HBIC COGNEURO WORKGROUP

HBIC COMPUTING RESOURCES

EXPANDING SYNAPSE CAPACITY

KU MEDICAL CENTER
The University of Kansas

SYNAPSE CAPACITY ISSUES AND SOLUTIONS

- Slow applications due to high load
- Incomplete preprocessing jobs due to killed processes
- Increasing capacity:
  - Added new compute servers
  - KU-L HPC cluster
  - Lab-owned resources
  - Web services

previous synapse

24 CPU @ 2.40 GHz
64 GB RAM

Physical hardware.
Exceeded service life
requiring decommission.

new hbic-synapse

12 CPU @ 2.40 GHz
24 GB RAM

Virtual machine.
Perpetual support,
in-place hardware upgrades

NEW HBIC SERVER LINEUP

- Added four servers
  - 8 CPUs / 12 GB RAM
- Main server
  - hbic-synapse

- Data Visualization
- Data management
- Small/short processing jobs
- Submit jobs to Compute nodes

- New Interactive terminals
  - hbic-alpha
  - hbic-beta

- Large interactive jobs
- Graphical applications

- New Compute nodes
  - hbic-gamma
  - hbic-delta

- Long processing scripts
- Dedicated CPU/RAM
- No ssh access

NEW TERMINALS: HBIC-ALPHA, HBIC-BETA

- Cloned hbic-synapse server: the same applications are on hbic-alpha and hbic-beta
- *NB: No syncing is done of user home directories or configuration files across servers. Keep shared files on R-Drive*
- Select a server after first checking current processing loads
- Connect the same way as on hbic-synapse:
  ```
  ssh hbic-alpha.kumc.edu
  ssh hbic-beta.kumc.edu
  ```
- These are an unmoderated shared resource – be considerate of your use
- Not suited for jobs that depend on dedicated resources for their duration (e.g., AFNI's SSWarper, freesurfer's recon-all)

NEW COMPUTE NODES: HBIC-GAMMA, HBIC-DELTA

- Uses SLURM (Simple Linux Utility for Resource Management) for job scheduling and resource management
- Synapse acts as the host. Login to Synapse, specify the script to run and its required CPU/RAM and submit it to the scheduler. Those resources are dedicated to your process for the job's duration
- If the requested resources are not available, the scheduler will wait to run it until those resources are freed up
- Request only as much CPU/RAM as you need to minimize wait times

SLURM USAGE

- Login to synapse and create a job script specifying the required resources (CPUs, memory, time) and commands to run.
- Use sinfo to check node status
- Submit the job using sbatch
- Use squeue to check the status of submitted jobs
- SLURM ensures jobs are queued and only start when the requested resources are available, preventing conflicts and killed processes due to exhausted memory

myjob.sh

```
#!/bin/bash
#SBATCH --job-name=fs_project      # Job name
#SBATCH --output=fs_%j.log          #logfile name
#SBATCH --error=fs_%j_error.log    #logfile name
#SBATCH --ntasks=1                      # Run a single task
#SBATCH --cpus-per-task=4           # Number of CPU cores per task
#SBATCH --mem=8G                      # Memory for the whole job
#SBATCH --time=1-00:00:00           # Time limit day-H:M:S
#SBATCH --mail-type=ALL              # Mail events (NONE, BEGIN, END,
FAIL, ALL)
#SBATCH –mail-user=jbartolotti2@kumc.edu  # mail address (kumc only)
load freesurfer
~/R-Drive/Bartolotti_J/project/2_freesurfer_recon.sh
```

