Minerva Town Hall
Oct 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
Yiyuan Liu, Bioinformatician
Catherine Mccaffrey, Project Manager
Ranjini Kottaiyan, MBA, OD, Senior Director, Finance

Oct. 26, 2023

Icahn
School of
Medicine at
Mount
Sinai

Outline

Metrics & Status

- April 2023 - Sep 2023 Minerva usage
- Over 1,500 publications have used Minerva since 2012!

Accomplishments & Updates

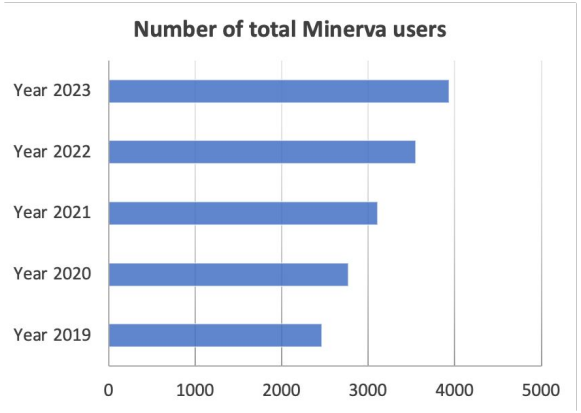
- Staffing
- Launched Open OnDemand: New graphic interface to Minerva via web browser
- LSF Operation: Changes on Resource Allocation Limit
- 2x New H100 GPU nodes in production this June
- Minerva refreshment in process
- Upgraded and migrated to new server: Posit Connect and MySQL database
- Migrating data from Archival storage LTO-5 tape solution to LTO-9
- Minerva PM (preventive maintenance)
- Documentation and training sessions
- Data Ark - Mount Sinai Data Commons

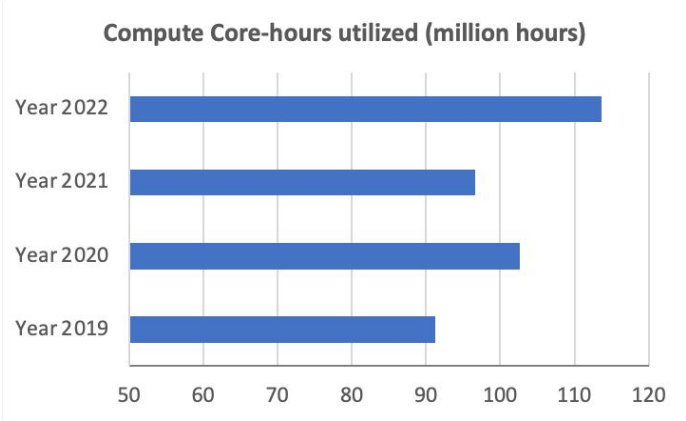
What's next?

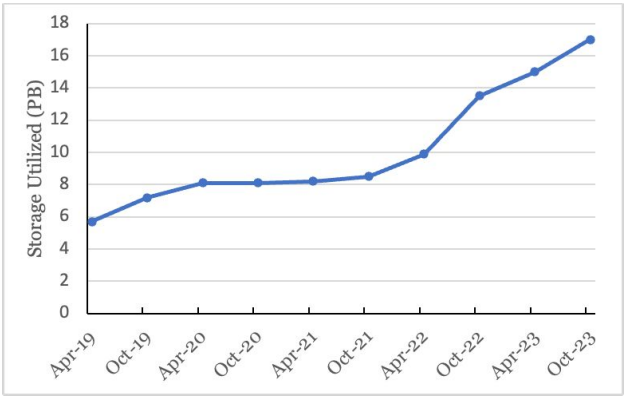
Q & A

Minerva Usage
(April 2023 - Sep 2023)

