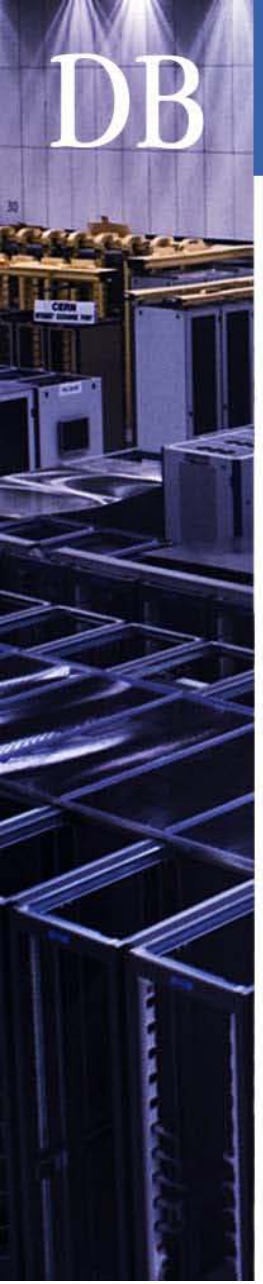


Ideas for evolution of replication technology @ CERN

Openlab Minor Review

December 14th, 2010

Zbigniew Baranowski, IT-DB

Outline

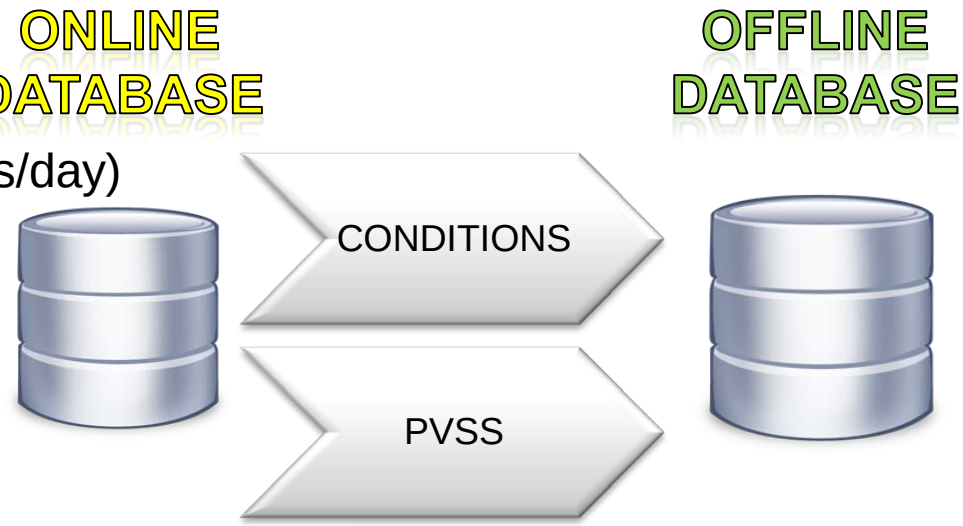
- Replication use cases at CERN
- Motivation for evolution of replication
- Oracle replication technologies
- Possible future replication solutions for LCG
- Summary

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Replication use cases: ONLINE - OFFLINE

CERN IT Department

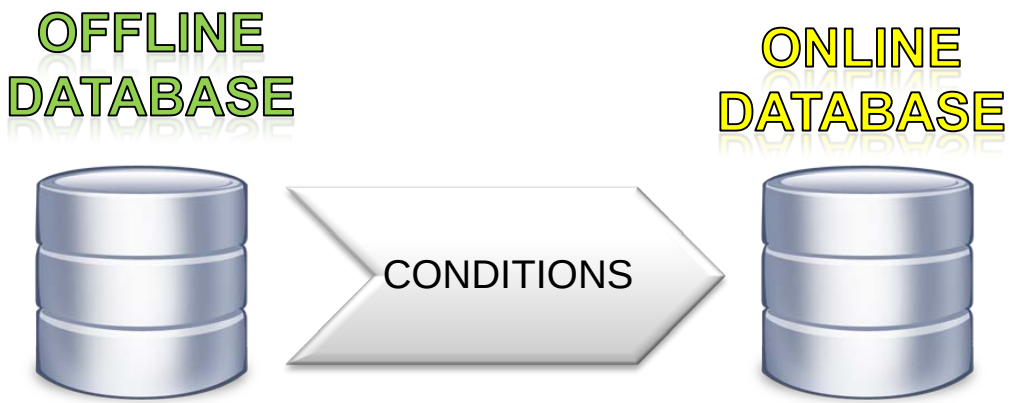
- **ATLAS**
  - CONDITIONS (4M LCRs/day)
  - PVSS (60M LCRs/day)
- **CMS**
  - CONDITIONS (6M LCRs/day)
  - PVSS (20M LCRs/day)
- **LHCb**
  - CONDITIONS (6K LCRs/day)
- **ALICE**
  - PVSS (4M LCRs/day)
- **COMPASS**
  - PVSS (4M LCRs/day)



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Replication use cases: OFFLINE - ONLINE

- **LHCb** (in addition to ONLINE-OFFLINE)
  - CONDITIONS (8K LCRs/day)

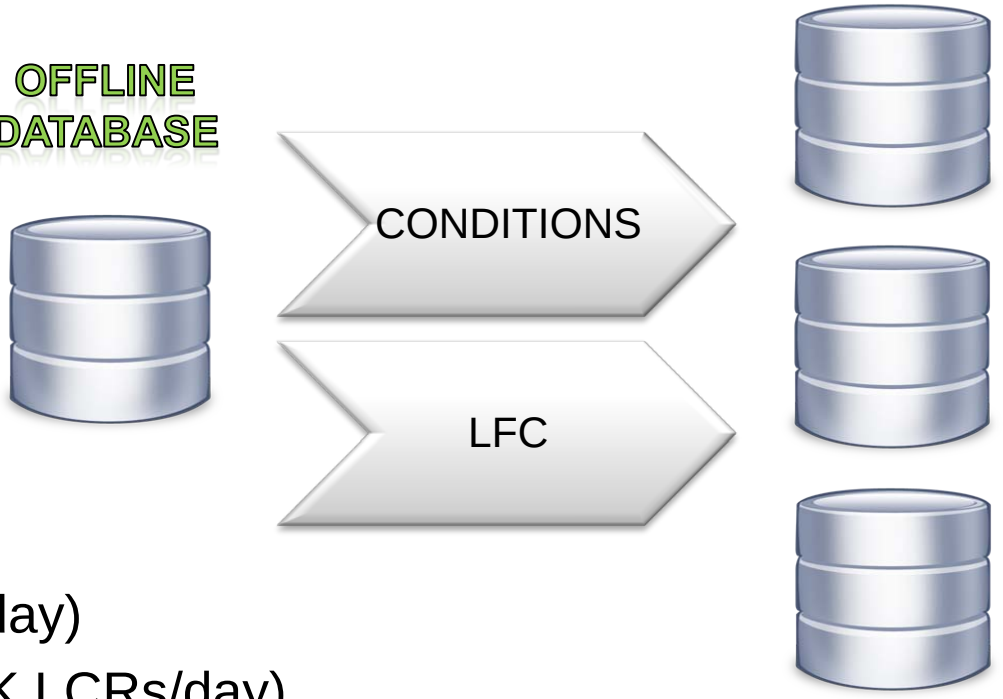


CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Replication use cases: OFFLINE – T1s

