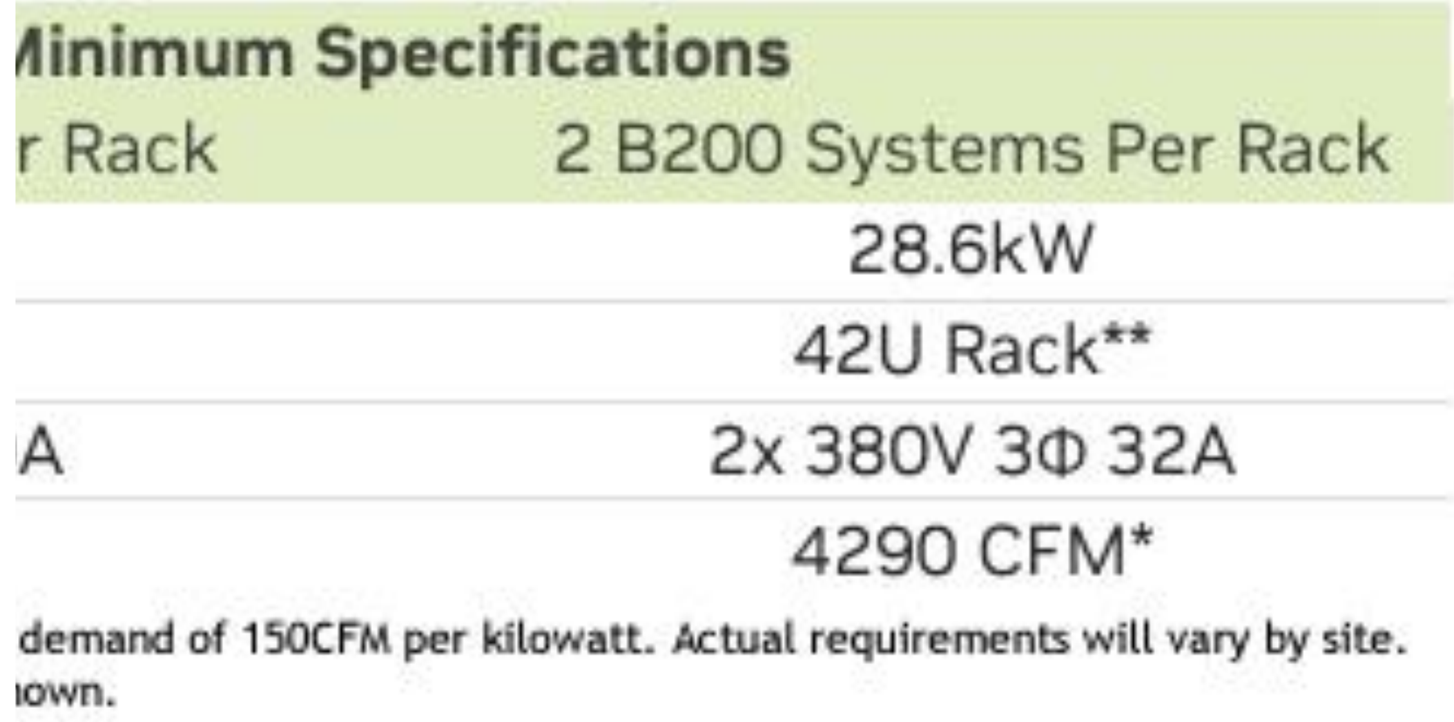
Example of proper cable management



Aisle Containment Systems



Computational Fluid Dynamics



linimum Specifications	
r Rack	2 B200 Systems Per Rack
	28.6kW
	42U Rack**
A	2x 380V 3Φ 32A
	4290 CFM*

demand of 150CFM per kilowatt. Actual requirements will vary by site.
iown.

