OSG' Planning work in progress…

Major stakeholder meeting – August 17th 2010
Council Meeting - August 18th 2010

Ruth Pordes
OSG Executive Director,
on behalf of the OSG Project
V4

Schedule (run past DOE and NSF)

- Funding available for current staff through to Sept 30th 2011.
- Develop draft proposal by October 15th. Need small working team for writing..
  - how detailed should it be?
  - Look at different scopes.
- Present to agency sponsors.
  - Have talked briefly with DOE OHEP, ASCR, NP, NSF MPS to request JOT or similar forums/fora in early November. (This is a better time for Alan Stone than October). NSF MPS are on board with this pending talking with OCI.
- Propose to have proposal at least to DOE and NSF by February 1 2011.
- Make a general enough model of the work and organization deliverables that we can respond to specifics quickly.

Preparatory Documents – seen before

- Software Hardening
  http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=866

- Campus https://twiki.grid.iu.edu/twiki/pub/Council/Agenda2009Aug11/NCI__the_Campuses_v3.pdf

- Workforce development
  http://osg-docdb.opensciencegrid.org/cgi-bin/ShowDocument?docid=864

- OSG-TG Principles of Collaboration
  https://twiki.grid.iu.edu/twiki/pub/Council/Agenda2009Aug11/OSG_TG_SharedPriciples_v4-3.docx

- Current Architecture (Draft)
  http://osg-docdb.opensciencegrid.org/0009/000966/001/OSG-Architecture-V2.pdf

- ATLAS/CMS Response to DOE/NSF program managers regarding OSG
  https://twiki.grid.iu.edu/twiki/pub/Council/Agenda2010Mar11/USLHC-OSG-Council-03-2010-1.ppt

- Definition of Satellites
  http://osg-docdb.opensciencegrid.org/0009/000913/002/OSG%20Interface%20to%20Satellite%20Projects%20v3.pdf

Annual Staff retreat in July

Core ET and area coordinators discussed program of work, ideas, out-of-band brainstorming, principles etc to help with an outline of a proposal to put before the Council. Here it is:

OSG 2011 – 2016

PI Input

Context

Vision

Science and Engineering Benefits

Customer profiles

Major accounts

Standard accounts

Cooperating Communities

Possible

Accomplishments to Date

Key Directions for 2011 – 2016

Organization

Project Management improvements

Program of Work

PI Input

- To be a success such that no one will dream to question the existence of OSG.
- We support open distributed science. Noone else can do it.

What we need to do:

- Become more efficient.
- Replace complex/difficult pieces of software stack from VDT and clean things up.
- Engage with first class of scientists.
- Do we solve new users/VO problems with technology or services?
- Start from and Preserve the local environment and then extend this to OSG.
- Success of institution A is also the success of institution B because we are all in OSG.

Satellites are key:

- Endorse and manage the overhead of satellites.
- More efficient to have multiple projects than running a 12-15 million $ project.
- OSG project should anchor satellites.

Data:

- HDFS, Consider approach/principles of contributing to open source community based software: e.g? hadoop, fedora, apache/squid....
- WLCG Demonstrators - are these the best approach with Caching on demand as the mantra.
- Need to consider the Theory of data and data management.
- Would be extremely successful if we could have offered a storage and backup solution to the biologists.
- Consider how to successfully address needs at the end of the WBS: Documentation. Education. Management.

Context

- Integrative collaboration. End-user researcher requirements drive activities.

- Sustaining quality production for active science societies

- Transformation and innovation in end-to-end distributed science

- Anchoring satellite projects for independent but synergistic contributions

- Focus on hundreds of commodity cluster resources of various types, connected via production and/or wide area experimental networks.

- Additionally providing access to specialized resources for specific tests and/or applications.

- Closer ties and partnerships with DOE ASCR facilities and NSF XD and CF21 programs.

