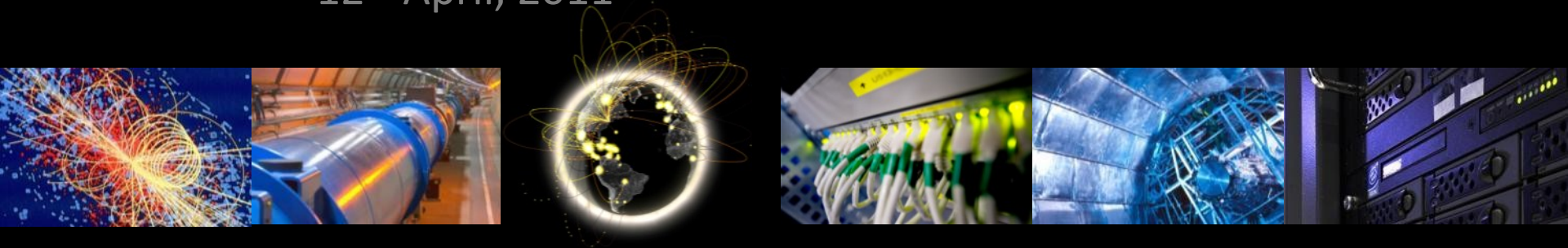


Project Status Report

Ian Bird

Computing Resource Review Board

12th April, 2011



Project status report

- Progress since last RRB
 - HI run, winter technical stop activities
- Reliabilities and readiness
- Service incidents
- Evolution of resource requests
 - Brief summary – more details from C-RSG
- Evolution
 - Data management evolution
 - LHCOPN and evolution
 - Tier 0



Worldwide LHC Computing Grid



>5GB/s to tape during HI
~ 2 PB/month to tape pp
~ 4 PB to tape in HI

LHCOPN TOTAL Traffic Flow 1 (Out-bound)

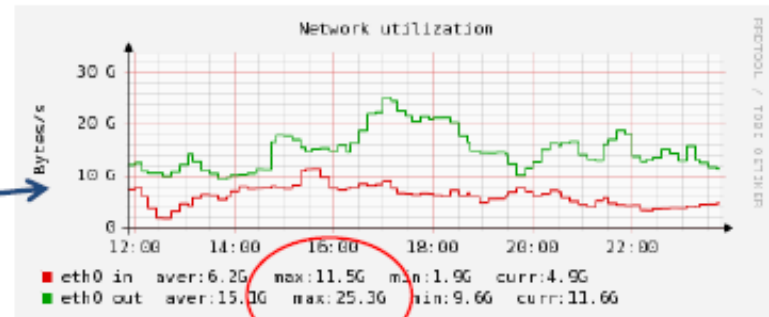
Legend:

- To CMAP
- To DE-KIT
- To IM2P3
- To NDGF
- To RAL
- To RLT1
- To ASGC
- To TRIUMF
- To BM
- To P...

Experiment	Avg	Max	Peak
To CMAP	759.61 M	4.98 G	9.89 G
To DE-KIT	1.36 G	6.88 G	9.23 G
To IM2P3	1.54 G	5.58 G	9.46 G
To NDGF	322.14 M	2.71 G	6.86 G
To RAL	1.03 G	5.55 G	8.83 G
To RLT1	768.05 M	6.22 G	9.46 G
To ASGC	872.12 M	4.11 G	7.93 G
To TRIUMF	567.20 M	2.12 G	N/A
To BM	1.55 G	6.81 G	N/A

