Computing Division

Scientific Computing Services

Unix Town Hall Meeting

Yemi Adesanya, Associate Director Scientific Computing Services

September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Unix Town Hall Meeting

Objectives:

- Communication
- Collaboration

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

email to: listserv@slac.stanford.edu

subscribe unix-community

Scientific Computing Services

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

Agenda:

- Conferences & Training
- Datacenter Winter Shutdown
- CentOS7 and Chef
- CentOS7 Demo with OpenStack
- Linux Container Deployment
- Storage & Data Management
- Solaris Phase-out
- Linux Desktop Support
- Server Hardware Lifecycle
- Cyber Security Projects
- Questions/Discussion

Conferences & Training

Scientific Computing Services

Yemi Adesanya, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Conferences and Training

- SC17
  - November 12th-17th Denver, CO
  - Joint Stanford/SLAC booth
  - Can you give a 15-20min presentation?
  - https://sc17.supercomputing.org
- LISA (Large Installation System Administration)
  - October 29th-November 3rd San Francisco
  - https://www.usenix.org/conference/lisa17

Datacenter Winter Shutdown

Scientific Computing Services

Shirley Gruber, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Datacenter Winter Shutdown

For new PSLB building, Facilities turning off power to Building 050 for 5 days during the Winter Shutdown period.

- HA (high-availability, generator-backed) systems remain in service
- https://portal.slac.stanford.edu/info/ITHelp/KB/HAServices.aspx
- Everything else goes down
- Provide minimum batch compute required for critical processing pipeline

- Tuesday, December 26 at 10AM – power off
- Saturday, December 30 at 5PM – power restored

A temporary, portable generator will be connected during the outage. To stay within the capacity of the temporary generator, a large number of systems will need to be powered down and stay down throughout the outage.

We would like to take all non-critical, test, and development systems down beginning on Friday, December 22 and keep them down until Monday, January 8. If that is not possible, we would like to take systems down on Dec 26 and bring them up again on Dec 30. Only critical systems will be kept up for the entire Winter Shutdown period.

Datacenter Winter Shutdown

Options:

1. Your systems can stay down between Dec 22 and Jan 8
2. Your systems can stay down between Dec 26 and Dec 30
3. Your systems are critical and must stay up for the entire Winter Holiday, including between Dec 26 and Dec 30

Option 3 notes:

- Systems put on temporary generator will be prioritized based on need and generator capacity.
- Systems must be taken down and brought back up on Dec 26 for the cutover from house to generator power.
- Systems must be taken down and brought back up again on Dec 30 for the transfer from generator power back to house power.

We appreciate your cooperation as we plan this outage.

Datacenter Winter Shutdown

Questions?

CentOS7 & Chef

Scientific Computing Services

Karl Amrhein, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Chef and CentOS 7

Chef is being used for configuration management starting with CentOS 7.

Taylor configuration management will stop with RHEL 6.

Configuration management is a “check and repair” process that runs on each centrally managed machine to ensure the appropriate operating system configuration.

Chef and CentOS 7

~100 Centrally managed CentOS 7 hosts using Chef:

- 67 servers, including LSST, Cyber, GPU, CryoEM, etc.
- 29 LSF batch nodes in the centos7 queue
- 24 SLAC OpenStack VMs - this number fluctuates due to the dynamic nature of the environment

Chef configures CentOS 7 so the machine is secure and centrally managed. Chef provides a modern workflow (CI/CD pipeline) with configurable gates, approvals, testing, and compliance scanning.

Chef and CentOS 7

List of SLAC Chef cookbooks for configuration management:

slac_yum
slac_rsyslog
slac_logrotate
slac_motd
slac_root
slac_krb5
slac_keytab
slac_sssd

slac_shells
slac_ssh
slac_remctl
slac_ntp
slac_sudo
slac_yumcron
slac_dns
slac_openafs

slac_lsf
slac_ganglia
slac_openldap
slac_monit
slac_telegraf2
slac_mailgateway

Questions?

