Overview of the Computer Resource Team (CRT)

Blaise Barney (LLNL)
Rob Cunningham (LANL)
Karen Haskell (Sandia)

PSAAP II Kickoff Meeting
December 9-10, 2013, Albuquerque, NM

LLNL-PRES-405061

Overview

- What is the CRT?
- What the CRT can do for you
- HPC customer support and problem tracking
- Training
- Documentation
- Communications
- Dedicated Runs
- Site visits
- Allocations
- LLNL compute resources
- LANL compute resources
- Sandia compute resources
- Getting accounts - SARAPE

What Is The CRT?

- The **Computer Resource Team (CRT)** is the component of the PSAAP II program that connects center researchers with the High Performance Computing (HPC) resources required to perform their work.

- The CRT is comprised of representatives from each NNSA Lab who are familiar with their Lab's computing resources, personnel and policies. The following individuals serve on the CRT:
  - Blaise Barney, LLNL *blaiseb@llnl.gov*
  - Rob Cunningham, LANL *rtc@lanl.gov*
  - Karen Haskell SNL *khaskel@sandia.gov*

- ***Our primary purpose is to provide assistance and guidance in all aspects related to the use of HPC resources located at LANL, LLNL, and Sandia.***

What the CRT Can Do For You?

- Assist with the establishment and use of computer accounts
- Assist with accessing compute resources
- Provide essential HPC user documentation
- Provide technical support and referral to in-depth consulting
- Conduct monthly telecons to keep PSAAP users up-to-date with account, access, policy, scheduling and technical issues, and to address issues with HPC platform usage
- Interface with other individuals and groups within the Labs, such as management, networking, system administration, storage, customer support, etc., to facilitate the effective support of PSAAP users
- Track and facilitate the resolution of problems reported to each Labs' customer support “hotline”
- Provide training opportunities

What the CRT Can Do For You? (cont)

- Collect and distribute monthly machine usage statistics
- Schedule and support special/dedicated runs
- Maintain a balance of machine usage between the PSAAP centers (if needed)
- Visit PSAAP centers to discuss HPC resources, user issues and to offer technical consultation and/or training
- Showcase PSAAP research in the NNSA/ASC or DOE research exhibit booth at the annual SC conference

National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

HPC Customer Support and Problem Tracking

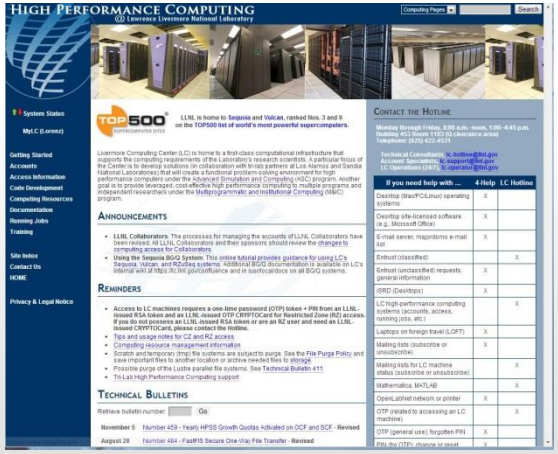
- All three labs offer HPC customer support via phone and email:
  - LLNL: Livermore Computing Hotline
  - LANL: ICN Consulting Office
  - Sandia: HPC OneStop Service Desk
- Includes support for user accounting issues and for technical assistance.
- Problems and questions are tracked via a customer support database application (varies with each Lab).
- Most problems/questions are handled directly by the customer support staff on duty.
- More in-depth issues are typically referred to local subject experts.
- The labs also coordinate with hardware and software vendors for issues that require outside support.
- CRT reps coordinate routinely with each other on Tri-lab user issues.

