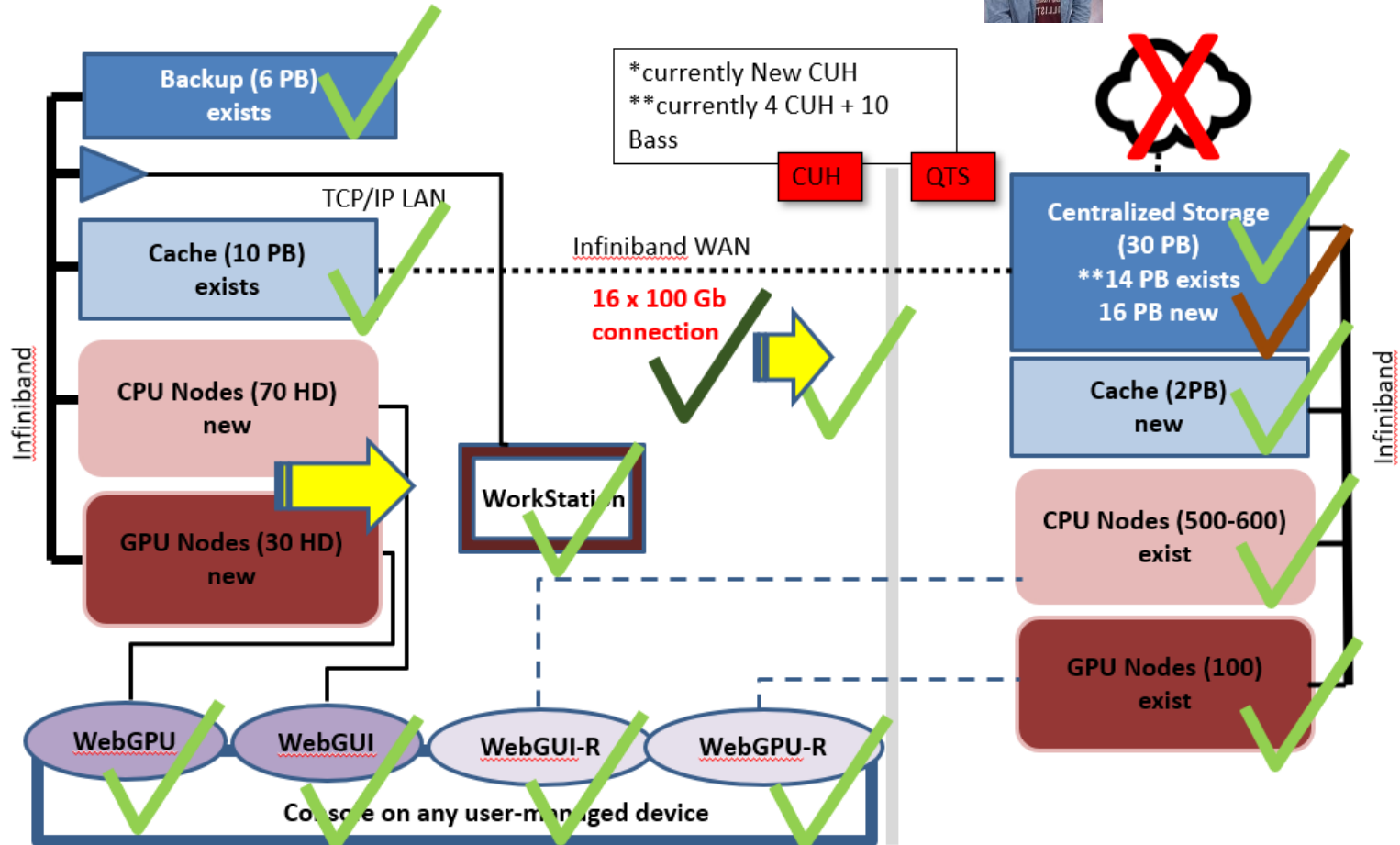
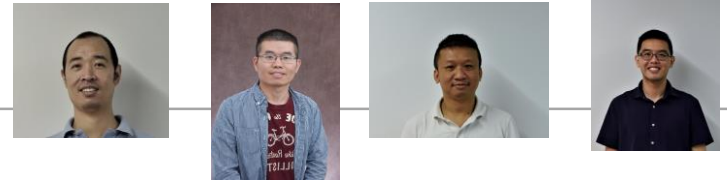


