Minerva Town Hall

Dec 2022

Patricia Kovatch, Dean for Scientific Computing and Data
Lili Gai, PhD, Director for High Performance Computing
Eugene Fluder, PhD, Senior Computational Scientist
Hyung Min Cho, PhD, Senior Computational Scientist
Jielin Yu, PhD, Computational Scientist
Wei Guo, PhD, High Performance Computing Architect (Leave of absence)
Kali McLennan, High Performance Computing Administrator
Jim Turner, High Performance Computing Architect (Contractor)
Catherine Mccaffrey, Project Manager
Ranjini Kottaiyan, MBA, OD, Senior Director,
Finance and Research Engagement

Mount Sinai

Dec. 7, 2022

Outline

2022 Minerva Usage

2022 Accomplishments (May 2022 - Current)

- Staffing
- TSM Archival Storage LTO-9 Tape Solution in Production
- Upgraded LSF & Integrated NVIDIA DCGM for GPU Job Statistics & GPU express queue
- Web Server for Hosting User Websites
- Minimized Minerva PM (preventive maintenance)
- Support Cloud Access Under NIH STRIDES
- Data Ark - Mount Sinai Data Commons
- Documents and Training sessions

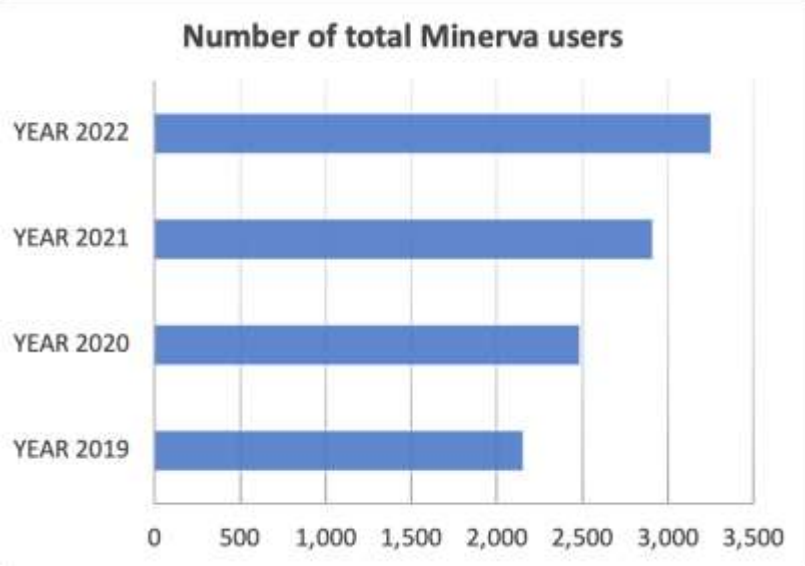
2023 Initiatives and Roadmap

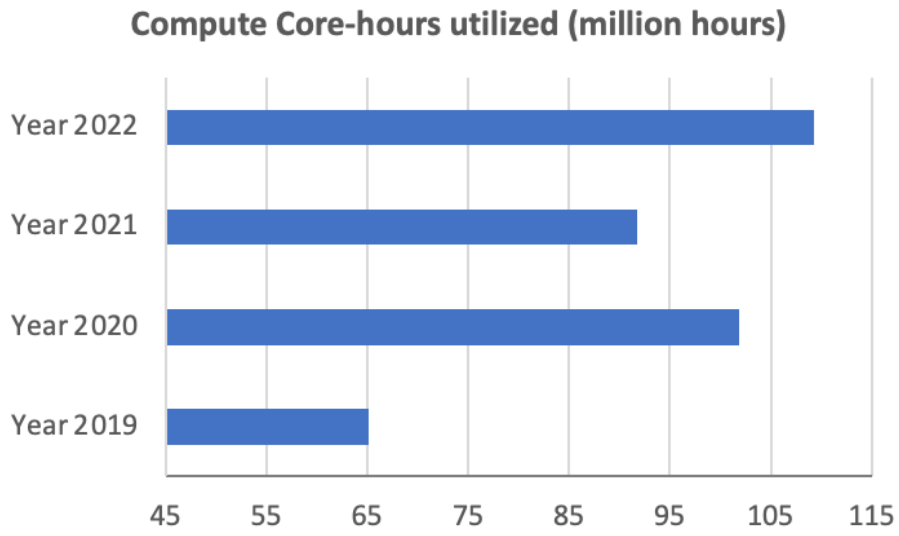
- Annual HIPAA Form Collection in Q1
- Annual HPC User Survey in Q1
- Open OnDemand
- Migrate database to new server
- Data Ark Expansion
- Upgrade GPFs of Arion to 5.1.x for better feature and support
- Mount Sinai Data Warehouse 2 OMOP on Minerva
- Image Research Warehouse on Minerva
- Chimera Refreshment
- More GPU resources

