
t2k.org
grid computing: past, present and future

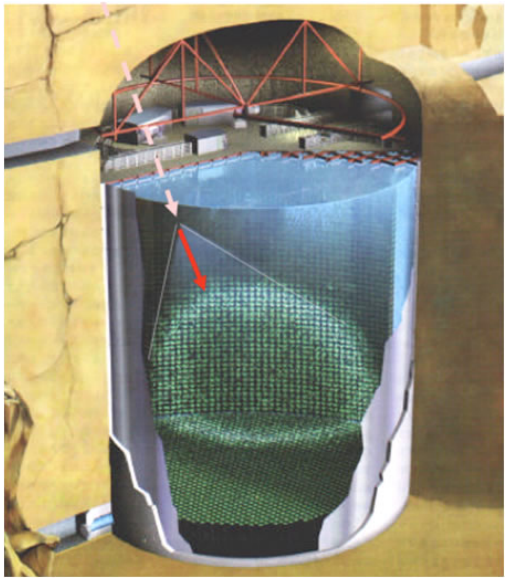
The University Of Sheffield.

The T2K experiment

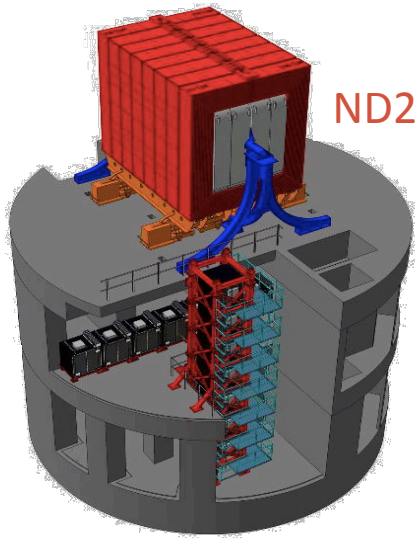
Tokai to Kamioka

$\nu_\mu \rightarrow \nu_e$ $\nu_\mu \rightarrow \nu_\mu$

~500 Collaborators
59 Institutions
11 Countries

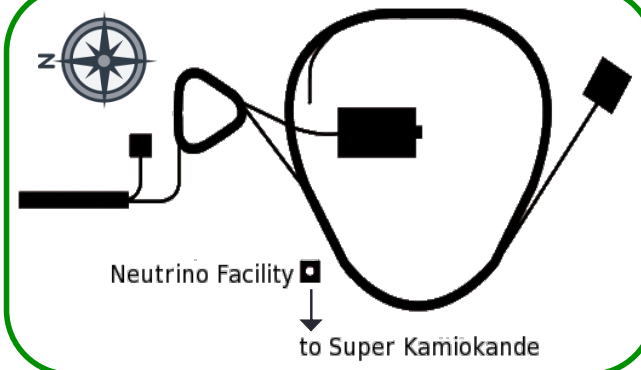


Super-KAMIOKANDE

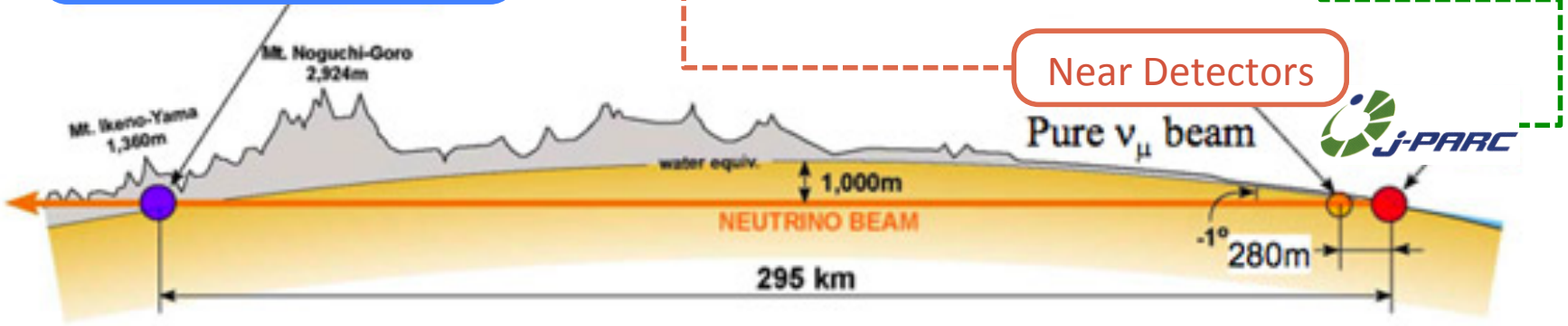


ND280

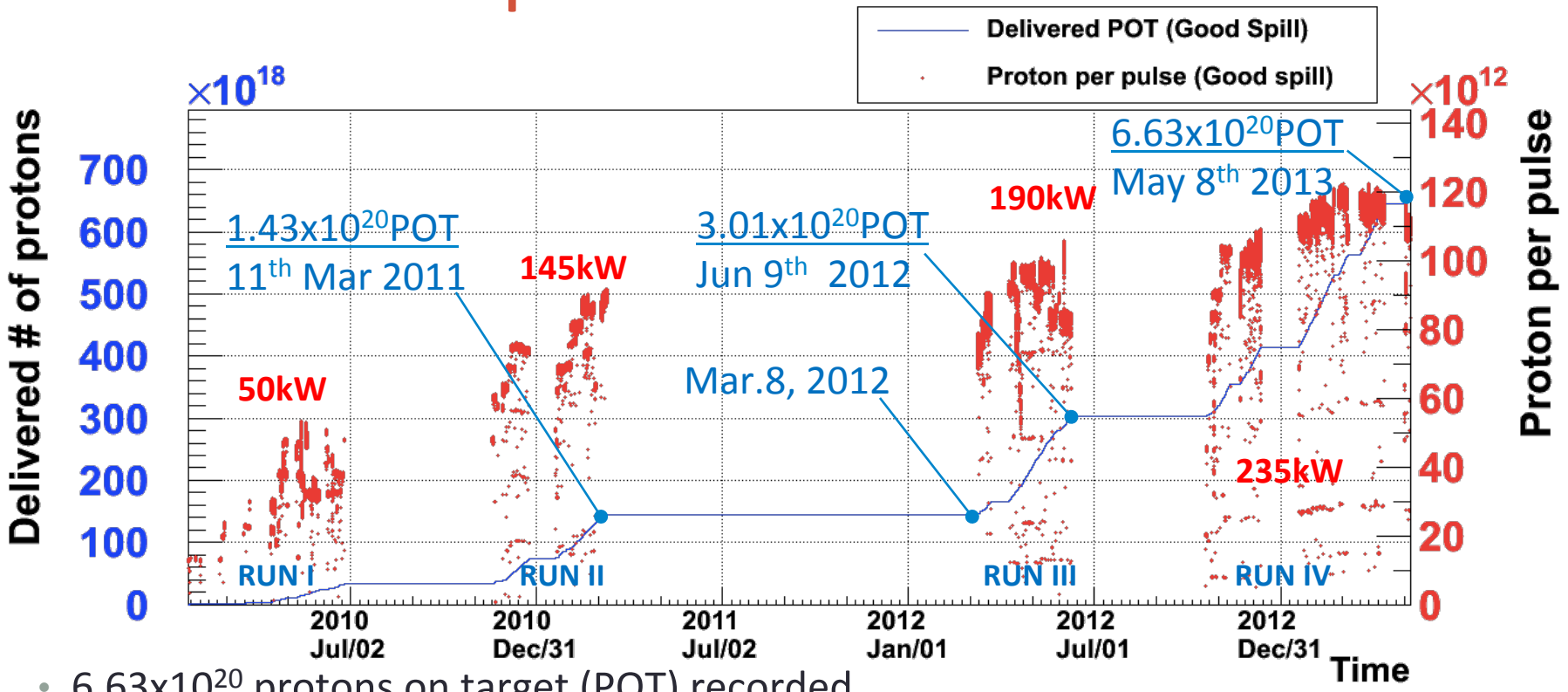
INGRID



Near Detectors



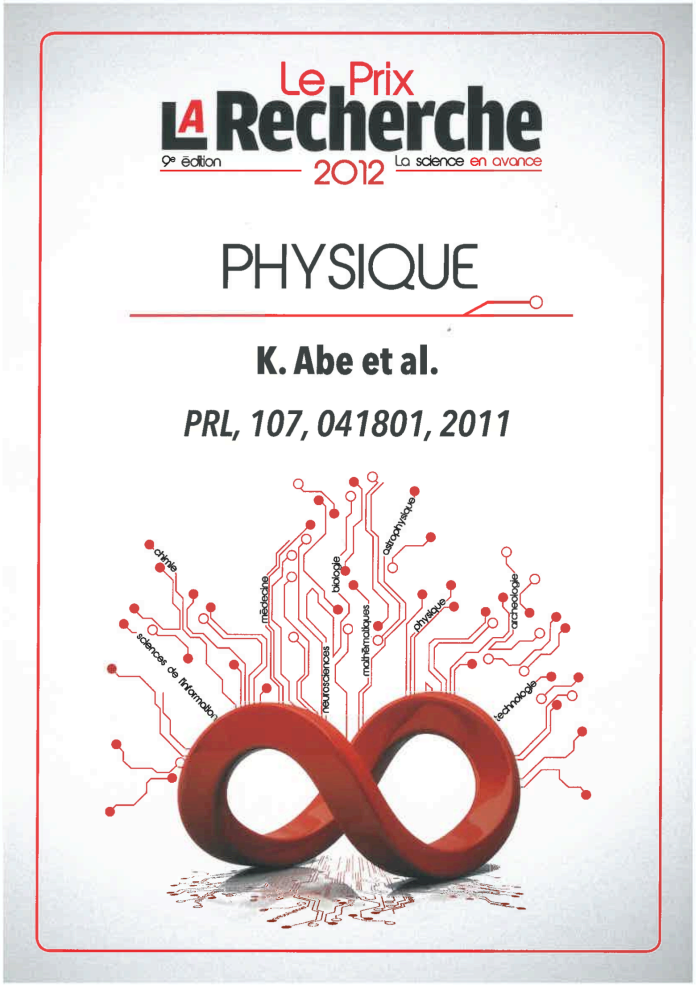
JPARC Beam performance



- 6.63x10^{20} protons on target (POT) recorded
  - <1mrad (~16MeV [2%]) beam stability for total period
- 8% of design goal POT so far
  - Achieved 1.2x10^{14} protons per pulse (WR)
- Summer shutdowns (aircon/humidity)
  - Great Tohoku Earthquake

The University Of Sheffield.

T2K recognition



- 2013 Suwa Prize awarded to T2K beam group
- T2K ν_e appearance paper recognised in Discovery Magazine's top 100 science stories of 2013
- T2K's first observation of ν_e in a pure ν_μ beam #7 in Physics World's top 10 breakthroughs in 2011

