

Distributed Database Operations Workshop Summary

CERN, November 26th and 27th 2009

Maria Girone, CERN IT



- 45 people participated in the workshop
 - ATLAS, CMS and LHCb coordinators and developers
 - 8 sites: ASGC, BNL, IN2P3, NDGF, KIT, PIC, SARA, RAL,
 - CNAF and TRIUMF could not participate
- Very lively & interactive workshop reflecting the well-established community: WLCG and beyond
 - e.g. ESA, GSI
- Summarize main conclusions from workshop as well as achievements of “3D” in last 2 years

- Review the Experiments and Database Operations
 - Experiments DB Strategies and Requests
 - FroNTier and Coral Server Status
 - Sites Status and Plans
- DB Services Readiness
 - Alarm and Problem Escalation and Handling
 - Storage and DB parameter configuration Review
 - DB and Streams Monitoring
 - Review the Backup & Recovery Strategy/Policies
 - Demonstrate the capability to recover production DBs with a recovery validation exercise
- Beyond the 2010 run: prepare for 11gR2



Material: [EVO details](#)

[Thursday 26 November 2009](#) | [Friday 27 November 2009](#) |

Thursday 26 November 2009

[top](#)↑

09:00->13:00 Experiments and Database Operations

- 09:15 Welcome and Workshop Goals (05) Maria Girone (CERN)
- 09:20 LHC Start-up (30) Roger Bailey (CERN)
- 09:50 Experiment Database Strategy (1h00) Elizabeth Gallas (Oxford)
 - ATLAS (20') vincenzo innocente (CERN)
 - CMS (20')
 - LHCb (15')
- 10:50 Coffee break (15')
- 11:05 ATLAS Frontier/Squid Status Update (20) Rodney Walker (SFU)
- 11:25 Coral Server Status (20) Andrea Valassi (CERN)
- 11:45 CERN Database Services for Physics: Status and Plans (15) Maria Girone (CERN)
- 12:00 CASTOR+SRM DB@Tier0 Status and Plans (10) Nilo Segura Chinchilla (CERN)
- 12:15 Tier1s Status and Plans (3D and CASTOR) (1h00) Site Representatives

13:00 Lunch break (1h00)















14:00->18:15 Tier0 and Tier1 Service Review

- 14:00 ASM Configuration Review and Set-up Example (1h00) Luca Canali (CERN)
- 15:00 Recovery Exercise (30) Jacek Wojcieszuk
- 15:30 Coffee break (15')
- 15:45 Backup configuration review (30) Jacek Wojcieszuk
- 16:15 RAL data loss incident review (20) Gordon Brown (STFC-RAL)
- 16:35 ASCG Incident and Recovery report (20) Jason Shih (ASGC)
- 16:55 Analysis of the recovery exercise (1h00) Jacek Wojcieszuk

19:00 Social Dinner (2h30) (TRD)







Friday 27 November 2009

[top](#)09:00->12:45   **Database Administration, Monitoring and Configuration**

09:00	  Automatic recovery tools (20')	Przemyslaw Adam Radowiecki
09:20	  Streams Service Review (30')	Eva Dafonte Perez
09:50	  Database Monitoring (40')	Dawid Wojcik
10:30	  Streams Monitoring Enhancements (20')	Zbigniew BARANOWSKI
10:50	Coffee break (20')	
11:10	  Security Review (20')	Katarzyna Maria Dzedziniewicz (CERN)
11:30	  Memory and process limit review (Tier1s Cond. DB for ATLAS) (20')	Gancho Dimitrov (DESY)
11:50	  Hardware Migration using Data Guard (50')	Carlos Gamboa, Carmine Cioffi, Jon Westgaard

12:45 Lunch break (1h15)

14:00->17:30   **Beyond 2010 Run: Oracle 11g New Features**

14:00	  RAC and ASM (40')	Dawid Wojcik
14:40	  Active Data Guard (40')	Svetozar Kapusta (Comenius University)
15:20	Coffee break (20')	
15:40	  Advanced Compression (45')	Jacek Wojcieszuk, Luca Canali
16:25	  Streams New Features Review (35')	Zbigniew BARANOWSKI, Eva Dafonte

Databases

A number of sites, including ASGC & RAL, have been unable to recover production databases from backups / recovery areas with major downtimes occurring as a result. A coordinated DB recovery validation exercise that is regularly tested should be considered to avoid such problems.

