Texas Digital Library Services

Preservation Network

Overview

- Mission and goals for project

- Initial test configuration

- Current status of project

- Extension of test systems

- Challenges and opportunities

- Questions

Texas Conference on Digital
Libraries

Objective

From TDL Website …

"The preservation of the digital assets stored in the Texas Digital Library is a fundamental concern to be addressed as part of all the TDL's services. A robust preservation network is being designed to provide long-term storage of digital information, geographically distributed storage nodes, and migration of supported formats to preserve the ability to access materials into the future."
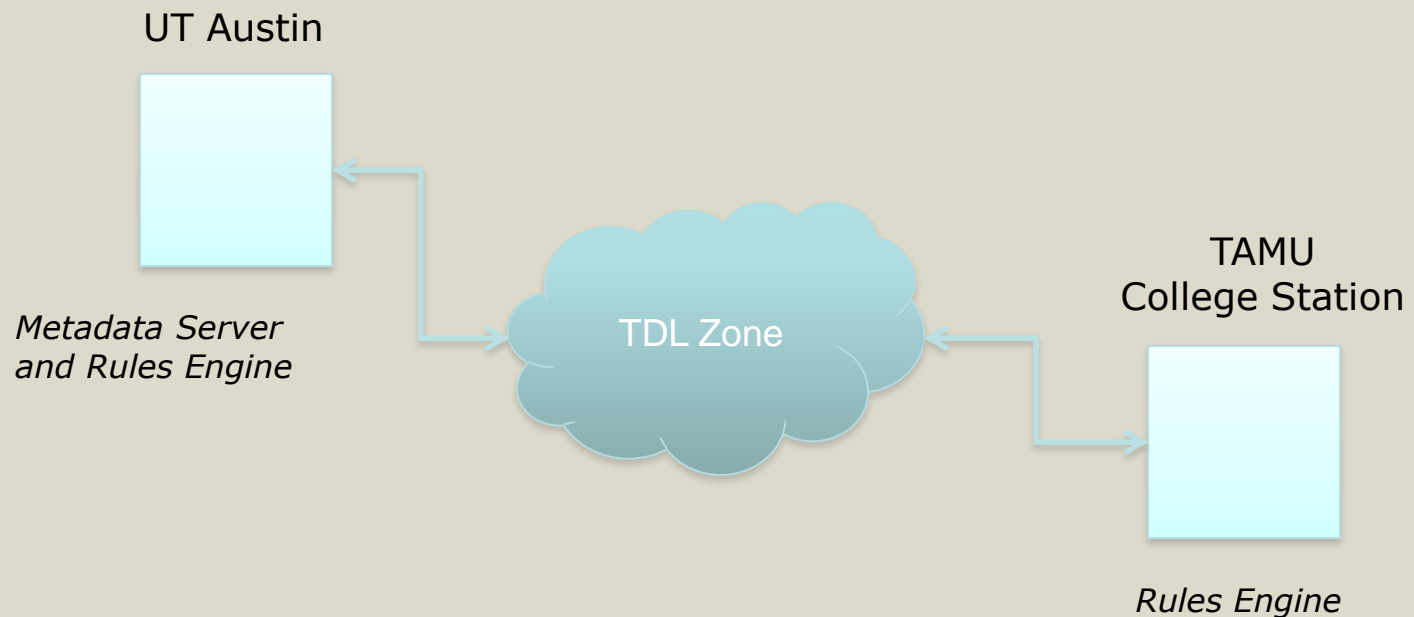
Preservation
is not
Backup

Texas Conference on Digital Libraries

Phase .5 Goals

- Establish initial test bed with multiple sites
- Explore iRODS as software platform
- Test transfer of files to multiple resource locations
- Test policy or rules driven configurations

