Data Tier Group

Oklahoma Workshop

Serban Protopopescu
July 9, 2002

The data tier group's main responsibility is to ensure that the output of the D0reco program is available in a form suitable for analysis and reprocessing subject to volume constraints.

The output from D0reco and RecoAnalyze which have been used until now are not sustainable in the long run, their volume will soon exceed the capacity of the STK robot and the available disk space.

Almost 2 years ago a data tier committee recommended that the output of D0reco be in two streams:

- DST:size < 120 Kb enough information to
  - Redo all calorimeter clustering
  - Redo track finding and track fitting (central and muon)
  - Redo vertex finding and vertex fitting
  - Redefine physics objects
  - All physics analysis
- Thumbnail: size<10Kb enough information for
  - Most physics analysis
  - Redo vertex finding and vertex fitting
  - Certify physics objects

DST files reside on tape, Thumbnail files reside on disk (for all events). Both managed by production group.

Day to day analysis is to be done with root files derived from thumbnail files. Managed by physics groups.

More details on proposal in
`http://www-d0.fnal.gov/~serban/data_tier`

Data Tier Group Serban	
Serban	DST (code development and testing)
E. Nagy	Thumbnail and TMB root tree code development and maintainance
Ia Iashvili	transition from root tuple to root tree

1st version of Thumbnail available in p11

0th order version of DST available in p11

1st version of TMB root tree available now in test releases, will be in p12

Must have well defined and debugged versions of all 3 in p13.

DST Requirements

1. size <120Kb
2. Event Display
   - Tracks with all associated hits
   - Unassociated hits
   - Calorimeter cells and clusters
   - Preshower clusters
   - Vertices
   - Physics Objects
3. Ability to rereconstruct
   - Tracks
   - Vertices
   - Physics Objects
4. Study detector performance
5. Remake thumbnail

Specifically left out of requirements

- Ability to recluster track hits
- Ability to recalibrate every detector component

Average size in Kbytes/event (p11) (MC SUSY & data)		
Chunk	MC (uncomp)	data
All	600 (1427)	350
RawDataChunk	105 (204)	8 (trig. block)
CPSDigiChunk	4	3
FPSDataChunk	8	
SFTDigiChunk	6	14
SMT data	134K	
MC data	40K	
CalDataChunk	46 (145)	12
CalClusterChunks	22 (45)	14
GTrackChunk	17 (23)	85
CPSClusterChunk	10.0 (29)	0.06
CFTClusterChunk	5.0 (14)	6.0
FPSClusterChunk	3.7 (14)	
VertexCollChunk	0.5	0.5
muon chunks	1.4 (5)	0.9
ThumbNailChunk	7.0 (43)	4.3
MCKineChunk	14 (23)	
TrackCalExtraChunk	25 (50)	10
SMTGlbl(B+D)CollectChunk	95 (242)	140

Adding all chunks in red:
MC: 126 Kb
data:131 Kb

Missing from above numbers:
D0PhysicsObj chunks:
$\approx$ 10 Kb
Compressed SMTGlbl(B+D)CollectChunk:
$\approx$ 25 Kb

TOTAL: $\approx$ 150 Kb

More compression needed if we want strip information.

In data gain 60Kb dropping extra GTrackChunk's

But get it all back lowering cal. threshold.