Minerva Town Hall
April 2022

Patricia Kovatch, Dean for Scientific Computing and Data
Lili Gai, PhD, Director for Scientific Computing and Data
Eugene Fluder, PhD, Senior Computational Scientist
Hyung Min Cho, PhD, Senior Computational Scientist
Wei Guo, PhD, High Performance Computing Architect
Wayne Westerhold, High Performance Computing Administrator
Kali McLennan, High Performance Computing Administrator
Wen Huang, Bioinformatician
Catherine Mccaffrey, Project Manager
Ranjini Kottaiyan, Finance Director

April 25, 2022

Mount
Sinai

Outline

Welcome and general comments

- 2021 - 2022 Minerva Usage
- 2022 User survey results
- 2022 Accomplishments
    - Staffing
    - $2M COVID and Translational Science (CATS) S10 Machine in Production
    - Large memory GPU nodes in production
    - Deploying TSM Archival Storage LTO-9 Tape Solution
    - PM
    - Data Ark - Mount Sinai Data Commons
    - Documents and Training sessions

2022 Initiatives and Roadmap

- Documents and Training sessions
- Migrate database to new server
- GPU Job Statistics with NVIDIA Data Center GPU Manager and LSF in Q2
- Mount Sinai Data Warehouse 2 OMOP on Minerva
- Image Research Warehouse on Minerva
- Digital Pathology Images
- User Channels: HPC - forum on Slack

Q&A

Minerva Usage
(June 2021 - April 2022)

Minerva usage summary (June 2021- April 2022)

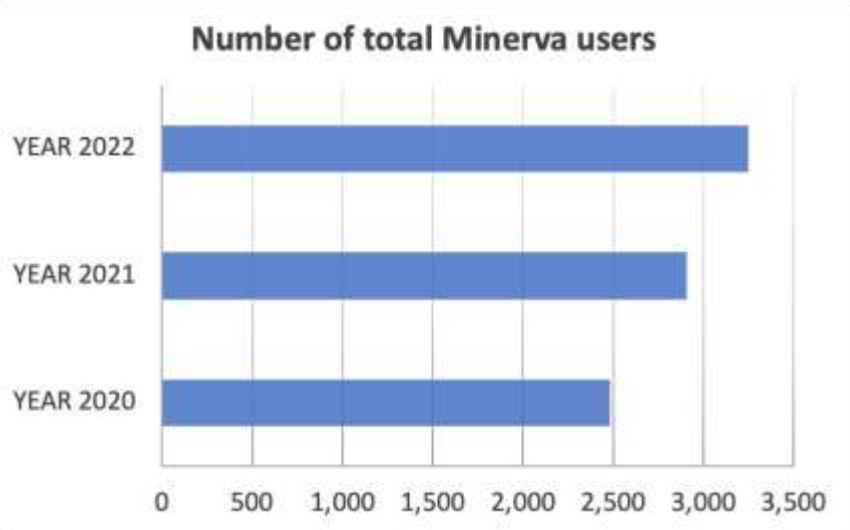
Accounts	
Number of new users	347
Number of active users	701
Number of total users	3,252
Number of project groups	408(318 active)

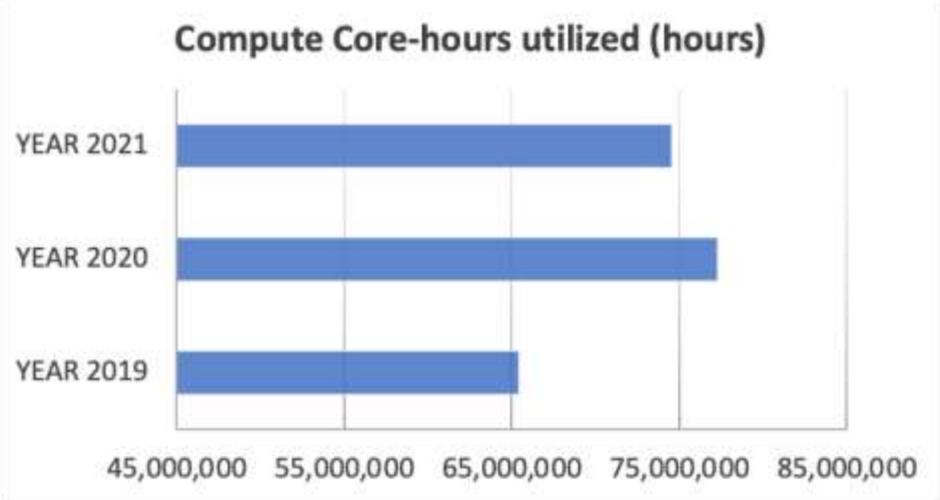
Storage	
High-speed storage used (Arion)	9.9 PiB (32% utilization) 5,460,412,794 files
Archival storage used	14.3 PiB