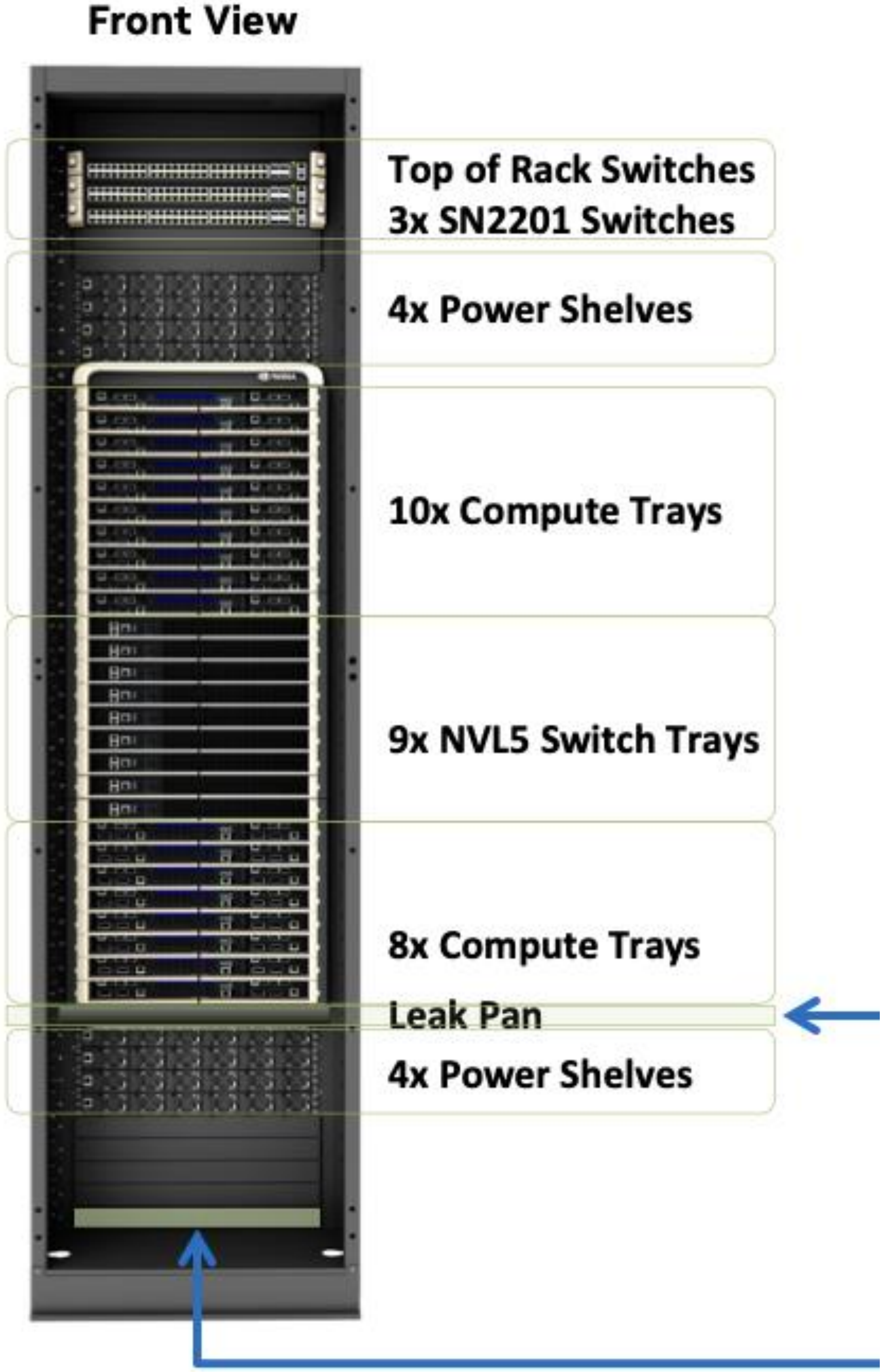
Blanking Panels

Brush Grommet

Planning Power and Cooling per Rack

Increasing Complex Datacenter Design for Exponential Growing Computation Requirements

- DGX GB200 Liquid-cooled, Datacenter-scale systems**
  - Scalable Units are 576 GPUs / 8 Racks
  - Storage
  - Networking
  - Management
- Complex Design and Operations requirement
  - Water Cooling Loops
  - Air Cooling Facilities
  - Networking Facilities
  - Etc.



