

Islandora and Fedora 4; The Atonement.

Project history and background

Fedora 4 Interest Group

Islandora Fedora 4 Interest Group

Convenors

- Nick Ruest (York University)
- Daniel Lamb (discoverygarden)

Terms of Reference

- The name of the IIG is the Islandora Fedora 4 Interest Group
- The purpose of the Islandora Fedora 4 Interest Group is to implement Fedora 4 in the context of Islandora.
 - [Islandora/Fedora 4 Prospectus](#)
 - [Islandora/Fedora 4 Project Plan](#)
- Specific goals, activities, outcomes may include:
 - Create a generic "upgration" document for Islandora and:
 - Identify pilot partners
 - [York University](#)
 - Reviewing Fedora 3 to 4 [Upgration Checklist](#) and:
 - Identifying what features are necessary in Fedora 4
 - Mapping Fedora 3 features to Fedora 4 features
 - Mapping Fedora 4 features to Islandora
 - Determining how Islandora will leverage new Fedora 4 features
 - Outreach and communication
 - Engaging other interested Islandora community members
 - Recruiting developers to work on integration tasks
- The interest group will meet once a month virtually (eg. via Skype). A call for agenda items will be posted to the Islandora Google Group 1 week prior to the meeting. The convenor will appoint a note taker for the meeting and meeting notes will be made available at some url.
- The convenors will produce a report to be submitted to the Islandora Roadmap Committee following the IF4IG's meeting.

Links

Thank you to our sponsors:



Atonement

Documentation

About Islandora[About this guide](#)[Installation](#)[Contributing](#)[Sponsors](#)[Maintainers](#)[Licensing](#)

About Islandora

[Islandora](#) is an open-source software framework designed to help institutions and organizations and their audiences collaboratively manage, and discover digital assets using a best-practices framework. Islandora was originally developed by the University of Prince Edward Island's Robertson Library, but is now implemented and contributed to by an ever-growing international community.

Islandora consists of:

- [Sync](#) - Event driven middleware based on Apache Camel that synchronizes a Fedora 4 JCR with a Drupal CMS.
- [Islandora](#) - Fedora 4 Repository module
- [Install](#) - This is a development environment virtual machine for the Islandora and Fedora 4 project. It should work on any operating system that supports VirtualBox and Vagrant.

About this guide

The [Technical Design documentation](#) will help you:

- Understand the Islandora 7.x-2.x design rationale
- Importance of using an integration framework
- Using camel
- Inversion of control and camel
- Camel and scripting languages
- Islandora Sync
- Solr and Triple store indexing
- Islandora (Drupal)

The [How to build documentation](#) provides an overview on how the documentation is created, built, and deployed.

Installation

The [installation section](#) provides an overview on how to create a virtual development environment.

Contributing

If you would like to contribute, please get involved with the [Islandora Fedora 4 Interest Group](#), and check out the [contributing section](#). We love to hear from you!

If you would like to contribute code to the project, you will need to be covered by an Islandora Foundation [Contributor License Agreement](#) or [Corporate Contributor License Agreement](#). Please see the [Contributors](#) page on Islandora.ca for more information.

Sponsors

- Discoverygarden

Contributing

<http://islandora-labs.github.io/islandora/contributing/contributing>

“All contributions are welcome: use-cases, documentation, code, patches, bug reports, feature requests, etc. You do not need to be a programmer to speak up!”

DevOps



VAGRANT

branch: 7.x-2.x islandora / install / +

This branch is 197 commits ahead, 113 commits behind Islandora:7.x

Pull Request Compare

Pom gardening. Removing hibernate dependencies.

daniel-dgi authored 2 days ago

latest commit 16c4a09252

..		
.vimrc	Setting default colorscheme to get rid of 'Highlight Normal not found...	2 months ago
README.md	Merge branch '7.x-2.x' of github.com:Islandora-Labs/Islandora into 7...	a month ago
Vagrantfile	For some reason this ordering helps the build.	21 days ago
bootstrap.sh	Update to Java8	a month ago
disable_node_access.php	Quick and dirty php script to disable node security so POC works with...	3 months ago
drupal.sh	Fedora specific info (like path and parent) are coming back asynchron...	3 days ago
fcrepo.sh	fcrepo version bump and bump up JAVA_OPTS	2 months ago
fuseki.sh	Update fuseki to 2.0.0 and deploy via war. Closes #13.	a month ago
post-install.sh	Initial services commit. Hibernate setup that inserts uid -> path	25 days ago
services.sh	Pom gardening. Removing hibernate dependencies.	2 days ago
solr.sh	Add Solr install script (4.10.3). Closes #2	3 months ago
sync.sh	Initial services commit. Hibernate setup that inserts uid -> path	25 days ago
vim.sh	Merge commit. More vim.sh mushing	a month ago

