UCLA Health: High-Performance Computing Build-Out

Paul Boutros, Paul Tung, Takafumi Yamaguchi

UCLA Jonsson Comprehensive Cancer Center

UCLA Health Institute for Precision Health

OHIA | Office of Health Informatics & Analytics

Pathway

- ❖ **Once upon a time in the snow**
- ❑ Arriving & Planning
- ❑ Big Ideas
- ❑ What Happened!

UCLA

2018: Happily Embedded in Toronto

Education:

- *University of Toronto (2014-2016)*
Executive MBA, Rotman School of Management

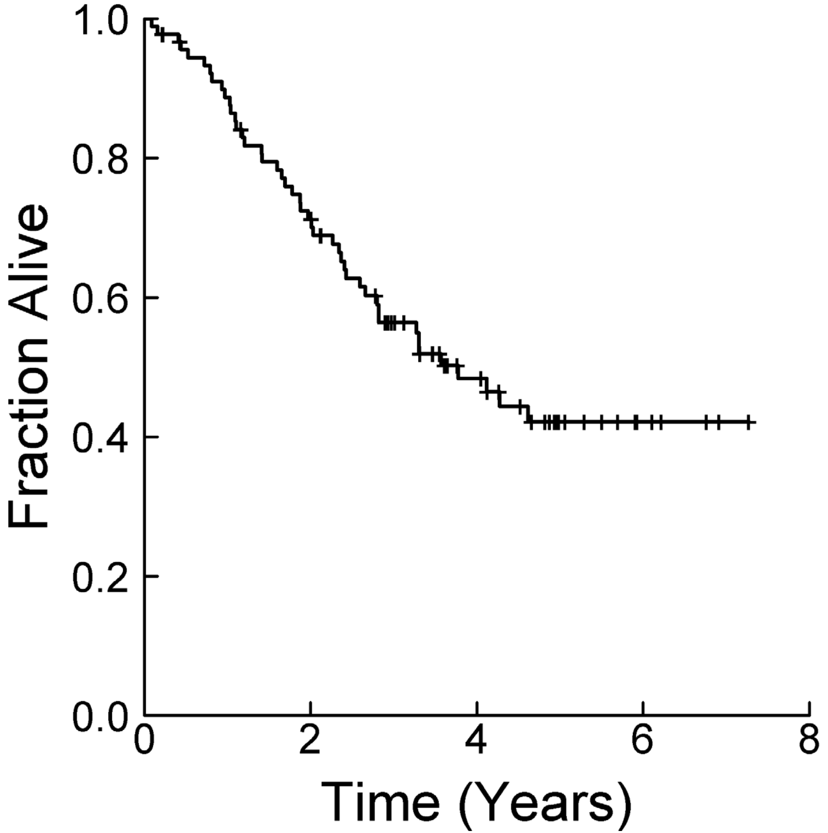
- *University of Toronto (2004-2008)*
PhD, Department of Medical Biophysics

- *University of Waterloo (2000-2004)*
B.Sc, Chemistry, Honors Co-operative Education

40-person data science team
~1 PB of Storage
>5000 CPU cluster
Dozens of active research projects
MTAs/DTAs/IP/etc./etc.

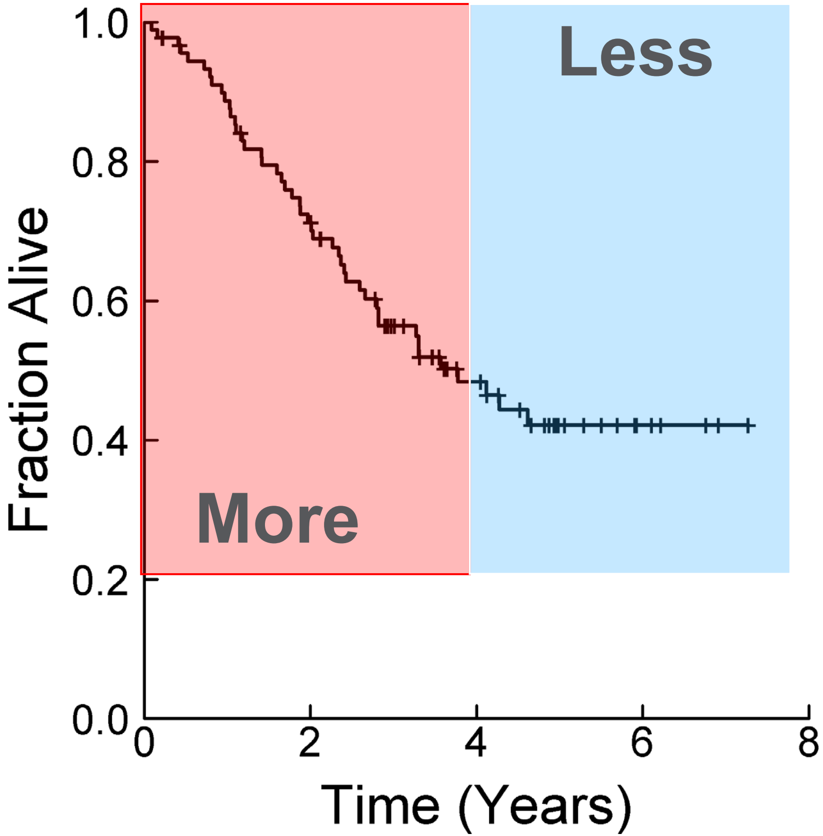
Toronto was home, my lab was running well so... No excitement planned for a few years!