288 GPU DGX GB200 SuperPOD

576 GPU DGX GB200 SuperPOD

1152 GPU DGX GB200 SuperPOD

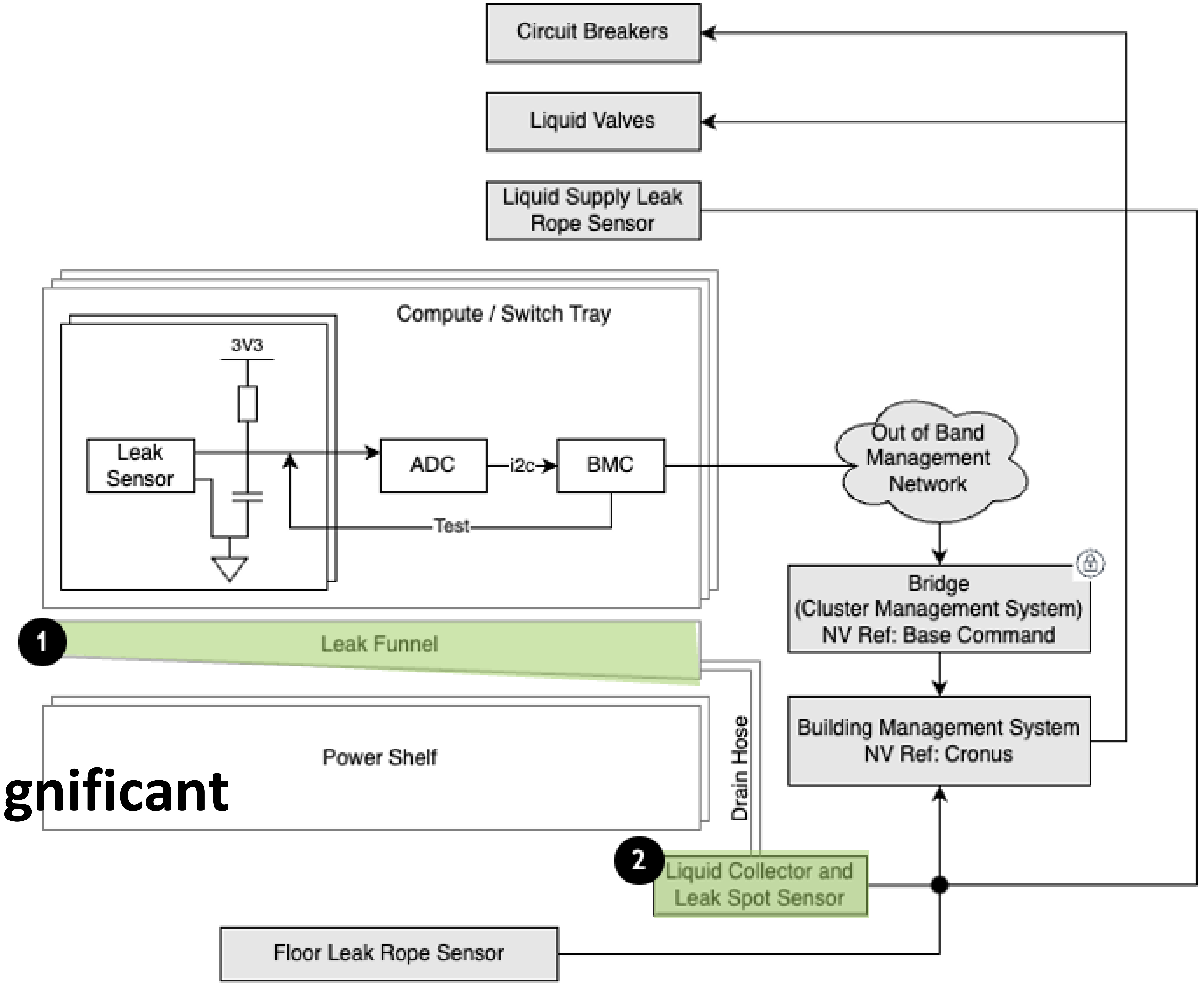
**Subject to change

Example: Leak Detection

Q: How to detect, report, and automatically react on leak event at node/rack/datacenter level?

Tasks:

- Report
- Close Valves
- Rescue Existing Jobs
- Damage Report
- (repair)
- Performance Validation
- Resume Jobs


Circuit Breakers
Liquid Valves
Liquid Supply Leak Rope Sensor
Compute / Switch Tray
3V3
Leak Sensor
ADC
i2c
BMC
Test
Out of Band Management Network
Bridge (Cluster Management System) NV Ref: Base Command
1
Leak Funnel
Power Shelf
Drain Hose
Building Management System NV Ref: Cronus
2
Liquid Collector and Leak Spot Sensor
Floor Leak Rope Sensor

Can we achieve these tasks without significant downtime?

Example: Complex Networking Components

400GbE Leaf/Spine/Super Spine Ethernet Switch
64 OSFP connectors (800G/connector)
2 400GbE ports per connector
Total of 128 400GbE ports. “Moose”

OSFP
Finned Top

Direct Attach Copper Cable
800G to 800G
OSFP to OSFP
MCP4Y10-N00A (0.5m) 30 AWG
MCP4Y10-N001 (1m) 30 AWG
MCP4Y10-N002 (2m) 26 AWG
Ultron

Twin port 800G Multi Mode Transceiver
MMA4Z00-NS (50m)
OSFP Finned-top
Dual MPO/APC
Louie

Multi-mode fibers:
MFP7E10-N0XX
(XX = 03-to-50) meters

400GbE Leaf/Spine/Super Spine Ethernet Switch
64 OSFP connectors (800G/connector)
2 400GbE ports per connector
Total of 128 400GbE ports. “Moose”

OSFP
Finned Top

Twin port 800G Mu
Mode Transceiver
MMA4Z00-NS (50m
OSFP Finned-top
Dual MPO/APC

Louie

400GbE/200GbE/NDR BlueFied3 3240 DPU
2x QSFP112 Connector (up to 400G/Connector)

