**High-Performance Computing (Resources and Capabilities)**

<https://github.com/wsphd/csun-hpc/>

**"High-Performance Computing (Resources and Capabilities)**  
**California State University, Northirdge (CSUN)**

**Friday, April 11, 2025**  
**CSUN Faculty Retreat - Odyssey Restaurant**

**Wayne Smith, Ph.D., *Department of Management*,** [**ws@csun.edu**](mailto:ws@csun.edu)

(Note: This page is also available in alternate formats - [.pdf](https://github.com/wsphd/csun-hpc/blob/main/alt-formats/csun-hpc.pdf), [.odt](https://github.com/wsphd/csun-hpc/blob/main/alt-formats/csun-hpc.odt), and [.docx](https://github.com/wsphd/csun-hpc/blob/main/alt-formats/csun-hpc.docx))

**Introduction/Background/Motivation**

* Some n i N needs are ≤ contemporary desktop/laptop and software
  + But double-check new methodologies and growth (and by extension, movement) of data
* Some n j N needs are > contemporary desktop/laptop and software
  + Essentially, "compute-intensive, data-intensive, or network-intensive"
  + Use primarily FOSS (Linux, Open Source, etc.) to complement COTS (Windows, SPSS, etc.)
* Private, "on-premises" servers (contact: [Zack Hillbruner](mailto:zack.hillbruner@csun.edu), *CSUN IT*)
  + Usually purchased by an individual faculty member or Dept. (often with a grant or project)
  + Usually located in the on-campus CSUN MDF
  + CSUN IT usually racks and networks the system; Users manage the system and applications
* Public Cloud (AWS, GCP, MS-Azure, OCI, Digital Ocean, etc.)
  + Use "free-tier" (still need to provide a credit card)
  + Purchase credits with a credit card
  + Scholars can ask for resources for substantive reseearch
* Or?
  + NSF-funded, multi-year, inter-institution, STEAM/SocialSTEM, R3s/CCC's too
  + [CSUN IT Technology Resources for Research](https://www.csun.edu/it/technology-resources-research)

**General Advanced Computing/Data Management**

* There are plenty of (non-HPC) advanced computing issues too (research and instruction).
* Ecosystem Transition: Compute
  + Beyond CPUs, there GPUs, FPGAs, DPUs, and others
  + COTS languages (e.g., SPSS, Stata, MPlus, Matlab, NVivo) -> FOSS languages (e.g., R, Python, Julia)
  + Single-threaded execution -> Parallel execution
  + COTS spreadsheets (e.g., Excel ) -> FOSS spreadsheets(e.g., LibreOffice).
  + Operating Systems (e.g., Windows/MacOs ) *plus* Linux, Excel -> LibreOffice, etc.
  + Beyond replication -> Reproducibility (not just 'A' journals)
* Ecosystem Transition: Data
  + "Big Data"
  + Research results can include output data (and perhaps even source data) too
  + Desktop sizes (e.g., GiB, TiB) -> Beyond-Desktop sizes (e.g., PiB, ExiB)
  + Human-readable file formats with no meta-data (e.g., CSV) -> Digital file formats with meta-data (e.g., Parquet)
  + Row-oriented databases (e.g., MariaSQL, PostgreSQL, SQLite, etc.) -> Column-oriented databases (e.g., DuckDB, MonetDB, TileDB)
  + Monolith APIs (e.g., REST) -> Robust APIs (e.g., GraphQL)
  + Single-file access (e.g., HTTPS) -> Multi-file 'buckets' (e.g., AWS/S3, GCS, Azure/Blog)
* Ecosystem Transition: Network
  + "High Throughput"
  + Big Data needs to be moved over fast, reliable networks
  + CSUN 'Science DMZ'
  + Los Angeles Public Library (LAPL) Northridge branch and Mid-Valley branch (>2x file transfer, symmetric)
* Example: Technology Trends
  + Campus Labs *plus* Home Labs, Open Science, Open Research, Open Data, Open anything...
* I'm happy to discuss these issues too but it's not the primary focus of this material.

