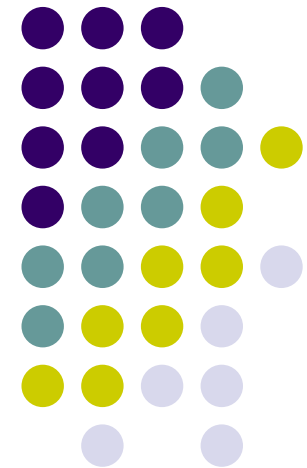


Data Access Patterns from ATLAS

Dietrich Liko IT/PSS



ATLAS activities on the Grid



- MC Production
- (Re)processing
- Distributed Analysis



MC Production

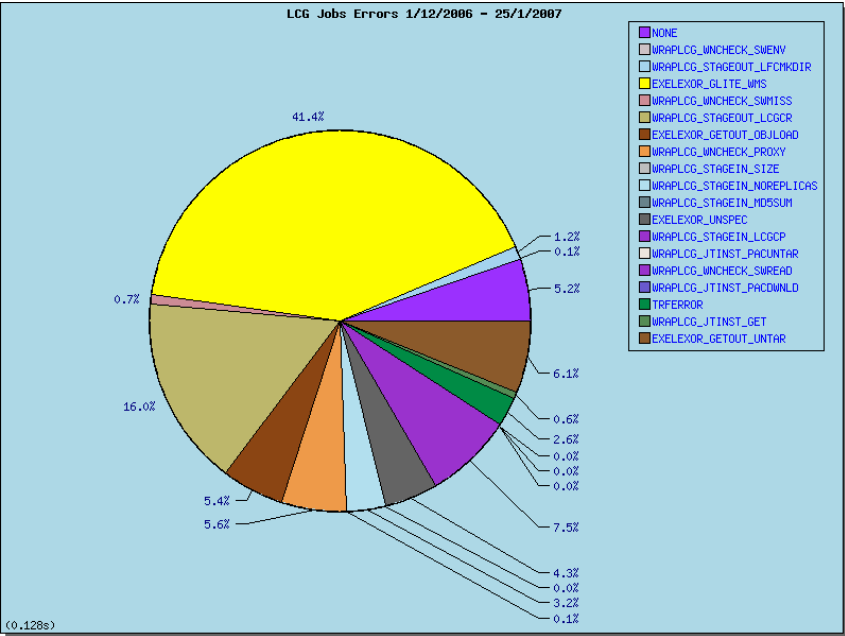
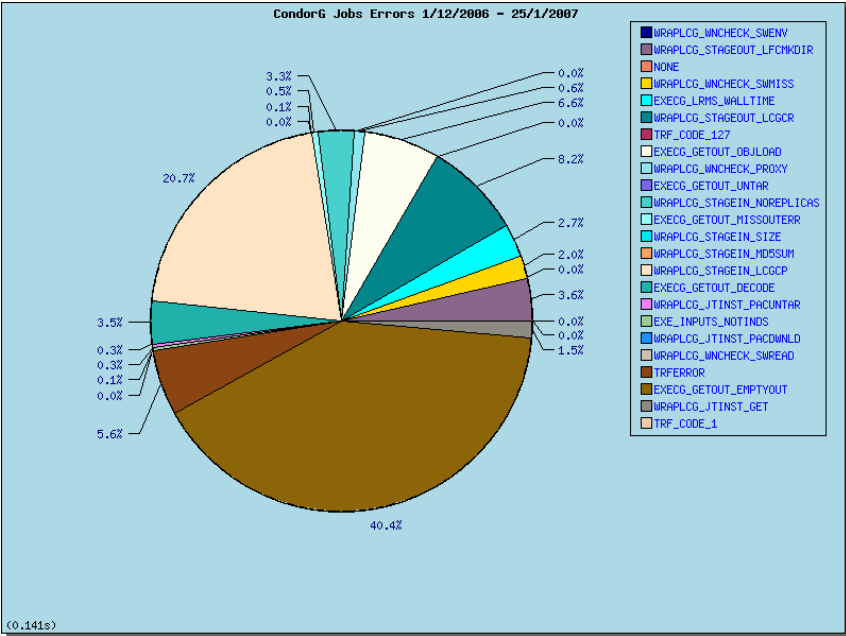
- Jobs are assigned to a Tier-1 cloud
 - Access to data via gridftp
 - Random access to data within the cloud
- Patterns
 - In some cases single files are accessed by many jobs at the same time and induce high load on the server
 - Some jobs download about 10 to 20 files
- This activity has the highest failure rate in ATLAS production after the submission system
- Often also difficulties in storing the results

