Computing Division

Scientific Computing Services

Unix Town Hall Meeting

Yemi Adesanya, September 22, 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Unix Town Hall Meeting

SLAC

Objectives:

- Communication
- Collaboration

Join our mailing list: unix-community@slac.stanford.edu

email to: listserv@slac.stanford.edu

subscribe unix-community

Scientific Computing Services

SLAC

Scientific Computing Services (confluence) page
https://confluence.slac.stanford.edu/display/SCSPub/Scientific+Computing+Services+Home

New web page under development
https://internal.slac.stanford.edu/computing/scientific-computing-services

unix-admin@slac.stanford.edu
support/questions

yemi@slac.stanford.edu
650-926-2863

Unix Town Hall Meeting

SLAC

Agenda:

- Announcements
- Unix Platform
- Storage & Data Management
- Cloud Compute
- Cyber Security update
- Questions/Discussion

Announcements

Scientific Computing Services

Yemi Adesanya, September 22, 2016

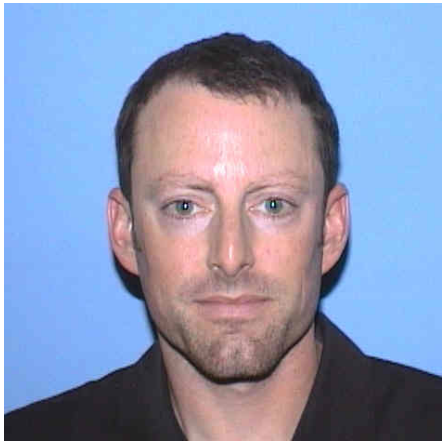
U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Welcome (Back)!

- Welcome, Christa Doane!
- Joined the SCS team in August
- Certified RHEL7 Admin

- Karl Amrhein is back in town!
- Working remotely from Denver, CO
- Linux, Cloud and Virtualization



Conferences and Training

SLAC

- CHEP 2016 hosted by SLAC and LBNL
  - October 10th-14th at the San Francisco Marriott Marquis
- HEPiX Fall 2016
  - October 17th-21st at LBNL
- NVIDIA providing 2-day GPU training for free
  - November 2nd & 3rd onsite at SLAC
  - Hands-on tutorials – bring your laptop
  - Registration details to follow via unix-community mailing list
- SC(SuperComputing)16 conference
  - November 13th-18th in Salt Lake City, UT
  - SLAC partnering with Campus HPC team to highlight research computing
  - Can you provide content for the booth? Contact Neal Adams neal@slac.stanford.edu

Unix Platform

Scientific Computing Services

Karl Amrhein, September 22, 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Unix Platform Update

SLAC

- Red Hat Enterprise Linux and CentOS
- Chef configuration management
- FastX
- OpenStack
- Science VMs on VMware or OpenStack?
- Amazon Web Services

Red Hat Enterprise Linux and CentOS

SLAC

- https://confluence/display/SCSPub/CentOS+7+and+Chef
- Unless RHEL 7 is required by your application for support, CentOS 7 is preferred and recommended instead
- Where is CentOS 7 rolled out
  - Interactive login pool, rhel7.slac.stanford.edu
  - LSF test queue – 'bsub –q rhel7 …'
  - LSST IR2, Zoox, OpenStack servers
  - Self service Openstack CentOS 7 VMs (eg, to use docker)
- Plans: reinstall RHEL5 infrastructure servers with CentOS 7
- RHEL5 End of Life, March 2017 (RHEL 6 EOL 2020)
- CentOS 7 desktop, ITDS portfolio of supported apps
  - standard portfolio of productivity apps like we have for Windows: mail client, web browser, ssh, office suite, etc.
  - CentOS 7 Desktops – for personal productivity, not servers

Chef configuration management

SLAC

- 32 CentOS 7 hosts under Chef config management
  - Most of those are test machines to verify Chef cookbooks
- Support contract with Chef: slack, office hours, etc.
- SLAC will be using Chef Automate for workflow.
  - Focus on standing Automate up and then use it to make and test configuration changes
  - Chef automate provides
    - prescriptive workflow (continuous deployment pipeline, git version control, automatically test all changes on full VMs
    - Visibility – dashboards, reports
    - Compliance – write compliance rules, view reports

