



Immutable All the Way Down

Using System Extensions
to ship Kubernetes



Containers Devroom, FOSDEM'25
February 1, 2025



Hello, I'm

Thilo



Thilo Fromm

Flatcar Maintainer

Github: [t-lo](#)

Mastodon: [@thilo@fromm.social](#)

Email: thilofromm@microsoft.com

Cluster Plumbing

Cluster Plumbing

Shipping Kubernetes

Cluster Plumbing

Shipping Kubernetes as immutable image

Cluster Plumbing

Shipping Kubernetes as immutable image

Separating Kubernetes from the OS

Cluster Plumbing

Shipping Kubernetes as immutable image

Separating Kubernetes from the OS

Live in-place updates

Before we start

Thilo, start provisioning our ClusterAPI cluster!



Extension Images

UAPI group Extension Image spec

https://uapi-group.org/specifications/specs/extension_image/

Application + lean metadata wrapped in **immutable** FS image

Merge into OS FS root, usually at boot

Tie-in with sysupdate, a lightweight update mechanism

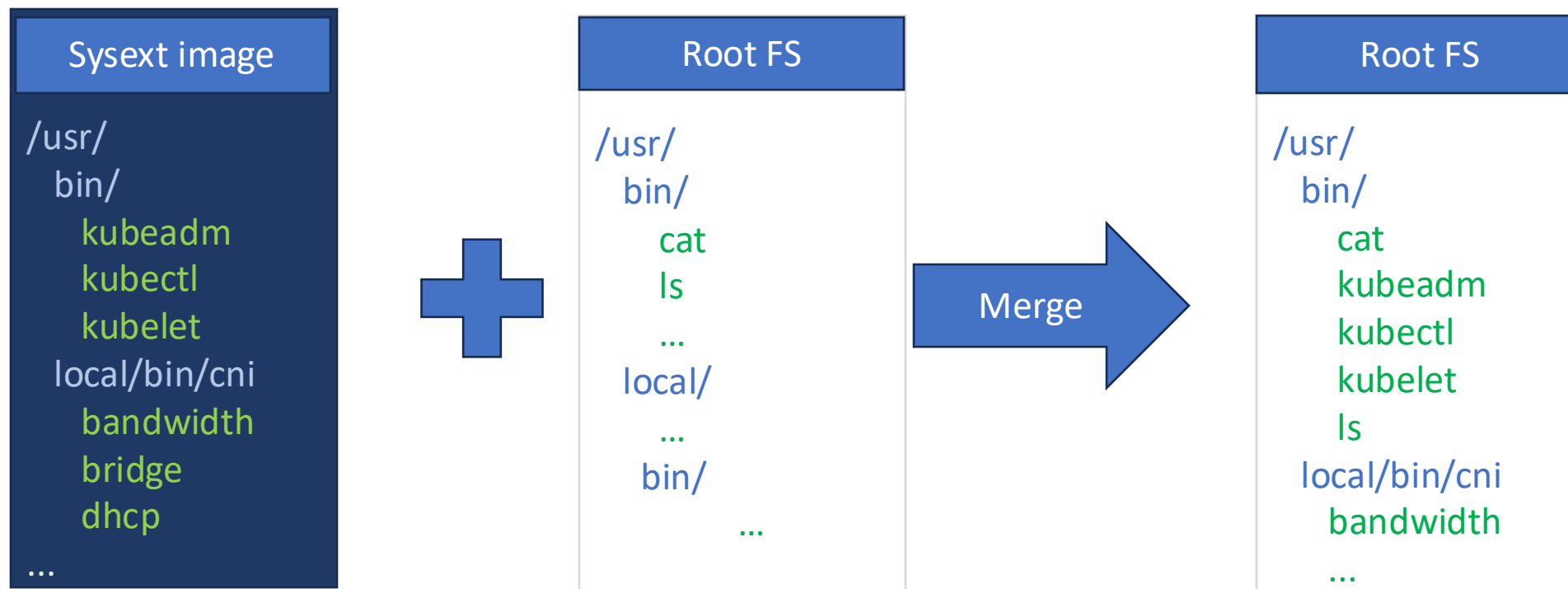
Application binaries and libraries

Static configuration files

Wrapper and helper scripts for automation

Systemd service units, timers, etc.

