Oracle technical update for CERN physics services

CERN openlab II quarterly review
29 January 2008

Dawid Wojcik

- **Streams Service Review**
- **Streams Environment Updates**
- **New Streams Features in Oracle 11g**
- **Service expansion – Q1 2008**
- **32-bit Linux to 64-bit migration**
- **New HA setup**
- **Joint Software Testing Programme**
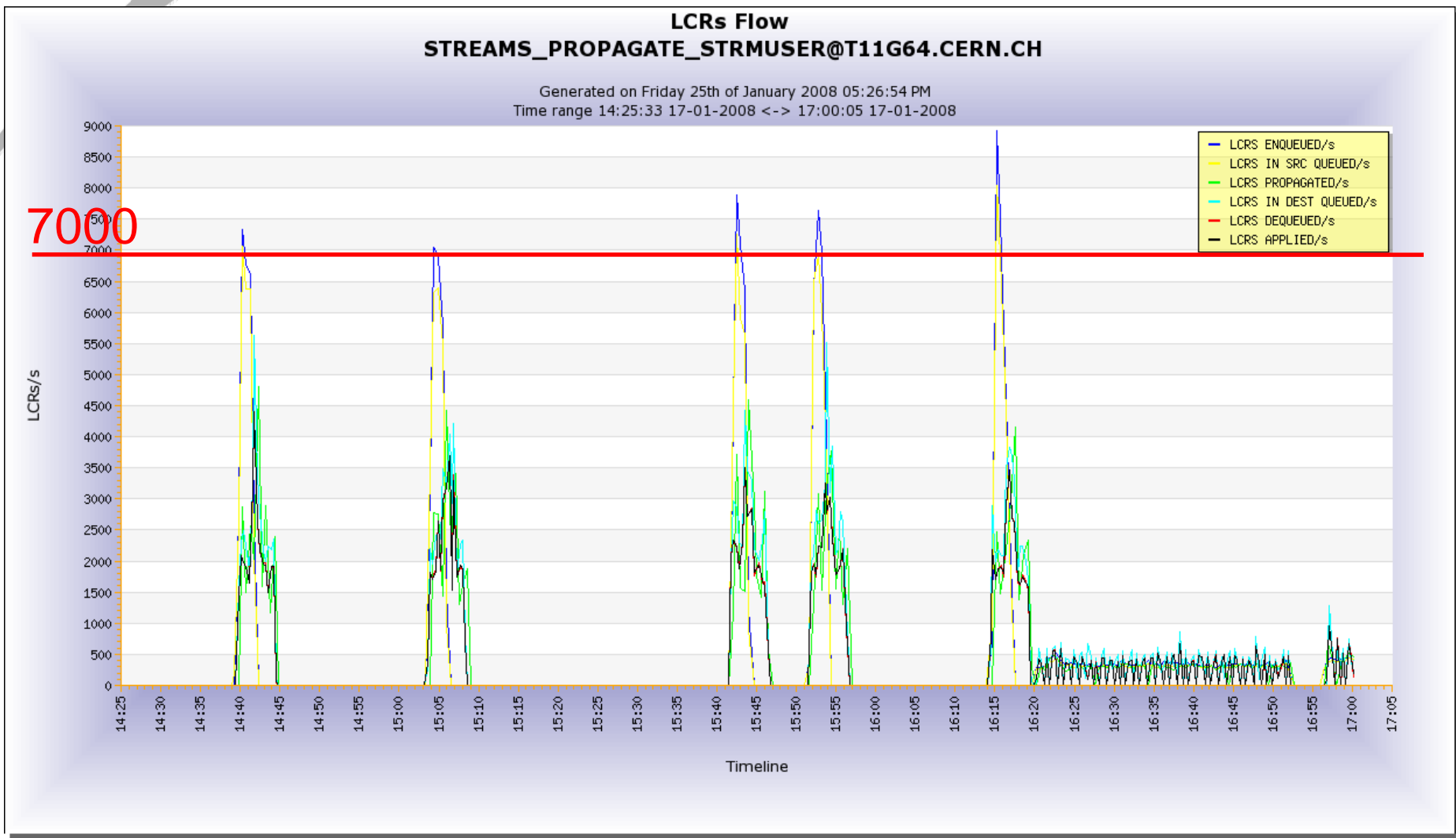- **Other Activities**
- **Future Work**