FastX

SLAC

- https://confluence.slac.stanford.edu/pages/viewpage.action?pageId=205985167
- Display remote Linux applications (X Clients) on your desktop or laptop
- Can use any standard web browser
- Can use client application
- FastX is a cluster of VMs in VMware
  - As demand necessitates, more VMs will be added

OpenStack

SLAC

- https://confluence.slac.stanford.edu/display/SCSPub/Using+the+OpenStack+cluster+at+SLAC
- 23 Projects
- 29 deft LSF batch bare metal servers
  - 58 VMs
- 12 bare metal servers for self service IaaS
  - 68 VMs
- Subnets available
  - SERV01-*, FARM04-*, FARM10-*
- 780 vCPUs, 4.3 TB Memory, 83.7 TB local disk

VMware on FARM network

SLAC

- Science high bandwidth network (vs. SERV networks)
- VMware FARM clusters - High Availability Row & non-HA
- To virtualize aging servers that are supporting science infrastructure (LSF, storage, etc.)
- Migrate Computing Division VMs on aging KVM infrastructure to FARM VMware (rhel6-64 and iris login pools)
  - VMware has built in lifecycle replacement via leasing
- Science groups no longer forced to buy dedicated KVM hypervisors for VM infrastructure
  - Many individual KVM hypervisors are more time consuming to manage without RHEV or oVirt
  - Better to put hypervisors under a management cluster for HA and increased utilization

Virtualization Technologies - where to run science VMs

SLAC

Requirement	OpenStack	VMware
I need a VM to run a critical service that requires generator-backed power in the datacenter		✓
I need to pick my own VM hostnames (with matching reverse DNS lookups)		✓
My service runs on a single VM with no built-in redundancy or failover across multiple VMs (aka, legacy scale-up application)		✓
I want to manage and change my VM configurations (vCPU, memory, storage, subnet, operating system)	✓	
I need a cloud GUI/API to create and destroy VMs in order to scale out my application automatically or manually	✓	
I want to install or reinstall my VM within minutes at will by myself	✓	
I want to be given a quota of resources (vCPU, memory, disk) and be able to provision and delete my own cluster of VMs	✓	
I want to take snapshots of my VM and manage those snapshot images myself, and revert back to a previous snapshot by myself	✓	

Amazon Web Services

SLAC

- NuSpective SLAC AWS account
- SCS SLAC research-only account
- Plans – VPN
- LSF resource connector
- Testing amazon glacier for storage, seamless LSF bursting, other applications (web, etc)

SLAC

Questions?

Storage & Data Management

Scientific Computing Services

Lance Nakata, September 22, 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Storage Updates

SLAC

- 16 T10000D 8TB tape drives scheduled to arrive next week. HPSS testing and deployment to begin in Q1FY17.
- 4 SSD-based servers have arrived for those tape drives.
- 3 SSD-based servers for AuriStor (AFS) service. Testing and deployment to begin in Q1FY17.

Storage Updates

SLAC

- Begin testing Spectrum Scale 4.2.1. Upgrade GPFS 3.5 to Spectrum Scale by 4/30/2017. As always, we will attempt to minimize disruption to your work.

- UNIX tape backups: will upgrade tsm1 app software and server hardware, plus switch from 1TB to 5TB tape drives.

- Reminder: StaaS chargeback begins 10/1/2016

End-Of-Life Storage Hardware

SLAC

- End-Of-Life = No longer supported by vendor and/or dropping off the SCS roadmap. EOL hardware:
  - Sun Thumpers/Thors (e.g., “kans, wains”)
  - NetApp filers (e.g., “surreys”)
  - Solaris SPARC storage (e.g., “sulkys”)
- Solaris 10 support will end 1/31/2018. Hardware phaseout will continue through 2017.
- GPFS/Spectrum Scale running on RHEL is the current supported storage platform.