**Faculty/Disciplines**

* CSUN is a big place--there are many faculty doing interesting things with HPC
* Ravi Absol (Chemistry)
  + NRP
  + Molecular Dynamics
  + undergraduate student (Anita) - uses AMBER software (with many GPUs, including NVIDIA A100s)
  + one protein combinatorics imulation took 3 months on a laptop -- now takes ~ 3 weeks
* Bingbing Li (Manufacturing Systems Engineering and Management)
  + NRP
  + Smart Manufacturing (e.g., Industrial AI, Multimodal Data Fusion for Autonomous System, Digital Twins, XR and Metaverse for Manufacturing)
* Xiyi Hang (Electrical and Computer Engineering)
  + NRP
  + Bioinfomatics/Bio-medical applications (e.g., Gene expression, data mining)
* Xunfei Jiang (Computer Science)
  + on-premises HPC and OSG/NRP
  + Cloud and Infrastructure design and management
  + Workforce development (full stack)
* Dr. Nicholas Kioussis (Physics)
  + CalTech Quantum Computing (and others)
  + Materials Science
* Wayne Smith (Management)
  + ACCESS and NRP
  + Federal Communications Commission (FCC) Universal Licensing System (ULS) data
  + Github - [R code](https://github.com/wsphd/fcculs)
  + Data output - [multi-file data output](https://www.qsl.net/n/n6lhv/scma/fcculs/)
* Other future CSUN projects
  + CTVA - UNREAL 3D imaging (e.g., digital twins, gaming, 3D mapping)
  + Digital Humanities
* Representative CSU projects
  + SDSU
  + CSUSB - [High Performance Computing](https://www.csusb.edu/faculty-center-for-excellence/idat/high-performance-computing)
  + Sonoma State
  + CalPolyHumboldt
* CSU help
  + Dr. Dung Vu (CSUSB)
  + Kyle Krick (SDSU)

**JupyterHub**

* [Multi-user, interactive notebooks (including R, Python, Julia, etc.)]<https://jupyter.org/hub>
* CSU Technology Infrastructure for Exploration ("TIDE")
  + *governance*: funded by NSF, provisioned by Cloudbank, managed/supported by UC Berkeley
  + *hardware*: basic access
  + *software*: R/RStudio, Python, Shiny, Linux terminal, generic notebooks
  + *log in*: Select "California State University, Northridge" and then log in is via CILOGON (just log in as usual)
  + Start here: [CAL-ICOR]<https://csun.cloudbank.2i2c.cloud/>
* CSU Technology Infrastructure for Exploration ("TIDE")
  + *governance*: funded by NSF, provisioned by NRP, managed/supported by SDSU
  + *hardware*: can have access to multiple GPUs, multiple cores, lots of RAM, some (shareable) storage
  + *software*: R/RStudio, Tensorflow, Datascience, generic notebooks, discipline-specific notebooks
  + *log in*: log in is via CILOGON (just select "California State University, Northridge" and log in as usual)
  + Start here: [CSUN TIDE]<https://tide.sdsu.edu/>
* CSUN myCSUNSoftware (aka "Apporto")
  + n/a (but other GUI and CLI software runs here)
  + Faculty will need to ask for it over time
  + Start here: [myCSUNSoftware]<https://www.csun.edu/it/software-services/software/all-software/mycsunsoftware>

**Jetstream2/ACCESS**

* Managed by Indiana University
* 100's of GiB of RAM, 10's of PB of disk, 10's of [GPUs](https://en.wikipedia.org/wiki/Graphics_processing_unit), fast networks
  + Best for new learners, data science projects (R, Python, Julia, etc.), large simulations, gateway to other systems, including several supercomputers around the country
* Need an "ACCESS ID"
  + Like an ORCID ID but for Research Computing
  + Have CV or Resume for upload (don't worry, your request will be approved)
* Be willing to learn:
  + How to ask (nicely and well, for more (incrementally) resources, and read a simple dashboard
  + the Command line and Linux
  + Webshell
  + SSH for logging int (and some learning curve for generating SSH keys and passphrases)
  + SCP for file transfer (after the SSH process is done)
* (Live demo...if possible)
* Start here:
  + [Jetstream ACCESS page]<https://jetstream-cloud.org/get-started/index.html>

