OpenZFS on Linux Development

OpenSFS Lustre Developer Meeting
Jan 22th, 2015

Brian Behlendorf
behlendorf1@llnl.gov
OpenZFS on Linux

- Current version 0.6.3 (released June 12th 2014)
- Easy to install packages for many distributions.
- Large enthusiastic user community.
  - zfs-discuss@zfsonlinux.org
  - #zfsonlinux on freenode.net
  - http://zfsonlinux.org
OpenZFS on Linux – Version 0.6.3

- Near feature parity with other OpenZFS implementations.
- Systematically addressing gaps in functionality
- Wide spectrum of users
- Used on diverse hardware
- Contributions (0.6.2-0.6.3)
  - 58 different developers
  - 301 commits

<table>
<thead>
<tr>
<th>Highlights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Updated Write Throttle</td>
</tr>
<tr>
<td>ARC Performance Improvements</td>
</tr>
<tr>
<td>POSIX ACLs</td>
</tr>
<tr>
<td>File Attributes (immutable, append-only)</td>
</tr>
<tr>
<td>Relatime style updates</td>
</tr>
<tr>
<td>SELinux Integration</td>
</tr>
<tr>
<td>Systemd Integration</td>
</tr>
<tr>
<td>ZFS Event Daemon (ZED)</td>
</tr>
<tr>
<td>Aarch64 and Sparc64 Support</td>
</tr>
<tr>
<td>Over 200 Bug fixes</td>
</tr>
</tbody>
</table>
OpenZFS on Linux – Version 0.6.4

- Continue to integrate OpenZFS features
- Continue to address known gaps
- Continue to improve Linux integration
- Contributions (0.6.3-HEAD)
  - 37 different developers
  - 141 commits

<table>
<thead>
<tr>
<th>Planned Highlights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feature Flag: Spacemap Histograms</td>
</tr>
<tr>
<td>Feature Flag: ZFS Bookmarks</td>
</tr>
<tr>
<td>Feature Flag: Hole Birth</td>
</tr>
<tr>
<td>Feature Flag: Embedded Data</td>
</tr>
<tr>
<td>Metaslab Improvements</td>
</tr>
<tr>
<td>Xattr Improvements</td>
</tr>
<tr>
<td>AIO and DirectIO Support</td>
</tr>
<tr>
<td>Falloate Hole Punching</td>
</tr>
<tr>
<td>Linux Tracepoints</td>
</tr>
<tr>
<td>NFS access to .zfs/snapshot</td>
</tr>
<tr>
<td>100 bug fixes and counting</td>
</tr>
</tbody>
</table>
The Road to Version 1.0.0

- Minor releases have a development focus
  - 0.6.x – Functionality / Integration
  - 0.7.x – Memory Management
  - 0.8.x – ZFS+SPL Consolidation
  - 0.9.x – Stable ABI / Hardening
  - 1.0.x – Feature Development / Performance

- Longer term roadmap to guide development
- Current development activities continue in parallel
- Features and bug fixes are merged when ready
OpenZFS on Linux - Version 0.7.x

- **Focus:** Memory Management
- **Goal:** ARC / page cache integration
  - Data buffers will be backed by page vectors
  - Data pages will be mapped in to the page cache
- **Benefits:**
  - Uses standard Linux memory accounting mechanisms
  - Uses standard Linux memory reclaim mechanisms
  - Eliminates fragmentation overhead
  - Eliminates `mmap` double caching
  - 32-Bit platform support (x86, ARM)
OpenZFS on Linux - Version 0.8.x

- Focus: ZFS+SPL Consolidation
- Goal: Merging ZFS+SPL git repositories
  - Existing layering is preserved
  - An opportunity to define the ZFS kernel ABI
- Benefits:
  - Simpler packaging for users and maintainers
  - Eliminates the kmod dependency problem
  - Eliminates the risk of mismatched versions
  - One source tree for developers
  - Improves portability for non-Linux platforms
OpenZFS on Linux - Version 0.9.x

- Focus: Stable ABI / Hardening
- Goal: Finalize a stable user/kernel ABI
- Benefits:
  -.Smopper upgrades / downgrades
  -.Establishes a clear a user/kernel ABI for Linux

- Goal: Hardening
  - Gracefully handle a wider range of potential failure modes
  - Fault management via the ZFS Event Daemon (ZED)

- Benefits:
  - Enables deployment of lower end commodity hardware
  - Even more robust operation
OpenZFS on Linux - Version 1.0.x

- Focus: Feature development / performance
- Goal: Mature high quality code base
- Benefits:
  - Semantic Versioning
  - New feature development
  - Performance analysis
Development Model

- Project hosted at Github
  - [https://github.com/zfsonlinux/](https://github.com/zfsonlinux/)
  - 1201 Watchers, 300 Forks

- illumos is tracked as upstream

- Independent of the Linux Kernel
  - Decouples ZFS from kernel updates
  - Linux 2.6.32 – 3.17 kernels supported
  - Enables use on non-Linux platforms
  - ZFS utilities / kmod can share code
  - Easier to integrate OpenZFS changes from Illumos/FreeBSD/OSX/OSV
Development Model – Issue Tracker

