



WLCG – January 2007



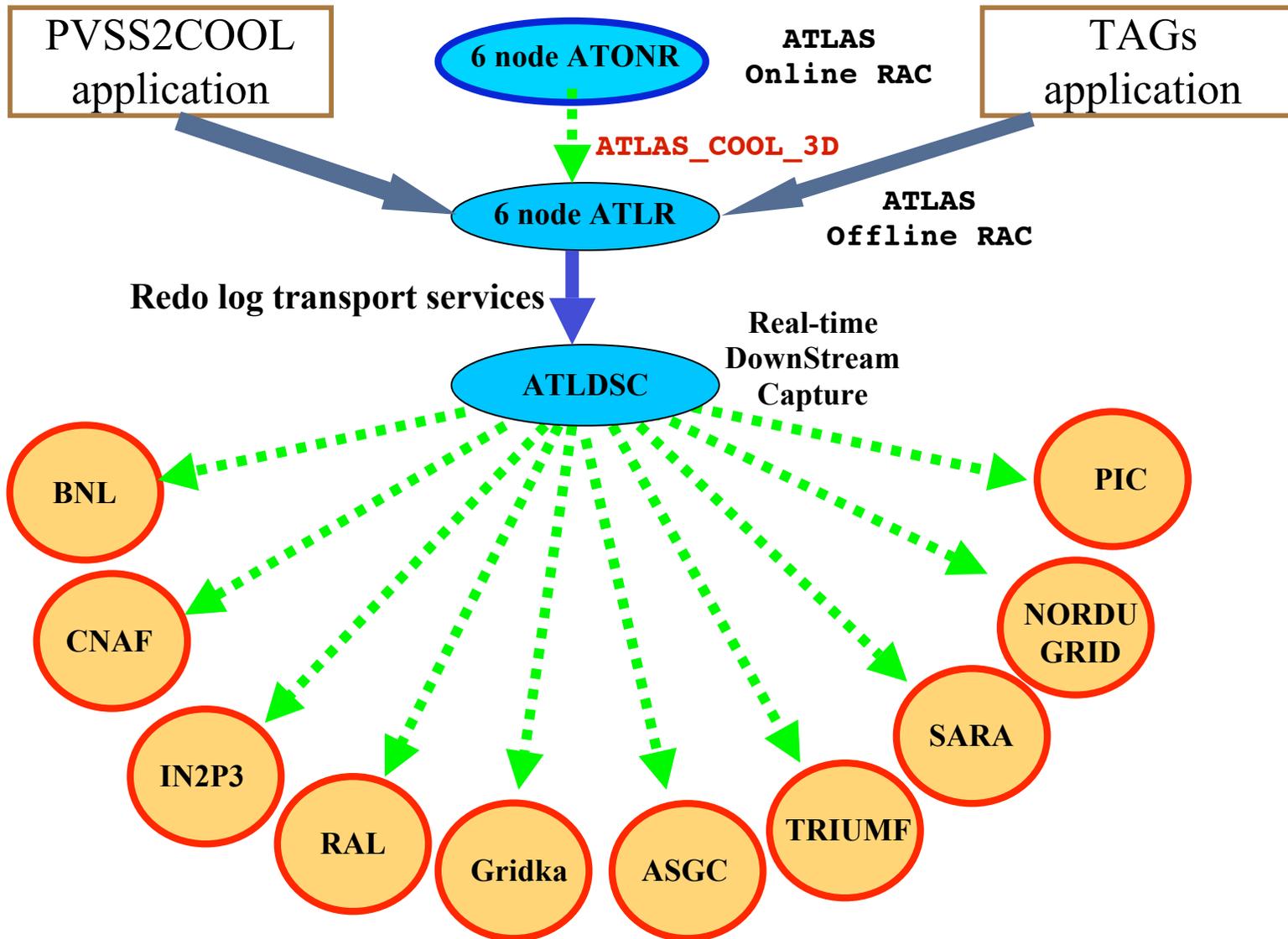
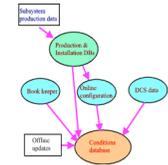
ATLAS COOL/TAGs (online, offline, T1)

