

D4Science Infrastructure - Task #12806

Support # 11605 (Closed): Set DataMiner queue length to 1

Task # 12614 (Closed): Need server with joblength 1 to use all 16 cores and a cue system per server.

Testing the new settings of the RAKIP infrastructure

Nov 02, 2018 04:25 PM - Lars Valentin

Status:	Closed	Start date:	Nov 02, 2018
Priority:	Urgent	Due date:	
Assignee:	Lars Valentin	% Done:	100%
Category:		Estimated time:	0.00 hour
Sprint:	Algorithms Integration		
Infrastructure:	Development, Production		
Description			
Hello Leonardo,			
I just published a workflow in KNIME/Dm15coresuse in https://aginfra.d4science.org/group/rakip_portal/data-miner , which is set up to use 16-1 cores in parallel to test the new setting. Could you or @andrea.dellamico@isti.cnr.it please check in htop if 15 cores are used while the workflow is running?			
Thanks a lot.			
Lars			

History

#1 - Nov 02, 2018 06:30 PM - Andrea Dell'Amico

- % Done changed from 0 to 90
- Assignee changed from Leonardo Candela to _InfraScience Systems Engineer
- Status changed from New to In Progress

Hi Lars,

we usually don't do that kind of support. We are in process of refining a set of dashboards to visualise our nodes metrics. The dashboards are very rough, but you can already obtain the data that you need.

Specifically for the dataminer cluster that you are going to use, the nodes are split between two data centers.

Five nodes you can find them here:

<https://grafana.d4science.org/d/ZFpM9qWmk/garr-ct1-multiple-hosts-server-metrics?orgId=1&var-node=dataminer-gw-4&var-node=dataminer-gw1&var-node=dataminer-gw2&var-node=dataminer-gw3&var-node=dataminer-gw5>

The other ten:

<https://grafana.d4science.org/d/-8vRr3Zmk/garr-pa1-multiple-hosts-server-metrics?orgId=1&var-node=dataminer-gw-1&var-node=dataminer-gw-10&var-node=dataminer-gw-2&var-node=dataminer-gw-3&var-node=dataminer-gw-4&var-node=dataminer-gw-5&var-node=dataminer-gw-6&var-node=dataminer-gw-7&var-node=dataminer-gw-8&var-node=dataminer-gw-9>

Keep in mind that's a work in progress, the dashboards are very heavy on the browser in their actual form and they have problems when more than 6 hosts are selected (you can always open more than a browser tab, with different hosts).

You should have access with the same credentials used to log into the portal, let us know if you get access.

#2 - Nov 02, 2018 07:22 PM - Lars Valentin

Hey Andrea,

thank you for your answer. I checked your links and I can see some nice graphs, but I am not sure if that really helps.

With the workflow I provided in the dataminer you could easily check if your changed settings work as expected, the possibility to use 16-1 cores at once.

I tried to access dataminer0-gw-proto which was used in one of the processes I run, but I could not find it.

I saw one graph in the linked servers which showed 3 peaks a 75%. That would fit to the old settings with 4 cores use max. According to that it might not have changed.

Best,
Lars

#3 - Nov 02, 2018 07:35 PM - Andrea Dell'Amico

Lars Valentin wrote:

Hey Andrea,

thank you for your answer. I checked your links and I can see some nice graphs, but I am not sure if that really helps.

With the workflow I provided in the dataminer you could easily check if your changed settings work as expected, the possibility to use 16-1 cores at once.

We did not change any settings on the nodes. We gave you access to a cluster that was already configured that way. The nodes were already there, but they were not meant to be directly accessible.

I tried to access dataminer0-gw-proto which was used in one of the processes I run, but I could not find it.

The dashboard for the generic worker prototypes nodes is <https://grafana.d4science.org/d/B63Xb3Gmk/dataminer-ct1-gw-proto?orgId=1>

I saw one graph in the linked servers which showed 3 peaks a 75%. That would fit to the old settings with 4 cores use max. According to that it might not have changed.

