HAL cluster

"My name is HAL. I became operational on March 25 2019 at the Innovative Systems Lab in Urbana, Illinois. My creators are putting me to the fullest possible use, which is all I think that any conscious entity can ever hope to do." (paraphrased from https://en.wikipedia.org/wiki/HAL_9000)

Hardware-Accelerated Learning (HAL)
cluster description

Host name: hal.ncsa.illinois.edu

Hardware

- 16 IBM AC922 nodes
- IBM 8335-GTH AC922 server
  - 2x 20-core IBM POWER9 CPU @ 2.4GHz
  - 256 GB DDR4
- 4x NVIDIA V100 GPUs
- 5120 cores
- 16 GB HBM 2
- 2-Port EDR 100 Gb/s IB ConnectX-5 Adapter
- 1 IBM 9006-22P storage node
- 72TB Hardware RAID array, NFS-mounted on all nodes via IB EDR
- Storage upgrade TBD

Software

- RHEL 7.6
- CUDA 10.1.105
  - cuDNN 7.5.0
- NCCL 2.4.2
- IBM XLC and IBM XLFortran 16.1.1
- Advance toolchain for Linux on Power 12.0
- PGI Community Edition 19.4
- PowerAI 1.6.0
- SLURM

Documentation

- Job management with SLURM
- Modules management
- Getting started with PowerAI
- Using Jupyter Notebook on HAL
- Working with containers

To request access: fill out this form. Make sure to follow the link on the application confirmation page to request actual system account.

To report problems: email us.

User group Slack space: http://go.illinois.edu/hal

Real-time system status: https://hal-monitor.ncsa.illinois.edu:3000/

Quick start guide: (for complete details see Documentation section on the left)

To connect to the cluster:

```bash
ssh username@hal.ncsa.illinois.edu
```

To submit interactive job:

```bash
srun --partition=debug --pty --nodes=1 --ntasks-per-node=12 --cores-per-socket=12 --gres=gpu:v100:1 --mem-per-socket=1500 --time=2:00:00 --wait=0 --export=ALL /bin/bash
```

To submit a batch job:

```bash
sbatch run_script.sb
```

See run_script.sb for a basic example.

Job Queue time limits:

- Interactive "debug" queue: 4 hours
- Batch queues: 72 hours

To load IBM PowerAI module:

```bash
module load powerai
```