Design - Modeling for non-LDPC objects

This page presents straw man data models for objects that may not need Linked Data Platform Containers (LDPCs).

The defining characteristics of this class of objects:

- They do not contain bitstreams, but rather, metadata about bitstreams held elsewhere
- They may contain pointer(s) to the referenced bitstreams (or bitstream objects)

The related concept in LDP (Linked Data Platform) terms: Non-LDPC LDPRs

- LDPR: Linked Data Platform Resource
- LDPC: Linked Data Platform Container

To what extent can/should this class of objects be modeled as LDPRs not contained within a LDPC? What are the benefits of such a model? What are the drawbacks or limitations?

- Background: LDP and its mapping to Fedora objects
- Proposal 1: Fedora container object with a child resource at a specified path redirecting to external content
 - LDP to Fedora mapping
 - Request 1: create container object
 - Request 2: create child external content object
 - The complete thing
- Proposal 2: LDP-RS as datastreams
 - LDP to Fedora mapping
 - Discussion
 - Use cases
 - Implementation notes
- Proposal 3: LDPRs as "Pointlike objects"
- LDP to Fedora mapping
- Proposal 4: LDPRs as "Pointlike objects" with no constraints on RDF statements.
 - Discussion
 - Use cases
 - Implementation notes

Background: LDP and its mapping to Fedora objects

The LDP specification lays out the following concepts, interaction modes, and constraints on content:

LDP Concept	LDP Interaction Modes	Constraints on Content
Container (LDPC)	LDPC, LDP-RS	RDF, must contain LDP membership triples for all members
Non-RDF Source (LDP-NR)	LDPR, LDP-NR	none
Non-Container RDF Source (LDP-RS-NC)	LDPR, LDP-RS	RDF

The last concept "Non-Container RDF Source" isn't given a specific name in the LDP spec, so we call it an LDP-RS-NC for the sake of discussion.

The defining characteristics of the interaction modes described in LDP are briefly summarized as follows

- LDPC interaction mode defines ability to create resources within the container via HTTP POST, remove them via HTTP DELETE, and places
 requirements on the HTTP headers of the LDPC resource to advertise its role as a container.
- LDPR interaction mode does not imply the creation new resources with HTTP POST on an LDPR
- LDP-NR interaction mode implies POST, PUT, and PATCH do not have any special semantics; a resource is nothing more than a binary blob with a media type
- LDP-RS interaction mode implies that POST, PUT, GET, and PATCH are logical RDF operations with enhanced semantics and behaviour:
 - LDP servers must interpret relative URIs as relative to the resource (for example, allowing null relative URIs "<>" to refer to the to-becreated resource URI)
 - LDP Servers may adopt standards such as SPARQL/Update for update semantics via methods such as PATCH
 - LDP Servers must provide, at minimum, text/turtle and application/ld+json representations when requested via GET.

At present, the mapping of Fedora concepts to LDP Concepts are as follows

LDP Concept	Fedora Concept	LDP Interaction Modes	Constraints on Content
LDPC	Object	LDPC, LDP-RS	RDF, must contain LDP membership triples, subjects must be repository objects, subjects must be the URI of the object
LDP-NR	Datastream	LDPR for properties, LDP-NR for bitstream	none

|--|

The issue at hand is whether Fedora should support the LDP construct of an RDF source that is not also a container (LDP-RS-NC), and how such an resource in LDP would map to a concept in Fedora.

Proposal 1: Fedora container object with a child resource at a specified path redirecting to external content

LDP to Fedora mapping

LDP Concept	Fedora Concept	LDP Interaction Modes	Constraints on Content
LDPC	Object	LDPC, LDP-RS	RDF, must contain LDP membership triples, subjects must be repository objects, subjects must be the URI of the object
LDP-NR	Datastream	LDPR for properties, LDP-NR for bitstream	none
LDP-RS-NC	-	-	-

Note that this proposal maintains the current behaviour in that LDP-RS-NC is not supported.

