

RELEASE NOTES

V12.2.4 LUMINOUS

This is the fourth bugfix release of Luminous v12.2.x long term stable release series. This was primarily intended to fix a few build, ceph-volume/ceph-disk and RGW issues. We recommend all the users of 12.2.x series to update.

NOTABLE CHANGES

- ceph-volume: adds support to zap encrypted devices ([issue#22878](#), [pr#20545](#), Andrew Schoen)
- ceph-volume: log the current running command for easier debugging ([issue#23004](#), [pr#20597](#), Andrew Schoen)
- ceph-volume: warn on mix of filestore and bluestore flags ([issue#23003](#), [pr#20568](#), Alfredo Deza)
- cmake: check bootstrap.sh instead before downloading boost ([issue#23071](#), [pr#20515](#), Kefu Chai)
- core: Backport of cache manipulation: issues #22603 and #22604 ([issue#22604](#), [issue#22603](#), [pr#20353](#), Adam C. Emerson)
- core: last-stat-seq returns 0 because osd stats are cleared ([issue#23093](#), [pr#20548](#), Sage Weil, David Zafman)
- core: Snapset inconsistency is detected with its own error ([issue#22996](#), [pr#20501](#), David Zafman)
- rgw: make init env methods return an error ([issue#23039](#), [pr#20564](#), Abhishek Lekshmanan)
- rgw: parse old rgw_obj with namespace correctly ([issue#22982](#), [pr#20566](#), Yehuda Sadeh)
- rgw: return valid Location element, CompleteMultipartUpload ([issue#22655](#), [pr#20266](#), Matt Benjamin)
- rgw: URL-decode S3 and Swift object-copy URLs ([issue#22121](#), [issue#22729](#), [pr#20236](#), Malcolm Lee, Matt Benjamin)
- rgw: use explicit index pool placement ([issue#22928](#), [pr#20565](#), Yehuda Sadeh)
- tools: ceph-disk: v12.2.2 unable to create bluestore osd using ceph-disk ([issue#22354](#), [pr#20563](#), Kefu Chai)
- tools: ceph-objectstore-tool: “\$OBJ get-omaphdr” and “\$OBJ list-omap” scan all pgs instead of using specific pg ([issue#21327](#), [pr#20283](#), David Zafman)

V12.2.3 LUMINOUS

This is the third bugfix release of Luminous v12.2.x long term stable release series. It contains a range of bug fixes and a few features across Bluestore, CephFS, RBD & RGW. We recommend all the users of 12.2.x series update.

NOTABLE CHANGES

- CephFS:*
 - The CephFS client now checks for older kernels’ inability to reliably clear dentries from the kernel dentry cache. The new option `client_die_on_failed_dentry_invalidate` (default: true) may be turned off to allow the client to proceed (dangerous!).

OTHER NOTABLE CHANGES

- bluestore: do not crash on over-large objects ([issue#22161](#), [pr#19630](#), Sage Weil)
- bluestore: OSD crash on boot with assert caused by Bluefs on flush write ([issue#21932](#), [pr#19047](#), Jianpeng Ma)
- build/ops: ceph-base symbols not stripped in debs ([issue#22640](#), [pr#19969](#), Sage Weil)
- build/ops: ceph-conf: dump parsed config in plain text or as json ([issue#21862](#), [pr#18842](#), Piotr Dałek)
- build/ops: ceph-mgr dashboard has dependency on python-jinja2 ([issue#22457](#), [pr#19865](#), John Spray)
- build/ops: ceph-volume fails when centos7 image doesn’t have lvm2 installed ([issue#22443](#), [issue#22217](#), [pr#20215](#), Nathan Cutler, Theofilos Mouratidis)
- build/ops: Default kernel.pid_max is easily exceeded during recovery on high OSD-count system ([issue#21929](#), [pr#19133](#), David Disseldorp, Kefu Chai)
- build/ops: install-deps.sh: revert gcc to the one shipped by distro ([issue#22220](#), [pr#19680](#), Kefu Chai)
- build/ops: luminous build fails with -without-radosgw ([issue#22321](#), [pr#19483](#), Jason Dillaman)
- build/ops: move ceph-* tool binaries out of ceph-test subpackage ([issue#22319](#), [issue#21762](#), [pr#19355](#), liuchang0812, Nathan Cutler, Kefu Chai, Sage Weil)
- build.ops: rpm: adjust ceph-{osdomap,kvstore,monstore}-tool feature move ([issue#22558](#), [pr#19839](#), Kefu Chai)
- ceph: cluster [ERR] Unhandled exception from module ‘balancer’ while running on mgr.x: ‘NoneType’ object has no attribute ‘iteritems’ in cluster log ([issue#22090](#), [pr#19023](#), Sage Weil)
- cephfs: cephfs-journal-tool: add “set pool_id” option ([issue#22631](#), [pr#20085](#), dongdong tao)
- cephfs: cephfs-journal-tool: tool would miss to report some invalid range ([issue#22459](#), [pr#19626](#), dongdong tao)

- cephfs: cephfs: potential adjust failure in lru_expire ([issue#22458](#), pr#19627, dongdong tao)
- cephfs: “ceph tell mds” commands result in “File exists” errors on client admin socket ([issue#21406](#), [issue#21967](#), pr#18831, Patrick Donnelly)
- cephfs: client: anchor Inode while trimming caps ([issue#22157](#), pr#19105, Patrick Donnelly)
- cephfs: client: avoid recursive lock in ll_get_vino ([issue#22629](#), pr#20086, dongdong tao)
- cephfs: client: dual client segfault with racing ceph_shutdown ([issue#21512](#), [issue#20988](#), pr#20082, Jeff Layton)
- cephfs: client: implement delegation support in userland cephfs ([issue#18490](#), pr#19480, Jeff Layton)
- cephfs: client: quit on failed remount during dentry invalidate test #19218 ([issue#22269](#), pr#19370, Patrick Donnelly)
- cephfs: List of filesystems does not get refreshed after a filesystem deletion ([issue#21599](#), pr#18730, John Spray)
- cephfs: MDS : Avoid the assert failure when the inode for the cap_export from other... ([issue#22610](#), pr#20300, Jianyu Li)
- cephfs: MDSMonitor: monitor gives constant “is now active in filesystem cephfs as rank” cluster log info messages ([issue#21959](#), pr#19055, Patrick Donnelly)
- cephfs: racy is_mounted() checks in libcephfs ([issue#21025](#), pr#17875, Jeff Layton)
- cephfs: src/mds/MDCache.cc: 7421: FAILED assert(CInode::count() == inode_map.size() + snap_inode_map.size()) ([issue#21928](#), pr#18912, “Yan, Zheng”)
- cephfs: vstart_runner: fixes for recent cephfs changes ([issue#22526](#), pr#19829, Patrick Donnelly)
- ceph-volume: adds a -destroy flag to ceph-volume lvm zap ([issue#22653](#), pr#20240, Andrew Schoen)
- ceph-volume: adds success messages for lvm prepare/activate/create ([issue#22307](#), pr#20238, Andrew Schoen)
- ceph-volume: dmcrypt support for lvm ([issue#22619](#), pr#20241, Alfredo Deza)
- ceph-volume dmcrypt support for simple ([issue#22620](#), pr#20350, Andrew Schoen, Alfredo Deza)
- ceph-volume: do not use -key during mkfs ([issue#22283](#), pr#20244, Kefu Chai, Sage Weil)
- ceph-volume: fix usage of the -osd-id flag ([issue#22642](#), [issue#22836](#), pr#20323, Andrew Schoen)
- ceph-volume Format correctly when vg/lv cannot be used ([issue#22299](#), pr#19527, Alfredo Deza)
- ceph-volume handle inline comments in the ceph.conf file ([issue#22297](#), pr#19532, Alfredo Deza)
- ceph-volume: handle leading whitespace/tabs in ceph.conf ([issue#22280](#), pr#19526, Alfredo Deza)
- ceph-volume: lvm zap will unmount osd paths used by zapped devices ([issue#22876](#), pr#20438, Andrew Schoen)
- ceph-volume: removed the explicit use of sudo ([issue#22282](#), pr#19525, Andrew Schoen)
- ceph-volume rollback on failed OSD prepare/create ([issue#22281](#), pr#20237, Alfredo Deza)
- ceph-volume should be able to handle multiple LVM (VG/LV) tags ([issue#22305](#), pr#19528, Alfredo Deza)
- ceph-volume use realpath when checking mounts ([issue#22988](#), pr#20429, Alfredo Deza)
- ceph-volume: warn on missing ceph.conf file ([issue#22326](#), pr#19530, Alfredo Deza)
- common: compute SimpleLRU’s size with contents.size() instead of lru.... ([issue#22613](#), pr#19977, Xuehan Xu)
- config: lower default omap entries recovered at once ([issue#21897](#), pr#19928, Josh Durgin)
- core: backoff causes out of order op ([issue#21407](#), pr#18747, Sage Weil)
- core: common/throttle: start using 64-bit values ([issue#22539](#), pr#19995, Igor Fedotov)
- core: fix broken use of streamstream::rdbuf() ([issue#22715](#), pr#20042, Sage Weil)
- core: possible deadlock in various maintenance operations ([issue#22120](#), pr#19123, Jason Dillaman)
- core: _read_bdev_label unable to decode label at offset ([issue#22285](#), pr#20326, Sage Weil)
- core: rocksdb: fixes early metadata spill over to slow device in ([issue#22264](#), pr#19257, Igor Fedotov)
- core: Various odd clog messages for mons ([issue#22082](#), pr#19031, John Spray)
- crush: balancer crush-compat sends “foo” command ([issue#22361](#), pr#19555, John Spray)
- doc: crush_ruleset is invalid command in luminous ([issue#20559](#), pr#19446, Nathan Cutler)
- doc: doc/rbd: tweaks for the LIO iSCSI gateway ([issue#21763](#), pr#20213, Ashish Singh, Mike Christie, Jason Dillaman)
- doc: man page for mount.fuse.ceph ([issue#21539](#), [issue#22595](#), pr#19449, Jos Collin)
- doc: misc fixes for CephFS best practices ([issue#22630](#), pr#19858, Jos Collin)
- doc: remove region from “INSTALL CEPH OBJECT GATEWAY” ([issue#21610](#), pr#18865, Orit Wasserman)
- doc: update Blacklisting and OSD epoch barrier ([issue#22542](#), pr#19741, Jos Collin)
- librbd: cannot clone all image-metas if we have more than 64 key/value pairs ([issue#21814](#), pr#19503, PCzhangPC)
- librbd: cannot copy all image-metas if we have more than 64 key/value pairs ([issue#21815](#), pr#19504, PCzhangPC)
- librbd: compare and write against a clone can result in failure ([issue#20789](#), pr#20211, Mykola Golub, Jason Dillaman)
- librbd: default to sparse-reads for any IO operation over 64K ([issue#21849](#), pr#20208, Jason Dillaman)
- librbd: fix snap create/rm may taking long time ([issue#22716](#), pr#20153, Song Shun)
- librbd: force removal of a snapshot cannot ignore dependent children ([issue#22791](#), pr#20135, Jason Dillaman)
- librbd: Image-meta should be dynamically refreshed ([issue#21529](#), pr#19447, Dongsheng Yang, Jason Dillaman)
- librbd: journal should ignore -EILSEQ errors from compare-and-write ([issue#21628](#), pr#20206, Jason Dillaman)
- librbd: refresh image after applying new/removing old metadata ([issue#21711](#), pr#19485, Jason Dillaman)
- librbd: set deleted parent pointer to null ([issue#22158](#), pr#20210, Jason Dillaman)
- luminous: ceph-fuse: ::rmdir() uses a deleted memory structure of dentry leads ... ([issue#22536](#), pr#19968, YunfeiGuan)
- mds: check for CEPH_OSDMAP_FULL is now wrong; cluster full flag is obsolete ([issue#22483](#), pr#19830, Patrick Donnelly)
- mds: don’t check gid when none specified in auth caps ([issue#22009](#), pr#18835, Douglas Fuller)
- mds: don’t delay processing completed requests in replay queue ([issue#22163](#), pr#19157, “Yan, Zheng”)
- mds: don’t report repaired backtraces in damagetable, write back after repair, clean up scrub log ([issue#18743](#), [issue#22058](#), pr#20341, “Yan, Zheng”, John Spray)
- mds: fix CDir::log_mark_dirty() ([issue#21584](#), pr#18008, “Yan, Zheng”)
- mds: fix dump last_sent ([issue#22562](#), pr#19959, dongdong tao)

- mds: fix MDS_FEATURE_INCOMPAT_FILE_LAYOUT_V2 definition ([issue#21985](#), [pr#18782](#), "Yan, Zheng")
- mds: fix return value of MDCache::dump_cache ([issue#22798](#), [pr#20121](#), "Yan, Zheng")
- mds: fix scrub crash ([issue#22730](#), [pr#20249](#), dongdong tao)
- mds: fix StrayManager::truncate() ([issue#21091](#), [pr#18019](#), "Yan, Zheng")
- mds: handle client reconnect gather race ([issue#22263](#), [pr#19326](#), "Yan, Zheng")
- mds: handle client session messages when mds is stopping ([issue#22460](#), [pr#19585](#), "Yan, Zheng")
- mds: handle 'inode gets queued for recovery multiple times' ([issue#22647](#), [pr#19982](#), "Yan, Zheng")
- mds: ignore export pin for unlinked directory ([issue#22219](#), [pr#19360](#), "Yan, Zheng")
- mds: limit size of subtree migration ([issue#21892](#), [pr#20339](#), "Yan, Zheng")
- mds: no assertion on inode being purging in find_ino_peers() ([issue#21722](#), [pr#18869](#), Zhi Zhang)
- mds: preserve order of requests during recovery of multimds cluster ([issue#21843](#), [pr#18871](#), "Yan, Zheng")
- mds: prevent filelock from being stuck at XSYN state ([issue#22008](#), [pr#20340](#), "Yan, Zheng")
- mds: properly eval locks after importing inode ([issue#22357](#), [pr#19646](#), "Yan, Zheng")
- mds: reduce debugging level for balancer messages ([issue#21853](#), [pr#19827](#), Patrick Donnelly)
- mds: respect mds_client_writeable_range_max_inc_objs config ([issue#22492](#), [pr#19776](#), "Yan, Zheng")
- mds: set higher priority for some perf counters ([issue#22776](#), [pr#20299](#), Shangzhong Zhu)
- mds: set PRIO_USEFUL on num_sessions counter ([issue#21927](#), [pr#18722](#), John Spray)
- mds: tell session ls returns vanila EINVAL when MDS is not active ([issue#21991](#), [pr#19505](#), Jos Collin)
- mds: track dirty dentries in separate list ([issue#19578](#), [pr#19775](#), "Yan, Zheng")
- mds: trim 'N' log segments according to how many log segments are there ([issue#21975](#), [pr#18783](#), "Yan, Zheng")
- mgr: ceph-mgr spuriously reloading OSD metadata on map changes ([issue#21159](#), [pr#18732](#), Yanhu Cao)
- mgr: disconnect unregistered service daemon when report received ([issue#22286](#), [pr#20089](#), Jason Dillaman)
- mgr: KeyError: ('name',) in balancer rm ([issue#22470](#), [pr#19624](#), Dan van der Ster)
- mgr: Manager daemon x is unresponsive. No standby daemons available ([issue#21147](#), [pr#19501](#), Sage Weil)
- mgr: mgr/balancer/upmap_max_iterations must be cast to integer ([issue#22429](#), [pr#19553](#), Dan van der Ster)
- mgr: mgr/dashboard: added iSCSI IOPS/throughput metrics ([issue#21391](#), [pr#20209](#), Jason Dillaman)
- mgr: mgr/dashboard: Fix PG status coloring ([issue#22615](#), [pr#19844](#), Wido den Hollander)
- mgr: mgr/prometheus: add missing 'deep' state to PG_STATES in ceph-mgr pro... ([issue#22116](#), [pr#19929](#), Jan Fajerski, Peter Woodman)
- mgr: mgr tests don't indicate failure if exception thrown from serve() ([issue#21999](#), [pr#18832](#), John Spray)
- mgr: mgr[zabbix] float division by zero (osd['kb'] = 0) ([issue#21904](#), [pr#19048](#), Ilja Slepnev)
- mgr: prometheus: added osd commit/apply latency metrics (#22718) ([issue#22718](#), [pr#20084](#), Konstantin Shalygin)
- mgr: pybind/mgr/dashboard: fix duplicated slash in html href ([issue#22851](#), [pr#20325](#), Shengjing Zhu)
- mgr: pybind/mgr/dashboard: fix reverse proxy support ([issue#22557](#), [pr#20182](#), Nick Erdmann, Kefu Chai)
- mgr: pybind/mgr/prometheus: fix metric type undef -> untyped ([issue#22313](#), [pr#19834](#), Ilya Margolin)
- mgr: restarting active ceph-mgr cause glitches in bps and iops metrics ([issue#21773](#), [pr#18735](#), Aleksei Gutikov, Kefu Chai)
- mgr: Services reported with blank hostname ([issue#20887](#), [issue#21687](#), [pr#17869](#), liuchang0812, Chang Liu)
- mon: do not use per_pool_sum_delta to show recovery summary ([issue#22727](#), [pr#20150](#), Chang Liu)
- mon: fix mgr using auth_client_required policy ([issue#22096](#), [pr#20156](#), John Spray)
- mon: MDSMonitor: reject misconfigured mds_blacklist_interval ([issue#21821](#), [pr#19871](#), John Spray)
- mon/MgrMonitor: limit mgrmap history ([issue#22257](#), [pr#19187](#), Sage Weil)
- mon: reenable timer to send digest when paxos is temporarily inactive ([issue#22142](#), [pr#19481](#), Jan Fajerski)
- msg: msg/async/AsyncConnection.cc: 1835: FAILED assert(state == STATE_CLOSED) ([issue#21883](#), [pr#18746](#), Haomai Wang)
- msg: msg/async: unregister connection failed when racing happened ([issue#22437](#), [pr#19552](#), Haomai Wang)
- osdc: "FAILED assert(bh->last_write_tid > tid)" in powercycle-wip-yuri-master-1.19.18-distro-basic-smithi ([issue#22741](#), [pr#20256](#), "Yan, Zheng")
- osdc/Journaler: add 'stopping' check to various finish callbacks ([issue#22360](#), [pr#19610](#), "Yan, Zheng")
- osdc/Objecter: objecter op_send_bytes perf counter always 0 ([issue#21982](#), [pr#19046](#), Jianpeng Ma)
- osd: do not check out-of-date osdmap for DESTROYED flag on start ([issue#22673](#), [pr#20068](#), Sage Weil)
- osd,mgr: report pending creating pg to mgr ([issue#22440](#), [pr#20204](#), Kefu Chai)
- osd: miscounting degraded objects and PG stuck in recovery_unfound ([issue#22145](#), [pr#20055](#), Sage Weil, David Zafman)
- osd: Objecter::C_ObjectOperation_sparse_read throws/catches exceptions on -ENOENT ([issue#21844](#), [pr#18744](#), Jason Dillaman)
- osd: Objecter::_send_op unnecessarily constructs costly hobject_t ([issue#21845](#), [pr#18745](#), Jason Dillaman)
- osd: On pg repair the primary is not favored as was intended ([issue#21907](#), [pr#19083](#), David Zafman)
- osd: OSD crushes with FAILED assert(used_blocks.size() > count) during the first start after upgrade 12.2.1 -> 12.2.2 ([issue#22535](#), [pr#19888](#), Igor Fedotov)
- osd: OSDMap cache assert on shutdown ([issue#21737](#), [pr#18749](#), Greg Farnum)
- osd: OSDService::recovery_need_sleep read+updated without locking ([issue#21566](#), [pr#18753](#), Neha Ojha)
- osd: "osd status" command exception if OSD not in pgmap stats ([issue#21707](#), [pr#19084](#), Yanhu Cao)
- osd, pg, mgr: make snap trim queue problems visible ([issue#22448](#), [pr#20098](#), Piotr Dałek)
- osd: Pool Compression type option doesn't apply to new OSDs ([issue#22419](#), [pr#20106](#), Kefu Chai)

- osd: replica read can trigger cache promotion ([issue#20919](#), [pr#19499](#), Sage Weil)
- osd/ReplicatedPG.cc: recover_replicas: object added to missing set for backfill, but is not in recovering, error! ([issue#21382](#), [issue#14513](#), [issue#18162](#), [pr#20081](#), David Zafman)
- osd: subscribe osdmmaps if any pending pgs ([issue#22113](#), [pr#19059](#), Kefu Chai)
- osd: “sudo cp /var/lib/ceph/osd/ceph-0/fsid ...” fails ([issue#20736](#), [pr#19631](#), Patrick Donnelly)
- os: fix 0-length zero semantics, test ([issue#21712](#), [pr#20049](#), Sage Weil)
- qa/tests: Applied PR 20053 to stress-split tests ([issue#22665](#), [pr#20451](#), Yuri Weinstein)
- rbd: abort in listing mapped nbd devices when running in a container ([issue#22012](#), [issue#22011](#), [pr#19051](#), Li Wang)
- rbd: [api] compare-and-write methods not properly advertised ([issue#22036](#), [pr#18834](#), Jason Dillaman)
- rbd: class rbd.Image discard---OSError: [errno 2147483648] error discarding region ([issue#21966](#), [pr#19058](#), Jason Dillaman)
- rbd: cluster resource agent ocf:ceph:rbd - wrong permissions ([issue#22362](#), [pr#19554](#), Nathan Cutler)
- rbd: disk usage on empty pool no longer returns an error message ([issue#22200](#), [pr#19107](#), Jason Dillaman)
- rbd: fix crash during map ([issue#21808](#), [pr#18698](#), Peter Keresztes Schmidt)
- rbd: [journal] tags are not being expired if no other clients are registered ([issue#21960](#), [pr#18840](#), Jason Dillaman)
- rbd: librbd: filter out potential race with image rename ([issue#18435](#), [pr#19853](#), Jason Dillaman)
- rbd-mirror: Allow a different data-pool to be used on the secondary cluster ([issue#21088](#), [pr#19305](#), Adam Wolfe Gordon)
- rbd-mirror: primary image should register in remote, non-primary image’s journal ([issue#21961](#), [issue#21561](#), [pr#20207](#), Jason Dillaman)
- rbd-mirror: sync image metadata when transferring remote image ([issue#21535](#), [pr#19484](#), Jason Dillaman)
- rbd: Python RBD metadata_get does not work ([issue#22306](#), [pr#19479](#), Mykola Golub)
- rbd: rbd ls -l crashes with SIGABRT ([issue#21558](#), [pr#19800](#), Jason Dillaman)
- rbd: [rbd-mirror] new pools might not be detected ([issue#22461](#), [pr#19625](#), Jason Dillaman)
- rbd: [rbd-nbd] Fedora does not register resize events ([issue#22131](#), [pr#19066](#), Jason Dillaman)
- rbd: [test] UpdateFeatures RPC message should be included in test_notify.py ([issue#21936](#), [pr#18838](#), Jason Dillaman)
- Revert “luminous: msg/async: unregister connection failed when racing happened” ([issue#22231](#), [pr#20247](#), Sage Weil)
- rgw: 501 is returned When init multipart is using V4 signature and chunk encoding ([issue#22129](#), [pr#19506](#), Jeegn Chen)
- rgw: add cors header rule check in cors option request ([issue#22002](#), [pr#19053](#), yuliyang)
- rgw: backport beast frontend and boost 1.66 update ([issue#22101](#), [issue#20935](#), [issue#21831](#), [issue#20048](#), [issue#22600](#), [issue#20971](#), [pr#19848](#), Casey Bodley, Jiaying Ren)
- rgw: bucket index object not deleted after radosgw-admin bucket rm -purge-objects -bypass-gc ([issue#22122](#), [issue#19959](#), [pr#19085](#), Aleksei Gutikov)
- rgw: bucket policy evaluation logical error ([issue#21901](#), [issue#21896](#), [pr#19810](#), Adam C. Emerson)
- rgw: bucket resharding should not update bucket ACL or user stats ([issue#22742](#), [issue#22124](#), [pr#20327](#), Orit Wasserman)
- rgw: check going_down() when lifecycle processing ([issue#22099](#), [pr#19088](#), Yao Zongyou)
- rgw: Dynamic bucket indexing, resharding and tenants seems to be broken ([issue#22046](#), [pr#19050](#), Orit Wasserman)
- rgw: file deadlock on lru evicting ([issue#22736](#), [pr#20075](#), Matt Benjamin)
- rgw: fix chained cache invalidation to prevent cache size growth ([issue#22410](#), [pr#19785](#), Mark Kogan)
- rgw: fix for empty query string in beast frontend ([issue#22797](#), [pr#20338](#), Casey Bodley)
- rgw: fix GET website response error code ([issue#22272](#), [pr#19489](#), Dmitry Plyakin)
- rgw: fix rewrite a versioning object create a new object bug ([issue#21984](#), [issue#22529](#), [pr#19787](#), Enming Zhang, Matt Benjamin)
- rgw: Fix swift object expiry not deleting objects ([issue#22084](#), [pr#18972](#), Pavan Rallabhandi)
- rgw: Fix swift object expiry not deleting objects ([issue#22084](#), [pr#19090](#), Pavan Rallabhandi)
- rgw: librgw: fix shutdown error with resources uncleared ([issue#22296](#), [pr#20073](#), Tao Chen)
- rgw: log keystone errors at a higher level ([issue#22151](#), [pr#19077](#), Abhishek Lekshmanan)
- rgw: make HTTP dechunking compatible with Amazon S3 ([issue#21015](#), [pr#19500](#), Radoslaw Zarzynski)
- rgw: modify s3 type subuser access permission fail ([issue#21983](#), [pr#18766](#), yuliyang)
- rgw: multisite: destination zone does not compress synced objects ([issue#21895](#), [pr#18867](#), Casey Bodley)
- rgw: multisite: ‘radosgw-admin sync error list’ contains temporary EBUSY errors ([issue#22473](#), [pr#19799](#), Casey Bodley)
- rgw: null instance mtime incorrect when enable versioning ([issue#21743](#), [pr#18870](#), Shasha Lu)
- rgw: Policy parser may or may not dereference uninitialized boost::optional sometimes ([issue#21962](#), [pr#18868](#), Adam C. Emerson)
- rgw: Possible deadlock in ‘list_children’ when refresh is required ([issue#21670](#), [pr#18564](#), Jason Dillaman)
- rgw: put bucket policy panics RGW process ([issue#22541](#), [pr#19847](#), Bingyin Zhang)
- rgw: radosgw-admin reshard command argument error ([issue#21723](#), [pr#19502](#), Yao Zongyou)
- rgw: radosgw-admin zonegroup get and zone get should return defaults when there is no realm ([issue#21615](#), [pr#19086](#), Lvshanchun)
- rgw: Random 500 errors in Swift PutObject (needs cache fixes) ([issue#22517](#), [issue#21560](#), [pr#19788](#), Adam C. Emerson)
- rgw: refuses upload when Content-Type missing from POST policy ([issue#20201](#), [pr#19867](#), Matt Benjamin)
- rgw: revert PR #8765 ([issue#22364](#), [pr#19434](#), fang.yuxiang)
- rgw: RGWCrashError: RGW will crash if a putting lc config request does not include an ID tag in the request xml

(issue#21980, issue#22006, pr#18765, Enming Zhang)

- rgw: rgw multisite: automated trimming for bucket index logs (issue#18229, pr#20062, Casey Bodley)
- rgw: RGW: S3 POST policy should not require Content-Type (issue#20201, pr#19784, Matt Benjamin)
- rgw: rgw segfaults after running radosgw-admin data sync init (issue#22083, pr#19071, Casey Bodley, Abhishek Lekshmanan)
- rgw: rgw usage trim only trims a few entries (issue#22234, pr#19636, Abhishek Lekshmanan)
- rgw: S3 API Policy Conditions IpAddress and NotIpAddress do not work (issue#20931, issue#20991, pr#19819, John Gibson, yuliyang, Casey Bodley, Abhishek Lekshmanan, Jiaying Ren)
- rgw: Segmentation fault when starting radosgw after reverting .rgw.root (issue#21996, pr#18764, Orit Wasserman, Casey Bodley)
- rgw: set sync_from_all as true when no value is seen (issue#22062, pr#19038, Abhishek Lekshmanan)
- rgw: unlink deleted bucket from bucket's owner (issue#22248, pr#20357, Casey Bodley)
- rgw: user stats increased after bucket reshards (issue#22124, pr#19538, Orit Wasserman)
- rgw: When a system object is created exclusively, do not distribute the (issue#22792, pr#20107, J. Eric Ivancich, Robin H. Johnson)
- tests: ceph_test_cls_log failures related to cls_cxx_subop_version() (issue#21964, pr#18715, Casey Bodley)
- tests: ceph_test_objectstore fails ObjectStore/StoreTest.Synthetic/1 (filestore) buffer content mismatch (issue#21712, issue#21818, pr#18742, Sage Weil)
- tests: configure zabbix properly before selftest (issue#22514, pr#19831, John Spray)
- tests: do not configure ec data pool with memstore (issue#22436, pr#19628, Patrick Donnelly)
- tests: force backfill test can conflict with pool removal (issue#22614, pr#19966, Sage Weil)
- tests: full flag not set on osdmap for tasks.cephfs.test_full (issue#22475, pr#19962, Patrick Donnelly)
- tests: increase osd count for ec testing (issue#22646, pr#19976, Patrick Donnelly)
- tests - Initial checkin for luminous point-to-point upgrade (issue#22048, pr#18771, Yuri Weinstein)
- tests: qa/workunits/rbd: simplify split-brain test to avoid potential race (issue#22485, pr#20205, Jason Dillaman)
- tests: qa/workunits/rbd: switch devstack to pike release (issue#22786, pr#20136, Jason Dillaman)
- tests: rbd_mirror_helpers.sh request_resync_image function saves image id to wrong variable (issue#21663, pr#19802, Jason Dillaman)
- tools/ceph_monstore_tool: include mgrmap in initial paxos epoch (issue#22266, pr#20116, Kefu Chai)
- tools: ceph-monstore-tool -readable mode doesn't understand FSMap, MgrMap (issue#21577, pr#18754, John Spray)
- tools: ceph-objectstore-tool: Add option dump-import to examine an export (issue#22086, pr#19487, David Zafman)
- tools: ceph_objectstore_tool: no flush before collection_empty() calls; ObjectStore/StoreTest.SimpleAttrTest/2 fails (issue#22409, pr#19967, Igor Fedotov)
- tools: ceph-objectstore-tool set-size should clear data-digest (issue#22112, pr#20069, David Zafman)
- tools/crushtool: skip device id if no name exists (issue#22117, pr#19039, Jan Fajerski)

V12.2.2 LUMINOUS

This is the second bugfix release of Luminous v12.2.x long term stable release series. It contains a range of bug fixes and a few features across Bluestore, CephFS, RBD & RGW. We recommend all the users of 12.2.x series update.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- Standby ceph-mgr daemons now redirect requests to the active messenger, easing configuration for tools & users accessing the web dashboard, restful API, or other ceph-mgr module services.
- The prometheus module has several significant updates and improvements.
- The new balancer module enables automatic optimization of CRUSH weights to balance data across the cluster.
- The ceph-volume tool has been updated to include support for BlueStore as well as FileStore. The only major missing ceph-volume feature is dm-crypt support.
- RGW's dynamic bucket index resharding is disabled in multisite environments, as it can cause inconsistencies in replication of bucket indexes to remote sites.

OTHER NOTABLE CHANGES

- build/ops: bump sphinx to 1.6 (issue#21717, pr#18167, Kefu Chai, Alfredo Deza)
- build/ops: macros expanding in spec file comment (issue#22250, pr#19173, Ken Dreyer)
- build/ops: python-numpy-devel build dependency for SUSE (issue#21176, pr#17692, Nathan Cutler)
- build/ops: selinux: Allow getattr on lnk sysfs files (issue#21492, pr#18650, Boris Ranto)
- build/ops: Ubuntu amd64 client can not discover the ubuntu arm64 ceph cluster (issue#19705, pr#18293, Kefu Chai)
- core: buffer: fix ABI breakage by removing list _mempool member (issue#21573, pr#18491, Sage Weil)

- core: Daemons(OSD, Mon...) exit abnormally at injectargs command ([issue#21365](#), [pr#17864](#), Yan Jun)
- core: Disable messenger logging (debug ms = 0/0) for clients unless overridden ([issue#21860](#), [pr#18529](#), Jason Dillaman)
- core: Improve OSD startup time by only scanning for omap corruption once ([issue#21328](#), [pr#17889](#), Luo Kexue, David Zafman)
- core: upmap does not respect osd reweights ([issue#21538](#), [pr#18699](#), Theofilos Mouratidis)
- dashboard: barfs on nulls where it expects numbers ([issue#21570](#), [pr#18728](#), John Spray)
- dashboard: OSD list has servers and osds in arbitrary order ([issue#21572](#), [pr#18736](#), John Spray)
- dashboard: the dashboard uses absolute links for filesystems and clients ([issue#20568](#), [pr#18737](#), Nick Erdmann)
- filestore: set default readahead and compaction threads for rocksdb ([issue#21505](#), [pr#18234](#), Josh Durgin, Mark Nelson)
- librbd: object map batch update might cause OSD suicide timeout ([issue#21797](#), [pr#18416](#), Jason Dillaman)
- librbd: snapshots should be created/removed against data pool ([issue#21567](#), [pr#18336](#), Jason Dillaman)
- mds: make sure snap inode's last matches its parent dentry's last ([issue#21337](#), [pr#17994](#), "Yan, Zheng")
- mds: sanitize mdsmap of removed pools ([issue#21945](#), [issue#21568](#), [pr#18628](#), Patrick Donnelly)
- mgr: bulk backport of ceph-mgr improvements ([issue#21594](#), [issue#17460](#), [issue#21197](#), [issue#21158](#), [issue#21593](#), [pr#18675](#), Benjeman Meekhof, Sage Weil, Jan Fajerski, John Spray, Kefu Chai, My Do, Spandan Kumar Sahu)
- mgr: ceph-mgr gets process called "exe" after respawn ([issue#21404](#), [pr#18738](#), John Spray)
- mgr: fix crashable DaemonStateIndex::get calls ([issue#17737](#), [pr#18412](#), John Spray)
- mgr: key mismatch for mgr after upgrade from jewel to luminous(dev) ([issue#20950](#), [pr#18727](#), John Spray)
- mgr: mgr status module uses base 10 units ([issue#21189](#), [issue#21752](#), [pr#18257](#), John Spray, Yanhu Cao)
- mgr: mgr[zabbix] float division by zero ([issue#21518](#), [pr#18734](#), John Spray)
- mgr: Prometheus crash when update ([issue#21253](#), [pr#17867](#), John Spray)
- mgr: prometheus module generates invalid output when counter names contain non-alphanum characters ([issue#20899](#), [pr#17868](#), John Spray, Jeremy H Austin)
- mgr: Quieten scary RuntimeError from restful module on startup ([issue#21292](#), [pr#17866](#), John Spray)
- mgr: Spurious ceph-mgr failovers during mon elections ([issue#20629](#), [pr#18726](#), John Spray)
- mon: Client client.admin marked osd.2 out, after it was down for 1504627577 seconds ([issue#21249](#), [pr#17862](#), John Spray)
- mon: DNS SRV default service name not used anymore ([issue#21204](#), [pr#17863](#), Kefu Chai)
- mon/MgrMonitor: handle cmd desc to/from disk in the absence of active mgr ([issue#21300](#), [pr#18038](#), Joao Eduardo Luis)
- mon/mgr: sync "mgr_command_descs","osd_metadata" and "mgr_metadata" prefixes to new mons ([issue#21527](#), [pr#18620](#), huanwen ren)
- mon: osd feature checks with 0 up osds ([issue#21471](#), [issue#20751](#), [pr#18364](#), Brad Hubbard, Sage Weil)
- mon,osd: fix "pg ls {forced_backfill, backfilling}" ([issue#21609](#), [pr#18236](#), Kefu Chai)
- mon/OSDMonitor: add option to fix up ruleset-* to crush-* for ec profiles ([issue#22128](#), [pr#18945](#), Sage Weil)
- mon, osd: per pool space-full flag support ([issue#21409](#), [pr#17730](#), xie xinguo)
- mon/PGMap: Fix %USED calculation ([issue#22247](#), [pr#19230](#), Xiaoxi Chen)
- mon: update get_store_prefixes implementations ([issue#21534](#), [pr#18621](#), John Spray, huanwen ren)
- msgr: messages/MOSDMap: do compat reencode of crush map, too ([issue#21882](#), [pr#18456](#), Sage Weil)
- msgr: src/messages/MOSDMap: reencode OSDMap for older clients ([issue#21660](#), [pr#18140](#), Sage Weil)
- os/bluestore/BlueFS: fix race with log flush during async log compaction ([issue#21878](#), [pr#18503](#), Sage Weil)
- os/bluestore: fix another aio stall/deadlock ([issue#21470](#), [pr#18127](#), Sage Weil)
- os/bluestore: fix SharedBlob unregistration ([issue#22039](#), [pr#18983](#), Sage Weil)
- os/bluestore: handle compressed extents in blob unsharing checks ([issue#21766](#), [pr#18501](#), Sage Weil)
- os/bluestore: replace 21089 repair with something online (instead of fsck) ([issue#21089](#), [pr#17734](#), Sage Weil)
- os/bluestore: set bitmap freelist resolution to min_alloc_size ([issue#21408](#), [pr#18050](#), Sage Weil)
- os/blueStore::umount will crash when the BlueStore is opened by start_kv_only() ([issue#21624](#), [pr#18750](#), Chang Liu)
- osd: additional protection for out-of-bounds EC reads ([issue#21629](#), [pr#18413](#), Jason Dillaman)
- osd: allow recovery preemption ([issue#21613](#), [pr#18025](#), Sage Weil)
- osd: build_past_intervals_parallel: Ignore new partially created PGs ([issue#21833](#), [pr#18673](#), David Zafman)
- osd: dump bluestore debug on shutdown if debug option is set ([issue#21259](#), [pr#18103](#), Sage Weil)
- osd: make stat_bytes and stat_bytes_used counters PRIO_USEFUL ([issue#21981](#), [pr#18723](#), Yao Zongyou)
- osd: make the PG's SORTBITWISE assert a more generous shutdown ([issue#20416](#), [pr#18132](#), Greg Farnum)
- osd: OSD metadata 'backend_filestore_dev_node' is unknown even for simple deployment ([issue#20944](#), [pr#17865](#), Sage Weil)
- rbd: [cli] mirror getter commands will fail if mirroring has never been enabled ([issue#21319](#), [pr#17861](#), Jason Dillaman)
- rbd: cls/journal: fixed possible infinite loop in expire_tags ([issue#21956](#), [pr#18626](#), Jason Dillaman)
- rbd: cls/journal: possible infinite loop within tag_list class method ([issue#21771](#), [pr#18417](#), Jason Dillaman)
- rbd: [rbd-mirror] asok hook names not updated when image is renamed ([issue#20860](#), [pr#17860](#), Mykola Golub)
- rbd: [rbd-mirror] forced promotion can result in incorrect status ([issue#21559](#), [pr#18337](#), Jason Dillaman)
- rbd: [rbd-mirror] peer cluster connections should filter out command line optionals ([issue#21894](#), [pr#18566](#), Jason Dillaman)
- rgw: add support for Swift's per storage policy statistics ([issue#17932](#), [issue#21506](#), [pr#17835](#), Radoslaw Zarzynski, Casey Bodley)

- rgw: add support for Swift's reversed account listings ([issue#21148](#), [pr#17834](#), Radoslaw Zarzynski)
- rgw: avoid logging keystone revocation failures when no keystone is configured ([issue#21400](#), [pr#18441](#), Abhishek Lekshmanan)
- rgw: disable dynamic resharding in multisite enviorment ([issue#21725](#), [pr#18432](#), Orit Wasserman)
- rgw: encryption: PutObj response does not include sse-kms headers ([issue#21576](#), [pr#18442](#), Casey Bodley)
- rgw: encryption: reject requests that don't provide all expected headers ([issue#21581](#), [pr#18429](#), Enming Zhang)
- rgw: expose -sync-stats via admin api ([issue#21301](#), [pr#18439](#), Nathan Johnson)
- rgw: failed CompleteMultipartUpload request does not release lock ([issue#21596](#), [pr#18430](#), Matt Benjamin)
- rgw_file: set s->obj_size from bytes_written ([issue#21940](#), [pr#18599](#), Matt Benjamin)
- rgw: fix a bug about inconsistent unit of comparison ([issue#21590](#), [pr#18438](#), gaosibei)
- rgw: fix bilog entries on multipart complete ([issue#21772](#), [pr#18334](#), Casey Bodley)
- rgw: fix error handling in ListBucketIndexesCR ([issue#21735](#), [pr#18591](#), Casey Bodley)
- rgw: fix refcnt issues ([issue#21819](#), [pr#18539](#), baixueyu)
- rgw: lc process only schdule the first item of lc objects ([issue#21022](#), [pr#17859](#), Shasha Lu)
- rgw: list bucket which enable versioning get wrong result when user marker ([issue#21500](#), [pr#18569](#), yuliyang)
- rgw: list_objects() honors end_marker regardless of namespace ([issue#18977](#), [pr#17832](#), Radoslaw Zarzynski)
- rgw: Multipart upload may double the quota ([issue#21586](#), [pr#18435](#), Sibe Gao)
- rgw: multisite: Get bucket location which is located in another zonegroup, will return 301 Moved Permanently ([issue#21125](#), [pr#17857](#), Shasha Lu)
- rgw: multisite: race between sync of bucket and bucket instance metadata ([issue#21990](#), [pr#18767](#), Casey Bodley)
- rgw: policy checks missing from Get/SetRequestPayment operations ([issue#21389](#), [pr#18440](#), Adam C. Emerson)
- rgw: radosgw-admin usage show loops indefinitely ([issue#21196](#), [pr#18437](#), Mark Kogan)
- rgw: rgw_file: explicit NFSv3 open() emulation ([issue#21854](#), [pr#18446](#), Matt Benjamin)
- rgw: rgw_file: fix write error when the write offset overlaps ([issue#21455](#), [pr#18004](#), Yao Zongyou)
- rgw: rgw file write error ([issue#21455](#), [pr#18433](#), Yao Zongyou)
- rgw: s3:GetBucketCORS/s3:PutBucketCORS policy fails with 403 ([issue#21578](#), [pr#18444](#), Adam C. Emerson)
- rgw: s3:GetBucketLocation bucket policy fails with 403 ([issue#21582](#), [pr#18443](#), Adam C. Emerson)
- rgw: s3:GetBucketWebsite/PutBucketWebsite fails with 403 ([issue#21597](#), [pr#18445](#), Adam C. Emerson)
- rgw: setxattr call leads to different mtimes for bucket index and object ([issue#21200](#), [pr#17856](#), Abhishek Lekshmanan)
- rgw: stop/join TokenCache revoke thread only if started ([issue#21666](#), [pr#18138](#), Karol Mroz)
- rgw: string_view instance points to expired memory in PrefixableSignatureHelper ([issue#21085](#), [pr#17858](#), Radoslaw Zarzynski)
- rgw: user creation can overwrite existing user even if different uid is given ([issue#21685](#), [pr#18436](#), Casey Bodley)
- rgw: We can't get torrents if objects are encrypted using SSE-C ([issue#21720](#), [pr#18431](#), Zhang Shaowen)
- rgw: wrong error message is returned when putting container with a name that is too long ([issue#17938](#), [issue#21169](#), [issue#17935](#), [issue#17934](#), [issue#17936](#), [pr#17811](#), Radoslaw Zarzynski)
- rgw: zone compression type is not validated ([issue#21775](#), [pr#18434](#), Casey Bodley)
- tools: ceph-disk create deprecation warnings ([issue#22154](#), [pr#18989](#), Alfredo Deza)
- tools: ceph-disk: fix '-runtime' omission for ceph-osd service ([issue#21498](#), [pr#17914](#), Carl Xiong)
- tools: ceph-disk flake8 test fails on very old, and very new, versions of flake8 ([issue#22207](#), [pr#19152](#), Nathan Cutler)
- tools: ceph-disk: retry on OSSError ([issue#21728](#), [pr#18189](#), Kefu Chai)
- tools: ceph-disk: unlocks dmcryptoed partitions when activating them ([issue#20488](#), [pr#18625](#), Kefu Chai, Felix Winterhalter)
- tools: ceph-kvstore-tool does not call bluestore's umount when exit ([issue#21625](#), [pr#18751](#), Chang Liu)
- tools: ceph_monstore_tool: rebuild initial mgrmap also ([issue#22266](#), [pr#19240](#), Kefu Chai)
- tools: ceph-objectstore-tool and ceph-bluestore-tool: backports from master ([issue#21272](#), [pr#17896](#), Sage Weil, David Zafman)
- tools: ceph_volume_client: add get, put, and delete object interfaces ([issue#21601](#), [pr#18037](#), Ramana Raja)
- tools: cli/crushtools/build.t sometimes fails in jenkins' make check run ([issue#21758](#), [pr#18398](#), Kefu Chai, Sage Weil)

V12.2.1 LUMINOUS

This is the first bugfix release of Luminous v12.2.x long term stable release series. It contains a range of bug fixes and a few features across CephFS, RBD & RGW. We recommend all the users of 12.2.x series update.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- Dynamic resharding is now enabled by default for RGW, RGW will now automatically reshards there bucket index once the index grows beyond `rgw_max_objs_per_shard`
- Limiting MDS cache via a memory limit is now supported using the new `mds_cache_memory_limit` config option (1GB by default). A cache reservation can also be specified using `mds_cache_reservation` as a percentage of the limit (5% by

default). Limits by inode count are still supported using mds_cache_size. Setting mds_cache_size to 0 (the default) disables the inode limit.

- The maximum number of PGs per OSD before the monitor issues a warning has been reduced from 300 to 200 PGs. 200 is still twice the generally recommended target of 100 PGs per OSD. This limit can be adjusted via the mon_max_pg_per_osd option on the monitors. The older mon_pg_warn_max_per_osd option has been removed.
- Creating pools or adjusting pg_num will now fail if the change would make the number of PGs per OSD exceed the configured mon_max_pg_per_osd limit. The option can be adjusted if it is really necessary to create a pool with more PGs.
- There was a bug in the PG mapping behavior of the new upmap feature. If you made use of this feature (e.g., via the ceph osd pg-upmap-items command), we recommend that all mappings be removed (via the ceph osd rm-pg-upmap-items command) before upgrading to this point release.
- A stall in BlueStore IO submission that was affecting many users has been resolved.

OTHER NOTABLE CHANGES

- bluestore: asyn cdeferred_try_submit deadlock ([issue#21207](#), [pr#17494](#), Sage Weil)
- bluestore: fix deferred write deadlock, aio short return handling ([issue#21171](#), [pr#17601](#), Sage Weil)
- bluestore: osd crash when change option bluestore_csum_type from none to CRC32 ([issue#21175](#), [pr#17497](#), xie xingguo)
- bluestore: os/bluestore/BlueFS.cc: 1255: FAILED assert(!log_file->fnode.extents.empty()) ([issue#21250](#), [pr#17562](#), Sage Weil)
- build/ops: ceph-fuse RPM should require fusermount ([issue#21057](#), [pr#17470](#), Ken Dreyer)
- build/ops: RHEL 7.3 Selinux denials at OSD start ([issue#19200](#), [pr#17468](#), Boris Ranto)
- build/ops: rocksdb,cmake: build portable binaries ([issue#20529](#), [pr#17745](#), Kefu Chai)
- cephfs: client/mds has wrong check to clear S_ISGID on chown ([issue#21004](#), [pr#17471](#), Patrick Donnelly)
- cephfs: get_quota_root sends lookupname op for every buffered write ([issue#20945](#), [pr#17473](#), Dan van der Ster)
- cephfs: MDCache::try_subtree_merge() may print N^2 lines of debug message ([issue#21221](#), [pr#17712](#), Patrick Donnelly)
- cephfs: MDS rank add/remove log messages say wrong number of ranks ([issue#21421](#), [pr#17887](#), John Spray)
- cephfs: MDS: standby-replay mds should avoid initiating subtree export ([issue#21378](#), [issue#21222](#), [pr#17714](#), "Yan, Zheng", Jianyu Li)
- cephfs: the standbys are not updated via ceph tell mds.* command ([issue#21230](#), [pr#17565](#), Kefu Chai)
- common: adding line break at end of some cli results ([issue#21019](#), [pr#17467](#), songweibin)
- core: [cls] metadata_list API function does not honor max_return parameter ([issue#21247](#), [pr#17558](#), Jason Dillaman)
- core: incorrect erasure-code space in command ceph df ([issue#21243](#), [pr#17724](#), liuchang0812)
- core: interval_set: optimize intersect_of insert operations ([issue#21229](#), [pr#17487](#), Zac Medico)
- core: osd crush rule rename not idempotent ([issue#21162](#), [pr#17481](#), xie xingguo)
- core: osd/PGLog: write only changed dup entries ([issue#21026](#), [pr#17378](#), Josh Durgin)
- doc: doc/rbd: iSCSI Gateway Documentation ([issue#20437](#), [pr#17381](#), Aron Gunn, Jason Dillaman)
- mds: fix 'dirfrag end' check in Server::handle_client_readdir ([issue#21070](#), [pr#17686](#), "Yan, Zheng")
- mds: support limiting cache by memory ([issue#20594](#), [pr#17711](#), "Yan, Zheng", Patrick Donnelly)
- mgr: 500 error when attempting to view filesystem data ([issue#20692](#), [pr#17477](#), John Spray)
- mgr: ceph mgr versions shows active mgr as Unknown ([issue#21260](#), [pr#17635](#), John Spray)
- mgr: Crash in MonCommandCompletion ([issue#21157](#), [pr#17483](#), John Spray)
- mon: mon/OSDMonitor: deleting pool while pgs are being created leads to assert(p != pools.end) in update_creating_pgs() ([issue#21309](#), [pr#17634](#), Joao Eduardo Luis)
- mon: OSDMonitor: osd pool application get support ([issue#20976](#), [pr#17472](#), xie xingguo)
- mon: rate limit on health check update logging ([issue#20888](#), [pr#17500](#), John Spray)
- osd: build_initial_pg_history doesn't update up/acting/etc ([issue#21203](#), [pr#17496](#), w11979, Sage Weil)
- osd: osd/PG: discard msgs from down peers ([issue#19605](#), [pr#17501](#), Kefu Chai)
- osd/PrimaryLogPG: request osdmap update in the right block ([issue#21428](#), [pr#17829](#), Josh Durgin)
- osd: PrimaryLogPG: sparse read won't trigger repair correctly ([issue#21123](#), [pr#17475](#), xie xingguo)
- osd: request new map from PG when needed ([issue#21428](#), [pr#17796](#), Josh Durgin)
- osd: Revert "osd/OSDMap: allow bidirectional swap of pg-upmap-items" ([issue#21410](#), [pr#17812](#), Sage Weil)
- osd: subscribe to new osdmap while waiting_for_healthy ([issue#21121](#), [pr#17498](#), Sage Weil)
- osd: update info only if new_interval ([issue#21203](#), [pr#17622](#), Kefu Chai)
- pybind: dashboard usage graph getting bigger and bigger ([issue#20746](#), [pr#17486](#), Yixing Yan)
- rbd: image-meta list does not return all entries ([issue#21179](#), [pr#17561](#), Jason Dillaman)
- rbd: some generic options can not be passed by rbd-nbd ([issue#20426](#), [pr#17557](#), Pan Liu)
- rbd: switch to new config option getter methods ([issue#20737](#), [pr#17464](#), Jason Dillaman)
- rbd: TestMirroringWatcher.ModeUpdated: periodic failure due to injected message failures ([issue#21029](#), [pr#17465](#), Jason Dillaman)
- rgw: bucket index sporadically reshards to 65521 shards ([issue#20934](#), [pr#17476](#), Aleksei Gutikov)
- rgw: bytes_send and bytes_recv in the msg of usage show returning is 0 in master branch ([issue#19870](#), [pr#17444](#), Marcus Watts)

- rgw: data encryption sometimes fails to follow AWS settings ([issue#21349](#), [pr#17642](#), hechuang)
- rgw: memory leak in MetadataHandlers ([issue#21214](#), [pr#17570](#), Luo Kexue, Jos Collin)
- rgw: multisite: objects encrypted with SSE-KMS are stored unencrypted in target zone ([issue#20668](#), [issue#20671](#), [pr#17446](#), Casey Bodley)
- rgw: need to stream metadata full sync init ([issue#18079](#), [pr#17448](#), Yehuda Sadeh)
- rgw: object copied from remote src acl permission become full-control issue ([issue#20658](#), [pr#17478](#), Enming Zhang)
- rgw: put lifecycle configuration fails if Prefix is not set ([issue#19587](#), [issue#20872](#), [pr#17479](#), Shasha Lu, Abhishek Lekshmanan)
- rgw: rgw_file: incorrect lane lock behavior in evict_block() ([issue#21141](#), [pr#17485](#), Matt Benjamin)
- rgw: send data-log list infinitely ([issue#20951](#), [pr#17445](#), fang.yuxiang)
- rgw: shadow objects are sometimes not removed ([issue#20234](#), [pr#17555](#), Yehuda Sadeh)
- rgw: usage of -inconsistent-index should require user confirmation and print a warning ([issue#20777](#), [pr#17488](#), Orit Wasserman)
- tools: [cli] rename of non-existent image results in seg fault ([issue#21248](#), [pr#17556](#), Jason Dillaman)

V12.2.0 LUMINOUS

This is the first release of Luminous v12.2.x long term stable release series. There have been major changes since Kraken (v11.2.z) and Jewel (v10.2.z), and the upgrade process is non-trivial. Please read these release notes carefully.

MAJOR CHANGES FROM KRAKEN

- *General:*
 - Ceph now has a simple, [built-in web-based dashboard](#) for monitoring cluster status.
- *RADOS:*
 - *BlueStore:*
 - The new *BlueStore* backend for *ceph-osd* is now stable and the new default for newly created OSDs. *BlueStore* manages data stored by each OSD by directly managing the physical HDDs or SSDs without the use of an intervening file system like XFS. This provides greater performance and features. See [Storage Devices](#) and [BlueStore Config Reference](#).
 - *BlueStore* supports [full data and metadata checksums](#) of all data stored by Ceph.
 - *BlueStore* supports [inline compression](#) using zlib, snappy, or LZ4. (Ceph also supports zstd for [RGW compression](#) but zstd is not recommended for *BlueStore* for performance reasons.)
 - *Erasure coded pools* now have [full support for overwrites](#), allowing them to be used with RBD and CephFS.
 - *ceph-mgr*:
 - There is a new daemon, *ceph-mgr*, which is a required part of any Ceph deployment. Although IO can continue when *ceph-mgr* is down, metrics will not refresh and some metrics-related calls (e.g., *ceph df*) may block. We recommend deploying several instances of *ceph-mgr* for reliability. See the notes on [Upgrading](#) below.
 - The *ceph-mgr* daemon includes a [REST-based management API](#). The API is still experimental and somewhat limited but will form the basis for API-based management of Ceph going forward.
 - *ceph-mgr* also includes a [Prometheus exporter](#) plugin, which can provide Ceph perfcounters to Prometheus.
 - *ceph-mgr* now has a [Zabbix](#) plugin. Using *zabbix_sender* it sends trapper events to a Zabbix server containing high-level information of the Ceph cluster. This makes it easy to monitor a Ceph cluster's status and send out notifications in case of a malfunction.
 - The overall *scalability* of the cluster has improved. We have successfully tested clusters with up to 10,000 OSDs.
 - Each OSD can now have a [device class](#) associated with it (e.g., *hdd* or *ssd*), allowing CRUSH rules to trivially map data to a subset of devices in the system. Manually writing CRUSH rules or manual editing of the CRUSH is normally not required.
 - There is a new [upmap](#) exception mechanism that allows individual PGs to be moved around to achieve a *perfect distribution* (this requires luminous clients).
 - Each OSD now adjusts its default configuration based on whether the backing device is an HDD or SSD. Manual tuning generally not required.
 - The prototype [mClock QoS queueing algorithm](#) is now available.
 - There is now a *backoff* mechanism that prevents OSDs from being overloaded by requests to objects or PGs that are not currently able to process IO.
 - There is a simplified [OSD replacement process](#) that is more robust.
 - You can query the supported features and (apparent) releases of all connected daemons and clients with [ceph features](#).
 - You can configure the oldest Ceph client version you wish to allow to connect to the cluster via *ceph osd set-require-min-compat-client* and Ceph will prevent you from enabling features that will break compatibility with those clients.
 - Several *sleep* settings, include *osd_recovery_sleep*, *osd_snap_trim_sleep*, and *osd_scrub_sleep* have been reimplemented to work efficiently. (These are used in some cases to work around issues throttling background

work.)

- Pools are now expected to be associated with the application using them. Upon completing the upgrade to Luminous, the cluster will attempt to associate existing pools to known applications (i.e. CephFS, RBD, and RGW). In-use pools that are not associated to an application will generate a health warning. Any unassociated pools can be manually associated using the new `ceph osd pool application enable` command. For more details see [associate pool to application](#) in the documentation.

- **RGW:**

- RGW *metadata search* backed by ElasticSearch now supports end user requests service via RGW itself, and also supports custom metadata fields. A query language a set of RESTful APIs were created for users to be able to search objects by their metadata. New APIs that allow control of custom metadata fields were also added.
- RGW now supports *dynamic bucket index sharding*. This has to be enabled via the `rgw dynamic resharding` configurable. As the number of objects in a bucket grows, RGW will automatically reshuffle the bucket index in response. No user intervention or bucket size capacity planning is required.
- RGW introduces *server side encryption* of uploaded objects with three options for the management of encryption keys: automatic encryption (only recommended for test setups), customer provided keys similar to Amazon SSE-C specification, and through the use of an external key management service (Openstack Barbican) similar to Amazon SSE-KMS specification. [Encryption](#)
- RGW now has preliminary AWS-like bucket policy API support. For now, policy is a means to express a range of new authorization concepts. In the future it will be the foundation for additional auth capabilities such as STS and group policy. [Bucket Policies](#)
- RGW has consolidated the several metadata index pools via the use of rados namespaces. [Pools](#)
- S3 Object Tagging API has been added; while APIs are supported for GET/PUT/DELETE object tags and in PUT object API, there is no support for tags on Policies & Lifecycle yet
- RGW multisite now supports for enabling or disabling sync at a bucket level.

- **RBD:**

- RBD now has full, stable support for *erasure coded pools* via the new `--data-pool` option to `rbd create`.
- RBD mirroring's `rbd-mirror` daemon is now highly available. We recommend deploying several instances of `rbd-mirror` for reliability.
- RBD mirroring's `rbd-mirror` daemon should utilize unique Ceph user IDs per instance to support the new mirroring dashboard.
- The default 'rbd' pool is no longer created automatically during cluster creation. Additionally, the name of the default pool used by the `rbd` CLI when no pool is specified can be overridden via a new `rbd default pool = <pool name>` configuration option.
- Initial support for deferred image deletion via new `rbd trash` CLI commands. Images, even ones actively in-use by clones, can be moved to the trash and deleted at a later time.
- New pool-level `rbd mirror pool promote` and `rbd mirror pool demote` commands to batch promote/demote all mirrored images within a pool.
- Mirroring now optionally supports a configurable replication delay via the `rbd mirroring replay delay = <seconds>` configuration option.
- Improved discard handling when the object map feature is enabled.
- `rbd` CLI import and copy commands now detect sparse and preserve sparse regions.
- Images and Snapshots will now include a creation timestamp.
- Specifying user authorization capabilities for RBD clients has been simplified. The general syntax for using RBD capability profiles is “`mon ‘profile rbd’ osd ‘profile rbd[-read-only][pool={pool-name}{, …}]’`”. For more details see [User Management](#) in the documentation.

- **CephFS:**

- *Multiple active MDS daemons* is now considered stable. The number of active MDS servers may be adjusted up or down on an active CephFS file system.
- CephFS *directory fragmentation* is now stable and enabled by default on new filesystems. To enable it on existing filesystems use “`ceph fs set <fs_name> allow_dirfrags`”. Large or very busy directories are sharded and (potentially) distributed across multiple MDS daemons automatically.
- Directory subtrees can be explicitly pinned to specific MDS daemons in cases where the automatic load balancing is not desired or effective.
- Client keys can now be created using the new `ceph fs authorize` command to create keys with access to the given CephFS file system and all of its data pools.
- When running ‘`df`’ on a CephFS filesystem comprising exactly one data pool, the result now reflects the file storage space used and available in that data pool (fuse client only).

- **Miscellaneous:**

- Release packages are now being built for *Debian Stretch*. Note that QA is limited to CentOS and Ubuntu (xenial and trusty). The distributions we build for now include:
 - CentOS 7 (x86_64 and aarch64)
 - Debian 8 Jessie (x86_64)
 - Debian 9 Stretch (x86_64)
 - Ubuntu 16.04 Xenial (x86_64 and aarch64)
 - Ubuntu 14.04 Trusty (x86_64)

- A first release of Ceph for FreeBSD is available which contains a full set of features, other than Bluestore. It will run everything needed to build a storage cluster. For clients, all access methods are available, albeit CephFS is only accessible through a Fuse implementation. RBD images can be mounted on FreeBSD systems through rbd-ggate. Ceph versions are released through the regular FreeBSD ports and packages system. The most current version is available as: net/ceph-devel. Once Luminous goes into official release, this version will be available as net/ceph. Future development releases will be available via net/ceph-devel. More details about this port are in:

[README.FreeBSD](#)

- *CLI changes:*

- The ceph -s or ceph status command has a fresh look.
- ceph mgr metadata will dump metadata associated with each mgr daemon.
- ceph versions or ceph {osd,mds,mon,mgr} versions summarize versions of running daemons.
- ceph {osd,mds,mon,mgr} count-metadata <property> similarly tabulates any other daemon metadata visible via the ceph {osd,mds,mon,mgr} metadata commands.
- ceph features summarizes features and releases of connected clients and daemons.
- ceph osd require-osd-release <release> replaces the old require_RELEASE_osds flags.
- ceph osd pg-upmap, ceph osd rm-pg-upmap, ceph osd pg-upmap-items, ceph osd rm-pg-upmap-items can explicitly manage upmap items (see [Using the pg-upmap](#)).
- ceph osd getcrushmap returns a crush map version number on stderr, and ceph osd setcrushmap [version] will only inject an updated crush map if the version matches. This allows crush maps to be updated offline and then reinjected into the cluster without fear of clobbering racing changes (e.g., by newly added osds or changes by other administrators).
- ceph osd create has been replaced by ceph osd new. This should be hidden from most users by user-facing tools like *ceph-disk*.
- ceph osd destroy will mark an OSD destroyed and remove its cephx and lockbox keys. However, the OSD id and CRUSH map entry will remain in place, allowing the id to be reused by a replacement device with minimal data rebalancing.
- ceph osd purge will remove all traces of an OSD from the cluster, including its cephx encryption keys, dm-crypt lockbox keys, OSD id, and crush map entry.
- ceph osd ls-tree <name> will output a list of OSD ids under the given CRUSH name (like a host or rack name). This is useful for applying changes to entire subtrees. For example, ceph osd down `ceph osd ls-tree rack1`.
- ceph osd {add,rm}-{noout,noin,nodown,noup} allow the *noout*, *noin*, *nodown*, and *noup* flags to be applied to specific OSDs.
- ceph osd safe-to-destroy <osd(s)> will report whether it is safe to remove or destroy OSD(s) without reducing data durability or redundancy.
- ceph osd ok-to-stop <osd(s)> will report whether it is okay to stop OSD(s) without immediately compromising availability (i.e., all PGs should remain active but may be degraded).
- ceph log last [n] will output the last n lines of the cluster log.
- ceph mgr dump will dump the MgrMap, including the currently active ceph-mgr daemon and any standbys.
- ceph mgr module ls will list active ceph-mgr modules.
- ceph mgr module {enable,disable} <name> will enable or disable the named mgr module. The module must be present in the configured *mgr_module_path* on the host(s) where *ceph-mgr* is running.
- ceph osd crush ls <node> will list items (OSDs or other CRUSH nodes) directly beneath a given CRUSH node.
- ceph osd crush swap-bucket <src> <dest> will swap the contents of two CRUSH buckets in the hierarchy while preserving the buckets' ids. This allows an entire subtree of devices to be replaced (e.g., to replace an entire host of FileStore OSDs with newly-imaged BlueStore OSDs) without disrupting the distribution of data across neighboring devices.
- ceph osd set-require-min-compat-client <release> configures the oldest client release the cluster is required to support. Other changes, like CRUSH tunables, will fail with an error if they would violate this setting. Changing this setting also fails if clients older than the specified release are currently connected to the cluster.
- ceph config-key dump dumps config-key entries and their contents. (The existing ceph config-key list only dumps the key names, not the values.)
- ceph config-key list is deprecated in favor of ceph config-key ls.
- ceph config-key put is deprecated in favor of ceph config-key set.
- ceph auth list is deprecated in favor of ceph auth ls.
- ceph osd crush rule list is deprecated in favor of ceph osd crush rule ls.
- ceph osd set-{full,nearfull,backfillfull}-ratio sets the cluster-wide ratio for various full thresholds (when the cluster refuses IO, when the cluster warns about being close to full, when an OSD will defer rebalancing a PG to itself, respectively).
- ceph osd reweightn will specify the *reweight* values for multiple OSDs in a single command. This is equivalent to a series of ceph osd reweight commands.
- ceph osd crush {set,rm}-device-class manage the new CRUSH device class feature. Note that manually creating or deleting a device class name is generally not necessary as it will be smart enough to be self-

- managed. `ceph osd crush class ls` and `ceph osd crush class ls-osd` will output all existing device classes and a list of OSD ids under the given device class respectively.
- `ceph osd crush rule create-replicated` replaces the old `ceph osd crush rule create-simple` command to create a CRUSH rule for a replicated pool. Notably it takes a *class* argument for the *device class* the rule should target (e.g., *ssd* or *hdd*).
- `ceph mon feature ls` will list monitor features recorded in the MonMap. `ceph mon feature set` will set an optional feature (none of these exist yet).
- `ceph tell <daemon> help` will now return a usage summary.
- `ceph fs authorize` creates a new client key with caps automatically set to access the given CephFS file system.
- The `ceph health` structured output (JSON or XML) no longer contains ‘timechecks’ section describing the time sync status. This information is now available via the ‘`ceph time-sync-status`’ command.
- Certain extra fields in the `ceph health` structured output that used to appear if the mons were low on disk space (which duplicated the information in the normal health warning messages) are now gone.
- The `ceph -w` output no longer contains audit log entries by default. Add a `--watch-channel=audit` or `--watch-channel=*` to see them.
- New “`ceph -w`” behavior - the “`ceph -w`” output no longer contains I/O rates, available space, pg info, etc. because these are no longer logged to the central log (which is what `ceph -w` shows). The same information can be obtained by running `ceph pg stat`; alternatively, I/O rates per pool can be determined using `ceph osd pool stats`. Although these commands do not self-update like `ceph -w` did, they do have the ability to return formatted output by providing a `--format=<format>` option.
- Added new commands `pg force-recovery` and `pg-force-backfill`. Use them to boost recovery or backfill priority of specified pgs, so they’re recovered/backfilled before any other. Note that these commands don’t interrupt ongoing recovery/backfill, but merely queue specified pgs before others so they’re recovered/backfilled as soon as possible. New commands `pg cancel-force-recovery` and `pg cancel-force-backfill` restore default recovery/backfill priority of previously forced pgs.

MAJOR CHANGES FROM JEWEL

- **RADOS:**
 - We now default to the `AsyncMessenger` (`ms type = async`) instead of the legacy `SimpleMessenger`. The most noticeable difference is that we now use a fixed sized thread pool for network connections (instead of two threads per socket with `SimpleMessenger`).
 - Some OSD failures are now detected almost immediately, whereas previously the heartbeat timeout (which defaults to 20 seconds) had to expire. This prevents IO from blocking for an extended period for failures where the host remains up but the `ceph-osd` process is no longer running.
 - The size of encoded OSDMaps has been reduced.
 - The OSDs now quiesce scrubbing when recovery or rebalancing is in progress.
- **RGW:**
 - RGW now supports the S3 multipart object copy-part API.
 - It is possible now to reshards an existing bucket offline. Offline bucket reshards currently requires that all IO (especially writes) to the specific bucket is quiesced. (For automatic online resharding, see the new feature in Luminous above.)
 - RGW now supports data compression for objects.
 - Civetweb version has been upgraded to 1.8
 - The Swift static website API is now supported (S3 support has been added previously).
 - S3 bucket lifecycle API has been added. Note that currently it only supports object expiration.
 - Support for custom search filters has been added to the LDAP auth implementation.
 - Support for NFS version 3 has been added to the RGW NFS gateway.
 - A Python binding has been created for `librgw`.
- **RBD:**
 - The `rbd-mirror` daemon now supports replicating dynamic image feature updates and image metadata key/value pairs from the primary image to the non-primary image.
 - The number of image snapshots can be optionally restricted to a configurable maximum.
 - The `rbd` Python API now supports asynchronous IO operations.
- **CephFS:**
 - `libcephfs` function definitions have been changed to enable proper uid/gid control. The library version has been increased to reflect the interface change.
 - Standby replay MDS daemons now consume less memory on workloads doing deletions.
 - Scrub now repairs backtrace, and populates `damage ls` with discovered errors.
 - A new `pg_files` subcommand to `cephfs-data-scan` can identify files affected by a damaged or lost RADOS PG.
 - The false-positive “failing to respond to cache pressure” warnings have been fixed.

UPGRADE FROM JEWEL OR KRAKEN

1. Ensure that the sortbitwise flag is enabled:

```
# ceph osd set sortbitwise
```

2. Make sure your cluster is stable and healthy (no down or recovering OSDs). (Optional, but recommended.)
3. Do not create any new erasure-code pools while upgrading the monitors.
4. You can monitor the progress of your upgrade at each stage with the ceph versions command, which will tell you what ceph version is running for each type of daemon.
5. Set the noout flag for the duration of the upgrade. (Optional but recommended.):

```
# ceph osd set noout
```

6. Verify that all RBD client users have sufficient caps to blacklist other client users. RBD client users with only "allow r" monitor caps should be updated as follows:

```
# ceph auth caps client.<ID> mon 'allow r, allow command "osd blacklist"' osd '<existing
```

7. Upgrade monitors by installing the new packages and restarting the monitor daemons. Note that, unlike prior releases, the ceph-mon daemons *must* be upgraded first:

```
# systemctl restart ceph-mon.target
```

Verify the monitor upgrade is complete once all monitors are up by looking for the *luminous* feature string in the mon map. For example:

```
# ceph mon feature ls
```

should include *luminous* under persistent features:

```
on current monmap (epoch NNN)
persistent: [kraken,luminous]
required: [kraken,luminous]
```

8. Add or restart ceph-mgr daemons. If you are upgrading from kraken, upgrade packages and restart ceph-mgr daemons with:

```
# systemctl restart ceph-mgr.target
```

If you are upgrading from kraken, you may already have ceph-mgr daemons deployed. If not, or if you are upgrading from jewel, you can deploy new daemons with tools like ceph-deploy or ceph-ansible. For example:

```
# ceph-deploy mgr create HOST
```

Verify the ceph-mgr daemons are running by checking ceph -s:

```
# ceph -s
...
services:
mon: 3 daemons, quorum foo,bar,baz
mgr: foo(active), standbys: bar, baz
...
```

9. Upgrade all OSDs by installing the new packages and restarting the ceph-osd daemons on all hosts:

```
# systemctl restart ceph-osd.target
```

You can monitor the progress of the OSD upgrades with the new ceph versions or ceph osd versions command:

```
# ceph osd versions
{
    "ceph version 12.2.0 (...) luminous (stable)": 12,
    "ceph version 10.2.6 (...)": 3,
}
```

10. Upgrade all CephFS daemons by upgrading packages and restarting daemons on all hosts:

```
# systemctl restart ceph-mds.target
```

11. Upgrade all radosgw daemons by upgrading packages and restarting daemons on all hosts:

```
# systemctl restart radosgw.target
```

12. Complete the upgrade by disallowing pre-luminous OSDs and enabling all new Luminous-only functionality:

```
# ceph osd require-osd-release luminous
```

If you set noout at the beginning, be sure to clear it with:

```
# ceph osd unset noout
```

13. Verify the cluster is healthy with ceph health.

UPGRADING FROM PRE-JEWEL RELEASES (LIKE HAMMER)

You *must* first upgrade to Jewel (10.2.z) before attempting an upgrade to Luminous.

UPGRADE COMPATIBILITY NOTES, JEWEL TO KRAKEN

These changes occurred between the Jewel and Kraken releases and will affect upgrades from Jewel to Luminous.

- The osd crush location config option is no longer supported. Please update your ceph.conf to use the crush location option instead. Be sure to update your config file to avoid any movement of OSDs from your customized location back to the default one.
- The OSDs now avoid starting new scrubs while recovery is in progress. To revert to the old behavior (and do not let recovery activity affect the scrub scheduling) you can set the following option:

```
osd scrub during recovery = true
```

- The list of monitor hosts/addresses for building the monmap can now be obtained from DNS SRV records. The service name used in when querying the DNS is defined in the “mon_dns_srv_name” config option, which defaults to “ceph-mon”.
- The ‘osd class load list’ config option is a list of object class names that the OSD is permitted to load (or '*' for all classes). By default it contains all existing in-tree classes for backwards compatibility.
- The ‘osd class default list’ config option is a list of object class names (or '*' for all classes) that clients may invoke having only the ‘*’, ‘x’, ‘class-read’, or ‘class-write’ capabilities. By default it contains all existing in-tree classes for backwards

compatibility. Invoking classes not listed in ‘osd class default list’ requires a capability naming the class (e.g. ‘allow class foo’).

- The ‘rgw rest getusage op compat’ config option allows you to dump (or not dump) the description of user stats in the S3 GetUsage API. This option defaults to false. If the value is true, the reponse data for GetUsage looks like:

```
"stats": {  
    "TotalBytes": 516,  
    "TotalBytesRounded": 1024,  
    "TotalEntries": 1  
}
```

If the value is false, the reponse for GetUsage looks as it did before:

```
{  
    516,  
    1024,  
    1  
}
```

- The ‘osd out ...’ and ‘osd in ...’ commands now preserve the OSD weight. That is, after marking an OSD out and then in, the weight will be the same as before (instead of being reset to 1.0). Previously the mons would only preserve the weight if the mon automatically marked an OSD out and then in, but not when an admin did so explicitly.
- The ‘ceph osd perf’ command will display ‘commit_latency(ms)’ and ‘apply_latency(ms)’. Previously, the names of these two columns were ‘fs_commit_latency(ms)’ and ‘fs_apply_latency(ms)’. We removed the prefix ‘fs_’ because the values are not filestore-specific.
- Monitors will no longer allow pools to be removed by default. The setting mon_allow_pool_delete has to be set to true (defaults to false) before they allow pools to be removed. This is a additional safeguard against pools being removed by accident.
- If you have manually specified the monitor user rocksdb via the mon keyvaluedb = rocksdb option, you will need to manually add a file to the mon data directory to preserve this option:

```
echo rocksdb > /var/lib/ceph/mon/ceph-`hostname`/kv_backend
```

New monitors will now use rocksdb by default, but if that file is not present, existing monitors will use leveldb. The mon keyvaluedb option now only affects the backend chosen when a monitor is created.

- The ‘osd crush initial weight’ option allows you to specify a CRUSH weight for a newly added OSD. Previously a value of 0 (the default) meant that we should use the size of the OSD’s store to weight the new OSD. Now, a value of 0 means it should have a weight of 0, and a negative value (the new default) means we should automatically weight the OSD based on its size. If your configuration file explicitly specifies a value of 0 for this option you will need to change it to a negative value (e.g., -1) to preserve the current behavior.
- The static libraries are no longer included by the debian development packages (lib*-dev) as it is not required per debian packaging policy. The shared (.so) versions are packaged as before.
- The libtool pseudo-libraries (.la files) are no longer included by the debian development packages (lib*-dev) as they are not required per <https://wiki.debian.org/ReleaseGoals/LAFileRemoval> and <https://www.debian.org/doc/manuals/maint-guide/advanced.en.html>.
- The jerasure and shec plugins can now detect SIMD instruction at runtime and no longer need to be explicitly configured for different processors. The following plugins are now deprecated: jerasure_generic, jerasure_sse3, jerasure_sse4, jerasure_neon, shec_generic, shec_sse3, shec_sse4, and shec_neon. If you use any of these plugins directly you will see a warning in the mon log file. Please switch to using just ‘jerasure’ or ‘shec’.
- The librados omap get_keys and get_vals operations include a start key and a limit on the number of keys to return. The OSD now imposes a configurable limit on the number of keys and number of total bytes it will respond with, which means that a librados user might get fewer keys than they asked for. This is necessary to prevent careless users from requesting an unreasonable amount of data from the cluster in a single operation. The new limits are configured with osd_max omap_entries_per_request, defaulting to 131,072, and osd_max omap_bytes_per_request, defaulting to 4MB.
- Calculation of recovery priorities has been updated. This could lead to unintuitive recovery prioritization during cluster

upgrade. In case of such recovery, OSDs in old version would operate on different priority ranges than new ones. Once upgraded, cluster will operate on consistent values.

UPGRADE COMPATIBILITY NOTES, KRAKEN TO LUMINOUS

- The configuration option `osd pool erasure code stripe width` has been replaced by `osd pool erasure code stripe unit`, and given the ability to be overridden by the erasure code profile setting `stripe_unit`. For more details see [Erasure code profiles](#).
- rbd and cephfs can use erasure coding with bluestore. This may be enabled by setting `allow_ec_overwrites` to true for a pool. Since this relies on bluestore's checksumming to do deep scrubbing, enabling this on a pool stored on filestore is not allowed.
- The `rados df` JSON output now prints numeric values as numbers instead of strings.
- The `mon_osd_max_op_age` option has been renamed to `mon_osd_warn_op_age` (default: 32 seconds), to indicate we generate a warning at this age. There is also a new `mon_osd_err_op_age_ratio` that is expressed as a multiple of `mon_osd_warn_op_age` (default: 128, for roughly 60 minutes) to control when an error is generated.
- The default maximum size for a single RADOS object has been reduced from 100GB to 128MB. The 100GB limit was completely impractical in practice while the 128MB limit is a bit high but not unreasonable. If you have an application written directly to librados that is using objects larger than 128MB you may need to adjust `osd_max_object_size`.
- The semantics of the `rados ls` and librados object listing operations have always been a bit confusing in that “whiteout” objects (which logically don’t exist and will return ENOENT if you try to access them) are included in the results. Previously whiteouts only occurred in cache tier pools. In luminous, logically deleted but snapshotted objects now result in a whiteout object, and as a result they will appear in `rados ls` results, even though trying to read such an object will result in ENOENT. The `rados listsnaps` operation can be used in such a case to enumerate which snapshots are present. This may seem a bit strange, but is less strange than having a deleted-but-snapshotted object not appear at all and be completely hidden from librados’s ability to enumerate objects. Future versions of Ceph will likely include an alternative object enumeration interface that makes it more natural and efficient to enumerate all objects along with their snapshot and clone metadata.
- The deprecated `crush_ruleset` property has finally been removed; please use `crush_rule` instead for the `osd pool get ...` and `osd pool set ...` commands.
- The `osd pool default crush replicated ruleset` option has been removed and replaced by the `osd pool default crush rule` option. By default it is -1, which means the mon will pick the first type replicated rule in the CRUSH map for replicated pools. Erasure coded pools have rules that are automatically created for them if they are not specified at pool creation time.
- We no longer test the FileStore ceph-osd backend in combination with btrfs. We recommend against using btrfs. If you are using btrfs-based OSDs and want to upgrade to luminous you will need to add the following to your `ceph.conf`:

```
enable experimental unrecoverable data corrupting features = btrfs
```

The code is mature and unlikely to change, but we are only continuing to test the Jewel stable branch against btrfs. We recommend moving these OSDs to FileStore with XFS or BlueStore.

- The `ruleset-*` properties for the erasure code profiles have been renamed to `crush-*` to move away from the obsolete ‘ruleset’ term and to be more clear about their purpose. There is also a new optional `crush-device-class` property to specify a CRUSH device class to use for the erasure coded pool. Existing erasure code profiles will be converted automatically when upgrade completes (when the `ceph osd require-osd-release luminous` command is run) but any provisioning tools that create erasure coded pools may need to be updated.
- The structure of the XML output for `osd crush tree` has changed slightly to better match the `osd tree` output. The top level structure is now `nodes` instead of `crush_map_roots`.
- When assigning a network to the public network and not to the cluster network the network specification of the public network will be used for the cluster network as well. In older versions this would lead to cluster services being bound to `0.0.0.0:<port>`, thus making the cluster service even more publicly available than the public services. When only specifying a cluster network it will still result in the public services binding to `0.0.0.0`.
- In previous versions, if a client sent an op to the wrong OSD, the OSD would reply with ENXIO. The rationale here is that the client or OSD is clearly buggy and we want to surface the error as clearly as possible. We now only send the ENXIO reply if the `osd_enxio_on_misdirected_op` option is enabled (it’s off by default). This means that a VM using librbd that previously would have gotten an EIO and gone read-only will now see a blocked/hung IO instead.

- The “journaler allow split entries” config setting has been removed.
- The ‘mon_warn_osd_usage_min_max_delta’ config option has been removed and the associated health warning has been disabled because it does not address clusters undergoing recovery or CRUSH rules that do not target all devices in the cluster.
- Added new configuration “public bind addr” to support dynamic environments like Kubernetes. When set the Ceph MON daemon could bind locally to an IP address and advertise a different IP address public_addr on the network.
- The crush choose_args encoding has been changed to make it architecture-independent. If you deployed Luminous dev releases or 12.1.0 rc release and made use of the CRUSH choose_args feature, you need to remove all choose_args mappings from your CRUSH map before starting the upgrade.
- *librados*:
 - Some variants of the omap_get_keys and omap_get_vals librados functions have been deprecated in favor of omap_get_vals2 and omap_get_keys2. The new methods include an output argument indicating whether there are additional keys left to fetch. Previously this had to be inferred from the requested key count vs the number of keys returned, but this breaks with new OSD-side limits on the number of keys or bytes that can be returned by a single omap request. These limits were introduced by kraken but are effectively disabled by default (by setting a very large limit of 1 GB) because users of the newly deprecated interface cannot tell whether they should fetch more keys or not. In the case of the standalone calls in the C++ interface (IoCtx::get_omap_{keys,vals}), librados has been updated to loop on the client side to provide a correct result via multiple calls to the OSD. In the case of the methods used for building multi-operation transactions, however, client-side looping is not practical, and the methods have been deprecated. Note that use of either the IoCtx methods on older librados versions or the deprecated methods on any version of librados will lead to incomplete results if/when the new OSD limits are enabled.
 - The original librados rados_objects_list_open (C) and objects_begin (C++) object listing API, deprecated in Hammer, has finally been removed. Users of this interface must update their software to use either the rados_nobjects_list_open (C) and nobjects_begin (C++) API or the new rados_object_list_begin (C) and object_list_begin (C++) API before updating the client-side librados library to Luminous. Object enumeration (via any API) with the latest librados version and pre-Hammer OSDs is no longer supported. Note that no in-tree Ceph services rely on object enumeration via the deprecated APIs, so only external librados users might be affected. The newest (and recommended) rados_object_list_begin (C) and object_list_begin (C++) API is only usable on clusters with the SORTBITWISE flag enabled (Jewel and later). (Note that this flag is required to be set before upgrading beyond Jewel.)
- *CephFS*:
 - When configuring ceph-fuse mounts in /etc/fstab, a new syntax is available that uses “ceph.<arg>=<val>” in the options column, instead of putting configuration in the device column. The old style syntax still works. See the documentation page “Mount CephFS in your file systems table” for details.
 - CephFS clients without the ‘p’ flag in their authentication capability string will no longer be able to set quotas or any layout fields. This flag previously only restricted modification of the pool and namespace fields in layouts.
 - CephFS will generate a health warning if you have fewer standby daemons than it thinks you wanted. By default this will be 1 if you ever had a standby, and 0 if you did not. You can customize this using ceph fs set <fs> standby_count_wanted <number>. Setting it to zero will effectively disable the health check.
 - The “ceph mds tell ...” command has been removed. It is superseded by “ceph tell mds.<id> ...”
 - The apply mode of cephfs-journal-tool has been removed

OTHER NOTABLE CHANGES

- async: Fixed compilation error when enable -DWITH_DPDK ([pr#12660](#), Pan Liu)
- async: fixed coredump when enable dpdk ([pr#12854](#), Pan Liu)
- async: fixed the error “Cause: Cannot create lock on ‘/var/run/. rte_c... ([pr#12860](#), Pan Liu)
- bluestore: avoid unnecessary copy with coll_t ([pr#12576](#), Yunchuan Wen)
- bluestore: BitAllocator: delete useless codes ([pr#13619](#), Jie Wang)
- bluestore: bluestore/BlueFS: pass string as const ref ([pr#16600](#), dingdangzhang)
- bluestore: bluestore, NVMEDEVICE: Specify the max io completion in conf ([pr#13799](#), optimistzy)
- bluestore: bluestore/NVMEDEVICE: update SPDK to version 17.03 ([pr#14585](#), optimistzy)
- bluestore: bluestore, NVMeDevice: use task’ own lock for (random) read ([pr#14094](#), optimistzy)
- bluestore,build/ops,performance: os/bluestore: enable SSE-assisted CRC32 calculations in RocksDB ([pr#13741](#), Radoslaw Zarzynski)
- bluestore: ceph-disk: add -filestore argument, default to -bluestore ([pr#15437](#), Loic Dachary, Sage Weil)
- bluestore: common/config: set rocksdb_cache_size to OPT_U64 ([pr#13995](#), liuhongtong)
- bluestore: common/options: make “blue{fs,store}_allocator” LEVEL_DEV ([issue#20660](#), [pr#16645](#), Kefu Chai)
- bluestore,common,performance: common/Finisher: Using queue(list<context*>) instead queue(context*) ([pr#8942](#), Jianpeng Ma)

- bluestore,common,performance: isa-l: update isa-l to v2.18 ([pr#15895](#), Ganesh Mahalingam, Tushar Gohad)
- bluestore,core: os/bluestore: fix statfs to not include DB partition in free space ([issue#18599](#), [pr#13140](#), Sage Weil)
- bluestore,core: os/bluestore: fix warning ([pr#15435](#), Sage Weil)
- bluestore,core: os/bluestore: improve mempool usage ([pr#15402](#), Sage Weil)
- bluestore,core: os/bluestore: write “mkfs_done” into disk only if we pass fsck() tests ([pr#15238](#), xie xingguo)
- bluestore,core: osd/OSDMap: should update input param if osd dne ([pr#14863](#), Kefu Chai)
- bluestore,core: os: remove experimental status for BlueStore ([pr#15177](#), Sage Weil)
- bluestore: fixed compilation error when enable spdk ([pr#12672](#), Pan Liu)
- bluestore: include/intarith: templatize ctz/clz/cbits helpers ([pr#14862](#), Kefu Chai)
- bluestore: luminous: os/bluestore: compensate for bad freelistmanager size/blocks metadata ([issue#21089](#), [pr#17273](#), Sage Weil)
- bluestore: NVMEDevice: add the spdk core mask check ([pr#14068](#), optimistyzy)
- bluestore: NVMEDevice: cleanup the logic in data_buf_next_sge ([pr#13056](#), optimistyzy)
- bluestore: NVMeDevice: fix the core id for rte_remote_launch ([pr#13896](#), optimistyzy)
- bluestore: NVMEDevice: fix bug in data_buf_next_sge ([pr#12812](#), optimistyzy)
- bluestore: NVMEDevice: minor error for get slave core ([pr#14012](#), Ziye Yang)
- bluestore: NVMEDevice: optimize sector_size usage ([pr#12780](#), optimistyzy)
- bluestore: NVMEDevice: remove unnecessary dpdk header file ([pr#14650](#), optimistyzy)
- bluestore: NVMEDevice: fix the I/O logic for read ([pr#13971](#), optimistyzy)
- bluestore: os/bluestore: add a debug option to bypass block device writes for bl... ([pr#12464](#), Igor Fedotov)
- bluestore: os/bluestore: Add bluestore pextent vector to mempool ([pr#12946](#), Igor Fedotvo, Igor Fedotov)
- bluestore: os/bluestore: add flush_store_cache cmd ([pr#13428](#), xie xingguo)
- bluestore: os/bluestore: add more perf_counters to BlueStore ([pr#13274](#), Igor Fedotov)
- bluestore: os/bluestore: add new garbage collector ([pr#12144](#), Igor Fedotov)
- bluestore: os/bluestore: add perf variable for throttle info in bluestore ([pr#12583](#), Pan Liu)
- bluestore: os/bluestore: add “_” prefix for internal methods ([pr#13409](#), xie xingguo)
- bluestore: os/bluestore: align reclaim size to bluefs_alloc_size ([pr#14744](#), Haomai Wang)
- bluestore: os/bluestore/Allocator: drop unused return value in release function ([pr#13913](#), wangzhengyong)
- bluestore: os/bluestore: allow multiple DeferredBatches in flight at once ([issue#20295](#), [pr#16769](#), Nathan Cutler, Sage Weil)
- bluestore: os/bluestore: allow multiple SPDK BlueStore OSD instances ([issue#16966](#), [pr#12604](#), Orlando Moreno)
- bluestore: os/bluestore: assert blob map returns success ([pr#14473](#), shiqi)
- bluestore: os/bluestore: avoid nullptr in bluestore_extent_ref_map_t::bound_encode ([pr#14073](#), Sage Weil)
- bluestore: os/bluestore: avoid unnecessary memory copy, use variable reference in BlockDevice::Open ([pr#12942](#), liuchang0812)
- bluestore: os/bluestore: better debug output on unsharing blobs ([issue#20227](#), [pr#15746](#), Sage Weil)
- bluestore: os/bluestore/BitAllocator: fix bug of checking required blocks ([pr#13470](#), wangzhengyong)
- bluestore: os/bluestore/BitMapAllocator: rm unused variable ([pr#13599](#), Jie Wang)
- bluestore: os/bluestore/BitmapFreelistManager: readability improvements ([pr#12719](#), xie xingguo)
- bluestore: os/bluestore/BlockDevice: support pmem device as bluestore backend ([pr#15102](#), Jianpeng Ma)
- bluestore: os/bluestore/BlueFS: clean up log_writer aios from compaction ([issue#20454](#), [pr#16017](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: clear current log entrys before dump all fnode ([pr#15973](#), Jianpeng Ma)
- bluestore: os/bluestore/BlueFS: fix reclaim_blocks ([issue#18368](#), [pr#12725](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: Rebuild memcpy for bufferlist::page_aligned_app... ([pr#15728](#), Jianpeng Ma, Sage Weil)
- bluestore: os/bluestore/BlueFS: .slow should be compared with dirname ([pr#15595](#), zhanglei)
- bluestore: os/bluestore/BlueStore: Avoid double counting state_kv_queued_lat ([pr#16374](#), Jianpeng Ma)
- bluestore: os/bluestore/BlueStore.cc:remove unuse code in _open_bdev() ([pr#13553](#), yonghengdexin735)
- bluestore: os/bluestore/BlueStore.cc: remove unused variable ([pr#12703](#), Li Wang)
- bluestore: os/bluestore/BlueStore: no device no symlink ([pr#15721](#), Jianpeng Ma)
- bluestore: os/bluestore/BlueStore: remove unused code ([pr#16522](#), Jianpeng Ma)
- bluestore: os/bluestore: cleanup BitAllocator ([pr#12661](#), xie xingguo)
- bluestore: os/bluestore: cleanup bluestore_types ([pr#15680](#), xie xingguo)
- bluestore: os/bluestore: clean up flush logic ([pr#14162](#), Jianpeng Ma)
- bluestore: os/bluestore: cleanup, got rid of table reference of 1<<x ([pr#13718](#), Adam Kupczyk)
- bluestore: os/bluestore: clean up Invalid return value judgment ([pr#14219](#), shiqi)
- bluestore: os/bluestore: cleanup min_alloc_size; some formatting nits ([pr#15826](#), xie xingguo)
- bluestore: os/bluestore: clear result in BlueRocksEnv::getChildren ([issue#20857](#), [pr#16683](#), liuchang0812)
- bluestore: os/bluestore: clear up redundant size assignment in KerenelDevice ([pr#16121](#), Shasha Lu)
- bluestore: os/bluestore: conditionally load crr option ([pr#12877](#), xie xingguo)
- bluestore: os/bluestore: configure rocksdb cache via bluestore_cache_kv_ratio ([pr#15580](#), Sage Weil)
- bluestore: os/bluestore: default journal media to store media if bluefs is disabled ([pr#16844](#), xie xingguo)
- bluestore: os/bluestore: _do_remove: dirty shard individually as each blob is unshared ([issue#20849](#), [pr#16822](#), Sage Weil)
- bluestore: os/blueStore: Failure retry for opening file ([pr#16237](#), Yankun Li)
- bluestore: os/bluestore: fix a bug in small write handling on sharded extents ([pr#13728](#), Igor Fedotov)

- bluestore: os/bluestore: fix Allocator::allocate() int truncation ([issue#18595](#), [pr#13010](#), Sage Weil)
- bluestore: os/bluestore: fix a typo about bleystore ([pr#15357](#), Dongsheng Yang)
- bluestore: os/bluestore: fix BitMapAllocator assert on out-of-bound hint value ([pr#15289](#), Igor Fedotov)
- bluestore: os/bluestore: fix buffers pinned by indefinitely deferred writes ([pr#15398](#), Sage Weil)
- bluestore: os/bluestore: fix bug for calc extent_avg in reshared function ([pr#13931](#), wangzhengyong)
- bluestore: os/bluestore: fix bug in aio_read() ([pr#13511](#), tangwenjun)
- bluestore: os/bluestore: fix bug in _open_alloc() ([pr#13577](#), yonghengdexin735)
- bluestore: os/bluestore: fix bug in _open_super_meta() ([pr#13559](#), Taeksang Kim)
- bluestore: os/bluestore: fix bugs in bluefs and bdev flush ([issue#19250](#), [issue#19251](#), [pr#13911](#), Sage Weil)
- bluestore: os/bluestore: fix coredump in register_ctrlr() ([pr#13556](#), tangwenjun)
- bluestore: os/bluestore: fix deferred_aio deadlock ([pr#16051](#), Sage Weil)
- bluestore: os/bluestore: fix deferred write race ([issue#19880](#), [pr#15004](#), Sage Weil)
- bluestore: os/bluestore: fix deferred writes vs collection split race ([issue#19379](#), [pr#14157](#), Sage Weil)
- bluestore: os/bluestore: fix dirty_range on _do_clone_range ([issue#20810](#), [pr#16738](#), Sage Weil)
- bluestore: os/bluestore: fix false assert in IOContext::aio_wake ([pr#15268](#), Igor Fedotov)
- bluestore: os/bluestore: fix false asserts in Cache::trim_all() ([pr#15470](#), xie xingguo)
- bluestore: os/bluestore: fix fsck deferred_replay ([pr#15295](#), Sage Weil)
- bluestore: os/bluestore: fix min_alloc_size at mkfs time ([pr#13192](#), Sage Weil)
- bluestore: os/bluestore: fix narrow osr->flush() race ([pr#14489](#), Sage Weil)
- bluestore: os/bluestore: fix NVMEDevice::open failure if serial number ends with a ... ([pr#12956](#), Hongtong Liu)
- bluestore: os/bluestore: fix OnodeSizeTracking testing ([issue#20498](#), [pr#12684](#), xie xingguo)
- bluestore: os/bluestore: fix perf counters ([pr#13965](#), Sage Weil)
- bluestore: os/bluestore: fix possible out of order shard(offset == 0); add sanity check ([pr#15658](#), xie xingguo)
- bluestore: os/bluestore: fix potential access violation ([pr#15657](#), xie xingguo)
- bluestore: os/bluestore: fix potential assert in cache_trim method ([pr#13234](#), Igor Fedotov)
- bluestore: os/bluestore: fix reclaim_blocks and clean up Allocator interface ([issue#18573](#), [pr#12963](#), Sage Weil)
- bluestore: os/bluestore: fix typo(s)/trasnaction/transaction/ ([pr#14890](#), xie xingguo)
- bluestore: os/bluestore: fix unsharing blob dirty_range args ([issue#20227](#), [pr#15766](#), Sage Weil)
- bluestore: os/bluestore: fix use after free race with aio_wait ([pr#14956](#), Sage Weil)
- bluestore: os/bluestore: fix wal-queue bytes-counter to keep pace with others ([pr#13382](#), xie xingguo)
- bluestore: os/bluestore: fsck: verify blob.unused field ([pr#14316](#), Sage Weil)
- bluestore: os/bluestore: handle rounding error in cache ratios ([pr#15672](#), Sage Weil)
- bluestore: os/bluestore: implement collect_metadata ([pr#14115](#), Sage Weil)
- bluestore: os/bluestore: include logical object offset in crc error ([pr#13074](#), Sage Weil)
- bluestore: os/bluestore: initialize finishers properly ([pr#15666](#), xie xingguo)
- bluestore: os/bluestore/KernelDevice: fix comments ([pr#15264](#), xie xingguo)
- bluestore: os/bluestore/KernelDevice: fix debug message ([pr#13135](#), Sage Weil)
- bluestore: os/bluestore/KernelDevice: helpful warning when aio limit exhausted ([pr#15116](#), Sage Weil)
- bluestore: os/bluestore/KernelDevice: kill zeros ([pr#12856](#), xie xingguo)
- bluestore: os/bluestore: kill BufferSpace.empty() ([pr#12871](#), xie xingguo)
- bluestore: os/bluestore: kill orphan declaration of do_write_check_depth() ([pr#12853](#), xie xingguo)
- bluestore: os/bluestore: leverage the type knowledge in BitMapAreaLeaf ([pr#13736](#), Radoslaw Zarzynski)
- bluestore: os/bluestore: Make BitmapFreelistManager kv iterator short lived ([pr#16243](#), Mark Nelson)
- bluestore: os/bluestore: make live changes for BlueStore throttle config work like initial config ([pr#14225](#), J. Eric Ivancich)
- bluestore: os/bluestore: miscellaneous fixes to BitAllocator ([pr#12696](#), xie xingguo)
- bluestore: os/bluestore: misc fix and cleanups ([pr#16315](#), Jianpeng Ma)
- bluestore: os/bluestore: misc fixes ([pr#14333](#), Sage Weil)
- bluestore: os/bluestore: move aio.h/cc from fs dir to bluestore dir ([pr#16409](#), Pan Liu)
- bluestore: os/bluestore: move object exist in assign nid ([pr#16117](#), Jianpeng Ma)
- bluestore: os/bluestore: move sharedblob to new collection in same shard ([issue#20358](#), [pr#15783](#), Sage Weil)
- bluestore: os/bluestore: narrow cache lock range; make sure min_alloc_size p2 aligned ([pr#15911](#), xie xingguo)
- bluestore: os/bluestore: “noid” is not always necessary in clone op ([pr#13769](#), wangzhengyong)
- bluestore: os/bluestore: nullptr in OmapIteratorImpl::valid ([pr#12900](#), Xinze Chi)
- bluestore: os/bluestore/NVMEDevice: Add multiple thread support for SPDK I/O thread ([pr#14420](#), Ziye Yang)
- bluestore: os/bluestore/NVMEDevice.cc: fix the random read issue ([pr#13055](#), optimistyzy)
- bluestore: os/bluestore/NVMEDevice: fix the compilation issue for collect_metadata ([pr#14455](#), optimistyzy)
- bluestore: os/bluestore/NVMEdevice: fix the unrelease segs issue ([pr#12862](#), optimistyzy)
- bluestore: os/bluestore: only submit deferred if there is any ([pr#16269](#), Sage Weil)
- bluestore: os/bluestore: preallocate object[extent_shard] key to avoid reallocate ([pr#12644](#), xie xingguo)
- bluestore: os/bluestore: pre-calculate number of ghost buffers to evict ([pr#15029](#), xie xingguo)
- bluestore: os/bluestore: put strings in mempool ([pr#12651](#), Allen Samuels, Sage Weil)
- bluestore: os/bluestore: Record l_bluestore_state_kv_queued_lat for sync_submit ([pr#14448](#), Jianpeng Ma)
- bluestore: os/bluestore: reduce some overhead for _do_clone_range() and _do_remove() ([pr#15944](#), xie xingguo)
- bluestore: os/bluestore: refactor BlueStore::_do_write; kill dead ExtentMap::find_lextent() method ([pr#15750](#), xie xingguo)

- bluestore: os/bluestore: refactor ExtentMap::update to avoid preceding db updat... ([pr#12394](#), Igor Fedotov)
- bluestore: os/bluestore: remove a never read value ([pr#12618](#), liuchang0812)
- bluestore: os/bluestore: Remove ExtentFreeListManager ([pr#14772](#), Jianpeng Ma)
- bluestore: os/bluestore: remove intermediate key var to avoid string copy ([pr#12643](#), xie xingguo)
- bluestore: os/bluestore: remove no use parameter in bluestore_blob_t::map_bh ([pr#13013](#), wangzhengyong)
- bluestore: os/bluestore: remove unneeded indirection in BitMapZone ([pr#13743](#), Radoslaw Zarzynski)
- bluestore: os/bluestore: remove unused condition variable ([pr#14973](#), Igor Fedotov)
- bluestore: os/bluestore: remove unused local variable “pos” ([pr#13715](#), wangzhengyong)
- bluestore: os/bluestore: remove unused variables ([pr#15718](#), zhanglei)
- bluestore: os/bluestore: rename/fix throttle options ([pr#14717](#), Sage Weil)
- bluestore: os/bluestore: roundoff bluefs allocs to bluefs_alloc_size ([pr#14876](#), Ramesh Chander)
- bluestore: os/bluestore: shrink buffer_map key into uint32_t ([pr#12850](#), xie xingguo)
- bluestore: os/bluestore: slightly refactor Blob::try_reuse_blob ([pr#15836](#), xie xingguo)
- bluestore: os/bluestore: some cleanup ([pr#13390](#), liuchang0812)
- bluestore: os/bluestore: space between func and contents ([pr#16804](#), xie xingguo)
- bluestore: os/bluestore: stop calculating bound if we must reshards; narrow shard combination condition ([pr#15631](#), xie xingguo)
- bluestore: os/bluestore/StupidAllocator: rounded down len to an align boundary ([issue#20660](#), [pr#16593](#), Zhu Shangzhong)
- bluestore: os/bluestore: target_bytes should scale with meta/data ratios ([pr#15708](#), Mark Nelson)
- bluestore: os/bluestore: _txc_release_alloc when do wal cleaning ([pr#12692](#), Xinze Chi)
- bluestore: os/bluestore: use bufferlist functions whenever possible ([pr#16158](#), Jianpeng Ma)
- bluestore: os/bluestore: use correct bound encode size for unused ([pr#14731](#), Haomai Wang)
- bluestore: os/bluestore: use reference to avoid string copy ([pr#16364](#), Pan Liu)
- bluestore: os: extend ObjectStore interface to dump store’s performance counters ([pr#13203](#), Igor Fedotov)
- bluestore,performance: common/config_opts.h: compaction readahead for bluestore/rocksdb ([pr#14932](#), Mark Nelson)
- bluestore,performance: kv/RocksDBStore: implement rm_range_keys operator interface and test ([pr#13855](#), Haomai Wang)
- bluestore,performance: os/aio: remove the redundant memset(struct iocb) ([pr#13662](#), Jianpeng Ma)
- bluestore,performance: os/bluestore: add bluestore_prefer_wal_size option ([pr#13217](#), Sage Weil)
- bluestore,performance: os/bluestore: avoid overloading extents during reshards; atomic deferred_batch_ops ([pr#15502](#), xie xingguo)
- bluestore,performance: os/bluestore: avoid the VTABLE-related burden in BitMapAllocator’s hotspot ([pr#14348](#), Radoslaw Zarzynski)
- bluestore,performance: os/bluestore: batch throttle ([pr#15284](#), Jianpeng Ma)
- bluestore,performance: os/bluestore/BlueFS: add bluefs_sync_write option ([pr#14510](#), Sage Weil)
- bluestore,performance: os/bluestore/BlueFS: optimize get_allocated ([pr#14121](#), Jianpeng Ma)
- bluestore,performance: os/bluestore/BlueFS: tune flushing of writes ([pr#13032](#), Sage Weil)
- bluestore,performance: os/bluestore/bluestore_types: drop std::bitset for blob unused ([pr#12569](#), Sage Weil)
- bluestore,performance: os/bluestore: cap rocksdb cache size ([pr#15786](#), Mark Nelson)
- bluestore,performance: os/bluestore: default 16KB min_alloc_size on ssd ([pr#14076](#), Sage Weil)
- bluestore,performance: os/bluestore: default cache size of 3gb ([pr#15976](#), Sage Weil)
- bluestore,performance: os/bluestore: differ default cache size for hdd/ssd backends ([pr#16157](#), xie xingguo)
- bluestore,performance: os/bluestore: do not balance bluefs on every kv_sync_thread iteration ([pr#14557](#), Sage Weil)
- bluestore,performance: os/bluestore: do not cache shard keys ([pr#12634](#), Sage Weil)
- bluestore,performance: os/bluestore: eliminate some excessive stuff ([pr#14675](#), Igor Fedotov)
- bluestore,performance: os/bluestore: fix deferred writes; improve flush ([pr#13888](#), Sage Weil)
- bluestore,performance: os/bluestore: generate same onode extent-shard keys in a more efficient way ([pr#12681](#), xie xingguo)
- bluestore,performance: os/bluestore: get rid off excessive lock at BitMapAllocator ([pr#14749](#), Igor Fedotov)
- bluestore,performance: os/blueStore: In osd_tp_thread, call _txc_finalize_kv ([pr#14709](#), Jianpeng Ma)
- bluestore,performance: os/bluestore: keep stats replica in RAM to avoid expensive KV retrieval ([pr#15309](#), Igor Fedotov)
- bluestore,performance: os/bluestore/KernelDevice: batch aio submit ([pr#16032](#), Haodong Tang)
- bluestore,performance: os/bluestore/KernelDevice: fix sync write vs flush ([pr#15034](#), Sage Weil)
- bluestore,performance: os/bluestore: kvdb histogram ([pr#12620](#), Varada Kari)
- bluestore,performance: os/bluestore: make bluestore_max_blob_size parameter hdd/ssd case dependant ([pr#14434](#), Igor Fedotov)
- bluestore,performance: os/bluestore: memory and dereference clean-up in the BitAllocator ([pr#13811](#), Radoslaw Zarzynski)
- bluestore,performance: os/bluestore: move cache_trim into MempoolThread ([pr#15380](#), xie xingguo)
- bluestore,performance: os/bluestore: optimize blob usage when doing appends/overwrites ([pr#13337](#), Igor Fedotov)
- bluestore,performance: os/bluestore: optimized (encode|decode)_escaped ([pr#15759](#), Piotr Dałek)
- bluestore,performance: os/bluestore: partial reshards support ([pr#13162](#), Sage Weil)
- bluestore,performance: os/bluestore: prevent lock for almost “flush” calls ([pr#12524](#), Haomai Wang)
- bluestore,performance: os/bluestore: put bluefs in the middle of the shared device ([pr#14873](#), Sage Weil)

- bluestore,performance: os/bluestore: refactor small write handling to reuse blob more effect... ([pr#14399](#), Igor Fedotov)
- bluestore,performance: os/bluestore: remove CephContext* from BmapEntry ([pr#13651](#), Radoslaw Zarzynski)
- bluestore,performance: os/bluestore: replace Blob ref_map with reference counting ([pr#12904](#), Igor Fedotov)
- bluestore,performance: os/bluestore: rewrite deferred write handling ([issue#16644](#), [pr#14491](#), Sage Weil)
- bluestore,performance: os/bluestore: separate kv_sync_thread into two parts ([pr#14035](#), Jianpeng Ma, Igor Fedotov, Sage Weil)
- bluestore,performance: os/bluestore: set cache meta ratio to .9 ([pr#12635](#), Sage Weil)
- bluestore,performance: os/bluestore: the exhausted check in BitMapZone can be lock-less ([pr#13653](#), Radoslaw Zarzynski)
- bluestore,performance: os/bluestore: try to unshare blobs for EC overwrite workload ([pr#14239](#), Sage Weil)
- bluestore,performance: os/bluestore: tune deferred_batch_ops separately for hdd and ssd ([pr#14435](#), Sage Weil)
- bluestore,performance: os/bluestore: unify throttling model ([issue#19542](#), [pr#14306](#), Sage Weil)
- bluestore,performance: os/bluestore: use aio for reads ([issue#19030](#), [pr#13066](#), Sage Weil)
- bluestore,performance: os/bluestore: use Best-Effort policy when evicting onode from cache ([pr#12876](#), xie xingguo)
- bluestore,performance: os/bluestore: use denc for varint encoding ([pr#14911](#), Piotr Dałek)
- bluestore,performance: os/bluestore: various onode changes to reduce its in-memory footprint ([pr#12700](#), Igor Fedotov)
- bluestore,performance: os/fs/aio: use small_vector for aio_t; clean up header location ([pr#14853](#), Sage Weil)
- bluestore: rocksdb: add option: writable_file_max_buffer_size = 0 ([pr#12562](#), Jianpeng Ma)
- bluestore,tests: ceph-dencoder: enable bluestore types ([pr#13595](#), Willem Jan Withagen, Kefu Chai)
- bluestore,tests: ceph_test_objectstore: match clone_range src and dst offset ([pr#13211](#), Sage Weil)
- bluestore,tests: qa/objectstore/bluestore*: fsck on mount ([pr#15785](#), Sage Weil)
- bluestore,tests: test/ceph-test-objectstore: Don't always include BlueStore code ([pr#13516](#), Willem Jan Withagen)
- bluestore,tests: test/objectstore/store_fixture.cc: Exclude bluestore code if required ([pr#14085](#), Willem Jan Withagen)
- bluestore,tests: test/store_test: add deferred test case setup to support explicit min... ([issue#18857](#), [pr#13415](#), Igor Fedotov)
- bluestore,tests: test/store_test: fix bluestore test cases disablement ([pr#14228](#), Igor Fedotov)
- bluestore,tests: test/unittest_bluefs: check whether add_block_device success ([pr#14013](#), shiqi)
- bluestore,tests: test/unittest_bluefs: When fsync ret is less than 0, fsync can not be... ([pr#15365](#), shiqi)
- bluestore,tests: unittest_alloc: add test_alloc_big ([issue#16662](#), [pr#14844](#), Sage Weil)
- bluestore,tools: ceph-bluestore-tool: rename from bluefs-tool; improve usage ([pr#14258](#), Sage Weil)
- bluestore,tools: ceph-kvstore-tool: allow 'bluestore-kv' as kvdb type; add escaping, compaction ([pr#14718](#), Sage Weil)
- bluestore: wrap blob id when it reaches maximum value of int16_t ([issue#19555](#), [pr#15654](#), Xiaoyan Li)
- build/ops: 12.0.3 ([pr#15600](#), Jenkins Build Slave User)
- build/ops: add 12.0.1 release tag in master ([pr#14690](#), Jenkins Build Slave User)
- build/ops: add psmisc dependency to ceph-base (deb and rpm) ([issue#19129](#), [pr#13744](#), Nathan Cutler)
- build/ops: add sanity checks to run-make-check.sh ([pr#12683](#), Nathan Cutler)
- build/ops: alpine: add alpine linux dev support ([pr#9853](#), John Coyle)
- build/ops: arch: fix build on PowerPC with FreeBSD ([pr#14378](#), Andrew Solomon)
- build/ops: arch: fix cmake's ARM CRC intrinsics test to handle duplicitous gcc 4.8.5 ([issue#19386](#), [pr#14132](#), Dan Mick)
- build/ops: arch: use __get_cpuid instead of do_cpuid ([issue#7869](#), [pr#14857](#), Jos Collin)
- build/ops: auth: Let's not use the deprecated cephx option ([pr#12721](#), Dave Chen)
- build/ops: build: Add Virtuozzo Linux support ([pr#14301](#), Andrey Parfenov)
- build/ops: build: build erasure-code isa lib without versions ([pr#16205](#), James Page)
- build/ops: build/cmake: provide asan, tsan, ubsan builds ([pr#12615](#), Matt Benjamin)
- build/ops: build: execute dh_systemd_{enable,start} after dh_install ([issue#19585](#), [pr#16218](#), James Page)
- build/ops: build: move bash_completion.d/ceph to ceph-common ([pr#15148](#), Leo Zhang)
- build/ops: build: remove ceph-disk-udev entirely ([pr#15259](#), Leo Zhang)
- build/ops: build: remove ceph-qa-suite directory ([pr#13880](#), Casey Bodley)
- build/ops: build: revert -Wvla from #15342 ([pr#15469](#), Willem Jan Withagen)
- build/ops: builds with dpdk v16.07 ([pr#12707](#), Kefu Chai)
- build/ops: build: Use .S suffix for ppc64le assembly files ([issue#20106](#), [pr#15373](#), Andrew Solomon)
- build/ops: ceph-disk: ability to use a different cluster name with dmcrypt ([issue#17821](#), [pr#11786](#), Sébastien Han, Erwan Velu)
- build/ops: ceph-disk: don't activate suppressed journal devices ([issue#19489](#), [pr#16123](#), David Disseldorp)
- build/ops: ceph.in: allow developer mode from outside build tree ([issue#20472](#), [pr#16055](#), Dan Mick)
- build/ops: ceph_release: we are in the 'rc' phase (12.1.z) ([pr#15957](#), Sage Weil)
- build/ops: ceph.spec.in, debian/control: Add bc to build dependencies ([issue#18876](#), [pr#13338](#), Kyr Shatskyy)
- build/ops: Clean up make check for persistent test nodes (like arm64) ([pr#16773](#), Dan Mick)
- build/ops: cmake,crc32c: conditionalize crc32c on different archs ([pr#14289](#), Kefu Chai)
- build/ops: CMakeLists.txt: boost_python.so requires libpython.*.so on FreeBSD ([pr#12763](#), Willem Jan Withagen)
- build/ops: CMakeLists.txt: don't do crypto/isa-l if not Intel ([pr#14721](#), Dan Mick)
- build/ops: CMakeLists.txt: suppress unneeded warning about jemalloc ([pr#13377](#), Willem Jan Withagen)
- build/ops,common: build: Adds C++ warning flag for C Variable-Length Arrays ([pr#15342](#), Jesse Williamson)
- build/ops,common: common/blkdev.cc: propagate get_device_by_fd to different OSes ([pr#15547](#), Willem Jan Withagen)

- build/ops: common/module.c: do not use strerror_r the GNU way ([pr#12363](#), Willem Jan Withagen)
- build/ops: compressor/zstd: add zstd to embedded ceph ([pr#13159](#), Bassam Tabbara)
- build/ops: conditionalize rgw Beast frontend so it isn't built on s390x architecture ([issue#20048](#), [pr#15225](#), Willem Jan Withagen, Nathan Cutler, Kefu Chai, Tim Serong, Casey Bodley)
- build/ops,core: build: let FreeBSD build ceph-fuse ([pr#14282](#), Willem Jan Withagen)
- build/ops,core: ceph-disk: use correct user in check_journal_req ([issue#18538](#), [pr#12947](#), Samuel Matzek)
- build/ops,core: common/freebsd_errno.cc: fix missing ([pr#15741](#), Willem Jan Withagen)
- build/ops,core: erasure-code: update ec_isa version + add missing AVX512 ISA-L sources ([pr#15636](#), Ganesh Mahalingam, Tushar Gohad)
- build/ops,core: os: allow offline conversion of filestore -> bluestore (or anything else) ([pr#14210](#), Sage Weil)
- build/ops,core: osd/OSD: auto class on osd start up ([pr#16014](#), xie xingguo)
- build/ops,core: osd/Pool: Disallow enabling ‘hashpspool’ option to a pool without ‘-yes-i-really-mean-it’ ([issue#18468](#), [pr#13406](#), Vikhyat Umrao)
- build/ops,core,tests: osd/dmclock/testing: reorganize testing, building now optional ([issue#19987](#), [pr#15375](#), J. Eric Ivancich)
- build/ops: debian: Add missing tp files in deb packaging ([pr#13526](#), Ganesh Mahalingam)
- build/ops: debian: ceph-mgr: fix package description ([pr#15513](#), Fabian Grünbichler)
- build/ops: debian/control: add ceph-base-dbg ([pr#13796](#), Sage Weil)
- build/ops: debian: drop boost build dependencies ([pr#13524](#), Kefu Chai)
- build/ops: debian: package ceph.logrotate properly ([issue#19390](#), [pr#14600](#), Kefu Chai)
- build/ops: debian: package crypto plugin only on amd64 ([pr#14820](#), Kefu Chai)
- build/ops: debian/rpm: move radosgw-admin to ceph-common ([issue#19577](#), [pr#14940](#), Ali Maredia)
- build/ops: debian/rules, ceph.spec.in: invoke cmake with -DBOOST_J ([pr#14114](#), Dan Mick)
- build/ops: debian: sync logrotate packaging with downstream ([issue#19938](#), [pr#15567](#), Fabian Grünbichler)
- build/ops: debian: workaround the bug in dpkg-maintscript-helper ([issue#20453](#), [pr#16072](#), Kefu Chai)
- build/ops: debian: wrap-and-sort all files ([pr#16110](#), James Page)
- build/ops: dmclock: error: ‘function’ in namespace ‘std’ does not name a template type ([pr#14909](#), Jos Collin)
- build/ops: dmclock: include missing <functional> header ([pr#14923](#), Jos Collin)
- build/ops: dmclock: initial commit of dmclock QoS library ([pr#14330](#), J. Eric Ivancich)
- build/ops: do_cmake.sh: enable ccache if installed ([pr#15274](#), Sage Weil)
- build/ops: do_cmake.sh: fix syntax for /bin/sh (doesn’t have +=) ([pr#16433](#), Dan Mick)
- build/ops: do_freebsd.sh: Remove ENODATA requirement ([pr#13626](#), Willem Jan Withagen)
- build/ops: drop libfcgi build dependency ([pr#15285](#), Nathan Cutler)
- build/ops: gitignore: Ignore rejects by patch ([pr#14405](#), Willem Jan Withagen)
- build/ops: include/assert: test c++ before using static_cast<> ([pr#16424](#), Kefu Chai)
- build/ops: init-ceph: add ceph libraries path to environment ([pr#14693](#), Mohamad Gebai)
- build/ops: init-ceph: fix ceph user args ([pr#13467](#), Sage Weil)
- build/ops: init-ceph: Make init-ceph work under FreeBSD for init-system ([pr#13209](#), Willem Jan Withagen)
- build/ops: install-deps.sh: add missing dependencies for FreeBSD ([pr#16545](#), Alan Somers)
- build/ops: install-deps.sh: workaround setuptools’ dependency on six ([pr#15406](#), Kefu Chai)
- build/ops: mailmap: Update OVH contributors ([pr#13063](#), Bartłomiej Święcki)
- build/ops: make package groups comply with openSUSE guidelines ([issue#19184](#), [pr#13781](#), Nathan Cutler)
- build/ops: make-srpm: Pass first parameter to make-dist for building SRPM ([pr#13480](#), Wido den Hollander)
- build/ops: merge v12.0.2 release tag ([pr#15091](#), Jenkins Build Slave User)
- build/ops,mgr: debian/ceph-base.dirs: create bootstrap-mgr dirs ([pr#14838](#), Sage Weil)
- build/ops: miscellaneous cleanups and fixes (run-make-check.sh, ceph.spec.in) ([issue#20091](#), [issue#20127](#), [pr#15399](#), Nathan Cutler)
- build/ops,mon: mon/ConfigKeyService: add ‘config-key dump’ to show keys and vals ([pr#14858](#), Dan Mick)
- build/ops,mon: systemd: Restart Mon after 10s in case of failure ([issue#18635](#), [pr#13057](#), Wido den Hollander)
- build/ops: msg/async/rdma: compile with rdma as default ([pr#13901](#), DanielBar-On)
- build/ops: os/bluestore: fix build errors when spdk is on ([pr#16118](#), Ilsoo Byun)
- build/ops: packaging: install libceph-common.so* not libceph-common.so.* ([issue#18692](#), [pr#13148](#), Kefu Chai)
- build/ops,performance: crc32c: Add crc32c function optimized for ppc architecture ([pr#13909](#), Andrew Solomon)
- build/ops,performance,rbd: byteorder: use gcc intrinsics for byteswap ([pr#15012](#), Kefu Chai)
- build/ops,rbd,rgw: CMakeLists: trim rbd/rgw forced dependencies ([pr#16574](#), Patrick Donnelly)
- build/ops,rbd,tests: test/librbd: decouple ceph_test_librbd_api from libceph-common ([issue#20175](#), [pr#15611](#), Kefu Chai)
- build/ops,rbd,tests: test/librbd: re-enable internal tests in ceph_test_librbd ([pr#16255](#), Mykola Golub)
- build/ops,rbd,tests: test: Need to exclude the fsx executable also on FreeBSD ([pr#13686](#), Willem Jan Withagen)
- build/ops: Revert “msg/async: increase worker reference with local listen table enabled backend” ([issue#20603](#), [pr#16323](#), Haomai Wang)
- build/ops: Revert “msg/async/rdma: Debug prints for ibv” ([pr#14245](#), Kefu Chai)
- build/ops,rgw: rgw_file: radosgw-admin can be built under FreeBSD ([pr#12191](#), Willem Jan Withagen)
- build/ops,rgw,tests,tools: vstart: allow to start multiple radosgw when RGW=x ([pr#15632](#), Adam Kupczyk)
- build/ops,rgw,tools: vstart: add -rgw_compression to set rgw compression plugin ([pr#15929](#), Casey Bodley)
- build/ops: rocksdb: build with ppc64 ([pr#12908](#), Kefu Chai)

- build/ops: rocksdb: sync with upstream ([pr#14456](#), Kefu Chai)
- build/ops: rocksdb: sync with upstream ([pr#14818](#), Nathan Cutler, Kefu Chai)
- build/ops: rpm: apply epoch only if %epoch macro is defined ([pr#15286](#), Nathan Cutler)
- build/ops: rpm: build ceph-resource-agents by default ([issue#17613](#), [pr#13515](#), Nathan Cutler)
- build/ops: rpm: bump epoch ahead of RHEL base ([issue#20508](#), [pr#16126](#), Ken Dreyer)
- build/ops: rpm,deb: fix ceph-volume ([issue#20915](#), [pr#16832](#), Sage Weil)
- build/ops: rpm: disable dwz to speed up valgrind ([issue#19099](#), [pr#13748](#), Kefu Chai)
- build/ops: rpm: drop boost build dependencies ([pr#13519](#), Nathan Cutler)
- build/ops: rpm: Drop legacy libxio support ([pr#16449](#), Nathan Cutler)
- build/ops: rpm: fix python-Sphinx package name for SUSE ([pr#15015](#), Nathan Cutler, Jan Matejek)
- build/ops: rpm: fix typo WTIH_BABELTRACE ([pr#16366](#), Nathan Cutler)
- build/ops: rpm: Fix undefined FIRST_ARG ([issue#20077](#), [pr#16208](#), Boris Ranto)
- build/ops: rpm: gperftools-devel >= 2.4 ([issue#13522](#), [pr#14870](#), Nathan Cutler)
- build/ops: rpm: make librbd1 %post scriptlet depend on coreutils ([issue#20052](#), [pr#15231](#), Giacomo Comes, Nathan Cutler)
- build/ops: rpm: move _epoch_prefix below Epoch definition ([pr#15417](#), Nathan Cutler)
- build/ops: rpm: move RDMA and python-prettytables build dependencies to distro-conditional section ([pr#15200](#), Nathan Cutler)
- build/ops: rpm: obsolete libcephfs1 ([pr#16074](#), Nathan Cutler)
- build/ops: rpm: package COPYING, move sample ceph.conf to ceph-common ([pr#15596](#), Nathan Cutler)
- build/ops: rpm: package crypto on x86_64 only ([pr#14779](#), Nathan Cutler)
- build/ops: rpm: put mgr python build dependencies in make_check bcond ([issue#20425](#), [pr#15940](#), Nathan Cutler, Tim Serong)
- build/ops: rpm: sane packaging of %{_docdir}/ceph directory ([pr#15900](#), Nathan Cutler)
- build/ops: script: adding contributor credits script ([pr#13251](#), Patrick McGarry)
- build/ops: script: drop the -x arg for credits script ([pr#14296](#), Abhishek Lekshmanan)
- build/ops: script/sepio_bt.sh: download packages from shaman not gitbuilder ([pr#12799](#), Kefu Chai)
- build/ops: script/sepio_bt.sh: get sha1,release from t.log if it's not in core ([pr#13620](#), Kefu Chai)
- build/ops: script/sepio_bt.sh: support xenial ([pr#13292](#), Kefu Chai)
- build/ops: selinux: Allow ceph daemons to read net stats ([issue#19254](#), [pr#13945](#), Boris Ranto)
- build/ops: selinux: Allow read on var_run_t ([issue#16674](#), [pr#15523](#), Boris Ranto)
- build/ops: selinux: Do parallel relabel on package install ([issue#20077](#), [pr#14871](#), Boris Ranto)
- build/ops: selinux: Install ceph-base before ceph-selinux ([issue#20184](#), [pr#15490](#), Boris Ranto)
- build/ops: Set subman cron attributes in spec file ([issue#20074](#), [pr#15270](#), Thomas Serlin)
- build/ops: spdk: upgrade spdk to v16.12 ([pr#12734](#), Pan Liu)
- build/ops: src/CMakeLists.txt: disable -Werror on rocksdb ([pr#12560](#), Willem Jan Withagen)
- build/ops: src/CMakeLists.txt: Move parse_secret_objs setting within definition block ([pr#12785](#), Willem Jan Withagen)
- build/ops: src/init-ceph.in: allow one((re)?start|stop) as commands ([pr#14560](#), Willem Jan Withagen)
- build/ops: sync luminous tag back to master ([pr#16758](#), Jenkins Build Slave User)
- build/ops: systemd: Add explicit Before=ceph.target ([pr#15835](#), Tim Serong)
- build/ops: systemd/ceph-disk: make it possible to customize timeout ([issue#18740](#), [pr#13197](#), Alexey Sheplyakov)
- build/ops: systemd/ceph-mgr: remove automagic mgr creation hack ([issue#19994](#), [pr#16023](#), Sage Weil)
- build/ops: systemd: remove ceph-create-keys from presets ([pr#14226](#), Sébastien Han)
- build/ops: systemd: Start OSDs after MONs ([issue#18516](#), [pr#13097](#), Boris Ranto)
- build/ops: test/fio_ceph_objectstore: fix fio plugin build failure caused by rec... ([pr#12655](#), Igor Fedotov)
- build/ops,tests: qa: make run-standalone work on FreeBSD ([pr#16595](#), Willem Jan Withagen)
- build/ops,tests: test/osd/CMakeLists.txt: osd-dup.sh require BlueStore/AIO ([pr#14387](#), Willem Jan Withagen)
- build/ops,tests: test/osd/osd-dup.sh: warn on low open file limit ([pr#14637](#), Piotr Dałek)
- build/ops,tests,tools: vstart.sh: Work around mgr restfull not available ([pr#15877](#), Willem Jan Withagen)
- build/ops: The Clangtastic Mr. Clocks ([pr#15186](#), Adam C. Emerson)
- build/ops: tool: add some ceph relate processes to ps-ceph.pl ([pr#12406](#), songbaisen)
- build/ops: tools/scripts:"FreeBSD getopt is not compatible, use the one from packages" ([pr#13260](#), Willem Jan Withagen)
- build/ops: tracing: Fix error in including all files in osd_tp ([pr#12501](#), Ganesh Mahalingam)
- build/ops: upstart: start radosgw-all according to runlevel ([issue#18313](#), [pr#12586](#), Ken Dreyer)
- build/ops: vstart: clean up usage a bit ([pr#13138](#), Sage Weil)
- build/ops: vstart: do not start mgr if not start_all ([pr#13974](#), Kefu Chai)
- build/ops: yasm-wrapper: filter -pthread ([pr#15249](#), Alessandro Barbieri)
- build/ops: yasm-wrapper: strip -E (stops ccache trashing source files) ([pr#14633](#), Tim Serong)
- cephfs: #11950: Persistent purge queue ([issue#11950](#), [pr#12786](#), John Spray)
- cephfs: #17980: MDS client blacklisting and blacklist on eviction ([issue#17980](#), [issue#9754](#), [pr#14610](#), John Spray)
- cephfs: #18600: Clear out tasks that don't make sense from multimds suite ([issue#18600](#), [pr#13089](#), John Spray)
- cephfs: ceph_fuse: fix daemonization when pid file is non-empty ([pr#13532](#), "Yan, Zheng")
- cephfs: ceph_fuse: pid_file default to empty ([issue#18309](#), [pr#12628](#), Nathan Cutler)
- cephfs: ceph-fuse: use user space permission check by default ([issue#19820](#), [pr#14907](#), "Yan, Zheng")
- cephfs: ceph: simplify CInode::maybe_export_pin() ([pr#15106](#), "Yan, Zheng")

- cephfs: client: avoid returning negative space available (issue#20178, pr#15481, John Spray)
- cephfs: client: call the lru_remove() twice, when trim cache (pr#15662, huanwen ren)
- cephfs: client: check for luminous MDS before sending FLUSH_MLOG (pr#15805, John Spray)
- cephfs: client/Client.cc: after reset session from MDS - reconnect (issue#18757, pr#13522, Henrik Korkuc)
- cephfs: client/Client.cc: prevent segfaulting (issue#9935, pr#12550, Michal Jarzabek)
- cephfs: client: client_quota no longer optional (pr#14978, Dan van der Ster)
- cephfs: client: don't request lookup parent if ino is root (pr#12478, huanwen ren)
- cephfs: client: drop cap snaps when auth mds session gets closed (issue#19022, pr#13579, "Yan, Zheng")
- cephfs: client: fix clang warn of "argument is an uninitialized value" (pr#12580, liuchang0812)
- cephfs: client: fix Client::handle_cap_flushsnap_ack() crash (issue#18460, pr#12859, Yan, Zheng)
- cephfs: client: fix Dentry::dump (pr#15779, huanwen ren)
- cephfs: client: fix display ino in the Idout (pr#15314, huanwen ren)
- cephfs: client: fix potential buffer overflow (pr#12515, Yunchuan Wen)
- cephfs: client: fix the cross-quota rename boundary check conditions (pr#12489, Greg Farnum)
- cephfs: client: fix UserPerm::gid_in_group() (issue#19903, pr#15039, "Yan, Zheng")
- cephfs: client: getattr before returning quota/layout xattrs (issue#17939, pr#14018, John Spray)
- cephfs: client/inode: fix the dump type of Inode::dump() (pr#15198, huanwen ren)
- cephfs: client: populate metadata during mount (issue#18361, pr#12915, John Spray)
- cephfs: client: priority to verify the correctness of the "flag" (pr#12897, huanwen ren)
- cephfs: client: refine fsync/close writeback error handling (pr#14589, John Spray)
- cephfs: client: remove dead log code (pr#13093, Patrick Donnelly)
- cephfs: client: remove request from session->requests when handling forward (issue#18675, pr#13124, "Yan, Zheng")
- cephfs: client: simplify remove_cap interface (pr#12161, John Spray)
- cephfs: client: specify inode in get_caps log message (pr#13966, John Spray)
- cephfs: client: wait for lastest osdmap when handling set file/dir layout (issue#18914, pr#13580, "Yan, Zheng")
- cephfs,common: common/MemoryModel: Bump int to long and drop mallinfo (pr#13453, Xiaoxi Chen)
- cephfs,common,core: librados,osdc: kill ack vs commit distinction (pr#12607, Sage Weil)
- cephfs,common: include/fs_types: fix unsigned integer overflow (pr#12440, runsisi)
- cephfs,common,rbd: blkin: librbd trace hooks (pr#15053, Victor Araujo, Jason Dillaman)
- cephfs,common,rbd: osdc: cache should ignore error bhs during trim (issue#18436, pr#12966, Jason Dillaman)
- cephfs,core: Add test for is_hacky_ecoverwrites in cephfs pool checks (pr#13466, John Spray)
- cephfs,core: cleanup: use std::make_shared to replace new (pr#12276, Yunchuan Wen)
- cephfs,core,mon: mon/MDSMonitor: fix segv when multiple MDSs raise same alert (pr#16302, Sage Weil)
- cephfs: fix mount point break off problem after mds switch occurred (issue#19437, pr#14267, Guan yunfei)
- cephfs: fix write_buf's _len overflow problem (issue#19033, pr#13587, Yang Honggang)
- cephfs: fs/ceph-fuse: normalize file open flags on the wire (pr#14822, Jan Fajerski)
- cephfs: libcephfs.cc: fix memory leak (pr#12557, Michal Jarzabek)
- cephfs: libcephfs: cleanups (pr#12830, huanwen ren)
- cephfs: libcephfs: fix cct refcount constructing from rados (pr#12831, John Spray)
- cephfs: mds/MDBalancer: remove useless check_targets and hit_targets logic from MDS balancer (issue#20131, pr#15407, Zhi Zhang)
- cephfs: mds/MDLog.cc Fix perf counter type for jlat (pr#13449, Xiaoxi Chen)
- cephfs: mds/Server.cc: Don't evict a slow client if (issue#17855, pr#12935, Michal Jarzabek)
- cephfs: mds/StrayManager: avoid reusing deleted inode in StrayManager::_purge_stray_logged (issue#18877, pr#13347, Zhi Zhang)
- cephfs,mgr: pybind/mgr/fsstatus: use mds_mem.dn as dentry counter (pr#15255, Zhi Zhang)
- cephfs: Mitigation for #16842, validate sessions after load (issue#16842, pr#14164, John Spray)
- cephfs: mon/FSCCommand: fix indentation (pr#15423, Sage Weil)
- cephfs: mon/MDSMonitor.cc:refuse fs new on pools with obj (issue#11124, pr#12825, Michal Jarzabek)
- cephfs: mon/MDSMonitor: respect mds_standby_for_rank config (pr#15129, "Yan, Zheng")
- cephfs: mount: do not print "unknown" option to kclient (issue#18159, pr#12465, John Spray)
- cephfs: osdc/Filer: truncate large file party by party (issue#19755, pr#14769, "Yan, Zheng")
- cephfs: osdc/Journaler: avoid executing on_safe contexts prematurely (issue#20055, pr#15240, "Yan, Zheng")
- cephfs: osdc/Journaler: fix memory leak in Journaler::_issue_read() (issue#20338, pr#15776, "Yan, Zheng")
- cephfs: osdc/Objecter: fix inflight_ops update (pr#15768, "Yan, Zheng")
- cephfs: osdc: remove journaler_allow_split_entries option (issue#19691, pr#14636, John Spray)
- cephfs,performance: client: make seeky readdir more efficiency (issue#19306, pr#14317, "Yan, Zheng")
- cephfs,performance: mds/server: skip unwanted dn in handle_client_readdir (pr#12870, Xiaoxi Chen)
- cephfs: Permit recovering metadata into a new RADOS pool (issue#15069, issue#15068, pr#10636, Douglas Fuller)
- cephfs: qa/cephfs: disable mds_bal_frag for TestStrays.test_purge_queue_op_rate (issue#19892, pr#15105, "Yan, Zheng")
- cephfs: qa/cephfs: Fix for test_data_scan (issue#19893, pr#15094, Douglas Fuller)
- cephfs: qa: fix race in Mount.open_background (issue#18661, pr#13137, John Spray)
- cephfs: qa/suites/fs: reserve more space for mds in full tests (issue#19891, pr#15026, "Yan, Zheng")
- cephfs: qa/tasks/cephfs: use getattr to guarantee inode is in client cache (issue#19912, pr#15062, "Yan, Zheng")

- cephfs: qa: unpin knfs from ubuntu ([issue#16397](#), [pr#13088](#), John Spray)
- cephfs: qa: update log whitelists for kcephfs suite ([pr#14922](#), "Yan, Zheng")
- cephfs: qa: update remaining ceph.com to download.ceph.com ([issue#18574](#), [pr#12964](#), John Spray)
- cephfs: qa: whitelist new fullness messages in fs tests ([issue#19253](#), [pr#13915](#), John Spray)
- cephfs: Remove "experimental" warnings from multimds ([pr#15154](#), John Spray, "Yan, Zheng")
- cephfs: Rewrite mount.fuse.ceph (to python) and move ceph-fuse options to fs_mntops ([pr#11448](#), Edgaras Lukosevicius)
- cephfs: tasks/cephfs: fix kernel force umount ([issue#18396](#), [pr#12833](#), Yan, Zheng)
- cephfs: test/libcephfs: avoid buffer overflow when testing ceph_getdents() ([issue#18941](#), [pr#13429](#), "Yan, Zheng")
- cephfs,tests: Add multimds:thrash sub-suite and fix bugs in thrasher for multimds ([issue#18690](#), [issue#10792](#), [pr#13262](#), Patrick Donnelly)
- cephfs,tests: ceph-object-corpus: mark MMDSSlaveRequest incompat change ([pr#15730](#), Sage Weil)
- cephfs,tests: Improve vstart_runner to (optionally) create its own cluster ([pr#12800](#), John Spray)
- cephfs,tests: qa: fix float parse error in test_fragment ([pr#15122](#), Patrick Donnelly)
- cephfs,tests: qa: fix test_standby_for_invalid_fscid with vstart_runner ([pr#14272](#), John Spray)
- cephfs,tests: qa: handle SSHException in logrotate ([pr#13359](#), John Spray)
- cephfs,tests: qa, mds: add checks for fragmentation, and enable it by default ([issue#16523](#), [pr#13862](#), John Spray, John Spray)
- cephfs,tests: qa: misc cephfs test improvements ([issue#20131](#), [pr#15411](#), John Spray)
- cephfs,tests: qa: re-enable ENOSPC tests for kclient ([issue#19550](#), [pr#14396](#), John Spray)
- cephfs,tests: qa: silence spurious insufficient standby health warnings ([pr#15035](#), Patrick Donnelly)
- cephfs,tests: qa: silence upgrade test failure ([issue#19934](#), [pr#15126](#), Patrick Donnelly)
- cephfs,tests: qa: simplify TestJournalRepair ([pr#15096](#), John Spray)
- cephfs,tests: qa/tasks: force umount during kclient teardown ([issue#18663](#), [pr#13099](#), John Spray)
- cephfs,tests: qa: Tidy up fs/ suite ([pr#14575](#), John Spray)
- cephfs,tests: qa/vstart_runner: amend ps invocation ([pr#14254](#), Ilya Dryomov)
- cephfs,tests: qa: whitelist another fullness log message ([issue#19253](#), [pr#14221](#), John Spray)
- cephfs,tests: tasks/cephfs: tear down on mount() failure ([pr#13282](#), John Spray)
- cephfs: tools/cephfs: remove apply mode of cephfs-journal-tool ([pr#15715](#), John Spray)
- cephfs: tools/cephfs: set dir_layout when injecting inodes ([issue#19406](#), [pr#14234](#), John Spray)
- ceph-volume: use unique logical volumes ([pr#17208](#), Alfredo Deza)
- cleanup: .gitignore: exclude rpm files ([pr#15745](#), Leo Zhang)
- cleanup: Move code from .h into .cc ([pr#12737](#), Amir Vadai)
- cleanup: resolve compiler warnings ([pr#13236](#), Adam C. Emerson)
- cleanup: src: put-to operator function - const input cleanup ([issue#3977](#), [pr#15364](#), Jos Collin)
- cmake: add "container" to required boost components ([pr#14850](#), Kefu Chai)
- cmake: Add -finstrument-functions flag to OSD code ([pr#15055](#), Mohamad Gebai)
- cmake: add RGW and MDS to libcephd ([pr#12345](#), Bassam Tabbara)
- cmake: Add simple recursive ctags target for Ceph source only ([pr#14334](#), Kefu Chai, Dan Mick)
- cmake: align cmake names of library packages ([issue#19853](#), [pr#14951](#), Nathan Cutler)
- cmake: Allow tests to build without NSS ([pr#13315](#), Daniel Grynewicz)
- cmake: build boost as an external project ([pr#15376](#), Kefu Chai)
- cmake: build tracepoint libraries for vstart target ([pr#14354](#), Mohamad Gebai)
- cmake: check the existence of gperf before using it ([pr#15164](#), Kefu Chai)
- cmake: cleanup the use of udev and blkid in target_link_lib() ([pr#12811](#), Willem Jan Withagen)
- cmake: disable -fvar-tracking-assignments for config.cc ([pr#16695](#), Kefu Chai)
- cmake: disable mallinfo for jemalloc ([pr#12469](#), Bassam Tabbara)
- cmake: do not add dependencies to INTERFACE library on cmake < 3.3 ([pr#15813](#), Kefu Chai)
- cmake: do not compile crush twice ([pr#14725](#), Kefu Chai)
- cmake: do not link libcommon against some libs ([pr#15340](#), Willem Jan Withagen)
- cmake: do not try to add submodule to exclude list if .git is not around ([pr#14495](#), Kefu Chai)
- cmake: enable cross-compilation of boost ([issue#18938](#), [pr#14881](#), Kefu Chai)
- cmake: exclude *.css while generating ctags ([pr#15663](#), Leo Zhang)
- cmake: explicitly call find_package(PythonInterp) first to fix build err ([pr#12385](#), Yixun Lan)
- cmake: fix boost components for WITH_SYSTEM_BOOST ([pr#15160](#), Bassam Tabbara)
- cmake: Fix broken async/rdma compilation since move to libceph-common ([pr#13122](#), Oren Duer)
- cmake: fix broken RDMA compilation after merge PR #12878 ([pr#13186](#), Oren Duer)
- cmake: fix hard coded boost python lib ([pr#12480](#), John Coyle)
- cmake: fix rpath on shared libraries and binaries targets ([pr#12927](#), Ricardo Dias)
- cmake: fix the build with -DWITH_ZFS=ON ([pr#15907](#), Kefu Chai)
- cmake: fix the linked lib reference of unittest_rgw_crypto ([pr#14869](#), Willem Jan Withagen)
- cmake: improved build speed by 5x when using ccache ([pr#15147](#), Bassam Tabbara)
- cmake: kill duplicated cmake commands ([pr#14948](#), liuchang0812)
- cmake: link against fcgi only if enabled ([pr#15425](#), Yao Zongyou)
- cmake: link ceph-{mgr,mon,mds,osd} against libcommon statically ([pr#12878](#), Kefu Chai)
- cmake: link consumers of libclient with libcommon ([issue#18838](#), [pr#13394](#), Kefu Chai)

- cmake: misc fixes for build on i386 ([pr#15516](#), James Page)
- cmake: pass -d0 to b2 if not CMAKE_VERBOSE_MAKEFILE ([pr#14651](#), Kefu Chai)
- cmake: remove Findpciaccess.cmake ([pr#12776](#), optimistzyz)
- cmake: Rewrite HAVE_BABELTRACE option to WITH ([pr#15305](#), Willem Jan Withagen)
- cmake: rgw: do not link against boost in a wholesale ([pr#15347](#), Nathan Cutler, Kefu Chai)
- cmake: search for Keyutils in default paths ([pr#12769](#), Pascal Bach)
- cmake: search for nspr include files for both suffixes: nspr4 and nspr ([issue#18535](#), [pr#12939](#), John Lin)
- cmake: should not compile crc32c_ppc.c on intel arch ([pr#14423](#), Kefu Chai)
- cmake: simplify find_package jemalloc ([pr#12468](#), Bassam Tabbara)
- cmake: support for external rocksdb ([pr#12467](#), Bassam Tabbara)
- cmake: support optional argument for overriding default ctag excludes ([pr#14379](#), Kefu Chai)
- cmake: turn libcommon into a shared library ([pr#12840](#), Kefu Chai)
- cmake: use CMAKE_INSTALL_INCLUDEDIR ([pr#16483](#), David Disseldorp)
- cmake: workaround ccache issue with .S assembly files ([pr#15142](#), Bassam Tabbara)
- common: add ceph::size() ([pr#15181](#), Kefu Chai)
- common: add override in common and misc ([issue#18922](#), [pr#13443](#), liuchang0812)
- common: add override in header file ([pr#13774](#), liuchang0812)
- common: add override in msg subsystem ([pr#13771](#), liuchang0812)
- common: auth: Enhancement for the supported auth methods ([pr#12937](#), Dave Chen)
- common: auth/RotatingKeyRing: use std::move() to set secrets ([pr#15866](#), Kefu Chai)
- common: avoid statically allocating configuration options ([issue#20869](#), [pr#16735](#), Jason Dillaman)
- common: Better handling for missing/inaccessible ceph.conf files ([issue#19658](#), [pr#14757](#), Dan Mick)
- common: bufferlist: cleanup semantical wrong for bufferlist::append ([pr#12247](#), Yankun Li)
- common: buffer: silence unused var warning on FreeBSD ([pr#16452](#), Willem Jan Withagen)
- common: ceph_osd: remove client message cap limit ([pr#14944](#), Haomai Wang)
- common: ceph: wait for maps before doing 'ceph tell ... help' ([issue#20113](#), [pr#16756](#), Sage Weil)
- common: cls/log/cls_log.cc: reduce logging noise ([issue#19835](#), [pr#14879](#), Willem Jan Withagen)
- common: cls: optimize header file dependency ([pr#15165](#), Brad Hubbard, Xiaowei Chen)
- common: cmdparse: more constness ([pr#15023](#), Kefu Chai)
- common: common/admin_socket: add config for admin socket permission bits ([pr#11684](#), runsisi)
- common: common/admin-socket: fix potential buffer overflow ([pr#12518](#), Yunchuan Wen)
- common: common/auth: add override in headers ([pr#13692](#), liuchang0812)
- common: common/BackTrace: add operator<< ([pr#9028](#), Kefu Chai)
- common: common/BackTrace: demangle on FreeBSD also ([pr#12992](#), Kefu Chai)
- common: common/buffer: close pipe fd if set nonblocking fails ([pr#12828](#), donglinpeng)
- common: common/buffer: off-by-one error in max iov length blocking ([issue#20907](#), [pr#16803](#), Dan Mick)
- common: common/ceph_context.cc: Use CEPH_DEV to reduce logfile noise ([pr#10384](#), Willem Jan Withagen)
- common: common/ceph_context: 'config diff get' option added ([pr#10736](#), Daniel Oliveira)
- common: common/ceph_context: fewer warnings about experimental features ([pr#14170](#), Sage Weil)
- common: common/ceph_context: fix leak of registered commands on exit ([pr#15302](#), xie xingguo)
- common: common/ceph_context: Show clear message if all features are enabled ([pr#12676](#), Dave Chen)
- common: common/cmdparse.cc: remove unused variable 'argnum' in dump_cmd_to_json() ([pr#16862](#), Luo Kexue)
- common: common/common_init: disable default dout logging for UTILITY_NODOUT too ([issue#20771](#), [pr#16578](#), Sage Weil)
- common: common/config: Add /usr/local/etc/ceph to default paths ([pr#14797](#), Willem Jan Withagen)
- common: common/config: eliminate config_t::set_val unsafe option ([issue#19106](#), [pr#13687](#), liuchang0812)
- common: common/config: fix return type of string::find and use string::npos ([pr#9924](#), Yan Jun)
- common: common,config: OPT_FLOAT and OPT_DOUBLE output format in config show ([issue#20104](#), [pr#15647](#), Yanhu Cao)
- common: common/config_opt: remove unused config ([pr#15874](#), alex.wu)
- common: common/config_opts: drop unused opt ([pr#15876](#), Yanhu Cao)
- common: common/config_opts.h: FreeBSD timing changed due to no SO_REUSEADDR ([pr#12594](#), Willem Jan Withagen)
- common: common/config_opts.h: Remove deprecated osd_compact_level_db_on_mount option ([issue#19318](#), [pr#14059](#), Vikhyat Umrao)
- common: common/config_opts.h: remove obsolete configuration option ([pr#12659](#), Li Wang)
- common: common/config_opts: Set the HDD throttle cost to 1.5M ([pr#14808](#), Mark Nelson)
- common: common/EventTrace: fix compiler warning ([pr#13659](#), Jianpeng Ma)
- common: common/Finisher: fix uninitialized variable warning ([pr#14958](#), Piotr Dałek)
- common: common/freebsd_errno.cc: fixed again a stupid typo ([pr#15742](#), Willem Jan Withagen)
- common: common/interval_set: return int64_t for size() ([pr#12898](#), Xinze Chi)
- common: common/iso_8601.cc: Make return expression Clang compatible ([pr#15336](#), Willem Jan Withagen)
- common: common/LogEntry: include EntityName in log entries ([pr#15395](#), Sage Weil)
- common: common/Mutex.cc: fixed the error in comment ([pr#16214](#), Pan Liu)
- common: common/options: refactors to set the properties in a more structured way ([pr#16482](#), Kefu Chai)
- common: common,osdc: remove atomic_t completely ([pr#15562](#), Kefu Chai)

- common: common/perf_counters: add average time for PERFCOUNTER_TIME (pr#15478, xie xingguo)
- common: common/perf_counters: fix race condition with atomic variables (pr#14227, J. Eric Ivancich)
- common: common/perf_counters: make schema more friendly and update docs (pr#14933, Sage Weil)
- common: common/perf_counters.: Remove unnecessary judgment (pr#10407, zhang.zezhu)
- common: common/simple_spin: use __ppc_yield() on all powerpc archs (pr#14310, Kefu Chai)
- common: common,test: migrate atomic_t to std::atomic (pr#14866, Jesse Williamson)
- common: common/Timer: do not add event if already shutdown (issue#20432, pr#16201, Kefu Chai)
- common: common/WorkQueue: use threadpoolname + threadaddr for heartbeat_han... (pr#16563, huangjun)
- common: common/xmlformatter: turn on underscored and add unittest (pr#12916, liuchang0812)
- common: compressor/zlib: remove g_ceph_context/g_conf from compressor plugin (pr#16245, Casey Bodley)
- common: compressor/zstd: add zstd compression plugin (pr#13075, Kefu Chai, Sage Weil)
- common: config: Improve warning for unobserved value (issue#18424, pr#12855, Brad Hubbard)
- common: config_opt: use bool instead of int for the default value of filestore_debug_omap_check (pr#15651, Leo Zhang)
- common,core: ceph_test_rados_api_misc: fix LibRadosMiscConnectFailure.ConnectFailure retry (issue#19901, pr#15522, Sage Weil)
- common: core/common: Fix ENODATA for FreeBSD with compat.h (issue#19883, pr#15685, Willem Jan Withagen)
- common,core: common, osd, tools: Add histograms to performance counters (pr#12829, Bartłomiej Święcki)
- common,core: common/pick_address.cc: Copy public_netw to cluster_netw if cluster empty (pr#12929, Willem Jan Withagen)
- common,core: common/TracepointProvider: add assert if dlopen error (pr#13430, Jianpeng Ma)
- common,core: common/TrackedOp: make TrackedOp::reset_desc() safe (issue#19110, pr#13702, Sage Weil)
- common,core: mempool: put bloom_filter in mempool (pr#13009, Sage Weil)
- common,core: osd,mds,mgr: do not dereference null rotating_keys (issue#20667, pr#16455, Sage Weil)
- common,core: osd,osdc: pg and osd-based backoff (pr#12342, Sage Weil)
- common,core: osd/OSDMap: make osd_state 32 bits wide (pr#15390, Sage Weil)
- common,core: osd/OSDMap: replace require_osds flags with a single require_osd_release field (pr#15068, Sage Weil)
- common,core: osd/OSDMap: replace string-based min_compat_client with a CEPH_RELEASE uint8_t (pr#15351, Sage Weil)
- common,core: osd/osd_types: add flag name (IGNORE_REDIRECT) (pr#15795, Myoungwon Oh)
- common,core: rados: we need to get the latest osdmap when pool does not exists (pr#13289, song baisen)
- common,core,tests: Wip cppcheck errors (pr#14446, Brad Hubbard)
- common: crc32c: include acconfig.h to fix ceph_crc32c_aarch64() (pr#15515, Kefu Chai)
- common: crush/CrushWrapper: fix has_incompat_choose_args (pr#15218, Sage Weil)
- common: crush/CrushWrapper: fix has_incompat_choose_args() (pr#15244, Sage Weil)
- common: crypto: cleanup NSPR in main thread (pr#14801, Kefu Chai)
- common: delete unused conf "filestore_debug_disable_sharded_check" (pr#13051, Chuanhong Wang)
- common: denc: add encode/decode for basic_sstring (pr#15135, Kefu Chai, Casey Bodley)
- common: do not print error when asok is closed (pr#14022, Patrick Donnelly)
- common: fix building against libcryptopp (pr#14949, Shengjing Zhu)
- common: Fix clang compilation (pr#13335, Bartłomiej Święcki)
- common: Fix heap buffer overflow in do_request (issue#19393, pr#14123, Brad Hubbard)
- common: fix lockdep vs recursive mutexes (pr#9940, Adam Kupczyk)
- common: fix log warnings (pr#16056, xie xingguo)
- common: fix Option set_long_description (pr#16668, Yan Jun)
- common: fix segfault in public IPv6 addr picking (issue#19371, pr#14124, Fabian Grünbichler)
- common: fix that \$host always expands to localhost instead of actual hostname (issue#11081, pr#12998, liuchang0812)
- common: fix typo in option of rados_mon_op_timeout's comment (pr#15681, Leo Zhang)
- common: Fix unused variable references warnings (pr#14790, Willem Jan Withagen)
- common: follow up to new options infrastructure (pr#16527, John Spray)

common: Forward-declare container I/O overloads

- common: get_process_name: use getprogname on bsd systems (pr#15338, Mykola Golub)
- common: get rid of "warning: ignoring return value of 'strerror_r'" (pr#12775, xie xingguo)
- common: global: we need to handle the init_on_startup return value when global_init (pr#13018, song baisen)
- common: Implements simple_spin_t in terms of std::atomic_flag (pr#14370, Jesse Williamson)
- common: Improved CRC calculation for zero buffers (pr#11966, Adam Kupczyk)
- common: include/ceph_features.h uses uint64_t, which is in sys/types.h (pr#13339, Willem Jan Withagen)
- common: include/denc: improvements (pr#12626, Adam C. Emerson)
- common: include/denc, kv: silence gcc warnings (pr#13458, Kefu Chai)
- common: include/denc: remove nullptr runtime magic boundedness check (pr#13889, Sage Weil)
- common: include/lru.h: add const to member functions (pr#15408, yonghengdexin735)
- common: include/mempool: fix typo in comments (pr#12772, huangjun)
- common: include/rados: Fix typo in rados_ioctlx_cct() doc (pr#15220, Jos Collin)
- common: include: Redo some includes for FreeBSD (issue#19883, pr#15337, Willem Jan Withagen)
- common: initialize array in struct BackTrace (pr#15864, Jos Collin)

- common: initialize _hash in LogEntryKey() ([pr#15615](#), Jos Collin)
- common: int_types.h: remove hacks to workaround old systems ([pr#15069](#), Kefu Chai)
- common: kv: resolve a crash issue in ~LevelDBStore() ([pr#16553](#), wumingqiao)
- common: librados,libradosstriper,test: migrate atomic_t to std::atomic (baragon) ([pr#14658](#), Jesse Williamson)
- common: librados, osd: clang fixes ([pr#13768](#), Kefu Chai)
- common: libradosstriper: Add example code ([pr#15350](#), Logan Blyth)
- common: libradosstriper: fix format injection vulnerability ([issue#20240](#), [pr#15674](#), Stan K)
- common: libradosstriper: fix MultiAioCompletion leaks on failure ([pr#15471](#), Kefu Chai)
- common: make attempts of auth rotating configurable ([pr#12563](#), xie xingguo)
- common: Make spinlock delay more conventional ([pr#14248](#), Brad Hubbard)
- common: mempool: improve dump; fix buffer accounting bugs ([pr#15403](#), Sage Weil)
- common: messages: fix return type name of MOSDMap ([pr#14382](#), Leo Zhang)
- common: mgr/PyFormatter: implement dump_format_va ([pr#15634](#), Sage Weil)
- common: misc cleanups in common, global, os, osd submodules ([pr#16321](#), Yan Jun)
- common: misc fixes detected by crypto shutdown assert ([pr#12925](#), Sage Weil)
- common,mon: crush,mon: add weight-set introspection and manipulation commands ([pr#16326](#), Sage Weil)
- common,mon: messenger,client,compressor: migrate atomic_t to std::atomic ([pr#14657](#), Jesse Williamson)
- common: mon/MonClient: scale backoff interval down when we have a healthy mon session ([issue#20371](#), [pr#16576](#), Kefu Chai, Sage Weil)
- common,mon: mon,crush: add 'osd crush swap-bucket' command ([pr#15072](#), Sage Weil)
- common: msg/async: add assert of ms_async_op_threads > 0 ([pr#15629](#), linbing)
- common: msg/async: assert if compiled code doesn't support the configured ms ([pr#12559](#), Avner BenHanoch)
- common: msg/async: fix crash that writing char to nonblock-fd gets EAGAIN in EventCenter::wakeup ([pr#13822](#), liuchang0812)
- common: msg/async: make recv_stamp more precise ([pr#15810](#), Pan Liu)
- common: msg/async/rdma: Add fork safe on RDMA ([pr#13740](#), Sarit Zubakov)
- common: msg/async/rdma: clean line endings ([pr#12688](#), Adir Lev)
- common: msg/async/rdma: Remove compilation warning ([pr#13142](#), Sarit Zubakov)
- common: msg/async/rdma: rename chunk_size to buffer_size ([pr#13666](#), Adir Lev)
- common: msg/async/rdma: Support for RoCE v2 and SL ([pr#12556](#), Oren Duer)
- common: msg/async/rdma: Update fix broken compilation ([pr#13940](#), Sarit Zubakov)
- common: msg/async: return right away in NetHandler::set_priority() if not supported ([pr#14795](#), Kefu Chai)
- common: msg/simple: call clear_pipe in wait() shutdown path ([issue#15784](#), [pr#12633](#), Sage Weil)
- common: msg/SimpleMessenger: error out misplace in set_socket_options ([pr#13961](#), wangzhengyong)
- common: msg/simple/Pipe: support IPv6 QoS ([issue#18887](#), [pr#13370](#), Robin H. Johnson)
- common: .organizationmap: Updated authors ([pr#14360](#), Jos Collin)
- common: osdc: fix osdc_osd_seesion perf counter ([pr#13478](#), Xiaoxi Chen)
- common: osdc/Objecter: fix bugs in explicit naming of op spg_t ([pr#13534](#), Sage Weil)
- common: osdc/Objecter: fix pool dne corner case ([issue#19552](#), [pr#14901](#), Sage Weil)
- common: osdc/Objecter: handle command target that goes down ([issue#19452](#), [pr#14302](#), Sage Weil)
- common: osdc/Objecter: release message if it's not handled ([issue#19741](#), [pr#15890](#), Kefu Chai)
- common: osdc/Objecter: resend RWORDERED ops on full ([issue#19133](#), [pr#13759](#), Sage Weil)
- common: osd/OSDMap: fix feature commit comment ([pr#15056](#), Sage Weil)
- common: osd/OSDMap: get_previous_up_osd_before() may run into endless loop ([pr#12976](#), Mingxin Liu)
- common: osd/OSDMap: print require_osd_release ([pr#15974](#), Sage Weil)
- common: osd/osd_types: clean up OSDOp printers ([pr#12980](#), Sage Weil)
- common: Passing null pointer option_name to operator << in md_config_t::parse_option() ([pr#15881](#), Jos Collin)
- common,performance: buffer: allow buffers to be accounted in arbitrary mempools ([pr#15352](#), Sage Weil)
- common,performance: common/Finisher: batch handle perfcounter && only send signal when waiter existed ([pr#14363](#), Jianpeng Ma)
- common,performance: crc32c: Add ppc64le fast zero optimized assembly ([pr#15100](#), Andrew Solomon)
- common,performance: inline_memory: optimized mem_is_zero for non-x64 ([pr#15307](#), Piotr Dałek)
- common,performance: kv/rocksdb: supports SliceParts interface ([pr#15058](#), Haomai Wang)
- common,performance: osd/OSDMap: make pg_temp more efficient ([pr#15291](#), Sage Weil)
- common: possible lockdep false alarm for ThreadPool lock ([issue#18819](#), [pr#13258](#), Mykola Golub)
- common: prevent unset_dumpable from generating warnings ([pr#16462](#), Willem Jan Withagen)
- common: rados: allow "rados purge" to delete objects when osd is full ([pr#13814](#), Pan Liu)
- common: rados: more info added to pool deletion error ([issue#19400](#), [pr#14235](#), Vedant Nanda)
- common,rbd: osdc/Objecter: unify disparate EAGAIN handling paths into one ([pr#16627](#), Sage Weil)
- common,rbd,rgw: common/escape: do not escape / in json ([pr#14130](#), Sage Weil)
- common,rbd,rgw: common/rgw/rbd: remove some unused variables ([pr#16690](#), Luo Kexue)
- common,rdma: msg/async/rdma: automatically set RDMAV_HUGE PAGES_SAFE according to conf ([pr#15755](#), DanielBar-On)
- common,rdma: msg/async/rdma: check ulimit ([pr#13655](#), Sarit Zubakov, Adir Lev)
- common,rdma: msg/async/rdma: Introduce Device.{cc,h} ([pr#14001](#), Amir Vadai)

- common,rdma: msg/async/rdma: Introduce RDMAConnMgr + Debug prints ([pr#14201](#), Amir Vadai)
- common,rdma: msg/async/rdma: Move resource handling to Device ([pr#14088](#), Sarit Zubakov, Amir Vadai)
- common,rdma: msg/async/rdma: RDMA-CM Initialize device on first connect ([pr#14179](#), Amir Vadai)
- common,rdma: msg/async/rdma: reduce number of rdma rx/tx buffers ([pr#13190](#), Adir Lev)
- common,rdma: msg/async/rdma: use lists properly ([pr#15908](#), Adir lev, Adir Lev)
- common: remove config opt conversion utility ([pr#16480](#), John Spray)
- common: remove n on clog messages ([pr#13794](#), Sage Weil)
- common: Remove redundant includes - 2 ([issue#19883](#), [pr#15169](#), Jos Collin)
- common: Remove redundant includes - 3 ([issue#19883](#), [pr#15204](#), Jos Collin)
- common: Remove redundant includes - 4 ([issue#19883](#), [pr#15251](#), Jos Collin)
- common: Remove redundant includes - 5 ([issue#19883](#), [pr#15267](#), Jos Collin)
- common: Remove redundant includes - 6 ([issue#19883](#), [pr#15299](#), Jos Collin)
- common: Remove redundant includes ([issue#19883](#), [pr#15003](#), Brad Hubbard)
- common: Remove redundant includes ([issue#19883](#), [pr#15019](#), Brad Hubbard)
- common: Remove redundant includes ([issue#19883](#), [pr#15042](#), Brad Hubbard)
- common: Remove redundant includes ([issue#19883](#), [pr#15086](#), Jos Collin)
- common: remove useless parameter ([pr#14096](#), baiyanchun)
- common: Revamp config option definitions ([issue#20627](#), [pr#16211](#), John Spray, Kefu Chai, Sage Weil)
- common,rgw: cls/refcount: store and use list of retired tags ([issue#20107](#), [pr#15673](#), Yehuda Sadeh)
- common: src/common/ceph_string: stringify new osd states ([pr#15751](#), xie xingguo)
- common: src/common: change last_work_queue to next_work_queue ([pr#14738](#), Pan Liu)
- common: support s390 and unknown architectures in spin-wait loop ([issue#19492](#), [pr#14337](#), Nathan Cutler)
- common,tests: ceph_test_rados_api_c_read_operations: do not assert per-op rval is correct ([issue#19518](#), [pr#16196](#), Sage Weil)
- common,tests: ceph_test_rados_api_list: more fix LibRadosListNP.ListObjectsError ([issue#19963](#), [pr#15138](#), Sage Weil)
- common,tests: test: Make screencandy optional for FreeBSD ([pr#15444](#), Willem Jan Withagen)
- common: the latency dumped by “ceph osd perf” is not real ([issue#20749](#), [pr#16512](#), Pan Liu)
- common,tools: osdmaptool: show all the pg map to osds info ([pr#9419](#), song baisen)
- common: tracing: Fix handle leak in TracepointProvider ([pr#12652](#), Brad Hubbard)
- common: tracing: fix segv ([issue#18576](#), [pr#14304](#), Anjaneya Chagam)
- common: Update the error string when res_nsearch() or res_search() fails ([pr#15878](#), huanwen ren)
- common: use ref to avoid unnecessary memory copy ([issue#19107](#), [pr#13689](#), liuchang0812)
- common: use std::move() for better performance ([pr#16620](#), Xinying Song)
- common: xio: migrate atomic_t to std::atomic<> ([pr#15230](#), Jesse Williamson)
- compressor: conditionalize on HAVE_LZ4 ([pr#17174](#), Kefu Chai)
- compressor: fix Mutex::Locker used is not correct ([pr#13935](#), hechuang)
- compressor: zlib: fix plugin for non-Intel arches ([pr#14947](#), Dan Mick)
- core: auth: ‘ceph auth import -i’ overwrites caps, if caps are not specified ([issue#18932](#), [pr#13468](#), Vikhyat Umrao)
- core: auth: Remove unused function in AuthSessionHandler ([pr#16666](#), Luo Kexue)
- core: ceph: allow ‘-’ with -i and -o for stdin/stdout ([pr#16359](#), Sage Weil)
- core: ceph-create-keys: Add connection timeouts ([pr#11995](#), Owen Synge)
- core: ceph-dencoder: Silence coverity CID 1412579 ([pr#15744](#), Brad Hubbard)
- core: ceph-detect-init: Add docker detection ([pr#13218](#), Guillaume Abrioux)
- core: ceph-disk: Adding retry loop in get_partition_dev() ([pr#14275](#), Erwan Velu)
- core: ceph-disk/ceph_disk/main.py: fix calling of the bsdrc init scripts ([pr#14476](#), Willem Jan Withagen)
- core: ceph-disk/ceph_disk/main.py: Replace ST_ISBLK() test by is_diskdevice() ([pr#15587](#), Willem Jan Withagen)
- core: ceph-disk: ceph-disk on FreeBSD should not use mpath-code ([pr#14837](#), Willem Jan Withagen)
- core: ceph-disk: dmcrypt activate must use the same cluster as prepare ([issue#17821](#), [pr#13573](#), Loic Dachary)
- core: ceph-disk: dmcrypt cluster must default to ceph ([issue#20893](#), [pr#16776](#), Loic Dachary)
- core: ceph-disk: do not setup_statedir on trigger ([issue#19941](#), [pr#15410](#), Loic Dachary)
- core: ceph-disk: enable directory backed OSD at boot time ([issue#19628](#), [pr#14546](#), Loic Dachary)
- core: ceph-disk: Fix getting wrong group name when -setgroup in bluestore ([issue#18955](#), [pr#13457](#), craigchi)
- core: ceph-disk: FreeBSD changes to get it working and passing tests ([pr#12086](#), Willem Jan Withagen)
- core: ceph-disk: implement prepare -no-locking ([pr#14728](#), Dan van der Ster, Loic Dachary)
- core: ceph_disk/main.py: Allow FreeBSD zap a OSD disk ([pr#15642](#), Willem Jan Withagen)
- core: ceph-disk,osd: add support for crush device classes ([issue#19513](#), [pr#14436](#), Loic Dachary)
- core: ceph-disk: Populate mount options when running “list” ([issue#17331](#), [pr#14293](#), Brad Hubbard)
- core: ceph-disk: Reporting /sys directory in get_partition_dev() ([pr#14080](#), Erwan Velu)
- core: ceph-disk: Revert “Revert “change get_dmcrypt_key test to support different cluster name”” ([pr#13600](#), Loic Dachary)
- core: ceph-disk: separate ceph-osd –check-needs-* logs ([issue#19888](#), [pr#15016](#), Loic Dachary)
- core: ceph-disk: set the default systemd unit timeout to 3h ([issue#20229](#), [pr#15585](#), Loic Dachary)
- core: ceph-disk: support osd new ([pr#15432](#), Loic Dachary, Sage Weil)
- core: ceph-disk: Write 10M to all partitions before zapping ([issue#18962](#), [pr#13766](#), Wido den Hollander)
- core: ceph: do not throw TypeError on connection failure ([pr#13268](#), Kefu Chai)

- core: ceph.in: Fix couple of minor issues on the messages ([pr#12797](#), Dave Chen)
- core: ceph-objectstore-tool: do not populate snapmapper with missing clones ([issue#19943](#), [pr#15787](#), Sage Weil)
- core: ceph-osd: fix auto detect which objectstore is currently running ([issue#20865](#), [pr#16717](#), Yanhu Cao)
- core: ceph-osd: -flush-journal: sporadic segfaults on exit ([issue#18820](#), [pr#13311](#), Alexey Sheplyakov)
- core: client/SyntheticClient.cc: Fix warning in random_walk ([issue#19445](#), [pr#14308](#), Brad Hubbard)
- core: cls/timeindex: clean up cls_timeindex_client.h|cc ([pr#13987](#), Shinobu Kinjo)
- core: common/options: remove mon_warn_osd_usage_min_max_delta from options.cc too ([pr#16488](#), Sage Weil)
- core: common/TrackedOp: allow dumping historic ops sorted by duration ([pr#14050](#), Piotr Dałek)
- core: compressor: add LZ4 support ([pr#15434](#), Haomai Wang)
- core: compressor: optimize header file dependency ([pr#15187](#), Brad Hubbard, Xiaowei Chen)
- core: Context: C_ContextsBase: delete enclosed contexts in dtor ([issue#20432](#), [pr#16159](#), Kefu Chai)
- core: crush/CrushWrapper: chooseargs encoding fix ([pr#15984](#), Ilya Dryomov)
- core: crush/CrushWrapper: make get_immediate_parent[_id] ignore per-class shadow hierarchy ([issue#20546](#), [pr#16221](#), Sage Weil)
- core: crush, mon: make jewel the lower bound for client/crush compat for new clusters ([pr#15370](#), Sage Weil)
- core: erasure-code: optimize header file dependency ([pr#15172](#), Brad Hubbard, Xiaowei Chen)
- core: erasure-code: Remove duplicate of isa-l files ([pr#15372](#), Ganesh Mahalingam)
- core: erasure-code: sync jerasure/gf-complete submodules ([pr#14424](#), Loic Dachary)
- core: filestore: migrate atomic_t to std::atomic<> ([pr#15228](#), Jesse Williamson)
- core: Give requested scrub work a higher priority ([issue#15789](#), [pr#14488](#), David Zafman)
- core: global: start removing g_ceph_context ([pr#12149](#), Adam C. Emerson)
- core: HashIndex.cc: add compat.h for ENODATA ([pr#16697](#), Willem Jan Withagen)
- core: include/denc: add {encode,decode}_nohead for denc_traits<basic_string> ([issue#18938](#), [pr#14099](#), Kefu Chai)
- core: include/mempool.h: fix Clangs complaint about types ([pr#13523](#), Willem Jan Withagen)
- core: include/types.h, introduce host_to_ceph_errno ([pr#15496](#), Willem Jan Withagen)
- core: Install Pecan for FreeBSD ([pr#15610](#), Willem Jan Withagen)
- core: introduce (and fix) code to pass errno to other OSes ([pr#15495](#), Willem Jan Withagen)
- core: introduce DirectMessenger ([pr#14755](#), Casey Bodley, Matt Benjamin)
- core: kv/RocksDBStore: abort if rocksdb EIO, don't return incorrect result ([pr#15862](#), Haomai Wang)
- core: kv/RocksDBStore: use vector instead of VLA for holding slices ([pr#16615](#), Kefu Chai)
- core: libradosstriper: Initialize member variable m_writeRc in WriteCompletionData ([pr#16780](#), amitkuma)
- core: luminous: Improve size scrub error handling and ignore system attrs in xattr checking ([issue#21051](#), [issue#18836](#), [issue#20243](#), [pr#17196](#), David Zafman)
- core: luminous: Include front/back interface names in OSD metadata ([issue#21048](#), [issue#20956](#), [pr#17193](#), John Spray)
- core: luminous: mon: bug in functon reweight_by_utilization ([issue#21079](#), [issue#20970](#), [pr#17198](#), xie xingguo)
- core: luminous: mon: "ceph osd crush rule rename" support ([pr#17260](#), xie xingguo)
- core: luminous: multisite: FAILED assert(prev_iter != pos_to_prev.end()) in RGWMetaSyncShardCR::collect_children() ([issue#21097](#), [issue#20906](#), [pr#17234](#), Casey Bodley)
- core: luminous: osd: osd_scrub_during_recovery only considers primary, not replicas ([issue#18206](#), [issue#21077](#), [pr#17195](#), David Zafman)
- core: luminous: src/common/LogClient.cc: 310: FAILED assert(num_unsent <= log_queue.size()) ([issue#20965](#), [issue#18209](#), [pr#17197](#), Sage Weil)
- core: make the conversion from wire error to host OS work ([pr#15780](#), Willem Jan Withagen)
- core: Merge pull request #16755 from ivancich/wip-pull-new-dmclock ([pr#16922](#), Gregory Farnum)
- core: messages: default-initialize MOSDPGRecoveryDelete[Reply] members ([pr#16584](#), Greg Farnum)
- core: messages: Initialize members in MMDSTableRequest ([pr#16810](#), amitkuma)
- core: messages: Initialize member variables ([pr#16819](#), amitkuma)
- core: messages: Initialize member variables ([pr#16839](#), amitkuma)
- core: messages: Initializing member variable in MMDSCacheRejoin ([pr#16791](#), amitkuma)
- core: messages/MOSDOP: fix pg_t decoding for version < 7 decoding ([issue#19005](#), [pr#13537](#), Sage Weil)
- core: messages/MOSDPGTrim: add the missed HEAD_VERSION AND COMPAT_VERSION ([issue#18266](#), [pr#12517](#), huangjun)
- core: messages/MOSDPing.h: drop unused fields ([pr#15843](#), Piotr Dałek)
- core: messages/MOSDPing: initialize MOSDPing padding ([issue#20323](#), [pr#15714](#), Sage Weil)
- core: messages/MOSDSubOp: Make encode_payload can be reentrant ([pr#12654](#), Haomai Wang)
- core: messages: remove compat cruft ([pr#14475](#), Sage Weil)
- core: mgr/MgrClient: do not attempt to access a global variable for config ([pr#16544](#), Jason Dillaman)
- core: mgr/MgrClient: use unique_ptr for MgrClient::session ([issue#19097](#), [pr#13685](#), Kefu Chai)
- core,mgr: mgr/DaemonServer: stop spamming log with pg stats ([pr#15487](#), Sage Weil)
- core,mgr: mgr,librados: service map ([pr#15858](#), Yehuda Sadeh, John Spray, Sage Weil)
- core,mgr,mon: mgr,mon: enable/disable mgr modules via 'ceph mgr module ...' commands ([pr#15958](#), Sage Weil)
- core,mgr,mon: mon,mgr: tag some commands for ceph-mgr ([pr#13617](#), Sage Weil)
- core,mgr,mon: mon/PGMap: fix osd_epoch update when removing osd_stat ([issue#20208](#), [pr#15573](#), Sage Weil)
- core,mgr: mon/PGMap: slightly better debugging around pgmap updates ([pr#15820](#), Sage Weil)
- core,mgr,tests: qa: flush out monc's dropped msgs on msgr failure injection ([issue#20371](#), [pr#16484](#), Joao Eduardo Luis)

- core,mgr,tests: qa/suites/rados/rest: test restful mgr module ([pr#15604](#), Sage Weil)
- core: misc: SCA fixes ([pr#14426](#), Danny Al-Gaaf)
- core,mon: crush, mon: simplify device class manipulation commands ([pr#16388](#), xie xingguo)
- core: mon,mgr: fix “ceph osd df”, add some tools to find untested commands ([issue#20256](#), [pr#15675](#), Greg Farnum)
- core: mon/MonClient: discard stray messages from non-acitve conns ([issue#19015](#), [pr#13656](#), Kefu Chai)
- core: mon/MonClient: don’t return zero global_id ([issue#19134](#), [pr#13853](#), “Yan, Zheng”, Kefu Chai)
- core: mon/MonClient: hunt monitors in parallel ([issue#16091](#), [pr#11128](#), Steven Dieffenbach, Kefu Chai)
- core: mon/MonClient: persist global_id across re-connecting ([issue#18968](#), [pr#13550](#), Kefu Chai)
- core: mon/MonClient: respect the priority in SRV RR ([issue#5249](#), [pr#15964](#), Kefu Chai)
- core,mon: mon/LogMonitor: ‘log last’ command ([pr#15497](#), Sage Weil)
- core: mon/MonmapMonitor: use `_func_` instead of hard code function name ([pr#16037](#), Yanhu Cao)
- core,mon: mon/MgrStatMonitor: avoid dup health warnings during luminous upgrade ([issue#20435](#), [pr#15986](#), Sage Weil)
- core,mon: mon/MgrStatMonitor: keep mgrstat version ahead of pgmon ([issue#20219](#), [pr#15584](#), Sage Weil)
- core,mon: mon,osd: add crush_version to OSDMap, and allow crush map updates to gate on crush_version ([pr#15533](#), Sage Weil)
- core,mon: mon,osd: decouple creating pgs from pgmap ([pr#13999](#), Kefu Chai)
- core,mon: mon, osd: misc fixes ([pr#16078](#), xie xingguo)
- core,mon: mon/OSDMonitor: cancel mapping job from update_from_paxos ([issue#20067](#), [pr#15320](#), Sage Weil)
- core,mon: mon/OSDMonitor: make ‘osd crush move ...’ work on osds ([issue#18587](#), [pr#12981](#), Sage Weil)
- core,mon: mon/OSDMonitor: make snaps on tier pool should not be allowed ([pr#9348](#), Mingxin Liu)
- core,mon: mon/OSDMonitor: use up set instead of acting set in reweight_by_utilization ([pr#13802](#), Mingxin Liu)
- core,mon: mon,osd: new mechanism for managing full and nearfull OSDs for luminous ([pr#13615](#), Sage Weil)
- core,mon: mon/PGMap: call blocked requests ERR not WARN ([pr#15501](#), Sage Weil)
- core: mon,osd: add require_min_compat_client setting to enforce and clarify client compatibility ([pr#14959](#), Sage Weil)
- core: mon,osd: luminous feature bits, require flags, upgrade gates ([pr#13278](#), Sage Weil)
- core: mon, osd: misc fixes and cleanups ([pr#16160](#), xie xingguo)
- core: mon, osd: misc fixes ([pr#16283](#), xie xingguo)
- core: mon/OSDMonitor: `_apply_remap -> _apply_upmap`; less code redundancy ([pr#15846](#), xie xingguo)
- core: mon/OSDMonitor: batch noup/noin osds support ([pr#15725](#), xie xingguo)
- core: mon/OSDMonitor: batch OSDs nodown/noout support ([pr#15381](#), xie xingguo)
- core: mon/OSDMonitor: change info in ‘osd failed’ messages ([pr#15321](#), Sage Weil)
- core: mon/OSDMonitor: do not allow crush device classes until luminous ([pr#16188](#), Sage Weil)
- core: mon/OSDMonitor: fixup sortbitwise flag warning ([pr#12682](#), huanwen ren)
- core: mon/OSDMonitor: make mapping job behave if `mon_osd_prime_pg_temp = false` ([issue#19020](#), [pr#13574](#), Sage Weil)
- core: mon/OSDMonitor: osd crush set-device-class ([issue#19307](#), [pr#14039](#), Loic Dachary)
- core: mon/OSDMonitor: set `last_force_op_resend` on overlay pool too ([issue#18366](#), [pr#12712](#), Sage Weil)
- core: mon/OSDMonitor: should propose osdmap update when cluster addr changed ([pr#11065](#), Mingxin Liu)
- core: mon/OSDMonitor: skip prime_pg_temp if mapping is prior to osdmap ([pr#14826](#), Kefu Chai)
- core: mon,osd/OSDMap: a couple pg-upmap fixes ([pr#15319](#), Sage Weil)
- core: mon/PGMap: factor `mon_osd_full_ratio` into MAX AVAIL calc ([issue#18522](#), [pr#12923](#), Sage Weil)
- core: mon/PGMonitor: fix wrongly report “pg stuck in inactive” ([pr#14391](#), Mingxin Liu)
- core,mon,rbd: mon,osd: new rbd-based cephx cap profiles ([pr#15991](#), Jason Dillaman)
- core: msg/async/AsyncConnection: keepalive objecter ping connection to avoid timeout ([pr#14009](#), Haomai Wang)
- core: msg/async/AsyncConnection: socket’s fd can be zero, avoid false assert ([pr#13080](#), Haomai Wang)
- core: msg/async: avoid requeue racing with handle_write ([issue#20093](#), [pr#15324](#), Haomai Wang)
- core: msg/async/dpdk: fix compile errors ([pr#12698](#), Haomai Wang)
- core: msg/async: fix deleted_conn is out of sync with conns ([issue#20230](#), [pr#15645](#), Haomai Wang)
- core: msg/async: fix the bug of inaccurate calculation of `l_msgr_send_bytes` ([pr#16526](#), Jin Cai)
- core: msg/async/rdma: add log to show correct destruct queuepair ([pr#13412](#), Haomai Wang)
- core: msg/async/rdma: add perf counters to RDMA backend ([pr#13484](#), Haomai Wang)
- core: msg/async/rdma: destroy QueuePair if needed ([pr#13810](#), Haomai Wang)
- core: msg/async/rdma: don’t need to delete event when tcp connection isn’t ... ([pr#13528](#), Haomai Wang)
- core: msg/async/rdma: fix `ceph_clock_now` calls ([pr#12711](#), Haomai Wang)
- core: msg/async/rdma: fix potential racing connection usage ([pr#13738](#), Haomai Wang)
- core: msg/async/rdma: make Infiniband can be forkable ([pr#13525](#), Haomai Wang)
- core: msg/async/rdm: fix leak when existing failure in ip network ([pr#13435](#), Haomai Wang)
- core: msg/async: set thread name for msgr worker ([pr#13699](#), Haomai Wang)
- core: msg/async/Stack.cc: use of `pthread_setname_np()` needs compat.h ([pr#13825](#), Willem Jan Withagen)
- core: msg/async: support IPv6 QoS ([issue#18887](#), [issue#18928](#), [pr#13418](#), Robin H. Johnson)
- core: msg/simple: fix missing unlock when already bind ([pr#13267](#), Haomai Wang)
- core: msg/simple/Pipe:the returned value for `do_recv` unequal to zero ([pr#10272](#), zhang.zezhu)
- core: objclass: modify `omap_get_{keys,vals}` api ([pr#16667](#), Yehuda Sadeh, Casey Bodley)
- core: objclass-sdk: use namespace ceph for bufferlist ([pr#15581](#), Neha Ojha)

- core: os/bluestore: do not use nullptr to calc the size of bluestore_pextent_t (pr#14030, Kefu Chai)
- core: os/bluestore rm unused variable in aio_read() (pr#13530, tangwenjun)
- core: os/bluestore: silence gcc warning (pr#14028, Kefu Chai)
- core: osdc: clean up osd_command/start_mon_command interfaces (pr#13727, John Spray)
- core: osdc/Objecter: fix possible OSDSession leak on wrong connection (pr#13365, xie xingguo)
- core: osdc/Objecter: resend pg commands on interval change (issue#18358, pr#12869, Samuel Just)
- core: osdc/Objecter: respect epoch barrier in _op_submit() (issue#19396, pr#14190, Ilya Dryomov)
- core: osd/: don't leak context for Blessed*Context or RecoveryQueueAsync (issue#18809, pr#13342, Samuel Just)
- core: OSD: drop parameter t from merge_log() (pr#13923, xie xingguo)
- core: osd/ECBackend: cleanup for unnecessary copy with pg_stat_t (pr#12564, Yunchuan Wen)
- core: osd/ECBackend: drop duplicated pending_commit field from << operator (pr#13665, xie xingguo)
- core: osd/ECBackend: only need check missing_loc when doing recovery (pr#12526, huangjun)
- core: osd/ECBackend: remove unused variable "ReadCB" (pr#12543, huangjun)
- core: osd/ECTransaction: cleanup the redundant check which works in overwrite IO context (pr#15765, tang.jin)
- core: osd/ECTransaction: only read partial stripes when below *original* object size (issue#19882, pr#15712, Sage Weil)
- core: osd/filestore: Revert "os/filestore: move ondisk in front (issue#20524, pr#16156, Kefu Chai)
- core: osd/librados: add manifest, redirect (pr#15325, Sage Weil)
- core: osd/librados: cmpext support (pr#14715, Zhengyong Wang, David Disseldorp, Mike Christie)
- core: osd/librados: remove clone_range and associated multi-object cruft (pr#13008, Samuel Just)
- core: osd, messages/MOSDPing: bunch of fixes related to ping inflation (pr#15727, Piotr Dałek)
- core: osd/mon/mds: fix config set tell command (issue#20803, pr#16700, John Spray)
- core: osd/mon: misc full fixes and cleanups (pr#13968, David Zafman)
- core: osd/OpRequest: dump both name and addr for the client op (pr#12691, runsisi)
- core: osd/OSD: bump up current version; conditionally encoding manifest into oi (pr#15687, xie xingguo)
- core: osd/osd_internal_types: wake snaptrimmer on put_read lock, too (issue#19131, pr#13755, Sage Weil)
- core: osd/OSDMap: bump encoding version for require_min_compat_client (pr#15046, "Yan, Zheng")
- core: osd/OSDMap: Change *pg_to_* to return void (pr#15684, Brad Hubbard)
- core: osd/OSDMap: don't set weight to IN when OSD is destroyed (issue#19119, pr#13730, Ilya Dryomov)
- core: osd/OSDMap: hide require_osd and sortbitwise flags (pr#14440, Sage Weil)
- core: osd/OSDMap: improve upmap calculation (issue#19818, pr#14902, Sage Weil)
- core: osd/OSDMap: Uncomment code to enable private default constructors (pr#12597, Brad Hubbard)
- core: osd/OSD: tolerate any 'set-device-class' error on OSD startup (pr#16812, xie xingguo)
- core: osd/osd_type: Fix logging output (pr#12778, Brad Hubbard)
- core: osd/osd_types: Move comment to more relevant position (pr#12779, Brad Hubbard)
- core: osd/osd_types: print notify-ack op properly (pr#12585, Sage Weil)
- core: osd/PG: add new have_unfound() function in MissingLoc (pr#12668, huangjun)
- core: osd/PG: Add two new mClock implementations of the PG sharded operator queue (pr#14997, J. Eric Ivancich)
- core: osd/PG.cc: Optimistic estimation on PG.last_active (pr#14799, Xiaoxi Chen)
- core: osd/PG.cc: unify the call of checking whether lock is held (pr#15013, Jin Cai)
- core: osd/PG: check the connection first in fulfill_log (pr#12579, huangjun)
- core: osd/PG: conditionally retry on receiving pg-notify when Primary is Incomplete (pr#13942, xie xingguo)
- core: osd/PG: drop pre-firefly compat_mode for choose_acting (pr#15057, Sage Weil)
- core: osd/PG: fix lost unfound + delete when there are no missing objects (issue#20904, pr#16809, Josh Durgin)
- core: osd/PG: fix possible overflow on unfound objects (pr#12669, huangjun)
- core: osd/PG: fix warning so we discard_event() on a no-op state change (pr#16655, Sage Weil)
- core: osd/PG: ignore CancelRecovery in NotRecovering (issue#20804, pr#16638, Sage Weil)
- core: osd/PGLog: avoid infinite loop if missing version is corrupted (pr#16798, Josh Durgin)
- core: osd/PGLog: fix inaccurate missing assert (issue#20753, pr#16539, Josh Durgin)
- core: osd/PGLog: fix index for parent and child log on split (issue#18975, pr#13493, Sage Weil)
- core: osd/pglog: remove loop through empty collection (pr#15121, J. Eric Ivancich)
- core: osd/PGLog: skip ERROR entires in _merge_object_divergent_entries (issue#20843, pr#16675, Jeegn Chen)
- core: osd/PG: make non-empty PastIntervals non-fatal (issue#20167, pr#15639, Sage Weil)
- core: osd/PG: only correct filestore collection bits on load (issue#19541, pr#14397, Sage Weil)
- core: osd/PG: publish PG stats when backfill-related states change (issue#18369, pr#12727, Sage Weil)
- core: osd/PG: reset the missing set when restarting backfill (issue#19191, pr#14053, Josh Durgin)
- core: osd/PG: restrict want_acting to up+acting on recovery completion (issue#18929, pr#13420, Sage Weil)
- core: osd/PG: set clean when last_epoch_clean is updated (issue#19023, pr#15555, Samuel Just)
- core: osd/PG: simplify the logic of backfill_targets checking (pr#12519, huangjun)
- core: osd/PG: some minor cleanups (pr#14133, runsisi)
- core: osd/PrimaryLogPG: clear oi from trim_object() (issue#19947, pr#15519, Sage Weil)
- core: osd/PrimaryLogPG: do not call on_shutdown() if (pg.deleting) (issue#19902, pr#15040, Kefu Chai)
- core: osd/PrimaryLogPG: do not expect FULL_TRY ops to get resent (issue#19430, pr#14255, Sage Weil)
- core: osd/PrimaryLogPG::failed_push: update missing as well (issue#18165, pr#12888, Samuel Just)
- core: osd/PrimaryLogPG: fix oi reset during trim_object (issue#19947, pr#15696, Sage Weil)
- core: osd/PrimaryLogPG: fix recovering hang when have unfound objects (pr#16558, huangjun)

- core: osd/PrimaryLogPG: optimal pick_newest_available ([pr#12695](#), huangjun)
- core: osd/PrimaryLogPG: record prior_version for DELETE events ([issue#20274](#), [pr#15649](#), Sage Weil)
- core: osd/PrimaryLogPG: remove duplicated code ([pr#13894](#), Jianpeng Ma)
- core: osd/PrimaryLogPG: set return value if sparse read failed ([pr#14093](#), huangjun)
- core: osd/PrimaryLogPG: skip deleted missing objects in pg[n]ls ([issue#20739](#), [pr#16490](#), Josh Durgin)
- core: osd/PrimaryLogPG solve cache tier osd high memory consumption ([issue#20464](#), [pr#16011](#), Peng Xie)
- core: osd/PrimaryLogPG::try_lock_for_read: give up if missing ([issue#18583](#), [pr#13087](#), Samuel Just)
- core: osd/PrimaryLogPG: unify the access to primary pg ([pr#12527](#), huangjun)
- core: osd/PrimayLogPG: update modified range to include the whole object size for write_full op ([pr#15021](#), runsisi)
- core: osd/ReplicatedBackend: clear pull source once we are done with it ([issue#19076](#), [pr#13879](#), Samuel Just)
- core: osd/ReplicatedBackend: remove MOSDSubOp cruft from repop_applied ([pr#14358](#), Jianpeng Ma)
- core: osd/ReplicatedBackend: reset thread heartbeat after every omap entry ... ([issue#20375](#), [pr#15823](#), Josh Durgin)
- core: osd/ReplicatedBackend: take read locks for clone sources during recovery ([issue#17831](#), [pr#12844](#), Samuel Just)
- core: os/filestore: call committed_thru when no journal entries are replayed ([pr#15781](#), Kuan-Kai Chiu)
- core: os/filestore: debug which omap keys are set ([issue#19067](#), [pr#13671](#), Sage Weil)
- core: os/filestore: do not free event if not added ([pr#16235](#), Kefu Chai)
- core: os/filestore/FileJournal: bufferlist rebuild ([pr#13980](#), Jianpeng Ma)
- core: os/filestore/FileJournal: FileJournal::open() close journal file before return error ([issue#20504](#), [pr#16120](#), Yang Honggang)
- core: os/filestore/FileStore.cc: remove a redundant judgement when get max latency ([pr#15961](#), Jianpeng Ma)
- core: os/filestore/FileStore.cc: remove unneeded loop ([pr#12177](#), Li Wang)
- core: os/filestore: fix clang static check warn “use-after-free” ([pr#12581](#), liuchang0812)
- core: os/filestore: fix infinit loops in fiemap() ([pr#14367](#), Ning Yao)
- core: os/filestore: handle error returned from write_fd() ([pr#10146](#), yonghengdexin735)
- core: os/filestore/HashIndex: be loud about splits ([issue#18235](#), [pr#12421](#), Dan van der Ster)
- core: os/filestore/JournalingObjectStore cleanup ([pr#12528](#), Li Wang)
- core: os/filestore: require experimental flag for btrfs ([pr#16086](#), Sage Weil)
- core: os/filestore: version will be uninitialized varible if store_version doesn't exist ([pr#12582](#), liuchang0812)
- core: os/fs/FS.cc: remove the redundant code ([pr#14362](#), Jianpeng Ma)
- core: os/FuseStore: include <functional> header in src/os/FuseStore.h for gcc 7.x ([pr#13454](#), Jos Collin)
- core,performance: common/config_opts: improve rdma buffer size to 128k ([pr#13510](#), Haomai Wang)
- core,performance: common/TrackedOp: various cleanups and optimizations ([pr#12537](#), Sage Weil)
- core,performance: kv/RocksDBStore: Table options for indexing and filtering ([pr#16450](#), Mark Nelson)
- core,performance: mon,osd: explicitly remap some pgs ([pr#13984](#), Sage Weil)
- core,performance: msg/async: avoid lossy connection sending ack message ([pr#13700](#), Haomai Wang)
- core,performance: msg/async/rdma: cleanup ([pr#13509](#), Haomai Wang)
- core,performance: msg/async/rdma: refactor tx handle flow to get rid of locks ([pr#13680](#), Haomai Wang)
- core,performance: msg/async: reduce write_lock contention ([pr#15092](#), Haomai Wang)
- core,performance: osd/ECBackend: Send write message to peers first, then do local write ([pr#12522](#), huangjun)
- core,performance: osd/OSD.h: requeue the scrub job with higher priority to shorten the blocking time of related requests ([pr#15552](#), Jin Cai)
- core,performance: osd, os: reduce fiemap burden ([pr#14640](#), Piotr Dałek)
- core,performance: osd/pg: bound the portion of the log we request in GetLog::GetLog() ([pr#12233](#), Jie Wang)
- core,performance: osd/PG: make prioritized recovery possible ([pr#13723](#), Piotr Dałek)
- core,performance: os/filestore: avoid unnecessary copy in filestore::do_transaction ([pr#12578](#), Yunchuan Wen)
- core,performance: os/filestore/HashIndex: randomize split threshold by a configurable amount ([issue#15835](#), [pr#15689](#), Josh Durgin)
- core,performance: os/filestore: queue ondisk completion before apply work ([pr#13918](#), Pan Liu)
- core,performance: os/filestore: use new sleep strategy when io_submit gets EAGAIN ([pr#14860](#), Pan Liu)
- core,performance: os/kstore: Added rocksdb bloom filter settings ([pr#13053](#), Ted-Chang)
- core,performance: src/OSD: add more useful perf counters for performance tuning ([pr#15915](#), Pan Liu)
- core: PGLog: store extra duplicate ops beyond the normal log entries ([pr#16172](#), Josh Durgin, J. Eric Ivancich)
- core: Prefix /proc/ with FreeBSD emulation ([pr#14290](#), Willem Jan Withagen)
- core: PrimaryLogPG: don't update digests for objects with mismatched names ([issue#18409](#), [pr#12788](#), Samuel Just)
- core: print more information when run ceph-osd cmd with ‘check options’ ([pr#16678](#), mychoxin)
- core: qa: do not restrict valgrind runs to centos ([issue#18126](#), [pr#15389](#), Greg Farnum)
- core,rbd: mon,osd: do not create rbd pool by default ([pr#15894](#), Greg Farnum, Sage Weil, David Zafman)
- core: ReplicatedBackend: don't queue Context outside of ObjectStore with obc ([issue#18927](#), [pr#13569](#), Samuel Just)
- core: Revert “PrimaryLogPG::failed_push: update missing as well” ([issue#18624](#), [pr#13090](#), David Zafman)
- core,rgw: misc: SCA and Coverity Fixes ([pr#13208](#), Danny Al-Gaaf)
- core,rgw: qa: Removed all ‘default_idle_timeout’ due to chnage in rgw task ([pr#15420](#), Yuri Weinstein)
- core,rgw,tests: qa/rgw_snaps: move default_idle_timeout config under the client ([issue#20128](#), [pr#15400](#), Yehuda Sadeh)
- core,rgw,tests: qa/suits/rados/basic/tasks/rgw_snaps: wait for pools to be created ([pr#16509](#), Sage Weil)
- core: rocksdb: sync with upstream ([issue#18464](#), [pr#13306](#), Kefu Chai)

- core: src/ceph.in: Use env(CEPH_DEV) to suppress noise from ceph ([pr#14746](#), Willem Jan Withagen)
- core: src/vstart.sh: kill dead upmap option ([pr#15848](#), xie xingguo)
- core:" Stringify needs access to << before reference" src/include/stringify.h ([pr#16334](#), Willem Jan Withagen)
- core: test, osd: fix some coverity issues ([pr#13293](#), liuchang0812)
- core: test/pybind/test_rados.py: tolerate TimedOut in test_ping_monitor ([issue#18529](#), [pr#12934](#), Samuel Just)
- core,tests: ceph-disk: sensible default for block.db ([pr#15576](#), Loic Dachary)
- core,tests: ceph-disk/tests: Certain partition types do not work on FreeBSD ([pr#13560](#), Willem Jan Withagen)
- core,tests: ceph-disk/tests/test_main.py: FreeBSD does not do multipath ([pr#13847](#), Willem Jan Withagen)
- core,tests: ceph_test_librados_api_misc: fix stupid LibRadosMiscConnectFailure.ConnectFailure test ([issue#15368](#), [pr#14261](#), Sage Weil)
- core,tests: ceph_test_rados_api_misc: avoid livelock from PoolCreationRace ([pr#13565](#), Sage Weil)
- core,tests: ceph_test_rados_api_misc: Fix trivial memory leak ([pr#12680](#), Brad Hubbard)
- core,tests: ceph_test_rados_api: wait for snap trim on ENOENT during cleanup ([issue#19948](#), [pr#15638](#), Sage Weil)
- core,tests: ceph_test_rados_api_watch_notify: flush after unwatch ([issue#20105](#), [pr#16402](#), Sage Weil)
- core,tests: ceph_test_rados_api_watch_notify: make LibRadosWatchNotify.Watch3Timeout tolerate thrashing ([issue#19433](#), [pr#14899](#), Sage Weil)
- core,tests: ceph_test_rados: max_stride_size must be more than min_stride_size ([issue#20775](#), [pr#16590](#), Lianne Wang)
- core,tests: c_write_operations.cc: Fix trivial memory leak ([pr#12663](#), Brad Hubbard)
- core,tests: do all valgrind runs on centos ([issue#20360](#), [issue#18126](#), [pr#16046](#), Sage Weil)
- core,tests: os: allow 'osd objectstore = random' to pick either filestore or bluestore ([pr#13754](#), Sage Weil)
- core,tests: qa: avoid map-gap tests for k=2 m=1 ([issue#20844](#), [pr#16789](#), Sage Weil)
- core,tests: qa: move ceph-helpers-based make check tests to qa/standalone; run via teuthology ([pr#16513](#), Sage Weil)
- core,tests: qa/objectstore/filestore-btrfs: test btrfs on trusty only ([issue#20169](#), [pr#15814](#), Sage Weil)
- core,tests: qa/objectstore: test bluestore with aggressive compression ([pr#14623](#), Sage Weil)
- core,tests: qa/rados/upgrade/jewel-x-singleton: run luminous.yaml at the end ([pr#13378](#), Sage Weil)
- core,tests: qa: stop testing btrfs ([issue#20169](#), [pr#16044](#), Sage Weil)
- core,tests: qa/suites/powercycle/osd/tasks/radosbench: consume less space ([issue#20302](#), [pr#15821](#), Sage Weil)
- core,tests: qa/suites/rados: at-end: ignore PG_{AVAILABILITY,DEGRADED} ([issue#20693](#), [pr#16575](#), Sage Weil)
- core,tests: qa/suites/rados/*/at-end: wait for healthy before scrubbing ([pr#15245](#), Sage Weil)
- core,tests: qa/suites/rados/basic: set low omap limit for rgw workload ([pr#13071](#), Sage Weil)
- core,tests: qa/suites/rados/basic/tasks/rados_python: POOL_APP_NOT_ENABLED ([pr#16827](#), Sage Weil)
- core,tests: qa/suites/rados/mgr/tasks/failover: whitelist ([pr#16795](#), Sage Weil)
- core,tests: qa/suites/rados/singleton/all/reg11184: whitelist health warnings ([pr#16306](#), Sage Weil)
- core,tests: qa/suites/rados/singleton-nomsg/health-warnings: behave on ext4 ([issue#20043](#), [pr#15207](#), Sage Weil)
- core,tests: qa/suites/rados: temporarily remove scrub_test from basic/ until post-luminous ([issue#19935](#), [pr#15202](#), Sage Weil)
- core,tests: qa/suites/rados/thrash/workload/*: enable rados.py cache tiering ops ([issue#11793](#), [pr#16244](#), Sage Weil)
- core,tests: qa/suites/upgrade/kraken-x: enable experimental for bluestore ([pr#15359](#), Sage Weil)
- core,tests: qa/tasks/ceph: enable rbd on rbd pool ([pr#16794](#), Sage Weil)
- core,tests: qa/tasks/ceph_manager: get osds all in after thrashing ([pr#15784](#), Sage Weil)
- core,tests: qa/tasks/ceph_manager: wait for osd to start after objectstore-tool sequence ([issue#20705](#), [pr#16454](#), Sage Weil)
- core,tests: qa/tasks/ceph_manager: wait longer for pg stats to flush ([pr#16322](#), Sage Weil)
- core,tests: qa/tasks/ceph: osd_scrub_pgs: reissue scrub requests in loop ([issue#20326](#), [pr#15747](#), Sage Weil)
- core,tests: qa/tasks/ceph.py: no osd id to 'osd create' command ([issue#20548](#), [pr#16233](#), Sage Weil)
- core,tests: qa/tasks/ceph.py: tolerate active+clean+something ([pr#15717](#), Sage Weil)
- core,tests: qa/tasks/ceph: simplify ceph deployment slightly ([pr#15853](#), Sage Weil)
- core,tests: qa/tasks/ceph: wait for mgr to activate and pg stats to flush in health() ([issue#20744](#), [pr#16514](#), Sage Weil)
- core,tests: qa/tasks/dump_stuck: fix dump_stuck test bug ([pr#16559](#), huangjun)
- core,tests: qa/tasks/dump_stuck: fix for active+clean+remapped ([issue#20431](#), [pr#15955](#), Sage Weil)
- core,tests: qa/tasks/radosbench: longer timeout ([pr#16213](#), Sage Weil)
- core,tests: qa/workunits/cephtool/test.sh: add sudo for daemon compact ([pr#16500](#), Sage Weil)
- core,tests: qa/workunits/cephtool/test.sh: fix osd full health detail grep ([issue#20187](#), [pr#15494](#), Sage Weil)
- core,tests: qa/workunits/rados/test_health_warning: misc fixes ([issue#19990](#), [pr#15201](#), Sage Weil)
- core,tests: qa/workunits/rest: use unique pool names for cephfs test ([pr#13188](#), Sage Weil)
- core,tests: Revert "qa: do not restrict valgrind runs to centos" ([issue#20360](#), [pr#15791](#), Sage Weil)
- core,tests: test: add separate ceph-helpers-based smoke test ([pr#16572](#), Sage Weil)
- core,tests: test/librados/cmd.cc: Fix trivial memory leaks ([pr#12671](#), Brad Hubbard)
- core,tests: test/librados/c_read_operations.cc: Fix trivial memory leak ([pr#12656](#), Brad Hubbard)
- core,tests: test/librados/c_read_operations.cc: Fix valgrind errors ([issue#18354](#), [pr#12657](#), Brad Hubbard)
- core,tests: test/librados: Silence Coverity memory leak warnings ([pr#12442](#), Brad Hubbard, Samuel Just)
- core,tests: test/librados/snapshots.cc: Fix memory leak ([pr#12690](#), Brad Hubbard)
- core,tests: test/librados/tier.cc: Fix valgrind errors ([issue#18360](#), [pr#12705](#), Brad Hubbard)
- core,tests: test/osd/TestRados.cc: run set-redirect test after finishing setup ([issue#20114](#), [pr#15385](#), Myoungwon Oh)
- core,tests: test_rados_watch_notify: Fix trivial memory leaks ([pr#12713](#), Brad Hubbard)