Problem: Localized Disease Outcomes Vary



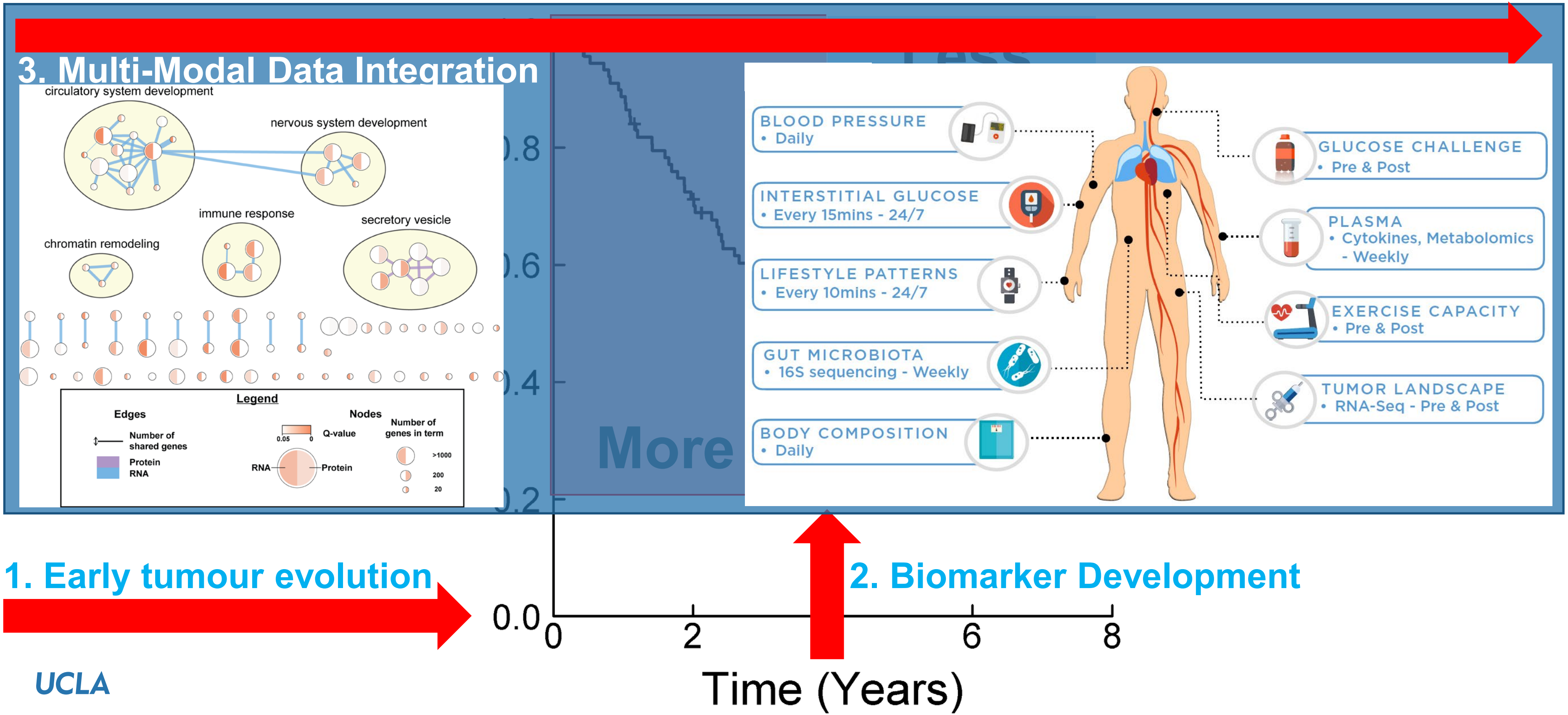
UCLA

A Solution: Prognostic & Predictive Biomarkers



UCLA

My Research: Turning Data Into Insight



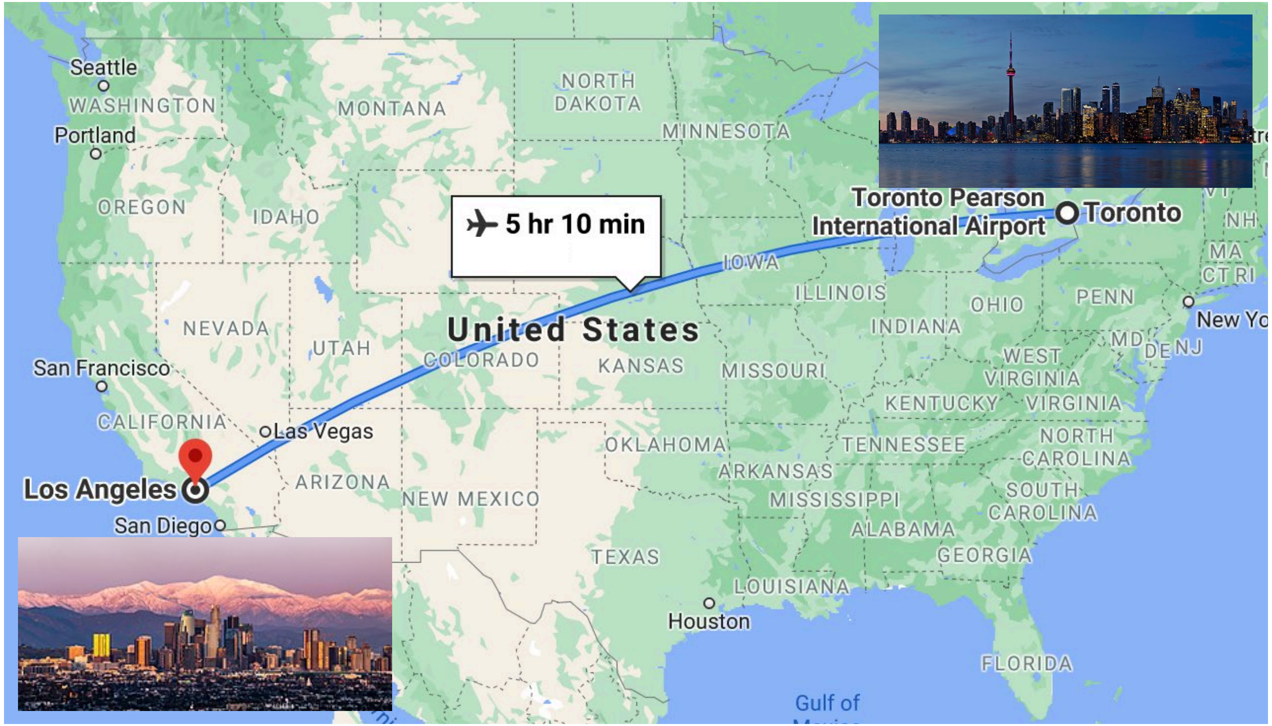
A Decision: Toronto → Los Angeles

Director Cancer Data Science – JCCC

- Create a framework to mine UCLA JCCC's Big Data
- HIPAA-compliant compute
- Cutting edge Machine-Learning
- World leaders in genomics & computational oncology

Associate Director, Cancer Informatics – IPH

- Drive big data innovation from research to clinic and back
- Health system scale, linked to economics & implementation research



UCLA

Pathway

- ✓ Once upon a time in the snow
- ❖ **Arriving & Planning**
- ❑ Big Ideas
- ❑ What Happened!

UCLA

Meeting #0: Compute Planning

Before I arrived

- Started monthly meetings around compute planning
  - Bi-weekly once I arrived
  - Weekly once my team arrived
- Started & completed capacity estimation
- Identifying all data, servers, code
- Thinking about IP – both patents & copyright

Once I arrived

- Meeting #1: DGIT & OHIA for compute planning!
- Priority #1: team coordination infrastructure
- Priority #2: storage to relocate data assets

Pathway

✓ Once upon a time in the snow

✓ Arriving & Planning

❖ **Big Ideas**

❑ What Happened!

UCLA

What Did We Want to Achieve?

Principle

Don’t treat this as a one-off. Set up the framework for a long-term consistent tempo of recruiting computational faculty.

Consequence

Need to productize throughout, and take the time to develop a cohort of skilled staff throughout UCLA & UCLA Health.

Principle

Keep a focus on security: in-depth, systematic, designed in.

Consequence

Close interactions with compliance, infrastructure, *etc.*

Principle

Plan for ongoing technological evolution.

Principle

Develop a close relationship with technology partners.

Specific Needs

Software Environment

- Team coordination software
  - Wiki, issue-tracking
- Software management
  - Import >1-million lines of code & history!
- Standardized software across environments
- Predictable upgrade tempo (quarterly)
- Framework for job-tracking, delegation, prioritization
- Open-source pipelines for standardized analyses
- Routine benchmarking & updating of methods
- Etc.

