

Fedora

CS 431 - April 17, 2006

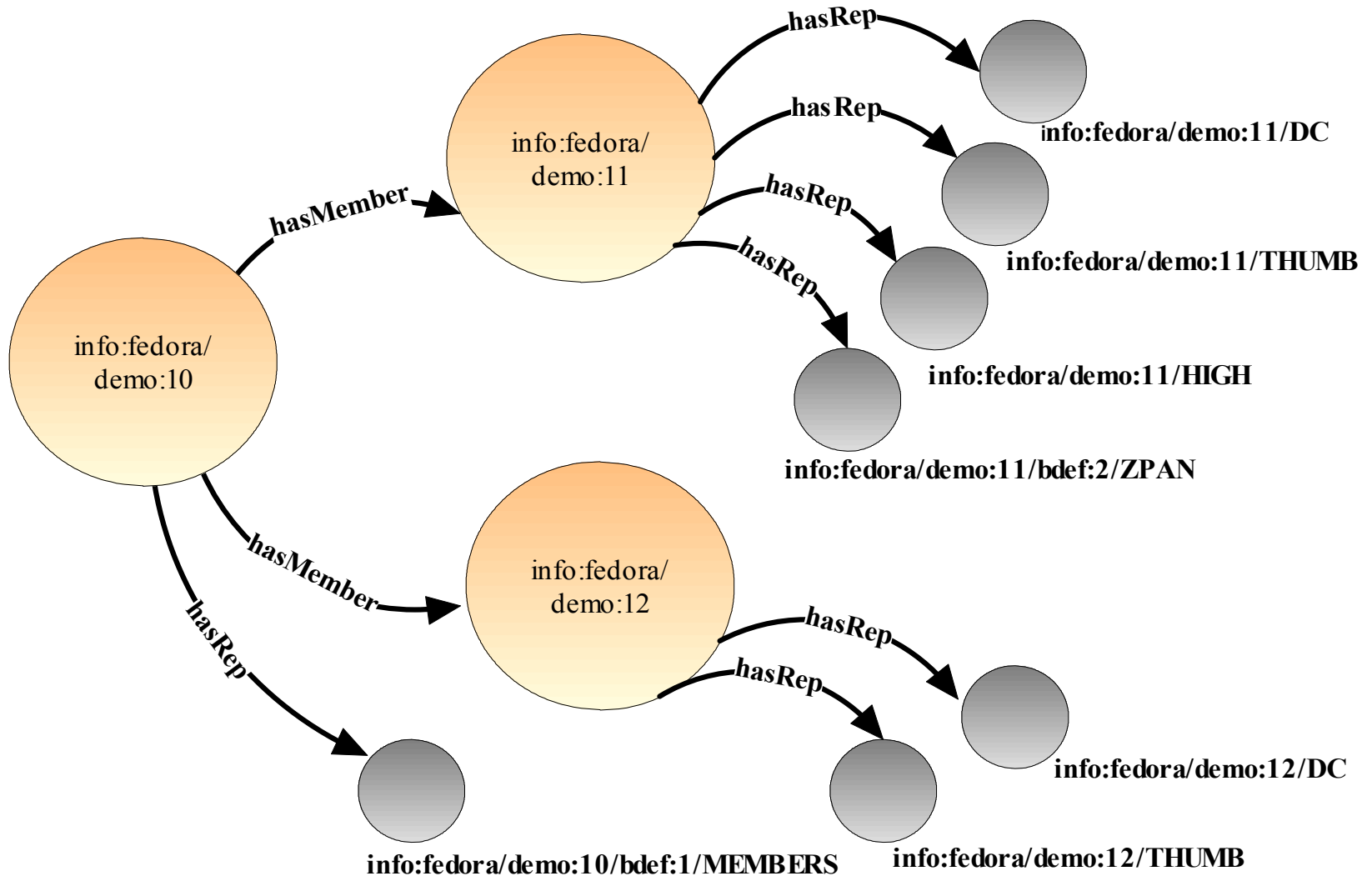
Carl Lagoze - Cornell University

Acknowledgements:

Sandy Payette (Cornell)

