

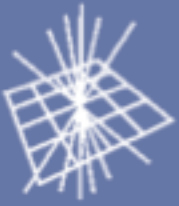
GridPP

UK Computing for Particle Physics

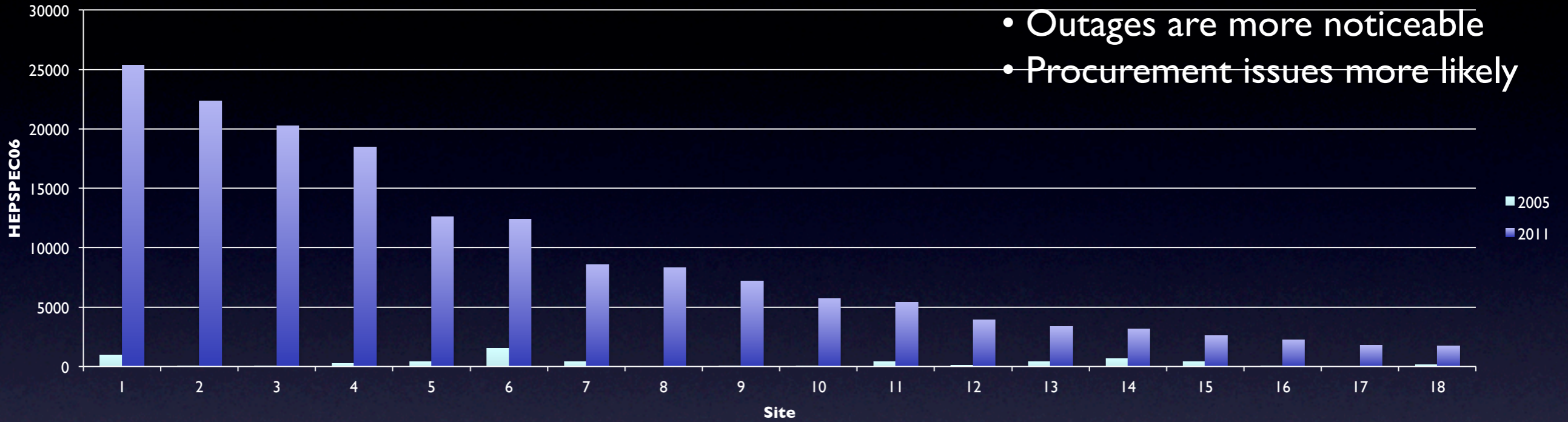
Operating the UK Particle Physics Grid in 2011

Jeremy Coles
ISGC2011 - Taipei, Taiwan
23rd March 2011

- The problems of scale
- Utilisation
- Large site hurdles – RAL TI
- Wider problems in the grid world
- Some upcoming challenges
- New communities
- Project changes
- Summary

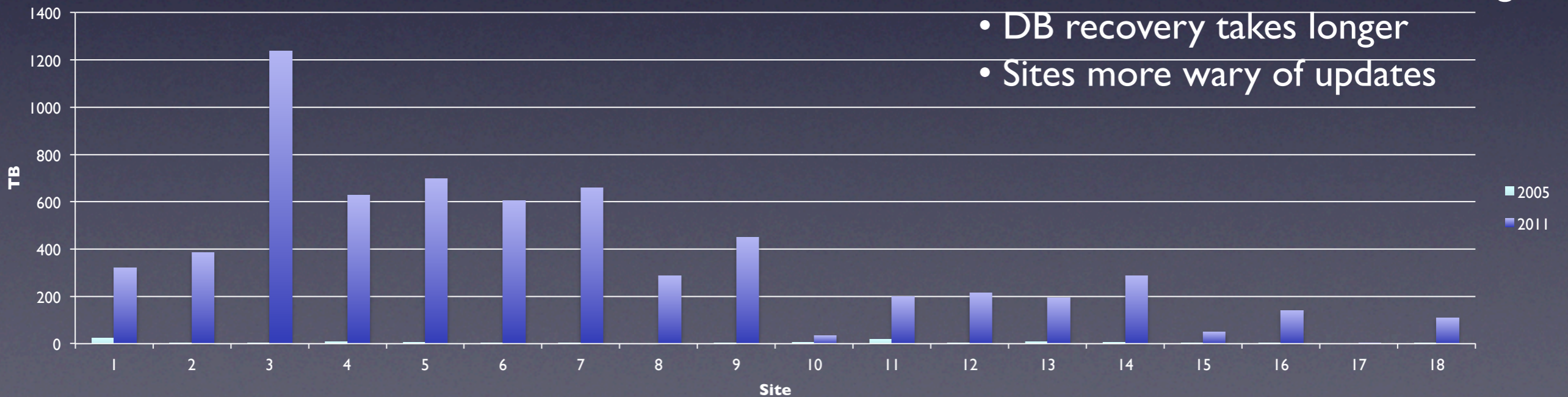


CPU



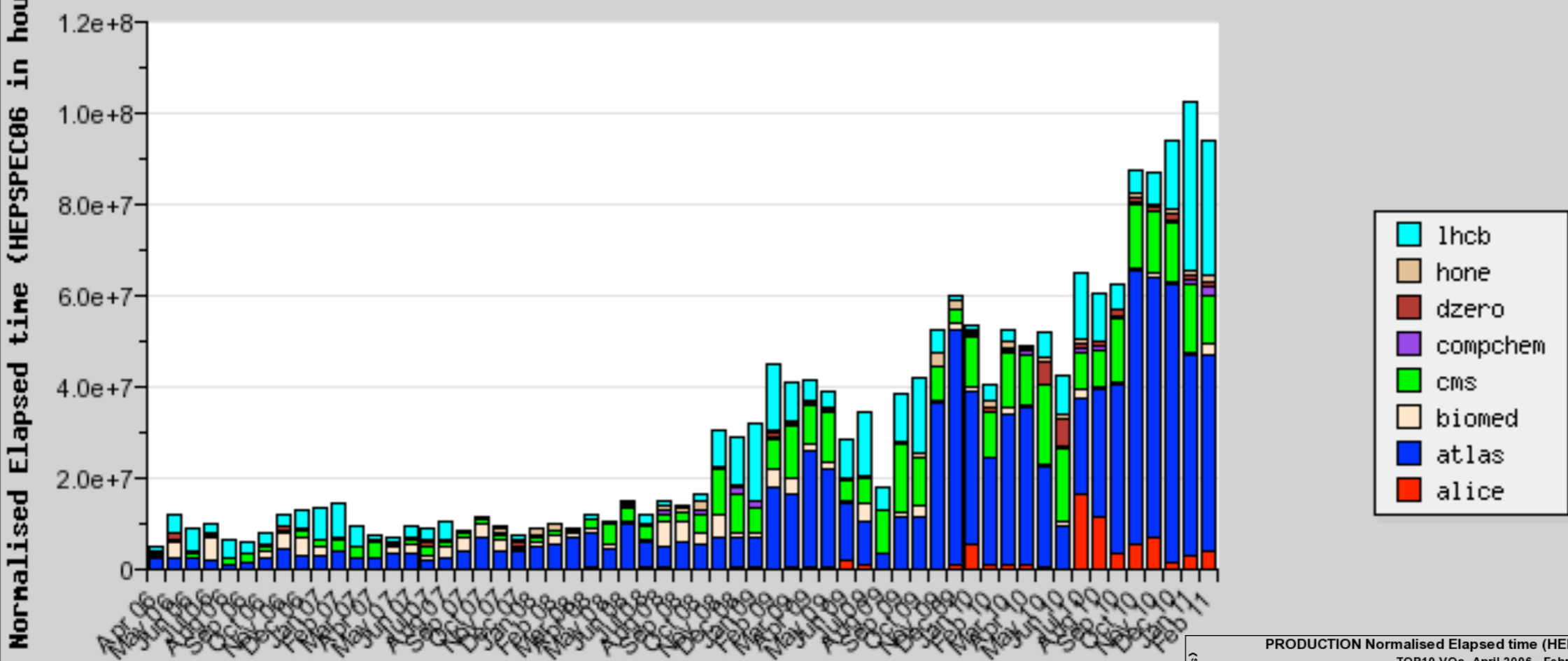
- Sites are larger – not manpower
- Outages are more noticeable
- Procurement issues more likely