- WLCG Service Reliability Workshop
- Review of known issues related to Steams
  - Bugs reported to Oracle fixed or being fixed
    - CERN openlab has set an excellent ground!
  - Workarounds implemented
- Review of the interventions
  - Working hours coverage
    - timezone may delay resolution
    - not ready for 24/7 service
- Streams support procedures included in the Oracle Tier0 Physics Database team's operations
- Documentation (TWiki page) with known problems and solutions prepared for Tier1 administrators
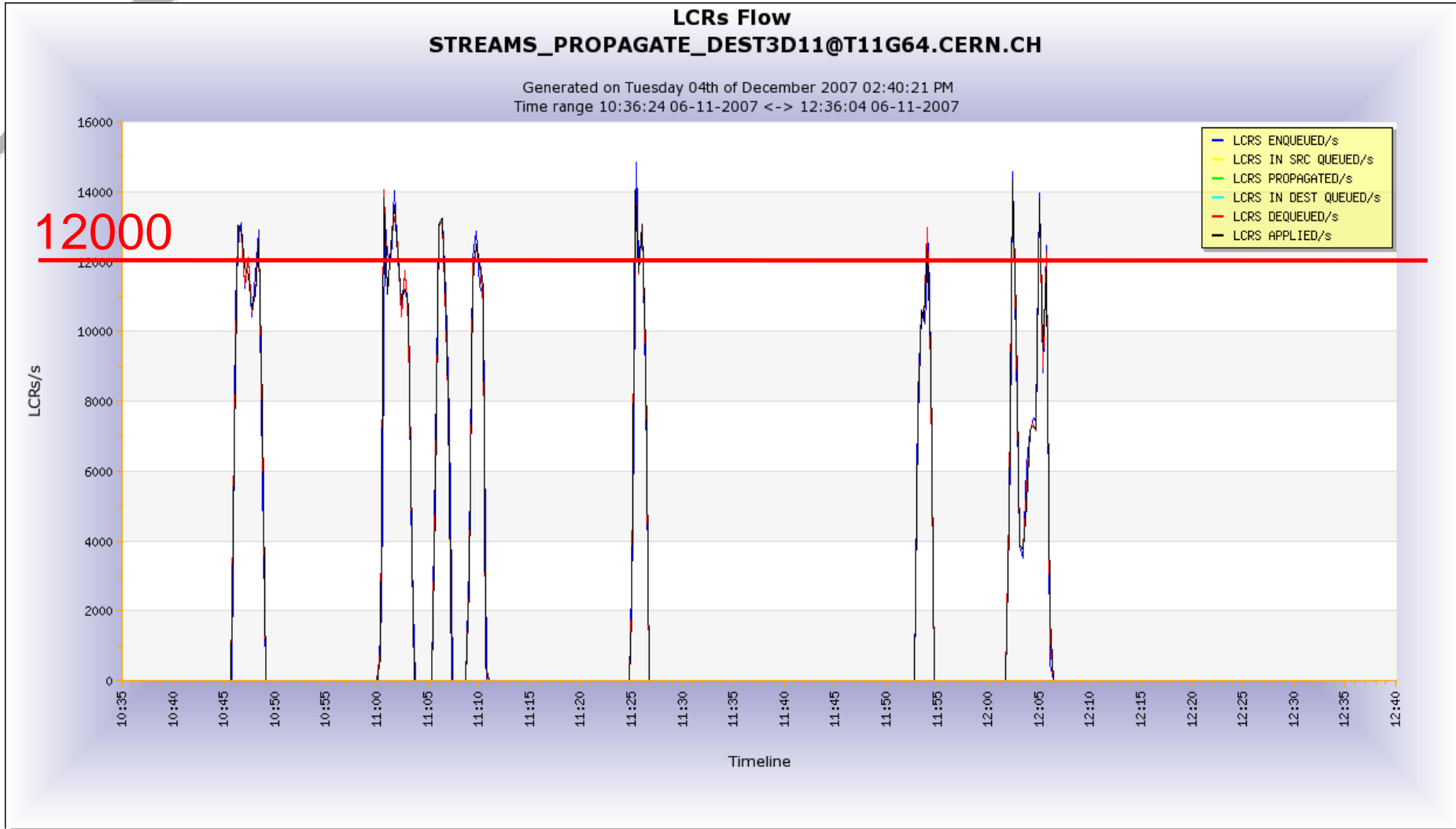
Streams Environment Updates