- core,tests,tools: Fixes: <http://tracker.ceph.com/issues/18533> (pr#13423, Samuel Just, David Zafman)
- core,tests: upgrade/jewel-x: a few fixes (pr#16830, Sage Weil)
- core: throttle: Minimal destructor fix for Luminous (pr#16661, Adam C. Emerson)
- core,tools: ceph: perfcounter priorities and daemonperf updates to use them (pr#14793, Sage Weil, Dan Mick)
- core,tools: kv: move 'bluestore-kv' hackery out of KeyValueDB into ceph-kvstore-tool (issue#19778, pr#14895, Sage Weil)
- core,tools: osdmaptool: require -upmap-save before modifying input osdmap (pr#15247, Sage Weil)
- core: vstart.sh: start mgr after mon, before osds (pr#16613, Sage Weil)
- core: Wip 20985 divergent handling luminous (issue#20985, pr#17001, Greg Farnum)
- create the ceph-volume and ceph-volume-systemd man pages (pr#17158, Alfredo Deza)
- crush: a couple of weight-set fixes (pr#16623, xie xingguo)
- crush: add devices class that rules can use as a filter (issue#18943, pr#13444, Loic Dachary)
- crush: add -dump to crushtool (pr#13726, Loic Dachary)
- crush: add missing tunable in tests (pr#15412, Loic Dachary)
- crush: allow uniform buckets with no items (pr#13521, Loic Dachary)
- crush: API documentation (pr#13205, Loic Dachary)
- crush: bucket: crush_add_uniform_bucket_item should check for uniformity (pr#14208, Sahid Orentino Ferdjaoui)
- crush: builder: clean the arguments of crush_reweight* methods (pr#14110, Sahid Orentino Ferdjaoui)
- crush: builder: creating crush map with optimal configurations (pr#14209, Sahid Orentino Ferdjaoui)
- crush: builder: legacy has chooseleaf_stable = 0 (pr#14695, Loic Dachary)
- crush: crush_init_workspace starts with struct crush_work (pr#14696, Loic Dachary)
- crush: detect and (usually) fix ruleset != rule id (pr#13683, Sage Weil)
- crush: document tunables and rule step set (pr#13722, Loic Dachary)
- crush: do is_out test only if we do not collide (pr#13326, xie xingguo)
- crush: encode can override weights with weight set (issue#19836, pr#15002, Loic Dachary)
- crush: enforce buckets-before-rules rule (pr#16453, Sage Weil)
- crush: fix CrushCompiler won't compile maps with empty shadow tree (pr#17228, xie xingguo)
- crush: fix dprintk compilation (pr#13424, Loic Dachary)
- crush: force rebuilding shadow hierarchy after swapping buckets (pr#17229, xie xingguo)
- crush: misc changes/fixes for device classes (issue#20845, pr#16805, Kefu Chai, xie xingguo, Sage Weil)
- crush: more class fixes (pr#16837, xie xingguo)
- crush: only encode class info if SERVER_LUMINOUS (issue#19361, pr#14131, Sage Weil)
- crush: optimize header file dependency (pr#9307, Xiaowei Chen)
- crush: silence warning from -Woverflow (pr#16329, Jos Collin)
- crush: s/ruleset/id/ in decompiled output; prevent compilation when ruleset != id (pr#16400, Sage Weil)
- crush: update choose_args when items are added/removed (pr#15311, Loic Dachary)
- crush: update documentation for negative choose step (pr#14970, Loic Dachary)
- crush: various weight-set fixes (pr#17214, xie xingguo)
- crush: verify weights is influenced by the number of replicas (issue#15653, pr#13083, Adam C. Emerson, Loic Dachary)
- crush: weight_set and id remapping (issue#15653, pr#14486, Loic Dachary)
- crush: when osd_location_hook does not exist, we should exit error (pr#12961, song baisen)
- doc: 12.1.0/release notes 2 (pr#15627, Abhishek Lekshmanan)
- doc: 12.1.1 & 12.1.2 release notes (pr#16377, Abhishek Lekshmanan)
- doc: add 0.94.10 and hammer EOL to releases.rst (pr#13069, Nathan Cutler)
- doc: add 12.0.1 release notes (pr#14106, Abhishek Lekshmanan)
- doc: Add amitkumar50 affiliation to .organizationmap (pr#16475, Amit Kumar)
- doc add ceph-volume and ceph-volume-systemd man pages to CMakeLists file (pr#17170, Alfredo Deza)
- doc: add changelog for v0.94.10 (pr#13572, Abhishek Lekshmanan)
- doc: add changelog for v10.2.6 Jewel release (pr#13839, Abhishek Lekshmanan)
- doc: add changelog for v10.2.7 (pr#14441, Abhishek Lekshmanan)
- doc: add descriptions for mon/mgr options (pr#15032, Kefu Chai)
- doc: add doc requirements on PR submitters (pr#16394, John Spray)
- doc: added mgr caps to manual deployment documentation (pr#16660, Nick Erdmann)
- doc: add FreeBSD manual install (pr#14941, Willem Jan Withagen)
- doc: add instructions for replacing an OSD (pr#16314, Kefu Chai)
- doc: add new cn ceph mirror to doc and mirroring (pr#15089, Shengjing Zhu)
- doc: add optional argument for build-doc (pr#14058, Kefu Chai)
- doc: add rados xattr commands to manpage (pr#15362, Andreas Gerstmayer)
- doc: add rbd new trash cli and cleanups in release-notes.rst (issue#20702, pr#16498, songweibin)
- doc: add README to dmclock subdir to inform developers it's a git subtree (pr#15386, J. Eric Ivancich)
- doc: add RGW Idap auth documentation (pr#14339, Harald Klein)
- doc: add some undocumented options to rbd-nbd (pr#14134, wangzhengyong)
- doc: add verbiage to rbdmap manpage (issue#18262, pr#12509, Nathan Cutler)
- doc: Add Zabbix ceph-mgr plugin to PendingReleaseNotes (pr#16412, Wido den Hollander)
- doc: AUTHORS: update CephFS PTL (pr#16399, Patrick Donnelly)

- doc: AUTHORS: update tech leads ([pr#14350](#), Patrick Donnelly)
- doc: AUTHORS: update with release manager, backport team ([pr#15391](#), Sage Weil)
- doc: build/install-deps.sh: Add sphinx package for building docs on FreeBSD ([pr#13223](#), Willem Jan Withagen)
- doc: ceph-disk: use '-' for feeding ceph cli with stdin ([pr#16362](#), Kefu Chai)
- doc: change osd_op_thread_timeout default value to 15 ([pr#14199](#), Andreas Gerstmayr)
- doc: Change the default values of some OSD options ([issue#20199](#), [pr#15566](#), Bara Ancincova)
- doc: clarify "ceph quorum" syntax ([issue#17802](#), [pr#11787](#), Nathan Cutler)
- doc: clarify SubmittingPatches.rst ([pr#12988](#), Nathan Cutler)
- doc: clarify that "ms bind ipv6" disables IPv4 ([pr#13317](#), Ken Dreyer)
- doc: clarify the path restriction mds cap example ([pr#12993](#), John Spray)
- doc: common/options.cc: document bluestore config options ([pr#16489](#), Sage Weil)
- doc: correct and improve add user capability section ([pr#14055](#), Chu, Hua-Rong)
- doc: correct arguments for ceph tell osd.N bench ([pr#14462](#), Patrick Dinnen)
- doc: Correcting the remove bucket example and adding bucket link/unlink examples ([pr#12460](#), Uday Mullangi)
- doc: correct S3 lifecycle support explain ([issue#18459](#), [pr#12827](#), liuchang0812)
- doc: correct the quota section ([issue#19397](#), [pr#14122](#), Chu, Hua-Rong)
- doc: crush: API documentation fixes ([pr#13589](#), Loic Dachary)
- doc: crush typo in algorithm description ([pr#13661](#), Loic Dachary)
- doc: deletes duplicated word and clarifies an example ([pr#13746](#), Tahia Khan)
- doc: describe CephFS max_file_size ([pr#15287](#), Ken Dreyer)
- doc: describe mark_events logging available via the OSD's OpTracker ([pr#15095](#), Greg Farnum)
- doc: Describe mClock's use within Ceph in great detail ([pr#16707](#), J. Eric Ivancich)
- doc: dev add a note about ccache ([pr#14478](#), Abhishek Lekshmanan)
- doc: dev: add notes on PR make check validation test ([pr#16079](#), Nathan Cutler)
- doc: dev guide: how to run s3-tests locally against vstart ([pr#14508](#), Nathan Cutler, Abhishek Lekshmanan)
- doc: dev improve the s3tests doc to reflect current scripts ([pr#15180](#), Abhishek Lekshmanan)
- doc: doc/cephfs: mention RADOS object size limit ([pr#15550](#), John Spray)
- doc: doc/dev: update log_based_pg.rst, fix some display problem ([pr#12730](#), liuchang0812)
- doc: Doc: Fixes Python Swift client commands ([issue#17746](#), [pr#12887](#), Ronak Jain)
- doc: doc/install/manual-deployment: update osd creation steps ([pr#16573](#), Sage Weil)
- doc: doc/mgr/dashboard: update dashboard docs to reflect new defaults ([pr#16241](#), Sage Weil)
- doc: doc/mon: fix ceph-authtool command in rebuild mon's sample ([pr#16503](#), huanwen ren)
- doc: doc/qa: cover config help command ([pr#16727](#), John Spray)
- doc: doc/rados.8: add offset option for put command ([pr#16155](#), Jianpeng Ma)
- doc: doc/rados: add page for health checks and update monitoring.rst ([pr#16566](#), John Spray)
- doc: doc/rados/configuration: document bluestore ([pr#16765](#), Sage Weil)
- doc: doc/radosgw/s3/cpp.rst: update usage of libS3 APIs to make the examples work ([pr#10851](#), Weibing Zhang)
- doc: doc/rados/operations/health-checks: osd section ([pr#16611](#), Sage Weil)
- doc: doc/release-notes: add Images creation timestamp note ([pr#15963](#), clove)
- doc: doc/release-notes: avoid 'production-ready' in describing kraken ([pr#13675](#), Sage Weil)
- doc: doc/release-notes: final kraken notes ([pr#12968](#), Sage Weil)
- doc: doc/release-notes: fix bluestore links ([pr#16787](#), Sage Weil)
- doc: doc/release-notes: fix links, formatting; add crush device class docs ([pr#16741](#), Sage Weil)
- doc: doc/release-notes: fix upmap and osd replacement links; add fixme ([pr#16730](#), Sage Weil)
- doc: doc/release-notes: Luminous release notes typo fixes "ceph config-key ls"->"ceph config-key list" ([pr#16330](#), scienceluo)
- doc: doc/release-notes: Luminous release notes typo fixes ([pr#16338](#), Luo Kexue)
- doc: doc/release-notes: sort release note changes into the right section ([pr#16764](#), Sage Weil)
- doc: doc/release-notes: update device class cli ([pr#16851](#), xie xingguo)
- doc: doc/release-notes: update luminous notes ([pr#15851](#), Sage Weil)
- doc: doc/release-notes: update which jewel version does sortbitwise warning ([pr#15209](#), Sage Weil)
- doc: doc/releases: Update releases from Feb 2017 to July 2017 ([pr#16303](#), Bryan Stillwell)
- doc: doc/rgw: instructions for changing multisite master zone ([pr#14089](#), Casey Bodley)
- doc: doc/rgw: remove fastcgi page and sample configs ([pr#15133](#), Casey Bodley)
- doc: doc/rgw: remove Federated Configuration, clean up multisite ([issue#19504](#), [issue#18082](#), [pr#15132](#), Casey Bodley)
- doc: docs: Clarify the relationship of min_size to EC pool recovery ([pr#14419](#), Brad Hubbard)
- doc: docs: Fix problems with example code ([pr#14007](#), Brad Hubbard)
- doc: docs: mgr dashboard ([pr#15920](#), Wido den Hollander)
- doc: [docs/quick-start]: update quick start to add a note for mgr create command for luminous+ builds ([pr#16350](#), Vasu Kulkarni)
- doc: Documentation Fixes for <http://tracker.ceph.com/issues/19879> ([issue#20057](#), [issue#19879](#), [pr#15606](#), Sameer Tiwari)
- doc: Documentation updates for July 2017 releases ([pr#16401](#), Bryan Stillwell)
- doc: document bluestore compression settings ([pr#16747](#), Kefu Chai)
- doc: document mClock related options ([pr#16552](#), Kefu Chai)

- doc: document *osd-agent-{max,low}-ops* options (pr#13648, Kefu Chai)
- doc: document perf histograms (pr#15150, Piotr Dałek)
- doc: document “rados cleanup” in rados manpage (issue#20894, pr#16777, Nathan Cutler)
- doc: document repair/scrub features (issue#15786, pr#9032, Kefu Chai, David Zafman)
- doc: Document RGW quota cache options (issue#18747, pr#13395, Daniel Gryniewicz)
- doc: Document that osd_heartbeat_grace applies to MON and OSD (pr#13098, Wido den Hollander)
- doc: document the setup of restful and dashboard plugins (issue#20239, pr#15707, Kefu Chai)
- doc: explain about logging levels (pr#12920, liuchang0812)
- doc: fio: update README.md so only the fio ceph engine is built (pr#15081, Kefu Chai)
- doc: fix a typo (pr#13930, Drunkard Zhang)
- doc: fix broken link in erasure-code.rst (issue#19972, pr#15143, MinSheng Lin)
- doc: fix document about rados mon (pr#12662, liuchang0812)
- doc: Fixed a typo in yum repo filename script (pr#16431, Jeff Green)
- doc: fixes a broken hyperlink to RADOS paper in architecture (pr#13682, Tahia Khan)
- doc: Fixes a typo (pr#13985, Edwin F. Boza)
- doc: Fixes parameter name in rbd configuration on openstack havana/icehouse (issue#17978, pr#13403, Michael Eischer)
- doc: Fixes radosgw-admin ex: in swift auth section (issue#16687, pr#12646, SirishaGuduru)
- doc: fixes to silence sphinx-build (pr#13997, Kefu Chai)
- doc: fix factual inaccuracy in doc/architecture.rst (pr#15235, Nathan Cutler, Sage Weil)
- doc: fixing an error in 12.0.3 release notes (pr#15195, Abhishek Lekshmanan)
- doc: fix link for ceph-mgr cephx authorization (pr#16246, Greg Farnum)
- doc: fix link that pointed to a nonexistent file (pr#14740, Peter Maloney)
- doc: fix syntax on code snippets in cephfs/multimds (pr#15499, John Spray)
- doc: fix the librados c api can not compile problem (pr#9396, song baisen)
- doc: fix the links to <http://ceph.com/docs> (issue#19090, pr#13976, Kefu Chai)
- doc: Fix typo and grammar in RGW config reference (pr#13356, Ruben Kerkhof)
- doc: fix typo in config.rst (pr#16721, Jos Collin)
- doc: fix typos in config.rst (pr#16681, Song Shun)
- doc: fix typos in radosgw-admin usage (pr#13936, Enming Zhang)
- doc: freshen mgr docs (pr#15690, John Spray)
- doc: hammer 0.94.10 release notes (pr#13152, Nathan Cutler)
- doc: Have install put manpages in the FreeBSD correct location (pr#13301, Willem Jan Withagen)
- doc: how to specify filesystem for cephfs clients (pr#14087, John Spray)
- doc: improve firewalld instructions (pr#13360, Ken Dreyer)
- doc: Indicate how to add multiple admin capabilities (pr#13956, Chu, Hua-Rong)
- doc: instructions and guidance for multimds (issue#19135, pr#13830, John Spray)
- doc: instructions for provisioning OpenStack VMs ad hoc (pr#13368, Nathan Cutler)
- doc: Jewel 10.2.6 release notes (pr#13835, Abhishek Lekshmanan)
- doc: Jewel v10.2.8 release notes (pr#16274, Nathan Cutler)
- doc: Jewel v10.2.9 release notes (pr#16318, Nathan Cutler)
- doc: kernel client os-recommendations update (pr#13369, John Spray, Ilya Dryomov)
- doc: kill some broken links (pr#15203, liuchang0812)
- doc: kill sphinx warnings (pr#16198, Kefu Chai)
- doc: luminous: doc: update rbd-mirroring documentation (issue#20701, pr#16912, Jason Dillaman)
- doc: Luminous release notes typo fixes (pr#15899, Abhishek Lekshmanan)
- doc: mailmap: add affiliation for Zhu Shangzhong (pr#16537, Zhu Shangzhong)
- doc: mailmap: add Alibaba into organization map (pr#14900, James Liu)
- doc: mailmap: add Myoungwon Oh's mailmap and affiliation (pr#15934, Myoungwon Oh)
- doc: mailmap for v12.0.2 (pr#14753, Abhishek Lekshmanan)
- doc: mailmap: Michal Koutny affiliation (pr#13036, Nathan Cutler)
- doc: mailmap, organizationmap: add affiliation for Tushar Gohad (pr#16081, Tushar Gohad)
- doc: .mailmap, .organizationmap: Update Fan Yang information and affiliation (pr#16067, Fan Yang)
- doc: .mailmap, .organizationmap: Update Song Weibin information and affiliation (pr#16311, songweibin)
- doc: .mailmap, .organizationmap: Update ztczll affiliation (pr#16038, zhanglei)
- doc: mailmap updates for v11.1.0 (pr#12335, Abhishek Lekshmanan)
- doc: mailmap updates (pr#13309, Loic Dachary)
- doc: mailmap: V12.0.1 credits (pr#14479, M Ranga Swami Reddy)
- doc: mailmap: Willem Jan Withagen affiliation (pr#13034, Willem Jan Withagen)
- doc: mailmap: ztczll affiliation (pr#15079, zhanglei)
- doc: man/8/ceph-disk: fix formatting (pr#13969, Kefu Chai)
- doc: mention certain conf vars should be in global (pr#15119, Ali Maredia)
- doc: mention ENXIO change in the 10.2.6 release notes (pr#13878, Nathan Cutler)
- doc: mention -show-mappings in crushtool manpage (issue#19649, pr#14599, Nathan Cutler, Loic Dachary)
- doc: mention teuthology-worker security group (pr#14748, Nathan Cutler)

- doc: Merge pull request from stiwari/wip-19879 ([issue#19879](#), [pr#15609](#), Sameer Tiwari)
- doc: mgr/restful: bind to :: and update docs ([pr#16267](#), Sage Weil)
- doc: minor changes in fuse client config reference ([pr#13065](#), Barbora Ančincová)
- doc: minor change to a cloud testing paragraph ([pr#13277](#), Jan Fajerski)
- doc: minor fixes in radosgw/ ([pr#15103](#), Drunkard Zhang)
- doc: min_size advice is not helpful ([pr#12936](#), Brad Hubbard)
- doc: misc minor fixes ([pr#13713](#), Drunkard Zhang)
- doc: Modify Configuring Cinder section ([issue#18840](#), [pr#13400](#), Shinobu Kinjo)
- doc: op queue and mclock related options ([pr#16662](#), J. Eric Ivancich)
- doc: organizationmap: add Xianxia Xiao to Kylin Cloud team ([pr#12718](#), Yunchuan Wen)
- doc: PendingReleaseNotes: “ceph -w” behavior has changed drastically ([pr#16425](#), Joao Eduardo Luis, Nathan Cutler)
- doc: PendingReleaseNotes: notes on whiteouts vs pgnls ([pr#15575](#), Sage Weil)
- doc: PendingReleaseNotes: note the fuse fstab format change ([pr#13259](#), John Spray)
- doc: PendingReleaseNotes: recent cephfs changes ([pr#14196](#), John Spray)
- doc: PendingReleaseNotes: warning about ‘osd rm ...’ and #19119 ([issue#19119](#), [pr#13731](#), Sage Weil)
- doc: peoplemap: add pdonnell alias ([pr#14352](#), Patrick Donnelly)
- doc: radosgw-admin: new ‘global quota’ commands update period config ([issue#19409](#), [pr#14252](#), Casey Bodley)
- doc: README.FreeBSD: update current status ([pr#12096](#), Willem Jan Withagen)
- doc: README.FreeBSD: Update the status ([pr#14406](#), Willem Jan Withagen)
- doc: README.md: fix build instructions inconsistent ([pr#14555](#), Yao Zongyou)
- doc: README.md: use github heading syntax to mark the headings ([pr#14591](#), Kefu Chai)
- doc: release-notes clarify about rgw encryption ([pr#14800](#), Abhishek Lekshmanan)
- doc: release notes for v10.2.7 Jewel ([pr#14295](#), Abhishek Lekshmanan)
- doc: release notes for v11.1.1 ([pr#12642](#), Abhishek Lekshmanan)
- doc: release notes for v12.0.3 (dev) ([pr#15090](#), Abhishek Lekshmanan)
- doc: releases update the luminous, hammer, jewel release dates ([pr#13584](#), Abhishek Lekshmanan)
- doc: remove deprecated subcommand in man/8/ceph.rst ([pr#14928](#), Drunkard Zhang)
- doc: remove docs on non-existent command ([pr#16616](#), Luo Kexue, Kefu Chai)
- doc: remove duplicated references ([pr#13396](#), Kefu Chai)
- doc: remove mentions about mon_osd_min_down_reports ([issue#19016](#), [pr#13558](#), Barbora Ančincová)
- doc: remove some non-existent and fix the default value according to ... ([pr#15664](#), Leo Zhang)
- doc: Remove “splitting” state ([pr#12636](#), Brad Hubbard)
- doc: reword mds deactivate docs; add optional fs_name argument ([issue#20607](#), [pr#16471](#), Jan Fajerski)
- doc: Re-word the warnings about using git subtrees ([pr#14999](#), J. Eric Ivancich)
- doc: rgw clarify limitations when creating tenant names ([pr#16418](#), Abhishek Lekshmanan)
- doc: rgw: Clean up create subuser parameters ([pr#14335](#), hrchu)
- doc: rgw: correct get usage parameter default value ([pr#14372](#), hrchu)
- doc: rgw: Get user usage needs to specify user ([pr#14804](#), hrchu)
- doc: rgw: make a note abt system users vs normal users ([issue#18889](#), [pr#13461](#), Abhishek Lekshmanan)
- doc: rgw: note rgw_enable_usage_log option in adminops guide ([pr#14803](#), hrchu)
- doc: rgw: remove mention of megabytes for quotas ([pr#14413](#), Hans van den Bogert)
- doc: rgw: Rewrite Java swift examples ([pr#14268](#), Chu, Hua-Rong)
- doc: rgw: Rewrite the key management ([pr#14384](#), hrchu)
- doc: rgw server-side encryption and barbican ([pr#13483](#), Adam Kupczyk, Casey Bodley)
- doc: script: build-doc/serve-doc fixes ([pr#14438](#), Abhishek Lekshmanan)
- doc: script: ceph-release-notes: use https instead of http ([pr#14103](#), Kefu Chai)
- docs: doc/cephfs/troubleshooting: fix broken bullet list ([pr#12894](#), Dan Mick)
- docs: doc/dev: add some info about FreeBSD ([pr#14503](#), Willem Jan Withagen)
- docs: doc/release-notes: fix ceph-deploy command ([pr#15987](#), Sage Weil)
- docs: doc/release-note: update release-note ([pr#15748](#), liuchang0812)
- docs: document “osd recovery max single start” setting ([issue#17396](#), [pr#15275](#), Ken Dreyer)
- docs: mailmap: fix Zhao Chao affiliation ([pr#13413](#), Zhao Chao)
- docs: mailmap: Leo Zhang infomation and affiliation ([pr#15145](#), Leo Zhang)
- docs: mailmap: Liu Yang affiliation ([pr#13427](#), LiuYang)
- docs: mailmap: shiqi affiliation ([pr#14361](#), shiqi)
- docs: mailmap: update organization info ([pr#14747](#), liuchang0812)
- docs: mailmap: Weibing Zhang mailmap affiliation ([pr#15076](#), Weibing Zhang)
- docs: PendingReleaseNotes: mention forced recovery ([pr#16775](#), Piotr Dałek)
- docs: Remove contractions from the documentation ([pr#16629](#), John Wilkins)
- doc: style fix for doc/cephfs/client-config-ref.rst ([pr#14840](#), Drunkard Zhang)
- doc: tools/cephfs: fix cephfs-journal-tool –help ([pr#15614](#), John Spray)
- doc: two minor fixes ([pr#14494](#), Drunkard Zhang)
- doc: typo fixes on hyperlink/words ([pr#15144](#), Drunkard Zhang)
- doc: typo fix in s3_compliance ([pr#12598](#), LiuYang)
- doc: typo in hit_set_search_last_n ([pr#14108](#), Sven Seeberg)

- doc: Update adminops.rst (pr#13893, Chu, Hua-Rong)
- doc: update ceph(8) man page with new sub-commands (pr#16437, Kefu Chai)
- doc: Update CephFS disaster recovery documentation (pr#12370, Wido den Hollander)
- doc: Update disk thread section to reflect that scrubbing is no longer... (pr#12621, Nick Fisk)
- doc: update intro, quick start docs (pr#16224, Sage Weil)
- doc: Update keystone.rst (pr#12717, Chu, Hua-Rong)
- doc: update links to point to ceph/qa instead of ceph-qa-suite (pr#13397, Jan Fajerski, Nathan Cutler)
- doc: Update .organizationmap (pr#16507, luokexue)
- doc: update packages mentioned by build-doc and related doc (pr#14649, Yu Shengzuo)
- doc: Update sample.ceph.conf (pr#13751, Saumay Agrawal)
- doc: update sample explaining "%" operator in test suites (pr#15511, Kefu Chai)
- doc: Update some RGW documentation (pr#15175, Jens Rosenboom)
- doc: update the pool names created by vstart.sh by default (pr#16652, Zhu Shangzhong)
- doc: update the rados namespace docs (pr#15838, Abhishek Lekshmanan)
- doc: update the support status of swift static website (pr#13824, Jing Wenjun)
- doc: update the usage of 'ceph-deploy purge' (pr#15080, Yu Shengzuo)
- doc: update to new ceph fs commands (pr#13346, Patrick Donnelly)
- doc: upmap docs; various missing links for release notes (pr#16637, Sage Weil)
- doc: use do_cmake.sh instead of cmake .. (pr#15110, Kefu Chai)
- doc: v12.0.0 release notes (pr#13281, Abhishek Lekshmanan)
- doc: v12.0.2 (dev) release notes (pr#14625, Abhishek Lekshmanan)
- doc: v12.1.0 release notes notable changes addition again (pr#15857, Abhishek Lekshmanan)
- doc: various fixes (pr#16723, Kefu Chai)
- doc: vstart: add -help documentation for rgw_num (pr#13817, Ali Maredia)
- doc: wip-doc-multisite ports downstream multisite document upstream (pr#14259, John Wilkins)
- doc: Wip osd discussion docs (pr#13344, Greg Farnum)
- filestore: os/filestore: Exclude BTRFS on FreeBSD (pr#16171, Willem Jan Withagen)
- filestore: os/filestore/FileJournal: Fix typo in the comment (pr#14493, Zhou Zhengping)
- filestore: os/filestore: use existing variable for same func (pr#13742, Pan Liu)
- filestore: os/filestore: when print log, use __func__ instead of hard code function name (pr#15261, mychoxin)
- filestore: os/filestore: zfs add get_name() (pr#15650, Yanhu Cao)

Fix full testing in cephtool/test.sh when used by rados suite @Jing-Scott updated, addressing @rzarzynski's change request

- librados: add log channel to rados_monitor_log2 callback (pr#15926, Sage Weil)
- librados: add missing implementations for C service daemon API methods (pr#16543, Jason Dillaman)
- librados: add override for librados (issue#18922, pr#13442, liuchang0812)
- librados: add override in headers (pr#13775, liuchang0812)
- librados: asynchronous selfmanaged_snap_create/selfmanaged_snap_remove APIs (issue#16180, pr#12050, Jason Dillaman)
- librados: do not expose non-public symbols (pr#13265, Kefu Chai)
- librados: fix compile errors from simplified aio completions (pr#12849, xie xingguo)
- librados: fix rados_pool_list when buf is null (pr#14859, Sage Weil)
- librados: redirect balanced reads to acting primary when targeting object isn't recovered (issue#17968, pr#15489, Xuehan Xu)
- librados: remove legacy object listing API, clean up newer api (pr#13149, Sage Weil)
- librados: replace the var name from onack to complete (pr#13857, Pan Liu)
- librados: set the flag CEPH_OSD_FLAG_FULL_TRY of Op in the right place (pr#14193, Pan Liu)
- librados: use cursor for nobjects listing (pr#13323, Yehuda Sadeh, Sage Weil)
- librbd: add compare and write API (pr#14868, Zhengyong Wang, Jason Dillaman)
- librbd: add create timestamp metadata for image (pr#15757, runsisi)
- librbd: added rbd_flatten_with_progress to API (issue#15824, pr#12905, Ricardo Dias)
- librbd: add LIBRBD_SUPPORTS_WRITESAME support (pr#16583, Xiubo Li)
- librbd: add override keyword in header files (issue#19012, pr#13536, liuchang0812)
- librbd: add SnapshotNamespace to ImageCtx (pr#12970, Victor Denisov)
- librbd: add writesame API (pr#12645, Mingxin Liu, Gui Hecheng)
- librbd: allow to open an image without opening the parent image (issue#18325, pr#12885, Ricardo Dias)
- librbd: asynchronous clone state machine (pr#12041, Dongsheng Yang)
- librbd: asynchronous image removal state machine (pr#12102, Dongsheng Yang, Venky Shankar)
- librbd: avoid possible recursive lock when racing acquire lock (issue#17447, pr#12991, Jason Dillaman)
- librbd: changed the return type of ImageRequestWQ::discard() (issue#18511, pr#14032, Jos Collin)
- librbd: cleanup logging code under librbd/io (pr#14975, runsisi)
- librbd: corrected resize RPC message backwards compatibility (issue#19636, pr#14615, Jason Dillaman)
- librbd: create fewer empty objects during copyup (issue#15028, pr#12326, Douglas Fuller, Venky Shankar)
- librbd: deferred image deletion (issue#18481, pr#13105, Ricardo Dias)

- librbd: delay mirror registration when creating clones ([issue#17993](#), [pr#12839](#), Jason Dillaman)
- librbd: discard related IO should skip op if object non-existent ([issue#19962](#), [pr#15239](#), Mykola Golub)
- librbd: do not instantiate templates while building tests ([issue#18938](#), [pr#14891](#), Kefu Chai)
- librbd: do not raise an error if trash list returns -ENOENT ([pr#15085](#), runsisi)
- librbd: don't continue to remove an image w/ incompatible features ([issue#18315](#), [pr#12638](#), Dongsheng Yang)
- librbd: eliminate compiler warnings ([pr#13729](#), Jason Dillaman)
- librbd: fail IO request when exclusive lock cannot be obtained ([pr#15860](#), Jason Dillaman)
- librbd: filter expected error codes from is_exclusive_lock_owner ([issue#20182](#), [pr#15483](#), Jason Dillaman)
- librbd: fix clang compilation error ([issue#19260](#), [pr#13926](#), Mykola Golub)
- librbd: fixed initializer list ordering ([pr#13042](#), Jason Dillaman)
- librbd: fix issues with image removal state machine ([pr#15734](#), Jason Dillaman)
- librbd: fix rbd_metadata_list and rbd_metadata_get ([issue#19588](#), [pr#14471](#), Mykola Golub)
- librbd: fix segfault on EOPNOTSUPP returned while fetching snapshot timestamp ([issue#18839](#), [pr#13287](#), Gui Hecheng)
- librbd: fix valgrind errors and ensure tests detect future leaks ([pr#15415](#), Jason Dillaman)
- librbd: fix valid coverity warnings ([pr#14023](#), Jason Dillaman)
- librbd: image create validates that pool supports overwrites ([issue#19081](#), [pr#13986](#), Jason Dillaman)
- librbd: image-extent cache needs to clip out-of-bounds read buffers ([pr#13679](#), Jason Dillaman)
- librbd: Include WorkQueue.h since we use it ([issue#18862](#), [pr#13322](#), Boris Ranto)
- librbd: initialize diff parent overlap to zero ([pr#13077](#), Gu Zhongyan)
- librbd: introduce new constants for tracking max block name prefix ([issue#18653](#), [pr#13141](#), Jason Dillaman)
- librbd: is_exclusive_lock_owner API should ping OSD ([issue#19287](#), [pr#14003](#), Jason Dillaman)
- librbd: managed lock refactoring ([pr#12922](#), Mykola Golub)
- librbd: metadata_set API operation should not change global config setting ([issue#18465](#), [pr#12843](#), Mykola Golub)
- librbd: minor fixes for image trash move ([pr#14834](#), runsisi)
- librbd: new API method to force break a peer's exclusive lock ([issue#18429](#), [issue#16988](#), [issue#18327](#), [pr#12639](#), Jason Dillaman)
- librbd: Notifier::notify API improvement ([pr#14072](#), Mykola Golub)
- librbd: optimize copy-up to add hints only once to object op ([issue#19875](#), [pr#15037](#), Mykola Golub)
- librbd: pass an uint64_t to clip_io() as the third param ([issue#18938](#), [pr#14159](#), Kefu Chai)
- librbd: permit removal of image being bootstrapped by rbd-mirror ([issue#16555](#), [pr#12549](#), Mykola Golub)
- librbd: possible deadlock with flush if refresh in-progress ([issue#18419](#), [pr#12838](#), Jason Dillaman)
- librbd: potential read IO hang when image is flattened ([issue#19832](#), [pr#15234](#), Jason Dillaman)
- librbd: potential use of uninitialised value in ImageWatcher ([pr#14091](#), Mykola Golub)
- librbd: prevent self-blacklisting during break lock ([issue#18666](#), [pr#13110](#), Jason Dillaman)
- librbd: race initializing exclusive lock and configuring IO path ([pr#13086](#), Jason Dillaman)
- librbd: random unit test failures due to shut down race ([issue#19389](#), [pr#14166](#), Jason Dillaman)
- librbd: rbd ack cleanup ([pr#13791](#), runsisi)
- librbd: reacquire lock should update lock owner client id ([issue#19929](#), [pr#15093](#), Jason Dillaman)
- librbd: reduce potential of erroneous blacklisting on image close ([issue#19970](#), [pr#15162](#), Jason Dillaman)
- librbd: refactor exclusive lock support into generic managed lock ([issue#17016](#), [pr#12846](#), Ricardo Dias, Jason Dillaman)
- librbd: relax "is parent mirrored" check when enabling mirroring for pool ([issue#19130](#), [pr#13752](#), Mykola Golub)
- librbd: remove redundant check for image id emptiness ([pr#14830](#), runsisi)
- librbd: remove unnecessary dependencies of ManagedLock ([pr#12982](#), Jason Dillaman)
- librbd: remove unused rbd_image_options_t ostream operator ([pr#15443](#), Mykola Golub)
- librbd: resolve static analyser warnings ([pr#12863](#), Jason Dillaman)
- librbd: scatter/gather support for the C API ([issue#13025](#), [pr#13447](#), Jason Dillaman)
- librbd: silence -Wunused-variable warning ([pr#14953](#), Kefu Chai)
- librbd: simplify image open/close semantics ([pr#13701](#), Jason Dillaman)
- librbd: support for shared locking in ManagedLock ([pr#12886](#), Ricardo Dias)
- librbd: support to list snapshot timestamp ([issue#808](#), [pr#12817](#), Pan Liu)
- librbd: Uninitialized variable used handle_refresh() ([pr#16724](#), amitkuma)
- librbd: use 'override' keyword instead of 'virtual' ([issue#18922](#), [pr#13437](#), liuchang0812)
- librbd: warning message for mirroring pool option ([issue#18125](#), [pr#12319](#), Gaurav Kumar Garg)
- log: use one write system call per message ([pr#11955](#), Patrick Donnelly)
- mds: add authority check for delay dirfrag split ([issue#18487](#), [pr#12994](#), "Yan, Zheng")
- mds: add override in headers ([pr#13691](#), liuchang0812)
- mds: add override in mds subsystem ([issue#18922](#), [pr#13438](#), liuchang0812)
- mds: add perf counters for file system operations ([pr#14938](#), Michael Sevilla)
- mds: automate MDS object count tracking ([pr#13591](#), Patrick Donnelly)
- mds: bump client_reply debug to match client_req ([pr#14036](#), Patrick Donnelly)
- mds: change_attr++ and set ctime for set_vxattr ([issue#19583](#), [pr#14726](#), Patrick Donnelly)
- mds: change the type of data_pools ([pr#15278](#), Vicente Cheng)
- mds: check export pin during replay ([issue#20039](#), [pr#15205](#), Patrick Donnelly)
- mds: check for errors decoding backtraces ([issue#18311](#), [pr#12588](#), John Spray)
- mds: Client syncfs is slow (waits for next MDS tick) ([issue#20129](#), [pr#15544](#), dongdong tao)

- mds: don't assert on read errors in RecoveryQueue ([issue#19282](#), [pr#14017](#), John Spray)
- mds: don't modify inode that is not projected ([issue#16768](#), [pr#13052](#), "Yan, Zheng")
- mds: drop partial entry and adjust write_pos when opening PurgeQueue ([issue#19450](#), [pr#14447](#), "Yan, Zheng")
- mds: explicitly output error msg for dump cache asok command ([pr#15592](#), Zhi Zhang)
- mds: extend 'p' auth cap to cover all vxattr stuff ([issue#19075](#), [pr#13628](#), John Spray)
- mds: finish clientreplay requests before requesting active state ([issue#18461](#), [pr#12852](#), Yan, Zheng)
- mds: fix bad iterator dereference reported by coverity ([issue#18830](#), [pr#13272](#), John Spray)
- mds: fix CDir::merge() for mds_debug_auth_pins ([issue#19946](#), [pr#15130](#), "Yan, Zheng")
- mds: fix client ID truncation ([pr#15258](#), Henry Chang)
- mds: fix handling very fast delete ops ([issue#19245](#), [pr#13899](#), John Spray)
- mds: fix hangs involving re-entrant calls to journaler ([issue#20165](#), [pr#15430](#), John Spray)
- mds: fix incorrect assertion in Server::_dir_is_nonempty() ([issue#18578](#), [pr#12973](#), Yan, Zheng)
- mds: fix IO error handling in SessionMap ([pr#13464](#), John Spray)
- mds: fix mantle script to not fail for last rank ([issue#19589](#), [pr#14704](#), Patrick Donnelly)
- mds: fix mgrc shutdown ([issue#19566](#), [pr#14505](#), John Spray)
- mds: fix null pointer dereference in Locker::handle_client_caps ([issue#18306](#), [pr#12808](#), Yan, Zheng)
- mds: fix stray creation/removal notification ([issue#19630](#), [pr#14554](#), "Yan, Zheng")
- mds: fix use-after-free in Locker::file_update_finish() ([issue#19828](#), [pr#14991](#), "Yan, Zheng")
- mds: ignore ENOENT on writing backtrace ([issue#19401](#), [pr#14207](#), John Spray)
- mds: ignore fs full check for CEPH_MDS_OP_SETFILELOCK ([issue#18953](#), [pr#13455](#), "Yan, Zheng")
- mds: improvements for stray reintegration ([pr#15548](#), "Yan, Zheng")
- mds: include advisory path field in damage ([issue#18509](#), [pr#14104](#), John Spray)
- mds: issue new caps when sending reply to client ([issue#19635](#), [pr#14743](#), "Yan, Zheng")
- mds: limit client writable range increment ([issue#19955](#), [pr#15131](#), "Yan, Zheng")
- mds: make C_MDSInternalNoop::complete() delete 'this' ([issue#19501](#), [pr#14347](#), "Yan, Zheng")
- mds: mds perf item 'l_mdl_expos' always behind journaler ([pr#15621](#), redickwang)
- mds: miscellaneous fixes ([issue#18646](#), [pr#12974](#), Yan, Zheng, "Yan, Zheng")
- mds: miscellaneous multimds fixes ([issue#19022](#), [pr#13698](#), "Yan, Zheng")
- mds: miscellaneous multimds fixes part2 ([pr#15125](#), "Yan, Zheng")
- mds: miscellaneous multimds fixes ([pr#14550](#), "Yan, Zheng")
- mds: misc multimds fixes ([issue#18717](#), [issue#18754](#), [pr#13227](#), "Yan, Zheng")
- mds: misc multimds fixes part2 ([pr#12794](#), Yan, Zheng)
- mds: misc multimds fixes ([pr#12274](#), Yan, Zheng)
- mds,mon: Clean issues detected by cppcheck ([pr#13199](#), Ilya Shipitsin)
- mds: multimds flock fixes ([pr#15440](#), "Yan, Zheng")
- mds: Pass empty string to clear mantle balancer ([issue#20076](#), [pr#15282](#), Zhi Zhang)
- mds: pretty json from tell commands ([pr#14105](#), John Spray)
- mds: print rank as int ([issue#19201](#), [pr#13816](#), Patrick Donnelly)
- mds: propagate error encountered during opening inode by number ([issue#18179](#), [pr#12749](#), Yan, Zheng)
- mds: properly create aux subtrees for pinned directory ([issue#20083](#), [pr#15300](#), "Yan, Zheng")
- mds: relocate PTRWAITER put near get ([pr#14921](#), Patrick Donnelly)
- mds: remove boost::pool usage and use tcmalloc directly ([issue#18425](#), [pr#12792](#), Zhi Zhang)
- mds: remove legacy "mds tell" command ([issue#19288](#), [pr#14015](#), John Spray)
- mds: remove "mds log" config option ([issue#18816](#), [pr#14652](#), John Spray)
- mds: remove some redundant object counters ([pr#13704](#), Patrick Donnelly)
- mds: replace C_VoidFn in MDSDaemon with lambdas ([pr#13465](#), John Spray)
- mds: Return error message instead of asserting ([pr#14469](#), Brad Hubbard)
- mds: save projected path into inode_t::stray_prior_path ([issue#20340](#), [pr#15800](#), "Yan, Zheng")
- mds: set ceph-mds name uncond for external tools ([issue#19291](#), [pr#14021](#), Patrick Donnelly)
- mds: shut down finisher before objecter ([issue#19204](#), [pr#13859](#), John Spray)
- mds: skip fragment space check for replayed request ([issue#18660](#), [pr#13095](#), "Yan, Zheng")
- mds: support export pinning on directories ([issue#17834](#), [pr#14598](#), "Yan, Zheng", Patrick Donnelly)
- mds: try to avoid false positive heartbeat timeouts ([issue#19118](#), [pr#13807](#), John Spray)
- mds: use debug_mds for most subsys ([issue#19734](#), [pr#15052](#), Patrick Donnelly)
- mds: use same inode count in health check as in trim ([issue#19395](#), [pr#14197](#), John Spray)
- mds: warn if insufficient standbys exist ([issue#17604](#), [pr#12074](#), Patrick Donnelly)
- mgr: add a get_version to the python interface ([pr#13669](#), John Spray)
- mgr: add machinery for python modules to send MCommands to daemons ([pr#14920](#), John Spray)
- mgr: add mgr allow * to client.admin ([pr#14864](#), huanwen ren)
- mgr: add override in headers ([pr#13772](#), liuchang0812)
- mgr: add override in mgr subsystem ([issue#18922](#), [pr#13436](#), liuchang0812)
- mgr: add per-DaemonState lock ([pr#16432](#), Sage Weil)
- mgr: always free allocated MgrPyModule ([issue#19590](#), [pr#14507](#), Kefu Chai)
- mgr: ceph-create-keys: update client.admin if it already exists ([issue#19940](#), [pr#15112](#), John Spray)
- mgr: ceph: introduce "tell x help" subcommand ([issue#19885](#), [pr#15111](#), liuchang0812)

- mgr: ceph-mgr: Implement new pecan-based rest api ([pr#14457](#), Boris Ranto)
- mgr: ceph-mgr: rotate logs on sighup ([issue#19568](#), [pr#14437](#), Dan van der Ster)
- mgr: clean up daemon start process ([issue#20383](#), [pr#16020](#), John Spray)
- mgr: clean up fsstatus module ([pr#15925](#), John Spray)
- mgr: cleanup, stop clients sending in perf counters ([pr#15578](#), John Spray)
- mgr: cluster log message on plugin load error ([pr#15927](#), John Spray)
- mgr: dashboard code cleanup ([pr#15577](#), John Spray)
- mgr: dashboard GUI module ([pr#14946](#), John Spray, Dan Mick)
- mgr: dashboard improvements ([pr#16043](#), John Spray)
- mgr: do shutdown using finisher so we can do it in the right order ([issue#19743](#), [pr#14835](#), Kefu Chai)
- mgr: do the shutdown in the right order ([issue#19813](#), [pr#14952](#), Kefu Chai)
- mgr: drop repeated log info. and unnecessary write permission ([pr#15896](#), Yan Jun)
- mgr: enable ceph_send_command() to send pg command ([pr#15865](#), Kefu Chai)
- mgr: fix bugs in init, beacons ([issue#19516](#), [issue#19502](#), [pr#14374](#), Sage Weil)
- mgr: fix crash on missing 'ceph_version' in daemon metadata (fixes #18764) ([issue#18764](#), [pr#14129](#), Tim Serong)
- mgr: fix crash on set_config from python module with insufficient caps ([issue#19629](#), [pr#14706](#), Tim Serong)
- mgr: fix lock cycle ([pr#16508](#), Sage Weil)
- mgr: fix metadata handling from old MDS daemons ([pr#14161](#), John Spray)
- mgr: fix MgrStandby eating messages ([pr#15716](#), John Spray)
- mgr: fix python module teardown & add tests ([issue#19407](#), [issue#19412](#), [issue#19258](#), [pr#14232](#), John Spray)
- mgr: fix session leak ([issue#19591](#), [pr#14720](#), Sage Weil)
- mgr: fix several init/re-init bugs ([issue#19491](#), [pr#14328](#), Sage Weil)
- mgr: handle "module.set_config(.., None)" correctly ([pr#16749](#), Kefu Chai)
- mgr: increase debug level for ticks 0 -> 10 ([pr#16301](#), Dan Mick)
- mgr: load modules in separate python sub-interpreters ([pr#14971](#), Tim Serong)
- mgr: luminous: mgr: add missing call to pick_addresses ([issue#20955](#), [issue#21049](#), [pr#17173](#), John Spray)
- mgr: Make stats period configurable ([issue#17449](#), [pr#12732](#), liuchang0812)
- mgr: Mark session connections down on shutdown ([issue#19900](#), [pr#15192](#), Brad Hubbard)
- mgr: mgr/ClusterState: do not mangle PGMap outside of Incremental ([issue#20208](#), [pr#16262](#), Sage Weil)
- mgr: mgr/DaemonServer.cc: log daemon type string as well as id ([pr#15560](#), Dan Mick)
- mgr: mgr/dashboard: add OSD list view ([pr#16373](#), John Spray)
- mgr: mgr/dashboard: fix type error in get_rate function ([issue#20276](#), [pr#15668](#), liuchang0812)
- mgr: mgr/dashboard: load log lines on startup, split out audit log ([pr#15709](#), John Spray)
- mgr: mgr/MgrClient: fix reconnect event leak ([issue#19580](#), [pr#14431](#), Sage Weil)
- mgr: mgr/MgrStandby: prevent use-after-free on just-shut-down Mgr ([issue#19595](#), [pr#15297](#), Sage Weil)
- mgr: mgr/MgrStandby: respawn when deactivated ([issue#19595](#), [issue#19549](#), [pr#15557](#), Sage Weil)
- mgr: mgr_module interface to report health alerts ([pr#16487](#), Sage Weil)
- mgr: mgr,osd: ceph-mgr -help, unify usage text of other daemons ([pr#15176](#), Tim Serong)
- mgr: mgr/PyState: shut up about get_config on nonexistent keys ([pr#16641](#), Sage Weil)
- mgr: mgr/status: row has incorrect number of values ([issue#20750](#), [pr#16529](#), liuchang0812)
- mgr: Misc. bug fixes ([issue#18994](#), [pr#14883](#), John Spray)
- mgr: mkdir bootstrap-mgr ([pr#14824](#), huanwen ren)
- mgr: mon/mgr: add detail error infomation ([pr#16048](#), Yan Jun)
- mgr,mon: mgr,mon: debug init and mgrdigest subscriptions ([issue#20633](#), [pr#16351](#), Sage Weil)
- mgr: mon/MgrMonitor: fix standby addition to mgrmap ([issue#20647](#), [pr#16397](#), Sage Weil)
- mgr,mon: mon/AuthMonitor: generate bootstrap-mgr key on upgrade ([issue#20666](#), [pr#16395](#), Joao Eduardo Luis)
- mgr,mon: mon,mgr: extricate PGmap from monitor ([issue#20067](#), [issue#20174](#), [issue#20050](#), [pr#15073](#), Kefu Chai, Sage Weil, Greg Farnum)
- mgr,mon: mon/MgrMonitor: add 'mgr dump [epoch]' command ([pr#15158](#), Sage Weil)
- mgr,mon: mon/MgrMonitor: only propose if we updated ([pr#14645](#), Sage Weil)
- mgr,mon: mon/MgrMonitor: reset mgrdigest timer with new subscription ([issue#20633](#), [pr#16582](#), Sage Weil)
- mgr,mon: mon,mgr: move reweight-by-* to mgr ([pr#14404](#), Kefu Chai)
- mgr,mon: mon,mgr: print pgmap reports to debug (not cluster) log ([pr#15740](#), Sage Weil)
- mgr,mon: mon,mgr: trim osdmap without the help of pgmap ([pr#14504](#), Kefu Chai)
- mgr: move 'osd perf' and 'osd blocked-by' to mgr ([pr#14303](#), Sage Weil)
- mgr: move "osd pool stats" to mgr ([pr#14365](#), Kefu Chai)
- mgr: optimization some judgment and adjust the debug remove value in register_new_pgs ([pr#14046](#), song baisen)
- mgr: optimize DaemonStateIndex::cull() a little bit ([pr#14967](#), Kefu Chai)
- mgr: pass through cluster log to plugins ([pr#13690](#), John Spray)
- mgr: perf schema fns/change notification and Prometheus plugin ([pr#16406](#), Dan Mick)
- mgr: print a more helpful error message for when users lack mgr ceph caps ([issue#20296](#), [pr#15697](#), Greg Farnum)
- mgr,pybind: luminous: mgr/dashboard: fix duplicate images listed on iSCSI status page ([issue#21017](#), [pr#17282](#), Jason Dillaman)
- mgr: pybind/mgr/dashboard: bind to :: by default ([pr#16223](#), Sage Weil)
- mgr: pybind/mgr/dashboard: monkeypatch os.exit to stop cherrypy from taking down mgr ([issue#20216](#), [pr#15588](#),

Sage Weil

- mgr: pybind/mgr: Delete rest module ([pr#15429](#), John Spray)
- mgr: pybind/mgr/rest: completely terminate cherrypy in shutdown ([pr#14995](#), Tim Serong)
- mgr: pybind/mgr/rest: don't set timezone to Chicago ([pr#14184](#), Tim Serong)
- mgr: pybind/mgr/restful: improve cert handling; work with vstart ([pr#15405](#), Sage Weil)
- mgr: pybind/mgr/zabbix: fix health in non-compat mode ([issue#20767](#), [pr#16580](#), Sage Weil)
- mgr,pybind,rbd: mgr/dashboard: show rbd image features ([pr#16468](#), Yanhu Cao)
- mgr: raise python exception on failure in send_command() ([pr#15704](#), Kefu Chai)
- mgr,rbd: mgr/dashboard: RBD iSCSI daemon status page ([pr#16547](#), Jason Dillaman)
- mgr,rbd: mgr/dashboard: rbd mirroring status page ([pr#16360](#), Jason Dillaman)
- mgr,rbd: pybind/mgr/dashboard: initial block integration ([pr#15521](#), Jason Dillaman)
- mgr: redirect python stdout,stderr to ceph log ([pr#14189](#), Kefu Chai, Tim Serong, Dan Mick)
- mgr: release allocated PyString ([pr#14716](#), Kefu Chai)
- mgr: remove default cert; disable both restful and dashboard by default ([pr#15601](#), Boris Ranto, Sage Weil)
- mgr: remove non-existent MDS daemons from FSMap ([issue#17453](#), [pr#14937](#), Spandan Kumar Sahu)
- mgr: remove unused function declarations ([pr#14366](#), Wei Jin)
- mgr: rm nonused main function ([pr#14313](#), Wei Jin)
- mgr: shutdown py_modules in Mgr::shutdown() ([issue#19258](#), [pr#14078](#), Kefu Chai)
- mgr,tests: qa/suites: move mgr tests into rados suite ([pr#14687](#), John Spray)
- mgr,tests: qa/upgrade/jewel-x/point-to-point: add a mgr during final upgrade ([pr#15637](#), Sage Weil)
- mgr: use unique_ptr for MgrStandby::active_mngr ([pr#13667](#), John Spray)
- mgr: various cleanups ([pr#14802](#), Kefu Chai)
- mgr: vstart.sh: fix mgr vs restful command startup race ([pr#16564](#), Sage Weil)
- mgr: Zabbix monitoring module ([pr#16019](#), Wido den Hollander)
- misc: fix code typos in header files ([pr#12716](#), Xianxia Xiao)
- misc: kill clang warnings ([pr#14549](#), Kefu Chai)
- misc: Warning Elimination ([pr#14439](#), Adam C. Emerson)
- mon: add crush type down health warnings ([pr#14914](#), Neha Ojha)
- mon: added bootstrap-rbd auth profile ([pr#16633](#), Jason Dillaman)
- mon: add force-create-pg back ([issue#20605](#), [pr#16353](#), Kefu Chai)
- mon: add mgr metdata commands, and overall 'versions' command for all daemon versions ([pr#16460](#), Sage Weil)
- mon: add mon_debug_no_require_luminous ([pr#14490](#), Sage Weil)
- mon: Add override for FsNewHandler::handle() ([pr#15331](#), yonghengdexin735)
- mon: add override in headers ([pr#13693](#), liuchang0812)
- mon: add override in mon subsystem ([issue#18922](#), [pr#13440](#), liuchang0812)
- mon: add support public_bind_addr option ([pr#16189](#), Bassam Tabbara)
- mon: add warn info for osds were removed from osdmap but still kept in crushmap ([pr#12273](#), song baisen)
- mon: a few health fixes ([pr#16415](#), xie xingguo)
- mon: a few more upmap (and other) fixes ([pr#16239](#), xie xingguo)
- mon: avoid segfault in wait_auth_rotating ([issue#19566](#), [pr#14430](#), John Spray)
- mon: avoid start election twice when quorum enter ([pr#10150](#), song baisen)
- mon: check is_shutdown() in timer callbacks ([issue#19825](#), [pr#14919](#), Kefu Chai)
- mon: clean up in ceph_mon.cc ([pr#14102](#), huanwen ren)
- mon: clean up some osdmon/pgmon interactions ([pr#12403](#), Sage Weil)
- mon: cleanups ([pr#15272](#), Kefu Chai)
- mon: collect mon metdata as part of the election ([issue#20434](#), [pr#16148](#), Sage Weil)
- mon: common/config_opts.h: kill mon_pg_create_interval ([pr#13800](#), xie xingguo)
- mon: 'config-key put' -> 'config-key set' ([pr#16569](#), Sage Weil)
- mon: crush straw_calc_version value is 0 or 1 not 0 to 2 ([pr#13554](#), song baisen)
- mon: debug session feature tracking ([issue#20475](#), [pr#16128](#), Sage Weil)
- mon: delete unused config opts of mon_sync_fs_threshold ([pr#15676](#), linbing)
- mon: delete useless function definition ([pr#15188](#), shiqi)
- mon: detect existing fs and duplicate name earlier ([issue#18964](#), [pr#13471](#), Patrick Donnelly)
- mon: DIVIDE_BY_ZERO in PGMapDigest::dump_pool_stats_full() ([pr#15622](#), Jos Collin)
- mon: Division by zero in PGMapDigest::dump_pool_stats_full() ([pr#15901](#), Jos Collin)
- mon: do crushtool test with fork and timeout, but w/o exec of crushtool ([issue#19964](#), [pr#16025](#), Sage Weil)
- mon: do not dereference empty mgr_commands ([pr#16501](#), Sage Weil)
- mon: do not prime_pg_temp creating pgs; clean up pg create conditions ([issue#19826](#), [pr#14913](#), Sage Weil)
- mon: don't call propose_pending in prepare_update() ([issue#19738](#), [pr#14711](#), John Spray)
- mon: don't kill MDSs unless some beacons are getting through ([issue#19706](#), [pr#15308](#), John Spray)
- mon: don't prefix mgr summary with epoch number ([pr#15512](#), John Spray)
- mon: don't set last_osd_report when the pg stats msg is ignored ([pr#12975](#), Zhiqiang Wang)
- mon: drop useless assignment statements ([pr#13958](#), wangzhengyong)
- mon: emit cluster log messages on MDS health changes ([issue#19551](#), [pr#14398](#), John Spray)
- mon: enable luminous monmap feature on full quorum ([pr#13379](#), Joao Eduardo Luis)

- mon: extensible output format for health checks ([pr#16701](#), John Spray)
- mon: Filter *log last* output by severity and channel ([pr#15924](#), John Spray)
- mon: fix accesing pending_fsmap from peon ([issue#20040](#), [pr#15213](#), John Spray)
- mon: fix a few bugs with the osd health reporting ([pr#15179](#), Sage Weil)
- mon: fix a few nits ([pr#12670](#), Sage Weil)
- mon: Fix deep_age copy paste error ([pr#16434](#), Brad Hubbard)
- mon: Fixed typo in function comment blocks and in other comments ([pr#15304](#), linbing)
- mon: Fixed typo in @post of *_active()* ([pr#15191](#), Linbing)
- mon: fix force_pg_create pg stuck in creating bug ([issue#18298](#), [pr#12539](#), Sage Weil)
- mon: fix hang on deprecated/removed ‘pg set_full_ratio’ commands ([issue#20600](#), [pr#16300](#), Sage Weil)
- mon: fix hiding mdsmonitor informative strings ([issue#16709](#), [pr#13904](#), John Spray)
- mon: fix kvstore type in mon compact command ([pr#15954](#), liuchang0812)
- mon: fix legacy health checks in ‘ceph status’ during upgrade; fix jewel-x upgrade combo ([pr#17176](#), Sage Weil)
- mon: fix mon_keyvaluedb application ([pr#15059](#), Sage Weil)
- mon: Fix output text and doc ([pr#16367](#), Yan Jun)
- mon: fix prime_pg_temp overrun ([issue#19874](#), [pr#14979](#), Sage Weil)
- mon: Fix status output warning for mon_warn_osd_usage_min_max_delta ([issue#20544](#), [pr#16220](#), David Zafman)
- mon: fix synchronise pgmap with others ([pr#14418](#), song baisen, z09440)
- mon: fix wrongly delete routed pgstats op ([issue#18458](#), [pr#12784](#), Mingxin Liu)
- mon: fix wrong mon-num counting logic of ‘ceph features’ command ([pr#17172](#), xie xingguo)
- mon: handle cases where store->get() may return error ([issue#19601](#), [pr#14678](#), Jos Collin)
- mon: include device class in tree view; hide shadow hierarchy ([pr#16016](#), Sage Weil)
- mon: Incorrect expression in PGMap::get_health() ([pr#15648](#), Jos Collin)
- mon: in output of “ceph osd df tree”, display “-”, not “0”, for pg amount of a bucket ([pr#13015](#), Chuanhong Hong)
- mon: it’s no need to get pg action_primary osd twice in pg scrub ([pr#15313](#), linbing)
- mon: ‘* list’ -> ‘* ls’ ([pr#16423](#), Sage Weil)
- mon: load mgr commands at runtime ([pr#16028](#), John Spray, Sage Weil)
- mon: logclient: use the seq id of the 1st log entry when resetting session ([issue#19427](#), [pr#14927](#), Kefu Chai)
- mon: Log errors at startup ([issue#14088](#), [pr#15723](#), Brad Hubbard)
- mon: luminous: mon/MonCommands: fix copy-and-paste error ([pr#17274](#), xie xingguo)
- mon: maintain the “cluster” PerfCounters when using ceph-mgr ([issue#20562](#), [pr#16249](#), Greg Farnum)
- mon: mark *osd create* as deprecated ([pr#15641](#), Joao Eduardo Luis)
- mon: mon/crush: create crush rules using device classes for replicated and ec pools via cli ([pr#16027](#), Sage Weil)
- mon: mon/HealthMonitor: avoid sending unnecessary MMonHealthChecks to leader ([pr#16478](#), xie xingguo)
- mon: mon/HealthMonitor: trigger a proposal if stat updated ([pr#16477](#), Kefu Chai)
- mon: mon/LogMonitor: don’t read list’s end() for log last ([pr#16376](#), Joao Eduardo Luis)
- mon: mon/MDSMonitor: close object section of formatter ([pr#16516](#), Chang Liu)
- mon: mon/MDSMonitor: remove create_new_fs from header ([pr#14019](#), Henrik Korkuc)
- mon: mon/MgrMonitor: only induce mgr epoch shortly after mkfs ([pr#16356](#), Sage Weil)
- mon: mon/MgrMonitor: send digests only if *is_active()* ([pr#15109](#), Kefu Chai)
- mon: mon/MgrStatMonitor: do not crash on luminous dev version upgrades ([pr#16287](#), Sage Weil)
- mon: mon/MonClient: cancel pending commands on shutdown ([issue#20051](#), [pr#15227](#), Kefu Chai, Sage Weil)
- mon: mon/MonClient: make *get_mon_log_message()* atomic ([issue#19427](#), [pr#14422](#), Kefu Chai)
- mon: mon/MonClient: random all ranks then pick first_n ([pr#13479](#), Mingxin Liu)
- mon: mon/Monitor.h: add const to member function ([pr#10412](#), Michal Jarzabek)
- mon: mon/Monitor: recreate mon session if features changed ([issue#20433](#), [pr#16230](#), Joao Eduardo Luis)
- mon: {mon,osd,mds} {versions,count-metadata} ([pr#15436](#), Sage Weil)
- mon: mon/OSDMonitor: a couple of upmap and other fixes ([pr#15917](#), xie xingguo)
- mon: mon/OSDMonitor: check *get()*’s return value instead of bl’s length ([pr#14805](#), Kefu Chai)
- mon: mon/OSDMonitor: check *last_osd_report* only when the whole cluster is lu... ([pr#14294](#), Kefu Chai)
- mon: mon/OSDMonitor: Clean up: delete extra S signature for plural ([pr#14174](#), Shinobu Kinjo)
- mon: mon/OSDMonitor: cleanup pending_created_pgs after done with it ([pr#14898](#), Kefu Chai)
- mon: mon/OSDMonitor: do not alter the “created” epoch of a pg ([issue#19787](#), [pr#14849](#), Kefu Chai)
- mon: mon/OSDMonitor: ensure UP is not set for newly-created OSDs ([issue#20751](#), [pr#16534](#), Sage Weil)
- mon: mon/OSDMonitor: fix dividing by zero in OSDUtilizationDumper ([pr#13531](#), Mingxin Liu)
- mon: mon/OSDMonitor: fix output func name in *can_mark_out* ([pr#14758](#), xie xingguo)
- mon: mon/OSDMonitor: fix process osd failure ([pr#12938](#), Mingxin Liu)
- mon: mon/OSDMonitor: guard ‘osd crush set-device-class’ ([pr#16217](#), Sage Weil)
- mon: mon/OSDMonitor: increase *last_epoch_clean*’s lower bound if possible ([pr#14855](#), Kefu Chai)
- mon: mon/OSDMonitor: issue pool application related warning ([pr#16520](#), xie xingguo)
- mon: mon/OSDMonitor: “osd crush class rename” support ([pr#15875](#), xie xingguo)
- mon: mon/OSDMonitor: remove trivial PGMap dependency for ‘osd primary-temp’ command ([pr#13616](#), Sage Weil)
- mon: mon/OSDMonitor: remove zeroed new_state updates ([issue#20751](#), [pr#16518](#), Sage Weil)
- mon: mon/OSDMonitor: sanity check osd before performing ‘osd purge’ ([pr#16838](#), xie xingguo)
- mon: mon/OSDMonitor: some cleanup for reweight-by-pg ([pr#13462](#), Haodong Tang)

- mon: mon/OSDMonitor: spinlock -> std::mutex ([pr#14269](#), Sage Weil)
- mon: mon/OSDMonitor: tolerate upgrade from post-kraken dev cluster ([pr#14442](#), Sage Weil)
- mon: mon/OSDMonitor: transit creating_pgs from pgmap when upgrading ([issue#19584](#), [pr#14551](#), Kefu Chai)
- mon: mon/OSDMonitor: two pool opts related fix ([pr#15968](#), xie xingguo)
- mon: mon/OSDMonitor: update creating epoch if target osd changed ([issue#19515](#), [pr#14386](#), Kefu Chai)
- mon: mon/OSDMonitor: update creating_pgs using pending_creatings ([issue#19814](#), [pr#14897](#), Kefu Chai)
- mon: mon/OSDMonitor: update pg_creatings even the new acting set is empty ([issue#19744](#), [pr#14730](#), Kefu Chai)
- mon: mon/PaxosService: use `_func_` instead of hard code function name ([pr#15863](#), Yanhu Cao)
- mon: mon/PGMap: add up_primary pg number field for pg-dump cmd ([pr#13451](#), xie xingguo)
- mon: mon/PGMap.cc: fix “osd_epochs” section of dump_basic ([pr#14996](#), xie xingguo)
- mon: mon/PGMap: make si units more readable in PGMap summary ([pr#14185](#), liuhong)
- mon: mon/PGMap: remove skewed utilization warning ([issue#20730](#), [pr#16461](#), Sage Weil)
- mon: mon/PGMap: show %used in formatted output ([issue#20123](#), [pr#15387](#), Joao Eduardo Luis)
- mon: mon/PGMonitor: clean up min/max span warning ([pr#14611](#), Sage Weil)
- mon: mon/PGMonitor: fix description for ceph pg ls ([pr#12807](#), runsisi)
- mon: mon/PGMonitor: rm nonused function ([pr#14033](#), Wei Jin)
- mon: move ‘pg map’ to OSDMonitor ([pr#14559](#), Sage Weil)
- mon: no delay for single message MSG_ALIVE and MSG_PGTEMP ([pr#12107](#), yaoning)
- mon: optracker’s initiated_at timestamp should not be NULL ([pr#12826](#), Mingxin Liu)
- mon: osd crush set crushmap need sanity check ([issue#19302](#), [pr#14029](#), Loic Dachary)
- mon: OSDMonitor add check only concern our self cluster command ([pr#10309](#), song baisen)
- mon: mon/OSDMonitor: add plain output for “crush class ls-osd” command ([pr#17230](#), xie xingguo)
- mon: mon/OSDMonitor: check creating_pgs.last_scan_epoch instead when sending creates ([issue#20785](#), [pr#17257](#), Kefu Chai)
- mon: OSDMonitor: check mon_max_pool_pg_num when set pool pg_num ([pr#16511](#), chenhg)
- mon: mon/OSDMonitor: do not send_pgCreates with stale info ([issue#20785](#), [pr#17191](#), Kefu Chai)
- mon: mon/OSDMonitor: fix improper input/testing range of crush somke testing ([pr#17232](#), xie xingguo)
- mon: mon/PGMonitor: always update pgmap with latest osdmap ([issue#19398](#), [pr#14777](#), Kefu Chai)
- mon: mon/pgmap: add objects prefix for unfound type ([issue#21127](#), [pr#17264](#), huanwen ren, Sage Weil)
- mon: mon/PGMap: fix “0 stuck requests are blocked > 4096 sec” warn ([pr#17215](#), xie xingguo)
- mon: mon/PGMonitor add check only concern our self cluster command ([pr#9976](#), song baisen)
- mon: mon/post-jewel cleanups ([pr#13150](#), Kefu Chai)
- mon: mon/prime_pg_temp and a few health warning fixes ([pr#16530](#), xie xingguo)
- mon: mon/refactor MDSMonitor command handling ([pr#13581](#), John Spray)
- mon: mon/Removed unnecessary function declaration in MDSMonitor.h ([pr#15374](#), yonghengdexin735)
- mon: mon/remove the redundant judgement in paxosservice is_writeable function ([pr#10240](#), song baisen)
- mon: mon/remove unnecessary function declaration ([pr#13762](#), liuchang0812)
- mon: mon/replace osds with osd destroy and osd new ([pr#14074](#), Joao Eduardo Luis, Sage Weil)
- mon: mon/restructure prime_pg_temp around a full pg mapping calculated on multiple CPUs ([pr#13207](#), Sage Weil)
- mon: mon/revamp health check/warning system ([pr#15643](#), John Spray, Sage Weil)
- mon: mon/revise “ceph status” output ([pr#15396](#), John Spray)
- mon: mon/show class in ‘osd crush tree’ output; sort output ([pr#16740](#), Sage Weil)
- mon: mon/show destroyed status in tree view; do not auto-out destroyed osds ([pr#16446](#), xie xingguo)
- mon: mon/show inactive % in ceph status ([pr#14810](#), Sage Weil)
- mon: mon/show io status quickly if no update in a long period ([pr#14176](#), Mingxin Liu)
- mon: mon/show the leader info on mon stat command ([pr#14178](#), song baisen)
- mon: mon/skip crush smoke test when running under valgrind ([issue#20602](#), [pr#16346](#), Sage Weil)
- mon: mon/smooth io/recovery stats over longer period ([pr#13249](#), Sage Weil)
- mon: mon/stop issuing not-[deep]-scrubbed warnings if disabled ([pr#16465](#), xie xingguo)
- mon: mon/support pool application metadata key/values ([pr#15763](#), Jason Dillaman)
- mon: mon/qa/suites: add test exercising workunits/mon/auth_caps.sh ([pr#15754](#), Kefu Chai)
- mon: mon/tests: test: Initialize pointer msg in MonClientHelper ([pr#16784](#), amitkuma)
- mon: mon/Tidy up removal of debug mon features ([pr#14467](#), Brad Hubbard)
- mon: mon/track features from connect clients, and use it to gate set-require-min-compat-client ([pr#15371](#), Sage Weil)
- mon: mon/trim the creating_pgs after updating it with pgmap ([issue#20067](#), [pr#15318](#), Kefu Chai)
- mon: mon/update mgrmap when active goes offline ([issue#19407](#), [pr#14220](#), John Spray)
- mon: mon/Update OSDMon.cc comments ([pr#13750](#), Saumay Agrawal)
- mon: mon/warn about using osd new instead of osd create ([pr#17302](#), Neha Ojha)
- msg: msg/allow different ms type for cluster network and public network ([pr#12023](#), Haomai Wang)
- msg: msg/always set header.version in encode_payload() ([issue#19939](#), [pr#16421](#), Kefu Chai)
- msg: msg/client bind ([pr#12901](#), Zengran Zhang, Haomai Wang)
- msg: msg/do not enable client-side binding by default ([issue#20049](#), [pr#15392](#), Jason Dillaman)
- msg: msg/don’t set msgr addr when disabling client bind ([pr#15243](#), Haomai Wang)
- msg: msg/end parameter in entity_addr_t::parse is optional ([pr#13650](#), Mykola Golub)
- msg: msg/Fix calls to Messenger::create with new parameter ([pr#13329](#), Sarit Zubakov)

- msg: Increase loglevels on some messages ([pr#14707](#), Willem Jan Withagen)
- msg: Initialize member variables in Infiniband ([pr#16781](#), amitkuma)
- msg: Initializing uninitialized members MMonGetVersion ([pr#16811](#), amitkuma)
- msg: Initializing uninitialized members MMonGetVersionReply ([pr#16813](#), amitkuma)
- msg: Initializing uninitialized members MMonPaxos ([pr#16814](#), amitkuma)
- msg: Initializing uninitialized members MMonProbe ([pr#16815](#), amitkuma)
- msg: Initializing uninitialized members module messages ([pr#16817](#), amitkuma)
- msg: Initializing uninitialized members MOSDALive ([pr#16816](#), amitkuma)
- msg: make listen backlog an option, increase from 128 to 512 ([issue#20330](#), [pr#15743](#), Haomai Wang)
- msg: messages: coverity fixes ([pr#13473](#), Kefu Chai)
- msg: msg/async: avoid atomic variable overhead ([pr#12809](#), Wei Jin)
- msg: msg/async: cleanup code ([pr#13304](#), Jianpeng Ma)
- msg: msg/async: cleanups ([pr#12832](#), Wei Jin)
- msg: msg/async: fix file description leak in NetHandler ([pr#13271](#), liuchang0812)
- msg: msg/async: increase worker reference with local listen table enabled backend ([issue#20390](#), [pr#15897](#), Haomai Wang)
- msg: msg/async: Lower down the AsyncMessenger's standby warning from debug ([pr#15242](#), Pan Liu)
- msg: msg/AsyncMessenger: remove unused method ([pr#10125](#), Michal Jarzabek)
- msg: msg/async/net_handler: errno should be stored before calling next function ([pr#14985](#), Zhou Zhengping)
- msg: msg/async/rdma: check if exp verbs avail ([pr#13391](#), Oren Duer, Adir Lev)
- msg: msg/async/rdma: check if fin message completed ([pr#15624](#), Alexander Mikheev, Adir Lev)
- msg: msg/async/rdma: Data path fixes ([pr#15903](#), Adir lev)
- msg: msg/async/rdma: Debug prints for ibv* ([pr#14249](#), Amir Vadai)
- msg: msg/async/rdma: Device::last_poll_dev must be positive ([pr#14250](#), Amir Vadai)
- msg: msg/async/rdma: Fix broken compilation ([pr#13603](#), Sarit Zubakov)
- msg: msg/async/rdma: Fix memory leak of OSD ([pr#13101](#), Sarit Zubakov)
- msg: msg/async/rdma: fix outstanding queuepair when destruct RDMAStack ([pr#13905](#), Haomai Wang)
- msg: msg/async/rdma: fix RoCE v2 deafault value ([pr#12648](#), Adir Lev, Oren Duer)
- msg: msg/async/rdma: Fix small memory leaks detected by valgrind ([pr#14288](#), Amir Vadai)
- msg: msg/async/rdma: handle buffers after close msg ([pr#15749](#), DanielBar-On, Alexander Mikheev, Adir Lev)
- msg: msg/async/rdma: move active_queue_pairs perf counter dec to polling ([pr#13716](#), DanielBar-On)
- msg: msg/async/rdma: Print error only on ENOMEM ([pr#13538](#), Sarit Zubakov)
- msg: msg/async/rdma: RDMA-CM, Pass specific ConnMgr info in constructor ([pr#14409](#), Amir Vadai)
- msg: msg/async/rdma: register buffer as continuous ([pr#15967](#), Adir Lev)
- msg: msg/async/rdma: remove assert from ibv_dealloc_pd in ProtectionDomain ([pr#15832](#), DanielBar-On)
- msg: msg/async/rdma: update destructor message ([pr#13539](#), Sarit Zubakov)
- msg: msg/async/rdma: zero wqe inline ([pr#13392](#), Adir Lev)
- msg: msg/async: remove false alert "assert" ([pr#15288](#), Haomai Wang)
- msg: msg/async: remove useless close function ([pr#13286](#), liuchang0812)
- msg: msg/async: rm nonused thread variable in posixworker ([pr#12777](#), Wei Jin)
- msg: msg/async: use auto iterator having more simple code and good performance ([pr#16524](#), dingdangzhang)
- msg: msg/Messenger.cc: add std::move ([pr#9760](#), Michal Jarzabek)
- msg: msg/MOSDOpReply: fix missing trace decode ([pr#15999](#), Yan Jun)
- msg: msg/RDMA: Fix broken compilation due to new argument in net.connect() ([pr#13096](#), Amir Vadai)
- msg: msg/simple: Remove dead code in pipe.cc ([issue#12684](#), [pr#12601](#), Rishabh Kumar)
- msg: msg/simple: use my addr when setting sock priority ([issue#19801](#), [pr#14878](#), Kefu Chai)
- msg: no need to pass supported features to Messenger::Policy ctor ([pr#13785](#), Sage Weil)
- msg: QueueStrategy::wait() joins all threads ([issue#20534](#), [pr#16194](#), Casey Bodley)
- msg: Remove unused variable perf_counter in RDMAStack ([pr#16783](#), amitkuma)
- msg: Revert the change from assert(0)-> ceph_abort() where is not applicable ([pr#12930](#), Dave Chen)
- msg: src/msg/async/AsyncConnect.cc: Use of sizeof() on a Pointer Type ([pr#14773](#), Svyatoslav)
- msg: src/msg/async: Update fix broken compilation for Posix ([pr#14336](#), Sarit Zubakov)
- msg: src/msg/simple/Pipe.cc: Fix the inclusion of '}' ([pr#14843](#), Willem Jan Withagen)
- os/bluestore: print leaked extents to debug output ([pr#17303](#), Sage Weil)
- osd: add asock command to dump the scrub queue ([issue#17861](#), [pr#12728](#), liuchang0812)
- osd: add default_device_class to metadata ([pr#16634](#), Neha Ojha)
- osd: add dump filter for tracked ops ([pr#16561](#), Yan Jun)
- osd: add "heap *" admin command ([issue#15475](#), [pr#13073](#), Jesse Williamson)
- osd: adding PerfCounters for backoff throttle ([pr#13017](#), Chuanhong Wang)
- osd: add is_split check before _start_split ([pr#13307](#), song baisen)
- osd: add override in headers files ([pr#13962](#), liuchang0812)
- osd: add override in osd subsystem ([issue#18922](#), [pr#13439](#), liuchang0812)
- osd: Add recovery sleep configuration option for HDDs and SSDs ([pr#16328](#), Neha Ojha)
- osd: add snap trim reservation and re-implement osd_snap_trim_sleep ([pr#13594](#), Samuel Just)
- osd: adjust osd_min_pg_log_entries ([issue#21026](#), [pr#17202](#), J. Eric Ivancich)

- osd: allow client throttler to be adjusted on-fly, without restart ([issue#18791](#), [pr#13213](#), Piotr Dałek)
- osd: bail from _committed_osd_maps inside osd_lock ([issue#20273](#), [pr#15710](#), Sage Weil)
- osd: Calculate degraded and misplaced more accurately ([issue#18619](#), [pr#13031](#), David Zafman)
- osd: change a few messages at level 0 and 1; change default level to 1/5 ([pr#13407](#), Sage Weil)
- osd: Check for and automatically repair object info soid during scrub ([issue#20471](#), [pr#16052](#), David Zafman)
- osd: check fsid is normal before osd mkfs ([pr#13898](#), song baisen)
- osd: check queue_transaction return value ([pr#15873](#), zhanglei)
- osd: Check snapset for validity when selecting authoritative shard ([issue#20186](#), [issue#18409](#), [pr#15559](#), David Zafman)
- osd: Check whether journal is rotational or not ([pr#16614](#), Neha Ojha)
- osd: clarify REQUIRE_LUMINOUS error message ([pr#13363](#), Josh Durgin)
- osd: clean nonused work queue ([pr#14990](#), Wei Jin)
- osd: Cleanup-Updated OSDMap.cc with C++11 style range-for loops ([pr#14381](#), Jos Collin)
- osd: cleanup: use string & to avoid unnecessary copy ([pr#12336](#), Yunchuan Wen)
- osd: clear_queued_recovery() in on_shutdown() ([issue#20432](#), [pr#16093](#), Kefu Chai)
- osd: cmpext operator should ignore -ENOENT on read ([pr#16622](#), Jason Dillaman)
- osd: combine conditional statements ([pr#16391](#), Yan Jun)
- osd: combine unstable stats with info.stats when publish stats to osd ([pr#14060](#), Mingxin Liu)
- osd: compact osd feature ([issue#19592](#), [pr#16045](#), liuchang0812)
- osd: condition object_info_t encoding on required (not up) features ([issue#18644](#), [pr#13114](#), Sage Weil)
- osd: constify OpRequest::get_req(); fix a few cases of operator<< vs mutated message races ([pr#13545](#), Sage Weil)
- osd: correct comment of perfcounter cached_crc in code ([pr#13256](#), lvshuhua)
- osd: correct epoch setting of osd boot msg ([pr#12623](#), Mingxin Liu)
- osd: correct the func name in execute_ctx() log messages ([pr#13582](#), Gu Zhongyan)
- osd: Corrupt objects stop snaptrim and mark pg snaptrim_error ([issue#13837](#), [pr#15635](#), David Zafman)
- osd: debug con in ms_handle_connect ([pr#13540](#), Sage Weil)
- osd: do not send ENXIO on misdirected op by default ([issue#18751](#), [pr#13206](#), Sage Weil)
- osd: do not send pg_created unless luminous ([issue#20785](#), [pr#16677](#), Kefu Chai)
- osd: do not try to boot until we've seen the first osdmap ([pr#15732](#), Sage Weil)
- osd: do not try to set device class before luminous ([issue#20850](#), [pr#16706](#), Josh Durgin)
- osd: don't leak pgrefs or reservations in SnapTrimmer ([issue#19931](#), [pr#15214](#), Greg Farnum)
- osd: don't share osdmap with objecter when preboot ([issue#15025](#), [pr#13946](#), Mingxin Liu)
- osd: don't use ORDERSNAP for flush; always request/send ondisk ack ([issue#18961](#), [pr#13570](#), Samuel Just)
- osd: drop support for listing objects at a given snap ([pr#13398](#), Sage Weil)
- osd: dump the field name of object watchers and cleanups ([pr#15946](#), Yan Jun)
- osd: EC read handling: don't grab an objectstore error to use as the read error ([pr#16663](#), David Zafman)
- osd: eliminate snapdir objects and move clone snaps vector into SnapSet ([pr#13610](#), Sage Weil)
- osd: Execute crush_location_hook when configured in ceph.conf ([pr#15951](#), Wido den Hollander)
- osd: _exit() instead of exit() for failure injection ([issue#18372](#), [pr#12726](#), Sage Weil)
- osd: extend OMAP_GETKEYS and GETVALS to include a 'more' output field ([pr#12950](#), Sage Weil)
- osd: fall back to failsafe threshold if osdmap doesn't set [near]full ([pr#14004](#), Sage Weil)
- osd: faster dispatch ([pr#13343](#), Sage Weil)
- osd: fix a couple bugs with persisting the missing set when it contains deletes ([issue#20704](#), [pr#16459](#), Josh Durgin)
- osd: fix argument-dependent lookup of swap() ([pr#15124](#), Casey Bodley)
- osd: fix a signed/unsigned warning in PG ([pr#13922](#), Greg Farnum)
- osd: fix comments about pg refs and lock ([pr#14279](#), tang.jin)
- osd: fix coverity warning for uninitialized members ([pr#12724](#), Li Wang)
- osd: fix func name in log produced by handle_pg_peering_evt() ([pr#13801](#), xie xingguo)
- osd: fix occasional MOSDMap leak ([issue#18293](#), [pr#14558](#), Sage Weil)
- osd: fix OpRequest and tracked op dump information ([pr#16504](#), Yan Jun)
- osd: fix past_intervals base case by adding epoch_pool_created to pg_history_t ([issue#19877](#), [pr#14989](#), Sage Weil)
- osd: fix pg ref leaks when osd shutdown ([issue#20684](#), [pr#16408](#), Yang Honggang)
- osd: fix some osd beacon bugs ([pr#14274](#), Sage Weil)
- osd: fix stat sum update of recovery pushing ([pr#13328](#), Zhiqiang Wang)
- osd: fix the setting of soid in sub_op_push ([pr#13353](#), Zhiqiang Wang)
- osd: fix typo in comment ([pr#13061](#), Gu Zhongyan)
- osd: Fix useless MAX(0, unsigned) to prevent out of wack misplaced ([issue#18718](#), [pr#13164](#), David Zafman)
- osd: have clients resend ops on pg split ([pr#13235](#), Sage Weil)
- osd: hdd vs ssd defaults for osd op thread pool ([pr#15422](#), Sage Weil)
- osd: heartbeat with packets large enough to require working jumbo frames ([issue#20087](#), [pr#15535](#), Greg Farnum)
- osd: Implement asynchronous recovery sleep ([pr#15212](#), Neha Ojha)
- osd: Implement asynchronous scrub sleep ([issue#19497](#), [pr#14886](#), Brad Hubbard)
- osd: Implement peering state timing ([pr#14627](#), Brad Hubbard)
- osd: improve error message when FileStore op fails due to EPERM ([issue#18037](#), [pr#12181](#), Nathan Cutler)
- osd: initialize waiting_for_pg_osdmap on startup ([issue#20748](#), [pr#16535](#), Sage Weil)

- osd: kill all remaining MOSDSubOp users ([pr#13401](#), Sage Weil)
- osd: kill sortbitwise ([pr#13321](#), Sage Weil)
- osd: Log audit ([pr#16281](#), Brad Hubbard)
- osd: make ec overwrites ready to use ([pr#14496](#), Josh Durgin)
- osd: moved OpFinisher logic from OSDOp to OpContext ([issue#20783](#), [pr#16617](#), Jason Dillaman)
- osd: Move scrub sleep timer to osdservice ([issue#19986](#), [pr#15217](#), Brad Hubbard)
- osd: never send rados ack (only commit) ([pr#12451](#), Sage Weil)
- osd: new op for calculating an extent checksum ([pr#14256](#), Jason Dillaman)
- osd: objclass sdk ([pr#14723](#), Neha Ojha)
- osd: Object level shard errors are tracked and used if no auth available ([issue#20089](#), [pr#15397](#), David Zafman)
- osd: On EIO from read recover the primary replica from another copy ([issue#18165](#), [pr#14760](#), David Zafman)
- osd: osdc/ObjectCacher: use state instead of get_state() ([pr#12544](#), huangjun)
- osd: osdc/Objecter: more constness ([pr#14819](#), Kefu Chai)
- osd: osdc: silence warning from -Wsign-compare ([pr#14729](#), Jos Collin)
- osd: osd does not use MPing Messages, do not include unused include ([pr#15833](#), linbing)
- osd: osd/OSDMap.cc: check if osd is out in subtree_type_is_down ([issue#19989](#), [pr#15250](#), Neha Ojha)
- osd: osd/OSDMap: require OSD features only of OSDs ([issue#18831](#), [pr#13275](#), Ilya Dryomov)
- osd: osd/PrimaryLogPG: nullptr not NULL ([pr#13973](#), Shinobu Kinjo)
- osd: 'osd tree inout|up|down' to filter tree results ([pr#15294](#), Sage Weil)
- osd: os/kstore: some error handling ([pr#13960](#), wangzhengyong)
- osd/PGBBackend: delete reply if fails to complete delete request ([issue#20913](#), [pr#17233](#), Kefu Chai)
- osd: pg: be more careful with locking around forced pg recovery ([issue#20808](#), [pr#16712](#), Greg Farnum)
- osd: pglog trimming fixes ([pr#12882](#), Zhiqiang Wang)
- osd: pglog: with config, don't assert in the presence of stale diverg... ([issue#17916](#), [pr#14648](#), Greg Farnum)
- osd: pg-remap -> pg-upmap ([pr#14556](#), Sage Weil)
- osd: populate last_epoch_split during build_initial_pg_history ([issue#20754](#), [pr#16519](#), Sage Weil)
- osd: Preserve OSDOp information for historic ops ([pr#15265](#), Guo-Fu Tseng)
- osd: PrimaryLogPG, PGBBackend: complete callback even if interval changes ([issue#20747](#), [pr#16536](#), Josh Durgin)
- osd: print pg_info_t::purged_snaps as array, not string ([issue#18584](#), [pr#14217](#), liuchang0812)
- osd: process deletes during recovery instead of peering ([issue#19971](#), [pr#15952](#), Josh Durgin)
- osd: put osdmap in mempool ([pr#14780](#), Sage Weil)
- osd: reduce buffer pinning from EC entries ([pr#15120](#), Sage Weil)
- osd: reduce map cache size ([pr#15292](#), Sage Weil)
- osd: reduce rados_max_object_size from 100 GB -> 128 MB ([pr#15520](#), Sage Weil)
- osd: remove copy-get-classic ([pr#13547](#), Sage Weil)
- osd: remove sortbitwise thrashing ([pr#13296](#), Sage Weil)
- osd: renamed the new vector name in OSDMap::build_simple_crush_map_from_conf ([pr#14583](#), Jos Collin)
- osd: rename osd -> osd_pglog; include pglog-related bufferlists ([pr#15531](#), Sage Weil)
- osd: rephrase "wrongly marked me down" clog message ([pr#16365](#), John Spray)
- osd: replace object_info_t::operator==() with decode() ([pr#13938](#), tang.jin)
- osd: ReplicatedBackend::prep_push() remove redundant variable assignments ([pr#14817](#), Jin Cai)
- osd: restart boot process if waiting for luminous mons ([issue#20631](#), [pr#16341](#), Sage Weil)
- osd: Return correct osd_objectstore in OSD metadata ([issue#18638](#), [pr#13072](#), Wido den Hollander)
- osd: Return early on shutdown ([issue#19900](#), [pr#15345](#), Brad Hubbard)
- osd: Reverse order of op_has_sufficient_caps and do_pg_op ([issue#19790](#), [pr#15354](#), Brad Hubbard)
- osd: sched_scrub() lock pg only if all scrubbing conditions are fulfilled ([pr#14968](#), Jin Cai)
- osd: scrub_to specifies clone ver, but transaction include head write... ([issue#20041](#), [pr#16404](#), David Zafman)
- osd: silence warning from -Wint-in-bool-context ([pr#16744](#), Jos Collin)
- osd: simplify past_intervals representation ([pr#14444](#), Samuel Just, Sage Weil)
- osd: small clear up and optimize on _recover_now and should_share_map function ([pr#13476](#), song baisen)
- osd: stop mgrc earlier in shutdown() ([issue#19638](#), [pr#14904](#), Kefu Chai)
- osd: stop MgrClient callbacks on shutdown ([issue#19638](#), [pr#14896](#), Sage Weil)
- osd: strip pglog op name ([pr#14764](#), liuchang0812)
- osd: support cmpext operation on EC-backed pools ([pr#15693](#), Zhengyong Wang, Jason Dillaman)
- osd: support dumping long ops ([pr#13019](#), Zhiqiang Wang)
- osd: switch filestore to default to rocksdb ([pr#14814](#), Neha Ojha)
- osd: tag fast dispatch messages with min_epoch ([pr#13681](#), Sage Weil)
- osd: take PGRef for recovery sleep wakeup event ([issue#20226](#), [pr#15582](#), Sage Weil)
- osd: the condition of last epoch <= superblock.newest_map epoch has been checked twice ([pr#15590](#), linbing)
- osd: the osd should not share map with others when it is in stopping state ([pr#13668](#), song baisen)
- osd: unlock sdata_op_ordering_lock with sdata_lock hold to avoid miss... ([pr#15891](#), Ming Lin)
- osd: use append(bufferlist &) to avoid unnecessary copy ([pr#12272](#), Yunchuan Wen)
- osd: use separate waitlist for scrub ([pr#13136](#), Sage Weil)
- osd: various changes for preventing internal ENOSPC condition ([issue#16878](#), [pr#13425](#), David Zafman)
- osd: we know the definite epoch of marking down ([pr#13121](#), Mingxin Liu)

- osd: when osd in not in failure_pending, we don't need to get osd inst from osdmap ([pr#15558](#), linbing)
- osd: When scrub finds an attr error mark shard inconsistent ([issue#20089](#), [pr#15368](#), David Zafman)
- osd: zipkin tracing ([pr#14305](#), Sage Weil, Marios-Evaggelos Kogias, Victor Araujo, Casey Bodley, Andrew Shewmaker, Chendi.Xue)
- performance: buffer, osd: add missing crc cache miss perf counter ([pr#14957](#), Piotr Dałek)
- performance: common/config_opts.h: Lower HDD throttle cost ([pr#15485](#), Mark Nelson)
- performance: crc32c: optimize aarch64 crc32c implementation ([pr#12977](#), wei xiao)
- performance: denc: add need_contiguous to denc_traits ([pr#15224](#), Kefu Chai)
- performance: osd, messenger, librados: lttng oid tracing ([pr#12492](#), Anjaneya Chagam)
- performance: osd/PG.cc: loop invariant code motion ([pr#12720](#), Li Wang)
- performance: osd/ReplicatedBackend: do not set omap header if it is empty ([pr#12612](#), fang yuxiang)
- performance,rgw: rgw_file: permit dirent offset computation ([pr#16275](#), Matt Benjamin)
- pybind: better error msg ([pr#14497](#), Kefu Chai)
- pybind: cephfs should be built without librados / python-rados ([pr#13431](#), Kefu Chai)
- pybind: ceph.in: Check return value when connecting ([pr#16130](#), Douglas Fuller)
- pybind: ceph-rest-api: Various REST API fixes ([pr#15910](#), Wido den Hollander)
- pybind: conditional compile the linux specific constant ([pr#12198](#), Kefu Chai)
- pybind: fix docstring for librbd Python binding ([pr#13977](#), runsisi)
- pybind: fix open flags calculation ([issue#19890](#), [pr#15018](#), "Yan, Zheng")
- pybind: pybind/ceph_argparse: fix empty string check ([issue#20135](#), [pr#15500](#), Sage Weil)
- pybind: pybind/ceph_daemon.py: fix Termsize.update ([pr#15253](#), Kefu Chai)
- pybind: pybind/ceph_daemon: use small chunk for recv ([pr#13804](#), Xiaoxi Chen)
- pybind: pybind/mgr/dashboard: fix get kernel_version error ([pr#16094](#), Peng Zhang)
- pybind: pybind/mgr/restful: fix typo ([pr#16560](#), Nick Erdmann)
- pybind: pybind/mgr/restful: use list to pass hooks to create a *Pecan* instance ([issue#20258](#), [pr#15646](#), Kefu Chai)
- pybind: pybind/rados: avoid call free() on invalid pointer ([pr#15159](#), Mingxin Liu)
- pybind: pybind/rados: use new APIs instead of deprecated ones ([pr#16684](#), Kefu Chai)
- pybind,rbd: pybind/rbd: OSError should be pickleable ([issue#20223](#), [pr#15574](#), Jason Dillaman)
- pybind: restore original API for backwards compatibility ([issue#20421](#), [pr#15932](#), Jason Dillaman)
- pybind: support mon target in pybind ([pr#15409](#), liuchang0812)
- qa: fix POOL_APP_NOT_ENABLED warning in krbd:unmap suite ([pr#17192](#), Ilya Dryomov)
- rbd: add default note info to size (create and resize) ([pr#15561](#), Zheng Yin)
- rbd: add error prompt when input command 'snap set limit' is incomplete ([pr#12945](#), Tang Jin)
- rbd: additional validation for 'bench' optional parameters ([pr#12697](#), Yunchuan Wen)
- rbd: bench-write should return error if io-size >= 4G ([issue#18422](#), [pr#12864](#), Gaurav Kumar Garg)
- rbd: cleanup: fix the typo in namespace comment ([pr#12858](#), Dongsheng Yang)
- rbd: cleanup: rbd: fix a typo in comment ([pr#14049](#), Dongsheng Yang)
- rbd: cls_rbd: default initialize snapshot namespace for legacy clients ([issue#19413](#), [pr#14903](#), Jason Dillaman)
- rbd: cls/rbd: silence warning from -Wunused-variable ([pr#16670](#), Yan Jun)
- rbd: cls/rbd: trash_list should be iterable ([issue#20643](#), [pr#16372](#), Jason Dillaman)
- rbd: common/bit_vector: utilize deep-copy during data decode ([issue#19863](#), [pr#15017](#), Jason Dillaman)
- rbd: correct coverity warnings ([pr#12954](#), Jason Dillaman)
- rbd: correct issues with image importing ([pr#14401](#), Jason Dillaman)
- rbd: demote/promote all mirrored images in a pool ([issue#18748](#), [pr#13758](#), Jason Dillaman)
- rbd: destination pool should be source pool if it is not specified ([issue#18326](#), [pr#13189](#), Gaurav Kumar Garg)
- rbd: do not attempt to load key if auth is disabled ([issue#19035](#), [pr#16024](#), Jason Dillaman)
- rbd: Drop unused member variable reopen in C_OpenComplete ([pr#16729](#), amitkuma)
- rbd: enable rbd on FreeBSD (without KRBD) ([pr#12798](#), Willem Jan Withagen)
- rbd: error out if import image format failed ([pr#13957](#), wangzhengyong)
- rbd: fixed coverity 'Argument cannot be negative' warning ([pr#16686](#), amitkuma)
- rbd: fix typo in Kernel.cc ([issue#19273](#), [pr#13983](#), Gaurav Kumar Garg)
- rbd: 'image-meta remove' for missing key does not return error ([issue#16990](#), [pr#16393](#), PCzhangPC)
- rbd: import-diff should discard any zeroed extents ([pr#14445](#), Jason Dillaman)
- rbd: import needs to sanity check auto-generated image name ([issue#19128](#), [pr#14754](#), Mykola Golub)
- rbd: import real thin-provision image ([issue#15648](#), [pr#12883](#), yaoning, Ning Yao)
- rbd: info command should indicate if parent is in trash ([pr#14875](#), Jason Dillaman)
- rbd: introduce v2 format for rbd export/import ([issue#13186](#), [pr#10487](#), Dongsheng Yang)
- rbd: journal: don't hold future lock during assignment ([issue#18618](#), [pr#13033](#), Jason Dillaman)
- rbd: journal: stop processing removal after error ([issue#18738](#), [pr#13193](#), Jason Dillaman)
- rbd: luminous: librbd: default localize parent reads to false ([issue#20941](#), [pr#16899](#), Jason Dillaman)
- rbd: luminous: librbd: remove consistency group rbd cli and API support ([pr#16875](#), Jason Dillaman)
- rbd: luminous: rbd-ggate: tool to map images on FreeBSD via GEOM Gate ([pr#16895](#), Mykola Golub)
- rbd: luminous: rbd-mirror: align use of uint64_t in service_daemon::AttributeType ([pr#16948](#), James Page)
- rbd: luminous: rbd-mirror: simplify notifications for image assignment ([issue#15764](#), [pr#16878](#), Jason Dillaman)
- rbd: luminous: rbd: parallelize rbd ls -l ([pr#16921](#), Piotr Dałek)

- rbd: make it more understandable when adding peer returns error ([pr#16313](#), songweibin)
- rbd-mirror: add support for active/pассиве daemon instances ([issue#17018](#), [issue#17019](#), [issue#17020](#), [pr#12948](#), Mykola Golub)
- rbd-mirror: assertion failure during pool replayer shut down ([issue#20644](#), [pr#16704](#), Jason Dillaman)
- rbd-mirror: avoid processing new events after stop requested ([issue#18441](#), [pr#12837](#), Jason Dillaman)
- rbd-mirror: check remote image mirroring state when bootstrapping ([issue#18447](#), [pr#12820](#), Mykola Golub)
- rbd-mirror: coordinate image syncs with leader ([issue#18789](#), [pr#14745](#), Mykola Golub)
- rbd-mirror: delayed replication support ([issue#15371](#), [pr#11879](#), Mykola Golub)
- rbd-mirror: deleting a snapshot during sync can result in read errors ([issue#18990](#), [pr#13568](#), Jason Dillaman)
- rbd-mirror: ensure missing images are re-synced when detected ([issue#19811](#), [pr#14945](#), Jason Dillaman)
- rbd-mirror: failover and fallback of unmodified image results in split-brain ([issue#19858](#), [pr#14963](#), Jason Dillaman)
- rbd-mirror: guard the deletion of non-primary images ([pr#16398](#), Jason Dillaman)
- rbd-mirror: ignore permission errors on rbd_mirroring object ([issue#20571](#), [pr#16264](#), Jason Dillaman)
- rbd-mirror: image deletions should be handled by assigned instance ([pr#14832](#), Jason Dillaman)
- rbd-mirror: initialize timer context pointer to null ([pr#16603](#), Jason Dillaman)
- rbd-mirror: InstanceWatcher watch/notify stub for leader/follower RPC ([issue#18783](#), [pr#13312](#), Mykola Golub)
- rbd-mirror: lock loss during sync should wait for in-flight copies ([pr#15532](#), Jason Dillaman)
- rbd-mirror: permit release of local image exclusive lock after force promotion ([issue#18963](#), [pr#15140](#), Jason Dillaman)
- rbd-mirror: pool watcher should track mirror uuid ([pr#14240](#), Jason Dillaman)
- rbd-mirror: remove tracking of image names from pool watcher ([pr#14712](#), Jason Dillaman)
- rbd-mirror: replace remote pool polling with add/remove notifications ([issue#15029](#), [pr#12364](#), Jason Dillaman)
- rbd-mirror: resolve admin socket path names collision ([issue#19907](#), [pr#15048](#), Mykola Golub)
- rbd-mirror: separate ImageReplayer handling from Replayer ([issue#18785](#), [pr#13803](#), Mykola Golub)
- rbd-mirror: Set the data pool correctly when creating images ([issue#20567](#), [pr#17023](#), Adam Wolfe Gordon)
- rbd-mirror: track images via global image id ([pr#13416](#), Jason Dillaman)
- rbd: modified some commands' description into imperative sentence ([pr#16694](#), songweibin)
- rbd-nbd: check /sys/block/nbdX/size to ensure kernel mapped correctly ([issue#18335](#), [pr#13229](#), Mykola Golub)
- rbd-nbd: clean up the doc and help information ([pr#14146](#), Pan Liu)
- rbd-nbd: create admin socket only for map command ([issue#17951](#), [pr#12433](#), Pan Liu)
- rbd-nbd: display pool/image/snap information in list output ([pr#15317](#), Pan Liu)
- rbd-nbd: don't ignore -read-only option in BLKROSET ioctl ([pr#13944](#), Pan Liu)
- rbd-nbd: ensure unmap returns error code ([pr#15593](#), guojiannan, chenfangxian)
- rbd-nbd: fix a typo "moudle" ([pr#13652](#), Pan Liu)
- rbd-nbd: fix typo in comment ([pr#14034](#), Pan Liu)
- rbd-nbd: no need to check image format any more ([pr#13389](#), Mykola Golub)
- rbd-nbd: relax size check for newer kernel versions ([issue#19871](#), [pr#14976](#), Mykola Golub)
- rbd-nbd: remove debug messages from do_unmap ([pr#14253](#), Pan Liu)
- rbd-nbd: s/cpp_error/cpp_strerror/ to fix FTBFS ([pr#14223](#), Kefu Chai)
- rbd-nbd: support signal handle for SIGHUP, SIGINT and SIGTERM ([issue#19349](#), [pr#14079](#), Pan Liu)
- rbd-nbd: update size only when NBD_SET_SIZE successful ([pr#14005](#), Pan Liu)
- rbd-nbd: warn when kernel parameters are ignored ([issue#19108](#), [pr#13694](#), Pan Liu)
- rbd: prevent adding multiple mirror peers to a single pool ([issue#19256](#), [pr#13919](#), Jason Dillaman)
- rbd: properly decode features when using image name optional ([issue#20185](#), [pr#15492](#), Jason Dillaman)
- rbd: pybind/rbd: add image metadata methods ([issue#19451](#), [pr#14463](#), Mykola Golub)
- rbd: pybind/rbd: fix crash if more than 1024 images in trash bin ([pr#15134](#), runsisi)
- rbd: rbd/bench: add notes of default values, it's easy to use ([pr#14762](#), Zheng Yin)
- rbd: rbd/bench: fix write gaps when doing sequential writes with io-threads > 1 ([pr#15206](#), Igor Fedotov)
- rbd: rbd, librbd: migrate atomic_t to std::atomic ([pr#14656](#), Jesse Williamson)
- rbd: rbd-mirror A/A: leader should track up/down rbd-mirror instances ([issue#18784](#), [pr#13571](#), Mykola Golub)
- rbd: rbd-mirror A/A: proxy InstanceReplayer APIs via InstanceWatcher RPC ([issue#18787](#), [pr#13978](#), Mykola Golub)
- rbd: recognize exclusive option ([pr#14785](#), Ilya Dryomov)
- rbd: removed hardcoded default pool ([pr#15518](#), Jason Dillaman)
- rbd: remove direct linking to static boost libraries ([pr#12962](#), Jason Dillaman)
- rbd: removed spurious error message from mirror pool commands ([pr#14935](#), Jason Dillaman)
- rbd: remove unused condition within group action handler ([pr#12723](#), Gaurav Kumar Garg)
- rbd,rgw,tools: tools/rbd, rgw: Removed unreachable returns ([pr#16308](#), Jos Collin)
- rbd: spell out image features unsupported by the kernel ([issue#19095](#), [pr#13812](#), Ilya Dryomov)
- rbd: stop indefinite thread waiting in krbd udev handling ([issue#17195](#), [pr#14051](#), Spandan Kumar Sahu)
- rbd: test: fix rbd unit test cases w/ striping feature ([issue#18888](#), [pr#13196](#), Venky Shankar)
- rbd,tests: luminous: qa/workunits/rbd: use command line option to specify watcher asok ([issue#20954](#), [pr#16946](#), Mykola Golub)
- rbd,tests: luminous: test/librbd: fix race condition with OSD map refresh ([issue#20918](#), [pr#16903](#), Jason Dillaman)
- rbd,tests: qa: add workunit to test krbd data-pool support ([pr#13482](#), Ilya Dryomov)
- rbd,tests: qa: integrate OpenStack 'gate-tempest-dsvm-full-devstack-plugin-ceph' ([issue#18594](#), [pr#13158](#), Jason Dillaman)

- rbd,tests: qa: krbd_data_pool.sh: account for rbd_info metadata object ([pr#14631](#), Ilya Dryomov)
- rbd,tests: qa: krbd discard/zeroout tests ([pr#15388](#), Ilya Dryomov)
- rbd,tests: qa: krbd write-after-checksum tests ([pr#14836](#), Ilya Dryomov)
- rbd,tests: qa/suites/krbd: unmap subsuite needs straw buckets ([pr#15290](#), Ilya Dryomov)
- rbd,tests: qa/suites/rbd: restrict python memcheck validation to CentOS ([pr#15923](#), Jason Dillaman)
- rbd,tests: qa/tasks/qemu: update default image url after ceph.com redesign ([issue#18542](#), [pr#12953](#), Jason Dillaman)
- rbd,tests: qa/tasks/rbd_fio: bump default fio version to 2.21 ([pr#16656](#), Ilya Dryomov)
- rbd,tests: qa/tasks: rbd-mirror daemon not properly run in foreground mode ([issue#20630](#), [pr#16340](#), Jason Dillaman)
- rbd,tests: qa: thrash tests for backoff and upmap ([pr#16428](#), Ilya Dryomov)
- rbd,tests: qa: update krbd_data_pool.sh to match the new rados ls behavior ([pr#15594](#), Ilya Dryomov)
- rbd,tests: qa/workunits: adjust path to ceph-helpers.sh ([pr#16599](#), Sage Weil)
- rbd,tests: qa/workunits: corrected issues with RBD cli test ([pr#14460](#), Jason Dillaman)
- rbd,tests: qa/workunits/rbd: diff.sh failed removing nonexistent file ([pr#14482](#), Mykola Golub)
- rbd,tests: qa/workunits/rbd: increased trash deferment period ([pr#14846](#), Jason Dillaman)
- rbd,tests: qa/workunits/rbd: resolve potential rbd-mirror race conditions ([issue#18935](#), [pr#13421](#), Jason Dillaman)
- rbd,tests: qa/workunits/rbd: test data pool is mirrored correctly ([pr#17077](#), Mykola Golub)
- rbd,tests: qa/workunits/rbd: tweak rbd-mirror config to speed up testing ([pr#13228](#), Mykola Golub)
- rbd,tests: qa/workunits: switch to OpenStack Ocata release for RBD testing ([pr#14465](#), Jason Dillaman)
- rbd,tests: test: correct language mode in file headers ([pr#12924](#), Jason Dillaman)
- rbd,tests: test: fix compile warning in ceph_test_cls_rbd ([pr#15919](#), Jason Dillaman)
- rbd,tests: test: fix failing rbd devstack teuthology test ([pr#15956](#), Jason Dillaman)
- rbd,tests: test/librados_test_stub: fixed cls_cxx_map_get_keys/vals return value ([issue#19597](#), [pr#14484](#), Jason Dillaman)
- rbd,tests: test/librbd: add break_lock test ([pr#12842](#), Mykola Golub)
- rbd,tests: test/librbd/CMakeLists.txt: ceph_test_librbd_fsx requires linux includes/libs ([pr#13630](#), Willem Jan Withagen)
- rbd,tests: test/librbd/fsx: Add break in case OP_WRITESAME and OP_COMPARE_AND_WRITE ([pr#16742](#), Luo Kexue)
- rbd,tests: test/librbd: move tests using non-public api to internal ([pr#13806](#), Venky Shankar)
- rbd,tests: test/librbd/test_librbd.cc: set *features even if RBD_FEATURES is unset ([issue#19865](#), [pr#14965](#), Dan Mick)
- rbd,tests: test/librbd/test_notify.py: don't disable feature in slave ([issue#19716](#), [pr#14751](#), Mykola Golub)
- rbd,tests: test/librbd: unit tests cleanup ([pr#15113](#), Mykola Golub)
- rbd,tests: test: rbd master/slave notify test should test active features ([issue#19692](#), [pr#14638](#), Jason Dillaman)
- rbd,tests: test/rbd_mirror: race in TestMockInstanceWatcher on destroy ([pr#14453](#), Mykola Golub)
- rbd,tests: test/rbd_mirror: race in TestMockLeaderWatcher.AcquireError ([issue#19405](#), [pr#14741](#), Mykola Golub)
- rbd,tests: test: remove hard-coded image name from RBD metadata test ([issue#19798](#), [pr#14848](#), Jason Dillaman)
- rbd,tests: test: support blacklisting within librados_test_stub ([pr#13737](#), Jason Dillaman)
- rbd,tests: test/unittest_librbd: remove unused variables ([pr#15720](#), shiqi)
- rbd,tests: test: use librados API to retrieve config params ([issue#18617](#), [pr#13076](#), Jason Dillaman)
- rbd,tools: rbdmap: consider /etc/ceph/rbdmap when unmapping images ([issue#18884](#), [pr#13361](#), David Disseldorp)
- rbd,tools: tools/rbd_mirror: initialize non-static class member m_do_resync in ImageReplayer ([pr#15889](#), Jos Collin)
- rbd,tools: tools/rbd_nbd: add -version show support ([pr#16254](#), Jin Cai)
- rbd: use concurrent writes for imports ([issue#19034](#), [pr#13782](#), Venky Shankar)
- rbd: use min<uint64_t>() explicitly ([issue#18938](#), [pr#14202](#), Kefu Chai)
- rbd: validate pool and snap name optionals ([issue#14535](#), [pr#13836](#), Gaurav Kumar Garg)
- rbd: warning, 'devno' may be used uninitialized in this function ([pr#14271](#), Jos Collin)
- rbd: When Ceph cluster becomes full, should allow user to remove rbd ... ([pr#12627](#), Pan Liu)
- rdma: msg/async: Postpone bind if network stack is not ready ([pr#14414](#), Amir Vadai, Haomai Wang)
- rdma: msg/async/rdma: Add DSCP support ([pr#15484](#), Sarit Zubakov)
- rdma: msg/async/rdma: add inqueue rx chunks perf counter ([pr#14782](#), Haomai Wang)
- rdma: msg/async/rdma: fix log line spacing ([pr#13263](#), Adir Lev)
- rdma: msg/async/rdma: Make poll_blocking() poll for async events in addition... ([pr#14320](#), Amir Vadai)
- rdma: msg/async/rdma: Make port number an attribute of the Connection not o... ([pr#14297](#), Amir Vadai)
- rdma: msg/async/rdma: RDMA-CM, get_device() by ibv_context ([pr#14410](#), Amir Vadai)
- rdma: msg/async: Revert RDMA-CM ([pr#15262](#), Amir Vadai)

Replace using sleep with new wait_for_health() bash function

- rgw: abort early when s->length empty during putobj ([pr#15682](#), Jiaying Ren)
- rgw: AbortMultipart request returns NoSuchUpload error if the meta obj doesn't exist ([pr#12793](#), Zhang Shaowen)
- rgw: acl grants num limit ([pr#16291](#), Enming Zhang)
- rgw: add a new error code for non-existed subuser ([pr#16095](#), Zhao Chao)
- rgw: add a new error code for non-existed user ([issue#20468](#), [pr#16033](#), Zhao Chao)
- rgw: add apis to support ragweed ([pr#13645](#), Yehuda Sadeh)
- rgw: add a separate configuration for data notify interval ([pr#16551](#), fang yuxiang)
- rgw: add bucket size limit check to radosgw-admin ([issue#17925](#), [pr#11796](#), Matt Benjamin)
- rgw: Added a globbing method for AWS Policies ([pr#12445](#), Pritha Srivastava)
- rgw: Added code for REST APIs for AWS Roles ([pr#12104](#), Pritha Srivastava)

- rgw: Added code to correctly account for bytes sent/ received during a ‘PUT’ operation ([pr#14042](#), Pritha Srivastava)
- rgw: Adding code to create tenanted user for s3 bucket policy tests ([pr#15028](#), Pritha Srivastava)
- rgw: add lifecycle validation according to S3 ([issue#18394](#), [pr#12750](#), Zhang Shaowen)
- rgw: add missing RGWPeriod::reflect() based on new atomic update_latest_epoch() ([issue#19816](#), [issue#19817](#), [pr#14915](#), Casey Bodley)
- rgw: add -num-zonegroups option for multi test ([pr#14216](#), lvshuhua)
- rgw: add override in header files mostly ([pr#13586](#), liuchang0812)
- rgw: add override in rgw subsystem ([issue#18922](#), [pr#13441](#), liuchang0812)
- rgw: add pool namespace to cache’s key so that system obj can have unique key ([issue#19372](#), [pr#14125](#), Zhang Shaowen)
- rgw: add radosclient finisher to perf counter ([issue#19011](#), [pr#13535](#), lvshuhua)
- rgw: add “rgw_verify_ssl” config ([pr#15301](#), Shasha Lu)
- rgw: add ‘state==SyncState::IncrementalSync’ condition when add item ... ([pr#14552](#), Shasha Lu)
- rgw: add support container and object levels of swift bulkupload ([pr#14775](#), Jing Wenjun)
- rgw: add support for delete marker expiration in s3 lifecycle ([issue#19730](#), [pr#14703](#), Zhang Shaowen)
- rgw: add support for FormPost of Swift API ([issue#17273](#), [pr#11179](#), Radoslaw Zarzynski, Orit Wasserman)
- rgw: add support for multipart upload expiration ([issue#19088](#), [pr#13622](#), Zhang Shaowen)
- rgw: add support for noncurrentversion expiration in s3 lifecycle ([issue#18916](#), [pr#13385](#), Zhang Shaowen)
- rgw: add support for Swift’s TempURLs with prefix-based scope ([issue#20398](#), [pr#16370](#), Radoslaw Zarzynski)
- rgw: add support for the BulkUpload of Swift API ([pr#12243](#), Radoslaw Zarzynski)
- rgw: add the remove-x-delete feature to cancel swift object expiration ([issue#19074](#), [pr#13621](#), Jing Wenjun)
- rgw: add the Vim’s modeline into rgw_orphan.cc ([pr#15431](#), Radoslaw Zarzynski)
- rgw: add variadic string join for s3 signature generation ([pr#15678](#), Casey Bodley)
- rgw: Add -zonegroup-new-name in usage ([pr#12084](#), Hans van den Bogert)
- rgw: allow larger payload for period commit ([issue#19505](#), [pr#14355](#), Casey Bodley)
- rgw: allow system users to read SLO parts ([issue#19027](#), [pr#13561](#), Casey Bodley)
- rgw: auto reshards old buckets ([pr#15665](#), Orit Wasserman)
- rgw: avoid listing user buckets for rgw_delete_user ([pr#13991](#), liuchang0812)
- rgw: avoid using null pointer in rgw_file.cc ([pr#14474](#), lihongjie)
- rgw: be aware about tenants on cls_user_bucket -> rgw_bucket conversion ([issue#18364](#), [issue#16355](#), [pr#13220](#), Radoslaw Zarzynski)
- rgw: bucket index check in radosgw-admin removes valid index ([issue#18470](#), [pr#12851](#), Zhang Shaowen)
- rgw: bucket stats display bucket index type ([pr#14466](#), fang yuxiang)
- rgw: change default chunk size to 4MB ([issue#18621](#), [issue#18622](#), [issue#18623](#), [pr#13035](#), Yehuda Sadeh)
- rgw: change loglevel to 20 for ‘System already converted’ message ([issue#18919](#), [pr#13399](#), Vikhyat Umrao)
- rgw: change loglevel to 5 in user’s quota sync ([issue#18921](#), [pr#13408](#), Zhang Shaowen)
- rgw: Changes for s3test config file, to add user under a tenant ([pr#15753](#), Pritha Srivastava)
- rgw: check placement existence when create bucket ([pr#16385](#), Jiaying Ren)
- rgw: check placement target existence during bucket creation ([pr#16384](#), Jiaying Ren)
- rgw: civetweb don’t go past the array index while calling mg_start ([issue#19749](#), [pr#14750](#), Abhishek Lekshmanan, Jesse Williamson)
- rgw: clean redundant code ([pr#13302](#), Yankun Li)
- rgw: clean unuse code in cls_statelog_check_state ([pr#10260](#), weiqiaomiao)
- rgw: clean-up error mapping in Swift’s authentication strategy ([pr#15756](#), Radoslaw Zarzynski)
- rgw: cleanup: fix variable name in RGWRados::create_pool() declaration ([pr#14547](#), Nathan Cutler)
- rgw: cleanup lc continuation ([pr#14906](#), Jiaying Ren)
- rgw: cleanup lifecycle managament ([pr#13820](#), Jiaying Ren)
- rgw: cleanup rgw-admin duplicated judge during OLH GET/READLOG ([pr#15700](#), Jiaying Ren)
- rgw: clean up the redundant assignment in last_entry_in_listing ([pr#13387](#), Jing Wenjun)
- rgw: clean up the unneeded rgw::io::ChunkingFilter::has_content_length ([pr#13504](#), Radoslaw Zarzynski)
- rgw: cleanup unused codes in rgw_admin.cc ([pr#15771](#), fang yuxiang)
- rgw: cleanup unused var in rgw/rgw_rest_s3.cc ([pr#13434](#), Jiaying Ren)
- rgw: clear master_zonegroup when resetting RGWPeriodMap ([issue#17239](#), [pr#12658](#), Orit Wasserman)
- rgw: clear old zone short ids on period update ([issue#15618](#), [pr#13949](#), Casey Bodley)
- rgw: cls: ceph::timespan tag_timeout wrong units ([issue#20380](#), [pr#16026](#), Matt Benjamin)
- rgw: cls/rgw: Clean up the “magic string” usage in the cls layer for RGW ([pr#12536](#), Ira Cooper)
- rgw: cls/rgw: list_plain_entries() stops before bi_log entries ([issue#19876](#), [pr#14981](#), Casey Bodley)
- rgw: cls/user: cls_user_bucket backward compatibility ([issue#19367](#), [pr#14128](#), Yehuda Sadeh)
- rgw: cls_user don’t clobber existing bucket stats when creating bucket ([issue#16357](#), [pr#10121](#), Abhishek Lekshmanan)
- rgw: complete versioning enablement after sending it to meta master ([issue#18003](#), [pr#12444](#), Orit Wasserman)
- rgw: Compress crash bug refactor ([issue#20098](#), [pr#15569](#), Adam Kupczyk)
- rgw: continuation of the auth rework - AWSv4 ([issue#18800](#), [pr#14885](#), Radoslaw Zarzynski, Javier M. Mellid)
- rgw: continuation of the auth rework ([pr#12893](#), Radoslaw Zarzynski, Matt Benjamin)
- rgw: Correcting the condition in ceph_assert while parsing an AWS Principal ([pr#15997](#), Pritha Srivastava)
- rgw: correct the debug info when unlink instance failed ([pr#13761](#), Zhang Shaowen)

- rgw: Correct the return codes for the health check feature ([issue#19025](#), [pr#13557](#), Pavan Rallabhandi)
- rgw: custom user data header ([issue#19644](#), [pr#14592](#), Pavan Rallabhandi)
- rgw: datalog trim and mdlog trim handles the result returned by osd incorrectly ([issue#20190](#), [pr#15507](#), Zhang Shaowen)
- rgw: data sync includes instance in rgw_obj_index_key ([pr#13948](#), Casey Bodley)
- rgw: deduplicate variants of rgw_make_bucket_entry_name() ([pr#14299](#), Radoslaw Zarzynski)
- rgw: delete non-empty buckets in slave zonegroup works not well ([issue#19313](#), [pr#14043](#), Zhang Shaowen)
- rgw: delete object in error path ([issue#20620](#), [pr#16324](#), Yehuda Sadeh)
- rgw: disable dynamic reshading for 1st L point release ([pr#16969](#), Matt Benjamin)
- rgw: display more info when using radosgw-admin bucket stats ([pr#15256](#), fang.yuxiang)
- rgw: Do not decrement stats cache when the cache values are zero ([issue#20661](#), [pr#16389](#), Pavan Rallabhandi)
- rgw: Do not fetch bucket stats by default upon bucket listing ([issue#20377](#), [pr#15834](#), Pavan Rallabhandi)
- rgw: do not log debug output at level 0 ([pr#15633](#), Jens Rosenboom)
- rgw: don't do unneccesary write if buffer with zero length ([pr#14925](#), fang yuxiang)
- rgw: don't init rgw_obj from rgw_obj_key when it's incorrect to do so ([issue#19096](#), [pr#13676](#), Yehuda Sadeh)
- rgw: don't log the env_map twice ([pr#13481](#), Abhishek Lekshmanan)
- rgw: don't read all user input for a few param requests ([pr#13815](#), Abhishek Lekshmanan)
- rgw: don't return skew time error in pre-signed url ([issue#18828](#), [pr#13354](#), liuchang0812)
- rgw: dont spawn error_repo until lease is acquired ([issue#19446](#), [pr#14714](#), Casey Bodley)
- rgw: don't specify a length when converting bl -> string ([issue#20037](#), [pr#15599](#), Abhishek Lekshmanan)
- rgw: don't use strlen in constexprs to not brake Clang builds ([pr#15688](#), Radoslaw Zarzynski)
- rgw: drop asio/{yield,coroutine}.hpp replacements ([pr#15413](#), Kefu Chai)
- rgw: Drop dump_usage_bucket_info() to silence warning from -Wunused-function ([pr#16497](#), Wei Qiaomiao)
- rgw: drop unused find_replacement() and some function docs ([pr#16386](#), Jiaying Ren)
- rgw: drop unused function RGWRemoteDataLog::get_shard_info() ([pr#16236](#), Shasha Lu)
- rgw: drop unused param "bucket" from select_bucket_placement ([pr#14390](#), Shasha Lu)
- rgw: drop unused port var ([pr#14412](#), Jiaying Ren)
- rgw: drop unused rgw_pool parameter, local variables and member variable ([pr#16154](#), Jiaying Ren)
- rgw: drop unused var header_ended ([pr#15686](#), Jiaying Ren)
- rgw: drop using std ns in header files and other cleanups ([pr#15137](#), Abhishek Lekshmanan)
- rgw: dynamic resharding ([pr#15493](#), Yehuda Sadeh, Orit Wasserman)
- rgw: enable to update acl of bucket created in slave zonegroup ([issue#16888](#), [pr#14082](#), Guo Zhandong)
- rgw: error_code in error log is not right when data sync fails ([issue#18437](#), [pr#12810](#), Zhang Shaowen)
- rgw: error more verbosely in RGWRados::create_pool ([pr#14642](#), Matt Benjamin)
- rgw: external auth engines of S3 honor rgw_keystone_implicit_tenants ([issue#17779](#), [pr#15572](#), Radoslaw Zarzynski)
- rgw: Fix a bug that multipart upload may exceed the quota ([issue#19602](#), [pr#12010](#), Zhang Shaowen)
- rgw: fix asctime when logging in rgw_lc ([pr#16422](#), Abhishek Lekshmanan)
- rgw: fix break inside of yield in RGWFetchAllMetaCR ([issue#17655](#), [pr#11586](#), Casey Bodley)
- rgw: fix broken /crossdomain.xml, /info and /healthcheck of Swift API ([issue#19520](#), [pr#14373](#), Radoslaw Zarzynski)
- rgw: fix build of conflict after auth rework ([pr#14203](#), Casey Bodley)
- rgw: fix configurable write obj window size ([pr#13934](#), hechuang)
- rgw: fix constexpr for string_size in clang ([pr#15738](#), Adam C. Emerson, Casey Bodley)
- rgw: fix disabling Swift's object versioning through empty X-Versions-Location ([issue#18852](#), [pr#13303](#), Jing Wenjun)
- rgw: Fix duplicate tag removal during GC ([issue#20107](#), [pr#15912](#), Jens Rosenboom)
- rgw: fix error code of inexistence of versions location in swift api ([issue#18880](#), [pr#13350](#), Jing Wenjun)
- rgw: fix error handling in get_params() of RGWPostObj_ObjStore_S3 ([pr#15670](#), Radoslaw Zarzynski)
- rgw: fix error handling in the link() method of RGWBucket ([issue#20279](#), [pr#15669](#), Radoslaw Zarzynski)
- rgw: fix error message in removing bucket with -bypass-gc flag ([issue#20688](#), [pr#16419](#), Abhishek Varshney)
- rgw: fix err when copy object in bucket with specified placement rule ([issue#20378](#), [pr#15837](#), fang yuxiang)
- rgw: fixes for AWSBrowserUploadAbstractor auth ([issue#20372](#), [pr#15882](#), Radoslaw Zarzynski, Casey Bodley)
- rgw: Fixes typo in rgw_admin.cc ([issue#19026](#), [pr#13576](#), Ronak Jain)
- rgw: fix for broken yields in RGWMetaSyncShardCR ([issue#18076](#), [pr#12223](#), Casey Bodley)
- rgw: fix for EINVAL errors on forwarded bucket put_acl requests ([pr#14376](#), Casey Bodley)
- rgw: fix for null version_id in fetch_remote_obj() ([pr#14375](#), Casey Bodley)
- rgw: Fix for Policy Parse exception in case of multiple statements ([pr#16689](#), Pritha Srivastava)
- rgw: fix forward request for bulkupload to be applied in multisite ([issue#19645](#), [pr#14601](#), Jing Wenjun)
- rgw: fix 'gc list -include-all' command infinite loop the first items ([issue#19978](#), [pr#12774](#), Shasha Lu, fang yuxiang)
- rgw: fix get bucket policy s3 compatible issue ([pr#15280](#), Enming Zhang)
- rgw: fix handling of -remote in radosgw-admin period commands ([issue#19554](#), [pr#14407](#), Casey Bodley)
- rgw: fix handling RGWUserInfo::system in RGWHandler_REST_SWIFT ([issue#18476](#), [pr#12865](#), Radoslaw Zarzynski)
- rgw: fix infinite loop in rest api for log list ([issue#20386](#), [pr#15983](#), xierui, Casey Bodley)
- rgw: fix init_bucket_for_sync retcode ([pr#13684](#), Shasha Lu)
- rgw: fix lc list failure when shards not be all created ([issue#19898](#), [pr#15025](#), Jiaying Ren)
- rgw: fix leaks with incomplete multipart ([issue#17164](#), [pr#15630](#), Abhishek Varshney)
- rgw: fix marker encoding problem ([issue#20463](#), [pr#15998](#), Marcus Watts)

- rgw: fix memory leak in copy_obj_to_remote_dest ([pr#9974](#), weiqiaomiao)
- rgw: fix memory leak in delete_obj_aio ([pr#13998](#), wangzhengyong)
- rgw: fix memory leak in RGWGetObjLayout ([pr#14014](#), liuchang0812)
- rgw: fix memory leaks during Swift Static Website's error handling ([issue#20757](#), [pr#16531](#), Radoslaw Zarzynski)
- rgw: fix not initialized vars which cause rgw crash with ec data pool ([issue#20542](#), [pr#16177](#), Aleksei Gutikov)
- rgw: fix off-by-one in RGWDataChangesLog::get_info ([issue#18488](#), [pr#12884](#), Casey Bodley)
- rgw: fix parse/eval of policy conditions with IfExists ([issue#20708](#), [pr#16463](#), Casey Bodley)
- rgw: fix period update crash ([issue#18631](#), [pr#13054](#), Orit Wasserman)
- rgw: fix potential null pointer dereference in rgw_admin ([pr#15667](#), Radoslaw Zarzynski)
- rgw: fix radosgw-admin data sync run crash ([issue#20423](#), [pr#15938](#), Shasha Lu)
- rgw: fix radosgw-admin retcode ([pr#15257](#), Shasha Lu)
- rgw: fix RadosGW hang during multi-chunk upload of AWSv4 ([issue#19754](#), [pr#14770](#), Radoslaw Zarzynski)
- rgw: fix radosgw will crash when service is restarted during lifecycl... ([issue#20756](#), [pr#16495](#), Wei Qiaomiao)
- rgw: fix response header of Swift API ([issue#19443](#), [pr#14280](#), tone-zhang)
- rgw: fix rest client's order of args in get_v2_signature ([pr#15731](#), Casey Bodley)
- rgw: fix rgw bucket policy IfExists position ([issue#20248](#), [pr#15607](#), yuliyang)
- rgw: fix rgw hang when do RGWRealmReloader::reload after go SIGHUP ([issue#20686](#), [pr#16417](#), fang.yuxiang)
- rgw: fix RGWPutBucketPolicy error when set BucketPolicy again without delete pre set Policy ([issue#20252](#), [pr#15617](#), yuliyang)
- rgw: fix s3 object uploads with chunked transfers and v4 signatures ([issue#20447](#), [pr#15965](#), Marcus Watts)
- rgw: fix segfault in RevokeThread during its shutdown procedure ([issue#19831](#), [pr#15033](#), Radoslaw Zarzynski)
- rgw: fix slave zonegroup cannot enable the bucket versioning ([issue#18003](#), [pr#12175](#), lvshuhua)
- rgw: fix SLO/DLO range requests ([pr#15060](#), Shasha Lu)
 - rgw: fix swift cannot disable object versioning
- rgw: fix swift default auth error after auth strategy refactoring ([pr#15711](#), Casey Bodley)
- rgw: fix test_multi.py default config file path ([pr#15306](#), Jiaying Ren)
- rgw: fix the bug that part's index can't be removed after completing multipart upload ([issue#19604](#), [pr#14500](#), Zhang Shaowen)
- rgw: fix the signature mismatch of FormPost in swift API ([issue#20220](#), [pr#15564](#), Jing Wenjun)
- rgw: fix the UTF8 check on bucket entry name in rgw_log_op() ([issue#20779](#), [pr#16604](#), Radoslaw Zarzynski)
- rgw: fix transition from full to incremental meta sync ([pr#13920](#), Casey Bodley)
- rgw: fix typo in comment ([pr#13578](#), liuchang0812)
- rgw: fix uninitialized fields ([pr#14120](#), wangzhengyong)
- rgw: fix upgrade from hammer when zone doesn't have zoneparams ([issue#19231](#), [pr#13900](#), Orit Wasserman)
- rgw: Fix up to 1000 entries at a time in check_bad_index_multipart ([issue#20772](#), [pr#16692](#), Orit Wasserman)
- rgw: fix use of marker in List::list_objects() ([issue#18331](#), [pr#13147](#), Yehuda Sadeh)
- rgw: fix versioned bucket data sync fail when upload is busy ([issue#18208](#), [pr#12357](#), lvshuhua)
- rgw: fix wrong error code for expired Swift TempURL's links ([issue#20384](#), [pr#15850](#), Radoslaw Zarzynski)
- rgw: fix X-Object-Meta-Static-Large-Object in SLO download ([issue#19951](#), [pr#15045](#), Shasha Lu)
- rgw: fix zone didn't update realm_id when added to zonegroup ([issue#17995](#), [pr#12139](#), Tianshan Qu)
- rgw: forward RGWPutBucketPolicy to meta master ([issue#20297](#), [pr#15736](#), Casey Bodley)
- rgw: get torrent request's parameter is not the same as amazon s3 ([issue#19136](#), [pr#13760](#), Zhang Shaowen)
- rgw: get wrong content when download object with specific range with compression ([issue#20100](#), [pr#15323](#), fang yuxiang)
- rgw: handle error return value in build_linked_oids_index ([pr#13955](#), wangzhengyong)
- rgw: http_client clarify the debug msg function call ([pr#13688](#), Abhishek Lekshmanan)
- rgw: if user.email is empty, dont try to delete ([issue#18980](#), [pr#13783](#), Casey Bodley)
- rgw: implement get/put object tags for S3 ([pr#13753](#), Abhishek Lekshmanan)
- rgw: improve handling of illformed Swift's container ACLs ([issue#18796](#), [pr#13248](#), Radoslaw Zarzynski)
- rgw: /info claims we do support Swift's accounts ACLs ([issue#20394](#), [pr#15887](#), Radoslaw Zarzynski)
- rgw: initialize non-static class members in ESQueryCompiler ([pr#15884](#), Jos Collin)
- rgw: initialize Non-static class member val in ESQueryNodeLeafVal_Int ([pr#15888](#), Jos Collin)
- rgw: initialize Non-static class member worker in RGWReshard ([pr#15886](#), Jos Collin)
- rgw: Initialize of member variable admin_specified in RGWUserAdminOpState ([pr#16847](#), amitkuma)
- rgw: Initialize pointer fields ([pr#16021](#), Jos Collin)
- rgw: LCWorker's worktime is not the same as config rgw_lifecycle_work_time ([issue#18087](#), [pr#11963](#), Zhang Shaowen)
- rgw: ldap: simple_bind() should set ldap version option on ldap ([pr#12616](#), Weibing Zhang)
- rgw: lease_stack: use reset method instead of assignment ([pr#16185](#), Nathan Cutler)
- rgw: Let the object stat command be shown in the usage ([issue#19013](#), [pr#13291](#), Pavan Rallabhandi)
- rgw: librgw shut ([issue#18585](#), [pr#12972](#), Matt Benjamin)
- rgw: lifecycle thread shouldn't process the bucket which has been deleted ([issue#20285](#), [pr#15677](#), Zhang Shaowen)
- rgw: lock is not released when set sync marker is failed ([issue#18077](#), [pr#12197](#), Zhang Shaowen)
- rgw: log_meta only for more than one zone ([issue#20357](#), [pr#15777](#), Orit Wasserman, Leo Zhang)
- rgw: lower some log's level in gc process ([pr#15426](#), Zhang Shaowen)

- rgw: luminous: rgw: Fix rgw not responding occasionally when receiving SIGHUP signal ([issue#20962](#), [pr#17113](#), Yao Zongyou)
- rgw: luminous: RGW: Get Bucket ACL does not honor the s3:GetBucketACL action ([issue#21013](#), [issue#21056](#), [pr#17117](#), Abhishek Lekshmanan)
- rgw: luminous: rgw: GetObject Tagging needs to exit earlier if the object has no attributes ([issue#21054](#), [issue#21010](#), [pr#17118](#), Abhishek Lekshmanan)
- rgw: luminous: rgw_lc: support for AWSv4 authentication ([pr#16914](#), Abhishek Lekshmanan)
- rgw: luminous: S3 v4 auth fails when query string contains ([issue#21000](#), [issue#21003](#), [issue#20501](#), [issue#21043](#), [pr#17114](#), Zhang Shaowen, Marcus Watts)
- rgw: luminous: rgw: Use namespace for lc_pool and roles_pool ([issue#20177](#), [issue#20967](#), [pr#16943](#), Orit Wasserman)
- rgw: make RGWEnv return a const ref. to its map ([pr#15269](#), Abhishek Lekshmanan)
- rgw: make sending Content-Length in 204 and 304 responses controllable ([issue#16602](#), [pr#10156](#), Radoslaw Zarzynski)
- rgw: make sync thread name clear ([issue#18860](#), [pr#13324](#), lvshuhua)
- rgw: match wildcards in StringLike policy conditions ([issue#20308](#), [pr#16491](#), Casey Bodley)
- rgw: metadata search part 2 ([pr#14351](#), Yehuda Sadeh)
- rgw: meta sync thread crash at RGWMetaSyncShardCR ([issue#20251](#), [pr#15660](#), fang.yuxiang)
- rgw: migrate atomic_t to std::atomic<> (ebirah) ([pr#14839](#), Jesse Williamson)
- rgw: migrate atomic_t to std::atomic<> ([pr#15001](#), Jesse Williamson)
- rgw: modify email to empty by admin RESTful api doesn't work ([pr#16309](#), fang.yuxiang)
- rgw: move the S3 anonymous auth handling to a dedicated engine ([pr#16485](#), Radoslaw Zarzynski)
- rgw: multipart copy-part remove '/' for s3 java sdk request header ([issue#20075](#), [pr#15283](#), root)
- rgw: multisite enabled over multiple clusters ([pr#12535](#), Ali Maredia)
- rgw: multisite: fixes for zonegroup redirect ([issue#19488](#), [pr#14319](#), Casey Bodley)
- rgw:multisite: fix RGWRadosRemoveOmapKeysCR and change cn to intrusive_ptr ([issue#20539](#), [pr#16197](#), Shasha Lu)
- rgw: never let http_redirect_code of RGWRedirectInfo to stay uninitialized ([issue#20774](#), [pr#16601](#), Radoslaw Zarzynski)
- rgw: omit X-Account-Access-Control if there is no grant to serialize ([issue#20395](#), [pr#15883](#), Radoslaw Zarzynski)
- rgw: only log metadata on metadata master zone ([issue#20244](#), [pr#15613](#), Casey Bodley)
- rgw: optimize data sync. Add zones_trace in log to avoid needless sync ([issue#19219](#), [pr#13851](#), Zhang Shaowen)
- rgw: optimize generating torrent file. Object data won't stay in memory now ([pr#15153](#), Zhang Shaowen)
- rgw: orphan: fix error messages ([pr#12782](#), Weibing Zhang)
- rgw: pass authentication domain to civetweb ([issue#17657](#), [pr#12861](#), Abhishek Lekshmanan)
- rgw: polymorphic error codes ([pr#10690](#), Pritha Srivastava, Marcus Watts)
- rgw: print is_admin as int instead of __u8 ([pr#12264](#), Casey Bodley)
- rgw: put object's acl can't work well on the latest object ([issue#18649](#), [pr#13078](#), Zhang Shaowen)
- rgw: radosgw-admin: use zone id when creating a zone ([issue#19498](#), [pr#14340](#), Orit Wasserman)
- rgw: radosgw-admin: warn that 'realm rename' does not update other clusters ([issue#19746](#), [pr#14722](#), Casey Bodley)
- rgw: radosgw, crypto: simplified code in handle_data functions ([pr#15598](#), Adam Kupczyk)
- rgw: radosgw: fix compilation with cryptopp ([pr#15960](#), Adam Kupczyk)
- rgw: raise debug level of meta sync logging ([pr#15524](#), Casey Bodley)
- rgw: raise debug level of RGWPostObj_ObjStore_S3::get_policy ([pr#16203](#), Shasha Lu)
- rgw: reject request if decoded URI contains 0 in the middle ([issue#20418](#), [pr#15953](#), Radoslaw Zarzynski)
- rgw: remove a redundant judgement in rgw_rados.cc:delete_obj ([pr#11124](#), Zhang Shaowen)
- rgw: Removed Unwanted headers ([pr#14183](#), Jos Collin)
- rgw: remove duplicate flush formatter ([pr#12437](#), Guo Zhandong)
- rgw: remove extra RGWMPObj in rgw_multi.h ([pr#14619](#), Casey Bodley)
- rgw: remove fastcgi from default rgw frontends ([pr#15098](#), Casey Bodley)
- rgw: remove invalid read size4 ([issue#18071](#), [pr#12767](#), Matt Benjamin)
 - rgw: Remove pessimizing move
- rgw: remove redundant codes in rgw_cache.h ([pr#13902](#), lihongjie)
- rgw: Remove spurious XML header for GetBucketPolicy ([issue#20247](#), [pr#15586](#), Adam C. Emerson)
- rgw: remove the useless output when listing zonegroups ([pr#16331](#), Zhang Shaowen)
- rgw: remove unused func in rgw_file.h ([pr#15698](#), lihongjie)
- rgw: remove useless -tier_type in radosgw-admin ([pr#13856](#), Zhang Shaowen)
- rgw: rename s3_code to err_code for swift ([pr#12300](#), Guo Zhandong)
- rgw: Replace get_zonegroup().is_master_zonegroup() with is_meta_master() in RGWBulkDelete::Deleter::delete_single() ([pr#16062](#), Fan Yang)
- rgw: req xml params size limitation error msg ([pr#16310](#), Enming Zhang)
- rgw: respect Swift's negative, HTTP referer-based ACL grants ([issue#18841](#), [pr#14344](#), Radoslaw Zarzynski)
- rgw: restore admin socket path in mrgw.sh ([pr#16540](#), Casey Bodley)
- rgw: return the version id in get object and object metadata request ([issue#19370](#), [pr#14117](#), Zhang Shaowen)
- rgw: rgw-admin: fix bucket limit check argparse, div(0) ([pr#15316](#), Matt Benjamin)
- rgw: rgw-admin: remove deprecated regionmap commands ([issue#18725](#), [pr#13963](#), Casey Bodley)
- rgw: rgw_common: use string::npos for the results of str.find ([pr#14341](#), Abhishek Lekshmanan)
- rgw: rgw_crypt: log error messages during failures ([pr#16726](#), Abhishek Lekshmanan)

- rgw: rgw_file: add compression interop to RGW NFS ([issue#20462](#), [pr#15989](#), Matt Benjamin)
- rgw: rgw_file: add lock protection for readdir against gc ([issue#20121](#), [pr#15329](#), Gui Hecheng)
- rgw: rgw_file: add service map registration ([pr#16251](#), Matt Benjamin)
- rgw: rgw_file: add timed namespace invalidation ([issue#18651](#), [pr#13038](#), Matt Benjamin)
- rgw: rgw_file: avoid a recursive lane lock in LRU drain ([issue#20374](#), [pr#15819](#), Matt Benjamin)
- rgw: rgw_file: avoid stranding invalid-name bucket handles in fhcache ([issue#19036](#), [pr#13590](#), Matt Benjamin)
- rgw: rgw_file cleanup names ([pr#15568](#), Gui Hecheng)
- rgw: rgw_file: cleanup virtual keyword on derived functions ([pr#14908](#), Gui Hecheng)
- rgw: rgw_file: ensure valid_s3_object_name for directories, too ([issue#19066](#), [pr#13614](#), Matt Benjamin)
- rgw: rgw_file: fix assert upon setattr on bucket ([issue#20287](#), [pr#15679](#), Gui Hecheng)
- rgw: rgw_file: fix double unref on rgw_fh for rename ([pr#13988](#), Gui Hecheng)
- rgw: rgw_file: fix flags set on unsuccessful unlink ([pr#15222](#), Gui Hecheng)
- rgw: rgw_file: fix fs_inst progression ([issue#19214](#), [pr#13832](#), Matt Benjamin)
- rgw: rgw_file: fix missing unlock in unlink ([issue#19435](#), [pr#14262](#), Gui Hecheng)
- rgw: rgw_file: fix misuse of make_key_name before make_fhk ([pr#15108](#), Gui Hecheng)
- rgw: rgw_file: fix non-negative return code for open operation ([pr#14045](#), Gui Hecheng)
- rgw: rgw_file: fix non-posix errcode EINVAL to ENAMETOOLONG ([pr#13764](#), Gui Hecheng)
- rgw: rgw_file: fix readdir after dirent-change ([issue#19634](#), [pr#14561](#), Matt Benjamin)
- rgw: rgw_file: fix reversed return value of getattr ([pr#13895](#), Gui Hecheng)
- rgw: rgw_file: fix RGWLibFS::setattr for directory objects ([issue#18808](#), [pr#13252](#), Matt Benjamin)
- rgw: rgw_file: fix up potential race condition ([pr#14553](#), Gui Hecheng)
- rgw: rgw_file: implement reliable has-children check (unlink dir) ([issue#19270](#), [pr#13953](#), Matt Benjamin)
- rgw: rgw_file: interned RGWFileHandle objects need parent refs ([issue#18650](#), [pr#13084](#), Matt Benjamin)
- rgw: rgw_file: posix style atime,ctime,mtime ([pr#13765](#), Gui Hecheng)
- rgw: rgw_file: pre-compute unix attrs in write_finish() ([issue#19653](#), [pr#14609](#), Matt Benjamin)
- rgw: rgw_file: prevent conflict of mkdir between restarts ([issue#20275](#), [pr#15655](#), Gui Hecheng)
- rgw: rgw_file: properly & or flags ([issue#20663](#), [pr#16448](#), Matt Benjamin)
- rgw: rgw_file: release rgw_fh lock and ref on ENOTEMPTY ([issue#20061](#), [pr#15246](#), Matt Benjamin)
- rgw: rgw_file: removed extra rele() on fs in rgw_umount() ([pr#15152](#), Gui Hecheng)
- rgw: rgw_file: remove hidden uxattr objects from buckets on delete ([issue#20045](#), [pr#15210](#), Matt Benjamin)
- rgw: rgw_file: remove post-unlink lookup check ([issue#20047](#), [pr#15216](#), Matt Benjamin)
- rgw: rgw_file: replace raw fs->fh_lru.unref with predefined fs->unref ([pr#15541](#), Gui Hecheng)
- rgw: rgw_file: RGWFileHandle dtor must also cond-unlink from FHCache ([issue#19112](#), [pr#13712](#), Matt Benjamin)
- rgw: rgw_file skip policy read for virtual components ([pr#16034](#), Gui Hecheng)
- rgw: rgw_file: split last argv on ws, if provided ([pr#12965](#), Matt Benjamin)
- rgw: rgw_file: store bucket uxattrs on the bucket ([issue#20082](#), [pr#15293](#), Matt Benjamin)
- rgw: rgw_file: support readdir cb type hints (plus fixes) ([issue#19623](#), [issue#19625](#), [issue#19624](#), [pr#14458](#), Matt Benjamin)
- rgw: rgw_file: use fh_hook::is_linked() to check residence ([issue#19111](#), [pr#13703](#), Matt Benjamin)
- rgw: rgw_file: v3: fix write-timer action ([issue#19932](#), [pr#15097](#), Matt Benjamin)
- rgw: rgw : fix race in RGWCompleteMultipart ([issue#20861](#), [pr#16732](#), Abhishek Varshney)
- rgw: rgw : fix s3 aws v2 signature priority between header['X-Amz-Date'] and header['Date'] ([issue#20176](#), [pr#15467](#), yuliyang)
- rgw: rgw: fix the subdir without slash of s3 website url ([issue#20307](#), [pr#15703](#), liuhong)
- rgw: rgw_lc: drop a bunch of unused headers ([pr#14342](#), Abhishek Lekshmanan)
- rgw: rgw_ldap: log the ldap err in case of bind failure ([pr#14781](#), Abhishek Lekshmanan)
- rgw: rgw/lifecycle: do not send lifecycle rules when GetLifeCycle failed ([issue#19363](#), [pr#14160](#), liuchang0812)
- rgw: RGWMetaSyncShardControlCR retries with backoff on all error codes ([issue#19019](#), [pr#13546](#), Casey Bodley)
- rgw: RGWMetaSyncShardCR drops stack refs on destruction ([issue#18412](#), [issue#18300](#), [pr#12605](#), Casey Bodley)
- rgw: rgw multisite: automated mdlog trimming ([pr#13111](#), Casey Bodley)
- rgw: rgw multisite: feature of bucket sync enable/disable ([pr#15801](#), Zhang Shaowen, Casey Bodley, Zengran Zhang)
- rgw: rgw multisite: fixes for meta sync across periods ([issue#18639](#), [pr#13070](#), Casey Bodley)
- rgw: rgw multisite: fix ref counting of completions ([issue#18414](#), [issue#18407](#), [pr#12841](#), Casey Bodley)
- rgw: rgw-multisite: fix the problem of rgw website configure 'RedirectAllRequestsTo' failed to sync to slave zone ([pr#15036](#), yuliyang)
- rgw: rgw-multisite: fix the problem of rgw website configure request not redirect to metadata master ([pr#15082](#), yuliyang)
- rgw: rgw multisite: remove the redundant post in OPT_ZONEGROUP_MODIFY ([pr#14359](#), Jing Wenjun)
- rgw: rgw/multisite: validate bucket location during bucket creation ([pr#15333](#), Jiaying Ren)
- rgw: RGW NFS: add nfs.rst to doc/radosgw ([pr#15789](#), Matt Benjamin)
- rgw: rgw_op: remove unused variable iter ([pr#14276](#), Weibing Zhang)
- rgw: RGWPeriodPusher spawns http thread before cr thread ([issue#19834](#), [pr#14936](#), Casey Bodley)
- rgw: rgw_rados: create sync module instances only if run_sync_thread is set ([issue#19830](#), [pr#14994](#), Abhishek Lekshmanan)
- rgw: rgw_rados drop deprecated global var ([pr#14411](#), Jiaying Ren)

- rgw: rgw_rados: initialize cur_shard ([pr#15735](#), Abhishek Lekshmanan)
- rgw: rgw realm set fixes ([issue#18333](#), [pr#12731](#), Orit Wasserman)
- rgw: rgw/rgw_frontend.h: Return negative value for empty uid in RGWLoadGenFrontend::init() ([pr#16204](#), jimifm)
- rgw: rgw/rgw_main.cc: fix parentheses and function result ([pr#12295](#), Willem Jan Withagen)
- rgw: rgw/rgw_op: Prevents memory leaks when calling func swift_versioning_copy() fails ([pr#15328](#), jimifm)
- rgw: rgw/rgw_rados: Remove duplicate calls in RGWRados::finalize() ([pr#15281](#), jimifm)
- rgw: rgw/rgw_string.h: FreeBSD would like errno.h included ([pr#15737](#), Willem Jan Withagen)
- rgw: rgw/rgw_swift_auth.cc: using string::back() instead as the C++11 recommend ([pr#14827](#), liuyuhong)
- rgw: rgw structures rework ([issue#17996](#), [issue#19249](#), [pr#11485](#), Yehuda Sadeh)
- rgw: rgw,test: fix rgw placement rule pool config option ([pr#16084](#), Jiaying Ren)
- rgw: S3 lifecycle now supports expiration date ([pr#15807](#), Zhang Shaowen)
- rgw: s3 server-side encryption (SSE-C, SSE-KMS) ([pr#11049](#), Adam Kupczyk, Casey Bodley, Radoslaw Zarzynski)
- rgw: segment fault when shard id out of range ([issue#19732](#), [pr#14389](#), redickwang)
- rgw: set dumpable flag after setuid post ff0e521 ([issue#19089](#), [pr#13657](#), Brad Hubbard)
- rgw: set FCGI_INCLUDE_DIR for ceph rgw_base ([issue#18918](#), [pr#13393](#), David Disseldorf)
- rgw: set object accounted size correctly ([issue#20071](#), [pr#14950](#), fang yuxiang)
- rgw: set placement rule properly ([pr#15221](#), fang.yuxiang)
- rgw: should delete in_stream_req if conn->get_obj(...) return not zero value ([pr#9950](#), weiqiaomiao)
- rgw: should not restrict location_constraint same when user not provide ([pr#16770](#), Tianshan Qu)
- rgw: should unlock when reshards_log->update() return non-zero in RGWB... ([pr#16502](#), Wei Qiaomiao)
- rgw: silence compile warning from -Wmaybe-uninitialized ([pr#15996](#), Jiaying Ren)
- rgw: silence warning from -Wmaybe-uninitialized ([pr#15949](#), Jos Collin)
- rgw: stat requests skip compression, manifest handling, etc ([pr#14109](#), Casey Bodley)
- rgw: Support certain archaic and antiquated distributions([pr#15498](#), Adam C. Emerson)
- rgw: swift: ability to update swift read and write acls separately ([issue#19289](#), [pr#14499](#), Marcus Watts)
- rgw: swift: disable revocation thread if sleep == 0 ([issue#19499](#), [issue#9493](#), [pr#14501](#), Marcus Watts)
- rgw: swift: fix anonymous user's error code of getting object ([issue#18806](#), [pr#13242](#), Jing Wenjun)
- rgw: swift: the http referer acl in swift API should be shown ([issue#18665](#), [pr#13003](#), Jing Wenjun)
- rgw: swift: The http referer should be parsed to compare in swift API ([issue#18685](#), [pr#13005](#), Jing Wenjun)
- rgw: switch from "timegm()" to "internal_timegm()" for better portability ([issue#12863](#), [pr#14327](#), Rishabh Kumar)
- rgw: switch to std::array in RGWBulkUploadOp due to C++11 and FreeBSD ([pr#14314](#), Radoslaw Zarzynski)
- rgw: sync status compares the current master period ([issue#18064](#), [pr#12907](#), Abhishek Lekshmanan)
- rgw: test,rgw: fix rgw placement rule pool config option ([pr#16380](#), Jiaying Ren)
- rgw,tests: luminous: qa/rgw: use 'ceph osd pool application enable' on created pools ([pr#17259](#), Casey Bodley)
- rgw,tests: qa/rgw: add cluster name to path when s3tests scans rgw log ([pr#14845](#), Casey Bodley)
- rgw,tests: qa/rgw: add configuration for server-side encryption tests ([pr#13597](#), Casey Bodley)
- rgw,tests: qa/rgw: add encryption config for s3tests under thrash ([pr#15694](#), Casey Bodley)
- rgw,tests: qa/rgw: add multisite suite to configure and run multisite tests ([pr#14688](#), Casey Bodley)
- rgw,tests: qa/rgw: disable lifecycle tests because of expiration failures ([pr#16760](#), Casey Bodley)
- rgw,tests: qa/rgw: don't scan radosgw logs for encryption keys on jewel upgrade test ([pr#14697](#), Casey Bodley)
- rgw,tests: qa/rgw: fix assertions in radosgw_admin task ([pr#14842](#), Casey Bodley)
- rgw,tests: qa/rgw: remove apache/fastcgi and radosgw-agent tests ([pr#15184](#), Casey Bodley)
- rgw,tests: qa/suites/rgw/thrash: add osd thrashing tests ([pr#13445](#), Sage Weil)
- rgw,tests: qa/tasks: S3A hadoop task to test s3a with Ceph ([pr#14624](#), Vasu Kulkarni)
- rgw,tests: test/rgw: add bucket acl and versioning tests to test_multi.py ([pr#12449](#), Casey Bodley)
- rgw,tests: test/rgw: add test for versioned object sync ([pr#12474](#), Casey Bodley)
- rgw,tests: test/rgw: fixes for test_multi_period_incremental_sync() ([pr#13067](#), Casey Bodley)
- rgw,tests: test/rgw: fix for empty lists as default arguments ([pr#14816](#), Casey Bodley)
- rgw,tests: test/rgw: refactor test_multi.py for use in qa suite ([pr#14433](#), Casey Bodley)
- rgw,tests: test/rgw: test_bucket_delete_notempty in test_multi.py ([pr#14090](#), Casey Bodley)
- rgw,tests: vstart: add rgw configuration needed to pass all s3tests ([pr#15782](#), Casey Bodley)
- rgw,tests: vstart: remove rgw_enable_static_website ([pr#15856](#), Casey Bodley)
- rgw: the swift container acl should support field .ref ([issue#18484](#), [pr#12874](#), Jing Wenjun)
- rgw: Turn off fcgi as a frontend ([issue#16784](#), [pr#15070](#), Thomas Serlin)
- rgw: Uninitialized member in LCRule ([pr#15827](#), Jos Collin)
- rgw: update Beast for streaming reads in asio frontend ([pr#14273](#), Casey Bodley)
- rgw: update bucket cors in secondary zonegroup should forward to master ([issue#16888](#), [pr#15260](#), Shasha Lu)
- rgw: update function doc in rgw_rados.h and rgw_rados.cc ([pr#15803](#), Jiaying Ren)
- rgw: update is_truncated in function rgw_read_user_buckets ([issue#19365](#), [pr#14343](#), liuchang0812)
- rgw: usage ([issue#16191](#), [pr#14287](#), Ji Chen, Orit Wasserman)
- rgw: use 64-bit offsets for compression ([issue#20231](#), [pr#15656](#), Adam Kupczyk, fang yuxiang)
- rgw: use a namespace for rgw reshards pool for upgrades as well ([issue#20289](#), [pr#16368](#), Karol Mroz, Abhishek Lekshmanan)
- rgw: Use comparison instead of assignment ([pr#16653](#), amitkuma)
- rgw: Use decoded URI when verifying TempURL ([issue#18590](#), [pr#13007](#), Michal Koutný)

- rgw: use get_data_extra_pool() when get extra pool ([issue#20064](#), [pr#15219](#), fang yuxiang)
- rgw: use pre-defined calls to replace raw flag operation ([pr#15107](#), Gui Hecheng)
- rgw: use rgw_zone_root_pool for region_map like is done in hammer ([issue#19195](#), [pr#13928](#), Orit Wasserman)
- rgw: use separate http_manager for read_sync_status ([issue#19236](#), [pr#13660](#), Shasha Lu)
- rgw: use uncompressed size for range_to_ofs() in slo/dlo ([pr#15931](#), Casey Bodley)
- rgw: using RGW_OBJ_NS_MULTIPART in check_bad_index_multipart ([pr#15774](#), Shasha Lu)
- rgw: using the same bucket num_shards as master zg when create bucket in secondary zg ([issue#19745](#), [pr#14388](#), Shasha Lu)
- rgw: validate tenant names during user create ([pr#16442](#), Abhishek Lekshmanan)
- rgw: verify md5 in post obj ([issue#19739](#), [pr#14961](#), Yehuda Sadeh)
- rgw: version id doesn't work in fetch_remote_obj ([pr#14010](#), Zhang Shaowen)
- rgw: VersionIdMarker and NextVersionIdMarker should be returned when listing object versions ([issue#19886](#), [pr#15014](#), Zhang Shaowen)
- rgw: warning, output may be truncated before the last format character ([pr#14194](#), Jos Collin)
- rgw: when create_bucket use the same num_shards with info.num_shards ([issue#19745](#), [pr#15010](#), Shasha Lu)
- rgw: wip dir orphan ([issue#18992](#), [issue#18989](#), [issue#19018](#), [issue#18991](#), [pr#13529](#), Matt Benjamin)
- rgw: Wip librgw refcnt ([pr#13405](#), Matt Benjamin)
- rgw: wip parentref ([issue#19060](#), [issue#19059](#), [pr#13607](#), Matt Benjamin)
- rgw: Wip rgw fix prefix list ([issue#19432](#), [pr#15916](#), Giovani Rinaldi, Orit Wasserman)
- rgw: Wip rgw openssl 7 ([issue#11239](#), [issue#16535](#), [pr#11776](#), Yehuda Sadeh, Marcus Watts)
- rgw: wip: rgw: rest_admin/user avoid double checking input args ([pr#13460](#), Abhishek Lekshmanan)
- tests: Add integration tests for admin socket output ([pr#15223](#), Brad Hubbard)
- tests: add MGR=1 so 'pg dump' won't be blocked ([pr#14266](#), Kefu Chai)
- tests: Add openstack requirements to smoke suite ([pr#12913](#), Zack Cerza)
- tests: add setup/teardown for asok dir ([pr#16523](#), Kefu Chai)
- tests: buildpackages: remove because it does not belong ([issue#18846](#), [pr#13297](#), Loic Dachary)
- tests: ceph-disk: add setting for external py-modules for tox-testing ([pr#15433](#), Willem Jan Withagen)
- tests: ceph-disk: use communicate() instead of wait() for output ([pr#16347](#), Kefu Chai)
- tests: ceph-helpers.sh reduce get_timeout_delays() verbosity ([pr#13257](#), Kefu Chai)
- tests: ceph_objectstore_tool.py: kill all daemons ([pr#14428](#), Kefu Chai)
- tests: ceph_test_objectstore: tolerate fsck EOPNOTSUPP too ([pr#13325](#), Sage Weil)
- tests: ceph_test_rados_api_tier: tolerate ENOENT from 'pg scrub' ([pr#14807](#), Sage Weil)
- tests: ceph_test_rados_api_watch_notify: move global variables into test class ([issue#18395](#), [pr#12751](#), Kefu Chai)
- tests: ceph_test_rados_api_watch_notify: test timeout using rados_wat... ([issue#19312](#), [pr#14061](#), Kefu Chai)
- tests: cephtool/test.sh error on full tests ([issue#19698](#), [pr#14647](#), Willem Jan Withagen, David Zafman)
- tests: cephtool/test.sh: Only delete a test pool when no longer needed ([pr#16443](#), Willem Jan Withagen)
- tests: cls_lock: move lock_info_t definition to cls_lock_types.h ([pr#16091](#), runsisi)
- tests: config_opts: drop unused opts ([pr#15031](#), Kefu Chai)
- tests: Decreased amount of jobs on master, kraken, luminous runs ([pr#17074](#), Yuri Weinstein)
- tests: Don't dump core when using EXPECT_DEATH ([pr#14821](#), Kefu Chai, Brad Hubbard)
- tests: drop buildpackages.py ([issue#18846](#), [pr#13319](#), Nathan Cutler)
- tests: drop obsolete Perl scripts ([pr#13951](#), Nathan Cutler)
- tests: drop rbd_cli_tests.pl and RbdLib.pm ([issue#14825](#), [pr#12821](#), Nathan Cutler)
- tests: drop unused rbd_functional_tests.pl script ([issue#14825](#), [pr#12818](#), Nathan Cutler)
- tests: fio_ceph_objectstore: fixes improper write request data lifetime ([pr#14338](#), Adam Kupczyk)
- tests: fix broken links in upgrade/hammer-jewel-x/stress-split ([issue#19793](#), [pr#14831](#), Nathan Cutler)
- tests: fix NULL references to be acceptable by Clang ([pr#12880](#), Willem Jan Withagen)
- tests: fix rados/upgrade/jewel-x-singleton and make workunit task handle repo URLs not ending in ".git" ([issue#20554](#), [issue#20368](#), [pr#16228](#), Nathan Cutler, Sage Weil)
- tests: fix regression in qa/tasks/ceph_master.py ([issue#16263](#), [pr#13279](#), Nathan Cutler, Kefu Chai)
- tests: fix template specialization of PromoteRequest class ([pr#12815](#), Ricardo Dias)
- tests: ignore bogus ceph-objectstore-tool error in ceph_manager ([issue#16263](#), [pr#13194](#), Nathan Cutler)
- tests: include/denc: support ENCODE_DUMP ([pr#14962](#), Sage Weil)
- tests: libradosstriper: do not assign garbage to returned value ([pr#15009](#), Kefu Chai)
- tests: luminous: tests: qa/standalone: misc fixes ([issue#20465](#), [issue#20921](#), [issue#20979](#), [pr#16985](#), David Zafman)
- tests: mgr,os,test: kill clang analyzer warnings ([pr#16227](#), Kefu Chai)
- tests: move swift.py task from teuthology to ceph, phase one (master) ([issue#20392](#), [pr#15859](#), Nathan Cutler, Sage Weil, Warren Usui, Greg Farnum, Ali Maredia, Tommi Virtanen, Zack Cerza, Sam Lang, Yehuda Sadeh, Joe Buck, Josh Durgin)
- tests: nosetests: use /usr/bin/env to find nosetests ([pr#12091](#), Willem Jan Withagen)
- tests: os: Argument cannot be negative ([pr#16688](#), amitkuma)
- tests: os/bluestore,test/ceph_test_objectstore: silence gcc warnings ([pr#13924](#), Kefu Chai)
- tests: osd-scrub-repair.sh disable scrub backoff in test ([pr#13334](#), Kefu Chai)
- tests: qa: Added luminous to the mix in schedule_subset.sh ([pr#16430](#), Yuri Weinstein)
- tests: qa/added overrides ([pr#14917](#), Yuri Weinstein)

- tests: qa: Add reboot case for systemd test ([issue#19717](#), [pr#14229](#), Vasu Kulkarni)
- tests: qa: add supported distros for ceph-ansible ([pr#13711](#), Tamil Muthamizhan)
- tests: qa: add task for dnsmasq configuration ([pr#15071](#), Casey Bodley)
- tests: [qa/ceph-deploy]: run create mgr nodes as well ([pr#16216](#), Vasu Kulkarni)
- tests: qa: Cleaned up distros to use *latest* versions ([pr#12804](#), Yuri Weinstein)
- tests: qa/distros: make centos_latest 7.3 ([pr#12944](#), Sage Weil)
- tests: qa/doc: document and fix tests for pool application warnings ([pr#16568](#), Sage Weil)
- tests: qa: do not mention ceph branch explicitly ([pr#13225](#), Tamil Muthamizhan)
- tests: qa: do not restrict valgrind runs to centos ([issue#18126](#), [pr#15893](#), Greg Farnum)
- tests: qa/erasure-code: override min_size to 2 ([issue#19770](#), [pr#14872](#), Kefu Chai)
- tests: qa: fixed distros links ([pr#12770](#), Yuri Weinstein)
- tests: qa/run-standalone.sh: fix the find option to be compatible with GNU find ([pr#16646](#), Kefu Chai)
- tests: qa: specify client for fs workunit ([pr#12914](#), Tamil Muthamizhan)
- tests: qa: split test_tiering into smaller pieces ([pr#15146](#), Kefu Chai)
- tests: qa/suite: Added a smoke suite for ceph-ansible ([pr#12610](#), Tamil Muthamizhan)
- tests: qa/suite: replace reference to fs/xfs.yaml ([pr#14756](#), Yehuda Sadeh)
- tests: qa/suites/ceph-ansible: removing fs workunit ([pr#12928](#), Tamil Muthamizhan)
- tests: qa/suites/{ceph-ansible,rest}: OpenStack volumes ([pr#13672](#), Zack Cerza)
- tests: qa/suites/ceph-deploy: Drop OpenStack volume count ([pr#13706](#), Zack Cerza)
- tests: qa/suites: drop 'fs' facet, and add 'objectstore' facet where missing ([pr#14198](#), Sage Weil)
- tests: qa/suites: escape the parenthesis of the whitelist text ([pr#16722](#), Kefu Chai)
- tests: qa/suites: fix upgrade tests vs cluster full thrashing ([pr#13852](#), Sage Weil)
- tests: qa/suites/fs: Add openstack volume configuration ([pr#13640](#), Zack Cerza)
- tests: qa/suites/jewel-x/point-to-point: don't scan for keys on second s3tests either ([pr#14788](#), Sage Weil)
- tests: qa/suites/kcephfs: Openstack volume configuration ([pr#13634](#), Zack Cerza)
- tests: qa/suites/{knfs,hadoop,samba}: OpenStack volume configuration ([pr#13637](#), Zack Cerza)
- tests: qa/suites/krbd: Add openstack volume configuration ([pr#13631](#), Zack Cerza)
- tests: qa/suites/powercycle/osd/whitelist_health: whitelist more ([pr#17306](#), Sage Weil)
- tests: qa/suites/powercycle: whitelist health for thrashing ([pr#16759](#), Sage Weil)
- tests: qa/suites/rados: a bit more whitelisting ([pr#16820](#), Sage Weil)
- tests: qa/suites/rados: fix ec thrashing ([pr#15087](#), Sage Weil)
- tests: qa/suites/rados/objectstore: enable experimental features for testing bluestore ([pr#13456](#), Kefu Chai, Dan Mick)
- tests: qa/suites/rados/singleton/all/erasure-code-nonregression: fix typo ([pr#16579](#), Sage Weil)
- tests: qa/suites/rados/singleton/all/mon-auth-caps: more osds so we can go clean ([pr#16225](#), Sage Weil)
- tests: qa/suites/rados/singleton-bluestore: concat settings ([pr#14884](#), Kefu Chai)
- tests: qa/suites/rados/singleton-nomsgr/all/multi-backfill-reject: sleep longer ([pr#16739](#), Sage Weil)
- tests: qa/suites/rados/singleton-nomsgr: fix syntax ([pr#15276](#), Sage Weil)
- tests: qa/suites/rados/thrash: make sure osds have map before legacy scrub ([pr#15117](#), Sage Weil)
- tests: qa/suites/rados/upgrade: restart mds ([pr#15517](#), Sage Weil)
- tests: qa/suites: Reduce fs combination tests for smoke, use bluestore ([pr#14854](#), Vasu Kulkarni)
- tests: qa/suites: Revert "qa/suites: add mon-reweight-min-pgs-per-osd = 4" ([pr#14584](#), Kefu Chai)
- tests: qa/suites/rgw: Add openstack volume configuration ([pr#13611](#), Zack Cerza)
- tests: qa/suites/upgarde/jewel-x/parallel: more whitelisting ([pr#16849](#), Sage Weil)
- tests: qa/suites/upgrade: add tiering test to hammer-jewel-x ([issue#19185](#), [pr#13805](#), Kefu Chai)
- tests: qa/suites/upgrade/hammer-jewel-x: add luminous.yaml ([issue#20342](#), [pr#15764](#), Kefu Chai)
- tests: qa/suites/upgrade/jewel-x: add mgr.x role ([pr#14689](#), Sage Weil)
- tests: qa/suites/upgrade/jewel-x: misc fixes for new health checks ([pr#16429](#), Sage Weil)
- tests: qa/suites/upgrade/kraken-x: do not thrash cluster full during upgrade ([issue#19232](#), [pr#13892](#), Dan Mick)
- tests: qa/suites/upgrade/kraken-x: misc fixes ([pr#14887](#), Sage Weil)
- tests: qa/suites/upgrade/kraken-x ([pr#13517](#), Sage Weil, Yuri Weinstein)
- tests: qa/suites/upgrade: set "sortbitwise" for jewel clusters ([pr#15661](#), Kefu Chai)
- tests: qa/suite/upgrade/jewel-x: various fixes ([pr#13734](#), Sage Weil)
- tests: qa/tasks: assert on pg status with a timeout ([issue#19594](#), [pr#14608](#), Kefu Chai)
- tests: qa/tasks/ceph: debug osd setup ([pr#16841](#), Sage Weil)
- tests: qa/tasks/ceph-deploy: create-keys explicitly ([pr#12867](#), Vasu Kulkarni)
- tests: qa/tasks/ceph-deploy: Fix bluestore options for ceph-deploy ([pr#16571](#), Vasu Kulkarni)
- tests: qa/tasks/ceph-deploy: use the new create option during instantiation ([pr#12892](#), Vasu Kulkarni)
- tests: qa/tasks/ceph: don't hard-code cluster name when copying fsid ([pr#16212](#), Jason Dillaman)
- tests: qa/tasks/ceph_manager: always fix pgp_num when done with thrashosd task ([issue#19771](#), [pr#14931](#), Kefu Chai)
- tests: qa/tasks/ceph_manager: 'ceph \$service tell ...' is obsolete ([pr#15252](#), Sage Weil)
- tests: qa/tasks/ceph.py: debug which pgs aren't scrubbing ([pr#13649](#), Sage Weil)
- tests: qa/tasks/ceph: raise exceptions if scrubbing fails or cannot proceed ([pr#15310](#), Sage Weil)
- tests: qa/tasks/ceph: should be "Waiting for all PGs", not "all osds" ([pr#16122](#), Kefu Chai)
- tests: qa/tasks/ceph: wait longer for scrub ([pr#16824](#), Sage Weil)
- tests: qa/tasks: few fixes to get ceph-deploy 1node to working state ([pr#14400](#), Vasu Kulkarni)

- tests: qa/tasks/radosbench: increase timeout ([pr#15885](#), Sage Weil)
- tests: qa/tasks/rebuild_mondb: grant "mgr:allow *" to client.admin ([issue#19439](#), [pr#14284](#), Kefu Chai)
- tests: qa/tasks/reg11184: use literal 'foo' instead pool_name ([pr#16451](#), Kefu Chai)
- tests: qa/tasks/repair_test: unset flags we set ([pr#15296](#), Sage Weil)
- tests: qa/tasks/rgw.py: start Apache before RadosGW ([pr#13846](#), Radoslaw Zarzynski)
- tests: qa/tasks/thrashosds-health.yaml: ignore MON_DOWN ([issue#20910](#), [pr#17003](#), Sage Weil)
- tests: qa/tasks: use sudo to check ceph health for systemd test ([pr#14464](#), Vasu Kulkarni)
- tests: qa/tasks/workunit.py: use "overrides" as the default settings of workunit ([issue#19429](#), [pr#14281](#), Kefu Chai)
- tests: qa/tasks/workunit: use ceph.git as an alternative of ceph-ci.git for cloning workunit ([pr#13663](#), Kefu Chai)
- tests: qa/tasks/workunit: use the suite repo for cloning workunit ([pr#13452](#), Kefu Chai)
- tests: qa/tasks/workunit: use the suite repo for cloning workunit ([pr#13625](#), Kefu Chai)
- tests: qa/test_rados_tool.sh: POSIX dd only accepts 'k' as multiplier ([pr#12699](#), Willem Jan Withagen)
- tests: qa: timeout when waiting for mgr to be available in healthy() ([pr#16797](#), Josh Durgin)
- tests: qa: Using centos 7.2 for *latest* version ([pr#12806](#), Yuri Weinstein)
- tests: qa/workunits/ceph-helpers: display rejected string ([issue#20344](#), [pr#14468](#), Kefu Chai)
- tests: qa/workunits/ceph-helpers: enable experimental features for osd ([pr#16319](#), Kefu Chai)
- tests: qa/workunits/ceph-helpers.sh: use syntax understood by jq 1.3 ([pr#15530](#), Kefu Chai)
- tests: qa/workunits/ceph-helpers: test wait_for_health_ok differently ([pr#16317](#), Kefu Chai)
- tests: qa/workunits/ceph-helpers: wait_for_clean() races with pg creation ([pr#12866](#), David Zafman)
- tests: qa/workunits/cephtool/test.sh: Be more liberal in testing health-output ([pr#14614](#), Willem Jan Withagen)
- tests: qa/workunits/cephtool/test.sh: "ceph osd stat" output changed, update accordingly ([pr#16444](#), Willem Jan Withagen, Kefu Chai)
- tests: qa/workunits/cephtool/test.sh: disable 'fs status' until bug is fixed ([issue#20761](#), [pr#16541](#), Sage Weil)
- tests: qa/workunits/cephtool/test.sh: fix test to watch audit channel ([pr#16470](#), Sage Weil)
- tests: qa/workunits/cephtool/test.sh: only include last line for epoch ([issue#20477](#), [pr#15770](#), Kefu Chai)
- tests: qa/workunits/rados/test.sh: print test name when it fails ([pr#13264](#), Kefu Chai)
- tests: rados: move cephtool.yaml to new singleton/bluestore subsuite ([issue#19797](#), [pr#14847](#), Nathan Cutler)
- tests: rbd/test_lock_fence.sh: fix rbdrw.py relative path ([issue#18388](#), [pr#12747](#), Nathan Cutler)
- tests: re-enable cephfs python tests on kclient ([issue#17193](#), [issue#18161](#), [pr#13200](#), Nathan Cutler)
- tests: remove temporary file ([pr#12919](#), Kefu Chai)
- tests: Revert "dummy: reduce run time, run user.yaml playbook" ([issue#18259](#), [pr#12506](#), Nathan Cutler)
- tests: Revert "qa/tasks/workunit: use the suite repo for cloning workunit" ([pr#13495](#), Sage Weil)
- tests: rgw.py: put client roles in a separate list ([issue#20417](#), [pr#15913](#), Nathan Cutler)
- tests: rgw.singleton: drop duplicate filestore-xfs.yaml ([pr#15959](#), Nathan Cutler)
- tests: set -x in suites/iozone.sh workunit ([issue#19740](#), [pr#14713](#), Nathan Cutler)
- tests: src/test/test_denc.cc: Fix errors in buffer overflow ([pr#12653](#), Willem Jan Withagen)
- tests: subst repo and branch in git.ceph.com URL in qa/tasks/cram.py and qa/tasks/qemu.py ([issue#18440](#), [pr#12816](#), Nathan Cutler)
- tests: tasks/workunit.py: when cloning, use -depth=1 ([pr#14214](#), Dan Mick)
- tests: test: add explicit braces to avoid ambiguous 'else' and to silence warnings ([pr#14472](#), Jos Collin)
- tests: test: add override in test submodule ([pr#13773](#), liuchang0812)
- tests: test: ceph osd stat out has changed, fix tests for that ([pr#16403](#), Willem Jan Withagen)
- tests: test:Check make_writeable() return value ([pr#15266](#), zhanglei)
- tests: test: clean up unused variable ([pr#12873](#), liuchang0812)
- tests: test/CMakeLists: disable test_pidfile.sh ([issue#20975](#), [pr#17241](#), Sage Weil)
- tests: test/compressor: disable isal tests if not available ([pr#14929](#), Kefu Chai)
- tests: test: c_read_operations.cc: silence warning from -Wsign-compare ([pr#14888](#), Jos Collin)
- tests: test: create asok files in a temp directory under \$TMPDIR ([issue#16895](#), [pr#16445](#), Kefu Chai)
- tests: test/crush: silence warnings from -Walloc-size-larger-than= and -Wstringop-overflow ([pr#15173](#), Jos Collin)
- tests: test: c_write_operations.cc: silence warning from -Wsign-compare ([pr#14889](#), Jos Collin)
- tests: test: Division by zero in Legacy::encode_n() ([pr#15902](#), Jos Collin)
- tests: test/encoding: fix readable.sh bugs; fix ceph-object-corpus ([pr#13678](#), Sage Weil)
- tests: test/fio_ceph_objectstore: fix fio plugin build failure by engine_data ([pr#15044](#), lisali)
- tests: test/fio: Fix assert in set_cache_shards in bluestore fio ([pr#15659](#), Xiaoyan Li)
- tests: test/fio: fix lack of setting for Sequencer::shard_hint ([pr#15571](#), Igor Fedotov)
- tests: test/fio: print all perfcounters rather than objectstore itself ([pr#16339](#), Jianpeng Ma)
- tests: test/fio: remove experimental option for bluestore & rocksdb ([pr#16263](#), Pan Liu)
- tests: test: Fixes for test_pidfile ([issue#20770](#), [pr#16587](#), David Zafman)
- tests: test: fixing assert that creates warning: comparison between signed and unsigned integer expressions ([pr#14794](#), Jos Collin)
- tests: test: Fix mismatched sign comparison in histogram test ([pr#13362](#), Adam C. Emerson)
- tests: test: Fix reg11184 test to remove extraneous pg ([pr#16265](#), David Zafman)
- tests: test: fix test_pidfile ([pr#13646](#), yaoning)
- tests: test/fsx: Remove the dead code associated with aio backend ([pr#14905](#), Zhou Zhengping)
- tests: test: Initialize pointer variables in TestMemIoCtxImpl ([pr#16785](#), amitkuma)

- tests: test/librados: Initialize member variables in aio.cc ([pr#16845](#), amitkuma)
- tests: test: librados_test_stub: tmap_update: return -ENOENT when removing nonexistent key ([pr#12667](#), Mykola Golub)
- tests: test: migrate atomic_t to std::atomic ([pr#14655](#), Jesse Williamson)
- tests: test,mon,msg: kill clang analyzer warnings ([pr#16320](#), Kefu Chai)
- tests: test/mon: silence warnings from -Wreorder ([pr#15692](#), Jos Collin)
- tests: test/msgr: fixed the hang issue for perf_msg_client ([pr#16358](#), Pan Liu)
- tests: test/msgr: silence warnings from -Wsign-compare ([pr#15356](#), Jos Collin)
- tests: test/msgr: silence warnings from -Wsign-compare ([pr#15570](#), Jos Collin)
- tests: test: objectstore: chain_xattr: fix wrong memset usage to fill buf ([pr#14277](#), Weibing Zhang)
- tests: test/objectstore: Check apply_transaction() return values ([pr#15171](#), zhanglei)
- tests: test/objectstore/: Check put_ref return value ([pr#15007](#), zhanglei)
- tests: test/old: Removed commented code ([pr#15366](#), Jos Collin)
- tests: test/osdc: fix comparison error and silence warning from -Wunused-value ([pr#15353](#), Willem Jan Withagen)
- tests: test/osd: kill compile warning ([pr#16669](#), Yan Jun)
- tests: test/osd/osd-dup.sh: lower wb fd throttle limits ([pr#14984](#), Dan Mick)
- tests: test/osd/osd-dup.sh: use wait_for_clean ([pr#15722](#), Dan Mick)
- tests: test/osd/osd-scrub-repair.sh: disable ec_overwrite tests on FreeBSD ([pr#15445](#), Willem Jan Withagen)
- tests: test/osd/osd-scrub-repair.sh: Fix diff options on FreeBSD ([pr#15914](#), Willem Jan Withagen)
- tests: test/osd: Removed Commented Code - 2 ([issue#20207](#), [pr#15540](#), Jos Collin)
- tests: test: osd/TestOSDMap.cc: fix Clang complain about promotion ([pr#15525](#), Willem Jan Withagen)
- tests: test/rados: fix wrong parameter order of RETURN1_IF_NOT_VAL ([pr#16589](#), Yan Jun)
- tests: test: reg11184 might not always find pg 2.0 prior to import ([pr#16610](#), David Zafman)
- tests: test: Rename FileJournal object to distinguish ([pr#15279](#), Jos Collin)
- tests: test: replace hard-code binary names with variables ([pr#12675](#), liuchang0812)
- tests: test: sed on FreeBSD requires “-i extension”, so use gsed ([pr#13903](#), Willem Jan Withagen)
- tests: test: s/osd_objectstore_type/osd_objectstore ([pr#16469](#), xie xinguo)
- tests: test: test_denc.cc: silence warning from -Wsign-compare ([pr#15355](#), Jos Collin)
- tests: test: Test fix for SnapSet change ([pr#15161](#), David Zafman)
- tests: test: test_pidfile running 2nd mon has unreliable log output ([pr#16635](#), David Zafman)
- tests: test: Thrasher: do not update pools_to_fix_pgp_num if nothing happens ([pr#13518](#), Kefu Chai)
- tests: test: Thrasher: update pgp_num of all expanded pools if not yet ([pr#13367](#), Kefu Chai)
- tests: test/unittest_bluefs: check whether mounted success ([pr#14988](#), shiqi)
- tests: test/unittest_bluefs: remove unused variable ([pr#14006](#), shiqi)
- tests: test: unittest_hostname compile error on freebsd ([pr#13739](#), liuchang0812)
- tests: test: update test_rados_tool.sh, use POOL and OBJ var ([pr#12706](#), liuchang0812)
- tests: test: use 7130 for crush-classes.sh ([pr#14783](#), Loic Dachary)
- tests: test/vstart_wrapper.sh: display_log on test failure ([pr#15620](#), Kefu Chai)
- tests: test: warning: comparison between signed and unsigned integer expressions ([pr#14705](#), Jos Collin)
- tests: Thrasher: eliminate a race between kill_osd and __init__ ([issue#18799](#), [pr#13237](#), Nathan Cutler)
- tests: Thrasher: handle “OSD has the store locked” gracefully ([issue#19556](#), [pr#14415](#), Nathan Cutler)
- tests,tools: script/find_dups_in_pg_log: scrip to find dup requests due to short pg logs ([pr#13417](#), Sage Weil)
- tests,tools: test, ceph-osdomap-tool: kill clang warnings ([pr#15905](#), Kefu Chai)
- tests,tools: test: kill warnings ([pr#14892](#), Kefu Chai)
- tests: update SUSE yaml facets in qa/distros/all ([issue#18856](#), [pr#13313](#), Nathan Cutler)
- tests: workunit: request branch when cloning ([pr#14260](#), Kefu Chai, Dan Mick)
- tools: add override in tool submodule ([pr#13776](#), liuchang0812)
- tools: brag: count the number of mds in fsmap not in mdsmmap ([issue#19192](#), [pr#13798](#), Peng Zhang)
- tools: ceph_common.sh: fix syntax error ([issue#17826](#), [pr#13419](#), Dan Mick)
- tools: ceph-conf: fix typo in usage: ‘mon add’ should be ‘mon addr’ ([pr#15935](#), Peng Zhang)
- tools: ceph-create-keys: add an argument to override default 10-minute timeout ([pr#16049](#), Douglas Fuller)
- tools: ceph-detect-init: adding Arch Linux support ([pr#12787](#), Jamin W. Collins)
- tools: ceph-detect-init: Adds Oracle Linux Server and Oracle VM Server detect ([pr#13917](#), Nikita Gerasimov)
- tools: ceph-detect-init: detect init system by poking the system ([issue#19884](#), [pr#15043](#), Kefu Chai)
- tools: ceph-disk: Add fix subcommand ([pr#13310](#), Boris Ranto)
- tools: ceph-disk: change the lockbox partition number to 5 ([issue#20556](#), [pr#16247](#), Shangzhong Zhu)
- tools: ceph-disk: command invocation needs all fields separate ([pr#15733](#), Willem Jan Withagen)
- tools: ceph-disk: convert none str to str before printing it ([issue#18371](#), [pr#12760](#), Kefu Chai)
- tools: ceph-disk: do not remove mount point if deactivate -once ([pr#16474](#), Song Shun)
- tools: ceph-disk: Fix for missing ‘not’ in *_is_diskdevice checks ([issue#20706](#), [pr#16481](#), Nikita Gerasimov)
- tools: ceph_disk/main.py: FreeBSD root has wheel for group ([pr#16609](#), Willem Jan Withagen)
- tools: ceph-disk: s/ceph_osd_mkfs/command_check_call/ ([issue#20685](#), [pr#16427](#), Zhu Shangzhong)
- tools: ceph.in: add help for locally-handled commands ([pr#13288](#), Dan Mick)
- tools: ceph.in: adjust usage width according to user’s tty ([pr#15190](#), Kefu Chai)
- tools: ceph.in: assert(state==connected) before help_for_target() ([pr#15156](#), Kefu Chai)
- tools: ceph.in: drop the compatibility to handle non json commands ([pr#15508](#), Kefu Chai)

- tools: ceph.in: filter out audit from ceph -w ([pr#16345](#), John Spray)
- tools: ceph.in, mgr: misc cleanups ([pr#16229](#), liuchang0812)
- tools: ceph.in: print return code when json_command failed ([pr#15378](#), liuchang0812)
- tools: ceph-objectstore-tool: Handle object names that are also valid json ([pr#12848](#), David Zafman)
- tools: ceph-release-notes: escape asterisks not for inline emphasis ([pr#16199](#), Kefu Chai)
- tools: ceph-release-notes: escape _ for unintended links ([issue#17499](#), [pr#16528](#), Kefu Chai)
- tools: ceph-release-notes: handle an edge case ([pr#16277](#), Nathan Cutler)
- tools: ceph-release-notes: ignore low-numbered PRs ([issue#18695](#), [pr#13151](#), Nathan Cutler)
- tools: ceph-release-notes: port it to py3 ([pr#16261](#), Kefu Chai)
- tools: ceph-release-notes: prefixes and pep8 compliance ([pr#14156](#), Nathan Cutler)
- tools: ceph-release-notes: refactor and fix regressions ([pr#16411](#), Nathan Cutler)
- tools: ceph-release-notes: strip trailing punctuation ([pr#14385](#), Nathan Cutler)
- tools: ceph-rest-api: be more tolerant on network failure ([issue#20115](#), [pr#15706](#), Kefu Chai)
- tools: ceph-volume: adds functional CI testing #16919 ([pr#16970](#), Andrew Schoen, Alfredo Deza)
- tools: ceph-volume docs ([pr#17124](#), Alfredo Deza)
- tools: ceph-volume: initial take on ceph-volume CLI tool ([pr#16632](#), Dan Mick, Alfredo Deza)
- tools: ceph-volume: Use a delimited CLI output parser instead of JSON ([pr#17123](#), Alfredo Deza)
- tools: change compare_exchange_weak to compare_exchange_strong ([pr#15030](#), Jesse Williamson)
- tools: Cleanup dead code in ceph-objectstore-tool ([pr#15812](#), David Zafman)
- tools: fio_ceph_objectstore: Print db_statistics when rocksdb_perf is enabled ([pr#15796](#), Xiaoyan Li)
- tools: init-ceph: print trailing n in "status" output ([pr#13351](#), Kefu Chai)
- tools: init-ceph: should have a space before "]" ([pr#14796](#), Kefu Chai)
- tools: os/bluestore/bluestore_tool: add sanity check to get rid of occasionally crash ([pr#16013](#), xie xingguo)
- tools: rados: check for negative return value of rados_create_with_context() as its comment put ([pr#10893](#), zhang.zezhu)
- tools: rados: fix typo in 'df' column name ([pr#15603](#), Ilya Dryomov)
- tools: rados: out json 'df' values as numbers, not strings ([issue#15546](#), [pr#14644](#), Sage Weil)
- tools: script: add docker core dump debugger ([pr#16375](#), Patrick Donnelly)
- tools: script: ceph-release-notes check orig. issue only for backports ([pr#12979](#), Abhishek Lekshmanan)
- tools: script/sephia_bt.sh: no need to pass version and sha1 anymore ([pr#13380](#), Kefu Chai)
- tools: src/ceph-disk/ceph_disk/main.py: Make 'ceph-disk list' work on FreeBSD ([pr#14483](#), Willem Jan Withagen)
- tools: stop.sh: boilerplate error (don't stop mon when stopping mgr) ([pr#14461](#), Dan Mick)
- tools: support hammer in rbd_recover_tool ([pr#12413](#), Bartłomiej Święcki)
- tools: tools/ceph_kvstore_tool: add "bluestore-kv" to usage ([pr#15326](#), xie xingguo)
- tools: tools/crushtool: replicated-rule API support ([pr#15011](#), xie xingguo)
- tools: tools/rados: add a parameter "-offset" to rados put command ([pr#12674](#), liuchang0812)
- tools: tools/rados: Check return value of connect ([issue#19319](#), [pr#14057](#), Brad Hubbard)
- tools: tools/rados: fixed typo in help information ([pr#15618](#), Pan Liu)
- tools: tools/rados: remove useless function declaration ([pr#12566](#), liuchang0812)
- tools: tools/rados: some cleanups ([pr#16147](#), Yan Jun)
- tools: vstart: "debug_ms=1" for mgr by default ([pr#15127](#), Kefu Chai)
- tools: vstart: don't configure rgw_dns_name ([pr#13411](#), Yehuda Sadeh)
- tools: vstart: don't create cluster by default ([pr#13891](#), Yehuda Sadeh)
- tools: vstart: print "start osd.\$id" instead of "start osd\$id" ([pr#15427](#), Kefu Chai)
- tools: vstart.sh: bind restful, dashboard to ::, not 127.0.0.1 ([pr#16349](#), Sage Weil)
- tools: vstart.sh: do not add host for mgr.* section if not \$overwrite_conf ([pr#13767](#), Kefu Chai)
- tools: warning, '%.16x' directive output truncated writing 16 bytes into a region of size 9 ([pr#14292](#), Jos Collin)
- tracing: don't include oid when tracing at dequeue_op() ([pr#13410](#), Yehuda Sadeh)

V11.2.1 KRAKEN

This is the first bugfix release for Kraken, and probably the last release of the Kraken series (Kraken will be declared "End Of Life" (EOL) when Luminous is declared stable). It contains a large number of bugfixes across all Ceph components.

We recommend that all v11.2.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- In previous versions, if a client sent an op to the wrong OSD, the OSD would reply with ENXIO. The rationale here is that the client or OSD is clearly buggy and we want to surface the error as clearly as possible. We now only send the ENXIO reply if the osd_enxio_on_misdirected_op option is enabled (it's off by default). This means that a VM using librbd that previously would have gotten an EIO and gone read-only will now see a blocked/hung IO instead.

- There was a bug introduced in Jewel (#19119) that broke the mapping behavior when an “out” OSD that still existed in the CRUSH map was removed with ‘osd rm’. This could result in ‘misdirected op’ and other errors. The bug is now fixed, but the fix itself introduces the same risk because the behavior may vary between clients and OSDs. To avoid problems, please ensure that all OSDs are removed from the CRUSH map before deleting them. That is, be sure to do:

```
ceph osd crush rm osd.123
```

before:

```
ceph osd rm osd.123
```

- This release greatly improves control and throttling of the snap trimmer. It introduces the “osd max trimming pgs” option (defaulting to 2), which limits how many PGs on an OSD can be trimming snapshots at a time. And it restores the safe use of the “osd snap trim sleep” option, which defaults to 0 but otherwise adds the given number of seconds in delay between every dispatch of trim operations to the underlying system.