Storage



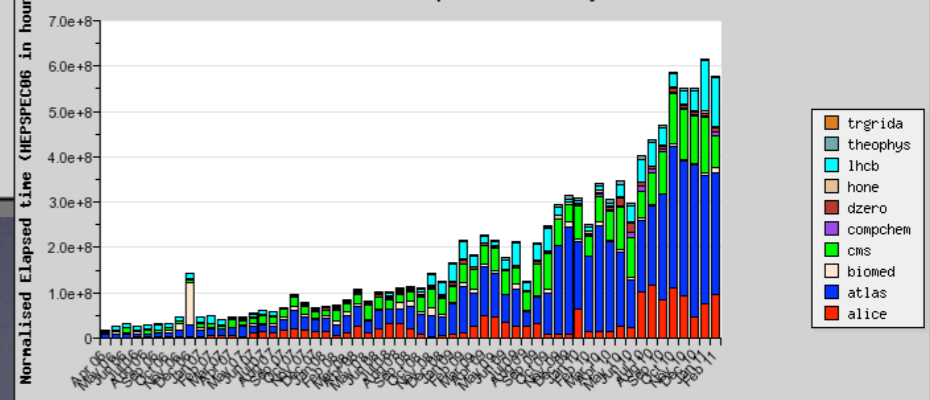
- Institutes rewarded for investing
- DB recovery takes longer
- Sites more wary of updates

UKI Normalised Elapsed time (HEPSPEC06) by VO and DATE
TOP10 VOs. April 2006 - February 2011



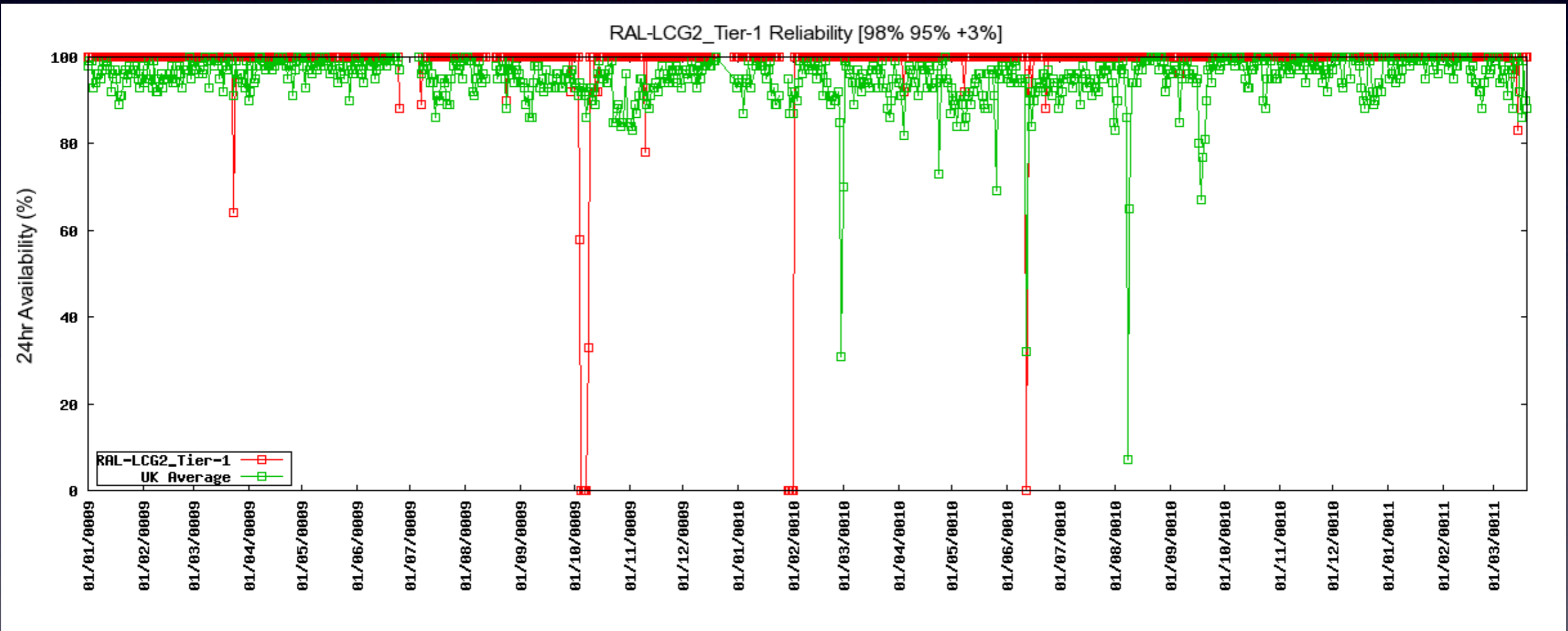
(C) CESGA 'EGI View': UKI / normelap-HEPSPEC06 / 2006:4-2011:2 / VO-DATE / top10 (x) / ACCBAR-LIN / i

PRODUCTION Normalised Elapsed time (HEPSPEC06) by VO and DATE
TOP10 VOs. April 2006 - February 2011