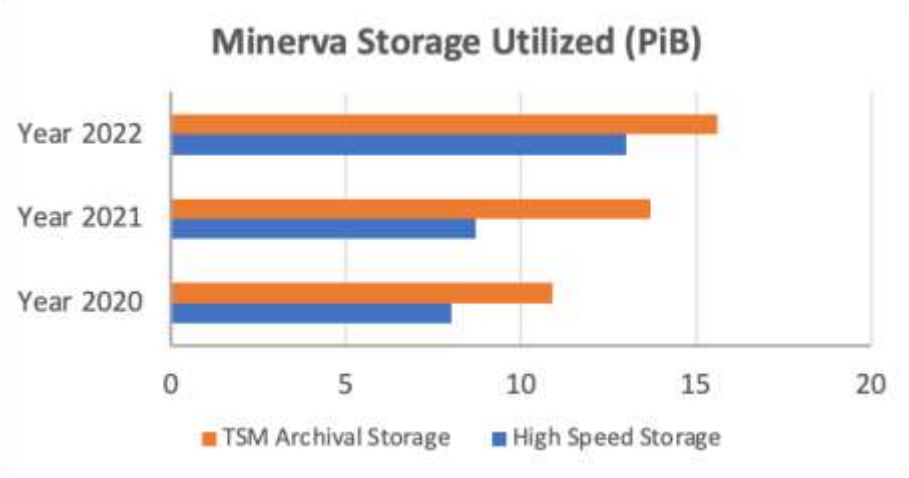
Q&A

Minerva Usage
(Jan 2022 - Oct 2022)

Minerva usage over years







Compute usage higher in year 2020

- COVID study surges in 2020
- Global job limits set on LSF in Mar 2021

Storage usage increased 3 PB in the last 6 months (1.5 PB for digital pathology)

Minerva usage summary (Jan 2022- Oct 2022)

