

CDF DB Project Status

Dmitry Litvintsev

January 28, 2003

Weekly Computing Division Project Status Reports



Introduction

➤ Project Leader

CDF DB Coordinator Alan Sill (Texas Tech University)
Deputy DB Coordinator Dmitry Litvintsev (CD/CDF)

➤ Applications and Coordinators

- Hardware Bill Badgett (PPD/CDF)
- RunConfiguration Bill Badgett (PPD/CDF)
- Trigger Tom Wright (University of Michigan)
Donatella Torretta (PPD/CDF)
- Calibrations Jack Cranshaw (Texas Tech University)
- Slow Control Margherita Vittone (CD)
Larry Kirsch (Brandeis University)
- Data File Catalog Dmitry Litvintsev (CD/CDF)
- SAM Jeff Tseng (CD/CDF Guest Scientist)

➤ Application is a relational database and API (C++ or Java) that allows to manipulate the contents of the database.



Project List

- Freeware DBMS
 - DB Monitoring
 - DB metering
 - Code Reviews
 - Schema Reviews
 - DBManager based API development for Trigger, Hardware and RunConfigurations
 - Codegen Rewrite
 - Calibration Task List

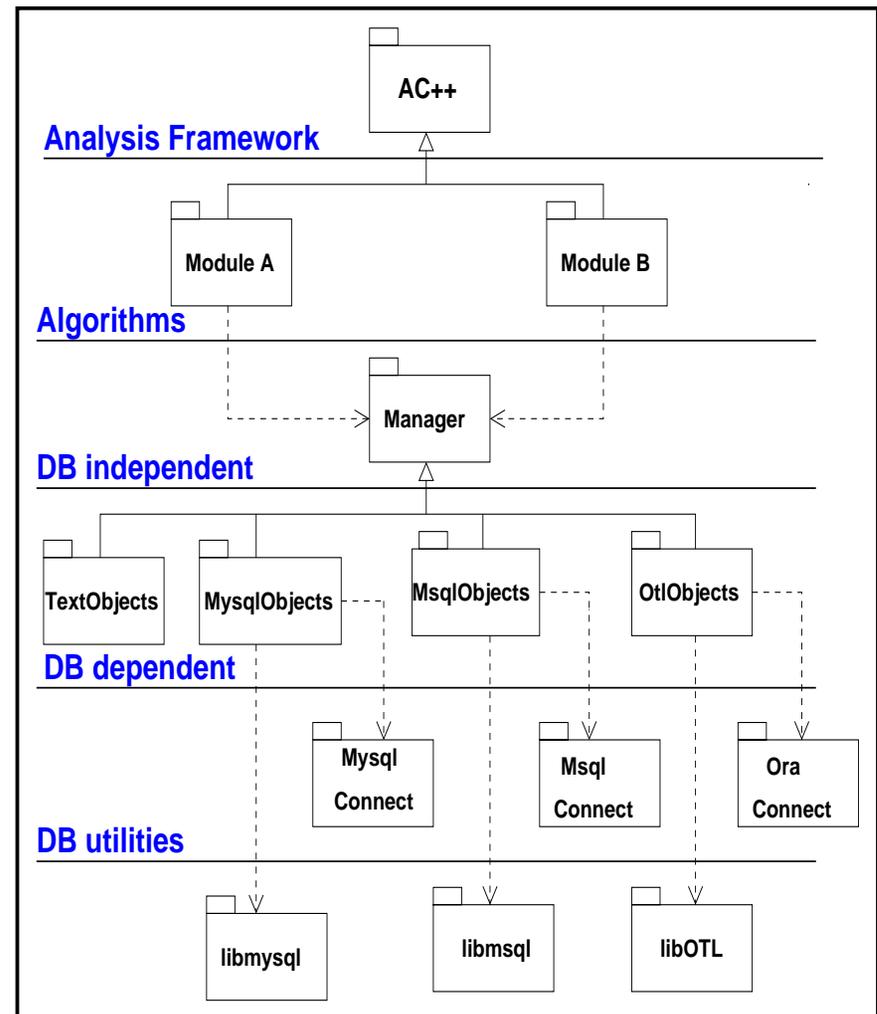
 - CDF SAM

 - DB Browser
- R.Hughes (OSU), D.Waters (UCL),
S.Lebedeva (Fermilab), J.Kowalkowski (Fermilab)
Y.Guo (Fermilab), J.Kowalkowski (Fermilab)
E.Wicklund (Fermilab)
D.Box (Fermilab), N.Stanfield (Fermilab)
A.Kreymer (Fermilab)
J.Kowalkowski (Fermilab), D.Box (Fermilab)
plus experienced CDF programmers
A.Kumar (Fermilab), N.Stanfield (Fermilab)
J.Trumbo (Fermilab)
CDF Helsinki group, overall up to 7 FTE
- D.Box (Fermilab), J.Cranshaw (TTU)
D.Box (Fermilab), J.Cranshaw (TTU)
Y.Guo (Fermilab), J.Kowalkowski (Fermilab)
L.Sexton-Kennedy (Fermilab)