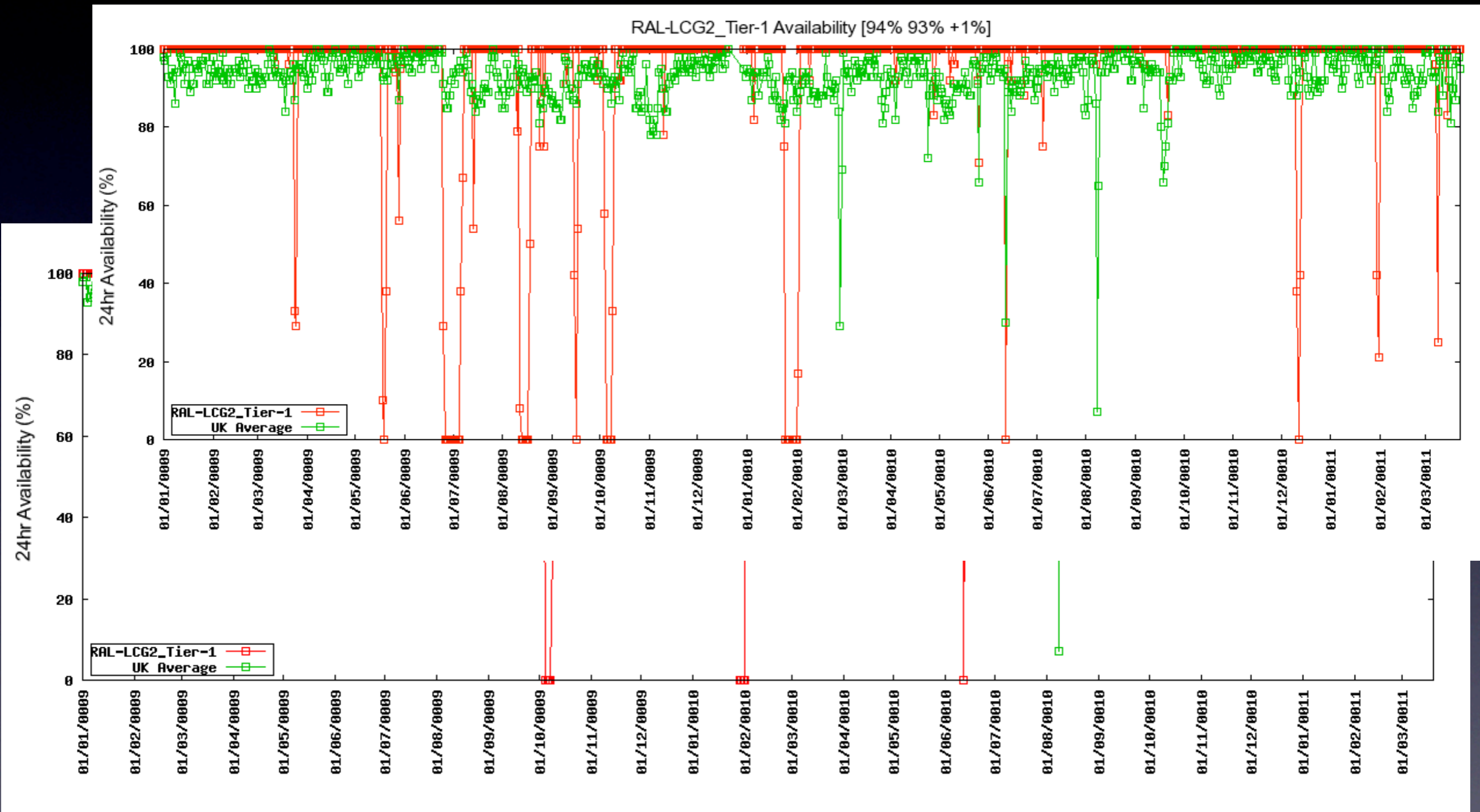


(C) CESGA 'EGI View': PRODUCTION / normelap-HEPSPEC06 / 2006:4-2011:2 / VO-DATE / top10 (x) / ACCBAR-LIN / i 2011-03-19 20:40 UTC

- Dips are usually monitoring issues not site problems
- The red line shows the RAL Tier-I reliability whereas the green shows the UK average over the period 2009 to 2011.

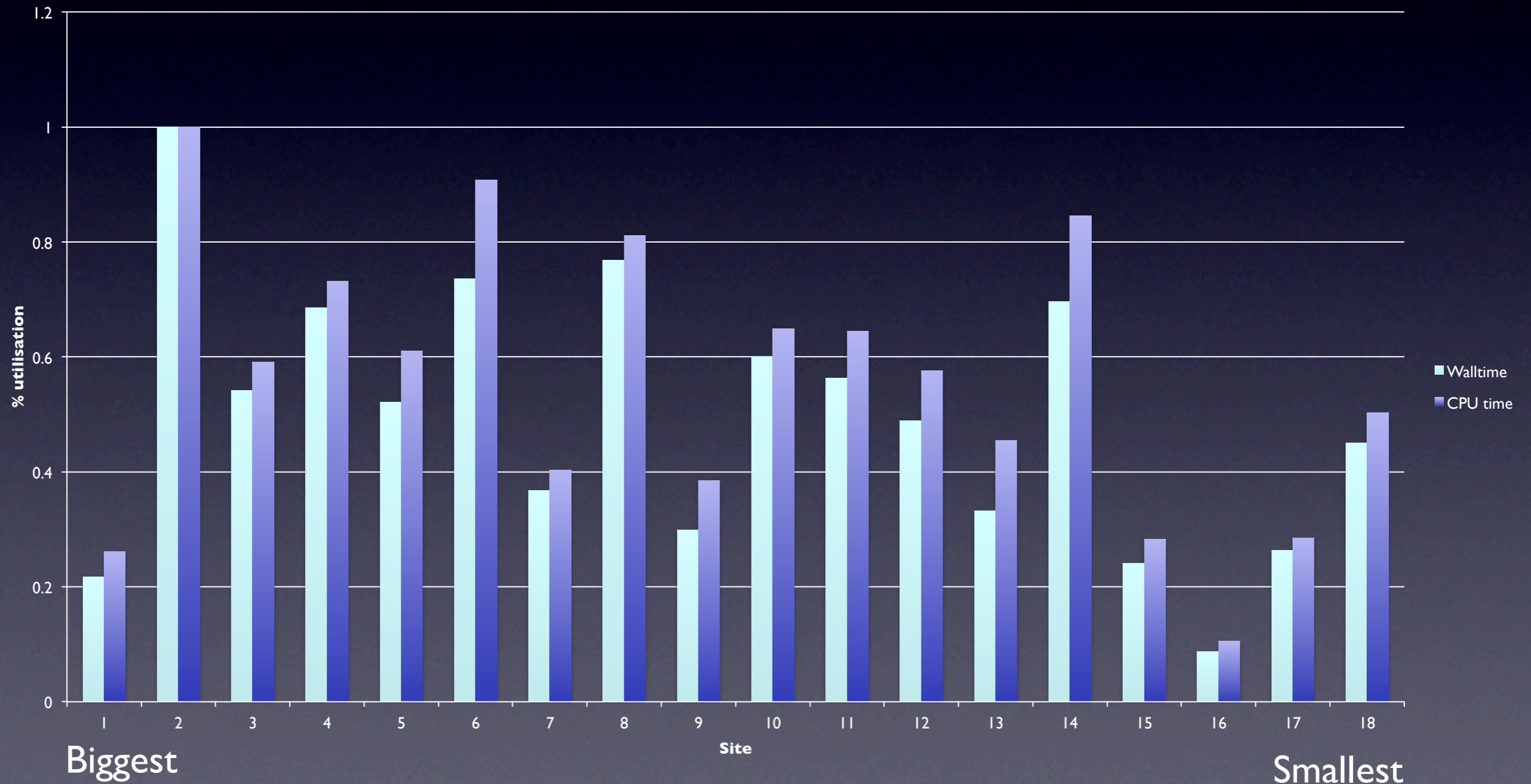


- Dips here are usually for scheduled maintenance. Seems to have reached a plateau.



We still do not see saturation of resources but 2011 LHC running may change this situation. It is interesting to see the differences in utilisation between sites of all sizes. This view covers Q4 2010.

Tier-2 CPU utilisation



- Ongoing changes for UPS

Leading power factor due to switch-mode PSUs in systems. Causes 3KHz ringing on current. Most kit stable but EMC AX4-5 FC arrays unpredictably detect supply failure and shut down arrays. Longer feed cable from UPS to PDU has made 50% reduction in distortion. Proposed long-term solution: isolation transformer

- WAN developments

Existing 10Gb/s site link to SuperJanet5 doubled to 20Gb/s. Failover capacity maintained LHCOPN failover link @ 10Gb/s

- Storage commissioning – disk acceptance issues (plus need to keep diverse suppliers)