Compute Environment

- HIPAA-compliant
- Scalable to arbitrary storage
- Support arbitrary compute (GPU, CPU, FPGA, etc.)
- Logging & traceability
- Lock-down of remote access to specific nodes
- Real-time monitoring
- Burstable to arbitrary size
- Productized: easy to reproduce for other uses
- Cost-conscious
- Etc.

Pathway

- ✓ Once upon a time in the snow
- ✓ Arriving & Planning
- ✓ Big Ideas
- ❖ **What Happened!**

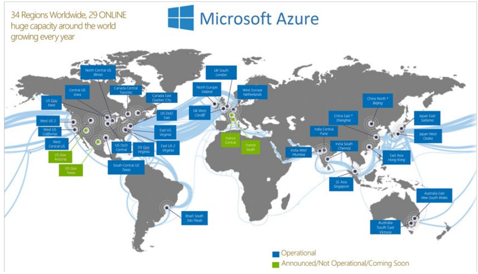
UCLA

Big Ideas to Big Infrastructure

Big Ideas meets Big Infrastructure

- Introduction
- Use Case: HPC Environment for Boutros Labs
- Ideal fit: Cloud (Azure)
- Checked off all the required boxes
- Microsoft acquired some of the best companies
    - Cycle Computing
    - Avere Systems

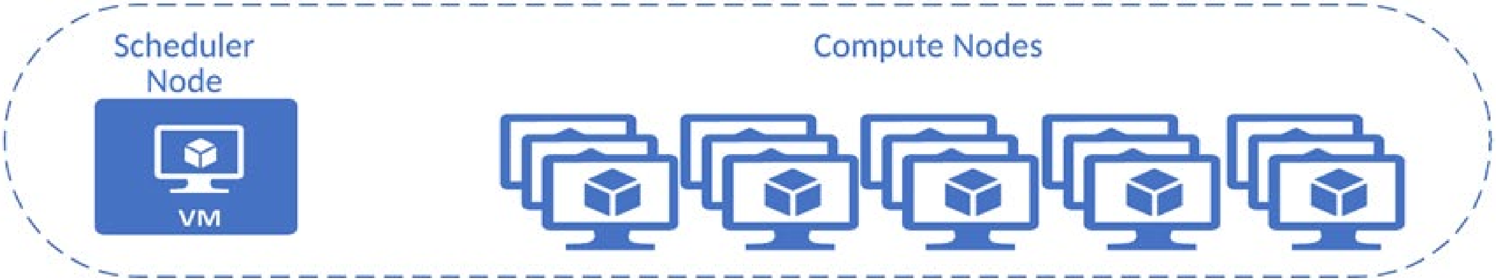
- Azure Cloud Infrastructure



UCLA

What is an HPC Cluster?

- Set of Computing nodes set up to work together to perform complex demanding tasks
- At a minimum, an HPC Cluster contains a scheduler node and compute nodes



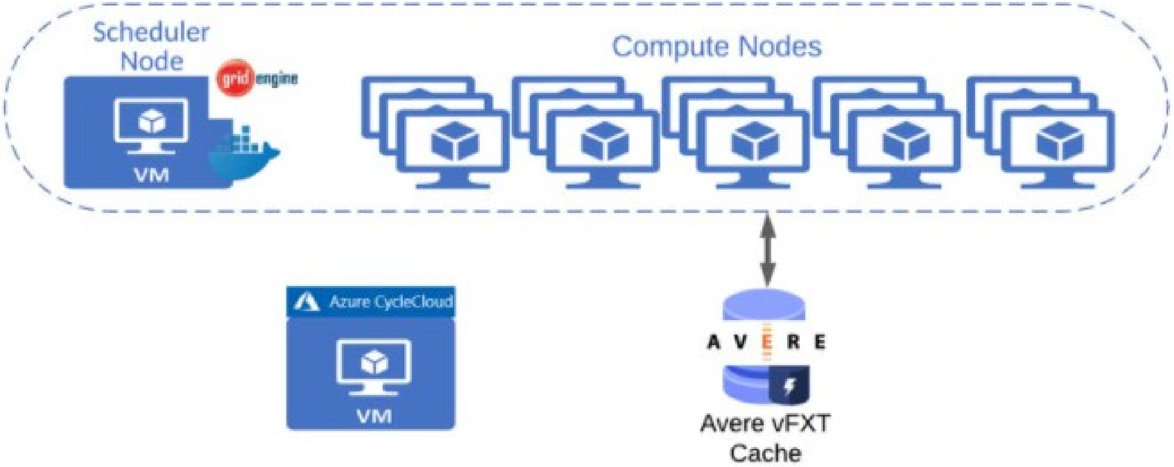
UCLA

Autoscaling

- Ability to:
  - Scale-out to any number of compute nodes required to complete a task
  - Scale-in when tasks are complete.
    - Cluster is down-sized
    - Compute nodes are de-provisioned
- Cost savings as you're only charged for the time the compute nodes actually provisioned.
- CycleCloud works in conjunction with the scheduler to autoscale compute nodes