**Nautilus/NRP**

* Managed by University of California, San Diego
* 100's of GiB of RAM, 10's of PB of disk, 100's of [GPUs](https://en.wikipedia.org/wiki/Graphics_processing_unit)/[FPGAs](https://en.wikipedia.org/wiki/Field-programmable_gate_array)/[TPUs](https://en.wikipedia.org/wiki/Tensor_Processing_Unit)/[DPUs](https://en.wikipedia.org/wiki/Data_processing_unit), very fast networks
  + Best for leading-edge science and engineering, especially w/ funded labs and staff
* Be willing to learn:
  + Must be comfortable with the Command Line, Open Source, and Linux
  + Kubernetes (open source client-server), you use the "kubectl" binary
  + You control just about everyting with ASCII ".yaml" files
* (Static demo -- sample .yaml configuration file with GPUs)

apiVersion: apps/v1

kind: Deployment

metadata:

name: deployment-amber-gpu-ws

labels:

k8s-app: deployment-amber-gpu-ws

spec:

replicas: 1

selector:

matchLabels:

k8s-app: deployment-amber-gpu-ws

template:

metadata:

labels:

k8s-app: deployment-amber-gpu-ws

spec:

containers:

- name: mypod-gpu

image: gitlab-registry.nrp-nautilus.io/prp/jupyter-stack/tensorflow

resources:

limits:

memory: 512Gi

cpu: 1500m

nvidia.com/gpu: 1

requests:

memory: 512i

cpu: 1500m

volumeMounts:

- name: mydata

mountPath: /mnt/myscratch

command: ["sh", "-c", "sleep infinity"]

volumes:

- name: mydata

emptyDir: {}

sizeLimit: 10Ti

nodeSelector:

nautilus.io/disktype: nvme

* Start here:
  + Send Wayne an email - [ws@csun.edu](mailto:ws@csun.edu)

**Additional Resources**

* Sometimes, researchers just need an unmanaged or managed (by students, supervised by faculty) resource to host public-facing files and applications
  + [Oregon State University Open Source Lab (OSL)](https://osuosl.org/)
* Recently, CSUN was added to the Cloudbank/2i2C JupyterHub resource (this complements CSUN Apporto and SDSU/CSUSB TIDE)
  + [Cal-ICOR JupyterHub Pilot](https://csun.cloudbank.2i2c.cloud/)
* Increasingly, Libretexts is moving beyond "texts" and becoming a complete LMS solution, including a JupyterHub resource
  + General System - [LibreTexts](https://libretexts.org/)
  + Specific Application - [JupyterHub](https://jupyter.libretexts.org/hub/login)
* Some researchers want to experiment with real Quantum resources
  + [D-Wave LEAP Quantum Launchpad/D-Wave Learn Program (D-Wave)](https://www.dwavequantum.com/learn/training/)
* Many researchers require an AI system that *is* open, transparent, and reproducible (built *top-down*)
  + [NSF National Artificial Intelligence Research Resource Pilot (NAIIR)](https://nairrpilot.org/)
* Some researchers desire an AI system that *is* open, transparent, and reproducible (built *bottom-up*)
  + [Non-Profit Personal AI Lab (Kwaai)](https://www.kwaai.ai/)

**Conferences/Fellowships**

* There are plenty of zero-cost and low-cost U.S. domestic events for learning about HPC resources at the *Application*-level.

| **Name** | **Venue** | **Cost** | **Timeframe** |
| --- | --- | --- | --- |
| [Practice & Experience in Advanced Research Computing (PEARC)](https://pearc.acm.org/) | varies | mid $ | late July |
| [Science Gateways (SGX3)](https://sciencegateways.org/gateways2025) | varies | $0 (NSF) | varies |
| [Confab (DOE)](https://confab25.es.net/) | varies | low $ | early April |
| [Institute for Mathematical and Statistical Innovation (IMSI)](https://www.imsi.institute/) | varies | $0 (NSF) | varies |
| [US-RSE Conference (US-RSE)](https://us-rse.org/usrse25/) | varies | $0 (Sloan) | early October |
| [IEEE e-science](https://www.escience-conference.org/2025/) | varies | mid $ | mid September |

* There are plenty of zero-cost and low-cost U.S. domestic events for learning about HPC resources at the *Infrastructure*-level.

| **Name** | **Venue** | **Cost** | **Timeframe** |
| --- | --- | --- | --- |
| [Research Computing at Smaller Institutions (RCSI)](https://rcsi.swarthmore.edu/) | Swarthmore, PA | $0 (NSF) | early June |
| [National Research Platform (NRP)](https://portal.nrp.ai/6nrp-workshop/) | UCSD, CA | $600 | late January |
| [Supercomputing (SC)](https://sc25.supercomputing.org/) | St. Louis, Denver, Atlanta | low $ | mid November |
| [Corporation for Networking and Research (CENIC)](https://cenic.org/events) | varies | low $ | late March |
| [Southern California Linux Expo (SCaLE)](https://www.socallinuxexpo.org/scale/22x) | Pasadena, CA | low $ | early March |

