

# Towards a HEP(++ ) SSC

WLCG Overview Board, June 2009

Jamie Shiers

Grid Support Group

IT Department, CERN

# Purpose

- The purpose of this presentation is to inform you of what is being done with respect to a proposed “support centre” for the HEP community and beyond in the “EGI era”
  - June 2010 for 3 years?
- Timeline: deadline for proposals November 2009
- Proposed methodology: “bottom-up” approach addressing requirements of LHC data taking & analysis (plus needs of other communities)

# Background

- As we have still not formed the “consortium” that would prepare the proposal, nor discussed in depth with the target communities, premature to give (too) concrete information on proposed activities
  - Think of them as (approximate) “continuation of EIS / ARDA” but refocused on top priorities of 2009 / 2010 and beyond
- **Your input needed to define the direction!**
- There have been discussions in EGI\_DS, EGEE NA4 etc on the roles and functions of an SSC in general
    - Most recently at a workshop in Athens in May;
    - Follow-up meeting in Orsay early July;
    - Barcelona during EGEE’09;
    - “Final” meeting before submission: CERN early November

# Preparation Plan

- Propose following steps outlined in Hyperion FP7 training course held recently at CERN
- A [hep-ssc-preparation@cern.ch](mailto:hep-ssc-preparation@cern.ch) mailing list exists and we have held 2 phone conferences
  - <http://indico.cern.ch/categoryDisplay.py?categId=2439>
- Tentatively, a 1-day workshop at CERN Friday 26<sup>th</sup> June (with EVO)
  - The training proposes an agenda for such a meeting, plus a timeline for preparing the proposal!
- Draft “internal” and “lobby” 1-pagers produced
- More calls / meetings will follow as we move forward with proposal...

# Agenda for Consortium Meeting

- Agree on Aim of Project (one page proposal)
- Refine Idea; address Evaluators' Questions
- Plan for Writing the Proposal
- Writing Team
- Prepare Proposal Review Plan
- [ Not the first time we have prepared an EU proposal! ]

# Introduction – Athens Slides

- From the CERN / WLCG perspective, a “grid support” team has been functioning for many years, focussing on the needs of the LHC experiments but also offering (in)valuable services to other HEP experiments, related disciplines and other projects
- From the EGEE perspective, a HEP cluster has existed for 5+ years
- **Clearly we are not starting from zero but building on what we have established over many years**

# A HEP(++ ) SSC

- CERN expects to be a lead partner in establishing such an SSC and hopes that other players can be identified very soon (if not already)
  - By end May 2009 would be optimal
  - Prague(?), DESY and INFN are understood to be interested – others?
  - **Input from Bob, others, this (Athens) workshop etc required!**
- We would expect to produce a draft – but already fairly complete – proposal no later than July 2009, with further iteration by the time of EGEE'09
  - Semi-public presentation in BCN; discussions with experiments from now
- Further iteration prior to final submission – as part of a complete proposal – for November 2009
- Our assumption is that any personnel would be in place from 1 May 2010 – continuity with the existing team(s) will be an important issue and will require “inventive solutions”
- From a technical perspective, this timeline looks readily achievable

++ includes both “related disciplines” as well as those using common solutions

# Main Theme(s)

- Centres of expertise (excellence?) at CERN and at other sites to match the needs of the user communities supported
  - There is a minimal size below which it does not make sense to go
- These teams should work together – sharing knowledge, tools and experience
  - Sizing based on existing experience: something like 2 people per LHC (or other) VO would really allow us (Europe) to fully exploit the “facilities” (Grid, LHC) that we have build up over O(decade) (or two)
  - It is expected that the people at CERN would be integrated into the existing Grid Support group: in other words CERN would match or most likely exceed the staffing available through “SSC”
- Contacts: myself for “management level” issues; Patricia for technical ones



# Possible Overview

Does HEP include GEANT?  
SIXT? Others?

WP1  
Project Management

WP2: discipline 1

...

WPn: HEP

WPx: photon science

WPy: fusion

...

...

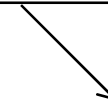
WPy: astro particle

...

...

...

WPz: international communities



EnviroGRIDS etc.

Operations & Support

Training

Dissemination

Common Tools

m/w  
release(s)

CIC  
GOCDDB  
Other OPS tools

GGUS

Ganga  
Diane  
Dashboard framework  
Amga  
Experiment frameworks?  
...

Per area?

