

mgr - Bug #43447

mgr/diskprediction: diskprediction module fails to initialize with newer SciPy versions

01/02/2020 05:09 PM - Volker Theile

Status:	New	% Done:	0%
Priority:	Normal		
Assignee:	Sage Weil		
Category:	diskprediction_cloud		
Target version:	v15.0.0		
Source:	Development	Reviewed:	
Tags:		Affected Versions:	
Backport:		ceph-qa-suite:	
Regression:	Yes	Pull request ID:	
Severity:	2 - major	Crash signature:	
Description			
In latest master it seems			
<pre>ceph -h</pre>			
does not list the manager module commands. Because of that it is not possible to start a vstart environment which infinitely waits for the dashboard module to be started.			
vstart is using			
<pre>ceph -h</pre>			
to check if a module is loaded, see https://github.com/ceph/ceph/blob/master/src/vstart.sh#L972 .			
<pre># start-ceph.sh === mon.c === Stopping Ceph mon.c on ceph-master-docker...done === mon.b === Stopping Ceph mon.b on ceph-master-docker...done === mon.a === Stopping Ceph mon.a on ceph-master-docker...done === mgr.x === Stopping Ceph mgr.x on ceph-master-docker...done ** going verbose ** rm -f core* hostname ceph-master-docker ip 192.168.178.24 port 40763 /ceph/build/bin/ceph-authtool --create-keyring --gen-key --name=mon. /ceph/build/keyring --cap mon 'allow *' creating /ceph/build/keyring /ceph/build/bin/ceph-authtool --gen-key --name=client.admin --cap mon 'allow *' --cap osd 'allow *' ' --cap mds 'allow *' --cap mgr 'allow *' /ceph/build/keyring /ceph/build/bin/monmaptool --create --clobber --addv a [v2:192.168.178.24:40763,v1:192.168.178.24: 40764] --addv b [v2:192.168.178.24:40765,v1:192.168.178.24:40766] --addv c [v2:192.168.178.24:4076 7,v1:192.168.178.24:40768] --print /tmp/ceph_monmap.30775 /ceph/build/bin/monmaptool: monmap file /tmp/ceph_monmap.30775 /ceph/build/bin/monmaptool: generated fsid 78ca53ba-40d0-4716-9d53-f24f254e2d0d epoch 0 fsid 78ca53ba-40d0-4716-9d53-f24f254e2d0d last_changed 2020-01-02T16:52:44.629708+0000 created 2020-01-02T16:52:44.629708+0000 min_mon_release 0 (unknown) 0: [v2:192.168.178.24:40763/0,v1:192.168.178.24:40764/0] mon.a 1: [v2:192.168.178.24:40765/0,v1:192.168.178.24:40766/0] mon.b</pre>			