Viegas, Florbela

Dimitrov, Gancho

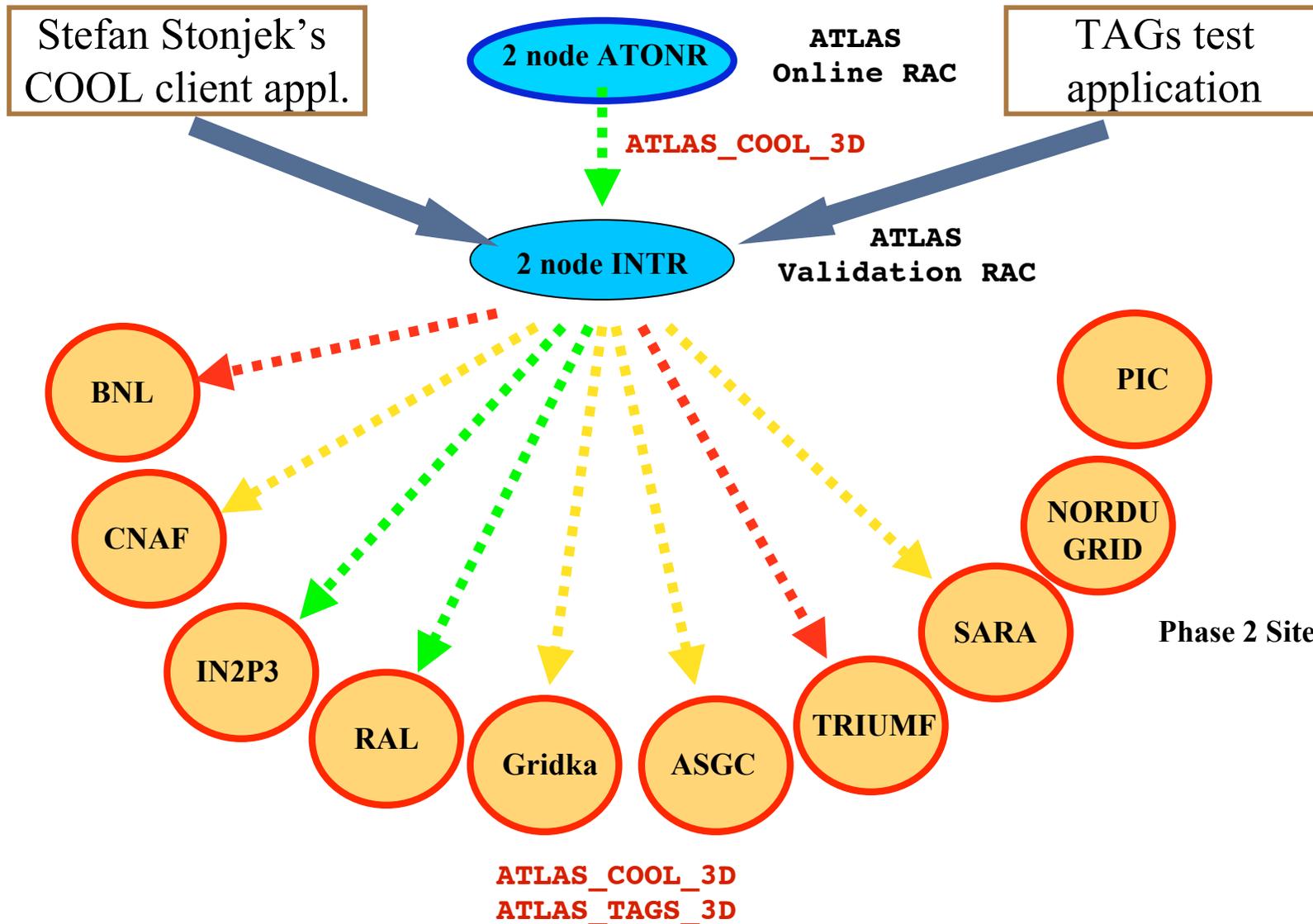
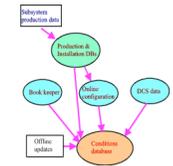


