OSG-LHC IRISHEP Retreat

Frank Würthwein
OSG Executive Director
UCSD/SDSC

April 27th 2021

Open Science Grid

UCSD

This Talk has 2 Parts

1) Draft slides for 30 month review
2) Plans for Y4 of IRISHEP

Open Science Grid

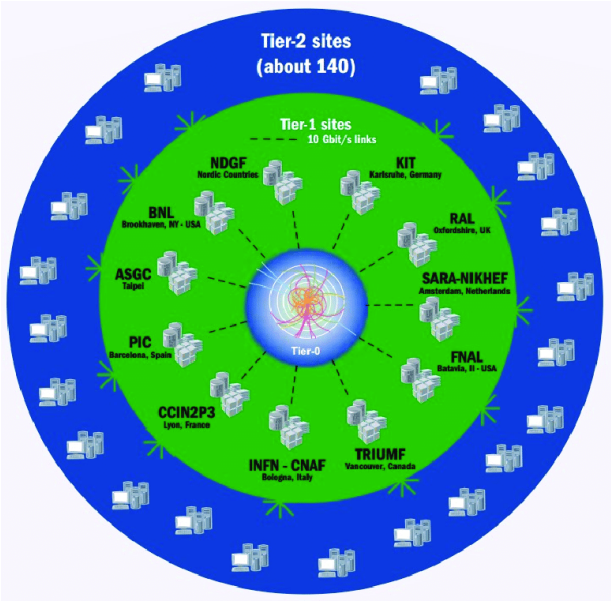
UCSD

Part 1: Draft Slides for 30 month review

(This is still very preliminary …. Am looking for feedback on pitch)

LHC and its global CI

- ATLAS & CMS are global collaborations of many thousands of scientists across more than 300 institutions in more than 50 countries.



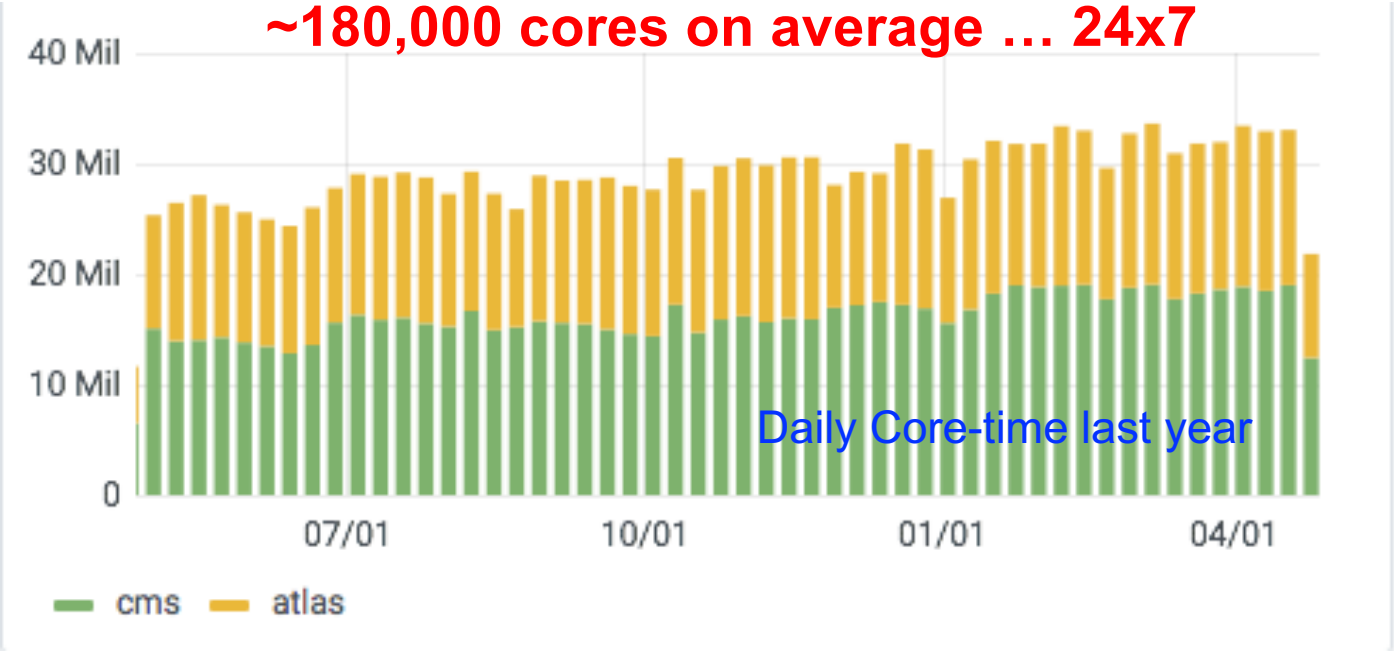
The LHC Science depends on a global CI for dHTC