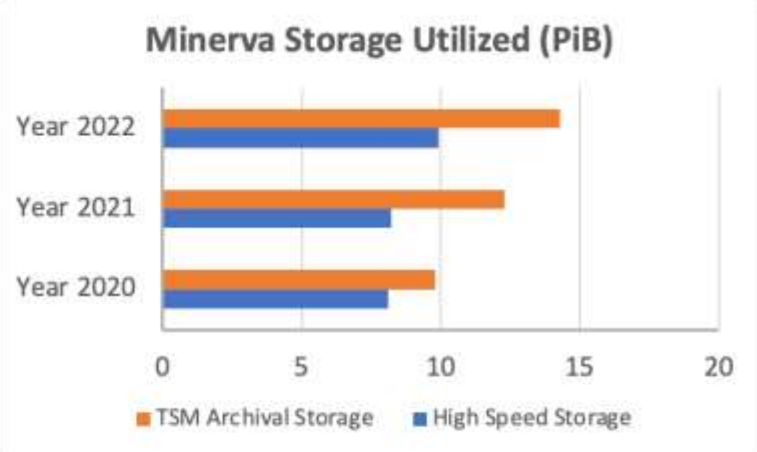
Compute	
Number of jobs run	25,938,066
Core-hours utilized	71,574,470 hrs

