




Mantis - Production**Viewing Issue Advanced Details**

ID: 4631	Category: Server	Severity: minor	Reproducibility: N/A	Date Submitted: 2011-06-06 20:35	Last Update: 2011-06-22 20:08
Reporter: James Morelli	Platform:				
Assigned To: James Morelli	OS:				
Priority: normal	OS Version:				
Status: in progress	Product Version:				
Product Build:	Resolution: open				
Projection: none	Fixed in Version:				
ETC:					
Summary: 0004631: Kernel Panic - dequeue_task_fair					
Description: Strange stack trace caused possibly by multiple servers accessing the same file on a gluster volume. Stack trace attached.					
Steps To Reproduce:					
Additional Information:					
Relationships related to 0004579 feedback Cullin J Wible 1.0.47 - Deployment Notes					
Attached Files:					
 dev-sg6-ny1-dmesg-2011-06-05.txt [^] (72,883 bytes) 2011-06-06 20:35 [Delete]					
 dmesg-dev-sg6-ny1-6-8-2011.txt [^] (54,156 bytes) 2011-06-08 17:55 [Delete]					
 dev-sg6-ny1-hung-processes.txt [^] (7,376 bytes) 2011-06-08 17:55 [Delete]					
Issue History					
Date Modified	Username	Field	Change		
2011-06-06 20:35	James Morelli	New Issue			
2011-06-06 20:35	James Morelli	File Added: dev-sg6-ny1-dmesg-2011-06-05.txt			
2011-06-06 20:36	James Morelli	Status	new => in progress		
2011-06-07 20:21	James Morelli	Note Added: 0012738			
2011-06-08 17:54	James Morelli	Note Added: 0012757			
2011-06-08 17:55	James Morelli	File Added: dmesg-dev-sg6-ny1-6-8-2011.txt			
2011-06-08 17:55	James Morelli	File Added: dev-sg6-ny1-hung-processes.txt			
2011-06-08 18:53	James Morelli	Note Added: 0012758			
2011-06-09 17:37	Cullin J Wible	Note Added: 0012782			
2011-06-16 14:37	James Morelli	Note Added: 0012917			
2011-06-16 15:33	James Morelli	Note Added: 0012919			
2011-06-17 21:36	Cullin J Wible	Assigned To	=> James Morelli		
2011-06-21 16:09	Cullin J Wible	Note Added: 0013021			
2011-06-22 20:07	Cullin J Wible	Relationship added	related to 0004579		
2011-06-22 20:08	Cullin J Wible	Target Version	=> V_1.0.47		
Notes					

<p>(0012738) James Morelli 2011-06-07 20:21</p>	<p>Started bonnie++ on gluster afr volume from dev-sg5-ny1 and dev-sg6-ny1. Both servers are using the same gluster client and server version (3.2). Each server is using the same output directory /var/db/mm/sv/sv3/bonnietest as the temp directory.</p> <p>Each server is running two bonnie++ processes. one process on each box logs to /var/db/mm/sv/sv3/bonnietest/sg5.log and the other to /var/db/mm/sv/sv3/bonnietest/sg6.log so that they must both access the same file.</p> <p>Waiting to see what happens.</p>
<p>(0012757) James Morelli 2011-06-08 17:54</p>	<p>new stack trace occurred on dev-sg6-ny1</p> <p>INFO: task mm-sg-account-m:12903 blocked for more than 120 seconds. "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message. mm-sg-account-m D ffff880105eca040 0 12903 12795 0x00000000 ffff880105eca040 0000000000000082 00000000000117c0 ffff880139befb98 ffff880139bef810 0000000000000000 0000000000000000 0000000000000000 0000000000000001 0000000000000001 0000000000000001 0000000000000002</p> <p>Call Trace: [<ffffffff812facd9>] ? schedule+0x379/0xa20 [<ffffffff81039882>] ? try_to_wake_up+0xd2/0x3d0 [<ffffffff810af4e0>] ? sync_page+0x0/0x50 [<ffffffff812fb4b6>] ? io_schedule+0x56/0x90 [<ffffffff810af51b>] ? sync_page+0x3b/0x50 [<ffffffff812fb80a>] ? __wait_on_bit_lock+0x4a/0x90 [<ffffffff810af4bf>] ? __lock_page+0x5f/0x70 [<ffffffff81062460>] ? wake_bit_function+0x0/0x30 [<ffffffff810b97e7>] ? pagevec_lookup+0x17/0x20 [<ffffffff810b9b35>] ? invalidate_inode_pages2_range+0x295/0x2c0 [<ffffffff8105477e>] ? recalc_sigpending+0xe/0x30 [<ffffffffffa0326f1b>] ? fuse_request_send+0x21b/0x280 [fuse] [<ffffffffffa03250c6>] ? fuse_request_init+0x36/0x40 [fuse] [<ffffffffffa032c161>] ? fuse_do_open+0x121/0x180 [fuse] [<ffffffffffa032bf23>] ? fuse_finish_open+0xe3/0xf0 [fuse] [<ffffffffffa032c253>] ? fuse_open_common+0x93/0xa0 [fuse] [<ffffffffffa032c260>] ? fuse_open+0x0/0x10 [fuse] [<ffffffff810efe01>] ? __dentry_open+0xf1/0x2f0 [<ffffffff810fc127>] ? inode_permission+0x47/0xc0 [<ffffffff810fc85e>] ? finish_open+0xce/0x180 [<ffffffffff810feb7d>] ? do_filp_open+0x25d/0x700 [<ffffffffff81109d3f>] ? alloc_fd+0x3f/0x130 [<ffffffffff810efc25>] ? do_sys_open+0x65/0x110 [<ffffffffff81002bbb>] ? system_call_fastpath+0x16/0x1b</p> <p>This same trace appeared several times all referencing the same process id. Finally the deque_task_fail process occurred.</p> <p>listed the contents of /proc/*/fd/ for all hung processes and logged into a file. file is attached as "dev-sg6-ny1-hung-processes.txt"</p> <p>dmesg is also attached</p>
<p>(0012758) James Morelli 2011-06-08 18:53</p>	<p>Rebooted servers. Downgraded to gluster-3.1.3</p>
<p>(0012782) Cullin J Wible 2011-06-09 17:37</p>	<p>This crash always seems to happen on the sg-account-meta process.</p> <p>In reviewing this process:</p> <ul style="list-style-type: none"> - I confirmed that the process-map is correct and the processes are not running on both servers. This eliminates any potential issues with concurrent file issues. - It is the only backend application that uses the lucene index. So I have a suspicion that it might be crashing if the account-meta process holds the index open while an index update is occurring. <p>If my suspicion is correct, the best way to reproduce this would be to reset a number of accounts just before the account-meta run intervals to ensure that indexing and account-meta are running at the same time.</p> <p>However, given that we have now downgraded to Gluster 3.1.3 we will let the system run for a few days and see how things perform. Then we will try resetting some indexes and see what happens.</p> <p>After that, we can try re-upgrading to 3.2 and test again.</p>
<p>(0012917) James Morelli 2011-06-16 14:37</p>	<p>Have been running the system on gluster 3.1.3 since 6/8/2011. Recently we cleared the id3 data from the tmp directory in all the accounts on this gluster volume. I then reset the indexing for all accounts and reran mm-sg-account-meta.php and mm-sg-update-lucene.php. I did this twice from</p>