Even if the configuration was wrong (and it's not), there are no hard limits on the cpu usage. So, if there's only one job running on a node, it's free to use all the available CPUs.

A usage of 75% means that all the jobs running on the node are not able to use more CPU than that, not that some policy is limiting them.

#4 - Nov 07, 2018 08:43 AM - Lars Valentin

Hi Andrea,

I still don't see a climb in the Grafana CPU fields with the links provided above after executing my testworkflow.

Apart from that there seems to be a difference in the version of our plugin installed and the developing version I use locally, which is causing that only 10 cores are used in the test. We are working on a new release version and as soon as we have it published I will ask you to install it in your infrastructure and I test again.

Best,
Lars

#5 - Nov 07, 2018 10:10 AM - Andrea Dell'Amico

Lars Valentin wrote:

Hi Andrea,

I still don't see a climb in the Grafana CPU fields with the links provided above after executing my testworkflow.

Apart from that there seems to be a difference in the version of our plugin installed and the developing version I use locally, which is causing that only 10 cores are used in the test. We are working on a new release version and as soon as we have it published I will ask you to install it in your infrastructure and I test again.

Your plugin is automatically upgraded every night, as you asked.

#6 - Nov 07, 2018 11:34 AM - Lars Valentin

Thats great! We will then have it right after the new version is published from our side.

But still, I can't even monitor the use of the current 10 cores. Therefore I need support on how to see that in the tool provided. I suggest that we do that together in the bilateral webmeeting which will be set up by CNR in the near future to get support in different issues.

Best,
Lars

#7 - Nov 21, 2018 07:07 AM - Lars Valentin

- Priority changed from Normal to Urgent

Hey [@andrea.dellamico@isti.cnr.it](mailto:andrea.dellamico@isti.cnr.it) ,

the new FSK version was installed fine and is running.

Unfortunately, I still have no feedback at all by the links you provided me.

I need to proof that each of the 16 cores is used while my 'challenge workflow' is executed, like in htop (eg. core 1 87%; core 2 77%; core 3 98% ...)
Could you please check that for me if it works as expected?

Furthermore, I am wondering why non of the CPU graphs reacted in any way while I was running the workflow, using your links and setting last 15 minutes. I assume they do not work properly or target other servers then those we are using.

Thank you!
Lars

#8 - Nov 21, 2018 01:28 PM - Andrea Dell'Amico

Lars Valentin wrote:

Unfortunately, I still have no feedback at all by the links you provided me.

I need to proof that each of the 16 cores is used while my 'challenge workflow' is executed, like in htop (eg. core 1 87%; core 2 77%; core 3 98% ...)
Could you please check that for me if it works as expected?

Furthermore, I am wondering why non of the CPU graphs reacted in any way while I was running the workflow, using your links and setting last 15 minutes. I assume they do not work properly or target other servers then those we are using.

They work properly, and they are related to the servers you are using.

I cannot spend time staring at a bunch of htop sessions on request by every users at any time, if you can run your workflow at a certain time I'll have a session ready to wach the execution. Later today, or in the late morning tomorrow. Otherwise we must go to the next week.

#9 - Nov 23, 2018 07:15 AM - Lars Valentin

They work properly, and they are related to the servers you are using.

I cannot spend time staring at a bunch of htop sessions on request by every users at any time, if you can run your workflow at a certain time I'll have a session ready to wach the execution. Later today, or in the late morning tomorrow. Otherwise we must go to the next week.

Good morning Andrea,

I expected that you are able to run the workflow yourself at any time it suits you, in order to check in htop if really all cores can be used by the model/workflow in your infrastructure. But if you need my help for that, for sure, I am very willing to provide that.

Today, I would like to invest time to finally solve that issue. If you don't mind we can have a webmeeting were we share screens and you show me that.

In addition, I would like to learn how I would have extracted this information from the grafana tool.