- Github issue tracker
  - Feature requests, bug reports, and milestones
    - Developers actively participate on the tracker
    - 539 open issues including 115 feature requests
  - Everything is as open and public as possible
  - Discussion by users and developers is encouraged
  - Issues are cross-linked to relevant git commits
# Development Model - Milestones

<table>
<thead>
<tr>
<th>Version</th>
<th>Status</th>
<th>Due Date</th>
<th>Last Updated</th>
<th>Description</th>
<th>Completion</th>
<th>Open</th>
<th>Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6.4</td>
<td>6 Open</td>
<td>No due date</td>
<td>Last updated 27 minutes ago</td>
<td>ZFS on Linux v0.6.4 - Functionality / Integration</td>
<td>74%</td>
<td>101</td>
<td>284</td>
</tr>
<tr>
<td>0.7.0</td>
<td>15 Closed</td>
<td>No due date</td>
<td>Last updated 3 minutes ago</td>
<td>ZFS on Linux v0.7.0 - Memory management</td>
<td>51%</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>0.8.0</td>
<td></td>
<td>No due date</td>
<td>Last updated 29 minutes ago</td>
<td>ZFS on Linux v0.8.0 - ZFS+SPL Consolidation</td>
<td>32%</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>0.9.0</td>
<td></td>
<td>No due date</td>
<td>Last updated 1 minute ago</td>
<td>ZFS on Linux v0.9.0 - Stable ABI / Hardening</td>
<td>17%</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>1.0.0</td>
<td></td>
<td>No due date</td>
<td>Last updated less than a minute ago</td>
<td>ZFS on Linux v1.0.0 - Feature Development / Performance</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Development Model – Pull Requests

- Github pull requests
  - Used to submit proposed code changes
  - All proposed changes must be reviewed
  - Continuous integration development model
    - Proposed changes are automatically tested using buildbot
    - Developers get quick feedback on any proposed change
    - Good test coverage (kernel, architecture, distribution, etc)
    - Changes are tested a second time after being merged
    - The master branch is *always* kept stable
  - Designed to make it easy for anyone to contribute
Development Model – Pull Requests

Linux tracepoint integration #2874

nedbass wants to merge 4 commits into fsonlinux:master from nedbass:b_tracepoints

Conversation 0  Commits 4  Files changed 18

nedbass commented a day ago

See #2406

Changes from the PR by @prakashsurya:

- Implement `dprintf()` as a tracepoint.
- Implement two new `DTRACE_PROBE`'s
- Refactor patch stack and make it cstyle-compliant
- Address feedback from @behlendorf on #2406

One thing I'm concerned about is the 256 byte stack buffer used by `dprintf()`. While short-lived this could potentially push stack usage over the limit. I've seen messages as long as 202 bytes to this isn't an unreasonable size, but we may want to consider dynamic allocation if this turns out to be problematic.

nedbass and others added some commits 17 days ago

- Move a few internal ARC structures to arc_impl.h
- Fix `dprintf` format specifiers
- cstyle: allow right paren on its own line
- Swap `DTRACE_PROBE*` with Linux tracepoints

- f772ef3
- 4d0e3fc
- 1be6e5c
- 7bde2d2

3 participants
Development Model - Buildbot

Console View

Branch: master

Legend: Passed | Failed | Warnings | Failed Again | Running | Prebuilt | Offline | No data

SP1bd3ted... Dan Swartzentruber

- Improve and symlink handling.
- Change the src helper program to replace any embedded spaces in the path or dataset names with `\` to ensure we have valid symlinks.
- The `\` character was chosen because it is not a valid character for a dataset name but it is allowed by udev. This ensures that all dataset names with an embedded space will be translated to a unique `/dev/zone` symlink.
- Signed-off-by: Dan Swartzentruber <dswartz@LBL.com>
- Signed-off-by: Dan Swartzentruber <dswartz@LBL.com>
- Signed-off-by: Brian Behlendorf <bbehendorf@LLNL.gov>
- Closed #3834

110022f95bc... Marcel Wysocki

- Add config/compile to config/ignore
- This file may be added by autotools and therefore should be added to config/ignore. For the full list of possible auxiliary programs see the full autotools documentation.
- http://www.gnu.org/software/automake/manual/automake.html#Auxiliary-Programs
- Signed-off-by: Marcel Wysocki <marcel.stipa@gmail.com>
- Signed-off-by: Brian Behlendorf <bbehendorf@LLNL.gov>
- Closed #38348

b03880c564... Alexander Pyshlov

- Fix modules installation directory
- When building src modules with kernel, compiled from dotsrc, the packaging process ends up installing the modules in the wrong place.
- Signed-off-by: Alexander Pyshlov <ap@phrase.com>
- Signed-off-by: Brian Behlendorf <bbehendorf@LLNL.gov>
- Closed #3822

b262070e23c... Richard Yao

- Make systemd-modules-load.service file directory configurable
- Installing outside of the prefix is not permitted under CentOS Prefix.
Join Us, Contributors Welcome

▪ If you are a developer…
  • Port a change from Illumos/FreeBSD/OSX/OSV
  • Review or comment on a proposed pull requests
  • Implement a requested feature or fix a known issue
  • Help us improve the automated testing

▪ If you are a user…
  • Open a new issue if you encounter a problem
  • Open pull requests even for trivial fixes
  • Help us rigorously test new features and bug fixes
Questions