CDF/DOC/COMP_UPG/PUBLIC/4498
version 1.02 February 21, 1998

ROOT-based data handling system for CDF in Run II.

2. Gedanken experiments for selected physics analyses

P.Murat

Abstract

This note presents a set of Gedanken experiments for the ROOT-based data handling and analysis system proposed in [1].

Following [2], we consider several physics analyses with very different data access patterns (top cross-section analysis, exotics missing-Et analysis, inclusive jet analysis, inclusive J/psi cross-section analysis and exclusive B reconstruction) and discuss the data handling aspect of these analyses.

We also discuss several typical offline reconstruction tasks which could be necessary in Run II: SVX retracking, full retracking and jet reclustering. Choice of the data handling model could have a very strong impact on the optimal strategy of these tasks.

1 Introduction

One of the goals of Gedanken experiments is to compare how different data handling models proposed to be used by CDF in Run II would be used by the physicists analysing the data. For this reason we will concentrate on the same examples of physics analyses considered in [2]

2 Top Cross-Section Analysis

2.1 Signal selection

Top cross-section analysis is an example of physics analysis which requires multi-step data selection.

The analysis starts from the inclusive high-Pt lepton datasets defined by Level3 triggers. In Run 1 inclusive high-Pt lepton data were processed in express stream (stream A). Overall number of events in Run II high-Pt lepton datasets could been estimated at about 20 million events (see [2]). The selection of the final dataset goes through the following intermediate steps:

- selection of inclusive high-P_t lepton (e- and μ-)datasets with the leptons passing so-called **tight** cuts. Data reduction factor at this step is about 5;
- selection of inclusive W- and Z- datasets with W-datasets containing about 90% of events. Another data size reduction by a factor of about 2.5 is expected at this step.
- finally inclusive W-dataset is split into several **W+n jets** datasets. W+3jets data sample, containing W and 3 or more high-Pt jets, is used to calculate the production cross-section of $t\bar{t}$-system.

ROOT-based data handling system [1] maintains all the data necessary for the analysis in the central disk repository all over the Run II. For this reason we expect that the strategy of the data analysis will not change from the beginning to the end of run - the analysis will be using only the disk-resident part of the data and will require tape mounts only for reading the Monte Carlo data, which are supposed to be tape-resident. Large size of the robotic storage dick cache would allow not to reread the same MC data twice.

There is however one technical issue, which would be very interesting to prototype. The data stored in the central repository in ROOT format could be used for analysis and histogramming purposes in exactly the same way as PAW ntuples. Due to this feature provided by ROOT, it might be possible not to create some of the secondary datasets (for tight leptons, for example), but to store selection results as a **list of run/event numbers** for the selected events. The size of the corresponding run/event list could be quite small. For example, for the whole Run II tight lepton dataset the size of the corresponding event/run list, specifying event and run numbers for each event in this dataset, would be of about 15-20 MBytes. If using run/event list for the next stages of analysis will appear to be efficient enough, one could avoid creating secondary datasets at intermediate levels and use event/run lists instead.

At the same time for the final datasets, which size is expected to be rather small (about 30-40 GBytes), it might be more convenient to keep them in separate files. This would significantly simplify distribution of these datasets to the remote institutions.

2.2 Studies of B-tagging efficiencies and background studies

2.3 B-tagging studies

A primary sourse of data used for determination of B-tagging efficiencies is an inclusive high-Pt lepton dataset. Disk-resident information maintained by ROOT-based system in the central repository is sufficient for the B-tagging studies, so no tape I/O is required.

SVX (re)alignment is an essential issue for all the B-tagging analyses, changes in alignment require the analyses to be repeated. As it will be shown below, with SVX cluster banks stored on disk, redoing SVX reconstruction could be organized without using tape-resident branches of event data.

2.4 Background Studies

One of the most important sources of background in top cross-section analysis are the mistags (false secondary vertices) found in non-$t\bar{t}$ events. Mistag studies in Run I have been mostly done with inclusive jet, MET and inclusive W samples. In Run II all these samples will be available from the central disk repository and we don't expect that accessing them will be any different from the accessing the inclusive lepton datasets. The specifics of the data access pattern here is that mistag studies do not require selection of secondary datasets.

