SEA ‘18 Containers in HPC Symposium

Modifying the Singularity Installation on Comet

Trevor Cooper, SDSC

SDSC Data Enabled Scientific Computing

- Visualization, High Performance Systems, User Services, Advanced Technology Lab, Scientific Computing Applications

- XSEDE Level 1 Service Provider (Comet)
- UCSD Campus Condo Cluster (TSCC)
- Private Cluster (Gordon, previously XSEDE)

Background

Assumptions

- Users are not trusted
- Container contents are not trusted
- Some form of privilege escalation is required

Comet Facts

- Not everybody needs full virtualization … *Comet* VC
- *Comet* deployed OS/kernel does not support user namespaces
- Singularity deployment makes use of SUID binaries
- SDSC security team restricts SUID binary usage to minimize risk

SDSC HPC Software Installation

- **Systems installed with Rocks® Cluster Distribution**
  - Appliance Types / Nodes defined in DB
  - Rocks® DAG selects RPMs and orders post-install configuration
- **Software is built into and managed with Rocks® Rolls**
  - **Basic…** configure, make, make install to produce RPMs
  - **Intermediate…** source patches, rpmbuild –rebuild, etc…
  - **Advanced…** bootstrap, roll dependencies, RPM.EXTRAS, etc…

singularity-roll

- Defines RPM destinations for various node types
- Pulls singularity source tarball from internal server
- Extracts singularity source and patches RPM spec file
- Builds RPMs for singularity using patched spec file
- Bundles RPMs and Rocks® node and graph definition files into ISO

sdscsec-roll

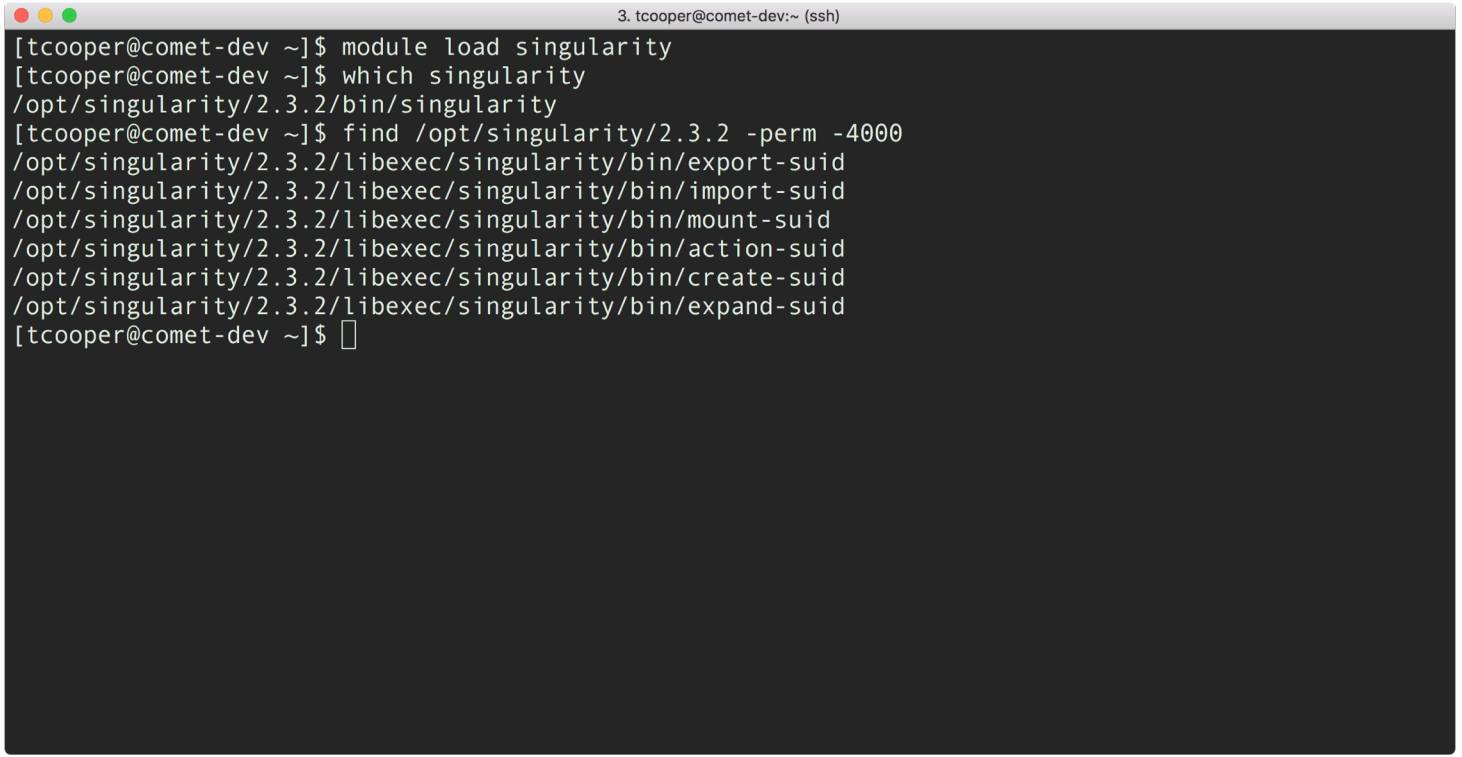
- Replaces distro provided software with custom builds (openssh, sudo, etc…)
- Deploys security related configuration (rsyslog, ssh{d}, certificates, sudo, etc…)
- Removes unwanted permissions (SUID binaries) during post-install and/or first boot

