# The DØ Monte Carlo Challenge

*G.E. Graham and Boaz Klima for the DØ Collaboration*

University of Maryland, College Park, MD, USA

**Abstract**

The DØ Monte Carlo Challenge (MCC) is an effort to test DØ software systems on a large scale in advance of Run II at Fermilab. The DØ GEANT Simulation of the Total Apparatus Response (DØGSTAR) is a GEANT 3.21 based simulation of the DØ detector wrapped in C++ code. Typical events are about 1.5 MB in size and require about 6 minutes of CPU time on an SGI R12000 processor. The goals of the MCC are to produce more than 500K events to be used in testing L3 software, testing reconstruction software, estimating trigger rates, testing the accuracy of the DØGSTAR program itself, and evaluating the physics potential of DØ in Run II. Important infrastructure, such as the Sequential Access Method (SAM) for data storage at Fermilab, and the integration of all these systems is also being tested. It is estimated that about 1 TB of Monte Carlo data will be stored during the monte carlo challenge. The MCC will take advantage of distributed computing resources around the world; including FNAL, Lyon, NIKHEF, Prague, the University of Texas (Arlington), and others.

Keywords: chep,infrastructure,monte carlo,systems

## 1 Introduction

The principal monte carlo simulation of the DØ detector is the DØ GEANT Simulation of the Total Apparatus Response (DØGSTAR). For the Tevatron Run II, DØGSTAR is based upon GEANT 3.21 and is a hybrid program written in Fortran and C++. In order to effectively utilize the higher luminosity of the Run II Tevatron program, new infrastructure and software tools must be deployed in order to effectively simulate the large amount of data expected in Run II. Our goal for Run II is to increase our capacity for generating, storing, and accessing large quantities of monte carlo data. These improvements will in turn affect how DØ is able to deal with the large amounts of data expected in Run II. The solution outlined here is a distributed processing and storage environment that takes advantage of resources physically located around the world to generate large numbers of monte carlo events.

## 2 The DØ Monte Carlo Challenge

The DØ Monte Carlo Challenge (MCC) is an effort to address the needs of DØ monte carlo generation in Run II and to test crucial infrastructure needed by DØ.The bedrock software tools include the monte carlo program (DØGSTAR) and the reconstruction program (DØRECO). The large quantities of events produced in the DØ monte carlo challenge will speed development of the reconstruction program through the ability to run on large quantities of realistic events. Also, the large monte carlo event sample will allow simulation of trigger rates and efficiencies in advance of the run. During this effort; our capabilities to generate monte carlo, store large amounts of monte carlo data, read the produced monte carlo data, perform full reconstruction of the events, and the integration of these systems will be tested.

Large scale event generation is expected to take place outside of Fermilab at remote processing sites around the world. Data is then transferred to a central repository at FNAL[1]. Infrastructure for generating monte carlo at distributed sites is being developed. The remote sites consist of either SGI machines running IRIX or Intel based machines running Linux.

Other important hardware infrastructure tested by the D∅ monte carlo challenge includes the primary event storage facility at D∅ : the Sequential Access Method (SAM)[1, 2]. SAM is a data management system maintained at FNAL that stores and catalogs large amounts of data. The stored data files can then be grouped and accessed according to different criteria; including event type, data stream, or conditions of generation. Although SAM has the ability to manage diverse storage media, the monte carlo challenge will mainly take advantage of a tape robot storage system using Mammoth-I tapes.

After generation, the reconstruction program D∅RECO will be run on the set of generated events on a PC Farm[4, 5] at Fermilab. The resulting reconstructed data is then also stored in SAM. The PC Farm system consists of a network of Intel based PCs running a special FNAL edition Linux[3] OS.

The monte carlo challenge was designed to be completed in stages reflecting the evolution of the software tools and hardware infrastructure. Phase I of the monte carlo challenge used a basic version of D∅GSTAR and was completed last year. About 100K of various events were generated on a 12 CPU SGI R10000 based machine running at 195 MHz located at FNAL. Generation was completed by January 1999, and reconstruction with a basic version of D∅RECO was completed by June 1999. Phase I focussed primarily on the infrastructure needed to run D∅GSTAR and D∅RECO, and on the accuracy and scope of the two programs.