The University Of Sheffield.

arXiv.org > full text search "T2K AND GridPP"

arXiv.org Full Text Search Results

Displaying hits 1 to 3 of 3. Reorder by score.

***T2K** Collaboration: K. Abe, J. Adam, H. Aihara et al., Measurement of the neutrino-oxygen neutral-current interaction cross section by observing nuclear de-excitation γ-rays (2014)*
abstract: ... at the Super-Kamiokande water Cherenkov detector. We use **T2K** data corresponding to $3.01 ...
http://arxiv.org/abs/1403.3140; Indexed Mar 13, 2014

***T2K** Collaboration: K. Abe, J. Adam, H. Aihara et al., Measurement of the intrinsic electron neutrino component in the **T2K** neutrino beam with the ND280 detector (2014)*
abstract: The **T2K** experiment has reported the first observation of the ... neutrinos originating from muon **and** kaon decays. In **T2K**, this component is expected to represent 1.2% of the total ... fluxes **and** neutrino cross section modeling used for **T2K** neutrino oscillation analyses. Electron neutrinos coming ...
http://arxiv.org/abs/1403.2552; Indexed Mar 11, 2014

***T2K** Collaboration: K. Abe, J. Adam, H. Aihara et al., Precise Measurement of the Neutrino Mixing Parameter \theta_{23} from Muon Neutrino Disappearance in an Off-axis Beam (2014)*
abstract: New data from the **T2K** neutrino oscillation experiment produce the most precise ... set corresponding to 6.57 x 10^{20} protons on target, **T2K** has fit the energy-dependent nu_mu oscillation probability ...
http://arxiv.org/abs/1403.1532; Indexed Mar 6, 2014

"Computations were performed on the LHC computing grid (LCG), supported in the UK by GridPP, and the supercomputers at the SciNet HPC Consortium."

The University Of Sheffield.

The (worldwide) t2k.org VO

- Use several GRIDs/farms
  - **LCG** [EU (mainly UK resources)]
  - **SciNet/WestGRID/Bugaboo** [USA and Canada]
  - **University of Colorado Batch** [USA]
  - … […]
- No sharing/transparency of EU and North American resources
  - Interface limited to ~mirroring between RAL and TRIUMF
  - Separate EU, American processing and MC productions
    - presently a user with a certificate from an EU CA cannot submit to North American queues

