

INFORMATION SERVICES

UC San Diego Health Sciences

Expanded Technical Services Handbook

***Please note that this handbook is geared toward individuals who have an advanced understanding of technical terms, technologies and services.*

Table of Contents

1. Information Services Tech Offerings

2. Research IT

3. High Performance Computing (HPC)

4. Advanced Network Troubleshooting

5. Advanced Storage

6. ACTRI

Information Services Tech Offerings

Please review our ITSS page [here](#). The link offers details regarding IS' technical services.

Please make specific requests [here](#) (login required; username@ucsd.edu format; Ensure you have DUO 2-Factor ready if you are not on the Health Sciences VPN).

Advanced IS Technology Requests

Servers

IS offers server virtualization including:

- **Windows Servers**
 - 2016
 - 2019
- **Linux Servers**
 - Red Hat Enterprise
 - 7.9 – 8.6
- **VMware Open Virtualization Application/Appliance (OVA)**
 - Infrastructure managed by IS
 - OS, patching, updates, vulnerabilities managed by app owners/vendors
- **VMware OVF Open Virtualization Format**
 - platform-independent, efficient, extensible, and open packaging and distribution format for virtual machines
 - Infrastructure managed by IS
 - OS, patching, updates, vulnerabilities managed by app owners/vendors
 - Information about [OVA/OVF](#)

Hint: When requesting a virtual machine, ensure you include the minimum tech specs required for your project (e.g. HD size, RAM, OS, who will need access etc).

To request this service, login to ServiceNow:

→ Request Something → Network & Server Requests → Technology Request

AWS Serverless Technologies

AWS offers technologies for running code, managing data, and integrating applications, all without managing servers. Serverless technologies feature automatic scaling, built-in high availability, and a pay-for-use billing model to

increase agility and optimize costs. These technologies also eliminate infrastructure management tasks like capacity provisioning and patching, so you can focus on writing code that serves your customers. Serverless applications start with AWS Lambda, an event-driven compute service natively integrated with over 200 AWS services and software as a service (SaaS) applications.

AWS Serverless Services

Compute Services include:

- AWS Lambda
- AWS Fargate

Application Integration:

- Amazon EventBridge
- AWS Step Functions
- Amazon SQS
- Amazon SNS
- Amazon API Gateway
- AWS AppSync

Data Store:

- Amazon S3
- Amazon DynamoDB
- Amazon RDS Proxy
- Amazon Aurora Serverless

Visit the Amazon [site](#) for technical details.

Storage

The storage strategies listed below are in addition to the “Data Backup/Storage Plans” already mentioned in the IT Handbook.

IS Storage

AWS

AWS Cloud offers both short-term and long-term storage strategies.

Amazon S3 Glacier

Amazon S3 Glacier (S3 Glacier) is a secure and durable service for low-cost data archiving and long-term backup.

With S3 Glacier, you can store your data cost effectively for months, years, or even decades. S3 Glacier helps you offload the administrative burdens of operating and scaling storage to AWS, so you don't have to worry about capacity planning, hardware provisioning, data replication, hardware failure detection and recovery, or time-consuming hardware migrations. For more information regarding any AWS service, submit a general request on the [IS Portal](#) (login required).

Amazon Simple Storage Service (Amazon S3) also provides three Amazon S3 Glacier archive storage classes. These storage classes are designed for different access patterns and storage duration. These storage classes differ as follows:

- **S3 Glacier Instant Retrieval** - Use for archiving data that is rarely accessed and requires milliseconds retrieval.
- **S3 Glacier Flexible Retrieval** (formerly the S3 Glacier storage class) - Use for archives where portions of the data might need to be retrieved in minutes. Data stored in the S3 Glacier Flexible Retrieval storage class can be accessed in as little as 1-5 minutes by using Expedited retrieval. You can also request free Bulk retrievals in up to 5-12 hours.
- **S3 Glacier Deep Archive** - Use for archiving data that rarely needs to be accessed. Data stored in the S3 Glacier Deep Archive storage class has a default retrieval time of 12 hours.

