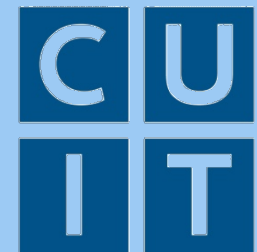




Habanero Operating Committee

January 25 2017



COLUMBIA UNIVERSITY
INFORMATION TECHNOLOGY

Habano Overview

1. Execute Nodes
2. Head Nodes
3. Storage
4. Network



Execute Nodes

Type	Quantity
Standard	176
High Memory	32
GPU*	14
Total	222



Execute Nodes

Standard Node	
CPU (2 per node)	E5-2650v4
Clock Speed	2.2 GHz
Cores	2 x12
Memory	128 GB
High Memory Node	
Memory	512 GB
GPU Node	
GPU (2 per node)	Nvidia K80
GPU Cores	2 x 4992



Execute Nodes

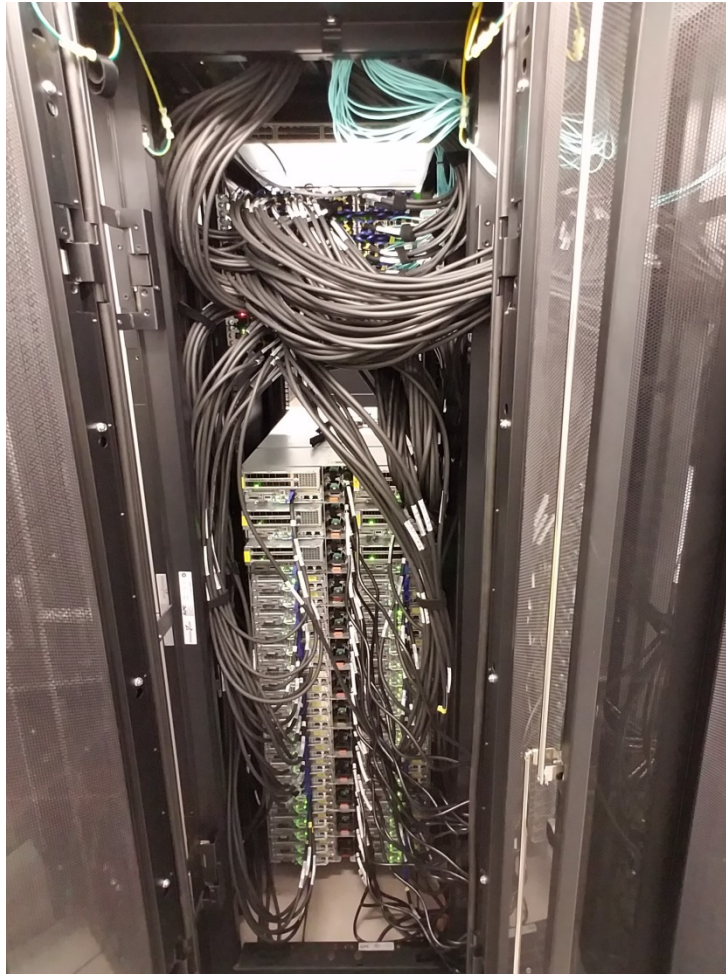


Execute Nodes



COLUMBIA UNIVERSITY
INFORMATION TECHNOLOGY

Execute Nodes



Head Nodes

Type	Quantity
Submit	2
Data Transfer	2
Management	2



Head Nodes



Storage

Type	Quantity
Model	DDN GS7K
File System	GPFS
Network	FDR Infiniband
Storage	407 TB



Storage



Network

Habanero	
EDR Infiniband	96 Gb/s
Yeti (for comparison)	
FDR Infiniband	54
1 Gb Ethernet	1
10 Gb Ethernet	10



Visualization Server

- Coming in February (probably)
- Remote GUI access to Habanero storage
- Reduce need to download data
- Same configuration as GPU node



Business Rules

- Business rules set by Habanero Operating Committee
- Habanero launched with rules similar to those used on Yeti



Nodes

For each account there are three types of execute nodes

1. Nodes owned by the account
2. Nodes owned by other accounts
3. Public nodes



Nodes

1. Nodes owned by the account
 - Fewest restrictions
 - Priority access for node owners



Nodes

2. Nodes owned by other accounts

- Most restrictions
- Priority access for node owners



Nodes

3. Public nodes

- Few restrictions
- No priority access



12 Hour Rule

- If your job asks for 12 hours of walltime or less, it can run on any node
- If your job asks for more than 12 hours of walltime, it can only run on nodes owned by its own account or public nodes



Job Partitions

- Jobs are assigned to one or more “partitions”
- Each account has 2 partitions
- There is a shared partition for short jobs



Job Partitions

Partition	Own Nodes	Others Nodes	Public Nodes	Priority?
<Account>1	Yes	No	No	Yes
<Account>2	Yes	No	Yes	No
short	Yes	Yes	Yes	No



Maximum Nodes in Use

Walltime	Maximum Nodes
12 hours or less	100
Between 12 hours and 5 days	50



Fair Share

- Every job is assigned a priority
- Two most important factors in priority
 1. Target share
 2. Recent use



Target Share

- Determined by number of nodes owned by account
- All members of account have same target share

Recent Use

- Number of cores*hours used “recently”
- Calculated at group and user level
- Recent use counts for more than past use
- Half-life weight currently set to two weeks



Job Priority

- If recent use is less than target share, job priority goes up
- If recent use is more than target share, job priority goes down
- Recalculated every scheduling iteration

Support Services

1. User support: hpc-support@columbia.edu
2. User documentation
3. Monthly Office Hours
4. Habanero Information Session
5. Group Information Sessions



User Documentation

- hpc.cc.columbia.edu
- Go to “HPC Support”
- Click on Habanero user documentation



Office Hours

HPC support staff are available to answer your Habanero questions in person on the first Monday of every month.

