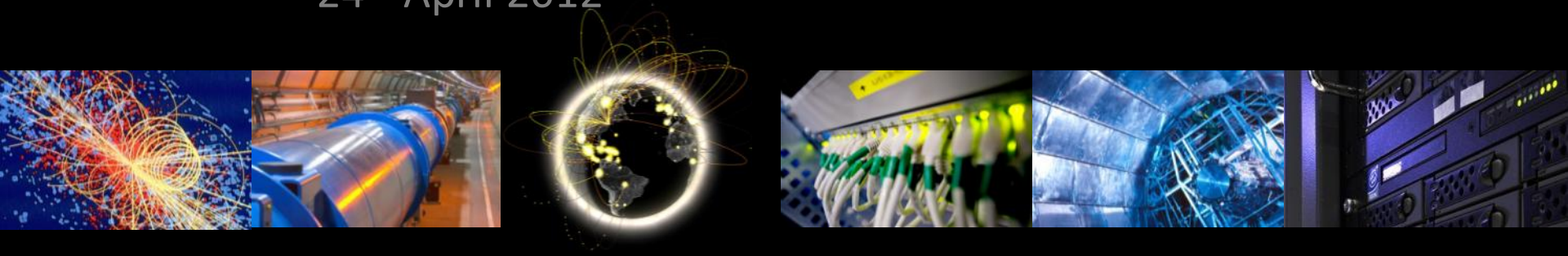
CERN-RRB-2012-042

Project Status Report

Ian Bird

Computing Resource Review Board

24th April 2012



WLCG Worldwide LHC Computing Grid

Outline

- WLCG Collaboration & MoU status
- WLCG status and usage
- Metrics reporting
- Resource pledges
- Funding & expenditure for WLCG at CERN
- Planning & evolution


CERN


US-BNL


Amsterdam/NIKHEF-SARA


Bologna/CNAF

WLCG Collaboration Status
Tier 0; 11 Tier 1s; 67 Tier 2 federations


Ca-TRIUMF


Taipei/ASGC

NDGF

Today we have 49 MoU signatories, representing 35 countries:

Australia, Austria, Belgium, Brazil, Canada, China, Czech Rep, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, India, Israel, Japan, Rep. Korea, Netherlands, Norway, Pakistan, Poland, Portugal, Romania, Russia, Slovenia, Spain, Sweden, Switzerland, Taipei, Turkey, UK, Ukraine, USA.


US-FNAL


De-FZK


Barcelona/PIC


Lyon/CCIN2P3


UK-RAL

WLCG MoU Status

- No additional signatures since last RRB meeting
  - Work with LLNL to enable signature on-going; meanwhile LLNL are reporting as a full Tier 2
- Reminder:
  - All Federations, sites, WLCG Collaboration Representative names and Funding Agencies are documented in MoU annex 1 and annex 2
  - Please check and ensure information is up to date
  - Signal any corrections to lcg.office@cern.ch

New Tier 1 sites

- Several discussions have been held regarding potential new Tier 1 sites
- A formal process has been documented and approved by the March 2012 WLCG Overview Board:
  - https://espace.cern.ch/WLCG-document-repository/Collaboration/New%20Tier1%20Process

Process

- Pre-requisite is that any such proposal must be supported by the experiment(s)
- Balance between encouragement of new sites/resources and reaching high standards of existing Tier 1 services
- Process:
  - Prepare with MB a detailed plan that shows how the site will demonstrate required functionality, performance, reliability; timeline and milestones
  - Present plan to OB: OB recommends acceptance (or not)
  - Site can sign MoU as an *Associate* Tier 1
  - MB monitors progress on ramp up, reports to OB
  - When milestones achieved as agreed by MB, final report to OB to recommend full Tier 1 status
    - This should normally take ~1 year

Requirements

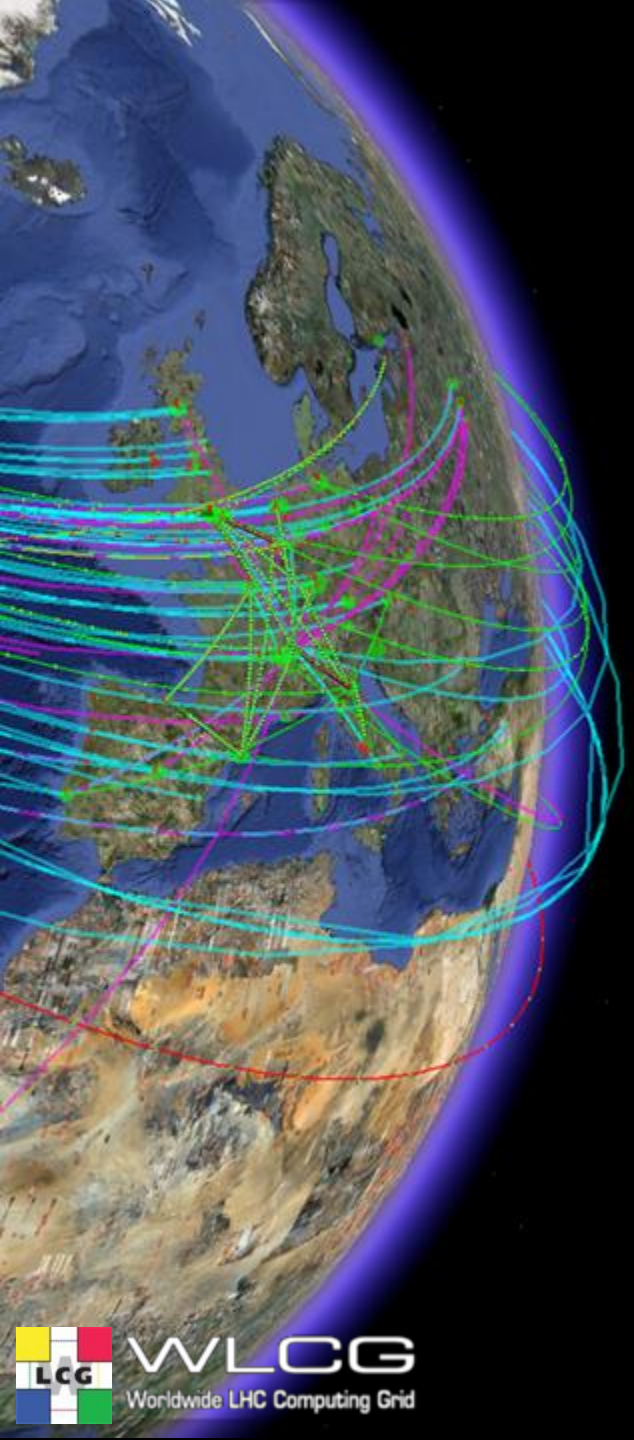
- Most elements are described in the MoU addenda
- Candidate site must achieve MoU requirements in terms of:
  - Level and performance of resources
  - Quality and reliability of services:
    - Set of services agreed with the experiments
    - Provide agreed levels of support – as in MoU. Typically on-call support year round
    - Availability and reliability: install agreed sensors, publish to WLCG monthly (as all other sites)
    - Interface to WLCG accounting, provide accounting data to be published monthly
    - Support for Tier 2s – in agreement with experiments. Data source and technical support for Tier 2s

New Tier 1s

- At the March OB; KISTI (S. Korea) presented an initial proposal as a Tier 1 for ALICE; the OB accepted KISTI as the first "Associate Tier1"
  - A full plan is now being prepared

- Also anticipated:
  - Russia has proposed providing Tier 1 for all 4 experiments
  - Discussions with Mexico for ALICE; and India for ALICE and CMS
  - All t.b.c.