Online Database, T0 and T1 Architecture





Test environment - the status as it was on 18-19.01.2007





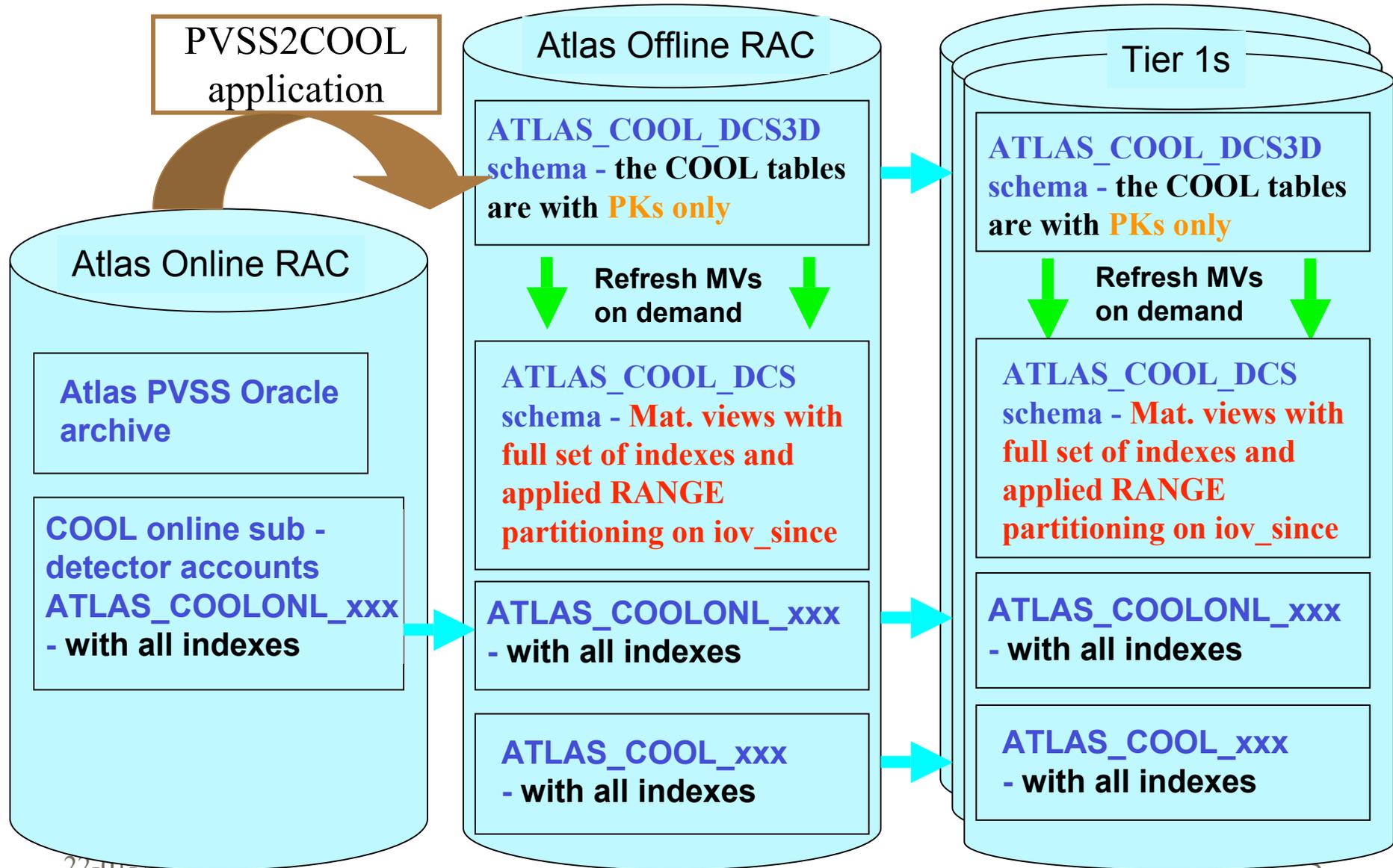
Replication Test Issues



- The tests that have been made, ‘support’(we still do not know what are the requirements for COOL) the throughput requirements on the following T1 sites: IN2P3, GRIDKA, CNAF, RAL, ASCG
- The following sites have throughput problems: TRIUMF and BNL
- Sites, being tested and connected now: SARA
- Sites, not ready yet: NORDUGRID and PIC

