

JIM DEPLOYMENT FOR THE CDF EXPERIMENT

Mòrag Burgon-Lyon¹, Andrew Baranovski², Valeria Bartsch³, Stefano Belforte⁴,
Gabriele Garzoglio², Randolph Herber², Robert Illingworth², Rob Kennedy¹,
Ulrich Kerzel⁵, Art Kreymer², Matt Leslie³, Lauri Loebel-Carpenter², Lee Lueking²,
Adam Lyon², Wyatt Merritt², Fedor Ratnikov⁶, Alan Sill⁷, Richard St. Denis¹,
Stefan Stonjek^{1,2}, Igor Terekhov², Julie Trumbo², Sinisa Veseli², Steve White²,
¹University of Glasgow, ²Fermi National Accelerator Laboratory, ³University of Oxford
⁴INFN / Trieste ⁵Universität Karlsruhe ⁶Rutgers University ⁷Texas Tech University

Abstract

JIM (Job and Information Management)[1] is a grid extension to the mature data handling system called SAM (Sequential Access via Metadata)[2] used by the CDF[3], DZero and Minos experiments based at Fermilab. JIM uses a thin client to allow job submissions from any computer with Internet access, provided the user has a valid certificate or kerberos ticket. On completion the job output can be downloaded using a web interface. The JIM execution site software can be installed in shared resources, such as ScotGrid[4], as it may be configured for any batch system and does not require exclusive control of the hardware. Resources that do not belong entirely to CDF and thus cannot run DCAF (Decentralised CDF Analysis Farm)[5], may therefore be accessed using JIM. We will report on the initial deployment of JIM for CDF and the steps taken to integrate JIM with DCAF.

INTRODUCTION

The CDF Experiment declared its intention to have 25% of computer processing power available from off-site resources by June 2004, with 50% by June 2005. The first target has been achieved with just over 35% or 1.8THz of CPU available at the time of writing. However the target is moving, and the total resource requirements increase over time. To continue to keep pace with the growing demand, a combined solution utilising SAM for data handling, CAF and DCAF software on the analysis farms and JIM to manage, distribute and execute jobs has begun implementation.

SAM station software can be easily installed using a single script on any individual PC or cluster. By communicating with a database based at Fermilab, SAM can be used to store or retrieve single files or entire datasets. Files are delivered by the fastest method from neighbouring sites, or from Fermilab. SAM can be used on its own or in conjunction with DCAF or JIM.

DCAF software has been installed on numerous clusters at sites remote to Fermilab and uses an FBSNG or Condor batch systems to control and monitor jobs sent to that cluster. The original two CAF installa-

tions at Fermilab coupled with the nine remote DCAF sites, now in production, provide sufficient processing power to meet the current needs of the experiment. However, this is not a true grid solution as users select the DCAF on which to run. In addition, growth in the number of DCAF sites is limited as complete control of the batch queues is required. Therefore shared resources, including those used by LHC experiments, can not be harnessed by CDF.

JIM provides job and information management, communicating with all JIM sites via the remote broker, and selecting the most appropriate location for job execution based on location[6] of the data and the load on the batch system. To submit a job to JIM, a user must have a valid kerberos ticket or grid certificate. Thus JIM meets the criteria of a grid solution. The software can be installed on a front-end node on a shared cluster, interfacing to local batch systems. Thereby peacefully co-existing with the other users and grid software on that cluster.

COMPONENTS OF JIM

JIM has been written as a grid extension to SAM and can be thought of as having two parts, Job Handling and Job Monitoring. JIM and SAM together are known as SAMGrid. From an installation perspective JIM is comprised of four site components: client; submission; monitoring; and execution. The thin client site component can be installed almost anywhere and provides the user with uniformity of submission, regardless of where the job is submitted or actually runs. The client is used to pass an authenticated user's job, with a valid kerberos ticket or grid certificate, to the regional submission site. The submission site queues the jobs using condor.

The submission site currently uses the Condor match making service to automatically select the job execution location. A site's resources are described using XML with subsequent projections onto the Condor Classified Advertisements (ClassAd) framework. Condor has been extended such that it communicates dataset distribution information with SAM to influence the choice of execution site[7]. SAM can be used

to transfer large volumes of data from any SAM cache to the JIM execution site, and also to store files back into the main SAM database. Depending on the location of the required data, and the resource availability JIM will send the job to the best suited execution site rather than strictly taking the job to the data or the data to the job.