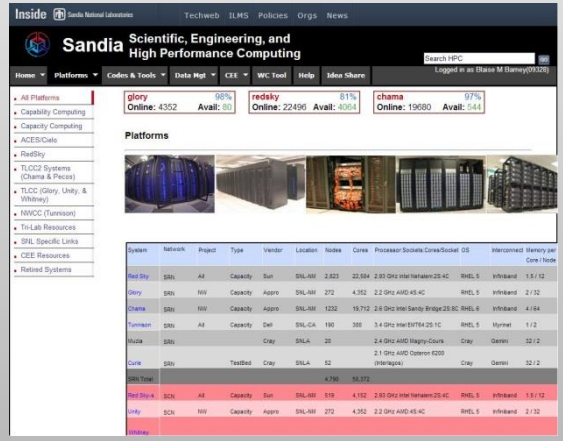
HPC Training

- Training is important – especially for new users
  - Online tutorials are available
  - Workshops conducted at the Labs are open to PSAAP center users
  - The CRT can deliver workshops/training at your center
  - The CRT can also assist with topic specific, customized workshops if that is of interest
- Topics include:
  - Getting started information
  - Compilers
  - Performance tools
  - Debuggers
  - Parallel programming (MPI, OpenMP, Pthreads…)
  - Batch schedulers
  - Machine architectures
  - Visualization tools
  - …

User Documentation

- Most of what users need to know is available online via web pages hosted by each of the Labs. Recommended starting points:
  - LLNL
    - computing.llnl.gov
    - computing.llnl.gov/tutorials/
  - LANL
    - www.lanl.gov/projects/computing/web_hpc.html
    - hpc.lanl.gov
  - Sandia
    - computing.sandia.gov
    - hpc.sandia.gov (open)
  - Note: most LLNL web pages are open – no authentication required. Most Sandia / LANL pages require authentication
- Quick Start Guide for new PSAAP users:
  - asc.llnl.gov/alliances/Alliance.Quickguide.pdf





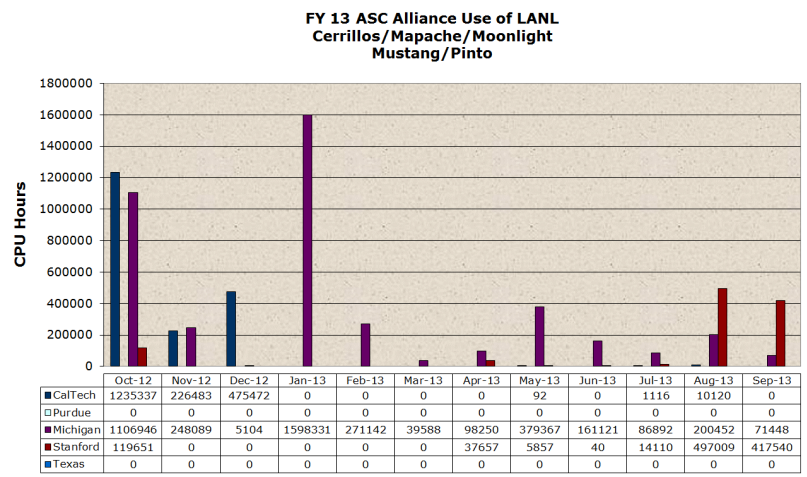
Communications

- Monthly telecons
  - Forum for discussion/questions on user topics such as accounts, access, technical issues, machine schedules, etc.
  - Open to all PSAAP center users
  - Highly recommended to have a Point-of-Contact (POC) person with some technical/computing experience from each Center attend. We'll be in touch with each Center's PI to find out who this might be.
  - First Wednesday of each month
  - Toll-free number hosted by the CRT
    - Dial-in Number: 866-914-3976 Participant Code: 187522#