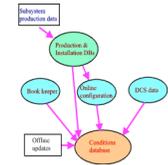


Special COOL setup 'to help' the Streams Apply process

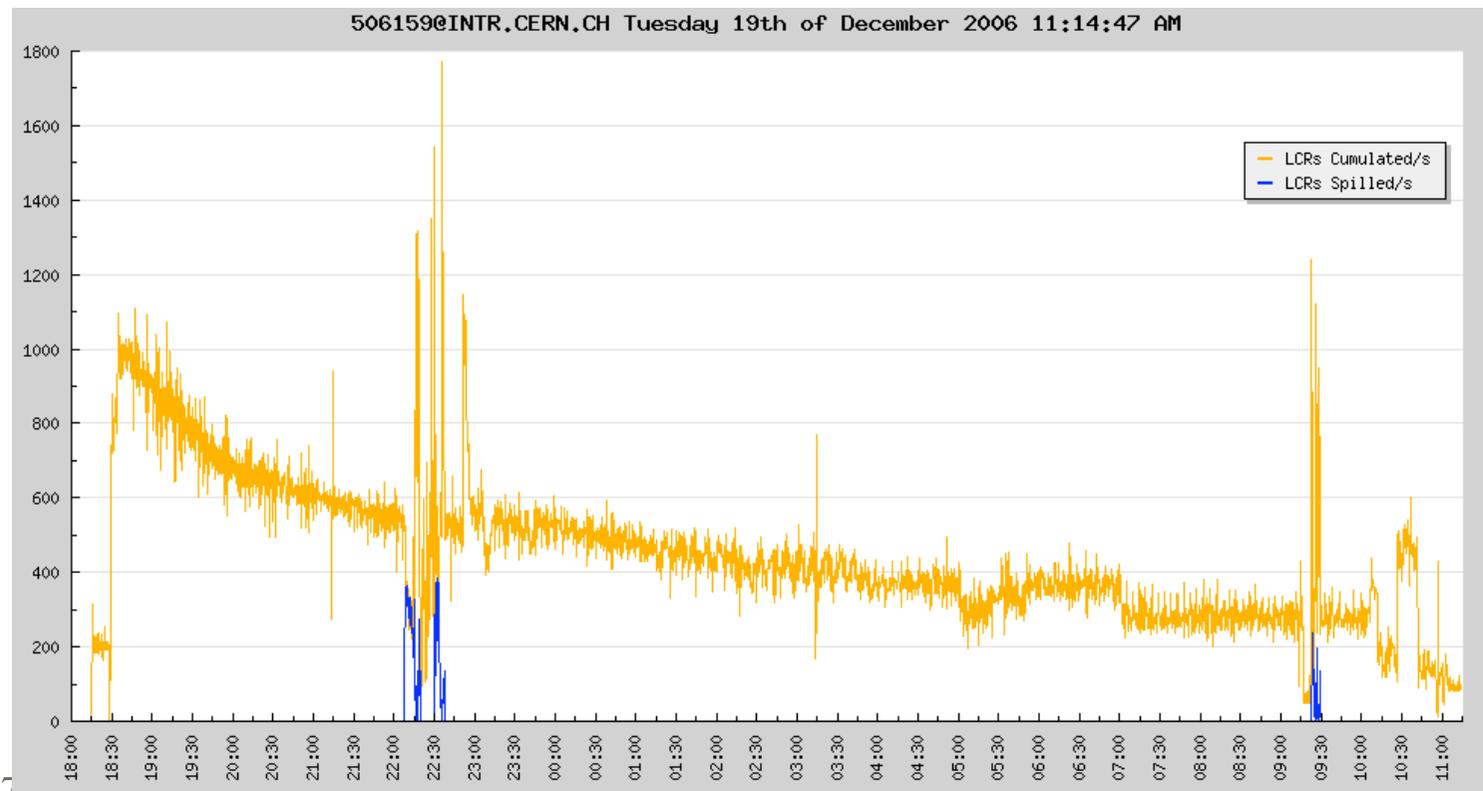




Best result so far ... (got on 19 Dec 2006)



- Participants: INTR (source) and RAL, GridKa, CNAF, IN2P3 (destinations)
- Tags insert rate - 200Hz (row length 1.3KB, means 15 MB per minute).
- COOL rate - 2,2 MB per minute at the beginning, but went down with the time, because of the lack of an index).
- For the test period of 12 hours, the latency was 1-2 seconds, except 2-3 small peaks (the picture below, y axis = LCRs/sec; x axis = timeline)



