CCRC'08 Planning Status

WLCG GDB, November 7th 2007

Jamie.Shiers@cern.ch

Proposed Schedule - Reminder

- Phase 1 - February 2008:
  - Possible scenario: blocks of functional tests,
    Try to reach 2008 scale for tests at…
    1. CERN: data recording, processing, CAF, data export
    2. Tier-1's: data handling (import, mass-storage, export), processing, analysis
    3. Tier-2's: Data Analysis, Monte Carlo, data import and export

➢ **More detail of these tests is needed for site planning, e.g. services & scaling factors involved in each of the above steps (see later slides)**

- Phase 2: Duration of challenge: 1 week setup, 4 weeks challenge

Ideas:

- Use **January** (pre-)GDB to review metric, tools to drive tests and monitoring tools
  - This means that we must preview the metric etc already in December meeting – more later!

➢ **Use March GDB to analysis CCRC phase 1**

➢ **Launch the challenge at the WLCG workshop (April 21-25, 2008)**

➢ **Schedule a mini-workshop after the challenge to summarize and extract lessons learned (tentatively June 12/13 in IT amphi or Council Chamber)**

➢ **Document performance and lessons learned within 4 weeks.**

CCRC'08 writeup

- Important to document the lessons learnt
- One possibility is via a paper, e.g. to Computing Physics Communications or IEEE Transactions on Nuclear Science

  *"… papers … that contain important information of **lasting value** may be submitted for review and publication in the Transactions on Nuclear Science (TNS). The TNS is a premier peer-reviewed journal with a significant distribution within the nuclear science and medical imaging communities."*

- All members of wlcg-ccrc08@cern.ch are co-authors!

CCRC08 Proposed Organization

Coordination: (1+4+nT1)

- WLCG overall coordination (1)
  - Maintains overall schedule
  - Coordinate the definition of goals and metrics
  - Coordinates regular preparation meetings
  - During the CCRC'08 coordinates operations meetings with experiments and sites
  - Coordinates the overall success evaluation

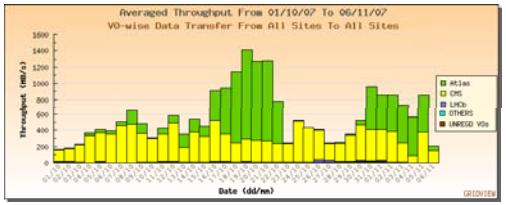
- Each Experiment: (4)
  - Coordinates the definition of the experiments goals and metrics
  - Coordinates experiments preparations
    - Applications for load driving
      (Certified and tested before the challenge)
  - During the CCRC'08 coordinates the experiments operations
  - Coordinates the experiments success evaluation

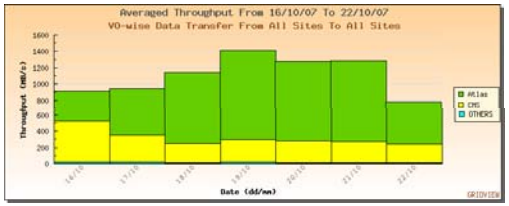
- Each Tier1 (nT1)
  - Coordinates the Tier1 preparation and the participation
  - Ensures the readiness of the center at the defined scale and schedule
  - Contributes to summary document

➢ **Tier2 coordinators have also been defined** New!

Possible Short-Term Schedule (F2Fs)

✓ Oct 9: CCRC'08 kick-off





[Excellent Tier0-Tier1 transfers from both ATLAS&CMS]

