

Shaheen III HPC Training

February 4, 2025

Location: Auditorium between bldgs. 2 & 3

KAUST Supercomputing core Lab (KSL)

IaaS: (Infrastructure as-a Service)

SaaS: (Software as a Service)



PaaS: (Platform as-a Service)

CaaS: (Collaboration/Consultancy as-a Service)

KAUST Supercomputing core Lab (KSL)

JYSOO LEE,
FACILITIES DIRECTOR,
RCCL



SABER FEKI,
LEAD, SENIOR
COMPUTATIONAL SCIENTIST



BILEL HADRI

HPC SW ENV.
AND TOOLS



KADIR AKBUDAK

WEATHER
AND CLIMATE
SUPPORT



ROOH KHURRAM

CFD SERVICES



ZHIYONG ZHU

MATERIAL SCIENCE/
CHEMISTRY



MOHSIN A. SHAIKH

AI/ML



**NAGARAJAN
KATHIRESAN**

BIO-SCIENCE

Agenda:

- **8:30am** **Welcome**
- **8:35am** **Shaheen III Hardware Overview**
- **8:55am** **How to apply on Shaheen III**
- **9:05am** **Getting Started on Shaheen III**
- **9:15am** **Software Environment**
- **9:35am** **Job Scheduling**
- **10:00am** **Coffee Break**
- **10:15am** **Storage overview & Best practices**
- **10:30am** **Applications software example: VASP workflow**
- **10:50 am** **Applications software example: CFD applications**
- **11:10 am** **Applications software example: Bio informatics workflow**
- **11:20-11.30am** **Q&A and Open Discussion**

Shaheen III HPC Training

Shaheen III Hardware Overview

February 4, 2025

Location: Auditorium between bldgs. 2 & 3

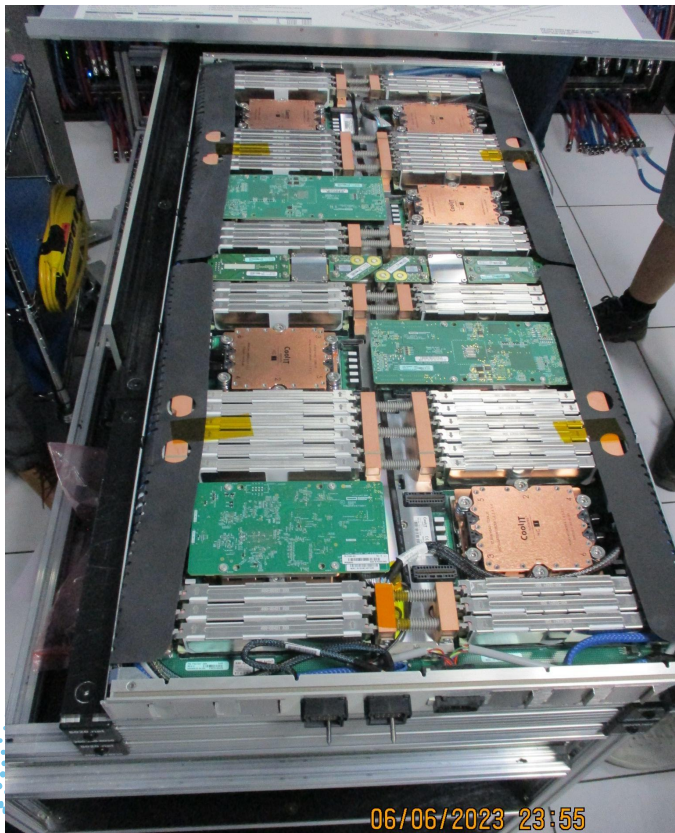
Shaheen III – CPU: #1 fastest supercomputer in the Middle East and #20 in the world in 2023



Shaheen III Hardware Specifications: Compute

Characteristics	Shaheen III Phase 1
Cabinets	18 x Cray EX4000
CPU Type	AMD Genoa 9654
#Socket X #Cores	2 x 96 = 192 cores
HPL Performance (TFlops/s)	6.87 TF/s per node
Memory	2X12X16GB DDR5@4800Mhz 384GB
Memory STREAM Bandwidth	800 GB/s per node
Total Number of Nodes	4608
Total Number of Cores	884,736
Theoretical/Sustained Peak (PFlops/s)	35.66/39.61 (90%)
Aggregated Theoretical Bandwidth	4.6 PB/s
Power (at highest load)	5.3 MW
Cooling	Direct Liquid Cooling