DOE ASCR facilities

- ESNET
  - Proposed a model of association for discussion.
  - What is in the "Facility Operations" for potential ?
    - Operational Support.
    - Operational Hosted Services.
    - Availability, accounting and monitoring of Registered (Site and other) Services.
    - Operational Security.
    - Support services – documentation, mail lists etc.
- NERSC
  - First discussion of synergistic areas with s/w

Strategic partnership with ESNET is in OSG's interest.
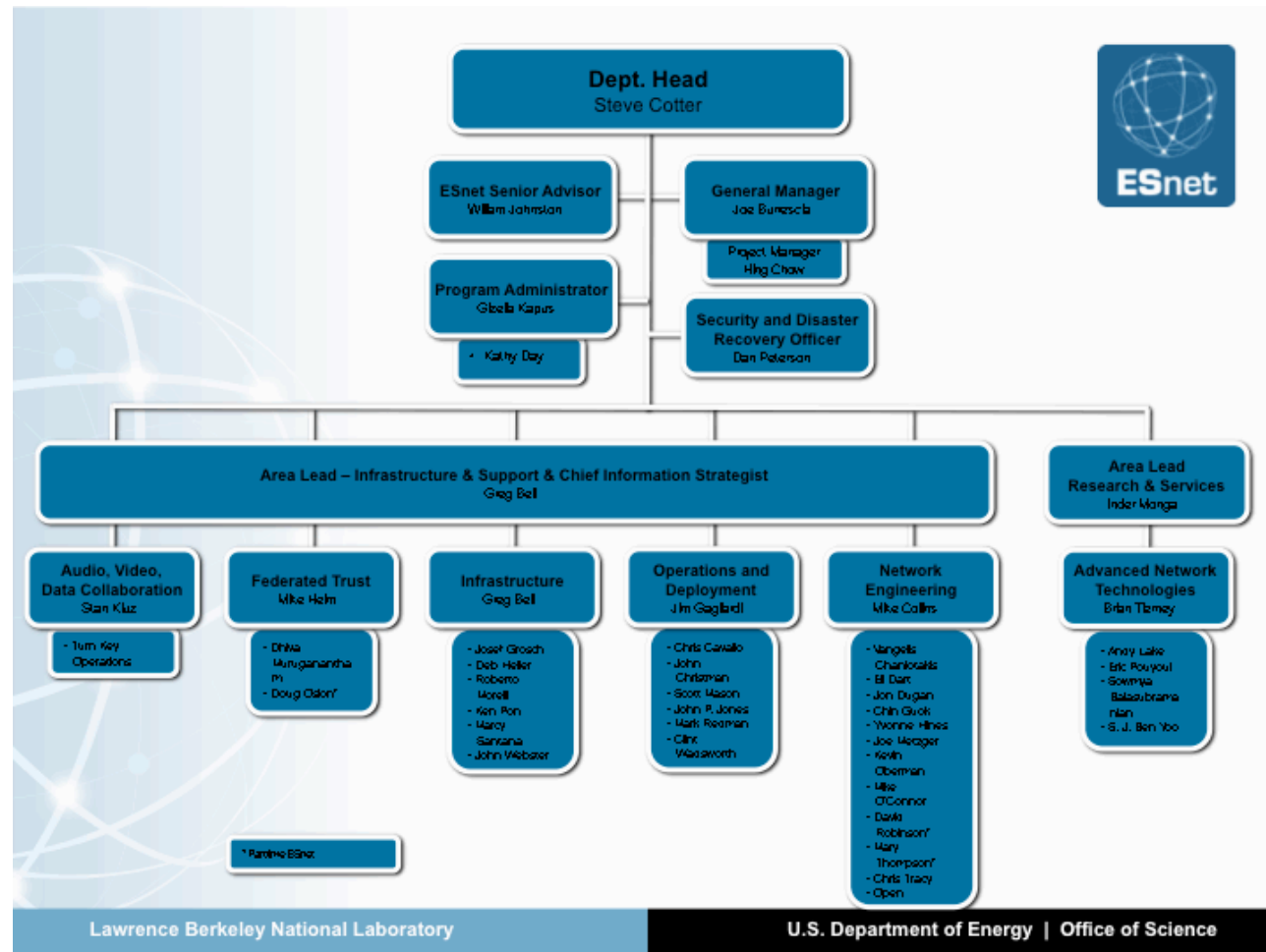
NSF XD

- 2 XD CMS proposals submitted July 16th.
- Separate letters of support/commitment from OSG
- First review mid-September.

NSF CF21

- SSI/SSE/S2I2
- Workshop to provide input to "Distributed Computing, Multidisciplinary Science and the NSF's Software Insitute Program," under the direction of Miron Livny, Ian Foster, Ruth Pordes.
https://sites.google.com/site/distributedcomputings2i2/ Workshops at end of August. Report in mid-september to NSF.