comet-config-roll

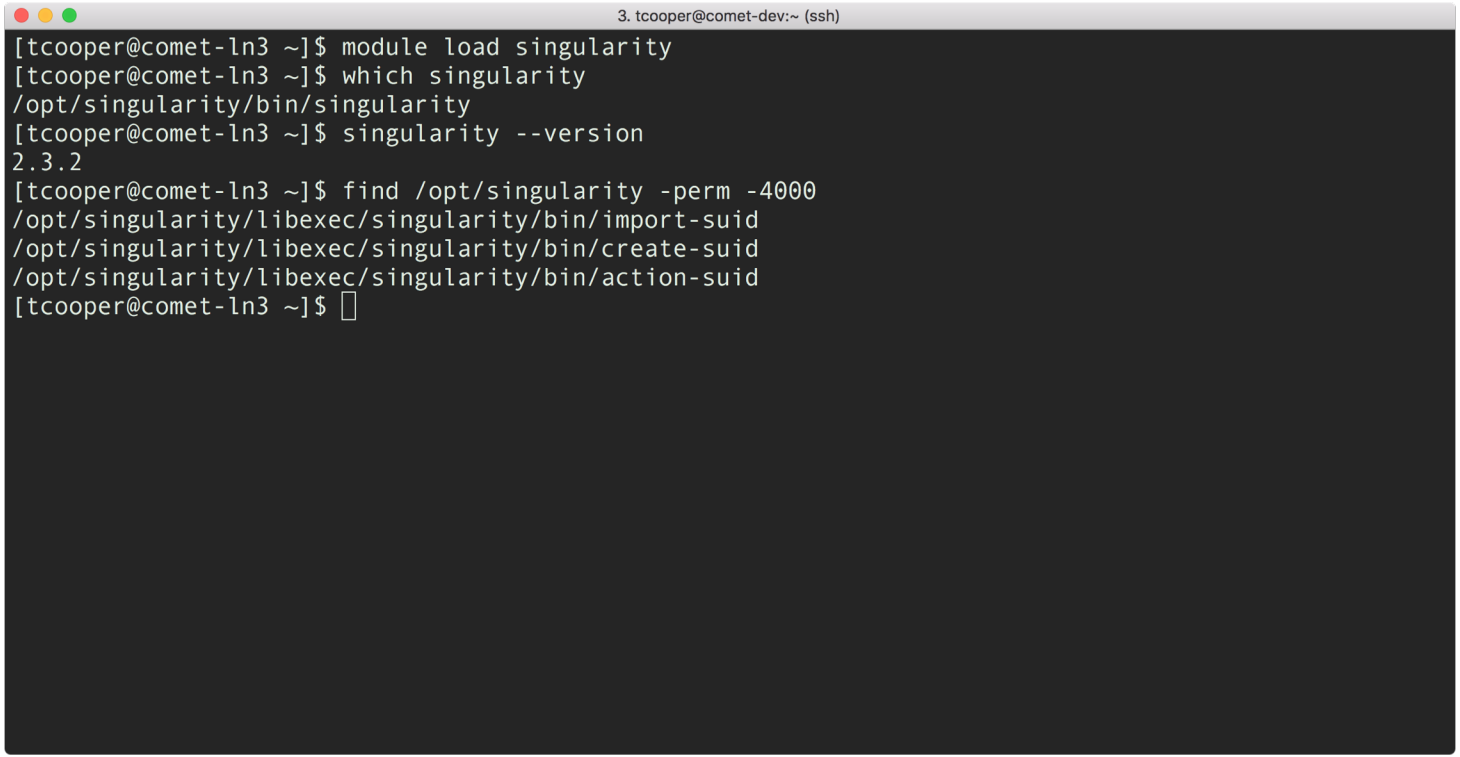
- Deploys Singularity configuration file to frontend
- Singularity configuration file is (re)distributed by Rocks® 411 process