# *Recommended Workplan Strategy*

- *Manage by workpackage, not by partner*
- *WP for coordination & exploitation / dissemination*
- *Identify leader + partners*
- *AVOID:*
  - *Everyone in every WP*
  - *One partner per WP*
  - *Doing work in all WPs*
  - *“Floating partners”*
- *Ideal: lead one, be in 2 others*

# A Specialised Support Centre for Large Internal Grid Communities (SLIC)

- INFRA-2010-1.2.1.2 (+ others??)
- The aim of this proposal is to establish a specialized support centre (SSC) focusing on large international communities that use or will use the EGEE → EGI Grid infrastructure.
- The work that will be carried out by the SSC will be predominantly support for production exploitation of the Grid, following on from the successful adaption of numerous major applications and their communities to the Grid environment, ensuring that the return on past investment is maximized.
- Ease of use, low cost of entry and of ownership will be key priorities.

# SLIC - Background

- The European Community has funded three phases of the Enabling Grids for E-science project, which has resulted in large scale production use of world-class Grid-based solutions by many key communities.
- It has established Europe's leadership in this area and – in the particular case of the Large Hadron Collider (LHC) at CERN and its computational and storage needs – brought us to the brink of scientific exploitation of this world facility.
- To capitalize on this investment and to ensure that Europe retains its leading role not only in Grid technologies but also in the scientific disciplines that are dependent on them, continued support for these communities is required.
- In the short to medium term it is expected that this will lead to significant advances in our basic understanding of the Universe around us, whereas in the longer term major spin-offs, both related to the advances in science as well as in Information Technology, can be expected.

# Expected Results + Lead Users

- The expected results of this work are a dramatic increase in the number of Grid users, as usage expands from the data processing activities that have dominated until now into the realm of data analysis, scientific discovery and publication.
- This can only be achieved by a significant simplification of an end-user's interaction with the Grid, through further adoption of existing tools such as Ganga, and by a flexible and scalable end-user support model. This includes the establishment of community support, whereby the communities are encouraged and enabled to be largely self-supporting, with expert guidance to establish and optimize the support structures and associated tools.
- This is essential not only to deal with the large expansion in terms of number of users but also for long-term sustainability.  
[ Numbers ??? ]
- ...
- **Expected Budget: €10,000,000**  
**Framework 7 contribution: €5,000,000**  
**Duration: 36 months**

# The Model – S.W.O.T. analysis

Past



now



Future

Strengths	This model has been proven to work over many years: “experiment integration” → (+) production support → (+) analysis support
Weaknesses	Funding is drying up; people are leaving; some have already gone and more in pipeline...
Opportunities	We have built up a worldwide production quality Grid system over many years: now is the time to exploit it fully – and to respond to Fabiola’s “challenge”.
Threats	We cannot afford to dilute the effort beyond reasonable limits: a team of ~2 FTEs is already on the low edge for an SSC. On the other hand, a team of ~2 “EIS FTEs” per LHC VO has been shown to be effective.

- [ The service ] *“should not limit ability of physicist to exploit performance of detectors nor LHC’s physics potential”*
- *“...whilst being stable, reliable and easy to use”*

# Next Steps (imho)

- We need to move ahead systematically and relatively rapidly in all aspects of preparing a proposal
- We need not only buy-in from the communities that we propose to support, but also their active involvement in all of the above
- This includes the sites / countries / regions involved
- A coherent view – balancing our needs in other areas (e.g. m/w, operations) is required

# Possible Funds (Extract)

- (Cal Loomis) had discussions with people at the European Commission concerning the FP7 Infrastructure "Capacities" call that will be targeted by the EGI and EGI-related proposals. As a reminder, the expected calls are:
  - INFRA-2010-1.2.1 Distributed Computing Infrastructure (DCI)
    - 1.2.1.1 Creation of EGI
    - 1.2.1.2 Service/user support for existing international ("heavy") grid users
    - 1.2.1.3 Middleware
    - 1.2.1.4 New user communities for DCIs
  - INFRA-2010-1.2.2 Simulation software and services
  - INFRA-2010-1.2.3 Virtual Research Communities
  - INFRA-2010-2.3.1 First impl. phase of the European High Performance Computing (HPC) service PRACE
  - INFRA-2010-3.3 Coordination actions, conferences, and studies supporting policy development, including international cooperation.
- The money available in the calls is:
  - 1.2.1.1 + 1.2.1.2 = 25M€
  - 1.2.1.3 + 1.2.1.4 = 25M€
  - 1.2.2 = 12M€
  - 1.2.3 = 23M€
  - 2.3.1 = 20M€
  - 3.3 = 10M€
- In any case, it looks like the SSCs will have to be "split" in some way to take advantage of all of the money in the various calls. Whether this split is purely on scientific discipline or on support activities is up to us. The EC was not concerned about having a scientific community appear in proposals in both calls so long as work was not duplicated between those proposals. I would suggest that for the short term we develop each SSC independently and worry about the placement/splitting/synergies until after the scope and work program of the individual SSCs is clear.



