UT Southwestern Medical Center | BioHPC

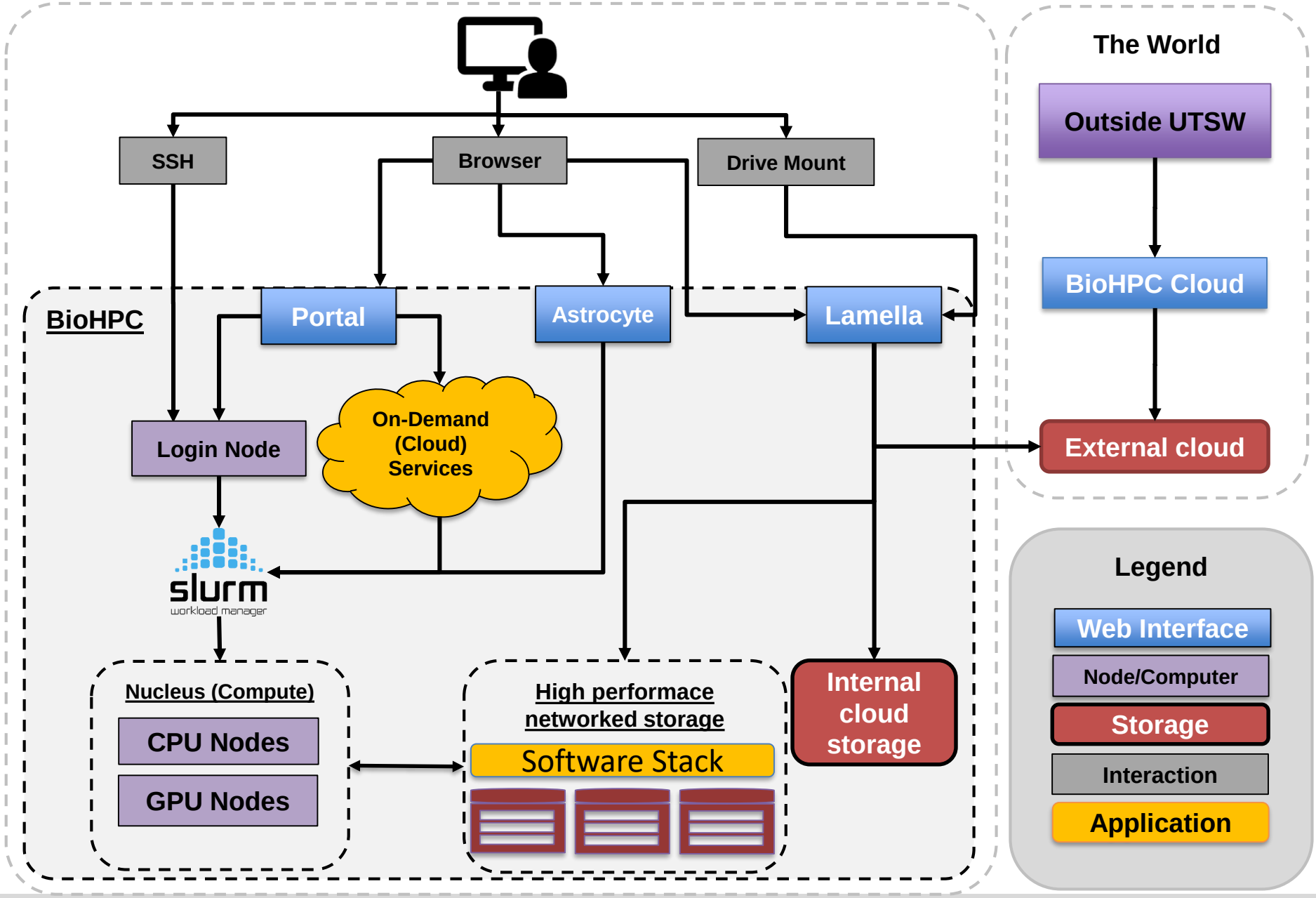
BioHPC OnDemand: Extending an HPC cluster to cloud infrastructures

Paniz Karbasi
BioHPC – UTSW
IBM Spectrum Scale User Group Meeting
SC 2022

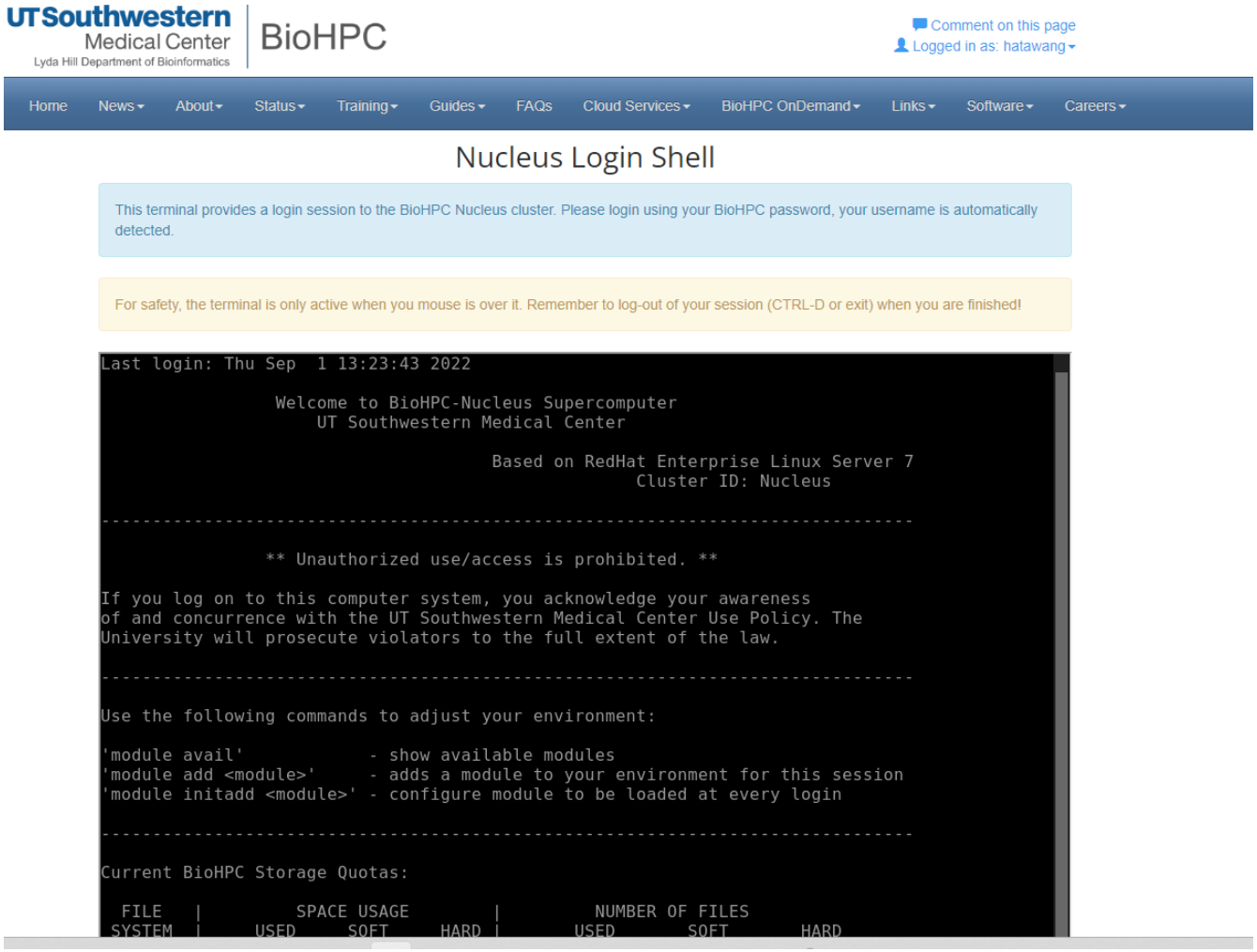
- Cluster name: Nucleus
- Cluster size: 800 nodes of cutting-edge GPU and CPU nodes
- 50 PB of high-throughput storage (360-400 GB/s)
- Two side HPC: 1.6 Tb/s RDMA
- 2000 users
  - Hot users 250 per months