Minerva usage over years







# of users increase by 300 - 400 every year

Compute usage higher in year 2020

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

Storage usage doubled in the last two years

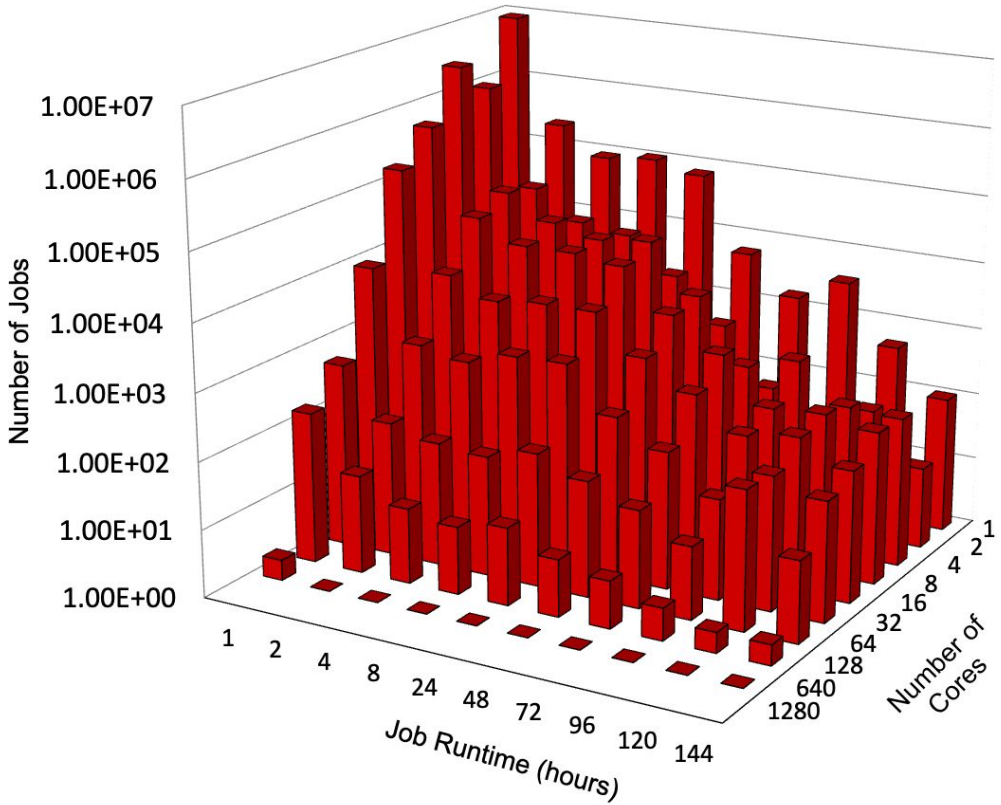
Minerva usage summary (April 2023- Sep 2023)

Accounts	
Number of active users	1,102
Number of total users	3,896
Number of project groups	486 (392 active)
Storage	
High-speed storage used (Arion)	17.0 PB (52% utilization) 7,464,377,393 Files
Archival storage used	16.6 PB
Compute	
Number of jobs run	14,258,808
Core-hours utilized	51,489,313 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	7,480,315	25,076,000	47.6 %
BODE2	3,446,208	9,484,848	63.2 %
Hi-memory nodes	2,007,609	4,839,953	78.6 %
CATS	1,270,666	9,491,198	64.9 %
GPU nodes	54,010	2,597,314	67.3 %
Total:	**14,258,808**	**51,489,313**	**55.8 %**

Job Mix


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
Roussos, Panos	Psychiatry	5,341,686	2,291,873
Schlessinger, Avner	Pharmacology	4,297,392	4,341
Sharp, Andrew	Genetics and Genomic Sciences	3,860,623	650,865
Charney, Alexander	Genetics and Genomic Sciences	3,307,426	2,145,447
Chowell, Diego	Precision Immunology Institute	2,784,709	24,402
Fang, Gang	Genetics and Genomic Sciences	2,606,403	140,949
Raj, Towfique	Genetics and Genomic Sciences	2,457,312	2,449,515
Goate, Alison	Genetics and Genomic Sciences	1,872,338	359,415
Pinto, Dalila	Genetics and Genomic Sciences	1,756,344	233,564
Buxbaum, Joseph	Genetics and Genomic Sciences	1,401,288	829,878

