DØ DATA REPROCESSING WITH SAM-GRID

J. Snow*, Langston University, Langston, OK 73050, USA
D. Wicke, Bergische Universität, Gaußstr. 20, 42097 Wuppertal, Germany
M. Diesburg, G. Garzoglio, FNAL, Batavia, IL 60510, USA
G. Davies, Imperial College London, London, UK

Abstract

Periodically an experiment will reprocess data taken previously to take advantage of advances in its reconstruction code and improved understanding of the detector. Within a period of 6 months the DØ experiment has reprocessed, on the grid, a large fraction (0.5fb^{-1}) of the Run II data. This corresponds to some 1 billion events or 250TB of data and used raw data as input, requiring remote database access. This is the largest HEP grid activity and has been a great success. SAM (Sequential Access to Metadata) has been in operation at DØ since the start of Run II and provides the data-grid (also enabling remote analysis). Job submission and management are provided by JIM. Together they form the middleware SAM-Grid, used for this activity. This massive task led to extensive developments in SAM-Grid, in a joint effort between the core developers and those carrying out the reprocessing at the remote sites. The resources used, corresponding to some 3500 GHz equivalent, were shared and include LCG and OSG facilities. This activity, including the development of SAM-Grid and the operational tools and procedures developed will be presented. Lessons learnt from carrying out such a task on the grid will be discussed.

INTRODUCTION

Improvements in understanding the detector, for example the calorimeter calibration, and improvements in the algorithms used in the analysis of the data periodically prompt the re-analysis of the accumulated raw data sample. Such was the case in 2005 when over a span of about six months 470pb^{-1}, corresponding to 1 billion events or 250 TB, of Run II data was reprocessed by the DØ collaboration [1]. Significantly this was accomplished primarily using SAM-Grid and remotely from FNAL, engaging both dedicated and non-dedicated facilities including OSG and LCG resources. The project was five times larger than the prior reprocessing effort [2] of 100pb^{-1} in 2003-2004 in which 30% was done with remote processing.

To accomplish this task in a timely manner it was determined that approximately 3400 CPU-GHz of computing power would be necessary for six months. With the DØ central farm of 1000 CPU's occupied with data taking remote processing was imperative. Building on the successful experience of running Monte Carlo simulation on the SAM-Grid [3] the reprocessing application was also designed to use it. Unlike simulation though access to information in databases at FNAL was necessary for the analysis, increasing the complexity of the task.

This paper presents the implementation of the reprocessing application, the certification of results, the tools used, the status at completion of bulk processing, and a conclusion.

* snow@fnal.gov

IMPLEMENTATION

The application work flow was organized as a two step process, with each step corresponding to a grid job. The production step started with raw data in mass storage at FNAL. The data were organized in datasets consisting of approximately one hundred files. A dataset was delivered to each execution site for processing via WAN transport. At each site the job would be parallelized by assigning each file in the dataset to a single batch job. The processing of a job required access to information in the database at FNAL. Each batch job produced an intermediate output file, called a TMB file, which was stored temporarily to a durable location on disk at the execution site. The durable storage is non-permanent storage lasting longer than the life of the grid job. This completed the production phase.

The merging step began after all files in a dataset were successfully processed. The intermediate TMB files stored at the site were merged together in a single job per dataset. The merging step was necessary to produce files with sizes of 1-2GB in order to make efficient use of tape storage and access. The merged TMB files were conveyed to mass storage at FNAL via WAN transport.

The inevitable crashes and failures from the analysis program, network outages, file delivery failures, batch system problems, worker node failures, file system corruption, etc. necessitate recovery procedures. For effective recovery it is critical to know what failed and what succeeded. Bookkeeping of successful jobs and files is needed to assure completion without duplication of events. Bookkeeping of failed jobs and files is needed to trace problems in order to fix bugs and assure efficiency.

