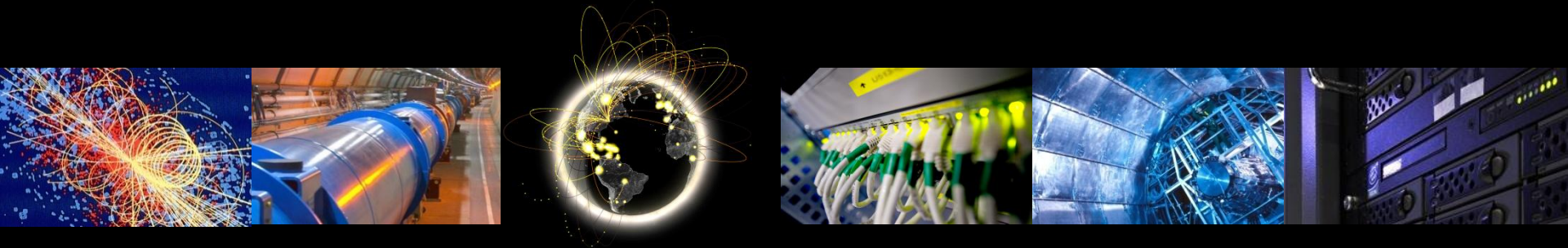# Storage Accounting

John Gordon, STFC

GDB March 2013

Outline

- History
- Timeline
- Current state
- Evidence of testing
- Current issues
- Open Issues
- Portal Views
- Reports?

- AOB
  - Data Retention
  - Portal Requirements

EMI StAR

- EMI recognised the requirement for storage accounting
- EMI created a proposal for a storage accounting record (StAR)
- 'Finalised' in June 2011
- Revised Autumn 2012 to include the WLCG requirements I described at the GDB last June.
- Published version: http://cds.cern.ch/record/1452920?ln=en
- Been through public comment in OGF and just resubmitted for approval
- implemented by EMI storage providers in EMI 3

Original EMI StAR timeline

- May 2011: StAR definition agreement

- Oct. 2011: Each SE to provide an implementation plan

- May 2012: Each SE to implement accounting sensors

- June 2012: Accounting record publishers (APEL and/or DGAS) to publish storage records

- December 2012: Each SE to test and deploy their accounting sensors (for EMI 3 release)

- January-April 2013: Bug-fixing in EMI 3 RC's

- April 2013: Accounting sensors released in EMI 3

Actual EMI Timeline

- February 2013 – first test storage records appeared from dcache and dpm

- February-March 2013 checking and revising.

- 11 March – Release of EMI 3.0

- Any changes required should be available in update 1.

Status of StAR implementations

- Dcache – published from a test system.
- Dpm - published from a test system.
- StoRM – did not make EMI 3.0 cutoff.
- NGI-Italy is working on publishing for other storage systems from the BDII.
- EMI produced tests to verify that StAR records agree with standard

  But that doesn't address requirements on semantics.

- StAR publishing uses the same SSM infrastructure as the new EMI cpu client, and the current OSG, DGAS, SGAS, CERN, Lyon, CPU publishing

EMI JRA1 Tech | OGF https://www.g | Inspire sa1 201 | Inspire jra1 201 | (2) Welcome! | | Overview of th

goc-accounting.grid-support.ac.uk/storagetest/

SSC | RCUK SSC Oracle Po... | BBC - Homepage | Intranet | Indico [LHC Compu... | STFC | STFC Intranet Home | » | Other bookmarks

EMI JRA1 Techni | OGF https://www.ggf | Inspire sa1 2013- | Inspire jra1 2013- | (2) Welcome! | Li | Overview of the

goc-accounting.grid-support.ac.uk/storagetest/

SSC | RCUK SSC Oracle Po... | BBC - Homepage | Intranet | Indico [LHC Compu... | STFC | STFC Intranet Home | » | Other bookmarks

