Replication solutions for Oracle database 11g

Zbigniew Baranowski

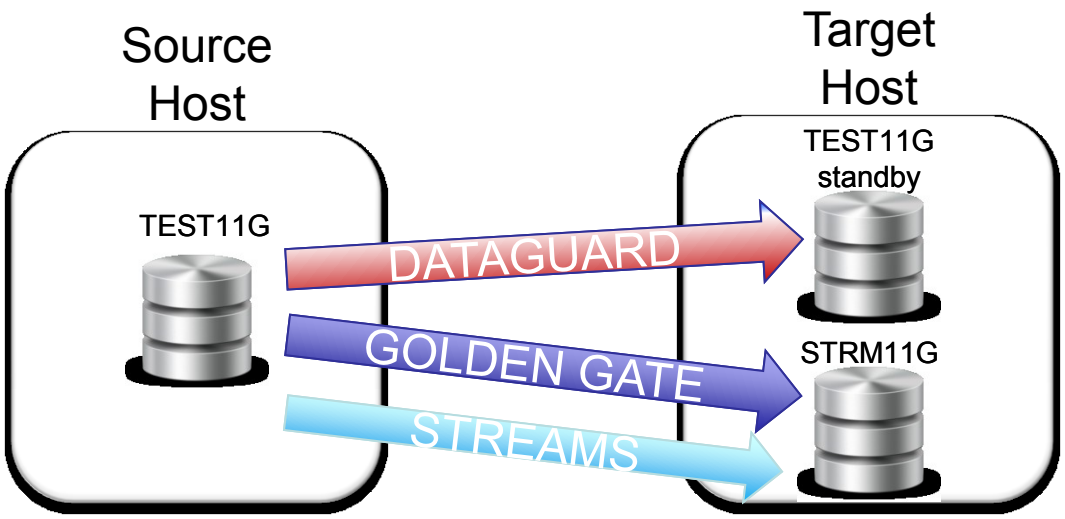
Agenda

- Testing of replication solutions with real production workload
  - Preparing the environment
  - Comparison of results
  - Conclusions
- 11g replication deployment for experiments databases
  - Proposals of changes for 2012
  - Deployment strategy

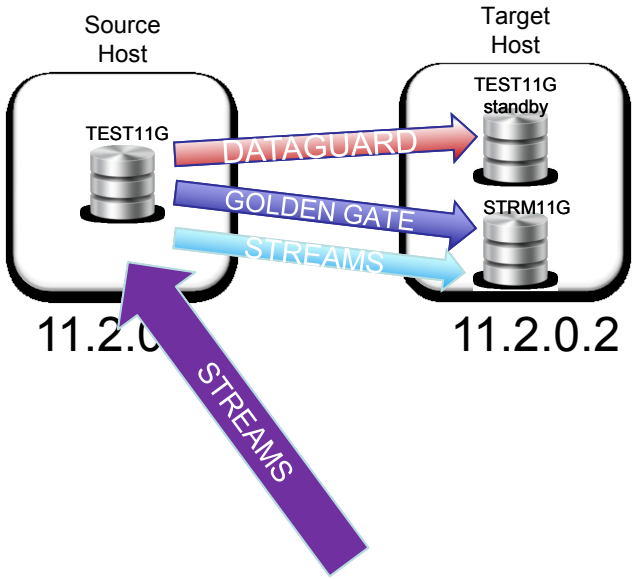
Motivation

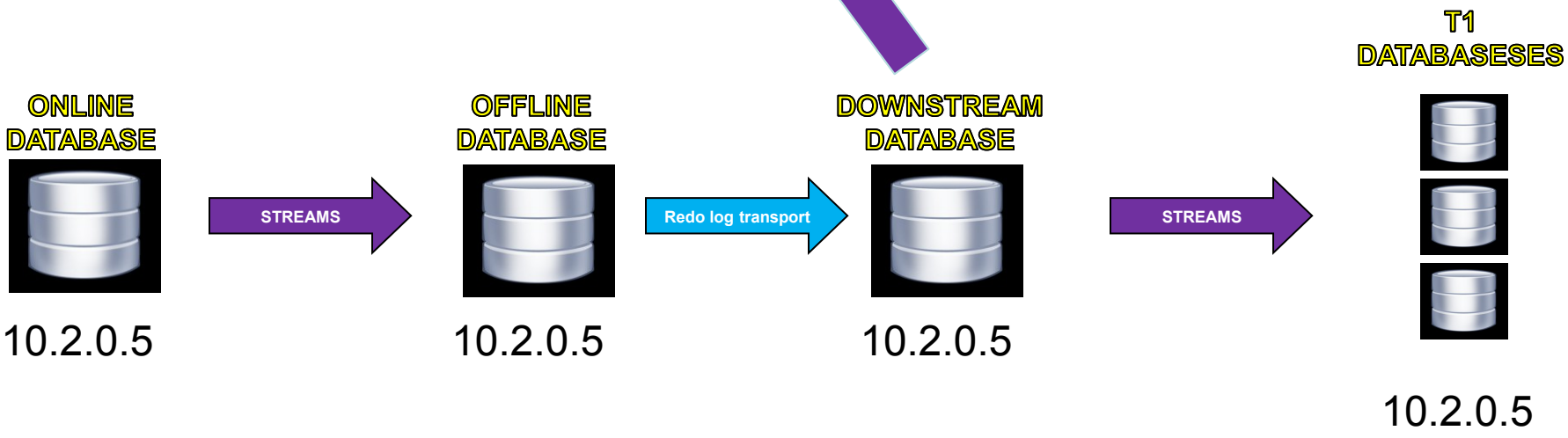
- Validation of technologies with real production data
  - Generic tests rarely uncover bugs and problems
  - Confidence when moving to production
  - Tests have no impact on production environment!
- Measurable and valuable comparison of available replication solutions
  - Exactly the same workload
  - HW&SW configuration equal to production
  - Isolated test environment