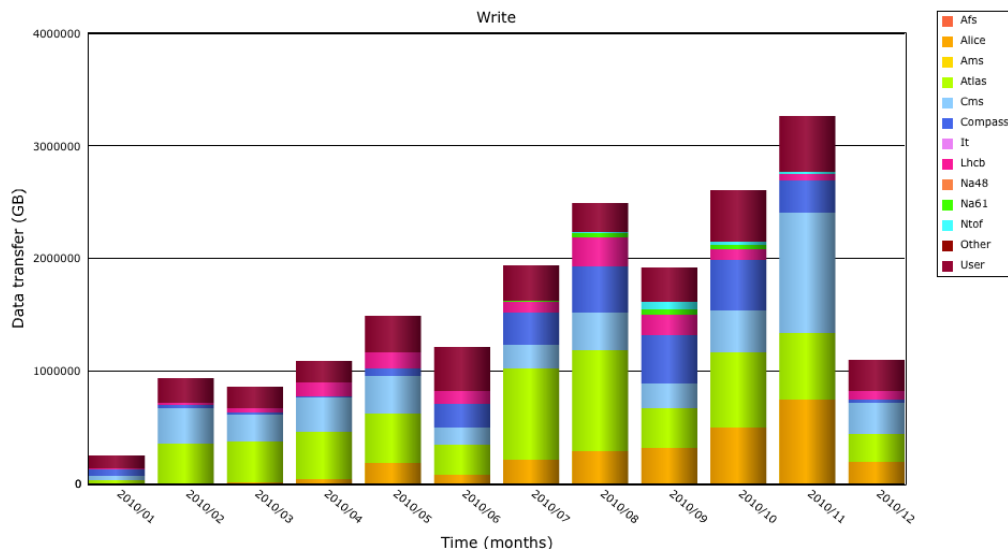
Total t
Total t

>5GB/s to tape during HI
~ 2 PB/month to tape pp



Summary of data archiving - 2010

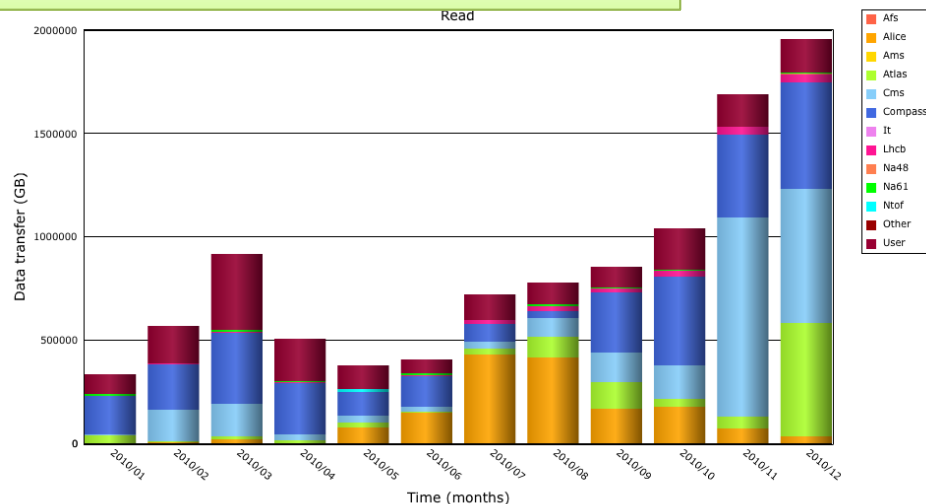
Data written into Castor – 2010: 19 PB



- Continuous evolution of software meant that castor was much more efficient :

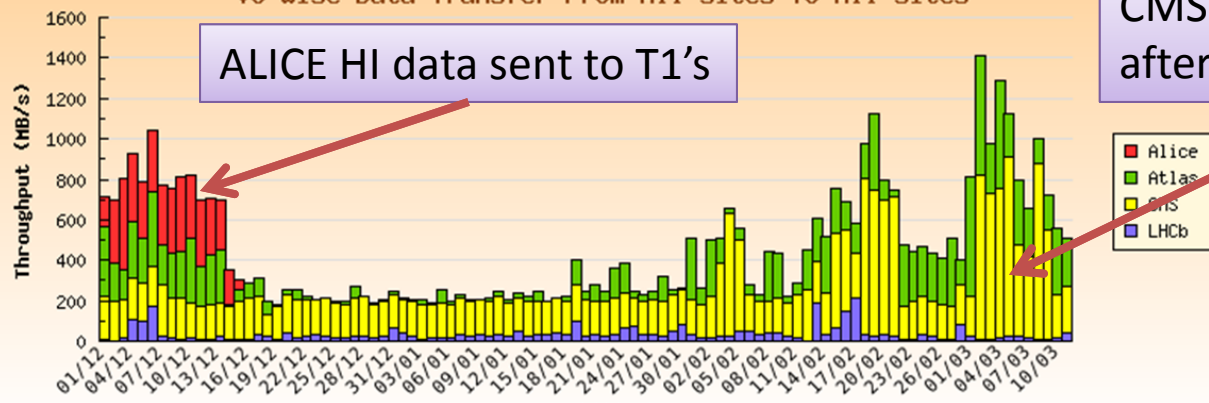
- improved scheduling
- Larger files written (in H1 – up to 40 GB files)
- Used only about 50 drives at peak rates (had been a concern that would need many more) – write speeds were close to native drive speeds

Data read from Castor – 2010: 10 PB



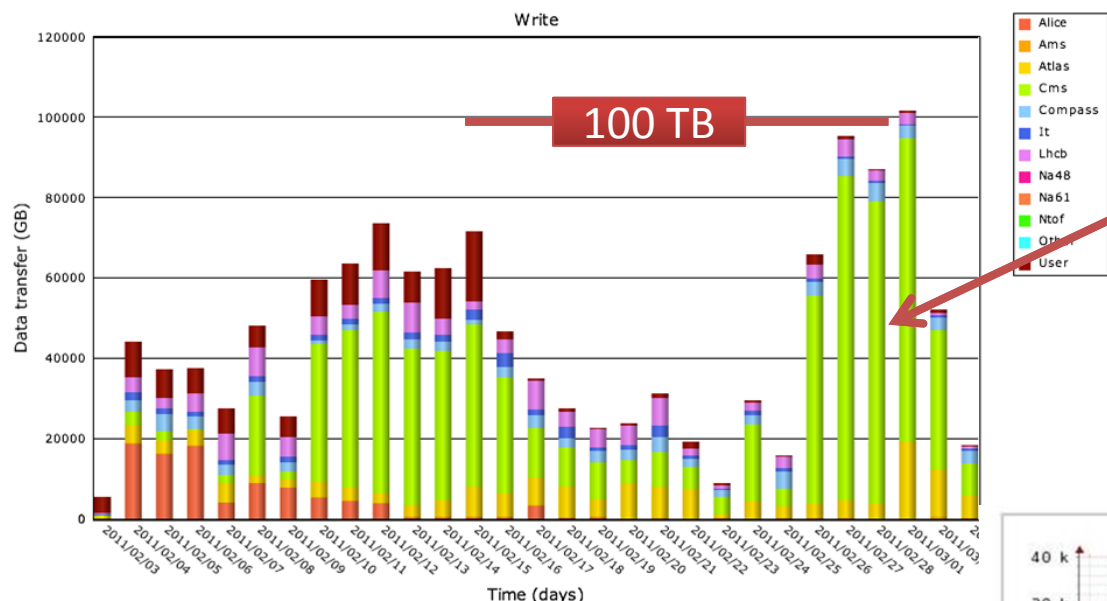
WLCG during winter break

Averaged Throughput From 01/12/10 To 11/03/11
VO-wise Data Transfer From All Sites To All Sites



