AI/ML computations on SDSC’s Voyager

Presented by Paul Rodriguez and Mahidhar Tatineni

Fifth National Research Platform (5NRP) Workshop

March 19, 2024

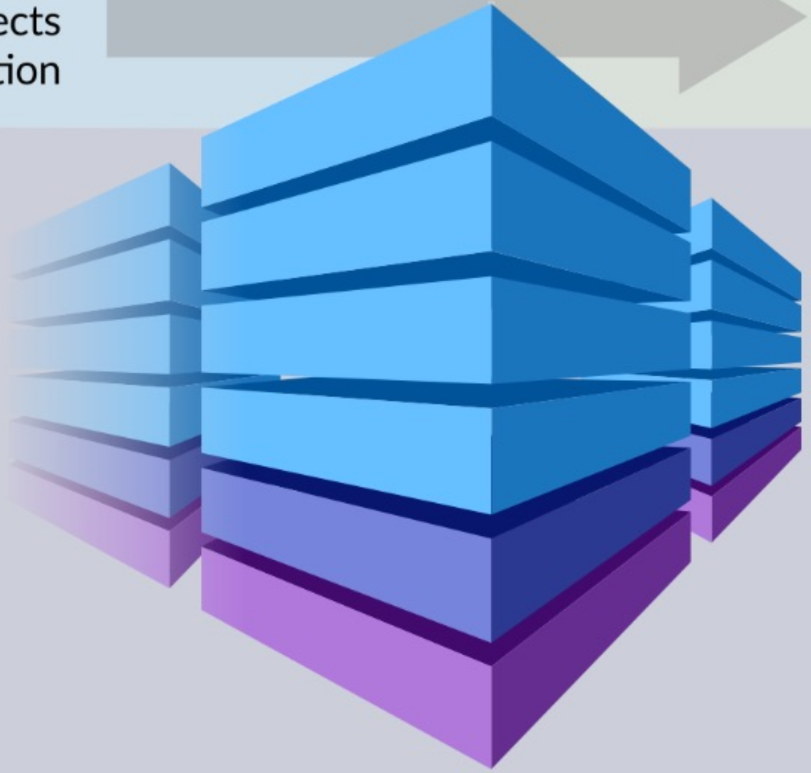
University of California San Diego & San Diego Supercomputer Center

VOYAGER

EXPLORING AI PROCESSORS in SCIENCE and ENGINEERING

3-YEAR TESTBED PHASE
Focused Select Projects
Workshops, Industry Interaction

2-YEAR ALLOCATIONS PHASE
NSF Allocations to the Broader Community
User Workshops



INNOVATIVE AI RESOURCE
Specialized Training Processors
Specialized Inference Processors
High-Performance Interconnect
X86 Standard Compute nodes
Rich Storage Hierarchy

OPTIMIZED AI SOFTWARE
Community Frameworks
Custom user-developed AI Applications
PyTorch, Tensorflow

IMPACT & ENGAGEMENT
Large-Scale Models
AI Architecture Advancement
Improved Performance of AI Applications
External Advisory Board of AI & HPC Experts
Wide Science & Engineering Community
Advanced Project Support & Training
Accelerating Scientific Discovery
Industrial Engagement

Category II System, NSF Award # 2005369

PI: Amit Majumdar (SDSC); Co PIs: Rommie Amaro (UCSD), Javier Duarte (UCSD), Mai Nguyen (SDSC), Robert Sinkovits (UCSD)

The project is structured in two phases: a 3-year testbed, followed by a 2-year allocations phase

- **Testbed Phase – started May 2022 (after successful NSF review in April 2022)**
  - Work closely with select research groups – deep user engagement
  - Evaluate *Voyager's* innovative DL hardware, software, libraries, ML application porting/performance
  - Semiannual workshops, user forums to share lessons learned, bring researchers together
  - Develop knowledge base, best use cases for future users, allocation policies
  - External Advisory Board to help recruit research groups, provide guidance to project

- Allocations Phase
  - Allocate via NSF-approved process
  - Lessons learned from Testbed phase inform documentation and training
  - Regular and advanced user support
  - Semiannual workshops
  - Industry engagement for similar technology evaluation

