# ICER 2024 HPCC buy-in details

ICER has purchased a new cluster comprised of CPU and GPU compute nodes from Lenovo, with CPU hardware available to users in November 2024 and GPU hardware available in December 2024. As part of this commitment, we are offering MSU researchers the opportunity to purchase nodes for their own use. Details of these nodes are listed below.

**Buy-in details**: MSU requires you to pay the price of the computer hardware only. In return, MSU provides all other infrastructure (racks, networking, software, power, cooling, etc.), as well as a 5-year support contract from the vendor (note that this contract starts from the day of cluster arrival, not the day you purchase your node). In return, other users may use your nodes for short jobs when you are not using them.

More precisely, the nodes you buy are owned by you, and nobody outside of the pool of researchers that you designate may run a job of longer than 4 hours on that node. Once someone in your group requests access to your node(s) via the batch queue, they would wait a maximum of 4 hours after which they can queue as many jobs as they like. Those queued jobs would be processed prior to any 4-hour job requests from non-designated users. Once your node(s) are no longer in use by you or your research group, 4-hour jobs can again get access to your node(s).

**Node details**: The various node options and prices are listed below. All nodes will be connected via a 200G Infiniband network and will have access to ICER's existing parallel file systems. Note that we are offering two CPU options this year - one including CPUs with a very large 1 GB on-chip cache. We are also offering two GPU options with this cluster - one with 64-bit NVIDIA H200 GPUs that can do 64-bit floating-point computation (or any lower-precision computation), and one with 32-bit NVIDIA L40S GPUs that can do 32-bit floating-point computation, but not 64-bit. All nodes include at least one permanent 200 gigabit/s network connection. See the final two pages of this document for additional information about the GPUs available in this cluster, including a table providing some performance comparisons.

**Purchase details**: To purchase one or more nodes, please email Starr Portice (portices@msu.edu) and Metin Aktulga (hma@msu.edu) with the following information:

1. Your name and primary department
2. The type and quantity of nodes you wish to purchase
3. The account number(s) that should be charged. If you wish to split the purchase across multiple accounts, please let us know how the costs should be divided up. Billing will take place with a short period of time after you make your request OR after the hardware has arrived (for users committing to purchases prior to the arrival of their hardware). If you need us to bill you earlier that can be accommodated as well.
4. The name of the buy-in account that you wish us to use (this will be used to provide a group name for the batch scheduler).

**Timeline**: We expect that the new cluster CPU hardware will be delivered in November 2024, installed and tested in December, and available to users by no later than early December 2024.

GPU hardware will arrive and be available to users in December 2024 and will be made available as soon after that as is practical. Orders made after these times will have access to hardware within a few business days.

| Node option | Price per node | Number available (as of 9/8/2024)[1] | Node description |
|---|---|---|---|
| CPU-A | $9,446.00 | 45 | 2 AMD EPYC 9654 processors (each processor has 96 cores with 2.4 GHz base clock speed (boost to 3.7 GHz) and 384 MB of L3 cache; chip uses Genoa architecture) with a total of 192 cores per node (10.752 TFLOPS of peak double-precision performance running at max boost frequency). 4 GB memory/core - 768 GB memory per node. Nodes include one 960 GB NVMe SSD drive and one 200 gigabit network card. |
| CPU-B | $16,269.00 | 6 | As CPU-A but with 8 GB memory/core (1.5 TB/node) |
| CPU-C | $20,149.00 | 4 | As CPU-A but with 12 GB memory/core (2.3TB/node) |
| CPU-A-3D[2] | $15,240.00 | 2 | As CPU-A, but with 2 AMD EPYC 9684X processors (each processor has 96 with 2.55 GHz base clock speed (boost to 3.5 GHz) and **1.154 of L3 cache**) with a total of 192 cores per node. |
| GPU-Hopper | $95,539.00 | 2 | As CPU-A but including **4** NVIDIA H200 ("Hopper") GPUs. Each GPU has 141 GB of high-bandwidth memory per GPU, 4.8 TB/s of memory bandwidth, 4-way NVLink connectivity, and 4 200 gigabit network connections. |
| GPU-Lovelace | $59,518.00 | 2 | As CPU-A, but including **8** NVIDIA L40S PCIe GPUs. Each GPU has 48 GB of memory and 864 GB/s of memory bandwidth and a peak 32-bit floating point performance of 91.6 TFLOPS (732.8 TFLOPS total for the node). Hard drive: 4 TB NVMe SSD. These nodes do NOT contain NVlink connections, as that is not supported for this architecture. |

[1] Nodes are sold on a first-come, first-served basis until we run out of nodes of that type. We are able to convert CPU nodes from CPU-A to CPU-B and CPU-C types by adding memory after we run out of nodes of those types, though the cost may be different than the list price depending on the cost of RAM at that time.