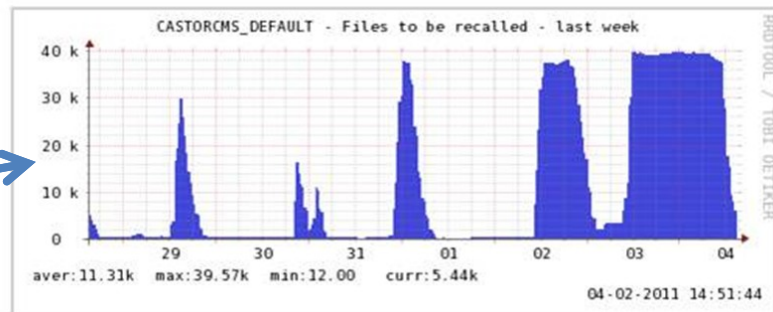
ALICE HI data sent to T1's

CMS HI data → FNAL after zero suppression



Data → tape at CERN in Feb/Mar: CMS HI data after 0 suppression

Huge recall of data (~2 M files) from Castor in very short time without problem.



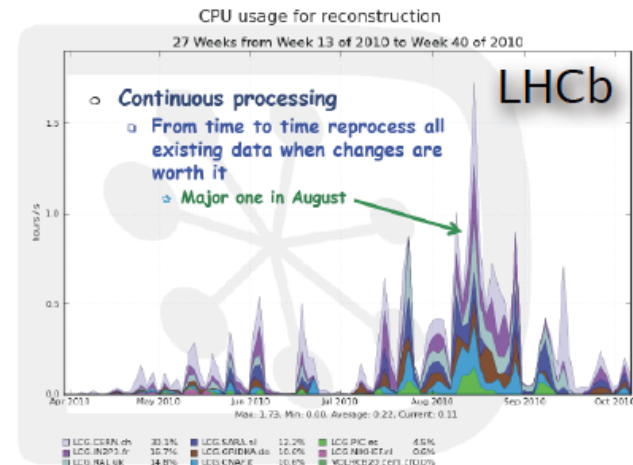
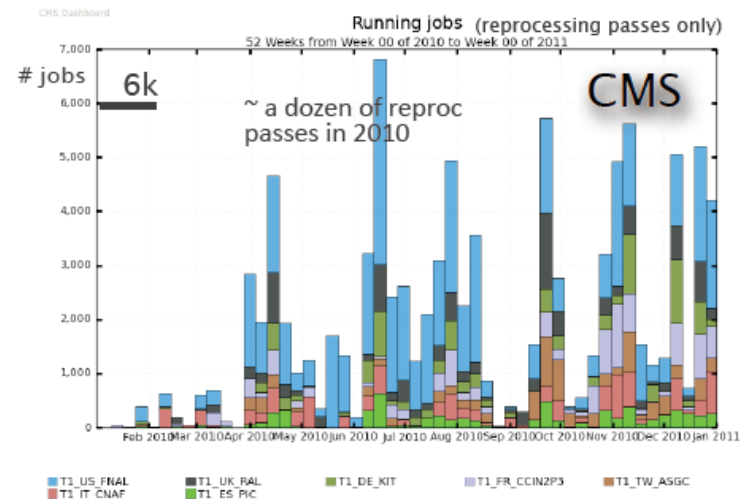
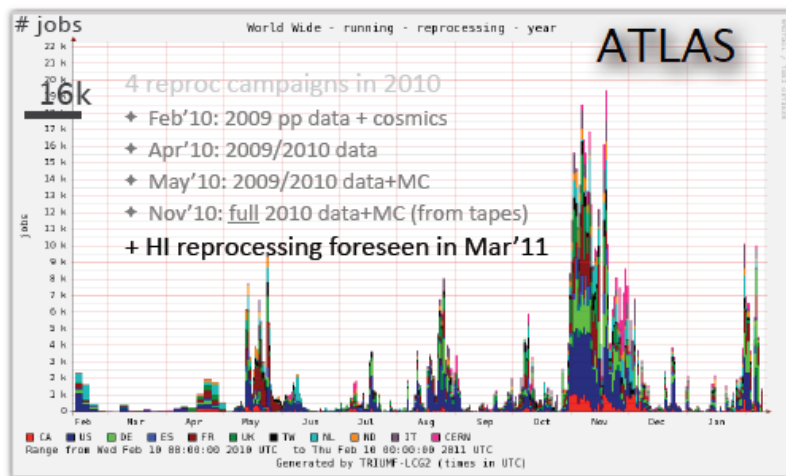
Activities during winter break

- ALICE:
 - Raw data copied to Tier 1s following run; reached max of 260 MB/s during HI run itself.
 - PbPb data have been reconstructed once, 2nd pass in progress
- ATLAS:
 - Full reprocessing of 2010 pp data and re-distribution completed by end 2010
 - HI data processed + re-processed
- CMS:
 - Full reprocessing of 2010 pp data by end 2010
 - Castor rates on 2GB/s in and 3-5 GB/s out to farm *on average* during HI run
 - Zero suppression of HI data under way – storing at Tier 0 and FNAL
- LHCb:
 - Full reprocessing of all 2010 data by end of 2010
 - Major MC production campaign under way
 - Disk clean up campaign; can store fewer copies of data since higher pile-up means event size is larger than anticipated

No break in computing activities ...