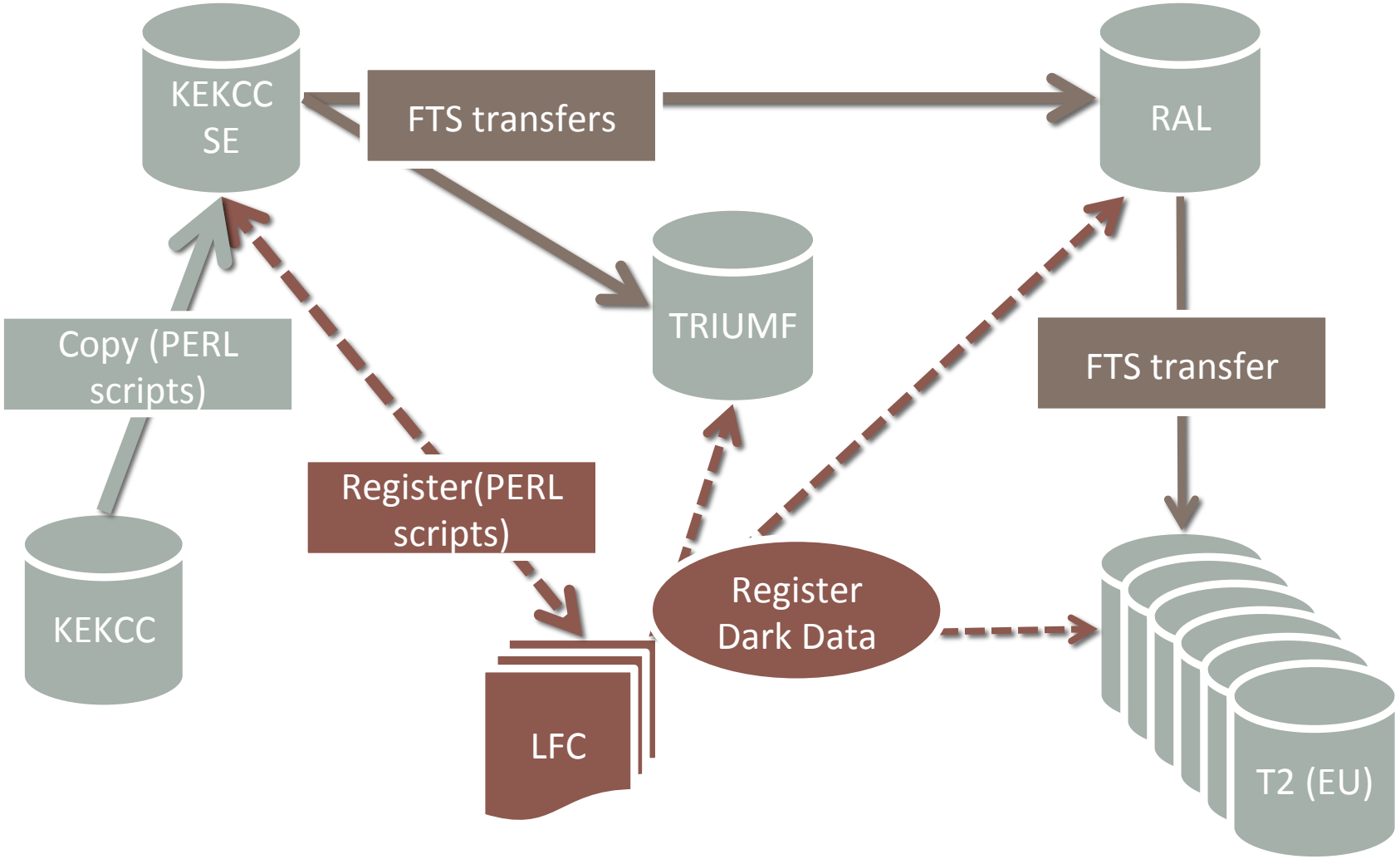
The University Of Sheffield.

t2k.org GRID model

- Storage
  - **T0** KEK *(data source, no GRID storage, offline data archive)*
  - **T1** RAL/TRIUMF *(primary archives [~mirrored])*
  - **T2** the rest *(cache data for/from processing only)*
- CPU
  - at **T1** and **T2**
  - s/w installed via GRID (build) job ahead of processing
    - CVMFS roll out almost complete – bringing more UK T2s online
- Submission/Monitoring/Book-keeping
  - use gLite out of the box (wrapped in python homebrew)
  - rely entirely on LFC to locate data
  - rely entirely on WMSs for submission
  - no dedicated monitoring interface
    - do have python scripts to query job statuses and resubmit failures

The University Of Sheffield.