3 Exotics Missing Et Analysis

Exotics missing Et analysis gives us a typical example of another distinct data access pattern with huge data volumes scanned in search for very few rare events. The absolute majority of the events could be rejected by very simple cuts, for example by cutting on missing Et in the event, number of high-Pt leptons in the event etc, which use only minor part of event data. This data access pattern is the one, where the advantages of ROOT data model could be seen very clearly: as it has been shown in [3], ROOT can process huge volumes of data in "dataset scan" mode (when only a minor part of event data is read in) 10 and more times faster then YBOS, which reads the data in sequential mode. This difference could become much larger in case if tape I/O is required for the sequential model. As all the necessary for the analysis data are available from the central repository, we do not expect any problems specific for this data access pattern.

Just to repeat, ROOT simultaneously provides tools for histogramming and data analysis, and as number of selected events is small, event/run list for the selected events could be easily used in this analysis instead of creating the secondary datasets (which, however, could also be is a reasonable option from the point of view of distributing the data to the remote institutions).

4 Inclusive Jet Analysis

Inclusive jet analysis gives an example of one more different data access pattern. In this analysis significant part of all the data enters the final histograms, so at all stages of the analysis, including the final one, physicists operate with the large number of events. Event selections applied in this analysis are primarily used not for reducing the size of the initial data sample but mostly for cleaning it up. Primary dataset maintained in the central disk repository could be directly used for histogramming in case of the ROOT-based data handling system. With all the data available from the central disk repository data access technique is pretty straighforward.

We could consider, however, an example of much less trivial from the point of view data handling system reconstruction task, which is related to the reconstrucion and could be of potential interest for the QCD group in Run II. This is a task of redoing the jet reconstruction with a new jet reconstruction algorithm, which may come up in the end of Run II. Data handling part of the problem consists in that in this case one needs to **add** more reconstructed data to **already reconstructed** event and to be able to store this information. Provided a sequential data model is being used, this task becomes something gigantic: it requires reprocessing **all the data**, rewriting **all the tapes** for inclusive jet data stream, changing **all the entries** in event/run/dataset databases etc. And we didn't even think yet about the shelf-resident (not kept on robotic tape storage) part of the data...

Within the ROOT-based data handling system the task seems to be quite feasible. First of all, the output of the reconstruction - a list of reconstructed jets - is a small object, so the size of the corresponding output container should also be small. Next, ROOT tree format allows a new branch to be added to a ROOT tree at any point in time. We store this new branch (containing list of jets, reconstructed with a new algorithm) in a separate chain of files. To do this some reserve of disk space in the central disk repository is required. As the overall size of the branch is small it could even be stored in the central repository. Storing this branch on tape requires only several hundred bytes per event to be written to robotic storage. Event/run/dataset database is updated accordingly. All the logical steps are well defined and all the necessary tools seem to be available.

However, this is a new technique and it has to be prototyped before the beginning of Run II.

5 B-physics Analyses

5.1 Inclusive J/Psi Cross-Section Analysis

From the point of view of its data access pattern the analysis aimed in calculating the inclusive J/psi production cross-section is rather similar to the inclusive jet analysis. The selections applied do not reduce dramatically the size of the primary data sample, but are intended to select the good quality data (good muons, good experimental runs etc).

The overall size of Run II inclusive J/psi dataset is estimated at about 80 million events. ROOT-based data handling system maintains all the PAD-level (30-35 Kbytes of uncompressed data per event) information available from the central disk repository. In Run I inclusive J/Psi data has been processed in express stream, which may not be the case for Run II. Depending on the time necessary to scan the whole dataset it may be possible to avoid creating the secondary, cleaned-up dataset. If we ignore for a moment the difference between Run I and Run II CDF silicon detectors, having all the data on disk makes it possible to redo RunI-type SVX reconstruction in case new SVX alignment constants are available without using the tape-resident data.

5.2 Exclusive B-reconstruction

Data access pattern in reconstruction of exlusive decay channels of beauty and charm hadrons is very close to that of missing Et-type analyses: the early selection (requiring, for example, an effective mass of J/Psi and a charged track to be close to B-meson mass) results in a relatively small dataset, so a problem with the size of a secondary dataset does not exist. The difference from missing Et-type analyis is that the selection procedure is sensitive to the performance of the track reconstruction code. When a new version of track reconstuction program (new pattern recognition, new calibration constants, new alignment constants etc) comes up, the primary dataset should be retracked and the secondary dataset should be recreated after that. Event/run lists of selected events could be quite appropriate for the analyses using the central disk repository. Distribution, however, is more convenient if the secondary datasets have been created as independent files (chains of files).