CERN IT Department

- **ATLAS**
  - CONDITIONS (4M LCRs/day)



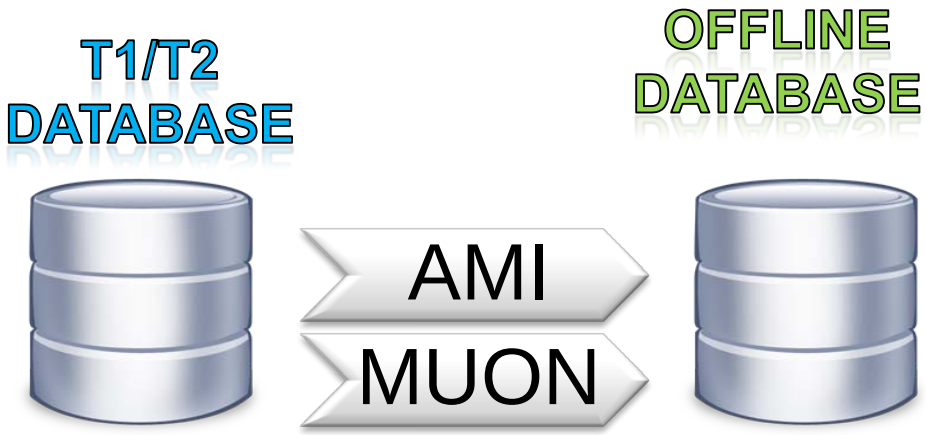
- **LHCb**
  - LFC (235K LCRs/day)
  - CONDITIONS (15K LCRs/day)

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Replication use cases: T1 - OFFLINE

CERN IT Department

- **ATLAS**
  - AMI (800K LCRs/day)
  - Muon (700K LCRs/day)



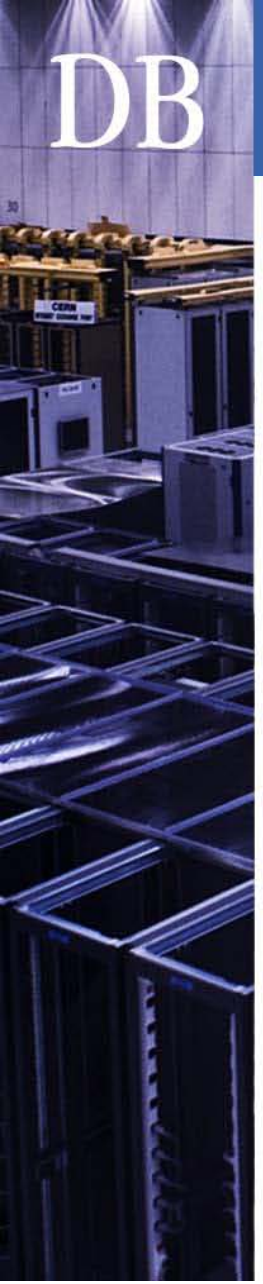
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Motivation for evolution of replication solutions

CERN IT Department

- Need of stable and reliable replication service
- Streams 10g require frequent interventions (at least once per week)
  - Consistency problems
  - Blocking sessions
  - Memory pools shortage
  - Logminer crashes
  - Users unsupported changes
- Streams administration is time consuming and requires expert knowledge
- Migration to 11gR2 in 2012

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it



Motivation for other replication solutions

- Is there a solution which can simplify maintenance of replication?
  - Satisfies physics data workload
  - Requires minimum maintenance effort
  - Is resilient to user's unsupported operations
  - Ensures replicated data consistency
  - Utilizes minimum amount of resources

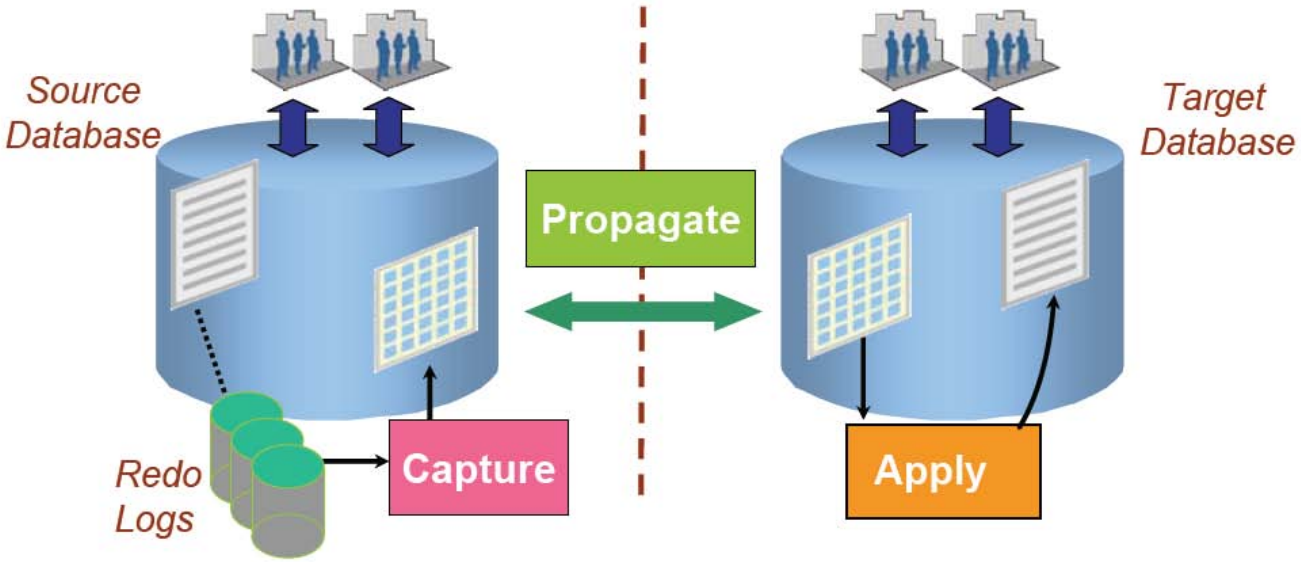
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Possible replication solutions