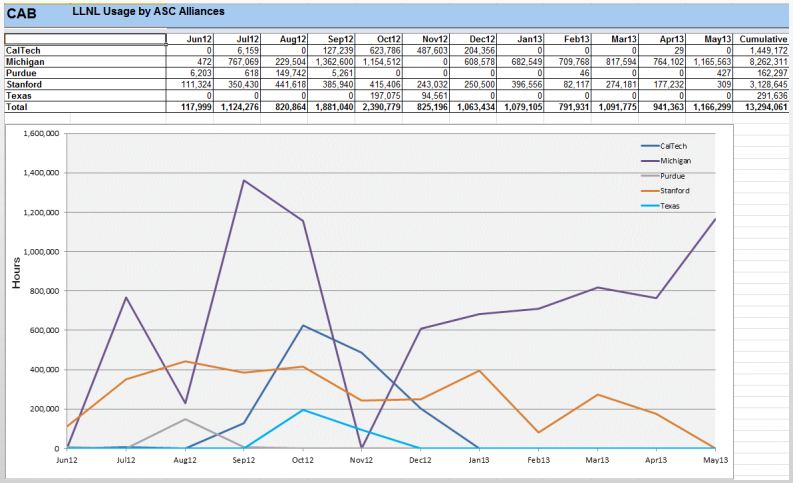
- Email list - asap-crt@lanl.gov
  - Telecon minutes and machine usage statistics are distributed via this list to all PSAAP Center PIs & POCs, ASC HQ and various staff & managers within the Labs
  - Let us know if you want anyone else at your Center added to our list - initially it includes only your PI and POC

Communications

- Usage stats
  - Collected by each Lab's computing center and distributed with the telecon minutes
  - Present both aggregate and detailed usage (down to the user level)



	Oct-12	Nov-12	Dec-12	Jan-13	Feb-13	Mar-13	Apr-13	May-13	Jun-13	Jul-13	Aug-13	Sep-13
CalTech	1235337	226483	475472	0	0	0	0	92	0	1116	10120	0
Purdue	0	0	0	0	0	0	0	0	0	0	0	0
Michigan	1106946	248089	5104	1598331	271142	39588	98250	379367	161121	86892	200452	71448
Stanford	119651	0	0	0	0	0	37657	5857	40	14110	497009	417540
Texas	0	0	0	0	0	0	0	0	0	0	0	0



CAB — LLNL Usage by ASC Alliances

	Jun12	Jul12	Aug12	Sep12	Oct12	Nov12	Dec12	Jan13	Feb13	Mar13	Apr13	May13	Cumulative
CalTech	0	6,159	0	127,239	623,786	487,603	204,356	0	0	0	29	0	1,449,172
Michigan	472	767,069	229,504	1,362,600	1,154,512	0	608,578	682,549	709,768	817,594	764,102	1,165,563	8,262,311
Purdue	6,203	618	149,742	5,261	0	0	0	0	46	0	0	427	162,297
Stanford	111,324	350,430	441,618	385,940	415,406	243,032	250,500	396,556	82,117	274,181	177,232	309	3,128,645
Texas	0	0	0	0	197,075	94,561	0	0	0	0	0	0	291,636
Total	117,999	1,124,276	820,864	1,881,040	2,390,779	825,196	1,063,434	1,079,105	791,931	1,091,775	941,363	1,166,299	13,294,061

- Email & phone
  - The CRT can be contacted directly by you and any of your Center's users:
    - Blaise Barney (LLNL) *blaiseb@llnl.gov* 925-422-2578
    - Rob Cunningham (LANL) *rtc@lanl.gov* 505-665-4444 x05704
    - Karen Haskell (Sandia) *khaskel@sandia.gov* 505-845-2087

National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

Dedicated Runs (DATs)