Shaheen III Hardware Specifications: Compute



A Cray EX 4000 CPU Blade with 4 Dual Socket AMD EPYC compute nodes



AMD EPYC 9654 Socket

Shaheen III Hardware Specifications: Compute

AMD Genoa CPU Microarchitecture

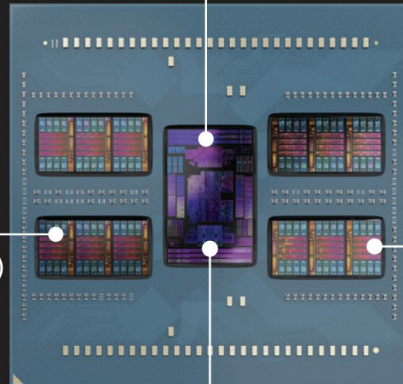
4TH GEN AMD EPYC PROCESSOR ARCHITECTURES

I/O die

- 12 memory controllers
- PCIe® Gen 5 controllers
- Infinity Fabric™ controllers
- SATA controllers
- CXL™ controllers
- AMD Secure Processor

CPU die

- Up to 16 cores per die (8 shown)
- Up to 12 dies per processor



AMD EPYC 9004 SERIES PROCESSORS (16–96 CORES)

'Zen 4' CPU die

(up to 12 per processor)

