

Log Reporting - Bug #8657

Log aggregation service not recording logs for the MNs - urn:node:TDAR and urn:node:OTS_NDC

2018-07-16 21:36 - Rushiraj Nenuji

Status:	New	Start date:	2018-07-16
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:	d1_log_aggregation	Estimated time:	0.00 hour
Target version:			
Story Points:			

Description

It like the Dataone Logagg service is not recording logs for The Digital Archaeological Record MN (urn:node:TDAR). I was able to find only 31 events for all the datasets for this MN. (that too from year 2016).

LogSolr Query:

[https://cn.dataone.org/cn/v2/query/logsolr/?q=+nodeId:%22urn%3Anode%3ATDAR%22+event:read+formatType:\(METADATA%20OR%20DATA\)+inPartialRobotList:false+isRepeatVisit:false&facet=true&facet.limit=-1&facet.mincount=1&facet.field=formatType&facet.range=dateLogged&facet.range.start=2000-01-01T00:00:00Z&facet.range.end=2018-06-06T06:00:00.000Z&facet.range.gap=%2B1MONTH&wt=json&rows=0](https://cn.dataone.org/cn/v2/query/logsolr/?q=+nodeId:%22urn%3Anode%3ATDAR%22+event:read+formatType:(METADATA%20OR%20DATA)+inPartialRobotList:false+isRepeatVisit:false&facet=true&facet.limit=-1&facet.mincount=1&facet.field=formatType&facet.range=dateLogged&facet.range.start=2000-01-01T00:00:00Z&facet.range.end=2018-06-06T06:00:00.000Z&facet.range.gap=%2B1MONTH&wt=json&rows=0)

Looks like the same thing is going on for urn:node:OTS_NDC MN as well.

Jing Tao did a investigation on this and found a bug in the code:

```
[ERROR] 2018-07-15 14:19:56,338 [LogAggregationQuartzScheduler_Worker-14] (LogAggregationHarvestJob:execute:105) Job-urn:node:TDAR - died: org.dataone.service.types.v1.Log cannot be cast to org.dataone.service.types.v2.Log
java.lang.ClassCastException: org.dataone.service.types.v1.Log cannot be cast to org.dataone.service.types.v2.Log
    at org.dataone.client.v2.impl.MultipartMNode.getLogRecords(MultipartMNode.java:306)
    at org.dataone.cn.batch.logging.v2.NodeCommunication.harvest(NodeCommunication.java:71)
    at org.dataone.cn.batch.logging.tasks.LogHarvesterTask.harvest(LogHarvesterTask.java:284)
    at org.dataone.cn.batch.logging.jobs.LogAggregationHarvestJob.execute(LogAggregationHarvestJob.java:85)
    at org.quartz.core.JobRunShell.run(JobRunShell.java:213)
    at org.quartz.simpl.SimpleThreadPool$WorkerThread.run(SimpleThreadPool.java:557)
```

Jing, also happened to see the error caused by the MN itself.

```
[ERROR] 2018-07-16 16:34:26,443 [LogAggregationQuartzScheduler_Worker-6] (LogAggregationHarvestJob:execute:105) Job-urn:node:TDAR - died: org.dataone.service.exceptions.ServiceFailure: 504: Gateway Time-out: parser for deserializing HTML not written yet. Providing message body:
```

```
<html>
<head><title>504 Gateway Time-out</title></head>
<body bgcolor="white">
<center><h1>504 Gateway Time-out</h1></center>
<hr><center>nginx/1.10.3 (Ubuntu)</center>
</body>
</html>
```

```
java.util.concurrent.ExecutionException: org.dataone.service.exceptions.ServiceFailure: 504: Gateway Time-out: parser for deserializing HTML not written yet. Providing message body:
```

```
<html>
<head><title>504 Gateway Time-out</title></head>
<body bgcolor="white">
<center><h1>504 Gateway Time-out</h1></center>
<hr><center>nginx/1.10.3 (Ubuntu)</center>
</body>
```

</html>