Infrastructure - Bug #3547

object sysmeta available from CN but not resolveable

2013-02-02 22:12 - Dave Vieglais

Status:	Closed	**Start date:**	2013-02-03
Priority:	Normal	**Due date:**	2013-02-06
Assignee:	Robert Waltz	**% Done:**	100%
Category:	mn.Merritt	**Estimated time:**	0.00 hour
Target version:	2013.6-Block.1.3		
Milestone:	None	**Story Points:**	
Product Version:	*		

Description

PID: ark:/13030/m5513zdc/1/cadwsap-s3410003-006-vuln.csv

Sysmeta (from cn-unm-1):

<?xml version="1.0" encoding="UTF-8"?>

1
ark:/13030/m5513zdc/1/cadwsap-s3410003-006-vuln.csv
text/csv
2229
a91d6f7f4c14b789037a44bb166a9022c75defc0
trusted
ark:/13030/j2z317tt

public
read

false
2012-07-06T18:44:41.277+00:00
2012-08-12T03:22:02.219+00:00
urn:node:CDL
urn:node:CDL

urn:node:CDL
completed
2012-08-12T00:00:00.000+00:00

/d1:systemMetadata

History

#1 - 2013-02-02 22:14 - Dave Vieglais

- Description updated

#2 - 2013-02-03 00:54 - Dave Vieglais

Seems to be generalizable to any data object on Merritt.

e.g. :

- ark:/13030/m53x85jw/1/cadwsap-s5810001-005-vuln.csv
- ark:/13030/m53x85jw/1/cadwsap-s5810001-005-main.csv

or search for "water" in Mercury, then any of the Merritt results, try and download data.

#3 - 2013-02-03 04:41 - Chris Jones

A couple more data points:
This exception is coming from d1_cn_rest: org.dataone.cn.rest.filter.v1.ResolveFilter, and

The object is available at CDL:
https://merritt.cdlib.org:8084/knb/d1/mn/v1/object/ark:/13030/m5513zdc/1/cadwsap-s3410003-006-vuln.csv

#4 - 2013-02-03 13:17 - Dave Vieglais

The base URL advertised by the merritt node is different to what is recorded by the CN:

Merritt: https://d1mn.cdlib.org:8443/knb/d1/mn

CNs: https://merritt.cdlib.org:8084/knb/d1/mn

However, the URL reported by Merritt times out with no connection, the URL used by the CNs works.

So... perhaps Merritt needs to fix their advertised BaseURL?

#5 - 2013-02-03 13:34 - Dave Vieglais

- Assignee changed from Ben Leinfelder to Robert Waltz

The LDAP entry for CDL on cn-unm-1 does not have any services listed. Perhaps related to the mismatch between reported baseURL and the one that actually works?

#6 - 2013-02-04 17:27 - Ben Leinfelder

I've previously talked with CDL about this and we decided that it really doesn't matter how their Metacat is configured in terms of what the CN "knows" about the baseURL. The CN should only be using what it has in the CN's nodeList and should not be consulting the MN. The MN can call CN.updateNodeCapabilities if it would like to change its baseURL.

So in this case, I think it's all on the CN resolve side of things.

#7 - 2013-02-05 01:09 - Robert Waltz

- Status changed from New to In Progress

- Target version set to 2013.6-Block.1.3

- Start date set to 2013-02-03

- Due date set to 2013-02-06

urn:node:CDL does not list any services available on the CN. So, from the CN perspective it has no MNRead API exposed and therefore no resolvable objects.

Not certain where the error occurred, but most likely in however Merrit was originally setup.

But, we shouldn't be harvesting any records from the CDL if there is no read service, so... put some check in sychronization to make certain that an approved node is at least Tier 1 before synchronizing any content. That way, we don't get into this situation again. We would rather not see any objects being synchronized and discover that the node does not have a MNRead defined in the first place.

To fix the current problem, I'll write a script that will update the node entry for CDL (considering their base url entry is currently incorrect on their node

definition).

#8 - 2013-02-05 19:28 - Robert Waltz

Updated cn-orc-1.dataone.com's LDAP with services. UCSB's ldap is not updating. Probably needs to restart slapd to pick up changes.

We will restart slapd after indexing has completed for UCSB and ORC.

UNM should get the changes during the upgrade.

#9 - 2013-02-08 02:00 - Robert Waltz

- Status changed from In Progress to Closed