OTHER NOTABLE CHANGES

- build/ops: ceph-base missing dependency for psmisc in Ubuntu Xenial ([issue#19129](#), [issue#19564](#), [pr#14425](#), Nathan Cutler)
- build/ops: logrotate is missing from debian package (kraken, master) ([issue#19670](#), [issue#19390](#), [pr#14734](#), Kefu Chai)
- build/ops: selinux: Do parallel relabel on package install ([issue#20077](#), [issue#20184](#), [issue#20191](#), [issue#20193](#), [pr#15509](#), Boris Ranto)
- build/ops: spec file mentions non-existent ceph-create-keys systemd unit file, causing ceph-mon units to not be enabled via preset ([issue#19460](#), [pr#14315](#), Sébastien Han)
- build/ops: systemd restarts Ceph Mon to quickly after failing to start ([issue#18635](#), [issue#18721](#), [pr#13185](#), Wido den Hollander)
- build/ops: systemd: Start OSDs after MONs ([issue#18907](#), [issue#18516](#), [pr#13494](#), Boris Ranto)
- ceph-disk: Add fix subcommand kraken back-port ([issue#19544](#), [pr#14345](#), Boris Ranto)
- ceph-disk: does not support cluster names different than ‘ceph’ ([issue#18973](#), [issue#17821](#), [pr#13497](#), Loic Dachary)
- ceph-disk: enable directory backed OSD at boot time ([issue#19628](#), [issue#19647](#), [pr#14604](#), Loic Dachary)
- ceph-disk: error on _bytes2str ([issue#18431](#), [issue#18371](#), [pr#13501](#), Kefu Chai)
- ceph-disk: fails if OSD udev rule triggers prior to mount of /var ([issue#20150](#), [issue#19941](#), [pr#16092](#), Loic Dachary)
- ceph-disk: Fix getting wrong group name when -setgroup in bluestore ([issue#18956](#), [pr#13488](#), craigchi)
- ceph-disk list reports mount error for OSD having mount options with SELinux context ([issue#19537](#), [issue#17331](#), [pr#14403](#), Brad Hubbard)
- ceph-disk prepare get wrong group name in bluestore ([issue#18997](#), [pr#13543](#), craigchi)
- ceph-disk: Racing between partition creation & device node creation ([issue#20034](#), [pr#16138](#), Erwan Velu)
- ceph-disk: separate ceph-osd -check-needs-* logs ([issue#20010](#), [issue#19888](#), [pr#16135](#), Loic Dachary)
- cephfs: buffer overflow in test LibCephFS.DirLs ([issue#18941](#), [issue#19045](#), [pr#14571](#), “Yan, Zheng”)
- cephfs: ceph-fuse crash during snapshot tests ([issue#18552](#), [issue#18460](#), [pr#14563](#), Yan, Zheng)
- cephfs: ceph-fuse does not recover after lost connection to MDS ([issue#19678](#), [issue#18757](#), [pr#16105](#), Henrik Korkuc)
- cephfs: client: fix the cross-quota rename boundary check conditions ([issue#18700](#), [pr#14567](#), Greg Farnum)
- cephfs: Deadlock on two ceph-fuse clients accessing the same file ([issue#20028](#), [issue#19635](#), [pr#16191](#), “Yan, Zheng”)
- cephfs: fragment space check can cause replayed request fail ([issue#18660](#), [issue#18706](#), [pr#14568](#), “Yan, Zheng”)
- cephfs: MDS crashes on missing metadata object ([issue#18179](#), [issue#18566](#), [pr#14565](#), Yan, Zheng)
- cephfs: MDS heartbeat timeout during rejoin, when working with large amount of caps/inodes ([issue#19118](#), [issue#19335](#), [pr#14572](#), John Spray)
- cephfs: mds is crushed, after I set about 400 64KB xattr kv pairs to a file ([issue#19674](#), [issue#19033](#), [pr#16103](#), Yang Honggang)
- cephfs: MDS server crashes due to inconsistent metadata ([issue#19406](#), [issue#19620](#), [pr#14574](#), John Spray)
- cephfs: mds/StrayManager: avoid reusing deleted inode in StrayManager::_purge_stray_logged ([issue#18950](#), [pr#14570](#), Zhi Zhang)
- cephfs: mount point break off problem after mds switch ([issue#19667](#), [issue#19437](#), [pr#16100](#), Guan yunfei, Sage Weil)
- cephfs: non-local quota changes not visible until some IO is done ([issue#17939](#), [issue#19763](#), [pr#16108](#), John Spray)
- cephfs: No output for ceph mds rmfailed 0 –yes-i-really-mean-it command ([issue#19483](#), [issue#16709](#), [pr#14573](#), John Spray)
- cephfs: normalize file open flags internally used by cephfs ([issue#19845](#), [pr#14998](#), Jan Fajerski)
- cephfs: segfault in handle_client_caps ([issue#18306](#), [issue#18616](#), [pr#14566](#), Yan, Zheng)
- cephfs: speed up readdir by skipping unwanted dn ([issue#18531](#), [pr#13028](#), Xiaoxi Chen)
- cephfs: src/test/pybind/test_cephfs.py fails ([issue#20500](#), [issue#19890](#), [pr#16114](#), “Yan, Zheng”)
- cephfs: test_client_recovery.TestClientRecovery fails ([issue#18562](#), [issue#18396](#), [pr#14564](#), Yan, Zheng)

- cephfs test failures (ceph.com/qa is broken, should be download.ceph.com/qa) ([issue#18574](#), [issue#18604](#), [pr#13024](#), John Spray)
- cephfs: Test failure: test_data_isolated (tasks.cephfs.test_volume_client.TestVolumeClient) ([issue#18914](#), [issue#19676](#), [pr#16104](#), "Yan, Zheng")
- cephfs: test_open_inode fails ([issue#18899](#), [issue#18661](#), [pr#14569](#), John Spray)
- client: populate metadata during mount ([issue#18361](#), [issue#18540](#), [pr#12951](#), John Spray)
- client: segfault on ceph_rmdir path / ([issue#18612](#), [issue#9935](#), [pr#13030](#), Michal Jarzabek)
- cls_rbd: default initialize snapshot namespace for legacy clients ([issue#19413](#), [issue#19833](#), [pr#14934](#), Jason Dillaman)
- cls/rgw: list_plain_entries() stops before bi_log entries ([issue#19876](#), [issue#20015](#), [pr#15384](#), Casey Bodley)
- common: monitor creation with IPv6 public network segfaults ([issue#19465](#), [issue#19371](#), [pr#14323](#), Fabian Grünbichler)
- common: possible lockdep false alarm for ThreadPool lock ([issue#18819](#), [issue#18894](#), [pr#13487](#), Mykola Golub)
- core: api_misc: [FAILED] LibRadosMiscConnectFailure.ConnectFailure ([issue#19561](#), [issue#15368](#), [pr#14733](#), Sage Weil)
- core: bluestore bdev: flush no-op optimization is racy ([issue#20495](#), [issue#19326](#), [issue#19327](#), [issue#19250](#), [issue#19251](#), [pr#14736](#), Sage Weil)
- core: improve control and throttling of the snap trimmer ([issue#19329](#), [issue#19931](#), [pr#14597](#), Samuel Just, Greg Farnum)
- core: two instances of omap_digest mismatch ([issue#19391](#), [pr#14200](#), Samuel Just, David Zafman)
- doc: PendingReleaseNotes: warning about 'osd rm ...' and #13733 ([issue#19119](#), [pr#14506](#), Sage Weil)
- doc: Python Swift client commands in Quick Developer Guide don't match configuration in vstart.sh ([issue#17746](#), [issue#18571](#), [pr#13044](#), Ronak Jain)
- doc: rgw: admin ops: fix the quota section ([issue#19397](#), [issue#19462](#), [pr#14521](#), Chu, Hua-Rong)
- fix: rgw crashed caused by shard id out of range when listing data log ([issue#20156](#), [issue#19732](#), [pr#16173](#), redickwang)
- fuse: TestVolumeClient.test_evict_client failure creating pidfile ([issue#18439](#), [issue#18309](#), [pr#12813](#), Nathan Cutler)
- librbd: allow to open an image without opening parent image ([issue#18609](#), [issue#18325](#), [pr#13132](#), Ricardo Dias)
- librbd: corrected resize RPC message backwards compatibility ([issue#19636](#), [issue#19659](#), [pr#14620](#), Jason Dillaman)
- librbd: Incomplete declaration for ContextWQ in librbd/Journal.h ([issue#18862](#), [issue#18892](#), [pr#14153](#), Boris Ranto)
- librbd: is_exclusive_lock_owner API should ping OSD ([issue#19467](#), [issue#19287](#), [pr#14480](#), Jason Dillaman)
- librbd: possible race in ExclusiveLock handle_peer_notification ([issue#19368](#), [pr#14163](#), Mykola Golub)
- librbd: prevent self-blacklisting during break lock ([issue#18703](#), [issue#18666](#), [pr#13201](#), Jason Dillaman)
- make check fails with Error EIO: load dlopen(build/lib/libec_FAKE.so): build/lib/libec_FAKE.so: cannot open shared object file: No such file or directory ([issue#20487](#), [issue#20345](#), [issue#18876](#), [pr#16069](#), Kefu Chai, Kyr Shatskyy)
- mds: assert fail when shutting down ([issue#19672](#), [issue#19204](#), [pr#16102](#), John Spray)
- mds: C_MDSInternalNoop::complete doesn't free itself ([issue#19664](#), [issue#19501](#), [pr#16099](#), "Yan, Zheng")
- mds: daemon goes readonly writing backtrace for a file whose data pool has been removed ([issue#19669](#), [issue#19401](#), [pr#16101](#), John Spray)
- mds: damage reporting by ino number is useless ([issue#18509](#), [issue#19680](#), [pr#16106](#), John Spray)
- mds: Decode errors on backtrace will crash MDS ([issue#18311](#), [issue#18463](#), [pr#12835](#), John Spray)
- mds: enable daemon to start when session ino info is corrupt ([issue#19710](#), [issue#16842](#), [pr#16107](#), John Spray)
- mds: failed filelock.can_read(-1) assertion in Server::_dir_is_nonempty ([issue#18707](#), [issue#18578](#), [pr#13555](#), Yan, Zheng)
- mds: finish clientreplay requests before requesting active state ([issue#18678](#), [issue#18461](#), [pr#13112](#), Yan, Zheng)
- mds: unresponsive when truncating a very large file ([issue#19755](#), [issue#20026](#), [pr#16190](#), "Yan, Zheng")
- mon: cache tiering: base pool last_force_resend not respected (racing read got wrong version) ([issue#18366](#), [issue#18403](#), [pr#13116](#), Sage Weil)
- mon crash on shutdown, lease_ack_timeout event ([issue#19928](#), [issue#19825](#), [pr#15084](#), Kefu Chai, Alexey Sheplyakov)
- mon: fail to form large quorum; msg/async busy loop ([issue#20230](#), [issue#20315](#), [pr#15729](#), Haomai Wang)
- mon: force_create_pg could leave pg stuck in creating state ([issue#19181](#), [issue#18298](#), [pr#13790](#), Adam C. Emerson, Sage Weil)
- mon/MonClient: make get_mon_log_message() atomic ([issue#19618](#), [issue#19427](#), [pr#14588](#), Kefu Chai)
- mon: 'osd crush move ...' doesnt work on osds ([issue#18682](#), [issue#18587](#), [pr#13500](#), Sage Weil)
- mon: osd crush set crushmap need sanity check ([issue#19302](#), [issue#20365](#), [pr#16143](#), Loic Dachary)
- mon: peon wrongly delete routed pg stats op before receive pg stats ack ([issue#18554](#), [issue#18458](#), [pr#13046](#), Mingxin Liu)
- mon/PGMap: factor mon_osd_full_ratio into MAX AVAIL calc ([issue#18522](#), [issue#20035](#), [pr#15237](#), Sage Weil)
- msg/simple/SimpleMessenger.cc: 239: FAILED assert(!cleared) ([issue#15784](#), [issue#18378](#), [pr#16133](#), Sage Weil)
- multisite: rest api fails to decode large period on 'period commit' ([issue#19505](#), [issue#19616](#), [issue#19614](#), [issue#20244](#), [issue#19488](#), [issue#19776](#), [issue#20293](#), [issue#19746](#), [pr#16161](#), Casey Bodley, Abhishek Lekshmanan)
- objecter: full_try behavior not consistent with osd ([issue#19560](#), [issue#19430](#), [pr#14732](#), Sage Weil)
- objecter: epoch_barrier isn't respected in _op_submit() ([issue#19396](#), [issue#19496](#), [pr#14331](#), Ilya Dryomov)
- os/bluestore: deep decode onode value ([issue#20366](#), [pr#15792](#), Sage Weil)
- os/bluestore: fix Allocator::allocate() int truncation ([issue#20884](#), [issue#18595](#), [pr#13011](#), Sage Weil)
- osd: allow client throttler to be adjusted on-fly, without restart ([issue#18791](#), [issue#18793](#), [pr#13216](#), Piotr Dałek)

- osd: An OSD was seen getting ENOSPC even with osdfailsafe_full_ratio passed ([issue#20544](#), [issue#16878](#), [issue#19340](#), [issue#19841](#), [issue#20672](#), [pr#16134](#), Sage Weil, David Zafman)
- osd: bogus assert when checking acting set on recovery completion in rados/upgrade ([issue#18999](#), [pr#13542](#), Sage Weil)
- osd: calc_clone_subsets misuses try_read_lock vs missing ([issue#18610](#), [issue#18583](#), [issue#18723](#), [issue#17831](#), [pr#14616](#), Samuel Just)
- osd: ceph degraded and misplaced status output inaccurate ([issue#18619](#), [issue#19480](#), [pr#14322](#), David Zafman)
- osd: condition object_info_t encoding on required (not up) features ([issue#18842](#), [issue#18831](#), [issue#18814](#), [pr#13485](#), Ilya Dryomov)
- osd: do not send ENXIO on misdirected op by default ([issue#19622](#), [pr#13253](#), Sage Weil)
- osd: FAILED assert(object_contexts.empty()) (live on master only from Jan-Feb 2017, all other instances are different) ([issue#20522](#), [issue#20523](#), [issue#18927](#), [issue#18809](#), [pr#16132](#), Samuel Just)
- osd: -flush-journal: sporadic segfaults on exit ([issue#18952](#), [issue#18820](#), [pr#13490](#), Alexey Sheplyakov)
- osd: Give requested scrubs a higher priority ([issue#19685](#), [issue#15789](#), [pr#14735](#), David Zafman)
- osd: Implement asynchronous scrub sleep ([issue#20033](#), [issue#19986](#), [issue#20173](#), [issue#19497](#), [pr#15526](#), Brad Hubbard)
- osd: leaked MOSDMap ([issue#19760](#), [issue#18293](#), [pr#14942](#), Sage Weil)
- osd: leveldb corruption leads to Operation not permitted not handled and assert ([issue#18037](#), [issue#18418](#), [pr#12790](#), Nathan Cutler)
- osd: metadata reports filestore when using bluestore ([issue#18677](#), [issue#18638](#), [pr#16083](#), Wido den Hollander)
- osd: New added OSD always down when full flag is set ([issue#19485](#), [pr#14321](#), Mingxin Liu)
- osd: Object level shard errors are tracked and used if no auth available ([issue#20089](#), [pr#15421](#), David Zafman)
- osd: os/bluestore: fix statfs to not include DB partition in free space ([issue#18599](#), [issue#18722](#), [pr#13284](#), Sage Weil)
- osd: osd/PrimaryLogPG: do not call on_shutdown() if (pg.deleting) ([issue#19902](#), [issue#19916](#), [pr#15066](#), Kefu Chai)
- osd: pg log split does not rebuild index for parent or child ([issue#19315](#), [issue#18975](#), [pr#14048](#), Sage Weil)
- osd: pglog: with config, don't assert in the presence of stale diverg... ([issue#17916](#), [issue#19702](#), [pr#14646](#), Greg Farnum)
- osd: publish PG stats when backfill-related states change ([issue#18497](#), [issue#18369](#), [pr#13295](#), Sage Weil)
- osd: Revert "PrimaryLogPG::failed_push: update missing as well" ([issue#18659](#), [pr#13091](#), David Zafman)
- osd: unlock sdata_op_ordering_lock with sdata_lock hold to avoid missing wakeup signal ([issue#20443](#), [pr#15962](#), Alexey Sheplyakov)
- pre-jewel "osd rm" incrementals are misinterpreted ([issue#19209](#), [issue#19119](#), [pr#13883](#), Ilya Dryomov)
- rbd: Add missing parameter feedback to 'rbd snap limit' ([issue#18601](#), [pr#14537](#), Tang Jin)
- rbd: [api] is_exclusive_lock_owner shouldn't return -EBUSY ([issue#20266](#), [issue#20182](#), [pr#16187](#), Jason Dillaman)
- rbd: [api] temporarily restrict (rbd_)mirror_peer_add from adding multiple peers ([issue#19256](#), [issue#19324](#), [pr#14545](#), Jason Dillaman)
- rbd: attempting to remove an image with incompatible features results in partial removal ([issue#18456](#), [issue#18315](#), [pr#13247](#), Dongsheng Yang)
- rbd: [cli] ensure positional arguments exist before casting ([issue#20264](#), [issue#20185](#), [pr#16186](#), Jason Dillaman)
- rbd: cli: map with cephx disabled results in error message ([issue#19035](#), [issue#20517](#), [pr#16298](#), Jason Dillaman)
- rbd: [FAILED] TestJournalTrimmer.RemoveObjectsWithOtherClient ([issue#18769](#), [issue#18738](#), [pr#14147](#), Jason Dillaman)
- rbd: Improve compatibility between librbd + krbd for the data pool ([issue#18771](#), [issue#18653](#), [pr#14539](#), Jason Dillaman)
- rbd: Issues with C API image metadata retrieval functions ([issue#19588](#), [issue#19611](#), [pr#15612](#), Mykola Golub)
- rbd: 'metadata_set' API operation should not change global config setting ([issue#18465](#), [issue#18549](#), [pr#14534](#), Mykola Golub)
- rbd-mirror: additional test stability improvements ([issue#18935](#), [issue#18947](#), [pr#14155](#), Jason Dillaman)
- rbd-mirror: deleting a snapshot during sync can result in read errors ([issue#19037](#), [issue#18990](#), [pr#14622](#), Jason Dillaman)
- rbd-mirror: ensure missing images are re-synced when detected ([issue#20022](#), [issue#19811](#), [pr#15486](#), Jason Dillaman)
- rbd-mirror: failover and fallback of unmodified image results in split-brain ([issue#19872](#), [issue#19858](#), [pr#14974](#), Jason Dillaman)
- rbd-mirror: potential race mirroring cloned image ([issue#18501](#), [issue#17993](#), [pr#14533](#), Jason Dillaman)
- rbd-mirror: sporadic image replayer shut down failure ([issue#18493](#), [issue#18441](#), [pr#14531](#), Jason Dillaman)
- rbd-nbd: add signal handler ([issue#19621](#), [issue#19349](#), [pr#16098](#), Kefu Chai, Pan Liu)
- rbd-nbd: check /sys/block/nbdX/size to ensure kernel mapped correctly ([issue#18970](#), [issue#17951](#), [issue#18910](#), [issue#18335](#), [pr#14540](#), Mykola Golub, Pan Liu)
- rbd: Possible deadlock performing a synchronous API action while refresh in-progress ([issue#18495](#), [issue#18419](#), [pr#14532](#), Jason Dillaman)
- rbd: Potential IO hang if image is flattened while read request is in-flight ([issue#19832](#), [issue#20154](#), [pr#16184](#), Jason Dillaman)
- rbd: [qa] crash in journal-enabled fsx run ([issue#18618](#), [issue#18632](#), [pr#14538](#), Jason Dillaman)
- rbd: qemu crash triggered by network issues ([issue#18776](#), [issue#18436](#), [pr#13245](#), Jason Dillaman)
- rbd: 'rbd bench-write' will crash if -io-size is 4G ([issue#18422](#), [issue#18557](#), [pr#14536](#), Gaurav Kumar Garg)

- rbd: rbd_clone_copy_on_read ineffective with exclusive-lock ([issue#19173](#), [issue#18888](#), [pr#14543](#), Venky Shankar)
- rbd: rbd -pool=x rename y z does not work ([issue#18777](#), [issue#18326](#), [pr#14149](#), Gaurav Kumar Garg)
- rbd: refuse to use an ec pool that doesn't support overwrites ([issue#19081](#), [issue#19336](#), [pr#16096](#), Jason Dillaman)
- rgw: add apis to support ragweed suite ([issue#19809](#), [pr#14852](#), Yehuda Sadeh)
- rgw: add the remove-x-delete feature to cancel swift object expiration ([issue#19472](#), [issue#19074](#), [pr#14522](#), Jing Wenjun)
- rgw: a few cases where rgw_obj is incorrectly initialized ([issue#19146](#), [issue#19096](#), [pr#13843](#), Yehuda Sadeh)
- rgw: anonymous user error code of getting object is not consistent with SWIFT ([issue#18806](#), [issue#19178](#), [pr#13877](#), Jing Wenjun)
- rgw: civetweb frontend segfaults in Luminous ([issue#19749](#), [issue#19840](#), [pr#16166](#), Abhishek Lekshmanan, Jesse Williamson)
- rgw: civetweb: move to post 1.8 version ([issue#19704](#), [pr#14960](#), Yehuda Sadeh)
- rgw: “cluster [WRN] bad locator @X on object @X....” in cluster log ([issue#19212](#), [issue#18980](#), [pr#14065](#), Casey Bodley)
- rgw: crash when updating period with placement group ([issue#18772](#), [issue#18631](#), [pr#14511](#), Orit Wasserman)
- rgw: Custom data header support ([issue#19843](#), [pr#15985](#), Pavan Rallabhandi)
- rgw: datalog trim can't work as expected ([issue#20263](#), [issue#20190](#), [pr#16175](#), Zhang Shaowen)
- rgw: DUMPABLE flag is cleared by setuid preventing core dumps ([issue#19147](#), [issue#19089](#), [pr#13845](#), Brad Hubbard)
- rgw: Error parsing xml when get bucket lifecycle ([issue#19363](#), [issue#19534](#), [pr#14528](#), liuchang0812)
- rgw: first write also tries to read object ([issue#18904](#), [issue#18622](#), [issue#18623](#), [issue#18621](#), [pr#14515](#), Yehuda Sadeh)
- rgw: fix break inside of yield in RGWFetchAllMetaCR ([issue#19322](#), [issue#17655](#), [pr#14067](#), Casey Bodley)
- rgw: fix handling RGWUserInfo::system in RGWHandler_REST_SWIFT ([issue#18476](#), [pr#13006](#), Radoslaw Zarzynski)
- rgw: fix RadosGW hang during multi-chunk upload of AWSv4 ([issue#19837](#), [issue#19754](#), [pr#14939](#), Radoslaw Zarzynski)
- rgw: fix use of marker in List::list_objects() ([issue#19047](#), [issue#18331](#), [pr#14517](#), Yehuda Sadeh)
- rgw: ‘gc list -include-all’ command infinite loop the first 1000 items ([issue#19978](#), [issue#20147](#), [pr#16139](#), Shasha Lu, fang yuxiang)
- rgw: get wrong content when download object with specific range when compression was enabled ([issue#20100](#), [issue#20268](#), [pr#16178](#), fang yuxiang)
- rgw: health check errors out incorrectly ([issue#19025](#), [issue#19157](#), [pr#13866](#), Pavan Rallabhandi)
- rgw: Lifecycle thread will still handle the bucket even if it has been removed ([issue#20285](#), [issue#20405](#), [pr#16183](#), Zhang Shaowen)
- rgw: make sending Content-Length in 204 and 304 controllable ([issue#18985](#), [issue#16602](#), [pr#13514](#), Radoslaw Zarzynski)
- rgw: meta sync thread crash at RGWMetaSyncShardCR ([issue#20251](#), [issue#20347](#), [pr#16180](#), Fang Yuxiang, Nathan Cutler)
- rgw: multisite: after CreateBucket is forwarded to master, local bucket may use different value for bucket index shards ([issue#19745](#), [issue#19759](#), [pr#16290](#), Shasha Lu)
- rgw: multisite: EPERM when trying to read SLO objects as system/admin user ([issue#19027](#), [issue#19475](#), [pr#14523](#), Casey Bodley)
- rgw: multisite: fetch_remote_obj() gets wrong version when copying from remote ([issue#19608](#), [pr#14606](#), Zhang Shaowen, Casey Bodley)
- rgw: multisite: RGWMetaSyncShardControlCR gives up on EIO ([issue#19160](#), [issue#19019](#), [pr#13868](#), Casey Bodley)
- rgw: multisite: segfault after changing value of rgw_data_log_num_shards ([issue#18488](#), [issue#18548](#), [pr#13181](#), Casey Bodley)
- rgw: multisite: some ‘radosgw-admin data sync’ commands hang ([issue#19236](#), [issue#19354](#), [pr#14142](#), Shasha Lu)
- rgw: multisite: some yields in RGWMetaSyncShardCR::full_sync() resume in incremental_sync() ([issue#19049](#), [issue#18076](#), [pr#13838](#), Casey Bodley)
- rgw: multisite: sync status reports master is on a different period ([issue#18709](#), [issue#18064](#), [pr#13176](#), Abhishek Lekshmanan)
- rgw: no http referer info in container metadata dump in swift API ([issue#18665](#), [issue#18898](#), [pr#13829](#), Jing Wenjun)
- rgw: “period update” does not remove short_zone_ids of deleted zones ([issue#15618](#), [issue#19342](#), [pr#14141](#), Casey Bodley)
- rgw: radosgw-admin: add the ‘object stat’ command to usage ([issue#19164](#), [issue#19013](#), [pr#13873](#), Pavan Rallabhandi)
- rgw: radosgw-admin period update reverts deleted zonegroup ([issue#18713](#), [issue#17239](#), [pr#13172](#), Orit Wasserman)
- rgw: ‘radosgw-admin usage show’ listing 0 bytes_sent/received ([issue#20261](#), [pr#16174](#), Pritha Srivastava)
- rgw: ‘radosgw-admin zone create’ command with specified zone-id creates a zone with different id ([issue#19524](#), [issue#19498](#), [pr#14526](#), Orit Wasserman)
- rgw: Realm set does not create a new period ([issue#18333](#), [issue#18499](#), [pr#14509](#), Orit Wasserman)
- rgw: reduce log level of ‘storing entry at’ in cls_log ([issue#19835](#), [issue#19839](#), [pr#16165](#), Willem Jan Withagen)
- rgw: Response header of swift API returned by radosgw does not contain x-openstack-request-id. But Swift returns it ([issue#19443](#), [issue#19573](#), [pr#14529](#), tone-zhang)
- rgw: rgw_file: fix marker computation ([issue#20158](#), [issue#19526](#), [issue#18989](#), [issue#19470](#), [issue#19471](#),

- issue#18651, issue#20195, issue#19059, issue#19112, issue#19018, issue#19036, issue#19154, issue#19170, issue#19663, issue#19661, issue#19111, issue#18992, issue#18650, issue#18991, issue#19623, issue#19149, issue#19270, issue#19723, issue#19625, issue#19624, issue#19060, issue#19166, issue#18810, issue#19168, issue#19162, issue#19066, issue#18808, issue#19634, issue#19435, issue#19144, issue#19229, issue#18902, pr#13871, Gui Hecheng, Matt Benjamin)
- rgw: S3 create bucket should not do response in json (issue#19172, issue#18889, pr#13875, Abhishek Lekshmanan)
- rgw: S3 v4 authentication issue with X-Amz-Expires (issue#19477, issue#18828, pr#14524, liuchang0812)
- rgw: S3 v4 authentication issue with X-Amz-Expires (issue#19725, issue#18828, pr#16162, liuchang0812)
- rgw: should parse the url to http host to compare with the container referer acl (issue#18896, issue#18685, pr#13780, Jing Wenjun)
- rgw: slave zonegroup cannot enable the bucket versioning (issue#18711, issue#18003, pr#13174, Orit Wasserman)
- rgw: Swift API: spurious newline after http body causes weird errors (issue#18780, issue#18473, pr#13224, Marcus Watts, Matt Benjamin)
- rgw: swift API: cannot disable object versioning with empty X-Versions-Location (issue#18852, issue#19175, pr#14519, Jing Wenjun)
- rgw: swift: disable revocation thread under certain circumstances (issue#19499, issue#9493, issue#19777, pr#16164, Marcus Watts)
- rgw: Swift's at-root features (/crossdomain.xml, /info, /healthcheck) are broken (issue#20031, issue#19520, pr#16168, Radoslaw Zarzynski)
- rgw: the swift container acl does not support field .ref (issue#18909, issue#19180, issue#18484, issue#18796, pr#14516, Jing Wenjun, Radoslaw Zarzynski)
- rgw: typo in rgw_admin.cc (issue#19156, issue#19026, pr#13864, Ronak Jain)
- rgw: unsafe access in RGWListBucket_ObjStore_SWIFT::send_response() (issue#19574, issue#19249, pr#14530, Yehuda Sadeh)
- rgw: upgrade to multisite v2 fails if there is a zone without zone info (issue#19331, issue#19231, pr#14137, Danny Al-Gaaf, Orit Wasserman)
- rgw: usage stats and quota are not operational for multi-tenant users (issue#18364, issue#18843, issue#16355, pr#14513, Radoslaw Zarzynski)
- rgw: Use decoded URI when verifying TempURL (issue#18590, issue#18627, pr#12986, Michal Koutný)
- rgw: VersionIdMarker and NextVersionIdMarker are not returned when listing object versions (issue#20363, issue#19886, pr#16181, Zhang Shaowen)
- rgw: when converting region_map we need to use rgw_zone_root_pool (issue#19195, issue#19356, pr#14144, Orit Wasserman)
- rgw: when uploading the objects continuously in the versioned bucket, some objects will not sync (issue#19766, issue#18208, pr#16163, lvshuhua)
- rgw: wrong object size after copy of uncompressed multipart objects (issue#20269, issue#20071, pr#16179, fang yuxiang)
- rgw: zonegroupmap set does not work (issue#18725, issue#19479, pr#14525, Casey Bodley)
- tests: AttributeError: Thrasher instance has no attribute 'ceph_objectstore_tool' (issue#19064, issue#18799, pr#13609, Nathan Cutler)
- tests: backport Sage's fixes to qa/suites/upgrade/jewel-x (issue#19651, pr#14612, Sage Weil)
- tests: ceph-object-corpus: kraken objects (issue#20878, pr#14983, Sage Weil)
- tests: CMakeLists.txt: disable memstore make check test (issue#17743, pr#16215, Sage Weil)
- tests: HEALTH_WARN pool rbd pg_num 244 > pgp_num 224 during upgrade (issue#19771, issue#20024, pr#16137, Kefu Chai)
- tests: ignore bogus ceph-objectstore-tool error in ceph_manager (issue#18805, issue#16263, pr#13239, Nathan Cutler, Kefu Chai)
- tests: insufficient timeout in radosbench task (issue#20497, pr#16111, Sage Weil)
- tests: LibRadosMiscConnectFailure.ConnectFailure hang (issue#20271, issue#19901, pr#16140, Sage Weil)
- tests: [librados_test_stub] cls_cxx_map_get_XYZ methods don't return correct value (issue#19597, issue#19609, pr#16097, Jason Dillaman)
- tests: move swift.py task from teuthology to ceph, phase one (kraken) (issue#20392, pr#15869, Nathan Cutler, Sage Weil, Warren Usui, Greg Farnum, Ali Maredia, Tommi Virtanen, Zack Cerza, Sam Lang, Yehuda Sadeh, Joe Buck, Josh Durgin)
- tests: ObjectStore/StoreTest.OnodeSizeTracking/2 fails on bluestore (issue#20499, pr#16112, xie xingguo)
- tests: qa: ceph-ansible test tweaks (issue#20882, pr#12984, pr#13618, Tamil Muthamizhan, Yuri Weinstein)
- tests: qa/suites/upgrade: add tiering test to hammer-jewel-x (issue#20879, issue#19185, pr#14692, Kefu Chai)
- tests: qa/tasks: misc systemd updates (issue#19719, pr#14702, Vasu Kulkarni)
- tests: qa/tasks: rbd-mirror daemon not properly run in foreground mode (issue#20638, issue#20630, issue#20634, pr#16342, Jason Dillaman)
- tests: qa/tasks: set pgp = pg num on thrashing finish (issue#20881, pr#13757, Kefu Chai)
- tests: qa/tasks/workunit: Backport repo fixes from master (issue#19429, issue#19531, pr#14487, Kefu Chai, Dan Mick)
- tests: remove hard-coded image name from TestLibRBD.Mirror (issue#18555, issue#19130, issue#19227, issue#18447, issue#19807, issue#19798, pr#16113, Mykola Golub, Jason Dillaman)
- tests: remove qa/suites/buildpackages (issue#18849, issue#18846, pr#13298, Loic Dachary)

- tests: run certain upgrade/jewel-x tests on Xenial only ([issue#20877](#), [pr#16493](#), Nathan Cutler)
- tests: run-rbd-unit-tests.sh assert in lockdep_will_lock, TestLibRBD.ObjectMapConsistentSnap ([issue#18822](#), [issue#17447](#), [pr#14151](#), Jason Dillaman)
- tests: SUSE yaml facets in qa/distros/all are out of date ([issue#18849](#), [issue#18870](#), [issue#18846](#), [issue#18856](#), [pr#13330](#), Nathan Cutler)
- tests: swift.py: clone the ceph-kraken branch ([issue#20520](#), [pr#16131](#), Nathan Cutler)
- tests: test/librbd: decouple ceph_test_librbd_api from libceph-common ([issue#20175](#), [issue#20351](#), [pr#16195](#), Kefu Chai)
- tests: test_notify.py: assert(not image.is_exclusive_lock_owner()) on line 147 ([issue#19716](#), [issue#19794](#), [pr#14833](#), Mykola Golub)
- tests: test_notify.py: rbd.InvalidArgument: error updating features for image test_notify_clone2 ([issue#19692](#), [issue#19693](#), [pr#14641](#), Jason Dillaman)
- tests: use ceph-kraken branch for s3tests ([issue#18387](#), [pr#12746](#), Nathan Cutler)
- tests: use librados API to retrieve config params ([issue#18668](#), [issue#18617](#), [pr#13102](#), Jason Dillaman)
- tests: various OpenStack tweaks ([issue#20882](#), [pr#13707](#), [pr#13641](#), [pr#13635](#), [pr#13633](#), [pr#13613](#), [pr#13283](#), [pr#13673](#), [pr#13638](#), [pr#14485](#), Zack Cerza)
- tools: ceph-brag fails to count “in” mds ([issue#19333](#), [issue#19192](#), [pr#14098](#), Peng Zhang)
- tools: ceph-disk prepare writes osd log 0 with root owner ([issue#18538](#), [issue#18606](#), [pr#13026](#), Samuel Matzek)
- tools: RadosImport::import should return an error if Rados::connect fails ([issue#19351](#), [issue#19319](#), [pr#14095](#), Brad Hubbard)

V11.2.0 KRAKEN

This is the first release of the Kraken series. It is a stable release that will be maintained with bugfixes and backports until the next stable release, Luminous, is completed in the Spring of 2017.

MAJOR CHANGES FROM JEWEL

- **RADOS:**
 - The new *BlueStore* backend now has a stable disk format and is passing our failure and stress testing. Although the backend is still flagged as experimental, we encourage users to try it out for non-production clusters and non-critical data sets.
 - RADOS now has experimental support for *overwrites on erasure-coded* pools. Because the disk format and implementation are not yet finalized, there is a special pool option that must be enabled to test the new feature. Enabling this option on a cluster will permanently bar that cluster from being upgraded to future versions.
 - We now default to the AsyncMessenger (ms type = async) instead of the legacy SimpleMessenger. The most noticeable difference is that we now use a fixed sized thread pool for network connections (instead of two threads per socket with SimpleMessenger).
 - Some OSD failures are now detected almost immediately, whereas previously the heartbeat timeout (which defaults to 20 seconds) had to expire. This prevents IO from blocking for an extended period for failures where the host remains up but the ceph-osd process is no longer running.
 - There is a new ceph-mgr daemon. It is currently collocated with the monitors by default, and is not yet used for much, but the basic infrastructure is now in place.
 - The size of encoded OSDMaps has been reduced.
 - The OSDs now quiesce scrubbing when recovery or rebalancing is in progress.
- **RGW:**
 - RGW now supports a new zone type that can be used for metadata indexing via ElasticSearch.
 - RGW now supports the S3 multipart object copy-part API.
 - It is possible now to reshards an existing bucket. Note that bucket resharding currently requires that all IO (especially writes) to the specific bucket is quiesced.
 - RGW now supports data compression for objects.
 - Civetweb version has been upgraded to 1.8
 - The Swift static website API is now supported (S3 support has been added previously).
 - S3 bucket lifecycle API has been added. Note that currently it only supports object expiration.
 - Support for custom search filters has been added to the LDAP auth implementation.
 - Support for NFS version 3 has been added to the RGW NFS gateway.
 - A Python binding has been created for librsgw.
- **RBD:**
 - RBD now supports images stored in an *erasure-coded* RADOS pool using the new (experimental) overwrite support. Images must be created using the new rbd CLI “-data-pool <ec pool>” option to specify the EC pool where the backing data objects are stored. Attempting to create an image directly on an EC pool will not be successful since the image’s backing metadata is only supported on a replicated pool.
 - The rbd-mirror daemon now supports replicating dynamic image feature updates and image metadata key/value

pairs from the primary image to the non-primary image.

- The number of image snapshots can be optionally restricted to a configurable maximum.
- The rbd Python API now supports asynchronous IO operations.

- **CephFS:**

- libcephfs function definitions have been changed to enable proper uid/gid control. The library version has been increased to reflect the interface change.
- Standby replay MDS daemons now consume less memory on workloads doing deletions.
- Scrub now repairs backtrace, and populates *damage ls* with discovered errors.
- A new *pg_files* subcommand to *cephfs-data-scan* can identify files affected by a damaged or lost RADOS PG.
- The false-positive “failing to respond to cache pressure” warnings have been fixed.

UPGRADING FROM KRAKEN RELEASE CANDIDATE 11.1.0

- The new *BlueStore* backend had an on-disk format change after 11.1.0. Any *BlueStore* OSDs created with 11.1.0 will need to be destroyed and recreated.

UPGRADING FROM JEWEL

- All clusters must first be upgraded to Jewel 10.2.z before upgrading to Kraken 11.2.z (or, eventually, Luminous 12.2.z).
- The *sortbitwise* flag must be set on the Jewel cluster before upgrading to Kraken. The latest Jewel (10.2.8+) releases issue a health warning if the flag is not set, so this is probably already set. If it is not, Kraken OSDs will refuse to start and will print an error message in their log.
- You may upgrade OSDs, Monitors, and MDSs in any order. RGW daemons should be upgraded last.
- When upgrading, new *ceph-mgr* daemon instances will be created automatically alongside any monitors. This will be true for Jewel to Kraken and Jewel to Luminous upgrades, but likely not be true for future upgrades beyond Luminous. You are, of course, free to create new *ceph-mgr* daemon instances and destroy the auto-created ones if you do not want them to be colocated with the *ceph-mon* daemons.

BLUESTORE

BlueStore is a new backend for managing data stored by each OSD on the directly hard disk or SSD. Unlike the existing *FileStore* implementation, which makes use of an XFS file system to store objects as files, *BlueStore* manages the underlying block device directly. Implements its own file system-like on-disk structure designed specifically for Ceph OSD workloads. Key features of *BlueStore* include:

- Checksums on all data written to disk, with checksum verifications on all reads, enabled by default.
- Inline compression support, which can be enabled on a per-pool or per-object basis via pool properties or client hints, respectively.
- Efficient journaling. Unlike *FileStore*, which writes *all* data to its journal device, *BlueStore* only journals metadata and (in some cases) small writes, reducing the size and throughput requirements for its journal. As with *FileStore*, the journal can be colocated on the same device as other data or allocated on a smaller, high-performance device (e.g., an SSD or NVMe device). *BlueStore* journals are only 512 MB by default.

The *BlueStore* on-disk format is expected to continue to evolve. However, we will provide support in the OSD to migrate to the new format on upgrade.

In order to enable *BlueStore*, add the following to *ceph.conf*:

```
enable experimental unrecoverable data corrupting features = bluestore
```

To create a *BlueStore* OSD, pass the *-bluestore* option to *ceph-disk* or *ceph-deploy* during OSD creation.

UPGRADE NOTES

- The OSDs now avoid starting new scrubs while recovery is in progress. To revert to the old behavior (and do not let recovery activity affect the scrub scheduling) you can set the following option:

```
osd scrub during recovery = true
```

- The list of monitor hosts/addresses for building the monmap can now be obtained from DNS SRV records. The service

name used when querying the DNS is defined in the “mon_dns_srv_name” config option, which defaults to “ceph-mon”.

- The ‘osd class load list’ config option is a list of object class names that the OSD is permitted to load (or '*' for all classes). By default it contains all existing in-tree classes for backwards compatibility.
- The ‘osd class default list’ config option is a list of object class names (or '*' for all classes) that clients may invoke having only the '*', 'x', 'class-read', or 'class-write' capabilities. By default it contains all existing in-tree classes for backwards compatibility. Invoking classes not listed in ‘osd class default list’ requires a capability naming the class (e.g. ‘allow class foo’).
- The ‘rgw rest getusage op compat’ config option allows you to dump (or not dump) the description of user stats in the S3 GetUsage API. This option defaults to false. If the value is true, the reponse data for GetUsage looks like:

```
"stats": {  
    "TotalBytes": 516,  
    "TotalBytesRounded": 1024,  
    "TotalEntries": 1  
}
```

If the value is false, the reponse for GetUsage looks as it did before:

```
{  
    516,  
    1024,  
    1  
}
```

- The ‘osd out ...’ and ‘osd in ...’ commands now preserve the OSD weight. That is, after marking an OSD out and then in, the weight will be the same as before (instead of being reset to 1.0). Previously the mons would only preserve the weight if the mon automatically marked an OSD out and then in, but not when an admin did so explicitly.
- The ‘ceph osd perf’ command will display ‘commit_latency(ms)’ and ‘apply_latency(ms)’. Previously, the names of these two columns are ‘fs_commit_latency(ms)’ and ‘fs_apply_latency(ms)’. We remove the prefix ‘fs_’, because they are not filestore specific.
- Monitors will no longer allow pools to be removed by default. The setting mon_allow_pool_delete has to be set to true (defaults to false) before they allow pools to be removed. This is a additional safeguard against pools being removed by accident.
- If you have manually specified the monitor user rocksdb via the mon keyvaluedb = rocksdb option, you will need to manually add a file to the mon data directory to preserve this option:

```
echo rocksdb > /var/lib/ceph/mon/ceph-`hostname`/kv_backend
```

New monitors will now use rocksdb by default, but if that file is not present, existing monitors will use leveldb. The mon keyvaluedb option now only affects the backend chosen when a monitor is created.

- The ‘osd crush initial weight’ option allows you to specify a CRUSH weight for a newly added OSD. Previously a value of 0 (the default) meant that we should use the size of the OSD’s store to weight the new OSD. Now, a value of 0 means it should have a weight of 0, and a negative value (the new default) means we should automatically weight the OSD based on its size. If your configuration file explicitly specifies a value of 0 for this option you will need to change it to a negative value (e.g., -1) to preserve the current behavior.
- The osd crush location config option is no longer supported. Please update your ceph.conf to use the crush location option instead.
- The static libraries are no longer included by the debian development packages (lib*-dev) as it is not required per debian packaging policy. The shared (.so) versions are packaged as before.
- The libtool pseudo-libraries (.la files) are no longer included by the debian development packages (lib*-dev) as they are not required per <https://wiki.debian.org/ReleaseGoals/LAFileRemoval> and <https://www.debian.org/doc/manuals/maint-guide/advanced.en.html>.
- The jerasure and shec plugins can now detect SIMD instruction at runtime and no longer need to be explicitly configured for different processors. The following plugins are now deprecated: jerasure_generic, jerasure_sse3, jerasure_sse4, jerasure_neon, shec_generic, shec_sse3, shec_sse4, and shec_neon. If you use any of these plugins directly you will see a

warning in the mon log file. Please switch to using just ‘jerasure’ or ‘shec’.

- The librados omap get_keys and get_vals operations include a start key and a limit on the number of keys to return. The OSD now imposes a configurable limit on the number of keys and number of total bytes it will respond with, which means that a librados user might get fewer keys than they asked for. This is necessary to prevent careless users from requesting an unreasonable amount of data from the cluster in a single operation. The new limits are configured with osd_max omap_entries_per_request, defaulting to 131,072, and osd_max omap_bytes_per_request, defaulting to 4MB.
- Calculation of recovery priorities has been updated. This could lead to unintuitive recovery prioritization during cluster upgrade. In case of such recovery, OSDs in the old version would operate on different priority ranges than new ones. Once upgraded, the cluster will operate on consistent values.

NOTABLE CHANGES

- bluestore: add counter to trace blob splitting ([pr#11718](#), xie xingguo)
- bluestore: a few more cleanups ([pr#11780](#), xie xingguo)
- bluestore: avoid polluting shard info if need resharding ([pr#11439](#), xie xingguo)
- bluestore: avoid unnecessary call to init_csum() ([pr#12015](#), xie xingguo)
- bluestore: ceph-disk: adjust bluestore default device sizes ([pr#12530](#), Sage Weil)
- bluestore: ceph_test_objectstore: smaller device ([pr#11591](#), Sage Weil)
- bluestore: clean up Allocator::dump ([issue#18054](#), [pr#12282](#), Sage Weil)
- bluestore: clear extent map on object removal ([pr#11603](#), Sage Weil)
- bluestore: compressor/ZLibCompressor: fix broken isal-l ([pr#11445](#), Igor Fedotov)
- bluestore: dedup if space overlap truly exists ([pr#11986](#), xie xingguo)
- bluestore: dedup omap_head, reuse nid instead ([pr#12275](#), xie xingguo)
- bluestore: deep fsck ([pr#11724](#), Sage Weil)
- bluestore: default bluestore_clone_cow=true ([pr#11540](#), Sage Weil)
- bluestore: drop inline_dirty from struct ExtentMap ([pr#11377](#), xie xingguo)
- bluestore: drop member “space” from Onode ([pr#12185](#), xie xingguo)
- bluestore: fix alloc release timing on sync submits ([pr#11983](#), Sage Weil)
- bluestore: fix bufferspace stats leak due to blob splitting ([pr#12039](#), xie xingguo)
- bluestore: fix collection_list end bound off-by-one ([pr#11771](#), Sage Weil)
- bluestore: fix compiler warnings ([pr#11905](#), xie xingguo)
- bluestore: fixes and cleanups ([pr#11761](#), xie xingguo)
- bluestore: fix escaping of chars > 0x80 ([pr#11502](#), Sage Weil)
- bluestore: fix extent shard span check ([pr#11725](#), Sage Weil)
- bluestore: fix has_aios ([pr#11317](#), Sage Weil)
- bluestore: Fix invalid compression statfs caused by clone op ([pr#11351](#), Igor Fedotov)
- bluestore: fix lack of resharding ([pr#11597](#), Igor Fedotov)
- bluestore: fix latency calculation ([pr#12040](#), Pan Liu)
- bluestore: fix onode vs extent key suffix ([pr#11452](#), Sage Weil)
- bluestore: fix potential memory leak ([pr#11893](#), xie xingguo)
- bluestore: fix race condition during blob splitting ([pr#11422](#), xiexingguo, xie xingguo)
- bluestore: fix remove_collection to properly detect collection e... ([pr#11398](#), Igor Fedotov)
- bluestore: fix _split_collections race with osr_reap ([pr#11748](#), Sage Weil)
- bluestore: fix up compression tests and debug output ([pr#11350](#), Sage Weil)
- bluestore: fix writes that span existing shard boundaries ([pr#11451](#), Sage Weil)
- bluestore: flush before enumerating omap values ([issue#18140](#), [pr#12328](#), Sage Weil)
- bluestore: formatting nits ([pr#11514](#), xie xingguo)
- bluestore: fsck: fix omap_head check ([pr#11726](#), Sage Weil)
- bluestore: GC infra refactor, more UTs and GC range calculation fixes ([pr#11482](#), Igor Fedotov)
- bluestore: KernelDevice: fix race in aio_thread vs aio_wait ([issue#17824](#), [pr#12204](#), Sage Weil)
- bluestore: kv: dump rocksdb stats ([pr#12287](#), Varada Kari, Jianpeng Ma, Sage Weil)
- bluestore: kv/rocksdb: enable rocksdb write path breakdown ([pr#11696](#), Haodong Tang)
- bluestore: kv/RocksDBStore: rename option ([pr#11769](#), Sage Weil)
- bluestore: less code redundancy ([pr#11740](#), xie xingguo)
- bluestore: make 2q cache kin/kout size tunable ([pr#11599](#), Haodong Tang)
- bluestore: mark ops that can’t tolerate ENOENT ([pr#12114](#), Sage Weil)
- bluestore: mempool: changes for bitmap allocator ([pr#11922](#), Ramesh Chander)
- bluestore: misc. fixes and cleanups ([pr#11964](#), xie xingguo)
- bluestore: move bluefs into its own mempool ([pr#11834](#), Sage Weil)
- bluestore: no garbage collection for uncompressed blobs ([pr#11539](#), Roushan Ali, Sage Weil)
- bluestore: optional debug mode to identify aio stalls ([pr#11818](#), Sage Weil)
- bluestore: os/bluestore: a few cleanups ([pr#11483](#), Sage Weil)

- bluestore: os/bluestore: avoid resharding if the last shard size fall below shar... ([pr#12447](#), Igor Fedotov)
- bluestore: os/bluestore: bitmap allocator dump functionality ([pr#12298](#), Ramesh Chander)
- bluestore: os/bluestore: bluestore_sync_submit_transaction = false ([pr#12367](#), Sage Weil)
- bluestore: os/bluestore: cleanup around Blob::ref_map ([pr#11896](#), Igor Fedotov)
- bluestore: os/bluestore: clear omap flag if parent has none ([pr#12351](#), xie xingguo)
- bluestore: os/bluestore: don't implicitly create the source object for clone ([pr#12353](#), xie xingguo)
- bluestore: os/bluestore: drop old bluestore preconditioning; replace with wal preextension of file size ([pr#12265](#), Sage Weil)
- bluestore: os/bluestore: fix global commit latency ([pr#12356](#), xie xingguo)
- bluestore: os/bluestore: fix ondisk encoding for blobs ([pr#12488](#), Varada Kari, Sage Weil)
- bluestore: os/bluestore: fix potential csum_order overflow ([pr#12333](#), xie xingguo)
- bluestore: os/bluestore: fix target_buffer value overflow in Cache::trim() ([pr#12507](#), Igor Fedotov)
- bluestore: os/bluestore: include modified objects in flush list even if onode unchanged ([pr#12541](#), Sage Weil)
- bluestore: os/bluestore: kill dead gc-related counters ([pr#12065](#), xie xingguo)
- bluestore: os/bluestore: kill overlay related options ([pr#11557](#), xie xingguo)
- bluestore: os/bluestore: misc coverity fixes/cleanups ([pr#12202](#), Sage Weil)
- bluestore: os/bluestore: preserve source collection cache during split ([pr#12574](#), Sage Weil)
- bluestore: os/bluestore: remove 'extents' from shard_info ([pr#12629](#), Sage Weil)
- bluestore: os/bluestore: simplified allocator interfaces to single apis ([pr#12355](#), Ramesh Chander)
- bluestore: os/bluestore: simplify allocator release flow ([pr#12343](#), Sage Weil)
- bluestore: os/bluestore: simplify can_split_at() ([pr#11607](#), xie xingguo)
- bluestore: os/bluestore: use iterator for erase() method directly ([pr#11490](#), xie xingguo)
- bluestore: os/kstore: rmcoll fix to satisfy store_test ([pr#11533](#), Igor Fedotov)
- bluestore: os: make filestore_blackhole -> objectstore_blackhole ([pr#11788](#), Sage Weil)
- bluestore: os: move_ranges_destroy_src ([pr#11237](#), Manali Kulkarni, Sage Weil)
- bluestore: readability improvements and doxygen fix ([pr#11895](#), xie xingguo)
- bluestore: reap collection after all pending ios done ([pr#11797](#), Haomai Wang)
- bluestore: reap ioc when stopping aio_thread. ([pr#11811](#), Haodong Tang)
- bluestore: refactor _do_write(); move initializaiton of csum out of loop ([pr#11823](#), xie xingguo)
- bluestore: remove duplicated namespace of tx state ([pr#11845](#), xie xingguo)
- bluestore: remove garbage collector staff ([pr#12042](#), Igor Fedotov)
- bluestore: set next object as gobject_t::get_max() when start.hobj.i... ([pr#11495](#), Xinze Chi, Haomai Wang)
- bluestore: simplify blob status checking for small writes ([pr#11366](#), xie xingguo)
- bluestore: some more cleanups ([pr#11910](#), xie xingguo)
- bluestore: spdk: a few fixes ([pr#11882](#), Yehuda Sadeh)
- bluestore: speed up omap-key generation for same onode ([pr#11807](#), xie xingguo)
- bluestore: traverse buffer_map in reverse order when spliting BufferSpace ([pr#11468](#), xie xingguo)
- bluestore: update cache logger after 'trim_cache' operation ([pr#11695](#), Haodong Tang)
- bluestore: use bitmap allocator for bluefs ([pr#12285](#), Sage Weil)
- bluestore: use std::unordered_map for SharedBlob lookup ([pr#11394](#), Sage Weil)
- build/ops: AArch64: Detect crc32 extension support from assembler ([issue#17516](#), [pr#11391](#), Alexander Graf)
- build/ops: boost: embedded ([pr#11817](#), Sage Weil, Matt Benjamin)
- build/ops: build: dump env during build ([issue#18084](#), [pr#12284](#), Sage Weil)
- build/ops: ceph-detect-init: FreeBSD introduction of bsdrc ([pr#11906](#), Willem Jan Withagen, Kefu Chai)
- build/ops: ceph-disk: enable -runtime ceph-osd systemd units ([issue#17889](#), [pr#12241](#), Loic Dachary)
- build/ops: ceph.spec: add pybind rgwfile ([pr#11847](#), Haomai Wang)
- build/ops,cleanup,bluestore: os/bluestore: remove build warning in a better way ([pr#11920](#), Igor Fedotov)
- build/ops: CMakeLists: add vstart-base target ([pr#12476](#), Sage Weil)
- build/ops: CMakeLists.txt: enable LTTNG by default ([pr#11500](#), Sage Weil)
- build/ops: common/buffer.cc: raw_pipe depends on splice(2) ([pr#11967](#), Willem Jan Withagen)
- build/ops,common: common/str_list.h: fix clang warning about std::move ([pr#12570](#), Willem Jan Withagen)
- build/ops,core: xio: fix build ([pr#11768](#), Matt Benjamin)
- build/ops: deb: add python dependencies where needed ([issue#17579](#), [pr#11507](#), Nathan Cutler, Kefu Chai)
- build/ops: deb: add python-rgw packages ([pr#11832](#), Sage Weil)
- build/ops: debian: apply dh_python to python-rgw also ([pr#12260](#), Kefu Chai)
- build/ops: deb: update python-rgw dependencies to librgw2 ([pr#11885](#), Casey Bodley)
- build/ops: do_freebsd.sh: Build with SYSTEM Boost on FreeBSD ([pr#11942](#), Willem Jan Withagen)
- build/ops: do_freebsd.sh: Do not use LTTNG on FreeBSD ([pr#11551](#), Willem Jan Withagen)
- build/ops: do_freebsd.sh: Set options for debug building. ([pr#11443](#), Willem Jan Withagen)
- build/ops: FreeBSD: do_freebsd.sh ([pr#12090](#), Willem Jan Withagen)
- build/ops: FreeBSD:test/encoding/readable.sh": fix nproc and ls -v calls ([pr#11522](#), Willem Jan Withagen)
- build/ops: FreeBSD: update require packages ([pr#11512](#), Willem Jan Withagen)
- build/ops: git-archive-all.sh: use an actually unique tmp dir ([pr#12011](#), Dan Mick)
- build/ops: include/enc: make clang happy ([pr#11638](#), Kefu Chai, Sage Weil)
- build/ops: install-deps.sh: allow building on SLES systems ([pr#11708](#), Nitin A Kamble)

- build/ops: install-deps.sh: JQ is needed in one script ([pr#12080](#), Willem Jan Withagen)
- build/ops: Log: Replace namespace log with logging ([pr#11650](#), Willem Jan Withagen)
- build/ops: Merging before make check because it clearly breaks the build and the build part is done ([pr#11924](#), Sage Weil)
- build/ops: ok, w/upstream acks, merging-jenkins build did succeed (this is a build-only change) ([pr#12008](#), Matt Benjamin)
- build/ops: qa: Add ceph-ansible installer. ([issue#16770](#), [pr#10402](#), Warren Usui)
- build/ops: rocksdb: do not build with -march=native ([pr#11677](#), Kefu Chai)
- build/ops: rocksdb: update to latest ([pr#12100](#), Kefu Chai)
- build/ops: rpm: Remove trailing whitespace in usermod command (SUSE) ([pr#10707](#), Tim Serong)
- build/ops: scripts/release-notes: allow title guesses from gh tags & description update ([pr#11399](#), Abhishek Lekshmanan)
- build/ops: systemd: Fix startup of ceph-mgr on Debian 8 ([pr#12555](#), Mark Korenberg)
- build/ops: tracing/objectstore_tp: add missing move_ranges_... tp ([pr#11484](#), Sage Weil)
- build/ops: upstart: fix ceph-crush-location default ([issue#6698](#), [pr#803](#), Jason Dillaman)
- build/ops: upstart: start ceph-all after static-network-up ([issue#17689](#), [pr#11631](#), Billy Olsen)
- cephfs: add gid to asok status ([pr#11487](#), Patrick Donnelly)
- cephfs: API cleanup for libcephfs interfaces ([issue#17911](#), [pr#12106](#), Jeff Layton)
- cephfs: ceph-fuse: start up log on parent process before shutdown ([issue#18157](#), [pr#12347](#), Greg Farnum)
- cephfs: ceph_fuse: use sizeof get the buf length ([pr#11176](#), LeoZhang)
- cephfs: cleanup: ceph-fuse: start up log on parent process before shutdown ([issue#18157](#), [pr#12358](#), Kefu Chai)
- cephfs: client: add pid to metadata ([issue#17276](#), [pr#11359](#), Patrick Donnelly)
- cephfs: client: Client.cc: remove duplicated op type checking against CEPH_MD... ([pr#11608](#), Weibing Zhang)
- cephfs: client: don't take extra target inode reference in ll_link ([pr#11440](#), Jeff Layton)
- cephfs: client: fix mutex name typos ([pr#12401](#), Yunchuan Wen)
- cephfs: client: get caller's uid/gid on every libcephfs operation ([issue#17591](#), [pr#11526](#), Yan, Zheng)
- cephfs: client: get gid from MonClient ([pr#11486](#), Patrick Donnelly)
- cephfs: client: improve failure messages/debugging ([pr#12110](#), Patrick Donnelly)
- cephfs: client/mds: Clear setuid bits when writing or truncating ([issue#18131](#), [pr#12412](#), Jeff Layton)
- cephfs: client: put CapSnap not ptr in cap_snaps map ([pr#12111](#), Patrick Donnelly)
- cephfs: client: remove redundant initialization ([pr#12028](#), Patrick Donnelly)
- cephfs: client: remove unnecessary bufferptr[] for writev ([pr#11836](#), Patrick Donnelly)
- cephfs: client: remove unneeded layout on MClientCaps ([pr#11790](#), John Spray)
- cephfs: client: set metadata["root"] from mount method when it's called with ... ([pr#12505](#), Jeff Layton)
- cephfs: client: trim_caps() do not dereference cap if it's removed ([pr#12145](#), Kefu Chai)
- cephfs: client: use unique_ptr ([pr#11837](#), Patrick Donnelly)
- cephfs: common/ceph_string: add ceph string constants for CEPH_SESSION_FORCE_RO ([pr#11516](#), Zhi Zhang)
- cephfs: Fix #17562 (backtrace check fails when scrubbing directory created by fsstress) ([issue#17562](#), [pr#11517](#), Yan, Zheng)
- cephfs: fix missing ll_get for ll_walk ([issue#18086](#), [pr#12061](#), Gui Hecheng)
- cephfs: get new fsmap after marking clusters down ([issue#7271](#), [issue#17894](#), [pr#1262](#), Patrick Donnelly)
- cephfs: Have ceph clear setuid/setgid bits on chown ([issue#18131](#), [pr#12331](#), Jeff Layton)
- cephfs: libcephfs: add ceph_fsetattr&&ceph_lchmod&&ceph_lutime ([pr#11191](#), huanwen ren)
- cephfs: libcephfs: add readlink function in cephfs.pyx ([pr#12384](#), huanwen ren)
- cephfs: libcephfs and test suite fixes ([issue#18013](#), [issue#17982](#), [pr#12228](#), Jeff Layton)
- cephfs: libcephfs client API overhaul and update ([pr#11647](#), Jeff Layton)
- cephfs: lua: use simpler lua_next traversal structure ([pr#11958](#), Patrick Donnelly)
- cephfs: mds/Beacon: move C_MDS_BeaconSender class to .cc ([pr#10940](#), Michal Jarzabek)
- cephfs: mds/CDir.cc: remove unneeded use of count ([pr#11613](#), Michal Jarzabek)
- cephfs: mds/CInode.h: remove unneeded use of count ([pr#11371](#), Michal Jarzabek)
- cephfs: mds/DamageTable.cc: move shared ptrs ([pr#11435](#), Michal Jarzabek)
- cephfs: mds/DamageTable.cc: remove unneeded use of count ([pr#11625](#), Michal Jarzabek)
- cephfs: mds/DamageTable: move classes to .cc file ([pr#11450](#), Michal Jarzabek)
- cephfs: mds/flock: add const to member functions ([pr#11692](#), Michal Jarzabek)
- cephfs: mds/FSMap.cc: remove unneeded use of count ([pr#11402](#), Michal Jarzabek)
- cephfs: mds/FSMapUser.h: remove copy ctr and assign op ([pr#11509](#), Michal Jarzabek)
- cephfs: mds/InfoTable.h: add override to virtual functs ([pr#11496](#), Michal Jarzabek)
- cephfs: mds/InoTable.h: add override to virtual functs ([pr#11604](#), Michal Jarzabek)
- cephfs: mds/Mantle.h: include correct header files ([pr#11886](#), Michal Jarzabek)
- cephfs: mds/Mantle: pass parameters by const ref ([pr#11713](#), Michal Jarzabek)
- cephfs: mds/MDCache.h: remove unneeded call to clear func ([pr#11954](#), Michal Jarzabek)
- cephfs: mds/MDCache.h: remove unused functions ([pr#11908](#), Michal Jarzabek)
- cephfs: mds/MDLog: add const to member functions ([pr#11663](#), Michal Jarzabek)
- cephfs: mds/MDSMap.h: add const to member functions ([pr#11511](#), Michal Jarzabek)
- cephfs: mds/MDSRank: add const to member functions ([pr#11752](#), Michal Jarzabek)
- cephfs: mds/MDSRank.h: add override to virtual function ([pr#11727](#), Michal Jarzabek)

- cephfs: mds/MDSRank.h: make destructor protected ([pr#11651](#), Michal Jarzabek)
- cephfs: mds/MDSTableClient.h: add const to member funct ([pr#11681](#), Michal Jarzabek)
- cephfs: mds/Migrator.cc: remove unneeded use of count ([pr#11523](#), Michal Jarzabek)
- cephfs: mds/Migrator.h: add const to member functions ([pr#11819](#), Michal Jarzabek)
- cephfs: mds/Migrator.h: remove unneeded use of count ([pr#11833](#), Michal Jarzabek)
- cephfs: mds/Mutation.h: add const to member functions ([pr#11670](#), Michal Jarzabek)
- cephfs: mds/Mutation.h: simplify constructors ([pr#11455](#), Michal Jarzabek)
- cephfs: MDS: reduce usage of context wrapper ([pr#11560](#), Yan, Zheng)
- cephfs: mds/ScrubHeader.h: pass string by const reference ([pr#11904](#), Michal Jarzabek)
- cephfs: mds/server: merge the snapshot request judgment ([pr#11150](#), huanwen ren)
- cephfs: mds/SessionMap: add const to member functions ([pr#11541](#), Michal Jarzabek)
- cephfs: mds/SessionMap.cc: avoid copying and add const ([pr#11297](#), Michal Jarzabek)
- cephfs: mds/SessionMap.cc:put classes in unnamed namespace ([pr#11316](#), Michal Jarzabek)
- cephfs: mds/SessionMap.cc: remove unneeded use of count ([pr#11338](#), Michal Jarzabek)
- cephfs: mds/SessionMap.h: remove unneeded function ([pr#11565](#), Michal Jarzabek)
- cephfs: mds/SessionMap.h: remove unneeded use of count ([pr#11358](#), Michal Jarzabek)
- cephfs: mds/SnapRealm: remove unneeded use of count ([pr#11609](#), Michal Jarzabek)
- cephfs: mds/SnapServer.h: add override to virtual functs ([pr#11380](#), Michal Jarzabek)
- cephfs: mds/SnapServer.h: add override to virtual functs ([pr#11583](#), Michal Jarzabek)
- cephfs: mon/MDSMonitor: fix iterating over mutated map ([issue#18166](#), [pr#12395](#), John Spray)
- cephfs: multimds: fix state check in Migrator::find_stale_export_freeze() ([pr#12098](#), Yan, Zheng)
- cephfs: osdc: After write try merge bh. ([issue#17270](#), [pr#11545](#), Jianpeng Ma)
- cephfs: Partial organization of mds/ header sections ([pr#11959](#), Patrick Donnelly)
- cephfs: Port/bootstrap ([pr#827](#), Yan, Zheng)
- cephfs: Revert “osdc: After write try merge bh.” ([issue#17270](#), [pr#11262](#), John Spray)
- cephfs: Small pile of random cephfs fixes and cleanup ([pr#11421](#), Jeff Layton)
- cephfs: src/mds: fix MDSMap upgrade decoding ([issue#17837](#), [pr#12097](#), John Spray)
- cephfs: systemd: add ceph-fuse service file ([pr#11542](#), Patrick Donnelly)
- cephfs: test fragment size limit ([issue#16164](#), [pr#1069](#), Patrick Donnelly)
- cephfs: test readahead is working ([issue#16024](#), [pr#1046](#), Patrick Donnelly)
- cephfs: test: temporarily remove fork(ing flock tests ([issue#16556](#), [pr#11211](#), John Spray)
- cephfs: tool/cephfs: displaying “list” in journal event mode ([pr#11236](#), huanwen ren)
- cephfs: tools/cephfs: add pg_files command ([issue#17249](#), [pr#11026](#), John Spray)
- cephfs: tools/cephfs: add scan_links command which fixes linkages errors ([pr#11446](#), Yan, Zheng)
- cephfs: update tests to enable multimds when needed ([pr#933](#), Greg Farnum)
- cleanup: build: The Light Clangtastic ([pr#11921](#), Adam C. Emerson)
- cleanup,common: common/blkdev: use realpath instead of readlink to resolve the recurs... ([pr#12462](#), Xinze Chi)
- cleanup,common: common/throttle: simplify Throttle::_wait() ([pr#11165](#), xie xingguo)
- cleanup,common: src/common: remove nonused config option ([pr#12311](#), Wei Jin)
- cleanup: coverity fix: fixing few coverity issue ([pr#9624](#), Gaurav Kumar Garg)
- cleanup: deprecate readdir_r() with readdir() ([pr#11805](#), Kefu Chai)
- cleanup: erasure-code: fix gf-complete warning ([pr#12150](#), Kefu Chai)
- cleanup: fix typos ([pr#12502](#), xianxiaoxiao)
- cleanup: mds/FSMap.cc: prevent unneeded copy of map entry ([pr#11798](#), Michal Jarzabek)
- cleanup: mds/FSMap.h: add const and reference ([pr#11802](#), Michal Jarzabek)
- cleanup: mds/FSMap: pass shared_ptr by const ref ([pr#11383](#), Michal Jarzabek)
- cleanup: mds/SnapServer: add const to member function ([pr#11688](#), Michal Jarzabek)
- cleanup: mon/MonCap.h: add std::move for std::string ([pr#10722](#), Michal Jarzabek)
- cleanup: mon/OSDMonitor: only show interesting flags in health warning ([issue#18175](#), [pr#12365](#), Sage Weil)
- cleanup: msg/async: assert(0) -> ceph_abort() ([pr#12339](#), Li Wang)
- cleanup: msg/AsyncMessenger: remove unneeded include ([pr#9846](#), Michal Jarzabek)
- cleanup: msg/async/rdma: fix disconnect log line ([pr#12254](#), Adir Lev)
- cleanup: msg/async: remove unused member variable ([pr#12387](#), Kefu Chai)
- cleanup: msg: fix format specifier for unsigned value id ([pr#11145](#), Weibing Zhang)
- cleanup: msg/Pipe: move DelayedDelivery class to cc file ([pr#10447](#), Michal Jarzabek)
- cleanup: msg/test: fix the guided compile-command to ceph_test_msgr ([pr#10490](#), Yan Jun)
- cleanup: osd/PGBackend: build_push_op segment fault ([pr#9357](#), Zengran Zhang)
- cleanup: osd/PG.h: change PGRecoveryStats struct to class ([pr#11178](#), Michal Jarzabek)
- cleanup: osd/PG.h: remove unneeded forward declaration ([pr#12135](#), Li Wang)
- cleanup: osd/ReplicatedPG: remove unneeded use of count ([pr#11251](#), Michal Jarzabek)
- cleanup: os/filestore: clean filestore perfcounters ([pr#11524](#), Wei Jin)
- cleanup: os/fs/FS.cc: condition on WITH_AIO for FreeBSD ([pr#11913](#), Willem Jan Withagen)
- cleanup,rbd: cls_rbd: silence compiler warnings ([pr#11363](#), xiexingguo)
- cleanup,rbd: journal: avoid logging an error when a watch is blacklisted ([issue#18243](#), [pr#12473](#), Jason Dillaman)
- cleanup,rbd: journal: prevent repetitive error messages after being blacklisted ([issue#18243](#), [pr#12497](#), Jason Dillaman)

- cleanup,rbd: librbd/ImageCtx: no need for virtual dtor ([pr#12220](#), Sage Weil)
- cleanup,rbd: rbd-mirror: configuration overrides for hard coded timers ([pr#11840](#), Dongsheng Yang)
- cleanup,rbd: rbd-mirror: set SEQUENTIAL and NOCACHE advise flags on image sync ([issue#17127](#), [pr#12280](#), Mykola Golub)
- cleanup: remove unneeded forward declaration ([pr#12257](#), Li Wang, Yunchuan Wen)
- cleanup: remove unused declaration ([pr#12466](#), Li Wang, Yunchuan Wen)
- cleanup,rgw: rgw multisite: move lease up to RunBucketSync instead of child crs ([pr#11598](#), Casey Bodley)
- cleanup,rgw: rgw/rest: don't print empty x-amz-request-id ([pr#10674](#), Marcus Watts)
- cleanup,rgw: verified: f23 ([pr#12103](#), Radoslaw Zarzynski)
- cleanup: src/common/perf_counters.h: fix wrong word ([pr#11690](#), zhang.zezhu)
- cleanup: Wip ctypos ([pr#12495](#), xianxiaoxiao)
- cleanup: xio: provide dout_prefix for XioConnection ([pr#9444](#), Avner BenHanoch)
- cleanup: yasm-wrapper: translate “-isystem \$1” to “-i \$1” ([pr#12093](#), Kefu Chai)
- cmake: add -Wno-unknown-pragmas to CMAKE_CXX_FLAGS ([pr#12128](#), Kefu Chai)
- cmake: check WITH_RADOSGW for fcgi and expat dependencies ([pr#11481](#), David Disseldorf)
- cmake: compile C code with c99 ([pr#12369](#), Kefu Chai)
- cmake: detect keyutils if WITH_LIBCEPHFS OR WITH_RBD ([pr#12359](#), Kefu Chai)
- cmake: do not link erasure tests again libosd ([pr#11738](#), Kefu Chai)
- cmake: find gperftools package for tcmalloc_minimal too ([pr#11403](#), Bassam Tabbara)
- cmake: fix boost build on ubuntu 16.10 yakkety ([pr#12143](#), Bassam Tabbara)
- cmake: Fix for cross compiling ([pr#11404](#), Bassam Tabbara)
- cmake: fix git version string, cleanup ([pr#11661](#), Sage Weil)
- cmake: librbd cleanup ([pr#11842](#), Kefu Chai)
- cmake: link tests against static librados ([issue#17260](#), [pr#11575](#), Kefu Chai)
- cmake: pass CMAKE_BUILD_TYPE down to rocksdb ([pr#11767](#), Kefu Chai)
- cmake: remove include/Makefile.am ([pr#11666](#), Kefu Chai)
- cmake: replace civetweb symlink w/file copy ([pr#11900](#), Matt Benjamin)
- cmake: should link against \${ALLOC_LIBS} ([pr#11978](#), Kefu Chai)
- cmake: src/test/CMakeLists.txt: Exclude test on HAVE_BLKID ([pr#12301](#), Willem Jan Withagen)
- cmake: Support for embedding Ceph Daemons ([pr#11764](#), Bassam Tabbara)
- cmake: use external project for rocksdb ([pr#11385](#), Bassam Tabbara)
- common: Add throttle_get_started perf counter ([pr#12163](#), Bartłomiej Święcki)
- common: assert(0) -> ceph_abort() ([pr#12031](#), Sage Weil)
- common: auth: fix NULL pointer access when trying to delete CryptoAESKeyHandler instance ([pr#11614](#), runsisi)
- common,bluestore: compressor: fixes and tests; disable zlib isal (it's broken) ([pr#11349](#), Sage Weil)
- common,bluestore: mempool: mempool infrastructure, bluestore changes to use it ([pr#11331](#), Allen Samuels, Sage Weil)
- common: buffer: add advance(unsigned) back ([issue#17809](#), [pr#11993](#), Kefu Chai)
- common: buffer: add copy(unsigned, ptr) back ([issue#17809](#), [pr#12246](#), Kefu Chai)
- common: client/Client.cc: fix/silence “logically dead code” CID-Error ([pr#291](#), Yehuda Sadeh)
- common: common/strtol.cc: Get error testing also to work on FreeBSD ([pr#12034](#), Willem Jan Withagen)
- common: fix clang compilation error ([pr#12565](#), Mykola Golub)
- common: FreeBSD/EventKqueue.{h,cc} Added code to restore events on (thread)fork ([pr#11430](#), Willem Jan Withagen)
- common: log/LogClient: fill seq & who for syslog and graylog ([issue#16609](#), [pr#10196](#), Xiaoxi Chen)
- common: make l_finisher_complete_lat more accurate ([pr#11637](#), Pan Liu)
- common: msg/simple/Acceptor.cc: replace shutdown() with selfpipe event in poll() (FreeBSD) ([pr#10720](#), Willem Jan Withagen)
- common: osdc/Objecter: fix relock race ([issue#17942](#), [pr#12234](#), Sage Weil)
- common: osdc/Objecter: handle race between calc_target and handle_osd_map ([issue#17942](#), [pr#12055](#), Sage Weil)
- common: osd/osdmap: fix divide by zero error ([pr#12521](#), Yunchuan Wen)
- common: release g_ceph_context before returns ([issue#17762](#), [pr#11733](#), Kefu Chai)
- common: Remove the runtime dependency on lsb_release ([issue#17425](#), [pr#11365](#), Brad Hubbard)
- common: test/fio: fix global CephContext life cycle ([pr#12245](#), Igor Fedotov)
- core: auth: tolerate missing MGR keys during upgrade ([pr#11401](#), Sage Weil)
- core,bluestore: os/bluestore: fix warning and uninit variable ([pr#12032](#), Sage Weil)
- core,bluestore: os: fix offsets for move_ranges operation ([pr#11595](#), Sage Weil)
- core,bluestore: os: remove move_ranges_destroy_src ([pr#11791](#), Sage Weil)
- core: ceph-disk: allow using a regular file as a journal ([issue#17662](#), [pr#11619](#), Jayashree Candadai, Loic Dachary)
- core: ceph-disk: resolve race conditions ([issue#17889](#), [issue#17813](#), [pr#12136](#), Loic Dachary)
- core,cephfs: osdc/ObjectCacher: wake up dirty stat waiters after removing buffers ([issue#17275](#), [pr#11593](#), Yan, Zheng)
- core: ceph.in: allow ‘flags’ to not be present in cmddescs ([issue#18297](#), [pr#12540](#), Dan Mick)
- core,cleanup: ceph-disk: do not create bluestore wal/db partitions by default ([issue#18291](#), [pr#12531](#), Loic Dachary)
- core,cleanup,common: common/TrackedOp: remove unused ‘now’ in _dump() ([pr#12007](#), John Spray)
- core,cleanup: FileStore: Only verify split when it has been really done and done correctly ([pr#11731](#), Li Wang)
- core,cleanup: kv: remove snapshot iterator ([pr#12049](#), Sage Weil)
- core,cleanup: mon/MonClient.h: remove repeated searching of map ([pr#10601](#), Michal Jarzabek)

- core,cleanup: msg: Fix typos in socket creation error message ([pr#11907](#), Brad Hubbard)
- core,cleanup: osd/command tell: check pgid at the right time ([pr#11547](#), Javeme)
- core,cleanup: osd/OSDMap.cc: fix duplicated assignment for new_blacklist_entries ([pr#11799](#), Ker Liu)
- core,cleanup: osd/PG.cc: prevent repeated searching of map/set ([pr#11203](#), Michal Jarzabek)
- core,cleanup: osd/ReplicatedPG: remove redundant check for balance/localize read ([pr#10209](#), runsisi)
- core,cleanup: osd/ReplicatedPG: remove unneeded use of count ([pr#11242](#), Michal Jarzabek)
- core,cleanup: os/filestore: handle EINTR returned by io_getevents() ([pr#11890](#), Pan Liu)
- core,cleanup: os/ObjectStore: remove legacy tbl support ([pr#11770](#), Jianpeng Ma)
- core,cleanup: scan build fixes ([pr#12148](#), Kefu Chai)
- core,cleanup: src: rename ReplicatedPG to PrimaryLogPG ([pr#12487](#), Samuel Just)
- core,cleanup: Wip scrub misc ([pr#11397](#), David Zafman)
- core,common: buffer: put buffers in buffer_{data,meta} mempools ([pr#11839](#), Sage Weil)
- core,common: msg: add entity_addr_t types; add new entity_addrvec_t type ([pr#9825](#), Zhao Junwang, Sage Weil)
- core,common: msg/simple/Pipe: handle addr decode error ([issue#18072](#), [pr#12221](#), Sage Weil)
- core: compress: Fix compilation failure from missing header ([pr#12108](#), Adam C. Emerson)
- core: denc: don't pass null instances into encoder fns ([issue#17636](#), [pr#11577](#), John Spray)
- core: erasure-code: synchronize with upstream gf-complete ([issue#18092](#), [pr#12382](#), Loic Dachary)
- core: FreeBSD/OSD.cc: add client_messenger to the avoid_ports set. ([pr#12463](#), Willem Jan Withagen)
- core: include/object: pass "snapid_t&" to bound_encode() ([pr#11552](#), Kefu Chai)
- core: kv/RocksDBStore: Don't update rocksdb perf_context if rocksdb_perf di... ([pr#12064](#), Jianpeng Ma)
- core: librados-dev: install inline_memory.h ([issue#17654](#), [pr#11730](#), Josh Durgin)
- core: messages/MForward: reencode forwarded message if target has differing features ([pr#11610](#), Sage Weil)
- core,mgr: messages: fix out of range assertion ([pr#11345](#), John Spray)
- core: mon,ceph-disk: add lockbox permissions to bootstrap-osd ([issue#17849](#), [pr#11996](#), Loic Dachary)
- core: mon: make it more clearly to debug for paxos state ([pr#12438](#), song baisen)
- core: mon/OSDMonitor: encode full osdmmaps with features all OSDs can understand ([pr#11284](#), Sage Weil)
- core: mon/OSDMonitor: encode OSDMap::Incremental with same features as OSDMap ([pr#11596](#), Sage Weil)
- core: mon/OSDMonitor: newly created osd should not be wrongly marked in ([pr#11795](#), runsisi)
- core: mon/OSDMonitor: remove duplicate jewel/kraken flag warning ([pr#11775](#), Josh Durgin)
- core: mon/PGMap: PGs can be stuck more than one thing ([issue#17515](#), [pr#11339](#), Sage Weil)
- core: mon: print the num_pools and num_objects in 'ceph -s -f json/json-p... ([issue#17703](#), [pr#11654](#), huangjun)
- core: msg/async/AsyncConnection: dispatch write handler on keepalive2 ([issue#17664](#), [pr#11601](#), Ilya Dryomov)
- core: msg/async: DPDKStack as AsyncMessenger backend ([pr#10748](#), Haomai Wang)
- core: msg/async/rdma: change log level: 0 -> 1 ([pr#12334](#), Avner BenHanoch)
- core: msg/async/rdma: don't use more buffers than what device capabilities ... ([pr#12263](#), Avner BenHanoch)
- core: msg/async/rdma: ensure CephContext existed ([pr#12068](#), Haomai Wang)
- core: msg/async/rdma: event polling thread can block on event ([pr#12270](#), Haomai Wang)
- core: msg/async/rdma: fixup memory free ([pr#12236](#), gongchuang)
- core: msg/async/rdma: set correct value to memory manager ([pr#12299](#), Adir Lev)
- core: msg/async: set nonce before starting the workers ([pr#12390](#), Kefu Chai)
- core: msg: make loopback Connection feature accurate all the time ([pr#11183](#), Sage Weil)
- core: msg: seed random engine used for ms_type="random" ([pr#11880](#), Casey Bodley)
- core: msg/simple/Pipe: avoid returning 0 on poll timeout ([issue#18184](#), [pr#12375](#), Sage Weil)
- core: msg/simple/Pipe::stop_and_wait: unlock pipe_lock for stop_fast_dispatching() ([issue#18042](#), [pr#12307](#), Samuel Just)
- core: msg/simple: save the errno in case being changed by subsequent codes ([pr#10297](#), Yan Jun)
- core: osd/ECTransaction: only write out the hinfo if not delete ([issue#17983](#), [pr#12141](#), Samuel Just)
- core: OSDMonitor: only reject MOSDBoot based on up_from if inst matches ([issue#17899](#), [pr#12003](#), Samuel Just)
- core: osd,mon: require sortbitwise flag to upgrade beyond jewel ([pr#11772](#), Sage Weil)
- core: osd/osd_types: fix the osd_stat_t::decode() ([pr#12235](#), Kefu Chai)
- core: osd/PG: add "down" pg state (distinct from down+peering) ([pr#12289](#), Sage Weil)
- core: osd/PGLog::proc_replica_log,merge_log: fix bound for last_update ([issue#18127](#), [pr#12340](#), Samuel Just)
- core: osd/ReplicatedPG: do_update_log_missing: take the pg lock in the callback ([issue#17789](#), [pr#11754](#), Samuel Just)
- core: osd/ReplicatedPG::record_write_error: don't leak orig_reply on cancel ([issue#18180](#), [pr#12450](#), Samuel Just)
- core: os/filestore: avoid to get the wrong hardlink number. ([pr#11841](#), huangjun)
- core: os/filestore/chain_xattr.h:uses ENODATA, so include compat.h ([pr#12279](#), Willem Jan Withagen)
- core: os/filestore: Fix erroneous WARNING: max attr too small ([issue#17420](#), [pr#11246](#), Brad Hubbard)
- core: os/FileStore: fix fiemap issue in xfs when #extents > 1364 ([pr#11554](#), Ning Yao)
- core: os/filestore: fix journal logger ([pr#12099](#), Wei Jin)
- core: os/filestore: fix potential result code overwriting ([pr#11491](#), xie xingguo)
- core: os/filestore/HashIndex: fix list_by_hash_* termination on reaching end ([issue#17859](#), [pr#11898](#), Sage Weil)
- core: os/ObjectStore: properly clear object map when replaying OP_REMOVE ([issue#17177](#), [pr#11388](#), Yan, Zheng)
- core,performance: msg/async: ibverbs/rdma support ([pr#11531](#), Haomai Wang, Zhi Wang)
- core,performance: osd/OSDMap.cc: remove unneeded use of count ([pr#11221](#), Michal Jarzabek)
- core,performance: osd/PrimaryLogPG: don't truncate if we don't have to for WRITEFULL ([pr#12534](#), Samuel Just)

- core,performance: os/fs/FS: optimize aio::pwritev which make caller provide length. ([pr#9062](#), Jianpeng Ma)
- core,pybind,common: python-rados: implement new aio_execute ([pr#12140](#), Iain Buclaw)
- core,rbd,bluestore,rgw,performance,cephfs: fast denc encoding ([pr#11027](#), Sage Weil)
- core: remove spurious executable permissions on source code files ([pr#1061](#), Samuel Just)
- core: ReplicatedPG::failed_push: release read lock on failure ([issue#17857](#), [pr#11914](#), Kefu Chai)
- core: rocksdb: update to latest, and make it the default for the mons ([pr#11354](#), Sage Weil)
- core: set dumpable flag after setuid ([issue#17650](#), [pr#11582](#), Patrick Donnelly)
- core: systemd/ceph-disk: reduce ceph-disk flock contention ([issue#18049](#), [issue#13160](#), [pr#12200](#), David Disseldorp)
- core: tchaikov ([issue#17713](#), [pr#11382](#), Haomai Wang)
- core,tests: ceph_test_rados_api_tier: dump hitset that we fail to decode ([issue#17945](#), [pr#12057](#), Sage Weil)
- core,tests: common osd: Improve scrub analysis, list-inconsistent-obj output and osd-scrub-repair test ([issue#18114](#), [pr#9613](#), Kefu Chai, David Zafman)
- core,tests: test,cmake: turn unit.h into unit.cc to speed up compilation ([pr#12194](#), Kefu Chai)
- core,tests: test/rados/list.cc: Memory leak in ceph_test_rados_api_list ([issue#18250](#), [pr#12479](#), Brad Hubbard)
- core,tests: workunits/ceph-helpers.sh: Fixes for FreeBSD ([pr#12085](#), Willem Jan Withagen)
- core,tools: Added append functionality to rados tool. ([pr#11036](#), Tomy Cheru)
- core,tools: Tested-by: Huawei Ren <ren.huanwen@zte.com.cn> ([issue#17400](#), [pr#11276](#), Kefu Chai)
- core,tools: vstart: decrease pool size if <3 OSDs ([pr#11528](#), John Spray)
- crush: make counting of choose_tries consistent ([issue#17229](#), [pr#10993](#), Vicente Cheng)
- crush: remove the crush_lock ([pr#11830](#), Adam C. Emerson)
- crush: Silence coverity warnings for test/crush/crush.cc ([pr#12436](#), Brad Hubbard)
- doc: Add doc about osd scrub {during recovery|chunk {min|max}| sleep} ([pr#12176](#), Paweł Sadowski)
- doc: Add docs about looking up Monitors through DNS ([issue#14527](#), [pr#10852](#), Wido den Hollander)
- doc: add docs for raw compression ([pr#12244](#), Casey Bodley)
- doc: Add documentation about mon_allow_pool_delete before pool remove ([pr#11943](#), Wido den Hollander)
- doc: add infernalis EOL date ([pr#11925](#), Ken Dreyer)
- doc: adding changelog for v10.2.4 ([pr#12346](#), Abhishek Lekshmanan)
- doc: Add MON docs about pool flags and pool removal config settings ([pr#10853](#), Wido den Hollander)
- doc: add python-rgw doc ([pr#11859](#), Kefu Chai)
- doc: change the osd_max_backfills default to 1 ([issue#17701](#), [pr#11658](#), huangjun)
- doc: clarify file deletion from OSD restricted pool behaviour ([issue#17937](#), [pr#12054](#), David Disseldorp)
- doc: clarify mds deactivate purpose ([pr#11957](#), Patrick Donnelly)
- doc: common/Throttle: fix typo for BackoffThrottle ([pr#12129](#), Wei Jin)
- doc: correcting the object name ([pr#12354](#), Uday Mullangi)
- doc: Correcting the sample python tempurl generation script. ([issue#15258](#), [pr#8712](#), Diwakar Goel)
- doc: Coverity and SCA fixes ([pr#7784](#), Danny Al-Gaaf)
- doc: doc/dev/osd_internals: add pgpool.rst ([pr#12500](#), Brad Hubbard)
- doc: doc/dev/perf: a few notes on perf ([pr#12168](#), Sage Weil)
- doc: doc/dev/perf: fix dittoigraphy ([pr#12317](#), xie xingguo)
- doc: doc/man: avoid file builtin to solve build error ([pr#11984](#), Patrick Donnelly)
- doc: doc/rados/configuration/ms-ref.rst: document a few async msgr options ([pr#12126](#), Piotr Dałek)
- doc: doc/rados/configuration/osd-config-ref.rst: document the fast mark down ([pr#12124](#), Piotr Dałek)
- doc: doc/release-notes: kraken release notes (draft) ([pr#12338](#), Sage Weil)
- doc: doc/releases: add links to kraken and v10.2.4 ([pr#12409](#), Kefu Chai)
- doc: doc/start/hardware-recommendations: cosmetic ([pr#10585](#), Zhao Junwang)
- doc: Documentation syntax cleanup ([pr#11784](#), John Spray)
- doc: document osd tell bench ([issue#5431](#), [pr#16](#), Sage Weil)
- doc: drop -journal-check from ceph-mds man page ([issue#17747](#), [pr#11912](#), Nathan Cutler)
- doc: explain rgw_fcgi_socket_backlog in rgw/config-ref.rst ([pr#12548](#), liuchang0812)
- doc: final additions to 11.1.0-rc release notes ([pr#12448](#), Abhishek Lekshmanan)
- doc: Fix broken link for caps ([issue#17587](#), [pr#11546](#), Uday Mullangi)
- doc: fix broken links ([issue#17587](#), [pr#11518](#), Uday Mullangi)
- doc: fix dead link "Hardware Recommendations" ([pr#11379](#), Leo Zhang)
- doc: fix dead link of "os-recommendations" in troubleshooting-osd ([pr#11454](#), Leo Zhang)
- doc: Fixed mapping error in legacy mds command ([pr#11668](#), Malte Fiala)
- doc: Fix for worker arguments to cephfs-data-scan tool ([pr#12360](#), Wido den Hollander)
- doc: fix grammar/spelling in RGW sections ([pr#12329](#), Ken Dreyer)
- doc: Fixing the broken hyperlinks by pointing to correct documentation. ([pr#11617](#), Uday Mullangi)
- doc: fix librados example programs ([pr#11302](#), Alexey Sheplyakov)
- doc: fix mgr literal block rST syntax ([pr#11652](#), Ken Dreyer)
- doc: fix start development cluster operation in index.rst ([pr#11233](#), Leo Zhang)
- doc: fix the script for rebuild monitor db ([pr#11962](#), Kefu Chai)
- doc: fix typos ([pr#8751](#), Li Peng)
- doc: Flag deprecated mds commands and omit deprecated mon commands in help output ([pr#11434](#), Patrick Donnelly)
- doc: mailmap: change personal info ([pr#12310](#), Wei Jin)

- doc: mailmap updates sept (pr#10955, Yann Dupont)
- doc: mds: fixup “mds bal mode” Description (pr#12127, huanwen ren)
- doc: mention corresponding libvirt section in nova.conf (pr#12584, Marc Koderer)
- doc: Modify documentation for mon_osd_down_out_interval (pr#12408, Brad Hubbard)
- doc: network-protocol typos (pr#9837, Zhao Junwang)
- doc: openstack glance mitaka uses show_multiple_locations (pr#12020, Sébastien Han)
- doc: README.FreeBSD: update to match the bimonthly FreeBSD status report (pr#11442, Willem Jan Withagen)
- doc: README: hint at where to look to diagnose test failures (pr#11903, Dan Mick)
- doc: reformat SubmittingPatches with more rst syntax (pr#11570, Kefu Chai)
- doc: release notes for 10.2.4 (pr#12053, Abhishek Lekshmanan)
- doc: release notes for 10.2.5 (issue#18207, pr#12410, Loic Dachary)
- doc: release notes for 11.0.2 (pr#11369, Abhishek Lekshmanan)
- doc: Remove duplicate command for Ubuntu (pr#12186, chrone)
- doc: reviewed-by: John Wilkins <jowilkin@redhat.com> (issue#17526, pr#11352, Loic Dachary)
- doc: reviewed-by: John Wilkins <jowilkin@redhat.com> (issue#17665, pr#11602, Jason Dillaman)
- doc: rgw: fix a typo in S3 java api example (pr#11762, Weibing Zhang)
- doc: rm “type=rpm-md” from yum repositories (pr#10248, Ken Dreyer)
- doc: Small styling fix to mirror documentation (pr#9714, Wido den Hollander)
- doc: src/doc: fix class names in exports.txt (pr#12000, John Spray)
- doc: standardize EPEL instructions (pr#11653, Ken Dreyer)
- doc: update cinder key permissions for mitaka (pr#12211, Sébastien Han)
- doc: Update crush-map.rst, fix a typo mistake (pr#11785, whu_liuchang)
- doc: Update filestore xattr config documentation. (pr#11826, Bartłomiej Świecki)
- doc: Update install-ceph-gateway.rst (pr#11432, Hans van den Bogert)
- doc: Update keystone doc about v3 options (pr#11392, Proskurin Kirill)
- doc: Update layout.rst, move commands to CODE block (pr#11987, liuchang0812)
- doc: we can now run multiple MDS, so qualify warning (issue#18040, pr#12184, Nathan Cutler)
- fs: add snapshot tests to mds thrashing (pr#1073, Yan, Zheng)
- fs: enable ceph-fuse permission checking for all pjd suites (pr#1187, Greg Farnum)
- fs: fix two frag_enable fragments (issue#6143, pr#656, Sage Weil)
- fs: fix up dd testing again (issue#10861, pr#373, Greg Farnum)
- fs: fuse_default_permissions = 0 for kernel build test (pr#1109, Patrick Donnelly)
- fs: Mantle: A Programmable Metadata Load Balancer (pr#10887, Michael Sevilla)
- fs: unify common parts of sub-suites (issue#1737, pr#1282, Patrick Donnelly)
- librados: Add rados_aio_exec to the C API (pr#11709, Iain Buclaw)
- librados: add timeout to watch/notify (pr#11378, Ryne Li)
- librados: do not request osd ack if no completed completion is set (pr#11204, Sage Weil)
- librados: For C-API, expose LIBRADOS_OPERATION_FULL_FORCE flag (pr#9172, Jianpeng Ma)
- librados: improvements async IO in librados and libradosstriper (pr#10049, Sébastien Ponce)
- librados: Memory leaks in object_list_begin and object_list_end (issue#18252, pr#12482, Brad Hubbard)
- librados: postpone cct deletion (pr#11659, Kefu Chai)
- librados: remove new setxattr overload to avoid breaking the C++ ABI (issue#18058, pr#12206, Josh Durgin)
- librados: remove unused bufferlist from rados_write_op_rmxattr (pr#12030, Piotr Dałek)
- librbd: add support for snapshot namespaces (pr#11160, Victor Denisov)
- librbd: API changes to support separate data pool (pr#11353, Jason Dillaman)
- librbd: batch object map updates during trim (issue#17356, pr#11510, Venky Shankar)
- librbd: bug fixes for optional data pool support (pr#11960, Venky Shankar)
- librbd: cannot access non-primary image when mirroring force disabled (issue#16740, issue#17588, pr#11568, Jason Dillaman)
- librbd: cls_rbd updates for separate data pool (issue#17422, pr#11327, Jason Dillaman)
- librbd: default features should be negotiated with the OSD (issue#17010, pr#11808, Mykola Golub)
- librbd: diffs to clone’s first snapshot should include parent diffs (issue#18068, pr#12218, Jason Dillaman)
- librbd: do not create empty object map object on image creation (issue#17752, pr#11704, Jason Dillaman)
- librbd: enabling/disabling rbd feature should report missing dependency (issue#16985, pr#12238, Gaurav Kumar Garg)
- librbd: ensure consistency groups will gracefully fail on older OSDs (pr#11623, Jason Dillaman)
- librbd: exclusive lock incorrectly initialized when switching to head revision (issue#17618, pr#11559, Jason Dillaman)
- librbd: fix rollback if failed to disable mirroring for image (pr#11260, runsisi)
- librbd: ignore error when object map is already locked by current client (issue#16179, pr#12484, runsisi)
- librbd: ignore notify errors on missing image header (issue#17549, pr#11395, Jason Dillaman)
- librbd: keep rbd_default_features setting as bitmask (issue#18247, pr#12486, Jason Dillaman)
- librbd: mark request as finished after failed refresh (issue#17973, pr#12160, Venky Shankar)
- librbd: minor cleanup (pr#12078, Dongsheng Yang)
- librbd: new API method to force break a peer’s exclusive lock (issue#18429, issue#16988, issue#18327, pr#12889, Jason Dillaman)
- librbd: parse rbd_default_features config option as a string (pr#11175, Alyona Kiseleva, Alexey Sheplyakov)

- librbd: possible assert failure creating image when using data pool ([pr#11641](#), Venky Shankar)
- librbd: proper check for get_data_pool compatibility ([issue#17791](#), [pr#11755](#), Mykola Golub)
- librbd: properly order concurrent updates to the object map ([issue#16176](#), [pr#12420](#), Jason Dillaman)
- librbd: release lock after demote ([issue#17880](#), [pr#11940](#), Mykola Golub)
- librbd: remove consistency group rbd cli and API support ([issue#18231](#), [pr#12475](#), Jason Dillaman)
- librbd: remove image header lock assertions ([issue#18244](#), [pr#12472](#), Jason Dillaman)
- librbd: remove unused local variable ([pr#12388](#), Yunchuan Wen)
- librbd: silence the unused variable warning ([pr#11678](#), Kefu Chai)
- librbd: snap_get_limit compatibility check ([pr#11766](#), Mykola Golub)
- librbd: update internals to use optional separate data pool ([pr#11356](#), Jason Dillaman)
- librbd: use proper snapshot when computing diff parent overlap ([issue#18200](#), [pr#12396](#), Xiaoxi Chen)
- log: optimize header file dependency ([pr#9768](#), Xiaowei Chen)
- mds: add debug assertion for issue #17636 ([pr#11576](#), Yan, Zheng)
- mds: add tests for mantle (programmable balancer) ([pr#1145](#), Michael Sevilla)
- mds: check if down mds is known ([issue#17670](#), [pr#11611](#), Patrick Donnelly)
- mds: don't access mdsmap from log submit thread ([issue#18047](#), [pr#12208](#), Yan, Zheng)
- mds: don't maintain bloom filters in standby replay ([issue#16924](#), [pr#12133](#), John Spray)
- mds: enable rmxattr on pool_namespace attrs ([issue#17797](#), [pr#11783](#), John Spray)
- mds: fix dropping events in standby replay ([issue#17954](#), [pr#12077](#), John Spray)
- mds: fix EMetaBlob::fullbit xattr dump ([pr#11536](#), Sage Weil)
- mds: fix false "failing to respond to cache pressure" warning ([pr#11373](#), Yan, Zheng)
- mds: force client flush snap data before truncating objects ([issue#17193](#), [pr#11994](#), Yan, Zheng)
- mds: handle bad standby_for_fscids in fsmmap ([issue#17466](#), [pr#11281](#), John Spray)
- mds: ignore 'session evict' when mds is replaying log ([issue#17801](#), [pr#11813](#), Yan, Zheng)
- mds: include legacy client fsid in FSMap print ([pr#11283](#), John Spray)
- mds: more deterministic timing on frag split/join ([issue#17853](#), [pr#12022](#), John Spray)
- mds: more unique_pointer changes ([pr#11635](#), Patrick Donnelly)
- mds: properly commit new dirfrag before splitting it ([issue#17990](#), [pr#12125](#), Yan, Zheng)
- mds: release pool allocator memory after exceeding size limit ([issue#18225](#), [pr#12443](#), John Spray)
- mds: remove duplicated log in handle_client_readdir ([pr#11806](#), Zhi Zhang)
- mds: remove "-journal-check" help text ([issue#17747](#), [pr#11739](#), Nathan Cutler)
- mds: remove unused EFragment::OP_ONESHOT ([pr#11887](#), John Spray)
- mds: repair backtraces during scrub ([issue#17639](#), [pr#11578](#), John Spray)
- mds: require MAY_SET_POOL to set pool_ns ([issue#17798](#), [pr#11789](#), John Spray)
- mds: respawn using /proc/self/exe ([issue#17531](#), [pr#11362](#), Patrick Donnelly)
- mds: revert "mds/Mutation: remove redundant _dump method" ([issue#17906](#), [pr#11985](#), Patrick Donnelly)
- mds: use parse_filesystem in parse_role to handle exceptions and reuse parsing code ([issue#17518](#), [pr#11357](#), Patrick Donnelly)
- mds: use projected path construction for access ([issue#17858](#), [pr#12063](#), Patrick Donnelly)
- mds: use unique_ptr to simplify resource mgmt ([pr#11543](#), Patrick Donnelly)
- mgr: doc/mgr: fix mgr how long to wait to failover ([pr#11550](#), huanwen ren)
- mgr: init() return when connection daemons failed && add some err info ([pr#11424](#), huanwen ren)
- mgr: misc minor changes ([issue#17455](#), [pr#11386](#), xie xingguo)
- mgr: PyModules.cc: remove duplicated if condition for fs_map ([pr#11639](#), Weibing Zhang)
- mgr: remove unnecessary C_StdFunction ([pr#11883](#), John Spray)
- mon: add missing space in warning message ([pr#11361](#), Patrick Donnelly)
- mon: clean legacy code ([pr#9643](#), Wei Jin)
- mon: clear duplicated logic in MDSMonitor ([pr#11209](#), Zhi Zhang)
- mon: Do not allow pools to be deleted by default ([pr#11665](#), Wido den Hollander)
- mon: fix "OSDs marked OUT wrongly after monitor failover" ([issue#17719](#), [pr#11664](#), Dong Wu)
- mon: Forbidden copy and assignment function in monoprequest ([pr#9513](#), song baisen)
- mon: have mon-specific features & rework internal monmap structures ([pr#10907](#), Joao Eduardo Luis)
- mon: if crushtool config is empty use internal crush test ([pr#11765](#), Bassam Tabbara)
- mon: make MDSMonitor tolerant of slow mon elections ([issue#17308](#), [pr#11167](#), John Spray)
- mon: MonmapMonitor: return success when monitor will be removed ([issue#17725](#), [pr#11747](#), Joao Eduardo Luis)
- mon: move case CEPH_MSG_POOLOP to OSDs group ([pr#11848](#), Javeme)
- mon: osdmap's epoch should be more than 0 ([pr#9859](#), Na Xie)
- mon: OSDMonitor: fix the check error of pg creating ([issue#17169](#), [pr#10916](#), DesmondS)
- mon: paxos add the timeout function when peon recovery ([pr#10359](#), song baisen)
- mon: preserve osd weight when marking osd out, then in ([pr#11293](#), Sage Weil)
- mon: prevent post-jewel OSDs from booting if require_jewel_osds is not set ([pr#11498](#), Sage Weil)
- mon: remove ceph-create-keys from mon startup ([issue#16036](#), [pr#9345](#), Owen Synge)
- mon: remove the redundant judgement in LogMonitor tick function ([pr#10474](#), song baisen)
- mon: remove utime_t param in _dump ([pr#12029](#), Patrick Donnelly)
- mon: send updated monmap to its subscribers ([issue#17558](#), [pr#11456](#), Kefu Chai)

- mon: small change on the HealthMonitor start_epoch function ([pr#10296](#), songbaisen)
- mon: support for building without leveldb + mon mkfs bug fix ([pr#11800](#), Bassam Tabbara)
- osd: add a pg_fastinfo attribute to reduce per-io metadata updates ([pr#11213](#), Sage Weil)
- osd: Add config option to disable new scrubs during recovery ([issue#17866](#), [pr#11874](#), Wido den Hollander)
- osd: a few fast dispatch optimizations ([pr#12052](#), Sage Weil)
- osd: cleanup C_CompleteSplits::finish() ([pr#12094](#), Jie Wang)
- osd: clean up PeeringWQ::_dequeue(), remove unnecessary variable ([pr#12117](#), Jie Wang)
- osd: clean up process_peering_events ([pr#12009](#), Jie Wang)
- osdc/Objecter: resend pg commands on interval change ([issue#18358](#), [pr#12910](#), Samuel Just)
- osd: condition OSDMap encoding on features ([pr#12166](#), Sage Weil)
- osd: default osd_scrub_during_recovery=false ([pr#12402](#), Sage Weil)
- osd: do not open pgs when the pg is not in pg_map ([issue#17806](#), [pr#11803](#), Xinze Chi)
- osd: drop stray debug message ([pr#11296](#), Sage Weil)
- osd: EC Overwrites ([issue#17668](#), [pr#11701](#), Tomy Cheru, Samuel Just)
- osd: enhance logging for osd network error ([pr#12458](#), liuchang0812)
- osd: fix CEPH OSD FLAG_RWORDERED ([pr#12603](#), Sage Weil)
- osd: fix duplicated id of incompat feature "fastinfo" ([pr#11588](#), xie xingguo)
- osd: fix ec scrub errors ([issue#17999](#), [pr#12306](#), Samuel Just)
- osd: fixes to make rbd on ec work ([pr#12305](#), Samuel Just)
- osd: Fix map gaps again (bug 15943) ([issue#15943](#), [pr#12571](#), Samuel Just)
- osd: fix memory leak from EC write workload ([issue#18093](#), [pr#12256](#), Sage Weil)
- osd: fix rados write op hang ([pr#11143](#), Yunchuan Wen)
- osd: Fix read error propagation in ECBBackend ([issue#17966](#), [pr#12142](#), Samuel Just)
- osd: fix scrub boundary to not include a SnapSet ([pr#11255](#), Samuel Just)
- osd: fix signed/unsigned comparison warning ([pr#12400](#), Greg Farnum)
- osd: fix typo in PG::clear_primary_state ([pr#11513](#), Brad Hubbard)
- osd: Fix typos in PG::find_best_info ([pr#11515](#), Brad Hubbard)
- osd: fix typos in "struct OSDOp" comments ([pr#12350](#), Chanyoung Park)
- osd: Flush Journal on shutdown ([pr#11249](#), Wido den Hollander)
- osd: force watch PING to be write ordered ([issue#18310](#), [pr#12590](#), Samuel Just)
- osd: handle EC recovery read errors ([issue#13937](#), [pr#9304](#), David Zafman)
- osd: heartbeat peers need to be updated when a new OSD added into an existed cluster ([issue#18004](#), [pr#12069](#), Pan Liu)
- osd: Increase priority for inactive PGs backfill ([pr#12389](#), Bartłomiej Święcki)
- osd: kill PG_STATE_SPLITTING ([pr#11824](#), xie xingguo)
- osd: mark queued flag for op ([pr#12352](#), Yunchuan Wen)
- osd: osdc: pass a string reference type to "osdmap->lookup_pg_pool_name" ([pr#12219](#), Leo Zhang)
- osd: osd/OSDMonitor: accept 'osd pool set ...' value as string ([pr#911](#), David Zafman)
- osd: PGLog: initialize writeout_from in PGLog constructor ([issue#12973](#), [pr#558](#), Sage Weil)
- osd/PrimaryLogPG: don't update digests for objects with mismatched names ([issue#18409](#), [pr#12803](#), Samuel Just)
- osd/PrimaryLogPG::failed_push: update missing as well ([issue#18165](#), [pr#12911](#), Samuel Just)
- osd: print log when osd want to kill self ([pr#9288](#), Haomai Wang)
- osd: Remove extra call to reg_next_scrub() during splits ([issue#16474](#), [pr#11206](#), David Zafman)
- osd: remove redundant call of heartbeat_check ([pr#12130](#), Pan Liu)
- osd: remove the lock heartbeat_update_lock, and change heartbeat_need_... ([pr#12461](#), Pan Liu)
- osd: remove the redundant clear method in consume_map function ([pr#10553](#), song bisen)
- osd: Remove unused '_lsb_release_' declarations ([pr#11364](#), Brad Hubbard)
- osd: replace hb_out and hb_in with a single hb_peers ([issue#18057](#), [pr#12178](#), Pan Liu)
- osd: ReplicatedPG: don't bless C OSD SendMessageOnConn ([issue#13304](#), [pr#669](#), Jason Dillaman)
- osd: set server-side limits on omap get operations ([pr#12059](#), Sage Weil)
- osd: When deep-scrub errors present upgrade regular scrubs ([pr#12268](#), David Zafman)
- performance,bluestore: kv/MemDB: making memdb code adapt to generic maps ([pr#11436](#), Ramesh Chander)
- performance,bluestore: os/bluestore: allow default to buffered write ([pr#11301](#), Sage Weil)
- performance,bluestore: os/bluestore: bluestore_cache_meta_ratio = .5 ([pr#11919](#), Sage Weil)
- performance,bluestore: os/bluestore: reduce Onode in-memory footprint ([pr#12568](#), Igor Fedotov)
- performance,bluestore: os/bluestore: refactor bluestore_sync_submit_transaction ([pr#11537](#), Sage Weil)
- performance,bluestore: os/bluestore: speed up omap-key generation for same onode(the read paths) ([pr#11894](#), xie xingguo)
- performance,bluestore: os/bluestore: speedup the performance of multi-replication flow by switc... ([pr#11844](#), Pan Liu)
- performance,cephfs: Fix long stalls when calling ceph_fsync() ([issue#17563](#), [pr#11710](#), Jeff Layton)
- performance,cleanup: Context: std::move the callback param in FunctionContext's ctor ([pr#11892](#), Kefu Chai)
- performance,cleanup: osd/PG.h: move shared ptr instead of copying it ([pr#11154](#), Michal Jarzabek)
- performance,common: common/config_opts.h: Optimized RocksDB WAL settings. ([pr#11530](#), Mark Nelson)
- performance,common: osd/OSDMap: improve the performance of pg_to_acting_osds ([pr#12190](#), Pan Liu)
- performance: msg/async: set ms_async_send_inline to false to improve small randread iops ([pr#11521](#), Mark Nelson)

- performance.tools: rados: add hints to rados bench ([pr#12169](#), Sage Weil)
- pybind: avoid “exception ‘int’ object is not iterable” ([pr#11532](#), Javeme)
- pybind,cephfs: ceph_volume_client: fix recovery from partial auth update ([issue#17216](#), [pr#11304](#), Ramana Raja)
- pybind,cephfs: ceph_volume_client: set an existing auth ID’s default mon caps ([issue#17800](#), [pr#11917](#), Ramana Raja)
- pybind: ceph-rest-api: understand the new style entity_addr_t representation ([issue#17742](#), [pr#11686](#), Kefu Chai)
- pybind: clean up mgr stuff for flake8 ([pr#11314](#), John Spray)
- pybind: fix build failure of rgwfile binding ([pr#11825](#), Kefu Chai)
- pybind: pybind/rados: add missing “length” requires for aio_execute() ([pr#12439](#), Kefu Chai)
- pybind: pybind/rados: Add @requires for all aio methods ([pr#12327](#), Iain Buclaw)
- qa: fixed distros links ([pr#12773](#), Yuri Weinstein)
- qa: Fixed link to centos distro ([pr#12768](#), Yuri Weinstein)
- qa/suites: switch from centos 7.2 to 7.x ([pr#12632](#), Sage Weil)
- qa/tasks/peer: update task based on current peering behavior ([issue#18330](#), [pr#12614](#), Sage Weil)
- qa/tasks/workunit: clear clone dir before retrying checkout ([issue#18336](#), [pr#12630](#), Sage Weil)
- qa: update Ubuntu image url after ceph.com refactor ([issue#18542](#), [pr#12960](#), Jason Dillaman)
- qa/workunits/rbd/test_lock_fence.sh fails ([issue#18388](#), [pr#12752](#), Nathan Cutler)
- rbd: added rbd-nbd fsx test case ([pr#1049](#), Jason Dillaman)
- rbd: add fsx journal replay test case ([pr#821](#), Jason Dillaman)
- rbd: add singleton to assert no rbdmap regression ([issue#14984](#), [pr#902](#), Nathan Cutler)
- rbd: add some missing workunits ([pr#870](#), Josh Durgin)
- rbd: add support for separate image data pool ([issue#17424](#), [pr#11355](#), Jason Dillaman)
- rbd: expose rbd unmap options ([issue#17554](#), [pr#11370](#), Ilya Dryomov)
- rbd: fix json formatting for image and journal status output ([issue#18261](#), [pr#12525](#), Mykola Golub)
- rbd: fix parsing of group and image specific pools ([pr#11632](#), Victor Denisov)
- rbd: journal: do not prematurely flag object recorder as closed ([issue#17590](#), [pr#11520](#), Jason Dillaman)
- rbd: krbd: kernel client expects ip[:port], not an entity_addr_t ([pr#11902](#), Ilya Dryomov)
- rbd: -max_part and -nbd_max options for nbd map ([issue#18186](#), [pr#12379](#), Pan Liu)
- rbd: move nbd test workload to separate client host from OSDs ([pr#1170](#), Jason Dillaman)
- rbd: provision volumes to format as XFS ([issue#6693](#), [pr#1028](#), Loic Dachary)
- rbd: rbd-mirror: fix sparse read optimization in image sync ([issue#18146](#), [pr#12368](#), Mykola Golub)
- rbd: rbd-mirror HA: move librbd::image_watcher::Notifier to librbd::object_watcher ([issue#17017](#), [pr#11290](#), Mykola Golub)
- rbd: rbd-mirror: recovering after split-brain ([issue#16991](#), [issue#18051](#), [pr#12212](#), Mykola Golub)
- rbd: rbd-mirror: snap protect of non-layered image results in split-brain ([issue#16962](#), [pr#11744](#), Mykola Golub)
- rbd: rbd-nbd: disallow mapping images >2TB in size ([issue#17219](#), [pr#11741](#), Mykola Golub)
- rbd: rbd-nbd: invalid error code for “failed to read nbd request” messages ([issue#18242](#), [pr#12483](#), Mykola Golub)
- rbd: rbd-nbd: restart parent process logger after forking ([issue#18070](#), [pr#12222](#), Jason Dillaman)
- rbd: rbd-nbd: support disabling auto-exclusive lock transition logic ([issue#17488](#), [pr#11438](#), Mykola Golub)
- rbd: rbd-nbd: support partition for rbd-nbd mapped raw block device ([issue#18115](#), [pr#12259](#), Pan Liu)
- rbd: tests with rbd_skip_partial_discard option enabled ([pr#1077](#), Mykola Golub)
- rbd,tools: rbd : make option -stripe-unit w/ B/K/M work ([pr#12407](#), Jianpeng Ma)
- rbd: updated tests to use new rbd default feature set ([pr#842](#), Jason Dillaman)
- rbd: use snap_remove implementation from internal ([pr#12035](#), Victor Denisov)
- rgw: add default zone name ([issue#7009](#), [pr#954](#), Orit Wasserman)
- rgw: add documentation for upgrading with rgw_region_root_pool ([pr#12138](#), Orit Wasserman)
- rgw: add option to log custom HTTP headers (rgw_log_http_headers) ([pr#7639](#), Matt Benjamin)
- rgw: add recovery procedure for upgrade to older version of jewel ([issue#17820](#), [pr#11827](#), Orit Wasserman)
- rgw: add rgw_compression_type=random for teuthology testing ([pr#11901](#), Casey Bodley)
- rgw: add sleep to let the sync agent init ([pr#1136](#), Orit Wasserman)
- rgw: add suport for creating S3 type subuser of admin rest api ([issue#16682](#), [pr#10325](#), snakeAngel2015)
- rgw: add support for the prefix parameter in account listing of Swift API ([issue#17931](#), [pr#12047](#), Radoslaw Zarzynski)
- rgw: allow fastcgi idle timeout to be adjusted ([pr#230](#), Sage Weil)
- rgw: also approve, passed teuthology (many false positives in several classes) ([issue#17985](#), [pr#12224](#), Yehuda Sadeh, Sage Weil)
- rgw: Anonymous users shouldn’t be able to access requester pays buckets. ([issue#17175](#), [pr#11719](#), Zhang Shaowen)
- rgw: aws4: add presigned url bugfix in runtime ([issue#16463](#), [pr#10160](#), Javier M. Mellid)
- rgw: bucket resharding ([issue#17550](#), [pr#11230](#), Yehuda Sadeh)
- rgw:bugfix for deleting objects name beginning and ending with underscores of one bucket using POST method of AWS’s js sdk. ([issue#17888](#), [pr#11982](#), root)
- rgw: Class member cookie is not initialized correctly in some coroutine’s constructor. ([pr#11673](#), Zhang Shaowen)
- rgw: clean up RGWShardedOmapCRManager on early return ([issue#17571](#), [pr#11505](#), Casey Bodley)
- rgw: clear data_sync_cr if RGWDataSyncControlCR fails ([issue#17569](#), [pr#11506](#), Casey Bodley)
- rgw: compilation of the ASIO front-end is enabled by default. ([pr#12073](#), Radoslaw Zarzynski)
- rgw: compression uses optional::emplace instead of in-place factories ([pr#12021](#), Radoslaw Zarzynski)
- rgw: conform to the standard usage of string::find ([pr#10086](#), Yan Jun)

- rgw: data_extra_pool is unique per zone ([issue#17025](#), [pr#1119](#), Orit Wasserman)
- rgw: delete entries_index in RGWFetchAllMetaCR ([issue#17812](#), [pr#11816](#), Casey Bodley)
- rgw: do not abort when accept a CORS request with short origin ([pr#12381](#), LiuYang)
- rgw: do not enable both tcp and uds for fastcgi ([issue#5797](#), [pr#479](#), Andrew Schoen)
- rgw: don't error out on empty owner when setting acls ([issue#6892](#), [pr#877](#), Loic Dachary, Nathan Cutler)
- rgw: Don't loop forever when reading data from 0 sized segment. ([issue#17692](#), [pr#11567](#), Marcus Watts)
- rgw: dont set CURLOPT_UPLOAD for GET requests ([issue#17822](#), [pr#12105](#), Casey Bodley)
- rgw: don't store empty chains in gc ([issue#17897](#), [pr#11969](#), Yehuda Sadeh)
- rgw: do quota tests on ubuntu ([issue#6382](#), [pr#635](#), Sage Weil)
- rgw: dump objects in RGWBucket::check_object_index() ([issue#14589](#), [pr#11324](#), Yehuda Sadeh)
- rgw: dump remaining coroutines when cr deadlock is detected ([pr#11580](#), Casey Bodley)
- rgw: extract host name from host:port string ([issue#17788](#), [pr#11751](#), Yehuda Sadeh)
- rgw: Fixed problem with PUT with x-amz-copy-source when source object is compressed. ([pr#12253](#), Adam Kupczyk)
- rgw: fixes for virtual hosting of buckets ([issue#17440](#), [issue#15975](#), [issue#17136](#), [pr#11280](#), Casey Bodley, Robin H. Johnson)
- rgw: fix etag in multipart complete ([issue#17794](#), [issue#6830](#), [issue#16129](#), [issue#17872](#), [pr#1269](#), Casey Bodley, Orit Wasserman)
- rgw: fix for bucket delete racing with mdlog sync ([issue#17698](#), [pr#11648](#), Casey Bodley)
- rgw: fix for passing temporary in InitBucketSyncStatus ([issue#17661](#), [pr#11594](#), Casey Bodley)
- rgw: fix for unsafe change of rgw_zonegroup ([issue#17962](#), [pr#12075](#), Casey Bodley)
- rgw: fix indentation for cache_pools ([issue#8295](#), [pr#251](#), Sage Weil)
- rgw: fix missing master zone for a single zone zonegroup ([issue#17364](#), [pr#11965](#), Orit Wasserman)
- rgw: fix osd crashes when execute "radosgw-admin bi list -max-entries=1" command ([issue#17745](#), [pr#11697](#), weiqiaomiao)
- rgw: fix put_acls for objects starting and ending with underscore ([issue#17625](#), [pr#11566](#), Orit Wasserman)
- rgw: fix RGWSimpleRadosLockCR set_description() ([pr#11961](#), Tianshan Qu)
- rgw: fix the field 'total_time' of log entry in log show opt ([issue#17598](#), [pr#11425](#), weiqiaomiao)
- rgw: fix uncompressed object size deduction in RGWRados::copy_obj_data. ([issue#17803](#), [pr#11794](#), Radoslaw Zarzynski)
- rgw: frontend subsystem rework ([pr#10767](#), Radoslaw Zarzynski, Casey Bodley, Matt Benjamin)
- rgw: ftw ([issue#17888](#), [pr#12262](#), Casey Bodley)
- rgw: get_system_obj does not use result of get_system_obj_state ([issue#17580](#), [pr#11444](#), Casey Bodley)
- rgw: get_zonegroup() uses "default" zonegroup if empty ([issue#17372](#), [pr#11207](#), Yehuda Sadeh)
- rgw: handle empty POST condition ([issue#17635](#), [pr#11581](#), Yehuda Sadeh)
- rgw: handle Swift auth errors in a way compatible with new Tempests. ([issue#16590](#), [pr#10021](#), Radoslaw Zarzynski)
- rgw: json encode/decode index_type, allow modification ([issue#17755](#), [pr#11707](#), Yehuda Sadeh)
- rgw: loses realm/period/zonegroup/zone data: period overwritten if somewhere in the cluster is still running Hammer ([issue#17371](#), [pr#11426](#), Orit Wasserman)
- rgw: make RGWLocalAuthApplier::is_admin_of() aware about system users. ([issue#18106](#), [pr#12283](#), Radoslaw Zarzynski)
- rgw: metadata sync info should be shown at master zone of slave zoneg... ([issue#18091](#), [pr#12187](#), Jing Wenjun)
- rgw: minor cleanup ([pr#10057](#), Yan Jun)
- rgw: move compression config into zone placement ([pr#12113](#), Casey Bodley)
- rgw: move xfs to a seperate directory ([pr#969](#), Orit Wasserman)
- rgw: multipart upload copy ([issue#12790](#), [pr#11269](#), Yehuda Sadeh, Javier M. Mellid)
- rgw: need to close_section in lc list op ([pr#12232](#), weiqiaomiao)
- rgw: policy acl format should be xml ([pr#946](#), Orit Wasserman)
- rgw: radosgw-admin: more on placement configuration ([issue#18078](#), [pr#12242](#), Casey Bodley)
- rgw: region conversion respects pre-existing rgw_region_root_pool ([issue#17963](#), [pr#12076](#), Casey Bodley)
- rgw: remove a redundant judgement when listng objects. ([pr#10849](#), zhangshaowen)
- rgw: remove circular reference in RGWAyncRadosRequest ([issue#17793](#), [issue#17792](#), [pr#11815](#), Casey Bodley)
- rgw: remove suggestion to upgrade libcurl ([pr#11630](#), Casey Bodley)
- rgw: remove unused variable "ostr" in rgw_b64.h and fix the comment ([pr#11329](#), Weibing Zhang)
- rgw: Replacing '+' with "%20" in canonical uri for s3 v4 auth. ([issue#17076](#), [pr#10919](#), Pritha Srivastava)
- rgw: revert unintentional change to civetweb ([pr#12004](#), Bassam Tabbara)
- rgw: rgw-admin: new commands to control placement ([issue#18078](#), [pr#12230](#), Yehuda Sadeh)
- rgw: RGWBucketSyncStatusManager uses existing async_rados ([issue#18083](#), [pr#12229](#), Casey Bodley)
- rgw: rgw_file: apply missed base64 try-catch ([issue#17663](#), [pr#11671](#), Matt Benjamin)
- rgw: RGWHTTPArgs::get_str() - return argument string that was set. ([pr#10672](#), Marcus Watts)
- rgw: rgw_multisite: fix the increamtal bucket sync init ([issue#17624](#), [pr#11553](#), Zengran Zhang)
- rgw: rgw_multisite: use a rados lock to coordinate data log trimming ([pr#10546](#), Casey Bodley)
- rgw: RGW Python bindings - use explicit array ([pr#11831](#), Daniel Grynewicz)
- rgw: rgw_rados.cc fix shard_num format for snprintf ([pr#11493](#), Weibing Zhang)
- rgw: rgw/rgw_file.cc: Add compat.h to allow CLOCK_MONOTONE ([pr#12309](#), Willem Jan Withagen)
- rgw: RGWSimpleRadosReadCR tolerates empty reads ([issue#17568](#), [pr#11504](#), Casey Bodley)

- rgw: [RGW] Wip rgw compression ([pr#11494](#), Alyona Kiseleva, Adam Kupczyk, Casey Bodley)
- rgw: set duration for lifecycle lease ([issue#17965](#), [pr#12231](#), Yehuda Sadeh)
- rgw: should assign ‘olh_bl’ to state.attrset[RGW_ATTR_OLH_ID_TAG] instead of ‘bl’ ([pr#10239](#), weiqiaomiao)
- rgw: skip empty http args in method parse() to avoid extra effort ([pr#11989](#), Guo Zhandong)
- rgw: split osd’s in 2 nodes ([issue#15612](#), [pr#1019](#), Vasu Kulkarni)
- rgw: support for x-robots-tag header ([issue#17790](#), [pr#11753](#), Yehuda Sadeh)
- rgw: sync modules, metadata search ([pr#10731](#), Yehuda Sadeh)
- rgw: Update version of civetweb to 1.8 ([pr#11343](#), Marcus Watts)
- rgw: use civetweb if no frontend was configured ([pr#958](#), Orit Wasserman)
- rgw: use explicit flag to cancel RGWCoroutinesManager::run() ([issue#17465](#), [pr#12452](#), Casey Bodley)
- rgw: valgrind fixes for kraken ([issue#18414](#), [issue#18407](#), [issue#18412](#), [issue#18300](#), [pr#12949](#), Casey Bodley)
- rgw: verified that failed check is in osd-scrub-repair.sh ([issue#17850](#), [pr#11881](#), Matt Benjamin)
- rgw: we don’t support btrfs any more ([pr#1132](#), Orit Wasserman)
- rgw: Wip rgwfile pybind ([pr#11624](#), Haomai Wang)
- tests,bluestore: os/bluestore: add UT for an estimation of Onode in-memory size ([pr#12532](#), Igor Fedotov)
- tests,bluestore: os/test/store_test: fix legacy bluestore cache settings application ([pr#11915](#), Igor Fedotov)
- tests: ceph-disk: force debug monc = 0 ([issue#17607](#), [pr#11534](#), Loic Dachary)
- tests: ceph_objectstore_tool.py: Don’t use btrfs on FreeBSD ([pr#10507](#), Willem Jan Withagen)
- tests: ceph_test_objectstore: fix Rename test ([pr#12261](#), Sage Weil)
- tests: check hostname -fqdn sanity before running make check ([issue#18134](#), [pr#12297](#), Nathan Cutler)
- tests,cleanup,rbd: test/librbd: in test_notify set object-map and fast-diff features by default ([pr#11821](#), Mykola Golub)
- tests,cleanup: test_bloom_filter.cc: Fix a mismatch for the random_seed parameter ([pr#11774](#), Willem Jan Withagen)
- tests,cleanup: test/osd/osd-fast-mark-down.sh: remove unnecessary teardown() calls ([pr#12101](#), Kefu Chai)
- tests,cleanup: test/osd-scrub-repair.sh: use repair() instead of “ceph pg repair” ([pr#12036](#), Kefu Chai)
- tests,cleanup: test/rados: remove unused bufferlist variable ([pr#10221](#), Yan Jun)
- tests,common: test: add perf-reset test in test/perf_counters.cc ([pr#8948](#), wangsongbo)
- tests: disable failing tests ([issue#17561](#), [issue#17757](#), [pr#11714](#), Loic Dachary)
- tests: disable the echo when running get_timeout_delays() ([pr#12180](#), Kefu Chai)
- tests: do not use memstore.test_temp_dir in two tests ([issue#17743](#), [pr#12281](#), Loic Dachary)
- tests: erasure-code: add k=2, m=2 for isa & jerasure ([issue#18188](#), [pr#12383](#), Loic Dachary)
- tests: facilitate background process debug in ceph-helpers.sh ([issue#17830](#), [pr#12183](#), Loic Dachary)
- tests: fix ceph-helpers.sh wait_for_clean delays ([issue#17830](#), [pr#12095](#), Loic Dachary)
- tests: fix osd-scrub-repair.sh ([pr#12072](#), David Zafman)
- tests: Fix racey test by setting noout flag (tracker 17757) ([issue#17757](#), [pr#11715](#), David Zafman)
- tests: merge ceph-qa-suite
- tests: Minor clean-ups ([pr#12048](#), David Zafman)
- tests: minor make check cleanup ([pr#12146](#), David Zafman)
- tests: no python3 tests for ceph-disk ([issue#17923](#), [pr#12025](#), Loic Dachary)
- tests: osd-crush.sh must retry crush dump ([issue#17919](#), [pr#12016](#), Loic Dachary)
- tests: osd-scrub-repair.sh abort if add_something fails ([pr#12172](#), Loic Dachary)
- tests: os/memstore: fix a mem leak in MemStore::Collection::create_object() ([pr#12201](#), Kefu Chai)
- tests: os/memstore, os/filestore: fix store_test’s to satisfy rm_coll behavi... ([pr#11558](#), Igor Fedotov)
- tests: paxos fixes ([issue#11913](#), [pr#457](#), John Spray)
- tests: pin flake8 to avoid behavior changes ([issue#17898](#), [pr#11971](#), Loic Dachary)
- tests: qa: fixed script to schedule rados and other suites with –subset option ([pr#12587](#), Yuri Weinstein)
- tests: qa/tasks/admin_socket: subst in repo name ([pr#12508](#), Sage Weil)
- tests: qa/tasks/ceph_deploy: use dev option instead of dev-commit ([pr#12514](#), Vasu Kulkarni)
- tests: qa/tasks/osd_backfill.py: wait for osd.[12] to start ([issue#18303](#), [pr#12577](#), Sage Weil)
- tests: qa/workunits/cephtool/test.sh: FreeBSD has no distro. ([pr#11702](#), Willem Jan Withagen)
- tests: qa/workunits: include extension for nose tests ([pr#12572](#), Sage Weil)
- tests: qa/workunits/rados/test_envlibrados_for_rocksdb: force librados-dev install ([pr#11941](#), Sage Weil)
- tests,rbd: qa/workunits/rbd: fix ([issue#18271](#), [pr#12511](#), Sage Weil)
- tests,rbd: qa/workunits/rbd: removed qemu-io test case 077 ([issue#10773](#), [pr#12366](#), Jason Dillaman)
- tests,rbd: qa/workunits/rbd: simplify running nbd test under build env ([pr#11781](#), Mykola Golub)
- tests,rbd: qa/workunits/rbd: use image id when probing for image presence ([issue#18048](#), [pr#12195](#), Mykola Golub)
- tests,rbd: qa/workunits/rbd: use more recent qemu-io tests that support Xenial ([issue#18149](#), [pr#12371](#), Jason Dillaman)
- tests,rbd: rbd-mirror: fix gmock warnings in bootstrap request unit tests ([issue#18156](#), [pr#12344](#), Mykola Golub)
- tests,rbd: rbd-mirror: improve resiliency of stress test case ([issue#17416](#), [pr#11326](#), Jason Dillaman)
- tests,rbd: test: new librbd discard after write test case ([pr#11645](#), Jason Dillaman)
- tests,rbd: test: skip TestLibRBD.DiscardAfterWrite if skip partial discard enabled ([issue#17750](#), [pr#11703](#), Jason Dillaman)
- tests,rbd: test: TestJournalReplay test cases need to wait for event commit ([issue#17566](#), [pr#11480](#), Jason Dillaman)
- tests: remove TestPGLog EXPECT_DEATH tests ([issue#18030](#), [pr#12361](#), Loic Dachary)
- tests: save 9 characters for asok paths ([issue#16014](#), [pr#12066](#), Loic Dachary)
- tests: sync ceph-erasure-code-corpus for using ‘arch’ not ‘uname -p’ ([pr#12024](#), Kefu Chai)

- tests: test/ceph_crypto: do not read ceph.conf in global_init() ([issue#18128](#), [pr#12318](#), Kefu Chai)
- tests: test: ceph-objectstore-tool: should import platform before using it ([pr#12038](#), Kefu Chai)
- tests: test/ceph_test_msgr: do not use Message::middle for holding transient... ([issue#17728](#), [pr#11680](#), Kefu Chai)
- tests: test: disable osd-scrub-repair and test-erasure-eio ([issue#17830](#), [pr#12058](#), Loic Dachary, Dan Mick)
- tests: test: disable osd-scrub-repair and test-erasure-eio ([pr#11979](#), Dan Mick)
- tests: test: Don't write to a poolid that this test might not have created ([pr#12378](#), David Zafman)
- tests: test: enable unittest_dns_resolve ([pr#12209](#), Kefu Chai)
- tests: test/encoding/readable.sh: fix shell script warning ([pr#11527](#), Willem Jan Withagen)
- tests: TestErasureCodePluginJerasure must stop the log thread ([issue#17561](#), [pr#11721](#), Loic Dachary)
- tests: test: fix test-erasure-eio and osd-scrub-repair races (17830) ([pr#11926](#), David Zafman)
- tests: test/osd-fast-mark-down.sh: wrong assumption on first subtest ([pr#12123](#), Piotr Dałek)
- tests: test/osd-fast-mark-down.sh: introduce large timeout ([issue#17918](#), [pr#12019](#), Piotr Dałek)
- tests: test/osd-scrub-repair.sh: Use test case specific object names to help... ([pr#11449](#), David Zafman)
- tests: test/store_test: fix errors on the whole test suite run caused by the... ([pr#11427](#), Igor Fedotov)
- tests: test_subman.sh: Don't use -tmpdir ([pr#11384](#), Willem Jan Withagen)
- tests: test: test-erasure-eio.sh fix recovery testing and enable it ([pr#12170](#), David Zafman)
- tests: The default changed to disallow pool delete as of #11665; the tests assume it's allowed. ([pr#11897](#), Sage Weil)
- tests: Turn off tests again due to Jenkins failures ([pr#12217](#), David Zafman)
- tests: unittest_throttle avoid ASSERT_DEATH ([issue#18036](#), [pr#12393](#), Loic Dachary)
- tests: update rbd singleton/all/formatted-output.yaml to support ceph-ci ([issue#18440](#), [pr#12823](#), Nathan Cutler)
- tests: use shorter directories for tests ([issue#16014](#), [pr#12046](#), Loic Dachary)
- tests: vstart.sh: fix bashism in the script ([pr#11889](#), Mykola Golub)
- tests: workunits/ceph-helpers.sh: FreeBSD returns a different errorstring. ([pr#12005](#), Willem Jan Withagen)
- tools: Adding ceph-lazy tool ([pr#11055](#), gcharot)
- tools: ceph-create-keys should not try forever to do things ([issue#17753](#), [issue#12649](#), [issue#16255](#), [pr#11749](#), Alfredo Deza)
- tools: ceph_detect_init: add support for Alpine ([pr#8316](#), John Coyle)
- tools: ceph-disk: fix flake8 errors ([issue#17898](#), [pr#11973](#), Ken Dreyer)
- tools: ceph-disk: prevent unnecessary tracebacks from subprocess.check_call ([issue#16125](#), [pr#12414](#), Alfredo Deza)
- tools: ceph-post-file: single command to upload a file to cephdrop ([pr#505](#), Dan Mick, Travis Rhoden)
- tools: cleanup phase of cephfs-data-scan ([pr#12337](#), Vishal Kanaujia)
- tools: osdmaptool: additional tests ([pr#1196](#), Sage Weil)
- tools: osdmaptool: fix divide by zero error ([pr#12561](#), Yunchuan Wen)
- tools: rados: fix segfaults when run without -pool ([issue#17684](#), [pr#11633](#), David Disseldorp)
- tools: rados: optionally support reading omap key from file ([issue#18123](#), [pr#12286](#), Jason Dillaman)
- tools: script/run-coverity: update ([pr#12162](#), Sage Weil)
- tools: script/sephia_bt.sh: a script to prepare for debugging on [teuthology@sephia](#) ([pr#12012](#), Kefu Chai)
- tools: src/vstart.sh: Only execute btrfs if it is available ([pr#11683](#), Willem Jan Withagen)
- tools: tools/ceph-monstore-update-crush.sh: FreeBSD getopt is not compatible... ([pr#11525](#), Willem Jan Withagen)

V11.0.2 KRAKEN

This development checkpoint release includes a lot of changes and improvements to Kraken. This is the first release introducing ceph-mgr, a new daemon which provides additional monitoring & interfaces to external monitoring/management systems. There are also many improvements to bluestore, RGW introduces sync modules, copy part for multipart uploads and metadata search via elastic search as a tech preview.

NOTABLE CHANGES

- bluestore: os/bluestore: misc fixes ([pr#10953](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: do not op_file_update deleted files ([pr#10686](#), Sage Weil)
- bluestore: bluestore/BitAllocator: Fix deadlock with musl libc ([pr#10634](#), John Coyle)
- bluestore: bluestore/BlueFS: revert direct IO for WRITER_WAL ([pr#11059](#), Mark Nelson)
- bluestore: ceph-disk: support creating block.db and block.wal with customized size for bluestore ([pr#10135](#), Zhi Zhang)
- bluestore: compressor/zlib: switch to raw deflate ([pr#11122](#), Piotr Dałek)
- bluestore: do not use freelist to track bluefs_extents ([pr#10698](#), Sage Weil)
- bluestore: initialize csum_order properly ([pr#10728](#), xie xingguo)
- bluestore: kv/rocksdb: dump transactions on error ([pr#11042](#), Somnath Roy)
- bluestore: kv: In memory keyvalue db implementation ([pr#9933](#), Ramesh Chander)
- bluestore: os/bluestore/BitAllocator: batch is_allocated bit checks ([pr#10704](#), Ramesh Chander)
- bluestore: os/bluestore/BlueFS: For logs of rocksdb & bluefs only use directio. ([pr#11012](#), Jianpeng Ma)
- bluestore: os/bluestore/BlueFS: async compaction ([pr#10717](#), Varada Kari, Sage Weil)
- bluestore: os/bluestore/BlueFS: do not hold internal lock while waiting for IO ([pr#9898](#), Varada Kari, Sage Weil)

- bluestore: os/bluestore/BlueFS: do not start racing async compaction ([pr#11010](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: don't inc l_bluefs_files_written_wal if overwrite. ([pr#10143](#), Jianpeng Ma)
- bluestore: os/bluestore/BlueFS: factor unflushed log into runway calculation ([pr#10966](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: fix async compaction logging bug ([pr#10964](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: log dirty files at sync time ([pr#11108](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: only extend extent on same bdev ([pr#11023](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: prevent concurrent async compaction ([pr#11095](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: release completed aios ([pr#11268](#), Sage Weil)
- bluestore: os/bluestore/BlueFS: use StupidAllocator; fix async compaction bug ([pr#11087](#), Sage Weil)
- bluestore: os/bluestore/bluefs: add file refs check ([pr#10863](#), xie xingguo)
- bluestore: os/bluestore/bluefs: use map to track dirty files ([pr#10923](#), xie xingguo)
- bluestore: os/bluestore/bluefs_types: fix extent operator<< ([pr#10685](#), Sage Weil)
- bluestore: os/bluestore/bluestore_types: uint64_t for ref_map ([pr#11267](#), Sage Weil)
- bluestore: os/bluestore: Hint based allocation in bitmap Allocator ([pr#10978](#), Ramesh Chander)
- bluestore: os/bluestore: Remove bit alloc Woverloaded-virtual warnings ([pr#10082](#), Ramesh Chander)
- bluestore: os/bluestore: a few cleanups ([pr#11192](#), xie xingguo)
- bluestore: os/bluestore: a few fixes about the global csum setting ([pr#11195](#), xie xingguo)
- bluestore: os/bluestore: add assert to compress_extent_map ([pr#11240](#), Sage Weil)
- bluestore: os/bluestore: add cache-related stats ([pr#10961](#), xie xingguo)
- bluestore: os/bluestore: add checks and kill unreachable code ([pr#11077](#), xie xingguo)
- bluestore: os/bluestore: add error injection ([pr#11151](#), Sage Weil)
- bluestore: os/bluestore: add max blob size; fix compressed min blob size logic ([pr#11239](#), Sage Weil)
- bluestore: os/bluestore: add multiple finishers to bluestore ([pr#10780](#), llsoo Byun)
- bluestore: os/bluestore: add perf counters for compression effectiveness and space utilization measurements ([pr#10449](#), Igor Fedotov)
- bluestore: os/bluestore: apply "small encoding" for onode_t::extents map ([pr#10018](#), Igor Fedotov)
- bluestore: os/bluestore: avoid blob_t reencode when unchanged ([pr#10768](#), Sage Weil)
- bluestore: os/bluestore: binary search specified shard ([pr#11245](#), xie xingguo)
- bluestore: os/bluestore: change algorithm of compression header from string to int ([pr#10137](#), xie xingguo)
- bluestore: os/bluestore: compaction fixes ([pr#11279](#), Sage Weil)
- bluestore: os/bluestore: drop redundant call of get_blob ([pr#11275](#), xie xingguo)
- bluestore: os/bluestore: drop unreferenced spanning blobs ([pr#11212](#), Sage Weil)
- bluestore: os/bluestore: fix a few leaks ([pr#11068](#), Sage Weil)
- bluestore: os/bluestore: fix a few memory utilization leaks and wasters ([pr#11011](#), Sage Weil)
- bluestore: os/bluestore: fix crash in decode_some() ([pr#11312](#), Sage Weil)
- bluestore: os/bluestore: fix decoding hash of bnode ([pr#10773](#), xie xingguo)
- bluestore: os/bluestore: fix fsck() won't catch stray shard sometimes ([pr#11219](#), xie xingguo)
- bluestore: os/bluestore: fix gc when blob extends past eof ([pr#11282](#), Sage Weil)
- bluestore: os/bluestore: fix improper local var variable in collection_list meth... ([pr#10680](#), Igor Fedotov)
- bluestore: os/bluestore: fix incorrect pool decoding of bnode ([pr#10117](#), xie xingguo)
- bluestore: os/bluestore: fix leak of result-checking of _fsck_check_extents ([pr#11040](#), xie xingguo)
- bluestore: os/bluestore: fix leaks in our use of rocksdb ([pr#11250](#), Sage Weil)
- bluestore: os/bluestore: fix memory leak during bit_alloc testing ([pr#9935](#), xie xingguo)
- bluestore: os/bluestore: fix offset bug in _do_write_small. ([pr#11030](#), amoxic)
- bluestore: os/bluestore: fix onode cache addition race ([pr#11300](#), Sage Weil)
- bluestore: os/bluestore: fix potential access violation ([pr#10362](#), xie xingguo)
- bluestore: os/bluestore: fix potential access violation during rename ([pr#11033](#), xie xingguo)
- bluestore: os/bluestore: fix shard_info::dump() ([pr#11061](#), xie xingguo)
- bluestore: os/bluestore: fix spanning blob leak from ~ExtentMap ([pr#11223](#), Somnath Roy)
- bluestore: os/bluestore: fix statfs tests ([pr#10910](#), Sage Weil)
- bluestore: os/bluestore: fix when block device is not a multiple of the block size ([pr#10844](#), Sage Weil)
- bluestore: os/bluestore: fix write_big counter and some more cleanups ([pr#11344](#), xie xingguo)
- bluestore: os/bluestore: fix/improve csum error message ([pr#10938](#), Sage Weil)
- bluestore: os/bluestore: garbage collect partially overlapped blobs ([pr#11232](#), Roushan Ali)
- bluestore: os/bluestore: get rid off "isa-l" type in ZLibCompressor ctor ([pr#10931](#), xie xingguo)
- bluestore: os/bluestore: gifting bluefs more carefully ([pr#10950](#), xie xingguo)
- bluestore: os/bluestore: honour allow-eio flag; use global compressor if possible ([pr#10970](#), xie xingguo)
- bluestore: os/bluestore: improve required compression threshold ([pr#10080](#), xie xingguo)
- bluestore: os/bluestore: include bluefs space in statfs result ([pr#10795](#), Sage Weil)
- bluestore: os/bluestore: introduce power 2 macros for block alignment and rounding ([pr#10128](#), xie xingguo)
- bluestore: os/bluestore: make assert conditional with macro for allocator ([pr#11014](#), Ramesh Chander)
- bluestore: os/bluestore: make cache settings process-wide ([pr#11295](#), Sage Weil)
- bluestore: os/bluestore: make clone_range copy-on-write ([pr#11106](#), Sage Weil)
- bluestore: os/bluestore: make onode keys more efficient (and sort correctly) ([pr#11009](#), xie xingguo, Sage Weil)
- bluestore: os/bluestore: make trim() of 2Q cache more fine-grained ([pr#9946](#), xie xingguo)

- bluestore: os/bluestore: make zone/span size of bitmap-allocator configurable (pr#10040, xie xingguo)
- bluestore: os/bluestore: misc cleanup and test fixes (pr#11346, Igor Fedotov)
- bluestore: os/bluestore: misc cleanups (pr#10201, xie xingguo)
- bluestore: os/bluestore: misc cleanups (pr#11197, Haomai Wang)
- bluestore: os/bluestore: misc fixes (pr#9999, xie xingguo)
- bluestore: os/bluestore: misc fixes (pr#10771, xie xingguo)
- bluestore: os/bluestore: misc. fixes (pr#11129, xie xingguo)
- bluestore: os/bluestore: more cleanups (pr#11235, xie xingguo)
- bluestore: os/bluestore: more cleanups and fixes (pr#11210, xie xingguo)
- bluestore: os/bluestore: narrow condition of sanity check when get_object_key() (pr#11149, xie xingguo)
- bluestore: os/bluestore: narrow lock scope for cache trim() (pr#10410, xie xingguo)
- bluestore: os/bluestore: optimize intrusive sets for size. (pr#11319, Mark Nelson)
- bluestore: os/bluestore: pack a few more in-memory types (pr#11328, Sage Weil)
- bluestore: os/bluestore: precondition rocksdb/bluefs during mkfs (pr#10814, Sage Weil)
- bluestore: os/bluestore: prevent extent merging across shard boundaries (pr#11216, Sage Weil)
- bluestore: os/bluestore: print bluefs_extents in hex (pr#10689, Sage Weil)
- bluestore: os/bluestore: proper handling for csum enable/disable settings (pr#10431, Igor Fedotov)
- bluestore: os/bluestore: refactor dirty blob tracking along with some related fixes (pr#10215, Igor Fedotov)
- bluestore: os/bluestore: remove cmake warning from extent alloc functions (issue#16766, pr#10492, Ramesh Chander)
- bluestore: os/bluestore: remove deferred_csum machinery (pr#11243, Sage Weil)
- bluestore: os/bluestore: remove some copy-pastes (pr#11017, Igor Fedotov)
- bluestore: os/bluestore: replace store with logger in Cache (pr#10969, xie xingguo)
- bluestore: os/bluestore: shard extent map (pr#10963, Sage Weil)
- bluestore: os/bluestore: simplify LRU Cache::trim() (pr#10109, xie xingguo)
- bluestore: os/bluestore: simplify calculation of collection key range (pr#11166, xie xingguo)
- bluestore: os/bluestore: sloppy reshards boundaries to avoid spanning blobs (pr#11263, Sage Weil)
- bluestore: os/bluestore: still more cleanups (pr#11274, xie xingguo)
- bluestore: os/bluestore: switch spanning_blob_map to std::map (pr#11336, Sage Weil)
- bluestore: os/bluestore: trim cache on reads (pr#10095, Sage Weil)
- bluestore: os/bluestore: try to split blobs instead of spanning them (pr#11264, Sage Weil)
- bluestore: os/bluestore: upgrade compression settings to atomics (pr#11244, xie xingguo)
- bluestore: os/bluestore: use small encoding for bluefs extent and fnode (pr#10375, xie xingguo)
- bluestore: os/bluestore: yet another statfs test fix (pr#10926, Igor Fedotov)
- bluestore: os/bluestore: Fix size calculation in bitallocator (pr#10377, Ramesh Chander)
- bluestore: os/bluestore: fix error handling of posix_fallocate() (pr#10277, xie xingguo)
- bluestore: os/bluestore: use BE for gifting and reclaiming from bluefs (pr#10294, xie xingguo)
- bluestore: os/bluestore: get rid off blob's ref_map for non-shared objects (pr#9988, Igor Fedotov)
- bluestore: kv/MemDB: fix wrong output target and add sanity checks (pr#10358, xie xingguo)
- bluestore: os/bluestore: add a boundary check of cache read (pr#10349, xie xingguo)
- bluestore: os/bluestore: fix bitmap allocating failure if max_alloc_size is 0 (pr#10379, xie xingguo)
- bluestore: os/bluestore: misc fixes (pr#10327, xie xingguo)
- bluestore: kv/MemDB: misc fixes and cleanups (pr#10295, xie xingguo)
- bluestore: rocksdb: pull up to master (4.12 + a few patches) (pr#11069, Sage Weil)
- bluestore: test/store_test: extend Bluestore compression test to verify compress... (pr#11080, Igor Fedotov)
- bluestore: test/store_test: fix statfs results check to consider SSD min_alloc_size (pr#11096, Igor Fedotov)
- bluestore: unittest_bluestore_types: a few more types for sizeof (pr#11323, Sage Weil)
- bluestore: ceph_test_objectstore: test clone_range and fix a few bugs (pr#11103, Sage Weil)
- bluestore: kv: fix some bugs in memdb (pr#10550, Haodong Tang)
- bluestore: os/bluestore/BlueFS: disable buffered io (pr#10766, Sage Weil)
- build/ops,bluestore: test/objectstore/CMakeLists.txt: fix libaio conditional (pr#11008, Sage Weil)
- build/ops,cephfs: client: added def for ACCESSPERMS when undefined (pr#9835, John Coyle)
- build/ops,cephfs: deb: merge ceph-fs-common into ceph-common (issue#16808, pr#10433, Nathan Cutler)
- build/ops,cephfs: man/Makefile-client.am: drop legacy cephfs tool (pr#10444, Nathan Cutler)
- build/ops,cephfs: test: break out librados-using cephfs test (issue#16556, pr#10452, John Spray)
- build/ops,common: common/dns_resolve: use ns_name_uncompress instead of ns_name_ntop (pr#9755, John Coyle)
- build/ops,common: msg/async/net_handler.cc: make it more compatible with BSDs (pr#10029, Willem Jan Withagen)
- build/ops,pybind: Include Python 3 bindings into the cmake build and make packages for them (pr#10208, Oleh Prypin)
- build/ops,rbd: systemd: add install section to rbdmap.service file (pr#10942, Jelle vd Kooij)
- build/ops,rbd: test: fix rbd-mirror workunit test cases for cmake (pr#10076, Jason Dillaman)
- build/ops,rgw: rgw-ldap: add ldap lib to rgw lib deps based on build config (pr#9852, John Coyle)
- build/ops:.gitignore: Add .pyc files globally (pr#11076, Brad Hubbard)
- build/ops: Allow compressor build without YASM (pr#10937, Daniel Gryniewicz)
- build/ops: CMake - stop pip checking for updates (pr#10161, Daniel Gryniewicz)
- build/ops: CMakeList.txt: link ceph_objectstore_tool against fuse only if WITH_FUSE (pr#10149, Willem Jan Withagen)
- build/ops: Cmake: fix using CMAKE_DL_LIBS instead of dl (pr#10317, Willem Jan Withagen)

- build/ops: CmakeLists.txt: use LIB_RESOLV instead of resolv. ([pr#10972](#), Willem Jan Withagen)
- build/ops: Enable builds without ceph-test subpackage ([issue#16776](#), [pr#10872](#), Ricardo Dias)
- build/ops: Fix libatomic_ops-devel in SUSE and specfile cleanup ([issue#16645](#), [pr#10363](#), Nathan Cutler)
- build/ops: FreeBSD: Define CLOCK_REALTIME_COARSE in compat.h ([pr#10506](#), Willem Jan Withagen)
- build/ops: Gentoo support for ceph-disk / ceph-detect-init; pip speedup ([pr#8317](#), Robin H. Johnson)
- build/ops: LTTng-UST disabled for openSUSE ([issue#16937](#), [pr#10592](#), Michel Normand)
- build/ops: Port ceph-brag to Python 3 (+ small fixes) ([pr#10064](#), Oleh Prypin)
- build/ops: Removes remaining reference to WITH_MDS ([pr#10286](#), J. Eric Ivancich)
- build/ops: Stop hiding errors from run-tox.sh ([issue#17267](#), [pr#11071](#), Dan Mick)
- build/ops: Wip kill warnings ([pr#10881](#), Kefu Chai)
- build/ops: autogen: Fix rocksdb error when make dist ([pr#10988](#), tianqing)
- build/ops: autotools: remove a few other remaining traces ([pr#11019](#), Sage Weil)
- build/ops: build scripts: Enable dnf for Fedora >= 22 ([pr#11105](#), Brad Hubbard)
- build/ops: build: drop dryrun of autogen.sh from run-cmake-check.sh script ([pr#11013](#), xie xingguo)
- build/ops: ceph-disk tests: Let missing python interpreters be non-fatal ([pr#11072](#), Dan Mick)
- build/ops: ceph-disk: Compatibility fixes for Python 3 ([pr#9936](#), Anirudha Bose)
- build/ops: ceph-disk: do not activate device that is not ready ([issue#15990](#), [pr#9943](#), Boris Ranto)
- build/ops: ceph-osd-prestart.sh: check existence of OSD data directory ([issue#17091](#), [pr#10809](#), Nathan Cutler)
- build/ops: ceph-osd-prestart.sh: drop Upstart-specific code ([issue#15984](#), [pr#9667](#), Nathan Cutler)
- build/ops: ceph-post-file replace DSA with RSA ssh key ([issue#14267](#), [pr#10800](#), David Galloway)
- build/ops: ceph.spec.in: don't try to package __pycache__ for SUSE ([issue#17106](#), [pr#10805](#), Tim Serong)
- build/ops: ceph.spec.in: fix rpm package building error ([pr#10115](#), runsisi)
- build/ops: changes for Clang and yasm ([pr#10417](#), Willem Jan Withagen)
- build/ops: cmake changes ([pr#10351](#), Kefu Chai)
- build/ops: cmake changes ([pr#10059](#), Kefu Chai)
- build/ops: cmake changes ([pr#10279](#), Kefu Chai)
- build/ops: cmake changes ([issue#16804](#), [pr#10391](#), Kefu Chai)
- build/ops: cmake changes ([pr#10361](#), Kefu Chai)
- build/ops: cmake changes ([pr#10112](#), Kefu Chai)
- build/ops: cmake changes ([pr#10489](#), Kefu Chai)
- build/ops: cmake changes ([pr#10283](#), Kefu Chai)
- build/ops: cmake changes ([issue#16504](#), [pr#9995](#), Kefu Chai, Sage Weil, Dan Mick)
- build/ops: cmake changes ([pr#9975](#), Kefu Chai)
- build/ops: cmake changes related to LTTng-UST ([pr#10917](#), Kefu Chai)
- build/ops: common/compressor: add libcommon as a dependency for zlib and snappy p... ([pr#11083](#), Igor Fedotov)
- build/ops: compat: add abstractions for non portable pthread name funcs ([pr#9763](#), John Coyle)
- build/ops: configure.ac: Use uname instead of arch. ([pr#9766](#), John Coyle)
- build/ops: configure.ac: add _LIBS variables for boost_system and boost_iostreams ([pr#9848](#), John Coyle)
- build/ops: configure.ac: fix res_query detection ([pr#9820](#), John Coyle)
- build/ops: debian and cmake cleanups ([pr#10788](#), Kefu Chai)
- build/ops: debian: bump compat to 9 ([issue#16744](#), [pr#10366](#), Kefu Chai)
- build/ops: debian: python related changes ([pr#10322](#), Kefu Chai)
- build/ops: debian: replace SysV rbdmap with systemd service ([pr#10435](#), Ken Dreyer)
- build/ops: debian: set libexec dir to correct value as autotools did ([pr#10096](#), Daniel Gryniewicz)
- build/ops: do_cmake.sh: set up initial plugin dir ([pr#10067](#), Sage Weil)
- build/ops: fix /etc/os-release parsing in install-deps.sh ([pr#10981](#), Nathan Cutler)
- build/ops: fix the rpm build for centos ([pr#10289](#), Oleh Prypin, Josh Durgin)
- build/ops: force Python 3 packages to build in SUSE ([issue#17106](#), [pr#10894](#), Dominique Leuenberger, Nathan Cutler)
- build/ops: install-deps.sh based on /etc/os-release ([issue#16522](#), [pr#10017](#), Jan Fajerski)
- build/ops: install-deps: exit non-zero when we cannot match distro ([pr#10941](#), Gregory Meno)
- build/ops: isa-l: add isa-l library as a submodule ([pr#10066](#), Alyona Kiseleva)
- build/ops: jerasure: include generic objects in neon jerasure lib (like sse3/4) ([pr#10879](#), Dan Mick)
- build/ops: logrotate: Run as root/ceph ([pr#10587](#), Boris Ranto)
- build/ops: lttng: build the tracepoint provider lib from .c files in repo ([pr#11196](#), Kefu Chai)
- build/ops: make-dist: generate ceph.spec ([issue#16501](#), [pr#9986](#), Sage Weil)
- build/ops: make-dist: set rpm_release correctly for release builds ([pr#11334](#), Dan Mick)
- build/ops: make-srpm.sh: A simple script to make the srpm for ceph. ([pr#11064](#), Ira Cooper)
- build/ops: makefile: change librgw_file_* as check_PROGRAMS ([issue#16646](#), [pr#10229](#), Brad Hubbard)
- build/ops: remove autotools ([pr#11007](#), Sage Weil)
- build/ops: rpm: Do not start targets on update ([pr#9968](#), Nathan Cutler, Boris Ranto)
- build/ops: rpm: ExclusiveArch for suse_version ([issue#16936](#), [pr#10594](#), Michel Normand)
- build/ops: rpm: Fix creation of mount.ceph symbolic link for SUSE distros ([pr#10353](#), Ricardo Dias)
- build/ops: rpm: add udev BuildRequires to provide /usr/lib/udev directory ([issue#16949](#), [pr#10608](#), Nathan Cutler)
- build/ops: rpm: build rpm with cmake ([pr#10016](#), Kefu Chai)
- build/ops: rpm: drop obsolete libs-compat and python-ceph-compat metapackages ([issue#16353](#), [pr#9757](#), Nathan

Cutler)

- build/ops: rpm: fix permissions for /etc/ceph/rbdmap ([issue#17395](#), [pr#11217](#), Ken Dreyer)
- build/ops: rpm: fix shared library devel package names and dependencies ([issue#16345](#), [issue#16346](#), [pr#9744](#), Nathan Cutler, Ken Dreyer)
- build/ops: rpm: move mount.ceph from ceph-base to ceph-common and add symlink in /sbin for SUSE ([issue#16598](#), [pr#10147](#), Nathan Cutler)
- build/ops: run-cmake-check.sh: Remove redundant calls ([pr#11116](#), Brad Hubbard)
- build/ops: script: improve ceph-release-notes regex ([pr#10729](#), Nathan Cutler)
- build/ops: src/CMakeLists.txt: remove double flag -Wno-invalid-offsetof ([pr#10443](#), Willem Jan Withagen)
- build/ops: src/CMakeLists.txt: remove unneeded libraries from ceph-dencoder target ([pr#10478](#), Willem Jan Withagen)
- build/ops: src/global/pidfile.cc: Assign elements in structures individually ([pr#10516](#), Willem Jan Withagen)
- build/ops: src/kv/CMakeLists.txt: force rocksdb/include to first include directory ([pr#11194](#), Willem Jan Withagen)
- build/ops: test/common/test_util.cc: FreeBSD does not have distro information ([pr#10547](#), Willem Jan Withagen)
- build/ops: test: make check using cmake ([pr#10116](#), Kefu Chai, Sage Weil)
- build/ops: verified f23 ([pr#10222](#), Kefu Chai)
- build/ops: yasm-wrapper: dont echo the yasm command line ([pr#10819](#), Casey Bodley)
- build/ops: .gitignore: exclude core dumps, logfiles and temporary testresults ([pr#8150](#), Willem Jan Withagen)
- build/ops: this fixes the broken build ([pr#9992](#), Haomai Wang)
- build/ops: mrgw: search for cmake build dir. ([pr#10180](#), Abhishek Lekshmanan)
- build/ops: mrun, mstart.sh, mstop.sh: search for cmake build directory ([pr#10097](#), Yehuda Sadeh)
- build/ops: arm64 fixes([pr#10438](#), Dan Mick)
- build/ops: Wip kill warnings ([pr#10934](#), Kefu Chai)
- build/ops: systemd: add osd id to service description ([pr#10091](#), Ruben Kerkhof)
- build/ops: fix wrong indent caused compile warning ([pr#10014](#), Wanlong Gao)
- build/ops: ceph-detect-init: fix the py3 test ([pr#10266](#), Kefu Chai)
- build/ops: ceph.spec: fix ceph-mgr version requirement ([pr#11285](#), Sage Weil)
- build/ops: make-dist/ceph.spec.in: Fix srpm build breakage. ([pr#10404](#), Ira Cooper)
- build/ops: master: remove SYSTEMD_RUN from initscript ([issue#16440](#), [issue#7627](#), [pr#9871](#), Vladislav Odintsov)
- build/ops: rocksdb: revert the change introduced by dc41731 ([pr#10595](#), Kefu Chai)
- build/ops: do_freebsd*.sh: rename do_freebsd-cmake.sh to do_freebsd.sh ([pr#11088](#), Kefu Chai)
- build/ops: gcc 6.1.1 complains about missing include: <random>. 4.8.3 does not c... ([pr#10747](#), Daniel Oliveira)
- build/ops: selinux: Allow ceph to manage tmp files ([issue#17436](#), [pr#11259](#), Boris Ranto)
- build/ops: selinux: allow read /proc/<pid>/cmdline ([issue#16675](#), [pr#10339](#), Kefu Chai)
- cephfs,common: osdc/Journaler: move C_DelayFlush class to .cc ([pr#10744](#), Michal Jarzabek)
- cephfs,core,rbd: ObjectCacher: fix bh_read_finish offset logic ([issue#16002](#), [pr#9606](#), Greg Farnum)
- cephfs,core,rbd: osdc/ObjectCacher: move C_ReadFinish, C_RetryRead ([pr#10781](#), Michal Jarzabek)
- cephfs: Add ceph_ll_setlk and ceph_ll_getlk ([pr#9566](#), Frank S. Filz)
- cephfs: CephFS: misc. cleanups and remove legacy cephfs tool ([issue#16195](#), [issue#16035](#), [issue#15923](#), [pr#10243](#), John Spray)
- cephfs: Clean up handling of "../" in ceph client ([pr#10691](#), Jeff Layton)
- cephfs: Client: fixup param type and return value ([pr#10463](#), gongchuang)
- cephfs: Client: pass "UserPerm" struct everywhere for security checks ([issue#16367](#), [issue#17368](#), [pr#11218](#), Greg Farnum)
- cephfs: First pile of statx patches ([pr#10922](#), Sage Weil, Jeff Layton)
- cephfs: Fix attribute handling at lookup time ([issue#16668](#), [pr#10386](#), Jeff Layton)
- cephfs: Inotable repair during forward scrub ([pr#10281](#), Vishal Kanaujia)
- cephfs: Server: drop locks and auth pins if wait for pending truncate ([pr#9716](#), xie xingguo)
- cephfs: Small interface cleanups for struct ceph_statx ([pr#11093](#), Jeff Layton)
- cephfs: build ceph-fuse on OSX ([pr#9371](#), Yan, Zheng)
- cephfs: ceph-fuse: link to libtcmalloc or jemalloc ([issue#16655](#), [pr#10258](#), Yan, Zheng)
- cephfs: ceph_volume_client: store authentication metadata ([issue#15406](#), [issue#15615](#), [pr#9864](#), John Spray, Ramana Raja)
- cephfs: client/barrier: move C_Block_Sync class to .cc ([pr#11001](#), Michal Jarzabek)
- cephfs: client/filer: cleanup the redundant judgments of _write&&_fallback ([pr#10062](#), huanwen ren)
- cephfs: client: add missing client_lock for get_root ([pr#10027](#), Patrick Donnelly)
- cephfs: client: discard mds map if it is identical to ours ([pr#9774](#), xie xingguo)
- cephfs: client: fast abort if underlying statsf() call failed; end scope of std::hex properly ([pr#9803](#), xie xingguo)
- cephfs: client: fix access violation ([pr#9793](#), xie xingguo)
- cephfs: client: fix readdir vs fragmentation race ([issue#17286](#), [pr#11147](#), Yan, Zheng)
- cephfs: client: fix segment fault in Client::_invalidate_kernel_dcache(). ([issue#17253](#), [pr#11170](#), Yan, Zheng)
- cephfs: client: fix shutdown with open inodes ([issue#16764](#), [pr#10419](#), John Spray)
- cephfs: client: include COMPLETE and ORDERED states in cache dump ([pr#10485](#), Greg Farnum)
- cephfs: client: kill compiling warning ([pr#9994](#), xie xingguo)
- cephfs: client: misc fixes ([pr#9838](#), xie xingguo)
- cephfs: client: move Inode specific cleanup to destructor ([pr#10168](#), Patrick Donnelly)

- cephfs: client: note order of member init in cons ([pr#10169](#), Patrick Donnelly)
- cephfs: client: properly set inode number of created inode in replay request ([issue#17172](#), [pr#10957](#), Yan, Zheng)
- cephfs: client: protect InodeRef with client_lock ([issue#17392](#), [pr#11225](#), Yan, Zheng)
- cephfs: doc/mds: fixup mds doc ([pr#10573](#), huanwen ren)
- cephfs: fuse_ll: fix incorrect error settings of fuse_ll_mkdir() ([pr#9809](#), xie xingguo)
- cephfs: include/ceph_fs.h: guard `#define CEPH_SETATTR_*` with `#ifndef` ([pr#10265](#), Kefu Chai)
- cephfs: libcephfs: Fix the incorrect integer conversion in libcephfs_jni.cc ([pr#10640](#), wenjunhuang)
- cephfs: libcephfs: add umount function in cephfs.pyx ([pr#10774](#), huanwen ren)
- cephfs: libcephfs: fix portability-related error settings ([pr#9794](#), xie xingguo)
- cephfs: libcephfs: kill compiling warning ([pr#10622](#), xie xingguo)
- cephfs: mds/CDir: remove the part of judgment for `_next_dentry_on_set` ([pr#10476](#), zhang.zezhu)
- cephfs: mds/CInode: fix potential fin hanging ([pr#9773](#), xie xingguo)
- cephfs: mds/MDBalancer: cleanup ([pr#10512](#), huanwen ren)
- cephfs: mds/MDCache: kill a compiler warning ([pr#11254](#), xie xingguo)
- cephfs: mds/MDSSMap default metadata pool to -1 (was: output None instead of 0 when no fs present.) ([issue#16588](#), [pr#10202](#), Xiaoxi Chen)
- cephfs: mds/MDSTable: add const to member functions ([pr#10846](#), Michal Jarzabek)
- cephfs: mds/SessionMap.h: change statement to assertion ([pr#11289](#), Michal Jarzabek)
- cephfs: mds/SnapRealm.h: add const to member functions ([pr#10878](#), Michal Jarzabek)
- cephfs: mds/server: clean up handle_client_open() ([pr#11120](#), huanwen ren)
- cephfs: mon/MDSSMonitor: move C_Updated class to .cc file ([pr#10668](#), Michal Jarzabek)
- cephfs: osdc/mds: fixup pos parameter in the journaler ([pr#10200](#), huanwen ren)
- cephfs: reduce unnecessary mds log flush ([pr#10393](#), Yan, Zheng)
- cephfs: tools/cephfs: Remove cephfs-data-scan tmap_upgrade ([issue#16144](#), [pr#10100](#), Douglas Fuller)
- cephfs: ceph_fuse: use sizeof get the buf length ([pr#11176](#), LeoZhang)
- cli: retry when the mon is not configured ([issue#16477](#), [pr#11089](#), Loic Dachary)
- cmake: Add -pie to CMAKE_EXE_LINKER_FLAGS ([pr#10755](#), Tim Serong)
- cmake: Fix FCGI include directory ([pr#9983](#), Tim Serong)
- cmake: Fix mismatched librgw VERSION / SOVERSION ([pr#10754](#), Tim Serong)
- cmake: FreeBSD specific excludes in CMakeLists.txt ([pr#10973](#), Willem Jan Withagen)
- cmake: FreeBSD specific excludes in CMakeLists.txt files ([pr#10517](#), Willem Jan Withagen)
- cmake: Really add FCGI_INCLUDE_DIR to include_directories for rgw ([pr#10139](#), Tim Serong)
- cmake: Removed README.cmake.md, edited README.md ([pr#10028](#), Ali Maredia)
- cmake: Support tcmalloc_minimal allocator ([pr#11111](#), Bassam Tabbara)
- cmake: add dependency from ceph_smallibenchrbd to cls libraries ([pr#10870](#), J. Eric Ivancich)
- cmake: add_subdirectory(include) ([pr#10360](#), Kefu Chai)
- cmake: ceph_test_rbd_mirror does not require librados_test_stub ([pr#10164](#), Jason Dillaman)
- cmake: cleanup Findgperftools.cmake ([pr#10670](#), Kefu Chai)
- cmake: correct ceph_test_librbd/ceph_test_rbd_mirror linkage ([issue#16882](#), [pr#10598](#), Jason Dillaman)
- cmake: disable -fvar-tracking-assignments for ceph_dencoder.cc ([pr#10275](#), Kefu Chai)
- cmake: disable unittest_async_compressor ([pr#10394](#), Kefu Chai)
- cmake: do not link against unused objects or libraries ([pr#10837](#), Kefu Chai)
- cmake: enable ccache for rocksdb too ([pr#11100](#), Bassam Tabbara)
- cmake: exclude non-public symbols in shared libraries ([issue#16556](#), [pr#10472](#), Kefu Chai)
- cmake: fix incorrect dependencies to librados ([pr#10145](#), Jason Dillaman)
- cmake: fix the FTBFS introduced by dc8b3ba ([pr#10282](#), Kefu Chai)
- cmake: fix the build of unittest_async_compressor ([pr#10400](#), Kefu Chai)
- cmake: fix the tracing header dependencies ([pr#10906](#), Kefu Chai)
- cmake: fix unittest_rbd_mirror failures under non-optimized builds ([pr#9990](#), Jason Dillaman)
- cmake: fix wrong path introduced by bb163e9 ([pr#10643](#), Kefu Chai)
- cmake: fixes ([pr#10092](#), Daniel Grynewicz)
- cmake: fixes for pypi changes ([pr#10204](#), Kefu Chai)
- cmake: include(SIMDExt) in src/CMakeLists.txt ([pr#11003](#), Kefu Chai)
- cmake: install ceph_test_cls_rgw ([pr#10025](#), Kefu Chai)
- cmake: install ceph_test_rados_striper_api_* ([pr#10541](#), Kefu Chai)
- cmake: install platlib into a subdir of build-base dir ([pr#10666](#), Kefu Chai)
- cmake: make py3 a nice-to-have ([issue#17103](#), [pr#11015](#), Kefu Chai)
- cmake: pass -DINTEL* to gf-complete cflags ([pr#10956](#), tone.zhang, Kefu Chai)
- cmake: pass cmake's compiler and flags to compile RocksDB into build ([pr#10418](#), Willem Jan Withagen)
- cmake: recompile erasure src for different variants ([pr#10772](#), Kefu Chai)
- cmake: remove WITH_MDS option ([pr#10186](#), Ali Maredia)
- cmake: remove more autotools hacks ([pr#11229](#), Sage Weil)
- cmake: remove unnecessary linked libs from libcephfs ([issue#16556](#), [pr#10081](#), Kefu Chai)
- cmake: rework NSS and SSL ([pr#9831](#), Matt Benjamin)
- cmake: set ARM_CRC_FLAGS from the CRC test rather than ARM_NEON_FLAGS ([issue#17250](#), [pr#11028](#), Dan Mick)

- cmake: specify distutils build path explicitly ([pr#10568](#), Kefu Chai)
- cmake: supress more warnings ([pr#10469](#), Willem Jan Withagen)
- cmake: use PERF_LOCAL_FLAGS only if defined ([issue#17104](#), [pr#10828](#), Michel Normand)
- cmake: use stock Find* modules. ([pr#10178](#), Kefu Chai)
- cmake: work to get initial FreeBSD stuff ([pr#10352](#), Willem Jan Withagen)
- cmake: find GIT_VER variables if there is no .git dir ([pr#11499](#), Ali Maredia)
- common,bluestore: Isa-I extention for zlib compression plugin ([pr#10158](#), Alyona Kiseleva, Dan Mick)
- common,bluestore: compressor/zlib: zlib wrapper fix ([pr#11079](#), Igor Fedotov)
- common: auth/cephx: misc fixes ([pr#9679](#), xie xingguo)
- common: common/PluginRegistry: improve error output for shared library load fa... ([pr#11081](#), Igor Fedotov)
- common: common/Throttle.h: remove unneeded class ([pr#10902](#), Michal Jarzabek)
- common: common/Timer.h: delete copy constr and assign op ([pr#11046](#), Michal Jarzabek)
- common: common/WorkQueue: add std move ([pr#9729](#), Michal Jarzabek)
- common: compressor: zlib compressor plugin cleanup ([pr#9782](#), Alyona Kiseleva)
- common: erasure-code: Runtime detection of SIMD for jerasure and shec ([pr#11086](#), Bassam Tabbara)
- common: global: log which process/command sent a signal ([pr#8964](#), song baisen)
- common: include/assert: clean up ceph assertion macros ([pr#9969](#), Sage Weil)
- common: instantiate strict_si_cast<long> not strict_si_cast<int64_t> ([issue#16398](#), [pr#9934](#), Kefu Chai)
- common: lockdep: verbose even if no logging is set ([pr#10576](#), Willem Jan Withagen)
- common: messages/MOSDMap: mark as enlighten OSDMap encoder ([pr#10843](#), Sage Weil)
- common: mon/Monitor.cc: replace lock/unlock with Mutex:Lockr ([pr#9792](#), Michal Jarzabek)
- common: msg/AsyncMessenger.cc: remove code duplication ([pr#10030](#), Michal Jarzabek)
- common: msg/async: less verbose debug messages at debug_ms=1 ([pr#11205](#), Sage Weil)
- common: msg/async: remove static member variable ([issue#16686](#), [pr#10440](#), Kefu Chai)
- common: only call crypto::init once per CephContext ([issue#17205](#), [pr#10965](#), Casey Bodley)
- common: osdc/ObjectCacher: change iterator to const_iterator and add const to member functions ([pr#9644](#), Michal Jarzabek)
- common: preforker: prevent call to ‘write’ on an fd that was already closed ([pr#10949](#), Avner BenHanoch)
- common: remove basename() dependency ([pr#9845](#), John Coyle)
- common: src/common/buffer.cc fix judgment for lseek ([pr#10130](#), zhang.zezhu)
- common: unknown hash type of judgment modification ([pr#9510](#), huanwen ren)
- common: Timer.cc: replace long types with auto ([pr#11067](#), Michal Jarzabek)
- common: TrackedOp: move ShardedTrackingData to .cc ([pr#10639](#), Michal Jarzabek)
- common: config_opts: fix comment(radio -> ratio) ([pr#10783](#), xie xingguo)
- common: src/common/dns_resolve.cc: reorder the includes ([pr#10505](#), Willem Jan Withagen)
- common: global/signal_handler: use sig_str instead of sys_siglist ([pr#10633](#), John Coyle)
- core,cephfs: Revert “osd/ReplicatedPG: for sync-read it don’t cacl l_osd_op_r_prep...” ([issue#16908](#), [pr#10875](#), Samuel Just)
- core,cephfs: mon/mds: add err info when load_metadata is abnormal ([pr#10176](#), huanwen ren)
- core,common: osd/OSD.cc: remove unneeded returns ([pr#11043](#), Michal Jarzabek)
- core,pybind: python-rados: extends ReadOp/WriteOp API ([pr#9944](#), Mehdi Abaakouk)
- core,pybind: python-rados: implement new aio_stat. ([pr#11006](#), Iain Buclaw)
- core,pybind: qa/workunits/rados/test_python.sh: Allow specifying Python executable ([pr#10782](#), Oleh Prypin)
- core: os/filestore/LFNIIndex: remove unused variable ‘subdir_path’ ([pr#8959](#), huangjun)
- core: Create ceph-mgr ([pr#10328](#), John Spray, Tim Serong)
- core: FileJournal: Remove obsolete _check_disk_write_cache function ([pr#11073](#), Brad Hubbard)
- core: Lua object class support ([pr#7338](#), Noah Watkins)
- core: OSD crash with Hammer to Jewel Upgrade: void FileStore::init_temp_collections() ([issue#16672](#), [pr#10565](#), David Zafman)
- core: OSD.cc: remove unneeded return ([pr#9701](#), Michal Jarzabek)
- core: OSD: avoid FileStore finisher deadlock in osd_lock when shutdown OSD ([pr#11052](#), Haomai Wang)
- core: ObjectCacher: fix last_write check in bh_write_adjacencies() ([issue#16610](#), [pr#10304](#), Yan, Zheng)
- core: ReplicatedPG: call op_applied for submit_log_entries based repops ([pr#9489](#), Samuel Just)
- core: Wip 16998 ([issue#16998](#), [pr#10688](#), Samuel Just)
- core: ceph-create-keys: add missing argument comma ([pr#11123](#), Patrick Donnelly)
- core: ceph-create-keys: fix existing-but-different case ([issue#16255](#), [pr#10415](#), John Spray)
- core: ceph-disk: partprobe should block udev induced BLKRRPART ([issue#15176](#), [pr#9330](#), Marius Vollmer, Loic Dachary)
- core: ceph-disk: timeout ceph-disk to avoid blocking forever ([issue#16580](#), [pr#10262](#), Loic Dachary)
- core: ceph-objectstore-tool: add a way to split filestore directories offline ([issue#17220](#), [pr#10776](#), Josh Durgin)
- core: ceph.in: python 3 compatibility of the ceph CLI ([pr#9702](#), Oleh Prypin)
- core: ceph_mon: use readdir() as readdir_r() is deprecated ([pr#11047](#), Kefu Chai)
- core: cephx: Fix multiple segfaults due to attempts to encrypt or decrypt ([issue#16266](#), [pr#9703](#), Brad Hubbard)
- core: <https://github.com/ceph/ceph/pull/11052> ([pr#10371](#), Yan Jun)
- core: include write error codes in the pg log ([issue#14468](#), [pr#10170](#), Josh Durgin)
- core: kv/MemDB: fix assert triggered by m_total_bytes underflow ([pr#10471](#), xie xingguo)

- core: kv/RocksDB: add perfcounter for submit_transaction_sync operation ([pr#9770](#), Haodong Tang)
- core: logmon: check is_leader() before doing any work on get_trim_to() ([pr#10342](#), song baisen)
- core: memstore: clone zero-fills holes from source range ([pr#11157](#), Casey Bodley)
- core: message: optimization for message priority strategy ([pr#8687](#), yaoning)
- core: messages/MForward: fix encoding features ([issue#17365](#), [pr#11180](#), Sage Weil)
- core: mgr/MgrClient: fix ms_handle_reset ([pr#11298](#), Sage Weil)
- core: mgr/MgrMap: initialize all fields ([issue#17492](#), [pr#11308](#), Sage Weil)
- core: mon/ConfigKeyService: pass strings by const ref ([pr#10618](#), Michal Jarzabek)
- core: mon/LogMonitor: move C_Log struct to cc file ([pr#10721](#), Michal Jarzabek)
- core: mon/MonClient.h: pass strings by const reference ([pr#10605](#), Michal Jarzabek)
- core: mon/MonDBStore: fix assert which never fires ([pr#10706](#), xie xingguo)
- core: mon/MonitorDBStore: do not use snapshot iterator; close on close ([pr#10102](#), Sage Weil)
- core: mon/OSDMonitor.cc: remove use of boost assign ([pr#11060](#), Michal Jarzabek)
- core: mon/PGMonitor: batch filter pg states; add sanity check ([pr#9394](#), xie xingguo)
- core: mon/PGMonitor: calc the %USED of pool using used/(used+avail) ([issue#16933](#), [pr#10584](#), Kefu Chai)
- core: mon/PGMonitor: move C_Stats struct to cc file ([pr#10719](#), Michal Jarzabek)
- core: mon/PaxosService: make the return value type inconsistent ([pr#10231](#), zhang.zezhu)
- core: mon/osdmonitor: fix incorrect output of "osd df" due to osd out ([issue#16706](#), [pr#10308](#), xie xingguo)
- core: msg/AsyncMessenger: change return type to void ([pr#10230](#), Michal Jarzabek)
- core: msg/Messenger: add const and override to function ([pr#10183](#), Michal Jarzabek)
- core: msg/async/AsyncConnection: replace Mutex with std::mutex for performance ([issue#16714](#), [issue#16715](#), [pr#10340](#), Haomai Wang)
- core: msg/async/Event: ensure not refer to member variable which may destroyed ([issue#16714](#), [pr#10369](#), Haomai Wang)
- core: msg/async/kqueue: avoid remove nonexistent kqueue event ([pr#9869](#), Haomai Wang)
- core: msg/async: Support close idle connection feature ([issue#16366](#), [pr#9783](#), Haomai Wang)
- core: msg/async: allow other async backend implementations ([pr#10264](#), Haomai Wang)
- core: msg/async: avoid set out of range ms_async_op_threads option ([pr#11200](#), Haomai Wang)
- core: msg/async: connect authorizer fix + recv_buf size ([pr#9784](#), Ilya Dryomov)
- core: msg/async: harden error logic handle ([pr#9781](#), Haomai Wang)
- core: msg/async: remove fd output in log prefix ([pr#11199](#), Haomai Wang)
- core: msg/async: remove file event lock ([issue#16554](#), [issue#16552](#), [pr#10090](#), Haomai Wang)
- core: msg/simple/Pipe: eliminating casts for the comparing of len and recv_max_prefetch ([pr#10273](#), zhang.zezhu)
- core: msg/simple: fix wrong condition checking of writing TAG_CLOSE on closing ([pr#10343](#), xie xingguo)
- core: msg/simple: wait dispatch_queue until all pipes closed ([issue#16472](#), [pr#9930](#), Haomai Wang)
- core: msg: make async backend default ([pr#10746](#), Haomai Wang)
- core: msg: mark daemons down on RST + ECONNREFUSED ([pr#8558](#), Piotr Dałek)
- core: os/FuseStore: fix several FuseStore issues ([pr#10723](#), Sage Weil)
- core: os/MemStore: move BufferlistObject to .cc file ([pr#10833](#), Michal Jarzabek)
- core: os/ObjectStore: fix return code of collection_empty() method ([pr#11050](#), xie xingguo)
- core: os/RocksDBStore: use effective Get API instead of iterator api ([pr#9411](#), Jianjian Huo, Haomai Wang, Mark Nelson)
- core: os/filestore/FDCache: fix bug when filestore_fd_cache_shards = 0 ([pr#11048](#), jimifm)
- core: os/filestore/FileJournal: error out if FileJournal is not a file ([issue#17307](#), [pr#11146](#), Kefu Chai)
- core: os/filestore: add sanity checks and cleanups for mount() process ([pr#9734](#), xie xingguo)
- core: os/filestore: disable use of splice by default ([pr#11113](#), Haomai Wang)
- core: osd/OSD.cc: remove repeated searching of map ([pr#10986](#), Michal Jarzabek)
- core: osd/OSD.cc: remove unneeded searching of maps ([pr#11039](#), Michal Jarzabek)
- core: osd/OSD.h: add const to member functions ([pr#11114](#), Michal Jarzabek)
- core: osd/OSD.h: move some members under private ([pr#11121](#), Michal Jarzabek)
- core: osd/OSD.h: remove unneeded line ([pr#8980](#), Michal Jarzabek)
- core: osd/OSDMonitor: misc. cleanups ([pr#10739](#), xie xingguo)
- core: osd/OSDMonitor: misc. fixes ([pr#10491](#), xie xingguo)
- core: osd/ReplicatedBackend: add sanity check during build_push_op() ([pr#9491](#), Yan Jun)
- core: osd/ReplicatedPG: for sync-read it don't call l_osd_op_r_prepare_lat. ([pr#10365](#), Jianpeng Ma)
- core: osd/ReplicatedPG: remove class redeclaration ([pr#11041](#), Michal Jarzabek)
- core: osd/ReplicatedPG: remove unused param "op" from generate_subop() ([pr#10811](#), jimifm)
- core: osd/Watch: add consts to member functions ([pr#10251](#), Michal Jarzabek)
- core: osd/osd_type: check if pool is gone during check_new_interval() ([pr#10859](#), xie xingguo)
- core: osd/osdmonitor: pool of objects and bytes beyond quota should all be warn ([pr#9085](#), huanwen ren)
- core: osdc/objecter: misc fixes ([pr#10826](#), xie xingguo)
- core: pass string by const ref and add override to virtual function ([pr#9082](#), Michal Jarzabek)
- core: qa/workunits/objectstore/test_fuse.sh: make test_fuse.sh work with filestore ([pr#11057](#), Sage Weil)
- core: rados: add option to include clones when doing flush or evict ([pr#9698](#), Mingxin Liu)
- core: subman: use replace instead of format ([issue#16961](#), [pr#10620](#), Loic Dachary)
- core: test/common/Throttle.cc: fix race in shutdown ([pr#10094](#), Samuel Just)

- core: test: add the necessary judgment ([pr#9694](#), huanwen ren)
- core: tox.ini: remove extraneous coverage -omit option ([pr#10943](#), Josh Durgin)
- core: udev: always populate /dev/disk/by-parttypeuuid ([issue#16351](#), [pr#9885](#), Loic Dachary)
- core: os/FuseStore: remove unneeded header file ([pr#10799](#), Michal Jarzabek)
- core: os/MemStore: move OmapIteratorImpl to cc file ([pr#10803](#), Michal Jarzabek)
- core: os/Memstore.h: add override to virtual functions ([pr#10801](#), Michal Jarzabek)
- core: os/Memstore: move PageSetObject class to .cc file ([pr#10817](#), Michal Jarzabek)
- core: os/bluestore: remove unused head file. ([pr#11186](#), Jianpeng Ma)
- core: safe_io: Improve portability by replacing loff_t type usage with off_t. ([pr#9767](#), John Coyle)
- core: src/kv/MemDB.cc: the type of the parameter of push_back() does not match the ops's value_type ([pr#10455](#), Willem Jan Withagen)
- core: msg/simple: apply prefetch policy more precisely ([pr#10344](#), xie xingguo)
- core: CompatSet.h: remove unneeded inline ([pr#10071](#), Michal Jarzabek)
- core: Objclass perm feedback ([pr#10313](#), Noah Watkins)
- core: arch/arm.c: remove unnecessary variable read for simplicity ([pr#10821](#), Weibing Zhang)
- crush: don't normalize input of crush_In iteratively ([pr#10935](#), Piotr Dałek)
- crush: reset bucket->h.items[i] when removing tree item ([issue#16525](#), [pr#10093](#), Kefu Chai)
- crush: CrushCompiler.cc:884 ([pr#10952](#), xu biao)
- crush: CrushCompiler: error out as long as parse fails ([issue#17306](#), [pr#11144](#), Kefu Chai)
- doc: Add documentation about snapshots ([pr#10436](#), Greg Farnum)
- doc: Add two options to radosgw-admin.rst manpage ([issue#17281](#), [pr#11134](#), Thomas Serlin)
- doc: Changed config parameter "rgw keystone make new tenants" in radosgw multitenancy ([issue#17293](#), [pr#11127](#), SirishaGuduru)
- doc: Modification for "TEST S3 ACCESS" section in "INSTALL CEPH OBJECT GATEWAY" page ([pr#9089](#), la-sguduru)
- doc: Update developer docs for cmake paths ([pr#11163](#), John Spray)
- doc: add "--orphan-stale-secs" to radosgw-admin(8) ([issue#17280](#), [pr#11097](#), Ken Dreyer)
- doc: add \$pid metavar conf doc ([pr#11172](#), Patrick Donnelly)
- doc: add Backporting section to Essentials chapter ([issue#15497](#), [pr#10457](#), Nathan Cutler)
- doc: add Prepare tenant section to Testing in the cloud chapter ([pr#10413](#), Nathan Cutler)
- doc: add Upload logs to archive server section... ([pr#10414](#), Nathan Cutler)
- doc: add client config ref ([issue#16743](#), [pr#10434](#), Patrick Donnelly)
- doc: add graphic for cap bit field ([pr#10897](#), Patrick Donnelly)
- doc: add missing PR to hammer 0.94.8 release notes ([pr#10900](#), Nathan Cutler)
- doc: add openSUSE instructions to quick-start-preflight ([pr#10454](#), Nathan Cutler)
- doc: add rgw_enable_usage_log option in Rados Gateway admin guide ([issue#16604](#), [pr#10159](#), Mike Hackett)
- doc: add troubleshooting steps for ceph-fuse ([pr#10374](#), Ken Dreyer)
- doc: admin/build-doc: bypass sanity check if building doc ([issue#16940](#), [pr#10623](#), Kefu Chai)
- doc: ceph-authtool man page option is -print-key not -print ([pr#9731](#), Brad Hubbard)
- doc: ceph-deploy mon add doesn't take multiple nodes ([pr#10085](#), Chengwei Yang)
- doc: clarify rbd size units ([pr#11303](#), Ilya Dryomov)
- doc: cleanup outdated radosgw description ([pr#11248](#), Jiaying Ren)
- doc: describe libvirt client logging ([pr#10542](#), Ken Dreyer)
- doc: do not list all major versions in get-packages.rst ([pr#10899](#), Nathan Cutler)
- doc: doc/cephfs: explain the various health messages ([pr#10244](#), John Spray)
- doc: doc/dev: Fix missing code section due to no lexer for "none" ([pr#9083](#), Brad Hubbard)
- doc: doc/radosgw: fix description of response elements 'Part' ([pr#10641](#), weiqiaomiao)
- doc: doc/radosgw: rename config.rst to config-fcgi.rst ([pr#10381](#), Nathan Cutler)
- doc: extend the CephFS troubleshooting guide ([pr#10458](#), Greg Farnum)
- doc: fix broken link in SHEC erasure code plugin ([issue#16996](#), [pr#10675](#), Albert Tu)
- doc: fix description for rsize and rasize ([pr#11101](#), Andreas Gerstmayer)
- doc: fix rados/configuration/osd-config-ref.rst ([pr#10619](#), Chengwei Yang)
- doc: fix singleton example in Developer Guide ([pr#10830](#), Nathan Cutler)
- doc: fix some nits in release notes and releases table ([pr#10903](#), Nathan Cutler)
- doc: fix standby replay config ([issue#16664](#), [pr#10268](#), Patrick Donnelly)
- doc: fix wrong osdkeepalive name in mount.ceph manpage ([pr#10840](#), Zhi Zhang)
- doc: fix/add changelog for 10.2.2, 0.94.7, 0.94.8 ([pr#10895](#), Sage Weil)
- doc: format 2 now is the default image format ([pr#10705](#), Chengwei Yang)
- doc: lgmt (build verified f23) ([pr#9745](#), weiqiaomiao)
- doc: mailmap updates for upcoming 11.0.0 ([pr#9301](#), Yann Dupont)
- doc: manual instructions to set up mds daemon ([pr#11115](#), Peter Maloney)
- doc: missing "make vstart" in quick_guide.rst ([pr#11226](#), Leo Zhang)
- doc: more details for pool deletion ([pr#10190](#), Ken Dreyer)
- doc: peering.rst, fix typo ([pr#10131](#), Brad Hubbard)
- doc: perf_counters.rst fix trivial typo ([pr#10292](#), Brad Hubbard)
- doc: rbdmap: specify bash shell interpreter ([issue#16608](#), [pr#10733](#), Jason Dillaman)

- doc: release-notes.rst: draft 0.94.8 release notes ([pr#10730](#), Nathan Cutler)
- doc: remove btrfs contradiction ([pr#9758](#), Nathan Cutler)
- doc: remove i386 from minimal hardware recommendations ([pr#10276](#), Kefu Chai)
- doc: remove old references to inktank premium support ([pr#11182](#), Alfredo Deza)
- doc: remove the description of deleted options ([issue#17041](#), [pr#10741](#), MinSheng Lin)
- doc: rgw, doc: fix formatting around Keystone-related options. ([pr#10331](#), Radoslaw Zarzynski)
- doc: rgw/doc: fix indent ([pr#10676](#), Yan Jun)
- doc: rm SysV instructions, add systemd ([pr#10184](#), Ken Dreyer)
- doc: silence sphinx warnings ([pr#10621](#), Kefu Chai)
- doc: small standby doc edits ([pr#10479](#), Patrick Donnelly)
- doc: update CephFS “early adopters” info ([pr#10068](#), John Spray)
- doc: update canonical tarballs URL ([pr#9695](#), Ken Dreyer)
- doc: update rbd glance configuration notes ([pr#10629](#), Jason Dillaman)
- doc: update s3 static webiste feature support status ([pr#10223](#), Jiaying Ren)
- doc: changelog: add v10.2.3 ([pr#11238](#), Abhishek Lekshmanan)
- doc: install: Use <https://> for download.ceph.com ([pr#10709](#), Colin Walters)
- doc: release-notes: v0.94.9 ([pr#10927](#), Sage Weil)
- doc: release-notes: v10.2.3 jewel ([pr#11234](#), Abhishek Lekshmanan)
- doc: Add UK mirror and update copyright ([pr#10531](#), Patrick McGarry)
- doc: README.md: replace package build instructions with tarball instructions ([pr#10829](#), Sage Weil)
- doc: Removed reference about pool ownership based on BZ#1368528 ([pr#11063](#), Bara Ancincova)
- librados: use bufferlist instead of buffer::list in public header ([pr#10632](#), Ryne Li)
- librados: Rados-stripper: Flexible string matching for not found attributes ([pr#10577](#), Willem Jan Withagen)
- librados: librados examples: link and include from current source tree by default. ([issue#15100](#), [pr#8189](#), Jesse Williamson)
- librbd: API methods to directly acquire and release the exclusive lock ([issue#15632](#), [pr#9592](#), Mykola Golub)
- librbd: add consistency groups operations with images ([pr#10034](#), Victor Denisov)
- librbd: add explicit shrink check while resizing images ([pr#9878](#), Vaibhav Bhembre)
- librbd: asynchronous v2 image creation ([issue#15321](#), [pr#9585](#), Venky Shankar)
- librbd: backward/forward compatibility for update_features ([issue#17330](#), [pr#11155](#), Jason Dillaman)
- librbd: block name prefix might overflow fixed size C-string ([issue#17310](#), [pr#11148](#), Jason Dillaman)
- librbd: cache was not switching to writeback after first flush ([issue#16654](#), [pr#10762](#), Jason Dillaman)
- librbd: corrected use-after-free in ImageWatcher ([issue#17289](#), [pr#11112](#), Jason Dillaman)
- librbd: deadlock when replaying journal during image open ([issue#17188](#), [pr#10945](#), Jason Dillaman)
- librbd: delay acquiring lock if image watch has failed ([issue#16923](#), [pr#10574](#), Jason Dillaman)
- librbd: discard hangs when ‘rbd_skip_partial_discard’ is enabled ([issue#16386](#), [pr#10060](#), Mykola Golub)
- librbd: extract group module from librbd/internal ([pr#11070](#), Victor Denisov)
- librbd: failed assertion after shrinking a clone image twice ([issue#16561](#), [pr#10072](#), Jason Dillaman)
- librbd: fix missing return statement if failed to get mirror image state ([pr#10136](#), runsisi)
- librbd: fix possible inconsistent state when disabling mirroring fails ([issue#16984](#), [pr#10711](#), Jason Dillaman)
- librbd: ignore partial refresh error when acquiring exclusive lock ([issue#17227](#), [pr#11044](#), Jason Dillaman)
- librbd: initial hooks for client-side, image-extent cache in IO path ([pr#9121](#), Jason Dillaman)
- librbd: interlock image refresh and exclusive lock operations ([issue#16773](#), [issue#17015](#), [pr#10770](#), Jason Dillaman)
- librbd: memory leak in MirroringWatcher::notify_image_updated ([pr#11306](#), Mykola Golub)
- librbd: optimize away unnecessary object map updates ([issue#16707](#), [issue#16689](#), [pr#10332](#), Jason Dillaman)
- librbd: optionally unregister “laggy” journal clients ([issue#14738](#), [pr#10378](#), Mykola Golub)
- librbd: permit disabling journaling if in corrupt state ([issue#16740](#), [pr#10712](#), Jason Dillaman)
- librbd: possible deadlock if cluster connection closed after image ([issue#17254](#), [pr#11037](#), Jason Dillaman)
- librbd: potential deadlock closing image with in-flight readahead ([issue#17198](#), [pr#11152](#), Jason Dillaman)
- librbd: potential double-unwatch of watch handle upon error ([issue#17210](#), [pr#10974](#), Jason Dillaman)
- librbd: potential seg fault when blacklisting an image client ([issue#17251](#), [pr#11034](#), Jason Dillaman)
- librbd: prevent creation of clone from non-primary mirrored image ([issue#16449](#), [pr#10123](#), Mykola Golub)
- librbd: prevent creation of v2 image ids that are too large ([issue#16887](#), [pr#10581](#), Jason Dillaman)
- mds: Add path filtering for dump cache ([issue#11171](#), [pr#9925](#), Douglas Fuller)
- mds: Kill C_SaferCond in evict_sessions() ([issue#16288](#), [pr#9971](#), Douglas Fuller)
- mds: Return “committing” rather than “committed” member in get_committing ([pr#10250](#), Greg Farnum)
- mds: Set mds_snap_max_uid to 4294967294 ([pr#11016](#), Wido den Hollander)
- mds: add assertion in handle_slave_rename_prep ([issue#16807](#), [pr#10429](#), John Spray)
- mds: add assertions for standby_daemons invariant ([issue#16592](#), [pr#10316](#), Patrick Donnelly)
- mds: add health warning for oversized cache ([issue#16570](#), [pr#10245](#), John Spray)
- mds: add maximum fragment size constraint ([issue#16164](#), [pr#9789](#), Patrick Donnelly)
- mds: add perf counters for MDLog replay and SessionMap ([pr#10539](#), John Spray)
- mds: catch duplicates in DamageTable ([issue#17173](#), [pr#11137](#), John Spray)
- mds: fix Session::check_access() ([issue#16358](#), [pr#9769](#), Yan, Zheng)
- mds: fix daemon selection when starting ranks ([pr#10540](#), John Spray)

- mds: fix shutting down mds timed-out due to deadlock ([issue#16396](#), [pr#9884](#), Zhi Zhang)
- mds: fix up _dispatch ref-counting semantics ([pr#10533](#), Greg Farnum)
- mds: fixup dump Formatter' type error; add path_ino and is_primary in the CDentry::dump() ([pr#10119](#), huanwen ren)
- mds: handle blacklisting during journal recovery ([issue#17236](#), [pr#11138](#), John Spray)
- mds: log path with CDir damage messages ([issue#16973](#), [pr#10996](#), John Spray)
- mds: move Finisher to unlocked shutdown ([issue#16042](#), [pr#10142](#), Patrick Donnelly)
- mds: populate DamageTable from scrub and log more quietly ([issue#16016](#), [pr#11136](#), John Spray)
- mds: remove fail-safe queueing replay request ([issue#17271](#), [pr#11078](#), Yan, Zheng)
- mds: remove max_mds config option ([issue#17105](#), [pr#10914](#), Patrick Donnelly)
- mds: remove unused MDSDaemon::objecter ([pr#10566](#), Patrick Donnelly)
- mds: snap failover fixes ([pr#9955](#), Yan, Zheng)
- mds: trim null dentries proactively ([issue#16919](#), [pr#10606](#), John Spray)
- mds: unuse Class and cleanup ([pr#10399](#), huanwen ren)
- mds: use reference to avoid copy ([pr#10191](#), Patrick Donnelly)
- mds: MDCache.h: remove unneeded access specifier ([pr#10901](#), Michal Jarzabek)
- mds: MDSDaemon: move C_MDS_Tick class to .cc file ([pr#11220](#), Michal Jarzabek)
- mgr: implement con reset handling ([pr#11299](#), Sage Weil)
- mgr: squash compiler warnings ([pr#11307](#), John Spray)
- mon: MonClient may hang on pinging an unresponsive monitor ([pr#9259](#), xie xingguo)
- mon: Monitor: validate prefix on handle_command() ([issue#16297](#), [pr#9700](#), You Ji)
- mon: OSDMonitor: Missing nearfull flag set ([pr#11082](#), Igor Podoski)
- mon: change osdmap flags set and unset messages ([issue#15983](#), [pr#9252](#), Vikhyat Umrao)
- mon: clear list in better way ([pr#9718](#), song baisen)
- mon: do not recalculate 'to_remove' when it's known ([pr#9717](#), song baisen)
- mon: misc cleanups ([pr#10591](#), xie xingguo)
- mon: remove the redundant cancel_probe_timeout function ([pr#10261](#), song baisen)
- mon: remove the redundant is_active judge in PaxosService ([pr#9749](#), song baisen)
- mon: tear down standby replays on MDS rank stop ([issue#16909](#), [pr#10628](#), John Spray)
- mon: use clearer code structure ([pr#10192](#), Patrick Donnelly)
- mon: validate states transmitted in beacons ([issue#16592](#), [pr#10428](#), John Spray)
- mon: wait 10m (not 5m) before marking down OSDs out ([pr#11184](#), Sage Weil)
- mon: write fsid use the right return value ([pr#10197](#), song baisen)
- mon: Elector:move C_ElectionExpire class to cc file ([pr#10416](#), Michal Jarzabek)
- mon: HealthMonitor: add override to virtual functs ([pr#10549](#), Michal Jarzabek)
- mon: HealthMonitor: remove unneeded include ([pr#10563](#), Michal Jarzabek)
- mon: MonClient.h: delete copy constr and assing op ([pr#10599](#), Michal Jarzabek)
- mon: MonClient: move C_CancelMonCommand to cc file ([pr#10392](#), Michal Jarzabek)
- mon: MonClient: move C_Tick struct to cc file ([pr#10383](#), Michal Jarzabek)
- mon: Monitor.h: add override to virtual functions ([pr#10515](#), Michal Jarzabek)
- mon: Monitor: move C_Scrub, C_ScrubTimeout to .cc ([pr#10513](#), Michal Jarzabek)
- mon: OSDMonitor.cc: remove unneeded casts ([pr#10575](#), Michal Jarzabek)
- mon: Paxos: move classes to .cc file ([pr#11215](#), Michal Jarzabek)
- mon: PaxosService: move classes to cc file ([pr#10529](#), Michal Jarzabek)
- mon: remove the redundant list swap in paxos commit_proposal ([pr#10011](#), song baisen)
- msgr: set close on exec flag ([issue#16390](#), [pr#9772](#), Kefu Chai)
- msgr: Acceptor.h: add override to virtual function ([pr#10422](#), Michal Jarzabek)
- msgr: Acceptor: move include to cc file ([pr#10441](#), Michal Jarzabek)
- msgr: AsyncConnection: add const to mem functions ([pr#10302](#), Michal Jarzabek)
- msgr: AsyncMessenger.cc: remove unneeded cast ([pr#10141](#), Michal Jarzabek)
- msgr: AsyncMessenger: add const to function ([pr#10114](#), Michal Jarzabek)
- msgr: AsyncMessenger: move C_handle_reap class to cc ([pr#10113](#), Michal Jarzabek)
- msgr: AsyncMessenger: move C_processor_accept class ([pr#9991](#), Michal Jarzabek)
- msgr: AsyncMessenger: remove unneeded include file ([pr#10195](#), Michal Jarzabek)
- msgr: AsyncMessenger: remove unused function ([pr#10163](#), Michal Jarzabek)
- msgr: EventKqueue.h: add override to virtual func ([pr#10318](#), Michal Jarzabek)
- msgr: EventPoll.h: add override to virtual functions ([pr#10314](#), Michal Jarzabek)
- msgr: EventSelect.h: add override to virtual funct ([pr#10321](#), Michal Jarzabek)
- msgr: EventSelect: move includes to cc file ([pr#10333](#), Michal Jarzabek)
- msgr: FastStrategy.h: add override to virtual funct ([pr#10482](#), Michal Jarzabek)
- msgr: Message.h: add const to member function ([pr#10354](#), Michal Jarzabek)
- msgr: Message.h: remove code duplication ([pr#10356](#), Michal Jarzabek)
- msgr: QueueStrategy: add override to virtual functs ([pr#10503](#), Michal Jarzabek)
- msgr: Stack.h: delete copy constr and assign op ([pr#11107](#), Michal Jarzabek)
- msgr: async/Event.h: add const to member function ([pr#10224](#), Michal Jarzabek)
- msgr: async: remove unused code. ([pr#11247](#), Jianpeng Ma)

- osd: bail out if transaction size overflows ([issue#16982](#), [pr#10753](#), Kefu Chai)
- osd: cleanup options and other redundancies ([pr#10450](#), xie xingguo)
- osd: drop unused variables/methods ([pr#10559](#), xie xingguo)
- osd: fix the mem leak of RepGather ([issue#16801](#), [pr#10423](#), Kefu Chai)
- osd: fixups to explicitly persistently missing sets ([pr#10405](#), Samuel Just)
- osd: increment stats on recovery pull also ([issue#16277](#), [pr#10152](#), Kefu Chai)
- osd: limit omap data in push op ([issue#16128](#), [pr#9894](#), Wanlong Gao)
- osd: minor performance improvements ([pr#10470](#), xie xingguo)
- osd: minor performance improvements and fixes ([pr#10526](#), xie xingguo)
- osd: misc fixes and cleanups ([pr#10610](#), xie xingguo)
- osd: miscellaneous fixes ([pr#10572](#), xie xingguo)
- osd: more cleanups ([pr#10548](#), xie xingguo)
- osd: object class loading and execution permissions ([pr#9972](#), Noah Watkins)
- osd: pass shared_ptr by const reference ([pr#11266](#), Michal Jarzabek)
- osd: persist the missing set explicitly ([pr#10334](#), Samuel Just)
- osd: remove dispatch queue check since we don't queue hb message to this ([pr#9947](#), Haomai Wang)
- osd: remove duplicated function ([pr#9117](#), Wei Jin)
- osd: replace ceph:atomic_t with std::atomic in osd module. ([pr#9138](#), Xiaowei Chen)
- osd: should not look up an empty pg ([issue#17380](#), [pr#11208](#), Kefu Chai, Loic Dachary)
- osd: small cleanups ([pr#9980](#), Wanlong Gao)
- osd: subscribe for old osdmaps when pause flag is set ([issue#17023](#), [pr#10725](#), Kefu Chai)
- osd: preserve allocation hint attribute during recovery ([pr#9452](#), yaoning)
- osd: osd-fast-mark-down.sh: fix typo in variable assignments ([pr#11224](#), Willem Jan Withagen)
- osd: OSD.cc: initialise variable at definition ([pr#11099](#), Michal Jarzabek)
- osd: OSD.cc: remove unneeded searching of map ([pr#11000](#), Michal Jarzabek)
- osd: OSD.h: make some members private ([pr#11085](#), Michal Jarzabek)
- osd: PG.cc: remove unneeded use of count ([pr#11228](#), Michal Jarzabek)
- osd: PGBackend.h: move structs to .cc file ([pr#10975](#), Michal Jarzabek)
- osd: ReplicatedBackend: move classes to cc file ([pr#10967](#), Michal Jarzabek)
- osd: ReplicatedPG.h: add override to virtual funct ([pr#11271](#), Michal Jarzabek)
- osd: ReplicatedPG: move classes to .cc file ([pr#10971](#), Michal Jarzabek)
- osd: ReplicatedPG:move C OSD_OnApplied class to cc ([pr#11288](#), Michal Jarzabek)
- osd: Watch.h: remove unneeded forward declaration ([pr#10269](#), Michal Jarzabek)
- osd: osdc/ObjectCacher.h: add const to member functions ([pr#9569](#), Michal Jarzabek)
- osd: osdc/ObjectCacher.h: add const to member functions ([pr#9652](#), Michal Jarzabek)
- osd: osdc/Objecter: move RequestStateHook class to .cc ([pr#10734](#), Michal Jarzabek)
- pybind: Port Python-based tests and remaining Python bindings to Python 3 ([pr#10177](#), Oleh Prypin)
- pybind: Rework cephfs/setup.py for PyPI ([pr#10315](#), Anirudha Bose)
- pybind: Rework rbd/setup.py for PyPI ([issue#16940](#), [pr#10376](#), Anirudha Bose)
- pybind: global/signal_handler: dump cmdline instead of arg[0] ([pr#10345](#), Kefu Chai)
- pybind: make rados ready for PyPI ([pr#9833](#), Anirudha Bose)
- pybind: pybind/ceph_argparse: handle non ascii unicode args ([issue#12287](#), [pr#8943](#), Kefu Chai)
- pybind: Python 3 compatibility for workunits ([pr#10815](#), Anirudha Bose)
- rbd: Allow user to remove snapshot with -force to auto flatten children ([pr#10087](#), Dongsheng Yang)
- rbd: Reviewed-off-by: Ilya Dryomov <idryomov@gmail.com> ([issue#16171](#), [pr#10481](#), Jason Dillaman)
- rbd: Reviewed-off-by: Ilya Dryomov <idryomov@gmail.com> ([issue#17030](#), [pr#10735](#), Jason Dillaman)
- rbd: bench io-size should not be larger than image size ([issue#16967](#), [pr#10708](#), Jason Dillaman)
- rbd: cleanup - Proxied operations shouldn't result in error messages if replayed ([issue#16130](#), [pr#9724](#), Vikhyat Umrao)
- rbd: cls_rbd: clean up status from rbd-mirror if image removed ([pr#11142](#), Huan Zhang)
- rbd: cls_rbd: set omap values in batch during image creation ([pr#9981](#), Dongsheng Yang)
- rbd: inherit the parent image features when cloning an image ([issue#15388](#), [pr#9334](#), Dongsheng Yang)
- rbd: journal: ensure in-flight ops are complete destroying journaler ([issue#17446](#), [pr#11257](#), Mykola Golub, Jason Dillaman)
- rbd: journal: increase concurrency/parallelism of journal recorder ([issue#15259](#), [pr#10445](#), Ricardo Dias)
- rbd: journal: move JournalTrimmer::C_RemoveSet struct ([pr#10912](#), Michal Jarzabek)
- rbd: qa/workunits/rbd: before removing image make sure it is not bootstrapped ([issue#16555](#), [pr#10155](#), Mykola Golub)
- rbd: qa/workunits/rbd: check status also in pool dir after asok commands ([pr#11291](#), Mykola Golub)
- rbd: qa/workunits/rbd: set image-meta on primary image and wait it is replicated ([pr#11294](#), Mykola Golub)
- rbd: qa/workunits/rbd: small fixup and improvements for rbd-mirror tests ([pr#10483](#), Mykola Golub)
- rbd: qa/workunits/rbd: wait for image deleted before checking health ([pr#10545](#), Mykola Golub)
- rbd: qa/workunits: support filtering cls_rbd unit test cases ([issue#16529](#), [pr#10714](#), Jason Dillaman)
- rbd: rbd-mirror: 'wait_for_scheduled_deletion' callback might deadlock ([issue#16491](#), [pr#9964](#), Jason Dillaman)
- rbd: rbd-mirror: Add sparse read for sync image ([issue#16780](#), [pr#11005](#), tianqing)
- rbd: rbd-mirror: add additional test scenarios ([pr#10488](#), lande1234)
- rbd: rbd-mirror: concurrent access of event might result in heap corruption ([issue#17283](#), [pr#11104](#), Jason Dillaman)

- rbd: rbd-mirror: force-promoted image will remain R/O until rbd-mirror daemon restarted ([issue#16974](#), [pr#11090](#), Jason Dillaman)
- rbd: rbd-mirror: gracefully fail if object map is unavailable ([issue#16558](#), [pr#10065](#), Jason Dillaman)
- rbd: rbd-mirror: gracefully handle being blacklisted ([issue#16349](#), [pr#9970](#), Jason Dillaman)
- rbd: rbd-mirror: image deleter should use pool id + global image uuid for key ([issue#16538](#), [issue#16227](#), [pr#10484](#), Jason Dillaman)
- rbd: rbd-mirror: improve split-brain detection logic ([issue#16855](#), [pr#10703](#), Jason Dillaman)
- rbd: rbd-mirror: include local pool id in resync throttle unique key ([issue#16536](#), [pr#10254](#), Jason Dillaman)
- rbd: rbd-mirror: non-primary image is recording journal events during image sync ([pr#10462](#), Jason Dillaman)
- rbd: rbd-mirror: potential IO stall when using asok flush request ([issue#16708](#), [pr#10432](#), Jason Dillaman)
- rbd: rbd-mirror: potential assertion failure during error-induced shutdown ([issue#16956](#), [pr#10613](#), Jason Dillaman)
- rbd: rbd-mirror: potential race condition during failure shutdown ([issue#16980](#), [pr#10667](#), Jason Dillaman)
- rbd: rbd-mirror: quiesce in-flight event commits before shut down ([issue#17355](#), [pr#11185](#), Jason Dillaman)
- rbd: rbd-mirror: reduce memory footprint during journal replay ([issue#16223](#), [pr#10341](#), Jason Dillaman)
- rbd: rbd-mirror: remove ceph_test_rbd_mirror_image_replay test case ([issue#16539](#), [pr#10083](#), Mykola Golub)
- rbd: rbd-mirror: replaying state should include flush action ([issue#16970](#), [pr#10627](#), Jason Dillaman)
- rbd: rbd-mirror: replicate dynamic feature updates ([issue#16213](#), [pr#10980](#), Mykola Golub)
- rbd: rbd-mirror: replicate image metadata settings ([issue#16212](#), [pr#11168](#), Mykola Golub)
- rbd: rbd-mirror: snap rename does not properly replicate to peers ([issue#16622](#), [pr#10249](#), Jason Dillaman)
- rbd: rbd-nbd does not properly handle resize notifications ([issue#15715](#), [pr#9291](#), Mykola Golub)
- rbd: rbd-nbd: fix kernel deadlock during teuthology testing ([issue#16921](#), [pr#10985](#), Jason Dillaman)
- rbd: recognize lock_on_read option ([pr#11313](#), Ilya Dryomov)
- rbd: return error if we specified a wrong image name for rbd du ([issue#16987](#), [pr#11031](#), Dongsheng Yang)
- rbd: test/librbd/fsx: enable exclusive-lock feature in krbd mode ([pr#10984](#), Ilya Dryomov)
- rbd: test/rbd: fix possible mock journal race conditions ([issue#17317](#), [pr#11153](#), Jason Dillaman)
- rbd: test: cmake related fixups for rbd tests ([pr#10124](#), Mykola Golub)
- rbd: test: run-rbd-tests test cmake fixup ([pr#10134](#), Mykola Golub)
- rbd: test: use wrapper that respects RBD_FEATURES when creating rbd image ([issue#16717](#), [pr#10348](#), Mykola Golub)
- rbd: udev: add krbd readahead placeholder ([pr#10841](#), Nick Fisk)
- rbd: rbd_mirror/ImageSyncThrottler: move struct to .cc ([pr#10928](#), Michal Jarzabek)
- rgw: (build verified, provably unused/not aliased) ([pr#9993](#), weiqiaomiao)
- rgw: Add documentation for the Multi-tenancy feature ([pr#9570](#), Pete Zaitcev)
- rgw: Clean up lifecycle thread ([pr#10480](#), Daniel Gryniewicz)
- rgw: Do not archive metadata by default ([issue#17256](#), [pr#11051](#), Pavan Rallabhandi)
- rgw: Fix Host->bucket fallback logic inversion ([issue#15975](#), [issue#17136](#), [pr#10873](#), Robin H. Johnson)
- rgw: Fix for using port 443 with pre-signed urls. ([issue#16548](#), [pr#10088](#), Pritha Srivastava)
- rgw: Fix incorrect content length and range for zero sized objects during range requests ([issue#16388](#), [pr#10207](#), Pavan Rallabhandi)
- rgw: Got rid of recursive mutex. ([pr#10562](#), Adam Kupczyk)
- rgw: RGW : setting socket backlog for via ceph.conf ([issue#16406](#), [pr#9891](#), Feng Guo)
- rgw: RGWMetaSyncCR holds refs to stacks instead of crs ([issue#16666](#), [pr#10301](#), Casey Bodley)
- rgw: Reviewed by: Pritha Srivastava <prsriwas@redhat.com> ([issue#16188](#), [pr#9584](#), Albert Tu)
- rgw: Rgw lifecycle testing ([pr#11131](#), Daniel Gryniewicz)
- rgw: Rgw nfs 28 ([pr#10611](#), Matt Benjamin)
- rgw: add configurables for {data,meta} sync error injection ([pr#10388](#), Yehuda Sadeh)
- rgw: add deadlock detection to RGWCoroutinesManager::run() ([pr#10032](#), Casey Bodley)
- rgw: add lc_pool when decode or encode struct RGWZoneParams ([pr#10439](#), weiqiaomiao)
- rgw: add missing master_zone when running with old default region config ([issue#16627](#), [pr#10205](#), Orit Wasserman)
- rgw: add pg_ver to tombstone_cache ([pr#9851](#), Casey Bodley)
- rgw: add reinit/rebind logic (ldap) ([pr#10532](#), Matt Benjamin)
- rgw: add return value checking to avoid possible subsequent parse exception ([pr#10241](#), Yan Jun)
- rgw: add suport for Swift-at-root dependent features of Swift API ([issue#16673](#), [pr#10280](#), Pritha Srivastava, Radoslaw Zarzynski)
- rgw: add support for Static Website of Swift API ([pr#9844](#), Radoslaw Zarzynski)
- rgw: add tenant support to multisite sync ([issue#16469](#), [pr#10075](#), Casey Bodley)
- rgw: back off bucket sync on failures, don't store marker ([issue#16742](#), [pr#10355](#), Yehuda Sadeh)
- rgw: better error message when user has no bucket created yet ([issue#16444](#), [pr#10162](#), Gaurav Kumar Garg)
- rgw: clean-up in the authentication infrastructure ([pr#10212](#), Radoslaw Zarzynski)
- rgw: clear realm watch on failed watch_restart ([issue#16817](#), [pr#10446](#), Casey Bodley)
- rgw: collect skips a specific coroutine stack ([issue#16665](#), [pr#10274](#), Yehuda Sadeh)
- rgw: cosmetic changes only-build verified, f23 ([pr#9931](#), Yan Jun)
- rgw: delete region map after upgrade to zonegroup map ([issue#17051](#), [pr#10831](#), Casey Bodley)
- rgw: do not try to encode or decode time_t and fix compiling warnings ([pr#10751](#), Kefu Chai)
- rgw: don't fail if lost race when setting acls ([issue#16930](#), [pr#11286](#), Yehuda Sadeh)
- rgw: drop create_bucket in fwd_request log message ([pr#10214](#), Abhishek Lekshmanan)

- rgw: eradicate dynamic memory allocation in RGWPostObj. ([pr#11054](#), Radoslaw Zarzynski)
- rgw: file setattr ([pr#8618](#), Matt Benjamin)
- rgw: finish error_repo cr in stop_spawned_services() ([issue#16530](#), [pr#10031](#), Yehuda Sadeh)
- rgw: fix RGWAccessControlPolicy_SWIFT::create return value check error ([issue#17090](#), [pr#10727](#), weiqiaomiao)
- rgw: fix compilation ([pr#10252](#), Josh Durgin)
- rgw: fix decoding of creation_time and last_update. ([issue#17167](#), [pr#11132](#), Orit Wasserman)
- rgw: fix error_repo segfault in data sync ([issue#16603](#), [pr#10157](#), Casey Bodley)
- rgw: fix failed to create bucket if a non-master zonegroup has a single zone ([pr#10991](#), weiqiaomiao)
- rgw: fix flush_read_list() error msg ([pr#10749](#), Jiaying Ren)
- rgw: fix for issue 16494 ([issue#16494](#), [pr#10077](#), Yehuda Sadeh)
- rgw: fix for s3tests failure when ldap auth is not applied ([pr#10669](#), Casey Bodley)
- rgw: fix get object instance returned NoSuchKey error ([issue#17111](#), [pr#10820](#), Yang Honggang)
- rgw: fix is_admin handling in RGWLDAuthEngine and introduce acct_privilege_t ([pr#10687](#), Radoslaw Zarzynski)
- rgw: fix issue 16435 ([issue#16435](#), [pr#10193](#), Yehuda Sadeh)
- rgw: fix multi-delete query param parsing. ([issue#16618](#), [pr#10187](#), Robin H. Johnson)
- rgw: fix period update -commit return error ([issue#17110](#), [pr#10836](#), weiqiaomiao)
- rgw: fix radosgw daemon core when reopen logs ([issue#17036](#), [pr#10737](#), weiqiaomiao)
- rgw: fix regression with handling double underscore ([issue#16856](#), [pr#10939](#), Orit Wasserman)
- rgw: fix rgw_bucket_dir_entry decode v ([pr#10918](#), Tianshan Qu)
- rgw: fix the error return variable in log message and cleanups ([pr#10138](#), Yan Jun)
- rgw: fix the missing return value ([pr#10122](#), Yan Jun)
- rgw: fix upgrade from old multisite to new multisite configuration ([issue#16751](#), [pr#10368](#), Orit Wasserman)
- rgw: fix wrong variable definition in cls_version_check func ([pr#10233](#), weiqiaomiao)
- rgw: fix wrong variable definition in rgw_cls_lc_set_entry function ([pr#10408](#), weiqiaomiao)
- rgw: for the create_bucket api, if the input creation_time is zero, we should set it to 'now' ([issue#16597](#), [pr#10118](#), weiqiaomiao)
- rgw: kill a compile warning for rgw_sync ([pr#10425](#), Casey Bodley, Abhishek Lekshmanan)
- rgw: lgtm ([pr#9941](#), weiqiaomiao)
- rgw: lgtm (build verified, f23) ([pr#9754](#), John Coyle)
- rgw: lgtm, build verified f23 ([pr#10035](#), Yan Jun)
- rgw: lgtm-build verified, f23 ([pr#10002](#), Yan Jun)
- rgw: lgtm-build verified, f23 ([pr#9985](#), Yan Jun)
- rgw: lgtm-should backport ([pr#9979](#), Yan Jun)
- rgw: log mp upload failures due to parts mismatch ([pr#10424](#), Abhishek Lekshmanan)
- rgw: merge setting flags operation together and cleanups ([pr#10203](#), Yan Jun)
- rgw: miscellaneous cleanups ([pr#10299](#), Yan Jun)
- rgw: multiple fixes for Swift's object expiration ([issue#16705](#), [issue#16684](#), [pr#10330](#), Radoslaw Zarzynski)
- rgw: need to 'open_object_section' before dump stats in 'RGWGetUsage...' ([issue#17499](#), [pr#11325](#), weiqiaomiao)
- rgw: obsolete 'radosgw-admin period prepare' command ([issue#17387](#), [pr#11278](#), Gaurav Kumar Garg)
- rgw: radosgw-admin: add "-orphan-stale-secs" to -help ([issue#17280](#), [pr#11098](#), Ken Dreyer)
- rgw: radosgw-admin: zone[group] modify can change realm id ([issue#16839](#), [pr#10477](#), Casey Bodley)
- rgw: raise log levels for common radosgw-admin errors ([issue#16935](#), [pr#10602](#), Shilpa Jagannath)
- rgw: register the correct handler for cls_user_complete_stats ([issue#16624](#), [pr#10151](#), Orit Wasserman)
- rgw: remove bucket index objects when deleting the bucket ([issue#16412](#), [pr#10120](#), Orit Wasserman)
- rgw: remove possible duplicate setting ([pr#10110](#), Yan Jun)
- rgw: remove the field ret from class RGWPutLC ([pr#10726](#), weiqiaomiao)
- rgw: remove unused bufferlist variable ([pr#10194](#), Yan Jun)
- rgw: remove unused realm from radosgw-admin zone modify ([issue#16632](#), [pr#10211](#), Orit Wasserman)
- rgw: remove unused variables ([pr#10589](#), Yan Jun)
- rgw: return "NoSuchLifecycleConfiguration" if lifecycle config does not exist ([pr#10442](#), weiqiaomiao)
- rgw: revert a commit that broke s3 signature validation ([issue#17279](#), [pr#11102](#), Casey Bodley)
- rgw: rgw file: remove busy-wait in RGWLibFS::gc() ([pr#10638](#), Matt Benjamin)
- rgw: rgw ldap: protect rgw::from_base64 from non-base64 input ([pr#10777](#), Matt Benjamin)
- rgw: rgw ldap: enforce simple_bind w/LDAPv3 ([pr#10593](#), Matt Benjamin)
- rgw: rgw multisite: RGWCoroutinesManager::run returns status of last cr ([issue#17047](#), [pr#10778](#), Casey Bodley)
- rgw: rgw multisite: RGWDataSyncCR fails on errors from RGWListBucketIndexesCR ([issue#17073](#), [pr#10779](#), Casey Bodley)
- rgw: rgw multisite: fix for assertion in RGWMetaSyncCR ([issue#17044](#), [pr#10743](#), Casey Bodley)
- rgw: rgw multisite: fixes for period puller ([issue#16939](#), [pr#10596](#), Casey Bodley)
- rgw: rgw multisite: trim data logs as peer zones catch up ([pr#10372](#), Casey Bodley)
- rgw: rgw nfs v3 completions ([pr#10745](#), Matt Benjamin)
- rgw: rgw-admin: allow unsetting user's email ([issue#13286](#), [pr#11340](#), Yehuda Sadeh, Weijun Duan)
- rgw: rgw/admin: fix some return values and indents ([pr#9170](#), Yan Jun)
- rgw: rgw/rados: remove confused error printout ([pr#9351](#), Yan Jun)
- rgw: rgw/rgw_common.cc: modify the end check in RGWHTTPArgs::sys_get ([pr#9136](#), zhao kun)

- rgw: rgw/rgw_lc.cc: fix sleep time according to the error message (pr#10930, Weibing Zhang)
- rgw: rgw/rgw_main: fix unnecessary variables defined (pr#10475, zhang.zezhu)
- rgw: rgw/swift: remove redundant assignment operation (pr#11292, Yan Jun)
- rgw: rgw_file: pre-assign times (issue#17367, pr#11181, Matt Benjamin)
- rgw: rgw_file: fix rename cases and unify unlink (pr#10271, Matt Benjamin)
- rgw: rgw_file: fix set_attrs operation (pr#11159, Matt Benjamin)
- rgw: rgw_file: refuse partial, out-of-order writes (pr#10284, Matt Benjamin)
- rgw: rgw_file: restore local definition of RGWLibFS gc interval (pr#10756, Matt Benjamin)
- rgw: rgw_file: unlock() must precede out label (pr#10635, Matt Benjamin)
- rgw: right parenthesis is missing in radosgw-admin help message on caps (pr#10947, Weibing Zhang)
- rgw: set correct instance on the object (issue#17443, pr#11270, Yehuda Sadeh)
- rgw: store oldest mdlog period in rados (issue#16894, pr#10558, Casey Bodley)
- rgw: test/multi.py add a destructive attr to tests (pr#10401, Abhishek Lekshmanan)
- rgw: test/rgw: add -gateways-per-zone to test_multi.py (pr#10742, Casey Bodley)
- rgw: test_multi.py avoid creating mds (pr#10174, Abhishek Lekshmanan)
- rgw: test_rgw_bencode: null terminate strings before checking (issue#16861, pr#10510, Yehuda Sadeh)
- rgw: use endpoints from master zone instead of zonegroup (issue#16834, pr#10456, Casey Bodley)
- rgw: use the standard usage of string.find (pr#10226, Yan Jun)
- rgw: verified: f23, subset of s3tests (pr#10448, Pritha Srivastava)
- rgw: verified (pr#10000, weiqiaomiao)
- rgw: verified non-regression (MS AD) (pr#10597, Pritha Srivastava)
- rgw: verified: autobuild (issue#16928, pr#10579, Robin H. Johnson)
- rgw: verified: MS AD (pr#10307, Pritha Srivastava)
- rgw: verified: f23 (pr#10882, Michal Jarzabek)
- rgw: verified: f23 (pr#10858, Weibing Zhang)
- rgw: verified: f23 (pr#10822, Yan Jun)
- rgw: verified: f23 (pr#10929, Weibing Zhang)
- rgw: wip: rgw multisite: preserve zone's extra pool (issue#16712, pr#10397, Abhishek Lekshmanan)
- rgw: work around curl_multi_wait bug with non-blocking reads (issue#15915, issue#16695, pr#10998, Casey Bodley)
- rgw: add a s3 API of make torrent for a object (pr#10396, zhouruisong)
- rgw: add a s3 API of make torrent for a object (pr#9589, zhouruisong)
- rgw: bucket check remove _multipart_ prefix (pr#6501, Weijun Duan)
- rgw: clean unuse bufferlist (pr#10232, weiqiaomiao)
- rgw: fix rgw boot failed after upgrade to master latest version (pr#10409, weiqiaomiao)
- rgw: lifecycle feature [rebased] (pr#9737, Ji Chen, Daniel Gryniewicz)
- rgw: rgw/rados.h: remove unneeded class C_Tick (pr#10954, Michal Jarzabek)
- rgw: ext_mime_map_init add string describing for error number (pr#9807, Yan Jun)
- tests: Add test for global static non-POD segfault (pr#10486, Brad Hubbard)
- tests: populate /dev/disk/by-partuuid for scsi_debug (issue#17100, pr#10824, Loic Dachary)
- tests: use a fixture for memstore clone testing (pr#11190, Kefu Chai)
- tests: run-*make-check.sh: Make DRY_RUN actually mean a dry run (pr#11074, Brad Hubbard)
- tests: run-cmake-check.sh: Actually run the tests (pr#11075, Brad Hubbard)
- tests: run-cmake-check.sh: Init submodules (pr#11091, Brad Hubbard)
- tests: run-make-check.sh: Make DRY_RUN actually do a dry run (pr#11092, Brad Hubbard)
- tests: run-make-check.sh: pass args to do_cmake.sh (pr#10701, John Coyle)
- tests: unittest_chain_xattr: account for existing xattrs (issue#16025, pr#11109, Dan Mick)
- tests: src/test/cli/* tests: POSIX Convert grep -P to grep -E (pr#10319, Willem Jan Withagen)
- test: ceph_test_msgr: fix circular locking dependency (issue#16955, pr#10612, Kefu Chai)
- test: cli/crushtool: fix the test of compile-decompile-recompile.t (issue#17306, pr#11173, Kefu Chai)
- test: libcephdfs: fix gcc sys/fcntl.h warnings (pr#10126, John Coyle)
- test: librados: rados_connect() should succeed (issue#17087, pr#10806, Kefu Chai)
- test: mds: add fs dump in test_ceph_argparse.py (pr#10347, huanwen ren)
- test: simple_dispatcher.cc: remove unused variable (pr#9932, Michal Jarzabek)
- test: store_test: tidy-up SyntheticWorkloadState class (pr#10775, xie xingguo)
- test: More portable use of mmap(MAP_ANON) (pr#10557, Willem Jan Withagen)
- test: Removeall merged after print_function commit needs a fix (pr#10535, David Zafman)
- test: ceph-disk.sh do not kill all daemons (issue#16729, pr#10346, Kefu Chai)
- test: cephtool/test.sh: fix expect_false() calls (pr#10133, Kefu Chai)
- test: fix usage info of omapbench (pr#10089, Wanlong Gao)
- test: remove ceph_test_rados_api_tmap_migrate (issue#16144, pr#10256, Kefu Chai)
- test: test_{compression_plugin,async_compressor}: do not copy plugins (pr#10153, Kefu Chai)
- test: test_rados_tool.sh: Make script work under cttest (pr#10166, Willem Jan Withagen)
- test: qa/workunits/cephtool/test.sh: fix omission of ceph-command (pr#10979, Willem Jan Withagen)
- test: qa/workunits/cephtool/test.sh: s/TMPDIR/TEMP_DIR/ (pr#10306, Kefu Chai)
- test: qa/workunits/cephtool/test.sh: use absolute path for TEMP_DIR (pr#10430, Kefu Chai)

- tools: New “removeall” used to remove head with snapshots ([pr#10098](#), David Zafman)
- tools: do not closed stdout ; fix overload of “<” operator ([pr#9290](#), xie xingguo)
- tools: fix the core dump when get the crushmap do not exist ([pr#10451](#), song baisen)
- tools: rebuild monstore ([issue#17179](#), [pr#10933](#), Kefu Chai)
- tools: use TextTable for “rados df” plain output ([pr#9362](#), xie xingguo)
- tools: fio engine for objectstore ([pr#10267](#), Casey Bodley, Igor Fedotov, Daniel Gollub)
- tools: rados/client: fix typo ([pr#10493](#), Yan Jun)
- tools: rados/client: fix waiting on the condition variable more efficient. ([pr#9939](#), Yan Jun)
- tools: tools/rebuild_mongodb: kill comipling warning and other fixes ([pr#11117](#), xie xingguo)
- tools: authtool: Enhance argument combinations validation ([issue#2904](#), [pr#9704](#), Brad Hubbard)
- tools: ceph-disk: change ownership of initfile to ceph:ceph ([issue#16280](#), [pr#9688](#), Shylesh Kumar)
- test: ceph_test_rados_api_tmap_migrate: remove test for tmap_upgrade ([pr#10234](#), Kefu Chai)

V10.2.10 JEWEL

This point release brings a number of important bugfixes in all major components of Ceph, we recommend all Jewel 10.2.x users to upgrade.