Our vision:

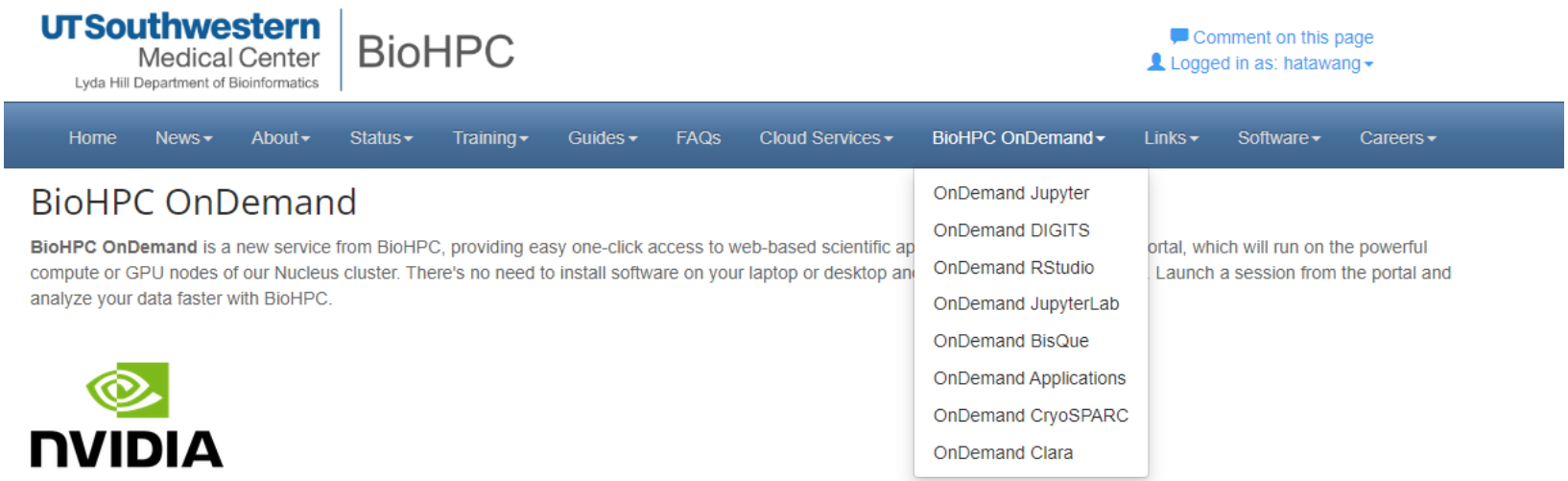
- Make science to be possible at UTSW
- Lower the barrier for scientists to get access to HPC recourses


The World
Outside UTSW
BioHPC Cloud
External cloud
SSH
Browser
Drive Mount
BioHPC
Portal
Astrocyte
Lamella
Login Node
On-Demand (Cloud) Services
slurm
workload manager
Nucleus (Compute)
CPU Nodes
GPU Nodes
High performace networked storage
Software Stack
Internal cloud storage
Legend
Web Interface
Node/Computer
Storage
Interaction
Application