SAM-Grid was used to implement this task on distributed systems. Its use provided the bookkeeping capabilities required in addition to a common environment for the analysis program and for common operation tools at all sites. The SAM-Grid project integrates standard grid technologies, such as the Globus Toolkit and Condor-G, for job and information management (JIM) [4, 5] with software developed at Fermilab for data handling, the Sequential Access via Metadata system (SAM) [6, 7].

SAM is a data handling system organized as a set of servers working together to store and retrieve files and associated metadata, including a complete record of the processing which has used the files. SAM is used for all data handling at DØ, moving petabytes per year.

SAM is designed for the following tasks:

- Track locations and comprehensive metadata for each file in the system.
- Provide storage utilities to add a file to a permanent storage location (in a tape library).
- Cache files on local disk for the duration of the requesting job or longer.
- Deliver files on request to systems that are SAM enabled.
- Utilize file location and system information for performance optimizations.
- Track processing information down to the level of per-file delivery and consumption status.

These properties enable SAM to know (a) from which RAW file(s) a given TMB file was created and (b) with which version of which program it was created. By checking (a) duplication of data in merging can be avoided. Using (a) and (b) finds those jobs/files that failed by asking for all RAW files minus those for which TMB's exist. SAM is sufficient do avoid data duplication and to create recovery jobs.

Job submission and distribution was handled via JIM which has these properties:

- JIM guarantees a uniform global interface to the system.
- Software releases are distributed via SAM (no pre-installation).
- All site peculiarities are parametrised in JIM.
- Provides a common software environment for the analysis program.
- All sites run the same scripts.
- Provides site independent methods of job submission.

JIM provides a local XML database at each site that contains information about:

- the definition and status of a grid job
- which batch jobs were created from a grid job
- status of each batch job
- which files were created by each batch job
- detailed error conditions in case of failure

As this information allows identification of errors, JIM's XML database was used to facilitate error recovery.

Processing RAW data requires the use of databases at FNAL. Direct database access from Europe was determined to be too slow. The problem was addressed by using database proxies at execution sites. The proxy communicated with the database servers at FNAL and maintained a data cache. The proxies were installed and tested at most sites and were proven to fix the problem.

CERTIFICATION

A certification process for remote sites was required prior to the site engaging in production. For certification these comparisons were made:

- SAM-Grid production to conventional production on the DØ Farm at FNAL
- SAM-Grid production at each site to production on the DØ Farm at FNAL
- Merged to unmerged TMB's at each site

Typical certification plots are shown in Figures 1 and 2. Figure 1 compares reconstruction done with the SAM-Grid infrastructure at the DØ farm and at Lyon. Figure 2 compares reconstruction done on the DØ farm with SAM-Grid and done conventionally. The agreement is so good that it is difficult to tell that there are two lines being plotted, except in the CPU time plot. Small differences may result from the random distribution of jobs across different CPU types. As a by-product of this procedure significant improvements in the certification of the analysis program were made.

OPERATIONAL TOOLS

A common package of scripts were used at all sites to manage the operation of the reprocessing project. The

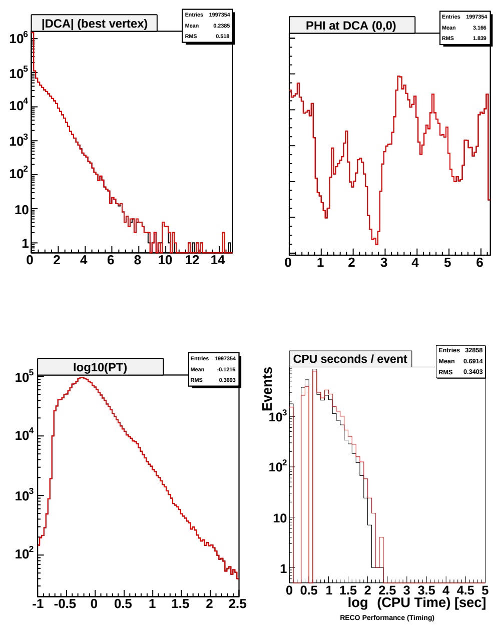

Figure 1: **Comparison of SAM-Grid reconstruction at the DØFarm vs production at Lyon. These plots are the superposition of two lines as seen in the CPU time plot.**