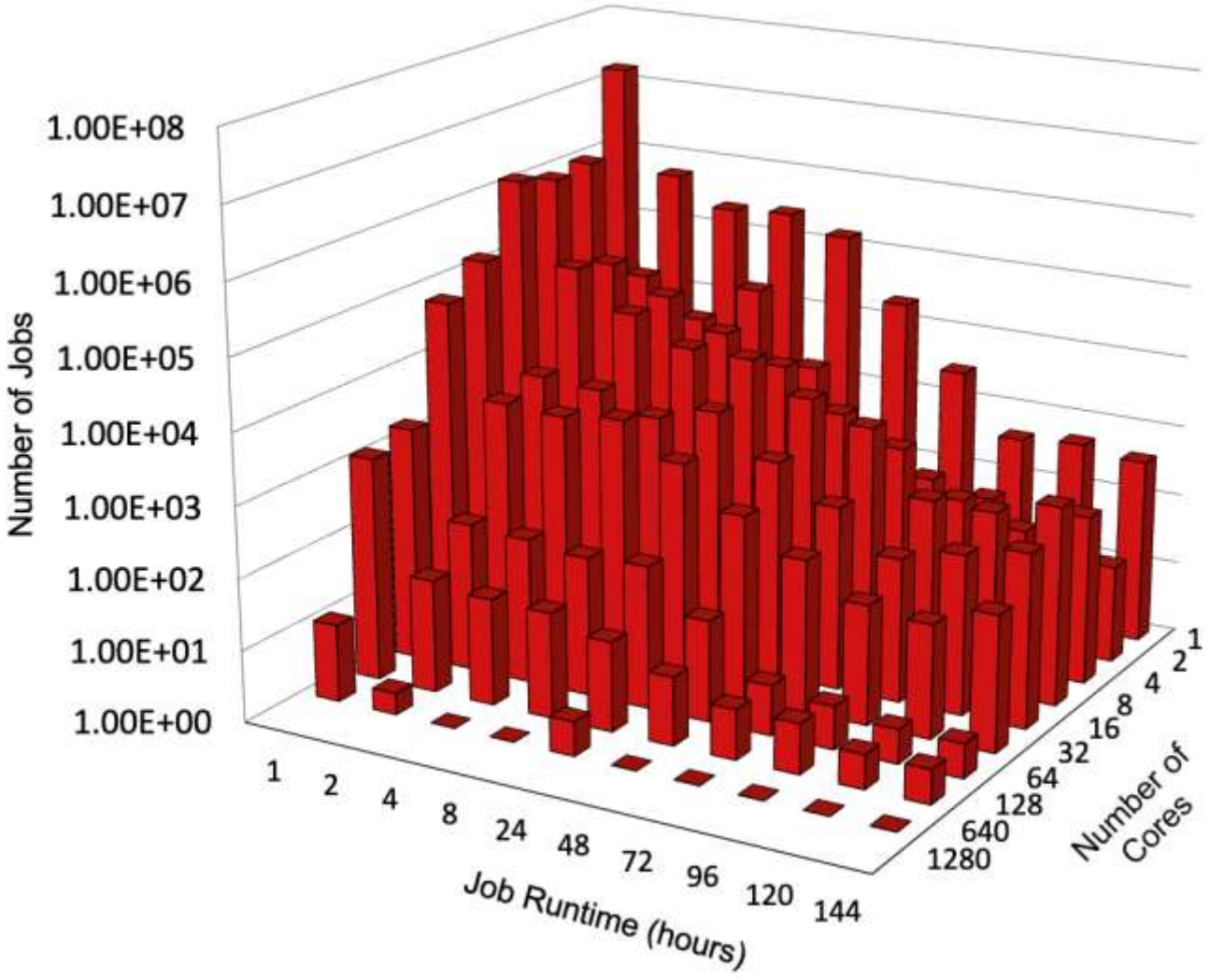
Accounts	
Number of new users	364
Number of active users	811
Number of total users	3,478
Number of project groups	428 (351 active)
Storage	
High-speed storage used (Arion)	13 PiB (41% utilization) 2,710,912,808 files
Archival storage used	15.6 PiB
Compute	
Number of jobs run	21,515,026
Core-hours utilized	79,661,539 hrs

System	
Number of maintenance sessions	1 preventative maintenance(99.9% uptime)

Jobs and compute core hours by partition

Compute	# Jobs	CPU-hours	Utilization
Chimera	8,486,158	44,498,266	53.4 %
BODE2	1,994,026	9,071,735	38.2 %
Hi-memory nodes	6,060,850	9,460,595	88.9 %
CATS	4,896,605	13,597,263	61.0 %
GPU nodes	77,387	3,033,681	54.1 %
Total:	**21,515,026**	**79,661,539**	**54.7 %**

Job Mix


1.00E+08
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
Number of Jobs
1
2
4
8
24
48
72
96
120
144
Job Runtime (hours)
1
2
4
8
16
32
64
128
640
1280
Number of Cores

Top 10 PIs - compute core hours

PI	Department	# Core-hours	# Jobs
Huckins, Laura	Genetics and Genomic Sciences	11,222,371	4,829,083
Chowell, Diego	Precision Immunology Institute	6,886,134	25,507
Zhang, Bin	Genetics and Genomic Sciences	5,021,146	120,442
Roussos, Panos	Psychiatry	4,545,762	905,390
Raj, Towfique	Neurosciences	3,778,817	720,663
Sharp, Andrew	Genetics and Genomic Sciences	3,574,771	1,937,778
Zhang, Zhongyang	Genetics and Genomic Sciences	2,645,500	288,053
Mezei, Mihaly	Structural and Chemical Biology	2,383,027	342
Bunyavanich, Supinda	Genetics and Genomic Sciences	2,332,390	159,008
Samstein, Robert	Radiation Oncology	1,883,510	240,225