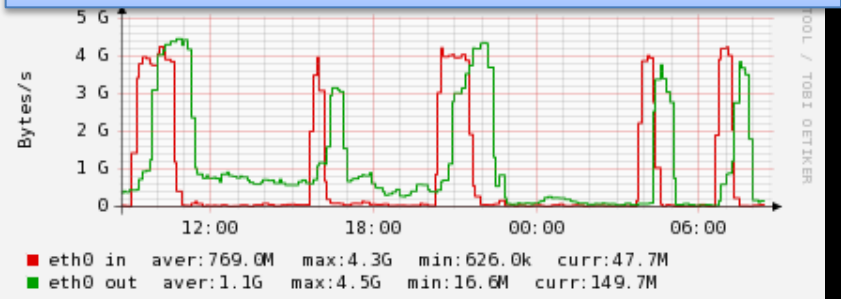
WLCG Status report



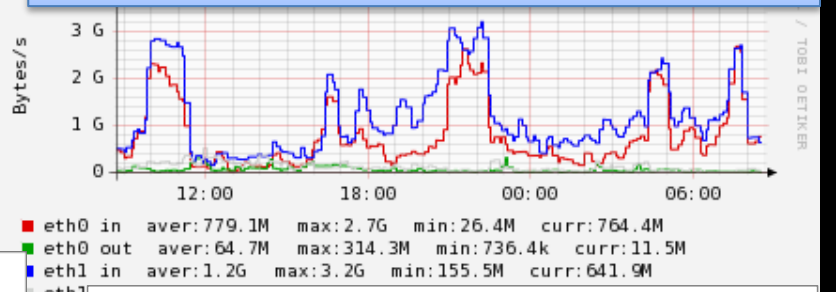
WLCG: Data in 2011

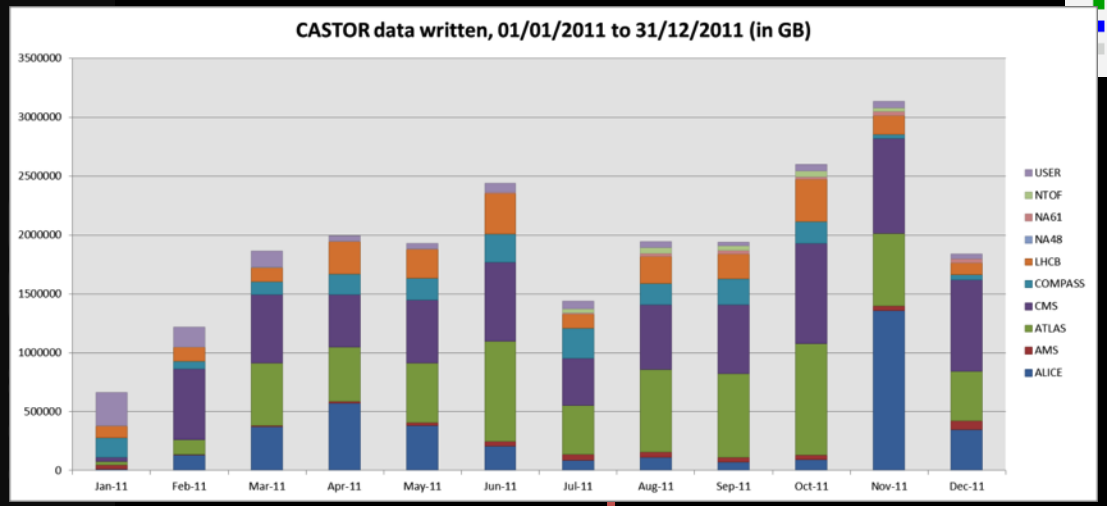
- Castor service at Tier 0 well adapted to the load:
  - Heavy Ions: more than 6 GB/s to tape (tests show that Castor can easily support >12 GB/s); Actual limit now is network from experiment to CC
  - Major improvements in tape efficiencies – tape writing at ~native drive speeds. Fewer drives needed
  - ALICE had x3 compression for raw data in HI runs

HI: ALICE data into Castor > 4 GB/s (red)

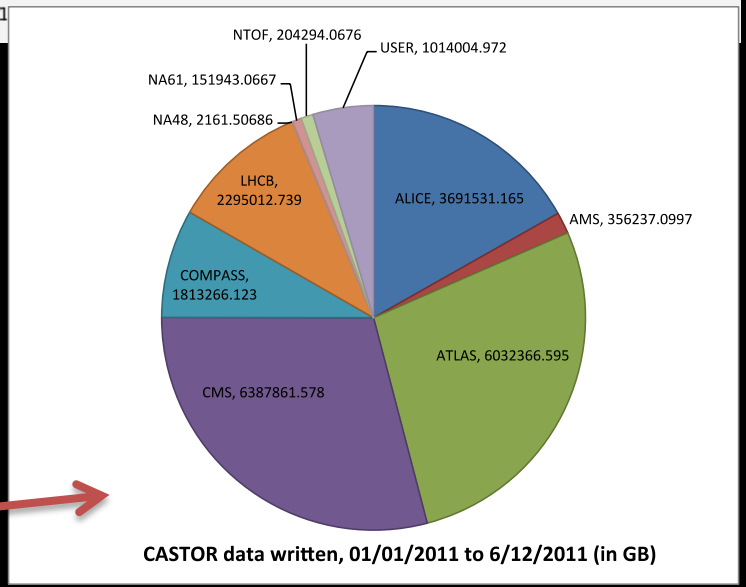


HI: Overall rates to tape > 6 GB/s (r+b)