The configuration:

```
# cat /ceph/build/ceph.conf
; generated by vstart.sh on Thu 02 Jan 2020 04:52:44 PM UTC
[client.vstart.sh]
    num mon = 3
    num osd = 3
    num mds = 3
    num mgr = 1
    num rgw = 1
    num ganesha = 0

[global]
    fsid = 55d5f021-9cff-41bb-861e-17dcb5df051e
    osd failsafe full ratio = .99
    mon osd full ratio = .99
    mon osd nearfull ratio = .99
    mon osd backfillfull ratio = .99
    erasure code dir = /ceph/build/lib
    plugin dir = /ceph/build/lib
    filestore fd cache size = 32
    run dir = /ceph/build/out
    crash dir = /ceph/build/out
    enable experimental unrecoverable data corrupting features = *
    osd_crush_chooseleaf_type = 0
    debug asok assert abort = true

    ms bind msgr2 = true
    ms bind msgr1 = true

    lockdep = true
    auth cluster required = cephx
    auth service required = cephx
    auth client required = cephx

[client]
    keyring = /ceph/build/keyring
    log file = /ceph/build/out/$name.$pid.log
    admin socket = /tmp/ceph-asok.014tIJ/$name.$pid.asok

    ; needed for s3tests
    rgw crypt s3 kms backend = testing
    rgw crypt s3 kms encryption keys = testkey-1=YmluCmJvb3N0CmJvb3N0LWJ1aWxkCmNlcGguY29uZgo=
testkey-2=aWIKTWFrZWZpbGUKbWFuCm91dApzcmMKVGVzdGluZwo=
    rgw crypt require ssl = false
    ; uncomment the following to set LC days as the value in seconds;
    ; needed for passing lc time based s3-tests (can be verbose)
    ; rgw lc debug interval = 10
    ; The following settings are for SSE-KMS with Vault
    ;rgw crypt s3 kms backend = vault
    ;rgw crypt vault auth = token
    ;rgw crypt vault token file = /ceph/build/vault.token
    ;rgw crypt vault addr = http://127.0.0.1:8200
    ;rgw crypt vault secret engine = kv
    ;rgw crypt vault prefix = /v1/kv/data
    ;rgw crypt vault secret engine = transit
    ;rgw crypt vault prefix = /v1/transit/export/encryption-key/

[cephfs-shell]
    debug shell = true

[client.rgw.8000]
    rgw frontends = beast port=8000
```

```

admin socket = /ceph/build/out/radosgw.8000.asok
[mds]

log file = /ceph/build/out/$name.log
admin socket = /tmp/ceph-asok.014tIJ/$name.asok
chdir = ""
pid file = /ceph/build/out/$name.pid
heartbeat file = /ceph/build/out/$name.heartbeat

mds data = /ceph/build/dev/mds.$id
mds root ino uid = 0
mds root ino gid = 0

[mgr]

mgr data = /ceph/build/dev/mgr.$id
mgr module path = /ceph/src/pybind/mgr
cephadm path = /ceph/src/cephadm/cephadm

log file = /ceph/build/out/$name.log
admin socket = /tmp/ceph-asok.014tIJ/$name.asok
chdir = ""
pid file = /ceph/build/out/$name.pid
heartbeat file = /ceph/build/out/$name.heartbeat

[osd]

log file = /ceph/build/out/$name.log
admin socket = /tmp/ceph-asok.014tIJ/$name.asok
chdir = ""
pid file = /ceph/build/out/$name.pid
heartbeat file = /ceph/build/out/$name.heartbeat

osd_check_max_object_name_len_on_startup = false
osd data = /ceph/build/dev/osd$id
osd journal = /ceph/build/dev/osd$id/journal
osd journal size = 100
osd class tmp = out
osd class dir = /ceph/build/lib
osd class load list = *
osd class default list = *
osd fast shutdown = false

filestore wbthrottle xfs ios start flusher = 10
filestore wbthrottle xfs ios hard limit = 20
filestore wbthrottle xfs inodes hard limit = 30
filestore wbthrottle btrfs ios start flusher = 10
filestore wbthrottle btrfs ios hard limit = 20
filestore wbthrottle btrfs inodes hard limit = 30
bluestore fsck on mount = true
bluestore block create = true
bluestore block db path = /ceph/build/dev/osd$id/block.db.file
bluestore block db size = 1073741824
bluestore block db create = true
bluestore block wal path = /ceph/build/dev/osd$id/block.wal.file
bluestore block wal size = 1048576000
bluestore block wal create = true

; kstore
kstore fsck on mount = true
osd objectstore = bluestore

[mon]

mgr initial modules = dashboard restful iostat

log file = /ceph/build/out/$name.log
admin socket = /tmp/ceph-asok.014tIJ/$name.asok
chdir = ""

```

```

pid file = /ceph/build/out/$name.pid
heartbeat file = /ceph/build/out/$name.heartbeat

debug mon = 20
debug paxos = 20
debug auth = 20
debug mgrc = 20
debug ms = 1

mon cluster log file = /ceph/build/out/cluster.mon.$id.log
osd pool default erasure code profile = plugin=jerasure technique=reed_sol_van k=2 m=1 cru
sh-failure-domain=osd
[mon.a]
    host = ceph-master-docker
    mon data = /ceph/build/dev/mon.a
[mon.b]
    host = ceph-master-docker
    mon data = /ceph/build/dev/mon.b
[mon.c]
    host = ceph-master-docker
    mon data = /ceph/build/dev/mon.c
[global]
    mon host = [v2:192.168.178.24:40763,v1:192.168.178.24:40764] [v2:192.168.178.24:40765,v1:
192.168.178.24:40766] [v2:192.168.178.24:40767,v1:192.168.178.24:40768]
[mgr.x]
    host = ceph-master-docker

```

The module is configured:

```

# /ceph/build/bin/ceph -c /ceph/build/ceph.conf -k /ceph/build/keyring mgr module ls
*** DEVELOPER MODE: setting PATH, PYTHONPATH and LD_LIBRARY_PATH ***
2020-01-02T16:55:29.438+0000 7f6dd3fa1700 -1 WARNING: all dangerous and experimental features are
enabled.
2020-01-02T16:55:29.462+0000 7f6dd3fa1700 -1 WARNING: all dangerous and experimental features are
enabled.
{
  "enabled_modules": [
    "dashboard",
    "iostat",
    "restful"
  ],
  "disabled_modules": []
}

```

Related issues:

Related to mgr - Bug #45147: Module 'diskprediction_local' takes forever to load

New

History

#1 - 01/02/2020 05:10 PM - Volker Theile

- File mgr.x.log added

#2 - 01/02/2020 05:17 PM - Volker Theile

- Assignee set to Sage Weil

@Sage I assigned the issue to you because you already worked on that issue some time ago:

<https://github.com/ceph/ceph/pull/30217/commits/81280455112b1c8f4495ef63b9a5e236e5c65465>

#3 - 01/03/2020 06:48 AM - Kiefer Chang

The mgr hangs when loading `diskprediction_local` module.