Vision

- Extend current "society" to include other communities, projects, partners.
- Fully integrate the campuses as part of the national cyberinfrastructure.
- Stronger partnerships with DOE distributed computing facilities and NSF XD, CF21, and future programs, advances and innovations in this area.
- Extend the current collaborations to interface to and benefit from satellites.
- Distributed computing laboratory for study and evolution
- Enable science and engineering – ie both pure and applied research.
- Be agile with respect to new capabilities needed by the stakeholders.
- Help with the design and planning, requirements and constraints on external development.
- Help with hardening, transition to production, integration and support.

Science and Engineering Benefits – Customer profiles – (1

Major accounts

Designated by the funding agencies rather than by OSG itself. Should they have a separate "group" for discussions/planning/decisions? Do we gain by formalizing it or not?

- LHC x 10 in data size and computation, may be significant variations in CPUsec/event due to detector.
- LIGO x 10 in data size and computation, open data, going global

Standard accounts – are these subdivided

Incubator? Production? Can we clarify "At –Large" "Collective" better for OSG'?

- Run II – during and after the run.
- Glue-X – part way on the path to dependency on OSG
- Alice – request to be at the table as a major stakeholder.
- STAR
- SBGrid – NIH – request to be at the table as a major stakeholder.
- Sum (0-> ∞) of the "independent investigators/small groups"

Science and Engineering Benefits – Customer profiles (2

Cooperating Communities

Multidisciplinary regional (campus, state, area), enablers, engaging local communities:

- HCC
- Glow
- "OSG"
- "South East/McGee"

…

Sociotechnologists

Computer science – basic and applied

Possibilities

- DUSEL/LBNE
- SNS/APS
- DES
- Networking projects (I2, ESNET)

Accomplishments to Date

- Met LHC deliverables
- Met LIGO Needs
- Enabled other physics and non-physics science
- Proven leading the (Virtual) Society organizational framework, principles, technologies and methods to include new entrants.
- Success in federated operation with EGEE (now EGI) and WLCG.
- Credited by DOE and NSF program offices and projects (including ESNET, I2 and Teragrid, letters of commitment to both XD teams) as part of the ongoing distributed computing landscape.
- Demonstrated that the whole is greater than the sum of the parts for existing members, new entrants and new science.
- Advanced the national interests in services, support, and software for distributed systems physics and beyond.
- Enabled science for multiple "SGs"
- Improved project management processes
- :

Key Directions for 2011 – 2016

- Meet LHC operational needs (see Assessment document)
- Meet LIGO software needs
- Growing the expertise; providing career paths for the staff and participants.
- Role and contributions of Satellites
- Support production, security, users, for an increasing set of communities.
- Continue and expand support for Software collections.
- Where does the testbed fit in?

Addressing LHC futures:

- Configuration Management
- Integration of Commercial and Scientific Clouds
- Usability for collaborative analysis
- Active management of shared capacity, utilization planning, accounting, reporting, change
- End-to-End Data Management challenges in light of advanced networks
- Ensure campus are "first class objects" in national CI
- Integration, hardening, support and broadening the usability to other VOs building on LHC data management/access demonstrators

Addressing LIGO futures:

- Software distributions fully integrated into the underlying OS?
- Reuse and broader support of LIGO security model?
- Better transparency between LDG and OSG and production support for additional analysis.

Run II, SBGRid, Glue-X, STAR… futures:

:

:

Organization (1

- Consortium
- Council
- Science Advisory Group
- Management Teams

- Project Core
- Satellites
- External projects
- Partnerships

Which other successful organizations are we most like (and might learn from) for parts of OSG, whereas as a whole OSG may not be similar to any: e.g.Lockheed Martin, eBay, Redhat/Ubuntu -.

As we grow our customer base, sphere of influence, leaders, capabilities and capacities there are natural lines of distinction between larger groupings of our work than the >14 areas that we now have. Does having these larger groupings – tied at the hip to the overall core OSG project –make us more effective and more efficient towards our pervasive national ambitions.

With a commitment to a broader set of core stakeholders – and a thrust to a deeper support of biologist and non-physics – need to revisit a broader Management structure/Executive Team and a more active, focused Council. With Satellites and/or cooperative partnerships as a more important part of the model will need additional coordination and communication effort across ET and Council. Can this be an action item from the August 18th Council meeting?

