

CentOS Stream at Facebook

DevConf.cz 2021

Davide Cavalca
Production Engineer

FACEBOOK Infrastructure

Agenda

Infrastructure

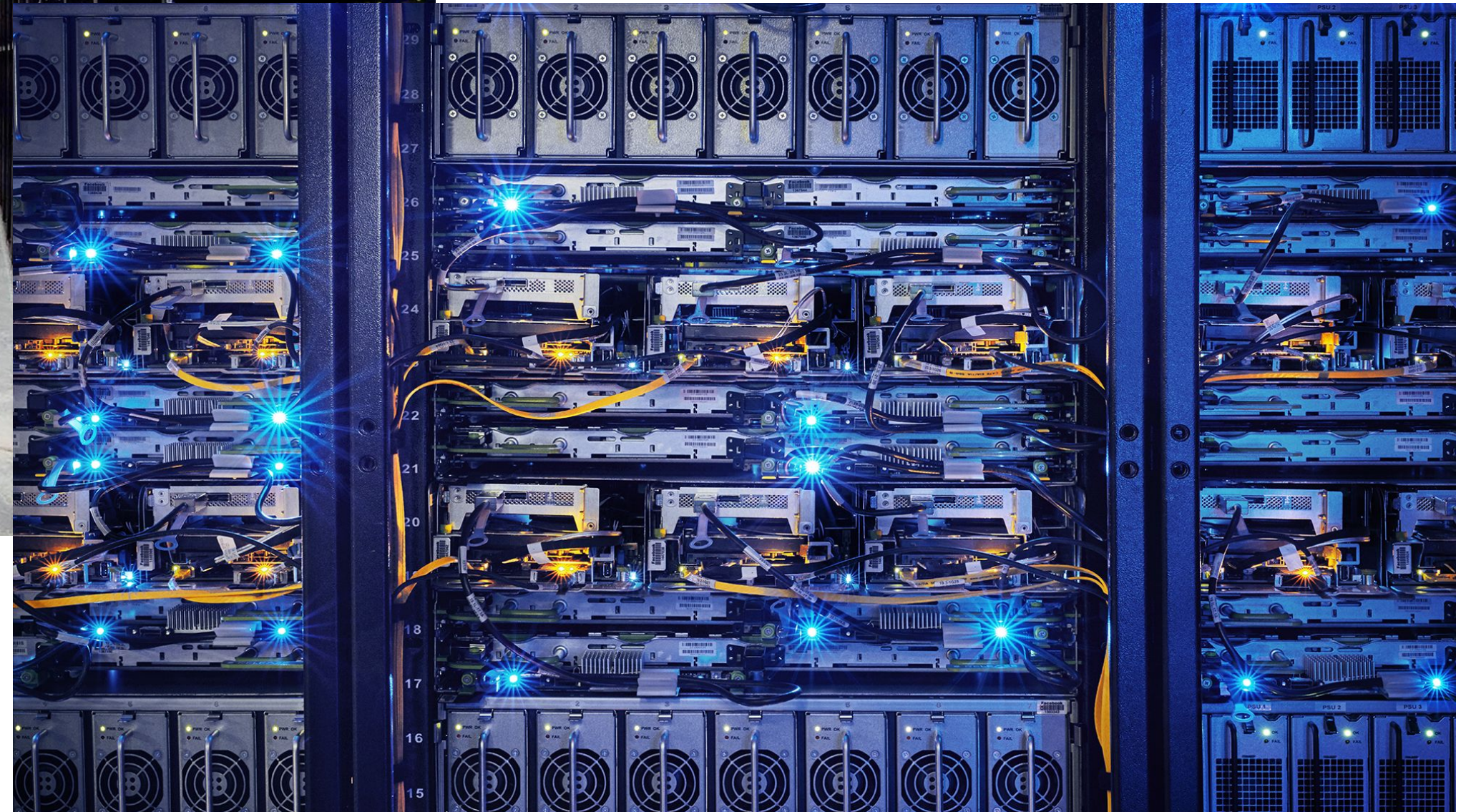
CentOS at Facebook

Contributing upstream

Deployment and management

Infrastructure

Infrastructure



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

CentOS at Facebook

FTL - Fast Thin Layer

- Backports from Fedora Rawhide for stuff we care about
- Mostly plumbing and low-level packages
- GitHub: [facebookincubator/rpm-backports](https://github.com/facebookincubator/rpm-backports)
- %facebook macro to gate internal stuff
- CentOS + FTL = stable distro, moving fast

CentOS at Facebook

Kernel

- Upstream kernel
- 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
- systemd-oomd: userspace OOM handling with PSI
- GitHub: facebookincubator/systemd-compatiblelibs
- GitHub: facebookincubator/pysystemd (also in Fedora and EPEL)
- All Systems Go 2019 talk: <https://tinyurl.com/v7lxmq3>

