Minerva Town Hall

May 2023

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
Kali McLennan, High Performance Computing Administrator
Jim Turner, High Performance Computing Architect (contractor)
Yiyuan Liu, PhD, Bioinformatician
Catherine Mccaffrey, Project Manager
Ranjini Kottaiyan, MBA, OD, Senior Director, Finance

Mount Sinai

May 3rd, 2023

Outline

Minerva Usage (Oct 2022 - March 2023)

2023 User Survey Results

2022- 2023 Accomplishments (Oct 2022 - March 2023)

- Staffing
- Minerva system upgrade including OS, networking stack and parallel file system
- New Annual HIPAA form launched in Jan 2023
- New H100 GPU nodes and LSF changes
- High-Availability (HA) cluster for MSDW OMOP servers
- Minerva PM (preventive maintenance)
- Data Ark - Mount Sinai Data Commons
- Documentation and training sessions

2023 Initiatives and Roadmap

- TSM archival storage LTO-5 tape solution out of support
- Minerva Refreshment
- Open OnDemand
- Migrate database to new server
- Migrate Minerva two factor authentication to Azure MFA
- Install/config new SSD storage server for user $HOME
- Data Ark Expansion

Q&A

Minerva Usage
(October 2022 - March 2023)

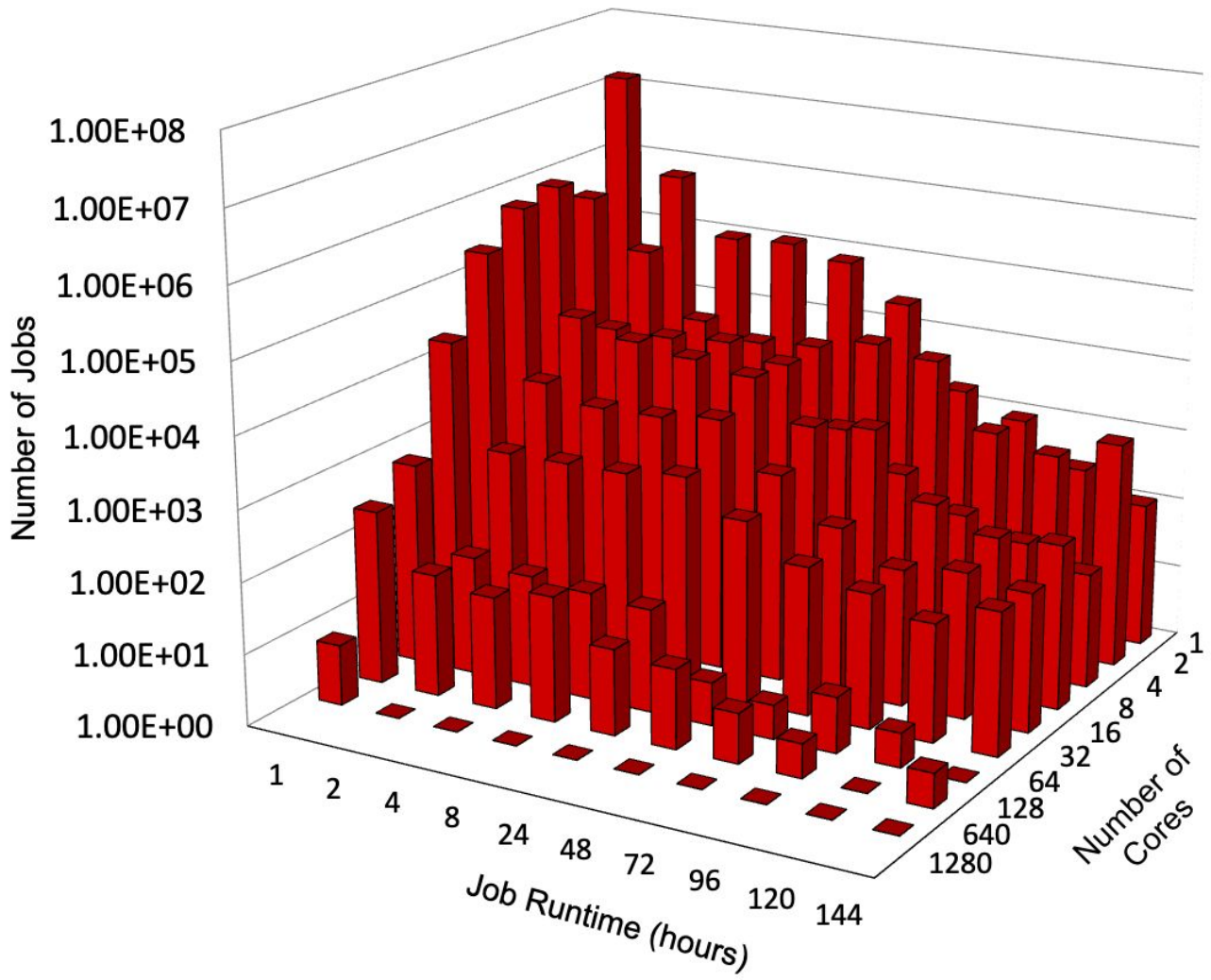
Minerva usage summary (Oct 2022- March 2023)