For a detailed list of changes refer to :download: *the complete changelog <changelog/v10.2.10txt>*

NOTABLE CHANGES

- build/ops: Add fix subcommand to ceph-disk, fix SELinux denials, and speed up upgrade from non-SELinux enabled ceph to an SELinux enabled one ([issue#20077](#), [issue#20184](#), [issue#19545](#), [pr#14346](#), Boris Ranto)
- build/ops: deb: Fix logrotate packaging ([issue#19938](#), [pr#15428](#), Nathan Cutler)
- build/ops: extended, customizable systemd ceph-disk timeout ([issue#18740](#), [pr#15051](#), Alexey Sheplyakov)
- build/ops: rpm: fix python-Sphinx package name for SUSE ([issue#19924](#), [pr#15196](#), Nathan Cutler, Jan Matejek)
- build/ops: rpm: set subman cron attributes in spec file ([issue#20074](#), [pr#15473](#), Thomas Serlin)
- cephfs: ceph-fuse segfaults at mount time, assert in ceph::log::Log::stop ([issue#18157](#), [pr#16963](#), Greg Farnum)
- cephfs: df reports negative disk “used” value when quota exceed ([issue#20178](#), [pr#16151](#), John Spray)
- cephfs: get_quota_root sends lookupname op for every buffered write ([issue#20945](#), [pr#17396](#), Dan van der Ster)
- cephfs: osdc/Filer: truncate large file party by party ([issue#19755](#), [pr#15442](#), “Yan, Zheng”)
- core: an OSD was seen getting ENOSPC even with osdfailsafe_full_ratio passed ([issue#20544](#), [issue#16878](#), [issue#19733](#), [issue#15912](#), [pr#15050](#), Sage Weil, David Zafman)
- core: disable skewed utilization warning by default ([issue#20730](#), [pr#17210](#), David Zafman)
- core: interval_set: optimize intersect_of insert operations ([issue#21229](#), [pr#17514](#), Zac Medico)
- core: kv: let ceph_logger destructed after db reset ([issue#21336](#), [pr#17626](#), wumingqiao)
- core: test_envlibrados_for_rocksdb.yaml fails on crypto restart ([issue#19741](#), [pr#16293](#), Kefu Chai)
- libradosstriper silently fails to delete empty objects in jewel ([issue#20325](#), [pr#15760](#), Stan K)
- librbd: fail IO request when exclusive lock cannot be obtained ([issue#20168](#), [issue#21251](#), [pr#17402](#), Jason Dillaman)
- librbd: prevent self-blacklisting during break lock ([issue#18666](#), [pr#17412](#), Jason Dillaman)
- librbd: reacquire lock should update lock owner client id ([issue#19929](#), [pr#17385](#), Jason Dillaman)
- mds: damage reporting by ino number is useless ([issue#18509](#), [issue#16016](#), [pr#14699](#), John Spray, Michal Jarzabek)
- mds: log rotation doesn’t work if mds has respawned ([issue#19291](#), [pr#14673](#), Patrick Donnelly)
- mds: save projected path into inode_t::stray_prior_path ([issue#20340](#), [pr#16150](#), “Yan, Zheng”)
- mon: crash on shutdown, lease_ack_timeout event ([issue#19825](#), [pr#15083](#), Kefu Chai, Michal Jarzabek, Alexey Sheplyakov)
- mon: Disallow enabling ‘hashpspool’ option to a pool without some kind of -i-understand-this-will-remap-all-pgs flag ([issue#18468](#), [pr#13507](#), Vikhyat Umrao)
- mon: factor mon_osd_full_ratio into MAX AVAIL calc ([issue#18522](#), [pr#15236](#), Sage Weil)
- mon: fail to form large quorum; msg/async busy loop ([issue#20230](#), [pr#15726](#), Haomai Wang, Michal Jarzabek)
- mon: fix force_pg_create pg stuck in creating bug ([issue#18298](#), [pr#17008](#), Alexey Sheplyakov)
- mon: osd crush set crushmap need sanity check ([issue#19302](#), [pr#16144](#), Loic Dachary)
- osd: Add heartbeat message for Jumbo Frames (MTU 9000) ([issue#20087](#), [issue#20323](#), [pr#16059](#), Piotr Dałek, Sage Weil, Greg Farnum)
- osd: fix infinite loops in fiemap ([issue#19996](#), [pr#15189](#), Sage Weil, Ning Yao)
- osd: leaked MOSDMap ([issue#18293](#), [pr#14943](#), Sage Weil)
- osd: objecter full_try behavior not consistent with osd ([issue#19430](#), [pr#15474](#), Sage Weil)
- osd: omap threadpool heartbeat is only reset every 100 values ([issue#20375](#), [pr#16167](#), Josh Durgin)
- osd: osd_internal_types: wake snaptrimmer on put_read lock, too ([issue#19131](#), [pr#16015](#), Sage Weil)
- osd: PrimaryLogPG: do not call on_shutdown() if (pg.deleting) ([issue#19902](#), [pr#15065](#), Kefu Chai)
- osd: rados ls on pool with no access returns no error ([issue#20043](#), [issue#19790](#), [pr#16473](#), Nathan Cutler, Kefu Chai, John Spray, Sage Weil, Brad Hubbard)

- osd: ReplicatedPG: solve cache tier osd high memory consumption ([issue#20464](#), [pr#16169](#), Peng Xie)
- osd: Reset() snaptrimmer on shutdown and do not default-abort on leaked pg refs ([issue#19931](#), [pr#15322](#), Greg Farnum)
- osd: scrub_to specifies clone ver, but transaction include head write ver ([issue#20041](#), [pr#16405](#), David Zafman)
- osd: unlock sdata_op_ordering_lock with sdata_lock hold to avoid missing wakeup signal ([issue#20427](#), [pr#15947](#), Alexey Sheplyakov)
- qa: add a sleep after restarting osd before "tell"ing it ([issue#16239](#), [pr#15475](#), Kefu Chai)
- rbd: api: is_exclusive_lock_owner shouldn't return -EBUSY ([issue#20182](#), [pr#16296](#), Jason Dillaman)
- rbd: cli: ensure positional arguments exist before casting ([issue#20185](#), [pr#16295](#), Jason Dillaman)
- rbd: cli: map with cephx disabled results in error message ([issue#19035](#), [pr#16297](#), Jason Dillaman)
- rbd: default features should be negotiated with the OSD ([issue#17010](#), [pr#14874](#), Mykola Golub, Jason Dillaman)
- rbd: Enabling mirroring for a pool with clones may fail ([issue#19798](#), [issue#19130](#), [pr#14663](#), Mykola Golub, Jason Dillaman)
- rbd-mirror: image sync should send NOCACHE advise flag ([issue#17127](#), [pr#16285](#), Mykola Golub)
- rbd: object-map: batch updates during trim operation ([issue#17356](#), [pr#15460](#), Mykola Golub, Venky Shankar, Nathan Cutler)
- rbd: Potential IO hang if image is flattened while read request is in-flight ([issue#19832](#), [pr#15464](#), Jason Dillaman)
- rbd: rbd_clone_copy_on_read ineffective with exclusive-lock ([issue#18888](#), [pr#16124](#), Nathan Cutler, Venky Shankar, Jason Dillaman)
- rbd: rbd-mirror: ensure missing images are re-synced when detected ([issue#19811](#), [pr#15488](#), Jason Dillaman)
- rbd: rbd-mirror: failover and fallback of unmodified image results in split-brain ([issue#19858](#), [pr#14977](#), Jason Dillaman)
- rbd: rbd-nbd: kernel reported invalid device size (0, expected 1073741824) ([issue#19871](#), [pr#15463](#), Mykola Golub)
- rgw: add the remove-x-delete feature to cancel swift object expiration ([issue#19074](#), [pr#14659](#), Jing Wenjun)
- rgw: aws4: add rgw_s3_auth_aws4_force_boto2_compat conf option ([issue#16463](#), [pr#17009](#), Javier M. Mellid)
- rgw: bucket index check in radosgw-admin removes valid index ([issue#18470](#), [pr#16856](#), Zhang Shaowen, Pavan Rallabhandi)
- rgw: cls: ceph::timespan tag_timeout wrong units ([issue#20380](#), [pr#16289](#), Matt Benjamin)
- rgw: Custom data header support ([issue#19644](#), [pr#15966](#), Pavan Rallabhandi)
- rgw: datalog trim can't work as expected ([issue#20190](#), [pr#16299](#), Zhang Shaowen)
- rgw: Delete non-empty bucket in slave zonegroup ([issue#19313](#), [pr#15477](#), Zhang Shaowen)
- rgw: Do not decrement stats cache when the cache values are zero ([issue#20661](#), [issue#20934](#), [pr#16720](#), Aleksei Gutikov, Pavan Rallabhandi)
- rgw: fix crash caused by shard id out of range when listing data log ([issue#19732](#), [pr#15465](#), redickwang)
- rgw: fix hangs in RGWRealmReloader::reload on SIGHUP ([issue#20686](#), [pr#17281](#), fang.yuxiang)
- rgw: fix infinite loop in rest api for log list ([issue#20386](#), [pr#15988](#), xierui, Casey Bodley)
- rgw: fix race in RGWCompleteMultipart ([issue#20861](#), [pr#16767](#), Abhishek Varshney, Matt Benjamin)
- rgw: Fix up to 1000 entries at a time in check_bad_index_multipart ([issue#20772](#), [pr#16880](#), Orit Wasserman, Matt Benjamin)
- rgw: folders starting with _ underscore are not in bucket index ([issue#19432](#), [pr#16276](#), Giovani Rinaldi, Orit Wasserman)
- rgw: 'gc list -include-all' command infinite loop the first 1000 items ([issue#19978](#), [pr#15719](#), Shasha Lu, fang yuxiang)
- rgw: meta sync thread crash at RGWMetaSyncShardCR ([issue#20251](#), [pr#16711](#), fang yuxiang, Nathan Cutler)
- rgw: multipart copy-part remove '/' for s3 java sdk request header ([issue#20075](#), [pr#16266](#), donglingpeng)
- rgw: multipart parts on versioned bucket create versioned bucket index entries ([issue#19604](#), [issue#17964](#), [pr#17278](#), Zhang Shaowen)
- rgw: multisite: after CreateBucket is forwarded to master, local bucket may use different value for bucket index shards ([issue#19745](#), [pr#15450](#), Shasha Lu)
- rgw: multisite: bucket zonegroup redirect not working ([issue#19488](#), [pr#15448](#), Casey Bodley)
- rgw: multisite: fixes for meta sync across periods ([issue#18639](#), [pr#15556](#), Casey Bodley)
- rgw: multisite: lock is not released when RGWMetaSyncShardCR::full_sync() fails to write marker ([issue#18077](#), [pr#17155](#), Zhang Shaowen)
- rgw: multisite: log_meta on secondary zone causes continuous loop of metadata sync ([issue#20357](#), [issue#20244](#), [pr#17148](#), Orit Wasserman, Casey Bodley)
- rgw: multisite: memory leak on failed lease in RGWDataSyncShardCR ([issue#19861](#), [issue#19834](#), [issue#19446](#), [pr#15457](#), Casey Bodley, weiqiaomiao)
- rgw: multisite: operating bucket's acl&cors is not restricted on slave zone ([issue#16888](#), [pr#15453](#), Casey Bodley, Shasha Lu, Guo Zhandong)
- rgw: multisite: realm rename does not propagate to other clusters ([issue#19746](#), [pr#15454](#), Casey Bodley)
- rgw: multisite: rest api fails to decode large period on "period commit" ([issue#19505](#), [pr#15447](#), Casey Bodley)
- rgw: multisite: RGWPeriodPuller does not call RGWPeriod::reflect() on new period ([issue#19816](#), [issue#19817](#), [pr#17167](#), Casey Bodley)
- rgw: multisite: RGWRadosRemoveOmapKeysCR::request_complete return val is wrong ([issue#20539](#), [pr#17156](#), Shasha Lu)
- rgw: not initialized pointer cause rgw crash with ec data pool ([issue#20542](#), [pr#17164](#), Aleksei Gutikov, fang yuxiang)
- rgw: radosgw-admin: bucket rm with -bypass-gc and without -purge-data doesn't throw error message ([issue#20688](#),

[pr#17159](#), Abhishek Varshney)

- rgw: radosgw-admin data sync run crash ([issue#20423](#), [pr#17165](#), Shasha Lu)
- rgw: radosgw-admin: fix bucket limit check argparse, div(0) ([issue#20966](#), [pr#16952](#), Matt Benjamin)
- rgw: reduce log level of 'storing entry at' in cls_log ([issue#19835](#), [pr#15455](#), Willem Jan Withagen)
- rgw: remove unnecessary 'error in read_id for object name: default' ([issue#19922](#), [pr#15197](#), weiqiaomiao)
- rgw: replace '+' with "%20" in canonical query string for s3 v4 auth ([issue#20501](#), [pr#16951](#), Zhang Shaowen, Matt Benjamin)
- rgw: rgw_common.cc: modify the end check in RGWHTTPArgs::sys_get ([issue#16072](#), [pr#16268](#), zhao kun)
- rgw: rgw_file: cannot delete bucket w/uxattrs ([issue#20061](#), [issue#20047](#), [issue#19214](#), [issue#20045](#), [pr#15459](#), Matt Benjamin)
- rgw: rgw_file: fix size and (c|m)time unix attrs in write_finish ([issue#19653](#), [pr#15449](#), Matt Benjamin)
- rgw: rgw_file: incorrect lane lock behavior in evict_block() ([issue#21141](#), [pr#17597](#), Matt Benjamin)
- rgw: rgw_file: prevent conflict of mkdir between restarts ([issue#20275](#), [pr#17147](#), Gui Hecheng)
- rgw: rgw_file: v3 write timer does not close open handles ([issue#19932](#), [pr#15456](#), Matt Benjamin)
- rgw: Segmentation fault when exporting rgw bucket in nfs-ganesha ([issue#20663](#), [pr#17285](#), Matt Benjamin)
- rgw: send data-log list infinitely ([issue#20951](#), [pr#17287](#), fang.yuxiang)
- rgw: set latest object's acl failed ([issue#18649](#), [pr#15451](#), Zhang Shaowen)
- rgw: Truncated objects ([issue#20107](#), [pr#17166](#), Yehuda Sadeh)
- rgw: uninitialized memory is accessed during creation of bucket's metadata ([issue#20774](#), [pr#17280](#), Radoslaw Zarzynski)
- rgw: usage logging on tenanted buckets causes invalid memory reads ([issue#20779](#), [pr#17279](#), Radoslaw Zarzynski)
- rgw: user quota did not work well on multipart upload ([issue#19285](#), [issue#19602](#), [pr#17277](#), Zhang Shaowen)
- rgw: VersionIdMarker and NextVersionIdMarker are not returned when listing object versions ([issue#19886](#), [pr#16316](#), Zhang Shaowen)
- rgw: when uploading objects continuously into a versioned bucket, some objects will not sync ([issue#18208](#), [pr#15452](#), lvshuhua)
- tools: ceph cli: Rados object in state configuring race ([issue#16477](#), [pr#15762](#), Loic Dachary)
- tools: ceph-disk: dmcrypt cluster must default to ceph ([issue#20893](#), [pr#16870](#), Loic Dachary)
- tools: ceph-disk: don't activate suppressed journal devices ([issue#19489](#), [pr#16703](#), David Disseldorp)
- tools: ceph-disk: separate ceph-osd -check-needs-* logs ([issue#19888](#), [pr#15503](#), Loic Dachary)
- tools: ceph-disk: systemd unit timesout too quickly ([issue#20229](#), [pr#17133](#), Loic Dachary)
- tools: ceph-disk: Use stdin for 'config-key put' command ([issue#21059](#), [pr#17084](#), Brad Hubbard, Loic Dachary, Sage Weil)
- tools: libradosstriper processes arbitrary printf placeholders in user input ([issue#20240](#), [pr#17574](#), Stan K)

V10.2.9 JEWEL

This point release fixes a regression introduced in v10.2.8.

We recommend that all Jewel users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- cephfs: Damaged MDS with 10.2.8 ([issue#20599](#), [pr#16282](#), Nathan Cutler)

V10.2.8 JEWEL

This point release brought a number of important bugfixes in all major components of Ceph. However, it also introduced a regression that could cause MDS damage, and a new release, v10.2.9, was published to address this. Therefore, Jewel users should *not* upgrade to this version - instead, we recommend upgrading directly to v10.2.9.

For more detailed information, see [the complete changelog](#).

OSD REMOVAL CAVEAT

There was a bug introduced in Jewel (#19119) that broke the mapping behavior when an "out" OSD that still existed in the CRUSH map was removed with 'osd rm'. This could result in 'misdirected op' and other errors. The bug is now fixed, but the fix itself introduces the same risk because the behavior may vary between clients and OSDs. To avoid problems, please ensure that all OSDs are removed from the CRUSH map before deleting them. That is, be sure to do:

```
ceph osd crush rm osd.123
```

before:

```
ceph osd rm osd.123
```

SNAP TRIMMER IMPROVEMENTS

This release greatly improves control and throttling of the snap trimmer. It introduces the “osd max trimming pgs” option (defaulting to 2), which limits how many PGs on an OSD can be trimming snapshots at a time. And it restores the safe use of the “osd snap trim sleep” option, which defaults to 0 but otherwise adds the given number of seconds in delay between every dispatch of trim operations to the underlying system.

OTHER NOTABLE CHANGES

- build/ops: “osd marked itself down” will not be recognised if host runs mon + osd on shutdown/reboot ([issue#18516](#), [pr#13492](#), Boris Ranto)
- build/ops: ceph-base package missing dependency for psmisc ([issue#19129](#), [pr#13786](#), Nathan Cutler)
- build/ops: enable build of ceph-resource-agents package on rpm-based os ([issue#17613](#), [issue#19546](#), [pr#13606](#), Nathan Cutler)
- build/ops: rbdmap.service not included in debian packaging (jewel-only) ([issue#19547](#), [pr#14383](#), Ken Dreyer)
- cephfs: Journaler may execute on_safe contexts prematurely ([issue#20055](#), [pr#15468](#), “Yan, Zheng”)
- cephfs: MDS assert failed when shutting down ([issue#19204](#), [pr#14683](#), John Spray)
- cephfs: MDS goes readonly writing backtrace for a file whose data pool has been removed ([issue#19401](#), [pr#14682](#), John Spray)
- cephfs: MDS server crashes due to inconsistent metadata ([issue#19406](#), [pr#14676](#), John Spray)
- cephfs: No output for ceph mds rmfailed 0 –yes-i-really-mean-it command ([issue#16709](#), [pr#14674](#), John Spray)
- cephfs: Test failure: test_data_isolated (tasks.cephfs.test_volume_client.TestVolumeClient) ([issue#18914](#), [pr#14685](#), “Yan, Zheng”)
- cephfs: Test failure: test_open_inode ([issue#18661](#), [pr#14669](#), John Spray)
- cephfs: The mount point break off when mds switch happened ([issue#19437](#), [pr#14679](#), Guan yunfei)
- cephfs: ceph-fuse does not recover after lost connection to MDS ([issue#16743](#), [issue#18757](#), [pr#14698](#), Kefu Chai, Henrik Korkuc, Patrick Donnelly)
- cephfs: client: fix the cross-quota rename boundary check conditions ([issue#18699](#), [pr#14667](#), Greg Farnum)
- cephfs: mds is crushed, after I set about 400 64KB xattr kv pairs to a file ([issue#19033](#), [pr#14684](#), Yang Honggang)
- cephfs: non-local quota changes not visible until some IO is done ([issue#17939](#), [pr#15466](#), John Spray, Nathan Cutler)
- cephfs: normalize file open flags internally used by cephfs ([issue#18872](#), [issue#19890](#), [pr#15000](#), Jan Fajerski, “Yan, Zheng”)
- common: monitor creation with IPv6 public network segfaults ([issue#19371](#), [pr#14324](#), Fabian Grünbichler)
- common: radosstriper: protect aio_write API from calls with 0 bytes ([issue#14609](#), [pr#13254](#), Sébastien Ponce)
- core: Objecter::epoch_barrier isn’t respected in _op_submit() ([issue#19396](#), [pr#14332](#), Ilya Dryomov)
- core: clear divergent_priors set off disk ([issue#17916](#), [pr#14596](#), Greg Farnum)
- core: improve snap trimming, enable restriction of parallelism ([issue#19241](#), [pr#14492](#), Samuel Just, Greg Farnum)
- core: os/filestore/HashIndex: be loud about splits ([issue#18235](#), [pr#13788](#), Dan van der Ster)
- core: os/filestore: fix clang static check warn use-after-free ([issue#19311](#), [pr#14044](#), liuchang0812, yaoning)
- core: transient jersasure unit test failures ([issue#18070](#), [issue#17762](#), [issue#18128](#), [issue#17951](#), [pr#14701](#), Kefu Chai, Pan Liu, Loic Dachary, Jason Dillaman)
- core: two instances of omap_digest mismatch ([issue#18533](#), [pr#14204](#), Samuel Just, David Zafman)
- doc: Improvements to crushtool manpage ([issue#19649](#), [pr#14635](#), Loic Dachary, Nathan Cutler)
- doc: PendingReleaseNotes: note about 19119 ([issue#19119](#), [pr#13732](#), Sage Weil)
- doc: admin ops: fix the quota section ([issue#19397](#), [pr#14654](#), Chu, Hua-Rong)
- doc: radosgw-admin: add the ‘object stat’ command to usage ([issue#19013](#), [pr#13872](#), Pavan Rallabhandi)
- doc: rgw S3 create bucket should not do response in json ([issue#18889](#), [pr#13874](#), Abhishek Lekshmanan)
- fs: Invalid error code returned by MDS is causing a kernel client WARNING ([issue#19205](#), [pr#13831](#), Jan Fajerski, xie xingguo)
- librbd: Incomplete declaration for ContextWQ in librbd/Journal.h ([issue#18862](#), [pr#14152](#), Boris Ranto)
- librbd: Issues with C API image metadata retrieval functions ([issue#19588](#), [pr#14666](#), Mykola Golub)
- librbd: Possible deadlock performing a synchronous API action while refresh in-progress ([issue#18419](#), [pr#13154](#), Jason Dillaman)
- librbd: is_exclusive_lock_owner API should ping OSD ([issue#19287](#), [pr#14481](#), Jason Dillaman)
- librbd: remove image header lock assertions ([issue#18244](#), [pr#13809](#), Jason Dillaman)
- mds: C_MDSInternalNoop::complete doesn’t free itself ([issue#19501](#), [pr#14677](#), “Yan, Zheng”)

- mds: Too many stat ops when trying to probe a large file ([issue#19955](#), [pr#15472](#), "Yan, Zheng")
- mds: avoid reusing deleted inode in StrayManager::_purge_stray_logged ([issue#18877](#), [pr#14670](#), Zhi Zhang)
- mds: enable start when session ino info is corrupt ([issue#19708](#), [issue#16842](#), [pr#14700](#), John Spray)
- mds: fragment space check can cause replayed request fail ([issue#18660](#), [pr#14668](#), "Yan, Zheng")
- mds: heartbeat timeout during rejoin, when working with large amount of caps/inodes ([issue#19118](#), [pr#14672](#), John Spray)
- mds: issue new caps when sending reply to client ([issue#19635](#), [pr#15438](#), "Yan, Zheng")
- mon: OSDMonitor: make 'osd crush move ...' work on osds ([issue#18587](#), [pr#13261](#), Sage Weil)
- mon: fix 'sortbitwise' warning on jewel ([issue#20578](#), [pr#15208](#), huanwen ren, Sage Weil)
- mon: make get_mon_log_message() atomic ([issue#19427](#), [pr#14587](#), Kefu Chai)
- mon: remove bad rocksdb option ([issue#19392](#), [pr#14236](#), Sage Weil)
- msg: IPv6 Heartbeat packets are not marked with DSCP QoS - simple messenger ([issue#18887](#), [pr#13450](#), Yan Jun, Robin H. Johnson)
- msg: set close on exec flag ([issue#16390](#), [pr#13585](#), Kefu Chai)
- osd: -flush-journal: sporadic segfaults on exit ([issue#18820](#), [pr#13477](#), Alexey Sheplyakov)
- osd: Give requested scrubs a higher priority ([issue#15789](#), [pr#14686](#), David Zafman)
- osd: Implement asynchronous scrub sleep ([issue#19986](#), [issue#19497](#), [pr#15529](#), Brad Hubbard)
- osd: Object level shard errors are tracked and used if no auth available ([issue#20089](#), [pr#15416](#), David Zafman)
- osd: ReplicatedPG: try with pool's use-gmt setting if hitset archive not found ([issue#19185](#), [pr#13827](#), Kefu Chai)
- osd: allow client throttler to be adjusted on-fly, without restart ([issue#18791](#), [pr#13214](#), Piotr Dałek)
- osd: bypass readonly ops when osd full ([issue#19394](#), [pr#14181](#), Jianpeng Ma, yaoning)
- osd: degraded and misplaced status output inaccurate ([issue#18619](#), [pr#14325](#), David Zafman)
- osd: new added OSD always down when full flag is set ([issue#15025](#), [pr#14326](#), Mingxin Liu)
- osd: pg_pool_t::encode(): be compatible with Hammer <= 0.94.6 ([issue#19508](#), [pr#14392](#), Alexey Sheplyakov)
- osd: pre-jewel "osd rm" incrementals are misinterpreted ([issue#19119](#), [pr#13884](#), Ilya Dryomov)
- osd: preserve allocation hint attribute during recovery ([issue#19083](#), [pr#13647](#), yaoning)
- osd: promote throttle parameters are reversed ([issue#19773](#), [pr#14791](#), Mark Nelson)
- osd: reindex properly on pg log split ([issue#18975](#), [pr#14047](#), Alexey Sheplyakov)
- osd: restrict want_acting to up+acting on recovery completion ([issue#18929](#), [pr#13541](#), Sage Weil)
- rbd-nbd: check /sys/block/nbdX/size to ensure kernel mapped correctly ([issue#18335](#), [pr#13932](#), Mykola Golub, Alexey Sheplyakov)
- rbd: [api] temporarily restrict (rbd_)mirror_peer_add from adding multiple peers ([issue#19256](#), [pr#14664](#), Jason Dillaman)
- rbd: qemu crash triggered by network issues ([issue#18436](#), [pr#13244](#), Jason Dillaman)
- rbd: rbd -pool=x rename y z does not work ([issue#18326](#), [pr#14148](#), Gaurav Kumar Garg)
- rbd: systemctl stop rbdmap unmaps all rbds and not just the ones in /etc/ceph/rbdmap ([issue#18884](#), [issue#18262](#), [pr#14083](#), David Disseldorp, Nathan Cutler)
- rgw: "cluster [WRN] bad locator @X on object @X...." in cluster log ([issue#18980](#), [pr#14064](#), Casey Bodley)
- rgw: 'radosgw-admin sync status' on master zone of non-master zonegroup ([issue#18091](#), [pr#13779](#), Jing Wenjun)
- rgw: Change loglevel to 20 for 'System already converted' message ([issue#18919](#), [pr#13834](#), Vikhyat Umrao)
- rgw: Use decoded URI when verifying TempURL ([issue#18590](#), [pr#13724](#), Alexey Sheplyakov)
- rgw: a few cases where rgw_obj is incorrectly initialized ([issue#19096](#), [pr#13842](#), Yehuda Sadeh)
- rgw: add apis to support ragweed suite ([issue#19804](#), [pr#14851](#), Yehuda Sadeh)
- rgw: add bucket size limit check to radosgw-admin ([issue#17925](#), [pr#14787](#), Matt Benjamin)
- rgw: allow system users to read SLO parts ([issue#19027](#), [pr#14752](#), Casey Bodley)
- rgw: don't return skew time in pre-signed url ([issue#18828](#), [issue#18829](#), [pr#14605](#), liuchang0812)
- rgw: failure to create s3 type subuser from admin rest api ([issue#16682](#), [pr#14815](#), snakeAngel2015)
- rgw: fix break inside of yield in RGWFetchAllMetaCR ([issue#17655](#), [pr#14066](#), Casey Bodley)
- rgw: fix failed to create bucket if a non-master zonegroup has a single zone ([issue#19756](#), [pr#14766](#), weiqiaomiao)
- rgw: health check errors out incorrectly ([issue#19025](#), [pr#13865](#), Pavan Rallabhandi)
- rgw: list_plain_entries() stops before bi_log entries ([issue#19876](#), [pr#15383](#), Casey Bodley)
- rgw: multisite: fetch_remote_obj() gets wrong version when copying from remote ([issue#19599](#), [pr#14607](#), Zhang Shaowen, Casey Bodley)
- rgw: multisite: some yields in RGWMetaSyncShardCR::full_sync() resume in incremental_sync() ([issue#18076](#), [pr#13837](#), Casey Bodley, Abhishek Lekshmanan)
- rgw: only append zonegroups to rest params if not empty ([issue#20078](#), [pr#15312](#), Yehuda Sadeh, Karol Mroz)
- rgw: pullup civet chunked ([issue#19736](#), [pr#14776](#), Matt Benjamin)
- rgw: rgw_file: fix event expire check, don't expire directories being read ([issue#19623](#), [issue#19270](#), [issue#19625](#), [issue#19624](#), [issue#19634](#), [issue#19435](#), [pr#14653](#), Gui Hecheng, Matt Benjamin)
- rgw: swift: disable revocation thread under certain circumstances ([issue#19499](#), [issue#9493](#), [pr#14789](#), Marcus Watts)
- rgw: the swift container acl does not support field .ref ([issue#18484](#), [pr#13833](#), Jing Wenjun)
- rgw: typo in rgw_admin.cc ([issue#19026](#), [pr#13863](#), Ronak Jain)
- rgw: unsafe access in RGWLListBucket_ObjStore_SWIFT::send_response() ([issue#19249](#), [pr#14661](#), Yehuda Sadeh)
- rgw: upgrade to multisite v2 fails if there is a zone without zone info ([issue#19231](#), [pr#14136](#), Danny Al-Gaaf, Orit Wasserman)

- rgw: use separate http_manager for read_sync_status ([issue#19236](#), [pr#14195](#), Casey Bodley, Shasha Lu)
- rgw: when converting region_map we need to use rgw_zone_root_pool ([issue#19195](#), [pr#14143](#), Orit Wasserman)
- rgw: zonegroupmap set does not work ([issue#19498](#), [issue#18725](#), [pr#14660](#), Orit Wasserman, Casey Bodley)
- rgw:fix memory leaks in data/md sync ([issue#20088](#), [pr#15382](#), weiqiaomiao)
- tests: ‘ceph auth import -i’ overwrites caps, should alert user before overwrite ([issue#18932](#), [pr#13544](#), Vikhyat Umrao)
- tests: New upgrade test for #19508 ([issue#19829](#), [issue#19508](#), [pr#14930](#), Nathan Cutler)
- tests: [FAILED] TestLibRBD.ImagePollIO in upgrade:client-upgrade-kraken-distro-basic-smithi ([issue#18617](#), [pr#13107](#), Jason Dillaman)
- tests: [librados_test_stub] cls_cxx_map_get_XYZ methods don’t return correct value ([issue#19597](#), [pr#14665](#), Jason Dillaman)
- tests: additional rbd-mirror test stability improvements ([issue#18935](#), [pr#14154](#), Jason Dillaman)
- tests: api_misc: [FAILED] LibRadosMiscConnectFailure.ConnectFailure ([issue#15368](#), [pr#14763](#), Sage Weil)
- tests: buffer overflow in test LibCephFS.DirLs ([issue#18941](#), [pr#14671](#), “Yan, Zheng”)
- tests: clone workunit using the branch specified by task ([issue#19429](#), [pr#14371](#), Kefu Chai, Dan Mick)
- tests: drop upgrade/hammer-jewel-x ([issue#20574](#), [pr#15933](#), Nathan Cutler)
- tests: dummy suite fails in OpenStack ([issue#18259](#), [pr#14070](#), Nathan Cutler)
- tests: eliminate race condition in Thrasher constructor ([issue#18799](#), [pr#13608](#), Nathan Cutler)
- tests: enable quotas for pre-luminous quota tests ([issue#20412](#), [pr#15936](#), Patrick Donnelly)
- tests: fix oversight in yaml comment ([issue#20581](#), [pr#14449](#), Nathan Cutler)
- tests: move swift.py task from teuthology to ceph, phase one (jewel) ([issue#20392](#), [pr#15870](#), Nathan Cutler, Sage Weil, Warren Usui, Greg Farnum, Ali Maredia, Tommi Virtanen, Zack Cerza, Sam Lang, Yehuda Sadeh, Joe Buck, Josh Durgin)
- tests: qa/Fixed upgrade sequence to 10.2.0 -> 10.2.7 -> latest -x (10.2.8) ([issue#20572](#), [pr#16089](#), Yuri Weinstein)
- tests: qa/suites/upgrade/hammer-x: set “sortbitwise” for jewel clusters ([issue#20342](#), [pr#15842](#), Nathan Cutler)
- tests: qa/workunits/rados/test-upgrade-*: whitelist tests for master (part 1) ([issue#20577](#), [pr#15360](#), Sage Weil)
- tests: qa/workunits/rados/test-upgrade-*: whitelist tests for master (part 2) ([issue#20576](#), [pr#15778](#), Kefu Chai)
- tests: qa/workunits/rados/test-upgrade-*: whitelist tests the right way ([issue#20575](#), [pr#15824](#), Kefu Chai)
- tests: rados: sleep before ceph tell osd.0 flush_pg_stats after restart ([issue#16239](#), [issue#20489](#), [pr#14710](#), Kefu Chai, Nathan Cutler)
- tests: run upgrade/client-upgrade on latest CentOS 7.3 ([issue#20573](#), [pr#16088](#), Nathan Cutler)
- tests: run-rbd-unit-tests.sh assert in lockdep_will_lock, TestLibRBD.ObjectMapConsistentSnap ([issue#17447](#), [pr#14150](#), Jason Dillaman)
- tests: systemd test backport to jewel ([issue#19717](#), [pr#14694](#), Vasu Kulkarni)
- tests: test/librados/tmap_migrate: g_ceph_context->put() upon return ([issue#20579](#), [pr#14809](#), Kefu Chai)
- tests: test_notify.py: rbd.InvalidArgument: error updating features for image test_notify_clone2 ([issue#19692](#), [pr#14680](#), Jason Dillaman)
- tests: upgrade/hammer-x failing with OSD has the store locked when Thrasher runs ceph-objectstore-tool on down PG ([issue#19556](#), [pr#14416](#), Nathan Cutler)
- tests: upgrade:hammer-x/stress-split-erasure-code-x86_64 fails in 10.2.8 integration testing ([issue#20413](#), [pr#15904](#), Nathan Cutler)
- tools: brag fails to count “in” mds ([issue#19192](#), [pr#14112](#), Oleh Prypin, Peng Zhang)
- tools: ceph-disk does not support cluster names different than ‘ceph’ ([issue#17821](#), [pr#14765](#), Loic Dachary)
- tools: ceph-disk: Racing between partition creation and device node creation ([issue#19428](#), [pr#14329](#), Erwan Velu)
- tools: ceph-disk: bluestore -setgroup incorrectly set with user ([issue#18955](#), [pr#13489](#), craigchi)
- tools: ceph-disk: ceph-disk list reports mount error for OSD having mount options with SELinux context ([issue#17331](#), [pr#14402](#), Brad Hubbard)
- tools: ceph-disk: do not setup_statedir on trigger ([issue#19941](#), [pr#15504](#), Loic Dachary)
- tools: ceph-disk: enable directory backed OSD at boot time ([issue#19628](#), [pr#14602](#), Loic Dachary)
- tools: rados: RadosImport::import should return an error if Rados::connect fails ([issue#19319](#), [pr#14113](#), Brad Hubbard)

V10.2.7 JEWEL

This point release fixes several important bugs in RBD mirroring, librbd & RGW.

We recommend that all v10.2.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- librbd: possible race in ExclusiveLock handle_peer_notification ([issue#19368](#), [pr#14233](#), Mykola Golub)
- osd: Increase priority for inactive PGs backfill ([issue#18350](#), [pr#13232](#), Bartłomiej Świecki)
- osd: Scrub improvements and other fixes ([issue#17857](#), [issue#18114](#), [issue#13937](#), [issue#18113](#), [pr#13146](#), Kefu Chai, David Zafman)
- osd: fix OSD network address in OSD heartbeat_check log message ([issue#18657](#), [pr#13108](#), Vikhyat Umrao)

- rbd-mirror: deleting a snapshot during sync can result in read errors ([issue#18990](#), [pr#13596](#), Jason Dillaman)
- rgw: ‘period update’ does not remove short_zone_ids of deleted zones ([issue#15618](#), [pr#14140](#), Casey Bodley)
- rgw: DUMPABLE flag is cleared by setuid preventing core dumps ([issue#19089](#), [pr#13844](#), Brad Hubbard)
- rgw: clear data_sync_cr if RGWDataSyncControlCR fails ([issue#17569](#), [pr#13886](#), Casey Bodley)
- rgw: fix openssl ([issue#11239](#), [issue#19098](#), [issue#16535](#), [pr#14215](#), Marcus Watts)
- rgw: fix swift cannot disable object versioning with empty X-Versions-Location ([issue#18852](#), [pr#13823](#), Jing Wenjun)
- rgw: librgw: RGWLibFS::setattr fails on directories ([issue#18808](#), [pr#13778](#), Matt Benjamin)
- rgw: make sending Content-Length in 204 and 304 controllable ([issue#16602](#), [pr#13503](#), Radoslaw Zarzynski, Matt Benjamin)
- rgw: multipart uploads copy part support ([issue#12790](#), [pr#13219](#), Yehuda Sadeh, Javier M. Mellid, Matt Benjamin)
- rgw: multisite: RGWMetaSyncShardControlCR gives up on EIO ([issue#19019](#), [pr#13867](#), Casey Bodley)
- rgw: radosgw/swift: clean up flush / newline behavior ([issue#18473](#), [pr#14100](#), Nathan Cutler, Marcus Watts, Matt Benjamin)
- rgw: radosgw/swift: clean up flush / newline behavior. ([issue#18473](#), [pr#13143](#), Marcus Watts, Matt Benjamin)
- rgw: rgw_fh: RGWFileHandle dtor must also cond-unlink from FHCache ([issue#19112](#), [pr#14231](#), Matt Benjamin)
- rgw: rgw_file: avoid interning .. in FHCache table and don’t ref for them ([issue#19036](#), [pr#13848](#), Matt Benjamin)
- rgw: rgw_file: interned RGWFileHandle objects need parent refs ([issue#18650](#), [pr#13583](#), Matt Benjamin)
- rgw: rgw_file: restore (corrected) fix for dir partial match (return of FLAG_EXACT_MATCH) ([issue#19060](#), [issue#18992](#), [issue#19059](#), [pr#13858](#), Matt Benjamin)
- rgw: rgw_file: FHCache residence check should be exhaustive ([issue#19111](#), [pr#14169](#), Matt Benjamin)
- rgw: rgw_file: ensure valid_s3_object_name for directories, too ([issue#19066](#), [pr#13717](#), Matt Benjamin)
- rgw: rgw_file: fix marker computation ([issue#19018](#), [issue#18989](#), [issue#18992](#), [issue#18991](#), [pr#13869](#), Matt Benjamin)
- rgw: rgw_file: wip dir orphan ([issue#18992](#), [issue#18989](#), [issue#19018](#), [issue#18991](#), [pr#14205](#), Gui Hecheng, Matt Benjamin)
- rgw: rgw_file: various fixes ([pr#14206](#), Matt Benjamin)
- rgw: rgw_file: expand argv ([pr#14230](#), Matt Benjamin)

V10.2.6 JEWEL

This point release fixes several important bugs in RBD mirroring, RGW multi-site, CephFS, and RADOS.

We recommend that all v10.2.x users upgrade.

For more detailed information, see [the complete changelog](#).

OSDS NO LONGER SEND ENXIO BY DEFAULT

In previous versions, if a client sent an op to the wrong OSD, the OSD would reply with ENXIO. The rationale here is that the client or OSD is clearly buggy and we want to surface the error as clearly as possible. We now only send the ENXIO reply if the osd_enxio_on_misdirected_op option is enabled (it’s off by default). This means that a VM using librbd that previously would have gotten an EIO and gone read-only will now see a blocked/hung IO instead.

OTHER NOTABLE CHANGES

- build/ops: add hostname sanity check to run-{c}make-check.sh ([issue#18134](#), [pr#12302](#), Nathan Cutler)
- build/ops: add ldap lib to rgw lib deps based on build config ([issue#17313](#), [pr#13183](#), Nathan Cutler)
- build/ops: ceph-create-keys loops forever ([issue#17753](#), [pr#11884](#), Alfredo Deza)
- build/ops: ceph daemons DUMPABLE flag is cleared by setuid preventing core dumps ([issue#17650](#), [pr#11736](#), Patrick Donnelly)
- build/ops: fixed compilation error when -with-radosgw=no ([issue#18512](#), [pr#12729](#), Pan Liu)
- build/ops: fixed the issue when -disable-server, compilation fails. ([issue#18120](#), [pr#12239](#), Pan Liu)
- build/ops: fix undefined crypto references with -with-xio ([issue#18133](#), [pr#12296](#), Nathan Cutler)
- build/ops: install-deps.sh based on /etc/os-release ([issue#18466](#), [issue#18198](#), [pr#12405](#), Jan Fajerski, Nitin A Kamble, Nathan Cutler)
- build/ops: Remove the runtime dependency on lsb_release ([issue#17425](#), [pr#11875](#), John Coyle, Brad Hubbard)
- build/ops: rpm: /etc/ceph/rbdmap is packaged with executable access rights ([issue#17395](#), [pr#11855](#), Ken Dreyer)
- build/ops: selinux: Allow ceph to manage tmp files ([issue#17436](#), [pr#13048](#), Boris Ranto)
- build/ops: systemd: Restart Mon after 10s in case of failure ([issue#18635](#), [pr#13058](#), Wido den Hollander)
- build/ops: systemd restarts Ceph Mon to quickly after failing to start ([issue#18635](#), [pr#13184](#), Wido den Hollander)
- ceph-disk: fix flake8 errors ([issue#17898](#), [pr#11976](#), Ken Dreyer)
- cephfs: fuse client crash when adding a new osd ([issue#17270](#), [pr#11860](#), John Spray)
- cli: ceph-disk: convert none str to str before printing it ([issue#18371](#), [pr#13187](#), Kefu Chai)

- client: Fix lookup of “/..” in jewel ([issue#18408](#), [pr#12766](#), Jeff Layton)
- client: fix stale entries in command table ([issue#17974](#), [pr#12137](#), John Spray)
- client: populate metadata during mount ([issue#18361](#), [pr#13085](#), John Spray)
- cli: implement functionality for adding, editing and removing omap values with binary keys ([issue#18123](#), [pr#12755](#), Jason Dillaman)
- common: Improve linux dcache hash algorithm ([issue#17599](#), [pr#11529](#), Yibo Cai)
- common: utime.h: fix timezone issue in round_to_* funcs. ([issue#14862](#), [pr#11508](#), Zhao Chao)
- doc: Python Swift client commands in Quick Developer Guide don’t match configuration in vstart.sh ([issue#17746](#), [pr#13043](#), Ronak Jain)
- librbd: allow to open an image without opening parent image ([issue#18325](#), [pr#13130](#), Ricardo Dias)
- librbd: metadata_set API operation should not change global config setting ([issue#18465](#), [pr#13168](#), Mykola Golub)
- librbd: new API method to force break a peer’s exclusive lock ([issue#15632](#), [issue#16773](#), [issue#17188](#), [issue#16988](#), [issue#17210](#), [issue#17251](#), [issue#18429](#), [issue#17227](#), [issue#18327](#), [issue#17015](#), [pr#12890](#), Danny Al-Gaaf, Mykola Golub, Jason Dillaman)
- librbd: properly order concurrent updates to the object map ([issue#16176](#), [pr#12909](#), Jason Dillaman)
- librbd: restore journal access when force disabling mirroring ([issue#17588](#), [pr#11916](#), Mykola Golub)
- mds: Cannot create deep directories when caps contain path=/somepath ([issue#17858](#), [pr#12154](#), Patrick Donnelly)
- mds: cephfs metadata pool: deep-scrub error omap_digest != best guess omap_digest ([issue#17177](#), [pr#12380](#), Yan, Zheng)
- mds: cephfs test failures (ceph.com/qa is broken, should be download.ceph.com/qa) ([issue#18574](#), [pr#13023](#), John Spray)
- mds: ceph-fuse crash during snapshot tests ([issue#18460](#), [pr#13120](#), Yan, Zheng)
- mds: ceph_volume_client: fix recovery from partial auth update ([issue#17216](#), [pr#11656](#), Ramana Raja)
- mds: ceph_volume_client.py : Error: Can’t handle arrays of non-strings ([issue#17800](#), [pr#12325](#), Ramana Raja)
- mds: Cleanly reject session evict command when in replay ([issue#17801](#), [pr#12153](#), Yan, Zheng)
- mds: client segfault on ceph_rmdir path / ([issue#9935](#), [pr#13029](#), Michal Jarzabek)
- mds: Clients without pool-changing caps shouldn’t be allowed to change pool_namespace ([issue#17798](#), [pr#12155](#), John Spray)
- mds: Decode errors on backtrace will crash MDS ([issue#18311](#), [pr#12836](#), Nathan Cutler, John Spray)
- mds: false failing to respond to cache pressure warning ([issue#17611](#), [pr#11861](#), Yan, Zheng)
- mds: finish clientreplay requests before requesting active state ([issue#18461](#), [pr#13113](#), Yan, Zheng)
- mds: fix incorrect assertion in Server::_dir_is_nonempty() ([issue#18578](#), [pr#13459](#), Yan, Zheng)
- mds: fix MDSMap upgrade decoding ([issue#17837](#), [pr#13139](#), John Spray, Patrick Donnelly)
- mds: fix missing ll_get for ll_walk ([issue#18086](#), [pr#13125](#), Gui Hecheng)
- mds: Fix mount root for ceph_mount users and change tarball format ([issue#18312](#), [issue#18254](#), [pr#12592](#), Jeff Layton)
- mds: fix null pointer dereference in Locker::handle_client_caps ([issue#18306](#), [pr#13060](#), Yan, Zheng)
- mds: lookup of ../ in returns -ENOENT ([issue#18408](#), [pr#12783](#), Jeff Layton)
- mds: MDS crashes on missing metadata object ([issue#18179](#), [pr#13119](#), Yan, Zheng)
- mds: mds fails to respawn if executable has changed ([issue#17531](#), [pr#11873](#), Patrick Donnelly)
- mds: MDS: false failing to respond to cache pressure warning ([issue#17716](#), [pr#11856](#), Yan, Zheng)
- mds: MDS goes damaged on blacklist (failed to read JournalPointer: -108 ((108) Cannot send after transport endpoint shutdown) ([issue#17236](#), [pr#11413](#), John Spray)
- mds: MDS long-time blocked ops. ceph-fuse locks up with getattr of file ([issue#17275](#), [pr#11858](#), Yan, Zheng)
- mds: speed up readdir by skipping unwanted dn ([issue#18519](#), [pr#12921](#), Xiaoxi Chen)
- mds: standby-replay daemons can sometimes miss events ([issue#17954](#), [pr#13126](#), John Spray)
- mon: cache tiering: base pool last_force_resend not respected (racing read got wrong version) ([issue#18366](#), [pr#13115](#), Sage Weil)
- mon: ceph osd down detection behaviour ([issue#18104](#), [pr#12677](#), xie xinguo)
- mon: Error EINVAL: removing mon.a at 172.21.15.16:6789/0, there will be 1 monitors ([issue#17725](#), [pr#11999](#), Joao Eduardo Luis)
- mon: health does not report pgs stuck in more than one state ([issue#17515](#), [pr#11660](#), Sage Weil)
- mon: monitor assertion failure when deactivating mds in (invalid) fscid 0 ([issue#17518](#), [pr#11862](#), Patrick Donnelly)
- mon: monitor cannot start because of FAILED assert(info.state == MDSMap::STATE_STANDBY) ([issue#18166](#), [pr#13123](#), John Spray, Patrick Donnelly)
- mon: osd flag health message is misleading ([issue#18175](#), [pr#13117](#), Sage Weil)
- mon: OSDMonitor: clear jewel+ feature bits when talking to Hammer OSD ([issue#18582](#), [pr#13131](#), Piotr Dałek)
- mon: OSDs marked OUT wrongly after monitor failover ([issue#17719](#), [pr#11947](#), Dong Wu)
- mon: peon wrongly delete routed pg stats op before receive pg stats ack ([issue#18458](#), [pr#13045](#), Mingxin Liu)
- mon: send updated monmap to its subscribers ([issue#17558](#), [pr#11743](#), Kefu Chai)
- msgr: don’t truncate message sequence to 32-bits ([issue#16122](#), [pr#12416](#), Yan, Zheng)
- msgr: msg/simple: clear_pipe when wait() is mopping up pipes ([issue#15784](#), [pr#13062](#), Sage Weil)
- msgr: msg/simple/Pipe: error decoding addr ([issue#18072](#), [pr#12291](#), Sage Weil)
- osd: Add config option to disable new scrubs during recovery ([issue#17866](#), [pr#11944](#), Wido den Hollander)
- osd: collection_list shadow return value # ([issue#17713](#), [pr#11737](#), Haomai Wang)
- osd: do not send ENXIO on misdirected op by default ([issue#18751](#), [pr#13255](#), Sage Weil)

- osd: FileStore: fiemap cannot be totally retrieved in xfs when the number of extents > 1364 ([issue#17610](#), [pr#11998](#), Kefu Chai, Ning Yao)
- osd: leveldb corruption leads to Operation not permitted not handled and assert ([issue#18037](#), [pr#12789](#), Nathan Cutler)
- osd: limit omap data in push op ([issue#16128](#), [pr#11991](#), Wanlong Gao)
- osd: osd crashes when radosgw-admin bi list -max-entries=1 command runing ([issue#17745](#), [pr#11758](#), weiqiaomiao)
- osd: osd_max_backfills default has changed, documentation should reflect that. ([issue#17701](#), [pr#11735](#), huangjun)
- osd: OSDMonitor: only reject MOSDBoot based on up_from if inst matches ([issue#17899](#), [pr#12868](#), Samuel Just)
- osd: osd/PG: publish PG stats when backfill-related states change ([issue#18369](#), [pr#12875](#), Alexey Sheplyakov, Sage Weil)
- osd: Remove extra call to reg_next_scrub() during splits ([issue#16474](#), [pr#11606](#), David Zafman)
- osd: Revert “Merge pull request #12978 from asheplyakov/jewel-18581” ([issue#18809](#), [pr#13280](#), Samuel Just)
- osd: update_log_missing does not order correctly with osd_ops ([issue#17789](#), [pr#11997](#), Samuel Just)
- qa/tasks: backport rbd_fio fixes to jewel ([issue#13512](#), [pr#13104](#), Ilya Dryomov)
- qa/tasks/workunits: backport misc fixes to jewel ([issue#18336](#), [pr#12912](#), Sage Weil)
- rados: crash adding snap to purged_snaps in ReplicatedPG::WaitingOnReplicas (part 2) ([issue#15943](#), [issue#18504](#), [pr#12791](#), Samuel Just)
- rados: Memory leaks in object_list_begin and object_list_end ([issue#18252](#), [pr#13118](#), Brad Hubbard)
- rados: The request lock RPC message might be incorrectly ignored ([issue#17030](#), [pr#10865](#), Jason Dillaman)
- rbd: add image id block name prefix APIs ([issue#18270](#), [pr#12529](#), Jason Dillaman)
- rbd: add max_part and nbds_max options in rbd nbd map, in order to keep consistent with ([issue#18186](#), [pr#12426](#), Pan Liu)
- rbd: Attempting to remove an image w/ incompatible features results in partial removal ([issue#18315](#), [pr#13156](#), Dongsheng Yang)
- rbd: bench-write will crash if -io-size is 4G ([issue#18422](#), [pr#13129](#), Gaurav Kumar Garg)
- rbd: diff calculate can hide parent extents when examining first snapshot in clone ([issue#18068](#), [pr#12322](#), Jason Dillaman)
- rbd: Exclusive lock improperly initialized on read-only image when using snap_set API ([issue#17618](#), [pr#11852](#), Jason Dillaman)
- rbd: FAILED assert(m_processing == 0) while running test_lock_fence.sh ([issue#17973](#), [pr#12323](#), Venky Shankar)
- rbd: Improve error reporting from rbd feature enable/disable ([issue#16985](#), [pr#13157](#), Gaurav Kumar Garg)
- rbd: JournalMetadata flooding with errors when being blacklisted ([issue#18243](#), [pr#12739](#), Jason Dillaman)
- rbd: librbd: use proper snapshot when computing diff parent overlap ([issue#18200](#), [pr#12649](#), Xiaoxi Chen)
- rbd: partition func should be enabled When load nbd.ko for rbd-nbd ([issue#18115](#), [pr#12754](#), Pan Liu)
- rbd: Potential race when removing two-way mirroring image ([issue#18447](#), [pr#13233](#), Mykola Golub)
- rbd: [qa] crash in journal-enabled fsx run ([issue#18618](#), [pr#13128](#), Jason Dillaman)
- rbd: ‘rbd du’ of missing image does not return error ([issue#16987](#), [pr#11854](#), Dongsheng Yang)
- rbd: rbd-mirror: gmock warnings in bootstrap request unit tests ([issue#18048](#), [issue#18012](#), [issue#18156](#), [issue#16991](#), [issue#18051](#), [pr#12425](#), Mykola Golub)
- rbd: rbd-mirror: image sync object map reload logs message ([issue#16179](#), [pr#12753](#), runsisi)
- rbd: rbd-mirror: snap protect of non-layered image results in split-brain ([issue#16962](#), [pr#11869](#), Mykola Golub)
- rbd: [rbd-mirror] sporadic image replayer shut down failure ([issue#18441](#), [pr#13155](#), Jason Dillaman)
- rbd: rbd-nbd: disallow mapping images >2TB in size ([issue#17219](#), [pr#11870](#), Mykola Golub)
- rbd: rbd-nbd: invalid error code for “failed to read nbd request” messages ([issue#18242](#), [pr#12756](#), Mykola Golub)
- rbd: status json format has duplicated/overwritten key ([issue#18261](#), [pr#12741](#), Mykola Golub)
- rbd: TestLibRBD.DiscardAfterWrite doesn’t handle rbd_skip_partial_discard = true ([issue#17750](#), [pr#11853](#), Jason Dillaman)
- rbd: truncate can cause unflushed snapshot data lose ([issue#17193](#), [pr#12324](#), Yan, Zheng)
- : ReplicatedBackend: take read locks for clone sources during recovery ([issue#17831](#), [issue#18583](#), [pr#12978](#), Samuel Just)
- rgw: add option to log custom HTTP headers (rgw_log_http_headers) ([issue#18891](#), [pr#12490](#), Matt Benjamin)
- rgw: add suport for Swift-at-root dependent features of Swift API ([issue#18526](#), [issue#16673](#), [pr#11497](#), Pritha Srivastava, Radoslaw Zarzynski, Pete Zaitcev, Abhishek Lekshmanan)
- rgw: add support for the prefix parameter in account listing of Swift API ([issue#17931](#), [pr#12258](#), Radoslaw Zarzynski)
- rgw: Add workaround for upgrade issues for older jewel versions ([issue#17820](#), [pr#12316](#), Orit Wasserman)
- rgw: be aware about tenants on cls_user_bucket -> rgw_bucket conversion ([issue#18364](#), [issue#16355](#), [pr#13276](#), Radoslaw Zarzynski)
- rgw: bucket check remove _multipart_prefix ([issue#13724](#), [pr#11470](#), Weijun Duan)
- rgw: bucket resharding ([issue#17549](#), [issue#17550](#), [pr#13341](#), Yehuda Sadeh, Robin H. Johnson)
- rgw: disable virtual hosting of buckets when no hostnames are configured ([issue#17440](#), [issue#15975](#), [issue#17136](#), [pr#11760](#), Casey Bodley, Robin H. Johnson)
- rgw: do not abort when accept a CORS request with short origin ([issue#18187](#), [pr#12397](#), LiuYang)
- rgw: don’t store empty chains in gc ([issue#17897](#), [pr#12174](#), Yehuda Sadeh)
- rgw:fix for deleting objects name beginning and ending with underscores of one bucket using POST method of js sdk. ([issue#17888](#), [pr#12320](#), Casey Bodley)
- rgw: fix period update crash ([issue#18631](#), [pr#13273](#), Orit Wasserman)

- rgw: fix put_acls for objects starting and ending with underscore ([issue#17625](#), [pr#11675](#), Orit Wasserman)
- rgw: fix use of marker in List::list_objects() ([issue#18331](#), [pr#13358](#), Yehuda Sadeh)
- rgw: for the create_bucket api, if the input creation_time is zero, we ... ([issue#16597](#), [pr#11990](#), weiqiaomiao)
- rgw: Have a flavor of bucket deletion in radosgw-admin to bypass garbage collection ([issue#15557](#), [pr#10661](#), Pavan Rallabhandi)
- rgw: json encode/decode of RGWBucketInfo missing index_type field ([issue#17755](#), [pr#11759](#), Yehuda Sadeh)
- rgw: ldap: enforce simple_bind w/LDAPv3 redux ([issue#18339](#), [pr#12678](#), Weibing Zhang)
- rgw: leak from RGWMetaSyncShardCR::incremental_sync ([issue#18412](#), [issue#18300](#), [pr#13004](#), Casey Bodley, Sage Weil)
- rgw: leak in RGWFetchAllMetaCR ([issue#17812](#), [pr#11872](#), Casey Bodley)
- rgw: librgw: objects created from s3 apis are not visible from nfs mount point ([issue#18651](#), [pr#13177](#), Matt Benjamin)
- rgw: log name instead of id for SystemMetaObj on failure ([issue#15776](#), [pr#12622](#), Wido den Hollander, Abhishek Lekshmanan)
- rgw: multimds: mds entering up:replay and processing down mds aborts ([issue#17670](#), [pr#11857](#), Patrick Donnelly)
- rgw: multipart upload copy ([issue#12790](#), [pr#13068](#), Yehuda Sadeh, Javier M. Mellid, Matt Benjamin)
- rgw: multisite: after finishing full sync on a bucket, incremental sync starts over from the beginning ([issue#17661](#), [issue#17624](#), [pr#11864](#), Zengran Zhang, Casey Bodley)
- rgw: multisite: assert(next) failed in RGWMetaSyncCR ([issue#17044](#), [pr#11477](#), Casey Bodley)
- rgw: multisite: coroutine deadlock assertion on error in FetchAllMetaCR ([issue#17571](#), [pr#11866](#), Casey Bodley)
- rgw: multisite: coroutine deadlock in RGWMetaSyncCR after ECANCELED errors ([issue#17465](#), [pr#12738](#), Casey Bodley)
- rgw: multisite doesn't retry RGWFetchAllMetaCR on failed lease ([issue#17047](#), [pr#11476](#), Casey Bodley)
- rgw: multisite: ECANCELED & 500 error on bucket delete ([issue#17698](#), [pr#12044](#), Casey Bodley)
- rgw: multisite: failed assertion in 'radosgw-admin bucket sync status' ([issue#18083](#), [pr#12314](#), Casey Bodley)
- rgw: multisite: fix ref counting of completions ([issue#17792](#), [issue#18414](#), [issue#17793](#), [issue#18407](#), [pr#13001](#), Casey Bodley)
- rgw: multisite: metadata master can get the wrong value for 'oldest_log_period' ([issue#16894](#), [pr#11868](#), Casey Bodley)
- rgw: multisite: obsolete 'radosgw-admin period prepare' command ([issue#17387](#), [pr#11574](#), Gaurav Kumar Garg)
- rgw: multisite: race between ReadSyncStatus and InitSyncStatus leads to EIO errors ([issue#17568](#), [pr#11865](#), Casey Bodley)
- rgw: multisite requests failing with '400 Bad Request' with civetweb 1.8 ([issue#17822](#), [pr#12313](#), Casey Bodley)
- rgw: multisite: segfault after changing value of rgw_data_log_num_shards ([issue#18488](#), [pr#13180](#), Casey Bodley)
- rgw: multisite: sync status reports master is on a different period ([issue#18064](#), [pr#13175](#), Abhishek Lekshmanan)
- rgw: multisite upgrade from hammer -> jewel ignores rgw_region_root_pool ([issue#17963](#), [pr#12156](#), Casey Bodley)
- rgw: radosgw-admin period update reverts deleted zonegroup ([issue#17239](#), [pr#13171](#), Orit Wasserman)
- rgw: Realm set does not create a new period ([issue#18333](#), [pr#13182](#), Orit Wasserman)
- rgw: remove spurious mount entries for RGW buckets ([issue#17850](#), [pr#12045](#), Matt Benjamin)
- rgw: Replacing '+' with "%20" in canonical uri for s3 v4 auth. ([issue#17076](#), [pr#12542](#), Pritha Srivastava)
- rgw: rgw-admin: missing command to modify placement targets ([issue#18078](#), [pr#12428](#), Yehuda Sadeh, Casey Bodley)
- rgw: RGWRados::get_system_obj() sends unnecessary stat request before read ([issue#17580](#), [pr#11867](#), Casey Bodley)
- rgw: rgw_rest_s3: apply missed base64 try-catch ([issue#17663](#), [pr#11672](#), Matt Benjamin)
- rgw: RGW will not list Argonaut-era bucket via HTTP (but radosgw-admin works) ([issue#17372](#), [pr#11863](#), Yehuda Sadeh)
- rgw: sends omap_getvals with (u64)-1 limit ([issue#17985](#), [pr#12419](#), Yehuda Sadeh, Sage Weil)
- rgw: slave zonegroup cannot enable the bucket versioning ([issue#18003](#), [pr#13173](#), Orit Wasserman)
- rgw: TempURL properly handles accounts created with the implicit tenant ([issue#17961](#), [pr#12079](#), Radoslaw Zarzynski)
- rgw: the value of total_time is wrong in the result of 'radosgw-admin log show' opt ([issue#17598](#), [pr#11876](#), weiqiaomiao)
- rgw: Unable to commit period zonegroup change ([issue#17364](#), [pr#12315](#), Orit Wasserman)
- rgw: valgrind "invalid read size 4" RGWGetObj ([issue#18071](#), [pr#12997](#), Matt Benjamin)
- rgw: work around curl_multi_wait bug with non-blocking reads ([issue#15915](#), [issue#16368](#), [issue#16695](#), [pr#11627](#), John Coyle, Casey Bodley)
- tests: add require_jewel_osds before upgrading last hammer node ([issue#18719](#), [pr#13161](#), Nathan Cutler)
- tests: add require_jewel_osds to upgrade/hammer-x/tiering ([issue#18920](#), [pr#13404](#), Nathan Cutler)
- tests: assertion failure in a radosgw-admin related task ([issue#17167](#), [pr#12764](#), Orit Wasserman)
- tests: Cannot reserve CentOS 7.2 smithi machines ([issue#18416](#), [issue#18401](#), [pr#13050](#), Nathan Cutler, Sage Weil, Yuri Weinstein)
- tests: ignore bogus ceph-objectstore-tool error in ceph_manager ([issue#16263](#), [pr#13240](#), Nathan Cutler, Kefu Chai)
- tests: objecter_requests workunit fails on wip branches ([issue#18393](#), [pr#12761](#), Sage Weil)
- tests: qa/suites/upgrade/hammer-x: break stress split ec symlinks ([issue#19006](#), [pr#13533](#), Nathan Cutler)
- tests: qa/suites/upgrade/hammer-x/stress-split: finish thrashing before final upgrade ([issue#19004](#), [pr#13222](#), Sage Weil)
- tests: qa/tasks/ceph_deploy.py: use dev option ([issue#18736](#), [pr#13106](#), Vasu Kulkarni)
- tests: qa/workunits/rbd: use more recent qemu-io tests that support Xenial ([issue#18149](#), [issue#10773](#), [pr#13103](#), Jason Dillaman)
- tests: remove qa/suites/buildpackages ([issue#18846](#), [pr#13299](#), Loic Dachary)
- tests: SUSE yaml facets in qa/distros/all are out of date ([issue#18856](#), [issue#18846](#), [pr#13331](#), Nathan Cutler)

- tests: update rbd singleton/all/formatted-output.yaml to support ceph-ci ([issue#18440](#), [pr#12822](#), Nathan Cutler, Venky Shankar)
- tests: update Ubuntu image url after ceph.com refactor ([issue#18542](#), [pr#12959](#), Jason Dillaman)
- tests: upgrade:hammer-x: install firefly only on Ubuntu 14.04 ([issue#18089](#), [pr#13153](#), Nathan Cutler)
- tests: use ceph-jewel branch for s3tests ([issue#18384](#), [pr#12745](#), Nathan Cutler)
- tests: Workunits needlessly wget from git.ceph.com ([issue#18336](#), [issue#18271](#), [issue#18388](#), [pr#12686](#), Nathan Cutler, Sage Weil)
- test: temporarily disable fork()'ing tests ([issue#16556](#), [issue#17832](#), [pr#11953](#), John Spray)
- test: test fails due to The UNIX domain socket path ([issue#16014](#), [pr#12151](#), Loic Dachary)
- tools: ceph-disk: ceph-disk@.service races with ceph-osd@.service ([issue#17889](#), [issue#17813](#), [pr#12147](#), Loic Dachary)
- tools: ceph-disk -dmcrypt create must not require admin key ([issue#17849](#), [pr#12033](#), Loic Dachary)
- tools: ceph-disk prepare writes osd log 0 with root owner ([issue#18538](#), [pr#13025](#), Samuel Matzek)
- tools: crushtool -compile is create output despite of missing item ([issue#17306](#), [pr#11410](#), Kefu Chai)
- tools: rados bench seq must verify the hostname ([issue#17526](#), [pr#13049](#), Loic Dachary)
- tools: snapshotted RBD extent objects can't be manually evicted from a cache tier ([issue#17896](#), [pr#11968](#), Mingxin Liu)
- tools: systemd/ceph-disk: reduce ceph-disk flock contention ([issue#18049](#), [issue#13160](#), [pr#12210](#), David Disseldorp)

V10.2.5 JEWEL

This point release fixes an important [regression introduced in v10.2.4](#).

We recommend that all v10.2.x users upgrade.

NOTABLE CHANGES

For more detailed information, see [the complete changelog](#).

- msg/simple/Pipe: avoid returning 0 on poll timeout ([issue#18185](#), [pr#12376](#), Sage Weil)

V10.2.4 JEWEL

This point release fixes several important bugs in RBD mirroring, RGW multi-site, CephFS, and RADOS.

We recommend that all v10.2.x users upgrade. Also note the following when upgrading from hammer

UPGRADING FROM HAMMER

When the last hammer OSD in a cluster containing jewel MONs is upgraded to jewel, as of 10.2.4 the jewel MONs will issue this warning: “all OSDs are running jewel or later but the ‘require_jewel_osds’ osdmap flag is not set” and change the cluster health status to HEALTH_WARN.

This is a signal for the admin to do “ceph osd set require_jewel_osds” - by doing this, the upgrade path is complete and no more pre-Jewel OSDs may be added to the cluster.

NOTABLE CHANGES

For more detailed information, see [the complete changelog](#).

- build/ops: aarch64: Compiler-based detection of crc32 extended CPU type is broken ([issue#17516](#), [pr#11492](#), Alexander Graf)
- build/ops: allow building RGW with LDAP disabled ([issue#17312](#), [pr#11478](#), Daniel Gryniewicz)
- build/ops: backport ‘logrotate: Run as root/ceph’ ([issue#17381](#), [pr#11201](#), Boris Ranto)
- build/ops: ceph installs stuff in %_udevrulesdir but does not own that directory ([issue#16949](#), [pr#10862](#), Nathan Cutler)
- build/ops: ceph-osd-prestart.sh fails confusingly when data directory does not exist ([issue#17091](#), [pr#10812](#), Nathan Cutler)
- build/ops: disable LTTng-UST in openSUSE builds ([issue#16937](#), [pr#10794](#), Michel Normand)
- build/ops: i386 tarball gitbuilder failure on master ([issue#16398](#), [pr#10855](#), Vikhyat Umrao, Kefu Chai)
- build/ops: include more files in “make dist” tarball ([issue#17560](#), [pr#11431](#), Ken Dreyer)
- build/ops: incorrect value of CINIT_FLAG_DEFER_DROP_PRIVILEGES ([issue#16663](#), [pr#10278](#), Casey Bodley)
- build/ops: remove SYSTEMD_RUN from initscript ([issue#7627](#), [issue#16441](#), [issue#16440](#), [pr#9872](#), Vladislav Odintsov)
- build/ops: systemd: add install section to rbdmap.service file ([issue#17541](#), [pr#11158](#), Jelle vd Kooij)
- common: Enable/Disable of features is allowed even the features are already enabled/disabled ([issue#16079](#), [pr#11460](#),

Lu Shi)

- common: Log.cc: Assign LOG_INFO priority to syslog calls ([issue#15808](#), [pr#11231](#), Brad Hubbard)
- common: Proxied operations shouldn't result in error messages if replayed ([issue#16130](#), [pr#11461](#), Vikhyat Umrao)
- common: Request exclusive lock if owner sends -ENOTSUPP for proxied maintenance op ([issue#16171](#), [pr#10784](#), Jason Dillaman)
- common: msgr/async: Messenger thread long time lock hold risk ([issue#15758](#), [pr#10761](#), Wei Jin)
- doc: fix description for rsize and rasize ([issue#17357](#), [pr#11171](#), Andreas Gerstmayr)
- filestore: can get stuck in an unbounded loop during scrub ([issue#17859](#), [pr#12001](#), Sage Weil)
- fs: Failure in snaptest-git-ceph.sh ([issue#17172](#), [pr#11419](#), Yan, Zheng)
- fs: Log path as well as ino when detecting metadata damage ([issue#16973](#), [pr#11418](#), John Spray)
- fs: client: FAILED assert(root_ancestor->qtree == __null) ([issue#16066](#), [issue#16067](#), [pr#10107](#), Yan, Zheng)
- fs: client: add missing client_lock for get_root ([issue#17197](#), [pr#10921](#), Patrick Donnelly)
- fs: client: fix shutdown with open inodes ([issue#16764](#), [pr#10958](#), John Spray)
- fs: client: nlink count is not maintained correctly ([issue#16668](#), [pr#10877](#), Jeff Layton)
- fs: multimds: allow_multimds not required when max_mds is set in ceph.conf at startup ([issue#17105](#), [pr#10997](#), Patrick Donnelly)
- librados: memory leaks from ceph::crypto (WITH_NSS) ([issue#17205](#), [pr#11409](#), Casey Bodley)
- librados: modify Pipe::connect() to return the error code ([issue#15308](#), [pr#11193](#), Vikhyat Umrao)
- librados: remove new setxattr overload to avoid breaking the C++ ABI ([issue#18058](#), [pr#12207](#), Josh Durgin)
- librbd: cannot disable journaling or remove non-mirrored, non-primary image ([issue#16740](#), [pr#11337](#), Jason Dillaman)
- librbd: discard after write can result in assertion failure ([issue#17695](#), [pr#11644](#), Jason Dillaman)
- librbd::Operations: update notification failed: (2) No such file or directory ([issue#17549](#), [pr#11420](#), Jason Dillaman)
- mds: Crash in Client::_invalidate_kernel_dcache when reconnecting during unmount ([issue#17253](#), [pr#11414](#), Yan, Zheng)
- mds: Duplicate damage table entries ([issue#17173](#), [pr#11412](#), John Spray)
- mds: Failure in dirfrag.sh ([issue#17286](#), [pr#11416](#), Yan, Zheng)
- mds: Failure in snaptest-git-ceph.sh ([issue#17271](#), [pr#11415](#), Yan, Zheng)
- mon: Ceph Status - Segmentation Fault ([issue#16266](#), [pr#11408](#), Brad Hubbard)
- mon: Display full flag in ceph status if full flag is set ([issue#15809](#), [pr#9388](#), Vikhyat Umrao)
- mon: Error EINVAL: removing mon.a at 172.21.15.16:6789/0, there will be 1 monitors ([issue#17725](#), [pr#12267](#), Joao Eduardo Luis)
- mon: OSDMonitor: only reject MOSDBoot based on up_from if inst matches ([issue#17899](#), [pr#12067](#), Samuel Just)
- mon: OSDMonitor: Missing nearfull flag set ([issue#17390](#), [pr#11272](#), Igor Podoski)
- mon: Upgrading 0.94.6 -> 0.94.9 saturating mon node networking ([issue#17365](#), [issue#17386](#), [pr#11679](#), Sage Weil, xie xinggauo)
- mon: ceph mon Segmentation fault after set crush_ruleset ceph 10.2.2 ([issue#16653](#), [pr#10861](#), song baisen)
- mon: crash: crush/CrushWrapper.h: 940: FAILED assert(successful_detach) ([issue#16525](#), [pr#10496](#), Kefu Chai)
- mon: don't crash on invalid standby_for_fscid ([issue#17466](#), [pr#11389](#), John Spray)
- mon: fix missing osd metadata (again) ([issue#17685](#), [pr#11642](#), John Spray)
- mon: osdmonitor: decouple adjust_heartbeat_grace and min_down_reporters ([issue#17055](#), [pr#10757](#), Zengran Zhang)
- mon: the %USED of ceph df is wrong ([issue#16933](#), [pr#10860](#), Kefu Chai)
- osd: condition OSDMap encoding on features ([issue#18015](#), [pr#12167](#), Sage Weil)
- osd: PG::_update_calc_stats wrong for CRUSH_ITEM_NONE up set items ([issue#16998](#), [pr#10883](#), Samuel Just)
- osd: PG::choose_acting valgrind error or ./common/hobject.h: 182: FAILED assert(!max || (*this == hobject_t(hobject_t::get_max()))) ([issue#13967](#), [pr#10885](#), Tao Chang)
- osd: Potential crash during journal::Replay shut down ([issue#16433](#), [pr#10645](#), Jason Dillaman)
- osd: add peer_addr in heartbeat_check log message ([issue#15762](#), [pr#9739](#), Vikhyat Umrao, Sage Weil)
- osd: adjust scrub boundary to object without SnapSet ([issue#17470](#), [pr#11311](#), Samuel Just)
- osd: ceph osd df does not show summarized info correctly if one or more OSDs are out ([issue#16706](#), [pr#10759](#), xie xinggauo)
- osd: journal: do not prematurely flag object recorder as closed ([issue#17590](#), [pr#11634](#), Jason Dillaman)
- osd: mark_all_unfound_lost() leaves unapplied changes ([issue#16156](#), [pr#10886](#), Samuel Just)
- osd: segfault in ObjectCacher::FlusherThread ([issue#16610](#), [pr#10864](#), Yan, Zheng)
- qa: remove EnumerateObjects from librados upgrade tests ([pr#11728](#), Josh Durgin)
- rbd: Disabling pool mirror mode with registered peers results orphaned mirrored images ([issue#16984](#), [pr#10857](#), Jason Dillaman)
- rbd: ImageWatcher: use after free within C_UnwatchAndFlush ([issue#17289](#), [issue#17254](#), [pr#11466](#), Jason Dillaman)
- rbd: Prevent the creation of a clone from a non-primary mirrored image ([issue#16449](#), [pr#10650](#), Mykola Golub)
- rbd: RBD should restrict mirror enable/disable actions on parents/clones ([issue#16056](#), [pr#11459](#), zhuangzeqiang)
- rbd: TestJournalReplay: sporadic assert(m_state == STATE_READY || m_state == STATE_STOPPING) failure ([issue#17566](#), [pr#11590](#), Jason Dillaman)
- rbd: bench io-size should not be larger than image size ([issue#16967](#), [pr#10796](#), Jason Dillaman)
- rbd: ceph 10.2.2 rbd status on image format 2 returns (2) No such file or directory ([issue#16887](#), [pr#10652](#), Jason Dillaman)
- rbd: helgrind: TestLibRBD.TestIOPP potential deadlock closing an image with read-ahead enabled ([issue#17198](#),

pr#11463, Jason Dillaman)

- rbd: image.stat() call in librbdpy fails sometimes ([issue#17310](#), [pr#11464](#), Jason Dillaman)
- rbd: krbd qa scripts and concurrent.sh test fix ([issue#17223](#), [pr#11018](#), Ilya Dryomov)
- rbd: krbd-related CLI patches ([issue#17554](#), [pr#11400](#), Ilya Dryomov)
- rbd: mirror: improve resiliency of stress test case ([issue#16855](#), [issue#16555](#), [issue#14738](#), [issue#15259](#), [issue#17446](#), [issue#17355](#), [issue#16538](#), [issue#16974](#), [issue#17283](#), [issue#17317](#), [issue#17416](#), [issue#16227](#), [pr#11433](#), Mykola Golub, Ricardo Dias, Jason Dillaman)
- rbd: rbd-nbd IO hang ([issue#16921](#), [pr#11467](#), Jason Dillaman)
- rbd: update_features API needs to support backwards/forward compatibility ([issue#17330](#), [pr#11462](#), Jason Dillaman)
- rgw: COPY broke multipart files uploaded under dumpling ([issue#16435](#), [pr#10866](#), Yehuda Sadeh)
- rgw: Config parameter rgw keystone make new tenants in radosgw multitenancy does not work ([issue#17293](#), [pr#11473](#), SirishaGuduru)
- rgw: Do not archive metadata by default ([issue#17256](#), [pr#11321](#), Pavan Rallabhandi, Matt Benjamin)
- rgw: ERROR: got unexpected error when trying to read object: -2 ([issue#17111](#), [pr#11472](#), Yang Honggang)
- rgw: Modification for TEST S3 ACCESS section in INSTALL CEPH OBJECT GATEWAY page ([issue#15603](#), [pr#11475](#), la-sguduru)
- rgw: RGW loses realm/period/zonegroup/zone data: period overwritten if somewhere in the cluster is still running Hammer ([issue#17371](#), [pr#11519](#), Orit Wasserman)
- rgw: RGWDataSyncCR fails on errors from RGWListBucketIndexesCR ([issue#17073](#), [pr#11330](#), Casey Bodley)
- rgw: S3 object versioning fails when applied on a non-master zone ([issue#16494](#), [pr#11367](#), Yehuda Sadeh)
- rgw: add orphan options to radosgw-admin -help and man page ([issue#17281](#), [issue#17280](#), [pr#11139](#), Ken Dreyer, Thomas Serlin)
- rgw: back off bucket sync on failures, don't store marker ([issue#16742](#), [pr#11021](#), Yehuda Sadeh)
- rgw: combined LDAP backports ([issue#17544](#), [issue#17185](#), [pr#11332](#), Harald Klein, Matt Benjamin)
- rgw: cors auto memleak ([issue#16564](#), [pr#10656](#), Yan Jun)
- rgw: default quota fixes ([issue#16410](#), [pr#10832](#), Pavan Rallabhandi, Daniel Gryniewicz)
- rgw: doc: description of multipart part entity is wrong ([issue#17504](#), [pr#11342](#), weiqiaomiao)
- rgw: don't loop forever when reading data from 0 sized segment. ([issue#17692](#), [pr#11626](#), Marcus Watts)
- rgw: fix put_acls for objects starting and ending with underscore ([issue#17625](#), [pr#11669](#), Orit Wasserman)
- rgw: fix regression with handling double underscore ([issue#17443](#), [issue#16856](#), [pr#11563](#), Yehuda Sadeh, Orit Wasserman)
- rgw: handle empty POST condition ([issue#17635](#), [pr#11662](#), Yehuda Sadeh)
- rgw: metadata sync can skip markers for failed/incomplete entries ([issue#16759](#), [pr#10657](#), Yehuda Sadeh)
- rgw: nfs backports ([issue#17393](#), [issue#17311](#), [issue#17367](#), [issue#17319](#), [issue#17321](#), [issue#17322](#), [issue#17323](#), [issue#17325](#), [issue#17326](#), [issue#17327](#), [pr#11335](#), Min Chen, Yan Jun, Weibing Zhang, Matt Benjamin)
- rgw: period commit loses zonegroup changes: region_map converted repeatedly ([issue#17051](#), [pr#10890](#), Casey Bodley)
- rgw: period commit return error when the current period has a zonegroup which doesn't have a master zone ([issue#17110](#), [pr#10867](#), weiqiaomiao)
- rgw: radosgw daemon core when reopen logs ([issue#17036](#), [pr#10868](#), weiqiaomiao)
- rgw: rgw file uses too much CPU in gc/idle thread ([issue#16976](#), [pr#10889](#), Matt Benjamin)
- rgw: s3tests-test-readwrite failing with 500 ([issue#16930](#), [pr#11471](#), Yehuda Sadeh)
- rgw: upgrade from old multisite to new multisite fails ([issue#16751](#), [pr#10891](#), Orit Wasserman)
- rgw: response information is error when getting token of swift account ([issue#15195](#), [pr#11474](#), Qiankun Zheng)
- rgw: user email can modify to empty when it has values ([issue#13286](#), [pr#11469](#), Yehuda Sadeh, Weijun Duan)
- tests: ceph-disk must ignore debug monc ([issue#17607](#), [pr#11548](#), Loic Dachary)
- tests: fix TestCIsRbd.mirror_image failure in upgrade:jewel-x-master-distro-basic-vps ([issue#16529](#), [pr#10888](#), Jason Dillaman)
- tests: scsi_debug fails /dev/disk/by-partuuid ([issue#17100](#), [pr#11411](#), Loic Dachary)
- tests: test/ceph_test_msgr: do not use Message::middle for holding transient... ([issue#17365](#), [issue#17728](#), [issue#16955](#), [pr#11742](#), Haomai Wang, Kefu Chai, Michal Jarzabek, Sage Weil)
- tools: Missing comma in ceph-create-keys causes concatenation of arguments ([issue#17815](#), [pr#11822](#), Patrick Donnelly)
- tools: add a tool to rebuild mon store from OSD ([issue#17179](#), [issue#17400](#), [pr#11126](#), Kefu Chai, xie xingguo)
- tools: ceph-create-keys: sometimes blocks forever if mds allow is set ([issue#16255](#), [pr#11417](#), John Spray)
- tools: ceph-disk should timeout when a lock cannot be acquired ([issue#16580](#), [pr#10758](#), Loic Dachary)
- tools: ceph-disk: expected systemd unit failures are confusing ([issue#15990](#), [pr#10884](#), Boris Ranto)
- tools: ceph-disk: using a regular file as a journal fails ([issue#16280](#), [issue#17662](#), [pr#11657](#), Jayashree Candadai, Anirudha Bose, Loic Dachary, Shylesh Kumar)
- tools: ceph-objectstore-tool crashes if --journal-path <a-directory> ([issue#17307](#), [pr#11407](#), Kefu Chai)
- tools: ceph-objectstore-tool: add a way to split filestore directories offline ([issue#17220](#), [pr#11252](#), Josh Durgin)
- tools: ceph-post-file: use new ssh key ([issue#14267](#), [pr#11746](#), David Galloway)

This point release fixes several important bugs in RBD mirroring, RGW multi-site, CephFS, and RADOS.

We recommend that all v10.2.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build/ops: 60-ceph-partuuid-workaround-rules still needed by debian jessie (udev 215-17) ([issue#16351](#), [pr#10653](#), runsisi, Loic Dachary)
- build/ops: ceph Resource Agent does not work with systemd ([issue#14828](#), [pr#9917](#), Nathan Cutler)
- build/ops: ceph-base requires parted ([issue#16095](#), [pr#10008](#), Ken Dreyer)
- build/ops: ceph-osd-prestart.sh contains Upstart-specific code ([issue#15984](#), [pr#10364](#), Nathan Cutler)
- build/ops: mount.ceph: move from ceph-base to ceph-common and add symlink in /sbin for SUSE ([issue#16598](#), [issue#16645](#), [pr#10357](#), Nathan Cutler, Dan Horák, Ricardo Dias, Kefu Chai)
- build/ops: need rocksdb commit 7ca731b12ce for ppc64le build ([issue#17092](#), [pr#10816](#), Nathan Cutler)
- build/ops: rpm: OBS needs ExclusiveArch ([issue#16936](#), [pr#10614](#), Michel Normand)
- cli: ceph command line tool chokes on ceph -w (the dash is unicode ‘en dash’ &ndash, copy-paste to reproduce) ([issue#12287](#), [pr#10420](#), Oleh Prypin, Kefu Chai)
- common: expose buffer const_iterator symbols ([issue#16899](#), [pr#10552](#), Noah Watkins)
- common: global-init: fixup chown of the run directory along with log and asok files ([issue#15607](#), [pr#8754](#), Karol Mroz)
- fs: ceph-fuse: link to libtcmalloc or jemalloc ([issue#16655](#), [pr#10303](#), Yan, Zheng)
- fs: client: crash in umount when fuse_use_invalidate_cb is enabled ([issue#16137](#), [pr#10106](#), Yan, Zheng)
- fs: client: fstat cap release ([issue#15723](#), [pr#9562](#), Yan, Zheng, Noah Watkins)
- fs: essential backports for OpenStack Manila ([issue#15406](#), [issue#15614](#), [issue#15615](#), [pr#10453](#), John Spray, Ramana Raja, Xiaoxi Chen)
- fs: fix double-unlock on shutdown ([issue#17126](#), [pr#10847](#), Greg Farnum)
- fs: fix mdsmap print_summary with standby replays ([issue#15705](#), [pr#9547](#), John Spray)
- fs: fuse mounted file systems fails SAMBA CTDB ping_pong rw test with v9.0.2 ([issue#12653](#), [issue#15634](#), [pr#10108](#), Yan, Zheng)
- librados: Missing export for rados_aio_get_version in src/include/rados/librados.h ([issue#15535](#), [pr#9574](#), Jim Wright)
- librados: osd: bad flags can crash the osd ([issue#16012](#), [pr#9997](#), Sage Weil)
- librbd: Close journal and object map before flagging exclusive lock as released ([issue#16450](#), [pr#10053](#), Jason Dillaman)
- librbd: Crash when utilizing advisory locking API functions ([issue#16364](#), [pr#10051](#), Jason Dillaman)
- librbd: ExclusiveLock object leaked when switching to snapshot ([issue#16446](#), [pr#10054](#), Jason Dillaman)
- librbd: FAILED assert(object_no < m_object_map.size()) ([issue#16561](#), [pr#10647](#), Jason Dillaman)
- librbd: Image removal doesn’t necessarily clean up all rbd_mirroring entries ([issue#16471](#), [pr#10009](#), Jason Dillaman)
- librbd: Object map/fast-diff invalidated if journal replays the same snap remove event ([issue#16350](#), [pr#10010](#), Jason Dillaman)
- librbd: Timeout sending mirroring notification shouldn’t result in failure ([issue#16470](#), [pr#10052](#), Jason Dillaman)
- librbd: Whitelist EBUSY error from snap unprotect for journal replay ([issue#16445](#), [pr#10055](#), Jason Dillaman)
- librbd: cancel all tasks should wait until finisher is done ([issue#16517](#), [pr#9752](#), Haomai Wang)
- librbd: delay acquiring lock if image watch has failed ([issue#16923](#), [pr#10827](#), Jason Dillaman)
- librbd: fix missing return statement if failed to get mirror image state ([issue#16600](#), [pr#10144](#), runsisi)
- librbd: flag image as updated after proxying maintenance op ([issue#16404](#), [pr#9883](#), Jason Dillaman)
- librbd: mkfs.xfs slow performance with discards and object map ([issue#16707](#), [issue#16689](#), [pr#10649](#), Jason Dillaman)
- librbd: potential use after free on refresh error ([issue#16519](#), [pr#9952](#), Mykola Golub)
- librbd: rbd-nbd does not properly handle resize notifications ([issue#15715](#), [pr#10679](#), Mykola Golub)
- librbd: the option ‘rbd_cache_writethrough_until_flush=true’ doesn’t work ([issue#16740](#), [issue#16386](#), [issue#16708](#), [issue#16654](#), [issue#16478](#), [pr#10797](#), Mykola Golub, xinxin shu, Xiaowei Chen, Jason Dillaman)
- mds: tell command blocks forever with async messenger (TestVolumeClient.test_evict_client failure) ([issue#16288](#), [pr#10501](#), Douglas Fuller)
- mds: Confusing MDS log message when shut down with stalled journaler reads ([issue#15689](#), [pr#9557](#), John Spray)
- mds: Deadlock on shutdown active rank while busy with metadata IO ([issue#16042](#), [pr#10502](#), Patrick Donnelly)
- mds: Failing file operations on kernel based cephfs mount point leaves unaccessible file behind on hammer 0.94.7 ([issue#16013](#), [pr#10199](#), Yan, Zheng)
- mds: Fix shutting down mds timed-out due to deadlock ([issue#16396](#), [pr#10500](#), Zhi Zhang)
- mds: MDSMonitor fixes ([issue#16136](#), [pr#9561](#), xie xinguo)
- mds: MDSMonitor::check_subs() is very buggy ([issue#16022](#), [pr#10103](#), Yan, Zheng)
- mds: Session::check_access() is buggy ([issue#16358](#), [pr#10105](#), Yan, Zheng)
- mds: StrayManager.cc: 520: FAILED assert(dnl->is_primary()) ([issue#15920](#), [pr#9559](#), Yan, Zheng)
- mds: enforce a dirfrag limit on entries ([issue#16164](#), [pr#10104](#), Patrick Donnelly)
- mds: fix SnapRealm::have_past_parents_open() ([issue#16299](#), [pr#10499](#), Yan, Zheng)
- mds: fix setattr starve setattr ([issue#16154](#), [pr#9560](#), Yan, Zheng)
- mds: wrongly treat symlink inode as normal file/dir when symlink inode is stale on kcephfs ([issue#15702](#), [pr#9405](#), Zhi

Zhang)

- mon: “mon metadata” fails when only one monitor exists ([issue#15866](#), [pr#10654](#), John Spray, Kefu Chai)
- mon: Monitor: validate prefix on handle_command() ([issue#16297](#), [pr#10036](#), You Ji)
- mon: OSDMonitor: drop pg temps from not the current primary ([issue#16127](#), [pr#9998](#), Samuel Just)
- mon: prepare_pgtemp needs to only update up_thru if newer than the existing one ([issue#16185](#), [pr#10001](#), Samuel Just)
- msgr: AsyncConnection::lockmsg/async lockdep cycle: AsyncMessenger::lock, MDSDaemon::mds_lock, AsyncConnection::lock ([issue#16237](#), [pr#10004](#), Haomai Wang)
- msgr: async messenger mon crash ([issue#16378](#), [issue#16418](#), [pr#9996](#), Haomai Wang)
- msgr: backports of all asyncmsgr fixes to jewel ([issue#15503](#), [issue#15372](#), [pr#9633](#), Yan Jun, Haomai Wang, Piotr Dałek)
- msgr: msg/async: connection race hang ([issue#15849](#), [pr#10003](#), Haomai Wang)
- osd: FileStore: umount hang because sync thread doesn’t exit ([issue#15695](#), [pr#9105](#), Kefu Chai)
- osd: Fixes for list-inconsistent-* ([issue#15766](#), [issue#16192](#), [issue#15719](#), [pr#9565](#), David Zafman)
- osd: New pools have bogus stuck inactive/unclean HEALTH_ERR messages until they are first active and clean ([issue#14952](#), [pr#10007](#), Sage Weil)
- osd: OSD crash with Hammer to Jewel Upgrade: void FileStore::init_temp_collections() ([issue#16672](#), [pr#10561](#), David Zafman)
- osd: OSD failed to subscribe skipped osdmaps after ceph osd pause ([issue#17023](#), [pr#10804](#), Kefu Chai)
- osd: ObjectCacher split BufferHead read fix ([issue#16002](#), [pr#10074](#), Greg Farnum)
- osd: ReplicatedBackend doesn’t increment stats on pull, only push ([issue#16277](#), [pr#10421](#), Kefu Chai)
- osd: Scrub error: 0/1 pinned ([issue#15952](#), [pr#9576](#), Samuel Just)
- osd: crash adding snap to purged_snaps in ReplicatedPG::WaitingOnReplicas ([issue#15943](#), [pr#9575](#), Samuel Just)
- osd: partprobe intermittent issues during ceph-disk prepare ([issue#15176](#), [pr#10497](#), Marius Vollmer, Loic Dachary)
- osd: saw valgrind issues in ReplicatedPG::new_repop ([issue#16801](#), [pr#10760](#), Kefu Chai)
- osd: sparse_read on ec pool should return extends with correct offset ([issue#16138](#), [pr#10006](#), kofiliu)
- osd:sched_time not actually randomized ([issue#15890](#), [pr#9578](#), xie xingguo)
- rbd: ImageReplayer::is_replaying does not include flush state ([issue#16970](#), [pr#10790](#), Jason Dillaman)
- rbd: Journal duplicate op detection can cause lockdep error ([issue#16363](#), [pr#10044](#), Jason Dillaman)
- rbd: Journal needs to handle duplicate maintenance op tids ([issue#16362](#), [pr#10045](#), Jason Dillaman)
- rbd: Unable to disable journaling feature if in unexpected mirror state ([issue#16348](#), [pr#10042](#), Jason Dillaman)
- rbd: bashism in src/rbdmap ([issue#16608](#), [pr#10786](#), Jason Dillaman)
- rbd: doc: format 2 now is the default image format ([issue#17026](#), [pr#10732](#), Chengwei Yang)
- rbd: hen journaling is enabled, a flush request shouldn’t flush the cache ([issue#15761](#), [pr#10041](#), Yuan Zhou)
- rbd: possible race condition during journal transition from replay to ready ([issue#16198](#), [pr#10047](#), Jason Dillaman)
- rbd: qa/workunits/rbd: respect RBD_CREATE_ARGS environment variable ([issue#16289](#), [pr#9721](#), Mykola Golub)
- rbd: rbd-mirror should disable proxied maintenance ops for non-primary image ([issue#16411](#), [pr#10050](#), Jason Dillaman)
- rbd: rbd-mirror: FAILED assert(m_local_image_ctx->object_map != nullptr) ([issue#16558](#), [pr#10646](#), Jason Dillaman)
- rbd: rbd-mirror: FAILED assert(m_on_update_status_finish == nullptr) ([issue#16956](#), [pr#10792](#), Jason Dillaman)
- rbd: rbd-mirror: FAILED assert(m_state == STATE_STOPPING) ([issue#16980](#), [pr#10791](#), Jason Dillaman)
- rbd: rbd-mirror: ensure replay status formatter has completed before stopping replay ([issue#16352](#), [pr#10043](#), Jason Dillaman)
- rbd: rbd-mirror: include local pool id in resync throttle unique key ([issue#16536](#), [issue#15239](#), [issue#16488](#), [issue#16491](#), [issue#16329](#), [issue#15108](#), [issue#15670](#), [pr#10678](#), Ricardo Dias, Jason Dillaman)
- rbd: rbd-mirror: potential race condition accessing local image journal ([issue#16230](#), [pr#10046](#), Jason Dillaman)
- rbd: rbd-mirror: reduce memory footprint during journal replay ([issue#16321](#), [issue#16489](#), [issue#16622](#), [issue#16539](#), [issue#16223](#), [issue#16349](#), [pr#10684](#), Mykola Golub, Jason Dillaman)
- rgw: A query on a static large object fails with 404 error ([issue#16015](#), [pr#9544](#), Radoslaw Zarzynski)
- rgw: Add zone rename to radosgw_admin ([issue#16934](#), [pr#10663](#), Shilpa Jagannath)
- rgw: Bucket index shards orphaned after bucket delete ([issue#16412](#), [pr#10525](#), Orit Wasserman)
- rgw: Bug when using port 443s in rgw. ([issue#16548](#), [pr#10664](#), Pritha Srivastava)
- rgw: Fallback to Host header for bucket name. ([issue#15975](#), [pr#10693](#), Robin H. Johnson)
- rgw: Fix civetweb IPv6 ([issue#16928](#), [pr#10580](#), Robin H. Johnson)
- rgw: Increase log level for messages occurring while running rgw admin command ([issue#16935](#), [pr#10765](#), Shilpa Jagannath)
- rgw: No Last-Modified, Content-Size and X-Object-Manifest headers if no segments in DLO manifest ([issue#15812](#), [pr#9265](#), Radoslaw Zarzynski)
- rgw: RGWPeriodPuller tries to pull from itself ([issue#16939](#), [pr#10764](#), Casey Bodley)
- rgw: Set Access-Control-Allow-Origin to a Asterisk if allowed in a rule ([issue#15348](#), [pr#9453](#), Wido den Hollander)
- rgw: Swift API returns double space usage and objects of account metadata ([issue#16188](#), [pr#10148](#), Albert Tu)
- rgw: account/container metadata not actually present in a request are deleted during POST through Swift API ([issue#15977](#), [issue#15779](#), [pr#9542](#), Radoslaw Zarzynski)
- rgw: add socket backlog setting for via ceph.conf ([issue#16406](#), [pr#10216](#), Feng Guo)
- rgw: add tenant support to multisite sync ([issue#16469](#), [issue#16121](#), [issue#16665](#), [pr#10845](#), Yehuda Sadeh, Josh Durgin, Casey Bodley, Pritha Srivastava)

- rgw: add_zone only clears master_zone if -master=false ([issue#15901](#), pr#9327, Casey Bodley)
- rgw: aws4 parsing issue ([issue#15940](#), [issue#15939](#), pr#9545, Yehuda Sadeh)
- rgw: aws4: add STREAMING-AWS4-HMAC-SHA256-PAYLOAD support ([issue#16146](#), pr#10167, Radoslaw Zarzynski, Javier M. Mellid)
- rgw: backport merge of static sites fixes ([issue#15555](#), [issue#15532](#), [issue#15531](#), pr#9568, Robin H. Johnson)
- rgw: can set negative max_buckets on RGWUserInfo ([issue#14534](#), pr#10655, Yehuda Sadeh)
- rgw: cleanup radosgw-admin temp command as it was deprecated ([issue#16023](#), pr#9390, Vikhyat Umrao)
- rgw: comparing return code to ERR_NOT_MODIFIED in rgw_rest_s3.cc (needs minus sign) ([issue#16327](#), pr#9790, Nathan Cutler)
- rgw: custom metadata aren't camelcased in Swift's responses ([issue#15902](#), pr#9267, Radoslaw Zarzynski)
- rgw: data sync stops after getting error in all data log sync shards ([issue#16530](#), pr#10073, Yehuda Sadeh)
- rgw: default zone and zonegroup cannot be added to a realm ([issue#16839](#), pr#10658, Casey Bodley)
- rgw: document multi tenancy ([issue#16635](#), pr#10217, Pete Zaitcev)
- rgw: don't unregister request if request is not connected to manager ([issue#15911](#), pr#9242, Yehuda Sadeh)
- rgw: failed to create bucket after upgrade from hammer to jewel ([issue#16627](#), pr#10524, Orit Wasserman)
- rgw: fix ldap bindpw parsing ([issue#16286](#), pr#10518, Matt Benjamin)
- rgw: fix multi-delete query param parsing. ([issue#16618](#), pr#10188, Robin H. Johnson)
- rgw: improve support for Swift's object versioning. ([issue#15925](#), pr#10710, Radoslaw Zarzynski)
- rgw: initial slashes are not properly handled in Swift's BulkDelete ([issue#15948](#), pr#9316, Radoslaw Zarzynski)
- rgw: master: build failures with boost > 1.58 ([issue#16392](#), [issue#16391](#), pr#10026, Abhishek Lekshmanan)
- rgw: multisite segfault on ~RGWRealmWatcher if realm was deleted ([issue#16817](#), pr#10660, Casey Bodley)
- rgw: multisite sync races with deletes ([issue#16222](#), [issue#16464](#), [issue#16220](#), [issue#16143](#), pr#10293, Yehuda Sadeh, Casey Bodley)
- rgw: multisite: preserve zone's extra pool ([issue#16712](#), pr#10537, Abhishek Lekshmanan)
- rgw: object expirer's hints might be trimmed without processing in some circumstances ([issue#16705](#), [issue#16684](#), pr#10763, Radoslaw Zarzynski)
- rgw: radosgw-admin failure for user create after upgrade from hammer to jewel ([issue#15937](#), pr#9294, Orit Wasserman, Abhishek Lekshmanan)
- rgw: radosgw-admin: EEXIST messages for create operations ([issue#15720](#), pr#9268, Abhishek Lekshmanan)
- rgw: radosgw-admin: inconsistency in uid/email handling ([issue#13598](#), pr#10520, Matt Benjamin)
- rgw: realm pull fails when using apache frontend ([issue#15846](#), pr#9266, Orit Wasserman)
- rgw: retry on bucket sync errors ([issue#16108](#), pr#9425, Yehuda Sadeh)
- rgw: s3website: x-amz-website-redirect-location header returns malformed HTTP response ([issue#15531](#), pr#9099, Robin H. Johnson)
- rgw: segfault in RGWOp_MDLog_Notify ([issue#16666](#), pr#10662, Casey Bodley)
- rgw: segmentation fault on error_repo in data sync ([issue#16603](#), pr#10523, Casey Bodley)
- rgw: selinux denials in RGW ([issue#16126](#), pr#10519, Boris Ranto)
- rgw: support size suffixes for -max-size in radosgw-admin command ([issue#16004](#), pr#9743, Vikhyat Umrao)
- rgw: updating CORS/ACLs might not work in some circumstances ([issue#15976](#), pr#9543, Radoslaw Zarzynski)
- rgw: use zone endpoints instead of zonegroup endpoints ([issue#16834](#), pr#10659, Casey Bodley)
- tests: improve rbd-mirror test case coverage ([issue#16197](#), pr#9631, Mykola Golub, Jason Dillaman)
- tests: rados/test.sh workunit timeout on OpenStack ([issue#15403](#), pr#8904, Loic Dachary)
- tools: ceph-disk: Accept bcache devices as data disks ([issue#13278](#), pr#8497, Peter Sabaini)
- tools: rados: Add cleanup message with time to rados bench output ([issue#15704](#), pr#9740, Vikhyat Umrao)
- tools: src/script/subman fails with KeyError: 'nband' ([issue#16961](#), pr#10625, Loic Dachary, Ali Maredia)

V10.2.2 JEWEL

This point release fixes several important bugs in RBD mirroring, RGW multi-site, CephFS, and RADOS.

We recommend that all v10.2.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- ceph: cli: exception when pool name has non-ascii characters ([issue#15913](#), pr#9320, Ricardo Dias)
- ceph-disk: workaround gperftool hang ([issue#13522](#), [issue#16103](#), pr#9427, Loic Dachary)
- cephfs: backports needed for Manila ([issue#15599](#), [issue#15417](#), [issue#15045](#), pr#9430, John Spray, Ramana Raja, Xiaoxi Chen)
- ceph.spec.in: drop support for RHEL<7 and SUSE<1210 in jewel and above ([issue#15725](#), [issue#15627](#), [issue#13445](#), [issue#15822](#), [issue#15472](#), [issue#15987](#), [issue#15516](#), [issue#15549](#), pr#8938, Boris Ranto, Sage Weil, Nathan Cutler, Lars Marowsky-Bree)
- ceph_test_librbd_fsx crashes during journal replay shut down ([issue#16123](#), pr#9556, Jason Dillaman)

- client: fix bugs accidentally disabling readahead ([issue#16024](#), [pr#9656](#), Patrick Donnelly, Greg Farnum)
- cls_journal: initialize empty commit position upon client register ([issue#15757](#), [pr#9376](#), runsisi, Venky Shankar)
- cls::rbd: mirror_image_status_list returned max 64 items ([pr#9069](#), Mykola Golub)
- cls_rbd: mirror image status summary should read full directory ([issue#16178](#), [pr#9608](#), Jason Dillaman)
- common: BackoffThrottle spins unnecessarily with very small backoff while the throttle is full ([issue#15953](#), [pr#9579](#), Samuel Just)
- common: Do not link ltng into libglobal ([pr#9194](#), Karol Mroz)
- debian: install systemd target files ([issue#15573](#), [pr#8815](#), Kefu Chai, Sage Weil)
- doc: update mirroring guide to include pool/image status commands ([issue#15746](#), [pr#9180](#), Mykola Golub)
- librbd: Disabling journaling feature results in "Transport endpoint is not connected" error ([issue#15863](#), [pr#9548](#), Yuan Zhou)
- librbd: do not shut down exclusive lock while acquiring' ([issue#16291](#), [issue#16260](#), [pr#9691](#), Jason Dillaman)
- librbd: Initial python APIs to support mirroring ([issue#15656](#), [pr#9550](#), Mykola Golub)
- librbd: journal IO error results in failed assertion in AioCompletion ([issue#16077](#), [issue#15034](#), [issue#15791](#), [pr#9611](#), Hector Martin, Jason Dillaman)
- librbd: journal: live replay might skip entries from previous object set ([issue#15864](#), [issue#15665](#), [pr#9217](#), Jason Dillaman)
- librbd: journal: support asynchronous shutdown ([issue#15949](#), [issue#14530](#), [issue#15993](#), [pr#9373](#), Jason Dillaman)
- librbd: Metadata config overrides are applied synchronously ([issue#15928](#), [pr#9318](#), Jason Dillaman)
- librbd: Object Map is showing as invalid, even when Object Map is disabled for that Image. ([issue#16076](#), [pr#9555](#), xinxin shu)
- librbd: prevent error messages when journal externally disabled ([issue#16114](#), [pr#9610](#), Zhiqiang Wang, Jason Dillaman)
- librbd: recursive lock possible when disabling journaling ([issue#16235](#), [pr#9654](#), Jason Dillaman)
- librbd: refresh image if needed in mirror functions ([issue#16096](#), [pr#9609](#), Jon Bernard)
- librbd: remove should ignore mirror errors from older OSDs ([issue#16268](#), [pr#9692](#), Jason Dillaman)
- librbd: reuse ImageCtx::finisher and SafeTimer for lots of images case ([issue#13938](#), [pr#9580](#), Haomai Wang)
- librbd: validate image metadata configuration overrides ([issue#15522](#), [pr#9554](#), zhuangzeqiang)
- mds: order directories by hash and fix simultaneous readdir races ([issue#15508](#), [pr#9655](#), Yan, Zheng, Greg Farnum)
- mon: Hammer (0.94.3) OSD does not delete old OSD Maps in a timely fashion (maybe at all?) ([issue#13990](#), [pr#9100](#), Kefu Chai)
- mon/Monitor: memory leak on Monitor::handle_ping() ([issue#15793](#), [pr#9270](#), xie xingguo)
- osd: acting_primary not updated on split ([issue#15523](#), [pr#8968](#), Sage Weil)
- osd: boot race with noup being set ([issue#15678](#), [pr#9101](#), Sage Weil)
- osd: deadlock in OSD::_committed_osd_maps ([issue#15701](#), [pr#9103](#), Xinze Chi)
- osd: hobject_t::get_max() vs is_max() discrepancy ([issue#16113](#), [pr#9614](#), Samuel Just)
- osd: LibRadosWatchNotifyPPTests/LibRadosWatchNotifyPP.WatchNotify2Timeout/1 segv ([issue#15760](#), [pr#9104](#), Sage Weil)
- osd: remove reliance on FLAG_OMAP for reads ([pr#9638](#), Samuel Just)
- osd valgrind invalid reads/writes ([issue#15870](#), [pr#9237](#), Samuel Just)
- pybind: rbd API should default features parameter to None ([issue#15982](#), [pr#9553](#), Mykola Golub)
- qa: dynamic_features.sh races with image deletion ([issue#15500](#), [pr#9552](#), Mykola Golub)
- qa/workunits: ensure replay has started before checking position ([issue#16248](#), [pr#9674](#), Jason Dillaman)
- qa/workunits/rbd: fixed rbd_mirror teuthology runtime errors ([pr#9232](#), Jason Dillaman)
- radosgw-admin: fix 'period push' handling of -url ([issue#15926](#), [pr#9210](#), Casey Bodley)
- rbd-mirror: Delete local image mirror when remote image mirroring is disabled ([issue#15916](#), [issue#14421](#), [pr#9372](#), runsisi, Mykola Golub, Ricardo Dias)
- rbd-mirror: do not propagate deletions when pool unavailable ([issue#16229](#), [pr#9630](#), Jason Dillaman)
- rbd-mirror: do not re-use image id from mirror directory if creating image ([issue#16253](#), [pr#9673](#), Jason Dillaman)
- rbd-mirror: FAILED assert(!m_status_watcher) ([issue#16245](#), [issue#16290](#), [pr#9690](#), Mykola Golub)
- rbd-mirror: fix deletion propagation edge cases ([issue#16226](#), [pr#9629](#), Jason Dillaman)
- rbd-mirror: fix journal shut down ordering ([issue#16165](#), [pr#9628](#), Jason Dillaman)
- rbd-mirror: potential crash during image status update ([issue#15909](#), [pr#9226](#), Mykola Golub, Jason Dillaman)
- rbd-mirror: refresh image after creating sync point ([issue#16196](#), [pr#9627](#), Jason Dillaman)
- rbd-mirror: replicate cloned images ([issue#14937](#), [pr#9423](#), Jason Dillaman)
- rbd-mirror should disable the rbd cache for local images ([issue#15930](#), [pr#9317](#), Jason Dillaman)
- rbd-mirror: support bootstrap canceling ([issue#16201](#), [pr#9612](#), Mykola Golub)
- rbd-mirror: support multiple replicated pools ([issue#16045](#), [pr#9409](#), Jason Dillaman)
- rgw: fix manager selection when APIs customized ([issue#15974](#), [issue#15973](#), [pr#9245](#), Robin H. Johnson)
- rgw: keep track of written_objs correctly ([issue#15886](#), [pr#9239](#), Yehuda Sadeh)
- rpm: ceph gid mismatch on upgrade from hammer with pre-existing ceph user (SUSE) ([issue#15869](#), [pr#9424](#), Nathan Cutler)
- systemd: ceph-{mds,mon,osd,radosgw} systemd unit files need wants=time-sync.target ([issue#15419](#), [pr#8802](#), Nathan Cutler)
- test: failure in journal.sh workunit test ([issue#16011](#), [pr#9377](#), Mykola Golub)
- tests: rm -fr /tmp/virtualenv ([issue#16087](#), [pr#9403](#), Loic Dachary)

This is the first bugfix release for Jewel. It contains several annoying packaging and init system fixes and a range of important bugfixes across RBD, RGW, and CephFS.

We recommend that all v10.2.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- cephfs: CephFSVolumeClient should isolate volumes by RADOS namespace ([issue#15400](#), [pr#8787](#), Xiaoxi Chen)
- cephfs: handle standby-replay nodes properly in upgrades ([issue#15591](#), [pr#8971](#), John Spray)
- ceph-{mds,mon,osd} packages need scriptlets with systemd code ([issue#14941](#), [pr#8801](#), Boris Ranto, Nathan Cutler)
- ceph_test_keyvaluedb: fix ([issue#15435](#), [pr#9051](#), Allen Samuels, Sage Weil)
- cmake: add missing source file to rbd_mirror/image_replayer ([pr#9052](#), Casey Bodley)
- cmake: fix rbd compile errors ([pr#9076](#), runsisi, Jason Dillaman)
- journal: incorrectly computed object offset within set ([issue#15765](#), [pr#9038](#), Jason Dillaman)
- librbd: client-side handling for incompatible object map sizes ([issue#15642](#), [pr#9039](#), Jason Dillaman)
- librbd: constrain size of AioWriteEvent journal entries ([issue#15750](#), [pr#9048](#), Jason Dillaman)
- librbd: does not crash if image header is too short ([pr#9044](#), Kefu Chai)
- librbd: Errors encountered disabling object-map while flatten is in-progress ([issue#15572](#), [pr#8869](#), Jason Dillaman)
- librbd: fix get/list mirror image status API ([issue#15771](#), [pr#9036](#), Mykola Golub)
- librbd: Parent image is closed twice if error encountered while opening ([issue#15574](#), [pr#8867](#), Jason Dillaman)
- librbd: possible double-free of object map invalidation request upon error ([issue#15643](#), [pr#8865](#), runsisi)
- librbd: possible race condition leads to use-after-free ([issue#15690](#), [pr#9009](#), Jason Dillaman)
- librbd: potential concurrent event processing during journal replay ([issue#15755](#), [pr#9040](#), Jason Dillaman)
- librbd: Potential double free of SetSnapRequest instance ([issue#15571](#), [pr#8803](#), runsisi)
- librbd: put the validation of image snap context earlier ([pr#9046](#), runsisi)
- librbd: reduce log level for image format 1 warning ([issue#15577](#), [pr#9003](#), Jason Dillaman)
- mds/MDSAuthCap parse no longer fails on paths with hyphens ([issue#15465](#), [pr#8969](#), John Spray)
- mds: MDS incarnation no longer gets lost after remove filesystem ([issue#15399](#), [pr#8970](#), John Spray)
- mon/OSDMonitor: avoid underflow in reweight-by-utilization if max_change=1 ([issue#15655](#), [pr#9006](#), Samuel Just)
- python: clone operation will fail if config overridden with “rbd default format = 1” ([issue#15685](#), [pr#8972](#), Jason Dillaman)
- radosgw-admin: add missing -zonegroup-id to usage ([issue#15650](#), [pr#9019](#), Casey Bodley)
- radosgw-admin: update usage for zone[group] modify ([issue#15651](#), [pr#9016](#), Casey Bodley)
- radosgw-admin: zonegroup remove command ([issue#15684](#), [pr#9015](#), Casey Bodley)
- rbd CLI to retrieve rbd mirror state for a pool / specific image ([issue#15144](#), [issue#14420](#), [pr#8868](#), Mykola Golub)
- rbd disk-usage CLI command should support calculating full image usage ([issue#14540](#), [pr#8870](#), Jason Dillaman)
- rbd: helpful error message on map failure ([issue#15721](#), [pr#9041](#), Venky Shankar)
- rbd: help message distinction between commands and aliases ([issue#15521](#), [pr#9004](#), Yongqiang He)
- rbd-mirror: admin socket commands to start/stop/restart mirroring ([issue#15718](#), [pr#9010](#), Mykola Golub, Josh Durgin)
- rbd-mirror can crash if start up is interrupted ([issue#15630](#), [pr#8866](#), Jason Dillaman)
- rbd-mirror: image sync needs to handle snapshot size and protection status ([issue#15110](#), [pr#9050](#), Jason Dillaman)
- rbd-mirror: lockdep error during bootstrap ([issue#15664](#), [pr#9008](#), Jason Dillaman)
- rbd-nbd: fix rbd-nbd aio callback error handling ([issue#15604](#), [pr#9005](#), Chang-Yi Lee)
- rgw: add AWS4 completion support for RGW_OP_SET_BUCKET_WEBSITE ([issue#15626](#), [pr#9018](#), Javier M. Mellid)
- rgw admin output ([issue#15747](#), [pr#9054](#), Casey Bodley)
- rgw: fix issue #15597 ([issue#15597](#), [pr#9020](#), Yehuda Sadeh)
- rgw: fix printing wrong X-Storage-Url in Swift’s TempAuth. ([issue#15667](#), [pr#9021](#), Radoslaw Zarzynski)
- rgw: handle stripe transition when flushing final pending_data_bl ([issue#15745](#), [pr#9053](#), Yehuda Sadeh)
- rgw: leak fixes ([issue#15792](#), [pr#9022](#), Yehuda Sadeh)
- rgw: multisite: Issues with Deleting Buckets ([issue#15540](#), [pr#8930](#), Abhishek Lekshmanan)
- rgw: period commit fix ([issue#15828](#), [pr#9081](#), Casey Bodley)
- rgw: period delete fixes ([issue#15469](#), [pr#9047](#), Casey Bodley)
- rgw: radosgw-admin zone set cuts pool names short if name starts with a period ([issue#15598](#), [pr#9029](#), Yehuda Sadeh)
- rgw: segfault at RGWASyncGetSystemObj ([issue#15565](#), [issue#15625](#), [pr#9017](#), Yehuda Sadeh)
- several backports ([issue#15588](#), [issue#15655](#), [pr#8853](#), Alexandre Derumier, xie xingguo, Alfredo Deza)
- systemd: fix typo in preset file ([pr#8843](#), Nathan Cutler)
- tests: make check fails on ext4 ([issue#15837](#), [pr#9063](#), Loic Dachary, Sage Weil)

This major release of Ceph is the foundation for the next long-term stable release series. There have been many major changes since the Infernalis (9.2.x) and Hammer (0.94.x) releases, and the upgrade process is non-trivial. Please read these release notes carefully.

MAJOR CHANGES FROM INFERNALIS

- *CephFS:*
 - This is the first release in which CephFS is declared stable! Several features are disabled by default, including snapshots and multiple active MDS servers.
 - The repair and disaster recovery tools are now feature-complete.
 - A new cephfs-volume-manager module is included that provides a high-level interface for creating “shares” for OpenStack Manila and similar projects.
 - There is now experimental support for multiple CephFS file systems within a single cluster.
- *RGW:*
 - The multisite feature has been almost completely rearchitected and rewritten to support any number of clusters/sites, bidirectional fail-over, and active/active configurations.
 - You can now access radosgw buckets via NFS (experimental).
 - The AWS4 authentication protocol is now supported.
 - There is now support for S3 request payer buckets.
 - The new multitenancy infrastructure improves compatibility with Swift, which provides a separate container namespace for each user/tenant.
 - The OpenStack Keystone v3 API is now supported. There are a range of other small Swift API features and compatibility improvements as well, including bulk delete and SLO (static large objects).
- *RBD:*
 - There is new support for mirroring (asynchronous replication) of RBD images across clusters. This is implemented as a per-RBD image journal that can be streamed across a WAN to another site, and a new rbd-mirror daemon that performs the cross-cluster replication.
 - The exclusive-lock, object-map, fast-diff, and journaling features can be enabled or disabled dynamically. The deep-flatten features can be disabled dynamically but not re-enabled.
 - The RBD CLI has been rewritten to provide command-specific help and full bash completion support.
 - RBD snapshots can now be renamed.
- *RADOS:*
 - BlueStore, a new OSD backend, is included as an experimental feature. The plan is for it to become the default backend in the K or L release.
 - The OSD now persists scrub results and provides a librados API to query results in detail.
 - We have revised our documentation to recommend *against* using ext4 as the underlying filesystem for Ceph OSD daemons due to problems supporting our long object name handling.

MAJOR CHANGES FROM HAMMER

- *General:*
 - Ceph daemons are now managed via systemd (with the exception of Ubuntu Trusty, which still uses upstart).
 - Ceph daemons run as ‘ceph’ user instead of ‘root’.
 - On Red Hat distros, there is also an SELinux policy.
- *RADOS:*
 - The RADOS cache tier can now proxy write operations to the base tier, allowing writes to be handled without forcing migration of an object into the cache.
 - The SHEC erasure coding support is no longer flagged as experimental. SHEC trades some additional storage space for faster repair.
 - There is now a unified queue (and thus prioritization) of client IO, recovery, scrubbing, and snapshot trimming.
 - There have been many improvements to low-level repair tooling (ceph-objectstore-tool).
 - The internal ObjectStore API has been significantly cleaned up in order to facilitate new storage backends like BlueStore.
- *RGW:*
 - The Swift API now supports object expiration.
 - There are many Swift API compatibility improvements.
- *RBD:*
 - The rbd du command shows actual usage (quickly, when object-map is enabled).
 - The object-map feature has seen many stability improvements.
 - The object-map and exclusive-lock features can be enabled or disabled dynamically.
 - You can now store user metadata and set persistent librbd options associated with individual images.
 - The new deep-flatten features allow flattening of a clone and all of its snapshots. (Previously snapshots could not be flattened.)

- The export-diff command is now faster (it uses aio). There is also a new fast-diff feature.
- The -size argument can be specified with a suffix for units (e.g., --size 64G).
- There is a new rbd status command that, for now, shows who has the image open/mapped.
- CephFS:
 - You can now rename snapshots.
 - There have been ongoing improvements around administration, diagnostics, and the check and repair tools.
 - The caching and revocation of client cache state due to unused inodes has been dramatically improved.
 - The ceph-fuse client behaves better on 32-bit hosts.

DISTRO COMPATIBILITY

Starting with Infernalis, we have dropped support for many older distributions so that we can move to a newer compiler toolchain (e.g., C++11). Although it is still possible to build Ceph on older distributions by installing backported development tools, we are not building and publishing release packages for ceph.com.

We now build packages for the following distributions and architectures:

- x86_64:
 - CentOS 7.x. We have dropped support for CentOS 6 (and other RHEL 6 derivatives, like Scientific Linux 6).
 - Debian Jessie 8.x. Debian Wheezy 7.x's g++ has incomplete support for C++11 (and no systemd).
 - Ubuntu Xenial 16.04 and Trusty 14.04. Ubuntu Precise 12.04 is no longer supported.
 - Fedora 22 or later.
- aarch64 / arm64:
 - Ubuntu Xenial 16.04.

UPGRADING FROM INFERNALIS OR HAMMER

- We now recommend against using ext4 as the underlying file system for Ceph OSDs, especially when RGW or other users of long RADOS object names are used. For more information about why, please see [Filesystem Recommendations](#).

If you have an existing cluster that uses ext4 for the OSDs but uses only RBD and/or CephFS, then the ext4 limitations will not affect you. Before upgrading, be sure add the following to ceph.conf to allow the OSDs to start:

```
osd max object name len = 256
osd max object namespace len = 64
```

Keep in mind that if you set these lower object name limits and later decide to use RGW on this cluster, it will have problems storing S3/Swift objects with long names. This startup check can also be disabled via the below option, although this is not recommended:

```
osd check max object name len on startup = false
```

- There are no major compatibility changes since Infernalis. Simply upgrading the daemons on each host and restarting all daemons is sufficient.
- The rbd CLI no longer accepts the deprecated ‘–image-features’ option during create, import, and clone operations. The ‘–image-feature’ option should be used instead.
- The rbd legacy image format (version 1) is deprecated with the Jewel release. Attempting to create a new version 1 RBD image will result in a warning. Future releases of Ceph will remove support for version 1 RBD images.
- The ‘send_pgCreates’ and ‘map_pgCreates’ mon CLI commands are obsolete and no longer supported.
- A new configure option ‘mon_election_timeout’ is added to specifically limit max waiting time of monitor election process, which was previously restricted by ‘mon_lease’.
- CephFS filesystems created using versions older than Firefly (0.80) must use the new ‘cephfs-data-scan tmap_upgrade’ command after upgrading to Jewel. See ‘Upgrading’ in the CephFS documentation for more information.
- The ‘ceph mds setmap’ command has been removed.
- The default RBD image features for new images have been updated to enable the following: exclusive lock, object map, fast-diff, and deep-flatten. These features are not currently supported by the RBD kernel driver nor older RBD clients. They can be disabled on a per-image basis via the RBD CLI, or the default features can be updated to the pre-Jewel

setting by adding the following to the client section of the Ceph configuration file:

```
rbd default features = 1
```

- The rbd legacy image format (version 1) is deprecated with the Jewel release.
- After upgrading, users should set the ‘sortbitwise’ flag to enable the new internal object sort order:

```
ceph osd set sortbitwise
```

This flag is important for the new object enumeration API and for new backends like BlueStore.

- The rbd CLI no longer permits creating images and snapshots with potentially ambiguous names (e.g. the ‘/’ and ‘@’ characters are disallowed). The validation can be temporarily disabled by adding “-rbd-validate-names=false” to the rbd CLI when creating an image or snapshot. It can also be disabled by adding the following to the client section of the Ceph configuration file:

```
rbd validate names = false
```

UPGRADING FROM HAMMER

- All cluster nodes must first upgrade to Hammer v0.94.4 or a later v0.94.z release; only then is it possible to upgrade to Jewel 10.2.z.
- For all distributions that support systemd (CentOS 7, Fedora, Debian Jessie 8.x, OpenSUSE), ceph daemons are now managed using native systemd files instead of the legacy sysvinit scripts. For example:

```
systemctl start ceph.target      # start all daemons
systemctl status ceph-osd@12     # check status of osd.12
```

The main notable distro that is *not* yet using systemd is Ubuntu trusty 14.04. (The next Ubuntu LTS, 16.04, will use systemd instead of upstart.)

- Ceph daemons now run as user and group ceph by default. The ceph user has a static UID assigned by Fedora and Debian (also used by derivative distributions like RHEL/CentOS and Ubuntu). On SUSE the same UID/GID as in Fedora and Debian will be used, *provided it is not already assigned*. In the unlikely event the preferred UID or GID is assigned to a different user/group, ceph will get a dynamically assigned UID/GID.

If your systems already have a ceph user, upgrading the package will cause problems. We suggest you first remove or rename the existing ‘ceph’ user and ‘ceph’ group before upgrading.

When upgrading, administrators have two options:

1. Add the following line to ceph.conf on all hosts:

```
setuser match path = /var/lib/ceph/$type/$cluster-$id
```

This will make the Ceph daemons run as root (i.e., not drop privileges and switch to user ceph) if the daemon’s data directory is still owned by root. Newly deployed daemons will be created with data owned by user ceph and will run with reduced privileges, but upgraded daemons will continue to run as root.

2. Fix the data ownership during the upgrade. This is the preferred option, but it is more work and can be very time consuming. The process for each host is to:

1. Upgrade the ceph package. This creates the ceph user and group. For example:

```
ceph-deploy install --stable jewel HOST
```

2. Stop the daemon(s):

```
service ceph stop          # fedora, centos, rhel, debian  
stop ceph-all              # ubuntu
```

3. Fix the ownership:

```
chown -R ceph:ceph /var/lib/ceph  
chown -R ceph:ceph /var/log/ceph
```

4. Restart the daemon(s):

```
start ceph-all             # ubuntu  
systemctl start ceph.target # debian, centos, fedora, rhel
```

Alternatively, the same process can be done with a single daemon type, for example by stopping only monitors and chowning only /var/lib/ceph/mon.

- The on-disk format for the experimental KeyValueStore OSD backend has changed. You will need to remove any OSDs using that backend before you upgrade any test clusters that use it.
- When a pool quota is reached, librados operations now block indefinitely, the same way they do when the cluster fills up. (Previously they would return -ENOSPC.) By default, a full cluster or pool will now block. If your librados application can handle ENOSPC or EDQUOT errors gracefully, you can get error returns instead by using the new librados OPERATION_FULL_TRY flag.
- The return code for librbd's rbd_aio_read and Image::aio_read API methods no longer returns the number of bytes read upon success. Instead, it returns 0 upon success and a negative value upon failure.
- ‘ceph scrub’, ‘ceph compact’ and ‘ceph sync force’ are now DEPRECATED. Users should instead use ‘ceph mon scrub’, ‘ceph mon compact’ and ‘ceph mon sync force’.
- ‘ceph mon_metadata’ should now be used as ‘ceph mon metadata’. There is no need to deprecate this command (same major release since it was first introduced).
- The *-dump-json* option of “osdmaptool” is replaced by *-dump json*.
- The commands of “pg ls-by-{pool,primary,osd}” and “pg ls” now take “recovering” instead of “recovery”, to include the recovering pgs in the listed pgs.

UPGRADING FROM FIREFLY

Upgrading directly from Firefly v0.80.z is not recommended. It is possible to do a direct upgrade, but not without downtime, as all OSDs must be stopped, upgraded, and then restarted. We recommend that clusters be first upgraded to Hammer v0.94.6 or a later v0.94.z release; only then is it possible to upgrade to Jewel 10.2.z for an online upgrade (see below).

To do an offline upgrade directly from Firefly, all Firefly OSDs must be stopped and marked down before any Jewel OSDs will be allowed to start up. This fencing is enforced by the Jewel monitor, so you should use an upgrade procedure like:

1. Upgrade Ceph on monitor hosts
2. Restart all ceph-mon daemons
3. Set noout::

```
ceph osd set noout
```

4. Upgrade Ceph on all OSD hosts
5. Stop all ceph-osd daemons
6. Mark all OSDs down with something like::

```
ceph osd down seq 0 1000
```

7. Start all ceph-osd daemons
8. Let the cluster settle and then unset noout::

```
ceph osd unset noout
```

9. Upgrade and restart any remaining daemons (ceph-mds, radosgw)

- admin/build-doc: depend on zlib1g-dev and graphviz ([pr#7522](#), Ken Dreyer)
- auth: fail if rotating key is missing (do not spam log) ([pr#6473](#), Qiankun Zheng)
- auth: fix crash when bad keyring is passed ([pr#6698](#), Dunrong Huang)
- auth: make keyring without mon entity type return -EACCES ([pr#5734](#), Xiaowei Chen)
- AUTHORS: update email ([pr#7854](#), Yehuda Sadeh)
- authtool: update --help and manpage to match code. ([pr#8456](#), Robin H. Johnson)
- bluestore: latest and greatest ([issue#14210](#), [issue#13801](#), [pr#6896](#), xie.xingguo, Jianpeng Ma, YiQiang Chen, Sage Weil, Ning Yao)
- buffer: add symmetry operator==() and operator!=() ([pr#7974](#), Kefu Chai)
- buffer: fix internal iterator invalidation on rebuild, get_contiguous ([pr#6962](#), Sage Weil)
- buffer: hide iterator_impl symbols ([issue#14788](#), [pr#7688](#), Kefu Chai)
- buffer: increment history alloc as well in raw_combined ([issue#14955](#), [pr#7910](#), Samuel Just)
- buffer: make usable outside of ceph source again ([pr#6863](#), Josh Durgin)
- buffer: raw_combined allocations buffer and ref count together ([pr#7612](#), Sage Weil)
- buffer: use move construct to append/push_back/push_front ([pr#7455](#), Haomai Wang)
- build: Adding build requires ([pr#7742](#), Erwan Velu)
- build: a few armhf (32-bit build) fixes ([pr#7999](#), Eric Lee, Sage Weil)
- build: allow jemalloc with rocksdb-static ([pr#7368](#), Somnath Roy)
- build: build internal plugins and classes as modules ([pr#6462](#), James Page)
- build: cmake check fixes ([pr#6787](#), Orit Wasserman)
- build: cmake tweaks ([pr#6254](#), John Spray)
- build: fix a few warnings ([pr#6847](#), Orit Wasserman)
- build: fix bz2-dev dependency ([pr#6948](#), Samuel Just)
- build: fix compiling warnings ([pr#8366](#), Dongsheng Yang)
- build: Fixing BTRFS issue at 'make check' ([pr#7805](#), Erwan Velu)
- build: fix Jenkins make check errors due to deep-scrub randomization ([pr#6671](#), David Zafman)
- build: fix the autotools and cmake build (the new fusestore needs libfuse) ([pr#7393](#), Kefu Chai)
- build: fix warnings ([pr#7197](#), Kefu Chai, xie xingguo)
- build: fix warnings ([pr#7315](#), Kefu Chai)
- build: FreeBSD related fixes ([pr#7170](#), Mykola Golub)
- build: Gentoo: _FORTIFY_SOURCE fix. ([issue#13920](#), [pr#6739](#), Robin H. Johnson)
- build: kill warnings ([pr#7397](#), Kefu Chai)
- build: misc make check fixes ([pr#7153](#), Sage Weil)
- build: more CMake package check fixes ([pr#6108](#), Daniel Gryniewicz)
- build: move libexec scripts to standardize across distros ([issue#14687](#), [issue#14705](#), [issue#14723](#), [pr#7636](#), Nathan Cutler, Kefu Chai)
- build/ops: enable CR in CentOS 7 ([issue#13997](#), [pr#6844](#), Loic Dachary)
- build/ops: rbd-replay moved from ceph-test-dbg to ceph-common-dbg ([issue#13785](#), [pr#6578](#), Loic Dachary)
- build/ops: systemd ceph-disk unit must not assume /bin/flock ([issue#13975](#), [pr#6803](#), Loic Dachary)
- build: Refrain from versioning and packaging EC testing plugins ([issue#14756](#), [issue#14723](#), [pr#7637](#), Nathan Cutler, Kefu Chai)
- build: Respect TMPDIR for virtualenv. ([pr#8457](#), Robin H. Johnson)
- build: spdk submodule; cmake ([pr#7503](#), Kefu Chai)
- build: workaround an automake bug for "make check" ([issue#14723](#), [pr#7626](#), Kefu Chai)
- ceph: bash auto complete for CLI based on mon command descriptions ([pr#7693](#), Adam Kupczyk)
- ceph_daemon.py: Resolved ImportError to work with python3 ([pr#7937](#), Sarthak Munshi)
- ceph-detect-init: add debian/jessie test ([pr#8074](#), Kefu Chai)
- ceph-detect-init: add missing test case ([pr#8105](#), Nathan Cutler)
- ceph-detect-init: fix py3 test ([pr#7025](#), Kefu Chai)
- ceph-detect-init: fix py3 test ([pr#7243](#), Kefu Chai)
- ceph_detect_init/_init__.py: remove shebang ([pr#7731](#), Nathan Cutler)
- ceph-detect-init: return correct value on recent SUSE distros ([issue#14770](#), [pr#7909](#), Nathan Cutler)
- ceph-detect-init/run-tox.sh: FreeBSD: No init detect ([pr#8373](#), Willem Jan Withagen)
- ceph-detect-init: Ubuntu >= 15.04 uses systemd ([pr#6873](#), James Page)
- ceph-disk: Add destroy and deactivate option ([issue#7454](#), [pr#5867](#), Vicente Cheng)
- ceph-disk: add -f flag for btrfs mkfs ([pr#7222](#), Darrell Enns)
- ceph-disk: Add -setuser and -setgroup options for ceph-disk ([pr#7351](#), Mike Shuey)
- ceph-disk: ceph-disk list fails on /dev/cciss!c0d0 ([issue#13970](#), [issue#14233](#), [issue#14230](#), [pr#6879](#), Loic Dachary)
- ceph-disk: compare parted output with the dereferenced path ([issue#13438](#), [pr#6219](#), Joe Julian)
- ceph-disk: deactivate / destroy PATH arg are optional ([pr#7756](#), Loic Dachary)
- ceph-disk: do not always fail when re-using a partition ([pr#8508](#), You Ji)
- ceph-disk: fix failures when preparing disks with udev > 214 ([issue#14080](#), [issue#14094](#), [pr#6926](#), Loic Dachary, Ilya Dryomov)
- ceph-disk: fix prepare --help ([pr#7758](#), Loic Dachary)
- ceph-disk: Fix trivial typo ([pr#7472](#), Brad Hubbard)

- ceph-disk: flake8 fixes ([pr#7646](#), Loic Dachary)
- ceph-disk: get Nonetype when ceph-disk list with -format plain on single device. ([pr#6410](#), Vicente Cheng)
- ceph-disk: Improving ‘make check’ for ceph-disk ([pr#7762](#), Erwan Velu)
- ceph-disk: key management support ([issue#14669](#), [pr#7552](#), Loic Dachary)
- ceph-disk: make some arguments as required if necessary ([pr#7687](#), Dongsheng Yang)
- ceph-disk: s/dmcrpyt/dmcrypt/ ([issue#14838](#), [pr#7744](#), Loic Dachary, Frode Sandholtbraaten)
- ceph-disk: support bluestore ([issue#13422](#), [pr#7218](#), Loic Dachary, Sage Weil)
- ceph-disk/test: fix test_prepare.py::TestPrepare tests ([pr#7549](#), Kefu Chai)
- ceph-disk: warn for prepare partitions with bad GUIDs ([issue#13943](#), [pr#6760](#), David Disseldorp)
- ceph: fix tell behavior ([pr#6329](#), David Zafman)
- cephfs-data-scan: scan_frags ([pr#5941](#), John Spray)
- cephfs-data-scan: scrub tag filtering (#12133 and #12145) ([issue#12133](#), [issue#12145](#), [pr#5685](#), John Spray)
- ceph-fuse: add process to ceph-fuse -help ([pr#6821](#), Wei Feng)
- ceph-fuse: fix double decreasing the count to trim caps ([issue#14319](#), [pr#7229](#), Zhi Zhang)
- ceph-fuse: fix double free of args ([pr#7015](#), Ilya Shipitsin)
- ceph-fuse: fix fsync() ([pr#6388](#), Yan, Zheng)
- ceph-fuse: Fix potential filehandle ref leak at umount ([issue#14800](#), [pr#7686](#), Zhi Zhang)
- ceph-fuse,libcephfs: Fix client handling of “lost” open directories on shutdown ([issue#14996](#), [pr#7994](#), Yan, Zheng)
- ceph-fuse,libcephfs: fix free fds being exhausted eventually because freed fds are never put back ([issue#14798](#), [pr#7685](#), Zhi Zhang)
- ceph-fuse:print usage information when no parameter specified ([pr#6868](#), Bo Cai)
- ceph-fuse: rotate log file ([pr#8485](#), Sage Weil)
- ceph-fuse: While starting ceph-fuse, start the log thread first ([issue#13443](#), [pr#6224](#), Wenjun Huang)
- ceph: improve the error message ([issue#11101](#), [pr#7106](#), Kefu Chai)
- ceph.in: avoid a broken pipe error when use ceph command ([issue#14354](#), [pr#7212](#), Bo Cai)
- ceph.in: correct dev python path for automake builds ([pr#8360](#), Josh Durgin)
- ceph.in: fix python libpath for automake as well ([pr#8362](#), Josh Durgin)
- ceph.in: Minor python3 specific changes ([pr#7947](#), Sarthak Munshi)
- ceph-kvstore-tool: handle bad out file on command line ([pr#6093](#), Kefu Chai)
- ceph-mds:add -help/-h ([pr#6850](#), Cilang Zhao)
- ceph_objectstore_bench: fix race condition, bugs ([issue#13516](#), [pr#6681](#), Igor Fedotov)
- ceph-objectstore-tool: fix -dry-run for many ceph-objectstore-tool operations ([pr#6545](#), David Zafman)
- ceph-rest-api: fix fs/flag/set ([pr#8428](#), Sage Weil)
- ceph.spec.in: add BuildRequires: systemd ([issue#13860](#), [pr#6692](#), Nathan Cutler)
- ceph.spec.in: add copyright notice ([issue#14694](#), [pr#7569](#), Nathan Cutler)
- ceph.spec.in: add license declaration ([pr#7574](#), Nathan Cutler)
- ceph.spec.in: disable lttng and babeltrace explicitly ([issue#14844](#), [pr#7857](#), Kefu Chai)
- ceph.spec.in: do not install Ceph RA on systemd platforms ([issue#14828](#), [pr#7894](#), Nathan Cutler)
- ceph.spec.in: fix openldap and openssl build dependencies for SUSE ([issue#15138](#), [pr#8120](#), Nathan Cutler)
- ceph.spec.in: limit _smp_mflags when lowmem_builder is set in SUSE’s OBS ([issue#13858](#), [pr#6691](#), Nathan Cutler)
- ceph.spec.in: terminate if statement in %pre scriptlet ([pr#8628](#), Nathan Cutler)
- ceph_test_libcephfs: tolerate duplicated entries in readdir ([issue#14377](#), [pr#7246](#), Yan, Zheng)
- ceph_test_msgr: reduce test size to fix memory size ([pr#8127](#), Haomai Wang)
- ceph_test_msgr: Use send_message instead of keepalive to wakeup connection ([pr#6605](#), Haomai Wang)
- ceph_test_rados: make long name ~300 chars, (not ~800) ([pr#8600](#), Sage Weil)
- ceph_test_rados_misc: shorten mount timeout ([pr#8209](#), Sage Weil)
- check-generated.sh: can’t source bash from sh ([pr#8521](#), Michal Jarzabek)
- cleanup ([pr#8058](#), Yehuda Sadeh, Orit Wasserman)
- cleanup: remove misc dead code ([pr#7201](#), Erwan Velu)
- client: a better check for MDS availability ([pr#6253](#), John Spray)
- client: add option to control how directory size is calculated ([pr#7323](#), Yan, Zheng)
- client: avoid creating orphan object in Client::check_pool_perm() ([issue#13782](#), [pr#6603](#), Yan, Zheng)
- client: check if Fh is readable when processing a read ([issue#11517](#), [pr#7209](#), Yan, Zheng)
- client: close mds sessions in shutdown() ([pr#6269](#), John Spray)
- client: don’t invalidate page cache when inode is no longer used ([pr#6380](#), Yan, Zheng)
- client: don’t mark_down on command reply ([pr#6204](#), John Spray)
- client: drop prefix from ints ([pr#6275](#), John Coyle)
- client: flush kernel pagecache before creating snapshot ([issue#10436](#), [pr#7495](#), Yan, Zheng)
- client: modify a word in log ([pr#6906](#), YongQiang He)
- client: properly trim unlinked inode ([issue#13903](#), [pr#7297](#), Yan, Zheng)
- client: removed unused Mutex from MetaRequest ([pr#7655](#), Greg Farnum)
- client: sys/file.h includes for flock operations ([pr#6282](#), John Coyle)
- client: use null snapc to check pool permission ([issue#13714](#), [pr#6497](#), Yan, Zheng)
- cls/cls_rbd.cc: fix misused metadata_name_from_key ([issue#13922](#), [pr#6661](#), Xiaoxi Chen)
- cls/cls_rbd: pass string by reference ([pr#7232](#), Jeffrey Lu)

- `cls_hello`: Fix grammatical error in description comment (pr#7951, Brad Hubbard)
- `cls_journal`: fix -EEXIST checking (pr#8413, runsisi)
- `cls_rbd`: add guards for error cases (issue#14316, issue#14317, pr#7165, xie xingguo)
- `cls_rbd`: change object_map_update to return 0 on success, add logging (pr#6467, Douglas Fuller)
- `cls_rbd`: enable object map checksums for object_map_save (issue#14280, pr#7149, Douglas Fuller)
- `cls_rbd`: fix -EEXIST checking in `cls::rbd::image_set` (pr#8371, runsisi)
- `cls_rbd`: fix the test for ceph-dencoder (pr#7793, Kefu Chai)
- `cls_rbd`: mirror_image_list should return global image id (pr#8297, Jason Dillaman)
- `cls_rbd`: mirroring directory (issue#14419, pr#7620, Josh Durgin)
- `cls_rbd`: pass WILLNEED fadvise flags during object map update (issue#15332, pr#8380, Jason Dillaman)
- `cls_rbd`: protect against excessively large object maps (issue#15121, pr#8099, Jason Dillaman)
- `cls_rbd`: read_peers: update last_read on next `cls_cxx_map_get_vals` (pr#8374, Mykola Golub)
- `cls/rgw`: fix FTBFS (pr#8142, Kefu Chai)
- `cls/rgw`: fix use of timespan (issue#15181, pr#8212, Yehuda Sadeh)
- `cmake`: add common/fs_types.cc to libcommon (pr#7898, Orit Wasserman)
- `cmake`: Add common/PluginRegistry.cc to CMakeLists.txt (pr#6805, Pete Zaitcev)
- `cmake`: Added new unittests to make check (pr#7572, Ali Maredia)
- `cmake`: Add ENABLE_GIT_VERSION to avoid rebuilding (pr#7171, Kefu Chai)
- `cmake`: add ErasureCode.cc to jerasure plugins (pr#7808, Casey Bodley)
- `cmake`: add FindOpenSSL.cmake (pr#8106, Marcus Watts, Matt Benjamin)
- `cmake`: add KernelDevice.cc to libos_srcs (pr#7507, Kefu Chai)
- `cmake`: add missing check for HAVE_EXECINFO_H (pr#7270, Casey Bodley)
- `cmake`: add missing librbd image_watcher sources (issue#14823, pr#7717, Casey Bodley)
- `cmake`: add missing librbd/MirrorWatcher.cc and librd/ObjectWatcher.cc (pr#8399, Orit Wasserman)
- `cmake`: add nss as a suffix for pk11pub.h (pr#6556, Samuel Just)
- `cmake`: add rgw_basic_types.cc to librgw.a (pr#6786, Orit Wasserman)
- `cmake`: add StandardPolicy.cc to librbd (pr#8368, Kefu Chai)
- `cmake`: add TracepointProvider.cc to libcommon (pr#6823, Orit Wasserman)
- `cmake`: avoid false-positive LDAP header detect (pr#8100, Matt Benjamin)
- `cmake`: Build cython modules and change paths to bin/, lib/ (pr#8351, John Spray, Ali Maredia)
- `cmake`: check for libsnappy in default path also (pr#7366, Kefu Chai)
- `cmake`: cleanups and more features from automake (pr#7103, Casey Bodley, Ali Maredia)
- `cmake`: define STRERROR_R_CHAR_P for GNU-specific strerror_r (pr#6751, Ilya Dryomov)
- `cmake`: detect bzip2 and lz4 (pr#7126, Kefu Chai)
- `cmake`: feb5 (pr#7541, Matt Benjamin)
- `cmake`: fix build with bluestore (pr#7099, John Spray)
- `cmake`: fix files list (pr#6539, Yehuda Sadeh)
- `cmake`: fix mrun to handle cmake build structure (pr#8237, Orit Wasserman)
- `cmake`: fix paths to various EC source files (pr#7748, Ali Maredia, Matt Benjamin)
- `cmake`: fix the build of test_rados_api_list (pr#8438, Kefu Chai)
- `cmake`: fix the build of tests (pr#7523, Kefu Chai)
- `cmake`: fix the build on trusty (pr#7249, Kefu Chai)
- `cmake`: For CMake version <= 2.8.11, use LINK_PRIVATE and LINK_PUBLIC (pr#7474, Tao Chang)
- `cmake`: For CMake version <= 2.8.11, use LINK_PRIVATE (pr#8422, Haomai Wang)
- `cmake`: let ceph-client-debug link with tcmalloc (pr#7314, Kefu Chai)
- `cmake`: librbd and libjournal build fixes (pr#6557, Ilya Dryomov)
- `cmake`: made rocksdb an imported library (pr#7131, Ali Maredia)
- `cmake`: make check (pr#7912, Kefu Chai, Ali Maredia)
- `cmake`: no need to run configure from run-cmake-check.sh (pr#6959, Orit Wasserman)
- `cmake` (pr#7849, Ali Maredia)
- `cmake/pybind`: fix include paths for cephfs (pr#8381, Josh Durgin)
- `cmake`: Remove duplicate find_package curl line. (pr#7972, Brad Hubbard)
- `cmake`: support ccache via a WITH_CCACHE build option (pr#6875, John Coyle)
- `cmake`: test_build_libcephfs needs \${ALLOC_LIBS} (pr#7300, Ali Maredia)
- `cmake`: update for recent librbd changes (pr#6715, John Spray)
- `cmake`: update for recent rbd changes (pr#6818, Mykola Golub)
- `cmake`: Use uname instead of arch. (pr#6358, John Coyle)
- `coc`: fix typo in the apt-get command (pr#6659, Chris Holcombe)
- `common`: add generic plugin infrastructure (pr#6696, Sage Weil)
- `common`: add latency perf counter for finisher (pr#6175, Xinze Chi)
- `common/address_help.cc`: fix the leak in entity_addr_from_url() (issue#14132, pr#6987, Qiankun Zheng)
- `common`: add thread names (pr#5882, Igor Podoski)
- `common`: add zlib compression plugin (pr#7437, Alyona Kiseleva, Kiseleva Alyona)
- `common`: admin socket commands for tcmalloc heap get/set operations (pr#7512, Samuel Just)
- `common`: ake ceph_time clocks work under BSD (pr#7340, Adam C. Emerson)

- common: allow enable/disable of optracker at runtime ([pr#5168](#), Jianpeng Ma)
- common: Allow OPT_INT settings with negative values ([issue#13829](#), [pr#7390](#), Brad Hubbard, Kefu Chai)
- common: assert: abort() rather than throw ([pr#6804](#), Adam C. Emerson)
- common: assert: __STRING macro is not defined by musl libc. ([pr#6210](#), John Coyle)
- common/bit_vector: use hard-coded value for block size ([issue#14747](#), [pr#7610](#), Jason Dillaman)
- common: buffer: add cached_crc and cached_crc_adjust counts to perf dump ([pr#6535](#), Ning Yao)
- common: buffer/assert minor fixes ([pr#6990](#), Matt Benjamin)
- common: buffer: put a guard for stat() syscall during read_file ([pr#7956](#), xie xingguo)
- common: buffer: remove unneeded list destructor ([pr#6456](#), Michal Jarzabek)
- common/buffer: replace RWLock with spinlocks ([pr#7294](#), Piotr Dałek)
- common/ceph_context.cc:fix order of initialisers ([pr#6838](#), Michal Jarzabek)
- common: change the type of counter total/unhealthy_workers ([pr#7254](#), Guang Yang)
- common: default cluster name to config file prefix ([pr#7364](#), Javen Wu)
- common: Deprecate or free up a bunch of feature bits ([pr#8214](#), Samuel Just)
- common: Do not use non-portable constants in mutex_debug ([pr#7766](#), Adam C. Emerson)
- common: don't reverse hobject_t hash bits when zero ([pr#6653](#), Piotr Dałek)
- common: fix OpTracker age histogram calculation ([pr#5065](#), Zhiqiang Wang)
- common: fix race during optracker switches between enabled/disabled mode ([pr#8330](#), xie xingguo)
- common: fix reset max in Throttle using perf reset command ([issue#13517](#), [pr#6300](#), Xinze Chi)
- common: fix time_t cast in decode ([issue#15330](#), [pr#8419](#), Adam C. Emerson)
- common/Formatter: avoid newline if there is no output ([pr#5351](#), Aran85)
- common: improve shared_cache and simple_cache efficiency with hash table ([pr#6909](#), Ning Yao)
- common/lockdep: increase max lock names ([pr#6961](#), Sage Weil)
- common: log: Assign LOG_DEBUG priority to syslog calls ([issue#13993](#), [pr#6815](#), Brad Hubbard)
- common: log: predict log message buffer allocation size ([pr#6641](#), Adam Kupczyk)
- common/MemoryModel: Added explicit feature check for mallinfo(). ([pr#6252](#), John Coyle)
- common: new timekeeping common code, and Objecter conversion ([pr#5782](#), Adam C. Emerson)
- common/obj_bencher.cc: bump the precision of bandwidth field ([pr#8021](#), Piotr Dałek)
- common/obj_bencher.cc: faster object name generation ([pr#7863](#), Piotr Dałek)
- common/obj_bencher.cc: fix verification crashing when there's no objects ([pr#5853](#), Piotr Dałek)
- common/obj_bencher.cc: make verify error fatal ([issue#14971](#), [pr#7897](#), Piotr Dałek)
- common: optimize debug logging code ([pr#6441](#), Adam Kupczyk)
- common: optimize debug logging ([pr#6307](#), Adam Kupczyk)
- common/page.cc: _page_mask has too many bits ([pr#7588](#), Dan Mick)
- common: perf counter for bufferlist history total alloc ([pr#6198](#), Xinze Chi)
- common: reduce CPU usage by making stringstream in stringify function thread local ([pr#6543](#), Evgeniy Firsov)
- common: re-enable backtrace support ([pr#6771](#), Jason Dillaman)
- common: set thread name from correct thread ([pr#7845](#), Igor Podoski)
- common: signal_handler: added support for using reentrant strsignal() implementations vs. sys_siglist[] ([pr#6796](#), John Coyle)
- common: snappy decompressor may assert when handling segmented input bufferlist ([issue#14400](#), [pr#7268](#), Igor Fedotov)
- common: string.h: return type from str_len(...) need not be const ([pr#7679](#), Matt Benjamin)
- common/str_map: cleanup: replaced get_str_map() function overloading by using default parameters for delimiters ([pr#7266](#), Sahithi R V)
- common/strtol.cc: fix the coverity warnings ([pr#7967](#), Kefu Chai)
- common: SubProcess: Avoid buffer corruption when calling err() ([issue#15011](#), [pr#8054](#), Erwan Velu)
- common: SubProcess: fix multiple definition bug ([pr#6790](#), Yunchuan Wen)
- common: Thread: move copy constructor and assignment op ([pr#5133](#), Michal Jarzabek)
- common: time: have skewing-now call non-skewing now ([pr#7466](#), Adam C. Emerson)
- common/TrackedOp: fix inaccurate counting for slow requests ([issue#14804](#), [pr#7690](#), xie xingguo)
- common: unit test for interval_set implementations ([pr#6](#), Igor Fedotov)
- common: use namespace instead of subclasses for buffer ([pr#6686](#), Michal Jarzabek)
- common: various fixes from SCA runs ([pr#7680](#), Danny Al-Gaaf)
- common: WeightedPriorityQueue Re-Add Round Robin for Classes ([pr#7984](#), Robert LeBlanc)
- common: WorkQueue: new PointerWQ base class for ContextWQ ([issue#13636](#), [pr#6525](#), Jason Dillaman)
- compat: use prefixed typeof extension ([pr#6216](#), John Coyle)
- config: add \$data_dir/config to config search path ([pr#7377](#), Sage Weil)
- config: complains when a setting is not tracked ([issue#11692](#), [pr#7085](#), Kefu Chai)
- config: fix osd_crush_initial_weight ([pr#7975](#), You Ji)
- config: increase default async op threads ([pr#7802](#), Piotr Dałek)
- config_opts: disable filestore throttle soft backoff by default ([pr#8265](#), Samuel Just)
- configure.ac: boost_iostreams is required, not optional ([pr#7816](#), Hector Martin)
- configure.ac: macro fix ([pr#6769](#), Igor Podoski)
- configure.ac: make “-with-librocksdb-static” default to ‘check’ ([issue#14463](#), [pr#7317](#), Dan Mick)

- configure.ac: update help strings for cython ([pr#7856](#), Josh Durgin)
- configure: Add -D_LARGEFILE64_SOURCE to Linux build. ([pr#8402](#), Ira Cooper)
- configure: detect bz2 and lz4 ([issue#13850](#), [issue#13981](#), [pr#7030](#), Kefu Chai)
- correct radosgw-admin command ([pr#7006](#), YankunLi)
- crush: add chooseleaf_stable tunable ([pr#6572](#), Sangdi Xu, Sage Weil)
- crush: add safety assert ([issue#14496](#), [pr#7344](#), songbaisen)
- crush: clean up whitespace removal ([issue#14302](#), [pr#7157](#), songbaisen)
- crush/CrushTester: check for overlapped rules ([pr#7139](#), Kefu Chai)
- crush/CrushTester: workaround a bug in boost::icl ([pr#7560](#), Kefu Chai)
- crush: fix cli tests for new crush tunables ([pr#8107](#), Sage Weil)
- crush: fix error log ([pr#8430](#), Wei Jin)
- crush: fix typo ([pr#8518](#), Wei Jin)
- crush: reply quickly from get_immediate_parent ([issue#14334](#), [pr#7181](#), song baisen)
- crushtool: Don't crash when called on a file that isn't a crushmap ([issue#8286](#), [pr#8038](#), Brad Hubbard)
- crushtool: improve usage/tip messages ([pr#7142](#), xie xingguo)
- crushtool: set type 0 name "device" for -build option ([pr#6824](#), Sangdi Xu)
- crush: update tunable docs. change default profile to jewel ([pr#7964](#), Sage Weil)
- crush: validate bucket id before indexing buckets array ([issue#13477](#), [pr#6246](#), Sage Weil)
- debian/changelog: Remove stray 'v' in version ([pr#7936](#), Dan Mick)
- debian/changelog: Remove stray 'v' in version ([pr#7938](#), Dan Mick)
- debian: include cpio in build-requires ([pr#7533](#), Rémi BUISSON)
- debian: package librgw_file* tests ([pr#7930](#), Ken Dreyer)
- debian: packaging fixes for jewel ([pr#7807](#), Ken Dreyer, Ali Maredia)
- debian/rpm split servers ([issue#10587](#), [pr#7746](#), Ken Dreyer)
- debian/rules: put init-ceph in /etc/init.d/ceph, not ceph-base ([issue#15329](#), [pr#8406](#), Dan Mick)
- debian: start ceph-mon-all and ceph-osd-all on package install ([issue#15098](#), [pr#8617](#), Sage Weil)
- doc: add ceph-detect-init(8) source to dist tarball ([pr#7933](#), Ken Dreyer)
- doc: add cinder backend section to rbd-openstack.rst ([pr#7923](#), RustShen)
- doc: adding "--allow-shrink" in decreasing the size of the rbd block to distinguish from the increasing option ([pr#7020](#), Yehua)
- doc: Adding documentation on how to use new dynamic throttle scheme ([pr#8069](#), Somnath Roy)
- doc: add orphans commands to radosgw-admin(8) ([issue#14637](#), [pr#7518](#), Ken Dreyer)
- doc: add v0.80.11 to the release timeline ([pr#6658](#), Loic Dachary)
- doc: admin/build-doc: add lxml dependencies on debian ([pr#6610](#), Ken Dreyer)
- doc: admin/build-doc: make paths absolute ([pr#7119](#), Dan Mick)
- doc: amend Fixes instructions in SubmittingPatches ([pr#8312](#), Nathan Cutler)
- doc: amend the rados.8 ([pr#7251](#), Kefu Chai)
- doc/architecture.rst: remove redundant word "across" ([pr#8179](#), Zhao Junwang)
- doc/cephfs posix: update ([pr#6922](#), Sage Weil)
- doc: Clarify usage on starting single osd/mds/mon. ([pr#7641](#), Patrick Donnelly)
- doc: CodingStyle: fix broken URLs ([pr#6733](#), Kefu Chai)
- doc: correct typo 'restared' to 'restarted' ([pr#6734](#), Yilong Zhao)
- doc: detailed description of bugfixing workflow ([pr#7941](#), Nathan Cutler)
- doc/dev: add "Deploy a cluster for manual testing" section ([issue#15218](#), [pr#8228](#), Nathan Cutler)
- doc/dev: add section on interrupting a running suite ([pr#8116](#), Nathan Cutler)
- doc/dev: continue writing Testing in the cloud chapter ([pr#7960](#), Nathan Cutler)
- doc: dev: document ceph-qa-suite ([pr#6955](#), Loic Dachary)
- doc/dev/index: refactor/reorg ([pr#6792](#), Nathan Cutler)
- doc/dev/index.rst: begin writing Contributing to Ceph ([pr#6727](#), Nathan Cutler)
- doc/dev/index.rst: fix headings ([pr#6780](#), Nathan Cutler)
- doc/dev: integrate testing into the narrative ([pr#7946](#), Nathan Cutler)
- doc: dev: introduction to tests ([pr#6910](#), Loic Dachary)
- doc/dev: various refinements ([pr#7954](#), Nathan Cutler)
- doc: document "readforward" and "readproxy" cache mode ([pr#7023](#), Kefu Chai)
- doc: download GPG key from download.ceph.com ([issue#13603](#), [pr#6384](#), Ken Dreyer)
- doc: draft notes for jewel ([pr#8211](#), Loic Dachary, Sage Weil)
- doc: file must be empty when writing layout fields of file use "setfattr" ([pr#6848](#), Cilang Zhao)
- doc: fix 0.94.4 and 0.94.5 ordering ([pr#7763](#), Loic Dachary)
- doc: fix dependencies ([pr#8587](#), Etienne Menguy)
- doc: Fixed incorrect name of a "List Multipart Upload Parts" Response Entity ([issue#14003](#), [pr#6829](#), Lenz Grimmer)
- doc: Fixes a CRUSH map step take argument ([pr#7327](#), Ivan Grcic)
- doc: Fixes a spelling error ([pr#6705](#), Jeremy Qian)
- doc: Fixes headline different font size and type ([pr#8328](#), scienceluo)
- doc: fixing image in section ERASURE CODING ([pr#7298](#), Rachana Patel)
- doc: fix misleading configuration guide on cache tiering ([pr#7000](#), Yuan Zhou)

- doc: fix “mon osd down out subtree limit” option name ([pr#7164](#), François Lafont)
- doc: fix outdated content in cache tier ([pr#6272](#), Yuan Zhou)
- doc: fix S3 C# example ([pr#7027](#), Dunrong Huang)
- doc: fix typo, duplicated content etc. for Jewel release notes ([pr#8342](#), xie xingguo)
- doc: fix typo in cephfs/quota ([pr#6745](#), Drunkard Zhang)
- doc: fix typo, indentation etc. ([pr#7829](#), xie xingguo)
- doc: fix typo in developer guide ([pr#6943](#), Nathan Cutler)
- doc: fix typo ([pr#7004](#), tianqing)
- doc: fix wrong type of hyphen ([pr#8252](#), xie xingguo)
- doc: initial draft of RBD mirroring admin documentation ([issue#15041](#), [pr#8169](#), Jason Dillaman)
- doc: INSTALL redirect to online documentation ([pr#6749](#), Loic Dachary)
- doc: little improvements for troubleshooting scrub issues ([pr#6827](#), Mykola Golub)
- doc: Modified a note section in rbd-snapshot doc. ([pr#6908](#), Nilamdyuti Goswami)
- doc: note that cephfs auth stuff is new in jewel ([pr#6858](#), John Spray)
- doc: osd-config Add Configuration Options for op queue. ([pr#7837](#), Robert LeBlanc)
- doc: osd: s/schedued/scheduled/ ([pr#6872](#), Loic Dachary)
- doc/rados/api/librados-intro.rst: fix typo ([pr#7879](#), xie xingguo)
- doc/rados/operations/crush: fix the formatting ([pr#8306](#), Kefu Chai)
- doc: recommend against ext4 ([pr#8556](#), Sage Weil)
- doc: reinstate accidentally removed section header ([pr#8569](#), Josh Durgin)
- doc: release-notes: draft v0.80.11 release notes ([pr#6374](#), Loic Dachary)
- doc: release-notes: draft v10.0.0 release notes ([pr#6666](#), Loic Dachary)
- doc/release-notes: fix indents ([pr#8345](#), Kefu Chai)
- doc/release-notes: jewel updates ([pr#8590](#), Nathan Cutler, Sage Weil)
- doc/release-notes: v9.1.0 ([pr#6281](#), Loic Dachary)
- doc/releases-notes: fix build error ([pr#6483](#), Kefu Chai)
- doc: Remove Ceph Monitors do lots of fsync() ([issue#15288](#), [pr#8327](#), Vikhyat Umrao)
- doc: remove redundant space in ceph-authtool/monmaptool doc ([pr#7244](#), Jiaying Ren)
- doc: remove toctree items under Create CephFS ([pr#6241](#), Jevon Qiao)
- doc: remove unnecessary period in headline ([pr#6775](#), Marc Koderer)
- doc: rename the “Create a Ceph User” section and add verbage about... ([issue#13502](#), [pr#6297](#), ritz303)
- doc: revise SubmittingPatches ([pr#7292](#), Kefu Chai)
- doc: rgw adding a very basic multisite doc ([issue#15247](#), [pr#8281](#), Abhishek Lekshmanan)
- doc: rgw admin uses “region list” not “regions list” ([pr#8517](#), Kris Jurka)
- doc: rgw explain keystone’s verify ssl switch ([pr#7862](#), Abhishek Lekshmanan)
- doc: rgw multisite, add pools section & minor cosmetic improvements ([pr#8653](#), Abhishek Lekshmanan)
- doc: rgw: port changes from downstream to upstream ([pr#7264](#), Bara Ancincova)
- doc: rgw_region_root_pool option should be in [global] ([issue#15244](#), [pr#8271](#), Vikhyat Umrao)
- doc: rst style fix for pools document ([pr#6816](#), Drunkard Zhang)
- doc: script and guidelines for mirroring Ceph ([pr#7384](#), Wido den Hollander)
- docs: Fix styling of newly added mirror docs ([pr#6127](#), Wido den Hollander)
- doc: small fixes ([pr#7813](#), xiexingguo)
- doc: standardize @param (not @parma, @parmam, @params) ([pr#7714](#), Nathan Cutler)
- doc: SubmittingPatches: there is no next; only jewel ([pr#6811](#), Nathan Cutler)
- doc, tests: update all <http://ceph.com/> to download.ceph.com ([pr#6435](#), Alfredo Deza)
- doc: Update ceph-disk manual page with new feature deactivate/destroy. ([pr#6637](#), Vicente Cheng)
- doc: Updated CloudStack RBD documentation ([pr#8308](#), Wido den Hollander)
- doc: update doc for with new pool settings ([pr#5951](#), Guang Yang)
- doc: Updated the rados command man page to include the -run-name opt... ([issue#12899](#), [pr#5900](#), ritz303)
- doc: update infernalis release notes ([pr#6575](#), vasukulkarni)
- doc: Update list of admin/build-doc dependencies ([issue#14070](#), [pr#6934](#), Nathan Cutler)
- doc: update radosgw-admin example ([pr#6256](#), YankunLi)
- doc: update the OS recommendations for newer Ceph releases ([pr#6355](#), ritz303)
- doc: use ‘ceph auth get-or-create’ for creating RGW keyring ([pr#6930](#), Wido den Hollander)
- doc: very basic doc on mstart ([pr#8207](#), Abhishek Lekshmanan)
- drop envz.h includes ([pr#6285](#), John Coyle)
- fix FTBFS introduced by d0af316 ([pr#7792](#), Kefu Chai)
- fix: use right init_flags to finish CephContext ([pr#6549](#), Yunchuan Wen)
- fs: be more careful about the “mds setmap” command to prevent breakage ([issue#14380](#), [pr#7262](#), Yan, Zheng)
- gobject_t: use # instead of ! as a separator ([pr#8055](#), Sage Weil)
- global: do not start two daemons with a single pid-file ([issue#13422](#), [pr#7075](#), shun song)
- global: do not start two daemons with a single pid-file (part 2) ([issue#13422](#), [pr#7463](#), Loic Dachary)
- global/global_init: expand metavariables in setuser_match_path ([issue#15365](#), [pr#8433](#), Sage Weil)
- global/signal_handler: print thread name in signal handler ([pr#8177](#), Jianpeng Ma)
- helgrind: additional race conditionslibrbd: journal replay should honor inter-event dependencies ([pr#7274](#), Jason

Dillaman)

- helgrind: fix real (and imaginary) race conditions ([issue#14163](#), [pr#7208](#), Jason Dillaman)
- include/encoding: do not try to be clever with list encoding ([pr#7913](#), Sage Weil)
- init-ceph: do umount when the path exists. ([pr#6866](#), Xiaoxi Chen)
- init-ceph.in: allow case-insensitive true in `osd crush update on start`` ([pr#7943](#), Eric Cook)
- init-ceph.in: skip ceph-disk if it is not present ([issue#10587](#), [pr#7286](#), Ken Dreyer)
- init-ceph: use getopt to make option processing more flexible ([issue#3015](#), [pr#6089](#), Nathan Cutler)
- journal: async methods to (un)register and update client ([pr#7832](#), Mykola Golub)
- journal: disconnect watch after watch error ([issue#14168](#), [pr#7113](#), Jason Dillaman)
- journal: fire replay complete event after reading last object ([issue#13924](#), [pr#6762](#), Jason Dillaman)
- journal: fix final result for `JournalTrimmer::C_RemoveSet` ([pr#8516](#), runsisi)
- journal: fix race condition between Future and journal shutdown ([issue#15364](#), [pr#8477](#), Jason Dillaman)
- journal: flush commit position on metadata shutdown ([pr#7385](#), Mykola Golub)
- journal: improve commit position tracking ([pr#7776](#), Jason Dillaman)
- journal: incremental improvements and fixes ([pr#6552](#), Mykola Golub)
- journal: possible race condition during live replay ([issue#15352](#), [pr#8494](#), Jason Dillaman)
- journal: prevent race injecting new records into overflowed object ([issue#15202](#), [pr#8220](#), Jason Dillaman)
- journal: reset `commit_position_task_ctx` pointer after task complete ([pr#7480](#), Mykola Golub)
- journal: re-use common threads between journalers ([pr#7906](#), Jason Dillaman)
- journal: support replaying beyond skipped splay objects ([pr#6687](#), Jason Dillaman)
- krbd: remove deprecated `-quiet` param from udevadm ([issue#13560](#), [pr#6394](#), Jason Dillaman)
- kv: fix bug in kv key optimization ([pr#6511](#), Sage Weil)
- kv: implement `value_as_ptr()` and use it in `.get()` ([pr#7052](#), Piotr Dałek)
- kv/KineticStore: fix broken `split_key` ([pr#6574](#), Haomai Wang)
- kv: optimize and clean up internal key/value interface ([pr#6312](#), Piotr Dałek, Sage Weil)
- libcephfs: fix python tests and fix `getcwd` on missing dir ([pr#7901](#), John Spray)
- libcephfs: Improve portability by replacing `loff_t` type usage with `off_t` ([pr#6301](#), John Coyle)
- libcephfs: only check file offset on glibc platforms ([pr#6288](#), John Coyle)
- libcephfs: update `LIBCEPHFS_VERSION` to indicate the interface was changed ([pr#7551](#), Jevon Qiao)
- librados: add c++ style osd/pg command interface ([pr#6893](#), Yunchuan Wen)
- librados: aix gcc librados port ([pr#6675](#), Rohan Mars)
- librados: avoid malloc(0) (which can return NULL on some platforms) ([issue#13944](#), [pr#6779](#), Dan Mick)
- librados: cancel aio notification linger op upon completion ([pr#8102](#), Jason Dillaman)
- librados: check connection state in `rados_monitor_log` ([issue#14499](#), [pr#7350](#), David Disseldorp)
- librados: clean up `Objecter.h` ([pr#6731](#), Jie Wang)
- librados: detect laggy ops with `objecter_timeout`, not `osd_timeout` ([pr#7629](#), Greg Farnum)
- librados: do cleanup ([pr#6488](#), xie xingguo)
- librados: do not clear handle for `aio_watch()` ([pr#7771](#), xie xingguo)
- librados: fix examples/librados/Makefile error. ([pr#6320](#), You Ji)
- librados: fix pool alignment API overflow issue ([issue#13715](#), [pr#6489](#), xie xingguo)
- librados: fix potential null pointer access when do `pool_snap_list` ([issue#13639](#), [pr#6422](#), xie xingguo)
- librados: fix PromoteOn2ndRead test for EC ([pr#6373](#), Sage Weil)
- librados: fix rare race where pool op callback may hang forever ([issue#13642](#), [pr#6426](#), xie xingguo)
- librados: fix several flaws introduced by the `enumeration_objects` API ([issue#14299](#), [issue#14301](#), [issue#14300](#), [pr#7156](#), xie xingguo)
- librados: fix test failure with new aio watch/unwatch API ([pr#7824](#), Jason Dillaman)
- librados: implement async watch/unwatch ([pr#7649](#), Haomai Wang)
- librados: include/rados/librados.h: fix typo ([pr#6741](#), Nathan Cutler)
- librados: init `crush_location` from config file. ([issue#13473](#), [pr#6243](#), Wei Luo)
- librados: mix lock cycle (un)registering asok commands ([pr#7581](#), John Spray)
- librados: move to c++11 concurrency types ([pr#5931](#), Adam C. Emerson)
- librados: new style (sharded) object listing ([pr#6405](#), John Spray, Sage Weil)
- librados: potential null pointer access in `list_(n)objects` ([issue#13822](#), [pr#6639](#), xie xingguo)
- librados: race condition on `aio_notify` completion handling ([pr#7864](#), Jason Dillaman)
- librados: remove duplicate definitions for `rados pool_stat_t` and `cluster_stat_t` ([pr#7330](#), Igor Fedotov)
- librados: Revert “rados: Add new field flags for `ceph_osd_op.copy_get.`” ([pr#8486](#), Sage Weil)
- librados: shutdown finisher in a more graceful way ([pr#7519](#), xie xingguo)
- librados: Solaris port ([pr#6416](#), Rohan Mars)
- librados: stat2 with higher time precision ([pr#7915](#), Yehuda Sadeh, Matt Benjamin)
- librados: Striper: Fix incorrect `push_front -> append_zero` change ([pr#7578](#), Haomai Wang)
- librados_test_stub: protect against notify/unwatch race ([pr#7540](#), Jason Dillaman)
- librados: wrongly passed in argument for stat command ([issue#13703](#), [pr#6476](#), xie xingguo)
- librbd: allocate new journal tag after acquiring exclusive lock ([pr#7884](#), Jason Dillaman)
- librbd: API: async open and close ([issue#14264](#), [pr#7259](#), Mykola Golub)
- librbd: automatically flush IO after blocking write operations ([issue#13913](#), [pr#6742](#), Jason Dillaman)

- librbd: Avoid create two threads per image ([pr#7400](#), Haomai Wang)
- librbd: avoid throwing error if mirroring is unsupported ([pr#8417](#), Jason Dillaman)
- librbd: better handling of exclusive lock transition period ([pr#7204](#), Jason Dillaman)
- librbd: block maintenance ops until after journal is ready ([issue#14510](#), [pr#7382](#), Jason Dillaman)
- librbd: block read requests until journal replayed ([pr#7627](#), Jason Dillaman)
- librbd: check for presence of journal before attempting to remove ([issue#13912](#), [pr#6737](#), Jason Dillaman)
- librbd: clear error when older OSD doesn't support image flags ([issue#14122](#), [pr#7035](#), Jason Dillaman)
- librbd: correct include guard in RenameRequest.h ([pr#7143](#), Jason Dillaman)
- librbd: correct issues discovered during teuthology testing ([issue#14108](#), [issue#14107](#), [pr#6974](#), Jason Dillaman)
- librbd: correct issues discovered via valgrind memcheck ([pr#8132](#), Jason Dillaman)
- librbd: correct issues discovered when cache is disabled ([issue#14123](#), [pr#6979](#), Jason Dillaman)
- librbd: correct race conditions discovered during unit testing ([issue#14060](#), [pr#6923](#), Jason Dillaman)
- librbd: deadlock while attempting to flush AIO requests ([issue#13726](#), [pr#6508](#), Jason Dillaman)
- librbd: differentiate journal replay flush vs shut down ([pr#7698](#), Jason Dillaman)
- librbd: disable copy-on-read when not exclusive lock owner ([issue#14167](#), [pr#7129](#), Jason Dillaman)
- librbd: disable image mirroring when image is removed ([issue#15265](#), [pr#8375](#), Ricardo Dias)
- librbd: disallow unsafe rbd_op_threads values ([issue#15034](#), [pr#8459](#), Josh Durgin)
- librbd: do not ignore self-managed snapshot release result ([issue#14170](#), [pr#7043](#), Jason Dillaman)
- librbd: enable/disable image mirroring automatically for pool mode ([issue#15143](#), [pr#8204](#), Ricardo Dias)
- librbd: ensure copy-on-read requests are complete prior to closing parent image ([pr#6740](#), Jason Dillaman)
- librbd: ensure librados callbacks are flushed prior to destroying ([issue#14092](#), [pr#7040](#), Jason Dillaman)
- librbd: exit if parent's snap is gone during clone ([issue#14118](#), [pr#6968](#), xie xingguo)
- librbd: fix enable objectmap feature issue ([issue#13558](#), [pr#6339](#), xinxin shu)
- librbd: fix handling of proxied maintenance operations during shut down ([issue#15471](#), [issue#15456](#), [pr#8565](#), Jason Dillaman)
- librbd: fix internal handling of dynamic feature updates ([pr#7299](#), Jason Dillaman)
- librbd: fix journal iohint ([pr#6917](#), Jianpeng Ma)
- librbd: fix known test case race condition failures ([issue#13969](#), [pr#6800](#), Jason Dillaman)
- librbd: fix merge-diff for >2GB diff-files ([issue#14030](#), [pr#6889](#), Yunchuan Wen)
- librbd: fix potential memory leak ([issue#14332](#), [issue#14333](#), [pr#7174](#), xie xingguo)
- librbd: fix snap_exists API return code overflow ([issue#14129](#), [pr#6986](#), xie xingguo)
- librbd: fix state machine race conditions during shut down ([pr#7761](#), Jason Dillaman)
- librbd: fix test case race condition for journaling ops ([pr#6877](#), Jason Dillaman)
- librbd: fix tracepoint parameter in diff_iterate ([pr#6892](#), Yunchuan Wen)
- librbd: flush and invalidate cache via admin socket ([issue#2468](#), [pr#6453](#), Mykola Golub)
- librbd: handle unregistering the image watcher when disconnected ([pr#8094](#), Jason Dillaman)
- librbd: image refresh code paths converted to async state machines ([pr#6859](#), Jason Dillaman)
- librbd: include missing header for bool type ([pr#6798](#), Mykola Golub)
- librbd: initial collection of state machine unit tests ([pr#6703](#), Jason Dillaman)
- librbd: integrate journaling for maintenance operations ([pr#6625](#), Jason Dillaman)
- librbd: integrate journaling support for IO operations ([pr#6541](#), Jason Dillaman)
- librbd: integrate journal replay with fsx testing ([pr#7583](#), Jason Dillaman)
- librbd: IO deadlock when dynamically enabling/disabling features ([issue#15102](#), [pr#8511](#), Jason Dillaman)
- librbd: journal framework for tracking exclusive lock transitions ([issue#13298](#), [pr#7529](#), Jason Dillaman)
- librbd: journaling-related lock dependency cleanup ([pr#6777](#), Jason Dillaman)
- librbd: journal replay needs to support re-executing maintenance ops ([issue#14822](#), [pr#7785](#), Jason Dillaman)
- librbd: journal replay should honor inter-event dependencies ([pr#7019](#), Jason Dillaman)
- librbd: journal shut down flush race condition ([issue#14434](#), [pr#7302](#), Jason Dillaman)
- librbd: not necessary to hold owner_lock while releasing snap id ([issue#13914](#), [pr#6736](#), Jason Dillaman)
- librbd: only send signal when AIO completions queue empty ([pr#6729](#), Jianpeng Ma)
- librbd: optionally validate new RBD pools for snapshot support ([issue#13633](#), [pr#6925](#), Jason Dillaman)
- librbd: partial revert of commit 9b0e359 ([issue#13969](#), [pr#6789](#), Jason Dillaman)
- librbd: perf counters might not be initialized on error ([issue#13740](#), [pr#6523](#), Jason Dillaman)
- librbd: perf section name: use hyphen to separate components ([issue#13719](#), [pr#6516](#), Mykola Golub)
- librbd: properly handle replay of snap remove RPC message ([issue#14164](#), [pr#7042](#), Jason Dillaman)
- librbd: reduce mem copies to user-buffer during read ([pr#7548](#), Jianpeng Ma)
- librbd: reduce verbosity of common error condition logging ([issue#14234](#), [pr#7114](#), Jason Dillaman)
- librbd: refresh image if required before replaying journal ops ([issue#14908](#), [pr#7978](#), Jason Dillaman)
- librbd: remove canceled tasks from timer thread ([issue#14476](#), [pr#7329](#), Douglas Fuller)
- librbd: remove duplicate read_only test in librbd::async_flatten ([pr#5856](#), runsisi)
- librbd: remove last synchronous librados calls from open/close state machine ([pr#7839](#), Jason Dillaman)
- librbd: replaying a journal op post-refresh requires locking ([pr#8028](#), Jason Dillaman)
- librbd: resize should only update image size within header ([issue#13674](#), [pr#6447](#), Jason Dillaman)
- librbd: retrieve image name when opening by id ([pr#7736](#), Mykola Golub)
- librbd: return error if we fail to delete object_map head object ([issue#14098](#), [pr#6958](#), xie xingguo)