desycertfb_cms__20130225T132005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 13:10:05	2013-02-25 13:20:05	692252908	692254956	None
desycertfb_cms__20130225T133006Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 13:20:05	2013-02-25 13:30:06	692252908	692254956	None
desycertfb_cms__20130225T134005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 13:30:06	2013-02-25 13:40:05	692252908	692254956	None
desycertfb_cms__20130225T135006Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 13:40:05	2013-02-25 13:50:06	692252908	692254956	None
desycertfb_cms__20130225T140005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 13:50:06	2013-02-25 14:00:05	692252908	692254956	None
desycertfb_cms__20130225T141005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 14:00:05	2013-02-25 14:10:05	692252908	692254956	None
desycertfb_cms__20130225T142006Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 14:10:05	2013-02-25 14:20:06	692252908	692254956	None
desycertfb_cms__20130225T143006Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 14:20:06	2013-02-25 14:30:06	692252908	692254956	None
desycertfb_cms__20130225T144005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 14:30:06	2013-02-25 14:40:05	692252908	692254956	None
desycertfb_cms__20130225T145004Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 14:40:05	2013-02-25 14:50:04	692252908	692254956	None
desycertfb_cms__20130225T150005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 14:50:04	2013-02-25 15:00:05	692252908	692254956	None
desycertfb_cms__20130225T151005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 15:00:05	2013-02-25 15:10:05	692252908	692254956	None
desycertfb_cms__20130225T152005Z	desycertfb	discordia.desy.de	None	None	None	8	cms	2013-02-25 15:10:05	2013-02-25 15:20:05	692252908	692254956	None
emi3rc-sl6-dpm.cern.ch-537f0bbc-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	1	atlas	2013-02-06 17:04:53	2013-03-08 17:04:53	0	0	0
emi3rc-sl6-dpm.cern.ch-537f1fb2-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	1	alice	2013-02-06 17:04:53	2013-03-08 17:04:53	0	0	0
emi3rc-sl6-dpm.cern.ch-537f5114-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	1	lhcb	2013-02-06 17:04:53	2013-03-08 17:04:53	0	0	0
emi3rc-sl6-dpm.cern.ch-537f42b2-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	1	cms	2013-02-06 17:04:53	2013-03-08 17:04:53	0	0	0
emi3rc-sl6-dpm.cern.ch-537f536a-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	1	dteam	2013-02-06 17:04:53	2013-03-08 17:04:53	0	0	0
emi3rc-sl6-dpm.cern.ch-537f63c8-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	2	ops	2013-02-06 17:04:53	2013-03-08 17:04:53	0	5242880	0
emi3rc-sl6-dpm.cern.ch-537f79bc-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	12	testers.eu-emi.eu	2013-02-06 17:04:53	2013-03-08 17:04:53	0	4768305152	0
emi3rc-sl6-dpm.cern.ch-537f8af6-707f-11e2-abac-00155dff44e0	None	emi3rc-sl6-dpm.cern.ch	None	disk	None	1	testers2.eu-emi.eu	2013-02-06 17:04:53	2013-03-08 17:04:53	0	0	0

WLCG
Worldwide LHC Computing Grid

Issues from Testing

- No GOCDB Sitename – may be artifact of test sites.

- No ResourceCapacityAllocated – just Used

- Under discussion

WLCG
Worldwide LHC Computing Grid

Open Issues

- Frequency of publishing

  Sites could collect data frequently and publish averages less frequently, say daily

  Repository could average daily records into months, like CPU.

- What is to be viewed in the portal?

Accounting Portal

- Summarise data by
- Site, VO, Month (Average, min, max)
- Views at any point in tree of
  Tier1, Tier2, Country, NGI
- View Average Storage for a time period and VO selection, for any two of (site/country/Tier2), VO, Date
- E.g. Country(date), VO(Site), VO(date), Site(date) (similar to CPU Portal Views.)
- WARNING! Examples do **not** show real data.

Portal View for Storage

Storage per country per month

- Storage used by LHC VOs at French sites.