Raw Data Replication

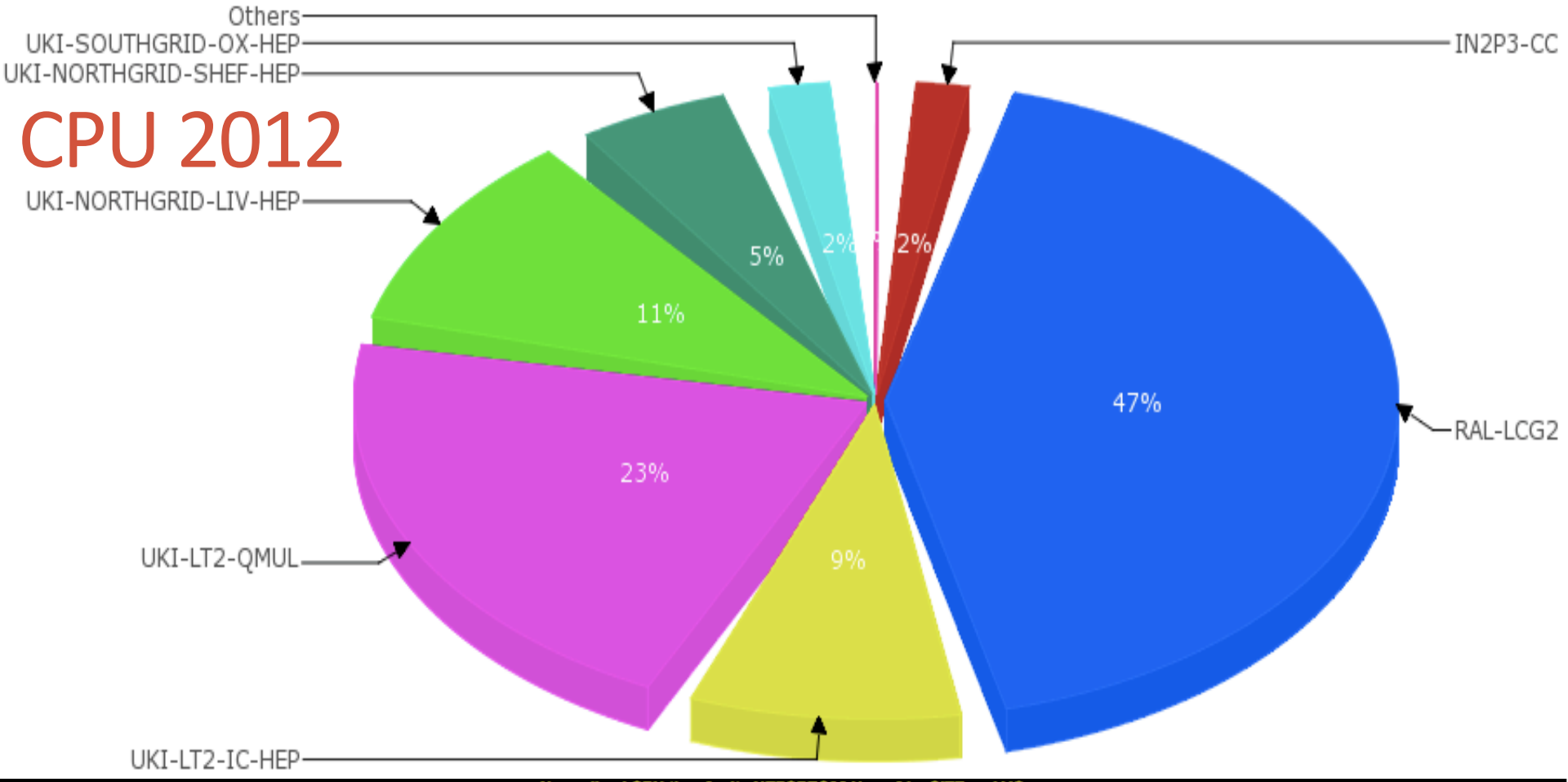


The University Of Sheffield.

t2k.org GRID model

- Present dataset ~36TB (8% of design [all sub-detectors])
  - push raw data [ND280 good spills only] and MC input (neutrino 4-vectors) to T2 ahead of processing
    - ~100TB MC output/cycle
- ***Discrete*** processing cycles
  - deposits output at site where processed
- Unified, global analysis file format – few 10s of MB
  - skimmed/reduced/analysed by different analysis sub groups
    - usually at local institutes,
- Archiving
  - ***retrospective*** clean-up (free-up **T2**, archive at **T1**)
    - delete intermediate files from previous cycle
  - data distribution is ***decoupled*** from data processing
- In general the LCG has been pretty stable over the past 1-2yrs
  - Feb/Mar 2014 in particular has seen excellent job throughput at RAL
    - LHC experiments relatively quiet