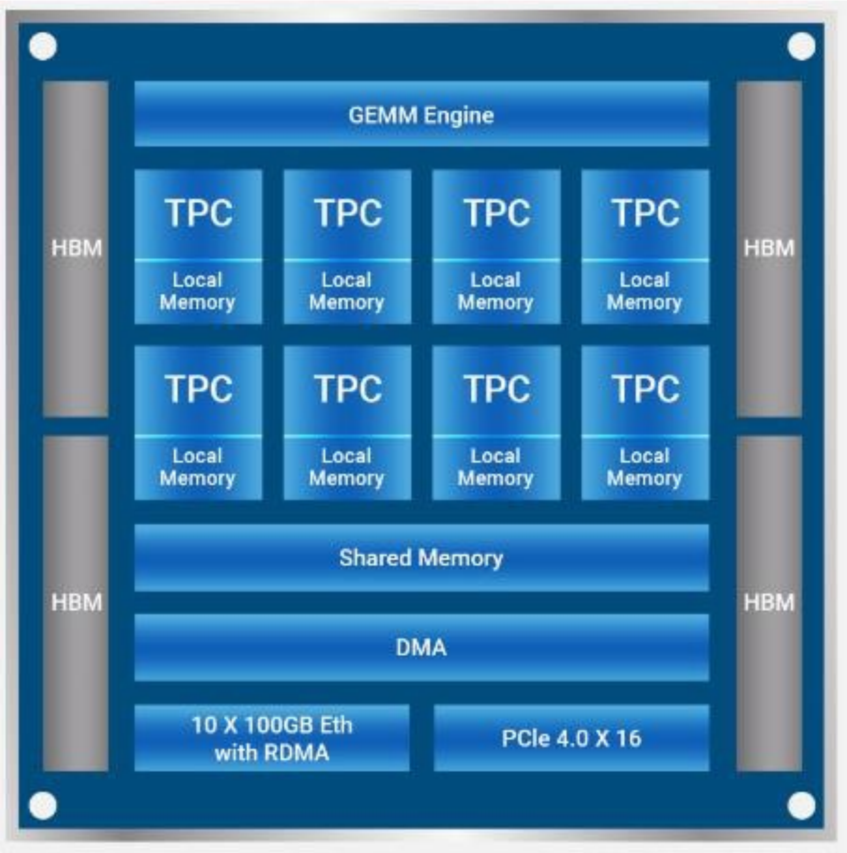
Voyager is a heterogeneous system designed to support complex AI workflow

- **42x Intel Habana Gaudi** training nodes, each with 8 training processors (**336 in total**); all-to-all network between processors on a node
- Gaudi processors feature specialized hardware units for AI, HBM2, and on-chip high-speed Ethernet
- **2x first generation inference nodes**, each with 8 inference processors (**16 in total**)
- **36x Intel x86 processors compute nodes** for general purpose computing and data processing
- **400 GbE interconnect** using RDMA over Converged Ethernet
- **3 PB Storage** system connected vis 25GbE. Deployed as Ceph, but open to others
- 324 TB HFS; connectivity to compute via 25GbE
- Machine integrated by Supermicro; includes Arista switch

System Component	*Configuration*
INTEL GAUDI TRAINING NODES	
Node count	42
Training processors/node	8
Host x86 processors/node	2
Memory/node	512 GB DDR4
Memory/training processor	32 GB HBM2
Local NVMe	6.4 TB
INTEL GOYA INFERENCE NODES	
Node count	2
Inference processors/node	8
Host x86 processors/node	2
Memory/node	512 GB DDR4
Memory/inference processor	16 GB DDR4
Local NVMe	3.2 TB
STANDARD COMPUTE NODES	
Node count	36
x86 processors/node	2
Memory capacity	384 GB
Local NVMe	3.2 TB
STORAGE SYSTEM	
High performance storage: HDD:NVMe	3 PB:140 TB
High performance filesystems	Ceph, Lustre
Home filesystem storage: HDD:NVMe	324 TB: 12.4 TB
File system	NFS

