Data Intensive Cyberinfrastructure Foundation

Supporting iRODS Open Source Community

Paul Tooby
Data Intensive Cyberinfrastructure Foundation
ptooby@diceresearch.org
Website: http://diceresearch.org
DICF Community Foundation

- iRODS Goals: Developing Useful Software
- Beyond software development...
- Long-term viability of iRODS open source s.w.
- Education
- iRODS copyrights
- iRODS Community
iRODS Goals

- Long-term viability of iRODS open source software
- Data Intensive Cyberinfrastructure Foundation
  - Nonprofit corporation
  - Startup in 2008.
- What are the benefits of a Foundation?
  - Long-term home for iRODS
    - Beyond any university, grant, individual developer, etc.
    - Assures community of long-term availability of iRODS.
  - iRODS Education
    - iRODS expertise and data management in general
  - Manage intellectual property
  - Community outreach
Data Intensive Cyberinfrastructure Foundation

Foundation Directors
- Reagan Moore, Arcot Rajasekar, Richard Marciano, Wayne Schroeder, Michael Wan

Foundation Officers
- Reagan Moore, President
- Paul Tooby, CFO and Secretary

Contact
- Data Intensive Cyberinfrastructure Foundation
- Website: http://diceresearch.org
Foundation Education

- iRODS Community Education
  - About iRODS and generic data management
  - Participation in collaborative education activities

- Using iRODS
  - iRODS is highly configurable and extensible
  - Showing the many ways of using iRODS
iRODS User Community

- A sampling of projects using iRODS...

- iRODS Development Collaborations
  - NARA TPAP Transcontinental Persistent Archive Prototype (NARA funded)
  - NSF SDCI Research in Adaptive Middleware Architecture Systems
  - SHAMAN Sustaining Heritage Access through Multivalent ArchiviNg
  - CC-IN2P3 France; ARCS Australia; UK e-Science data grid; CeRch KC London; KEK Japan; Academica Sinica, Taiwan; Fedora Commons; Dspace, HDF5; OpenCobalt; Taverna, etc.

- Communities Using DICE Technologies, including Biology, Environment, Psychology, Human Subjects
  - iPlant Collaborative (NSF)
  - TDLC Temporal Dynamics of Learning Center (NSF funded)
  - National Center for Microscopy and Imaging Research (NCMIR CCDB)

- Physical Sciences, including astronomy, physics, etc.
  - NOAO National Optical Astronomy Observatories data grid (NSF)
  - NVO National Virtual Observatory (NSF)
iRODS User Community

- Physical Sciences (continued)
  - Observatoire de Strasbourg, France, VOSpace Interface
  - CADAC Computational Astrophysics Data Analysis Center (NSF)
  - BaBar high energy physics data grid (DOE)

- Persistent Archives and Digital Preservation / Humanities Uses
  - NARA TPAP Transcontinental Persistent Archive Prototype
  - e-Legacy Preserving the Geospatial Data of the State of California
  - Distributed Custodial Archival Preservation Environments (NHPRC)
  - T-RACES Testbed for Redlining Archives of CA Exclusionary Spaces (IMLS)

- Geosciences Uses
  - OOI Ocean Observatories Initiative (NSF)
  - SCEC Southern California Earthquake Center (NSF)

- High Performance and Grid Computing
  - NSF TeraGrid, DOE, NASA Center for Computational Sciences (NCCS)

- And more, growing all the time...
iRODS Use Cases

☐ Help others find out about your iRODS project...

☐ Are you included in our projects list?
  ■ In the DICF Website check the list on "IRODS Uses"
    ☐ http://diceresearch.org

☐ Send link to your project and we'll include you!

☐ Send us a bit more information...
  ■ Run the "zoneinfo" script and send us results
    ☐ Helps build use cases
    ☐ Gives you overview of your zone, helps troubleshooting
    ☐ Helps iRODS developers
iRODS Use Cases

- Example of zoneinfo script information
  - Tue Sep 15 14:32:26 PDT 2009
  - Local zone is: newZone
  - Remote zones are:
  - Number of users: 5
  - Number of user groups: 2
  - Number of resources: 2
  - Number of collections: 1,077
  - Number of data-objects: 1,050,639 (1 million)
  - Total data size (bytes): 299,137,241 (299 million)
  - Data distribution by resource:
    - 299137157 bytes in 1050637 files in 'demoResc'
  - Using config file: ./../config/config.mk
  - OS=solaris
  - no GSI
  - no Kerberos
  - no FUSE
  - MODULES= properties
  - PostgreSQL ICAT
  - SvrInfo:
    - RCAT_ENABLED relVersion=rods2.1 apiVersion=d rodsZone=newZone up 0 days, 5:31
  - Tue Sep 15 14:32:48 PDT 2009
  - Please enter yes or y to send the above information to the iRODS team:y
Learning About iRODS

iRODS Use Cases

- Outreach to non-technical audiences including scientists, archivists, others showing how they can use iRODS. Also shows benefits to funding agencies...
- Raising the level of "data management literacy"

Use case for your project... made easy

- Send us link to your project and one sentence descrip.
- Run and send zoneinfo script
- Leverage your existing information:
  - Send news & overviews - Web text, ppt, press release, reports.
  - And we'll send your Use Case to funding agencies
    - NSF, etc., your project receives wider recognition.
Learning About iRODS

- Training programs
  - Customized training
  - Data architecting, etc.

- Books
  - *iRODS Primer*, Morgan & Claypool, 2010

- Web resources
  - Use cases
  - Online video
    - e.g. Networking and Information Technology Research and Development (NITRD) demo
Long-term iRODS viability

- Managing iRODS intellectual property
  - Copyrights in iRODS code
  - Contributor's Agreement
  - Trademark, etc.

- Intellectual Property and Developers
  - Open source ≠ public domain: Someone owns copyrights.
  - Default: you (or employer) own copyrights in code you write.
iRODS and Copyrights

- iRODS Copyrights and DICF Foundation
  - Transferring copyrights in code contributions to DICF lets us distribute iRODS as open source BSD license software
    - Ensures iRODS remains open source over long term
  - Defragments copyrights, Foundation can protect them

- iRODS Contributor's Agreement
  - Normal in open source projects e.g. Apache, etc.
  - Download from wiki, contributor of code (or employer) completes
  - Transfers copyrights in code contributions to Foundation
  - Also gives developer license to do anything they want with their code. Benefits iRODS with no loss of freedom.
iRODS Collaborations

- Collaborations with other projects and foundations
  - Education projects, development, etc.
  - Outreach e.g. DuraSpace Foundation, preservation community, etc.

- Partner Program
  - Vendors bundling, supporting iRODS, etc.; large users, etc.
Summing up...

- Building iRODS community...
- Send us a link to your project
- Run zoneinfo script & send brief info.
- Contributors
  - iRODS Contributor's Agreement
- Collaborations with projects & foundations
Thank you!

☐ Where to find out more
  ■ DICF website http://diceresearch.org

☐ Questions?
  ■ ptooby@diceresearch.org