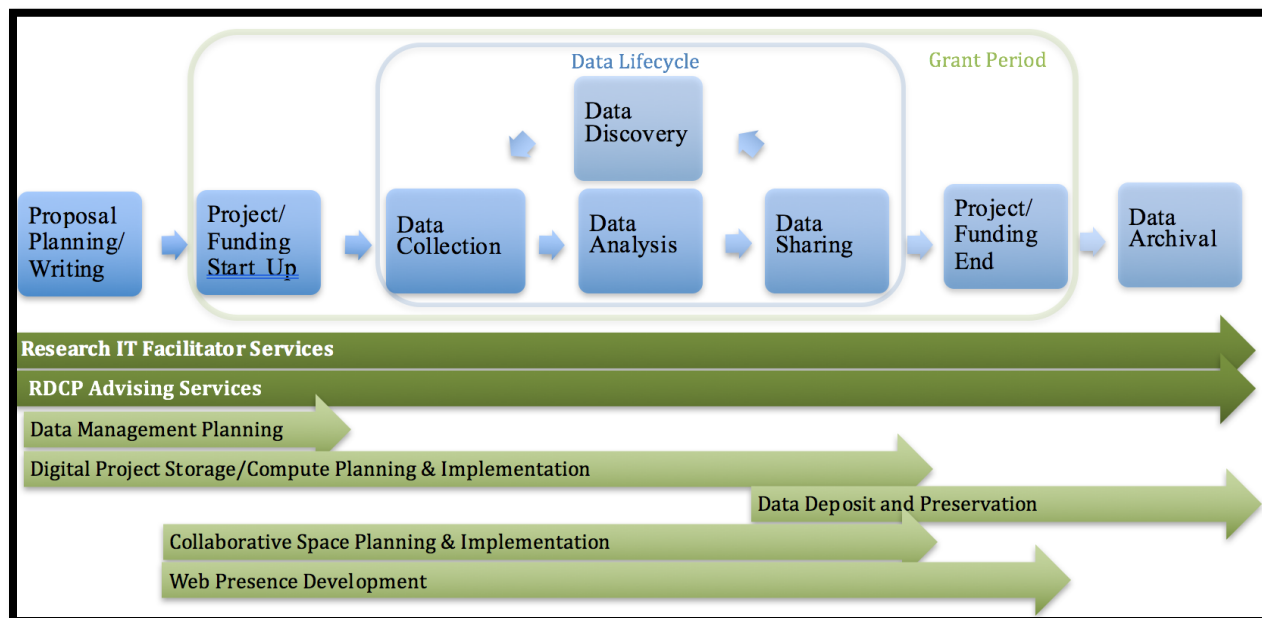
Amazon DynamoDB

DynamoDB is a key-value and document database service, delivering single-digit millisecond performance at any scale.

For more information regarding the above and additional storage services, please visit AWS [here](#).

Research IT

Research IT is a Campus team offering free Research Computing and Data (RCD) support and services to faculty, research staff, postdoctoral, graduate, and undergraduate students throughout all academic departments, schools, and divisions. A non-exhaustive list of our services include direct engagement and resource recommendations to faculty, assistance in integrating new technologies in the lab, development of demos or proof-of-concepts for researchers, and consulting to provide solutions toward advanced cyberinfrastructure and cloud service needs.



Research IT facilitator services

Research facilitation & digital technology advising helps to guide you to services that meet your research needs. Contact Research IT at research-it@ucsd.edu for personalized recommendations. Facilitation provided at no cost to the researcher.

Technical consultations & demo/ proof-of-concept services

Need to solve a difficult technical issue for your project? Need support to put together a demo for a funding agency? Need help assembling a proof-of-concept to strengthen your grant proposal? Research IT can help! Our system integration engineers offer a limited amount of project support free to UC San Diego faculty.

For general information about facilitation & consulting services, email research-it@ucsd.edu.

High-Performance Computing (HPC)

Offered through San Diego Supercomputer Center (SDSC), HPC systems offer cutting-edge high-performance computing. Whether you're looking to expand computing beyond your lab or a business looking for that competitive advantage, SDSC's HPC experts will guide potential users in selecting the right resource, thereby reducing time to solution while taking your science to the next level.

HPC Expanse

In the Fall of 2020, SDSC launched its newest National Science Foundation (NSF)-funded supercomputer, *Expanse*. At over twice the performance of SDSC's *Comet* system, *Expanse* supports SDSC's vision of 'Computing without Boundaries' by increasing the capacity and performance for thousands of users of batch-oriented and science gateway computing, and by providing new capabilities that will enable research increasingly dependent upon heterogeneous and distributed resources composed into integrated and highly usable cyberinfrastructure.

