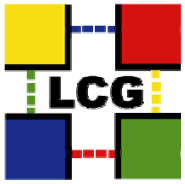
The Worldwide LHC Computing Grid

The WLCG Service Challenges

Closeout Review, September 2006

Jamie Shiers, CERN, Geneva, Switzerland



Worldwide LHC Computing Grid

Distributed Production Environment for Physics data Processing

Recommendations & Actions

- Site monitoring of local services still needs **considerable further improvement** – many issues that could be spotted locally are still first found by the central Service Coordination Team or – worse still – by the users;
- Sites are encouraged to share their monitoring tools and experience. To this end, a focussed discussion on monitoring is foreseen at the Service Challenge Technical Day, September 15th at CERN.
- **Done – much more available than previously understood: will build on this! (See also next slide)**
- **More coordination of the different groups and approaches would be valuable…**
- ¿ **Work on this over coming months → WLCG Collaboration Workshop in January?**

Statistics concerning all the transfers performed during last week

Time frame concerned: Between **2006-09-13 00:00:00 +00:00** and **2006-09-20 00:00:00 +00:00**

Channel Name	VO Name	Total	% Failures	# Succ.	# Fail.	1st Failure Reason	% 1st Failure Reason	2nd Failure Reason	% 2nd Failure Reason	Avg. Size (GB)	Avg. Duration (sec)	Avg. Tx Rate (MB/sec)	Eff. Tx Bytes (GB)	Tx Bytes (GB)
INFN-CERN	[All]	28309	94.52	1551	26758	Source SRM	88.26	Dest SRM	9.72	0.27	129.38	2.12	832.71	924.95
CERN-SARA	[All]	33820	92.22	2632	31188	Other	97.77	Source SRM	1.31	0.34	47.32	7.37	2681.42	2692.43
FNAL-CERN	[All]	4918	91.85	401	4517	Source SRM	90.95	Dest SRM	7.9	1.96	868.14	2.31	786.73	892.01
AACHEN-CERN	[All]	4820	90.93	437	4383	Source SRM	98.86	Dest SRM	0.64	1.78	289.62	6.31	779.29	784.49
SARA-CERN	[All]	27745	89.21	2993	24752	Other	59.62	Dest SRM	38.76	0.12	154.83	0.8	362.19	446.56
RAL-CERN	[All]	25846	88.99	2846	23000	Source SRM	60.36	Dest SRM	30.76	0.16	78.45	2.09	910.21	1045.27
PISA-CERN	[All]	418	78.95	88	330	Other	100			1.97	257.64	7.84	173.68	173.68
CERN-INFN	[All]	30901	77.41	6980	23921	Dest SRM	67.01	Source SRM	16.98	0.11	76.65	1.51	3139.71	4872.26
	alice	**7105**	**74.61**	**1804**	**5301**	**Source SRM**	**45.92**	**Dest SRM**	**44.95**	**0.23**	**89.13**	**2.64**	**412.71**	**418.09**
	atlas	**1023**	**94.82**	**53**	**970**	**Other**	**62.27**	**Source SRM**	**30.93**	**0.41**	**150.7**	**2.79**	**21.75**	**34.96**
	cms	**18410**	**74.33**	**4725**	**13685**	**Dest SRM**	**72.03**	**Other**	**12.42**	**0.57**	**409.3**	**1.42**	**2683.61**	**4393.81**
	lhcb	**4363**	**90.88**	**398**	**3965**	**Dest SRM**	**94.17**	**Transfer**	**3.08**	**0.05**	**94.2**	**0.59**	**21.65**	**25.39**
BNL-CERN	[All]	3400	75.62	829	2571	Dest SRM	79.19	Other	12.99	0.13	165.63	0.81	107.55	178.75
GRIDKA-CERN	[All]	13396	74.22	3454	9942	Dest SRM	68.53	Other	26.54	0.12	96	1.3	839.17	933.02
CERN-CERN	[All]	4251	73.91	1109	3142	Dest SRM	72.76	Source SRM	22.63	0.15	110.27	1.37	163.82	210.47
IN2P3-CERN	[All]	14830	69.12	4580	10250	Dest SRM	85.94	Other	12.38	0.06	87.99	0.68	534.31	656.54
STAR-CERN	[All]	29205	68.38	9236	19969	Dest SRM	73.87	Source SRM	17.04	0.19	134.55	1.45	3448.66	3895.67
TRIUMF-CERN	[All]	1531	67.28	501	1030	Dest SRM	59.13	Source SRM	35.83	0.15	178.23	0.84	73.57	113.85
UNL-CERN	[All]	1016	62.7	379	637	Transfer	96.08	Other	3.92	1.75	4698.35	0.38	0	0
CERN-TRIUMF	[All]	3395	58.56	1407	1988	Other	54.18	Dest SRM	44.87	0.49	181.29	2.75	1366.75	1567.93
CERN-IN2P3	[All]	35688	56.23	15621	20067	Dest SRM	47.91	Other	43.15	0.13	33.52	4.11	10450.14	10450.14
PIC-CERN	[All]	619	53.31	289	330	Dest SRM	70	Other	16.36	0.91	734.09	1.26	261.74	320.06
CERN-ASCC	[All]	25741	48.79	13182	12559	Dest SRM	63.06	Other	31.61	0.66	372.51	1.8	17307.79	19303.83
CERN-RAL	[All]	17841	44.39	9922	7919	Other	64.15	Source SRM	30.82	0.05	122.03	0.4	1412.93	1904.6
CERN-BNL	[All]	24136	42.69	13833	10303	Other	61.76	Source SRM	34.09	0.6	238.85	2.56	16531.26	16946.37
CERN-PIC	[All]	22091	38.24	13643	8448	Dest SRM	58.07	Other	30.02	0.15	134.54	1.15	8202.38	9126.42
KBFI-CERN	[All]	1493	36.64	946	547	Source SRM	55.94	Other	29.8	1.76	351.11	5.12	1660.55	1712.82
LNL-CERN	[All]	1249	32.11	848	401	Source SRM	86.53	Dest SRM	12.47	1.79	313.61	5.83	1514.93	1526.18
ULBVUB-CERN	[All]	1790	30.34	1247	543	Source SRM	54.33	Other	35.91	1.74	1001.78	1.78	2169.91	2276.63
DESY-CERN	[All]	330	28.79	235	95	Dest SRM	57.89	Source SRM	40	1.87	437.33	4.37	438.95	438.95