End-Of-Life Storage Hardware

SLAC

Science Area	Total Disk Hardware (TB)	Healthy Disk Hardware (TB)	EOL Disk Hardware (TB)
BaBar	2277	1307	970
Fermi	5534	2880	2654
EXO	304	120	184
KIPAC/DES	2612	1200	1412
LSST	46	0	46
CDMS	236	144	92
LCD	113	0	113
EPP Theory	44	36	8
MCC/EED	415	240	175

- ~5.5PB of EPP+PAC storage capacity on EOL hardware
- How do we migrate data off EOL hardware in 2017?

One Solution: Storage as a Service (StaaS)

SLAC

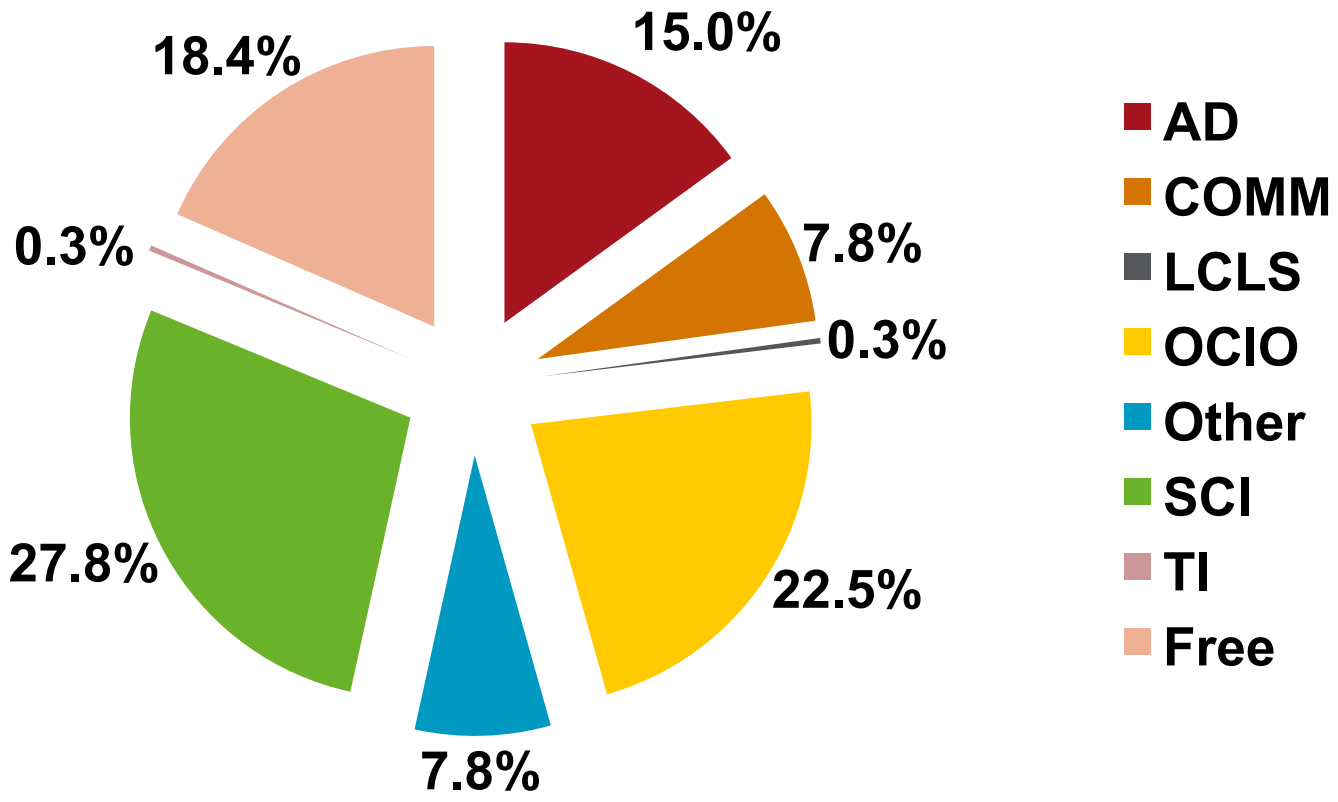
- Improved performance via aggregated storage/parallel file system
- Higher availability via redundant servers
- Multi-PB acquisition = Lower cost per TB
- Automatic tiering to further reduce costs (we plan to gather access pattern statistics)
- Fast storage provisioning (1 day vs. weeks)
- OCIO can lease the hardware and manage lifecycle internally