- Logical (SQL based) replication
  - Streams11gR2
  - GoldenGate
- Physical (block-level) replication
  - Active DataGuard11gR2
- Combinations of physical and logical replication

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Streams 11gR2



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

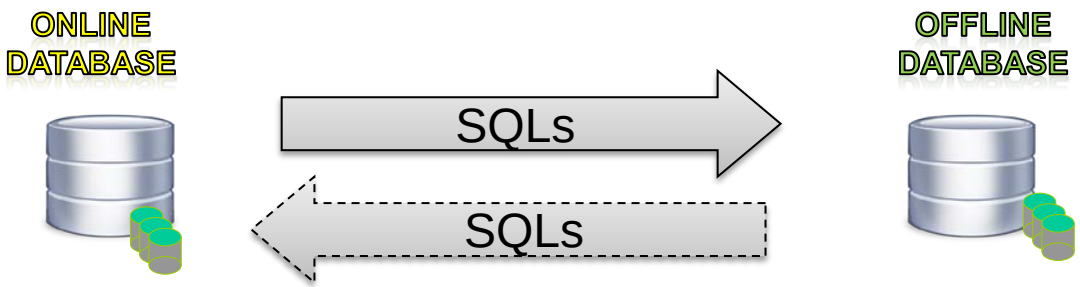
Streams11gR2 solution

- Technology features
  - ☹ Considerable maintenance effort
    - but in 11g should be less than in 10g
  - ☺ No additional license required
  - ☺ Many improvements
    - stability, management, monitoring, verification of data consistency
  - ☺ Very good performance (30K-40K LCRs/s)
  - ☺ Best practices identified – a lot of experience
  - ☺ Source and destination database fully accessible for reads and writes

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Streams11gR2 solution

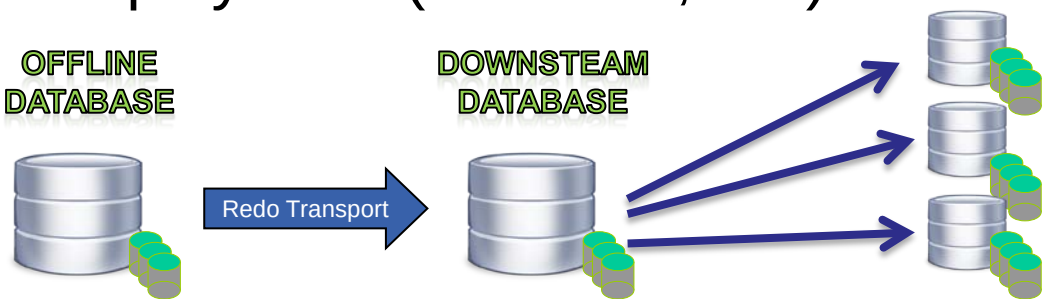
- As ONLINE – OFFLINE replication
  - ☹ Users and data content can abort the replication
  - ☹ streams processes may affect performance of online database
  - ☺ no extra hardware needed
  - ☺ bi-directional replication



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

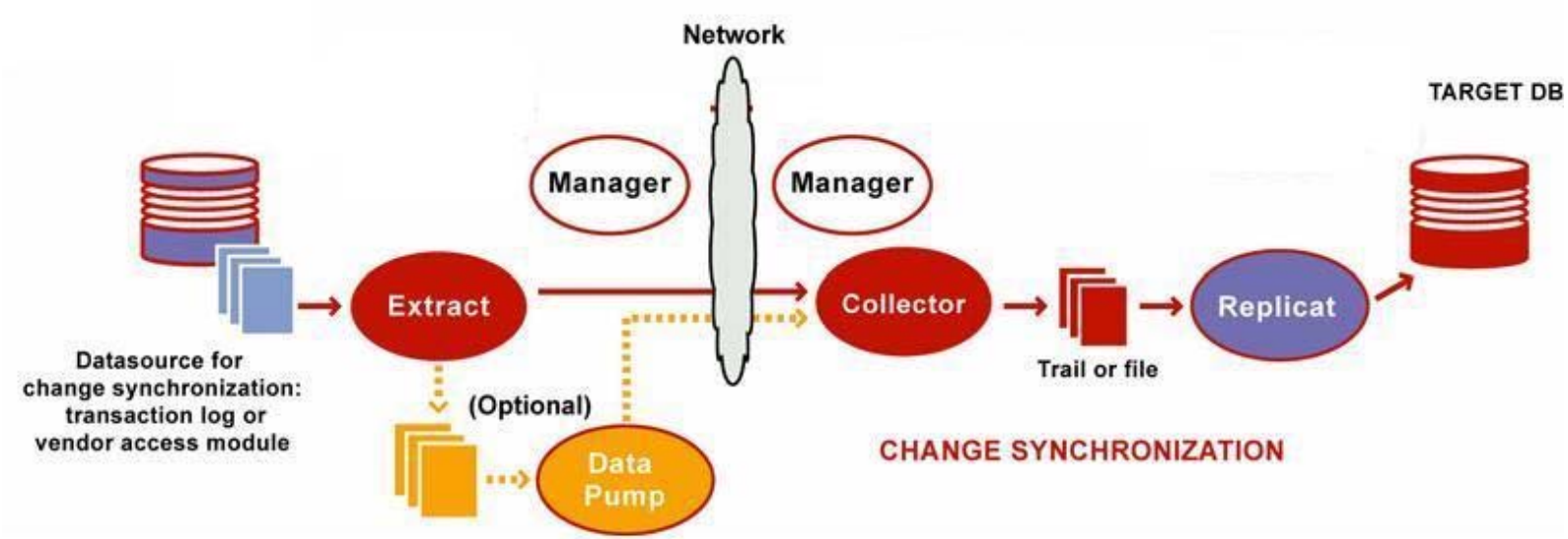
Streams11gR2

- As OFFLINE – T1s
  - Recovery of replica requires
    - ☹ coordination between T1 and other T1, T0
    - ☹ expert knowledge of procedures
  - Downstream capture
    - ☹ additional hardware required
    - ☺ complete isolation from OFFLINE database
    - ☺ standby database can be source of replication
  - ☺ T1s databases is read/write accessible
  - ☺ Good monitoring for distributed streams deployment (strmmon, EM)



