

D4Science Infrastructure - Support #2134

Enhancing Statistical Manager performance

Feb 04, 2016 11:38 AM - Gianpaolo Coro

<b>Status:</b>	Closed	<b>Start date:</b>	Feb 04, 2016
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	_InfraScience Systems Engineer	<b>% Done:</b>	100%
<b>Category:</b>	High-Throughput-Computing	<b>Estimated time:</b>	0.00 hour
<b>Infrastructure:</b>	Development		
<b>Description</b>			
Several requests by BlueBRIDGE partners require running processes on the Statistical Manager (SM) machines. We are going to test the requirements of these algorithms, which are CPU and memory intensive and this activity requires enhancing the machines resources.			
For example, an IRD process usually requiring 15 minutes on a modern Desktop machine, is taking more than one hour on one SM machine. During the process, I see the following statistics using the "top" command:			
<pre>22066 gcube      20    0  208m 120m 1340 R   100   1.5  24:15.13 vpa-2box.out 490 syslog      20    0  243m 4336  928 S    0    0.1  15:45.74 rsyslogd ...</pre>			
(Note: I'm running the process directly, without using the service at this stage)			
Since the processes stress one CPU at time and require large memory, I wonder if better processors could be assigned to the dev statistical manager machine for this assessment phase. Also, 16GB of RAM memory are necessary for this phase.			
<b>Related issues:</b>			
Related to D4Science Infrastructure - Task #3070: Provide a way of starting a...		Closed	Mar 30, 2016

History

#2 - Feb 09, 2016 06:20 PM - Tommaso Piccioli

- Assignee changed from Tommaso Piccioli to \_InfraScience Systems Engineer

#3 - Mar 09, 2016 08:50 PM - Pasquale Pagano

If this request has been analyzed, please answer to the request and schedule a task for moving the VM to better hw.

#4 - Mar 09, 2016 09:10 PM - Gianpaolo Coro

I have not received feedback about this yet.

#5 - Mar 23, 2016 05:57 PM - Gianpaolo Coro

Please, could you update me about this ticket? Is there any possibility to have more powerful machines for Dataminer and StatMan? At least for Dataminer these are required.

We are currently in the paradoxal scenario in which memory demanding computations should be better run on personal computers (as some user are already doing)!

#6 - Mar 23, 2016 06:14 PM - Andrea Dell'Amico

Waiting for Tom for the operations.

But we really need some data, because your top output is completely useless.

- How many CPUs?
- How do you elaborated that 16 GB for each host are needed, and what amout of that memory should be assigned as Java heap?

dataminer2-d-d4s and dataminer2-p-d4s are already running on the fastest hardware we have. If you don't see any significant difference from the jobs run on that machines and the other, there's nothing we can do other that increment the assigned RAM (but we aren't using all the memory already available, as ganglia shows:

[http://monitoring.research-infrastructures.eu/ganglia/?r=hour&cs=&ce=&m=mem\\_report&s=by+name&c=D4science+Dataminer+production+cluster&h=&host\\_regex=&max\\_graphs=0&tab=m&vn=&sh=1&z=small&hc=4](http://monitoring.research-infrastructures.eu/ganglia/?r=hour&cs=&ce=&m=mem_report&s=by+name&c=D4science+Dataminer+production+cluster&h=&host_regex=&max_graphs=0&tab=m&vn=&sh=1&z=small&hc=4)

#7 - Mar 23, 2016 11:09 PM - Gianpaolo Coro

16 CPUs or cores should be sufficient if these were averagely free.

16 GB of RAM comes from the fact that one of the algorithms we are going to integrate (Ecopath with Ecosym - Global Ocean model) requires at least 10GB of RAM.

But we are still very far from this requirement: we are currently doing experiments with regional geospatial datasets, requiring to keep 300G real numbers in memory. Although this is managed by my laptop (8GB RAM), Dataminer (4.8GB RAM on the machine, I don't know how much set for the service) goes obviously out-of-memory.

