

CephFS - Feature #8358

client: opportunistically update backtraces on files

05/14/2014 03:23 PM - Greg Farnum

Status:	New	% Done:	0%
Priority:	Normal		
Assignee:			
Category:	Performance/Resource Usage		
Target version:			
Source:	Development	Affected Versions:	
Tags:		Component(FS):	Client, Common/Protocol, MDS
Backport:		Labels (FS):	
Reviewed:		Pull request ID:	
Description			
<p>I've seen a few reports that backtrace updates are causing issues due to the increased write load (and the bursty nature of the current implementation; see #8230).</p> <p>Since most backtraces are simply set on creation and never updated, it should be possible for clients to write backtraces out if they're already touching the object and tell the MDS they've done so.</p>			

History

#1 - 07/02/2015 09:27 PM - Greg Farnum

I think the steps to do this are:

- 1) Have MDS provide a bufferlist to the client whenever an inode is created (as an optional part of the InodeStat)
- 2) If the client ever writes to the first object, it writes this bufferlist to a known xattr.
- 3) Tell the MDS that the bufferlist was written as part of a caps update. (This is the part I'm divided on: should it be some kind of full cap, or just a boolean we set?)
- 4) In the MDS, call `CInode::clear_dirty_parent()` if we get a response from the client indicating they set the backtrace.

#2 - 07/13/2016 12:02 AM - Greg Farnum

- Category changed from 46 to Performance/Resource Usage

- Component(FS) Client, Common/Protocol, MDS added