200GbE/ ConnectX-7 HCA/NIC
2x QSFP112 Connector (up to 200G/Connecto

Single Port NDR
Multimode 400G QSFP112 Transceiver
QSFP112/Flat top
MMA1Z00-NS400 (50m)
Single MPO/APC
QLouie

ConnectX-7 400G HCA OSFP
2x OSFP Connector (up to 400G/Connector)

Single port 400G Transceiver – Multi Mode
MMA4Z00-NS400 (50m)
OSFP Flat-top
Single MPO/APC
Louie 400

OSFP
Finned Top

Direct Attach Cable
2x 200G to 4x 100G
OSFP to 4x QSFP56
MCP7Y70-H001 (1m)
MCP7Y70-H01A (1.5m)
MCP7Y70-H002 (2m)

Hulk BC

QSFP56

ConnectX-6
QSFP56

OSFP
Finned Top

Up to 100 meters

Up to 500-meters

Twin port 800G Single mode Transceiver
MMS4X00-NS (100m)
MMS4X00-NM (500m)
OSFP Finned-top
Dual MPO/APC
Bagheera

100G DR1 Single mode Transceiver
MMS1V70-CM *(500m)*
QSFP28 Flat-top
Single LC
1x100G-PAM4 optical to 4x25G-NRZ electrical

1x100G-PAM4 optical

Fiber splitter cable not supplied by NVIDIA.
Reference mgrs can be provided.
Reach is limited to 100-meters maximum of
Twin-port 100-meter 800G transceiver
500-meters with only -NM transceiver

QSFP28
4x25G-NRZ

Customer Edge
Network

1GbE /100GbE TOR/Leaf/OOB Ethernet Switch
4x QSFP28 Connector (100G/connector)
48x RJ45 Connector (1GbE/connector)

Alternative Solution:
100GbE QSFP28 CWDM4
MMA1L30-CM *(2KM)*
QSFP28 Flat-top
single mode 4- channel (CWDM4)

Ethernet Fabric – DGX GB200 Single SU

Customer Edge Network
SN5600 Spine
SN5600 Spine
SN5600 Leaf
SN2201 Leaf
C-UFM-2
C-UFM-1
NFS Storage
...
Management
High Performance Storage
PDUs, iDRAC, iLO, etc. etc.
OOB Connections, BF3, COMe
Out-of-Band SN2201
GB200-Rack-5
GB200-Rack-6
GB200-Rack-7
GB200-Rack-8
GB200-Rack-3
GB200-Rack-4
GB200-Rack-1
GB200-Rack-2
1x 100 Gbps
6x 800 Gbps
4x 800 Gbps
2x 100 Gbps
Varies
1 x 1 Gbps
1x 200 Gbps

Ethernet Fabric - DGX GB200 Single SU

Customer Edge Network
SN5600 Spine
SN5600 Spine
SN5600 Leaf
SN2201 Leaf
C-UFM-2
C-UFM-1
NFS Storage
Flapping Link
...
Management
High Performance Storage
Out-of-Band SN2201
PDUs, iDRAC, iLO, etc. etc.
OOB Connections, BF3, COMe
GB200-Rack-5
GB200-Rack-6
GB200-Rack-7
GB200-Rack-8
GB200-Rack-3
GB200-Rack-4
GB200-Rack-1
GB200-Rack-2
1x 100 Gbps
6x 800 Gbps
4x 800 Gbps
2x 100 Gbps
Varies
1 x 1 Gbps
1x 200 Gbps

Ethernet Fabric - DGX GB200 Single SU

Customer Edge Network
SN5600 Spine
SN5600 Spine
SN5600 Leaf
SN2201 Leaf
Flapping Link
MPI Timeout
C-UFM-2
C-UFM-1
NFS Storage
...
Management
High Performance Storage
PDUs, iDRAC, iLO, etc. etc.
OOB Connections, BF3, COMe
Out-of-Band SN2201
GB200-Rack-5
GB200-Rack-6
GB200-Rack-7
GB200-Rack-8
GB200-Rack-3
GB200-Rack-4
GB200-Rack-1
GB200-Rack-2
1x 100 Gbps
6x 800 Gbps
4x 800 Gbps
2x 100 Gbps
Varies
1 x 1 Gbps
1x 200 Gbps

Ethernet Fabric - DGX GB200 Single SU

Customer Edge Network
SN5600 Spine
SN5600 Spine
SN5600 Leaf
SN2201 Leaf
Flapping Link
MPI Timeout
Job Failure
C-UFM-2
C-UFM-1
NFS Storage
Management
High Performance Storage
PDUs, iDRAC, iLO, etc. etc.
OOB Connections, BF3, COMe
Out-of-Band SN2201
GB200-Rack-5
GB200-Rack-6
GB200-Rack-7
GB200-Rack-8
GB200-Rack-3
GB200-Rack-4
GB200-Rack-1
GB200-Rack-2
1x 100 Gbps
6x 800 Gbps
4x 800 Gbps
2x 100 Gbps
Varies
1 x 1 Gbps
1x 200 Gbps