Failure rates



CondorG

gLite





Failure rates

- Submission system: 40 %
- Stagein: 7.5 to 20.7 %
- Stageout: 8 to 16 %



Failure reasons

- Information system
 - Filecatalog
 - SRM
 - Gridftp
-
- We do not know what is the underlying reason
 - Discussed in the System Monitoring WG (Julia Andreeva)



(Re) processing

- Local access to data stored on tapes
- Essential aspect is the prestaging of data in a meaningful way
- Up to now it has been a more theoretical exercise



Distributed Analysis

- Data is being placed on the grid
- Jobs are send to the data
- We plan to read data only locally

- Its of advantage to use posix-like IO system
 - Jobs might read only part of the event data
 - Jobs might read only a subset of the events



Potential problems

- ROOT IO plugins used by Athena out of data
 - dCache, RFIO, GFAL
 - We have backported some of them
 - Fons promised me to do in the future ...
- Should local access use SRM ????????
- No clear technical understanding from our side
- As of today no technical experience with GFAL



Some other problems

- Problems using new Castor-2 with GFAL
 - srm-durable-atlas
- Wrongly configured sites
 - Access to Castor in IFIC, Valencia
- DPM access lists wrongly configured
 - Use of “Production” role



Systematic tests

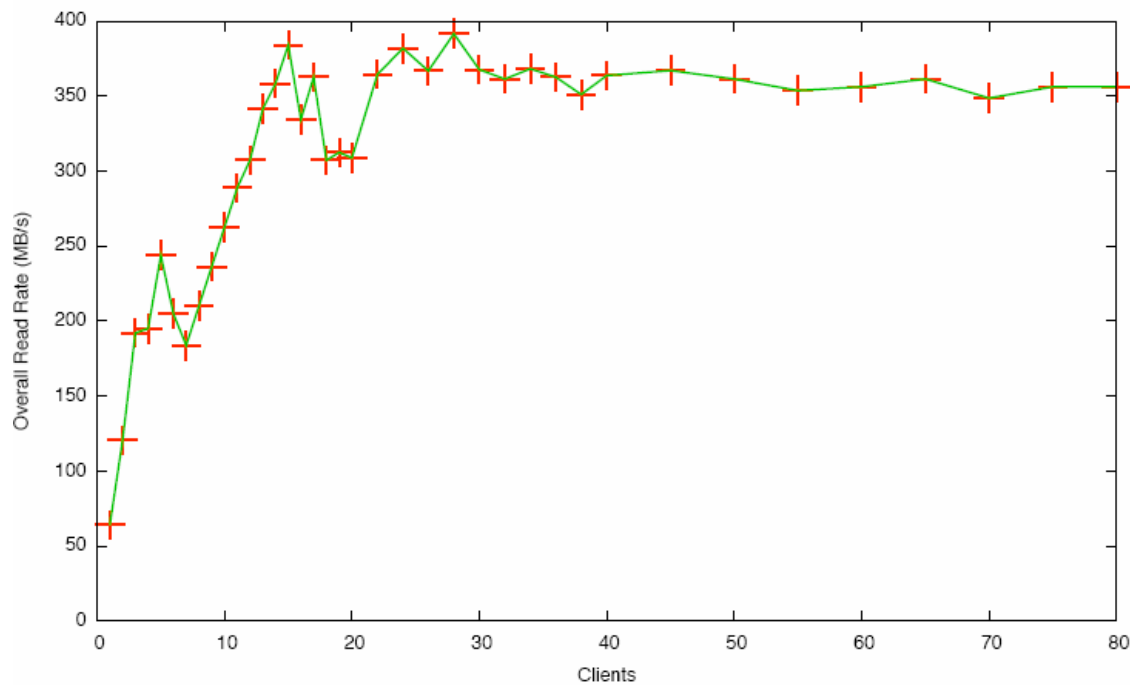
- SAM tests should be extended to test also the Posix IO on SE on LCG
 - GFAL
 - Using the specific protocol (rfio, dcap)



Performance

- First test in Glasglow on a DPM
 - Graeme Stewart & Greig Cowan

overall read rate





More realistic examples

- Athena 2 MB/sec max
- TAG based analysis (only single events are read) will increase the load on the SEs
- Standard Analysis Example
 - 10 files a 130 MB
 - Local 14:02 min
 - DPM 16:30 min
 - Castor-2 20:29 min

Summary



- MC production:
 - We should understand better why stagein and stageout is so error prone
- (Re)-processing
 - We have not yet started
- Distributed Analysis
 - We want to use posix-like IO from analysis program
 - We need to solve all the detail problems
 - We need also to verify the performance of the SE