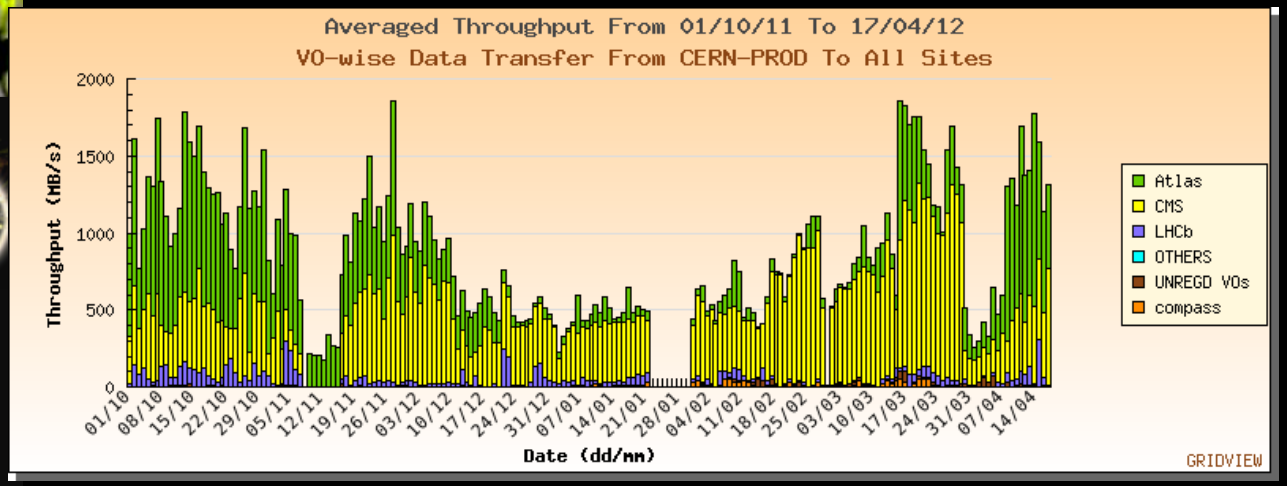


23 PB data written in 2011

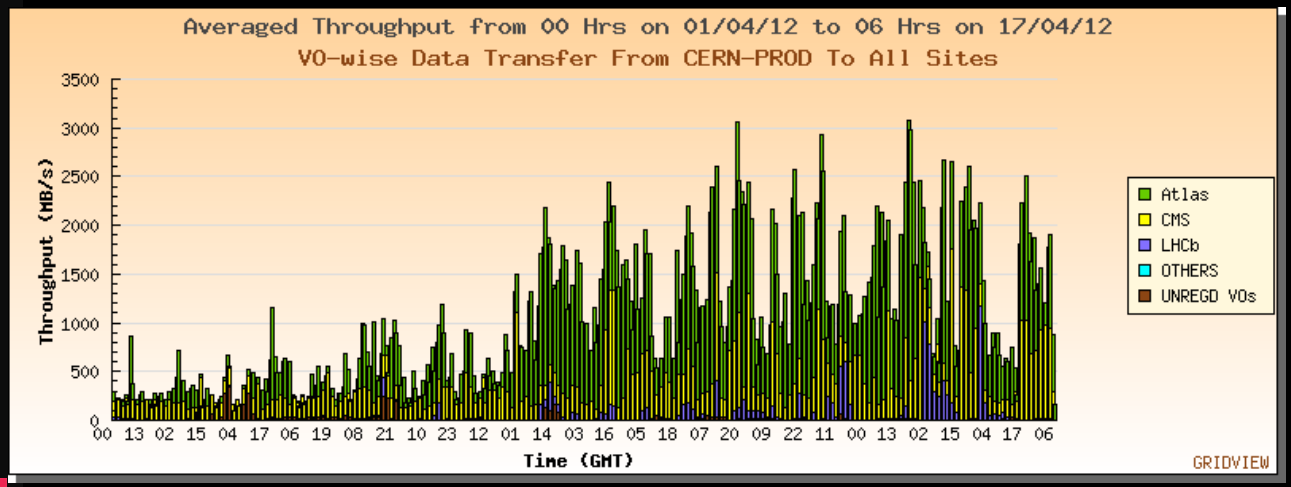


Data transfers…

Already back to “normal” levels for accelerator running



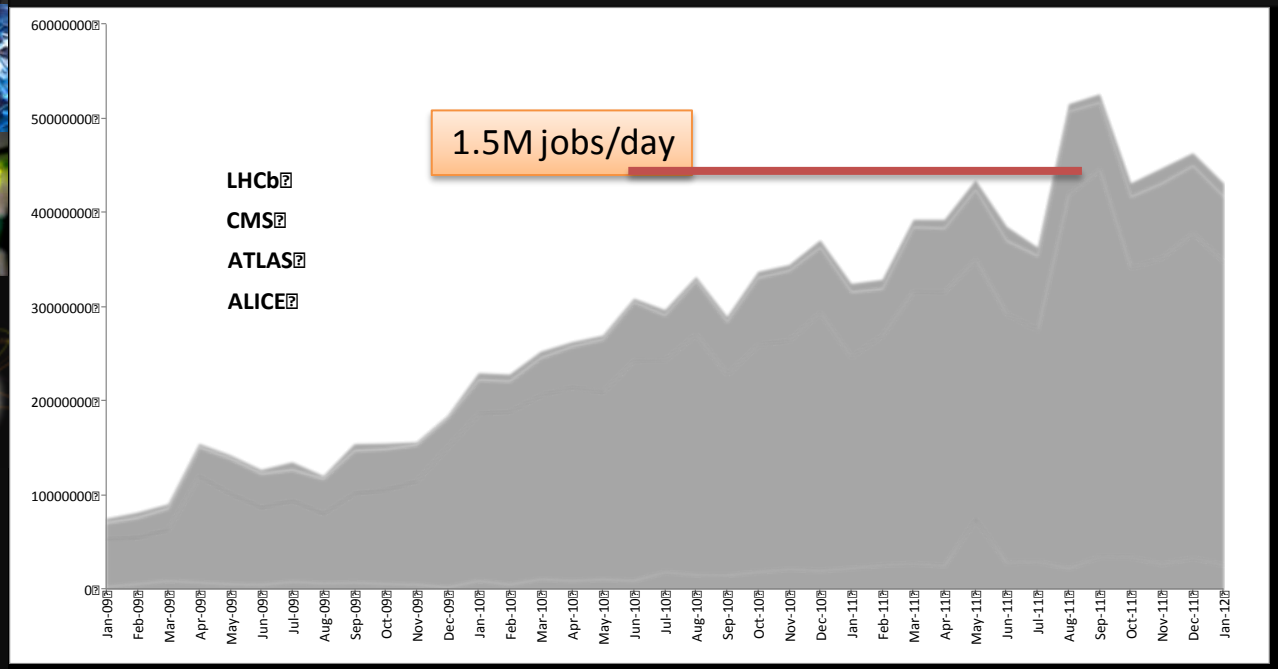
Since last RRB



2012 Data

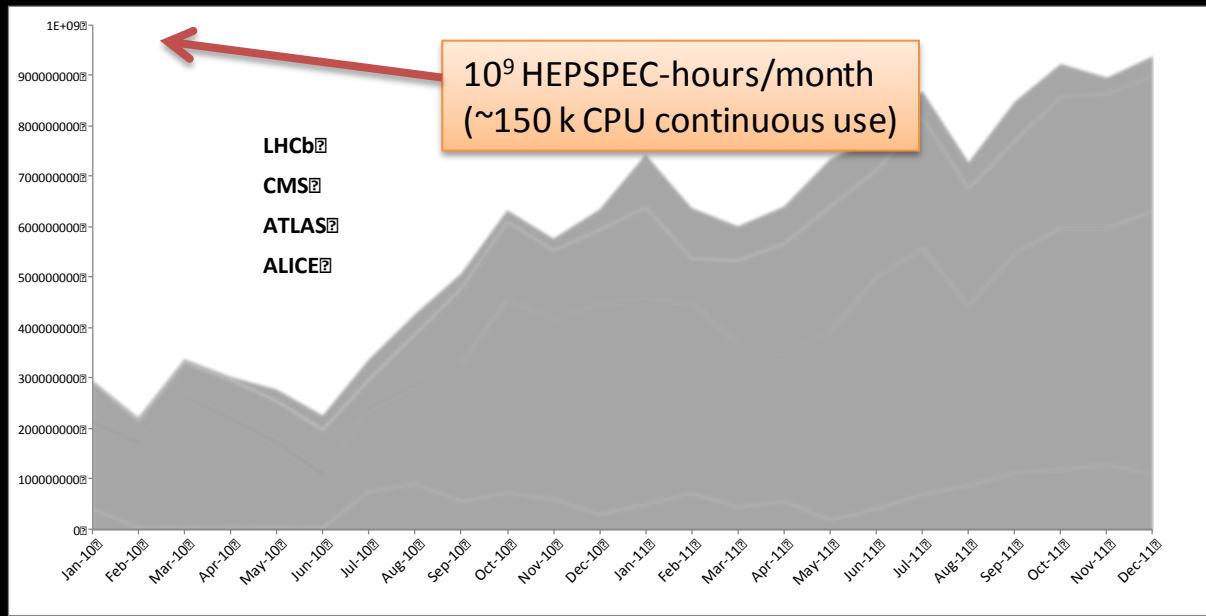
WLCG Worldwide LHC Computing Grid

Overall use of WLCG



Usage continues to grow even over end of year technical stop

- # jobs/day
- CPU usage



WLCG Worldwide LHC Computing Grid

WLCG – no stop for computing

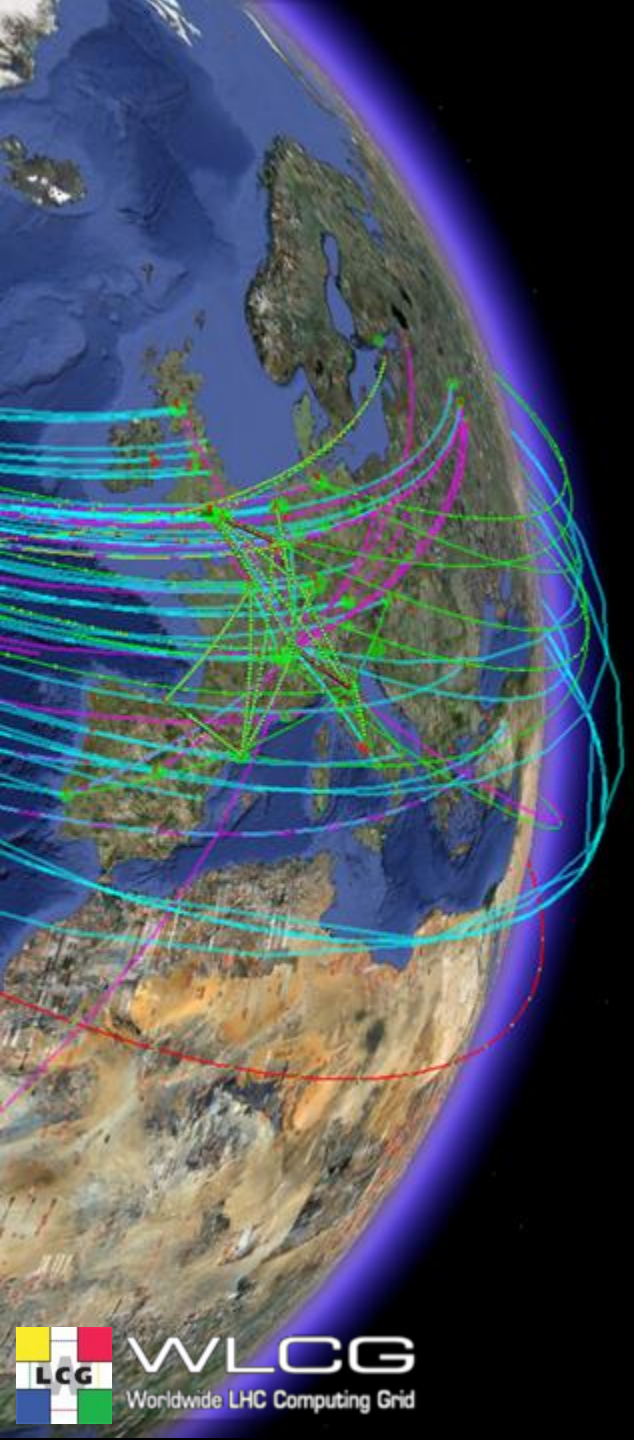
WLCG
Worldwide LHC Computing Grid

Main features of recent use:

- Continued growth in overall usage levels
- High levels of analysis use, particularly in preparation for winter conferences
- Resources → fully occupied
- Full reprocessing runs of full 2011 data samples
  - Achieved by end of the year
- HI: complete processing of 2011 samples
- Large simulation campaigns for 2011 data and in preparation for 8 TeV run
- Disk clean up campaigns in preparation for 2012 data

Other comments

- ATLAS & CMS:
  - Significant work to improve software performance particularly for high pile up conditions
  - Both achieved large factors improvements in reconstruction speed and memory use
- ALICE:
  - Software and infrastructure improvements to deal with low CPU efficiencies seen in 2011
  - Achieved copy of HI data to Tier 1s in a few days (300 MB/s)
- LHCb:
  - Use of Tier 2s for data reprocessing
  - Start to test use of online farm for offline use