Ethernet Fabric - DGX GB200 Single SU

Flapping Link → MPI Timeout → Job Failure → Lost Money

Customer Edge Network
SN5600 Spine
SN5600 Spine
SN5600 Leaf
SN2201 Leaf
NFS Storage
C-UFM-2
C-UFM-1
Management
High Performance Storage
Out-of-Band SN2201
PDUs, iDRAC, iLO, etc. etc.
OOB Connections, BF3, COMe
GB200-Rack-5
GB200-Rack-6
GB200-Rack-7
GB200-Rack-8
GB200-Rack-3
GB200-Rack-4
GB200-Rack-1
GB200-Rack-2
1x 100 Gbps
6x 800 Gbps
4x 800 Gbps
2x 100 Gbps
Varies
1 x 1 Gbps
1x 200 Gbps

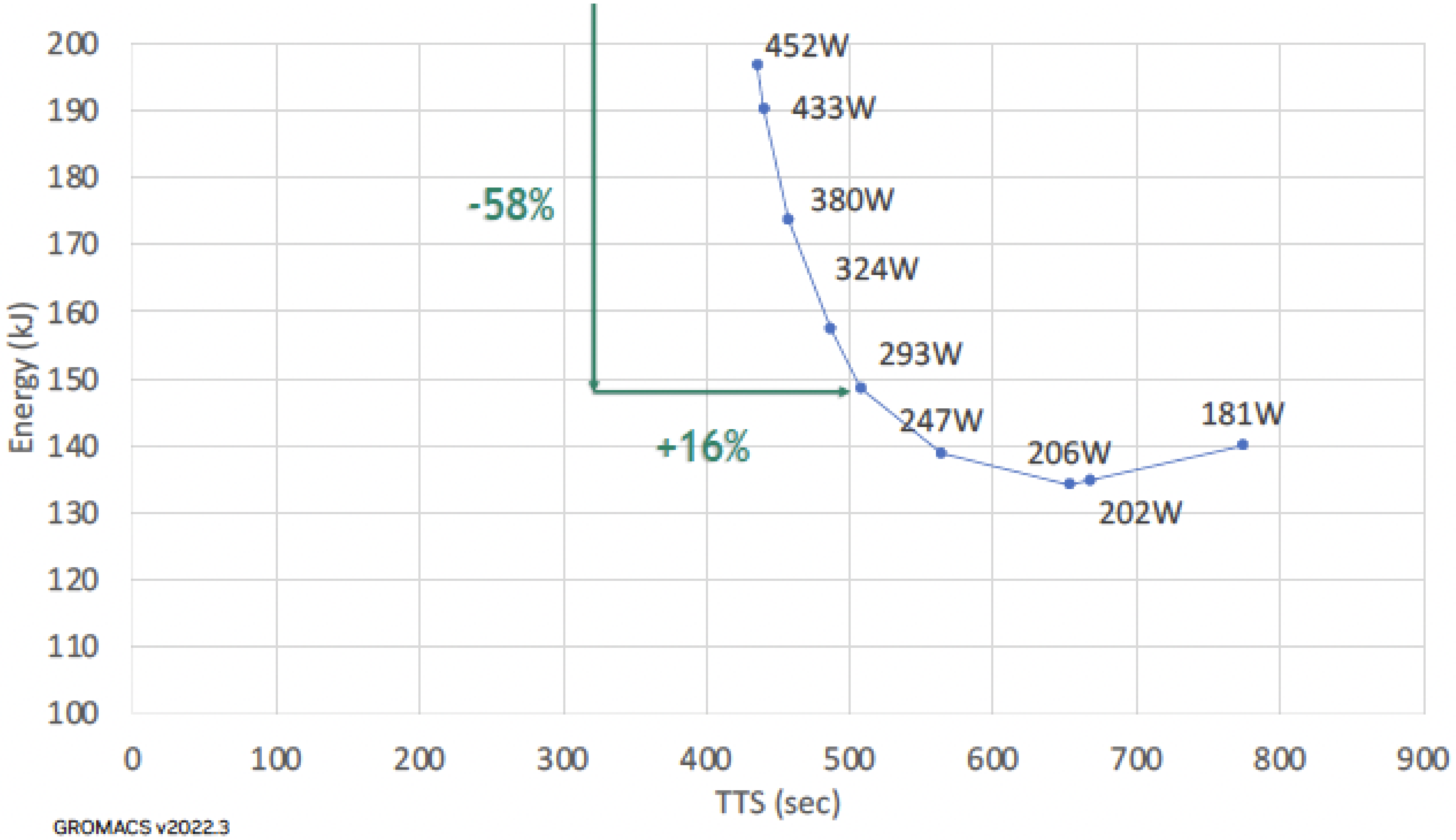
Ethernet Fabric - DGX GB200 Single SU

Customer Edge Network
SN5600 Spine
SN5600 Spine
SN5600 Leaf
SN2201 Leaf
NFS Storage
C-UFM-2
C-UFM-1
...
Management
High Performance Storage
PDUs, iDRAC, iLO, etc. etc.
OOB Connections, BF3, COMe
Out-of-Band SN2201
1x 100 Gbps
6x 800 Gbps
4x 800 Gbps
2x 100 Gbps
Varies
1 x 1 Gbps
1x 200 Gbps
GB200-Rack-5
GB200-Rack-6
GB200-Rack-7
GB200-Rack-8
GB200-Rack-3
GB200-Rack-4
GB200-Rack-1
GB200-Rack-2

Flapping Link → MPI Timeout → Job Failure → Lost Money

Can we reduce application downtime?

Example: Finding Optimal Power Configuration



GROMACS v2022.3
STMV, 1000K simulation steps
Single pre-production H100-SXM5 with 132 SMs and HBM3, with single-socket 16-core Intel Xeon Silver 4314 CPU (h100-sxm5-hbm3-preprod partition on computelab).

