Data Tier Proposal

April 21, 2000

U. Heintz, J. Hobbs, B. Klima, J.Krane, L. Lueking, S. Protopopescu, M. Strovink, A. Zylberstejn

The main motivation for having events with different amounts of information is speed of access. The original charge to the committee was to consider the contents of EDU250, EDU50 and EDU5. The data tier committee would like to propose to change the charge somewhat and refer to 5 different formats for event data, only three of which will be available for all events. In this note we will try to justify why we wish to amend the charge. The 5 different data formats are:

1. **Raw Data**: That is the original information from the detector. The data will be written to tape, processed by a reconstruction program and archived. There will be rapid access to a limited amount of Raw Data but as time goes on it will be a smaller and smaller % of the total. Event size is expected to be 250 Kb.
2. **EDU250**: This is basically the complete output of the reconstruction program. The size for planning purposes was set to 250 Kb, but present indications are that this is likely to be closer to 500Kb. We will refer to this format as *DBG* (for debug) from now on.
3. **EDU50**: This is some subset of DBG. The 50Kb event size was set by the requirement that all the data be stored on the robot for faster tape access but data are still accessed sequentially. However, expected robot capacity indicates this size could be doubled. The 50Kbyte size is too restrictive and we propose to keep as much as 120Kbytes/event (half the Raw Data size). We will refer to this set as *DST* from now on. This set should contain all information needed for analysis and post-reconstruction processing.
4. **EDU5**: This is a subset of DST which we should keep strictly to 5 Kb. This set contains information needed for rapid event selection, and resides on disk allowing random access to each event. We will refer to it as *TMB* (for thumbnail).

5. **Analysis sets**: These are subsets whose format is defined by the physics groups, must reside on disk, and are used for physics analysis. They could be subsets of any format: RawData, DBG, DST, and Root tuples.

1. Raw Data

Most of the raw data are expected to be stored in one chunk, **RawDataChunk**, with all the information the detector produces during a beam crossing that satisfies trigger requirements. Information from L1 and L2 triggers are stored in the RawDataChunk. There may be an additional chunk generated by L3.

The information is unpacked for event reconstruction; but there are blocks of information (particularly trigger information) which are needed for analysis and are not used much by the reconstruction program.

2. DBG

This is supposed to contain enough event information to follow in detail how the raw data was processed besides the information needed for physics analysis. We propose that this format not be used to store all events. It should be used only for small subsets which need to be investigated in detail, and for event display. Data in this format should be created on demand by rerunning the reconstruction program on Raw Data.

The data can be thought to consist roughly of 6 categories, listed in order of dependency:

1. Trigger Information: some of this information may be used in the reconstruction program for chosing algorithms. All of it needs to be passed on for later analysis.
2. Hits: these are unpacked data with calibration corrections applied. If there is a new set of calibration constants for the same data set, hits need to be recalculated.

3. Clusters: Hits are collected into clusters. This is the first stage where multiple algorithms may be applied. The clusters usually include spatial information, thus the algorithms need geometry and alignment information. Changes in alignment may require redoing the clustering.
4. Tracks: Clusters in tracking chambers are used to make tracks. Track finding in the CFT and SMT are, by a large factor, the most cpu intensive activities in the reconstruction program and likely to be done only once for the bulk of the events (but many times for some subsets, particularly at the start of data taking). After track finding follows track fitting to improve errors on track parameters.
5. Vertices: made from fitted tracks. Primary vertices are a critical component of all physics objects, and thus need to be found prior to object identification. Secondary vertices are critical for specific physics channels, and require the highest possible resolution. It is almost certain that in some cases secondary vertices will be refitted after object identification. The refitting is not likely to be part of D0reco.
6. Physics objects: These are the objects most heavily used in physics analysis (electrons, photons, jets, etc.). The reconstruction program is only the first pass. The D0reco goal is to identify candidates with high efficiency while keeping backgrounds manageable. Determination of quality is left to the analysis stage.

The complete output of the D0reco program has been designed mostly for ease of code development, not necessarily for ease of analysis.

The bulk of the complete D0reco output is in the hits and clusters ($\approx$ 75%), followed by the tracks ($\approx$ 22%). The tracking information has an enormous amount of duplication, and can easily be reduced by a factor of 4.

3. DST

This format should contain all the information needed to analyze events. It should be the standard output of D0reco. The bulk of the data in DBG are hits, clusters, and GTracks (global tracks) from central tracking. Most of the CPU in D0reco is in finding tracks so, if one needs to refind tracks, little is gained trying to redo this step with already processed data if the event size is the same or bigger than the Raw Data. We propose to limit the DST to <120Kb (1/2 the Raw Data size). This requires that fitted tracks and clusters be kept in the most compact form possible while removing raw data and hits from the central tracking. The most compact form for the data from any subdetector is Raw Data, so we propose to keep up to 60-70Kbytes in calibrated, packed pseudo-"Raw Data" format (or equivalent packing) on DST from subdetectors other than central tracking. Having these data available gives the flexibility of redoing most aspects of the reconstruction. In the remainder (50-60Kbytes) we must keep all information from physics objects, tracks and vertices. That information should be complete enough that the information in Raw Data format and compacted clusters will not usually be unpacked, and thus will not contribute significantly to the average CPU time for processing events in DST format.

All the reconstructed data should be stored in this format. It is expected that we will be cycling through these data frequently, and thus, the freight train approach may be the most efficient way to handle them. However, we should keep in mind that there may be two different types of programs trying to access these data:

1. Making Root tuples (or equivalent) for physics analysis using pre-selected lists of events.
2. Reprocessing aspects of reconstruction before making the Root tuples.

The CPU requirements are likely to be rather different for 1) and 2), and a single freight train may not be optimal.

4. TMB

This format should be kept as small as possible for two reasons:

1. Very rapid cycling through events
2. Free disk space for physics groups

The main purpose is for event selection with rather minimal analysis. We envisage using them for making lists of events after applying some selection criteria, and lists can be used as input for making more restrictive lists. For efficient selection the full data set must exist on disk with random access to events. The generated lists can then be used for selecting events with larger format to make analysis sets. We must resist the temptation of trying to make the contents of this format sufficient for analysis. It will be impossible to keep the event sizes small if we do.

5. Analysis sets

The analysis sets, which must reside on disk, should be defined by the physics groups. Each physics group should be allocated a large amount of disk space which they should be free to fill with data in whatever mixture of formats is most convenient for their analyses. Clearly for storing large number of events they may choose a Root format, but for smaller data sets they can also keep DBG or DST. There is no need to store TMB separately, only the lists of events they are interested in. The experiment as a whole will benefit from having a well defined Root tuple core common to most physics groups but each physics group should be given the freedom to add to that core additional information that is of interest to specific analyses.