Accounts	
Number of active users	832
Number of total users	3,673
Number of project groups	467(339 active)
Storage	
High-speed storage used (Arion)	15PB (46% utilization) 7,110,736,556 files
Archival storage used	17.7 PB
Compute	
Number of jobs run	16,923,121
Core-hours utilized	54,733,818 hrs
System	
Number of maintenance sessions	No preventative maintenance (99.6% uptime)

Jobs and compute core hours by partition

Compute	# Jobs	CPU-hours	Utilization
Chimera	5,787,548	21,662,635	37.7 %
BODE2	2,426,618	11,205,641	68.5 %
Hi-memory nodes	3,770,975	5,787,886	79.7 %
CATS	4,813,399	13,350,893	86.9%
GPU nodes	124,581	2,726,763	70.0 %
Total:	**16,923,121**	**54,733,818**	**54.5 %**

Job Mix


Number of Jobs
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
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
Charney, Alexander	Genetics and Genomic Sciences	8,844,682	2,288,302
Bunyavanich, Supinda	Genetics and Genomic Sciences	8,111,176	125,972
Sharp, Andrew	Genetics and Genomic Sciences	4,809,884	2,006,906
Roussos, Panos	Psychiatry	3,304,671	856,675
Zhang, Bin	Genetics and Genomic Sciences	2,788,860	95,829
Goate, Alison	Genetics and Genomic Sciences	2,661,290	189,998
Buxbaum, Joseph	Genetics and Genomic Sciences	2,431,900	693,169
Klein, Robert	Genetics and Genomic Sciences	1,717,080	6,963
Raj, Towfique	Neurosciences	1,511,968	796,602
Schadt, Eric	Genetics and Genomic Sciences	1,503,468	25,051

Top 10 PIs GPFS high speed storage

PI	Department	Storage usage
Thomas Fuchs	AI and Human Health	1.7 petabytes
Bin Zhang	Genetics and Genomic Sciences	1.3 petabytes
Alexander Charney	Genetics and Genomic Sciences	1.0 petabytes
Panagiotis Roussos	Genetics and Genomic Sciences	991 terabytes
Robert Sebra	Genetics and Genomic Sciences	962 terabytes
Towfique Raj	Neurosciences	669 terabytes
Stuart Sealfon	Neurology	554 terabytes
Joseph Buxbaum	Psychiatry	486 terabytes
Alison Goate	Genetics and Genomic Sciences	422 terabytes
Samir Parekh	Oncological Sciences	339 terabytes

Top compute and storage usage department/institute

Department/Institute	Compute Core Hours
Genetics and Genomic Sciences	37,403,333
Psychiatry	7,206,416
Neurosciences	3,832,045
Oncological Sciences	2,023,745
AI and Human Health	1,141,662
Structural and Chemical Biology	1,091,584
Medicine	866,623
Precision Immunology Institute	809,616
Neurology	635,653
Institute for Genomic Health	487,952

Department/Institute	Storage (terabytes)
Genetics and Genomic Sciences	6,144
Psychiatry	1,781
AI and Human Health	1,741
Oncological Sciences	925
Neurosciences	847
Neurology	605
Precision Immunology Institute	228
Structural and Chemical Biology	220
Institute for Genomic Health	217
Microbiology	181

Top 10 PIs - GPU usage hours

PI	Department	GPU hours	# Jobs
Fuchs, Thomas	AI and Human Health	59,706	16,256
Filizola, Marta	Structural and Chemical Biology	48,576	3,419
Raj, Towfique	Neurosciences	24,961	5,292
Sumowski, James	Neurology	21,304	894
Charney, Alexander	Structural and Chemical Biology	12,067	56,565
Shen, Li	Neurosciences	9,548	2,064
Beck, Erin	Neurology	5,512	317
Schlessinger, Avner	Pharmacology	5,116	19,498
Kim-schulze, Seunghee	Oncological Sciences	4,858	1,243
Nadkarni, Girish	Medicine	4,002	914

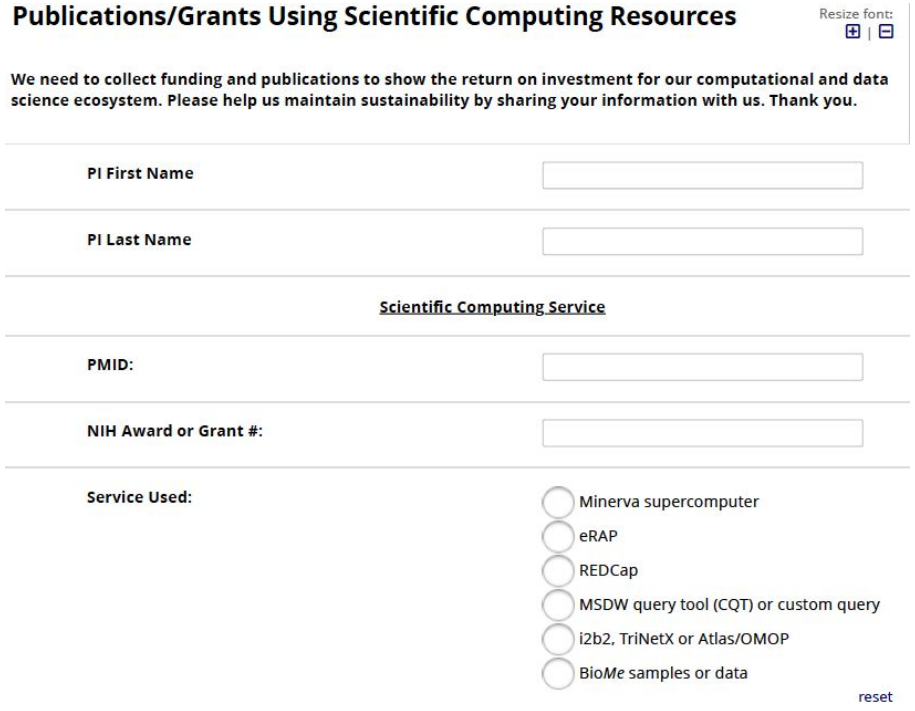
Total TSM Archival Storage Usage (Oct 2022- Mar 2023)

