



Report on Findings of the CASTOR2 Readiness Review

LCG Milestone : CAS-1 15.03.06

Review Committee:

- | | |
|---------------|------------------------------|
| Miguel Branco | - CERN PH/ATC |
| John Harvey | - CERN PH/SFT (chair) |
| Don Petravick | - FNAL CD-COMP & COMP FABRIC |
| Shaun de Witt | - RAL e-Science Centre |



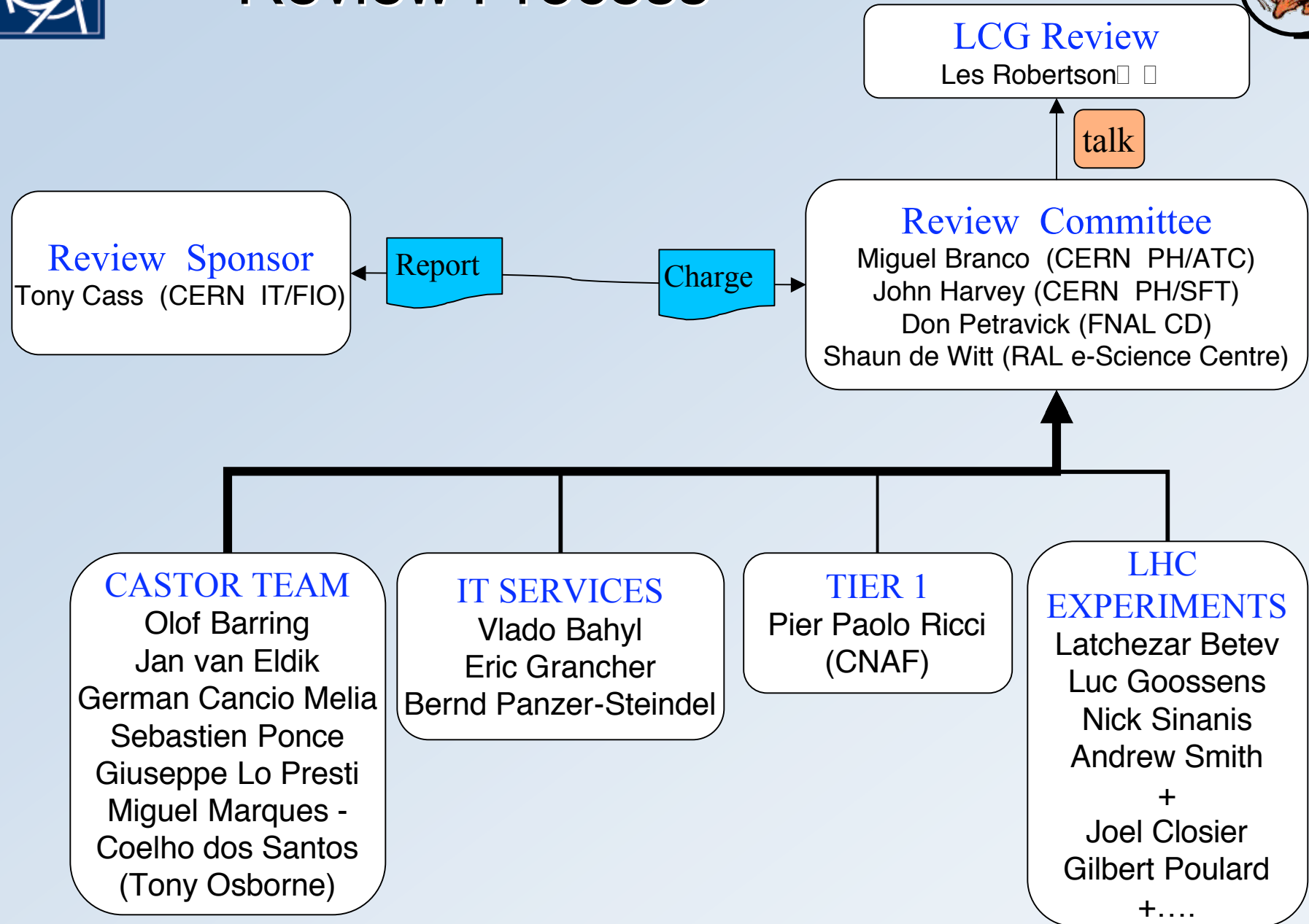
Outline



- ❖ Introduction
- ❖ Summary of findings and Recommendations
- ❖ Concluding Remarks



Review Process





Charge Assess the readiness of the Castor2 service for LHC production



- ❖ Does the software have the required functionality? If not, is the functionality at least adequate? And what are the priorities for development?
- ❖ How mature/stable is the software? And the software development process?
- ❖ Is the service infrastructure, and particularly the database server hardware, appropriate?
- ❖ Are the necessary operational procedures in place (with adequate trained personnel) to ensure the MoU service reliability commitments can be met?
- ❖ Does the overall service have the required performance for the Tier0? For the CAF? What aspects of the software or system limit performance? Are these performance limits a concern?
- ❖ Is the current policy for the support of Castor at other sites sustainable?



Introduction



- ❖ The scope of the review is core CASTOR project. The scope of the review does not extend to the SRM interface.



Overview Comments



- ❖ Overall the committee believes that the project is on the right track, that the structure of the software shows promise.
- ❖ With a vigorous development team, and careful planning, we can see development and deployment at CERN keeping up to meet the complete requirements of the LHC experiments.
- ❖ We can foresee many years continuous work in the system.
- ❖ We can foresee periods of operational distress associated with this process.



