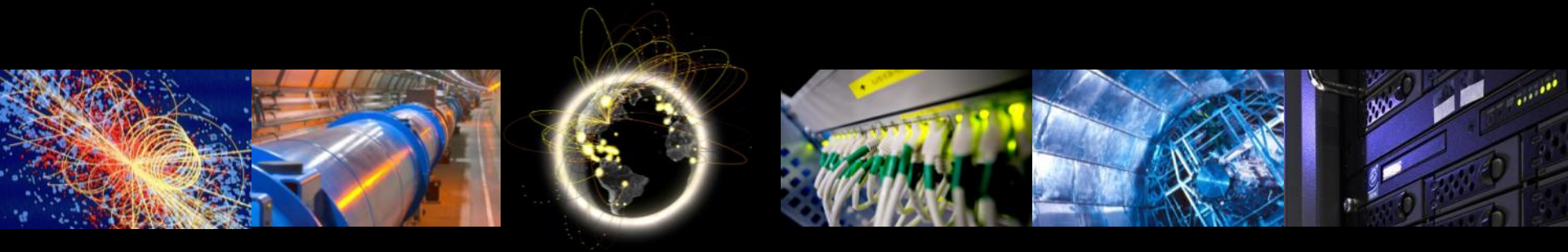# Summary of RSG/RRB

Ian Bird

GDB 9th May 2012

- Slides taken from C-RSG report to the RRB

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

2012 rates

- Real/nominal rates in 2011 and expectations for 2012:

ALICE: 380/100 Hz   (200 Hz for PbPb, expected in 2012: 400Hz, for pPb 560 Hz)
ATLAS: 340/200 Hz   (expected in 2012: 400 Hz)
CMS: 375/300 Hz   (includes 25% overlap, expected in 2012: 400 Hz and up to 600Hz)
LHCb: 3000/2000 Hz   (expected in 2012: 4500 Hz)

- Running time during 2012 expected to be very similar to 2011. No running in 2013

Pile-up is expected to increase up to 25-30 events per crossing.

Experiments plan to reprocess all data since 2010 in 2013 as well as analyze low priority streams (`parked data')

ALICE – 1

Comments on the scrutiny for the 2013 requests

ALICE

- Presents a request in line with the expected resources and describes in detail possible new contributors in the short term.

- Some unpledged resources are listed and acounted for, helping to bridge the gap with requests.

- Stays within the `natural envelope' of resources.

- Low CPU efficiency is the major concern. We ask the collaboration to use the 2013+ period to reformulate some of the computing strategies aiming to reach efficiencies comparable to the other experiments as much as possible.

- Some unknowns concerning the pPb run at the end of 2012.

ALICE – 2

ALICE

Resource	Site(s)	2013
CPU/kHS06	T0+CAF	*125*
	T1	*95*
	T2	*195*
Disk/PB	T0+CAF	*13.4*
	T1	*10.9*
	T2	*19.4*
Tape/PB	T0+CAF	*23.5*
	T1	*19.1*

ATLAS – 1

ATLAS

- Plans to record data @ 400 Hz and `park' the less relevant part for later analysis.

- Plans to make intensive use of the DAQ farm and T0 resources.

- Makes an intensive use of all resources available. They were able to make much more simulation than originally envisaged and can use MC production as a lever.

- Submitted a `revised' 2012 estimate with increased requests. Requests for 2013 are even larger and appear to the CRSG unrealistic in view of the existing spending profile and the availability of free resources in 2013+.

- 2012 will be very much similar to 2011 as data taking is concerned, except that pile-up will increase.

- Taking into account the LHCC recommendations and having the previous considerations in mind we conclude that the committed resources should match the **revised** 2012 ones.

- This is a tentative scrutiny; the final one will be provided in the October 2012 C-RRB where the present estimates can be revised if deemed necessary.

ATLAS – 2

ATLAS

CPU [kHS06]	2013 (this scrutiny)	2013 (previous estimate)
CERN	*111*	*111*
Tier-1	*297*	*273*
Tier-2	*319*	*281*
Disk [PB]		
CERN	*10*	*10*
Tier-1	*29*	*30*
Tier-2	*49*	*53*
Tape [PB]		
CERN	*19*	*18*
Tier-1	*34*	*33*

CMS – 1

CMS

- Plans to record up to 600 Hz, 400 Hz on average, and `park' the less relevant data for later analysis.