Gaudi: Architected for performance and efficiency

- *Fully programmable Tensor Processing Cores (TPC) with tools & libraries*
- *Configurable Matrix Math Engine (GEMM)*
- *Multi-stage memory hierarchy with 32GB HBM2 memory*
- *Integrated 10 x 100 Gigabit Ethernet for multi-chip scale-out training*

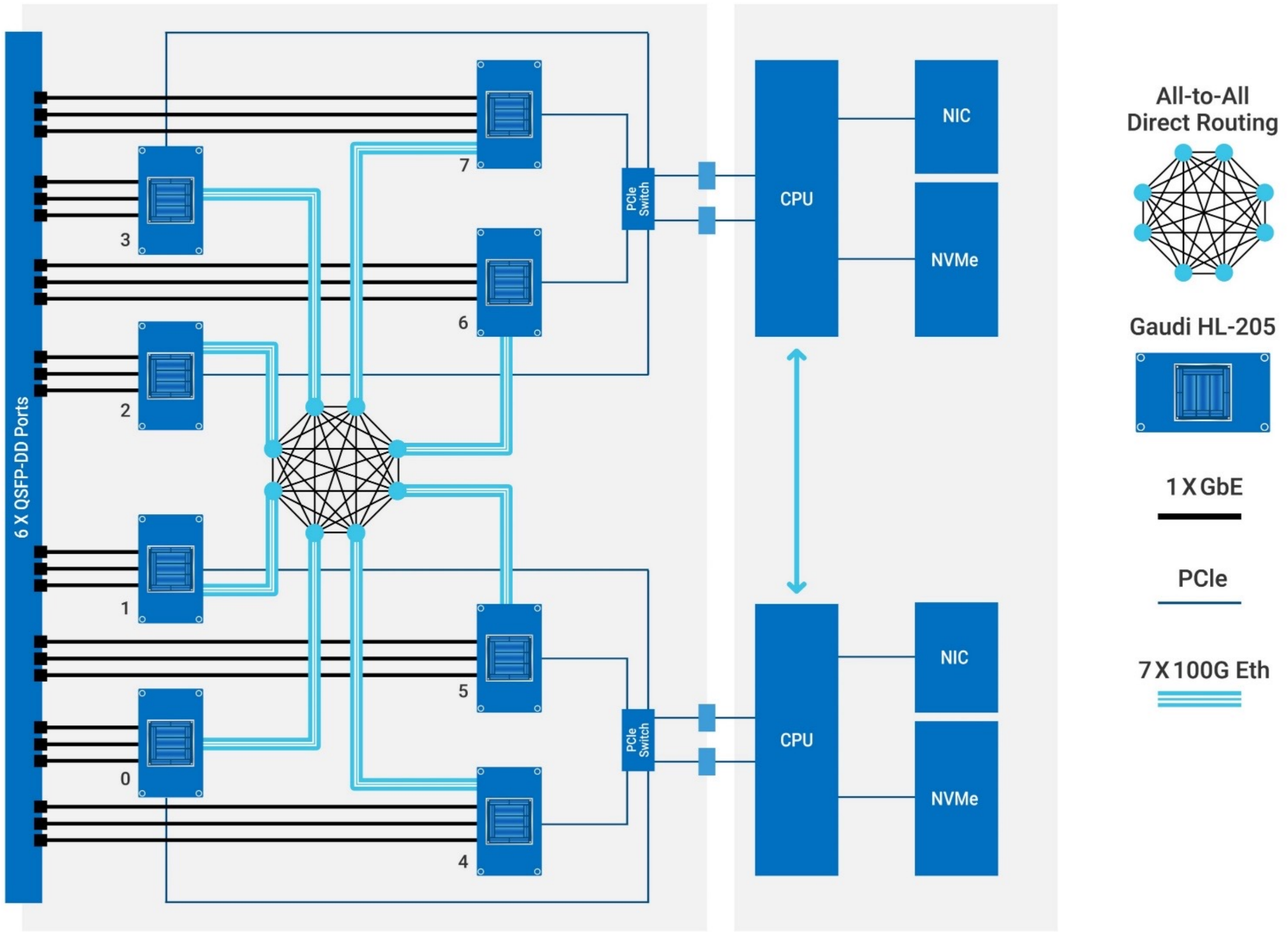


Designed for flexible and easy model migration

Ease of use	*Customization*	*Balanced compute & memory*
Integrated with TensorFlow and PyTorch; minimal code changes to get started *→ SynapseAI maps model topology onto Gaudi devices*	*SynapseAI TPC SDK facilitates development of custom kernels*	*32GB HBM2 memories similar to GPUs, so existing DL models will fit into Gaudi memory*