R.Herber (Fermilab), A.Kreymer (Fermilab)
R.Kennedy (Fermilab), D.Litvintsev (Fermilab)
J.Tseng (Fermilab)
people from Glasgow, Oxford, UCL
TTU, Karlsruhe University, Rutgers University
R.Herber (Fermilab)



DB Access

- data stored in DBs are available to reconstruction and analysis algorithms via DB independent database management layer designed by Jim Kowalkowski (DBManager)
- DBManager provides two APIs:
 - back-end transient to persistent mapping API, *IoPackage*
 - template based front-end *Manager<OBJ,KEY>* that provides common put/get methods on transient objects.
- transient object instances can be cached by key value to configurable depth
- transient object definitions, *Mapper*, *Manager<>* and *Handle<>* classes are generated from simple Java description of persistent data using *CodeGen* suite.
- statically linked executables contain a lot of unused code (code bloat)
- problem of keeping *Mapper* objects in sync for all DB implementations



CDF code accesses Oracle servers directly



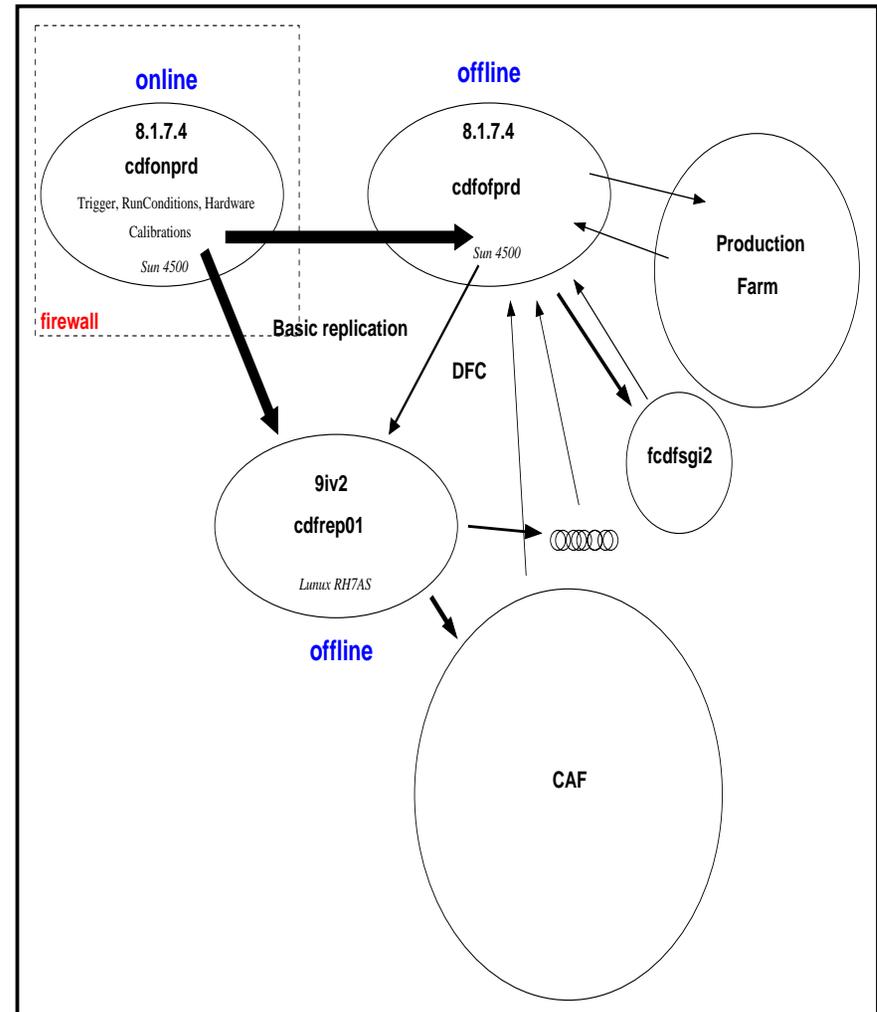
DB Access

- in principle we are satisfied with DB access API
- existing API allows to change underlying DB access mechanism w/o any changes to client code.
- this opens a window of opportunities
 - ☞ should CDF become interested in middle tier DB access it can be accommodated by providing corresponding back-end
 - ☞ CDF is interested in substituting multiple freeware back-ends (mSQL or mySQL) with single interface to ODBC
 - ☞ we want to retain OTL layer to continue to rely on Oracle specific features like triggers



