CDF Run II Computing
Status & Plan

Dan Amidei
University of Michigan
July 26, 1999

Big Offline Goal for Summer 99

- **Deliver a beta-release which**
  - has firmly in place enough of the major structure...
  - that it can sensibly be used...
  - to engage the collaboration in the next step...
  - of moving the Offline from technical development...
  - to an informed evolution which addresses our physics driven needs...
  - and utilizes our vast distributed physics expertise.

  (*and also suck in a small army of postdocs and students that we badly need to get to work on all the technical micro-detail that goes into it)*

- **Ingredients**
  - the “May Integration”
    - ok, make that July
  - the CDF Run II Physics Workshop

The July Integration: focus on infrastructure

- **Event Data Model (EDM)**
  - the representation of the data **IS** an OO natural
  - pass C++ objects between modules
  - persistify using ROOT I/O
  - backward compatible using StorableBanks
  - features to implement policy re bookkeeping etc
  - we have an implementation, and we have examples of “high level physics objects”
    - **JET, MET, ELE, MUON, CDFTRACK**
  - we did not get to complete and latch-key integration, but its useable
  - benchmarked with Farms group: 10MByte/cpu-sec
    - OK for farms
    - OK for L3
  - **evaluation and evolution of the EDM and “high level objects” based on input from physics groups and collaboration at large may be most important issue for the beta test**

The July Integration (cont)

- **unified geometry**
  - enforce consistency across Sim and Reco
  - design and implementation of base geometry class complete: **CDFDetector**
  - tracking compliant, calorimeter in progress
  - but still coming:
    - muons
    - Geant3 interface being tested
    - Sim to use ReadoutGeometry
    - move initialization out of headers to real DB
- **database access**
  - C++ interface to ORACLE, gets, puts, validity checks
  - Java interface for online
  - automatic generation tools using Java
  - support of mSQL as possibility for remote sites
  - used in Calor commissioning, and in Calor Reco
  - user examples on Web
  - **milestone met!**
- **framework (ac++)**
  - a number of improvements in function and interface

Other Progress on Code

- **Reconstruction**
  - L3 Regional tracking, Outside-In and r-z for silicon
  - Calorimetry retrofit, first pass Electrons
  - Muons: c++ version of stubs/extrapolation underway
- **Simulation**
  - more generators: AC++ VECBOS, QQ
  - luminosity monitor, TOF, improved silicon hit model
  - detailed validation
- **Event Display**
  - ROOT based prototype taking off
  - unheard-of functionality !!
  - eager for feedback and input

- **Its All pulled together as Offline 3.1.0**

Retrospective on May Integration

- **Offline 3.1.0: A somewhat imperfect beta version**
  - huge code efforts mounted by single individuals will lead to schedule and mission slip
  - nevertheless
    - we have a release
    - **main elements of the design there in enough detail to allow serious user test and feedback**
- **So Consider it A PROPOSAL**
  - a real offline prototype for your consideration and comment well in advance of data taking
- **The ball is in YOUR court**
  - contact your convener today

Meanwhile: Data Handling

- **Prototype cluster**
  - small Sun + small SGI + cdfaix+ linux (coming)
  - FC disk
  - AIX home areas on fileserver
- **Run II central analysis systems**
  - **first Run II analysis node on floor at FCC**
    - SGI ORIGIN 2000 wins bid
    - 64 R12000 processors (300MHz)
    - 30,000 VUPS = 10,000 Spec Int = 6 x cdfsga
    - CDF environs, users in September
  - **CDF Tape Robot**
    - EMASS/ADIC " technology independent"
    - D0 expelled, 4th "quadro-tower" added
    - capacity of 20,000 tapes = 1 Pbyte for 50G cartridge
  - **Disk**
    - RFP out for > 5TB FC/SCSI disk
    - on big SGI by October

Data Handling (cont.)

