Enabling Collection Interoperability & Preservation

Using iRODS and the OAI-PMH

Jewel H. Ward

Doctoral Student, UNC-CH SILS

Digital Library Federation
Spring Forum 2009

May 4-6, 2009
Raleigh, NC USA

Discussion **Points**

Motivation & Introduction

Background

Method

Results & Discussion

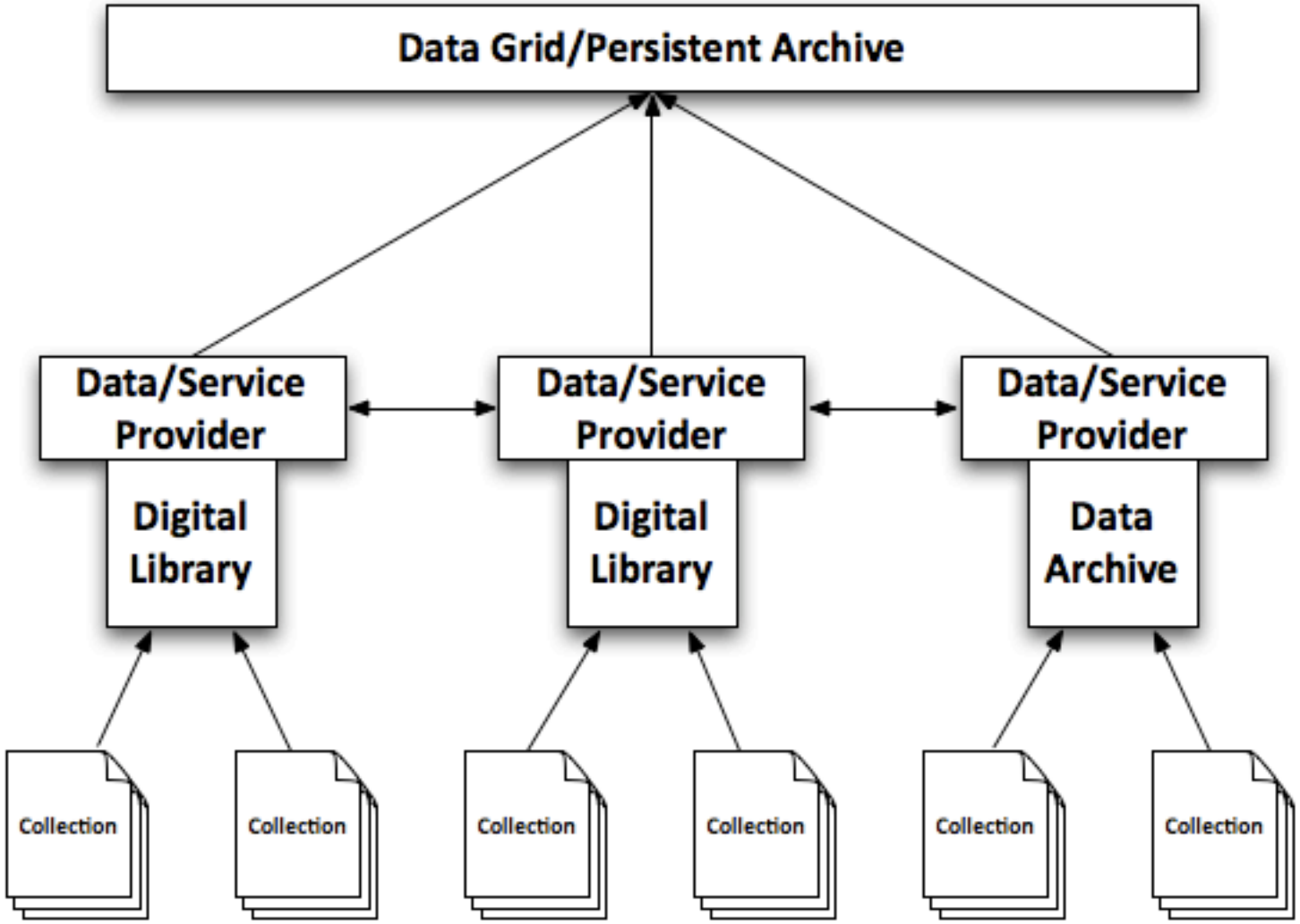
Conclusions

Current & Future Work

Acknowledgements & Questions


40
25
15
15

Motivation &
Introduction


Data Grid/Persistent Archive
Data/Service Provider
Data/Service Provider
Data/Service Provider
Digital Library
Digital Library
Data Archive
Collection
Collection
Collection
Collection
Collection
Collection

Interoperability

Archivist's Preservation Models

Digital Librarians' Information Models

Data Grid Archival Storage Technology

Archival **Model**

*Appraisal| Accession

Arrangement | Description

Preservation| Access

*R. Moore, *The Preservation of Data, Information & Knowledge*.

Digital **Libraries**

*Federated Servers| Multiple Clients

Collection Owned Data| Access by Logical Name

Web-based | Federated Name Spaces

*R. Moore, *Evolution of Data Grid Concepts*.