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

Minerva usage over years







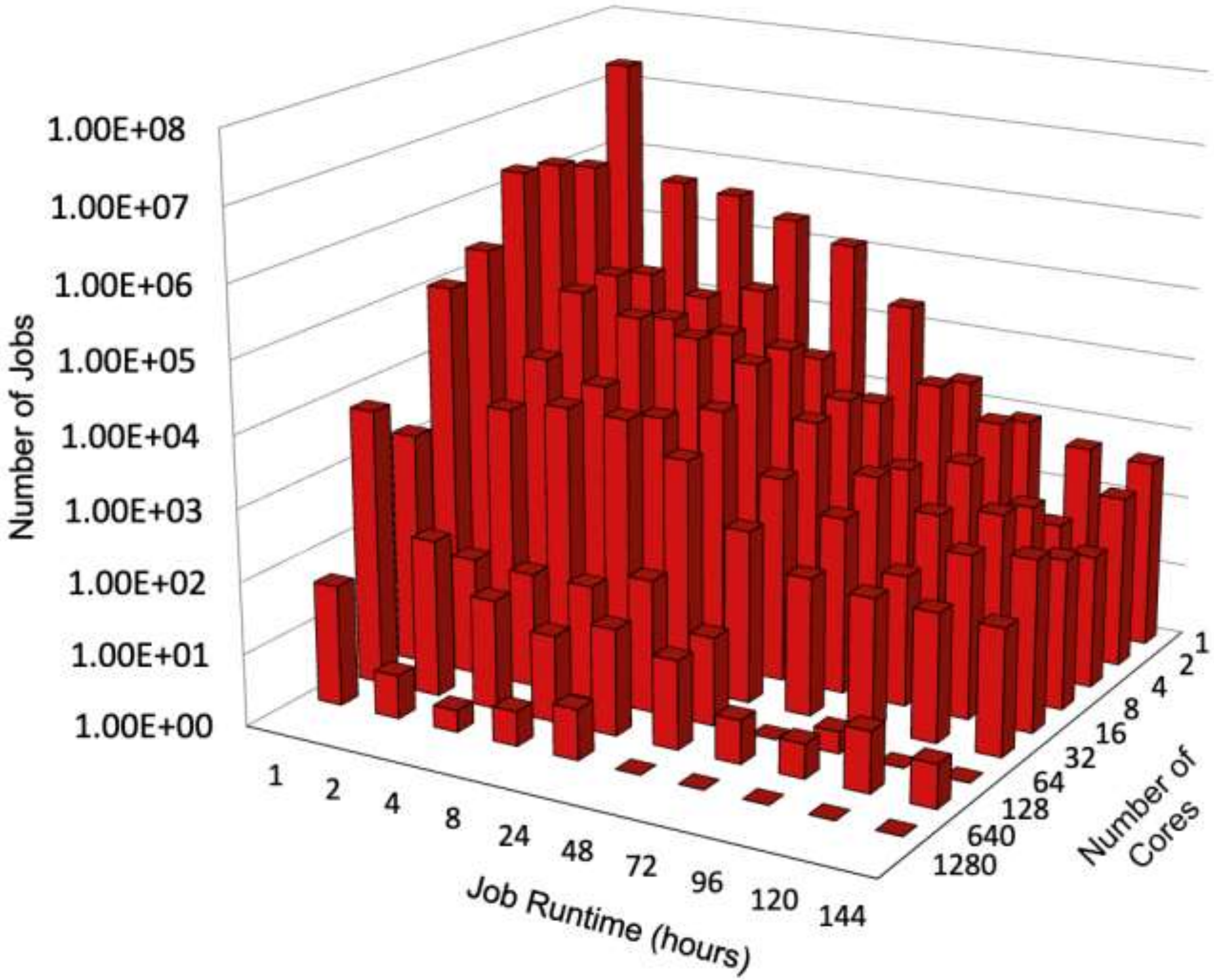
Compute usage higher in year 2020

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

Jobs and compute core hours by partition

Compute	# Jobs	CPU-hours	Utilization
Chimera	15,592,708	42,674,794	44.6 %
BODE2	3,650,129	10,907,959	40.1 %
Hi-memory nodes	5,300,124	11,631,071	95.2 %
CATS	1,315,626	3,716,690	67.7%
GPU nodes	79,479	2,643,956	47.3 %
Total:	**25,938,066**	**71,574,471**	**49.0 %**

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 users compute core hours

PI	Department	# Core-hours	# Jobs
Huckins, Laura	Genetics and Genomic Sciences	7,827,185	7,344,716
Chowell, Diego	Precision Immunology Institute	4,022,801	30,923
Reva, Boris	Genetics and Genomic Sciences	3,351,279	70,137
Raj, Towfique	Genetics and Genomic Sciences	3,238,573	756,766
Sharp, Andrew	Genetics and Genomic Sciences	3,217,987	746,586
Buxbaum, Joseph	Genetics and Genomic Sciences	3,139,192	410,812
Zhang, Bin	Genetics and Genomic Sciences	3,079,200	115,843
Cho, Judy	Genetics and Genomic Sciences	2,784,418	159,745
Roussos, Panos	Psychiatry	2,482,515	1,816,003
Charney, Alexander	Genetics and Genomic Sciences	2,466,961	5,027,639

Top 10 PIs GPFS high speed storage

User	Department	Storage usage
Bin Zhang	Genetics and Genomic Sciences	928 terabytes
Robert Sebra	Genetics and Genomic Sciences	896 terabytes
Panagiotis Roussos	Genetics and Genomic Sciences	862 terabytes
Laura Huckins	Genetics and Genomic Sciences	578 terabytes
Alexander Charney	Genetics and Genomic Sciences	384 terabytes
Judy Cho	Genetics and Genomic Sciences	367 terabytes
Towfique Raj	Neurosciences	353 terabytes
Alison Goate	Neurosciences	321 terabytes
Stuart Sealfon	Neurology	293 terabytes
Diego Chowell	Precision Immunology Institute	255 terabytes

Top compute and storage usage department/institute

Department/Institute	Compute Core Hours
Genetics and Genomic Sciences	37,794,805
Psychiatry	7,555,143
Precision Immunology Institute	4,022,801
Neurosciences	3,991,429
Oncological Sciences	3,656,635
Medicine	2,929,874
Structural and Chemical Biology	1,210,146
HPIMS	612,467
Institute for Genomic Health	576,296
Mindich Child Health and Development Institute	417,036

Department/Institute	Storage (Terabytes)
Genetics and Genomic Sciences	5,300
Psychiatry	1,100
Neurosciences	786
Oncological Sciences	589
Neurology	312
Precision Immunology Institute	228
Structural and Chemical Biology	182
Medicine	122
Surgery	114
BioMedical Engineering and Imaging Institute	99

Total TSM Archival Storage Usage (June 2021- April 2022)

Current archive storage usage	
Archived data	14.3 PiB
Total data with offsite copy	28.6 PiB
Number of tapes used	21,024

Statistics of June 2021 - April 2022			
Amount of archived data	2,407 TB	Amount of retrieved data	486 TiB
# of users who have issued archive commands	79	# of users who have issued retrieve operations	53

2022 User Survey Results

Survey results and discussion

We asked five questions:

Q1: Overall, how satisfied are you with the LSF queue structure, compute and storage resources (GPUs, high-memory nodes, TSM, etc)?

Q2: Please rate current software environment (packages and services such as database, data transfer, container etc).

Q3: Please rate your satisfaction with operations (ticket system, responsiveness of staff, documentation, user support etc).

Q4: Which of the following would you most prefer for future Minerva expansion?