UT Southwestern
Medical Center
Lyda Hill Department of Bioinformatics

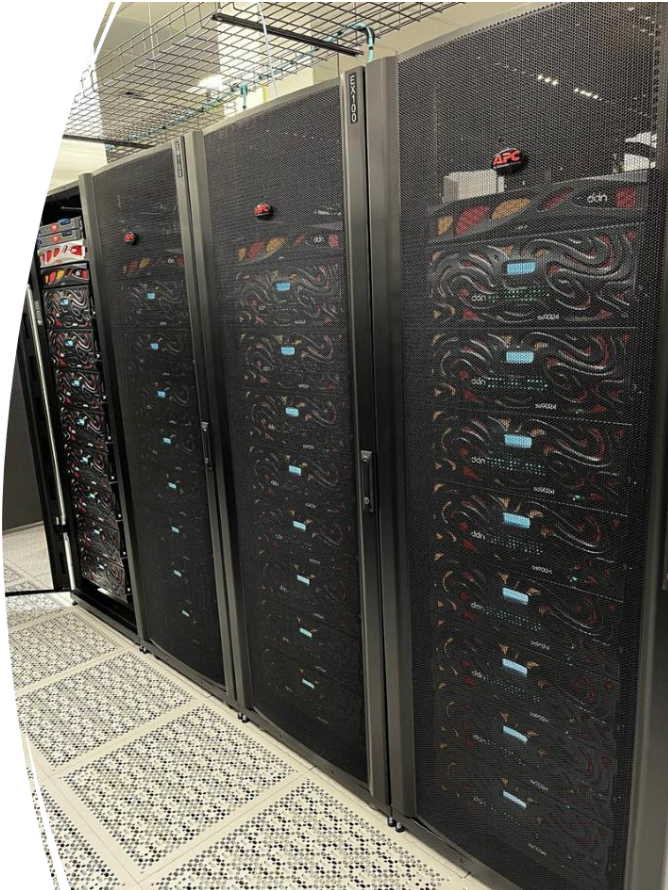
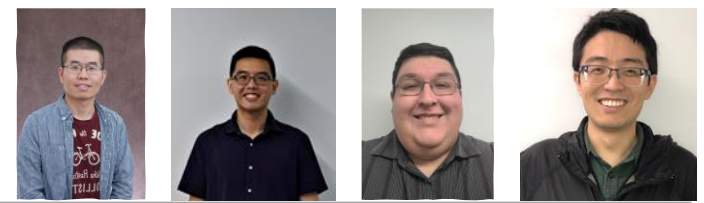
BioHPC