- When running at max frequency on H100, GROMACS only draws 452W on average.
- It is possible to enforce power limit of 500W without any effect on performance.
- 400W power limit only reduces performance by 4%, and average power goes to 390W.

Can my cluster select the best power profile for me to reduce energy cost?

DGX Software Stack **

HW + OS + NV AI Enterprise + Datacenter Management → DGX B200 SPOD or DGX GB200 SPOD

- DGX OS 6 / 7
- DGX Firmware
- Cumulus OS
- Quantum / Spectrum Firmware

- NIMs
- Microservices
- CUDA-X
- Base Command Manager

- UFM
- PaaS – Run: Ai, DGX Ready Software Partners
- Datacenter Observability and Operations
- Resiliency
- Power Optimizations
- Full stack upgrade

Grey Areas: Increasing Interest in today’s AI Datacenter

** Subject to change

Current NVIDIA Effort

Full-stack resiliency and efficiency

Application

Scheduler

Networking

Compute

Slurm/ Kubernetes Job Queues

Fragmented Error Reporting

Manual Error attribution

Training or Inference

Application

Scheduler

Networking

Compute

Workload orchestration with Fast checkpoint and Restart

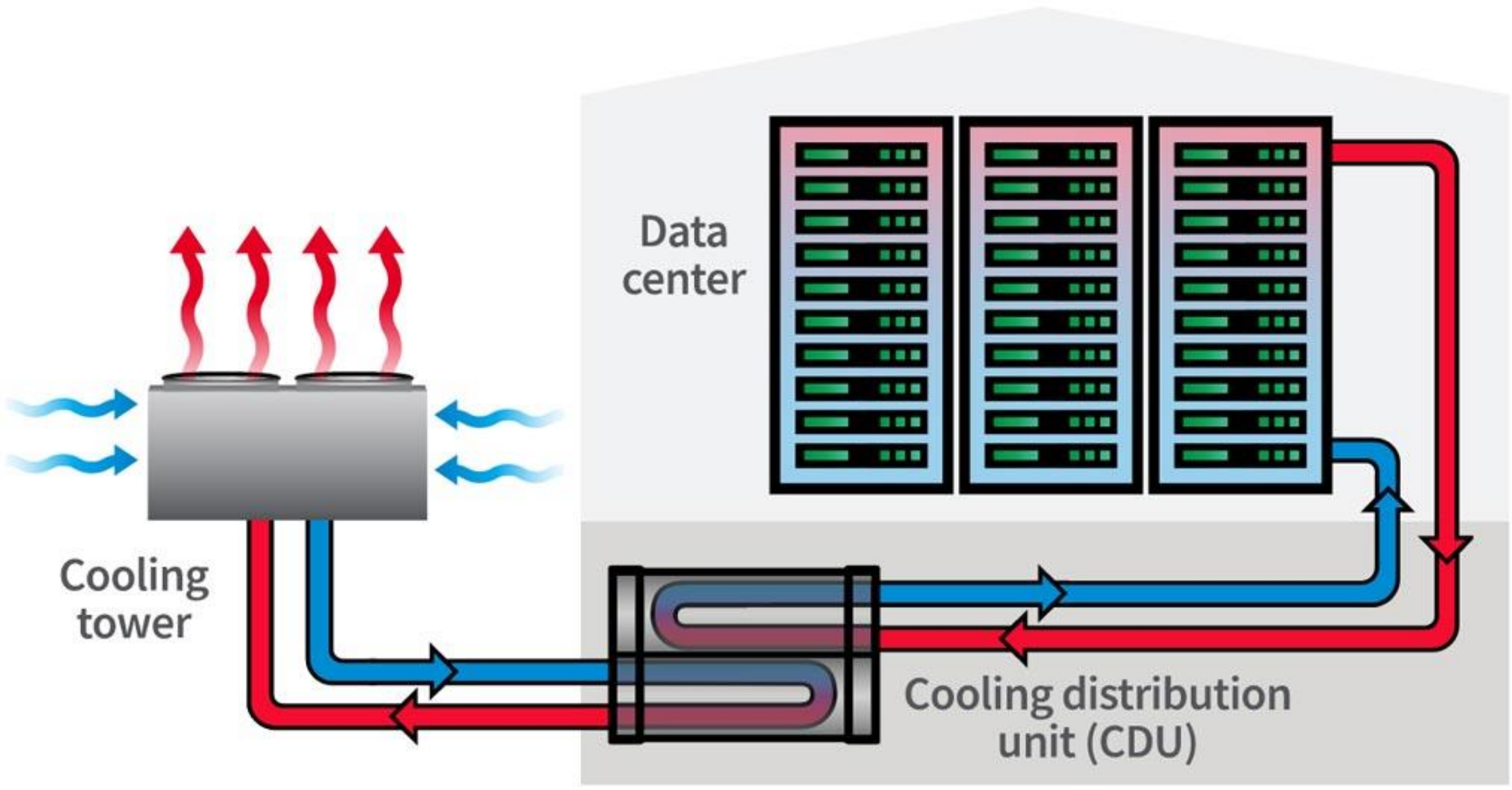
Advanced Cluster Information Autoscaling, Topology Manager, etc.