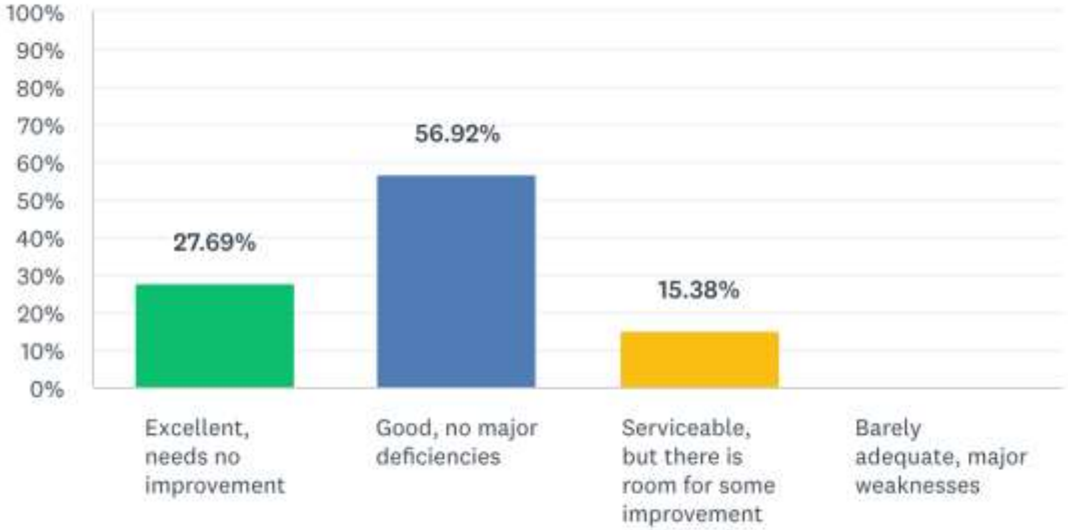
Q5: What suggestions do you have for improving our service? (17 comments)

We received 65 responses and 64 comments.

Thank you for your feedback! This is the motivation for our 2022 Minerva Roadmap!

2022 Survey results question 1

Q1: Overall, how satisfied are you with the LSF queue structure, compute and storage resources (GPUs, high-memory nodes, TSM, etc)?



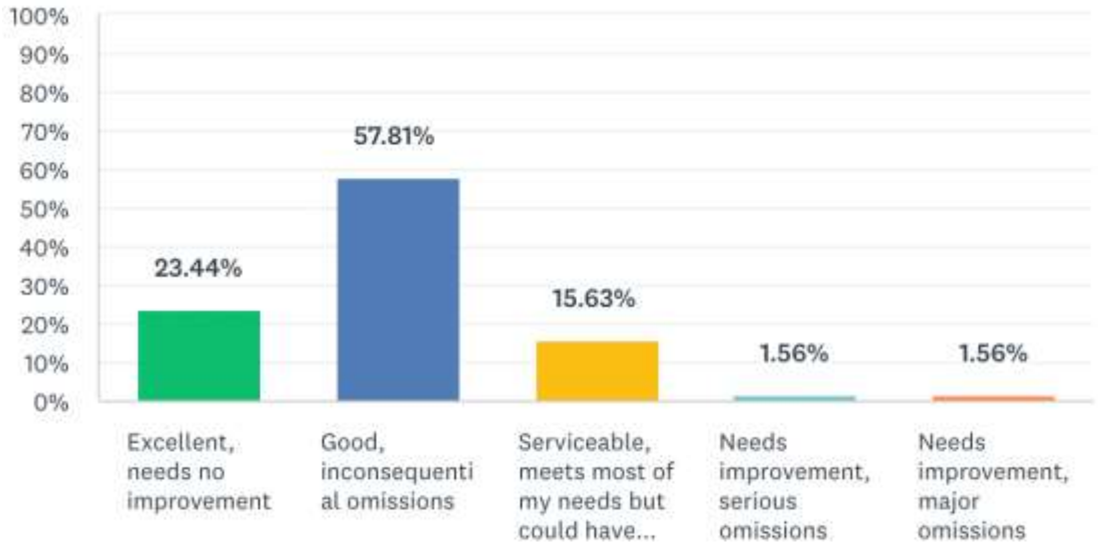
User satisfaction(>=Good)
2021: 85%
2020: 81%
2019: 65%
2018: 54%

<u>**Comments(18):**</u>

- **Login nodes**: Not responding or slow
- **LSF queuing**: long waiting time; Need another express queue with a smaller limit on the # of jobs
- **Nodes**: more gpu nodes, and high memory nodes
- **TSM**: need backup solution

2022 Survey results question 2

Q2: Please rate the current software environment (packages and services such as database, data transfer, container etc)



