LHCb QR-3

Umberto Marconi
INFN

MB meeting, 6/11/2007

Computing Model

CERN
(Tier-0 + Tier-1)

Generate RAW data
Reconstruction
Stripping
User Analysis

Tier-1s

CNAF | FZK | IN2P3 | NIKHEF | PIC | RAL

Reconstruction
Stripping
User Analysis

Tier-2s

Monte-Carlo
Production
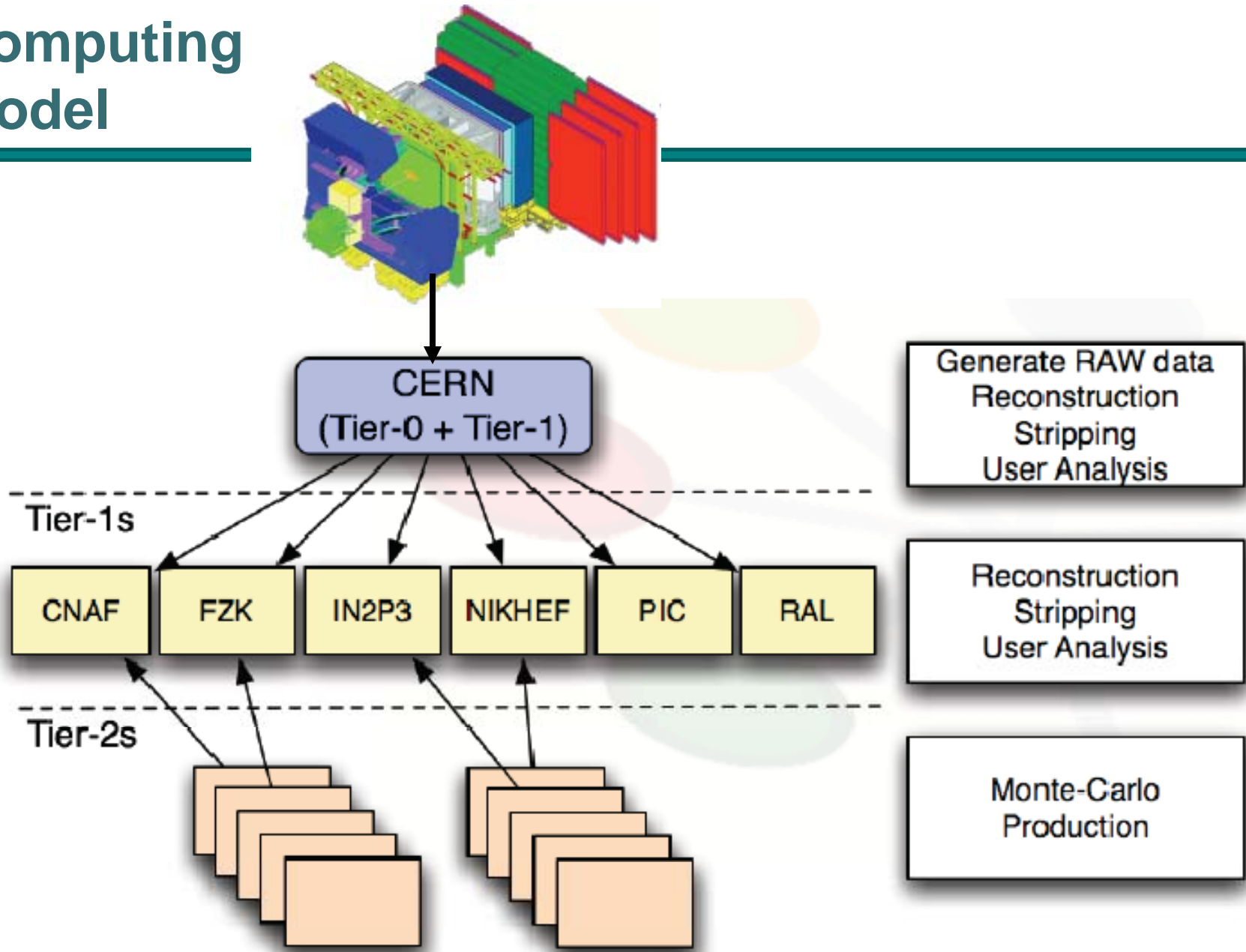
Data flow and distribution

```
@Tier-2

                          Detector (40 MHz)

Simulation                Trigger (2 kHz)

                                RAW

                          Reconstruction

                                rDST

@Tier-1                     Stripping

                             DST, TAG

                          User Analysis

                             Root-tuple
```