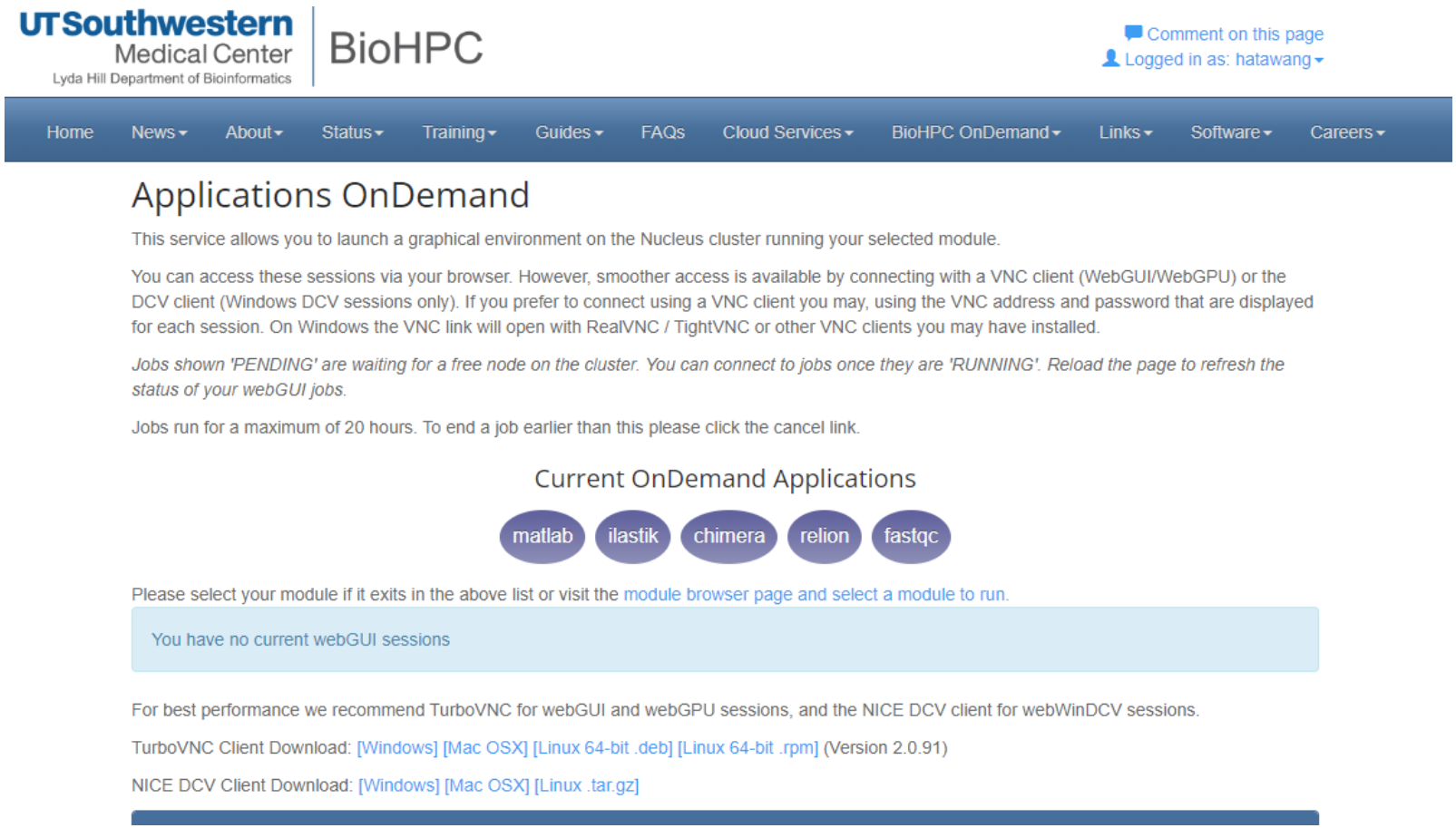
Terminal OnDemand



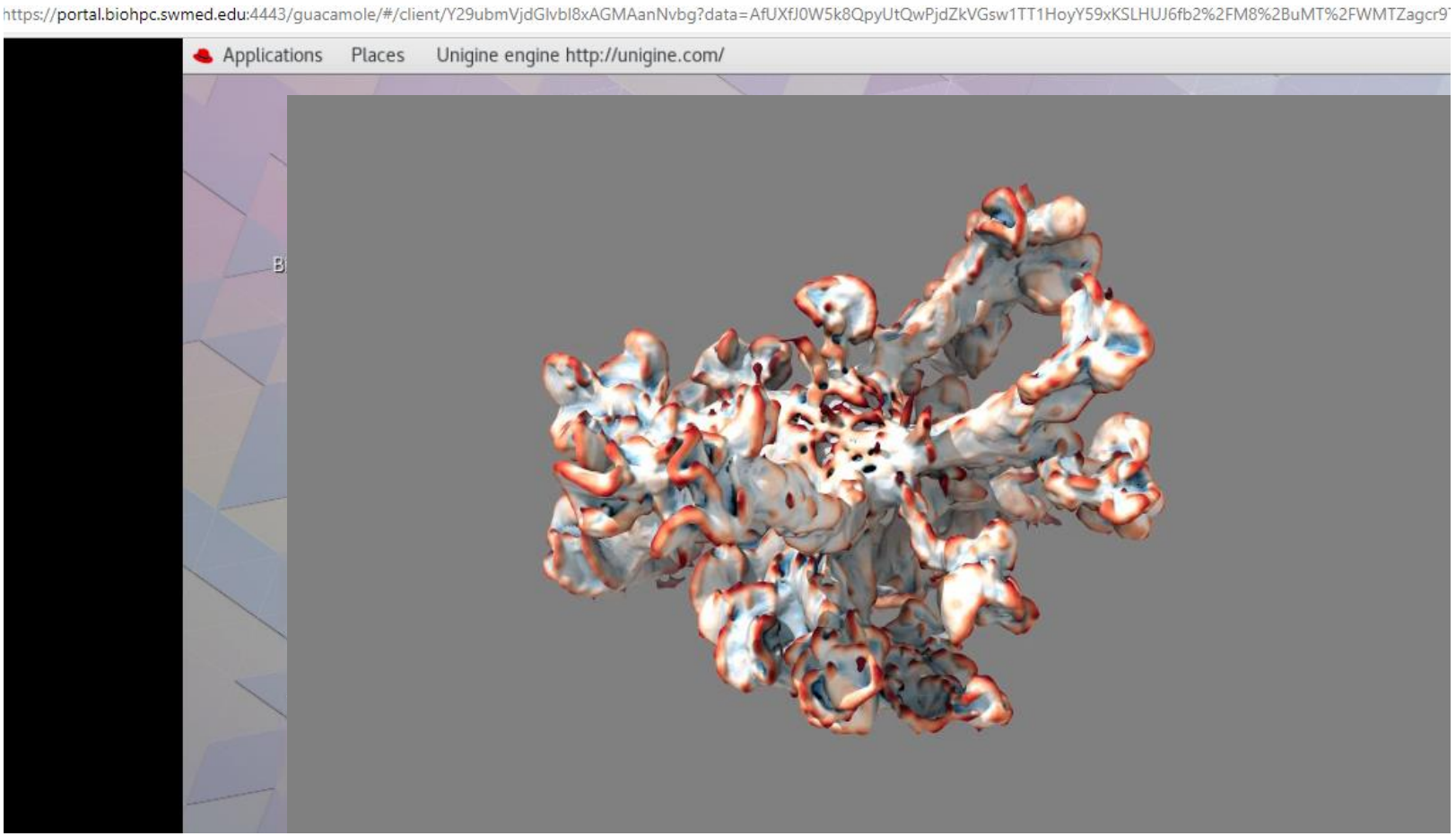
Web application OnDemand



Application OnDemand