Current archive storage usage	
Archived data	17.7 PB (LTO5: 14.1 PB, LTO9: 3.6 PB)
Total data with offsite copy	35.4 PB (LTO5: 28.2 PB, LTO9: 7.2 PB)
Number of tapes used	21,205

Statistics of Oct 2022 - Mar 2023			
Amount of archived data	1, 437 TB	Amount of retrieved data	215 TB
# of users who have issued archive commands	61	# of users who have issued retrieve operations	40

Minerva Publications > 1,500 since 2012!!

We collect publications twice a year. Thank you!!!



Publications/Grants Using Scientific Computing Resources

Resize font:

We need to collect funding and publications to show the return on investment for our computational and data science ecosystem. Please help us maintain sustainability by sharing your information with us. Thank you.

PI First Name

PI Last Name

Scientific Computing Service

PMID:

NIH Award or Grant #:

Service Used:

- Minerva supercomputer
- eRAP
- REDCap
- MSDW query tool (CQT) or custom query
- i2b2, TriNetX or Atlas/OMOP
- Bio*Me* samples or data

reset

Year	# pubs
2012	55
2013	59
2014	62
2015	115
2016	150
2017	169
2018	141
2019	193
2020	159
2021	233
2022	143
2023*	21

* As of Feb. 2023

Kovatch P, Gai L, Cho H, Fluder E, Jiang D, Optimizing High-Performance Computing Systems for Biomedical Workloads, The 19th International Workshop on High Performance Computational Biology (HiCOMB), IPDPS, IEEE International Parallel and Distributed Processing Symposium, **May 2020**.

Kovatch P, Costa A, Giles Z, Fluder E, Cho H, and Mazurkova S, Big Omic Data Experience, SC'15: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis, **November 2015**.

2023 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?

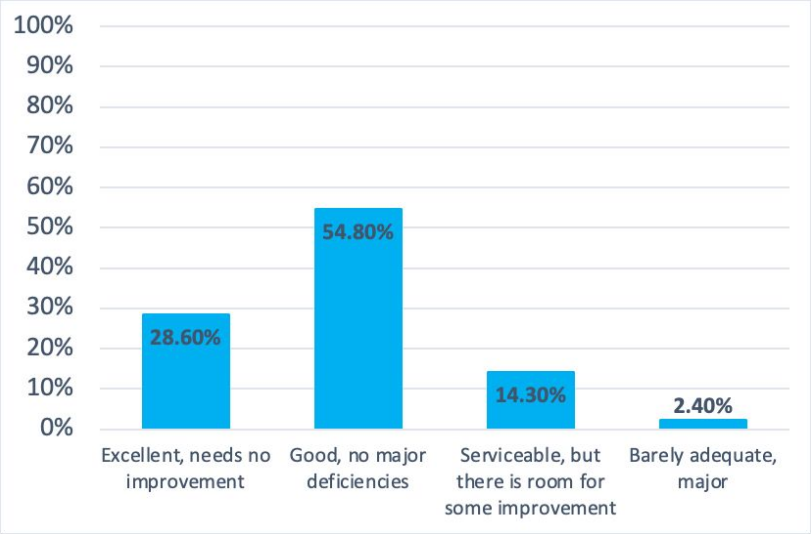
Q5: What suggestions do you have for improving our service?

We received 44 responses (5.2% response rate) and 31 comments.

Thank you for your feedback! This is the motivation for our 2023 Minerva Roadmap!

2023 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)?



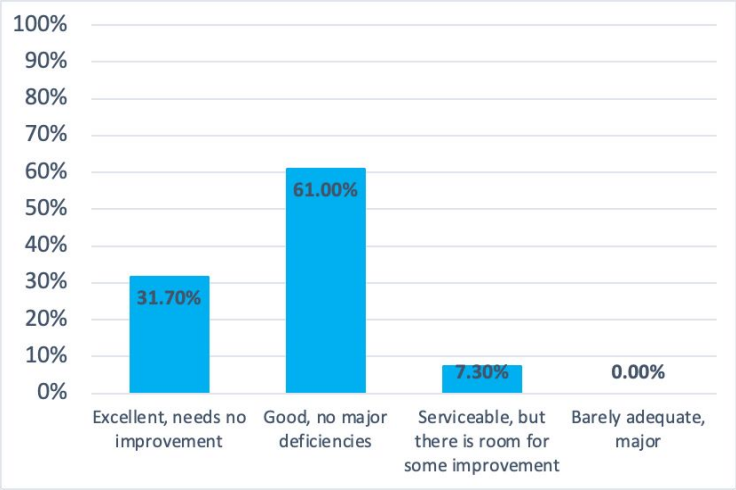
User satisfaction(>=Good)
2022: 84%
2021: 85%
2020: 81%
2019: 65%
2018: 54%

Raw comments(8):

- **LSF queuing**: Occasionally there are users running 1,000s of jobs at once and making it difficult to secure resources; the gpu queue is often not efficient. Balance the scheduler for shorter GPU jobs and distribute over more users to prevent single users hogging GPUs for days or weeks.
- **Nodes**: Need more GPUs, this is a rate limiting step for my research.

