

OpenNebula - Bug #3390

mm_sched stops scheduling new vms

12/01/2014 10:31 AM - Jan Horacek

Status: Closed	Start date: 12/01/2014
Priority: Normal	Due date:
Assignee:	% Done: 0%
Category: Scheduler	Estimated time: 0.00 hour
Target version: Release 5.0	Pull request:
Resolution: fixed	
Affected Versions: OpenNebula 4.10, OpenNebula 4.8	

Description

in our production environment i recurently met problem with mm_sched, which is running, but stops scheduling new vms.

symptom is, that vms stuck in PENDING state.

sched.log does not contain any new records after last sucessfull schedule.

no hung transactions in the mysql db, no running processes.

after 'service opennebula restart', scheduler is restarted and vms are scheduled.

firs hit this bug in 4.8 and now in 4.10.1 it is still happening.

Related issues:

Related to Bug # 4284: Scheduler stops working during night	Closed	01/11/2016
---	---------------	-------------------

Associated revisions

Revision c6076a17 - 03/24/2016 01:07 AM - Ruben S. Montero

bug #3390: Reimplements the xmlrpc client using the complex one. Adds TIMEOUT to client requests. This can be configured in sched.conf/oned.conf.

History

#1 - 12/09/2014 03:09 PM - Ruben S. Montero

- Target version changed from Release 4.10.1 to Release 4.10.2

#2 - 12/09/2014 03:12 PM - Ruben S. Montero

Are there any other symptoms for the mm_sched process? Do the size of the mm_sched grows during the execution? I've just moved to the next update version.

#3 - 12/10/2014 09:53 AM - Ruben S. Montero

- Category set to Scheduler

#4 - 12/10/2014 09:53 AM - Ruben S. Montero

- Assignee set to Carlos Martín

#5 - 12/10/2014 10:03 AM - Jan Horacek

i found one anomaly in the templates, it has defined network AND bridge too. that is a mistake, but i hope this should not be the reason.

...

```
NIC=[IP="...",BRIDGE="vbr0",NETWORK_ID="4"]
```

...

#6 - 12/12/2014 05:56 PM - Carlos Martín

Hi Jan,

Next time it happens, can you please check if you have the periodic vm pool info calls in oned.conf?

```
Fri Dec 12 17:38:12 2014 [Z0][ReM][D]: Req:8256 UID:0 VirtualMachinePoolInfo invoked , -2, -1, -1, -1
```

Also, please save the list of pending VMs when the scheduler stops working (onevm list -x). It might be that a VM is causing a parsing error of some kind.

#7 - 01/10/2015 09:44 PM - Ruben S. Montero

- Target version deleted (Release 4.10.2)

moving this out of 4.10.2, till we have the needed input to debug this

#8 - 03/04/2015 12:09 AM - Ruben S. Montero

- Status changed from Pending to Closed

- Resolution set to worksforme

#9 - 10/20/2015 01:05 PM - Matej Kovac

Hello, I believe I'm hit by this bug too on fresh install of OpenNebula version 4.12.3.

I've noticed that VM's stay PENDING when I defined IP="..." in template. I was stuck for hours when I restarted scheduler looked around for a known bug.

Scheduler and oned messages appear as follows:

```
Oct 20 10:52:02 one oned[559]: [Z0][ReM][D]: Req:4832 UID:0 VirtualMachinePoolInfo invoked , -2, -1, -1, -1
Oct 20 10:52:02 one oned[559]: [Z0][ReM][D]: Req:4832 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL></VM_POOL>"
Oct 20 10:52:02 one mm_sched[578]: [Z0][SCHED][D]: Getting VM and Host information. Total time: 0.00s
Oct 20 10:56:31 one oned[559]: [Z0][ReM][D]: Req:9264 UID:3 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
Oct 20 10:56:31 one oned[559]: [Z0][ReM][D]: Req:9264 UID:3 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL></VM_POOL>"
Oct 20 10:56:38 one oned[559]: [Z0][ReM][D]: Req:6976 UID:3 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
Oct 20 10:56:38 one oned[559]: [Z0][ReM][D]: Req:6976 UID:3 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL></VM_POOL>"
Oct 20 11:16:13 one oned[559]: [Z0][ReM][D]: Req:7248 UID:0 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
Oct 20 11:16:13 one oned[559]: [Z0][ReM][D]: Req:7248 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL></VM_POOL>"
Oct 20 11:23:03 one oned[559]: [Z0][ReM][D]: Req:5472 UID:0 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
Oct 20 11:23:03 one oned[559]: [Z0][ReM][D]: Req:5472 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL><VM><ID>28<..."
Oct 20 11:23:28 one oned[559]: [Z0][ReM][D]: Req:7632 UID:0 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
```

```
Oct 20 11:23:28 one oned[559]: [Z0][ReM][D]: Req:7632 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL><VM><ID>28<...>"
Oct 20 11:23:28 one oned[559]: [Z0][ReM][D]: Req:4016 UID:0 VirtualMachineInfo result SUCCESS, "<VM><ID>28</ID><UID>..."
Oct 20 11:23:52 one oned[559]: [Z0][ReM][D]: Req:7728 UID:0 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
```