BioHPC Roadmap

Two datahall sites



New /project file system and storage roadmap



/home2

user profile

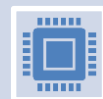
two backups per week



/work

important data

one backup per week



/project

scratch data

performance**



/archive

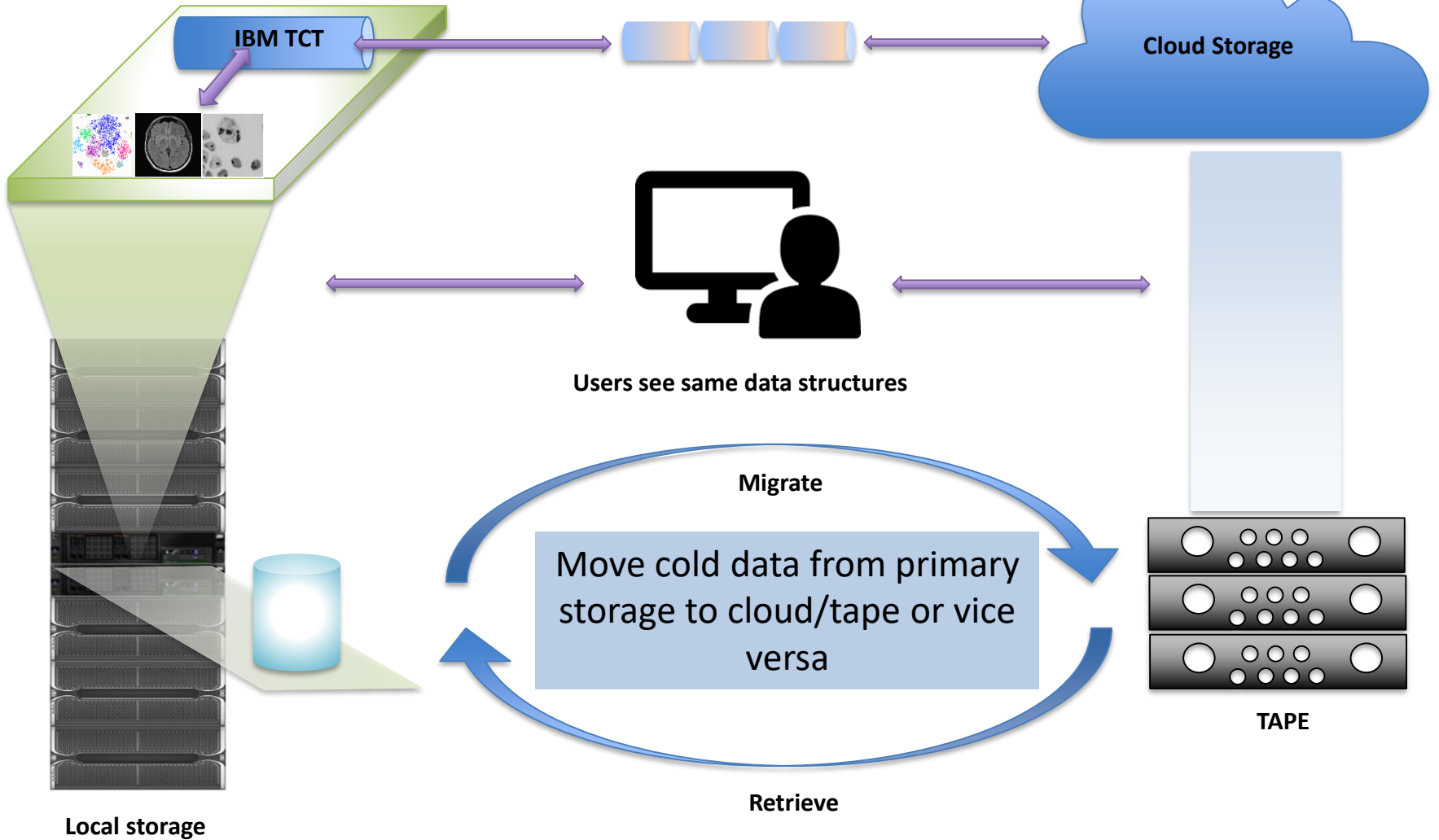
long term archive data

multiple tiers##

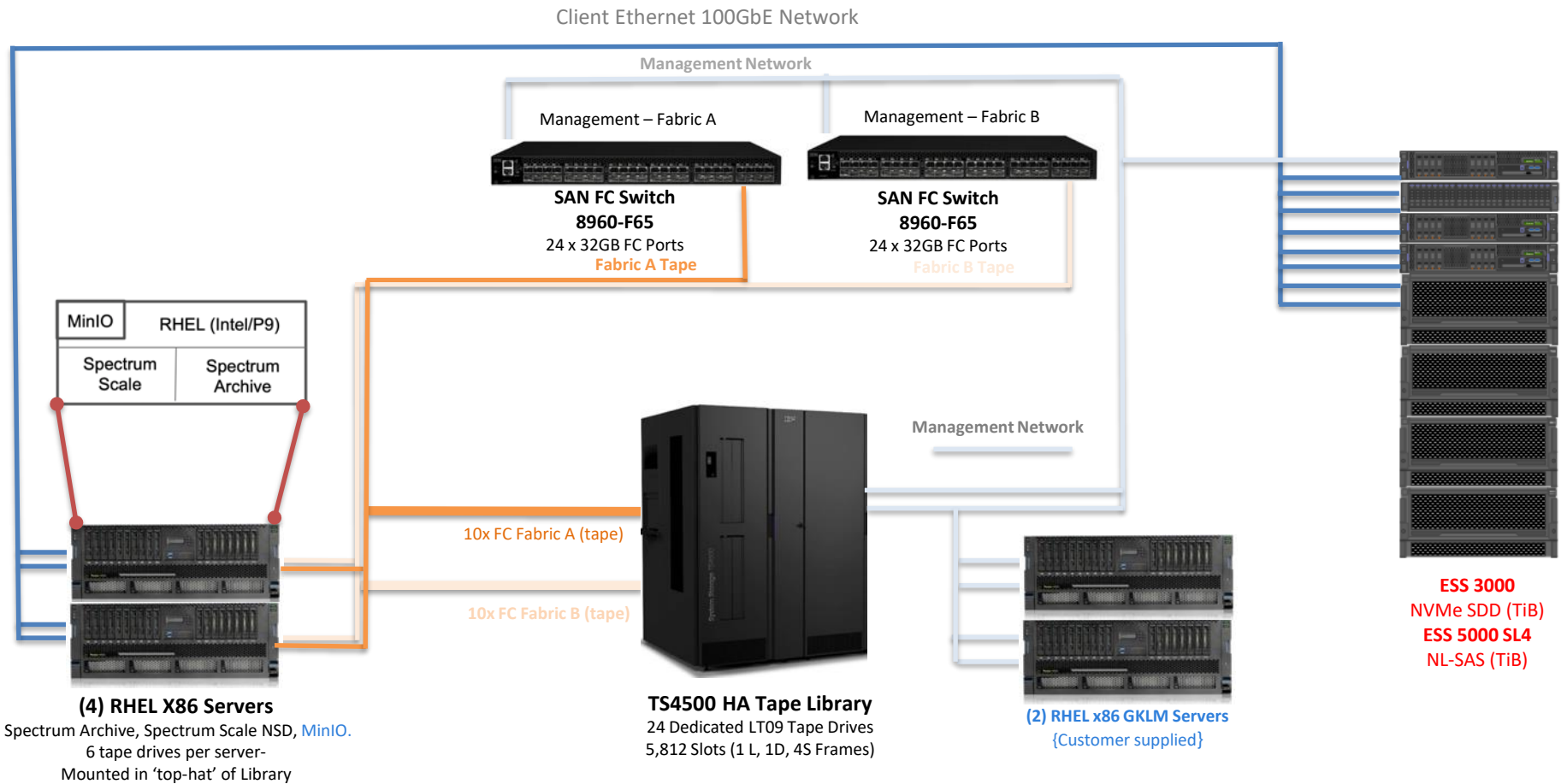
** 20 times faster than /work and /archive); No backup (but PI can request)