Setting up the environment



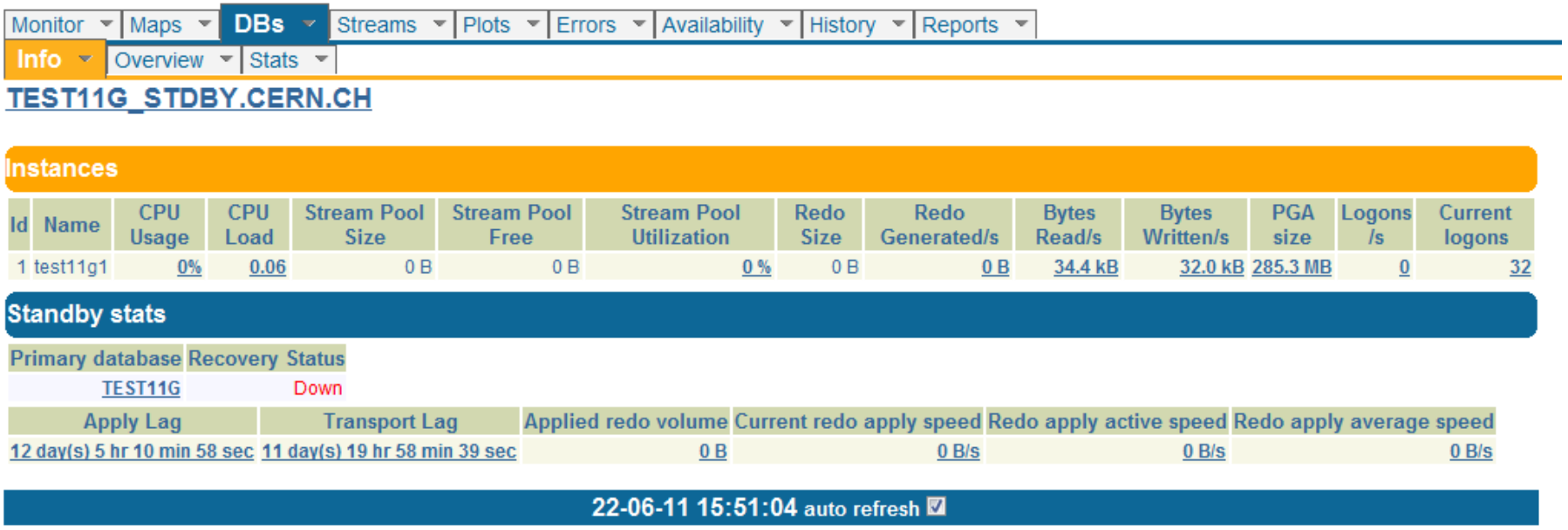
Setting up the environment





Performance Monitoring

- Streams 11g
  - STRMMON – compatible with 11g
- GoldenGate
  - Current out of the box monitoring does not offer much



Monitor | Maps | DBs | Streams | Plots | Errors | Availability | History | Reports

Info | Overview | Stats

TEST11G_STDBY.CERN.CH

Instances

Id	Name	CPU Usage	CPU Load	Stream Pool Size	Stream Pool Free	Stream Pool Utilization	Redo Size	Redo Generated/s	Bytes Read/s	Bytes Written/s	PGA size	Logons /s	Current logons
1	test11g1	0%	0.06	0 B	0 B	0 %	0 B	0 B	34.4 kB	32.0 kB	285.3 MB	0	32

Standby stats

Primary database	Recovery Status
TEST11G	Down

Apply Lag	Transport Lag	Applied redo volume	Current redo apply speed	Redo apply active speed	Redo apply average speed
12 day(s) 5 hr 10 min 58 sec	11 day(s) 19 hr 58 min 39 sec	0 B	0 B/s	0 B/s	0 B/s

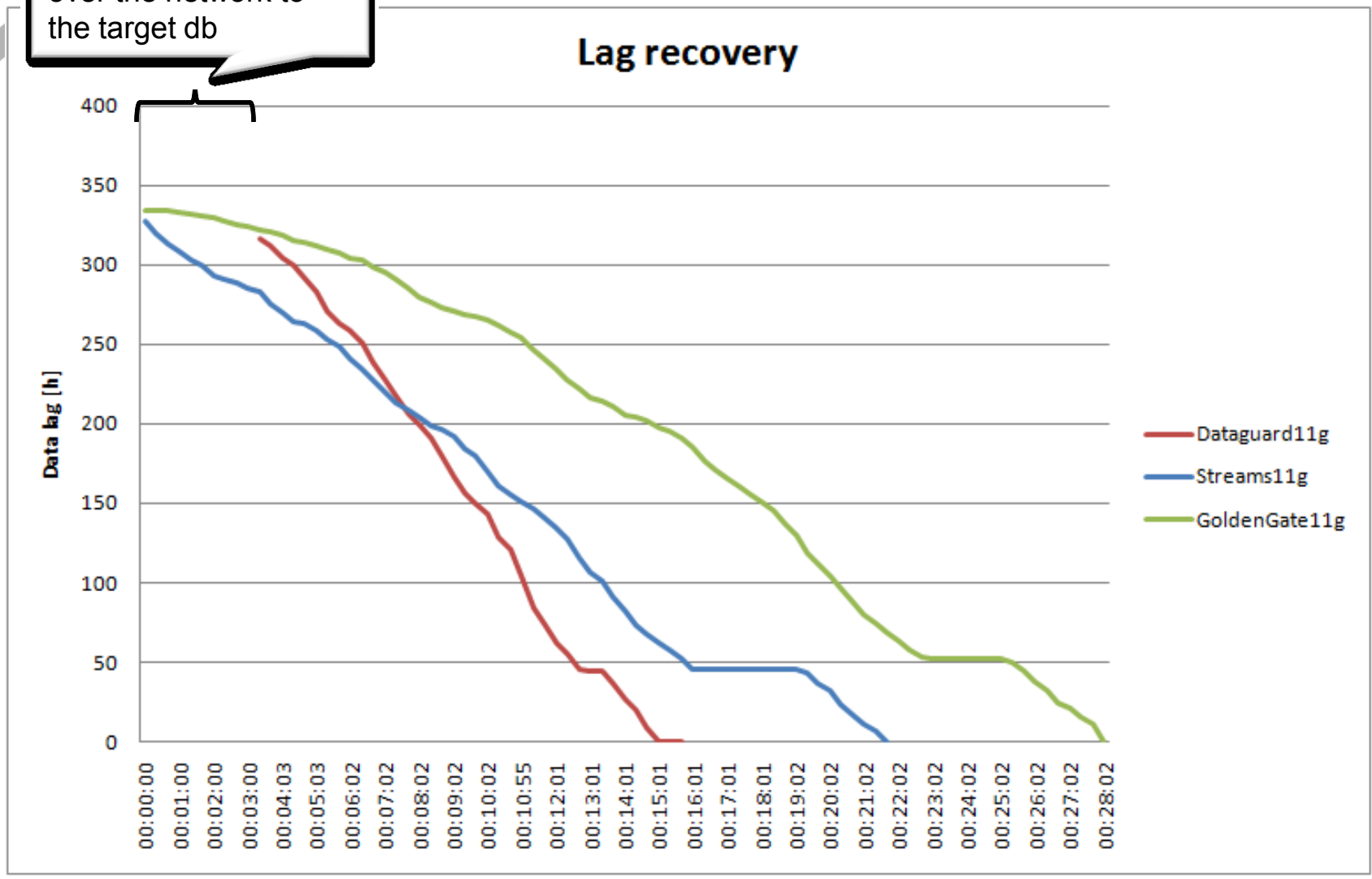
22-06-11 15:51:04 auto refresh ☑

Workload description

- Replication software configuration
  - Default SW configuration
- Workload set #1
  - LHC file catalogue (LHCb)
  - Data window of 13 days (28/05 – 10/06)
  - 15GB of redo logs generated on source
  - ~4M of operations (statements)
    - 1,5M of inserts
    - 1,5M of updates
    - 1M of deletes
  - 7M of row changes