NGI-France TB Allocated

alice — atlas — cms — lhcb

NGI-France TB Used

alice — atlas — cms — lhcb
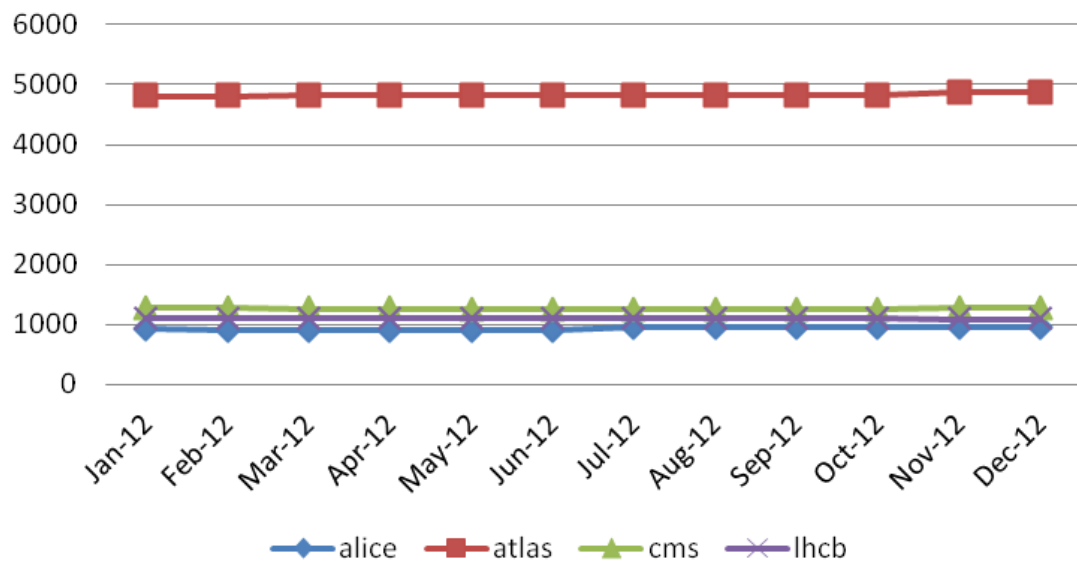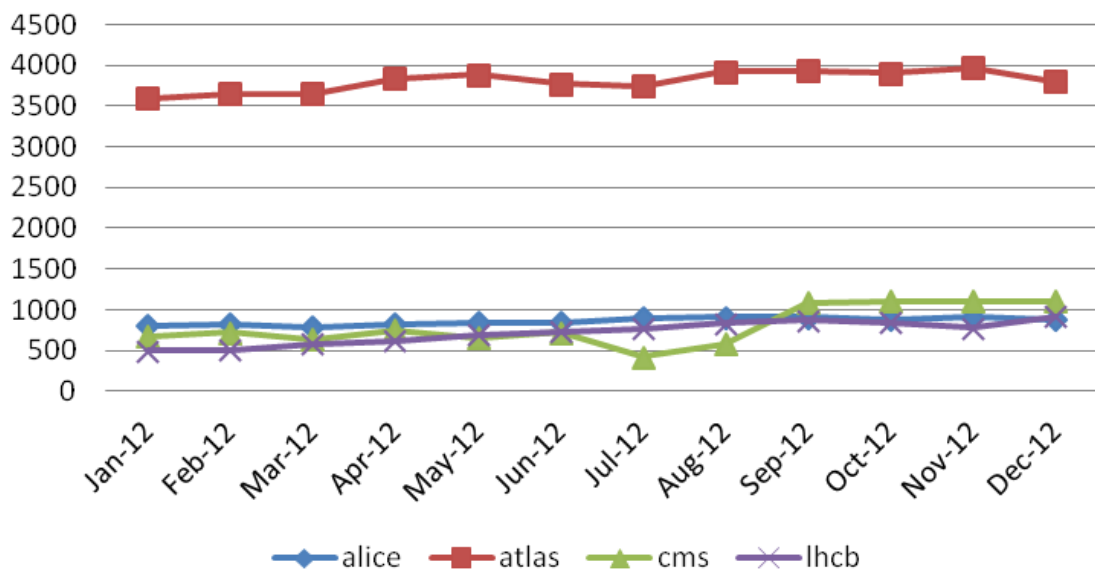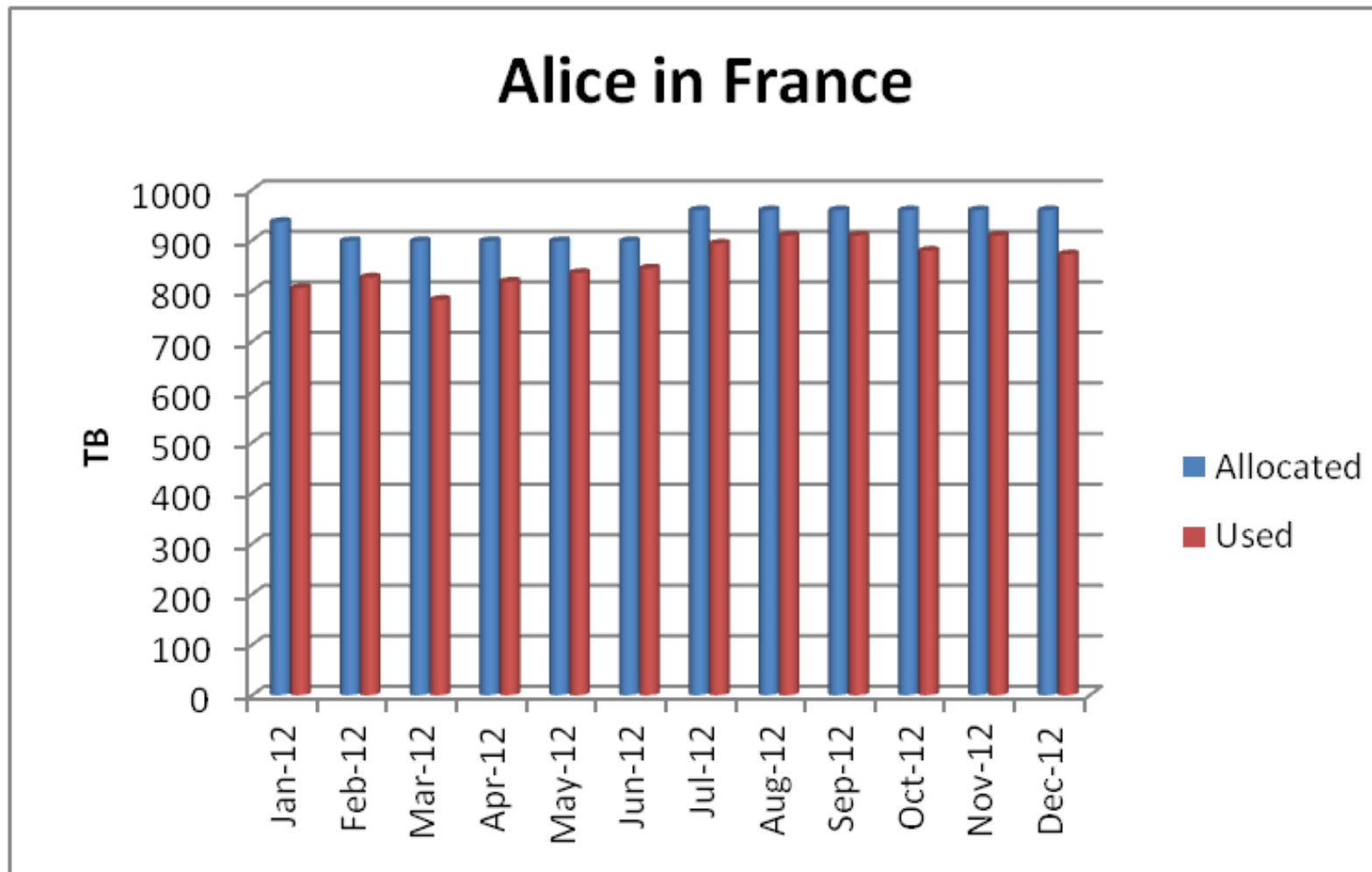