

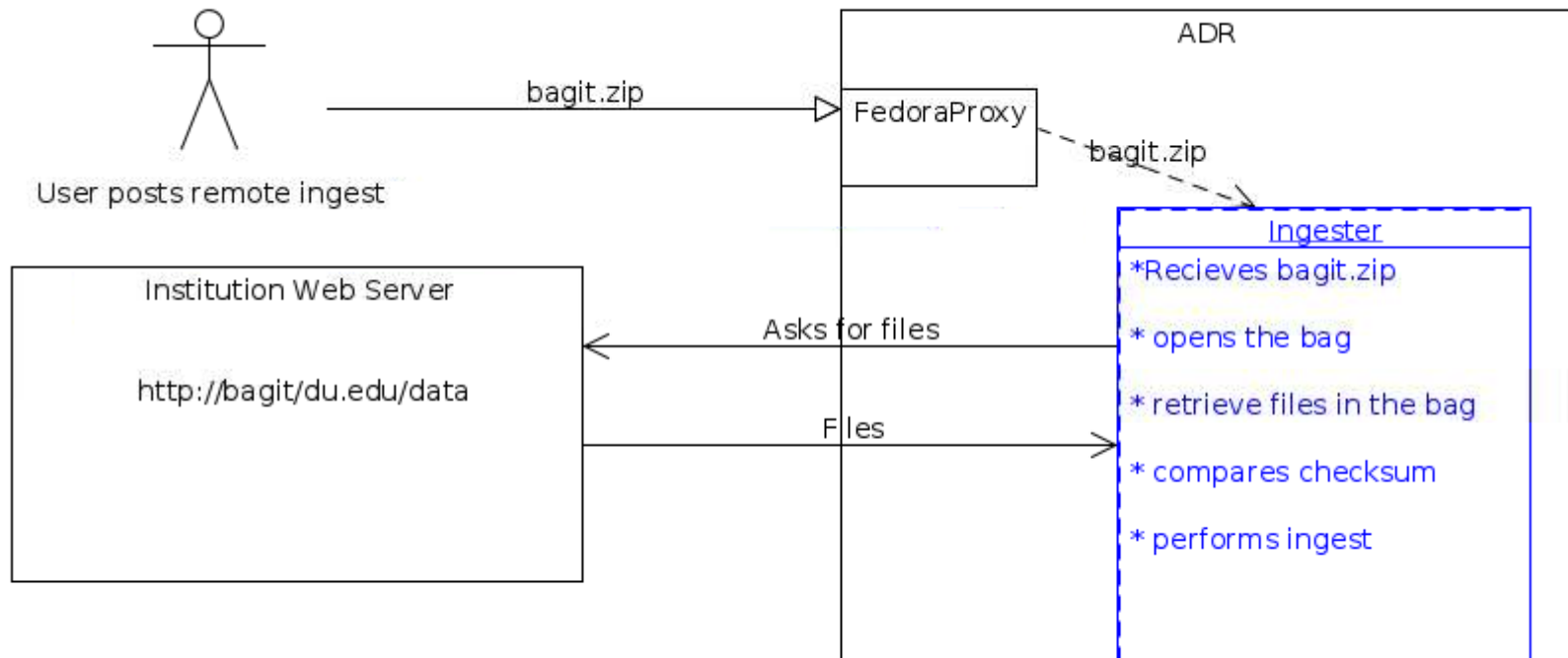
DU

REMOTE

Bagit Ingest

Nov 2012

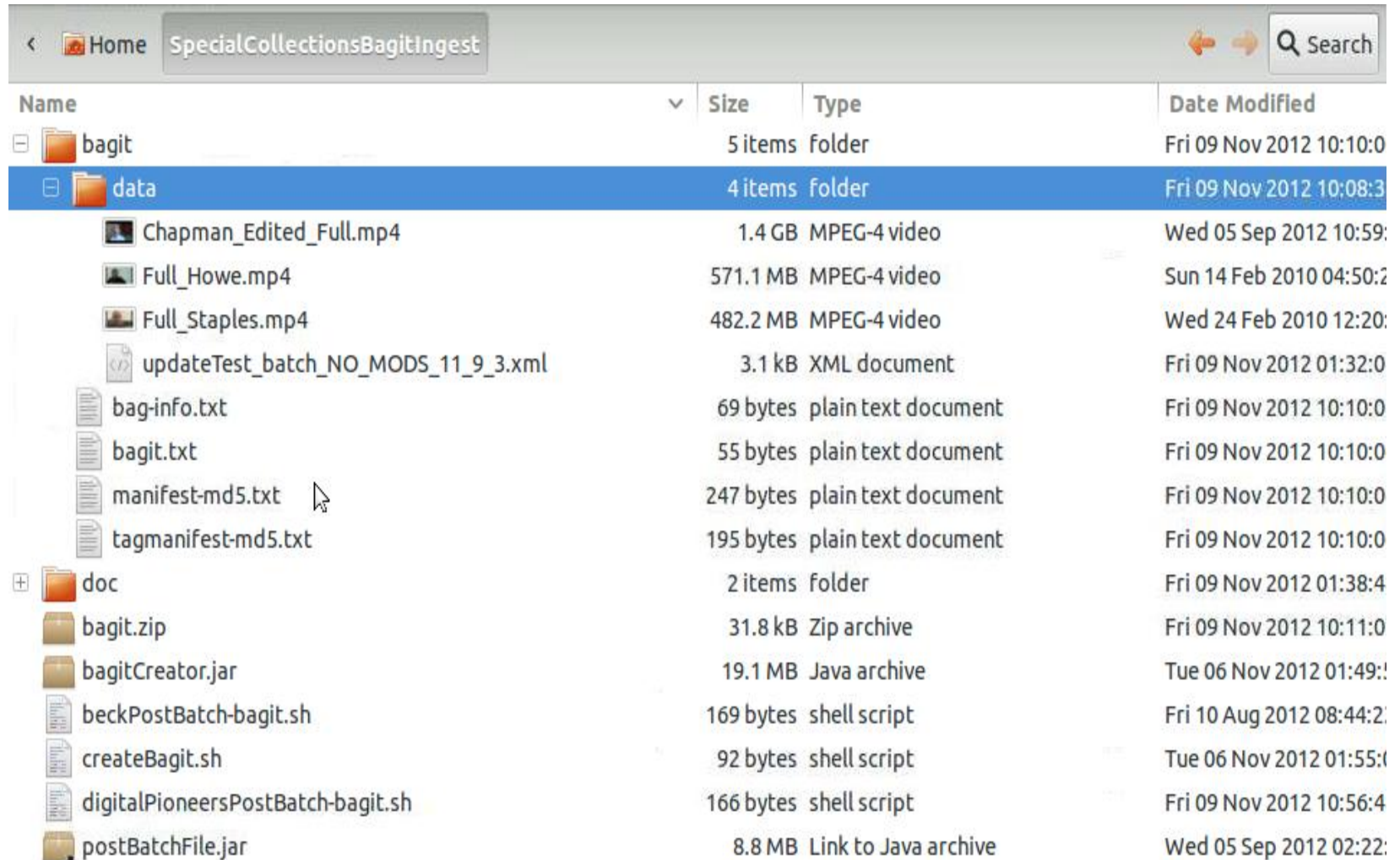
Diagram of a bagit ingest



Steps to perform a remote ingest

- Copy files to be ingested into the bagit data directory (data directory is assessable with <http://bagit.du.edu/data>)
- Create the bagit.zip file
- Post the bagit.zip file to the ADR server

