BABAR MC AND Data Analysis Center

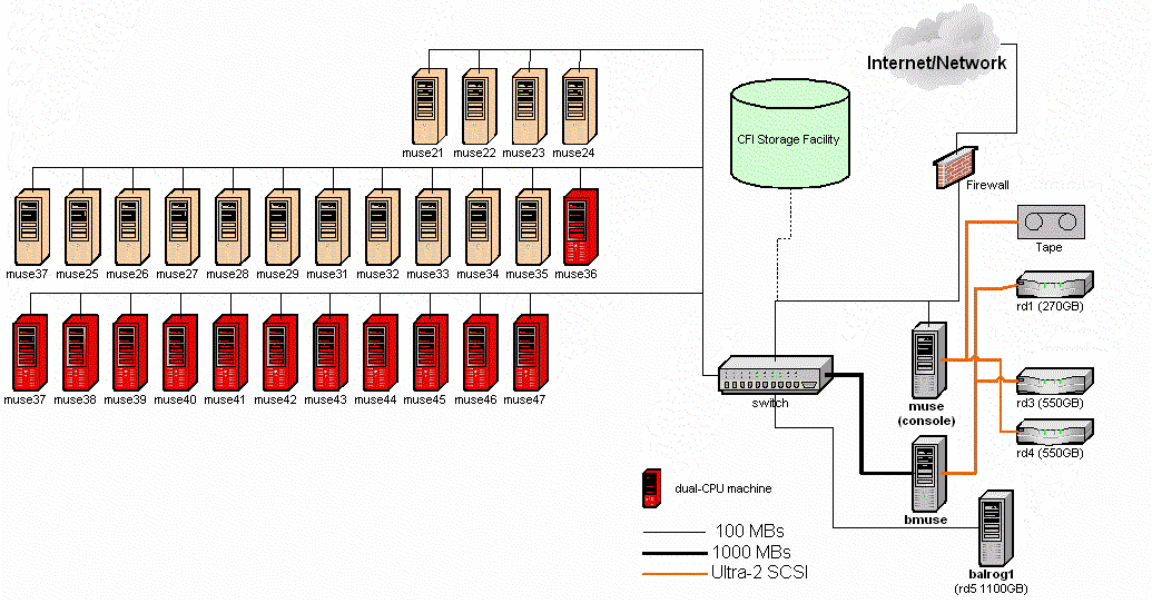
by
Ashok Agarwal
University of Victoria, Victoria, BC, Canada

- Details of Linux cluster for the BaBar MC Production and Data Analysis Center.
- Highlights of the Monte Carlo (MC) production.
- Setup of the BaBar Data Analysis Center.
- Summary and future plans.

DETAILS OF THE LINUX CLUSTER FOR BABAR MC PRODUCTION AND DATA ANALYSIS CENTER

- Consists of one console (muse) 450 Mhz PIII dual processor and one file server (bmuse) 450 MHz PIII dual processor.
- 12 high speed dual CPU nodes mixture of both Athlon and Intel machines.
  - 4 Intel nodes : 1.0 GHz Pentium III, 768 MB RAM
  - 8 Athlon nodes : 1.5 GHz MP, 768 MB RAM
- 2 SCSI Raid-5 towers : rd1 (270 GB) and rd3 (550 GB) attached to the file server (bmuse).
- 1 Raid tower and 1 file server : rd4 (550 GB) and rd5 (1000 GB) attached to the console (muse).
- Linux cluster is protected by putting firewall between the outside world and the console (muse).
- Condor batch system is used to submit jobs on Linux cluster.

Babar- The University of Victoria Beowulf Cluster for Data Analysis and MC Production



muse 36-39	Dual processor 1Ghz Pentium III, 768 MB RAM
muse 40-46	Dual Processor 1.5Ghz Athlon MP, 768 MB RAM
switch	HP Procurve 4000M with 7 8x100 modules, 1 1000 module
rd1,rd2	SCSI Raid-5 Tower, 270GB
rd3,rd4	SCSI Raid-5 Tower, 550GB

muse	console machine. dual processor 450MHz PIII with 512MB of RAM
bmuse	file server. dual processor 450MHz PIII with 512MB of RAM
balrog1	file server. dual processor 1.3GHz Athlon MP with 512MB of RAM

Diagram by Jan Van Uytven
(wyvern@uvic.ca)

HIGHLIGHTS OF THE MONTE CARLO PRODUCTION

- What is Monte Carlo (MC) Simulation ?
- Requirements for installation of BABAR software.
- Setup of Directory structure as at SLAC.
- Recipe to produce MC events.

MONTE CARLO SIMULATION

Consists of following 3 steps:

- BgsApp – constructs the GHits using the conditions/configurations databases from a snapshot.
    - Conditions databases: calibration constant, alignments etc.
    - Configurations databases: trigger configurations, chip threshold settings etc.
    - Snapshot: conditions/configurations files
- SimApp – digitises raw information in GHits and constructs the Digi Objects. Background triggers may be mixed to make Digi Objects.
- Bear – runs the reconstruction subsystem modules (SVT, DCH, EMC, DIRC, IFR) using Digi Objects.

