Monte Carlo production for the CLIC CDR

Our experience

S. Poss, C.B. Lam, J. Strube, C. Grefe

CERN

September 8, 2011

Table of contents

Introduction

- CLIC CDR is submitted to reviewers!
- 6 benchmark analysis:
  - $e^+e^- \to h\nu_e\overline{\nu}_e$,
  - $e^+e^- \to H^+H^-$, $e^+e^- \to H^0A$,
  - $e^+e^- \to \widetilde{q}_R\overline{\widetilde{q}}_R$,
  - $e^+e^- \to \widetilde{\ell}\overline{\widetilde{\ell}}\,(\ell = e, \mu)$,
  - $e^+e^- \to \widetilde{\chi}^{\pm}_i\widetilde{\chi}^{\mp}_j$, $e^+e^- \to \widetilde{\chi}^0{}_i\widetilde{\chi}^0{}_j$,
  - $e^+e^- \to t\bar{t}$ (500 GeV).
- Plus all the backgrounds

Total number of events generated in the last 6 months: $17 \cdot 10^6$, events processed: $\sim 100 \cdot 10^6$.

Part I

Framework

Production framework

DIRAC has a Transformation System:

- Create a workflow object (XML representation of a job), and let the system create the jobs for you
- Automatic resubmission of failed jobs: minor monitoring needed

It saved us a huge amount of time and man power.

Production framework

The DIRAC transformation system comes with a useful set of API commands: get properties, change them, update inputs, etc.

Full support for the ILD and SiD software frameworks, plus support for WHIZARD and PYTHIA code. Additionnaly we have support for generator level cuts (StdhepCut tool from Lars Weuste), Overlay input, etc.

Software now installed in the Shared areas of the sites when/where possible. Does not use SAM framework (yet) as DIRAC takes care of the software installation and removal.

Defining steps

- Due to uncertain CPU usage, it was decided to split all steps in different jobs
- Implies dedicated productions for WHIZARD/PYTHIA, Mokka/SLIC, Marlin/LCSIM-SLICPandora-LCSIM
- Linking of productions is straightforward, more on that later

Replica catalog

- Use of LFC is somewhat limited in performance sometimes (more true for the LHC though)
- DIRAC provides a catalog to overcome this
- Has also the possibility to write and read in the LFC via the same interface.

In the futur (coming month), when adding a file in the DIRAC FC, it will also end up in the DESY LFC. It would be nice that the files be always added in both catalogs.

Metadata catalog

- ILC stores the metadata encoded in the file names.
- Somehow restricts the number of meta info that one can add to a given file once the file is created.
- DIRAC FC comes with a metadata catalog, that sets the metadata at directory level (e.g. production ID, software packages, cross section, polarisation, etc.) and at File level (number of events, etc.).
- Flexible as one can add (and remove) metadata on the fly.
- Searchable: find EvtType=tt Datatype=SIM
- Possibility to set ACLs.
- Also comes with convenient API commands.

Web interface is under development, should be available for tests in the coming weeks.

Part II

CLIC Production

File names conventions

- Base path is /ilc/prod to separate from /ilc/user data.
- Path holds most relevant info (machine, energy, process, detector, data type (gen, SIM, REC, DST), prodID)
- File name is process_prodID_jobID.ext (ext=stdhep,slcio)

Defining a production

1. Define the input data (if any) via a metadata query (e.g. meta['ProdID']=186)
2. Define the application and versions you want to use
3. Define the output files (if any)

Paths are resolved according to the input data query.

Steering files are installed on the sites like applications (dependency relation), so no need to pass them in the input sandbox of the job.

Data driven procedure: if new files are added to the catalog that correspond to the input data query, jobs are automatically created.

Monitoring

Monitoring done with web site:

- Overview of production statuses: % of the files processed, failure rate
- Access to jobs a given production in one click
- Monitoring and accounting plots also available

Overlay handling

CLIC detector benchmark require to reconstruct the events with overlaid $\gamma\gamma \rightarrow had$. 60 BX with 3.2 interactions per bunch Xings are considered for every signal event.

Files simulated contain 100 events each (more would be better but CPU is a constrain). For a 10 signal events file, one need $\sim$ 2000 bkg events, or 20 files per job.

This did not work: storages (CERN in particular, but also IN2P3) do not cope with such a load (2000 jobs trying to access 20 random files each represent a huge number of queries). Solution: merge the files randomly once to reduce the number of files needed: merge them by packs of 200.

Works but a bit tedious: requires to have all the files locally once. Plus one needs a very large number of files to have enough combinaisons.

Needed solution: use random access (LCIO v1.51 has it, but not LCSIM). Then only one big file is needed on every site with direct access. Then no more transfers are needed.

Data validation

When simulating data with 2 frameworks, it's necessary to validate it before running. For that purpose TOMATO was designed: convert the stdhep to slcio (stdhepjob supported by DIRAC), and run TOMATO (Marlin processor) to create histograms of "significant" distributions.

"Significant" is specific to the different analysis, so every working group needs to provide the variables to be plotted.

Was used for the $t\bar{t}$ analysis.

