RAL CASTOR Administration

Shaun de Witt

GDB

2008-09-09
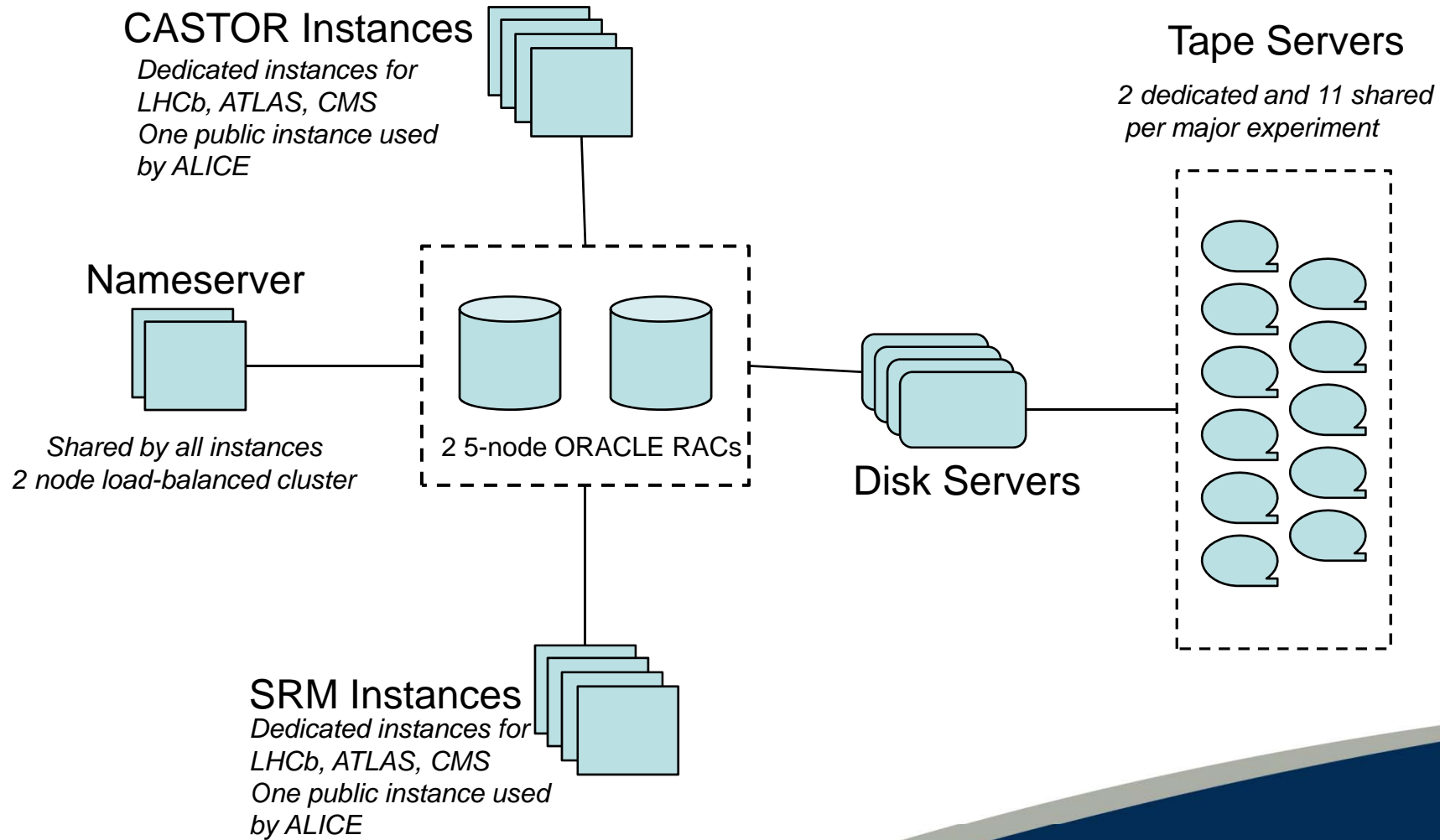
Introduction

- Current Infrastructure
- Problems Accessing Data
- System Administration
- Configuration

