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

ONLINE DATABASE

10.2.0.5 ➔ STREAMS ➔ OFFLINE DATABASE

10.2.0.5 ➔ Redo log transport ➔ DOWNSTREAM DATABASE

10.2.0.5 ➔ STREAMS ➔ T1 DATABASESES

10.2.0.5
Performance Monitoring

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

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

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 53 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

Lag recovery graph comparing different methods:
- Dataguard 11g
- Streams 11g
- GoldenGate 11g
Performance results with workload #1

- Database writer (DBW) process was a bottleneck due to CPU (100% on one core).
- Random access to database data files was a bottleneck.
- Slow down due to big transaction processing.
- Log mining of redo logs is very efficient.
- Random access to data files was a bottleneck.
Resource utilization by workload

- **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 parallelism)

- **Insignificant load on source and target**

- **Insignificant load on source**
  - High reads rates during log mining

- **Small amount of reads**
  - High writes rate (~20MB/s)

- **Moderate I/O rates**
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

BatchSQL disabled
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

Data processing
Without BatchSQL load is lower and performance better

Target load increased by 1 but performance did not improve
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 discovered

- Streams DDL replication incompatibilities 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
Databases
- Online (CMSONR)
  - Size 3486 GB
  - 588 schemas
- Offline (CMSR)
  - Size 3459 GB

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

Everybody is happy with ADG – we can drop streams

PVSS, COND, MISC readers sessions

Other OFFLINE readers sessions
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)
Everybody is happy with ADG – we can drop streams.

- PVSS replica schemas can be dropped.
- Some storage space can be reclaimed on PDBR.
Future plans

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