8 'Zen 4' cores

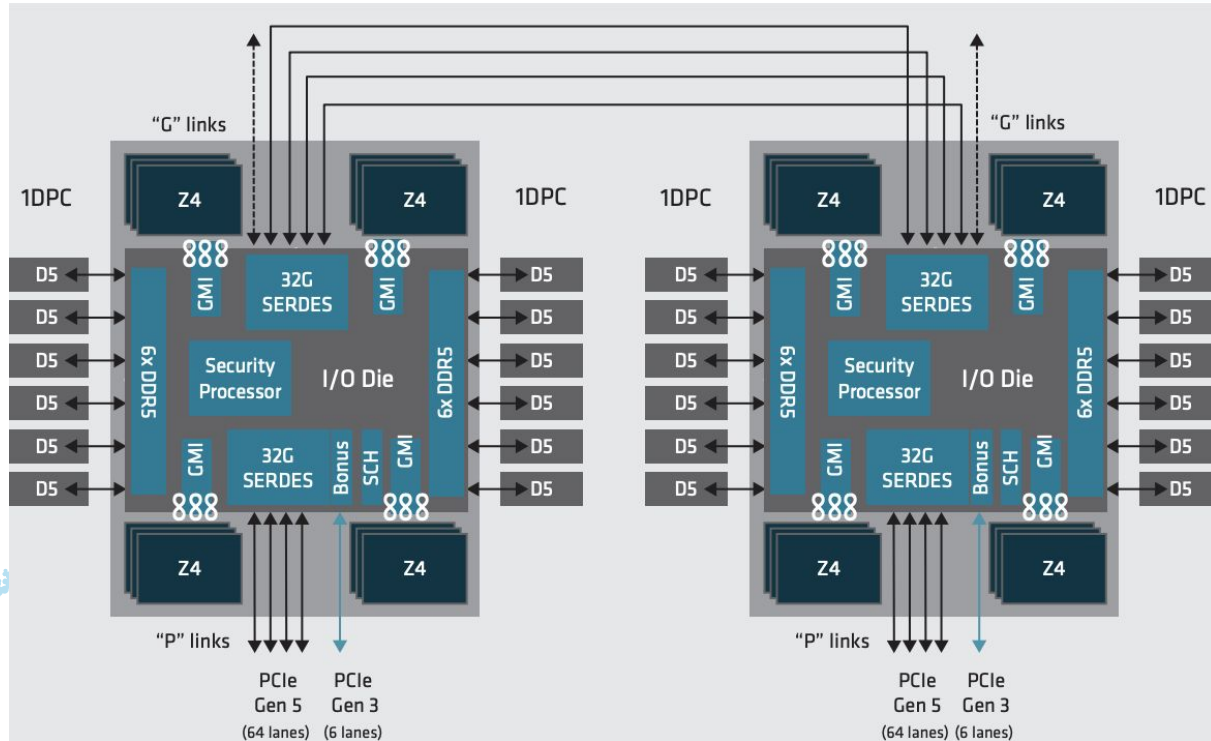
1 MB L2 cache per core

Shared 32 MB L3 cache



Shaheen III Hardware Specifications: Compute

AMD Genoa CPU node architecture



Shaheen III Hardware Specifications: Compute

Characteristics	Shaheen III Phase 2
Cabinets	7 x Cray EX4000
#Nodes	700 x 4 X GH Superchip
GPUs	H100 SXM 96GB HBM3 @ ~3TB/s
Host CPU Type	4 x NARM Grace - 72 ARM Cores each
Host Memory	4X 128GB = 512GB (LPDDR5)
CPU-GPU Interconnect	Memory Coherency
NICS	4 x Cassini 200 Gbps, 1 per superchip
GPU Perf Peak FP64 Tensor	54.5 TF/s @550W 67 TF/s @700W
Total Perf /Peak FP64/ efficiency)	152.6 PF/s / 100PF/s (65.5%) 900 PF/s HPL-AI
Power HPL in MW	2.22 MW
GF/W	44.4

Shaheen III Hardware Specifications: Interconnect

Characteristics	Shaheen III High Speed Network
Type	Slingshot-11
Topology	Dragonfly, multi-level all-to-all
Bandwidth	200 Gbps per link
Latency	Up to 2.6 μ s for max 3 hops
Injections per node	CPU Nodes: 1 injection GPU Nodes: 4 injections
Features	Adaptive routing, RDMA, Decongestion, Ethernet compatible

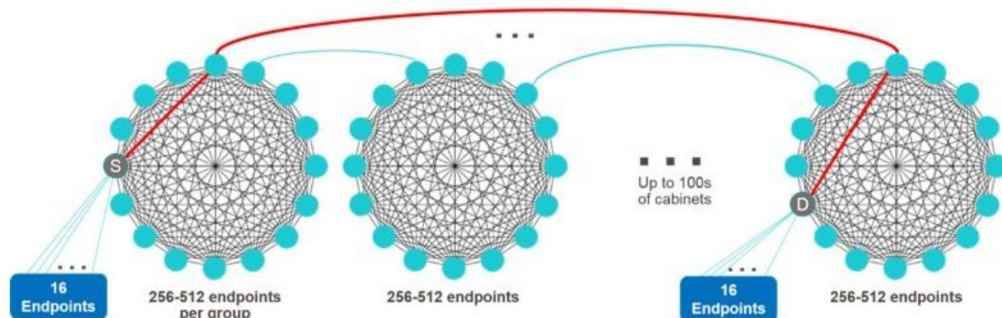
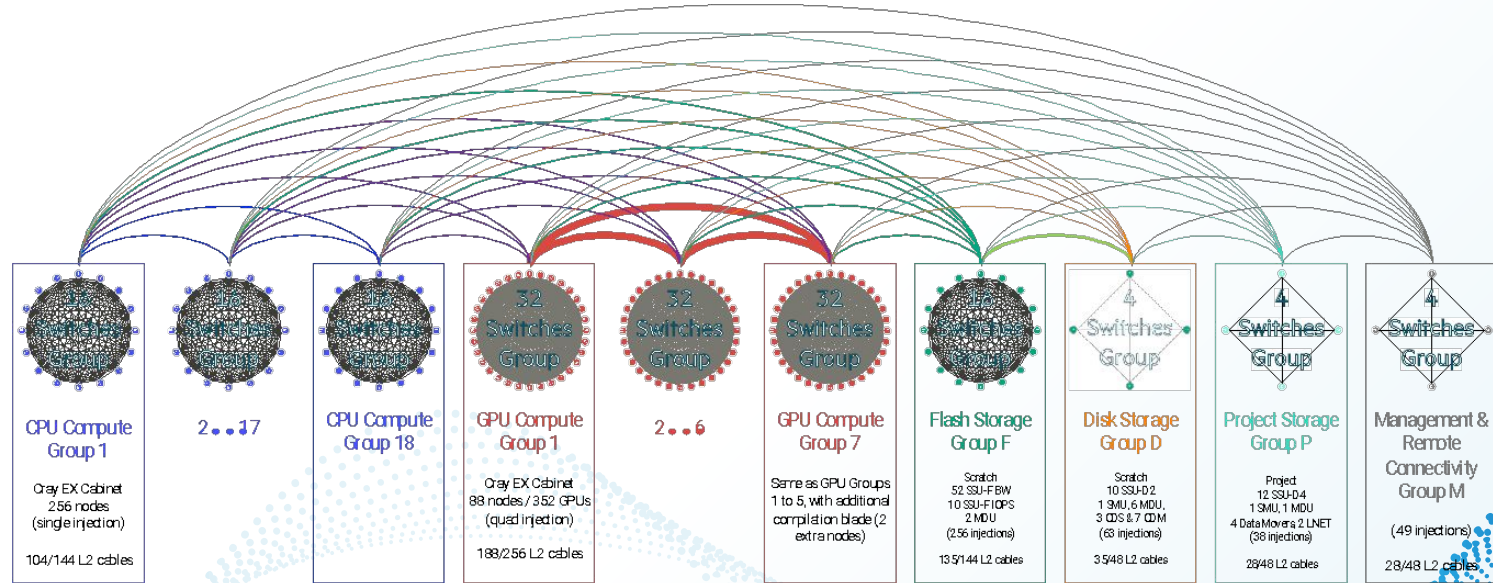


