EMC Contributions: Mainlining Lustre Client & Patches

EMC Contributions to OpenSFS

Peng Tao, Sorin Faibish, John Bent, and Sassan Teymouri

LUG Meting
April 2013, San Diego

EMC² where information lives

Lustre Client for Linux Kernel

- A collaboration of EMC with Intel/Whamcloud
- Other contributors
  - Intel, Cray, ORNL, LLNL, SUSE, Gentoo
- EMC contribution in 2012
  - 10+ LU tickets (LU-709, LU-1347, LU-1337, LU-1214, LU-1756, LU-1113, LU-1994, LU-2850, LU-2335, etc.)
  - 80+ patches merged
  - 20+ pending review (Please help!)
  - ~20K LOC touched

Lustre Patches/Fixes by EMC

Peng Tao and Xuezhao Liu; EMC China

- Over 50 patches to mainline since LUG 2012
  - libcfs cleanup: 2 patches
  - server/client split: 5 patches
  - compat25.h cleanup: 7 patches
  - auto-configure macro cleanup: 16 patches
  - Linux Next 3.9 kernel support: 22 patches

Current Status of Lustre Client to Linux next

- So far
  - Coding style change (done with checkpatch.pl.)
  - Server/client code split (done)
  - Ptlrpc cleanup(done)
  - Obsolete macro cleanup (done)
  - Ported to latest upstream kernel (done)
  - Old code removal (done with Coan scripts)
  - Kernel Kbuild/Kconfig support (done with scripts)
  - Client kernel code extraction (done with scripts)
- Remaining
  - Libcfs cleanup (half done)
  - Obdclass cleanup (half done)
  - Other minor cleanups

Availability

- Lustre master branch
  - Apply patches in LU-2335
  - Build lustre client code as external modules –or-
  - Copy Lustre client code into kernel tree
- Git
  - Latest Linux kernel tree patched with Lustre client source
  - git clone git://github.com/bergwolf/linux.git
  - git checkout –b 3.9-rc3-lustre
  - export CONFIG_LUSTRE_FS; make

LSF-2012 Recommendations

Peng Tao (Bergwolf) - EMC made the case

- Check in staging area to clean ready
  - Greg KH offered help
  - Must be compliable
  - Must be properly approved
- Recommended Steps (Intel Linux team support):
  - Huge patches merged in staging tree first
  - Put Lustre code in fs directory and depend on staging
  - Small and incremental patches to cleanup
  - New feature patches are allowed (flash support)
  - Each patch properly signed-off by key reviewers
  - Merge code into main stream Linux kernel

Future Work

Need Support of Lustre Community

- Submit code to staging tree
- Continued maintenance and hardening
  - Need to merge patches between upstream and Lustre master back and forth.
- Splitclient utility build
  - Or an option to build utility only
    - i.e. without building client/server kernel modules

Questions?

- Visit EMC booth for more info about this and our other Lustre/HPC initiatives
  - Burst buffer appliances for Lustre acceleration
  - vHPC for virtual storage and advanced HPC scheduling
  - Burst buffer research with DAOS for DOE FastForward
  - Massive small file parallel IOPS for Lustre
  - Umbrella file systems for MDS balancing
  - Suitability of ATMOS for long-term archival storage
  - Contributions to Openbenchmark
  - pNFS Lustre layout draft submitted to IETF

THANK YOU

EMC² where information lives

THANK YOU