if data didn't R/W for one year, move to tape storage system; No backup (but PI can request)

/archive storage -- Single namespace



/Archive storage -- multi-tiers



NVIDIA H100 GPU



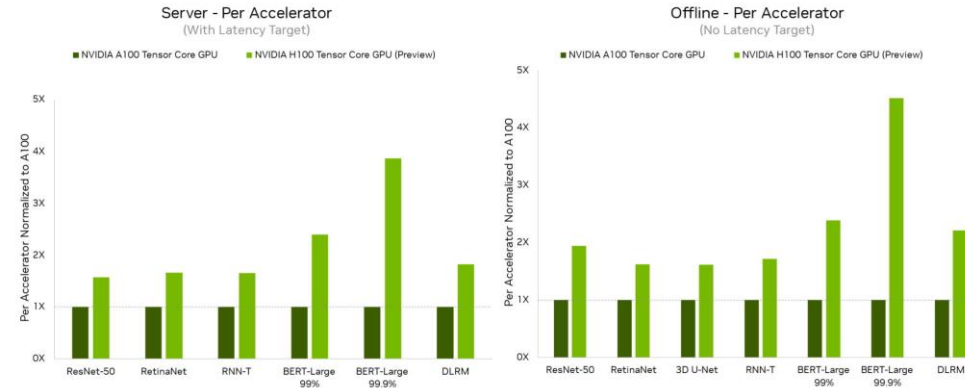
Specs:

Clock speed: 1095 MHz
(boosted up to 1755 MHz)

Memory: 80 GB

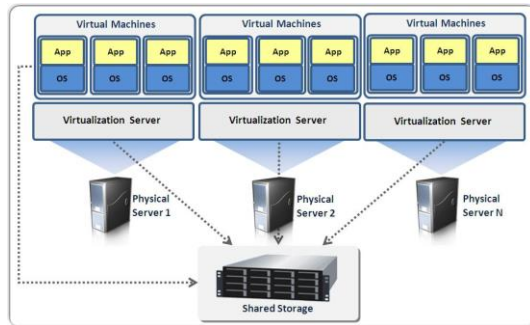
Bandwidth: 2039 GB/s

H100 Supercharges NVIDIA AI Up to 4.5X Faster than A100

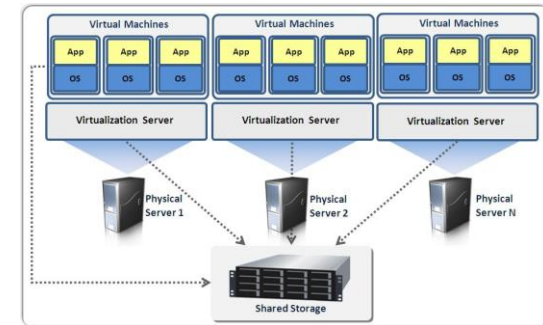
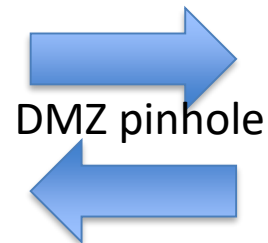


MLPerf Inference v2.1 Cloud: Per accelerator performance derived from the best MLPerf results for respective submissions using reported accelerator count. Format: Submitter, MLPerfID, # of N1, NVIDIA 2.1.0.086, NVIDIA 2.1.0.122, (Benchmark: NVIDIA 2.1.0.086, NVIDIA 2.1.0.122) (RNN-T: NVIDIA 2.1.0.086, NVIDIA 2.1.0.122) (BERT-Large 99%: NVIDIA 2.1.0.086, NVIDIA 2.1.0.122) (BERT-Large 99.9%: NVIDIA 2.1.0.086, NVIDIA 2.1.0.122) (DLRM: NVIDIA 2.1.0.086, NVIDIA 2.1.0.122). NVIDIA 2.1.0.086, NVIDIA 2.1.0.122.