Figure 26: Slingshot Extreme Scale and Performance with Dragonfly Topology

Shaheen III Hardware Specifications: Interconnect



CPU Compute Group 1

256 nodes
(single injection)

104/144 L2 cables

2...17

CPU Compute Group 18

32 Switches Group

GPU Compute Group 1

Cray EX Cabinet
88 nodes / 352 GPUs
(quad injection)

188/256 L2 cables

32 Switches Group

2...6

32 Switches Group

GPU Compute Group 7

Same as GPU Groups 1 to 5, with additional compilation blade (2 extra nodes)

Flash Storage Group F

Scratch
52 SSU-F BW
10 SSU-F OPS
2 MDU
(256 injections)

135/144 L2 cables

4 Switches Group

Disk Storage Group D

Scratch
10 SSU-D
1 SMU, 6 MDU,
3 ODS & 7 ODM
(68 injections)

35/48 L2 cables

4 Switches Group

Project Storage Group P

Project
12 SSU-D4
1 SMU, 1 MDU
4 Data Movers 2 LNET
(38 injections)

28/48 L2 cables

4 Switches Group

Management & Reroute Connectivity Group M

(49 injections)

28/48 L2 cables

- Bundle = 4 - Between CPU Compute Groups
- Bundle = 18 - Between GPU Compute Groups
- Bundle = 4 - CPU groups to GPU Groups
- Bundle = 5 - Flash Storage Group to Compute Groups
- Bundle = 1 - Disk Storage Group to Compute Groups
- Bundle = 8 - Disk Storage Group to Flash Storage Group
- Bundle = 1 - Project Storage Group to all others
- Bundle = 1 - Management Group to all others

Shaheen III Hardware Specifications: Storage

Characteristics	Shaheen III /scratch storage
Total Capacity (usable)	32 PB
Capacity tier (HDD)	25 PB
Capacity tier perf Read/Write	330/260 GB/s
BW tier capacity	6.8PB (20.9%)
BW Perf. tier Read/Write	3750/2500 GB/s
IOPS tier capacity	338 TB
IOPS tier IOPS (Read/Write)	10+M IOPS

- I/O500: #3 in Overall Production, #7 in Bandwidth, The fastest Lustre on the list
- /project storage will be disconnected from Shaheen II
- Coming Soon on Shaheen III with upgraded capacity and performance !

Shaheen III Hardware Specifications

C O M P U T E	CPU Nodes	Processor type: AMD EPYC Genoa	2 CPU sockets per node, 96 processors cores per CPU
		4608 Nodes	884,736 cores
		384 GB of memory per node	Over 1.770 PB total memory
	GPU Nodes	Host: Grace Hopper Superchip	4 X 72 ARM Cores directly attached to NVIDIA H100 GPUs
		700 Nodes	2800 H100 GPUs
	Weight/Size	More than 100 metrics tons	18+7 Cray EX Compute cabinets
	Speed	HPL: 35Pflops/s #20 HPCG: 651.5 TF #16	GPUs with more than 100 Pflops/s sustained HPL performance
Network	Cray Slingshot interconnect	Dragonfly topology with a max of 3 hops	
S T O R E	Scratch	E1000 Lustre appliance	32 Petabytes of usable storage including a performance and IOPS optimized tiers
	Project	E1000 Lustre appliance	Same as Shaheen II (future upgrade in 2025)
	Archiving	HPE Data Management Framework (DMF) for data backup	100+ PB of tape storage, using a spectra logic tape library.