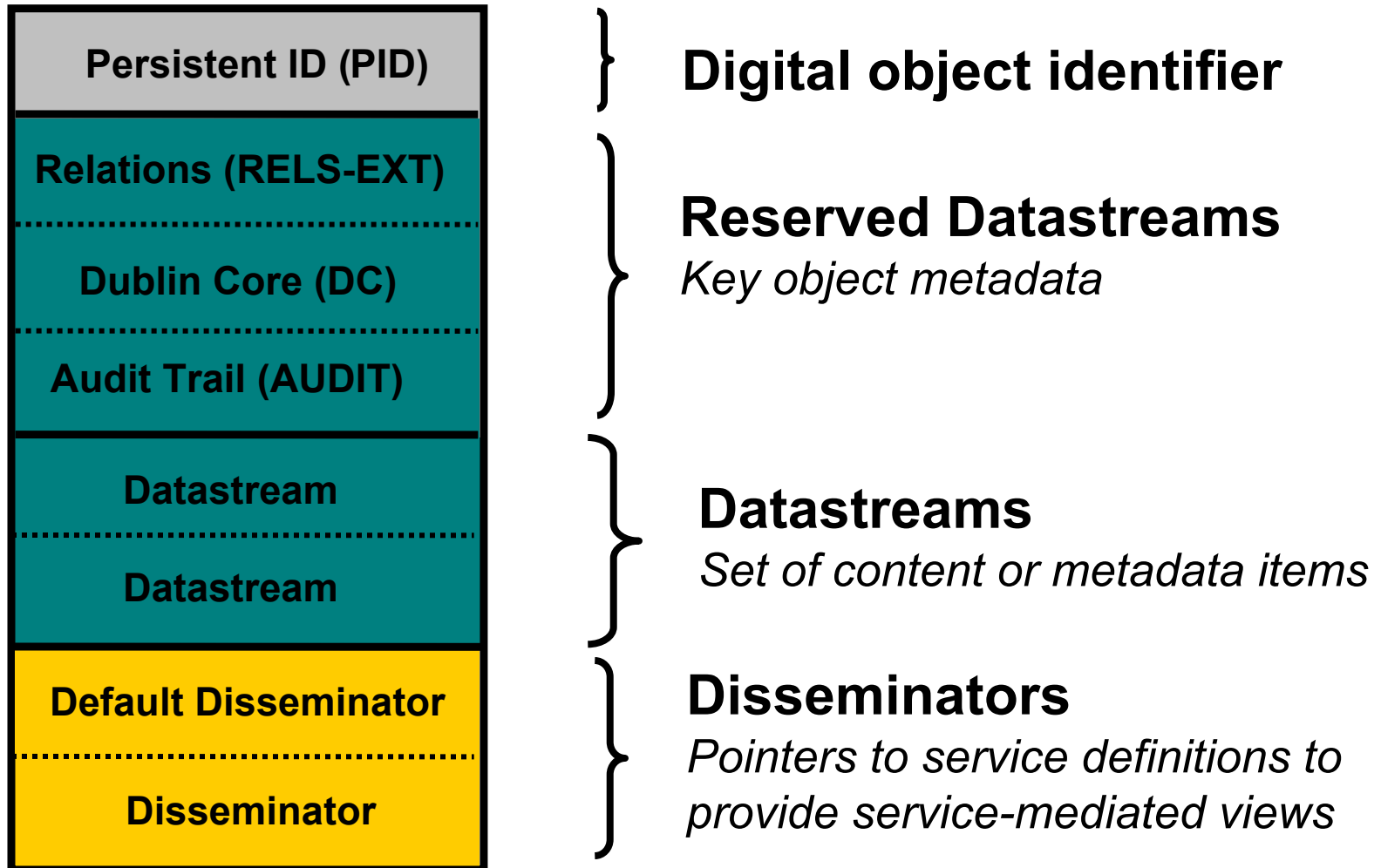
Digital Object Model

"Graph" View of Fedora Objects



Fedora Digital Object Model

Component View



The Datastream Component

4 Classifications for Datastreams

Inline XML

Fedora stores a name-spaced block of XML content within the Fedora digital object XML file.

Managed Content

Fedora stores and manages the content bytestream (non-XML content)