Functionality, Priorities (1)



- ❖ CASTOR 2 is an upgrade project to CASTOR-1
- ❖ One goal of the project is to make the daemon code stateless, so that daemons may be stopped and started, and also to scale the system by making the daemons replicable.
- ❖ The stateless architecture is commonly used in high availability systems within industry.
- ❖ However, the stateless architecture does require a great deal of effort in testing and development.
- ❖ ***The chosen architecture induces an intrinsic reliance on performance of the relational database,*** since it holds all significant state.



Functionality, Priorities (2)



- ❖ It seems to the reviewers, LSF may become a major performance bottleneck on CASTOR-2, particularly the response time when opening a file.
 - **The committee recommends that flexibility and loose coupling to LSF should be kept**
- ❖ Evaluation of the experiments' interests and motivations for XROOTD must take place.
- ❖ While out of scope
 - The project is not able to maintain its SRM V1 interface.
- ❖ Service-level monitoring for many aspects of the service need to be completed and exposed to the experiments.
 - Information that will allow problem resolution for the services relying on CASTOR. e..g. what transfers are pending, which are active, etc.



Functionality, Priorities (3)



- ❖ The experiments have raised important issues concerning the deployment of new releases of the CASTOR client library, Libshift.
 - The experiments therefore request to have control on the version that is used for building the software.
 - the decision was taken to install Libshift in the External Software Service area managed by the LCG Applications Area along with other external libraries and this has now been done.
 - It is important that this procedure is followed in the future.



Functionality, Priorities (4)



❖ Common code

- The project makes use of the DPM code base for development of the Nameserver although there is no common DPM/CASTOR-2 CVS repository yet.
- Two implementations of the rfio library also exist and this is causing some confusion for users.

❖ Project should work toward common code library.



Service Infrastructure and hardware(1)



- ❖ The project has a critical dependency on ORACLE performance.
 - “Anything but the best database hardware...is a significant risk to the operations of CASTOR.”
 - N.b many CASTOR instances are planned for CERN.
 - Adequate performance induces reliance on ORACLE-specific features.
 - Requires specific ORACLE skills.
 - Project needed much more time from skilled ORACLE DBAs than first anticipated.
 - **Project requires continued access to such a skilled person to assure its success.**
 - N.B. The project has had to rely on its own heuristics to estimate performance requirements. (opens/TB)



Service Infrastructure and hardware(2)



- ❖ Quattor and Lemon are an aid to CASTOR operators and maintainers.
 - There are still many tasks which must be done by hand.
 - Effort will need to go into codifying fault recovery scripts, which are in the Lemon framework.
- ❖ Significant effort has gone into finding workarounds to problems whenever feasible rather than fixing bugs.
 - The priority of the development team has been given to meeting requests of the LHC experiments.
 - **The committee recommends that the development and operations teams urgently review the list of workarounds and now give full attention to improving the reliability and robustness of the CASTOR software.**



Service Infrastructure and hardware(3)



- ❖ The committee recommends that the project reviews and streamlines use of the various mechanisms (mailing lists, Remedy, Savannah) for reporting and following-up on problems.
- ❖ The committee recommends that the project organizes a meeting with the 4 LHC experiments in order to reach agreement on a set of requirements and priorities for Service Monitoring Tools and a plan for their implementation.



Performance



- ❖ See prior comments about ORACLE and LSF.
- ❖ Internal benchmarks exist indicating that 80 new files/sec could be handled.
- ❖ CERN will deploy a separate CASTOR instance for each LHC experiment.
- ❖ **The committee recommends that the project should aim to validate experimentally the scalability of the performance of CASTOR2 to ensure that it will meet the requirements of the LHC experiments under sustained highest operational loads.**



Support of CASTOR at other sites



- ❖ External sites contributing to castor development, enabling local expertise at remote sites, is a good model.
- ❖ The model for sustaining this collaboration envisages a community effort providing mutual support using an external operations mailing list.
- ❖ **The committee recommends that the project should aim to clarify the policy for the support of CASTOR with its collaborators with a view to specifying the minimal effort and expertise required to operate a reliable CASTOR service and the level of support CERN can offer.**
 - Note comments about ORACLE servers, hardware.



Manpower and Planning



❖ Manpower

- The project has recently undergone a difficult transition with several developers and operations staff, leaving.
- The loss of effort has been compensated by the recruitment of replacements who are currently gaining experience.
- The project's engineering practices will help it cope with turnover.
- The project needs assigned ORACLE experience.
- The project could benefit from dedicate project planning effort.

❖ Planning

- Only recently has the project made a WBS and schedule.
- The schedule has not existed long enough to establish its predictive power.

❖ **The committee recommends that the project management organize a 'delta review' in 6 months to review the project plan and progress made.**



Future directions



- ❖ The committee believes the CASTOR project should be the main effort for storage at CERN.
 - With a vigorous development team, and careful planning, we can see development and deployment at CERN, keeping up to meet the complete requirements of the LHC experiments.
 - We can foresee many years continuous work in the system.



Concluding Remarks



- ❖ The committee would like to once again thank all the participants to the review for all their hard work and for all the help that has been given to the committee members.
- ❖ The committee will deliver its final report to its sponsors in one week.