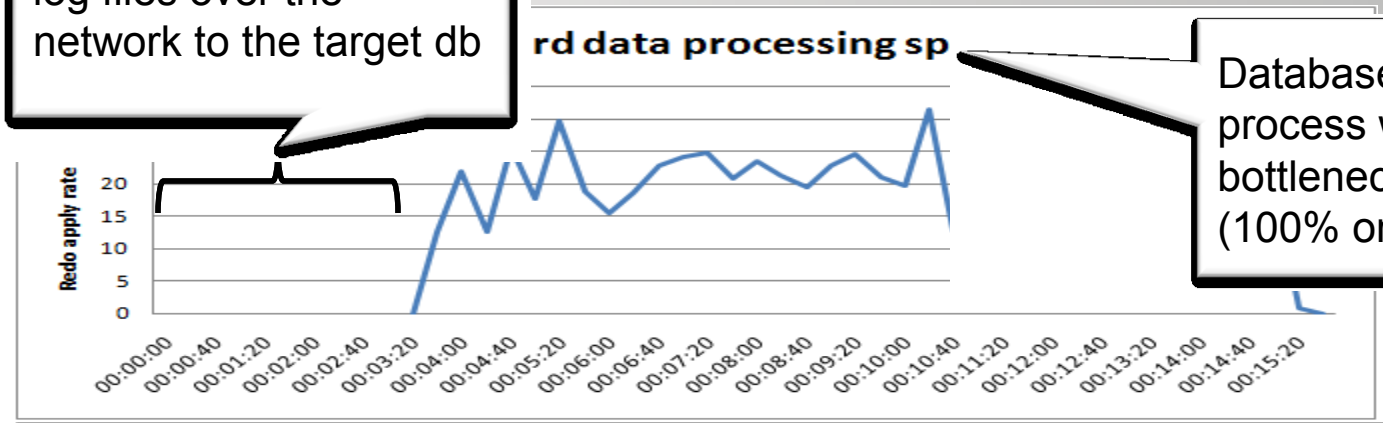
Results for workload #1

Dataguard: Shipping database redo log files over the network to the target db

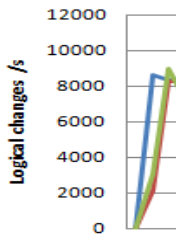


ance results with workload #1

Shipping database redo log files over the network to the target db



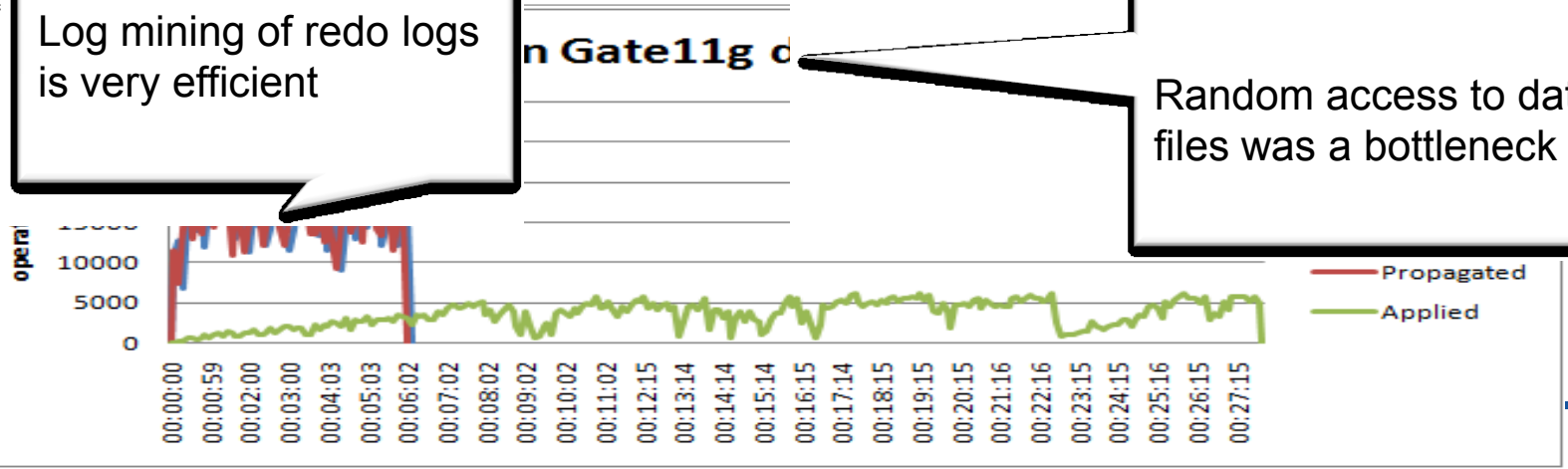
Database writer (DBW) process was a bottleneck due to CPU (100% on one core)



Slow down due to big transaction processing

Random access to database data files was a bottleneck

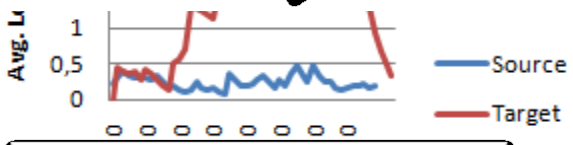
Log mining of redo logs is very efficient



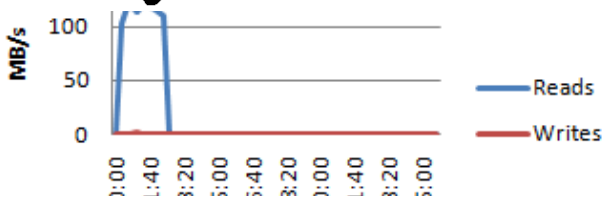
Random access to data files was a bottleneck

Res on b

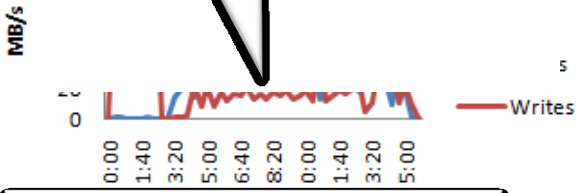
Almost no load on source
Insignificant load on target



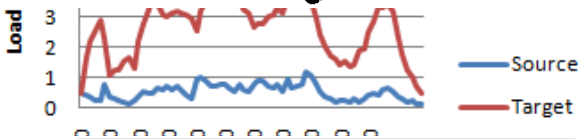
I/O (reads) only for shipping redo logs over the network
No writes on the source