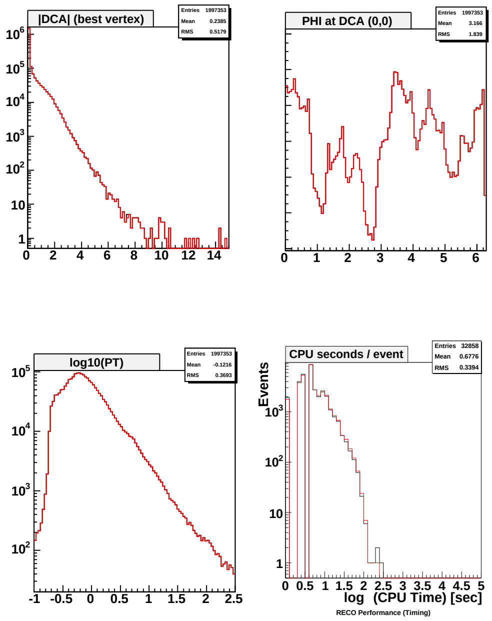

Figure 2: **Comparison of Standard vs SAM-Grid reconstruction at the DØ Farm. These plots are the superposition of two lines as seen in the CPU time plot.**

scripts were used for job submission and recovery for both production and merge job types. The package provided scripts to determine the status of both job types and the ability to manually modify the status of jobs. The typical work flow for dataset reprocessing would be to (1) submit a production job and investigate or retry in case of failures; (2) submit a merge job after production is finished and retry if failed; (3) manually set status to finished in case of unrecoverable failures.

Additional functionality was provided in the script package to investigate the status of all active requests, clean completed and finished datasets' status and submission history files, display the status of all active requests, and suggest to the operator appropriate actions. Suggested actions were to recover production if less than 5% failed; submit merge job if unmerged files exist and the last job was production; optionally approved additional production jobs (one per automatic merge submission).

These suggested actions, which are based on the status of a dataset processing request, combined with the scripts' ability to run commands that implement the suggested actions, enabled an "Autopilot" mode for operating the reprocessing project. The "Autopilot" could be run in a loop with 1 or 2 hours delay. "Autopilot" was built on the experience of reprocessing and significantly reduced the work load of operations. More that 90% of the operational work was to chase and fix failures. Reliable bookkeeping provided by SAM was a prerequisite to implementing these tools.

The demands of the reprocessing project stimulated the development of improved scalability and reliability for SAM-Grid. In particular the developments included implementation of data queues and database access through the use of proxies as noted above. Another concept that was implemented by SAM-Grid was the application-aware grid services, that is to configure different applications with different policies, *e.g.* for the use of the storage or data queues [8].

COMPLETION STATUS

The reprocessing effort started on March 25 and by November 24, 2005 all remote sites finished reprocessing. The number of events completed were 958.7 million out of 986.7 million or 97.2%. The rate of reprocessing started to slow after 5.5 months as sites finished their assigned datasets (see Figure 3). Bulk production was completed on schedule. The production rate began at at about 2.5 million events per day and reached a plateau of 10 million events per day two months later. With efficiency defined as the number of batch jobs that produced a file over the number of started jobs, the average job efficiency was about 90%. The failures were dominated by failures of services. Based on metadata in SAM the rate of unrecoverable failures was 3.0%. Recovery procedures are being undertaken with improved operational tools to recover the remainder. It is expected that about half of the failures can be attributed to missing capabilities of the operational tools and the other half due to bad input data.

Table 1 shows the planned and actual contributions to the reprocessing task. Numbers in parentheses were not part of the original plan. The DØ farm was committed to primary processing in the original plan but was able to contribute using both standard and SAM-Grid methods. Total production on-site at FNAL was 25% and 70% was done off-site. SAM-Grid processed events accounted for 76% of the total. The discrepancy between planned and actual reflects the approximate nature of the original estimate and the alternate final use of facilities.