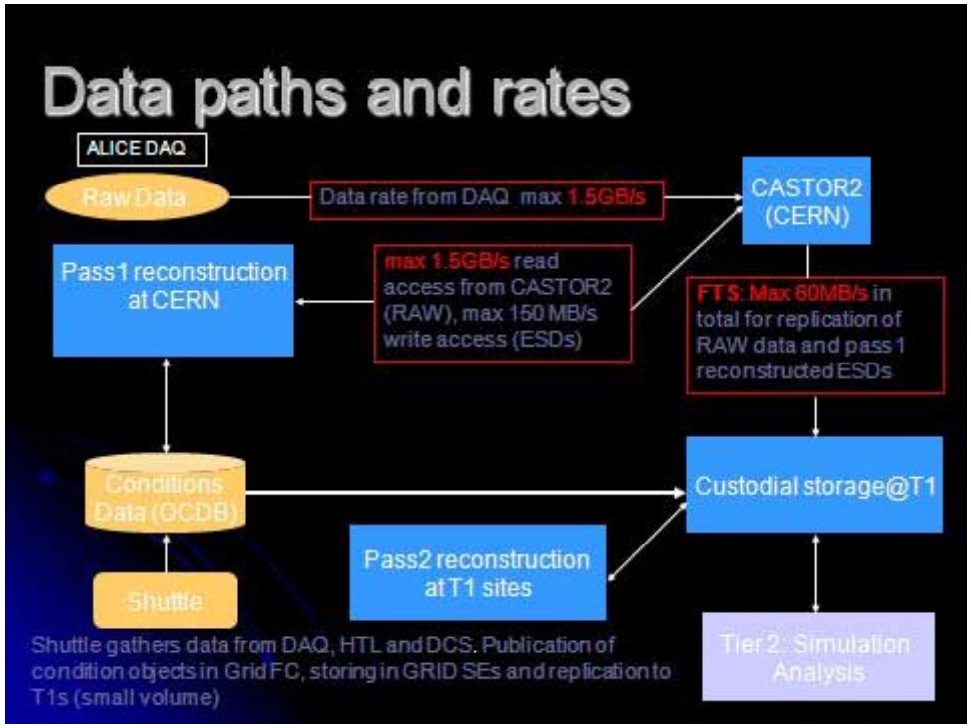
➢ Nov 6: agreement on key services & goals – including with sites; draft schedule for component testing; check-point on Explicit Requirements (ERs)

- Dec 4: progress with component testing; plans for integration testing; remaining ERs; status of site services
  - Not clear that a ½ day F2F will be enough – there are quite a lot of details that need to be worked through this year (in time for February challenge) – the same will be true for the May challenge

New!

- Jan 8: review metric, tools to drive tests and monitoring tools; progress with integration

💣 **Feb 12: mid-challenge! (2 more F2Fs!)**



ATLAS numbers needed asap!

(not just transfer rates)

Basic Scaling Items to Check in CSA08

Service	CSA08 Goal	CSA07 Goal	CSA06 Goal	Status 2006
Tier-0 Reco Rate (Hz)	150 - 300	100Hz	50Hz	Achieved
Network Transfers between T0-T1	600MB/s	300MB/s	150MB/s	Achieved All (6/7 continuous)
Network Transfers between T1-T2	50-500 MB/s	20-200 MB/s	10-100 MB/s	Achieved (15 sites)
Network Transfers T1-T1	100MB/s	50MB/s	NA	NA
Job Submission to Tier-1s	50k jobs/d	25k jobs/d	12k jobs/d	3k jobs/d
Job Submissions to Tier-2s	150k jobs/d	75k jobs/d	48k jobs/d	Achieved
MC Simulation	1.5 10^9 /year =100M /month	50M per month	NA	Not Attempted

October 9, 2007 MKasemann CCRC 07planning meeting 1/10

"Scaling" table

Service	Goal
CERN+T1 recons rate	(see job submission)
T0-T1 rate	35+6x1 MB/s
T1-T0 rate	6 MB/s
T1-T1 rate	9 MB/s per typical T1
Job submission to CERN	0.3k jobs/day
Job submission to Tier-1s	1.7k jobs/day
Analysis job to CERN/T1 (May only)	0.1-0.5k jobs/day

All production jobs ~24 hours in duration
These rates are same for May but sustained for longer duration

Some questions… (Tier0)

- How much tape space is required for the test in February ?
- How much of the space will be kept after the test (i.e. can we delete it all) ?
- How much tape space is required for the test in April (pronounced “May”) ?
- How much of the space will be kept after the test ?

- Are there any indications of the data rate required
  - to tape (i.e. writing) ?
  - from tape (i.e. processing of data which was not cached at the disk layer) ?