*Developers can enjoy the **same abstraction** they are accustomed to today*	*Developers can **customize** models to extract best performance*	*Developers can spend **less effort** to port their models to Gaudi*

Gaudi servers supports all-to-all connectivity

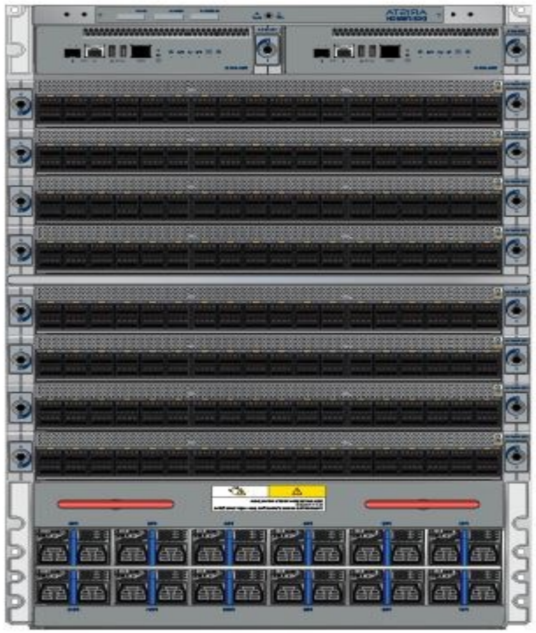
- 8 Gaudi OCP OAM cards
- 24 x 100GbE RDMA RoCE for scale-out
- Non-blocking, all-2-all internal interconnect across Gaudi AI processors
- Separate PCIe ports for external Host CPU traffic



Example of Integrated Server with eight Gaudi AI processors, two Xeon CPU and multiple Ethernet Interfaces

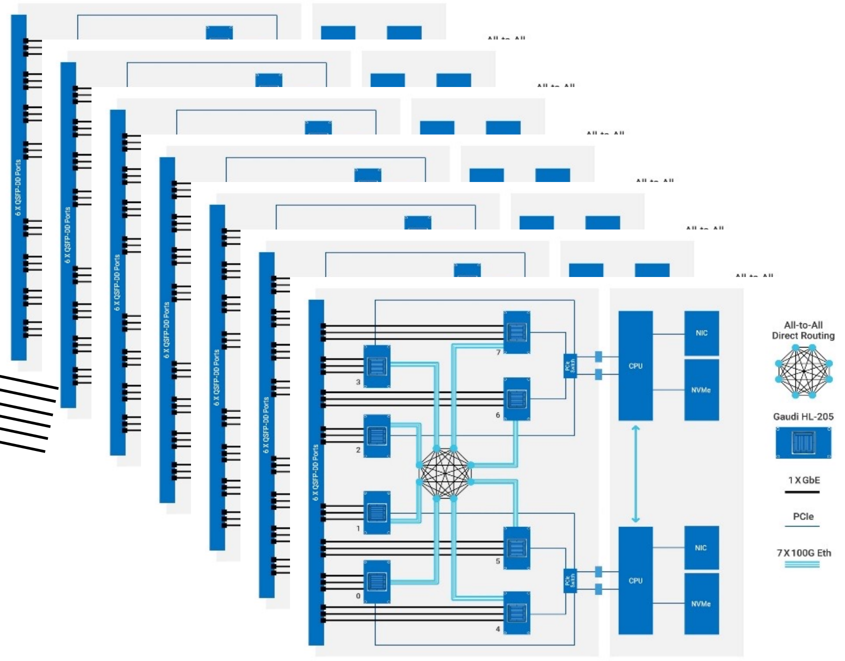
Gaudi design enables highly efficient scaling