CERN   Tier1's

RAW x 2

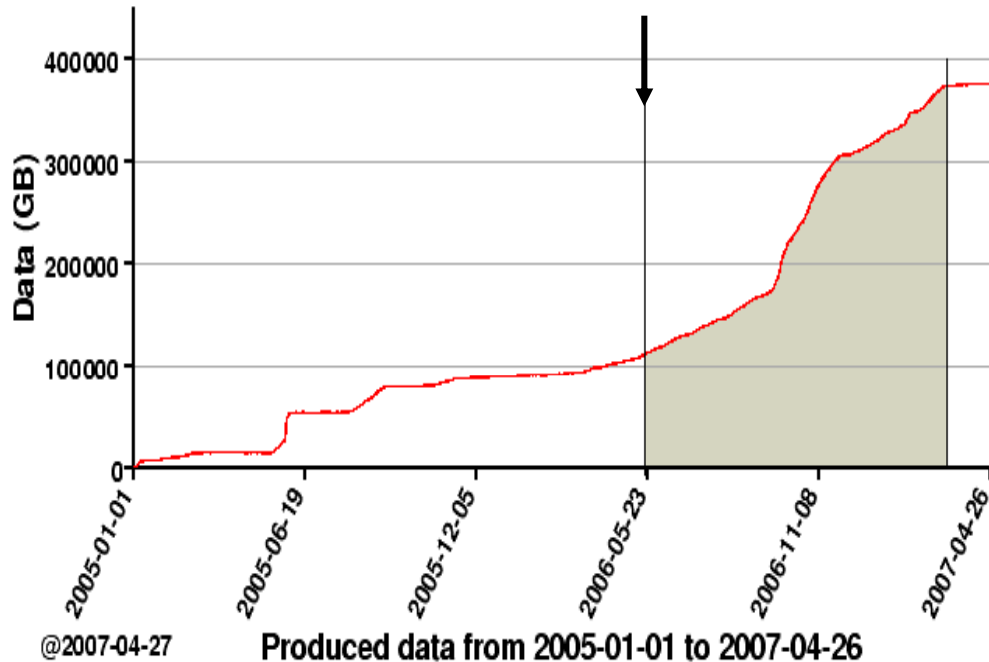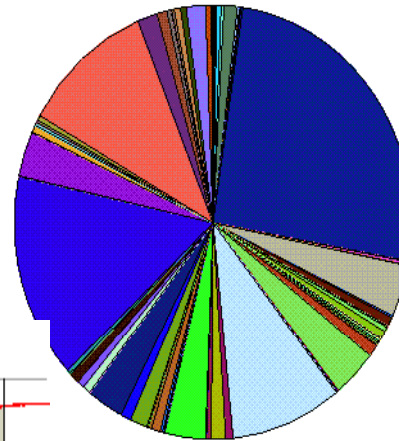rDST