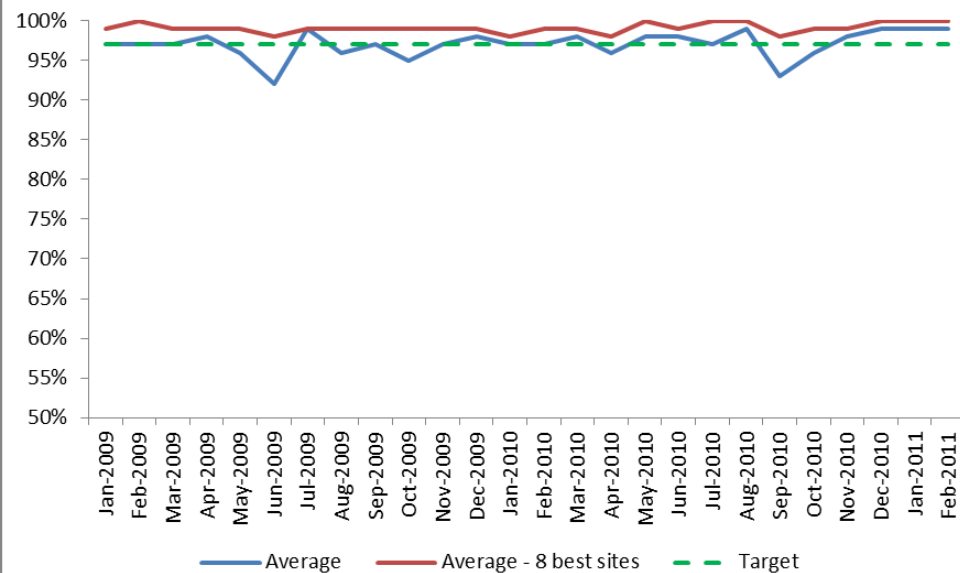
Reprocessing....

- Data at Tier 1s – reprocessed as needed
 - New calibrations, improved software, new data formats

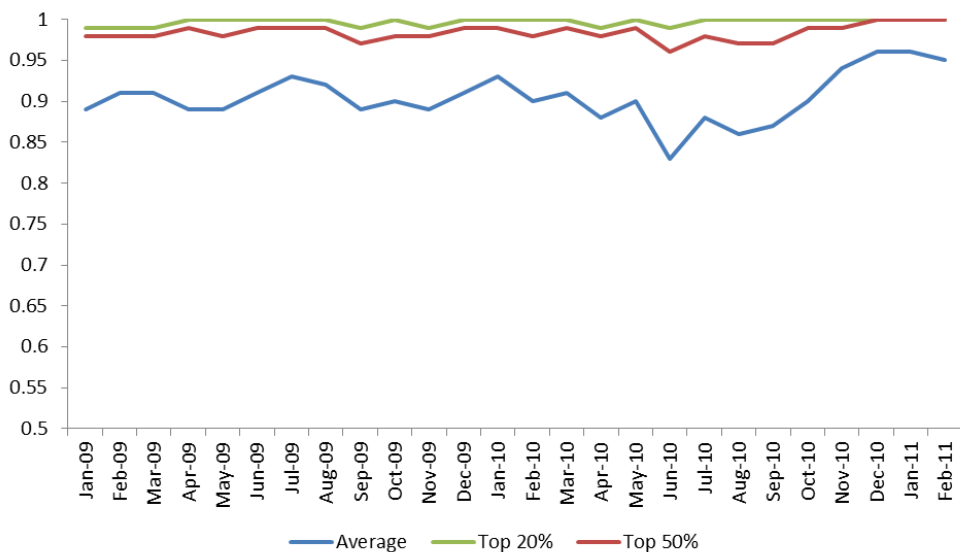


Reliabilities

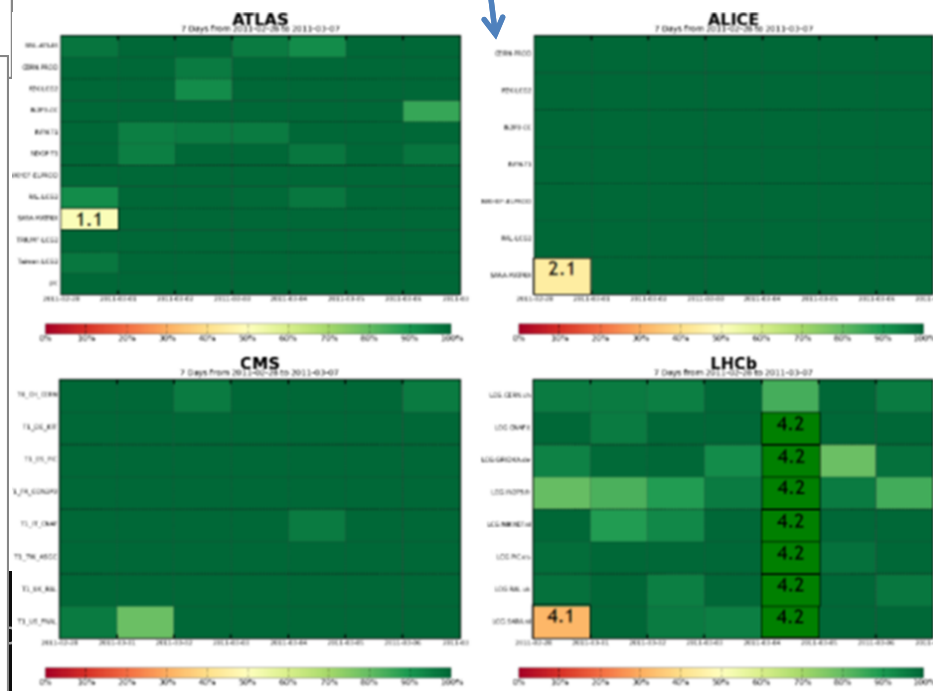
Site Reliability: CERN + Tier 1s



Tier 2 Reliabilities



Experiment-measured site readiness:
includes experiment specific tests as well as availability
testing