External Referenced

Fedora stores a reference (URL) to the content

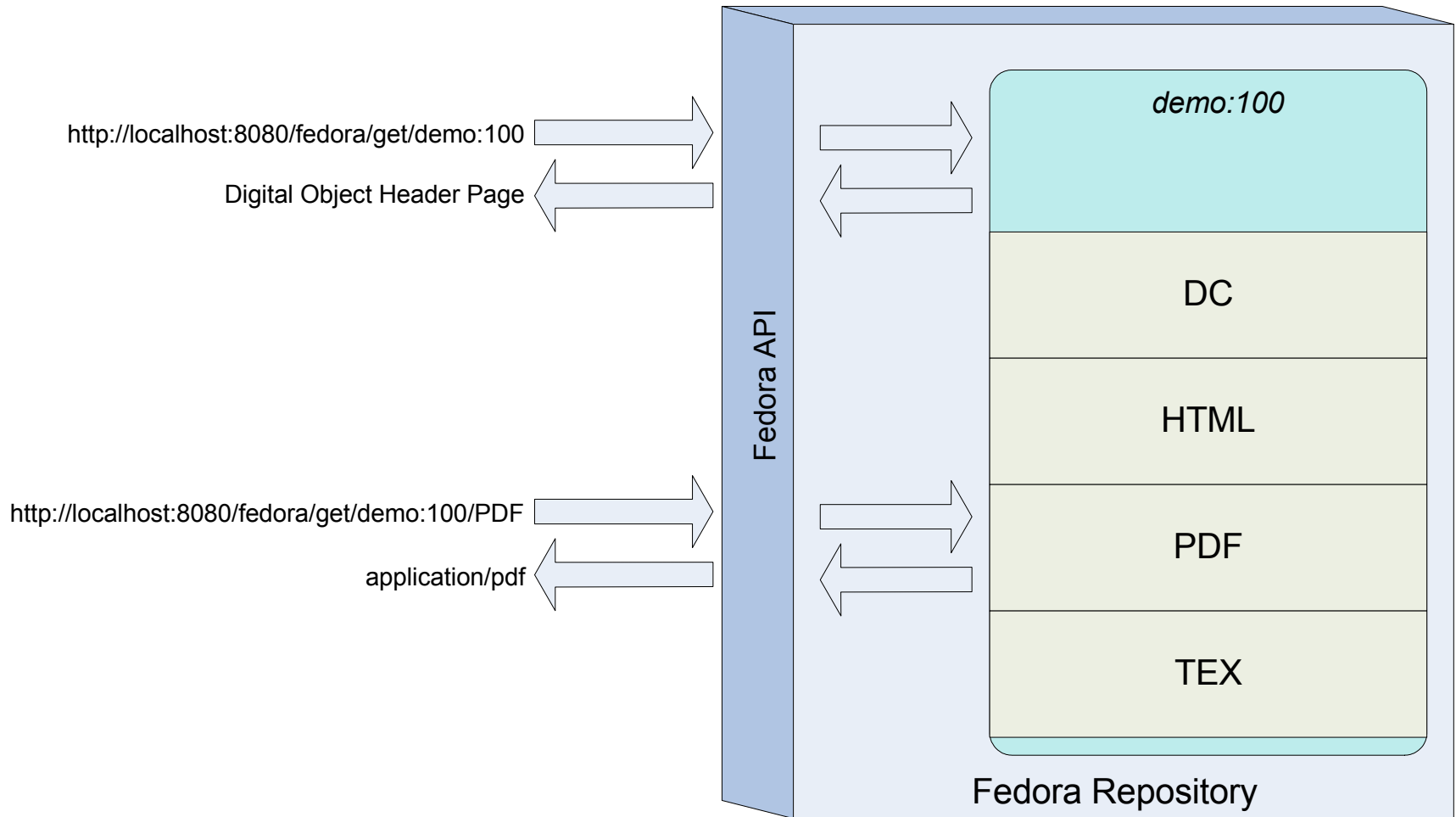
External Redirected

Fedora stores a reference (URL) to the content, but will not mediate access to content. (Optimized for streaming)

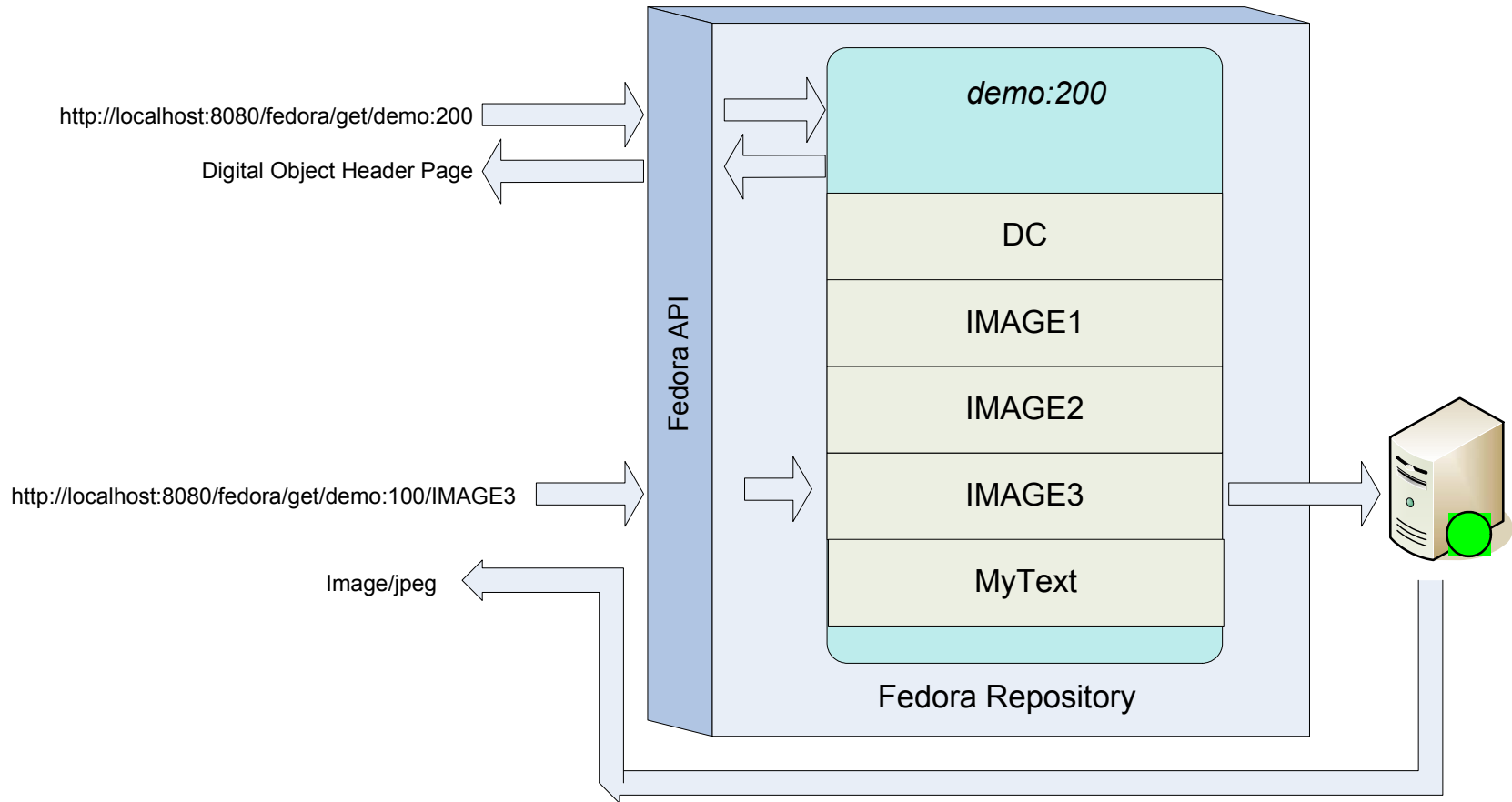
Simple Fedora model for aggregating static content

- Representations map to datastreams
- Datastreams may be local or surrogates (redirect) to remote data
- REST URL's give client access to representations