Top 10 PIs GPFS high speed storage

PI	Department	Storage usage
Thomas Fuchs	AI and Human Health	2.5 PiB
Bin Zhang	Genetics and Genomic Sciences	1.3 PiB
Alexander Charney	Genetics and Genomic Sciences	1.2 PiB
Panagiotis Roussos	Genetics and Genomic Sciences	1.1 PiB
Towfique Raj	Neurosciences	807 TiB
Robert Sebra	Genetics and Genomic Sciences	794 TiB
Stuart Sealfon	Neurology	668 TiB
Alison Goate	Genetics and Genomic Sciences	518 TiB
Joseph Buxbaum	Psychiatry	514 TiB
Girish Nadkarni|Alexander Charney	Genetics and Genomic Sciences	510 TiB

Top 10 PIs - GPU usage hours

PI	Department	GPU hours	# Jobs
Fuchs, Thomas	AI and Human Health	673,894	16,391
Filizola, Marta	Structural and Chemical Biology	308,558	1,004
Schlessinger, Avner	Pharmacology	275,943	2,153
Raj, Towfique	Neurosciences	233,724	11,932
Beck, Erin	Neurology	157,576	724
Nadkarni, Girish	Medicine	141,677	2,130
Kenny, Paul	Neuroscience	71,742	46
Kim-schulze, Seunghee	Oncological Sciences	71,359	921
Roussos, Panagiotis	Psychiatry	70,137	1,105
Fang, Gang	Genetics and Genomic Sciences	69,979	1,284

Total TSM Archival Storage Usage (April 2023- Sep 2023)

Current archive storage usage	
Archived data	16.6 PB (LTO-5: 11.0 PB, LTO-9: 5.7 PB)
Total data with offsite copy	34.0 PB (LTO-5: 22.7 PB, LTO-9: 11.3 PB)
Number of tapes used	17,476 (16,753 LTO-5 + 723 LTO-9)

Statistics of April 2023 - Sep 2023			
Amount of archived data	1,299 TB	Amount of retrieved data	702 TB
# of users who have issued archive commands	98	# of users who have issued retrieve operations	67

Minerva Publications > 1,500 since 2012!!

We collect publications twice a year (Jan & June). Thank you!!!

We sent email to PIs and delegate

Year	# pubs
2012	55
2013	59
2014	62
2015	115
2016	149
2017	165
2018	133
2019	178
2020	146
2021	234
2022	174
2023	69

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**.

Accomplishments & Updates

Accomplishments Summary (April 2023 - Sep 2023)

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

- ✔ Continued to support Minerva users through ticketing system (closed 3,103 tickets) and in-person meetings
- ✔ Staffing
- ✔ Launched Open OnDemand: New graphic interface to Minerva via web browser!!!
- ✔ LSF Operation: Changes on Resource Allocation Limit
- ✔ 2x New H100 GPU nodes in production this June
- ✔ Minerva Refreshment and Expansion in Progress
- ✔ Upgraded and migrated to new server: Posit Connect and MySQL database
- ✔ Migrating data from Archival storage LTO-5 tape solution to LTO-9
- ✔ No cluster-wide Minerva PM (preventive maintenance)
- ✔ Updated the documentation and presented 6 tutorial sessions
- ✔ Data Ark Updates - Mount Sinai Data Commons (18 data sets in total currently)

Details will be presented in the following slides.

Thank you very much for the feedback!

Staffing

The HPC team consists of three senior/computational scientists

- **Goodbye and huge thanks to Eugene Fluder, PhD**
  - Worked here for 11.5 years; Retiring at Oct 31 2023
- Hyung Min Cho, PhD
- Jielin Yu, PhD

and six HPC architects/admins positions

- Wei Guo, PhD
- Kali McLennan
- **Four open positions:**
  - Associate Director
  - Lead HPC Architect
  - HPC Architect
  - HPC admin

… and one Bioinformatician for Data Ark

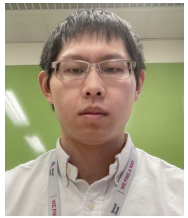
- Yiyuan Liu, PhD

With the Sinai hiring freeze, responses to users' requests will be delayed up to 72 hours due to being short-staffed (5 open positions)