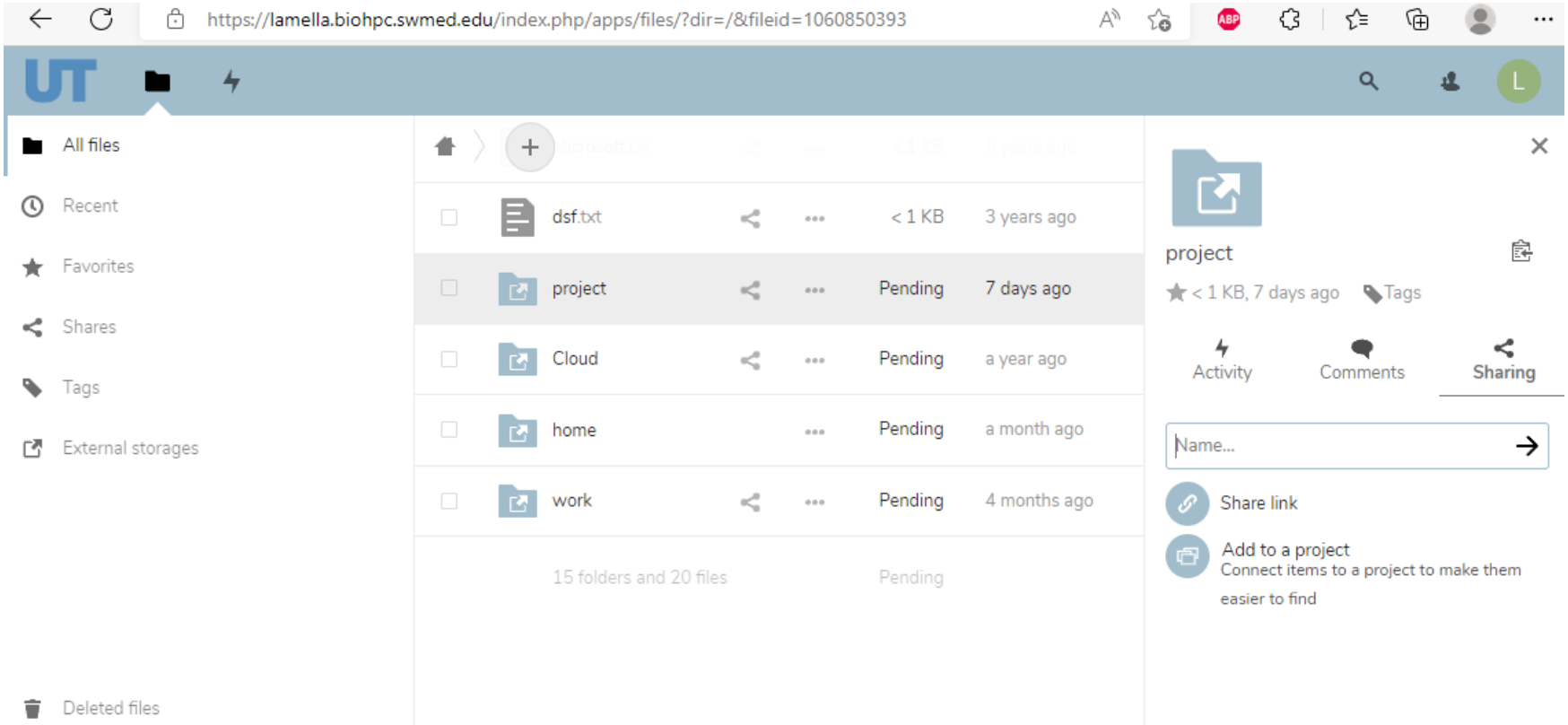


Remote webGPU OnDemand



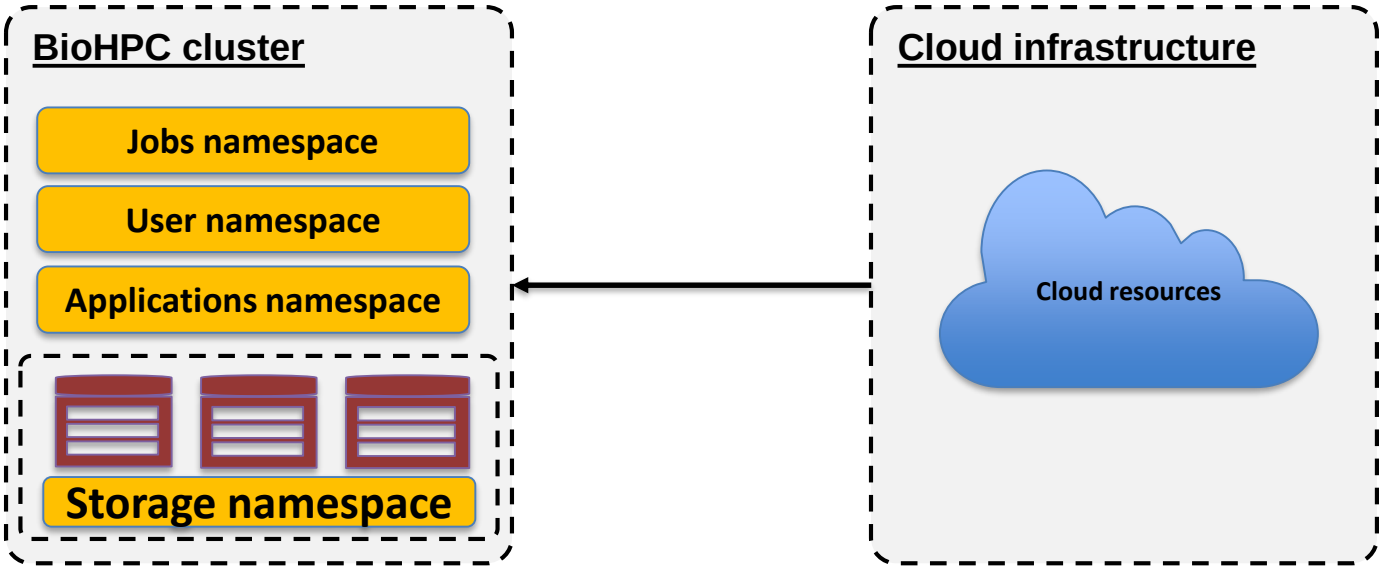
Reference: Drs. Gaudenz Danuser and Meghan Driscoll @ UTSW