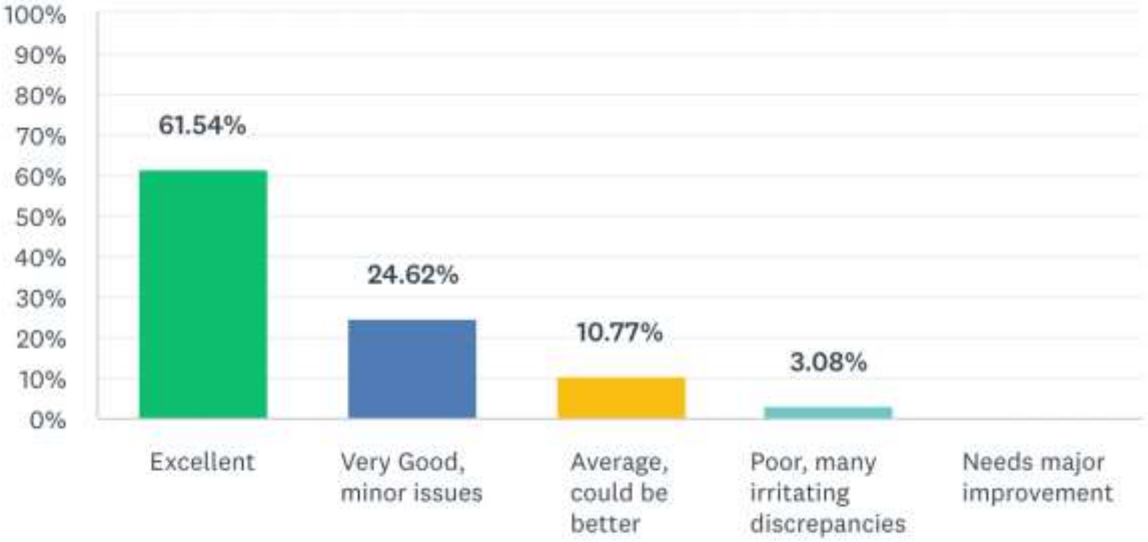
User satisfaction(>=Good)
2021: 81%
2020: 80%
2019: 80%
2018: 67%

Comments(13):

- More up to date versions of python and pip for virtual environment work
- Integration with workflow managers like Snakemake and other apps like RstudioWeb is a bit clunky and a drain on productivity. Still, there are usually workarounds, it is just time intensive
- Certain modules are broken, a reassessment of modules' functionality may be useful.
- The current RStudio-on-the-fly makes it a little difficult to run computations that require a long time, and sometimes multiple attempts are needed to open a session.

2022 Survey results question 3

Q3: Please rate your satisfaction with operations (ticket system, responsiveness of staff, documentation, user support etc)



User satisfaction(>=Good)
2021: 86%
2020: 91%
2019: 73%
2018: 80%

Comments (11):

- **Documentation:** Documentation needs to be better. The PDFs of training powerpoints is useful to an extent, but an actual documentation system is needed.
- **Tickets:** occasional missed tickets but very rare
- **PMs:** Please schedule maintenance on evenings and/or weekends. It is extremely disruptive to do during the working day.
- **Staff:** very responsive and have no issues

2022 Survey results and discussion

Q4: Which of the following would you most prefer for future Minerva expansion?

ANSWER CHOICES	RESPONSES	
More GPU Nodes	29.69%	19
More High Memory Nodes	39.06%	25
Higher Number of Core Counts Per Node	26.56%	17
Faster Cores	40.63%	26
Compute Nodes with Local SSDs	9.38%	6
No need	18.75%	12
Others (please specify):	7.81%	5
Total Respondents: 64		

With the response, we will keep this mind when expanding Minerva in future

2022 Survey results summary

Thank you for your feedback! This is the motivation for our 2022 Minerva Roadmap

Some raw comments:

- I've found the slurm scheduler to be more preferable to LSF, don't know if you guys agree with that but I would prefer slurm to LSF.
- If there could be a common place for labs to share basic helper scripts, documentation on tasks, and/or templates, I would love to contribute, and I'm sure other labs would too.

- Remove the recently implemented two-factor authentication requirement to SSH into Minerva from within the MSSM network: this is already required to join the MSSM network in the first place to obtain a login prompt, so there is NO real added security as a result of this policy, regardless of your interpretation of HIPAA requirements
- Again, a spot for backing up some of the data. Clusters can be dangerous if data is lost and not recoverable.

- The service we get, response time, attending to specific needs - all are exceptional!!!!!
- Staff is *the best*.
- I want to start by saying Minerva is an incredibly helpful and necessary system with so much to offer. I love using it,
- I want to reiterate that Minerva is excellent, and my research would be nowhere close to where it is without it. Your team is so responsive and works so hard, and this system is incredibly

We have all the responses to the comments here.

Updates

Accomplishments Summary

June 2021-April 2022

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

- ✓ Surpassed over **1,300** publications that utilized Minerva!!
- ✓ Hired 1 HPC System Administrator
- ✓ Hired a bioinformatician for Data Ark in conjunction with GGS
- ✓ Expanded Minerva with more high memory nodes and storage with $2M NIH award
- ✓ Minimized cluster-wide PMs (only one in 2022)
- ✓ Announced Mount Sinai Data Commons - Data Ark on Minerva
- ✓ Set up new MSDW2/OMOP database on Minerva VM
- ✓ Updated the documentation and presented 4 tutorial sessions
- ✓ Continued to support Minerva users through ticketing system (closed **2,660** tickets) and in-person meetings

Details will be presented in the following slides.