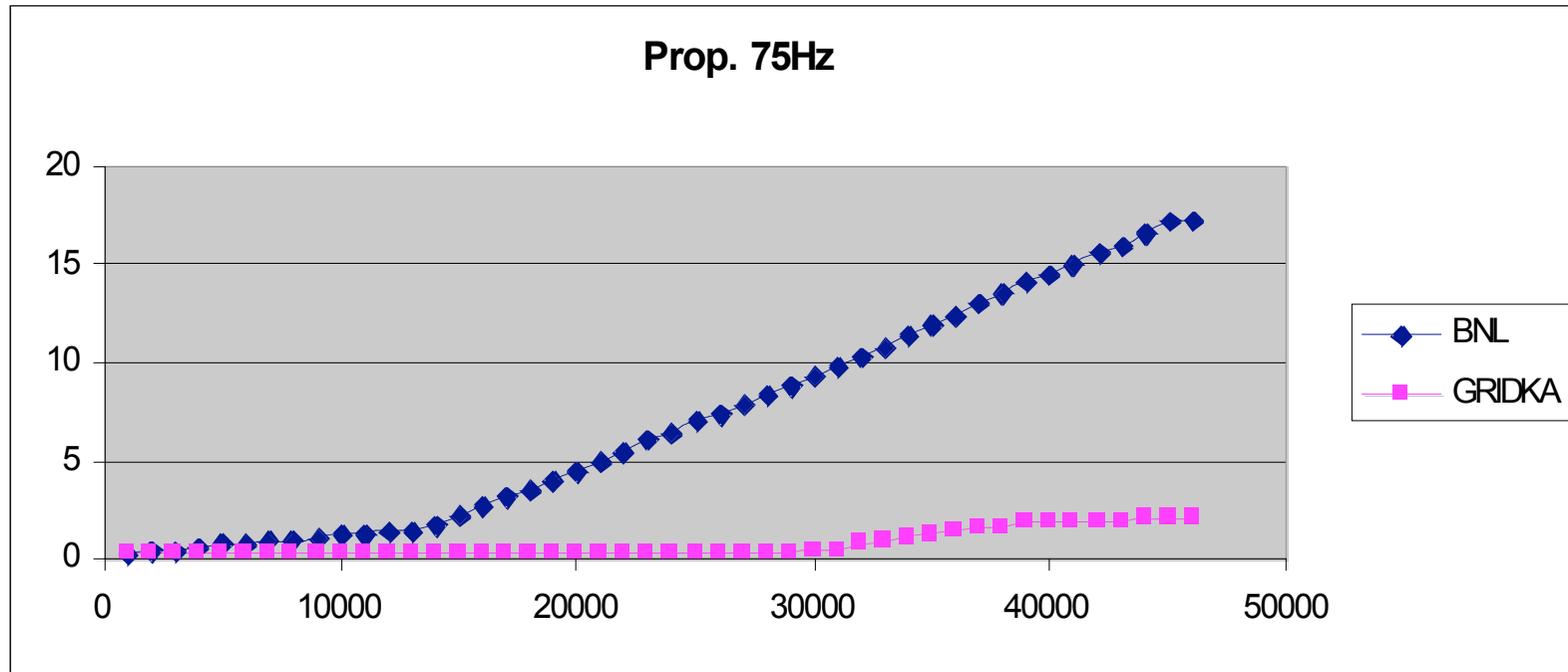
22-01-2007



Non conformity of throughput -BNL



BNL has the following throughput:



The maximum TAGS rate supported is 50Hz, which translates into
 $1.3\text{kB} \times 60 \times 50 = 3.9 \text{ MB/minute}$.