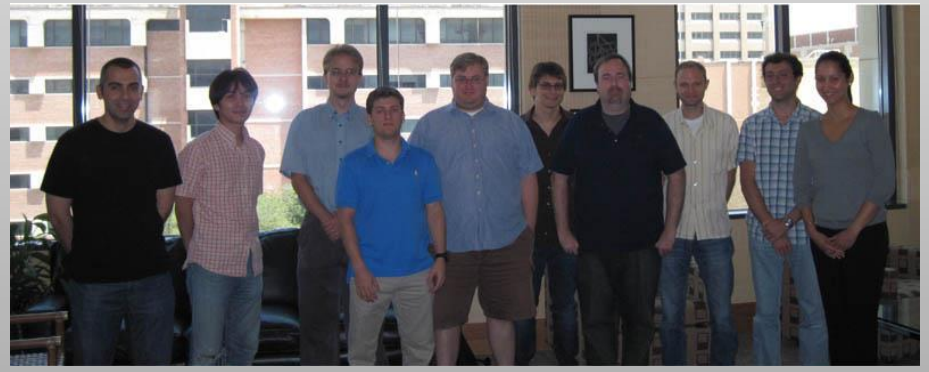
- Normally, users submit their jobs to a batch system using a standard Tri-lab Moab batch scheduler interface:
  - Jobs are then queued to wait their turn for execution.
  - There are limits on the number of nodes and number of hours that a job can use.

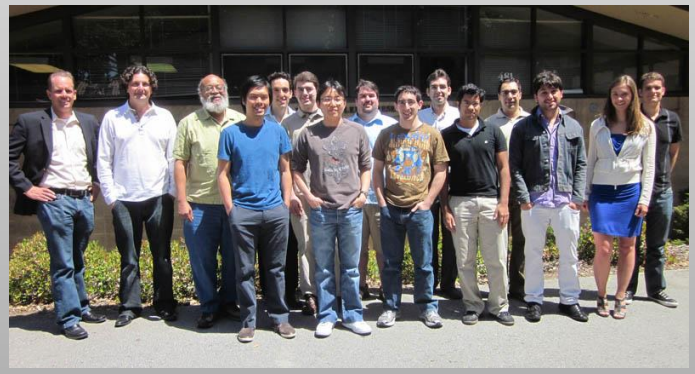
- Dedicated application time - DAT: can be requested by any PSAAP user:
  - Overrides normal node limits - up to the full machine
  - Overrides normal time limits
  - Typically conducted on weekends at LLNL, weekdays at LANL/Sandia

- How to request a DAT:
  - LLNL: computing.llnl.gov/?set=forms&page=ASC_dat_forml
  - LANL: email to consult@lanl.gov
  - Sandia: email to hpc-help@sandia.gov

Site Visits

- The CRT is available to visit PSAAP centers:
  - Visits (2-4 hrs) by the CRT and other Lab staff
  - Focus is on the center's *users* of HPC computing resources
  - Updates on architectures, policies, future plans at the Labs
  - Forum for discussing user issues, problems, questions
  - Not to be confused with TST visits or reviews
- We can also include a technical "training" session if desired
- We'd like to talk with you during the kickoff meeting! To discuss setting up an initial visit - after your users have accounts.





National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

Allocations - Production Machines

- PSAAP centers have allocations on eight ASC funded, Tri-lab production machines:
  - PSAAP allocations vary by machine; shared at the group level
  - Able to use more than allocation if there are available cycles

Machine	Lab	Total available CPU hr/mo	PSAAP allocation	PSAAP total hr/mo	Per Center hr/mo
CAB	LLNL	14,016,000	3.75%	525,600	87,600
CHAMA	Sandia	14,097,760	5.0%	704,888	117,481
GLORY	Sandia	2,990,080	5.0%	149,504	24,917
MAPACHE	LANL	3,457,280	55.7%	1,926,198	321,033
MOONLIGHT	LANL	3,597,440	64.3%	2,312,640	385,440
MUSTANG	LANL	28,032,000	8.6%	2,402,742	400,457
PINTO	LANL	1,798,720	4.3%	77,088	12,848
VULCAN	LLNL	287,047,680	5.0%	14,352,384	2,392,064
TOTAL		**355,036,960**	**10.5%**	**22,451,044**	**3,741,840**