README.md

Introduction

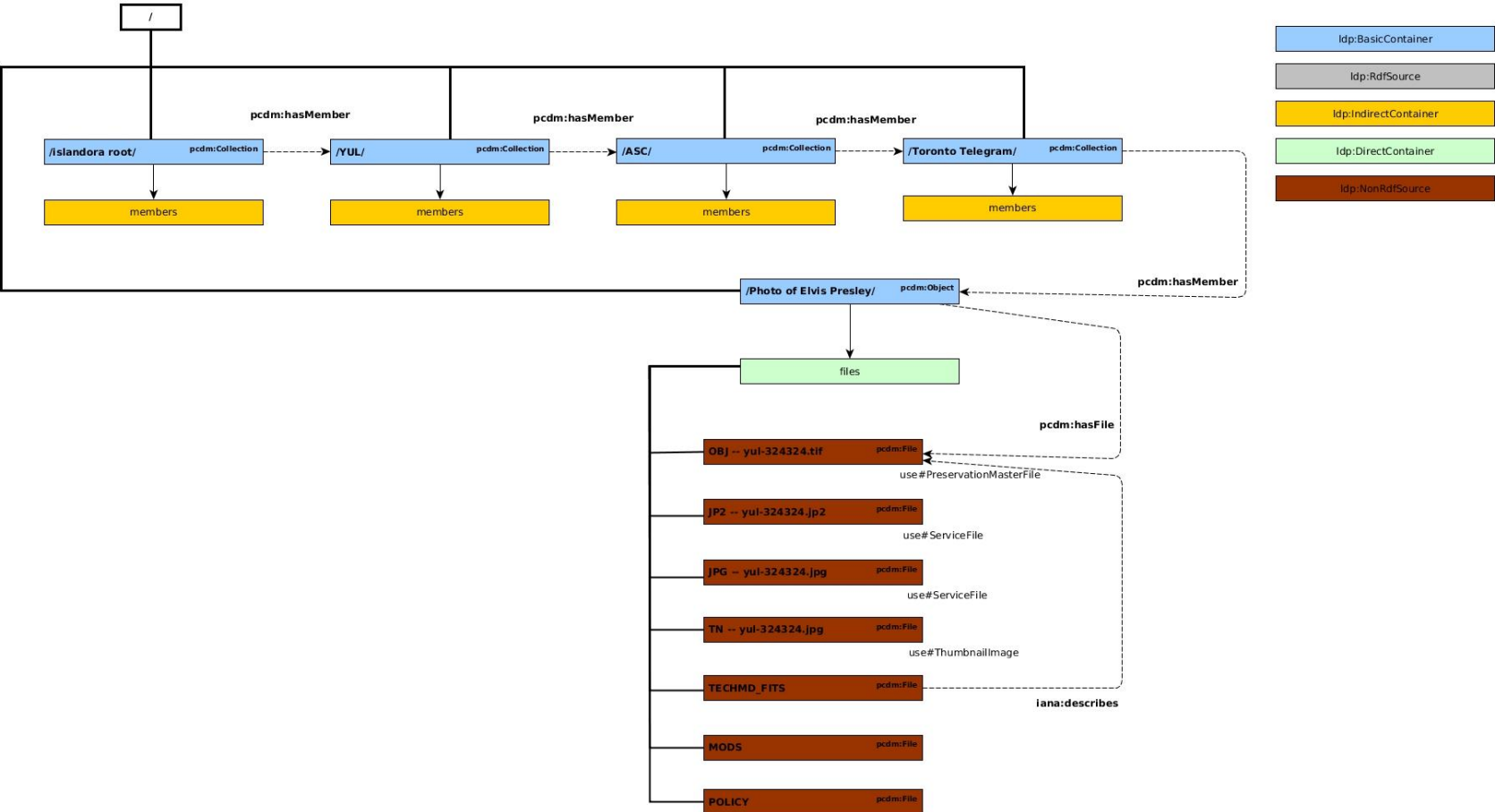
The is a development environment virtual machine for the Islandora and Fedora 4 project. It should work on any operating system that supports VirtualBox and Vagrant.

vagrant up

PCDM

Fedora, Hydra, Islandora... Repository
communities unite!