Please let me know what time. I think Leonardo can setup a GOTO meeting for us.

Best,
Lars

#10 - Nov 23, 2018 11:48 AM - Andrea Dell'Amico

- Assignee changed from *_InfraScience Systems Engineer* to *Giancarlo Panichi*

I'm passing this activity to [@g.panichi@isti.cnr.it](mailto:g.panichi@isti.cnr.it), he knows dataminer much better than me.

#11 - Nov 23, 2018 02:18 PM - Giancarlo Panichi

- Assignee changed from *Giancarlo Panichi* to *Lars Valentin*

Hi [@lars.valentin@bfr.bund.de](mailto:lars.valentin@bfr.bund.de), please could you give me the Equivalent Get Request that is given by the DataMiner when you invoke the algorithm? Or alternatively, can you tell me the name of the algorithm and with which parameters it must be invoked? In this way I can take a test and see how your algorithm works.
Thanks

#13 - Nov 23, 2018 02:54 PM - Lars Valentin

- Assignee changed from *Lars Valentin* to *Giancarlo Panichi*

#14 - Nov 23, 2018 03:00 PM - Lars Valentin

Hi [@g.panichi@isti.cnr.it](mailto:g.panichi@isti.cnr.it)

looks like my previous message got lost while uploading.

Thank you for taking over! The workflow was already mentioned in the ticket opening. Here is the get request from SHOW:

<http://dataminer-genericworkers-proto.d4science.org/wps/WebProcessingService?request=Execute&service=WPS&Version=1.0.0&gcube-token=USERTOKEN&lang=en-US&Identifier=org.gcube.dataanalysis.wps.statisticalmanager.synchserver.mappedclasses.transducerers.DM15CORESUSE&DataInputs=dummy=dummy>;

Since many tickets are public I removed the user token.

As already explained above, all I need to know is how many cores are really used by the workflow on the server. I usually easily check that with HTOP.

HTH,

Lars

#15 - Nov 23, 2018 03:08 PM - Giancarlo Panichi

- Assignee changed from Giancarlo Panichi to Lars Valentin
- File Screenshot_2018-11-23_7.png added
- File Screenshot_2018-11-23_6.png added
- File Screenshot_2018-11-23_5.png added
- File Screenshot_2018-11-23_4.png added
- File Screenshot_2018-11-23_3.png added
- File Screenshot_2018-11-23_2.png added
- File Screenshot_2018-11-23_1.png added

Hi @lars.valentin@bfr.bund.de , I tried your algorithm Dm15coresuse with the default parameter, and is actually using all the cores. Furthermore, I have noticed that the process is stable on the amount of memory used. It seems to me that there is a problem in the termination phase of the algorithm, because although the server has finished, the algorithm does not end and remains there. I think you should check your algorithm, however, the algorithm ends correctly. See the screenshots.

#16 - Nov 23, 2018 03:18 PM - Giancarlo Panichi

- % Done changed from 90 to 100

Sorry @lars.valentin@bfr.bund.de , I had not noticed the amount of files produced. Well, then the execution of your algorithm is correct.

#17 - Nov 23, 2018 03:29 PM - Lars Valentin

- Status changed from In Progress to Closed

Hi @g.panichi@isti.cnr.it ,

thanks a lot for helping me. Looks like the multicore use does now behave as expected in your infrastructure. I will close the ticket.

Best,
Lars

#18 - Nov 23, 2018 04:03 PM - Leonardo Candela

- Status changed from Closed to In Progress

Lars Valentin wrote:

Hi @g.panichi@isti.cnr.it ,

thanks a lot for helping me. Looks like the multicore use does now behave as expected in your infrastructure. I will close the ticket.

Best,
Lars

Hi all, thanks for completing this investigation.

To avoid future misunderstandings or misleading messages, I kindly invite [@lars.valentin@bfr.bund.de](mailto:lars.valentin@bfr.bund.de) to clarify what he means with "Looks like the multicore use does now behave as expected in your infrastructure."