Top 10 PIs - GPFS high speed storage

User	Department	Storage usage
Thomas Fuchs	AI and Human Health	1,550 terabytes
Bin Zhang	Genetics and Genomic Sciences	1,247 terabytes
Girish Nadkarni | Alexander Charney	Genetics and Genomic Sciences	1,192 terabytes
Panagiotis Roussos	Psychiatry	923 terabytes
Robert Sebra	Genetics and Genomic Sciences	851 terabytes
Joseph Buxbaum	Psychiatry	436 terabytes
Towfique Raj	Neurosciences	432 terabytes
Alison Goate	Genetics and Genomic Sciences	390 terabytes
Stuart Sealfon	Neurology	381 terabytes
Samir Parekh	Oncological Sciences	326 terabytes

Top compute and storage usage department/institute

Department/Institute	Compute Core Hours
Genetics and Genomic Sciences	37,490,634
Psychiatry	9,047,739
Precision Immunology Institute	6,886,134
Neurosciences	4,225,103
Structural and Chemical Biology	3,234,707
Oncological Sciences	2,917,445
Radiation Oncology	1,883,510
Medicine	1,601,550
HPIMS	526,094
Mindich Child Health and Development Institute	405,308

Department/Institute	Storage (Terabytes)
Genetics and Genomic Sciences	5,363
Psychiatry	1,645
AI and Human Health	1,550
Neurosciences	987
Oncological Sciences	790
Neurology	418
Institute for Genomic Health	210
Structural and Chemical Biology	205
Precision Immunology Institute	161
Medicine	141

Top compute and storage usage department/institute

Department/Institute	Compute Core Hours
Genetics and Genomic Sciences	37,490,634
Psychiatry	9,047,739
Precision Immunology Institute	6,886,134
Neurosciences	4,225,103
Structural and Chemical Biology	3,234,707
Oncological Sciences	2,917,445
Radiation Oncology	1,883,510
Medicine	1,601,550
HPIMS	526,094
Mindich Child Health and Development Institute	405,308

Department/Institute	Storage (Terabytes)
Genetics and Genomic Sciences	5,363
Psychiatry	1,645
AI and Human Health	1,550
Neurosciences	987
Oncological Sciences	790
Neurology	418
Institute for Genomic Health	210
Structural and Chemical Biology	205
Precision Immunology Institute	161
Medicine	141

Total TSM Archival Storage Usage (Jan 2022 - Oct 2022)

Current archive storage usage	
Archived data	15.6 PiB (13.7 PiB LTO5, 1.9 PiB LTO9)
Total data with offsite copy	31.2 PiB (27.4 PiB LTO5, 3.8 PiB LTO9)
Number of tapes used	20,462 (LTO5) 226 (LTO9)

Statistics of Jan 2022 - Oct 2022			
Amount of archived data	2,689 TB	Amount of retrieved data	597 TiB
# of users who have issued archive commands	67	# of users who have issued retrieve operations	48

2022 Minerva Accomplishments

Accomplishments Summary
May 2022 - Current

Actions we took (in response to the user survey and our last roadmap):

- ✓ Surpassed over **1,400** publications that utilized Minerva!!
- ✓ Hired 1 computational scientist and 1 HPC system admin (contractor)
- ✓ Deployed TSM Archival Storage LTO-9 Tape Solution
- ✓ Upgraded LSF & Integrated NVIDIA DCGM for GPU Job Statistics
- ✓ Set up new GPU express queue
- ✓ Set up new procedures for Hosting User Websites on Minerva web server
- ✓ Minimized cluster-wide PMs (only one in 2022)
- ✓ Supported Cloud Access Under NIH STRIDES
- ✓ Expanded Data Ark - Mount Sinai Data Commons (16 data sets in total currently)
- ✓ Set up new MSDW2/OMOP database on Minerva VM
- ✓ Updated the documentation and presented 5 tutorial sessions
- ✓ Continued to support Minerva users through ticketing system (closed **3,021** tickets) and in-person meetings

Details will be presented in the following slides.

Thank you very much for the feedback!

Staffing

Thanks to our staff for keeping Minerva function during the shortage!!! 

The HPC team consists of three senior/computational scientists

- Eugene Fluder, PhD
- Hyung Min Cho, PhD
- Jielin Yu, PhD joined in June 2022

...and five HPC architects/admins

- Kali McLennan joined in April 2022
- Wei Guo, PhD - on LOA (leave of absence for 6 months)
- Jim Turner - part-time contractor started Oct. 2022
- **Three open positions:**
    - **Lead HPC Architect**
    - **HPC Architect**
    - **HPC Admin**

… and one Bioinformatician for Data Ark

- One open position
    - Wen left for SUNY Downstate medical school on Aug 30th.
    - Jielin Yu, PhD is helping during the transition.