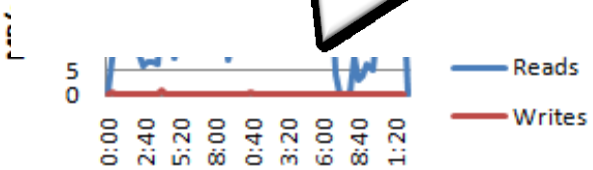
Besides redo download quite significant load of I/O system (reads and writes)



Almost no load on source
Some load on target (apply paralellism)

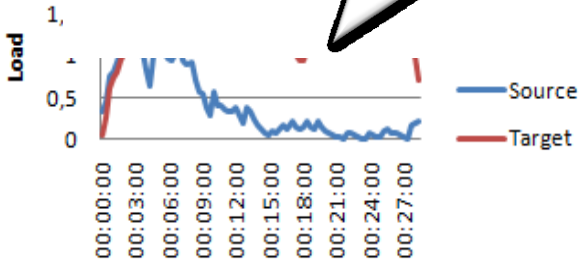


No writes on source.
Quite a lot of reads (logmining)

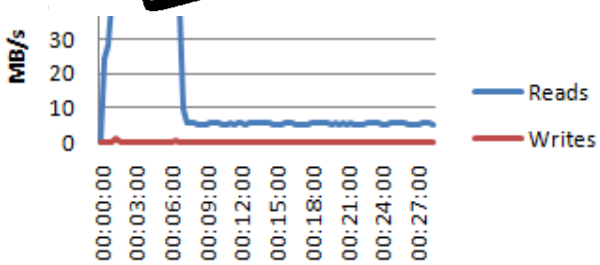


Small amount of reads.
High writes rato (~20MB/s)

