

LCG Service Challenges - Milestones Document

Jamie Shiers, CERN-IT-GD

8 February 2005

Agenda

- **Setting the Scene**
 - An overview of recent activities
- **Draft high-level milestones**
 - Goal is to present these today...
 - Then go back and discuss schedule / detail
 - Then produce detailed MS Project plan
- **Nothing can be done without your help!**
- **Valuable input already received from Tier 1 sites**

Setting the Scene

Progress So Far...

- High Level Overview document of Service Challenge goals
 - The "Mandate" of the activity?
 - Useful as basis for initial discussions (and to boot-strap me...)
 - Re-cycled for LCG TDR ☺
- "Requirements" document, based on Summary of Computing Models and re-calculation of base network rates
 - Will summarize these in a minute...
- Global Milestones document being prepared (next)
 - To allow partners (other IT groups, equivalents at other sites, network providers, experiments) to understand what is required of them and to produce their own plans
- Once global milestones agreed, will work on detailed Project plan together with partners
 - T0/T1 sites, experiments, T2s, etc.
- Current versions of all above documents linked into today's agenda

Other Activities...

- In addition to planned GDB meetings, Service Challenge Meetings, Network Meetings etc:
- Visits to all Tier1 sites (initially)
 - Goal is to meet as many of the players as possible
 - Not just GDB representatives! Equivalents of Vlado etc.
- Current Schedule:
 - Aim to complete many of European sites by Easter
 - "Round world" trip to BNL / FNAL / Triumpf / ASCC in April
- Need to address also Tier2s
 - Cannot be done in the same way!
 - Work through existing structures, e.g.
 - HEPiX, national and regional bodies etc.
 - e.g. GridPP (12)
- Talking of SC Update at May HEPiX (FZK) with more extensive programme at Fall HEPiX (SLAC)
 - Maybe some sort of North American T2-fest around this?

More on next slide

Tier2 Plans

- SC3 should include a couple of T2s
- SC4 should complete with essentially all T2s on board
- How many? 50 - 100? [[Draft compilation](#) - Kors]
- Cannot use 'T1 model' for adding these
- Suggestion:
 - Work through bodies such as GridPP and INFN
 - Use this experience to provide guidance for adding others
 - Use HEPiX, regional / national events and workshops

➤ Do not leave until last minute!

	ALICE	ATLAS	CMS	LHCb
Parameters:				
Number of Tier-1s	4	6	6	5
Number of Tier-2s	20	24	25	15

File Edit View Favorites Tools Help

Back Forward Stop Refresh Home Search Favorites History Mail Print Edit Discuss Research Messenger

Address <http://goc.grid.sinica.edu.tw/gstat/>

Links CERN SMS Conference Rooms EDH Jamie La Clusaz Weather MeteoSuisse Geneva Weather SIMBA

Google Search Web 356 blocked AutoFill Options

GIS Monitor 15:51:12 01/25/05 GMT

[home](#) [alert](#) [table](#) [service](#) [regional](#) [rgma](#) [access](#) [service](#) [metrics](#) [misc](#) [pages](#) [help](#)