System	Performance	Key Features
Expanse Expanse Home User Guide	5 Pflop/s peak; 93,184 CPU cores; 208 NVIDIA GPUs; 220 TB total DRAM; 810 TB total NVMe	Standard Compute Nodes (728 total) AMD EPYC 7742 (Rome) Compute Nodes; 2.25 GHz; 128 cores per node; 1 TB NVMe per node; 256 GB DRAM per node GPU Nodes (52 total) NVIDIA V100s SMX2 with 4 GPUs per node; 40 6248 Xeon CPU cores per node; 384 GB CPU DRAM per node; 2.5 GHz CPU clock speed; 32 GB memory per GPU; 1.6 TB NVMe per node; connected via NVLINK Large-memory Nodes (4 total) AMD Rome nodes; 2 TB DRAM per node; 3.2 TB SSD memory per node; 128 cores per node; 2.25 GHz Interconnect HDR InfiniBand, Hybrid Fat-Tree topology; 100 Gb/s (bidirectional) link bandwidth; 1.17-x.xx μ s MPI latency Storage Systems Access to Lustre (12 PB) and Ceph (7 PB) storage SDSC Scalable Compute Units (13 total) Entire system organized as 13 complete SSCUs, consisting of 56 standard nodes and four GPU nodes connected with 100 GB/s HDR InfiniBand

For an EXPANSE consultation click [here](#).

HPC Comet

Comet is a petascale supercomputer designed to transform advanced scientific computing by expanding access and capacity among traditional as well as non-traditional research domains. The result of a \$21.6 million National Science Foundation award, *Comet* is capable of an overall peak performance of 2.76 petaflops, or 2.76 quadrillion operations per second.

System	Performance	Key Features
Comet Summary User Guide	2.76 Pflop/s peak; 48,784 CPU cores; 288 NVIDIA GPUs; 247 TB total memory; 634 TB total flash memory	<p>Standard Compute Nodes (1944 total) Intel Xeon E5-2680v3 2.5 GHz dual socket, 12 cores/socket; 320 GB flash memory; 120 GB/s memory bandwidth</p> <p>GPU Nodes (72 total) 36 K80 nodes: 2 NVIDIA K80 GPUs per node; dual socket, 12 cores/socket; 128 GB DDR4 DRAM; 120GB/s memory bandwidth; 320 GB flash memory 36 P100 nodes: 4 NVIDIA P100 GPUs; dual socket, 14 cores/socket; 128 GB DDR4 DRAM; 150GB/s memory bandwidth; 400 GB flash memory</p> <p>Large-memory Nodes (4 total) 1.5 TB total memory; 4 sockets, 16 cores/socket; 2.2 GHz</p> <p>Interconnect Hybrid Fat-Tree topology; 56 Gb/s (bidirectional) link bandwidth; 1.03-1.97 μs MPI latency</p> <p>7.6 PB Lustre-based Parallel File System Access to Data Oasis</p> <p>High-performance virtualization</p>

For a COMET consultation, click [here](#).

HPC Triton Shared Computing Cluster (TSCC)

In mid-2013 SDSC launched the *Triton Shared Computing Cluster (TSCC)* after recognizing that UC San Diego investigators could benefit from an HPC system dedicated to their needs and with near-immediate access and reasonable wait times instead of accessing a national system entailing competitive proposals and often longer wait times. Following an extensive study of successful research computing programs across the country, SDSC selected the “condo computing” model as the main business model for *TSCC*. Condo computing is a shared ownership model in which researchers use equipment purchase funds from grants or other sources to purchase and contribute compute “nodes” (servers) to the system. The result is a researcher-owned computing resource of medium to large proportions.

Important information regarding TSCC Computing “Condo” can be found [here](#). You’ll find details including condo plan summary, usage model, hotel vs condo and more.

System	Performance	Key Features
<p>TSCC: Triton Shared Computing Cluster</p> <p>TSCC Home</p> <p>User Guide</p> <p>Summary</p>	80+ Tflop/s (and growing!)	<p>General Computing Nodes Dual-socket, 12-core, 2.5GHz Intel Xeon E5-2680 (coming) and Dual-socket, 8-core, 2.6GHz Intel Xeon E5-2670</p> <p>GPU Nodes Host Processors: Dual-socket, 6-core, 2.6GHz Intel Xeon E5-2630v2 GPUs: 4 NVIDIA GeForce GTX 980</p> <p>Interconnect 10GbE (QDR InfiniBand optional)</p> <p>Lustre-based Parallel File System Access to Data Oasis</p>

For a free trial, email tscc-info@ucsd.edu and provide your:

- Name
- Department
- UC San Diego affiliation (grad student, post-doc, faculty etc)

Trial accounts are 250 core-hours valid for 90 days.