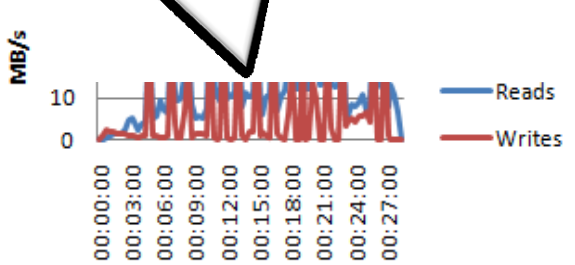
Insignificant load on source and target



High reads rates during logmining



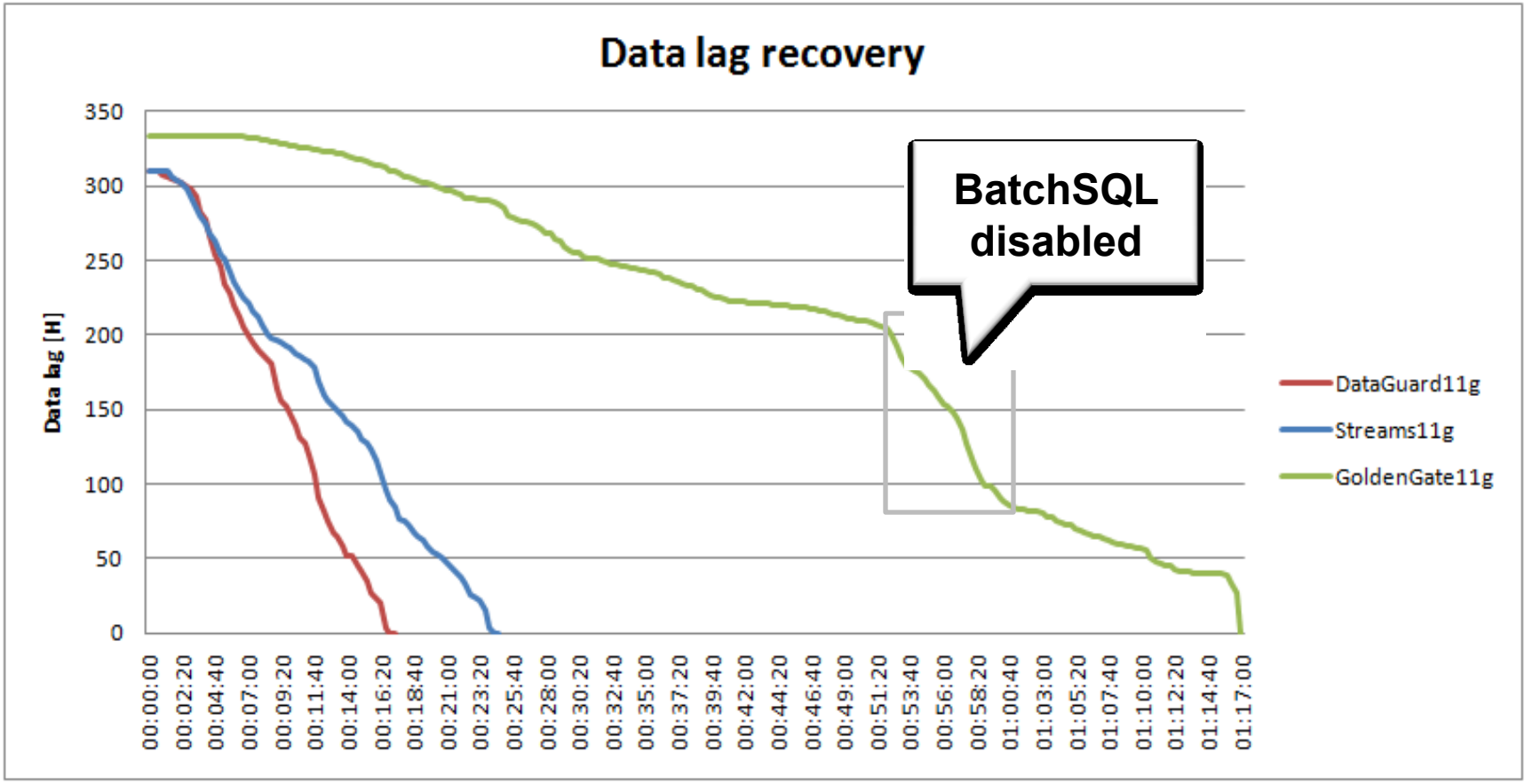
Modrate I/O reates



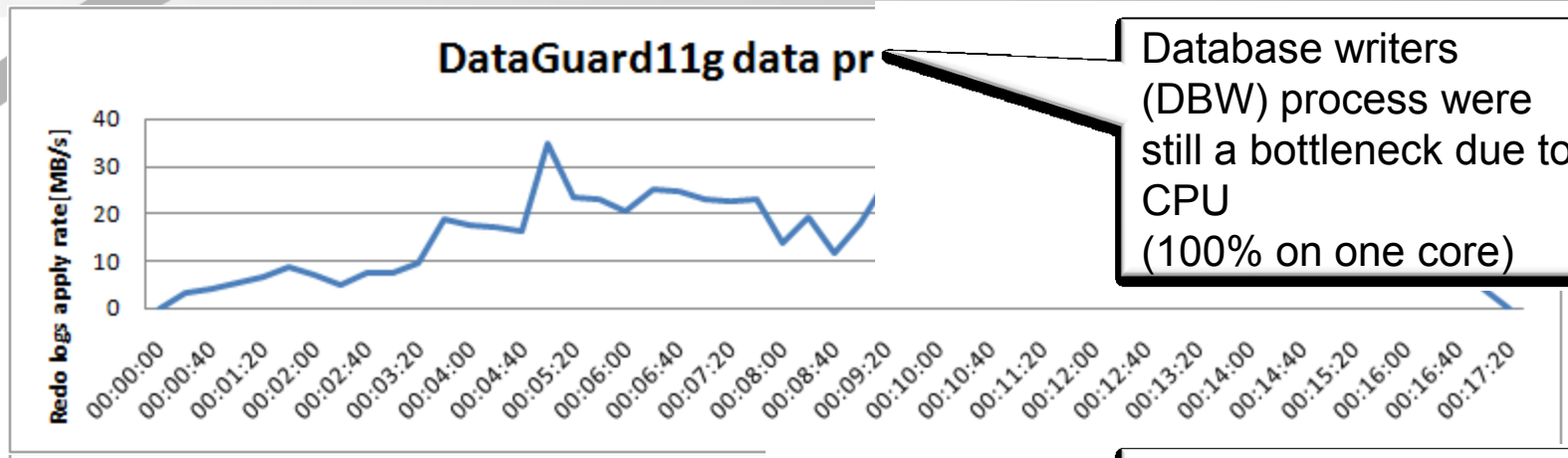
Workload description

- Replication software configuration
  - DataGuard: 2x database writers
  - GoldenGate: SQLBatch optimization enabled
- Workload set #2
  - LHC file catalogue (LHCb)
  - Data window of 13 days (10/06 – 23/06)
  - 17GB of redo logs generated on source
  - ~6M of operations (statements)
    - 2,5M of inserts
    - 2M of updates
    - 1,5M of deletes
  - 3229273 transactions in total (~ 10M of row changes)

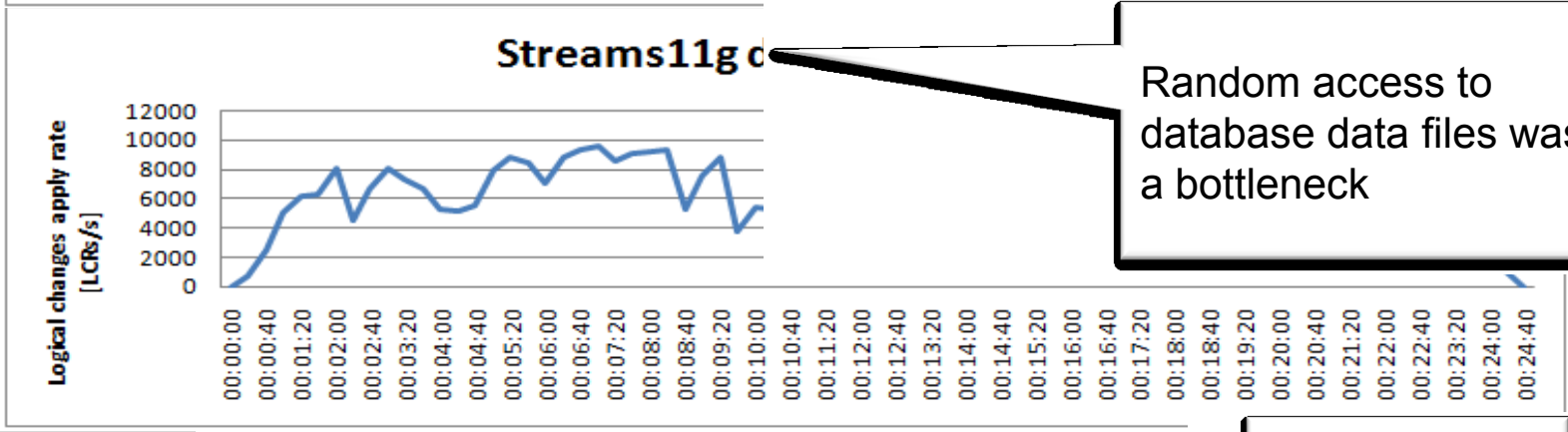
Results for workload #2



Performance results with workload #2

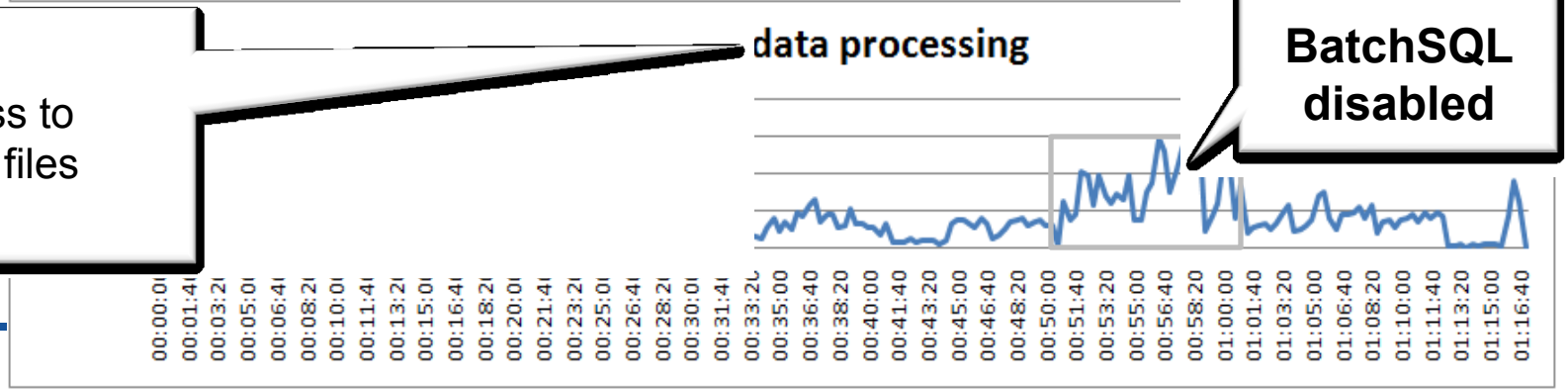


Database writers (DBW) process were still a bottleneck due to CPU (100% on one core)



