The ATLAS Computing Model

Dario Barberis

(CERN & Genoa University/INFN)

Background & History

- Computing Model studies started at the time of the Computing Technical Proposal (CTP) in 1996
- Modelling continued in the late 90's in the context of the MONARC studies
- More quantitative estimates of needed resources were presented at the "Hoffmann" review in 2000/2001
- The document "Principles of Cost Sharing for the ATLAS Offline Computing Resources", approved by the ATLAS Collaboration in 2002, is the basis of the current version of the Computing model
- Since then, discussion has concentrated mostly on data distribution to Tier-1 facilities and balance between disk and tape resident data

Computing Model Working Group (1)

- This group (chaired by Roger Jones) is:
  - assembling existing information and digesting it
  - acting as contact point for input into the Computing Model from all ATLAS members
  - preparing a "running" Computing Model document with up-to-date information to be used for resource bids etc. (end 2003)
  - preparing the Computing Model Report for the LHCC/LCG by end 2004
  - contributing the Computing Model section of the Computing TDR (2005)
- The goal is to come up with a coherent model for:
  - physical hardware configuration
    - **e.g. how much disk should be located at experiment hall between the Event Filter & Prompt Reconstruction Farm**
  - data flows
  - processing stages
  - latencies
  - resources needed at CERN and in Tier-1 and Tier-2 facilities

Computing Model Working Group (2)

- Areas of investigation:
  - Software and database implications of models
  - Top down resources issues
  - Services, operations, problem tracking and reporting
  - Grid services, role-based access, Grid instrumentation, modelling
  - Online, networking, data-flow
  - Distributed physics analysis
  - Alignment and calibration needs, production at Tier-1 sites

Event Data Flow from Online to Offline

- The trigger system will reduce the event rate from 40 MHz to:
  - 20-30 kHz after the Level-1 trigger (muons and calorimetry)
  - ~2000 Hz after the Level-2 trigger (several algorithms in parallel, running independently for each subdetector)
  - ~160 Hz after the Event Filter ("offline" algorithms on full event)
    - only ~140 Hz will be real "physics" triggers, ~20 Hz are calibration and monitor triggers
- These rates are almost independent of luminosity:
  - there is more "interesting" physics than 160 Hz even at low luminosity
  - trigger thresholds will be adjusted to follow the luminosity
- The "nominal" event size is 1.6 MB
  - initially it may be much larger (7-8 MB) until data compression in the calorimetry is switched on
- The nominal rate from online to offline is therefore 250 MB/s

Parameters of the Computing Model

- Data Sizes:
  - Simulated Event Data 2.0 MB
  - Raw Data 1.6 MB
  - Event Summary Data 0.5 MB
  - Analysis Object Data 10 kB
  - TAG Data 0.5 kB
- Other parameters:
  - Total Trigger Rate 160 Hz
  - Physics Trigger Rate 140 Hz
  - Nominal year 10^7 s
  - Time/event for Simul. 30 kSI2k s (*)
  - Time/event for Recon. 6.4 kSI2k s (*)

(*) numbers from Geant3-based simulations

Operation of Tier-0

- The Tier-0 facility at CERN will have to:
  - hold a copy of all raw data to tape
  - keep calibration data on disk
  - run first-pass reconstruction
  - distribute ESD's to external Tier-1's (1/3 to each one of 6 Tier-1's)
- Currently under discussion:
  - copy in real time all raw data to Tier-1's (second copy useful also for later reprocessing)
  - reprocess raw data on tier-0 or on Tier-1's
  - "shelf" vs "automatic" tapes
  - archiving of simulated data
  - sharing of facilities between HLT and Tier-0
- Tier-0 will have to be a dedicated facility, where the CPU power and network bandwidth match the real time event rate

Operation of Tier-1's and Tier-2's

- We envisage ~6 Tier-1's for ATLAS. Each one will:
  - keep on disk 1/3 of the ESD's and a full copy of AOD's and TAG's
  - (possibly) keep on tape 1/6 of Raw Data
  - keep on disk 1/3 of currently simulated ESD's and on tape 1/6 of previous versions
  - provide facilities (CPU and disk space) for user analysis (~200 users/Tier-1)
  - run simulation, calibration and/or reprocessing of real data
- We estimate ~4 Tier-2's for each Tier-1. Each one will:
  - keep on disk a full copy of AOD's and TAG's
  - (possibly) keep on disk a selected sample of ESD's
  - provide facilities (CPU and disk space) for user analysis (~50 users/Tier-2)
  - run simulation and/or calibration procedures

Analysis on Tier-2's and Tier-3's

- This area is under the most active change
  - Capturing resource usage and usage patterns from recent Physics Workshop
- Assume about ~10 selected large AOD datasets, one for each physics analysis group
- Assume that each large local centre will have full TAG to allow simple selections
  - Using these, jobs submitted to T1 cloud to select on full ESD
  - New collection or ntuple-equivalent returned to local resource
- Distributed analysis systems under development
  - Metadata integration, event navigation, database designs are all at top priority
  - ARDA may help, but will be late in the day for DC2 (risk of interference with DC2 developments)

Resources needed for first year

- Assuming 100 days in first run, spread over year break 2007/08:

Summary of Resources Requirements

	Raw + Cal	CERN T0	CERN T1	CERN (tot.)	Each T1	Total
Autom.Tape (TB)	3216	800	60	4076	233	**5476**
Shelf Tape	2816	0	0	2816	0	**2816**
Disk (TB)	40	800	427	1267	421	**3791**
CPU (MSI2k)	0	2.5	2.7	5.2	0.96	**10.9**

- This does not include the 1/6 RAW at each external T1 now being considered

DC2: test of the Computing Model (1)

- ATLAS Data Challenge 2 will be the first real test of the Computing Model.
- April-June 2004 (Phase 1):
  - simulation of >10^7 events ("one day of data-taking in ATLAS") with Geant4, output in POOL
  - pile-up and digitization in Athena, output in POOL and in ByteStream (Raw Data) format
  - copy of the ByteStream to CERN "Tier-0"
- June-July 2004 (Phase 2):
  - "prompt" reconstruction in real time of events on Tier-0, distribution of ESD's to Tier-1's
    - this phase will need a dedicated facility
  - target rate is 10% of final, i.e. reconstruct all events in a continuous operation running for 10 days

DC2: test of the Computing Model (2)

- July-September 2004 (Phase 3):
  - distributed analysis of reconstructed data
  - test and use as much as possible of Grid tools and infrastructure available at that point in time (LCG-1/2/3..., NorduGrid, US-Grid, (ARDA?) ...)
  - this is where the Grid becomes necessary
    - earlier phases were organized distributed productions
- End 2004: Computing Model Document for the LCG Project
  - basis for LCG TDR
  - also for Computing MoU's
- Spring 2005: LCG and ATLAS Computing TDR

2005 and beyond

- 2nd half 2005: DC3
  - larger scale operation
  - full s/w chain including Trigger and Event Filter
  - complete calibration/alignment algorithms and infrastructure
- mid-2006: start of commissioning run
  - this is real life!
  - rate from cosmic ray trigger can be comparable to real collisions
    - need CPU's for event processing
  - useful for preliminary calibrations and alignments of detector as built
  - full s/w chain must be available, robust, in operation
  - data lifetime may be short
    - not much requested as permanent storage before 2007