# CTA: CERN Tape Archive
# Adding front-ends and back-ends
# Status report

Vlado Bahyl, German Cancio, Eric Cano, Michael Davis, Cristina Moraru and Steven Murray

# Adding disk front-ends

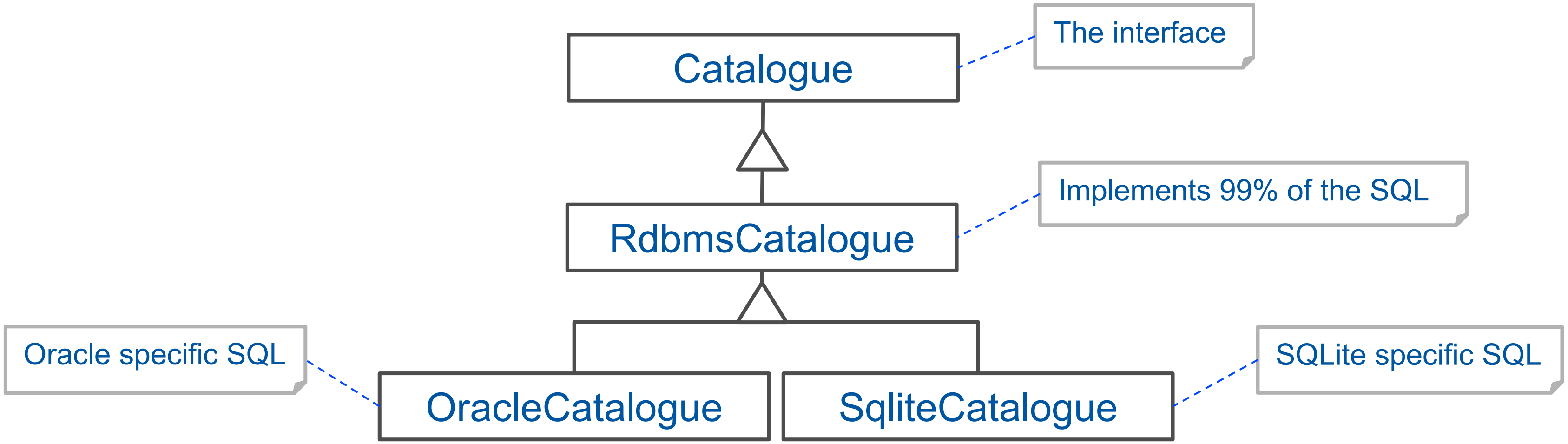
- The disk front-end must:
  - Allow users to tag directories with tape storage classes
  - Allow users to asynchronously request files be brought on-line (retrieved) from tape
  - Allow users to query if a file is on tape (The “m bit” of CASTOR)
  - Automatically send requests to CTA to archive files, retrieve files, delete files and update file metadata
  - Allow CTA to callback when a file is safely stored on tape
  - Distinguish the difference between a user creating a new file and a tape server retrieving an existing file
  - Automatically garbage collect disk files that are safely stored on tape
  - Allow CTA to scan the disk namespace without negatively impacting user operations

# Adding database back-ends part 1

- CTA catalogue avoids vendor lock in
  - No PL/SQL
  - No advanced queuing
- Two differences in RDBMS technologies influenced the design of the CTA catalogue
  - Differences in SQL syntax
  - Differences in programming APIs
- CTA currently uses two RDBMS technologies
  - Oracle for production
  - SQLite for “in memory” unit-tests

# Adding database back-ends part 2

## Design of the CTA catalogue and RDBMS library



Catalogue

The interface

RdbmsCatalogue

Implements 99% of the SQL

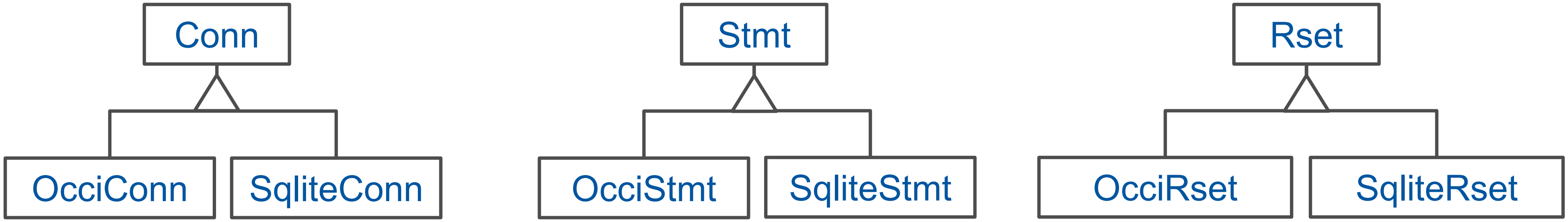
OracleCatalogue

Oracle specific SQL

SqliteCatalogue

SQLite specific SQL

### RDBMS library



Conn

OcciConn

SqliteConn

Stmt

OcciStmt

SqliteStmt

Rset

OcciRset

SqliteRset

# Team status

- Cristina
  - Recommended Access Order for CASTOR and CTA
  - Low-level tape developments
- Michael
  - The EOS/CTA interface
  - The CTA front-end
- Julien
  - Continuous Integration system
  - System test infrastructure
- Eric
  - Main architect
  - Tape server developments
- Steve
  - Project leader
  - CTA catalogue developments
- Anastasia – Joins in September 2017
  - Tape library control software

# Release 0 status

- Archive additional copies of AFS backups
- Vlado is in charge of the EOS/CTA installation
- Performance requirements
  - Average archival rate = 0.3 Hz
  - Average data throughput = 780 MB/s
  - Peak archival rate = 2.4 Hz
  - Peak data throughput = 2.4 GB/s
- Planned to start by 1st June 2017

# Development work status

- Using the CI features of CERN GitLab service
  - Automatic builds, unit-tests and a system test
- Automated system test infrastructure
  - Kubernetes based
  - Can be extended to use real tape hardware
- Currently adding planned performance improvements to CTA
- Currently smoothing some rough edges to help in the deployment of CTA
- Developing RAO for CASTOR and CTA