LLNL PSAAP Compute Resources

- **CAB**
  - Intel Xeon E5-2670 (Sandy Bridge-EP) architecture; 2.6 GHz
  - 1,296 nodes with 16 cores each; 20,736 total cores
  - 32 GB memory per node; 41.5 TB total memory
  - QDR QLogic Infiniband interconnect; ~5.0 GB/sec pt2pt MPI
  - 431 Tflops peak
  - Details at computing.llnl.gov/tutorials/linux_clusters

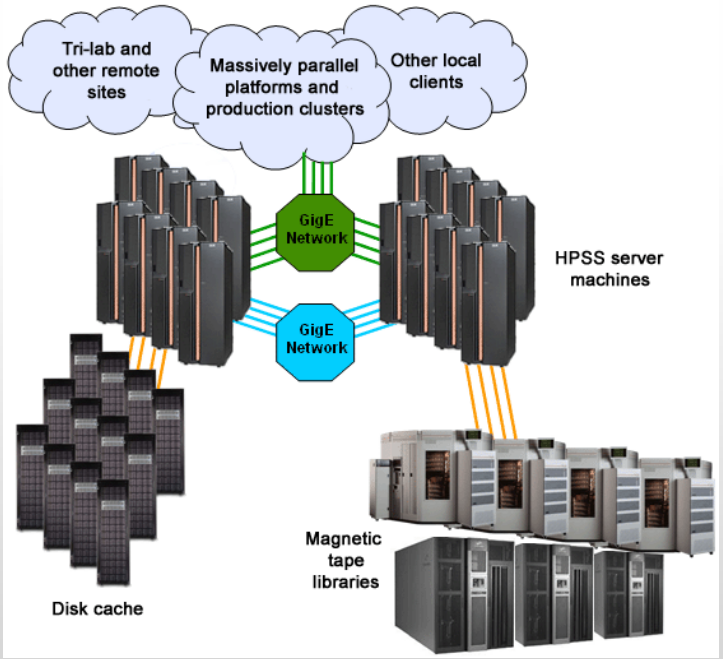
National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

LLNL PSAAP Compute Resources

- **VULCAN**
  - IBM Blue Gene/Q architecture; 1.6 GHz PowerPC A2
  - 24,576 nodes with 16 cores each; 393,216 total cores
  - 4 hardware threads per core; 1.57 million threads total
  - 16 GB memory per node; 393 TB total memory
  - 5D torus interconnect; ~40 GB/sec per node
  - 5 Pflops peak
  - Details at computing.llnl.gov/tutorials/bgq

LLNL PSAAP Compute Resources

- **HPSS Archival Storage**
  - Mounted from all unclassified production clusters
  - Users automatically get an HPSS account with their machine accounts
  - 23 PB in use as of Dec 2013
  - 1.1 PB disk cache



- **Visualization Cluster - EDGE**
  - Intel Xeon EP X5660 @ 2.8 GHz
  - 216 nodes, 12 cores/node
  - 2 NVIDIA Tesla M2050 GPUs/node
  - 48-96 GB memory/node
  - Available upon request

LANL PSAAP Compute Resources

Rob Cunningham

LANL PSAAP Compute Resources

- **MOONLIGHT -** hybrid ASC cluster
  - Intel Xeon E5-2670 (Sandy Bridge-EP) architecture; 2.6 GHz
  - 308 nodes with 16 cores each; 4,928 total cores
  - 32 GB memory per node; 9.9 TB total memory
  - QDR QLogic Infiniband interconnect
  - 2 NVIDIA Tesla M2090 GPGPUs per node
  - 488 Tflops peak

National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

LANL PSAAP Compute Resources

- **MUSTANG** - larger scale
  - AMD Opteron 6176 architecture; 2.3 GHz
  - 1,600 nodes with 24 cores each; 38,400 total cores
  - 64 GB memory per node; 102 TB total memory
  - Mellanox Infiniscale IV 4x-QDR interconnect
  - 353 Tflops peak