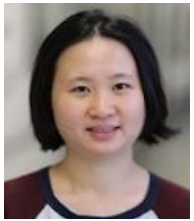
The HPC Team

Eugene Fluder
Senior Computational Scientist

Hyung Min Cho PhD
Senior Computational Scientist



Jielin Yu
Computational Scientist



Yiyuan Liu PhD
Bioinformatician

Wei Guo PhD
HPC Architect

Kali McLennan
Senior HPC Admin

Open OnDemand is available: New graphic interface to Minerva via web browser

Purpose: Providing easy graphical access to Minerva without Linux command needed

Status: Open OnDemand service portal to access Minerva in production on Aug 2 2023!!

- There are >100 users utilizing this currently
- Extra supported needed on this new service
- A training session provided on Open OnDemand on October 11

This product offers a fully-compliant job management and desktop portal requiring minimal knowledge of Linux high-performance computing (HPC) environments with no end-user installation requirements other than an up-to-date web browser (Chrome or Firefox recommended)

The service portal is accessed at URL: https://ondemand.hpc.mssm.edu
Documentation is available at https://hpc.mssm.edu Documentation>Open OnDemand or here

LSF Operation: Changes on Resource Allocation Limit

Goal:

- Stop users (needing heavy resource & misusing resource) from blocking the shared cluster resource and make better balance between users
- Maintain high throughput for Minerva users

Action:

- We set two global limits (keep monitoring and tune as needed):
  - ✔ Max number of running job slots is 4,000 per user
  - ✔ Max requested memory of all jobs is 15TB per user (added on 08/25/2023)
- We are also working with users to use Minerva appropriately, i.e., requesting resources in accordance with their job needs and job type. So far > 100 users contacted (~10% of active* users).
  - If you need help to understand and characterize your job, please reach out to our support team at hpchelp@hpc.mssm.edu immediately. You can also refer to our LSF documentation at HERE.
  - We are actively monitoring the resource usage of user jobs on Minerva and contacting the users who are misusing the resources. Please respond to us in time and modify your job submission script as needed.
  - Those who are not respondent will have their jobs removed and their access to Minerva will be limited till the issue is addressed accordingly

