Outline

- Critical service incident
  - Atlas 3D
  - Castor DB corruption
- Procedures and event logs
- Plan
Atlas 3D incident - I

- Sept 27th,
  - First observed by Eva ORA-12514 error code, and propagation cannot be re-started normally. verbose exception msg is 'TNS:listener does not currently know of service requested in connect descriptor'

Streams Monitor Error Report
Report date: 2009-09-27 02:21:49
Affected Site: CERN-PROD
Affected Database: ATLDSC.CERN.CH
Process Name: STRM_PROPAGATION_ASGC
Error Time: 27-09-2009 02:21:48
Error Message: ORA-12514: TNS:listener does not currently know of service requested in connect descriptor
ORA-06512: at "SYS.DBMS_AQADM_SYS", line 1087
ORA-06512: at "SYS.DBMS_AQADM_SYS", line 7642
ORA-06512: at "SYS.DBMS_AQADM", line 631
ORA-06512: at line 1
Atlas 3D incident - II

• 28-Sep-09 07:17 - CERN-PROD : Process error report
  STRM_PROPAGATION_ASGC@ATLDSC.CERN.CH, service degradation of listener is suspicious.

• 28-Sep-09 13:40 - Eva confirm the same error code "ORA-12528: TNS: listener: all appropriate instances are blocking new connections"
  • Instance down, err code 'ERROR at line 1: ORA-01507: database not mounted'
  • 28-Sep-09 14:05 - all lsnr service are blocked, and confirm the db is not open properly except with the fs mounted.
Atlas 3D incident - events

- 28-Sep-09 - the problem seems affect by the recent power surge and cause hardware degradation of two blade chassis. Consider performing point-in-time restore for the 3D database
  - propagation is disable now and we confirm LCRs activities already via stream monitor page
- 29-Sep-09 14:05, problem persist after point-in-time recovery. In fact, it is not possible to access to some old data so this makes impossible to try any kind of resynchronization using Streams with the current database status.
- 30-Sep-09 - perform point-in-time restore to Sept 21, where we confirm to have full backup
  - 30-Sep-09 13:41 - Eva confirm with some errors when try to access the data and this is why we suspect there is data corrupted (12 entries being selected with owner like 'ATLAS_COOLEFL_DCS' and object_name like '%_F0028_%', and the missing object observed: strmadmin@ASGC3D> select * from ATLAS_COOLEFL_DCS.COMP200_F0028_IOVS_SEQ;
    - ERROR at line 1:
    - ORA-08103: object no longer exists
- 5-Oct-09 13:41, the datafiles are nolonger valid (the point-in-time restored refer to Sept 21 while streams might not be able to recover from a 2 weeks backlog or not having archived log files on disk
Atlas 3D incident - recovery

RMAN> run {
SET UNTIL SCN 6077727611290;
#SET UNTIL TIME "TO DATE('Sep 21 2009 18:00:00', 'Mon DD YYYY HH24:MI:SS')";
restore database;
recover database;
}

RMAN-03002: failure of restore command at 09/29/2009 19:21:45
RMAN-20207: UNTIL TIME or RECOVERY WINDOW is before RESETLOGS time

channel ORA_DISK_1: restored backup piece 1
piece handle=+ATLAS_BACKUP/asgc3d/datafiles/uakpn18k_1_1
    tag=TAG20090920T090003
channel ORA_DISK_1: restore complete, elapsed time: 00:00:46

Tue Sep 29 19:57:47 2009
alter database recover logfile
'+ATLAS_DATA/asgc3d/archivelog/2009_09_28/thread_2_seq_4571.879.698789657'
Tue Sep 29 19:57:47 2009
Media Recovery Log
+ATLAS_DATA/asgc3d/archivelog/2009_09_28/thread_2_seq_4571.879.698789657
Tue Sep 29 19:57:48 2009
Errors in file /u01/app/oracle/admin/asgc3d/bdump/asgc3d1_p000_11562.trc:
ORA-00600: internal error code, arguments: [4552], [2], [0], [], [], [], [], [], []
[, []
Tue Sep 29 19:57:49 2009
Errors in file /u01/app/oracle/admin/asgc3d/bdump/asgc3d1_p000_11562.trc:
ORA-00308: cannot open archived log
'+ATLAS_DATA/asgc3d/online/log/group_15.291.677251137'
Castor DB incident

- DB lock critical alarm
  - Date: Wed Oct 21 00:16:52 UTC 2009
  - Error: OracleLocks CRITICAL - CASTORDB2: 10 are waiting for lock out of the 146 sessions are connected
    - url: https://nagios.grid.sinica.edu.tw/asgc_ops/cgi-bin/status.cgi?host=w-rac02&servicestatustypes=28
Castor DB incident - analysis

• Alert log in 2nd instance indicating the block corruption:

```
# tail /u01/app/oracle/admin/castordb/bdump/alert_castordb2.log
Wed Oct 21 10:00:38 2009
Recovery of Online Redo Log: Thread 2 Group 21 Seq 53451 Reading mem 0
   Mem# 0: /u02/oradata/castordb/redo2101.log
Block recovery completed at rba 53451.178442.16, scn 2.3721027505
Wed Oct 21 10:00:39 2009
Corrupt Block Found
   TSN = 10, TSNAME = STAGER_DATA
   RFN = 10, BLK = 1967881, RDBA = 43910921
   OBJN = 84545, OBJD = 84545, OBJECT = SUBREQUEST, SUBOBJECT =
   P_STATUS_6
   SEGMENT OWNER = STAGER, SEGMENT TYPE = Table Partition
```
**Castor DB incident – recovery**

- **Oct 29th – Nov 3rd,**
  - Hardware preparation
    - New blade servers, pass-through parts, fabric switch
    - Backend storage
    - 5 nodes in total, split into two clusters
  - Clean cluster preparation
    - collect any additional patches suggested by CASTOR DB team at CERN
  - CERN lead
    - install additional patches and the current security patch
    - start DB and check proper server functioning
    - initiate cloning of binaries to second ASGC cluster (SRM, Stager, DLF)
    - import recovered name server content into ASGC cluster 1
    - bring up cluster 2 and import the other recovered schemata
Castor DB incident – service restart

- **Nov 4th,**
  - Recreate dlf schema which is less critical
  - confirm all core services up and running
  - Upgrading SRM to 2.8
    - SRM service resume Nov 5th
- **Nov 6th,**
  - OP revert the wrong srm config and cause the service degraded start from Fri.
  - Situation clarified:
    - Stager/VO mapping
    - Incorrect CNS host definition
    - Full function since Mon.
      - Exp. confirm with FT and production transfers
Plan

- Weekly Tape backup
- Spare hardware for standby cluster
- Enforcement of recovery exercises
Acknowledgment

- Atlas 3D resync
  - Carlos from BNL
  - Eva/Maria from CERN
- CastorDB recovery
  - Eric, Luca, Jacek, Dawid, Nilo, Giuseppe, Maria and Dirk.
  - Castor support team