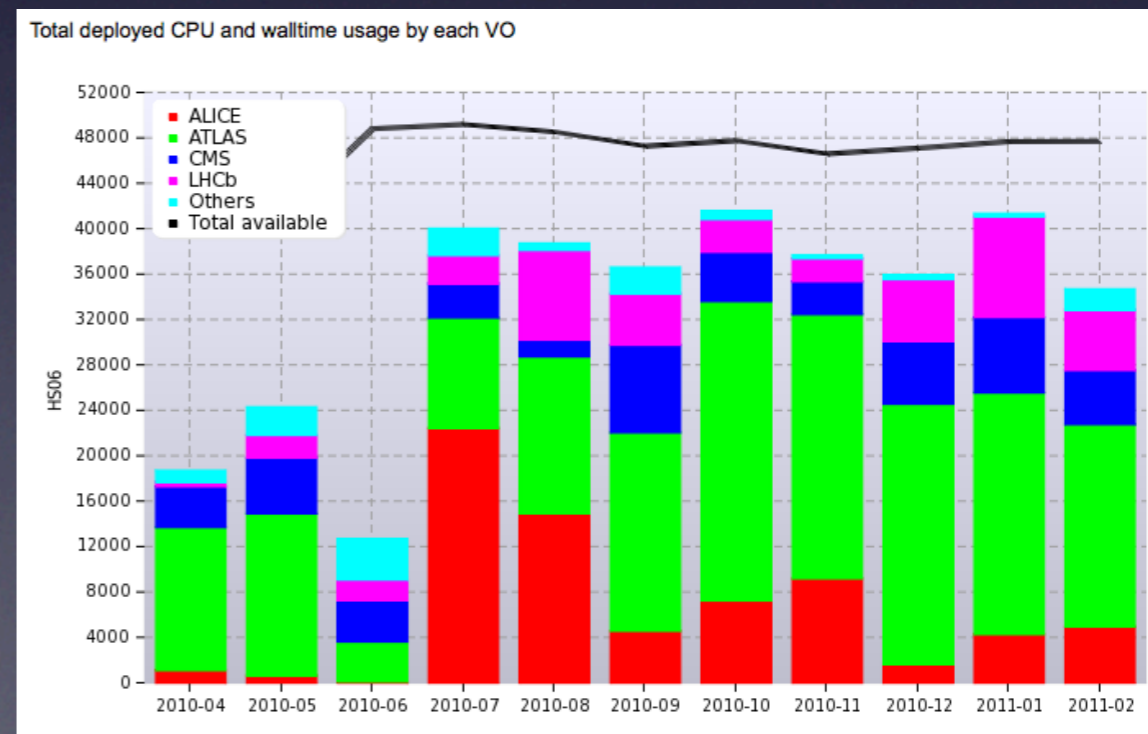
- Dust in machine room from lagging. – still a concern for the tape systems

Ongoing work:

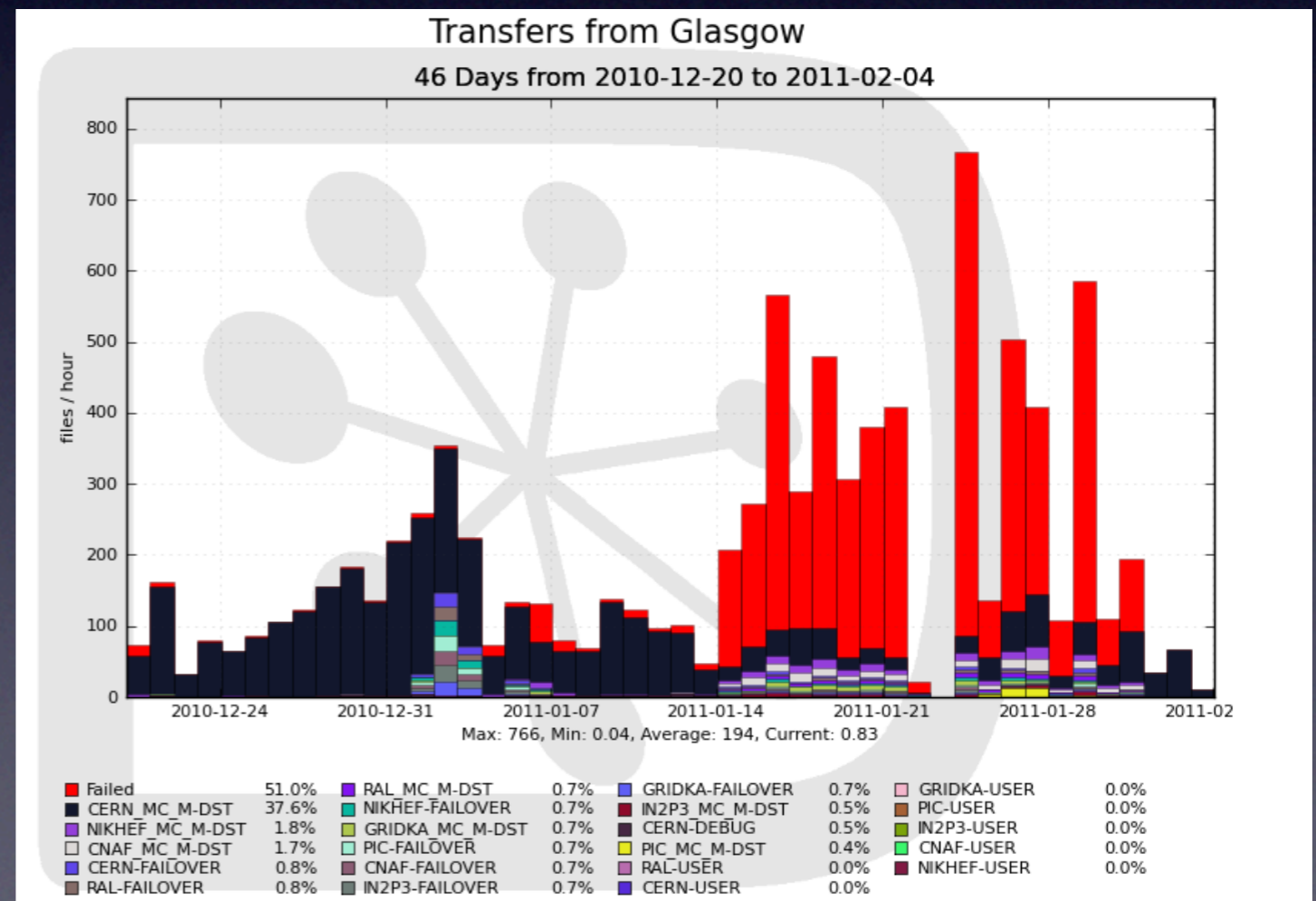
- Quattor (fabric management. ~5200 cores, 700+ systems)
- Virtualisation study of Hyper-V as possible virtualisation platform
- CVMFS (A caching, http based read-only filesystem. Removes the need for local software install jobs at every site)
- Whole node reservation
- File system efficiency and using many core nodes (48-core being tested at some sites). The problem of meta-data/data/ bandwidth and I/O limits (e.g. cache) and tuning.



<http://www.gridpp.rl.ac.uk/capacity/index.php?p=home>



- VO user jobs using remote SEs in a manner that overloads site WAN links
- Lcg-infosites too old on some sites leading to reduction in supporting sites
- Re(chasing) the same problems! (See diagram). NAT SACK/DSACK configuration changes reset!
- Confidence & security – dealing with Multi-user pilot jobs (glexec etc.)
- Convergence of monitoring/publishing accuracy
- Full disk tokens at many sites (ATLAS)
- TI impacts (site access router & LFC)
- Network problem ownership
- Various WMS issues (inc. pheno impact)
- Patching for vulnerabilities
- Adapting as computing models change...
 - Higher trigger rates wanted. Larger pileup.
 - Network: Flatter hierarchy; Dynamic data caching; remote data access.