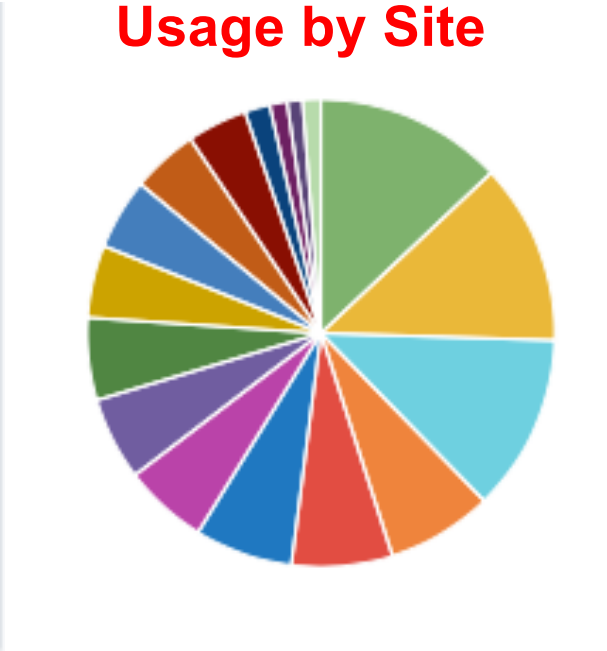
OSG-LHC coordinates the US contribution ~ 1/3 of the total.

Global CI has some unique challenges

- OSG-LHC provides:
  - An integrated software stack, coordinated with the global community.
    - Software lifecycle management, including ingestion, releases, “orphanage”, and retirement process.
  - A cybersecurity Context
    - From policies to operational security
  - Operations
    - Accounting and Runtime environment
  - Network performance monitoring
  - Various coordination functions with the Worldwide LHC Computing Grid (WLCG), the US LHC Operations program, and other science communities.

OSG-LHC by Numbers





Largest sites in US are ~10% of US total => distributed HTC

US LHC Operations Programs are the primary hardware providers, and thus customer of OSG. Lot's of coordination activities, at many levels.

- Annual workshop of US LHC facilities programs co-located with OSG AHM
- OSG-LHC staff attending relevant ATLAS and CMS meetings in US & CERN
- US LHC Ops program management part of OSG management
- …. and many more ….

The People in OSG-LHC


Tim Cartwright


Zalak Shah

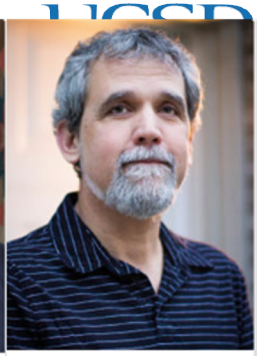

Mike Stanfield


Huijun Zhu


Marian Zvada


Mátyás (Mat) Selmeci


Tim Theisen


Shawn McKee


Brian Lin


Derek Weitzel


Carl Edquist

Operations = UNL = 1FTE
Security = Indiana University = 0.8FTE
Software = U. Wisconsin – Madison = 3FTE
Networking = U. Michigan = 0.4FTE

For full team see: https://opensciencegrid.org/about/team

Anybody who is missing here, please send me picture and name, preferably in a format that fits into this slide nicely.