Virtualizing machine farm



Campus



DMZ

- For flexible projects that don't require full nodes
- Getting your research tools out into the publicly-accessible network
- Improving service reliability, frequency of updates, shortening downtimes
- Virtual GPUs, other passthrough hardware available
- **Please ask IR security for the Approval**

Public-accessible workflow system -- Astrocyte



UT Southwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

Astrocyte 2.0.0

Logged in as: s190450

[Astrocyte Home](#) [My Projects](#) [Browse Workflows](#)

[Documentation](#)

UT Southwestern
Medical Center
Lyda Hill Department of Bioinformatics

BioHPC

BioHPC is UT Southwestern's high performance computing (HPC) group, providing access to HPC systems, support and training to contributing departments. BioHPC maintains a 112 node compute cluster, Nucleus, and large-scale storage systems to support the wide range of computational research at the institution.

A significant focus of BioHPC is providing easy-to-use web based access to our resources, tailored to the needs and experience of a wide range of users. Astrocyte is our in-house workflow platform that allows bioinformaticians to easily deploy their workflows to the web, making use of the nucleus cluster to run them while presenting a simple interface to end users.

BioHPC provides example astrocyte workflows here, as well as working closely with other groups to bring their workflows to astrocyte.

Read more about BioHPC on our [web portal](#).

Please email biohpc-help@utsouthwestern.edu with questions or comments about these workflows.

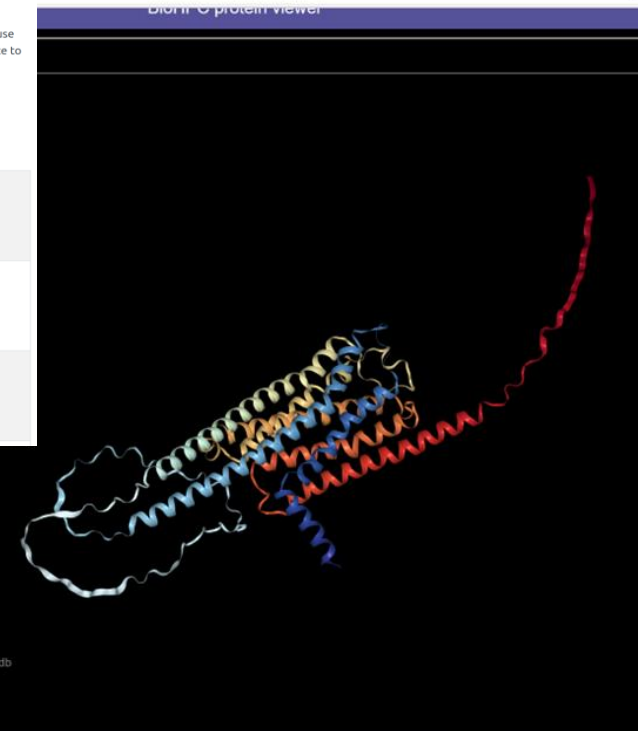
Astrocyte Autodock Workflow Run molecular docking with Autodock and/or Autodock Vina.	Current Version: astrocyte_autodock - 1.2.1 Author: Peng Lian Contact: biohpc-help@utsouthwestern.edu	▶ Run Workflow 📄 Documentation ☰ View Versions
Astrocyte CHARMM-GUI MD Workflow Run MD simulations based on the input files from CHARMM-GUI at https://charmm-gui.org . Build trajectory.vmd for visualizing results.	Current Version: astrocyte_charmmgui_md - 1.2.0 Author: Peng Lian Contact: biohpc-help@utsouthwestern.edu	▶ Run Workflow 📄 Documentation ☰ View Versions
Astrocyte AlphaFold Workflow This workflow is based on AlphaFold 2.1.1 that supports multimer	Current Version: astrocyte_alphafold - 0.0.3 Author: Peng Lian, Xiaochu Lou, Yingfei Chen Contact: biohpc-help@utsouthwestern.edu	▶ Run Workflow 📄 Documentation ☰ View Versions