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

GoldenGate



Source: Oracle.com

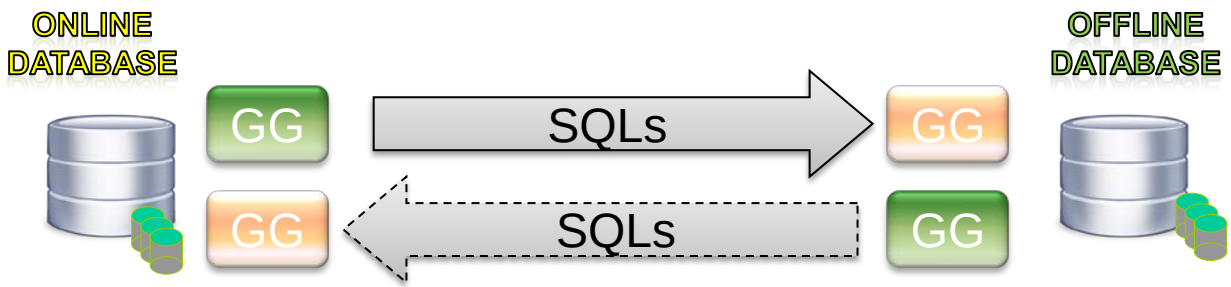
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

GoldenGate

- Technology features
  - ☺ source and destination database fully accessible for reads and writes
  - ☺ good quality of software (very stable, free of locks, almost transparent for databases)
  - ☺ good performance (comparable to Streams11g)
  - ☹ additional license required
  - ☹ standby database cannot be used as source
  - ☹ no in-house experience
  - ☹ additional dedicated disk space required for trail files
  - ☹ additional software to be installed and maintained on database's machines

GoldenGate solution

- As ONLINE-OFFLINE replication
  - ☺ no extra hardware needed
  - ☺ possible loops back in replication
  - ☺ minor impact on source database
  - ☹ users and data content can abort the replication

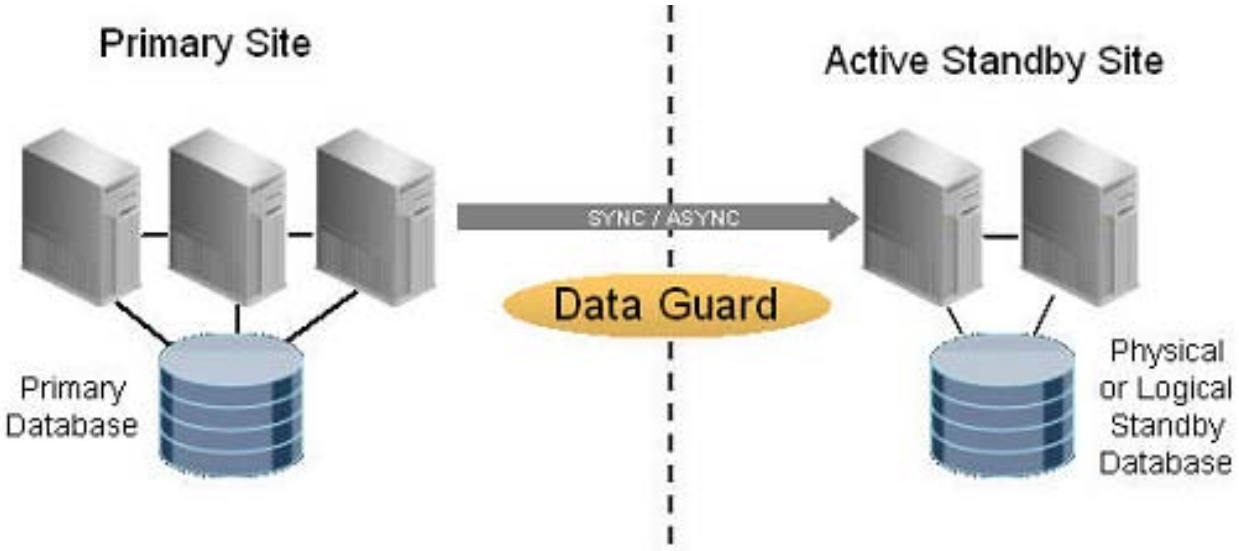


GoldenGate solution

- As OFFLINE – T1s
  - ☺ easier maintenance
    - No side effects on source when target is down
    - No split of replication required
    - Trail files can be used for T1 recovery
  - ☹ no remote administration - access to nodes required
  - ☹ no monitoring for distributed environment
  - ☹ cannot use standby database (i.e. Active Dataguard) as a source of replication

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Active DataGuard 11gR2



Source: Oracle.com

Active DataGuard 11gR2

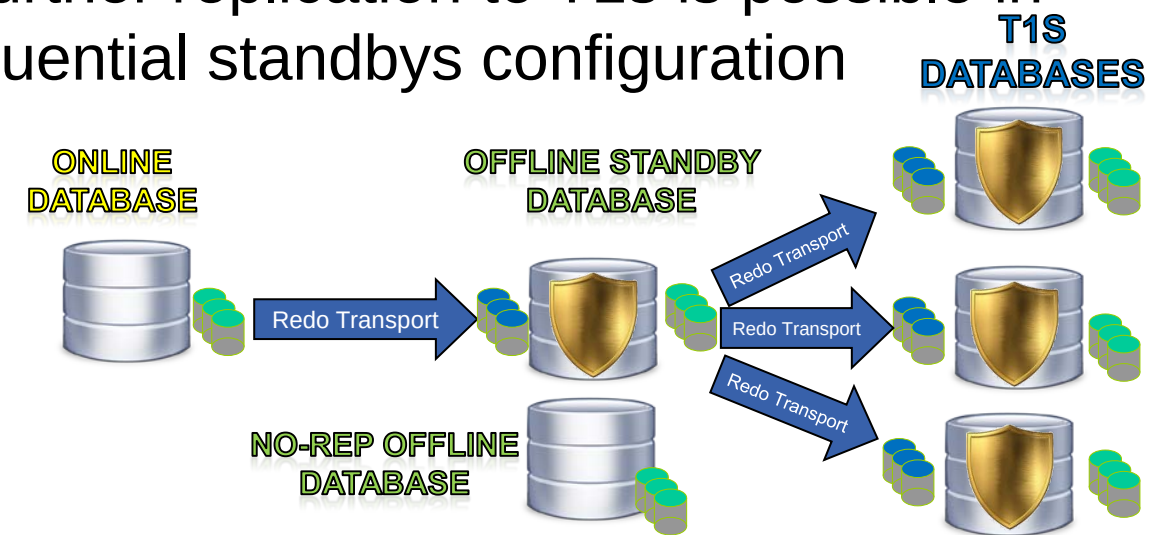
CERN IT Department

- Technology features
  - ☺ Physical replication
    - identical copy
  - ☺ Minimum maintenance effort
  - ☺ Outperforms other replication technologies
    - Oracle claims 200 MB/s of redo processing
  - ☺ Improved data reliability of primary database
    - failover
    - automatic recovery of corrupted blocks
  - ☺ Fast recovery with RMAN
  - ☹ Additional license required
  - ☹ Target/standby database is read only