Organization (2 – rationalize current areas..)

Project Staff:

Architecture and technology, communications and external relations, documentation, (security officer), program and project management, WLCG interface, LIGO interface, assessment and metrics administration (Council + Project)

- Security
- User Support
- Software
- Production and Operation
- Medium term technology investigations
- Campus infrastructures
- Extensions as closely aligned Satellites

Council staff

- Estimate that Council chair effort is ~1 FTE : 0.5 Chair + 0.5 FTE administrative support. Include this in the cross-cutting project staff.

Project Management improvements

- Plan, decide, act, assess; Tradeoffs, training;
- Methodologies: cmm-sei, agile software, test first "development"
- Relationship to transformation and innovation
- Solid and robust and/or principled evolution

Program of Work

US ATLAS, US CMS, (US ALICE), WLCG Operations (see assessment document). LIGO, standard customer operations.

Production and User Support.

Software & Security.

Create campus societies, core from the Tier-3s.

Provide access to greater diversity of resources as they become available including commercial and scientific clouds, federated infrastructures.

International collaborations..

Distributed and layered data center of unique scale and reach: (>100 universities, >5 national labs, >5 organizational peers).

- Adopt recognized methodology (ITIL, eTOM, ISO20K, other..).
- Continuous improvement.
- Management of change.
- Transform and innovate.
- Investigate outsourcing by the end of the project

Collaborate with/contribute to open source projects we. (do or might) depend on:

Additional capabilities seen as on the community roadmap (not yet explicit):

- Integration with dynamic network provisioning and overlays.
- More complex workflows.
- Data curation.
- Integration with other security /privacy/management services.

International Collaborations through existing OSG communities: e.g. European clients

Satellites

Contributing specific services and deliverables to the mission of the OSG Consortium.

Existing:

- CI team
- CorralWMS,
- ExTENCI,
- HTPC,
- ANI,
- LIGO
- PIF,
- DISUN

Proposed

- Condor/ATLAS/CMS/LIGO (submitted)
- ?Nano@UNL (submitted)
- Education/AIM/Summer School
- Integrated Networks
- SBGrid
- ISGTW
- LIGO PIF'
- DISUN'
- OSG Americas
- SE Regional Grid

Supplements

- Condor
- Build and Test

Overhead to the Cross-organizational staff effort (at Council or Project layer) <~1 month/Satellite.

From Aug 2009 & Mar 2010

Core Mission - 1

Continue operations of the production infrastructure:

Improve the usability and reduce the effort to operate at all levels.

Maintain a usable, secure, robust, distributed facility open to all contributors (resources and applications) in the scientific, research, and education domains. (No commercial use).

Continue support for the OSG Virtual Data Toolkit for the OSG Consortium and other projects, including the evolution of WLCG and EGI.

Solidify presence and usability of shared cyberinfrastructure within the US University campuses and transparent use of local and wide area cyberinfrastructure nationally.

Core Misson - 2

Extend:

The capabilities and capacities offered for the stakeholders at the table, including those represented on the OSG Consortium Council, VOs represented by the "at-large" Council member, and users through the Campus, Engagement, Education and Outreach activities.

Strengthen organizational, operational and technical gateways between OSG and other peer infrastructures including the TeraGrid and XD, and EGI and NGIs in Europe, Campus and regional grids in the US and off-shore, and commercial cloud resource providers.

Workforce training for the future…TBA.

Operations Services, Security, Sites

- **Critical operations center services** – problem recording, triaging, ticketing and ownership; gateways and publishing to the WLCG services.
- Provide frameworks and exemplars for general utility to the communities within the operations framework
- Lower the barrier to use, improve usability of all services and installations.
- Increase automation, alarming.
- Reduce the impact of site-to-site variability and lower the barrier to use.
- Operational support for US LHC Tier 3s and small sites in general.
- Active monitoring and notification of software bugs in s/w.
- Support for Virtual Machines across all operational services impacted (STAR).
- Improve usability of end-to-end security infrastructure – identity token acquisition and management.
- Support for MPI (CIGI, LIGO?, SBGrid)
- Increase the resources accessible for sharing (CDF, D0)
- Help grow sustainable and usable (lower the barrier to entry) regional grids (NYSGrid).
- Campus grids. (GPN, GROW)