ALBERTA-LCG2 il	BEIJING-LCG2 ok	BG01-IPP ok	BHAM-LCG2 ok	BITLab-LCG ct	BNL-LCG2 ct
BUDAPEST ok	CARLETONU-LCG2 il	CAVENDISH-LCG2 il	CCIN2P3-LCG2 ct	CEA-DAPNIA-SACLAY na	CERN-LCG2 ok
CESGA-EGEE il	CGG-LCG2 ok	CIEMAT-LCG2 ok	CNAF-LCG2 wa	CNB-LCG2 il	CSCS-LCG2 na
csTCDie il	CY01-LCG2 il	CYFRONET-LCG2 ok	DESYPRO ct	Durham il	ekplcg2 ok
FNAL-LCG2 ct	FZK-LCG2 ok	GR-01-AUTH ok	GSI-LCG2 il	HEPHY-UIBK ok	HG-01-GRNET ok
HPC2N ok	HPTC-LCG2 il	HU-BERLIN na	IC-LCG2 ok	IFCA-LCG2 il	IFIC-LCG2 ok
IISAS-Bratislava il	IN2P3-CPPM na	IN2P3-LAL ok	IN2P3-LAPP na	IN2P3-LPC il	INFN-BARI na
INFN-BOLOGNA na	INFN-BOLOGNA-CMS na	INFN-CAGLIARI na	INFN-CATANIA il	INFN-CNAF il	INFN-FERRARA na
INFN-FRASCATI ok	INFN-LNL-LCG il	INFN-MILANO-LCG2 ok	INFN-NAPOLI na	INFN-NAPOLI-ATLAS il	INFN-PADOVA il
INFN-PERUGIA na	INFN-PISA na	INFN-ROMA1 ok	INFN-ROMA1-VIRGO na	INFN-ROMA2 na	INFN-TORINO-LCG2 ct
INTA-CAB ok	IPSL-IPGP-LCG2 ok	ITEP ok	JINR-LCG2 il	Lancs-LCG2 ct	LIP-LCG2 ok
LivHEP-LCG2 il	ManHEP-LCG2 ok	NCP-LCG2 il	NIKHEF_NL ok	NSC ct	OXFORD-01-LCG2 ok
PIC-LCG2 ok	POZNAN-LCG2 ct	Prague-CESNET ok	Prague-LCG2 il	QMUL-eScience ok	RAL-LCG2 ok
RALPP-LCG ok	RHUL-LCG2 il	ROGRID-ICI il	RU-Moscow-KIAM-LCG2 ct	ru-Moscow-SINP-LCG2 il	RU-Protvino-IHEP ok
ru-PSN-LCG2 il	RWTH-Aachen-Test il	SARA-LCG2 ok	SCAI ok	ScotGRID-Edinburgh ok	scotgrid-gla il
SHEFFIELD-LCG2 na	Taiwan-IPAS-LCG2 ok	Taiwan-LCG2 ok	TAU-LCG2 ok	TIFR-LCG2 na	TOKYO-LCG2 il
TORONTO-LCG2 ok	TRIUMF-GC-LCG2 il	TRIUMF-LCG2 ok	TW-NCUHEP na	UAM-LCG2 ct	UB-LCG2 ok
UCL-CCC ct	UCL-HEP na	UIOWA-LCG2 ok	Umontreal-LCG2 ok	Uni-Wuppertal ok	UPV-GRyCAP il
USC-LCG2 ok	WARSAW-LCG2 il	WEIZMANN-LCG2 il			

	sites	countries	totalCPU	freeCPU	runJob	waitJob	seAvail TB	seUsed TB	maxCPU	avgCPU
Total	111	31	9884	5074	2095	789	4784.84	1930.36	19600	9120

Code	Color	Alert Severity
0	#FFFFFF	-
10	#FAFFFA	OK
20	#EEFFFF	INFO
30	#EDEFEE	NOTE
40	#FFFCC	WARN

<http://goc.grid.sinica.edu.tw/gstat/>

Mailing Lists

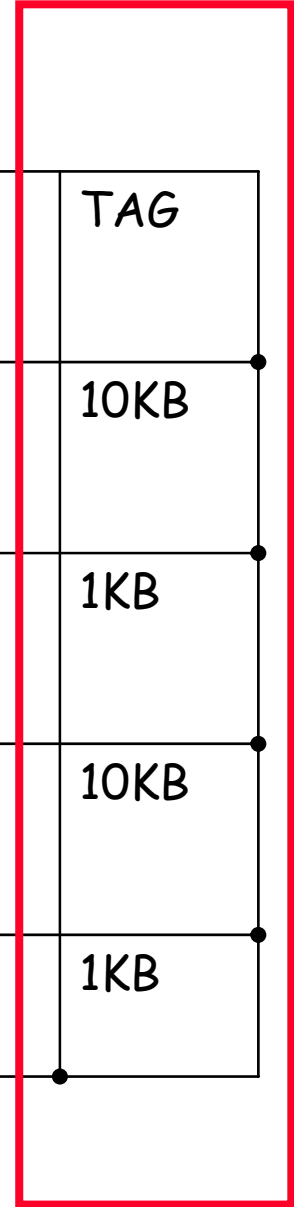
3 new mailing lists created:

- **service-challenge-man**
 - For 'management level' discussions
- **service-challenge-tech**
 - For 'technical level' discussions
- **service-challenge-cern**
 - For discussions within the CERN team