Storage OnDemand: Lamella



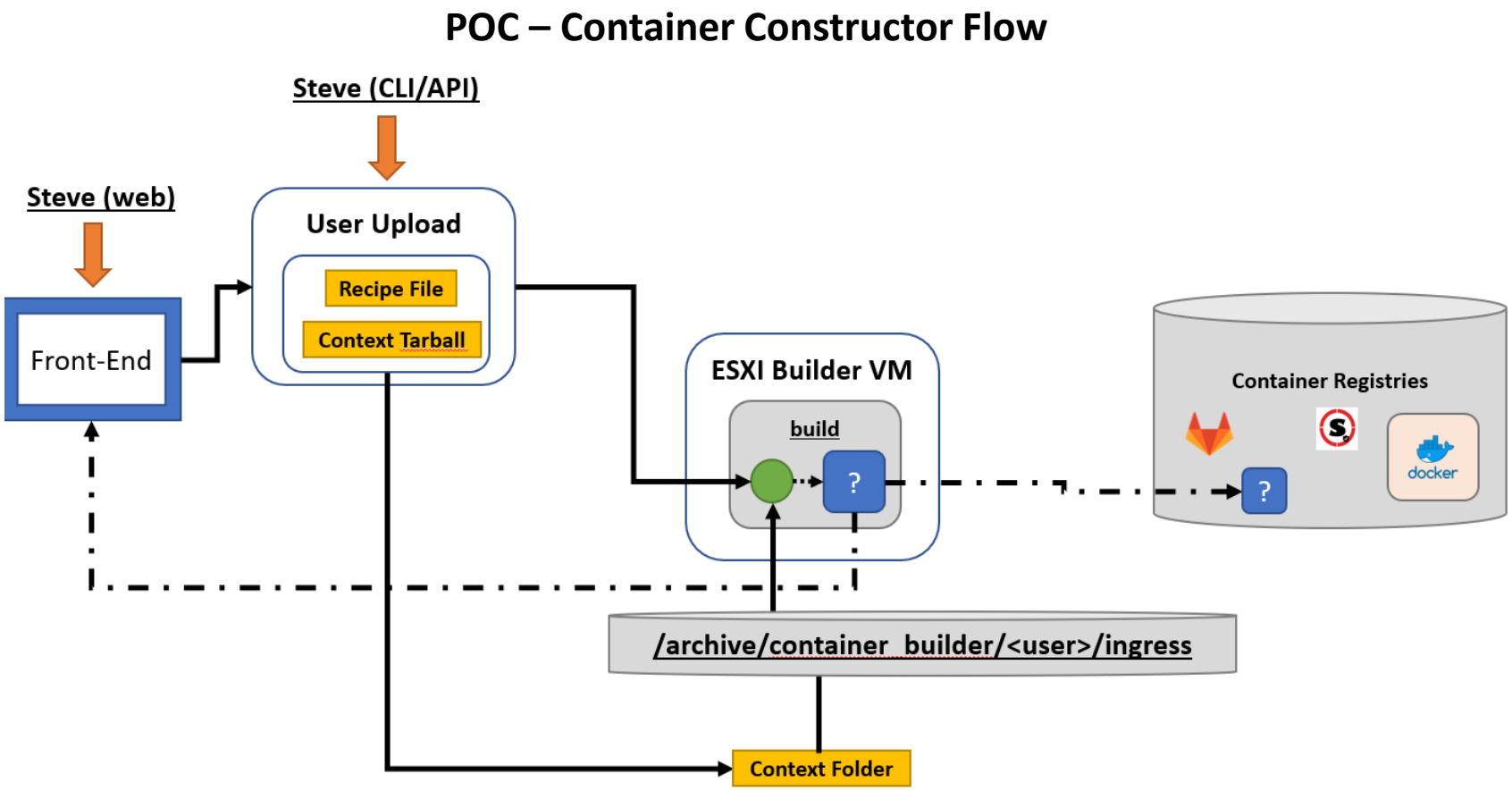
PM: Xiaochu Lou

Extend BioHPC OnDemand to cloud infrastructures



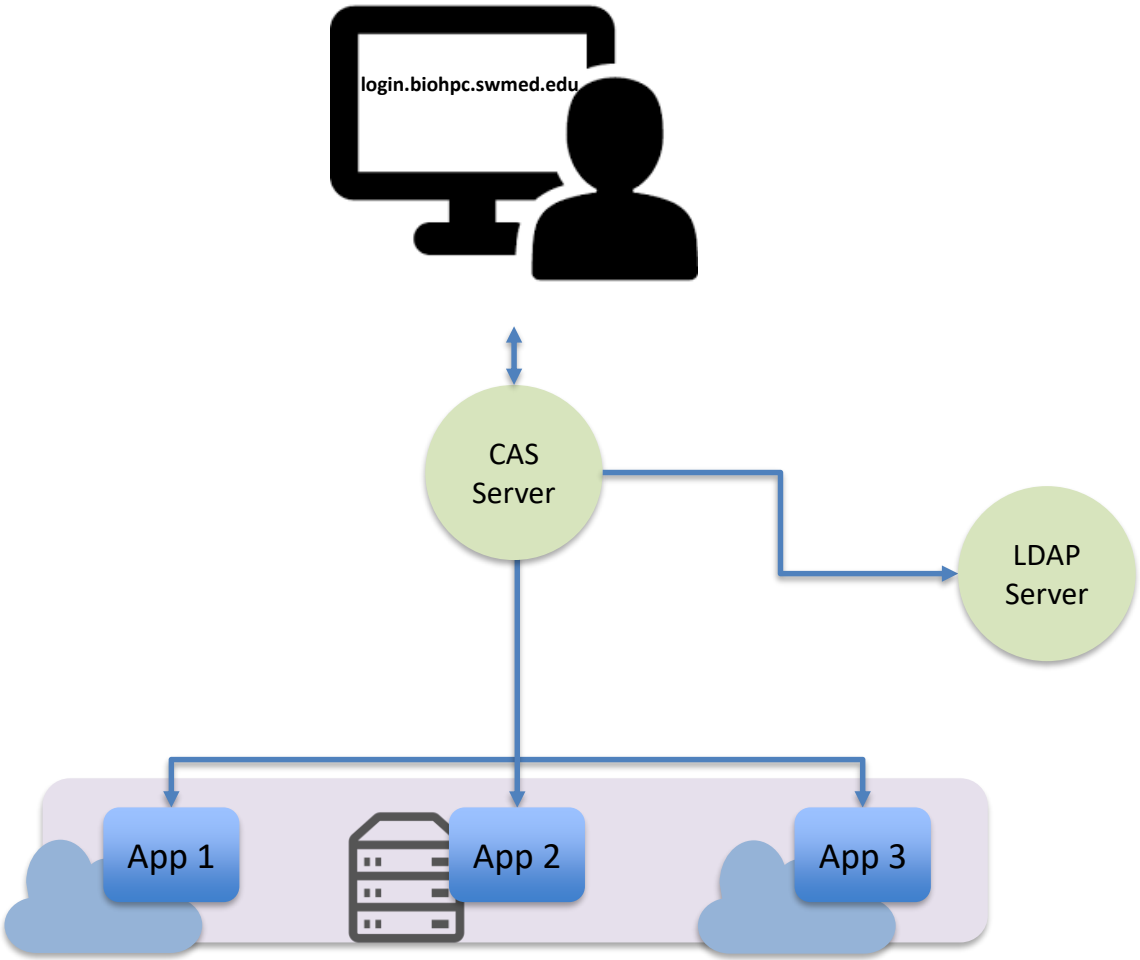
Extend on-premise cluster to outside infrastructure with minimizing refactoring the existing workflow.

Extend BioHPC OnDemand to cloud infrastructures: Application namespace mapping



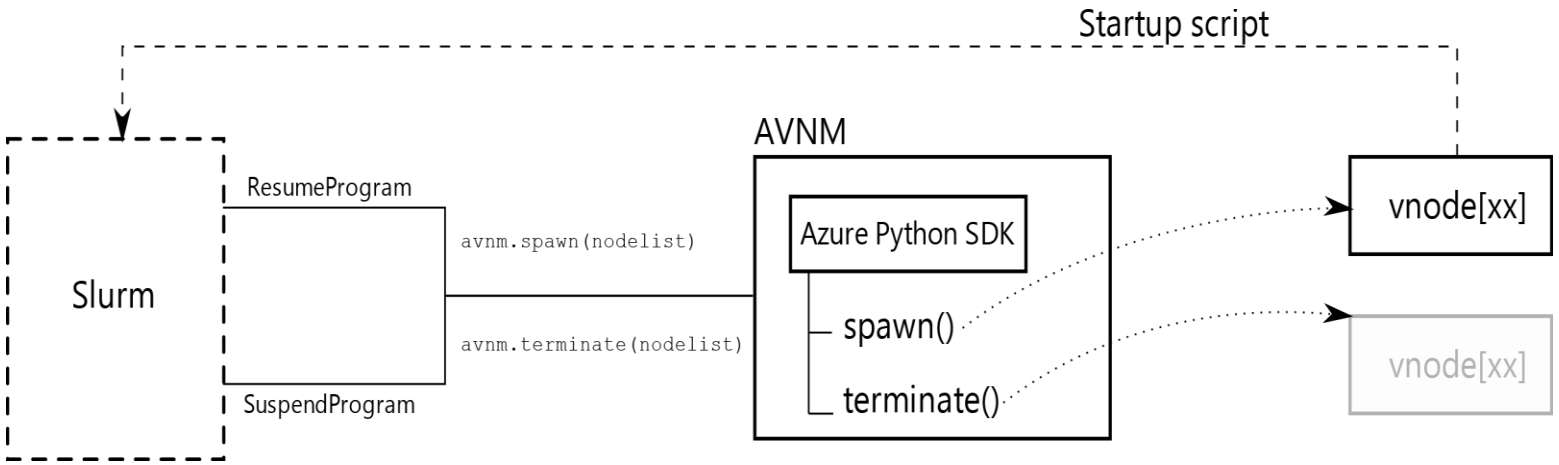
PM: Devin O'Kelly

POC – CAS Integration with LDAP



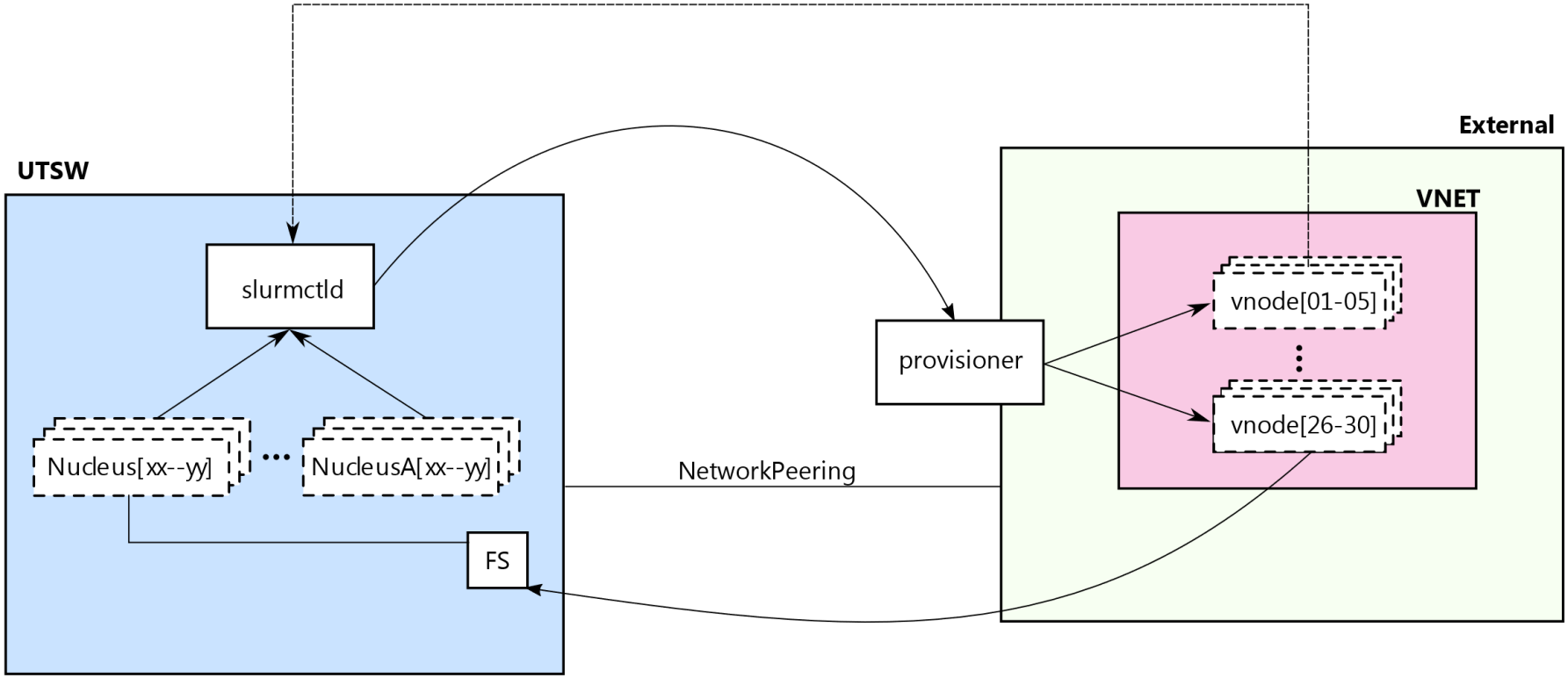
Extend BioHPC OnDemand to cloud infrastructures: Jobs namespace mapping