Copy files into the bagit data directory



Name	Size	Type	Date Modified
bagit	5 items	folder	Fri 09 Nov 2012 10:10:0
data	4 items	folder	Fri 09 Nov 2012 10:08:3
Chapman_Edited_Full.mp4	1.4 GB	MPEG-4 video	Wed 05 Sep 2012 10:59:
Full_Howe.mp4	571.1 MB	MPEG-4 video	Sun 14 Feb 2010 04:50:2
Full_Staples.mp4	482.2 MB	MPEG-4 video	Wed 24 Feb 2010 12:20:
updateTest_batch_NO_MODS_11_9_3.xml	3.1 kB	XML document	Fri 09 Nov 2012 01:32:0
bag-info.txt	69 bytes	plain text document	Fri 09 Nov 2012 10:10:0
bagit.txt	55 bytes	plain text document	Fri 09 Nov 2012 10:10:0
manifest-md5.txt	247 bytes	plain text document	Fri 09 Nov 2012 10:10:0
tagmanifest-md5.txt	195 bytes	plain text document	Fri 09 Nov 2012 10:10:0
doc	2 items	folder	Fri 09 Nov 2012 01:38:4
bagit.zip	31.8 kB	Zip archive	Fri 09 Nov 2012 10:11:0
bagitCreator.jar	19.1 MB	Java archive	Tue 06 Nov 2012 01:49:!
beckPostBatch-bagit.sh	169 bytes	shell script	Fri 10 Aug 2012 08:44:2
createBagit.sh	92 bytes	shell script	Tue 06 Nov 2012 01:55:!
digitalPioneersPostBatch-bagit.sh	166 bytes	shell script	Fri 09 Nov 2012 10:56:4
postBatchFile.jar	8.8 MB	Link to Java archive	Wed 05 Sep 2012 02:22:

Create the bagit.zip file

chet@lib-ram:

