

Refining the DC04 Milestones

Oct 2003

David.Stickland@cern.ch

Introduction

As defined to the LHCC, the Level-2 milestone consists of:

CS-1041 1 April 2004 5% Data challenge complete (Now called DC04)

The purpose of this milestone is to demonstrate the validity of the software baseline to be used for the Physics TDR and in the preparation of the Computing TDR.

The challenge comprises the completion of a “5% data challenge”, which successfully copes with a sustained data-taking rate equivalent to 25Hz at a luminosity of $0.2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ for a period of 1 month (approximately 5×10^7 events).

The emphasis of the challenge is on the validation of the deployed grid model on a sufficient number of Tier-0, Tier-1, and Tier-2 sites. We assume that 2-3 of the Tier-1 centers and 5-10 of the Tier-2 centers intending to supply computing to CMS in the 2007 first LHC run would participate to this challenge.

Context

Readiness for data taking and analysis in 2007 is clearly the level-0 milestone of the CCS project (and CPT). The immediate requirement is the completion of the Computing TDR at end of 2004 that will describe the plan to reach that readiness (This being a Level-1 Milestone). Thus a successful conclusion of DC04 would be one from which enough information could be derived to adequately inform the planning required for the TDR. It is not necessary that all work undertaken in DC04 is an unmitigated success, only that the lessons that need to be drawn can be drawn and that a realistic plan for the Computing TDR can be formulated.

We propose therefore to establish a set of Level-3 milestones/work-packages each of which can be separately monitored. The completion of the Level-1 DC04 milestone will then depend on the performance of each of these sub-milestones. The successful completion of DC04 would be that a sufficiently complete set of sub-milestones were satisfied such that an adequately tested plan could be proposed in the Computing TDR.

Pre-Challenge Phase Milestones

The PCP is currently underway.

Estimation of Resources required for Simulation		
SI2k Conversion (for P3 on our codes)	800 Mhz	297 SI2K
SI2k Conversion (for P4 on our codes)	2800 MHz	900 SI2K
Time for 1 CMSIM event (800Mhz)	350 s	103.95 kSI2K.s
Size of a Simulated event	2 MB	
GEANT4 /CMSIM Processing time	2.3	
Number of events required	5.E+07	
Size of Simulated data Sample	100 TB	
Processing power required	4612 kSI2K.months	

Estimation of Resources required for Digitization		
Time for 1 Event Digitization (800Mhz P3)	20 s	5.94 kSI2K.s
Size of a Digitized event	1.5 MB	
Number of events required	5.E+07	
Size of Digitized data Sample	75 TB	
Processing power required	115 kSI2K.months	
Pileup / Event	15 MB	
Time to Digitize (Proposal)	2 months	
Pileup bandwidth required	145 MB/s	

PCP-1. Generation of approximately 50 million Monte-Carlo events. The choice of event types is made by the PRS groups to be of the most use in their ongoing work. The events are mostly chosen such that they will satisfy HLT selections and thus represent the types of events that will indeed flow through the offline farm, though the mix of event types is not required to be realistic.

PCP-2. Simulation of the events with either CMSIM or OSCAR. This simulation can happen anywhere using any local and/or grid facilities. The simulated events will be stored at Tier-1 centers. The actual mix of CMSIM/OSCAR is not relevant to the CCS PCP requirements, but is decided by the PRS groups taking into account the software status.

PCP-2a. At least a fraction x of the 50M events simulated with CMSIM. These events must be Hit-Formatted by ORCA and stored in the POOL format. Catalog and Database tracking must be in place to allow later processing of datasets and adequate information on the event provenance.

PCP-2b. At least a fraction $(1-x)$ of the 50M events simulated with OSCAR and directly stored in the same POOL format as in (a) with the same cataloging and reference information available. (From the perspective of later processing events simulated via CMSIM or OSCAR must be equally accessible)

PCP-3. The Digitization of the 50M events at an effective luminosity of $2 \times 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ (under the 25ns bunch spacing assumption). The Digitization can be carried out at any sites as required.

