BaBar Analysis and Computing Model & Long Term Analysis Plans

H. Neal
BaBar Computing Coordinator
SLAC *National Accelerator Laboratory*

January 26th , 2009
for the
Data Preservation and Long Term Analysis Workshop @ *DESY* (Hamburg, DE)

The BaBar Data

- 22 billion events (~540/fb) recorded
- ~9 billion events reconstructed
- ~10 billion events simulated with latest reconstruction code ... so far

BaBar Analysis Environment

- analysis work distributed among 13 independent Analysis Working Groups hosted at five TierA sites

Tier-A	AWG
SLAC	AllEventsSkim, Breco, LeptBC, PID, SemiLep, Tracking, TauQED, Y({n<>4}S)
Bologna	AllEventsSkim, Charm
CCIN2P3	AllEventsSkim, Charmonium, PartSpec, RadPenguin, ChlsTwoBody, ChlsThreeBody, Quasi2Body, LeptBC
GridKa	TDBC
UVic	Tau

- >150 analyses currently in progress

Analysis Activity

- level of analysis work NOT foreseen to decrease until after 2010
  - ~120 analyses possible
- will need to have centralized services until end of 2012 to support
  - ~70 analyses depending on available manpower
- archival system only after 2012 – will need to support ~35 potential analyses

Collaboration with sister experiments

- until now there has been little collaboration with other experiments but we have now entered a phase where inter-collaboration collaboration is being seriously considered

The Event Store

- analysts can either use the output of reconstruction directly or from skims which are streams of data allowing access to specific classes of events needed by certain analyses. There is an association between skims, analysis working groups and where they are hosted.

- the skims inflate the data store but greatly increase the efficiency of access to the data and reduce latencies from competition for accessing the same bits of data

Components of the BaBar data

- micro – all quantities needed by most analyses
- mini – detailed hit and cluster information
- skims can either just reference the relevant events in the reconstruction output or include the micro and in some cases include the mini
- information about any quality issues with the direct reconstruction output is propagated to the skims in the bookkeeping database

