OAK RIDGE National Laboratory

Upstream efforts progress

Report of milestones, May 2026

James Simmons

Storage Systems Engineer

Oak Ridge National Laboratory

ORNL is managed by UT-Battelle LLC for the US Department of Energy

U.S. DEPARTMENT OF ENERGY

Project summary

- Heavily demanded project
- large community support
  - Amazon
  - Aeon
  - ORNL
  - HPE
  - Lustre collective
- Work done using two trees
  - git clone git://git.whamcloud.com/fs/lustre-release.git
  - git clone git://git.github.com/jasimmons1973/linux.git
    - creating lustre branch

What is needed for upstreaming

- Need a key for the Web of Trust
- Validate native client work on OpenSFS branch
  - Most of this work is done
- Support for building upstream code in OpenSFS tree
- Sync native client to OpenSFS tree.
  - GSS is missing. Working on restoring it.
  - Update to latest kernel
    - Fix up fscrypt support
    - Folio is a requirement
- Migrate to vger might be necessary

OpenSFS kernel support state

- Support of latest kernel (6.19 / 7.0)
- All pre 5.0 kernel support is being removed
- o2iblnd uses GDS not supported upstream
  - p2pdma is in the works
- Folio support is in the works
- dcache fixes can now be worked on
- Native fscrypt will soon be supported
- Only client side ldlm still doesn't use rhashtables
- Only md_stats and pools proc files left for the client
  - Support Secure Boot Linux using Netlink / YAML / pyYAML + go + others

Finish IPv6

- Was done for upstreaming efforts
- Several pieces still missing
  - IB IPv6 support
  - LNet selftest
  - UDSP support
  - Proper lustre UUID to NIDs mapping
  - IPv6 support for GSS
  - Other regressions

OpenSFS source tree changes

- Automake cleanup
  - Moving to kernel makefile
- Separate kernel and non kernel items. Example Documentation
- Creation of lustre_compact
  - include/lustre_compact
    - #include <lustre_compact/linux/*.h>
  - Lustre_compact *.c code
- Breakup libcfs and compat split
  - Still complied as one module
- Potential reduction of modules

Side products of this work

- Separate mount targets. Please use mount –t lustre_tgt
- Move to fs_context API. Generic server mount handling
- Use of ktest to validate OpenSFS tree against latest kernels
- New debug work
  - Use new upstream memory leak framework (/proc/allocinfo)
  - Embedded BTF in lustre modules
  - LNet eBPF tracing
- Profiling changes (stats etc)
  - YAML Netlink based instead of file open in proc or debugfs
  - Python, go, rust wrappers
    - lljobstats and llobdstat will be examples

Special thanks

- Native client support is a true community effort
  - Neil Brown (SuSE)
  - Arshad Hussain (Aeon computing)
  - Shaun Tancheff (HPE)
  - Timothy Day (lustre collective)
  - Whamcloud team

Conclusion

- Preparation for working with Linux Lustre tree is progressing
- Important milestones are within reach
- New features and improvements from this work that benefits the user community
- Healthy momentum to support this work

Acknowledgments

This work was performed under the auspices of the U.S. DOE by Oak Ridge Leadership Computing Facility at ORNL under contract DE-AC05-00OR22725.