- **Software**
  - **Recall:**
    - CDF catalog in ORACLE
    - LSF batch for CPU and I/O
    - DiskInventoryManager/Stager proofed in Jan
  - **DH prototype version grows out of DIM proof**
    - AC++ input module, access CDF catalog using 'standard' ORACLE API, invoke DIM
    - to be reviewed in August, installed thereafter
    - preliminary model for output defined
  - **Also in progress**
    - tape deamons
    - CDF catalog web browser
- **Farms**
  - CDF production running on prototype
  - Detailed design for "Production Software"
    - to be reviewed in August
  - RFP out for first big worker node purchase
  - I/O test of EDM

Next:

- **September**
  - Code
    - Run II Workshop Meeting for Global critique?
    - proposal for EDM branches
    - checklist of production routines
    - farms production prototype
    - L3 exe prototype
  - Data Handling
    - coding DIM/stager continues
    - prototype output function
    - disk and tape on SGI
    - proto DIM/stage/batch on SGI

And then on to October...

October Integration

- **Code**
  - focus on completeness of algorithms and access
  - implement improvements identified by beta test
  - establish level of ROOT connectivity
- **Data handling**
  - full system for Production
  - complete hardware model
    - new SGI
    - TB FC disk
    - EMASS + chosen drives
    - farm1
  - complete software (v 1.0)
    - DIM
    - tape stage
    - catalog
    - input
    - output
  - **"Physics Job" runs on Production output**

Then:

- **Mock Data Challenge 1 (ca 15 Oct)**
  - data from L3 through to Production Output
  - a “connectivity test”
- **Cosmic Ray Data (ca 15 Nov)**
- **Mock Data Challenge 2 (ca Feb 2000)**
  - add DB server
  - run system at full speed/capacity from L3 to Production
  - big MC and user jobs
  - a “stress test”
- **Integration 2000**
  - pull it all together in time for first collision data

Total CDF Computing

- **Committee on Run II Trailer Computing**
  - Agreed to serve:

    Mark Leininger and Kevin McFarland, co-chairs

    Stephano Belforte

    Chris Green

    Richard Hughes

    Stephan Lammel

    Petar Maksimovic

    Armin Reichold

    M. Shimojima

    JJ Schmidt
  - Charge, *e.g.:*
    - How tightly should we specify what hardware configurations will be supported?
    - Is there NSF cross mounting of disks and what are the implications of this policy
    - Who has system priv?
    - What is the strategy for backup of the local PC disks?
  - Report end September

  - Note: support manpower in trailers is a significant issue
    - Organize around FCC + Trailers as **integrated system**

Total CDF Computing (cont.)

- **FNAL Security Plan**
  - "strong authentication", KERBEROS
  - or "secure portal" w/ non-reuseable password
  - reset alll ID's, accounts, remote access
  - see
    - www-dcd.fnal.gov/computersecurity/strongauth/
    - Seminar: **Tuesday, 6/27 1:00-2:00 in WH 1West**
    - Comp Reps Meet
- **Run II Tape Technology**
  - Exabyte Mammoth-II is "front runner"
    - Pending delivery of test modules
  - Sony AIT-2 also considered
  - Details at Comp. Reps meeting
- **VMS migration**
  - Round 1: Run I data disks to CDFSGA, kill satellites
  - Round 2: all non-essential cycles off FNALD by 2000 AD
    - that means CDFNEWS and **MAIL**
    - details, inc. proposal for new Web-based CDFNEWS at Comp Reps

CDF Week activities

- **Monday pm:**
  - More Offline Overview
    - M. Shapiro on the Offline Design
    - R. Hughes on Experience and Examples with Run II Offline
    - Session on EDM and Physics Objects
- **Tues**
  - am: Offline Planning meeting
  - pm: Tracking/Muons/Discussion, trgsim++
- **Wed**
  - am: Computing Reps meeting, L3
  - pm: Calorimetry/Jets/Electrons/Discussion
- **Thurs**
  - pm: Online Monitors, High level Physics Objects Continued
- **Tutorials**
  - many sign-ups: schedule to be posted **Tues am**