D4Science Infrastructure - Incident #6575

Infra Gateway-1 node Garbage Collection (GC) process out of memory

Jan 24, 2017 06:50 PM - Massimiliano Assante

Status:	Closed		**Start date:**	Jan 24, 2017
Priority:	Urgent		**Due date:**	Jan 25, 2017
Assignee:	Massimiliano Assante		**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	UnSprintable			
Infrastructure:	Production			

Description

infra-gateway1 went out of memory today around 16 CET, the stack trace is in attachment. At a glance it is unclear what caused the problem, though many exception are reported from couch base.

Alessandro can you check?

History

#1 - Jan 24, 2017 07:20 PM - Massimiliano Assante

from Nagios:

```
Service: Memory status
Host: infra-gateway1.d4science.org
Address: infra-gateway1.d4science.org
State: WARNING

Date/Time: Tue Jan 24 15:28:59 CET 2017

Additional Info:

CHECK_MEMORY WARNING - 0G free
Notification Type: PROBLEM

Service: Haproxy backends
Host: infra-lb.d4science.org
Address: infra-lb.d4science.org
State: CRITICAL

Date/Time: Tue Jan 24 16:01:59 CET 2017

Additional Info:

Check haproxy CRITICAL - server: infra_backend:infra2 is DOWN 1/3 (check status: layer 7 (HTTP/SMTP) timeout):
```

#2 - Jan 25, 2017 10:45 AM - Massimiliano Assante

- Status changed from New to In Progress

#3 - Jan 25, 2017 10:48 AM - Alessandro Pieve

I checked the log file and I found about 20 accounting exceptions.
Considering that this "out of memory" issue do not occur on other nodes with thousands of exceptions, I am quite confident about the fact that the number of exceptions did not cause the issue.
In any case those errors were related to requests made by me to check the arrival of the data from the registry and the collector, while the cluster was also executing aggregation.
In the file I also noticed other errors and problems with disconnections yesterday, perhaps there have been other issues.
However, I can increase the request time out limit of the Couchbase library.

#4 - Jan 25, 2017 11:15 AM - Pasquale Pagano

- % Done changed from 0 to 20

this incident can be hardly fixed with the current information. It is was important to look at the logs but if you don't identify a trace to follow, please comment and close this issue.

#5 - Jan 25, 2017 11:31 AM - Massimiliano Assante

We have a leak, clearly and we must identify what happened yesterday otherwise the issue will happen again we cannot close the ticket.

#6 - Jan 25, 2017 11:47 AM - Pasquale Pagano

Yes, I agree but keeping the ticket open does not solve the issue. The ticket is assigned to Alessandro and I understood from his post that the issue is not related, or very hardly is related, to the accounting. So, what next? Who has to perform additional investigations?

#7 - Jan 25, 2017 11:53 AM - Massimiliano Assante

- Assignee changed from Alessandro Pieve to Massimiliano Assante

#8 - Jan 25, 2017 04:30 PM - Massimiliano Assante

- Status changed from In Progress to Closed

- % Done changed from 20 to 50

By looking at the logs I can only confirm that between 14.00 and 16.00 CET the only exceptions where due to the accounting, nothing else weird in the logs. Both cluster nodes were restarted last Saturday from me for fixing a problem with email notification, so the leak is indeed occurred between, Sunday and Tuesday PM.

Unfortunately the portal cluster nodes memory occupation was not monitored as they were not on Ganglia (for an error occurred) nor on Munin (they were never added). Tickets have been opened to fix the lacking of monitoring.

Nothing else can be done until the leak appears again.

Files

catalina.out.zip	861 KB	Jan 24, 2017	Massimiliano Assante