- Recovery exercise performed and results analyzed
 - One setup at RAL, three at CERN
- Recent incidents at ASGC and RAL addressed
- Extensive review recommendations of ASM configurations and backup policies well received
- All sites managed perform a point-in-time recovery – strong agreement that sites need to regularly repeat exercise

- Experiments very satisfied with level of service and coordination for online + offline DBs;
- Standby-DBs – introduced prior to 2008 run – have provided additional redundancy and have been extremely useful for recovery (human error, failover during maintenance etc.)
- Strong interest in Coral Server – currently used by ATLAS online
- Archive DBs – introduced prior to 2009 for ATLAS and CMS for R/O applications (TAGS, conditions snap-shots)

		ASGC	CNAF	GridKa	IN2P3	SARA	BNL	RAL	PIC	TRIUMF	NDGF	CERN
FTS	RAC	Y	Y	Y	Y	Y	Y	Y	Y	Y	n/a	Y
	ASM	N	Y	Y	Y	Y	Y	Y	Y	Y	n/a	Y
	ASM: #Disk Arrays		2	2	1	1	2	2	1	1	n/a	8
	ASM: #Failgroups		1	2	1	1	1	2	1	1	n/a	8
	ASM: Redundancy		Ext	Ext	Ext	Ext	Ext	Normal	Ext	Normal	n/a	Normal
	RAID	6	10	5	5	5	10	6	6	None	n/a	None
	Flash copy	N	N	N	N	N	Y	N	Y	N	n/a	Y
	Backup to tape	N	Q2/10	N	Y	Y	Y	Y	N	N	n/a	Y
	Backup to disk		Y		N	Y	Y	Y	Y	Y	n/a	N
	Data Guard	N	N	N	N	N	N	N	N	N	n/a	Y

		ASGC	CNAF	GridKa	IN2P3	SARA	BNL	RAL	PIC	TRIUMF	NDGF	CERN
LFC	RAC	Y	Y	Y	Y	Y	Y	Y	Y	n/a	n/a	Y
	ASM	N	Y	Y	Y	Y	Y	Y	Y	n/a	n/a	Y
	ASM: #Disk Arrays		2	2	1	1	2	2	1	n/a	n/a	8
	ASM: #Failgroups		1	2	1	1	1	2	1	n/a	n/a	8
	ASM: Redundancy		Ext	Ext	Ext	Ext	Ext	Normal	Ext	n/a	n/a	Normal
	RAID	6	10	5	5	5	10	6	6	n/a	n/a	None
	Flash copy	N	N	N	N	N	Y	N	Y	n/a	n/a	Y
	Backup to tape	N	Q2/10	N	Y	Y	Y	Y	N	n/a	n/a	Y
	Backup to disk		Y		N	Y	Y	Y	Y	n/a	n/a	N
	Data Guard	N	N	N	N	N	N	N	N	n/a	n/a	Y

		ASGC	CNAF	GridKa	IN2P3	SARA	BNL	RAL	PIC	TRIUMF	NDGF	CERN
CASTOR	RAC	Y	Y	n/a	n/a	n/a	n/a	Y	n/a	n/a	n/a	Y
	ASM	Y	Y	n/a	n/a	n/a	n/a	Y	n/a	n/a	n/a	N
	ASM: #Disk Arrays	1	2	n/a	n/a	n/a	n/a	2	n/a	n/a	n/a	NAS
	ASM: #Failgroups	1	1	n/a	n/a	n/a	n/a	2	n/a	n/a	n/a	n/a
	ASM: Redundancy	Ext	Ext	n/a	n/a	n/a	n/a	Normal	n/a	n/a	n/a	n/a
	RAID	6	10	n/a	n/a	n/a	n/a	None	n/a	n/a	n/a	6
	Flash copy	N	N	n/a	n/a	n/a	n/a	N	n/a	n/a	n/a	N
	Backup to tape	N	Q2/10	n/a	n/a	n/a	n/a	Y	n/a	n/a	n/a	Y
	Backup to disk		Y	n/a	n/a	n/a	n/a	Y	n/a	n/a	n/a	N
	Data Guard	N	N	n/a	n/a	n/a	n/a	N	n/a	n/a	n/a	N



- ✓ Built community across CERN online and offline and with WLCG Tier1 sites: proven during CCRC'08, STEP'09 and recent data taking
- ✓ Fully integrated into overall WLCG operations, including WLCG workshops and daily phone calls
- ✓ Sharing of architecture, knowledge and procedures – important for minimizing manpower costs and demonstrated benefits in helping sites to recover
- ✓ Project extended to cover coordination of CASTOR & SRM DBs at Tier1s: ASGC DB now configured following recommendations from 3D + IT-DM
- ✓ Status: data taking!