Service Incidents

<u>Site</u>	<u>Service Area</u>	<u>Date</u>	<u>Duration</u>	<u>Service</u>	<u>Impact</u>
CERN	DB	11-Mar-11	5h	CMS offline production db	The database was completely down for ~2 hours and partially not available for 5 hours
IN2P3	Infrastructure	Feb 25-26	13h	Batch system	85% of batch system unavailable, jobs lost
IN2P3	Storage	Feb 13	3 h	Storage service	Storage services degraded, no big impact on jobs
PIC	Storage	21-Jan-11 to 08-Feb-11	18 days	Storage service	250TB of ATLAS data partially unavailable
KIT	infrastructure	28-Jan-11 to 02-Feb-11	5 days	Batch system, job submission	batch system degraded, reduced # of job slots available
CERN	DB	25-Jan-11	8h	FTS, LFC, SAM dashboards	LHCb jobs use area intensively at start; eventually fixed with new AFS client, cache tuning on many-core boxes
IN2P3	infrastructure	8-Jul-10 to 7-Jan-11	6 months	shared s/w area	Not really solved ... also shows problems in "3 rd party" problem resolution. Better defined procedure has resulted, backup circuit now "primary"; only just closed.
CNAF-BNL	network	23 Aug-10 to 20-Jan-11	months	primary OPN circuit	
<u>Site</u>	<u>Date</u>	<u>Duration</u>	<u>Service</u>	<u>Area</u>	<u>Impact</u>
CERN	18 Dec	5 days	DB	DB	Service interruption: ATLARC DB following the power cut at CERN CC
CERN	18 Dec	26 hours for services with weight > 50	power	infrastructure	Interruption of physics services following power cut
CERN	16 Dec	2.5h	DB	DB	ATLR database affected (degradation then complete outage) by FC switch replacement

LHCOPN – better monitoring

LHCOPN Dashboard

Home

Current status

Last update: Wed, 30 Mar 2011 13:32:55 GMT

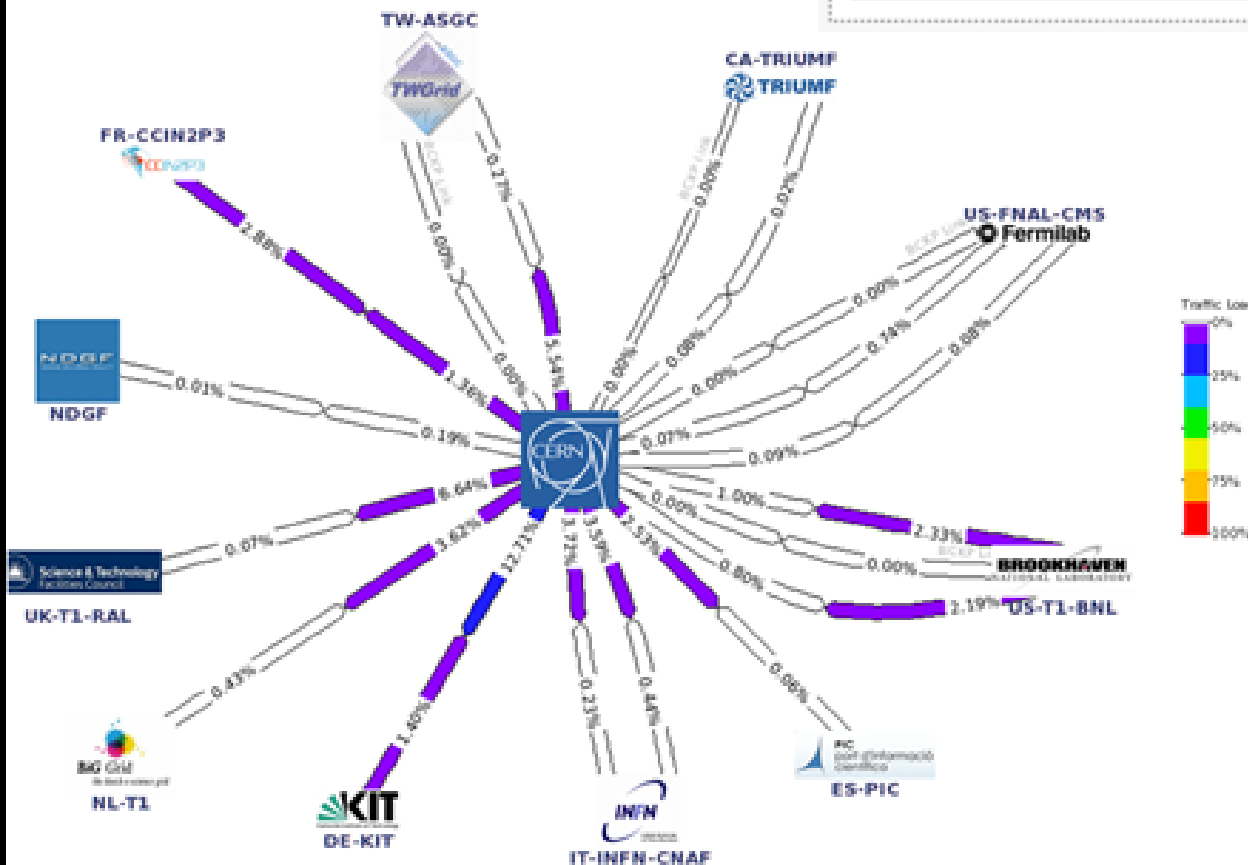
From / To	CA-TRIUMF	CH-CERN	DE-KIT	ES-PIC	FR-CCIN2P3	IT-INFN-CNAF	NDGF	NL-T1	TW-ASGC	UK-T1-RAL	US-FNAL-CMS	US-T1-BNL
CA-TRIUMF												
CH-CERN												
DE-KIT												
ES-PIC												
FR-CCIN2P3												
IT-INFN-CNAF												
NDGF												
NL-T1												
TW-ASGC												
UK-T1-RAL												
US-FNAL-CMS												
US-T1-BNL												