DST x 2

MC x 2

MSS
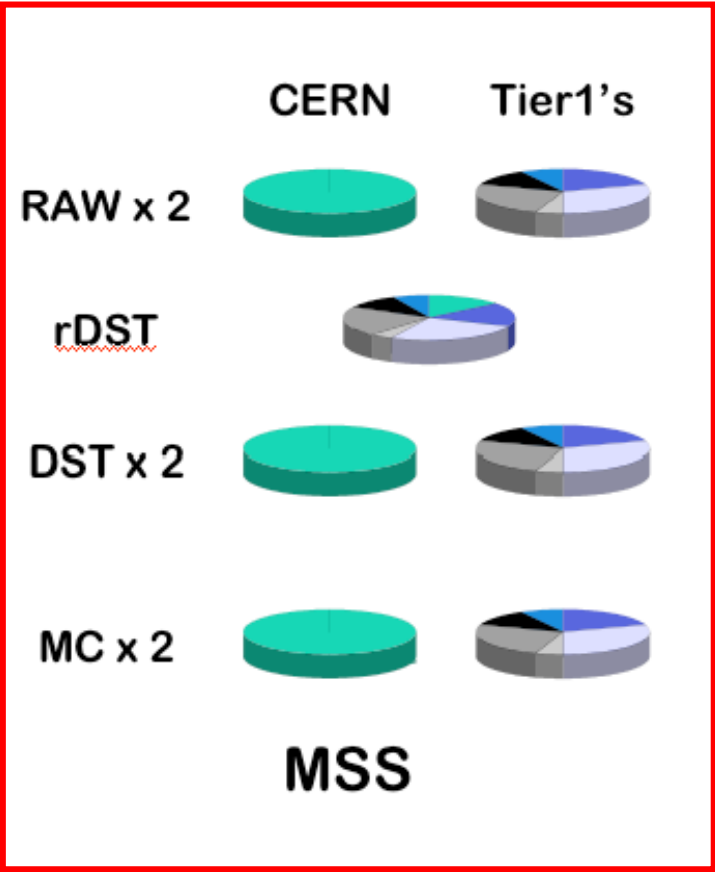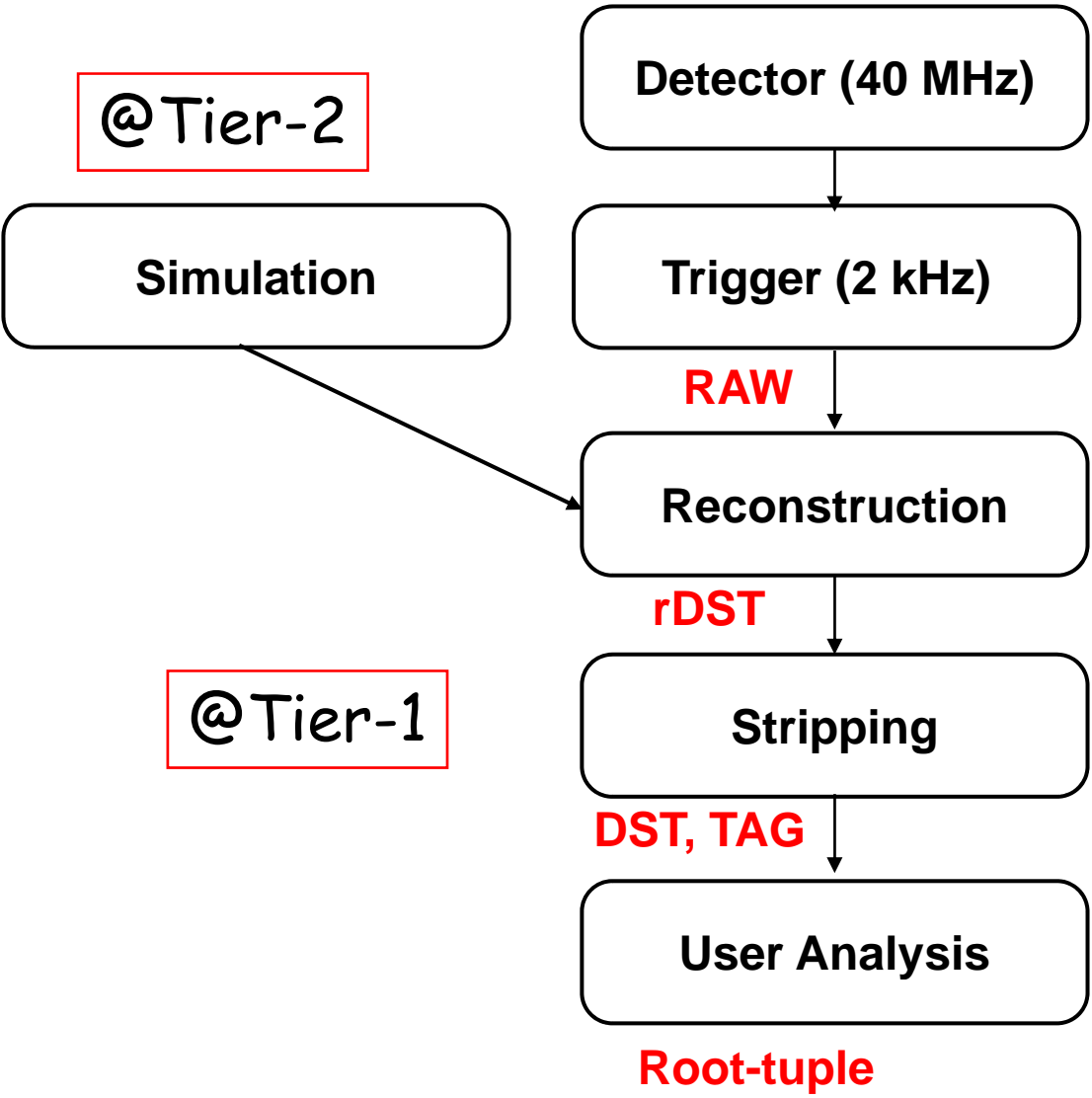
DC'06