**#8 - Mar 24, 2016 12:22 PM - Andrea Dell'Amico**

Gianpaolo Coro wrote:

16 CPUs or cores should be sufficient if these were averagely free.

$(16 * (2 \text{ dataminer nodes})) + (16 * (4 \text{ statistical nodes})) = 96 \text{ CPUs}$ . It's 2 1/2 out of our most powerful machines and we have three of them, running a total of more than 100 VMs.

16 GB of RAM comes from the fact that one of the algorithms we are going to integrate (Ecopath with Ecosym - Global Ocean model) requires at least 10GB of RAM.

But we are still very far from this requirement: we are currently doing experiments with regional geospatial datasets, requiring to keep 300G real numbers in memory. Although this is managed by my laptop (8GB RAM), Dataminer (4.8GB RAM on the machine, I don't know how much set for the service) goes obviously out-of-memory.

We could rise the requested resources if we had a more dynamic environment, but after the experience with the generic worker nodes I don't trust any request as not permanent anymore. We can work in two directions:

- calculate what amount of memory is really necessary to run the biggest jobs (300GB is out of discussion) to resize the two dataminer nodes (the production dataminer nodes run with 6GB of ram, 5GB assigned to the JDK heap)
- Exploit the EGI resources to run bigger dataminer tasks on demand. We have potentially a lot of usable resources out there and I made a very big effort so that we are able to use them

**#9 - Mar 24, 2016 12:30 PM - Gianpaolo Coro**

Let's start rising RAM to 16 GB

**#10 - Mar 24, 2016 12:38 PM - Andrea Dell'Amico**

On what? production dataminer nodes? devel dataminer nodes? and what about the JDK heap? Do the jobs run inside or outside the JDK heap space?

**#11 - Mar 24, 2016 12:42 PM - Gianpaolo Coro**

On the dataminer nodes (both dev and prod, see my comment above). Most of the jobs run in the JDK heap thus this should potentially access all the available memory.

**#12 - Mar 24, 2016 12:49 PM - Andrea Dell'Amico**

Gianpaolo Coro wrote:

On the dataminer nodes (both dev and prod, see my comment above). Most of the jobs run in the JDK heap thus this should potentially access all the available memory.

16GB on all the four nodes is way too much, really. We can go up to 8GB on the dev nodes and 16 on the production ones.

**#13 - Mar 24, 2016 01:00 PM - Gianpaolo Coro**

So we are stuck. Our platforms has been involved in projects that require embedding software produced by other people, who did not designed them to be parallelised.

One of the most important processes we have to run requires at least 10GB of RAM and may also run concurrently. For other processes, like CMSY, I have not managed to convince other scientists to use our e-Infrastructure because their laptops have more RAM (and sometimes can use more CPU cores).

Thus, what I'm asking for is **mandatory** for our projects scenarios, it is not a "kind request".

If we cannot satisfy these requirements, then the European projects we are involved in should be officially alerted...and we should move to other kinds of activities.

**#14 - Mar 24, 2016 01:13 PM - Andrea Dell'Amico**

Gianpaolo Coro wrote:

So we are stuck. Our platforms has been involved in projects that require embedding software produced by other people, who did not designed them to be parallelised.

One of the most important processes we have to run requires at least 10GB of RAM and may also run concurrently. For other processes, like CMSY, I have not managed to convince other scientists to use our e-Infrastructure because their laptops have more RAM (and sometimes can use more CPU cores).

Thus, what I'm asking for is **mandatory** for our projects scenarios, it is not a "kind request".

If we cannot satisfy these requirements, then the European projects we are involved in should be officially alerted...and we should move to other kinds of activities.

I don't know why discussing with you always needs to be so exhausting.

Do you see that 10 is very different from 16? If so, we maybe reach an agreement on 12GB on all the dataminer nodes?

And why the dev nodes matter in this context?

And, last: why the use of VMs on the EGI environment is not an option?

