ATLAS
Data Management Planning

- Status of present data taking and computing
- Review of short/medium term perspectives
- (remarks on the) Organisation of the medium-long term future of data, access, policies etc.

- implicitely I make use of what Elizabeth presented yesterday on data volumes, CPU, ...

Hans von der Schmitt on behalf of the ATLAS collaboration

Motivation

Physics motivation is obvious - see DPHEP-2009-001 and this workshop:

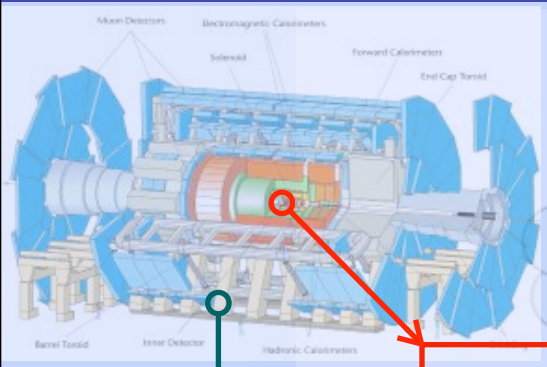
Make sure our data remain available for future research purposes. This payed dividends already when reusing data from former HEP experiments - OPAL, JADE, even Gargamelle (which was easy).

But already today in the early LHC phase, preservation of inputs, results and knowledge is important and has difficult technical and organizational aspects.

- e.g. we have low-pileup data from 2010 which may not come again
- present data taking is for 7 TeV c.m.s. energy which will not come again
- LHC is progressing fast. We will have a complete 7 TeV experiment finished by the 2013-14 shutdown with many fb^{-1} luminosity. Then we will concentrate on 14 TeV.
- LHC events are very rich. Re-use for future science ideas likely... and may have to wait a while for the next accelerator.

This week I'm here to learn, building on DPHEP-2009-001...

Status: Data sources and paths



Simulated data (files)

- *Generated/reconstructed on Tiers, to Tiers*

Event data (files, TAG also on DB)

- *RAW from detector reconstructed to ESD, AOD, TAG, NTUP at T0, to Tiers; re-reconstructed at Tier1s, to Tiers*

non-event Metadata (DB, Pool files; in event files)

- *trigger conditions (in Trigger DB)*
- *running conditions / data quality (in COOL DB)*
- *luminosity, backgrounds (COOL)*

- *software (SVN)*
- *software tags used in processing (AMI DB)*
- *less-structured input (logbooks, doc, knowledge)*

Physics analyses

Publications

Status: Policy

- Excerpts from some of the ATLAS policy documents (here, regarding publications):

  "...

  Any physics result published by ATLAS must be reproducible by the collaboration at any point in time *without the need to obtain information from the original authors* of the analysis.

  ...

  The code used to operate on the tertiary DPD and produce the final histograms used for publication should be publicly available in the ATLAS CVS repository, validated and *should be simple enough* for the analysis reviewer to reproduce the histograms.

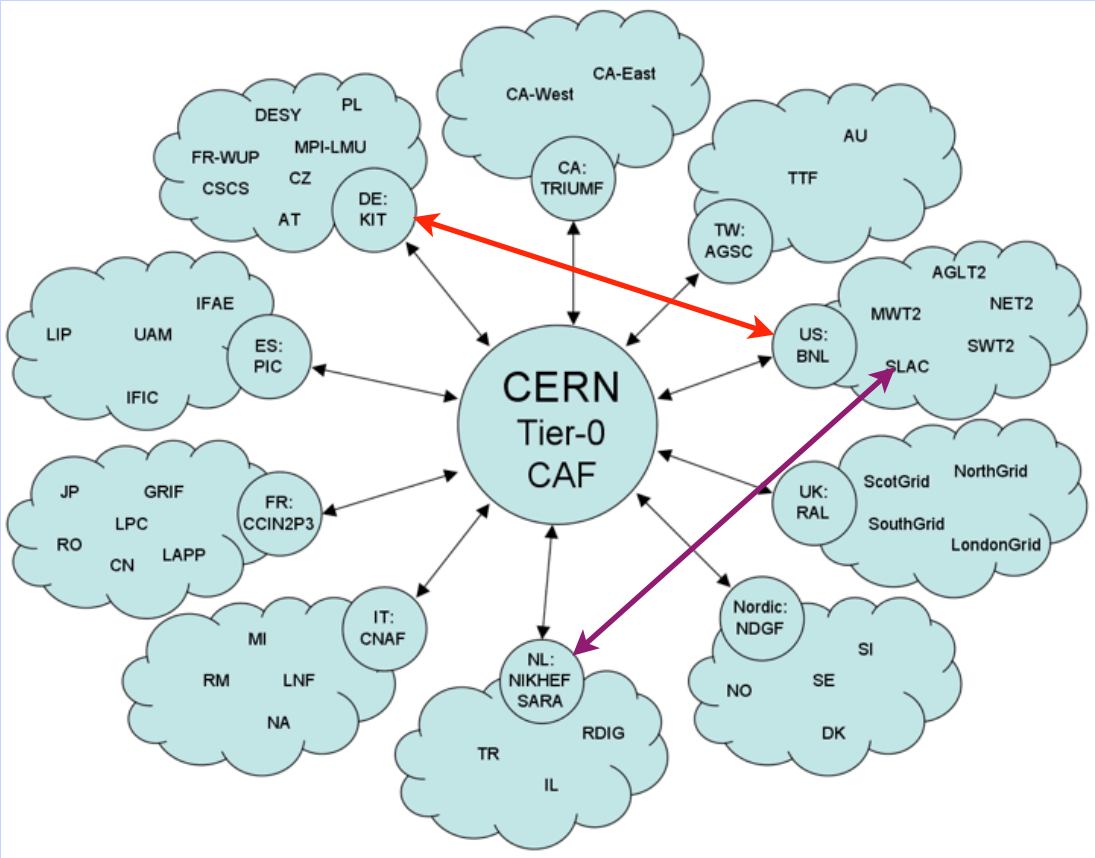
  ..."