[@lars.valentin@bfr.bund.de](mailto:lars.valentin@bfr.bund.de) please help me in understanding what is your expectation and where the observed behaviour is mismatching it.

#19 - Nov 23, 2018 10:09 PM - Lars Valentin

- Assignee changed from Lars Valentin to Leonardo Candela

- Status changed from In Progress to Feedback

Hi Leonardo,

as a developer of services, I have to prove that the outcome of my work is behaving as expected. Therefore, it is my duty to test that for example a workflow which should be able to use all cores of a server in your infrastructure really does that. Since, I could not do that as easy in the VRE as in my own server, I did ask for help to check HTOP. Unfortunately, this seemed to be more difficult than I thought. But Giancarlo was so kind to finally check HTOP and I know now, that all present cores can be used in parallel via one single model. Meaning a graph like added above, showing the CPU use for each core of a single server. My work in this task is completed now and I can focus on the other tasks.

Regarding Grafana, this tool did not show me the information I needed. And after some hours wasting time with the tool and having the ticket open for 21 one days for a 5 minutes support task, you might understand that I needed a working solution.

HTH,
Lars

#20 - Nov 27, 2018 12:28 PM - Leonardo Candela

[@lars.valentin@bfr.bund.de](mailto:lars.valentin@bfr.bund.de) thanks for the explanation. What was puzzling me was the comment "multicore use does **now** behave as expected" ... this sound like CNR changed something to make the cluster working as expected ... this is not the case since the beginning of this thread.

BTW, I'm happy to close this ticket and to consider the activity complete.

#21 - Nov 27, 2018 12:28 PM - Leonardo Candela

- Status changed from Feedback to Closed

#22 - Nov 27, 2018 12:42 PM - Lars Valentin

- Status changed from Closed to In Progress

Leonardo Candela wrote:

[@lars.valentin@bfr.bund.de](mailto:lars.valentin@bfr.bund.de) thanks for the explanation. What was puzzling me was the comment "multicore use does **now** behave as expected" ... this sound like CNR changed something to make the cluster working as expected ... this is not the case since the beginning of this thread.

BTW, I'm happy to close this ticket and to consider the activity complete.

Hey Leonardo,

this is a big misunderstanding here. Since we have spent many months and many tickets to finally persuade you to allow us to use all cores of a

server (16) with a joblength = 1 you now say it was always like that.

This is totally confusing to me since it was always the stated fact by you and your colleagues that only 4 cores can be used by 1 job (joblength =4) and you finally were so kind a short time ago to change that only for us and explicit only for the RAKIP Dataminer. Therefore there is the huge gap between what was said by CNR all the time and what you say now. I would kindly ask you to explain that in detail to avoid further complaining.

Thank you

#23 - Nov 27, 2018 12:46 PM - Lars Valentin

I add examples of what CNR said in the past:

<https://support.d4science.org/issues/12614#note-5>

<https://support.d4science.org/issues/12614#note-8>

#24 - Nov 27, 2018 12:57 PM - Leonardo Candela

- Status changed from *In Progress* to *Feedback*

- *Infrastructure Production* added

Hey @lars.valentin@bfr.bund.de

yes there is a big misunderstanding here (and not only here I would say). The overall attitude should be reconsidered, we are not here to fight we are here to collaborate ... questioning is fine, going beyond that is not useful.

My sentence is "... this sound like CNR changed something to make the cluster working as expected ... this is not the case since **the beginning of this thread.**" ... **THIS THREAD** is the discussion we are having in this ticket ... I confirm that in the last 25 days nothing has been changed in RAKIP.

Are we on the same page?

#25 - Nov 27, 2018 12:58 PM - Leonardo Candela

- Assignee changed from *Leonardo Candela* to *Lars Valentin*

#26 - Nov 27, 2018 01:07 PM - Lars Valentin

- Assignee changed from *Lars Valentin* to *Leonardo Candela*

Why do you say I am fighting? I was kindly asking why we have this gap in what CNR said before and does say now.