Part III

Additional considerations

Additionnal considerations

Dropped pilot mode as some sites did not support it, using private pilot mode

Testing running in the US sites: when it runs, we'll need to see for disk space here and there

CERN disk requests will be made: for the moment, only 40Tb are available on disk (as mush as needed on TAPE though).

Part IV

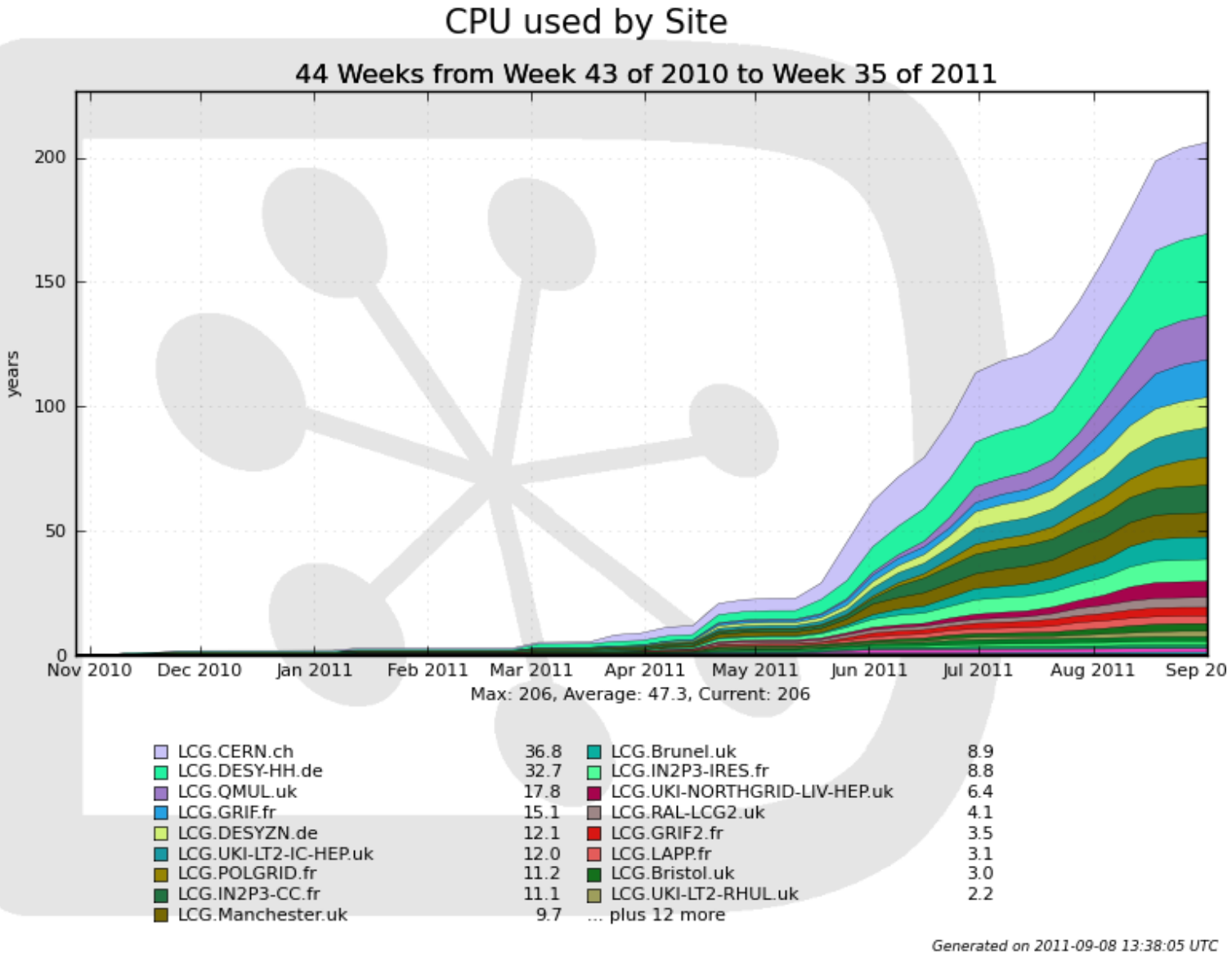
Conclusion

Conclusion

- Produced in 6 month $\approx$ 100 million events in 9 months
- Used $2 \cdot 10^{10}$ KSI2K seconds of CPU in $\sim$ 20 sites
- Data produced correspond to $\sim$ 130 TB
- Most problems were due to storage access (overlay file)
- ILC DBD could benefit from DIRAC usage

At least, the data produced should be stored in both catalogs.

Resources used: CPU



Resources used: Storage

Site	Production (TB)	User (TB)
CERN	128	20
IN2P3	4	9
RAL	4	28
KEK	0.02	0
IMPERIAL	1.6	0
TAU	$4 \cdot 10^{-4}$	0

Sites that can be used in addition: DESY, Bristol, BONN, RALPP

Services availability

100 million events correspond to $\sim$ 4 million jobs. On average 1000 concurrently running.

File catalog is the most expensive service in terms of CPU and simultaneous queries to DB. Had to duplicate service (not DB).

It would be useful to replicate the DB and other services to have more instances available (more VO boxes needed).

Same problem with JobManager with high load.