TSM Archival Storage LTO-9 Tape Solution

TSM LTO-9 Tape Solution entered production in May

Deploy LTO-9 Tape Solution with increased capacity, performance, and reliability for TSM Archive

- LTO-9 tape capacity is 18TB per tape compared to 1.5 TB on current LTO-5 tapes
- LTO-9 tape transfer rate is 400 MB/sec compared to 140 MB/sec on current LTO-5 tapes.
- LTO-9 tape library is expandable such that **ALL** onsite tapes will be physically in the library all times. This greatly simplifies Lab Operator interactions, is less prone to mistakes, and improves TSM operation completion time.

GPU Changes

NVIDIA DCGM is integrated within LSF on Aug. 15

- The LSF queuing system was upgraded and configured to gather GPU resource usage using NVIDIA Data Center GPU Manager (DCGM).
- User can view GPU usage information for complete LSF jobs with simple commands here

GPU express queue was set up on Nov. 7 with max walltime 15 hours

- To reduce the pending time
- Please use "`#BSUB -q gpuexpress`" for jobs < 15 hours

Will set GPU number limits per user during periods of high usage (Jan 16 2023)

- Heavy users block the GPU queue, resulting in longer pending time for other researchers
- Users will need to **CHANGE** their script for requesting GPU resources (new syntax is simplified)
  - LSF new syntax for GPU resources request will be enabled

Web hosting on shared server web01 - Security

Web01 public access was isolated on Aug 28th due to security concern, and was resumed on Sep. 28.

Actions we took/are taking to enhance security

- Separate the web hosting into two different Domain Name System
- Remediated the global vulnerabilities on web01
- Every web application is scanned by InfoSec before public releasing - DTP
- Establishing web application firewall (WAF) on web01 - DTP
- Will further insulate web01 server by unmounting central file system for public websites, and will provide a different storage solution for file sharing 2023 Q1

New procedures for public website

- By default, all the websites are generated at userid.u.hpc.mssm.edu with only internal access (campus network or VPN)
- Public website URL userid.dmz.hpc.mssm.edu for public websites **[Action needed]**
  - If you need public websites for your research, please fill out the form at https://redcap.link/g08ytzki. The DTP security team will scan the web application. If no critical/high vulnerabilities reported, we will move the webpage to userid.dmz.hpc.mssm.edu for public access.
  - **Timeline: It may take a week to open your website to public**

One Cluster-Wide PM in 2022

With all the updates we had on the system, our system admins managed to do it with only 1 short cluster-wide PM (2 hours).

