SLAC Scientific Computing Services

Yemi Adesanya

October 2016

v. 3.0

U.S. DEPARTMENT OF ENERGY
Office of Science

SLAC NATIONAL ACCELERATOR LABORATORY

About Yemi Adesanya…..

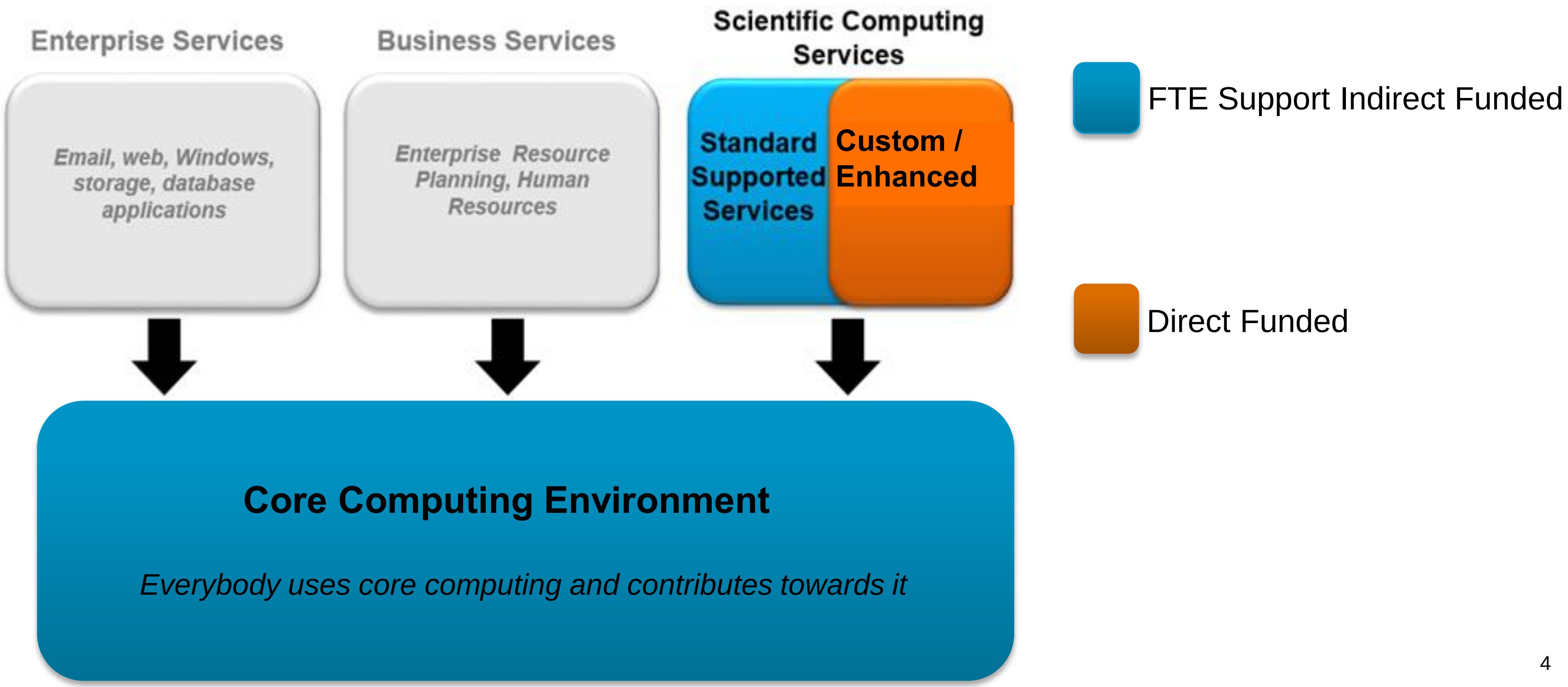
- In my previous life: CERN fellow and technical student
- 18 years at SLAC
  - BaBar computing
  - Object-Oriented frameworks for data analysis (C++, Java)
  - Monitoring and visualization
- Director for Scientific Computing Services since March 2015
  - Be Strategic
  - Drive the computing innovation
  - Build a roadmap for central scientific computing
  - Build relationships

SLAC Scientific Computing Services

- Scientific Computing Services (SCS) is committed to providing a portfolio of Scientific Computing Services to the SLAC research community
- Currently 9 FTEs to support numerous experiments and facilities:
    - ACD, Beam Physics, Klystron, EED, SPEAR
    - ATLAS, BaBar, (Super)CDMS, DES, EXO, FGST, HPS, KIPAC, LCD, LSST, Theory
    - LCLS, PULSE, SIMES, SUNCAT
    - SSRL
  - More than 4000 users
  - Centralized management of 3000 servers (~19K cores)
- SCS is a Service provider
- It takes an OCIO village to support Scientific Computing: Networking, Apps, DBs, Datacenter, Procurement, ITDS, Accounts……

The SCS Funding Model

- SCS FTEs are largely funded through lab indirect
- Justification: SCS services support science across the entire lab
- Exception: “Custom/Enhanced” work is funded directly by experiments



Local Scientific Computing

- Majority of SCS-managed Unix infrastructure and servers are housed in the Computing Division Datacenter (building 50)
- SLAC (HEP) computing is traditionally very data-intensive
  - Workloads are sensitive to I/O latency
  - Requires high throughput to multi-petabyte storage
  - Long term Data management – SCS is the data custodian or archivist
- “Quasi-online” projects with critical processing pipelines
- Provide small to mid-range solutions
  - A stepping stone between analysis on the desktop and a DOE Computing Center
  - Be (relatively) flexible!