ATLAS Sonar Tests: Large Files, All tests since 2011-01-01

Source Site	Destination Site										Average
	RAL-LCG2_DATADISK	FZK-LCG2_DATADISK	INFN-T1_DATADISK	PIC_DATADISK	NDGF-T1_DATADISK	TAIWAN-LCG2_DATADISK	SARA-MATRIX_DATADISK	TRIUMF-LCG2_DATADISK	IN2P3-CC_DATADISK	BNL-OSG2_DATADISK	
UKI-SCOTGRID-GLASGOW_DATADISK	14.07	5.18	7.68	3.01	10.37	0.90	7.88	1.91	8.87	1.95	6.18
UKI-NORTHGRID-MAN-HEP_DATADISK	3.06	3.48	3.48	2.00	0.59	0.67	4.13	1.78	6.30	1.20	2.67
UKI-NORTHGRID-LANCS-HEP_DATADISK	5.87	2.78	3.20	2.51	1.99	0.57	5.36	1.25	4.81	0.96	2.93
UKI-NORTHGRID-SHEF-HEP_DATADISK	9.77	6.48	8.03	5.25	1.81	1.56	9.09	2.76	15.88	6.16	6.68
UKI-NORTHGRID-LIV-HEP_DATADISK	8.26	5.64	7.80	5.40	4.13	2.16	9.08	2.79	14.10	5.27	6.46
UKI-SOUTHGRID-CAM-HEP_DATADISK	6.99	3.93	5.78	3.53	2.82	0.79	8.50	1.82	8.92	2.96	4.60
UKI-SOUTHGRID-OX-HEP_DATADISK	10.44	6.18	7.18	4.90	5.05	1.13	9.80	2.30	11.40	4.32	6.27
UKI-SOUTHGRID-RALPP_DATADISK	19.64	5.43	6.75	6.54	1.61	1.88	10.62	2.96	18.84	5.02	7.93
UKI-LT2-QMUL_DATADISK	14.24	3.31	12.75	11.73	14.90	4.36	31.18	10.47	21.41	14.47	13.88
Average	10.26	4.71	6.96	4.99	4.81	1.56	10.63	3.11	12.28	4.70	6.40

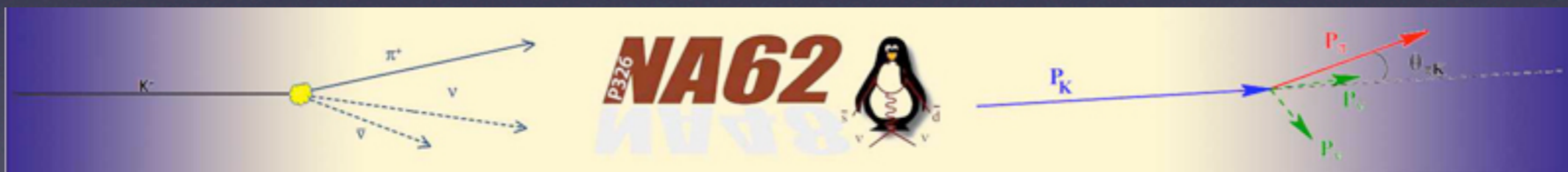
Results do not suggest network bandwidth limitations but hint at the need for improved tuning across the transfer chain (e.g. buffer sizes, FTS settings)

Source Site	Destination Site										Average
	UKI-SCOTGRID-GLASGOW_DATADISK	UKI-NORTHGRID-MAN-HEP_DATADISK	UKI-NORTHGRID-LANCS-HEP_DATADISK	UKI-NORTHGRID-SHEF-HEP_DATADISK	UKI-NORTHGRID-LIV-HEP_DATADISK	UKI-SOUTHGRID-CAM-HEP_DATADISK	UKI-SOUTHGRID-OX-HEP_DATADISK	UKI-SOUTHGRID-RALPP_DATADISK	UKI-LT2-QMUL_DATADISK	Average	
RAL-LCG2_DATADISK	3.44	40.21	18.12	9.00	5.00	13.51	10.87	32.72	9.85	15.86	
FZK-LCG2_DATADISK	5.76	9.07	11.23	4.57	11.94	12.23	10.22	4.18	5.37	8.28	
INFN-T1_DATADISK	1.24	5.47	17.15	3.76	2.15	3.52	3.84	0.61	3.76	4.61	
PIC_DATADISK	1.07	1.41	2.12	1.04	1.14	1.77	1.54	1.34	1.95	1.49	
NDGF-T1_DATADISK	3.86	21.58	26.70	4.54	1.65	4.05	6.73	1.71	17.48	9.81	
TAIWAN-LCG2_DATADISK	1.50	3.62	3.98	1.02	1.76	3.39	1.64	0.56	3.99	2.39	
SARA-MATRIX_DATADISK	0.63	20.47	38.96	1.31	0.98	7.97	7.38	0.97	19.53	10.91	
TRIUMF-LCG2_DATADISK	1.11	3.70	5.53	1.09	1.42	3.22	1.43	0.64	2.08	2.25	
IN2P3-CC_DATADISK	2.13	24.20	28.31	3.04	2.13	8.29	6.02	5.31	11.37	10.09	
BNL-OSG2_DATADISK	0.94	9.60	14.68	1.96	0.88	1.59	3.70		5.40	4.84	
Average	2.17	13.93	16.68	3.13	2.90	5.96	5.34	5.34	8.08	7.05	

Neutrino facility - SNO+ will be a new kilo-tonne scale liquid scintillator detector that will study neutrinos. The experiment will be located approximately 2km underground in VALE's Creighton mine near Sudbury, Ontario, Canada. <http://snoplus.phy.queensu.ca/Home.html>



CERN@school gives secondary schools across the UK the opportunity to become engaged in university level physics research. It makes use of Medipix detector chips to observe cosmic rays. <http://194.81.239.119/>



NA62 proposes to measure the very rare kaon decay $K^+ \rightarrow \pi^+ \nu \bar{\nu}$ at the CERN SPS to extract a 10% measurement of the Cabibbo–Kobayashi–Maskawa matrix parameter $|V_{td}|$. <http://na62.web.cern.ch/na62/>



National **e**-Infrastructure for **S**ocial **S**imulation

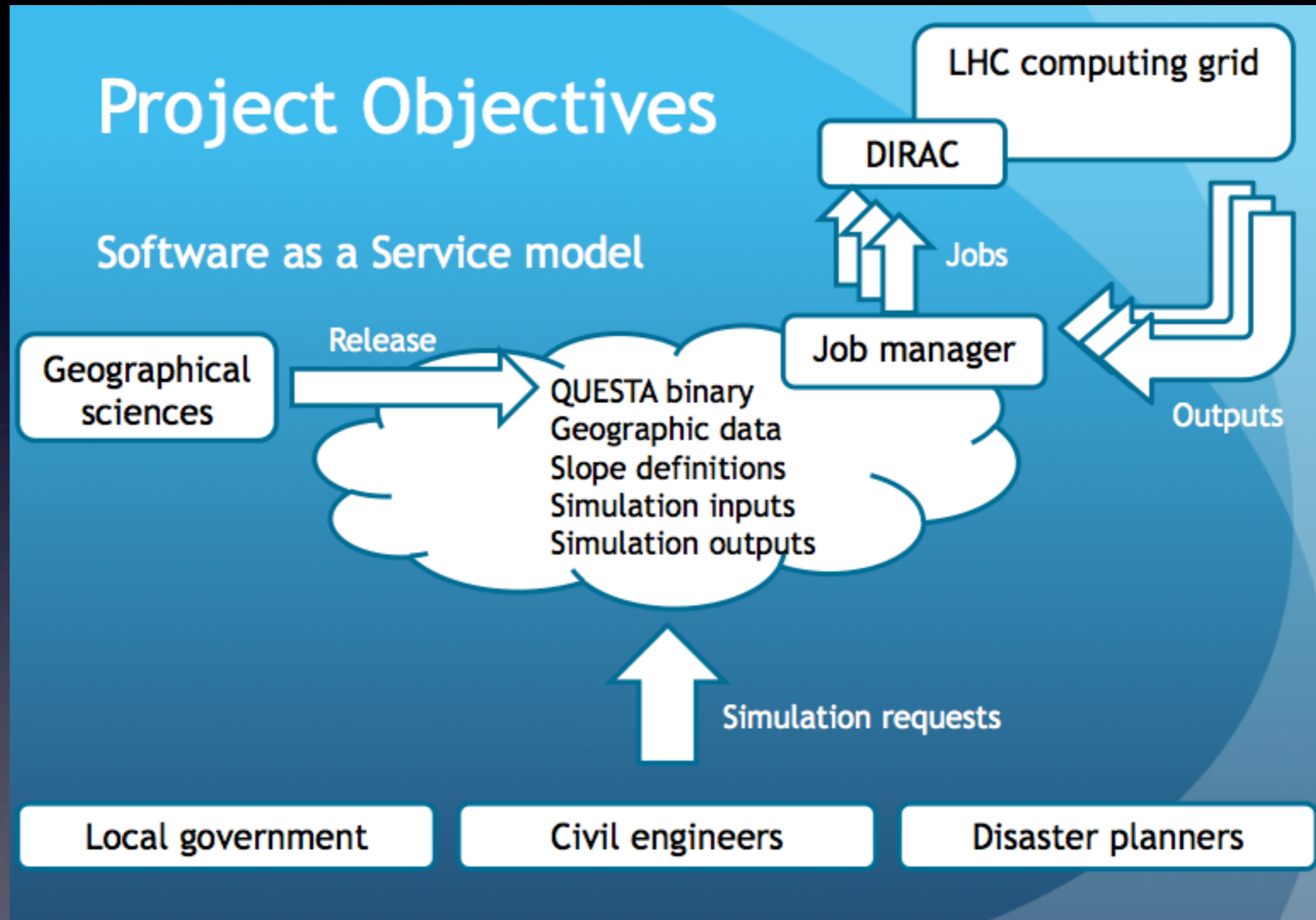
NeISS project will build a generic production quality social simulation e-Infrastructure covering the social simulation lifecycle. Social Simulation is an expanding field due to its forecasting applications for scenarios in transport, housing, education, healthcare etc.

