LHCb: CCRC'08 Update

Nick Brook

University of BRISTOL
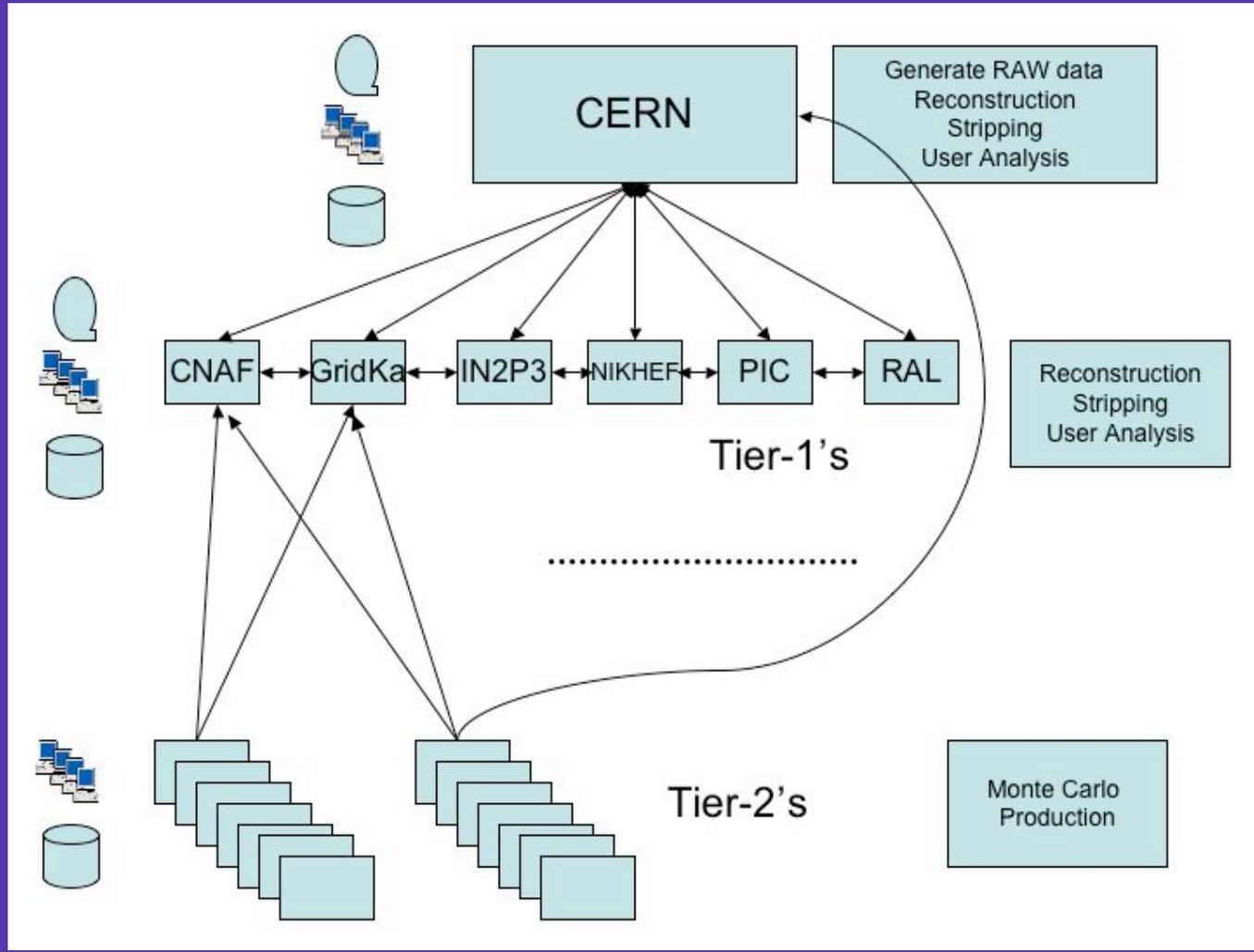
Use of computing centres

