

Member Nodes - Task #3774

MNDeployment # 3118 (Operational): Dryad Member Node

return NotFound exceptions for /meta and /object calls using "near miss" identifiers

2013-05-22 18:01 - Rob Nahf

<b>Status:</b>	New	<b>Start date:</b>	2013-05-22
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Ryan Scherle	<b>% Done:</b>	0%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	Operational		
<b>Story Points:</b>			
<b>Description</b>			
minor variations of known identifiers are returning content instead of NotFound exception.			
There are 2 or 3 parts to it:			
1. multiple consecutive slashes in the identifier are treated as one. So http://dx.doi.org/foo and http://dx.doi.org/foo will return the same thing.			
In some cases, the identifier uses http://dx.doi.org, in other cases <a href="http://dx.doi.org">http://dx.doi.org</a> . ( <a href="http://datadryad.org/mn/v1/object/http:%2Fdx.doi.org%2F10.5061%2Fdryad.8665">http://datadryad.org/mn/v1/object/http:%2Fdx.doi.org%2F10.5061%2Fdryad.8665</a> )			
2. the resource-map-indicator-suffix "?format=d1rem" will return a resource map, but so will "?format=d1remarkable" or "?format=d1remfooooooooooooo". The /object call returns a resource map with the "official" resource map identifier, but the /meta call uses the identifier from the api call as the identifier.			
3. altering the "ver={date}" suffix to the resource map indicator generates a unique variant of the resource map, incorporating that identifier into the RDF.			
So for example, the identifiers:			
<a href="http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;ver=2011-08-02T16:00:05.530-0400foobar">http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;ver=2011-08-02T16:00:05.530-0400foobar</a>			
<a href="http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;ver=2011-08-02T16:00:05.530-0411">http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;ver=2011-08-02T16:00:05.530-0411</a>			
<a href="http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;ver=2011-08-02T16:00:05.530-0412">http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;ver=2011-08-02T16:00:05.530-0412</a>			
<a href="http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;beer=2011-08-02T16:00:05.530-0400">http://dx.doi.org/F10.5061/dryad.12?format=d1rem&amp;beer=2011-08-02T16:00:05.530-0400</a>			
all return different resource maps (with different checksums and potentially sizes) that aggregate and document the same metadata and data objects related to the DOI. The differences reside in the URIs of the resource map and aggregation elements.			
The main point is that it's also important for the rest endpoints to return NotFound when the identifier requested doesn't exactly equal a known identifier.			
I'm not sure how that affects how you provision resource maps dynamically, but I assume that you keep track of resource map identifiers that are put into listObjects? maybe a check against that list before evaluating any regex's would do the trick.			
<b>Related issues:</b>			
Related to Member Nodes - Task #3854: Troubleshoot Dryad science metadata cre...		Closed	2013-06-28

History

#1 - 2013-05-24 18:48 - Matthew Jones

- Target version changed from Deploy by RSV to Deploy by end of Y4Q4

#2 - 2013-06-19 20:53 - Rob Nahf

- Description updated

#3 - 2013-08-01 19:50 - Bruce Wilson

- Target version changed from Deploy by end of Y4Q4 to Deploy by end of Y5Q2

**#4 - 2013-09-24 21:01 - Ryan Scherle**

This problem stems from the way Dryad generates records on-the-fly for DataONE. We have plans to revamp this system in phase 2 of the DataONE implementation. Dryad is tracking this issue at <https://trello.com/c/IYF9y6gk>

**#5 - 2014-02-03 16:03 - Laura Moyers**

- *Target version changed from Deploy by end of Y5Q2 to Deploy by end of Y5Q3*

**#6 - 2014-03-14 18:22 - Laura Moyers**

- *Target version changed from Deploy by end of Y5Q3 to Operational*