Digital Object Aggregating Local Content



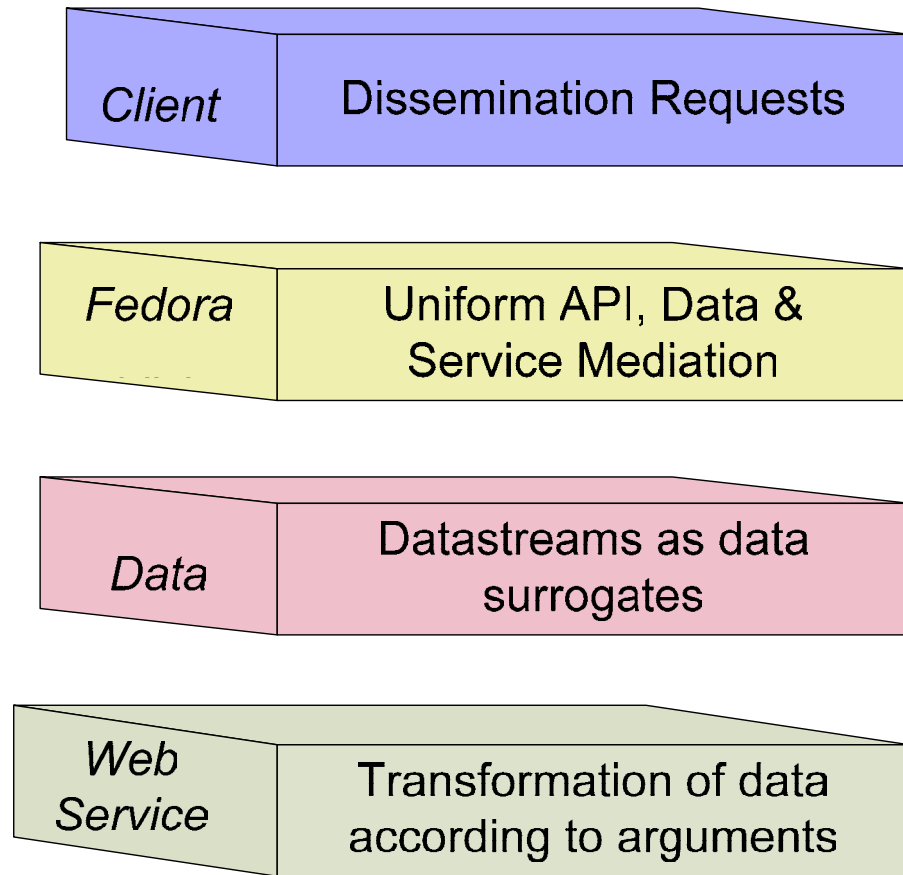
Digital Object for Local and Remote Content



Fedora for dynamic content

- Representations map to service-based transforms of data (in addition to static datastreams)
- Opaque to REST based access (client see only representations, not how they are produced)
- Motivating examples
 - Canonical XML metadata format - XSLT to Dublin Core
 - Document source in TeX, programmatic transform to PDF, PS, HTML, etc.

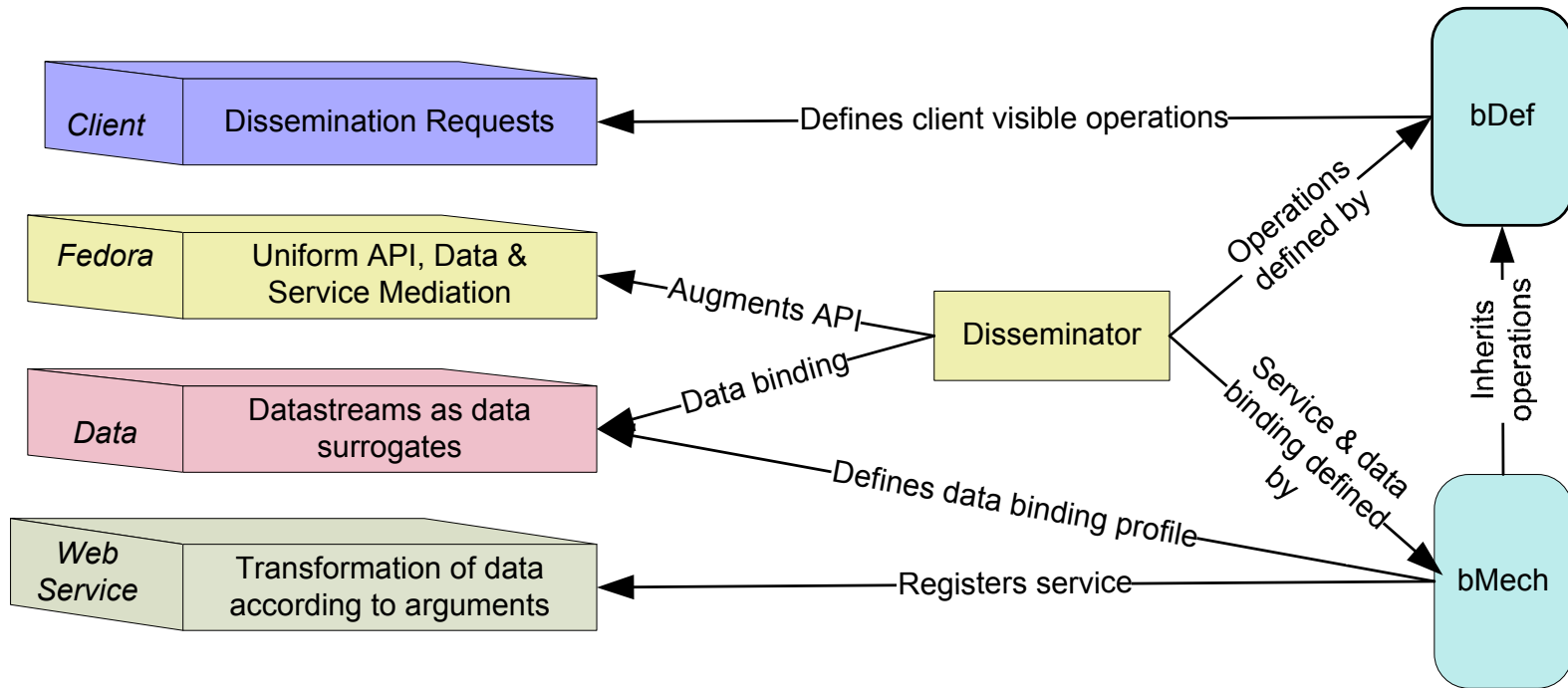
Understanding Dynamic Disseminations (1)



