



# **CMS Castor2 Experience**

*On behalf of CMS Computing*

**Nick J. Sinanis**  
**CERN**

**Castor Review - CERN**  
June 8th, 2006



# Castor2 for CMS

- **January '05 to Summer '05**
  - ◆ First Castor2 tests with PhEDEx
  
- **Summer '05 to December '05**
  - ◆ First tests for SC3 (Lassi Tuura)
    - Lots of problems
  - ◆ Production (Nikolay Darmenov)
    - Smooth when not hitting server problems
  - ◆ Transfers & Production - the 1<sup>st</sup> migrated activity
  
- **December '05 to March '06**
  - ◆ Towards full migration
    - Had to understand the issues of
      - Applications (@cern and not at @cern)
      - Configuration



# Castor2 Migration

- **Mid-March**
  - ◆ First CMS DS in place
    - WAN Pool
    - CMSPROD (relocated from castorgridsc)
  
- **Migration was finalized 2<sup>nd</sup> of May**
  - ◆ Most delicate part migrating all default accesses
    - Env Switch
  - ◆ Delayed not to interfere with Physics TDR II
  - ◆ Some problems with
    - Libshift
    - Castor Env for Grid Jobs



## Castor2 Migration (II)

- **Strong and early buy-in from CMS**
  - ◆ Started testing at the earliest possibility
  - ◆ Both sides learnt a lot, even at a cost of time
  - ◆ Had to coordinate the Castor2 migration while
    - Having a changing Software environment
    - intense PTDR activities were taking place
      - Running old Software
    - Had to respond to urgent production and transfer requests for the PTDR
  
- **Yet, we lack the full experience that could make us fully confident**
  - ◆ More to come on this soon...



# Castor2 Issues

- ◆ **Configuration**
- ◆ **Performance and Scalability**
- ◆ **Operations**
- ◆ **Support**



# Configuration

- **Libshift.so**
  - ◆ **New ABI SO versioning policy requires change management!**
    - Was easier with Castor1
    - Castor2 is expected to release libshift versions more frequently
  
  - ◆ **This raises important issues as:**
    - Forward compatibility
    - How patches are handled?
    - Validation by experiments?
    - Effective distribution to sites?
    - Part of the OS or like any other external?



# Performance and Scalability

- **Our sole experience is with SC's**
  - ◆ Numbers look good so far
    - It was an enormous amount of work
  - ◆ They are not representative though
    - No opportunity to confirm SC3 rerun numbers
    - Neither T0 nor CAF access patterns have been tried yet!
  
- **CMS *real test* of Castor2 is CSA '06 !**



## CSA '06

- **Computing, Software & Analysis Challenge 2006**
  - ◆ Targeted for October and for one month
    - **Early start mid-September**
  - ◆ 20-40 Hz exercise from Tier-0 down to Tier-2
  - ◆ Simulate DAQ with 50Mevts on Tier-0 disk
  - ◆ Run Calibration jobs
  - ◆ Run Tier-0 PR
  - ◆ Write to Tape
  - ◆ Ship Full Event to Tier-1s
    - **300 MB/s from Tier-0 to Tier-1s**
  
- **CSA '06 Demonstrates the Workflow**
  - ◆ SC4 is a validation step for CSA '06





# Operations

## ■ Service Side

- ◆ Several reconfigurations required so far
  - Maturity?
  - Or just flexibility?
  - Can Tier-1's deploying Castor2 follow all this?
- ◆ Many new components
  - Usually takes time to understand
  - Usually users manage to break things
- ◆ Validation of Significant Service Reconfigurations
  - Need to be part of the experiment's validation
    - Aim for this at the start up



# Operations (II)

- **User Side**
  - ◆ This mainly means “Special Users”
  - ◆ CSA '06 is planned as an operations aware exercise
    - Will need to have effective channels for this
  
- **Interface to Large Scale DM Tools**
  - ◆ Administration tools
  - ◆ Monitoring is crucial
  - ◆ System level monitoring is good and very detailed
    - But not what we actually need!



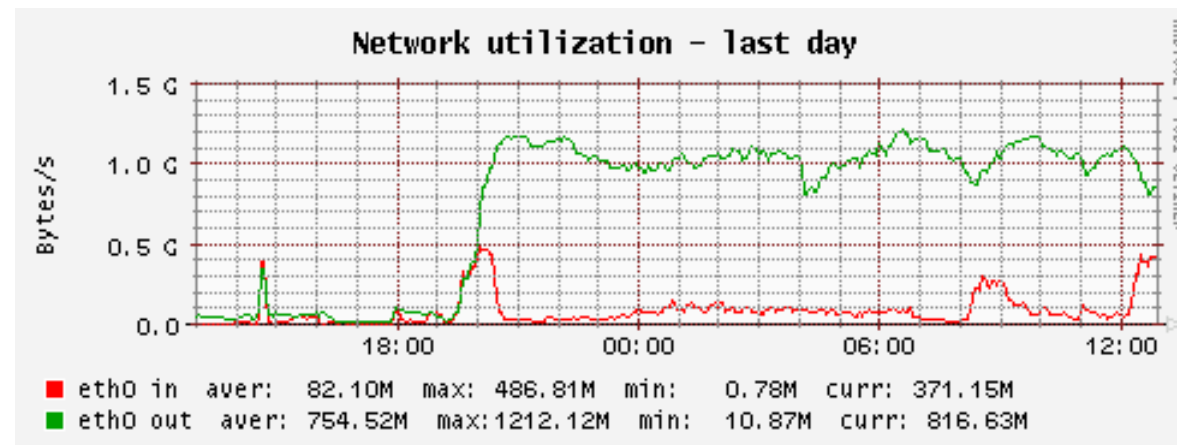
# Service Monitoring

- **We need to couple monitoring to**
  - ◆ User application (CAF)
  - ◆ Tier-0
    - Merge and Export buffers
  
- **But also to understand**
  - ◆ Disk retention and GC policies
  - ◆ File replication
  - ◆ File pinning
  - ◆ Staging latencies
  - ◆ Active Users (top 20)



# Monitoring Example

- **What is all this traffic?**
  - ◆ CMS \_default cluster yesterday





# SRM

- **SRM is a core strategy for CMS**
  - ◆ Used since spring 2004 (DC04)
  - ◆ Part of C-TDR baseline
  - ◆ Practical interoperability critical
    - V1, V2 and recently agreed extensions
    - "SRM-as-used-by-X" insufficient
  
- **Vast majority of our Castor-2 problems had to do with Castor/SRM or interoperability**
  - ◆ SRM request corruption and state confusion
  - ◆ Lack of "safe-for-everyday-use" `srmCopy()`
  - ◆ Inability to delete files (advisory delete)
  
- **We lack confidence in the SRM side**



# Support

- **It has been almost immediate!**
  - ◆ Evident effort to dig into rather complex problems
    - e.g., SRM problems
  - ◆ Many thanks go to Olof, Jan et al., behind the scenes
  
- **Support People in the middle of**
  - ◆ Systems
  - ◆ Users
  - ◆ Policies, etc.
  
- **Is it sufficient?**
  - ◆ Aren't we quickly burning precious resources?
  - ◆ Will support level scale as we approach real experiment tests and startup?
  
- **What happened to the user manual?**



# Summary

- **Strong buy-in from CMS**
- **Successful migration**
- **Pending confirmation of performance and scalability in CSA'06**
- **Support has been fine**
- **Inevitably we need to see the big picture**
  - ◆ **SRM**
  - ◆ **Service Monitoring tools**



# Questions