- librbd: Revert “librbd: use task finisher per CephContext” ([issue#14780](#), [pr#7667](#), Josh Durgin)
- librbd: send notifications for mirroring status updates ([pr#8355](#), Jason Dillaman)
- librbd: several race conditions discovered under single CPU environment ([pr#7653](#), Jason Dillaman)
- librbd: simplify IO method signatures for 32bit environments ([pr#6700](#), Jason Dillaman)
- librbd: small fixes for error messages and readahead counter ([issue#14127](#), [pr#6983](#), xie xingguo)
- librbd: start perf counters after id is initialized ([issue#13720](#), [pr#6494](#), Mykola Golub)
- librbd: support eventfd for AIO completion notifications ([pr#5465](#), Haomai Wang)
- librbd: truncate does not need to mark the object as existing in the object map ([issue#14789](#), [pr#7772](#), xinxin shu)
- librbd: uninitialized state in snap remove state machine ([pr#6982](#), Jason Dillaman)
- librbd: update of mirror pool mode and mirror peer handling ([pr#7718](#), Jason Dillaman)
- librbd: use async librados notifications ([pr#7668](#), Jason Dillaman)
- log: do not repeat errors to stderr ([issue#14616](#), [pr#7983](#), Sage Weil)
- log: fix stack overflow when flushing large log lines ([issue#14707](#), [pr#7599](#), Igor Fedotov)
- log: segv in a portable way ([issue#14856](#), [pr#7790](#), Kefu Chai)
- log: use delete[] ([pr#7904](#), Sage Weil)
- mailmap: Abhishek Lekshmanan affiliation ([pr#8615](#), Abhishek Lekshmanan)
- mailmap: add UMCLOUD affiliation ([pr#6820](#), Jiaying Ren)
- mailmap for 10.0.4 ([pr#7932](#), Abhishek Lekshmanan)
- mailmap: hange organization for Dongmao Zhang ([pr#7173](#), Dongmao Zhang)
- mailmap: Igor Podoski affiliation ([pr#7219](#), Igor Podoski)
- mailmap: Jewel updates ([pr#6750](#), Abhishek Lekshmanan)
- mailmap: modify member info ([pr#6468](#), Xiaowei Chen)
- mailmap: revise organization ([pr#6519](#), Li Wang)
- mailmap: Ubuntu Kylin name changed to Kylin Cloud ([pr#6532](#), Loic Dachary)
- mailmap: update .organizationmap ([pr#6565](#), chenji-kael)
- mailmap update ([pr#7210](#), M Ranga Swami Reddy)
- mailmap update ([pr#8522](#), M Ranga Swami Reddy)
- mailmap update ([pr#8608](#), M Ranga Swami Reddy)
- mailmap: updates for infernalis. ([pr#6495](#), Yann Dupont)
- mailmap: updates ([pr#6258](#), M Ranga Swami Reddy)
- mailmap: updates ([pr#6594](#), chenji-kael)
- mailmap updates ([pr#6992](#), Loic Dachary)
- mailmap updates ([pr#7189](#), Loic Dachary)
- mailmap updates ([pr#7528](#), Yann Dupont)
- mailmap updates ([pr#8256](#), Loic Dachary)
- mailmap: Xie Xingguo affiliation ([pr#6409](#), Loic Dachary)
- Makefile-env.am: set a default for CEPH_BUILD_VIRTUALENV (part 2) ([pr#8320](#), Loic Dachary)
- makefile: fix rbdmap manpage ([pr#8310](#), Kefu Chai)
- makefile: remove libedit from libclient.la ([pr#7284](#), Kefu Chai)
- makefiles: remove bz2-dev from dependencies ([issue#13981](#), [pr#6939](#), Piotr Dałek)
- man/8/ceph-disk: fix formatting issue ([pr#8003](#), Sage Weil)
- man/8/ceph-disk: fix formatting issue ([pr#8012](#), Sage Weil)
- man: document listwatchers cmd in “rados” manpage ([pr#7021](#), Kefu Chai)
- mds: A few more snapshot fixes, mostly around snapshotted inode/dentry tracking ([pr#7798](#), Yan, Zheng)
- mds: Add cmapv to ESessions default constructor initializer list ([pr#8403](#), John Coyle)
- mds: add ‘p’ flag in auth caps to control setting pool in layout ([pr#6567](#), John Spray)
- mds: advance clientreplay when replying ([issue#14357](#), [pr#7216](#), John Spray)
- mds: allow client to request caps when opening file ([issue#14360](#), [pr#7952](#), Yan, Zheng)
- mds: Be more careful about directory fragmentation and scrubbing ([issue#15167](#), [pr#8180](#), Yan, Zheng)
- mds, client: add namespace to file_layout_t (previously ceph_file_layout) ([pr#7098](#), Yan, Zheng, Sage Weil)
- mds, client: fix locking around handle_conf_change ([issue#14365](#), [issue#14374](#), [pr#7312](#), John Spray)
- mds: don’t double-shutdown the timer when suiciding ([issue#14697](#), [pr#7616](#), Greg Farnum)
- mds: expose state of recovery to status ASOK command ([issue#14146](#), [pr#7068](#), Yan, Zheng)
- mds: filelock deadlock ([pr#7713](#), Yan, Zheng)
- mds: fix client capabilities during reconnect (client.XXXX isn’t responding to mclientcaps(revoke)) ([issue#11482](#), [pr#6432](#), Yan, Zheng)
- mds: fix client cap/message replay order on restart ([issue#14254](#), [issue#13546](#), [pr#7199](#), Yan, Zheng)
- mds: fix file_layout_t legacy encoding snafu ([pr#8455](#), Sage Weil)
- mds: fix fsmap decode ([pr#8063](#), Greg Farnum)
- mds: fix FSMap upgrade with daemons in the map ([pr#8073](#), John Spray, Greg Farnum)
- mds: fix inode_t::compare() ([issue#15038](#), [pr#8014](#), Yan, Zheng)
- mds: fix scrub_path ([pr#6684](#), John Spray)
- mds: fix setxattr (broken in a536d114) ([issue#14029](#), [pr#6941](#), John Spray)
- mds: fix standby replay thread creation ([issue#14144](#), [pr#7132](#), John Spray)
- mds: fix stray purging in ‘stripe_count > 1’ case ([issue#15050](#), [pr#8040](#), Yan, Zheng)

- mds: FSmap decode fix ([pr#8604](#), Greg Farnum)
- mds/FSMap: use _ in key name ([pr#8614](#), Sage Weil)
- mds: function parameter ‘df’ should be passed by reference ([pr#7490](#), Na Xie)
- mds: implement snapshot rename ([pr#5645](#), xinxin shu)
- mds: judgment added to avoid the risk of visiting the NULL pointer ([pr#7358](#), Kongming Wu)
- mds: messages/MOSDOp: cast in assert to eliminate warnings ([issue#13625](#), [pr#6414](#), David Zafman)
- mds: Minor fixes around data scan in some scenarios ([pr#8115](#), Yan, Zheng)
- mds: Multi-filesystem support ([issue#14952](#), [pr#6953](#), John Spray, Sage Weil)
- mds: new filtered MDS tell commands for sessions ([pr#6180](#), John Spray)
- mds: properly set STATE_STRAY/STATE_ORPHAN for stray dentry/inode ([issue#13777](#), [pr#6553](#), Yan, Zheng)
- mds: repair the command option “–hot-standby” ([pr#6454](#), Wei Feng)
- mds: ScrubStack and “tag path” command ([pr#5662](#), Yan, Zheng, John Spray, Greg Farnum)
- mds/Session: use projected parent for auth path check ([issue#13364](#), [pr#6200](#), Sage Weil)
- mds: tear down connections from tell commands ([issue#14048](#), [pr#6933](#), John Spray)
- mds: we should wait messenger when MDSDaemon suicide ([pr#6996](#), Wei Feng)
- messages/MOSDOp: clear reqid inc for v6 encoding ([issue#15230](#), [pr#8299](#), Sage Weil)
- mirrors: Change contact e-mail address for se.ceph.com ([pr#8007](#), Wido den Hollander)
- mirrors: Updated scripts and documentation for mirrors ([pr#7847](#), Wido den Hollander)
- misc: use make_shared while creating shared_ptr ([pr#7769](#), Somnath Roy)
- mon: add an independent option for max election time ([pr#7245](#), Sangdi Xu)
- mon: add osd blacklist clear ([pr#6945](#), John Spray)
- mon: add RAW USED column to ceph df detail ([pr#7087](#), Ruifeng Yang)
- mon: block ‘ceph osd pg-temp ...’ if pg_temp update is already pending ([pr#6704](#), Sage Weil)
- mon: cleanup set-quota error msg ([pr#7371](#), Abhishek Lekshmanan)
- monclient: avoid key renew storm on clock skew ([issue#12065](#), [pr#8258](#), Alexey Sheplyakov)
- mon: compact full epochs also ([issue#14537](#), [pr#7396](#), Kefu Chai)
- mon: consider pool size when creating pool ([issue#14509](#), [pr#7359](#), songbaisen)
- mon: consider the pool size when setting pool crush rule ([issue#14495](#), [pr#7341](#), song baisen)
- mon: degrade a log message to level 2 ([pr#6929](#), Kongming Wu)
- mon: do not send useless pg_create messages for split pgs ([pr#8247](#), Sage Weil)
- mon: don’t require OSD W for MRemoveSnaps ([issue#13777](#), [pr#6601](#), John Spray)
- mon: drop useless rank init assignment ([issue#14508](#), [pr#7321](#), huanwen ren)
- mon: enable ‘mon osd prime pg temp’ by default ([pr#7838](#), Robert LeBlanc)
- mon: fix calculation of %USED ([pr#7881](#), Adam Kupczyk)
- mon: fix ceph df pool available calculation for 0-weighted OSDs ([pr#6660](#), Chengyuan Li)
- mon: fix coding-style on PG related Monitor files ([pr#6881](#), Wido den Hollander)
- mon: fixes related to mondbstore->get() changes ([pr#6564](#), Piotr Dałek)
- mon: fix keyring permissions ([issue#14950](#), [pr#7880](#), Owen Synge)
- mon: fix locking in preinit error paths ([issue#14473](#), [pr#7353](#), huanwen ren)
- mon: fix monmap creation stamp ([pr#7459](#), duanweijun)
- mon: fix reuse of osd ids (clear osd info on osd deletion) ([issue#13988](#), [pr#6900](#), Loic Dachary, Sage Weil)
- mon: fix routed_request_tids leak ([pr#6102](#), Ning Yao)
- mon: fix sync of config-key data ([pr#7363](#), Xiaowei Chen)
- mon: fix the can’t change subscribe level bug in monitoring log ([pr#7031](#), Zhiqiang Wang)
- mon: go into ERR state if multiple PGs are stuck inactive ([issue#13923](#), [pr#7253](#), Wido den Hollander)
- mon: initialize last_* timestamps on new pgs to creation time ([issue#14952](#), [pr#7980](#), Sage Weil)
- mon: initialize recorded election epoch properly even when standalone ([issue#13627](#), [pr#6407](#), huanwen ren)
- mon: make clock skew checks sane ([issue#14175](#), [pr#7141](#), Joao Eduardo Luis)
- mon: mark_down_pgs in lockstep with pg_map’s osdmap epoch ([pr#8208](#), Sage Weil)
- mon/MDSMonitor: add confirmation to “ceph mds rmfailed” ([issue#14379](#), [pr#7248](#), Yan, Zheng)
- mon/MDSMonitor.cc: properly note beacon when health metrics changes ([issue#14684](#), [pr#7757](#), Yan, Zheng)
- mon: modify a dout level in OSDMonitor.cc ([pr#6928](#), Yongqiang He)
- mon/MonClient: avoid null pointer error when configured incorrectly ([issue#14405](#), [pr#7276](#), Bo Cai)
- mon/MonClient: fix shutdown race ([issue#13992](#), [pr#8335](#), Sage Weil)
- mon/monitor: some clean up ([pr#7520](#), huanwen ren)
- mon: MonmapMonitor: don’t expose uncommitted state to client ([pr#6854](#), Joao Eduardo Luis)
- mon/OSDMonitor: osdmap laggy set a maximum limit for interval ([pr#7109](#), Zengran Zhang)
- mon: osd [test-]reweight-by-{pg,utilization} command updates ([pr#7890](#), Dan van der Ster, Sage Weil)
- mon: paxos is_recovering calc error ([pr#7227](#), Weijun Duan)
- mon/PGMap: show rd/wr iops separately in status reports ([pr#7072](#), Cilang Zhao)
- mon: PGMonitor: acting primary diff with cur_stat, should not set pg to stale ([pr#7083](#), Xiaowei Chen)
- mon/PGMonitor: reliably mark PGs state ([pr#8089](#), Sage Weil)
- mon: PG Monitor should report waiting for backfill ([issue#12744](#), [pr#7398](#), Abhishek Lekshmanan)
- mon/pgmonitor: use appropriate forced conversions in get_rule_avail ([pr#7705](#), huanwen ren)
- mon: reduce CPU and memory manager pressure of pg health check ([pr#7482](#), Piotr Dałek)

- mon: remove ‘mds setmap’ ([issue#15136](#), [pr#8121](#), Sage Weil)
- mon: remove remove_legacy_versions() ([pr#8324](#), Kefu Chai)
- mon: remove unnecessary comment for update_from_paxos ([pr#8400](#), Qinghua Jin)
- mon: remove unused variable ([issue#15292](#), [pr#8337](#), Javier M. Mellid)
- mon: revert MonitorDBStore’s WholeStoreIteratorImpl::get ([issue#13742](#), [pr#6522](#), Piotr Dałek)
- mon: should not set isvalid = true when cephx_verify_authorizer return false ([issue#13525](#), [pr#6306](#), Rui Feng Yang)
- mon: show the pool quota info on ceph df detail command ([issue#14216](#), [pr#7094](#), song baisen)
- mon: some cleanup in MonmapMonitor.cc ([pr#7418](#), huanwen ren)
- mon: standardize Ceph removal commands ([pr#7939](#), Dongsheng Yang)
- mon: support min_down_reporter by subtree level (default by host) ([pr#6709](#), Xiaoxi Chen)
- mon: unconfuse object count skew message ([pr#7882](#), Piotr Dałek)
- mon: unregister command on shutdown ([pr#7504](#), huanwen ren)
- mon: warn if pg(s) not scrubbed ([issue#13142](#), [pr#6440](#), Michal Jarzabek)
- mount.ceph: memory leaks ([pr#6905](#), Qiankun Zheng)
- mount.fuse.ceph: better parsing of arguments passed to mount.fuse.ceph by mount command ([issue#14735](#), [pr#7607](#), Florent Bautista)
- mrun: update path to cmake binaries ([pr#8447](#), Casey Bodley)
- msg: add override to virutal methods ([pr#6977](#), Michal Jarzabek)
- msg: add thread safety for “random” Messenger + fix wrong usage of random functions ([pr#7650](#), Avner BenHanoch)
- msg/async: AsyncConnection: avoid debug log in cleanup_handler ([pr#7547](#), Haomai Wang)
- msg/async: AsyncConnection: avoid is_connected require connection’s lock ([issue#15440](#), [pr#8520](#), Haomai Wang)
- msg/async: AsyncMessenger: fix several bugs ([pr#7831](#), Haomai Wang)
- msg/async: AsyncMessenger: fix valgrind leak ([pr#7725](#), Haomai Wang)
- msg/async: avoid log spam on throttle ([issue#15031](#), [pr#8263](#), Kefu Chai)
- msg/async: bunch of fixes ([pr#7379](#), Piotr Dałek)
- msg/async: cleanup dead connection and misc things ([pr#7158](#), Haomai Wang)
- msg/async: don’t calculate msg header crc when not needed ([pr#7815](#), Piotr Dałek)
- msg/async: don’t use shared_ptr to manage EventCallback ([pr#7028](#), Haomai Wang)
- msg/async: Event: fix clock skew problem ([pr#7949](#), Wei Jin)
- msg/async: fix array boundary ([pr#7451](#), Wei Jin)
- msg: async: fix perf counter description and simplify _send_keepalive_or_ack ([pr#8046](#), xie xingguo)
- msg/async: fix potential race condition ([pr#7453](#), Haomai Wang)
- msg/async: fix send closed local_connection message problem ([pr#7255](#), Haomai Wang)
- msg/async: let receiver ack message ASAP ([pr#6478](#), Haomai Wang)
- msg/async: reduce extra tcp packet for message ack ([pr#7380](#), Haomai Wang)
- msg/async: remove experiment feature ([pr#7820](#), Haomai Wang)
- msg: async: small cleanups ([pr#7871](#), xie xingguo)
- msg/async: smarter MSG_MORE ([pr#7625](#), Piotr Dałek)
- msg: async: start over after failing to bind a port in specified range ([issue#14928](#), [issue#13002](#), [pr#7852](#), xie xingguo)
- msg/async: support of non-block connect in async messenger ([issue#12802](#), [pr#5848](#), Jianhui Yuan)
- msg/async: _try_send trim already sent for outcoming_bl more efficient ([pr#7970](#), Yan Jun)
- msg/async: will crash if enabling async msg because of an assertion ([pr#6640](#), Zhi Zhang)
- msg: filter out lo addr when bind osd addr ([pr#7012](#), Ji Chen)
- msg: removed unneeded includes from Dispatcher ([pr#6814](#), Michal Jarzabek)
- msg: remove duplicated code - local_delivery will now call ‘enqueue’ ([pr#7948](#), Avner BenHanoch)
- msg: remove unneeded inline ([pr#6989](#), Michal Jarzabek)
- msgr: fix large message data content length causing overflow ([pr#6809](#), Jun Huang, Haomai Wang)
- msg: significantly reduce minimal memory usage of connections ([pr#7567](#), Piotr Dałek)
- msg/simple: pipe: memory leak when signature check failed ([pr#7096](#), Rui Feng Yang)
- msg/simple: remove unneeded friend declarations ([pr#6924](#), Michal Jarzabek)
- msg/xio: fix compilation ([pr#7479](#), Roi Dayan)
- msg/xio: fixes ([pr#7603](#), Roi Dayan)
- msg/xio: simple fixes ([pr#8555](#), Avner BenHanoch)
- mstart: start rgw on different ports as well ([pr#8167](#), Abhishek Lekshmanan)
- nfs for rgw (Matt Benjamin, Orit Wasserman) ([pr#7634](#), Yehuda Sadeh, Matt Benjamin)
- objecter: avoid recursive lock of Objecter::rwlock ([pr#7343](#), Yan, Zheng)
- organizationmap: modify org mail info. ([pr#7240](#), Xiaowei Chen)
- os/bluestore: a few fixes ([pr#8193](#), Sage Weil)
- os/bluestore/BlueFS: Before reap ioc, it should wait io complete ([pr#8178](#), Jianpeng Ma)
- os/bluestore/BlueStore: Don’t leak trim overlay data before write. ([pr#7895](#), Jianpeng Ma)
- os/bluestore: ceph-bluefs-tool fixes ([issue#15261](#), [pr#8292](#), Venky Shankar)
- os/bluestore: clone overlay data ([pr#7860](#), Jianpeng Ma)
- os/bluestore: fix assert ([issue#14436](#), [pr#7293](#), xie xingguo)
- os/bluestore: fix a typo in SPDK path parsing ([pr#7601](#), Jianjian Huo)
- os/bluestore: fix bluestore_wal_transaction_t encoding test ([pr#7342](#), Kefu Chai)

- os/bluestore: fix bluestore_wal_transaction_t encoding test (pr#7419, Kefu Chai, Brad Hubbard)
- os/bluestore: fix shortened bluefs paths in env mirror case (pr#8498, Sage Weil)
- os/bluestore: insert new onode to the front position of onode LRU (pr#7492, Jianjian Huo)
- os/bluestore/KernelDevice: force block size (pr#8006, Sage Weil)
- os/bluestore: make bluestore_sync_transaction = true can work. (pr#7674, Jianpeng Ma)
- os/bluestore/NVMEDevice: make IO thread using dpdk launch (pr#8160, Haomai Wang)
- os/bluestore/NVMEDevice: refactor probe/attach codes and support zero command (pr#7647, Haomai Wang)
- os/bluestore: revamp BlueFS bdev management and add perfcounters (issue#15376, pr#8431, Sage Weil)
- os/bluestore: small fixes in bluestore StupidAllocator (pr#8101, Jianjian Huo)
- os/bluestore: use intrusive_ptr for Dir (pr#7247, Igor Fedotov)
- osd: add cache hint when pushing raw clone during recovery (pr#7069, Zhiqiang Wang)
- osd: Add config option osd_read_ec_check_for_errors for testing (pr#5865, David Zafman)
- osd: add missing newline to usage message (pr#7613, Willem Jan Withagen)
- osd: add osd op queue latency perfcounter (pr#5793, Haomai Wang)
- osd: add pin/unpin support to cache tier (11066) (pr#6326, Zhiqiang Wang)
- osd: add ‘proxy’ cache mode (issue#12814, pr#8210, Sage Weil)
- osd: add scrub persist/query API (issue#13505, pr#6898, Kefu Chai, Samuel Just)
- osd: add the support of per pool scrub priority (pr#7062, Zhiqiang Wang)
- osd: a fix for HeartbeatDispatcher and cleanups (pr#7550, Kefu Chai)
- osd: Allow repair of history.last_epoch_started using config (pr#6793, David Zafman)
- osd: auto repair EC pool (issue#12754, pr#6196, Guang Yang)
- osd: avoid calculating crush mapping for most ops (pr#6371, Sage Weil)
- osd: avoid debug std::string initialization in PG::get/put (pr#7117, Evgeniy Firsov)
- osd: avoid double-check for replaying and can_checkpoint() in FileStore::_check_replay_guard (pr#6471, Ning Yao)
- osd: avoid duplicate op->mark_started in ReplicatedBackend (pr#6689, Jacek J. Łakis)
- osd: avoid FORCE updating digest been overwritten by MAYBE when comparing scrub map (pr#7051, Zhiqiang Wang)
- osd: avoid osd_op_thread suicide because osd_scrub_sleep (pr#7009, Jianpeng Ma)
- osd: bail out of _committed_osd_maps if we are shutting down (pr#8267, Samuel Just)
- osd: blockdevice: avoid implicit cast and add guard (pr#7460, xie xingguo)
- osd: bluefs: fix alignment for odd page sizes (pr#7900, Dan Mick)
- osd: bluestore: add ‘override’ to virtual functions (pr#7886, Michal Jarzabek)
- osd: bluestore: allow _dump_onode dynamic accept log level (pr#7995, Jianpeng Ma)
- osd: bluestore/blockdevice: use std::mutex et al (pr#7568, Sage Weil)
- osd: bluestore: bluefs: fix several small bugs (issue#14344, issue#14343, pr#7200, xie xingguo)
- osd: bluestore/BlueFS: initialize super block_size earlier in mkfs (pr#7535, Sage Weil)
- osd: bluestore: don’t include when building without libaio (issue#14207, pr#7169, Mykola Golub)
- osd: bluestore: fix bluestore onode_t attr leak (pr#7125, Ning Yao)
- osd: bluestore: fix bluestore_wal_transaction_t encoding test (pr#7168, Kefu Chai)
- osd: bluestore: fix check for write falling within the same extent (issue#14954, pr#7892, Jianpeng Ma)
- osd: BlueStore: fix fsck and blockdevice read-relevant issue (pr#7362, xie xingguo)
- osd: BlueStore: fix null pointer access (issue#14561, pr#7435, xie xingguo)
- osd: bluestore: fix several bugs (issue#14259, issue#14353, issue#14260, issue#14261, pr#7122, xie xingguo)
- osd: bluestore: fix space rebalancing, collection split, buffered reads (pr#7196, Sage Weil)
- osd: bluestore: for overwrite a extent, allocate new extent on min_alloc_size write (pr#7996, Jianpeng Ma)
- osd: bluestore: improve fs-type verification and tidy up (pr#7651, xie xingguo)
- osd: bluestore, kstore: fix nid overwritten logic (issue#14407, issue#14433, pr#7283, xie xingguo)
- osd: bluestore: misc fixes (pr#7658, Jianpeng Ma)
- osd: bluestore: more fixes (pr#7130, Sage Weil)
- osd: BlueStore/NVMEDevice: fix compiling and fd leak (pr#7496, xie xingguo)
- osd: bluestore: NVMEDevice: fix error handling (pr#7799, xie xingguo)
- osd: bluestore: remove unneeded includes (pr#7870, Michal Jarzabek)
- osd: bluestore: Revert NVMEDevice task ctor and refresh interface changes (pr#7729, Haomai Wang)
- osd: bluestore updates, scrub fixes (pr#8035, Sage Weil)
- osd: bluestore: use btree_map for allocator (pr#7269, Igor Fedotov, Sage Weil)
- osd: cache tier: add config option for eviction check list size (pr#6997, Yuan Zhou)
- osd: call on_new_interval on newly split child PG (issue#13962, pr#6778, Sage Weil)
- osd: cancel failure reports if we fail to rebind network (pr#6278, Xinze Chi)
- osdc: Fix race condition with tick_event and shutdown (issue#14256, pr#7151, Adam C. Emerson)
- osd: change mutex to spinlock to optimize thread context switch. (pr#6492, Xiaowei Chen)
- osd: check do_shutdown before do_restart (pr#6547, Xiaoxi Chen)
- osd: check health state before pre_booting (issue#14181, pr#7053, Xiaoxi Chen)
- osd: clarify the scrub result report (pr#6534, Li Wang)
- osd/ClassHandler: only dlclose() the classes not missing (pr#8354, Kefu Chai)
- osd: clean up CMPXATTR checks (pr#5961, Jianpeng Ma)
- osd: clean up temp object if copy-from fails (pr#8487, Sage Weil)

- osd: clear pg_stat_queue after stopping pgs ([issue#14212](#), [pr#7091](#), Sage Weil)
- osd: clear requeue_scrub_on_unblock flag after requeue scrub ([pr#8570](#), xie xingguo)
- osdc/Objecter: dout log after assign tid ([pr#8202](#), Xinze Chi)
- osdc/Objecter: fix narrow race with tid assignment ([issue#14364](#), [pr#7981](#), Sage Weil)
- osdc/Objecter: use full pgid hash in PGNLS ops ([pr#8378](#), Sage Weil)
- osd: consider high/low mode when putting agent to sleep ([issue#14752](#), [pr#7631](#), Sage Weil)
- osd: correctly handle small osd_scrub_interval_randomize_ratio ([pr#7147](#), Samuel Just)
- osdc/Striper: reduce assemble_result log level ([pr#8426](#), Jason Dillaman)
- osd: defer decoding of MOSDRepOp/MOSDRepOpReply ([pr#6503](#), Xinze Chi)
- osd: delay populating in-memory PG log hashmaps ([pr#6425](#), Piotr Dałek)
- osd: disable filestore_xfs_extsize by default ([issue#14397](#), [pr#7265](#), Ken Dreyer)
- osd: do not keep ref of old osdmap in pg ([issue#13990](#), [pr#7007](#), Kefu Chai)
- osd: don't do random deep scrubs for user initiated scrubs ([pr#6673](#), David Zafman)
- osd: don't update epoch and rollback_info objects attrs if there is no need ([pr#6555](#), Ning Yao)
- osd: drop deprecated removal pg type ([pr#6970](#), Igor Podoski)
- osd: drop fiemap len=0 logic ([pr#7267](#), Sage Weil)
- osd: drop the interim set from load_pgs() ([pr#6277](#), Piotr Dałek)
- osd: dump number of missing objects for each peer with pg query ([pr#6058](#), Guang Yang)
- osd: duplicated clear for peer_missing ([pr#8315](#), Ning Yao)
- osd: enable perfcounters on sharded work queue mutexes ([pr#6455](#), Jacek J. Łakis)
- osd: ensure new osdmaps commit before publishing them to pgs ([issue#15073](#), [pr#8096](#), Sage Weil)
- osd: FileJournal: _fdump wrongly returns if journal is currently unreadable. ([issue#13626](#), [pr#6406](#), xie xingguo)
- osd: FileJournal: fix return code of create method ([issue#14134](#), [pr#6988](#), xie xingguo)
- osd: FileJournal: reduce locking scope in write_aio_bl ([issue#12789](#), [pr#5670](#), Zhi Zhang)
- osd: filejournal: report journal entry count ([pr#7643](#), tianqing)
- osd: FileJournal: support batch peak and pop from writeq ([pr#6701](#), Xinze Chi)
- osd: FileStore: add a field indicate xattr only one chunk for set xattr. ([pr#6244](#), Jianpeng Ma)
- osd: FileStore: Added O_DSYNC write scheme ([pr#7752](#), Somnath Roy)
- osd: FileStore: add error check for object_map->sync() ([pr#7281](#), Chendi Xue)
- osd: FileStore: cleanup: remove obsolete option "filestore_xattr_use_omap" ([issue#14356](#), [pr#7217](#), Vikhyat Umrao)
- osd: FileStore: conditional collection of drive metadata ([pr#6956](#), Somnath Roy)
- osd: filestore: FALLOC_FL_PUNCH_HOLE must be used with FALLOC_FL_KEEP_SIZE ([pr#7768](#), xinxin shu)
- osd: filestore: fast abort if statfs encounters ENOENT ([pr#7703](#), xie xingguo)
- osd: FileStore: fix initialization order for m_disable_wbthrottle ([pr#8067](#), Samuel Just)
- osd: filestore: fix race condition with split vs collection_move_rename and long object names ([issue#14766](#), [pr#8136](#), Samuel Just)
- osd: filestore: fix result code overwritten for clone ([issue#14817](#), [issue#14827](#), [pr#7711](#), xie xingguo)
- osd: filestore: fix wrong scope of result code for error cases during mkfs ([issue#14814](#), [pr#7704](#), xie xingguo)
- osd: filestore: fix wrong scope of result code for error cases during mount ([issue#14815](#), [pr#7707](#), xie xingguo)
- osd: FileStore: LFNIndex: remove redundant local variable 'obj'. ([issue#13552](#), [pr#6333](#), xie xingguo)
- osd: FileStore: modify the format of colon ([pr#7333](#), Donghai Xu)
- osd: FileStore:: optimize Ifn_unlink ([pr#6649](#), Jianpeng Ma)
- osd: FileStore: potential memory leak if _fgetattrs fails ([issue#13597](#), [pr#6377](#), xie xingguo)
- osd: FileStore: print file name before osd assert if read file failed ([pr#7111](#), Ji Chen)
- osd: FileStore: remove __SWORD_TYPE dependency ([pr#6263](#), John Coyle)
- osd: FileStore: remove unused local variable 'handle' ([pr#6381](#), xie xingguo)
- osd: filestore: restructure journal and op queue throttling ([pr#7767](#), Samuel Just)
- osd: FileStore: support multiple ondisk finish and apply finishers ([pr#6486](#), Xinze Chi, Haomai Wang)
- osd: FileStore: use pwritev instead of lseek+writev ([pr#7349](#), Haomai Wang, Tao Chang)
- osd: fix bogus scrub results when missing a clone ([issue#12738](#), [issue#12740](#), [pr#5783](#), David Zafman)
- osd: fix broken balance / localized read handling ([issue#13491](#), [pr#6364](#), Jason Dillaman)
- osd: fix bug in last_* PG state timestamps ([pr#6517](#), Li Wang)
- osd: fix bugs for omap ops ([pr#8230](#), Jianpeng Ma)
- osd: fix ClassHandler::ClassData::get_filter() ([pr#6747](#), Yan, Zheng)
- osd: fix/clean up full map request handling ([pr#8446](#), Sage Weil)
- osd: fix debug message in OSD::is_healthy ([pr#6226](#), Xiaoxi Chen)
- osd: fix dirtying info without correctly setting dirty_info field ([pr#8275](#), xie xingguo)
- osd: fix dump_ops_in_flight races ([issue#8885](#), [pr#8044](#), David Zafman)
- osd: fix epoch check in handle_pg_create ([pr#8382](#), Samuel Just)
- osd: fixes for several cases where op result code was not checked or set ([issue#13566](#), [pr#6347](#), xie xingguo)
- osd: fix failure report handling during ms_handle_connect() ([pr#8348](#), xie xingguo)
- osd: fix FileStore::_destroy_collection error return code ([pr#6612](#), Ruifeng Yang)
- osd: fix forced prmootion for CALL ops ([issue#14745](#), [pr#7617](#), Sage Weil)
- osd: fix fusestore hanging during stop/quit ([issue#14786](#), [pr#7677](#), xie xingguo)
- osd: fix inaccurate counter and skip over queueing an empty transaction ([pr#7754](#), xie xingguo)

- osd: fix incorrect throttle in WBThrottle ([pr#6713](#), Zhang Huan)
- osd: fix invalid list traversal in process_copy_chunk ([pr#7511](#), Samuel Just)
- osd: fix lack of object unblock when flush fails ([issue#14511](#), [pr#7584](#), Igor Fedotov)
- osd: fix log info ([pr#8273](#), Wei Jin)
- osd: fix MOSDOp encoding ([pr#6174](#), Sage Weil)
- osd: fix MOSDRepScrub reference counter in replica_scrub ([pr#6730](#), Jie Wang)
- osd: fix null pointer access and race condition ([issue#14072](#), [pr#6916](#), xie xingguo)
- osd: fix overload of '==' operator for pg_stat_t ([issue#14921](#), [pr#7842](#), xie xingguo)
- osd: fix race condition for heartbeat_need_update ([issue#14387](#), [pr#7739](#), xie xingguo)
- osd: fix reactivate (check OSDSuperblock in mkfs() when we already have the superblock) ([issue#13586](#), [pr#6385](#), Vicente Cheng)
- osd: fix reference count, rare race condition etc. ([pr#8254](#), xie xingguo)
- osd: fix return value from maybe_handle_cache_detail() ([pr#7593](#), Igor Fedotov)
- osd: fix rollback_info_trimmed_to before index() ([issue#13965](#), [pr#6801](#), Samuel Just)
- osd: fix scrub start hobject ([pr#7467](#), Sage Weil)
- osd: fix sparse-read result code checking logic ([issue#14151](#), [pr#7016](#), xie xingguo)
- osd: fix temp object removal after upgrade ([issue#13862](#), [pr#6976](#), David Zafman)
- osd: fix tick relevant issues ([pr#8369](#), xie xingguo)
- osd: fix trivial scrub bug ([pr#6533](#), Li Wang)
- osd: fix two scrub relevant issues ([pr#8462](#), xie xingguo)
- osd: fix unnecessary object promotion when deleting from cache pool ([issue#13894](#), [pr#7537](#), Igor Fedotov)
- osd: fix watch reconnect race ([issue#15441](#), [pr#8586](#), Sage Weil)
- osd: fix wip (l_osd_op_wip) perf counter and remove repop_map ([pr#7077](#), Xinze Chi)
- osd: fix wrongly placed assert and some cleanups ([pr#6766](#), xiexingguo, xie xingguo)
- osd: fix wrong return type of find_osd_on_ip() ([issue#14872](#), [pr#7812](#), xie xingguo)
- osd: fix wrong use of right parenthesis in localized read logic ([pr#6566](#), Jie Wang)
- osd: ghobject_t: use ! instead of @ as a separator ([pr#7595](#), Sage Weil)
- osd: handle dup pg_create that races with pg deletion ([pr#8033](#), Sage Weil)
- osd: improve temperature calculation for cache tier agent ([pr#4737](#), MingXin Liu)
- osd: initialize last recalibrate field at construction ([pr#8071](#), xie xingguo)
- osd: init started to 0 ([issue#13206](#), [pr#6107](#), Sage Weil)
- osd: KeyValueStore: don't queue NULL context ([pr#6783](#), Haomai Wang)
- osd: KeyValueStore: fix return code of mkfs ([pr#7036](#), xie xingguo)
- osd: KeyValueStore: fix the name's typo of keyvaluestore_default_strip_size ([pr#6375](#), Zhi Zhang)
- osd: KeyValueStore: fix wrongly placed assert ([issue#14176](#), [issue#14178](#), [pr#7047](#), xie xingguo)
- osd: kstore: fix a race condition in _txc_finish() ([pr#7804](#), Jianjian Huo)
- osd: kstore: latency breakdown ([pr#7850](#), James Liu)
- osd: kstore: several small fixes ([issue#14351](#), [issue#14352](#), [pr#7213](#), xie xingguo)
- osd: kstore: small fixes to kstore ([issue#14204](#), [pr#7095](#), xie xingguo)
- osd: kstore: sync up kstore with recent bluestore updates ([pr#7681](#), Jianjian Huo)
- osd: make backend and block device code a bit more generic ([pr#6759](#), Sage Weil)
- osd: make list_missing query missing_loc.needs_recovery_map ([pr#6298](#), Guang Yang)
- osdmap: remove unused local variables ([pr#6864](#), luo kexue)
- osdmap: rm nonused variable ([pr#8423](#), Wei Jin)
- osd: memstore: fix alignment of Page for test_pageset ([pr#7587](#), Casey Bodley)
- osd: memstore: fix two bugs ([pr#6963](#), Casey Bodley, Sage Weil)
- osd: merge local_t and op_t txn to single one ([pr#6439](#), Xinze Chi)
- osd: min_write_recency_for_promote & min_read_recency_for_promote are tiering only ([pr#8081](#), huanwen ren)
- osd: misc FileStore fixes ([issue#14192](#), [issue#14188](#), [issue#14194](#), [issue#14187](#), [issue#14186](#), [pr#7059](#), xie xingguo)
- osd: misc optimization for map utilization ([pr#6950](#), Ning Yao)
- osd, mon: fix exit issue ([pr#7420](#), Jiaying Ren)
- osd,mon: log leveldb and rocksdb to ceph log ([pr#6921](#), Sage Weil)
- osd: more fixes for incorrectly dirtying info; resend reply for duplicated scrub-reserve req ([pr#8291](#), xie xingguo)
- osd: move newest decode version of MOSDOp and MOSDOpReply to the front ([pr#6642](#), Jacek J. Łakis)
- osd: new and delete ObjectStore::Transaction in a function is not necessary ([pr#6299](#), Ruifeng Yang)
- osd: newstore: misc updates (including kv and os/fs stuff) ([pr#6609](#), Sage Weil)
- osd: note down the number of missing clones ([pr#6654](#), Kefu Chai)
- osd: Omap small bugs adapted ([pr#6669](#), Jianpeng Ma, David Zafman)
- osd: optimize clone write path if object-map is enabled ([pr#6403](#), xinxin shu)
- osd: optimize get_object_context ([pr#6305](#), Jianpeng Ma)
- osd: optimize MOSDOp/do_op/handle_op ([pr#5211](#), Jacek J. Lakis)
- osd: optimize scrub subset_last_update calculation ([pr#6518](#), Li Wang)
- osd: optimize the session_handle_reset function ([issue#14182](#), [pr#7054](#), songbaisen)
- osd: os/chain_xattr: On linux use linux/limits.h for XATTR_NAME_MAX. ([pr#6343](#), John Coyle)
- osd/OSD.cc: finish full_map_request every MOSDMap message. ([issue#15130](#), [pr#8147](#), Xiaoxi Chen)

- osd/OSD: fix build_past_intervals_parallel ([pr#8215](#), David Zafman)
- osd/OSDMap: fix typo in summarize_mapping_stats ([pr#8088](#), Sage Weil)
- osd: OSDMap: reset osd_primary_affinity shared_ptr when deepish_copy_from ([issue#14686](#), [pr#7553](#), Xinze Chi)
- osd: OSDService: Fix typo in osdmap comment ([pr#7275](#), Brad Hubbard)
- osd: os: skip checking pg_meta object existance in FileStore ([pr#6870](#), Ning Yao)
- osd: partial revert of “ReplicatedPG: result code not correctly set in some cases.” ([issue#13796](#), [pr#6622](#), Sage Weil)
- osd: PG::activate(): handle unexpected cached_removed_snaps more gracefully ([issue#14428](#), [pr#7309](#), Alexey Sheplyakov)
- osd/PG: indicate in pg query output whether ignore_history_les would help ([pr#8156](#), Sage Weil)
- osd: PGLog: clean up read_log ([pr#7092](#), Jie Wang)
- osd/PGLog: fix warning ([pr#8057](#), Sage Weil)
- osd/PG: pg down state blocked by osd.x, lost osd.x cannot solve peering stuck ([issue#13531](#), [pr#6317](#), Xiaowei Chen)
- osd: pg_pool_t: add dictionary for pool options ([issue#13077](#), [pr#6081](#), Mykola Golub)
- osd: PG: set epoch_created and parent_split_bits for child pg ([issue#15426](#), [pr#8552](#), Kefu Chai)
- osd: prevent osd_recovery_sleep from causing recovery-thread suicide ([pr#7065](#), Jianpeng Ma)
- osd: probabilistic cache tier promotion throttling ([pr#7465](#), Sage Weil)
- osd: randomize deep scrubbing ([pr#6550](#), Dan van der Ster, Herve Rousseau)
- osd: reduce memory consumption of some structs ([pr#6475](#), Piotr Dałek)
- osd: reduce string use in coll_t::calc_str() ([pr#6505](#), Igor Podoski)
- osd: refuse to start if object name limits are not possible ([pr#8560](#), Sage Weil)
- osd: release related sources when scrub is interrupted ([pr#6744](#), Jianpeng Ma)
- osd: release the message throttle when OpRequest unregistered ([issue#14248](#), [pr#7148](#), Samuel Just)
- osd: remove __SWORD_TYPE dependency ([pr#6262](#), John Coyle)
- osd: remove unused OSDMap::set_weightf() ([issue#14369](#), [pr#7231](#), huanwen ren)
- osd: remove up_thru_pending field, which is never used ([pr#7991](#), xie xingguo)
- osd: reorder bool fields in PGLog struct ([pr#6279](#), Piotr Dałek)
- osd: Replace sprintf with faster implementation in eversion_t::get_key_name ([pr#7121](#), Evgeniy Firsov)
- osd/ReplicatedPG: be more careful about calling publish_stats_to_osd() ([issue#14962](#), [pr#8039](#), Greg Farnum)
- osd: replicatedpg: break out loop if we encounter fatal error during do_pg_op() ([issue#14922](#), [pr#7844](#), xie xingguo)
- osd: ReplicatedPG: clean up unused function ([pr#7211](#), Xiaowei Chen)
- osd/ReplicatedPG: clear watches on change after applying repops ([issue#15151](#), [pr#8163](#), Sage Weil)
- osd/ReplicatedPG: fix implementation of register_on_success() ([pr#8612](#), xie xingguo)
- osd/ReplicatedPG: fix promotion recency logic ([issue#14320](#), [pr#6702](#), Sage Weil)
- osd/ReplicatedPG: make handle_watch_timeout no-op if !active ([issue#15391](#), [pr#8505](#), Sage Weil)
- osd: ReplicatedPG: remove unused local variables ([issue#13575](#), [pr#6360](#), xiexingguo)
- osd/ReplicatedPG::__rollback_to: update the OMAP flag ([issue#14777](#), [pr#8495](#), Samuel Just)
- osd: repop and lost-unfound overhaul ([pr#7765](#), Samuel Just)
- osd: reset primary and up_primary when building a new past_interval. ([issue#13471](#), [pr#6240](#), xiexingguo)
- osd: resolve boot vs NOUP set + clear race ([pr#7483](#), Sage Weil)
- osd: scrub: do not assign value if read error ([pr#6568](#), Li Wang)
- osd/ScrubStore: remove unused function ([pr#8045](#), Kefu Chai)
- osd: several small cleanups ([pr#7055](#), xie xingguo)
- osd: shut down if we flap too many times in a short period ([pr#6708](#), Xiaoxi Chen)
- osd: skip promote for writefull w/ FADVISE_DONTNEED/NOCACHE ([pr#7010](#), Jianpeng Ma)
- osd: slightly reduce actual size of pg_log_entry_t ([pr#6690](#), Piotr Dałek)
- osd: small fixes to memstore ([issue#14228](#), [issue#14229](#), [issue#14227](#), [pr#7107](#), xie xingguo)
- osd: some debug output for old osdmap trimming ([pr#8613](#), Sage Weil)
- osd: support pool level recovery_priority and recovery_op_priority ([pr#5953](#), Guang Yang)
- osd: try evicting after flushing is done ([pr#5630](#), Zhiqiang Wang)
- osd: use atomic to generate ceph_tid ([pr#7017](#), Evgeniy Firsov)
- osd: use optimized is_zero in object_stat_sum_t.is_zero() ([pr#7203](#), Piotr Dałek)
- osd: use pg id (without shard) when referring the PG ([pr#6236](#), Guang Yang)
- osd: utime_t, eversion_t, osd_stat_sum_t encoding optimization ([pr#6902](#), Xinze Chi)
- osd: WeightedPriorityQueue: move to intrusive containers ([pr#7654](#), Robert LeBlanc)
- osd: write file journal optimization ([pr#6484](#), Xinze Chi)
- os/filestore/FileJournal: set block size via config option ([pr#7628](#), Sage Weil)
- os/filestore: fix punch hole usage in _zero ([pr#8050](#), Sage Weil)
- os/filestore: fix result handling logic of destroy_collection ([pr#7721](#), xie xingguo)
- os/filestore: fix return type mismatch for Ifn parse ([pr#8545](#), xie xingguo)
- os/filestore: force Ifn attrs to be written atomically, restructure name length limits ([pr#8496](#), Samuel Just)
- os/filestore: require offset == length == 0 for full object read; add test ([pr#7957](#), Jianpeng Ma)
- os/fs: fix io_getevents argument ([pr#7355](#), Jingkai Yuan)
- os/fusestore: add error handling ([pr#7395](#), xie xingguo)
- os/keyvaluestore: kill KeyValueStore ([pr#7320](#), Haomai Wang)
- os/kstore: insert new onode to the front position of onode LRU ([pr#7505](#), xie xingguo)

- os/ObjectStore: add custom move operations for ObjectStore::Transaction (pr#7303, Casey Bodley)
- os/ObjectStore: add noexcept to ensure move ctor is used (pr#8421, Kefu Chai)
- os/ObjectStore: fix _update_op for split dest_cid (pr#8364, Sage Weil)
- os/ObjectStore: implement more efficient get_encoded_bytes() (pr#7775, Piotr Dałek)
- os/ObjectStore: make device uuid probe output something friendly (pr#8418, Sage Weil)
- os/ObjectStore: try_move_rename in transaction append and add coverage to store_test (issue#15205, pr#8359, Samuel Just)
- packaging: add build dependency on python devel package (pr#7205, Josh Durgin)
- packaging: make infernalis -> jewel upgrade work (issue#15047, pr#8034, Nathan Cutler)
- packaging: move cephfs repair tools to ceph-common (issue#15145, pr#8133, Boris Ranto, Ken Dreyer)
- pybind: add ceph_volume_client interface for Manila and similar frameworks (pr#6205, John Spray)
- pybind: add flock to libcephfs python bindings (pr#7902, John Spray)
- pybind/cephfs: add symlink and its unit test (pr#6323, Shang Ding)
- pybind: decode empty string in conf_parse_argv() correctly (pr#6711, Josh Durgin)
- pybind: Ensure correct python flags are passed (pr#7663, James Page)
- pybind: fix build failure, remove extraneous semicolon in method (issue#14371, pr#7235, Abhishek Lekshmanan)
- pybind: flag an RBD image as closed regardless of result code (pr#8005, Jason Dillaman)
- pybind: Implementation of rados_ioctx_snapshot_rollback (pr#6878, Florent Manens)
- pybind/Makefile.am: Prevent race creating CYTHON_BUILD_DIR (issue#15276, pr#8356, Dan Mick)
- pybind: move cephfs to Cython (pr#7745, John Spray, Mehdi Abaakouk)
- pybind: port the rbd bindings to Cython (issue#13115, pr#6768, Hector Martin)
- pybind/rados: fix object lifetime issues and other bugs in aio (pr#7778, Hector Martin)
- pybind/rados: fix omap method return values (pr#8603, Roland Mechler)
- pybind/rados: python3 fix (pr#8331, Mehdi Abaakouk)
- pybind/rados: use __dealloc__ since __del__ is ignored by cython (pr#7692, Mehdi Abaakouk)
- pybind: remove next() on iterators (pr#7706, Mehdi Abaakouk)
- pybind: replace __del__ with __dealloc__ for rbd (pr#7708, Josh Durgin)
- pybind: support ioctx:exec (pr#6795, Noah Watkins)
- pybind/test_rbd: fix test_create_defaults (issue#14279, pr#7155, Josh Durgin)
- pybind: use correct subdir for rados install-exec rule (pr#7684, Josh Durgin)
- python binding of librados with cython (pr#7621, Mehdi Abaakouk)
- python: use pip instead of python setup.py (pr#7605, Loic Dachary)
- qa: add workunit to run ceph_test_rbd_mirror (pr#8221, Josh Durgin)
- qa: disable rbd/qemu-iotests test case 055 on RHEL/CentOSlibrbd: journal replay should honor inter-event dependencies (issue#14385, pr#7272, Jason Dillaman)
- qa: erasure-code benchmark plugin selection (pr#6685, Loic Dachary)
- qa/krbd: Expunge generic/247 (pr#6831, Douglas Fuller)
- qa: update rest test cephfs calls (issue#15309, pr#8372, John Spray)
- qa: update rest test cephfs calls (part 2) (issue#15309, pr#8393, John Spray)
- qa/workunits/cephtool/test.sh: false positive fail on /tmp/obj1. (pr#6837, Robin H. Johnson)
- qa/workunits/cephtool/test.sh: no ./ (pr#6748, Sage Weil)
- qa/workunits/cephtool/test.sh: wait longer in ceph_watch_start() (issue#14910, pr#7861, Kefu Chai)
- qa/workunits: merge_diff shouldn't attempt to use striping (issue#14165, pr#7041, Jason Dillaman)
- qa/workunits/rados: Test exit values on test.sh, fix tier.cc (issue#15165, pr#8266, Samuel Just)
- qa/workunits/rados/test.sh: capture stderr too (pr#8004, Sage Weil)
- qa/workunits/rados/test.sh: test tmap_migrate (pr#8114, Sage Weil)
- qa/workunits/rbd: do not use object map during read flag testing (pr#8104, Jason Dillaman)
- qa/workunits/rbd: new online maintenance op tests (pr#8216, Jason Dillaman)
- qa/workunits/rbd: rbd-nbd test should use sudo for map/unmap ops (issue#14221, pr#7101, Jason Dillaman)
- qa/workunits/rbd: use POSIX function definition (issue#15104, pr#8068, Nathan Cutler)
- qa/workunits/rest/test.py: add confirmation to 'mds setmap' (issue#14606, pr#7982, Sage Weil)
- qa/workunits/rest/test.py: don't use newfs (pr#8191, Sage Weil)
- qa/workunits/snaps: move snap tests into fs sub-directory (pr#6496, Yan, Zheng)
- rados: add ceph:: namespace to bufferlist type (pr#8059, Noah Watkins)
- rados: bench: fix off-by-one to avoid writing past object_size (pr#6677, Tao Chang)
- rados: fix bug for write bench (pr#7851, James Liu)
- radosgw-admin: allow 'period pull -url' without full multisite config (pr#8567, Casey Bodley)
- radosgw-admin: allow (pr#8529, Orit Wasserman)
- radosgw-admin: Checking the legality of the parameters (issue#13018, pr#5879, Qiankun Zheng)
- radosgw-admin: Create -secret-key alias for -secret (issue#5821, pr#5335, Yuan Zhou)
- radosgw-admin: fix for 'realm pull' (pr#8404, Casey Bodley)
- radosgw-admin: metadata list user should return an empty list when user pool is empty (issue#13596, pr#6465, Orit Wasserman)
- radosgw-admin: 'period commit' supplies user-readable error messages (pr#8264, Casey Bodley)
- rados: implement rm -force option to force remove when full (pr#6202, Xiaowei Chen)

- rados: new options for write benchmark ([pr#6340](#), Joaquim Rocha)
- rbd: accept -user, refuse -i command-line optionals ([pr#6590](#), Ilya Dryomov)
- rbd: additional validation for striping parameters ([pr#6914](#), Na Xie)
- rbd: add missing command aliases to refactored CLI ([issue#13806](#), [pr#6606](#), Jason Dillaman)
- rbd: add -object-size option, deprecate -order ([issue#12112](#), [pr#6830](#), Vikhyat Umrao)
- rbd: add pool name to disambiguate rbd admin socket commands ([pr#6904](#), wuxiangwei)
- rbd: add RBD pool mirroring configuration API + CLI ([pr#6129](#), Jason Dillaman)
- rbd: add support for mirror image promotion/demotion/resync ([pr#8138](#), Jason Dillaman)
- rbd: allow librados to prune the command-line for config overrides ([issue#15250](#), [pr#8282](#), Jason Dillaman)
- rbd: clone operation should default to image format 2 ([pr#8119](#), Jason Dillaman)
- rbd: correct an output string for merge-diff ([pr#7046](#), Kongming Wu)
- rbd: deprecate image format 1 ([pr#7841](#), Jason Dillaman)
- rbd: dynamically generated bash completion ([issue#13494](#), [pr#6316](#), Jason Dillaman)
- rbd: fix build with “-without-rbd” ([issue#14058](#), [pr#6899](#), Piotr Dałek)
- rbd: fix clone isssue ([issue#13553](#), [pr#6334](#), xinxin shu)
- rbd: fixes for refactored CLI and related tests ([pr#6738](#), Ilya Dryomov)
- rbd: fix init-rbdmap CMDPARAMS ([issue#13214](#), [pr#6109](#), Sage Weil)
- rbd: fix static initialization ordering issues ([pr#6978](#), Mykola Golub)
- rbd-fuse: image name can not include snap name ([pr#7044](#), Yongqiang He)
- rbd-fuse: implement mv operation ([pr#6938](#), wuxiangwei)
- rbd: journal: configuration via conf, cli, api and some fixes ([pr#6665](#), Mykola Golub)
- rbd: journal reset should disable/re-enable journaling feature ([issue#15097](#), [pr#8490](#), Jason Dillaman)
- rbd: loosen image spec validation requirements ([pr#8606](#), Jason Dillaman)
- rbd: make config changes actually apply ([pr#6520](#), Mykola Golub)
- rbdmap: add manpage ([issue#15212](#), [pr#8224](#), Nathan Cutler)
- rbdmap: systemd support ([issue#13374](#), [pr#6479](#), Boris Ranto)
- rbd: merge_diff test should use new -object-size parameter instead of -order ([issue#14106](#), [pr#6972](#), Na Xie, Jason Dillaman)
- rbd-mirror: asok commands to get status and flush on Mirror and Replayer level ([pr#8235](#), Mykola Golub)
- rbd-mirror: enabling/disabling pool mirroring should update the mirroring directory ([issue#15217](#), [pr#8261](#), Ricardo Dias)
- rbd-mirror: fix image replay test failures ([pr#8158](#), Jason Dillaman)
- rbd-mirror: fix long termination due to 30sec wait in main loop ([pr#8185](#), Mykola Golub)
- rbd-mirror: fix missing increment for iterators ([pr#8352](#), runsisi)
- rbd-mirror: ImageReplayer async start/stop ([pr#7944](#), Mykola Golub)
- rbd-mirror: ImageReplayer improvements ([pr#7759](#), Mykola Golub)
- rbd-mirror: implement ImageReplayer ([pr#7614](#), Mykola Golub)
- rbd-mirror: initial failover / fallback support ([pr#8287](#), Jason Dillaman)
- rbd-mirror: integrate with image sync state machine ([pr#8079](#), Jason Dillaman)
- rbd-mirror: make remote context respect env and argv config params ([pr#8182](#), Mykola Golub)
- rbd-mirror: minor fix-ups for initial skeleton implementation ([pr#7958](#), Mykola Golub)
- rbd-mirror: prevent enabling/disabling an image’s mirroring when not in image mode ([issue#15267](#), [pr#8332](#), Ricardo Dias)
- rbd-mirror: remote to local cluster image sync ([pr#7979](#), Jason Dillaman)
- rbd-mirror: switch fsid over to mirror uuid ([issue#15238](#), [pr#8280](#), Ricardo Dias)
- rbd-mirror: use pool/image names in asok commands ([pr#8159](#), Mykola Golub)
- rbd-mirror: use the mirroring directory to detect candidate images ([issue#15142](#), [pr#8162](#), Ricardo Dias)
- rbd-mirror: workaround for intermingled lockdep singletons ([pr#8476](#), Jason Dillaman)
- rbd: must specify both of stripe-unit and stripe-count when specifying stripingv2 feature ([pr#7026](#), Donghai Xu)
- rbd-nbd: add copyright ([pr#7166](#), Li Wang)
- rbd-nbd: fix up return code handling ([pr#7215](#), Mykola Golub)
- rbd-nbd: network block device (NBD) support for RBD ([pr#6657](#), Yunchuan Wen, Li Wang)
- rbd-nbd: small improvements in logging and forking ([pr#7127](#), Mykola Golub)
- rbd: output formatter may not be closed upon error ([issue#13711](#), [pr#6706](#), xie xingguo)
- rbd: rbdmap improvements ([pr#6445](#), Boris Ranto)
- rbd: rbd order will be place in 22, when set to 0 in the config_opt ([issue#14139](#), [issue#14047](#), [pr#6886](#), huanwen ren)
- rbd: refactor cli command handling ([pr#5987](#), Jason Dillaman)
- rbd/run_cli_tests.sh: Reflect test failures ([issue#14825](#), [pr#7781](#), Zack Cerza)
- rbd: stripe unit/count set incorrectly from config ([pr#6593](#), Mykola Golub)
- rbd: striping parameters should support 64bit integers ([pr#6942](#), Na Xie)
- rbd: support for enabling/disabling mirroring on specific images ([issue#13296](#), [pr#8056](#), Ricardo Dias)
- rbd: support negative boolean command-line optionals ([issue#13784](#), [pr#6607](#), Jason Dillaman)
- rbd: unbreak rbd map + cephx_sign_messages option ([pr#6583](#), Ilya Dryomov)
- rbd: update default image features ([pr#7846](#), Jason Dillaman)
- rbd: use default order from configuration when not specified ([pr#6965](#), Yunchuan Wen)
- release-notes: draft v0.94.4 release notes ([pr#5907](#), Loic Dachary)

- release-notes: draft v0.94.4 release notes ([pr#6195](#), Loic Dachary)
- release-notes: draft v0.94.4 release notes ([pr#6238](#), Loic Dachary)
- release-notes: draft v0.94.6 release notes ([issue#13356](#), [pr#7689](#), Abhishek Varshney, Loic Dachary)
- release-notes: draft v10.0.3 release notes ([pr#7592](#), Loic Dachary)
- release-notes: draft v10.0.4 release notes ([pr#7966](#), Loic Dachary)
- release-notes: draft v9.2.1 release notes ([issue#13750](#), [pr#7694](#), Abhishek Varshney)
- release-notes: v10.1.2 release notes ([pr#8594](#), Loic Dachary)
- releases: what is merged where and when ? ([pr#8358](#), Loic Dachary)
- rgw: accept data only at the first time in response to a request ([pr#8084](#), sunspot)
- rgw: add a few more help options in admin interface ([pr#8410](#), Abhishek Lekshmanan)
- rgw: add a method to purge all associate keys when removing a subuser ([issue#12890](#), [pr#6002](#), Sangdi Xu)
- rgw: add a missing cap type ([pr#6774](#), Yehuda Sadeh)
- rgw: add an inspection to the field of type when assigning user caps ([pr#6051](#), Kongming Wu)
- rgw: Add a test for multi-tenancy ([pr#8592](#), Pete Zaitcev)
- rgw: add bucket request payment feature usage statistics integration ([issue#13834](#), [pr#6656](#), Javier M. Mellid)
- rgw: add compat header for TEMP_FAILURE_RETRY ([pr#6294](#), John Coyle)
- rgw: add default quota config ([pr#6400](#), Daniel Gryniewicz)
- rgw: add LifeCycle feature ([pr#6331](#), Ji Chen)
- rgw: add missing error code for admin op API ([pr#7037](#), Dunrong Huang)
- rgw: adds the radosgw-admin sync status command that gives a human readable status of the sync process at a specific zone ([pr#8030](#), Yehuda Sadeh)
- rgw: add support for caching of Keystone admin token. ([pr#7630](#), Radoslaw Zarzynski)
- rgw: add support for “end_marker” parameter for GET on Swift account. ([issue#10682](#), [pr#4216](#), Radoslaw Zarzynski)
- rgw: add support for getting Swift’s DLO without manifest handling ([pr#6206](#), Radoslaw Zarzynski)
- rgw: add support for metadata upload during PUT on Swift container. ([pr#8002](#), Radoslaw Zarzynski)
- rgw: add support for Static Large Objects of Swift API ([issue#12886](#), [issue#13452](#), [pr#6643](#), Yehuda Sadeh, Radoslaw Zarzynski)
- rgw: add support for system requests over Swift API ([pr#7666](#), Radoslaw Zarzynski)
- rgw: add zone delete to rgw-admin help ([pr#8184](#), Abhishek Lekshmanan)
- rgw: adjust error code when bucket does not exist in copy operation ([issue#14975](#), [pr#7916](#), Yehuda Sadeh)
- rgw: adjust the request_uri to support absoluteURI of http request ([issue#12917](#), [pr#7675](#), Wenjun Huang)
- rgw: admin api for retrieving usage info (Ji Chen) ([pr#8031](#), Yehuda Sadeh, Ji Chen)
- rgw_admin: improve period update errors ([issue#15251](#), [pr#8564](#), Abhishek Lekshmanan)
- rgw_admin: orphans finish segfaults ([pr#6652](#), Igor Fedotov)
- rgw-admin: remove unused iterator and fix error message ([pr#8507](#), Karol Mroz)
- rgw_admin: remove unused parent_period arg ([pr#8411](#), Abhishek Lekshmanan)
- rgw: Allow an implicit tenant in case of Keystone ([pr#8139](#), Pete Zaitcev)
- rgw: allow authentication keystone with self signed certs ([issue#14853](#), [issue#13422](#), [pr#7777](#), Abhishek Lekshmanan)
- rgw: approximate AmazonS3 HostId error field. ([pr#7444](#), Robin H. Johnson)
- rgw: aws4: handle UNSIGNED-PAYLOAD under header auth ([issue#15499](#), [pr#8601](#), Javier M. Mellid)
- rgw: aws4 subdomain calling bugfix ([issue#15369](#), [pr#8472](#), Javier M. Mellid)
- rgw: bucket link now set the bucket.instance acl (bug fix) ([issue#11076](#), [pr#8037](#), Zengran Zhang)
- rgw: bucket request payment support ([issue#13427](#), [pr#6214](#), Javier M. Mellid)
- rgw: Bug fix for mtime anomalies in RadosGW and other places ([pr#7328](#), Adam C. Emerson, Casey Bodley)
- rgw: build-related fixes ([pr#8076](#), Yehuda Sadeh, Matt Benjamin)
- rgw: calculate payload hash in RGWPutObj_ObjStore only when necessary. ([pr#7869](#), Radoslaw Zarzynski)
- rgw: call rgw_log_usage_finalize() on reconfiguration ([pr#8585](#), Casey Bodley)
- rgw: Check return code in RGWFileHandle::write ([pr#7875](#), Brad Hubbard)
- rgw: check the return value when call fe->run() ([issue#14585](#), [pr#7457](#), wei qiaomiao)
- rgw: clarify the error message when trying to create an existed user ([pr#5938](#), Zeqiang Zhuang)
- rgw: cleanups to comments and messages ([pr#7633](#), Pete Zaitcev)
- rgw: content length ([issue#13582](#), [pr#6975](#), Yehuda Sadeh)
- rgw: convert plain object to versioned (with null version) when removing ([issue#15243](#), [pr#8268](#), Yehuda Sadeh)
- rgw: delete default zone ([pr#7005](#), YankunLi)
- rgw: do not abort radowgw server when using admin op API with bad parameters ([issue#14190](#), [issue#14191](#), [pr#7063](#), Dunrong Huang)
- rgw: Do not send a Content-Type on a ‘304 Not Modified’ response ([issue#15119](#), [pr#8253](#), Wido den Hollander)
- rgw: don’t use s->bucket for metadata api path entry ([issue#14549](#), [pr#7408](#), Yehuda Sadeh)
- rgw: Drop a debugging message ([pr#7280](#), Pete Zaitcev)
- rgw: drop permissions of rgw/civetweb after startup ([issue#13600](#), [pr#8019](#), Karol Mroz)
- rgw: Drop unused usage_exit from rgw_admin.cc ([pr#7632](#), Pete Zaitcev)
- rgw: Ensure xmlns is consistent on S3 responses. ([pr#8526](#), Robin H. Johnson)
- rgw: extend rgw_extended_http_attrs to affect Swift accounts and containers as well ([pr#5969](#), Radoslaw Zarzynski)
- rgw: fcgi should include acconfig ([pr#7760](#), Abhishek Lekshmanan)
- rgw_file: set owner uid, gid, and Unix mode on new objects ([pr#8321](#), Matt Benjamin)

- rgw: file unix ([pr#8563](#), Matt Benjamin)
- rgw: fix a glaring syntax error ([pr#6888](#), Pavan Rallabandi)
- rgw: fix a typo in error message ([pr#8434](#), Abhishek Lekshmanan)
- rgw: fix a typo in init-radosgw ([pr#6817](#), Zhi Zhang)
- rgw: fix binfo_cache, RGWShardCollectCR ([pr#8619](#), Yehuda Sadeh)
- rgw: fix compilation warning ([pr#7160](#), Yehuda Sadeh)
- rgw: fix compiling error ([pr#8394](#), xie xingguo)
- rgw: fixes for per-period metadata logs ([pr#7827](#), Casey Bodley)
- rgw: fix lockdep false positive ([pr#8284](#), Yehuda Sadeh)
- rgw: fix mdlog ([pr#8183](#), Orit Wasserman)
- rgw: fix objects can not be displayed which object name does not cont... ([issue#12963](#), [pr#5738](#), Weijun Duan)
- rgw: fix openssl linkage ([pr#6513](#), Yehuda Sadeh)
- rgw: fix partial read issue in rgw_admin and rgw_tools ([pr#6761](#), Jiaying Ren)
- rgw: fix problem deleting objects begining with double underscores ([issue#15318](#), [pr#8488](#), Orit Wasserman)
- rgw: fix reload on non Debian systems. ([pr#6482](#), Hervé Rousseau)
- rgw: fix response of delete expired objects ([issue#13469](#), [pr#6228](#), Yuan Zhou)
- rgw: Fix subuser harder with tenants ([pr#7618](#), Pete Zaitcev)
- rgw: fix swift API returning incorrect account metadata ([issue#13140](#), [pr#6047](#), Sangdi Xu)
- rgw: fix the build failure ([pr#6927](#), Kefu Chai)
- rgw: fix typo in RGWHTTPClient::process error message ([pr#6424](#), Brad Hubbard)
- rgw: fix unsafe c_str() usage ([issue#15463](#), [pr#8559](#), Sage Weil)
- rgw: fix wrong check for parse() return ([pr#6797](#), Dunrong Huang)
- rgw: fix wrong etag calculation during POST on S3 bucket. ([issue#11241](#), [pr#6030](#), Radoslaw Zarzynski)
- rgw: fix wrong handling of limit=0 during listing of Swift account. ([issue#14903](#), [pr#7821](#), Radoslaw Zarzynski)
- rgw: handle no current_period in is_syncing_bucket_meta ([pr#8597](#), Orit Wasserman)
- rgw: highres time stamps ([pr#8108](#), Yehuda Sadeh, Adam C. Emerson, Matt Benjamin)
- rgw: improve error handling in S3/Keystone integration ([pr#7597](#), Radoslaw Zarzynski)
- rgw: increase verbosity level on RGWObjManifest line ([pr#7285](#), magicrobotmonkey)
- rgw: indexless ([pr#7786](#), Yehuda Sadeh)
- rgw: jewel nfs fixes ([pr#8460](#), Matt Benjamin)
- rgw: keystone v3 ([pr#7719](#), Mark Barnes, Radoslaw Zarzynski)
- rgw: ldap fixes ([pr#8168](#), Matt Benjamin)
- rgw_ldap: make ldap.h inclusion conditional ([pr#8500](#), Matt Benjamin)
- rgw: ldap (Matt Benjamin) ([pr#7985](#), Matt Benjamin)
- rgw: leaks fixes ([pr#8636](#), Yehuda Sadeh)
- rgw: let radosgw-admin bucket stats return a standard json ([pr#7029](#), Ruifeng Yang)
- rgw: link against system openssl (instead of dlopen at runtime) ([pr#6419](#), Sage Weil)
- rgw: link civetweb with openssl (Sage, Marcus Watts) ([pr#7825](#), Marcus Watts, Sage Weil)
- rgw: link payer info to usage logging ([pr#7918](#), Yehuda Sadeh, Javier M. Mellid)
- rgw: mdlog trim add usage prompt ([pr#6059](#), Weijun Duan)
- rgw: modify command stucking when operating radosgw-admin metadata list user ([pr#7032](#), Peiyang Liu)
- rgw: modify documents and help infos' descriptions to the usage of option date when executing command "log show" ([pr#6080](#), Kongming Wu)
- rgw: modify the conditional statement in parse_metadata_key method. ([pr#5875](#), Zengran Zhang)
- rgw: move signal.h dependency from rgw_front.h ([pr#7678](#), Matt Benjamin)
- rgw: Multipart ListPartsResult ETag quotes ([issue#15334](#), [pr#8387](#), Robin H. Johnson)
- rgw: multiple improvements regarding etag calculation for SLO/DLO of Swift API. ([pr#7764](#), Radoslaw Zarzynski)
- rgw: multiple Swift API compliance improvements for TempURL (Radoslaw Zarzynski) ([issue#14806](#), [issue#11163](#), [pr#7891](#), Radoslaw Zarzynski)
- rgw: multisite fixes ([pr#8013](#), Yehuda Sadeh)
- rgw: multitenancy support ([pr#6784](#), Yehuda Sadeh, Pete Zaitcev)
- rgw: new multisite merge ([issue#14549](#), [pr#7709](#), Yehuda Sadeh, Orit Wasserman, Casey Bodley, Daniel Gryniewicz)
- rgw: Parse -subuser better ([pr#7279](#), Pete Zaitcev)
- rgw: prevent anonymous user from reading bucket with authenticated read ACL ([issue#13207](#), [pr#6057](#), root)
- rgw: radosgw-admin bucket check -fix not work ([pr#7093](#), Weijun Duan)
- rgw-rados: return RGWSSystemMetaObj init directly ([pr#8534](#), Karol Mroz)
- rgw: refuse to calculate digest when the s3 secret key is empty ([issue#13133](#), [pr#6045](#), Sangdi Xu)
- rgw : region to zonegroup related fixes ([pr#8620](#), Yehuda Sadeh)
- rgw: remove duplicated code in RGWRados::get_bucket_info() ([pr#7413](#), liyankun)
- rgw: remove extra check in RGWGetObj::execute ([issue#12352](#), [pr#5262](#), Javier M. Mellid)
- rgw: Remove unused code in PutMetadataAccount::execute ([pr#6668](#), Pete Zaitcev)
- rgw: remove unused variable in RGWPutMetadataBucket::execute ([pr#6735](#), Radoslaw Zarzynski)
- rgw: remove unused vector ([pr#7990](#), Na Xie)
- rgw: reset return code in when iterating over the bucket the objects ([issue#14826](#), [pr#7803](#), Orit Wasserman)
- rgw: retry RGWRemoteMetaLog::read_log_info() while master is down ([pr#8453](#), Casey Bodley)

- rgw: Revert “rgw ldap” ([pr#8075](#), Yehuda Sadeh)
- rgw/rgw_admin:fix bug about list and stats command ([pr#8200](#), Qiankun Zheng)
- rgw/rgw_common.h: fix the RGWBucketInfo decoding ([pr#8154](#), Kefu Chai)
- rgw/rgw_common.h: fix the RGWBucketInfo decoding ([pr#8165](#), Kefu Chai)
- rgw: RGWLib::env is not used so remove it ([pr#7874](#), Brad Hubbard)
- rgw/rgw_orphan: check the return value of save_state ([pr#7544](#), Boris Ranto)
- rgw: RGWPeriod::reflect() sets master zonegroup as default ([pr#8566](#), Casey Bodley)
- rgw/rgw_resolve: fallback to res_query when res_nquery not implemented ([pr#6292](#), John Coyle)
- rgw: RGWZoneParams::create should not handle -EEXIST error ([pr#7927](#), Orit Wasserman)
- rgw: S3: set EncodingType in ListBucketResult ([pr#7712](#), Victor Makarov)
- rgw: signature mismatch with escaped characters in url query portion ([issue#15358](#), [pr#8445](#), Javier M. Mellid)
- rgw: static large objects (Radoslaw Zarzynski, Yehuda Sadeh)
- rgw: store system object meta in cache when creating it ([issue#14678](#), [pr#7615](#), Yehuda Sadeh)
- rgw: support core file limit for radosgw daemon ([pr#6346](#), Guang Yang)
- rgw: support for aws authentication v4 (Javier M. Mellid) ([issue#10333](#), [pr#7720](#), Yehuda Sadeh, Javier M. Mellid)
- rgw: support json format for admin policy API (Dunrong Huang) ([issue#14090](#), [pr#8036](#), Yehuda Sadeh, Dunrong Huang)
- rgw: swift bulk delete (Radoslaw Zarzynski)
- rgw: swift use Civetweb ssl can not get right url ([issue#13628](#), [pr#6408](#), Weijun Duan)
- rgw: swift versioning disabled ([pr#8066](#), Yehuda Sadeh, Radoslaw Zarzynski)
- rgw: sync fixes 3 ([pr#8170](#), Yehuda Sadeh)
- rgw: sync fixes 4 ([pr#8190](#), Yehuda Sadeh)
- rgw sync fixes ([pr#8095](#), Yehuda Sadeh)
- rgw: the map ‘headers’ is assigned a wrong value ([pr#8481](#), weiqiaomiao)
- rgw: try to parse Keystone token in order appropriate to configuration. ([pr#7822](#), Radoslaw Zarzynski)
- rgw: try to use current period id in a few more cases ([pr#8588](#), Yehuda Sadeh)
- rgw:Use count fn in RGWUserBuckets for quota check ([pr#8294](#), Abhishek Lekshmanan)
- rgw: use pimpl pattern for RGWPeriodHistory ([pr#7809](#), Casey Bodley)
- rgw: user quota may not adjust on bucket removal ([issue#14507](#), [pr#7586](#), root)
- rgw: use smart pointer for C_Reinitwatch ([pr#6767](#), Orit Wasserman)
- rgw: warn on suspicious civetweb frontend parameters ([pr#6944](#), Matt Benjamin)
- rocksdb: build with PORTABLE=1 ([pr#6311](#), Sage Weil)
- rocksdb: remove ldb source files from dist tarball ([issue#13554](#), [pr#6379](#), Kefu Chai)
- rocksdb: remove ldb sources from dist tarball ([issue#13554](#), [pr#7105](#), Venky Shankar)
- rocksdb: use native rocksdb makefile (and our autotools) ([pr#6290](#), Sage Weil)
- rpm: ceph.spec.in: correctly declare systemd dependency for SLE/openSUSE ([pr#6114](#), Nathan Cutler)
- rpm: ceph.spec.in: fix libs-compat / devel-compat conditional ([issue#12315](#), [pr#5219](#), Ken Dreyer)
- rpm,deb: remove conditional BuildRequires for btrfs-progs ([issue#15042](#), [pr#8016](#), Erwan Velu)
- rpm: move %post(un) ldconfig calls to ceph-base ([issue#14940](#), [pr#7867](#), Nathan Cutler)
- rpm: move runtime dependencies to ceph-base and fix other packaging issues ([issue#14864](#), [pr#7826](#), Nathan Cutler)
- rpm: prefer UID/GID 167 when creating ceph user/group ([issue#15246](#), [pr#8277](#), Nathan Cutler)
- rpm: remove sub-package dependencies on “ceph” ([issue#15146](#), [pr#8137](#), Ken Dreyer)
- rpm: rhel 5.9 librados compile fix, moved blkid to RBD check/compilation ([issue#13177](#), [pr#5954](#), Rohan Mars)
- script: add missing stop_rgw variable to stop.sh script ([pr#7959](#), Karol Mroz)
- scripts: adjust mstart and mstop script to run with cmake build ([pr#6920](#), Orit Wasserman)
- scripts: release_notes can track original issue ([pr#6009](#), Abhishek Lekshmanan)
- script: subscription-manager support ([issue#14972](#), [pr#7907](#), Loic Dachary)
- selinux: allow log files to be located in /var/log/radosgw ([pr#7604](#), Boris Ranto)
- selinux: Update policy to grant additional access ([issue#14870](#), [pr#7971](#), Boris Ranto)
- set 128MB tcmalloc cache size by bytes ([pr#8427](#), Star Guo)
- stringify outputted error code and fix unmatched parentheses. ([pr#6998](#), xie.xingguo, xie xingguo)
- submodules: revert an accidental change ([pr#7929](#), Yehuda Sadeh)
- systemd: correctly escape block device paths ([issue#14706](#), [pr#7579](#), James Page)
- systemd: drop any systemd imposed process/thread limits ([pr#8450](#), James Page)
- systemd: fix typos ([pr#6679](#), Tobias Suckow)
- systemd: set up environment in rbdmap unit file ([issue#14984](#), [pr#8222](#), Nathan Cutler)
- systemd: start/stop/restart ceph services by daemon type ([issue#13497](#), [pr#6276](#), Zhi Zhang)
- systemd: Use the same restart limits as upstart ([pr#8188](#), Boris Ranto)
- sysvinit: allow custom cluster names ([pr#6732](#), Richard Chan)
- test: add missing shut_down mock method ([pr#8125](#), Jason Dillaman)
- test/bufferlist: Avoid false-positive tests ([pr#7955](#), Erwan Velu)
- test: ceph_test_rados: use less CPU ([pr#7513](#), Samuel Just)
- test/cli-integration/rbd: disable progress output ([issue#14931](#), [pr#7858](#), Josh Durgin)
- test: correct librbd errors discovered with unoptimized cmake build ([pr#7914](#), Jason Dillaman)
- test: create pools for rbd tests with different prefix ([pr#7738](#), Mykola Golub)
- test: enable test for bug #2339 which has been resolved. ([pr#7743](#), You Ji)

- test/encoding/readable.sh fix (pr#6714, Igor Podoski)
- test/encoding/readable: use [for “test” not (((pr#8574, Kefu Chai)
- test: fix issues discovered via the rbd permissions test case (pr#8129, Jason Dillaman)
- test: fix osd-scrub-snaps.sh (pr#6697, Xinze Chi)
- test: Fix test to run with btrfs which has snap_### dirs (issue#15347, pr#8420, David Zafman)
- test: fixup and improvements for rbd-mirror test (pr#8090, Mykola Golub)
- test: fix ut test failure caused by lfn change (issue#15464, pr#8544, xie xingguo)
- test: fix valgrind memcheck issues for rbd-mirror test cases (issue#15354, pr#8493, Jason Dillaman)
- test: handle exception thrown from close during rbd lock test (pr#8124, Jason Dillaman)
- test: image replayer needs dummy remote mirror peer uuid (pr#8584, Jason Dillaman)
- test/libcephfs/flock: add sys/file.h include for flock operations (pr#6310, John Coyle)
- test/librados/test.cc: clean up EC pools’ crush rules too (issue#13878, pr#6788, Loic Dachary, Dan Mick)
- test/librbd/fsx: Use c++11 std::mt19937 generator instead of random_r() (pr#6332, John Coyle)
- test/mon/osd-erasure-code-profile: pick new mon port (pr#7161, Sage Weil)
- test: more debug logging for TestWatchNotify (pr#7737, Mykola Golub)
- test: new librbd flatten test case (pr#7609, Jason Dillaman)
- test/osd: Relax the timing intervals in osd-markdown.sh (pr#7899, Dan Mick)
- test_pool_create.sh: put test files in the test dir so they are cleaned up (pr#8219, Josh Durgin)
- test/pybind/test_ceph_argparse: fix reweight-by-utilization tests (pr#8027, Kefu Chai, Sage Weil)
- test/radosgw-admin: update the expected usage outputs (pr#7723, Kefu Chai)
- test: rbd-mirror: add “switch to the next tag” test (pr#8149, Mykola Golub)
- test: rbd-mirror: compare positions using all fields (pr#8172, Mykola Golub)
- test: rbd-mirror: script improvements for manual testing (pr#8325, Mykola Golub)
- test: reproducer for writeback CoW deadlock (pr#8009, Jason Dillaman)
- test/rgw: add multisite test for meta sync across periods (pr#7887, Casey Bodley)
- test_rgw_admin: use freopen for output redirection. (pr#6303, John Coyle)
- tests: add const for ec test (pr#6911, Michal Jarzabek)
- tests: add Ubuntu 16.04 xenial dockerfile (pr#8519, Loic Dachary)
- tests: allow docker-test.sh to run under root (issue#13355, pr#6173, Loic Dachary)
- tests: allow object corpus readable test to skip specific incompat instances (pr#6932, Igor Podoski)
- tests: centos7 needs the Continuous Release (CR) Repository enabled for (issue#13997, pr#6842, Brad Hubbard)
- tests: ceph-disk.sh: should use “readlink -f” instead (pr#7594, Kefu Chai)
- tests: ceph-disk.sh: use “readlink -f” instead for fullpath (pr#7606, Kefu Chai)
- tests: ceph-disk workunit uses configobj (pr#6342, Loic Dachary)
- tests: ceph-helpers assert success getting backfills (pr#6699, Loic Dachary)
- tests: ceph_test_keyvaluedb_iterators: fix broken test (pr#6597, Haomai Wang)
- tests: concatenate test_rados_test_tool from src and qa (issue#13691, pr#6464, Loic Dachary)
- tests: configure with rocksdb by default (issue#14220, pr#7100, Loic Dachary)
- tests: destroy testprofile before creating one (issue#13664, pr#6446, Loic Dachary)
- test: set a default \$CEPH_ROOT env variable (pr#8645, Kefu Chai)
- tests: fix a few build warnings (pr#7608, Sage Weil)
- tests: fix failure for osd-scrub-snap.sh (issue#13986, pr#6890, Loic Dachary, Ning Yao)
- tests: Fix for make check. (pr#7102, David Zafman)
- tests: Fixing broken test/cephtool-test-mon.sh test (pr#8429, Erwan Velu)
- tests: fix race condition testing auto scrub (issue#13592, pr#6724, Xinze Chi, Loic Dachary)
- tests: fix test_rados_tools.sh rados lookup (issue#13691, pr#6502, Loic Dachary)
- tests: fix typo in TestCIsRbd.snapshots test case (issue#13727, pr#6504, Jason Dillaman)
- tests: flush op work queue prior to destroying MockImageCtx (issue#14092, pr#7002, Jason Dillaman)
- tests: ignore test-suite.log (pr#6584, Loic Dachary)
- tests: Improving ‘make check’ execution time (pr#8131, Erwan Velu)
- tests: notification slave needs to wait for master (issue#13810, pr#7220, Jason Dillaman)
- tests: -osd-scrub-load-threshold=2000 for more consistency (issue#14027, pr#6871, Loic Dachary)
- tests: osd-scrub-snaps.sh to display full osd logs on error (issue#13986, pr#6857, Loic Dachary)
- tests: port uniqueness reminder (pr#6387, Loic Dachary)
- tests: restore run-cli-tests (pr#6571, Loic Dachary, Sage Weil, Jason Dillaman)
- tests: snap rename and rebuild object map in client update test (pr#7224, Jason Dillaman)
- tests: sync ceph-erasure-code-corpus for mktemp -d (pr#7596, Loic Dachary)
- tests: test/librados/test.cc must create profile (issue#13664, pr#6452, Loic Dachary)
- tests: test_pidfile.sh lingering processes (issue#14834, pr#7734, Loic Dachary)
- tests: unittest_bufferlist: fix hexdump test (pr#7152, Sage Weil)
- tests: unittest_ipaddr: fix segv (pr#7154, Sage Weil)
- test/system/rados_list_parallel: print oid if rados_write fails (issue#15240, pr#8309, Kefu Chai)
- test/system/*: use dynamically generated pool name (issue#15240, pr#8318, Kefu Chai)
- test/test-erasure-code.sh: disable pg_temp priming (issue#15211, pr#8260, Sage Weil)
- test: TestMirroringWatcher test cases were not closing images (pr#8435, Jason Dillaman)

- test/TestPGLog: fix the FTBFS ([issue#14930](#), [pr#7855](#), Kefu Chai)
- test/test_pool_create.sh: fix port ([pr#8361](#), Sage Weil)
- test/time: no need to abs(uint64_t) for comparing ([pr#7726](#), Kefu Chai)
- test: update rbd integration cram tests for new default features ([pr#8001](#), Jason Dillaman)
- test: use sequential journal_tid for object cacher test ([issue#13877](#), [pr#6710](#), Josh Durgin)
- tools: add cephfs-table-tool ‘take_inos’ ([pr#6655](#), John Spray)
- tools/cephfs: add tmap_upgrade ([pr#7003](#), John Spray)
- tools/cephfs: fix overflow writing header to fixed size buffer (#13816) ([pr#6617](#), John Spray)
- tools/cephfs: fix overflow writing header to fixed size buffer (#13816) ([pr#6617](#), John Spray)
- tools/cephfs: fix tmap_upgrade ([issue#15135](#), [pr#8128](#), John Spray)
- tools: ceph_monstore_tool: add inflate-pgmap command ([issue#14217](#), [pr#7097](#), Kefu Chai)
- tools: ceph-monstore-update-crush: add “-test” when testing crushmap ([pr#6418](#), Kefu Chai)
- tools: Fix layout handing in cephfs-data-scan (#13898) ([pr#6719](#), John Spray)
- tools: monstore: add ‘show-versions’ command. ([pr#7073](#), Cilang Zhao)
- tools/rados: reduce “rados put” memory usage by op_size ([pr#7928](#), Piotr Dałek)
- tools:remove duplicate references ([pr#5917](#), Bo Cai)
- tools: support printing part cluster map in readable fashion ([issue#13079](#), [pr#5921](#), Bo Cai)
- unittest_compression_zlib: do not assume buffer will be null terminated ([pr#8064](#), Sage Weil)
- unittest_erasure_code_plugin: fix deadlock (Alpine) ([pr#8314](#), John Coyle)
- unittest_osdmap: default crush tunables now firefly ([pr#8098](#), Sage Weil)
- vstart: fix up cmake paths when VSTART_DEST is given ([pr#8363](#), Casey Bodley)
- vstart: grant full access to Swift testing account ([pr#6239](#), Yuan Zhou)
- vstart: make -k with optional mon_num. ([pr#8251](#), Jianpeng Ma)
- vstart: set cephfs root uid/gid to caller ([pr#6255](#), John Spray)
- vstart.sh: add mstart, mstop, mrun wrappers for running multiple vstart-style test clusters out of src tree ([pr#6901](#), Yehuda Sadeh)
- vstart.sh: avoid race condition starting rgw via vstart.sh ([issue#14829](#), [pr#7727](#), Javier M. Mellid)
- vstart.sh: silence a harmless msg where btrfs is not found ([pr#7640](#), Patrick Donnelly)
- xio: add prefix to xio msgr logs ([pr#8148](#), Roi Dayan)
- xio: fix compilation against latest accelio ([pr#8022](#), Roi Dayan)
- xio: fix incorrect ip being assigned in case of multiple RDMA ports ([pr#7747](#), Subramanyam Varanasi)
- xio: remove duplicate assignment of peer addr ([pr#8025](#), Roi Dayan)
- xio: remove redundant magic methods ([pr#7773](#), Roi Dayan)
- xio: remove unused variable ([pr#8023](#), Roi Dayan)
- xio: xio_init needs to be called before any other xio function ([pr#8227](#), Roi Dayan)
- xxhash: use clone of xxhash.git; add .gitignore ([pr#7986](#), Sage Weil)

NOTABLE CHANGES SINCE HAMMER

- aarch64: add optimized version of crc32c (Yazen Ghannam, Steve Capper)
- auth: cache/reuse crypto lib key objects, optimize msg signature check (Sage Weil)
- auth: reinit NSS after fork() (#11128 Yan, Zheng)
- autotools: fix out of tree build (Krzysztof Kosinski)
- autotools: improve make check output (Loic Dachary)
- buffer: add invalidate_crc() (Piotr Dalek)
- buffer: fix zero bug (#12252 Haomai Wang)
- buffer: some cleanup (Michał Jarzabek)
- build: allow tcmalloc-minimal (Thorsten Behrens)
- build: C++11 now supported
- build: cmake: fix nss linking (Danny Al-Gaaf)
- build: cmake: misc fixes (Orit Wasserman, Casey Bodley)
- build: disable LTTNG by default (#11333 Josh Durgin)
- build: do not build ceph-dencoder with tcmalloc (#10691 Boris Ranto)
- build: fix junit detection on Fedora 22 (Ira Cooper)
- build: fix pg ref disabling (William A. Kennington III)
- build: fix ppc build (James Page)
- build: install-deps: misc fixes (Loic Dachary)
- build: install-deps.sh improvements (Loic Dachary)
- build: install-deps: support OpenSUSE (Loic Dachary)
- build: make_dist_tarball.sh (Sage Weil)
- build: many cmake improvements
- build: misc cmake fixes (Matt Benjamin)
- build: misc fixes (Boris Ranto, Ken Dreyer, Owen Synge)
- build: OSX build fixes (Yan, Zheng)

- build: remove rest-bench
- ceph-authtool: fix return code on error (Gerhard Muntingh)
- ceph-detect-init: added Linux Mint (Michal Jarzabek)
- ceph-detect-init: robust init system detection (Owen Synge)
- ceph-disk: ensure ‘zap’ only operates on a full disk (#11272 Loic Dachary)
- ceph-disk: fix zap sgdisk invocation (Owen Synge, Thorsten Behrens)
- ceph-disk: follow ceph-osd hints when creating journal (#9580 Sage Weil)
- ceph-disk: handle re-using existing partition (#10987 Loic Dachary)
- ceph-disk: improve parted output parsing (#10983 Loic Dachary)
- ceph-disk: install pip > 6.1 (#11952 Loic Dachary)
- ceph-disk: make suppression work for activate-all and activate-journal (Dan van der Ster)
- ceph-disk: many fixes (Loic Dachary, Alfredo Deza)
- ceph-disk: fixes to respect init system (Loic Dachary, Owen Synge)
- ceph-disk: pass –cluster arg on prepare subcommand (Kefu Chai)
- ceph-disk: support for multipath devices (Loic Dachary)
- ceph-disk: support NVMe device partitions (#11612 Ilja Slepnev)
- ceph: fix ‘df’ units (Zhe Zhang)
- ceph: fix parsing in interactive cli mode (#11279 Kefu Chai)
- cephfs-data-scan: many additions, improvements (John Spray)
- ceph-fuse: do not require successful remount when unmounting (#10982 Greg Farnum)
- ceph-fuse, libcephfs: don’t clear COMPLETE when trimming null (Yan, Zheng)
- ceph-fuse, libcephfs: drop inode when rmdir finishes (#11339 Yan, Zheng)
- ceph-fuse, libcephfs: fix uninlne (#11356 Yan, Zheng)
- ceph-fuse, libcephfs: hold exclusive caps on dirs we “own” (#11226 Greg Farnum)
- ceph-fuse: mostly behave on 32-bit hosts (Yan, Zheng)
- ceph: improve error output for ‘tell’ (#11101 Kefu Chai)
- ceph-monstore-tool: fix store-copy (Huangjun)
- ceph: new ‘ceph daemonperf’ command (John Spray, Mykola Golub)
- ceph-objectstore-tool: many many improvements (David Zafman)
- ceph-objectstore-tool: refactoring and cleanup (John Spray)
- ceph-post-file: misc fixes (Joey McDonald, Sage Weil)
- ceph_test_rados: test pipelined reads (Zhiqiang Wang)
- client: avoid sending unnecessary FLUSHSNAP messages (Yan, Zheng)
- client: exclude setfilelock when calculating oldest tid (Yan, Zheng)
- client: fix error handling in check_pool_perm (John Spray)
- client: fsync waits only for inode’s caps to flush (Yan, Zheng)
- client: invalidate kernel dcache when cache size exceeds limits (Yan, Zheng)
- client: make fsync wait for unsafe dir operations (Yan, Zheng)
- client: pin lookup dentry to avoid inode being freed (Yan, Zheng)
- common: add descriptions to perfcounters (Kiseleva Alyona)
- common: add perf counter descriptions (Alyona Kiseleva)
- common: bufferlist performance tuning (Piotr Dalek, Sage Weil)
- common: detect overflow of int config values (#11484 Kefu Chai)
- common: fix bit_vector extent calc (#12611 Jason Dillaman)
- common: fix json parsing of utf8 (#7387 Tim Serong)
- common: fix leak of pthread_mutexattr (#11762 Ketor Meng)
- common: fix LTTNG vs fork issue (Josh Durgin)
- common: fix throttle max change (Henry Chang)
- common: make mutex more efficient
- common: make work queue addition/removal thread safe (#12662 Jason Dillaman)
- common: optracker improvements (Zhiqiang Wang, Jianpeng Ma)
- common: PriorityQueue tests (Kefu Chai)
- common: some async compression infrastructure (Haomai Wang)
- crush: add –check to validate dangling names, max osd id (Kefu Chai)
- crush: cleanup, sync with kernel (Ilya Dryomov)
- crush: fix crash from invalid ‘take’ argument (#11602 Shiva Rkreddy, Sage Weil)
- crush: fix divide-by-2 in straw2 (#11357 Yann Dupont, Sage Weil)
- crush: fix has_v4_buckets (#11364 Sage Weil)
- crush: fix subtree base weight on adjust_subtree_weight (#11855 Sage Weil)
- crush: respect default replicated ruleset config on map creation (Ilya Dryomov)
- crushtool: fix order of operations, usage (Sage Weil)
- crypto: fix NSS leak (Jason Dillaman)
- crypto: fix unbalanced init/shutdown (#12598 Zheng Yan)
- deb: fix rest-bench-dbg and ceph-test-dbg dependendies (Ken Dreyer)
- debian: minor package reorg (Ken Dreyer)

- deb, rpm: move ceph-objectstore-tool to ceph (Ken Dreyer)
- doc: docuemnt object corpus generation (#11099 Alexis Normand)
- doc: document region hostnames (Robin H. Johnson)
- doc: fix gender neutrality (Alexandre Maragone)
- doc: fix install doc (#10957 Kefu Chai)
- doc: fix sphinx issues (Kefu Chai)
- doc: man page updates (Kefu Chai)
- doc: mds data structure docs (Yan, Zheng)
- doc: misc updates (Fracois Lafont, Ken Dreyer, Kefu Chai, Owen Synge, Gael Fenet-Garde, Loic Dachary, Yannick Atchyd-Dalama, Jiaying Ren, Kevin Caradant, Robert Maxime, Nicolas Yong, Germain Chipaux, Arthur Gorjux, Gabriel Sentucq, Clement Lebrun, Jean-Remi Deveaux, Clair Massot, Robin Tang, Thomas Laumondais, Jordan Dorne, Yuan Zhou, Valentin Thomas, Pierre Chaumont, Benjamin Troquereau, Benjamin Sesia, Vikhyat Umrao, Nilamdyuti Goswami, Vartika Rai, Florian Haas, Loic Dachary, Simon Guinot, Andy Allan, Alistair Israel, Ken Dreyer, Robin Rehu, Lee Revell, Florian Marsylle, Thomas Johnson, Bosse Klykken, Travis Rhoden, Ian Kelling)
- doc: swift tempurls (#10184 Abhishek Lekshmanan)
- doc: switch doxygen integration back to breathe (#6115 Kefu Chai)
- doc: update release schedule docs (Loic Dachary)
- erasure-code: cleanup (Kefu Chai)
- erasure-code: improve tests (Loic Dachary)
- erasure-code: shec: fix recovery bugs (Takanori Nakao, Shotaro Kawaguchi)
- erasure-code: update ISA-L to 2.13 (Yuan Zhou)
- gmock: switch to submodule (Danny Al-Gaaf, Loic Dachary)
- hadoop: add terasort test (Noah Watkins)
- init-radosgw: merge with sysv version; fix enumeration (Sage Weil)
- java: fix libcephfs bindings (Noah Watkins)
- libcephfs: add pread, pwrite (Jevon Qiao)
- libcephfs,ceph-fuse: cache cleanup (Zheng Yan)
- libcephfs,ceph-fuse: fix request resend on cap reconnect (#10912 Yan, Zheng)
- librados: add config observer (Alistair Strachan)
- librados: add FULL_TRY and FULL_FORCE flags for dealing with full clusters or pools (Sage Weil)
- librados: add src_fadvise_flags for copy-from (Jianpeng Ma)
- librados: define C++ flags from C constants (Josh Durgin)
- librados: fadvise flags per op (Jianpeng Ma)
- librados: fix last_force_resent handling (#11026 Jianpeng Ma)
- librados: fix memory leak from C_TwoContexts (Xiong Yiliang)
- librados: fix notify completion race (#13114 Sage Weil)
- librados: fix striper when stripe_count = 1 and stripe_unit != object_size (#11120 Yan, Zheng)
- librados, libcephfs: randomize client nonces (Josh Durgin)
- librados: op perf counters (John Spray)
- librados: pybind: fix binary omap values (Robin H. Johnson)
- librados: pybind: fix write() method return code (Javier Guerra)
- librados: respect default_crush_ruleset on pool_create (#11640 Yuan Zhou)
- libradosstriper: fix leak (Danny Al-Gaaf)
- librbd: add const for single-client-only features (Josh Durgin)
- librbd: add deep-flatten operation (Jason Dillaman)
- librbd: add purge_on_error cache behavior (Jianpeng Ma)
- librbd: allow additional metadata to be stored with the image (Haomai Wang)
- librbd: avoid blocking aio API methods (#11056 Jason Dillaman)
- librbd: better handling for dup flatten requests (#11370 Jason Dillaman)
- librbd: cancel in-flight ops on watch error (#11363 Jason Dillaman)
- librbd: default new images to format 2 (#11348 Jason Dillaman)
- librbd: fadvise for copy, export, import (Jianpeng Ma)
- librbd: fast diff implementation that leverages object map (Jason Dillaman)
- librbd: fix fast diff bugs (#11553 Jason Dillaman)
- librbd: fix image format detection (Zhiqiang Wang)
- librbd: fix lock ordering issue (#11577 Jason Dillaman)
- librbd: fix reads larger than the cache size (Lu Shi)
- librbd: fix snapshot creation when other snap is active (#11475 Jason Dillaman)
- librbd: flatten/copyup fixes (Jason Dillaman)
- librbd: handle NOCACHE fadvise flag (Jianpeng Ma)
- librbd: lockdep, helgrind validation (Jason Dillaman, Josh Durgin)
- librbd: metadata filter fixes (Haomai Wang)
- librbd: misc aio fixes (#5488 Jason Dillaman)
- librbd: misc rbd fixes (#11478 #11113 #11342 #11380 Jason Dillaman, Zhiqiang Wang)
- librbd: new diff_iterate2 API (Jason Dillaman)

- librbd: object map rebuild support (Jason Dillaman)
- librbd: only update image flags while hold exclusive lock (#11791 Jason Dillaman)
- librbd: optionally disable allocation hint (Haomai Wang)
- librbd: prevent race between resize requests (#12664 Jason Dillaman)
- librbd: readahead fixes (Zhiqiang Wang)
- librbd: return result code from close (#12069 Jason Dillaman)
- librbd: store metadata, including config options, in image (Haomai Wang)
- librbd: tolerate old osds when getting image metadata (#11549 Jason Dillaman)
- librbd: use write_full when possible (Zhiqiang Wang)
- log: fix data corruption race resulting from log rotation (#12465 Samuel Just)
- logrotate.d: prefer service over invoke-rc.d (#11330 Win Hierman, Sage Weil)
- mds: add 'damaged' state to MDSMap (John Spray)
- mds: add nicknames for perfcounters (John Spray)
- mds: avoid emitting cap warnigns before evicting session (John Spray)
- mds: avoid getting stuck in XLOCKDONE (#11254 Yan, Zheng)
- mds: disable problematic rstat propagation into snap parents (Yan, Zheng)
- mds: do not add snapped items to bloom filter (Yan, Zheng)
- mds: expose frags via asok (John Spray)
- mds: fix expected holes in journal objects (#13167 Yan, Zheng)
- mds: fix handling for missing mydir dirfrag (#11641 John Spray)
- mds: fix integer truncateion on large client ids (Henry Chang)
- mds: fix mydir replica issue with shutdown (#10743 John Spray)
- mds: fix out-of-order messages (#11258 Yan, Zheng)
- mds: fix rejoin (Yan, Zheng)
- mds: fix setting entire file layout in one setxattr (John Spray)
- mds: fix shutdown (John Spray)
- mds: fix shutdown with strays (#10744 John Spray)
- mds: fix SnapServer crash on deleted pool (John Spray)
- mds: fix snapshot bugs (Yan, Zheng)
- mds: fix stray reintegration (Yan, Zheng)
- mds: fix stray handling (John Spray)
- mds: fix suicide beacon (John Spray)
- mds: flush immediately in do_open_truncate (#11011 John Spray)
- mds: handle misc corruption issues (John Spray)
- mds: improve dump methods (John Spray)
- mds: many fixes (Yan, Zheng, John Spray, Greg Farnum)
- mds: many snapshot and stray fixes (Yan, Zheng)
- mds: misc fixes (Jianpeng Ma, Dan van der Ster, Zhang Zhi)
- mds: misc journal cleanups and fixes (#10368 John Spray)
- mds: misc repair improvements (John Spray)
- mds: misc snap fixes (Zheng Yan)
- mds: misc snapshot fixes (Yan, Zheng)
- mds: new SessionMap storage using omap (#10649 John Spray)
- mds: persist completed_requests reliably (#11048 John Spray)
- mds: reduce memory consumption (Yan, Zheng)
- mds: respawn instead of suicide on blacklist (John Spray)
- mds: separate safe_pos in Journaler (#10368 John Spray)
- mds: snapshot rename support (#3645 Yan, Zheng)
- mds: store layout on header object (#4161 John Spray)
- mds: throttle purge stray operations (#10390 John Spray)
- mds: tolerate clock jumping backwards (#11053 Yan, Zheng)
- mds: warn when clients fail to advance oldest_client_tid (#10657 Yan, Zheng)
- misc cleanups and fixes (Danny Al-Gaaf)
- misc coverity fixes (Danny Al-Gaaf)
- misc performance and cleanup (Nathan Cutler, Xinxin Shu)
- mon: add cache over MonitorDBStore (Kefu Chai)
- mon: add 'mon_metadata <id>' command (Kefu Chai)
- mon: add 'node ls ...' command (Kefu Chai)
- mon: add NOFORWARD, OBSOLETE, DEPRECATE flags for mon commands (Joao Eduardo Luis)
- mon: add PG count to 'ceph osd df' output (Michal Jarzabek)
- mon: 'ceph osd metadata' can dump all osds (Haomai Wang)
- mon: clean up, reorg some mon commands (Joao Eduardo Luis)
- monclient: flush_log (John Spray)
- mon: detect kv backend failures (Sage Weil)
- mon: disallow >2 tiers (#11840 Kefu Chai)

- mon: disallow ec pools as tiers (#11650 Samuel Just)
- mon: do not deactivate last mds (#10862 John Spray)
- mon: fix average utilization calc for 'osd df' (Mykola Golub)
- mon: fix CRUSH map test for new pools (Sage Weil)
- mon: fix log dump crash when debugging (Mykola Golub)
- mon: fix mds beacon replies (#11590 Kefu Chai)
- mon: fix metadata update race (Mykola Golub)
- mon: fix min_last_epoch_clean tracking (Kefu Chai)
- mon: fix 'pg ls' sort order, state names (#11569 Kefu Chai)
- mon: fix refresh (#11470 Joao Eduardo Luis)
- mon: fix variance calc in 'osd df' (Sage Weil)
- mon: improve callout to crushtool (Mykola Golub)
- mon: make blocked op messages more readable (Jianpeng Ma)
- mon: make osd get pool 'all' only return applicable fields (#10891 Michal Jarzabek)
- mon: misc scaling fixes (Sage Weil)
- mon: normalize erasure-code profile for storage and comparison (Loic Dachary)
- mon: only send mon metadata to supporting peers (Sage Weil)
- mon: optionally specify osd id on 'osd create' (Mykola Golub)
- mon: 'osd tree' fixes (Kefu Chai)
- mon: periodic background scrub (Joao Eduardo Luis)
- mon: prevent bucket deletion when referenced by a crush rule (#11602 Sage Weil)
- mon: prevent pgp_num > pg_num (#12025 Xinxin Shu)
- mon: prevent pool with snapshot state from being used as a tier (#11493 Sage Weil)
- mon: prime pg_temp when CRUSH map changes (Sage Weil)
- mon: refine check_remove_tier checks (#11504 John Spray)
- mon: reject large max_mds values (#12222 John Spray)
- mon: remove spurious who arg from 'mds rm ...' (John Spray)
- mon: streamline session handling, fix memory leaks (Sage Weil)
- mon: upgrades must pass through hammer (Sage Weil)
- mon: warn on bogus cache tier config (Jianpeng Ma)
- msgr: add ceph_perf_msgr tool (Hoamai Wang)
- msgr: async: fix seq handling (Hoamai Wang)
- msgr: async: many many fixes (Hoamai Wang)
- msgr: simple: fix clear_pipe (#11381 Haomai Wang)
- msgr: simple: fix connect_seq assert (Haomai Wang)
- msgr: xio: fastpath improvements (Raju Kurunkad)
- msgr: xio: fix ip and nonce (Raju Kurunkad)
- msgr: xio: improve lane assignment (Vu Pham)
- msgr: xio: sync with accellio v1.4 (Vu Pham)
- msgr: xio: misc fixes (#10735 Matt Benjamin, Kefu Chai, Danny Al-Gaaf, Raju Kurunkad, Vu Pham, Casey Bodley)
- msg: unit tests (Haomai Wang)
- objectcacher: misc bug fixes (Jianpeng Ma)
- osd: add latency perf counters for tier operations (Xinze Chi)
- osd: add misc perfcounters (Xinze Chi)
- osd: add simple sleep injection in recovery (Sage Weil)
- osd: allow SEEK_HOLE/SEEK_DATA for sparse read (Zhiqiang Wang)
- osd: avoid dup omap sets for in pg metadata (Sage Weil)
- osd: avoid multiple hit set insertions (Zhiqiang Wang)
- osd: avoid transaction append in some cases (Sage Weil)
- osd: break PG removal into multiple iterations (#10198 Guang Yang)
- osd: cache proxy-write support (Zhiqiang Wang, Samuel Just)
- osd: check scrub state when handling map (Jianpeng Ma)
- osd: clean up some constness, privateness (Kefu Chai)
- osd: clean up temp object if promotion fails (Jianpeng Ma)
- osd: configure promotion based on write recency (Zhiqiang Wang)
- osd: constrain collections to meta and PGs (normal and temp) (Sage Weil)
- osd: don't send dup MMonGetOSDMap requests (Sage Weil, Kefu Chai)
- osd: EIO injection (David Zhang)
- osd: elminiate txn apend, ECSubWrite copy (Samuel Just)
- osd: erasure-code: drop entries according to LRU (Andreas-Joachim Peters)
- osd: erasure-code: fix SHEC floating point bug (#12936 Loic Dachary)
- osd: erasure-code: update to ISA-L 2.14 (Yuan Zhou)
- osd: filejournal: cleanup (David Zafman)
- osd: filestore: clone using splice (Jianpeng Ma)
- osd: filestore: fix recursive lock (Xinxin Shu)

- osd: fix check_for_full (Henry Chang)
- osd: fix dirty accounting in make_writeable (Zhiqiang Wang)
- osd: fix dup promotion lost op bug (Zhiqiang Wang)
- osd: fix endless repair when object is unrecoverable (Jianpeng Ma, Kefu Chai)
- osd: fix hitset object naming to use GMT (Kefu Chai)
- osd: fix misc memory leaks (Sage Weil)
- osd: fix negative degraded stats during backfill (Guang Yang)
- osd: fix osdmap dump of blacklist items (John Spray)
- osd: fix peek_queue locking in FileStore (Xinze Chi)
- osd: fix pg resurrection (#11429 Samuel Just)
- osd: fix promotion vs full cache tier (Samuel Just)
- osd: fix replay requeue when pg is still activating (#13116 Samuel Just)
- osd: fix scrub stat bugs (Sage Weil, Samuel Just)
- osd: fix snap flushing from cache tier (again) (#11787 Samuel Just)
- osd: fix snap handling on promotion (#11296 Sam Just)
- osd: fix temp-clearing (David Zafman)
- osd: force promotion for ops EC can't handle (Zhiqiang Wang)
- osd: handle log split with overlapping entries (#11358 Samuel Just)
- osd: ignore non-existent osds in unfound calc (#10976 Mykola Golub)
- osd: improve behavior on machines with large memory pages (Steve Capper)
- osd: include a temp namespace within each collection/pgid (Sage Weil)
- osd: increase default max open files (Owen Synge)
- osd: keyvaluestore: misc fixes (Varada Kari)
- osd: low and high speed flush modes (Mingxin Liu)
- osd: make suicide timeouts individually configurable (Samuel Just)
- osd: merge multiple setattr calls into a setattrs call (Xinxin Shu)
- osd: misc fixes (Ning Yao, Kefu Chai, Xinze Chi, Zhiqiang Wang, Jianpeng Ma)
- osd: move scrub in OpWQ (Samuel Just)
- osd: newstore prototype (Sage Weil)
- osd: ObjectStore internal API refactor (Sage Weil)
- osd: peer_features includes self (David Zafman)
- osd: pool size change triggers new interval (#11771 Samuel Just)
- osd: prepopulate needs_recovery_map when only one peer has missing (#9558 Guang Yang)
- osd: randomize scrub times (#10973 Kefu Chai)
- osd: recovery, peering fixes (#11687 Samuel Just)
- osd: refactor scrub and digest recording (Sage Weil)
- osd: refuse first write to EC object at non-zero offset (Jianpeng Ma)
- osd: relax reply order on proxy read (#11211 Zhiqiang Wang)
- osd: require firefly features (David Zafman)
- osd: set initial crush weight with more precision (Sage Weil)
- osd: SHEC no longer experimental
- osd: skip promotion for flush/evict op (Zhiqiang Wang)
- osd: stripe over small xattrs to fit in XFS's 255 byte inline limit (Sage Weil, Ning Yao)
- osd: sync object_map on syncfs (Samuel Just)
- osd: take excl lock of op is rw (Samuel Just)
- osd: throttle evict ops (Yunchuan Wen)
- osd: upgrades must pass through hammer (Sage Weil)
- osd: use a temp object for recovery (Sage Weil)
- osd: use blkid to collection partition information (Joseph Handzik)
- osd: use SEEK_HOLE / SEEK_DATA for sparse copy (Xinxin Shu)
- osd: WBThrottle cleanups (Jianpeng Ma)
- osd: write journal header on clean shutdown (Xinze Chi)
- osdc/Objecter: allow per-pool calls to op_cancel_writes (John Spray)
- os/filestore: enlarge getxattr buffer size (Jianpeng Ma)
- pybind: pep8 cleanups (Danny Al-Gaaf)
- pycephfs: many fixes for bindings (Haomai Wang)
- qa: fix filelock_interrupt.py test (Yan, Zheng)
- qa: improve ceph-disk tests (Loic Dachary)
- qa: improve docker build layers (Loic Dachary)
- qa: run-make-check.sh script (Loic Dachary)
- rados: add -striper option to use libradosstriper (#10759 Sebastien Ponce)
- rados: bench: add -no-verify option to improve performance (Piotr Dalek)
- rados bench: misc fixes (Dmitry Yatsushkevich)
- rados: fix error message on failed pool removal (Wido den Hollander)
- radosgw-admin: add 'bucket check' function to repair bucket index (Yehuda Sadeh)

- radosgw-admin: fix subuser modify output (#12286 Guce)
- rados: handle -snapid arg properly (Abhishek Lekshmanan)
- rados: improve bench buffer handling, performance (Piotr Dalek)
- rados: misc bench fixes (Dmitry Yatsushkevich)
- rados: new pool import implementation (John Spray)
- rados: translate errno to string in CLI (#10877 Kefu Chai)
- rbd: accept map options config option (Ilya Dryomov)
- rbd: add disk usage tool (#7746 Jason Dillaman)
- rbd: allow unmapping by spec (Ilya Dryomov)
- rbd: cli: fix arg parsing with -io-pattern (Dmitry Yatsushkevich)
- rbd: deprecate -new-format option (Jason Dillman)
- rbd: fix error messages (#2862 Rajesh Nambiar)
- rbd: fix link issues (Jason Dillaman)
- rbd: improve CLI arg parsing, usage (Ilya Dryomov)
- rbd: rbd-replay-prep and rbd-replay improvements (Jason Dillaman)
- rbd: recognize queue_depth kernel option (Ilya Dryomov)
- rbd: support G and T units for CLI (Abhishek Lekshmanan)
- rbd: update rbd man page (Ilya Dryomov)
- rbd: update xfstests tests (Douglas Fuller)
- rbd: use image-spec and snap-spec in help (Vikhyat Umrao, Ilya Dryomov)
- rest-bench: misc fixes (Shawn Chen)
- rest-bench: support https (#3968 Yuan Zhou)
- rgw: add max multipart upload parts (#12146 Abhishek Dixit)
- rgw: add missing headers to Swift container details (#10666 Ahmad Faheem, Dmytro Iurchenko)
- rgw: add stats to headers for account GET (#10684 Yuan Zhou)
- rgw: add Trasnaction-Id to response (Abhishek Dixit)
- rgw: add X-Timestamp for Swift containers (#10938 Radoslaw Zarzynski)
- rgw: always check if token is expired (#11367 Anton Aksola, Riku Lehto)
- rgw: conversion tool to repair broken multipart objects (#12079 Yehuda Sadeh)
- rgw: document layout of pools and objects (Pete Zaitcev)
- rgw: do not enclose bucket header in quotes (#11860 Wido den Hollander)
- rgw: do not prefetch data for HEAD requests (Guang Yang)
- rgw: do not preserve ACLs when copying object (#12370 Yehuda Sadeh)
- rgw: do not set content-type if length is 0 (#11091 Orit Wasserman)
- rgw: don't clobber bucket/object owner when setting ACLs (#10978 Yehuda Sadeh)
- rgw: don't use end_marker for namespaced object listing (#11437 Yehuda Sadeh)
- rgw: don't use rgw_socket_path if frontend is configured (#11160 Yehuda Sadeh)
- rgw: enforce Content-Length for POST on Swift cont/obj (#10661 Radoslaw Zarzynski)
- rgw: error out if frontend did not send all data (#11851 Yehuda Sadeh)
- rgw: expose the number of unhealthy workers through admin socket (Guang Yang)
- rgw: fail if parts not specified on multipart upload (#11435 Yehuda Sadeh)
- rgw: fix assignment of copy obj attributes (#11563 Yehuda Sadeh)
- rgw: fix broken stats in container listing (#11285 Radoslaw Zarzynski)
- rgw: fix bug in domain/subdomain splitting (Robin H. Johnson)
- rgw: fix casing of Content-Type header (Robin H. Johnson)
- rgw: fix civetweb max threads (#10243 Yehuda Sadeh)
- rgw: fix Connection: header handling (#12298 Wido den Hollander)
- rgw: fix copy metadata, support X-Copied-From for swift (#10663 Radoslaw Zarzynski)
- rgw: fix data corruptions race condition (#11749 Wuxingyi)
- rgw: fix decoding of X-Object-Manifest from GET on Swift DLO (Radslow Rzarzynski)
- rgw: fix GET on swift account when limit == 0 (#10683 Radoslaw Zarzynski)
- rgw: fix handling empty metadata items on Swift container (#11088 Radoslaw Zarzynski)
- rgw: fix JSON response when getting user quota (#12117 Wuxingyi)
- rgw: fix locator for objects starting with _ (#11442 Yehuda Sadeh)
- rgw: fix log rotation (Wuxingyi)
- rgw: fix multipart upload in retry path (#11604 Yehuda Sadeh)
- rgw: fix quota enforcement on POST (#11323 Sergey Arkhipov)
- rgw: fix reset_loc (#11974 Yehuda Sadeh)
- rgw: fix return code on missing upload (#11436 Yehuda Sadeh)
- rgw: fix sysvinit script
- rgw: fix sysvinit script w/ multiple instances (Sage Weil, Pavan Rallabhandi)
- rgw: force content_type for swift bucket stats requests (#12095 Orit Wasserman)
- rgw: force content type header on responses with no body (#11438 Orit Wasserman)
- rgw: generate Date header for civetweb (#10873 Radoslaw Zarzynski)
- rgw: generate new object tag when setting attrs (#11256 Yehuda Sadeh)

- rgw: improve content-length env var handling (#11419 Robin H. Johnson)
- rgw: improved support for swift account metadata (Radoslaw Zarzynski)
- rgw: improve handling of already removed buckets in expirer (Radoslaw Rzarzynski)
- rgw: issue aio for first chunk before flush cached data (#11322 Guang Yang)
- rgw: log to /var/log/ceph instead of /var/log/radosgw
- rgw: make init script wait for radosgw to stop (#11140 Dmitry Yatsushkevich)
- rgw: make max put size configurable (#6999 Yuan Zhou)
- rgw: make quota/gc threads configurable (#11047 Guang Yang)
- rgw: make read user buckets backward compat (#10683 Radoslaw Zarzynski)
- rgw: merge manifests properly with prefix override (#11622 Yehuda Sadeh)
- rgw: only scan for objects not in a namespace (#11984 Yehuda Sadeh)
- rgw: orphan detection tool (Yehuda Sadeh)
- rgw: pass in civetweb configurables (#10907 Yehuda Sadeh)
- rgw: rectify 202 Accepted in PUT response (#11148 Radoslaw Zarzynski)
- rgw: remove meta file after deleting bucket (#11149 Orit Wasserman)
- rgw: remove trailing :port from HTTP_HOST header (Sage Weil)
- rgw: return 412 on bad limit when listing buckets (#11613 Yehuda Sadeh)
- rgw: rework X-Trans-Id header to conform with Swift API (Radoslaw Rzarzynski)
- rgw: s3 encoding-type for get bucket (Jeff Weber)
- rgw: send ETag, Last-Modified for swift (#11087 Radoslaw Zarzynski)
- rgw: set content length on container GET, PUT, DELETE, HEAD (#10971, #11036 Radoslaw Zarzynski)
- rgw: set max buckets per user in ceph.conf (Vikhyat Umrao)
- rgw: shard work over multiple librados instances (Pavan Rallabhandi)
- rgw: support end marker on swift container GET (#10682 Radoslaw Zarzynski)
- rgw: support for Swift expiration API (Radoslaw Rzarzynski, Yehuda Sadeh)
- rgw: swift: allow setting attributes with COPY (#10662 Ahmad Faheem, Dmytro lurchenko)
- rgw: swift: do not override sent content type (#12363 Orit Wasserman)
- rgw: swift: enforce Content-Type in response (#12157 Radoslaw Zarzynski)
- rgw: swift: fix account listing (#11501 Radoslaw Zarzynski)
- rgw: swift: fix metadata handling on copy (#10645 Radoslaw Zarzynski)
- rgw: swift: send Last-Modified header (#10650 Radoslaw Zarzynski)
- rgw: swift: set Content-Length for account GET (#12158 Radoslav Zarzynski)
- rgw: swift: set content-length on keystone tokens (#11473 Herv Rousseau)
- rgw: update keystone cache with token info (#11125 Yehuda Sadeh)
- rgw: update to latest civetweb, enable config for IPv6 (#10965 Yehuda Sadeh)
- rgw: use attrs from source bucket on copy (#11639 Javier M. Mellid)
- rgw: use correct oid for gc chains (#11447 Yehuda Sadeh)
- rgw: user rm is idempotent (Orit Wasserman)
- rgw: use unique request id for civetweb (#10295 Orit Wasserman)
- rocksdb: add perf counters for get/put latency (Xinxin Shu)
- rocksdb, leveldb: fix compact_on_mount (Xiaoxi Chen)
- rocksdb: pass options as single string (Xiaoxi Chen)
- rocksdb: update to latest (Xiaoxi Chen)
- rpm: add suse firewall files (Tim Serong)
- rpm: always rebuild and install man pages for rpm (Owen Synge)
- rpm: loosen ceph-test dependencies (Ken Dreyer)
- rpm: many spec file fixes (Owen Synge, Ken Dreyer)
- rpm: misc fixes (Boris Ranto, Owen Synge, Ken Dreyer, Ira Cooper)
- rpm: misc systemd and SUSE fixes (Owen Synge, Nathan Cutler)
- selinux policy (Boris Ranto, Milan Broz)
- systemd: logrotate fixes (Tim Serong, Lars Marowsky-Bree, Nathan Cutler)
- systemd: many fixes (Sage Weil, Owen Synge, Boris Ranto, Dan van der Ster)
- systemd: run daemons as user ceph
- sysvinit compat: misc fixes (Owen Synge)
- test: misc fs test improvements (John Spray, Loic Dachary)
- test: python tests, linter cleanup (Alfredo Deza)
- tests: fixes for rbd xstests (Douglas Fuller)
- tests: fix tiering health checks (Loic Dachary)
- tests for low-level performance (Haomai Wang)
- tests: many ec non-regression improvements (Loic Dachary)
- tests: many many ec test improvements (Loic Dachary)
- upstart: throttle restarts (#11798 Sage Weil, Greg Farnum)

This is the third (and likely final) release candidate for Jewel.

NOTABLE CHANGES SINCE V10.1.1

- rgw: Jewel nfs fixes 3 ([pr#8460](#), Matt Benjamin)
- osd/PG: set epoch_created and parent_split_bits for child pg ([issue#15426](#), [pr#8552](#), Kefu Chai)
- rgw: signature mismatch with escaped characters in url query portion ([issue#15358](#), [pr#8445](#), Javier M. Mellid)
- authtool: update --help and manpage to match code. ([pr#8456](#), Robin H. Johnson)
- build: Respect TMPDIR for virtualenv. ([pr#8457](#), Robin H. Johnson)
- ceph-disk: do not always fail when re-using a partition ([pr#8508](#), You Ji)
- ceph-fuse: rotate log file ([pr#8485](#), Sage Weil)
- ceph-rest-api: fix fs/flag/set ([pr#8428](#), Sage Weil)
- check-generated.sh: can't source bash from sh ([pr#8521](#), Michal Jarzabek)
- common: buffer: put a guard for stat() syscall during read_file ([pr#7956](#), xie xingguo)
- common: fix time_t cast in decode ([issue#15330](#), [pr#8419](#), Adam C. Emerson)
- crush: fix typo ([pr#8518](#), Wei Jin)
- doc: rgw admin uses “region list” not “regions list” ([pr#8517](#), Kris Jurka)
- journal: fix final result for JournalTrimmer::C_RemoveSet ([pr#8516](#), runsi)
- journal: fix race condition between Future and journal shutdown ([issue#15364](#), [pr#8477](#), Jason Dillaman)
- librados: Revert “rados: Add new field flags for ceph_osd_op.copy_get.” ([pr#8486](#), Sage Weil)
- librbd: disallow unsafe rbd_op_threads values ([issue#15034](#), [pr#8459](#), Josh Durgin)
- mailmap update ([pr#8522](#), M Ranga Swami Reddy)
- mds: Add cmapv to ESessions default constructor initializer list ([pr#8403](#), John Coyle)
- mds: fix file_layout_t legacy encoding snafu ([pr#8455](#), Sage Weil)
- mds: Protect a number of unstable/experimental features behind durable flags ([pr#8383](#), Greg Farnum)
- mds: Extend the existing pool access checking to include specific RADOS namespaces. ([pr#8444](#), Yan, Zheng)
- os/ObjectStore: make device uuid probe output something friendly ([pr#8418](#), Sage Weil)
- os/bluestore: revamp BlueFS bdev management and add perfcounters ([issue#15376](#), [pr#8431](#), Sage Weil)
- os/filestore: force lfn attrs to be written atomically, restructure name length limits ([pr#8496](#), Samuel Just)
- osd/ReplicatedPG::_rollback_to: update the OMAP flag ([issue#14777](#), [pr#8495](#), Samuel Just)
- osd: clean up temp object if copy-from fails ([pr#8487](#), Sage Weil)
- osd: fix two scrub relevant issues ([pr#8462](#), xie xingguo)
- osd: fix/clean up full map request handling ([pr#8446](#), Sage Weil)
- osdc/Objecter: fix narrow race with tid assignment ([issue#14364](#), [pr#7981](#), Sage Weil)
- radosgw-admin: allow ([pr#8529](#), Orit Wasserman)
- rbd-mirror: workaround for intermingled lockdep singletons ([pr#8476](#), Jason Dillaman)
- rbd: journal reset should disable/re-enable journaling feature ([issue#15097](#), [pr#8490](#), Jason Dillaman)
- rgw-admin: remove unused iterator and fix error message ([pr#8507](#), Karol Mroz)
- rgw: aws4 subdomain calling bugfix ([issue#15369](#), [pr#8472](#), Javier M. Mellid)
- rgw: fix a typo in error message ([pr#8434](#), Abhishek Lekshmanan)
- rgw: fix problem deleting objects beginning with double underscores ([issue#15318](#), [pr#8488](#), Orit Wasserman)
- rgw: retry RGWRemoteMetaLog::read_log_info() while master is down ([pr#8453](#), Casey Bodley)
- rgw: the map ‘headers’ is assigned a wrong value ([pr#8481](#), weiqiaomiao)
- rgw_ldap: make ldap.h inclusion conditional ([pr#8500](#), Matt Benjamin)
- systemd: drop any systemd imposed process/thread limits ([pr#8450](#), James Page)
- test: fix ut test failure caused by lfn change ([issue#15464](#), [pr#8544](#), xie xingguo)
- test: fix valgrind memcheck issues for rbd-mirror test cases ([issue#15354](#), [pr#8493](#), Jason Dillaman)
- tests: add Ubuntu 16.04 xenial dockerfile ([pr#8519](#), Loic Dachary)

V10.1.1 JEWEL (RELEASE CANDIDATE)

This is the first release build that includes arm64/aarch64 packages for Ubuntu Xenial 16.04.

KNOW ISSUES WITH THIS RELEASE CANDIDATE

There are a few known issues to watch out for:

- Old CephFS clusters will mangle the layouts with this release; the fix was committed just after it was cut. Wait for the next RC or the release if you’re upgrading a cluster that has a CephFS data pool as pool 0.
- The upstart ceph-mds-all.conf file is missing.

NOTABLE CHANGES SINCE V10.1.0

- Adding documentation on how to use new dynamic throttle scheme ([pr#8069](#), Somnath Roy)
- Be more careful about directory fragmentation and scrubbing ([issue#15167](#), [pr#8180](#), Yan, Zheng)
- CMake: For CMake version <= 2.8.11, use LINK_PRIVATE ([pr#8422](#), Haomai Wang)
- Makefile-env.am: set a default for CEPH_BUILD_VIRTUALENV (part 2) ([pr#8320](#), Loic Dachary)
- Minor fixes around data scan in some scenarios ([pr#8115](#), Yan, Zheng)
- PG: pg down state blocked by osd.x, lost osd.x cannot solve peering stuck ([issue#13531](#), [pr#6317](#), Xiaowei Chen)
- osd/PG: indicate in pg query output whether ignore_history_les would help ([pr#8156](#), Sage Weil)
- Striper: reduce assemble_result log level ([pr#8426](#), Jason Dillaman)
- Test exit values on test.sh, fix tier.cc ([issue#15165](#), [pr#8266](#), Samuel Just)
- ceph.in: correct dev python path for automake builds ([pr#8360](#), Josh Durgin)
- osd/OSD: fix build_past_intervals_parallel ([pr#8215](#), David Zafman)
- rgw: sync fixes 4 ([pr#8190](#), Yehuda Sadeh)
- [rgw] Check return code in RGWFileHandle::write ([pr#7875](#), Brad Hubbard)
- build: fix compiling warnings ([pr#8366](#), Dongsheng Yang)
- ceph-detect-init/run-tox.sh: FreeBSD: No init detect ([pr#8373](#), Willem Jan Withagen)
- ceph.in: fix python libpath for automake as well ([pr#8362](#), Josh Durgin)
- ceph: bash auto complete for CLI based on mon command descriptions ([pr#7693](#), Adam Kupczyk)
- cls_journal: fix -EEXIST checking ([pr#8413](#), runsisi)
- cls_rbd: fix -EEXIST checking in cls::rbd::image_set ([pr#8371](#), runsisi)
- cls_rbd: mirror_image_list should return global image id ([pr#8297](#), Jason Dillaman)
- cls_rbd: pass WILLNEED fadvise flags during object map update ([issue#15332](#), [pr#8380](#), Jason Dillaman)
- cls_rbd: read_peers: update last_read on next cls_cxx_map_get_vals ([pr#8374](#), Mykola Golub)
- cmake: Build cython modules and change paths to bin/, lib/ ([pr#8351](#), John Spray, Ali Maredia)
- cmake: add FindOpenSSL.cmake ([pr#8106](#), Marcus Watts, Matt Benjamin)
- cmake: add StandardPolicy.cc to librbd ([pr#8368](#), Kefu Chai)
- cmake: add missing librbd/MirrorWatcher.cc and librd/ObjectWatcher.cc ([pr#8399](#), Orit Wasserman)
- cmake: fix mrun to handle cmake build structure ([pr#8237](#), Orit Wasserman)
- cmake: fix the build of test_rados_api_list ([pr#8438](#), Kefu Chai)
- common: fix race during otracker switches between enabled/disabled mode ([pr#8330](#), xie xingguo)
- config_opts: disable filestore throttle soft backoff by default ([pr#8265](#), Samuel Just)
- configure: Add -D_LARGEFILE64_SOURCE to Linux build. ([pr#8402](#), Ira Cooper)
- crush: fix error log ([pr#8430](#), Wei Jin)
- crushtool: Don't crash when called on a file that isn't a crushmap ([issue#8286](#), [pr#8038](#), Brad Hubbard)
- debian/rules: put init-ceph in /etc/init.d/ceph, not ceph-base ([issue#15329](#), [pr#8406](#), Dan Mick)
- doc/dev: add "Deploy a cluster for manual testing" section ([issue#15218](#), [pr#8228](#), Nathan Cutler)
- doc/rados/operations/crush: fix the formatting ([pr#8306](#), Kefu Chai)
- doc/release-notes: fix indents ([pr#8345](#), Kefu Chai)
- doc: Fixes headline different font size and type ([pr#8328](#), scienceluo)
- doc: Remove Ceph Monitors do lots of fsync() ([issue#15288](#), [pr#8327](#), Vikhyat Umrao)
- doc: Updated CloudStack RBD documentation ([pr#8308](#), Wido den Hollander)
- doc: amend Fixes instructions in SubmittingPatches ([pr#8312](#), Nathan Cutler)
- doc: draft notes for jewel ([pr#8211](#), Loic Dachary, Sage Weil)
- doc: fix typo, duplicated content etc. for Jewel release notes ([pr#8342](#), xie xingguo)
- doc: fix wrong type of hyphen ([pr#8252](#), xie xingguo)
- doc: rgw_region_root_pool option should be in [global] ([issue#15244](#), [pr#8271](#), Vikhyat Umrao)
- doc: very basic doc on mstart ([pr#8207](#), Abhishek Lekshmanan)
- global/global_init: expand metavariables in setuser_match_path ([issue#15365](#), [pr#8433](#), Sage Weil)
- global/signal_handler: print thread name in signal handler ([pr#8177](#), Jianpeng Ma)
- libcephfs: fix python tests and fix.getcwd on missing dir ([pr#7901](#), John Spray)
- librbd: avoid throwing error if mirroring is unsupported ([pr#8417](#), Jason Dillaman)
- librbd: disable image mirroring when image is removed ([issue#15265](#), [pr#8375](#), Ricardo Dias)
- librbd: send notifications for mirroring status updates ([pr#8355](#), Jason Dillaman)
- mailmap updates ([pr#8256](#), Loic Dachary)
- makefile: fix rbdmap manpage ([pr#8310](#), Kefu Chai)
- mds: allow client to request caps when opening file ([issue#14360](#), [pr#7952](#), Yan, Zheng)
- messages/MOSDOp: clear reqid inc for v6 encoding ([issue#15230](#), [pr#8299](#), Sage Weil)
- mon/MonClient: fix shutdown race ([issue#13992](#), [pr#8335](#), Sage Weil)
- mon: do not send useless pg_create messages for split pgs ([pr#8247](#), Sage Weil)
- mon: mark_down_pgs in lockstep with pg_map's osdmap epoch ([pr#8208](#), Sage Weil)
- mon: remove remove_legacy_versions() ([pr#8324](#), Kefu Chai)
- mon: remove unnecessary comment for update_from_paxos ([pr#8400](#), Qinghua Jin)
- mon: remove unused variable ([issue#15292](#), [pr#8337](#), Javier M. Mellid)

- mon: show the pool quota info on ceph df detail command ([issue#14216](#), [pr#7094](#), song baisen)
- monclient: avoid key renew storm on clock skew ([issue#12065](#), [pr#8258](#), Alexey Sheplyakov)
- mrun: update path to cmake binaries ([pr#8447](#), Casey Bodley)
- msg/async: avoid log spam on throttle ([issue#15031](#), [pr#8263](#), Kefu Chai)
- msg/async: remove experiment feature ([pr#7820](#), Haomai Wang)
- os/ObjectStore: add noexcept to ensure move ctor is used ([pr#8421](#), Kefu Chai)
- os/ObjectStore: fix_update_op for split dest_cid ([pr#8364](#), Sage Weil)
- os/ObjectStore: try_move_rename in transaction append and add coverage to store_test ([issue#15205](#), [pr#8359](#), Samuel Just)
- os/bluestore: a few fixes ([pr#8193](#), Sage Weil)
- os/bluestore: ceph-bluefs-tool fixes ([issue#15261](#), [pr#8292](#), Venky Shankar)
- osd/ClassHandler: only dlclose() the classes not missing ([pr#8354](#), Kefu Chai)
- osd/OSD.cc: finish full_map_request every MOSDMap message. ([issue#15130](#), [pr#8147](#), Xiaoxi Chen)
- osd: add ‘proxy’ cache mode ([issue#12814](#), [pr#8210](#), Sage Weil)
- osd: add the support of per pool scrub priority ([pr#7062](#), Zhiqiang Wang)
- osd: bail out of _committed_osd_maps if we are shutting down ([pr#8267](#), Samuel Just)
- osd: duplicated clear for peer_missing ([pr#8315](#), Ning Yao)
- osd: fix bugs for omap ops ([pr#8230](#), Jianpeng Ma)
- osd: fix dirtying info without correctly setting dirty_info field ([pr#8275](#), xie xingguo)
- osd: fix dump_ops_in_flight races ([issue#8885](#), [pr#8044](#), David Zafman)
- osd: fix epoch check in handle_pg_create ([pr#8382](#), Samuel Just)
- osd: fix failure report handling during ms_handle_connect() ([pr#8348](#), xie xingguo)
- osd: fix log info ([pr#8273](#), Wei Jin)
- osd: fix reference count, rare race condition etc. ([pr#8254](#), xie xingguo)
- osd: fix tick relevant issues ([pr#8369](#), xie xingguo)
- osd: more fixes for incorrectly dirtying info; resend reply for duplicated scrub-reserve req ([pr#8291](#), xie xingguo)
- osdc/Objecter: dout log after assign tid ([pr#8202](#), Xinze Chi)
- osdc/Objecter: use full pgid hash in PGNLS ops ([pr#8378](#), Sage Weil)
- osdmap: rm nonused variable ([pr#8423](#), Wei Jin)
- pybind/Makefile.am: Prevent race creating CYTHON_BUILD_DIR ([issue#15276](#), [pr#8356](#), Dan Mick)
- pybind/rados: python3 fix ([pr#8331](#), Mehdi Abaakouk)
- pybind: add flock to libcephfs python bindings ([pr#7902](#), John Spray)
- qa: update rest test cephfs calls ([issue#15309](#), [pr#8372](#), John Spray)
- qa: update rest test cephfs calls (part 2) ([issue#15309](#), [pr#8393](#), John Spray)
- radosgw-admin: ‘period commit’ supplies user-readable error messages ([pr#8264](#), Casey Bodley)
- radosgw-admin: fix for ‘realm pull’ ([pr#8404](#), Casey Bodley)
- rbd-mirror: asok commands to get status and flush on Mirror and Replayer level ([pr#8235](#), Mykola Golub)
- rbd-mirror: enabling/disabling pool mirroring should update the mirroring directory ([issue#15217](#), [pr#8261](#), Ricardo Dias)
- rbd-mirror: fix missing increment for iterators ([pr#8352](#), runsisi)
- rbd-mirror: initial failover / fallback support ([pr#8287](#), Jason Dillaman)
- rbd-mirror: prevent enabling/disabling an image’s mirroring when not in image mode ([issue#15267](#), [pr#8332](#), Ricardo Dias)
- rbd-mirror: switch fsid over to mirror uuid ([issue#15238](#), [pr#8280](#), Ricardo Dias)
- rbd: allow librados to prune the command-line for config overrides ([issue#15250](#), [pr#8282](#), Jason Dillaman)
- rbdmap: add manpage ([issue#15212](#), [pr#8224](#), Nathan Cutler)
- releases: what is merged where and when ? ([pr#8358](#), Loic Dachary)
- rgw/rgw_admin:fix bug about list and stats command ([pr#8200](#), Qiankun Zheng)
- rgw: Do not send a Content-Type on a ‘304 Not Modified’ response ([issue#15119](#), [pr#8253](#), Wido den Hollander)
- rgw: Multipart ListPartsResult ETag quotes ([issue#15334](#), [pr#8387](#), Robin H. Johnson)
- rgw: S3: set EncodingType in ListBucketResult ([pr#7712](#), Victor Makarov)
- rgw: accept data only at the first time in response to a request ([pr#8084](#), sunspot)
- rgw: add a few more help options in admin interface ([pr#8410](#), Abhishek Lekshmanan)
- rgw: add zone delete to rgw-admin help ([pr#8184](#), Abhishek Lekshmanan)
- rgw: convert plain object to versioned (with null version) when removing ([issue#15243](#), [pr#8268](#), Yehuda Sadeh)
- rgw: fix compiling error ([pr#8394](#), xie xingguo)
- rgw: fix lockdep false positive ([pr#8284](#), Yehuda Sadeh)
- rgw:Use count fn in RGWUserBuckets for quota check ([pr#8294](#), Abhishek Lekshmanan)
- rgw_admin: remove unused parent_period arg ([pr#8411](#), Abhishek Lekshmanan)
- rgw_file: set owner uid, gid, and Unix mode on new objects ([pr#8321](#), Matt Benjamin)
- rpm: prefer UID/GID 167 when creating ceph user/group ([issue#15246](#), [pr#8277](#), Nathan Cutler)
- script: subscription-manager support ([issue#14972](#), [pr#7907](#), Loic Dachary)
- set 128MB tcmalloc cache size by bytes ([pr#8427](#), Star Guo)
- systemd: set up environment in rbdmap unit file ([issue#14984](#), [pr#8222](#), Nathan Cutler)
- test/system/*: use dynamically generated pool name ([issue#15240](#), [pr#8318](#), Kefu Chai)
- test/system/rados_list_parallel: print oid if rados_write fails ([issue#15240](#), [pr#8309](#), Kefu Chai)

- test/test-erasure-code.sh: disable pg_temp priming ([issue#15211](#), [pr#8260](#), Sage Weil)
- test/test_pool_create.sh: fix port ([pr#8361](#), Sage Weil)
- test: Fix test to run with btrfs which has snap_### dirs ([issue#15347](#), [pr#8420](#), David Zafman)
- test: TestMirroringWatcher test cases were not closing images ([pr#8435](#), Jason Dillaman)
- test: rbd-mirror: script improvements for manual testing ([pr#8325](#), Mykola Golub)
- tests: Fixing broken test/cephtool-test-mon.sh test ([pr#8429](#), Erwan Velu)
- tests: Improving ‘make check’ execution time ([pr#8131](#), Erwan Velu)
- unittest_erasure_code_plugin: fix deadlock (Alpine) ([pr#8314](#), John Coyle)
- vstart: fix up cmake paths when VSTART_DEST is given ([pr#8363](#), Casey Bodley)
- vstart: make -k with optional mon_num. ([pr#8251](#), Jianpeng Ma)
- xio: add prefix to xio msgr logs ([pr#8148](#), Roi Dayan)
- xio: fix compilation against latest accelio ([pr#8022](#), Roi Dayan)
- xio: xio_init needs to be called before any other xio function ([pr#8227](#), Roi Dayan)
- ceph.spec.in: disable ltng and babeltrace explicitly ([issue#14844](#), [pr#7857](#), Kefu Chai)

V10.1.0 JEWEL (RELEASE CANDIDATE)

There are a few known issues with this release candidate; see below.

KNOWN ISSUES WITH V10.1.0

- While running a mixed version cluster of jewel and infernalis or hammer monitors, any MDSMap updates will cause the pre-jewel monitors to crash. Workaround is to simply upgrade all monitors. There is a fix but it is still being tested.
- Some of the rbd-mirror functionality for switching between active and replica images is not yet merged.

NOTABLE CHANGES SINCE V10.0.4

- ceph.spec.in: do not install Ceph RA on systemd platforms ([issue#14828](#), [pr#7894](#), Nathan Cutler)
- mdsa: A few more snapshot fixes, mostly around snapshotted inode/dentry tracking ([pr#7798](#), Yan, Zheng)
- AUTHORS: update email ([pr#7854](#), Yehuda Sadeh)
- ceph-disk: Add -setuser and -setgroup options for ceph-disk ([pr#7351](#), Mike Shuey)
- build: Adding build requires ([pr#7742](#), Erwan Velu)
- msg/async: AsyncMessenger: fix several bugs ([pr#7831](#), Haomai Wang)
- msg/async: AsyncMessenger: fix valgrind leak ([pr#7725](#), Haomai Wang)
- doc: Clarify usage on starting single osd/mds/mon. ([pr#7641](#), Patrick Donnelly)
- common: Deprecate or free up a bunch of feature bits ([pr#8214](#), Samuel Just)
- msg/async: Event: fix clock skew problem ([pr#7949](#), Wei Jin)
- osd: FileStore: Added O_DSYNC write scheme ([pr#7752](#), Somnath Roy)
- osd: FileStore: fix initialization order for m_disable_wbthrottle ([pr#8067](#), Samuel Just)
- build: Fixing BTRFS issue at ‘make check’ ([pr#7805](#), Erwan Velu)
- build: FreeBSD related fixes ([pr#7170](#), Mykola Golub)
- ceph-disk: Improving ‘make check’ for ceph-disk ([pr#7762](#), Erwan Velu)
- rgw: increase verbosity level on RGWObjManifest line ([pr#7285](#), magicrobotmonkey)
- build: workaround an automake bug for “make check” ([issue#14723](#), [pr#7626](#), Kefu Chai)
- ceph-fuse,libcephfs: Fix client handling of “lost” open directories on shutdown ([issue#14996](#), [pr#7994](#), Yan, Zheng)
- mds: Multi-filesystem support ([issue#14952](#), [pr#6953](#), John Spray, Sage Weil)
- os/bluestore/NVMEDevice: refactor probe/attach codes and support zero command ([pr#7647](#), Haomai Wang)
- librados: detect laggy ops with objecter_timeout, not osd_timeout ([pr#7629](#), Greg Farnum)
- ceph.spec.in: fix openldap and openssl build dependencies for SUSE ([issue#15138](#), [pr#8120](#), Nathan Cutler)
- osd: repop and lost-unfound overhaul ([pr#7765](#), Samuel Just)
- librbd: Revert “librbd: use task finisher per CephContext” ([issue#14780](#), [pr#7667](#), Josh Durgin)
- rgw: Fix subuser harder with tenants ([pr#7618](#), Pete Zaitcev)
- rgw: ldap fixes ([pr#8168](#), Matt Benjamin)
- rgw: check the return value when call fe->run() ([issue#14585](#), [pr#7457](#), wei qiaomiao)
- rgw: Revert “rgw ldap” ([pr#8075](#), Yehuda Sadeh)
- librados: Striper: Fix incorrect push_front -> append_zero change ([pr#7578](#), Haomai Wang)
- rgw: build-related fixes ([pr#8076](#), Yehuda Sadeh, Matt Benjamin)
- mirrors: Updated scripts and documentation for mirrors ([pr#7847](#), Wido den Hollander)
- misc: use make_shared while creating shared_ptr ([pr#7769](#), Somnath Roy)
- cmake ([pr#7849](#), Ali Maredia)
- mds: filelock deadlock ([pr#7713](#), Yan, Zheng)
- mds: fix fsmap decode ([pr#8063](#), Greg Farnum)

- rgw: fix mdlog ([pr#8183](#), Orit Wasserman)
- rgw: highres time stamps ([pr#8108](#), Yehuda Sadeh, Adam C. Emerson, Matt Benjamin)
- rgw: swift versioning disabled ([pr#8066](#), Yehuda Sadeh, Radoslaw Zarzynski)
- rgw: sync fixes 3 ([pr#8170](#), Yehuda Sadeh)
- msg/xio: fixes ([pr#7603](#), Roi Dayan)
- ceph-fuse,libcephfs: fix free fds being exhausted eventually because freed fds are never put back ([issue#14798](#), [pr#7685](#), Zhi Zhang)
- rgw: RGWLib::env is not used so remove it ([pr#7874](#), Brad Hubbard)
- build: a few armhf (32-bit build) fixes ([pr#7999](#), Eric Lee, Sage Weil)
- osd: add scrub persist/query API ([issue#13505](#), [pr#6898](#), Kefu Chai, Samuel Just)
- rgw: adds the radosgw-admin sync status command that gives a human readable status of the sync process at a specific zone ([pr#8030](#), Yehuda Sadeh)
- scripts: adjust mstart and mstop script to run with cmake build ([pr#6920](#), Orit Wasserman)
- buffer: add symmetry operator==() and operator!=() ([pr#7974](#), Kefu Chai)
- buffer: hide iterator_impl symbols ([issue#14788](#), [pr#7688](#), Kefu Chai)
- buffer: increment history alloc as well in raw_combined ([issue#14955](#), [pr#7910](#), Samuel Just)
- buffer: raw_combined allocations buffer and ref count together ([pr#7612](#), Sage Weil)
- ceph-detect-init: add debian/jessie test ([pr#8074](#), Kefu Chai)
- ceph-detect-init: add missing test case ([pr#8105](#), Nathan Cutler)
- ceph-detect-init: fix py3 test ([pr#7243](#), Kefu Chai)
- ceph-detect-init: return correct value on recent SUSE distros ([issue#14770](#), [pr#7909](#), Nathan Cutler)
- ceph-disk: deactivate / destroy PATH arg are optional ([pr#7756](#), Loic Dachary)
- ceph-disk: fix prepare -help ([pr#7758](#), Loic Dachary)
- ceph-disk: flake8 fixes ([pr#7646](#), Loic Dachary)
- ceph-disk: key management support ([issue#14669](#), [pr#7552](#), Loic Dachary)
- ceph-disk: make some arguments as required if necessary ([pr#7687](#), Dongsheng Yang)
- ceph-disk: s/dmcrpyt/dmcrypt/ ([issue#14838](#), [pr#7744](#), Loic Dachary, Frode Sandholbraaten)
- ceph-fuse: Fix potential filehandle ref leak at umount ([issue#14800](#), [pr#7686](#), Zhi Zhang)
- ceph.in: Minor python3 specific changes ([pr#7947](#), Sarthak Munshi)
- ceph_daemon.py: Resolved ImportError to work with python3 ([pr#7937](#), Sarthak Munshi)
- ceph_detect_init/_init__.py: remove shebang ([pr#7731](#), Nathan Cutler)
- ceph_test_msgr: reduce test size to fix memory size ([pr#8127](#), Haomai Wang)
- ceph_test_rados_misc: shorten mount timeout ([pr#8209](#), Sage Weil)
- cleanup ([pr#8058](#), Yehuda Sadeh, Orit Wasserman)
- client: flush kernel pagecache before creating snapshot ([issue#10436](#), [pr#7495](#), Yan, Zheng)
- client: removed unused Mutex from MetaRequest ([pr#7655](#), Greg Farnum)
- cls/rgw: fix FTBFS ([pr#8142](#), Kefu Chai)
- cls/rgw: fix use of timespan ([issue#15181](#), [pr#8212](#), Yehuda Sadeh)
- cls_hello: Fix grammatical error in description comment ([pr#7951](#), Brad Hubbard)
- cls_rbd: fix the test for ceph-dencoder ([pr#7793](#), Kefu Chai)
- cls_rbd: mirroring directory ([issue#14419](#), [pr#7620](#), Josh Durgin)
- cls_rbd: protect against excessively large object maps ([issue#15121](#), [pr#8099](#), Jason Dillaman)
- cmake: Remove duplicate find_package libcurl line. ([pr#7972](#), Brad Hubbard)
- cmake: add ErasureCode.cc to jerasure plugins ([pr#7808](#), Casey Bodley)
- cmake: add common/fs_types.cc to libcommon ([pr#7898](#), Orit Wasserman)
- cmake: add missing librbd image_watcher sources ([issue#14823](#), [pr#7717](#), Casey Bodley)
- cmake: avoid false-positive LDAP header detect ([pr#8100](#), Matt Benjamin)
- cmake: fix paths to various EC source files ([pr#7748](#), Ali Maredia, Matt Benjamin)
- cmake: fix the build of tests ([pr#7523](#), Kefu Chai)
- common/TrackedOp: fix inaccurate counting for slow requests ([issue#14804](#), [pr#7690](#), xie xingguo)
- common/bit_vector: use hard-coded value for block size ([issue#14747](#), [pr#7610](#), Jason Dillaman)
- common/obj_bencher.cc: bump the precision of bandwidth field ([pr#8021](#), Piotr Dałek)
- common/obj_bencher.cc: faster object name generation ([pr#7863](#), Piotr Dałek)
- common/obj_bencher.cc: make verify error fatal ([issue#14971](#), [pr#7897](#), Piotr Dałek)
- common/page.cc: _page_mask has too many bits ([pr#7588](#), Dan Mick)
- common/strtol.cc: fix the coverity warnings ([pr#7967](#), Kefu Chai)
- common: Do not use non-portable constants in mutex_debug ([pr#7766](#), Adam C. Emerson)
- common: SubProcess: Avoid buffer corruption when calling err() ([issue#15011](#), [pr#8054](#), Erwan Velu)
- common: default cluster name to config file prefix ([pr#7364](#), Javen Wu)
- common: set thread name from correct thread ([pr#7845](#), Igor Podoski)
- common: various fixes from SCA runs ([pr#7680](#), Danny Al-Gaaf)
- config: fix osd_crush_initial_weight ([pr#7975](#), You Ji)
- config: increase default async op threads ([pr#7802](#), Piotr Dałek)
- configure.ac: boost_iostreams is required, not optional ([pr#7816](#), Hector Martin)
- configure.ac: update help strings for cython ([pr#7856](#), Josh Durgin)

- crush/CrushTester: workaround a bug in boost::icl ([pr#7560](#), Kefu Chai)
- crush: fix cli tests for new crush tunables ([pr#8107](#), Sage Weil)
- crush: update tunable docs. change default profile to jewel ([pr#7964](#), Sage Weil)
- debian/changelog: Remove stray 'v' in version ([pr#7936](#), Dan Mick)
- debian/changelog: Remove stray 'v' in version ([pr#7938](#), Dan Mick)
- debian: include cpio in build-requires ([pr#7533](#), Rémi BUISSON)
- debian: package librgw_file* tests ([pr#7930](#), Ken Dreyer)
- doc/architecture.rst: remove redundant word "across" ([pr#8179](#), Zhao Junwang)
- doc/dev: add section on interrupting a running suite ([pr#8116](#), Nathan Cutler)
- doc/dev: continue writing Testing in the cloud chapter ([pr#7960](#), Nathan Cutler)
- doc/dev: integrate testing into the narrative ([pr#7946](#), Nathan Cutler)
- doc/dev: various refinements ([pr#7954](#), Nathan Cutler)
- doc/rados/api/librados-intro.rst: fix typo ([pr#7879](#), xie xingguo)
- doc: add ceph-detect-init(8) source to dist tarball ([pr#7933](#), Ken Dreyer)
- doc: add cinder backend section to rbd-openstack.rst ([pr#7923](#), RustShen)
- doc: detailed description of bugfixing workflow ([pr#7941](#), Nathan Cutler)
- doc: fix 0.94.4 and 0.94.5 ordering ([pr#7763](#), Loic Dachary)
- doc: fix typo, indentation etc. ([pr#7829](#), xie xingguo)
- doc: initial draft of RBD mirroring admin documentation ([issue#15041](#), [pr#8169](#), Jason Dillaman)
- doc: osd-config Add Configuration Options for op queue. ([pr#7837](#), Robert LeBlanc)
- doc: rgw explain keystone's verify ssl switch ([pr#7862](#), Abhishek Lekshmanan)
- doc: small fixes ([pr#7813](#), xiexingguo)
- doc: standardize @param (not @parma, @parmam, @params) ([pr#7714](#), Nathan Cutler)
- fix FTBFS introduced by d0af316 ([pr#7792](#), Kefu Chai)
- ghobject_t: use # instead of ! as a separator ([pr#8055](#), Sage Weil)
- include/encoding: do not try to be clever with list encoding ([pr#7913](#), Sage Weil)
- init-ceph.in: allow case-insensitive true in `osd crush update on start`` ([pr#7943](#), Eric Cook)
- init-ceph.in: skip ceph-disk if it is not present ([issue#10587](#), [pr#7286](#), Ken Dreyer)
- journal: async methods to (un)register and update client ([pr#7832](#), Mykola Golub)
- journal: improve commit position tracking ([pr#7776](#), Jason Dillaman)
- journal: prevent race injecting new records into overflowed object ([issue#15202](#), [pr#8220](#), Jason Dillaman)
- librados: cancel aio notification linger op upon completion ([pr#8102](#), Jason Dillaman)
- librados: check connection state in rados_monitor_log ([issue#14499](#), [pr#7350](#), David Disseldorp)
- librados: do not clear handle for aio_watch() ([pr#7771](#), xie xingguo)
- librados: fix test failure with new aio watch/unwatch API ([pr#7824](#), Jason Dillaman)
- librados: implement async watch/unwatch ([pr#7649](#), Haomai Wang)
- librados: mix lock cycle (un)registering asok commands ([pr#7581](#), John Spray)
- librados: race condition on aio_notify completion handling ([pr#7864](#), Jason Dillaman)
- librados: stat2 with higher time precision ([pr#7915](#), Yehuda Sadeh, Matt Benjamin)
- librbd: allocate new journal tag after acquiring exclusive lock ([pr#7884](#), Jason Dillaman)
- librbd: block read requests until journal replayed ([pr#7627](#), Jason Dillaman)
- librbd: correct issues discovered via valgrind memcheck ([pr#8132](#), Jason Dillaman)
- librbd: differentiate journal replay flush vs shut down ([pr#7698](#), Jason Dillaman)
- librbd: enable/disable image mirroring automatically for pool mode ([issue#15143](#), [pr#8204](#), Ricardo Dias)
- librbd: fix state machine race conditions during shut down ([pr#7761](#), Jason Dillaman)
- librbd: handle unregistering the image watcher when disconnected ([pr#8094](#), Jason Dillaman)
- librbd: integrate journal replay with fsx testing ([pr#7583](#), Jason Dillaman)
- librbd: journal replay needs to support re-executing maintenance ops ([issue#14822](#), [pr#7785](#), Jason Dillaman)
- librbd: reduce mem copies to user-buffer during read ([pr#7548](#), Jianpeng Ma)
- librbd: refresh image if required before replaying journal ops ([issue#14908](#), [pr#7978](#), Jason Dillaman)
- librbd: remove last synchronous librados calls from open/close state machine ([pr#7839](#), Jason Dillaman)
- librbd: replaying a journal op post-refresh requires locking ([pr#8028](#), Jason Dillaman)
- librbd: retrieve image name when opening by id ([pr#7736](#), Mykola Golub)
- librbd: several race conditions discovered under single CPU environment ([pr#7653](#), Jason Dillaman)
- librbd: truncate does not need to mark the object as existing in the object map ([issue#14789](#), [pr#7772](#), xinxin shu)
- librbd: update of mirror pool mode and mirror peer handling ([pr#7718](#), Jason Dillaman)
- librbd: use async librados notifications ([pr#7668](#), Jason Dillaman)
- log: do not repeat errors to stderr ([issue#14616](#), [pr#7983](#), Sage Weil)
- log: fix stack overflow when flushing large log lines ([issue#14707](#), [pr#7599](#), Igor Fedotov)
- log: segv in a portable way ([issue#14856](#), [pr#7790](#), Kefu Chai)
- log: use delete[] ([pr#7904](#), Sage Weil)
- mailmap for 10.0.4 ([pr#7932](#), Abhishek Lekshmanan)
- mailmap updates ([pr#7528](#), Yann Dupont)
- man/8/ceph-disk: fix formatting issue ([pr#8012](#), Sage Weil)
- man/8/ceph-disk: fix formatting issue ([pr#8003](#), Sage Weil)

- mds, client: add namespace to file_layout_t (previously ceph_file_layout) ([pr#7098](#), Yan, Zheng, Sage Weil)
- mds: don't double-shutdown the timer when suiciding ([issue#14697](#), [pr#7616](#), Greg Farnum)
- mds: fix FSSMap upgrade with daemons in the map ([pr#8073](#), John Spray, Greg Farnum)
- mds: fix inode_t::compare() ([issue#15038](#), [pr#8014](#), Yan, Zheng)
- mds: fix stray purging in 'stripe_count > 1' case ([issue#15050](#), [pr#8040](#), Yan, Zheng)
- mds: function parameter 'df' should be passed by reference ([pr#7490](#), Na Xie)
- mirrors: Change contact e-mail address for se.ceph.com ([pr#8007](#), Wido den Hollander)
- mon/PGMonitor: reliably mark PGs state ([pr#8089](#), Sage Weil)
- mon/monitor: some clean up ([pr#7520](#), huanwen ren)
- mon/pgmonitor: use appropriate forced conversions in get_rule_avail ([pr#7705](#), huanwen ren)
- mon: cleanup set-quota error msg ([pr#7371](#), Abhishek Lekshmanan)
- mon: consider pool size when creating pool ([issue#14509](#), [pr#7359](#), songbaisen)
- mon: enable 'mon osd prime pg temp' by default ([pr#7838](#), Robert LeBlanc)
- mon: fix calculation of %USED ([pr#7881](#), Adam Kupczyk)
- mon: fix keyring permissions ([issue#14950](#), [pr#7880](#), Owen Synge)
- mon: initialize last_* timestamps on new pgs to creation time ([issue#14952](#), [pr#7980](#), Sage Weil)
- mon: make clock skew checks sane ([issue#14175](#), [pr#7141](#), Joao Eduardo Luis)
- mon: osd [test-]reweight-by-{pg,utilization} command updates ([pr#7890](#), Dan van der Ster, Sage Weil)
- mon: remove 'mds setmap' ([issue#15136](#), [pr#8121](#), Sage Weil)
- mon: standardize Ceph removal commands ([pr#7939](#), Dongsheng Yang)
- mon: unconfuse object count skew message ([pr#7882](#), Piotr Dałek)
- mon: unregister command on shutdown ([pr#7504](#), huanwen ren)
- mount.fuse.ceph: better parsing of arguments passed to mount.fuse.ceph by mount command ([issue#14735](#), [pr#7607](#), Florent Bautista)
- msg/async: _try_send trim already sent for outcoming_bl more efficient ([pr#7970](#), Yan Jun)
- msg/async: don't calculate msg header crc when not needed ([pr#7815](#), Piotr Dałek)
- msg/async: smarter MSG_MORE ([pr#7625](#), Piotr Dałek)
- msg: add thread safety for "random" Messenger + fix wrong usage of random functions ([pr#7650](#), Avner BenHanoch)
- msg: async: fix perf counter description and simplify _send_keepalive_or_ack ([pr#8046](#), xie xingguo)
- msg: async: small cleanups ([pr#7871](#), xie xingguo)
- msg: async: start over after failing to bind a port in specified range ([issue#14928](#), [issue#13002](#), [pr#7852](#), xie xingguo)
- msg: remove duplicated code - local_delivery will now call 'enqueue' ([pr#7948](#), Avner BenHanoch)
- msg: significantly reduce minimal memory usage of connections ([pr#7567](#), Piotr Dałek)
- mstart: start rgw on different ports as well ([pr#8167](#), Abhishek Lekshmanan)
- nfs for rgw (Matt Benjamin, Orit Wasserman) ([pr#7634](#), Yehuda Sadeh, Matt Benjamin)
- os/ObjectStore: implement more efficient get_encoded_bytes() ([pr#7775](#), Piotr Dałek)
- os/bluestore/BlueFS: Before reap iocb, it should wait io complete ([pr#8178](#), Jianpeng Ma)
- os/bluestore/BlueStore: Don't leak trim overlay data before write. ([pr#7895](#), Jianpeng Ma)
- os/bluestore/KernelDevice: force block size ([pr#8006](#), Sage Weil)
- os/bluestore/NVMEDevice: make IO thread using dpdk launch ([pr#8160](#), Haomai Wang)
- os/bluestore: clone overlay data ([pr#7860](#), Jianpeng Ma)
- os/bluestore: fix a typo in SPDK path parsing ([pr#7601](#), Jianjian Huo)
- os/bluestore: make bluestore_sync_transaction = true can work. ([pr#7674](#), Jianpeng Ma)
- os/bluestore: small fixes in bluestore StupidAllocator ([pr#8101](#), Jianjian Huo)
- os/filestore/FileJournal: set block size via config option ([pr#7628](#), Sage Weil)
- os/filestore: fix punch hole usage in _zero ([pr#8050](#), Sage Weil)
- os/filestore: fix result handling logic of destroy_collection ([pr#7721](#), xie xingguo)
- os/filestore: require offset == length == 0 for full object read; add test ([pr#7957](#), Jianpeng Ma)
- osd/OSDMap: fix typo in summarize_mapping_stats ([pr#8088](#), Sage Weil)
- osd/PGLog: fix warning ([pr#8057](#), Sage Weil)
- osd/ReplicatedPG: be more careful about calling publish_stats_to_osd() ([issue#14962](#), [pr#8039](#), Greg Farnum)
- osd/ReplicatedPG: clear watches on change after applying repops ([issue#15151](#), [pr#8163](#), Sage Weil)
- osd/ScrubStore: remove unused function ([pr#8045](#), Kefu Chai)
- osd: BlueStore/NVMEDevice: fix compiling and fd leak ([pr#7496](#), xie xingguo)
- osd: FileStore: use pwritev instead of lseek+writev ([pr#7349](#), Haomai Wang, Tao Chang)
- osd: OSDMap: reset osd_primary_affinity shared_ptr when deepish_copy_from ([issue#14686](#), [pr#7553](#), Xinze Chi)
- osd: Replace sprintf with faster implementation in eversion_t::get_key_name ([pr#7121](#), Evgeniy Firsov)
- osd: WeightedPriorityQueue: move to intrusive containers ([pr#7654](#), Robert LeBlanc)
- osd: a fix for HeartbeatDispatcher and cleanups ([pr#7550](#), Kefu Chai)
- osd: add missing newline to usage message ([pr#7613](#), Willem Jan Withagen)
- osd: avoid FORCE updating digest been overwritten by MAYBE when comparing scrub map ([pr#7051](#), Zhiqiang Wang)
- osd: bluesfs: fix alignment for odd page sizes ([pr#7900](#), Dan Mick)
- osd: bluestore updates, scrub fixes ([pr#8035](#), Sage Weil)
- osd: bluestore/blockdevice: use std::mutex et al ([pr#7568](#), Sage Weil)
- osd: bluestore: NVMEDevice: fix error handling ([pr#7799](#), xie xingguo)

- osd: bluestore: Revert NVMEDevice task cstor and refresh interface changes ([pr#7729](#), Haomai Wang)
- osd: bluestore: add 'override' to virtual functions ([pr#7886](#), Michal Jarzabek)
- osd: bluestore: allow _dump_onode dynamic accept log level ([pr#7995](#), Jianpeng Ma)
- osd: bluestore: fix check for write falling within the same extent ([issue#14954](#), [pr#7892](#), Jianpeng Ma)
- osd: bluestore: for overwrite a extent, allocate new extent on min_alloc_size write ([pr#7996](#), Jianpeng Ma)
- osd: bluestore: improve fs-type verification and tidy up ([pr#7651](#), xie xingguo)
- osd: bluestore: misc fixes ([pr#7658](#), Jianpeng Ma)
- osd: bluestore: remove unneeded includes ([pr#7870](#), Michal Jarzabek)
- osd: clean up CMPXATTR checks ([pr#5961](#), Jianpeng Ma)
- osd: consider high/low mode when putting agent to sleep ([issue#14752](#), [pr#7631](#), Sage Weil)
- osd: ensure new osdmmaps commit before publishing them to pgs ([issue#15073](#), [pr#8096](#), Sage Weil)
- osd: filejournal: report journal entry count ([pr#7643](#), tianqing)
- osd: filestore: FALLOC_FL_PUNCH_HOLE must be used with FALLOC_FL_KEEP_SIZE ([pr#7768](#), xinxin shu)
- osd: filestore: fast abort if statfs encounters ENOENT ([pr#7703](#), xie xingguo)
- osd: filestore: fix race condition with split vs collection_move_rename and long object names ([issue#14766](#), [pr#8136](#), Samuel Just)
- osd: filestore: fix result code overwritten for clone ([issue#14817](#), [issue#14827](#), [pr#7711](#), xie xingguo)
- osd: filestore: fix wrong scope of result code for error cases during mkfs ([issue#14814](#), [pr#7704](#), xie xingguo)
- osd: filestore: fix wrong scope of result code for error cases during mount ([issue#14815](#), [pr#7707](#), xie xingguo)
- osd: filestore: restructure journal and op queue throttling ([pr#7767](#), Samuel Just)
- osd: fix forced prmootion for CALL ops ([issue#14745](#), [pr#7617](#), Sage Weil)
- osd: fix fusestore hanging during stop/quit ([issue#14786](#), [pr#7677](#), xie xingguo)
- osd: fix inaccurate counter and skip over queuing an empty transaction ([pr#7754](#), xie xingguo)
- osd: fix lack of object unblock when flush fails ([issue#14511](#), [pr#7584](#), Igor Fedotov)
- osd: fix overload of '==' operator for pg_stat_t ([issue#14921](#), [pr#7842](#), xie xingguo)
- osd: fix race condition for heartbeat_need_update ([issue#14387](#), [pr#7739](#), xie xingguo)
- osd: fix return value from maybe_handle_cache_detail() ([pr#7593](#), Igor Fedotov)
- osd: fix unnecessary object promotion when deleting from cache pool ([issue#13894](#), [pr#7537](#), Igor Fedotov)
- osd: fix wrong return type of find_osd_on_ip() ([issue#14872](#), [pr#7812](#), xie xingguo)
- osd: gobject_t: use ! instead of @ as a separator ([pr#7595](#), Sage Weil)
- osd: handle dup pg_create that races with pg deletion ([pr#8033](#), Sage Weil)
- osd: initialize last recalibrate field at construction ([pr#8071](#), xie xingguo)
- osd: kstore: fix a race condition in _txc_finish() ([pr#7804](#), Jianjian Huo)
- osd: kstore: latency breakdown ([pr#7850](#), James Liu)
- osd: kstore: sync up kstore with recent bluestore updates ([pr#7681](#), Jianjian Huo)
- osd: memstore: fix alignment of Page for test_pageset ([pr#7587](#), Casey Bodley)
- osd: min_write_recency_for_promote & min_read_recency_for_promote are tiering only ([pr#8081](#), huanwen ren)
- osd: probabilistic cache tier promotion throttling ([pr#7465](#), Sage Weil)
- osd: remove up_thru_pending field, which is never used ([pr#7991](#), xie xingguo)
- osd: replicatedpg: break out loop if we encounter fatal error during do_pg_op() ([issue#14922](#), [pr#7844](#), xie xingguo)
- osd: resolve boot vs NOUP set + clear race ([pr#7483](#), Sage Weil)
- packaging: make infernalis -> jewel upgrade work ([issue#15047](#), [pr#8034](#), Nathan Cutler)
- packaging: move cephfs repair tools to ceph-common ([issue#15145](#), [pr#8133](#), Boris Ranto, Ken Dreyer)
- pybind/rados: fix object lifetime issues and other bugs in aio ([pr#7778](#), Hector Martin)
- pybind/rados: use __dealloc__ since __del__ is ignored by cython ([pr#7692](#), Mehdi Abaakouk)
- pybind: Ensure correct python flags are passed ([pr#7663](#), James Page)
- pybind: flag an RBD image as closed regardless of result code ([pr#8005](#), Jason Dillaman)
- pybind: move cephfs to Cython ([pr#7745](#), John Spray, Mehdi Abaakouk)
- pybind: remove next() on iterators ([pr#7706](#), Mehdi Abaakouk)
- pybind: replace __del__ with __dealloc__ for rbd ([pr#7708](#), Josh Durgin)
- pybind: use correct subdir for rados install-exec rule ([pr#7684](#), Josh Durgin)
- python binding of librados with cython ([pr#7621](#), Mehdi Abaakouk)
- python: use pip instead of python setup.py ([pr#7605](#), Loic Dachary)
- qa/workunits/cephtool/test.sh: wait longer in ceph_watch_start() ([issue#14910](#), [pr#7861](#), Kefu Chai)
- qa/workunits/rados/test.sh: capture stderr too ([pr#8004](#), Sage Weil)
- qa/workunits/rados/test.sh: test tmap_migrate ([pr#8114](#), Sage Weil)
- qa/workunits/rbd: do not use object map during read flag testing ([pr#8104](#), Jason Dillaman)
- qa/workunits/rbd: new online maintenance op tests ([pr#8216](#), Jason Dillaman)
- qa/workunits/rbd: use POSIX function definition ([issue#15104](#), [pr#8068](#), Nathan Cutler)
- qa/workunits/rest/test.py: add confirmation to 'mds setmap' ([issue#14606](#), [pr#7982](#), Sage Weil)
- qa/workunits/rest/test.py: don't use newfs ([pr#8191](#), Sage Weil)
- qa: add workunit to run ceph_test_rbd_mirror ([pr#8221](#), Josh Durgin)
- rados: add ceph:: namespace to bufferlist type ([pr#8059](#), Noah Watkins)
- rados: fix bug for write bench ([pr#7851](#), James Liu)
- rbd-mirror: ImageReplayer async start/stop ([pr#7944](#), Mykola Golub)

- rbd-mirror: ImageReplayer improvements ([pr#7759](#), Mykola Golub)
- rbd-mirror: fix image replay test failures ([pr#8158](#), Jason Dillaman)
- rbd-mirror: fix long termination due to 30sec wait in main loop ([pr#8185](#), Mykola Golub)
- rbd-mirror: implement ImageReplayer ([pr#7614](#), Mykola Golub)
- rbd-mirror: integrate with image sync state machine ([pr#8079](#), Jason Dillaman)
- rbd-mirror: minor fix-ups for initial skeleton implementation ([pr#7958](#), Mykola Golub)
- rbd-mirror: remote to local cluster image sync ([pr#7979](#), Jason Dillaman)
- rbd-mirror: use pool/image names in asok commands ([pr#8159](#), Mykola Golub)
- rbd-mirror: use the mirroring directory to detect candidate images ([issue#15142](#), [pr#8162](#), Ricardo Dias)
- rbd/run_cli_tests.sh: Reflect test failures ([issue#14825](#), [pr#7781](#), Zack Cerza)
- rbd: add support for mirror image promotion/demotion/resync ([pr#8138](#), Jason Dillaman)
- rbd: clone operation should default to image format 2 ([pr#8119](#), Jason Dillaman)
- rbd: deprecate image format 1 ([pr#7841](#), Jason Dillaman)
- rbd: support for enabling/disabling mirroring on specific images ([issue#13296](#), [pr#8056](#), Ricardo Dias)
- release-notes: draft v0.94.6 release notes ([issue#13356](#), [pr#7689](#), Abhishek Varshney, Loic Dachary)
- release-notes: draft v10.0.3 release notes ([pr#7592](#), Loic Dachary)
- release-notes: draft v10.0.4 release notes ([pr#7966](#), Loic Dachary)
- release-notes: draft v9.2.1 release notes ([issue#13750](#), [pr#7694](#), Abhishek Varshney)
- rgw: Idap (Matt Benjamin) ([pr#7985](#), Matt Benjamin)
- rgw: multisite fixes ([pr#8013](#), Yehuda Sadeh)
- rgw: support for aws authentication v4 (Javier M. Mellid) ([issue#10333](#), [pr#7720](#), Yehuda Sadeh, Javier M. Mellid)
- rgw sync fixes ([pr#8095](#), Yehuda Sadeh)
- rgw/rgw_common.h: fix the RGWBucketInfo decoding ([pr#8165](#), Kefu Chai)
- rgw/rgw_common.h: fix the RGWBucketInfo decoding ([pr#8154](#), Kefu Chai)
- rgw/rgw_orphan: check the return value of save_state ([pr#7544](#), Boris Ranto)
- rgw: Allow an implicit tenant in case of Keystone ([pr#8139](#), Pete Zaitcev)
- rgw: Drop unused usage_exit from rgw_admin.cc ([pr#7632](#), Pete Zaitcev)
- rgw: RGWZoneParams::create should not handle -EEXIST error ([pr#7927](#), Orit Wasserman)
- rgw: add bucket request payment feature usage statistics integration ([issue#13834](#), [pr#6656](#), Javier M. Mellid)
- rgw: add support for caching of Keystone admin token. ([pr#7630](#), Radoslaw Zarzynski)
- rgw: add support for metadata upload during PUT on Swift container. ([pr#8002](#), Radoslaw Zarzynski)
- rgw: add support for system requests over Swift API ([pr#7666](#), Radoslaw Zarzynski)
- rgw: adjust the request_uri to support absoluteURI of http request ([issue#12917](#), [pr#7675](#), Wenjun Huang)
- rgw: admin api for retrieving usage info (Ji Chen) ([pr#8031](#), Yehuda Sadeh, Ji Chen)
- rgw: allow authentication keystone with self signed certs ([issue#14853](#), [issue#13422](#), [pr#7777](#), Abhishek Lekshmanan)
- rgw: approximate AmazonS3 HostId error field. ([pr#7444](#), Robin H. Johnson)
- rgw: calculate payload hash in RGWPutObj_ObjStore only when necessary. ([pr#7869](#), Radoslaw Zarzynski)
- rgw: cleanups to comments and messages ([pr#7633](#), Pete Zaitcev)
- rgw: don't use s->bucket for metadata api path entry ([issue#14549](#), [pr#7408](#), Yehuda Sadeh)
- rgw: drop permissions of rgw/civetweb after startup ([issue#13600](#), [pr#8019](#), Karol Mroz)
- rgw: fcgi should include acconfig ([pr#7760](#), Abhishek Lekshmanan)
- rgw: fix wrong handling of limit=0 during listing of Swift account. ([issue#14903](#), [pr#7821](#), Radoslaw Zarzynski)
- rgw: fixes for per-period metadata logs ([pr#7827](#), Casey Bodley)
- rgw: improve error handling in S3/Keystone integration ([pr#7597](#), Radoslaw Zarzynski)
- rgw: link civetweb with openssl (Sage, Marcus Watts) ([pr#7825](#), Marcus Watts, Sage Weil)
- rgw: link payer info to usage logging ([pr#7918](#), Yehuda Sadeh, Javier M. Mellid)
- rgw: move signal.h dependency from rgw_front.h ([pr#7678](#), Matt Benjamin)
- rgw: multiple Swift API compliance improvements for TempURL (Radoslaw Zarzynsk) ([issue#14806](#), [issue#11163](#), [pr#7891](#), Radoslaw Zarzynski)
- rgw: multiple improvements regarding etag calculation for SLO/DLO of Swift API. ([pr#7764](#), Radoslaw Zarzynski)
- rgw: remove duplicated code in RGWRados::get_bucket_info() ([pr#7413](#), liyankun)
- rgw: remove unused vector ([pr#7990](#), Na Xie)
- rgw: reset return code in when iterating over the bucket the objects ([issue#14826](#), [pr#7803](#), Orit Wasserman)
- rgw: store system object meta in cache when creating it ([issue#14678](#), [pr#7615](#), Yehuda Sadeh)
- rgw: support json format for admin policy API (Dunrong Huang) ([issue#14090](#), [pr#8036](#), Yehuda Sadeh, Dunrong Huang)
- rgw: try to parse Keystone token in order appropriate to configuration. ([pr#7822](#), Radoslaw Zarzynski)
- rgw: use pimpl pattern for RGWPeriodHistory ([pr#7809](#), Casey Bodley)
- rgw: user quota may not adjust on bucket removal ([issue#14507](#), [pr#7586](#), root)
- rgw:bucket link now set the bucket.instance acl (bug fix) ([issue#11076](#), [pr#8037](#), Zengran Zhang)
- rpm,deb: remove conditional BuildRequires for btrfs-progs ([issue#15042](#), [pr#8016](#), Erwan Velu)
- rpm: remove sub-package dependencies on "ceph" ([issue#15146](#), [pr#8137](#), Ken Dreyer)
- script: add missing stop_rgw variable to stop.sh script ([pr#7959](#), Karol Mroz)
- selinux: Update policy to grant additional access ([issue#14870](#), [pr#7971](#), Boris Ranto)
- selinux: allow log files to be located in /var/log/radosgw ([pr#7604](#), Boris Ranto)
- sstring.hh: return type from str_len(...) need not be const ([pr#7679](#), Matt Benjamin)

- submodules: revert an accidental change ([pr#7929](#), Yehuda Sadeh)
- systemd: correctly escape block device paths ([issue#14706](#), [pr#7579](#), James Page)
- test/TestPGLog: fix the FTBFS ([issue#14930](#), [pr#7855](#), Kefu Chai)
- test/bufferlist: Avoid false-positive tests ([pr#7955](#), Erwan Velu)
- test/cli-integration/rbd: disable progress output ([issue#14931](#), [pr#7858](#), Josh Durgin)
- test/osd: Relax the timing intervals in osd-markdown.sh ([pr#7899](#), Dan Mick)
- test/pybind/test_ceph_argparse: fix reweight-by-utilization tests ([pr#8027](#), Kefu Chai, Sage Weil)
- test/radosgw-admin: update the expected usage outputs ([pr#7723](#), Kefu Chai)
- test/rgw: add multisite test for meta sync across periods ([pr#7887](#), Casey Bodley)
- test/time: no need to abs(uint64_t) for comparing ([pr#7726](#), Kefu Chai)
- test: add missing shut_down mock method ([pr#8125](#), Jason Dillaman)
- test: correct librbd errors discovered with unoptimized cmake build ([pr#7914](#), Jason Dillaman)
- test: create pools for rbd tests with different prefix ([pr#7738](#), Mykola Golub)
- test: enable test for bug #2339 which has been resolved. ([pr#7743](#), You Ji)
- test: fix issues discovered via the rbd permissions test case ([pr#8129](#), Jason Dillaman)
- test: fixup and improvements for rbd-mirror test ([pr#8090](#), Mykola Golub)
- test: handle exception thrown from close during rbd lock test ([pr#8124](#), Jason Dillaman)
- test: more debug logging for TestWatchNotify ([pr#7737](#), Mykola Golub)
- test: new librbd flatten test case ([pr#7609](#), Jason Dillaman)
- test: rbd-mirror: add “switch to the next tag” test ([pr#8149](#), Mykola Golub)
- test: rbd-mirror: compare positions using all fields ([pr#8172](#), Mykola Golub)
- test: reproducer for writeback CoW deadlock ([pr#8009](#), Jason Dillaman)
- test: update rbd integration cram tests for new default features ([pr#8001](#), Jason Dillaman)
- test_pool_create.sh: put test files in the test dir so they are cleaned up ([pr#8219](#), Josh Durgin)
- tests: ceph-disk.sh: should use “readlink -f” instead ([pr#7594](#), Kefu Chai)
- tests: ceph-disk.sh: use “readlink -f” instead for fullpath ([pr#7606](#), Kefu Chai)
- tests: fix a few build warnings ([pr#7608](#), Sage Weil)
- tests: sync ceph-erasure-code-corpus for mktemp -d ([pr#7596](#), Loic Dachary)
- tests: test_pidfile.sh lingering processes ([issue#14834](#), [pr#7734](#), Loic Dachary)
- tools/cephfs: add tmap_upgrade ([pr#7003](#), John Spray)
- tools/cephfs: fix tmap_upgrade ([issue#15135](#), [pr#8128](#), John Spray)
- tools/rados: reduce “rados put” memory usage by op_size ([pr#7928](#), Piotr Dałek)
- unittest_compression_zlib: do not assume buffer will be null terminated ([pr#8064](#), Sage Weil)
- unittest_osdmap: default crush tunables now firefly ([pr#8098](#), Sage Weil)
- vstart.sh: avoid race condition starting rgw via vstart.sh ([issue#14829](#), [pr#7727](#), Javier M. Mellid)
- vstart.sh: silence a harmless msg where btrfs is not found ([pr#7640](#), Patrick Donnelly)
- xio: fix incorrect ip being assigned in case of multiple RDMA ports ([pr#7747](#), Subramanyam Varanasi)
- xio: remove duplicate assignment of peer addr ([pr#8025](#), Roi Dayan)
- xio: remove redundant magic methods ([pr#7773](#), Roi Dayan)
- xio: remove unused variable ([pr#8023](#), Roi Dayan)
- xxhash: use clone of xxhash.git; add .gitignore ([pr#7986](#), Sage Weil)
- rbd: update default image features ([pr#7846](#), Jason Dillaman)
- rbd-mirror: make remote context respect env and argv config params ([pr#8182](#), Mykola Golub)
- journal: re-use common threads between journalers ([pr#7906](#), Jason Dillaman)
- client: add option to control how directory size is calculated ([pr#7323](#), Yan, Zheng)
- rgw: keystone v3 ([pr#7719](#), Mark Barnes, Radoslaw Zarzynski)
- rgw: new multisite merge ([issue#14549](#), [pr#7709](#), Yehuda Sadeh, Orit Wasserman, Casey Bodley, Daniel Gryniewicz)
- rgw: adjust error code when bucket does not exist in copy operation ([issue#14975](#), [pr#7916](#), Yehuda Sadeh)
- rgw: indexless ([pr#7786](#), Yehuda Sadeh)

V10.0.5

This is identical to v10.0.4 and was only created because of a git tagging mistake.

V10.0.4

This is the fifth and last development release before Jewel. The next release will be a release candidate with the final set of features. Big items include RGW static website support, librbd journal framework, fixed mon sync of config-key data, C++11 updates, and bluestore/kstore.

NOTABLE CHANGES

- admin/build-doc: depend on zlib1g-dev and graphviz ([pr#7522](#), Ken Dreyer)
- buffer: use move construct to append/push_back/push_front ([pr#7455](#), Haomai Wang)
- build: allow jemalloc with rocksdb-static ([pr#7368](#), Somnath Roy)
- build: fix the autotools and cmake build (the new fusestore needs libfuse) ([pr#7393](#), Kefu Chai)
- build: fix warnings ([pr#7197](#), Kefu Chai, xie xingguo)
- build: fix warnings ([pr#7315](#), Kefu Chai)
- build: kill warnings ([pr#7397](#), Kefu Chai)
- build: move libexec scripts to standardize across distros ([issue#14687](#), [issue#14705](#), [issue#14723](#), [pr#7636](#), Nathan Cutler, Kefu Chai)
- build: Refrain from versioning and packaging EC testing plugins ([issue#14756](#), [issue#14723](#), [pr#7637](#), Nathan Cutler, Kefu Chai)
- build: spdk submodule; cmake ([pr#7503](#), Kefu Chai)
- ceph-disk: support bluestore ([issue#13422](#), [pr#7218](#), Loic Dachary, Sage Weil)
- ceph-disk/test: fix test_prepare.py::TestPrepare tests ([pr#7549](#), Kefu Chai)
- cleanup: remove misc dead code ([pr#7201](#), Erwan Vélu)
- cls/cls_rbd: pass string by reference ([pr#7232](#), Jeffrey Lu)
- cmake: Added new unittests to make check ([pr#7572](#), Ali Maredia)
- cmake: add KernelDevice.cc to libos_srcs ([pr#7507](#), Kefu Chai)
- cmake: check for libsnappy in default path also ([pr#7366](#), Kefu Chai)
- cmake: feb5 ([pr#7541](#), Matt Benjamin)
- cmake: For CMake version <= 2.8.11, use LINK_PRIVATE and LINK_PUBLIC ([pr#7474](#), Tao Chang)
- cmake: let ceph-client-debug link with tcmalloc ([pr#7314](#), Kefu Chai)
- cmake: support ccache via a WITH_CCACHE build option ([pr#6875](#), John Coyle)
- common: add zlib compression plugin ([pr#7437](#), Alyona Kiseleva, Kiseleva Alyona)
- common: admin socket commands for tcmalloc heap get/set operations ([pr#7512](#), Samuel Just)
- common: ake ceph_time clocks work under BSD ([pr#7340](#), Adam C. Emerson)
- common: Allow OPT_INT settings with negative values ([issue#13829](#), [pr#7390](#), Brad Hubbard, Kefu Chai)
- common/buffer: replace RWLock with spinlocks ([pr#7294](#), Piotr Dałek)
- common: change the type of counter total/unhealthy_workers ([pr#7254](#), Guang Yang)
- common: snappy decompressor may assert when handling segmented input bufferlist ([issue#14400](#), [pr#7268](#), Igor Fedotov)
- common/str_map: cleanup: replaced get_str_map() function overloading by using default parameters for delimiters ([pr#7266](#), Sahithi R V)
- common: time: have skewing-now call non-skewing now ([pr#7466](#), Adam C. Emerson)
- common: unit test for interval_set implementations ([pr#6](#), Igor Fedotov)
- config: add \$data_dir/config to config search path ([pr#7377](#), Sage Weil)
- configure.ac: make “–with-librocksdb-static” default to ‘check’ ([issue#14463](#), [pr#7317](#), Dan Mick)
- crush: add safety assert ([issue#14496](#), [pr#7344](#), songbaisen)
- crush: reply quickly from get_immediate_parent ([issue#14334](#), [pr#7181](#), song bisen)
- debian: packaging fixes for jewel ([pr#7807](#), Ken Dreyer, Ali Maredia)
- debian/rpm split servers ([issue#10587](#), [pr#7746](#), Ken Dreyer)
- doc: add orphans commands to radosgw-admin(8) ([issue#14637](#), [pr#7518](#), Ken Dreyer)
- doc: amend the rados.8 ([pr#7251](#), Kefu Chai)
- doc: Fixes a CRUSH map step take argument ([pr#7327](#), Ivan Grcic)
- doc: fixing image in section ERASURE CODING ([pr#7298](#), Rachana Patel)
- doc: fix misleading configuration guide on cache tiering ([pr#7000](#), Yuan Zhou)
- doc: fix S3 C# example ([pr#7027](#), Dunrong Huang)
- doc: remove redundant space in ceph-authtool/monmaptool doc ([pr#7244](#), Jiaying Ren)
- doc: revise SubmittingPatches ([pr#7292](#), Kefu Chai)
- doc: rgw: port changes from downstream to upstream ([pr#7264](#), Bara Ancincova)
- doc: script and guidelines for mirroring Ceph ([pr#7384](#), Wido den Hollander)
- doc: use ‘ceph auth get-or-create’ for creating RGW keyring ([pr#6930](#), Wido den Hollander)
- global: do not start two daemons with a single pid-file ([issue#13422](#), [pr#7075](#), shun song)
- global: do not start two daemons with a single pid-file (part 2) ([issue#13422](#), [pr#7463](#), Loic Dachary)
- journal: flush commit position on metadata shutdown ([pr#7385](#), Mykola Golub)
- journal: reset commit_position_task_ctx pointer after task complete ([pr#7480](#), Mykola Golub)
- libcephfs: update LIBCEPHFS_VERSION to indicate the interface was changed ([pr#7551](#), Jevon Qiao)
- librados: move to c++11 concurrency types ([pr#5931](#), Adam C. Emerson)
- librados: remove duplicate definitions for rados pool_stat_t and cluster_stat_t ([pr#7330](#), Igor Fedotov)
- librados: shutdown finisher in a more graceful way ([pr#7519](#), xie xingguo)
- librados_test_stub: protect against notify/unwatch race ([pr#7540](#), Jason Dillaman)
- librbd: API: async open and close ([issue#14264](#), [pr#7259](#), Mykola Golub)
- librbd: Avoid create two threads per image ([pr#7400](#), Haomai Wang)
- librbd: block maintenance ops until after journal is ready ([issue#14510](#), [pr#7382](#), Jason Dillaman)
- librbd: fix internal handling of dynamic feature updates ([pr#7299](#), Jason Dillaman)

- librbd: journal framework for tracking exclusive lock transitions ([issue#13298](#), [pr#7529](#), Jason Dillaman)
- librbd: journal shut down flush race condition ([issue#14434](#), [pr#7302](#), Jason Dillaman)
- librbd: remove canceled tasks from timer thread ([issue#14476](#), [pr#7329](#), Douglas Fuller)
- makefile: remove libedit from libclient.la ([pr#7284](#), Kefu Chai)
- mds, client: fix locking around handle_conf_change ([issue#14365](#), [issue#14374](#), [pr#7312](#), John Spray)
- mds: judgment added to avoid the risk of visiting the NULL pointer ([pr#7358](#), Kongming Wu)
- mon: add an independent option for max election time ([pr#7245](#), Sangdi Xu)
- mon: compact full epochs also ([issue#14537](#), [pr#7396](#), Kefu Chai)
- mon: consider the pool size when setting pool crush rule ([issue#14495](#), [pr#7341](#), song baisen)
- mon: drop useless rank init assignment ([issue#14508](#), [pr#7321](#), huanwen ren)
- mon: fix locking in preinit error paths ([issue#14473](#), [pr#7353](#), huanwen ren)
- mon: fix monmap creation stamp ([pr#7459](#), duanweijun)
- mon: fix sync of config-key data ([pr#7363](#), Xiaowei Chen)
- mon: go into ERR state if multiple PGs are stuck inactive ([issue#13923](#), [pr#7253](#), Wido den Hollander)
- mon/MDSMonitor.cc: properly note beacon when health metrics changes ([issue#14684](#), [pr#7757](#), Yan, Zheng)
- mon/MonClient: avoid null pointer error when configured incorrectly ([issue#14405](#), [pr#7276](#), Bo Cai)
- mon: PG Monitor should report waiting for backfill ([issue#12744](#), [pr#7398](#), Abhishek Lekshmanan)
- mon: reduce CPU and memory manager pressure of pg health check ([pr#7482](#), Piotr Dałek)
- mon: some cleanup in MonmapMonitor.cc ([pr#7418](#), huanwen ren)
- mon: warn if pg(s) not scrubbed ([issue#13142](#), [pr#6440](#), Michal Jarzabek)
- msg/async: AsyncConnection: avoid debug log in cleanup_handler ([pr#7547](#), Haomai Wang)
- msg/async: bunch of fixes ([pr#7379](#), Piotr Dałek)
- msg/async: fix array boundary ([pr#7451](#), Wei Jin)
- msg/async: fix potential race condition ([pr#7453](#), Haomai Wang)
- msg/async: fix send closed local_connection message problem ([pr#7255](#), Haomai Wang)
- msg/async: reduce extra tcp packet for message ack ([pr#7380](#), Haomai Wang)
- msg/xio: fix compilation ([pr#7479](#), Roi Dayan)
- organizationmap: modify org mail info. ([pr#7240](#), Xiaowei Chen)
- os/bluestore: fix assert ([issue#14436](#), [pr#7293](#), xie xingguo)
- os/bluestore: fix bluestore_wal_transaction_t encoding test ([pr#7342](#), Kefu Chai)
- os/bluestore: insert new onode to the front position of onode LRU ([pr#7492](#), Jianjian Huo)
- os/bluestore: use intrusive_ptr for Dir ([pr#7247](#), Igor Fedotov)
- osd: blockdevice: avoid implicit cast and add guard ([pr#7460](#), xie xingguo)
- osd: bluestore/BlueFS: initialize super block_size earlier in mkfs ([pr#7535](#), Sage Weil)
- osd: BlueStore: fix fsck and blockdevice read-relevant issue ([pr#7362](#), xie xingguo)
- osd: BlueStore: fix null pointer access ([issue#14561](#), [pr#7435](#), xie xingguo)
- osd: bluestore, kstore: fix nid overwritten logic ([issue#14407](#), [issue#14433](#), [pr#7283](#), xie xingguo)
- osd: bluestore: use btree_map for allocator ([pr#7269](#), Igor Fedotov, Sage Weil)
- osd: drop fiemap len=0 logic ([pr#7267](#), Sage Weil)
- osd: FileStore: add error check for object_map->sync() ([pr#7281](#), Chendi Xue)
- osd: FileStore: cleanup: remove obsolete option “filestore_xattr_use_omap” ([issue#14356](#), [pr#7217](#), Vikhyat Umrao)
- osd: FileStore: modify the format of colon ([pr#7333](#), Donghai Xu)
- osd: FileStore: print file name before osd assert if read file failed ([pr#7111](#), Ji Chen)
- osd: fix invalid list traversal in process_copy_chunk ([pr#7511](#), Samuel Just)
- osd, mon: fix exit issue ([pr#7420](#), Jiaying Ren)
- osd: PG::activate(): handle unexpected cached_removed_snaps more gracefully ([issue#14428](#), [pr#7309](#), Alexey Sheplyakov)
- os/fs: fix io_getevents argument ([pr#7355](#), Jingkai Yuan)
- os/fusestore: add error handling ([pr#7395](#), xie xingguo)
- os/keyvaluestore: kill KeyValueStore ([pr#7320](#), Haomai Wang)
- os/kstore: insert new onode to the front position of onode LRU ([pr#7505](#), xie xingguo)
- os/ObjectStore: add custom move operations for ObjectStore::Transaction ([pr#7303](#), Casey Bodley)
- rgw: Bug fix for mtime anomalies in RadosGW and other places ([pr#7328](#), Adam C. Emerson, Casey Bodley)
- rpm: move %post(un) Idconfig calls to ceph-base ([issue#14940](#), [pr#7867](#), Nathan Cutler)
- rpm: move runtime dependencies to ceph-base and fix other packaging issues ([issue#14864](#), [pr#7826](#), Nathan Cutler)
- test: ceph_test_rados: use less CPU ([pr#7513](#), Samuel Just)

V10.0.3

This is the fourth development release for Jewel. Several big pieces have been added this release, including BlueStore (a new backend for OSD to replace FileStore), many ceph-disk fixes, a new CRUSH tunable that improves mapping stability, a new librados object enumeration API, and a whole slew of OSD and RADOS optimizations.