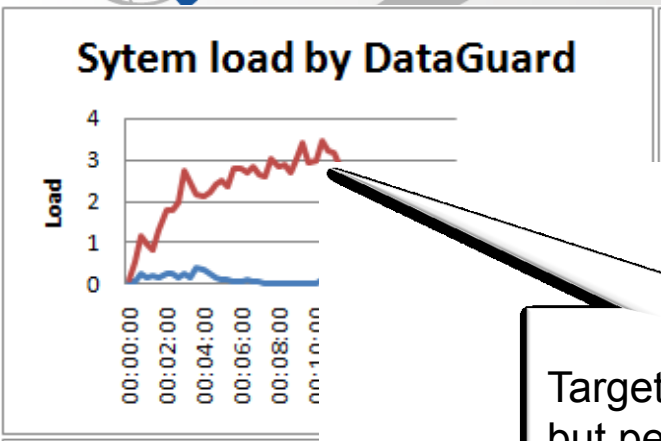
Random access to database data files was a bottleneck

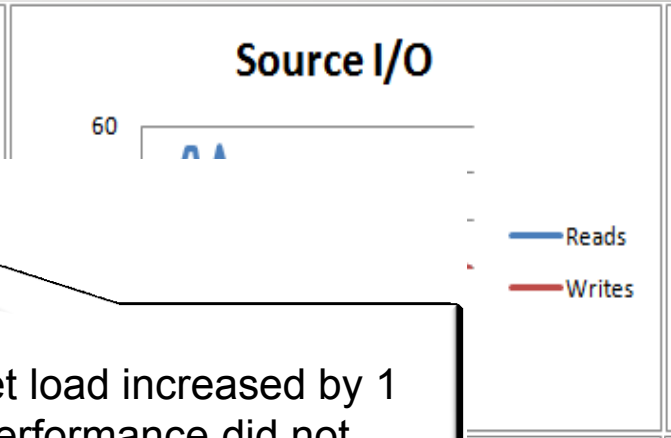
Random access to database data files

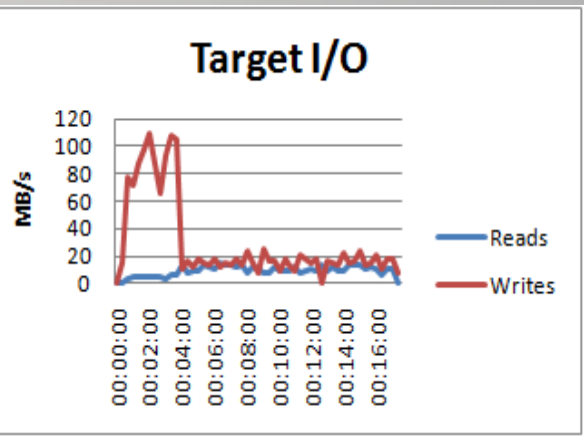


BatchSQL disabled

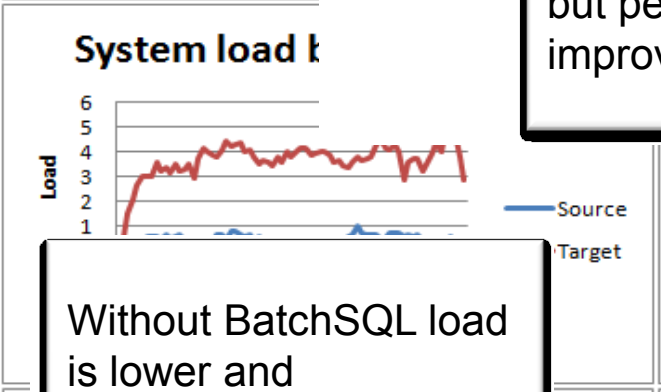
Resource utilization by workload#2

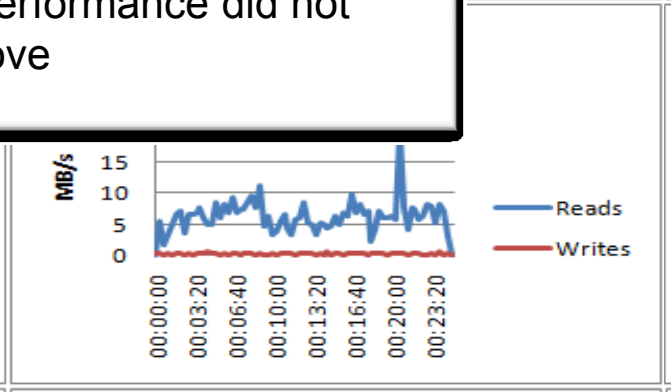


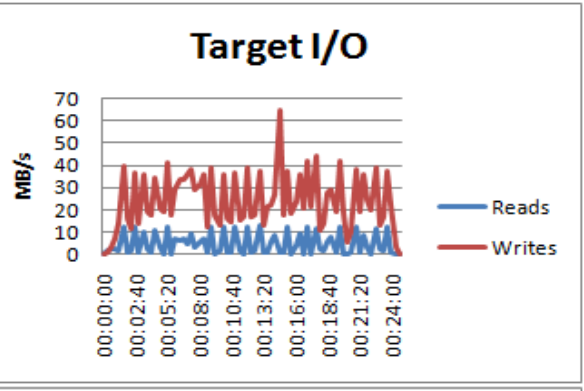




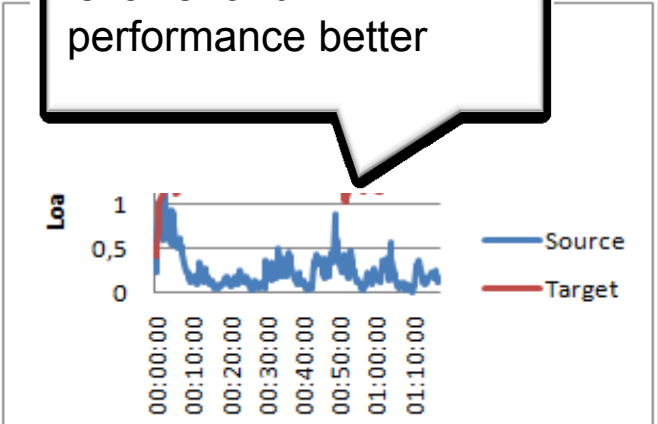
Target load increased by 1 but performance did not improve

