



Lustre 2.0 Alpha-2 Summary

Summary

The focus of the second Alpha release (Alpha-2) of 2.0 was to continue improve stability of Lustre while landing queued fixes to HEAD and completing additional bug fixes. 85 total fixes were landed during this cycle. This release was tested on RHEL5/x86_64 and SLES 10/x86_64 with both IB and TCP connectivity. Known failures are documented in the HEAD Daily Testing Document available from lustre.org. Additionally, IOR and simul runs were executed iteratively on a modest cluster (approx 75 clients) and run for 18 hours successfully with iSCSI attached HW in file per process mode. IOR and Simul were tested on a larger direct attach cluster and IOR has run successfully for 4 hours.

Timeline

The Alpha-2 bug fix period was from April 11, 2009 – May 7, 2009. The build on May 7 marks the Lustre 2.0 Alpha-2 release which was announced and available for download May 12, 2009.

Fixes Landed

Below is a summary listing of all unique landings to HEAD during the Lustre 2.0 Alpha-2 Milestone period. This is inclusive of all landings -- code fixes as well as test framework fixes.

1. 14109 Making new e2fsprogs releases *Attachment 17803*
2. 15983 IB allocations causing fragmentation and eventual OOM *Attachment 21686*
3. 18983 MDS Kernel panic: RIP: ptlrpc:target_committed_to_req+0x18/0x120 *Attachment 22821*
4. 19123 test-framework.sh:get_mds_dir failed to find MDS index *Attachment 22835*
5. 18844 Deadlock in usocklnd *Attachment 22805*
6. 17931 conf-sanity test_33a timed out *Attachment 22907*
7. 18988 kernel BUG at mm/filemap.c:2370! *Attachment 22896*
8. 18951 sanity.sh test 43a is hanging *Attachment 22636*
9. 19140 PLTDEBUG / SUBSYSTEM not accept string value. *Attachment 22971*
10. 19175 test-framework: check_catastrophe is noisy and broken *Attachment 22994*
11. 18582 MDS failover with quotas on resulted in lustre becoming unresponsive *Attachment 22828*
12. 18810 IB allocations causing fragmentation and eventual OOM *Attachment 22984*
13. 19116 start_client_loads assume no "-" in hostname *Attachment 22996*
14. 16890 small fixtures for quota *Attachment 22535*
15. 18711 kernel: LustreError: 987:0:(lu_object.c:1072:key_fini()) ASSERTION(atomic_read(&key->lct_used) > 1) failed *Attachment 22743*
16. 18689 ll_log_commit_* thread name too long *Attachment 22591*
17. 16715 NFS cannot open files in HEAD, MDS crashing *Attachment 22255*
18. 15977 conf-sanity test_39: @@@@@ FAIL: memory leak detected *Attachment 22855*

