SquashFS & FUSE for Better HPC Containers

Authors: Megan Phinney, Iowa State University; Anna Chernikov, NC State University

Mentors: Jordan Ogas, Reid Priedhorsky

LA-UR 21-27656

Charliecloud is a light-weight container implementation for high performance computing. The typical filesystem image formats for Charliecloud are SquashFS and tar archives. SquashFS is a compressed, read-only filesystem that unprivileged users can mount in user space with SquashFUSE; it is the preferred image format due to its various efficiencies. The current SquashFS workflow is non-ideal due to user complexity and difficulties with HPC job schedulers.

We have designed a new workflow that requires us to link SquashFUSE to Charliecloud to enable the new mount/unmount procedure of only needing a single user command. Also, an additional persistent process is needed to service the FUSE requests called the FUSE loop: once the containerized application process finishes, it unmounts the SquashFS and ends the FUSE loop. Last summer, we created a working prototype with a modified version of the SquashFUSE code. This summer I am converting our prototype for production use in Charliecloud using SquashFUSE's new shared library. Our new SquashFS workflow is more user friendly, cleans up after itself and is more compatible with HPC job schedulers. We were able to reduce user commands from 3 to 1, increase reliability and decrease mount/unmount time by more than 50%.