Metrics Reporting

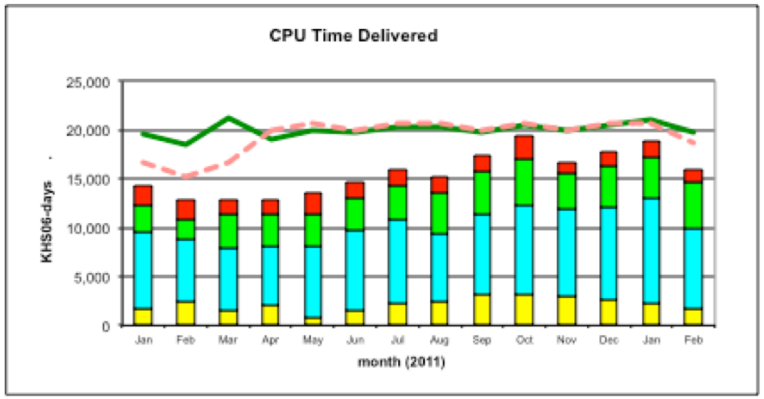
WLCG
Worldwide LHC Computing Grid

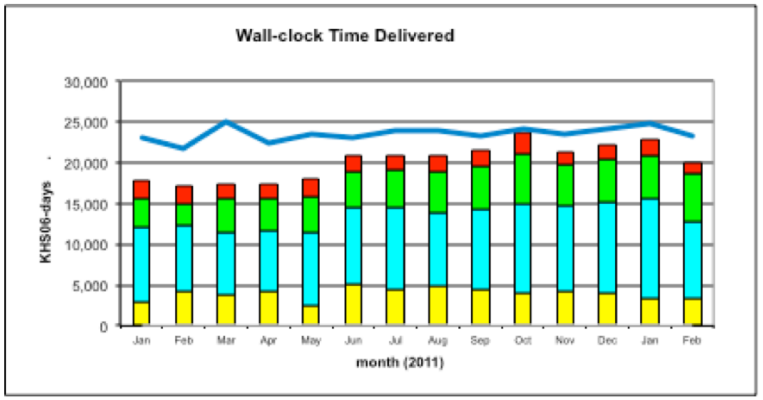
CERN

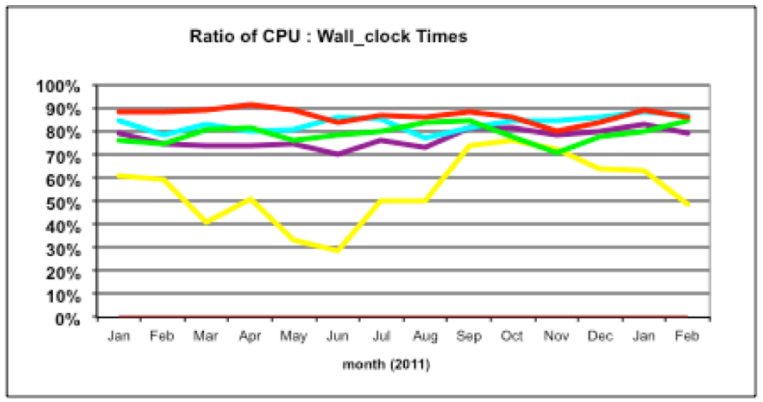
Accounting

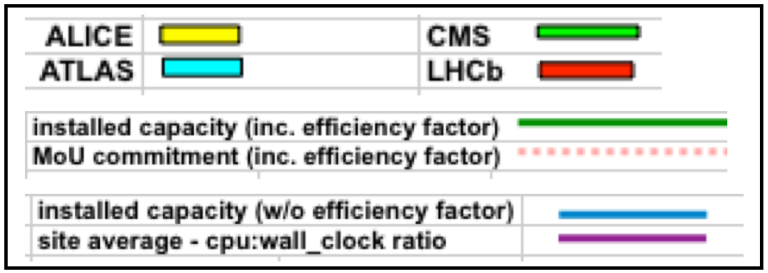
- Comments:
  - Almost all Tier 2s are now reporting correctly, except GSI who are still not reporting accounting information
- Following plots taken directly from monthly accounting reports

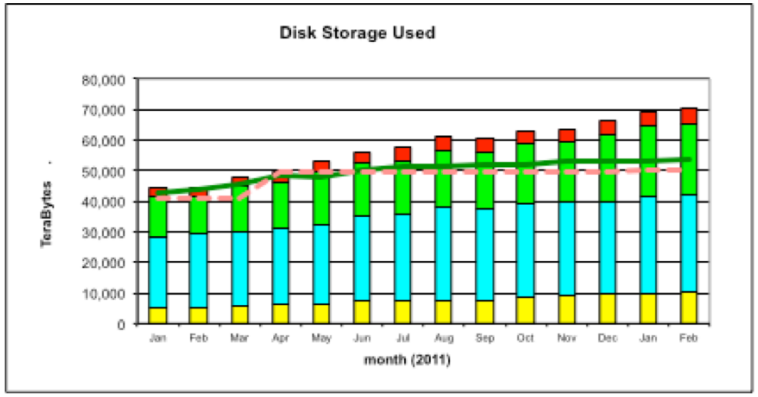
CERN & Tier 1 Accounting

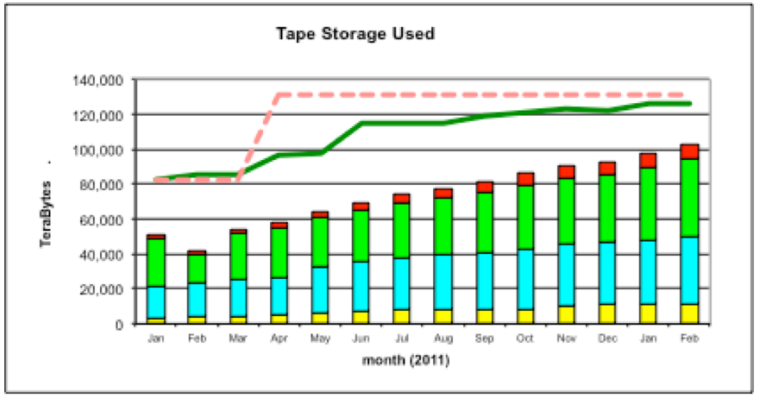












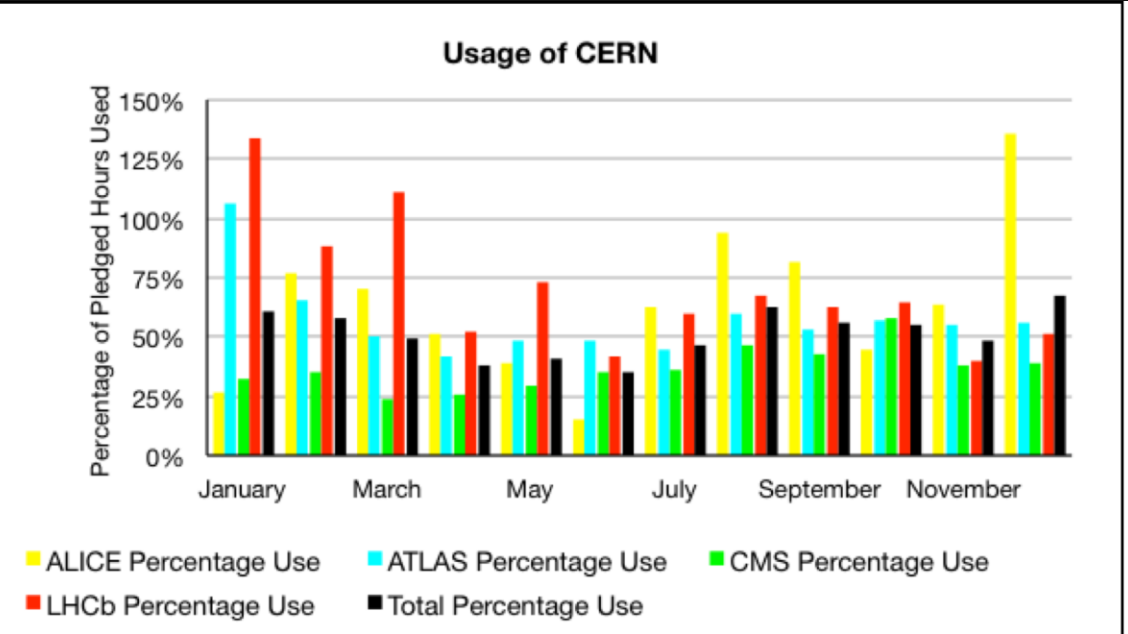
Use of T0 + T1 resources

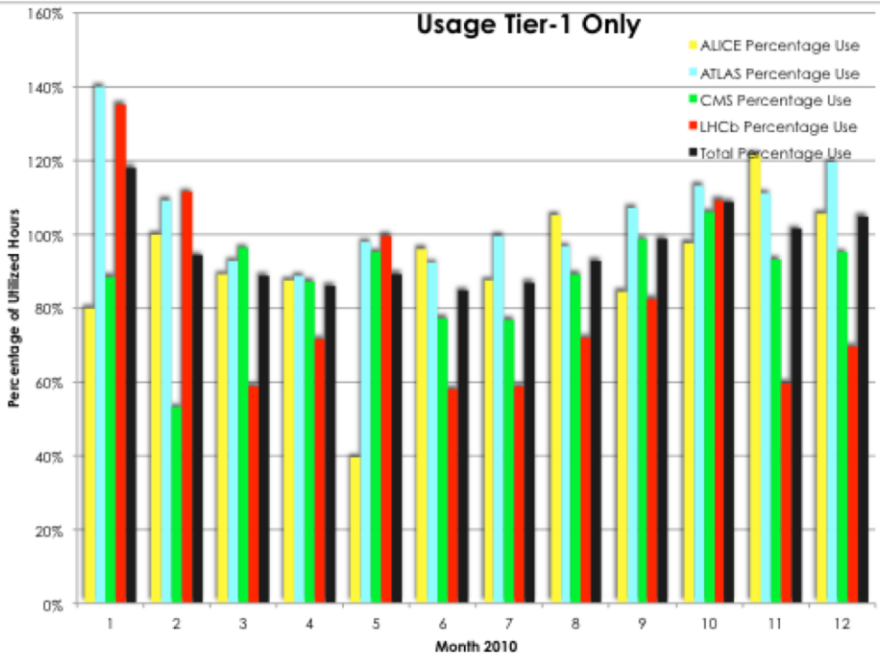
Comparison between use per experiment and pledges
- For Tier 0 alone
- For sum of Tier 1s