19. 18676 recovery-small test_60: FAIL: Changelog not empty: 30005 entries Attachment 22552
20. 13908 Distribute patchless client packages Attachment 22951
21. 18192 replay-single test_70b: @@@@@@ FAIL: rundbench load on clients failed Attachment 23018
22. 15332 Lack of congestion control in LNET causing timeouts in routed configuration Attachment 22843
23. 17974 add lazystatfs mount option to allow statfs(2) to skip down OSTs Attachment 23031
24. 18818 mds_set_info_client misuse of RMF_SETINFO_KEY Attachment 22857
25. 19047 use do_nodes() instead of looping on all nodes Add llog_reader man page (no attachment)
26. 19280 conf-sanity test_50e: error : glob /proc/{fs,sys}/{lnet,lustre}/osc/lustre-OST0000-osc/ost_server_uuid: No such process Attachment 23313
27. 19190 replay-single: rm: cannot lstat `/mnt/lustre/f65b': No such file or directory Attachment 23085
28. 13908 Distribute patchless client packages Attachment 23193
29. 19104 liblustre: Assertion `info->clt_counters[i].ctc_nr_locks_locked == 0' failed Attachment 22873
30. 19149 sanity-quota test_6: @@@@@@ FAIL: write filea failure, expect success Attachment 22919
31. 16312 build the lustre kernel the way the vendor does Attachment 22713
32. 16312 build the lustre kernel the way the vendor does Attachment 23158
33. 16312 build the lustre kernel the way the vendor does Attachment 23214
34. 18818 mds_set_info_client misuse of RMF_SETINFO_KEY Attachment 23131 (reverse patch - Attachment 22857)
35. 19262 test-framework.sh: ./functions.sh: No such file or directory Attachment 23065
36. 18818 mds_set_info_client misuse of RMF_SETINFO_KEY Attachment 22857 (Patch reapplied)
37. 18735 e2fsck mmp bug Landed directly, no attachment. Description: use O_DIRECT to read uncached MMP block from disk; Details : if the MMP block has been read from disk previously, use; O_DIRECT to ensure that it is re-read from the disk instead of from the page cache; Description: clear MMP e2fsck marker if e2fsck is interrupted; Details : clear MMP e2fsck marker from the SIGINT and SIGTERM signal handlers, or if e2fsck hits a fatal error. Description: add "clear-mmp" option to tune2fs to reset the MMP e2fsck marker; Details : add an option to run "tune2fs -f -E clear-mmp" to reset the MMP block, which is easier than disabling and re-enabling it. This is less necessary now that e2fsck will clear it itself in most cases, but may still be needed if e.g. e2fsck crashes.
38. 18735 e2fsck mmp bug Landed directly, no attachment. Fix duplicate error codes with upstream e2fsck. Fix patch context for problem.c and problem.h so that the patches apply cleanly and the hunks are inserted in the right part of the error table.
39. 18735 e2fsck mmp bug; Landed directly, no attachment. Add a regression test for the problem.h and problem.c error tables, to ensure there are not duplicate problem codes. This can result in hard to diagnose errors, wrong questions asked of the user, etc.
40. 18683 OSS crashes when I try to use tunefs.lustre to view an OST which another OSS owns. Attachment 22238
41. 19223 client doesn't try to reconnect Attachment 23039
42. 16839 ll_sync thread stay in waiting mds<=>ost recovery finished Attachment 23023
43. 19260 output from "e2fsck --mdsdb" may contain garbage characters following filenames Attachment 23143
44. 19154 replay-single test_20b: @@@@@@ FAIL: after 2707848 > before 2701704 Attachment 23008
45. 18280 fix ptlrpc_register_bulk printing of xid Attachment 23180
46. 18073 lctl get/set_param: wrong resolution for '.' in param name Attachment 21846; Attachment 21847
47. 18690 smaller hash bucket sizes, cleanups Attachment 22085
48. 18982 (fid_handler.c:222: __seq_server_alloc_meta()) ASSERTION(in->lsr_end > space-