During DC06
80% of the LHCb
whole production

10^4 concurrent
jobs running on LCG

DC06 starts : June '06

Total Running Jobs: 9715
DIRAC: 0.29% LCG: 99.71%

DIRAC.Zurich-MH.ch	0.14%	LCG.IRB.hr	0.11%
DIRAC.Zurich.ch	0.14%	LCG.ITEP.ru	0.21%
LCG.ACAD.bg	0.36%	LCG.KFKI.hu	0.73%
LCG.AUVER.fr	0.24%	LCG.KIAE.ru	0.29%
LCG.Barcelona.es	1.08%	LCG.KIAM.ru	0.09%
LCG.Bari.it	0.17%	LCG.Krakow.pl	1.41%
LCG.BHAM-HEP.uk	0.03%	LCG.LAL.fr	0.84%
LCG.Bologna.it	0.01%	LCG.Lancashire.uk	2.99%
LCG.Cagliari.it	0.06%	LCG.LAPP.fr	0.02%
LCG.Catania.it	0.15%	LCG.LISA.nl	0.63%
LCG.CERN.ch	25.42%	LCG.Liverpool.uk	0.61%
LCG.CESGA.es	0.04%	LCG.LPC.fr	0.75%
LCG.CGG.fr	0.27%	LCG.LPN-fails.fr	0.10%
LCG.CNAF-GRIDIT.it	0.04%	LCG.LPN.fr	0.31%
LCG.CNAF.it	3.87%	LCG.Manchester.uk	15.35%
LCG.CNB.es	0.15%	LCG.Napoli-Atlas.it	0.02%
LCG.CPPM.fr	0.69%	LCG.Napoli.it	0.03%
LCG.CSCS.ch	0.12%	LCG.NIKHEF.nl	3.32%
LCG.Dortmund.de	0.36%	LCG.OU.il	0.04%
LCG.Durham.uk	0.61%	LCG.Oxford.uk	0.54%
LCG.Edinburgh.uk	0.04%	LCG.Padova.it	0.24%
LCG.EELA-CIEMAT.es	0.21%	LCG.Pisa.it	0.20%
LCG.ETF-RIH.lv	0.14%	LCG.PNPI.ru	0.37%
LCG.Ferrara.it	0.19%	LCG.QMUL.uk	10.66%
LCG.FESB.hr	0.08%	LCG.RAL-HEP.uk	1.61%
LCG.FORTH.gr	0.89%	LCG.RHUL.uk	0.81%
LCG.Glasgow.uk	3.63%	LCG.SINP.ru	0.34%
LCG.GR-01.gr	0.09%	LCG.Sofia.bg	0.06%
LCG.GR-05.gr	0.10%	LCG.TCD.ie	0.21%
LCG.GRIDKA.de	9.04%	LCG.Torino.it	0.51%
LCG.HG-02.gr	0.59%	LCG.ULAKBIM.tr	0.46%
LCG.HG-06.gr	1.16%	LCG.USC.es	1.55%
LCG.HPC2N.se	0.39%	LCG.WARSAW.pl	0.47%
LCG.Imperial.uk	0.07%	LCG.WCSS.pl	0.11%
LCG.IN2P3.fr	3.31%	LCG.WEIZMANN.il	0.02%
LCG.IPSL-IPGP.fr	0.06%		

8 2007, 08:40

@2007-04-27   Produced data from 2005-01-01 to 2007-04-26

RAW collected at CERN and distributed to the Tier-1s emulating real data taking for reconstruction and stripping at sites

4

Event re-reconstruction and stripping

- ## February 2007 onwards:

  - Events reconstruction at Tier1s of RAW data files no longer on cache, to be recalled from tape.

    - A reconstruction job uses 20 MC RAW data files as input.