UCSD Research Cluster

In collaboration with the Data Science and Machine Learning Platform (DSMLP) which is provided for instructional use, Research IT offers the UCSD Research Cluster to provide faculty and other researchers CPU/GPU compute resources for projects that are non-instructional and data intense. Contact research-it@ucsd.edu to set up a consultation.

Advanced Network Troubleshooting

UCSD's Research Network Extensions

UCSD is home to two NSF-supported research network extensions:

- Prism@UCSD, a science DMZ providing 40G or greater cross-campus connections to big data projects,
- and to CHERuB (maintained jointly by ACT and SDSC), which gives UCSD an additional 100G of configurable connectivity for experimental research.

UCSD High-Performance Research Partners

In addition to our Internet connectivity, UCSD connects directly at high speed to many nearby research institutions. You can [download a PDF of the current connections to external partners](#).

Troubleshooting

What's slowing my connection?

- **Check your area:** campus network connections are rarely impacted (connect to ITS' [UCSD Network Weathermap](#) and log in with your Active Directory username/password to see current conditions).
- Your wall jack, provided by the NGN (Next Generation Network) program at UCSD, defaults to a **1Gbps standard**, which matches most current Ethernet adapters (NICs).
- In many cases, tuning your kernel or TCP/IP stack will improve your performance more than a network upgrade.
 - Check out a UCSD SysWiki article on [tuning your TCP/IP stack](#) [Note: this link can only be accessed from on campus or via the VPN]
 - ESnet offers an extensive site on performance tuning.
- [Consult with ITS](#), who can test your connection path and find the source of the problem you are experiencing.

I've identified a bottleneck

If your bottleneck is your *subnet* or *building network*, and/or if you're planning on purchasing 10Gbps or greater NICs, you have several options to improve your connection. Both require purchase of a network switch and may have cabling or other one-time costs:

- Upgrade to a [10Gbps production network](#) connection (40G available in some areas)

This NGN-maintained connection will be yours alone until it reaches the building edge, where it will share a 10Gbps path across campus. (Regular

campus backbone traffic rarely exceeds 1Gbps.) There is no special configuration required to use these connections.

- Upgrade to a 40Gbps Prism@UCSD research network connection, where available

This NSF-supported connectivity allows 40Gbps connections that are yours alone all the way to the Prism@UCSD switch, where 120Gbps in connectivity provides access to SDSC and 2x40Gbps connections into the CHERuB 100Gbps switch. If you need connectivity via CHERuB to the Internet, you will need to know your Internet destination so that a pathway can be negotiated and set up for your use. Subject to fiber availability. Note that Prism@UCSD is not supported by IT Services production network staff.

IPv6 is available on both types of connection where needed.

Contact research-it@ucsd.edu to discuss which is right for your needs.

ACTRI

[ACTRI](#) is located in the Altman Clinical and Translational Research Institute Building at UC San Diego near Thornton Pavilion in La Jolla. This seven-story facility serves as a bridge between UC San Diego's clinical and research enterprises, and houses ACTRI's administrative offices, with approximately 50 staff members.

The ACTRI facilitates the research of others by providing assistance with proposal development, study design, regulatory issues and submissions, biostatistics and informatics, and offers ethics consultation, clinical trial coordination, genomics technologies, biorepository services, facilitation with an affiliated laboratory network, and training for medical students and early career investigators. Nearly 200 clinical trials are run through ACTRI every year. The ACTRI does not conduct its own independent research; we exist to facilitate the research needs of our faculty and investigators.

ACTRI has [several interactive functional areas](#) that provide resources, services, and guidance to assist researchers with their projects:

1. Center for Clinical Research (CCR)
2. Translational & Workforce Development (TWD)
3. Biomedical Informatics (BMI)
4. Translational Research Partnership (TSP)
5. Translational Technology Center (TTC)
6. Community Engagement (CE)
7. Center for Excellence in Immunogenomics (CETI)
8. Device Acceleration Center (DAC)
9. Dissemination & Implementation Science Center (DISC)
10. Center for Life Course and Vulnerable Population Research (CLVR)
11. Evaluation

Biomedical Informatics

Informatics

The ACTRI supports software for managing clinical trials and observational studies, as well as tools to query clinical records.

