BlueHive Tutorial

Department of Political Science
University of Rochester
Overview

- What is BlueHive?
- Connecting to BlueHive
- Sessions, Partitions, and Resource Constraints
- Department Nodes and Ground Rules
- Available Software
- Storage and Transferring Files
What is BlueHive?

▶ Consider a typical desktop computer:
  - ~ 2-4 CPU cores
  - ~ 16-32 GB RAM

▶ Consider a desktop computer on steroids:
  - ~ 24-36 CPU cores
  - ~ 128-256 GB RAM
What is BlueHive?

- BlueHive is a cluster (network) of such powerful computers (402 “nodes”):
  - 9,648 CPU cores
  - 44 TB RAM

- Standard user allowance: 120 CPU cores at a time (more resources may be available on request)
Connecting to BlueHive

- You must be connected to the UR network:
  - If you’re on campus, you must use a wired (Ethernet) connection or the *UR_Connected* wireless connection
  - If you’re off campus, you must connect to the UR Virtual Private Network (VPN):
    http://tech.rochester.edu/services/remote-access-vpn/
Connecting to BlueHive: Three Alternatives

- **Using the Linux Command Line**
  - Most versatile but requires familiarity with the UNIX/Linux language

- **Using X2Go**
  - Much more user-friendly, can be downloaded and installed for free on Mac/Windows/Linux machines
  - Graphical user interface (GUI) that creates a remote desktop session to visually interact with BlueHive

- **FastX**
  - Same as X2Go but browser-based and robust to loss of network connection
Connecting to BlueHive: FastX

- Go to https://bluehive.circ.rochester.edu/

- Log in with your NetID credentials
  - You will receive a Duo prompt for authentication

- Click on Launch Session to start a new session
  (or select an existing session to resume)
Default Session

- Limited resources but long duration (30 days)
  - After 30 days, session will terminate automatically and any unsaved work will be lost
  - But user can terminate and start a fresh default session anytime

- Ideal for managing files, writing and testing code, or any other task that is not computationally intensive

- Session can be resumed anytime, even if browser is closed or network connection is lost
Interactive Sessions and Node Partitions

- Flexible resources: e.g.,
  - `--time=8:00:00` (8 hours) or `--time=2-00:00:00` (2 days)
  - `--cpus-per-task=12` (number of cores)
  - `--mem=24g` (total RAM)
  - `-p standard` (BlueHive partition)

- BlueHive nodes are grouped into partitions:
  [https://info.circ.rochester.edu/BlueHive/Compute_Nodes.html](https://info.circ.rochester.edu/BlueHive/Compute_Nodes.html)

- Interactive session is launched on a node in specified partition
  - Thus, resource limits are determined by partition
  - E.g., `debug` partition limits: 1 hour, 24 cores, 62GB RAM

- Multiple simultaneous interactive sessions are possible (up to 16 nodes or 120 cores)
Resource Availability

- Interactive session will not launch until requested resources become available.

- It is possible to check node availability in each partition (but not, unfortunately, when resources in use will become available):
  - In any active session, go to
    \( \textit{Applications} > \textit{System Tools} > \textit{MATE Terminal} \)
  - Once Terminal window opens, type: \texttt{sinfo -p [partition name]} (e.g., \texttt{sinfo -p standard})
  - Available nodes will be shown as \textit{idle} or \textit{mix} (partially idle)
Political Science Partition

- Our department has priority access to partition *pscd*, which has global limits: 14 days and 40 cores per user.

- Department nodes are accessible via “-p pscd -w [node name]”

- Node *bhd0042* has 24 cores @ 3.2GHz (processor speed) with 278GB RAM

- Node *bhd0043* has 40 cores @ 2.4GHz with 371GB RAM, ideal for parallelization
Political Science Partition: Ground Rules

- Please use public Bluehive resources (e.g., your default session) to write/test/debug code

- Be economical: don’t request more resources (cores or memory) than needed for your job(s)

- Graduate students:
  - Please limit yourselves to no more than 8 cores or 64GB of memory at a time
  - If you need more resources or are planning to run a job that takes longer than 7 days, please contact the department computing supervisor beforehand
Available Software (go to Applications)

- R (and RStudio)
- Stata
- Matlab
- Mathematica
- Knitro (optimization)
- Gambit (computes Nash equilibria of finite games)
- Python, Jupyter, Julia, ...

- Some (e.g., Knitro and Gambit) require additional steps to load and use (covered in separate tutorial)
Storage

- Every BlueHive user is given three file directories: 
  
  `/home/[NetID]`, `/scratch/[NetID]`, and `/public/[NetID]`

- To access your directories:
  
  - Click on the *Computer* icon
  - Go to *File System* > *home* (or *scratch* or *public*) > `[NetID]`

- The quota for home is 20GB
  
  - This directory is backed up securely every 24 hrs

- The quota for scratch and public is 200GB (combined)
  
  - These directories are not backed up

- Home and scratch are only accessible by owner, while public is readable (but not writeable) by all BlueHive users
Transferring Files: Three Alternatives

- Using the Linux Command Line
  - Most versatile but requires familiarity with the UNIX/Linux language

- Using a GUI:
  - *Fetch* for Mac:
    https://info.circ.rochester.edu/Getting_Started/Connecting_using_Mac_OS_X/Transferring_Files.html
  - *WinSCP* for Windows:
    https://info.circ.rochester.edu/Getting_Started/Connecting_using_Windows/Transferring_Files.html

- Using, e.g., *Dropbox* from a browser window (Firefox) within BlueHive session
Additional Information

- Go to https://info.circ.rochester.edu for additional documentation (including FAQs) on CIRC systems and software.

- For tech support, email circ@rochester.edu.

- Our department website has additional tutorials and documentation: https://www.sas.rochester.edu/psc/research/computing.html.