CentOS7/OpenStack Demo

Scientific Computing Services

Karl Amrhein, September 28th, 2017

CentOS7/OpenStack Demo

Questions?

Linux Container Deployment

Scientific Computing Services

Renata Dart & Wei Yang, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Singularity Containers

Singularity is the clear choice for portable Scientific Computing:

- 25+ more newest LSF batch nodes with CentOS 7 (as of this week!)
  - Singularity 2.3.1 and 1.9TB /scratch, NFS and AFS
  - TODO: cvmfs, GPFS
- Users can run their own singularity container images
  - **FAT image**: package your Python libraries in container image
  - Instead of storing on NFS / GPFS !
  - Can significantly reduce the load on NFS servers and GPFS servers
- SCS ***may*** provide standard CentOS 6/7 images with scientific libraries, etc
  - To make the environment similar to the RHEL6 batch nodes
  - As a way to avoid heavy customization on the batch node
  - We are open to suggestions on the above choices
- Still need: a few CentOS 7 login nodes with similar environment

Docker Containers

Do we need Docker on batch nodes? If so

- We need to integrate with LSF
  - To provide basic defense against privilege escalation.
  - No such requirement for Singularity
- We also need infrastructure to host user images
  - How can we be sure that the user images are harmless?
- We don’t envision large scale deployment on batch nodes
  - So many aspects haven’t been carefully thought about