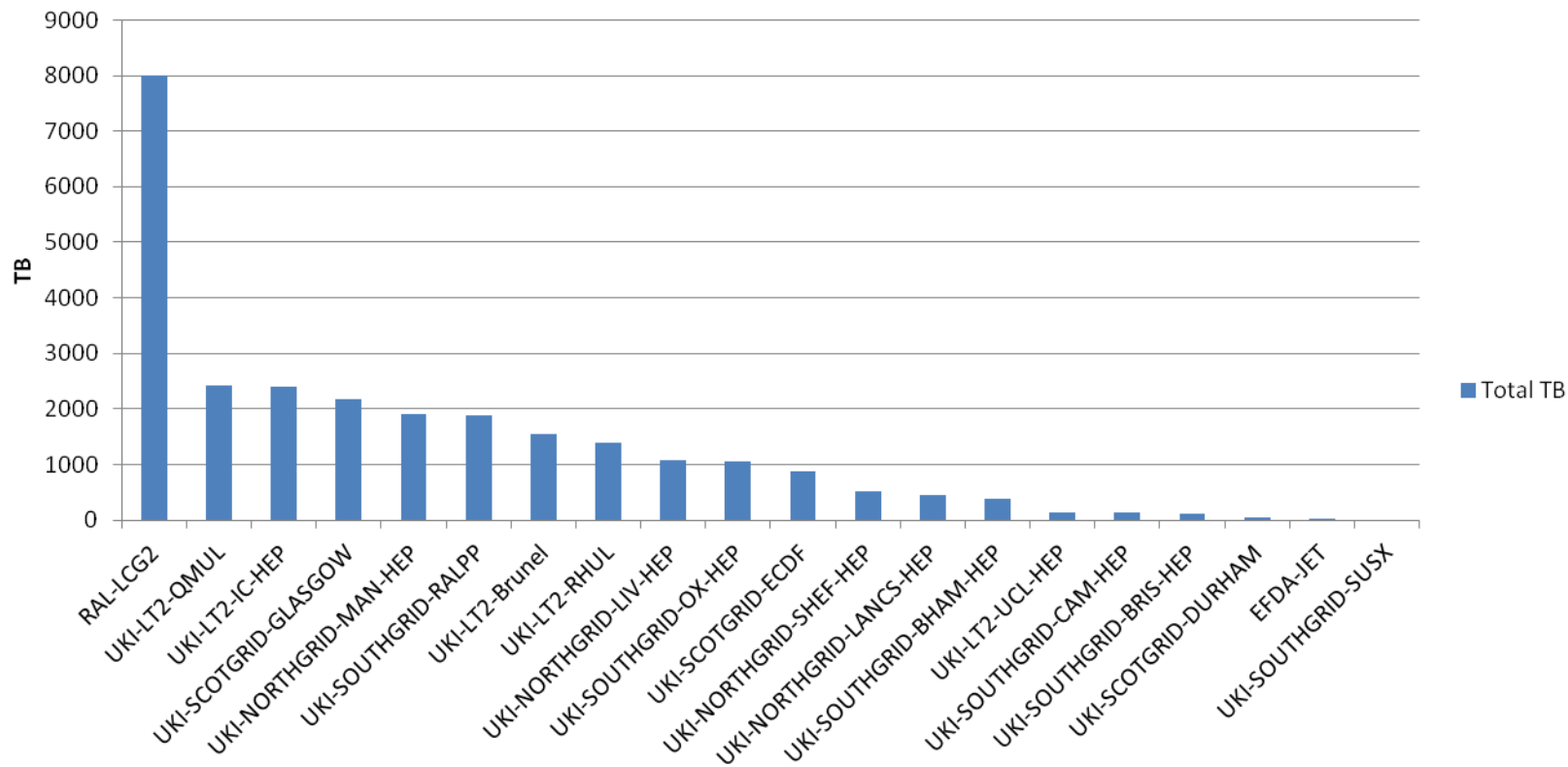
Storage for Alice in France
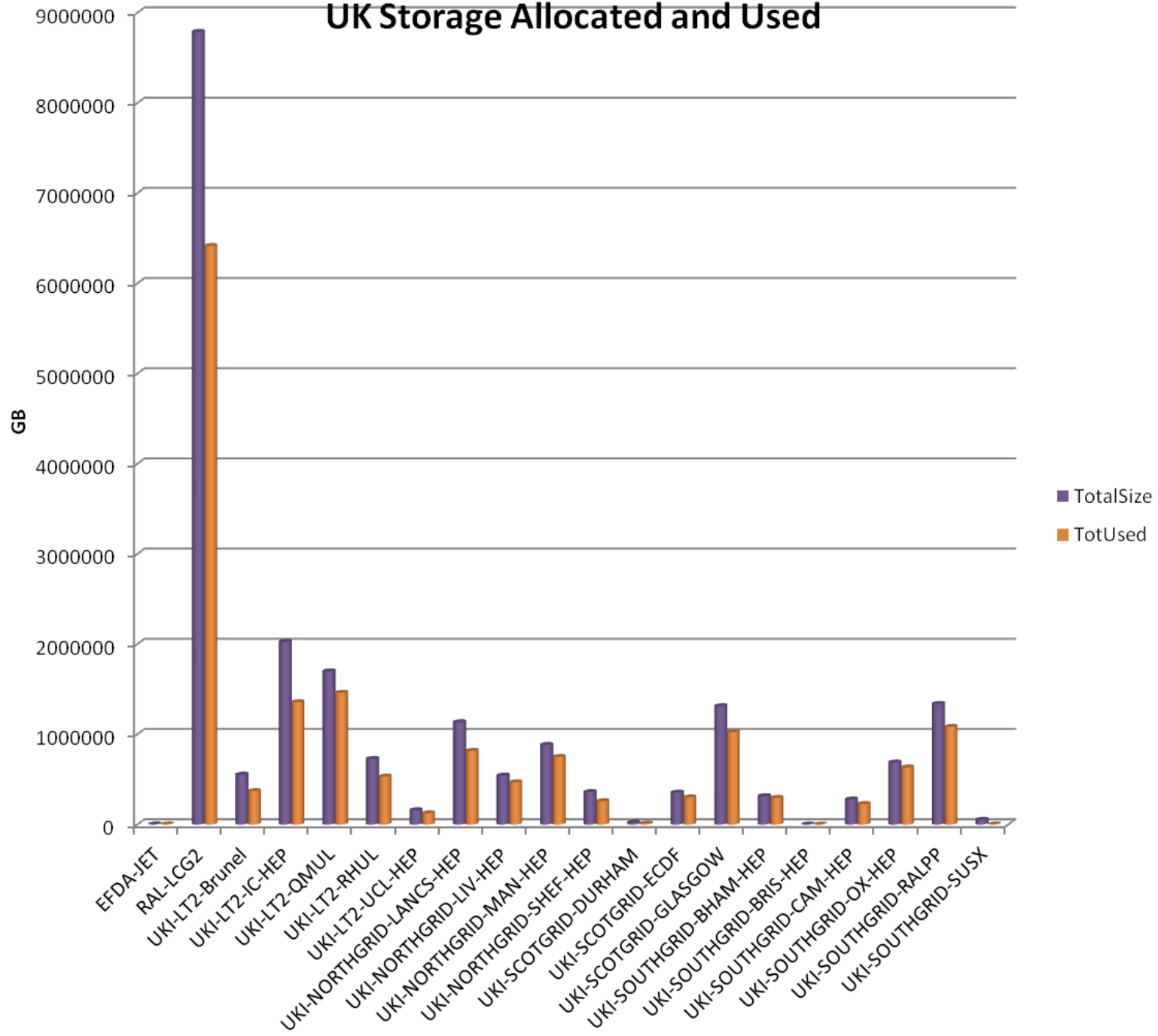
Allocated and used.