CentOS at Facebook

Packaging

- Standard packaging stack: rpm, dnf/yum
- rpmdb at scale
 - dcrpm to mitigate corruption and remediate
 - GitHub: facebookincubator/dcrpm (also in Fedora and EPEL)
 - Using alternate rpmdb: ndb
- Work in progress
 - DNF and RPM Copy-on-Write for faster package installs
 - CentOS Dojo FOSDEM 2020 talk: <https://tinyurl.com/y3jr3qv5>
 - SQLite rpmdb evaluation

CentOS at Facebook

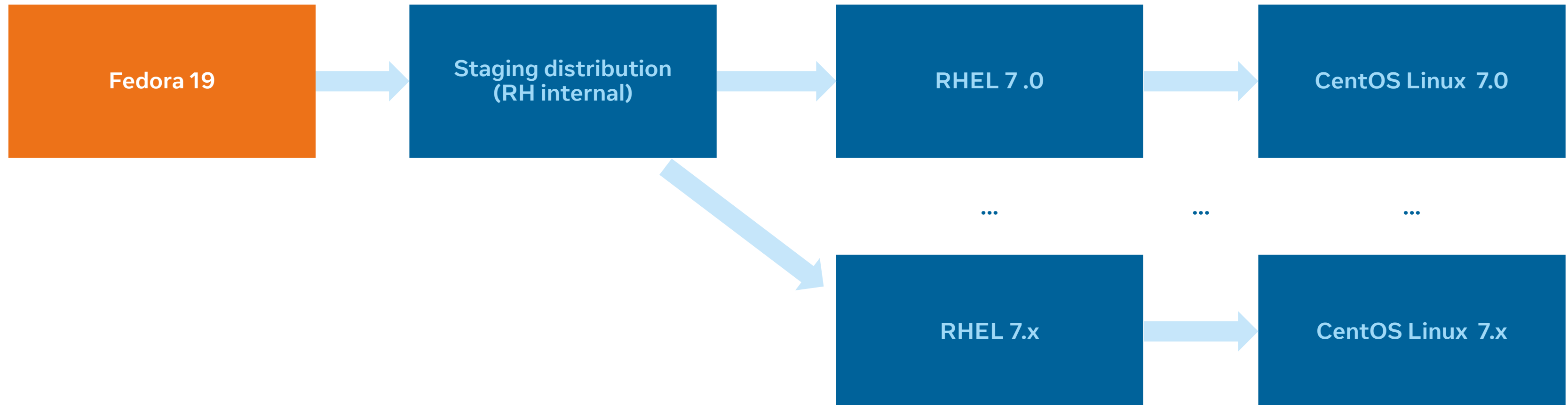
Policy deviations

- cgroup2 by default
- btrfs on / by default
- iptables: legacy backend instead of nftables
- networking: network-scripts instead of NetworkManager
- CentOS Dojo Brussels 2020 talk: <https://tinyurl.com/yqwhr4j8>