Main user analysis supported at CERN+6 "Tier-1" centres

Tier-2 centres essentially Monte Carlo production facilities

Plan to make use of LHCb online farm for re-processing

Planned tasks

- Raw data distribution from pit $\rightarrow$ T0 centre
  - Use of rfcp into CASTOR from pit - T1D0
- Raw data distribution from T0 $\rightarrow$ T1 centres
  - Use of FTS - T1D0
- Recons of raw data at CERN & T1 centres
  - Production of rDST data - T1D0
  - Use of SRM 2.2
- Stripping of data at CERN & T1 centres
  - Input data: RAW & rDST - T1D0
  - Output data: DST - T1D1
  - Use SRM 2.2
- Distribution of DST data to all other centres
  - Use of FTS - T0D1 (except CERN T1D1)

All tasks envisaged during data taking in 2008

Planned tasks

- February activities were
  - Maintain equivalent of 2 weeks data taking
  - Assuming a 50% machine cycle efficiency
  - Used quasi fake RAW files of ~2 GB
- May activities
  - Maintain equivalent of 1 month data taking
  - Assuming a 50% machine cycle efficiency
  - Run fake analysis activity in parallel to production type activities

Activities across the sites

- Breakdown of processing activities (CPU needs)

Site	Fraction (%)
CERN	14
FZK	11
IN2P3	25
CNAF	9
NIKHEF/SARA	26
PIC	4
RAL	11

NB: Full production activities envisaged plus analysis activities

Pit -> Tier 0

- Use of rfcp to copy data from pit to CASTOR
  - rfcp is the recommended approach from IT
  - Will have 25 TB capacity available online
  - A file sent every ~30 sec
  - Data remains on online disk until CASTOR migration
  - Rate to CASTOR - ~70MB/s

Tier 0 -> Tier 1

- FTS from CERN to Tier-1 centres
  - Transfer of RAW will only occur once data has migrated to tape & checksum is verified
  - Rate out of CERN - ~35MB/s averaged over the period
  - Registration in LHCb bookkeeping & LFC occurs at same time as request to remove data from online disk is issued

Tier 1 <-> Tier 1

- FTS used to transfer data between from CERN/T1 to CERN/T1 of stripped files
  - LHCb mechanisms in place to do this but never been used in anger using FTS
    - This operation not tested during February phase

Data Access

- Will need SRM 2.2 SE with correct space tokens
  - LHCb space tokens are:
    - LHCb_RAW (T1D0)
    - LHCb_RDST (T1D0)
    - LHCb_M-DST (T1D1)
    - LHCb_DST (T0D1)
    - LHCb_MC_M-DST (T1D1)
    - LHCb_MC_DST (T0D1)
    - LHCb_FAILOVER (T0D1)
    - LHCb_USER (T0D1)
  - Space token names should not be encode in namespace at site
- See comments on experience later...

Databases

- Conditions DB at CERN & Tier-1 centres
  - No plans to test replication of conditions DB Pit $\leftrightarrow$ Tier-0 (and beyond) during CCRC08
    - Standalone tests at a later date
- LFC
  - Use "streaming" to populate the read-only instance at T1 from CERN

May's Activities

- 84 TB of data from pit to CERN T0
  - Corresponding ~52k files
- Same 52k RAW files from CERN to be distributed over T1 centres
- 14% of rDST production at CERN, remaining 86% at T1 centres (see table on earlier <u>slide</u>)
  - LHCb responsibility to ensure unique files are recons across CERN & T1 centres
  - Additional ~52k (rDST) files produced (integrated across all sites)
  - Corresponds to an additional 42 TB of data

May's Activities

- Stripping on rDST files
  - 17.5k DST files produced during the process (and stored on T1D1) - corresponds to 16TB of data
  - All files are distributed to other sites
    - 7x17.5k files
    - 7x16 TB
- Total number of jobs accessing the data
  - Recons: ~52k
  - Stripping: 17.5k

Amount of data (per storage class)

Storage (TB)	LHCb_RAW (T1D0)	LHCb_RDST (T1D0)	LHCb_M-DST (T1D1)	LHCb_DST (T0D1)	LHCb_USER (T0D1)
CERN	84	6	16	0	3.3
FZK	10.5	4.5	1.7	14.3	1.3
IN2P3	24.5	10.5	4	12	2.9
CNAF	9.3	4	1.6	14.4	1.1
NIKHEF	25.4	10.9	4.2	11.8	3
PIC	4.2	1.8	0.7	15.3	0.5
RAL	10.5	4.5	1.7	14.3	1.3