Global schedule: M*, FDR & CCRC'08

- FDR must test the full ATLAS data flow system, end to end
  - SFO → Tier-0 → calib/align/recon → Tier-1s → Tier-2s → analyse
  - Stage-in (Tier-1s) → reprocess → Tier-2s → analyse
  - Simulate (Tier-2s) → Tier-1s → Tier-2s → analyse
- The SFO→Tier-0 tests interfere with cosmic data-taking
- We must decouple these tests from the global data distribution and distributed operation tests as much as possible
- CCRC'08 must test the full distributed operations at the same time for all LHC experiments
  - As requested by Tier-1 centres to check their own infrastructure
- Proposal:
  - Decouple CCRC'08 from M* and FDR
    - CCRC'08 has to have fixed timescales as many people are involved
    - CCRC'08 can use any datasets prepared for the FDR, starting from Tier-0 disks
    - CCRC'08 can then run in parallel with cosmic data-taking
      - Tier-0 possible interference and total load has to be checked
      - Cosmic data distribution can be done in parallel as data flow is irregular and on average much lower than nominal rates

LHCb CCRC'08 Aims

- Test the full chain: from DAQ to Tier-0 to Tier-1's.
- Test data transfer and data access running concurrently (current tests have tested individual components).
- Test DB services at sites: conditions DB and LFC replicas.
- Tests in May will include the analysis component.
- Test the LHCb prioritisation approach to balance production and analysis at the Tier-1 centres.
- Test sites response to "chaotic activity" going on in parallel to the scheduled production activity.

LHCb CCRC'08 Planned Tasks

- RAW data distribution from the pit to the Tier-0 centre.
  - Use of rfcp into CASTOR from pit to the T1D0 storage class.
- RAW data distribution from the Tier-0 to the Tier-1 centres.
  - Use of FTS. Storage class T1D0.
- Reconstruction of the RAW data at CERN and at the Tier-1s for the production of rDST data.
  - Use of SRM 2.2. Storage class T1D0
- Stripping of data at CERN and at T1 centres.
  - Input data: RAW and rDST on T1D0.
  - Output data: DST on T1D1
  - Use SRM 2.2.
- Distribution of DST data to all other centres
  - Use of FTS - T0D1 (except CERN T1D1)

LHCb CCRC'08 Planned Tasks (II)

- Preparation of RAW data will occur over the next few month
  - We need to merge existing MC datasets into ~2GB files
- February activities:
  - Maintain the equivalent of 2 weeks data taking.
- May activities:
  - Maintain equivalent of 1 month data taking.
  - Run fake analysis activity in parallel to production type activities using generic agents.
  - Generic agents are the LHCb baseline. It needs to be integral part of CCRC'08.

“Explicit Requirements”

- “Conventional wisdom” from the SC days tell us that services need to be in place at least 2 months prior to a given “challenge”
- **For February, that basically means now**
  - **Xmas shutdown = 2 weeks lost**
- This is also true for resources – very little time for any additional deployment < February!

Explicit Requirements

Service	Experiments	Comments
SRM v2.2	ATLAS, CMS, LHCb	Roll-out schedule defined and now in progress. Expect to be at ~all Tier1s < end 2007, ~1/2 Tier2s by end January 2008, ~all Tier2s by end March 2008.
xrootd i/f	ALICE	See draft document on support for this
R/O LFC	LHCb	Developments for R/O replicas done – to be packaged and released. **Patch to be submitted this week(?) EMT 15/10**
Generic agents (aka "pilot jobs")	LHCb	See this week's discussions at MB & GDB
Commissioned links	CMS	According to CMS definition & measurement
Conditions DB	ATLAS, LHCb	In production. Tested at CCRC'08 scale?

SRM v2.2 use by Experiments

- SRM v2.2 was explicitly listed by 3 experiments (#1) as a pre-requisite for CCRC'08
  - Implicitly by ALICE – required for Tier0-Tier1 FTS
- But CCRC'08 is ***not*** the time to adapt to SRM v2.2 interface / functionality
- This must be done beforehand!
- Proposed during this morning's GSSD session that this issue be included as part of CCRC'08 phone / F2F meetings New!

What do we need to discuss in December?