POC - Preparation

- Worked with different departments within UCLA, DGIT and ISS Security to ensure we had the right resources for the POC
- Microsoft met and worked with us on-site to set up a POC with a CycleCloud Server and Avere Storage Cache in Azure
- Initial Setup
  - ~~SGE~~->Slurm Scheduler
  - Docker enabled
  - 10 compute nodes
  - Small set of users

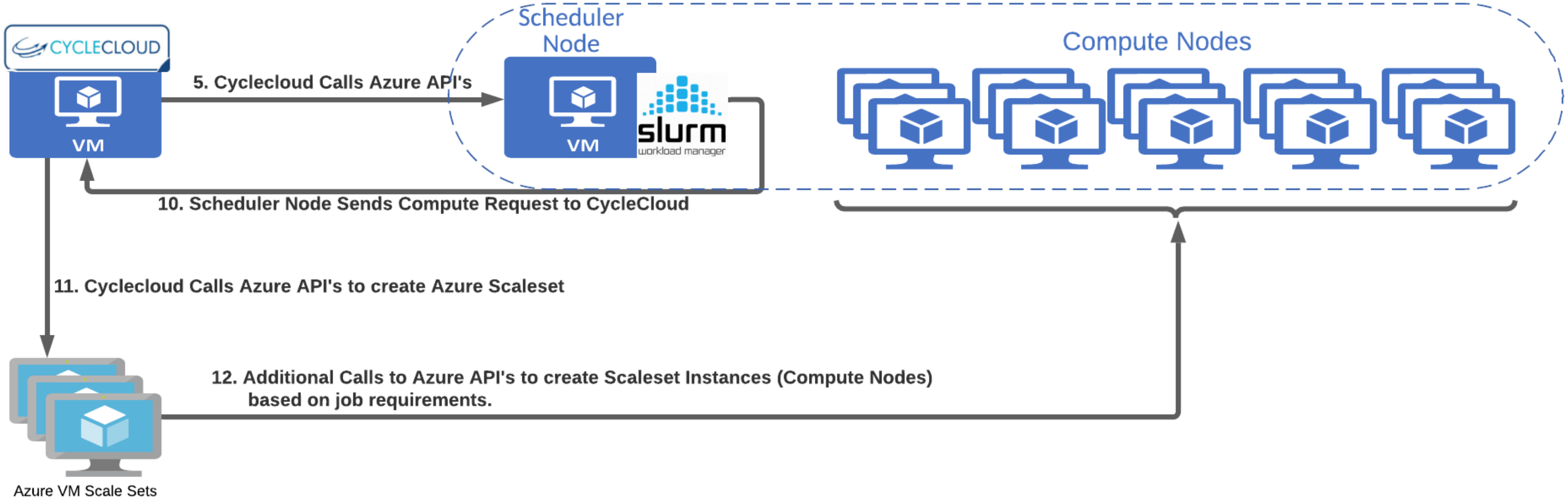


HPC Cluster – Behind the scenes

1. Create Slurm Cluster Template
2. Create Cluster Definition from Template
3. Define Parameters and values
4. Start Cluster

6. Azure Provisions VM for the Scheduler Node
7. Cyclecloud Installs Slurm Scheduler Server
8. Additional software is installed (NextFlow, etc.)
9. User Submits Job to the Scheduler Node

13. Cyclecloud Installs Slurm Compute Software
14. Additional software is installed (NextFlow, etc.)
15. Slurm assigns and executes job scripts each individual Compute Node
16. When job is finished, Compute node is de-provisioned