<http://www.geog.leeds.ac.uk/projects/neiss/about.php>

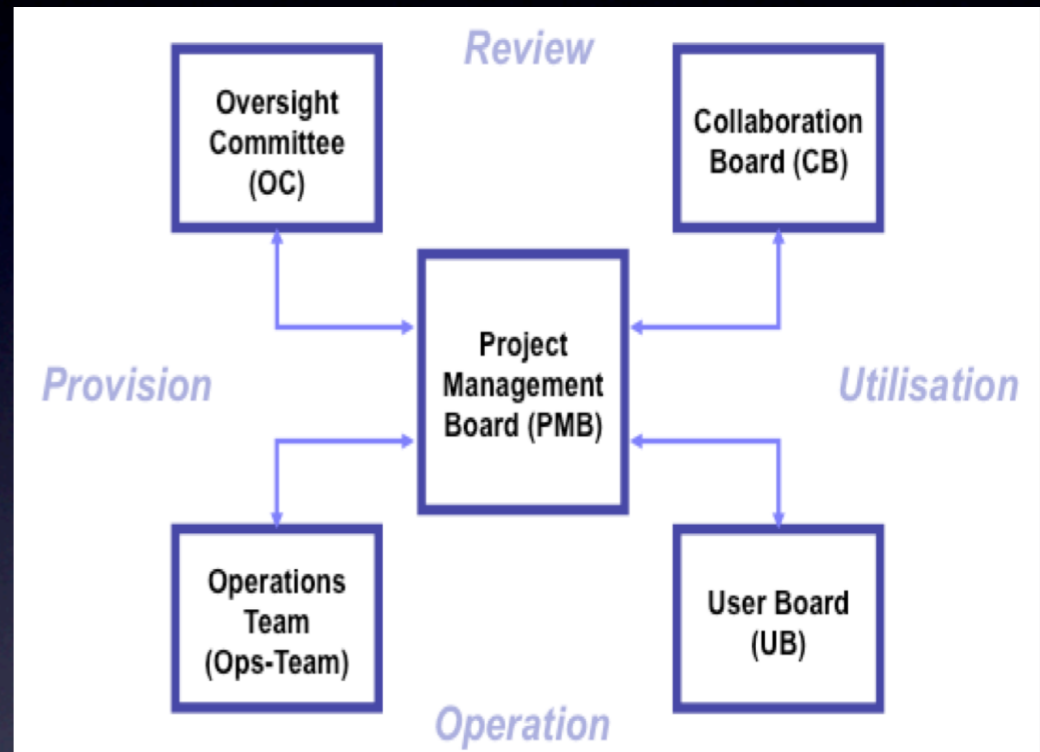
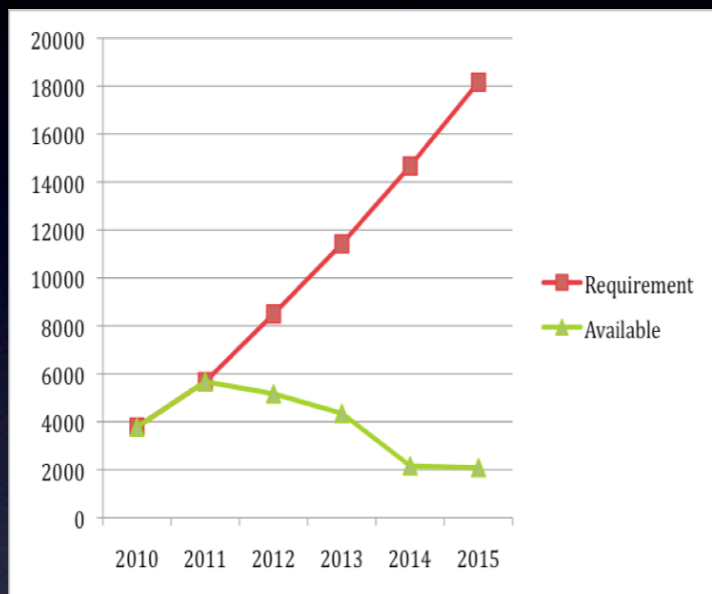
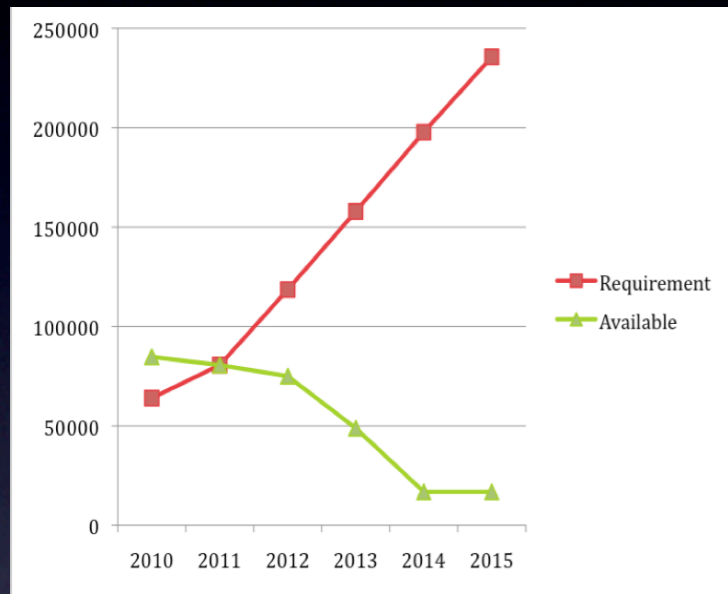
Landslides: Field evidence shows the complex relationships between precipitation, soil properties, vegetation and slope geometries in respect of potential failure. These can be modelled to help assess risks.



<http://www.mossaic.org/landslide/>



- GridPP is funded in (approximately) 4 year cycles. GridPP3 finishes this month. The collaboration has received funding for GridPP4 “Computing in the LHC era”.