POC – Cluster Extension to Azure



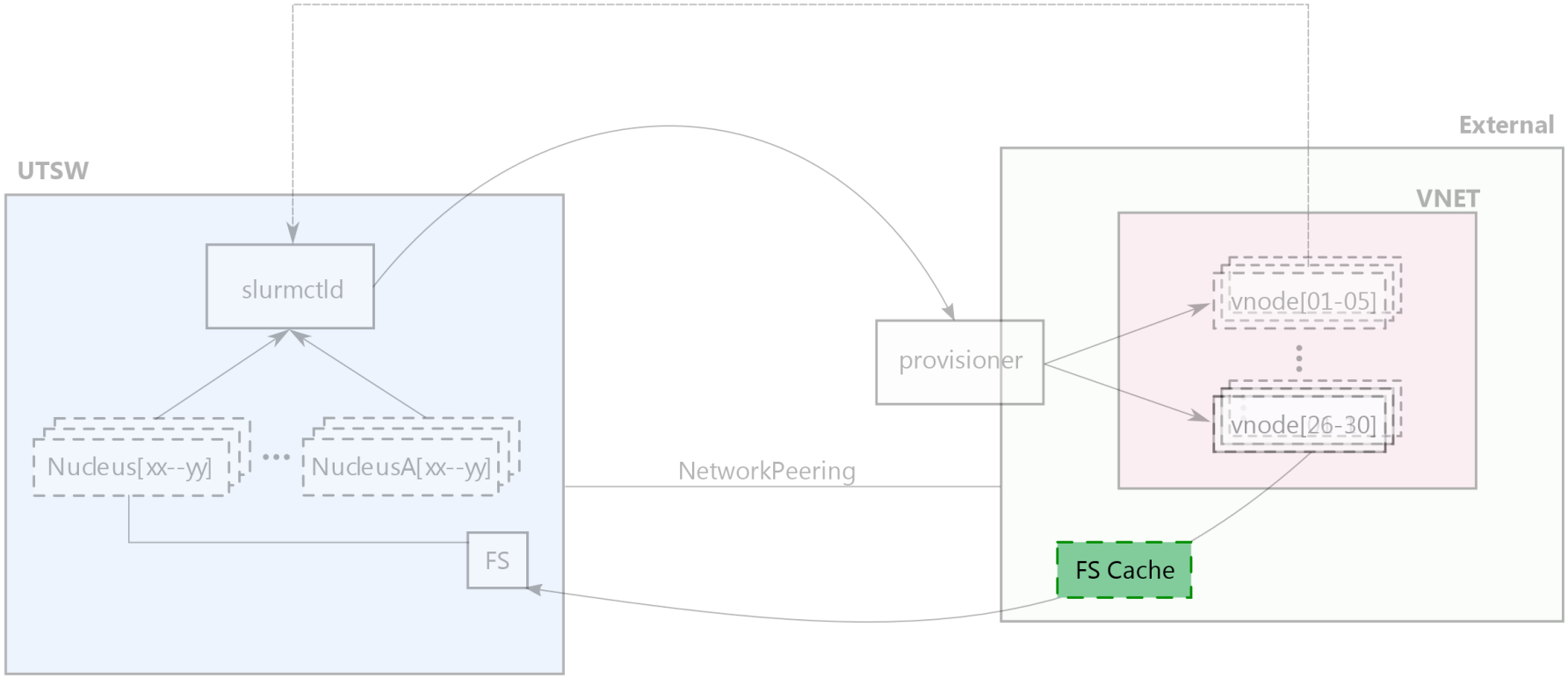
POC - Cluster Extension to any external location

Extend on-premise cluster to outside infrastructure with minimizing refactoring the existing workflow.



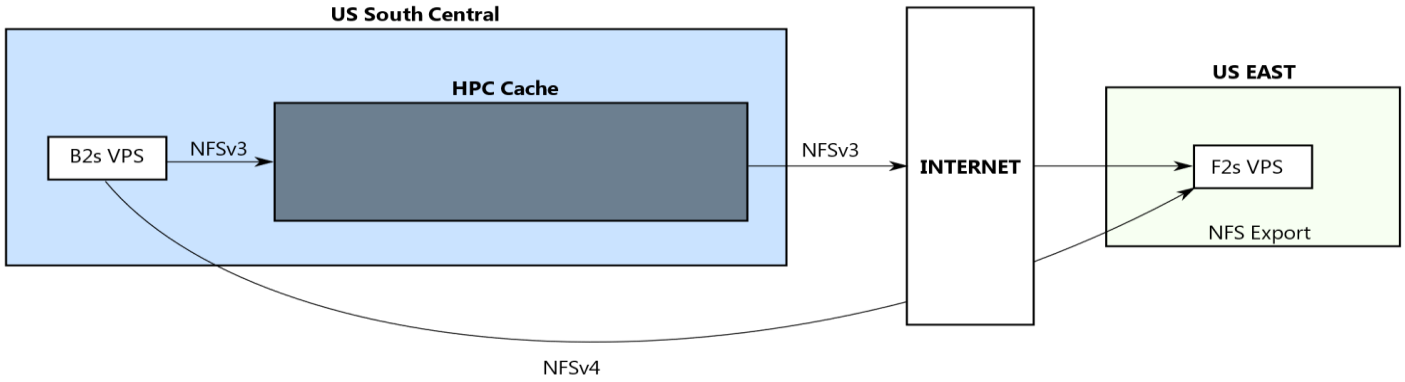
Extend BioHPC OnDemand to cloud infrastructures: Storage namespace mapping

Extend on-premise cluster to outside infrastructure with minimizing refactoring the existing workflow.