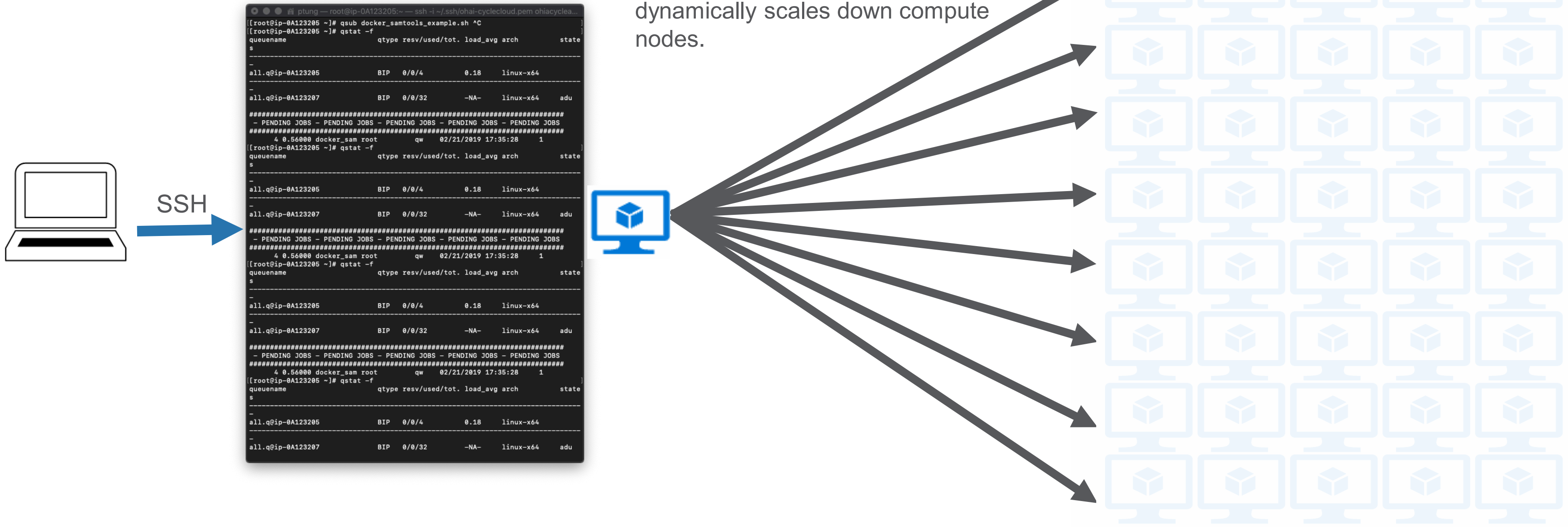


User Experience - Scheduler and Worker Nodes

1. Users Log in to the Scheduler Node

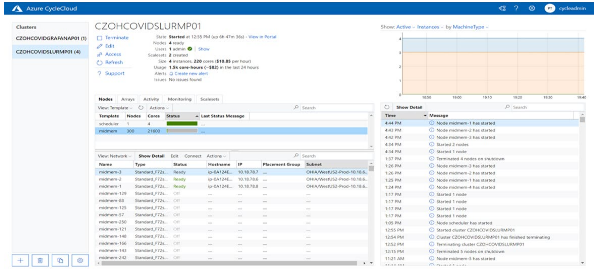
2. Users submit a job to the cluster

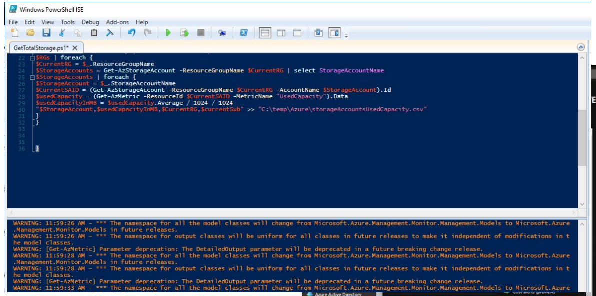
3. CycleCloud dynamically scales up compute nodes to process jobs, then dynamically scales down compute nodes.

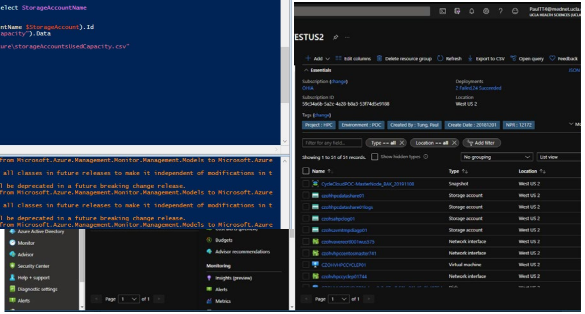


Admin Experience – Cluster Management

- Cluster Management and Monitoring:
  - Cyclecloud
    - Admin Portal
    - Cyclecloud CLI
- Slurm Administration: SSH Client
- Supporting infrastructure (HPC Cache, MariaDB,etc) :
  - Azure portal
  - Azure CLI