National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

LANL PSAAP Compute Resources

- **MAPACHE** - 65% ASC allocation
  - SGI XE1300 architecture; 2.6 GHz
  - 592 nodes with 8 cores each; 4,736 total cores
  - 24 GB memory per node; 14.2 TB total memory
  - Mellanox InfiniBand 4x-QDR
  - 50.4 Tflops peak

LANL PSAAP Compute Resources

- **PINTO** - small scale
  - Intel Xeon E5-2670 (Sandy Bridge-EP) architecture; 2.6 GHz
  - 154 nodes with 16 cores each; 2,464 total cores
  - 32 GB memory per node; 4.9 TB total memory
  - QDR QLogic Infiniband interconnect
  - 51.3 Tflops peak

LANL PSAAP Compute Resources

- **LIGHTSHOW** - visualization cluster
  - Intel Xeon 5650 architecture; 2.7 GHz
  - 16 render nodes with 12 cores each; 192 total cores
  - 96 GB memory per node plus 2 x nVidia Quadro 6000 cards
  - QDR Mellanox Infiniband interconnect
  - 4 Tflops peak + GPUs

National Nuclear Security Administration | ASC | Los Alamos National Laboratory | Sandia National Laboratories | Lawrence Livermore National Laboratory

LANL PSAAP Compute Resources

- **GPFS Archival Storage**
  - Accessible from gpfst1-fe.lanl.gov through gpfst7-fe
  - Mount both the huge scratch spaces and also the tape library
  - http://hpc.lanl.gov/turquoise_archive

- **Fast File Transfer Service**
  - File Transfer Agents, Internal: ftaint01 and ftaint02
  - Utilize scp and parallel bbcp, up to 300 MB/s measured
  - Uses a 2-hop method
  - http://hpc.lanl.gov/turquoise_transfer

Sandia PSAAP Compute Resources

Karen Haskell

Sandia PSAAP Compute Resources



- **CHAMA**
  - Intel Xeon E5-2670 (Sandy Bridge-EP) architecture; 2.6 GHz
  - 1,232 nodes with 16 cores each; 19,712 total cores
  - 64 GB memory per node; 78.9 TB total memory
  - QDR QLogic Infiniband interconnect
  - 410 Tflops peak
  - In SARAPE, request "SRN Capacity Clusters"

Sandia PSAAP Compute Resources

- **GLORY**
  - AMD Opteron architecture; 2.2 GHz
  - 288 nodes with 16 cores each; 4,608 total cores
  - 32 GB memory per node; 9.2 TB total memory
  - DDR Voltaire/Mellanox Infiniband interconnect
  - 40.6 Tflops peak
  - In SARAPE, request "SRN Capacity Clusters"

Sandia PSAAP Compute Resources

- **Sandia Mass Storage System (SMSS)**
  - High Performance Storage System (HPSS) provides high-end near-line storage for HPC systems
  - 123 TB disk cache (short-term, fast)
  - 31PB tape storage (long-term, stable)

- **LYNX**
  - Sandia's HPC file transfer agent nodes
  - Mounts the same parallel file systems available on the compute clusters
  - 10 gigabit and Infiniband (IB) technology to optimize file transfer performance to/from Sandia SMSS
  - High performance data transfer tools HSI, HTAR, and PFTP

Sandia PSAAP Compute Resources

- **Testbed systems**
  - Request accounts through SARAPE (sarape.sandia.gov)
  - Research systems from several vendors
  - Some require NDAs
  - Not for production computing
  - Systems come and go, lifetime usually a year or less
  - Users cannot expect mature hardware/software
  - Support available via Sandia HPC OneStop (e.g. email to <*machine-name*>-help@sandia.gov
  - In-depth support typically provided by research teams

Sandia PSAAP Compute Resources

Tri-lab Advanced System Technology Testbeds - A Co-Design Tool

- Not for production computing cycles - but can be provided to Test Pilot users
- Both hardware and software are intended to be highly dynamic
- Closer to prototypes and technology development drivers
- Multiple nodes available but more important to explore a diverse set of architectural alternatives, than push large scale