Thank you very much for the feedback from user survey!

Staffing

The HPC team consists of three senior/computational scientists...

- Eugene Fluder, PhD
- Hyung Min Cho, PhD
- **One computational scientist vacant position** will be joining in June

...and five HPC architects/admins:

- Wayne Westerhold
- Wei Guo, PhD
- Kali McLennan - joined us at April 18 2022
- **Two open positions:**
    - **Lead HPC Architect**
    - **HPC Architect**

… <u>and one Bioinformatician for Data Ark</u>

Wen(Sofia) Huang, joined us in Jan 2022

Minerva Expansion with $2M NIH award - CATS

In production Jan 2022!!!

55 high memory nodes (**1.5 TB** of memory per node)

- 64 Intel Xeon Platinum 8358 2.6 GHz Processors per node, a total of 3,520 cores
- This Intel Ice Lake processor is ~15% - 20% performance improve compared to our current Cascade Lake and Skylake processors.

12 petabytes of usable storage (16 PB raw)

- Already integrated into Arion file system, for a total of 32 PB of usable storage
  - Updated GPFs server version to 5.0.5.7

```
$ df -h /sc/arion
Filesystem      Size  Used Avail Use% Mounted on
arion            32P  9.9P   22P  32% /sc/arion
```

- Apply for storage allocation here

Minerva has a total of 497 nodes for a total of 24,024 cores, and 32 petabytes of usable storage after this addition

New 2 TB memory GPU nodes in production

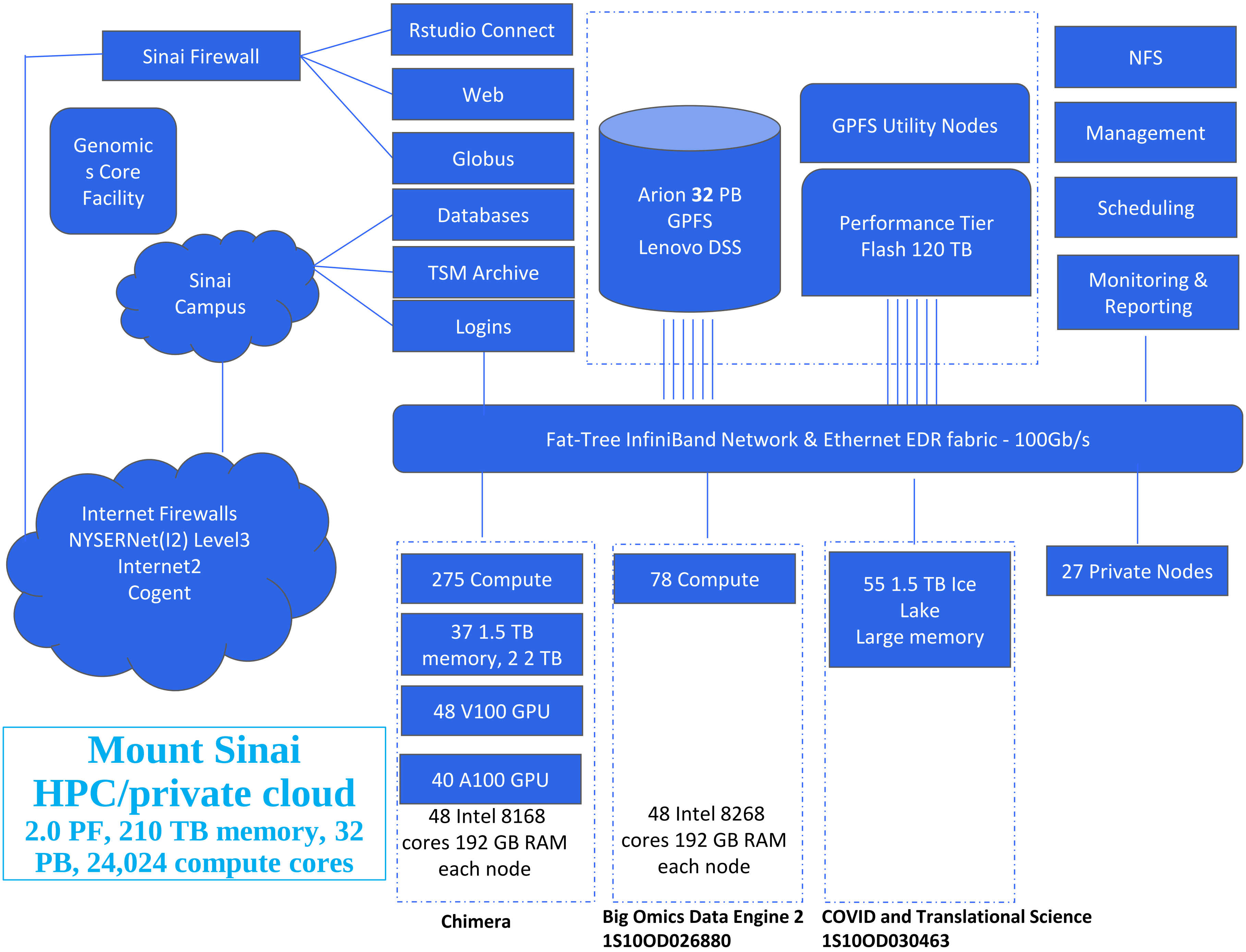
In production April 2022!!!