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Active DataGuard 11gR2

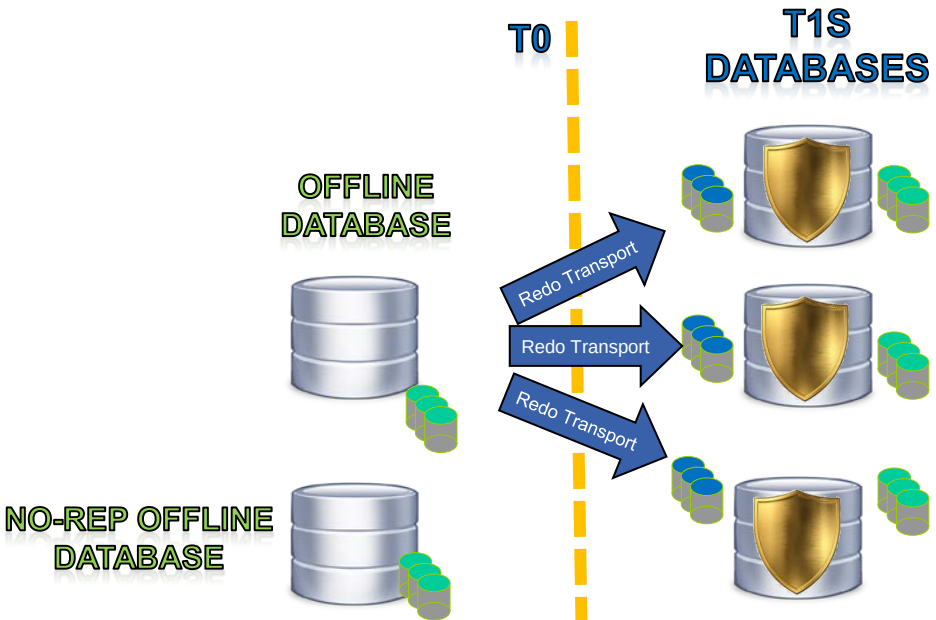
- As ONLINE – OFFLINE replication
  - ☹ additional database installations needed for no replicated data (split of OFFLINE)
  - ☹ same version of software required (installation, upgrades)
  - ☺ online database is protected with another standby database
  - ☺ further replication to T1s is possible in sequential standbys configuration



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Active DataGuard 11gR2

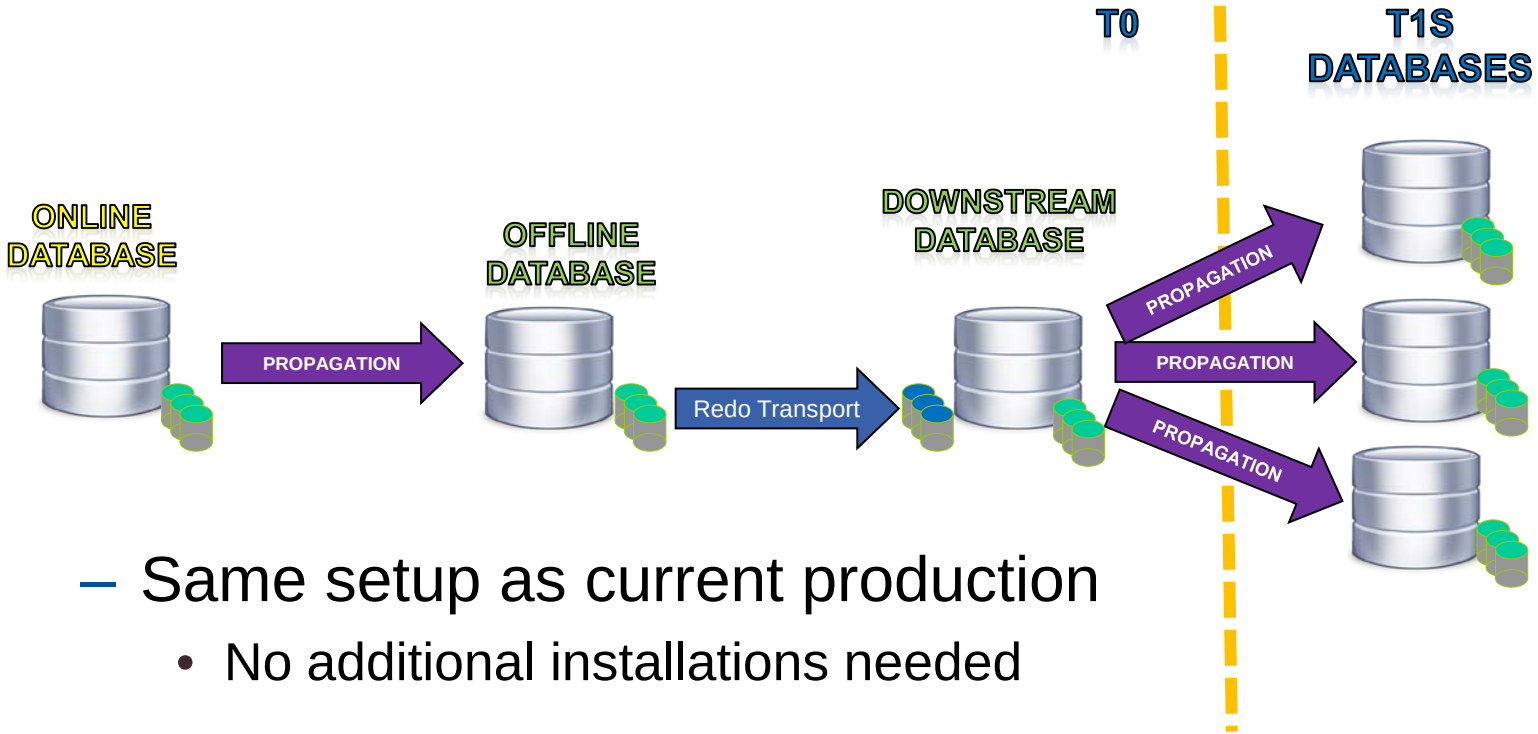
- As OFFLINE – T1s
  - ☹ same version required on all T1s DBs
    - Coordination of interventions becomes critical
  - ☹ T1 database is read only
  - ☹ additional database installations needed for no replicated data (split of OFFLINE)
  - ☺ Physical replication: lower maintenance effort
  - ☺ No downstream needed