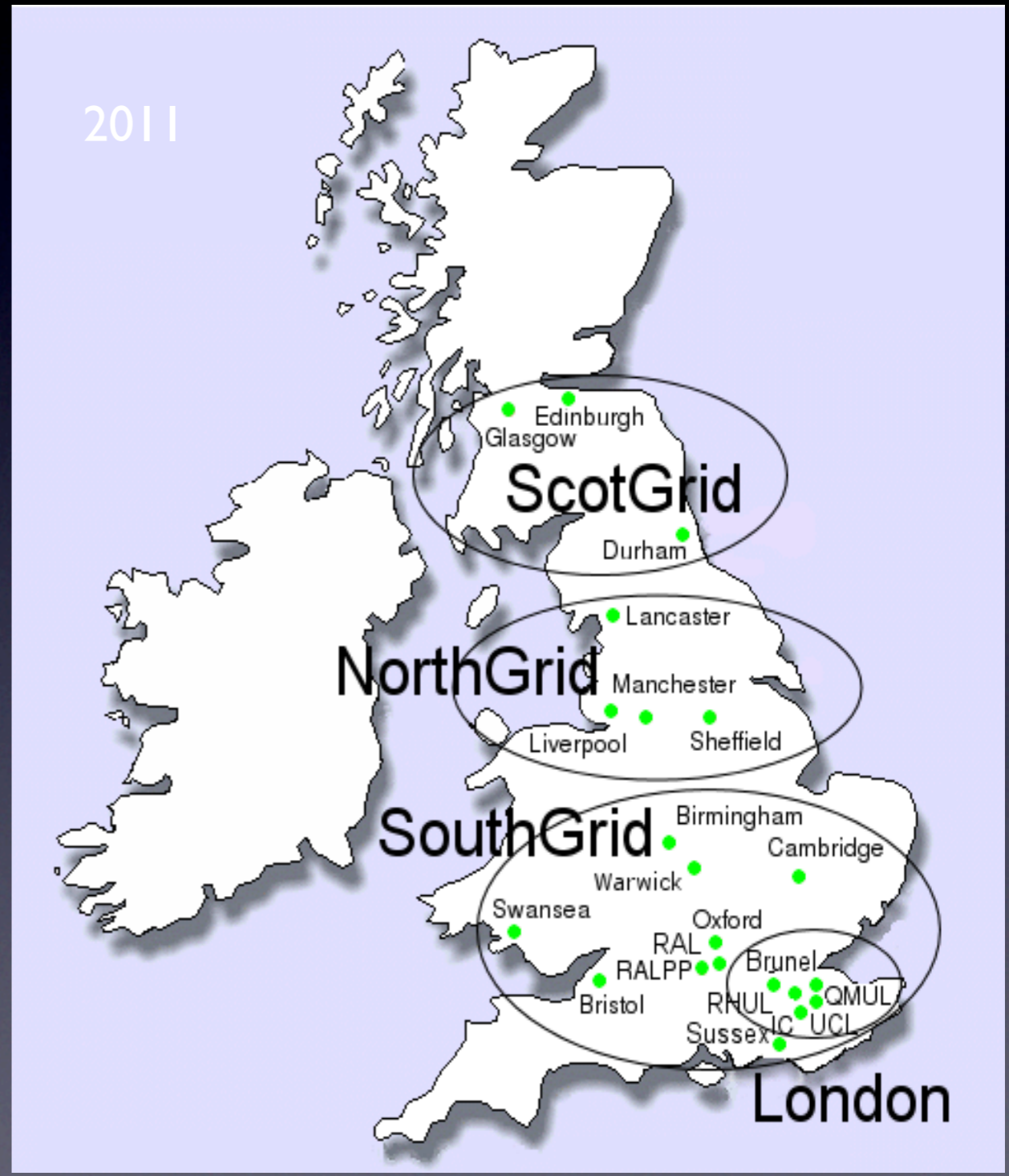
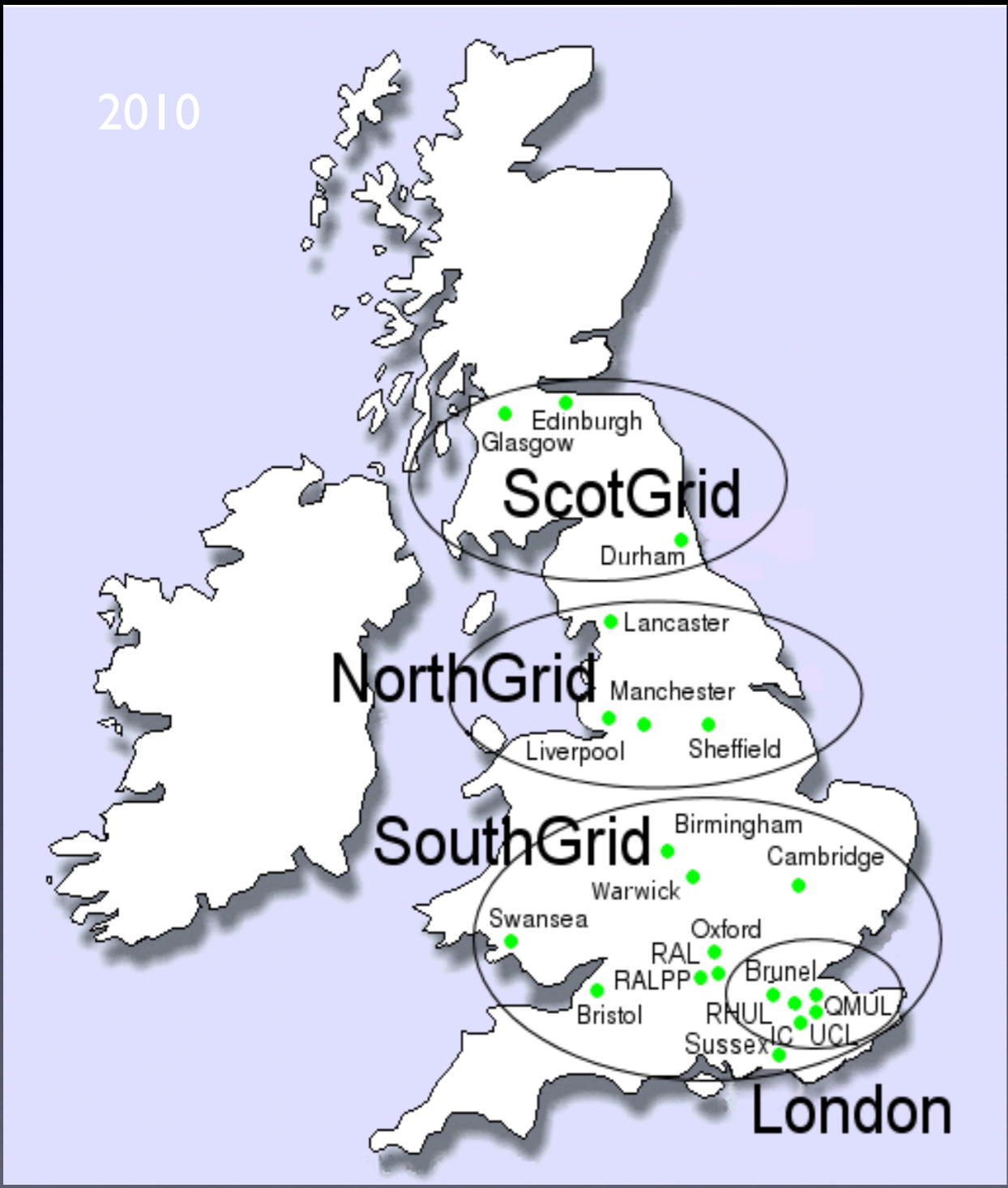


Contains 6 work packages (T1.T2. Deployment ops & support. Expt. Support. Management. Impact.)