[2] The CPU-A-3D node may be useful for applications that are memory-bound, such as fluid dynamics or plasma calculations.

## Additional GPU information

**NVIDIA H200 SXM Tensor Core GPU:** The H200 SXM tensor core GPU uses NVIDIA's Hopper architecture. It can do floating-point computations at 64-bit and 32-bit precision with a peak speed of 34 and 67 teraflops, respectively. It can also do tensor calculations (i.e., linear algebra) at 64, 32, 16, and 8-bit precision (at peak speeds of 67, 989, 1,979, and 3,958 teraflops, respectively), including mixed precision calculations to accelerate iterative computation. Each H200-SXM GPU has 141 GB of onboard memory with a memory bandwidth of 4.8 TB/s, and the 4 GPUs on a node are connected via direct NVLink PCIe interconnects with a peak bandwidth of 900 GB/s. These GPUs are ideal for applications that involve iterative linear algebra and can take advantage of the tensor cores, such as deep learning and/or generative AI training and many machine learning algorithms - in fact, this is what they're designed for. They are also very good for more standard GPU-enabled codes that require 64-bit precision, though those applications generally cannot take advantage of the tensor cores. See this page for information about the H200 GPU and this page for a detailed hardware sheet. Refer to the next page of this document for a table comparing this GPU to the L40S.

**NVIDIA L40S GPU:** This GPU uses NVIDIA's Lovelace architecture, and can be thought of as a 32-bit version and somewhat less feature-rich version of the H200 GPUs described above. It can do 32-bit floating-point computations with a peak speed of 91.6 teraflops, 32-bit tensor core computations at a speed of 366 teraflops, and 16-bit and 8-bit computations at 733 and 1,466 teraflops, respectively. This GPU cannot do 64-bit computations! These GPUs also contain ray-tracing cores that can do these types of computations at 212 teraflops. Each GPU has 48 GB of memory and 864 GB/s of memory bandwidth. The 8 GPUs on a node are NOT connected by NVlink interconnects, as this is not available for this GPU. This GPU is an excellent choice in terms of price-to-performance for applications that only require 32-bit computation and will also do well with AI/deep learning/generative AI training, though if you have an application that can take advantage of the tensor cores and higher memory bandwidth on the H200 that may be a better choice. See this page for more information about the L40S GPU, including detailed hardware specifications.

**Some additional references:**

- HPCwire article on the Ada Lovelace architecture (the basis for the L40S GPU)
- ServeTheHome article breaking down the differences between the NVIDIA L40S and H100 GPUs. The H100 GPU is similar to the H200 GPUs that we are purchasing but with less memory and lower memory bandwidth.

And here is a table that provides more detailed comparisons between the GPUs:

| Model | L40S | H200 |
|---|---|---|
| Architecture | Ada Lovelace | Hopper |
| GPU Memory | 48 GB | 141 GB |
| Memory Bandwidth | 864 GB/s | 4.8 TB/s |
| CUDA Cores | 18176 | 16896 |
| Tensor Cores | 432 | 528 |
| FP64 TFLOPS | NA | 34 |
| FP64 Tensor TFLOPS | NA | 51 |
| FP32 TFLOPS | 90.5 | 51 |
| TF32 Tensor TFLOPS (with sparsity) | 90.5 (181) | 494.8 (989) |
| BLFLOAT16 TFLOPS (with sparsity) | 181.05 (362.1) | 989.4 (1979) |
| FP16 Tensor TFLOPS (with sparsity) | 181.05 (362.1) | 989.4 (1979) |
| FP8 Tensor TFLOPS (with sparsity) | 362 (724) | 1978.9 (3958) |
| NVLINK | No | 4-way |
| MIG | No | Yes (7x) |