RADOS - Bug #37871

Ceph cannot connect to any monitors if one of them has a DNS resolution problem

01/11/2019 11:39 AM - Jairo Llopis

Status:	New	**Start date:**	01/11/2019
Priority:	Normal	**Due date:**	
Assignee:		**% Done:**	0%
Category:	Administration/Usability	**Estimated time:**	0.00 hour
Target version:	v13.2.4	**Spent time:**	0.00 hour
Source:	Community (user)	**Affected Versions:**	v13.2.4
Tags:		**ceph-qa-suite:**	
Backport:		**Component(RADOS):**	MonClient
Regression:	No	**Pull request ID:**	
Severity:	4 - irritation	**Crash signature:**	
Reviewed:			

Description

My ceph cluster is configured with this:

```
mon host = mon1,mon2,mon3
```

If I remove the DNS entry for mon2 and, from mon1, I get status, it raises an error:

```
$ ceph -s
server name not found: mon2 (Name or service not known)
unable to parse addrs in 'mon1,mon2,mon3'
2019-01-11 11:31:10.269 7f9ec32dc700 -1 monclient: get_monmap_and_config cannot identify monitors
to contact
[errno 22] error connecting to the cluster
```

According to the docs:

the mon host configuration option only needs to be sufficiently up to date such that a client can reach one monitor that is currently online.

The above configuration matches that requirement, since both mon1 and mon3 can still be resolved.

An additional detail is that if I replace that config line by the actual IP addresses and then check, it properly connects to a monitor and returns a status:

```
mon host = 172.21.0.3,172.21.0.5,172.21.0.7
```

```
$ ceph -s
  cluster:
    id:     7060741a-8aad-5f55-b64e-c3f527e322f8
    health: HEALTH_WARN
            1/3 mons down, quorum mon1,mon3

  services:
    mon: 3 daemons, quorum mon1,mon3, out of quorum: mon2
```

```
   mgr: mon3(active), standbys: mon1
   osd: 4 osds: 0 up, 0 in

 data:
   pools:   0 pools, 0 pgs
   objects: 0   objects, 0 B
   usage:   0 B used, 0 B / 0 B avail
   pgs:
```

So, of course, as a workaround, I'm gonna start writing the raw IP address list there. But I still think this is a bug, because it should only fail to contact the cluster in case all DNS entries fail, not in case just one fails, the same way it fails in case all IP addresses cannot be contacted, not when just one fails. That's the whole point of resilience, isn't it?

History

#1 - 01/21/2019 10:21 PM - Greg Farnum

- Project changed from Ceph to RADOS

- Category changed from Monitor to Administration/Usability

- Component(RADOS) MonClient added

#2 - 01/22/2019 03:32 AM - Kefu Chai

i think the unresolvable address(es) is more of a configuration issue. and we should not ignore this. it's quite different from monitor which is not reachable, but its name can be resolved.

#3 - 01/22/2019 09:24 AM - Jairo Llopis

In practical terms, what's the difference between not being able to connect because the host name cannot be resolved, and not being able to connect because the host is down?

At the end of the day, you cannot connect to that server, but can still connect to others, so as long as Ceph can still work, I don't see a reason for it to stop doing it...