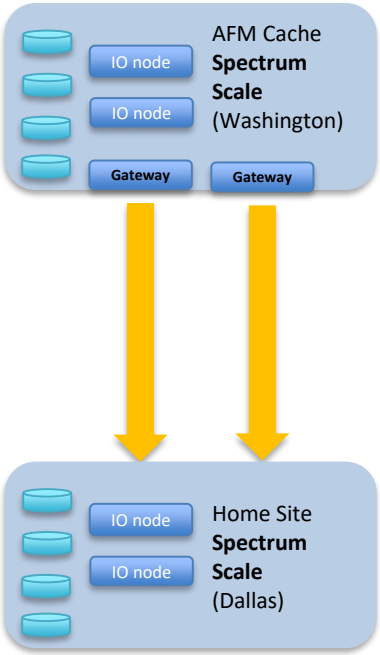
Extend BioHPC OnDemand to cloud infrastructures: Storage namespace mapping via MS HPC cache

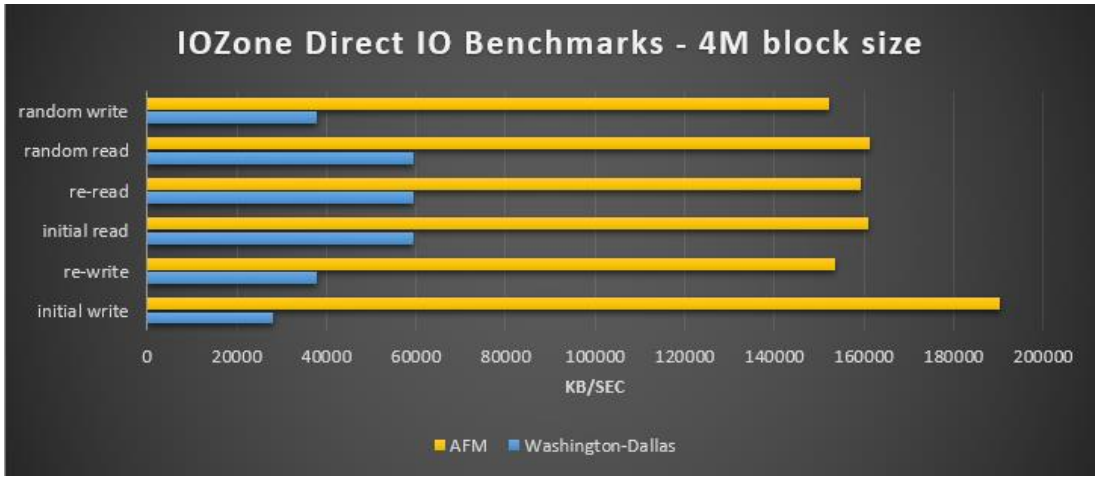
POC – Microsoft HPC Cache



Direct Mount	MS HPC Cache
16.95 MB/s	4.25 MB/s

POC – IBM Spectrum Scale AFM as HPC/FS Cache





IBM Cloud sites: Dallas (primary) and Washington

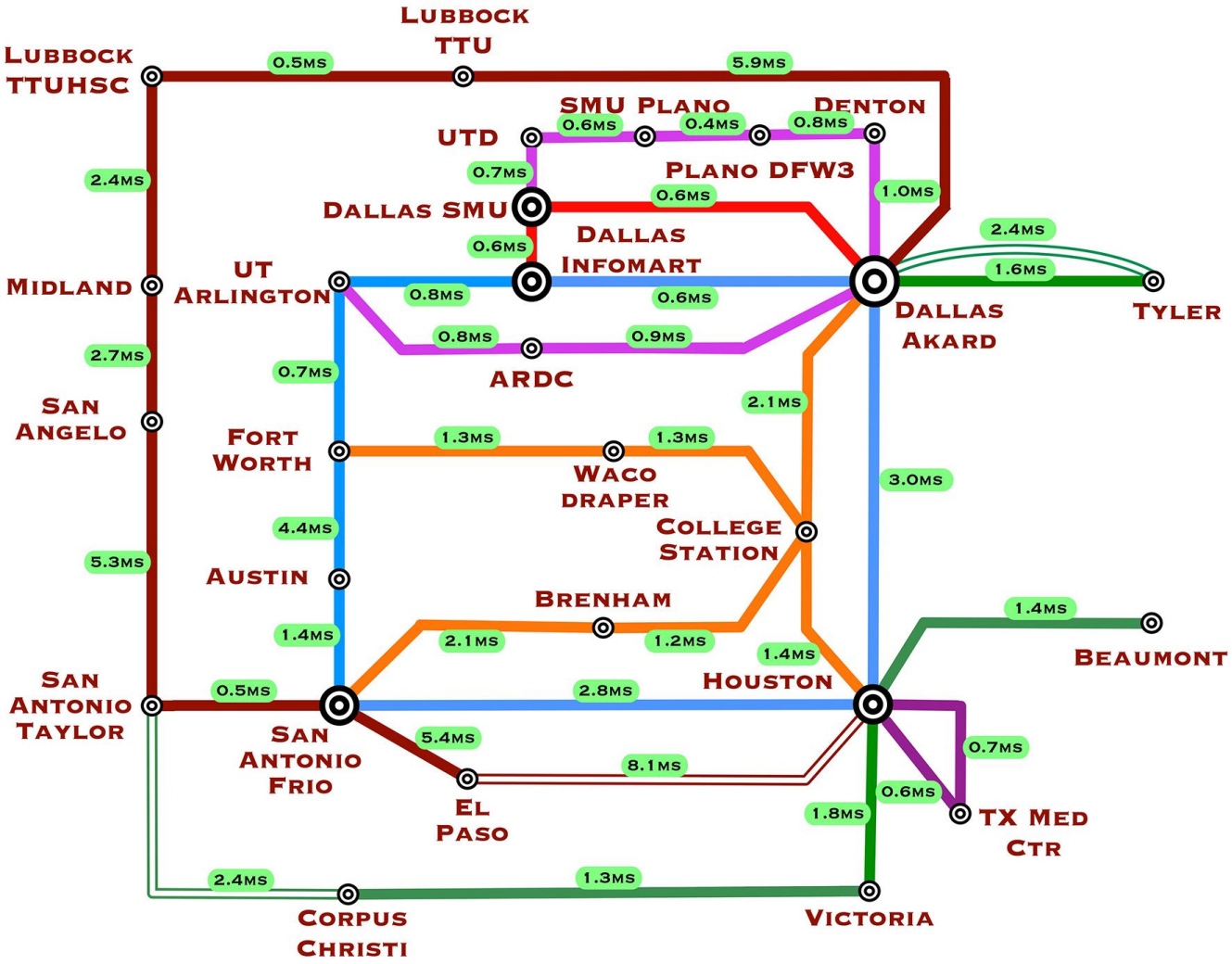
How about the network extension to cloud