- If we generalize such policy to cover also future analyses, after the active ATLAS lifetime, we will be in good shape for DPHEP goals
  - of course, quite a few technical issues to solve until we are there !
    - many data sources enter an analysis and need to be preserved
    - software usage isn't really trivial yet
  - as well as many organizational issues.

Status: Data

- Custodial data:
  - archive copy, must always be kept
  - e.g. for RAW data - two copies on tape, distributed
- Primary data:
  - must be kept while in this state but may be re-classified as secondary
  - e.g. AOD, DESD from most recent (re)processings
  - reprocessing reasons: new sw release, new calibration, new data quality assessment
- Secondary data:
  - may be deleted when space is needed (notification, possibly archiving)
  - e.g. derived data from older (re)processings
  - some fraction of RAW data on disk
- Event sizes, annual volumes for a single copy:
  - RAW not too large, compressed to 0.8MB/ev, O(1PB)/yr
  - AOD still high with pileup, aim at the 150kB/ev of our Computing Model, O(200TB)/yr
  - TAG ~2kB/ev, O(3TB)/yr: ~300 variables + pointers to full event data, easy access
  - note: LEP/OPAL's lifetime data volume was some O(100TB).

Status: Computing on the Grid



- Hierarchy no longer as strict as originally conceived:
  - T1 - T1 traffic
  - some T2s can connect to T1s of different cloud
  - Dynamic data placement T1 to T2, also T1 to T1: no longer pushing final #copies to T1s and T2s - only minimal #copies from T0 to T1s. And dynamic deletion!

Status: Software

- Mostly c++, and a lot of Python as well
- Organized in Projects (e.g. core, reconstruction, trigger, analysis) with some 2000 Packages in total
- Packages contain c++ headers, c++ sources, python scripts, requirements file which steers build process
- Stored/versioned in SVN, using AMI for tagging
- Nightly builds - several in parallel
- A few widely distributed releases per year
- Documentation: partly in-line using Doxygen, partly separate, much use of TWikis

- Subset of the projects is used in the High Level Trigger

Some collected remarks

Collection of remarks/concerns raised during discussions in ATLAS, preparing for this workshop:

here, regarding the present analysis workflow and the preservation of analysis code

- Need to cross-check earlier analyses
  - the policy is there (slide 3) but is not enforced throughout yet
- Examples for building upon earlier studies exist
  - run new signal through existing analysis to reinterpret the result
  - extend similar analysis to include more luminosity
  - use as starting point for similar analyses
- Problems with preserving an analysis
  - requires preservation of a given DPD format, but no central solution (SVN) for that yet
  - need to keep track of AMI tags used for simulation, reconstruction, DPD making, of the Good Run Lists (data quality related), and of event selections
  - how to preserve systematic studies done in multiple steps by many people
- ...and much more.

Short/medium term perspectives

- Grid computing will be simplified
  - R&D ongoing to move towards Cloud computing, to use fewer HEP specific ingredients
  - more R&D - e.g. transition to noSQL where Oracle hits scalability limits
- Reconstruction/analysis software needs to be simplified...
  - for maintainability
  - for easier use
  - hope to do this "as if" we were going for Preservation Level 4
  - aim at SW improvements available for the 2013/14 shutdown
- ...and speeded up
  - identify opportunities for substantial speedup in our framework and application code
- Use smaller set of storage types for input data
  - we are using relational and other DBs for metadata, and several specific systems for event data
  - R&D is underway to simplify both areas
- Organisational
  - check availability of Data Conservation Centres to experiment, use experience of libraries
  - discuss archivist position
  - discuss open access... to check ease of access, 2000 physicists are already a good test
    - ease of TAG access gives some hope

Medium to long term future

- Confident that the technological progress should make reliable long-term storage of multi-PB volumes a non-issue
- CPU required to re-analyse from RAW (ideally; else from derived data) or to re-simulate events in ATLAS, should be abundant/affordable
  - probably only if software can run on parallel architectures
  - however, fancy parallelism may may break on future HW architectures
  - so we may stay with athenaMP: fork in event loop to share memory, copy-on-write
- What about software technology ?
  - languages - will future c++ compilers and Python still digest ATLAS sources
  - operating systems - will today's virtual machines still run on future platforms
- ...and specific HEP applications ?
  - will present-day MC generators and the complex ATLAS geometry still run
- Already learned a lot during this workshop (InSpire, support given by computing centres, continuous test/update of SW). Hope to learn more...