**Islandora, LDP, PCDM, and Fedora 4
(Islandora Large Image Object)**



$$7.x - 2.x$$

Design Goals

- Utilize Fedora 4 to its fullest
- Easier to install
- Easier to develop and contribute
- Easier to use
 - Enhanced user experience
 - Faster searching, display, and ingest
 - Expose more control of your repository through UI
- Easier to scale

How Do We Achieve These Goals?

- Installation bash scripts
- Vagrant development environment
- Tight Drupal integration
- Asynchronous interactions (when possible)
- Transformation logic in the UI
- Decoupling
 - Decoupling... Decoupling... Decoupling...

**And How Do We Achieve THOSE
Goals?**

Apache Camel

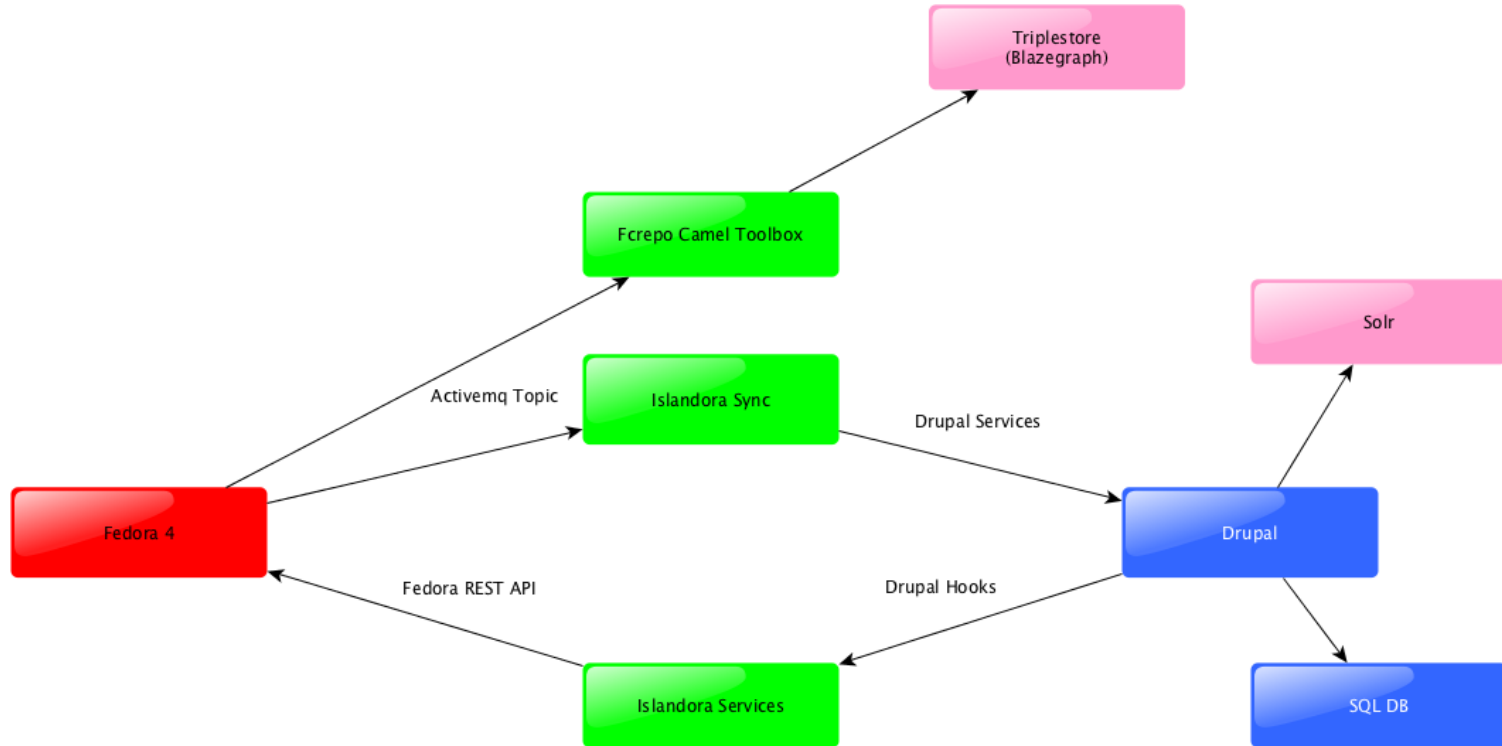
Why Camel?

- Islandora is middleware!
 - Camel is a framework for building middleware
 - Provides workflow-like structure for the code
 - It's already solved a lot of really hard problems
 - Message splitting, transformations, concurrent processing, error handling / redelivery...
 - Works out of the box with everything you can think of
 - Filesystems, emails, web services, databases, and now... Fedora (thanks Aaron Coburn!)