Oracle Servers

- 1st Linux/Oracle install on fcd-flnx1, RH7AS machine. September, 2002
- currently contents of on-line db server are replicated using basic Oracle replication (except Slow Control) to *cdfofprd* and *cdfrep01*
- off-line db server serves CDF Production Farms
- all other read-only users are directed to off-line replica *cdfrep01* with fail-over to *cdfofprd* as specified in *tnsnames.ora*
- write access (to DFC) on *cdfofprd*
- off-line Production db is unloaded
- issues with Basic Replication:
 - limited scalability, since cannot replicate replicated DBs.
 - Firewall issues (*cdfonprd* is behind a firewall)
 - DDL modifications do not propagate automatically
- Oracle 9i addresses these issues by introducing replication using Oracle Streams
- *cdfondev* and *cdfonint* migrated to Oracle 9.2.0.2 in October 2002



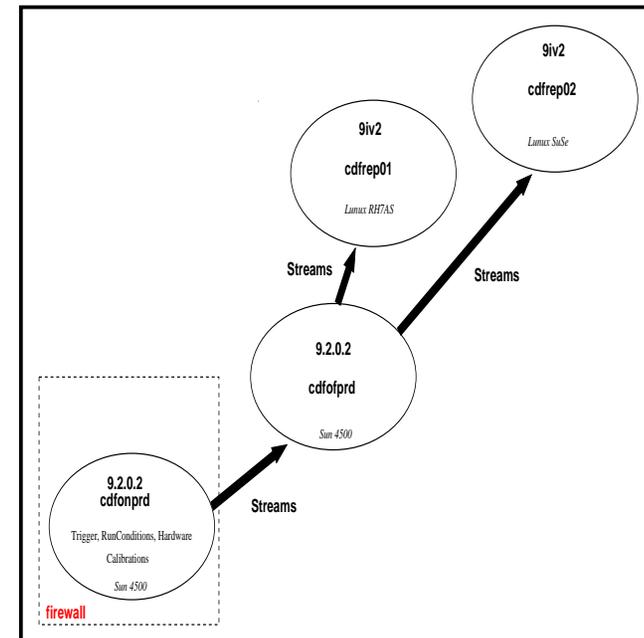
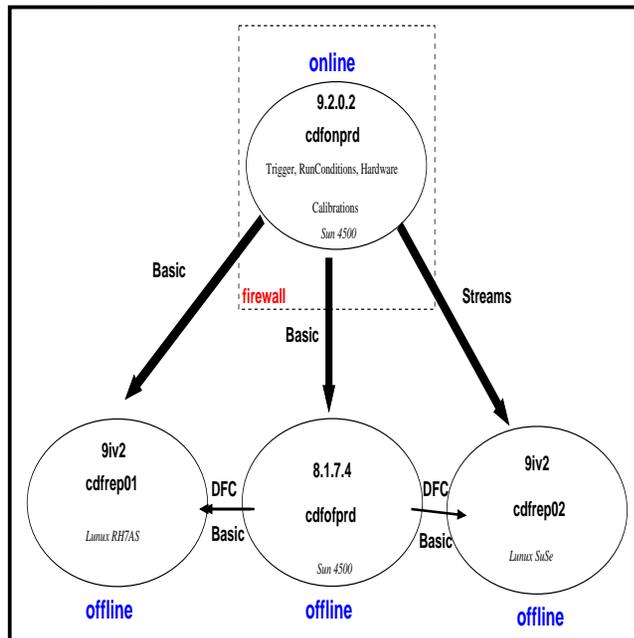


Projects

Replication using Oracle Streams

Personnel: A.Kumar, N.Stanfield

- during January shutdown *cdfonprd* has been upgraded to Oracle 9.2.0.2
- ready to start setting up stream replication
- fully implement stream replication by 1 Apr, 2003