Data Grids

*Library-based Access to Servers| Personally Owned Data

Access by Physical File Name | Application Driven Processing

Global Name Space | Data Moves to Process

Management of Distributed State Information

*R. Moore, *Evolution of Data Grid Concepts*.

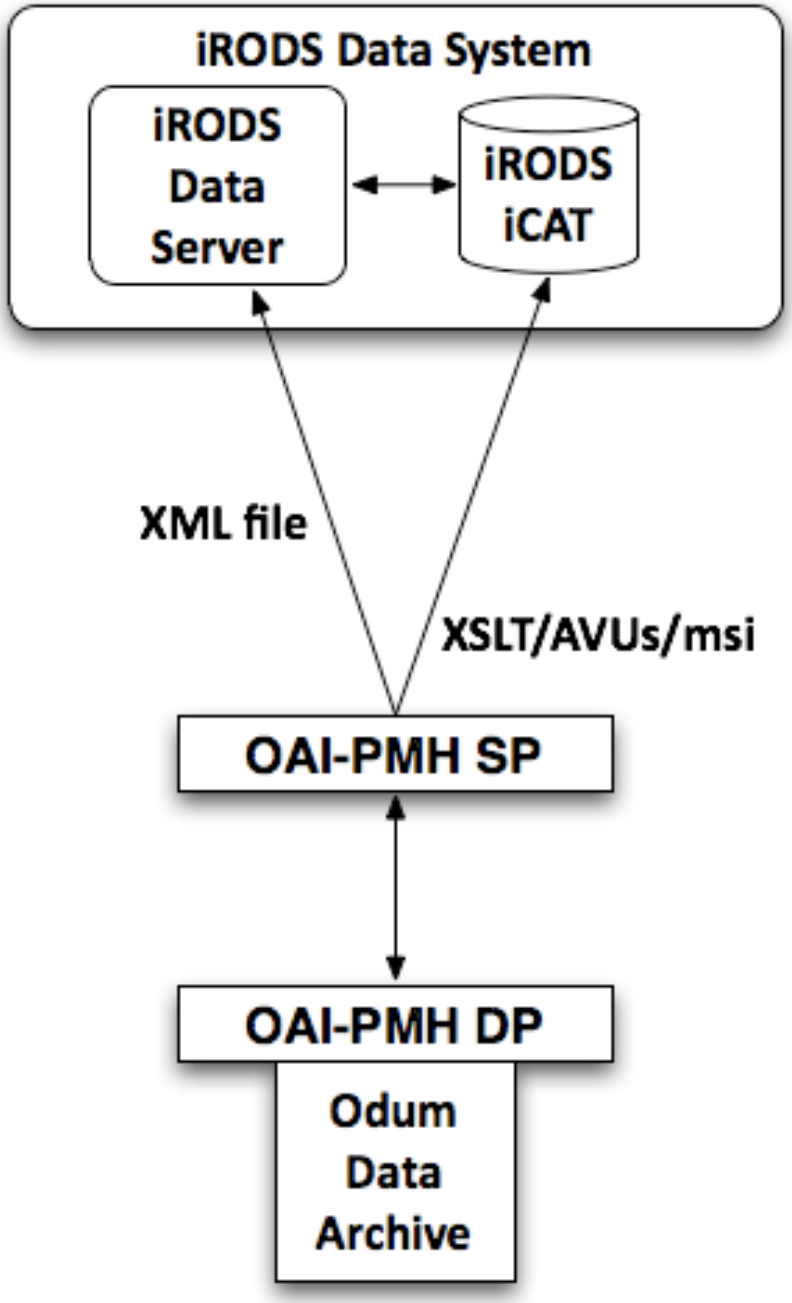
Background

Open Archives Initiative | OAI-PMH

Odum Institute Digital Archive | DVN

Data Intensive Cyber Environments | iRODS

National Archives and Records Administration | TPAP


iRODS Data System
iRODS Data Server
iRODS iCAT
XML file
XSLT/AVUs/msi
OAI-PMH SP
OAI-PMH DP
Odum Data Archive

Method

Results & Discussion

Conclusions & Future Work

Acknowledgements

Antoine de Torcy (DICE, UNC-CH)

Jennifer Mantooth (SILS, UNC-CH)

Mason Chua (Odum, UNC-CH)

Jon Crabtree (Odum, UNC-CH)

Acknowledgements

This work is funded by the NSF grant OCI-0848296 and is a collaboration with NARA on the development of the "NARA Transcontinental Persistent Archive Prototype". The initial work on this project was funded by the NARA supplement to NSF SCI 0438741, “Cyberinfrastructure; from Vision to Reality” – Transcontinental Persistent Archive Prototype (TPAP) (2005-2008).

Content, in the form of reusable and often vary large data sets and data bases – numeric, textual, visual – is an **integral** part of advanced **information technology** also.

~Clifford Lynch