- Upcoming StaaS upgrades:
  - SSDs for 1) metadata (faster file create/delete/ls) and 2) high IOPS storage pool
  - Expand NLSAS disk tier (~260TB out of 320TB allocated)
  - Add Clustered NFS nodes to better distribute the NFS load
  - (Separate) High Availability (HA) StaaS (in planning stages)

StaaS Usage

SLAC

Storage as a Service
About 261 TB allocated out of 320 TB total



StaaS Rates

SLAC

- SSD, Disk, Tape differ in price and performance, but we need stable costs and a straightforward chargeback scheme
- Current StaaS pricing is $100/TB/year for disk, based on an OCIO 5-year lease of 320TB increments
- We believe we can lease PBs of disk to lower StaaS cost per TB (bulk order; capacity config, not performance)
- One possible approach for tiered StaaS pricing options:
  A. $X / TB / year for data that can be migrated to tape
  B. $Y / TB / year for data that must stay on SSD/disk
- Consensus on retention periods helps reduce price points
  - Example: Groups agree that data not accessed in 6 months or 1 year can go to tape => two price points for option A

StaaS Expansion Proposal

SLAC

- SCS will investigate a multi-PB expansion for StaaS that will combine SSD, disk and tape to address the EOL disk storage for PAC, EPP and other groups while maintaining reasonable performance levels
- Chargeback would reflect tape vs. no-tape storage tiering option
- Ideal: a multi-year commitment from groups to cover the storage lease costs. SCS is also a StaaS customer.

- We are your non-profit computing partner and want a win-win solution, so please share your ideas and feedback!

SLAC

Questions?

Cloud Compute

Scientific Computing Services

Yemi Adesanya, September 22, 2016

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

HPC in the Cloud

SLAC

A quote from a Cray exec. at SC15 conference:

> "The future of HPC is partly cloudy"
>
> Cloud will not replace our entire datacenter (today)

- Why Cloud?
  - Scalability and flexibility
  - Use it when you need it for bursty workloads
- Contributing factors for SLAC cloud services
  - Underutilized static, bare-metal compute clusters
    - Shared clusters averaging 87% ✔
    - Private clusters may be as low as 5% ✖
  - Funding agencies may question requests for hardware purchases
    - "Can the researcher use DOE compute center, NERSC, AWS, etc."

Seamless Cloud

SLAC

- SLAC compute is largely high-throughput, data-intensive
- Compute Models traditionally designed for batch job submission, not for cloud infrastructure
- Let's leverage cloud resources through the existing LSF job scheduler
- LSF Resource Connector = API for interfacing LSF cluster with an external resource provider
  - OpenStack Resource Connector
  - AWS Resource Connector
- Job queues can be configured to burst into OpenStack or AWS when baremetal clusters are busy
- We are not replicating SLAC data in the cloud (for now…)

Seamless Cloud

SLAC

- Seamless Cloud bursting with LSF OpenStack resource connector is working as advertised, AWS is next
  - Opens up possibilities for ‘express cloud’ job queues – why wait for bullets/hequs/etc. to become idle?
  - OpenStack hosts are local – no performance penalty when accessing SLAC filesystems
  - AWS instances will run on a VPN with limited bandwidth – consider setting up a proxy storage service in the cloud (ATLAS)
- Docker container support with LSF is coming
  - Docker supported in RHEL7
  - Try encapsulating your runtime environment in a portable Docker image

Funding Model for Cloud

SLAC

- Introduce Cloud Compute-as-a-Service in FY18
- An alternative to groups buying baremetal compute hardware
- Leverage OpenStack, AWS, VMware
- Try out the service before you commit funds
- We need a straightforward chargeback model:
  - Chargeback per-group or per-project, not per individual
  - Offer a select number of service tiers or bundles
  - Quota based, not actual utilization minutes
- Any new service funding model will have to be reviewed with senior management and business managers

SLAC

Questions?