Note that, due to general developer busyness, we aren't building official release packages for this dev release. You can fetch

NOTABLE CHANGES

- bluestore: latest and greatest ([issue#14210](#), [issue#13801](#), [pr#6896](#), xie.xingguo, Jianpeng Ma, YiQiang Chen, Sage Weil, Ning Yao)
- buffer: fix internal iterator invalidation on rebuild, get_contiguous ([pr#6962](#), Sage Weil)
- build: fix a few warnings ([pr#6847](#), Orit Wasserman)
- build: misc make check fixes ([pr#7153](#), Sage Weil)
- ceph-detect-init: fix py3 test ([pr#7025](#), Kefu Chai)
- ceph-disk: add -f flag for btrfs mkfs ([pr#7222](#), Darrell Enns)
- ceph-disk: ceph-disk list fails on /dev/cciss!c0d0 ([issue#13970](#), [issue#14233](#), [issue#14230](#), [pr#6879](#), Loic Dachary)
- ceph-disk: fix failures when preparing disks with udev > 214 ([issue#14080](#), [issue#14094](#), [pr#6926](#), Loic Dachary, Ilya Dryomov)
- ceph-disk: Fix trivial typo ([pr#7472](#), Brad Hubbard)
- ceph-disk: warn for prepare partitions with bad GUIDs ([issue#13943](#), [pr#6760](#), David Disseldorp)
- ceph-fuse: fix double decreasing the count to trim caps ([issue#14319](#), [pr#7229](#), Zhi Zhang)
- ceph-fuse: fix double free of args ([pr#7015](#), Ilya Shipitsin)
- ceph-fuse: fix fsync() ([pr#6388](#), Yan, Zheng)
- ceph-fuse: print usage information when no parameter specified ([pr#6868](#), Bo Cai)
- ceph: improve the error message ([issue#11101](#), [pr#7106](#), Kefu Chai)
- ceph.in: avoid a broken pipe error when use ceph command ([issue#14354](#), [pr#7212](#), Bo Cai)
- ceph.spec.in: add copyright notice ([issue#14694](#), [pr#7569](#), Nathan Cutler)
- ceph.spec.in: add license declaration ([pr#7574](#), Nathan Cutler)
- ceph_test_libcephfs: tolerate duplicated entries in readdir ([issue#14377](#), [pr#7246](#), Yan, Zheng)
- client: check if Fh is readable when processing a read ([issue#11517](#), [pr#7209](#), Yan, Zheng)
- client: properly trim unlinked inode ([issue#13903](#), [pr#7297](#), Yan, Zheng)
- cls_rbd: add guards for error cases ([issue#14316](#), [issue#14317](#), [pr#7165](#), xie xingguo)
- cls_rbd: enable object map checksums for object_map_save ([issue#14280](#), [pr#7149](#), Douglas Fuller)
- cmake: Add ENABLE_GIT_VERSION to avoid rebuilding ([pr#7171](#), Kefu Chai)
- cmake: add missing check for HAVE_EXECINFO_H ([pr#7270](#), Casey Bodley)
- cmake: cleanups and more features from automake ([pr#7103](#), Casey Bodley, Ali Maredia)
- cmake: detect bzip2 and lz4 ([pr#7126](#), Kefu Chai)
- cmake: fix build with bluestore ([pr#7099](#), John Spray)
- cmake: fix the build on trusty ([pr#7249](#), Kefu Chai)
- cmake: made rocksdb an imported library ([pr#7131](#), Ali Maredia)
- cmake: no need to run configure from run-cmake-check.sh ([pr#6959](#), Orit Wasserman)
- cmake: test_build_libcephfs needs \${ALLOC_LIBS} ([pr#7300](#), Ali Maredia)
- common/address_help.cc: fix the leak in entity_addr_from_url() ([issue#14132](#), [pr#6987](#), Qiankun Zheng)
- common: add thread names ([pr#5882](#), Igor Podoski)
- common: assert: abort() rather than throw ([pr#6804](#), Adam C. Emerson)
- common: buffer/assert minor fixes ([pr#6990](#), Matt Benjamin)
- common/Formatter: avoid newline if there is no output ([pr#5351](#), Aran85)
- common: improve shared_cache and simple_cache efficiency with hash table ([pr#6909](#), Ning Yao)
- common/lockdep: increase max lock names ([pr#6961](#), Sage Weil)
- common: new timekeeping common code, and Objecter conversion ([pr#5782](#), Adam C. Emerson)
- common: signal_handler: added support for using reentrant strsignal() implementations vs. sys_siglist[] ([pr#6796](#), John Coyle)
- config: complains when a setting is not tracked ([issue#11692](#), [pr#7085](#), Kefu Chai)
- configure: detect bz2 and lz4 ([issue#13850](#), [issue#13981](#), [pr#7030](#), Kefu Chai)
- correct radosgw-admin command ([pr#7006](#), YankunLi)
- crush: add chooseleaf_stable tunable ([pr#6572](#), Sangdi Xu, Sage Weil)
- crush: clean up whitespace removal ([issue#14302](#), [pr#7157](#), songbaisen)
- crush/CrushTester: check for overlapped rules ([pr#7139](#), Kefu Chai)
- crushtool: improve usage/tip messages ([pr#7142](#), xie xingguo)
- crushtool: set type 0 name “device” for –build option ([pr#6824](#), Sangdi Xu)
- doc: adding “–allow-shrink” in decreasing the size of the rbd block to distinguish from the increasing option ([pr#7020](#), Yehua)
- doc: admin/build-doc: make paths absolute ([pr#7119](#), Dan Mick)
- doc: dev: document ceph-qa-suite ([pr#6955](#), Loic Dachary)
- doc: document “readforward” and “readproxy” cache mode ([pr#7023](#), Kefu Chai)
- doc: fix “mon osd down out subtree limit” option name ([pr#7164](#), François Lafont)
- doc: fix typo ([pr#7004](#), tianqing)
- doc: Updated the rados command man page to include the –run-name opt... ([issue#12899](#), [pr#5900](#), ritz303)

- fs: be more careful about the “mds setmap” command to prevent breakage ([issue#14380](#), [pr#7262](#), Yan, Zheng)
- helgrind: additional race conditionslibrbd: journal replay should honor inter-event dependencies ([pr#7274](#), Jason Dillaman)
- helgrind: fix real (and imaginary) race conditions ([issue#14163](#), [pr#7208](#), Jason Dillaman)
- kv: implement value_as_ptr() and use it in .get() ([pr#7052](#), Piotr Dałek)
- librados: add c++ style osd/pg command interface ([pr#6893](#), Yunchuan Wen)
- librados: fix several flaws introduced by the enumeration_objects API ([issue#14299](#), [issue#14301](#), [issue#14300](#), [pr#7156](#), xie xingguo)
- librados: new style (sharded) object listing ([pr#6405](#), John Spray, Sage Weil)
- librados: potential null pointer access in list_(n)objects ([issue#13822](#), [pr#6639](#), xie xingguo)
- librbd: exit if parent’s snap is gone during clone ([issue#14118](#), [pr#6968](#), xie xingguo)
- librbd: fix potential memory leak ([issue#14332](#), [issue#14333](#), [pr#7174](#), xie xingguo)
- librbd: fix snap_exists API return code overflow ([issue#14129](#), [pr#6986](#), xie xingguo)
- librbd: journal replay should honor inter-event dependencies ([pr#7019](#), Jason Dillaman)
- librbd: return error if we fail to delete object_map head object ([issue#14098](#), [pr#6958](#), xie xingguo)
- librbd: small fixes for error messages and readahead counter ([issue#14127](#), [pr#6983](#), xie xingguo)
- librbd: uninitialized state in snap remove state machine ([pr#6982](#), Jason Dillaman)
- mailmap: hange organization for Dongmao Zhang ([pr#7173](#), Dongmao Zhang)
- mailmap: Igor Podoski affiliation ([pr#7219](#), Igor Podoski)
- mailmap update ([pr#7210](#), M Ranga Swami Reddy)
- mailmap updates ([pr#6992](#), Loic Dachary)
- mailmap updates ([pr#7189](#), Loic Dachary)
- man: document listwatchers cmd in “rados” manpage ([pr#7021](#), Kefu Chai)
- mds: advance clientreplay when replying ([issue#14357](#), [pr#7216](#), John Spray)
- mds: expose state of recovery to status ASOK command ([issue#14146](#), [pr#7068](#), Yan, Zheng)
- mds: fix client cap/message replay order on restart ([issue#14254](#), [issue#13546](#), [pr#7199](#), Yan, Zheng)
- mds: fix standby replay thread creation ([issue#14144](#), [pr#7132](#), John Spray)
- mds: we should wait messenger when MDSDaemon suicide ([pr#6996](#), Wei Feng)
- mon: add *osd blacklist clear* ([pr#6945](#), John Spray)
- mon: add RAW USED column to ceph df detail ([pr#7087](#), Ruifeng Yang)
- mon: degrade a log message to level 2 ([pr#6929](#), Kongming Wu)
- mon: fix coding-style on PG related Monitor files ([pr#6881](#), Wido den Hollander)
- mon: fixes related to mondbstore->get() changes ([pr#6564](#), Piotr Dałek)
- mon: fix reuse of osd ids (clear osd info on osd deletion) ([issue#13988](#), [pr#6900](#), Loic Dachary, Sage Weil)
- mon: fix the can’t change subscribe level bug in monitoring log ([pr#7031](#), Zhiqiang Wang)
- mon/MDSMonitor: add confirmation to “ceph mds rmfailed” ([issue#14379](#), [pr#7248](#), Yan, Zheng)
- mon: modify a dout level in OSDMonitor.cc ([pr#6928](#), Yongqiang He)
- mon: MonmapMonitor: don’t expose uncommitted state to client ([pr#6854](#), Joao Eduardo Luis)
- mon/OSDMonitor: osdmap laggy set a maximum limit for interval ([pr#7109](#), Zengran Zhang)
- mon: paxos is_recovering calc error ([pr#7227](#), Weijun Duan)
- mon/PGMap: show rd/wr iops separately in status reports ([pr#7072](#), Cilang Zhao)
- mon: PGMonitor: acting primary diff with cur_stat, should not set pg to stale ([pr#7083](#), Xiaowei Chen)
- msg: add override to virutal methods ([pr#6977](#), Michal Jarzabek)
- msg/async: cleanup dead connection and misc things ([pr#7158](#), Haomai Wang)
- msg/async: don’t use shared_ptr to manage EventCallback ([pr#7028](#), Haomai Wang)
- msg: filter out lo addr when bind osd addr ([pr#7012](#), Ji Chen)
- msg: removed unneeded includes from Dispatcher ([pr#6814](#), Michal Jarzabek)
- msg: remove unneeded inline ([pr#6989](#), Michal Jarzabek)
- msgr: fix large message data content length causing overflow ([pr#6809](#), Jun Huang, Haomai Wang)
- msg/simple: pipe: memory leak when signature check failed ([pr#7096](#), Ruifeng Yang)
- msg/simple: remove unneeded friend declarations ([pr#6924](#), Michal Jarzabek)
- objecter: avoid recursive lock of Objecter::rwlock ([pr#7343](#), Yan, Zheng)
- os/bluestore: fix bluestore_wal_transaction_t encoding test ([pr#7419](#), Kefu Chai, Brad Hubbard)
- osd: add cache hint when pushing raw clone during recovery ([pr#7069](#), Zhiqiang Wang)
- osd: avoid debug std::string initialization in PG:::get/put ([pr#7117](#), Evgeniy Firsov)
- osd: avoid osd_op_thread suicide because osd_scrub_sleep ([pr#7009](#), Jianpeng Ma)
- osd: bluestore: bluefs: fix several small bugs ([issue#14344](#), [issue#14343](#), [pr#7200](#), xie xingguo)
- osd: bluestore: don’t include when building without libaio ([issue#14207](#), [pr#7169](#), Mykola Golub)
- osd: bluestore: fix bluestore onode_t attr leak ([pr#7125](#), Ning Yao)
- osd: bluestore: fix bluestore_wal_transaction_t encoding test ([pr#7168](#), Kefu Chai)
- osd: bluestore: fix several bugs ([issue#14259](#), [issue#14353](#), [issue#14260](#), [issue#14261](#), [pr#7122](#), xie xingguo)
- osd: bluestore: fix space rebalancing, collection split, buffered reads ([pr#7196](#), Sage Weil)
- osd: bluestore: more fixes ([pr#7130](#), Sage Weil)
- osd: cache tier: add config option for eviction check list size ([pr#6997](#), Yuan Zhou)
- osdc: Fix race condition with tick_event and shutdown ([issue#14256](#), [pr#7151](#), Adam C. Emerson)

- osd: check health state before pre_booting ([issue#14181](#), [pr#7053](#), Xiaoxi Chen)
- osd: clear pg_stat_queue after stopping pgs ([issue#14212](#), [pr#7091](#), Sage Weil)
- osd: delay populating in-memory PG log hashmaps ([pr#6425](#), Piotr Dałek)
- osd: disable filestore_xfs_extsize by default ([issue#14397](#), [pr#7265](#), Ken Dreyer)
- osd: do not keep ref of old osdmap in pg ([issue#13990](#), [pr#7007](#), Kefu Chai)
- osd: drop deprecated removal pg type ([pr#6970](#), Igor Podoski)
- osd: FileJournal: fix return code of create method ([issue#14134](#), [pr#6988](#), xie xingguo)
- osd: FileJournal: support batch peak and pop from writeq ([pr#6701](#), Xinze Chi)
- osd: FileStore: conditional collection of drive metadata ([pr#6956](#), Somnath Roy)
- osd: FileStore:: optimize lfn_unlink ([pr#6649](#), Jianpeng Ma)
- osd: fix null pointer access and race condition ([issue#14072](#), [pr#6916](#), xie xingguo)
- osd: fix scrub start hobject ([pr#7467](#), Sage Weil)
- osd: fix sparse-read result code checking logic ([issue#14151](#), [pr#7016](#), xie xingguo)
- osd: fix temp object removal after upgrade ([issue#13862](#), [pr#6976](#), David Zafman)
- osd: fix wip (l_osd_op_wip) perf counter and remove repop_map ([pr#7077](#), Xinze Chi)
- osd: fix wrongly placed assert and some cleanups ([pr#6766](#), xiexingguo, xie xingguo)
- osd: KeyValueStore: fix return code of mkfs ([pr#7036](#), xie xingguo)
- osd: KeyValueStore: fix wrongly placed assert ([issue#14176](#), [issue#14178](#), [pr#7047](#), xie xingguo)
- osd: kstore: several small fixes ([issue#14351](#), [issue#14352](#), [pr#7213](#), xie xingguo)
- osd: kstore: small fixes to kstore ([issue#14204](#), [pr#7095](#), xie xingguo)
- osd: make list_missing query missing_loc.needs_recovery_map ([pr#6298](#), Guang Yang)
- osdmap: remove unused local variables ([pr#6864](#), luo kexue)
- osd: memstore: fix two bugs ([pr#6963](#), Casey Bodley, Sage Weil)
- osd: misc FileStore fixes ([issue#14192](#), [issue#14188](#), [issue#14194](#), [issue#14187](#), [issue#14186](#), [pr#7059](#), xie xingguo)
- osd: misc optimization for map utilization ([pr#6950](#), Ning Yao)
- osd,mon: log leveldb and rocksdb to ceph log ([pr#6921](#), Sage Weil)
- osd: Omap small bugs adapted ([pr#6669](#), Jianpeng Ma, David Zafman)
- osd: optimize the session_handle_reset function ([issue#14182](#), [pr#7054](#), songbaisen)
- osd: OSDService: Fix typo in osdmap comment ([pr#7275](#), Brad Hubbard)
- osd: os: skip checking pg_meta object existance in FileStore ([pr#6870](#), Ning Yao)
- osd: PGLog: clean up read_log ([pr#7092](#), Jie Wang)
- osd: prevent osd_recovery_sleep from causing recovery-thread suicide ([pr#7065](#), Jianpeng Ma)
- osd: reduce string use in coll_t::calc_str() ([pr#6505](#), Igor Podoski)
- osd: release related sources when scrub is interrupted ([pr#6744](#), Jianpeng Ma)
- osd: remove unused OSDMap::set_weightf() ([issue#14369](#), [pr#7231](#), huanwen ren)
- osd: ReplicatedPG: clean up unused function ([pr#7211](#), Xiaowei Chen)
- osd/ReplicatedPG: fix promotion recency logic ([issue#14320](#), [pr#6702](#), Sage Weil)
- osd: several small cleanups ([pr#7055](#), xie xingguo)
- osd: shut down if we flap too many times in a short period ([pr#6708](#), Xiaoxi Chen)
- osd: skip promote for writefull w/ FADVISE_DONTNEED/NOCACHE ([pr#7010](#), Jianpeng Ma)
- osd: small fixes to memstore ([issue#14228](#), [issue#14229](#), [issue#14227](#), [pr#7107](#), xie xingguo)
- osd: try evicting after flushing is done ([pr#5630](#), Zhiqiang Wang)
- osd: use atomic to generate ceph_tid ([pr#7017](#), Evgeniy Firsov)
- osd: use optimized is_zero in object_stat_sum_t.is_zero() ([pr#7203](#), Piotr Dałek)
- osd: utime_t, eversion_t, osd_stat_sum_t encoding optimization ([pr#6902](#), Xinze Chi)
- pybind: add ceph_volume_client interface for Manila and similar frameworks ([pr#6205](#), John Spray)
- pybind: fix build failure, remove extraneous semicolon in method ([issue#14371](#), [pr#7235](#), Abhishek Lekshmanan)
- pybind/test_rbd: fix test_create_defaults ([issue#14279](#), [pr#7155](#), Josh Durgin)
- qa: disable rbd/qemu-iotests test case 055 on RHEL/CentOSlibrbd: journal replay should honor inter-event dependencies ([issue#14385](#), [pr#7272](#), Jason Dillaman)
- qa/workunits: merge_diff shouldn't attempt to use striping ([issue#14165](#), [pr#7041](#), Jason Dillaman)
- qa/workunits/snaps: move snap tests into fs sub-directory ([pr#6496](#), Yan, Zheng)
- rados: implement rm -force option to force remove when full ([pr#6202](#), Xiaowei Chen)
- rbd: additional validation for striping parameters ([pr#6914](#), Na Xie)
- rbd: add pool name to disambiguate rbd admin socket commands ([pr#6904](#), wuxiangwei)
- rbd: correct an output string for merge-diff ([pr#7046](#), Kongming Wu)
- rbd: fix static initialization ordering issues ([pr#6978](#), Mykola Golub)
- rbd-fuse: image name can not include snap name ([pr#7044](#), Yongqiang He)
- rbd-fuse: implement mv operation ([pr#6938](#), wuxiangwei)
- rbd: must specify both of stripe-unit and stripe-count when specifying stripingv2 feature ([pr#7026](#), Donghai Xu)
- rbd-nbd: add copyright ([pr#7166](#), Li Wang)
- rbd-nbd: fix up return code handling ([pr#7215](#), Mykola Golub)
- rbd-nbd: small improvements in logging and forking ([pr#7127](#), Mykola Golub)
- rbd: rbd order will be place in 22, when set to 0 in the config_opt ([issue#14139](#), [issue#14047](#), [pr#6886](#), huanwen ren)
- rbd: striping parameters should support 64bit integers ([pr#6942](#), Na Xie)

- rbd: use default order from configuration when not specified ([pr#6965](#), Yunchuan Wen)
- rgw: add a method to purge all associate keys when removing a subuser ([issue#12890](#), [pr#6002](#), Sangdi Xu)
- rgw: add missing error code for admin op API ([pr#7037](#), Dunrong Huang)
- rgw: add support for “end_marker” parameter for GET on Swift account. ([issue#10682](#), [pr#4216](#), Radoslaw Zarzynski)
- rgw_admin: orphans finish segfaults ([pr#6652](#), Igor Fedotov)
- rgw: content length ([issue#13582](#), [pr#6975](#), Yehuda Sadeh)
- rgw: delete default zone ([pr#7005](#), YankunLi)
- rgw: do not abort radowgw server when using admin op API with bad parameters ([issue#14190](#), [issue#14191](#), [pr#7063](#), Dunrong Huang)
- rgw: Drop a debugging message ([pr#7280](#), Pete Zaitcev)
- rgw: fix a typo in init-radosgw ([pr#6817](#), Zhi Zhang)
- rgw: fix compilation warning ([pr#7160](#), Yehuda Sadeh)
- rgw: fix wrong check for parse() return ([pr#6797](#), Dunrong Huang)
- rgw: let radosgw-admin bucket stats return a standard json ([pr#7029](#), Ruifeng Yang)
- rgw: modify command stucking when operating radosgw-admin metadata list user ([pr#7032](#), Peiyang Liu)
- rgw: modify documents and help infos’ descriptions to the usage of option date when executing command “log show” ([pr#6080](#), Kongming Wu)
- rgw: Parse -subuser better ([pr#7279](#), Pete Zaitcev)
- rgw: radosgw-admin bucket check -fix not work ([pr#7093](#), Weijun Duan)
- rgw: warn on suspicious civetweb frontend parameters ([pr#6944](#), Matt Benjamin)
- rocksdb: remove rdb sources from dist tarball ([issue#13554](#), [pr#7105](#), Venky Shankar)
- stringify outputted error code and fix unmatched parentheses. ([pr#6998](#), xie.xingguo, xie xingguo)
- test/librbd/fsx: Use c++11 std::mt19937 generator instead of random_r() ([pr#6332](#), John Coyle)
- test/mon/osd-erasure-code-profile: pick new mon port ([pr#7161](#), Sage Weil)
- tests: add const for ec test ([pr#6911](#), Michal Jarzabek)
- tests: configure with rocksdb by default ([issue#14220](#), [pr#7100](#), Loic Dachary)
- tests: Fix for make check. ([pr#7102](#), David Zafman)
- tests: notification slave needs to wait for master ([issue#13810](#), [pr#7220](#), Jason Dillaman)
- tests: snap rename and rebuild object map in client update test ([pr#7224](#), Jason Dillaman)
- tests: unittest_bufferlist: fix hexdump test ([pr#7152](#), Sage Weil)
- tests: unittest_ipaddr: fix segv ([pr#7154](#), Sage Weil)
- tools: ceph_monstore_tool: add inflate-pgmap command ([issue#14217](#), [pr#7097](#), Kefu Chai)
- tools: monstore: add ‘show-versions’ command. ([pr#7073](#), Cilang Zhao)

V10.0.2

This development release includes a raft of changes and improvements for Jewel. Key additions include CephFS scrub/repair improvements, an AIX and Solaris port of librados, many librbd journaling additions and fixes, extended per-pool options, and NBD driver for RBD (rbd-nbd) that allows librbd to present a kernel-level block device on Linux, multitenancy support for RGW, RGW bucket lifecycle support, RGW support for Swift static large objects (SLO), and RGW support for Swift bulk delete.

There are also lots of smaller optimizations and performance fixes going in all over the tree, particular in the OSD and common code.

NOTABLE CHANGES

- auth: fail if rotating key is missing (do not spam log) ([pr#6473](#), Qiankun Zheng)
- auth: fix crash when bad keyring is passed ([pr#6698](#), Dunrong Huang)
- auth: make keyring without mon entity type return -EACCES ([pr#5734](#), Xiaowei Chen)
- buffer: make usable outside of ceph source again ([pr#6863](#), Josh Durgin)
- build: cmake check fixes ([pr#6787](#), Orit Wasserman)
- build: fix bz2-dev dependency ([pr#6948](#), Samuel Just)
- build: Gentoo: _FORTIFY_SOURCE fix. ([issue#13920](#), [pr#6739](#), Robin H. Johnson)
- build/ops: systemd ceph-disk unit must not assume /bin/flock ([issue#13975](#), [pr#6803](#), Loic Dachary)
- ceph-detect-init: Ubuntu >= 15.04 uses systemd ([pr#6873](#), James Page)
- cephfs-data-scan: scan_frags ([pr#5941](#), John Spray)
- cephfs-data-scan: scrub tag filtering (#12133 and #12145) ([issue#12133](#), [issue#12145](#), [pr#5685](#), John Spray)
- ceph-fuse: add process to ceph-fuse -help ([pr#6821](#), Wei Feng)
- ceph-kvstore-tool: handle bad out file on command line ([pr#6093](#), Kefu Chai)
- ceph-mds: add -help/-h ([pr#6850](#), Cilang Zhao)
- ceph_objectstore_bench: fix race condition, bugs ([issue#13516](#), [pr#6681](#), Igor Fedotov)
- ceph.spec.in: add BuildRequires: systemd ([issue#13860](#), [pr#6692](#), Nathan Cutler)
- client: a better check for MDS availability ([pr#6253](#), John Spray)

- client: close mds sessions in shutdown() ([pr#6269](#), John Spray)
- client: don't invalidate page cache when inode is no longer used ([pr#6380](#), Yan, Zheng)
- client: modify a word in log ([pr#6906](#), YongQiang He)
- cls/cls_rbd.cc: fix misused metadata_name_from_key ([issue#13922](#), [pr#6661](#), Xiaoxi Chen)
- cmake: Add common/PluginRegistry.cc to CMakeLists.txt ([pr#6805](#), Pete Zaitcev)
- cmake: add rgw_basic_types.cc to librgw.a ([pr#6786](#), Orit Wasserman)
- cmake: add TracepointProvider.cc to libcommon ([pr#6823](#), Orit Wasserman)
- cmake: define STRERROR_R_CHAR_P for GNU-specific strerror_r ([pr#6751](#), Ilya Dryomov)
- cmake: update for recent librbd changes ([pr#6715](#), John Spray)
- cmake: update for recent rbd changes ([pr#6818](#), Mykola Golub)
- common: add generic plugin infrastructure ([pr#6696](#), Sage Weil)
- common: add latency perf counter for finisher ([pr#6175](#), Xinze Chi)
- common: buffer: add cached_crc and cached_crc_adjust counts to perf dump ([pr#6535](#), Ning Yao)
- common: buffer: remove unneeded list destructor ([pr#6456](#), Michal Jarzabek)
- common/ceph_context.cc:fix order of initialisers ([pr#6838](#), Michal Jarzabek)
- common: don't reverse hobject_t hash bits when zero ([pr#6653](#), Piotr Dałek)
- common: log: Assign LOG_DEBUG priority to syslog calls ([issue#13993](#), [pr#6815](#), Brad Hubbard)
- common: log: predict log message buffer allocation size ([pr#6641](#), Adam Kupczyk)
- common: optimize debug logging code ([pr#6441](#), Adam Kupczyk)
- common: perf counter for bufferlist history total alloc ([pr#6198](#), Xinze Chi)
- common: reduce CPU usage by making stringstream in stringify function thread local ([pr#6543](#), Evgeniy Firsov)
- common: re-enable backtrace support ([pr#6771](#), Jason Dillaman)
- common: SubProcess: fix multiple definition bug ([pr#6790](#), Yunchuan Wen)
- common: use namespace instead of subclasses for buffer ([pr#6686](#), Michal Jarzabek)
- configure.ac: macro fix ([pr#6769](#), Igor Podoski)
- doc: admin/build-doc: add lxml dependencies on debian ([pr#6610](#), Ken Dreyer)
- doc/cephfs posix: update ([pr#6922](#), Sage Weil)
- doc: CodingStyle: fix broken URLs ([pr#6733](#), Kefu Chai)
- doc: correct typo 'restared' to 'restarted' ([pr#6734](#), Yilong Zhao)
- doc/dev/index: refactor/reorg ([pr#6792](#), Nathan Cutler)
- doc/dev/index.rst: begin writing Contributing to Ceph ([pr#6727](#), Nathan Cutler)
- doc/dev/index.rst: fix headings ([pr#6780](#), Nathan Cutler)
- doc: dev: introduction to tests ([pr#6910](#), Loic Dachary)
- doc: file must be empty when writing layout fields of file use "setfattr" ([pr#6848](#), Cilang Zhao)
- doc: Fixed incorrect name of a "List Multipart Upload Parts" Response Entity ([issue#14003](#), [pr#6829](#), Lenz Grimmer)
- doc: Fixes a spelling error ([pr#6705](#), Jeremy Qian)
- doc: fix typo in cephfs/quota ([pr#6745](#), Drunkard Zhang)
- doc: fix typo in developer guide ([pr#6943](#), Nathan Cutler)
- doc: INSTALL redirect to online documentation ([pr#6749](#), Loic Dachary)
- doc: little improvements for troubleshooting scrub issues ([pr#6827](#), Mykola Golub)
- doc: Modified a note section in rbd-snapshot doc. ([pr#6908](#), Nilamdyuti Goswami)
- doc: note that cephfs auth stuff is new in jewel ([pr#6858](#), John Spray)
- doc: osd: s/schedued/scheduled/ ([pr#6872](#), Loic Dachary)
- doc: remove unnecessary period in headline ([pr#6775](#), Marc Koderer)
- doc: rst style fix for pools document ([pr#6816](#), Drunkard Zhang)
- doc: Update list of admin/build-doc dependencies ([issue#14070](#), [pr#6934](#), Nathan Cutler)
- init-ceph: do umount when the path exists. ([pr#6866](#), Xiaoxi Chen)
- journal: disconnect watch after watch error ([issue#14168](#), [pr#7113](#), Jason Dillaman)
- journal: fire replay complete event after reading last object ([issue#13924](#), [pr#6762](#), Jason Dillaman)
- journal: support replaying beyond skipped splay objects ([pr#6687](#), Jason Dillaman)
- librados: aix gcc librados port ([pr#6675](#), Rohan Mars)
- librados: avoid malloc(0) (which can return NULL on some platforms) ([issue#13944](#), [pr#6779](#), Dan Mick)
- librados: clean up Objecter.h ([pr#6731](#), Jie Wang)
- librados: include/rados/librados.h: fix typo ([pr#6741](#), Nathan Cutler)
- librbd: automatically flush IO after blocking write operations ([issue#13913](#), [pr#6742](#), Jason Dillaman)
- librbd: better handling of exclusive lock transition period ([pr#7204](#), Jason Dillaman)
- librbd: check for presence of journal before attempting to remove ([issue#13912](#), [pr#6737](#), Jason Dillaman)
- librbd: clear error when older OSD doesn't support image flags ([issue#14122](#), [pr#7035](#), Jason Dillaman)
- librbd: correct include guard in RenameRequest.h ([pr#7143](#), Jason Dillaman)
- librbd: correct issues discovered during teuthology testing ([issue#14108](#), [issue#14107](#), [pr#6974](#), Jason Dillaman)
- librbd: correct issues discovered when cache is disabled ([issue#14123](#), [pr#6979](#), Jason Dillaman)
- librbd: correct race conditions discovered during unit testing ([issue#14060](#), [pr#6923](#), Jason Dillaman)
- librbd: disable copy-on-read when not exclusive lock owner ([issue#14167](#), [pr#7129](#), Jason Dillaman)
- librbd: do not ignore self-managed snapshot release result ([issue#14170](#), [pr#7043](#), Jason Dillaman)
- librbd: ensure copy-on-read requests are complete prior to closing parent image ([pr#6740](#), Jason Dillaman)

- librbd: ensure librados callbacks are flushed prior to destroying ([issue#14092](#), [pr#7040](#), Jason Dillaman)
- librbd: fix journal iohint ([pr#6917](#), Jianpeng Ma)
- librbd: fix known test case race condition failures ([issue#13969](#), [pr#6800](#), Jason Dillaman)
- librbd: fix merge-diff for >2GB diff-files ([issue#14030](#), [pr#6889](#), Yunchuan Wen)
- librbd: fix test case race condition for journaling ops ([pr#6877](#), Jason Dillaman)
- librbd: fix tracepoint parameter in diff_iterate ([pr#6892](#), Yunchuan Wen)
- librbd: image refresh code paths converted to async state machines ([pr#6859](#), Jason Dillaman)
- librbd: include missing header for bool type ([pr#6798](#), Mykola Golub)
- librbd: initial collection of state machine unit tests ([pr#6703](#), Jason Dillaman)
- librbd: integrate journaling for maintenance operations ([pr#6625](#), Jason Dillaman)
- librbd: journaling-related lock dependency cleanup ([pr#6777](#), Jason Dillaman)
- librbd: not necessary to hold owner_lock while releasing snap id ([issue#13914](#), [pr#6736](#), Jason Dillaman)
- librbd: only send signal when AIO completions queue empty ([pr#6729](#), Jianpeng Ma)
- librbd: optionally validate new RBD pools for snapshot support ([issue#13633](#), [pr#6925](#), Jason Dillaman)
- librbd: partial revert of commit 9b0e359 ([issue#13969](#), [pr#6789](#), Jason Dillaman)
- librbd: properly handle replay of snap remove RPC message ([issue#14164](#), [pr#7042](#), Jason Dillaman)
- librbd: reduce verbosity of common error condition logging ([issue#14234](#), [pr#7114](#), Jason Dillaman)
- librbd: simplify IO method signatures for 32bit environments ([pr#6700](#), Jason Dillaman)
- librbd: support eventfd for AIO completion notifications ([pr#5465](#), Haomai Wang)
- mailmap: add UMCLOUD affiliation ([pr#6820](#), Jiaying Ren)
- mailmap: Jewel updates ([pr#6750](#), Abhishek Lekshmanan)
- makefiles: remove bz2-dev from dependencies ([issue#13981](#), [pr#6939](#), Piotr Dałek)
- mds: add 'p' flag in auth caps to control setting pool in layout ([pr#6567](#), John Spray)
- mds: fix client capabilities during reconnect (client.XXXX isn't responding to mclientcaps(revoke)) ([issue#11482](#), [pr#6432](#), Yan, Zheng)
- mds: fix setxattr (broken in a536d114) ([issue#14029](#), [pr#6941](#), John Spray)
- mds: repair the command option “hot-standby” ([pr#6454](#), Wei Feng)
- mds: tear down connections from tell commands ([issue#14048](#), [pr#6933](#), John Spray)
- mon: fix ceph df pool available calculation for 0-weighted OSDs ([pr#6660](#), Chengyuan Li)
- mon: fix routed_request_tids leak ([pr#6102](#), Ning Yao)
- mon: support min_down_reporter by subtree level (default by host) ([pr#6709](#), Xiaoxi Chen)
- mount.ceph: memory leaks ([pr#6905](#), Qiankun Zheng)
- osd: add osd op queue latency perfcounter ([pr#5793](#), Haomai Wang)
- osd: Allow repair of history.last_epoch_started using config ([pr#6793](#), David Zafman)
- osd: avoid duplicate op->mark_started in ReplicatedBackend ([pr#6689](#), Jacek J. Łakis)
- osd: cancel failure reports if we fail to rebind network ([pr#6278](#), Xinze Chi)
- osd: correctly handle small osd_scrub_interval_randomize_ratio ([pr#7147](#), Samuel Just)
- osd: defer decoding of MOSDRepOp/MOSDRepOpReply ([pr#6503](#), Xinze Chi)
- osd: don't update epoch and rollback_info objects attrs if there is no need ([pr#6555](#), Ning Yao)
- osd: dump number of missing objects for each peer with pg query ([pr#6058](#), Guang Yang)
- osd: enable perfcounters on sharded work queue mutexes ([pr#6455](#), Jacek J. Łakis)
- osd: FileJournal: reduce locking scope in write_aio_bl ([issue#12789](#), [pr#5670](#), Zhi Zhang)
- osd: FileStore: remove __SWORD_TYPE dependency ([pr#6263](#), John Coyle)
- osd: fix FileStore::_destroy_collection error return code ([pr#6612](#), Ruifeng Yang)
- osd: fix incorrect throttle in WBThrottle ([pr#6713](#), Zhang Huan)
- osd: fix MOSDRepScrub reference counter in replica_scrub ([pr#6730](#), Jie Wang)
- osd: fix rollback_info_trimmed_to before index() ([issue#13965](#), [pr#6801](#), Samuel Just)
- osd: fix trivial scrub bug ([pr#6533](#), Li Wang)
- osd: KeyValueStore: don't queue NULL context ([pr#6783](#), Haomai Wang)
- osd: make backend and block device code a bit more generic ([pr#6759](#), Sage Weil)
- osd: move newest decode version of MOSDOp and MOSDOpReply to the front ([pr#6642](#), Jacek J. Łakis)
- osd: pg_pool_t: add dictionary for pool options ([issue#13077](#), [pr#6081](#), Mykola Golub)
- osd: reduce memory consumption of some structs ([pr#6475](#), Piotr Dałek)
- osd: release the message throttle when OpRequest unregistered ([issue#14248](#), [pr#7148](#), Samuel Just)
- osd: remove __SWORD_TYPE dependency ([pr#6262](#), John Coyle)
- osd: slightly reduce actual size of pg_log_entry_t ([pr#6690](#), Piotr Dałek)
- osd: support pool level recovery_priority and recovery_op_priority ([pr#5953](#), Guang Yang)
- osd: use pg id (without shard) when referring the PG ([pr#6236](#), Guang Yang)
- packaging: add build dependency on python-devel package ([pr#7205](#), Josh Durgin)
- pybind/cephfs: add symlink and its unit test ([pr#6323](#), Shang Ding)
- pybind: decode empty string in conf_parse_argv() correctly ([pr#6711](#), Josh Durgin)
- pybind: Implementation of rados_ioctx_snapshot_rollback ([pr#6878](#), Florent Manens)
- pybind: port the rbd bindings to Cython ([issue#13115](#), [pr#6768](#), Hector Martin)
- pybind: support ioctx:exec ([pr#6795](#), Noah Watkins)
- qa: erasure-code benchmark plugin selection ([pr#6685](#), Loic Dachary)

- qa/krbd: Expunge generic/247 ([pr#6831](#), Douglas Fuller)
- qa/workunits/cephtool/test.sh: false positive fail on /tmp/obj1. ([pr#6837](#), Robin H. Johnson)
- qa/workunits/cephtool/test.sh: no ./ ([pr#6748](#), Sage Weil)
- qa/workunits/rbd: rbd-nbd test should use sudo for map/unmap ops ([issue#14221](#), [pr#7101](#), Jason Dillaman)
- rados: bench: fix off-by-one to avoid writing past object_size ([pr#6677](#), Tao Chang)
- rbd: add -object-size option, deprecate -order ([issue#12112](#), [pr#6830](#), Vikhyat Umrao)
- rbd: add RBD pool mirroring configuration API + CLI ([pr#6129](#), Jason Dillaman)
- rbd: fix build with “-without-rbd” ([issue#14058](#), [pr#6899](#), Piotr Datek)
- rbd: journal: configuration via conf, cli, api and some fixes ([pr#6665](#), Mykola Golub)
- rbd: merge_diff test should use new -object-size parameter instead of -order ([issue#14106](#), [pr#6972](#), Na Xie, Jason Dillaman)
- rbd-nbd: network block device (NBD) support for RBD ([pr#6657](#), Yunchuan Wen, Li Wang)
- rbd: output formatter may not be closed upon error ([issue#13711](#), [pr#6706](#), xie xingguo)
- rgw: add a missing cap type ([pr#6774](#), Yehuda Sadeh)
- rgw: add an inspection to the field of type when assigning user caps ([pr#6051](#), Kongming Wu)
- rgw: add LifeCycle feature ([pr#6331](#), Ji Chen)
- rgw: add support for Static Large Objects of Swift API ([issue#12886](#), [issue#13452](#), [pr#6643](#), Yehuda Sadeh, Radoslaw Zarzynski)
- rgw: fix a glaring syntax error ([pr#6888](#), Pavan Rallabhandi)
- rgw: fix the build failure ([pr#6927](#), Kefu Chai)
- rgw: multitenancy support ([pr#6784](#), Yehuda Sadeh, Pete Zaitcev)
- rgw: Remove unused code in PutMetadataAccount:execute ([pr#6668](#), Pete Zaitcev)
- rgw: remove unused variable in RGWPutMetadataBucket::execute ([pr#6735](#), Radoslaw Zarzynski)
- rgw/rgw_resolve: fallback to res_query when res_nquery not implemented ([pr#6292](#), John Coyle)
- rgw: static large objects (Radoslaw Zarzynski, Yehuda Sadeh)
- rgw: swift bulk delete (Radoslaw Zarzynski)
- systemd: start/stop/restart ceph services by daemon type ([issue#13497](#), [pr#6276](#), Zhi Zhang)
- sysvinit: allow custom cluster names ([pr#6732](#), Richard Chan)
- test/encoding/readable.sh fix ([pr#6714](#), Igor Podoski)
- test: fix osd-scrub-snaps.sh ([pr#6697](#), Xinze Chi)
- test/librados/test.cc: clean up EC pools’ crush rules too ([issue#13878](#), [pr#6788](#), Loic Dachary, Dan Mick)
- tests: allow object corpus readable test to skip specific incompat instances ([pr#6932](#), Igor Podoski)
- tests: ceph-helpers assert success getting backfills ([pr#6699](#), Loic Dachary)
- tests: ceph_test_keyvaluedb_iterators: fix broken test ([pr#6597](#), Haomai Wang)
- tests: fix failure for osd-scrub-snap.sh ([issue#13986](#), [pr#6890](#), Loic Dachary, Ning Yao)
- tests: fix race condition testing auto scrub ([issue#13592](#), [pr#6724](#), Xinze Chi, Loic Dachary)
- tests: flush op work queue prior to destroying MockImageCtx ([issue#14092](#), [pr#7002](#), Jason Dillaman)
- tests: -osd-scrub-load-threshold=2000 for more consistency ([issue#14027](#), [pr#6871](#), Loic Dachary)
- tests: osd-scrub-snaps.sh to display full osd logs on error ([issue#13986](#), [pr#6857](#), Loic Dachary)
- test: use sequential journal_tid for object cache test ([issue#13877](#), [pr#6710](#), Josh Durgin)
- tools: add cephfs-table-tool ‘take_inos’ ([pr#6655](#), John Spray)
- tools: Fix layout handing in cephfs-data-scan (#13898) ([pr#6719](#), John Spray)
- tools: support printing part cluster map in readable fashion ([issue#13079](#), [pr#5921](#), Bo Cai)
- vstart.sh: add mstart, mstop, mrun wrappers for running multiple vstart-style test clusters out of src tree ([pr#6901](#), Yehuda Sadeh)

V10.0.1

This is the second development release for the Jewel cycle. Highlights include some KeyValueDB interface optimizations, initial journaling support in librbd, MDS scrubbing progress, and a bunch of OSD optimizations and improvements (removal of an unnecessary CRUSH calculation in the IO path, sharding for FileStore completions, improved temperature calculation for cache tiering, fixed randomization of scrub times, and various optimizations).

NOTABLE CHANGES

- build: build internal plugins and classes as modules ([pr#6462](#), James Page)
- build: fix Jenkins make check errors due to deep-scrub randomization ([pr#6671](#), David Zafman)
- build/ops: enable CR in CentOS 7 ([issue#13997](#), [pr#6844](#), Loic Dachary)
- build/ops: rbd-replay moved from ceph-test-dbg to ceph-common-dbg ([issue#13785](#), [pr#6578](#), Loic Dachary)
- ceph-disk: Add destroy and deactivate option ([issue#7454](#), [pr#5867](#), Vicente Cheng)
- ceph-disk: compare parted output with the dereferenced path ([issue#13438](#), [pr#6219](#), Joe Julian)
- ceph-objectstore-tool: fix -dry-run for many ceph-objectstore-tool operations ([pr#6545](#), David Zafman)
- ceph.spec.in: limit _smp_mflags when lowmem_builder is set in SUSE’s OBS ([issue#13858](#), [pr#6691](#), Nathan Cutler)

- ceph_test_msgr: Use send_message instead of keepalive to wakeup connection ([pr#6605](#), Haomai Wang)
- client: avoid creating orphan object in Client::check_pool_perm() ([issue#13782](#), [pr#6603](#), Yan, Zheng)
- client: use null snapc to check pool permission ([issue#13714](#), [pr#6497](#), Yan, Zheng)
- cmake: add nss as a suffix for pk11pub.h ([pr#6556](#), Samuel Just)
- cmake: fix files list ([pr#6539](#), Yehuda Sadeh)
- cmake: librbd and libjournal build fixes ([pr#6557](#), Ilya Dryomov)
- coc: fix typo in the apt-get command ([pr#6659](#), Chris Holcombe)
- common: allow enable/disable of optracker at runtime ([pr#5168](#), Jianpeng Ma)
- common: fix reset max in Throttle using perf reset command ([issue#13517](#), [pr#6300](#), Xinze Chi)
- doc: add v0.80.11 to the release timeline ([pr#6658](#), Loic Dachary)
- doc: release-notes: draft v0.80.11 release notes ([pr#6374](#), Loic Dachary)
- doc: release-notes: draft v10.0.0 release notes ([pr#6666](#), Loic Dachary)
- doc: SubmittingPatches: there is no next; only jewel ([pr#6811](#), Nathan Cutler)
- doc: Update ceph-disk manual page with new feature deactivate/destroy. ([pr#6637](#), Vicente Cheng)
- doc: update infernalis release notes ([pr#6575](#), vasukulkarni)
- fix: use right init_flags to finish CephContext ([pr#6549](#), Yunchuan Wen)
- init-ceph: use getopt to make option processing more flexible ([issue#3015](#), [pr#6089](#), Nathan Cutler)
- journal: incremental improvements and fixes ([pr#6552](#), Mykola Golub)
- krbd: remove deprecated -quiet param from udevadm ([issue#13560](#), [pr#6394](#), Jason Dillaman)
- kv: fix bug in kv key optimization ([pr#6511](#), Sage Weil)
- kv/KineticStore: fix broken split_key ([pr#6574](#), Haomai Wang)
- kv: optimize and clean up internal key/value interface ([pr#6312](#), Piotr Dałek, Sage Weil)
- librados: do cleanup ([pr#6488](#), xie xingguo)
- librados: fix pool alignment API overflow issue ([issue#13715](#), [pr#6489](#), xie xingguo)
- librados: fix potential null pointer access when do pool_snap_list ([issue#13639](#), [pr#6422](#), xie xingguo)
- librados: fix PromoteOn2ndRead test for EC ([pr#6373](#), Sage Weil)
- librados: fix rare race where pool op callback may hang forever ([issue#13642](#), [pr#6426](#), xie xingguo)
- librados: Solaris port ([pr#6416](#), Rohan Mars)
- librbd: flush and invalidate cache via admin socket ([issue#2468](#), [pr#6453](#), Mykola Golub)
- librbd: integrate journaling support for IO operations ([pr#6541](#), Jason Dillaman)
- librbd: perf counters might not be initialized on error ([issue#13740](#), [pr#6523](#), Jason Dillaman)
- librbd: perf section name: use hyphen to separate components ([issue#13719](#), [pr#6516](#), Mykola Golub)
- librbd: resize should only update image size within header ([issue#13674](#), [pr#6447](#), Jason Dillaman)
- librbd: start perf counters after id is initialized ([issue#13720](#), [pr#6494](#), Mykola Golub)
- mailmap: revise organization ([pr#6519](#), Li Wang)
- mailmap: Ubuntu Kylin name changed to Kylin Cloud ([pr#6532](#), Loic Dachary)
- mailmap: update .organizationmap ([pr#6565](#), chenji-kael)
- mailmap: updates for infernalis. ([pr#6495](#), Yann Dupont)
- mailmap: updates ([pr#6594](#), chenji-kael)
- mds: fix scrub_path ([pr#6684](#), John Spray)
- mds: properly set STATE_STRAY/STATE_ORPHAN for stray dentry/inode ([issue#13777](#), [pr#6553](#), Yan, Zheng)
- mds: ScrubStack and “tag path” command ([pr#5662](#), Yan, Zheng, John Spray, Greg Farnum)
- mon: block ‘ceph osd pg-temp ...’ if pg_temp update is already pending ([pr#6704](#), Sage Weil)
- mon: don’t require OSD W for MRemoveSnaps ([issue#13777](#), [pr#6601](#), John Spray)
- mon: initialize recorded election epoch properly even when standalone ([issue#13627](#), [pr#6407](#), huanwen ren)
- mon: revert MonitorDBStore’s WholeStoreIteratorImpl::get ([issue#13742](#), [pr#6522](#), Piotr Dałek)
- msg/asynchronous: let receiver ack message ASAP ([pr#6478](#), Haomai Wang)
- msg/asynchronous: support of non-block connect in async messenger ([issue#12802](#), [pr#5848](#), Jianhui Yuan)
- msg/asynchronous: will crash if enabling async msg because of an assertion ([pr#6640](#), Zhi Zhang)
- osd: avoid calculating crush mapping for most ops ([pr#6371](#), Sage Weil)
- osd: avoid double-check for replaying and can_checkpoint() in FileStore::_check_replay_guard ([pr#6471](#), Ning Yao)
- osd: call on_new_interval on newly split child PG ([issue#13962](#), [pr#6778](#), Sage Weil)
- osd: change mutex to spinlock to optimize thread context switch. ([pr#6492](#), Xiaowei Chen)
- osd: check do_shutdown before do_restart ([pr#6547](#), Xiaoxi Chen)
- osd: clarify the scrub result report ([pr#6534](#), Li Wang)
- osd: don’t do random deep scrubs for user initiated scrubs ([pr#6673](#), David Zafman)
- osd: FileStore: support multiple ondisk finish and apply finishers ([pr#6486](#), Xinze Chi, Haomai Wang)
- osd: fix broken balance / localized read handling ([issue#13491](#), [pr#6364](#), Jason Dillaman)
- osd: fix bug in last_* PG state timestamps ([pr#6517](#), Li Wang)
- osd: fix ClassHandler::ClassData::get_filter() ([pr#6747](#), Yan, Zheng)
- osd: fixes for several cases where op result code was not checked or set ([issue#13566](#), [pr#6347](#), xie xingguo)
- osd: fix reactivate (check OSDSuperblock in mkfs() when we already have the superblock) ([issue#13586](#), [pr#6385](#), Vicente Cheng)
- osd: fix wrong use of right parenthesis in localized read logic ([pr#6566](#), Jie Wang)
- osd: improve temperature calculation for cache tier agent ([pr#4737](#), MingXin Liu)

- osd: merge local_t and op_t txn to single one ([pr#6439](#), Xinze Chi)
- osd: newstore: misc updates (including kv and os/fs stuff) ([pr#6609](#), Sage Weil)
- osd: note down the number of missing clones ([pr#6654](#), Kefu Chai)
- osd: optimize clone write path if object-map is enabled ([pr#6403](#), xinxin shu)
- osd: optimize scrub subset_last_update calculation ([pr#6518](#), Li Wang)
- osd: partial revert of “ReplicatedPG: result code not correctly set in some cases.” ([issue#13796](#), [pr#6622](#), Sage Weil)
- osd: randomize deep scrubbing ([pr#6550](#), Dan van der Ster, Herve Rousseau)
- osd: scrub: do not assign value if read error ([pr#6568](#), Li Wang)
- osd: write file journal optimization ([pr#6484](#), Xinze Chi)
- rbd: accept -user, refuse -i command-line optionals ([pr#6590](#), Ilya Dryomov)
- rbd: add missing command aliases to refactored CLI ([issue#13806](#), [pr#6606](#), Jason Dillaman)
- rbd: dynamically generated bash completion ([issue#13494](#), [pr#6316](#), Jason Dillaman)
- rbd: fixes for refactored CLI and related tests ([pr#6738](#), Ilya Dryomov)
- rbd: make config changes actually apply ([pr#6520](#), Mykola Golub)
- rbd: refactor cli command handling ([pr#5987](#), Jason Dillaman)
- rbd: stripe unit/count set incorrectly from config ([pr#6593](#), Mykola Golub)
- rbd: support negative boolean command-line optionals ([issue#13784](#), [pr#6607](#), Jason Dillaman)
- rbd: unbreak rbd map + cephx_sign_messages option ([pr#6583](#), Ilya Dryomov)
- rgw: bucket request payment support ([issue#13427](#), [pr#6214](#), Javier M. Mellid)
- rgw: extend rgw_extended_http_attrs to affect Swift accounts and containers as well ([pr#5969](#), Radoslaw Zarzynski)
- rgw: fix openssl linkage ([pr#6513](#), Yehuda Sadeh)
- rgw: fix partial read issue in rgw_admin and rgw_tools ([pr#6761](#), Jiaying Ren)
- rgw: fix reload on non Debian systems. ([pr#6482](#), Hervé Rousseau)
- rgw: fix response of delete expired objects ([issue#13469](#), [pr#6228](#), Yuan Zhou)
- rgw: fix swift API returning incorrect account metadata ([issue#13140](#), [pr#6047](#), Sangdi Xu)
- rgw: link against system openssl (instead of dlopen at runtime) ([pr#6419](#), Sage Weil)
- rgw: prevent anonymous user from reading bucket with authenticated read ACL ([issue#13207](#), [pr#6057](#), root)
- rgw: use smart pointer for C_Reinitwatch ([pr#6767](#), Orit Wasserman)
- systemd: fix typos ([pr#6679](#), Tobias Suckow)
- tests: centos7 needs the Continuous Release (CR) Repository enabled for ([issue#13997](#), [pr#6842](#), Brad Hubbard)
- tests: concatenate test_rados_test_tool from src and qa ([issue#13691](#), [pr#6464](#), Loic Dachary)
- tests: fix test_rados_tools.sh rados lookup ([issue#13691](#), [pr#6502](#), Loic Dachary)
- tests: fix typo in TestClisRbd.snapshots test case ([issue#13727](#), [pr#6504](#), Jason Dillaman)
- tests: ignore test-suite.log ([pr#6584](#), Loic Dachary)
- tests: restore run-cli-tests ([pr#6571](#), Loic Dachary, Sage Weil, Jason Dillaman)
- tools/cephfs: fix overflow writing header to fixed size buffer (#13816) ([pr#6617](#), John Spray)

V10.0.0

This is the first development release for the Jewel cycle.

NOTABLE CHANGES

- build: cmake tweaks ([pr#6254](#), John Spray)
- build: more CMake package check fixes ([pr#6108](#), Daniel Gryniewicz)
- ceph-disk: get Nonetype when ceph-disk list with -format plain on single device. ([pr#6410](#), Vicente Cheng)
- ceph: fix tell behavior ([pr#6329](#), David Zafman)
- ceph-fuse: While starting ceph-fuse, start the log thread first ([issue#13443](#), [pr#6224](#), Wenjun Huang)
- client: don't mark_down on command reply ([pr#6204](#), John Spray)
- client: drop prefix from ints ([pr#6275](#), John Coyle)
- client: sys/file.h includes for flock operations ([pr#6282](#), John Coyle)
- cls_rbd: change object_map_update to return 0 on success, add logging ([pr#6467](#), Douglas Fuller)
- cmake: Use uname instead of arch. ([pr#6358](#), John Coyle)
- common: assert: __STRING macro is not defined by musl libc. ([pr#6210](#), John Coyle)
- common: fix OpTracker age histogram calculation ([pr#5065](#), Zhiqiang Wang)
- common/MemoryModel: Added explicit feature check for mallinfo(). ([pr#6252](#), John Coyle)
- common/obj_bencher.cc: fix verification crashing when there's no objects ([pr#5853](#), Piotr Dałek)
- common: optimize debug logging ([pr#6307](#), Adam Kupczyk)
- common: Thread: move copy constructor and assignment op ([pr#5133](#), Michal Jarzabek)
- common: WorkQueue: new PointerWQ base class for ContextWQ ([issue#13636](#), [pr#6525](#), Jason Dillaman)
- compat: use prefixed typeof extension ([pr#6216](#), John Coyle)
- crush: validate bucket id before indexing buckets array ([issue#13477](#), [pr#6246](#), Sage Weil)
- doc: download GPG key from download.ceph.com ([issue#13603](#), [pr#6384](#), Ken Dreyer)

- doc: fix outdated content in cache tier ([pr#6272](#), Yuan Zhou)
- doc/release-notes: v9.1.0 ([pr#6281](#), Loic Dachary)
- doc/releases-notes: fix build error ([pr#6483](#), Kefu Chai)
- doc: remove tocTree items under Create CephFS ([pr#6241](#), Jevon Qiao)
- doc: rename the “Create a Ceph User” section and add verbage about... ([issue#13502](#), [pr#6297](#), ritz303)
- docs: Fix styling of newly added mirror docs ([pr#6127](#), Wido den Hollander)
- doc, tests: update all <http://ceph.com/> to download.ceph.com ([pr#6435](#), Alfredo Deza)
- doc: update doc for with new pool settings ([pr#5951](#), Guang Yang)
- doc: update radosgw-admin example ([pr#6256](#), YankunLi)
- doc: update the OS recommendations for newer Ceph releases ([pr#6355](#), ritz303)
- drop envz.h includes ([pr#6285](#), John Coyle)
- libcephfs: Improve portability by replacing loff_t type usage with off_t ([pr#6301](#), John Coyle)
- libcephfs: only check file offset on glibc platforms ([pr#6288](#), John Coyle)
- librados: fix examples/librados/Makefile error. ([pr#6320](#), You Ji)
- librados: init crush_location from config file. ([issue#13473](#), [pr#6243](#), Wei Luo)
- librados: wrongly passed in argument for stat command ([issue#13703](#), [pr#6476](#), xie xingguo)
- librbd: deadlock while attempting to flush AIO requests ([issue#13726](#), [pr#6508](#), Jason Dillaman)
- librbd: fix enable objectmap feature issue ([issue#13558](#), [pr#6339](#), xinxin shu)
- librbd: remove duplicate read_only test in librbd::async_flatten ([pr#5856](#), runsisi)
- mailmap: modify member info ([pr#6468](#), Xiaowei Chen)
- mailmap: updates ([pr#6258](#), M Ranga Swami Reddy)
- mailmap: Xie Xingguo affiliation ([pr#6409](#), Loic Dachary)
- mds: implement snapshot rename ([pr#5645](#), xinxin shu)
- mds: messages/MOSDOp: cast in assert to eliminate warnings ([issue#13625](#), [pr#6414](#), David Zafman)
- mds: new filtered MDS tell commands for sessions ([pr#6180](#), John Spray)
- mds/Session: use projected parent for auth path check ([issue#13364](#), [pr#6200](#), Sage Weil)
- mon: should not set isvalid = true when cephx_verify_authorizer return false ([issue#13525](#), [pr#6306](#), Ruifeng Yang)
- osd: Add config option osd_read_ec_check_for_errors for testing ([pr#5865](#), David Zafman)
- osd: add pin/unpin support to cache tier (11066) ([pr#6326](#), Zhiqiang Wang)
- osd: auto repair EC pool ([issue#12754](#), [pr#6196](#), Guang Yang)
- osd: drop the interim set from load_pgs() ([pr#6277](#), Piotr Dałek)
- osd: FileJournal: _fdump wrongly returns if journal is currently unreadable. ([issue#13626](#), [pr#6406](#), xie xingguo)
- osd: FileStore: add a field indicate xattr only one chunk for set xattr. ([pr#6244](#), Jianpeng Ma)
- osd: FileStore: LFNIndex: remove redundant local variable ‘obj’. ([issue#13552](#), [pr#6333](#), xie xingguo)
- osd: FileStore: potential memory leak if _fgetattrs fails ([issue#13597](#), [pr#6377](#), xie xingguo)
- osd: FileStore: remove unused local variable ‘handle’ ([pr#6381](#), xie xingguo)
- osd: fix bogus scrub results when missing a clone ([issue#12738](#), [issue#12740](#), [pr#5783](#), David Zafman)
- osd: fix debug message in OSD::is_healthy ([pr#6226](#), Xiaozi Chen)
- osd: fix MOSDOp encoding ([pr#6174](#), Sage Weil)
- osd: init started to 0 ([issue#13206](#), [pr#6107](#), Sage Weil)
- osd: KeyValueStore: fix the name’s typo of keyvaluestore_default_strip_size ([pr#6375](#), Zhi Zhang)
- osd: new and delete ObjectStore::Transaction in a function is not necessary ([pr#6299](#), Ruifeng Yang)
- osd: optimize get_object_context ([pr#6305](#), Jianpeng Ma)
- osd: optimize MOSDOp/do_op/handle_op ([pr#5211](#), Jacek J. Lakis)
- osd: os/chain_xattr: On linux use linux/limits.h for XATTR_NAME_MAX. ([pr#6343](#), John Coyle)
- osd: reorder bool fields in PGLog struct ([pr#6279](#), Piotr Dałek)
- osd: ReplicatedPG: remove unused local variables ([issue#13575](#), [pr#6360](#), xie xingguo)
- osd: reset primary and up_primary when building a new past_interval. ([issue#13471](#), [pr#6240](#), xie xingguo)
- radosgw-admin: Checking the legality of the parameters ([issue#13018](#), [pr#5879](#), Qiankun Zheng)
- radosgw-admin: Create -secret-key alias for -secret ([issue#5821](#), [pr#5335](#), Yuan Zhou)
- radosgw-admin: metadata list user should return an empty list when user pool is empty ([issue#13596](#), [pr#6465](#), Orit Wasserman)
- rados: new options for write benchmark ([pr#6340](#), Joaquim Rocha)
- rbd: fix clone isssue ([issue#13553](#), [pr#6334](#), xinxin shu)
- rbd: fix init-rbdmap CMDPARAMS ([issue#13214](#), [pr#6109](#), Sage Weil)
- rbdmap: systemd support ([issue#13374](#), [pr#6479](#), Boris Ranto)
- rbd: rbdmap improvements ([pr#6445](#), Boris Ranto)
- release-notes: draft v0.94.4 release notes ([pr#5907](#), Loic Dachary)
- release-notes: draft v0.94.4 release notes ([pr#6195](#), Loic Dachary)
- release-notes: draft v0.94.4 release notes ([pr#6238](#), Loic Dachary)
- rgw: add compat header for TEMP_FAILURE_RETRY ([pr#6294](#), John Coyle)
- rgw: add default quota config ([pr#6400](#), Daniel Gryniewicz)
- rgw: add support for getting Swift’s DLO without manifest handling ([pr#6206](#), Radoslaw Zarzynski)
- rgw: clarify the error message when trying to create an existed user ([pr#5938](#), Zeqiang Zhuang)
- rgw: fix objects can not be displayed which object name does not cont... ([issue#12963](#), [pr#5738](#), Weijun Duan)

- rgw: fix typo in RGWHTTPClient::process error message ([pr#6424](#), Brad Hubbard)
- rgw: fix wrong etag calculation during POST on S3 bucket. ([issue#11241](#), [pr#6030](#), Radoslaw Zarzynski)
- rgw: mdlog trim add usage prompt ([pr#6059](#), Weijun Duan)
- rgw: modify the conditional statement in parse_metadata_key method. ([pr#5875](#), Zengran Zhang)
- rgw: refuse to calculate digest when the s3 secret key is empty ([issue#13133](#), [pr#6045](#), Sangdi Xu)
- rgw: remove extra check in RGWGetObj::execute ([issue#12352](#), [pr#5262](#), Javier M. Mellid)
- rgw: support core file limit for radosgw daemon ([pr#6346](#), Guang Yang)
- rgw: swift use Civetweb ssl can not get right url ([issue#13628](#), [pr#6408](#), Weijun Duan)
- rocksdb: build with PORTABLE=1 ([pr#6311](#), Sage Weil)
- rocksdb: remove rdb source files from dist tarball ([issue#13554](#), [pr#6379](#), Kefu Chai)
- rocksdb: use native rocksdb makefile (and our autotools) ([pr#6290](#), Sage Weil)
- rpm: ceph.spec.in: correctly declare systemd dependency for SLE/openSUSE ([pr#6114](#), Nathan Cutler)
- rpm: ceph.spec.in: fix libs-compat / devel-compat conditional ([issue#12315](#), [pr#5219](#), Ken Dreyer)
- rpm: rhel 5.9 librados compile fix, moved blkid to RBD check/compilation ([issue#13177](#), [pr#5954](#), Rohan Mars)
- scripts: release_notes can track original issue ([pr#6009](#), Abhishek Lekshmanan)
- test/libcephfs/flock: add sys/file.h include for flock operations ([pr#6310](#), John Coyle)
- test_rgw_admin: use freopen for output redirection. ([pr#6303](#), John Coyle)
- tests: allow docker-test.sh to run under root ([issue#13355](#), [pr#6173](#), Loic Dachary)
- tests: ceph-disk workunit uses configobj ([pr#6342](#), Loic Dachary)
- tests: destroy testprofile before creating one ([issue#13664](#), [pr#6446](#), Loic Dachary)
- tests: port uniqueness reminder ([pr#6387](#), Loic Dachary)
- tests: test/librados/test.cc must create profile ([issue#13664](#), [pr#6452](#), Loic Dachary)
- tools/cephfs: fix overflow writing header to fixed size buffer (#13816) ([pr#6617](#), John Spray)
- tools: ceph-monstore-update-crush: add “-test” when testing crushmap ([pr#6418](#), Kefu Chai)
- tools:remove duplicate references ([pr#5917](#), Bo Cai)
- vstart: grant full access to Swift testing account ([pr#6239](#), Yuan Zhou)
- vstart: set cephfs root uid/gid to caller ([pr#6255](#), John Spray)

V9.2.1 INFERNALIS

This Infernalis point release fixes several packagins and init script issues, enables the librbd objectmap feature by default, a few librbd bugs, and a range of miscellaneous bug fixes across the system.

We recommend that all infernalis v9.2.0 users upgrade.

For more detailed information, see [the complete changelog](#).

UPGRADING

- Some symbols wrongly exposed by the C++ interface for librados in v9.1.0 and v9.2.0 were removed. If you compiled your own application against librados shipped with these releases, it is very likely referencing these removed symbols. So you will need to recompile it.

NOTABLE CHANGES

- build/ops: Ceph daemon failed to start, because the service name was already used. ([issue#13474](#), [pr#6833](#), Chuanhong Wang)
- build/ops: ceph upstart script rbdmap.conf incorrectly processes parameters ([issue#13214](#), [pr#6396](#), Sage Weil)
- build/ops: libunwind package missing on CentOS 7 ([issue#13997](#), [pr#6845](#), Loic Dachary)
- build/ops: rbd-replay-* moved from ceph-test-dbg to ceph-common-dbg as well ([issue#13785](#), [pr#6628](#), Loic Dachary)
- build/ops: systemd/ceph-disk@.service assumes /bin/flock ([issue#13975](#), [pr#6852](#), Loic Dachary)
- build/ops: systemd: no rbdmap systemd unit file ([issue#13374](#), [pr#6500](#), Boris Ranto)
- common: auth/cephx: large amounts of log are produced by osd ([issue#13610](#), [pr#6836](#), Qiankun Zheng)
- common: log: Log.cc: Assign LOG_DEBUG priority to syslog calls ([issue#13993](#), [pr#6993](#), Brad Hubbard)
- crush: crash if we see CRUSH_ITEM_NONE in early rule step ([issue#13477](#), [pr#6626](#), Sage Weil)
- fs: Ceph file system is not freeing space ([issue#13777](#), [pr#7431](#), Yan, Zheng, John Spray)
- fs: Ceph-fuse won't start correctly when the option log_max_new in ceph.conf set to zero ([issue#13443](#), [pr#6395](#), Wenjun Huang)
- fs: Segmentation fault accessing file using fuse mount ([issue#13714](#), [pr#6853](#), Yan, Zheng)
- librbd: Avoid re-writing old-format image header on resize ([issue#13674](#), [pr#6630](#), Jason Dillaman)
- librbd: ImageWatcher shouldn't block the notification thread ([issue#14373](#), [pr#7406](#), Jason Dillaman)
- librbd: QEMU hangs after creating snapshot and stopping VM ([issue#13726](#), [pr#6632](#), Jason Dillaman)
- librbd: Verify self-managed snapshot functionality on image create ([issue#13633](#), [pr#7080](#), Jason Dillaman)

- librbd: [FAILED] TestLibRBD.SnapRemoveViaLockOwner ([issue#14164](#), [pr#7079](#), Jason Dillaman)
- librbd: enable feature objectmap ([issue#13558](#), [pr#6477](#), xinxin shu)
- librbd: fix merge-diff for >2GB diff-files ([issue#14030](#), [pr#6981](#), Jason Dillaman)
- librbd: flattening an rbd image with active IO can lead to hang ([issue#14092](#), [issue#14483](#), [pr#7484](#), Jason Dillaman)
- mds: fix client capabilities during reconnect (client.XXXX isn't responding to mclientcaps warning) ([issue#11482](#), [pr#6752](#), Yan, Zheng)
- mon: Ceph Pools' MAX AVAIL is 0 if some OSDs' weight is 0 ([issue#13840](#), [pr#6907](#), Chengyuan Li)
- mon: should not set isvalid = true when cephx_verify_authorizer return... ([issue#13525](#), [pr#6392](#), Ruifeng Yang)
- objecter: pool op callback may hang forever. ([issue#13642](#), [pr#6627](#), xie xingguo)
- objecter: potential null pointer access when do pool_snap_list. ([issue#13639](#), [pr#6840](#), xie xingguo)
- osd: FileStore: potential memory leak if getattrs fails. ([issue#13597](#), [pr#6846](#), xie xingguo)
- osd: OSD::build_past_intervals_parallel() shall reset primary and up_primary when begin a new past_interval. ([issue#13471](#), [pr#6397](#), xie xingguo)
- osd: call on_new_interval on newly split child PG ([issue#13962](#), [pr#6849](#), Sage Weil)
- osd: ceph-disk list fails on /dev/cciss!c0d0 ([issue#13970](#), [issue#14230](#), [pr#6880](#), Loic Dachary)
- osd: ceph-disk: use blkid instead of sgdisk -i ([issue#14080](#), [pr#7001](#), Loic Dachary, Ilya Dryomov)
- osd: fix race condition during send_failures ([issue#13821](#), [pr#6694](#), Sage Weil)
- osd: osd/PG.cc: 288: FAILED assert(info.last_epoch_started >= info.history.last_epoch_started) ([issue#14015](#), [pr#6851](#), David Zafman)
- osd: pgs stuck inconsistent after infernalis upgrade ([issue#13862](#), [pr#7421](#), David Zafman)
- rbd: TaskFinisher::cancel should remove event from SafeTimer ([issue#14476](#), [pr#7426](#), Douglas Fuller)
- rbd: cls_rbd: object_map_save should enable checksums ([issue#14280](#), [pr#7428](#), Douglas Fuller)
- rbd: misdirected op in rbd balance-reads test ([issue#13491](#), [pr#6629](#), Jason Dillaman)
- rbd: pure virtual method called ([issue#13636](#), [pr#6633](#), Jason Dillaman)
- rbd: rbd clone issue ([issue#13553](#), [pr#6474](#), xinxin shu)
- rbd: rbd-replay does not check for EOF and goes to endless loop ([issue#14452](#), [pr#7427](#), Mykola Golub)
- rbd: unknown argument -quiet in udevadm settle ([issue#13560](#), [pr#6634](#), Jason Dillaman)
- rgw: init script reload doesn't work on EL7 ([issue#13709](#), [pr#6650](#), Hervé Rousseau)
- rgw: radosgw-admin -help doesn't show the orphans find command ([issue#14516](#), [pr#7543](#), Yehuda Sadeh)
- tests: ceph-disk workunit uses configobj ([issue#14004](#), [pr#6828](#), Loic Dachary)
- tests: fsx failed to compile ([issue#14384](#), [pr#7429](#), Greg Farnum)
- tests: notification slave needs to wait for master ([issue#13810](#), [pr#7225](#), Jason Dillaman)
- tests: rebuild exclusive lock test should acquire exclusive lock ([issue#14121](#), [pr#7038](#), Jason Dillaman)
- tests: testprofile must be removed before it is re-created ([issue#13664](#), [pr#6449](#), Loic Dachary)
- tests: verify it is possible to reuse an OSD id ([issue#13988](#), [pr#6882](#), Loic Dachary)

V9.2.0 INFERNALIS

This major release will be the foundation for the next stable series. There have been some major changes since v0.94.x Hammer, and the upgrade process is non-trivial. Please read these release notes carefully.

MAJOR CHANGES FROM HAMMER

- *General:*
 - Ceph daemons are now managed via systemd (with the exception of Ubuntu Trusty, which still uses upstart).
 - Ceph daemons run as 'ceph' user instead root.
 - On Red Hat distros, there is also an SELinux policy.
- *RADOS:*
 - The RADOS cache tier can now proxy write operations to the base tier, allowing writes to be handled without forcing migration of an object into the cache.
 - The SHEC erasure coding support is no longer flagged as experimental. SHEC trades some additional storage space for faster repair.
 - There is now a unified queue (and thus prioritization) of client IO, recovery, scrubbing, and snapshot trimming.
 - There have been many improvements to low-level repair tooling (ceph-objectstore-tool).
 - The internal ObjectStore API has been significantly cleaned up in order to facilitate new storage backends like NewStore.
- *RGW:*
 - The Swift API now supports object expiration.
 - There are many Swift API compatibility improvements.
- *RBD:*
 - The rbd du command shows actual usage (quickly, when object-map is enabled).
 - The object-map feature has seen many stability improvements.
 - Object-map and exclusive-lock features can be enabled or disabled dynamically.

- You can now store user metadata and set persistent librbd options associated with individual images.
- The new deep-flatten features allows flattening of a clone and all of its snapshots. (Previously snapshots could not be flattened.)
- The export-diff command command is now faster (it uses aio). There is also a new fast-diff feature.
- The -size argument can be specified with a suffix for units (e.g., --size 64G).
- There is a new rbd status command that, for now, shows who has the image open/mapped.
- *CephFS*:
 - You can now rename snapshots.
 - There have been ongoing improvements around administration, diagnostics, and the check and repair tools.
 - The caching and revocation of client cache state due to unused inodes has been dramatically improved.
 - The ceph-fuse client behaves better on 32-bit hosts.

DISTRO COMPATIBILITY

We have decided to drop support for many older distributions so that we can move to a newer compiler toolchain (e.g., C++11). Although it is still possible to build Ceph on older distributions by installing backported development tools, we are not building and publishing release packages for ceph.com.

We now build packages for:

- CentOS 7 or later. We have dropped support for CentOS 6 (and other RHEL 6 derivatives, like Scientific Linux 6).
- Debian Jessie 8.x or later. Debian Wheezy 7.x's g++ has incomplete support for C++11 (and no systemd).
- Ubuntu Trusty 14.04 or later. Ubuntu Precise 12.04 is no longer supported.
- Fedora 22 or later.

UPGRADING FROM FIREFLY

Upgrading directly from Firefly v0.80.z is not recommended. It is possible to do a direct upgrade, but not without downtime. We recommend that clusters are first upgraded to Hammer v0.94.4 or a later v0.94.z release; only then is it possible to upgrade to Infernalis 9.2.z for an online upgrade (see below).

To do an offline upgrade directly from Firefly, all Firefly OSDs must be stopped and marked down before any Infernalis OSDs will be allowed to start up. This fencing is enforced by the Infernalis monitor, so use an upgrade procedure like:

1. Upgrade Ceph on monitor hosts
2. Restart all ceph-mon daemons
3. Upgrade Ceph on all OSD hosts
4. Stop all ceph-osd daemons
5. Mark all OSDs down with something like:

```
ceph osd down `seq 0 1000`
```

6. Start all ceph-osd daemons
7. Upgrade and restart remaining daemons (ceph-mds, radosgw)

UPGRADING FROM HAMMER

- All cluster nodes must first upgrade to Hammer v0.94.4 or a later v0.94.z release; only then is it possible to upgrade to Infernalis 9.2.z.
- For all distributions that support systemd (CentOS 7, Fedora, Debian Jessie 8.x, OpenSUSE), ceph daemons are now managed using native systemd files instead of the legacy sysvinit scripts. For example:

```
systemctl start ceph.target      # start all daemons
systemctl status ceph-osd@12    # check status of osd.12
```

The main notable distro that is *not* yet using systemd is Ubuntu trusty 14.04. (The next Ubuntu LTS, 16.04, will use systemd instead of upstart.)

- Ceph daemons now run as user and group ceph by default. The ceph user has a static UID assigned by Fedora and Debian (also used by derivative distributions like RHEL/CentOS and Ubuntu). On SUSE the ceph user will currently get a dynamically assigned UID when the user is created.

If your systems already have a ceph user, upgrading the package will cause problems. We suggest you first remove or rename the existing ‘ceph’ user and ‘ceph’ group before upgrading.

When upgrading, administrators have two options:

1. Add the following line to ceph.conf on all hosts:

```
setuser match path = /var/lib/ceph/$type/$cluster-$id
```

This will make the Ceph daemons run as root (i.e., not drop privileges and switch to user ceph) if the daemon’s data directory is still owned by root. Newly deployed daemons will be created with data owned by user ceph and will run with reduced privileges, but upgraded daemons will continue to run as root.

2. Fix the data ownership during the upgrade. This is the preferred option, but it is more work and can be very time consuming. The process for each host is to:

1. Upgrade the ceph package. This creates the ceph user and group. For example:

```
ceph-deploy install --stable infernalis HOST
```

2. Stop the daemon(s):

```
service ceph stop          # fedora, centos, rhel, debian
stop ceph-all              # ubuntu
```

3. Fix the ownership:

```
chown -R ceph:ceph /var/lib/ceph
chown -R ceph:ceph /var/log/ceph
```

4. Restart the daemon(s):

```
start ceph-all           # ubuntu
systemctl start ceph.target # debian, centos, fedora, rhel
```

Alternatively, the same process can be done with a single daemon type, for example by stopping only monitors and chowning only /var/lib/ceph/mon.

- The on-disk format for the experimental KeyValueStore OSD backend has changed. You will need to remove any OSDs using that backend before you upgrade any test clusters that use it.
- When a pool quota is reached, librados operations now block indefinitely, the same way they do when the cluster fills up. (Previously they would return -ENOSPC). By default, a full cluster or pool will now block. If your librados application can handle ENOSPC or EDQUOT errors gracefully, you can get error returns instead by using the new librados OPERATION_FULL_TRY flag.
- The return code for librbd’s rbd_aio_read and Image::aio_read API methods no longer returns the number of bytes read upon success. Instead, it returns 0 upon success and a negative value upon failure.
- ‘ceph scrub’, ‘ceph compact’ and ‘ceph sync force’ are now DEPRECATED. Users should instead use ‘ceph mon scrub’, ‘ceph mon compact’ and ‘ceph mon sync force’.
- ‘ceph mon_metadata’ should now be used as ‘ceph mon metadata’. There is no need to deprecate this command (same major release since it was first introduced).
- The *-dump-json* option of “osdmaptool” is replaced by *-dump json*.
- The commands of “pg ls-by-{pool,primary,osd}” and “pg ls” now take “recovering” instead of “recovery”, to include the recovering pgs in the listed pgs.

NOTABLE CHANGES SINCE HAMMER

- aarch64: add optimized version of crc32c (Yazen Ghannam, Steve Capper)
- auth: cache/reuse crypto lib key objects, optimize msg signature check (Sage Weil)
- auth: reinit NSS after fork() (#11128 Yan, Zheng)
- autotools: fix out of tree build (Krysztof Kosinski)
- autotools: improve make check output (Loic Dachary)
- buffer: add invalidate_crc() (Piotr Dalek)
- buffer: fix zero bug (#12252 Haomai Wang)
- buffer: some cleanup (Michal Jarzabek)
- build: allow tcmalloc-minimal (Thorsten Behrens)
- build: C++11 now supported
- build: cmake: fix nss linking (Danny Al-Gaaf)
- build: cmake: misc fixes (Orit Wasserman, Casey Bodley)
- build: disable LTTNG by default (#11333 Josh Durgin)
- build: do not build ceph-dencoder with tcmalloc (#10691 Boris Ranto)
- build: fix junit detection on Fedora 22 (Ira Cooper)
- build: fix pg ref disabling (William A. Kennington III)
- build: fix ppc build (James Page)
- build: install-deps: misc fixes (Loic Dachary)
- build: install-deps.sh improvements (Loic Dachary)
- build: install-deps: support OpenSUSE (Loic Dachary)
- build: make_dist_tarball.sh (Sage Weil)
- build: many cmake improvements
- build: misc cmake fixes (Matt Benjamin)
- build: misc fixes (Boris Ranto, Ken Dreyer, Owen Synge)
- build: OSX build fixes (Yan, Zheng)
- build: remove rest-bench
- ceph-authtool: fix return code on error (Gerhard Muntingh)
- ceph-detect-init: added Linux Mint (Michal Jarzabek)
- ceph-detect-init: robust init system detection (Owen Synge)
- ceph-disk: ensure ‘zap’ only operates on a full disk (#11272 Loic Dachary)
- ceph-disk: fix zap sgdisk invocation (Owen Synge, Thorsten Behrens)
- ceph-disk: follow ceph-osd hints when creating journal (#9580 Sage Weil)
- ceph-disk: handle re-using existing partition (#10987 Loic Dachary)
- ceph-disk: improve parted output parsing (#10983 Loic Dachary)
- ceph-disk: install pip > 6.1 (#11952 Loic Dachary)
- ceph-disk: make suppression work for activate-all and activate-journal (Dan van der Ster)
- ceph-disk: many fixes (Loic Dachary, Alfredo Deza)
- ceph-disk: fixes to respect init system (Loic Dachary, Owen Synge)
- ceph-disk: pass –cluster arg on prepare subcommand (Kefu Chai)
- ceph-disk: support for multipath devices (Loic Dachary)
- ceph-disk: support NVMe device partitions (#11612 Ilja Slepnev)
- ceph: fix ‘df’ units (Zhe Zhang)
- ceph: fix parsing in interactive cli mode (#11279 Kefu Chai)
- cephfs-data-scan: many additions, improvements (John Spray)
- ceph-fuse: do not require successful remount when unmounting (#10982 Greg Farnum)
- ceph-fuse, libcephfs: don’t clear COMPLETE when trimming null (Yan, Zheng)
- ceph-fuse, libcephfs: drop inode when rmdir finishes (#11339 Yan, Zheng)
- ceph-fuse, libcephfs: fix uninode (#11356 Yan, Zheng)
- ceph-fuse, libcephfs: hold exclusive caps on dirs we “own” (#11226 Greg Farnum)
- ceph-fuse: mostly behave on 32-bit hosts (Yan, Zheng)
- ceph: improve error output for ‘tell’ (#11101 Kefu Chai)
- ceph-monstore-tool: fix store-copy (Huangjun)
- ceph: new ‘ceph daemonperf’ command (John Spray, Mykola Golub)
- ceph-objectstore-tool: many many improvements (David Zafman)
- ceph-objectstore-tool: refactoring and cleanup (John Spray)
- ceph-post-file: misc fixes (Joey McDonald, Sage Weil)
- ceph_test_rados: test pipelined reads (Zhiqiang Wang)
- client: avoid sending unnecessary FLUSHSNAP messages (Yan, Zheng)
- client: exclude setfilelock when calculating oldest tid (Yan, Zheng)
- client: fix error handling in check_pool_perm (John Spray)
- client: fsync waits only for inode’s caps to flush (Yan, Zheng)
- client: invalidate kernel dcache when cache size exceeds limits (Yan, Zheng)

- client: make fsync wait for unsafe dir operations (Yan, Zheng)
- client: pin lookup dentry to avoid inode being freed (Yan, Zheng)
- common: add descriptions to perfcounters (Kiseleva Alyona)
- common: add perf counter descriptions (Alyona Kiseleva)
- common: bufferlist performance tuning (Piotr Dalek, Sage Weil)
- common: detect overflow of int config values (#11484 Kefu Chai)
- common: fix bit_vector extent calc (#12611 Jason Dillaman)
- common: fix json parsing of utf8 (#7387 Tim Serong)
- common: fix leak of pthread_mutexattr (#11762 Keton Meng)
- common: fix LTTNG vs fork issue (Josh Durgin)
- common: fix throttle max change (Henry Chang)
- common: make mutex more efficient
- common: make work queue addition/removal thread safe (#12662 Jason Dillaman)
- common: optracker improvements (Zhiqiang Wang, Jianpeng Ma)
- common: PriorityQueue tests (Kefu Chai)
- common: some async compression infrastructure (Haomai Wang)
- crush: add -check to validate dangling names, max osd id (Kefu Chai)
- crush: cleanup, sync with kernel (Ilya Dryomov)
- crush: fix crash from invalid 'take' argument (#11602 Shiva Rkreddy, Sage Weil)
- crush: fix divide-by-2 in straw2 (#11357 Yann Dupont, Sage Weil)
- crush: fix has_v4_buckets (#11364 Sage Weil)
- crush: fix subtree base weight on adjust_subtree_weight (#11855 Sage Weil)
- crush: respect default replicated ruleset config on map creation (Ilya Dryomov)
- crushtool: fix order of operations, usage (Sage Weil)
- crypto: fix NSS leak (Jason Dillaman)
- crypto: fix unbalanced init/shutdown (#12598 Zheng Yan)
- deb: fix rest-bench-dbg and ceph-test-dbg dependendies (Ken Dreyer)
- debian: minor package reorg (Ken Dreyer)
- deb, rpm: move ceph-objectstore-tool to ceph (Ken Dreyer)
- doc: docuemnt object corpus generation (#11099 Alexis Normand)
- doc: document region hostnames (Robin H. Johnson)
- doc: fix gender neutrality (Alexandre Maragone)
- doc: fix install doc (#10957 Kefu Chai)
- doc: fix sphinx issues (Kefu Chai)
- doc: man page updates (Kefu Chai)
- doc: mds data structure docs (Yan, Zheng)
- doc: misc updates (Fracois Lafont, Ken Dreyer, Kefu Chai, Owen Synge, Gael Fenet-Garde, Loic Dachary, Yannick Atchydala, Jiaying Ren, Kevin Caradant, Robert Maxime, Nicolas Yong, Germain Chipaux, Arthur Gorjux, Gabriel Sentucq, Clement Lebrun, Jean-Remi Deveaux, Clair Massot, Robin Tang, Thomas Laumondais, Jordan Dorne, Yuan Zhou, Valentin Thomas, Pierre Chaumont, Benjamin Troquereau, Benjamin Sesia, Vikhyat Umrao, Nilamdyuti Goswami, Vartika Rai, Florian Haas, Loic Dachary, Simon Guinot, Andy Allan, Alistair Israel, Ken Dreyer, Robin Rehu, Lee Revell, Florian Marsylle, Thomas Johnson, Bosse Klykken, Travis Rhoden, Ian Kelling)
- doc: swift tempurls (#10184 Abhishek Lekshmanan)
- doc: switch doxygen integration back to breathe (#6115 Kefu Chai)
- doc: update release schedule docs (Loic Dachary)
- erasure-code: cleanup (Kefu Chai)
- erasure-code: improve tests (Loic Dachary)
- erasure-code: shec: fix recovery bugs (Takanori Nakao, Shotaro Kawaguchi)
- erasure-code: update ISA-L to 2.13 (Yuan Zhou)
- gmock: switch to submodule (Danny Al-Gaaf, Loic Dachary)
- hadoop: add terasort test (Noah Watkins)
- init-radosgw: merge with sysv version; fix enumeration (Sage Weil)
- java: fix libcephfs bindings (Noah Watkins)
- libcephfs: add pread, pwrite (Jevon Qiao)
- libcephfs,ceph-fuse: cache cleanup (Zheng Yan)
- libcephfs,ceph-fuse: fix request resend on cap reconnect (#10912 Yan, Zheng)
- librados: add config observer (Alistair Strachan)
- librados: add FULL_TRY and FULL_FORCE flags for dealing with full clusters or pools (Sage Weil)
- librados: add src_fadvise_flags for copy-from (Jianpeng Ma)
- librados: define C++ flags from C constants (Josh Durgin)
- librados: fadvise flags per op (Jianpeng Ma)
- librados: fix last_force_resent handling (#11026 Jianpeng Ma)
- librados: fix memory leak from C_TwoContexts (Xiong Yiliang)
- librados: fix notify completion race (#13114 Sage Weil)
- librados: fix striper when stripe_count = 1 and stripe_unit != object_size (#11120 Yan, Zheng)

- librados, libcephfs: randomize client nonces (Josh Durgin)
- librados: op perf counters (John Spray)
- librados: pybind: fix binary omap values (Robin H. Johnson)
- librados: pybind: fix write() method return code (Javier Guerra)
- librados: respect default_crush_ruleset on pool_create (#11640 Yuan Zhou)
- libradosstriper: fix leak (Danny Al-Gaaf)
- librbd: add const for single-client-only features (Josh Durgin)
- librbd: add deep-flatten operation (Jason Dillaman)
- librbd: add purge_on_error cache behavior (Jianpeng Ma)
- librbd: allow additional metadata to be stored with the image (Haomai Wang)
- librbd: avoid blocking aio API methods (#11056 Jason Dillaman)
- librbd: better handling for dup flatten requests (#11370 Jason Dillaman)
- librbd: cancel in-flight ops on watch error (#11363 Jason Dillaman)
- librbd: default new images to format 2 (#11348 Jason Dillaman)
- librbd: fadvise for copy, export, import (Jianpeng Ma)
- librbd: fast diff implementation that leverages object map (Jason Dillaman)
- librbd: fix fast diff bugs (#11553 Jason Dillaman)
- librbd: fix image format detection (Zhiqiang Wang)
- librbd: fix lock ordering issue (#11577 Jason Dillaman)
- librbd: fix reads larger than the cache size (Lu Shi)
- librbd: fix snapshot creation when other snap is active (#11475 Jason Dillaman)
- librbd: flatten/copyup fixes (Jason Dillaman)
- librbd: handle NOCACHE fadvise flag (Jianpeng Ma)
- librbd: lockdep, helgrind validation (Jason Dillaman, Josh Durgin)
- librbd: metadata filter fixes (Haomai Wang)
- librbd: misc aio fixes (#5488 Jason Dillaman)
- librbd: misc rbd fixes (#11478 #11113 #11342 #11380 Jason Dillaman, Zhiqiang Wang)
- librbd: new diff_iterate2 API (Jason Dillaman)
- librbd: object map rebuild support (Jason Dillaman)
- librbd: only update image flags while hold exclusive lock (#11791 Jason Dillaman)
- librbd: optionally disable allocation hint (Haomai Wang)
- librbd: prevent race between resize requests (#12664 Jason Dillaman)
- librbd: readahead fixes (Zhiqiang Wang)
- librbd: return result code from close (#12069 Jason Dillaman)
- librbd: store metadata, including config options, in image (Haomai Wang)
- librbd: tolerate old osds when getting image metadata (#11549 Jason Dillaman)
- librbd: use write_full when possible (Zhiqiang Wang)
- log: fix data corruption race resulting from log rotation (#12465 Samuel Just)
- logrotate.d: prefer service over invoke-rc.d (#11330 Win Hierman, Sage Weil)
- mds: add 'damaged' state to MDSMap (John Spray)
- mds: add nicknames for perfcounters (John Spray)
- mds: avoid emitting cap warnigns before evicting session (John Spray)
- mds: avoid getting stuck in XLOCKDONE (#11254 Yan, Zheng)
- mds: disable problematic rstat propagation into snap parents (Yan, Zheng)
- mds: do not add snapped items to bloom filter (Yan, Zheng)
- mds: expose frags via asok (John Spray)
- mds: fix expected holes in journal objects (#13167 Yan, Zheng)
- mds: fix handling for missing mydir dirfrag (#11641 John Spray)
- mds: fix integer truncateion on large client ids (Henry Chang)
- mds: fix mydir replica issue with shutdown (#10743 John Spray)
- mds: fix out-of-order messages (#11258 Yan, Zheng)
- mds: fix rejoin (Yan, Zheng)
- mds: fix setting entire file layout in one setxattr (John Spray)
- mds: fix shutdown (John Spray)
- mds: fix shutdown with strays (#10744 John Spray)
- mds: fix SnapServer crash on deleted pool (John Spray)
- mds: fix snapshot bugs (Yan, Zheng)
- mds: fix stray reintegartion (Yan, Zheng)
- mds: fix stray handling (John Spray)
- mds: fix suicide beacon (John Spray)
- mds: flush immediately in do_open_truncate (#11011 John Spray)
- mds: handle misc corruption issues (John Spray)
- mds: improve dump methods (John Spray)
- mds: many fixes (Yan, Zheng, John Spray, Greg Farnum)
- mds: many snapshot and stray fixes (Yan, Zheng)

- mds: misc fixes (Jianpeng Ma, Dan van der Ster, Zhang Zhi)
- mds: misc journal cleanups and fixes (#10368 John Spray)
- mds: misc repair improvements (John Spray)
- mds: misc snap fixes (Zheng Yan)
- mds: misc snapshot fixes (Yan, Zheng)
- mds: new SessionMap storage using omap (#10649 John Spray)
- mds: persist completed_requests reliably (#11048 John Spray)
- mds: reduce memory consumption (Yan, Zheng)
- mds: respawn instead of suicide on blacklist (John Spray)
- mds: separate safe_pos in Journaler (#10368 John Spray)
- mds: snapshot rename support (#3645 Yan, Zheng)
- mds: store layout on header object (#4161 John Spray)
- mds: throttle purge stray operations (#10390 John Spray)
- mds: tolerate clock jumping backwards (#11053 Yan, Zheng)
- mds: warn when clients fail to advance oldest_client_tid (#10657 Yan, Zheng)
- misc cleanups and fixes (Danny Al-Gaaf)
- misc coverity fixes (Danny Al-Gaaf)
- misc performance and cleanup (Nathan Cutler, Xinxin Shu)
- mon: add cache over MonitorDBStore (Kefu Chai)
- mon: add 'mon_metadata <id>' command (Kefu Chai)
- mon: add 'node ls ...' command (Kefu Chai)
- mon: add NOFORWARD, OBSOLETE, DEPRECATE flags for mon commands (Joao Eduardo Luis)
- mon: add PG count to 'ceph osd df' output (Michal Jarzabek)
- mon: 'ceph osd metadata' can dump all osds (Haomai Wang)
- mon: clean up, reorg some mon commands (Joao Eduardo Luis)
- monclient: flush_log (John Spray)
- mon: detect kv backend failures (Sage Weil)
- mon: disallow >2 tiers (#11840 Kefu Chai)
- mon: disallow ec pools as tiers (#11650 Samuel Just)
- mon: do not deactivate last mds (#10862 John Spray)
- mon: fix average utilization calc for 'osd df' (Mykola Golub)
- mon: fix CRUSH map test for new pools (Sage Weil)
- mon: fix log dump crash when debugging (Mykola Golub)
- mon: fix mds beacon replies (#11590 Kefu Chai)
- mon: fix metadata update race (Mykola Golub)
- mon: fix min_last_epoch_clean tracking (Kefu Chai)
- mon: fix 'pg ls' sort order, state names (#11569 Kefu Chai)
- mon: fix refresh (#11470 Joao Eduardo Luis)
- mon: fix variance calc in 'osd df' (Sage Weil)
- mon: improve callout to crushtool (Mykola Golub)
- mon: make blocked op messages more readable (Jianpeng Ma)
- mon: make osd get pool 'all' only return applicable fields (#10891 Michal Jarzabek)
- mon: misc scaling fixes (Sage Weil)
- mon: normalize erasure-code profile for storage and comparison (Loic Dachary)
- mon: only send mon metadata to supporting peers (Sage Weil)
- mon: optionally specify osd id on 'osd create' (Mykola Golub)
- mon: 'osd tree' fixes (Kefu Chai)
- mon: periodic background scrub (Joao Eduardo Luis)
- mon: prevent bucket deletion when referenced by a crush rule (#11602 Sage Weil)
- mon: prevent pgp_num > pg_num (#12025 Xinxin Shu)
- mon: prevent pool with snapshot state from being used as a tier (#11493 Sage Weil)
- mon: prime pg_temp when CRUSH map changes (Sage Weil)
- mon: refine check_remove_tier checks (#11504 John Spray)
- mon: reject large max_mds values (#12222 John Spray)
- mon: remove spurious who arg from 'mds rm ...' (John Spray)
- mon: streamline session handling, fix memory leaks (Sage Weil)
- mon: upgrades must pass through hammer (Sage Weil)
- mon: warn on bogus cache tier config (Jianpeng Ma)
- msgr: add ceph_perf_msgr tool (Haomai Wang)
- msgr: async: fix seq handling (Haomai Wang)
- msgr: async: many many fixes (Haomai Wang)
- msgr: simple: fix clear_pipe (#11381 Haomai Wang)
- msgr: simple: fix connect_seq assert (Haomai Wang)
- msgr: xio: fastpath improvements (Raju Kurunkad)
- msgr: xio: fix ip and nonce (Raju Kurunkad)

- msgr: xio: improve lane assignment (Vu Pham)
- msgr: xio: sync with accelio v1.4 (Vu Pham)
- msgr: xio: misc fixes (#10735 Matt Benjamin, Kefu Chai, Danny Al-Gaaf, Raju Kurunkad, Vu Pham, Casey Bodley)
- msg: unit tests (Haomai Wang)
- objectcacher: misc bug fixes (Jianpeng Ma)
- osd: add latency perf counters for tier operations (Xinze Chi)
- osd: add misc perfcounters (Xinze Chi)
- osd: add simple sleep injection in recovery (Sage Weil)
- osd: allow SEEK_HOLE/SEEK_DATA for sparse read (Zhiqiang Wang)
- osd: avoid dup omap sets for in pg metadata (Sage Weil)
- osd: avoid multiple hit set insertions (Zhiqiang Wang)
- osd: avoid transaction append in some cases (Sage Weil)
- osd: break PG removal into multiple iterations (#10198 Guang Yang)
- osd: cache proxy-write support (Zhiqiang Wang, Samuel Just)
- osd: check scrub state when handling map (Jianpeng Ma)
- osd: clean up some constness, privateness (Kefu Chai)
- osd: clean up temp object if promotion fails (Jianpeng Ma)
- osd: configure promotion based on write recency (Zhiqiang Wang)
- osd: constrain collections to meta and PGs (normal and temp) (Sage Weil)
- osd: don't send dup MMonGetOSDMap requests (Sage Weil, Kefu Chai)
- osd: EIO injection (David Zhang)
- osd: elminiate txn apend, ECSubWrite copy (Samuel Just)
- osd: erasure-code: drop entries according to LRU (Andreas-Joachim Peters)
- osd: erasure-code: fix SHEC floating point bug (#12936 Loic Dachary)
- osd: erasure-code: update to ISA-L 2.14 (Yuan Zhou)
- osd: filejournal: cleanup (David Zafman)
- osd: filestore: clone using splice (Jianpeng Ma)
- osd: filestore: fix recursive lock (Xinxin Shu)
- osd: fix check_for_full (Henry Chang)
- osd: fix dirty accounting in make_writeable (Zhiqiang Wang)
- osd: fix dup promotion lost op bug (Zhiqiang Wang)
- osd: fix endless repair when object is unrecoverable (Jianpeng Ma, Kefu Chai)
- osd: fix hitset object naming to use GMT (Kefu Chai)
- osd: fix misc memory leaks (Sage Weil)
- osd: fix negative degraded stats during backfill (Guang Yang)
- osd: fix osdmap dump of blacklist items (John Spray)
- osd: fix peek_queue locking in FileStore (Xinze Chi)
- osd: fix pg resurrection (#11429 Samuel Just)
- osd: fix promotion vs full cache tier (Samuel Just)
- osd: fix replay requeue when pg is still activating (#13116 Samuel Just)
- osd: fix scrub stat bugs (Sage Weil, Samuel Just)
- osd: fix snap flushing from cache tier (again) (#11787 Samuel Just)
- osd: fix snap handling on promotion (#11296 Sam Just)
- osd: fix temp-clearing (David Zafman)
- osd: force promotion for ops EC can't handle (Zhiqiang Wang)
- osd: handle log split with overlapping entries (#11358 Samuel Just)
- osd: ignore non-existent osds in unfound calc (#10976 Mykola Golub)
- osd: improve behavior on machines with large memory pages (Steve Capper)
- osd: include a temp namespace within each collection/pgid (Sage Weil)
- osd: increase default max open files (Owen Synge)
- osd: keyvaluestore: misc fixes (Varada Kari)
- osd: low and high speed flush modes (Mingxin Liu)
- osd: make suicide timeouts individually configurable (Samuel Just)
- osd: merge multiple setattr calls into a setattrs call (Xinxin Shu)
- osd: misc fixes (Ning Yao, Kefu Chai, Xinze Chi, Zhiqiang Wang, Jianpeng Ma)
- osd: move scrub in OpWQ (Samuel Just)
- osd: newstore prototype (Sage Weil)
- osd: ObjectStore internal API refactor (Sage Weil)
- osd: peer_features includes self (David Zafman)
- osd: pool size change triggers new interval (#11771 Samuel Just)
- osd: prepopulate needs_recovery_map when only one peer has missing (#9558 Guang Yang)
- osd: randomize scrub times (#10973 Kefu Chai)
- osd: recovery, peering fixes (#11687 Samuel Just)
- osd: refactor scrub and digest recording (Sage Weil)
- osd: refuse first write to EC object at non-zero offset (Jianpeng Ma)

- osd: relax reply order on proxy read (#11211 Zhiqiang Wang)
- osd: require firefly features (David Zafman)
- osd: set initial crush weight with more precision (Sage Weil)
- osd: SHEC no longer experimental
- osd: skip promotion for flush/evict op (Zhiqiang Wang)
- osd: stripe over small xattrs to fit in XFS's 255 byte inline limit (Sage Weil, Ning Yao)
- osd: sync object_map on syncfs (Samuel Just)
- osd: take excl lock of op is rw (Samuel Just)
- osd: throttle evict ops (Yunchuan Wen)
- osd: upgrades must pass through hammer (Sage Weil)
- osd: use a temp object for recovery (Sage Weil)
- osd: use blkid to collection partition information (Joseph Handzik)
- osd: use SEEK_HOLE / SEEK_DATA for sparse copy (Xinxin Shu)
- osd: WBThrottle cleanups (Jianpeng Ma)
- osd: write journal header on clean shutdown (Xinze Chi)
- osdc/Objecter: allow per-pool calls to op_cancel_writes (John Spray)
- os/filestore: enlarge getxattr buffer size (Jianpeng Ma)
- pybind: pep8 cleanups (Danny Al-Gaaf)
- pycephfs: many fixes for bindings (Haomai Wang)
- qa: fix filelock_interrupt.py test (Yan, Zheng)
- qa: improve ceph-disk tests (Loic Dachary)
- qa: improve docker build layers (Loic Dachary)
- qa: run-make-check.sh script (Loic Dachary)
- rados: add -striper option to use libradosstriper (#10759 Sebastien Ponce)
- rados: bench: add -no-verify option to improve performance (Piotr Dalek)
- rados bench: misc fixes (Dmitry Yatsushkevich)
- rados: fix error message on failed pool removal (Wido den Hollander)
- radosgw-admin: add 'bucket check' function to repair bucket index (Yehuda Sadeh)
- radosgw-admin: fix subuser modify output (#12286 Guce)
- rados: handle -snapid arg properly (Abhishek Lekshmanan)
- rados: improve bench buffer handling, performance (Piotr Dalek)
- rados: misc bench fixes (Dmitry Yatsushkevich)
- rados: new pool import implementation (John Spray)
- rados: translate errno to string in CLI (#10877 Kefu Chai)
- rbd: accept map options config option (Ilya Dryomov)
- rbd: add disk usage tool (#7746 Jason Dillaman)
- rbd: allow unmapping by spec (Ilya Dryomov)
- rbd: cli: fix arg parsing with -io-pattern (Dmitry Yatsushkevich)
- rbd: deprecate -new-format option (Jason Dillman)
- rbd: fix error messages (#2862 Rajesh Nambiar)
- rbd: fix link issues (Jason Dillaman)
- rbd: improve CLI arg parsing, usage (Ilya Dryomov)
- rbd: rbd-replay-prep and rbd-replay improvements (Jason Dillaman)
- rbd: recognize queue_depth kernel option (Ilya Dryomov)
- rbd: support G and T units for CLI (Abhishek Lekshmanan)
- rbd: update rbd man page (Ilya Dryomov)
- rbd: update xfstests tests (Douglas Fuller)
- rbd: use image-spec and snap-spec in help (Vikhyat Umrao, Ilya Dryomov)
- rest-bench: misc fixes (Shawn Chen)
- rest-bench: support https (#3968 Yuan Zhou)
- rgw: add max multipart upload parts (#12146 Abhishek Dixit)
- rgw: add missing headers to Swift container details (#10666 Ahmad Faheem, Dmytro Iurchenko)
- rgw: add stats to headers for account GET (#10684 Yuan Zhou)
- rgw: add Transaction-Id to response (Abhishek Dixit)
- rgw: add X-Timestamp for Swift containers (#10938 Radoslaw Zarzynski)
- rgw: always check if token is expired (#11367 Anton Aksola, Riku Lehto)
- rgw: conversion tool to repair broken multipart objects (#12079 Yehuda Sadeh)
- rgw: document layout of pools and objects (Pete Zaitcev)
- rgw: do not enclose bucket header in quotes (#11860 Wido den Hollander)
- rgw: do not prefetch data for HEAD requests (Guang Yang)
- rgw: do not preserve ACLs when copying object (#12370 Yehuda Sadeh)
- rgw: do not set content-type if length is 0 (#11091 Orit Wasserman)
- rgw: don't clobber bucket/object owner when setting ACLs (#10978 Yehuda Sadeh)
- rgw: don't use end_marker for namespaced object listing (#11437 Yehuda Sadeh)
- rgw: don't use rgw_socket_path if frontend is configured (#11160 Yehuda Sadeh)

- rgw: enforce Content-Length for POST on Swift cont/obj (#10661 Radoslaw Zarzynski)
- rgw: error out if frontend did not send all data (#11851 Yehuda Sadeh)
- rgw: expose the number of unhealthy workers through admin socket (Guang Yang)
- rgw: fail if parts not specified on multipart upload (#11435 Yehuda Sadeh)
- rgw: fix assignment of copy obj attributes (#11563 Yehuda Sadeh)
- rgw: fix broken stats in container listing (#11285 Radoslaw Zarzynski)
- rgw: fix bug in domain/subdomain splitting (Robin H. Johnson)
- rgw: fix casing of Content-Type header (Robin H. Johnson)
- rgw: fix civetweb max threads (#10243 Yehuda Sadeh)
- rgw: fix Connection: header handling (#12298 Wido den Hollander)
- rgw: fix copy metadata, support X-Copied-From for swift (#10663 Radoslaw Zarzynski)
- rgw: fix data corruptions race condition (#11749 Wuxingyi)
- rgw: fix decoding of X-Object-Manifest from GET on Swift DLO (Radoslaw Rzarzynski)
- rgw: fix GET on swift account when limit == 0 (#10683 Radoslaw Zarzynski)
- rgw: fix handling empty metadata items on Swift container (#11088 Radoslaw Zarzynski)
- rgw: fix JSON response when getting user quota (#12117 Wuxingyi)
- rgw: fix locator for objects starting with _ (#11442 Yehuda Sadeh)
- rgw: fix log rotation (Wuxingyi)
- rgw: fix multipart upload in retry path (#11604 Yehuda Sadeh)
- rgw: fix quota enforcement on POST (#11323 Sergey Arkhipov)
- rgw: fix reset_loc (#11974 Yehuda Sadeh)
- rgw: fix return code on missing upload (#11436 Yehuda Sadeh)
- rgw: fix sysvinit script
- rgw: fix sysvinit script w/ multiple instances (Sage Weil, Pavan Rallabhandi)
- rgw: force content_type for swift bucket stats requests (#12095 Orit Wasserman)
- rgw: force content type header on responses with no body (#11438 Orit Wasserman)
- rgw: generate Date header for civetweb (#10873 Radoslaw Zarzynski)
- rgw: generate new object tag when setting attrs (#11256 Yehuda Sadeh)
- rgw: improve content-length env var handling (#11419 Robin H. Johnson)
- rgw: improved support for swift account metadata (Radoslaw Zarzynski)
- rgw: improve handling of already removed buckets in expirer (Radoslaw Rzarzynski)
- rgw: issue aio for first chunk before flush cached data (#11322 Guang Yang)
- rgw: log to /var/log/ceph instead of /var/log/radosgw
- rgw: make init script wait for radosgw to stop (#11140 Dmitry Yatsushkevich)
- rgw: make max put size configurable (#6999 Yuan Zhou)
- rgw: make quota/gc threads configurable (#11047 Guang Yang)
- rgw: make read user buckets backward compat (#10683 Radoslaw Zarzynski)
- rgw: merge manifests properly with prefix override (#11622 Yehuda Sadeh)
- rgw: only scan for objects not in a namespace (#11984 Yehuda Sadeh)
- rgw: orphan detection tool (Yehuda Sadeh)
- rgw: pass in civetweb configurables (#10907 Yehuda Sadeh)
- rgw: rectify 202 Accepted in PUT response (#11148 Radoslaw Zarzynski)
- rgw: remove meta file after deleting bucket (#11149 Orit Wasserman)
- rgw: remove trailing :port from HTTP_HOST header (Sage Weil)
- rgw: return 412 on bad limit when listing buckets (#11613 Yehuda Sadeh)
- rgw: rework X-Trans-Id header to conform with Swift API (Radoslaw Rzarzynski)
- rgw: s3 encoding-type for get bucket (Jeff Weber)
- rgw: send ETag, Last-Modified for swift (#11087 Radoslaw Zarzynski)
- rgw: set content length on container GET, PUT, DELETE, HEAD (#10971, #11036 Radoslaw Zarzynski)
- rgw: set max buckets per user in ceph.conf (Vikhyat Umrao)
- rgw: shard work over multiple librados instances (Pavan Rallabhandi)
- rgw: support end marker on swift container GET (#10682 Radoslaw Zarzynski)
- rgw: support for Swift expiration API (Radoslaw Rzarzynski, Yehuda Sadeh)
- rgw: swift: allow setting attributes with COPY (#10662 Ahmad Faheem, Dmytro Iurchenko)
- rgw: swift: do not override sent content type (#12363 Orit Wasserman)
- rgw: swift: enforce Content-Type in response (#12157 Radoslaw Zarzynski)
- rgw: swift: fix account listing (#11501 Radoslaw Zarzynski)
- rgw: swift: fix metadata handling on copy (#10645 Radoslaw Zarzynski)
- rgw: swift: send Last-Modified header (#10650 Radoslaw Zarzynski)
- rgw: swift: set Content-Length for account GET (#12158 Radoslaw Zarzynski)
- rgw: swift: set content-length on keystone tokens (#11473 Herv Rousseau)
- rgw: update keystone cache with token info (#11125 Yehuda Sadeh)
- rgw: update to latest civetweb, enable config for IPv6 (#10965 Yehuda Sadeh)
- rgw: use attrs from source bucket on copy (#11639 Javier M. Mellid)
- rgw: use correct oid for gc chains (#11447 Yehuda Sadeh)

- rgw: user rm is idempotent (Orit Wasserman)
- rgw: use unique request id for civetweb (#10295 Orit Wasserman)
- rocksdb: add perf counters for get/put latency (Xinxin Shu)
- rocksdb, leveldb: fix compact_on_mount (Xiaoxi Chen)
- rocksdb: pass options as single string (Xiaoxi Chen)
- rocksdb: update to latest (Xiaoxi Chen)
- rpm: add suse firewall files (Tim Serong)
- rpm: always rebuild and install man pages for rpm (Owen Synge)
- rpm: loosen ceph-test dependencies (Ken Dreyer)
- rpm: many spec file fixes (Owen Synge, Ken Dreyer)
- rpm: misc fixes (Boris Ranto, Owen Synge, Ken Dreyer, Ira Cooper)
- rpm: misc systemd and SUSE fixes (Owen Synge, Nathan Cutler)
- selinux policy (Boris Ranto, Milan Broz)
- systemd: logrotate fixes (Tim Serong, Lars Marowsky-Bree, Nathan Cutler)
- systemd: many fixes (Sage Weil, Owen Synge, Boris Ranto, Dan van der Ster)
- systemd: run daemons as user ceph
- sysvinit compat: misc fixes (Owen Synge)
- test: misc fs test improvements (John Spray, Loic Dachary)
- test: python tests, linter cleanup (Alfredo Deza)
- tests: fixes for rbd xstests (Douglas Fuller)
- tests: fix tiering health checks (Loic Dachary)
- tests for low-level performance (Haomai Wang)
- tests: many ec non-regression improvements (Loic Dachary)
- tests: many many ec test improvements (Loic Dachary)
- upstart: throttle restarts (#11798 Sage Weil, Greg Farnum)

V9.1.0 INFERNALIS RELEASE CANDIDATE

This is the first Infernalis release candidate. There have been some major changes since Hammer, and the upgrade process is non-trivial. Please read carefully.

GETTING THE RELEASE CANDIDATE

The v9.1.0 packages are pushed to the development release repositories:

```
http://download.ceph.com/rpm-testing
http://download.ceph.com/debian-testing
```

For for info, see:

```
http://docs.ceph.com/docs/master/install/get-packages/
```

Or install with ceph-deploy via:

```
ceph-deploy install --testing HOST
```

KNOWN ISSUES

- librbd and librados ABI compatibility is broken. Be careful installing this RC on client machines (e.g., those running qemu). It will be fixed in the final v9.2.0 release.

MAJOR CHANGES FROM HAMMER

- *General*:
 - Ceph daemons are now managed via systemd (with the exception of Ubuntu Trusty, which still uses upstart).
 - Ceph daemons run as ‘ceph’ user instead of root.
 - On Red Hat distros, there is also an SELinux policy.
- *RADOS*:

- The RADOS cache tier can now proxy write operations to the base tier, allowing writes to be handled without forcing migration of an object into the cache.
- The SHEC erasure coding support is no longer flagged as experimental. SHEC trades some additional storage space for faster repair.
- There is now a unified queue (and thus prioritization) of client IO, scrubbing, and snapshot trimming.
- There have been many improvements to low-level repair tooling (ceph-objectstore-tool).
- The internal ObjectStore API has been significantly cleaned up in order to facilitate new storage backends like NewStore.
- *RGW*:
 - The Swift API now supports object expiration.
 - There are many Swift API compatibility improvements.
- *RBD*:
 - The rbd du command shows actual usage (quickly, when object-map is enabled).
 - The object-map feature has seen many stability improvements.
 - Object-map and exclusive-lock features can be enabled or disabled dynamically.
 - You can now store user metadata and set persistent librbd options associated with individual images.
 - The new deep-flatten feature allows flattening of a clone and all of its snapshots. (Previously snapshots could not be flattened.)
 - The export-diff command is now faster (it uses aio). There is also a new fast-diff feature.
 - The -size argument can be specified with a suffix for units (e.g., --size 64G).
 - There is a new rbd status command that, for now, shows who has the image open/mapped.
- *CephFS*:
 - You can now rename snapshots.
 - There have been ongoing improvements around administration, diagnostics, and the check and repair tools.
 - The caching and revocation of client cache state due to unused inodes has been dramatically improved.
 - The ceph-fuse client behaves better on 32-bit hosts.

DISTRO COMPATIBILITY

We have decided to drop support for many older distributions so that we can move to a newer compiler toolchain (e.g., C++11). Although it is still possible to build Ceph on older distributions by installing backported development tools, we are not building and publishing release packages for them on ceph.com.

In particular,

- CentOS 7 or later; we have dropped support for CentOS 6 (and other RHEL 6 derivatives, like Scientific Linux 6).
- Debian Jessie 8.x or later; Debian Wheezy 7.x's g++ has incomplete support for C++11 (and no systemd).
- Ubuntu Trusty 14.04 or later; Ubuntu Precise 12.04 is no longer supported.
- Fedora 22 or later.

UPGRADING FROM FIREFLY

Upgrading directly from Firefly v0.80.z is not possible. All clusters must first upgrade to Hammer v0.94.4 or a later v0.94.z release; only then is it possible to do online upgrade to Infernalis 9.2.z.

User can upgrade to latest hammer v0.94.z from gitbuilder with(also refer the hammer release notes for more details):

```
ceph-deploy install --release hammer HOST
```

UPGRADING FROM HAMMER

- All cluster nodes must first upgrade to Hammer v0.94.4 or a later v0.94.z release; only then is it possible to do online upgrade to Infernalis 9.2.z.
- For all distributions that support systemd (CentOS 7, Fedora, Debian Jessie 8.x, OpenSUSE), ceph daemons are now managed using native systemd files instead of the legacy sysvinit scripts. For example:

```
systemctl start ceph.target          # start all daemons
systemctl status ceph-osd@12         # check status of osd.12
```

The main notable distro that is *not* yet using systemd is Ubuntu trusty 14.04. (The next Ubuntu LTS, 16.04, will use systemd instead of upstart.)

- Ceph daemons now run as user and group ceph by default. The ceph user has a static UID assigned by Fedora and Debian (also used by derivative distributions like RHEL/CentOS and Ubuntu). On SUSE the ceph user will currently get a dynamically assigned UID when the user is created.

If your systems already have a ceph user, the package upgrade process will usually fail with an error. We suggest you first remove or rename the existing ‘ceph’ user and then upgrade.

When upgrading, administrators have two options:

1. Add the following line to ceph.conf on all hosts:

```
setuser match path = /var/lib/ceph/$type/$cluster-$id
```

This will make the Ceph daemons run as root (i.e., not drop privileges and switch to user ceph) if the daemon’s data directory is still owned by root. Newly deployed daemons will be created with data owned by user ceph and will run with reduced privileges, but upgraded daemons will continue to run as root.

2. Fix the data ownership during the upgrade. This is the preferred option, but is more work. The process for each host would be to:

1. Upgrade the ceph package. This creates the ceph user and group. For example:

```
ceph-deploy install --stable infernalis HOST
```

2. Stop the daemon(s):

```
service ceph stop          # fedora, centos, rhel, debian  
stop ceph-all             # ubuntu
```

3. Fix the ownership:

```
chown -R ceph:ceph /var/lib/ceph  
chown -R ceph:ceph /var/log/ceph
```

4. Restart the daemon(s):

```
start ceph-all           # ubuntu  
systemctl start ceph.target # debian, centos, fedora, rhel
```

- The on-disk format for the experimental KeyValueStore OSD backend has changed. You will need to remove any OSDs using that backend before you upgrade any test clusters that use it.

UPGRADE NOTES

- When a pool quota is reached, librados operations now block indefinitely, the same way they do when the cluster fills up. (Previously they would return -ENOSPC). By default, a full cluster or pool will now block. If your librados application can handle ENOSPC or EDQUOT errors gracefully, you can get error returns instead by using the new librados OPERATION_FULL_TRY flag.

NOTABLE CHANGES

NOTE: These notes are somewhat abbreviated while we find a less time-consuming process for generating them.

- build: C++11 now supported
- build: many cmake improvements
- build: OSX build fixes (Yan, Zheng)
- build: remove rest-bench
- ceph-disk: many fixes (Loic Dachary)
- ceph-disk: support for multipath devices (Loic Dachary)

- ceph-fuse: mostly behave on 32-bit hosts (Yan, Zheng)
- ceph-objectstore-tool: many improvements (David Zafman)
- common: bufferlist performance tuning (Piotr Dalek, Sage Weil)
- common: make mutex more efficient
- common: some async compression infrastructure (Haomai Wang)
- librados: add FULL_TRY and FULL_FORCE flags for dealing with full clusters or pools (Sage Weil)
- librados: fix notify completion race (#13114 Sage Weil)
- librados, libcephfs: randomize client nonces (Josh Durgin)
- librados: pybind: fix binary omap values (Robin H. Johnson)
- librbd: fix reads larger than the cache size (Lu Shi)
- librbd: metadata filter fixes (Haomai Wang)
- librbd: use write_full when possible (Zhiqiang Wang)
- mds: avoid emitting cap warnigs before evicting session (John Spray)
- mds: fix expected holes in journal objects (#13167 Yan, Zheng)
- mds: fix SnapServer crash on deleted pool (John Spray)
- mds: many fixes (Yan, Zheng, John Spray, Greg Farnum)
- mon: add cache over MonitorDBStore (Kefu Chai)
- mon: ‘ceph osd metadata’ can dump all osds (Haomai Wang)
- mon: detect kv backend failures (Sage Weil)
- mon: fix CRUSH map test for new pools (Sage Weil)
- mon: fix min_last_epoch_clean tracking (Kefu Chai)
- mon: misc scaling fixes (Sage Weil)
- mon: streamline session handling, fix memory leaks (Sage Weil)
- mon: upgrades must pass through hammer (Sage Weil)
- msg/async: many fixes (Haomai Wang)
- osd: cache proxy-write support (Zhiqiang Wang, Samuel Just)
- osd: configure promotion based on write recency (Zhiqiang Wang)
- osd: don’t send dup MMonGetOSDMap requests (Sage Weil, Kefu Chai)
- osd: erasure-code: fix SHEC floating point bug (#12936 Loic Dachary)
- osd: erasure-code: update to ISA-L 2.14 (Yuan Zhou)
- osd: fix hitset object naming to use GMT (Kefu Chai)
- osd: fix misc memory leaks (Sage Weil)
- osd: fix peek_queue locking in FileStore (Xinze Chi)
- osd: fix promotion vs full cache tier (Samuel Just)
- osd: fix replay requeue when pg is still activating (#13116 Samuel Just)
- osd: fix scrub stat bugs (Sage Weil, Samuel Just)
- osd: force promotion for ops EC can’t handle (Zhiqiang Wang)
- osd: improve behavior on machines with large memory pages (Steve Capper)
- osd: merge multiple setattr calls into a setattrs call (Xinxin Shu)
- osd: newstore prototype (Sage Weil)
- osd: ObjectStore internal API refactor (Sage Weil)
- osd: SHEC no longer experimental
- osd: throttle evict ops (Yunchuan Wen)
- osd: upgrades must pass through hammer (Sage Weil)
- osd: use SEEK_HOLE / SEEK_DATA for sparse copy (Xinxin Shu)
- rbd: rbd-replay-prep and rbd-replay improvements (Jason Dillaman)
- rgw: expose the number of unhealthy workers through admin socket (Guang Yang)
- rgw: fix casing of Content-Type header (Robin H. Johnson)
- rgw: fix decoding of X-Object-Manifest from GET on Swift DLO (Radoslaw Rzarzynski)
- rgw: fix sysvinit script
- rgw: fix sysvinit script w/ multiple instances (Sage Weil, Pavan Rallabhandi)
- rgw: improve handling of already removed buckets in expirer (Radoslaw Rzarzynski)
- rgw: log to /var/log/ceph instead of /var/log/radosgw
- rgw: rework X-Trans-Id header to be conform with Swift API (Radoslaw Rzarzynski)
- rgw: s3 encoding-type for get bucket (Jeff Weber)
- rgw: set max buckets per user in ceph.conf (Vikhyat Umrao)
- rgw: support for Swift expiration API (Radoslaw Rzarzynski, Yehuda Sadeh)
- rgw: user rm is idempotent (Orit Wasserman)
- selinux policy (Boris Ranto, Milan Broz)
- systemd: many fixes (Sage Weil, Owen Synge, Boris Ranto, Dan van der Ster)
- systemd: run daemons as user ceph

This is the second to last batch of development work for the Infernalis cycle. The most intrusive change is an internal (non user-visible) change to the OSD's ObjectStore interface. Many fixes and improvements elsewhere across RGW, RBD, and another big pile of CephFS scrub/repair improvements.

UPGRADING

- The return code for librbd's rbd_aio_read and Image::aio_read API methods no longer returns the number of bytes read upon success. Instead, it returns 0 upon success and a negative value upon failure.
- 'ceph scrub', 'ceph compact' and 'ceph sync force' are now deprecated. Users should instead use 'ceph mon scrub', 'ceph mon compact' and 'ceph mon sync force'.
- 'ceph mon_metadata' should now be used as 'ceph mon metadata'.
- The *-dump-json* option of "osdmaptool" is replaced by *-dump json*.
- The commands of 'pg ls-by-{pool,primary,osd}' and 'pg ls' now take 'recovering' instead of 'recovery' to include the recovering pgs in the listed pgs.

NOTABLE CHANGES

- autotools: fix out of tree build (Krzysztof Kosinski)
- autotools: improve make check output (Loic Dachary)
- buffer: add invalidate_crc() (Piotr Dalek)
- buffer: fix zero bug (#12252 Haomai Wang)
- build: fix junit detection on Fedora 22 (Ira Cooper)
- ceph-disk: install pip > 6.1 (#11952 Loic Dachary)
- cephfs-data-scan: many additions, improvements (John Spray)
- ceph: improve error output for 'tell' (#11101 Kefu Chai)
- ceph-objectstore-tool: misc improvements (David Zafman)
- ceph-objectstore-tool: refactoring and cleanup (John Spray)
- ceph_test_rados: test pipelined reads (Zhiqiang Wang)
- common: fix bit_vector extent calc (#12611 Jason Dillaman)
- common: make work queue addition/removal thread safe (#12662 Jason Dillaman)
- common: optracker improvements (Zhiqiang Wang, Jianpeng Ma)
- crush: add -check to validate dangling names, max osd id (Kefu Chai)
- crush: cleanup, sync with kernel (Ilya Dryomov)
- crush: fix subtree base weight on adjust_subtree_weight (#11855 Sage Weil)
- crypto: fix NSS leak (Jason Dillaman)
- crypto: fix unbalanced init/shutdown (#12598 Zheng Yan)
- doc: misc updates (Kefu Chai, Owen Syngue, Gael Fenet-Garde, Loic Dachary, Yannick Atchy-Dalama, Jiaying Ren, Kevin Caradant, Robert Maxime, Nicolas Yong, Germain Chipaux, Arthur Gorjux, Gabriel Sentucq, Clement Lebrun, Jean-Remi Deveaux, Clair Massot, Robin Tang, Thomas Laumondais, Jordan Dorne, Yuan Zhou, Valentin Thomas, Pierre Chaumont, Benjamin Troquereau, Benjamin Sesia, Vikhyat Umrao)
- erasure-code: cleanup (Kefu Chai)
- erasure-code: improve tests (Loic Dachary)
- erasure-code: shec: fix recovery bugs (Takanori Nakao, Shotaro Kawaguchi)
- libcephfs: add pread, pwrite (Jevon Qiao)
- libcephfs,ceph-fuse: cache cleanup (Zheng Yan)
- librados: add src_fadvise_flags for copy-from (Jianpeng Ma)
- librados: respect default_crush_ruleset on pool_create (#11640 Yuan Zhou)
- librbd: fadvise for copy, export, import (Jianpeng Ma)
- librbd: handle NOCACHE fadvise flag (Jianpeng Ma)
- librbd: optionally disable allocation hint (Haomai Wang)
- librbd: prevent race between resize requests (#12664 Jason Dillaman)
- log: fix data corruption race resulting from log rotation (#12465 Samuel Just)
- mds: expose frags via asok (John Spray)
- mds: fix setting entire file layout in one setxattr (John Spray)
- mds: fix shutdown (John Spray)
- mds: handle misc corruption issues (John Spray)
- mds: misc fixes (Jianpeng Ma, Dan van der Ster, Zhang Zhi)
- mds: misc snap fixes (Zheng Yan)
- mds: store layout on header object (#4161 John Spray)
- misc performance and cleanup (Nathan Cutler, Xinxin Shu)
- mon: add NOFORWARD, OBSOLETE, DEPRECATE flags for mon commands (Joao Eduardo Luis)
- mon: add PG count to 'ceph osd df' output (Michal Jarzabek)
- mon: clean up, reorg some mon commands (Joao Eduardo Luis)

- mon: disallow >2 tiers (#11840 Kefu Chai)
- mon: fix log dump crash when debugging (Mykola Golub)
- mon: fix metadata update race (Mykola Golub)
- mon: fix refresh (#11470 Joao Eduardo Luis)
- mon: make blocked op messages more readable (Jianpeng Ma)
- mon: only send mon metadata to supporting peers (Sage Weil)
- mon: periodic background scrub (Joao Eduardo Luis)
- mon: prevent pgp_num > pg_num (#12025 Xinxin Shu)
- mon: reject large max_mds values (#12222 John Spray)
- msgr: add ceph_perf_msgr tool (Hoamai Wang)
- msgr: async: fix seq handling (Haomai Wang)
- msgr: xio: fastpath improvements (Raju Kurunkad)
- msgr: xio: sync with accellio v1.4 (Vu Pham)
- osd: clean up temp object if promotion fails (Jianpeng Ma)
- osd: constrain collections to meta and PGs (normal and temp) (Sage Weil)
- osd: filestore: clone using splice (Jianpeng Ma)
- osd: filestore: fix recursive lock (Xinxin Shu)
- osd: fix dup promotion lost op bug (Zhiqiang Wang)
- osd: fix temp-clearing (David Zafman)
- osd: include a temp namespace within each collection/pgid (Sage Weil)
- osd: low and high speed flush modes (Mingxin Liu)
- osd: peer_features includes self (David Zafman)
- osd: recovery, peering fixes (#11687 Samuel Just)
- osd: require firefly features (David Zafman)
- osd: set initial crush weight with more precision (Sage Weil)
- osd: use a temp object for recovery (Sage Weil)
- osd: use blkid to collection partition information (Joseph Handzik)
- rados: add --striper option to use libradosstriper (#10759 Sebastien Ponce)
- radosgw-admin: fix subuser modify output (#12286 Guce)
- rados: handle -snapid arg properly (Abhishek Lekshmanan)
- rados: improve bench buffer handling, performance (Piotr Dalek)
- rados: new pool import implementation (John Spray)
- rbd: fix link issues (Jason Dillaman)
- rbd: improve CLI arg parsing, usage (Ilya Dryomov)
- rbd: recognize queue_depth kernel option (Ilya Dryomov)
- rbd: support G and T units for CLI (Abhishek Lekshmanan)
- rbd: use image-spec and snap-spec in help (Vikhyat Umrao, Ilya Dryomov)
- rest-bench: misc fixes (Shawn Chen)
- rest-bench: support https (#3968 Yuan Zhou)
- rgw: add max multipart upload parts (#12146 Abhishek Dixit)
- rgw: add Transaction-Id to response (Abhishek Dixit)
- rgw: document layout of pools and objects (Pete Zaitcev)
- rgw: do not preserve ACLs when copying object (#12370 Yehuda Sadeh)
- rgw: fix Connection: header handling (#12298 Wido den Hollander)
- rgw: fix data corruptions race condition (#11749 Wuxingyi)
- rgw: fix JSON response when getting user quota (#12117 Wuxingyi)
- rgw: force content_type for swift bucket stats requests (#12095 Orit Wasserman)
- rgw: improved support for swift account metadata (Radoslaw Zarzynski)
- rgw: make max put size configurable (#6999 Yuan Zhou)
- rgw: orphan detection tool (Yehuda Sadeh)
- rgw: swift: do not override sent content type (#12363 Orit Wasserman)
- rgw: swift: set Content-Length for account GET (#12158 Radoslav Zarzynski)
- rpm: always rebuild and install man pages for rpm (Owen Synge)
- rpm: misc fixes (Boris Ranto, Owen Synge, Ken Dreyer, Ira Cooper)
- systemd: logrotate fixes (Tim Seron, Lars Marowsky-Bree, Nathan Cutler)
- sysvinit compat: misc fixes (Owen Synge)
- test: misc fs test improvements (John Spray, Loic Dachary)
- test: python tests, linter cleanup (Alfredo Deza)

V9.0.2

This development release features more of the OSD work queue unification, randomized osd scrub times, a huge pile of librbd fixes, more MDS repair and snapshot fixes, and a significant amount of work on the tests and build infrastructure.

NOTABLE CHANGES

- buffer: some cleanup (Michal Jarzabek)
- build: cmake: fix nss linking (Danny Al-Gaaf)
- build: cmake: misc fixes (Orit Wasserman, Casey Bodley)
- build: install-deps: misc fixes (Loic Dachary)
- build: make_dist_tarball.sh (Sage Weil)
- ceph-detect-init: added Linux Mint (Michal Jarzabek)
- ceph-detect-init: robust init system detection (Owen Synge, Loic Dachary)
- ceph-disk: ensure ‘zap’ only operates on a full disk (#11272 Loic Dachary)
- ceph-disk: misc fixes to respect init system (Loic Dachary, Owen Synge)
- ceph-disk: support NVMe device partitions (#11612 Ilja Slepnev)
- ceph: fix ‘df’ units (Zhe Zhang)
- ceph: fix parsing in interactive cli mode (#11279 Kefu Chai)
- ceph-objectstore-tool: many many changes (David Zafman)
- ceph-post-file: misc fixes (Joey McDonald, Sage Weil)
- client: avoid sending unnecessary FLUSHSNAP messages (Yan, Zheng)
- client: exclude setfilelock when calculating oldest tid (Yan, Zheng)
- client: fix error handling in check_pool_perm (John Spray)
- client: fsync waits only for inode’s caps to flush (Yan, Zheng)
- client: invalidate kernel dcache when cache size exceeds limits (Yan, Zheng)
- client: make fsync wait for unsafe dir operations (Yan, Zheng)
- client: pin lookup dentry to avoid inode being freed (Yan, Zheng)
- common: detect overflow of int config values (#11484 Kefu Chai)
- common: fix json parsing of utf8 (#7387 Tim Serong)
- common: fix leak of pthread_mutexattr (#11762 Ketor Meng)
- crush: respect default replicated ruleset config on map creation (Ilya Dryomov)
- deb, rpm: move ceph-objectstore-tool to ceph (Ken Dreyer)
- doc: man page updates (Kefu Chai)
- doc: misc updates (#11396 Nilamdyuti, Francois Lafont, Ken Dreyer, Kefu Chai)
- init-radosgw: merge with sysv version; fix enumeration (Sage Weil)
- librados: add config observer (Alistair Strachan)
- librbd: add const for single-client-only features (Josh Durgin)
- librbd: add deep-flatten operation (Jason Dillaman)
- librbd: avoid blocking aio API methods (#11056 Jason Dillaman)
- librbd: fix fast diff bugs (#11553 Jason Dillaman)
- librbd: fix image format detection (Zhiqiang Wang)
- librbd: fix lock ordering issue (#11577 Jason Dillaman)
- librbd: flatten/copyup fixes (Jason Dillaman)
- librbd: lockdep, helgrind validation (Jason Dillaman, Josh Durgin)
- librbd: only update image flags while hold exclusive lock (#11791 Jason Dillaman)
- librbd: return result code from close (#12069 Jason Dillaman)
- librbd: tolerate old osds when getting image metadata (#11549 Jason Dillaman)
- mds: do not add snapped items to bloom filter (Yan, Zheng)
- mds: fix handling for missing mydir dirfrag (#11641 John Spray)
- mds: fix rejoin (Yan, Zheng)
- mds: fix stra reintegration (Yan, Zheng)
- mds: fix suicide beason (John Spray)
- mds: misc repair improvements (John Spray)
- mds: misc snapshot fixes (Yan, Zheng)
- mds: respawn instead of suicide on blacklist (John Spray)
- misc coverity fixes (Danny Al-Gaaf)
- mon: add ‘mon_metadata <id>’ command (Kefu Chai)
- mon: add ‘node ls ...’ command (Kefu Chai)
- mon: disallow ec pools as tiers (#11650 Samuel Just)
- mon: fix mds beacon replies (#11590 Kefu Chai)
- mon: fix ‘pg ls’ sort order, state names (#11569 Kefu Chai)
- mon: normalize erasure-code profile for storage and comparison (Loic Dachary)
- mon: optionally specify osd id on ‘osd create’ (Mykola Golub)
- mon: ‘osd tree’ fixes (Kefu Chai)
- mon: prevent pool with snapshot state from being used as a tier (#11493 Sage Weil)
- mon: refine check_remove_tier checks (#11504 John Spray)
- mon: remove spurious who arg from ‘mds rm ...’ (John Spray)
- msgr: async: misc fixes (Haomai Wang)

- msgr: xio: fix ip and nonce (Raju Kurunkad)
- msgr: xio: improve lane assignment (Vu Pham)
- msgr: xio: misc fixes (Vu Pham, Cosey Bodley)
- osd: avoid transaction append in some cases (Sage Weil)
- osdc/Objecter: allow per-pool calls to op_cancel_writes (John Spray)
- osd: elminiate txn apend, ECSubWrite copy (Samuel Just)
- osd: filejournal: cleanup (David Zafman)
- osd: fix check_for_full (Henry Chang)
- osd: fix dirty accounting in make_writeable (Zhiqiang Wang)
- osd: fix osdmap dump of blacklist items (John Spray)
- osd: fix snap flushing from cache tier (again) (#11787 Samuel Just)
- osd: fix snap handling on promotion (#11296 Sam Just)
- osd: handle log split with overlapping entries (#11358 Samuel Just)
- osd: keyvaluestore: misc fixes (Varada Kari)
- osd: make suicide timeouts individually configurable (Samuel Just)
- osd: move scrub in OpWQ (Samuel Just)
- osd: pool size change triggers new interval (#11771 Samuel Just)
- osd: randomize scrub times (#10973 Kefu Chai)
- osd: refactor scrub and digest recording (Sage Weil)
- osd: refuse first write to EC object at non-zero offset (Jianpeng Ma)
- osd: stripe over small xattrs to fit in XFS's 255 byte inline limit (Sage Weil, Ning Yao)
- osd: sync object_map on syncfs (Samuel Just)
- osd: take excl lock of op is rw (Samuel Just)
- osd: WBThrottle cleanups (Jianpeng Ma)
- pycephfs: many fixes for bindings (Haomai Wang)
- rados: bench: add -no-verify option to improve performance (Piotr Dalek)
- rados: misc bench fixes (Dmitry Yatsushkevich)
- rbd: add disk usage tool (#7746 Jason Dillaman)
- rgw: alwasy check if token is expired (#11367 Anton Aksola, Riku Lehto)
- rgw: conversion tool to repair broken multipart objects (#12079 Yehuda Sadeh)
- rgw: do not enclose bucket header in quotes (#11860 Wido den Hollander)
- rgw: error out if frontend did not send all data (#11851 Yehuda Sadeh)
- rgw: fix assignment of copy obj attributes (#11563 Yehuda Sadeh)
- rgw: fix reset_loc (#11974 Yehuda Sadeh)
- rgw: improve content-length env var handling (#11419 Robin H. Johnson)
- rgw: only scan for objects not in a namespace (#11984 Yehuda Sadeh)
- rgw: remove trailing :port from HTTP_HOST header (Sage Weil)
- rgw: shard work over multiple librados instances (Pavan Rallabhandi)
- rgw: swift: enforce Content-Type in response (#12157 Radoslaw Zarzynski)
- rgw: use attrs from source bucket on copy (#11639 Javier M. Mellid)
- rocksdb: pass options as single string (Xiaoxi Chen)
- rpm: many spec file fixes (Owen Synge, Ken Dreyer)
- tests: fixes for rbd xtests (Douglas Fuller)
- tests: fix tiering health checks (Loic Dachary)
- tests for low-level performance (Haomai Wang)
- tests: many ec non-regression improvements (Loic Dachary)
- tests: many many ec test improvements (Loic Dachary)
- upstart: throttle restarts (#11798 Sage Weil, Greg Farnum)

V9.0.1

This development release is delayed a bit due to tooling changes in the build environment. As a result the next one (v9.0.2) will have a bit more work than is usual.

Highlights here include lots of RGW Swift fixes, RBD feature work surrounding the new object map feature, more CephFS snapshot fixes, and a few important CRUSH fixes.

NOTABLE CHANGES

- auth: cache/reuse crypto lib key objects, optimize msg signature check (Sage Weil)
- build: allow tcmalloc-minimal (Thorsten Behrens)
- build: do not build ceph-dencoder with tcmalloc (#10691 Boris Ranto)
- build: fix pg ref disabling (William A. Kennington III)

- build: install-deps.sh improvements (Loic Dachary)
- build: misc fixes (Boris Ranto, Ken Dreyer, Owen Synge)
- ceph-authtool: fix return code on error (Gerhard Muntingh)
- ceph-disk: fix zap sgdisk invocation (Owen Synge, Thorsten Behrens)
- ceph-disk: pass -cluster arg on prepare subcommand (Kefu Chai)
- ceph-fuse, libcephfs: drop inode when rmdir finishes (#11339 Yan, Zheng)
- ceph-fuse,libcephfs: fix uninode (#11356 Yan, Zheng)
- ceph-monstore-tool: fix store-copy (Huangjun)
- common: add perf counter descriptions (Alyona Kiseleva)
- common: fix throttle max change (Henry Chang)
- crush: fix crash from invalid ‘take’ argument (#11602 Shiva Rkreddy, Sage Weil)
- crush: fix divide-by-2 in straw2 (#11357 Yann Dupont, Sage Weil)
- deb: fix rest-bench-dbg and ceph-test-dbg dependendies (Ken Dreyer)
- doc: document region hostnames (Robin H. Johnson)
- doc: update release schedule docs (Loic Dachary)
- init-radosgw: run radosgw as root (#11453 Ken Dreyer)
- librados: fadvise flags per op (Jianpeng Ma)
- librbd: allow additional metadata to be stored with the image (Haomai Wang)
- librbd: better handling for dup flatten requests (#11370 Jason Dillaman)
- librbd: cancel in-flight ops on watch error (#11363 Jason Dillaman)
- librbd: default new images to format 2 (#11348 Jason Dillaman)
- librbd: fast diff implementation that leverages object map (Jason Dillaman)
- librbd: fix snapshot creation when other snap is active (#11475 Jason Dillaman)
- librbd: new diff_iterate2 API (Jason Dillaman)
- librbd: object map rebuild support (Jason Dillaman)
- logrotate.d: prefer service over invoke-rc.d (#11330 Win Hierman, Sage Weil)
- mds: avoid getting stuck in XLOCKDONE (#11254 Yan, Zheng)
- mds: fix integer truncateion on large client ids (Henry Chang)
- mds: many snapshot and stray fixes (Yan, Zheng)
- mds: persist completed_requests reliably (#11048 John Spray)
- mds: separate safe_pos in Journaler (#10368 John Spray)
- mds: snapshot rename support (#3645 Yan, Zheng)
- mds: warn when clients fail to advance oldest_client_tid (#10657 Yan, Zheng)
- misc cleanups and fixes (Danny Al-Gaaf)
- mon: fix average utilization calc for ‘osd df’ (Mykola Golub)
- mon: fix variance calc in ‘osd df’ (Sage Weil)
- mon: improve callout to crushtool (Mykola Golub)
- mon: prevent bucket deletion when referenced by a crush rule (#11602 Sage Weil)
- mon: prime pg_temp when CRUSH map changes (Sage Weil)
- monclient: flush_log (John Spray)
- msgr: async: many many fixes (Haomai Wang)
- msgr: simple: fix clear_pipe (#11381 Haomai Wang)
- osd: add latency perf counters for tier operations (Xinze Chi)
- osd: avoid multiple hit set insertions (Zhiqiang Wang)
- osd: break PG removal into multiple iterations (#10198 Guang Yang)
- osd: check scrub state when handling map (Jianpeng Ma)
- osd: fix endless repair when object is unrecoverable (Jianpeng Ma, Kefu Chai)
- osd: fix pg resurrection (#11429 Samuel Just)
- osd: ignore non-existent osds in unfound calc (#10976 Mykola Golub)
- osd: increase default max open files (Owen Synge)
- osd: prepopulate needs_recovery_map when only one peer has missing (#9558 Guang Yang)
- osd: relax reply order on proxy read (#11211 Zhiqiang Wang)
- osd: skip promotion for flush/evict op (Zhiqiang Wang)
- osd: write journal header on clean shutdown (Xinze Chi)
- qa: run-make-check.sh script (Loic Dachary)
- rados bench: misc fixes (Dmitry Yatsushkevich)
- rados: fix error message on failed pool removal (Wido den Hollander)
- radosgw-admin: add ‘bucket check’ function to repair bucket index (Yehuda Sadeh)
- rbd: allow unmapping by spec (Ilya Dryomov)
- rbd: deprecate -new-format option (Jason Dillman)
- rgw: do not set content-type if length is 0 (#11091 Orit Wasserman)
- rgw: don’t use end_marker for namespaced object listing (#11437 Yehuda Sadeh)
- rgw: fail if parts not specified on multipart upload (#11435 Yehuda Sadeh)
- rgw: fix GET on swift account when limit == 0 (#10683 Radoslaw Zarzynski)
- rgw: fix broken stats in container listing (#11285 Radoslaw Zarzynski)

- rgw: fix bug in domain/subdomain splitting (Robin H. Johnson)
- rgw: fix civetweb max threads (#10243 Yehuda Sadeh)
- rgw: fix copy metadata, support X-Copied-From for swift (#10663 Radoslaw Zarzynski)
- rgw: fix locator for objects starting with _ (#11442 Yehuda Sadeh)
- rgw: fix multipart upload in retry path (#11604 Yehuda Sadeh)
- rgw: fix quota enforcement on POST (#11323 Sergey Arkhipov)
- rgw: fix return code on missing upload (#11436 Yehuda Sadeh)
- rgw: force content type header on responses with no body (#11438 Orit Wasserman)
- rgw: generate new object tag when setting attrs (#11256 Yehuda Sadeh)
- rgw: issue aio for first chunk before flush cached data (#11322 Guang Yang)
- rgw: make read user buckets backward compat (#10683 Radoslaw Zarzynski)
- rgw: merge manifests properly with prefix override (#11622 Yehuda Sadeh)
- rgw: return 412 on bad limit when listing buckets (#11613 Yehuda Sadeh)
- rgw: send ETag, Last-Modified for swift (#11087 Radoslaw Zarzynski)
- rgw: set content length on container GET, PUT, DELETE, HEAD (#10971, #11036 Radoslaw Zarzynski)
- rgw: support end marker on swift container GET (#10682 Radoslaw Zarzynski)
- rgw: swift: fix account listing (#11501 Radoslaw Zarzynski)
- rgw: swift: set content-length on keystone tokens (#11473 Herv Rousseau)
- rgw: use correct oid for gc chains (#11447 Yehuda Sadeh)
- rgw: use unique request id for civetweb (#10295 Orit Wasserman)
- rocksdb, leveldb: fix compact_on_mount (Xiaoxi Chen)
- rocksdb: add perf counters for get/put latency (Xinxin Shu)
- rpm: add suse firewall files (Tim Serong)
- rpm: misc systemd and suse fixes (Owen Synge, Nathan Cutler)

V9.0.0

This is the first development release for the Infernalis cycle, and the first Ceph release to sport a version number from the new numbering scheme. The “9” indicates this is the 9th release cycle-I (for Infernalis) is the 9th letter. The first “0” indicates this is a development release (“1” will mean release candidate and “2” will mean stable release), and the final “0” indicates this is the first such development release.

A few highlights include:

- a new ‘ceph daemonperf’ command to watch perfcounter stats in realtime
- reduced MDS memory usage
- many MDS snapshot fixes
- librbd can now store options in the image itself
- many fixes for RGW Swift API support
- OSD performance improvements
- many doc updates and misc bug fixes

NOTABLE CHANGES

- aarch64: add optimized version of crc32c (Yazen Ghannam, Steve Capper)
- auth: reinit NSS after fork() (#11128 Yan, Zheng)
- build: disable LTTNG by default (#11333 Josh Durgin)
- build: fix ppc build (James Page)
- build: install-deps: support OpenSUSE (Loic Dachary)
- build: misc cmake fixes (Matt Benjamin)
- ceph-disk: follow ceph-osd hints when creating journal (#9580 Sage Weil)
- ceph-disk: handle re-using existing partition (#10987 Loic Dachary)
- ceph-disk: improve parted output parsing (#10983 Loic Dachary)
- ceph-disk: make suppression work for activate-all and activate-journal (Dan van der Ster)
- ceph-disk: misc fixes (Alfredo Deza)
- ceph-fuse, libcephfs: don’t clear COMPLETE when trimming null (Yan, Zheng)
- ceph-fuse, libcephfs: hold exclusive caps on dirs we “own” (#11226 Greg Farnum)
- ceph-fuse: do not require successful remount when unmounting (#10982 Greg Farnum)
- ceph: new ‘ceph daemonperf’ command (John Spray, Mykola Golub)
- common: PriorityQueue tests (Kefu Chai)
- common: add descriptions to perfcounters (Kiseleva Alyona)
- common: fix LTTNG vs fork issue (Josh Durgin)
- crush: fix has_v4_buckets (#11364 Sage Weil)
- crushtool: fix order of operations, usage (Sage Weil)

- debian: minor package reorg (Ken Dreyer)
- doc: docuemnt object corpus generation (#11099 Alexis Normand)
- doc: fix gender neutrality (Alexandre Maragone)
- doc: fix install doc (#10957 Kefu Chai)
- doc: fix sphinx issues (Kefu Chai)
- doc: mds data structure docs (Yan, Zheng)
- doc: misc updates (Nilamdyuti Goswami, Vartika Rai, Florian Haas, Loic Dachary, Simon Guinot, Andy Allan, Alistair Israel, Ken Dreyer, Robin Rehu, Lee Revell, Florian Marsylle, Thomas Johnson, Bosse Klykken, Travis Rhoden, Ian Kelling)
- doc: swift tempurls (#10184 Abhishek Lekshmanan)
- doc: switch doxygen integration back to breathe (#6115 Kefu Chai)
- erasure-code: update ISA-L to 2.13 (Yuan Zhou)
- gmock: switch to submodule (Danny Al-Gaaf, Loic Dachary)
- hadoop: add terasort test (Noah Watkins)
- java: fix libcephfs bindings (Noah Watkins)
- libcephfs,ceph-fuse: fix request resend on cap reconnect (#10912 Yan, Zheng)
- librados: define C++ flags from C constants (Josh Durgin)
- librados: fix last_force_resent handling (#11026 Jianpeng Ma)
- librados: fix memory leak from C_TwoContexts (Xiong Yiliang)
- librados: fix striper when stripe_count = 1 and stripe_unit != object_size (#11120 Yan, Zheng)
- librados: op perf counters (John Spray)
- librados: pybind: fix write() method return code (Javier Guerra)
- libradosstriper: fix leak (Danny Al-Gaaf)
- librbd: add purge_on_error cache behavior (Jianpeng Ma)
- librbd: misc aio fixes (#5488 Jason Dillaman)
- librbd: misc rbd fixes (#11478 #11113 #11342 #11380 Jason Dillaman, Zhiqiang Wang)
- librbd: readahead fixes (Zhiqiang Wang)
- librbd: store metadata, including config options, in image (Haomai Wang)
- mds: add ‘damaged’ state to MDSMap (John Spray)
- mds: add nicknames for perfcounters (John Spray)
- mds: disable problematic rstat propagation into snap parents (Yan, Zheng)
- mds: fix mydir replica issue with shutdown (#10743 John Spray)
- mds: fix out-of-order messages (#11258 Yan, Zheng)
- mds: fix shutdown with strays (#10744 John Spray)
- mds: fix snapshot fixes (Yan, Zheng)
- mds: fix stray handling (John Spray)
- mds: flush immediately in do_open_truncate (#11011 John Spray)
- mds: improve dump methods (John Spray)
- mds: misc journal cleanups and fixes (#10368 John Spray)
- mds: new SessionMap storage using omap (#10649 John Spray)
- mds: reduce memory consumption (Yan, Zheng)
- mds: throttle purge stray operations (#10390 John Spray)
- mds: tolerate clock jumping backwards (#11053 Yan, Zheng)
- misc coverity fixes (Danny Al-Gaaf)
- mon: do not deactivate last mds (#10862 John Spray)
- mon: make osd get pool ‘all’ only return applicable fields (#10891 Michal Jarzabek)
- mon: warn on bogus cache tier config (Jianpeng Ma)
- msg/async: misc bug fixes and updates (Haomai Wang)
- msg/simple: fix connect_seq assert (Haomai Wang)
- msg/xio: misc fixes (#10735 Matt Benjamin, Kefu Chai, Danny Al-Gaaf, Raju Kurunkad, Vu Pham)
- msg: unit tests (Haomai Wang)
- objectcacher: misc bug fixes (Jianpeng Ma)
- os/filestore: enlarge getattr buffer size (Jianpeng Ma)
- osd: EIO injection (David Zhang)
- osd: add misc perfcounters (Xinze Chi)
- osd: add simple sleep injection in recovery (Sage Weil)
- osd: allow SEEK_HOLE/SEEK_DATA for sparse read (Zhiqiang Wang)
- osd: avoid dup omap sets for in pg metadata (Sage Weil)
- osd: clean up some constness, privateness (Kefu Chai)
- osd: erasure-code: drop entries according to LRU (Andreas-Joachim Peters)
- osd: fix negative degraded stats during backfill (Guang Yang)
- osd: misc fixes (Ning Yao, Kefu Chai, Xinze Chi, Zhiqiang Wang, Jianpeng Ma)
- pybind: pep8 cleanups (Danny Al-Gaaf)
- qa: fix filelock_interrupt.py test (Yan, Zheng)
- qa: improve ceph-disk tests (Loic Dachary)
- qa: improve docker build layers (Loic Dachary)

- rados: translate erno to string in CLI (#10877 Kefu Chai)
- rbd: accept map options config option (Ilya Dryomov)
- rbd: cli: fix arg parsing with -io-pattern (Dmitry Yatsushkevich)
- rbd: fix error messages (#2862 Rajesh Nambiar)
- rbd: update rbd man page (Ilya Dryomov)
- rbd: update xfstests tests (Douglas Fuller)
- rgw: add X-Timestamp for Swift containers (#10938 Radoslaw Zarzynski)
- rgw: add missing headers to Swift container details (#10666 Ahmad Faheem, Dmytro Iurchenko)
- rgw: add stats to headers for account GET (#10684 Yuan Zhou)
- rgw: do not prefech data for HEAD requests (Guang Yang)
- rgw: don't clobber bucket/object owner when setting ACLs (#10978 Yehuda Sadeh)
- rgw: don't use rgw_socket_path if frontend is configured (#11160 Yehuda Sadeh)
- rgw: enforce Content-Lenth for POST on Swift cont/obj (#10661 Radoslaw Zarzynski)
- rgw: fix handling empty metadata items on Swift container (#11088 Radoslaw Zarzynski)
- rgw: fix log rotation (Wuxingyi)
- rgw: generate Date header for civetweb (#10873 Radoslaw Zarzynski)
- rgw: make init script wait for radosgw to stop (#11140 Dmitry Yatsushkevich)
- rgw: make quota/gc threads configurable (#11047 Guang Yang)
- rgw: pass in civetweb configurables (#10907 Yehuda Sadeh)
- rgw: rectify 202 Accepted in PUT response (#11148 Radoslaw Zarzynski)
- rgw: remove meta file after deleting bucket (#11149 Orit Wasserman)
- rgw: swift: allow setting attributes with COPY (#10662 Ahmad Faheem, Dmytro Iurchenko)
- rgw: swift: fix metadata handling on copy (#10645 Radoslaw Zarzynski)
- rgw: swift: send Last-Modified header (#10650 Radoslaw Zarzynski)
- rgw: update keystone cache with token info (#11125 Yehuda Sadeh)
- rgw: update to latest civetweb, enable config for IPv6 (#10965 Yehuda Sadeh)
- rocksdb: update to latest (Xiaoxi Chen)
- rpm: loosen ceph-test dependencies (Ken Dreyer)

V0.94.10 HAMMER

This Hammer point release fixes several bugs and adds two new features.

We recommend that all hammer v0.94.x users upgrade.

For more detailed information, see [the complete changelog](#).

NEW FEATURES

ceph-objectstore-tool and ceph-monstore-tool now enable user to rebuild the monitor database from OSDs. (This feature is especially useful when all monitors fail to boot due to leveldb corruption.)

In RADOS Gateway, it is now possible to reshards an existing bucket's index using an off-line tool.

Usage:

```
$ radosgw-admin bucket reshards -bucket=<bucket_name> -num_shards=<num_shards>
```

This will create a new linked bucket instance that points to the newly created index objects. The old bucket instance still exists and currently it's up to the user to manually remove the old bucket index objects. (Note that bucket resharding currently requires that all IO (especially writes) to the specific bucket is quiesced.)

OTHER NOTABLE CHANGES

- build/ops: ceph-create-keys loops forever ([issue#17753](#), [pr#12805](#), Alfredo Deza)
- build/ops: improve ceph.in error message ([issue#11101](#), [pr#10905](#), Kefu Chai)
- build/ops: make stop.sh more portable ([issue#16918](#), [pr#10569](#), Mykola Golub)
- build/ops: remove SYSTEMD_RUN from initscript ([issue#16440](#), [issue#7627](#), [pr#9873](#), Vladislav Odintsov)
- cephx: Fix multiple segfaults due to attempts to encrypt or decrypt ([issue#16266](#), [pr#11930](#), Brad Hubbard)
- common: SIGABRT in TrackedOp::dump() via dump_ops_in_flight() ([issue#8885](#), [pr#12121](#), Jianpeng Ma, Zhiqiang Wang, David Zafman)
- common: os/ObjectStore: fix _update_op for split dest_cid ([issue#15345](#), [pr#12071](#), Sage Weil)
- crush: reset bucket->h.items[i] when removing tree item ([issue#16525](#), [pr#10724](#), Kefu Chai)
- doc: add "Upgrading to Hammer" section ([issue#17386](#), [pr#11372](#), Kefu Chai)

- doc: add orphan options to radosgw-admin -help and man page ([issue#17281](#), [issue#17280](#), pr#11140, Abhishek Lekshmanan, Casey Bodley, Ken Dreyer, Thomas Serlin)
- doc: clarify that RGW bucket object versioning is supported ([issue#16574](#), pr#10437, Yuan Zhou, shawn chen)
- librados: bad flags can crash the osd ([issue#16012](#), pr#11936, Jianpeng Ma, Sage Weil)
- librbd: ceph 10.2.2 rbd status on image format 2 returns "(2) No such file or directory" ([issue#16887](#), pr#10987, Jason Dillaman)
- librbd: diffs to clone's first snapshot should include parent diffs ([issue#18068](#), pr#12446, Jason Dillaman)
- librbd: image.stat() call in librbdpy fails sometimes ([issue#17310](#), pr#11949, Jason Dillaman)
- librbd: request exclusive lock if current owner cannot execute op ([issue#16171](#), pr#12018, Mykola Golub)
- mds: fix cephfs-java ftruncate unit test failure ([issue#11258](#), pr#11939, Yan, Zheng)
- mon: %USED of ceph df is wrong ([issue#16933](#), pr#11934, Kefu Chai)
- mon: MonmapMonitor should return success when MON will be removed ([issue#17725](#), pr#12006, Joao Eduardo Luis)
- mon: OSDMonitor: Missing nearfull flag set ([issue#17390](#), pr#11273, Igor Podoski)
- mon: OSDs marked OUT wrongly after monitor failover ([issue#17719](#), pr#11946, Dong Wu)
- mon: fix memory leak in prepare_beacon ([issue#17285](#), pr#10238, Igor Podoski)
- mon: osd flag health message is misleading ([issue#18175](#), pr#12687, Sage Weil)
- mon: prepare_pgtemp needs to only update up_thru if newer than the existing one ([issue#16185](#), pr#11937, Samuel Just)
- mon: return size_t from MonitorDBStore::Transaction::size() ([issue#14217](#), pr#10904, Kefu Chai)
- mon: send updated monmap to its subscribers ([issue#17558](#), pr#11457, Kefu Chai)
- msgr: OpTracker needs to release the message throttle in _unregistered ([issue#14248](#), pr#11938, Samuel Just)
- msgr: simple/Pipe: error decoding addr ([issue#18072](#), pr#12266, Sage Weil)
- osd: PG::update_calc_stats wrong for CRUSH_ITEM_NONE up set items ([issue#16998](#), pr#11933, Samuel Just)
- osd: PG::choose_acting valgrind error or ./common/hobject.h: 182: FAILED assert(!max || (*this == hobject_t(hobject_t::get_max()))) ([issue#13967](#), pr#11932, Tao Chang)
- osd: ReplicatedBackend::build_push_op: add a second config to limit omap entries/chunk independently of object data ([issue#16128](#), pr#12417, Wanlong Gao)
- osd: crash on EIO during deep-scrubbing ([issue#16034](#), pr#11935, Nathan Cutler)
- osd: filestore: FALLOC_FL_PUNCH_HOLE must be used with FALLOC_FL_KEEP_SIZE ([issue#18446](#), pr#13041, xinxin shu)
- osd: fix cached_removed_snaps bug in PGPool::update after map gap ([issue#18628](#), [issue#15943](#), pr#12906, Samuel Just)
- osd: fix collection_list shadow return value ([issue#17713](#), pr#11927, Haomai Wang)
- osd: fix fiemap issue in xfs when #extents > 1364 ([issue#17610](#), pr#11615, Kefu Chai, Ning Yao)
- osd: update PGPool to detect map gaps and reset cached_removed_snaps ([issue#15943](#), pr#11676, Samuel Just)
- rbd: export diff should open image as read-only ([issue#17671](#), pr#11948, liyankun)
- rbd: fix parameter check ([issue#18237](#), pr#12312, Yankun Li)
- rbd: fix possible rbd data corruption ([issue#16002](#), pr#11618, Yan, Zheng, Greg Farnum)
- rgw: Anonymous user is able to read bucket with authenticated read ACL ([issue#13207](#), pr#11045, rahul.1aggarwal@gmail.com)
- rgw: COPY broke multipart files uploaded under dumpling ([issue#16435](#), pr#11950, Yehuda Sadeh)
- rgw: TempURL in radosgw behaves now like its Swift's counterpart. ([issue#18316](#), pr#12619, Radoslaw Zarzynski)
- rgw: default quota fixes ([issue#16410](#), pr#10839, Pavan Rallabandi, Daniel Gryniewicz)
- rgw: do not abort when accept a CORS request with short origin ([issue#18187](#), pr#12398, LiuYang)
- rgw: do not omap_getvals with (u64)-1 max ([issue#17985](#), pr#12418, Yehuda Sadeh, Sage Weil)
- rgw: fix crash when client posts object with null condition ([issue#17635](#), pr#11809, Yehuda Sadeh)
- rgw: fix inconsistent uid/email handling in radosgw-admin ([issue#13598](#), pr#11952, Matt Benjamin)
- rgw: implement offline resharding command ([issue#17745](#), pr#12227, Yehuda Sadeh, Orit Wasserman, weiqiaomiao)
- rgw: swift: ranged request on a DLO provides wrong values in Content-Range HTTP header ([issue#13452](#), pr#11951, Radoslaw Zarzynski)
- rgw: the value of total_time is wrong in the result of 'radosgw-admin log show' opt ([issue#17598](#), pr#11899, weiqiaomiao)
- tests: Cannot clone ceph/s3-tests.git (missing branch) ([issue#18384](#), pr#12744, Orit Wasserman)
- tests: Cannot reserve CentOS 7.2 smithi machines ([issue#18401](#), pr#12762, Nathan Cutler)
- tests: OSDs commit suicide in rbd suite when testing on btrfs ([issue#18397](#), pr#12758, Nathan Cutler)
- tests: Workunits needlessly wget from git.ceph.com ([issue#18336](#), [issue#18271](#), [issue#18388](#), pr#12685, Sage Weil, Nathan Cutler)
- tests: cephfs test failures (ceph.com/qa is broken, should be download.ceph.com/qa) ([issue#18574](#), pr#13022, John Spray)
- tests: merge ceph-qa-suite (pr#12455, Sage Weil)
- tests: objecter_requests workunit fails on wip branches ([issue#18393](#), pr#12759, Sage Weil)
- tests: populate mnt_point in qa/tasks/ceph.py ([issue#18383](#), pr#12743, Nathan Cutler)
- tests: qemu/tests/qemu-iotests/077 fails in dumpling, hammer, and jewel ([issue#10773](#), pr#12423, Jason Dillaman)
- tests: run fs/thrash on xfs instead of btrfs ([issue#17151](#), pr#13039, Nathan Cutler)
- tests: update Ubuntu image url after ceph.com refactor ([issue#18542](#), pr#12957, Jason Dillaman)
- tests: update rbd/singleton/all/formatted-output.yaml to support ceph-ci * ([issue#18440](#), pr#12824 *, Venky Shankar,

Nathan Cutler)

- tools: add a tool to rebuild mon store from OSD ([issue#17179](#), [issue#17400](#), [pr#11125](#), Kefu Chai, xie xingguo)
- tools: ceph-objectstore-tool crashes if -journal-path <a-directory> ([issue#17307](#), [pr#11929](#), Kefu Chai)
- tools: ceph-objectstore-tool: add a way to split filestore directories offline ([issue#17220](#), [pr#11253](#), Josh Durgin)
- tools: crushtool -compile generates output despite missing item ([issue#17306](#), [pr#11931](#), Kefu Chai)

V0.94.9 HAMMER

This Hammer point release fixes a build issue present in 0.94.8 that prevented us from generating packages for Ubuntu Precise and CentOS 6.x.

We recommend all users of v0.94.7 or older upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build/ops: revert: boost uuid makes valgrind complain ([pr#10913](#), Sage Weil)

V0.94.8 HAMMER

This Hammer point release fixes several bugs.

We recommend that all hammer v0.94.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build/ops: rocksdb do not link against tcmalloc if it's disabled ([issue#14799](#), [pr#10750](#), Sage Weil, Kefu Chai)
- build/ops: Add -D_LARGEFILE64_SOURCE to Linux build. ([issue#16611](#), [pr#10182](#), Ira Cooper)
- build/ops: boost uuid makes valgrind complain ([issue#12736](#), [pr#9741](#), Sage Weil, Rohan Mars)
- build/ops: ceph-disk s/by-parttype-uuid/by-parttypeuuid/ ([issue#15867](#), [pr#9107](#), Nathan Cutler)
- common: add units to rados bench output and clean up formatting ([issue#12248](#), [pr#8960](#), Dmitry Yatsushkevich, Brad Hubbard, Gu Zhongyan)
- common: config set with negative value results in "error setting 'filestore_merge_threshold' to '-40': (22) Invalid argument" ([issue#13829](#), [pr#10291](#), Brad Hubbard, Kefu Chai)
- common: linking to -lrbd causes process startup times to balloon ([issue#15225](#), [pr#8538](#), Richard W.M. Jones)
- doc: fix by-parttypeuuid in ceph-disk(8) nroff ([issue#15867](#), [pr#10699](#), Ken Dreyer)
- fs: double decreased the count to trim caps which will cause failing to respond to cache pressure ([issue#14319](#), [pr#8804](#), Zhi Zhang)
- log: do not repeat errors to stderr ([issue#14616](#), [pr#10227](#), Sage Weil)
- mds: failing file operations on kernel based cephfs mount point leaves unaccessible file behind on hammer 0.94.7 ([issue#16013](#), [pr#10198](#), Yan, Zheng)
- mds: fix stray purging in 'stripe_count > 1' case ([issue#15050](#), [pr#8042](#), Yan, Zheng)
- mds: wrongly treat symlink inode as normal file/dir when symlink inode is stale on kcephfs ([issue#15702](#), [pr#9404](#), Zhi Zhang)
- mon: LibRadosMiscConnectFailure.ConnectFailure (not so intermittent) failure in upgrade/hammer-x ([issue#13992](#), [pr#8806](#), Sage Weil)
- mon: Monitor: validate prefix on handle_command() ([issue#16297](#), [pr#10038](#), You Ji)
- mon: drop pg temps from not the current primary in OSDMonitor ([issue#16127](#), [pr#9893](#), Samuel Just)
- mon: fix calculation of %USED ([issue#15641](#), [pr#9125](#), Ruifeng Yang, David Zafman)
- mon: improve reweight_by_utilization() logic ([issue#15686](#), [pr#9416](#), xie xingguo)
- mon: pool quota alarm is not in effect ([issue#15478](#), [pr#8593](#), Danny Al-Gaaf)
- mon: wrong ceph get mdsmap assertion ([issue#14681](#), [pr#7542](#), Vicente Cheng)
- msgr: ceph-osd valgrind invalid reads/writes ([issue#15870](#), [pr#9238](#), Samuel Just)
- objecter: LibRadosWatchNotifyPPTests/LibRadosWatchNotifyPP.WatchNotify2Timeout/1 segv ([issue#15760](#), [pr#9400](#), Sage Weil)
- osd: OSD reporting ENOTEMPTY and crashing ([issue#14766](#), [pr#9277](#), Samuel Just)
- osd: When generating past intervals due to an import end at pg epoch and fix build_past_intervals_parallel ([issue#12387](#), [issue#14438](#), [pr#8464](#), David Zafman)
- osd: acting_primary not updated on split ([issue#15523](#), [pr#9001](#), Sage Weil)

- osd: assert(!actingbackfill.empty()): old watch timeout tries to queue repop on replica ([issue#15391](#), [pr#8665](#), Sage Weil)
- osd: assert(rollback_info_trimmed_to == head) in PGLog ([issue#13965](#), [pr#8849](#), Samuel Just)
- osd: delete one of the repeated op->mark_started in ReplicatedBackend::sub_op_modify_impl ([issue#16572](#), [pr#9977](#), shun-s)
- osd: fix omap digest compare when scrub ([issue#16000](#), [pr#9271](#), Xinze Chi)
- osd: is_split crash in handle_pg_create ([issue#15426](#), [pr#8805](#), Kefu Chai)
- osd: objects unfound after repair (fixed by repeering the pg) ([issue#15006](#), [pr#7961](#), Jianpeng Ma, Loic Dachary, Kefu Chai)
- osd: rados cppool omap to ec pool crashes osd ([issue#14695](#), [pr#8845](#), Jianpeng Ma)
- osd: remove all stale osdmmaps in handle_osd_map() ([issue#13990](#), [pr#9090](#), Kefu Chai)
- osd: send write and read sub ops on behalf of client ops at normal priority in ECBackend ([issue#14313](#), [pr#8573](#), Samuel Just)
- rbd: snap rollback: restore the link to parent ([issue#14512](#), [pr#8535](#), Alexey Sheplyakov)
- rgw: S3: set EncodingType in ListBucketResult ([issue#15896](#), [pr#8987](#), Victor Makarov, Robin H. Johnson)
- rgw: backport rgwx-copy-if-newer for radosgw-agent ([issue#16262](#), [pr#9671](#), Yehuda Sadeh)
- rgw: bucket listing following object delete is partial ([issue#14826](#), [pr#10555](#), Orit Wasserman)
- rgw: convert plain object to versioned (with null version) when removing ([issue#15243](#), [pr#8755](#), Yehuda Sadeh)
- rgw: fix multi-delete query param parsing. ([issue#16618](#), [pr#10189](#), Robin H. Johnson)
- rgw: have a flavor of bucket deletion to bypass GC and to trigger ([issue#15557](#), [pr#10509](#), Pavan Rallabhandi)
- rgw: keep track of written_objs correctly ([issue#15886](#), [pr#9240](#), Yehuda Sadeh)
- rgw: multipart ListPartsResult has missing quotes on ETag ([issue#15334](#), [pr#8475](#), xie xingguo, Robin H. Johnson)
- rgw: no Last-Modified, Content-Size and X-Object-Manifest headers if no segments in DLO manifest ([issue#15812](#), [pr#9402](#), Radoslaw Zarzynski)
- rgw: radosgw server abort when user passed bad parameters to set quota ([issue#14190](#), [issue#14191](#), [pr#8313](#), Dunrong Huang)
- rgw: radosgw-admin region-map set is not reporting the bucket quota correctly ([issue#16815](#), [pr#10554](#), Yehuda Sadeh, Orit Wasserman)
- rgw: refrain from sending Content-Type/Content-Length for 304 responses ([issue#16327](#), [issue#13582](#), [issue#15119](#), [issue#14005](#), [pr#8379](#), Yehuda Sadeh, Nathan Cutler, Wido den Hollander)
- rgw: remove bucket index objects when deleting the bucket ([issue#16412](#), [pr#10530](#), Orit Wasserman)
- rgw: set Access-Control-Allow-Origin to an asterisk if allowed in a rule ([issue#15348](#), [pr#8528](#), Wido den Hollander)
- rgw: subset of uploaded objects via radosgw are unretrievable when using EC pool ([issue#15745](#), [pr#9407](#), Yehuda Sadeh)
- rgw: subuser rm fails with status 125 ([issue#14375](#), [pr#9961](#), Orit Wasserman)
- rgw: the swift key remains after removing a subuser ([issue#12890](#), [issue#14375](#), [pr#10718](#), Orit Wasserman, Sangdi Xu)
- rgw: user quota may not adjust on bucket removal ([issue#14507](#), [pr#8113](#), Edward Yang)
- tests: be more generous with test timeout ([issue#15403](#), [pr#8470](#), Loic Dachary)
- tests: qa/workunits/rbd: respect RBD_CREATE_ARGS environment variable ([issue#16289](#), [pr#9722](#), Mykola Golub)

V0.94.7 HAMMER

This Hammer point release fixes several minor bugs. It also includes a backport of an improved ‘ceph osd reweight-by-utilization’ command for handling OSDs with higher-than-average utilizations.

We recommend that all hammer v0.94.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- auth: keyring permissions for mon deamon ([issue#14950](#), [pr#8049](#), Owen Syngle)
- auth: PK11_DestroyContext() is called twice if PK11_DigestFinal() fails ([issue#14958](#), [pr#7922](#), Brad Hubbard, Dunrong Huang)
- auth: use libnss more safely ([issue#14620](#), [pr#7488](#), Sage Weil)
- ceph-disk: use blkid instead of sgdisk -i ([issue#14080](#), [issue#14094](#), [pr#7475](#), Ilya Dryomov, Loic Dachary)
- ceph-fuse: fix ceph-fuse writing to stale log file after log rotation ([issue#12350](#), [pr#7110](#), Zhi Zhang)
- ceph init script unconditionally sources /lib/lsb/init-functions ([issue#14402](#), [pr#7797](#), Yan, Zheng)
- ceph.in: Notify user that ‘tell’ can’t be used in interactive mode ([issue#14773](#), [pr#7656](#), David Zafman)
- ceph-objectstore-tool, osd: Fix import handling ([issue#10794](#), [issue#13382](#), [pr#7917](#), Sage Weil, David Zafman)
- client: added permission check based on getgrouplist ([issue#13268](#), [pr#6604](#), Yan, Zheng, Danny Al-Gaaf)
- client: inoderef ([issue#13729](#), [pr#6551](#), Yan, Zheng)
- common: clock skew report is incorrect by ceph health detail command ([issue#14175](#), [pr#8051](#), Joao Eduardo Luis)
- global/pidfile: do not start two daemons with a single pid-file ([issue#13422](#), [pr#7671](#), Loic Dachary, shun song)

- librados: segfault in Objecter::handle_watch_notify ([issue#13805](#), pr#7992, Sage Weil)
- librbd: flattening an rbd image with active IO can lead to hang ([issue#14092](#), [issue#14483](#), pr#7485, Jason Dillaman)
- librbd: possible QEMU deadlock after creating image snapshots ([issue#14988](#), pr#8011, Jason Dillaman)
- mon: Bucket owner isn't changed after unlink/link ([issue#11076](#), pr#8583, Zengran Zhang)
- monclient: avoid key renew storm on clock skew ([issue#12065](#), pr#8398, Alexey Sheplyakov)
- mon: implement reweight-by-utilization feature ([issue#15054](#), pr#8026, Kefu Chai, Dan van der Ster, Sage Weil)
- mon/LogMonitor: use the configured facility if log to syslog ([issue#13748](#), pr#7648, Kefu Chai)
- mon: mon sync does not copy config-key ([issue#14577](#), pr#7576, Xiaowei Chen)
- mon/OSDMonitor: avoid underflow in reweight-by-utilization if max_change=1 ([issue#15655](#), pr#8979, Samuel Just)
- osd: consume_maps clearing of waiting_for_pg needs to check the spg_t shard for acting set membership ([issue#14278](#), pr#7577, Samuel Just)
- osd: log inconsistent shard sizes ([issue#14009](#), pr#6946, Loic Dachary)
- osd: OSD coredumps with leveldb compact on mount = true ([issue#14748](#), pr#7645, Xiaoxi Chen)
- osd/OSDMap: reset osd_primary_affinity shared_ptr when deepish_copy_from ([issue#14686](#), pr#7590, Xinze Chi)
- osd: Protect against excessively large object map sizes ([issue#15121](#), pr#8401, Jason Dillaman)
- osd/ReplicatedPG: do not proxy read and process op locally ([issue#15171](#), pr#8187, Sage Weil)
- osd: scrub bogus results when missing a clone ([issue#14875](#), [issue#14874](#), [issue#14877](#), [issue#10098](#), [issue#14878](#), [issue#14881](#), [issue#14882](#), [issue#14883](#), [issue#14879](#), [issue#10290](#), [issue#12740](#), [issue#12738](#), [issue#14880](#), [issue#11135](#), [issue#14876](#), [issue#10809](#), [issue#12193](#), [issue#11237](#), pr#7702, Xinze Chi, Sage Weil, John Spray, Kefu Chai, Mykola Golub, David Zafman)
- osd: Unable to bring up OSD's after dealing with FULL cluster (OSD assert with /include/interval_set.h: 386: FAILED assert(_size >= 0)) ([issue#14428](#), pr#7415, Alexey Sheplyakov)
- osd: use GMT time for the object name of hitsets ([issue#13192](#), [issue#9732](#), [issue#12968](#), pr#7883, Kefu Chai, David Zafman)
- qa/workunits/post-file.sh: sudo ([issue#14586](#), pr#7456, Sage Weil)
- qa/workunits: remove 'mds setmap' from workunits (pr#8123, Sage Weil)
- rgw: default quota params ([issue#12997](#), pr#7188, Daniel Gryniewicz)
- rgw: make rgw_fronends more forgiving of whitespace ([issue#12038](#), pr#7414, Matt Benjamin)
- rgw: radosgw-admin bucket check -fix not work ([issue#14215](#), pr#7185, Weijun Duan)
- rpm package building fails if the build machine has lttng and babeltrace development packages installed locally ([issue#14844](#), pr#8440, Kefu Chai)
- rpm: redhat-lsb-core dependency was dropped, but is still needed ([issue#14906](#), pr#7876, Nathan Cutler)
- test_bit_vector.cc uses magic numbers against #defines that vary ([issue#14747](#), pr#7672, Jason Dillaman)
- test/librados/tier.cc doesn't completely clean up EC pools ([issue#13878](#), pr#8052, Loic Dachary, Dan Mick)
- tests: bufferlist: do not expect !is_page_aligned() after unaligned rebuild ([issue#15305](#), pr#8272, Kefu Chai)
- tools: fix race condition in seq/rand bench (part 1) ([issue#14968](#), [issue#14873](#), pr#7896, Alexey Sheplyakov, Piotr Dałek)
- tools: fix race condition in seq/rand bench (part 2) ([issue#14873](#), pr#7817, Alexey Sheplyakov)
- tools/rados: add bench smoke tests ([issue#14971](#), pr#7903, Piotr Dałek)
- tools, test: Add ceph-objectstore-tool to operate on the meta collection ([issue#14977](#), pr#7911, David Zafman)
- unittest_crypto: benchmark 100,000 CryptoKey::encrypt() calls ([issue#14863](#), pr#7801, Sage Weil)

V0.94.6 HAMMER

This Hammer point release fixes a range of bugs, most notably a fix for unbounded growth of the monitor's leveldb store, and a workaround in the OSD to keep most xattrs small enough to be stored inline in XFS inodes.

We recommend that all hammer v0.94.x users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build/ops: Ceph daemon failed to start, because the service name was already used. ([issue#13474](#), pr#6832, Chuanhong Wang)
- build/ops: LTTng-UST tracing should be dynamically enabled ([issue#13274](#), pr#6415, Jason Dillaman)
- build/ops: ceph upstart script rbdmap.conf incorrectly processes parameters ([issue#13214](#), pr#6159, Sage Weil)
- build/ops: ceph.spec.in License line does not reflect COPYING ([issue#12935](#), pr#6680, Nathan Cutler)
- build/ops: ceph.spec.in libcephfs_jni1 has no %post and %postun ([issue#12927](#), pr#5789, Owen Synge)
- build/ops: configure.ac: no use to add "+" before ac_ext=c ([issue#14330](#), pr#6973, Kefu Chai, Robin H. Johnson)
- build/ops: deb: strip tracepoint libraries from Wheezy/Precise builds ([issue#14801](#), pr#7316, Jason Dillaman)
- build/ops: init script reload doesn't work on EL7 ([issue#13709](#), pr#7187, Hervé Rousseau)
- build/ops: init-rbdmap uses distro-specific functions ([issue#12415](#), pr#6528, Boris Ranto)
- build/ops: logrotate reload error on Ubuntu 14.04 ([issue#11330](#), pr#5787, Sage Weil)

- build/ops: miscellaneous spec file fixes ([issue#12931](#), [issue#12994](#), [issue#12924](#), [issue#12360](#), [pr#5790](#), Boris Ranto, Nathan Cutler, Owen Synge, Travis Rhoden, Ken Dreyer)
- build/ops: pass tcmalloc env through to ceph-os ([issue#14802](#), [pr#7365](#), Sage Weil)
- build/ops: rbd-replay-* moved from ceph-test-dbg to ceph-common-dbg as well ([issue#13785](#), [pr#6580](#), Loic Dachary)
- build/ops: unknown argument -quiet in udevadm settle ([issue#13560](#), [pr#6530](#), Jason Dillaman)
- common: Objecter: pool op callback may hang forever. ([issue#13642](#), [pr#6588](#), xie xingguo)
- common: Objecter: potential null pointer access when do pool_snap_list. ([issue#13639](#), [pr#6839](#), xie xingguo)
- common: ThreadPool add/remove work queue methods not thread safe ([issue#12662](#), [pr#5889](#), Jason Dillaman)
- common: auth/cephx: large amounts of log are produced by osd ([issue#13610](#), [pr#6835](#), Qiankun Zheng)
- common: client nonce collision due to unshared pid namespaces ([issue#13032](#), [pr#6151](#), Josh Durgin)
- common: common/Thread:pthread_attr_destroy(thread_attr) when done with it ([issue#12570](#), [pr#6157](#), Piotr Dałek)
- common: log: Log.cc: Assign LOG_DEBUG priority to syslog calls ([issue#13993](#), [pr#6994](#), Brad Hubbard)
- common: objecter: cancellation bugs ([issue#13071](#), [pr#6155](#), Jianpeng Ma)
- common: pure virtual method called ([issue#13636](#), [pr#6587](#), Jason Dillaman)
- common: small probability sigabrt when setting rados_osd_op_timeout ([issue#13208](#), [pr#6143](#), Ruifeng Yang)
- common: wrong conditional for boolean function KeyServer::get_auth() ([issue#9756](#), [issue#13424](#), [pr#6213](#), Nathan Cutler)
- crush: crash if we see CRUSH_ITEM_NONE in early rule step ([issue#13477](#), [pr#6430](#), Sage Weil)
- doc: man: document listwatchers cmd in "rados" manpage ([issue#14556](#), [pr#7434](#), Kefu Chai)
- doc: regenerate man pages, add orphans commands to radosgw-admin(8) ([issue#14637](#), [pr#7524](#), Ken Dreyer)
- fs: CephFS restriction on removing cache tiers is overly strict ([issue#11504](#), [pr#6402](#), John Spray)
- fs: fsstress.sh fails ([issue#12710](#), [pr#7454](#), Yan, Zheng)
- librados: LibRadosWatchNotify.WatchNotify2Timeout ([issue#13114](#), [pr#6336](#), Sage Weil)
- librbd: ImageWatcher shouldn't block the notification thread ([issue#14373](#), [pr#7407](#), Jason Dillaman)
- librbd: diff_iterate needs to handle holes in parent images ([issue#12885](#), [pr#6097](#), Jason Dillaman)
- librbd: fix merge-diff for >2GB diff-files ([issue#14030](#), [pr#6980](#), Jason Dillaman)
- librbd: invalidate object map on error even w/o holding lock ([issue#13372](#), [pr#6289](#), Jason Dillaman)
- librbd: reads larger than cache size hang ([issue#13164](#), [pr#6354](#), Lu Shi)
- mds: ceph mds add_data_pool check for EC pool is wrong ([issue#12426](#), [pr#5766](#), John Spray)
- mon: MonitorDBStore: get_next_key() only if prefix matches ([issue#11786](#), [pr#5361](#), Joao Eduardo Luis)
- mon: OSDMonitor: do not assume a session exists in send_incremental() ([issue#14236](#), [pr#7150](#), Joao Eduardo Luis)
- mon: check for store writeability before participating in election ([issue#13089](#), [pr#6144](#), Sage Weil)
- mon: compact full epochs also ([issue#14537](#), [pr#7446](#), Kefu Chai)
- mon: include min_last_epoch_clean as part of PGMap::print_summary and PGMap::dump ([issue#13198](#), [pr#6152](#), Guang Yang)
- mon: map_cache can become inaccurate if osd does not receive the osdmmaps ([issue#10930](#), [pr#5773](#), Kefu Chai)
- mon: should not set isvalid = true when cephx_verify_authorizer return false ([issue#13525](#), [pr#6391](#), Ruifeng Yang)
- osd: Ceph Pools' MAX AVAIL is 0 if some OSDs' weight is 0 ([issue#13840](#), [pr#6834](#), Chengyuan Li)
- osd: FileStore calls syncfs(2) even it is not supported ([issue#12512](#), [pr#5530](#), Kefu Chai)
- osd: FileStore: potential memory leak if getattr fails. ([issue#13597](#), [pr#6420](#), xie xingguo)
- osd: IO error on kvm/rbd with an erasure coded pool tier ([issue#12012](#), [pr#5897](#), Kefu Chai)
- osd: OSD::build_past_intervals_parallel() shall reset primary and up_primary when begin a new past_interval. ([issue#13471](#), [pr#6398](#), xie xingguo)
- osd: ReplicatedBackend: populate recovery_info.size for clone (bug symptom is size mismatch on replicated backend on a clone in scrub) ([issue#12828](#), [pr#6153](#), Samuel Just)
- osd: ReplicatedPG: wrong result code checking logic during sparse_read ([issue#14151](#), [pr#7179](#), xie xingguo)
- osd: ReplicatedPG::hit_set_trim osd/ReplicatedPG.cc: 11006: FAILED assert(obc) ([issue#13192](#), [issue#9732](#), [issue#12968](#), [pr#5825](#), Kefu Chai, Zhiqiang Wang, Samuel Just, David Zafman)
- osd: avoid multi set osd_op.outdata in tier pool ([issue#12540](#), [pr#6060](#), Xinze Chi)
- osd: bug with cache/tiering and snapshot reads ([issue#12748](#), [pr#6589](#), Kefu Chai)
- osd: ceph osd pool stats broken in hammer ([issue#13843](#), [pr#7180](#), BJ Lougee)
- osd: ceph-disk prepare fails if device is a symlink ([issue#13438](#), [pr#7176](#), Joe Julian)
- osd: check for full before changing the cached obc (hammer) ([issue#13098](#), [pr#6918](#), Alexey Sheplyakov)
- osd: config_opts: increase suicide timeout to 300 to match recovery ([issue#14376](#), [pr#7236](#), Samuel Just)
- osd: disable filestore_xfs_extsize by default ([issue#14397](#), [pr#7411](#), Ken Dreyer)
- osd: do not cache unused memory in attrs ([issue#12565](#), [pr#6499](#), Xinze Chi, Ning Yao)
- osd: dumpling incrementals do not work properly on hammer and newer ([issue#13234](#), [pr#6132](#), Samuel Just)
- osd: filestore: fix peek_queue for OpSequencer ([issue#13209](#), [pr#6145](#), Xinze Chi)
- osd: hit set clear repops fired in same epoch as map change - segfault since they fall into the new interval even though the repops are cleared ([issue#12809](#), [pr#5890](#), Samuel Just)
- osd: object_info_t::decode() has wrong version ([issue#13462](#), [pr#6335](#), David Zafman)
- osd: osd/OSD.cc: 2469: FAILED assert(pg_stat_queue.empty()) on shutdown ([issue#14212](#), [pr#7178](#), Sage Weil)
- osd: osd/PG.cc: 288: FAILED assert(info.last_epoch_started >= info.history.last_epoch_started) ([issue#14015](#), [pr#7177](#), David Zafman)
- osd: osd/PG.cc: 3837: FAILED assert(0 == "Running incompatible OSD") ([issue#11661](#), [pr#7206](#), David Zafman)

- osd: osd/ReplicatedPG: Recency fix ([issue#14320](#), pr#7207, Sage Weil, Robert LeBlanc)
- osd: pg stuck in replay ([issue#13116](#), pr#6401, Sage Weil)
- osd: race condition detected during send_failures ([issue#13821](#), pr#6755, Sage Weil)
- osd: randomize scrub times ([issue#10973](#), pr#6199, Kefu Chai)
- osd: requeue_scrub when kick_object_context_blocked ([issue#12515](#), pr#5891, Xinze Chi)
- osd: revert: use GMT time for hitsets ([issue#13812](#), pr#6644, Loic Dachary)
- osd: segfault in agent_work ([issue#13199](#), pr#6146, Samuel Just)
- osd: should recalc the min_last_epoch_clean when decode PGMap ([issue#13112](#), pr#6154, Kefu Chai)
- osd: smaller object_info_t xattrs ([issue#14803](#), pr#6544, Sage Weil)
- osd: we do not ignore notify from down osds ([issue#12990](#), pr#6158, Samuel Just)
- rbd: QEMU hangs after creating snapshot and stopping VM ([issue#13726](#), pr#6586, Jason Dillaman)
- rbd: TaskFinisher::cancel should remove event from SafeTimer ([issue#14476](#), pr#7417, Douglas Fuller)
- rbd: avoid re-writing old-format image header on resize ([issue#13674](#), pr#6585, Jason Dillaman)
- rbd: fix bench-write ([issue#14225](#), pr#7183, Sage Weil)
- rbd: rbd-replay does not check for EOF and goes to endless loop ([issue#14452](#), pr#7416, Mykola Golub)
- rbd: rbd-replay-prep and rbd-replay improvements ([issue#13221](#), [issue#13220](#), [issue#13378](#), pr#6286, Jason Dillaman)
- rbd: verify self-managed snapshot functionality on image create ([issue#13633](#), pr#7182, Jason Dillaman)
- rgw: Make RGW_MAX_PUT_SIZE configurable ([issue#6999](#), pr#7441, Vladislav Odintsov, Yuan Zhou)
- rgw: Setting ACL on Object removes ETag ([issue#12955](#), pr#6620, Brian Felton)
- rgw: backport content-type casing ([issue#12939](#), pr#5910, Robin H. Johnson)
- rgw: bucket listing hangs on versioned buckets ([issue#12913](#), pr#6352, Yehuda Sadeh)
- rgw: fix wrong etag calculation during POST on S3 bucket. ([issue#11241](#), pr#7442, Vladislav Odintsov, Radoslaw Zarzynski)
- rgw: get bucket location returns region name, not region api name ([issue#13458](#), pr#6349, Yehuda Sadeh)
- rgw: missing handling of encoding-type=url when listing keys in bucket ([issue#12735](#), pr#6527, Jeff Weber)
- rgw: orphan tool should be careful about removing head objects ([issue#12958](#), pr#6351, Yehuda Sadeh)
- rgw: orphans finish segfaults ([issue#13824](#), pr#7186, Igor Fedotov)
- rgw: rgw-admin: document orphans commands in usage ([issue#14516](#), pr#7526, Yehuda Sadeh)
- rgw: swift API returns more than real object count and bytes used when retrieving account metadata ([issue#13140](#), pr#6512, Sangdi Xu)
- rgw: swift use Civetweb ssl can not get right url ([issue#13628](#), pr#6491, Weijun Duan)
- rgw: value of Swift API's X-Object-Manifest header is not url_decoded during segment look up ([issue#12728](#), pr#6353, Radoslaw Zarzynski)
- tests: fixed broken Makefiles after integration of ttng into rados ([issue#13210](#), pr#6322, Sebastien Ponce)
- tests: fsx failed to compile ([issue#14384](#), pr#7501, Greg Farnum)
- tests: notification slave needs to wait for master ([issue#13810](#), pr#7226, Jason Dillaman)
- tests: qa: remove legacy OS support from rbd/qemu-iotests ([issue#13483](#), [issue#14385](#), pr#7252, Vasu Kulkarni, Jason Dillaman)
- tests: testprofile must be removed before it is re-created ([issue#13664](#), pr#6450, Loic Dachary)
- tools: ceph-monstore-tool must do out_store.close() ([issue#10093](#), pr#7347, huangjun)
- tools: heavy memory shuffling in rados bench ([issue#12946](#), pr#5810, Piotr Dałek)
- tools: race condition in rados bench ([issue#12947](#), pr#6791, Piotr Dałek)
- tools: tool for artificially inflate the leveldb of the mon store for testing purposes ([issue#10093](#), [issue#11815](#), [issue#14217](#), pr#7412, Cilang Zhao, Bo Cai, Kefu Chai, huangjun, Joao Eduardo Luis)

V0.94.5 HAMMER

This Hammer point release fixes a critical regression in librbd that can cause QEMU/KVM to crash when caching is enabled on images that have been cloned.

All v0.94.4 Hammer users are strongly encouraged to upgrade.

NOTABLE CHANGES

- librbd: potential assertion failure during cache read ([issue#13559](#), pr#6348, Jason Dillaman)
- osd: osd/ReplicatedPG: remove stray debug line ([issue#13455](#), pr#6362, Sage Weil)
- tests: qemu workunit refers to apt-mirror.front.sepia.ceph.com ([issue#13420](#), pr#6330, Yuan Zhou)

For more detailed information, see [the complete changelog](#).

V0.94.4 HAMMER

This Hammer point release fixes several important bugs in Hammer, as well as fixing interoperability issues that are required before an upgrade to Infernalis. That is, all users of earlier version of Hammer or any version of Firefly will first need to upgrade to hammer v0.94.4 or later before upgrading to Infernalis (or future releases).

All v0.94.x Hammer users are strongly encouraged to upgrade.

NOTABLE CHANGES

- build/ops: ceph.spec.in: 50-rbd.rules conditional is wrong ([issue#12166](#), [pr#5207](#), Nathan Cutler)
- build/ops: ceph.spec.in: ceph-common needs python-argparse on older distros, but doesn't require it ([issue#12034](#), [pr#5216](#), Nathan Cutler)
- build/ops: ceph.spec.in: radosgw requires apache for SUSE only - makes no sense ([issue#12358](#), [pr#5411](#), Nathan Cutler)
- build/ops: ceph.spec.in: rpm: cephfs_java not fully conditionalized ([issue#11991](#), [pr#5202](#), Nathan Cutler)
- build/ops: ceph.spec.in: rpm: not possible to turn off Java ([issue#11992](#), [pr#5203](#), Owen Synge)
- build/ops: ceph.spec.in: running fdupes unnecessarily ([issue#12301](#), [pr#5223](#), Nathan Cutler)
- build/ops: ceph.spec.in: snappy-devel for all supported distros ([issue#12361](#), [pr#5264](#), Nathan Cutler)
- build/ops: ceph.spec.in: SUSE/openSUSE builds need libbz2-devel ([issue#11629](#), [pr#5204](#), Nathan Cutler)
- build/ops: ceph.spec.in: useless %py_requires breaks SLE11-SP3 build ([issue#12351](#), [pr#5412](#), Nathan Cutler)
- build/ops: error in ext_mime_map_init() when /etc/mime.types is missing ([issue#11864](#), [pr#5385](#), Ken Dreyer)
- build/ops: upstart: limit respawn to 3 in 30 mins (instead of 5 in 30s) ([issue#11798](#), [pr#5930](#), Sage Weil)
- build/ops: With root as default user, unable to have multiple RGW instances running ([issue#10927](#), [pr#6161](#), Sage Weil)
- build/ops: With root as default user, unable to have multiple RGW instances running ([issue#11140](#), [pr#6161](#), Sage Weil)
- build/ops: With root as default user, unable to have multiple RGW instances running ([issue#11686](#), [pr#6161](#), Sage Weil)
- build/ops: With root as default user, unable to have multiple RGW instances running ([issue#12407](#), [pr#6161](#), Sage Weil)
- cli: ceph: cli throws exception on unrecognized errno ([issue#11354](#), [pr#5368](#), Kefu Chai)
- cli: ceph tell: broken error message / misleading hinting ([issue#11101](#), [pr#5371](#), Kefu Chai)
- common: arm: all programs that link to librados2 hang forever on startup ([issue#12505](#), [pr#5366](#), Boris Ranto)
- common: buffer: critical bufferlist::zero bug ([issue#12252](#), [pr#5365](#), Haomai Wang)
- common: ceph-object-corpus: add 0.94.2-207-g88e7ee7 hammer objects ([issue#13070](#), [pr#5551](#), Sage Weil)
- common: do not insert empty ptr when rebuild empty bufferlist ([issue#12775](#), [pr#5764](#), Xinze Chi)
- common: [FAILED] TestLibRBD.BlockingAIO ([issue#12479](#), [pr#5768](#), Jason Dillaman)
- common: LibCephFS.GetPoold failure ([issue#12598](#), [pr#5887](#), Yan, Zheng)
- common: Memory leak in Mutex.cc, pthread_mutexattr_init without pthread_mutexattr_destroy ([issue#11762](#), [pr#5378](#), Ketor Meng)
- common: object_map_update fails with -EINVAL return code ([issue#12611](#), [pr#5559](#), Jason Dillaman)
- common: Pipe: Drop connect_seq increase line ([issue#13093](#), [pr#5908](#), Haomai Wang)
- common: recursive lock of md_config_t (0) ([issue#12614](#), [pr#5759](#), Josh Durgin)
- crush: ceph osd crush reweight-subtree does not reweight parent node ([issue#11855](#), [pr#5374](#), Sage Weil)
- doc: update docs to point to download.ceph.com ([issue#13162](#), [pr#6156](#), Alfredo Deza)
- fs: ceph-fuse 0.94.2-1trusty segfaults / aborts ([issue#12297](#), [pr#5381](#), Greg Farnum)
- fs: segfault launching ceph-fuse with bad -name ([issue#12417](#), [pr#5382](#), John Spray)
- librados: Change radosgw pools default crush ruleset ([issue#11640](#), [pr#5754](#), Yuan Zhou)
- librbd: correct issues discovered via lockdep / helgrind ([issue#12345](#), [pr#5296](#), Jason Dillaman)
- librbd: Crash during TestInternal.MultipleResize ([issue#12664](#), [pr#5769](#), Jason Dillaman)
- librbd: deadlock during cooperative exclusive lock transition ([issue#11537](#), [pr#5319](#), Jason Dillaman)
- librbd: Possible crash while concurrently writing and shrinking an image ([issue#11743](#), [pr#5318](#), Jason Dillaman)
- mon: add a cache layer over MonitorDBStore ([issue#12638](#), [pr#5697](#), Kefu Chai)
- mon: fix crush testing for new pools ([issue#13400](#), [pr#6192](#), Sage Weil)
- mon: get pools health'info have error ([issue#12402](#), [pr#5369](#), renhwztetecs)
- mon: implicit erasure code crush ruleset is not validated ([issue#11814](#), [pr#5276](#), Loic Dachary)
- mon: PaxosService: call post_refresh() instead of post_paxos_update() ([issue#11470](#), [pr#5359](#), Joao Eduardo Luis)
- mon: pgmonitor: wrong at/near target max" reporting ([issue#12401](#), [pr#5370](#), huangjun)
- mon: register_new_pgs() should check ruleno instead of its index ([issue#12210](#), [pr#5377](#), Xinze Chi)
- mon: Show osd as NONE in ceph osd map <pool> <object> output ([issue#11820](#), [pr#5376](#), Shylesh Kumar)
- mon: the output is wrong when running ceph osd reweight ([issue#12251](#), [pr#5372](#), Joao Eduardo Luis)
- osd: allow peek_map_epoch to return an error ([issue#13060](#), [pr#5892](#), Sage Weil)
- osd: cache agent is idle although one object is left in the cache ([issue#12673](#), [pr#5765](#), Loic Dachary)
- osd: copy-from doesn't preserve truncate_{seq,size} ([issue#12551](#), [pr#5885](#), Samuel Just)
- osd: crash creating/deleting pools ([issue#12429](#), [pr#5527](#), John Spray)
- osd: fix repair when recorded digest is wrong ([issue#12577](#), [pr#5468](#), Sage Weil)
- osd: include/ceph_features: define HAMMER_0_94_4 feature ([issue#13026](#), [pr#5687](#), Sage Weil)
- osd: is_new_interval() fixes ([issue#10399](#), [pr#5691](#), Jason Dillaman)
- osd: is_new_interval() fixes ([issue#11771](#), [pr#5691](#), Jason Dillaman)

- osd: long standing slow requests: connection->session->waiting_for_map->connection ref cycle ([issue#12338](#), [pr#5761](#), Samuel Just)
- osd: Mutex Assert from PipeConnection::try_get_pipe ([issue#12437](#), [pr#5758](#), David Zafman)
- osd: pg_interval_t::check_new_interval - for ec pool, should not rely on min_size to determine if the PG was active at the interval ([issue#12162](#), [pr#5373](#), Guang G Yang)
- osd: PGLog.cc: 732: FAILED assert(log.log.size() == log_keys_debug.size()) ([issue#12652](#), [pr#5763](#), Sage Weil)
- osd: PGLog::proc_replica_log: correctly handle case where entries between olog.head and log.tail were split out ([issue#11358](#), [pr#5380](#), Samuel Just)
- osd: read on chunk-aligned xattr not handled ([issue#12309](#), [pr#5367](#), Sage Weil)
- osd: suicide timeout during peering - search for missing objects ([issue#12523](#), [pr#5762](#), Guang G Yang)
- osd: WBThrottle::clear_object: signal on cond when we reduce throttle values ([issue#12223](#), [pr#5757](#), Samuel Just)
- rbd: crash during shutdown after writeback blocked by IO errors ([issue#12597](#), [pr#5767](#), Jianpeng Ma)
- rgw: add delimiter to prefix only when path is specified ([issue#12960](#), [pr#5860](#), Sylvain Baubéau)
- rgw: create a tool for orphaned objects cleanup ([issue#9604](#), [pr#5717](#), Yehuda Sadeh)
- rgw: don't preserve acls when copying object ([issue#11563](#), [pr#6039](#), Yehuda Sadeh)
- rgw: don't preserve acls when copying object ([issue#12370](#), [pr#6039](#), Yehuda Sadeh)
- rgw: don't preserve acls when copying object ([issue#13015](#), [pr#6039](#), Yehuda Sadeh)
- rgw: Ensure that swift keys don't include backslashes ([issue#7647](#), [pr#5716](#), Yehuda Sadeh)
- rgw: GWWatcher::handle_error -> common/Mutex.cc: 95: FAILED assert(r == 0) ([issue#12208](#), [pr#6164](#), Yehuda Sadeh)
- rgw: HTTP return code is not being logged by CivetWeb ([issue#12432](#), [pr#5498](#), Yehuda Sadeh)
- rgw: init_rados failed leads to repeated delete ([issue#12978](#), [pr#6165](#), Xiaowei Chen)
- rgw: init some manifest fields when handling explicit objs ([issue#11455](#), [pr#5732](#), Yehuda Sadeh)
- rgw: Keystone Fernet tokens break auth ([issue#12761](#), [pr#6162](#), Abhishek Lekshmanan)
- rgw: region data still exist in region-map after region-map update ([issue#12964](#), [pr#6163](#), dwj192)
- rgw: remove trailing :port from host for purposes of subdomain matching ([issue#12353](#), [pr#6042](#), Yehuda Sadeh)
- rgw: rest-bench common/WorkQueue.cc: 54: FAILED assert(_threads.empty()) ([issue#3896](#), [pr#5383](#), huangjun)
- rgw: returns requested bucket name raw in Bucket response header ([issue#12537](#), [pr#5715](#), Yehuda Sadeh)
- rgw: segmentation fault when rgw_gc_max_objs > HASH_PRIME ([issue#12630](#), [pr#5719](#), Ruifeng Yang)
- rgw: segments are read during HEAD on Swift DLO ([issue#12780](#), [pr#6160](#), Yehuda Sadeh)
- rgw: setting max number of buckets for user via ceph.conf option ([issue#12714](#), [pr#6166](#), Vikhyat Umrao)
- rgw: Swift API: X-Trans-Id header is wrongly formatted ([issue#12108](#), [pr#5721](#), Radoslaw Zarzynski)
- rgw: testGetContentType and testHead failed ([issue#11091](#), [pr#5718](#), Radoslaw Zarzynski)
- rgw: testGetContentType and testHead failed ([issue#11438](#), [pr#5718](#), Radoslaw Zarzynski)
- rgw: testGetContentType and testHead failed ([issue#12157](#), [pr#5718](#), Radoslaw Zarzynski)
- rgw: testGetContentType and testHead failed ([issue#12158](#), [pr#5718](#), Radoslaw Zarzynski)
- rgw: testGetContentType and testHead failed ([issue#12363](#), [pr#5718](#), Radoslaw Zarzynski)
- rgw: the arguments 'domain' should not be assigned when return false ([issue#12629](#), [pr#5720](#), Ruifeng Yang)
- tests: qa/workunits/cephtool/test.sh: don't assume crash_replay_interval=45 ([issue#13406](#), [pr#6172](#), Sage Weil)
- tests: TEST_crush_rule_create_erasure consistently fails on i386 builder ([issue#12419](#), [pr#6201](#), Loic Dachary)
- tools: ceph-disk zap should ensure block device ([issue#11272](#), [pr#5755](#), Loic Dachary)

For more detailed information, see [the complete changelog](#).

V0.94.3 HAMMER

This Hammer point release fixes a critical (though rare) data corruption bug that could be triggered when logs are rotated via SIGHUP. It also fixes a range of other important bugs in the OSD, monitor, RGW, RGW, and CephFS.

All v0.94.x Hammer users are strongly encouraged to upgrade.

UPGRADING

- The pg ls-by-{pool,primary,osd} commands and pg ls now take the argument recovering instead of recovery in order to include the recovering pgs in the listed pgs.