- ## ATLAS
  - ### Patch applied due to a problem found using Streams and compressed tables (PVSS)
- ## LHCb / LFC
  - ### LFC software upgrade
    - new tables consistently added to the replication
  - ### Hardware migration @CNAF
    - Streams reconfiguration
  - ### New test environment for VOMS

New 11g Streams Features

- **Performance improvements**
  - minimizes disk I/O and reduce capture latency
  - reduces CPU consumption

- **Combined Capture and Apply**
  - condition – only 1 propagation job running
  - capture sends LCRs directly to apply
  - configuration detected automatically
  - source and target 11g databases
  - big performance improvement
    - rate: 14.000 LCRs/sec (before 5.000 LCRs/sec)
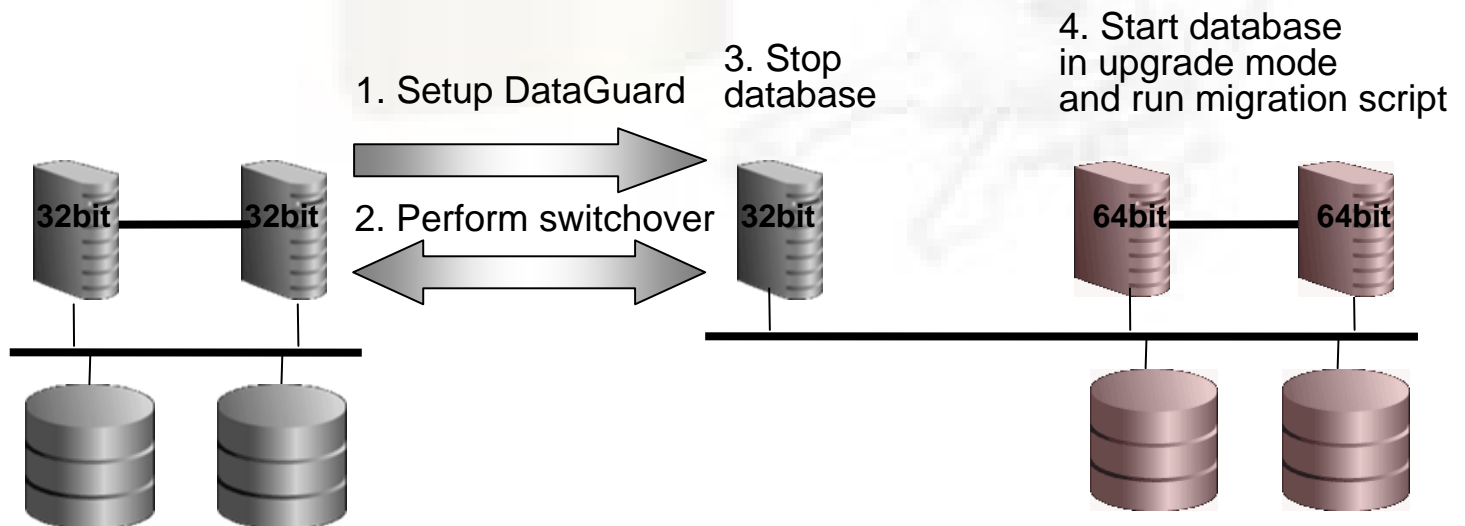
11g Streams test performance

