

fs - Bug #18743

Scrub considers dirty backtraces to be damaged, puts in damage table even though it repairs

01/31/2017 11:59 AM - John Spray

Status:	Resolved	Start date:	01/31/2017
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:	fsck/damage handling	Estimated time:	0.00 hour
Target version:	v12.0.0	Affected Versions:	
Source:		ceph-qa-suite:	
Tags:		Component(FS):	MDS
Backport:	luminous	Labels (FS):	
Regression:	No	Pull request ID:	
Severity:	3 - minor	Crash signature:	
Reviewed:			

Description

Two things are wrong here:

- When running scrub_path /teuthology-archive on the lab filesystem, I get a flurry of "bad backtrace" on very recently written files, which makes me 99% certain these are false positives for dirty metadata
- The scrub code goes ahead and repairs backtraces, but proceeds to put an entry in the damage table anyway.

For example:

```
2017-01-31 11:50:45.992113 7f1be3c6c700 0 log_channel(cluster) log [WRN] : Scrub error on inode [
inode 100221723c2 [2,head] /teuthology-archive/teuthology-2017-01-31_11:30:02-hadoop-kraken---basi
c-mira/768757/data/mon.2.tgz auth v8 dirtyparent s=119028 n(v0 b119028 l=1+0) (iversion lock) | di
rtyparent=1 scrubqueue=0 dirty=1 0x7f1bfb354998] (/teuthology-archive/teuthology-2017-01-31_11:30:
02-hadoop-kraken---basic-mira/768757/data/mon.2.tgz) see mds.mira049 log for details
2017-01-31 11:50:45.992150 7f1be3c6c700 -1 mds.0.scrubstack _validate_inode_done scrub error on in
ode [inode 100221723c2 [2,head] /teuthology-archive/teuthology-2017-01-31_11:30:02-hadoop-kraken--
-basic-mira/768757/data/mon.2.tgz auth v8 dirtyparent s=119028 n(v0 b119028 l=1+0) (iversion lock)
| dirtyparent=1 scrubqueue=0 dirty=1 0x7f1bfb354998]: {"performed_validation":true,"passed_valida
tion":false,"backtrace":{"checked":true,"passed":false,"read_ret_val":-61,"ondisk_value": "(-1)0:[
\\\\"/","memoryvalue": "(0)100221723c2:[<100221723bf\\mon.2.tgz v8>,<100221707e0\\data v283>,<1002217
07df\\768757 v582>,<100011cf577\\teuthology-2017-01-31_11:30:02-hadoop-kraken---basic-mira v294521
364>,<1\\teuthology-archive v139611246>]\\\\\\"/","error_str":"failed to read off disk; see retval"},"
raw_stats":{"checked":false,"passed":false,"read_ret_val":0,"ondisk_value.dirrstat":"f()","ondisk_v
alue.rstat":"n()","memory_value.dirrstat":"f()","memory_value.rstat":"n()","error_str":""},"return
_code":-61}
2017-01-31 11:50:46.012429 7f1be3c6c700 0 log_channel(cluster) log [WRN] : bad backtrace on inode
[inode 100221723c3 [...2,head] /teuthology-archive/teuthology-2017-01-31_11:30:02-hadoop-kraken--
-basic-mira/768757/remote/ auth v416 dirtyparent f(v0 m2017-01-31 11:42:21.894689 3=0+3) n(v13 rc2
017-01-31 11:43:03.254448 b2671453 108=98+10)/n(v12 rc2017-01-31 11:43:03.254448 b2671453 108=98+1
0) (inest lock dirty) (iversion lock) | dirtyscattered=1 lock=0 dirfrag=1 dirtyrstat=1 dirtyparent
=1 scrubqueue=0 dirty=1 authpin=0 0x7f1bfb354f80], rewriting it at /teuthology-archive/teuthology-
2017-01-31_11:30:02-hadoop-kraken---basic-mira/768757/remote
```

I suspect we would reproduce this rather easily if scrubbing during normal workloads in our automated testing (<http://tracker.ceph.com/issues/17856>)

Related issues:

Copied to fs - Backport #22089: luminous: Scrub considers dirty backtraces to...

Resolved

History

#1 - 01/31/2017 07:30 PM - Dan Mick

- *Description updated*

#2 - 10/25/2017 01:50 PM - John Spray

- *Status changed from New to Need Review*

- *Backport set to luminous*

<https://github.com/ceph/ceph/pull/18538>

Inspired to fix this from working on today's "[ceph-users] MDS damaged" thread.

#3 - 11/03/2017 09:41 PM - Patrick Donnelly

- *Status changed from Need Review to Pending Backport*

#4 - 11/08/2017 10:45 PM - Nathan Cutler

- *Copied to Backport #22089: luminous: Scrub considers dirty backtraces to be damaged, puts in damage table even though it repairs added*

#5 - 02/08/2018 05:51 PM - Nathan Cutler

- *Status changed from Pending Backport to Resolved*