None Spin Rock

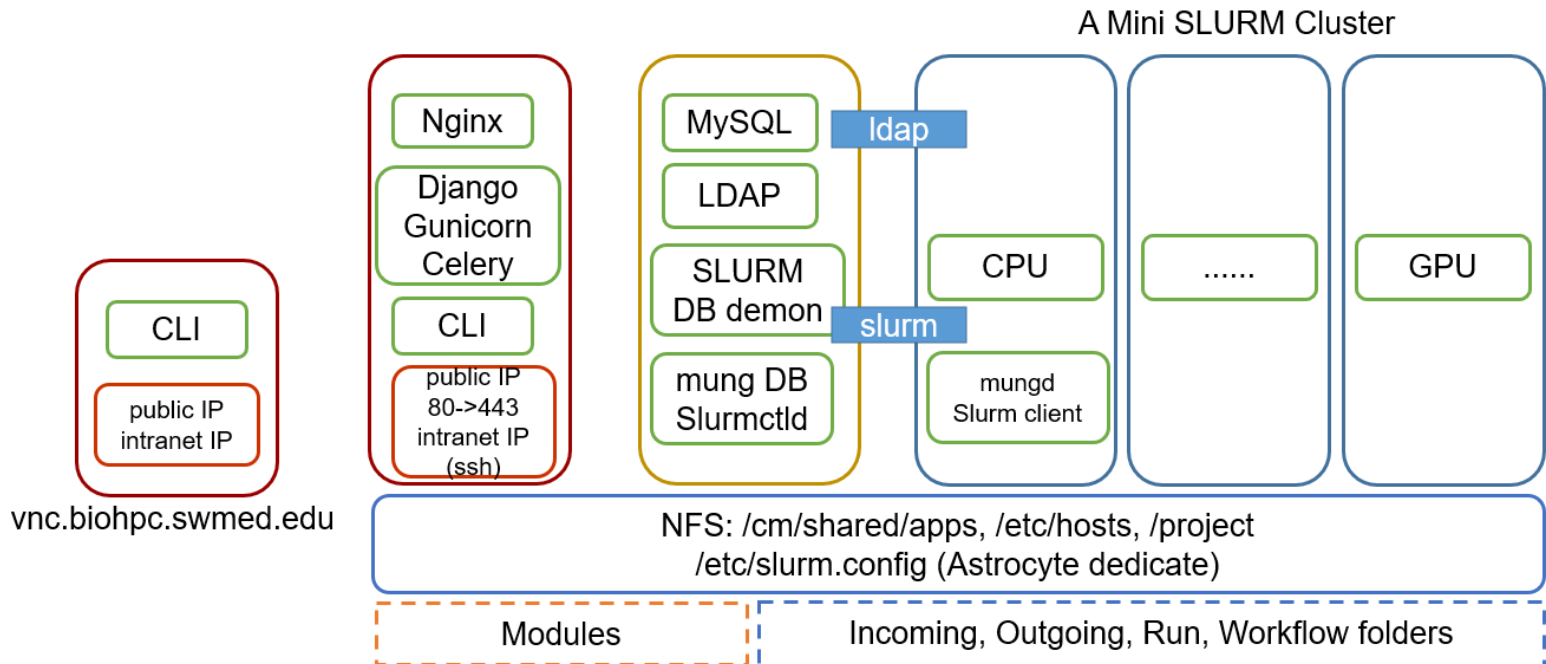
Sequence

Fullscreen

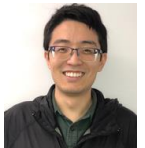
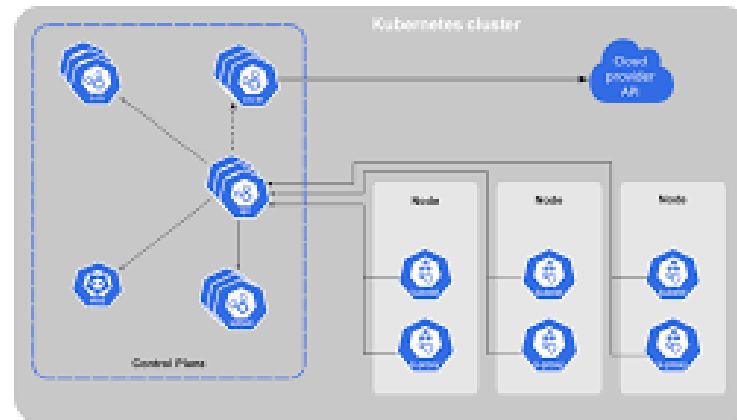
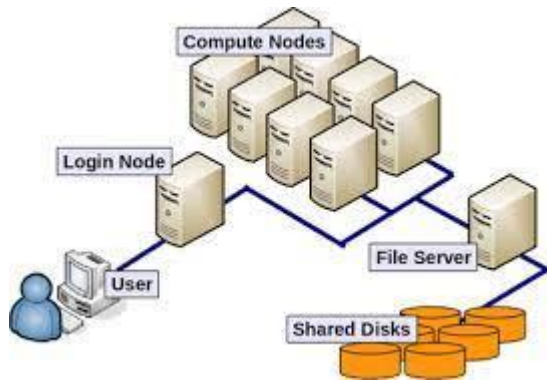
Showing file: output/5HT1A.fasta/5HT1A/ranked_0.pdb