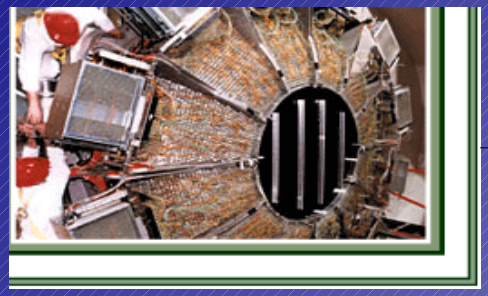
Event Sizes

- avg. over BBbar, ccbar, uds, tau, μ-pairs and Bhabha evts

simulation data	Kbytes/event
micro	5,67
mini	9,67
skims	21,12

reconstructed detector data	
micro	3,07
mini	7,05
skims	8,87

Production Chain



Prompt Calibration

SLAC

Event Reconstruction

SLAC & Padova

background events

Simulation Production

@ 20 sites

Skim Production

SLAC, GridKa, UK

skims are distributed to TierA sites based on the AWGs being hosted at the site

Calibrations

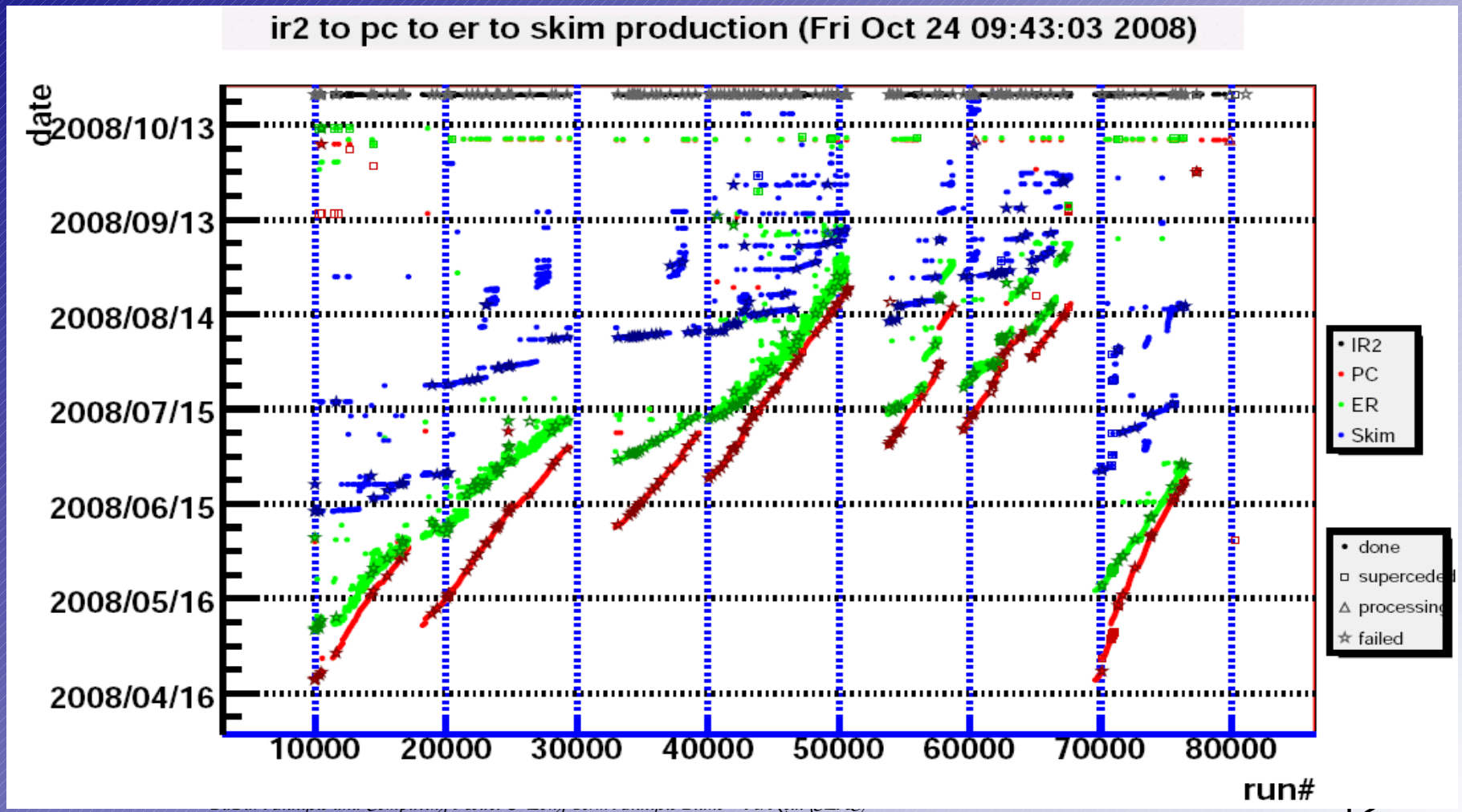
- **manual calibrations**
  - source calibrations
  - alignment calibrations
- **rolling calibrations**
  - boost, energy, IP position, timing, etc....
  - determined by a prompt calibration pass before event reconstruction

Databases

- production databases (simulation and skim)
- physics databases (documents and supporting materials)
- detector run database
- reconstruction database
- bookkeeping database
- conditions databases

All of this needs to be consolidated for the archival system

BaBar recent production cycle



BaBar's Simulation Production Distributed over 20 sites

- manager at SLAC allocates production requests to the sites after validating the site with the needed code release and conditions database
- the simulated data from the various sites is exported to SLAC
- recorded on the HPSS tape system @ SLAC and registered in the data bookkeeping database
- the simulated data is then exported to ccin2p3 and as requested to skim production sites

ALL:

Week Beginning	TOTAL	caltech	ccin2p3	cu-boulder	fzk	gridx1	infn	infngrid	osu	ral	slac	slac2	tud	udo	uofl	utd	utenn	uvic2	westgrid
21-9-2008	191970	0	8488	0	0	0	2218	744	832	3720	86896	63608	2648	3808	0	6424	0	12536	48
14-9-2008	846163	6889	168816	20216	4736	0	6312	1144	14016	126496	206264	221994	4288	15208	0	8352	0	33528	7904
7-9-2008	796994	11898	165760	7352	17392	0	3736	7006	9168	77976	217416	219392	2000	14418	0	1176	0	35920	6384

Last Modified: Last Modified: Tue Sep 23 02:18:38 PDT - slightly behind last week (some loss due to validations)