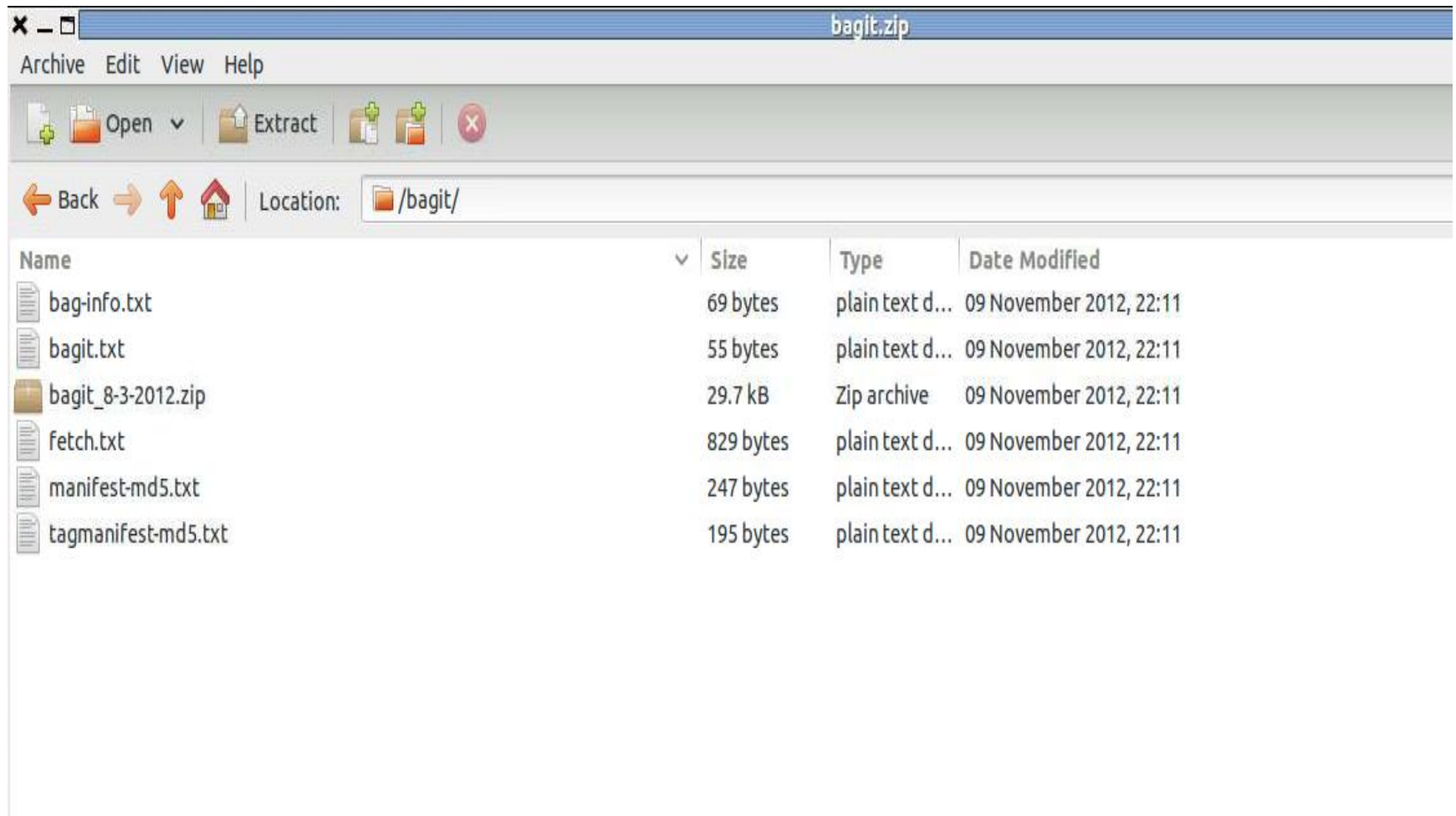
cd SpecialCollectionsBagitIngest/

./createBagit.sh

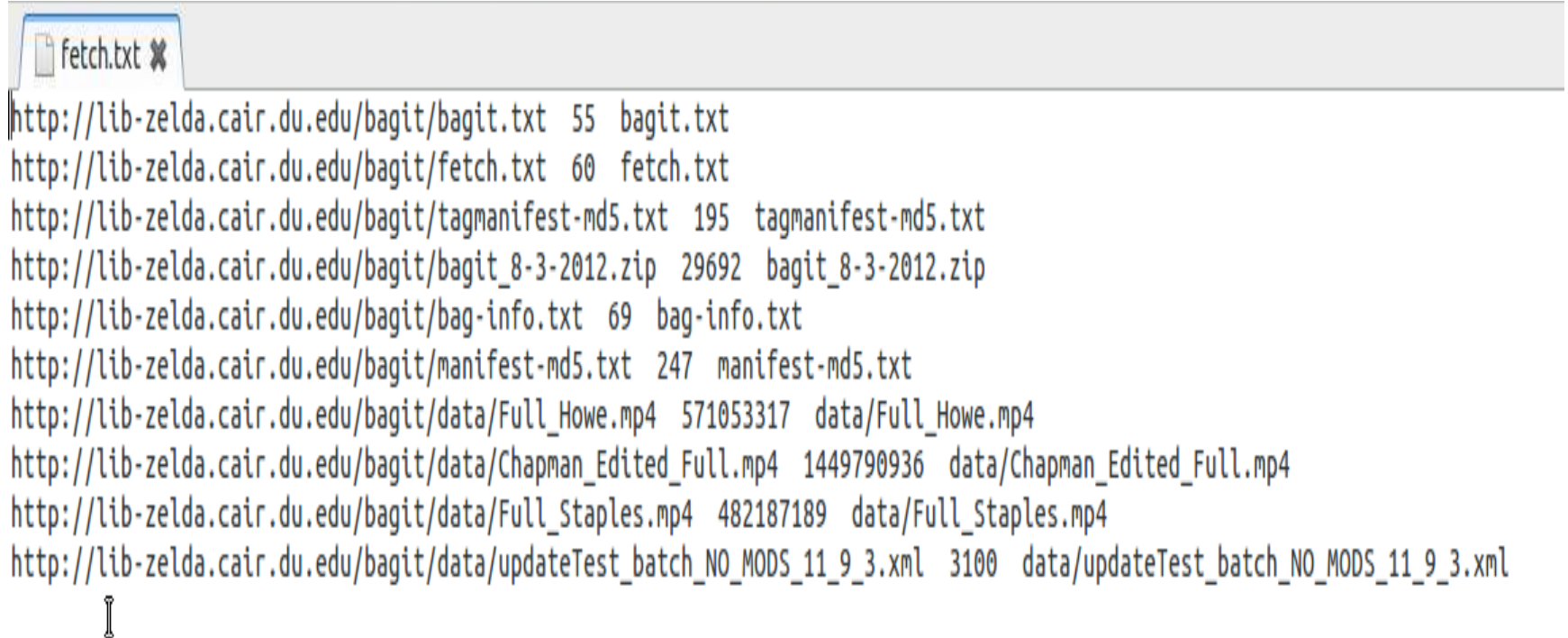
```
[INFO,DefaultBag,main] gov.loc.repository.bagger.bagit.impl.DefaultBag: DefaultBag.init file: null, version: 0.96
[INFO,PreBagImpl,main] Making a bag in place at ./bagit
[INFO,FileSystemWriter,main] Writing bag
[INFO,HolePuncherImpl,main] Making bag holey
[INFO,FetchTxtImpl,main] Creating new fetch.txt.
[INFO,CompleteVerifierImpl,main] Completed verification that bag is complete.
[INFO,CompleteVerifierImpl,main] Note that this a verification of completeness, not validity. A bag may be complete without being valid, though a valid bag must be complete.
[INFO,CompleteVerifierImpl,main] Result of verification that complete: Result is true. missingAndInvalidFiles: []
[INFO,ValidVerifierImpl,main] Completed verification that bag is valid.
[INFO,ValidVerifierImpl,main] Validity check: Result is true. missingAndInvalidFiles: []
[INFO,DefaultBag,main] Bag-Info to write: {Bag-Size=2.3 GB, Bagging-Date=2012-11-09, Payload-Oxum=2503034542.4}
[INFO,ZipWriter,main] Writing bag
[INFO,FetchTxtImpl,main] Creating fetch.txt.
[INFO,DefaultBag,main] gov.loc.repository.bagger.bagit.impl.DefaultBag: DefaultBag.init file: ./bagit, version: 0.96
[INFO,CompleteVerifierImpl,main] Completed verification that bag is complete.
[INFO,CompleteVerifierImpl,main] Note that this a verification of completeness, not validity. A bag may be complete without being valid, though a valid bag must be complete.
[INFO,CompleteVerifierImpl,main] Result of verification that complete: Result is true. missingAndInvalidFiles: []
[INFO,ValidVerifierImpl,main] Completed verification that bag is valid.
[INFO,ValidVerifierImpl,main] Validity check: Result is true. missingAndInvalidFiles: []
```