* And the list of *International* events for learning about HPR resources is growing quickly.

| **Name** | **Venue** | **Cost** | **Timeframe** |
| --- | --- | --- | --- |
| [CINI HPC Summer School (CINI)](https://www.hpcsummerschool.it/) | Naples, Italy | N/A | mid June |
| [International Conference on Scalable Scientific Data Management (SSDBM)](https://ssdbm.org/2025/) | varies | N/A | late June |

* The following are some of the Fellowships available:
  + [ICICLE: Intelligent CI with Computational Learning in the Environment (ICICLE)](https://icicle.osu.edu/education-and-outreach/icicle-educational-fellows-program)

**National Workshops**

* There are plenty of *in-person* events for learning about HPR resources.

| **Name** | **Venue** | **Cost** | **Timeframe** |
| --- | --- | --- | --- |
| [Minority-Serving Cyberinfrastructure Capabilities (MS-CC)](https://ms-cc.org/) | varies | $0 (NSF) | late May |
| [Open Science Grig (OSG)](https://osg-htc.org/school-2025/) | U of Wisconsin-Madion, WI | $0 (NSF) | late June |
| [HPC and Data Science Summer Institute (SDSC)](https://na.eventscloud.com/website/83760/) | UCSD, CA | $350 | late July - early August |
| [NERSC International HPC Summer School (NERSC)](https://www.nersc.gov/users/training/events/2025/international-hpc-summer-school-july-2025/) | varies | $0 (DOE) | early July |
| [KNIT (FABRIC)](https://knit.fabric-testbed.net/) | varies | $0 (NSF) | mid March |
| [INTERSECT Research Software training (INTERSECT)](https://intersect-training.org/index.html) | Princeton/Alabama | $0 (NSF) | mid July |
| [SHINE Workshop (SHINE)](https://helioshine.org/the-developing-heliophysics-standards-and-cross-science-collaborations-workshop/) | Iowa City | $0 fees | mid August |

* There are plenty of *virtual* events for learning about HPR resources.

| **Name** | **Venue** | **Cost** | **Timeframe** |
| --- | --- | --- | --- |
| [OU Supercomputing Center for Education & Research (OU)](https://www.oscer.ou.edu/virtualresidency2024.php) | virtual, synchronous | $0 (NSF) | late June |
| [HPC Pathways (NCSA)](https://www.hpc-training.org/moodle/enrol/index.php?id=101) | virtual, asynchronous | $0 | on-going |
| [Cornell Roadmaps](https://cvw.cac.cornell.edu/roadmaps) | virtual | $0 | asynchronous, on-going |
| [HPC Carpentry](https://www.hpc-carpentry.org/) | in-person and virtual, synchronous | $0 | varies |
| [Ecosystem for Research Networking (ERN) Summit](https://docs.google.com/forms/d/e/1FAIpQLSfK9U0vi7IuUETzW3EsRpsC7GfwKdNlkNoa1XuK6Q5NhtbQaA/viewform) | virtual, synchronous only | $0 | late April |
| [Oklahoma University ACI-REF Virtual Residency (Henry Neeman)](https://shareok.org/collections/f2d21d11-57f4-461d-b69b-ed80c93c632b) | virtual | $0 | recurring |
| [(comprehensive, searchable list of resources](https://campuschampions.cyberinfrastructure.org/knowledge-base/resources) | N/A | $0 | varies |