Details

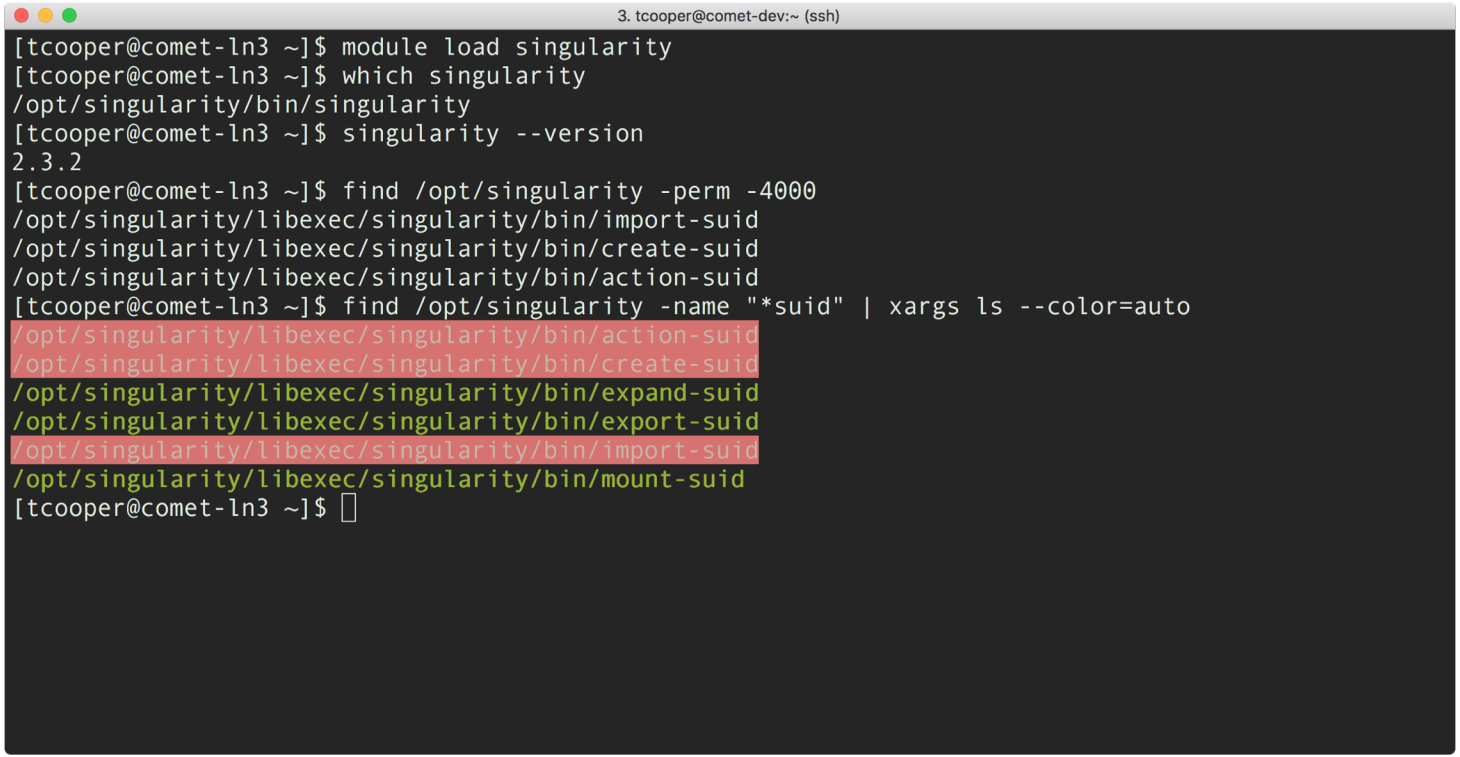
Where are the SUID Components (dev)?



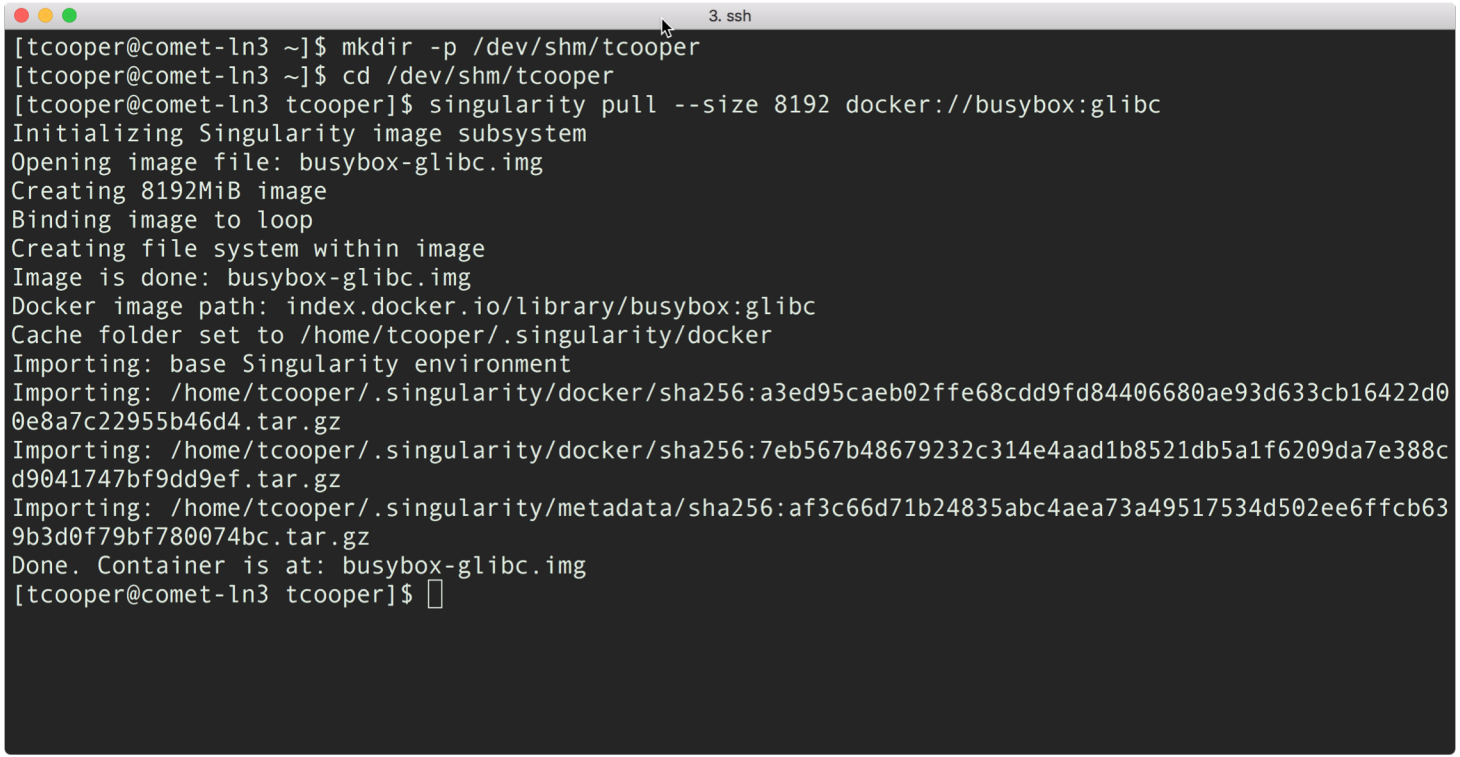
Where are the SUID Components (prod)?