Request 1: create container object

```
curl -i -X PUT "http://localhost:8080/rest/my_external_datastream_obj"
HTTP/1.1 201 Created
ETag: "f93bbf77fee0fa16da4872102a54240e7135eaf2"
Last-Modified: Thu, 12 Mar 2015 20:56:08 GMT
Content-Length: 53
Location: http://localhost:8080/rest/my_external_datastream_obj
Content-Type: text/plain
Server: Jetty(8.1.12.v20130726)
curl -H "Accept: text/turtle" "http://localhost:8080/rest/my_external_datastream_obj"
<http://localhost:8080/rest/my_external_datastream_obj> a ldp:RDFSource , ldp:Container , <http://www.jcp.org</pre>
/jcr/nt/1.0folder> , <http://www.jcp.org/jcr/nt/1.0hierarchyNode> , <http://www.jcp.org/jcr/nt/1.0base> ,
<http://www.jcp.org/jcr/mix/1.0created> , fedora:Container , fedora:Resource , fedora:Resource , <http://www.</pre>
jcp.org/jcr/mix/1.0created> , <http://www.jcp.org/jcr/mix/1.0lastModified> , <http://www.jcp.org/jcr/mix/1.</pre>
Oreferenceable> ;
        fedora:lastModifiedBy "bypassAdmin"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:uuid "fe7a8497-d826-490d-a786-ba601e039055"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:createdBy "bypassAdmin"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:primaryType "nt:folder"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:created "2015-03-12T20:56:08.468Z"^^<http://www.w3.org/2001/XMLSchema#dateTime> ;
        fedora:mixinTypes "fedora:Container"^<http://www.w3.org/2001/XMLSchema#string> , "fedora:Resource"
^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:lastModified "2015-03-12T20:56:08.468Z"^^<http://www.w3.org/2001/XMLSchema#dateTime> ;
        fedora:writable "true"^^<http://www.w3.org/2001/XMLSchema#boolean> ;
        fedora:hasParent <http://localhost:8080/rest/> ;
        fedora:exportsAs <http://localhost:8080/rest/my_external_datastream_obj/fcr:export?format=jcr/xml> .
<http://localhost:8080/rest/my_external_datastream_obj/fcr:export?format=jcr/xml> dc:format <http://fedora.info</pre>
/definitions/v4/repository#jcr/xml> .
<http://fedora.info/definitions/v4/repository#jcr/xml> rdfs:label "jcr/xml"^^<http://www.w3.org/2001</pre>
/XMLSchema#string> .
```

Request 2: create child external content object

```
curl -i -X PUT -H"Content-Type: message/external-body; access-type=URL; URL=\"http://en.wikipedia.org/wiki/File:
Harry_Whittier_Frees_-_What%27s_Delaying_My_Dinner.jpg\"" "http://localhost:8080/rest/my_external_datastream_obj
/extds"
HTTP/1.1 201 Created
ETaq: "cbc2d06dd76fe1b3f6379e5dc0b8212ab056443e"
Last-Modified: Thu, 12 Mar 2015 21:08:33 GMT
Link: <http://localhost:8080/rest/my_external_datastream_obj/extds/fcr:metadata>; rel="describedby"
Content-Length: 59
Location: http://localhost:8080/rest/my_external_datastream_obj/extds
Content-Type: text/plain
Server: Jetty(8.1.12.v20130726)
http://localhost:8080/rest/my_external_datastream_obj/extds
curl -i "http://localhost:8080/rest/my_external_datastream_obj/extds"
HTTP/1.1 307 Temporary Redirect
ETag: "cbc2d06dd76fe1b3f6379e5dc0b8212ab056443e"
Last-Modified: Thu, 12 Mar 2015 21:08:33 GMT
Content-Type: message/external-body; access-type=URL; url="http://en.wikipedia.org/wiki/File:
Harry_Whittier_Frees_-_What%27s_Delaying_My_Dinner.jpg"
Content-Length: 0
Accept-Ranges: bytes
Content-Disposition: attachment; filename=""; creation-date="Thu, 12 Mar 2015 21:08:33 GMT"; modification-date="
Thu, 12 Mar 2015 21:08:33 GMT"; size=0
Link: <http://www.w3.org/ns/ldp#Resource>;rel="type"
Link: <http://www.w3.org/ns/ldp#NonRDFSource>;rel="type"
Link: <http://localhost:8080/rest/my_external_datastream_obj/extds/fcr:metadata>; rel="describedby"
Allow: DELETE, HEAD, GET, PUT, OPTIONS
Location: http://en.wikipedia.org/wiki/File:Harry_Whittier_Frees_-_What%27s_Delaying_My_Dinner.jpg
Server: Jetty(8.1.12.v20130726)
```