2 high memory GPU nodes (**2 TB** of memory per node)

- 64 Intel Xeon Platinum 8358 2.6 GHz Processors per node, a total of 128 cores
- 8 * 80GB A100 in total
- NVLinked,
- 7TB of local NVME SSD, which can deliver a sustained read-write speed of 3.5 GB/s in contrast with SATA SSDs that limit at 600 MB/s
- Request by “ **`-R a10080g`**”

Mount Sinai HPC/private cloud

2.0 PF, 210 TB memory, 32 PB, 24,024 compute cores



Sinai Firewall

Rstudio Connect

Web

Globus

Databases

TSM Archive

Logins

Genomics Core Facility

Sinai Campus

Internet Firewalls
NYSERNet(I2) Level3
Internet2
Cogent

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

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

Chimera

78 Compute

48 Intel 8268 cores 192 GB RAM each node

Big Omics Data Engine 2
1S10OD026880

55 1.5 TB Ice Lake
Large memory

COVID and Translational Science
1S10OD030463

27 Private Nodes

Deploying TSM Archival Storage LTO-9 Tape Solution

We are at the final stage upgrading the TSM server

Will be in 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.

Only 1 Cluster-Wide PM in 2022

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

- Some short windows on specific servers, network upgrade, and TSM upgrade
- Well-prepared worksheet by system admin before changes made on system

Data Ark Data Sets

Launched March 2021!

Public data sets(unrestricted): NO DUA form required

- ▶1,000 Genomes–Whole Genome Sequencing (WGS) data on ~1,000 individuals of mixed ancestry
- ▶GWAS Summary Stats –Genome Wide Association Studies (GWAS) results in standardized format across 1,000s of outcomes
- ▶GTEx–Gene expression data on hundreds of individuals across ~50 tissues.
- ▶gnomAD–The Genome Aggregation Database (gnomAD)is a resource that for variation interpretation.

Public data sets(restricted)

- ▶UK Biobank –Genetic data (genotype/WES) from the UK Biobank data on 500,000 individuals

Mount Sinai-generated data

- ▶STOP COVID NYC Cohort
  - –Symptom and behavior on COVID-19 on ~50,000 New York City residents surveyed via phone apps in April 2020
  - –Gene expression data on hundreds of individuals across ~50 tissues
- ▶Mount Sinai Data Warehouse COVID-19 Electronic Health Record Data Set
  - –De-identified clinical data on patients from Caboodle with or suspected of COVID-19 containing 350 data fields
- ▶Mount Sinai COVID-19 Biobank
  - –Blood samples from hundreds of COVID-19 patients hospitalized at Mount Sinai, with genotype/WGS data

Data Ark website: https://labs.icahn.mssm.edu/minervalab/data-ark/

Data Ark Updates

- Issued two surveys to Minerva users for Data Ark and Computational Genomics, and distributed the results on website and to Minerva users. Results are here
- New GGS-funded staff:
  - Bioinformatician Wen Huang, MS, MPH started on January 1, 2022
  - Maria Julia Castro, project manager, started on February 9, 2022
- Developing policies to add/delete data sets in conjunction with the DA Advisory Board
- Exploring the addition of heavily-used portions of TCGA (The Cancer Genome Atlas)
  - RNAseq data counts+MAF(Mutation Annotation Format)+ cBioportal

Documentation updates and training sessions

- 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

- Offered training sessions in person/Zoom:
  - Two sets of training sessions in spring and fall
  - Topics include “Introduction to Minerva” and “LSF job scheduler”
  - Uploading the course on PEAK

- Offered HPC Town Hall in person/Zoom

- Documentation updated on the website:
  - Our website at https://labs.icahn.mssm.edu/minervalab/
  - We will add newer pages/articles as needed
  - We provided additional training material (including slides) online

2022 Initiatives and Roadmap

What’s next?