Biostatistics

The Biostatistics faculty and staff advise and collaborate with ACTRI researchers on biostatistics, bioinformatics, and epidemiology. In addition, to common statistical tools, the faculty and staff have expertise in areas including

reproductive research reporting, R Bioconductor, and high-performance computing. Visit [Biostatistics](#) for more info.

Center for Computational Biology and Bioinformatics

The Center for Computational Biology and Bioinformatics (CCBB) provides analysis in genomics technologies including Next-Generation Sequencing (NGS) and the interpretation of genomics data. We offer analysis and consultation on epigenomics as well. Our long-term plans include expansion of our service to proteomics and metabolomics. For details, visit [CCBB](#).

CCBB Requests can be submitted [here](#).

CCBB Analytical Services and Rates can be found [here](#).

General CTRI Service Requests, submit [here](#).

To request Clinical Research Services (Clinic, Coordinators, Regulatory Support, Pharmacy, MRI, Data Safety Monitoring Board), please use the new [Clinical Research Service Request Form](#)

To request Center for Computational Biology and Bioinformatics (CCBB) Services, please use the new [CCBB Service Request Form](#)

Velos/REDCap Information

Velos eResearch is an integrated software system designed to assist Investigators in managing their clinical research trials/studies. The software links to the UCSD Health's Epic Electronic Medical Record System to provide improved information and integration for clinical research projects. A robust support team assists investigators in implementing their protocols, study calendars, coverage analysis and more. Velos eResearch is a web-based system that offers the following functionality:

- **Study Management**
 - Creation of protocols
 - Creation of detailed study calendars based on protocol guidelines
 - Creation of coverage analysis
 - Creation of study budgets
 - Tracking of various study statuses
- **Patient Management**
 - Ability to register patient in the Velos Database
 - Track visits / events as they occur
 - Electronic data capture
 - Ability to track Adverse Events
- **eSample Biospecimen Tracking**
 - Creation of specimen record
 - Creation of child samples

- Ability to build “virtual” storages to track where specimens are stored
- Ability to track multiple specimen statuses
- Ability to capture additional data elements on a custom form
- **Electronic Data Capture**
 - Custom form creation
 - Data query capability
- **Financial Management**
 - Creation of billing milestones
 - Creation of invoices
 - Tracking of payments received
- **Reporting**
 - Report Central -standard and custom reposts available
 - Ad-Hoc Query Reporting- ability to query database and export data
 - CTRP Reporting-Clinical Trial Reporting Program for NIH

[Velos Wiki](#) - The Velos Wiki is a web-based support tool, allowing users access to training information, including videos and a manual, as well as the ability to post questions to the UC San Diego Velos community and to receive updates about Velos. In addition, new users can download the User Access Request Form and review the upcoming training schedule.

Velos Access and Training Information

Click [here](#) to request access.

REDCap

REDCap is a secure web application for building and managing online surveys and databases. REDCap is available to all UC San Diego faculty and staff, and to users outside the organization who have sponsorship from UC San Diego faculty.

Click [here](#) for further details to learn more.

Additional ACTRI Services

ACTRI provides a range of services and applications to assist your research:

- [UCSD COVID-19 Registries](#)
- [Virtual Research Desktop \(VRD\)](#)
- [Data Extraction Concierge Service \(DECS\)](#)
- [Accrual for Clinical Trials \(ACT\)](#)
- [Tableau](#)

ACTRI IT Support

IT Support/Help Desk

ACTRI IT Support: ctri-support@ucsd.edu

ACT help desk (creation / deletion of new email accounts):
acthelp@ucsd.edu | (858) 543-4357

Telecom (phone issues, new phone, activate wall jack):
(858) 534-2930

Facilities:
(858) 534-2930

Mailing Address:
Help IT Center
9500 Gilman Drive MC 0990
La Jolla, CA 92093

Physical Address:
ACTRI
9452 Medical Center Drive
La Jolla, CA 92037

Please visit the ACTRI Website [here](#) to see details on the above & more!