Secure Linux Containers

Application Sandboxes

- Isolate general purpose applications
- Target specific use cases
- Variety of approaches
  - Seccomp – Linux syscall restriction
  - Java VM – bytecode verification
  - SELinux – MCS isolation
  - Virtualization – OS separation
- Multiple layers of defense

Linux Containers

- What is a container?
  - Most people think LXC
    - We will use libvirt-lxc rather then lxc command set.
  - Linux namespaces

Namespaces

- pam_namespace - RHEL5/Fedora 6
- SELinux sandbox - RHEL6/Fedora 8
- Systemd - Fedora 17
  - UnitFile: PrivateTmp, PrivateNetwork
- Openshift - RHEL6
  - Pam_namespace : Private /tmp

Linux Namespaces

- Mount : mounting/unmounting filesystems
- UTS : hostname, domainname
- IPC : SysV message queues, semaphore/shared memory segments
- Network: IPv4/IPv6 stacks, routing, firewall, proc/net /sys/class/net directory trees, sock
- Pid: Own set of pids
- UID: Not implemented yet.

libvirt

- Standard, simple, secure C API
- API bindings
  - Perl, Python, Java, etc
- Mapping to object models
  - SNMP, GObject, CIM, QMF
- Remote RPC access
  - SSH, TLS, GSSAPI

libvirt-lxc

- Container virtualization
- Boot “init” binary
- sVirt SELinux TE + MCS
- Firewall ebtables/ip[6]tables
- Host FS passthrough bind mounts
- CGroups resource control

libvirt-sandbox API

- Based on GObject object system
- Uses libvirt-{glib,gconfig,gobject}
- Accessible from non-C via introspection
- All CLI tools built on top of the API

Example: Server Virtual Hosting

- Goal:
  - Deploy multiple Apache virtual hosts
  - Strong isolation between virtual hosts
- Solution:
  - One apache instance per virtual host
  - Run apache inside a sandbox

virt-sandbox-service

- virt-sandbox-service create -C -u httpd.service apache1
  - Config /etc/libvirt-sandbox/service/apache1.sandbox
  - Multiple unit files allowed
  - SystemD unit file
    - /etc/systemd/system/httpd@apache1.service
  - Create state directories or image
    - /var/lib/libvirt/filesystem/apache1
      - Chroot type directory
    - Examines rpm payload
    - Clone - /var and /etc config
    - Share /usr
  - Allocate unique MCS security label

virt-sandbox-service

- virt-sandbox-service start apache1
  - Starts service from config
- virt-sandbox-service stop apache1
  - Stop service
- virt-sandbox-service connect apache1
  - Connect admin debug shell to container
- Virt-sandbox-service execute -C ifconfig apache1
  - Execute command within container
  - virt-sandbox-service.logrotate
    - /usr/bin/virt-sandbox-service execute -C /etc/cron.daily/logrotate $i

Systemd

- systemctl start httpd@apache1.service
- systemctl reload httpd.service
  - Should trigger reload in all httpd@ services
  - ReloadPropagatedFrom=httpd.service
- Systemctl start httpd@.service
  - Should start all httpd services

Other Use cases

- Run mock within a container
- Run customer services on gluster nodes
- Run mysql within a container
- OpenShift work loads

Demo

- libvirt-0.10.0-0rc0.2.fc18.x86_64
- libvirt-sandbox-0.1.0-1.fc18.x86_64
- selinux-policy-3.11.1-7.fc18.noarch