

RADOS - Bug #38945

osd: leaked pg refs on shutdown

03/27/2019 01:54 AM - Zengran Zhang

Status:	Resolved	Start date:	03/27/2019
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Source:		Reviewed:	
Tags:		Affected Versions:	
Backport:	luminous,mimic,nautilus	ceph-qa-suite:	
Regression:	No	Component(RADOS):	
Severity:	3 - minor	Pull request ID:	27206

Description

recovery_request_timer may hold some QueuePeeringEvs which PGRef, if we dont shutdown it earlier, it potentially cause the PGRef leak when kicking pg.

key log:

...

```
2019-03-26 06:27:52.492825 7fd0d1d1d700 10 osd.23 pg_epoch: 10415 pg[14.1b1s0( v 10356'133210
(9988'131707,10356'133210) local-lis/les=10411/10413 n=3676 ec=595/595 lis/c 10411/10326 les/c/f 10413/10327/0
10411/10411/10411) [23,13,4,16,2147483647,2147483647]p23(0) r=0 lpr=10411 pi=[10326,10411)/2 crt=10356'133210 lcod 0'0
mlcod 0'0 recovering+undersized+degraded+peered] handle_peering_event: epoch_sent: 10415 epoch_requested: 10415
DeferRecovery: delay 30
2019-03-26 06:27:52.492837 7fd0d1d1d700 10 osd.23 pg_epoch: 10415 pg[14.1b1s0( v 10356'133210
(9988'131707,10356'133210) local-lis/les=10411/10413 n=3676 ec=595/595 lis/c 10411/10326 les/c/f 10413/10327/0
10411/10411/10411) [23,13,4,16,2147483647,2147483647]p23(0) r=0 lpr=10411 pi=[10326,10411)/2 crt=10356'133210 lcod 0'0
mlcod 0'0 recovering+undersized+degraded+peered] state<Started/Primary/Active/Recovering>: defer recovery, retry delay 30
2019-03-26 06:28:22.204344 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] on_shutdown
2019-03-26 06:28:22.204367 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] cancel_copy_ops
2019-03-26 06:28:22.204383 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] cancel_flush_ops
2019-03-26 06:28:22.204397 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] cancel_proxy_ops
2019-03-26 06:28:22.204418 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] clear_backoffs
2019-03-26 06:28:22.204444 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] on_change
2019-03-26 06:28:22.204486 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] clear_async_reads
2019-03-26 06:28:22.204501 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
```

```

local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] clear_primary_state
2019-03-26 06:28:22.204523 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 luod=0'0 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] release_backoffs [14:8d800000:::head,14:8e000000:::head)
2019-03-26 06:28:22.204551 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 luod=0'0 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] cancel_recovery
2019-03-26 06:28:22.204568 7fd0c6d07700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210 (9988'131707,10356'133210)
local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0 10421/10421/10411)
[23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 luod=0'0 crt=10356'133210 lcod 0'0 mlcod 0'0
active+undersized+degraded] clear_recovery_state
2019-03-26 06:28:22.376456 7fd0c6d07700 -1 osd.23 10425 pgid 14.1b1s0 has ref count of 2
2019-03-26 06:28:22.492955 7fd0e3540700 10 osd.23 pg_epoch: 10425 pg[14.1b1s0( v 10356'133210
(9988'131707,10356'133210) local-lis/les=10421/10422 n=3676 ec=595/595 lis/c 10421/10326 les/c/f 10422/10327/0
10421/10421/10411) [23,13,4,16,24,2147483647]p23(0) r=0 lpr=10421 pi=[10326,10421)/2 luod=0'0 crt=10356'133210 lcod 0'0
mlcod 0'0 active+undersized+degraded] old_peering_msg reply_epoch 10415 query_epoch 10415 last_peering_reset 10421

```

Related issues:

Copied to RADOS - Backport #39204: luminous: osd: leaked pg refs on shutdown	Resolved
Copied to RADOS - Backport #39205: nautilus: osd: leaked pg refs on shutdown	Resolved
Copied to RADOS - Backport #39206: mimic: osd: leaked pg refs on shutdown	Resolved

History

#1 - 03/28/2019 07:43 AM - Kefu Chai

- Status changed from New to Need Review
- Pull request ID set to 27206

#2 - 03/28/2019 07:50 AM - Kefu Chai

- Backport set to luminous,mimc,nautilus

please note, in luminous, we also need to stop snap_sleep_timer and scrub_sleep_timer into OSDService::shutdown(). but this is relatively low priority, as it only impacts the shutdown behavior of OSD.

#3 - 03/28/2019 12:40 PM - Nathan Cutler

- Backport changed from luminous,mimc,nautilus to luminous,mimic,nautilus

#4 - 04/01/2019 01:48 PM - Kefu Chai

- Status changed from Need Review to Pending Backport

#5 - 04/10/2019 09:05 PM - Nathan Cutler

- Copied to Backport #39204: luminous: osd: leaked pg refs on shutdown added

#6 - 04/10/2019 09:05 PM - Nathan Cutler

- Copied to Backport #39205: nautilus: osd: leaked pg refs on shutdown added

#7 - 04/10/2019 09:05 PM - Nathan Cutler

- Copied to Backport #39206: mimic: osd: leaked pg refs on shutdown added

#8 - 06/20/2019 10:01 AM - Nathan Cutler

- Status changed from Pending Backport to Resolved