The last message from mm_sched comes on 10:52:02. Then after restart it looks like:

```
Oct 20 14:25:00 one oned[559]: [Z0][ReM][D]: Req:7744 UID:0 VirtualMachinePoolInfo invoked , -2, 0, -2000, -1
Oct 20 14:25:00 one oned[559]: [Z0][ReM][D]: Req:7744 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL><VM><ID>30<...>"
Oct 20 14:26:50 one mm_sched[29807]: [Z0][SCHED][I]: Init Scheduler Log system
Oct 20 14:26:50 one mm_sched[29807]: [Z0][SCHED][I]: Starting Scheduler Daemon
```

Should it happen again, I post onevm list -x. IP= argument in template works as expected and VMs are scheduled.

#10 - 10/23/2015 07:53 AM - Carlos Martín

- Status changed from Closed to Pending

Moving it to pending so we don't forget to evaluate Matej's comment

#11 - 01/19/2016 02:25 PM - Laurent Grawet

Hi,

We are having the same problem with OpenNebula 4.14.2. We have to restart the scheduler twice a week!

I can see the periodic vm pool info calls in oned.log.

Thanks for helping,

Laurent

#12 - 01/19/2016 04:33 PM - Laurent Grawet

Here is a strace of mm_sched process and childs :

```
# strace -fp 21353
Process 21353 attached with 2 threads
[pid 21356] pselect6(5, [4], [], [], {1, 446753853}, {NULL, 8} <unfinished ...>
[pid 21353] futex(0x7fcc456c39d0, FUTEX_WAIT, 21356, NULL <unfinished ...>
[pid 21356] <... pselect6 resumed> ) = 0 (Timeout)
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328977, 836510062}) = 0
[pid 21356] rt_sigaction(SIGPIPE, NULL, {SIG_DFL, [], SA_RESTORER, 0x7fcc4b6238d0}, 8) = 0
[pid 21356] rt_sigaction(SIGPIPE, {SIG_IGN, [], SA_RESTORER, 0x7fcc4b6238d0}, NULL, 8) = 0
[pid 21356] poll([{fd=4, events=POLLIN|POLLPR|POLLRDNORM|POLLRDBAND}], 1, 0) = 0 (Timeout)
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328977, 836906002}) = 0
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328977, 836977072}) = 0
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328977, 837081010}) = 0
[pid 21356] rt_sigaction(SIGPIPE, {SIG_DFL, [], SA_RESTORER, 0x7fcc4b6238d0}, NULL, 8) = 0
[pid 21356] gettimeofday({1453220993, 211927}, NULL) = 0
```

```
[pid 21356] pselect6(5, [4], [], [], {3, 0}, {NULL, 8}) = 0 (Timeout)
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328980, 840576822}) = 0
[pid 21356] rt_sigaction(SIGPIPE, NULL, {SIG_DFL, [], SA_RESTORER, 0x7fcc4b6238d0}, 8) = 0
[pid 21356] rt_sigaction(SIGPIPE, {SIG_IGN, [], SA_RESTORER, 0x7fcc4b6238d0}, NULL, 8) = 0
[pid 21356] poll([{fd=4, events=POLLIN|POLLPRIO|POLLRDNORM|POLLRDBAND}], 1, 0) = 0 (Timeout)
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328980, 840930402}) = 0
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328980, 841019069}) = 0
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328980, 841099969}) = 0
[pid 21356] rt_sigaction(SIGPIPE, {SIG_DFL, [], SA_RESTORER, 0x7fcc4b6238d0}, NULL, 8) = 0
[pid 21356] gettimeofday({1453220996, 215968}, NULL) = 0
[pid 21356] pselect6(5, [4], [], [], {3, 0}, {NULL, 8}) = 0 (Timeout)
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328983, 844564028}) = 0
[pid 21356] rt_sigaction(SIGPIPE, NULL, {SIG_DFL, [], SA_RESTORER, 0x7fcc4b6238d0}, 8) = 0
[pid 21356] rt_sigaction(SIGPIPE, {SIG_IGN, [], SA_RESTORER, 0x7fcc4b6238d0}, NULL, 8) = 0
[pid 21356] poll([{fd=4, events=POLLIN|POLLPRIO|POLLRDNORM|POLLRDBAND}], 1, 0) = 0 (Timeout)
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328983, 844911695}) = 0
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328983, 844980758}) = 0
[pid 21356] clock_gettime(CLOCK_MONOTONIC, {4328983, 845078520}) = 0
[pid 21356] rt_sigaction(SIGPIPE, {SIG_DFL, [], SA_RESTORER, 0x7fcc4b6238d0}, NULL, 8) = 0
[pid 21356] gettimeofday({1453220999, 219959}, NULL) = 0
[pid 21356] pselect6(5, [4], [], [], {3, 0}, {NULL, 8})
```