CONCLUSION

The 2005 data reprocessing effort was fives times bigger than the prior one in 2003/4. A total of 250 TB of raw data was reprocessed and used 16 CPU years. This constituted the largest distributed HEP effort to date and was fully gridified. The project used a common set of operational tools at the 11 participating sites on three continents. Bulk production was done on schedule with the recovery of the remaining 3% ongoing.

The 470pb^{-1} of reprocessed data doubled the dataset available for physics. All of the data available has been processed with the most recent reconstruction code and calibrations.

The reprocessing project stimulated significant improvements in the reliability and scalability (factor of 500 to 1000!) of SAM-Grid as well as improving the reconstruction code certification process. Important to the success of the project was the bookkeeping capabilities of the

Table 1: Planned and actual contributions to reprocessing effort. The planned is shown in CPU-GHz and the actual is percentage of total events.

Site	Planned	Actual
DØFarm (Fermilab), Std:	0	13.75%
SAM-Grid:		8.75%
CMS Farm (Fermilab)	300	2.75%
CCIN2P3 (Lyon)	400	27.0%
Westgrid (Vancouver),	600	26.25%
FZU (Prague)	200	5.75%
GridKa (Karlsruhe)	500	4.25%
UTA (Arlington)	230	3.0%
OSCER (Oklahoma)	(140)	1.5%
Wisconsin	30	1.25%
SPRACE (São Paolo)	(140)	0.75%
UK-RAC (UK)	500	0.5%
External	~3040	

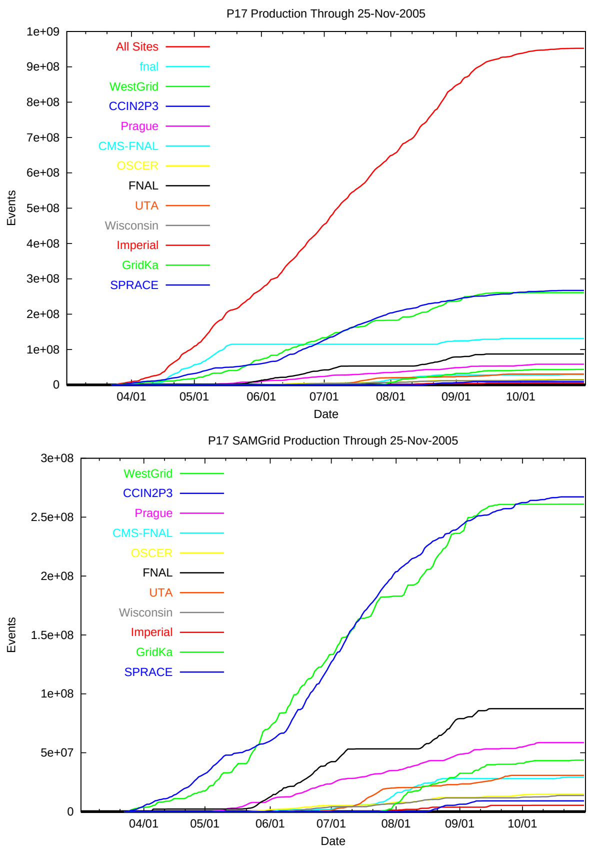

Figure 3: **Total and SAM-Grid processed integrated events.**

SAM/JIM system which allowed for accurate and automated recovery procedures.

For DØ the grid is providing a big return on investment in terms of scalability of operations and advantageous use of resources. While the infrastructure is person power intensive to set up there are significant advantages. These include a common environment for the analysis program, common operational tools, and efficient use of manpower and resources. The flexibility to use non-dedicated resources was exploited by DØ as most of the sites were shared facilities including both OSG and LCG sites. This greatly bolstered the computing power brought to bear. These resources together with the improved scalability of SAM-Grid were crucial in completing the bulk processing on time with a relatively small cadre of operators. The SAM-Grid infrastructure deployed in reprocessing is now being used for Monte Carlo production, which is benefiting from the scalability and reliability advances made to SAM-Grid during the reprocessing effort. Plans are underway to do primary processing of raw data using SAM-Grid in this manner.