Where: Science & Engineering Library, NWC Building

When: 3-5 pm first Monday of the month

Next session: 3-5 pm Monday February 6



Habanero Information Session

Introduction to Habanero

Tuesday January 31, 1:00 pm - 3:00 pm

Science & Engineering Library, NWC Building

Mostly a repeat of session held in December

- Cluster overview
- Using slurm to run jobs
- Business rules



Group Information Sessions

HPC support staff can come and talk to your group

Topics can be general and introductory or tailored to your group.

Contact [hpc-support](#) to discuss setting up a session.



Benchmarks

High Performance LINPACK (HPL) measures compute performance and is used to build the TOP500 list.

Nodes	Gflops	Gflops / Node
1	864	864
4	3041	762
10	7380	738
219	134900	616

Intel MPI is a set of MPI performance measurements for communication operations for a range of message sizes.

- Bandwidth: 96 Gbit/s average Infiniband bandwidth measured between nodes.
- Latency: 1.3 microseconds



Benchmarks (continued)

IOR measures parallel file system I/O performance.

- Mean Write: 9.9 GB/s
- Mean Read: 1.46 GB/s

mdtest measures performance of file system metadata operations.

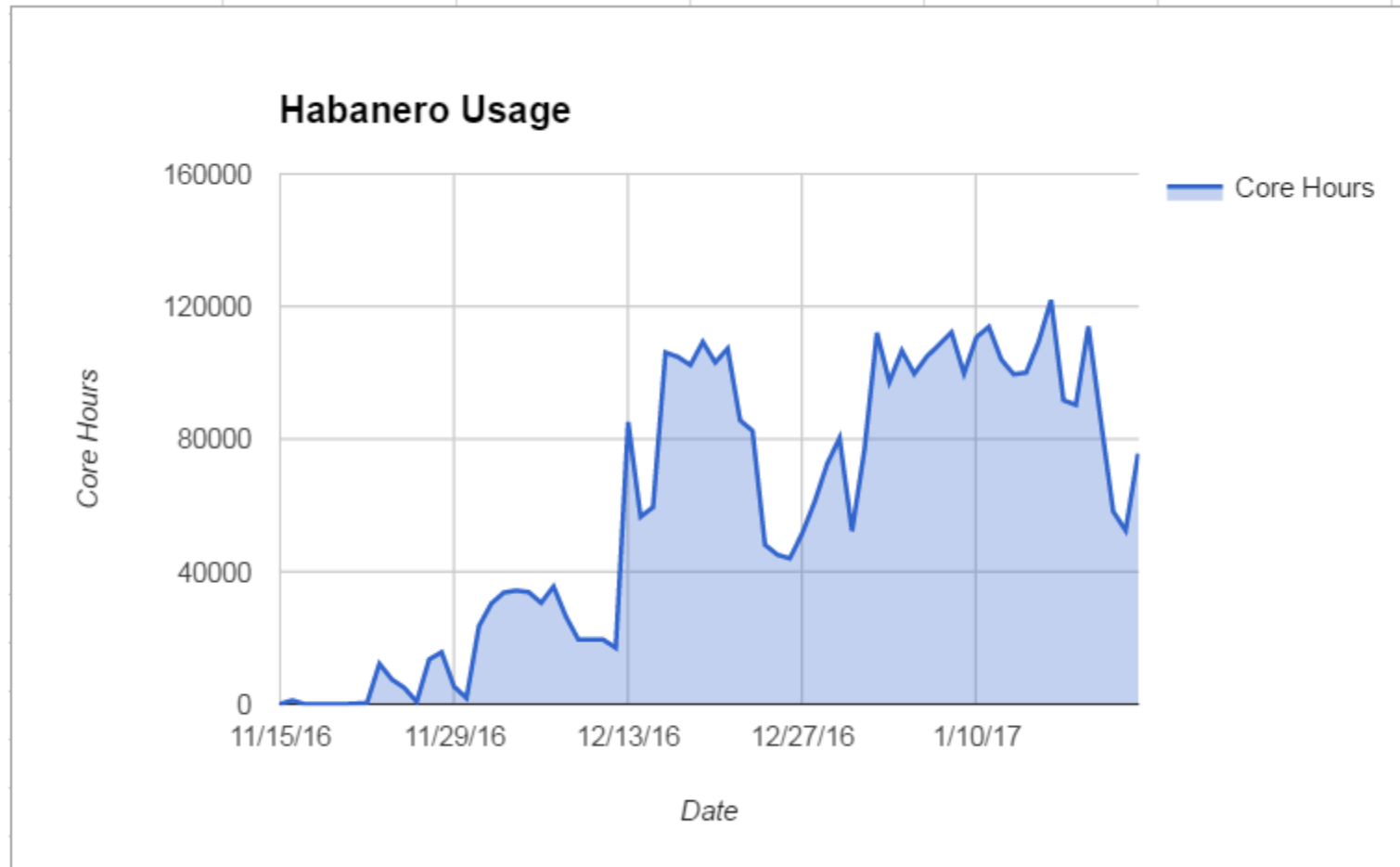
- Create: 41044 OPS
- Remove: 21572 OPS
- Read: 29880 OPS

STREAM measures sustainable memory bandwidth and helps detect issues with memory modules.

- Memory Bandwidth/core: 6.9 GB/s



Usage



End of Slides

Questions?

User support: hpc-support@columbia.edu