Early in year, pledges start to be installed – can be used

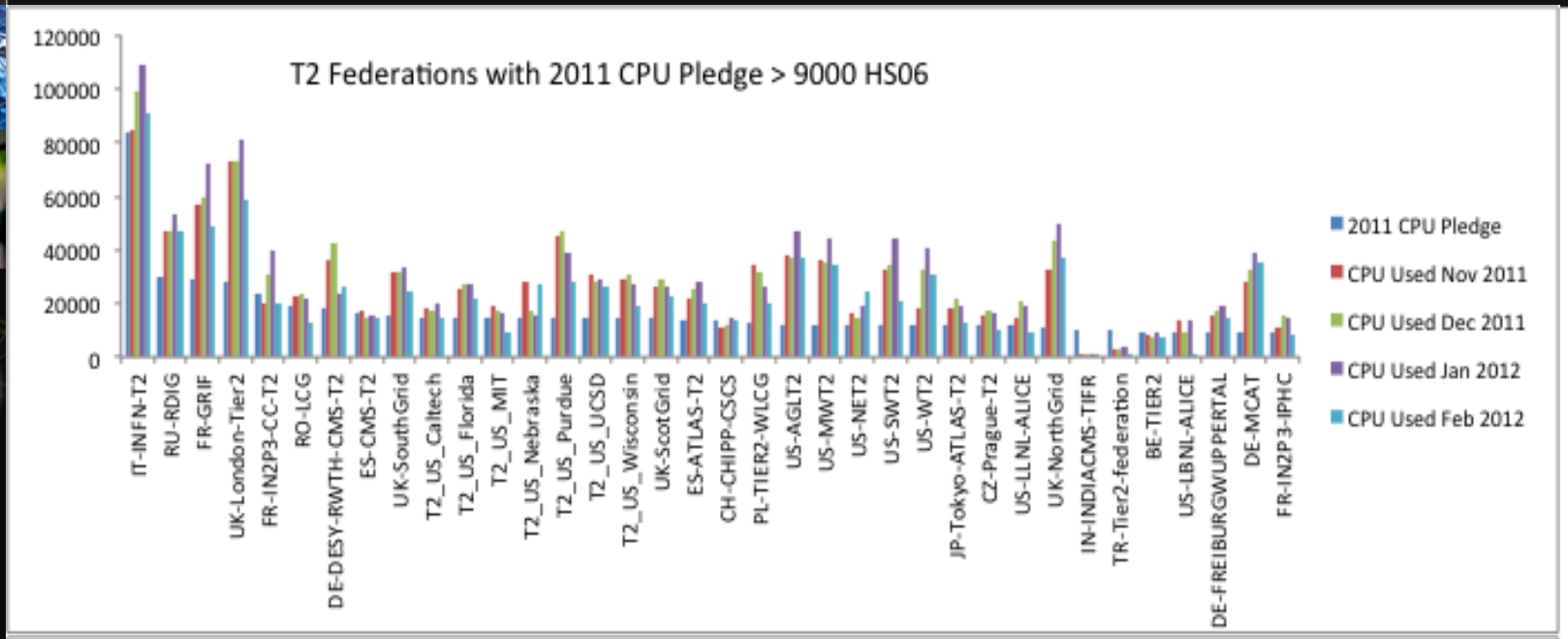
Tier 1 use – close to full

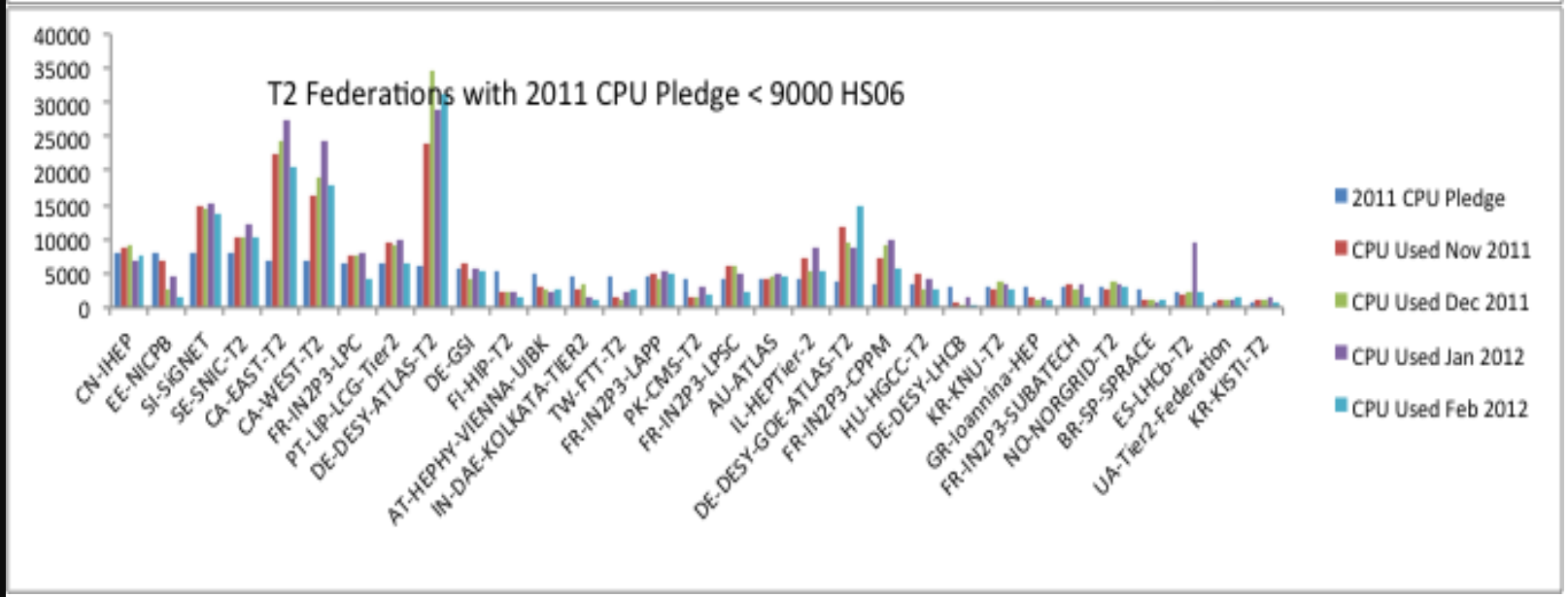
Can make use of capacity share not used, esp. ALICE & LHCb