2022 Survey results question 2

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



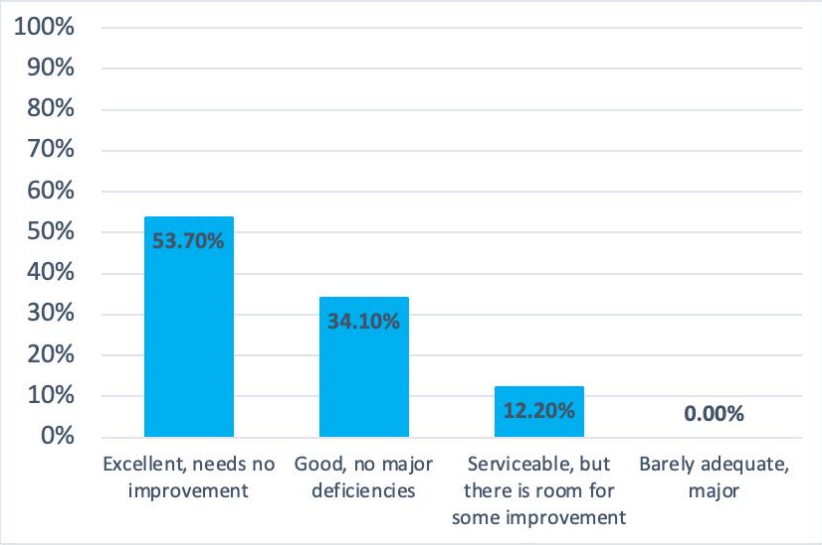
User satisfaction(>=Good)
2022: 93%
2021: 81%
2020: 80%
2019: 80%
2018: 67%

Raw comments(7):

- Some packages need to be updated.
- It would be useful to explore and support more container technologies such rootless docker, Shifter, Charliecloud and Podman.
- Ability to use Globus to transfer directly to/from MSSM OneDrive accounts would be very useful.
- Using jupyter notebooks on Minerva requires better documentation for dummies like me.

2022 Survey results question 3

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



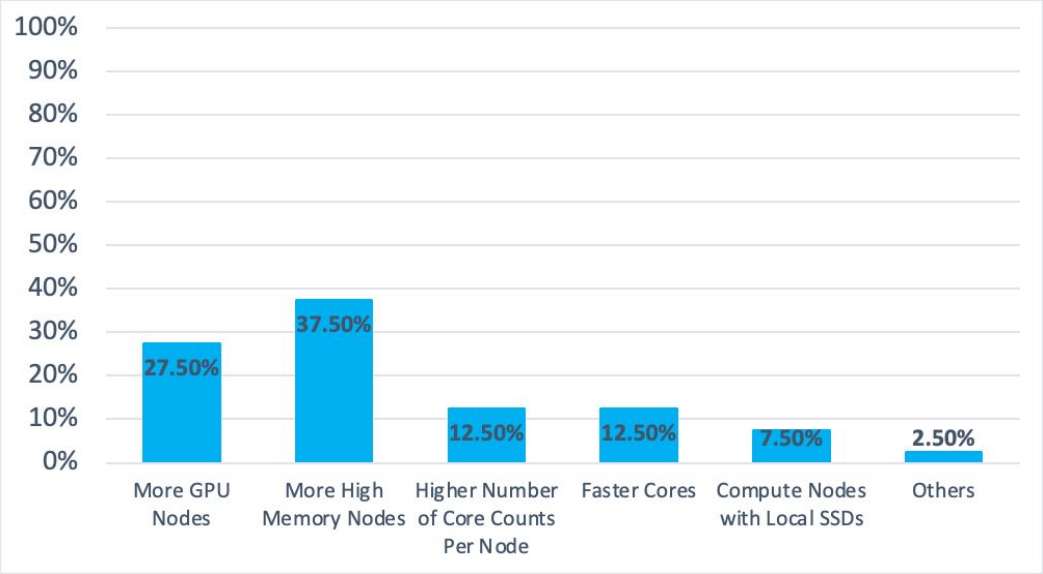
User satisfaction(>=Good)
2022: 88%
2021: 86%
2020: 91%
2019: 73%
2018: 80%

Raw comments (7):

- **Documentation:** More detailed documentation regarding software, databases, supported web servers and various advanced operations with the job scheduler would cut down the need to contact HPC for enquiries
- **Tickets:** Sometime it takes a week or more to receive an answer to a ticket.
- **Staff:** The staff is quite responsive to requests. But unaware of any previous history of dealing with a specific user which leads to endless repetitions of assessments already performed.