Updating Documentation on Minerva in Q2

- 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

GPU Job Statistics with NVIDIA Data Center GPU Manager and LSF in Q2

- Upgrade LSF and get NVIDIA Data Center GPU Manager (DCGM) integrated

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

Mount Sinai Data Warehouse 2 OMOP on Minerva

- De-identified and identified clinical data warehouse on Minerva

Imaging Research Warehouse 1.0 soon to be available on Minerva

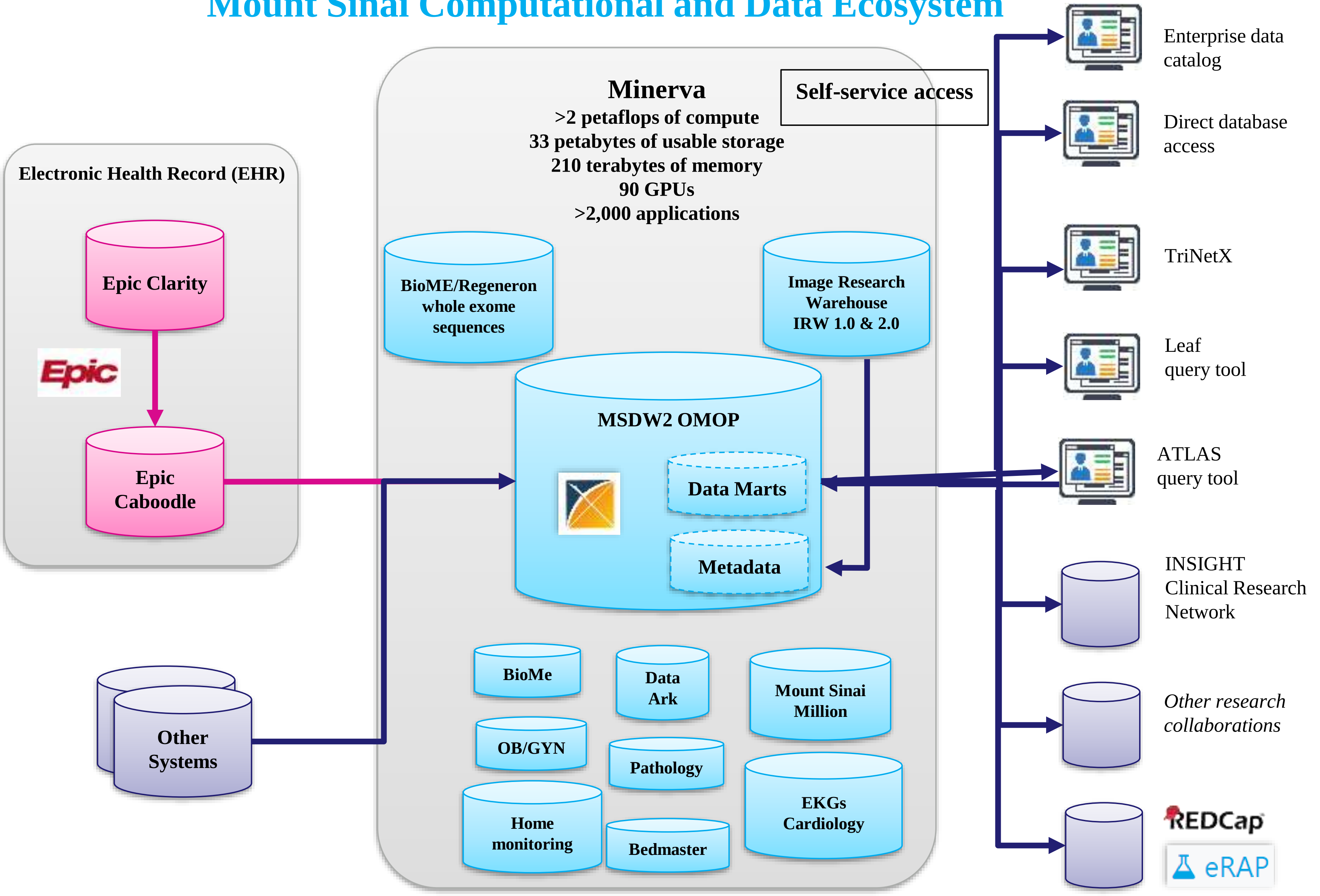
- Develop cohorts with the Leaf self-service cohort query tool
- Access multi-modal Radiology images served via XNAT metadata server

Working to move digitized pathology slides to Minerva

New User Channels: HPC - forum on Slack

- For Minerva Users to communicate, share and collaborate

Mount Sinai Computational and Data Ecosystem



Electronic Health Record (EHR)

Epic Clarity

Epic Caboodle

Minerva
>2 petaflops of compute
33 petabytes of usable storage
210 terabytes of memory
90 GPUs
>2,000 applications

Self-service access

BioME/Regeneron whole exome sequences

Image Research Warehouse IRW 1.0 & 2.0

MSDW2 OMOP

Data Marts

Metadata

BioMe

Data Ark

Mount Sinai Million

OB/GYN

Pathology

Home monitoring

Bedmaster

EKGs Cardiology

Other Systems

Enterprise data catalog

Direct database access

TriNetX

Leaf query tool

ATLAS query tool

INSIGHT Clinical Research Network

Other research collaborations

REDCap

eRAP

Thank you!