NOTABLE CHANGES

- librbd: aio calls may block ([issue#11770](#), [pr#4875](#), Jason Dillaman)
- osd: make the all osd/filestore thread pool suicide timeouts separately configurable ([issue#11701](#), [pr#5159](#), Samuel Just)
- mon: ceph fails to compile with boost 1.58 ([issue#11982](#), [pr#5122](#), Kefu Chai)
- tests: TEST_crush_reject_empty must not run a mon ([issue#12285](#), [issue#11975](#), [pr#5208](#), Kefu Chai)
- osd: FAILED assert(!old_value.deleted()) in upgrade:giant-x-hammer-distro-basic-multi run ([issue#11983](#), [pr#5121](#), Samuel Just)

- build/ops: linking ceph to tcmalloc causes segfault on SUSE SLE11-SP3 ([issue#12368](#), [pr#5265](#), Thorsten Behrens)
- common: utf8 and old gcc breakage on RHEL6.5 ([issue#7387](#), [pr#4687](#), Kefu Chai)
- crush: take crashes due to invalid arg ([issue#11740](#), [pr#4891](#), Sage Weil)
- rgw: need conversion tool to handle fixes following #11974 ([issue#12502](#), [pr#5384](#), Yehuda Sadeh)
- rgw: Swift API: support for 202 Accepted response code on container creation ([issue#12299](#), [pr#5214](#), Radoslaw Zarzynski)
- common: Log::reopen_log_file: take m_flush_mutex ([issue#12520](#), [pr#5405](#), Samuel Just)
- rgw: Properly respond to the Connection header with Civetweb ([issue#12398](#), [pr#5284](#), Wido den Hollander)
- rgw: multipart list part response returns incorrect field ([issue#12399](#), [pr#5285](#), Henry Chang)
- build/ops: ceph.spec.in: 95-ceph-osd.rules, mount.ceph, and mount.fuse.ceph not installed properly on SUSE ([issue#12397](#), [pr#5283](#), Nathan Cutler)
- rgw: radosgw-admin dumps user info twice ([issue#12400](#), [pr#5286](#), guce)
- doc: fix doc build ([issue#12180](#), [pr#5095](#), Kefu Chai)
- tests: backport 11493 fixes, and test, preventing ec cache pools ([issue#12314](#), [pr#4961](#), Samuel Just)
- rgw: does not send Date HTTP header when civetweb frontend is used ([issue#11872](#), [pr#5228](#), Radoslaw Zarzynski)
- mon: pg ls is broken ([issue#11910](#), [pr#5160](#), Kefu Chai)
- librbd: A client opening an image mid-resize can result in the object map being invalidated ([issue#12237](#), [pr#5279](#), Jason Dillaman)
- doc: missing man pages for ceph-create-keys, ceph-disk-* ([issue#11862](#), [pr#4846](#), Nathan Cutler)
- tools: ceph-post-file fails on rhel7 ([issue#11876](#), [pr#5038](#), Sage Weil)
- build/ops: rcceph script is buggy ([issue#12090](#), [pr#5028](#), Owen Synge)
- rgw: Bucket header is enclosed by quotes ([issue#11874](#), [pr#4862](#), Wido den Hollander)
- build/ops: packaging: add SuSEfirewall2 service files ([issue#12092](#), [pr#5030](#), Tim Serong)
- rgw: Keystone PKI token expiration is not enforced ([issue#11722](#), [pr#4884](#), Anton Aksola)
- build/ops: debian/control: ceph-common (>> 0.94.2) must be >= 0.94.2-2 ([issue#12529](#), [issue#11998](#), [pr#5417](#), Loic Dachary)
- mon: Clock skew causes missing summary and confuses Calamari ([issue#11879](#), [pr#4868](#), Thorsten Behrens)
- rgw: rados objects wrongly deleted ([issue#12099](#), [pr#5117](#), wuxingyi)
- tests: kernel_untar_build fails on EL7 ([issue#12098](#), [pr#5119](#), Greg Farnum)
- fs: Fh ref count will leak if readahead does not need to do read from osd ([issue#12319](#), [pr#5427](#), Zhi Zhang)
- mon: OSDMonitor: allow addition of cache pool with non-empty snaps with co... ([issue#12595](#), [pr#5252](#), Samuel Just)
- mon: MDSMonitor: handle MDSBeacon messages properly ([issue#11979](#), [pr#5123](#), Kefu Chai)
- tools: ceph-disk: get_partition_type fails on /dev/cciss... ([issue#11760](#), [pr#4892](#), islepnev)
- build/ops: max files open limit for OSD daemon is too low ([issue#12087](#), [pr#5026](#), Owen Synge)
- mon: add an “osd crush tree” command ([issue#11833](#), [pr#5248](#), Kefu Chai)
- mon: mon crashes when “ceph osd tree 85 -format json” ([issue#11975](#), [pr#4936](#), Kefu Chai)
- build/ops: ceph / ceph-dbg steal ceph-objectstore-tool from ceph-test / ceph-test-dbg ([issue#11806](#), [pr#5069](#), Loic Dachary)
- rgw: DragonDisk fails to create directories via S3: MissingContentLength ([issue#12042](#), [pr#5118](#), Yehuda Sadeh)
- build/ops: /usr/bin/ceph from ceph-common is broken without installing ceph ([issue#11998](#), [pr#5206](#), Ken Dreyer)
- build/ops: systemd: Increase max files open limit for OSD daemon ([issue#11964](#), [pr#5040](#), Owen Synge)
- build/ops: rgw/logrotate.conf calls service with wrong init script name ([issue#12044](#), [pr#5055](#), wuxingyi)
- common: OPT_INT option interprets 3221225472 as -1073741824, and crashes in Throttle::Throttle() ([issue#11738](#), [pr#4889](#), Kefu Chai)
- doc: doc/release-notes: v0.94.2 ([issue#11492](#), [pr#4934](#), Sage Weil)
- common: admin_socket: close socket descriptor in destructor ([issue#11706](#), [pr#4657](#), Jon Bernard)
- rgw: Object copy bug ([issue#11755](#), [pr#4885](#), Javier M. Mellid)
- rgw: empty json response when getting user quota ([issue#12245](#), [pr#5237](#), wuxingyi)
- fs: cephfs Dumper tries to load whole journal into memory at once ([issue#11999](#), [pr#5120](#), John Spray)
- rgw: Fix tool for #11442 does not correctly fix objects created via multipart uploads ([issue#12242](#), [pr#5229](#), Yehuda Sadeh)
- rgw: Civetweb RGW appears to report full size of object as downloaded when only partially downloaded ([issue#12243](#), [pr#5231](#), Yehuda Sadeh)
- osd: stuck incomplete ([issue#12362](#), [pr#5269](#), Samuel Just)
- osd: start_flush: filter out removed snaps before determining snapc’s ([issue#11911](#), [pr#4899](#), Samuel Just)
- librbd: internal.cc: 1967: FAILED assert(watchers.size() == 1) ([issue#12239](#), [pr#5243](#), Jason Dillaman)
- librbd: new QA client upgrade tests ([issue#12109](#), [pr#5046](#), Jason Dillaman)
- librbd: [FAILED] TestLibRBD.ExclusiveLockTransition ([issue#12238](#), [pr#5241](#), Jason Dillaman)
- rgw: Swift API: XML document generated in response for GET on account does not contain account name ([issue#12323](#), [pr#5227](#), Radoslaw Zarzynski)
- rgw: keystone does not support chunked input ([issue#12322](#), [pr#5226](#), Hervé Rousseau)
- mds: MDS is crashed (mds/CDir.cc: 1391: FAILED assert(!is_complete())) ([issue#11737](#), [pr#4886](#), Yan, Zheng)
- cli: ceph: cli interactive mode does not understand quotes ([issue#11736](#), [pr#4776](#), Kefu Chai)
- librbd: add valgrind memory checks for unit tests ([issue#12384](#), [pr#5280](#), Zhiqiang Wang)
- build/ops: admin/build-doc: script fails silently under certain circumstances ([issue#11902](#), [pr#4877](#), John Spray)
- osd: Fixes for rados ops with snaps ([issue#11908](#), [pr#4902](#), Samuel Just)

- build/ops: ceph.spec.in: ceph-common subpackage def needs tweaking for SUSE/openSUSE ([issue#12308](#), [pr#4883](#), Nathan Cutler)
- fs: client: reference counting ‘struct Fh’ ([issue#12088](#), [pr#5222](#), Yan, Zheng)
- build/ops: ceph.spec: update OpenSUSE BuildRequires ([issue#11611](#), [pr#4667](#), Loic Dachary)

For more detailed information, see [the complete changelog](#).

V0.94.2 HAMMER

This Hammer point release fixes a few critical bugs in RGW that can prevent objects starting with underscore from behaving properly and that prevent garbage collection of deleted objects when using the Civetweb standalone mode.

All v0.94.x Hammer users are strongly encouraged to upgrade, and to make note of the repair procedure below if RGW is in use.

UPGRADING FROM PREVIOUS HAMMER RELEASE

Bug #11442 introduced a change that made rgw objects that start with underscore incompatible with previous versions. The fix to that bug reverts to the previous behavior. In order to be able to access objects that start with an underscore and were created in prior Hammer releases, following the upgrade it is required to run (for each affected bucket):

```
$ radosgw-admin bucket check --check-head-obj-locator \
    --bucket=<bucket> [--fix]
```

NOTABLE CHANGES

- build: compilation error: No high-precision counter available (armhf, powerpc..) (#11432, James Page)
- ceph-dencoder links to libtcmalloc, and shouldn't (#10691, Boris Ranto)
- ceph-disk: disk zap sgdisk invocation (#11143, Owen Syngue)
- ceph-disk: use a new disk as journal disk,ceph-disk prepare fail (#10983, Loic Dachary)
- ceph-objectstore-tool should be in the ceph server package (#11376, Ken Dreyer)
- librados: can get stuck in redirect loop if osdmap epoch == last_force_op_resend (#11026, Jianpeng Ma)
- librbd: A retransmit of proxied flatten request can result in -EINVAL (Jason Dillaman)
- librbd: ImageWatcher should cancel in-flight ops on watch error (#11363, Jason Dillaman)
- librbd: Objectcacher setting max object counts too low (#7385, Jason Dillaman)
- librbd: Periodic failure of TestLibRBD.DiffIterateStress (#11369, Jason Dillaman)
- librbd: Queued AIO reference counters not properly updated (#11478, Jason Dillaman)
- librbd: deadlock in image refresh (#5488, Jason Dillaman)
- librbd: notification race condition on snap_create (#11342, Jason Dillaman)
- mds: Hammer uclient checking (#11510, John Spray)
- mds: remove caps from revoking list when caps are voluntarily released (#11482, Yan, Zheng)
- messenger: double clear of pipe in reaper (#11381, Haomai Wang)
- mon: Total size of OSDs is a magnitude less than it is supposed to be. (#11534, Zhe Zhang)
- osd: don't check order in finish_proxy_read (#11211, Zhiqiang Wang)
- osd: handle old semi-deleted pgs after upgrade (#11429, Samuel Just)
- osd: object creation by write cannot use an offset on an erasure coded pool (#11507, Jianpeng Ma)
- rgw: Improve rgw HEAD request by avoiding read the body of the first chunk (#11001, Guang Yang)
- rgw: civetweb is hitting a limit (number of threads 1024) (#10243, Yehuda Sadeh)
- rgw: civetweb should use unique request id (#10295, Orit Wasserman)
- rgw: critical fixes for hammer (#11447, #11442, Yehuda Sadeh)
- rgw: fix swift COPY headers (#10662, #10663, #11087, #10645, Radoslaw Zarzynski)
- rgw: improve performance for large object (multiple chunks) GET (#11322, Guang Yang)
- rgw: init-radosgw: run RGW as root (#11453, Ken Dreyer)
- rgw: keystone token cache does not work correctly (#11125, Yehuda Sadeh)
- rgw: make quota/gc thread configurable for starting (#11047, Guang Yang)
- rgw: make swift responses of RGW return last-modified, content-length, x-trans-id headers. (#10650, Radoslaw Zarzynski)
- rgw: merge manifests correctly when there's prefix override (#11622, Yehuda Sadeh)
- rgw: quota not respected in POST object (#11323, Sergey Arkhipov)
- rgw: restore buffer of multipart upload after EEXIST (#11604, Yehuda Sadeh)
- rgw: shouldn't need to disable rgw_socket_path if frontend is configured (#11160, Yehuda Sadeh)
- rgw: swift: Response header of GET request for container does not contain X-Container-Object-Count, X-Container-Bytes-Used and x-trans-id headers (#10666, Dmytro Iurchenko)

- rgw: swift: Response header of POST request for object does not contain content-length and x-trans-id headers (#10661, Radoslaw Zarzynski)
- rgw: swift: response for GET/HEAD on container does not contain the X-Timestamp header (#10938, Radoslaw Zarzynski)
- rgw: swift: response for PUT on /container does not contain the mandatory Content-Length header when FCGI is used (#11036, #10971, Radoslaw Zarzynski)
- rgw: swift: wrong handling of empty metadata on Swift container (#11088, Radoslaw Zarzynski)
- tests: TestFlatIndex.cc races with TestLFNIndex.cc (#11217, Xinze Chi)
- tests: ceph-helpers kill_daemons fails when kill fails (#11398, Loic Dachary)

For more detailed information, see [the complete changelog](#).

V0.94.1 HAMMER

This bug fix release fixes a few critical issues with CRUSH. The most important addresses a bug in feature bit enforcement that may prevent pre-hammer clients from communicating with the cluster during an upgrade. This only manifests in some cases (for example, when the ‘rack’ type is in use in the CRUSH map, and possibly other cases), but for safety we strongly recommend that all users use 0.94.1 instead of 0.94 when upgrading.

There is also a fix in the new straw2 buckets when OSD weights are 0.

We recommend that all v0.94 users upgrade.

NOTABLE CHANGES

- crush: fix divide-by-0 in straw2 (#11357 Sage Weil)
- crush: fix has_v4_buckets (#11364 Sage Weil)
- osd: fix negative degraded objects during backfilling (#7737 Guang Yang)

For more detailed information, see [the complete changelog](#).

V0.94 HAMMER

This major release is expected to form the basis of the next long-term stable series. It is intended to supersede v0.80.x Firefly.

Highlights since Giant include:

- *RADOS Performance*: a range of improvements have been made in the OSD and client-side librados code that improve the throughput on flash backends and improve parallelism and scaling on fast machines.
- *Simplified RGW deployment*: the ceph-deploy tool now has a new ‘ceph-deploy rgw create HOST’ command that quickly deploys a instance of the S3/Swift gateway using the embedded Civetweb server. This is vastly simpler than the previous Apache-based deployment. There are a few rough edges (e.g., around SSL support) but we encourage users to try [the new method](#).
- *RGW object versioning*: RGW now supports the S3 object versioning API, which preserves old version of objects instead of overwriting them.
- *RGW bucket sharding*: RGW can now shard the bucket index for large buckets across, improving performance for very large buckets.
- *RBD object maps*: RBD now has an object map function that tracks which parts of the image are allocating, improving performance for clones and for commands like export and delete.
- *RBD mandatory locking*: RBD has a new mandatory locking framework (still disabled by default) that adds additional safeguards to prevent multiple clients from using the same image at the same time.
- *RBD copy-on-read*: RBD now supports copy-on-read for image clones, improving performance for some workloads.
- *CephFS snapshot improvements*: Many many bugs have been fixed with CephFS snapshots. Although they are still disabled by default, stability has improved significantly.
- *CephFS Recovery tools*: We have built some journal recovery and diagnostic tools. Stability and performance of single-MDS systems is vastly improved in Giant, and more improvements have been made now in Hammer. Although we still recommend caution when storing important data in CephFS, we do encourage testing for non-critical workloads so that we can better guage the feature, usability, performance, and stability gaps.
- *CRUSH improvements*: We have added a new straw2 bucket algorithm that reduces the amount of data migration required when changes are made to the cluster.
- *Shingled erasure codes (SHEC)*: The OSDs now have experimental support for shingled erasure codes, which allow a small amount of additional storage to be traded for improved recovery performance.
- *RADOS cache tiering*: A series of changes have been made in the cache tiering code that improve performance and reduce latency.
- *RDMA support*: There is now experimental support the RDMA via the Accelio (libxio) library.

- *New administrator commands:* The ‘ceph osd df’ command shows pertinent details on OSD disk utilizations. The ‘ceph pg ls ...’ command makes it much simpler to query PG states while diagnosing cluster issues.

Other highlights since Firefly include:

- *CephFS:* we have fixed a raft of bugs in CephFS and built some basic journal recovery and diagnostic tools. Stability and performance of single-MDS systems is vastly improved in Giant. Although we do not yet recommend CephFS for production deployments, we do encourage testing for non-critical workloads so that we can better gauge the feature, usability, performance, and stability gaps.
- *Local Recovery Codes:* the OSDs now support an erasure-coding scheme that stores some additional data blocks to reduce the IO required to recover from single OSD failures.
- *Degraded vs misplaced:* the Ceph health reports from ‘ceph -s’ and related commands now make a distinction between data that is degraded (there are fewer than the desired number of copies) and data that is misplaced (stored in the wrong location in the cluster). The distinction is important because the latter does not compromise data safety.
- *Tiering improvements:* we have made several improvements to the cache tiering implementation that improve performance. Most notably, objects are not promoted into the cache tier by a single read; they must be found to be sufficiently hot before that happens.
- *Monitor performance:* the monitors now perform writes to the local data store asynchronously, improving overall responsiveness.
- *Recovery tools:* the ceph-objectstore-tool is greatly expanded to allow manipulation of an individual OSDs data store for debugging and repair purposes. This is most heavily used by our QA infrastructure to exercise recovery code.

I would like to take this opportunity to call out the amazing growth in contributors to Ceph beyond the core development team from Inktank. Hammer features major new features and improvements from Intel, Fujitsu, UnitedStack, Yahoo, UbuntuKylin, CohortFS, Mellanox, CERN, Deutsche Telekom, Mirantis, and SanDisk.

DEDICATION

This release is dedicated in memoriam to Sandon Van Ness, aka Houkouonchi, who unexpectedly passed away a few weeks ago. Sandon was responsible for maintaining the large and complex Sepia lab that houses the Ceph project’s build and test infrastructure. His efforts have made an important impact on our ability to reliably test Ceph with a relatively small group of people. He was a valued member of the team and we will miss him. H is also for Houkouonchi.

UPGRADING

- If your existing cluster is running a version older than v0.80.x Firefly, please first upgrade to the latest Firefly release before moving on to Giant. We have not tested upgrades directly from Emperor, Dumpling, or older releases.

We have tested:

- Firefly to Hammer
- Giant to Hammer
- Dumpling to Firefly to Hammer
- Please upgrade daemons in the following order:
 1. Monitors
 2. OSDs
 3. MDSs and/or radosgw

Note that the relative ordering of OSDs and monitors should not matter, but we primarily tested upgrading monitors first.

- The ceph-osd daemons will perform a disk-format upgrade improve the PG metadata layout and to repair a minor bug in the on-disk format. It may take a minute or two for this to complete, depending on how many objects are stored on the node; do not be alarmed if they do not mark “up” by the cluster immediately after starting.
- If upgrading from v0.93, set

```
osd enable degraded writes = false
```

on all osds prior to upgrading. The degraded writes feature has been reverted due to 11155.
- The LTTNG tracing in librbd and librados is disabled in the release packages until we find a way to avoid violating distro security policies when linking libust.

UPGRADING FROM V0.87.X GIANT

- librbd and librados include lttng tracepoints on distros with liblttng 2.4 or later (only Ubuntu Trusty for the ceph.com packages). When running a daemon that uses these libraries, i.e. an application that calls fork(2) or clone(2) without exec(3), you must set LD_PRELOAD=liblttng-ust-fork.so.0 to prevent a crash in the lttng atexit handler when the process exits. The only ceph tool that requires this is rbd-fuse.
- If rgw_socket_path is defined and rgw_frontends defines a socket_port and socket_host, we now allow the rgw_frontends settings to take precedence. This change should only affect users who have made non-standard changes to their radosgw configuration.
- If you are upgrading specifically from v0.92, you must stop all OSD daemons and flush their journals (ceph-osd -i NNN --flush-journal) before upgrading. There was a transaction encoding bug in v0.92 that broke compatibility. Upgrading from v0.93, v0.91, or anything earlier is safe.
- The experimental ‘keyvaluestore-dev’ OSD backend has been renamed ‘keyvaluestore’ (for simplicity) and marked as experimental. To enable this untested feature and acknowledge that you understand that it is untested and may destroy data, you need to add the following to your ceph.conf:

```
enable experimental unrecoverable data corrupting features = keyvaluestore
```

- The following librados C API function calls take a ‘flags’ argument whose value is now correctly interpreted:

```
rados_write_op_operate() rados_aio_write_op_operate() rados_read_op_operate() rados_aio_read_op_operate()
```

The flags were not correctly being translated from the librados constants to the internal values. Now they are. Any code that is passing flags to these methods should be audited to ensure that they are using the correct LIBRADOS_OP_FLAG_* constants.

- The ‘rados’ CLI ‘copy’ and ‘cppool’ commands now use the copy-from operation, which means the latest CLI cannot run these commands against pre-firefly OSDs.
- The librados watch/notify API now includes a watch_flush() operation to flush the async queue of notify operations. This should be called by any watch/notify user prior to rados_shutdown().
- The ‘category’ field for objects has been removed. This was originally added to track PG stat summations over different categories of objects for use by radosgw. It is no longer has any known users and is prone to abuse because it can lead to a pg_stat_t structure that is unbounded. The librados API calls that accept this field now ignore it, and the OSD no longer tracks the per-category summations.
- The output for ‘rados df’ has changed. The ‘category’ level has been eliminated, so there is now a single stat object per pool. The structure of the JSON output is different, and the plaintext output has one less column.
- The ‘rados create <objectname> [category]’ optional category argument is no longer supported or recognized.
- rados.py’s Rados class no longer has a __del__ method; it was causing problems on interpreter shutdown and use of threads. If your code has Rados objects with limited lifetimes and you’re concerned about locked resources, call Rados.shutdown() explicitly.
- There is a new version of the librados watch/notify API with vastly improved semantics. Any applications using this interface are encouraged to migrate to the new API. The old API calls are marked as deprecated and will eventually be removed.
- The librados rados_unwatch() call used to be safe to call on an invalid handle. The new version has undefined behavior when passed a bogus value (for example, when rados_watch() returns an error and handle is not defined).
- The structure of the formatted ‘pg stat’ command is changed for the portion that counts states by name to avoid using the ‘+’ character (which appears in state names) as part of the XML token (it is not legal).
- Previously, the formatted output of ‘ceph pg stat -f ...’ was a full pg dump that included all metadata about all PGs in the system. It is now a concise summary of high-level PG stats, just like the unformatted ‘ceph pg stat’ command.
- All JSON dumps of floating point values were incorrectly surrounding the value with quotes. These quotes have been removed. Any consumer of structured JSON output that was consuming the floating point values was previously having to interpret the quoted string and will most likely need to be fixed to take the unquoted number.
- New ability to list all objects from all namespaces that can fail or return incomplete results when not all OSDs have been upgraded. Features rados -all ls, rados cppool, rados export, rados cache-flush-evict-all and rados cache-try-flush-evict-all can also fail or return incomplete results.
- Due to a change in the Linux kernel version 3.18 and the limits of the FUSE interface, ceph-fuse needs to be mounted as root on at least some systems. See issues #9997, #10277, and #10542 for details.

- The client-side caching for librbd is now enabled by default (rbd cache = true). A safety option (rbd cache writethrough until flush = true) is also enabled so that writeback caching is not used until the library observes a ‘flush’ command, indicating that the librbd users is passing that operation through from the guest VM. This avoids potential data loss when used with older versions of qemu that do not support flush.

```
leveldb_write_buffer_size = 8*1024*1024 = 33554432 // 8MB leveldb_cache_size = 512*1024*1204 =
536870912 // 512MB leveldb_block_size = 64*1024 = 65536 // 64KB leveldb_compression = false leveldb_log
= ""
```

OSDs will still maintain the following osd-specific defaults:

```
leveldb_log = ""
```

- The ‘rados getxattr …’ command used to add a gratuitous newline to the attr value; it now does not.
- The *_kb perf counters on the monitor have been removed. These are replaced with a new set of *_bytes counters (e.g., cluster_osd_kb is replaced by cluster_osd_bytes).
- The rd_kb and wr_kb fields in the JSON dumps for pool stats (accessed via the ceph df detail -f json-pretty and related commands) have been replaced with corresponding *_bytes fields. Similarly, the total_space, total_used, and total_avail fields are replaced with total_bytes, total_used_bytes, and total_avail_bytes fields.
- The rados df --format=json output read_bytes and write_bytes fields were incorrectly reporting ops; this is now fixed.
- The rados df --format=json output previously included read_kb and write_kb fields; these have been removed. Please use read_bytes and write_bytes instead (and divide by 1024 if appropriate).
- The experimental keyvaluestore-dev OSD backend had an on-disk format change that prevents existing OSD data from being upgraded. This affects developers and testers only.
- mon-specific and osd-specific leveldb options have been removed. From this point onward users should use the leveldb_* generic options and add the options in the appropriate sections of their configuration files. Monitors will still maintain the following monitor-specific defaults:

```
leveldb_write_buffer_size = 8*1024*1024 = 33554432 // 8MB leveldb_cache_size = 512*1024*1204 =
536870912 // 512MB leveldb_block_size = 64*1024 = 65536 // 64KB leveldb_compression = false leveldb_log
= ""
```

OSDs will still maintain the following osd-specific defaults:

```
leveldb_log = ""
```

- CephFS support for the legacy anchor table has finally been removed. Users with file systems created before firefly should ensure that inodes with multiple hard links are modified *prior* to the upgrade to ensure that the backtraces are written properly. For example:

```
sudo find /mnt/cephfs -type f -links +1 -exec touch \{\}\ \;
```

- We disallow nonsensical ‘tier cache-mode’ transitions. From this point onward, ‘writeback’ can only transition to ‘forward’ and ‘forward’ can transition to 1) ‘writeback’ if there are dirty objects, or 2) any if there are no dirty objects.

NOTABLE CHANGES SINCE V0.93

- build: a few cmake fixes (Matt Benjamin)
- build: fix build on RHEL/CentOS 5.9 (Rohan Mars)
- build: reorganize Makefile to allow modular builds (Boris Ranto)
- ceph-fuse: be more forgiving on remount (#10982 Greg Farnum)
- ceph: improve CLI parsing (#11093 David Zafman)
- common: fix cluster logging to default channel (#11177 Sage Weil)
- crush: fix parsing of straw2 buckets (#11015 Sage Weil)
- doc: update man pages (David Zafman)
- librados: fix leak in C_TwoContexts (Xiong Yiliang)
- librados: fix leak in watch/notify path (Sage Weil)
- librbd: fix and improve AIO cache invalidation (#10958 Jason Dillaman)

- librbd: fix memory leak (Jason Dillaman)
- librbd: fix ordering/queueing of resize operations (Jason Dillaman)
- librbd: validate image is r/w on resize/flatten (Jason Dillaman)
- librbd: various internal locking fixes (Jason Dillaman)
- lttng: tracing is disabled until we streamline dependencies (Josh Durgin)
- mon: add bootstrap-rgw profile (Sage Weil)
- mon: do not pollute mon dir with CSV files from CRUSH check (Loic Dachary)
- mon: fix clock drift time check interval (#10546 Joao Eduardo Luis)
- mon: fix units in store stats (Joao Eduardo Luis)
- mon: improve error handling on erasure code profile set (#10488, #11144 Loic Dachary)
- mon: set {read,write}_tier on 'osd tier add-cache ...' (Jianpeng Ma)
- ms: xio: fix misc bugs (Matt Benjamin, Vu Pham)
- osd: DBObjectMap: fix locking to prevent rare crash (#9891 Samuel Just)
- osd: fix and document last_epoch_started semantics (Samuel Just)
- osd: fix divergent entry handling on PG split (Samuel Just)
- osd: fix leak on shutdown (Kefu Chai)
- osd: fix recording of digest on scrub (Samuel Just)
- osd: fix whiteout handling (Sage Weil)
- rbd: allow v2 striping parameters for clones and imports (Jason Dillaman)
- rbd: fix formatted output of image features (Jason Dillaman)
- rbd: update eman page (Ilya Dryomov)
- rgw: don't overwrite bucket/object owner when setting ACLs (#10978 Yehuda Sadeh)
- rgw: enable IPv6 for civetweb (#10965 Yehuda Sadeh)
- rgw: fix sysvinit script when rgw_socket_path is not defined (#11159 Yehuda Sadeh, Dan Mick)
- rgw: pass civetweb configurables through (#10907 Yehuda Sadeh)
- rgw: use new watch/notify API (Yehuda Sadeh, Sage Weil)
- osd: reverted degraded writes feature due to 11155

NOTABLE CHANGES SINCE V0.87.X GIANT

- add experimental features option (Sage Weil)
- arch: fix NEON feaeture detection (#10185 Loic Dachary)
- asyncmsg: misc fixes (Haomai Wang)
- buffer: add 'shareable' construct (Matt Benjamin)
- buffer: add list::get contiguous (Sage Weil)
- buffer: avoid rebuild if buffer already contiguous (Jianpeng Ma)
- build: CMake support (Ali Maredia, Casey Bodley, Adam Emerson, Marcus Watts, Matt Benjamin)
- build: a few cmake fixes (Matt Benjamin)
- build: aarch64 build fixes (Noah Watkins, Haomai Wang)
- build: adjust build deps for yasm, virtualenv (Jianpeng Ma)
- build: fix 'make check' races (#10384 Loic Dachary)
- build: fix build on RHEL/CentOS 5.9 (Rohan Mars)
- build: fix pkg names when libkeyutils is missing (Pankag Garg, Ken Dreyer)
- build: improve build dependency tooling (Loic Dachary)
- build: reorganize Makefile to allow modular builds (Boris Ranto)
- build: support for jemalloc (Shishir Gowda)
- ceph-disk: Scientific Linux support (Dan van der Ster)
- ceph-disk: allow journal partition re-use (#10146 Loic Dachary, Dav van der Ster)
- ceph-disk: call partx/partprobe consistency (#9721 Loic Dachary)
- ceph-disk: do not re-use partition if encryption is required (Loic Dachary)
- ceph-disk: fix dmcrypt key permissions (Loic Dachary)
- ceph-disk: fix umount race condition (#10096 Blaine Gardner)
- ceph-disk: improved systemd support (Owen Synge)
- ceph-disk: init=none option (Loic Dachary)
- ceph-disk: misc fixes (Christos Stavrakakis)
- ceph-disk: respect -statedir for keyring (Loic Dachary)
- ceph-disk: set guid if reusing journal partition (Dan van der Ster)
- ceph-disk: support LUKS for encrypted partitions (Andrew Bartlett, Loic Dachary)
- ceph-fuse, libcephfs: POSIX file lock support (Yan, Zheng)
- ceph-fuse, libcephfs: allow xattr caps in inject_release_failure (#9800 John Spray)
- ceph-fuse, libcephfs: fix I_COMPLETE_ORDERED checks (#9894 Yan, Zheng)
- ceph-fuse, libcephfs: fix cap flush overflow (Greg Farnum, Yan, Zheng)
- ceph-fuse, libcephfs: fix root inode xattrs (Yan, Zheng)
- ceph-fuse, libcephfs: preserve dir ordering (#9178 Yan, Zheng)

- ceph-fuse, libcephfs: trim inodes before reconnecting to MDS (Yan, Zheng)
- ceph-fuse,libcephfs: add support for O_NOFOLLOW and O_PATH (Greg Farnum)
- ceph-fuse,libcephfs: resend requests before completing cap reconnect (#10912 Yan, Zheng)
- ceph-fuse: be more forgiving on remount (#10982 Greg Farnum)
- ceph-fuse: fix dentry invalidation on 3.18+ kernels (#9997 Yan, Zheng)
- ceph-fuse: fix kernel cache trimming (#10277 Yan, Zheng)
- ceph-fuse: select kernel cache invalidation mechanism based on kernel version (Greg Farnum)
- ceph-monstore-tool: fix shutdown (#10093 Loic Dachary)
- ceph-monstore-tool: fix/improve CLI (Joao Eduardo Luis)
- ceph-objectstore-tool: fix import (#10090 David Zafman)
- ceph-objectstore-tool: improved import (David Zafman)
- ceph-objectstore-tool: many improvements and tests (David Zafman)
- ceph-objectstore-tool: many many improvements (David Zafman)
- ceph-objectstore-tool: misc improvements, fixes (#9870 #9871 David Zafman)
- ceph.spec: package rbd-replay-prep (Ken Dreyer)
- ceph: add 'ceph osd df [tree]' command (#10452 Mykola Golub)
- ceph: do not parse injectargs twice (Loic Dachary)
- ceph: fix 'ceph tell ...' command validation (#10439 Joao Eduardo Luis)
- ceph: improve 'ceph osd tree' output (Mykola Golub)
- ceph: improve CLI parsing (#11093 David Zafman)
- ceph: make 'ceph -s' output more readable (Sage Weil)
- ceph: make 'ceph -s' show PG state counts in sorted order (Sage Weil)
- ceph: make 'ceph tell mon.* version' work (Mykola Golub)
- ceph: new 'ceph tell mds.\$name_or_rank_or_gid' (John Spray)
- ceph: show primary-affinity in 'ceph osd tree' (Mykola Golub)
- ceph: test robustness (Joao Eduardo Luis)
- ceph_objectstore_tool: behave with sharded flag (#9661 David Zafman)
- cephfs-journal-tool: add recover_dentries function (#9883 John Spray)
- cephfs-journal-tool: fix journal import (#10025 John Spray)
- cephfs-journal-tool: skip up to expire_pos (#9977 John Spray)
- cleanup rados.h definitions with macros (Ilya Dryomov)
- common: add 'perf reset ...' admin command (Jianpeng Ma)
- common: add TableFormatter (Andreas Peters)
- common: add newline to flushed json output (Sage Weil)
- common: check syncfs() return code (Jianpeng Ma)
- common: do not unlock rwlock on destruction (Federico Simoncelli)
- common: filtering for 'perf dump' (John Spray)
- common: fix Formatter factory breakage (#10547 Loic Dachary)
- common: fix block device discard check (#10296 Sage Weil)
- common: make json-pretty output prettier (Sage Weil)
- common: remove broken CEPH_LOCKDEP optoin (Kefu Chai)
- common: shared_cache unit tests (Cheng Cheng)
- common: support new gperftools header locations (Key Dreyer)
- config: add \$cctid meta variable (Adam Crume)
- crush: fix buffer overrun for poorly formed rules (#9492 Johnu George)
- crush: fix detach_bucket (#10095 Sage Weil)
- crush: fix parsing of straw2 buckets (#11015 Sage Weil)
- crush: fix several bugs in adjust_item_weight (Rongze Zhu)
- crush: fix tree bucket behavior (Rongze Zhu)
- crush: improve constness (Loic Dachary)
- crush: new and improved straw2 bucket type (Sage Weil, Christina Anderson, Xiaoxi Chen)
- crush: straw bucket weight calculation fixes (#9998 Sage Weil)
- crush: update tries stats for indep rules (#10349 Loic Dachary)
- crush: use larger choose_tries value for erasure code rulesets (#10353 Loic Dachary)
- crushtool: add -location <id> command (Sage Weil, Loic Dachary)
- debian,rpm: move RBD udev rules to ceph-common (#10864 Ken Dreyer)
- debian: split python-ceph into python-{rbd,rados,cephfs} (Boris Ranto)
- default to libnss instead of crypto++ (Federico Gimenez)
- doc: CephFS disaster recovery guidance (John Spray)
- doc: CephFS for early adopters (John Spray)
- doc: add build-doc guidlines for Fedora and CentOS/RHEL (Nilamdyuti Goswami)
- doc: add dumpling to firefly upgrade section (#7679 John Wilkins)
- doc: ceph osd reweight vs crush weight (Laurent Guerby)
- doc: do not suggest dangerous XFS nobarrier option (Dan van der Ster)
- doc: document erasure coded pool operations (#9970 Loic Dachary)

- doc: document the LRC per-layer plugin configuration (Yuan Zhou)
- doc: enable rbd cache on openstack deployments (Sebastien Han)
- doc: erasure code doc updates (Loic Dachary)
- doc: file system osd config settings (Kevin Dalley)
- doc: fix OpenStack Glance docs (#10478 Sebastien Han)
- doc: improved installation notes on CentOS/RHEL installs (John Wilkins)
- doc: key/value store config reference (John Wilkins)
- doc: misc cleanups (Adam Spiers, Sebastien Han, Nilamdyuti Goswami, Ken Dreyer, John Wilkins)
- doc: misc improvements (Nilamdyuti Goswami, John Wilkins, Chris Holcombe)
- doc: misc updates (#9793 #9922 #10204 #10203 Travis Rhoden, Hazem, Ayari, Florian Coste, Andy Allan, Frank Yu, Baptiste Veuillez-Mainard, Yuan Zhou, Armando Segnini, Robert Jansen, Tyler Brekke, Viktor Suprun)
- doc: misc updates (Alfredo Deza, VRan Liu)
- doc: misc updates (Nilamdyuti Goswami, John Wilkins)
- doc: new man pages (Nilamdyuti Goswami)
- doc: preflight doc fixes (John Wilkins)
- doc: replace cloudfiles with swiftclient Python Swift example (Tim Freund)
- doc: update PG count guide (Gerbent Meijer, Laurent Guerby, Loic Dachary)
- doc: update man pages (David Zafman)
- doc: update openstack docs for Juno (Sebastien Han)
- doc: update release descriptions (Ken Dreyer)
- doc: update sepio hardware inventory (Sandon Van Ness)
- erasure-code: add mSHEC erasure code support (Takeshi Miyamae)
- erasure-code: improved docs (#10340 Loic Dachary)
- erasure-code: set max_size to 20 (#10363 Loic Dachary)
- fix cluster logging from non-mon daemons (Sage Weil)
- init-ceph: check for systemd-run before using it (Boris Ranto)
- install-deps.sh: do not require sudo when root (Loic Dachary)
- keyvaluestore: misc fixes (Haomai Wang)
- keyvaluestore: performance improvements (Haomai Wang)
- libcephfs,ceph-fuse: add 'status' asok (John Spray)
- libcephfs,ceph-fuse: fix getting zero-length xattr (#10552 Yan, Zheng)
- libcephfs: fix dirfrag trimming (#10387 Yan, Zheng)
- libcephfs: fix mount timeout (#10041 Yan, Zheng)
- libcephfs: fix test (#10415 Yan, Zheng)
- libcephfs: fix use-afer-free on umount (#10412 Yan, Zheng)
- libcephfs: include ceph and git version in client metadata (Sage Weil)
- librados, osd: new watch/notify implementation (Sage Weil)
- librados: add blacklist_add convenience method (Jason Dillaman)
- librados: add rados_pool_get_base_tier() call (Adam Crume)
- librados: add watch_flush() operation (Sage Weil, Haomai Wang)
- librados: avoid memcpy on getxattr, read (Jianpeng Ma)
- librados: cap buffer length (Loic Dachary)
- librados: create ioctx by pool id (Jason Dillaman)
- librados: do notify completion in fast-dispatch (Sage Weil)
- librados: drop 'category' feature (Sage Weil)
- librados: expose rados_{read|write}_op_assert_version in C API (Kim Vandry)
- librados: fix infinite loop with skipped map epochs (#9986 Ding Dinghua)
- librados: fix iterator operator= bugs (#10082 David Zafman, Yehuda Sadeh)
- librados: fix leak in C_TwoContexts (Xiong Yiliang)
- librados: fix leak in watch/notify path (Sage Weil)
- librados: fix null deref when pool DNE (#9944 Sage Weil)
- librados: fix objecter races (#9617 Josh Durgin)
- librados: fix pool deletion handling (#10372 Sage Weil)
- librados: fix pool name caching (#10458 Radoslaw Zarzynski)
- librados: fix resource leak, misc bugs (#10425 Radoslaw Zarzynski)
- librados: fix some watch/notify locking (Jason Dillaman, Josh Durgin)
- librados: fix timer race from recent refactor (Sage Weil)
- librados: new fadvise API (Ma Jianpeng)
- librados: only export public API symbols (Jason Dillaman)
- librados: remove shadowed variable (Kefu Chain)
- librados: translate op flags from C APIs (Matthew Richards)
- libradosstriper: fix remove() (Dongmao Zhang)
- libradosstriper: fix shutdown hang (Dongmao Zhang)
- libradosstriper: fix stat strtoll (Dongmao Zhang)
- libradosstriper: fix trunc method (#10129 Sebastien Ponce)

- libradosstriper: fix write_full when ENOENT (#10758 Sebastien Ponce)
- libradosstriper: misc fixes (Sebastien Ponce)
- librbd: CRC protection for RBD image map (Jason Dillaman)
- librbd: add missing python docstrings (Jason Dillaman)
- librbd: add per-image object map for improved performance (Jason Dillaman)
- librbd: add readahead (Adam Crume)
- librbd: add support for an “object map” indicating which objects exist (Jason Dillaman)
- librbd: adjust internal locking (Josh Durgin, Jason Dillaman)
- librbd: better handling of watch errors (Jason Dillaman)
- librbd: complete pending ops before closing image (#10299 Josh Durgin)
- librbd: coordinate maint operations through lock owner (Jason Dillaman)
- librbd: copy-on-read (Min Chen, Li Wang, Yunchuan Wen, Cheng Cheng, Jason Dillaman)
- librbd: differentiate between R/O vs R/W features (Jason Dillaman)
- librbd: don’t close a closed parent in failure path (#10030 Jason Dillaman)
- librbd: enforce write ordering with a snapshot (Jason Dillaman)
- librbd: exclusive image locking (Jason Dillaman)
- librbd: fadvise API (Ma Jianpeng)
- librbd: fadvise-style hints; add misc hints for certain operations (Jianpeng Ma)
- librbd: fix and improve AIO cache invalidation (#10958 Jason Dillaman)
- librbd: fix cache tiers in list_children and snap_unprotect (Adam Crume)
- librbd: fix coverity false-positives (Jason Dillaman)
- librbd: fix diff test (#10002 Josh Durgin)
- librbd: fix list_children from invalid pool iocxs (#10123 Jason Dillaman)
- librbd: fix locking for readahead (#10045 Jason Dillaman)
- librbd: fix memory leak (Jason Dillaman)
- librbd: fix ordering/queueing of resize operations (Jason Dillaman)
- librbd: fix performance regression in ObjectCacher (#9513 Adam Crume)
- librbd: fix snap create races (Jason Dillaman)
- librbd: fix write vs import race (#10590 Jason Dillaman)
- librbd: flush AIO operations asynchronously (#10714 Jason Dillaman)
- librbd: gracefully handle deleted/renamed pools (#10270 Jason Dillaman)
- librbd: lttng tracepoints (Adam Crume)
- librbd: make async versions of long-running maint operations (Jason Dillaman)
- librbd: misc fixes (Xinxin Shu, Jason Dillaman)
- librbd: mock tests (Jason Dillaman)
- librbd: only export public API symbols (Jason Dillaman)
- librbd: optionally blacklist clients before breaking locks (#10761 Jason Dillaman)
- librbd: prevent copyup during shrink (Jason Dillaman)
- librbd: refactor unit tests to use fixtures (Jason Dillaman)
- librbd: validate image is r/w on resize/flatten (Jason Dillaman)
- librbd: various internal locking fixes (Jason Dillaman)
- many coverity fixes (Danny Al-Gaaf)
- many many coverity cleanups (Danny Al-Gaaf)
- mds: ‘flush journal’ admin command (John Spray)
- mds: ENOSPC and OSDMap epoch barriers (#7317 John Spray)
- mds: a whole bunch of initial scrub infrastructure (Greg Farnum)
- mds: add cephfs-table-tool (John Spray)
- mds: asok command for fetching subtree map (John Spray)
- mds: avoid sending traceless replies in most cases (Yan, Zheng)
- mds: constify MDSCacheObjects (John Spray)
- mds: dirfrag buf fix (Yan, Zheng)
- mds: disallow most commands on inactive MDS’s (Greg Farnum)
- mds: drop dentries, leases on deleted directories (#10164 Yan, Zheng)
- mds: export dir asok command (John Spray)
- mds: fix MDLog IO callback deadlock (John Spray)
- mds: fix compat_version for MClientSession (#9945 John Spray)
- mds: fix deadlock during journal probe vs purge (#10229 Yan, Zheng)
- mds: fix race trimming log segments (Yan, Zheng)
- mds: fix reply snapbl (Yan, Zheng)
- mds: fix sessionmap lifecycle bugs (Yan, Zheng)
- mds: fix stray/purge perfcounters (#10388 John Spray)
- mds: handle heartbeat_reset during shutdown (#10382 John Spray)
- mds: handle zero-size xattr (#10335 Yan, Zheng)
- mds: initialize root inode xattr version (Yan, Zheng)
- mds: introduce auth caps (John Spray)

- mds: many many snapshot-related fixes (Yan, Zheng)
- mds: misc bugs (Greg Farnum, John Spray, Yan, Zheng, Henry Change)
- mds: refactor, improve Session storage (John Spray)
- mds: store backtrace for stray dir (Yan, Zheng)
- mds: subtree quota support (Yunchuan Wen)
- mds: verify backtrace when fetching dirfrag (#9557 Yan, Zheng)
- memstore: free space tracking (John Spray)
- misc cleanup (Danny Al-Gaaf, David Anderson)
- misc coverity fixes (Danny Al-Gaaf)
- misc coverity fixes (Danny Al-Gaaf)
- misc: various valgrind fixes and cleanups (Danny Al-Gaaf)
- mon: ‘osd crush reweight-all’ command (Sage Weil)
- mon: add ‘ceph osd rename-bucket ...’ command (Loic Dachary)
- mon: add bootstrap-rgw profile (Sage Weil)
- mon: add max pgs per osd warning (Sage Weil)
- mon: add noforward flag for some mon commands (Mykola Golub)
- mon: allow adding tiers to fs pools (#10135 John Spray)
- mon: allow full flag to be manually cleared (#9323 Sage Weil)
- mon: clean up auth list output (Loic Dachary)
- mon: delay failure injection (Joao Eduardo Luis)
- mon: disallow empty pool names (#10555 Wido den Hollander)
- mon: do not deactivate last mds (#10862 John Spray)
- mon: do not pollute mon dir with CSV files from CRUSH check (Loic Dachary)
- mon: drop old ceph_mon_store_converter (Sage Weil)
- mon: fix ‘ceph pg dump_stuck degraded’ (Xinxin Shu)
- mon: fix ‘mds fail’ for standby MDSs (John Spray)
- mon: fix ‘osd crush link’ id resolution (John Spray)
- mon: fix ‘profile osd’ use of config-key function on mon (#10844 Joao Eduardo Luis)
- mon: fix _ratio units and types (Sage Weil)
- mon: fix JSON dumps to dump floats as flots and not strings (Sage Weil)
- mon: fix MDS health status from peons (#10151 John Spray)
- mon: fix caching for min_last_epoch_clean (#9987 Sage Weil)
- mon: fix clock drift time check interval (#10546 Joao Eduardo Luis)
- mon: fix compatset initialization during mkfs (Joao Eduardo Luis)
- mon: fix error output for add_data_pool (#9852 Joao Eduardo Luis)
- mon: fix feature tracking during elections (Joao Eduardo Luis)
- mon: fix formatter ‘pg stat’ command output (Sage Weil)
- mon: fix mds gid/rank/state parsing (John Spray)
- mon: fix misc error paths (Joao Eduardo Luis)
- mon: fix paxos off-by-one corner case (#9301 Sage Weil)
- mon: fix paxos timeouts (#10220 Joao Eduardo Luis)
- mon: fix stashed monmap encoding (#5203 Xie Rui)
- mon: fix units in store stats (Joao Eduardo Luis)
- mon: get canonical OSDMap from leader (#10422 Sage Weil)
- mon: ignore failure reports from before up_from (#10762 Dan van der Ster, Sage Weil)
- mon: implement ‘fs reset’ command (John Spray)
- mon: improve error handling on erasure code profile set (#10488, #11144 Loic Dachary)
- mon: improved corrupt CRUSH map detection (Joao Eduardo Luis)
- mon: include entity name in audit log for forwarded requests (#9913 Joao Eduardo Luis)
- mon: include pg_temp count in osdmap summary (Sage Weil)
- mon: log health summary to cluster log (#9440 Joao Eduardo Luis)
- mon: make ‘mds fail’ idempotent (John Spray)
- mon: make pg dump {sum,pgs,pgs_brief} work for format=plain (#5963 #6759 Mykola Golub)
- mon: new ‘ceph pool ls [detail]’ command (Sage Weil)
- mon: new pool safety flags nodelete, nopgchange, nosizechange (#9792 Mykola Golub)
- mon: new, friendly ‘ceph pg ls ...’ command (Xinxin Shu)
- mon: paxos: allow reads while proposing (#9321 #9322 Joao Eduardo Luis)
- mon: prevent MDS transition from STOPPING (#10791 Greg Farnum)
- mon: propose all pending work in one transaction (Sage Weil)
- mon: remove pg_temps for nonexistent pools (Joao Eduardo Luis)
- mon: require mon_allow_pool_delete option to remove pools (Sage Weil)
- mon: respect down flag when promoting standbys (John Spray)
- mon: set globalid prealloc to larger value (Sage Weil)
- mon: set {read,write}_tier on ‘osd tier add-cache ...’ (Jianpeng Ma)
- mon: skip zeroed osd stats in get_rule_avail (#10257 Joao Eduardo Luis)

- mon: validate min_size range (Jianpeng Ma)
- mon: wait for writeable before cross-proposing (#9794 Joao Eduardo Luis)
- mount.ceph: fix suprious error message (#10351 Yan, Zheng)
- ms: xio: fix misc bugs (Matt Benjamin, Vu Pham)
- msgr: async: bind threads to CPU cores, improved poll (Haomai Wang)
- msgr: async: many fixes, unit tests (Haomai Wang)
- msgr: async: several fixes (Haomai Wang)
- msgr: asyncmessenger: add kqueue support (#9926 Haomai Wang)
- msgr: avoid useless new/delete (Haomai Wang)
- msgr: fix RESETSESSION bug (#10080 Greg Farnum)
- msgr: fix crc configuration (Mykola Golub)
- msgr: fix delay injection bug (#9910 Sage Weil, Greg Farnum)
- msgr: misc unit tests (Haomai Wang)
- msgr: new AsymcMessenger alternative implementation (Haomai Wang)
- msgr: prefetch data when doing recv (Yehuda Sadeh)
- msgr: simple: fix rare deadlock (Greg Farnum)
- msgr: simple: retry binding to port on failure (#10029 Wido den Hollander)
- msgr: xio: XioMessenger RDMA support (Casey Bodley, Vu Pham, Matt Benjamin)
- objectstore: deprecate collection attrs (Sage Weil)
- osd, librados: fadvise-style librados hints (Jianpeng Ma)
- osd, librados: fix xattr_cmp_u64 (Dongmao Zhang)
- osd, librados: revamp PG listing API to handle namespaces (#9031 #9262 #9438 David Zafman)
- osd, mds: 'ops' as shorthand for 'dump_ops_in_flight' on asok (Sage Weil)
- osd, mon: add checksums to all OSDMaps (Sage Weil)
- osd, mon: send intial pg create time from mon to osd (#9887 David Zafman)
- osd,mon: add 'norebalance' flag (Kefu Chai)
- osd,mon: specify OSD features explicitly in MOSDBoot (#10911 Sage Weil)
- osd: DBObjectMap: fix locking to prevent rare crash (#9891 Samuel Just)
- osd: EIO on whole-object reads when checksum is wrong (Sage Weil)
- osd: add erasure code corpus (Loic Dachary)
- osd: add fadvise flags to ObjectStore API (Jianpeng Ma)
- osd: add get_latest_osdmap asok command (#9483 #9484 Mykola Golub)
- osd: add misc tests (Loic Dachary, Danny Al-Gaaf)
- osd: add option to prioritize heartbeat network traffic (Jian Wen)
- osd: add support for the SHEC erasure-code algorithm (Takeshi Miyamae, Loic Dachary)
- osd: allow deletion of objects with watcher (#2339 Sage Weil)
- osd: allow recovery while below min_size (Samuel Just)
- osd: allow recovery with fewer than min_size OSDs (Samuel Just)
- osd: allow sparse read for Push/Pull (Haomai Wang)
- osd: allow whiteout deletion in cache pool (Sage Weil)
- osd: allow writes to degraded objects (Samuel Just)
- osd: allow writes to degraded objects (Samuel Just)
- osd: avoid publishing unchanged PG stats (Sage Weil)
- osd: batch pg log trim (Xinze Chi)
- osd: cache pool: ignore min flush age when cache is full (Xinze Chi)
- osd: cache recent ObjectContexts (Dong Yuan)
- osd: cache reverse_nibbles hash value (Dong Yuan)
- osd: clean up internal ObjectStore interface (Sage Weil)
- osd: cleanup boost optionals (William Kennington)
- osd: clear cache on interval change (Samuel Just)
- osd: do no proxy reads unless target OSDs are new (#10788 Sage Weil)
- osd: do not abort deep scrub on missing hinfo (#10018 Loic Dachary)
- osd: do not update digest on inconsistent object (#10524 Samuel Just)
- osd: don't record digests for snapdirs (#10536 Samuel Just)
- osd: drop upgrade support for pre-dumpling (Sage Weil)
- osd: enable and use posix_fadvise (Sage Weil)
- osd: erasure coding: allow bench.sh to test ISA backend (Yuan Zhou)
- osd: erasure-code: encoding regression tests, corpus (#9420 Loic Dachary)
- osd: erasure-code: enforce chunk size alignment (#10211 Loic Dachary)
- osd: erasure-code: jerasure support for NEON (Loic Dachary)
- osd: erasure-code: relax cauchy w restrictions (#10325 David Zhang, Loic Dachary)
- osd: erasure-code: update gf-complete to latest upstream (Loic Dachary)
- osd: expose non-journal backends via ceph-osd CLI (Hoamai Wang)
- osd: filejournal: don't cache journal when not using direct IO (Jianpeng Ma)
- osd: fix JSON output for stray OSDs (Loic Dachary)

- osd: fix OSDCap parser on old (el6) boost::spirit (#10757 Kefu Chai)
- osd: fix OSDCap parsing on el6 (#10757 Kefu Chai)
- osd: fix ObjectStore::Transaction encoding version (#10734 Samuel Just)
- osd: fix WBTHrottle perf counters (Haomai Wang)
- osd: fix and document last_epoch_started semantics (Samuel Just)
- osd: fix auth object selection during repair (#10524 Samuel Just)
- osd: fix backfill bug (#10150 Samuel Just)
- osd: fix bug in pending digest updates (#10840 Samuel Just)
- osd: fix cancel_proxy_read_ops (Sage Weil)
- osd: fix cleanup of interrupted pg deletion (#10617 Sage Weil)
- osd: fix divergent entry handling on PG split (Samuel Just)
- osd: fix ghobject_t formatted output to include shard (#10063 Loic Dachary)
- osd: fix ioprio option (Mykola Golub)
- osd: fix ioprio options (Loic Dachary)
- osd: fix journal shutdown race (Sage Weil)
- osd: fix journal wrapping bug (#10883 David Zafman)
- osd: fix leak in SnapTrimWQ (#10421 Kefu Chai)
- osd: fix leak on shutdown (Kefu Chai)
- osd: fix memstore free space calculation (Xiaoxi Chen)
- osd: fix mixed-version peering issues (Samuel Just)
- osd: fix object age eviction (Zhiqiang Wang)
- osd: fix object atime calculation (Xinze Chi)
- osd: fix object digest update bug (#10840 Samuel Just)
- osd: fix occasional peering stalls (#10431 Sage Weil)
- osd: fix ordering issue with new transaction encoding (#10534 Dong Yuan)
- osd: fix osd peer check on scrub messages (#9555 Sage Weil)
- osd: fix past_interval display bug (#9752 Loic Dachary)
- osd: fix past_interval generation (#10427 #10430 David Zafman)
- osd: fix pgls filter ops (#9439 David Zafman)
- osd: fix recording of digest on scrub (Samuel Just)
- osd: fix scrub delay bug (#10693 Samuel Just)
- osd: fix scrub vs try-flush bug (#8011 Samuel Just)
- osd: fix short read handling on push (#8121 David Zafman)
- osd: fix stderr with -f or -d (Dan Mick)
- osd: fix transaction accounting (Jianpeng Ma)
- osd: fix watch reconnect race (#10441 Sage Weil)
- osd: fix watch timeout cache state update (#10784 David Zafman)
- osd: fix whiteout handling (Sage Weil)
- osd: flush snapshots from cache tier immediately (Sage Weil)
- osd: force promotion of watch/notify ops (Zhiqiang Wang)
- osd: handle no-op write with snapshot (#10262 Sage Weil)
- osd: improve idempotency detection across cache promotion/demotion (#8935 Sage Weil, Samuel Just)
- osd: include activating peers in blocked_by (#10477 Sage Weil)
- osd: jerasure and gf-complete updates from upstream (#10216 Loic Dachary)
- osd: journal: check fsync/fdatasync result (Jianpeng Ma)
- osd: journal: fix alignment checks, avoid useless memmove (Jianpeng Ma)
- osd: journal: fix hang on shutdown (#10474 David Zafman)
- osd: journal: fix header.committed_up_to (Xinze Chi)
- osd: journal: fix journal zeroing when direct IO is enabled (Xie Rui)
- osd: journal: initialize throttle (Ning Yao)
- osd: journal: misc bug fixes (#6003 David Zafman, Samuel Just)
- osd: journal: update committed_thru after replay (#6756 Samuel Just)
- osd: keyvaluestore: cleanup dead code (Ning Yao)
- osd: keyvaluestore: fix getattr semantics (Haomai Wang)
- osd: keyvaluestore: fix key ordering (#10119 Haomai Wang)
- osd: keyvaluestore_dev: optimization (Chendi Xue)
- osd: limit in-flight read requests (Jason Dillaman)
- osd: log when scrub or repair starts (Loic Dachary)
- osd: make misdirected op checks robust for EC pools (#9835 Sage Weil)
- osd: memstore: fix size limit (Xiaoxi Chen)
- osd: misc FIEMAP fixes (Ma Jianpeng)
- osd: misc cleanup (Xinze Chi, Yongyue Sun)
- osd: misc optimizations (Xinxin Shu, Zhiqiang Wang, Xinze Chi)
- osd: misc scrub fixes (#10017 Loic Dachary)
- osd: new 'activating' state between peering and active (Sage Weil)

- osd: new optimized encoding for ObjectStore::Transaction (Dong Yuan)
- osd: optimize Finisher (Xinze Chi)
- osd: optimize WBThrottle map with unordered_map (Ning Yao)
- osd: optimize filter_snapc (Ning Yao)
- osd: preserve reqids for idempotency checks for promote/demote (Sage Weil, Zhiqiang Wang, Samuel Just)
- osd: proxy read support (Zhiqiang Wang)
- osd: proxy reads during cache promote (Zhiqiang Wang)
- osd: remove dead locking code (Xinxin Shu)
- osd: remove legacy classic scrub code (Sage Weil)
- osd: remove unused fields in MOSDSubOp (Xiaoxi Chen)
- osd: removed some dead code (Xinze Chi)
- osd: replace MOSDSubOp messages with simpler, optimized MOSDRepOp (Xiaoxi Chen)
- osd: restrict scrub to certain times of day (Xinze Chi)
- osd: rocksdb: fix shutdown (Hoamai Wang)
- osd: store PG metadata in per-collection objects for better concurrency (Sage Weil)
- osd: store whole-object checksums on scrub, write_full (Sage Weil)
- osd: support for discard for journal trim (Jianpeng Ma)
- osd: use FIEMAP_FLAGS_SYNC instead of fsync (Jianpeng Ma)
- osd: verify kernel is new enough before using XFS extsize ioctl, enable by default (#9956 Sage Weil)
- pybind: fix memory leak in librados bindings (Billy Olsen)
- pyrados: add object lock support (#6114 Mehdi Abaakouk)
- pyrados: fix misnamed wait_* routings (#10104 Dan Mick)
- pyrados: misc cleanups (Kefu Chai)
- qa: add large auth ticket tests (Ilya Dryomov)
- qa: fix mds tests (#10539 John Spray)
- qa: fix osd create dup tests (#10083 Loic Dachary)
- qa: ignore duplicates in rados ls (Josh Durgin)
- qa: improve hadoop tests (Noah Watkins)
- qa: many 'make check' improvements (Loic Dachary)
- qa: misc tests (Loic Dachary, Yan, Zheng)
- qa: parallelize make check (Loic Dachary)
- qa: reorg fs quota tests (Greg Farnum)
- qa: tolerate nearly-full disk for make check (Loic Dachary)
- rados: fix put of /dev/null (Loic Dachary)
- rados: fix usage (Jianpeng Ma)
- rados: parse command-line arguments more strictly (#8983 Adam Crume)
- rados: use copy-from operation for copy, cppool (Sage Weil)
- radosgw-admin: add replicalog update command (Yehuda Sadeh)
- rbd-fuse: clean up on shutdown (Josh Durgin)
- rbd-fuse: fix memory leak (Adam Crume)
- rbd-replay-many (Adam Crume)
- rbd-replay: -anonymize flag to rbd-replay-prep (Adam Crume)
- rbd: add 'merge-diff' function (MingXin Liu, Yunchuan Wen, Li Wang)
- rbd: allow v2 striping parameters for clones and imports (Jason Dillaman)
- rbd: fix 'rbd diff' for non-existent objects (Adam Crume)
- rbd: fix buffer handling on image import (#10590 Jason Dillaman)
- rbd: fix error when striping with format 1 (Sebastien Han)
- rbd: fix export for image sizes over 2GB (Vicente Cheng)
- rbd: fix formatted output of image features (Jason Dillaman)
- rbd: leave exclusive lockin goff by default (Jason Dillaman)
- rbd: updat eman page (Ilya Dryomov)
- rbd: update init-rbdmap to fix dup mount point (Karel Striegel)
- rbd: use IO hints for import, export, and bench operations (#10462 Jason Dillaman)
- rbd: use rolling average for rbd bench-write throughput (Jason Dillaman)
- rbd_recover_tool: RBD image recovery tool (Min Chen)
- rgw: S3-style object versioning support (Yehuda Sadeh)
- rgw: add location header when object is in another region (VRan Liu)
- rgw: change multipart upload id magic (#10271 Yehuda Sadeh)
- rgw: check keystone auth for S3 POST requests (#10062 Abhishek Lekshmanan)
- rgw: check timestamp on s3 keystone auth (#10062 Abhishek Lekshmanan)
- rgw: conditional PUT on ETag (#8562 Ray Lv)
- rgw: create subuser if needed when creating user (#10103 Yehuda Sadeh)
- rgw: decode http query params correction (#10271 Yehuda Sadeh)
- rgw: don't overwrite bucket/object owner when setting ACLs (#10978 Yehuda Sadeh)
- rgw: enable IPv6 for civetweb (#10965 Yehuda Sadeh)

- rgw: extend replica log API (purge-all) (Yehuda Sadeh)
- rgw: fail S3 POST if keystone not configured (#10688 Valery Tschopp, Yehuda Sadeh)
- rgw: fix If-Modified-Since (VRan Liu)
- rgw: fix XML header on get ACL request (#10106 Yehuda Sadeh)
- rgw: fix bucket removal with data purge (Yehuda Sadeh)
- rgw: fix content length check (#10701 Axel Dunkel, Yehuda Sadeh)
- rgw: fix content-length update (#9576 Yehuda Sadeh)
- rgw: fix disabling of max_size quota (#9907 Dong Lei)
- rgw: fix error codes (#10334 #10329 Yehuda Sadeh)
- rgw: fix incorrect len when len is 0 (#9877 Yehuda Sadeh)
- rgw: fix object copy content type (#9478 Yehuda Sadeh)
- rgw: fix partial GET in swift (#10553 Yehuda Sadeh)
- rgw: fix replica log indexing (#8251 Yehuda Sadeh)
- rgw: fix shutdown (#10472 Yehuda Sadeh)
- rgw: fix swift metadata header name (Dmytro lurchenko)
- rgw: fix sysvinit script when rgw_socket_path is not defined (#11159 Yehuda Sadeh, Dan Mick)
- rgw: fix user stags in get-user-info API (#9359 Ray Lv)
- rgw: include XML ns on get ACL request (#10106 Yehuda Sadeh)
- rgw: index swift keys appropriately (#10471 Yehuda Sadeh)
- rgw: make sysvinit script set ulimit -n properly (Sage Weil)
- rgw: misc fixes (#10307 Yehuda Sadeh)
- rgw: only track cleanup for objects we write (#10311 Yehuda Sadeh)
- rgw: pass civetweb configurables through (#10907 Yehuda Sadeh)
- rgw: prevent illegal bucket policy that doesn't match placement rule (Yehuda Sadeh)
- rgw: remove multipart entries from bucket index on abort (#10719 Yehuda Sadeh)
- rgw: remove swift user manifest (DLO) hash calculation (#9973 Yehuda Sadeh)
- rgw: respond with 204 to POST on containers (#10667 Yuan Zhou)
- rgw: return timestamp on GET/HEAD (#8911 Yehuda Sadeh)
- rgw: reuse fcgx connection struct (#10194 Yehuda Sadeh)
- rgw: run radosgw as apache with systemd (#10125 Loic Dachary)
- rgw: send explicit HTTP status string (Yehuda Sadeh)
- rgw: set ETag on object copy (#9479 Yehuda Sadeh)
- rgw: set length for keystone token validation request (#7796 Yehuda Sadeh, Mark Kirkwood)
- rgw: support X-Storage-Policy header for Swift storage policy compat (Yehuda Sadeh)
- rgw: support multiple host names (#7467 Yehuda Sadeh)
- rgw: swift: dump container's custom metadata (#10665 Ahmad Faheem, Dmytro lurchenko)
- rgw: swift: support Accept header for response format (#10746 Dmytro lurchenko)
- rgw: swift: support for X-Remove-Container-Meta-{key} (#10475 Dmytro lurchenko)
- rgw: tweak error codes (#10329 #10334 Yehuda Sadeh)
- rgw: update bucket index on attr changes, for multi-site sync (#5595 Yehuda Sadeh)
- rgw: use rn for http headers (#9254 Yehuda Sadeh)
- rgw: use gc for multipart abort (#10445 Aaron Bassett, Yehuda Sadeh)
- rgw: use new watch/notify API (Yehuda Sadeh, Sage Weil)
- rpm: misc fixes (Key Dreyer)
- rpm: move rgw logrotate to radosgw subpackage (Ken Dreyer)
- systemd: better systemd unit files (Owen Synge)
- sysvinit: fix race in 'stop' (#10389 Loic Dachary)
- test: fix bufferlist tests (Jianpeng Ma)
- tests: ability to run unit tests under docker (Loic Dachary)
- tests: centos-6 dockerfile (#10755 Loic Dachary)
- tests: improve docker-based tests (Loic Dachary)
- tests: unit tests for shared_cache (Dong Yuan)
- udev: fix rules for CentOS7/RHEL7 (Loic Dachary)
- use clock_gettime instead of gettimeofday (Jianpeng Ma)
- vstart.sh: set up environment for s3-tests (Luis Pabon)
- vstart.sh: work with cmake (Yehuda Sadeh)

V0.93

This is the first release candidate for Hammer, and includes all of the features that will be present in the final release. We welcome and encourage any and all testing in non-production clusters to identify any problems with functionality, stability, or performance before the final Hammer release.

We suggest some caution in one area: librbd. There is a lot of new functionality around object maps and locking that is disabled

by default but may still affect stability for existing images. We are continuing to shake out those bugs so that the final Hammer release (probably v0.94) will be rock solid.

Major features since Giant include:

- cephfs: journal scavenger repair tool (John Spray)
- crush: new and improved straw2 bucket type (Sage Weil, Christina Anderson, Xiaoxi Chen)
- doc: improved guidance for CephFS early adopters (John Spray)
- librbd: add per-image object map for improved performance (Jason Dillaman)
- librbd: copy-on-read (Min Chen, Li Wang, Yunchuan Wen, Cheng Cheng)
- librados: fadvise-style IO hints (Jianpeng Ma)
- mds: many many snapshot-related fixes (Yan, Zheng)
- mon: new ‘ceph osd df’ command (Mykola Golub)
- mon: new ‘ceph pg ls ...’ command (Xinxin Shu)
- osd: improved performance for high-performance backends
- osd: improved recovery behavior (Samuel Just)
- osd: improved cache tier behavior with reads (Zhiqiang Wang)
- rgw: S3-compatible bucket versioning support (Yehuda Sadeh)
- rgw: large bucket index sharding (Guang Yang, Yehuda Sadeh)
- RDMA “xio” messenger support (Matt Benjamin, Vu Pham)

UPGRADING

- If you are upgrading from v0.92, you must stop all OSD daemons and flush their journals (`ceph-osd -i NNN --flush-journal`) before upgrading. There was a transaction encoding bug in v0.92 that broke compatibility. Upgrading from v0.91 or anything earlier is safe.
- No special restrictions when upgrading from firefly or giant.

NOTABLE CHANGES

- build: CMake support (Ali Maredia, Casey Bodley, Adam Emerson, Marcus Watts, Matt Benjamin)
- ceph-disk: do not re-use partition if encryption is required (Loic Dachary)
- ceph-disk: support LUKS for encrypted partitions (Andrew Bartlett, Loic Dachary)
- ceph-fuse,libcephfs: add support for O_NOFOLLOW and O_PATH (Greg Farnum)
- ceph-fuse,libcephfs: resend requests before completing cap reconnect (#10912 Yan, Zheng)
- ceph-fuse: select kernel cache invalidation mechanism based on kernel version (Greg Farnum)
- ceph-objectstore-tool: improved import (David Zafman)
- ceph-objectstore-tool: misc improvements, fixes (#9870 #9871 David Zafman)
- ceph: add ‘ceph osd df [tree]’ command (#10452 Mykola Golub)
- ceph: fix ‘ceph tell ...’ command validation (#10439 Joao Eduardo Luis)
- ceph: improve ‘ceph osd tree’ output (Mykola Golub)
- cephfs-journal-tool: add recover_dentries function (#9883 John Spray)
- common: add newline to flushed json output (Sage Weil)
- common: filtering for ‘perf dump’ (John Spray)
- common: fix Formatter factory breakage (#10547 Loic Dachary)
- common: make json-pretty output prettier (Sage Weil)
- crush: new and improved straw2 bucket type (Sage Weil, Christina Anderson, Xiaoxi Chen)
- crush: update tries stats for indep rules (#10349 Loic Dachary)
- crush: use larger choose_tries value for erasure code rulesets (#10353 Loic Dachary)
- debian,rpm: move RBD udev rules to ceph-common (#10864 Ken Dreyer)
- debian: split python-ceph into python-{rbd,rados,cephfs} (Boris Ranto)
- doc: CephFS disaster recovery guidance (John Spray)
- doc: CephFS for early adopters (John Spray)
- doc: fix OpenStack Glance docs (#10478 Sebastien Han)
- doc: misc updates (#9793 #9922 #10204 #10203 Travis Rhoden, Hazem, Ayari, Florian Coste, Andy Allan, Frank Yu, Baptiste Veuillez-Mainard, Yuan Zhou, Armando Segnini, Robert Jansen, Tyler Brekke, Viktor Suprun)
- doc: replace cloudfiles with swiftclient Python Swift example (Tim Freund)
- erasure-code: add mSHEC erasure code support (Takeshi Miyamae)
- erasure-code: improved docs (#10340 Loic Dachary)
- erasure-code: set max_size to 20 (#10363 Loic Dachary)
- libcephfs,ceph-fuse: fix getting zero-length xattr (#10552 Yan, Zheng)
- librados: add blacklist_add convenience method (Jason Dillaman)
- librados: expose rados_{read|write}_op_assert_version in C API (Kim Vandry)
- librados: fix pool name caching (#10458 Radoslaw Zarzynski)

- librados: fix resource leak, misc bugs (#10425 Radoslaw Zarzynski)
- librados: fix some watch/notify locking (Jason Dillaman, Josh Durgin)
- libradosstriper: fix write_full when ENOENT (#10758 Sebastien Ponce)
- librbd: CRC protection for RBD image map (Jason Dillaman)
- librbd: add per-image object map for improved performance (Jason Dillaman)
- librbd: add support for an “object map” indicating which objects exist (Jason Dillaman)
- librbd: adjust internal locking (Josh Durgin, Jason Dillaman)
- librbd: better handling of watch errors (Jason Dillaman)
- librbd: coordinate maint operations through lock owner (Jason Dillaman)
- librbd: copy-on-read (Min Chen, Li Wang, Yunchuan Wen, Cheng Cheng, Jason Dillaman)
- librbd: enforce write ordering with a snapshot (Jason Dillaman)
- librbd: fadvise-style hints; add misc hints for certain operations (Jianpeng Ma)
- librbd: fix coverity false-positives (Jason Dillaman)
- librbd: fix snap create races (Jason Dillaman)
- librbd: flush AIO operations asynchronously (#10714 Jason Dillaman)
- librbd: make async versions of long-running maint operations (Jason Dillaman)
- librbd: mock tests (Jason Dillaman)
- librbd: optionally blacklist clients before breaking locks (#10761 Jason Dillaman)
- librbd: prevent copyup during shrink (Jason Dillaman)
- mds: add cephfs-table-tool (John Spray)
- mds: avoid sending traceless replies in most cases (Yan, Zheng)
- mds: export dir asok command (John Spray)
- mds: fix stray/purge perfcounters (#10388 John Spray)
- mds: handle heartbeat_reset during shutdown (#10382 John Spray)
- mds: many many snapshot-related fixes (Yan, Zheng)
- mds: refactor, improve Session storage (John Spray)
- misc coverity fixes (Danny Al-Gaaf)
- mon: add noforward flag for some mon commands (Mykola Golub)
- mon: disallow empty pool names (#10555 Wido den Hollander)
- mon: do not deactivate last mds (#10862 John Spray)
- mon: drop old ceph_mon_store_converter (Sage Weil)
- mon: fix ‘ceph pg dump_stuck degraded’ (Xinxin Shu)
- mon: fix ‘profile osd’ use of config-key function on mon (#10844 Joao Eduardo Luis)
- mon: fix compatset initialization during mkfs (Joao Eduardo Luis)
- mon: fix feature tracking during elections (Joao Eduardo Luis)
- mon: fix mds gid/rank/state parsing (John Spray)
- mon: ignore failure reports from before up_from (#10762 Dan van der Ster, Sage Weil)
- mon: improved corrupt CRUSH map detection (Joao Eduardo Luis)
- mon: include pg_temp count in osdmap summary (Sage Weil)
- mon: log health summary to cluster log (#9440 Joao Eduardo Luis)
- mon: make ‘mds fail’ idempotent (John Spray)
- mon: make pg dump {sum,pgs,pgs_brief} work for format=plain (#5963 #6759 Mykola Golub)
- mon: new pool safety flags nodelete, nopgchange, nosizechange (#9792 Mykola Golub)
- mon: new, friendly ‘ceph pg ls ...’ command (Xinxin Shu)
- mon: prevent MDS transition from STOPPING (#10791 Greg Farnum)
- mon: propose all pending work in one transaction (Sage Weil)
- mon: remove pg_temps for nonexistent pools (Joao Eduardo Luis)
- mon: require mon_allow_pool_delete option to remove pools (Sage Weil)
- mon: set globalid prealloc to larger value (Sage Weil)
- mon: skip zeroed osd stats in get_rule_avail (#10257 Joao Eduardo Luis)
- mon: validate min_size range (Jianpeng Ma)
- msgr: async: bind threads to CPU cores, improved poll (Haomai Wang)
- msgr: fix crc configuration (Mykola Golub)
- msgr: misc unit tests (Haomai Wang)
- msgr: xio: XioMessenger RDMA support (Casey Bodley, Vu Pham, Matt Benjamin)
- osd, librados: fadvise-style librados hints (Jianpeng Ma)
- osd, librados: fix xattr_cmp_u64 (Dongmao Zhang)
- osd,mon: add ‘norebalance’ flag (Kefu Chai)
- osd,mon: specify OSD features explicitly in MOSDBoot (#10911 Sage Weil)
- osd: add option to prioritize heartbeat network traffic (Jian Wen)
- osd: add support for the SHEC erasure-code algorithm (Takeshi Miyamae, Loic Dachary)
- osd: allow recovery while below min_size (Samuel Just)
- osd: allow recovery with fewer than min_size OSDs (Samuel Just)
- osd: allow writes to degraded objects (Samuel Just)
- osd: allow writes to degraded objects (Samuel Just)

- osd: avoid publishing unchanged PG stats (Sage Weil)
- osd: cache recent ObjectContexts (Dong Yuan)
- osd: clear cache on interval change (Samuel Just)
- osd: do no proxy reads unless target OSDs are new (#10788 Sage Weil)
- osd: do not update digest on inconsistent object (#10524 Samuel Just)
- osd: don't record digests for snapdirs (#10536 Samuel Just)
- osd: fix OSDCap parser on old (el6) boost::spirit (#10757 Kefu Chai)
- osd: fix OSDCap parsing on el6 (#10757 Kefu Chai)
- osd: fix ObjectStore::Transaction encoding version (#10734 Samuel Just)
- osd: fix auth object selection during repair (#10524 Samuel Just)
- osd: fix bug in pending digest updates (#10840 Samuel Just)
- osd: fix cancel_proxy_read_ops (Sage Weil)
- osd: fix cleanup of interrupted pg deletion (#10617 Sage Weil)
- osd: fix journal wrapping bug (#10883 David Zafman)
- osd: fix leak in SnapTrimWQ (#10421 Kefu Chai)
- osd: fix memstore free space calculation (Xiaoxi Chen)
- osd: fix mixed-version peering issues (Samuel Just)
- osd: fix object digest update bug (#10840 Samuel Just)
- osd: fix ordering issue with new transaction encoding (#10534 Dong Yuan)
- osd: fix past_interval generation (#10427 #10430 David Zafman)
- osd: fix short read handling on push (#8121 David Zafman)
- osd: fix watch timeout cache state update (#10784 David Zafman)
- osd: force promotion of watch/notify ops (Zhiqiang Wang)
- osd: improve idempotency detection across cache promotion/demotion (#8935 Sage Weil, Samuel Just)
- osd: include activating peers in blocked_by (#10477 Sage Weil)
- osd: jerasure and gf-complete updates from upstream (#10216 Loic Dachary)
- osd: journal: check fsync/fdatasync result (Jianpeng Ma)
- osd: journal: fix hang on shutdown (#10474 David Zafman)
- osd: journal: fix header.committed_up_to (Xinze Chi)
- osd: journal: initialize throttle (Ning Yao)
- osd: journal: misc bug fixes (#6003 David Zafman, Samuel Just)
- osd: misc cleanup (Xinze Chi, Yongyue Sun)
- osd: new 'activating' state between peering and active (Sage Weil)
- osd: preserve reqids for idempotency checks for promote/demote (Sage Weil, Zhiqiang Wang, Samuel Just)
- osd: remove dead locking code (Xinxin Shu)
- osd: restrict scrub to certain times of day (Xinze Chi)
- osd: rocksdb: fix shutdown (Hoamai Wang)
- pybind: fix memory leak in librados bindings (Billy Olsen)
- qa: fix mds tests (#10539 John Spray)
- qa: ignore duplicates in rados ls (Josh Durgin)
- qa: improve hadoop tests (Noah Watkins)
- qa: reorg fs quota tests (Greg Farnum)
- rados: fix usage (Jianpeng Ma)
- radosgw-admin: add replicalog update command (Yehuda Sadeh)
- rbd-fuse: clean up on shutdown (Josh Durgin)
- rbd: add 'merge-diff' function (MingXin Liu, Yunchuan Wen, Li Wang)
- rbd: fix buffer handling on image import (#10590 Jason Dillaman)
- rbd: leave exclusive lockin goff by default (Jason Dillaman)
- rbd: update init-rbdmap to fix dup mount point (Karel Striegel)
- rbd: use IO hints for import, export, and bench operations (#10462 Jason Dillaman)
- rbd_recover_tool: RBD image recovery tool (Min Chen)
- rgw: S3-style object versioning support (Yehuda Sadeh)
- rgw: check keystone auth for S3 POST requests (#10062 Abhishek Lekshmanan)
- rgw: extend replica log API (purge-all) (Yehuda Sadeh)
- rgw: fail S3 POST if keystone not configured (#10688 Valery Tschopp, Yehuda Sadeh)
- rgw: fix XML header on get ACL request (#10106 Yehuda Sadeh)
- rgw: fix bucket removal with data purge (Yehuda Sadeh)
- rgw: fix replica log indexing (#8251 Yehuda Sadeh)
- rgw: fix swift metadata header name (Dmytro Iurchenko)
- rgw: remove multipart entries from bucket index on abort (#10719 Yehuda Sadeh)
- rgw: respond with 204 to POST on containers (#10667 Yuan Zhou)
- rgw: reuse fcgx connection struct (#10194 Yehuda Sadeh)
- rgw: support multiple host names (#7467 Yehuda Sadeh)
- rgw: swift: dump container's custom metadata (#10665 Ahmad Faheem, Dmytro Iurchenko)
- rgw: swift: support Accept header for response format (#10746 Dmytro Iurchenko)

- rgw: swift: support for X-Remove-Container-Meta-{key} (#10475 Dmytro Iurchenko)
- rpm: move rgw logrotate to radosgw subpackage (Ken Dreyer)
- tests: centos-6 dockerfile (#10755 Loic Dachary)
- tests: unit tests for shared_cache (Dong Yuan)
- vstart.sh: work with cmake (Yehuda Sadeh)

V0.92

This is the second-to-last chunk of new stuff before Hammer. Big items include additional checksums on OSD objects, proxied reads in the cache tier, image locking in RBD, optimized OSD Transaction and replication messages, and a big pile of RGW and MDS bug fixes.

UPGRADING

- The experimental ‘keyvaluestore-dev’ OSD backend has been renamed ‘keyvaluestore’ (for simplicity) and marked as experimental. To enable this untested feature and acknowledge that you understand that it is untested and may destroy data, you need to add the following to your ceph.conf:

```
enable experimental unrecoverable data corrupting featuers = keyvaluestore
```

- The following librados C API function calls take a ‘flags’ argument whose value is now correctly interpreted:

```
rados_write_op_operate() rados_aio_write_op_operate() rados_read_op_operate() rados_aio_read_op_operate()
```

The flags were not correctly being translated from the librados constants to the internal values. Now they are. Any code that is passing flags to these methods should be audited to ensure that they are using the correct LIBRADOS_OP_FLAG_* constants.

- The ‘rados’ CLI ‘copy’ and ‘cppool’ commands now use the copy-from operation, which means the latest CLI cannot run these commands against pre-firefly OSDs.
- The librados watch/notify API now includes a `watch_flush()` operation to flush the async queue of notify operations. This should be called by any watch/notify user prior to `rados_shutdown()`.

NOTABLE CHANGES

- add experimental features option (Sage Weil)
- build: fix ‘make check’ races (#10384 Loic Dachary)
- build: fix pkg names when libkeyutils is missing (Pankag Garg, Ken Dreyer)
- ceph: make ‘ceph -s’ show PG state counts in sorted order (Sage Weil)
- ceph: make ‘ceph tell mon.* version’ work (Mykola Golub)
- ceph-monstore-tool: fix/improve CLI (Joao Eduardo Luis)
- ceph: show primary-affinity in ‘ceph osd tree’ (Mykola Golub)
- common: add TableFormatter (Andreas Peters)
- common: check syncfs() return code (Jianpeng Ma)
- doc: do not suggest dangerous XFS nobarrier option (Dan van der Ster)
- doc: misc updates (Nilamdyuti Goswami, John Wilkins)
- install-deps.sh: do not require sudo when root (Loic Dachary)
- libcephfs: fix dirfrag trimming (#10387 Yan, Zheng)
- libcephfs: fix mount timeout (#10041 Yan, Zheng)
- libcephfs: fix test (#10415 Yan, Zheng)
- libcephfs: fix use-afer-free on umount (#10412 Yan, Zheng)
- libcephfs: include ceph and git version in client metadata (Sage Weil)
- librados: add `watch_flush()` operation (Sage Weil, Haomai Wang)
- librados: avoid memcpy on getxattr, read (Jianpeng Ma)
- librados: create iocx by pool id (Jason Dillaman)
- librados: do notify completion in fast-dispatch (Sage Weil)
- librados: remove shadowed variable (Kefu Chain)
- librados: translate op flags from C APIs (Matthew Richards)
- librbd: differentiate between R/O vs R/W features (Jason Dillaman)
- librbd: exclusive image locking (Jason Dillaman)
- librbd: fix write vs import race (#10590 Jason Dillaman)

- librbd: gracefully handle deleted/renamed pools (#10270 Jason Dillaman)
- mds: asok command for fetching subtree map (John Spray)
- mds: constify MDSCacheObjects (John Spray)
- misc: various valgrind fixes and cleanups (Danny Al-Gaaf)
- mon: fix ‘mds fail’ for standby MDSs (John Spray)
- mon: fix stashed monmap encoding (#5203 Xie Rui)
- mon: implement ‘fs reset’ command (John Spray)
- mon: respect down flag when promoting standbys (John Spray)
- mount.ceph: fix suprious error message (#10351 Yan, Zheng)
- msgr: async: many fixes, unit tests (Haomai Wang)
- msgr: simple: retry binding to port on failure (#10029 Wido den Hollander)
- osd: add fadvise flags to ObjectStore API (Jianpeng Ma)
- osd: add get_latest_osdmap asok command (#9483 #9484 Mykola Golub)
- osd: EIO on whole-object reads when checksum is wrong (Sage Weil)
- osd: filejournal: don’t cache journal when not using direct IO (Jianpeng Ma)
- osd: fix ioprio option (Mykola Golub)
- osd: fix scrub delay bug (#10693 Samuel Just)
- osd: fix watch reconnect race (#10441 Sage Weil)
- osd: handle no-op write with snapshot (#10262 Sage Weil)
- osd: journal: fix journal zeroing when direct IO is enabled (Xie Rui)
- osd: keyvaluestore: cleanup dead code (Ning Yao)
- osd, mds: ‘ops’ as shorthand for ‘dump_ops_in_flight’ on asok (Sage Weil)
- osd: memstore: fix size limit (Xiaoxi Chen)
- osd: misc scrub fixes (#10017 Loic Dachary)
- osd: new optimized encoding for ObjectStore::Transaction (Dong Yuan)
- osd: optimize filter_snapc (Ning Yao)
- osd: optimize WBThrottle map with unordered_map (Ning Yao)
- osd: proxy reads during cache promote (Zhiqiang Wang)
- osd: proxy read support (Zhiqiang Wang)
- osd: remove legacy classic scrub code (Sage Weil)
- osd: remove unused fields in MOSDSubOp (Xiaoxi Chen)
- osd: replace MOSDSubOp messages with simpler, optimized MOSDRepOp (Xiaoxi Chen)
- osd: store whole-object checksums on scrub, write_full (Sage Weil)
- osd: verify kernel is new enough before using XFS extsize ioctl, enable by default (#9956 Sage Weil)
- rados: use copy-from operation for copy, cpool (Sage Weil)
- rgw: change multipart upload id magic (#10271 Yehuda Sadeh)
- rgw: decode http query params correction (#10271 Yehuda Sadeh)
- rgw: fix content length check (#10701 Axel Dunkel, Yehuda Sadeh)
- rgw: fix partial GET in swift (#10553 Yehuda Sadeh)
- rgw: fix shutdown (#10472 Yehuda Sadeh)
- rgw: include XML ns on get ACL request (#10106 Yehuda Sadeh)
- rgw: misc fixes (#10307 Yehuda Sadeh)
- rgw: only track cleanup for objects we write (#10311 Yehuda Sadeh)
- rgw: tweak error codes (#10329 #10334 Yehuda Sadeh)
- rgw: use gc for multipart abort (#10445 Aaron Bassett, Yehuda Sadeh)
- sysvinit: fix race in ‘stop’ (#10389 Loic Dachary)
- test: fix bufferlist tests (Jianpeng Ma)
- tests: improve docker-based tests (Loic Dachary)

V0.91

We are quickly approaching the Hammer feature freeze but have a few more dev releases to go before we get there. The headline items are subtree-based quota support in CephFS (ceph-fuse/libcephfs client support only for now), a rewrite of the watch/notify librados API used by RBD and RGW, OSDMap checksums to ensure that maps are always consistent inside the cluster, new API calls in librados and librbd for IO hinting modeled after posix_fadvise, and improved storage of per-PG state.

We expect two more releases before the Hammer feature freeze (v0.93).

UPGRADING

- The ‘category’ field for objects has been removed. This was originally added to track PG stat summations over different categories of objects for use by radosgw. It is no longer has any known users and is prone to abuse because it can lead to a pg_stat_t structure that is unbounded. The librados API calls that accept this field now ignore it, and the OSD no longer

tracks the per-category summations.

- The output for ‘rados df’ has changed. The ‘category’ level has been eliminated, so there is now a single stat object per pool. The structure of the JSON output is different, and the plaintext output has one less column.
- The ‘rados create <objectname> [category]’ optional category argument is no longer supported or recognized.
- rados.py’s Rados class no longer has a `__del__` method; it was causing problems on interpreter shutdown and use of threads. If your code has Rados objects with limited lifetimes and you’re concerned about locked resources, call `Rados.shutdown()` explicitly.
- There is a new version of the librados watch/notify API with vastly improved semantics. Any applications using this interface are encouraged to migrate to the new API. The old API calls are marked as deprecated and will eventually be removed.
- The librados `rados_unwatch()` call used to be safe to call on an invalid handle. The new version has undefined behavior when passed a bogus value (for example, when `rados_watch()` returns an error and handle is not defined).
- The structure of the formatted ‘pg stat’ command is changed for the portion that counts states by name to avoid using the ‘+’ character (which appears in state names) as part of the XML token (it is not legal).

NOTABLE CHANGES

- `asyncmsgc`: misc fixes (Haomai Wang)
- `buffer`: add ‘shareable’ construct (Matt Benjamin)
- `build`: aarch64 build fixes (Noah Watkins, Haomai Wang)
- `build`: support for jemalloc (Shishir Gowda)
- `ceph-disk`: allow journal partition re-use (#10146 Loic Dachary, Dav van der Ster)
- `ceph-disk`: misc fixes (Christos Stavrakakis)
- `ceph-fuse`: fix kernel cache trimming (#10277 Yan, Zheng)
- `ceph-objectstore-tool`: many many improvements (David Zafman)
- `common`: support new gperftools header locations (Key Dreyer)
- `crush`: straw bucket weight calculation fixes (#9998 Sage Weil)
- `doc`: misc improvements (Nilamdyuti Goswami, John Wilkins, Chris Holcombe)
- `libcephfs,ceph-fuse`: add ‘status’ asok (John Spray)
- `librados, osd`: new watch/notify implementation (Sage Weil)
- `librados`: drop ‘category’ feature (Sage Weil)
- `librados`: fix pool deletion handling (#10372 Sage Weil)
- `librados`: new fadvise API (Ma Jianpeng)
- `librados_striper`: fix remove() (Dongmao Zhang)
- `librbd`: complete pending ops before closing image (#10299 Josh Durgin)
- `librbd`: fadvise API (Ma Jianpeng)
- `mds`: ENOSPC and OSDMap epoch barriers (#7317 John Spray)
- `mds`: dirfrag buf fix (Yan, Zheng)
- `mds`: disallow most commands on inactive MDS’s (Greg Farnum)
- `mds`: drop dentries, leases on deleted directories (#10164 Yan, Zheng)
- `mds`: handle zero-size xattr (#10335 Yan, Zheng)
- `mds`: subtree quota support (Yunchuan Wen)
- `memstore`: free space tracking (John Spray)
- `misc` cleanup (Danny Al-Gaaf, David Anderson)
- `mon`: ‘osd crush reweight-all’ command (Sage Weil)
- `mon`: allow full flag to be manually cleared (#9323 Sage Weil)
- `mon`: delay failure injection (Joao Eduardo Luis)
- `mon`: fix paxos timeouts (#10220 Joao Eduardo Luis)
- `mon`: get canonical OSDMap from leader (#10422 Sage Weil)
- `msgr`: fix RESETSESSION bug (#10080 Greg Farnum)
- `objectstore`: deprecate collection attrs (Sage Weil)
- `osd, mon`: add checksums to all OSDMaps (Sage Weil)
- `osd`: allow deletion of objects with watcher (#2339 Sage Weil)
- `osd`: allow sparse read for Push/Pull (Haomai Wang)
- `osd`: cache reverse_nibbles hash value (Dong Yuan)
- `osd`: drop upgrade support for pre-dumpling (Sage Weil)
- `osd`: enable and use posix_fadvise (Sage Weil)
- `osd: erasure-code`: enforce chunk size alignment (#10211 Loic Dachary)
- `osd: erasure-code`: jerasure support for NEON (Loic Dachary)
- `osd: erasure-code`: relax cauchy w restrictions (#10325 David Zhang, Loic Dachary)
- `osd: erasure-code`: update gf-complete to latest upstream (Loic Dachary)
- `osd`: fix WBTHrottle perf counters (Haomai Wang)
- `osd`: fix backfill bug (#10150 Samuel Just)
- `osd`: fix occasional peering stalls (#10431 Sage Weil)

- osd: fix scrub vs try-flush bug (#8011 Samuel Just)
- osd: fix stderr with -f or -d (Dan Mick)
- osd: misc FIEMAP fixes (Ma Jianpeng)
- osd: optimize Finisher (Xinze Chi)
- osd: store PG metadata in per-collection objects for better concurrency (Sage Weil)
- pyrados: add object lock support (#6114 Mehdi Abaakouk)
- pyrados: fix misnamed wait_* routings (#10104 Dan Mick)
- pyrados: misc cleanups (Kefu Chai)
- qa: add large auth ticket tests (Ilya Dryomov)
- qa: many ‘make check’ improvements (Loic Dachary)
- qa: misc tests (Loic Dachary, Yan, Zheng)
- rgw: conditional PUT on ETag (#8562 Ray Lv)
- rgw: fix error codes (#10334 #10329 Yehuda Sadeh)
- rgw: index swift keys appropriately (#10471 Yehuda Sadeh)
- rgw: prevent illegal bucket policy that doesn’t match placement rule (Yehuda Sadeh)
- rgw: run radosgw as apache with systemd (#10125 Loic Dachary)
- rgw: support X-Storage-Policy header for Swift storage policy compat (Yehuda Sadeh)
- rgw: use rn for http headers (#9254 Yehuda Sadeh)
- rpm: misc fixes (Key Dreyer)

V0.90

This is the last development release before Christmas. There are some API cleanups for librados and librbd, and lots of bug fixes across the board for the OSD, MDS, RGW, and CRUSH. The OSD also gets support for discard (potentially helpful on SSDs, although it is off by default), and there are several improvements to ceph-disk.

The next two development releases will be getting a slew of new functionality for hammer. Stay tuned!

UPGRADING

- Previously, the formatted output of ‘ceph pg stat -f ...’ was a full pg dump that included all metadata about all PGs in the system. It is now a concise summary of high-level PG stats, just like the unformatted ‘ceph pg stat’ command.
- All JSON dumps of floating point values were incorrectly surrounding the value with quotes. These quotes have been removed. Any consumer of structured JSON output that was consuming the floating point values was previously having to interpret the quoted string and will most likely need to be fixed to take the unquoted number.

NOTABLE CHANGES

- arch: fix NEON feaeture detection (#10185 Loic Dachary)
- build: adjust build deps for yasm, virtualenv (Jianpeng Ma)
- build: improve build dependency tooling (Loic Dachary)
- ceph-disk: call partx/partprobe consistency (#9721 Loic Dachary)
- ceph-disk: fix dmcrypt key permissions (Loic Dachary)
- ceph-disk: fix umount race condition (#10096 Blaine Gardner)
- ceph-disk: init=none option (Loic Dachary)
- ceph-monstore-tool: fix shutdown (#10093 Loic Dachary)
- ceph-objectstore-tool: fix import (#10090 David Zafman)
- ceph-objectstore-tool: many improvements and tests (David Zafman)
- ceph.spec: package rbd-replay-prep (Ken Dreyer)
- common: add ‘perf reset ...’ admin command (Jianpeng Ma)
- common: do not unlock rlock on destruction (Federico Simoncelli)
- common: fix block device discard check (#10296 Sage Weil)
- common: remove broken CEPH_LOCKDEP optoin (Kefu Chai)
- crush: fix tree bucket behavior (Rongze Zhu)
- doc: add build-doc guidlines for Fedora and CentOS/RHEL (Nilamdyuti Goswami)
- doc: enable rbd cache on openstack deployments (Sebastien Han)
- doc: improved installation nots on CentOS/RHEL installs (John Wilkins)
- doc: misc cleanups (Adam Spiers, Sebastien Han, Nilamdyuti Goswami, Ken Dreyer, John Wilkins)
- doc: new man pages (Nilamdyuti Goswami)
- doc: update release descriptions (Ken Dreyer)
- doc: update sepia hardware inventory (Sandon Van Ness)
- librados: only export public API symbols (Jason Dillaman)

- libradosstriper: fix stat strtol (Dongmao Zhang)
- libradosstriper: fix trunc method (#10129 Sebastien Ponce)
- librbd: fix list_children from invalid pool iocctxs (#10123 Jason Dillaman)
- librbd: only export public API symbols (Jason Dillaman)
- many coverity fixes (Danny Al-Gaaf)
- mds: ‘flush journal’ admin command (John Spray)
- mds: fix MDLog IO callback deadlock (John Spray)
- mds: fix deadlock during journal probe vs purge (#10229 Yan, Zheng)
- mds: fix race trimming log segments (Yan, Zheng)
- mds: store backtrace for stray dir (Yan, Zheng)
- mds: verify backtrace when fetching dirfrag (#9557 Yan, Zheng)
- mon: add max pgs per osd warning (Sage Weil)
- mon: fix _ratio units and types (Sage Weil)
- mon: fix JSON dumps to dump floats as flots and not strings (Sage Weil)
- mon: fix formatter ‘pg stat’ command output (Sage Weil)
- msgr: async: several fixes (Haomai Wang)
- msgr: simple: fix rare deadlock (Greg Farnum)
- osd: batch pg log trim (Xinze Chi)
- osd: clean up internal ObjectStore interface (Sage Weil)
- osd: do not abort deep scrub on missing hinfo (#10018 Loic Dachary)
- osd: fix ghobject_t formatted output to include shard (#10063 Loic Dachary)
- osd: fix osd peer check on scrub messages (#9555 Sage Weil)
- osd: fix pgls filter ops (#9439 David Zafman)
- osd: flush snapshots from cache tier immediately (Sage Weil)
- osd: keyvaluestore: fix getattr semantics (Haomai Wang)
- osd: keyvaluestore: fix key ordering (#10119 Haomai Wang)
- osd: limit in-flight read requests (Jason Dillaman)
- osd: log when scrub or repair starts (Loic Dachary)
- osd: support for discard for journal trim (Jianpeng Ma)
- qa: fix osd create dup tests (#10083 Loic Dachary)
- rgw: add location header when object is in another region (VRan Liu)
- rgw: check timestamp on s3 keystone auth (#10062 Abhishek Lekshmanan)
- rgw: make sysvinit script set ulimit -n properly (Sage Weil)
- systemd: better systemd unit files (Owen Synge)
- tests: ability to run unit tests under docker (Loic Dachary)

V0.89

This is the second development release since Giant. The big items include the first batch of scrub patches from Greg for CephFS, a rework in the librados object listing API to properly handle namespaces, and a pile of bug fixes for RGW. There are also several smaller issues fixed up in the performance area with buffer alignment and memory copies, osd cache tiering agent, and various CephFS fixes.

UPGRADING

- New ability to list all objects from all namespaces can fail or return incomplete results when not all OSDs have been upgraded. Features rados -all ls, rados cppool, rados export, rados cache-flush-evict-all and rados cache-try-flush-evict-all can also fail or return incomplete results.

NOTABLE CHANGES

- buffer: add list::get_contiguous (Sage Weil)
- buffer: avoid rebuild if buffer already contiguous (Jianpeng Ma)
- ceph-disk: improved systemd support (Owen Synge)
- ceph-disk: set guid if reusing journal partition (Dan van der Ster)
- ceph-fuse, libcephfs: allow xattr caps in inject_release_failure (#9800 John Spray)
- ceph-fuse, libcephfs: fix I_COMPLETE_ORDERED checks (#9894 Yan, Zheng)
- ceph-fuse: fix dentry invalidation on 3.18+ kernels (#9997 Yan, Zheng)
- crush: fix detach_bucket (#10095 Sage Weil)
- crush: fix several bugs in adjust_item_weight (Rongze Zhu)
- doc: add dumpling to firefly upgrade section (#7679 John Wilkins)
- doc: document erasure coded pool operations (#9970 Loic Dachary)

- doc: file system osd config settings (Kevin Dalley)
- doc: key/value store config reference (John Wilkins)
- doc: update openstack docs for Juno (Sebastien Han)
- fix cluster logging from non-mon daemons (Sage Weil)
- init-ceph: check for systemd-run before using it (Boris Ranto)
- librados: fix infinite loop with skipped map epochs (#9986 Ding Dinghua)
- librados: fix iterator operator= bugs (#10082 David Zafman, Yehuda Sadeh)
- librados: fix null deref when pool DNE (#9944 Sage Weil)
- librados: fix timer race from recent refactor (Sage Weil)
- libradosstriper: fix shutdown hang (Dongmao Zhang)
- librbd: don't close a closed parent in failure path (#10030 Jason Dillaman)
- librbd: fix diff test (#10002 Josh Durgin)
- librbd: fix locking for readahead (#10045 Jason Dillaman)
- librbd: refactor unit tests to use fixtures (Jason Dillaman)
- many many coverity cleanups (Danny Al-Gaaf)
- mds: a whole bunch of initial scrub infrastructure (Greg Farnum)
- mds: fix compat_version for MClientSession (#9945 John Spray)
- mds: fix reply snapbl (Yan, Zheng)
- mon: allow adding tiers to fs pools (#10135 John Spray)
- mon: fix MDS health status from peons (#10151 John Spray)
- mon: fix caching for min_last_epoch_clean (#9987 Sage Weil)
- mon: fix error output for add_data_pool (#9852 Joao Eduardo Luis)
- mon: include entity name in audit log for forwarded requests (#9913 Joao Eduardo Luis)
- mon: paxos: allow reads while proposing (#9321 #9322 Joao Eduardo Luis)
- msgr: asyncmessenger: add kqueue support (#9926 Haomai Wang)
- osd, librados: revamp PG listing API to handle namespaces (#9031 #9262 #9438 David Zafman)
- osd, mon: send intial pg create time from mon to osd (#9887 David Zafman)
- osd: allow whiteout deletion in cache pool (Sage Weil)
- osd: cache pool: ignore min flush age when cache is full (Xinze Chi)
- osd: erasure coding: allow bench.sh to test ISA backend (Yuan Zhou)
- osd: erasure-code: encoding regression tests, corpus (#9420 Loic Dachary)
- osd: fix journal shutdown race (Sage Weil)
- osd: fix object age eviction (Zhiqiang Wang)
- osd: fix object atime calculation (Xinze Chi)
- osd: fix past_interval display bug (#9752 Loic Dachary)
- osd: journal: fix alignment checks, avoid useless memmove (Jianpeng Ma)
- osd: journal: update committed_thru after replay (#6756 Samuel Just)
- osd: keyvaluestore_dev: optimization (Chendi Xue)
- osd: make misdirected op checks robust for EC pools (#9835 Sage Weil)
- osd: removed some dead code (Xinze Chi)
- qa: parallelize make check (Loic Dachary)
- qa: tolerate nearly-full disk for make check (Loic Dachary)
- rgw: create subuser if needed when creating user (#10103 Yehuda Sadeh)
- rgw: fix If-Modified-Since (VRan Liu)
- rgw: fix content-length update (#9576 Yehuda Sadeh)
- rgw: fix disabling of max_size quota (#9907 Dong Lei)
- rgw: fix incorrect len when len is 0 (#9877 Yehuda Sadeh)
- rgw: fix object copy content type (#9478 Yehuda Sadeh)
- rgw: fix user stags in get-user-info API (#9359 Ray Lv)
- rgw: remove swift user manifest (DLO) hash calculation (#9973 Yehuda Sadeh)
- rgw: return timestamp on GET/HEAD (#8911 Yehuda Sadeh)
- rgw: set ETag on object copy (#9479 Yehuda Sadeh)
- rgw: update bucket index on attr changes, for multi-site sync (#5595 Yehuda Sadeh)

V0.88

This is the first development release after Giant. The two main features merged this round are the new AsyncMessenger (an alternative implementation of the network layer) from Haomai Wang at UnitedStack, and support for POSIX file locks in ceph-fuse and libcephfs from Yan, Zheng. There is also a big pile of smaller items that re merged while we were stabilizing Giant, including a range of smaller performance and bug fixes and some new tracepoints for LTTNG.

NOTABLE CHANGES

- ceph-disk: Scientific Linux support (Dan van der Ster)
- ceph-disk: respect -statedir for keyring (Loic Dachary)
- ceph-fuse, libcephfs: POSIX file lock support (Yan, Zheng)
- ceph-fuse, libcephfs: fix cap flush overflow (Greg Farnum, Yan, Zheng)
- ceph-fuse, libcephfs: fix root inode xattrs (Yan, Zheng)
- ceph-fuse, libcephfs: preserve dir ordering (#9178 Yan, Zheng)
- ceph-fuse, libcephfs: trim inodes before reconnecting to MDS (Yan, Zheng)
- ceph: do not parse injectargs twice (Loic Dachary)
- ceph: make 'ceph -s' output more readable (Sage Weil)
- ceph: new 'ceph tell mds.\$name_or_rank_or_gid' (John Spray)
- ceph: test robustness (Joao Eduardo Luis)
- ceph_objectstore_tool: behave with sharded flag (#9661 David Zafman)
- cephfs-journal-tool: fix journal import (#10025 John Spray)
- cephfs-journal-tool: skip up to expire_pos (#9977 John Spray)
- cleanup rados.h definitions with macros (Ilya Dryomov)
- common: shared_cache unit tests (Cheng Cheng)
- config: add \$cctid meta variable (Adam Crume)
- crush: fix buffer overrun for poorly formed rules (#9492 John George)
- crush: improve constness (Loic Dachary)
- crushtool: add -location <id> command (Sage Weil, Loic Dachary)
- default to libnss instead of crypto++ (Federico Gimenez)
- doc: ceph osd reweight vs crush weight (Laurent Guerby)
- doc: document the LRC per-layer plugin configuration (Yuan Zhou)
- doc: erasure code doc updates (Loic Dachary)
- doc: misc updates (Alfredo Deza, VRan Liu)
- doc: preflight doc fixes (John Wilkins)
- doc: update PG count guide (Gerbent Meijer, Laurent Guerby, Loic Dachary)
- keyvaluestore: misc fixes (Haomai Wang)
- keyvaluestore: performance improvements (Haomai Wang)
- librados: add rados_pool_get_base_tier() call (Adam Crume)
- librados: cap buffer length (Loic Dachary)
- librados: fix objecter races (#9617 Josh Durgin)
- libradosstriper: misc fixes (Sebastien Ponce)
- librbd: add missing python docstrings (Jason Dillaman)
- librbd: add readahead (Adam Crume)
- librbd: fix cache tiers in list_children and snap_unprotect (Adam Crume)
- librbd: fix performance regression in ObjectCacher (#9513 Adam Crume)
- librbd: lttng tracepoints (Adam Crume)
- librbd: misc fixes (Xinxin Shu, Jason Dillaman)
- mds: fix sessionmap lifecycle bugs (Yan, Zheng)
- mds: initialize root inode xattr version (Yan, Zheng)
- mds: introduce auth caps (John Spray)
- mds: misc bugs (Greg Farnum, John Spray, Yan, Zheng, Henry Change)
- misc coverity fixes (Danny Al-Gaaf)
- mon: add 'ceph osd rename-bucket ...' command (Loic Dachary)
- mon: clean up auth list output (Loic Dachary)
- mon: fix 'osd crush link' id resolution (John Spray)
- mon: fix misc error paths (Joao Eduardo Luis)
- mon: fix paxos off-by-one corner case (#9301 Sage Weil)
- mon: new 'ceph pool ls [detail]' command (Sage Weil)
- mon: wait for writeable before cross-proposing (#9794 Joao Eduardo Luis)
- msgr: avoid useless new/delete (Haomai Wang)
- msgr: fix delay injection bug (#9910 Sage Weil, Greg Farnum)
- msgr: new AsymcMessenger alternative implementation (Haomai Wang)
- msgr: prefetch data when doing recv (Yehuda Sadeh)
- osd: add erasure code corpus (Loic Dachary)
- osd: add misc tests (Loic Dachary, Danny Al-Gaaf)
- osd: cleanup boost optionals (William Kennington)
- osd: expose non-journal backends via ceph-osd CLI (Hoamai Wang)
- osd: fix JSON output for stray OSDs (Loic Dachary)
- osd: fix ioprio options (Loic Dachary)
- osd: fix transaction accounting (Jianpeng Ma)
- osd: misc optimizations (Xinxin Shu, Zhiqiang Wang, Xinze Chi)
- osd: use FIEMAP_FLAGS_SYNC instead of fsync (Jianpeng Ma)
- rados: fix put of /dev/null (Loic Dachary)

- rados: parse command-line arguments more strictly (#8983 Adam Crume)
- rbd-fuse: fix memory leak (Adam Crume)
- rbd-replay-many (Adam Crume)
- rbd-replay: -anonymize flag to rbd-replay-prep (Adam Crume)
- rbd: fix ‘rbd diff’ for non-existent objects (Adam Crume)
- rbd: fix error when striping with format 1 (Sebastien Han)
- rbd: fix export for image sizes over 2GB (Vicente Cheng)
- rbd: use rolling average for rbd bench-write throughput (Jason Dillaman)
- rgw: send explicit HTTP status string (Yehuda Sadeh)
- rgw: set length for keystone token validation request (#7796 Yehuda Sadeh, Mark Kirkwood)
- udev: fix rules for CentOS7/RHEL7 (Loic Dachary)
- use clock_gettime instead of gettimeofday (Jianpeng Ma)
- vstart.sh: set up environment for s3-tests (Luis Pabon)

V0.87.2 GIANT

This is the second (and possibly final) point release for Giant.

We recommend all v0.87.x Giant users upgrade to this release.

NOTABLE CHANGES

- ceph-objectstore-tool: only output unsupported features when incompatible (#11176 David Zafman)
- common: do not implicitly unlock rwlock on destruction (Federico Simoncelli)
- common: make wait timeout on empty queue configurable (#10818 Samuel Just)
- crush: pick ruleset id that matches and rule id (Xiaoxi Chen)
- crush: set_choose_tries = 100 for new erasure code rulesets (#10353 Loic Dachary)
- librados: check initialized atomic safely (#9617 Josh Durgin)
- librados: fix failed tick_event assert (#11183 Zhiqiang Wang)
- librados: fix looping on skipped maps (#9986 Ding Dinghua)
- librados: fix op submit with timeout (#10340 Samuel Just)
- librados: pybind: fix memory leak (#10723 Billy Olsen)
- librados: pybind: keep reference to callbacks (#10775 Josh Durgin)
- librados: translate operation flags from C APIs (Matthew Richards)
- libradosstriper: fix write_full on ENOENT (#10758 Sebastien Ponce)
- libradosstriper: use strtoll instead of strtol (Dongmao Zhang)
- mds: fix assertion caused by system time moving backwards (#11053 Yan, Zheng)
- mon: allow injection of random delays on writes (Joao Eduardo Luis)
- mon: do not trust small osd epoch cache values (#10787 Sage Weil)
- mon: fail non-blocking flush if object is being scrubbed (#8011 Samuel Just)
- mon: fix division by zero in stats dump (Joao Eduardo Luis)
- mon: fix get_rule_avail when no osds (#10257 Joao Eduardo Luis)
- mon: fix timeout rounds period (#10546 Joao Eduardo Luis)
- mon: ignore osd failures before up_from (#10762 Dan van der Ster, Sage Weil)
- mon: paxos: reset accept timeout before writing to store (#10220 Joao Eduardo Luis)
- mon: return if fs exists on ‘fs new’ (Joao Eduardo Luis)
- mon: use EntityName when expanding profiles (#10844 Joao Eduardo Luis)
- mon: verify cross-service proposal preconditions (#10643 Joao Eduardo Luis)
- mon: wait for osdmon to be writeable when requesting proposal (#9794 Joao Eduardo Luis)
- mount.ceph: avoid spurious error message about /etc/mtab (#10351 Yan, Zheng)
- msg/simple: allow RESETSESSION when we forget an endpoint (#10080 Greg Farnum)
- msg/simple: discard delay queue before incoming queue (#9910 Sage Weil)
- osd: clear_primary_state when leaving Primary (#10059 Samuel Just)
- osd: do not ignore deleted pgs on startup (#10617 Sage Weil)
- osd: fix FileJournal wrap to get header out first (#10883 David Zafman)
- osd: fix PG leak in SnapTrimWQ (#10421 Kefu Chai)
- osd: fix journalq population in do_read_entry (#6003 Samuel Just)
- osd: fix operator== for op_queue_age_hit and fs_perf_stat (#10259 Samuel Just)
- osd: fix rare assert after split (#10430 David Zafman)
- osd: get pgid ancestor from last_map when building past intervals (#10430 David Zafman)
- osd: include rollback_info_trimmed_to in {read,write}_log (#10157 Samuel Just)
- osd: lock header_lock in DBObjectMap::sync (#9891 Samuel Just)
- osd: requeue blocked op before flush it was blocked on (#10512 Sage Weil)

- osd: tolerate missing object between list and attr get on backfill (#10150 Samuel Just)
- osd: use correct atime for eviction decision (Xinze Chi)
- rgw: flush XML header on get ACL request (#10106 Yehuda Sadeh)
- rgw: index swift keys appropriately (#10471 Hemant Bruman, Yehuda Sadeh)
- rgw: send cancel for bucket index pending ops (#10770 Baijiaruo, Yehuda Sadeh)
- rgw: swift: support X_Remove_Container-Meta-{key} (#01475 Dmytro lurchenko)

For more detailed information, see [the complete changelog](#).

V0.87.1 GIANT

This is the first (and possibly final) point release for Giant. Our focus on stability fixes will be directed towards Hammer and Firefly.

We recommend that all v0.87 Giant users upgrade to this release.

UPGRADING

- Due to a change in the Linux kernel version 3.18 and the limits of the FUSE interface, ceph-fuse needs be mounted as root on at least some systems. See issues #9997, #10277, and #10542 for details.

NOTABLE CHANGES

- build: disable stack-execute bit on assembler objects (#10114 Dan Mick)
- build: support boost 1.57.0 (#10688 Ken Dreyer)
- ceph-disk: fix dmcrypt file permissions (#9785 Loic Dachary)
- ceph-disk: run partprobe after zap, behave with partx or partprobe (#9665 #9721 Loic Dachary)
- cephfs-journal-tool: fix import for aged journals (#9977 John Spray)
- cephfs-journal-tool: fix journal import (#10025 John Spray)
- ceph-fuse: use remount to trim kernel dcache (#10277 Yan, Zheng)
- common: add cctid meta variable (#6228 Adam Crume)
- common: fix dump of shard for gobject_t (#10063 Loic Dachary)
- crush: fix bucket weight underflow (#9998 Pawel Sadowski)
- erasure-code: enforce chunk size alignment (#10211 Loic Dachary)
- erasure-code: regression test suite (#9420 Loic Dachary)
- erasure-code: relax caucy w restrictions (#10325 Loic Dachary)
- libcephfs,ceph-fuse: allow xattr caps on inject_release_failure (#9800 John Spray)
- libcephfs,ceph-fuse: fix cap flush tid comparison (#9869 Greg Farnum)
- libcephfs,ceph-fuse: new flag to indicated sorted dcache (#9178 Yan, Zheng)
- libcephfs,ceph-fuse: prune cache before reconnecting to MDS (Yan, Zheng)
- librados: limit number of in-flight read requests (#9854 Jason Dillaman)
- libradospy: fix thread shutdown (#8797 Dan Mick)
- libradosstriper: fix locking issue in truncate (#10129 Sebastien Ponce)
- librbd: complete pending ops before closing mage (#10299 Jason Dillaman)
- librbd: fix error path on image open failure (#10030 Jason Dillaman)
- librbd: gracefully handle deleted/renamed pools (#10270 Jason Dillaman)
- librbd: handle errors when creating iocxt while listing children (#10123 Jason Dillaman)
- mds: fix compat version in MClientSession (#9945 John Spray)
- mds: fix journaler write error handling (#10011 John Spray)
- mds: fix locking for file size recovery (#10229 Yan, Zheng)
- mds: handle heartbeat_reset during shutdown (#10382 John Spray)
- mds: store backtrace for straydir (Yan, Zheng)
- mon: allow tiers for FS pools (#10135 John Spray)
- mon: fix caching of last_epoch_clean, osdmap trimming (#9987 Sage Weil)
- mon: fix 'fs ls' on peons (#10288 John Spray)
- mon: fix MDS health status from peons (#10151 John Spray)
- mon: fix paxos off-by-one (#9301 Sage Weil)
- msgr: simple: do not block on takeover while holding global lock (#9921 Greg Farnum)
- osd: deep scrub must not abort if hinfo is missing (#10018 Loic Dachary)
- osd: fix misdirected op detection (#9835 Sage Weil)
- osd: fix past_interval display for acting (#9752 Loic Dachary)
- osd: fix PG peering backoff when behind on osdmaps (#10431 Sage Weil)
- osd: handle no-op write with snapshot case (#10262 Sage Weil)

- osd: use fast-dispatch (Sage Weil, Greg Farnum)
- rados: fix write to /dev/null (Loic Dachary)
- radosgw-admin: create subuser when needed (#10103 Yehuda Sadeh)
- rbd: avoid invalidating aio_write buffer during image import (#10590 Jason Dillaman)
- rbd: fix export with images > 2GB (Vicente Cheng)
- rgw: change multipart upload id magic (#10271 Georgios Dimitrakakis, Yehuda Sadeh)
- rgw: check keystone auth for S3 POST (#10062 Abhishek Lekshmanan)
- rgw: check timestamp for S3 keystone auth (#10062 Abhishek Lekshmanan)
- rgw: fix partial GET with swift (#10553 Yehuda Sadeh)
- rgw: fix quota disable (#9907 Dong Lei)
- rgw: fix rare corruption of object metadata on put (#9576 Yehuda Sadeh)
- rgw: fix S3 object copy content-type (#9478 Yehuda Sadeh)
- rgw: headers end with rn (#9254 Benedikt Fraunhofer, Yehuda Sadeh)
- rgw: remove swift user manifest DLO hash calculation (#9973 Yehuda Sadeh)
- rgw: return correct len when len is 0 (#9877 Yehuda Sadeh)
- rgw: return X-Timestamp field (#8911 Yehuda Sadeh)
- rgw: run radosgw as apache with systemd (#10125)
- rgw: sent ETag on S3 object copy (#9479 Yehuda Sadeh)
- rgw: sent HTTP status reason explicitly in fastcgi (Yehuda Sadeh)
- rgw: set length for keystone token validation (#7796 Mark Kirkwood, Yehuda Sadeh)
- rgw: set ulimit -n on sysvinit before starting daemon (#9587 Sage Weil)
- rgw: update bucket index on set_attrs (#5595 Yehuda Sadeh)
- rgw: update swift subuser permission masks when authenticating (#9918 Yehuda Sadeh)
- rgw: URL decode HTTP query params correction (#10271 Georgios Dimitrakakis, Yehuda Sadeh)
- rgw: use cached attrs while reading object attrs (#10307 Yehuda Sadeh)
- rgw: use strict_strtoll for content length (#10701 Axel Dunkel, Yehuda Sadeh)

For more detailed information, see [the complete changelog](#).

V0.87 GIANT

This release will form the basis for the stable release Giant, v0.87.x. Highlights for Giant include:

- **RADOS Performance:** a range of improvements have been made in the OSD and client-side librados code that improve the throughput on flash backends and improve parallelism and scaling on fast machines.
- **CephFS:** we have fixed a raft of bugs in CephFS and built some basic journal recovery and diagnostic tools. Stability and performance of single-MDS systems is vastly improved in Giant. Although we do not yet recommend CephFS for production deployments, we do encourage testing for non-critical workloads so that we can better gauge the feature, usability, performance, and stability gaps.
- **Local Recovery Codes:** the OSDs now support an erasure-coding scheme that stores some additional data blocks to reduce the IO required to recover from single OSD failures.
- **Degraded vs misplaced:** the Ceph health reports from ‘ceph -s’ and related commands now make a distinction between data that is degraded (there are fewer than the desired number of copies) and data that is misplaced (stored in the wrong location in the cluster). The distinction is important because the latter does not compromise data safety.
- **Tiering improvements:** we have made several improvements to the cache tiering implementation that improve performance. Most notably, objects are not promoted into the cache tier by a single read; they must be found to be sufficiently hot before that happens.
- **Monitor performance:** the monitors now perform writes to the local data store asynchronously, improving overall responsiveness.
- **Recovery tools:** the ceph_objectstore_tool is greatly expanded to allow manipulation of an individual OSDs data store for debugging and repair purposes. This is most heavily used by our QA infrastructure to exercise recovery code.

UPGRADE SEQUENCING

- If your existing cluster is running a version older than v0.80.x Firefly, please first upgrade to the latest Firefly release before moving on to Giant. We have not tested upgrades directly from Emperor, Dumpling, or older releases.

We have tested:

- Firefly to Giant
- Dumpling to Firefly to Giant

- Please upgrade daemons in the following order:

1. Monitors

- 2. OSDs
- 3. MDSs and/or radosgw

Note that the relative ordering of OSDs and monitors should not matter, but we primarily tested upgrading monitors first.

UPGRADING FROM V0.80X FIREFLY

- The client-side caching for librbd is now enabled by default (rbd cache = true). A safety option (rbd cache writethrough until flush = true) is also enabled so that writeback caching is not used until the library observes a ‘flush’ command, indicating that the librbd users is passing that operation through from the guest VM. This avoids potential data loss when used with older versions of qemu that do not support flush.

```
leveldb_write_buffer_size = 8*1024*1024 = 33554432 // 8MB leveldb_cache_size = 512*1024*1204 =
536870912 // 512MB leveldb_block_size = 64*1024 = 65536 // 64KB leveldb_compression = false leveldb_log
= ""
```

OSDs will still maintain the following osd-specific defaults:

```
leveldb_log = ""
```

- The ‘rados getxattr ...’ command used to add a gratuitous newline to the attr value; it now does not.
- The *_kb perf counters on the monitor have been removed. These are replaced with a new set of *_bytes counters (e.g., cluster_osd_kb is replaced by cluster_osd_bytes).
- The rd_kb and wr_kb fields in the JSON dumps for pool stats (accessed via the ceph df detail -f json-pretty and related commands) have been replaced with corresponding *_bytes fields. Similarly, the total_space, total_used, and total_avail fields are replaced with total_bytes, total_used_bytes, and total_avail_bytes fields.
- The rados df --format=json output read_bytes and write_bytes fields were incorrectly reporting ops; this is now fixed.
- The rados df --format=json output previously included read_kb and write_kb fields; these have been removed. Please use read_bytes and write_bytes instead (and divide by 1024 if appropriate).
- The experimental keyvaluestore-dev OSD backend had an on-disk format change that prevents existing OSD data from being upgraded. This affects developers and testers only.
- mon-specific and osd-specific leveldb options have been removed. From this point onward users should use the leveldb_* generic options and add the options in the appropriate sections of their configuration files. Monitors will still maintain the following monitor-specific defaults:

```
leveldb_write_buffer_size = 8*1024*1024 = 33554432 // 8MB leveldb_cache_size = 512*1024*1204 =
536870912 // 512MB leveldb_block_size = 64*1024 = 65536 // 64KB leveldb_compression = false leveldb_log
= ""
```

OSDs will still maintain the following osd-specific defaults:

```
leveldb_log = ""
```

- CephFS support for the legacy anchor table has finally been removed. Users with file systems created before firefly should ensure that inodes with multiple hard links are modified *prior* to the upgrade to ensure that the backtraces are written properly. For example:

```
sudo find /mnt/cephfs -type f -links +1 -exec touch \{\}\;
```

- We disallow nonsensical ‘tier cache-mode’ transitions. From this point onward, ‘writeback’ can only transition to ‘forward’ and ‘forward’ can transition to 1) ‘writeback’ if there are dirty objects, or 2) any if there are no dirty objects.

NOTABLE CHANGES SINCE V0.86

- ceph-disk: use new udev rules for centos7/rhel7 (#9747 Loic Dachary)
- libcephfs-java: fix fstat mode (Noah Watkins)
- librados: fix deadlock when listing PG contents (Guang Yang)
- librados: misc fixes to the new threading model (#9582 #9706 #9845 #9873 Sage Weil)
- mds: fix inotify initialization (Henry C Chang)
- mds: gracefully handle unknown lock type in flock requests (Yan, Zheng)

- mon: add read-only, read-write, and role-definer roles (Joao Eduardo Luis)
- mon: fix mon cap checks (Joao Eduardo Luis)
- mon: misc fixes for new paxos async writes (#9635 Sage Weil)
- mon: set scrub timestamps on PG creation (#9496 Joao Eduardo Luis)
- osd: erasure code: fix buffer alignment (Janne Grunau, Loic Dachary)
- osd: fix alloc hint induced crashes on mixed clusters (#9419 David Zafman)
- osd: fix backfill reservation release on rejection (#9626, Samuel Just)
- osd: fix ioprio option parsing (#9676 #9677 Loic Dachary)
- osd: fix memory leak during snap trimming (#9113 Samuel Just)
- osd: misc peering and recovery fixes (#9614 #9696 #9731 #9718 #9821 #9875 Samuel Just, Guang Yang)

NOTABLE CHANGES SINCE V0.80.X FIREFLY

- bash completion improvements (Wido den Hollander)
- brag: fixes, improvements (Loic Dachary)
- buffer: improve rebuild_page_aligned (Ma Jianpeng)
- build: fix build on alpha (Michael Cree, Dmitry Smirnov)
- build: fix CentOS 5 (Gerben Meijer)
- build: fix yasm check for x32 (Daniel Schepler, Sage Weil)
- ceph-brag: add tox tests (Alfredo Deza)
- ceph-conf: flush log on exit (Sage Weil)
- ceph.conf: update sample (Sebastien Han)
- ceph-dencoder: refactor build a bit to limit dependencies (Sage Weil, Dan Mick)
- ceph-disk: add Scientific Linux support (Dan van der Ster)
- ceph-disk: do not inadvertently create directories (Owne Synge)
- ceph-disk: fix dmcrypt support (Sage Weil)
- ceph-disk: fix dmcrypt support (Stephen Taylor)
- ceph-disk: handle corrupt volumes (Stuart Longlang)
- ceph-disk: linter cleanup, logging improvements (Alfredo Deza)
- ceph-disk: partprobe as needed (Eric Eastman)
- ceph-disk: show information about dmcrypt in ‘ceph-disk list’ output (Sage Weil)
- ceph-disk: use partition type UUIDs and blkid (Sage Weil)
- ceph: fix for non-default cluster names (#8944, Dan Mick)
- ceph-fuse, libcephfs: asok hooks for handling session resets, timeouts (Yan, Zheng)
- ceph-fuse, libcephfs: fix crash in trim_caps (John Spray)
- ceph-fuse, libcephfs: improve cap trimming (John Spray)
- ceph-fuse, libcephfs: improve traceless reply handling (Sage Weil)
- ceph-fuse, libcephfs: virtual xattrs for rstat (Yan, Zheng)
- ceph_objectstore_tool: vastly improved and extended tool for working offline with OSD data stores (David Zafman)
- ceph.spec: many fixes (Erik Logtenberg, Boris Ranto, Dan Mick, Sandon Van Ness)
- ceph.spec: split out ceph-common package, other fixes (Sandon Van Ness)
- ceph_test_librbd_fsx: fix RNG, make deterministic (Ilya Dryomov)
- cephtool: fix help (Yilong Zhao)
- cephtool: refactor and improve CLI tests (Joao Eduardo Luis)
- cephtool: test cleanup (Joao Eduardo Luis)
- clang build fixes (John Spray, Danny Al-Gaaf)
- client: improved MDS session dumps (John Spray)
- common: add config diff admin socket command (Joao Eduardo Luis)
- common: add rlock assertion checks (Yehuda Sadeh)
- common: fix dup log messages (#9080, Sage Weil)
- common: perfcounters now use atomics and go faster (Sage Weil)
- config: support G, M, K, etc. suffixes (Joao Eduardo Luis)
- coverity cleanups (Danny Al-Gaaf)
- crush: clean up CrushWrapper interface (Xiaoqi Chen)
- crush: include new tunables in dump (Sage Weil)
- crush: make ruleset ids unique (Xiaoqi Chen, Loic Dachary)
- crush: only require rule features if the rule is used (#8963, Sage Weil)
- crushtool: send output to stdout, not stderr (Wido den Hollander)
- doc: cache tiering (John Wilkins)
- doc: CRUSH updates (John Wilkins)
- doc: document new upstream wireshark dissector (Kevin Cox)
- doc: improve manual install docs (Francois Lafont)
- doc: keystone integration docs (John Wilkins)
- doc: librados example fixes (Kevin Dalley)

- doc: many doc updates (John Wilkins)
- doc: many install doc updates (John Wilkins)
- doc: misc updates (John Wilkins, Loic Dachary, David Moreau Simard, Wido den Hollander, Volker Voigt, Alfredo Deza, Stephen Jahl, Dan van der Ster)
- doc: osd primary affinity (John Wilkins)
- doc: pool quotas (John Wilkins)
- doc: pre-flight doc improvements (Kevin Dalley)
- doc: switch to an unencumbered font (Ross Turk)
- doc: updated simple configuration guides (John Wilkins)
- doc: update erasure docs (Loic Dachary, Venky Shankar)
- doc: update openstack docs (Josh Durgin)
- filestore: disable use of XFS hint (buggy on old kernels) (Samuel Just)
- filestore: fix xattr spillout (Greg Farnum, Haomai Wang)
- fix hppa arch build (Dmitry Smirnov)
- fix i386 builds (Sage Weil)
- fix struct vs class inconsistencies (Thorsten Behrens)
- global: write pid file even when running in foreground (Alexandre Oliva)
- hadoop: improve tests (Huamin Chen, Greg Farnum, John Spray)
- hadoop: update hadoop tests for Hadoop 2.0 (Haumin Chen)
- init-ceph: continue starting other daemons on crush or mount failure (#8343, Sage Weil)
- journaler: fix locking (Zheng, Yan)
- keyvaluestore: fix hint crash (#8381, Haomai Wang)
- keyvaluestore: header cache (Haomai Wang)
- libcephfs-java: build against older JNI headers (Greg Farnum)
- libcephfs-java: fix gcj-jdk build (Dmitry Smirnov)
- librados: fix crash on read op timeout (#9362 Matthias Kiefer, Sage Weil)
- librados: fix lock leaks in error paths (#9022, Paval Rallabhandi)
- librados: fix pool existence check (#8835, Pavan Rallabhandi)
- librados: fix rados_pool_list bounds checks (Sage Weil)
- librados: fix shutdown race (#9130 Sage Weil)
- librados: fix watch/notify test (#7934 David Zafman)
- librados: fix watch reregistration on acting set change (#9220 Samuel Just)
- librados: give Objecter fine-grained locks (Yehuda Sadeh, Sage Weil, John Spray)
- librados: lttng tracepoints (Adam Crume)
- librados, osd: return ETIMEDOUT on failed notify (Sage Weil)
- librados: pybind: fix reads when 0 is present (#9547 Mohammad Salehe)
- librados_striper: striping library for librados (Sebastien Ponce)
- librbd, ceph-fuse: reduce cache flush overhead (Haomai Wang)
- librbd: check error code on cache invalidate (Josh Durgin)
- librbd: enable caching by default (Sage Weil)
- librbd: enforce cache size on read requests (Jason Dillaman)
- librbd: fix crash using clone of flattened image (#8845, Josh Durgin)
- librbd: fix error path when opening image (#8912, Josh Durgin)
- librbd: handle blacklisting during shutdown (#9105 John Spray)
- librbd: lttng tracepoints (Adam Crume)
- librbd: new libkrbd library for kernel map/unmap/showmapped (Ilya Dryomov)
- librbd: store and retrieve snapshot metadata based on id (Josh Durgin)
- libs3: update to latest (Danny Al-Gaaf)
- log: fix derr level (Joao Eduardo Luis)
- logrotate: fix osd log rotation on ubuntu (Sage Weil)
- lttng: tracing infrastructure (Noah Watkins, Adam Crume)
- mailmap: many updates (Loic Dachary)
- mailmap: updates (Loic Dachary, Abhishek Lekshmanan, M Ranga Swami Reddy)
- Makefile: fix out of source builds (Stefan Eilemann)
- many many coverity fixes, cleanups (Danny Al-Gaaf)
- mds: adapt to new Objecter locking, give types to all Contexts (John Spray)
- mds: add file system name, enabled flag (John Spray)
- mds: add internal health checks (John Spray)
- mds: add min/max UID for snapshot creation/deletion (#9029, Wido den Hollander)
- mds: avoid tight mon reconnect loop (#9428 Sage Weil)
- mds: boot refactor, cleanup (John Spray)
- mds: cephfs-journal-tool (John Spray)
- mds: fix crash killing sessions (#9173 John Spray)
- mds: fix ctime updates (#9514 Greg Farnum)
- mds: fix journal conversion with standby-replay (John Spray)

- mds: fix replay locking (Yan, Zheng)
- mds: fix standby-replay cache trimming (#8648 Zheng, Yan)
- mds: fix xattr bug triggered by ACLs (Yan, Zheng)
- mds: give perfcounters meaningful names (Sage Weil)
- mds: improve health reporting to monitor (John Spray)
- mds: improve Journaler on-disk format (John Spray)
- mds: improve journal locking (Zheng, Yan)
- mds, libcephfs: use client timestamp for mtime/ctime (Sage Weil)
- mds: make max file recoveries tunable (Sage Weil)
- mds: misc encoding improvements (John Spray)
- mds: misc fixes for multi-mds (Yan, Zheng)
- mds: multi-mds fixes (Yan, Zheng)
- mds: OPTracker integration, dump_ops_in_flight (Greg Farnum)
- mds: prioritize file recovery when appropriate (Sage Weil)
- mds: refactor beacon, improve reliability (John Spray)
- mds: remove legacy anchor table (Yan, Zheng)
- mds: remove legacy discover ino (Yan, Zheng)
- mds: restart on EBLACKLISTED (John Spray)
- mds: separate inode recovery queue (John Spray)
- mds: session ls, evict commands (John Spray)
- mds: submit log events in async thread (Yan, Zheng)
- mds: track RECALL progress, report failure (#9284 John Spray)
- mds: update segment references during journal write (John Spray, Greg Farnum)
- mds: use client-provided timestamp for user-visible file metadata (Yan, Zheng)
- mds: use meaningful names for clients (John Spray)
- mds: validate journal header on load and save (John Spray)
- mds: warn clients which aren't revoking caps (Zheng, Yan, John Spray)
- misc build errors/warnings for Fedora 20 (Boris Ranto)
- misc build fixes for OS X (John Spray)
- misc cleanup (Christophe Courtaut)
- misc integer size cleanups (Kevin Cox)
- misc memory leaks, cleanups, fixes (Danny Al-Gaaf, Sahid Ferdjaoui)
- misc suse fixes (Danny Al-Gaaf)
- misc word size fixes (Kevin Cox)
- mon: add audit log for all admin commands (Joao Eduardo Luis)
- mon: add cluster fingerprint (Sage Weil)
- mon: add get-quota commands (Joao Eduardo Luis)
- mon: add 'osd blocked-by' command to easily see which OSDs are blocking peering progress (Sage Weil)
- mon: add 'osd reweight-by-pg' command (Sage Weil, Guang Yang)
- mon: add perfcounters for paxos operations (Sage Weil)
- mon: avoid creating unnecessary rule on pool create (#9304 Loic Dachary)
- monclient: fix hang (Sage Weil)
- mon: create default EC profile if needed (Loic Dachary)
- mon: do not create file system by default (John Spray)
- mon: do not spam log (Aanchal Agrawal, Sage Weil)
- mon: drop mon- and osd- specific leveldb options (Joao Eduardo Luis)
- mon: ec pool profile fixes (Loic Dachary)
- mon: fix bug when no auth keys are present (#8851, Joao Eduardo Luis)
- mon: fix 'ceph df' output for available space (Xiaoxi Chen)
- mon: fix compat version for MForward (Joao Eduardo Luis)
- mon: fix crash on loopback messages and paxos timeouts (#9062, Sage Weil)
- mon: fix default replication pool ruleset choice (#8373, John Spray)
- mon: fix divide by zero when pg_num is adjusted before OSDs are added (#9101, Sage Weil)
- mon: fix double-free of old MOSDBoot (Sage Weil)
- mon: fix health down messages (Sage Weil)
- mon: fix occasional memory leak after session reset (#9176, Sage Weil)
- mon: fix op write latency perfcounter (#9217 Xinxin Shu)
- mon: fix 'osd perf' reported latency (#9269 Samuel Just)
- mon: fix quorum feature check (#8738, Greg Farnum)
- mon: fix ruleset/ruleid bugs (#9044, Loic Dachary)
- mon: fix set cache_target_full_ratio (#8440, Geoffrey Hartz)
- mon: fix store check on startup (Joao Eduardo Luis)
- mon: include per-pool 'max avail' in df output (Sage Weil)
- mon: make paxos transaction commits asynchronous (Sage Weil)
- mon: make usage dumps in terms of bytes, not kB (Sage Weil)

- mon: ‘osd crush reweight-subtree ...’ (Sage Weil)
- mon, osd: relax client EC support requirements (Sage Weil)
- mon: preload erasure plugins (#9153 Loic Dachary)
- mon: prevent cache pools from being used directly by CephFS (#9435 John Spray)
- mon: prevent EC pools from being used with cephfs (Joao Eduardo Luis)
- mon: prevent implicit destruction of OSDs with ‘osd setmaxosd ...’ (#8865, Anand Bhat)
- mon: prevent nonsensical cache-mode transitions (Joao Eduardo Luis)
- mon: restore original weight when auto-marked out OSDs restart (Sage Weil)
- mon: restrict some pool properties to tiered pools (Joao Eduardo Luis)
- mon: some instrumentation (Sage Weil)
- mon: use msg header tid for MMonGetVersionReply (Ilya Dryomov)
- mon: use user-provided ruleset for replicated pool (Xiaoxi Chen)
- mon: verify all quorum members are contiguous at end of Paxos round (#9053, Sage Weil)
- mon: verify available disk space on startup (#9502 Joao Eduardo Luis)
- mon: verify erasure plugin version on load (Loic Dachary)
- msgr: avoid big lock when sending (most) messages (Greg Farnum)
- msgr: fix logged address (Yongyue Sun)
- msgr: misc locking fixes for fast dispatch (#8891, Sage Weil)
- msgr: refactor to cleanly separate SimpleMessenger implemenetation, move toward Connection-based calls (Matt Benjamin, Sage Weil)
- objecter: flag operations that are redirected by caching (Sage Weil)
- objectstore: clean up KeyValueDB interface for key/value backends (Sage Weil)
- osd: account for hit_set_archive bytes (Sage Weil)
- osd: add ability to prehash filestore directories (Guang Yang)
- osd: add ‘dump_reservations’ admin socket command (Sage Weil)
- osd: add feature bit for erasure plugins (Loic Dachary)
- osd: add header cache for KeyValueStore (Haomai Wang)
- osd: add ISA erasure plugin table cache (Andreas-Joachim Peters)
- osd: add local_mtime for use by cache agent (Zhiqiang Wang)
- osd: add local recovery code (LRC) erasure plugin (Loic Dachary)
- osd: add prototype KineticStore based on Seagate Kinetic (Josh Durgin)
- osd: add READFORWARD caching mode (Luis Pabon)
- osd: add superblock for KeyValueStore backend (Haomai Wang)
- osd: add support for Intel ISA-L erasure code library (Andreas-Joachim Peters)
- osd: allow map cache size to be adjusted at runtime (Sage Weil)
- osd: avoid refcounting overhead by passing a few things by ref (Somnath Roy)
- osd: avoid sharing PG info that is not durable (Samuel Just)
- osd: bound osdmap epoch skew between PGs (Sage Weil)
- osd: cache tier flushing fixes for snapped objects (Samuel Just)
- osd: cap hit_set size (#9339 Samuel Just)
- osd: clean up shard_id_t, shard_t (Loic Dachary)
- osd: clear FDCache on unlink (#8914 Loic Dachary)
- osd: clear slow request latency info on osd up/down (Sage Weil)
- osd: do not evict blocked objects (#9285 Zhiqiang Wang)
- osd: do not skip promote for write-ordered reads (#9064, Samuel Just)
- osd: fix agent early finish looping (David Zafman)
- osd: fix ambiguous encoding order for blacklisted clients (#9211, Sage Weil)
- osd: fix bogus assert during OSD shutdown (Sage Weil)
- osd: fix bug with long object names and rename (#8701, Sage Weil)
- osd: fix cache flush corner case for snapshotted objects (#9054, Samuel Just)
- osd: fix cache full -> not full requeueing (#8931, Sage Weil)
- osd: fix clone deletion case (#8334, Sam Just)
- osd: fix clone vs cache_evict bug (#8629 Sage Weil)
- osd: fix connection reconnect race (Greg Farnum)
- osd: fix crash from duplicate backfill reservation (#8863 Sage Weil)
- osd: fix dead peer connection checks (#9295 Greg Farnum, Sage Weil)
- osd: fix discard of old/obsolete subop replies (#9259, Samuel Just)
- osd: fix discard of peer messages from previous intervals (Greg Farnum)
- osd: fix dump of open fds on EMFILE (Sage Weil)
- osd: fix dumps (Joao Eduardo Luis)
- osd: fix erasure-code lib initialization (Loic Dachary)
- osd: fix extent normalization (Adam Crume)
- osd: fix filestore removal corner case (#8332, Sam Just)
- osd: fix flush vs OpContext (Samuel Just)
- osd: fix gating of messages from old OSD instances (Greg Farnum)

- osd: fix hang waiting for osdmap (#8338, Greg Farnum)
- osd: fix interval check corner case during peering (#8104, Sam Just)
- osd: fix ISA erasure alignment (Loic Dachary, Andreas-Joachim Peters)
- osd: fix journal dump (Ma Jianpeng)
- osd: fix journal-less operation (Sage Weil)
- osd: fix keyvaluestore scrub (#8589 Haomai Wang)
- osd: fix keyvaluestore upgrade (Haomai Wang)
- osd: fix loopback msgr issue (Ma Jianpeng)
- osd: fix LSB release parsing (Danny Al-Gaaf)
- osd: fix MarkMeDown and other shutdown races (Sage Weil)
- osd: fix memstore bugs with collection_move_rename, lock ordering (Sage Weil)
- osd: fix min_read_recency_for_promote default on upgrade (Zhiqiang Wang)
- osd: fix mon feature bit requirements bug and resulting log spam (Sage Weil)
- osd: fix mount/remount sync race (#9144 Sage Weil)
- osd: fix PG object listing/ordering bug (Guang Yang)
- osd: fix PG stat errors with tiering (#9082, Sage Weil)
- osd: fix purged_snap initialization on backfill (Sage Weil, Samuel Just, Dan van der Ster, Florian Haas)
- osd: fix race condition on object deletion (#9480 Somnath Roy)
- osd: fix recovery chunk size usage during EC recovery (Ma Jianpeng)
- osd: fix recovery reservation deadlock for EC pools (Samuel Just)
- osd: fix removal of old xattrs when overwriting chained xattrs (Ma Jianpeng)
- osd: fix requesting queueing on PG split (Samuel Just)
- osd: fix scrub vs cache bugs (Samuel Just)
- osd: fix snap object writeback from cache tier (#9054 Samuel Just)
- osd: fix trim of hitsets (Sage Weil)
- osd: force new xattrs into leveldb if fs returns E2BIG (#7779, Sage Weil)
- osd: implement alignment on chunk sizes (Loic Dachary)
- osd: improved backfill priorities (Sage Weil)
- osd: improve journal shutdown (Ma Jianpeng, Mark Kirkwood)
- osd: improve locking for KeyValueStore (Haomai Wang)
- osd: improve locking in OpTracker (Pavan Rallabhandi, Somnath Roy)
- osd: improve prioritization of recovery of degraded over misplaced objects (Sage Weil)
- osd: improve tiering agent arithmetic (Zhiqiang Wang, Sage Weil, Samuel Just)
- osd: include backend information in metadata reported to mon (Sage Weil)
- osd: locking, sharding, caching improvements in FileStore's FDCache (Somnath Roy, Greg Farnum)
- osd: lttng tracepoints for filestore (Noah Watkins)
- osd: make blacklist encoding deterministic (#9211 Sage Weil)
- osd: make tiering behave if hit_sets aren't enabled (Sage Weil)
- osd: many important bug fixes (Samuel Just)
- osd: many many core fixes (Samuel Just)
- osd: many many important fixes (#8231 #8315 #9113 #9179 #9293 #9294 #9326 #9453 #9481 #9482 #9497 #9574 Samuel Just)
- osd: mark pools with incomplete clones (Sage Weil)
- osd: misc erasure code plugin fixes (Loic Dachary)
- osd: misc locking fixes for fast dispatch (Samuel Just, Ma Jianpeng)
- osd, mon: add rocksdb support (Xinxin Shu, Sage Weil)
- osd, mon: config sanity checks on start (Sage Weil, Joao Eduardo Luis)
- osd, mon: distinguish between "misplaced" and "degraded" objects in cluster health and PG state reporting (Sage Weil)
- osd, msgr: fast-dispatch of OSD ops (Greg Farnum, Samuel Just)
- osd, objecter: resend ops on last_force_op_resend barrier; fix cache overlay op ordering (Sage Weil)
- osd: preload erasure plugins (#9153 Loic Dachary)
- osd: prevent old rados clients from using tiered pools (#8714, Sage Weil)
- osd: reduce OpTracker overhead (Somnath Roy)
- osd: refactor some ErasureCode functionality into command parent class (Loic Dachary)
- osd: remove obsolete classic scrub code (David Zafman)
- osd: scrub PGs with invalid stats (Sage Weil)
- osd: set configurable hard limits on object and xattr names (Sage Weil, Haomai Wang)
- osd: set rollback_info_completed on create (#8625, Samuel Just)
- osd: sharded threadpool to improve parallelism (Somnath Roy)
- osd: shard OpTracker to improve performance (Somnath Roy)
- osd: simple io prioritization for scrub (Sage Weil)
- osd: simple scrub throttling (Sage Weil)
- osd: simple snap trimmer throttle (Sage Weil)
- osd: tests for bench command (Loic Dachary)
- osd: trim old EC objects quickly; verify on scrub (Samuel Just)

- osd: use FIEMAP to inform copy_range (Haomai Wang)
- osd: use local time for tiering decisions (Zhiqiang Wang)
- osd: use xfs hint less frequently (Ilya Dryomov)
- osd: verify erasure plugin version on load (Loic Dachary)
- osd: work around GCC 4.8 bug in journal code (Matt Benjamin)
- pybind/rados: fix small timeouts (John Spray)
- qa: xfstests updates (Ilya Dryomov)
- rados: allow setxattr value to be read from stdin (Sage Weil)
- rados bench: fix arg order (Kevin Dalley)
- rados: drop gratuitous n from getxattr command (Sage Weil)
- rados: fix bench write arithmetic (Jiangheng)
- rados: fix {read,write}_ops values for df output (Sage Weil)
- rbd: add rbdmap pre- and post post- hooks, fix misc bugs (Dmitry Smirnov)
- rbd-fuse: allow exposing single image (Stephen Taylor)
- rbd-fuse: fix unlink (Josh Durgin)
- rbd: improve option default behavior (Josh Durgin)
- rbd: parallelize rbd import, export (Jason Dillaman)
- rbd: rbd-replay utility to replay captured rbd workload traces (Adam Crume)
- rbd: use write-back (not write-through) when caching is enabled (Jason Dillaman)
- removed mkcephfs (deprecated since dumpling)
- rest-api: fix help (Ailing Zhang)
- rgw: add civetweb as default frontend on port 7490 (#9013 Yehuda Sadeh)
- rgw: add -min-rewrite-stripe-size for object restriper (Yehuda Sadeh)
- rgw: add powerdns hook for dynamic DNS for global clusters (Wido den Hollander)
- rgw: add S3 bucket get location operation (Abhishek Lekshmanan)
- rgw: allow : in S3 access key (Roman Haritonov)
- rgw: automatically align writes to EC pool (#8442, Yehuda Sadeh)
- rgw: bucket link uses instance id (Yehuda Sadeh)
- rgw: cache bucket info (Yehuda Sadeh)
- rgw: cache decoded user info (Yehuda Sadeh)
- rgw: check entity permission for put_metadata (#8428, Yehuda Sadeh)
- rgw: copy object data is target bucket is in a different pool (#9039, Yehuda Sadeh)
- rgw: do not try to authenticate CORS preflight requests (#8718, Robert Hubbard, Yehuda Sadeh)
- rgw: fix admin create user op (#8583 Ray Lv)
- rgw: fix civetweb URL decoding (#8621, Yehuda Sadeh)
- rgw: fix crash on swift CORS preflight request (#8586, Yehuda Sadeh)
- rgw: fix log filename suffix (#9353 Alexandre Marangone)
- rgw: fix memory leak following chunk read error (Yehuda Sadeh)
- rgw: fix memory leaks (Andrey Kuznetsov)
- rgw: fix multipart object attr regression (#8452, Yehuda Sadeh)
- rgw: fix multipart upload (#8846, Silvain Munaut, Yehuda Sadeh)
- rgw: fix radosgw-admin 'show log' command (#8553, Yehuda Sadeh)
- rgw: fix removal of objects during object creation (Patrycja Szablowska, Yehuda Sadeh)
- rgw: fix striping for copied objects (#9089, Yehuda Sadeh)
- rgw: fix test for identify whether an object has a tail (#9226, Yehuda Sadeh)
- rgw: fix URL decoding (#8702, Brian Rak)
- rgw: fix URL escaping (Yehuda Sadeh)
- rgw: fix usage (Abhishek Lekshmanan)
- rgw: fix user manifest (Yehuda Sadeh)
- rgw: fix when stripe size is not a multiple of chunk size (#8937, Yehuda Sadeh)
- rgw: handle empty extra pool name (Yehuda Sadeh)
- rgw: improve civetweb logging (Yehuda Sadeh)
- rgw: improve delimited listing of bucket, misc fixes (Yehuda Sadeh)
- rgw: improve -h (Abhishek Lekshmanan)
- rgw: many fixes for civetweb (Yehuda Sadeh)
- rgw: misc civetweb fixes (Yehuda Sadeh)
- rgw: misc civetweb frontend fixes (Yehuda Sadeh)
- rgw: object and bucket rewrite functions to allow restriping old objects (Yehuda Sadeh)
- rgw: powerdns backend for global namespaces (Wido den Hollander)
- rgw: prevent multiobject PUT race (Yehuda Sadeh)
- rgw: send user manifest header (Yehuda Sadeh)
- rgw: subuser creation fixes (#8587 Yehuda Sadeh)
- rgw: use systemd-run from sysvinit script (Juanjose Galvez)
- rpm: do not restart daemons on upgrade (Alfredo Deza)
- rpm: misc packaging fixes for rhel7 (Sandon Van Ness)

- rpm: split ceph-common from ceph (Sandon Van Ness)
- systemd: initial systemd config files (Federico Simoncelli)
- systemd: wrap started daemons in new systemd environment (Sage Weil, Dan Mick)
- sysvinit: add support for non-default cluster names (Alfredo Deza)
- sysvinit: less sensitive to failures (Sage Weil)
- test_librbd_fsx: test kbd as well as librbd (Ilya Dryomov)
- unit test improvements (Loic Dachary)
- upstart: increase max open files limit (Sage Weil)
- vstart.sh: fix/improve rgw support (Luis Pabon, Abhishek Lekshmanan)

V0.86

This is a release candidate for Giant, which will hopefully be out in another week or two. We did a feature freeze about a month ago and since then have been doing only stabilization and bug fixing (and a handful of low-risk enhancements). A fair bit of new functionality went into the final sprint, but it's baked for quite a while now and we're feeling pretty good about it.

Major items include:

- librados locking refactor to improve scaling and client performance
- local recovery code (LRC) erasure code plugin to trade some additional storage overhead for improved recovery performance
- LTTNG tracing framework, with initial tracepoints in librados, librbd, and the OSD FileStore backend
- separate monitor audit log for all administrative commands
- asynchronous monitor transaction commits to reduce the impact on monitor read requests while processing updates
- low-level tool for working with individual OSD data stores for debugging, recovery, and testing
- many MDS improvements (bug fixes, health reporting)

There are still a handful of known bugs in this release, but nothing severe enough to prevent a release. By and large we are pretty pleased with the stability and expect the final Giant release to be quite reliable.

Please try this out on your non-production clusters for a preview

NOTABLE CHANGES

- buffer: improve rebuild_page_aligned (Ma Jianpeng)
- build: fix CentOS 5 (Gerben Meijer)
- build: fix build on alpha (Michael Cree, Dmitry Smirnov)
- build: fix yasm check for x32 (Daniel Schepler, Sage Weil)
- ceph-disk: add Scientific Linux support (Dan van der Ster)
- ceph-fuse, libcephfs: fix crash in trim_caps (John Spray)
- ceph-fuse, libcephfs: improve cap trimming (John Spray)
- ceph-fuse, libcephfs: virtual xattrs for rstat (Yan, Zheng)
- ceph.conf: update sample (Sebastien Han)
- ceph.spec: many fixes (Erik Logtenberg, Boris Ranto, Dan Mick, Sandon Van Ness)
- ceph_objectstore_tool: vastly improved and extended tool for working offline with OSD data stores (David Zafman)
- common: add config diff admin socket command (Joao Eduardo Luis)
- common: add rwlock assertion checks (Yehuda Sadeh)
- crush: clean up CrushWrapper interface (Xiaoxi Chen)
- crush: make ruleset ids unique (Xiaoxi Chen, Loic Dachary)
- doc: improve manual install docs (Francois Lafont)
- doc: misc updates (John Wilkins, Loic Dachary, David Moreau Simard, Wido den Hollander, Volker Voigt, Alfredo Deza, Stephen Jahl, Dan van der Ster)
- global: write pid file even when running in foreground (Alexandre Oliva)
- hadoop: improve tests (Huamin Chen, Greg Farnum, John Spray)
- journaler: fix locking (Zheng, Yan)
- librados, osd: return ETIMEDOUT on failed notify (Sage Weil)
- librados: fix crash on read op timeout (#9362 Matthias Kiefer, Sage Weil)
- librados: fix shutdown race (#9130 Sage Weil)
- librados: fix watch reregistration on acting set change (#9220 Samuel Just)
- librados: fix watch/notify test (#7934 David Zafman)
- librados: give Objecter fine-grained locks (Yehuda Sadeh, Sage Weil, John Spray)
- librados: lttng tracepoitns (Adam Crume)
- librados: pybind: fix reads when 0 is present (#9547 Mohammad Salehe)
- librbd: enforce cache size on read requests (Jason Dillaman)

- librbd: handle blacklisting during shutdown (#9105 John Spray)
- librbd: Ittng tracepoints (Adam Crume)
- Ittng: tracing infrastructure (Noah Watkins, Adam Crume)
- mailmap: updates (Loic Dachary, Abhishek Lekshmanan, M Ranga Swami Reddy)
- many many coverity fixes, cleanups (Danny Al-Gaaf)
- mds: adapt to new Objecter locking, give types to all Contexts (John Spray)
- mds: add internal health checks (John Spray)
- mds: avoid tight mon reconnect loop (#9428 Sage Weil)
- mds: fix crash killing sessions (#9173 John Spray)
- mds: fix ctime updates (#9514 Greg Farnum)
- mds: fix replay locking (Yan, Zheng)
- mds: fix standby-replay cache trimming (#8648 Zheng, Yan)
- mds: give perfcounters meaningful names (Sage Weil)
- mds: improve health reporting to monitor (John Spray)
- mds: improve journal locking (Zheng, Yan)
- mds: make max file recoveries tunable (Sage Weil)
- mds: prioritize file recovery when appropriate (Sage Weil)
- mds: refactor beacon, improve reliability (John Spray)
- mds: restart on EBLACKLISTED (John Spray)
- mds: track RECALL progress, report failure (#9284 John Spray)
- mds: update segment references during journal write (John Spray, Greg Farnum)
- mds: use meaningful names for clients (John Spray)
- mds: warn clients which aren't revoking caps (Zheng, Yan, John Spray)
- mon: add 'osd reweight-by-pg' command (Sage Weil, Guang Yang)
- mon: add audit log for all admin commands (Joao Eduardo Luis)
- mon: add cluster fingerprint (Sage Weil)
- mon: avoid creating unnecessary rule on pool create (#9304 Loic Dachary)
- mon: do not spam log (Aanchal Agrawal, Sage Weil)
- mon: fix 'osd perf' reported latency (#9269 Samuel Just)
- mon: fix double-free of old MOSDBoot (Sage Weil)
- mon: fix op write latency perfcounter (#9217 Xinxin Shu)
- mon: fix store check on startup (Joao Eduardo Luis)
- mon: make paxos transaction commits asynchronous (Sage Weil)
- mon: preload erasure plugins (#9153 Loic Dachary)
- mon: prevent cache pools from being used directly by CephFS (#9435 John Spray)
- mon: use user-provided ruleset for replicated pool (Xiaoxi Chen)
- mon: verify available disk space on startup (#9502 Joao Eduardo Luis)
- mon: verify erasure plugin version on load (Loic Dachary)
- msgr: fix logged address (Yongyue Sun)
- osd: account for hit_set_archive bytes (Sage Weil)
- osd: add ISA erasure plugin table cache (Andreas-Joachim Peters)
- osd: add ability to prehash filestore directories (Guang Yang)
- osd: add feature bit for erasure plugins (Loic Dachary)
- osd: add local recovery code (LRC) erasure plugin (Loic Dachary)
- osd: cap hit_set size (#9339 Samuel Just)
- osd: clear FDCache on unlink (#8914 Loic Dachary)
- osd: do not evict blocked objects (#9285 Zhiqiang Wang)
- osd: fix ISA erasure alignment (Loic Dachary, Andreas-Joachim Peters)
- osd: fix clone vs cache_evict bug (#8629 Sage Weil)
- osd: fix crash from duplicate backfill reservation (#8863 Sage Weil)
- osd: fix dead peer connection checks (#9295 Greg Farnum, Sage Weil)
- osd: fix keyvaluestore scrub (#8589 Haomai Wang)
- osd: fix keyvaluestore upgrade (Haomai Wang)
- osd: fix min_read_recency_for_promote default on upgrade (Zhiqiang Wang)
- osd: fix mount/remount sync race (#9144 Sage Weil)
- osd: fix purged_snap initialization on backfill (Sage Weil, Samuel Just, Dan van der Ster, Florian Haas)
- osd: fix race condition on object deletion (#9480 Somnath Roy)
- osd: fix snap object writeback from cache tier (#9054 Samuel Just)
- osd: improve journal shutdown (Ma Jianpeng, Mark Kirkwood)
- osd: improve locking in OpTracker (Pavan Rallabhandi, Somnath Roy)
- osd: improve tiering agent arithmetic (Zhiqiang Wang, Sage Weil, Samuel Just)
- osd: Ittng tracepoints for filestore (Noah Watkins)
- osd: make blacklist encoding deterministic (#9211 Sage Weil)
- osd: many many important fixes (#8231 #8315 #9113 #9179 #9293 #9294 #9326 #9453 #9481 #9482 #9497 #9574 Samuel Just)

- osd: misc erasure code plugin fixes (Loic Dachary)
- osd: preload erasure plugins (#9153 Loic Dachary)
- osd: shard OpTracker to improve performance (Somnath Roy)
- osd: use local time for tiering decisions (Zhiqiang Wang)
- osd: verify erasure plugin version on load (Loic Dachary)
- rados: fix bench write arithmetic (Jiangheng)
- rbd: parallelize rbd import, export (Jason Dillaman)
- rbd: rbd-replay utility to replay captured rbd workload traces (Adam Crume)
- rbd: use write-back (not write-through) when caching is enabled (Jason Dillaman)
- rgw: add S3 bucket get location operation (Abhishek Lekshmanan)
- rgw: add civetweb as default frontend on port 7490 (#9013 Yehuda Sadeh)
- rgw: allow : in S3 access key (Roman Haritonov)
- rgw: fix admin create user op (#8583 Ray Lv)
- rgw: fix log filename suffix (#9353 Alexandre Marangone)
- rgw: fix usage (Abhishek Lekshmanan)
- rgw: many fixes for civetweb (Yehuda Sadeh)
- rgw: subuser creation fixes (#8587 Yehuda Sadeh)
- rgw: use systemd-run from sysvinit script (JuanJose Galvez)
- unit test improvements (Loic Dachary)
- vstart.sh: fix/improve rgw support (Luis Pabon, Abhishek Lekshmanan)

V0.85

This is the second-to-last development release before Giant that contains new functionality. The big items to land during this cycle are the messenger refactoring from Matt Benjmain that lays some groundwork for RDMA support, a performance improvement series from SanDisk that improves performance on SSDs, lots of improvements to our new standalone civetweb-based RGW frontend, and a new ‘osd blocked-by’ mon command that allows admins to easily identify which OSDs are blocking peering progress. The other big change is that the OSDs and Monitors now distinguish between “misplaced” and “degraded” objects: the latter means there are fewer copies than we’d like, while the former simply means the are not stored in the locations where we want them to be.

Also of note is a change to librbd that enables client-side caching by default. This is coupled with another option that makes the cache write-through until a “flush” operations is observed: this implies that the librbd user (usually a VM guest OS) supports barriers and flush and that it is safe for the cache to switch into writeback mode without compromising data safety or integrity. It has long been recommended practice that these options be enabled (e.g., in OpenStack environments) but until now it has not been the default.

We have frozen the tree for the looming Giant release, and the next development release will be a release candidate with a final batch of new functionality.

UPGRADING

- The client-side caching for librbd is now enabled by default (rbd cache = true). A safety option (rbd cache writethrough until flush = true) is also enabled so that writeback caching is not used until the library observes a ‘flush’ command, indicating that the librbd users is passing that operation through from the guest VM. This avoids potential data loss when used with older versions of qemu that do not support flush.

```
leveldb_write_buffer_size = 32*1024*1024 = 33554432 // 32MB leveldb_cache_size = 512*1024*1204 =
536870912 // 512MB leveldb_block_size = 64*1024 = 65536 // 64KB leveldb_compression = false leveldb_log
= ""
```

OSDs will still maintain the following osd-specific defaults:

```
leveldb_log = ""
```

- The ‘rados getxattr ...’ command used to add a gratuitous newline to the attr value; it now does not.

NOTABLE CHANGES

- ceph-disk: do not inadvertently create directories (Owne Syngue)
- ceph-disk: fix dmcrypt support (Sage Weil)
- ceph-disk: linter cleanup, logging improvements (Alfredo Deza)
- ceph-disk: show information about dmcrypt in ‘ceph-disk list’ output (Sage Weil)
- ceph-disk: use partition type UUIDs and blkid (Sage Weil)

- ceph: fix for non-default cluster names (#8944, Dan Mick)
- doc: document new upstream wireshark dissector (Kevin Cox)
- doc: many install doc updates (John Wilkins)
- librados: fix lock leaks in error paths (#9022, Paval Rallabhandi)
- librados: fix pool existence check (#8835, Pavan Rallabhandi)
- librbd: enable caching by default (Sage Weil)
- librbd: fix crash using clone of flattened image (#8845, Josh Durgin)
- librbd: store and retrieve snapshot metadata based on id (Josh Durgin)
- mailmap: many updates (Loic Dachary)
- mds: add min/max UID for snapshot creation/deletion (#9029, Wido den Hollander)
- misc build errors/warnings for Fedora 20 (Boris Ranto)
- mon: add ‘osd blocked-by’ command to easily see which OSDs are blocking peering progress (Sage Weil)
- mon: add perfcounters for paxos operations (Sage Weil)
- mon: create default EC profile if needed (Loic Dachary)
- mon: fix crash on loopback messages and paxos timeouts (#9062, Sage Weil)
- mon: fix divide by zero when pg_num is adjusted before OSDs are added (#9101, Sage Weil)
- mon: fix occasional memory leak after session reset (#9176, Sage Weil)
- mon: fix ruleset/ruleid bugs (#9044, Loic Dachary)
- mon: make usage dumps in terms of bytes, not kB (Sage Weil)
- mon: prevent implicit destruction of OSDs with ‘osd setmaxosd ...’ (#8865, Anand Bhat)
- mon: verify all quorum members are contiguous at end of Paxos round (#9053, Sage Weil)
- msgr: refactor to cleanly separate SimpleMessenger implemenetation, move toward Connection-based calls (Matt Benjamin, Sage Weil)
- objectstore: clean up KeyValueDB interface for key/value backends (Sage Weil)
- osd: add local_mtime for use by cache agent (Zhiqiang Wang)
- osd: add superblock for KeyValueStore backend (Haomai Wang)
- osd: add support for Intel ISA-L erasure code library (Andreas-Joachim Peters)
- osd: do not skip promote for write-ordered reads (#9064, Samuel Just)
- osd: fix ambiguous encoding order for blacklisted clients (#9211, Sage Weil)
- osd: fix cache flush corner case for snapshotted objects (#9054, Samuel Just)
- osd: fix discard of old/obsolete subop replies (#9259, Samuel Just)
- osd: fix discard of peer messages from previous intervals (Greg Farnum)
- osd: fix dump of open fds on EMFILE (Sage Weil)
- osd: fix journal dump (Ma Jianpeng)
- osd: fix mon feature bit requirements bug and resulting log spam (Sage Weil)
- osd: fix recovery chunk size usage during EC recovery (Ma Jianpeng)
- osd: fix recovery reservation deadlock for EC pools (Samuel Just)
- osd: fix removal of old xattrs when overwriting chained xattrs (Ma Jianpeng)
- osd: fix requesting queueing on PG split (Samuel Just)
- osd: force new xattrs into leveldb if fs returns E2BIG (#7779, Sage Weil)
- osd: implement alignment on chunk sizes (Loic Dachary)
- osd: improve prioritization of recovery of degraded over misplaced objects (Sage Weil)
- osd: locking, sharding, caching improvements in FileStore’s FDCache (Somnath Roy, Greg Farnum)
- osd: many important bug fixes (Samuel Just)
- osd, mon: add rocksdb support (Xinxin Shu, Sage Weil)
- osd, mon: distinguish between “misplaced” and “degraded” objects in cluster health and PG state reporting (Sage Weil)
- osd: refactor some ErasureCode functionality into command parent class (Loic Dachary)
- osd: set rollback_info_completed on create (#8625, Samuel Just)
- rados: allow setxattr value to be read from stdin (Sage Weil)
- rados: drop gratuitous n from getxattr command (Sage Weil)
- rgw: add -min-rewrite-stripe-size for object restriper (Yehuda Sadeh)
- rgw: add powerdns hook for dynamic DNS for global clusters (Wido den Hollander)
- rgw: copy object data is target bucket is in a different pool (#9039, Yehuda Sadeh)
- rgw: do not try to authenticate CORS preflight requests (#8718, Robert Hubbard, Yehuda Sadeh)
- rgw: fix civetweb URL decoding (#8621, Yehuda Sadeh)
- rgw: fix removal of objects during object creation (Patrycja Szablowska, Yehuda Sadeh)
- rgw: fix striping for copied objects (#9089, Yehuda Sadeh)
- rgw: fix test for identify whether an object has a tail (#9226, Yehuda Sadeh)
- rgw: fix when stripe size is not a multiple of chunk size (#8937, Yehuda Sadeh)
- rgw: improve civetweb logging (Yehuda Sadeh)
- rgw: misc civetweb frontend fixes (Yehuda Sadeh)
- sysvinit: add support for non-default cluster names (Alfredo Deza)

The next Ceph development release is here! This release contains several meaty items, including some MDS improvements for journaling, the ability to remove the CephFS file system (and name it), several mon cleanups with tiered pools, several OSD performance branches, a new “read forward” RADOS caching mode, a prototype Kinetic OSD backend, and various radosgw improvements (especially with the new standalone civetweb frontend). And there are a zillion OSD bug fixes. Things are looking pretty good for the Giant release that is coming up in the next month.

UPGRADING

- The *_kb perf counters on the monitor have been removed. These are replaced with a new set of *_bytes counters (e.g., cluster_osd_kb is replaced by cluster_osd_bytes).
- The rd_kb and wr_kb fields in the JSON dumps for pool stats (accessed via the ceph df detail -f json-pretty and related commands) have been replaced with corresponding *_bytes fields. Similarly, the total_space, total_used, and total_avail fields are replaced with total_bytes, total_used_bytes, and total_avail_bytes fields.
- The rados df --format=json output read_kb and write_kb fields were incorrectly reporting ops; this is now fixed.
- The rados df --format=json output previously included read_kb and write_kb fields; these have been removed. Please use read_bytes and write_bytes instead (and divide by 1024 if appropriate).

NOTABLE CHANGES

- ceph-conf: flush log on exit (Sage Weil)
- ceph-dencoder: refactor build a bit to limit dependencies (Sage Weil, Dan Mick)
- ceph.spec: split out ceph-common package, other fixes (Sandon Van Ness)
- ceph_test_librbd_fsx: fix RNG, make deterministic (Ilya Dryomov)
- cephtool: refactor and improve CLI tests (Joao Eduardo Luis)
- client: improved MDS session dumps (John Spray)
- common: fix dup log messages (#9080, Sage Weil)
- crush: include new tunables in dump (Sage Weil)
- crush: only require rule features if the rule is used (#8963, Sage Weil)
- crushtool: send output to stdout, not stderr (Wido den Hollander)
- fix i386 builds (Sage Weil)
- fix struct vs class inconsistencies (Thorsten Behrens)
- hadoop: update hadoop tests for Hadoop 2.0 (Haumin Chen)
- librbd, ceph-fuse: reduce cache flush overhead (Haomai Wang)
- librbd: fix error path when opening image (#8912, Josh Durgin)
- mds: add file system name, enabled flag (John Spray)
- mds: boot refactor, cleanup (John Spray)
- mds: fix journal conversion with standby-replay (John Spray)
- mds: separate inode recovery queue (John Spray)
- mds: session ls, evict commands (John Spray)
- mds: submit log events in async thread (Yan, Zheng)
- mds: use client-provided timestamp for user-visible file metadata (Yan, Zheng)
- mds: validate journal header on load and save (John Spray)
- misc build fixes for OS X (John Spray)
- misc integer size cleanups (Kevin Cox)
- mon: add get-quota commands (Joao Eduardo Luis)
- mon: do not create file system by default (John Spray)
- mon: fix ‘ceph df’ output for available space (Xiaoxi Chen)
- mon: fix bug when no auth keys are present (#8851, Joao Eduardo Luis)
- mon: fix compat version for MForward (Joao Eduardo Luis)
- mon: restrict some pool properties to tiered pools (Joao Eduardo Luis)
- msgr: misc locking fixes for fast dispatch (#8891, Sage Weil)
- osd: add ‘dump_reservations’ admin socket command (Sage Weil)
- osd: add READFORWARD caching mode (Luis Pabon)
- osd: add header cache for KeyValueStore (Haomai Wang)
- osd: add prototype KineticStore based on Seagate Kinetic (Josh Durgin)
- osd: allow map cache size to be adjusted at runtime (Sage Weil)
- osd: avoid refcounting overhead by passing a few things by ref (Somnath Roy)
- osd: avoid sharing PG info that is not durable (Samuel Just)
- osd: clear slow request latency info on osd up/down (Sage Weil)
- osd: fix PG object listing/ordering bug (Guang Yang)
- osd: fix PG stat errors with tiering (#9082, Sage Weil)

- osd: fix bug with long object names and rename (#8701, Sage Weil)
- osd: fix cache full -> not full requeueing (#8931, Sage Weil)
- osd: fix gating of messages from old OSD instances (Greg Farnum)
- osd: fix memstore bugs with collection_move_rename, lock ordering (Sage Weil)
- osd: improve locking for KeyValueStore (Haomai Wang)
- osd: make tiering behave if hit_sets aren't enabled (Sage Weil)
- osd: mark pools with incomplete clones (Sage Weil)
- osd: misc locking fixes for fast dispatch (Samuel Just, Ma Jianpeng)
- osd: prevent old rados clients from using tiered pools (#8714, Sage Weil)
- osd: reduce OpTracker overhead (Somnath Roy)
- osd: set configurable hard limits on object and xattr names (Sage Weil, Haomai Wang)
- osd: trim old EC objects quickly; verify on scrub (Samuel Just)
- osd: work around GCC 4.8 bug in journal code (Matt Benjamin)
- rados bench: fix arg order (Kevin Dalley)
- rados: fix {read,write}_ops values for df output (Sage Weil)
- rbd: add rbdmap pre- and post post- hooks, fix misc bugs (Dmitry Smirnov)
- rbd: improve option default behavior (Josh Durgin)
- rgw: automatically align writes to EC pool (#8442, Yehuda Sadeh)
- rgw: fix crash on swift CORS preflight request (#8586, Yehuda Sadeh)
- rgw: fix memory leaks (Andrey Kuznetsov)
- rgw: fix multipart upload (#8846, Silvain Munaut, Yehuda Sadeh)
- rgw: improve -h (Abhishek Lekshmanan)
- rgw: improve delimited listing of bucket, misc fixes (Yehuda Sadeh)
- rgw: misc civetweb fixes (Yehuda Sadeh)
- rgw: powerdns backend for global namespaces (Wido den Hollander)
- systemd: initial systemd config files (Federico Simoncelli)

V0.83

Another Ceph development release! This has been a longer cycle, so there has been quite a bit of bug fixing and stabilization in this round. There is also a bunch of packaging fixes for RPM distros (RHEL/CentOS, Fedora, and SUSE) and for systemd. We've also added a new librados-striper library from Sébastien Ponce that provides a generic striping API for applications to code to.

UPGRADING

- The experimental keyvaluestore-dev OSD backend had an on-disk format change that prevents existing OSD data from being upgraded. This affects developers and testers only.
- mon-specific and osd-specific leveldb options have been removed. From this point onward users should use the *leveldb_** generic options and add the options in the appropriate sections of their configuration files. Monitors will still maintain the following monitor-specific defaults:

```
leveldb_write_buffer_size = 32*1024*1024 = 33554432 // 32MB leveldb_cache_size = 512*1024*1204 =
536870912 // 512MB leveldb_block_size = 64*1024 = 65536 // 64KB leveldb_compression = false leveldb_log
= ""
```

OSDs will still maintain the following osd-specific defaults:

```
leveldb_log = ""
```

NOTABLE CHANGES

- ceph-disk: fix dmcrypt support (Stephen Taylor)
- cephtool: fix help (Yilong Zhao)
- cephtool: test cleanup (Joao Eduardo Luis)
- doc: librados example fixes (Kevin Dalley)
- doc: many doc updates (John Wilkins)
- doc: update erasure docs (Loic Dachary, Venky Shankar)
- filestore: disable use of XFS hint (buggy on old kernels) (Samuel Just)
- filestore: fix xattr spillout (Greg Farnum, Haomai Wang)
- keyvaluestore: header cache (Haomai Wang)
- librados_striper: striping library for librados (Sébastien Ponce)

- libs3: update to latest (Danny Al-Gaaf)
- log: fix derr level (Joao Eduardo Luis)
- logrotate: fix osd log rotation on ubuntu (Sage Weil)
- mds: fix xattr bug triggered by ACLs (Yan, Zheng)
- misc memory leaks, cleanups, fixes (Danny Al-Gaaf, Sahid Ferdjaoui)
- misc suse fixes (Danny Al-Gaaf)
- misc word size fixes (Kevin Cox)
- mon: drop mon- and osd- specific leveldb options (Joao Eduardo Luis)
- mon: ec pool profile fixes (Loic Dachary)
- mon: fix health down messages (Sage Weil)
- mon: fix quorum feature check (#8738, Greg Farnum)
- mon: ‘osd crush reweight-subtree ...’ (Sage Weil)
- mon, osd: relax client EC support requirements (Sage Weil)
- mon: some instrumentation (Sage Weil)
- objecter: flag operations that are redirected by caching (Sage Weil)
- osd: clean up shard_id_t, shard_t (Loic Dachary)
- osd: fix connection reconnect race (Greg Farnum)
- osd: fix dumps (Joao Eduardo Luis)
- osd: fix erasure-code lib initialization (Loic Dachary)
- osd: fix extent normalization (Adam Crume)
- osd: fix loopback msgr issue (Ma Jianpeng)
- osd: fix LSB release parsing (Danny Al-Gaaf)
- osd: improved backfill priorities (Sage Weil)
- osd: many many core fixes (Samuel Just)
- osd, mon: config sanity checks on start (Sage Weil, Joao Eduardo Luis)
- osd: sharded threadpool to improve parallelism (Somnath Roy)
- osd: simple io prioritization for scrub (Sage Weil)
- osd: simple scrub throttling (Sage Weil)
- osd: tests for bench command (Loic Dachary)
- osd: use xfs hint less frequently (Ilya Dryomov)
- pybind/rados: fix small timeouts (John Spray)
- qa: xfstests updates (Ilya Dryomov)
- rgw: cache bucket info (Yehuda Sadeh)
- rgw: cache decoded user info (Yehuda Sadeh)
- rgw: fix multipart object attr regression (#8452, Yehuda Sadeh)
- rgw: fix radosgw-admin ‘show log’ command (#8553, Yehuda Sadeh)
- rgw: fix URL decoding (#8702, Brian Rak)
- rgw: handle empty extra pool name (Yehuda Sadeh)
- rpm: do not restart daemons on upgrade (Alfredo Deza)
- rpm: misc packaging fixes for rhel7 (Sandon Van Ness)
- rpm: split ceph-common from ceph (Sandon Van Ness)
- systemd: wrap started daemons in new systemd environment (Sage Weil, Dan Mick)
- sysvinit: less sensitive to failures (Sage Weil)
- upstart: increase max open files limit (Sage Weil)

V0.82

This is the second post-firefly development release. It includes a range of bug fixes and some usability improvements. There are some MDS debugging and diagnostic tools, an improved ‘ceph df’, and some OSD backend refactoring and cleanup.

NOTABLE CHANGES

- ceph-brag: add tox tests (Alfredo Deza)
- common: perfcounters now use atomics and go faster (Sage Weil)
- doc: CRUSH updates (John Wilkins)
- doc: osd primary affinity (John Wilkins)
- doc: pool quotas (John Wilkins)
- doc: pre-flight doc improvements (Kevin Dalley)
- doc: switch to an unencumbered font (Ross Turk)
- doc: update openstack docs (Josh Durgin)
- fix hppa arch build (Dmitry Smirnov)
- init-ceph: continue starting other daemons on crush or mount failure (#8343, Sage Weil)
- keyvaluestore: fix hint crash (#8381, Haomai Wang)

- libcephfs-java: build against older JNI headers (Greg Farnum)
- librados: fix rados_pool_list bounds checks (Sage Weil)
- mds: cephfs-journal-tool (John Spray)
- mds: improve Journaler on-disk format (John Spray)
- mds, libcephfs: use client timestamp for mtime/ctime (Sage Weil)
- mds: misc encoding improvements (John Spray)
- mds: misc fixes for multi-mds (Yan, Zheng)
- mds: OPTacker integration, dump_ops_in_flight (Greg Farnum)
- misc cleanup (Christophe Courtaut)
- mon: fix default replication pool ruleset choice (#8373, John Spray)
- mon: fix set cache_target_full_ratio (#8440, Geoffrey Hartz)
- mon: include per-pool ‘max avail’ in df output (Sage Weil)
- mon: prevent EC pools from being used with cephfs (Joao Eduardo Luis)
- mon: restore original weight when auto-marked out OSDs restart (Sage Weil)
- mon: use msg header tid for MMonGetVersionReply (Ilya Dryomov)
- osd: fix bogus assert during OSD shutdown (Sage Weil)
- osd: fix clone deletion case (#8334, Sam Just)
- osd: fix filestore removal corner case (#8332, Sam Just)
- osd: fix hang waiting for osdmap (#8338, Greg Farnum)
- osd: fix interval check corner case during peering (#8104, Sam Just)
- osd: fix journal-less operation (Sage Weil)
- osd: include backend information in metadata reported to mon (Sage Weil)
- rest-api: fix help (Ailing Zhang)
- rgw: check entity permission for put_metadata (#8428, Yehuda Sadeh)

V0.81

This is the first development release since Firefly. It includes a lot of work that we delayed merging while stabilizing things. Lots of new functionality, as well as several fixes that are baking a bit before getting backported.

UPGRADING

- CephFS support for the legacy anchor table has finally been removed. Users with file systems created before firefly should ensure that inodes with multiple hard links are modified *prior* to the upgrade to ensure that the backtraces are written properly. For example:

```
sudo find /mnt/cephfs -type f -links +1 -exec touch \{\}\ \;
```

- Disallow nonsensical ‘tier cache-mode’ transitions. From this point onward, ‘writeback’ can only transition to ‘forward’ and ‘forward’ can transition to 1) ‘writeback’ if there are dirty objects, or 2) any if there are no dirty objects.

NOTABLE CHANGES

- bash completion improvements (Wido den Hollander)
- brag: fixes, improvements (Loic Dachary)
- ceph-disk: handle corrupt volumes (Stuart Longlang)
- ceph-disk: partprobe as needed (Eric Eastman)
- ceph-fuse, libcephfs: asok hooks for handling session resets, timeouts (Yan, Zheng)
- ceph-fuse, libcephfs: improve traceless reply handling (Sage Weil)
- clang build fixes (John Spray, Danny Al-Gaaf)
- config: support G, M, K, etc. suffixes (Joao Eduardo Luis)
- coverity cleanups (Danny Al-Gaaf)
- doc: cache tiering (John Wilkins)
- doc: keystone integration docs (John Wilkins)
- doc: updated simple configuration guides (John Wilkins)
- libcephfs-java: fix gcj-jdk build (Dmitry Smirnov)
- librbd: check error code on cache invalidate (Josh Durgin)
- librbd: new libkrbd library for kernel map/unmap/showmapped (Ilya Dryomov)
- Makefile: fix out of source builds (Stefan Eilemann)
- mds: multi-mds fixes (Yan, Zheng)
- mds: remove legacy anchor table (Yan, Zheng)

- mds: remove legacy discover ino (Yan, Zheng)
- monclient: fix hang (Sage Weil)
- mon: prevent nonsensical cache-mode transitions (Joao Eduardo Luis)
- msgr: avoid big lock when sending (most) messages (Greg Farnum)
- osd: bound osdmap epoch skew between PGs (Sage Weil)
- osd: cache tier flushing fixes for snapped objects (Samuel Just)
- osd: fix agent early finish looping (David Zafman)
- osd: fix flush vs OpContext (Samuel Just)
- osd: fix MarkMeDown and other shutdown races (Sage Weil)
- osd: fix scrub vs cache bugs (Samuel Just)
- osd: fix trim of hitsets (Sage Weil)
- osd, msgr: fast-dispatch of OSD ops (Greg Farnum, Samuel Just)
- osd, objecter: resend ops on last_force_op_resend barrier; fix cache overlay op ordering (Sage Weil)
- osd: remove obsolete classic scrub code (David Zafman)
- osd: scrub PGs with invalid stats (Sage Weil)
- osd: simple snap trimmer throttle (Sage Weil)
- osd: use FIEMAP to inform copy_range (Haomai Wang)
- rbd-fuse: allow exposing single image (Stephen Taylor)
- rbd-fuse: fix unlink (Josh Durgin)
- removed mkcephfs (deprecated since dumpling)
- rgw: bucket link uses instance id (Yehuda Sadeh)
- rgw: fix memory leak following chunk read error (Yehuda Sadeh)
- rgw: fix URL escaping (Yehuda Sadeh)
- rgw: fix user manifest (Yehuda Sadeh)
- rgw: object and bucket rewrite functions to allow restriping old objects (Yehuda Sadeh)
- rgw: prevent multiobject PUT race (Yehuda Sadeh)
- rgw: send user manifest header (Yehuda Sadeh)
- test_librbd_fsx: test krbd as well as librbd (Ilya Dryomov)

V0.80.11 FIREFLY

This is a bugfix release for Firefly. This Firefly 0.80.x is nearing its planned end of life in January 2016 it may also be the last.

We recommend that all Firefly users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build/ops: /etc/init.d/radosgw restart does not work correctly ([issue#11140](#), [pr#5831](#), Dmitry Yatsushkevich)
- build/ops: Fix -Wno-format and -Werror=format-security options clash ([issue#13417](#), [pr#6207](#), Boris Ranto)
- build/ops: ceph-common needs python-argparse on older distros, but doesn't require it ([issue#12034](#), [pr#5217](#), Nathan Cutler)
- build/ops: ceph.spec.in running fdupes unnecessarily ([issue#12301](#), [pr#5224](#), Nathan Cutler)
- build/ops: ceph.spec.in: 50-rbd.rules conditional is wrong ([issue#12166](#), [pr#5225](#), Nathan Cutler)
- build/ops: ceph.spec.in: useless %py_requires breaks SLE11-SP3 build ([issue#12351](#), [pr#5394](#), Nathan Cutler)
- build/ops: fedora21 has junit, not junit4 ([issue#10728](#), [pr#6203](#), Ken Dreyer, Loic Dachary)
- build/ops: upstart: configuration is too generous on restarts ([issue#11798](#), [pr#5992](#), Sage Weil)
- common: Client admin socket leaks file descriptors ([issue#11535](#), [pr#4633](#), Jon Bernard)
- common: FileStore calls syncfs(2) even it is not supported ([issue#12512](#), [pr#5529](#), Danny Al-Gaaf, Kefu Chai, Jianpeng Ma)
- common: HeartBeat: include types ([issue#13088](#), [pr#6038](#), Sage Weil)
- common: Malformed JSON command output when non-ASCII strings are present ([issue#7387](#), [pr#4635](#), Kefu Chai, Tim Serong)
- common: Memory leak in Mutex.cc, pthread_mutexattr_init without pthread_mutexattr_destroy ([issue#11762](#), [pr#5403](#), Ketor Meng)
- common: Thread:pthread_attr_destroy(thread_attr) when done with it ([issue#12570](#), [pr#6325](#), Piotr Dałek, Zheng Qiankun)
- common: ThreadPool add/remove work queue methods not thread safe ([issue#12662](#), [pr#5991](#), Jason Dillaman)
- common: buffer: critical bufferlist::zero bug ([issue#12252](#), [pr#5388](#), Haomai Wang)
- common: log: take mutex while opening fd ([issue#12465](#), [pr#5406](#), Samuel Just)
- common: recursive lock of md_config_t (0) ([issue#12614](#), [pr#5814](#), Josh Durgin)
- crush: take crashes due to invalid arg ([issue#11602](#), [pr#4769](#), Sage Weil)
- doc: backport v0.80.10 release notes to firefly ([issue#11090](#), [pr#5307](#), Loic Dachary, Sage Weil)

- doc: update docs to point to download.ceph.com ([issue#13162](#), pr#[5993](#), Alfredo Deza)
- fs: MDSMonitor: handle MDSBeacon messages properly ([issue#11590](#), pr#[5199](#), Kefu Chai)
- fs: client nonce collision due to unshared pid namespaces ([issue#13032](#), pr#[6087](#), Josh Durgin, Sage Weil)
- librbd: Objectcacher setting max object counts too low ([issue#7385](#), pr#[4639](#), Jason Dillaman)
- librbd: aio calls may block ([issue#11056](#), pr#[4854](#), Haomai Wang, Sage Weil, Jason Dillaman)
- librbd: internal.cc: 1967: FAILED assert(watchers.size() == 1) ([issue#12176](#), pr#[5171](#), Jason Dillaman)
- mon: Clock skew causes missing summary and confuses Calamari ([issue#11877](#), pr#[4867](#), Thorsten Behrens)
- mon: EC pools are not allowed as cache pools, disallow in the mon ([issue#11650](#), pr#[5389](#), Samuel Just)
- mon: Make it more difficult to delete pools in firefly ([issue#11800](#), pr#[4788](#), Sage Weil)
- mon: MonitorDBStore: get_next_key() only if prefix matches ([issue#11786](#), pr#[5360](#), Joao Eduardo Luis)
- mon: PaxosService: call post_refresh() instead of post_paxos_update() ([issue#11470](#), pr#[5358](#), Joao Eduardo Luis)
- mon: add a cache layer over MonitorDBStore ([issue#12638](#), pr#[5698](#), Kefu Chai)
- mon: adding existing pool as tier with --force-nonempty clobbers removed_snaps ([issue#11493](#), pr#[5236](#), Sage Weil, Samuel Just)
- mon: ceph fails to compile with boost 1.58 ([issue#11576](#), pr#[5129](#), Kefu Chai)
- mon: does not check for IO errors on every transaction ([issue#13089](#), pr#[6091](#), Sage Weil)
- mon: get pools health'info have error ([issue#12402](#), pr#[5410](#), renhwztetecs)
- mon: increase globalid default for firefly ([issue#13255](#), pr#[6010](#), Sage Weil)
- mon: pgmonitor: wrong at/near target max" reporting ([issue#12401](#), pr#[5409](#), huangjun)
- mon: register_new_pgs() should check ruleno instead of its index ([issue#12210](#), pr#[5404](#), Xinze Chi)
- mon: scrub error (osdmap encoding mismatch?) upgrading from 0.80 to ~0.80.2 ([issue#8815](#), [issue#8674](#), [issue#9064](#), pr#[5200](#), Sage Weil, Zhiqiang Wang, Samuel Just)
- mon: the output is wrong when running ceph osd reweight ([issue#12251](#), pr#[5408](#), Joao Eduardo Luis)
- objecter: can get stuck in redirect loop if osdmap epoch == last_force_op_resend ([issue#11026](#), pr#[4597](#), Jianpeng Ma, Sage Weil)
- objecter: pg listing can deadlock when throttling is in use ([issue#9008](#), pr#[5043](#), Guang Yang)
- objecter: resend linger ops on split ([issue#9806](#), pr#[5062](#), Josh Durgin, Samuel Just)
- osd: Cleanup boost optionals for boost 1.56 ([issue#9983](#), pr#[5039](#), William A. Kennington III)
- osd: LibRadosTwoPools[EC]PP.PromoteSnap failure ([issue#10052](#), pr#[5050](#), Sage Weil)
- osd: Mutex Assert from PipeConnection::try_get_pipe ([issue#12437](#), pr#[5815](#), David Zafman)
- osd: PG stuck with remapped ([issue#9614](#), pr#[5044](#), Guang Yang)
- osd: PG::handle_advance_map: on_pool_change after handling the map change ([issue#12809](#), pr#[5988](#), Samuel Just)
- osd: PGLog: split divergent priors as well ([issue#11069](#), pr#[4631](#), Samuel Just)
- osd: PGLog::proc_replica_log: correctly handle case where entries between olog.head and log.tail were split out ([issue#11358](#), pr#[5287](#), Samuel Just)
- osd: WBThrottle::clear_object: signal on cond when we reduce throttle values ([issue#12223](#), pr#[5822](#), Samuel Just)
- osd: cache full mode still skips young objects ([issue#10006](#), pr#[5051](#), Xinze Chi, Zhiqiang Wang)
- osd: crash creating/deleting pools ([issue#12429](#), pr#[5526](#), John Spray)
- osd: explicitly specify OSD features in MOSDBoot ([issue#10911](#), pr#[4960](#), Sage Weil)
- osd: is_new_interval() fixes ([issue#11771](#), [issue#10399](#), pr#[5726](#), Samuel Just, Jason Dillaman)
- osd: make the all osd/filestore thread pool suicide timeouts separately configurable ([issue#11439](#), pr#[5823](#), Samuel Just)
- osd: object creation by write cannot use an offset on an erasure coded pool ([issue#11507](#), pr#[4632](#), Jianpeng Ma, Loic Dachary)
- osd: os/FileJournal: Fix journal write fail, align for direct io ([issue#12943](#), pr#[5619](#), Xie Rui)
- osd: osd/PGLog.cc: 732: FAILED assert(log.log.size() == log_keys_debug.size()) ([issue#12652](#), pr#[5820](#), Sage Weil)
- osd: read on chunk-aligned xattr not handled ([issue#12309](#), pr#[5235](#), Sage Weil)
- rgw: Change variable length array of std::strings (not legal in C++) to std::vector<std::string> ([issue#12467](#), pr#[4583](#), Daniel J. Hofmann)
- rgw: Civetweb RGW appears to report full size of object as downloaded when only partially downloaded ([issue#11851](#), pr#[5234](#), Yehuda Sadeh)
- rgw: Keystone PKI token expiration is not enforced ([issue#11367](#), pr#[4765](#), Anton Aksola)
- rgw: Object copy bug ([issue#11639](#), pr#[4762](#), Javier M. Mellid)
- rgw: RGW returns requested bucket name raw in "Bucket" response header ([issue#11860](#), [issue#12537](#), pr#[5730](#), Yehuda Sadeh, Wido den Hollander)
- rgw: Swift API: response for PUT on /container does not contain the mandatory Content-Length header when FCGI is used ([issue#11036](#), pr#[5170](#), Radoslaw Zarzynski)
- rgw: content length parsing calls strtol() instead of strtoll() ([issue#10701](#), pr#[5997](#), Yehuda Sadeh)
- rgw: delete bucket does not remove .bucket.meta file ([issue#11149](#), pr#[4641](#), Orit Wasserman)
- rgw: doesn't return 'x-timestamp' in header which is used by 'View Details' of OpenStack ([issue#8911](#), pr#[4584](#), Yehuda Sadeh)
- rgw: init some manifest fields when handling explicit objs ([issue#11455](#), pr#[5729](#), Yehuda Sadeh)
- rgw: logfile does not get chowned properly ([issue#12073](#), pr#[5233](#), Thorsten Behrens)
- rgw: logrotate.conf calls service with wrong init script name ([issue#12043](#), pr#[5390](#), wuxingyi)
- rgw: quota not respected in POST object ([issue#11323](#), pr#[4642](#), Sergey Arkhipov)
- rgw: swift smoke test fails on TestAccountUTF8 ([issue#11091](#), [issue#11438](#), [issue#12939](#), [issue#12157](#), [issue#12158](#),

issue#12363, pr#5532, Radoslaw Zarzynski, Orit Wasserman, Robin H. Johnson)

- rgw: use correct objv_tracker for bucket instance (issue#11416, pr#4535, Yehuda Sadeh)
- tests: ceph-fuse crash in test_client_recovery (issue#12673, pr#5813, Loic Dachary)
- tests: kernel_untar_build fails on EL7 (issue#11758, pr#6000, Greg Farnum)
- tests: qemu workunit refers to apt-mirror.front.sepia.ceph.com (issue#13420, pr#6328, Yuan Zhou, Sage Weil)
- tools: src/ceph-disk : disk zap sgdisk invocation (issue#11143, pr#4636, Thorsten Behrens, Owen Synge)
- tools: ceph-disk: sometimes the journal symlink is not created (issue#10146, pr#5541, Dan van der Ster)
- tools: ceph-disk: support NVMe device partitions (issue#11612, pr#4771, Ilja Slepnev)
- tools: ceph-post-file fails on rhel7 (issue#11836, pr#5037, Joseph McDonald, Sage Weil)
- tools: ceph_argparse_flag has no regular 3rd parameter (issue#11543, pr#4582, Thorsten Behrens)
- tools: use a new disk as journal disk,ceph-disk prepare fail (issue#10983, pr#4630, Loic Dachary)

V0.80.10 FIREFLY

This is a bugfix release for Firefly.

We recommend that all Firefly users upgrade.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build/ops: ceph.spec.in: package mkcephfs on EL6 (issue#11955, pr#4924, Ken Dreyer)
- build/ops: debian: ceph-test and rest-bench debug packages should require their respective binary packages (issue#11673, pr#4766, Ken Dreyer)
- build/ops: run RGW as root (issue#11453, pr#4638, Ken Dreyer)
- common: messages/MWatchNotify: include an error code in the message (issue#9193, pr#3944, Sage Weil)
- common: Rados.shutdown() dies with Illegal instruction (core dumped) (issue#10153, pr#3963, Federico Simoncelli)
- common: SimpleMessenger: allow RESETSESSION whenever we forget an endpoint (issue#10080, pr#3915, Greg Farnum)
- common: WorkQueue: make wait timeout on empty queue configurable (issue#10817, pr#3941, Samuel Just)
- crush: set_choose_tries = 100 for erasure code rulesets (issue#10353, pr#3824, Loic Dachary)
- doc: backport ceph-disk man page to Firefly (issue#10724, pr#3936, Nilamdyuti Goswami)
- doc: Fix ceph command manpage to match ceph -h (issue#10676, pr#3996, David Zafman)
- fs: mount.ceph: avoid spurious error message (issue#10351, pr#3927, Yan, Zheng)
- librados: Fix memory leak in python rados bindings (issue#10723, pr#3935, Josh Durgin)
- librados: fix resources leakage in RadosClient::connect() (issue#10425, pr#3828, Radoslaw Zarzynski)
- librados: Translate operation flags from C APIs (issue#10497, pr#3930, Matt Richards)
- librbd: acquire cache_lock before refreshing parent (issue#5488, pr#4206, Jason Dillaman)
- librbd: snap_remove should ignore -ENOENT errors (issue#11113, pr#4245, Jason Dillaman)
- mds: fix assertion caused by system clock backwards (issue#11053, pr#3970, Yan, Zheng)
- mon: ignore osd failures from before up_from (issue#10762, pr#3937, Sage Weil)
- mon: MonCap: take EntityName instead when expanding profiles (issue#10844, pr#3942, Joao Eduardo Luis)
- mon: Monitor: fix timecheck rounds period (issue#10546, pr#3932, Joao Eduardo Luis)
- mon: OSDMonitor: do not trust small values in osd epoch cache (issue#10787, pr#3823, Sage Weil)
- mon: OSDMonitor: fallback to json-pretty in case of invalid formatter (issue#9538, pr#4475, Loic Dachary)
- mon: PGMonitor: several stats output error fixes (issue#10257, pr#3826, Joao Eduardo Luis)
- objecter: fix map skipping (issue#9986, pr#3952, Ding Dinghua)
- osd: cache tiering: fix the atime logic of the eviction (issue#9915, pr#3949, Zhiqiang Wang)
- osd: cancel_pull: requeue waiters (issue#11244, pr#4415, Samuel Just)
- osd: check that source OSD is valid for MOSDRepScrub (issue#9555, pr#3947, Sage Weil)
- osd: DBObjectMap: lock header_lock on sync() (issue#9891, pr#3948, Samuel Just)
- osd: do not ignore deleted pgs on startup (issue#10617, pr#3933, Sage Weil)
- osd: ENOENT on clone (issue#11199, pr#4385, Samuel Just)
- osd: erasure-code-profile set races with erasure-code-profile rm (issue#11144, pr#4383, Loic Dachary)
- osd: FAILED assert(soid < scrubber.start || soid >= scrubber.end) (issue#11156, pr#4185, Samuel Just)
- osd: FileJournal: fix journalq population in do_read_entry() (issue#6003, pr#3960, Samuel Just)
- osd: fix negative degraded objects during backfilling (issue#7737, pr#4021, Guang Yang)
- osd: get the currently atime of the object in cache pool for eviction (issue#9985, pr#3950, Sage Weil)
- osd: load_pgs: we need to handle the case where an upgrade from earlier versions which ignored non-existent pgs resurrects a pg with a prehistoric osdmap (issue#11429, pr#4556, Samuel Just)
- osd: ObjectStore: Don't use largest_data_off to calc data_align. (issue#10014, pr#3954, Jianpeng Ma)
- osd: osd_types: op_queue_age_hist and fs_perf_stat should be in osd_stat_t::... (issue#10259, pr#3827, Samuel Just)
- osd: PG::actingset should be used when checking the number of acting OSDs for... (issue#11454, pr#4453, Guang Yang)

- osd: PG::all_unfound_are_queried_or_lost for non-existent osds ([issue#10976](#), [pr#4416](#), Mykola Golub)
- osd: PG: always clear_primary_state ([issue#10059](#), [pr#3955](#), Samuel Just)
- osd: PGLog.h: 279: FAILED assert(log.log.size() == log_keys_debug.size()) ([issue#10718](#), [pr#4382](#), Samuel Just)
- osd: PGLog: include rollback_info_trimmed_to in (read|write)_log ([issue#10157](#), [pr#3964](#), Samuel Just)
- osd: pg stuck stale after create with activation delay ([issue#11197](#), [pr#4384](#), Samuel Just)
- osd: ReplicatedPG: fail a non-blocking flush if the object is being scrubbed ([issue#8011](#), [pr#3943](#), Samuel Just)
- osd: ReplicatedPG::on_change: clean up callbacks_for_degraded_object ([issue#8753](#), [pr#3940](#), Samuel Just)
- osd: ReplicatedPG::scan_range: an object can disappear between the list and t... ([issue#10150](#), [pr#3962](#), Samuel Just)
- osd: requeue blocked op before flush it was blocked on ([issue#10512](#), [pr#3931](#), Sage Weil)
- rgw: check for timestamp for s3 keystone auth ([issue#10062](#), [pr#3958](#), Abhishek Lekshmanan)
- rgw: civetweb should use unique request id ([issue#11720](#), [pr#4780](#), Orit Wasserman)
- rgw: don't allow negative / invalid content length ([issue#11890](#), [pr#4829](#), Yehuda Sadeh)
- rgw: fail s3 POST auth if keystone not configured ([issue#10698](#), [pr#3966](#), Yehuda Sadeh)
- rgw: flush xml header on get acl request ([issue#10106](#), [pr#3961](#), Yehuda Sadeh)
- rgw: generate new tag for object when setting object attrs ([issue#11256](#), [pr#4571](#), Yehuda Sadeh)
- rgw: generate the "Date" HTTP header for civetweb. ([issue#11871,11891](#), [pr#4851](#), Radoslaw Zarzynski)
- rgw: keystone token cache does not work correctly ([issue#11125](#), [pr#4414](#), Yehuda Sadeh)
- rgw: merge manifests correctly when there's prefix override ([issue#11622](#), [pr#4697](#), Yehuda Sadeh)
- rgw: send appropriate op to cancel bucket index pending operation ([issue#10770](#), [pr#3938](#), Yehuda Sadeh)
- rgw: shouldn't need to disable rgw_socket_path if frontend is configured ([issue#11160](#), [pr#4275](#), Yehuda Sadeh)
- rgw: Swift API. Dump container's custom metadata. ([issue#10665](#), [pr#3934](#), Dmytro lurchenko)
- rgw: Swift API. Support for X-Remove-Container-Meta-{key} header. ([issue#10475](#), [pr#3929](#), Dmytro lurchenko)
- rgw: use correct objv_tracker for bucket instance ([issue#11416](#), [pr#4379](#), Yehuda Sadeh)
- tests: force checkout of submodules ([issue#11157](#), [pr#4079](#), Loic Dachary)
- tools: Backport ceph-objectstore-tool changes to firefly ([issue#12327](#), [pr#3866](#), David Zafman)
- tools: ceph-objectstore-tool: Output only unsupported features when incompatible ([issue#11176](#), [pr#4126](#), David Zafman)
- tools: ceph-objectstore-tool: Use exit status 11 for incompatible import attemp... ([issue#11139](#), [pr#4129](#), David Zafman)
- tools: Fix do_autogen.sh so that -L is allowed ([issue#11303](#), [pr#4247](#), Alfredo Deza)

V0.80.9 FIREFLY

This is a bugfix release for firefly. It fixes a performance regression in librbd, an important CRUSH misbehavior (see below), and several RGW bugs. We have also backported support for flock/fcntl locks to ceph-fuse and libcephfs.

We recommend that all Firefly users upgrade.

For more detailed information, see [the complete changelog](#).

ADJUSTING CRUSH MAPS

- This point release fixes several issues with CRUSH that trigger excessive data migration when adjusting OSD weights. These are most obvious when a very small weight change (e.g., a change from 0 to .01) triggers a large amount of movement, but the same set of bugs can also lead to excessive (though less noticeable) movement in other cases.

However, because the bug may already have affected your cluster, fixing it may trigger movement *back* to the more correct location. For this reason, you must manually opt-in to the fixed behavior.

In order to set the new tunable to correct the behavior:

```
ceph osd crush set-tunable straw_calc_version 1
```

Note that this change will have no immediate effect. However, from this point forward, any 'straw' bucket in your CRUSH map that is adjusted will get non-buggy internal weights, and that transition may trigger some rebalancing.

You can estimate how much rebalancing will eventually be necessary on your cluster with:

```
ceph osd getcrushmap -o /tmp/cm
crushtool -i /tmp/cm --num-rep 3 --test --show-mappings > /tmp/a 2>&1
crushtool -i /tmp/cm --set-straw-calc-version 1 -o /tmp/cm2
crushtool -i /tmp/cm2 --reweight -o /tmp/cm2
crushtool -i /tmp/cm2 --num-rep 3 --test --show-mappings > /tmp/b 2>&1
wc -l /tmp/a                                # num total mappings
diff -u /tmp/a /tmp/b | grep -c ^+      # num changed mappings
```

Divide the number of changed lines by the total number of lines [in](#) /tmp/a. We've found that most clusters are under 10%.

You can force [all](#) of this rebalancing to happen at once [with](#):

```
ceph osd crush reweight-all
```

Otherwise, it will happen at some unknown point [in](#) the future when CRUSH weights are [next](#) adjusted.