# Discussion

# Tentative Conclusions

- In the above scenario, CERN would expect to lead the HEP WP and to be involved in (at least) the International Communities one
- It would also coordinate the overall effort – at least up to the stage of proposal submission – of the agreed WPs (+ i/f to Cal & other SSCs?)
  - This is arguably enough! (Other WP leads?)
- Need to agree partners and PoW rather soon!
- Checkpoints: July, September, then monthly
  - We still have target of draft prior to Paris meeting! (Clearly not all details but key directions & goals)

**BACKUP SLIDES**

# Without a HEP SSC?

- We have already cut the number of staff supporting the LHC experiments – including in the critical area of analysis support – and further cuts are in the pipeline (budget imposed CERN-wide)
- The EIS team – generously supported in significant part by INFN – would most likely collapse into the most rudimentary support
- The reality is: we will drop from 8 (2008) to 2.5 (early 2010) to <1 (post EGEE III)
- There is nowhere else that we can make up these resources – other key areas are also subject to the same level of drastic reduction!
- Whilst as an enabling factor (using the grid, empowering the VO, ...) this has to be one of the best ROI areas (IMHO...)
  - You can and should design for resilience; you can and should automate; you cannot replace or substitute this area with other than the most skilled and dedicated personnel – such as those we have today
  - (This was not the original timeline / planning...)

# With a (HEP) SSC

- We can really deliver on the promises we have made to our funding agencies – and to our user communities
- We can – at extremely low (or no) cost – assist other important / visible projects
- We can deliver the “added value” – the real(?) reason that much of this research is funded (“science & society”)
- There are many examples of the “added value” of a “SSC-like” team at CERN: and the costs are very modest compared to other types of “petascale computing”

# Pre-GDB – May 12

- Short 10' presentation by each Tier 1 (and CERN): CERN, Italy, UK, Germany, France, Spain, Netherlands, Nordic
- Which services you currently provide for WLCG (via EGEE) that you will commit to continue to support (see attached slide) – what is the level of effort you currently provide for these (separated into operation, maintenance, and development)
- Which services you will not be able to continue to support, or where the level of effort may be significantly decreased that may slow developments, bug fixes, etc.
- What is the state of the planning for the NGI:
  - Will it be in place (and fully operational!) by the end of EGEE-III?
  - What is the management structure of the NGI?, and
  - How do the Tier 1 and Tier 2s fit into that structure?
  - How the effort that today is part of the ROCs (e.g. COD, TPM, etc) for supporting the WLCG operations evolve? How will daily operations support be provided?
  - Does the country intend to sign the Letter of Intent and MoU expressing the intention to be a full member of EGI?
  - Which additional services could the Tier 1 offer if other Tier 1s are unable to provide them?
  - Other issues particular to the country, or general problems to be addressed.
  - What are the plans to maintain the WLCG service if the NGI is not in place by May 2010, or if EGI.org is not in place.
- For ASGC and Triumf it would be useful to hear on their plans in the absence of EGEE ROC support – i.e. do they have plans to continue or build local support centres.
- For BNL and FNAL I assume that nothing will really change on the timescale of the next year.
- If there are other countries with Tier 2s that would like to mention the state of their planning that would also be welcome within the constraints of the time.
- Other issues that need to be discussed include how the support for non-EU, non-US sites will be managed. For example sites in Latin America and others which are currently supported by the CERN ROC.



# EGEE Services needed by WLCG (Plan B)

- **GGUS**
  - Relies on connections to local support ticket systems – today in ROCs and sites
    - → Tier1 and Tier2 sites?
  - COD, TPM
- **Operations and Service coordination**
  - CERN + EGEE ROCs
- **ROCs:**
  - Support effort (TPM, COD) → moves to Tier 1s?
- **EIS team – CERN (largely LCG funded)**
- **ENOC**
  - Coordination of OPN operations- currently by IN2P3
- **Deployment support:**
  - m/w deployment/testing/rollout/support
  - Pre-production testing – effort and resources
- **Operational Security coordination**
- **Policy development**
- **Accounting:**
  - APEL – infrastructure/DB and service
    - NB Italy uses DGAS and publishes into APEL; OSG + ARC publish into APEL
  - Portal – CESGA
- **GOCDB: configuration DB**
  - Important for all configurations and definitions of sites and services
- **CIC Portal:**
  - Contact information, VO-ID cards, broadcast tool, Automated reporting,
- **Availability/Reliability:**
  - SAM framework (and migration to Nagios); SAM tests
  - Gridview/Algorithms etc:
  - GridMap:
  - MSG
- **Dashboards**
  - Service, framework and common services
  - Experiment-specifics
- **Middleware ...**