Fedora CoreOS, a container focused OS to securely deploy and run applications

Pass the SALT 2021

Timothée Ravier
CoreOS engineer at Red Hat

[m] @siosm:matrix.org
@siosm
travier@redhat.com
tim.siosm.fr

Agenda

- What is Fedora CoreOS?
- What makes it great to securely run containers?
- Demos: Single node, Nomad and Kubernetes
- Questions!

travier@redhat.com

What is Fedora CoreOS?

An emerging Fedora edition

- Came from the **merging** of two communities:
  - CoreOS Inc's Container Linux
  - Project Atomic's Atomic Host

- Incorporates **Container Linux**
  - Philosophy
  - Provisioning Stack
  - Cloud Native Expertise

- Incorporates **Atomic Host**
  - Fedora Foundation
  - Update Stack
  - SELinux Enhanced Security

travier@redhat.com

Philosophy behind Fedora CoreOS

- **Automatic updates**
  - No interaction for administrators
- **Automated provisioning**
  - All nodes start from **~same starting point**
  - Use Ignition to provision a node on **first boot**
- **Immutable infrastructure**
  - **Automate** deployment and system configuration
  - Update configs and **re-provision** to apply changes
- User software runs in **containers**
  - Makes host updates more **reliable**

travier@redhat.com

Supported platforms and architectures

- Available for a plethora of **cloud/virt platforms**
  - Alibaba, AWS, Azure, DigitalOcean, Exoscale, GCP, IBM Cloud, Openstack, Vultr, VMWare, QEMU/KVM
  - Directly launchable on AWS & GCP

- Several options for **Bare Metal**
  - Classic ISO
  - PXE (network) boot
  - 4K native disks images

- Currently **x86_64** only (**aarch64** support coming soon)

travier@redhat.com

Making it easy to run containers securely

Software has bugs (!)

- **Memory safety** issues, logic bugs
- **Linux kernel** vulnerabilities
- CVEs & **non** CVEs fixes
- etc.

travier@redhat.com

Reducing the OS footprint

- First step in security hardening: **reducing** attack surface
  - Less software to **track** for security and bug fixes
- Built from **Fedora Linux packages**
  - Plus some small configuration additions
- Only **essential** system services and administration tools
- **Container runtimes** (podman & moby-engine)
- Bash is the **only interpreter**. No Python, etc.

travier@redhat.com

Building with safer languages

- Using **memory safe languages** for most of Fedora CoreOS specific additions:
  - **Go**: Butane, Ignition, toolbox, container engines (podman & moby-engine)
  - **Rust**: Afterburn, Zincati, coreos-installer, bootupd, rpm-ostree (in progress)

travier@redhat.com

OS versioning and filesystem layout

- Based on **rpm-ostree** to manage and update the system
- rpm-ostree: Hybrid image/package system
  - “Like **git** for your Operating System”
  - A single identifier for a given system version
    - Example: 32.20200615.2.0 - 86c0246
- Uses **read-only** filesystem mounts
  - Prevents accidental OS corruption (rm -rf)
  - Prevents novice attacks from modifying system
- **Clear distinction** between /usr, /etc and /var

travier@redhat.com

Automatic updates by default

- **Automatic** updates → **Reliable** updates
  - User software runs in containers

- **Extensive tests** in automated CI pipelines
  - Several update streams to **preview** what's coming
    - Users run various streams to help find issues
  - **Managed upgrade rollouts** over several days
    - Halt the rollout if issues are found

- For when things go wrong
  - rpm-ostree **rollback** can be used to go back

travier@redhat.com

Multiple update streams

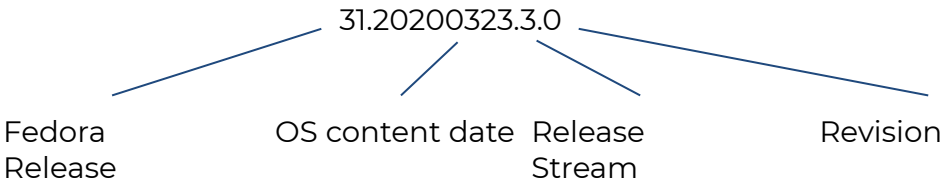
- Offered update streams with automatic updates
  - **next** - experimental features, Fedora major rebases
  - **testing** - preview of what's coming to stable
    - point in time snapshot of Fedora stable rpm content
  - **stable** - most reliable stream offered
    - promotion of testing stream after some bake time