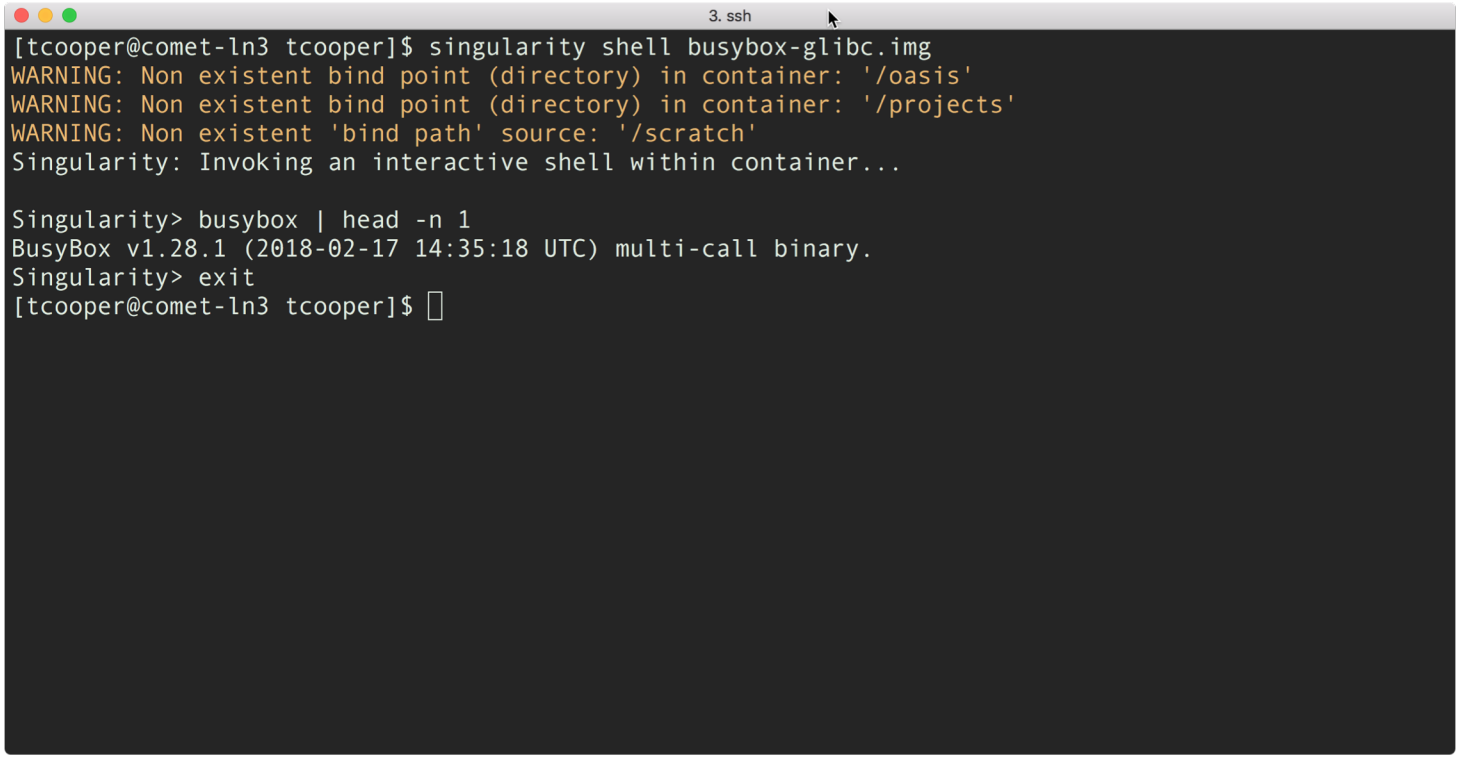
Where are the SUID Components (prod)?