What is Camel's Role in Islandora?

- **Dirty Work**
 - Data transformation
 - Calling out to command line tools
 - Interacting with other APIs
- **Doing the Dirty Work Sanely**
 - Transactions, redelivery, dead letter channels
- **Man in the Middle**
 - Fedora and Drupal never talk to each other

What would a system with Camel at its core look like?



Islandora Sync

- Event driven middleware that synchronizes Fedora resources with Drupal nodes
- Does 3 things
 - Deletes things that no longer exist
 - Upserts everything else
 - Transforms Id+json RDF from Fedora into Drupal node JSON and issues PUT requests.
 - Generates derivatives

Islandora Services

- Web services invoked in response to Drupal hooks
- Wraps the Fedora 4 REST API
 - Accepts data in Drupal's formats
 - Constructs PCDM-compliant resources in Fedora
 - Transforms Drupal node JSON into SPARQL Update queries
- Doesn't wait around

Benefits of This Approach

- Better Drupal Experience
 - Content is Drupal nodes
 - Developers work with Drupal content, not Fedora content
 - Take full advantage of drupal hooks and theming
 - UI controlled content modeling (fields)
 - UI controlled display (site building)
 - Views!

Benefits of This Approach (Cont'd)

- 3rd Party Modules
 - apachesolr
 - rdfx
 - services
 - xml_field
 - xpath_field
 - field_permission
 - and many many more....

Benefits of This Approach (Cont'd)

- RDF / Field synchronization
- RDFa enriched output
- All exposed through Drupal UI

PCDM:HASMEMBER RDF MAPPING

RDF Predicates

Enter a comma-separated list of predicates for *pcdm:hasMember* u:

Attribute Type

For fields containing literals—things such as plain text, html, numb or [rev](#) attribute.

Save settings

Benefits of This Approach (Cont'd)

- XML Fields
- XPath Fields
- Extract metadata from XML and map to RDF
- Also exposed through Drupal UI

XPath value

Enter the XPath to select a fragment from the XML Field instance.

Only use text content

Return the text contents of the node specified by the XPath query

MODS TITLE RDF MAPPING

RDF Predicates

Enter a comma-separated list of predicates for *MODS Title* using C

Attribute Type

For fields containing literals—things such as plain text, html, num or [rev](#) attribute.

Datatype

Enter the datatype for *MODS Title* using CURIE syntax. For a list of

Benefits of This Approach (Cont'd)

MODS XML

```
<?xml version="1.0"?>
<mods xmlns:mods="http://www.loc.gov/mods/v3" xmlns:xsi="http://www.w3.org/2001/X
<titleInfo>
  <title>Oh, you're an engineer?</title>
  <subTitle>Please continue telling me why my career is a waste of time</subTitle>
</titleInfo>
<name type="personal">
  <namePart>The Internet</namePart>
```

A MODS record for the Fedora resource

MODS Title (Derived with /mods/titleInfo/title)

Oh, you're an engineer?

foaf: name

admin

modsrdf: title

Oh, you're an engineer?

nfo: uuid

2bc4d9af-7ae7-440e-a7f7-d7b59adb32ca

Benefits of This Approach (Cont'd)

- Now using drupal.org solr modules
- No more GSearch
 - No more XSLTs
- Maintaining much less code

Related Projects

Projects that are co-maintained in combination with this one

[Apache Solr Attachments](#) (7.x & 6.x)

Send your files to Apache Tika and allow Drupal to search in your attachments/files

[Apache Solr Multisite search](#) (7.x & 6.x)

Combine any combination of 6.x-3.x and 7.x-1.x Drupal sites to let them all search and index to the same Apache Solr core index

[Apache Solr Devel](#)(7.x-1.x)

Allows to fetch your entity from the solr index in a similar way as you can show the entity object in the devel tab of any entity. Very useful to find out what values you have in Solr for each field.

[Rich Snippets](#)(7.x-1.x)

By using RDF and schema.org semantic data for your structured content you can now show similar snippets as the ones Google provides

[Apache Solr Sort](#) (7.x-1.x)

Adds support for the Solr Grouping feature and adds an UI to enable/disable sort fields

[Apache Solr API Examples](#) (7.x-1.x, 6.x-3.x)

Shows you, with real examples, how to program Drupal modules using the Apache Solr Search Integration API

[Apache Solr Parallel Indexing](#) (7.x-1.x, 6.x-3.x)

Using PHP socket streams we can execute a number of php processes at the same time to make indexing speed faster. Speed increase is limited by the number of CPU's

Next Steps

Questions/Discussion