- Schedule for “explicit requirements” – delivery as production services + experiment adaptation / testing
- Look carefully at all relevant milestones
- Walk-through of “functional blocks” by experiment, highlighting services (experiment, WLCG) and “scaling factors”
- This will give us a pre-view of the “metric” for agreement (for February run) at January meeting
- Confirmation from sites of resources available for production for February → scope / scale
  - e.g. will there be enough resources deployed at all sites for full re-processing loop?
- Agreement on how we view what is going on
  - The dashboards (VO view) + monitoring / accounting (service status and resources delivered)
- Example follows…

Grid Services (1)

CMS Service Requirements

Draft March 21, 2007

Service	Activities	Ramification of service interruption	Service Level
Central Services			
Oracle DB	Used by DBS	Stops creation of new analysis and re-reconstruction request. Jobs already submitted continue	
	Frontier/Calibration	Stops loading new calibration from offline database. Calibrations in cache should be accessible. Periodic cache refresh will fail	Critical after 24 hours
	PhEDEx	Stops all transfers between sites for all CMS	Critical Service
CMS RB and BDII	Used by CRAB and ProdAgent for submission for EGEE sites	No new submissions to EGEE sites and running jobs will fail. Looking at direct submission techniques as well	
FTS at CERN	Used by CERN transfers to and from Tier-1s	Transfers from CE the Tier-0 and the headroom for recc	
	Used to send data to and from Russian Tier-2 sites	Simulation proces buffers. The comp be temporarily arc	
SRM at CERN	Used by CERN transfers to and from Tier-1s	Transfers from CE the Tier-0 and the headroom for recc	
	Used to send data to and from Russian Tier-2 sites	Simulation proces buffers. The comp be temporarily arc	

LCG Service Challenges – Deploying the Service

WLCG Tier1 Services[1]

i. acceptance of an agreed share of raw data from the Tier0 Centre, keeping up with data acquisition;
ii. acceptance of an agreed share of first-pass reconstructed data from the Tier0 Centre;
iii. acceptance of processed and simulated data from other centres of the WLCG;
iv. recording and archival storage of the accepted share of raw data (distributed back-up);
v. recording and maintenance of processed and simulated data on permanent mass storage;
vi. provision of managed disk storage providing permanent and temporary data storage for files and databases;
vii. provision of access to the stored data by other centres of the WLCG and by named AF's as defined in paragraph X of this MoU;
viii. operation of a data-intensive analysis facility;
ix. provision of other services according to agreed Experiment requirements;
x. ensure high-capacity network bandwidth and services for data exchange with the Tier0 Centre, as part of an overall plan agreed amongst the Experiments, Tier1 and Tier0 Centres;
xi. ensure network bandwidth and services for data exchange with Tier1 and Tier2 Centres, as part of an overall plan agreed amongst the Experiments, Tier1 and Tier2 Centres;
xii. administration of databases required by Experiments at Tier1 Centres.
- All storage and computational services shall be "grid enabled" according to standards agreed between the LHC Experiments and the regional centres.

[1] WLCG Memorandum of Understanding (signed by each T0/T1/T2)

Specific Challenges

- HLT (High Level Trigger)
  - Startup time for Cal/Ali < 10 seconds.
  - Simultaneous
  - Uses hierarchy of squid caches
- Tier0 (Prompt Reconstruct)
  - Startup time for conditions load < 1% of total job time.
  - Usually staggered
  - DNS Round Robin should scale to 8 squids

Parameter	HLT	Tier0
# Nodes	2000	1000
# Processes	~16k	~3k
Startup	<10 sec all clients	<100 sec per client
Client Access	Simultaneous	Staggered
Cache Load	< 1 Min	N/A
Tot Obj Size	100 MB*	150 MB*
New Objects	100% / run*	100% / run*
# Squids	1 per node	Scalable (2-8)

* Worst case scenario

“Functional blocks”

1. CERN: data recording, processing, CAF, data export
2. Tier-1’s: data handling (import, mass-storage, export), processing, analysis
3. Tier-2’s: Data Analysis, Monte Carlo, data import and export

December meeting

- It seems to me that there is enough for a full day meeting
- But experiment usage of production SRM v2.2 services will also be an important issue requiring time for discussion and agreement!
- Pre-pre-GDB? Or Pre- & Post- GDBs?
  - Post-GDB slots now taken for ATLAS T1 & LCG OPN meetings….
- Comments from experiments and sites?

Summary

- We are making progress with the definition of the CCRC'08 challenges in both February & May

💣A lot of detail is still not defined, including the schedule for testing elements of the full chain << February

- It is unclear how much progress we will make on this up to and including the December meeting
- Maybe an outcome of the February run will be a better understanding of the level of detail that is required???
- Not all of the "explicit requirements" can be fully in place for February (let alone 2 months in advance)….

☺ Its going to be fun…