Texas Conference on Digital Libraries

Topology

- Built on LEARN
- Simple, Single Zone

UT Austin

TAMU
College Station

TDL Zone

*Metadata Server
and Rules Engine*

Rules Engine

Software

- iRODS

  - Rules engine
  - Metadata server

- Windows Explorer client
- Web Interface

iRODS

- Policy driven software
- Driven by metadata
- Extends across entire zone

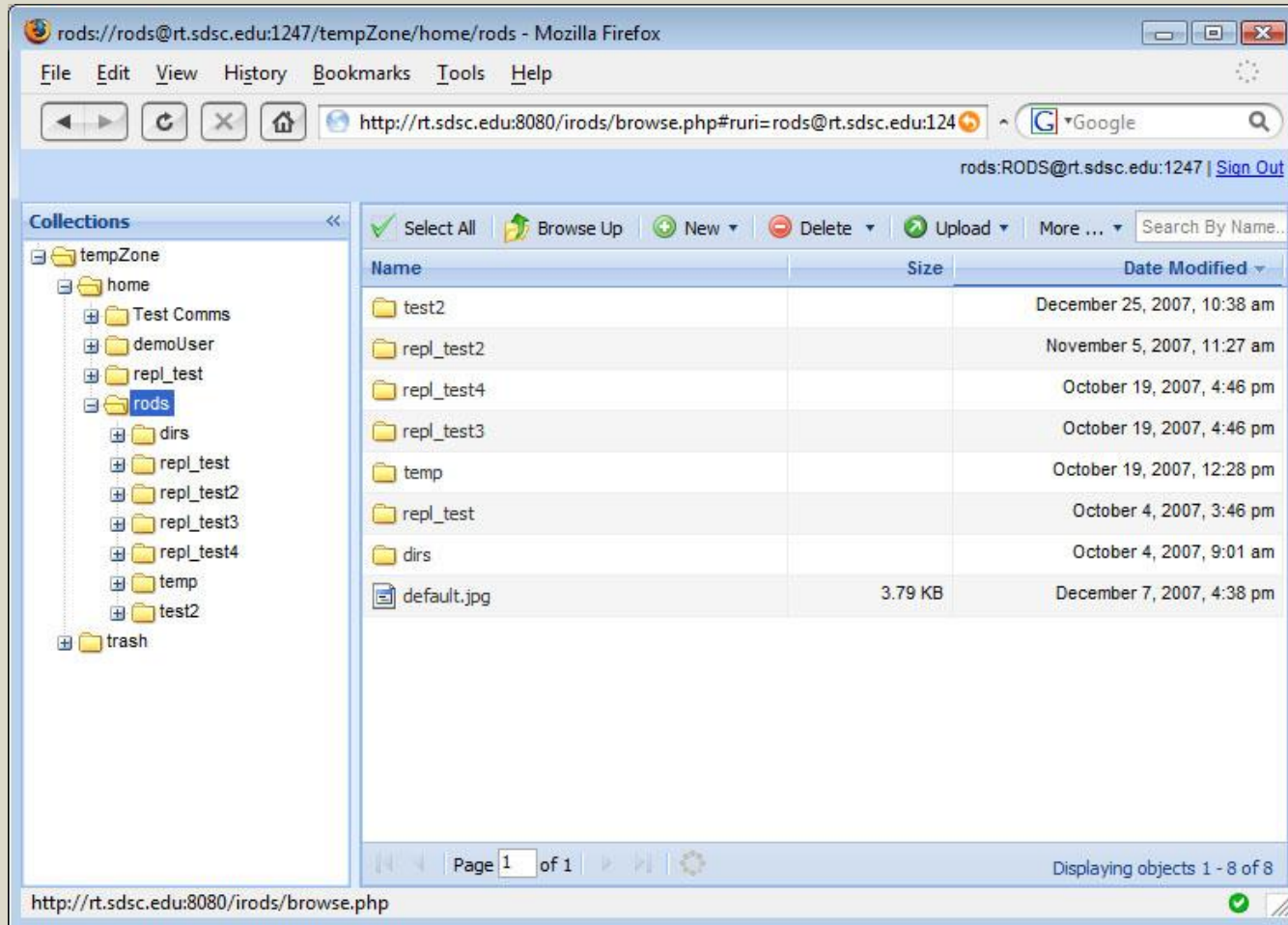*"make a duplicate copy of file upon initial store in this resource"*

Current Status

- Test configuration operational

  - Austin to College Station
  - Rather small configuration at this time

- Imported multiple files
- Tested transfer of files
- Replicated files between nodes