The Astrocyte DMZ Cluster



Hybrid HPC

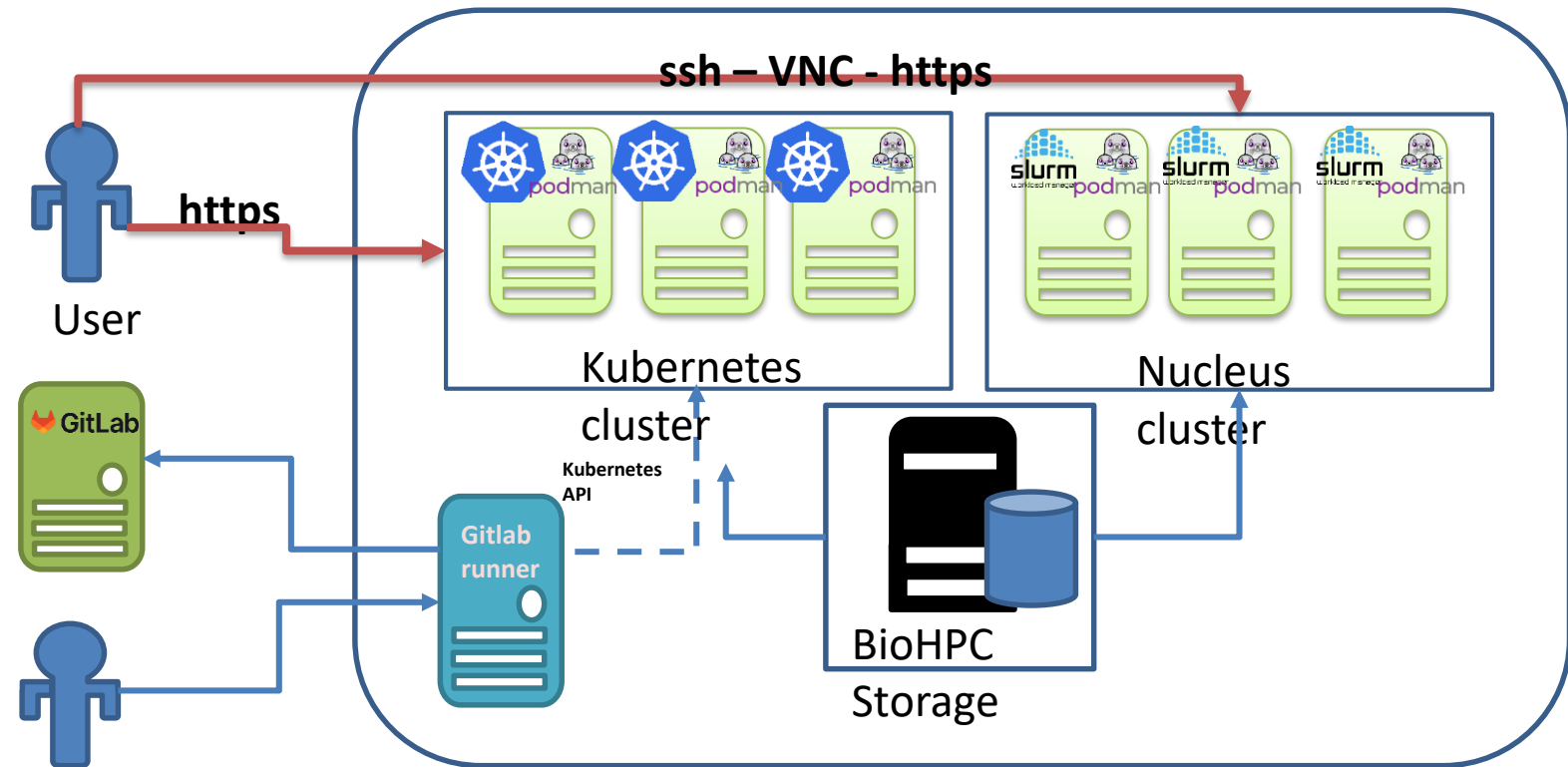


- API server
- Cloud controller manager (platform)
- Controller manager
- etcd (distributed key-value)
- kubelet
- kube-proxy
- Scheduler
- Control plane
- Node