JSON feed

Legend: OK Deviation from Baseline Critical Unknown



LHCOPN

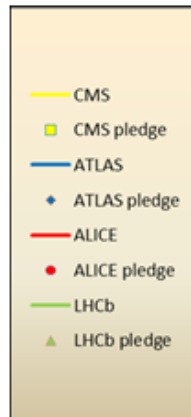
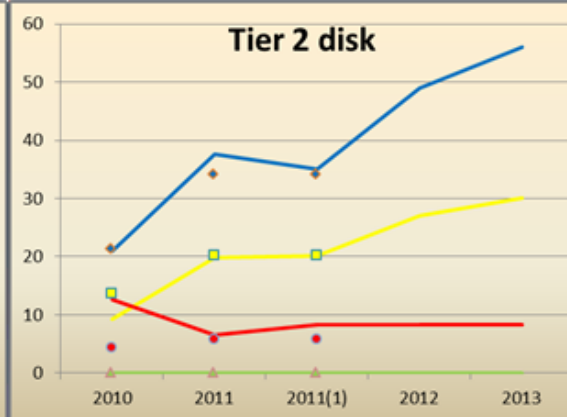
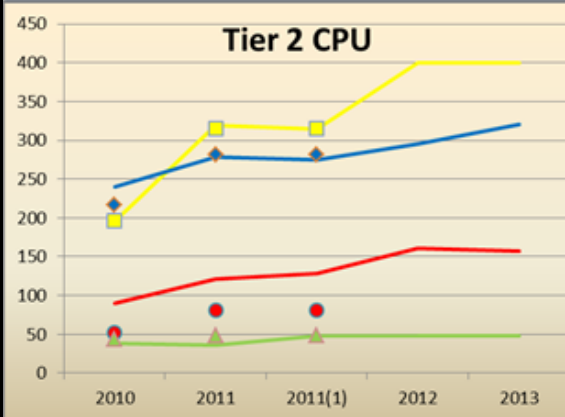
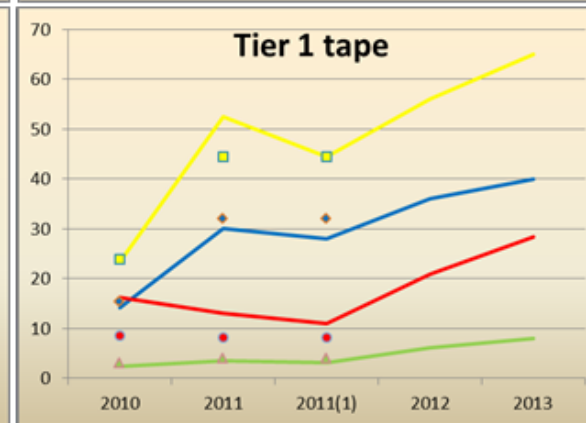
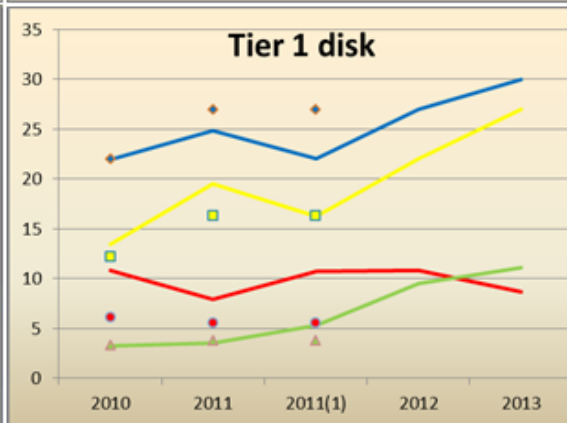
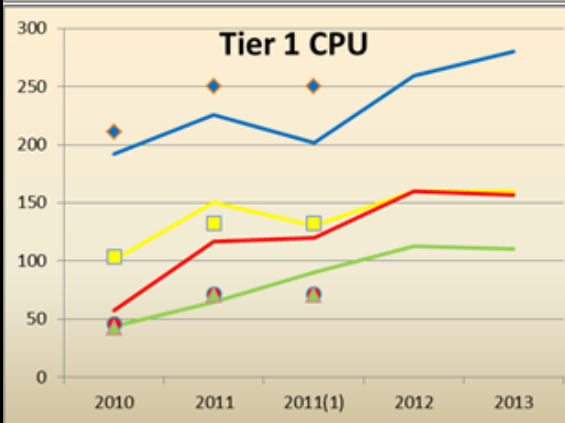
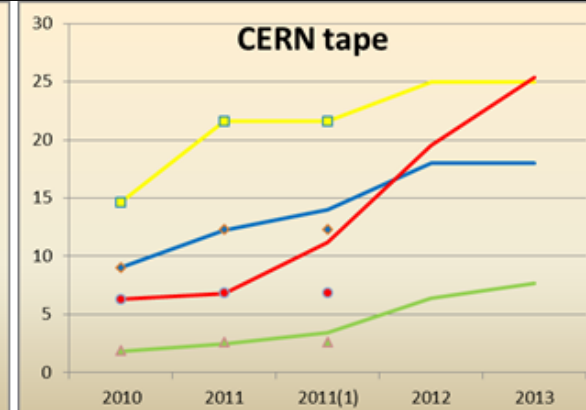
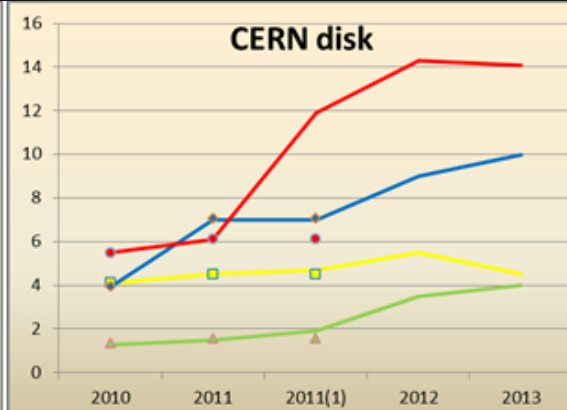
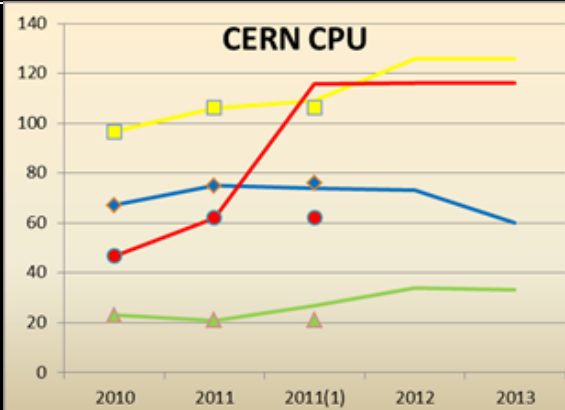