Phase II of the monte carlo challenge utilizes an improved and fully functional version of the D∅GSTAR program. The goals of Phase II include those of Phase I and in addition the goal of estimating the physics performance of the D∅ detector. The sample generated so far in Phase II includes a wide variety of physics events which can be used to test many aspects of the reconstruction program; to place the integrated systems of generation, storage, and access under a stress test; and to estimate the physics potential of D∅. These are shown in table 2. Phase II reached initial monte carlo production goals of 500K events generated with D∅GSTAR in January 2000, and the reconstruction with an improved D∅RECO will begin soon after. The physics events are generated using the Pythia monte carlo program, version 6.125.

Phase III will begin in June 2000 and incorporate more improvements. The D∅GSTAR will be further optimized and the D∅RECO program will be upgraded both in optimization and performance. The goals of Phase III will focus primarily on the physics performance of the D∅ detector in Run II and its trigger system.

## 3 The D∅GSTAR Program

D∅GSTAR is a Fortran based monte carlo simulation of the D∅ detector based on GEANT 3.21 with C++ wrappers to allow D∅GSTAR to run in the D∅ Run II software environment. It is supported on the IRIX 6 platform and the Linux[2] platform. D∅GSTAR models the entire D∅ detector including the silicon vertex detector, liquid Argon calorimeter, central fiber tracker, the preshower detectors,

[1]It is also possible that data will be warehoused at the remote sites. In this case, the data are still cataloged at the FNAL central storage site.

[2]The distribution is Fermi RedHat 5.2 running kernel version 2.0.X, soon to be updated to Fermi RedHat 6.X running kernel version 2.2.X.

**Table I:** The production goals in numbers of events of phase II of the D∅ monte carlo challenge broken down by physics process. All physics events are generated using the Pythia monte carlo program version 6.125 and subsequently processed with D∅GSTAR. The total number of events will be somewhat higher than 500000 since we have been able to exceed our goals in a number of areas.

| Process | Limits | Goal (# events) |
|---|---|---|
| QCD Dijet | $E_T > 2, 5, 10, 20, 40, 80, 160$ GeV | 50000 each of 7 bins |
| $t\bar{t}$ | | 50000 |
| $bJ/\Psi$ | $J/\Psi \rightarrow e^+e^-, \mu^+\mu^-$ | 50000 |
| $Z \rightarrow e^+e^-, \mu^+\mu^-, \tau^+\tau^-$ | | 10000 each of 3 processes |
| $Z \rightarrow b\bar{b}$ | | 5000 |
| $W \rightarrow e\nu, \mu\nu, \tau\nu$ | | 1000 each of 3 processes |
| $\Upsilon \rightarrow e^+e^-$ | | 1000 |
| Other | | 11000 |
| Totals | | 500000 |

and the muon system. Typical processing times on an SGI R12000 processor running at 300 MHz are shown in table II along with typical event sizes. The D∅GSTAR program has a 350 MB virtual

**Table II:** The event generation times and sizes after compression ($\simeq 50\%$) for typical processes are shown. An Poisson distributed additional number of minimum bias events (Pythia Dijet events) are added to each event. The times shown are for an SGI R12000 processor running at 300 MHz.

| Process | $N_{\mathrm{minbi}}$ | Size (Mb per event) | Time (Minutes per event) |
|---|---|---|---|
| $Z \rightarrow e^+e^-$ | 0 | 0.5 | 2 |
| $t\bar{t}$ | 0 | 1.0 | 3 |
| $Z \rightarrow e^+e^-$ | 1 | 1.0 | 6 |
| $t\bar{t}$ | 1 | 1.5 | 7 |

memory footprint [3] on the SGI platforms.

## 4 Monte Carlo Production and Storage

Monte carlo production is distributed among several production sites around the world. The hardware utilized in the monte carlo challenge falls into two categories : machines based on SGI processors R10000/R12000 and Intel chip based PC hardware. The current monte carlo challenge processing sites include FNAL; the Institute of Nuclear and Particle Physics in Lyon, France (IN2P3); NIKHEF in the Netherlands, the Institute of Physics of the Academy of Sciences of the Czech Republic (FzU) ; and the University of Texas at Arlington (UTA).

The production site at FNAL currently uses a 48 processor Origin 2000[4] with 250 Mb of memory per processor. These processors are SGI R12000 running at 300 MHz. The platform is IRIX 6.5. Since this site is designated as multi-use, monte carlo production has been restricted to a maximum of 17 concurrently running monte carlo challenge jobs at any one time. Files generated at FNAL are imported directly into SAM.

