

Summary of the WLCG Agreements on *xrootd* Support for ALICE

This note specifies the support for *xrootd* that has been agreed between ALICE and WLCG and states my understanding of the support agreed by third parties.

The *xrootd* system

The *xrootd* system has been developed by a collaboration between SLAC and INFN (M.Morandin, Padova section), in the context of the BaBar computing model evolution called CM2 (2002-2003). The project leader was P.Elmer (SLAC) with A. Hanushevsky as technical leader and A. Dorigo and F. Furano as main core developers. CERN has provided a significant contribution to *xrootd* via the ROOT project (PH/SFT, in particular G.Ganis). Within WLCG, *xrootd* is used by ALICE at CERN and at other WLCG sites. F. Furano, has recently joined the IT/PSS/ED section to support ALICE including its use of *xrootd*. At the moment this position is funded by the EGEE project until September 2008, with a possible further extension to 2010 depending on funding becoming available.

F.Carminati, on behalf of the ALICE Computing Project, has discussed the question of support of *xrootd* with R. Mount, Computing Director at SLAC. R. Mount stated that the SLAC science program is heavily dependent on *xrootd* and SLAC can therefore assure that *xrootd* will be maintained by SLAC for at least five years. *xrootd* is an open source product and its code will remain freely available. In view of the fact that *xrootd* brings valuable and currently unique capabilities in scalable high-volume data analysis, SLAC considers as part of its mission to encourage wide use of developments like *xrootd* where they can benefit national and international science programs.

SLAC staff supporting *xrootd* will be encouraged to examine, as time permits, problems and suggestions submitted by users who are not connected with the SLAC program. Non-trivial work in response to such submissions would require that work be also beneficial to SLAC's use of *xrootd*. ALICE is very comfortable with this model of collaboration, that is very similar to the one successfully adopted since ten years with the ROOT project.

The ALICE Computing Project and the SLAC Computing Centre are particularly in favour of a collaborative approach to maintaining and developing products like *xrootd*. This approach promotes wide use, and creates a situation where effort spent on issues raised by a collaborator could also be considered valuable to SLAC. F. Furano will collaborate closely with the *xrootd* developers with the primary aim of achieving the goals of CERN and ALICE. However he might work on general *xrootd* issues if this is compatible with his primary commitments.

Castor

CERN/IT provides support for the integration in Castor of the *xrootd* plug-in developed by A.Hanushevsky, and for its operation at CERN for the ALICE experiment. The support initially is for read-only data, within the performance limitations of the current implementation, planned to provide access from the ALICE PROOF facility to Castor disk pools. The system is available at CERN but has not yet been widely used by ALICE. ALICE will also test the functionality and performance of the plug-in to read RAW data from and write ESDs and AODs to the T0 disk pools. The transfer of RAW data between the DAQ disk buffer at ALICE P2 and the T0 disk pools (CDR) will be done through *rfc*.

The plug-in is part of the standard Castor deployment. The *xrootd* services required to use the plug-in are installed and operated by ALICE. The plug-in code is supported by SLAC on a best efforts basis subject to the limitations described above.

Use of the *xrootd* plug-in by-passes the usual load balancing features of Castor and so risks to perturb the overall performance of the Castor disk storage spaces. Taking this into account, and with the experience gained through

**Note: Only change in version 3a – All references to LCG changed to WLCG.

thorough testing, ALICE and CERN/IT will agree on the portion of the traffic that could go through the *xrootd* Castor interface, in particular to and from the T0 disk pool.

DPM

An *xrootd* plug-in for DPM has been developed by CERN/IT with help from A.Hanushevsky. This plug-in is now part of the WLCG deployment. It uses the ALICE security model and so is suitable only for use by the ALICE collaboration. The *xrootd* plug-in is packaged as an rpm and provided together with the other DPM rpms. The plug-in will be supported by CERN/IT at any WLCG site that has agreed with ALICE to deploy it. Use of the plug-in requires the prior installation of *xrootd* services at the site. It is the responsibility of ALICE to agree with the site on the deployment and support for these services.

dCache

The dCache implementation of the *xrootd* protocol has been developed by the dCache project and is part of the standard dCache deployment. Support for the next year, within the performance characteristics of the current implementation, is funded by the D-Grid project. F.Carminati, on behalf of the ALICE Computing Project, has discussed with P.Fuhrmann about future support. Further support can be provided by the dCache project, provided funding can be assured. The dCache project also requires effective communication channels to be established with ALICE, and a formal validation procedure to be defined and operated. This is detailed in the document attached to this note.