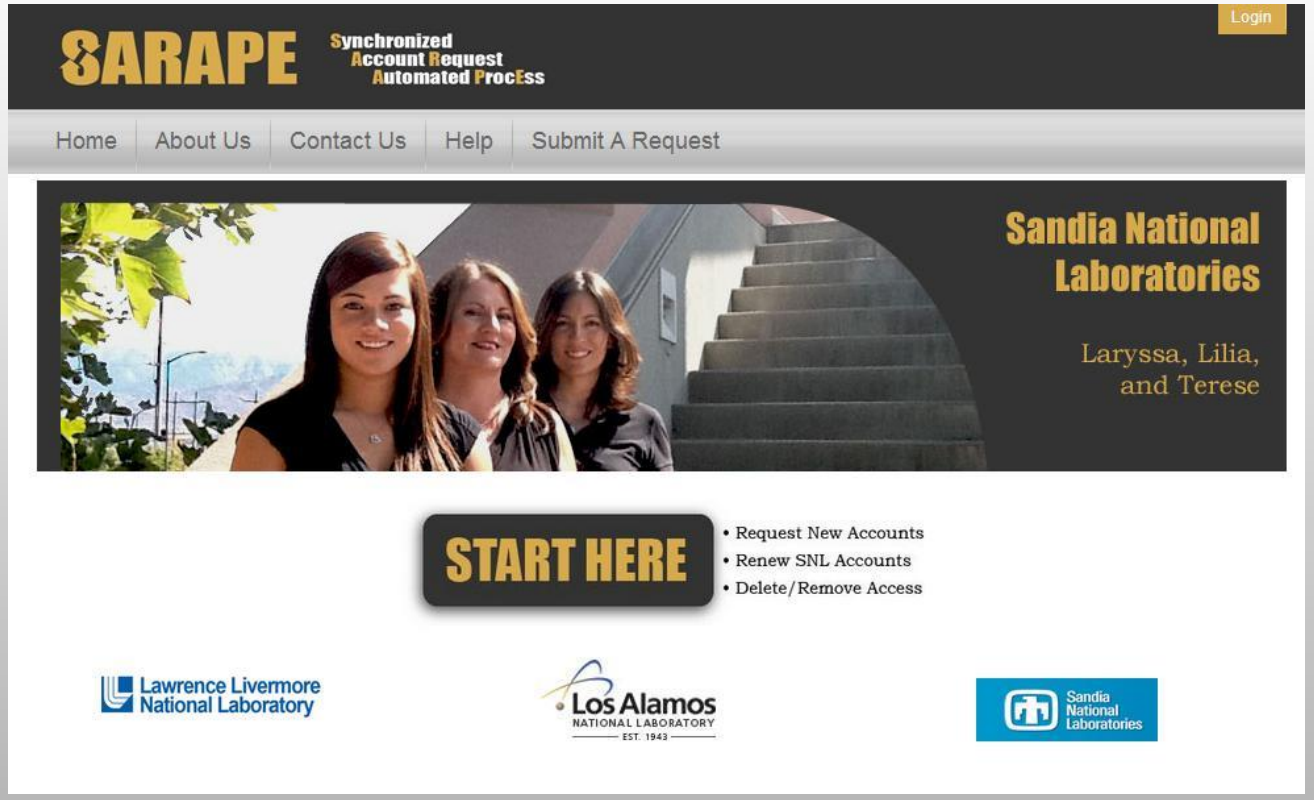
- Available for PSAAP access, via SARAPE.