2022 Survey results and discussion

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



With the response, we will keep this mind with this upcoming RFP

- More GPUs ✔
- More High memory nodes ✔
- Higher # of cores per node ✔
- Faster cores ✔
- Some compute nodes with local SSD ✔

2023 Survey results summary

Thank you for your feedback! This is the motivation for our 2023 Minerva Roadmap

More raw comments:

- Can always be stronger, faster…
- Need more GPUs, this is a rate limiting step for my research. Significantly more GPUs are needed. Please get more GPU nodes or even the new intel Gaudi2 processors (https://habana.ai/training/gaudi2/) if GPUs are expensive. This would really help train much larger models that could be very useful to the genomics community. Having more GPUs available for training could dramatically speed this up and enable research that isn't currently being done at research universities.
- Occasionally there are users running 1000s of jobs at once and making it difficult to secure resources. It would be useful to explore resources such as the LSF job scheduling could be fairly or equally assigned to users/groups.

- I work with Minerva since 2017. I witnessed major improvements in service over the last few years as well as a substantial increase in resources. Fees for space usage are also very affordable. Communication has never been a problem, the team replies quite fast and, more importantly, always tries to accommodate our needs. Thank you.
- Excellent, needs no improvement
- The staff is quite responsive to requests.

We posted all the responses on our website.

Accomplishments & Updates

Accomplishments Summary (Oct 2022 - Mar 2023)

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

- ✔ Surpassed over **1,500** publications that utilized Minerva!!
- ✔ Hired 1 HPC system admin and 1 bioinformatician
- ✔ Minerva system upgrade including OS, networking stack and parallel file system
- ✔ New Annual HIPAA form launched in Jan 2023
- ✔ Purchased new H100 GPU nodes and LSF changes in GPUs
- ✔ High-Availability (HA) cluster for MSDW OMOP servers
- ✔ No cluster-wide Minerva PM (preventive maintenance)
- ✔ Expanded Data Ark - Mount Sinai Data Commons (18 data sets in total currently)
- ✔ Updated the documentation and presented 6 tutorial sessions
- ✔ Continued to support Minerva users through ticketing system (closed 1,714 tickets) and in-person meetings

Details will be presented in the following slides.

Thank you very much for the feedback!

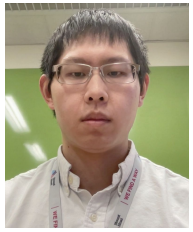
Staffing

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





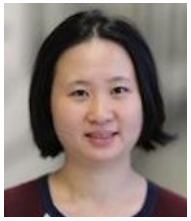
...and five HPC architects/admins positions

- Wei Guo, PhD (came back from 6 months leave of absence)
- Kali McLennan
- One HPC admin joining this June
- Jim Turner - part-time contractor
- **Two open positions:**
  - Lead HPC Architect
  - HPC Architect


We are actively hiring!!!

… and one Bioinformatician for Data Ark

- Yiyuan Liu joined May 1st



Minerva System Upgrade without Outages!

We performed a rolling upgrade of the operating system, networking stack and parallel file system software for Minerva during Dec. 2022 - Feb. 2023 and prevented any outages!

Why upgrade?

This is necessary to apply security patches, improve stability and benefit from the better features and support of the new Spectrum Scale version.

What is the plan?

We have tested the changes and will roll out the changes to the compute nodes to avoid a cluster-wide Preventive Maintenance (PM). The admins will drain a set of compute nodes out of the queue at a time (~15% of all nodes or less) based on the job loads and release them after the upgrade has been completed (the upgrade itself will take only a few minutes per node) .

The upgrade includes the base image to Centos7.9 (with kernel to 3.10.0-1160.el7.x86_64), the high-speed network software stack to (OpenFabrics Enterprise Distribution (OFED) 4.9) and the Spectrum Scale version (to 5.1.4.1).

What may you experience?

You may experience some delays in the job dispatch since there are less nodes available during the draining of the nodes.

New Annual HIPAA Form Launched in Jan 2023

Issues solved: ALL PIs can login and sign the online HIPAA form using DTP's active directory authentication

We launched a new HIPAA form this January with the authentication configured against ALL Mount Sinai ID including School and Hospital AD so ALL PIs can sign the online form

- Authenticate using your Mount Sinai login ID (not your email address) and password. *VIP Token is not needed.*
- From the green Login Options menu please choose either "Login with School Account" or "Login with Hospital Account".
- *A local Minerva account is no longer required for login.*

GPU Nodes and LSF Changes

To respond to users' feedback about limited GPU resources, we

Purchased 2 x New H100 GPU nodes - Planned in production May 2023

- 64 Intel Xeon Platinum 8358 2.6 GHz Processors per node
- 512 GB of memory per node
- 4 * 80 GB H100
- 3.84 TB 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
- PCIe