Computing Model Summaries

Overview of pp running

Experiment	SIM	SIMESD	RAW	Trigger	RECO	AOD	TAG
ALICE	400KB	40KB	1MB	100Hz	200KB	50KB	10KB
ATLAS	2MB	500KB	1.6MB	200Hz	500KB	100KB	1KB
CMS	2MB	400KB	1.5MB	150Hz	250KB	50KB	10KB
LHCb		400KB	25KB	2KHz	75KB	25KB	1KB



pp questions / uncertainties

- Trigger rates essentially independent of luminosity
 - Explicitly stated in both ATLAS and CMS CM docs
- Uncertainty (at least in my mind) on issues such as zero suppression, compaction etc of raw data sizes
 - Discussion of these factors in CMS CM doc p22:
- RAW data size ~300kB (Estimated from MC)
 - Multiplicative factors drawn from CDF experience
 - MC Underestimation factor 1.6
 - HLT Inflation of RAW Data, factor 1.25
 - Startup, thresholds, zero suppression,.... Factor 2.5
 - Real initial event size more like 1.5MB
 - Could be anywhere between 1 and 2 MB
 - Hard to deduce when the even size will fall and how that will be compensated by increasing Luminosity
- i.e. total factor = 5 for CMS raw data
- N.B. must consider not only Data Type (e.g. result of Reconstruction) but also how it is used
 - e.g. compare how Data Types are used in LHCb compared to CMS
- All this must be plugged into the meta-modell!

Overview of Heavy Ion running

Experiment	SIM	SIMESD	RAW	Trigger	RECO	AOD	TAG
ALICE	300MB	2.1MB	12.5MB	100Hz	2.5MB	250KB	10KB
ATLAS			5MB	50Hz			
CMS			7MB	50Hz	1MB	200KB	TBD
LHCb	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Heavy Ion Questions / Uncertainties

- Heavy Ion computing models less well established than for pp running
- I am concerned about model for 1st/2nd/3rd pass reconstruction and data distribution
- *"We therefore require that these data (Pb-Pb) are reconstructed at the CERN TO and exported over a four-month period after data taking. This should leave enough time for a second and third reconstruction pass at the Tier 1's" (ALICE)*
- Heavy Ion model has major impact on those Tier1's supporting these experiments
 - All bar LHCb!
- Critical to clarify these issues as soon as possible...

Nominal	These are the raw figures produced by multiplying e.g. event size x trigger rate.
Headroom	A factor of 1.5 that is applied to cater for peak rates.
Efficiency	A factor of 2 to ensure networks run at less than 50% load.
Recovery	A factor of 2 to ensure that backlogs can be cleared within 24 - 48 hours and to allow the load from a failed Tier1 to be switched over to others.
Total Requirement	<p>A factor of 6 must be applied to the nominal values to obtain the bandwidth that must be provisioned.</p> <p>Arguably this is an over-estimate, as "Recovery" and "Peak load" conditions are presumably relatively infrequent, and can also be smoothed out using appropriately sized transfer buffers.</p> <p>But as there may be under-estimates elsewhere...</p>

2007 / 8 Running (Fabiola Gianotti, CHEP)

- Software and Computing Model developed for steady-state operation
 - ≥ 2009 ?
- But : at the beginning they will be confronted with most atypical (and stressful) situations
 - cosmic and beam-halo muons used to calibrate detectors during machine commissioning
 - machine backgrounds ; higher-than-expected trigger rates
 - fast/frequent reprocessing of part of data
 - e.g. special calibration streams)
 - O(10 O(10³) physicists) physicists in panic-mode using and modifying the Software and accessing the database, GRID ...
- it is time for the Software/Computing to address the early phase of LHC operation,
- not to hinder the fast delivery of physics results (and a possible early discovery ...)

My Guess...

- Many of the detector experts will be at CERN...
- But the processing power available at CERN will not be enough
- So we will need to extensively exercise file transfer service and exploitation of processing power at all Grid sites
- Don't let's fall into the trap of 'taking it easy'...

LCG Service Challenges - Overview

- LHC will enter production (physics) in April 2007
 - Will generate an enormous volume of data
 - Will require huge amount of processing power
- LCG 'solution' is a world-wide Grid
 - Many components understood, deployed, tested..
- But...
 - Unprecedented scale
 - Humungous challenge of getting large numbers of institutes and individuals, all with existing, sometimes conflicting commitments, to work together
- LCG must be ready at full production capacity, functionality and reliability in less than 2 years from now
 - Issues include h/w acquisition, personnel hiring and training, vendor rollout schedules etc.
- Should not limit ability of physicist to exploit performance of detectors nor LHC's physics potential
 - Whilst being stable, reliable and easy to use

Key Principles

- Service challenges results in a series of services that exist in parallel with baseline production service
- Rapidly and successively approach production needs of LHC
- Initial focus: core (data management) services
- Swiftly expand out to cover full spectrum of production and analysis chain
- Must be as realistic as possible, including end-end testing of key experiment use-cases over extended periods with recovery from glitches and longer-term outages
- Necessary resources and commitment pre-requisite to success!
- Should not be under-estimated!