Amount of data/site

Storage (TB)	T1D0	T1D1	T0D1	Tape	Disk
CERN	90	16	3.3	106	19.3
FZK	15	1.7	15.6	16.7	17.3
IN2P3	35	4	14.9	39	18.9
CNAF	13.3	1.6	15.5	14.9	17.1
NIKHEF	36.3	4.2	14.8	40.5	19
PIC	6	0.7	15.8	6.7	16.5
RAL	15	1.7	15.6	16.7	17.3

CPU Needs

1 event is ~2.4kSI2k.s to reconstruct

1 file $\equiv$ 50k events: 120 MSI2k.s

CPU(kSI2k.days)	Recons	Stripping	Total
CERN	10417	2026	12443
FZK	7739	1505	9244
IN2P3	18172	3534	21706
CNAF	6873	1337	8210
NIKHEF	18884	3672	22556
PIC	3096	602	3698
RAL	7739	1505	9244
Total	72920	14181	87101

No CPU efficiency factor of 0.85 included a la TDRs

Nos of jobs/site

	Total Jobs			Simultaneous jobs		
	Recons	Strip	Total	Recons	Strip	Total
CERN	7500	2500	10000	268	90	358
FZK	5600	1900	7500	200	68	268
IN2P3	13100	4400	17500	468	158	626
CNAF	4900	1600	6500	175	58	233
NIKHEF	13600	4500	18100	486	161	647
PIC	2200	700	2900	79	25	104
RAL	5600	1900	7500	200	68	268
Total	52500	17500	70000	1875	625	2500

Job Details

Job type	Duration (hrs on 2.2kSI2k)	i/p files (from tape buffer)
Recons	24	1x~1.6 GB
Strip	6	3x~1.0GB +3x~1.6 GB

Cache for tape

Cache(TB)	LHCb_RAW (T1D0)	LHCb_RDST (T1D0)
CERN	1.8	0.5
FZK	1.3	0.4
IN2P3	3.1	0.8
CNAF	1.2	0.3
NIKHEF	3.2	0.8
PIC	0.5	0.2
RAL	1.3	0.4

This is for processing - CERN needs cache adding for pit transfer

CERN CASTOR cache for data taking

- Guaranteed 8 hour migration time

  - More likely 3 hours

- 2 TB cache

SRM & Data Access

- Continue to be plagued with data access from SE's
  - Not in general an issue with SRM …
  - … seem more of an issue at dCache sites
    - Though CASTOR sites weren't immune to problems
- Question to be addressed:
  - Is data access directly from SE feasible?
    - Probably not currently
    - Fighting these issues for 2 years - data taking just around the corner
  - Test copying data locally in May phase
    - Timescale for other solutions too tight?
      - Mechanism in place for LHCb
      - Site configurable
    - Estimate need 15 GB per core on the WN

DIRAC3 & Feb CCRC08

- CCRC08 will run under a new version of DIRAC
  - Major factorisation and upgrade of code
  - Major lessons learnt during Feb CCRC'08
- Critical items to fix/check
  - Data integrity checking
    - Problems with file size checking lcg utils
  - Thorough checking of all space tokens @ all sites
  - Check number of simultaneous connections to SE at sites
  - Production management tools
    - Including extensive testing of workflow definitions
  - Porting of stager service to DIRAC3
  - Failover request management
  - Improved error handling
  - Operation tools for non-experts
    - Monitoring: both production, services & site improved
  - Interface of GANGA to DIRAC3
- Timescale for testing - extremely tight!

Summary of Feb'08 CCRC08

- Focus on what need to be improved
  - Pre-testing of whole chain
  - Better involvement of non-experts/production team
  - Data access remains major concern
  - Pro-active planning and better use of meetings
- Feb CCRC08
  - In the circumstances some notably successes
  - Data transfer pit-T0-T1
    - All went extremely smoothly
  - Last week of challenge recons ramped up to several 100's of jobs running concurrently across all Tier-1's

Middleware

- ## Glexec
  - Were planning to during May phase - early statement in CCRC08 planning
    - Won't be rolled out in time
  - Mechanisms in DIRAC tested out during Feb with "production" role
    - Further test of mechanisms during May phase