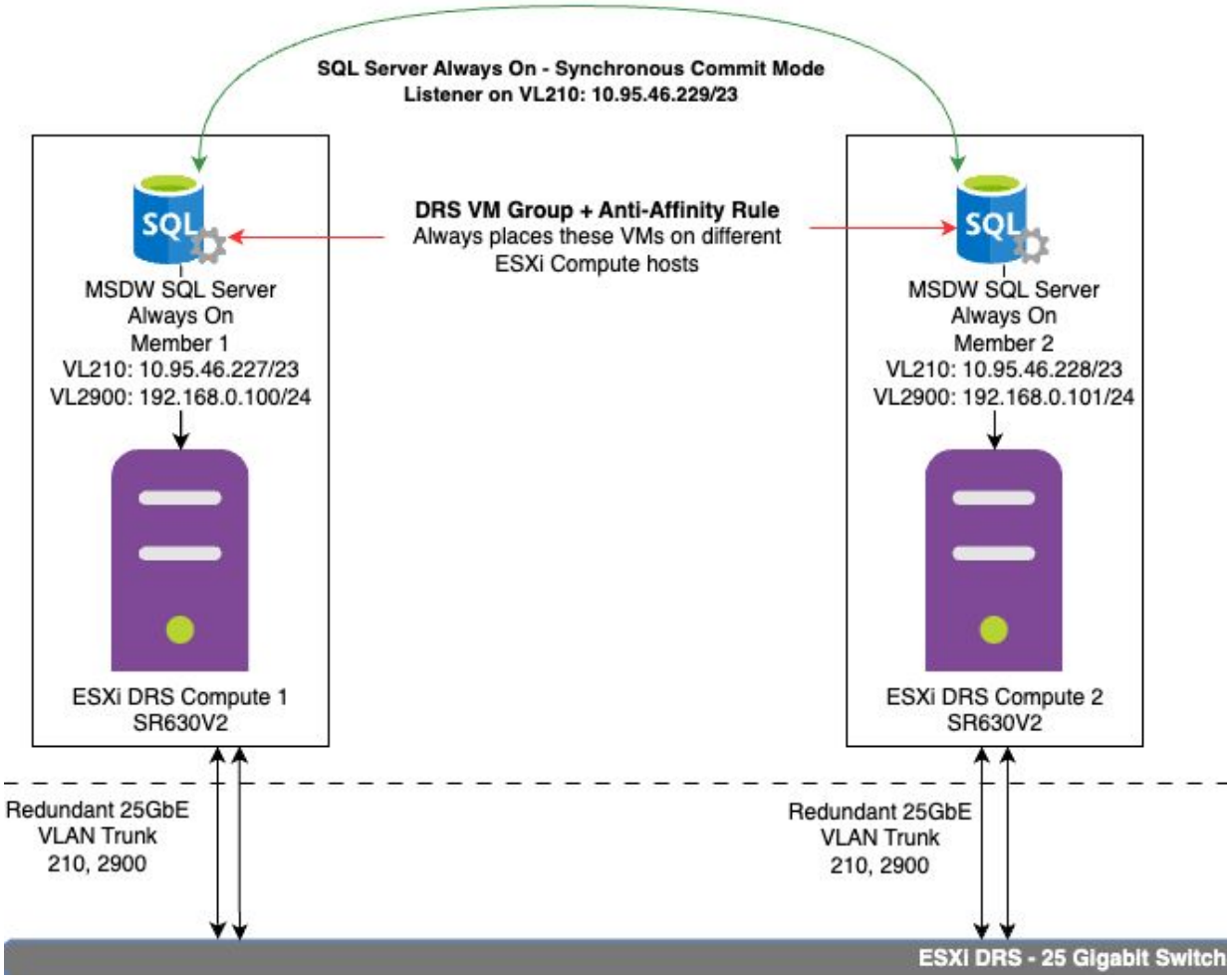
Set Global GPU number limit per user in March 15, 2023

- Set GPU number limits per user to 10
- Enable more fair sharing of GPUs

High Availability (HA) for MSDW SQL Servers

To provide protection against a failed ESXi host for SQL servers, we designed the DRS(Distributed Resource Scheduler) cluster with HA enabled.

- 2 x SQL production servers Always On and on different hypervisors



No Cluster-Wide PM in last six months

With all the updates we had on the system, system admins managed to do it with rolling upgrade without system-wide downtime

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

An unexpected outage on Minerva login in 11/02/2022

- Caused by DTP unannounced changes on a security certificate with mistakes
- We will work with the IT team for better change management to avoid this from happening again.

Thanks to our admins!!!

Data Ark Data Commons: Streamlined Data Access within 24 hours

There are 18 data sets hosted under Data Ark currently

- Immediate access to 11 public-unrestricted data sets
- Access within 24 hours to 5 Mount Sinai-generated data sets