>lsr_start) failed Attachment 22997
49. 18733 node crashes with general protection fault: 0000 [1] Attachment 22471
50. 18571 Kernel panic - not syncing: Out of memory and no killable processes... on OSS when iozone Attachment 23033
51. 16312 build the lustre kernel the way the vendor does Attachment 23183
52. 18851 replay-single 61b timed out Attachment 23178
53. 18790 Build failure on HEAD, RHEL 5, PPC64 Client Attachment 23034
54. 19171 Oops: EIP is at lov_quota_ctl+0x54a/0xeb0 [lquota] Attachment 22999
55. 19085 replay-single test_70b HEAD failures Attachment 23159
56. 18577 1.6.5 mdsrate performance is slower than 1.4.11/12 (MDS is not cpu bound!) Attachment 22326
57. 19051 HEAD liblustre t50b :liblustre is about to shutdown Attachment 22956
58. 19184 kmem_cache_create: duplicate cache jbd_4k Attachment 22906
59. 19058 ext3_remount() return code when mmp fails Attachment 22746
60. 15691 sanity.sh: FAIL: test_123a statahead is slow! failure seen on v1_6_5_RC1 with acc-sm Attachment 22883
61. 18096 Memory leak in obdecho Attachment 21038; Attachment 23165
62. 19143 fix t-f to support different pri- and sec- mds shared storage dev names Attachment 23014
63. 17671 o2iblnd with OFED1.4 support; I don't see landed flags set but bonsai has record of landing. Enable Lustre to be built with OFED 1.4.1.
64. 17402 LustreError: 5450:0:(lprocfs_status.c:779:lprocfs_free_client_stats()) ASSERTION(client_stat->nid_exp_ref_count == 0) failed:count -1 Attachment 23094
65. 19153 access freed memory in "llu_lov_setstripe_ea_info()" Attachment 22903
66. 19142 performance-sanity: client eviction, enospc Attachment 23218
67. 16818 recovery-small test_26b: @@@@ @@@@ FAIL: client not evicted from MDS Attachment 23155
68. 18674 clients not allowed to reconnect because a glimpse callback cannot be sent Attachment 22170
69. 18674 clients not allowed to reconnect because a glimpse callback cannot be sent Attachment 22126
70. 15332 Lack of congestion control in LNET causing timeouts in routed configuration Attachment 22843
71. 19319 multiop uses wrong open() prototype with O_CREAT flag Attachment 23153
72. 19308 sanity-quota test_14a: @@@@ @@@@ FAIL: unexpected: user(quota_usr) write a file successfully! Attachment 23235
73. 19058 ext3_remount() return code when mmp fails Attachment 22746
74. 19107 improve acc-sm liblustre test to use -e opt Attachment 22967
75. 18921 integrate type2 tests to acc-sm Attachment 23312
76. 16721 4*racer.sh: osd_object_ref_del() ASSERTION(inode->i_nlink > 0) Attachment 23244
77. 16721 4*racer.sh: osd_object_ref_del() ASSERTION(inode->i_nlink > 0) Attachment 23142
78. 16721 4*racer.sh: osd_object_ref_del() ASSERTION(inode->i_nlink > 0) Attachment 23102
79. 16721 4*racer.sh: osd_object_ref_del() ASSERTION(inode->i_nlink > 0) Attachment 22989
80. 19400 sanity test_160: @@@@ @@@@ FAIL: user index should be 6; is 1 Attachment 23278
81. 19264 sanity-quota: multiple failures: FAIL: System is error when query quota for block (u:quota_usr) Attachment 23274
82. 19280 conf-sanity test_50e: error : glob /proc/{fs,sys}/{lnet,lustre}/osc/lustre-OST0000-osc/ost_server_uuid: No such process Attachment 23313
83. 19151 performance-sanity test_3: @@@@ @@@@ FAIL: test_3 failed with 1 Attachment 23327
84. 17584 integrate kDMU with lustre build system Attachment 22293

85. 19450 sanity-quota test_18c hung Attachment 23379

Results

Please review the Lustre 2.0 Alpha Test Plan document and Head Daily Testing Results spreadsheet (both available on lustre.org) for detailed plans and daily testing results.

The Lustre 2.0 Alpha-2 (build) has the following characteristics:

- Testing completed for RHEL5/x86_64 and SLES10x85_64 using IB and TCP
- Passing consistently: ior, simul, runtests, sanity, dbench, bonnie, iozone, fsx, sanityn, recovery small, replay_dual, insanity, sanity_sec, replay_vbr, replay_ost_single
- Known Failures: liblustre (18721), replay single (18382,19356), sanity_quota (19450), performance_sanity (18755, 19151), conf_sanity(19471,19458,19459)
- IOR and simul passed on a 74 client cluster; run iteratively for 18 hours iSCSI, file per process. In direct attach on larger cluster, runs for up to 4 hours
- CMD, Quotas, Recovery, and 1.8->2.0 upgrades not yet tested

Lustre 2.0 Alpha-2 was delivered on May 12, 2009 and is available from:

http://downloads.lustre.org/public/lustre/Lustre_2.0_Alpha2/