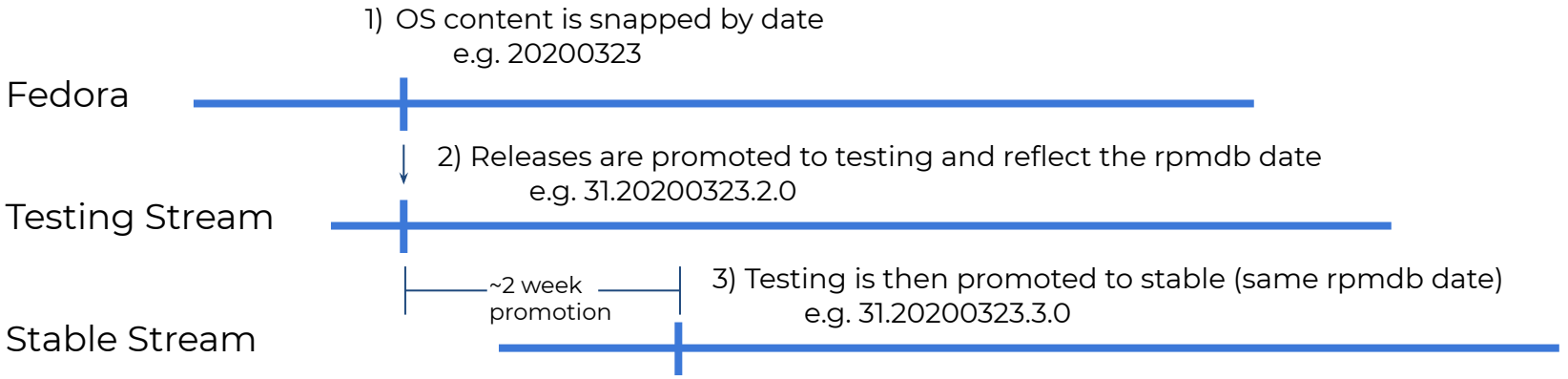
- Goals
  - New releases approximately **every two weeks**
  - Find issues in next/testing **before they hit stable**

travier@redhat.com

Fedora CoreOS release promotion

Release Nomenclature





travier@redhat.com

Everything else runs in containers

- Two container runtimes available:
  - **podman** and **moby-engine (docker)**
- Confinement with **SELinux**:
  - **Confined** system services (targeted policy)
  - **Isolation** between containers and container ↔ host
  - Already **blocked** several real vulnerabilities in runc:
    - CVE-2019-5736: Latest container exploit (runc) can be blocked by SELinux
    - CVE-2021-30465: Mitigated by Default in OpenShift

travier@redhat.com

Automated provisioning on first boot

- Fedora CoreOS uses Ignition to **automate** provisioning
- Any logic for machine lifetime is **encoded** in the config
  - Very easy to automatically **re-provision** nodes
- **Same starting point** whether on bare metal or cloud
  - Use Ignition **everywhere** as opposed to kickstart for bare metal and cloud-init for cloud

travier@redhat.com

Ignition configs

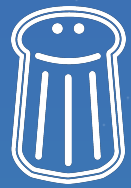
- **Declarative** JSON documents provided via user data
- Runs **exactly once**, during the initramfs stage on **first boot**
- Can write files and systemd units, create users and groups, partition disks, create RAID arrays, format filesystems
- **If provisioning fails, the boot fails** (no half provisioned systems)
- Ignition configs are **machine-friendly** (JSON)

```
{
 "ignition": {
  "config": {},
  "timeouts": {},
  "version": "3.0.0"
 },
 "passwd": {
  "users": [
   {
    "name": "core",
    "passwordHash": "$6$43y3tkl...",
    "sshAuthorizedKeys": [
     "ssh-ed25519 ..."
    ]
   }
  ]
 },
 "storage": {
  ...
 },
 "systemd": {
  ...
 }
}
```