The complete thing

```
curl -H "Accept: text/turtle" -H "Prefer: return=representation; include=\"http://fedora.info/definitions/v4
/repository#EmbedResources\";" "http://localhost:8080/rest/my_external_datastream_obj"
<http://localhost:8080/rest/my_external_datastream_obj> a ldp:RDFSource , ldp:Container , <http://www.jcp.org</pre>
/jcr/nt/1.0folder> , <http://www.jcp.org/jcr/nt/1.0hierarchyNode> , <http://www.jcp.org/jcr/nt/1.0base>
<http://www.jcp.org/jcr/mix/1.0created> , fedora:Container , fedora:Resource , fedora:Resource , <http://www.</pre>
jcp.org/jcr/mix/1.0created> , <http://www.jcp.org/jcr/mix/1.0lastModified> , <http://www.jcp.org/jcr/mix/1.
0referenceable> ;
        fedora:lastModifiedBy "bypassAdmin"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:uuid "fe7a8497-d826-490d-a786-ba601e039055" * * < http://www.w3.org/2001/XMLSchema#string > ;
        fedora:createdBy "bypassAdmin"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:primaryType "nt:folder"^^<http://www.w3.org/2001/XMLSchema#string> ;
       fedora:created "2015-03-12T20:56:08.468Z"^^<http://www.w3.org/2001/XMLSchema#dateTime> ;
        fedora:mixinTypes "fedora:Container"^^<http://www.w3.org/2001/XMLSchema#string> , "fedora:Resource"
^^<http://www.w3.org/2001/XMLSchema#string> ;
       fedora:lastModified "2015-03-12T21:08:33.787Z"^<http://www.w3.org/2001/XMLSchema#dateTime> ;
        fedora:writable "true"^^<http://www.w3.org/2001/XMLSchema#boolean> ;
        fedora:hasParent <http://localhost:8080/rest/> ;
        ldp:contains <http://localhost:8080/rest/my_external_datastream_obj/extds> .
<http://localhost:8080/rest/my_external_datastream_obj/extds> a <http://www.jcp.org/jcr/nt/1.0resource> ,
<http://www.jcp.org/jcr/mix/1.0mimeType> , <http://www.jcp.org/jcr/mix/1.0mimeType> , <http://www.jcp.org/jcr/mix/1.</pre>
0lastModified> , fedora:Binary , fedora:Resource , <http://www.jcp.org/jcr/mix/1.0created> , <http://www.jcp.org</pre>
/jcr/mix/1.0lastModified> , <http://www.jcp.org/jcr/mix/1.0referenceable> ;
        fedora:lastModifiedBy "bypassAdmin"^^<http://www.w3.org/2001/XMLSchema#string> ;
       fedora:uuid "86aa4107-768d-4e2d-9064-78be072c0888"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:mimeType "message/external-body; access-type=URL; url=\"http://en.wikipedia.org/wiki/File:
Harry_Whittier_Frees_-_What%27s_Delaying_My_Dinner.jpg\""^^<http://www.w3.org/2001/XMLSchema#string> ;
       fedora:digest <urn:sha1:da39a3ee5e6b4b0d3255bfef95601890afd80709> ;
       premis:hasSize "0"^^<http://www.w3.org/2001/XMLSchema#long> ;
        fedora:createdBy "bypassAdmin"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:primaryType "nt:resource"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:created "2015-03-12T21:08:33.787Z"^^<http://www.w3.org/2001/XMLSchema#dateTime> ;
        fedora:mixinTypes "fedora:Binary"^^<http://www.w3.org/2001/XMLSchema#string> ;
        fedora:lastModified "2015-03-12T21:08:33.787Z"^^<http://www.w3.org/2001/XMLSchema#dateTime> ;
       premis:hasOriginalName ""^^<http://www.w3.org/2001/XMLSchema#string> .
<http://localhost:8080/rest/my_external_datastream_obj> fedora:exportsAs <http://localhost:8080/rest</pre>
/my_external_datastream_obj/fcr:export?format=jcr/xml> .
<http://localhost:8080/rest/my_external_datastream_obj/fcr:export?format=jcr/xml> dc:format <http://fedora.info</pre>
/definitions/v4/repository#jcr/xml> .
```

```
<http://fedora.info/definitions/v4/repository#jcr/xml> rdfs:label "jcr/xml"^<http://www.w3.org/2001
/XMLSchema#string> .
```

Proposal 2: LDP-RS as datastreams

In this proposal, an LDP-RS-NC maps to a Fedora datastream. In the fedora object model, they are no different from the datastreams that already exist, and have no inherent characteristics that make them different in any way. The difference lies in the HTTP API, in that they support the LDP-RS interaction mode as described in the background section of this document.