Science & Technology Facilities Council
Rutherford Appleton Laboratory

High Level Infrastructure

CASTOR Instances

Dedicated instances for LHCb, ATLAS, CMS One public instance used by ALICE

Tape Servers

2 dedicated and 11 shared per major experiment

Nameserver

Shared by all instances 2 node load-balanced cluster

2 5-node ORACLE RACs

Disk Servers

SRM Instances

Dedicated instances for LHCb, ATLAS, CMS One public instance used by ALICE

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Single Instance Infrastructure (CMS)

3 diskpools, 5 service classes

cmsWanIn, cmsWaninTest
cmsWanOut, cmsWanOutTest
cmsFarmRead

3 CASTOR head nodes

Nameserver

ORACLE RAC

SRM

2 load balanced front ends
1 back end

Science & Technology Facilities Council
Rutherford Appleton Laboratory
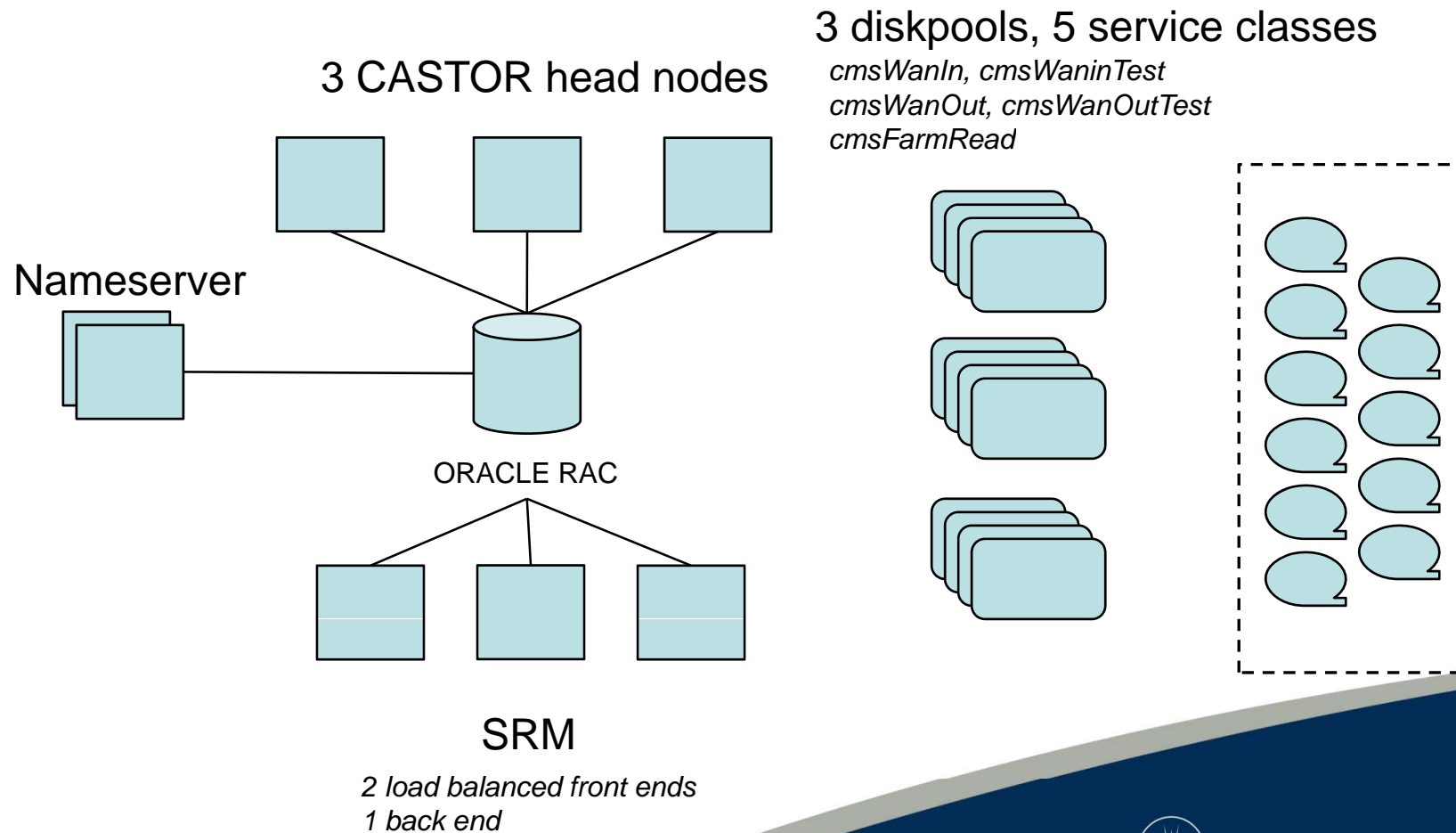
Monitoring

