Demonstration of Policy Reification in iRODS

Antoine de Torcy and Mike Conway

University of North Carolina at Chapel Hill, Chapel Hill, NC

Abstract

The DICE group has been working with the National Archives and Records Administration (NARA) on building preservation environments to store and manage NARA's digital records. In particular, NARA has been interested in using iRODS to define and automatically enforce digital preservation policies. Using iRODS' built-in Rule Engine, it is possible to apply such policies by mapping their contents to computer-actionable Rules executed by the iRODS system.

To validate this principle, a prototype of a policy-based preservation environment has been built and presented at the Society of American Archivists' 2010 annual meeting. It consists of iRODS and two software tools designed by DICE primarily for the digital preservation community: Arch, an archivist's interface that allows the definition and management of Policies via a user-friendly graphical interface; and iDrop, another user-friendly graphical user interface that enables users to easily ingest records into iRODS.

Finally, two repositories were created inside iRODS to store the Preservation Policy documents themselves and their mappings to iRODS Rules. Using the contents of these repositories the system then allows an archivist to define a policy from an available set of capabilities and to bind this policy to record series, collections, or any arrangements of records.

We demonstrated this prototype policy-based preservation environment by implementing a simple set of requirements likely to be part of accession policies and related to the integrity and authenticity of the records.

By filling in a form in Arch, an archivist can compose a policy that enforces any (or all) of the following actions on incoming records: physical replication, checksum validation, and scan for malware. A retention period can also be attached to the records.

The policy is then stored in XML format in a Policy Repository inside iRODS. Arch also allows the definition of Record Series, at which point it is possible to select (via a menu) from the list of available Policies in the Repository, which Policies should be bound to the Record Series. We showed with this system that adding new records to a series that is bound to a policy results in the policy being applied to the new records.

In addition, we used the iDrop interface to show how the presence of certain metadata flags can modify the display of records, therefore reflecting compliance or non-compliance with their associated polices.

This prototype demonstrates that computer-actionable Rules can automatically enforce desired preservation properties. Considerable work remains to be done in enabling the defining of such Rules and mapping them to Preservation Policies, so that archivists may create policies without needing extensive knowledge of the underlying preservation system software.

Index Keyword Terms — digital records, preservation, policies, iRODS, iDrop, Arch.