An LDP-RS-NC (datastream) shall be created when the client specifies an LDPR interaction model for newly created resources as per section 5.2.3.4 of the LDP spec.

LDP to Fedora mapping

LDP Concept	Fedora Concept	LDP Interaction Modes	Constraints on Content
LDPC	Object	LDPC, LDP-RS	RDF, must contain LDP membership triples, subjects must be repository objects, subjects must be the URI of the object
LDP-NR	Datastream	LDPR for properties, LDP-NR for bitstream	none
LDP-RS-NC	Special datastream	LDPR, LDP-RS	RDF

Note that objects in the Fedora model map to the LDPC interaction model, while any resource that provides a datastream's properties and the new special type of datastream introduced here both map to the LDPR interaction model. The distinguishing factor of a Fedora object, as exposed by LDP, is that it is a container. The distinguishing factor of a special datastream, exposed in LDP, is that it is not a container (i.e. it is a leaf node in the LDP model).

Discussion

Mapping the LDP-RS-NC (non-container RDF source resources) to datastreams has several consequences as a result of the LDP-RS interaction mode defined by LDP, or as a result of their nature as a datastream

- 1. The Fedora object model does not change at all with this proposal
- The LDP-RS-NCs created by Fedora may contain arbitrary RDF. This is a consequence of their 'datastream' nature, in that their content is largely
 opaque to the Fedora model. They are blobs, so there is no inherent need to constrain their content other than to verify that they contain RDF as
 implied by an LDP-RS.
 - a. No inherent restriction on allowable subjects, predicates, or objects in the triples
 - b. No inherent restriction on constructs such as blank nodes
 - c. This implies an unknown but high cost for implementation.
- 3. The LDP-RS-NCs created to by Fedora may be 'pure' in that they are logically unmodified from client requests. Again, this is due to their nature as an opaque datastream as far as the fedora model is concerned.
 - a. No server managed triples are added by Fedora
- 4. The LDP-RS-NCs benefit from the LDP-RS interaction model as defined by LDP on these resources:
 - Retrieve triples in specified representation (xml, turtle, ntriples, etc); as per LDP spec text/turtle and application/ld+json representations are always available.
 - b. SPARQL Update semantics (PATCH)
 - c. Interpretation of relative URIs in RDF content as per the LDP spec (for example, allowing null relative URIs "<>" to refer to the to-becreated resource URI)
 - d. This also implies an unknown but high cost for implementation.

If it helps, this proposal can be also be conceptualized as a kind of 'service' provided on top of a datastream to provide an enhanced RDF CRUD API that neatly maps to an LDP concept that Fedora does not currently support, but most LDP servers do.

Use cases

- Arbitrary RDF documents as 'assets' managed by Fedora, yet accessible via the LDP-RS API. For example, Tim Berners-Lee's FOAF file.
 It contains multiple subjects, for example when providing details of linked blog posts, papers, or conferences.
- Named-graph provenance
- · A place to put statements that are not allowable in Fedora objects, as a pragmatic workaround (e.g. assertions about an external resource)

Implementation notes

- 1. Fedora may describe these resources using slightly different metadata in their corresponding 'fcr:metadata' resources.
 - a. For example, "fedora:mixinTypes "fedora:NonRdfSourceDescription"^<<u>http://www.w3.org/2001/XMLSchema#string</u>>" is probably not appropriate.
- 2. Basic validation as well-formed RDF or SPARQL/update statements is a prerequisite for accepting content.
- 3. The binary representation of an LDP-RS-NC as persisted in Fedora as datastream content may potentially be opaque (a decision would need to be made on this)
 - a. The serialization format may be specified in advance. For example, Fedora may chose to serialize as turtle, always
 - b. Fedora may perhaps chose to store the exact binary content supplied by the client via PUT, and record metadata indicating which RDF serialization had been persisted.
 - c. Likewise, the content may be entirely opaque, leaving it an implementation decision that can change at any time.