	<p>both bricks in the gluster volume.</p> <p>On both machines the processes managed to complete successfully.</p>
<p>(0012919) James Morelli 2011-06-16 15:33</p>	<p>Gluster is upgraded to 3.2.0. Will wait a day or two and then put some more stress on the gluster volume.</p>
<p>(0013021) Cullin J Wible 2011-06-21 16:09</p>	<p>This just happened running gluster 3.2.0-01 on a new host - dev-sg3-ny1.</p> <p>Given that this didn't happen for over a week when we were running 3.1.3 I think it's safe to say this is a 3.2 issue.</p> <pre>[root@dev-sg3-ny1 ~]# dmesg less "echo 0 > /proc/sys/kernel/hung_task_timeout_secs" disables this message. mm-sg-account-m D ffff88013b0aa7f0 0 19539 19532 0x00000000 ffff88013b0aa7f0 0000000000000082 0000000000117c0 ffff88013b04a6c8 ffff88013b04a340 ffff88013b04a378 ffff880bfd917c0 ffffffff8103e933 ffff88013b04a340 ffffffff81034fcd ffff880bfd91828 0000000000000002 Call Trace: [<fffffff8103e933>] ? dequeue_task_fair+0xe3/0x170 [<fffffff81034fcd>] ? __dequeue_entity+0x3d/0x50 [<fffffff812facd9>] ? schedule+0x379/0xa20 [<fffffff810af4e0>] ? sync_page+0x0/0x50 [<fffffff812fb4b6>] ? io_schedule+0x56/0x90 [<fffffff810af51b>] ? sync_page+0x3b/0x50 [<fffffff812fb80a>] ? __wait_on_bit_lock+0x4a/0x90 [<fffffff810af4bf>] ? __lock_page+0x5f/0x70 [<fffffff81062460>] ? wake_bit_function+0x0/0x30 [<fffffff810b97e7>] ? pagevec_lookup+0x17/0x20 [<fffffff810b9b35>] ? invalidate_inode_pages2_range+0x295/0x2c0 [<fffffff8105477e>] ? recalc_sigpending+0xe/0x30 [<fffffff8a0353f1b>] ? fuse_request_send+0x21b/0x280 [fuse] [<fffffff8a03520c6>] ? fuse_request_init+0x36/0x40 [fuse] [<fffffff8a0359161>] ? fuse_do_open+0x121/0x180 [fuse] [<fffffff8a0358f23>] ? fuse_finish_open+0xe3/0xf0 [fuse] [<fffffff8a0359253>] ? fuse_open_common+0x93/0xa0 [fuse] [<fffffff8a0359260>] ? fuse_open+0x0/0x10 [fuse] [<fffffff810efe01>] ? __dentry_open+0xf1/0x2f0 [<fffffff810fc127>] ? inode_permission+0x47/0xc0 [<fffffff810fc85e>] ? finish_open+0xce/0x180 [<fffffff810feb7d>] ? do_filp_open+0x25d/0x700 [<fffffff81109d3f>] ? alloc_fd+0x3f/0x130 [<fffffff810efc25>] ? do_sys_open+0x65/0x110 [<fffffff81002bbb>] ? system_call_fastpath+0x16/0x1b INFO: task mm-sg-account-m:19541 blocked for more than 120 seconds.</pre>