UCLA

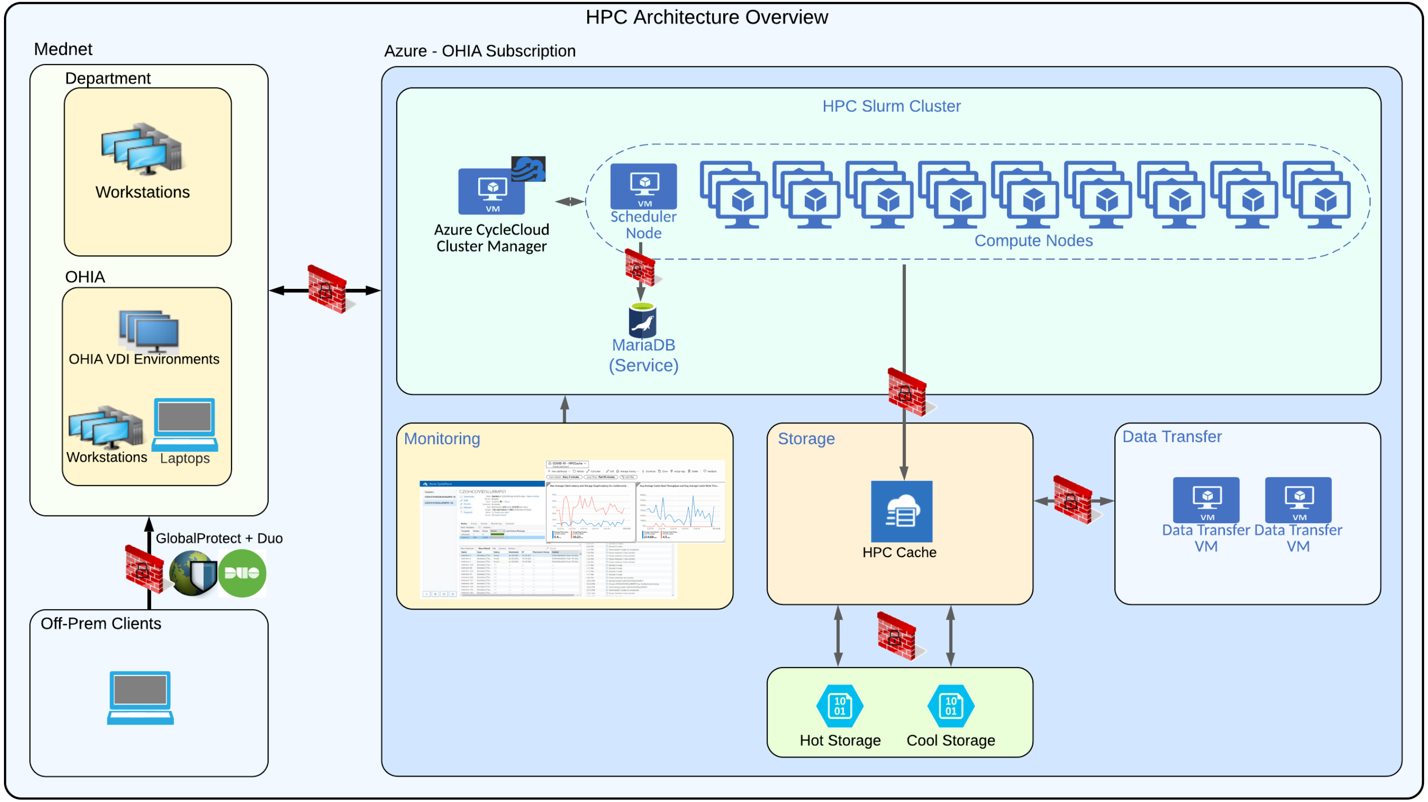
Platform Features

- Choice of VM sizes depending on workload – Multiple VM sizes in the same cluster (Partitions)
- Integrated ability to use source repos and Docker
- Custom software suites and installations
- Ability to handle scaling to extreme capacities (21,600 cores)
- Multiple storage mounts and tiered storage
- Custom scratch space and RAID arrays for better performance
- Performance Monitoring Dashboards
- Stability and efficiency enhancements

Where we’re going in the future

- Blueprint for HPC Platform
  - Repeatable
  - Scalable
  - Modular
- Scaling to sizes beyond 21,600 cores
- Automating the deployment of the entire infrastructure
- Working on CI/CD processes to create custom images:
  - hyper-streamlined image vs general purpose

HPC Platform – Blueprint



UCLA

Perspective

Entity	Type Prod/Non-Prod	Total Clusters	Cluster Type	Compute	Date
UCLA Health HPC Clusters	• 1/0	• 1	Cloud	• 720 Cores • Storage – 250TB	• February 2020
UCLA Health HPC Clusters	• 3/1	• 4	Cloud	• 60,000+ Cores • Storage – 2PB	• May 2021
UCLA Hoffman HPC Cluster	• 1/0	• 1	Physical	• 21,000 Cores • Storage - 50TB	
Berkeley Research Computing HPC	• 1/0	• 1	Physical	• 15,300 Cores • Storage – 20TB	

UCLA

UCLA Health HPC Platform

Use Case	Prod Clusters	Test/Dev Clusters	Description	Compute Needs
Boutros Lab Production	2	1	• Slurm • Azure HPC Cache	1. 12,000 cores target 2. Spread across two computing clusters 3. 20% annual expansion (2022 and beyond)
Boutros Lab COVID Use Case	1	0	• Slurm • Azure HPC Cache	1. 14,400-21,600 cores target 2. <1 year
IPH Regeneron	1	1	• Slurm • Azure HPC Cache	1. 5,000-10,000 cores target
MDL HPC	1	0	• Slurm • Azure HPC Cache	1. 100 cores target
IPH HPC	1	1	• Slurm • Azure HPC Cache	1. 5,000-10,000 cores target
DGC HPC Analytics Platform	1	1	• Slurm • Azure HPC Cache	1. 5,000-10,000 cores target

Conclusion

- Always on the cutting edge, pushing the limits of the technology.
- Along the way, we found a lot of opportunities to partner with Microsoft to improve service offerings to enhance the user experience and technology
- We'll also continue to refine our technology offerings in partnership with our customers.