Contributing
upstream

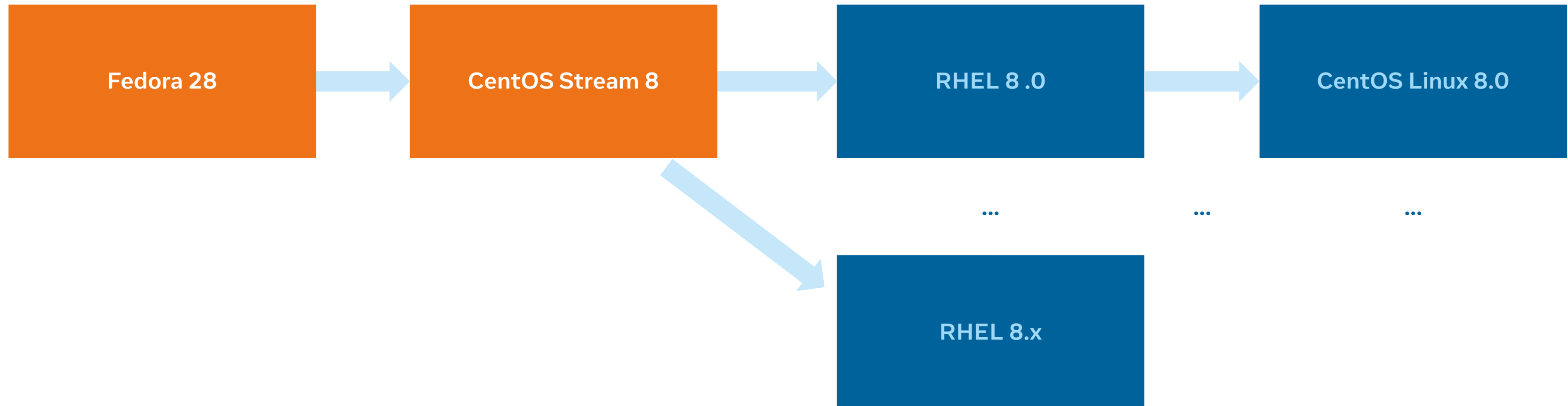
Contributing upstream

CentOS Linux 7



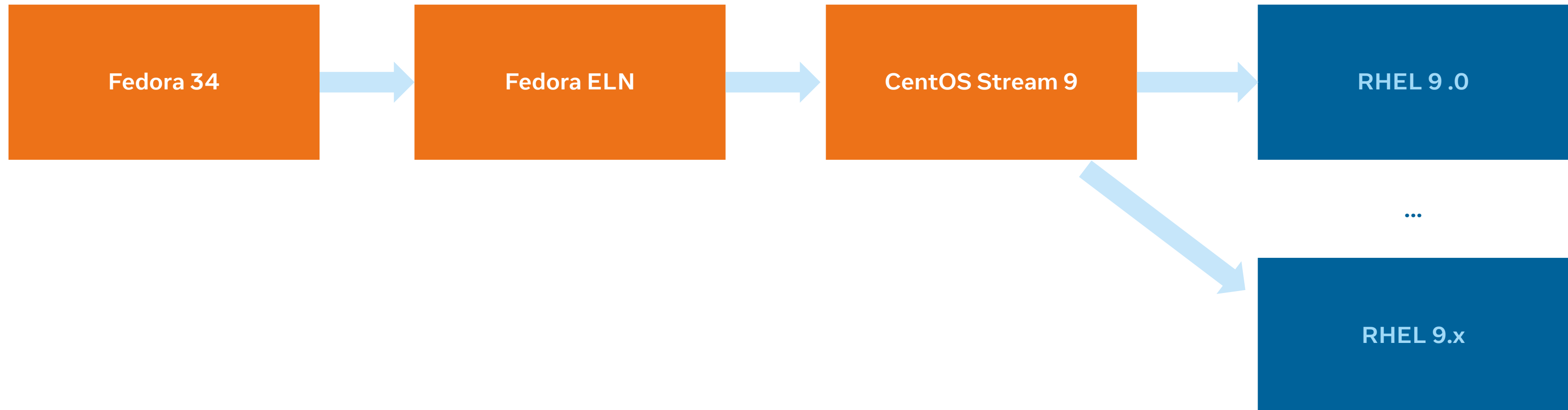
Contributing upstream

CentOS Linux 8 and CentOS Stream 8



Contributing upstream

CentOS Stream 9



CentOS Stream blog post: <https://tinyurl.com/ycn29k2c>

Contributing upstream

What can you do

- Fedora
 - Influences the next CentOS Stream major release
 - File and fix bugs, maintain packages, drive Changes, etc.
- Fedora ELN
 - Assists in the bringup of the next CentOS Stream major release
 - Join the meetup tomorrow (<https://tinyurl.com/45h8qp8y>)
- CentOS Stream
 - Continuously delivered distribution tracking the next minor release of RHEL
 - File and fix bugs, send pull requests, etc.
 - Join or create a SIG

Contributing upstream

Fedora and EPEL

- Fedora packaging
 - Facebook stack and OSS projects
 - Feature enablement
 - Rust packaging
- Change proposals
 - F33: Btrfs by default
 - F34: Btrfs with zstd compression by default
 - F34: systemd-oomd by default
 - ~~F34~~ F35: DNF RPM Copy-on-Write
- EPEL Packagers SIG
- FOSDEM 2021 talk: <https://tinyurl.com/1osbj4b>

Contributing upstream

Hyperscale SIG

- CentOS Stream focus
- Large scale infrastructure
- Foster cross-company collaboration on packaging and tooling
- Bring in-house development out in the open
- Open to anybody interested in working in this space
- <https://wiki.centos.org/SpecialInterestGroup/Hyperscale>
- CentOS Dojo FOSDEM 2021 talk: <https://tinyurl.com/nw9wehi9>

Contributing upstream

Hyperscale SIG

- Faster-moving package backports
 - systemd, grep, dwarves, libvirt, rasdaemon, ...
- Policy and configuration alternatives
 - iptables
- Large-scale testing
 - RPM Copy-on-Write
- Kernel
 - LTS-based, btrfs and cgroup2 support

Contributing upstream

Hyperscale SIG

- Right now
 - c8s-sig-hyperscale branches on git.centos.org
 - Main package repository
 - `dnf install centos-release-hyperscale`
- In the future
 - Experimental package repository
 - Kernel
 - Cloud images

Deployment and management

Deployment and management

Chef

- Chef for config management
- Philosophy: <https://tinyurl.com/mqxb923>
 - 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 (also in Fedora)
 - GitHub: facebook/taste-tester
 - GitHub: facebook/grocery-delivery
- Cookbooks
 - GitHub: facebook/chef-cookbooks

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
- ‘dnf 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

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

Major OS upgrades

- CentOS Linux 5 -> 6 (~2013-2016)
- CentOS Linux 6 -> 7 (2016-2018)
- CentOS Linux 7 -> CentOS Stream 8 (2018-2021)
- DevConf.cz 2020 talk: <https://tinyurl.com/52hqdp6t>
- Current status
 - >85% of the fleet on CentOS Stream 8
 - Long tail: switches, storage, containers
 - Coming up: CentOS Stream 9!

Thank you!

Questions?

FACEBOOK Infrastructure