Unified Error collection & Reporting

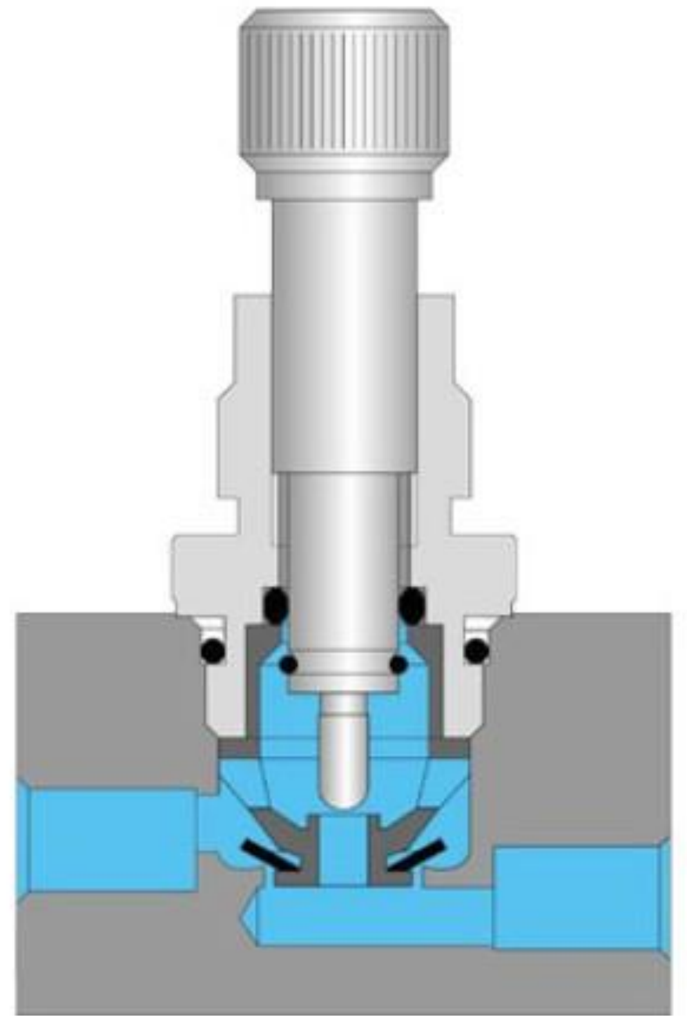
Automated Error Attribution, Containment, & Anomaly Isolation Power Optimization

API-Driven, Partner-Driven Datacenter Integration

NVIDIA DGX SuperPOD Datacenter Design Guidelines



Coolant Distribution Unit (CDUs)

