

Infrastructure - Feature #4039

Update ResolveSolrField to pull metacat properties for running as MN

2013-10-04 15:28 - Skye Roseboom

Status:	Closed	Start date:	2013-10-04
Priority:	Normal	Due date:	2014-04-12
Assignee:	Jing Tao	% Done:	100%
Category:	d1_indexer	Estimated time:	0.00 hour
Target version:	2014.14-Block.2.3	Story Points:	
Milestone:	None		
Product Version:	*		

Description

D1 index processing has been integrated into metacat MN stack as as search provider.

This class assumes CN operations:

- using /etc/dataone/node.properties to configure the hostname part of the resolve URL
- using 'cn' in the resolve url as part of the rest api (hostname/cn/v1/resolve/pid)

Possible to refactor the class to optionally look for hostname and cn/mn from metacat configuration if dataone configuration file is not available. This would allow the ResolveSolrField to operate in both a CN and MN metacat environment.

Matt mentioned metacat.properties as a source for this information:
{metacat web app context}/WEB-INF/metacat.properties
it contains properties:
server.name
dataone.nodeType
May also need base url info to append to host name before the /mn/v1/resolve/ path part?

In the search index - the 'dataUrl' and 'fileID' fields are filled by this class.

History

#1 - 2013-10-04 15:29 - Skye Roseboom

- translation missing: en.field_remaining_hours set to 0.0
- Due date set to 2013-10-26
- Tracker changed from Task to Feature

#2 - 2013-10-04 17:58 - Skye Roseboom

- Description updated

#3 - 2013-10-04 19:23 - Skye Roseboom

- Description updated

#4 - 2013-10-18 23:22 - Skye Roseboom

- Assignee changed from Skye Roseboom to Ben Leinfelder

Ben, can you document the desired content for these fields ('dataUrl' and 'fileID')? Sounded like maybe they should just stay empty? If there are more properties to be pulled from metacat.properties than listed above - could you document the name of each property? Thanks!

#5 - 2013-10-18 23:58 - Ben Leinfelder

If not blank, then I think the safest thing is to use this url for both fields:
://[[:port]]/d1/v1/object/{pid}

relevant metacat.properties:

server.name=localhost
server.httpPort=80
server.httpSSLPort=443
application.context=metacat
dataone.nodeType=mn

Unfortunately, selecting the protocol right now in Metacat is somewhat convoluted - if server.httpPort is set to 443, then https is used. I'd much rather there were a server.protocol property. Some day.

#6 - 2014-01-06 18:17 - Chris Jones

- Target version changed from 2013.42-Block.5.4 to 2014.2-Block.1.1
- Due date changed from 2013-10-26 to 2014-01-18

#7 - 2014-03-14 17:24 - Chris Jones

- Due date changed from 2014-01-18 to 2014-03-29
- Target version changed from 2014.2-Block.1.1 to 2014.12-Block.2.2
- Assignee changed from Ben Leinfelder to Jing Tao

#8 - 2014-03-14 17:26 - Skye Roseboom

Could be implemented as a separate 'metacat' specific implementation class for the resolve solr field. Does not necessarily need to live in the d1 code base. Runtime class is specified in the index related spring application context configuration - where this class could be specified over the dataone version of resolve.

#9 - 2014-03-31 17:57 - Jing Tao

- Due date changed from 2014-03-29 to 2014-04-12
- Target version changed from 2014.12-Block.2.2 to 2014.14-Block.2.3

#10 - 2014-04-11 21:43 - Jing Tao

I talked with Skye. He said the reason he wrote the bug is he saw the field in the Metacat MN solr index is not correct (pointing to cn). The only field which use the ResolveSolrField class in metacat-solr index is eml.fileID. The field is not used by any clients. And if the metacat mn deployment is changed (e. g., different host name or different context name), the eml.fileID, which is a direct url, should be reindexed. So i am thinking just remove the eml.fileID from the application-context-eml-base.xml.

#11 - 2014-04-15 16:38 - Jing Tao

- Status changed from New to Closed

Removed the eml.fileID from the application-context-eml-base.xml in metacat-index project.

#12 - 2014-04-19 00:39 - Jing Tao

- Status changed from Closed to In Progress

We only removed the fileID. but the dataUrl is defined in the application-context-systemmeta100.xml in d1_cn_index_processor.

#13 - 2014-04-28 23:57 - Jing Tao

- Status changed from In Progress to Closed

I copied application-context-systemmeta100.xml from the d1_cn_index_processor to the metacat-index and also removed the bean for the dataUrl.