Network latency

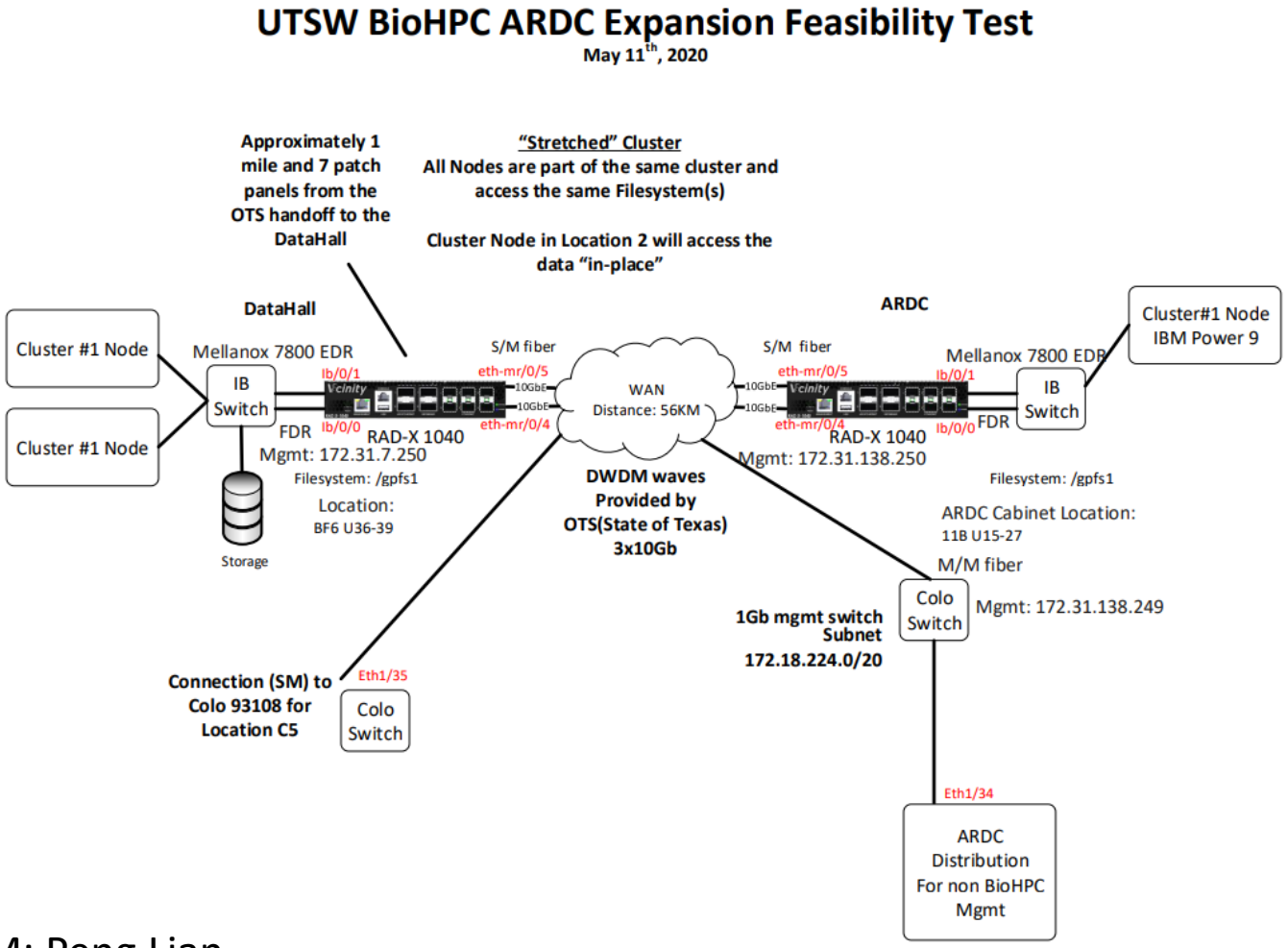
Unit (s)	Multiple	Symbol
10^{-12}	1 picosecond	ps
10^{-9}	1 nanosecond	**ns**
10^{-6}	1 microsecond	µs
10^{-3}	1 millisecond	**ms**

1.3us + 5us/km over fiber
(every kilometer of fiber)

LEARN - Texas State DWDM network



Latency testing with Nvidia – RDMA via DWDM

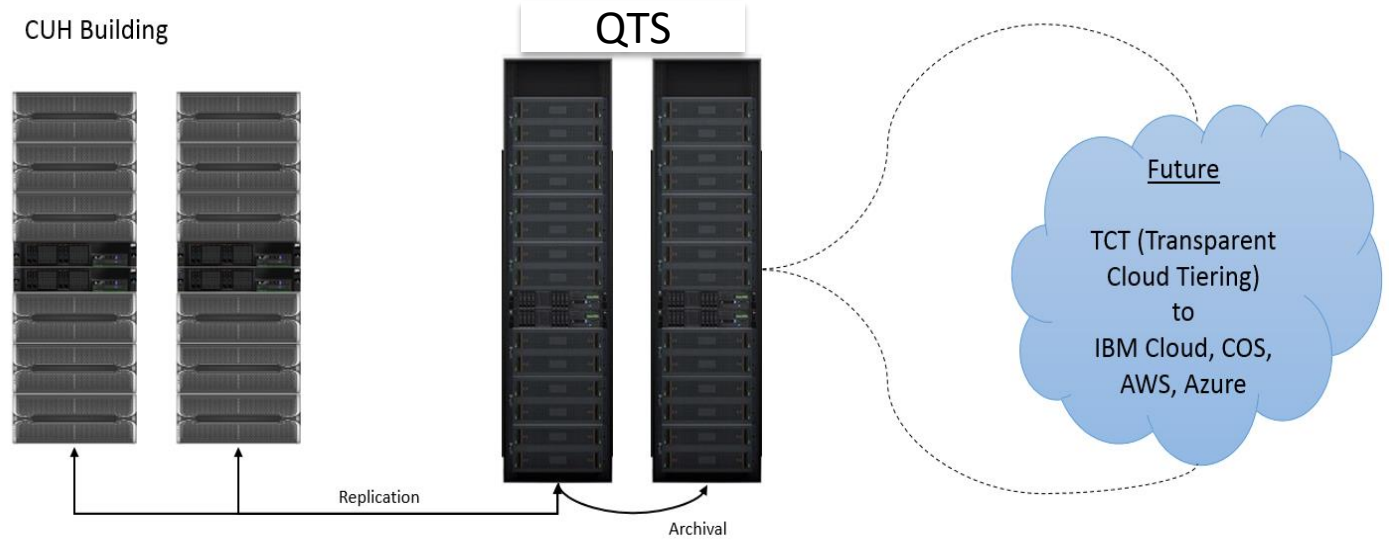


PM: Peng Lian

ARDC Benchmark

Networks	bandwidth tests			latency tests		
	iperf	ib_read_bw	ib_write_bw	ping (64B-64kB)	ib_read_lat (2KB - 1MB)	ib_write_lat (2KB - 1MB)
ARDC						
ib/0/0.0	9.21 Gbits/sec	9.03 Gbits/sec	N/A	N/A	N/A	N/A
ib/0/1.0	9.24 Gbits/Sec	9.06 Gbits/sec	9.06 Gbit/sec	1.42-1.8 ms	1.2-1.7ms	0.6 - 1.6 ms
UTSW						
10GbE	9.86 Gbits/sec	N/A	N/A	0.21 - 0.37 ms	N/A	N/A
Bass IB	43.4 Gbit/sec	48.2 Gbit/sec	46.3 Gbit/sec	0.12 - 0.47 ms	0.0027 - 0.19 ms	0.0014 - 0.18 ms
Bass-CUH IB	21.2 Gbit/sec	36.6 Gbit/sec	36.6 Gbit/sec	0.25 - 0.51 ms	0.033 - 0.26 ms	0.013 - 0.23 ms

POC – Spectrum Scale Transparent Cloud Tiering



Existing Scale/GPFS Environment
- 5PB Standard Licensing
- No Replication or Encryption

Upgrades
- Upgrade to DME
- Add Replication and Encryption

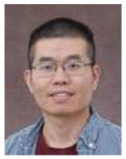
Additional Scale/GPFS Environment
- Additional 10PB Capacity
- Replication and Encryption
- Spectrum Discover (MetaOcean)
  - MetaData Management

Meet the BioHPC team

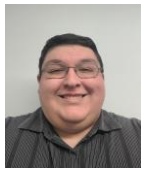
Liqiang Wang
Director

Paniz Karbasi
Ph.D. Electrical and computer engineering

Peng Lian *Ph.D. Biomedical Engineering, Bioinformatics*



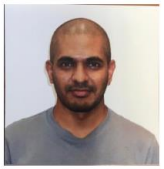
Xiaochu Lou - *Ph.D. Biochemistry*



Felix Perez
M.S, Computer Science

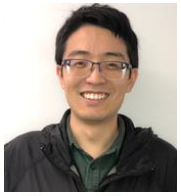
Devin O'Kelly
Ph.D., Biomedical and Molecular Imaging

Hung Le
Ph.D., Computational Chemistry



Suresh Pannerselvam
Ph.D., Medical Sciences

Kenny Huynh
B.S., Computer Science



Xueyan Li
M.S., Computational Geophysics

Find out more about us!

Thank you.