Recommendations & Actions

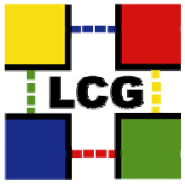
- A **regular** (**3**-4/year?) **WLCG Service Coordination meeting**, where the Tier0 and all Tier1+Tier2 "clouds" (federations) as well as the experiments are represented, should be established. This should review the services delivered by that federation, main issues encountered and plans to resolve them, possibly following the model used by GridPP for their collaboration meetings
  - See, for example Deployment Metrics and Planning, presented at GridPP16).
- It should also cover the experiments' plans for the coming quarter in more detail than can be achieved at the weekly joint operations meetings (which nevertheless could cover any updates). This meeting should not require physical presence, but would require the reports / presentations to be submitted in advance;

Recommendations & Actions

- A "Service Coordinator (On Duty – SCOD)" – a rotating, full-time activity for the length of an LHC run (but almost certainly required also outside data taking) should be established as soon as possible. The person assuming this activity would, for their period on duty:
  - Attend the daily and weekly operations meetings, relevant experiment planning and operations meetings, CASTOR deployment meetings;
  - Liaise with site and experiment contacts (**MOD, SMOD, GMOD, DBMOD**, …);
  - Maintain a daily log of on-going events, problems and their resolution;
  - Act as a single point of contact for all immediate WLCG service issues;
  - Escalate problems as appropriate to sites, experiments and / or management;
  - Write (**and present**) a detailed 'run report' at the end of the period on duty.
- It is proposed that this rota be staffed by the Tier0 and Tier1 sites, each site manning ~2 2-week periods per year (or 4 1-week periods);
- Could also be supplemented by "Yet-Another-Grid" projects (TT)

SC4 Review

Discussion of Service Levels, Intervention Times & Availability Targets



Worldwide LHC Computing Grid

Distributed Production Environment for Physics data Processing

WLCG Service Availability Targets - CERN

- Based on experience of Service Phases of SC3 & SC4, where do we stand with respect to the Service Availability targets in the MoU?