Deployment at Sites

There are no formal WLCG milestones about deployment of *xrootd* support at sites. This is organized through bilateral agreements between a site and ALICE, as in the case for CERN and Castor.

Inter-Site Data Transfer

It should be noted that the only WLCG supported mechanism for transferring data between sites is FTS, where each end of the transfer uses one of the supported storage systems with SRM 1.1 or SRM 2.2.

Support proposal for the *xrootd* protocol implementation in dCache for the LHC Alice Experiment

Patrick Fuhrmann, 22 November 2007

Background

The LHC Alice experiment decided to use the *xrootd* protocol to access data stored in their WLCG Storage Elements for local and wide area transfers. The major advantage of using this protocol for ALICE is to access different Storage Element implementations by means of a single protocol. While DPM and Castor have been integrating the original implementation of the *xrootd* server, dCache.org decided to implement the *xrootd* protocol inside dCache. Mainly two arguments have been driving this decision.

- a) dCache strictly separates the file name spaces from the data storage spaces, so that the name of a file is not known on the storage nodes, which would be required by the *xrootd* server on each single data node.
- b) Secondly, dCache keeps track of open files in order to do proper load sharing, to replicate files on detection of hot spots and to migrate and restore datasets from/to tape.

This difference between the DPM, Castor and the dCache implementation, and the fact that dCache.org is not part of the WLCG management structure requires an agreement on how and to which extent the dCache people are supposed and able to support the *xrootd* implementation in dCache.

General Support

The DESY part of the dCache collaboration agrees to support the *xrootd* protocol in dCache for ALICE, on a best effort bases in the framework of the already established dCache support mechanisms, as we do for any other dCache component. In this context, support means that we continue to maintain the already existing software. In case additional development is required, please see the section on development below. There are however some more constraints to consider and some requirements from the dCache group to WLCG or the ALICE experiment to be able to provide a reliable support.

Funding

Currently the *xrootd* protocol implementation and support in dCache is funded by the HEPCG d-grid project. This funding will end at the end of 2008. Any support beyond this point depends on whether or not funding from other sources can be established. If no funding can be acquired, dCache.org offers a collaboration with ALICE so that ALICE may use its own IT resources in order to maintain and support the *xrootd* component in dCache, as we do with other groups. dCache.org would provide source code and documentation.

Level of support

dCache.org is not able to provide support for individual ALICE users. We expect ALICE or WLCG to implement a support level infrastructure, so that only identified issues or official ALICE requests are reported to support@dcache.org

Improved information exchange and release certification

dCache.org can only provide reasonable support for its xrootd implementation if we have access to a WLCG/ALICE contact person, who is sufficiently familiar with the xrootd client and server implementations, as well as with the ALICE security model.

There must be an agreed ALICE or WLCG certification process of the xrootd implementation in dCache to make sure that the certified dCache version is compatible with the xrootd client applications and libraries, with major focus to the use ALICE makes of them and to the versions deployed on ALICE sites. dCache.org will only formally verify the xrootd compatibility but it will not perform the testing against the whole ALICE framework.

Further development

dCache will consider further developments of its xrootd implementation, provided the following conditions are fulfilled, besides the ones mentioned above for normal support.

- * Although the xrootd protocol specification is freely available, the way ALICE is using the protocol needs detailed expertise. This is especially true for the ALICE security model. dCache.org requires a WLCG/ALICE contact person, which is sufficiently familiar with the xrootd client and server implementation as well as with the non-standard Alice security model.
- * dCache.org needs to be involved in the discussion of any changes in the
 - xrootd protocol specification (used by ALICE).
 - the methods ALICE client software is using out of the protocol specification.
 - the way ALICE client software is using the protocol.
 - Alice security model.
- * Sufficient resources must be available.
- * If the amount of resources dedicated to xrootd support in dCache.org is not enough to perform the requested development, and in particular if funding for this activity has ended, dCache.org may refuse or delay further development.

In case this development is critical for the ALICE experiment, ALICE should

- Establish a collaboration with dCache.org, where ALICE will provide the resources to further develop the xrootd component in dCache (as we do with other groups). In this case dCache.org would provide source code and documentation.
- Provide funding to dCache.org to perform the additional development in question.