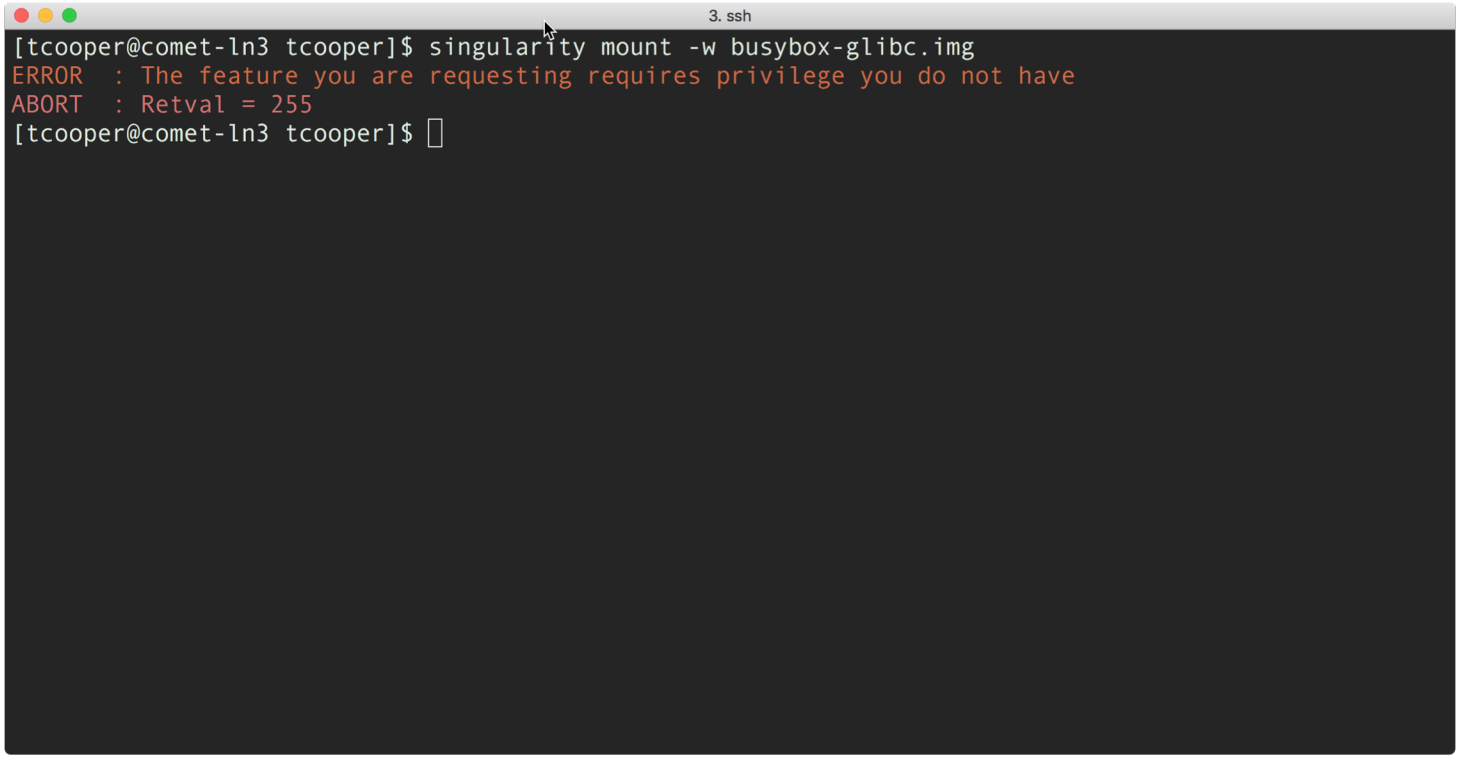
create-suid + import-suid together enable container pull…



action-suid enables container exec|run|shell…



…but you cannot mount or modify a singularity image…



Modifications

security-roll SUID whitelist (v.2.3)

3. root@comet-ln3:~ (ssh)

```
#!/bin/bash

# suid-filter.sh - removes setuid bit from files that are not in the whitelist
# below.
# This will only check filesystems of the types listed below.

MOUNTLIST=`mount | egrep -i ' type ext3 | type ext4 | type ext2 | type xfs | type tmpfs ' | sed 's/^
.* on \(.*\) type .*$/\1/'`

find -H $MOUNTLIST -mount -type f -perm /4000 -a ! \( \
-wholename /bin/ping6 -o \
...snip...
-wholename /opt/singularity/libexec/singularity/sexec-suid -o \
-wholename /opt/singularity/libexec/singularity/bin/action-suid -o \
...snip...
\) \
-exec logger "$0 Removing setuid bit from {}" \; \
-exec echo "Removing setuid bit from {}" \; \
-exec chmod u-s {} \;
~
~
~
~
~
                                                              1,1           All
```

Patching RPM spec during build (v.2.3)

3. tcooper@comet-dev:~ (ssh)