As you know, this thread was opened because you suggested "open a new ticket declaring the testing / assessment plan (e.g. algorithms to be integrated and tested, load, concurrency, number of users served in parallel) and use the new ticket to report on the effectiveness of the proposed solution." here <https://support.d4science.org/issues/12614#note-8> .

All I wanted to do was following your suggestion and trying to test if the NEW setting is working as expected. And I can't say that CNR really supportive in that topic for a long time.

So please start to consider the ticket in the context it was and is, I am just following your advise.

#27 - Nov 27, 2018 01:12 PM - Panagiotis Zervas

Lars Valentin wrote:

Why do you say I am fighting? I was kindly asking why we have this gap in what CNR said before and does say now.

As you know, this thread was opened because you suggested "open a new ticket declaring the testing / assessment plan (e.g. algorithms to be integrated and tested, load, concurrency, number of users served in parallel) and use the new ticket to report on the effectiveness of the proposed solution." here <https://support.d4science.org/issues/12614#note-8> .

All I wanted to do was following your suggestion and trying to test if the NEW setting is working as expected. And I can't say that CNR was not really supportive in that topic for a long time.

So please start to consider the ticket in the context it was and is, I am just following your advise.

@leonardo.candela@isti.cnr.it and @lars.valentin@bfr.bund.de we have a meeting in 20 minutes can we discuss there the issue and resolve any misunderstanding? I don't think this is a fruitful discussion

#28 - Nov 27, 2018 02:38 PM - Leonardo Candela

- Assignee changed from Leonardo Candela to Lars Valentin

- Status changed from Feedback to Closed

No need to continue keeping this ticket open, there was a misunderstanding and that's all.

#29 - Nov 27, 2018 04:41 PM - Pasquale Pagano

I will add a bit of technical information since I believe it is not yet clear.

D4Science operated by CNR supports two configurations for the computing clusters assigned to the VREs. We shortly called: (joblength =4) and (joblength =1).

With the first (joblength =4), each server allocated to the cluster accepts up to 4 jobs at a time. This does mean that

-- each server may run 1, 2, 3, 4 jobs at a time. 4 jobs are the maximum number of jobs accepted by the server before queueing.

-- each job may use up to 15 cores. Clearly, if 4 jobs are running at the same time, they will compete to access the same set of cores and the OS will change contexts several times. This will not be efficient but still supported.

With the first (joblength =1), each server allocated to the cluster accepts up to 1 job at a time. Each job may use up to 15 cores without competing for the computing resources.

Both configurations have pros and cons and this is why we support both. It really depends on the use cases. The first case makes a better use of the computational resources, while with the second configuration we risk to unexploit computational resources that are kept unallocated.

#30 - Nov 29, 2018 07:39 AM - Lars Valentin

@pasquale.pagano@isti.cnr.it

Thanks a lot for the detailed explanation!

In the past I always understood that the `joblength=4` is somehow setup with the limitation that 1 job can use only 4 cores max. Now I know it just behaves like a normal system. I see both advantages and disadvantages and will take them into account for the further development plans.

Best,
Lars

Files

Screenshot_2018-11-23_1.png	304 KB	Nov 23, 2018	Giancarlo Panichi
Screenshot_2018-11-23_2.png	312 KB	Nov 23, 2018	Giancarlo Panichi
Screenshot_2018-11-23_3.png	319 KB	Nov 23, 2018	Giancarlo Panichi
Screenshot_2018-11-23_4.png	328 KB	Nov 23, 2018	Giancarlo Panichi
Screenshot_2018-11-23_5.png	327 KB	Nov 23, 2018	Giancarlo Panichi
Screenshot_2018-11-23_6.png	321 KB	Nov 23, 2018	Giancarlo Panichi
Screenshot_2018-11-23_7.png	315 KB	Nov 23, 2018	Giancarlo Panichi