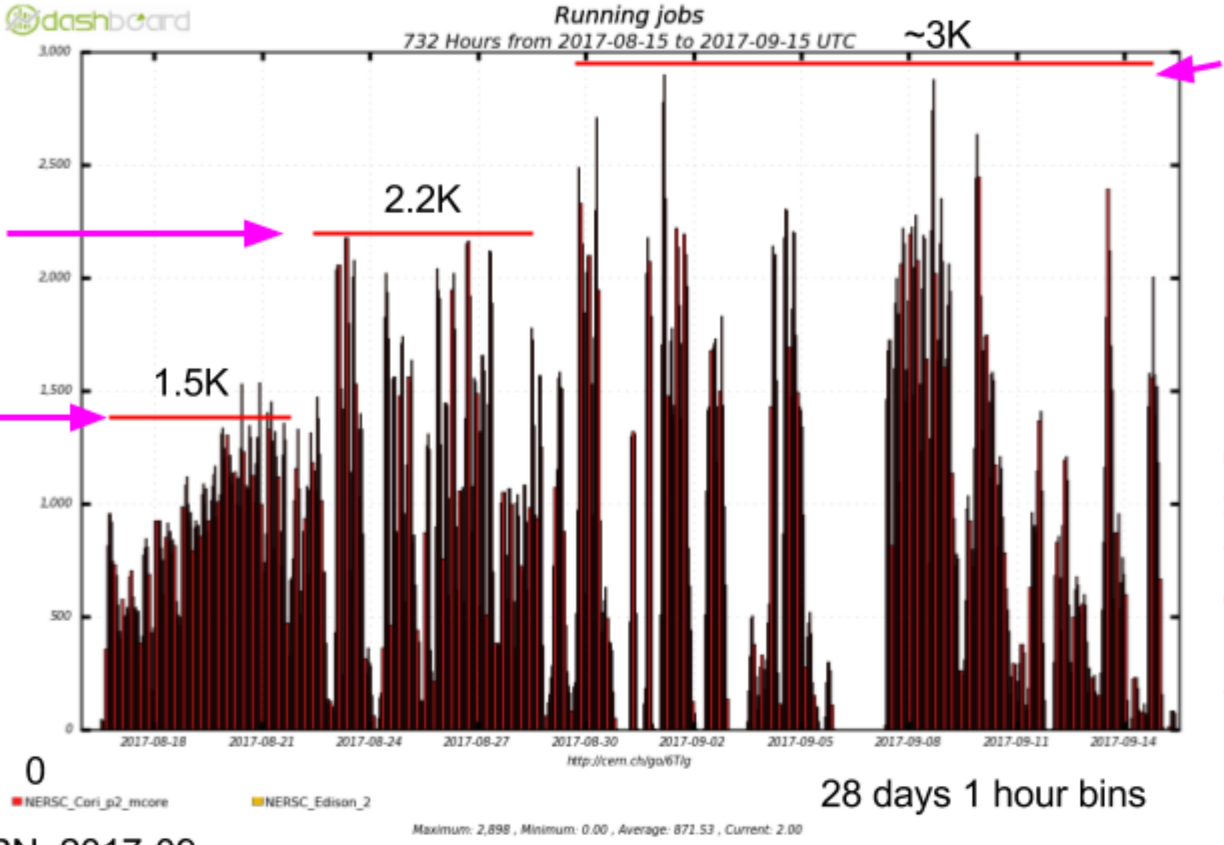
FAT Containers

Effect of using FAT containers on NERSC Cori 2

Running jobs/nodes

Containers and output to Shared file system

Running in Shared File system only - at limit



28 days 1 hour bins

Containers and loopback filesystem for output

Example of effect of 1) using the Shared File System alone, 2) using Containers and Shared file system for output, and 3) Containers and loopback file system for output

Source:
Douglas Benjamin
Wei Yang
ATLAS TIM @ CERN, 2017-09

Linux Container Deployment

Questions?

Storage & Data Management

Scientific Computing Services

Lance Nakata, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Storage Update

- Storage as a Service (StaaS) upgrades since 3/2/2017
  - All metadata now on SSDs; lower latency and less NL-SAS contention
  - Clustered NFS servers now bare metal with 2x cores and memory
  - Plan: increase StaaS data capacity by 50% (Q1FY18)
  - Plan: enhance read performance with Local Read-Only Cache (LROC) on CNFS (requires Spectrum Scale 4.1.1 or later)
  - Plan: create small SSD data area for special needs

- GPFS 3.5 to Spectrum Scale 4.1.1 upgrades ongoing
  - bullet cluster now running 4.1.1 client code
  - Two-step upgrade process from 3.5 to 4.1.1 then later to 4.2.x to reduce scheduled downtime

Storage Update (2)

- 3 SSD-based server for AuriStor service now in production
  - Better AFS performance
  - Server code has many bug fixes
- tsm1 tape backup server upgrade (Q1FY18)
  - SSDs for increased database performance
  - Move from 1TB to 5TB tape drives
  - Likely target for Spectrum Scale HSM tape usage
- IBM TS1155 enterprise tape drive
  - 15TB native capacity, 360MB/s native speed
  - Higher capacity than upcoming LTO 8 (12.5TB) tape drive
  - JD tape possibly writeable at higher density when TS1160 ships in 2018
  - Won’t work in our SL8500 tape silos, but is an option for future tape libraries

Storage and Solaris End-Of-Life

- End-Of-Life = No longer supported by vendor. EOL hardware:
  - Sun Thumpers/Thors (e.g., “kans, wains”)
  - Solaris SPARC storage (e.g., “sulkys”)
  - LSI Engenio disk arrays 9/30/2017 (Fermi, KIPAC, SIMES)
- Solaris 10 support will end 1/31/2018. Hardware phaseout will continue through 2017.
- Working on contingency plans for hardware that won’t make the various support cut-off dates
- Spectrum Scale/GPFS running on RHEL is the current supported storage platform.

Questions?