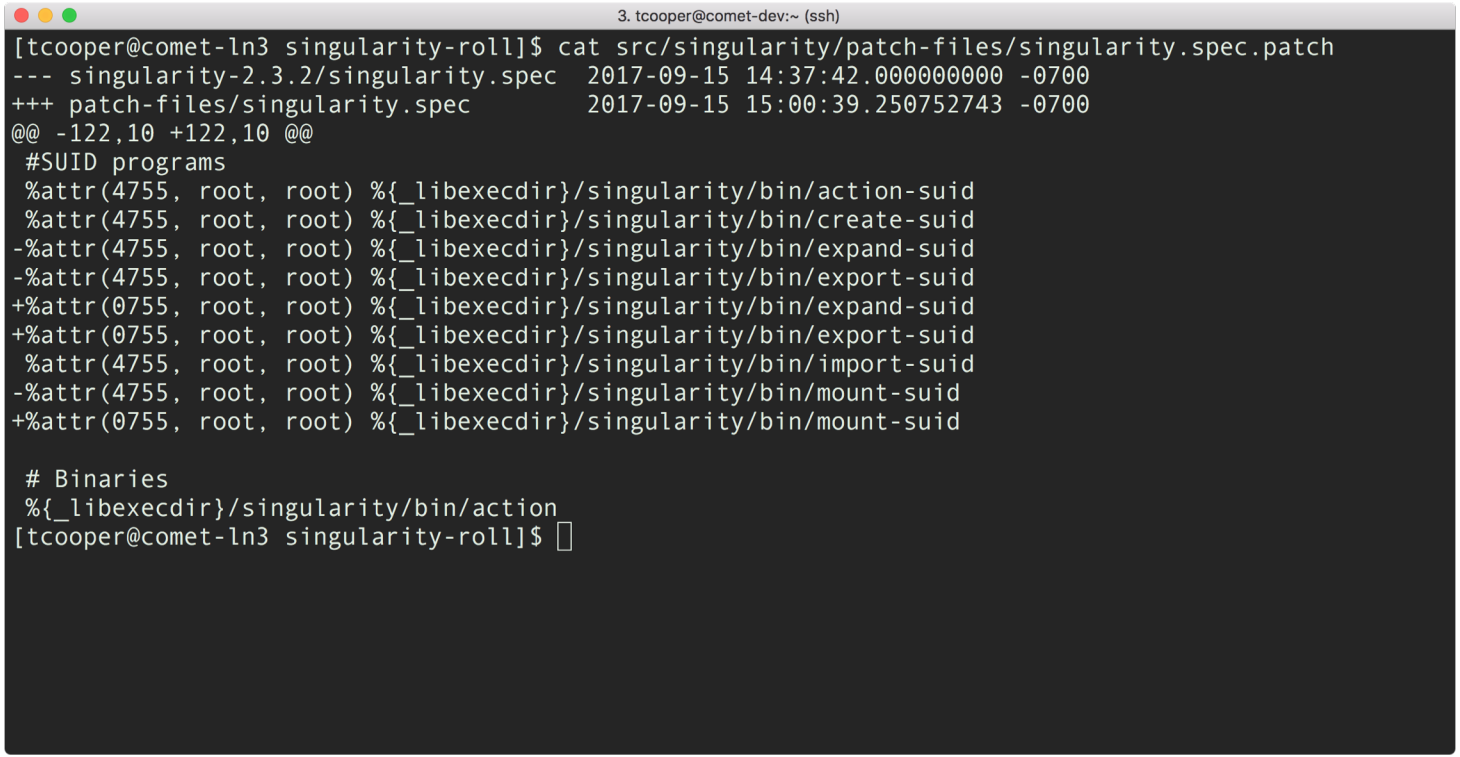
```
[tcooper@comet-ln3 singularity-roll]$ cat src/singularity/patch-files/singularity.spec.patch
--- singularity.spec    2017-06-15 15:02:45.000000000 -0700
+++ patch-files/singularity.spec        2017-06-15 15:03:35.000000000 -0700
@@ -121,12 +121,12 @@

 #SUID programs
 %attr(4755, root, root) %{_libexecdir}/singularity/bin/action-suid
-%attr(4755, root, root) %{_libexecdir}/singularity/bin/create-suid
-%attr(4755, root, root) %{_libexecdir}/singularity/bin/copy-suid
-%attr(4755, root, root) %{_libexecdir}/singularity/bin/expand-suid
-%attr(4755, root, root) %{_libexecdir}/singularity/bin/export-suid
-%attr(4755, root, root) %{_libexecdir}/singularity/bin/import-suid
-%attr(4755, root, root) %{_libexecdir}/singularity/bin/mount-suid
+%attr(0755, root, root) %{_libexecdir}/singularity/bin/create-suid
+%attr(0755, root, root) %{_libexecdir}/singularity/bin/copy-suid
+%attr(0755, root, root) %{_libexecdir}/singularity/bin/expand-suid
+%attr(0755, root, root) %{_libexecdir}/singularity/bin/export-suid
+%attr(0755, root, root) %{_libexecdir}/singularity/bin/import-suid
+%attr(0755, root, root) %{_libexecdir}/singularity/bin/mount-suid

 # Binaries
 %{_libexecdir}/singularity/bin/action
[tcooper@comet-ln3 singularity-roll]$
```

security-roll SUID whitelist (v.2.3.2)