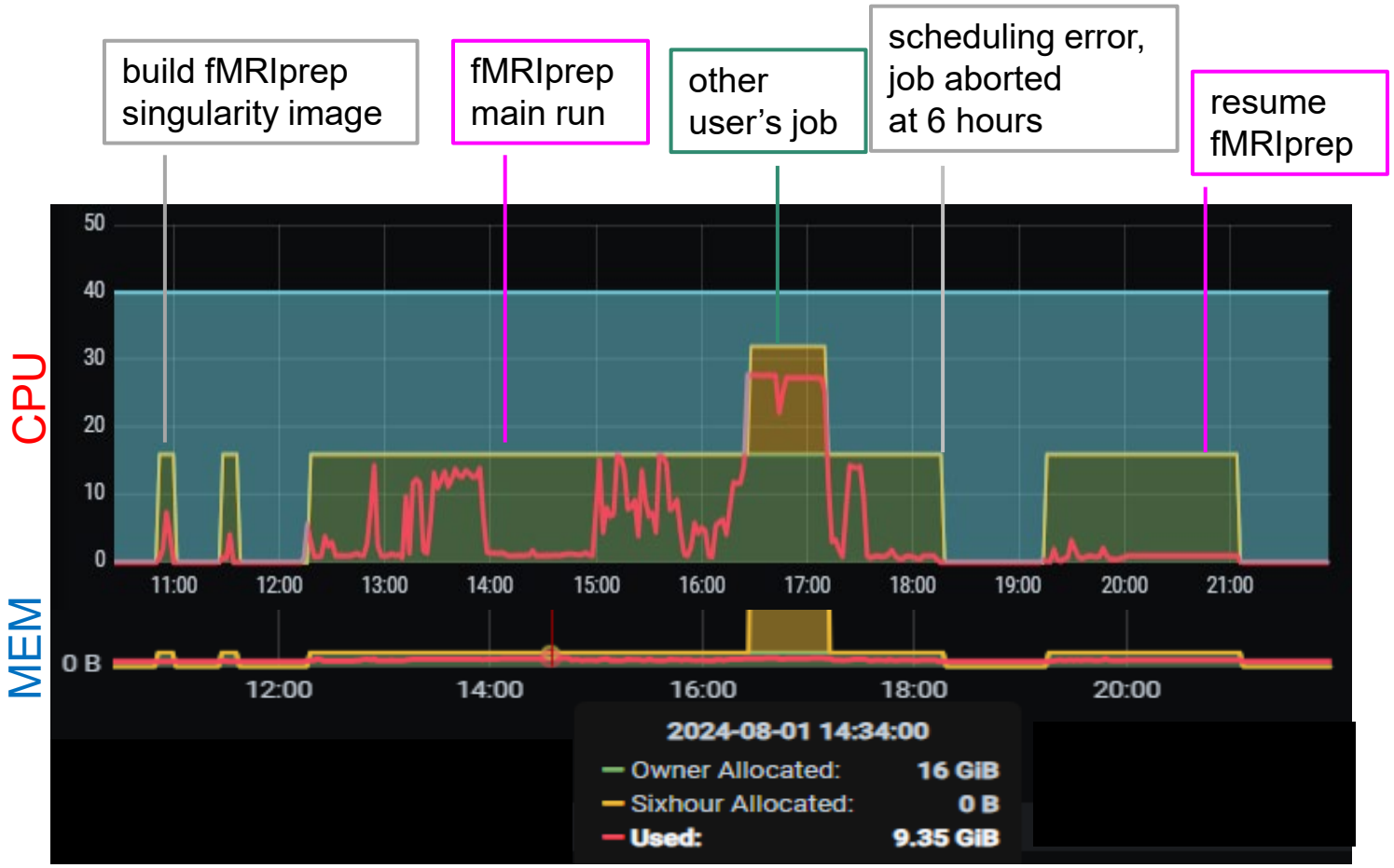
```
[sa-j186b025@hbic-synapse ~]$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
hbic         up   infinite      1   idle hbic-delta.kumc.edu
hbic         up   infinite      1   idle hbic-gamma.kumc.edu
[sa-j186b025@hbic-synapse]$ sbatch myjob.sh
```

KUMC COMMUNITY CLUSTER

- High Performance Computing (HPC) Cluster that aggregates hardware purchased by different KU-L and KUMC researchers
- Hardware and software administered by the KU Center for Research Computing Staff with a 5-year warranty, ~$11,000 upfront cost
- Access is restricted to hardware owners or individuals sponsored by an owner
- Data processing jobs can be run on your own node, or a special “six-hour” time-limited partition comprising pooled inactive resources from all nodes
- Jobs submitted using SLURM