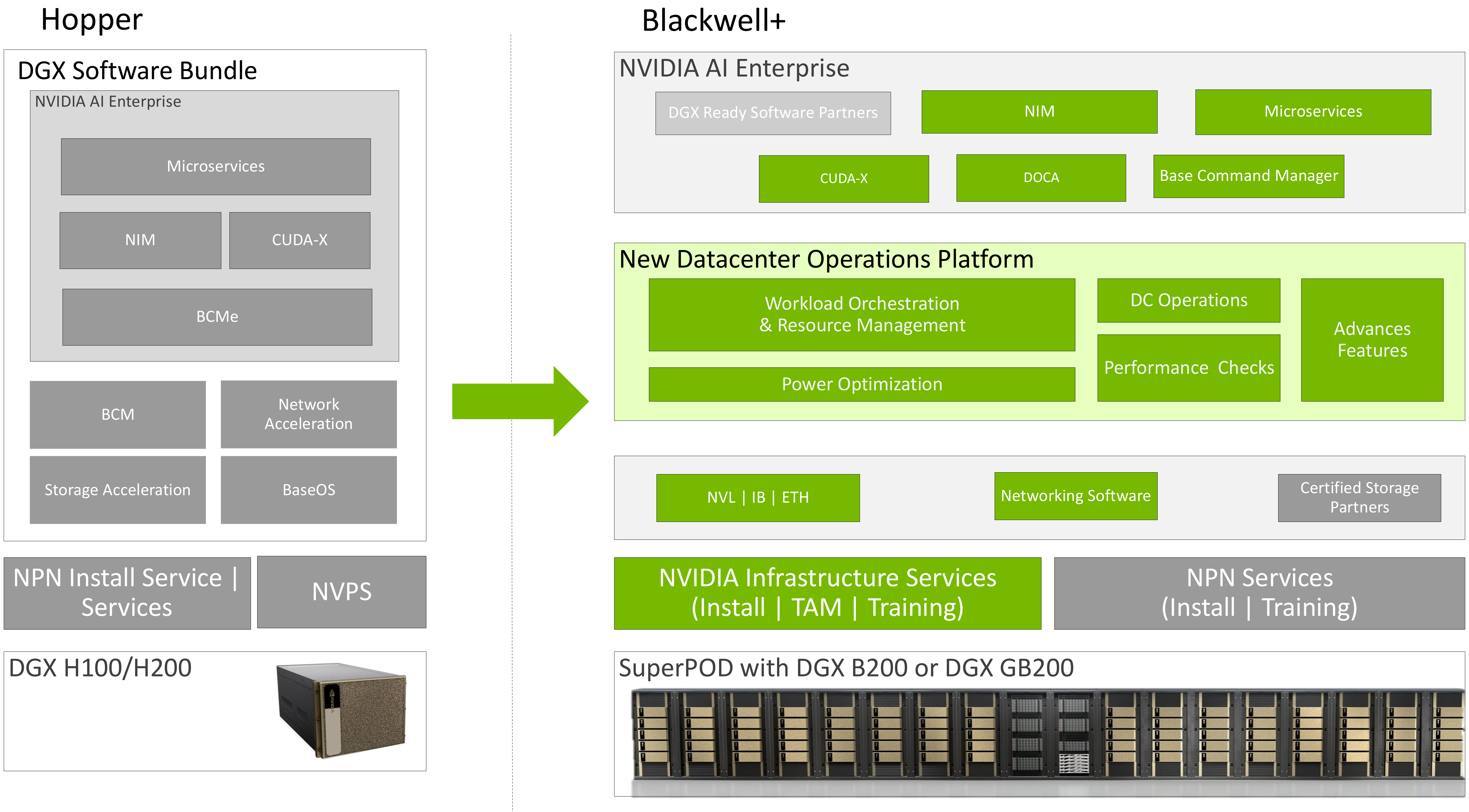


Flow Control Valves

Circuit Breakers

NEW DGX SOFTWARE STACK

Developer productivity software that augments best in class infrastructure



Hopper

DGX Software Bundle

NVIDIA AI Enterprise

- Microservices
- NIM
- CUDA-X
- BCMe

- BCM
- Network Acceleration
- Storage Acceleration
- BaseOS

NPN Install Service | Services

NVPS

DGX H100/H200

Blackwell+

NVIDIA AI Enterprise

- DGX Ready Software Partners
- NIM
- Microservices
- CUDA-X
- DOCA
- Base Command Manager

New Datacenter Operations Platform

- Workload Orchestration & Resource Management
- Power Optimization
- DC Operations
- Performance Checks
- Advances Features

- NVL | IB | ETH
- Networking Software
- Certified Storage Partners

NVIDIA Infrastructure Services (Install | TAM | Training)

NPN Services (Install | Training)

SuperPOD with DGX B200 or DGX GB200

NVIDIA®