PCP-4. At least the Digitized data (not necessarily with the MC truth information) must be transferred to the CERN Mass Storage, together with the adequate catalog information for its later processing.

PCP-5. To be able to run 75% of the PCP Simulation production in an LCG Grid Environment. This is not an integrated 75% over the PCP period, but a demonstration that for a period of, say, a week we can reach this instantaneous level. (For example the CMISM simulation rate is (currently) about 4M events/week, Completion of this milestone would be that ~3M events could be simulated using grid resources in a week.) This would not have to be part of the PCP but could be executed in for example December 2003. (Question all LCG or LCG and DPE?)

There are related Software Base milestones.

SWBASE-1. CMSIM version compete for PRS requirements

SWBASE-2. OSCAR version complete for PRS requirements

SWBASE-3. Storage of MC truth and hits in POOL. Readability guaranteed for xxx months/years.

SWBASE-4. Digitization code in ORCA/COBRA complete for PRS requirements.

SWBASE-5. At least a single-site complete catalog of all produced files relevant to later processing. Meta-data adequate to permit dataset/collection processing

Tier 0 Milestones

The Tier-0 activity is a classical single-site computing challenge. It will depend on efficient use of the computing resources available and on an adequately complete base of CMS software.

Estimation of Resources required for T0 Reconstruction		
Time for 1 Event Reconstruction (800Mhz P3)	30 s	8.91 kSI2K.s
Size of a Reconstructed event	0.5 MB	
Size of an AOD event?	20 kB	
Number of events required	5.E+07	
Size of Reconstructed Data Sample	25 TB	
Size of AOD	1 TB	
Processing power required	172 kSI95.months	
Data Challenge Duration	1 month	
Number of CPU (@ 500SI2K/CPU)	344	

TIER0-1. Data serving pool. Capable to serve Digitized events at 25Hz to the computing farm with 20/24 hour operation. Adequate buffer space (Digitized data set expected to be of order 100TB, aim to keep 1/4 of this in the disk buffer). Pre-staging software. File locking while in use, buffer cleaning and restocking as files have been processed.

TIER0-2. Computing Farm operating for 30 days . Approximately 350 running jobs 20/24 hours. Files in buffer locked till successful job completion. (500 events per job implies 3 hour batch jobs)

TIER0-3. Output products of Tier-0 production stored to CERN MSS.

TIER0-4. Secure and complete catalog of all data input/products maintained at CERN.

TIER0-5. Data catalog is accessible and/or replicable to the other computing centers.

The reconstruction step requires definition:

RPPROM-1. Tier-0 reconstruction software defined

RPPROM-2. DST persistent classes defined

RPPROM-3. Reconstruction and persistency code complete

RPPROM-4. TAG/NTUPLE defined. (Information to characterize the events and allow efficient selection in later analysis)

RPPROM-5. TAG/NTUPLE production coded, including cataloging information to allow at least ROOT and ORCA processing of TAG/NTUPLE collections

Data Analysis Issues

We propose that the model to test is that the Tier-0 is only available for the initial reconstruction. Analysis steps must be carried out at Tier-1 and Tier-2 centers. If possible we should test the use of the LCG submission system to make

a simple optimization of the computing/data sites (at least at selected Tier-1 and “associated” Tier-2 centers) as well as a more classical analysis model where the data is pre-positioned and the computing sites chosen accordingly. True distributed analysis is beyond the scope of DC04. However the data itself and its wide distribution can be useful in the following years work. We may choose to run TAG/NTUPLE production at the Tier-0 and/or at the Tier-1 sites. One could imagine running the HLT code at the Tier-0 and creating an HLT result TAG, and then running the separate PRS group TAG/NTUPLE productions at the Tier-1 sites. This both offloads the CERN site and requires the Tier-1 sites to operate some productions. We propose that TAG/NTUPLES be analyzable at the Tier-1 and more importantly the Tier-2 sites. The output of the analysis step is made of summary tables and/or plots.

