Data Handling Report

Paul Hubbard, Mark Leininger, Dimitri Litvintsev, Fedor Ratnikov, Terry Watts, Eric Wicklund, and Stephan Lammel

Rutgers University and Fermilab-CD

- **What was done**
- **Current Status**
- **Some Statistics**
- **Remaining Things**

What was done

Late start, at the last review, there was Terry Watts and myself.

- Terry put a lot of effort in the evaluation of the prototype as requested by the last review to proof the feasability of the proposed system. During the last month he also worked through all the issues of secondary dataset creation.

Eric Wicklund put a first version for the CDF Run II Data File Catalogue together.

- He put a lot of work into Oracle to make this a feasable option for both calibrarion and DFC.

Fedor Ratnikov joined end of 1999 and took over the data handling I/O modules in ac++.

- He worked also on the persistency and ac++ event I/O in general, making the current version of DH modules very feature rich and user friendly.

What was done, cont

Paul Hubbard joined in early 2000 working on the disk inventory manager.

- There were a lot of general design discussions with the offline management in the first half of 2000. In autumn a first client server version was available. After a relative short upgrade period, he was overrun by operational issues.

Dimitri Litvintsev joined in summer 2000.

- Shortage of expertise in the event display delayed his start on the Data File Catalogue further. He has done extensive work on understanding all interactions, revising the schema, designing new API and command line interfaces plus implementing all of this.

Mark Leininger joined December of 2000.

- He worked on the procurement for the central analysis system until spring of this year. He is very much involved in the day to day operation.

Current Status

The tape subsystem is pretty much complete and in maintenance mode. (The exception is the check out/testing of the AIT-2 on Linux.)

The disk inventory manager is still in a rather early stage and significant work remains. Missing parts include robustness and support for secondary dataset creation.

The daemon to cluster files into filesets had to be kept in its temporary toolkit style implementation because of performance problems with the online disk subsystems.

The CDF Data File Catalogue is in good shape and has the schema and C++ API for all the data elements/table complete. There is an elaborate web interface that now also includes access to the calibration and other CDF databases.

The data handling I/O modules for ac++ are in very good shape. The current version has the full interface to the Data File Catalogue, supports rather complex input specification, and can sort events on input.

The batch system setup on our machines is an out of the box LSF setup. We need a few customizations/enhancements for Run II.

Current Status, cont

Data archiving works well.

Staging data from tape back onto disk is fully functioning.

Accessing data in the disk inventory manager from ac++ as datasets is all there.

Managing and querying the metadata is available via both command line and web interface.

Sharing data (while it flows through the system) works as designed.

Writing secondary datasets is not yet possible.

All components for tape copy to export data are in hand but need to be setup.

Some Statistics

The system is running for online data archiving and farm input + output quasi non-stop since middle of 2000.

- over 1100 tapes written
- about 50 TBytes

User data access is up and running on fcdfsgi2 since December of 2000.

- first months very smooth (coordinated data staging from tape)
- starting April very rough for several months (user driven tape access)
  - exercising not well tested parts of DH software
  - subtle bugs in server causing big problems
  - IRIX scheduling behaviour under heavier I/O load
  - known batch system deficiencies

Last two weeks the server was pretty stable:

- 1.8 M client connections
- 1.3 k filesets staged in from tape 11 TBytes
- 3.5 k filesets read by users 30 TBytes

Remaining problems:

- user inflicted (using files in disk inventory managed space but not informing the disk inventory manager)
- batch system (bkills of jobs in tape queue)