#13 - 01/20/2016 10:07 PM - Ruben S. Montero

So it seems that the mm_sched is getting a SIGPIPE, probably related to a socket operation in the libxmlrpc-c library. Although the default is to terminate the process, all the signals are blocked from the main thread:

<https://github.com/OpenNebula/one/blob/master/src/scheduler/src/sched/Scheduler.cc#L368>

Is the mm_sched process of that trace actually terminating ?

#14 - 01/22/2016 08:56 AM - Laurent Grawet

SIGTERM (one stop-sched) doesn't terminate the process, I have to send SIGKILL

#15 - 02/16/2016 09:25 AM - Michal Leinweber

Same problem with 4.14.2 on Debian 8.

#16 - 02/16/2016 02:22 PM - Ruben S. Montero

- Target version set to Release 5.0

- Resolution deleted (worksforme)

Ok thanks, reopening the issue. We'll try to reproduce it in a long-running sched process.

#17 - 02/17/2016 02:52 PM - Laurent Grawet

Hi,

The scheduler is not responding again on Debian 8.2. How can we provide more useful information ?

```
# strace -fp 19045
Process 19045 attached with 2 threads
[pid 19048] pselect6(5, [4], [], [], {1, 241349476}, {NULL, 8} <unfinished ...>
[pid 19045] rt_sigtimedwait([INT TERM], NULL, NULL, 8 <unfinished ...>
[pid 19048] <... pselect6 resumed> ) = 0 (Timeout)
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737222, 32835984}) = 0
[pid 19048] rt_sigaction(SIGPIPE, NULL, {SIG_DFL, [], SA_RESTORER, 0x7f50149d48d0}, 8) = 0
[pid 19048] rt_sigaction(SIGPIPE, {SIG_IGN, [], SA_RESTORER, 0x7f50149d48d0}, NULL, 8) = 0
[pid 19048] poll([{fd=4, events=POLLIN|POLLPRI|POLLRDNORM|POLLRDBAND}], 1, 0) = 0 (Timeout)
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737222, 33455909}) = 0
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737222, 33621189}) = 0
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737222, 33757970}) = 0
[pid 19048] rt_sigaction(SIGPIPE, {SIG_DFL, [], SA_RESTORER, 0x7f50149d48d0}, NULL, 8) = 0
[pid 19048] gettimeofday({1455629237, 408716}, NULL) = 0
[pid 19048] pselect6(5, [4], [], [], {3, 0}, {NULL, 8}) = 0 (Timeout)
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737225, 37418780}) = 0
[pid 19048] rt_sigaction(SIGPIPE, NULL, {SIG_DFL, [], SA_RESTORER, 0x7f50149d48d0}, 8) = 0
[pid 19048] rt_sigaction(SIGPIPE, {SIG_IGN, [], SA_RESTORER, 0x7f50149d48d0}, NULL, 8) = 0
[pid 19048] poll([{fd=4, events=POLLIN|POLLPRI|POLLRDNORM|POLLRDBAND}], 1, 0) = 0 (Timeout)
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737225, 37932210}) = 0
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737225, 38041341}) = 0
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737225, 38147489}) = 0
[pid 19048] rt_sigaction(SIGPIPE, {SIG_DFL, [], SA_RESTORER, 0x7f50149d48d0}, NULL, 8) = 0
[pid 19048] gettimeofday({1455629240, 413040}, NULL) = 0