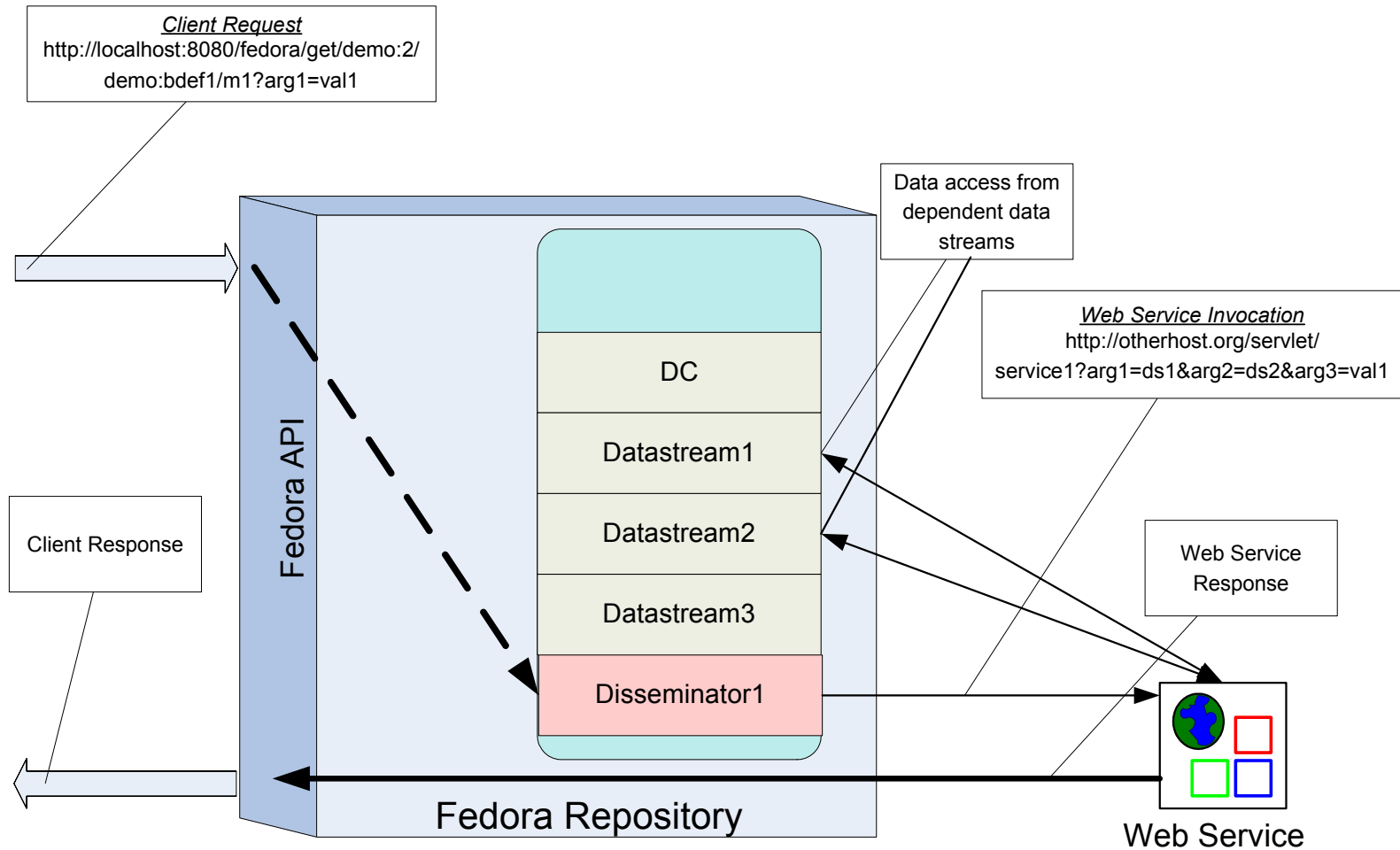
Understanding Dynamic Disseminations (2)

- Behavior Definitions (bDef)
 - Special digital object defining client side functionality (method template)
- Behavior Mechanism (bMech)
 - Special digital object that refines a bDef by defining:
 - Data profile: set of datastreams required for execution
 - Service binding: where the work is performed
 - May be many bMechs for a bDef
- Disseminator
 - Association of a bMech/bDef with a digital object endowing it with bDef-defined functionality (methods)
 - A digital object may have multiple disseminators (polymorphic typing)

