

## D4Science Infrastructure - Support #10072

### Test the TwitterMonitor application in SoBigDataLab scope

Oct 26, 2017 02:00 PM - Salvatore Minutoli

<b>Status:</b>	Closed	<b>Start date:</b>	Oct 26, 2017
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Salvatore Minutoli	<b>% Done:</b>	100%
<b>Category:</b>	Application	<b>Estimated time:</b>	0.00 hour
<b>Infrastructure:</b>	Production		
<b>Description</b>			
<p>I'm trying to test the TwitterMonitor application in SoBigDataLab scope. I use the TwitterMonitor DataMiner algorithm at <a href="https://sobigdata.d4science.org/group/sobigdatalab/method-engine">https://sobigdata.d4science.org/group/sobigdatalab/method-engine</a>.</p> <p>To check the application I need to see the logs of the DataMiner algorithm. I checked the log files in <a href="http://dataminer3-p-d4s.d4science.org/gcube-logs/">http://dataminer3-p-d4s.d4science.org/gcube-logs/</a> following Andrea Dell'Amico suggestion. I read ghn.log, catalina.log and analysis.log and found no logs of my algorithm. In analysis.log and ghn.log I see that my algorithm has been called and that it terminates, probably due to some error. Could it be a problem of log configuration? Currently I use: private static Logger logger = AnalysisLogger.getLogger(); logger.debug(".....");</p> <p>The DataMiner algorithm must access a postgres database through the "TwitterMonitorDatabase" ServiceEndpoint. Is there some way to see the database contents to check if the records have been correctly inserted?</p>			

#### History

##### #1 - Oct 26, 2017 02:30 PM - Andrea Dell'Amico

Salvatore Minutoli wrote:

I'm trying to test the TwitterMonitor application in SoBigDataLab scope.  
I use the TwitterMonitor DataMiner algorithm at <https://sobigdata.d4science.org/group/sobigdatalab/method-engine>.

To check the application I need to see the logs of the DataMiner algorithm.  
I checked the log files in <http://dataminer3-p-d4s.d4science.org/gcube-logs/> following Andrea Dell'Amico suggestion. I read ghn.log, catalina.log and analysis.log and found no logs of my algorithm. In analysis.log and ghn.log I see that my algorithm has been called and that it terminates, probably due to some error.  
Could it be a problem of log configuration?  
Currently I use:  
private static Logger logger = AnalysisLogger.getLogger();  
logger.debug(".....");

The DataMiner algorithm must access a postgres database through the "TwitterMonitorDatabase" ServiceEndpoint. Is there some way to see the database contents to check if the records have been correctly inserted?

On the contrary to what I wrote by email, the dataminer servers assigned to the VRE have already access to the twittermon database:

```
2017-10-25 17:41:46.584 CEST [2305] twittermon@twittermon LOG: disconnection: session time: 0:00:00.574 user=twittermon database=twittermon host=dataminer3-p-d4s.d4science.org port=49443
2017-10-25 17:47:30.174 CEST [2375] [unknown]@[unknown] LOG: connection received: host=twittermon1.d4science.org port=46870
2017-10-25 17:47:30.176 CEST [2375] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 18:04:10.220 CEST [2375] twittermon@twittermon LOG: disconnection: session time: 0:16:40.047 user=twittermon database=twittermon host=twittermon1.d4science.org port=46870
2017-10-25 18:23:21.281 CEST [5275] [unknown]@[unknown] LOG: connection received: host=twittermon1.d4science.org port=47013
2017-10-25 18:23:21.283 CEST [5275] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 18:35:32.279 CEST [5426] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 18:35:33.184 CEST [5426] twittermon@twittermon LOG: disconnection: session time: 0:00:00.911 user=twittermon database=twittermon host=dataminer5-p-d4s.d4science.org port=38815
2017-10-25 18:37:30.212 CEST [5275] twittermon@twittermon LOG: disconnection: session time: 0:14:08.932 user=twittermon database=twittermon host=twittermon1.d4science.org port=47013
```

```

2017-10-25 18:53:41.936 CEST [5647] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 18:53:42.367 CEST [5647] twittermon@twittermon LOG: disconnection: session time: 0:00:00.434 user=twittermon database=twittermon host=dataminer3-p-d4s.d4science.org port=52635
2017-10-25 18:56:29.444 CEST [5685] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 18:56:30.206 CEST [5685] twittermon@twittermon LOG: disconnection: session time: 0:00:00.765 user=twittermon database=twittermon host=dataminer5-p-d4s.d4science.org port=39740
2017-10-25 19:05:42.240 CEST [5802] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 19:05:42.915 CEST [5802] twittermon@twittermon LOG: disconnection: session time: 0:00:00.679 user=twittermon database=twittermon host=dataminer5-p-d4s.d4science.org port=40159
2017-10-25 19:09:55.494 CEST [5853] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-25 19:09:56.168 CEST [5853] twittermon@twittermon LOG: disconnection: session time: 0:00:00.677 user=twittermon database=twittermon host=dataminer5-p-d4s.d4science.org port=40353
2017-10-26 01:24:21.289 CEST [10505] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon SSL enabled (protocol=TLSv1.2, cipher=ECDHE-ECDHSA-AES256-GCM-SHA384, compression=off)
2017-10-26 01:24:21.655 CEST [10505] twittermon@twittermon LOG: disconnection: session time: 0:00:00.373 user=twittermon database=twittermon host=localhost port=51744
2017-10-26 11:16:53.791 CEST [20516] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-26 11:16:54.365 CEST [20516] twittermon@twittermon LOG: disconnection: session time: 0:00:00.578 user=twittermon database=twittermon host=dataminer3-p-d4s.d4science.org port=42098
2017-10-26 11:24:35.125 CEST [20617] [unknown]@[unknown] LOG: connection received: host=twittermon1.d4science.org port=49882
2017-10-26 11:24:35.127 CEST [20617] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-26 11:26:44.003 CEST [20650] [unknown]@[unknown] LOG: connection received: host=twittermon1.d4science.org port=49911
2017-10-26 11:26:44.005 CEST [20650] twittermon@twittermon LOG: connection authorized: user=twittermon database=twittermon
2017-10-26 11:26:44.819 CEST [20617] twittermon@twittermon LOG: disconnection: session time: 0:02:09.696 user=twittermon database=twittermon host=twittermon1.d4science.org port=49882
2017-10-26 11:38:21.684 CEST [20650] twittermon@twittermon LOG: disconnection: session time: 0:11:37.683 user=twittermon database=twittermon host=twittermon1.d4science.org port=49911