Big Infrastructure to Big Data

Big Data needs Big Infrastructure

- Our mission is to cure cancer using “Big Data”
    - Develop biomarkers and personalized treatment options
- HPC is our key infrastructure
    - Download, process, analyze and store "Big Data"
    - Use large-scale high-throughput molecular datasets
- “Big Data” gives us efficiency
    - Understand and optimize processes
    - Make high confident decisions
    - Answer important biological questions

Boutros Lab "Big Data"

- ~250TB (February, 2020) -> ~2,000TB (May, 2021)
- Various high-throughput molecular datasets including
    - Whole genome sequencing (WGS)
    - Whole exome sequencing (WXS)
    - RNA-seq (whole transcriptome sequencing)
    - Epigenomic and proteomics data
- Cancer data ~300GB per patient
- COVID-19 data ~100GB per patient

UCLA

How do we obtain “Big Data”?

- Generate sequencing data
  - ~2,000GB of sequencing data in 2 days!

llumina NovaSeq 6000

- Download open/controlled access datasets to the cluster storage

NIH NATIONAL CANCER INSTITUTE Genomic Data Commons

EUROPEAN GENOME-PHENOME ARCHIVE

ICGC Data Portal

How do we process “Big Data”?

- Nextflow pipelines
  - Scalable and reproducible parallel workflows on clouds and clusters
  - Boutros lab DNA Nextflow pipelines
  - Boutros lab RNA Nextflow pipelines

- The pipelines can process not only cancer datasets but also any DNA-seq/RNA-seq datasets (e.g. COVID-19 patient WGS)

COVID-19 Genetic Predisposition Study

Investigate genetic role in COVID-19 susceptibility and severity

- Program with five UC medical centers
  - UCLA, UCSF, UCSD, UCD, and UCI
- Analyzing WGS data of 702 patients with diverse backgrounds
  - To identify genetic risk and protection factors linked with symptoms
  - To predict risk for infection and develop new treatment
- Integrating the data into two global consortia
  - COVID Human Genetic Effort
  - COVID-19 Host Genetics Initiative

COVID-19 WGS samples

- Received 702 WGS samples (~50TB) from the five UC centers

UC	# FASTQ	# Patient
UCLA	4272	534
UCSF	608	76
UCSD	336	42
UCD	208	26
UCI	192	24

Cluster scaling

- We scaled up the cluster size on an on-demand basis
  - Reserved 300 worker nodes (= 21,600 CPUs)
  - Adjusted HPC Cache size and throughput
- Why scaled up the cluster?
  - Urgency
    - To fight against the pandemic and process all samples as soon as possible
  - Cost efficiency
    - To reduce the total cost of the whole HPC environment by saving running cost

UCLA

Cluster scaling result

- We successfully
  - deployed 300 worker nodes = 21,600 CPUs
  - processed the 702 WGS samples using the DNA pipelines in ~10 days
  - generated germline data for the 702 samples
    - Single Nucleotide Polymorphisms (SNP)
    - Structural Variations (SV) and Copy Number Variations (CNV)
    - Mitochondria SNPs
  - Identified features to improve our cluster infrastructure and implemented updates

UCLA

Remarkable Progress: Lots More Fun to Come!

Software Environment

- ✓ Team coordination software
  - ✓ Wiki, issue-tracking
- ✓ Software management
  - ✓ Import >1-million lines of code & history!
- ❖ Standardized software across environments
- ❖ Predictable upgrade tempo (quarterly)
- ✓ Framework for job-tracking, delegation, prioritization
- ✓ Open-source pipelines for standardized analyses
- ✓ Routine benchmarking & updating of methods
- ❖ Etc.

Compute Environment

- ✓ HIPAA-compliant
- ✓ Scalable to arbitrary storage
- ✓ Support arbitrary compute (GPU, CPU, FPGA, etc.)
- ❖ Logging & traceability
- ❖ Lock-down of remote access to specific nodes
- ❖ Real-time monitoring
- ✓ Burstable to arbitrary size
- ✓ Productized: easy to reproduce for other uses
- ✓ Cost-conscious
- ❖ Etc.