Projected Production Output

2009	2010	2011	2012
460 Tbytes	208 Tbytes	60 Tbytes	33 Tbytes

- Mostly simulation and skim production output
- Some detector data reconstruction output in 2009 & 2010 only

Platforms Currently in Use @ SLAC

350	RH Enterprise Linux 3 (32-bit kernel), GCC 3.2.3, glibc-2.3.2	Sun Fire V20z dual dual-core 2.0GHz Opteron 270 CPUs, 4GB memory
78	RH Enterprise Linux 4 (64-bit kernel), GCC 3.4.6, glibc-2.3.4	Sun Fire X2200M2 dual dual-core 2.6GHz Opteron 2218 CPUs, 8GB memory
156	RH Enterprise Linux 3 (32-bit kernel), GCC 3.2.3, glibc-2.3.2 RH Enterprise Linux 4 (64-bit kernel), GCC 3.4.6, glibc-2.3.4	Sun Fire X4100 dual dual-core 2.2GHz Opteron 275 CPUs, 4GB memory
252	RH Enterprise Linux 4 (64-bit kernel), GCC 3.4.6, glibc-2.3.4	Sun Fire X2200M2 dual dual-core 2.6GHz Opteron 2218 CPUs, 8GB memory
126	RH Enterprise Linux 4 (64-bit kernel), GCC 3.4.6, glibc-2.3.4	Dell Poweredge 1950 dual quad-core 2.66GHz Xeon CPUs, 16GB memory

Current Planning for Long Term Data Access

- **Long Term Data Access**
  - major effort to ensure that BaBar data will be accessible and useable into the far future
  - effort now will profit from existing expertise and reduce headaches later
  - aim towards a simple cost effective system

Work towards the Archival System

- ***MIGRATION***: porting of the BaBar code to a more current platform so as to insure that the platform will be one that will continue to be supported in terms of security updates and other patches until the end of the steady analysis period

- ***SIMPLIFICATION***: reducing dependence on third party software, switching from oracle to mysql, all data on disk.

- ***VIRTUALIZATION***: everything running on one or a few many core machines with a large disk attached and the BaBar operating environment protected by a virtualization container that will allow the code to continue to run on future platforms.

What's been achieved so far

- a task force has been formed
- they have succeeded in getting the BaBar code to work with ROOT 5.20 (was ROOT 5.14). The data store access has been demonstrated to work with ROOT 5.20 but we have not proved that it will work at the production level.
- they have succeeded in migrating the BaBar code to work with SL5. Until now, all production has used SL3 32bit builds but we also build our releases on SL4, Solaris5.8 and Solaris5.10.
- working on making code releases now with the updates and then we will test extensively
- first tests have started with using RedHat Virtualization

Performance Concerns

- per core performance not foreseen to greatly increase over the coming years
- #cores per box is expected to grow
- using mostly 8 and 4 core AMD and Intel boxes now
- expect > 100 core boxes to be normal by the time we prepare our final archival system
- i/o is a concern – how to efficiently pipe data from a multitude of cores in one box at a similar rate as a multitude of boxes each with there own internet connection
- increasing interest in parallel computing

Mass Storage

- Currently BaBar is using old silos/tapes at SLAC that will soon become unsupported and unsupportable. The old tapes are also costly compared to the newer T10000 tapes.
- 2 petabytes of essential data being carried forward
- 460 terabytes of this will be from new production this year
- Eventually only the newest release of data will be kept
  - total amount of disk storage needed for the far future will be ~2 petabytes
- currently we use 450 Tbytes of disk space at SLAC for storage of the active production output – it is a cyclic buffer

  +

  230 Tbytes for storage of user data

  total among the sites for prod + user storage is ~1.8 petabytes

The Future of the Collaboration

- the collaboration is foreseen to have centralized management until the end of 2012
  - data taking concluded April 7th , 2008
  - after 2012 there will still be important analyses to be performed and for that there will be the archival system
  - for the expected scientific output, see the analysis activity slide (#3)
- in 2012 we project still having > 100 physicists active in BaBar
- we are open to collaboration between experiments and ultimately open access to the data
  - initial discussions occurring now concerning the possibility of a B-factory legacy book that would result from collaboration with Belle