valid bag

Contents of the bag



Contents of the fetch.txt file



```
http://lib-zelda.cair.du.edu/bagit/bagit.txt 55 bagit.txt
http://lib-zelda.cair.du.edu/bagit/fetch.txt 60 fetch.txt
http://lib-zelda.cair.du.edu/bagit/tagmanifest-md5.txt 195 tagmanifest-md5.txt
http://lib-zelda.cair.du.edu/bagit/bagit_8-3-2012.zip 29692 bagit_8-3-2012.zip
http://lib-zelda.cair.du.edu/bagit/bag-info.txt 69 bag-info.txt
http://lib-zelda.cair.du.edu/bagit/manifest-md5.txt 247 manifest-md5.txt
http://lib-zelda.cair.du.edu/bagit/data/Full_Howe.mp4 571053317 data/Full_Howe.mp4
http://lib-zelda.cair.du.edu/bagit/data/Chapman_Edited_Full.mp4 1449790936 data/Chapman_Edited_Full.mp4
http://lib-zelda.cair.du.edu/bagit/data/Full_Staples.mp4 482187189 data/Full_Staples.mp4
http://lib-zelda.cair.du.edu/bagit/data/updateTest_batch_NO_MODS_11_9_3.xml 3100 data/updateTest_batch_NO_MODS_11_9_3.xml
```

Send the bag to the ADR

chet@lib-ram:

`./digitalPioneersPostBatch-bagit.sh`

ADR receives bag and extracts its contents

Fedora Batch Ingest - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Connecting...

localhost:7080/FedoraApp/viewRunningIngests.htm

Google

Rocky Mountain Smal... Weebly - Create a fre... ADR Basic Document... Air Care Colorado :: V... The Ocean Explorer

