Member Nodes - Task #3500

MNDeployment # 2564 (Deprecated): ORNL DAAC

Archiving ORNL DAAC pids

2013-01-18 18:11 - Robert Waltz

Status:	Closed		**Start date:**	2013-01-18
Priority:	Normal		**Due date:**	
Assignee:	Chris Jones		**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	Operational			
Story Points:				

Description

This script -should- might work for archiving pids. Run it as root

If it does not run correctly, then it would be easy to rewrite this up using d1-client-java

Dave, reassign once pids are added in to whom every you choose.

Add in each PID on a newline between EOF tags, run on cn-ucsb-1 or if another server, change script to point to appropriate certificate

#!/bin/bash

2012-07-13

Script to archive old PIDs from DAAC.

Need to run this as a user with access to the certificiate

curl -E "/etc/dataone/client/private/urn:node:CNUNM1.pem" -X PUT "https://cn.dataone.org/cn/v1/archive/somerecordpid"

```
HOST="cn-ucsb-1.dataone.org"
CERT="/etc/dataone/client/private/urn:node:CNUCSB1.pem"

PIDS=$(cat << EOF
PIDX
PIDY
PIDZ
EOF)

for PID in ${PIDS}; do
echo ${PID};
curl -E "${CERT}" -X PUT "https://${HOST}/cn/v1/archive/${PID}"
done
```

History

#1 - 2013-02-11 18:38 - Skye Roseboom

- Target version deleted (2013.2-Block.1.1)

#2 - 2013-02-11 18:41 - Skye Roseboom

- Target version deleted (2013.2-Block.1.1)

- Parent task set to #3568

#3 - 2013-04-19 22:40 - Skye Roseboom

- Target version set to Operational

#4 - 2013-04-19 22:40 - Skye Roseboom

- Project changed from Infrastructure to Member Nodes

- Category deleted (Support Operations)

- Parent task changed from #3568 to #2564

#5 - 2013-04-19 22:43 - Skye Roseboom

Production CN contains un-archived system metadata for record*.xml metadata records which do not exist on the ORNL DAAC.

https://cn-ucsb-1.dataone.org/cn/v1/query/solr/?rows=1000&fl=id&q=datasource:urn\:node\:ORNLDAAC%20AND%20formatType:METADATA%20AND%20id:record*

579 objects.

Ex:
https://cn-ucsb-1.dataone.org/cn/v1/meta/record847.xml
http://mercury-ops2.ornl.gov/ornldaac/mn/v1/meta/record847.xml

#6 - 2013-04-23 19:34 - Skye Roseboom

- Assignee changed from Dave Vieglais to Ranjeet Devarakonda

Spoke with Chris Jones and Ranjeet today about resolving these issues:

We are going to attempt a different solution than is presented above. This issue is similar to issue on ORNL Daac as well - with some content that is not resolvable any longer. So we are going to try the same solution:

USGS CSAS MN will update the system metadata to set the 'archive' flag and the 'dateSysMetadataModified'. This will trigger the CN to re-sync the system metadata with the archive information updated.   This will trigger the search index to remove these records from the search index.

#7 - 2013-06-04 14:45 - Skye Roseboom

Hi Ranjeet - If these documents have been deleted from ORNL DAAC, we will need a list of pids to 'archive' for ornl daac.

Can you confirm that all 579 objects identified by this query should be archived:

https://cn-ucsb-1.dataone.org/cn/v1/query/solr/?rows=1000&q=datasource:urn\:node\:ORNLDAAC%20AND%20formatType:METADATA%20AND%20id:record*

#8 - 2013-06-04 15:23 - Ranjeet Devarakonda

Hi Skye - Yes, all identifiers that start with "record*" should be set to archived. Your list looks correct.

#9 - 2013-06-05 15:51 - Skye Roseboom

- Assignee changed from Ranjeet Devarakonda to Chris Jones

Chris- Ranjeet has confirmed the list of pids to be archived. I believe we spoke about using your system metadata update script to archive these documents.

This solr query will return a list of pids that are from ORNL Daac and have a pid that begins with 'record'*.

https://cn-ucsb-1.dataone.org/cn/v1/query/solr/?rows=1000&q=datasource:urn\:node\:ORNLDAAC%20AND%20formatType:METADATA%20AND%20id:record*&fl=id&wt=csv

579 total docs

#10 - 2013-06-24 16:01 - Chris Jones

- Status changed from New to Closed

- translation missing: en.field_remaining_hours set to 0.0

I've changed the system metadata on the CNs for the record* pids listed above in the Solr query to 'archived', so they should no longer show up in the ONEMercury interface search. I'll confirm this with Skye.

#11 - 2013-07-17 22:45 - Chris Jones

- Status changed from Closed to In Progress

- Estimated time set to 0.00

Re-opening this issue since we have 3 more pids to archive that were removed from the MN, but still remain on the CN:

scimeta_415.xml, scimeta_650.xml, scimeta_704.xml

#12 - 2013-07-17 23:45 - Chris Jones

- Status changed from In Progress to Closed

I've archived the last three pids above.