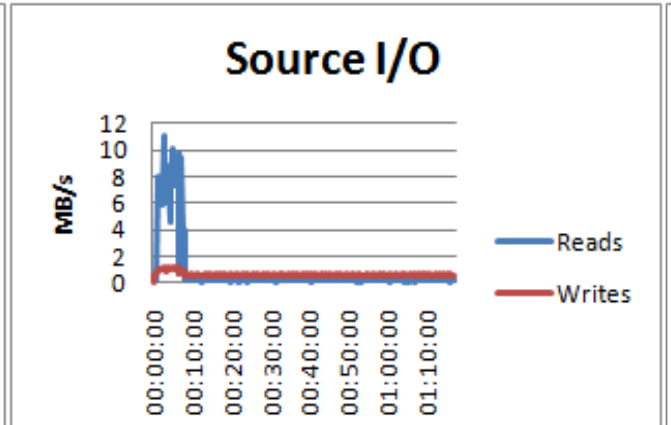


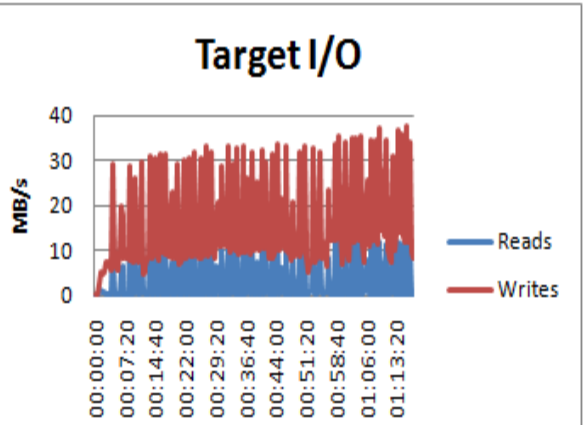




Without BatchSQL load is lower and performance better







Conclusions

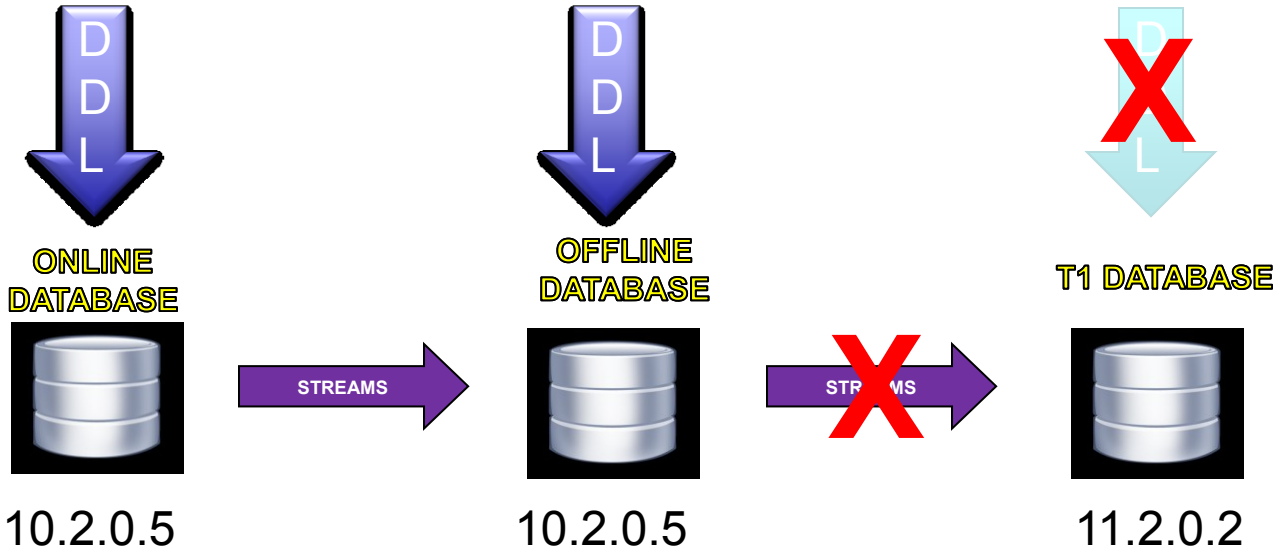
- DataGuard was the fastest technology
  - Streams were slower by 1.5
  - GoldenGate was slower by 2
- Optimization of DG did not bring significant effects
- BatchSQL optimization brought down GG performance by 3 introducing overhead additionally
- No more SW optimization from Streams and Goldengate

Conclusions

- Resource utilization
  - Insignificant load on source system (all)
  - DataGuard while using log buffers does not need to read anything from disks on source db
  - DataGuard writes data to replica most efficient (lowest write rates)
  - Streams introduce highest load on target system (parallel composition of data changes)
  - Streams required a lot of reads on source (~15MB/s) system but less on target (use of buffered queues)

Bug discoverd

- Streams DDL replication incomatibiities between RDBMS 10.2.0.5 and 11.2.0.2

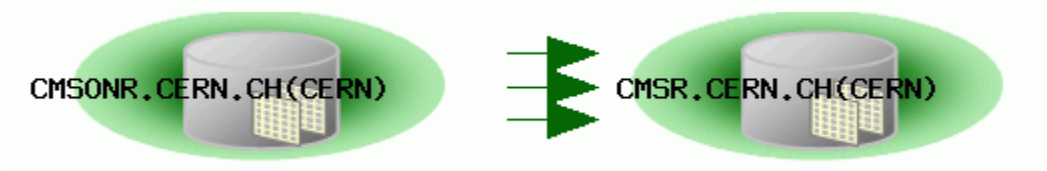


Replication plans for 2012

- No changes (Streams11g)
  - ATLAS (foreseen in 2013)
  - LHCb
  - COMPASS
- Streams replacement with ADG
  - CMS
  - ALICE