Fedora Ingest Client 1.60

[Home](#) **Ingest Tasks (manual, _REMOTE and background _TASK)** 5 Second Refresh

BatchSetName	Institution	BatchSet	Status
demo_fedoraAppDemoCollection_REMOTE	demo	fedoraAppDemoCollection	Extacting bag batch_bagit_REMOTE.zip
codu_mixed_TASK	codu	mixed	Checking to see if there is anything to ingest

ADR reads the contents of fetch.txt and fetches all needed files from the Institutions server

Fedora Batch Ingest - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Connecting...

localhost:7080/FedoraApp/viewRunningIngests.htm

Google

Rocky Mountain Smal... Weebly - Create a fre... ADR Basic Document... Air Care Colorado :: V... The Ocean Explorer

Fedora Ingest Client 1.60

[Home](#) **Ingest Tasks (manual, _REMOTE and background_TASK)** 5 Second Refresh

BatchSetName	Institution	BatchSet	Status
demo_fedoraAppDemoCollection_REMOTE	demo	fedoraAppDemoCollection	Fetching http://localhost:9080/bagit/data/diver-with-manatee-trichechus-manatus-latirostris_access.jpg
codu_mixed_TASK	codu	mixed	Checking to see if there is anything to ingest

Ingest retries until unable to retrieve files

Fedora Batch Status - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Fedora Batch Status

localhost:7080/fedoraApp/edu.du.penrose.systems.fedoraApp.web.gwt.batchIngest.Bat

Google

Rocky Mountain Smal... Weebly - Create a fre... ADR Basic Document... Air Care Colorado :: V... The Ocean Explorer

[Home](#) | [Start New Ingest](#) | [View Running Ingests](#)

Added: 0
Failed: 0
Updated:
Failed:

Collection: demo:fedoraAppCollection
Content Model: demo:fedoraAppCM

| Institution: demo Batch Set: demo_fedoraAppDemoCollection_REMOTE

Status: **ERROR: Unable to fetch File:http://localhost:9080/bagit/data/diver-with-manatee-trichechus-manatus-latirostris_access.jpg-Connection refused**

Stop Ingest

v0.26; Start time: fri, nov 9, 2012, 9:50 pm

WARNING: Access to Handle Server is DISABLED.

Retrieving Files

Fedora Batch Status - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Fedora Batch Status aquatic-animal-mammal-un...

localhost:7080/fedoraApp/edu.du.penrose.systems.fedoraApp.web.gwt.batchIngest.Bat Google

Rocky Mountain Smal... Weebly - Create a fre... ADR Basic Document... Air Care Colorado :: V... The Ocean Explorer

[Home](#) | [Start New Ingest](#) | [View Running Ingests](#)

Added: 0
Failed: 0
Updated: 0
Failed: 0
Collection: demo:fedoraAppCollection
Content Model: demo:fedoraAppCM

/ Institution: **demo** Batch Set: **demo_fedoraAppDemoCollection_REMOTE**

Status: **Fetching http://localhost:9080/bagit/data/aquatic-animal-mammal-underwater-manatee_access.jpg**

Stop Ingest

v0.26; Start time: fri, nov 9, 2012, 9:50 pm

WARNING: Access to Handle Server is DISABLED.

Validating the complete bagit.zip which is now at the ADR

The screenshot shows a web browser window titled "Fedora Batch Status - Mozilla Firefox". The address bar displays the URL "localhost:7080/fedoraApp/edu.du.penrose.systems.fedoraApp.web.gwt.batchIngest.Bat". The browser's tab bar shows two tabs: "Fedora Batch Status" and "aquatic-animal-mammal-un...". The browser's toolbar includes a search bar with "Google" and several bookmarked sites: "Rocky Mountain Smal...", "Weebly - Create a fre...", "ADR Basic Document...", "Air Care Colorado :: V...", and "The Ocean Explorer".