January 28, 2003

Dmitry Litvintsev, Fermilab CD/CDF



Projects

Freeware DBMS

Motivation

People: Richard Hughes (OSU), Jim Kowalkowski (Fermilab)
Svetlana Lebedeva (Fermilab), David Waters(UCL)

- ✚ users doing analysis at remote institutions would clearly benefit from being able to read from local database
- ✚ one of the requirements to CDF Data Base access layer was to provide back-end to freeware DB implementations. The choice at the time was mSQL
- ✚ all CDF DB applications accessible from off-line reconstruction or analysis code support mSQL
- ✚ there is a positive experience with running off mSQL database at Rutgers University (DFC only)
- ✚ database export using mSQL has been worked out by Mark Lancaster but is now in limbo
- ✚ mySQL back end is very similar to mSQL back end.
- ✚ mySQL is widely used database
- ✚ advantage of using public domain products to populate mySQL from Oracle
- ✚ advantage of having replication features built into recent versions of mySQL (not the case with mSQL)



Projects

Freeware DBMS

Timeline

- ✚ end 2001 -- beginning 2002, introduced mySQL Database Utility classes
- ✚ Jan-Mar 2002, CalibDB and DataFileDB APIs have mySQL back-end implemented
- ✚ beginning 2002, Java program that generates mySQL DLL and populates mySQL database from Oracle:
 - applies to all databases applications
 - Oracle data types are mapped properly
 - cron job driven incremental updates are possible
 - customized output, allows to select entries the client is interested in
- ✚ Jan-Feb 2002 mySQL replication from OSU to UK has been setup and has been demonstrated to work
- ✚ November-December 2002, Download, configure and compiling different versions of Mysql and Postgres for following platform Linux+2.2/+2.4, SunOS+5/+5.6, IRIX+6.5.
- ✚ December 2002, Research and test different freeware features, configuration options, server/client environment, configuring many servers on the same machine, administration freeware server, and main set of Mysql and Postgres commands.
- ✚ December 2002, Mysql and Postgres DBMS are in KITS



☞ January 2003,

Working on next generation of Java codes: created property file corresponding to each used-case user can define, handle Oracle views. The Oracle views are created as Mysql tables and then populationg data by executing the Oracle view. There are some options to populate Mysql data:

- truncate Mysql tables, get Oracle record set using JDBC to INSERT data into Mysql tables.
- truncate Mysql table, create flat files one per Oracle table, and use Mysql LOAD INFILE command to load data from flat file into corresponding Mysql tables.
- Do not truncate Mysql tables, just to append data according to user criteria.

☞ January - February 2003, initial tests of running AC++ code, debugging mySQL back-ends.

☞ March 2003 full scale tests, simulation loading of mySQL db, torture tests

☞ work with OSU and other interested parties on setting up mySQL replicas



Projects

DB monitoring

People: Yuyi Guo (Fermilab), Jim Kowalkowski (Fermilab)

- ☛ we have had major problems with robustness and reliability of DB connection code, e.g. dangling connection problems.
- ☛ last fall Jim proposed very effective solution to collect information about queries done from within CDF analysis code.
- ☛ Error reporting code has been added to DBManager
- ☛ Based on ErrorLogger it reports to separate logging server
- ☛ User control of detail level on per-job basis.
- ☛ This will be our primary tool for finding out out connection usage patterns
- ☛ Needs input from CDF on what kind of information and how to be displayed on the server. We need one additional FTE to work on this project. We want Eric Wicklund to work with Yuyi on this. Approximate time scale is three months.

DB metering code

People: Dennis Box(Fermilab), Art Kreymer(Fermilab)
Nelly Stanfield(Fermilab)

- ☛ to contain off-line DB servers overload due to buggy code that generated many connections the DB metering patch has been proposed (Fall 2002)
- ☛ installed USER_LIMITS table. December 2002
- ☛ Dennis committed the metering patch December 2002



Projects

Code reviews

People: Ken Bloom (University of Michigan), Dmitry Litvintsev (Fermilab)
Jim Kowalkowski (Femilab), Liz Sexton-Kennedy (Femrilab), Rick Snider (Fermilab)

- the instabilities in database connection code, resulting in dangling idle connections and crashes prompted review of this code.
- code review committee was formed late September 2002
- October 10, 2002 committee presented its findings and recommendations (CDF 6179)
- Dennis Box implemented these recommendations during the following month
- changes in connection code uncovered inappropriate use of connection authentication mechanism. In particular a lot of effort of Dennis and Dmitry has been spend on understanding and fixing code. This is so called *TRIGSIM++* problem.