  - The rDST data output has to be uploaded locally to the Tier1.

- ## June 2007 onwards:

  - Events stripping at Tier1s.

    - A stripping job uses 2 rDST files as input.

    - Accesses to the 40 corresponding MC RAW files for the full reconstruction of the selected events.

  - DST files have to be distributed to the Tier1s.

Feedbacks

- Reconstruction it is easy for first prompt processing while it is painful for re-reprocessing, when files have to be staged.

  - Jobs are put in the DIRAC central queue only when files are staged.

- Too many instabilities in SEs.

  - Full time job checking availability, enabling/disabling SEs in the DMS.

- Staging at some sites is extremely slow. Problems with the SE software? Problems with the configuration (number of servers, number of tape drives)?

- Some files are not retrievable from tape: registered in our LFC, found using srm-get-metadata but fail to get a tURL (error in lcg-gt).

Feedbacks (II)

- Shortage problems encountered with Disk1TapeX.

  - Need to clean up datasets to get space: painful with SRM v1.

- Not easy to monitor the storage usage:

  - Developed a specific agent reporting every day from LFC.

  - Agents checking integrity between SEs and catalogs.

  - VOBOX helps but needs guidance to avoid DoS.

- Need to establish a protocol to get warning from site to set a flag in LFC indicating the replica is temporarily unavailable (not used for matching jobs).

- On our side it may help to tune the number of stage requests issued in one go trying to optimise the recall from tape.

Feedbacks (III)

- Inconsistencies between SRM tURLs and root access.

- Problems with ROOT finding the HOME directory at RAL, fixed by providing an additional library (compatibility mode on SLC4).

- Unreliability of rfio, problems with rootd protocol authentication on the Grid (now fixed by ROOT).

  - lcg-gt returning a tURL on dCache but not staging files, workaround with dccp, then fixed by dCache.

Tests and Developments

- SLC4 migration

    - Straightforward for LHCb applications. Problems were with middleware clients used by them: dCache, gfal, lfc, etc.

- Essential to test sites permanently with the SAM framework: CE, SE, SRM.

- SRM v2 tests done successfully.

- Several plans for SE migration: RAL, PIC, CNAF, SARA (to NIKHEF).

    - It is a the large effort we have to put, in particular concerning the changes of replicas in the LFC.

- We hadn't the required VOMS set of groups/roles. With a default set there are still difficulties to have the proper mapping, in particular for SGM and PRD. It induces difficulties in LFC registration (impossible for us to modify the internal mapping of DNs and FQANs, having to go through the administrators).

Distributed Analysis
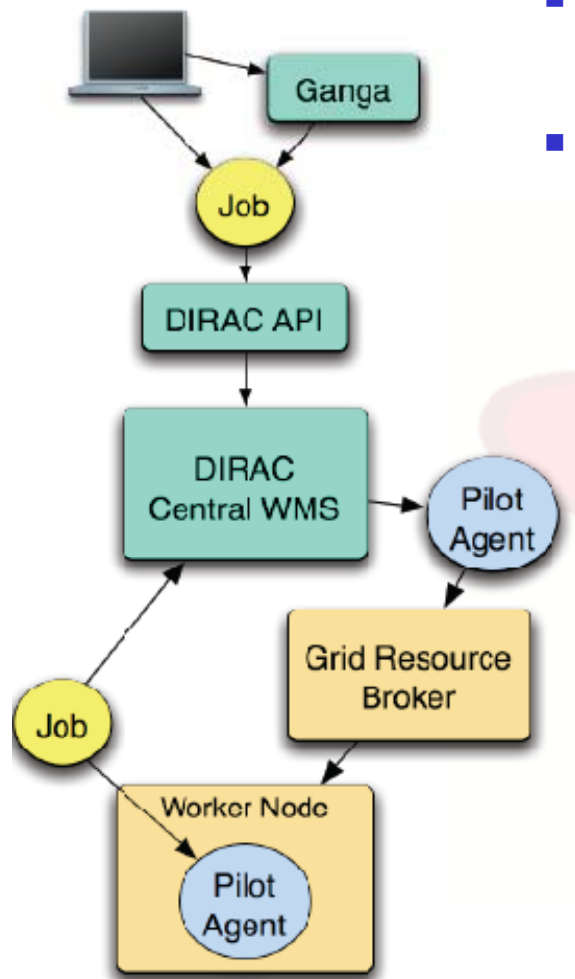
- GANGA is the UI used to submit the analysis job to the Grid.