When development first began on the JIM extension to SAM, the grid tools available were too limited for the task. Initially the JIM developers extended the Condor match making service to provide sufficient functionality to allow JIM to operate with SAM as a full grid service. As grid technologies have advanced, the JIM development team have made changes to use these new and improved technologies. An example of this has been the updating of SAM-Grid software for compatibility with Globus 2.4.3.

The Virtual Data Toolkit (VDT)[8] is an ensemble of grid middleware that can be easily installed and configured. VDT supports the LHC Computing Grid (LCG) and the Particle Physics Data Grid (PPDG). Distribution of SAM-Grid is now available using VDT, helping to ensure the continued ability of the SAM-Grid system to co-exist on LHC shared resources.

IMPLEMENTATION ISSUES

Earlier this year, the JIM development team focused on Monte Carlo (MC) production for the D0 experiment. The D0 success rate for MC is now over 99%. Efforts are underway to emulate these results for CDF. A script that makes a tarball from the CDF software environment has been used to manually run CDF MC on D0 computing facilities at Wisconsin. This first step means that the CDF software environment can be transferred around the grid, preventing problems with differing code versions on execution sites. This ensures that shared resources can be used fully for CDF jobs without application version issues. This approach is not required for DCAF submissions as up to date CDF code is available at each DCAF site. Integration of JIM and DCAF has been successful using an earlier version of the JIM software. However the sandbox size that can be used for a job submitted to DCAF is limited, and before JIM can be deployed on existing DCAFs consideration must be given to the challenge of ensuring that all sites, available for job submission using JIM, have access to the CDF software environment without overloading the DCAF systems with an overly large sandbox. A possible solution is to ensure up to date CDF software is maintained on the same front-end node as the JIM software.

MC can be split into three separate parts, each taking the output from the previous step as input. When a job is first submitted, a list of run numbers is given. The run number list is filtered using a database query to ensure only runs for which the luminosity makes

sense will be used. The output is passed through simulation, reconstruction and production.

The MC methods differ between the CDF and D0 experiments, thus extra development effort is in place to change both JIM and CDF MC code. This will allow the collation of information from earlier parts of the MC generation to ensure that any failed jobs will not adversely affect the end results. Resubmissions of failed jobs can be generated using an appropriate random number seed.

SAM station installations have been vastly simplified by the creation of a script. The once timely and difficult process can now be completed within a couple of hours, largely unattended. Simplifying the installation procedure was a critical step to allow the quick roll-out of SAM, a crucial element of the CDF Grid software setup. A similar script is under development to provide the same ease of installation for JIM.

The JIM installation script is more complex than the SAM script. The requirement of a host, server and user certificate coupled with the fact root access is required for parts of the installation that deal with grid security makes scripting difficult. One of the tailoring scripts requires information about the system architecture, such as the location of the scratch area, which the installer has to decide. When the installation assumes a default value that is not available on a shared system (e.g. using the top level of the /opt directory) the installation will fail. In addition, there are four separate types of installation, giving several possible combinations of site, although in practise however, only three types of installation are common (client only; execution and monitoring; and all four at submission sites). As JIM can interface to many different batch systems configuration of the local system is required, which cannot be readily scripted[9].

CDF uses local databases to handle product installations, avoiding the need for root access. The installations of SAM and JIM software are segregated into two separate databases. The SAM Station script is presently used regularly by participating sites to update SAM installations, keeping them current. By using a separate products area for JIM, the potential pitfall of overwriting SAM configuration with JIM updates or visa-versa is avoided.

The area of tailoring presenting the most challenges is the execution site. As JIM will work with any batch system, automating the installation has proven difficult. For example, as yet there is no way to query the system to discover the appropriate queue name automatically. Efforts to simplify the iJIM installation process by the development team are coupled with the deployment preparation of simplified scripts.