- *Natively integrated RoCE on Gaudi processor*
- *6x Quad-100 GbE per node (8x Gaudi)*
- *7808 Arista 400 GbE switch*



7808 Arista 400GbE

6x 400GbE per Gaudi Node



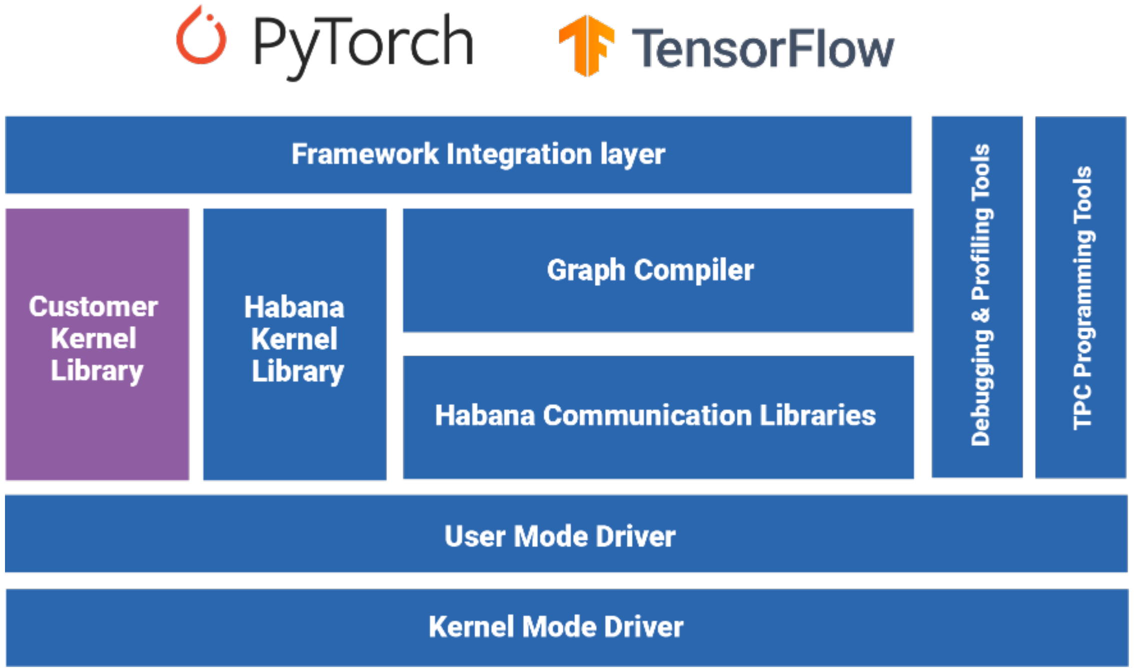
6x Gaudi nodes per rack

Voyager’s storage hierarchy supports AI workflows

	Filesystem	Capacity	Connectivity	Use cases
Node-local NVMe	XFS	3.2 TB 6.4 TB	PCEe 4.0	Large, node-local NVMe drives provide ephemeral storage and excellent performance for workload that don’t need shared data.
Home file system	NFS	324 TB	50 Gb Ethernet	High-Availability Network Files System (NFS) Cluster for user home directory storage
Project storage	Ceph; will investigate other options during Testbed (e.g. VAST being tested)	3 PB	50 Gb Ethernet	Large data, project storage

SynapseAI Software Suite: Designed for Performance and Ease of Use

SynapseAI®



- Shared software suite for training and inference
- Start running on Habana accelerators with minimal code changes
- Integrated with PyTorch and TensorFlow
- Rich library of performance-optimized kernels
- Advanced users can write their custom kernels
- Docker container images and Kubernetes orchestration
- Habana Developer Site & HabanaAI GitHub
- Habana Developer Forum

Habana Deep Learning Software Ecosystem

– Development on the latest models, more quickly, more easily