NOTABLE CHANGES

- ceph-fuse: flock, fcntl lock support (Yan, Zheng, Greg Farnum)
- crush: fix straw bucket weight calculation, add straw_calc_version tunable (#10095 Sage Weil)
- crush: fix tree bucket (Rongzu Zhu)
- crush: fix underflow of tree weights (Loic Dachary, Sage Weil)
- crushtool: add -reweight (Sage Weil)
- librbd: complete pending operations before losing image (#10299 Jason Dillaman)
- librbd: fix read caching performance regression (#9854 Jason Dillaman)
- librbd: gracefully handle deleted/renamed pools (#10270 Jason Dillaman)
- mon: fix dump of chooseleaf_vary_r tunable (Sage Weil)
- osd: fix PG ref leak in snaptrimmer on peering (#10421 Kefu Chai)
- osd: handle no-op write with snapshot (#10262 Sage Weil)
- radosgw-admin: create subuser when creating user (#10103 Yehuda Sadeh)
- rgw: change multipart uplaod id magic (#10271 Georgio Dimitrakakis, Yehuda Sadeh)
- rgw: don't overwrite bucket/object owner when setting ACLs (#10978 Yehuda Sadeh)
- rgw: enable IPv6 for embedded civetweb (#10965 Yehuda Sadeh)
- rgw: fix partial swift GET (#10553 Yehuda Sadeh)
- rgw: fix quota disable (#9907 Dong Lei)
- rgw: index swift keys appropriately (#10471 Hemant Burman, Yehuda Sadeh)
- rgw: make setattrs update bucket index (#5595 Yehuda Sadeh)
- rgw: pass civetweb configurables (#10907 Yehuda Sadeh)
- rgw: remove swift user manifest (DLO) hash calculation (#9973 Yehuda Sadeh)
- rgw: return correct len for 0-len objects (#9877 Yehuda Sadeh)
- rgw: S3 object copy content-type fix (#9478 Yehuda Sadeh)
- rgw: send ETag on S3 object copy (#9479 Yehuda Sadeh)
- rgw: send HTTP status reason explicitly in fastcgi (Yehuda Sadeh)
- rgw: set ulimit -n from sysvinit (el6) init script (#9587 Sage Weil)
- rgw: update swift subuser permission masks when authenticating (#9918 Yehuda Sadeh)
- rgw: URL decode query params correctly (#10271 Georgio Dimitrakakis, Yehuda Sadeh)
- rgw: use attrs when reading object attrs (#10307 Yehuda Sadeh)
- rgw: use rn for http headers (#9254 Benedikt Fraunhofer, Yehuda Sadeh)

V0.80.8 FIREFLY

This is a long-awaited bugfix release for firefly. It has several important (but relatively rare) OSD peering fixes, performance issues when snapshots are trimmed, several RGW fixes, a paxos corner case fix, and some packaging updates.

We recommend that all users for v0.80.x firefly upgrade when it is convenient to do so.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build: remove stack-execute bit from assembled code sections (#10114 Dan Mick)
- ceph-disk: fix dmcrypt key permissions (#9785 Loic Dachary)
- ceph-disk: fix keyring location (#9653 Loic Dachary)
- ceph-disk: make partition checks more robust (#9721 #9665 Loic Dachary)
- ceph: cleanly shut down librados context on shutdown (#8797 Dan Mick)
- common: add \$cctid config metavariable (#6228 Adam Crume)
- crush: align rule and ruleset ids (#9675 Xiaoxi Chen)
- crush: fix negative weight bug during create_or_move_item (#9998 Paweł Sadowski)
- crush: fix potential buffer overflow in erasure rules (#9492 Johnu George)
- debian: fix python-ceph -> ceph file movement (Sage Weil)

- libcephfs,ceph-fuse: fix flush tid wraparound bug (#9869 Greg Farnum, Yan, Zheng)
- libcephfs: close fd before umount (#10415 Yan, Zheng)
- librados: fix crash from C API when read timeout is enabled (#9582 Sage Weil)
- librados: handle reply race with pool deletion (#10372 Sage Weil)
- librbd: cap memory utilization for read requests (Jason Dillaman)
- librbd: do not close a closed parent image on failure (#10030 Jason Dillaman)
- librbd: fix diff tests (#10002 Josh Durgin)
- librbd: protect list_children from invalid pools (#10123 Jason Dillaman)
- make check improvements (Loic Dachary)
- mds: fix ctime updates (#9514 Greg Farnum)
- mds: fix journal import tool (#10025 John Spray)
- mds: fix rare NULL deref in cap flush handler (Greg Farnum)
- mds: handle unknown lock messages (Yan, Zheng)
- mds: store backtrace for straydir (Yan, Zheng)
- mon: abort startup if disk is full (#9502 Joao Eduardo Luis)
- mon: add paxos instrumentation (Sage Weil)
- mon: fix double-free in rare OSD startup path (Sage Weil)
- mon: fix osdmap trimming (#9987 Sage Weil)
- mon: fix paxos corner cases (#9301 #9053 Sage Weil)
- osd: cancel callback on blacklisted watchers (#8315 Samuel Just)
- osd: cleanly abort set-alloc-hint operations during upgrade (#9419 David Zafman)
- osd: clear rollback PG metadata on PG deletion (#9293 Samuel Just)
- osd: do not abort deep scrub if hinfo is missing (#10018 Loic Dachary)
- osd: erasure-code regression tests (Loic Dachary)
- osd: fix distro metadata reporting for SUSE (#8654 Danny Al-Gaaf)
- osd: fix full OSD checks during backfill (#9574 Samuel Just)
- osd: fix ioprio parsing (#9677 Loic Dachary)
- osd: fix journal direct-io shutdown (#9073 Mark Kirkwood, Ma Jianpeng, Somnath Roy)
- osd: fix journal dump (Ma Jianpeng)
- osd: fix occasional stall during peering or activation (Sage Weil)
- osd: fix past_interval display bug (#9752 Loic Dachary)
- osd: fix rare crash triggered by admin socket dump_ops_in_filght (#9916 Dong Lei)
- osd: fix snap trimming performance issues (#9487 #9113 Samuel Just, Sage Weil, Dan van der Ster, Florian Haas)
- osd: fix snapdir handling on cache eviction (#8629 Sage Weil)
- osd: handle map gaps in map advance code (Sage Weil)
- osd: handle undefined CRUSH results in interval check (#9718 Samuel Just)
- osd: include shard in JSON dump of ghobject (#10063 Loic Dachary)
- osd: make backfill reservation denial handling more robust (#9626 Samuel Just)
- osd: make misdirected op checks handle EC + primary affinity (#9835 Samuel Just, Sage Weil)
- osd: mount XFS with inode64 by default (Sage Weil)
- osd: other misc bugs (#9821 #9875 Samuel Just)
- rgw: add .log to default log path (#9353 Alexandre Marangone)
- rgw: clean up fcgi request context (#10194 Yehuda Sadeh)
- rgw: convert header underscores to dashes (#9206 Yehuda Sadeh)
- rgw: copy object data if copy target is in different pool (#9039 Yehuda Sadeh)
- rgw: don't try to authenticate CORS preflight request (#8718 Robert Hubbard, Yehuda Sadeh)
- rgw: fix civetweb URL decoding (#8621 Yehuda Sadeh)
- rgw: fix hash calculation during PUT (Yehuda Sadeh)
- rgw: fix misc bugs (#9089 #9201 Yehuda Sadeh)
- rgw: fix object tail test (#9226 Sylvain Munaut, Yehuda Sadeh)
- rgw: make sysvinit script run rgw under systemd context as needed (#10125 Loic Dachary)
- rgw: separate civetweb log from rgw log (Yehuda Sadeh)
- rgw: set length for keystone token validations (#7796 Mark Kirkwood, Yehuda Sadeh)
- rgw: subuser creation fixes (#8587 Yehuda Sadeh)
- rpm: misc packaging improvements (Sandon Van Ness, Dan Mick, Erik Logthenberg, Boris Ranto)
- rpm: use standard udev rules for CentOS7/RHEL7 (#9747 Loic Dachary)

V0.80.7 FIREFLY

This release fixes a few critical issues with v0.80.6, particularly with clusters running mixed versions.

We recommend that all v0.80.x Firefly users upgrade to this release.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- osd: fix invalid memory reference in log trimming (#9731 Samuel Just)
- osd: fix use-after-free in cache tiering code (#7588 Sage Weil)
- osd: remove bad backfill assertion for mixed-version clusters (#9696 Samuel Just)

V0.80.6 FIREFLY

This is a major bugfix release for firefly, fixing a range of issues in the OSD and monitor, particularly with cache tiering. There are also important fixes in librados, with the watch/notify mechanism used by librbd, and in radosgw.

A few pieces of new functionality have been backported, including improved ‘ceph df’ output (view amount of writeable space per pool), support for non-default cluster names when using sysvinit or systemd, and improved (and fixed) support for dmcrypt.

We recommend that all v0.80.x Firefly users upgrade to this release.

For more detailed information, see [the complete changelog](#).

NOTABLE CHANGES

- build: fix atomic64_t on i386 (#8969 Sage Weil)
- build: fix build on alpha (Michael Cree, Dmitry Smirnov)
- build: fix build on hppa (Dmitry Smirnov)
- build: fix yasm detection on x32 arch (Sage Weil)
- ceph-disk: fix ‘list’ function with dmcrypt (Sage Weil)
- ceph-disk: fix dmcrypt support (Alfredo Deza)
- ceph: allow non-default cluster to be specified (#8944)
- common: fix dup log messages to mon (#9080 Sage Weil)
- global: write pid file when -f is used (systemd, upstart) (Alexandre Oliva)
- librados: fix crash when read timeout is enabled (#9362 Matthias Kiefer, Sage Weil)
- librados: fix lock leaks in error paths (#9022 Pavan Rallabhandi)
- librados: fix watch resend on PG acting set change (#9220 Samuel Just)
- librados: python: fix aio_read handling with 0 (Mohammad Salehe)
- librbd: add interface to invalidate cached data (Josh Durgin)
- librbd: fix crash when using clone of flattened image (#8845 Josh Durgin)
- librbd: fix error path cleanup on open (#8912 Josh Durgin)
- librbd: fix null pointer check (Danny Al-Gaaf)
- librbd: limit dirty object count (Haomai Wang)
- mds: fix rstats for root and mdsdir (Yan, Zheng)
- mon: add ‘get’ command for new cache tier pool properties (Joao Eduardo Luis)
- mon: add ‘osd pool get-quota’ (#8523 Joao Eduardo Luis)
- mon: add cluster fingerprint (Sage Weil)
- mon: disallow nonsensical cache-mode transitions (#8155 Joao Eduardo Luis)
- mon: fix cache tier rounding error on i386 (Sage Weil)
- mon: fix occasional memory leak (#9176 Sage Weil)
- mon: fix reported latency for ‘osd perf’ (#9269 Samuel Just)
- mon: include ‘max avail’ in ‘ceph df’ output (Sage Weil, Xioaxi Chen)
- mon: persistently mark pools where scrub may find incomplete clones (#8882 Sage Weil)
- mon: preload erasure plugins (Loic Dachary)
- mon: prevent cache-specific settings on non-tier pools (#8696 Joao Eduardo Luis)
- mon: reduce log spam (Aanchal Agrawal, Sage Weil)
- mon: warn when cache pools have no hit_sets enabled (Sage Weil)
- msgr: fix trivial memory leak (Sage Weil)
- osd: automatically scrub PGs with invalid stats (#8147 Sage Weil)
- osd: avoid sharing PG metadata that is not durable (Samuel Just)
- osd: cap hit_set size (#9339 Samuel Just)
- osd: create default erasure profile if needed (#8601 Loic Dachary)
- osd: dump tid as JSON int (not string) where appropriate (Joao Eduardo Luis)
- osd: encode blacklist in deterministic order (#9211 Sage Weil)
- osd: fix behavior when cache tier has no hit_sets enabled (#8982 Sage Weil)
- osd: fix cache tier flushing of snapshots (#9054 Samuel Just)
- osd: fix cache tier op ordering when going from full to non-full (#8931 Sage Weil)
- osd: fix crash on dup recovery reservation (#8863 Sage Weil)
- osd: fix division by zero when pg_num adjusted with no OSDs (#9052 Sage Weil)

- osd: fix hint crash in experimental keyvaluestore_dev backend (Hoamai Wang)
- osd: fix leak in copyfrom cancellation (#8894 Samuel Just)
- osd: fix locking for copyfrom finish (#8889 Sage Weil)
- osd: fix long filename handling in backend (#8701 Sage Weil)
- osd: fix min_size check with backfill (#9497 Samuel Just)
- osd: fix mount/remount sync race (#9144 Sage Weil)
- osd: fix object listing + erasure code bug (Guang Yang)
- osd: fix race on reconnect to failed OSD (#8944 Greg Farnum)
- osd: fix recovery reservation deadlock (Samuel Just)
- osd: fix tiering agent arithmetic for negative values (#9082 Karan Singh)
- osd: improve shutdown order (#9218 Sage Weil)
- osd: improve subop discard logic (#9259 Samuel Just)
- osd: introduce optional sleep, io priority for scrub and snap trim (Sage Weil)
- osd: make scrub check for and remove stale erasure-coded objects (Samuel Just)
- osd: misc fixes (#9481 #9482 #9179 Sameul Just)
- osd: mix keyvaluestore_dev improvements (Haomai Wang)
- osd: only require CRUSH features for rules that are used (#8963 Sage Weil)
- osd: preload erasure plugins on startup (Loic Dachary)
- osd: prevent PGs from falling behind when consuming OSDMaps (#7576 Sage Weil)
- osd: prevent old clients from using tiered pools (#8714 Sage Weil)
- osd: set min_size on erasure pools to data chunk count (Sage Weil)
- osd: trim old erasure-coded objects more aggressively (Samuel Just)
- rados: enforce erasure code alignment (Lluis Pamies-Juarez)
- rgw: align object stripes with erasure pool alignment (#8442 Yehuda Sadeh)
- rgw: don't send error body on HEAD for civetweb (#8539 Yehuda Sadeh)
- rgw: fix crash in CORS preflight request (Yehuda Sadeh)
- rgw: fix decoding of + in URL (#8702 Brian Rak)
- rgw: fix object removal on object create (#8972 Patrycja Szabowska, Yehuda Sadeh)
- systemd: use systemd-run when starting radosgw (JuanJose Galvez)
- sysvinit: support non-default cluster name (Alfredo Deza)

V0.80.5 FIREFLY

This release fixes a few important bugs in the radosgw and fixes several packaging and environment issues, including OSD log rotation, systemd environments, and daemon restarts on upgrade.

We recommend that all v0.80.x Firefly users upgrade, particularly if they are using upstart, systemd, or radosgw.

NOTABLE CHANGES

- ceph-dencoder: do not needlessly link to librgw, librados, etc. (Sage Weil)
- do not needlessly link binaries to leveldb (Sage Weil)
- mon: fix mon crash when no auth keys are present (#8851, Joao Eduardo Luis)
- osd: fix cleanup (and avoid occasional crash) during shutdown (#7981, Sage Weil)
- osd: fix log rotation under upstart (Sage Weil)
- rgw: fix multipart upload when object has irregular size (#8846, Yehuda Sadeh, Sylvain Munaut)
- rgw: improve bucket listing S3 compatibility (#8858, Yehuda Sadeh)
- rgw: improve delimited bucket listing (Yehuda Sadeh)
- rpm: do not restart daemons on upgrade (#8849, Alfredo Deza)

For more detailed information, see [the complete changelog](#).

V0.80.4 FIREFLY

This Firefly point release fixes an potential data corruption problem when ceph-osd daemons run on top of XFS and service Firefly librbd clients. A recently added allocation hint that RBD utilizes triggers an XFS bug on some kernels (Linux 3.2, and likely others) that leads to data corruption and deep-scrub errors (and inconsistent PGs). This release avoids the situation by disabling the allocation hint until we can validate which kernels are affected and/or are known to be safe to use the hint on.

We recommend that all v0.80.x Firefly users urgently upgrade, especially if they are using RBD.

NOTABLE CHANGES

- osd: disable XFS extsize hint by default (#8830, Samuel Just)
- rgw: fix extra data pool default name (Yehuda Sadeh)

For more detailed information, see [the complete changelog](#).

V0.80.3 FIREFLY

This is the third Firefly point release. It includes a single fix for a radosgw regression that was discovered in v0.80.2 right after it was released.

We recommend that all v0.80.x Firefly users upgrade.

NOTABLE CHANGES

- radosgw: fix regression in manifest decoding (#8804, Sage Weil)

For more detailed information, see [the complete changelog](#).

V0.80.2 FIREFLY

This is the second Firefly point release. It contains a range of important fixes, including several bugs in the OSD cache tiering, some compatibility checks that affect upgrade situations, several radosgw bugs, and an irritating and unnecessary feature bit check that prevents older clients from communicating with a cluster with any erasure coded pools.

One someone large change in this point release is that the ceph RPM package is separated into a ceph and ceph-common package, similar to Debian. The ceph-common package contains just the client libraries without any of the server-side daemons.

We recommend that all v0.80.x Firefly users skip this release and use v0.80.3.

NOTABLE CHANGES

- ceph-disk: better debug logging (Alfredo Deza)
- ceph-disk: fix preparation of OSDs with dmcrypt (#6700, Stephen F Taylor)
- ceph-disk: partprobe on prepare to fix dm-crypt (#6966, Eric Eastman)
- do not require ERASURE_CODE feature from clients (#8556, Sage Weil)
- libcephfs-java: build with older JNI headers (Greg Farnum)
- libcephfs-java: fix build with gcj-jdk (Dmitry Smirnov)
- librados: fix osd op tid for redirected ops (#7588, Samuel Just)
- librados: fix rados_pool_list buffer bounds checks (#8447, Sage Weil)
- librados: resend ops when pool overlay changes (#8305, Sage Weil)
- librbd, ceph-fuse: reduce CPU overhead for clean object check in cache (Haomai Wang)
- mon: allow deletion of cephfs pools (John Spray)
- mon: fix default pool ruleset choice (#8373, John Spray)
- mon: fix health summary for mon low disk warning (Sage Weil)
- mon: fix ‘osd pool set <pool> cache_target_full_ratio’ (Geoffrey Hartz)
- mon: fix quorum feature check (Greg Farnum)
- mon: fix request forwarding in mixed firefly+dumpling clusters 9#8727, Joao Eduardo Luis)
- mon: fix rule vs ruleset check in ‘osd pool set ... crush_ruleset’ command (John Spray)
- mon: make osd ‘down’ count accurate (Sage Weil)
- mon: set ‘next commit’ in primary-affinity reply (Ilya Dryomov)
- mon: verify CRUSH features are supported by all mons (#8738, Greg Farnum)
- msgr: fix sequence negotiation during connection reset (Guang Yang)
- osd: block scrub on blocked objects (#8011, Samuel Just)
- osd: call XFS hint ioctl less often (#8241, Ilya Dryomov)
- osd: copy xattr spill out marker on clone (Haomai Wang)
- osd: fix flush of snapped objects (#8334, Samuel Just)
- osd: fix hashindex restart of merge operation (#8332, Samuel Just)
- osd: fix osdmap subscription bug causing startup hang (Greg Farnum)
- osd: fix potential null deref (#8328, Sage Weil)
- osd: fix shutdown race (#8319, Sage Weil)
- osd: handle ‘none’ in CRUSH results properly during peering (#8507, Samuel Just)
- osd: set no spill out marker on new objects (Greg Farnum)

- osd: skip op ordering debug checks on tiered pools (#8380, Sage Weil)
- rados: enforce ‘put’ alignment (Lluis Pamies-Juarez)
- rest-api: fix for ‘rx’ commands (Ailing Zhang)
- rgw: calc user manifest etag and fix check (#8169, #8436, Yehuda Sadeh)
- rgw: fetch attrs on multipart completion (#8452, Yehuda Sadeh, Sylvain Munaut)
- rgw: fix buffer overflow for long instance ids (#8608, Yehuda Sadeh)
- rgw: fix entity permission check on metadata put (#8428, Yehuda Sadeh)
- rgw: fix multipart retry race (#8269, Yehuda Sadeh)
- rpm: split ceph into ceph and ceph-common RPMs (Sandon Van Ness, Dan Mick)
- sysvinit: continue startin daemons after failure doing mount (#8554, Sage Weil)

For more detailed information, see [the complete changelog](#).

V0.80.1 FIREFLY

This first Firefly point release fixes a few bugs, the most visible being a problem that prevents scrub from completing in some cases.

NOTABLE CHANGES

- osd: revert incomplete scrub fix (Samuel Just)
- rgw: fix stripe calculation for manifest objects (Yehuda Sadeh)
- rgw: improve handling, memory usage for abort reads (Yehuda Sadeh)
- rgw: send Swift user manifest HTTP header (Yehuda Sadeh)
- libcephfs, ceph-fuse: expose MDS session state via admin socket (Yan, Zheng)
- osd: add simple throttle for snap trimming (Sage Weil)
- monclient: fix possible hang from ill-timed monitor connection failure (Sage Weil)
- osd: fix trimming of past HitSets (Sage Weil)
- osd: fix whiteouts for non-writeback cache modes (Sage Weil)
- osd: prevent divide by zero in tiering agent (David Zafman)
- osd: prevent busy loop when tiering agent can do no work (David Zafman)

For more detailed information, see [the complete changelog](#).

V0.80 FIREFLY

This release will form the basis for our long-term supported release Firefly, v0.80.x. The big new features are support for erasure coding and cache tiering, although a broad range of other features, fixes, and improvements have been made across the code base. Highlights include:

- *Erasure coding*: support for a broad range of erasure codes for lower storage overhead and better data durability.
- *Cache tiering*: support for creating ‘cache pools’ that store hot, recently accessed objects with automatic demotion of colder data to a base tier. Typically the cache pool is backed by faster storage devices like SSDs.
- *Primary affinity*: Ceph now has the ability to skew selection of OSDs as the “primary” copy, which allows the read workload to be cheaply skewed away from parts of the cluster without migrating any data.
- *Key/value OSD backend* (experimental): An alternative storage backend for Ceph OSD processes that puts all data in a key/value database like leveldb. This provides better performance for workloads dominated by key/value operations (like radosgw bucket indices).
- *Standalone radosgw* (experimental): The radosgw process can now run in a standalone mode without an apache (or similar) web server or fastcgi. This simplifies deployment and can improve performance.

We expect to maintain a series of stable releases based on v0.80 Firefly for as much as a year. In the meantime, development of Ceph continues with the next release, Giant, which will feature work on the CephFS distributed file system, more alternative storage backends (like RocksDB and f2fs), RDMA support, support for pyramid erasure codes, and additional functionality in the block device (RBD) like copy-on-read and multisite mirroring.

UPGRADE SEQUENCING

- If your existing cluster is running a version older than v0.67 Dumpling, please first upgrade to the latest Dumpling release before upgrading to v0.80 Firefly. Please refer to the [Dumpling upgrade](#) documentation.
- We recommend adding the following to the [mon] section of your ceph.conf prior to upgrade:

```
mon warn on legacy crush tunables = false
```

This will prevent health warnings due to the use of legacy CRUSH placement. Although it is possible to rebalance existing data across your cluster (see the upgrade notes below), we do not normally recommend it for production environments as a large amount of data will move and there is a significant performance impact from the rebalancing.

- Upgrade daemons in the following order:

1. Monitors
2. OSDs
3. MDSs and/or radosgw

If the ceph-mds daemon is restarted first, it will wait until all OSDs have been upgraded before finishing its startup sequence. If the ceph-mon daemons are not restarted prior to the ceph-osd daemons, they will not correctly register their new capabilities with the cluster and new features may not be usable until they are restarted a second time.

- Upgrade radosgw daemons together. There is a subtle change in behavior for multipart uploads that prevents a multipart request that was initiated with a new radosgw from being completed by an old radosgw.

UPGRADING FROM V0.79

- OSDMap's json-formatted dump changed for keys 'full' and 'nearfull'. What was previously being outputted as 'true' or 'false' strings are now being outputted 'true' and 'false' booleans according to json syntax.
- HEALTH_WARN on 'mon osd down out interval == 0'. Having this option set to zero on the leader acts much like having the 'noout' flag set. This warning will only be reported if the monitor getting the 'health' or 'status' request has this option set to zero.
- Monitor 'auth' commands now require the mon 'x' capability. This matches dumpling v0.67.x and earlier, but differs from emperor v0.72.x.
- A librados WATCH operation on a non-existent object now returns ENOENT; previously it did not.
- Librados interface change: As there are no partial writes, the rados_write() and rados_append() operations now return 0 on success like rados_write_full() always has. This includes the C++ interface equivalents and AIO return values for the aio variants.
- The radosgw init script (sysvinit) now requires that the 'host = ...' line in ceph.conf, if present, match the short hostname (the output of 'hostname -s'), not the fully qualified hostname or the (occasionally non-short) output of 'hostname'. Failure to adjust this when upgrading from emperor or dumpling may prevent the radosgw daemon from starting.

UPGRADING FROM V0.72 EMPEROR

- See notes above.
- The 'ceph -s' or 'ceph status' command's 'num_in_osds' field in the JSON and XML output has been changed from a string to an int.
- The recently added 'ceph mds set allow_new_snaps' command's syntax has changed slightly; it is now 'ceph mds set allow_new_snaps true'. The 'unset' command has been removed; instead, set the value to 'false'.
- The syntax for allowing snapshots is now 'mds set allow_new_snaps <true|false>' instead of 'mds <set unset> allow_new_snaps'.
- 'rbd ls' on a pool which never held rbd images now exits with code 0. It outputs nothing in plain format, or an empty list in non-plain format. This is consistent with the behavior for a pool which used to hold images, but contains none. Scripts relying on this behavior should be updated.
- The MDS requires a new OSD operation TMAP2OMAP, added in this release. When upgrading, be sure to upgrade and restart the ceph-osd daemons before the ceph-mds daemon. The MDS will refuse to start if any up OSDs do not support the new feature.
- The 'ceph mds set_max_mds N' command is now deprecated in favor of 'ceph mds set max_mds N'.
- The 'osd pool create ...' syntax has changed for erasure pools.
- The default CRUSH rules and layouts are now using the 'bobtail' tunables and defaults. Upgraded clusters using the old values will now present with a health WARN state. This can be disabled by adding 'mon warn on legacy crush tunables = false' to ceph.conf and restarting the monitors. Alternatively, you can switch to the new tunables with 'ceph osd crush tunables firefly,' but keep in mind that this will involve moving a *significant* portion of the data already stored in the cluster and in a large cluster may take several days to complete. We do not recommend adjusting tunables on a production cluster.
- We now default to the 'bobtail' CRUSH tunable values that are first supported by Ceph clients in bobtail (v0.56) and Linux kernel version v3.9. If you plan to access a newly created Ceph cluster with an older kernel client, you should use 'ceph osd crush tunables legacy' to switch back to the legacy behavior. Note that making that change will likely result in some data movement in the system, so adjust the setting before populating the new cluster with data.
- We now set the HASHPSPOOL flag on newly created pools (and new clusters) by default. Support for this flag first

appeared in v0.64; v0.67 Dumpling is the first major release that supports it. It is first supported by the Linux kernel version v3.9. If you plan to access a newly created Ceph cluster with an older kernel or clients (e.g., librados, librbd) from a pre-dumpling Ceph release, you should add ‘osd pool default flag hashspool = false’ to the ‘[global]’ section of your ‘ceph.conf’ prior to creating your monitors (e.g., after ‘ceph-deploy new’ but before ‘ceph-deploy mon create ...’).

- The configuration option ‘osd pool default crush rule’ is deprecated and replaced with ‘osd pool default crush replicated ruleset’. ‘osd pool default crush rule’ takes precedence for backward compatibility and a deprecation warning is displayed when it is used.
- As part of fix for #6796, ‘ceph osd pool set <pool> <var> <arg>’ now receives <arg> as an integer instead of a string. This affects how ‘hashspool’ flag is set/unset: instead of ‘true’ or ‘false’, it now must be ‘0’ or ‘1’.
- The behavior of the CRUSH ‘indep’ choose mode has been changed. No ceph cluster should have been using this behavior unless someone has manually extracted a crush map, modified a CRUSH rule to replace ‘firstn’ with ‘indep’, recompiled, and reinjected the new map into the cluster. If the ‘indep’ mode is currently in use on a cluster, the rule should be modified to use ‘firstn’ instead, and the administrator should wait until any data movement completes before upgrading.
- The ‘osd dump’ command now dumps pool snaps as an array instead of an object.

UPGRADING FROM V0.67 DUMPLING

- See notes above.
- ceph-fuse and radosgw now use the same default values for the admin socket and log file paths that the other daemons (ceph-osd, ceph-mon, etc.) do. If you run these daemons as non-root, you may need to adjust your ceph.conf to disable these options or to adjust the permissions on /var/run/ceph and /var/log/ceph.
- The MDS now disallows snapshots by default as they are not considered stable. The command ‘ceph mds set allow_snaps’ will enable them.
- For clusters that were created before v0.44 (pre-argonaut, Spring 2012) and store radosgw data, the auto-upgrade from TMAP to OMAP objects has been disabled. Before upgrading, make sure that any buckets created on pre-argonaut releases have been modified (e.g., by PUTing and then DELETEing an object from each bucket). Any cluster created with argonaut (v0.48) or a later release or not using radosgw never relied on the automatic conversion and is not affected by this change.
- Any direct users of the ‘tmap’ portion of the librados API should be aware that the automatic tmap -> omap conversion functionality has been removed.
- Most output that used K or KB (e.g., for kilobyte) now uses a lower-case k to match the official SI convention. Any scripts that parse output and check for an upper-case K will need to be modified.
- librados::Rados::pool_create_async() and librados::Rados::pool_delete_async() don’t drop a reference to the completion object on error, caller needs to take care of that. This has never really worked correctly and we were leaking an object
- ‘ceph osd crush set <id> <weight> <loc..>’ no longer adds the osd to the specified location, as that’s a job for ‘ceph osd crush add’. It will however continue to work just the same as long as the osd already exists in the crush map.
- The OSD now enforces that class write methods cannot both mutate an object and return data. The rbd.assign_bid method, the lone offender, has been removed. This breaks compatibility with pre-bobtail librbd clients by preventing them from creating new images.
- librados now returns on commit instead of ack for synchronous calls. This is a bit safer in the case where both OSDs and the client crash, and is probably how it should have been acting from the beginning. Users are unlikely to notice but it could result in lower performance in some circumstances. Those who care should switch to using the async interfaces, which let you specify safety semantics precisely.
- The C++ librados AioComplete::get_version() method was incorrectly returning an int (usually 32-bits). To avoid breaking library compatibility, a get_version64() method is added that returns the full-width value. The old method is deprecated and will be removed in a future release. Users of the C++ librados API that make use of the get_version() method should modify their code to avoid getting a value that is truncated from 64 to to 32 bits.

NOTABLE CHANGES SINCE V0.79

- ceph-fuse, libcephfs: fix several caching bugs (Yan, Zheng)
- ceph-fuse: trim inodes in response to mds memory pressure (Yan, Zheng)
- librados: fix inconsistencies in API error values (David Zafman)
- librados: fix watch operations with cache pools (Sage Weil)
- librados: new snap rollback operation (David Zafman)
- mds: fix respawn (John Spray)
- mds: misc bugs (Yan, Zheng)
- mds: misc multi-mds fixes (Yan, Zheng)
- mds: use shared_ptr for requests (Greg Farnum)
- mon: fix peer feature checks (Sage Weil)
- mon: require ‘x’ mon caps for auth operations (Joao Luis)
- mon: shutdown when removed from mon cluster (Joao Luis)

- msgr: fix locking bug in authentication (Josh Durgin)
- osd: fix bug in journal replay/restart (Sage Weil)
- osd: many many many bug fixes with cache tiering (Samuel Just)
- osd: track omap and hit_set objects in pg stats (Samuel Just)
- osd: warn if agent cannot enable due to invalid (post-split) stats (Sage Weil)
- rados bench: track metadata for multiple runs separately (Guang Yang)
- rgw: fixed subuser modify (Yehuda Sadeh)
- rpm: fix redhat-lsb dependency (Sage Weil, Alfredo Deza)

NOTABLE CHANGES SINCE V0.72 EMPEROR

- buffer: some zero-copy groundwork (Josh Durgin)
- build: misc improvements (Ken Dreyer)
- ceph-conf: stop creating bogus log files (Josh Durgin, Sage Weil)
- ceph-crush-location: new hook for setting CRUSH location of osd daemons on start)
- ceph-disk: avoid fd0 (Loic Dachary)
- ceph-disk: generalize path names, add tests (Loic Dachary)
- ceph-disk: misc improvements for puppet (Loic Dachary)
- ceph-disk: several bug fixes (Loic Dachary)
- ceph-fuse: fix race for sync reads (Sage Weil)
- ceph-fuse, libcephfs: fix several caching bugs (Yan, Zheng)
- ceph-fuse: trim inodes in response to mds memory pressure (Yan, Zheng)
- ceph-kvstore-tool: expanded command set and capabilities (Joao Eduardo Luis)
- ceph.spec: fix build dependency (Loic Dachary)
- common: bloom filter improvements (Sage Weil)
- common: check preexisting admin socket for active daemon before removing (Loic Dachary)
- common: fix aligned buffer allocation (Loic Dachary)
- common: fix authentication on big-endian architectures (Dan Mick)
- common: fix config variable substitution (Loic Dachary)
- common: portability changes to support libc++ (Noah Watkins)
- common: switch to unordered_map from hash_map (Noah Watkins)
- config: recursive metavariable expansion (Loic Dachary)
- crush: default to bobtail tunables (Sage Weil)
- crush: fix off-by-one error in recent refactor (Sage Weil)
- crush: many additional tests (Loic Dachary)
- crush: misc fixes, cleanups (Loic Dachary)
- crush: new rule steps to adjust retry attempts (Sage Weil)
- crush, osd: s/rep/replicated/ for less confusion (Loic Dachary)
- crush: refactor descend_once behavior; support set_choose*_tries for replicated rules (Sage Weil)
- crush: usability and test improvements (Loic Dachary)
- debian: change directory ownership between ceph and ceph-common (Sage Weil)
- debian: integrate misc fixes from downstream packaging (James Page)
- doc: big update to install docs (John Wilkins)
- doc: many many install doc improvements (John Wilkins)
- doc: many many updates (John Wilkins)
- doc: misc fixes (David Moreau Simard, Kun Huang)
- erasure-code: improve buffer alignment (Loic Dachary)
- erasure-code: rewrite region-xor using vector operations (Andreas Peters)
- init: fix startup ordering/timeout problem with OSDs (Dmitry Smirnov)
- libcephfs: fix resource leak (Zheng Yan)
- librados: add C API coverage for atomic write operations (Christian Marie)
- librados: fix inconsistencies in API error values (David Zafman)
- librados: fix throttle leak (and eventual deadlock) (Josh Durgin)
- librados: fix watch operations with cache pools (Sage Weil)
- librados: new snap rollback operation (David Zafman)
- librados, osd: new TMAP2OMAP operation (Yan, Zheng)
- librados: read directly into user buffer (Rutger ter Borg)
- librbd: fix use-after-free aio completion bug #5426 (Josh Durgin)
- librbd: localize/distribute parent reads (Sage Weil)
- librbd: skip zeroes/holes when copying sparse images (Josh Durgin)
- mailmap: affiliation updates (Loic Dachary)
- mailmap updates (Loic Dachary)
- many portability improvements (Noah Watkins)
- many unit test improvements (Loic Dachary)

- mds: always store backtrace in default pool (Yan, Zheng)
- mds: cope with MDS failure during creation (John Spray)
- mds: fix cap migration behavior (Yan, Zheng)
- mds: fix client session flushing (Yan, Zheng)
- mds: fix crash from client sleep/resume (Zheng Yan)
- mds: fix many many multi-mds bugs (Yan, Zheng)
- mds: fix readdir end check (Zheng Yan)
- mds: fix Resetter locking (Alexandre Oliva)
- mds: fix respawn (John Spray)
- mds: inline data support (Li Wang, Yunchuan Wen)
- mds: misc bugs (Yan, Zheng)
- mds: misc fixes for directory fragments (Zheng Yan)
- mds: misc fixes for larger directories (Zheng Yan)
- mds: misc fixes for multiple MDSs (Zheng Yan)
- mds: misc multi-mds fixes (Yan, Zheng)
- mds: remove .ceph directory (John Spray)
- mds: store directories in omap instead of tmap (Yan, Zheng)
- mds: update old-format backtraces opportunistically (Zheng Yan)
- mds: use shared_ptr for requests (Greg Farnum)
- misc cleanups from coverity (Xing Lin)
- misc coverity fixes, cleanups (Danny Al-Gaaf)
- misc coverity fixes (Xing Lin, Li Wang, Danny Al-Gaaf)
- misc portability fixes (Noah Watkins, Alan Somers)
- misc portability fixes (Noah Watkins, Christophe Courtaut, Alan Somers, huanjun)
- misc portability work (Noah Watkins)
- mon: add erasure profiles and improve erasure pool creation (Loic Dachary)
- mon: add ‘mon getmap EPOCH’ (Joao Eduardo Luis)
- mon: allow adjustment of cephfs max file size via ‘ceph mds set max_file_size’ (Sage Weil)
- mon: allow debug quorum_{enter,exit} commands via admin socket
- mon: ‘ceph osd pg-temp ...’ and primary-temp commands (Ilya Dryomov)
- mon: change mds allow_new_snaps syntax to be more consistent (Sage Weil)
- mon: clean up initial crush rule creation (Loic Dachary)
- mon: collect misc metadata about osd (os, kernel, etc.), new ‘osd metadata’ command (Sage Weil)
- mon: do not create erasure rules by default (Sage Weil)
- mon: do not generate spurious MDSMaps in certain cases (Sage Weil)
- mon: do not use keyring if auth = none (Loic Dachary)
- mon: fix peer feature checks (Sage Weil)
- mon: fix pg_temp leaks (Joao Eduardo Luis)
- mon: fix pool count in ‘ceph -s’ output (Sage Weil)
- mon: handle more whitespace (newline, tab) in mon capabilities (Sage Weil)
- mon: improve (replicate or erasure) pool creation UX (Loic Dachary)
- mon: infrastructure to handle mixed-version mon cluster and cli/rest API (Greg Farnum)
- mon: MForward tests (Loic Dachary)
- mon: mkfs now idempotent (Loic Dachary)
- mon: only seed new osdmmaps to current OSDs (Sage Weil)
- mon, osd: create erasure style crush rules (Loic Dachary, Sage Weil)
- mon: ‘osd crush show-tunables’ (Sage Weil)
- mon: ‘osd dump’ dumps pool snaps as array, not object (Dan Mick)
- mon, osd: new ‘erasure’ pool type (still not fully supported)
- mon: persist quorum features to disk (Greg Farnum)
- mon: prevent extreme changes in pool pg_num (Greg Farnum)
- mon: require ‘x’ mon caps for auth operations (Joao Luis)
- mon: shutdown when removed from mon cluster (Joao Luis)
- mon: take ‘osd pool set ...’ value as an int, not string (Joao Eduardo Luis)
- mon: track osd features in OSDMap (Joao Luis, David Zafman)
- mon: trim MDSMaps (Joao Eduardo Luis)
- mon: warn if crush has non-optimal tunables (Sage Weil)
- mount.ceph: add -n for autofs support (Steve Stock)
- msgr: fix locking bug in authentication (Josh Durgin)
- msgr: fix messenger restart race (Xihui He)
- msgr: improve connection error detection between clients and monitors (Greg Farnum, Sage Weil)
- osd: add/fix CPU feature detection for jerasure (Loic Dachary)
- osd: add HitSet tracking for read ops (Sage Weil, Greg Farnum)
- osd: avoid touching leveldb for some xattrs (Haomai Wang, Sage Weil)
- osd: backfill to multiple targets (David Zafman)

- osd: backfill to osds not in acting set (David Zafman)
- osd: cache pool support for snapshots (Sage Weil)
- osd: client IO path changes for EC (Samuel Just)
- osd: default to 3x replication
- osd: do not include backfill targets in acting set (David Zafman)
- osd: enable new hashpspool layout by default (Sage Weil)
- osd: erasure plugin benchmarking tool (Loic Dachary)
- osd: fix and cleanup misc backfill issues (David Zafman)
- osd: fix bug in journal replay/restart (Sage Weil)
- osd: fix copy-get omap bug (Sage Weil)
- osd: fix linux kernel version detection (Ilya Dryomov)
- osd: fix memstore segv (Haomai Wang)
- osd: fix object_info_t encoding bug from emperor (Sam Just)
- osd: fix omap_clear operation to not zap xattrs (Sam Just, Yan, Zheng)
- osd: fix several bugs with tier infrastructure
- osd: fix throttle thread (Haomai Wang)
- osd: fix XFS detection (Greg Farnum, Sushma Gurram)
- osd: generalize scrubbing infrastructure to allow EC (David Zafman)
- osd: handle more whitespace (newline, tab) in osd capabilities (Sage Weil)
- osd: ignore num_objects_dirty on scrub for old pools (Sage Weil)
- osd: improved scrub checks on clones (Sage Weil, Sam Just)
- osd: improve locking in fd lookup cache (Samuel Just, Greg Farnum)
- osd: include more info in pg query result (Sage Weil)
- osd, librados: fix full cluster handling (Josh Durgin)
- osd: many erasure fixes (Sam Just)
- osd: many many many bug fixes with cache tiering (Samuel Just)
- osd: move to jerasure2 library (Loic Dachary)
- osd: new 'chassis' type in default crush hierarchy (Sage Weil)
- osd: new keyvaluestore-dev backend based on leveldb (Haomai Wang)
- osd: new OSDMap encoding (Greg Farnum)
- osd: new tests for erasure pools (David Zafman)
- osd: preliminary cache pool support (no snaps) (Greg Farnum, Sage Weil)
- osd: reduce scrub lock contention (Guang Yang)
- osd: requery unfound on stray notify (#6909) (Samuel Just)
- osd: some PGBackend infrastructure (Samuel Just)
- osd: support for new 'memstore' (memory-backed) backend (Sage Weil)
- osd: track erasure compatibility (David Zafman)
- osd: track omap and hit_set objects in pg stats (Samuel Just)
- osd: warn if agent cannot enable due to invalid (post-split) stats (Sage Weil)
- rados: add 'crush location', smart replica selection/balancing (Sage Weil)
- rados bench: track metadata for multiple runs separately (Guang Yang)
- rados: some performance optimizations (Yehuda Sadeh)
- rados tool: fix listomapvals (Josh Durgin)
- rbd: add 'rbdfmap' init script for mapping rbd images on book (Adam Twardowski)
- rbd: add rbdfmap support for upstart (Laurent Barbe)
- rbd: expose kernel rbd client options via 'rbd map' (Ilya Dryomov)
- rbd: fix bench-write command (Hoamai Wang)
- rbd: make 'rbd list' return empty list and success on empty pool (Josh Durgin)
- rbd: prevent deletion of images with watchers (Ilya Dryomov)
- rbd: support for 4096 mapped devices, up from ~250 (Ilya Dryomov)
- rest-api: do not fail when no OSDs yet exist (Dan Mick)
- rgw: add 'status' command to sysvinit script (David Moreau Simard)
- rgw: allow multiple frontends (Yehuda Sadeh)
- rgw: allow use of an erasure data pool (Yehuda Sadeh)
- rgw: convert bucket info to new format on demand (Yehuda Sadeh)
- rgw: fixed subuser modify (Yehuda Sadeh)
- rgw: fix error setting empty owner on ACLs (Yehuda Sadeh)
- rgw: fix fastcgi deadlock (do not return data from librados callback) (Yehuda Sadeh)
- rgw: fix many-part multipart uploads (Yehuda Sadeh)
- rgw: fix misc CORS bugs (Robin H. Johnson)
- rgw: fix object placement read op (Yehuda Sadeh)
- rgw: fix reading bucket policy (#6940)
- rgw: fix read_user_buckets 'max' behavior (Yehuda Sadeh)
- rgw: fix several CORS bugs (Robin H. Johnson)
- rgw: fix use-after-free when releasing completion handle (Yehuda Sadeh)

- rgw: improve swift temp URL support (Yehuda Sadeh)
- rgw: make multi-object delete idempotent (Yehuda Sadeh)
- rgw: optionally defer to bucket ACLs instead of object ACLs (Liam Monahan)
- rgw: prototype mongoose frontend (Yehuda Sadeh)
- rgw: several doc fixes (Alexandre Marangone)
- rgw: support for password (instead of admin token) for keystone authentication (Christophe Courtault)
- rgw: switch from mongoose to civetweb (Yehuda Sadeh)
- rgw: user quotas (Yehuda Sadeh)
- rpm: fix redhat-lsb dependency (Sage Weil, Alfredo Deza)
- specfile: fix RPM build on RHEL6 (Ken Dreyer, Derek Yarnell)
- specfile: ship libdir/ceph (Key Dreyer)
- sysvinit, upstart: prevent both init systems from starting the same daemons (Josh Durgin)

NOTABLE CHANGES SINCE V0.67 DUMPLING

- build cleanly under clang (Christophe Courtault)
- build: Makefile refactor (Roald J. van Loon)
- build: fix [/usr]/sbin locations (Alan Somers)
- ceph-disk: fix journal preallocation
- ceph-fuse, radosgw: enable admin socket and logging by default
- ceph-fuse: fix problem with readahead vs truncate race (Yan, Zheng)
- ceph-fuse: trim deleted inodes from cache (Yan, Zheng)
- ceph-fuse: use newer fuse api (Jianpeng Ma)
- ceph-kvstore-tool: new tool for working with leveldb (copy, crc) (Joao Luis)
- ceph-post-file: new command to easily share logs or other files with ceph devs
- ceph: improve parsing of CEPH_ARGS (Benoit Knecht)
- ceph: make -h behave when monitors are down
- ceph: parse CEPH_ARGS env variable
- common: bloom_filter improvements, cleanups
- common: cache crc32c values where possible
- common: correct SI is kB not KB (Dan Mick)
- common: fix looping on BSD (Alan Somers)
- common: migrate SharedPtrRegistry to use boost::shared_ptr<> (Loic Dachary)
- common: misc portability fixes (Noah Watkins)
- crc32c: fix optimized crc32c code (it now detects arch support properly)
- crc32c: improved intel-optimized crc32c support (~8x faster on my laptop!)
- crush: fix name caching
- doc: erasure coding design notes (Loic Dachary)
- hadoop: removed old version of shim to avoid confusing users (Noah Watkins)
- librados, mon: ability to query/ping out-of-quorum monitor status (Joao Luis)
- librados: fix async aio completion wakeup
- librados: fix installed header #includes (Dan Mick)
- librados: get_version64() method for C++ API
- librados: hello_world example (Greg Farnum)
- librados: sync calls now return on commit (instead of ack) (Greg Farnum)
- librbd python bindings: fix parent image name limit (Josh Durgin)
- librbd, ceph-fuse: avoid some sources of ceph-fuse, rbd cache stalls
- mds: avoid leaking objects when deleting truncated files (Yan, Zheng)
- mds: fix F_GETLK (Yan, Zheng)
- mds: fix LOOKUPSNAP bug
- mds: fix heap profiler commands (Joao Luis)
- mds: fix locking deadlock (David Disseldorp)
- mds: fix many bugs with stray (unlinked) inodes (Yan, Zheng)
- mds: fix many directory fragmentation bugs (Yan, Zheng)
- mds: fix mds rejoin with legacy parent backpointer xattrs (Alexandre Oliva)
- mds: fix rare restart/failure race during fs creation
- mds: fix standby-replay when we fall behind (Yan, Zheng)
- mds: fix stray directory purging (Yan, Zheng)
- mds: notify clients about deleted files (so they can release from their cache) (Yan, Zheng)
- mds: several bug fixes with clustered mds (Yan, Zheng)
- mon, osd: improve osdmap trimming logic (Samuel Just)
- mon, osd: initial CLI for configuring tiering
- mon: a few ‘ceph mon add’ races fixed (command is now idempotent) (Joao Luis)
- mon: allow (un)setting HASHPSPOOL flag on existing pools (Joao Luis)

- mon: allow cap strings with . to be unquoted
- mon: allow logging level of cluster log (/var/log/ceph/ceph.log) to be adjusted
- mon: avoid rewriting full osdmaps on restart (Joao Luis)
- mon: continue to discover peer addr info during election phase
- mon: disallow CephFS snapshots until ‘ceph mds set allow_new_snaps’ (Greg Farnum)
- mon: do not expose uncommitted state from ‘osd crush {add,set} ...’ (Joao Luis)
- mon: fix ‘ceph osd crush reweight ...’ (Joao Luis)
- mon: fix ‘osd crush move ...’ command for buckets (Joao Luis)
- mon: fix byte counts (off by factor of 4) (Dan Mick, Joao Luis)
- mon: fix paxos corner case
- mon: kv properties for pools to support EC (Loic Dachary)
- mon: make ‘osd pool rename’ idempotent (Joao Luis)
- mon: modify ‘auth add’ semantics to make a bit more sense (Joao Luis)
- mon: new ‘osd perf’ command to dump recent performance information (Samuel Just)
- mon: new and improved ‘ceph -s’ or ‘ceph status’ command (more info, easier to read)
- mon: some auth check cleanups (Joao Luis)
- mon: track per-pool stats (Joao Luis)
- mon: warn about pools with bad pg_num
- mon: warn when mon data stores grow very large (Joao Luis)
- monc: fix small memory leak
- new wireshark patches pulled into the tree (Kevin Jones)
- objecter, librados: redirect requests based on cache tier config
- objecter: fix possible hang when cluster is unpause (Josh Durgin)
- osd, librados: add new COPY_FROM rados operation
- osd, librados: add new COPY_GET rados operations (used by COPY_FROM)
- osd: ‘osd recover clone overlap limit’ option to limit cloning during recovery (Samuel Just)
- osd: COPY_GET on-wire encoding improvements (Greg Farnum)
- osd: add ‘osd heartbeat min healthy ratio’ configurable (was hard-coded at 33%)
- osd: add option to disable pg log debug code (which burns CPU)
- osd: allow cap strings with . to be unquoted
- osd: automatically detect proper xattr limits (David Zafman)
- osd: avoid extra copy in erasure coding reference implementation (Loic Dachary)
- osd: basic cache pool redirects (Greg Farnum)
- osd: basic whiteout, dirty flag support (not yet used)
- osd: bloom_filter encodability, fixes, cleanups (Loic Dachary, Sage Weil)
- osd: clean up and generalize copy-from code (Greg Farnum)
- osd: cls_hello OSD class example
- osd: erasure coding doc updates (Loic Dachary)
- osd: erasure coding plugin infrastructure, tests (Loic Dachary)
- osd: experimental support for ZFS (zfsonlinux.org) (Yan, Zheng)
- osd: fix RWORDER flags
- osd: fix exponential backoff of slow request warnings (Loic Dachary)
- osd: fix handling of racing read vs write (Samuel Just)
- osd: fix version value returned by various operations (Greg Farnum)
- osd: generalized temp object infrastructure
- osd: ghobject_t infrastructure for EC (David Zafman)
- osd: improvements for compatset support and storage (David Zafman)
- osd: infrastructure to copy objects from other OSDs
- osd: instrument peering states (David Zafman)
- osd: misc copy-from improvements
- osd: opportunistic crc checking on stored data (off by default)
- osd: properly enforce RD/WR flags for rados classes
- osd: reduce blocking on backing fs (Samuel Just)
- osd: refactor recovery using PGBackend (Samuel Just)
- osd: remove old magical tmap->omap conversion
- osd: remove old pg log on upgrade (Samuel Just)
- osd: revert xattr size limit (fixes large rgw uploads)
- osd: use fdatasync(2) instead of fsync(2) to improve performance (Sam Just)
- pybind: fix blacklisting nonce (Loic Dachary)
- radosgw-agent: multi-region replication/DR
- rgw: complete in-progress requests before shutting down
- rgw: default log level is now more reasonable (Yehuda Sadeh)
- rgw: fix S3 auth with response-* query string params (Sylvain Munaut, Yehuda Sadeh)
- rgw: fix a few minor memory leaks (Yehuda Sadeh)
- rgw: fix acl group check (Yehuda Sadeh)

- rgw: fix inefficient use of std::list::size() (Yehuda Sadeh)
- rgw: fix major CPU utilization bug with internal caching (Yehuda Sadeh, Mark Nelson)
- rgw: fix ordering of write operations (preventing data loss on crash) (Yehuda Sadeh)
- rgw: fix ordering of writes for multipart upload (Yehuda Sadeh)
- rgw: fix various CORS bugs (Yehuda Sadeh)
- rgw: fix/improve swift COPY support (Yehuda Sadeh)
- rgw: improve help output (Christophe Courtault)
- rgw: misc fixes to support DR (Josh Durgin, Yehuda Sadeh)
- rgw: per-bucket quota (Yehuda Sadeh)
- rgw: validate S3 tokens against keystone (Roald J. van Loon)
- rgw: wildcard support for keystone roles (Christophe Courtault)
- rpm: fix junit dependencies (Alan Grosskurth)
- sysvinit radosgw: fix status return code (Danny Al-Gaaf)
- sysvinit rbdmap: fix error 'service rbdmap stop' (Laurent Barbe)
- sysvinit: add conrestart command (Dan van der Ster)
- sysvinit: fix shutdown order (mons last) (Alfredo Deza)

V0.79

This release is intended to serve as a release candidate for firefly, which will hopefully be v0.80. No changes are being made to the code base at this point except those that fix bugs. Please test this release if you intend to make use of the new erasure-coded pools or cache tiers in firefly.

This release fixes a range of bugs found in v0.78 and streamlines the user experience when creating erasure-coded pools. There is also a raft of fixes for the MDS (multi-mds, directory fragmentation, and large directories). The main notable new piece of functionality is a small change to allow radosgw to use an erasure-coded pool for object data.

UPGRADING

- Erasure pools created with v0.78 will no longer function with v0.79. You will need to delete the old pool and create a new one.
- A bug was fixed in the authentication handshake with big-endian architectures that prevent authentication between big- and little-endian machines in the same cluster. If you have a cluster that consists entirely of big-endian machines, you will need to upgrade all daemons and clients and restart.
- The 'ceph.file.layout' and 'ceph.dir.layout' extended attributes are no longer included in the listxattr(2) results to prevent problems with 'cp -a' and similar tools.
- Monitor 'auth' read-only commands now expect the user to have 'rx' caps. This is the same behavior that was present in dumpling, but in emperor and more recent development releases the 'r' cap was sufficient. The affected commands are:

```
ceph auth export
ceph auth get
ceph auth get-key
ceph auth print-key
ceph auth list
```

NOTABLE CHANGES

- ceph-conf: stop creating bogus log files (Josh Durgin, Sage Weil)
- common: fix authentication on big-endian architectures (Dan Mick)
- debian: change directory ownership between ceph and ceph-common (Sage Weil)
- init: fix startup ordering/timeout problem with OSDs (Dmitry Smirnov)
- librbd: skip zeroes/holes when copying sparse images (Josh Durgin)
- mds: cope with MDS failure during creation (John Spray)
- mds: fix crash from client sleep/resume (Zheng Yan)
- mds: misc fixes for directory fragments (Zheng Yan)
- mds: misc fixes for larger directories (Zheng Yan)
- mds: misc fixes for multiple MDSs (Zheng Yan)
- mds: remove .ceph directory (John Spray)
- misc coverity fixes, cleanups (Danny Al-Gaaf)
- mon: add erasure profiles and improve erasure pool creation (Loic Dachary)

- mon: ‘ceph osd pg-temp ...’ and primary-temp commands (Ilya Dryomov)
- mon: fix pool count in ‘ceph -s’ output (Sage Weil)
- msgr: improve connection error detection between clients and monitors (Greg Farnum, Sage Weil)
- osd: add/fix CPU feature detection for jerasure (Loic Dachary)
- osd: improved scrub checks on clones (Sage Weil, Sam Just)
- osd: many erasure fixes (Sam Just)
- osd: move to jerasure2 library (Loic Dachary)
- osd: new tests for erasure pools (David Zafman)
- osd: reduce scrub lock contention (Guang Yang)
- rgw: allow use of an erasure data pool (Yehuda Sadeh)

V0.78

This development release includes two key features: erasure coding and cache tiering. A huge amount of code was merged for this release and several additional weeks were spent stabilizing the code base, and it is now in a state where it is ready to be tested by a broader user base.

This is *not* the firefly release. Firefly will be delayed for at least another sprint so that we can get some operational experience with the new code and do some additional testing before committing to long term support.

Note: Please note that while it is possible to create and test erasure coded pools in this release, the pools will not be usable when you upgrade to v0.79 as the OSDMap encoding will subtly change. Please do not populate your test pools with important data that can't be reloaded.

UPGRADING

- Upgrade daemons in the following order:

1. Monitors
2. OSDs
3. MDSs and/or radosgw

If the ceph-mds daemon is restarted first, it will wait until all OSDs have been upgraded before finishing its startup sequence. If the ceph-mon daemons are not restarted prior to the ceph-osd daemons, they will not correctly register their new capabilities with the cluster and new features may not be usable until they are restarted a second time.

- Upgrade radosgw daemons together. There is a subtle change in behavior for multipart uploads that prevents a multipart request that was initiated with a new radosgw from being completed by an old radosgw.
- CephFS recently added support for a new ‘backtrace’ attribute on file data objects that is used for lookup by inode number (i.e., NFS reexport and hard links), and will later be used by fsck repair. This replaces the existing anchor table mechanism that is used for hard link resolution. In order to completely phase that out, any inode that has an outdated backtrace attribute will get updated when the inode itself is modified. This will result in some extra workload after a legacy CephFS file system is upgraded.
- The per-op return code in librados’ ObjectWriteOperation interface is now filled in.
- The librados cmpxattr operation now handles xattrs containing null bytes as data rather than null-terminated strings.
- Compound operations in librados that create and then delete the same object are now explicitly disallowed (they fail with -EINVAL).
- The default leveldb cache size for the ceph-osd daemon has been increased from 4 MB to 128 MB. This will increase the memory footprint of that process but tends to increase performance of omap (key/value) objects (used for CephFS and the radosgw). If memory in your deployment is tight, you can preserve the old behavio by adding:

```
leveldb write buffer size = 0
leveldb cache size = 0
```

to your ceph.conf to get back the (leveldb) defaults.

NOTABLE CHANGES

- ceph-brag: new client and server tools (Sebastien Han, Babu Shanmugam)

- ceph-disk: use partx on RHEL or CentOS instead of partprobe (Alfredo Deza)
- ceph: fix combination of ‘tell’ and interactive mode (Joao Eduardo Luis)
- ceph-fuse: fix bugs with inline data and multiple MDSs (Zheng Yan)
- client: fix getcwd() to use new LOOKUPPARENT operation (Zheng Yan)
- common: fall back to json-pretty for admin socket (Loic Dachary)
- common: fix ‘config dump’ debug prefix (Danny Al-Gaaf)
- common: misc coverity fixes (Danny Al-Gaaf)
- common: throttler, shared_cache performance improvements, TrackedOp (Greg Farnum, Samuel Just)
- crush: fix JSON schema for dump (John Spray)
- crush: misc cleanups, tests (Loic Dachary)
- crush: new vary_r tunable (Sage Weil)
- crush: prevent invalid buckets of type 0 (Sage Weil)
- keyvaluestore: add perfcounters, misc bug fixes (Haomai Wang)
- keyvaluestore: portability improvements (Noah Watkins)
- libcephfs: API changes to better support NFS reexport via Ganesha (Matt Benjamin, Adam Emerson, Andrey Kuznetsov, Casey Bodley, David Zafman)
- librados: API documentation improvements (John Wilkins, Josh Durgin)
- librados: fix object enumeration bugs; allow iterator assignment (Josh Durgin)
- librados: streamline tests (Josh Durgin)
- librados: support for atomic read and omap operations for C API (Josh Durgin)
- librados: support for osd and mon command timeouts (Josh Durgin)
- librbd: pass allocation hints to OSD (Ilya Dryomov)
- logrotate: fix bug that prevented rotation for some daemons (Loic Dachary)
- mds: avoid duplicated discovers during recovery (Zheng Yan)
- mds: fix file lock owner checks (Zheng Yan)
- mds: fix LOOKUPPARENT, new LOOKUPNAME ops for reliable NFS reexport (Zheng Yan)
- mds: fix xattr handling on setxattr (Zheng Yan)
- mds: fix xatrs in getattr replies (Sage Weil)
- mds: force backtrace updates for old inodes on update (Zheng Yan)
- mds: several multi-mds and dirfrag bug fixes (Zheng Yan)
- mon: encode erasure stripe width in pool metadata (Loic Dachary)
- mon: erasure code crush rule creation (Loic Dachary)
- mon: erasure code plugin support (Loic Dachary)
- mon: fix bugs in initial post-mkfs quorum creation (Sage Weil)
- mon: fix error output to terminal during startup (Joao Eduardo Luis)
- mon: fix legacy CRUSH tunables warning (Sage Weil)
- mon: fix osd_epochs lower bound tracking for map trimming (Sage Weil)
- mon: fix OSDMap encoding features (Sage Weil, Aaron Ten Clay)
- mon: fix ‘pg dump’ JSON output (John Spray)
- mon: include dirty stats in ‘ceph df detail’ (Sage Weil)
- mon: list quorum member names in quorum order (Sage Weil)
- mon: prevent addition of non-empty cache tier (Sage Weil)
- mon: prevent deletion of CephFS pools (John Spray)
- mon: warn when cache tier approaches ‘full’ (Sage Weil)
- osd: allocation hint, with XFS support (Ilya Dryomov)
- osd: erasure coded pool support (Samuel Just)
- osd: fix bug causing slow/stalled recovery (#7706) (Samuel Just)
- osd: fix bugs in log merging (Samuel Just)
- osd: fix/clarify end-of-object handling on read (Loic Dachary)
- osd: fix impolite mon session backoff, reconnect behavior (Greg Farnum)
- osd: fix SnapContext cache id bug (Samuel Just)
- osd: increase default leveldb cache size and write buffer (Sage Weil, Dmitry Smirnov)
- osd: limit size of ‘osd bench ...’ arguments (Joao Eduardo Luis)
- osdmaptool: new -test-map-pgs mode (Sage Weil, Ilya Dryomov)
- osd, mon: add primary-affinity to adjust selection of primaries (Sage Weil)
- osd: new ‘status’ admin socket command (Sage Weil)
- osd: simple tiering agent (Sage Weil)
- osd: store checksums for erasure coded object stripes (Samuel Just)
- osd: tests for objectstore backends (Haomai Wang)
- osd: various refactoring and bug fixes (Samuel Just, David Zafman)
- rados: add ‘set-alloc-hint’ command (Ilya Dryomov)
- rbd-fuse: fix enumerate_images overflow, memory leak (Ilya Dryomov)
- rbdmap: fix upstart script (Stephan Renatus)
- rgw: avoid logging system events to usage log (Yehuda Sadeh)
- rgw: fix Swift range response (Yehuda Sadeh)

- rgw: improve scalability for manifest objects (Yehuda Sadeh)
- rgw: misc fixes for multipart objects, policies (Yehuda Sadeh)
- rgw: support non-standard MultipartUpload command (Yehuda Sadeh)

V0.77

This is the final development release before the Firefly feature freeze. The main items in this release include some additional refactoring work in the OSD IO path (include some locking improvements), per-user quotas for the radosgw, a switch to civetweb from mongoose for the prototype radosgw standalone mode, and a prototype leveldb-based backend for the OSD. The C librados API also got support for atomic write operations (read side transactions will appear in v0.78).

UPGRADING

- The ‘ceph -s’ or ‘ceph status’ command’s ‘num_in_osds’ field in the JSON and XML output has been changed from a string to an int.
- The recently added ‘ceph mds set allow_new_snaps’ command’s syntax has changed slightly; it is now ‘ceph mds set allow_new_snaps true’. The ‘unset’ command has been removed; instead, set the value to ‘false’.
- The syntax for allowing snapshots is now ‘mds set allow_new_snaps <true|false>’ instead of ‘mds <set unset> allow_new_snaps’.

NOTABLE CHANGES

- osd: client IO path changes for EC (Samuel Just)
- common: portability changes to support libc++ (Noah Watkins)
- common: switch to unordered_map from hash_map (Noah Watkins)
- rgw: switch from mongoose to civetweb (Yehuda Sadeh)
- osd: improve locking in fd lookup cache (Samuel Just, Greg Farnum)
- doc: many many updates (John Wilkins)
- rgw: user quotas (Yehuda Sadeh)
- mon: persist quorum features to disk (Greg Farnum)
- mon: MForward tests (Loic Dachary)
- mds: inline data support (Li Wang, Yunchuan Wen)
- rgw: fix many-part multipart uploads (Yehuda Sadeh)
- osd: new keyvaluestore-dev backend based on leveldb (Haomai Wang)
- rbd: prevent deletion of images with watchers (Ilya Dryomov)
- osd: avoid touching leveldb for some xattrs (Haomai Wang, Sage Weil)
- mailmap: affiliation updates (Loic Dachary)
- osd: new OSDMap encoding (Greg Farnum)
- osd: generalize scrubbing infrastructure to allow EC (David Zafman)
- rgw: several doc fixes (Alexandre Marangone)
- librados: add C API coverage for atomic write operations (Christian Marie)
- rgw: improve swift temp URL support (Yehuda Sadeh)
- rest-api: do not fail when no OSDs yet exist (Dan Mick)
- common: check preexisting admin socket for active daemon before removing (Loic Dachary)
- osd: handle more whitespace (newline, tab) in osd capabilities (Sage Weil)
- mon: handle more whitespace (newline, tab) in mon capabilities (Sage Weil)
- rgw: make multi-object delete idempotent (Yehuda Sadeh)
- crush: fix off-by-one error in recent refactor (Sage Weil)
- rgw: fix read_user_buckets ‘max’ behavior (Yehuda Sadeh)
- mon: change mds allow_new_snaps syntax to be more consistent (Sage Weil)

V0.76

This release includes another batch of updates for firefly functionality. Most notably, the cache pool infrastructure now support snapshots, the OSD backfill functionality has been generalized to include multiple targets (necessary for the coming erasure pools), and there were performance improvements to the erasure code plugin on capable processors. The MDS now properly utilizes (and seamlessly migrates to) the OSD key/value interface (aka omap) for storing directory objects. There continue to be many other fixes and improvements for usability and code portability across the tree.

UPGRADING

- ‘rbd ls’ on a pool which never held rbd images now exits with code 0. It outputs nothing in plain format, or an empty list in non-plain format. This is consistent with the behavior for a pool which used to hold images, but contains none. Scripts relying on this behavior should be updated.
- The MDS requires a new OSD operation TMAP2OMAP, added in this release. When upgrading, be sure to upgrade and restart the ceph-osd daemons before the ceph-mds daemon. The MDS will refuse to start if any up OSDs do not support the new feature.
- The ‘ceph mds set_max_mds N’ command is now deprecated in favor of ‘ceph mds set max_mds N’.

NOTABLE CHANGES

- build: misc improvements (Ken Dreyer)
- ceph-disk: generalize path names, add tests (Loic Dachary)
- ceph-disk: misc improvements for puppet (Loic Dachary)
- ceph-disk: several bug fixes (Loic Dachary)
- ceph-fuse: fix race for sync reads (Sage Weil)
- config: recursive metavariable expansion (Loic Dachary)
- crush: usability and test improvements (Loic Dachary)
- doc: misc fixes (David Moreau Simard, Kun Huang)
- erasure-code: improve buffer alignment (Loic Dachary)
- erasure-code: rewrite region-xor using vector operations (Andreas Peters)
- librados, osd: new TMAP2OMAP operation (Yan, Zheng)
- mailmap updates (Loic Dachary)
- many portability improvements (Noah Watkins)
- many unit test improvements (Loic Dachary)
- mds: always store backtrace in default pool (Yan, Zheng)
- mds: store directories in omap instead of tmap (Yan, Zheng)
- mon: allow adjustment of cephfs max file size via ‘ceph mds set max_file_size’ (Sage Weil)
- mon: do not create erasure rules by default (Sage Weil)
- mon: do not generate spurious MDSMaps in certain cases (Sage Weil)
- mon: do not use keyring if auth = none (Loic Dachary)
- mon: fix pg_temp leaks (Joao Eduardo Luis)
- osd: backfill to multiple targets (David Zafman)
- osd: cache pool support for snapshots (Sage Weil)
- osd: fix and cleanup misc backfill issues (David Zafman)
- osd: fix omap_clear operation to not zap xattrs (Sam Just, Yan, Zheng)
- osd: ignore num_objects_dirty on scrub for old pools (Sage Weil)
- osd: include more info in pg query result (Sage Weil)
- osd: track erasure compatibility (David Zafman)
- rbd: make ‘rbd list’ return empty list and success on empty pool (Josh Durgin)
- rgw: fix object placement read op (Yehuda Sadeh)
- rgw: fix several CORS bugs (Robin H. Johnson)
- specfile: fix RPM build on RHEL6 (Ken Dreyer, Derek Yarnell)
- specfile: ship libdir/ceph (Key Dreyer)

V0.75

This is a big release, with lots of infrastructure going in for firefly. The big items include a prototype standalone frontend for radosgw (which does not require apache or fastcgi), tracking for read activity on the osds (to inform tiering decisions), preliminary cache pool support (no snapshots yet), and lots of bug fixes and other work across the tree to get ready for the next batch of erasure coding patches.

For comparison, here are the diff stats for the last few versions:

v0.75 291 files changed, 82713 insertions(+), 33495 deletions(-)
v0.74 192 files changed, 17980 insertions(+), 1062 deletions(-)
v0.73 148 files changed, 4464 insertions(+), 2129 deletions(-)

UPGRADING

- The ‘osd pool create ...’ syntax has changed for erasure pools.
- The default CRUSH rules and layouts are now using the latest and greatest tunables and defaults. Clusters using the old values will now present with a health WARN state. This can be disabled by adding ‘mon warn on legacy crush tunables =

false' to ceph.conf.

NOTABLE CHANGES

- common: bloom filter improvements (Sage Weil)
- common: fix config variable substitution (Loic Dachary)
- crush, osd: s/rep/replicated/ for less confusion (Loic Dachary)
- crush: refactor descend_once behavior; support set_choose*_tries for replicated rules (Sage Weil)
- librados: fix throttle leak (and eventual deadlock) (Josh Durgin)
- librados: read directly into user buffer (Rutger ter Borg)
- librbd: fix use-after-free aio completion bug #5426 (Josh Durgin)
- librbd: localize/distribute parent reads (Sage Weil)
- mds: fix Resetter locking (Alexandre Oliva)
- mds: fix cap migration behavior (Yan, Zheng)
- mds: fix client session flushing (Yan, Zheng)
- mds: fix many many multi-mds bugs (Yan, Zheng)
- misc portability work (Noah Watkins)
- mon, osd: create erasure style crush rules (Loic Dachary, Sage Weil)
- mon: 'osd crush show-tunables' (Sage Weil)
- mon: clean up initial crush rule creation (Loic Dachary)
- mon: improve (replicate or erasure) pool creation UX (Loic Dachary)
- mon: infrastructure to handle mixed-version mon cluster and cli/rest API (Greg Farnum)
- mon: mkfs now idempotent (Loic Dachary)
- mon: only seed new osdmmaps to current OSDs (Sage Weil)
- mon: track osd features in OSDMap (Joao Luis, David Zafman)
- mon: warn if crush has non-optimal tunables (Sage Weil)
- mount.ceph: add -n for autofs support (Steve Stock)
- msgr: fix messenger restart race (Xihui He)
- osd, librados: fix full cluster handling (Josh Durgin)
- osd: add HitSet tracking for read ops (Sage Weil, Greg Farnum)
- osd: backfill to osds not in acting set (David Zafman)
- osd: enable new hashpspool layout by default (Sage Weil)
- osd: erasure plugin benchmarking tool (Loic Dachary)
- osd: fix XFS detection (Greg Farnum, Sushma Gurram)
- osd: fix copy-get omap bug (Sage Weil)
- osd: fix linux kernel version detection (Ilya Dryomov)
- osd: fix memstore segv (Haomai Wang)
- osd: fix several bugs with tier infrastructure
- osd: fix throttle thread (Haomai Wang)
- osd: preliminary cache pool support (no snaps) (Greg Farnum, Sage Weil)
- rados tool: fix listomapvals (Josh Durgin)
- rados: add 'crush location', smart replica selection/balancing (Sage Weil)
- rados: some performance optimizations (Yehuda Sadeh)
- rbd: add rbdmap support for upstart (Laurent Barbe)
- rbd: expose kernel rbd client options via 'rbd map' (Ilya Dryomov)
- rbd: fix bench-write command (Haomai Wang)
- rbd: support for 4096 mapped devices, up from ~250 (Ilya Dryomov)
- rgw: allow multiple frontends (Yehuda Sadeh)
- rgw: convert bucket info to new format on demand (Yehuda Sadeh)
- rgw: fix misc CORS bugs (Robin H. Johnson)
- rgw: prototype mongoose frontend (Yehuda Sadeh)

V0.74

This release includes a few substantial pieces for Firefly, including a long-overdue switch to 3x replication by default and a switch to the “new” CRUSH tunables by default (supported since bobtail). There is also a fix for a long-standing radosgw bug (stalled GET) that has already been backported to emperor and dumpling.

UPGRADING

- We now default to the ‘bobtail’ CRUSH tunable values that are first supported by Ceph clients in bobtail (v0.56) and Linux kernel version v3.9. If you plan to access a newly created Ceph cluster with an older kernel client, you should use ‘ceph

osd crush tunables legacy' to switch back to the legacy behavior. Note that making that change will likely result in some data movement in the system, so adjust the setting before populating the new cluster with data.

- We now set the HASHPSPOOL flag on newly created pools (and new clusters) by default. Support for this flag first appeared in v0.64; v0.67 Dumpling is the first major release that supports it. It is first supported by the Linux kernel version v3.9. If you plan to access a newly created Ceph cluster with an older kernel or clients (e.g., librados, librbd) from a pre-dumpling Ceph release, you should add 'osd pool default flag hashpspool = false' to the '[global]' section of your 'ceph.conf' prior to creating your monitors (e.g., after 'ceph-deploy new' but before 'ceph-deploy mon create ...').
- The configuration option 'osd pool default crush rule' is deprecated and replaced with 'osd pool default crush replicated ruleset'. 'osd pool default crush rule' takes precedence for backward compatibility and a deprecation warning is displayed when it is used.

NOTABLE CHANGES

- buffer: some zero-copy groundwork (Josh Durgin)
- ceph-disk: avoid fd0 (Loic Dachary)
- crush: default to bobtail tunables (Sage Weil)
- crush: many additional tests (Loic Dachary)
- crush: misc fixes, cleanups (Loic Dachary)
- crush: new rule steps to adjust retry attempts (Sage Weil)
- debian: integrate misc fixes from downstream packaging (James Page)
- doc: big update to install docs (John Wilkins)
- libcephfs: fix resource leak (Zheng Yan)
- misc coverity fixes (Xing Lin, Li Wang, Danny Al-Gaaf)
- misc portability fixes (Noah Watkins, Alan Somers)
- mon, osd: new 'erasure' pool type (still not fully supported)
- mon: add 'mon getmap EPOCH' (Joao Eduardo Luis)
- mon: collect misc metadata about osd (os, kernel, etc.), new 'osd metadata' command (Sage Weil)
- osd: default to 3x replication
- osd: do not include backfill targets in acting set (David Zafman)
- osd: new 'chassis' type in default crush hierarchy (Sage Weil)
- osd: requery unfound on stray notify (#6909) (Samuel Just)
- osd: some PGBackend infrastructure (Samuel Just)
- osd: support for new 'memstore' (memory-backed) backend (Sage Weil)
- rgw: fix fastcgi deadlock (do not return data from librados callback) (Yehuda Sadeh)
- rgw: fix reading bucket policy (#6940)
- rgw: fix use-after-free when releasing completion handle (Yehuda Sadeh)

V0.73

This release, the first development release after emperor, includes many bug fixes and a few additional pieces of functionality. The first batch of larger changes will be landing in the next version, v0.74.

UPGRADING

- As part of fix for #6796, 'ceph osd pool set <pool> <var> <arg>' now receives <arg> as an integer instead of a string. This affects how 'hashpspool' flag is set/unset: instead of 'true' or 'false', it now must be '0' or '1'.
- The behavior of the CRUSH 'indep' choose mode has been changed. No ceph cluster should have been using this behavior unless someone has manually extracted a crush map, modified a CRUSH rule to replace 'firstn' with 'indep', recompiled, and reinjected the new map into the cluster. If the 'indep' mode is currently in use on a cluster, the rule should be modified to use 'firstn' instead, and the administrator should wait until any data movement completes before upgrading.
- The 'osd dump' command now dumps pool snaps as an array instead of an object.
- The radosgw init script (sysvinit) now requires that the 'host = ...' line in ceph.conf, if present, match the short hostname (the output of 'hostname -s'), not the fully qualified hostname or the (occasionally non-short) output of 'hostname'. Failure to adjust this when upgrading from emperor or dumpling may prevent the radosgw daemon from starting.

NOTABLE CHANGES

- ceph-crush-location: new hook for setting CRUSH location of osd daemons on start
- ceph-kvstore-tool: expanded command set and capabilities (Joao Eduardo Luis)
- ceph.spec: fix build dependency (Loic Dachary)

- common: fix aligned buffer allocation (Loic Dachary)
- doc: many many install doc improvements (John Wilkins)
- mds: fix readdir end check (Zheng Yan)
- mds: update old-format backtraces opportunistically (Zheng Yan)
- misc cleanups from coverity (Xing Lin)
- misc portability fixes (Noah Watkins, Christophe Courtaut, Alan Somers, huanjun)
- mon: ‘osd dump’ dumps pool snaps as array, not object (Dan Mick)
- mon: allow debug quorum_{enter,exit} commands via admin socket
- mon: prevent extreme changes in pool pg_num (Greg Farnum)
- mon: take ‘osd pool set ...’ value as an int, not string (Joao Eduardo Luis)
- mon: trim MDSMaps (Joao Eduardo Luis)
- osd: fix object_info_t encoding bug from emperor (Sam Just)
- rbd: add ‘rbdmap’ init script for mapping rbd images on book (Adam Twardowski)
- rgw: add ‘status’ command to sysvinit script (David Moreau Simard)
- rgw: fix error setting empty owner on ACLs (Yehuda Sadeh)
- rgw: optionally defer to bucket ACLs instead of object ACLs (Liam Monahan)
- rgw: support for password (instead of admin token) for keystone authentication (Christophe Courtaut)
- sysvinit, upstart: prevent both init systems from starting the same daemons (Josh Durgin)

V0.72.3 EMPEROR (PENDING RELEASE)

UPGRADING

- Monitor ‘auth’ read-only commands now expect the user to have ‘rx’ caps. This is the same behavior that was present in dumpling, but in emperor and more recent development releases the ‘r’ cap was sufficient. Note that this backported security fix will break mon keys that are using the following commands but do not have the ‘x’ bit in the mon capability:

```
ceph auth export
ceph auth get
ceph auth get-key
ceph auth print-key
ceph auth list
```

V0.72.2 EMPEROR

This is the second bugfix release for the v0.72.x Emperor series. We have fixed a hang in radosgw, and fixed (again) a problem with monitor CLI compatibility with mixed version monitors. (In the future this will no longer be a problem.)

UPGRADING

- The JSON schema for the ‘osd pool set ...’ command changed slightly. Please avoid issuing this particular command via the CLI while there is a mix of v0.72.1 and v0.72.2 monitor daemons running.
- As part of fix for #6796, ‘ceph osd pool set <pool> <var> <arg>’ now receives <arg> as an integer instead of a string. This affects how ‘hashpspool’ flag is set/unset: instead of ‘true’ or ‘false’, it now must be ‘0’ or ‘1’.

CHANGES

- mon: ‘osd pool set ...’ syntax change
- osd: added test for missing on-disk HEAD object
- osd: fix osd bench block size argument
- rgw: fix hang on large object GET
- rgw: fix rare use-after-free
- rgw: various DR bug fixes
- rgw: do not return error on empty owner when setting ACL
- sysvinit, upstart: prevent starting daemons using both init systems

For more detailed information, see [the complete changelog](#).

V0.72.1 EMPEROR

IMPORTANT NOTE

When you are upgrading from Dumpling to Emperor, do not run any of the “ceph osd pool set” commands while your monitors are running separate versions. Doing so could result in inadvertently changing cluster configuration settings that exhaust compute resources in your OSDs.

CHANGES

- osd: fix upgrade bug #6761
- ceph_filestore_tool: introduced tool to repair errors caused by #6761

This release addresses issue #6761. Upgrading to Emperor can cause reads to begin returning ENFILE (too many open files). v0.72.1 fixes that upgrade issue and adds a tool `ceph_filestore_tool` to repair osd stores affected by this bug.

To repair a cluster affected by this bug:

1. Upgrade all osd machines to v0.72.1
2. Install the `ceph-test` package on each osd machine to get `ceph_filestore_tool`
3. Stop all osd processes
4. To see all lost objects, run the following on each osd with the osd stopped and the osd data directory mounted:

```
ceph_filestore_tool --list-lost-objects=true --filestore-path=<path-to-osd-filestore> --j
```

5. To fix all lost objects, run the following on each osd with the osd stopped and the osd data directory mounted:

```
ceph_filestore_tool --fix-lost-objects=true --list-lost-objects=true --filestore-path=<pa
```

6. Once lost objects have been repaired on each osd, you can restart the cluster.

Note, the `ceph_filestore_tool` performs a scan of all objects on the osd and may take some time.

V0.72 EMPEROR

This is the fifth major release of Ceph, the fourth since adopting a 3-month development cycle. This release brings several new features, including multi-datacenter replication for the radosgw, improved usability, and lands a lot of incremental performance and internal refactoring work to support upcoming features in Firefly.

IMPORTANT NOTE

When you are upgrading from Dumpling to Emperor, do not run any of the “ceph osd pool set” commands while your monitors are running separate versions. Doing so could result in inadvertently changing cluster configuration settings that exhaust compute resources in your OSDs.

HIGHLIGHTS

- common: improved crc32c performance
- librados: new example client and class code
- mds: many bug fixes and stability improvements
- mon: health warnings when pool pg_num values are not reasonable
- mon: per-pool performance stats
- osd, librados: new object copy primitives
- osd: improved interaction with backend file system to reduce latency
- osd: much internal refactoring to support ongoing erasure coding and tiering support
- rgw: bucket quotas
- rgw: improved CORS support

- rgw: performance improvements
- rgw: validate S3 tokens against Keystone

Coincident with core Ceph, the Emperor release also brings:

- radosgw-agent: support for multi-datacenter replication for disaster recovery
- tgt: improved support for iSCSI via upstream tgt

Packages for both are available on ceph.com.

UPGRADE SEQUENCING

There are no specific upgrade restrictions on the order or sequence of upgrading from 0.67.x Dumpling. However, you cannot run any of the “ceph osd pool set” commands while your monitors are running separate versions. Doing so could result in inadvertently changing cluster configuration settings and exhausting compute resources in your OSDs.

It is also possible to do a rolling upgrade from 0.61.x Cuttlefish, but there are ordering restrictions. (This is the same set of restrictions for Cuttlefish to Dumpling.)

1. Upgrade ceph-common on all nodes that will use the command line ‘ceph’ utility.
2. Upgrade all monitors (upgrade ceph package, restart ceph-mon daemons). This can happen one daemon or host at a time. Note that because cuttlefish and dumpling monitors can’t talk to each other, all monitors should be upgraded in relatively short succession to minimize the risk that an untimely failure will reduce availability.
3. Upgrade all osds (upgrade ceph package, restart ceph-osd daemons). This can happen one daemon or host at a time.
4. Upgrade radosgw (upgrade radosgw package, restart radosgw daemons).

UPGRADING FROM V0.71

- ceph-fuse and radosgw now use the same default values for the admin socket and log file paths that the other daemons (ceph-osd, ceph-mon, etc.) do. If you run these daemons as non-root, you may need to adjust your ceph.conf to disable these options or to adjust the permissions on /var/run/ceph and /var/log/ceph.

UPGRADING FROM V0.67 DUMPLING

- ceph-fuse and radosgw now use the same default values for the admin socket and log file paths that the other daemons (ceph-osd, ceph-mon, etc.) do. If you run these daemons as non-root, you may need to adjust your ceph.conf to disable these options or to adjust the permissions on /var/run/ceph and /var/log/ceph.
- The MDS now disallows snapshots by default as they are not considered stable. The command ‘ceph mds set allow_snaps’ will enable them.
- For clusters that were created before v0.44 (pre-argonaut, Spring 2012) and store radosgw data, the auto-upgrade from TMAP to OMAP objects has been disabled. Before upgrading, make sure that any buckets created on pre-argonaut releases have been modified (e.g., by PUTing and then DELETEing an object from each bucket). Any cluster created with argonaut (v0.48) or a later release or not using radosgw never relied on the automatic conversion and is not affected by this change.
- Any direct users of the ‘tmap’ portion of the librados API should be aware that the automatic tmap -> omap conversion functionality has been removed.
- Most output that used K or KB (e.g., for kilobyte) now uses a lower-case k to match the official SI convention. Any scripts that parse output and check for an upper-case K will need to be modified.
- librados::Rados::pool_create_async() and librados::Rados::pool_delete_async() don’t drop a reference to the completion object on error, caller needs to take care of that. This has never really worked correctly and we were leaking an object
- ‘ceph osd crush set <id> <weight> <loc..>’ no longer adds the osd to the specified location, as that’s a job for ‘ceph osd crush add’. It will however continue to work just the same as long as the osd already exists in the crush map.
- The OSD now enforces that class write methods cannot both mutate an object and return data. The rbd.assign_bid method, the lone offender, has been removed. This breaks compatibility with pre-bobtail librbd clients by preventing them from creating new images.
- librados now returns on commit instead of ack for synchronous calls. This is a bit safer in the case where both OSDs and the client crash, and is probably how it should have been acting from the beginning. Users are unlikely to notice but it could result in lower performance in some circumstances. Those who care should switch to using the async interfaces, which let you specify safety semantics precisely.
- The C++ librados AioComplete::get_version() method was incorrectly returning an int (usually 32-bits). To avoid breaking library compatibility, a get_version64() method is added that returns the full-width value. The old method is deprecated and will be removed in a future release. Users of the C++ librados API that make use of the get_version() method should modify their code to avoid getting a value that is truncated from 64 to 32 bits.

NOTABLE CHANGES SINCE V0.71

- build: fix [/usr]/sbin locations (Alan Somers)
- ceph-fuse, radosgw: enable admin socket and logging by default
- ceph: make -h behave when monitors are down
- common: cache crc32c values where possible
- common: fix looping on BSD (Alan Somers)
- librados, mon: ability to query/ping out-of-quorum monitor status (Joao Luis)
- librbd python bindings: fix parent image name limit (Josh Durgin)
- mds: avoid leaking objects when deleting truncated files (Yan, Zheng)
- mds: fix F_GETLK (Yan, Zheng)
- mds: fix many bugs with stray (unlinked) inodes (Yan, Zheng)
- mds: fix many directory fragmentation bugs (Yan, Zheng)
- mon: allow (un)setting HASHPSPOOL flag on existing pools (Joao Luis)
- mon: make ‘osd pool rename’ idempotent (Joao Luis)
- osd: COPY_GET on-wire encoding improvements (Greg Farnum)
- osd: bloom_filter encodability, fixes, cleanups (Loic Dachary, Sage Weil)
- osd: fix handling of racing read vs write (Samuel Just)
- osd: reduce blocking on backing fs (Samuel Just)
- radosgw-agent: multi-region replication/DR
- rgw: fix/improve swift COPY support (Yehuda Sadeh)
- rgw: misc fixes to support DR (Josh Durgin, Yehuda Sadeh)
- rgw: per-bucket quota (Yehuda Sadeh)
- rpm: fix junit dependencies (Alan Grosskurth)

NOTABLE CHANGES SINCE V0.67 DUMPLING

- build cleanly under clang (Christophe Courtault)
- build: Makefile refactor (Roald J. van Loon)
- build: fix [/usr]/sbin locations (Alan Somers)
- ceph-disk: fix journal preallocation
- ceph-fuse, radosgw: enable admin socket and logging by default
- ceph-fuse: fix problem with readahead vs truncate race (Yan, Zheng)
- ceph-fuse: trim deleted inodes from cache (Yan, Zheng)
- ceph-fuse: use newer fuse api (Jianpeng Ma)
- ceph-kvstore-tool: new tool for working with leveldb (copy, crc) (Joao Luis)
- ceph-post-file: new command to easily share logs or other files with ceph devs
- ceph: improve parsing of CEPH_ARGS (Benoit Knecht)
- ceph: make -h behave when monitors are down
- ceph: parse CEPH_ARGS env variable
- common: bloom_filter improvements, cleanups
- common: cache crc32c values where possible
- common: correct SI is kB not KB (Dan Mick)
- common: fix looping on BSD (Alan Somers)
- common: migrate SharedPtrRegistry to use boost::shared_ptr<> (Loic Dachary)
- common: misc portability fixes (Noah Watkins)
- crc32c: fix optimized crc32c code (it now detects arch support properly)
- crc32c: improved intel-optimized crc32c support (~8x faster on my laptop!)
- crush: fix name caching
- doc: erasure coding design notes (Loic Dachary)
- hadoop: removed old version of shim to avoid confusing users (Noah Watkins)
- librados, mon: ability to query/ping out-of-quorum monitor status (Joao Luis)
- librados: fix async aio completion wakeup
- librados: fix installed header #includes (Dan Mick)
- librados: get_version64() method for C++ API
- librados: hello_world example (Greg Farnum)
- librados: sync calls now return on commit (instead of ack) (Greg Farnum)
- librbd python bindings: fix parent image name limit (Josh Durgin)
- librbd, ceph-fuse: avoid some sources of ceph-fuse, rbd cache stalls
- mds: avoid leaking objects when deleting truncated files (Yan, Zheng)
- mds: fix F_GETLK (Yan, Zheng)
- mds: fix LOOKUPSNAP bug
- mds: fix heap profiler commands (Joao Luis)
- mds: fix locking deadlock (David Disseldorp)

- mds: fix many bugs with stray (unlinked) inodes (Yan, Zheng)
- mds: fix many directory fragmentation bugs (Yan, Zheng)
- mds: fix mds rejoin with legacy parent backpointer xattrs (Alexandre Oliva)
- mds: fix rare restart/failure race during fs creation
- mds: fix standby-replay when we fall behind (Yan, Zheng)
- mds: fix stray directory purging (Yan, Zheng)
- mds: notify clients about deleted files (so they can release from their cache) (Yan, Zheng)
- mds: several bug fixes with clustered mds (Yan, Zheng)
- mon, osd: improve osdmap trimming logic (Samuel Just)
- mon, osd: initial CLI for configuring tiering
- mon: a few ‘ceph mon add’ races fixed (command is now idempotent) (Joao Luis)
- mon: allow (un)setting HASHPSPOOL flag on existing pools (Joao Luis)
- mon: allow cap strings with . to be unquoted
- mon: allow logging level of cluster log (/var/log/ceph/ceph.log) to be adjusted
- mon: avoid rewriting full osdmmaps on restart (Joao Luis)
- mon: continue to discover peer addr info during election phase
- mon: disallow CephFS snapshots until ‘ceph mds set allow_new_snaps’ (Greg Farnum)
- mon: do not expose uncommitted state from ‘osd crush {add,set} ...’ (Joao Luis)
- mon: fix ‘ceph osd crush reweight ...’ (Joao Luis)
- mon: fix ‘osd crush move ...’ command for buckets (Joao Luis)
- mon: fix byte counts (off by factor of 4) (Dan Mick, Joao Luis)
- mon: fix paxos corner case
- mon: kv properties for pools to support EC (Loic Dachary)
- mon: make ‘osd pool rename’ idempotent (Joao Luis)
- mon: modify ‘auth add’ semantics to make a bit more sense (Joao Luis)
- mon: new ‘osd perf’ command to dump recent performance information (Samuel Just)
- mon: new and improved ‘ceph -s’ or ‘ceph status’ command (more info, easier to read)
- mon: some auth check cleanups (Joao Luis)
- mon: track per-pool stats (Joao Luis)
- mon: warn about pools with bad pg_num
- mon: warn when mon data stores grow very large (Joao Luis)
- monc: fix small memory leak
- new wireshark patches pulled into the tree (Kevin Jones)
- objecter, librados: redirect requests based on cache tier config
- objecter: fix possible hang when cluster is unpause (Josh Durgin)
- osd, librados: add new COPY_FROM rados operation
- osd, librados: add new COPY_GET rados operations (used by COPY_FROM)
- osd: ‘osd recover clone overlap limit’ option to limit cloning during recovery (Samuel Just)
- osd: COPY_GET on-wire encoding improvements (Greg Farnum)
- osd: add ‘osd heartbeat min healthy ratio’ configurable (was hard-coded at 33%)
- osd: add option to disable pg log debug code (which burns CPU)
- osd: allow cap strings with . to be unquoted
- osd: automatically detect proper xattr limits (David Zafman)
- osd: avoid extra copy in erasure coding reference implementation (Loic Dachary)
- osd: basic cache pool redirects (Greg Farnum)
- osd: basic whiteout, dirty flag support (not yet used)
- osd: bloom_filter encodability, fixes, cleanups (Loic Dachary, Sage Weil)
- osd: clean up and generalize copy-from code (Greg Farnum)
- osd: cls_hello OSD class example
- osd: erasure coding doc updates (Loic Dachary)
- osd: erasure coding plugin infrastructure, tests (Loic Dachary)
- osd: experimental support for ZFS (zfsonlinux.org) (Yan, Zheng)
- osd: fix RWORDER flags
- osd: fix exponential backoff of slow request warnings (Loic Dachary)
- osd: fix handling of racing read vs write (Samuel Just)
- osd: fix version value returned by various operations (Greg Farnum)
- osd: generalized temp object infrastructure
- osd: ghobject_t infrastructure for EC (David Zafman)
- osd: improvements for compatset support and storage (David Zafman)
- osd: infrastructure to copy objects from other OSDs
- osd: instrument peering states (David Zafman)
- osd: misc copy-from improvements
- osd: opportunistic crc checking on stored data (off by default)
- osd: properly enforce RD/WR flags for rados classes
- osd: reduce blocking on backing fs (Samuel Just)

- osd: refactor recovery using PGBackend (Samuel Just)
- osd: remove old magical tmap->omap conversion
- osd: remove old pg log on upgrade (Samuel Just)
- osd: revert xattr size limit (fixes large rgw uploads)
- osd: use fdatasync(2) instead of fsync(2) to improve performance (Sam Just)
- pybind: fix blacklisting nonce (Loic Dachary)
- radosgw-agent: multi-region replication/DR
- rgw: complete in-progress requests before shutting down
- rgw: default log level is now more reasonable (Yehuda Sadeh)
- rgw: fix S3 auth with response-* query string params (Sylvain Munaut, Yehuda Sadeh)
- rgw: fix a few minor memory leaks (Yehuda Sadeh)
- rgw: fix acl group check (Yehuda Sadeh)
- rgw: fix inefficient use of std::list::size() (Yehuda Sadeh)
- rgw: fix major CPU utilization bug with internal caching (Yehuda Sadeh, Mark Nelson)
- rgw: fix ordering of write operations (preventing data loss on crash) (Yehuda Sadeh)
- rgw: fix ordering of writes for multipart upload (Yehuda Sadeh)
- rgw: fix various CORS bugs (Yehuda Sadeh)
- rgw: fix/improve swift COPY support (Yehuda Sadeh)
- rgw: improve help output (Christophe Courtault)
- rgw: misc fixes to support DR (Josh Durgin, Yehuda Sadeh)
- rgw: per-bucket quota (Yehuda Sadeh)
- rgw: validate S3 tokens against keystone (Roald J. van Loon)
- rgw: wildcard support for keystone roles (Christophe Courtault)
- rpm: fix junit dependencies (Alan Grosskurth)
- sysvinit radosgw: fix status return code (Danny Al-Gaaf)
- sysvinit rbdmap: fix error 'service rbdmap stop' (Laurent Barbe)
- sysvinit: add conrestart command (Dan van der Ster)
- sysvinit: fix shutdown order (mons last) (Alfredo Deza)

V0.71

This development release includes a significant amount of new code and refactoring, as well as a lot of preliminary functionality that will be needed for erasure coding and tiering support. There are also several significant patch sets improving this with the MDS.

UPGRADING

- The MDS now disallows snapshots by default as they are not considered stable. The command ‘ceph mds set allow_snaps’ will enable them.
- For clusters that were created before v0.44 (pre-argonaut, Spring 2012) and store radosgw data, the auto-upgrade from TMAP to OMAP objects has been disabled. Before upgrading, make sure that any buckets created on pre-argonaut releases have been modified (e.g., by PUTing and then DELETEing an object from each bucket). Any cluster created with argonaut (v0.48) or a later release or not using radosgw never relied on the automatic conversion and is not affected by this change.
- Any direct users of the ‘tmap’ portion of the librados API should be aware that the automatic tmap -> omap conversion functionality has been removed.
- Most output that used K or KB (e.g., for kilobyte) now uses a lower-case k to match the official SI convention. Any scripts that parse output and check for an upper-case K will need to be modified.

NOTABLE CHANGES

- build: Makefile refactor (Roald J. van Loon)
- ceph-disk: fix journal preallocation
- ceph-fuse: trim deleted inodes from cache (Yan, Zheng)
- ceph-fuse: use newer fuse api (Jianpeng Ma)
- ceph-kvstore-tool: new tool for working with leveldb (copy, crc) (Joao Luis)
- common: bloom_filter improvements, cleanups
- common: correct SI is kB not KB (Dan Mick)
- common: misc portability fixes (Noah Watkins)
- hadoop: removed old version of shim to avoid confusing users (Noah Watkins)
- librados: fix installed header #includes (Dan Mick)
- librbd, ceph-fuse: avoid some sources of ceph-fuse, rbd cache stalls

- mds: fix LOOKUPSNAP bug
- mds: fix standby-replay when we fall behind (Yan, Zheng)
- mds: fix stray directory purging (Yan, Zheng)
- mon: disallow CephFS snapshots until ‘ceph mds set allow_new_snaps’ (Greg Farnum)
- mon, osd: improve osdmap trimming logic (Samuel Just)
- mon: kv properties for pools to support EC (Loic Dachary)
- mon: some auth check cleanups (Joao Luis)
- mon: track per-pool stats (Joao Luis)
- mon: warn about pools with bad pg_num
- osd: automatically detect proper xattr limits (David Zafman)
- osd: avoid extra copy in erasure coding reference implementation (Loic Dachary)
- osd: basic cache pool redirects (Greg Farnum)
- osd: basic whiteout, dirty flag support (not yet used)
- osd: clean up and generalize copy-from code (Greg Farnum)
- osd: erasure coding doc updates (Loic Dachary)
- osd: erasure coding plugin infrastructure, tests (Loic Dachary)
- osd: fix RWORDER flags
- osd: fix exponential backoff of slow request warnings (Loic Dachary)
- osd: generalized temp object infrastructure
- osd: gobject_t infrastructure for EC (David Zafman)
- osd: improvements for compatset support and storage (David Zafman)
- osd: misc copy-from improvements
- osd: opportunistic crc checking on stored data (off by default)
- osd: refactor recovery using PGBackend (Samuel Just)
- osd: remove old magical tmap->omap conversion
- pybind: fix blacklisting nonce (Loic Dachary)
- rgw: default log level is now more reasonable (Yehuda Sadeh)
- rgw: fix acl group check (Yehuda Sadeh)
- sysvinit: fix shutdown order (mons last) (Alfredo Deza)

V0.70

UPGRADING

- librados::Rados::pool_create_async() and librados::Rados::pool_delete_async() don’t drop a reference to the completion object on error, caller needs to take care of that. This has never really worked correctly and we were leaking an object
- ‘ceph osd crush set <id> <weight> <loc..>’ no longer adds the osd to the specified location, as that’s a job for ‘ceph osd crush add’. It will however continue to work just the same as long as the osd already exists in the crush map.

NOTABLE CHANGES

- mon: a few ‘ceph mon add’ races fixed (command is now idempotent) (Joao Luis)
- crush: fix name caching
- rgw: fix a few minor memory leaks (Yehuda Sadeh)
- ceph: improve parsing of CEPH_ARGS (Benoit Knecht)
- mon: avoid rewriting full osdmmaps on restart (Joao Luis)
- crc32c: fix optimized crc32c code (it now detects arch support properly)
- mon: fix ‘ceph osd crush reweight ...’ (Joao Luis)
- osd: revert xattr size limit (fixes large rgw uploads)
- mds: fix heap profiler commands (Joao Luis)
- rgw: fix inefficient use of std::list::size() (Yehuda Sadeh)

V0.69

UPGRADING

- The sysvinit /etc/init.d/ceph script will, by default, update the CRUSH location of an OSD when it starts. Previously, if the monitors were not available, this command would hang indefinitely. Now, that step will time out after 10 seconds and the ceph-osd daemon will not be started.
- Users of the librados C++ API should replace users of get_version() with get_version64() as the old method only returns a

32-bit value for a 64-bit field. The existing 32-bit get_version() method is now deprecated.

- The OSDs are now more picky that request payload match their declared size. A write operation across N bytes that includes M bytes of data will now be rejected. No known clients do this, but because the server-side behavior has changed it is possible that an application misusing the interface may now get errors.
- The OSD now enforces that class write methods cannot both mutate an object and return data. The rbd.assign_bid method, the lone offender, has been removed. This breaks compatibility with pre-bobtail librbd clients by preventing them from creating new images.
- librados now returns on commit instead of ack for synchronous calls. This is a bit safer in the case where both OSDs and the client crash, and is probably how it should have been acting from the beginning. Users are unlikely to notice but it could result in lower performance in some circumstances. Those who care should switch to using the async interfaces, which let you specify safety semantics precisely.
- The C++ librados AioComplete::get_version() method was incorrectly returning an int (usually 32-bits). To avoid breaking library compatibility, a get_version64() method is added that returns the full-width value. The old method is deprecated and will be removed in a future release. Users of the C++ librados API that make use of the get_version() method should modify their code to avoid getting a value that is truncated from 64 to 32 bits.

NOTABLE CHANGES

- build cleanly under clang (Christophe Courtault)
- common: migrate SharedPtrRegistry to use boost::shared_ptr<> (Loic Dachary)
- doc: erasure coding design notes (Loic Dachary)
- improved intel-optimized crc32c support (~8x faster on my laptop!)
- librados: get_version64() method for C++ API
- mds: fix locking deadlock (David Disseldorp)
- mon, osd: initial CLI for configuring tiering
- mon: allow cap strings with . to be unquoted
- mon: continue to discover peer addr info during election phase
- mon: fix ‘osd crush move ...’ command for buckets (Joao Luis)
- mon: warn when mon data stores grow very large (Joao Luis)
- objecter, librados: redirect requests based on cache tier config
- osd, librados: add new COPY_FROM rados operation
- osd, librados: add new COPY_GET rados operations (used by COPY_FROM)
- osd: add ‘osd heartbeat min healthy ratio’ configurable (was hard-coded at 33%)
- osd: add option to disable pg log debug code (which burns CPU)
- osd: allow cap strings with . to be unquoted
- osd: fix version value returned by various operations (Greg Farnum)
- osd: infrastructure to copy objects from other OSDs
- osd: use fdatasync(2) instead of fsync(2) to improve performance (Sam Just)
- rgw: fix major CPU utilization bug with internal caching (Yehuda Sadeh, Mark Nelson)
- rgw: fix ordering of write operations (preventing data loss on crash) (Yehuda Sadeh)
- rgw: fix ordering of writes for multipart upload (Yehuda Sadeh)
- rgw: fix various CORS bugs (Yehuda Sadeh)
- rgw: improve help output (Christophe Courtault)
- rgw: validate S3 tokens against keystone (Roald J. van Loon)
- rgw: wildcard support for keystone roles (Christophe Courtault)
- sysvinit radosgw: fix status return code (Danny Al-Gaaf)
- sysvinit rbdmap: fix error ‘service rbdmap stop’ (Laurent Barbe)

V0.68

UPGRADING

- ‘ceph osd crush set <id> <weight> <loc..>’ no longer adds the osd to the specified location, as that’s a job for ‘ceph osd crush add’. It will however continue to work just the same as long as the osd already exists in the crush map.
- The OSD now enforces that class write methods cannot both mutate an object and return data. The rbd.assign_bid method, the lone offender, has been removed. This breaks compatibility with pre-bobtail librbd clients by preventing them from creating new images.
- librados now returns on commit instead of ack for synchronous calls. This is a bit safer in the case where both OSDs and the client crash, and is probably how it should have been acting from the beginning. Users are unlikely to notice but it could result in lower performance in some circumstances. Those who care should switch to using the async interfaces, which let you specify safety semantics precisely.
- The C++ librados AioComplete::get_version() method was incorrectly returning an int (usually 32-bits). To avoid breaking

library compatibility, a `get_version64()` method is added that returns the full-width value. The old method is deprecated and will be removed in a future release. Users of the C++ librados API that make use of the `get_version()` method should modify their code to avoid getting a value that is truncated from 64 to 32 bits.

NOTABLE CHANGES

- ceph-fuse: fix problem with readahead vs truncate race (Yan, Zheng)
- ceph-post-file: new command to easily share logs or other files with ceph devs
- ceph: parse CEPH_ARGS env variable
- librados: fix async aio completion wakeup
- librados: hello_world example (Greg Farnum)
- librados: sync calls now return on commit (instead of ack) (Greg Farnum)
- mds: fix mds rejoin with legacy parent backpointer xattrs (Alexandre Oliva)
- mds: fix rare restart/failure race during fs creation
- mds: notify clients about deleted files (so they can release from their cache) (Yan, Zheng)
- mds: several bug fixes with clustered mds (Yan, Zheng)
- mon: allow logging level of cluster log (`/var/log/ceph/ceph.log`) to be adjusted
- mon: do not expose uncommitted state from ‘osd crush {add,set} ...’ (Joao Luis)
- mon: fix byte counts (off by factor of 4) (Dan Mick, Joao Luis)
- mon: fix paxos corner case
- mon: modify ‘auth add’ semantics to make a bit more sense (Joao Luis)
- mon: new ‘osd perf’ command to dump recent performance information (Samuel Just)
- mon: new and improved ‘ceph -s’ or ‘ceph status’ command (more info, easier to read)
- monc: fix small memory leak
- new wireshark patches pulled into the tree (Kevin Jones)
- objecter: fix possible hang when cluster is unpause (Josh Durgin)
- osd: ‘osd recover clone overlap limit’ option to limit cloning during recovery (Samuel Just)
- osd: cls_hello OSD class example
- osd: experimental support for ZFS (zfsonlinux.org) (Yan, Zheng)
- osd: instrument peering states (David Zafman)
- osd: properly enforce RD/WR flags for rados classes
- osd: remove old pg log on upgrade (Samuel Just)
- rgw: complete in-progress requests before shutting down
- rgw: fix S3 auth with response-* query string params (Sylvain Munaut, Yehuda Sadeh)
- sysvinit: add conrestart command (Dan van der Ster)

V0.67.12 “DUMPLING” (DRAFT)

This stable update for Dumpling fixes a few longstanding issues with backfill in the OSD that can lead to stalled IOs. There is also a fix for memory utilization for reads in librbd when caching is enabled, and then several other small fixes across the rest of the system.

Dumpling users who have encountered IO stalls during backfill and who do not expect to upgrade to Firefly soon should upgrade. Everyone else should upgrade to Firefly already. This is likely to be the last stable release for the 0.67.x Dumpling series.

NOTABLE CHANGES

- buffer: fix buffer rebuild alignment corner case (#6614 #6003 Loic Dachary, Samuel Just)
- ceph-disk: reprobe partitions after zap (#9665 #9721 Loic Dachary)
- ceph-disk: use partx instead of partprobe when appropriate (Loic Dachary)
- common: add \$cctid meta variable (#6228 Adam Crume)
- crush: fix get_full_location_ordered (Sage Weil)
- crush: pick ruleset id that matches rule_id (#9675 Xiaoxi Chen)
- libcephfs: fix tid wrap bug (#9869 Greg Farnum)
- libcephfs: get osd location on -1 should return EINVAL (Sage Weil)
- librados: fix race condition with C API and op timeouts (#9582 Sage Weil)
- librbd: constrain max number of in-flight read requests (#9854 Jason Dillaman)
- librbd: enforce cache size on read requests (Jason Dillaman)
- librbd: fix invalid close in image open failure path (#10030 Jason Dillaman)
- librbd: fix read hang on sparse files (Jason Dillaman)
- librbd: gracefully handle deleted/renamed pools (#10270 #10122 Jason Dillaman)

- librbd: protect list_children from invalid child pool iocxts (#10123 Jason Dillaman)
- mds: fix ctime updates from clients without dirty caps (#9514 Greg Farnum)
- mds: fix rare NULL dereference in cap update path (Greg Farnum)
- mds: fix assertion caused by system clock backwards (#11053 Yan, Zheng)
- mds: store backtrace on straydir (Yan, Zheng)
- osd: fix journal committed_thru update after replay (#6756 Samuel Just)
- osd: fix memory leak, busy loop on snap trim (#9113 Samuel Just)
- osd: fix misc peering, recovery bugs (#10168 Samuel Just)
- osd: fix purged_snap field on backfill start (#9487 Sage Weil, Samuel Just)
- osd: handle no-op write with snapshot corner case (#10262 Sage Weil, Loic Dachary)
- osd: respect RWORDERED rados flag (Sage Weil)
- osd: several backfill fixes and refactors (Samuel Just, David Zafman)
- rgw: send http status reason explicitly in fastcgi (Yehuda Sadeh)

V0.67.11 “DUMPLING”

This stable update for Dumpling fixes several important bugs that affect a small set of users.

We recommend that all Dumpling users upgrade at their convenience. If none of these issues are affecting your deployment there is no urgency.

NOTABLE CHANGES

- common: fix sending dup cluster log items (#9080 Sage Weil)
- doc: several doc updates (Alfredo Deza)
- libcephfs-java: fix build against older JNI headers (Greg Farnum)
- librados: fix crash in op timeout path (#9362 Matthias Kiefer, Sage Weil)
- librbd: fix crash using clone of flattened image (#8845 Josh Durgin)
- librbd: fix error path cleanup when failing to open image (#8912 Josh Durgin)
- mon: fix crash when adjusting pg_num before any OSDs are added (#9052 Sage Weil)
- mon: reduce log noise from paxos (Aanchal Agrawal, Sage Weil)
- osd: allow scrub and snap trim thread pool IO priority to be adjusted (Sage Weil)
- osd: fix mount/remount sync race (#9144 Sage Weil)

V0.67.10 “DUMPLING”

This stable update release for Dumpling includes primarily fixes for RGW, including several issues with bucket listings and a potential data corruption problem when multiple multi-part uploads race. There is also some throttling capability added in the OSD for scrub that can mitigate the performance impact on production clusters.

We recommend that all Dumpling users upgrade at their convenience.

NOTABLE CHANGES

- ceph-disk: partprobe before settle, fixing dm-crypt (#6966, Eric Eastman)
- librbd: add invalidate cache interface (Josh Durgin)
- librbd: close image if remove_child fails (Ilya Dryomov)
- librbd: fix potential null pointer dereference (Danny Al-Gaaf)
- librbd: improve writeback checks, performance (Haomai Wang)
- librbd: skip zeroes when copying image (#6257, Josh Durgin)
- mon: fix rule(set) check on ‘ceph pool set ... crush_ruleset ...’ (#8599, John Spray)
- mon: shut down if mon is removed from cluster (#6789, Joao Eduardo Luis)
- osd: fix filestore perf reports to mon (Sage Weil)
- osd: force any new or updated xattr into leveldb if E2BIG from XFS (#7779, Sage Weil)
- osd: lock snapdir object during write to fix race with backfill (Samuel Just)
- osd: option sleep during scrub (Sage Weil)
- osd: set io priority on scrub and snap trim threads (Sage Weil)
- osd: ‘status’ admin socket command (Sage Weil)
- rbd: tolerate missing NULL terminator on block_name_prefix (#7577, Dan Mick)
- rgw: calculate user manifest (#8169, Yehuda Sadeh)
- rgw: fix abort on chunk read error, avoid using extra memory (#8289, Yehuda Sadeh)
- rgw: fix buffer overflow on bucket instance id (#8608, Yehuda Sadeh)

- rgw: fix crash in swift CORS preflight request (#8586, Yehuda Sadeh)
- rgw: fix implicit removal of old objects on object creation (#8972, Patrycja Szablowska, Yehuda Sadeh)
- rgw: fix MaxKeys in bucket listing (Yehuda Sadeh)
- rgw: fix race with multiple updates to a single multipart object (#8269, Yehuda Sadeh)
- rgw: improve bucket listing with delimiter (Yehuda Sadeh)
- rgw: include NextMarker in bucket listing (#8858, Yehuda Sadeh)
- rgw: return error early on non-existent bucket (#7064, Yehuda Sadeh)
- rgw: set truncation flag correctly in bucket listing (Yehuda Sadeh)
- sysvinit: continue starting daemons after pre-mount error (#8554, Sage Weil)

For more detailed information, see [the complete changelog](#).

V0.67.9 "DUMPLING"

This Dumpling point release fixes several minor bugs. The most prevalent in the field is one that occasionally prevents OSDs from starting on recently created clusters.

We recommend that all Dumpling users upgrade at their convenience.

NOTABLE CHANGES

- ceph-fuse, libcephfs: client admin socket command to kick and inspect MDS sessions (#8021, Zheng Yan)
- monclient: fix failure detection during mon handshake (#8278, Sage Weil)
- mon: set tid on no-op PGStatsAck messages (#8280, Sage Weil)
- msgr: fix a rare bug with connection negotiation between OSDs (Guang Yang)
- osd: allow snap trim throttling with simple delay (#6278, Sage Weil)
- osd: check for splitting when processing recover/backfill reservations (#6565, Samuel Just)
- osd: fix backfill position tracking (#8162, Samuel Just)
- osd: fix bug in backfill stats (Samuel Just)
- osd: fix bug preventing OSD startup for infant clusters (#8162, Greg Farnum)
- osd: fix rare PG resurrection race causing an incomplete PG (#7740, Samuel Just)
- osd: only complete replicas count toward min_size (#7805, Samuel Just)
- rgw: allow setting ACLs with empty owner (#6892, Yehuda Sadeh)
- rgw: send user manifest header field (#8170, Yehuda Sadeh)

For more detailed information, see [the complete changelog](#).

V0.67.8 "DUMPLING"

This Dumpling point release fixes several non-critical issues since v0.67.7. The most notable bug fixes are an auth fix in librbd (observed as an occasional crash from KVM), an improvement in the network failure detection with the monitor, and several hard to hit OSD crashes or hangs.

We recommend that all users upgrade at their convenience.

UPGRADING

- The ‘rbd ls’ function now returns success and returns an empty when a pool does not store any rbd images. Previously it would return an ENOENT error.
- Ceph will now issue a health warning if the ‘mon osd down out interval’ config option is set to zero. This warning can be disabled by adding ‘mon warn on osd down out interval zero = false’ to ceph.conf.

NOTABLE CHANGES

- all: improve keepalive detection of failed monitor connections (#7888, Sage Weil)
- ceph-fuse, libcephfs: pin inodes during readahead, fixing rare crash (#7867, Sage Weil)
- librbd: make cache writeback a bit less aggressive (Sage Weil)
- librbd: make symlink for qemu to detect librbd in RPM (#7293, Josh Durgin)
- mon: allow ‘hashspool’ pool flag to be set and unset (Loic Dachary)
- mon: commit paxos state only after entire quorum acks, fixing rare race where prior round state is readable (#7736, Sage Weil)
- mon: make elections and timeouts a bit more robust (#7212, Sage Weil)

- mon: prevent extreme pool split operations (Greg Farnum)
- mon: wait for quorum for get_version requests to close rare pool creation race (#7997, Sage Weil)
- mon: warn on ‘mon osd down out interval = 0’ (#7784, Joao Luis)
- msgr: fix byte-order for auth challenge, fixing auth errors on big-endian clients (#7977, Dan Mick)
- msgr: fix occasional crash in authentication code (usually triggered by librbd) (#6840, Josh Durgin)
- msgr: fix rebind() race (#6992, Xihui He)
- osd: avoid timeouts during slow PG deletion (#6528, Samuel Just)
- osd: fix bug in pool listing during recovery (#6633, Samuel Just)
- osd: fix queue limits, fixing recovery stalls (#7706, Samuel Just)
- osd: fix rare peering crashes (#6722, #6910, Samuel Just)
- osd: fix rare recovery hang (#6681, Samuel Just)
- osd: improve error handling on journal errors (#7738, Sage Weil)
- osd: reduce load on the monitor from OSDMap subscriptions (Greg Farnum)
- osd: rery GetLog on peer osd startup, fixing some rare peering stalls (#6909, Samuel Just)
- osd: reset journal state on remount to fix occasional crash on OSD startup (#8019, Sage Weil)
- osd: share maps with peers more aggressively (Greg Farnum)
- rbd: make it harder to delete an rbd image that is currently in use (#7076, Ilya Drymov)
- rgw: deny writes to secondary zone by non-system users (#6678, Yehuda Sadeh)
- rgw: do’t log system requests in usage log (#6889, Yehuda Sadeh)
- rgw: fix bucket recreation (#6951, Yehuda Sadeh)
- rgw: fix Swift range response (#7099, Julien Calvet, Yehuda Sadeh)
- rgw: fix URL escaping (#8202, Yehuda Sadeh)
- rgw: fix whitespace trimming in http headers (#7543, Yehuda Sadeh)
- rgw: make multi-object deletion idempotent (#7346, Yehuda Sadeh)

For more detailed information, see [the complete changelog](#).

V0.67.7 “DUMPLING”

This Dumpling point release fixes a few critical issues in v0.67.6.

All v0.67.6 users are urgently encouraged to upgrade. We also recommend that all v0.67.5 (or older) users upgrade.

UPGRADING

- Once you have upgraded a radosgw instance or OSD to v0.67.7, you should not downgrade to a previous version.

NOTABLE CHANGES

- ceph-disk: additional unit tests
- librbd: revert caching behavior change in v0.67.6
- osd: fix problem reading xattrs due to incomplete backport in v0.67.6
- radosgw-admin: fix reading object policy

For more detailed information, see [the complete changelog](#).

V0.67.6 “DUMPLING”

This Dumpling point release contains a number of important fixed for the OSD, monitor, and radosgw. Most significantly, a change that forces large object attributes to spill over into leveldb has been backported that can prevent objects and the cluster from being damaged by large attributes (which can be induced via the radosgw). There is also a set of fixes that improves data safety and RADOS semantics when the cluster becomes full and then non-full.

We recommend that all 0.67.x Dumpling users skip this release and upgrade to v0.67.7.

UPGRADING

- The OSD has long contained a feature that allows large xattrs to spill over into the leveldb backing store in situations where not all local file systems are able to store them reliably. This option is now enabled unconditionally in order to avoid rare cases where storing large xattrs renders the object unreadable. This is known to be triggered by very large multipart objects, but could be caused by other workloads as well. Although there is some small risk that performance for certain workloads will degrade, it is more important that data be retrievable. Note that newer versions of Ceph (e.g.,

firefly) do some additional work to avoid the potential performance regression in this case, but that is current considered too complex for backport to the Dumpling stable series.

- It is very dangerous to downgrade from v0.67.6 to a prior version of Dumpling. If the old version does not have ‘filestore xattr use omap = true’ it may not be able to read all xattrs for an object and can cause undefined behavior.

NOTABLE CHANGES

- ceph-disk: misc bug fixes, particularly on RHEL (Loic Dachary, Alfredo Deza, various)
- ceph-fuse, libcephfs: fix crash from read over certain sparseness patterns (Sage Weil)
- ceph-fuse, libcephfs: fix integer overflow for sync reads racing with appends (Sage Weil)
- ceph.spec: fix udev rule when building RPM under RHEL (Derek Yarnell)
- common: fix crash from bad format from admin socket (Loic Dachary)
- librados: add optional timeouts (Josh Durgin)
- librados: do not leak budget when resending localized or redirected ops (Josh Durgin)
- librados, osd: fix and improve full cluster handling (Josh Durgin)
- librbd: fix use-after-free when updating perfcounters during image close (Josh Durgin)
- librbd: remove limit on objects in cache (Josh Durgin)
- mon: avoid on-disk full OSDMap corruption from pg_temp removal (Sage Weil)
- mon: avoid stray pg_temp entries from pool deletion race (Joao Eduardo Luis)
- mon: do not generate spurious MDSMaps from laggy daemons (Joao Eduardo Luis)
- mon: fix error code from ‘osd rm|down|out|in ...’ commands (Loic Dachary)
- mon: include all health items in summary output (John Spray)
- osd: fix occasional race/crash during startup (Sage Weil)
- osd: ignore stray OSDMap messages during init (Sage Weil)
- osd: unconditionally let xattrs overflow into leveldb (David Zafman)
- rados: fix a few error checks for the CLI (Josh Durgin)
- rgw: convert legacy bucket info objects on demand (Yehuda Sadeh)
- rgw: fix bug causing system users to lose privileges (Yehuda Sadeh)
- rgw: fix CORS bugs related to headers and case sensitivity (Robin H. Johnson)
- rgw: fix multipart object listing (Yehuda Sadeh)
- rgw: fix racing object creations (Yehuda Sadeh)
- rgw: fix racing object put and delete (Yehuda Sadeh)
- rgw: fix S3 auth when using response-* query string params (Sylvain Munaut)
- rgw: use correct secret key for POST authentication (Robin H. Johnson)

For more detailed information, see [the complete changelog](#).

V0.67.5 “DUMPLING”

This release includes a few critical bug fixes for the radosgw, including a fix for hanging operations on large objects. There are also several bug fixes for radosgw multi-site replications, and a few backported features. Also, notably, the ‘osd perf’ command (which dumps recent performance information about active OSDs) has been backported.

We recommend that all 0.67.x Dumpling users upgrade.

NOTABLE CHANGES

- ceph-fuse: fix crash in caching code
- mds: fix looping in populate_mydir()
- mds: fix standby-replay race
- mon: accept ‘osd pool set ...’ as string
- mon: backport: ‘osd perf’ command to dump recent OSD performance stats
- osd: add feature compat check for upcoming object sharding
- osd: fix osd bench block size argument
- rbd.py: increase parent name size limit
- rgw: backport: allow wildcard in supported keystone roles
- rgw: backport: improve swift COPY behavior
- rgw: backport: log and open admin socket by default
- rgw: backport: validate S3 tokens against keystone
- rgw: fix bucket removal
- rgw: fix client error code for chunked PUT failure
- rgw: fix hang on large object GET
- rgw: fix rare use-after-free

- rgw: various DR bug fixes
- sysvinit, upstart: prevent starting daemons using both init systems

For more detailed information, see [the complete changelog](#).

V0.67.4 "DUMPLING"

This point release fixes an important performance issue with radosgw, keystone authentication token caching, and CORS. All users (especially those of rgw) are encouraged to upgrade.

NOTABLE CHANGES

- crush: fix invalidation of cached names
- crushtool: do not crash on non-unique bucket ids
- mds: be more careful when decoding LogEvents
- mds: fix heap check debugging commands
- mon: avoid rebuilding old full osdmaps
- mon: fix 'ceph crush move ...'
- mon: fix 'ceph osd crush reweight ...'
- mon: fix writeout of full osdmaps during trim
- mon: limit size of transactions
- mon: prevent both unmanaged and pool snaps
- osd: disable xattr size limit (prevents upload of large rgw objects)
- osd: fix recovery op throttling
- osd: fix throttling of log messages for very slow requests
- rgw: drain pending requests before completing write
- rgw: fix CORS
- rgw: fix inefficient list::size() usage
- rgw: fix keystone token expiration
- rgw: fix minor memory leaks
- rgw: fix null termination of buffer

For more detailed information, see [the complete changelog](#).

V0.67.3 "DUMPLING"

This point release fixes a few important performance regressions with the OSD (both with CPU and disk utilization), as well as several other important but less common problems. We recommend that all production users upgrade.

NOTABLE CHANGES

- ceph-disk: partprobe after creation journal partition
- ceph-disk: specify fs type when mounting
- ceph-post-file: new utility to help share logs and other files with ceph developers
- libcephfs: fix truncate vs readahead race (crash)
- mds: fix flock/fcntl lock deadlock
- mds: fix rejoin loop when encountering pre-dumpling backpointers
- mon: allow name and addr discovery during election stage
- mon: always refresh after Paxos store_state (fixes recovery corner case)
- mon: fix off-by-4x bug with osd byte counts
- osd: add and disable 'pg log keys debug' by default
- osd: add option to disable throttling
- osd: avoid leveldb iterators for pg log append and trim
- osd: fix readdir_r invocations
- osd: use fdatasync instead of sync
- radosgw: fix sysvinit script return status
- rbd: relicense as LGPL2
- rgw: flush pending data on multipart upload
- rgw: recheck object name during S3 POST
- rgw: reorder init/startup
- rpm: fix debuginfo package build

V0.67.2 “DUMPLING”

This is an important point release for Dumpling. Most notably, it fixes a problem when upgrading directly from v0.56.x Bobtail to v0.67.x Dumpling (without stopping at v0.61.x Cuttlefish along the way). It also fixes a problem with the CLI parsing of the CEPH_ARGS environment variable, high CPU utilization by the ceph-osd daemons, and cleans up the radosgw shutdown sequence.

NOTABLE CHANGES

- objecter: resend linger requests when cluster goes from full to non-full
- ceph: parse CEPH_ARGS environment variable
- librados: fix small memory leak
- osd: remove old log objects on upgrade (fixes bobtail -> dumpling jump)
- osd: disable PGLog::check() via config option (fixes CPU burn)
- rgw: drain requests on shutdown
- rgw: misc memory leaks on shutdown

For more detailed information, see [the complete changelog](#).

V0.67.1 “DUMPLING”

This is a minor point release for Dumpling that fixes problems with OpenStack and librbd hangs when caching is disabled.

NOTABLE CHANGES

- librados, librbd: fix constructor for python bindings with certain usages (in particular, that used by OpenStack)
- librados, librbd: fix aio_flush wakeup when cache is disabled
- librados: fix locking for aio completion refcounting
- fixes ‘ceph -admin-daemon ...’ command error code on error
- fixes ‘ceph daemon ... config set ...’ command for boolean config options.

For more detailed information, see [the complete changelog](#).

V0.67 “DUMPLING”

This is the fourth major release of Ceph, code-named “Dumpling.” The headline features for this release include:

- Multi-site support for radosgw. This includes the ability to set up separate “regions” in the same or different Ceph clusters that share a single S3/Swift bucket/container namespace.
- RESTful API endpoint for Ceph cluster administration. `ceph-rest-api`, a wrapper around `ceph_rest_api.py`, can be used to start up a test single-threaded HTTP server that provides access to cluster information and administration in very similar ways to the `ceph` commandline tool. `ceph_rest_api.py` can be used as a WSGI application for deployment in a more-capable web server. See `ceph-rest-api.8` for more.
- Object namespaces in librados.

UPGRADE SEQUENCING

It is possible to do a rolling upgrade from Cuttlefish to Dumpling.

1. Upgrade `ceph-common` on all nodes that will use the command line ‘`ceph`’ utility.
2. Upgrade all monitors (upgrade `ceph` package, restart `ceph-mon` daemons). This can happen one daemon or host at a time. Note that because cuttlefish and dumpling monitors can’t talk to each other, all monitors should be upgraded in relatively short succession to minimize the risk that an untimely failure will reduce availability.
3. Upgrade all osds (upgrade `ceph` package, restart `ceph-osd` daemons). This can happen one daemon or host at a time.
4. Upgrade radosgw (upgrade `radosgw` package, restart `radosgw` daemons).

UPGRADING FROM V0.66

- There is monitor internal protocol change, which means that v0.67 ceph-mon daemons cannot talk to v0.66 or older daemons. We recommend upgrading all monitors at once (or in relatively quick succession) to minimize the possibility of downtime.
- The output of ‘ceph status -format=json’ or ‘ceph -s -format=json’ has changed to return status information in a more structured and usable format.
- The ‘ceph pg dump_stuck [threshold]’ command used to require a -threshold or -t prefix to the threshold argument, but now does not.
- Many more ceph commands now output formatted information; select with ‘-format=<format>’, where <format> can be ‘json’, ‘json-pretty’, ‘xml’, or ‘xml-pretty’.
- The ‘ceph pg <pgid> ...’ commands (like ‘ceph pg <pgid> query’) are deprecated in favor of ‘ceph tell <pgid> ...’. This makes the distinction between ‘ceph pg <command> <pgid>’ and ‘ceph pg <pgid> <command>’ less awkward by making it clearer that the ‘tell’ commands are talking to the OSD serving the placement group, not the monitor.
- The ‘ceph -admin-daemon <path> <command ...>’ used to accept the command and arguments as either a single string or as separate arguments. It will now only accept the command spread across multiple arguments. This means that any script which does something like:

```
ceph --admin-daemon /var/run/ceph/ceph-osd.0.asok 'config set debug_ms 1'
```

needs to remove the quotes. Also, note that the above can now be shortened to:

```
ceph daemon osd.0 config set debug_ms 1
```

- The radosgw caps were inconsistently documented to be either ‘mon = allow r’ or ‘mon = allow rw’. The ‘mon = allow rw’ is required for radosgw to create its own pools. All documentation has been updated accordingly.
- The radosgw copy object operation may return extra progress info during the operation. At this point it will only happen when doing cross zone copy operations. The S3 response will now return extra <Progress> field under the <CopyResult> container. The Swift response will now send the progress as a json array.
- In v0.66 and v0.65 the HASHPSPOOL pool flag was enabled by default for new pools, but has been disabled again until Linux kernel client support reaches more distributions and users.
- ceph-osd now requires a max file descriptor limit (e.g., ulimit -n ...) of at least filestore_wbthrottle_(xfs|btrfs)_inodes_hard_limit (5000 by default) in order to accomodate the new write back throttle system. On Ubuntu, upstart now sets the fd limit to 32k. On other platforms, the sysvinit script will set it to 32k by default (still overrideable via max_open_files). If this field has been customized in ceph.conf it should likely be adjusted upwards.

UPGRADING FROM V0.61 “CUTTLEFISH”

In addition to the above notes about upgrading from v0.66:

- There has been a huge revamp of the ‘ceph’ command-line interface implementation. The ceph-common client library needs to be upgrade before ceph-mon is restarted in order to avoid problems using the CLI (the old ceph client utility cannot talk to the new ceph-mon).
- The CLI is now very careful about sending the ‘status’ one-liner output to stderr and command output to stdout. Scripts relying on output should take care.
- The ‘ceph osd tell ...’ and ‘ceph mon tell ...’ commands are no longer supported. Any callers should use:

```
ceph tell osd.<id or *> ...
ceph tell mon.<id or name or *> ...
```

The ‘ceph mds tell ...’ command is still there, but will soon also transition to ‘ceph tell mds.<id or name or *> ...’

- The ‘ceph osd crush add ...’ command used to take one of two forms:

```
ceph osd crush add 123 osd.123 <weight> <location ...>
ceph osd crush add osd.123 <weight> <location ...>
```

This is because the id and crush name are redundant. Now only the simple form is supported, where the osd name/id can either be a bare id (integer) or name (osd.<id>):

```
ceph osd crush add osd.123 <weight> <location ...>
ceph osd crush add 123 <weight> <location ...>
```

- There is now a maximum RADOS object size, configurable via ‘osd max object size’, defaulting to 100 GB. Note that this has no effect on RBD, CephFS, or radosgw, which all stripe over objects. If you are using librados and storing objects larger than that, you will need to adjust ‘osd max object size’, and should consider using smaller objects instead.
- The ‘osd min down {reporters|reports}’ config options have been renamed to ‘mon osd min down {reporters|reports}’, and the documentation has been updated to reflect that these options apply to the monitors (who process failure reports) and not OSDs. If you have adjusted these settings, please update your ceph.conf accordingly.

NOTABLE CHANGES SINCE V0.66

- mon: sync improvements (performance and robustness)
- mon: many bug fixes (paxos and services)
- mon: fixed bugs in recovery and io rate reporting (negative/large values)
- mon: collect metadata on osd performance
- mon: generate health warnings from slow or stuck requests
- mon: expanded -format=<json|xml|...> support for monitor commands
- mon: scrub function for verifying data integrity
- mon, osd: fix old osdmap trimming logic
- mon: enable leveldb caching by default
- mon: more efficient storage of PG metadata
- ceph-rest-api: RESTful endpoint for administer cluster (mirrors CLI)
- rgw: multi-region support
- rgw: infrastructure to support georeplication of bucket and user metadata
- rgw: infrastructure to support georeplication of bucket data
- rgw: COPY object support between regions
- rbd: /etc/ceph/rbdmap file for mapping rbd images on startup
- osd: many bug fixes
- osd: limit number of incremental osdmaps sent to peers (could cause osds to be wrongly marked down)
- osd: more efficient small object recovery
- osd, librados: support for object namespaces
- osd: automatically enable xattrs on leveldb as necessary
- mds: fix bug in LOOKUPINO (used by nfs reexport)
- mds: fix O_TRUNC locking
- msgr: fixed race condition in inter-osd network communication
- msgr: fixed various memory leaks related to network sessions
- ceph-disk: fixes for unusual device names, partition detection
- hypertable: fixes for hypertable CephBroker bindings
- use SSE4.2 crc32c instruction if present

NOTABLE CHANGES SINCE V0.61 “CUTTLEFISH”

- add ‘config get’ admin socket command
- ceph-conf: -show-config-value now reflects daemon defaults
- ceph-disk: add ‘[un]suppress-active DEV’ command
- ceph-disk: avoid mounting over an existing osd in /var/lib/ceph/osd/*
- ceph-disk: fixes for unusual device names, partition detection
- ceph-disk: improved handling of odd device names
- ceph-disk: many fixes for RHEL/CentOS, Fedora, wheezy
- ceph-disk: simpler, more robust locking
- ceph-fuse, libcephfs: fix a few caps revocation bugs
- ceph-fuse, libcephfs: fix read zeroing at EOF
- ceph-fuse, libcephfs: fix request refcounting bug (hang on shutdown)
- ceph-fuse, libcephfs: fix truncatation bug on >4MB files (Yan, Zheng)
- ceph-fuse, libcephfs: fix for cap release/hang
- ceph-fuse: add ioctl support
- ceph-fuse: fixed long-standing O_NOATIME vs O_LAZY bug
- ceph-rest-api: RESTful endpoint for administer cluster (mirrors CLI)

- ceph, librados: fix resending of commands on mon reconnect
- daemons: create /var/run/ceph as needed
- debian wheezy: fix udev rules
- debian, specfile: packaging cleanups
- debian: fix upstart behavior with upgrades
- debian: rgw: stop daemon on uninstall
- debian: stop daemons on uninstall; fix dependencies
- hypertable: fixes for hypertable CephBroker bindings
- librados python binding cleanups
- librados python: fix xattrs > 4KB (Josh Durgin)
- librados: configurable max object size (default 100 GB)
- librados: new calls to administer the cluster
- librbd: ability to read from local replicas
- librbd: locking tests (Josh Durgin)
- librbd: make default options/features for newly created images (e.g., via qemu-img) configurable
- librbd: parallelize delete, rollback, flatten, copy, resize
- many many fixes from static code analysis (Danny Al-Gaaf)
- mds: fix O_TRUNC locking
- mds: fix bug in LOOKUPINO (used by nfs reexport)
- mds: fix rare hang after client restart
- mds: fix several bugs (Yan, Zheng)
- mds: many backpointer improvements (Yan, Zheng)
- mds: many fixes for mds clustering
- mds: misc stability fixes (Yan, Zheng, Greg Farnum)
- mds: new robust open-by-ino support (Yan, Zheng)
- mds: support robust lookup by ino number (good for NFS) (Yan, Zheng)
- mon, ceph: huge revamp of CLI and internal admin API. (Dan Mick)
- mon, osd: fix old osdmap trimming logic
- mon, osd: many memory leaks fixed
- mon: better trim/compaction behavior
- mon: collect metadata on osd performance
- mon: enable leveldb caching by default
- mon: expanded -format=<json|xml|...> support for monitor commands
- mon: fix election timeout
- mon: fix leveldb compression, trimming
- mon: fix start fork behavior
- mon: fix units in ‘ceph df’ output
- mon: fix validation of mds ids from CLI commands
- mon: fixed bugs in recovery and io rate reporting (negative/large values)
- mon: generate health warnings from slow or stuck requests
- mon: many bug fixes (paxos and services, sync)
- mon: many stability fixes (Joao Luis)
- mon: more efficient storage of PG metadata
- mon: new -extract-monmap to aid disaster recovery
- mon: new capability syntax
- mon: scrub function for verifying data integrity
- mon: simplify PaxosService vs Paxos interaction, fix readable/writeable checks
- mon: sync improvements (performance and robustness)
- mon: tuning, performance improvements
- msgr: fix various memory leaks
- msgr: fixed race condition in inter-osd network communication
- msgr: fixed various memory leaks related to network sessions
- osd, librados: support for object namespaces
- osd, mon: optionally dump leveldb transactions to a log
- osd: automatically enable xattrs on leveldb as necessary
- osd: avoid osd flapping from asymmetric network failure
- osd: break blacklisted client watches (David Zafman)
- osd: close narrow journal race
- osd: do not use fadvise(DONTNEED) on XFS (data corruption on power cycle)
- osd: fix for an op ordering bug
- osd: fix handling for split after upgrade from bobtail
- osd: fix incorrect mark-down of osds
- osd: fix internal heartbeat timeouts when scrubbing very large objects
- osd: fix memory/network inefficiency during deep scrub
- osd: fixed problem with front-side heartbeats and mixed clusters (David Zafman)

- osd: limit number of incremental osdmaps sent to peers (could cause osds to be wrongly marked down)
- osd: many bug fixes
- osd: monitor both front and back interfaces
- osd: more efficient small object recovery
- osd: new writeback throttling (for less bursty write performance) (Sam Just)
- osd: pg log (re)writes are now vastly more efficient (faster peering) (Sam Just)
- osd: ping/heartbeat on public and private interfaces
- osd: prioritize recovery for degraded PGs
- osd: re-use partially deleted PG contents when present (Sam Just)
- osd: recovery and peering performance improvements
- osd: resurrect partially deleted PGs
- osd: verify both front and back network are working before rejoining cluster
- rados: clonedata command for cli
- radosgw-admin: create keys for new users by default
- rbd: /etc/ceph/rbdmap file for mapping rbd images on startup
- rgw: COPY object support between regions
- rgw: fix CORS bugs
- rgw: fix locking issue, user operation mask,
- rgw: fix radosgw-admin buckets list (Yehuda Sadeh)
- rgw: fix usage log scanning for large, untrimmed logs
- rgw: handle deep uri resources
- rgw: infrastructure to support georeplication of bucket and user metadata
- rgw: infrastructure to support georeplication of bucket data
- rgw: multi-region support
- sysvinit: fix enumeration of local daemons
- sysvinit: fix osd crush weight calculation when using -a
- sysvinit: handle symlinks in /var/lib/ceph/osd/*
- use SSE4.2 crc32c instruction if present

V0.66

UPGRADING

- There is now a configurable maximum rados object size, defaulting to 100 GB. If you are using librados and storing objects larger than that, you will need to adjust 'osd max object size', and should consider using smaller objects instead.

NOTABLE CHANGES

- osd: pg log (re)writes are now vastly more efficient (faster peering) (Sam Just)
- osd: fixed problem with front-side heartbeats and mixed clusters (David Zafman)
- mon: tuning, performance improvements
- mon: simplify PaxosService vs Paxos interaction, fix readable/writeable checks
- rgw: fix radosgw-admin buckets list (Yehuda Sadeh)
- mds: support robust lookup by ino number (good for NFS) (Yan, Zheng)
- mds: fix several bugs (Yan, Zheng)
- ceph-fuse, libcephfs: fix truncation bug on >4MB files (Yan, Zheng)
- ceph/librados: fix resending of commands on mon reconnect
- librados python: fix xattrs > 4KB (Josh Durgin)
- librados: configurable max object size (default 100 GB)
- msgr: fix various memory leaks
- ceph-fuse: fixed long-standing O_NOATIME vs O_LAZY bug
- ceph-fuse, libcephfs: fix request refcounting bug (hang on shutdown)
- ceph-fuse, libcephfs: fix read zeroing at EOF
- ceph-conf: -show-config-value now reflects daemon defaults
- ceph-disk: simpler, more robust locking
- ceph-disk: avoid mounting over an existing osd in /var/lib/ceph/osd/*
- sysvinit: handle symlinks in /var/lib/ceph/osd/*

V0.65

UPGRADING

- Huge revamp of the ‘ceph’ command-line interface implementation. The ceph-common client library needs to be upgraded before ceph-mon is restarted in order to avoid problems using the CLI (the old ceph client utility cannot talk to the new ceph-mon).
- The CLI is now very careful about sending the ‘status’ one-liner output to stderr and command output to stdout. Scripts relying on output should take care.
- The ‘ceph osd tell ...’ and ‘ceph mon tell ...’ commands are no longer supported. Any callers should use:

```
ceph tell osd.<id or *> ...
ceph tell mon.<id or name or *> ...
```

The ‘ceph mds tell ...’ command is still there, but will soon also transition to ‘ceph tell mds.<id or name or *> ...’

- The ‘ceph osd crush add ...’ command used to take one of two forms:

```
ceph osd crush add 123 osd.123 <weight> <location ...>
ceph osd crush add osd.123 <weight> <location ...>
```

This is because the id and crush name are redundant. Now only the simple form is supported, where the osd name/id can either be a bare id (integer) or name (osd.<id>):

```
ceph osd crush add osd.123 <weight> <location ...>
ceph osd crush add 123 <weight> <location ...>
```

- There is now a maximum RADOS object size, configurable via ‘osd max object size’, defaulting to 100 GB. Note that this has no effect on RBD, CephFS, or radosgw, which all stripe over objects.

NOTABLE CHANGES

- mon, ceph: huge revamp of CLI and internal admin API. (Dan Mick)
- mon: new capability syntax
- osd: do not use fadvise(DONTNEED) on XFS (data corruption on power cycle)
- osd: recovery and peering performance improvements
- osd: new writeback throttling (for less bursty write performance) (Sam Just)
- osd: ping/heartbeat on public and private interfaces
- osd: avoid osd flapping from asymmetric network failure
- osd: re-use partially deleted PG contents when present (Sam Just)
- osd: break blacklisted client watches (David Zafman)
- mon: many stability fixes (Joao Luis)
- mon, osd: many memory leaks fixed
- mds: misc stability fixes (Yan, Zheng, Greg Farnum)
- mds: many backpointer improvements (Yan, Zheng)
- mds: new robust open-by-ino support (Yan, Zheng)
- ceph-fuse, libcephfs: fix a few caps revocation bugs
- librados: new calls to administer the cluster
- librbd: locking tests (Josh Durgin)
- ceph-disk: improved handling of odd device names
- ceph-disk: many fixes for RHEL/CentOS, Fedora, wheezy
- many many fixes from static code analysis (Danny Al-Gaaf)
- daemons: create /var/run/ceph as needed

V0.64

UPGRADING

- New pools now have the HASHPSPOOL flag set by default to provide better distribution over OSDs. Support for this feature was introduced in v0.59 and Linux kernel version v3.9. If you wish to access the cluster from an older kernel, set

the ‘osd pool default flag hashpspool = false’ option in your ceph.conf prior to creating the cluster or creating new pools. Note that the presence of any pool in the cluster with the flag enabled will make the OSD require support from all clients.

NOTABLE CHANGES

- osd: monitor both front and back interfaces
- osd: verify both front and back network are working before rejoining cluster
- osd: fix memory/network inefficiency during deep scrub
- osd: fix incorrect mark-down of osds
- mon: fix start fork behavior
- mon: fix election timeout
- mon: better trim/compaction behavior
- mon: fix units in ‘ceph df’ output
- mon, osd: misc memory leaks
- librbd: make default options/features for newly created images (e.g., via qemu-img) configurable
- mds: many fixes for mds clustering
- mds: fix rare hang after client restart
- ceph-fuse: add ioctl support
- ceph-fuse/libcephfs: fix for cap release/hang
- rgw: handle deep uri resources
- rgw: fix CORS bugs
- ceph-disk: add ‘[un]suppress-active DEV’ command
- debian: rgw: stop daemon on uninstall
- debian: fix upstart behavior with upgrades

V0.63

UPGRADING

- The ‘osd min down {reporters|reports}’ config options have been renamed to ‘mon osd min down {reporters|reports}’, and the documentation has been updated to reflect that these options apply to the monitors (who process failure reports) and not OSDs. If you have adjusted these settings, please update your ceph.conf accordingly.

NOTABLE CHANGES

- librbd: parallelize delete, rollback, flatten, copy, resize
- librbd: ability to read from local replicas
- osd: resurrect partially deleted PGs
- osd: prioritize recovery for degraded PGs
- osd: fix internal heartbeat timeouts when scrubbing very large objects
- osd: close narrow journal race
- rgw: fix usage log scanning for large, untrimmed logs
- rgw: fix locking issue, user operation mask,
- initscript: fix osd crush weight calculation when using -a
- initscript: fix enumeration of local daemons
- mon: several fixes to paxos, sync
- mon: new -extract-monmap to aid disaster recovery
- mon: fix leveldb compression, trimming
- add ‘config get’ admin socket command
- rados: clonedata command for cli
- debian: stop daemons on uninstall; fix dependencies
- debian wheezy: fix udev rules
- many many small fixes from coverity scan

V0.62

NOTABLE CHANGES

- mon: fix validation of mds ids from CLI commands

- osd: fix for an op ordering bug
- osd, mon: optionally dump leveldb transactions to a log
- osd: fix handling for split after upgrade from bobtail
- debian, specfile: packaging cleanups
- radosgw-admin: create keys for new users by default
- librados python binding cleanups
- misc code cleanups

V0.61.9 "CUTTLEFISH"

This point release resolves several low to medium-impact bugs across the code base, and fixes a performance problem (CPU utilization) with radosgw. We recommend that all production cuttlefish users upgrade.

NOTABLE CHANGES

- ceph, ceph-authtool: fix help (Danny Al-Gaaf)
- ceph-disk: partprobe after creating journal partition
- ceph-disk: specific fs type when mounting (Alfredo Deza)
- ceph-fuse: fix bug when compiled against old versions
- ceph-fuse: fix use-after-free in caching code (Yan, Zheng)
- ceph-fuse: misc caching bugs
- ceph.spec: remove incorrect mod_fcgi dependency (Gary Lowell)
- crush: fix name caching
- librbd: fix bug when unpausing cluster (Josh Durgin)
- mds: fix LAZYIO lock hang
- mds: fix bug in file size recovery (after client crash)
- mon: fix paxos recovery corner case
- osd: fix exponential backoff for slow request warnings (Loic Dachary)
- osd: fix readdir_r usage
- osd: fix startup for long-stopped OSDs
- rgw: avoid std::list::size() to avoid wasting CPU cycles (Yehuda Sadeh)
- rgw: drain pending requests during write (fixes data safety issue) (Yehuda Sadeh)
- rgw: fix authenticated users group ACL check (Yehuda Sadeh)
- rgw: fix bug in POST (Yehuda Sadeh)
- rgw: fix sysvinit script 'status' command, return value (Danny Al-Gaaf)
- rgw: reduce default log level (Yehuda Sadeh)

For more detailed information, see [the complete changelog](#).

V0.61.8 "CUTTLEFISH"

This release includes a number of important issues, including rare race conditions in the OSD, a few monitor bugs, and fixes for RBD flush behavior. We recommend that production users upgrade at their convenience.

NOTABLE CHANGES

- librados: fix async aio completion wakeup
- librados: fix aio completion locking
- librados: fix rare deadlock during shutdown
- osd: fix race when queueing recovery operations
- osd: fix possible race during recovery
- osd: optionally preload rados classes on startup (disabled by default)
- osd: fix journal replay corner condition
- osd: limit size of peering work queue batch (to speed up peering)
- mon: fix paxos recovery corner case
- mon: fix rare hang when monmap updates during an election
- mon: make 'osd pool mksnap ...' avoid exposing uncommitted state
- mon: make 'osd pool rmsnap ...' not racy, avoid exposing uncommitted state
- mon: fix bug during mon cluster expansion
- rgw: fix crash during multi delete operation
- msgr: fix race conditions during osd network reinitialization

- ceph-disk: apply mount options when remounting

For more detailed information, see [the complete changelog](#).

V0.61.7 "CUTTLEFISH"

This release fixes another regression preventing monitors to start after undergoing certain upgrade sequences, as well as some corner cases with Paxos and support for unusual device names in ceph-disk/ceph-deploy.

NOTABLE CHANGES

- mon: fix regression in latest full osdmap retrieval
- mon: fix a long-standing bug in a paxos corner case
- ceph-disk: improved support for unusual device names (e.g., /dev/cciss/c0d0)

For more detailed information, see [the complete changelog](#).

V0.61.6 "CUTTLEFISH"

This release fixes a regression in v0.61.5 that could prevent monitors from restarting. This affects any cluster that was upgraded from a previous version of Ceph (and not freshly created with v0.61.5).

All users are strongly recommended to upgrade.

NOTABLE CHANGES

- mon: record latest full osdmap
- mon: work around previous bug in which latest full osdmap is not recorded
- mon: avoid scrub while updating

For more detailed information, see [the complete changelog](#).

V0.61.5 "CUTTLEFISH"

This release most improves stability of the monitor and fixes a few bugs with the ceph-disk utility (used by ceph-deploy). We recommend that all v0.61.x users upgrade.

UPGRADING

- This release fixes a 32-bit vs 64-bit arithmetic bug with the feature bits. An unfortunate consequence of the fix is that 0.61.4 (or earlier) ceph-mon daemons can't form a quorum with 0.61.5 (or later) monitors. To avoid the possibility of service disruption, we recommend you upgrade all monitors at once.

NOTABLE CHANGES

- mon: misc sync improvements (faster, more reliable, better tuning)
- mon: enable leveldb cache by default (big performance improvement)
- mon: new scrub feature (primarily for diagnostic, testing purposes)
- mon: fix occasional leveldb assertion on startup
- mon: prevent reads until initial state is committed
- mon: improved logic for trimming old osdmmaps
- mon: fix pick_addresses bug when expanding mon cluster
- mon: several small paxos fixes, improvements
- mon: fix bug osdmap trim behavior
- osd: fix several bugs with PG stat reporting
- osd: limit number of maps shared with peers (which could cause domino failures)
- rgw: fix radosgw-admin buckets list (for all buckets)
- mds: fix occasional client failure to reconnect
- mds: fix bad list traversal after unlink
- mds: fix underwater dentry cleanup (occasional crash after mds restart)

- libcephfs, ceph-fuse: fix occasional hangs on umount
- libcephfs, ceph-fuse: fix old bug with O_LAZY vs O_NOATIME confusion
- ceph-disk: more robust journal device detection on RHEL/CentOS
- ceph-disk: better, simpler locking
- ceph-disk: do not inadvertently mount over existing osd mounts
- ceph-disk: better handling for unusual device names
- sysvinit, upstart: handle symlinks in /var/lib/ceph/*

For more detailed information, see [the complete changelog](#).

V0.61.4 "CUTTLEFISH"

This release resolves a possible data corruption on power-cycle when using XFS, a few outstanding problems with monitor sync, several problems with ceph-disk and ceph-deploy operation, and a problem with OSD memory usage during scrub.

UPGRADING

- No issues.

NOTABLE CHANGES

- mon: fix daemon exit behavior when error is encountered on startup
- mon: more robust sync behavior
- osd: do not use sync_file_range(2), posix_fadvise(...DONTNEED) (can cause data corruption on power loss on XFS)
- osd: avoid unnecessary log rewrite (improves peering speed)
- osd: fix scrub efficiency bug (problematic on old clusters)
- rgw: fix listing objects that start with underscore
- rgw: fix deep URI resource, CORS bugs
- librados python binding: fix truncate on 32-bit architectures
- ceph-disk: fix udev rules
- rpm: install sysvinit script on package install
- ceph-disk: fix OSD start on machine reboot on Debian wheezy
- ceph-disk: activate OSD when journal device appears second
- ceph-disk: fix various bugs on RHEL/CentOS 6.3
- ceph-disk: add 'zap' command
- ceph-disk: add '[un]suppress-activate' command for preparing spare disks
- upstart: start on runlevel [2345] (instead of after the first network interface starts)
- ceph-fuse, libcephfs: handle mds session reset during session open
- ceph-fuse, libcephfs: fix two capability revocation bugs
- ceph-fuse: fix thread creation on startup
- all daemons: create /var/run/ceph directory on startup if missing

For more detailed information, see [the complete changelog](#).

V0.61.3 "CUTTLEFISH"

This release resolves a number of problems with the monitors and leveldb that users have been seeing. Please upgrade.

UPGRADING

- There is one known problem with mon upgrades from bobtail. If the ceph-mon conversion on startup is aborted or fails for some reason, we do not correctly error out, but instead continue with (in certain cases) odd results. Please be careful if you have to restart the mons during the upgrade. A 0.61.4 release with a fix will be out shortly.
- In the meantime, for current cuttlefish users, v0.61.3 is safe to use.

NOTABLE CHANGES

- mon: paxos state trimming fix (resolves runaway disk usage)
- mon: finer-grained compaction on trim
- mon: discard messages from disconnected clients (lowers load)

- mon: leveldb compaction and other stats available via admin socket
- mon: async compaction (lower overhead)
- mon: fix bug incorrectly marking osds down with insufficient failure reports
- osd: fixed small bug in pg request map
- osd: avoid rewriting pg info on every osdmap
- osd: avoid internal heartbeta timeouts when scrubbing very large objects
- osd: fix narrow race with journal replay
- mon: fixed narrow pg split race
- rgw: fix leaked space when copying object
- rgw: fix iteration over large/untrimmed usage logs
- rgw: fix locking issue with ops log socket
- rgw: require matching version of librados
- librbd: make image creation defaults configurable (e.g., create format 2 images via qemu-img)
- fix units in ‘ceph df’ output
- debian: fix prerm/postinst hooks to start/stop daemons appropriately
- upstart: allow uppercase daemons names (and thus hostnames)
- sysvinit: fix enumeration of local daemons by type
- sysvinit: fix osd weight calcuation when using -a
- fix build on unsigned char platforms (e.g., arm)

For more detailed information, see [the complete changelog](#).

V0.61.2 “CUTTLEFISH”

This release disables a monitor debug log that consumes disk space and fixes a bug when upgrade some monitors from bobtail to cuttlefish.

NOTABLE CHANGES

- mon: fix conversion of stores with duplicated GV values
- mon: disable ‘mon debug dump transactions’ by default

For more detailed information, see [the complete changelog](#).

V0.61.1 “CUTTLEFISH”

This release fixes a problem when upgrading a bobtail cluster that had snapshots to cuttlefish.

NOTABLE CHANGES

- osd: handle upgrade when legacy snap collections are present; repair from previous failed restart
- ceph-create-keys: fix race with ceph-mon startup (which broke ‘ceph-deploy gatherkeys ...’)
- ceph-create-keys: gracefully handle bad response from ceph-osd
- sysvinit: do not assume default osd_data when automatically weighting OSD
- osd: avoid crash from ill-behaved classes using getomapvals
- debian: fix squeeze dependency
- mon: debug options to log or dump leveldb transactions

For more detailed information, see [the complete changelog](#).

V0.61 “CUTTLEFISH”

UPGRADING FROM V0.60

- The ceph-deploy tool is now the preferred method of provisioning new clusters. For existing clusters created via mkcephfs that would like to transition to the new tool, there is a migration path, documented at [Transitioning to ceph-deploy](#).
- The sysvinit script (/etc/init.d/ceph) will now verify (and, if necessary, update) the OSD’s position in the CRUSH map on startup. (The upstart script has always worked this way.) By default, this ensures that the OSD is under a ‘host’ with a name that matches the hostname (hostname -s). Legacy clusters create with mkcephfs do this by default, so this should

not cause any problems, but legacy clusters with customized CRUSH maps with an alternate structure should set osd crush update on start = false.

- radosgw-admin now uses the term zone instead of cluster to describe each instance of the radosgw data store (and corresponding collection of radosgw daemons). The usage for the radosgw-admin command and the 'rgw zone root pool' config options have changed accordingly.
- rbd progress indicators now go to standard error instead of standard out. (You can disable progress with -no-progress.)
- The 'rbd resize ...' command now requires the -allow-shrink option when resizing to a smaller size. Expanding images to a larger size is unchanged.
- Please review the changes going back to 0.56.4 if you are upgrading all the way from bobtail.
- The old 'ceph stop_cluster' command has been removed.
- The sysvinit script now uses the ceph.conf file on the remote host when starting remote daemons via the '-a' option. Note that if '-a' is used in conjunction with '-c path', the path must also be present on the remote host (it is not copied to a temporary file, as it was previously).

UPGRADING FROM V0.56.4 "BOBTAIL"

Please see [Upgrading from Bobtail to Cuttlefish](#) for details.

- The ceph-deploy tool is now the preferred method of provisioning new clusters. For existing clusters created via mkcephfs that would like to transition to the new tool, there is a migration path, documented at [Transitioning to ceph-deploy](#).
- The sysvinit script (/etc/init.d/ceph) will now verify (and, if necessary, update) the OSD's position in the CRUSH map on startup. (The upstart script has always worked this way.) By default, this ensures that the OSD is under a 'host' with a name that matches the hostname (hostname -s). Legacy clusters create with mkcephfs do this by default, so this should not cause any problems, but legacy clusters with customized CRUSH maps with an alternate structure should set osd crush update on start = false.
- radosgw-admin now uses the term zone instead of cluster to describe each instance of the radosgw data store (and corresponding collection of radosgw daemons). The usage for the radosgw-admin command and the 'rgw zone root pool' config options have changed accordingly.
- rbd progress indicators now go to standard error instead of standard out. (You can disable progress with -no-progress.)
- The 'rbd resize ...' command now requires the -allow-shrink option when resizing to a smaller size. Expanding images to a larger size is unchanged.
- Please review the changes going back to 0.56.4 if you are upgrading all the way from bobtail.
- The old 'ceph stop_cluster' command has been removed.
- The sysvinit script now uses the ceph.conf file on the remote host when starting remote daemons via the '-a' option. Note that if '-a' is used in conjunction with '-c path', the path must also be present on the remote host (it is not copied to a temporary file, as it was previously).
- The monitor is using a completely new storage strategy and intra-cluster protocol. This means that cuttlefish and bobtail monitors do not talk to each other. When you upgrade each one, it will convert its local data store to the new format. Once you upgrade a majority, the quorum will be formed using the new protocol and the old monitors will be blocked out until they too get upgraded. For this reason, we recommend not running a mixed-version cluster for very long.
- ceph-mon now requires the creation of its data directory prior to -mkfs, similarly to what happens on ceph-osd. This directory is no longer automatically created, and custom scripts should be adjusted to reflect just that.
- The monitor now enforces that MDS names be unique. If you have multiple daemons start with the same id (e.g., mds.a) the second one will implicitly mark the first as failed. This makes things less confusing and makes a daemon restart faster (we no longer wait for the stopped daemon to time out) but existing multi-mds configurations may need to be adjusted accordingly to give daemons unique names.
- The 'ceph osd pool delete <poolname>' and 'rados rmpool <poolname>' now have safety interlocks with loud warnings that make you confirm pool removal. Any scripts currently rely on these functions zapping data without confirmation need to be adjusted accordingly.

NOTABLE CHANGES FROM V0.60

- rbd: incremental backups
- rbd: only set STRIPINGV2 feature if striping parameters are incompatible with old versions
- rbd: require -allow-shrink for resizing images down
- librbd: many bug fixes
- rgw: management REST API
- rgw: fix object corruption on COPY to self
- rgw: new sysvinit script for rpm-based systems
- rgw: allow buckets with '_'
- rgw: CORS support
- mon: many fixes

- mon: improved trimming behavior
- mon: fix data conversion/upgrade problem (from bobtail)
- mon: ability to tune leveldb
- mon: config-keys service to store arbitrary data on monitor
- mon: ‘osd crush add|link|unlink|add-bucket ...’ commands
- mon: trigger leveldb compaction on trim
- osd: per-rados pool quotas (objects, bytes)
- osd: tool to export, import, and delete PGs from an individual OSD data store
- osd: notify mon on clean shutdown to avoid IO stall
- osd: improved detection of corrupted journals
- osd: ability to tune leveldb
- osd: improve client request throttling
- osd, librados: fixes to the LIST_SNAPS operation
- osd: improvements to scrub error repair
- osd: better prevention of wedging OSDs with ENOSPC
- osd: many small fixes
- mds: fix xattr handling on root inode
- mds: fixed bugs in journal replay
- mds: many fixes
- librados: clean up snapshot constant definitions
- libcephfs: calls to query CRUSH topology (used by Hadoop)
- ceph-fuse, libcephfs: misc fixes to mds session management
- ceph-fuse: disabled cache invalidation (again) due to potential deadlock with kernel
- sysvinit: try to start all daemons despite early failures
- ceph-disk: new ‘list’ command
- ceph-disk: hotplug fixes for RHEL/CentOS
- ceph-disk: fix creation of OSD data partitions on >2TB disks
- osd: fix udev rules for RHEL/CentOS systems
- fix daemon logging during initial startup

NOTABLE CHANGES FROM V0.56 “BOBTAIL”

- always use installed system leveldb (Gary Lowell)
- auth: ability to require new cephx signatures on messages (still off by default)
- buffer unit testing (Loic Dachary)
- ceph tool: some CLI interface cleanups
- ceph-disk: improve multicluster support, error handling (Sage Weil)
- ceph-disk: support for dm-crypt (Alexandre Marangone)
- ceph-disk: support for sysvinit, directories or partitions (not full disks)
- ceph-disk: fix mkfs args on old distros (Alexandre Marangone)
- ceph-disk: fix creation of OSD data partitions on >2TB disks
- ceph-disk: hotplug fixes for RHEL/CentOS
- ceph-disk: new ‘list’ command
- ceph-fuse, libcephfs: misc fixes to mds session management
- ceph-fuse: disabled cache invalidation (again) due to potential deadlock with kernel
- ceph-fuse: enable kernel cache invalidation (Sam Lang)
- ceph-fuse: fix statfs(2) reporting
- ceph-fuse: session handling cleanup, bug fixes (Sage Weil)
- crush: ability to create, remove rules via CLI
- crush: update weights for all instances of an item, not just the first (Sage Weil)
- fix daemon logging during initial startup
- fixed log rotation (Gary Lowell)
- init-ceph, mkcephfs: close a few security holes with -a (Sage Weil)
- libcephfs: calls to query CRUSH topology (used by Hadoop)
- libcephfs: many fixes, cleanups with the Java bindings
- libcephfs: new topo API requests for Hadoop (Noah Watkins)
- librados: clean up snapshot constant definitions
- librados: fix linger bugs (Josh Durgin)
- librbd: fixed flatten deadlock (Josh Durgin)
- librbd: fixed some locking issues with flatten (Josh Durgin)
- librbd: many bug fixes
- librbd: optionally wait for flush before enabling writeback (Josh Durgin)
- many many cleanups (Danny Al-Gaaf)
- mds, ceph-fuse: fix bugs with replayed requests after MDS restart (Sage Weil)

- mds, ceph-fuse: manage layouts via xattrs
- mds: allow xattr on root
- mds: fast failover between MDSs (enforce unique mds names)
- mds: fix xattr handling on root inode
- mds: fixed bugs in journal replay
- mds: improve session cleanup (Sage Weil)
- mds: many fixes (Yan Zheng)
- mds: misc bug fixes with clustered MDSs and failure recovery
- mds: misc bug fixes with readdir
- mds: new encoding for all data types (to allow forward/backward compatibility) (Greg Farnum)
- mds: store and update backpointers/traces on directory, file objects (Sam Lang)
- mon: ‘osd crush add|link|unlink|add-bucket ...’ commands
- mon: ability to tune leveldb
- mon: approximate recovery, IO workload stats
- mon: avoid marking entire CRUSH subtrees out (e.g., if an entire rack goes offline)
- mon: config-keys service to store arbitrary data on monitor
- mon: easy adjustment of crush tunables via ‘ceph osd crush tunables ...’
- mon: easy creation of crush rules via ‘ceph osd rule ...’
- mon: fix data conversion/upgrade problem (from bobtail)
- mon: improved trimming behavior
- mon: many fixes
- mon: new ‘ceph df [detail]’ command
- mon: new checks for identifying and reporting clock drift
- mon: rearchitected to utilize single instance of paxos and a key/value store (Joao Luis)
- mon: safety check for pool deletion
- mon: shut down safely if disk approaches full (Joao Luis)
- mon: trigger leveldb compaction on trim
- msgr: fix comparison of IPv6 addresses (fixes monitor bringup via ceph-deploy, chef)
- msgr: fixed race in connection reset
- msgr: optionally tune TCP buffer size to avoid throughput collapse (Jim Schutt)
- much code cleanup and optimization (Danny Al-Gaaf)
- osd, librados: ability to list watchers (David Zafman)
- osd, librados: fixes to the LIST_SNAPS operation
- osd, librados: new listsnaps command (David Zafman)
- osd: a few journaling bug fixes
- osd: ability to tune leveldb
- osd: add ‘noscrub’, ‘nodeepscrub’ osdmap flags (David Zafman)
- osd: better prevention of wedging OSDs with ENOSPC
- osd: ceph-filestore-dump tool for debugging
- osd: connection handling bug fixes
- osd: deep-scrub omap keys/values
- osd: default to libaio for the journal (some performance boost)
- osd: fix hang in ‘journal aio = true’ mode (Sage Weil)
- osd: fix pg log trimming (avoids memory bloat on degraded clusters)
- osd: fix udev rules for RHEL/CentOS systems
- osd: fixed bug in journal checksums (Sam Just)
- osd: improved client request throttling
- osd: improved handling when disk fills up (David Zafman)
- osd: improved journal corruption detection (Sam Just)
- osd: improved detection of corrupted journals
- osd: improvements to scrub error repair
- osd: make tracking of object snapshot metadata more efficient (Sam Just)
- osd: many small fixes
- osd: misc fixes to PG split (Sam Just)
- osd: move pg info, log into leveldb (== better performance) (David Zafman)
- osd: notify mon on clean shutdown to avoid IO stall
- osd: per-rados pool quotas (objects, bytes)
- osd: refactored watch/notify infrastructure (fixes protocol, removes many bugs) (Sam Just)
- osd: support for improved hashing of PGs across OSDs via HASHPSPOOL pool flag and feature
- osd: tool to export, import, and delete PGs from an individual OSD data store
- osd: trim log more aggressively, avoid appearance of leak memory
- osd: validate snap collections on startup
- osd: verify snap collections on startup (Sam Just)
- radosgw: ACL grants in headers (Caleb Miles)
- radosgw: ability to listen to fastcgi via a port (Guilhem Lettron)

- radosgw: fix object copy onto self (Yehuda Sadeh)
- radosgw: misc fixes
- rbd-fuse: new tool, package
- rbd: avoid FIEMAP when importing from file (it can be buggy)
- rbd: incremental backups
- rbd: only set STRIPINGV2 feature if striping parameters are incompatible with old versions
- rbd: require -allow-shrink for resizing images down
- rbd: udevadm settle on map/unmap to avoid various races (Dan Mick)
- rbd: wait for udev to settle in strategic places (avoid spurious errors, failures)
- rgw: CORS support
- rgw: allow buckets with '_'
- rgw: fix Content-Length on 32-bit machines (Jan Harkes)
- rgw: fix log rotation
- rgw: fix object corruption on COPY to self
- rgw: fixed >4MB range requests (Jan Harkes)
- rgw: new sysvinit script for rpm-based systems
- rpm/deb: do not remove /var/lib/ceph on purge (v0.59 was the only release to do so)
- sysvinit: try to start all daemons despite early failures
- upstart: automatically set osd weight based on df (Guilhem Lettron)
- use less memory for logging by default

V0.60

UPGRADING

- Please note that the recently added librados ‘list_snaps’ function call is in a state of flux and is changing slightly in v0.61. You are advised not to make use of it in v0.59 or v0.60.

NOTABLE CHANGES

- osd: make tracking of object snapshot metadata more efficient (Sam Just)
- osd: misc fixes to PG split (Sam Just)
- osd: improve journal corruption detection (Sam Just)
- osd: improve handling when disk fills up (David Zafman)
- osd: add ‘noscrub’, ‘nodeepscrub’ osdmap flags (David Zafman)
- osd: fix hang in ‘journal aio = true’ mode (Sage Weil)
- ceph-disk-prepare: fix mkfs args on old distros (Alexandre Marangone)
- ceph-disk-activate: improve multicluster support, error handling (Sage Weil)
- librbd: optionally wait for flush before enabling writeback (Josh Durgin)
- crush: update weights for all instances of an item, not just the first (Sage Weil)
- mon: shut down safely if disk approaches full (Joao Luis)
- rgw: fix Content-Length on 32-bit machines (Jan Harkes)
- mds: store and update backpointers/traces on directory, file objects (Sam Lang)
- mds: improve session cleanup (Sage Weil)
- mds, ceph-fuse: fix bugs with replayed requests after MDS restart (Sage Weil)
- ceph-fuse: enable kernel cache invalidation (Sam Lang)
- libcephfs: new topo API requests for Hadoop (Noah Watkins)
- ceph-fuse: session handling cleanup, bug fixes (Sage Weil)
- much code cleanup and optimization (Danny Al-Gaaf)
- use less memory for logging by default
- upstart: automatically set osd weight based on df (Guilhem Lettron)
- init-ceph, mkcephfs: close a few security holes with -a (Sage Weil)
- rpm/deb: do not remove /var/lib/ceph on purge (v0.59 was the only release to do so)

V0.59

UPGRADING

- The monitor is using a completely new storage strategy and intra-cluster protocol. This means that v0.59 and pre-v0.59 monitors do not talk to each other. When you upgrade each one, it will convert its local data store to the new format.

Once you upgrade a majority, the quorum will be formed using the new protocol and the old monitors will be blocked out until they too get upgraded. For this reason, we recommend not running a mixed-version cluster for very long.

- ceph-mon now requires the creation of its data directory prior to -mkfs, similarly to what happens on ceph-osd. This directory is no longer automatically created, and custom scripts should be adjusted to reflect just that.

NOTABLE CHANGES

- mon: rearchitected to utilize single instance of paxos and a key/value store (Joao Luis)
- mon: new 'ceph df [detail]' command
- osd: support for improved hashing of PGs across OSDs via HASHPSPOOL pool flag and feature
- osd: refactored watch/notify infrastructure (fixes protocol, removes many bugs) (Sam Just)
- osd, librados: ability to list watchers (David Zafman)
- osd, librados: new listsnaps command (David Zafman)
- osd: trim log more aggressively, avoid appearance of leak memory
- osd: misc split fixes
- osd: a few journaling bug fixes
- osd: connection handling bug fixes
- rbd: avoid FIEMAP when importing from file (it can be buggy)
- librados: fix linger bugs (Josh Durgin)
- librbd: fixed flatten deadlock (Josh Durgin)
- rgw: fixed >4MB range requests (Jan Harkes)
- rgw: fix log rotation
- mds: allow xattrs on root
- ceph-fuse: fix statfs(2) reporting
- msgr: optionally tune TCP buffer size to avoid throughput collapse (Jim Schutt)
- consume less memory for logging by default
- always use system leveldb (Gary Lowell)

V0.58

UPGRADING

- The monitor now enforces that MDS names be unique. If you have multiple daemons start with with the same id (e.g., mds.a) the second one will implicitly mark the first as failed. This makes things less confusing and makes a daemon restart faster (we no longer wait for the stopped daemon to time out) but existing multi-mds configurations may need to be adjusted accordingly to give daemons unique names.

NOTABLE CHANGES

- librbd: fixed some locking issues with flatten (Josh Durgin)
- rbd: udevadm settle on map/unmap to avoid various races (Dan Mick)
- osd: move pg info, log into leveldb (== better performance) (David Zafman)
- osd: fix pg log trimming (avoids memory bloat on degraded clusters)
- osd: fixed bug in journal checksums (Sam Just)
- osd: verify snap collections on startup (Sam Just)
- ceph-disk-prepare/activate: support for dm-crypt (Alexandre Marangone)
- ceph-disk-prepare/activate: support for sysvinit, directories or partitions (not full disks)
- msgr: fixed race in connection reset
- msgr: fix comparison of IPv6 addresses (fixes monitor bringup via ceph-deploy, chef)
- radosgw: fix object copy onto self (Yehuda Sadeh)
- radosgw: ACL grants in headers (Caleb Miles)
- radosgw: ability to listen to fastcgi via a port (Guilhem Lettron)
- mds: new encoding for all data types (to allow forward/backward compatibility) (Greg Farnum)
- mds: fast failover between MDSS (enforce unique mds names)
- crush: ability to create, remove rules via CLI
- many many cleanups (Danny Al-Gaaf)
- buffer unit testing (Loic Dachary)
- fixed log rotation (Gary Lowell)

V0.57

This development release has a lot of additional functionality accumulated over the last couple months. Most of the bug fixes (with the notable exception of the MDS related work) has already been backported to v0.56.x, and is not mentioned here.

UPGRADING

- The ‘ceph osd pool delete <poolname>’ and ‘rados rm pool <poolname>’ now have safety interlocks with loud warnings that make you confirm pool removal. Any scripts currently rely on these functions zapping data without confirmation need to be adjusted accordingly.

NOTABLE CHANGES

- osd: default to libaio for the journal (some performance boost)
- osd: validate snap collections on startup
- osd: ceph-filestore-dump tool for debugging
- osd: deep-scrub omap keys/values
- ceph tool: some CLI interface cleanups
- mon: easy adjustment of crush tunables via ‘ceph osd crush tunables ...’
- mon: easy creation of crush rules via ‘ceph osd rule ...’
- mon: approximate recovery, IO workload stats
- mon: avoid marking entire CRUSH subtrees out (e.g., if an entire rack goes offline)
- mon: safety check for pool deletion
- mon: new checks for identifying and reporting clock drift
- radosgw: misc fixes
- rbd: wait for udev to settle in strategic places (avoid spurious errors, failures)
- rbd-fuse: new tool, package
- mds, ceph-fuse: manage layouts via xattrs
- mds: misc bug fixes with clustered MDSs and failure recovery
- mds: misc bug fixes with readdir
- libcephfs: many fixes, cleanups with the Java bindings
- auth: ability to require new cephx signatures on messages (still off by default)

V0.56.7 “BOBTAIL”

This bobtail update fixes a range of radosgw bugs (including an easily triggered crash from multi-delete), a possible data corruption issue with power failure on XFS, and several OSD problems, including a memory “leak” that will affect aged clusters.

NOTABLE CHANGES

- ceph-fuse: create finisher flags after fork()
- debian: fix preinst/postinst hooks; do not restart daemons on upgrade
- librados: fix async aio completion wakeup (manifests as rbd hang)
- librados: fix hang when osd becomes full and then not full
- librados: fix locking for aio completion refcounting
- librbd python bindings: fix stripe_unit, stripe_count
- librbd: make image creation default configurable
- mon: fix validation of mds ids in mon commands
- osd: avoid excessive disk updates during peering
- osd: avoid excessive memory usage on scrub
- osd: avoid heartbeat failure/suicide when scrubbing
- osd: misc minor bug fixes
- osd: use fdatasync instead of sync_file_range (may avoid xfs power-loss corruption)
- rgw: escape prefix correctly when listing objects
- rgw: fix copy attrs
- rgw: fix crash on multi delete
- rgw: fix locking/crash when using ops log socket
- rgw: fix usage logging
- rgw: handle deep uri resources

For more detailed information, see [the complete changelog](#).

V0.56.6 “BOBTAIL”

NOTABLE CHANGES

- rgw: fix garbage collection
- rpm: fix package dependencies

For more detailed information, see [the complete changelog](#).

V0.56.5 "BOBTAIL"

UPGRADING

- ceph-disk[-prepare,-activate] behavior has changed in various ways. There should not be any compatibility issues, but chef users should be aware.

NOTABLE CHANGES

- mon: fix recording of quorum feature set (important for argonaut -> bobtail -> cuttlefish mon upgrades)
- osd: minor peering bug fixes
- osd: fix a few bugs when pools are renamed
- osd: fix occasionally corrupted pg stats
- osd: fix behavior when broken v0.56[.0] clients connect
- rbd: avoid FIEMAP ioctl on import (it is broken on some kernels)
- librbd: fixes for several request/reply ordering bugs
- librbd: only set STRIPINGV2 feature on new images when needed
- librbd: new async flush method to resolve qemu hangs (requires QEMU update as well)
- librbd: a few fixes to flatten
- ceph-disk: support for dm-crypt
- ceph-disk: many backports to allow bobtail deployments with ceph-deploy, chef
- sysvinit: do not stop starting daemons on first failure
- udev: fixed rules for redhat-based distros
- build fixes for raring

For more detailed information, see [the complete changelog](#).

V0.56.4 "BOBTAIL"

UPGRADING

- There is a fix in the syntax for the output of 'ceph osd tree -format=json'.
- The MDS disk format has changed from prior releases *and* from v0.57. In particular, upgrades to v0.56.4 are safe, but you cannot move from v0.56.4 to v0.57 if you are using the MDS for CephFS; you must upgrade directly to v0.58 (or later) instead.

NOTABLE CHANGES

- mon: fix bug in bringup with IPv6
- reduce default memory utilization by internal logging (all daemons)
- rgw: fix for bucket removal
- rgw: reopen logs after log rotation
- rgw: fix multipat upload listing
- rgw: don't copy object when copied onto self
- osd: fix caps parsing for pools with - or _
- osd: allow pg log trimming when degraded, scrubbing, recovering (reducing memory consumption)
- osd: fix potential deadlock when 'journal aio = true'
- osd: various fixes for collection creation/removal, rename, temp collections
- osd: various fixes for PG split
- osd: deep-scrub omap key/value data
- osd: fix rare bug in journal replay

- osd: misc fixes for snapshot tracking
- osd: fix leak in recovery reservations on pool deletion
- osd: fix bug in connection management
- osd: fix for op ordering when rebalancing
- ceph-fuse: report file system size with correct units
- mds: get and set directory layout policies via virtual xattrs
- mds: on-disk format revision (see upgrading note above)
- mkcephfs, init-ceph: close potential security issues with predictable filenames

For more detailed information, see [the complete changelog](#).

V0.56.3 “BOBTAIL”

This release has several bug fixes surrounding OSD stability. Most significantly, an issue with OSDs being unresponsive shortly after startup (and occasionally crashing due to an internal heartbeat check) is resolved. Please upgrade.

UPGRADING

- A bug was fixed in which the OSDMap epoch for PGs without any IO requests was not recorded. If there are pools in the cluster that are completely idle (for example, the data and metadata pools normally used by CephFS), and a large number of OSDMap epochs have elapsed since the ceph-osd daemon was last restarted, those maps will get reprocessed when the daemon restarts. This process can take a while if there are a lot of maps. A workaround is to ‘touch’ any idle pools with IO prior to restarting the daemons after packages are upgraded:

```
rados bench 10 write -t 1 -b 4096 -p {POOLNAME}
```

This will typically generate enough IO to touch every PG in the pool without generating significant cluster load, and also cleans up any temporary objects it creates.

NOTABLE CHANGES

- osd: flush peering work queue prior to start
- osd: persist osdmap epoch for idle PGs
- osd: fix and simplify connection handling for heartbeats
- osd: avoid crash on invalid admin command
- mon: fix rare races with monitor elections and commands
- mon: enforce that OSD reweights be between 0 and 1 (NOTE: not CRUSH weights)
- mon: approximate client, recovery bandwidth logging
- radosgw: fixed some XML formatting to conform to Swift API inconsistency
- radosgw: fix usage accounting bug; add repair tool
- radosgw: make fallback URI configurable (necessary on some web servers)
- librbd: fix handling for interrupted ‘unprotect’ operations
- mds, ceph-fuse: allow file and directory layouts to be modified via virtual xattrs

For more detailed information, see [the complete changelog](#).

V0.56.2 “BOBTAIL”

This release has a wide range of bug fixes, stability improvements, and some performance improvements. Please upgrade.

UPGRADING

- The meaning of the ‘osd scrub min interval’ and ‘osd scrub max interval’ has changed slightly. The min interval used to be meaningless, while the max interval would only trigger a scrub if the load was sufficiently low. Now, the min interval option works the way the old max interval did (it will trigger a scrub after this amount of time if the load is low), while the max interval will force a scrub regardless of load. The default options have been adjusted accordingly. If you have customized these in ceph.conf, please review their values when upgrading.
- CRUSH maps that are generated by default when calling ceph-mon --mkfs directly now distribute replicas across hosts instead of across OSDs. Any provisioning tools that are being used by Ceph may be affected, although probably for the better, as distributing across hosts is a much more commonly sought behavior. If you use mkcephfs to create the cluster,

the default CRUSH rule is still inferred by the number of hosts and/or racks in the initial ceph.conf.

NOTABLE CHANGES

- osd: snapshot trimming fixes
- osd: scrub snapshot metadata
- osd: fix osdmap trimming
- osd: misc peering fixes
- osd: stop heartbeating with peers if internal threads are stuck/hung
- osd: PG removal is friendlier to other workloads
- osd: fix recovery start delay (was causing very slow recovery)
- osd: fix scheduling of explicitly requested scrubs
- osd: fix scrub interval config options
- osd: improve recovery vs client io tuning
- osd: improve ‘slow request’ warning detail for better diagnosis
- osd: default CRUSH map now distributes across hosts, not OSDs
- osd: fix crash on 32-bit hosts triggered by librbd clients
- librbd: fix error handling when talking to older OSDs
- mon: fix a few rare crashes
- ceph command: ability to easily adjust CRUSH tunables
- radosgw: object copy does not copy source ACLs
- rados command: fix omap command usage
- sysvinit script: set ulimit -n properly on remote hosts
- msgr: fix narrow race with message queuing
- fixed compilation on some old distros (e.g., RHEL 5.x)

For more detailed information, see [the complete changelog](#).

V0.56.1 “BOBTAIL”

This release has two critical fixes. Please upgrade.

UPGRADING

- There is a protocol compatibility problem between v0.56 and any other version that is now fixed. If your radosgw or RBD clients are running v0.56, they will need to be upgraded too. If they are running a version prior to v0.56, they can be left as is.

NOTABLE CHANGES

- osd: fix commit sequence for XFS, ext4 (or any other non-btrfs) to prevent data loss on power cycle or kernel panic
- osd: fix compatibility for CALL operation
- osd: process old osdmmaps prior to joining cluster (fixes slow startup)
- osd: fix a couple of recovery-related crashes
- osd: fix large io requests when journal is in (non-default) aio mode
- log: fix possible deadlock in logging code

For more detailed information, see [the complete changelog](#).

V0.56 “BOBTAIL”

Bobtail is the second stable release of Ceph, named in honor of the *Bobtail Squid*: https://en.wikipedia.org/wiki/Bobtail_squid.

KEY FEATURES SINCE V0.48 “ARGONAUT”

- Object Storage Daemon (OSD): improved threading, small-io performance, and performance during recovery
- Object Storage Daemon (OSD): regular “deep” scrubbing of all stored data to detect latent disk errors
- RADOS Block Device (RBD): support for copy-on-write clones of images.
- RADOS Block Device (RBD): better client-side caching.
- RADOS Block Device (RBD): advisory image locking

- Rados Gateway (RGW): support for efficient usage logging/scraping (for billing purposes)
- Rados Gateway (RGW): expanded S3 and Swift API coverage (e.g., POST, multi-object delete)
- Rados Gateway (RGW): improved striping for large objects
- Rados Gateway (RGW): OpenStack Keystone integration
- RPM packages for Fedora, RHEL/CentOS, OpenSUSE, and SLES
- mkcephfs: support for automatically formatting and mounting XFS and ext4 (in addition to btrfs)

UPGRADING

Please refer to the document [Upgrading from Argonaut to Bobtail](#) for details.

- Cephx authentication is now enabled by default (since v0.55). Upgrading a cluster without adjusting the Ceph configuration will likely prevent the system from starting up on its own. We recommend first modifying the configuration to indicate that authentication is disabled, and only then upgrading to the latest version:

```
auth client required = none
auth service required = none
auth cluster required = none
```

- Ceph daemons can be upgraded one-by-one while the cluster is online and in service.
- The ceph-osd daemons must be upgraded and restarted *before* any radosgw daemons are restarted, as they depend on some new ceph-osd functionality. (The ceph-mon, ceph-osd, and ceph-mds daemons can be upgraded and restarted in any order.)
- Once each individual daemon has been upgraded and restarted, it cannot be downgraded.
- The cluster of ceph-mon daemons will migrate to a new internal on-wire protocol once all daemons in the quorum have been upgraded. Upgrading only a majority of the nodes (e.g., two out of three) may expose the cluster to a situation where a single additional failure may compromise availability (because the non-upgraded daemon cannot participate in the new protocol). We recommend not waiting for an extended period of time between ceph-mon upgrades.
- The ops log and usage log for radosgw are now off by default. If you need these logs (e.g., for billing purposes), you must enable them explicitly. For logging of all operations to objects in the .log pool (see `radosgw-admin log ...`):

```
rgw enable ops log = true
```

For usage logging of aggregated bandwidth usage (see `radosgw-admin usage ...`):

```
rgw enable usage log = true
```

- You should not create or use “format 2” RBD images until after all ceph-osd daemons have been upgraded. Note that “format 1” is still the default. You can use the new `ceph osd ls` and `ceph tell osd.N version` commands to doublecheck your cluster. `ceph osd ls` will give a list of all OSD IDs that are part of the cluster, and you can use that to write a simple shell loop to display all the OSD version strings:

```
for i in $(ceph osd ls); do
    ceph tell osd.${i} version
done
```

COMPATIBILITY CHANGES

- The ‘`ceph osd create [<uuid>]`’ command now rejects an argument that is not a UUID. (Previously it would take an optional integer OSD id.) This correct syntax has been ‘`ceph osd create [<uuid>]`’ since v0.47, but the older calling convention was being silently ignored.
- The CRUSH map root nodes now have type root instead of type pool. This avoids confusion with RADOS pools, which are not directly related. Any scripts or tools that use the `ceph osd crush ...` commands may need to be adjusted accordingly.
- The `ceph osd pool create <poolname> <pgnum>` command now requires the pgnum argument. Previously this was optional, and would default to 8, which was almost never a good number.

- Degraded mode (when there fewer than the desired number of replicas) is now more configurable on a per-pool basis, with the min_size parameter. By default, with min_size 0, this allows I/O to objects with N - floor(N/2) replicas, where N is the total number of expected copies. Argonaut behavior was equivalent to having min_size = 1, so I/O would always be possible if any completely up to date copy remained. min_size = 1 could result in lower overall availability in certain cases, such as flapping network partitions.
- The sysvinit start/stop script now defaults to adjusting the max open files ulimit to 16384. On most systems the default is 1024, so this is an increase and won't break anything. If some system has a higher initial value, however, this change will lower the limit. The value can be adjusted explicitly by adding an entry to the ceph.conf file in the appropriate section. For example:

```
[global]
max open files = 32768
```

- 'rbd lock list' and 'rbd showmapped' no longer use tabs as separators in their output.
- There is configurable limit on the number of PGs when creating a new pool, to prevent a user from accidentally specifying a ridiculous number for pg_num. It can be adjusted via the 'mon max pool pg num' option on the monitor, and defaults to 65536 (the current max supported by the Linux kernel client).
- The osd capabilities associated with a rados user have changed syntax since 0.48 argonaut. The new format is mostly backwards compatible, but there are two backwards-incompatible changes:
 - specifying a list of pools in one grant, i.e. 'allow r pool=foo,bar' is now done in separate grants, i.e. 'allow r pool=foo, allow r pool=bar'.
 - restricting pool access by pool owner ('allow r uid=foo') is removed. This feature was not very useful and unused in practice.

The new format is documented in the ceph-authtool man page.

- 'rbd cp' and 'rbd rename' use rbd as the default destination pool, regardless of what pool the source image is in. Previously they would default to the same pool as the source image.
- 'rbd export' no longer prints a message for each object written. It just reports percent complete like other long-lasting operations.
- 'ceph osd tree' now uses 4 decimal places for weight so output is nicer for humans
- Several monitor operations are now idempotent:
 - ceph osd pool create
 - ceph osd pool delete
 - ceph osd pool mksnap
 - ceph osd rm
 - ceph pg <pgid> revert

NOTABLE CHANGES

- auth: enable cephx by default
- auth: expanded authentication settings for greater flexibility
- auth: sign messages when using cephx
- build fixes for Fedora 18, CentOS/RHEL 6
- ceph: new 'osd ls' and 'osd tell <osd.N> version' commands
- ceph-debugpack: misc improvements
- ceph-disk-prepare: creates and labels GPT partitions
- ceph-disk-prepare: support for external journals, default mount/mkfs options, etc.
- ceph-fuse/libcephfs: many misc fixes, admin socket debugging
- ceph-fuse: fix handling for .. in root directory
- ceph-fuse: many fixes (including memory leaks, hangs)
- ceph-fuse: mount helper (mount.fuse.ceph) for use with /etc/fstab
- ceph.spec: misc packaging fixes
- common: thread pool sizes can now be adjusted at runtime
- config: \$pid is now available as a metavariable
- crush: default root of tree type is now 'root' instead of 'pool' (to avoid confusiong wrt rados pools)
- crush: fixed retry behavior with chooseleaf via tunable
- crush: tunables documented; feature bit now present and enforced

- libcephfs: java wrapper
- librados: several bug fixes (rare races, locking errors)
- librados: some locking fixes
- librados: watch/notify fixes, misc memory leaks
- librbd: a few fixes to ‘discard’ support
- librbd: fine-grained striping feature
- librbd: fixed memory leaks
- librbd: fully functional and documented image cloning
- librbd: image (advisory) locking
- librbd: improved caching (of object non-existence)
- librbd: ‘flatten’ command to sever clone parent relationship
- librbd: ‘protect’/‘unprotect’ commands to prevent clone parent from being deleted
- librbd: clip requests past end-of-image.
- librbd: fixes an issue with some windows guests running in qemu (remove floating point usage)
- log: fix in-memory buffering behavior (to only write log messages on crash)
- mds: fix ino release on abort session close, relative getattr path, mds shutdown, other misc items
- mds: misc fixes
- mkcephfs: fix for default keyring, osd data/journal locations
- mkcephfs: support for formatting xfs, ext4 (as well as btrfs)
- init: support for automatically mounting xfs and ext4 osd data directories
- mon, radosgw, ceph-fuse: fixed memory leaks
- mon: improved ENOSPC, fs error checking
- mon: less-destructive ceph-mon -mkfs behavior
- mon: misc fixes
- mon: more informative info about stuck PGs in ‘health detail’
- mon: information about recovery and backfill in ‘pg <pgid> query’
- mon: new ‘osd crush create-or-move ...’ command
- mon: new ‘osd crush move ...’ command lets you rearrange your CRUSH hierarchy
- mon: optionally dump ‘osd tree’ in json
- mon: configurable cap on maximum osd number (mon max osd)
- mon: many bug fixes (various races causing ceph-mon crashes)
- mon: new on-disk metadata to facilitate future mon changes (post-bobtail)
- mon: election bug fixes
- mon: throttle client messages (limit memory consumption)
- mon: throttle osd flapping based on osd history (limits osdmap ‘thrashing’ on overloaded or unhappy clusters)
- mon: ‘report’ command for dumping detailed cluster status (e.g., for use when reporting bugs)
- mon: osdmap flags like noup, noin now cause a health warning
- msgr: improved failure handling code
- msgr: many bug fixes
- osd, mon: honor new ‘nobackfill’ and ‘norecover’ osdmap flags
- osd, mon: use feature bits to lock out clients lacking CRUSH tunables when they are in use
- osd: backfill reservation framework (to avoid flooding new osds with backfill data)
- osd: backfill target reservations (improve performance during recovery)
- osd: better tracking of recent slow operations
- osd: capability grammar improvements, bug fixes
- osd: client vs recovery io prioritization
- osd: crush performance improvements
- osd: default journal size to 5 GB
- osd: experimental support for PG “splitting” (pg_num adjustment for existing pools)
- osd: fix memory leak on certain error paths
- osd: fixed detection of EIO errors from fs on read
- osd: major refactor of PG peering and threading
- osd: many bug fixes
- osd: more/better dump info about in-progress operations
- osd: new caps structure (see compatibility notes)
- osd: new ‘deep scrub’ will compare object content across replicas (once per week by default)
- osd: new ‘lock’ rados class for generic object locking
- osd: optional ‘min’ pg size
- osd: recovery reservations
- osd: scrub efficiency improvement
- osd: several out of order reply bug fixes
- osd: several rare peering cases fixed
- osd: some performance improvements related to request queuing
- osd: use entire device if journal is a block device
- osd: use syncfs(2) when kernel supports it, even if glibc does not

- osd: various fixes for out-of-order op replies
- rados: ability to copy, rename pools
- rados: bench command now cleans up after itself
- rados: ‘cppool’ command to copy rados pools
- rados: ‘rm’ now accepts a list of objects to be removed
- radosgw: POST support
- radosgw: REST API for managing usage stats
- radosgw: fix bug in bucket stat updates
- radosgw: fix copy-object vs attributes
- radosgw: fix range header for large objects, ETag quoting, GMT dates, other compatibility fixes
- radosgw: improved garbage collection framework
- radosgw: many small fixes, cleanups
- radosgw: openstack keystone integration
- radosgw: stripe large (non-multipart) objects
- radosgw: support for multi-object deletes
- radosgw: support for swift manifest objects
- radosgw: vanity bucket dns names
- radosgw: various API compatibility fixes
- rbd: import from stdin, export to stdout
- rbd: new ‘ls -l’ option to view images with metadata
- rbd: use generic id and keyring options for ‘rbd map’
- rbd: don’t issue usage on errors
- udev: fix symlink creation for rbd images containing partitions
- upstart: job files for all daemon types (not enabled by default)
- wireshark: ceph protocol dissector patch updated

V0.54

UPGRADING

- The osd capabilities associated with a rados user have changed syntax since 0.48 argonaut. The new format is mostly backwards compatible, but there are two backwards-incompatible changes:
 - specifying a list of pools in one grant, i.e. ‘allow r pool=foo,bar’ is now done in separate grants, i.e. ‘allow r pool=foo’, ‘allow r pool=bar’.
 - restricting pool access by pool owner (‘allow r uid=foo’) is removed. This feature was not very useful and unused in practice.

The new format is documented in the `ceph-authtool` man page.

- Bug fixes to the new osd capability format parsing properly validate the allowed operations. If an existing rados user gets permissions errors after upgrading, its capabilities were probably misconfigured. See the `ceph-authtool` man page for details on osd capabilities.
- ‘rbd lock list’ and ‘rbd showmapped’ no longer use tabs as separators in their output.

V0.48.3 “ARGONAUT”

This release contains a critical fix that can prevent data loss or corruption after a power loss or kernel panic event. Please upgrade immediately.

UPGRADING

- If you are using the undocumented `ceph-disk-prepare` and `ceph-disk-activate` tools, they have several new features and some additional functionality. Please review the changes in behavior carefully before upgrading.
- The .deb packages now require `xfsprogs`.

NOTABLE CHANGES

- filestore: fix op_seq write order (fixes journal replay after power loss)
- osd: fix occasional indefinitely hung “slow” request

- osd: fix encoding for pool_snap_info_t when talking to pre-v0.48 clients
- osd: fix heartbeat check
- osd: reduce log noise about rbd watch
- log: fixes for deadlocks in the internal logging code
- log: make log buffer size adjustable
- init script: fix for ‘ceph status’ across machines
- radosgw: fix swift error handling
- radosgw: fix swift authentication concurrency bug
- radosgw: don’t cache large objects
- radosgw: fix some memory leaks
- radosgw: fix timezone conversion on read
- radosgw: relax date format restrictions
- radosgw: fix multipart overwrite
- radosgw: stop processing requests on client disconnect
- radosgw: avoid adding port to url that already has a port
- radosgw: fix copy to not override ETAG
- common: make parsing of ip address lists more forgiving
- common: fix admin socket compatibility with old protocol (for collectd plugin)
- mon: drop dup commands on paxos reset
- mds: fix loner selection for multiclient workloads
- mds: fix compat bit checks
- ceph-fuse: fix segfault on startup when keyring is missing
- ceph-authtool: fix usage
- ceph-disk-activate: misc backports
- ceph-disk-prepare: misc backports
- debian: depend on xfsprogs (we use xfs by default)
- rpm: build rpms, some related Makefile changes

For more detailed information, see [the complete changelog](#).

V0.48.2 “ARGONAUT”

UPGRADING

- The default search path for keyring files now includes /etc/ceph/ceph.\$name.keyring. If such files are present on your cluster, be aware that by default they may now be used.
- There are several changes to the upstart init files. These have not been previously documented or recommended. Any existing users should review the changes before upgrading.
- The ceph-disk-prepare and ceph-disk-active scripts have been updated significantly. These have not been previously documented or recommended. Any existing users should review the changes before upgrading.

NOTABLE CHANGES

- mkcephfs: fix keyring generation for mds, osd when default paths are used
- radosgw: fix bug causing occasional corruption of per-bucket stats
- radosgw: workaround to avoid previously corrupted stats from going negative
- radosgw: fix bug in usage stats reporting on busy buckets
- radosgw: fix Content-Range: header for objects bigger than 2 GB.
- rbd: avoid leaving watch acting when command line tool errors out (avoids 30s delay on subsequent operations)
- rbd: friendlier use of -pool/-image options for import (old calling convention still works)
- librbd: fix rare snapshot creation race (could “lose” a snap when creation is concurrent)
- librbd: fix discard handling when spanning holes
- librbd: fix memory leak on discard when caching is enabled
- objecter: misc fixes for op reordering
- objecter: fix for rare startup-time deadlock waiting for osdmap
- ceph: fix usage
- mon: reduce log noise about “check_sub”
- ceph-disk-activate: misc fixes, improvements
- ceph-disk-prepare: partition and format osd disks automatically
- upstart: start everyone on a reboot
- upstart: always update the osd crush location on start if specified in the config
- config: add /etc/ceph/ceph.\$name.keyring to default keyring search path

- ceph.spec: don't package crush headers

For more detailed information, see [the complete changelog](#).

V0.48.1 "ARGONAUT"

UPGRADING

- The radosgw usage trim function was effectively broken in v0.48. Earlier it would remove more usage data than what was requested. This is fixed in v0.48.1, but the fix is incompatible. The v0.48 radosgw-admin tool cannot be used to initiate the trimming; please use the v0.48.1 version.
- v0.48.1 now explicitly indicates support for the CRUSH_TUNABLES feature. No other version of Ceph requires this, yet, but future versions will when the tunables are adjusted from their historical defaults.
- There are no other compatibility changes between v0.48.1 and v0.48.

NOTABLE CHANGES

- mkcephfs: use default 'keyring', 'osd data', 'osd journal' paths when not specified in conf
- msgr: various fixes to socket error handling
- osd: reduce scrub overhead
- osd: misc peering fixes (past_interval sharing, pgs stuck in 'peering' states)
- osd: fail on EIO in read path (do not silently ignore read errors from failing disks)
- osd: avoid internal heartbeat errors by breaking some large transactions into pieces
- osd: fix osdmap catch-up during startup (catch up and then add daemon to osdmap)
- osd: fix spurious 'misdirected op' messages
- osd: report scrub status via 'pg ... query'
- rbd: fix race when watch registrations are resent
- rbd: fix rbd image id assignment scheme (new image data objects have slightly different names)
- rbd: fix perf stats for cache hit rate
- rbd tool: fix off-by-one in key name (crash when empty key specified)
- rbd: more robust udev rules
- rados tool: copy object, pool commands
- radosgw: fix in usage stats trimming
- radosgw: misc API compatibility fixes (date strings, ETag quoting, swift headers, etc.)
- ceph-fuse: fix locking in read/write paths
- mon: fix rare race corrupting on-disk data
- config: fix admin socket 'config set' command
- log: fix in-memory log event gathering
- debian: remove crush headers, include librados-config
- rpm: add ceph-disk-{activate, prepare}

For more detailed information, see [the complete changelog](#).

V0.48 "ARGONAUT"

UPGRADING

- This release includes a disk format upgrade. Each ceph-osd daemon, upon startup, will migrate its locally stored data to the new format. This process can take a while (for large object counts, even hours), especially on non-btrfs file systems.
- To keep the cluster available while the upgrade is in progress, we recommend you upgrade a storage node or rack at a time, and wait for the cluster to recover each time. To prevent the cluster from moving data around in response to the OSD daemons being down for minutes or hours, you may want to:

```
ceph osd set noout
```

This will prevent the cluster from marking down OSDs as "out" and re-replicating the data elsewhere. If you do this, be sure to clear the flag when the upgrade is complete:

```
ceph osd unset noout
```

- There is a encoding format change internal to the monitor cluster. The monitor daemons are careful to switch to the new format only when all members of the quorum support it. However, that means that a partial quorum with new code may move to the new format, and a recovering monitor running old code will be unable to join (it will crash). If this occurs, simply upgrading the remaining monitor will resolve the problem.
- The ceph tool's -s and -w commands from previous versions are incompatible with this version. Upgrade your client tools at the same time you upgrade the monitors if you rely on those commands.
- It is not possible to downgrade from v0.48 to a previous version.

NOTABLE CHANGES

- osd: stability improvements
- osd: capability model simplification
- osd: simpler/safer -mkfs (no longer removes all files; safe to re-run on active osd)
- osd: potentially buggy FIEMAP behavior disabled by default
- rbd: caching improvements
- rbd: improved instrumentation
- rbd: bug fixes
- radosgw: new, scalable usage logging infrastructure
- radosgw: per-user bucket limits
- mon: streamlined process for setting up authentication keys
- mon: stability improvements
- mon: log message throttling
- doc: improved documentation (ceph, rbd, radosgw, chef, etc.)
- config: new default locations for daemon keyrings
- config: arbitrary variable substitutions
- improved 'admin socket' daemon admin interface (ceph -admin-daemon ...)
- chef: support for multiple monitor clusters
- upstart: basic support for monitors, mds, radosgw; osd support still a work in progress.

The new default keyring locations mean that when enabling authentication (auth supported = cephx), keyring locations do not need to be specified if the keyring file is located inside the daemon's data directory (/var/lib/ceph/\$type/ceph-\$id by default).

There is also a lot of librbd code in this release that is laying the groundwork for the upcoming layering functionality, but is not actually used. Likewise, the upstart support is still incomplete and not recommended; we will backport that functionality later if it turns out to be non-disruptive.