➢ Take 2 concrete examples:

1. Event reconstruction;
2. Distribution of data to Tier1s during run.

– What are the main WLCG & VO-specific services involved?

– How can targets be met? Implications?

Services Examples

1st pass processing, data export

- These two services are characterised by strong dependence on both VO and IT provided services
- Data export introduces a further coupling to storage services at Tier1 sites
- Cannot meet targets without on-call services!
- Typical interruptions:
  - 02:00 weekdays until 10:00
  - 14:00 Saturday until Monday 10:00

WLCG MoU Targets

Service	*Maximum delay in responding to operational problems*			*Average availability[1] measured on an annual basis*	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Raw data recording	4 hours	6 hours	6 hours	99%	n/a
Event reconstruction or distribution of data to Tier-1 Centres during accelerator operation	6 hours	6 hours	12 hours	99%	n/a
Networking service to Tier-1 Centres during accelerator operation	6 hours	6 hours	12 hours	99%	n/a
All other Tier-0 services	12 hours	24 hours	48 hours	98%	98%
All other services[2] – prime service hours[3]	1 hour	1 hour	4 hours	98%	98%
All other services[2] – outside prime service hours[3]	12 hours	24 hours	48 hours	97%	97%

[1] (time running)/(scheduled up-time)
[2] Services essential to the running of the Centre and to those who are using it.
[3] Prime service hours for the Host Laboratory: 08:00-18:00 in the time zone of the Host Laboratory, Monday-Friday, except public holidays and scheduled laboratory closures.

Event Reconstruction

- It is assumed that event reconstruction is performed using the local batch system, i.e. LSF
- Other services involved include the conditions database service used by the experiment in question (an Oracle-based application for all except ALICE), the experiment-specific book-keeping system(s) (typically based on Oracle and/or MySQL), the LFC (either as a file catalog or as the basis of the CMS DLS), as well as CASTOR2;
  - In the recent ATLAS Tier0 exercise, DDM/LFC operations were decoupled leaving dependencies only on CASTOR, LSF and AFS;
  - In this exercise, AFS was the primary bottleneck and cause of job failures. This is being followed up (e.g. by the use of volume replication);
  - Overall LSF performed worse than in the previous test – leading to the suggestion that a dedicated instance for first pass processing might be needed;
  - CASTOR exceeded the goal of 1 week of stable operation but with a pool 2-times over-dimensioned and Atlas wasted time trying to understand its performance;
- In summary, steps are being taken to ensure reliable services, although coupling to CASTOR, LSF and AFS (and presumably experiment-specific services) remains. All of these services are complex and problems typically require 'the expert' to be solved;

Distribution of Data (1/2)

- This activity is loosely coupled to the former, in that it requires the output of the reconstruction phase. It is, by definition, tightly coupled to the storage management services of the host laboratory (CASTOR + SRM, hence also Oracle and LSF), as well as the FTS (which also depends on Oracle), the experiment-specific framework that drives the FTS, as well as the corresponding storage management services at all of the Tier1 sites supporting a given VO;
- Except in the case of failure or severe degradation of host laboratory services, problems with a single site can, in principle, be tolerated (provided that the site in question has the proven ability to rapidly catch up with a backlog, however caused (e.g. source/sink error, or both));
- On the assumption that recovery from backlogs is demonstrated, expert coverage can probably be limited to ~12-16 hours per day. Although inter-site problems typically require dialog between experts on both sides, more than 2/3 of the data is sent to European sites, where the maximum time difference is 1 hour;
- (Sites must still respond to site-local problems as per MoU)

Distribution of Data (2/2)

- In the case of data export to the Tier1 sites, corresponding on-call services are required at the Tier1s as well, together with inter-site contacts and escalation procedures;
- We note that GGUS and COD currently provide a service during office hours (of the site in question) only, but should provide the primary problem reporting route during such periods. This requires that realistic VO-specific transfer tests are provide in the SAM (or equivalent) framework, together with the appropriate documentation and procedures;
- The list of contacts and the procedures for handling out-of-hours problems will be elaborated by the WLCG Service Coordination team and presented to the Management Board for approval. These procedures will be constructed to facilitate their eventual adoption by standard operations teams, should extended cover ever be provided. We note that such a service may address problem determination, but will not, with the current structures, provide problem resolution.
  - Detailed proposal for service monitoring enhancements also available