Projects

DB schema review

People: A.Kumar, N.Stanfield, J.Trumbo, A.Varganov, T.Wright

- we want to initiate a review of Trigger application starting two weeks from now.
- review should be completed withing 7-14 days
- review recommendations implementation will take another 2 to 3 weeks

Trigger API code review

People: to be identified

- following schema review the code will have to be modified
- this will be a good opportunity to review the code as well
- time for review and implementation is 2 - 3 weeks



Projects

New API development

- not all DB applications are completely ''dressed'' with API code
- this creates major issue with users bypassing the API creating idle connections

TriggerDB and RunConfigurations and Hardware

People: Jari Soudunsaari (Rovaniemi Polytechnic)

Hannu Kahkola (Rovaniemi Polytechnic)

Risto Kivilaht plus 3-4 students

- Finnish colleagues agreed to take on development of DBManager based API for TriggerDB, RunConfigurations and Hardware DB applications
- they will arrive May 2003 and will be permanently here for 6-7 months.
- it is anticipated that working API for listed applications will be delivered during this time frame



Projects

CodeGen Rewrite

People: Dennis Box (Fermilab)

- CodeGen rewrite is complete and repositied in cvs
- This is OTL and Text versions of CodeGen
- pending global test of API. Need assistance from Art Kreymer or Chuck De-Baun to use cdfpca and rawhide release to do the test. This has to be complete this week. Possible fixes will follow next week.
- CDF is interested in having CodeGen work with ODBC. Dennis expressed interest in coding. This is a project for 1-2 months depending on his other engagements
- CDF is trying to attract computer professional from participating institution to help to work on it.



Projects

Calibration Task List

People: Dennis Box (Fermilab), Jack Cranshaw (TTU), Jim Kowalkowski (Fermilab)

- ☞ fill SET_SET_MAPS. The table exists but there is no API. This is needed to retrieve list of parent valid sets that were used to create used set, 10 days
- ☞ fetch by Process Name, 2wk, need full spec.
- ☞ latest Run/Jobset for consumers
- ☞ multi-run fetch. Needed for beamlines
- ☞ storable Calibrations, 1wk
- ☞ code bloat analysis, 2wk
- ☞ split CalibrationManager into DataBaseManager and CalibrationManager, the former responsible for handling DB connection and the latter responsible for retrieving process_name, pass, jobset etc. 2wk
- ☞ better iomap aliases , 1d
- ☞ optimiation tests -- some SQL queries are inefficient, 2wk (DB monitoring tool is of help)
- ☞ Data subselect (subset of channels)

complete document is here:

http://www-cdf.fnal.gov/internal/upgrades/computing/calibration_db/calibtasks.html

We expect to be short on manpower here. Therefore completing this task list as soon as possible is high priority.



Projects

CDF SAM

People: Randy Herber, Art Kreymer
Robert Kennedy, Dmitri Litvintsev, Jeff Tseng
people from Glasgow, Oxford, UCL (GridPP), TTU, Karlsruhe Univ, Rutgers

- ☞ DFC to SAM migration (3mo):
 - finalize DB design
 - cut in CDF dev
 - CDF adaptations to DataFileDB; dbserver changes
 - SAM population from CDF DFC
 - integration refresh
 - production cut
- ☞ storage of MC into SAM
- ☞ dCache+SAM (queued behind dCache stability)
- ☞ streamlined bootstrap, configuration, versions (2mo)
 - software versions
 - supported station configurations need to be defined, with security constraints
- ☞ SAM parallel jobs (2mo)
- ☞ SAM for remote institutions (1mo after SAM parallel jobs)
- ☞ integration of dCache+SAM and SAM parallel jobs CAF (1mo after dCache+SAM)
- ☞ SAM services migration from fndaut1 to fcdfsun1 (1mo)



Conclusion

- CDF is successfully collecting and analysing data
- Database contains less than 0.1% of information from experiment but it is crucial for proper analysis
- CDF DB API layer does not create major operational problems or adversely affects reproducibility of physics results
- CDF enjoys strong support from CDF computing staff
- programming support from Dennis and Yuyi and design support from Jim and others is required to both:
 - bring development projects to completion
 - keep existing applications in good shape. This code is being heavily pounded upon by users and there is always a need for fixes or extended of functionality
- DBA services provided by Anil, Nelly and Julie are essential for:
 - keep replication working
 - setting up streams replication
 - 24x7 support of main production servers
 - table analysis
 - training new people