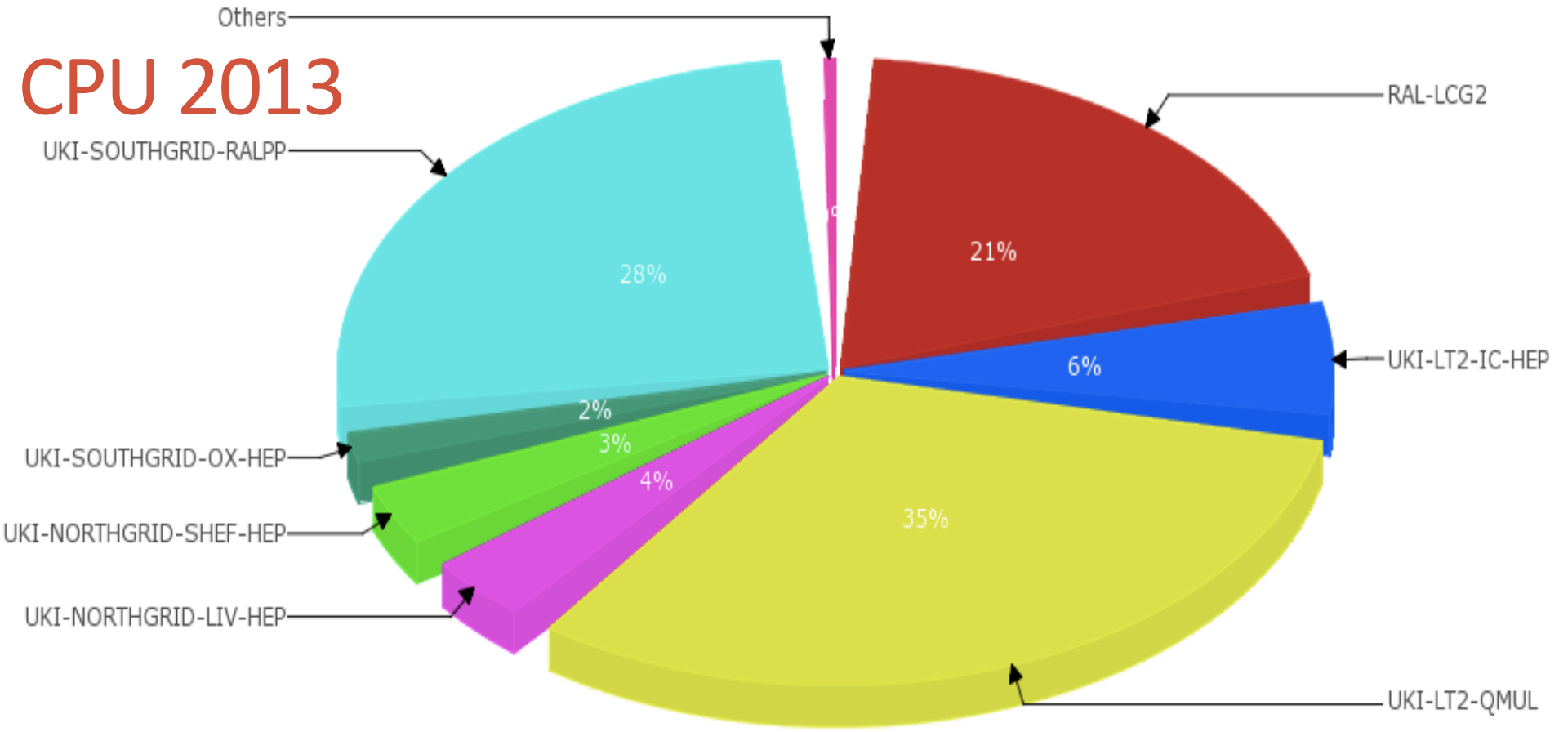
CPU 2012



Normalised CPU time [units HEPSPEC06.Hours] by SITE and VO

SITE	t2k.org	Total	%
GRIF	0	0	0.00%
IFIC-LCG2	0	0	0.00%
IN2P3-CC	169,744	169,744	1.84%
pic	236	236	0.00%
RAL-LCG2	4,320,988	4,320,988	46.95%
UKI-LT2-IC-HEP	862,520	862,520	9.37%
UKI-LT2-QMUL	2,136,232	2,136,232	23.21%
UKI-NORTHGRID-LANCS-HEP	1,964	1,964	0.02%
UKI-NORTHGRID-LIV-HEP	1,035,552	1,035,552	11.25%
UKI-NORTHGRID-MAN-HEP	0	0	0.00%
UKI-NORTHGRID-SHEF-HEP	479,952	479,952	5.21%
UKI-SOUTHGRID-OX-HEP	196,568	196,568	2.14%
UKI-SOUTHGRID-RALPP	32	32	0.00%
Total	9,203,788	9,203,788	

CPU 2013



Normalised CPU time [units HEPSPEC06.Hours] by SITE and VO

SITE	t2k.org	Total	%
GRIF	88	88	0.00%
IFIC-LCG2	52	52	0.00%
IN2P3-CC	2,504	2,504	0.01%
pic	12	12	0.00%
RAL-LCG2	5,819,360	5,819,360	21.10%
SFU-LCG2	12,012	12,012	0.04%
UKI-LT2-IC-HEP	1,778,824	1,778,824	6.45%
UKI-LT2-QMUL	9,679,396	9,679,396	35.09%
UKI-NORTHGRID-LANCS-HEP	116,164	116,164	0.42%
UKI-NORTHGRID-LIV-HEP	1,034,908	1,034,908	3.75%
UKI-NORTHGRID-MAN-HEP	8	8	0.00%
UKI-NORTHGRID-SHEF-HEP	945,332	945,332	3.43%
UKI-SOUTHGRID-OX-HEP	452,752	452,752	1.64%
UKI-SOUTHGRID-RALPP	7,740,276	7,740,276	28.06%
Total	27,581,688	27,581,688	

The University Of Sheffield.

Data usage snapshot – `lcg-infosites --vo t2k.org space`

TAG	SE	Free (TB)	Used(TB)
-	bohr3226.tier2.hep.manchester.ac.uk	16.62	3.69
T2KORGDISK	fal-pygrid-30.lancs.ac.uk	0.61	16.57
-	fal-pygrid-30.lancs.ac.uk	12.06	5.91
-	gfe02.grid.hep.ph.ic.ac.uk	64.02	33.50
T2KORGDISK	gfe02.grid.hep.ph.ic.ac.uk	77.32	33.04
T2KLIVERPOOLDISK	hepgrid11.ph.liv.ac.uk	2.74	6.93
-	hepgrid11.ph.liv.ac.uk	0.70	4.64
-	hepgrid11.ph.liv.ac.uk	0.94	20.31
-	heplnx204.pp.rl.ac.uk	0.00	0.00
T2K_DEFAULT	heplnx204.pp.rl.ac.uk	0.10	0.00
-	lcgse0.shef.ac.uk	0.85	5.29
T2KORGDISK	se03.esc.qmul.ac.uk	44.21	53.45
T2KORGDISK	srm-t2k.gridpp.rl.ac.uk	101.84	471.40
-	t2ksrm.nd280.org	83.83	390.81
T2KORGDISK	t2se01.physics.ox.ac.uk	2.15	0.00
-	t2se01.physics.ox.ac.uk	12.28	70.78

The University Of Sheffield.