One Example Innovation in Infrastructure Software and what it takes to change the world

Rather than list all such examples superficially, we decided to discuss one example in some details to explain the context we live in, and the work we do.

From Identity to Capability

- Old:
  - A credential representing a person travels with the workflow around the world to write output to storage via the WAN.
- New:
  - Person receives capability token that allows writing into storage by authenticating at submit point.
  - Capability token travels around the world.
- Key differences:
  - Personal identity does not travel, and is thus less exposed to risk of theft.
  - Actions that travelling credential is allowed can be much more restrictive.

Replacing the concept & software implementation of authz affects almost all the infrastructure software we use globally.

Past Timeline

- Summer 2017:
  - Globus announced end of support of GTC, including GSI & GridFTP for 1/2018.
  - OSG coordinates the international community around developing a replacement strategy.
    - June 6 2017 Announcement on OSG website to calm down global community
    - Lead to creation of Grid Community Forum as code repo for "orphanage" of GTC
  - SciToken (NSF 1738962) awarded July 12 2017.
    - Lucky coincidence of timing
- October 2017: SciToken concept introduced to WLCG-GDB
- July 2018 (CHEP): International HEP Community agrees on globus replacement strategy "in principle" but without any schedule.
- September 2018: First release of Grid Community Toolkit (GCT is "copy" of GTC) available in GitHub and EPEL.
- August 2019:
  - OSG Release 3.5
    - Last OSG Release that includes GSI & GridFTP
    - First OSG release to include SciToken support in HTCondor
  - OSG released schedule for transition (details in google doc)
- March 2020: Proof of concept of SciToken based site
- February 2021: OSG Release 3.6 – First release without GSI and GridFTP

IRIS-HEP has driven this transition

Sprint to the finish line

- April 2021: OSG transitioning SciToken into production for PATh supported science.
- June 2021: OSG organized token hackathon for middleware developers.
- October 2021: Week of workshops to assess status of transition across communities and middleware
  - Pre-GDB meeting for LHC & PATh organized community events
- February 2022: End of Support of OSG 3.5, and thus end of support for GSI & GridFTP

Sprint is joint effort between IRIS-HEP & PATh

PATh's scope is all of open science other than LHC.
It's the perfect partner for IRIS-HEP to impact
dHTC infrastructure software in the wider community.

Summary & Conclusion

- OSG coordinates the infrastructure software and services for a national distributed High Throughput Computing (dHTC) platform for all of Open Science
  - OSG is structured as a consortium
  - OSG depends on contributions from funded projects
    - Dominant contributors: PATh, IRIS-HEP, US LHC Ops, NSF CC* programs, various in kind contributions from the US R&E community at large.
- IRIS-HEP provides the effort to address LHC challenges for the HL-LHC in this context.
  - Manage the needed changes while maintaining a production dHTC platform.
    - GSI => Capability Tokens
    - GridFTP => HTTPS
    - Introduction of containerized services model & remote operations teams
    - Caching as an essential infrastructure component
- IRIS-HEP collaborates with PATh such that PATh provides services for the wider community beyond LHC, and HTCondor software used by the LHC.

IRIS-HEP is an integral part of the wider open science dHTC community

Open Science Grid

UCSD

Part 2: Plans for Year 4

At a High Level

- Complete the HTTPS & SciToken transition
- Execute role of IRIS-HEP in Data Challenge
  - Follow up on what we decide that role should be.
- Complete the XrootD monitoring upgrade into production
  - Special focus on Xcache but including all else.
- Bring an IRIS-HEP supported service through integration to production
  - ServiceX …. Not much to do for OSG-LHC
  - SkyHook

Questions & Comments ?