Alice in France

Total Storage at UK Sites

UK Storage Allocated and Used

Reports

- When storage accounting is deployed at most sites then:-

    The Tier1 monthly report can be autofilled with storage data.

    The Tier2 report can be extended to cover storage as well as cpu.

Next Steps

- Verify that storage providers are publishing consistent records.

- EMI 3.0 will allow records to be published to extend testing of client, transport, and repository

- Resultant fixes get incorporated in a future EMI3 update (March or April).

- Wider coverage of storage solutions

   StoRM and BDII

- Finalise what is shown in portal

- Implement Portal view and reports.

Conclusions

- Storage Accounting is making (slow) progress.
- There will not be much data to show the CRRB but it can be reported that it is happening.
- Are sites and experiments happy with suggestions for portal displays?
- Are there suggestions of other views or reports?
- Questions?

APEL – Data Retention

- Last year EGI discussed the APEL Data Retention Policy which stated that user data should not be kept for longer than 12 months. We agreed to extend this to 18 months.

- We are just getting round to implementing this and want to clarify the implications.

- **We keep the anonymous summaries of all records from the beginning of time (2004) and will continue to do so.** This is what you can view today in the Accounting Portal

- In future we will **delete individual Job Record data older than 18 months** at RAL. i.e. We do not keep anonymous job records without the UserDN for longer.

- Highlighting this in case there is any misunderstanding. Today, if you have the privileges, youcan see UserDN data going back more than a year.

- UserDN data older than 18 months will be deleted on 1st July and every month thereafter.

... Just one more thing.

- EGI takes requests for features in the Accounting Portal.

- If you have any requests for different views or reports, please discuss them with me, or circulate round GDB mailing list and we can work up the requirement.