```

The DB is almost empty, there are three tables containing one line each. The only crawlers table entry was inserted two days ago: 2017-10-24 15:12:48.418475

## #2 - Oct 26, 2017 04:19 PM - Salvatore Minutoli

Thanks Andrea for the information about the DB contents.

In previous tests no new record was added to the DB due to incorrect user parameters.

Using correct parameters a new record has been correctly added and starting a SmartExecutor plugin by means of a client, the processing was started. Now the problem is that the SmartExecutor.stop function gives an Exception and seems to not stop the plugin.

I already asked @roberto.cirillo@isti.cnr.it to stop and restart the system to avoid filling up all the disk space.

With incorrect user parameters the algorithm should return an error to the user: this does not happen. I would need to see the logs of the DataMiner algorithm to understand why.

## #3 - Oct 27, 2017 02:58 PM - Roberto Cirillo

- Status changed from New to Feedback

- Assignee changed from Roberto Cirillo to Salvatore Minutoli

## #5 - Oct 27, 2017 05:22 PM - Salvatore Minutoli

- Assignee changed from Salvatore Minutoli to Roberto Cirillo

The problem I reported about the logs is that in some log files (ghn.log, catalina.log and analysis.log) I can see that the Infrastructure actually calls my algorithm.

However I don't see the logs written by my algorithm (the class name is TwMonScheduler).

From an example taken in the online documentation I'm using:

```

import org.gcube.contentmanagement.lexicalmatcher.utils.AnalysisLogger;
private static Logger logger = AnalysisLogger.getLogger();
logger.debug("TwMonScheduler:init");

```

and it worked when I made the previous tests in RPrototypingLab (I was able to see my algorithm logs in analysis.log)

My IDE reports the AnalysisLogger class as deprecated, so my doubt is if I should switch to the same logging system I use in SmartExecutor plugins, i.e.

```
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;
private static Logger logger = LoggerFactory.getLogger(TwMonScheduler.class);
logger.debug(" ..... "),
```

**#6 - Oct 27, 2017 06:17 PM - Roberto Cirillo**

- Status changed from Feedback to In Progress

OK @salvatore.minutoli@iit.cnr.it thanks for your feedback. You are using a dedicated logger named "AnalysisLogger". I can try to change the default logback configuration in order to enable this kind of logger.

**#7 - Oct 27, 2017 06:23 PM - Roberto Cirillo**

- Status changed from In Progress to Feedback

- Assignee changed from Roberto Cirillo to Salvatore Minutoli

- % Done changed from 0 to 100

I've changed the logback configuration. Please could you re-check?

**#8 - Oct 27, 2017 06:31 PM - Roberto Cirillo**

I've added the following appender in the logback configuration on the dataminer\_bigdata cluster:

```
<logger name="AnalysisLogger" level="DEBUG">
  <appender-ref ref="ANALYSIS" />
</logger>
```

This change has been done by ansible. If it works well, I'm going to commit this change to the ansible playbook.

**#9 - Oct 30, 2017 02:16 PM - Salvatore Minutoli**

It seems my logs are not written. I started a computation and received a link from dataminer3-p-4ds.d4science.org.

I accessed <http://dataminer3-p-4ds.d4science.org/gcube-logs/> but the only file changed is catalina.out that contains a line written from my plugin, but only because I forgot a System.out.println in my code. I don't see any other log from my file and I cannot even find any log from the Infrastructure calling my DataMiner algorithm (TwMonScheduler).

Can you please tell me the name of the file that should contain the logs?

**#10 - Oct 31, 2017 09:47 AM - Roberto Cirillo**

The logger you are using is declared in the logback configuration and you should see your logs in the analysis.log file.

Yesterday, we have upgraded the bigdata cluster, so maybe something is changed, I hope. Please could you recheck if you see your logs now? Otherwise we should do further test in RprototypingLab environment.

**#11 - Oct 31, 2017 01:16 PM - Salvatore Minutoli**

Now I see the logs of my DataMiner Algorithm in analysis.log .

**#12 - Oct 31, 2017 01:59 PM - Salvatore Minutoli**

- Status changed from Feedback to Closed