Timeline

- Official target date for first collisions in LHC: April 2007
- Including ski-week(s), this is only 2 years away!
- But the real target is even earlier!
- Must be ready 6 months prior to data taking
- And data taking starts earlier than colliding beams!
- Cosmics (ATLAS in a few months), calibrations, single beams, ...

Draft High-Level Milestones

- Two sorts of milestones:
 - "Generic milestones"
 - Not tied to a specific service challenge
 - SC milestones
 - Tied to a specified challenge
- Will present by 'due date'
- Regular internal CERN planning meetings started
 - e.g. Wednesday 9th February at 11:00

February Milestones

M2.01	Choice of data management components for SC2	February GDB
	The data management components and their versions that will be deployed at CERN will be defined, together with the plan for acceptance tests and service deployment. CASTOR SRM (CERN, INFN), dCache SRM (FNAL, RAL, IN2P3, FZK), SARA SRM(?) + RADIANT, no file catalog(s)	
M2.02	T0 Configuration for SC2	February
	The hardware and network configuration to be used at CERN in SC2 should be finalized, together with the schedule for putting it in place.	
M2.03	Choice of Tier 1 sites to participate in SC2	February GDB
	The sites that will participate in SC2 should confirm their commitment	
M2.04	Choice of 2 Tier 1 sites for 500MB/s for SC2	February GDB
	The sites with which 500MB/s data transfers will be attempted should be confirmed.	
M2.05	Plans for SC2	February GDB
	All Tier 1 centres involved in SC2 should present their plans for the SC2 challenge, including the foreseen data rates that they expect to be able to support. The plans should also detail the data management software components and versions that will be deployed.	
M2.06	Choice of 2 Tier 1 sites for 500MB/s for SC2	February GDB
	The sites with which 500MB/s data transfers will be attempted should be confirmed.	

March Milestones

M2.07	Acceptance tests for Data Management s/w and configuration for SC2	March GDB
	[to be defined in detail]	
M3.01	Draft list of Tier 2 centres	March GDB
	All experiments should come with a draft list of their potential Tier 2 centres with an outline of the likely network topology.	
Mg.01	Heavy ion models and data rates	March GDB
	The required data rates between T0 and T1 sites should be presented, based on revised calculations from the relevant computing models.	
Mg.02	Synchronization of service challenge and experiment milestones	March GDB
	The milestones for experiment-specific challenges should be synchronized with the global service challenge milestones so that the former build on the latter.	

April Milestones (1/2)

M2.08	SC2 complete	April SC meeting
	SC2 is complete having achieved 100MB/s T0-each participating T1, together with 500MB/s aggregate (T0) plus 500MB/s to FZK and FNAL.	
Mg.03	Plans for 10Gbit network connectivity	April SC meeting
	Each Tier1 should present its plan for obtaining 10Gbit connectivity to CERN.	
Mg.04	T1-T1 and T1-T2 data rates	April SC meeting
	The required data rates between T1 and T2 sites should be presented, based on the T2 list from the March SC meeting	
Mg.05	Draft network topology and usage schedule	April SC meeting
	A first draft of the expected network topology and usage schedule, including T0 and T1 sites should be presented,	

April Milestones (2/2)

M3.02	SC3 detailed milestones	April SC meeting
	The detailed milestones for SC3 should be presented, including the list of T1 and T2 sites that will be involved, the schedule and the choice of experiments that will initially take part.	
M3.04	SC3 plan by experiment	April SC meeting
	The key features of the experiments computing models that will be stressed during SC3 should be identified together with corresponding milestones.	
Mg.06	Plan for remaining T1 sites to join Service Challenges	April SC meeting
	The plans for the remaining T1 sites to actively participate in the Service Challenges should be presented.	
Mg.07	Initial plan for including T2 sites to SC3	April SC meeting
	An initial plan for adding T2 sites to SC3 should be presented. This should include the foreseen resources at the given T2, its primary and alternative T1, including network routing considerations, as well as contact names.	

Future Milestones

M3.05	Confirmation of sites for SC3	May SC meeting
	The list of sites that will participate in SC3 should be confirmed, together with detailed schedule and configuration plans	
		June SC meeting
		July SC meeting
		September SC meeting
		October SC meeting
		November SC meeting

Conclusions

- To be ready to fully exploit LHC, significant resources need to be allocated to a series of Service Challenges by all concerned parties
- These challenges should be seen as an essential on-going and long-term commitment to achieving production LCG
- The countdown has started - we are already in (pre-)production mode
- Next stop: 2020