Created: Mar 30 2011 15:34:03

Evolution of requirements

- Change in LHC schedule, run in 2011 and 2012, break in 2013
- Considerations based on real experience from 2010:
 - Much higher pile-up than anticipated for this stage of LHC
 - Drives event sizes and thus reconstruction times and storage space needs – in some cases factors of 2 or more
- More discussion in C-RSG report

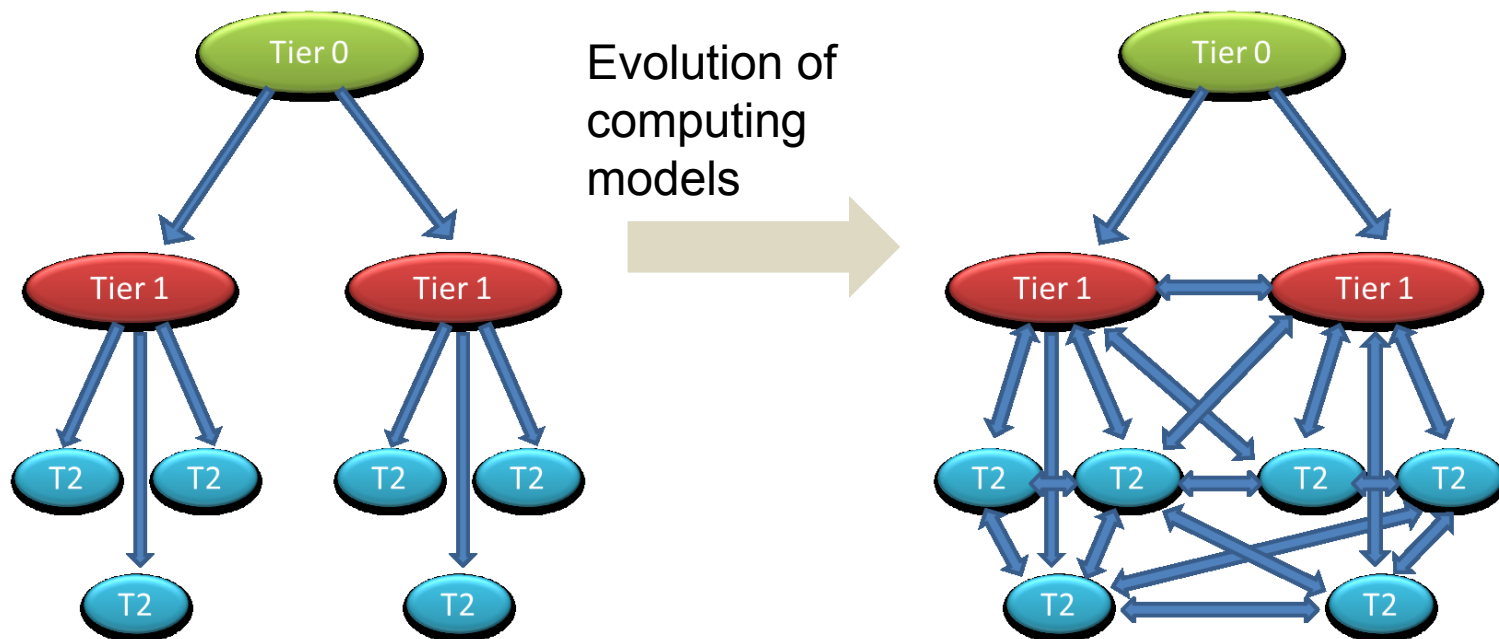
Evolution of requirements



Data management – evolution

- Overall goal to improve data access and availability, especially for analysis:
- Follow up of prototyping activities
 - Started in Amsterdam meeting last June
 - several “demonstrators” proposed
 - Follow up in January – to wrap up process as agreed
- ~10 demonstrators actively being investigated or followed by one or more experiments
 - High level of commonality
 - Based on real needs
 - Several are related and will likely converge