[3] This figure is cited for the SGI platform. For reasons not fully explored, the memory footprint on the Linux platform is measured to be somewhat smaller.

[4] to be expanded to 64 processors

The hardware at the Lyon (IN2P3) site includes a PC farm system consisting of 15 dual Pentium processor PCs. Roughly half are Pentium II based at 450 MHz and the rest are Pentium III based running at 500 MHz. There are 256 MB per processor. The platform is the version 5.2 of Fermilab RedHat Linux, and data is relayed back to FNAL via ftp connection into a special SAM import area. The number of nodes will be increased in the future.

The production site at the NIKHEF Institute (and the NCF computing facility) in the Netherlands uses R10000 processors in an Origin 2000 machine with 128 processors and 450 Mb of memory per processor. These processors run at 250 MHz. The platform is IRIX 6.5. The current setup at NIKHEF is limited to 6 concurrently running jobs at any one time; although this will be expanded in the future. Finished data files are transferred to FNAL over the network to a special SAM import area via ftp. The measured performance in transferring files is 4 Mb/s on average.

The hardware at the Prague site includes 3 PCs based on the Pentium III chip. Interestingly, the Prague site is itself distributed; consisting of a 500 MHz machine with 128 Mb of memory located at FNAL, a 500 MHz machine with 256 Mb of memory located at MFF UK (Charles University), and a dual CPU machine with 450 MHz processors and 128 Mb of memory at FzU in Prague. Each PC runs version 5.2 of the Fermilab RedHat Linux OS. There are no restrictions on usage; nor is there a batch system. The Prague site also relays data back to FNAL via ftp network connection with an average performance of about 100 Kb/sec.

The hardware at the UTA site includes 7 PCs which are each dual Pentium III processors running at 500 MHz with Linux kernels supporting SMP. There is 250 Mb total memory per processor. The platform is version 5.2 of the Fermilab RedHat Linux OS with SMP support compiled into the kernel. The UTA site can relay data back to FNAL via ftp network connection; however, UTA also has a single Mammoth-I tape drive that can be used to write data directly in SAM format. Such tapes can then easily be added into the tape robot administered by SAM for the monte carlo challenge.

The PC Farm at FNAL[4, 5], currently used to run the D∅ reconstruction, may also be used in the future to generate monte carlo. The farm system has 50 Dual Pentium III machines running at 500 MHz. Each processor has 256 MB of memory.

Other production sites are planned in the future; for example, at Lancaster in the UK, at Nijmegen, etc. The site at Lancaster will be a similar PC farm site with 120 worker nodes with two processors each.

The storage system for all of the monte carlo produced is a tape robot with Mammoth-I tapes administered through SAM. Each Mammoth-I tape holds 18 GB of data. SAM is configurable so that our storage options may change in the future in response to storage needs.

## 5 Results

As of January 2000, more than 500K events have been generated around the world for the D∅ monte carlo challenge. This data has been successfully generated in a distributed environment and collected in the FNAL central storage site, SAM. Table III shows the results of production broken down by production center. The monte carlo challenge has thus far been a success, and many groups in D∅ are already processing the data from Phase II. We have shown that large monte carlo data samples can be generated at distributed sites and managed effectively. Finally, SAM was exercised and has so far stored about 75% of the data from monte carlo challenge Phase II equal to about 0.6 TB of data. This data is currently being used as input to the D∅RECO program, and will also be used in upcoming trigger studies.

**Table III:** The number of monte carlo challenge events generated so far broken down by production site.

| Center | Number Generated |
|---|---|
| FNAL | 240000 |
| Lyon | 210000 |
| NIKHEF | 30000 |
| Prague | 20000 |
| UTA | (Just Starting) |
| Total | 500000 |

## References


1. L. Lueking, “The Data Access Layer for D∅ Run II: Design and Features of SAM”, CHEP 2000, Padua, Italy, Winter 2000
2. D. Petravick, et al.“Fermilab Computing Division Systems for Scientific Data Storage and Movement”, CHEP 2000, Padua, Italy, Winter 2000
3. C. Sieh, et al. “Linux Support at Fermilab”, CHEP’98, Chicago, USA, Autumn 1998.
4. H. Schellman, et al. “Large Scale Test of D∅ Reconstruction Farm”, CHEP 2000, Padua, Italy, Winter 2000.
5. J. Fromm, et al. “Processing Farms Plans for CDF and D∅ for Run II”, CHEP’98, Chicago, USA, Autumn 1998.