Linux Desktop Support

Scientific Computing Services

Yemi Adesanya, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

NATIONAL ACCELERATOR LABORATORY

Linux Desktop Support

- We need to define the scope of SLAC Linux Desktop Support
- What is a “Supported Linux Desktop”?
- We need to distinguish between Desktop and Server environments (not simply based on the hardware form factor)
- Desktop:
  - Mostly single-user
  - In a personal office/workspace
  - Intended for personal productivity
  - A gateway to access central Linux services
- Server
  - Multiple users/accounts
  - Kiosk or shared workstation in a lab or control room
  - Treated as production infrastructure

Linux Desktop Support

- Target Operating Systems
  - Ubuntu 16.04 LTS (5-year Long Term Support)
  - CentOS7 (10-year support, 6/2024 EOL)
- Chef Configuration Management
- Windows AD for Authentication
  - Reduce dependence on Unix Heimdal Kerberos
  - We want to converge on a single Authentication System
- Comply with SLAC MinSec (Minimum Security) policies
- Support utilities to interface with central Unix services
  - X11 Graphical terminals via FastX
  - Secure file copy and SSHFS
  - OpenStack Cloud
- Users can still get privs (sudo) to configure their Linux desktops
- Users will be responsible for supporting and patching their custom desktop configurations that fall outside the supported scope
- Desktop Support (SUIT) will support Linux Desktops
- SCS will continue to support Linux servers

Questions?

Server Hardware Lifecycle

Scientific Computing Services

Yemi Adesanya, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Server Hardware Lifecycle

- Hardware Lifecycle is essential for cost-effective computing
- Consider the Total Cost of Ownership in terms of Datacenter overhead and maintenance
- Decrease in reliability over time (> 5 years) impacts uptimes and service availability
- Operating costs increase as the hardware ages
- Our server hardware support policies must align with our lifecycle goals

Server Hardware Lifecycle

- 5-year Lifecycle Objective
- Computing Division will pay for the labor to repair recommended server hardware <= 5 years old
- System Owners must purchase a 5-year warranty with all new hardware
- For existing hardware purchased with 4-year warranty:
  - If hardware failures occur during the 5th year (non-warranty) of the 5-year hardware lifecycle, the owner can choose to:
    - Decommission the system
    - Pay for the cost of parts to repair the system
  - If the user elects to pay for parts in the 5th year, OCIO/SUIT will make a courtesy check for available, used hardware or components in our Data Center and offer it to the user for free, without a warranty

Server Hardware Lifecycle

Questions?

Cyber Security Projects Retrospective & Roadmap

SLAC CISO

Erwin Lopez, September 28th, 2017

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

Collaborative Accomplishments

Multifactor Authentication

- DUO, PIV-I
- Bifurcation of the Network “low” and “moderate” enclaves

Cyber program documentation

- Risk Management Approach updated
- Cyber Security Management Plan updated
- Continuous Monitoring Plan documented
- Common Infrastructure System Security Plan updated

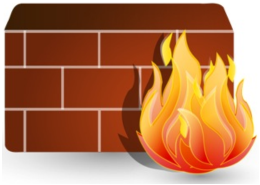
Vulnerability scanning expansion

- 100% of SLAC subnets are either being scanned or are documented as exceptions

Nessus®
vulnerability scanner

Network perimeter tightening

- All networks behind the Palo Alto firewall have a policy associated with them; DMZ firewall deployed



Expanded network monitoring

- Monitoring extended to LCLS, SSRL, and AD enclaves

Continuous monitoring program

- Splunk dashboard, reports, and alerts implemented according to the Continuous Monitoring Plan

splunk>

Make machine data accessible, usable and valuable to everyone.

Current Cyber Security Projects: Looking to the Future

Upcoming Cyber Security Projects

What we're tackling next…

- Unmanaged Systems and Posture Checking
- EnCase Deployment
- Web Application Scanning
- More Cyber Documentation
  - Science enclave System Security Plans
  - Business Impact Assessment
  - Threat Assessment

EnCase Deployment

Problem

- Currently no forensic visibility.
- When an incident or security event happens, don't have access to the memory and can't take a snapshot of processes or ports being utilized.
- Current process is to disconnect the system and bring it to cyber for analysis.