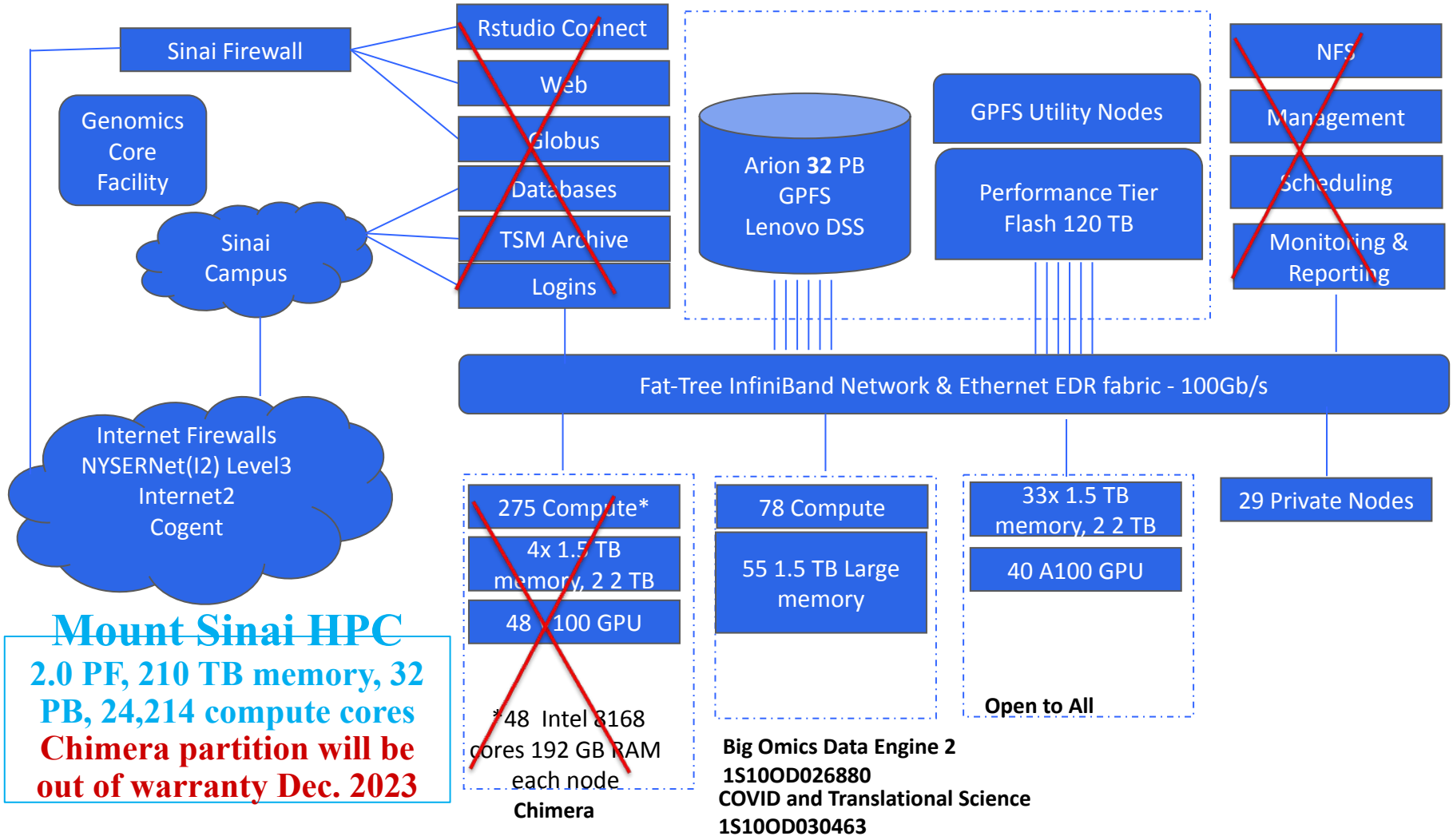
*Active means user has activity in the last 6 months.

New 2x H100 GPU nodes in production

In production June 2023!!!

2x latest H100 GPU nodes

- **8x H100** Tensor Core GPUs in 2 nodes and check its datasheet here; 80 GB of memory for each GPU, for a total 320 GB per node
- 64 Intel Xeon Platinum 8358 2.6 GHz Processors per node, a total of 128 cores; 0.5TB memory per node
- 7 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
- The H100 is connected via PCIE
- Request by “ `-R h10080g`”


Sinai Firewall
Genomics Core Facility
Sinai Campus
Internet Firewalls
NYSERNet(I2) Level3
Internet2
Cogent
Rstudio Connect
Web
Globus
Databases
TSM Archive
Logins
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*
4x 1.5 TB memory, 2 2 TB
48 100 GPU
*48 Intel 8168 cores 192 GB RAM each node
Chimera
78 Compute
55 1.5 TB Large memory
Big Omics Data Engine 2
1S10OD026880
COVID and Translational Science
1S10OD030463
33x 1.5 TB memory, 2 2 TB
40 A100 GPU
Open to All
29 Private Nodes
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 and Expansion in Progress

Budget "Approved": $10M equipementin HPC

- Enable more GPUs and high memory resources (2 TB large shared memory nodes) to our users

Status: Delayed due to unexpected financial considerations for the cooling/power

- It will take 6-9 months for facility to upgrade the Data Center with interim cooling capacity
- It will take 12-24 months for facility to upgrade the Data Center with long term cooling capacity

The Refreshment Processes:

- RFP issued for bidding on May 8th ✔
- Proposal from all the vendors received and revision on Aug 9th ✔
- Expected Down Select to an RFP Award Finalist on Aug 18 - Delayed
- Negotiations with Selected Vendor and Process the Order
- System delivery and Install
- Acceptance Test
- Soft open to friendly users
- Open to public

Service Upgraded: Posit Connect and MySQL database

- MySQL/MariaDB database
  - The service is migrated to a new host and upgrade to the new MariaDB 11.0.1 release for better performance and service
  - Please connect as usual with hostnames **data1.chimera.hpc.mssm.edu** (IP 172.28.7.78, on cluster) or **data1.hpc.mssm.edu** (10.95.46.229, on campus network).