- Use of the generic agents also with the analysis job, downloaded from the DIRAC central queues.

Analysis job vs time
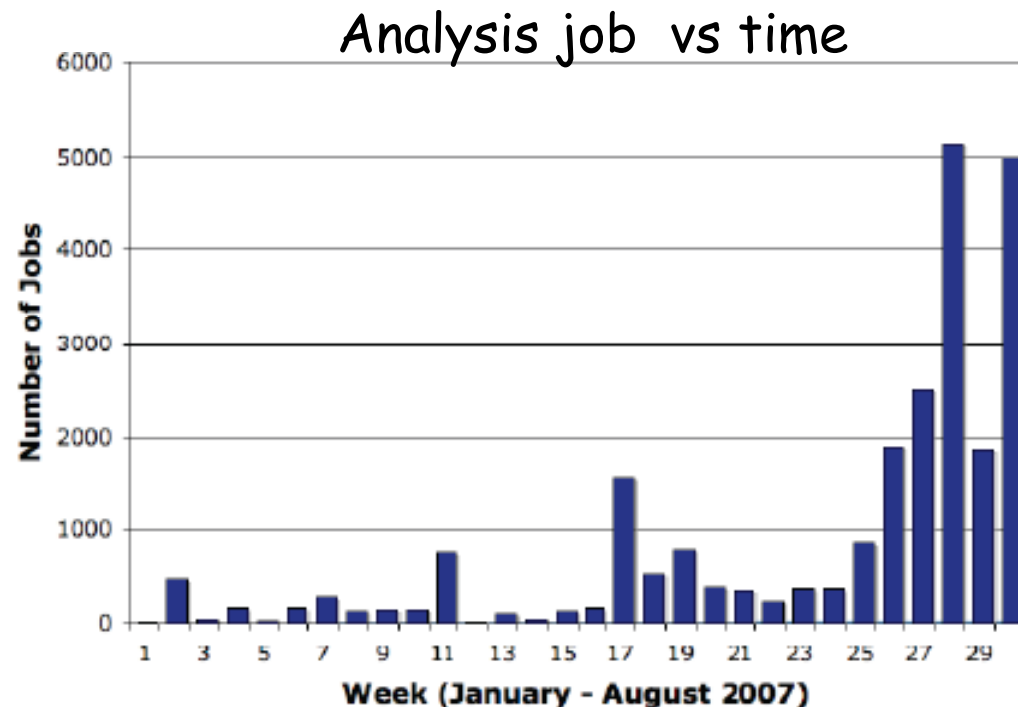
CCRC'08 Aims

- Test the full chain: from DAQ to Tier-0 to Tier-1's.

- Test data transfer and data access running concurrently (current tests have tested individual components).

- Test DB services at sites: conditions DB and LFC replicas.

- Tests in May will include the analysis component.

- Test the LHCb prioritisation approach to balance production and analysis at the Tier-1 centres.

- Test sites response to "chaotic activity" going on in parallel to the scheduled production activity.

CCRC'08 Planned Tasks

- RAW data distribution from the pit to the Tier-0 centre.

  - Use of rfcp into CASTOR from pit to the T1D0 storage class.

- RAW data distribution from the Tier-0 to the Tier-1 centres.

  - Use of FTS. Storage class T1D0.

- Reconstruction of the RAW data at CERN and at the Tier-1s for the production of rDST data.

  - Use of SRM 2.2. Storage class T1D0

- Stripping of data at CERN and at T1 centres.

  - Input data: RAW and rDST on T1D0.

  - Output data: DST on T1D1

  - Use SRM 2.2.

- Distribution of DST data to all other centres

  - Use of FTS - T0D1 (except CERN T1D1)

CCRC'08 Planned Tasks (II)

- Preparation of RAW data will occur over the next few month

    - We need to merge existing MC datasets into ~2GB files

- February activities:

    - Maintain the equivalent of 2 weeks data taking.

- May activities:

    - Maintain equivalent of 1 month data taking.

    - Run fake analysis activity in parallel to production type activities using generic agents.

    - Generic agents are the LHCb baseline. It needs to be integral part of CCRC'08.