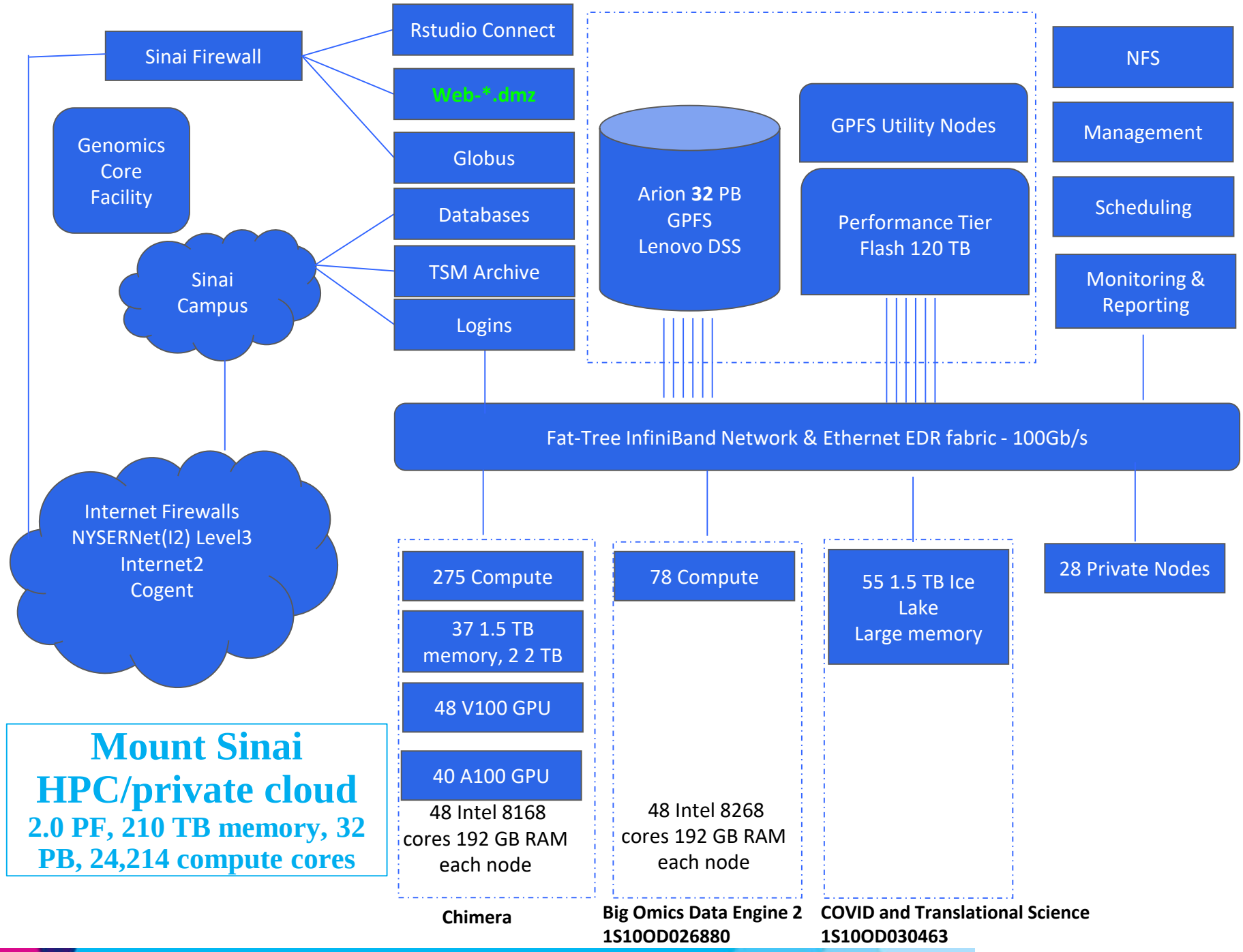
- Well-prepared worksheet by system admins before changes made on system
- Actively monitor the system

Some short windows on specific servers such as TSM upgrade, web server, globus, LSF

An unexpected outage on Minerva login during 5PM Nov. 2nd - 10AM Nov. 3rd

- Caused by an outage on DTP symantec VIP servers.

Thanks to our admins!!!


Sinai Firewall
Rstudio Connect
Web-*.dmz
Globus
Databases
TSM Archive
Logins
Genomics Core Facility
Sinai Campus
Arion 32 PB GPFS Lenovo DSS
GPFS Utility Nodes
Performance Tier Flash 120 TB
NFS
Management
Scheduling
Monitoring & Reporting
Fat-Tree InfiniBand Network & Ethernet EDR fabric - 100Gb/s
Internet Firewalls
NYSERNet(I2) Level3
Internet2
Cogent
275 Compute
37 1.5 TB memory, 2 2 TB
48 V100 GPU
40 A100 GPU
48 Intel 8168 cores 192 GB RAM each node
78 Compute
48 Intel 8268 cores 192 GB RAM each node
55 1.5 TB Ice Lake Large memory
28 Private Nodes
Mount Sinai HPC/private cloud
2.0 PF, 210 TB memory, 32 PB, 24,214 compute cores
Chimera
Big Omics Data Engine 2
1S10OD026880
COVID and Translational Science
1S10OD030463

Cloud Access Under NIH STRIDES

Starting Nov. 2022, we help researchers to gain access to major Cloud Providers such as Google Cloud Platform (GCP), Amazon Web Services (AWS), and Microsoft Azure under the NIH Science and Technology Research Infrastructure for Discovery, Experimentation, and Sustainability (STRIDES) Initiative. NIH's STRIDES Initiative provides cost-effective access/discount to industry-leading partners to help advance biomedical research. More Info

Intake process with HPC Team

- For getting access and benefits inquiry, please open a Minerva ticket at hpchelp@hpc.mssm.edu.