DIRAC WMS with generic Pilot Agents
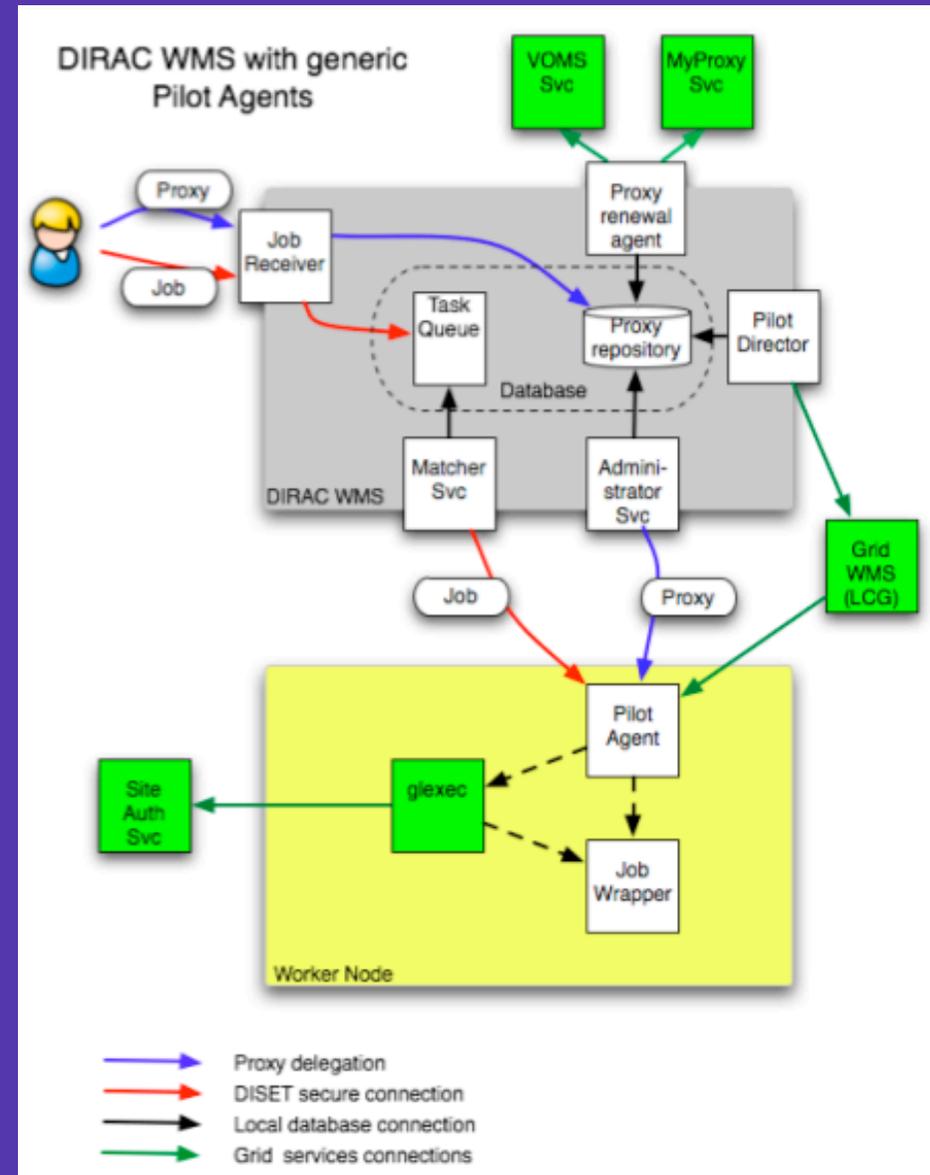
Calendar of Activites

	Week-2 (21st Apr)	Week-1 (28th Apr)	Week0 (5th May)	Week1 (12th May)	Week2 (19th May)	Week3 (26th May)
Raw data distribution	Low scale full chain test					
Re-construction	Data copied to WN; conditions DB access					
Stripping	V low level tests started	Castor sites dCache sites				
DST dist^{bn}		(part of stripping tests)				
Analysis	Testing of GANGA/DIRAC3					

Summary

- Resource requirements updated
- LHCb readiness
  - Plan to meet challenge as originally outlined in 2007
  - Very concerned about data access
  - Timescale extremely aggressive (again!)

  https://twiki.cern.ch/twiki/bin/view/LHCb/CCRC08