REQUIREMENTS FOR BABAR SOFTWARE AT REMOTE SITE

- Install Andrew File System (AFS) as client.
- Set up CVS repository at SLAC for check-ins and check-outs over AFS.
- Install Objectivity version 6.1 for Linux.
- Install Rogue Wave Software for Linux.
- Set up Cern libraries, binaries, include files.
- Install root software.
- Set up BABAR software releases, libraries, and binaries.

SETUP OF DIRECTORY STRUCTURE AS AT SLAC

Environment Variable	Typical Value	Contains
BFROOT	/homes/babar	The top of the BaBar directory tree
	$BFROOT/bin	Scripts common to all releases. Can be left empty.
	$BFROOT/bin.<arch>	Binaries for specific architecture common to all releases. The BaBar version of CVS is typically installed here.
	$BFROOT/man	Manual pages.
	$BFROOT/package	3rd party software specific to BaBar like Rogue Wave tools and Objectivity.
BFOVERRRIDE	$BFROOT/etc	Local overrides for the position of libraries etc.
BFDIST	$BFROOT/dist	The software releases.
	$BFDIST/packages	The checked out source for the installed releases.
	$BFDIST/releases	The libraries and binaries for the installed releases.

RECIPE TO PRODUCE MC EVENTS

- Select the release directory (10.3.0e).

```
newrel  -s  [dir]  -t  10.3.0e  t1030e
```

- Change to release directory.

```
cd  t1030e
```

- Set up correct path and variables.

```
srtpath  10.3.0e   Linux2
```

- Obtain the conditions/configurations databases and background triggers.
  - Create .bbobjy file.

```
FD_NUMBER = 7878
```

  - Create your own test Federation.

```
setboot
```

- Build your executable.
  - Check out all of the packages.

```
addpkg  BgsApp
addpkg  SimApp
addpkg  EvtMix
addpkg  Bear
```

- Make lib and bin for all.

```
gmake lib
gmake BgsApp.bin
gmake SimApp.bin
gmake Bear.bin
```

- Set up work directory.

```
addpkg  workdir
gmake  workdir.setup
cd  workdir
```

- Set the run time variables.

```
setenv  RUNNUM  200000
setenv  UDECAY  tau_generic.dec
setenv  CONDALIAS  Feb2002
setenv  NEVENT 2000
```

- Run the executables.

```
BgsApp../BgsApp/BgsUserExampleEvtGen.tcl > b.log
```

where Xx is one of: EvtGen, Koralb, Bhwide or Bkqed

```
SimAppApp  ../SimApp/SimProduction.tcl  >  s.log
BearApp  /Bear/BearProduction.tcl  >  b2.log
```

- Transfer databases to SLAC using MocaEspresso.

SETUP OF BABAR DATA ANALYSIS CENTER

- Mirror the ORACLE database at SLAC using the relatively simple and free software MYSQL.
- ORACLE database has meta information about the Objy and Kanga files.
- Mirrored database at remote sites have an additional "import-status" flag.
  - 2 means files are yet "to be imported" (default).
  - 1 means files are "to be imported."
  - 0 means files are on "local disk."
- Install following 4 Perl scripts to maintain and update local ORACLE database.
  - skimSqlMirror : for updating database.
  - skimSqlSelect : for selecting stream.
  - skimImport : for importing data from SLAC.
  - skimDelete : for deleting stream on local disk.

RECIPE TO RUN OBJY OR KANGA ANALYSIS

- Prepare the input tcl file using the skimDataRemote command.

```
skimDataRemote –stream < Desired_skim  /Stream > -g
<Desired_runs>  50k  -- tcl
```

- Produce the executable for running your application.
  - Select the release directory (10.4.0a-physics-1).

```
newrel -s [dir] -t 10.4.0a-physics-1   t1040a
```

  - Change to release directory.

```
cd  t1040a
```

  - Set up correct path and variables.

```
srtpath 10.4.0a Linux2
```

  - Add your analysis package.

```
addpkg BetaUser
```

  - Add workdir package.

```
addpkg workdir

gmake workdir.setup
```

  - Compile and link.

```
gmake lib

gmake BetaUser.bin
```

- Set up the desired environment variables.
- Run the executable using the input tcl file to get either hbook or root file.

```
cd workdir

BetaApp kanga.tcl >  kanga.log
```

SUMMARY AND FUTURE PLANS

- Both Monte Carlo Production and Data Analysis Center are running successfully.
- Approximately 1.6 Mevents/month (800 GB) is exported to SLAC.
- More dual CPU nodes and storage disk space will be added in the Linux cluster to increase the production.
- MC production is moving from 3 executables (BgsApp, SimApp, BearApp) to single executable called MooseApp.
- Grid-based MC production will start soon as BaBar grid is being setup at SLAC using Globus tools.
- Once CFI storage facility is up and running, it will be utilized by the data analysis center to store large Objy and Kanga files.