Texas Conference on Digital Libraries

User Interface

- Command Line Interface

- Web browser

- Windows Explorer

- Developed in new software or added to existing systems

User Interface

Texas Conference on Digital Libraries
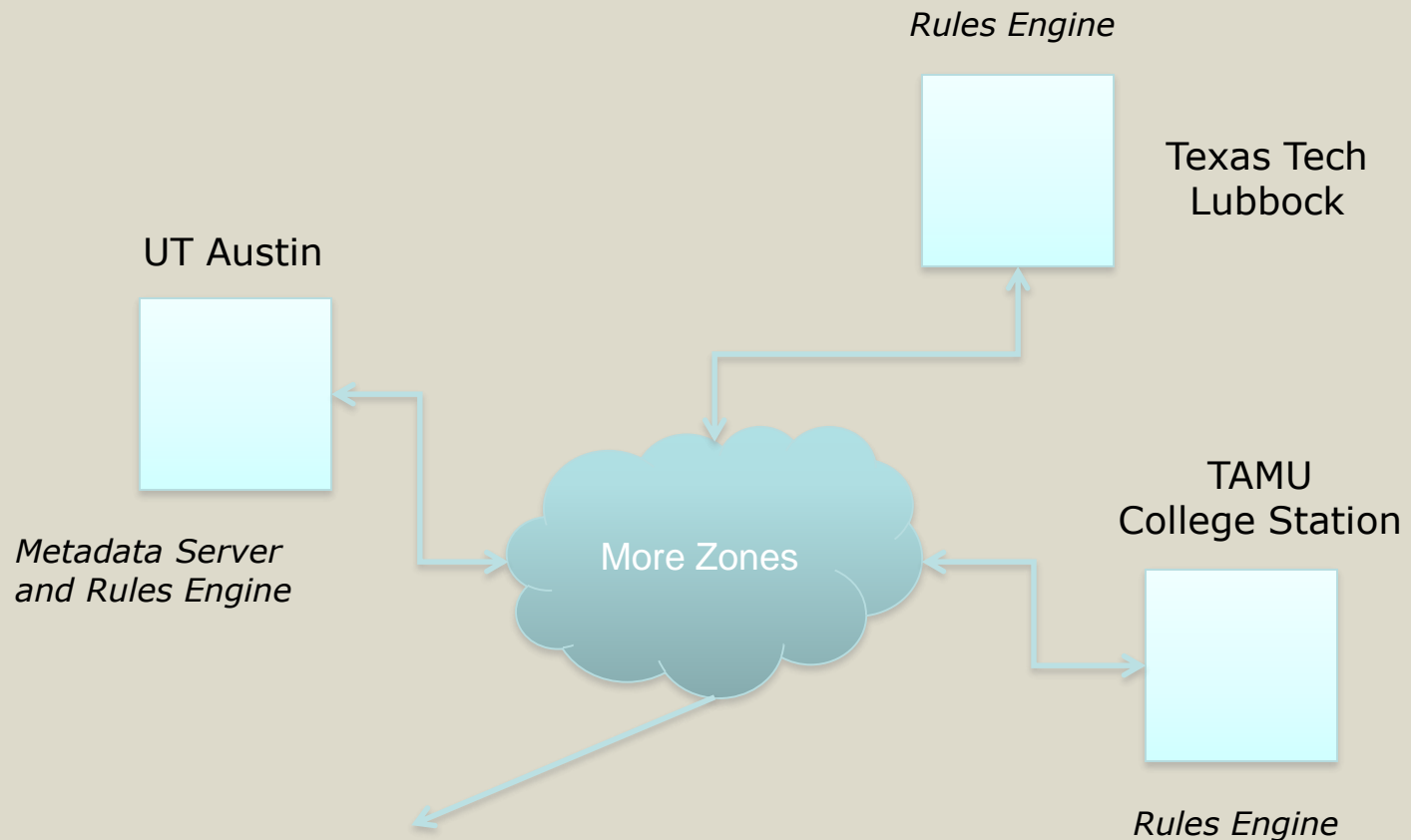
Extension of Test

- Add Texas Tech

- Develop rules and policies

- Extend metadata to include custom defined elements

- Test additional zones

Extension of System

Rules Engine

Texas Tech
Lubbock

UT Austin

*Metadata Server
and Rules Engine*

More Zones

TAMU
College Station

Rules Engine

How we use System

- Institutional Repository will be primary interface

  The IR will be the primary storage and search mechanism for scholarly works.

- Little need to ever access iRODS servers directly

Open Repositories '09

- Learned a great deal

- Opportunities for collaborations

- Study what others have done and are doing

- UK and Europe strong leaders in preservation and archiving

- Made a number of contacts

Roadmap

- Design configuration with geographically disperse nodes

- Look at deep archive

- Explore use of repository for archive management

- Develop and refine policies for storage, retention, data migration

- More fully develop metadata

Challenges

XAM

MODS

PREMIS

MADS

EDTF

SNIA

RDF

METS

CLOUDS

XACML

PACKAGES

Discussion/Questions