Proposal 3: LDPRs as "Pointlike objects"

In this view, LDP-RS-NCs appear in the repository as objects without datastreams or contained objects. They would possess all the functions of other objects, except that attempts to add children to them (datastreams or child objects) would fail under all circumstances. They could support the interaction described in section 5.2.3.4 of the LDP specification, if a client so specifies in the appropriate way when creating one. Their serializations would be no different from those of any other object, nor would they offer a special binary representation of any kind, other than as provided for by any object in the repository.

LDP to Fedora mapping

LDP Concept	Fedora Concept	LDP Interaction Modes	Constraints on Content
LDPC	Object	LDPC, LDP-RS	RDF, must contain LDP membership triples, subjects must be repository objects, subjects must be the URI of the object
LDP-NR	Datastream	LDPR for properties, LDP-NR for bitstream	none
LDP-RS-NC	Object	LDPR, LDP-RS	RDF, subjects must be repository objects, subjects must be the URI of the object

Note that objects in the Fedora model map to the LDP-RS interaction model, while datastreams in Fedora map to the LDPR and LDP-NR interaction models.

Proposal 4: LDPRs as "Pointlike objects" with no constraints on RDF statements.

(This is included largely for completeness sake, but it is very interesting)

This is like the "LDPRs as pointlike objects" proposal, except that it removes Fedora's existing constraints on triples associated with objects in general. In this view, an LDP-RS can be seen as a bag of triples with an identity (named graph). Fedora enhances the value of these resources by defining an ontology, and maintaining a set of triples that express its identity in the Fedora object model according to this ontology. This bestows LDP-RS resources with the identity and structure of Fedora objects. Triples within an LDP-RS that are not part of the Fedora model, and violate Fedora's assumptions (such as "subjects must be the URI of the resource") are preserved as part of the resource and named by its URI, but may be ignored by services that are build on top of the Fedora object model. Users of Fedora can embrace or extend the Fedora model, or ignore it completely.

LDP Concept	Fedora Concept	LDP Interaction Modes	Constraints on Content
LDPC	Object	LDPC, LDP-RS	RDF, must contain LDP membership triples
LDP-NR	Datastream	LDPR for properties, LDP-NR for bitstream	none
LDP-RS-NC	Object	LDPR, LDP-RS	RDF

Note that 'Objects' in the Fedora model map to the LDP-RS interaction model, while 'Datastreams' in Fedora map to the LDP-NR interaction model. The distinguishing factor of a Fedora object, as exposed by LDP, is that is an RDF source. The distinguishing factor of the resource that provides a datastream's bitstream is that it is a non-RDF source.

Discussion

In this proposal, Fedora's relationship with LDP can be seen as a "value add service:" in creating an LDP-RS, Fedora enhances the value of these LDP resources by managing a set of triples that give it an identity in the Fedora model, describe it as such, and enable services on it such as fixity, indexing, etc.

- Any statements outside this ontology are preserved, but largely ignored by the Fedora service stack (except, perhaps general-purpose services such as triple store indexing).
- To an outside user, Fedora would appear just like a fully-featured LDP service that happens to have a lot of server-managed triples. • Fedora would be completely usable as a generic LDP service without any knowledge of the Fedora model
 - An outside user would not need to understand the Fedora model in order to understand why an LDP request failed.
- To a current user of Fedora 4, there would be no observable difference in an object and how it behaves, other than the fact that triples will not be
 rejected unless they contradict the Fedora ontology

Use cases

- Allow Fedora to be used with models that make use of assertions about non-repository objects (e.g. Hydra use case, tamsin johnson?)
- Allow Fedora to coexist with models that are unrelated to the Fedora ontology
- Decouple Fedora from LDP
- Fedora could be thought of as a model and set of services that exists on top of LDP, rather than as part of it.
- Named graphs for provenance would be a bit of a challenge unless there were a way to distinctly separate triples from the Fedora ontology/model from those provided by the user
 - This is where the mailing list suggestion of "a way to exclude server managed triples" when GETting RDF from an object may be relevant, *if* support of that use case natively is a goal.

Implementation notes

- According to JIRA and the mailing list, this would be difficult to implement, given the way Fedora objects are currently expressed as JCR nodes
 - I believe triples are currently expressed as properties in the JCR model. Do Fedora objects (not datastreams) have any JCR binary content? If not, could that be leveraged to implement something like this (e.g. serialize all statements that do not have a repository subject as binary)