

CMS Monte Carlo Processing Service - ITB Application Integration Plan

Greg Graham FNAL CD/CMS

OSG Integration Workshop at UC

Feb 15-18, 2005

CMS MCPS

- MCPS (Monte Carlo Processing Service) is a processing job builder and submission client
 - processing Monte Carlo data on the CMS User Analysis Facility
 - Accesses experiment services to locate datasets and store results of processing

- ITB Success Table

| Task/Location | CMS ITB Site | non-CMS ITB site |
|-------------------------|---------------------|-------------------------|
| Install CMS Code | Phase 1 | Phase 3 |
| Run Jobs | Phase 2 | Phase 4 |

- Potential issue: Gathering information needed for job submission.
 - Explore using GridCat for this at the workshop.

MCPS Description

- MCPS Workflow
 - User executes a general purpose macro script to create jobs that MC data and/or process existing data
 - Needs access to experiment services to discover and access data
 - Needs access to OSG services to submit jobs
 - GridCat to discover information about state of the grid
 - MOP to submit jobs through Condor-G
 - MonaLisa to discover information about the state of the jobs
 - SRM to transfer files to/from ITB site
- The validation test will consist of MCPS developers installing CMS software and submitting jobs to DAG cluster at FNAL and other ITB sites.

Installation and Configuration

- MCPS/MOP needs only be installed at the site of job creation/submission respectively.
- We need to explore with Bockjoo the ability to store MOP information variables in GridCat.
 - These are currently in a CMS tool called ConfMon.
 - MonaLisa is not yet integrated into MCPS jobs, so no special integration work is needed yet.
- Delivery method
 - N/A
- Configuration high points
 - Getting MOP information variables into GridCat
 - Backup is to use the old information files
 - but this requires information to be collected beforehand.

Test and Validate

- The service is considered tested after:
 - Phase 1: MOP successfully deploys CMS Application to DAG cluster site running VDT 1.3.1
 - Phase 2: MCPS/MOP successfully submits jobs to DAG cluster, they are seen locally to be running, and results are returned to submission site.
 - Phase 3: MOP successfully deploys CMS Application to ITB (not DAG) cluster site running VDT 1.3.1
 - Phase 4: MCPS/MOP successfully submits jobs to ITB (not DAG) cluster, they are seen by MonaLisa to be running, and results are returned to submission site by srm.
 - Scale: DAG cluster - about 10 jobs
 - Scale: non-DAG cluster - about 10 jobs
 - Higher scale tests should be scheduled as a planned OSG activity, because we would like to go to >5000 jobs.

Support and Documentation

- MCPS
 - <http://www.uscms.org/SoftwareComputing/Grid/MCPS/>
- MOP
 - http://home.fnal.gov/~anzar/MOP_MASTER/MOP_Reference.html
- Runjob Project
 - <http://projects.fnal.gov/runjob>
- Contact information for support of the service in the integration phase
 - MCPS/MOP: Dave Evans evansde@fnal.gov
 - MOP: Craig Prescott prescott@phys.ufl.edu and Vijay Sekhri sekhri@fnal.gov
 - Coordination: Greg Graham ggraham@fnal.gov

Pending Issues and Next Steps

- MOP variables are not mapped to GridCat information
- MOP submission traditionally requires operator to choose a site
 - If MOP variables general enough, ITB sites should be able to be characterized completely through GridCat information
 - Automation
 - Revisit Dan Bradley's prototype
- Rough timeline
 - Up to CMS project management in conjunction with PPDG coordinators and OSG coordinators