**#15 - Mar 24, 2016 01:18 PM - Andrea Dell'Amico**

I lied. The last question is: why you expose the mandatory requirement as argument only after days of discussion?

And I'm not entering in the "my laptop has more ram so its faster" argument because it's a so big nonsense that it's unbelievable that an adult person can raise it.

**#16 - Mar 24, 2016 01:38 PM - Gianpaolo Coro**

You're right, this is not the place to talk about the mandates of the projects we are involved in, you should have already known them.

The real problem is that I'm reporting a requirement coming after discussions with other participants, several meetings and tests, and they are being treated like my personal requirements, just because I cannot report in one ticket all the explanations this information comes from. That's crazy!

If an important requirement cannot be satisfied we should discuss it in a separate meeting, giving the requirement its dignity and importance.

After many tickets and email, the EGI environment has been demonstrated not to be working for our purposes yet. Further, it is not a prototyping environment like the one we need at first glance. Former infra users, in fact, are now developers too and they need to test their processes on dev machines.

In summary, either we clarify this situation (possibly with a meeting) or I'm going to close this ticket and report the status to the BlueBRIDGE project management, to make them understand our contingency.

**#17 - Mar 24, 2016 01:51 PM - Andrea Dell'Amico**

Gianpaolo Coro wrote:

You're right, this is not the place to talk about the mandates of the projects we are involved in, you should have already known them.

How?

The real problem is that I'm reporting a requirement coming after discussions with other participants, several meetings and tests, and they are being treated like my personal requirements, just because I cannot report in one ticket all the explanations this information comes from. That's crazy!

No. I'm just trying to get a list of valid requirements. You started asking for 16GB of RAM on the statistical manager nodes showing two lines from top were a process was using 200MB RAM.

If an important requirement cannot be satisfied we should discuss it in a separate meeting, giving the requirement its dignity and importance.

I'm trying to satisfy it, and I made more than one proposal.

Further, it is not a prototyping environment like the one we need at first glance. Former infra users, in fact, are now developers too and they need to test their processes on dev machines.

In a place where resources are not unlimited (and still greatly wasted, see the ganglia reports about the memory usage), the dev environment cannot be sized as the production one. It has to work obviously, so: if 10GB is the least amount of memory that can make the dev environment work, is 10GB of ram for the dev dataminer and 16 for the production ones a viable solution?

**#18 - Mar 24, 2016 03:41 PM - Pasquale Pagano**

I believe that the static allocation of resources to services is something is going to constraint us too much. Hopefully in a not so far future we could manage our site with more elasticity and reduce the wasted resources in favor of the services that need them for the time they need them. Saying this, I believe that we can shutdown in production

- either 10 WNs i
- or one instance of the SM if we have not enough memory and allocate this memory to the production DataMiner.

Alternatively we should look at the development infrastructure where for sure there are VMs that are not exploited.

**#19 - Mar 25, 2016 04:41 PM - Gianpaolo Coro**

Today we have had a short meeting and it seems that a self service tool can be built to allow people turn machines on and off.

This would allow having some sort of dynamic resources allocation, where machines would be activated only when really needed, for example to manage a class or a large computation.

As a first step, we could shut 50 workers (out of 60) down and also dataminer2-d. Possibly, after talking with @roberto.cirillo@isti.cnr.it , we could shut two Statistical Managers down and activate them when needed.

This requires synchronizing ourselves, especially during the release phases. Could we talk about this in a proper meeting?

**#20 - Mar 30, 2016 11:03 AM - Roberto Cirillo**

It's ok for me

**#21 - Mar 30, 2016 02:50 PM - Andrea Dell'Amico**

- *Related to Task #3070: Provide a way of starting and stopping a list of VMs by some users added*

**#22 - Mar 30, 2016 02:50 PM - Andrea Dell'Amico**

- *Status changed from New to Closed*

- *% Done changed from 0 to 100*

An agreement was reached, see <https://support.d4science.org/issues/3070>