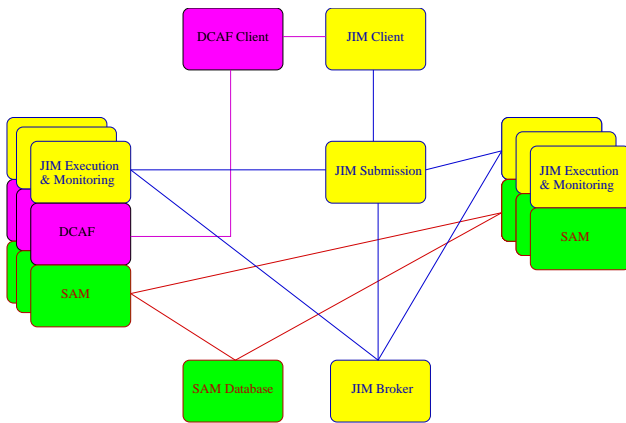


Figure 1: How CDF Grid fits together

DEPLOYMENT CONSIDERATIONS

The above diagram, figure 1 shows how elements of the CDF Grid will fit together once JIM is deployed. Users currently submit jobs from their terminal to DCAF. Although this will still be an option, once JIM is fully deployed, users will be encouraged to submit their jobs through the JIM client software. JIM client passes the job to the area submission site for queueing. After communicating with the broker, the job will be sent to an execution site, which may have a DCAF. The job will be executed, using SAM to transfer files, and DCAF or the local batch system (e.g. PBS) to execute the job.

JIM deployment is planned for existing SAM stations using the JIM installation script. The script is currently being tested and refined using an eleven node Dell cluster based at The University and Oxford. The installations of the latest version of SAM-Grid are planned for sites at the University of Glasgow, using a PC frontend to ScotGrid, and an second Dell cluster.

The steps to full CDF grid deployment entail testing a complete JIM and SAM installation at the Universities of Glasgow and Oxford; addition of JIM to a site running DCAF, and finally roll-out of JIM software using the JIM installation script. Testing will focus on Monte Carlo and Analysis jobs and will also include Reconstruction as these are the main job types used for the experiment.

The deployment will be separated into physical regions (e.g. Scotland, Southern England) where each physical region will have one submission site installation. Locations wishing to use the JIM interface to CDF Grid will require a JIM client installation. Local installations of the small JIM client software will be configured to pass submitted jobs to the regional submission site. Monitoring site installations will be required at execution sites. Each site that runs DCAF and SAM will be expanded with at least the JIM execution site installation. As some sites will require all four installation types, the JIM installation script will

allow the local system administrator to pass over installations the site does not require.

Due to the issues presented, the planned deployment of JIM software has slipped. As these issues are being addressed, the deployment is now planned for the Autumn.

SUMMARY

CDF have sufficient DCAF installations to allow 35% of computing to be done on site. To achieve the goal of 50% by June 2005 a grid solution is required to harness shared resources where DCAF installations are not possible.

JIM deployment has been delayed due to difficulties simplifying the installation procedure. Due to these problems, investigation is underway into possible alternatives to some parts of JIM, particularly the execution site. Deployment of JIM or a grid3/JIM combined solution is still planned for the near future.

REFERENCES

- [1] JIM home page
<http://www.d0.fnal.gov/computing/grid/>
- [2] SAM home page
<http://projects.fnal.gov/samgrid/cdf/cdf.html>
- [3] CDF home page
<http://www.cdf.fnal.gov/>
- [4] ScotGrid home page
<http://www.scotgrid.ac.uk/>
- [5] DCAF home page
<http://cdfcaf.fnal.gov/>
- [6] M. Burgon-Lyon, et al., Experience using grid tools for CDF Physics, Advanced Computing and Analysis Techniques in Physics Research (ACAT) 2003.
- [7] A. Baranovski, et al., Management of Grid Jobs and Information within SAMGrid, Computing for High Energy Physics (CHEP) 2003, La Jolla, CA
- [8] Virtual Data Toolkit
<http://www.cs.wisc.edu/vdt/>
- [9] G. Garzoglio, I. Terekhov, A. Baranovski, S. Veseli, L. Lueking, P. Mhashikar, V. Murthi, The SAM-Grid Fabric Services, Advanced Computing and Analysis Techniques in Physics Research (ACAT) 2003.