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Possible solutions

- Streams11gR2 replication at all Tiers

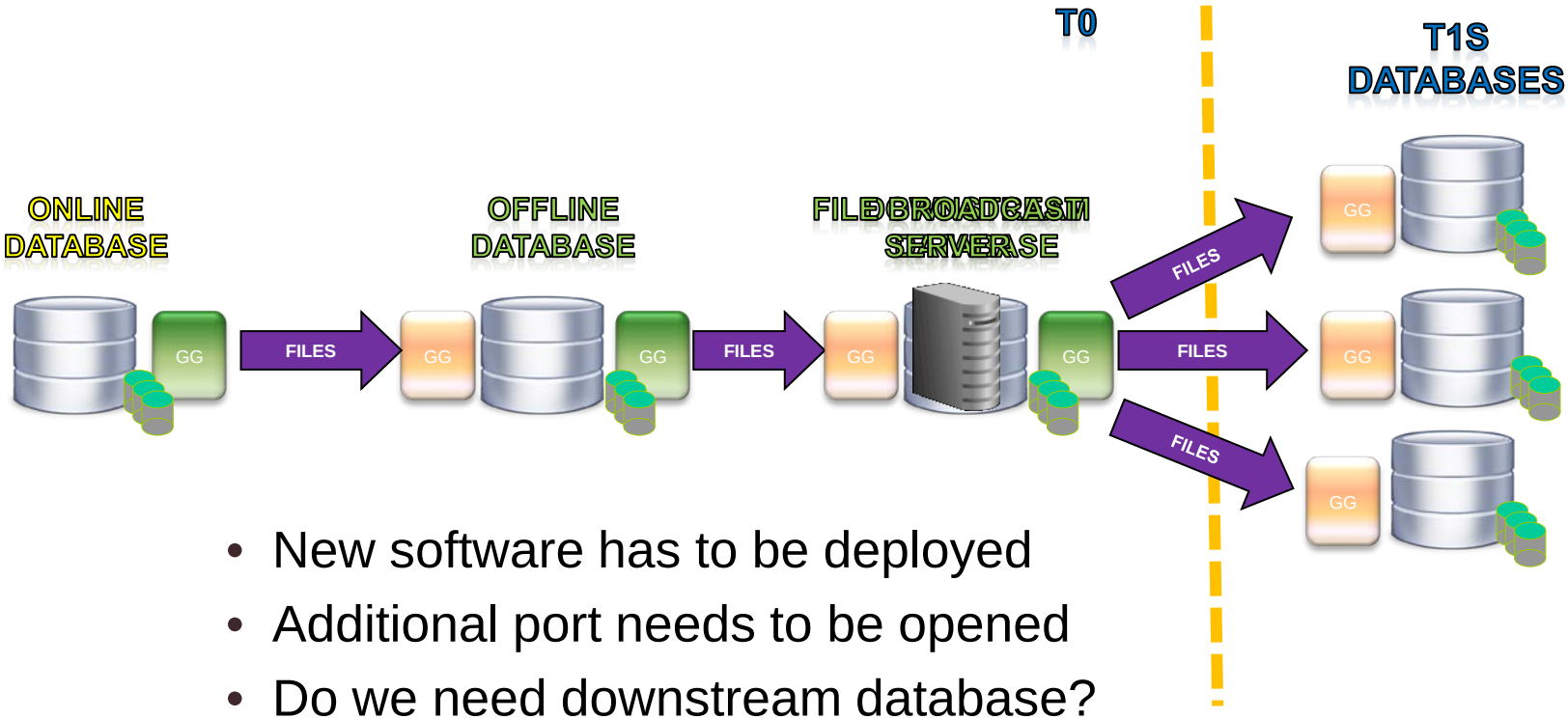


  - Same setup as current production
    - No additional installations needed

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Possible solutions

- GoldenGate replication at all Tiers

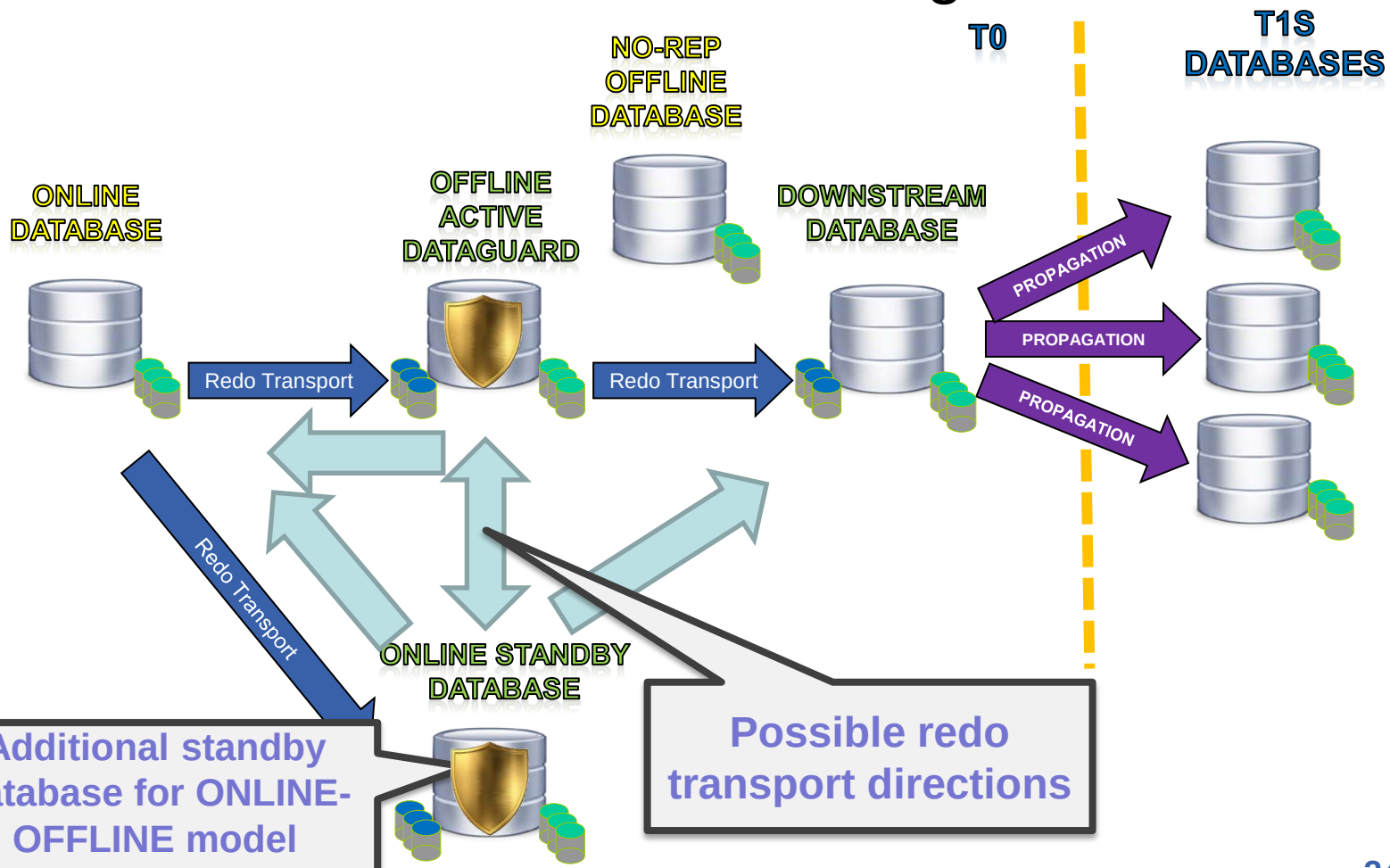


- New software has to be deployed
- Additional port needs to be opened
- Do we need downstream database?