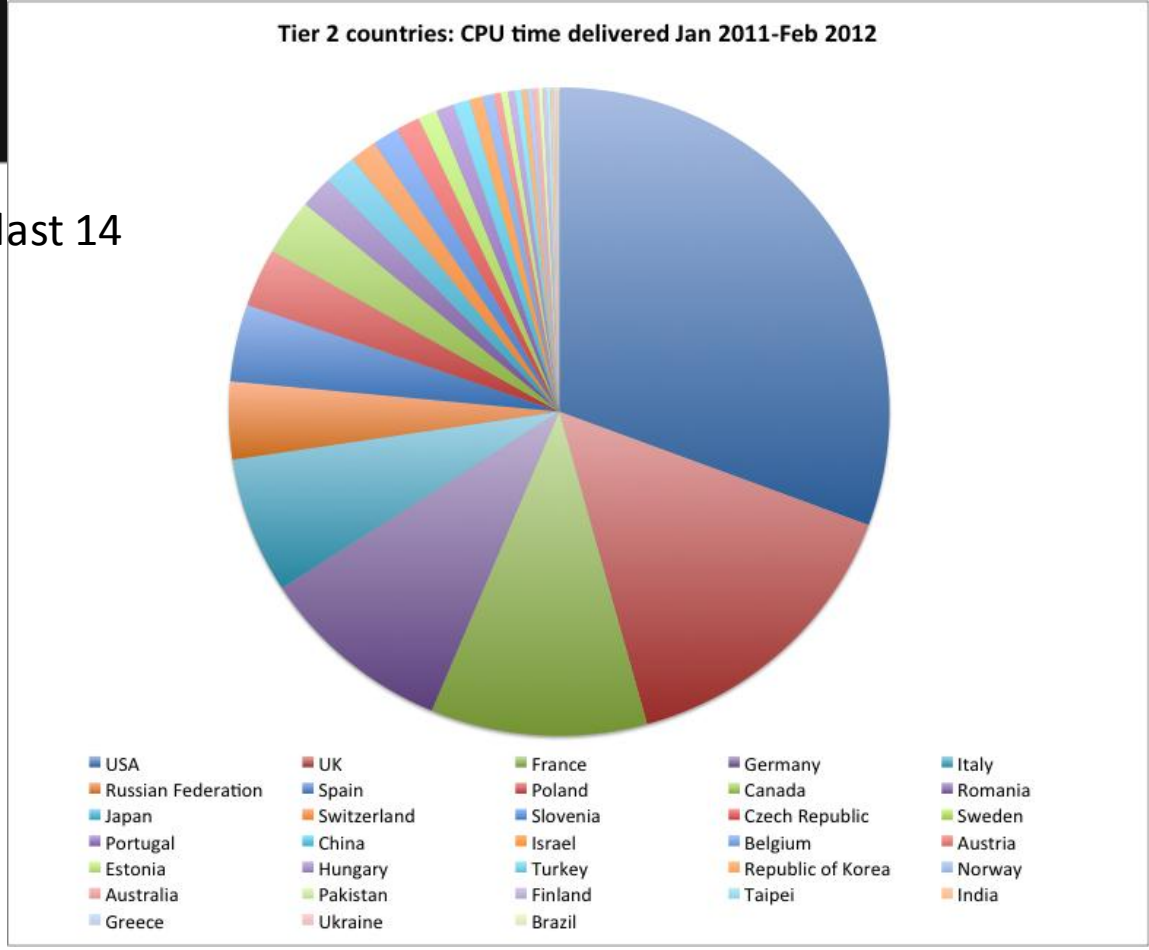
Tier 2 accounting



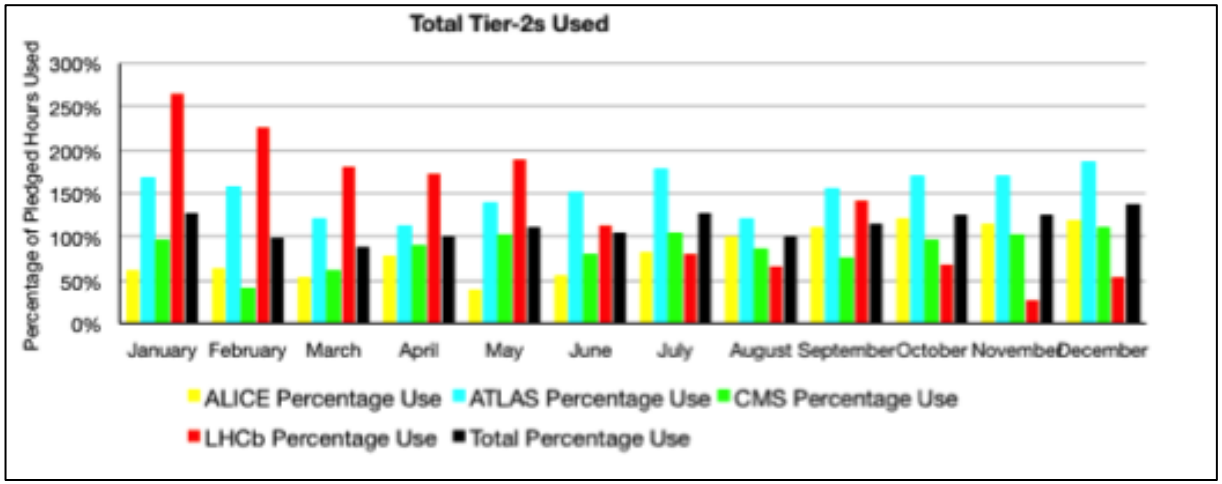


Tier 2 usage

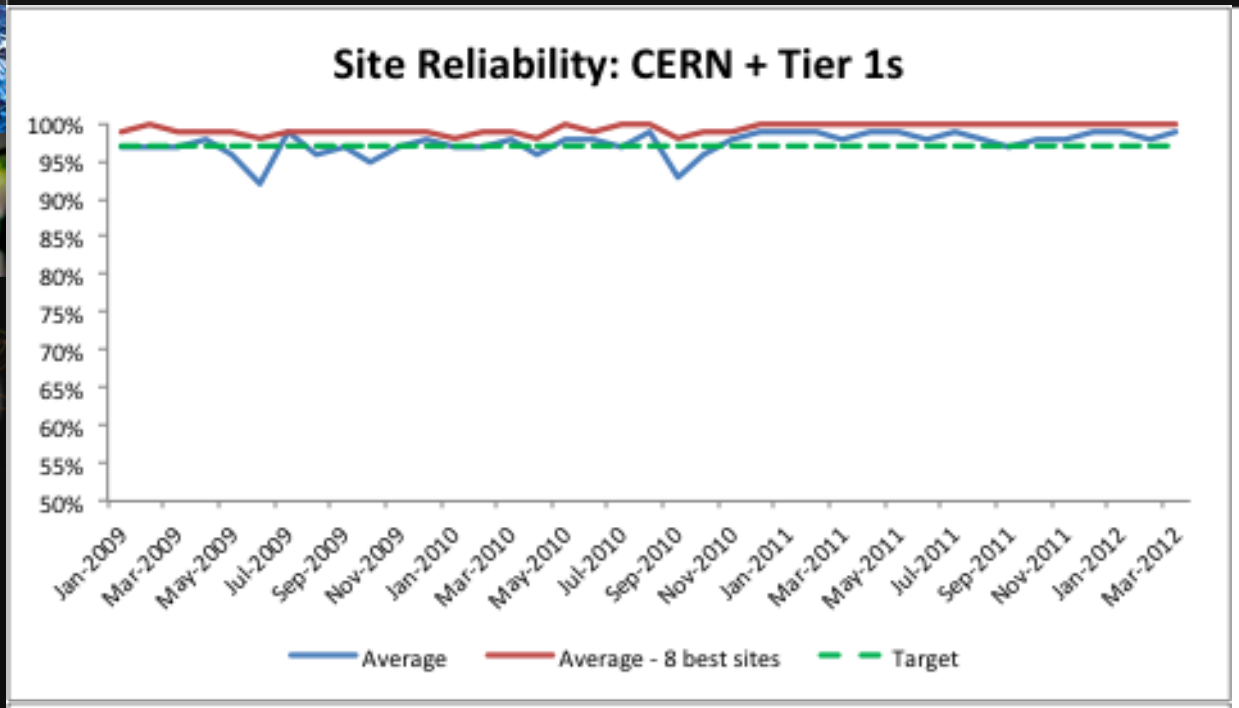
Tier 2 CPU delivered last 14 months – by country