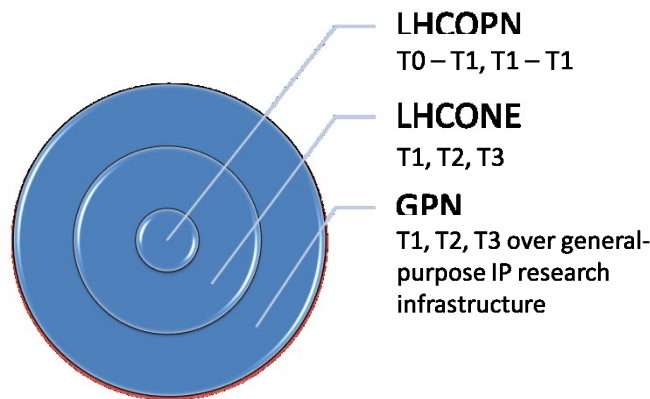
Network evolution



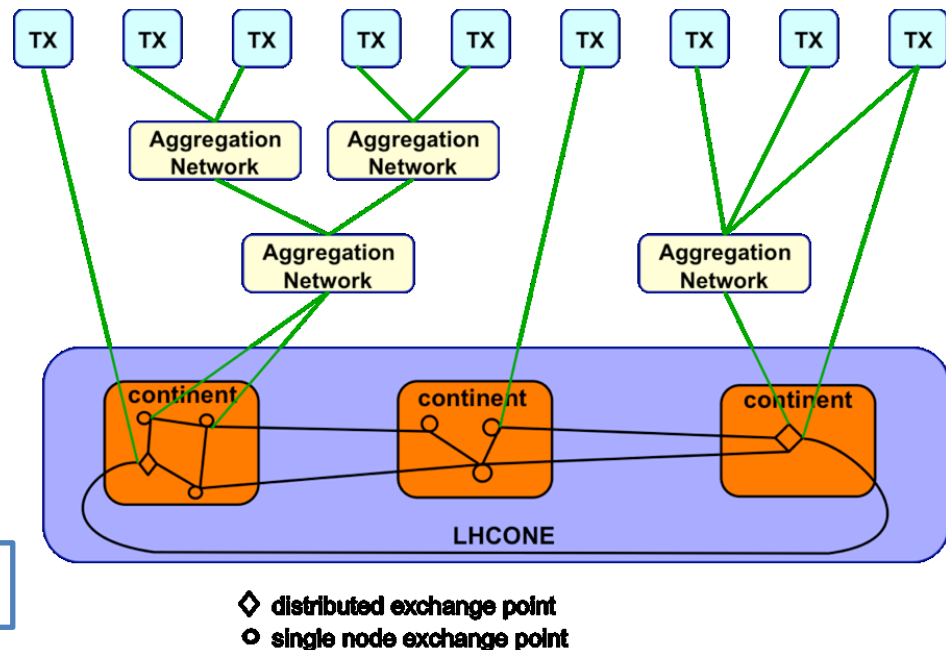
- On-demand will augment data pre-placement
- Network usage will increase & be more dynamic (less predictable)
 - need to enable high-volume data transport between T1s, T2s, and T3s.
- General-purpose R&E networks should not be swamped with Tier1/2/3 LHC traffic

- LHCOPN works well - so don't break it!
- T1/2/3s are not all equal -> architecture must be all-inclusive
- No central structure - distributed management & operations
- Design for agility and expandability (new technologies; changes in traffic volumes)
- Must be appealing to funding agencies
- Need good monitoring

Proposed solution → LHConc



LHCONC → LHC Open Network Environment



- LHCONC builds on the idea of exchange points
- Exchange points will be built in carrier-neutral facilities so that any connector can connect with their own fiber or using circuits provided by any telecom provider.
- LHCONC enables T2s and T3s to obtain their data from any T1 or T2
- Use of LHCONC will alleviate the general R&E IP infrastructure
- LHCONC provides a framework for future network requirements, such as the European Grid, ESnet, CANARIE, etc.
- Next steps:
 - Solicit community input
 - Build a prototype (first switch installed at CERN)
 - Refine architecture document and work on governance, operations model, monitoring...

NB: Network requirements should in future be (again) included in resource requests, and budgeted for in pledges

Tier 0 Evolution

- Following invitation for interested countries to submit informal proposals for a remote Tier 0
 - Around 25 were received by the end of November
 - Many of them are very interesting, not all can be directly compared
 - Visits and discussions ongoing with a number of these
- More formal steps not likely until later this year
- Given the evolution of the LHC schedule and resources, the requirement is to have additional capacity available by 2014.

Summary

- 2010 was a very successful start for LHC computing
 - Data rates (especially in HI running) exceeded by far those anticipated and tested
 - The full system (Tier 0-1-2) managed these without problem
- Resource usage reached a peak in the latter part of the year with Tier 1 and Tier 2 sites reaching full capacity
 - In 2011 we will need to manage resource contention ...
- Activities continued without break over the LHC technical stop ...
- Based on experience and better understanding of LHC conditions, computing models are evolving; important that commonalities are exploited as far as possible