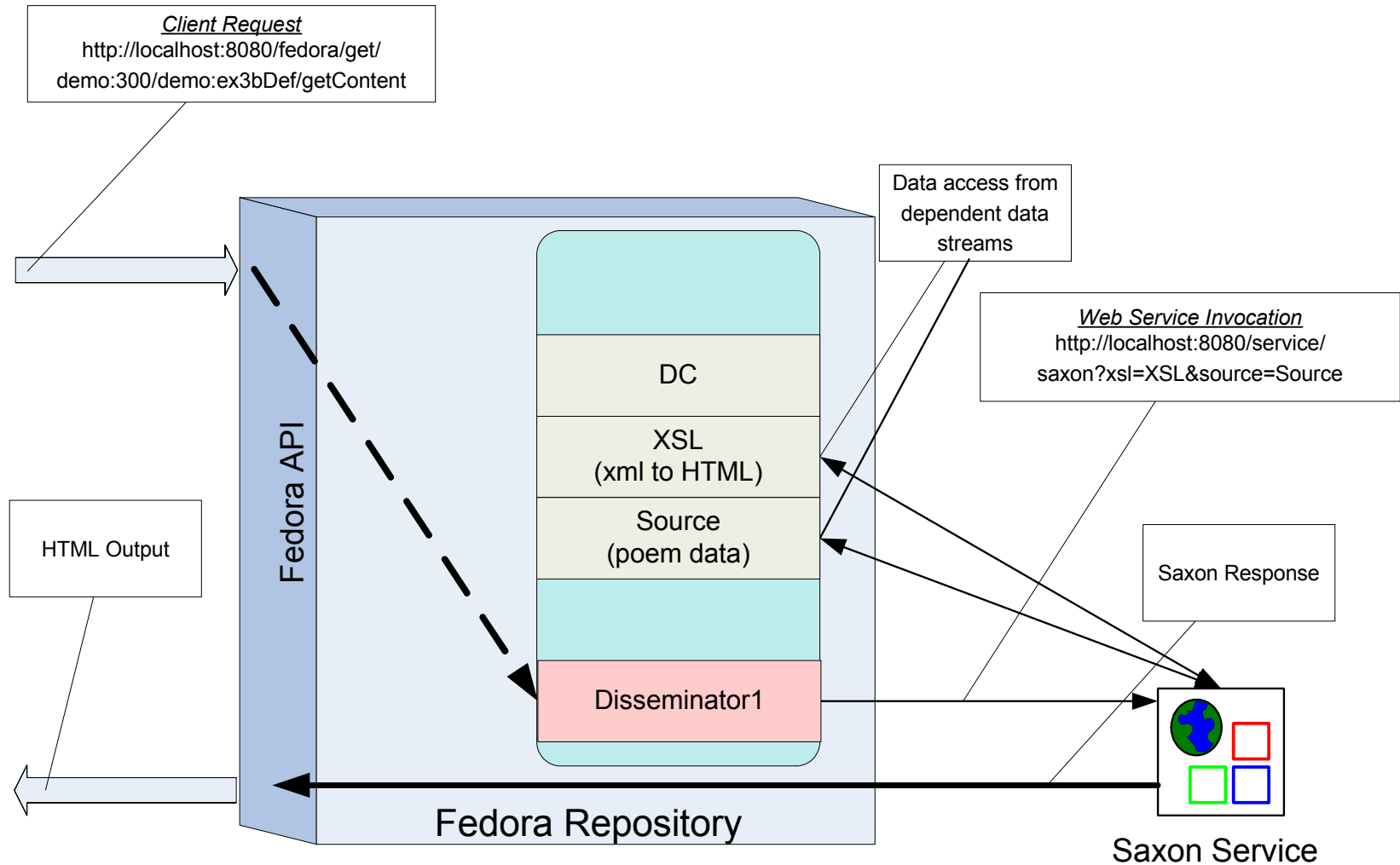
Understanding Dynamic Disseminations (3)



Dynamic Dissemination Access



Dynamic Dissemination Example



Fedora - XML for digital objects

- **FOXML (Fedora Object XML)**
 - Simple XML format directly expresses Fedora object model
 - Easily adapts to Fedora new and planned features
 - Easily translated to other well-known formats
 - Internal storage format for objects in repository
- **XML-based Ingest/Export of objects**
 - FOXML, METS (Fedora extension)
 - Extensible to accommodate new XML formats
 - Planned: METS 1.4, MPEG21 DIDL

