

# FIFE: Onboarding New Experiments Best Practices

Gabriele Garzoglio, SCD / SCS/ GCS

Oct 31, 2013

# Onboarding Method and Governance

- FIFE will consult with each experiment. FIFE will support the experiment in porting its computations to OSG.
- The porting activity will progress in phases of increased scope and complexity
  - proof-of-principle, production-demo, production-ready
- The experiment should identify a representative for the porting project.
  - S/he will organize the effort from the experiment side.
- FIFE will identify a representative for managing the activity (coordinate communication, record problems / follow up on resolution).
  - S/he will be responsible to SCD for the success of the project.
  - S/he will organize the communication channels and define with the experiment the scope and deadlines for the phase

# Communication

- For every porting phase, we should have
  - Onboarding face to face meetings with the experiment to identify the goals of the phase
    - Scientific goals, target application, application workflow, size (# jobs, IO per job, total jobs / IO), scope (total duration, operational needs and complexity), deadlines, representatives, etc.
  - Regular meetings
    - weekly, daily stand-up towards the deadline, etc.
  - Web context
    - To record the goals of the phase, important architectural decisions, links to GOC/Fermilab tickets, links to documentation
  - Dedicated archived mailing list
  - Instant messaging room (TBD)
  - Regular reporting to management
  - Final report with lessons learned
  - OSG Newsletter article (or equivalent) at the end of the phase to document and advertise the work.

# Operational Environment

- Create accounts on submission machine (access to jobsub client)
  - [https://cdcvs.fnal.gov/redmine/projects/ifront/wiki/Steps\\_to\\_adding\\_a\\_new\\_experiment\\_to\\_gpsn01](https://cdcvs.fnal.gov/redmine/projects/ifront/wiki/Steps_to_adding_a_new_experiment_to_gpsn01)
- Create dedicated gWMS Front End group
- Interact with gWMS OSG Factory admins to add initial set of OSG sites
- Establish app distribution mechanism; if required, request space on OASIS for CVMFS / app distribution
- Identify archival storage and verify integration with IFDH
- Identify and set up other specific FIFE services (calibration distribution, SAM, on-demand, ...)
- Engage dedicated and opportunistic OSG sites. When interacting with sites, we should consider requesting authorization for more than one experiment at once (CE/SE authZ, storage areas, cvmfs mount points, etc.)

# Code Portability

- The experiment and FIFE representatives will work to enable the portability of the target application code
  - Remove hard coded paths
  - Improve error checking in script / failure modes
  - Arrange workflows “vertically” to take advantage of data locality at the worker node i.e. app chain running all at one worker node, when possible
  - Tune application parameters to fit the resource constraints of the OSG WN (local scratch, memory, time limits, etc.)
  - Integrate storage management for I/O
  - Setup running environment
- The FIFE and experiment representatives will run tests of portability on a WN-like environment deployed locally and with interactive access until the application is ready for a new grid-enabled release.

# Test Cycles

- The FIFE expert will coordinate periodic cycles of generic and specific tests towards the goal of the phase. This will help identify problems and the maturity of the activity.
  - FIFE will be responsible to develop a set of tests to verify that the relevant OSG services are supported
    - AuthZ at CE/SE at FNAL and sites; availability of local resources at sites; Software in CVMFS; ...
  - The FIFE expert will work with the experiment contact to define short jobs that test the application on OSG end-to-end.
    - These should be based on tagged stable grid-enabled releases as much as possible.
    - These jobs will evolve to full-fledged jobs as the infrastructure matures.
- The FIFE expert will be responsible for collecting problems, opening tickets, and driving resolutions.

# Phase close out

- The phase will be considered closed when the experiment contact will have run operations of the target application on OSG to produce / process the agreed amount of data.
  - Growing in scope and complexity, these engagement phases will prepare the experiment to run their computations on the OSG, without consuming undue resources in the engagement process.
- Reports / lessons learned, Newsletter articles, celebrations, etc.