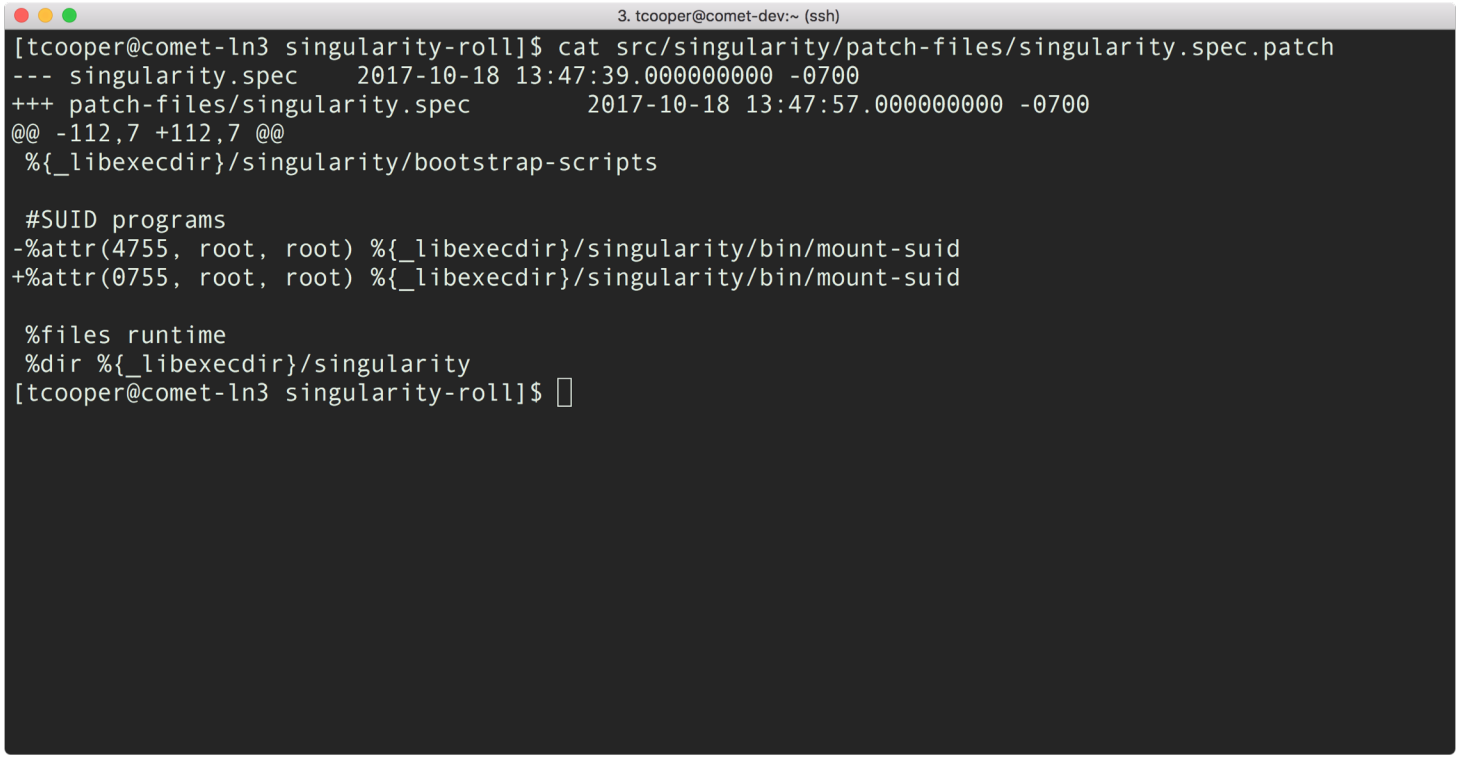
3. root@comet-ln3:~ (ssh)

```
#!/bin/bash

# suid-filter.sh - removes setuid bit from files that are not in the whitelist
# below.
# This will only check filesystems of the types listed below.

MOUNTLIST=`mount | egrep -i ' type ext3 | type ext4 | type ext2 | type xfs | type tmpfs ' | sed 's/^
.* on \(.*\) type .*$/\1/'`

find -H $MOUNTLIST -mount -type f -perm /4000 -a ! \( \
-wholename /bin/ping6 -o \
...snip...
-wholename /opt/singularity/libexec/singularity/sexec-suid -o \
-wholename /opt/singularity/libexec/singularity/bin/action-suid -o \
-wholename /opt/singularity/libexec/singularity/bin/create-suid -o \
-wholename /opt/singularity/libexec/singularity/bin/import-suid -o \
...snip...
\) \
-exec logger "$0 Removing setuid bit from {}" \; \
-exec echo "Removing setuid bit from {}" \; \
-exec chmod u-s {} \;
~
~
~
                                                                1,1           All
```

Patching RPM spec during build (v.2.3.2)



Patching RPM spec during build (v.2.4)