Immediate Access

Public Data Sets

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

Access within 24 hours

Mount Sinai Generated Data

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

Restricted Access

Public Data Sets

- UK Biobank

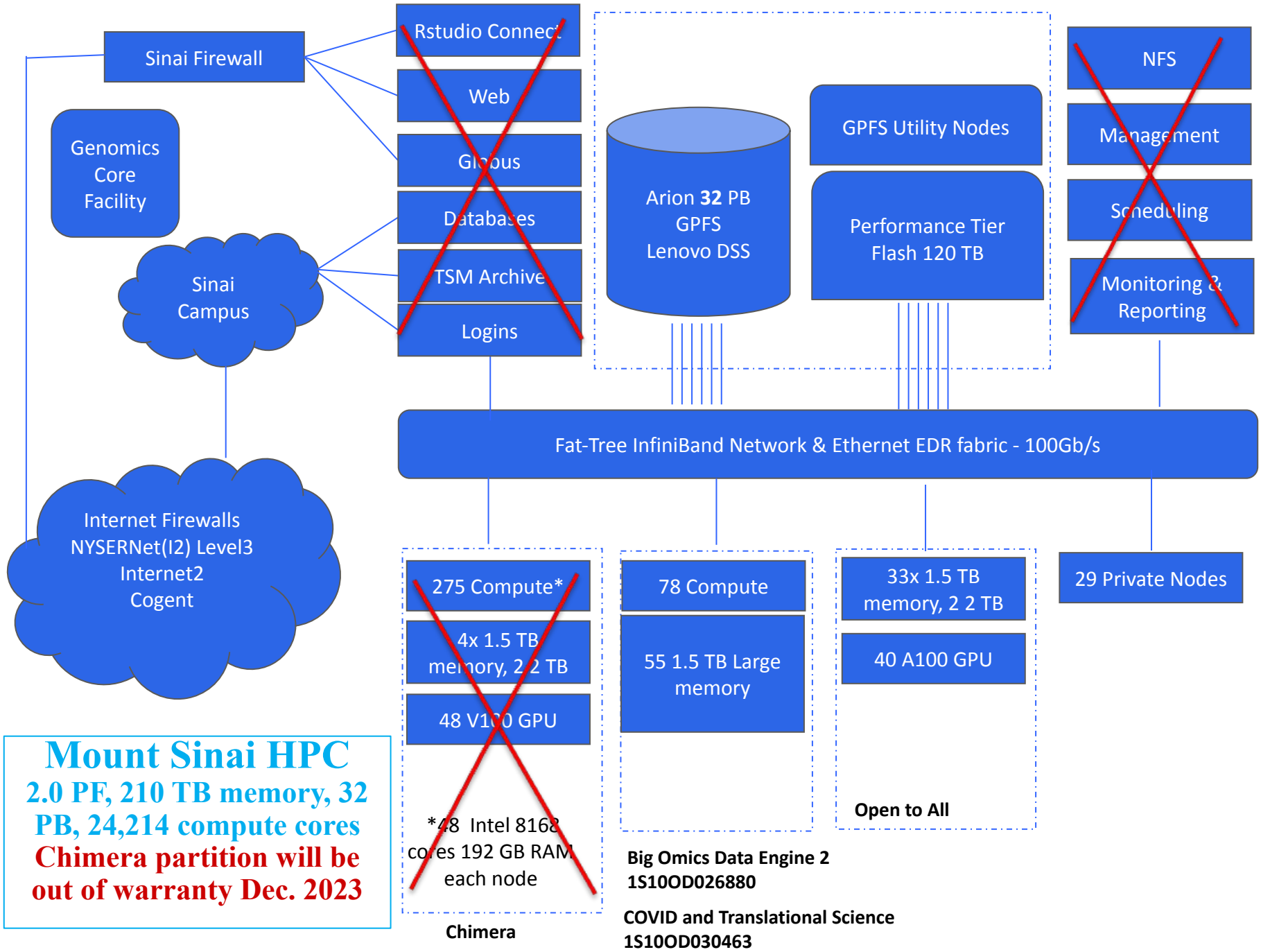
School-Acquired Data Sets

- MarketScan®

Documentation and Training Sessions

- Documentation updated
  - Our website at https://labs.icahn.mssm.edu/minervalab/
  - We provided additional training material (including slides & recording) online
- Offered training sessions in person/Zoom:
  - Six training sessions in spring and fall
  - Topics include “Introduction to Minerva”, “LSF job scheduler” & “Singularity Container” & “Running Jupyter Notebook and RStudio on Minerva”
  - **60 -100 participants** per session
- HPC Town Hall in person/Zoom
  - Twice yearly. Next on May 3rd
- 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



Mount Sinai HPC

2.0 PF, 210 TB memory, 32 PB, 24,214 compute cores

Chimera partition will be out of warranty Dec. 2023

Minerva Refreshment RFP (Request for Proposal)

Plan to release RFP to vendors in early May

Challenge in deployment - no available cooling capacity for new nodes in Mount Sinai data centers

We are investigating cooling with

- **Additional Chiller to Hess**
- **Remote data centers**
  - ➔ Expensive annual fee and considerable time for set up plus slow networking between Sinai campus and the remote data center
  - ➔ Don't want to move the 15 PB non-backed up file system (GPFS) to the remote data center

Researching possible options on how to operate GPFs for current and new nodes

1. Serve GPFS from Sinai campus to the remote data center over a high-bandwidth link
2. Copy GPFS to the remote data center

TSM Archival Storage LTO-5 Tape Solution

Minerva TSM Archival Storage LTO-5 Tape Solution will be out of support on 12/31/2023

Status:

Two generations of TSM Archival Storage operating on Minerva including LTO-5 tape solution deployed in 2012 and LTO-9 tape solution deployed in 2022. LTO-9 is used for all data archived/backup since then, while LTO-5 is only serving for the old data retrieval.

Upcoming Changes

IBM is discontinuing service and support on LTO-5 tape solution on 12/31/2023 with details here. As a result, we will no longer be able to guarantee access to the data archived on those LTO-5 tapes after that. Any failures within that library will be uncorrectable.

