Running CentOS on the Facebook fleet

Davide Cavalca
Production Engineer - Operating Systems
Agenda

- Infrastructure
- CentOS at Facebook
- Deployment and management
Infrastructure

How does it work?

- OS team manages the fleet bare metal experience
- OS as a platform
- Individual teams are responsible for their own hosts
- Built on an Open Source foundation
- Linux, CentOS, rpm/yum/dnf, Chef, systemd
Infrastructure

Upstream first

- Community sets the direction
- We move fast; Open Source often moves faster
- We don’t need to write everything ourselves
- Sharing our code means sharing the maintenance and having others extend it
- DevConf.CZ 2017 talk: https://tinyurl.com/y7gx6nro
CentOS at Facebook
**CentOS at Facebook**

Why CentOS?

- Stable releases
- Binary compatibility
- Security updates
- Mature and well understood tooling
- EPEL
- Close relationship with Fedora
Backports from Fedora Rawhide for stuff we care about
Mostly plumbing and low-level packages
  · GitHub: facebookincubator/rpm-backports
  · %facebook macro to gate internal stuff
CentOS + FTL = stable distro, moving fast
CentOS at Facebook

Kernel

- Upstream kernel (or as close as possible)
- Development in master
- Internal stable and development branches
- btrfs, cgroup2, PSI, eBPF
- Testing and rollout automation
- Blog: https://tinyurl.com/sftvy7v
CentOS at Facebook

systemd

- systemd backport tracking upstream
- Internal CI/CD pipeline for regression testing
- Feature development
- GitHub: facebookincubator/systemd-compat-libs
- GitHub: facebookincubator/pystemd
- All Systems Go 2019 talk: https://tinyurl.com/v7lxmq3
CentOS at Facebook
Resource control

- cgroup2 in systemd
  - Memory / CPU / IO controls
  - Man page: systemd.resource-control(5)
- Userspace OOM handling with PSI
  - GitHub: facebookincubator/oomd
  - GitHub: facebookincubator/senpai
- resctl, oomd, senpai talks at All Systems Go 2018 / 19
CentOS at Facebook
Network management

- Using network-scripts fleetwide
  - GitHub: fedora-sysv/initscripts
- Hook support
- Control on network restarts
- Troubleshooting
- Feature development and bug fixes
CentOS at Facebook
Packaging

- Standard packaging stack: rpm, dnf/yum
- rpmdb at scale
  - dcrpm to mitigate corruption and remediate
  - GitHub: facebookincubator/dcrpm
- Using an alternate rpmdb: ndb
- Feature development
  - RPM: fsync, CoW experiments with btrfs
  - Yum: Torrent plugin, package cleanup
**CentOS at Facebook**

**Packaging**

- Two stacks: system and fb-runtime
- System packages
  - Yummy: mock wrapper with SCM integration
  - Store specfiles and patches in SCM, blobs in NFS
  - Automation to sync to GitHub
- fb-runtime packages
  - Packman: rpmbuild wrapper for simple RPMs
  - Integrates with buck, generate specfiles from YAML
CentOS at Facebook

Provisioning

- Hosts are netbooted
- Installer: Anaconda => Image-based installer
- Leverage the containers build system
- Chef is used to configure the image
CentOS at Facebook

Containers

- More workloads moving to containers over the years
- FB-specific container manager
- cgroup2, btrfs, systemd
- Blog: https://tinyurl.com/y68evmjjy
Deployment and management
Deployment and management
Chef

- Chef for config management
- Philosophy: https://tinyurl.com/mgxb923
  - Layered configuration through attribute-based APIs
  - Separation of policy and mechanism
  - Idempotency
  - Configuration as programming
- Cookbooks in source control
- Develop locally, test on real machines
Deployment and management

Chef

• Documentation, best practices and tooling
  • GitHub: facebook/chef-utils
  • GitHub: facebookincubator/go2chef
  • GitHub: facebook/taste-tester
  • GitHub: facebook/grocery-delivery
• Cookbooks
  • GitHub: facebook/chef-cookbooks
Deployment and management

Major OS upgrades

- Reprovisioning for OS upgrades
- Clean slate
  - Deprecated unwanted features
  - Policy changes coupling
  - Leverage the general host maintenance window
- Tooling and automation for rollouts
Deployment and management

Minor OS upgrades

- Incremental Rolling OS upgrades
- Every two weeks we sync down the latest updates…
- …and roll them out over two weeks
- ‘yum upgrade’ kicked off via fb_yum in Chef
- High level monitoring of rollout health
- Easy stop button and opt out for individual packages
Deployment and management

Where are we now

- Most of the fleet is on CentOS 7
- Rolling out CentOS Stream
  - Easier rolling updates
  - Contributing to upstream
- CentOS 7 EOL targeted by end of the year
- DevConf.CZ 2020 talk: https://tinyurl.com/tu5gbvk
Questions?
Thank you!

Davide Cavalca <dcavalca@fb.com>