

Ceph - Bug #13089

mon: check for store writeability before participating in election

09/14/2015 07:36 PM - Samuel Just

Status:	Resolved	Start date:	09/14/2015
Priority:	Normal	Due date:	
Assignee:	Sage Weil	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Source:	other	Reviewed:	
Tags:		Affected Versions:	
Backport:	hammer,firefly	ceph-qa-suite:	
Regression:	No	Pull request ID:	
Severity:	3 - minor		
Description			
Related issues:			
Duplicated by Ceph - Bug #13200: mon: does not check for IO errors on every t...		Duplicate	09/22/2015
Copied to Ceph - Backport #13339: mon: check for store writeability before p...		Resolved	
Copied to Ceph - Backport #13252: mon: does not check for IO errors on every ...		Resolved	09/22/2015

Associated revisions

Revision 2fb7b1f0 - 09/23/2015 02:58 PM - Sage Weil

mon/MonitorDBStore: assert/crash if there is a write error

Do this globally instead of relying on the zillion mon callers to check the error code. There are no cases where we want to tolerate a commit failure.

Fixes: #13089

Signed-off-by: Sage Weil <sage@redhat.com>

Revision d2bfa190 - 09/27/2015 08:43 PM - Sage Weil

mon/MonitorDBStore: assert/crash if there is a write error

Do this globally instead of relying on the zillion mon callers to check the error code. There are no cases where we want to tolerate a commit failure.

Fixes: #13089

Signed-off-by: Sage Weil <sage@redhat.com>

(cherry picked from commit 2fb7b1f0e33ada7c9a1be3de2f7708eb0760cef)

Revision 20a4c0c4 - 10/02/2015 10:49 AM - Sage Weil

mon/MonitorDBStore: assert/crash if there is a write error

Do this globally instead of relying on the zillion mon callers to check the error code. There are no cases where we want to tolerate a commit failure.

Fixes: #13089

Signed-off-by: Sage Weil <sage@redhat.com>

(cherry picked from commit 2fb7b1f0e33ada7c9a1be3de2f7708eb0760fcef)

History

#1 - 09/14/2015 09:53 PM - Greg Farnum

Mmm, I'm not sure we want to do the naive thing (write a file and fsync) on every election as that will slow them down a lot in the common case. Probably if we just check store accessibility on bootup, flip a flag if we ever fail one, and check for it on every election?

#2 - 09/22/2015 02:58 PM - Sage Weil

Greg Farnum wrote:

Mmm, I'm not sure we want to do the naive thing (write a file and fsync) on every election as that will slow them down a lot in the common case. Probably if we just check store accessibility on bootup, flip a flag if we ever fail one, and check for it on every election?

Yeah I think a check on boot is sufficient.

#3 - 09/22/2015 02:59 PM - Kefu Chai

- Assignee set to Kefu Chai

#4 - 09/22/2015 04:28 PM - Joao Eduardo Luis

Can we have a bit more context please? What's the rationale behind this? The monitor should not start if leveldb is not open, and since it's not open it's not writeable. Is this in case the disk is full or the fs is read-only? Or something else I'm missing?

#5 - 09/23/2015 02:40 AM - Kefu Chai

- Status changed from New to Need Review

- Assignee changed from Kefu Chai to Sage Weil

<https://github.com/ceph/ceph/pull/6040>

#6 - 09/23/2015 02:50 AM - Kefu Chai

- Tags set to hammer

#7 - 09/23/2015 02:50 AM - Kefu Chai

- *Tags deleted (hammer)*
- *Backport set to hammer*

#8 - 09/25/2015 12:54 PM - Sage Weil

- *Status changed from Need Review to Pending Backport*

#9 - 10/02/2015 11:22 AM - Loic Dachary

- *Backport changed from hammer to hammer,firefly*

#10 - 11/10/2015 11:15 AM - Loic Dachary

- *Status changed from Pending Backport to Resolved*

#11 - 02/05/2016 09:59 AM - Kefu Chai

please note that the background compaction fails, if leveldb can not create new tables or remove merged tables.

so, if we "chmod -w store.db". and leveldb complains in this case, but it does not return error on Write() unless "leveldb_paranoid" is true, but this option is "false" by default. so "failed to write to db" assertion is not triggered, even if compaction fails with the default settings.