SCS as a Service Provider

- We need centrally-managed shared resources
- A Service encapsulates all resources required to meet a science computing requirement: hardware, software, licenses, labor
- Business Model reflects total cost, not just the price of a server
- Users care about the level of service, they don’t have to own hardware
- A Service Catalog is being developed and will be continually refreshed:
  - We don’t run hardware indefinitely
  - We don’t support an OS indefinitely
  - We must innovate to provide more and meet future needs

Storage as a Service (StaaS)

- Users pay $/TB/year rate for GPFS storage
- Now in production with the first billing cycle (Q1 FY2017)
- Improved performance via aggregated storage/parallel file system
- Higher availability via redundant servers
- Multi-PB acquisition = Lower cost per TB
- Automatic tape 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)

Cloud Compute as a Service

- Red Hat RDO OpenStack infrastructure using Dell M630 + R620 servers

- OpenStack IaaS (self-service) cloud
  - More than 32 OpenStack projects
  - Running LSST's Qserv distributed MySQL DB
  - OpenStack GPU VMs are coming

- LSF interface to OpenStack with seamless bursting
  - Running LSF 9.1.x in production and 10.1 test cluster
  - LSF Resource Connector will spin up OpenStack VMs to meet demand
  - AWS with VPN coming soon (but we're not replicating data)

- Investigate non-batch analysis frameworks (eg. Jupyter)
- Charge for Compute based on utilization quota (AKA fairshare)

Update on Central Unix Infrastructure

- 100 Gb/s networking with new Nexus 77 core switches
- CentOS7/RHEL7 for Servers and Desktops with CHEF configuration management
- Migrate AFS to Auristor
  - Kerberos upgrade completed 6/2016
  - Seek to reduce dependencies on AFS/Auristor
- FastX for displaying remote Linux apps on the desktop (with support from Desktop team)
- Phase out Solaris by Q1 2018
- VMware with High-Availability power for hosting critical VMs

Preparing for LCLS-II

- LCLS-II will drive networking, compute and storage innovation
- Goal of 1Tb/sec bandwidth by 2024
- Strong partnership with NERSC is essential
- LCLS-II will require more SLAC datacenter capability
- Upgrading HPSS tape capacity with 8TB Oracle T10KD drives (additional silo is also under consideration)
- Joint development with LCLS-II and Zoox (Autonomous Vehicles) on next-generation GPU cluster for Machine Learning

Challenges to Implementing Chargeback

- There will always be a degree of uncertainty surrounding future funding for experiments
- Science groups may purchase compute and storage in irregular increments
- SLAC finance prefer to chargeback on a regular billing cycle
- Central Computing cannot function as a bank – all money received for services must be spent on resources within the same fiscal year

- Example of how we would like chargeback to work: A group has $50K to spend on storage. SCS is able to take the $50K as a lump sum and map it to an X terabyte disk quota for the next Y years. SCS does not profit – the $50K goes back into storage hardware, licenses, etc.

SCS as a Scientific Computing Facilitator

Wikipedia: *"They help a group of people understand their common objectives and assists them to plan how to achieve these objectives; in doing so, the facilitator remains "neutral" meaning he/she does not take a particular position in the discussion"*

- SCS exists to support all lab science
- SCS has a history of running central resources and providing access to multiple user groups
- We need to reach out to the silos of Computing Expertise
  - Efforts to build a GPU / ML community
  - Free NVIDIA GPU training in November
- Close partnership and alignment with:
  - Scientific Software developers
  - New SLAC Computer Science Department (Prof. Alex Aiken)
- Develop services and expertise to complement other compute facilities
  - NERSC
  - SRCF
- Centralized Scientific Computing must provide an individual with more value/capability than they could achieve independently

Lots of exciting stuff that we need to communicate

SLAC

- Socialization = Share our strategy with management and users across the lab
- Seek buy-in by gathering feedback and revising strategy and proposals where appropriate
- The strategy takes a lab-wide view of Computing so socialization should really start from the top (ALDs and PIs)
- Quarterly Unix Town Hall meetings are an opportunity for SCS to interact with the user community
- Unix-community mailing list

The End

Questions?

SCS as a Service Provider

- Once upon a time, SLAC was a Single-Purpose facility
  - BaBar was the primary experiment and funding source
  - BaBar funded the core computing infrastructure
  - Primary function of SCS (formerly known as “Systems Group”): Run BaBar-owned Unix hardware and Storage
- SLAC became multi-purpose, supporting groups from several Science disciplines
  - More funding sources
  - Stanford-affiliated PIs
  - Wider range of computing models and requirements
  - SCS took on management of several dedicated compute ‘silos’

End-Of-Life Storage Hardware

SLAC

EOL hardware = No longer supported by vendor or no longer on the SCS service roadmap

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