[pid 19048] pselect6(5, [4], [], [], {3, 0}, {NULL, 8}) = 0 (Timeout)
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737228, 41704583}) = 0
[pid 19048] rt_sigaction(SIGPIPE, NULL, {SIG_DFL, [], SA_RESTORER, 0x7f50149d48d0}, 8) = 0
[pid 19048] rt_sigaction(SIGPIPE, {SIG_IGN, [], SA_RESTORER, 0x7f50149d48d0}, NULL, 8) = 0
[pid 19048] poll([{fd=4, events=POLLIN|POLLPRI|POLLRDNORM|POLLRDBAND}], 1, 0) = 0 (Timeout)
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737228, 42125591}) = 0
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737228, 42227923}) = 0
[pid 19048] clock_gettime(CLOCK_MONOTONIC, {6737228, 42331549}) = 0
[pid 19048] rt_sigaction(SIGPIPE, {SIG_DFL, [], SA_RESTORER, 0x7f50149d48d0}, NULL, 8) = 0
[pid 19048] gettimeofday({1455629243, 417220}, NULL) = 0
[pid 19048] pselect6(5, [4], [], [], {3, 0}, {NULL, 8})^CProcess 19045 detached
Process 19048 detached
<detached ...>
```

#18 - 02/18/2016 01:08 AM - Ruben S. Montero

Laurent Grawet wrote:

Hi,

The scheduler is not responding again on Debian 8.2. How can we provide more useful information ?

[...]

09/21/2020

5/9

I think this is a bug in xml-rpc library. This is happening in Debian 8 systems (at least), so have you installed OpenNebula from source, or are you using the precompiled packages? What version of the xml-rpc library is installed in your system? Just to estimate the length of the messages exchange, can you send the output of "onevm list -x | wc -c" and "onehost list -x | wc -c"

Cheers

#19 - 02/18/2016 09:00 AM - Laurent Grawet

We use the precompiled packages.

libxmlrpc-c++8 - 1.33.14-0.2

onevm list -x | wc -c ----> 2080144

onehost list -x | wc -c ----> 1290998

Thanks for your support

#20 - 02/18/2016 05:47 PM - Michal Leinweber

OpenNebule from precompiled packages

Library is the same: libxmlrpc-c++8 1.33.14-0.2

My lists are much smaller:

onevm list -x | wc -c

267490

onehost list -x | wc -c

96498

#21 - 02/26/2016 08:42 AM - Michal Leinweber

Is there any progress? Can I help with other information?

#22 - 03/01/2016 08:08 AM - Michal Leinweber

sched.log ends with:

Fri Feb 26 16:08:09 2016 [Z0][SCHED][D]: Getting scheduled actions information. Total time: 0.11s

Fri Feb 26 16:08:09 2016 [Z0][SCHED][D]: Getting VM and Host information. Total time: 0.10s

and in the same time in oned.log:

Fri Feb 26 16:08:09 2016 [Z0][ReM][D]: Req:3472 UID:0 VirtualMachinePoolInfo invoked , -2, -1, -1, -1

Fri Feb 26 16:08:09 2016 [Z0][ReM][D]: Req:3472 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL><VM><ID>171..."

Fri Feb 26 16:08:09 2016 [Z0][ReM][D]: Req:2416 UID:0 VirtualMachinePoolInfo invoked , -2, -1, -1, -1

Fri Feb 26 16:08:09 2016 [Z0][ReM][D]: Req:2416 UID:0 VirtualMachinePoolInfo result SUCCESS, "<VM_POOL><VM><ID>171..."