Objective

- Implement a tool to provide visibility to perform forensic analysis over the network on a compromised system.

Approach

- Deploy EnCase Enterprise cyber security forensics software and the agents to centrally managed Windows and Mac systems, and Linux(Optional). **(Windows Endpoint Sept 30, Mac Endpoints 10/15/2017, Windows Servers 10/30)**
- Develop documentation and a repeatable process to deploy future agent updates.

Benefits

- Save time on cyber incident investigations.
- Minimize downtime for the system owner.
- Enhance operations situational verification when investigating a system due to an alert or malicious activity.

Web Application Scanning

Problem

- Currently no vulnerability scanning of web applications. 2015 EA audit finding.

Objective

- Create a process/service to scan high risk moderate applications and externally accessible applications looking for web vulnerabilities.
- Create a process to manage these vulnerabilities and/or create an exception for them based on risk factors.

Approach

- Define a software development life cycle for **new** applications that includes vulnerability management
- Establish a process to check for vulnerabilities in **existing** applications. Follow through on remediation plans
- Implement a dashboard showing progress of vulnerability remediation on:
  - Existing applications
  - New applications coming through the development process
- Establish process strategy
  - Determine tools to use
  - Develop an approach for breadth (externally accessible apps) and depth scans (internal)
  - Develop a risk-based approach for accepting vulnerabilities
  - Perform an initial scan and associated remediation
  - **External scans by 11/21/2017**
  - **Internal scans by Feb 2018**

Unmanaged Systems

Problem

- 2015 DOE Enterprise Assessments audit finding:
  - Insecure and potentially compromised devices can connect to SLAC Network
- US Government FISMA targets
  - All devices have a posture assessment before connecting to VPN
- Stanford University
  - All devices must comply with minimum security standards before connecting
- VPN and Visitor wireless network are used for SLAC business on unmanaged devices (SLAC-owned or employee-owned)
- **Visitor wireless is a mission-critical capability used to deliver services to facility users**

Objective

- Reduce the risk posed by unmanaged systems

Approach

- VPN **(Q1 FY18)**
  - Enable Cisco ASA feature to interrogate systems attempting to get on the VPN
  - Exact approach, TBD
- Visitor Wireless **(Q4 FY18)**
  - Do NOT interrogate systems belonging to facilities users
  - Create a workflow that addresses systems used for SLAC business (SLAC-
  - Exact approach, TBD

Authorization to Operate

- Cyber Documentation
  - Science enclave System Security Plans
    - Haven’t been updated since 2011 or earlier
    - Required to obtain Authorization to Operate
  - Business Impact Assessment
  - Threat Assessment

Back Up Slides

Multifactor Authentication

New Authentication Landscape at SLAC:

- LoA3 (Duo) for remote access
  - ~2,600 Outlook Web accounts
  - ~1,885 VPN accounts
  - ~1,390 Citrix accounts
- LoA3 (Duo) on moderate endpoints of standard users
  - ~300 endpoints

- LoA4 (PIV-I) for privileged users
  - 67 accounts
- Bifurcation of network into “low” and “moderate” enclaves

Multifactor Authentication

Lessons Learned

What went well?

- Socialization with executive team
- Broad communications with stakeholders
- Duo Systems – Windows endpoint modifications; CEO support
- Support from Stanford CISO
- “Soft” rollout of remote access Duo

What could be done better?

- Internal communications and engagement (within Computing)
- Identification of cross-functional dependencies
- Data management to support deployment decisions
- Operational and maintenance planning

Mac Life Cycle Management

Problem

- Older Macs do not receive security updates

Objective

- Eliminate old Macs in use at SLAC
- Ensure SLAC-owned Macs meet minimum security requirements

Approach

- FY17: Develop policy, replacement plan, and cost estimates; begin outreach
- FY18: Implement the policy and plan