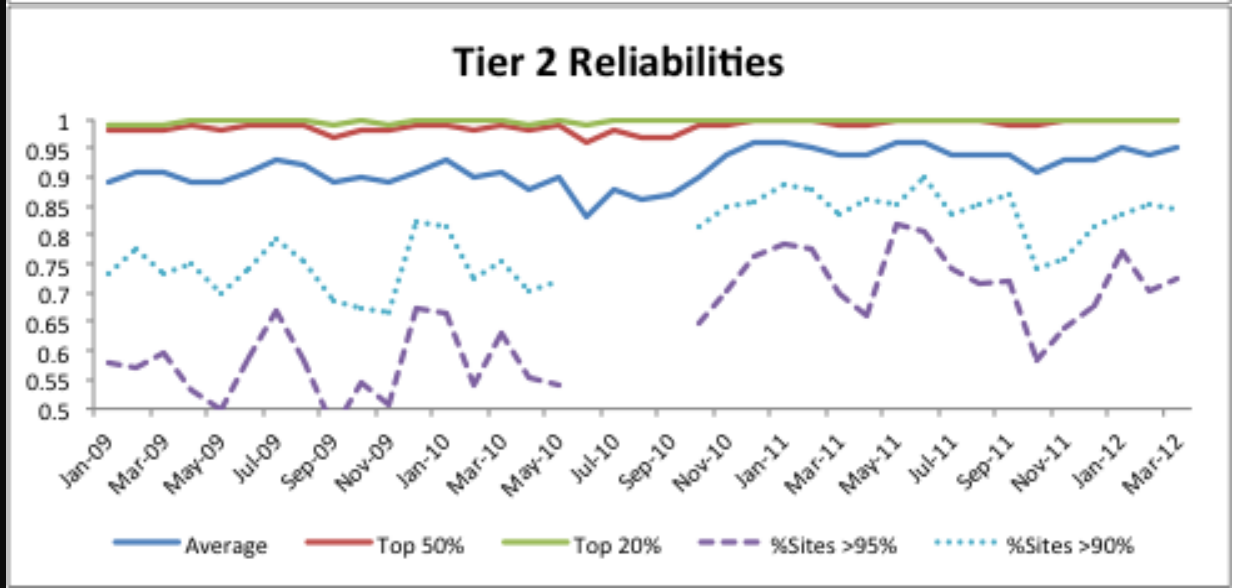


Comparison use & pledges



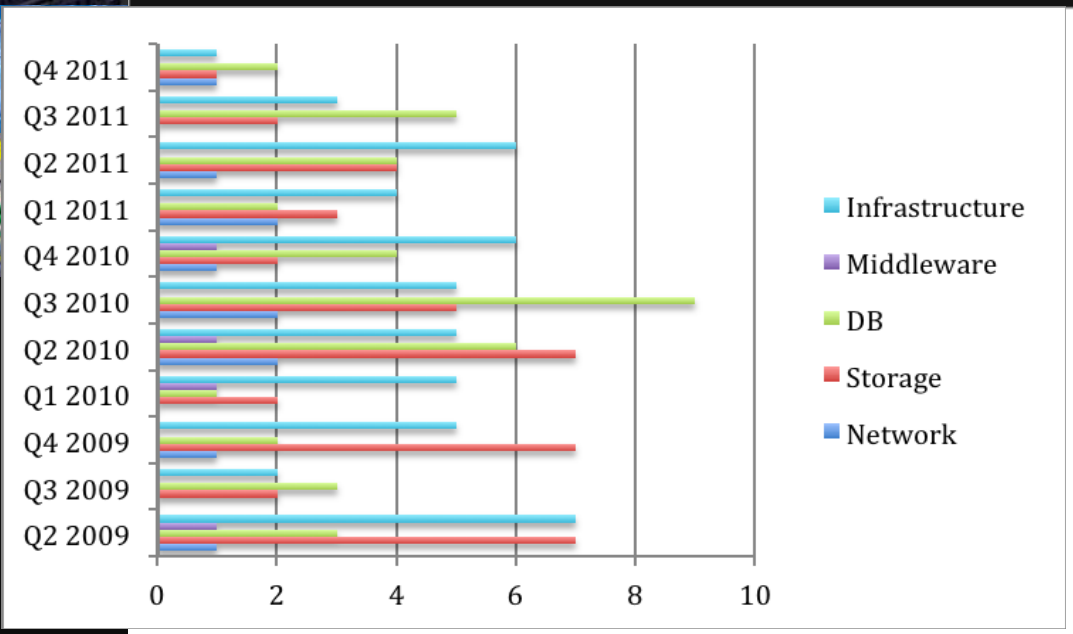
Reliabilities

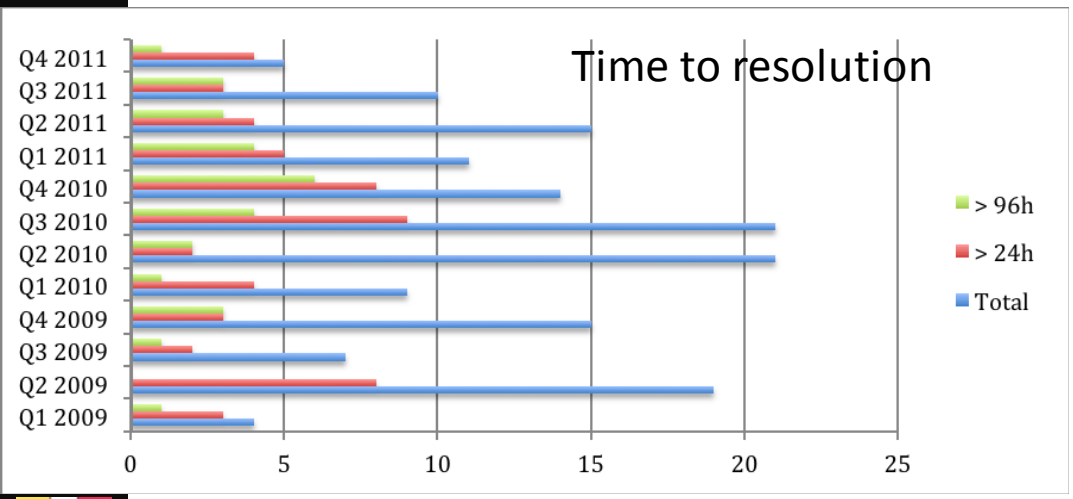




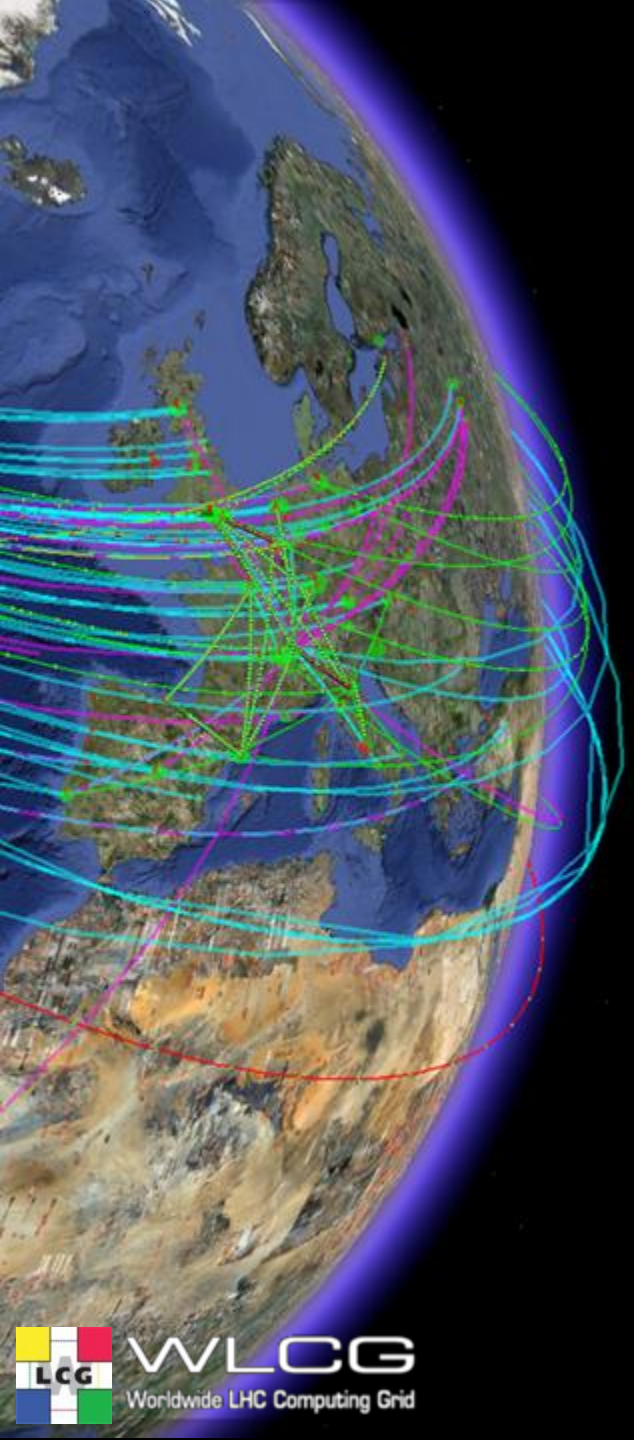
- No real issue now
- Plots show “ops” reports
- Also published monthly are experiment-specific measured reliabilities: since Feb new report allows use of “arbitrary” experiment tests

Service incidents





- Fewer incidents in general
- But longer lasting (or most difficult to resolve)
- Q4 2011 all except 1 took >24 hr to resolve



Resource pledges

Pledge vs Requirements: 2012

Tier	Pledge Type	ALICE	Required	Balance	ATLAS	Required	Balance	CMS	Required	Balance	LHCb	Required	Balance	SUM	Required	Balance
Tier 0	CPU (HEP-SPEC06)	90000	116000	-22%	111000	111000	0%	121000	120000	1%	34000	34000	0%	356000	381000	-7%
Tier 0	Disk (Tbytes)	8100	14300	-43%	9000	9000	0%	7000	7000	0%	3500	3500	0%	27600	33800	-18%
Tier 0	Tape (Tbytes)	20000	20000	0%	18000	18000	0%	23000	23000	0%	6400	6400	0%	67400	67400	0%
Tier 1	CPU (HEP-SPEC06)	94507	160000	-41%	288472	259000	11%	137085	145000	-5%	90567	113000	-20%	610631	677000	-10%
Tier 1	Disk (Tbytes)	7030	10800	-35%	30548	27000	13%	20882	22000	-5%	7360	9500	-23%	65820	69300	-5%
Tier 1	Tape (Tbytes)	11523	21000	-45%	39108	29000	35%	46531	45000	3%	5572	6200	-10%	102734	101200	2%
Tier 2	CPU (HEP-SPEC06)	128688	145000	-11%	328237	266000	23%	320373	315000	2%	47335	43000	10%	824633	769000	7%
Tier 2	Disk (Tbytes)	9109	8300	10%	45059	47000	-4%	26520	26000	2%	296	0	0%	80984	81300	0%

- Table shows situation as at last RRB
- But:
  - Some changes announced;
  - ALICE:

CPU (KHEP06)	T0	CAF	T1s	T2s
Required	90.0	35.0	95.0	207
Pledged	90		95	115 (194)
Difference	0%		0%	-80%

Disk (PB)	T0	CAF	T1s	T2s
Required	7.6	0.24	7.0	12.4
Pledged	8.1		7.22	9.11 (12.9)
Difference	6%		3%	-36%

Tape (PB)	T0	T1
Required	17.1	11.3
Pledged	20.0	11.5
Difference	14%	2%

Pledge installation status: T0/1

Site	CPU	DISK	TAPE
CERN			
TRIUMF			
IN2P3			
KIT			
INFN CNAF			
NL-T1			
NDGF			
PIC			
ASGC			
RAL			
US-BNL			
US-FNAL			

	All OK
	> 75% or slight delay
	< 75% or delay > 3 months

Pledge installation status: Tier 2

Country	CPU	DISK	Country	CPU	DISK
Canada			Turkey		
France			Hungary		
Germany			Czech Rep.		
Italy			Japan		
Nordic			Austria		
Spain			Estonia		
Taipei			Pakistan		
UK			Romania		
US-ATLAS			Ukraine		
US-CMS			Australia		
Greece			Slovenia		
Israel			China		
India				All OK	
S.Korea				> 75% or slight delay	
Switzerland				< 75% or delay > 3 months	

