Ceph - Bug #2364

Bug # 2221 (Resolved): Monitor setup bugs

mon: can't specify monitor to join with -m

04/30/2012 03:48 PM - Greg Farnum

Status:	Won't Fix	**% Done:**	0%
Priority:	Normal	**Spent time:**	0.00 hour
Assignee:	Greg Farnum		
Category:	Monitor		
Target version:	v0.47		
Source:	Development	**Reviewed:**	
Tags:		**Affected Versions:**	
Backport:		**ceph-qa-suite:**	
Regression:	No	**Pull request ID:**	
Severity:	3 - minor	**Crash signature:**	

Description

```
gregf@kai:~/ceph/src [wip-mon-setup]$ ./ceph-mon -i b --mkfs --mon-data dev/mon.b -m 192.168.106.2
25:6789
./ceph-mon: mon.noname-a 192.168.106.225:6789/0 is local, renaming to mon.b
./ceph-mon: generated monmap has no fsid; use '--fsid <uuid>'
```

Yet the documentation says if you specify them that way, you will join. I'm not sure if it's an implementation or documentation bug yet.

History

#1 - 04/30/2012 04:09 PM - Greg Farnum

Oh, and if you do specify the fsid in the above step (apparently required, my bad) and try to start up the mon:

```
gregf@kai:~/ceph/src [wip-mon-setup]$ ./ceph-mon -i b --mon-data dev/mon.b -m 192.168.106.225:6789
starting mon.b rank 0 at 192.168.106.225:6789/0 mon_data dev/mon.b fsid eda79780-a9ea-4a1e-80aa-e3c4f12a811b
accepter.bind unable to bind to 192.168.106.225:6789: Address already in use
```

If you **don't** use -m in the mkfs step, it goes a little better when trying to start the actual mon daemon:

```
gregf@kai:~/ceph/src [wip-mon-setup]$ ./ceph-mon -i b --mon-data dev/mon.b -m 192.168.106.225:6789
2012-04-30 15:59:38.980387 7f2d98485780 -1 no public_addr or public_network specified, and mon.b not present i
n monmap or ceph.conf
```

And if you stick the public-addr on the end you get

```
gregf@kai:~/ceph/src [wip-mon-setup]$ ./ceph-mon -i b --mon-data dev/mon.b -m 192.168.106.225:6789 --public-ad
dr 192.168.106.225:6889
starting mon.b rank -1 at 192.168.106.225:6889/0 mon_data dev/mon.b fsid eda79780-a9ea-4a1e-80aa-e3c4f12a811b
```

And that appears to work properly.

#2 - 04/30/2012 04:39 PM - Greg Farnum

- Status changed from New to In Progress

Of course, at that point the -m is essentially ignored.

If I merge in my no-conf-necessary changes and run without a conf file, there's a big problem:

```
gregf@kai:~/ceph [wip-mon-setup]$ ./ceph-mon -i b --mkfs --mon-data src/dev/mon.b/ --fsid 969d3aef-0c89-440c-a
1ae-6c7c26b11219
2012-04-30 16:30:14.414642 7fc51ac58780 -1 did not load config file, using default settings.
unable to find any monitors in conf. please specify monitors via -m monaddr or -c ceph.conf
./ceph-mon: error generating initial monmap: (2) No such file or directory
usage: ceph-mon -i monid [--mon-data=pathtodata] [flags]
  --debug_mon n
        debug monitor level (e.g. 10)
  --mkfs
        build fresh monitor fs
--conf/-c       Read configuration from the given configuration file
-d              Run in foreground, log to stderr.
-f              Run in foreground, log to usual location.
--id/-i         set ID portion of my name
--name/-n       set name (TYPE.ID)
--version       show version and quit

   --debug_ms N
        set message debug level (e.g. 1)
2012-04-30 16:30:14.414988 7fc51ac58780 -1 asok(0x13b3000) AdminSocketConfigObs::init: failed: AdminSocket::bi
nd_and_listen: failed to bind the UNIX domain socket to '/var/run/ceph/ceph.name.asok': (2) No such file or di
rectory
2012-04-30 16:30:14.415001 7fc51ac58780 -1 asok(0x13b3000) AdminSocketConfigObs: failed to start AdminSocket
```

If you're having trouble parsing all that, it won't proceed because you aren't specifying any other monitors to connect to. But if you try and specify other monitors manually using -m, then as previously, it tries to bind to their address (at least if they're local — I haven't tried a multi-machine setup yet).

#3 - 04/30/2012 04:58 PM - Greg Farnum

It does work if you have a monmap, though (although it's noisy for things like lack of keyrings, admin socket locations, etc). And once you've provided the monmap on mkfs it appears to be okay if you use the -m flag when starting up the new mon. But it doesn't need it at all.

So while I'm inclined to call this a documentation bug, it would be nice if you could just use -m instead of a conf file or monmap. I think this is a

cleanup I need to do.

#4 - 05/01/2012 04:01 PM - Greg Farnum

I've investigated this and it only tries to bind to existing monitor addrs if they have a local IP in the list provided. I discussed this with Sage and he seemed to want to leave it that way, and I checked on a multi-machine setup and did not have an issue.

So I will clean up the docs a bit and talk to Carl to see if this could be the cause of what he was seeing, but I don't think it will result in any code cleanups.

#5 - 05/01/2012 04:49 PM - Greg Farnum

- Status changed from In Progress to Won't Fix

I didn't have any troubles doing any of this with multiple machines so it does appear to only be a problem if you're running multiple monitors on a node. That's a dev problem, not a user problem, so I just added a warning to the docs.