The analysis step requires definition:

RPR0M-6. TAG/NTUPLE analysis code ready from physics groups.

A low-priority activity is the execution of private user code from selected users on the distributed data. Private code doesn't need to be made available in advance.

Data Distribution Milestones

- DATA-1. Replication of the DST at one or more Tier-1 centers. (Approximately 25TB in total) possibly using the LCG replication tools.
- DATA-2. Replication of at least those parts of the catalog that have been imported to each Tier-1. In case TIER0-5 is met this milestone is automatically achieved.
- DATA-3. Transparent access of jobs at the Tier-1 sites to the local data whether in MSS or on disk buffer.
- DATA-4. Defined linkage between Tier-1 and Tier-2 sites. Tier-2 sites to access the data only via the peer Tier-1 site. (This is a linkage just for the duration of DC04 and subsequent analysis, not a commitment for all time)
- DATA-5. Replication of the full Tier-0 TAG/NTUPLE at each Tier1 and further replication from the Tier-1 sites to Tier-2 sites (Whichever parts of the TAG/NTUPLE are required by the Tier-2 sites).
- DATA-6. Replication of any TAG/NTUPLES produced at the Tier-1 sites to the other Tier-1 sites and interested Tier-2 sites

Calibration Challenge Milestones

At least one site will undertake a calibration challenge. This will require the import of a possible special “DST” containing the information (for example just the high-pt tracks, or a fraction of the raw data for a sub-sample of the events) and the operation of a simple calibration task (for example alignment residual calculation) and publication of the results via a web interface. Ideally the results of the challenge will be a set of new alignment constants (presumably with a zero-mean) that can be published back to the Tier-0 reconstruction and made use of in subsequent reconstruction. It is not intended that the reconstructed data will actually be affected by these ‘calibrations’, the goal is to demonstrate the round trip ability and the ability to switch between (even identical) calibrations while production continues.

- CALIB-1. Definition of the pseudo-calibration task and identification of the site and people responsible for it.
- CALIB-2. Definition and coding of the calibration “DST” to be used, incorporation of the code in the Tier-0 executable.
- CALIB-3. Rapid distribution of the calibration “DST” (within hours at most) to the Tier-1 site and automatically scheduled jobs to process the data as it arrives.

CALIB-4. Publication of the results in an appropriate form that can be returned to the Tier-0 for incorporation in the calibration “database”

CALIB-5. Ability to switch calibration “database” at the Tier-0 on the fly and to be able to track from the meta-data which calibration table has been used.

Tier-1 Analysis Milestones

TIER1-1. All data distributed from Tier-0 safely inserted to local storage

TIER1-2. Management and publication of a local catalog indicating status of locally resident data

TIER1-3. Operation of the PRS TAG/NTUPLE productions on the imported data. Different Tier-1 sites may import different data and run different PRS group code according to their “contract” with the PRS users.

TIER1-4. Local computing facilities made available to Tier-2 users, possibly via the LCG job submission system.

TIER1-5. Export of the PRS TAG/NTUPLE to requesting sites (Tier-0, -1 or -2)

TIER1-6. Operation of a scheduled Analysis service, for example publication of plots associated with the PRS TAG/NTUPLES for each dataset processed.

TIER1-7. Tier-1 data catalogue (either produced locally or replicated from the Tier-0) accessible remotely and made available at least to the “associated” Tier-2 centers.

TIER1-8. Register the data produced locally to the Tier-0 catalog and make them available to at least selected sites via the LCG replication tools.

Tier-2 Analysis Milestones

TIER2-1. Pulling of data from peered Tier-1 sites as defined by the local Tier-2 activities

TIER2-2. Analysis on the local TAG/NTUPLE produces plots and/or summary tables.

TIER2-3. Analysis on distributed TAG/NTUPLE or DST available at least at the reference Tier-1 and “associated” Tier-2 centers. Results are made available to selected remote users possibly via the LCG data replication tools.

TIER2-4. Private analysis on distributed TAG/NTUPLE or DST is outside DC04 scope but will be kept as a low-priority milestone.