Distributed data reprocessing

Draft (*Revision* : 1.5) by Daniel Wicke

March 10, 2004

Abstract

The DØ-experiment has rereconstructed all recorded data to create a homogeneous dataset for physics analysis based on an up to date understanding of the detector.

To finalise this effort in time for the winter conferences new computing resources had to be exploited. DØ therefore performed the task in a fully distributed manner utilising 7 computing centres in 6 countries. The diversity of these computing centres posed strong constraints on the software. DØ-tools and Grid-tools were used ....

This document shall describe the various aspects of a distributed reprocessing, discuss the problems that occured and the experiences gained during the process.

1 Introduction

The improved understanding of the DØ detector lead to significant improvements of the physics reconstruction software. Its procedures and parameters and no longer based on design specification, but on the actual detector layout and its performance. To benefit from these new developments the recorded data had to be reprocessed withthisnew recontruction softwaqre in order to create a homogeneous dataset for physics analysis.

At the time when the new software was deployed to the processing famt about **300M (needs check)** events were already recorded by DØ and reconstructed with older software versions. Geven a reconstruction speed of around 50s/Event the estimated need for CPU time was 4 million hours or 5800 months on a 1GHz PIII machine. With a desired completion time of 3 months nearly 2000 CPUs were required.

As the DØ recontruction farm is dimensioned to keep pace with datataking it could only be used during the foreseen shutdown of 8 weeks. With its size equivalent to 1000 1GHz PIII processors this farm was expected to handle only around one third of the events. Thus other resources had to be used to achive a timely completion of tasks.

Such additional resource are available to DØ through the participating institutes. Most of these resources are shared with other experiments. Each of these computing clusters has a different hardware and software configuration.

To expoit these resources the procedure and the software needed to be set up to suit the many different cluster setups.

In the following this document will describe...

2 What existed

2.1 DØ software environment and requirements

All DØ software is managed with the UPS/UPD package managing system [**?**]. This system allows to install several versions of the same package on one system. The required version is selected with special setup commands. Within the UPS/UPD system all software is installed under a special account usually called `products`. The directories holding the actual software installations have to be readable from all compute nodes of a given cluster.

The central piece of software needed for accessing data within DØ is SAM (Sequential Access through Metadata) [**?**]. On fully installed DØ-systems a set of programms, called SAM-station, is running. This SAM-station serves requests for datasets of the applications on that cluster. The SAM-station takes care of any necessary file transport from tape or from a remote disk to the local SAM-cache from which the application will read the file. The files corresponding to the requested datasets are given to the application in an order which minimises the necessary tape mounts and transports.

The DØ reconstruction software, `d0reco`, ... database access ...

To combine the different production steps into a single job, DØ uses a workflow system called Runjob. Before the reprocessing Runjob has been used in MC production for quite some time.

2.2 Configuration of participating clusters

2.2.1 GridKa

The German Grid Computing Centre in Karlsruhe (GridKa) [**?**] is a cluster shared by 8 high energy physics experiments. All nodes share common file areas and have additional local scratch space. For each experiment a so called head node is setup which can hold experment specific services and serves as a login node for the users.

The DØ head node runs a SAM-station which is capable of accessing the local TSM tape robot. Its cache is shared by all nodes.

At the time of the reprocessing GridKa had around 460 CPUs installed with clock frequencies between 1.2 and 2.6GHz. DØ's share on the system was 6%.

2.2.2 Lyon

2.2.3 Nikhef and EDG

2.2.4 SAR

2.2.5 UK: Lancester, London, Manchester

2.2.6 UMICH

2.2.7 Westgrid

3 Planning

The planning for the first DØ reprocessing began in Beaune.

At the end of the workshop a number of required steps were identfied and accommpanied with deadlines needed to achieve a completion by end of November.
* list **

By middle of September none of the required were achieved.

Two reasons: Holiday season. d0reco not available in version without db access. Therefore remote testing wasn't possible.

4 Operation

Data distribution (SAM)
Job distribution (manual, JIM)
Monitoring (manual)
Bookkeeping (Daniel, Lyon, NIKHEF, manual)
Cross-checks (In runjob, sam store, on d0mino)

5 Conclusions

5.1 DØ

In which order should we tackle the problems we have.

5.2 LCG

What would we imagine how things should work in an ideal Grid. (Compared to what exists in LCG)

6 Summary