Butane configs

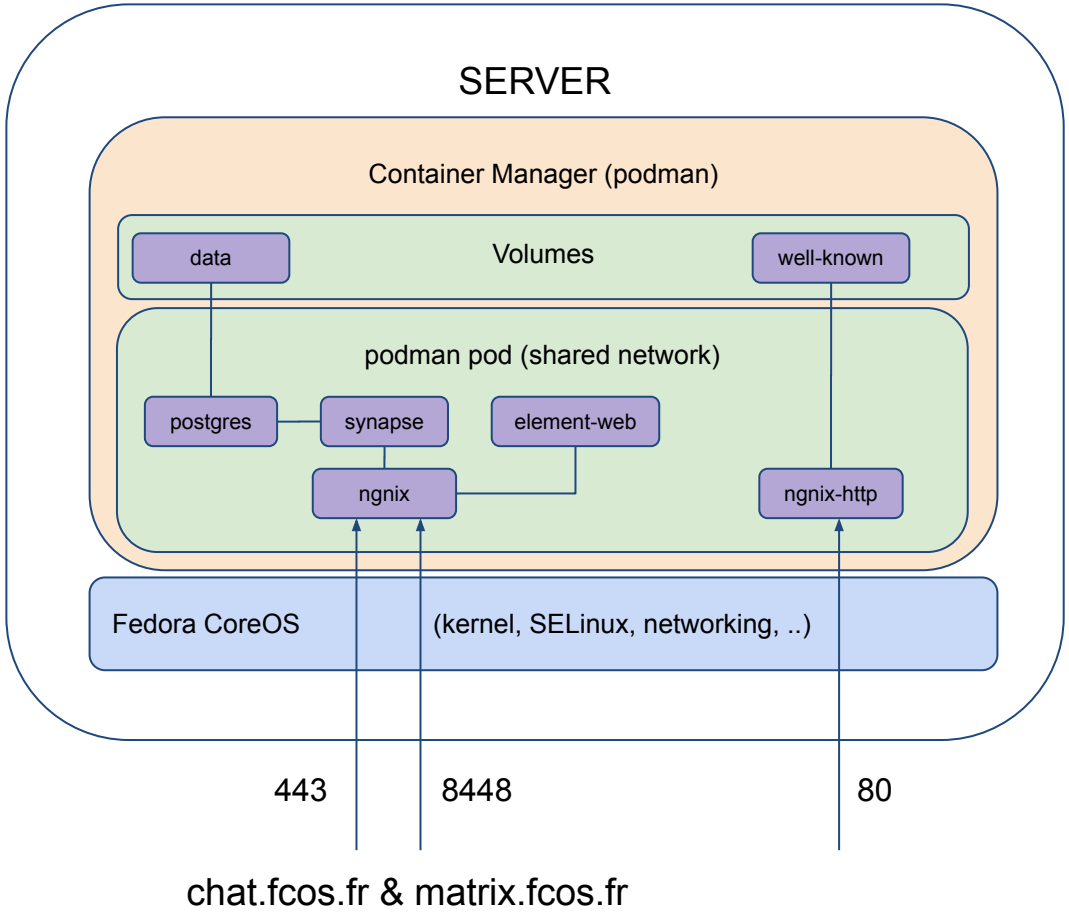
- **Butane** is a configuration transpiler
- **Converts** Butane configs to Ignition configs
- Butane configs are **Human friendly** (YAML)
- Ignition semantics, plus **sugar** for common operations
- Transpiler catches common errors at **build time**

```
variant: fcos
version: 1.3.0
passwd:
 users:
   - name: core
     ssh_authorized_keys:
       - ssh-ed25519 ...
systemd:
 units:
   - name: docker.service
     enabled: false
     mask: true
   - name: docker.socket
     enabled: false
     mask: true
   - name:
storage:
 files:
   - path: /etc/chrony.conf
     overwrite: yes
     mode: 0644
     contents:
       local: chrony.conf
```



Demos!

Matrix demo



chat.fcos.fr & matrix.fcos.fr

travier@redhat.com

https://github.com/travier/fedora-coreos-matrix

Nomad demo

https://github.com/travier/fedora-coreos-nomad

travier@redhat.com

OKD demo

https://github.com/openshift/okd#getting-started

travier@redhat.com

Get involved!

- Web: https://getfedora.org/coreos
- Issues: https://github.com/coreos/fedora-coreos-tracker/issues
- Forum: https://discussion.fedoraproject.org/c/server/coreos
- Mailing list: coreos@lists.fedoraproject.org
- IRC: Libera.chat #fedora-coreos
- Other talks to get started:
  - Fedora CoreOS Introduction (Jul 13, 2020)
  - Getting Started with Fedora CoreOS (Mar 17, 2021)