

Petabyte Data Collection and Processing at CDF

High energy physics experiments collect massive amounts of data.

A detector consists of many hundreds of thousands of potential pieces of information for each collision of particles. The online and trigger systems are used to select the "best" events, based on the signals seen in the detector. The information so filtered is saved and processed later.

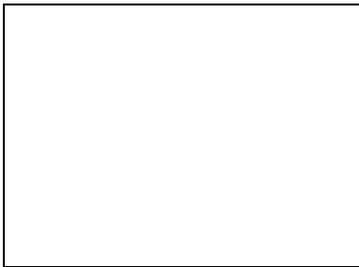
Title: CDF Detector Roll-in



CDF Run II LOGO

Loop on screen: CDF detector roll-in, assembly

Title: CDF Event Display



On Screen: Many views of "live" CDF events

Caption: The CDF event display shows the characteristics of events that were triggered by the online and trigger system.

Title: CDF Online System



On Screen: CDF logbook, camera in CDF control room

Caption: The online system is used to keep those events which are most interesting for further analysis
The reduction factor is from:

Collision rate: 2.5 MHz

L1 : xxx KHz

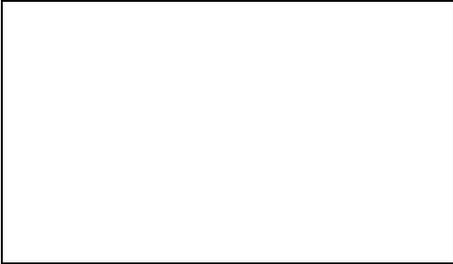
L2: 300 Hz

L3: 75 Hz

The average event size is 250 KByte

The output rate is about 20 MB/s

Title: CDF offline production farms



On screen: Views of farms, displays from
real-time farms monitoring

Caption: The CDF farms are arrays of PC's running Linux which are used to analyze the raw data from the detector and reorganize it into quantities which can be used by the physicists to understand the details of each event. Each PC analyzes two 1 GByte files simultaneously (these are dual PC's). Each file takes about 5 hours to process.