**Upskilling - Professional Associations/Societies**

* Faculty - These HPC resources should be of use to *Faculty* over time.
  + [R OpenSci (ROpenSci)](https://ropensci.org/)
  + [PyOpenSci (pyOpenSci)](https://www.pyopensci.org/)
  + [JuliaCon (annual Summer conference abstracts, proceedings)](https://juliacon.org/2025/)
  + [Framework for Open and Reproducible Research Training (FORRT)](https://forrt.org/)
  + [Open Accelerated Computing (OpenACC) (C/C++ optimizations for research, annual Summer conference)](https://www.openacc.org/)
  + [NumFOCUS (NumFOCUS) (open resource software practices](https://numfocus.org/)
  + [Consortium for the Advancement of Scientific Software (CASS) (DoE sponsored)](https://cass.community/)
  + [Center for Open-Source Research Software Stewardship and Advancement (CORSA)](https://corsa.center/)
  + [LF AI & Data (AI Innovation)](https://lfaidata.foundation/)
  + [US Research Software Sustainability Institute (URSSI)](https://urssi.us/)
  + [Open Molecular Software Foundation (OMSF)](https://omsf.io/)
  + [ZENODO (open data/scholarship/publication repository - managed by CERN and OpenAIRE)](https://zenodo.org/)
  + [Software Carpentries (software engineering)](https://software-carpentry.org/)
  + (and check your discipline's pre-conference workshops and related conference themes for HPC events)
* Staff - These HPC resources should be of use to *Staff* over time.
  + [US Research Software Engineering Association (US-RSE)](https://us-rse.org/)
  + [Campus Research Computing Consortium (CaRCC)](https://carcc.org/)
  + [Campus Champions](https://campuschampions.cyberinfrastructure.org/)
  + [OpenOnDemand](https://openondemand.org/)
  + [Internet2 Research Engagement](https://internet2.edu/community/research-engagement/research-community/)
  + [Internet2 NET+](https://internet2.edu/cloud/internet2-net-plus-services/)
  + [EduCause Research Computing and Data Community Group](https://www.educause.edu/community/research-computing-and-data-community-group)
* Administration - These HPC resources should be of use to *Administration* over time.
  + [Coalition for Academic Scientific Computing (CASC)](https://casc.org/)
  + [Research Software Alliance (ReSA)](https://www.researchsoft.org/)
* Sundry - These HPC resources related to *networking* should be of use to various individuals over time.
  + [ES NET (DOE)](https://www.es.net/)
  + [Globus](https://www.globus.org/)
  + [The Quilt](https://www.thequilt.net/)
  + [Fabric](https://portal.fabric-testbed.net/about/about-fabric)
* Sundry - These miscellaneous open source research-related resources should be of use to various individuals over time.
  + [Research Data Alliance (RDA)](https://www.rd-alliance.org/)
  + [Center for Open Science - Open Software Foundation (OSF)](https://osf.io/)
  + [Digital Management Plan Tool (DMPTOOL)](https://dmptool.org/)
  + [UC Open Source Program Offices)](https://www.socallinuxexpo.org/scale/22x/presentations/building-network-open-source-program-offices-university-california)
  + [Professional Development for Instructors Interested in Student Participation in Humanitarian Free and Open Source Software (POSSE)](https://teachingopensource.org/POSSE)
  + [HPC Social (HPC community development)](https://hpc.social/)
* Sundry - These resources related to *domestic* open science/reproducibility should be of use to various individuals over time.
  + [Science Philanthropy Alliance](https://sciencephilanthropyalliance.org/)
  + [Crossref (Crossref](https://www.crossref.org/)
* Sundry - These resources related to *international* open science/reproducibility should be of use to various individuals over time.
  + [Coalition for Advancing Research Assessment (CoARA)](https://coara.eu/)
  + [UNESCO Open Science (UNESCO)](https://www.unesco.org/en/open-science)
  + [Global Research Council (GRC)](https://globalresearchcouncil.org/)
  + [Society of Software Research Engineering (SocRSE) (provides £1,000 for conference travel too)](https://society-rse.org/)
  + [Science for Life Laboratory (SciLifeLab)](https://www.scilifelab.se/)
  + [Turing Way (Turing Way)](https://book.the-turing-way.org/)
  + [Digital Research Alliance of Canada](https://www.alliancecan.ca/en)
  + [International Conference on Scalable Scientific Data Management](https://ssdbm.org/2025/)
  + [Reach the World (K-12 supercomputing)](https://reachtheworld.org/supercomputing-and-future-ai/journal/coding-high-school-building-movement?page=2)
* Sundry - These charitable organizations provide regular funding for HPC-related and scientific software.
  + [Chan/Zuckberberg Initiative (Essential Open Source Software for Science (Cycle 6))](https://chanzuckerberg.com/rfa/essential-open-source-software-for-science/)
  + [Kavli Foundation (general science)](https://www.kavlifoundation.org/funding-opportunities)
  + [Wellcome Trust (generally, health-related research software](https://wellcome.org/grant-funding)
  + [Simons Foundation (generally, math and physical sciences software)](https://wellcome.org/grant-funding)
  + [Sloan Foundation (generally, emerging technology of any type)](https://sloan.org/programs/digital-technology)
  + [Code for Science and Society (generally, open source software)](https://www.codeforsociety.org/projects)
* Sundry - Parody.
  + [Journal of Astrological Big Data Ecology (like the old "Journal of Irreproducible Results")](https://jabde.com/)