	Compton	Curie	Shannon	Teller	Volta
CPU	Dual Socket Intel E5-2670 (Sandy Bridge) 2.6 GHz 8-core	Interlagos	Sandy Bridge	AMD A10-5800K (Piledriver) 3.8GHz Quad-core	Intel Xeon E5-2695 V2 dual socket (Ivy Bridge) total 24 cores 2.4 GHz
Accelerator	Pre-Production Intel Xeon Phi Knights Corner (KNC) 2 per node	Nvidia Kepler K20X	Nvidia Kepler K20X 2 per node	Radeon Northern Islands (on die integration)	None
GPU cores	57 1.1GHz cores	2688 732 MHz cores	2688 732 MHz cores	384 800MHz cores	N/A
Nodes	42	52	32	104	56
Interconnect	Mellanox QDR IB	Gemini	Mellanox QDR IB	QLogic QDR IB	Aires
Other	80GB SSD per node	Full featured RAS sys On SRN	Full PCI Gen 3 NVIDIA GPU Direct	Integrated CPU/GPU+ 256GB SSD/node power monitoring capability	Full featured RAS system including power monitoring and control capabilities

http://www.sandia.gov/asc/computational_systems/HAAPS.html

Computer Accounts - SARAPE

- Each Lab has its own policies, forms and procedures, however there is a single entry portal (**sarape.sandia.gov**) for requesting an account at any of the 3 Labs.



Computer Accounts - SARAPE

- Centers need at least one *account authorizer* (also called a GPA). This can be a PI, POC and/or a trustworthy, knowledgeable designee.
- Account authorizers are responsible for overseeing and approving (via SARAPE) the accounts for all of their Center's users. Here's who we have as of today:
  - Stanford – Rika Bosmans
  - Utah – Dav de St. Germain
  - Illinois – Jill Peckham
  - Florida – S. Balachandar “Bala”
  - Notre Dame – Scott Hampton
  - Texas A&M – William Hawkins, Timmie Smith
- Having a “backup” authorizer is important if the primary authorizer is often not available.

Computer Accounts - SARAPE

- After a SARAPE account request is approved by the Center's account authorizer, it is routed to the appropriate Lab for review, collection of additional information and approval.
- Typically takes 1-2 weeks for US citizens
- Account processing for non-US citizens requires additional time and “paperwork” - allow 30-90 days (plan ahead!).
- Lilia Martinez (Sandia: lmartin@sandia.gov 505-845-7967) manages the sarape.sandia.gov user accounts portal and has already contacted the PIs of each PSAAP II Center to get the account request process going.
- Questions? Contact Lilia Martinez or one of the CRT representatives.

Questions?

Backup

HPC Compute Resources Available To PSAAP

Site	System	Architecture	Nodes	Cores / node	Cores Total	Memory / node	Tflop
LANL	moonlight	Intel Xeon Nvidia Tesla	308	16 + GPUs	4,928 Xeon 315,392 GPU	32 GB	488
LANL	mustang	AMD Opteron	1,600	24	38,400	64 GB	353
LANL	pinto	Intel Xeon	154	16	2,464	32 GB	51
LANL	mapache	Intel Xeon	592	8	4736	24 GB	50
LLNL	vulcan	IBM BG/Q	24,576	16	393,216	16 GB	5,033
LLNL	cab	Intel Xeon	1,296	16	20,736	32	431
Sandia	chama	Intel Xeon	1,232	16	19,712	32	410
Sandia	glory	AMD Opteron	288	16	4,608	32	41

National Nuclear Security Administration ASC Los Alamos NATIONAL LABORATORY Sandia National Laboratories Lawrence Livermore National Laboratory

Computer Access

- As with accounts, each lab has its own access policies and procedures.

- All 3 labs require a valid computer account, ssh and use of a password generating device (cryptocard / one-time token).

Queues

- The are limits on the number of nodes and number of hours a job may run.
  - Every machine is different
  - Limits may vary between weekday and weekend
  - Subject to change
- For example, PSAAP resources at LLNL:

Machine	Queue	Node Limit Weekday	Time Limit Weekday	Node Limit Weekend	Time Limit Weekend
CAB	pdebug	16	30 min	16	2 hr
	pbatch	258	16 hr	258	24 hr
VULCAN	pdebug	2K	2 hr	-	-
	psmall	1K	1 hr	-	-
	pbatch	8K	12 hr	-	-