podman

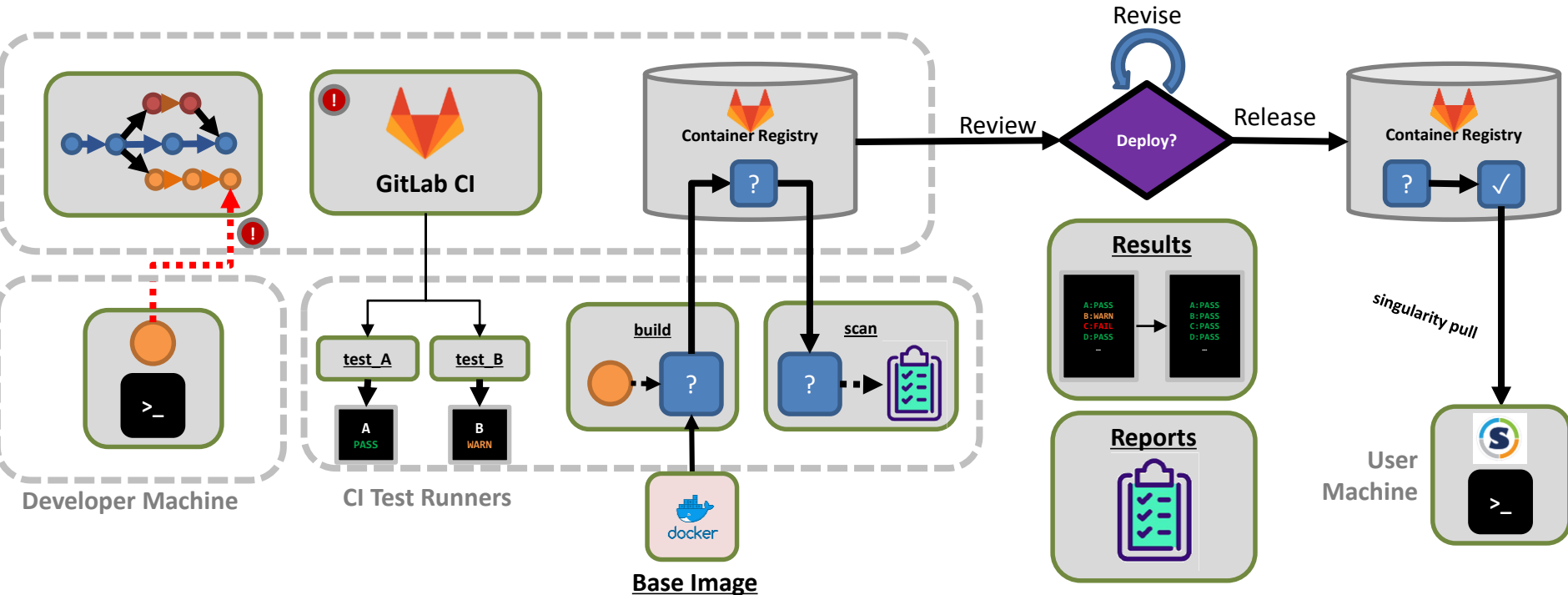
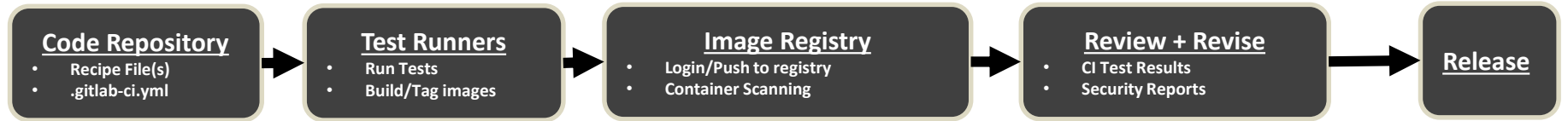
A modern biomedical data management and processing workflow using Kubernetes cluster: Flywheel project



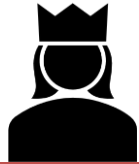
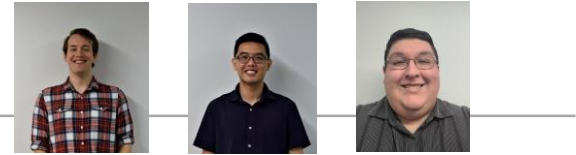
Developer

Hybrid Slurm and K8s cluster with OnDemand access to cluster nodes via browser

Extending Container Infrastructure



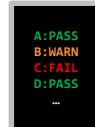
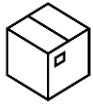
Container Constructor (Coming soon!)



User Browser



Container 'recipe'



Container + Reports

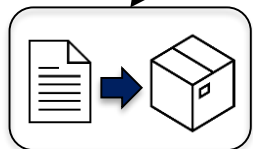


Constructor Service

build

test

scan



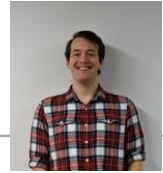
Web-based front end

- Like Web Job Submission or Astrocyte
- Enable users to familiarize themselves with containers

Over time:

- Debugging interface
- Export to e.g. DockerHub, other image registries

BioHPC internship and fellowship program



	Fellowship	Internship
Intended Eligibility	UTSW graduate students, post-doc and trainee scientists – HPC-enabled research projects	Outside (non-UTSW) candidates – less experienced (BS/MS/fresh PhD)
Focus	Developing computational skills/tools to further their lab's research	Exploring and gaining more work experience in HPC – engineering projects
Time	1 day/wk with BioHPC	Full time
Expected Duration	1 year, annual renewal	3 to 6 months

Introducing BioHPC v.2.0

- Townhall for Department Chairs and Administrators Institutional perspective of BioHPC, including revised cost sharing model (Joan C)
 - Founder's perspective of BioHPC, what it is and is not meant for (Gaudenz D)
 - Technical Roadmap (Liqiang W.)
 - BioHPC vis-à-vis RaaS (Russ P.)
 - Administrative procedures (Rebekah C.)

Thank you

