11 May 2021

Dave Dykstra, Fermilab

The Transition from Singularity to Apptainer

The Singularity saga

- singularity-2 was written in C by Greg Kurtzer at LBNL in 2015, with many community contributions
  - WLCG got involved in 2016, after Brian Bockelman submitted some enhancements to make it work better from pilot jobs
- Greg found an investor that provided funding, and he founded Sylabs in 2017
  - Greg hired ~10 software developers, who wrote singularity-3 from scratch in Go and added many new features
- In May 2020, Greg left Sylabs and founded new startup company CIQ with different investors and at the same time founded a new open community organization HPCng
  - The two most knowledgeable developers went with him to CIQ, including the one who wrote most of the Go core
  - Greg was permitted to retain leadership of the singularity project, and moved the github repository to https://github.com/hpcng/singularity instead of https://github.com/sylabs/singularity
- In May 2021, Sylabs left the community project and created their own fork, moving it back in github
  - The developers at CIQ weren't contributing much, focusing on other things, and Sylabs had two remaining active developers who were frustrated by lack of progress of HPCng and perceived lack of support by Greg
  - Greg also founded Rocky Linux in December 2020 and so was spread very thin
  - This split created much confusion, especially because Sylabs kept the same name, calling their open source project SingularityCE which many people understandably could not distinguish from the main Singularity project
  - At this time I got more involved in the main Singularity project, getting commit privileges and becoming the leader of the security team
    - Serious consideration was given to abandoning the main project and letting Sylabs own it

The move to Linux Foundation

- In November 2021, the Singularity project joined the Linux Foundation (LF), which required a name change so they could trademark the brand
  - The name “Apptainer” was chosen from a few options by those in the community who volunteered to be involved
  - The primary reason for that choice was that it was a mostly unused word and so was easier to protect
- The LF required a charter including a Technical Steering Committee, not 1 leader
  - The TSC members are the people with commit/approval privileges on the code
  - Started with 5 members: Greg, the two former Sylabs employees that work for him, a supporter from LBNL (who has since left LBNL), and me
- The LF does not automatically come with many resources other than legal, but Greg hired a contractor to help with the transition and to move on to develop new improvements
  - Greg also has recently obtained additional funding and expects to dedicate 1 or 2 people for Apptainer development
  - I have been making sure that most of the changes that Sylabs puts into SingularityCE get imported into Apptainer, generally filling in where needed, and doing some development

Apptainer differences

- The command name changed to “apptainer”, but a ”singularity” symbolic link is included in the package
- The container format is unchanged, both the “sandbox” form we use and SIF image files
- The LF did not want any interfaces included that were not open and were only for support of a single company
  - Sylabs had invented a “container library” protocol and sells services for storing containers, and that was initially expected to be removed
    - There was however a separate, open source server implementation, and I led an effort to create a standards organization for that protocol so it was allowed to remain
    - We did, however, remove the Sylabs server as the default “remote” for reading and writing with the ”library://” url
  - Sylabs also sells access to a remote builder (accessed with `singularity build –remote`) which was convenient for people who didn’t have root access on any machine, since building containers requires privileges for security reasons
    - There was no separate implementation of that service, so that was removed
  - The third service associated with a “remote” was a PGP key server used for signing SIF images, and the default for that was replaced with keys.openpgp.org

Apptainer differences (2)

- Version number restarted at 1.0.0
- Configuration moved to /etc/apptainer/apptainer.conf and ~/.apptainer
  - The system configuration is not automatically converted, but there's a warning if /etc/singularity exists
  - ~/.apptainer is automatically imported from ~/.singularity if it doesn't exist
- Environment variables prefer APPTAINER_ prefix
  - SINGULARITY_ prefixes are also supported, but if only those are set it prints a deprecation warning
  - If both APPTAINER_ and SINGULARITY_ prefix variables set to same value, there is no warning
    - If they are set to different values there is a different warning, and the APPTAINER_ value is used
    - Singularity wrapper scripts will need to especially be aware of this, because if they only set SINGULARITY_ prefixes but the environment already has APPTAINER_ prefixes, the wrapper script settings may be ignored
- Includes most SingularityCE 3.9.5 changes except where the Apptainer TSC disagreed
  - In particular, Apptainer does not allow running the new –nvccli option in privileged mode because it would run an external program as root
- Has an additional new feature of supporting checkpointing through DMTCP (Distributed MultiThreaded CheckPointing)

Upcoming focus

- Focus for the next Apptainer minor release (1.1.0) is to be by default unprivileged, which is not so important for HTC but it is important for HPC and strategic for the continued health of the project because of the podman alternative
  - SIF files mounted with squashfuse in unprivileged user namespace
  - Unprivileged overlay with recent kernel overlayfs or fuse-overlayfs
  - Unprivileged cgroups v2
  - The apptainer rpm will not have a setuid program; it will be in an apptainer-suid rpm
  - Planned to be released in EPEL, and yum update will replace singularity rpm
- Sylabs is instead focused on OCI/docker/podman compatibility
  - Apptainer TSC is reserving judgment on that and does not intend to change any defaults
    - suspicious of OCI due to need for setuid newuidmap & newgidmap and per-user config in /etc/subuid and /etc/subgid, and default use of unprivileged network namespaces (basis for all recent unprivileged namespace CVEs)
  - SingularityCE latest release candidate added an experimental option to mount SIF with squashfuse outside of a namespace (requires setuid fusermount), claiming that will fit better with future OCI compatibility