Effect

All data archived/backed up to LTO-5 tapes (i.e., prior to 05/10/2022) may be affected.

~~***Your action needed (revised May 2nd)***~~

~~**If you have data archived/backed up to LTO-5 tapes (i.e., prior to 05/10/2022), you will need to retrieve and re-archive any data that you still wish to retain.** See TSM Guide on how to retrieve and archive. It will take some time to complete this process for large data, so please plan your activity accordingly.~~

Our admin will help migrate the data to LTO-9 tapes.

2023 Roadmap Continued

Launch Open OnDemand as better visualization portal to access Minerva through web browsers in Q2

- Finalizing application supports
- Planned for July 1st. Slowed due to limited system admins

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

- Upgrade MariaDB and migrate the storage to local SSD

Migrate Minerva two factor authentication to Azure MFA

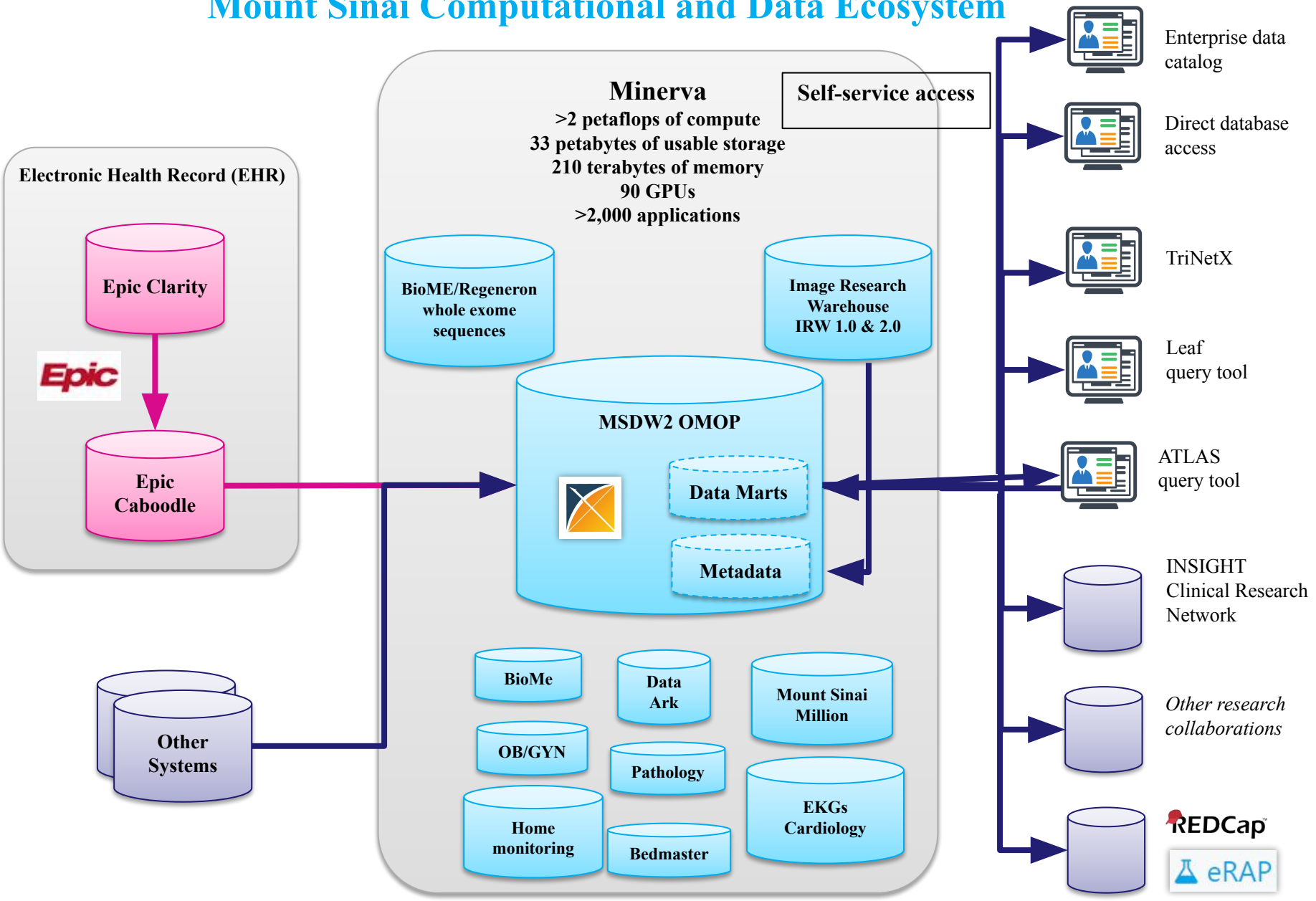
- The current Symantec VIP will be gradually deprecated starting in July

Install and config new storage server for user $HOME for improved performance and stability

Data Ark Expansion

- MSDW de-identified OMOP dataset (entire copy)
- De-identified digitized pathology slides linked to slide metadata in MSDW
- Imaging Research Warehouse 2.0
- Bedmaster patient monitoring data

Mount Sinai Computational and Data Ecosystem



Acknowledge CTSA

Please acknowledge CTSA in your publications

- Supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Advancing Translational Sciences, National Institutes of Health.

Thank you!