Some experiments:

- Move diskprediction_local module folder out of source tree to avoid loading it --> mgr starts.
- The new-created container has scipy 1.4.1. I temporarily replace it with scipy 1.3.2 --> mgr starts.

May be related to [#42764](#).

#4 - 01/03/2020 08:18 AM - Volker Theile

Kiefer, you made my day. I can confirm that removing the src/pybind/mgr/diskprediction_local/ directory fixes the issue. I assume that downgrading scipy will also fix it, too.

#5 - 01/15/2020 12:59 PM - Patrick Seidensal

I ran into this today and removing `diskprediction_local` from `src/pybind/mgr` resolved the issue for me.

#6 - 01/15/2020 01:01 PM - Patrick Seidensal

@Kiefer Chang

That explains why a new build of my container seems to have broken vstart! Thanks for the investigation.

#7 - 01/29/2020 01:12 PM - Lenz Grimmer

FWIW, I ran into this issue myself a few days ago and was not able to determine the root cause. Downgrading scipy to version 1.3.2 fixed the issue for me.

We should investigate why newer versions of that module cause the diskprediction module to get stuck in initialization.

#8 - 01/29/2020 01:17 PM - Lenz Grimmer

- Project changed from Ceph to mgr

- Subject changed from vstart: ceph -h does not show manager module commands to mgr/diskprediction: diskprediction module fails to initialize with newer SciPy versions

- Category changed from ceph cli to diskprediction_cloud

- Regression changed from No to Yes

- Severity changed from 3 - minor to 2 - major

Updated component and severity according to the latest findings.

#9 - 03/12/2020 08:10 AM - Kiefer Chang

gdb call trace of ceph-mgr when using with scipy 1.4.1.

```
(gdb) bt
#0 0x00007f1eefe96aaa in pthread_cond_timedwait@@GLIBC_2.3.2 () from /lib64/libpthread.so.0
#1 0x00007f1ef0582f6f in PyEval_RestoreThread () from /usr/lib64/libpython3.8.so.1.0
#2 0x00007f1ef0608c59 in PyGILState_Ensure () from /usr/lib64/libpython3.8.so.1.0
#3 0x00007f1ec3003e7d in ?? () from /usr/lib64/python3.8/site-packages/scipy/fft/_pocketfft/pypocketfft.cpython-38-x86_64-linux-gnu.so
#4 0x00007f1ec2f7dcb5 in ?? () from /usr/lib64/python3.8/site-packages/scipy/fft/_pocketfft/pypocketfft.cpython-38-x86_64-linux-gnu.so
#5 0x00007f1ec30027ee in ?? () from /usr/lib64/python3.8/site-packages/scipy/fft/_pocketfft/pypocketfft.cpython-38-x86_64-linux-gnu.so
#6 0x00007f1ec2ff204d in ?? () from /usr/lib64/python3.8/site-packages/scipy/fft/_pocketfft/pypocketfft.cpython-38-x86_64-linux-gnu.so
```

```
#7 0x00007f1ec2ff4f78 in PyInit_pypocketfft () from /usr/lib64/python3.8/site-packages/scipy/fft/_pocketfft/pypocketfft.cpython-38-x86_64-linux-gnu.so
#8 0x00007f1ef060e3cf in _PyImport_LoadDynamicModuleWithSpec () from /usr/lib64/libpython3.8.so.1.0
#9 0x00007f1ef060fc1d in ?? () from /usr/lib64/libpython3.8.so.1.0
#10 0x00007f1ef0516d97 in ?? () from /usr/lib64/libpython3.8.so.1.0
#11 0x00007f1ef055db52 in PyVectorcall_Call () from /usr/lib64/libpython3.8.so.1.0
```

NOTE: scipy.fft is a new submodule since scipy 1.4.0 ([release note](#))

#10 - 04/29/2020 09:42 AM - Kiefer Chang

- Related to Bug #45147: Module 'diskprediction_local' takes forever to load added

Files

mgr.x.log	103 KB	01/02/2020	Volker Theile
-----------	--------	------------	---------------