FOXML - Object Properties

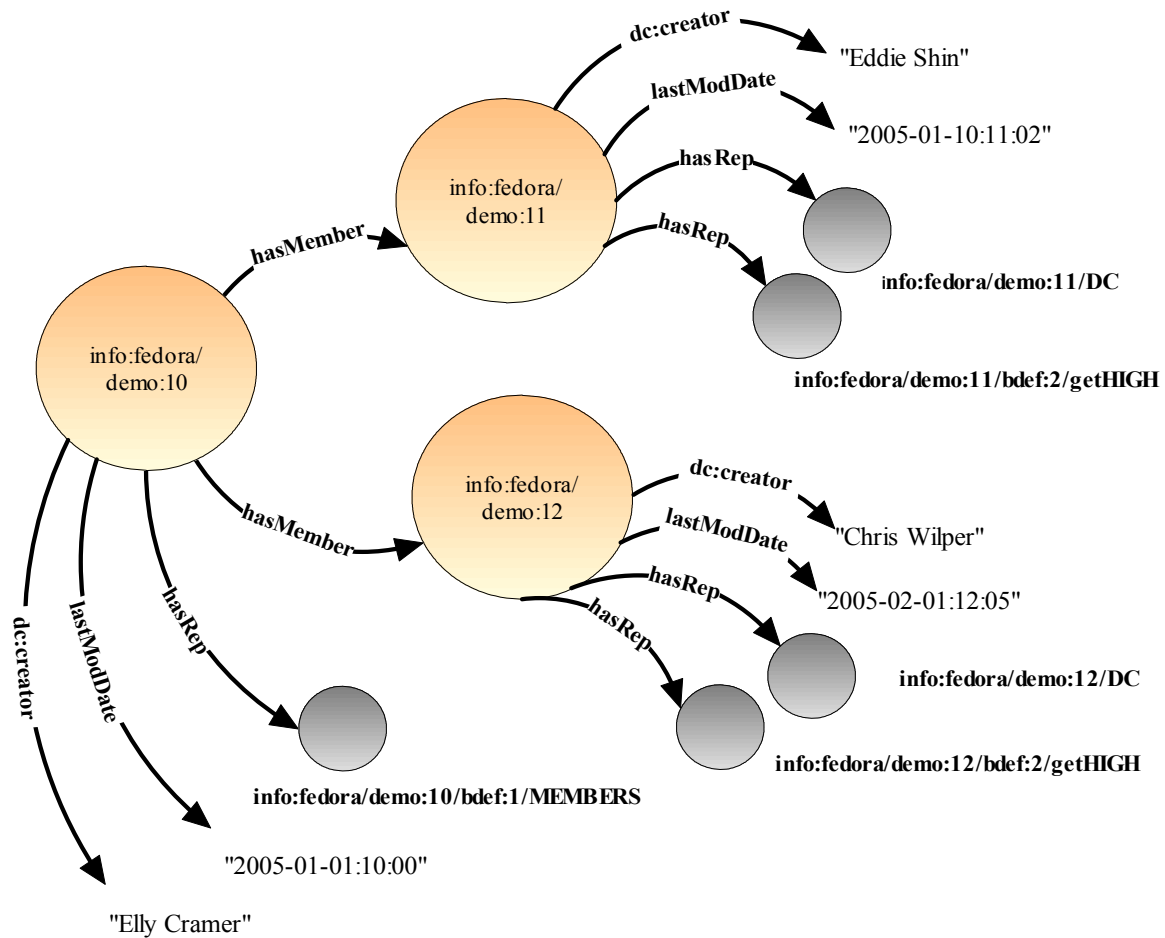
```
<foxml:objectProperties>  
  <foxml:property NAME="http://www.w3.org/1999/02/22-rdf-syntax-ns#type" VALUE="FedoraObject"/>  
  <foxml:property NAME="info:fedora/fedora-system:def/model#state" VALUE="A" />  
  <foxml:property NAME="info:fedora/fedora-system:def/model#label" VALUE="Sandy's Test Object"/>  
  <foxml:property NAME="info:fedora/fedora-system:def/model#contentModel" VALUE="TEST"/>  
</foxml:objectProperties>
```


Fedora Resource Index:

Using RDF and ontologies

Fedora Digital Objects

Resource Index View



Fedora 2.0 and RDF

- **Object-to-object and object-to-literal Relationships**
 - Ontology of common relationships (RDF schema)
 - Relationships stored in special datastream (RELS-EXT)
- **Resource Index (RI)**
 - RDF-based index of repository (Kowari triple-store)
 - Graph-based index includes:
 - Object properties and Dublin Core
 - Object Relationships
 - Object Disseminations
- **RI Search**
 - Powerful querying of graph of inter-related objects
 - REST-based query interface (using RDQL or ITQL)
 - Results in different formats (triples, tuples, sparql)

Uses of Object Relationships

- Define collections (e.g., collection objects)
- Assert critical relationships among object for management purposes
- Enable network overlay
 - Surrogate objects referring to external entities
 - Assert relationships among them
 - Assert other relationships (e.g., annotations)
- Enable navigation of repository (as tree or graph)

Fedora Relationship Ontology (RDFS)

- isPartOf / hasPart
- isMemberOf / hasMember
- isDescriptionOf / hasDescription
- hasEquivalent
- ... others