Fri Feb 26 16:08:10 2016 [Z0][InM][E]: Information driver crashed, recovering...

Where to find more information for the last error row in oned.log?

#23 - 03/04/2016 09:26 AM - Ruben S. Montero

I've running the scheduler for 3 days now with 1000VMs PENDING VMs and 50 hosts for 3 days now, till now it is not hanging... The env uses the same libxmlrpc libraries. I'll keep updated with any progress.

#24 - 03/04/2016 09:28 AM - Ruben S. Montero

Michal Leinweber wrote:

sched.log ends with:

[...]

and in the same time in oned.log:

[...]

Where to find more information for the last error row in oned.log?

This is the information manager driver process dying, it should not have nothing to do with the scheduler

#25 - 03/05/2016 03:09 PM - Michal Leinweber

Ruben S. Montero wrote:

This is the information manager driver process dying, it should not have nothing to do with the scheduler

Are you sure it is not related? I have it there every time scheduler stops working.

#26 - 03/08/2016 08:13 AM - Michal Leinweber

Where to find more debug messages for this error:

`[Z0][InM][E]: Information driver crashed, recovering...`

I have debug level 3 and no details information for this error in oned.log.

#27 - 03/15/2016 09:05 AM - Ruben S. Montero

Still not able to reproduce this :(, shceduler running for ~1000 VMs more than 10 days...

The message is because of the information manager driver died, this is a ruby script so there may be a problem with the ruby interpreter, related

libraries or the driver itself. Maybe the condition that makes the IM driver fail is the same that makes the scheduler stuck, but they are separated process one in ruby the other in C++.

#28 - 03/15/2016 09:07 AM - Ruben S. Montero

- *Related to Bug #4284: Scheduler stops working during night added*

#29 - 03/15/2016 09:09 AM - Ruben S. Montero

I'll investigate the suggestion of Tobias, to look for any systemd interaction... Note also that the scheduler does not keep any file open, so logrotates should not be a problem...

#30 - 03/23/2016 05:07 PM - Ruben S. Montero

Good/Bad news, we were able to finally reproduce it...

1. Scheduler periodically calls oned to get info about VMs, hosts etc..
2. Calls use the SimpleClient included in xmlrpc-c library
3. This call is synchronous and timeout cannot be set
4. We see process hang in a pselect loop, somehow the server throw the connection and the client side logic does not consider the transaction finished.
- 5.- As a result the scheduler blocks in this loop

I think the cleanest way to fix this is to get rid of the SimpleClient class and use the Asynchronous client that would let us cope with these situations.

The bad news is this is not a couple-of-lines fix, we are addressing this for the next release...the current workaround is to restart the scheduler :(

#31 - 03/24/2016 01:17 AM - Ruben S. Montero

- *Assignee deleted (Carlos Martín)*

I've just uploaded a new client that timeouts requests and should fix this bug. Closing this one... I'll check if a patch can be created for the 4.14 scheduler.

#32 - 03/24/2016 01:20 AM - Ruben S. Montero

- *Status changed from Pending to Closed*

- *Resolution set to fixed*

#33 - 05/02/2016 02:58 PM - Laurent Grawet

Hi,

Is there any chance to obtain a patch against 4.14.2 ?

Thanks,

Laurent

#34 - 05/05/2016 08:22 AM - Ruben S. Montero

Laurent Grawet wrote:

Hi,

Is there any chance to obtain a patch against 4.14.2 ?

Thanks,

Laurent

Hi Laurent,

This is not an easy one, we can consider this after 5.0 release. Please ping here :)

Cheers

#35 - 07/15/2016 03:35 PM - Vladimir Babichev

Hi,

5.0 has been released a month ago, so Ping :)

PS: I have the same problem with 4.14.2

Cheers

#36 - 07/18/2016 08:09 AM - Ruben S. Montero

No plans for the backport in the short-term, sorry

#37 - 09/19/2016 07:44 AM - Michael Ivanov

this bug happens to me on CentOS 7 and OpenNebula 4.14.2 few times a week. Please backport to 4.14.2 CentOS repo, I can make a donation if needed :)