Create accounts and Billing

- Accounts will be managed by DTP.
- User will be billed based on actual usage of the Cloud resources and will be handled by DTP.

Data Ark Data Sets

Launched March 2021!

The Data Ark is located on Minerva and the number, type, and diversity of data sets on the Data Ark will increase substantially in the coming months. The Data Ark consists of public data sets, Mount Sinai generate data sets and School-Acquired data sets. There are also some data supplements provided via Data Ark. Core sets hosted on Minerva include:

Public Data Sets (unrestricted)

- 1,000 Genomes Project
- GTEx
- GWAS Summary Stats
- gnomAD
- The Cancer Genome Atlas (TCGA)
- eQTLGen
- UKBB-LD
- LDSCORE
- BLAST
- Reference Genome

Public Data Sets (restricted)

- UK Biobank

Mount Sinai Generated Data

- STOP COVID NYC Cohort
- MSDW COVID-19 EHR Data Set
- Mount Sinai COVID-19 Biobank
- The Living Brain Project

School-Acquired Data Sets (restricted)

- IBM® MarketScan®

Data Set Supplements – Minerva-hosted, Open Soon Through Data Ark (restricted)

- The Imaging Research Warehouse (IRW) 1.0
- The BioMe BioBank Program
- CIB (Cancer Institute Biorepository)

Data Ark also provides resources on helpful links to external data sets.

Helpful External Data Sets: All of Us

Data Ark Updates

Finalized Data Ark's data set onboarding & retention policy and posted on website

Nine new data sets added this year

- gnomAD
- TCGA (partial)
- eQTLGen (Expression Quantitative Trait Locus)
- UKBB-LD (UK Biobank Linkage Disequilibrium)
- Reference genome
- LDSCORE
- BLAST
- The Living Brain Project (LBP)
- IBM MarketScan - claims data on 200 million people

Data Sets Pending

- BioMe
- Human Cell Atlas

Added Data Ark to weekly Digital Concierge Service (every Wednesday from 3:30-4:30 PM)

Documentation and training sessions

- Documentation updated on the website in Sep 2022!!!
  - ➔ Logging in
  - ➔ Storage and file permission
  - ➔ Software Environment: Lmod
  - ➔ Job scheduling: LSF
  - ➔ GPGPU and GPU Etiquette
  - ➔ Restart your LSF Jobs:
  - ➔ Running Container: Singularity
  - ➔ R and Rstudio
  - ➔ Python and Jupyter Notebook
  - ➔ MATLAB
  - ➔ Schrodinger Suite
  - ➔ Long-Term Data Archive: TSM
  - ➔ File transfer: Globus and others
  - ➔ Web application: Web server
  - ➔ Database: MariaDB
  - ➔ Software build and compile
  - Our website at https://labs.icahn.mssm.edu/minervalab/
  - We provided additional training material (including slides & recording) online
- Offered training sessions in person/Zoom:
  - Five training sessions in spring and fall
  - Topics include “Introduction to Minerva”, “LSF job scheduler” & “Singularity Container”
  - ~40 participants per session
- For most recent announcement and updates:
  - Join our mail-list: hpcusers@mssm.edu
  - Minerva user group meetings will be scheduled as needed
  - Message Of The Day on Minerva

2023 Initiatives and Roadmap

What's next?

Annual HIPAA Form Collection in Q1

Annual HPC User Survey in Q1

Visualization Portal

- Set up Open OnDemand as better visualization portal to access Minerva through web browsers
- Slowed due to short of system admins

Migrate the old database to new server for better performance and stability

Upgrade GPFs of Arion to 5.1.x for better feature and support

Data Ark Expansion in Q1

- Digitized pathology slides will be available on Minerva, integrated with MSDW
- BioMe WES Data will be hosted under Data Ark
- Data from BedMaster servers

Mount Sinai Data Warehouse 2 OMOP on Minerva

- De-identified and identified clinical data warehouse on Minerva
- High-Availability (HA) infrastructure for MSDW OMOP

Imaging Research Warehouse 2.0 soon to be available on Minerva

- Develop cohorts with the Leaf self-service cohort query tool for 1.0 (5 years of images now)
- Access multi-modal Radiology images served via XNAT metadata server in Q2 2023

What's next?