Example fMRIprep processing
- 16 CPU cores
- ~8 GB RAM (max 9.35)
- 7-8 hours runtime

Allocated: 16 CPU / 16 GB



KU CLUSTER AND KUMC SERVER SPECS

KU Cluster

~$11,000 Standard Compute Unit (5 year warranty)

- 48 cores @ 2.6GHz *Dual Intel Xeon 6442Y CPU*
- 256GB RAM *4800MT/s DDR5*
- 480GB SSD Hard Disk
- 1 Gb/s Ethernet & 100 Gb/s Infiniband
- 5 Year Hardware Warranty
  after 5 years owner pays for repairs; unit remains in cluster as long as space is available

KUMC Large Server (e.g., Synapse)

$6,000 KUMC server (over 5 years)

- 12 CPU @ 2.20 GHz
- 24 GB RAM
- 500 GB Hard Disk
- 10 Gb/s Ethernet
- $1,200 annual fee
  hardware maintained by IT, no repair costs to VM owner

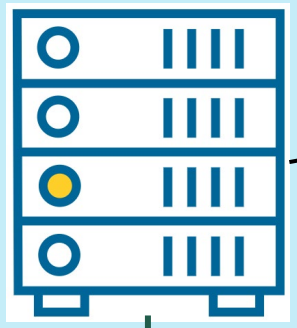
KUMC Typical Server

$3,950 KUMC server (over 5 years)

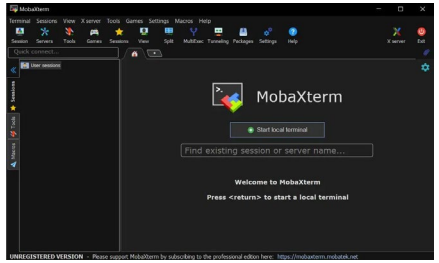
- 8 CPU @ 2.20 GHz
- 12 GB RAM
- 500 GB Hard Disk
- 10 Gb/s Ethernet
- $790 annual fee
  hardware maintained by IT, no repair costs to VM owner

USAGE

KU Cluster



ssh



ssh

KUMC hbic-synapse

$HOME
50 GB/user
NO PHI

$WORK
1TB/node
NO PHI

$SCRATCH
temp storage
NO PHI

R-Drive
NO PHI

P-Drive

Typical Workflow

Copy Raw Data
XNAT –> P-Drive
Synapse: De-identify data
Synapse: Copy De-id data
P/R Drive –> $WORK
KU Cluster: Preprocess data
KU Cluster: Group analysis script
Synapse: Move processed data $WORK -> P/R Drive
Synapse: Interactive data visualization

Pros
- Access to more compute resources
- Access to “6-hour” partition (pooled resources from other nodes, 6 hour time limit per job)
- Guaranteed minimum CPU/RAM per job

Cons
- Limited interactive terminal / graphical apps
- Limited software installation (use containers)
- No PHI data storage or processing

Pros
- Interactive terminal, no time limits
- Administrated by HBIC, flexible software installation
- Can mount and interact with P-Drive (PHI data)
- Can mount KU Cluster drives (HOME, WORK, SCRATCH)
- Run web servers (e.g., XNAT)

Cons
- Limited, shared compute resources
- Jobs killed if CPU/RAM exhausted partway through

OTHER RESOURCES

- Lab-owned server
  - KUMC IT supports Windows or Linux (RedHat) servers.
    IT provides secure access, backup, and security/OS updates.
  - Application installation and support is performed by an application admin in your team, *not* IT
  - https://kumed.sharepoint.com/sites/mykumc/catalog/tech/Lists/Services/DispForm.aspx?ID=85
- Cloud computing (e.g. Amazon Web Services)
  - Highly scalable
  - Suitable for preprocessing very large datasets in parallel
  - Current lack of KUMC/HBIC support

Typical Analysis Server

8 CPU @ 2.20 GHz
12 GB RAM
500 GB Hard Disk
Disk backed up by IT
$65.91/month
$790 annual
Disk not backed up by IT
$23.33/month
$280 annual