6 Retracking of B-group Datasets

In this section we'd like to discuss in more detail a very important for many analyses (top analyses involving B-tagging, all the B-physics analyses etc) issue of retracking. Data management aspect of this task is a very important issue, ignoring it could cost weeks and months of wasted time.

Looking at the history of Run I we could see that in Run I after a new set of SVX alignment constants has been derived, only a minor part of all the data (including important yet low-statistics top datasets and J/psi dataset) was reprocessed in a centralized way: primary data reprocessed, results saved, secondary datasets recreated. Many other also important for physics but high-statistics data samples (like low-Pt dimuon data and inclusive low-Pt electron data) have never been retracked (more accurately, refitted) with the last version of SVX' alignment constants. The only reason for this was that with sequentially stored data reprocessing of thousands of **shelf-resident** tapes and creating new DST's was out of any reasonable time scale. As a result, each job running on these data samples has been redoing track reconstruction in SVX, which took and still takes seconds of CPU time per event ... and weeks for the whole dataset.

On the contrary, provided a ROOT-based data handling system is used, we don't see any showstoppers. Assuming the branch containing the reconstructed SVX clusters (equivalent of QSVC) is disk resident, it is only necessary to reprocess the data once and to replace all the references to the old location of the equivalent of SVXS branch with its new location in the event/run/dataset databases. This could be done without a single tape mount. However, this is an example of the task where one needs some reserve of disk space, so a new branch would be created without changing the existing oranization of the data. This is exactly the reason why in [1] we proposed to keep about 10% of all the space in central disk repository reserved. Retracking will supposedly take some non-negligible time, during which we don't want to disable on-going data analyses. After the retracking is complete, it is necessary to update the event/run/dataset databases. Updating the databases is a fast operation which (if prepared well enough) could be done within one hour, during which the access to the central disk repository could be disabled.

And, of course, the results of the new reconstruction pass have be stored on tape. This operation involves writing out an equivalent of QTRK branch, replacing the old tapes on robotic storage with the new ones and updating the dataset database. Event/Run database is not changed. All the data corresponding to the last reconstruction pass are available in automatic mode (disk+robotic storage).

We should notice, however, that retracking impacts practically everything in the offline reconstruction: given a new list of tracks is available one would like to redo electron reconstruction, muon (but not the muon stub) reconstruction etc... This issue has more to do with physics analysis than with the data acess itself. One of the possible solutions consists in reconstructing high-level physics objects (electrons, muons, gammas) at run-time each time the analysis job is running, instead of using their parameters stored on disk. In this case new list of reconstructed tracks automatically means new list of muons, new list of electrons etc.

Finally, let's consider an issue of full retracking of some high-statistics dataset, for example, inclusive low-Pt electrons. The data access pattern is similar to the data access pattern of redoing the drift chamber calibration constants for the whole run (this methodic task has also to be foreseen!). Supposedly COTD banks (COTD branch) will reside on tape. Taking for this estimate the characteristic size of COTD bank to be 20-30 KBytes (10 KBytes pointers + 10-20 KBytes hit data), one definitely doesn't find enough disk space to keep all the COT hit data in the central repository. However, all the tape-resident branches are kept on robotic storage and only about 10% of all the event data - only COTD branch - has to be reread. Large size of robotic storage cache will allow to keep COTD data on disk for resonably long period of time. In fact 10 TBytes of SILO cache correspond to about 30% of all the Run II COTD branch data volume, so for each single production output stream it is possible to cache all its COTD branch on disk. Output of the reconstruction - equivalent of QTRK bank - is small and saving the results of reprocessing onto robotic storage should be a minor job.

We didn't find obvious logical flaws in the procedure decribed above. However it is based on using a principally new data access model. Therefore we consider it very important to prototype data handling aspects of retracking before the beginning of Run II.

7 Acknowledgements

Author would like to thank J.Lewis, M.Shapiro and S.Tkaczyk for many clarifying discussions.

References

[1] R.Brun, P.Murat, F.Rademakers, ROOT-based data handling system for CDF in Run II. 1. Specifications, Data Flow Analysis,Initial Implementation Plan
CDF internal note 4497

[2] E.Berman, ... S.Lammel et al, Gedanken Experiments for Selected CDF Analyses for the Light–Weight Database Approach,
CDF internal note 4380

[3] P.Murat, On the possibility to use ROOT Data Model for CDF data storage needs in Run II, CDF internal note 4412