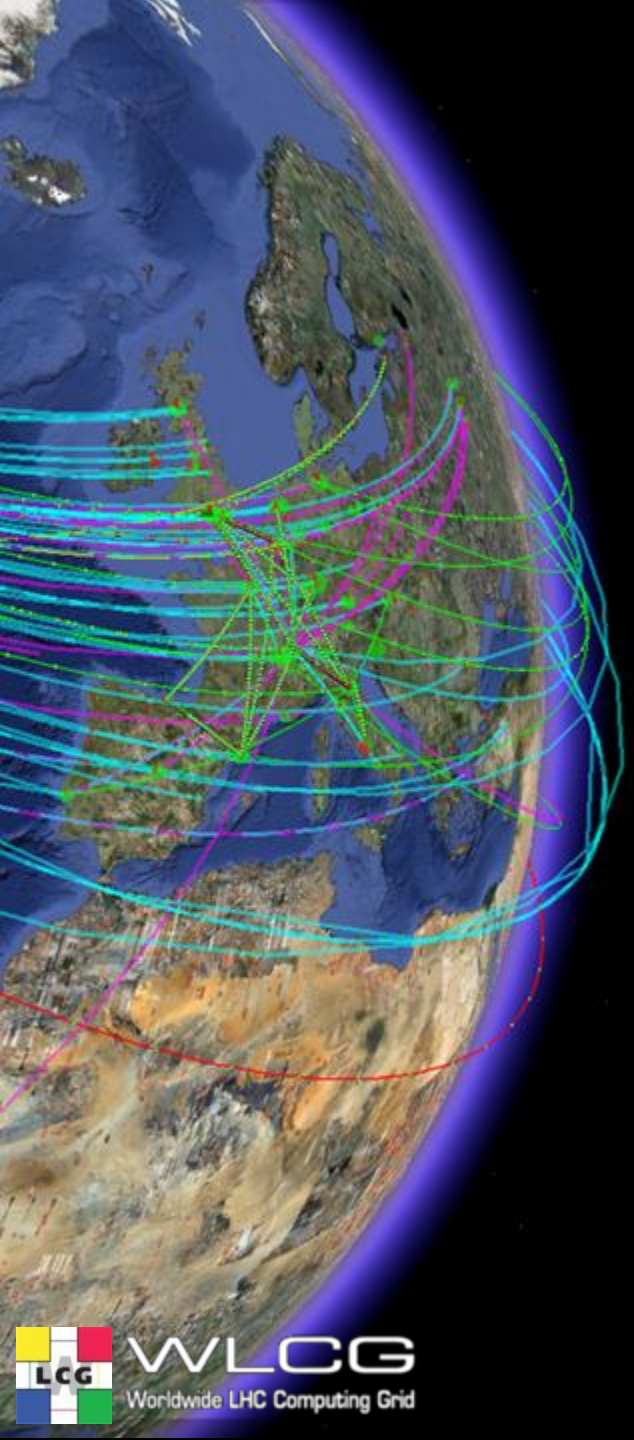
Ian.Bird@

Comments on previous RSG proposals

- Collection of installed capacity data – particularly for Tier 2s
  - Automated collection is too complex given the complex environments
  - Will use REBUS to gather this information
- The Tier 2 efficiency factor (60%→67%→70%) for CPU, has been taken into account in requirements
  - Will be updated in Accounting reports from April 2012 (accounting year boundary)
- Suggestion to disentangle "chaotic" from "organised" analysis work to determine this efficiency is not possible from the infrastructure point of view:
  - A site does not (cannot) know if a given job is "organised" or "chaotic"
  - Only the experiments themselves have this possibility

Additional data planned

- ATLAS, CMS, and LHCb all intending to take additional triggers in 2012
  - Will only be processed in 2013/14
- ATLAS:
  - Will take 400 Hz in physics streams in 2012
  - Additional 75 Hz delayed streams – to be processed in 2013: mostly for B physics
    - Additional 200 TB raw data on tape (*2 copies) + 100 TB DAODs (x2 copies)
- CMS:
  - Will take additional data and "park" it
  - Estimate +20% resources (T1) and +15% (T2) than previous estimates for 2013
- LHCb:
  - Will add +1.5 kHz Charm triggers (total 4.5 kHz); what cannot be processed in 2012 will be "locked" until resources available in 2013/14 (by re-stripping with additional channels)
- LHCC discussion generally supported these initiatives, with the proviso that priorities should be set in order according to the availability of resources



Funding & expenditure

Funding & expenditure for WLCG at CERN

- Materials planning based on current LCG resource plan
  - Provisional requirements evolve frequently
  - Currently understood accelerator schedule
  - Plan for remote Tier 0 has been evolving – now cost plan should become firmer
- Personnel – plan kept up to date with APT planning tool used for cost estimates of current contracts, planned replacements, and on-going recruitment
- Impact for 2012 & beyond:
  - Personnel: balanced situation foreseen
  - Materials: reasonably balanced given inherent uncertainties; rely on ability to carry-forward to manage delays (e.g. in CC consolidation, remote T0 costs)

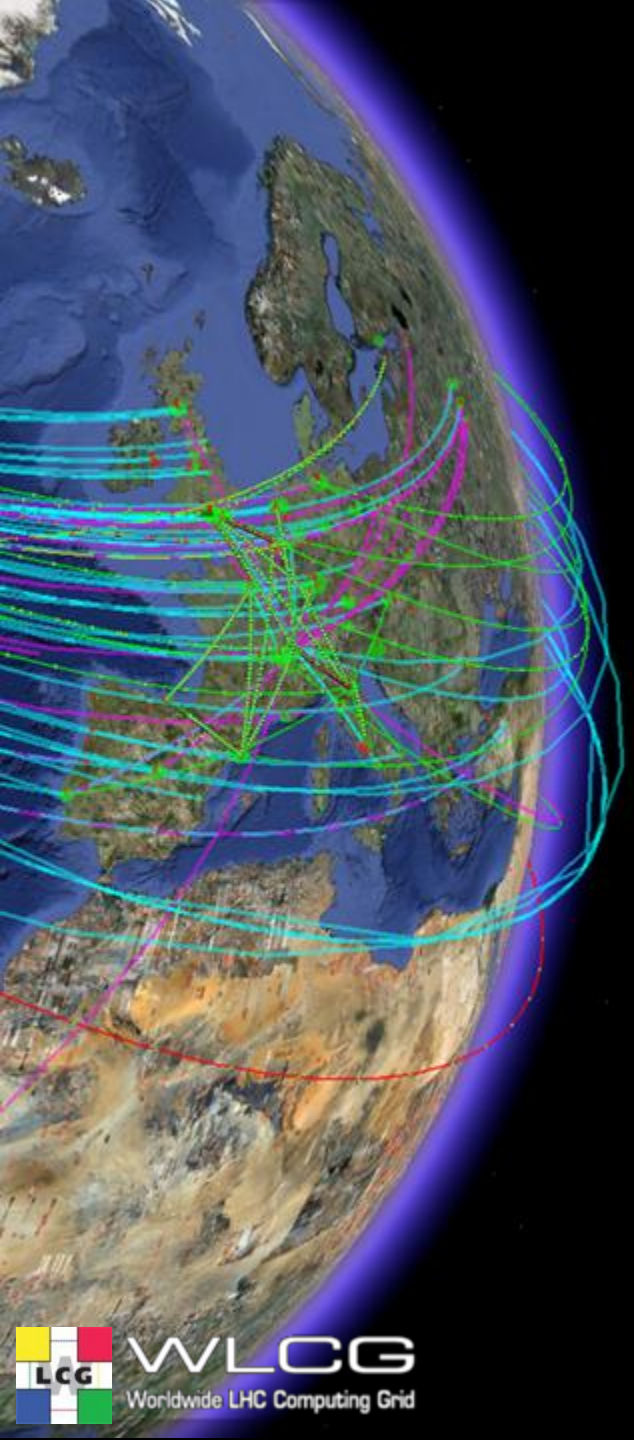
WLCG funding and expenditure

LHC Future Computing Funding and Expenditure Estimates
(all figures in MCHF)

	2012	2013	2014	2015	2016	2017	TOTAL
Funding							
From CERN Budget							
- Personnel	16.5	17.0	17.0	16.9	17.0	17.0	**101.4**
- Materials *	25.8	23.0	23.3	21.3	20.3	20.3	**134.0**
Contributions via Team Accounts**							
- Personnel	1.0	0.5					**1.5**
- Materials							
Total							
- Personnel	**17.5**	**17.5**	**17.0**	**16.9**	**17.0**	**17.0**	**102.9**
- Materials	**25.8**	**23.0**	**23.3**	**21.3**	**20.3**	**20.3**	**134.0**
Total Funding	**43.4**	**40.5**	**40.3**	**38.2**	**37.3**	**37.3**	**236.8**
Expenditure							
- Personnel ***	**17.0**	**17.7**	**17.2**	**17.1**	**17.0**	**16.7**	**102.7**
- Materials	**26.4**	**23.1**	**22.3**	**22.7**	**21.2**	**19.8**	**135.5**
Total Planned Expenditure	**43.4**	**40.8**	**39.6**	**39.8**	**38.3**	**36.5**	**238.3**
Balance Personnel	**0.5**	**-0.2**	**-0.2**	**-0.2**	**-0.1**	**0.3**	**0.1**
Balance Materials	**-0.5**	**-0.1**	**0.9**	**-1.4**	**-0.9**	**0.5**	**-1.5**
Balance	**0.0**	**-0.3**	**0.7**	**-1.6**	**-1.0**	**0.8**	**-1.4**

* Includes 4.9 MCHF carry-forward from 2011 to 2012-2015
** As planned to be pledged in the WLCG MoU (Annex 6.6)
*** Excluding EGI/EMI funded personnel and Computer Centre Operators

Planning & evolution



Tier 0 evolution

- Consolidation of existing centre at CERN
  - Project ongoing to add additional critical power in the "barn" & consolidate UPS capacity
  - Scheduled to complete Oct 2012
- Remote Tier 0
  - Tendering completed, adjudication done in March Finance Committee
  - Wigner Inst., Budapest, Hungary selected
  - Anticipate
    - Testing and first equipment installed in 2013
    - Production 2014 in time for end of LS1
  - Will be true extension of Tier 0
    - Anticipate 2*100 Gb/s connectivity, and (eventually) LHCOPN, LHCOne, IP connectivity direct from Budapest (not in 1st years)
    - Capacity to ramp up
    - Use model – as dynamic as feasible (avoid pre-allocation of experiments or types of work)

Technical Evolution

- 6 working groups set up last Autumn
- They have now reported (last week!) – reports available in WLCG Document repository
- Recommendations and proposals being digested and discussed
  - Very good opportunities for gaining commonality between experiments, & interest in doing so
  - Also between EGI, OSG, etc grid flavours
  - https://espace.cern.ch/WLCG-document-repository/Boards/MB (in Technical Evolution Strategy folder).
- Consolidated "executive summary" to be produced in near future as guidance of future priorities for work and collaborative activities
  - Goal – discussion in WLCG workshop at CHEP
- Expectation of ongoing sub-groups on specific topics, under the umbrella of the GDB

Conclusions

- Smooth WLCG operations since last RRB: HI data taking, end of year technical stop, restart of data taking
- Good use of resources at all Tiers, full pledges made use of
- Intention to take additional data in 2012 for processing in LS1
- Planning for the future ongoing and active
  - Tier 0 extension, technical work, etc.