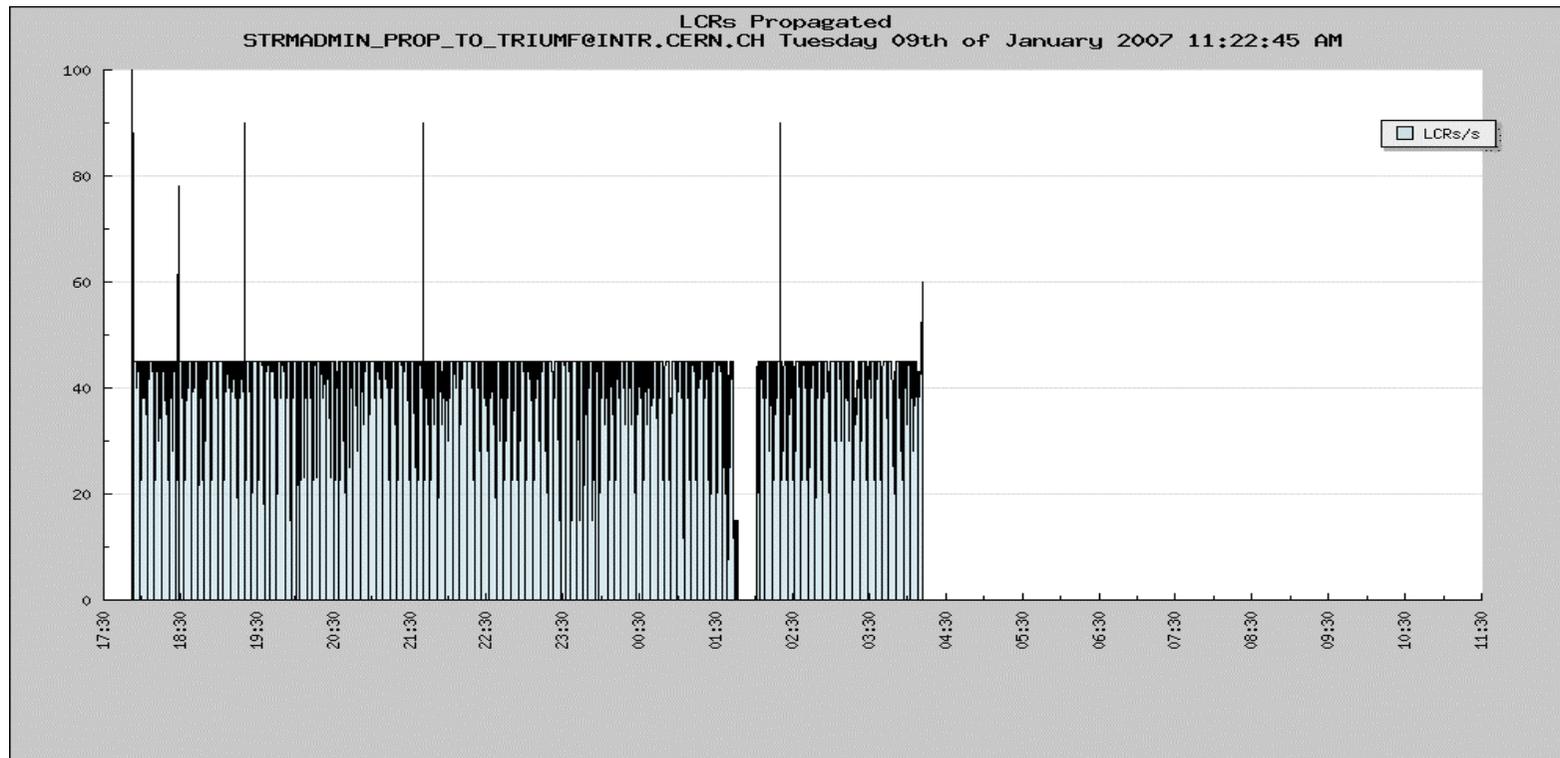
(legend y axis=minutes of delay; x axis number of records sent)



Non conformity of throughput -TRIUMF



TRIUMF has the following throughput:



The maximum TAGS rate supported is 100Hz, which translates into $1.3\text{kB} \times 60 \times 100 = 7.8 \text{ MB/minute}$.

(y axis = LCRs/sec, x axis=timeline)



Issues encountered in tests



- Dropping propagations causes ORA-600 errors, which means that if a site is down and it has to be removed from the streaming current, it can't. We have to drop the whole capture, which might render the destinations out of sync. We seek a better solution for this problem.
- When we put rules for filtering out tables, we find that the capture slows down considerably, even for a simple rule.
- There is a « library cache lock » problem we are experiencing in INTR which is an obstacle for the administration of the propagation schedules
- Capture process lag due to very high production of archive logs in the source database.
- Undo SQL using ROWIDs gets translated at destination with a column by column comparison, which is not usable in production for recovery.



Proposal for optimizing the Capture



- Since the capture gets delayed when there is a large transaction volume on the database, we propose to split the source instances (Online and Offline) into two separate databases
- One has only the applications that are to be streamed, and, so the archive logs have the meaningful transactions only.
- Another is dedicated to non-streamed applications, so its load does not interfere with streams performance.



Agreement of the Proposal from IT for Backup Strategy



- Backup Strategy is the same for Online Database, Offline(T0) Database and T1 sites
- The backup on disk is kept for 48 hours, so recovery can be quick for any point in time in between.
- The backup on tape is kept for 31 days, so recovery can be done for data from 31 days ago. Recovery from tape takes 5 hours in a 300Gb database.
- ATLAS is in agreement, with the following provisions:
 - **UNDO_RETENTION** should be set for the databases to safeguard against errors. We propose 24 hours as a beginning value.
 - Provision has to be made for historical, read-only data, to be stored in a special pool of tapes or disks, out of the retention policy for RMAN.



Plans



- Continue with the “high-rate” streaming tests
- Helping in resolving problems that pop up at Tier 1s
- Demonstrate reliable service over an extended time period. Athena jobs running at Tier-1s, which read data back that has been generated at Tier-0. Tests with real COOL data have already started by Richard and Stefan. (RAL has been chosen as destination)
- Go into production mode at the end of March (when the ATLAS offline release 13 will be available)