FULL AUTHOR LIST AND ACKNOWLEDGEMENTS

J. Snow[1], D. Wicke[2], M. Diesburg[3], G. Garzoglio[3], C. Ay[4], A. Barnovski[3], I. Bertram[5], Y. Coadou[6], G. Davies[7], L. Duflot[8], D. Evans[5], D. Gillberg[6], E. Gregores[9], M. Hildreth[10], V. Hynek[11], R. Illingworth[3], T. Kurča[12], D. Lamb[10], P. Lebrun[12], S. Lietti[9], Z. Liu[6], P. Love[5], P. McGuigan[13], P. Mercadante[9], J. Meyer[14], P. Mhashilkar[3], T. Nunnemann[15], D. O'Neil[6], S. Salih[7], J. Steele[16], T. Stewart[6], F. Villeneuve-Seguier[7], J. Yu[13]

[1]Langston University, Langston, Oklahoma 73050, USA; [2]University of Wuppertal, Wuppertal, Germany; [3]Fermi National Accelerator Laboratory, Batavia, IL 60510, USA; [4]Universität Mainz, Mainz, Germany; [5]Lancaster University, Lancaster, United Kingdom; [6]University of Alberta, Edmonton, Alberta, Canada, Simon Fraser University, Burnaby, British Columbia, Canada, York University, Toronto, Ontario, Canada, and McGill University, Montreal, Quebec, Canada; [7]Imperial College, London, United Kingdom; [8]IN2P3-CNRS, Laboratoire de l'Accélérateur Linéaire, Orsay, France; [9]Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil; [10]University of Notre Dame, Notre Dame, Indiana 46556, USA; [11]Center for Particle Physics, Charles University, Prague, Czech Republic; [12]Republic Institut de Physique Nucléaire de Lyon; IN2P3-CNRS, Université Claude Bernard, Villeurbanne, France; [13]University of Texas, Arlington, Texas 76019, USA; [14]Physikalisches Institut, Universität Bonn, Bonn, Germany; [15]Ludwig-Maximilians-Universität München, München, Germany; [16]Louisiana Tech University, Ruston, Louisiana 71272, USA

This task required the assistance of many beyond those listed, both at Fermilab and at the remote sites, and we thank them for helping make this project the success that it was.

REFERENCES

[1] DØ Reprocessing page: http://www-d0.fnal.gov/computing/reprocessing/p17/

[2] M. Diesburg and D. Wicke, "Data Reprocessing on Worldwide Distributed Systems", in Proceedings of Computing in High Energy and Nuclear Physics (CHEP04), Interlaken, Switzerland, September, 2004.

[3] G. Garzoglio, *et al.*, "Experience Producing Simulated Events for the DZero Experiment on the SAM-Grid", in Proceedings of Computing in High Energy and Nuclear Physics (CHEP04), Interlaken, Switzerland, September, 2004.

[4] I. Terekhov, *et al.*, "Grid Job and Information Management for the FNAL Run II Experiments", in Proceedings of Computing in High Energy and Nuclear Physics (CHEP03), La Jolla, Ca, USA, March 2003.

[5] G. Garzoglio, *et al.*, "The SAM-GRID project: architecture and plan.", in Nuclear Instruments and Methods in Physics Research, Section A, NIMA14225, vol. 502/2-3 pp 423 - 425

[6] I. Terekhov *et al.*, "Meta-Computing at DØ", in Nuclear Instruments and Methods in Physics Research, Section A, NIMA14225, vol. 502/2-3 pp 402 406

[7] I. Terekhov, *et al.*, "Distributed Data Access and Resource Management in the DØ SAM System", in Proceedings of the 10th IEEE International Symposium on High Performance Distributed Computing (HPDC-10), San Francisco, California, Aug. 2001

[8] G. Garzoglio, *et al.*, "A Case for Application-Aware Grid Services", in Proceedings of Computing in High Energy and Nuclear Physics (CHEP06), Mumbai, India, February, 2006.