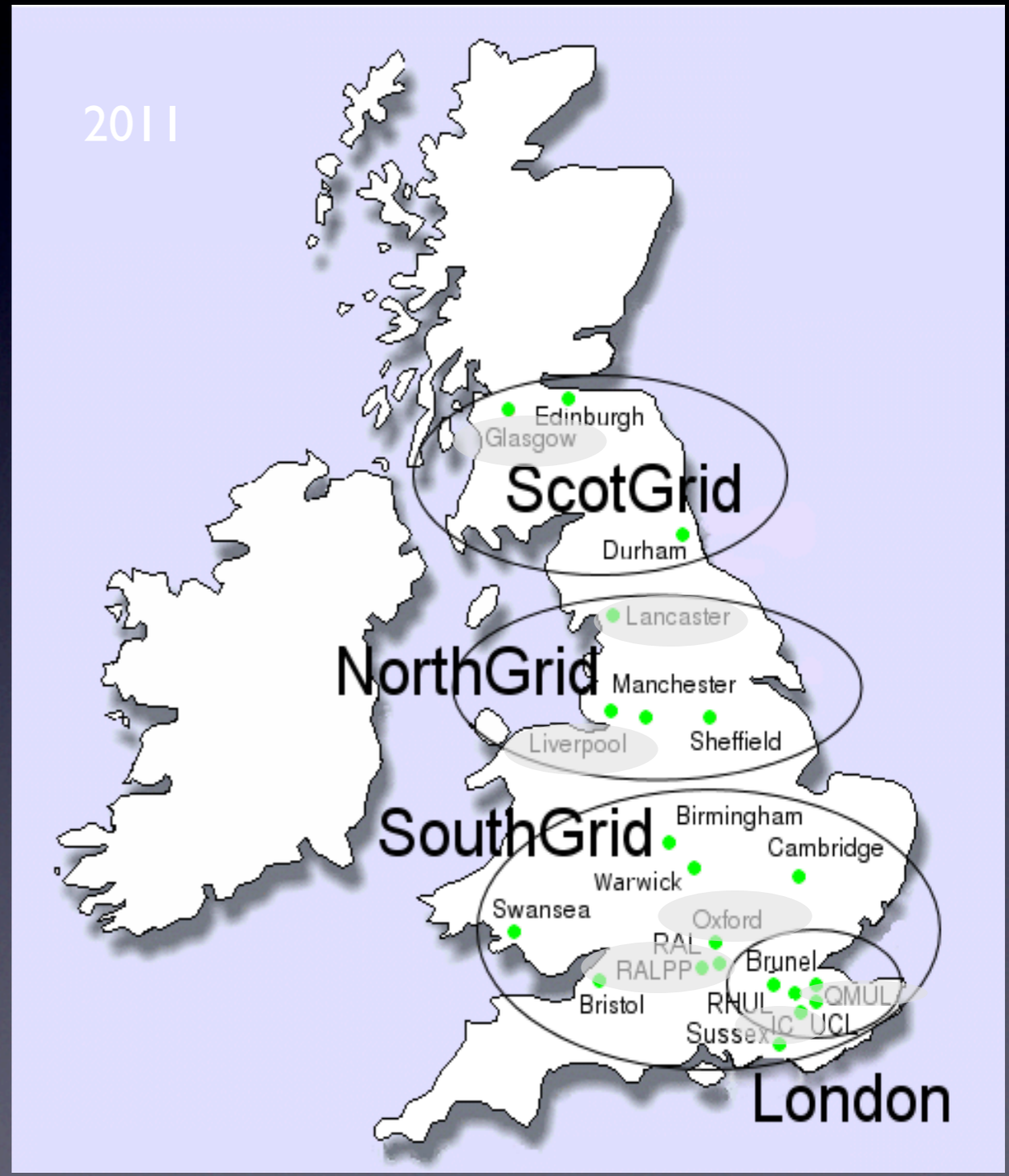
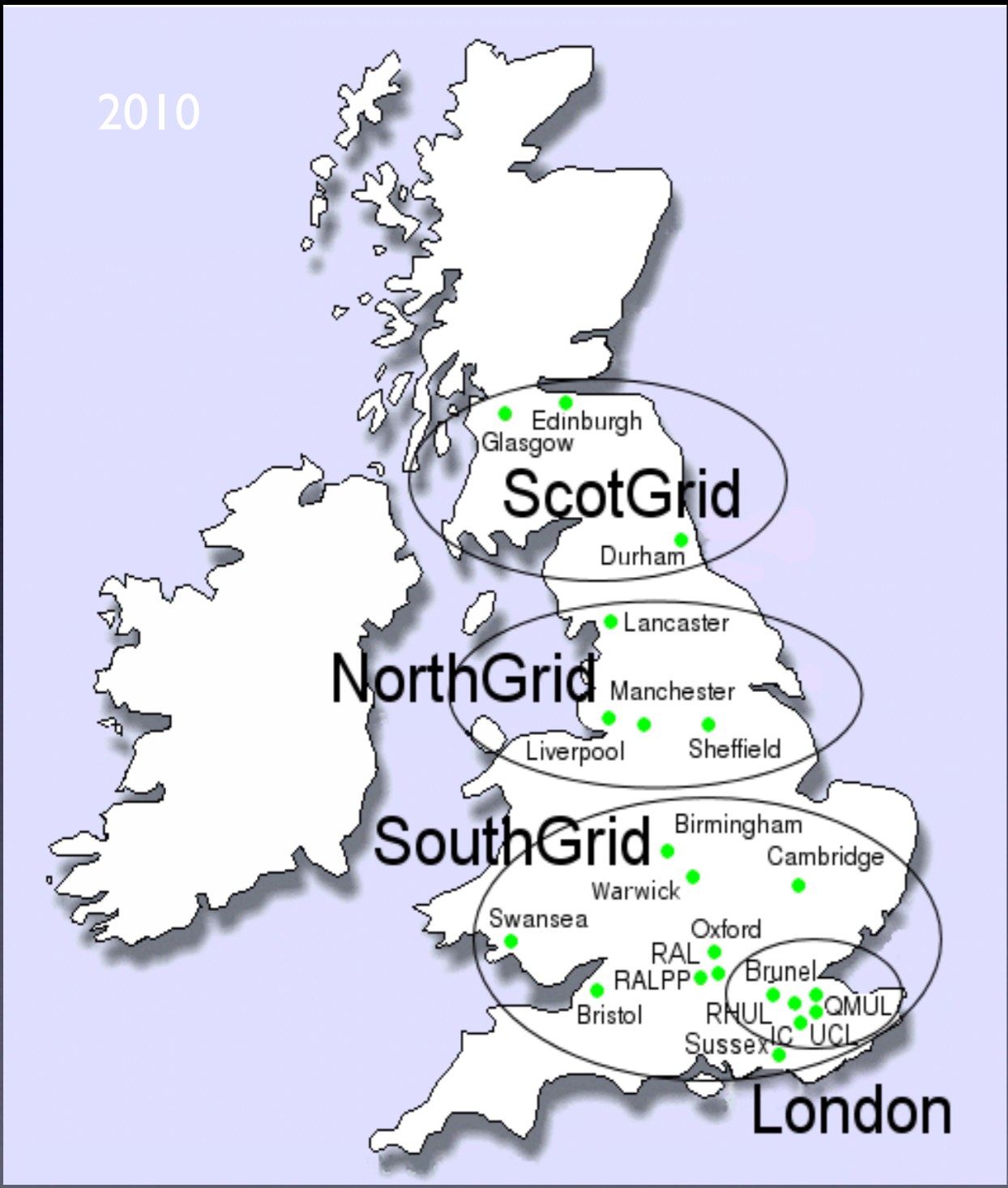
But after funding agreed...hardware costs coming in more expensive recently than expected on historical projections.... LHC running in 2012 has increased demands....

Funding is more targeted based on experiment requirements for stability/specialist sites. We are currently developing an algorithm to keep funding allocations transparent.

- Spot the difference!



- In GridPP4 we differentiate sites based on their use for analysis



- Sites become larger and face new challenges
- Still no saturation but that may change very soon
- Many problems remain for sites from both middleware and hardware changes
- Interest in opportunistic use of resources is growing
- The UK physics grid has secured funding for the next project phase