LCRs Flow
STREAMS_PROPAGATE_DEST3D11@T11G64.CERN.CH

Generated on Tuesday 04th of December 2007 02:40:21 PM
Time range 10:36:24 06-11-2007 <-> 12:36:04 06-11-2007

12000

Legend:
- LCRS ENQUEUED/s
- LCRS IN SRC QUEUED/s
- LCRS PROPAGATED/s
- LCRS IN DEST QUEUED/s
- LCRS DEQUEUED/s
- LCRS APPLIED/s

New 11g Streams Features

- ## Split/Merge of Streams
  - maintains high performance for all replicas
  - automated, fast "catch-up" for unavailable replica

- ## Cross-database LCR tracking
  - trace Streams messages from start to finish in single view

- ## Source and Target data compare & converge
  - compare rows in an object at 2 databases
  - converge objects in case of differences

- New hardware acquisitions next year
  - 60 SAN diskservers (16 disks x 400GB)
  - 35 mid-range servers (2 x Intel quad core, 16 GB RAM)

- Moving from 32-bit Linux to 64-bit Linux
  - Migration using Oracle DataGuard
    - minimum downtime required (independent of database size)
    - easy to rollback if something goes wrong

- **Setup DataGuard**

- **Perform switchover**

- **Stop intermediate database**

- **Perform upgrade (*utlirp.sql*)**

1. Setup DataGuard

3. Stop database

4. Start database in upgrade mode and run migration script

2. Perform switchover

32bit   32bit   32bit

64bit   64bit

- Migrated backup/monitoring machine to cluster configuration (active-pasive mode)
    - file system mounted on demand on the node that is active
    - cron jobs schedule backup, monitoring and PKI access management on active node
    - machine behind VIP to provide HA in case the current node fails
    - LDAP setup as cluster service for TNS names resolution

- **PVSS improvements**
  - space waste observed (result of direct path inserts)
  - direct path insert removed from PL/SQL procedures
    - improved space usage
    - satisfactory performance observed
  - Testing compression of PVSS historic data
    - compressing of history tables resulted in ~ 2 times decrease of space utilization and twofold increase of performance (full scans)
    - streams bug encountered – resolved with help of Oracle Support (existing bug fix ported from 64-bit to 32-bit architecture)

- Presentations on WLCG Service Reliability Workshop
  - DB application design issues
  - DB Monitoring tools
  - Streams service review

- Monitoring improvements
  - Streams monitoring moved to HA architecture
  - RAC monitoring now synchronizes blackouts with OEM
  - Better unavailability tracing (overview of scheduled and unscheduled interventions) – one place to enter intervention data

- **Oracle Streams and Data Replication Services**
  - Study the use of transportable tablespaces to synchronize Tier1 sites when recovery window is exceeded
  - Automate more the split-merge procedure when one site has to be dropped / resynchronized
  - Implement failover for the downstream capture component and the monitoring system
- **Highly available database services based on RAC/ASM**
  - Big migration of services to 64-bit Linux using DataGuard
    - 35 new mid-range servers (280 cores) – 175% of current CPU power dedicated to databases (160 cores)
    - investigating possible ways of improving installation procedures

- **Oracle Streams and Data Replication Services**
  - Single Point of Contact: E. Dafonte Perez (CERN) – G. Kerr (Oracle)
  - Participants: D. Duellman (CERN) – P.McElroy (Oracle), A. Downing (Oracle)

- **Streams and RAC monitoring**
  - Participants: D. Wojcik (CERN)

- **Oracle Enterprise Manager**
  - Single Point of Contact: C. Lambert (CERN) – A. Bulloch (Oracle)
  - Participants: D. Wojcik, A. Wiecek (CERN) – G. Kerr (Oracle)

- **Oracle Data Guard**
  - Single Point of Contact: A. Topurov (CERN) – G. Kerr (Oracle)
  - Participants: E. Dafonte, M. Girone, E. Grancher, D. Wojcik (CERN)

- **Highly available database services based on RAC/ASM**
  - Single Point of Contact: D. Wojcik (CERN) – G. Kerr (Oracle)
  - Participants: M. Girone (CERN)

Q & A