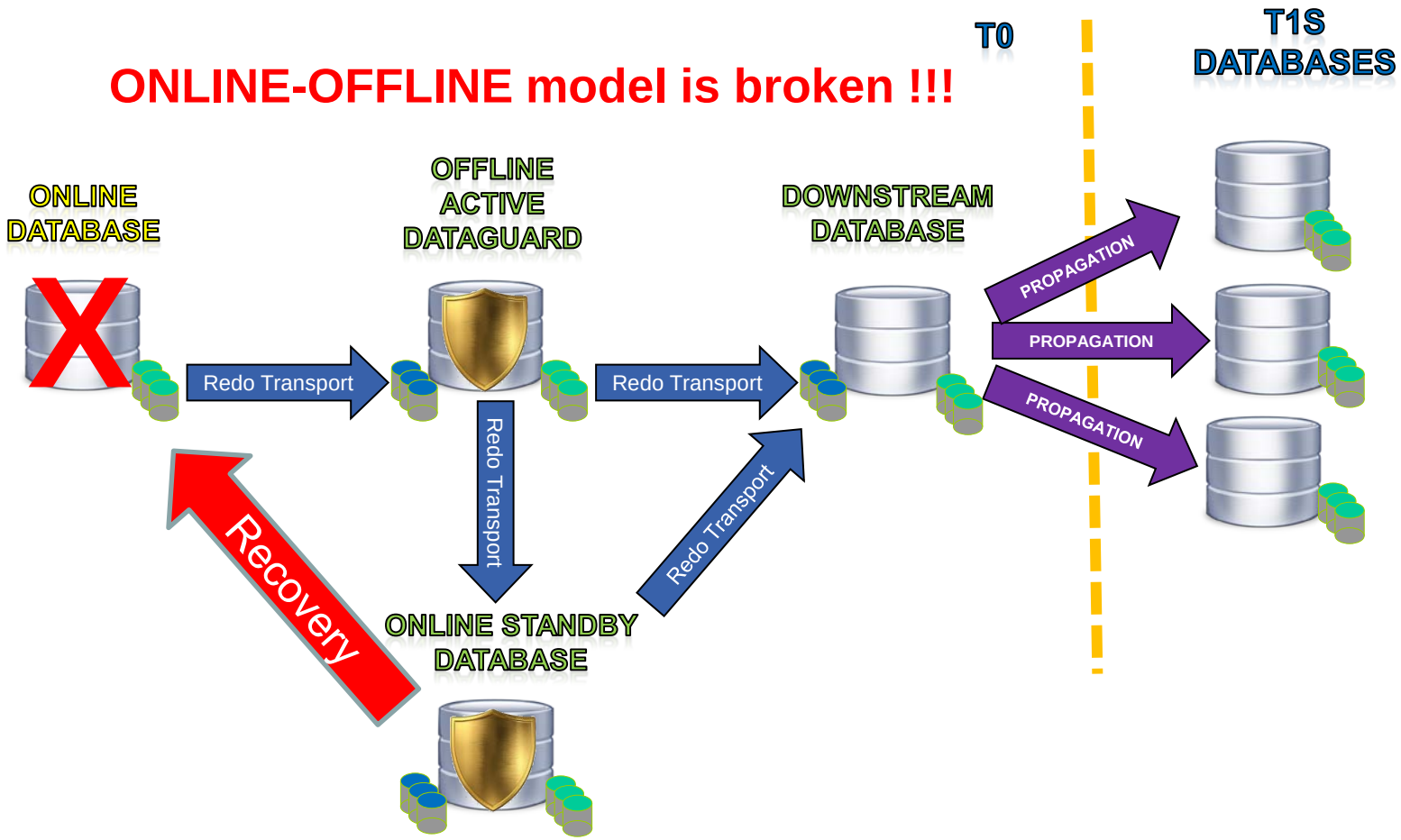
CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Possible solutions