- **Nagios alerts**
  - Alerts for problems on SE and CE
  - Over 350 alerts graded by severity
- **SURE System**
  - Callouts generated on subset of alerts
    - about 175 Nagios alerts generate call-outs
  - Notifies Primary On Call
  - Primary can all on other experts
    - CASTOR on Call
    - DBA On call

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Problems Accessing Data

- Common problems:
  - Tape drives being marked down
    - CASTOR very sensitive to tape problems
    - Leads to being unable to recall files in a timely manner
  - Database slow-downs
    - Sometimes database execution plan needs a 'nudge' to follow right path
    - Leads to timeouts in srm and stager

Science & Technology Facilities Council
Rutherford Appleton Laboratory

More problems accessing data

- Less Common Problems
  - SRM back end failures
    - Infrequent (about 1 every 2 months)
    - On call out, problems resolved quickly.
  - DiskServer Hardware problems
    - If Disk1Tape0, files are temporarily unavailable
    - SRM does not report this correctly to user.
    - Typically fixed in < 2 days

Science & Technology Facilities Council
Rutherford Appleton Laboratory

System Administration Issues

- Non-trivial to add new diskservers
  - recently largely overcome by use of "Puppet"
- Upgrading can causes unexpected problems
  - e.g. 'checkreplicas' script broken when upgrading to 2.1.7
  - Other internal RAL scripts require modification after upgrade
  - Some Nagios checks break

Science & Technology Facilities Council
Rutherford Appleton Laboratory

More Sys Admin Issues

- Identifying causes of problems
  - Lack of 'end-to-end' unique identifier between middleware components
  - Some error messages are somewhat cryptic
- Identifying tapes that need repacking
  - Tool to look for sparse tapes would be useful
- Automated 'draining' tools
  - Tool to automatically copy out files from draining disk servers

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Configuration and Tuning

- Setting up job slots per disk server critical
  - Too many slots leads to heavy I/O load, and reduced throughput
  - Too few means we can't meet MoU agreement
  - Depending on use of service class, job slots need to be tuned differently
    - Slots on WAN servers << Slots on LAN

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Configuration and Tuning

- Tuning FTS to each instance is important
  - Failure to throttle can lead to overload in CASTOR
  - Results in timeouts reported to user, but jobs can still execute
  - Would be nice to be able to configure channels by space token…

Science & Technology Facilities Council
Rutherford Appleton Laboratory

Improving Stability…

- Moved to ORACLE RAC
  - Automated database fallover in case of hardware failure

- Load balancing
  - Load balanced nameserver and SRM front end

- Increasing number of monitoring scripts
  - more experience, more to watch!

Science & Technology Facilities Council
Rutherford Appleton Laboratory