Issues

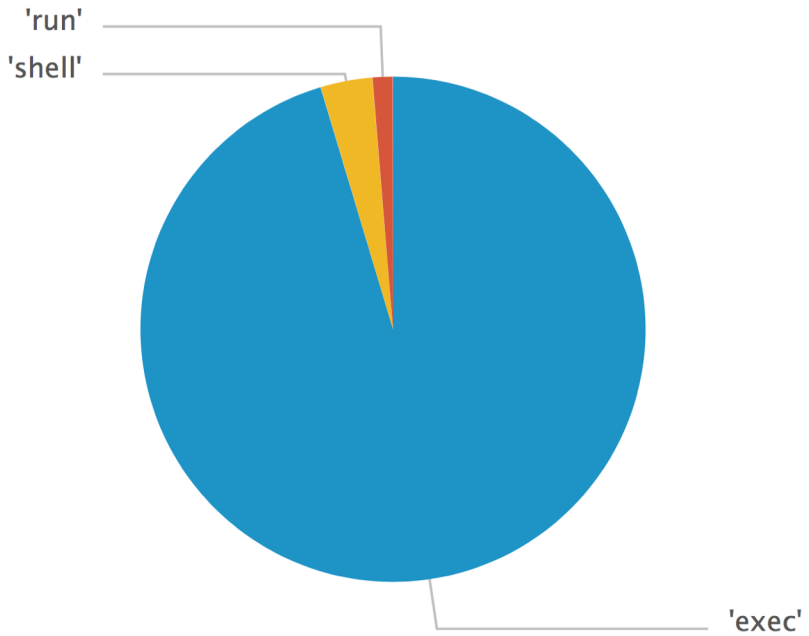
- **For each Singularity version we've needed to…**
  - Examine behavior of new SUID binaries
  - Update sdscsec-roll SUID whitelist in advance of rollout on multiple systems
  - Update singularity-roll build as SUID binaries have changed
- **Security updates…**
  - Early notification from Singularity Team
  - Double updates

Future

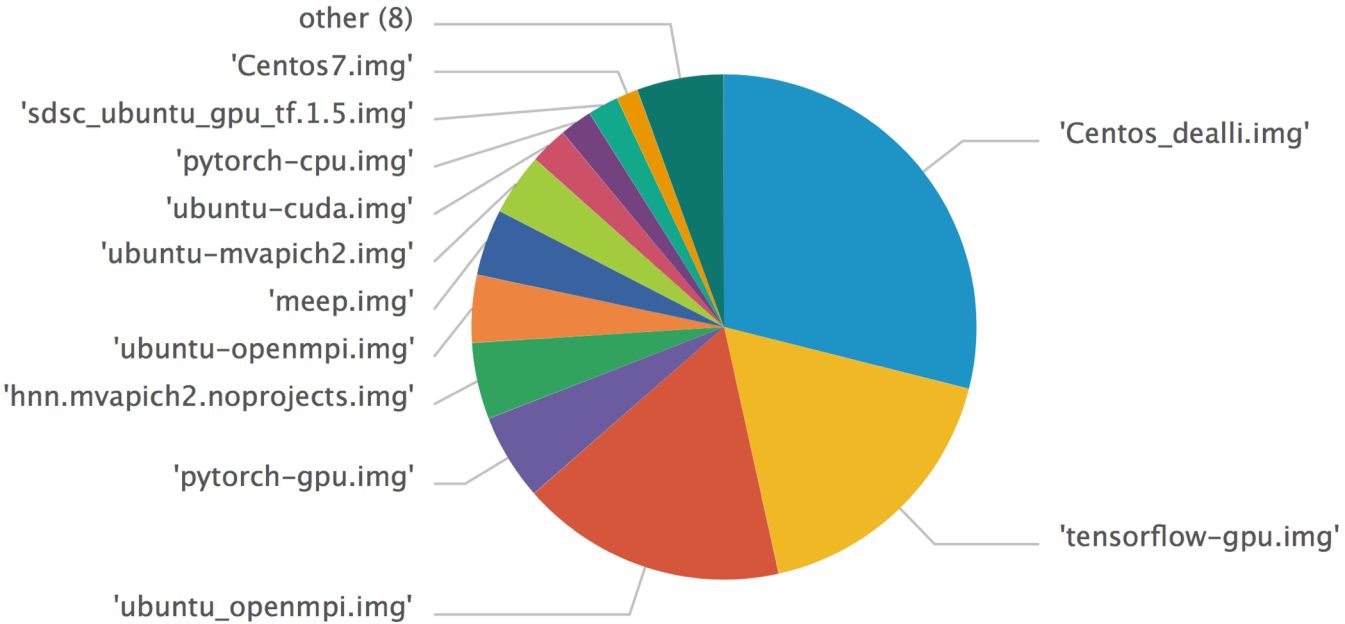
- **Singularity 3.x…**
  - Portions of Python and Bash code base will be replaced by Go
  - SUID binaries are C based and should remain
- **Comet upgrade to ROCKS7 / CentOS 7.4+**
  - User Namespace Support???

THANK YOU

Comet container usage (30 days)



Comet container usage (30 days)



Comet container usage (30 days)