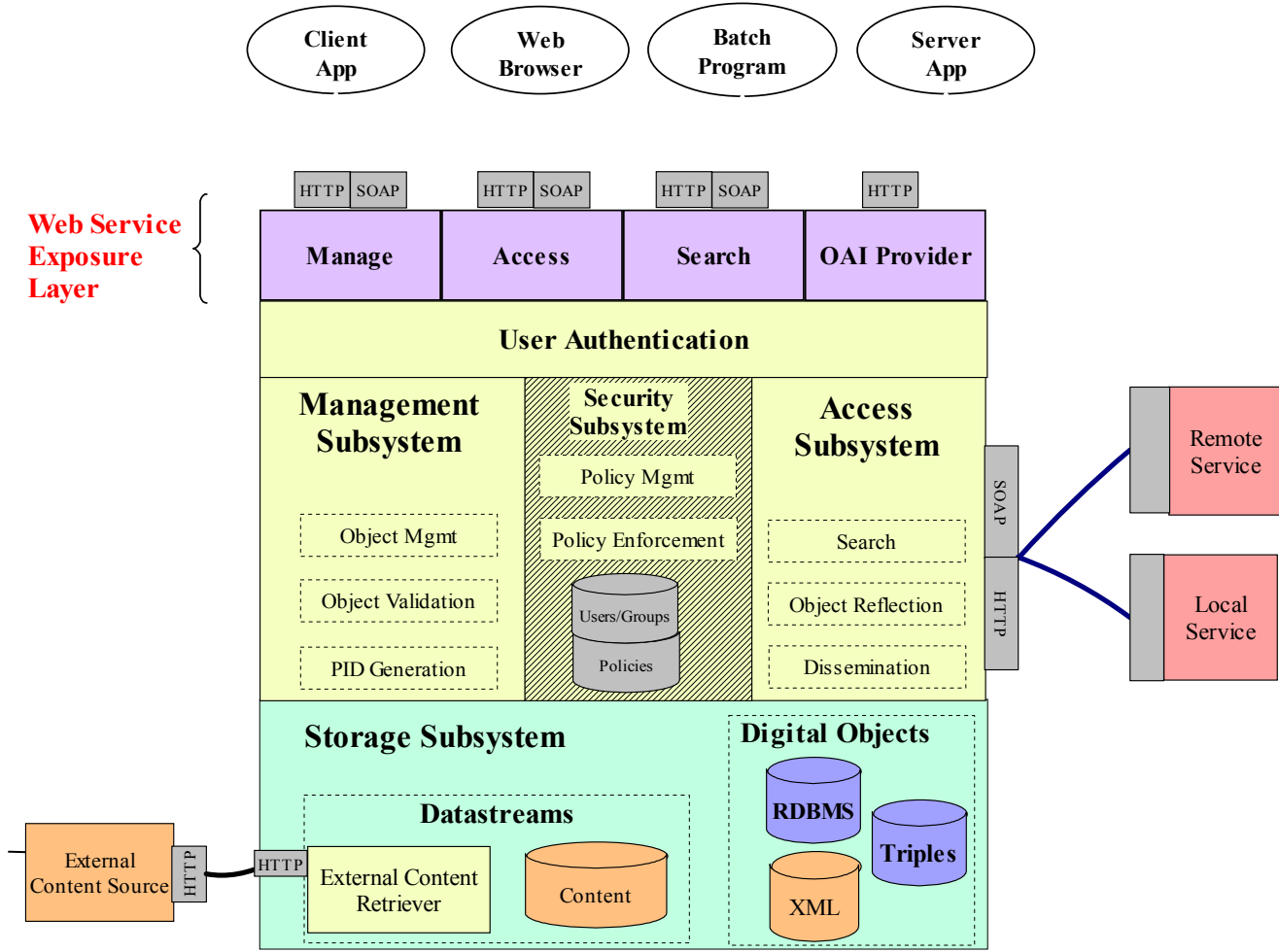
Demo:

Collection - Member Relationships

- Collection Object [[smiley](#)]
 - Datastream containing a query to Resource Index for all members of collection
- Image Objects [[brush](#)]
 - Use RELS-EXT datastream to assert relationship to collection object

Fedora Repository Service

Fedora Repository Service



Fedora Repository: 3 Layers

<h2>1. Interfaces</h2>	<ul style="list-style-type: none">• Access/Search Service• Management Service• OAI Provider Service• Resource Index Service
<h2>2. Modules</h2>	<p>Configurable modules that implement all repository functionality in terms of the Fedora digital object model.</p>
<h2>3. Persistent Store</h2>	<ul style="list-style-type: none">• RDBMS<ul style="list-style-type: none">- Digital object registry- Object "cache" for performance• File System<ul style="list-style-type: none">- XML object serializations- Managed Content (Datastreams)

Fedora Web Service APIs in a Nutshell

- **Management Service (API-M)**
 - Ingest Object
 - Export Object
 - Get Object XML
 - Purge Object
 - Modify Object
 - Get Next PID

 - Get Datastream(s)
 - Get DatastreamHistory
 - **Get DisseminatorHistory**
 - Get Disseminator(s)

 - Add/modify/purge Datastream
 - Add/modify/purge Disseminator
 - Set State

Fedora Web Service APIs in a Nutshell

- **Access Service (API-A and API-A-LITE)**
 - Describe Repository
 - Get Object Profile
 - Get Object History
 - Get Datastream
 - Get Dissemination

 - Find Objects
 - Resume Find Objects

Fedora Web Service APIs in a Nutshell

- **API-A-Lite**
 - **Repository-level operations:**
 - **fedora/describe** - Describe Repository
 - **fedora/search** - methods to locate objects via the default repository index
 - **Object-level operations:**
 - **fedora/get** - method to get object profile
 - **fedora/get/..** - method to "disseminate" a view of an object's content
 - **Fedora/getMethods** - methods get information about all disseminations available on object
- **OAI-PMH Provider Service**
 - **All OAI-PMH methods to harvest OAI-DC from each object**

Fedora 2.0 - Clients

Fedora Administrator (via Fedora SOAP interfaces)

- Java Swing client
- Ingest/Export objects
- Batch creation and modification of objects
- One-up creation and modification of objects
- Search repository
- Wizards for creating BDEF/BMECH objects

• Web Browser (via Fedora REST interfaces)

- Access, Search,
- OAI
- Resource Index
- Selected management operations

• Command Line Utilities

- Ingest, export, purge
- Migration

Fedora Software Distribution

- **Open Source (Mozilla Public License)**
- **100% Java (Sun Java J2SDK1.4)**
- **Supporting Technologies**
 - Apache Tomcat and Apache Axis (SOAP)
 - Xerces for XML parsing and validation
 - Saxon for XSLT transformation
 - Schematron for validation
 - MySQL and Mckoi relational database
 - Oracle 9i support
 - Kowari for triple-store
- **Deployment Platforms**
 - Windows 2000, NT, XP
 - Solaris
 - Linux
 - Mac OSX

Fedora 2.1.1 (April 2006)

- Authentication plug-ins
 - HTTP basic authentication and SSL
 - Plug-in #1 : user/password file
 - Plug-in #2 : LDAP tie-in
 - Plug-in #3 : Radius Authentication
- Authorization module
 - XACML policy enforcement for API operations
- New OAI Provider (stand-alone service)
- Support for MPEG21-DIDL (ingest/export/oai)
- Performance testing and improvements
- Integration with Storage Resource Borker