

Tales From ATLAS DDM Operations

Dimitrios Christidis



UNIVERSITY OF
TEXAS
ARLINGTON

Rucio Works

- Successfully handles the data for ATLAS
 - 470+ PB (1 billion files)
 - 1,5 PB worth of transfers per day (1,8 million files)
- Mature and flexible
- Abstracts the complexity of a diverse infrastructure
- Truly open and community-driven
- Predictable release schedule





You must be *this* tall
to ride the donkey



Operator's Manual

- Lots of experience with operations, but stored in people's heads
- Occasionally, consulting the source code becomes necessary
- What's missing:
 - Operational guidelines
 - Procedures and recipes for reoccurring tasks
- Some challenges
 - Rucio is still evolving rapidly
 - External dependencies (e.g. FTS)
- ATLAS will invest in this
 - No clear timeline yet
 - Community participation?



Scalability

- Deciding on number of servers and daemons
- Dynamic scaling based on work load
 - Kubernetes (see previous talk by T. Beermann)
 - In some cases, the database is the bottleneck
- Scenario: deletion of a moderate number of jumbo datasets (more than 50K files per dataset)
- Spreading over a few days might be necessary
 - But how many?
 - The monitoring might require some interpretation



Further Topics

- Managing the storage space
 - Not a technical problem
- Investigating what happened to a dataset
 - Time-consuming
- Site issues
 - Unavoidable
 - Operational Intelligence
- Site decommissioning
 - A tedious procedure
 - Automation is on the way
- R & D
 - New developments shouldn't affect daily operations
- User support
 - Improved user-oriented documentation would help



Questions & Discussion