- ONLINE –> OFFLINE: Active DataGuard
- OFFLINE –> T1s: Streams11g

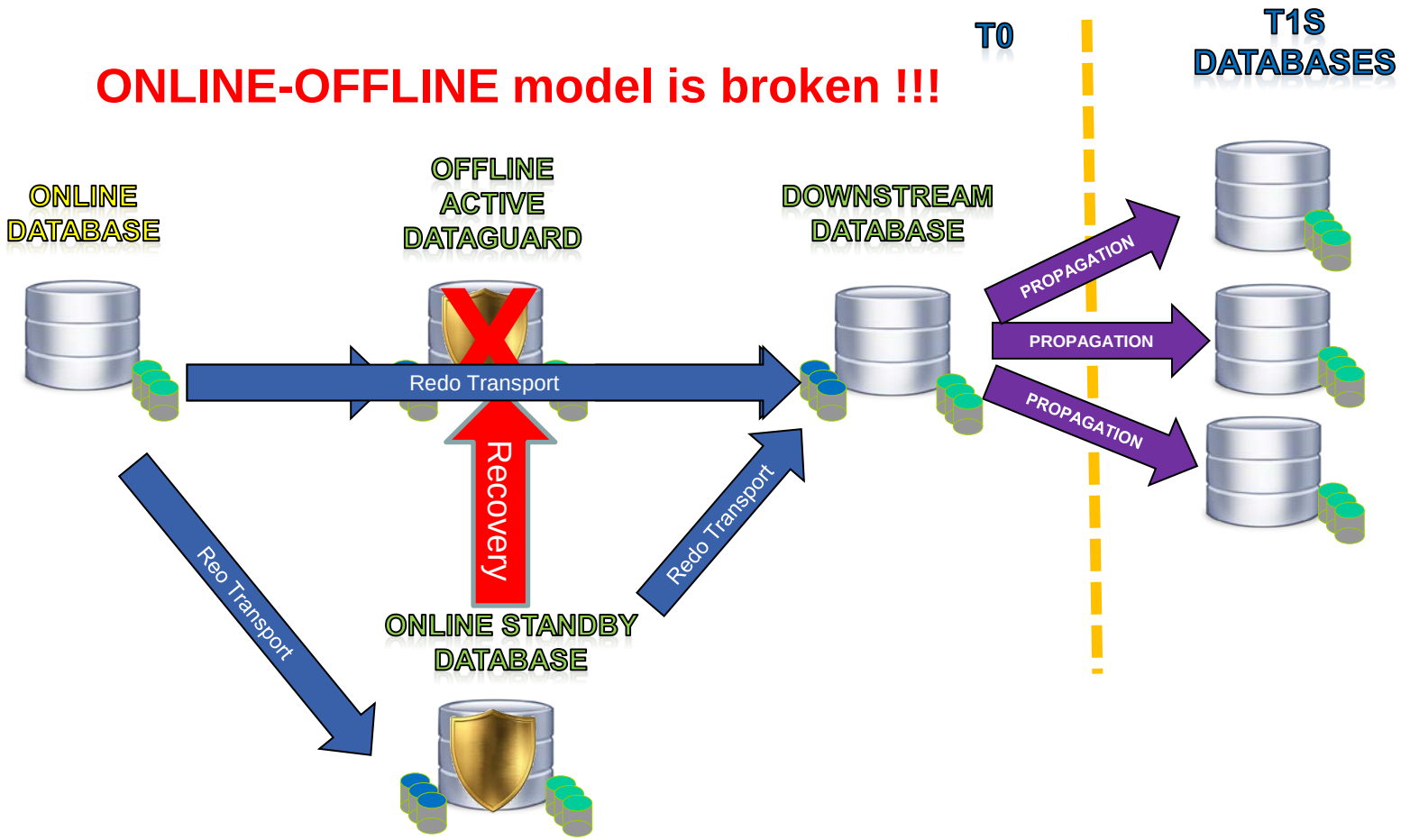


Online database failover and recovery with ADG11gR2



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

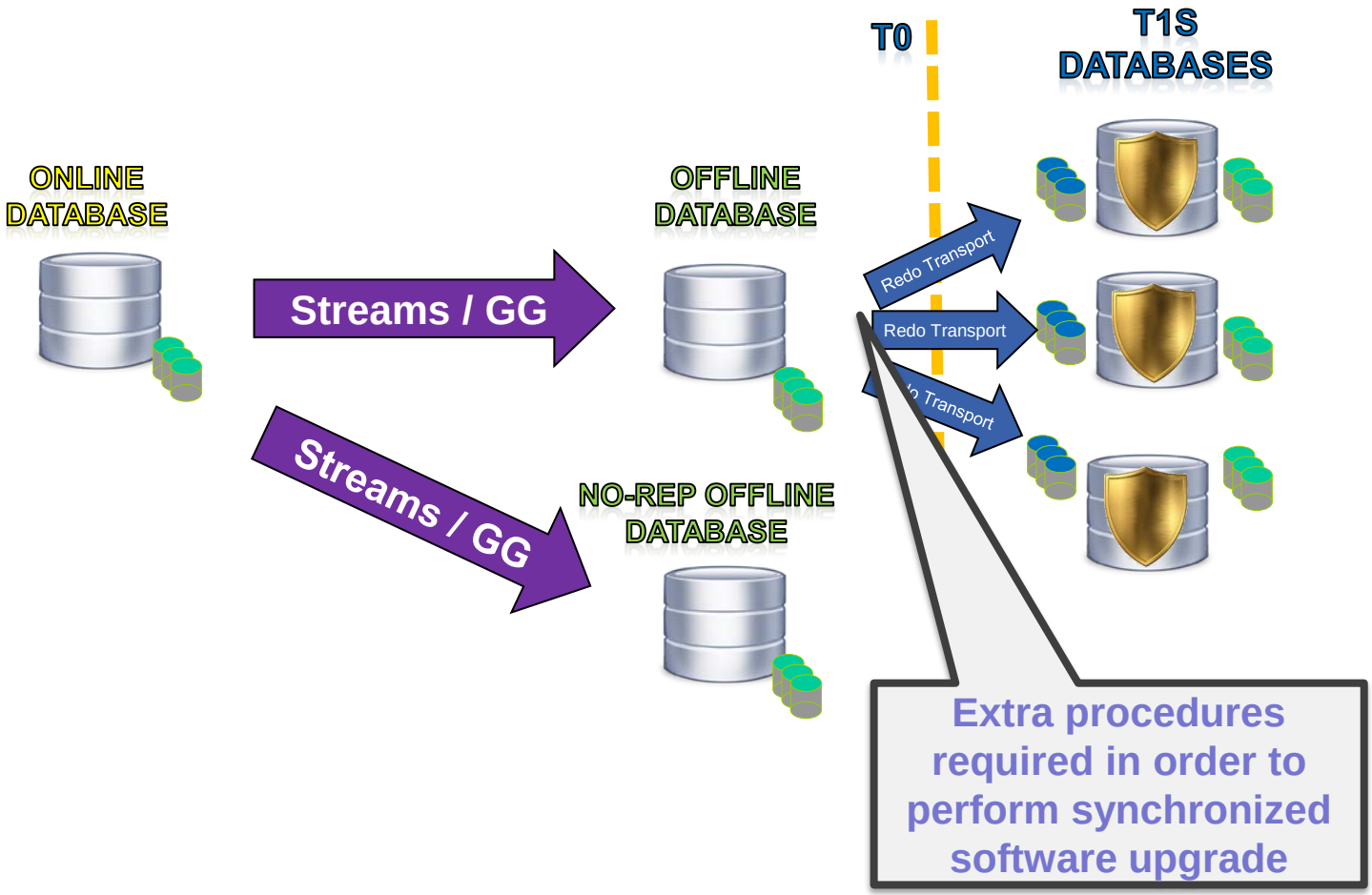
Offline database failover and recovery with ADG11gR2



CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Possible solutions

- ONLINE –> OFFLINE: Streams11g / GG
- OFFLINE –> T1s: Streams11g



Summary

- Migration to the new database versions (2012) gives an opportunity to re-design and improve the replication service
- Three candidate technologies are being investigated
  - Streams11gR 2
  - GoldenGate
  - Active DataGuard

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Acknowledgements

- Many thanks to all Physics DBAs, especially:
  - Luca
  - Jacek
  - Dawid
- Consultancy
  - Gancho
  - Stephen Balousek (Oracle)
  - Jagdev Dhillon (Oracle)

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it

Questions?

CERN IT Department
CH-1211 Geneva 23
Switzerland
www.cern.ch/it