The main content area of the web application displays the following information:

- Navigation links: [Home](#) | [Start New Ingest](#) | [View Running Ingests](#)
- Statistics:
 - Added: 0
 - Failed: 0
 - Updated:
 - Failed:
- Collection: demo:fedoraAppCollection
- Content Model: demo:fedoraAppCM

Below the statistics, the following information is displayed:

- * Institution: demo Batch Set: demo_fedoraAppDemoCollection_REMOTE
- Status: Validating Bag

At the bottom of the main content area, there are two buttons:

- [View Ingest Report](#)
- [View PID Report](#)

At the very bottom of the page, a box contains the following text:

v0.26; Start time: fri, nov 9, 2012, 9:50 pm
Stop time: fri, nov 9, 2012, 9:56 pm

The Ingest is ready to start

Fedora Batch Ingest - Mozilla Firefox

File Edit View History Bookmarks Tools Help

Fedora Batch Ingest Batch Ingest Results aquatic-animal-mammal-un...

localhost:7080/fedoraApp/viewRunningIngests.htm Google

Rocky Mountain Smal... Weebly - Create a fre... ADR Basic Document... Air Care Colorado :: V... The Ocean Explorer

Fedora Ingest Client 1.60

[Home](#) **Ingest Tasks (manual, _REMOTE and background _TASK)** 5 Second Refresh

BatchSetName	Institution	BatchSet	Status
demo_fedoraAppDemoCollection_REMOTE	demo	fedoraAppDemoCollection	Start batch ingest
codu_mixed_TASK	codu	mixed	Checking to see if there is anything to ingest

The ingest is complete,
objects are now in Fedora
and an email of results is sent



The screenshot shows a Mozilla Firefox browser window titled "Fedora Batch Ingest - Mozilla Firefox". The address bar displays "localhost:7080/fedoraApp/viewRunningIngests.htm". The browser's tab bar shows several tabs, including "Problem loading page", "Fedora Batch Ingest", "aquatic-animal-mammal-un...", and "Error 404 - Not Found". The main content area of the browser displays the "Fedora Ingest Client 1.60" interface. At the top of this interface is a navigation bar with a link to "Home" and a section titled "Ingest Tasks (manual, _REMOTE and background _TASK) 5 Second Refresh". Below this is a table with four columns: "BatchSetName", "Institution", "BatchSet", and "Status". The table contains one row of data where the "BatchSetName" is "demo_fedoraAppDemoCollection_REMOTE", the "Institution" is "demo", the "BatchSet" is "fedoraAppDemoCollection", and the "Status" is "Completed".

BatchSetName	Institution	BatchSet	Status
demo_fedoraAppDemoCollection_REMOTE	demo	fedoraAppDemoCollection	Completed

Sample Batch File (inside bagit.zip)

```
<?xml version="1.0" encoding="UTF-8"?>
<batch version="2">
  <ingestControl command="u" type="pco">

    <mets:mets xmlns:mets="http://www.loc.gov/METS/" OBJID="codu:63667" LABEL="Howe, Kay Digital Pioneers Oral History">

      <mets:fileSec>
        <mets:fileGrp ID="GR1" USE="master">
          <mets:file MIMETYPE="video/mp4" ID="Full_Howe.mp4">
            <mets:FLocat xmlns:xlink="http://www.w3.org/1999/xlink" LOCTYPE="URL" xlink:href="file:Full_Howe.mp4"></mets:FLocat>
          </mets:file>
        </mets:fileGrp>
      </mets:fileSec>

      <mets:structMap TYPE="logical">
        <mets:div TYPE="section" LABEL="Howe, Kay Digital Pioneers Oral History" ID="Full_Howe.id">
          <mets:fptr FILEID="Full_Howe.mp4 "></mets:fptr>
        </mets:div>
      </mets:structMap>

    </mets:mets>

  </batch>
```


Contents of script to post the bag to the ADR

- `#!/bin/sh`
- `java -jar postBatchFile.jar UserName UserPassword http://fedora.coalliance.org:9080/fedoraProxy/du/codu/digitalPioneers/ingest.it bagit.zip`
- `java -jar postBatchFile.jar UserName UserPassword
http://fedora.coalliance.org:9080/fedoraProxy/du/codu/digitalPioneers/ingest.it
bagit.zip`