Request from GridPP to forecast for 2015-19

- Last guestimate for storage
  - 1-2PB at T1 by the end of 2015
  - average 60-80TB per T2 (presently varies from ~few to ~55 TB)
- CPU (based on HEPSPEC06 as reported by EGI-portal)

```
2010-2011        0.18M hrs
2011-2012        1.40M hrs
2012-2013        2.20M hrs
2013-present     2.00M hrs
```

- Prediction for 2015

```
2013-2014     ~3-4 M hrs
2014-2015     ~4-6 M hrs
```

- Discontinuous beam schedule muddies the water

The University Of Sheffield.

Request from GridPP to forecast for 2015-19

- based on a 10x increase in our dataset
  - (as scheduled by the end of 2019)
- (somewhat crudely) scaling resource requirements by same factor
  - tentatively say we'll hit
    - 3-5PB T1
    - 600-800TB T2
    - 10-30Mhrs T1+T2 (HEPSPEC06) CPU
      - [derived from EGI portal accounting].

The University Of Sheffield.

Outstanding storage issues

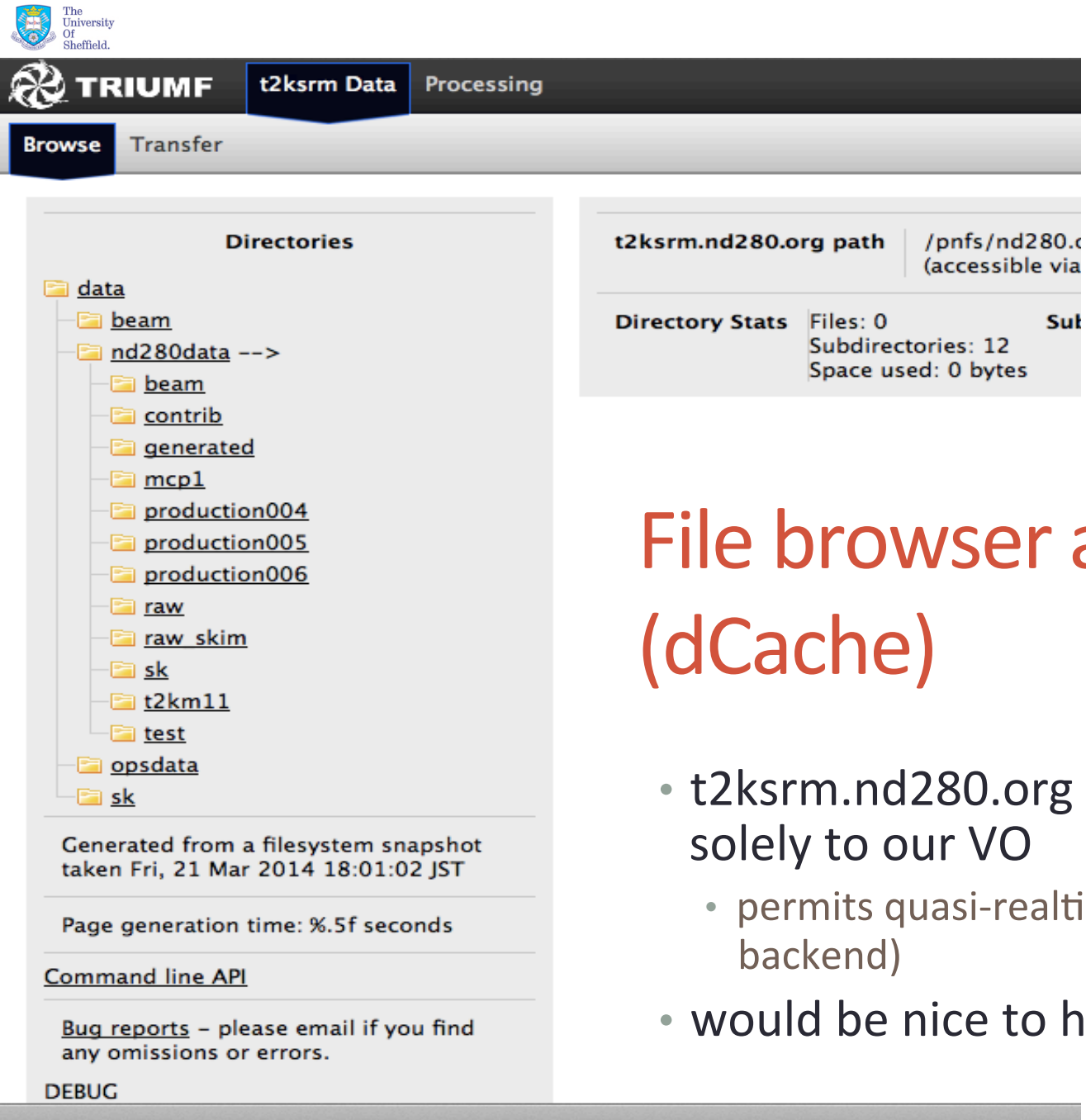
- FTS doesn't register copies
  - may be addressed by FTS3.2.0
    - have a development RPM but not tested it yet
- Have used FTS3 successfully many times
  - very quick
    - compared to FTS2 channels with low numbers of parallel transfers enabled
  - monitoring page clunky
    - frequent timeouts/slow loading
    - browse by VO functionality not as friendly as FTS2 counterpart*
  - not migrated T0-T1 transfers to FTS3 yet
    - hold out for FTS3.2 and replica registration
    - hope for monitoring improvements
      - presently FTS2 channel based view more informative*