- Rstudio Posit connect:
  - The service is now migrated to a new host and upgrade to 2023.06 version of Posit Connect for better performance and service
  - To login to the web: rstudio-connect.hpc.mssm.edu, please use your minerva username and password joined with <6 digit vip token>, like ssh in to the login node.

TSM Archival Storage LTO-5 Tape Solution

Updates: 20% of all archived files have been migrated from old LTO-5 tapes to new LTO-9 tapes (14 PB in total) since June.

(We are migrating everyone's files to the new tapes for maximal user convenience).Extra effort from our system admin.

Minerva TSM Archival Storage LTO-5 Tape Solution will be out of support soon.

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.

No Cluster-Wide PM in last six months

System admins try to minimize the system-wide downtime

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

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:
  - Four training sessions offered this fall
  - Topics include “Introduction to Minerva”, “LSF job scheduler” & “Open OnDemand” & “Data Ark/UKBB”
- HPC Town Hall in person/Zoom
  - Twice yearly
- 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

Data Ark Data Commons Datasets

There are 18 datasets hosted under Data Ark currently

- Immediate access to public-unrestricted datasets
  - **Datasets in gray are subject to removal**
- Access within 24 hours to Mount Sinai-generated datasets

Immediate Access

Public Data Sets

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

Access within 24 hours

Mount Sinai Generated Data

- The CBIPM-BioMe Data Set (coming soon pending on IRB approval)
- MSDW De-identified OMOP 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 (changes to be released soon)

User Group-Acquired Data Sets

- MarketScan®

Data Ark Updates

- Welcome our new clinical science director Dr. Cyrus Hedvat
- Welcome our new member of data ark advisory board Dr. Alexis Zebrowski
- Reviewed Policy on Public Data Sets and requesting approval from original owner
- Submitted an IRB to host the Biome data
- A user group model built for MarketScan data set
- New dataset onboarded - MSDW OMOP data mart
- Research Engagement: Presented on TCI Committee meeting, Sinai CTSA Research Day, GGS faculty, and training session

What’s next?

What's next?

Continue working on Minerva refreshment and expansion

- Preparation for the deployment of new machines including new network, new OS and management servers including NFS servers, xcat, mail server, proxy server, LSF, LDAP, etc

Continue to migrate data from Archival storage LTO-5 tape solution to LTO-9

Survey users on job schedulers LSF vs SLURM

Annual Form Collection (HIPAA and CTSA/S10 acknowledgement form) in Q1 2024

Annual HPC User Survey in Q1 2024

Migrate Minerva two factor authentication to Azure MFA - Q4 2024

Data Ark Expansion

- De-identified digitized pathology slides linked to slide metadata in MSDW

Acknowledging S10/CTSA in your publications

CTSA Clinical & Translational Science Awards®

Minerva are funded by CTSA & NIH S10 - crediting CTSA/S10 support is imperative

Annual user agreement required for everyone starting January 1, 2024

- With logging into one site, you will see a list of form for you to review and sign annually

You must acknowledge the support from Scientific Computing and Data at the Icahn School of Medicine at Mount Sinai by including the following acknowledgement in a publication of any material, whether copyrighted or not, based on or developed with Minerva HPC resources:

"This work was supported in part through the computational resources and staff expertise provided by Scientific Computing and Data at the Icahn School of Medicine at Mount Sinai and supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Advancing Translational Sciences."

Your must further acknowledge the NIH S10 awards by including the following acknowledgement in a publication of any material, whether copyrighted or not, based on or developed with NIH S10 funded equipment (BODE2-Big Omics Data Engine 2 and CATS-COVID and Translational Science Supercomputer):

"Research reported in this paper was supported by the Office of Research Infrastructure of the National Institutes of Health under award number S10OD026880 and S10OD030463. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health."

Only 4 publications acknowledged CATS and 12 acknowledged CTSA in 2023

Thank you!