CMS overview

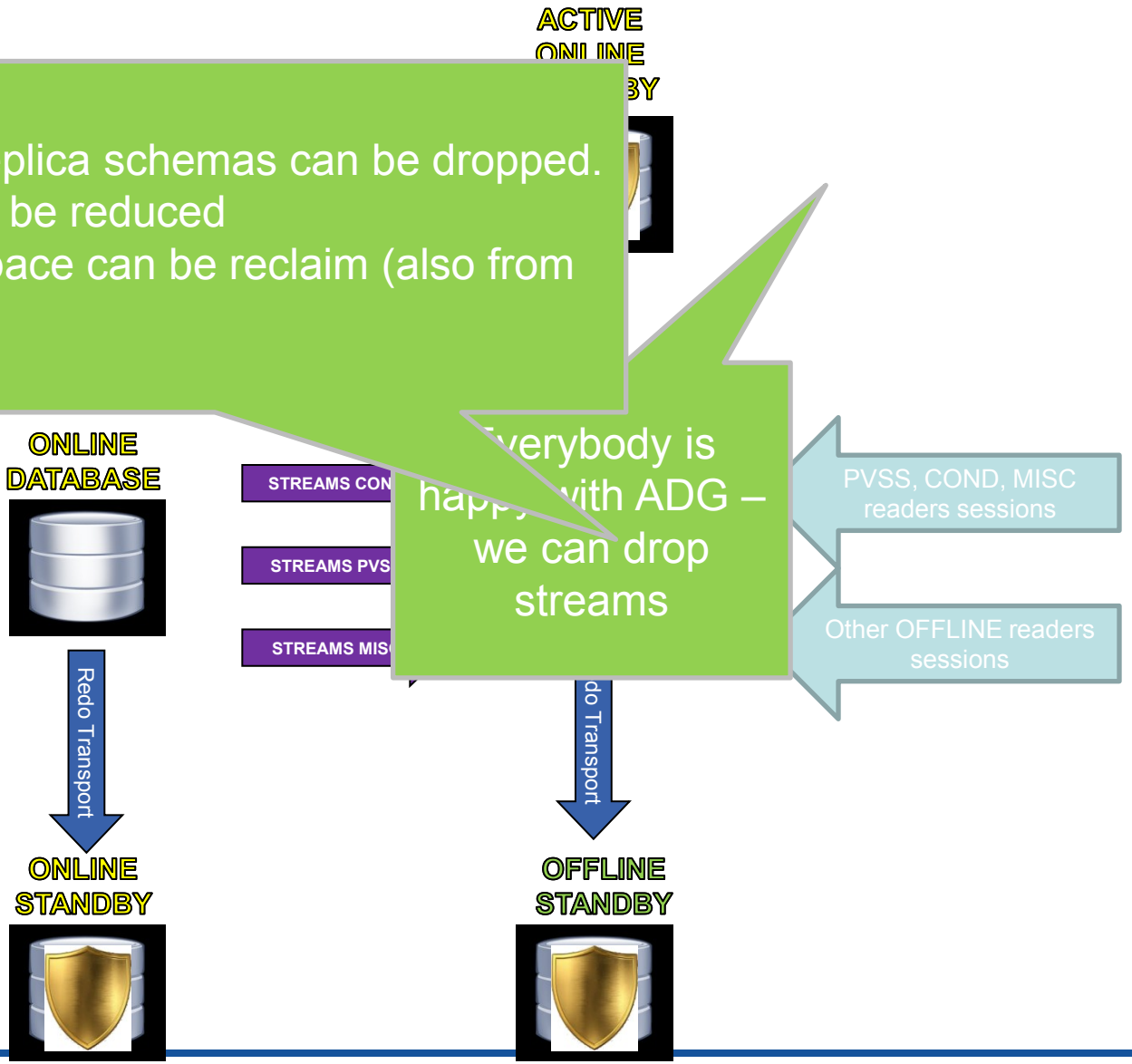
- Databases
  - Online (CMSONR)
    - Size 3486 GB
    - 588 schemas
  - Offline (CMSR)
    - Size 3459GB



- Replications
  - 3 streamings (conditions, pvss and misc)
  - 60% of CMSONR database size is replicated
    - 92 schemas, 2078 GB
  - many DDL updates
  - many cross schema dependencies
  - high workload - around 500 LCR/s (periodical data delivery latency on PVSS replication)

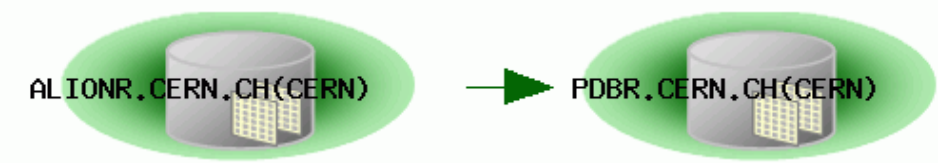
Safe streams replacement with ADG for CMS

- PVSS, COND replica schemas can be dropped.
- CMSR RAC can be reduced
- some storage space can be reclaim (also from STANDBY)



ALICE overview

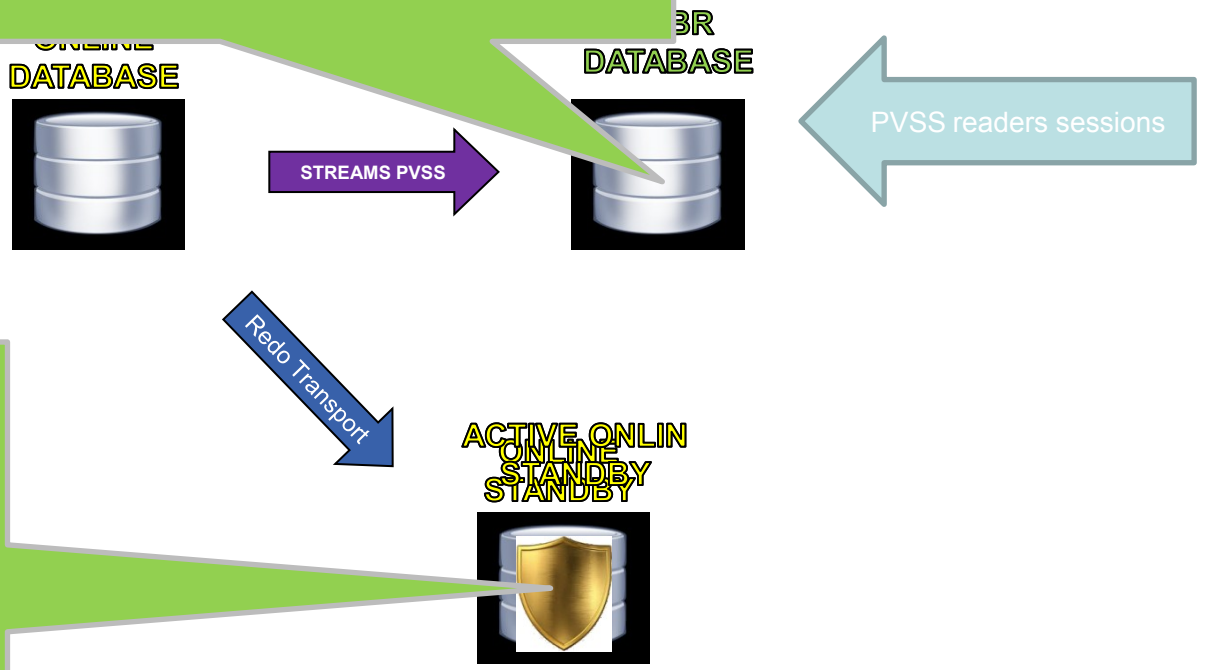
- Databases
  - Online (ALIONR)
    - size: 2788GB
    - 97 schemas
  - Offline (PDBR)
    - Size 1174GB



- Replications
  - PVSS
    - 2 schemas ,431 GB (15% of ALIONR), avg workload 50 LCRs/s
    - Replica size on PDBR: 349 GB (30% of db size)

Safe streams replacement with ADG for ALICE

- PVSS replica schemas can be dropped.
- some storage space can be reclaim on PDBR



Everybody is happy with ADG – we can drop streams

Future plans

- Switchover and failover tests in double standby database environment
- Validation of active standby with CMS applications
- GodenGate 11.2 beta testing?