*these are personal viewpoints

The University Of Sheffield.

Outstanding storage issues

- Auditing of disk usage still a chore
  - no global view(er), usually spider-bot style crawl of LFC instead
    - misses dark data
    - slow
  - disk usage reporting on bdii untrustworthy
    - occasionally misquotes
- Syncat dumps…
  - XML needs parsing to be useful
  - ATLAS do this, but their code uses dq2 and hence is not generic
    - https://twiki.cern.ch/twiki/bin/view/Atlas/DDMOperationsScripts



File browser at TRIUMF (dCache)

- t2ksrm.nd280.org SE is dedicated solely to our VO
  - permits quasi-realtime file browsing (SQL backend)
- would be nice to have elsewhere

The University Of Sheffield.

TRIUMF t2ksrm Data Processing

Browse Transfer

Directories

- data
  - beam
  - nd280data
    - beam
    - contrib
    - generated
    - mcp1
    - production004
    - production005
    - production006
    - raw
      - ND280
        - ECAL
        - FGD
        - ND280
          - 00000000_00000999
          - 00001000_00001999
          - 00002000_00002999
          - 00003000_00003999
          - 00004000_00004999
          - 00005000_00005999
          - 00006000_00006999
          - 00007000_00007999
          - 00008000_00008999 -->
          - 00009000_00009999

Query File name: may include 1 wildcard at end i.e. oa_nd_cos_* Filter Reset

t2ksrm.nd280.org path /pnfs/nd280.org/data/nd280data/raw/ND280/ND280/00008000_00008999
(accessible via dCache or Grid – more information)

Directory Stats	Subtree Stats	Display
Files: 11328	Files: 11328	Filename: hide
Subdirectories: 0	Subdirectories: 0	Size: hide
Space used: 9.8 TB	Space used: 9.8 TB	Last modified: hide

Filename	Size	Last modified
nd280_00008000_0000.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:18:19 PST
nd280_00008000_0001.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:19:29 PST
nd280_00008000_0002.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:17:33 PST
nd280_00008000_0003.daq.mid.gz	187 MB	Mon, 16 Jan 2012 10:10:22 PST
nd280_00008002_0000.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:18:21 PST
nd280_00008002_0001.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:17:31 PST
nd280_00008002_0002.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:18:28 PST
nd280_00008002_0003.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:17:44 PST
nd280_00008002_0004.daq.mid.gz	256 MB	Mon, 16 Jan 2012 10:13:28 PST
nd280_00008003_0000.daq.mid.gz	54 MB	Mon, 16 Jan 2012 10:15:15 PST
nd280_00008004_0000.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:26:51 PST
nd280_00008004_0001.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:27:48 PST
nd280_00008004_0002.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:27:46 PST
nd280_00008004_0003.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:28:08 PST
nd280_00008004_0004.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:27:34 PST
nd280_00008004_0005.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:27:18 PST
nd280_00008004_0006.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:27:29 PST
nd280_00008004_0007.daq.mid.gz	955 MB	Mon, 16 Jan 2012 10:28:10 PST
nd280_00008004_0008.daq.mid.gz	954 MB	Mon, 16 Jan 2012 10:35:27 PST

The University Of Sheffield.

In the pipeline

- Presently have 0.5 grad students working on GANGA
  - has been working with developers
    - ideally could do with more manpower
      - GANGA *could* be a the single homogenous layer above our software (local/batch/grid/…)
- Have submitted some jobs via new DIRAC server at ICL
  - presently there is a clunky workaround due to existing format of LFC SQL
    - again, resources for this development are thin on the ground
    - what is the future landscape for DIRAC, is it a worthwhile endeavour for T2K?
- MySQL based Beam Summary Data implementation
  - previously locally cached on CE
    - now imposes too high disk burden
- xrootd implementation underway
  - so far tested at RAL only
    - requires additional compiler flags for ROOT
- Hyper-K is coming…

The University Of Sheffield.

thanks