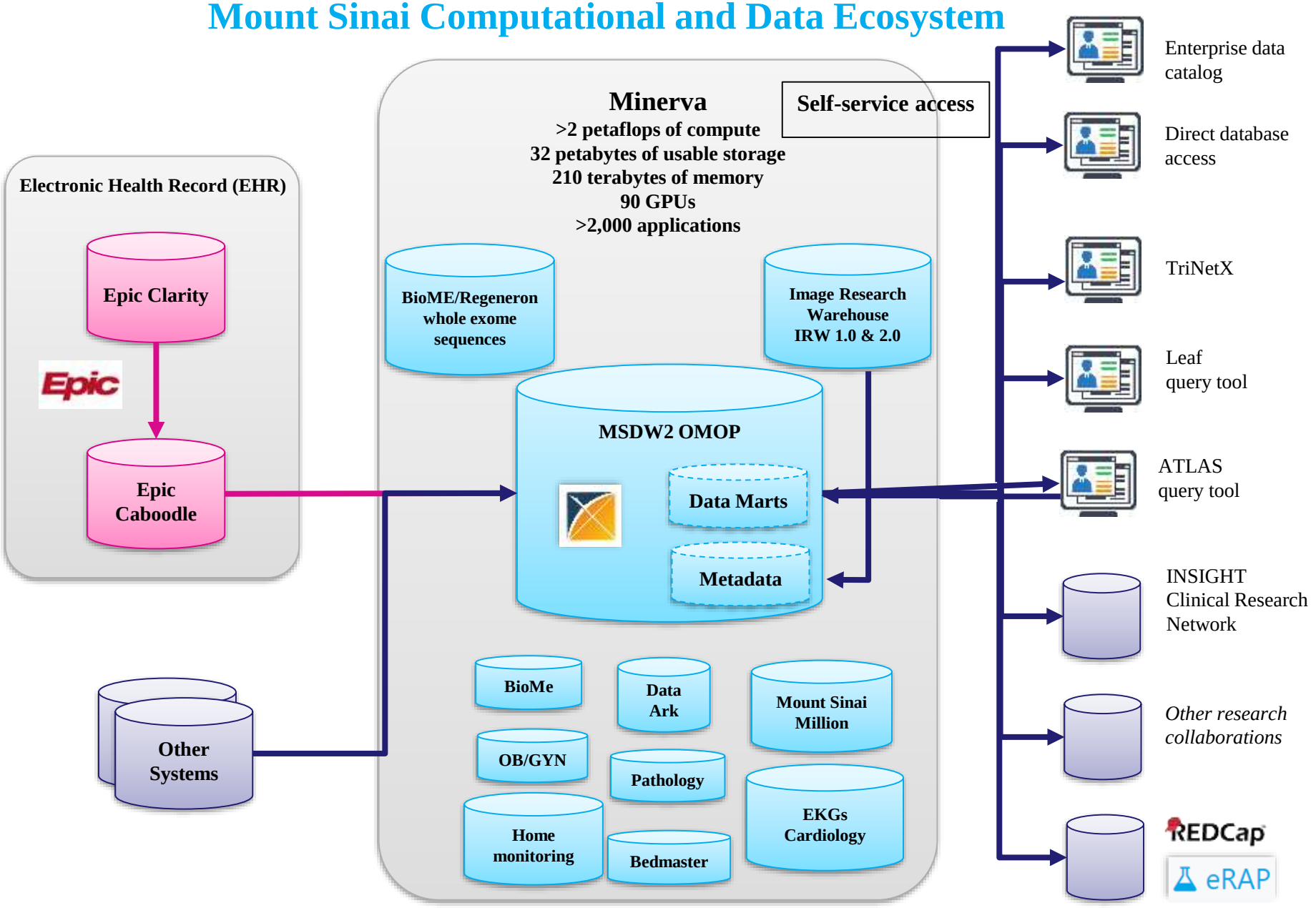
Chimera Refreshment

- Maintenance on Chimera (14,000 intel Skylake cores and 48 V100 GPUs) is expiring Dec 2023
- We need to start the RFP process in Jan 2023 to replace this functionality if budget approved

More GPU resources?

- We plan to submit a NIH S10 gpu proposal (Due June 1 2023)
- WE NEED YOUR HELP!
    - We need > 10 PIs with NIH funding

Mount Sinai Computational and Data Ecosystem



Thank you!