- The use of HTL farm unclear. Its use is strongly encouraged by the CRSG.

- Makes an intensive use of all resources available. They were also able to make much more simulation than originally envisaged. During 2011 experienced problems with the memory footprint that reduced their CERN usage, hopefully partly solved.

- Also submitted a `revised' 2012 estimate with increased requests. Requests for 2013 are even larger, particularly on CPU @T1 and T2. The CRSG cannot endorse this large request.

- 2012 will be very much similar to 2011 as data taking is concerned, except that pile-up will increase. Taking into account the LHCC recommendations and having the previous considerations in mind we also concluded in this case that the committed resources should match the revised 2012 ones. Some additional disk @T1 appears justified.

- This is a tentative scrutiny; the final one will be provided in the October 2012 C-RRB where the present estimates can be revised if deemed necessary.

-

CMS – 2

CMS

CPU [kHS06]	2013 (this scrutiny)	2013 (previous estimate)
CERN	*121*	*120*
Tier-1	*145*	*145*
Tier-2	*350*	*306*
Disk [PB]		
CERN	*7*	*7*
Tier-1	*26*	*27*
Tier-2	*26*	*26*
Tape [PB]		
CERN (including HI)	*23*	*23*
Tier-1	*45*	*59*

LHCb

- Plans to record data @ 4500 Hz, justified on basis of the revised charm physics program

- Plans to make intensive use of the on-line farm.

- LHCb computing is very mature but a clear underuse of their CERN usage has been observed which had to be compensated by redistribution of tasks, particularle in the T2. The model has shown good flexibility in adapting to tighter resources.

- While the total computing power is OK, the CRSG is of the opinion that some rethinking of the model may be necessary.

- The 2013 request is flat with respect to previous requests.

- A substantial amount of unpledged resources will help LHCb to fulfill their new ambitious physics program.

LHCb

Site	kHS06	Disk (PB)	Tape (PB)
CERN	*21*	*3.5*	*6.2*
Tier-1	*55*	*7.6*	*6.1*
Tier-2	*47*	*0*	*0*
Unpledged	*(54)*	*--*	*--*
Total	*123 (177)*	*11.1*	*12.3*

General recommendations – 1

Recommendations for 2013

- We recommend the use of the on-line farms during 2013 for reprocessing and simulated data production. This would entail "parking" some fraction of the data for later processing. The T0 resources should also be available in 2013. These two resources should permit an approximately flat profile for CPU requests in 2013 compared to 2012.

- We recommend to smoothen out the CPU needs throughout the year and consider the possibility of using external resources for very localized demands, particularly for MonteCarlo production.

- We remind the collaborations that, while they are welcome to write data at increased rates they cannot expect that resources automatically increase to match these rates. Therefore they should be selective in the kind of `dark' or `parked' data they plan to collect. Maintaining a reasonably flat profile is essential for the sustainability of the WLCG

- The CRSG would like to keep a balanced usage of the different Tiers. Ensuring such a balance will maintain a healthy WLCG collaboration and so ensure long term success.

General recommendations – 2

Recommendations for 2013 (continued)

- We recommend keeping the request for new disk under close scrutiny. Some collaborations have enlarged their physics scope and this may justify some increases but others have not fully justified the usage of existing disk resources yet.
If possible, the collaborations should present data access statistics to better understand and demonstrate that the data placement policies are meaningful and effective.

- The CRSG encourages close collaboration of the different centres with the experiments to continue the implementation of intelligent storage management policies to allow efficient and cost-effective access to data. In particular the implications on network bandwidth for best-use of resources should be considered. We consider this issue very relevant for the operation of the LHC experiments after 2014.

- We encourage the experimental collaborations to continue working on realistic estimates for the computing needs in 2014 and beyond, keeping the budgetary constraints in mind and working with the CRSG as necessary.

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
- Need to progress with storage accounting
- 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