Merging Extension Images



Download during node provisioning

Composed into base OS at boot

```
storage:  
  files:  
    - path: /etc/extensions/kubernetes.raw  
      contents:  
        source: https://github.com/flatcar/sysexst-bakery/releases/download/latest/kubernetes-v1.31.3-x86-64.raw
```

Show me!

Single node sysext live demo

Updating Extension Images



Systemd-sysupdate: complementary service for updating images

Can consume from HTTP[S]

Sysupdate supports updating symlinks

[Source]

Type=url-file

Path=<https://github.com/flatcar/sysexst-bakery/releases/latest/download/>

MatchPattern=kubernetes-v1.31.@v-%a.raw

[Target]

InstancesMax=3

Type=regular-file

Path=/opt/extensions/kubernetes/

CurrentSymlink=/etc/extensions/kubernetes.raw

Producing and hosting Extension Images



Run mksquashfs on directory w/ sysext root

Host on HTTP[S] w/ index file

= => GitOps!

```
6565c5f8912076cb7c966b7019f512ca639263266924bc983bd21e82885e80f26
11d8f9898c2bb7ccb4ececbb1dca142874cc7b3d82c91268c0c2be2e80af4318
393ce3b2bd45043b9f9af20d48ea10bdf2ee179da14df2729e6b8ae3fb199c1c
db33e645887b75143341458bedc3a67071861d492cdb0757ec43951c3d44c8a9
98a98c0241ba4e141181ecbd102c76bf1aa2963e316c55aa154e904a9352028b
cdb15c905147b8985ed0ce50dfc1b4c71fc93fe02c235fc56e8ca1fe077c63d1
3aa9d51232052cd44b5cb125aa82fe6b8aa62fdc1d6ea4ececbb0374d94be4a98
03cdeba64af248e021e82c3a3c4c1fee55b8033fccacfe971973222fe0e5e277
f2a09e3162c42c4c35a6bf1397cca168c46318ad1004dac168ba59d53634c409
899146b4dbc386fbcc90e555f4f325169b8339931ab59c2f1652aff4d9df237
f245ee2dfb59691943bb5898a3bbccd02a1b2637b77b4ddccc896958e60d8e86
b7941ef076736029718594b7b0124d039c88bf18229c21638281fede4af172e
5a185397a0c0458ef2b2d1ac9b7c8c9cca809afd542eb27addb0150576631683
866f2741a619fb3371796bc04924ecb7b078fac4ae0a69eb5b0af15cef151982
3ab9d5ad92476270806fbc53de2c0bc7343bab53cfff894dcab2075f608cf6933
1d380d23d5ffa99d4b93cc9281b5e8162445308c3582c062753cf13ffba17203
4516c74e30a02b9486a7ab0e83c923b07186472ca9e34a7bb9fc3badfdb8cfb9
5f9066d7e71e2bd8d4439a5adf320a841b03903370cb066957047c07190be75c
67a5755f2136d9c160bc5ae36683e2558328a731ee6e279d0c57101d69567b4f
f2d6ad55deb81695336817003268bd41d7f9edc642d8d7b894c6b451f207249c
b90b4da1ca309e66921698ae83870176524e9631b9b7e5a5f485d2abf4c61446
46c5705d7468ca3b9d6d2c103c4598f034e1957af01c69b0cb39905a590b66e1
95d645ced1b639d509fba3eb78d46913f03fe3afbd46d61e88371e1da737a2ed
```

```
kubernetes-v1.28.14-
kubernetes-v1.28.14-
kubernetes-v1.28.5-
kubernetes-v1.28.5-
kubernetes-v1.28.6-
kubernetes-v1.28.6-
kubernetes-v1.28.7-
kubernetes-v1.28.7-
kubernetes-v1.28.8-
kubernetes-v1.28.8-
kubernetes-v1.28.9-
kubernetes-v1.28.9-
kubernetes-v1.29.1-
kubernetes-v1.29.1-
kubernetes-v1.29.10-
kubernetes-v1.29.10-
kubernetes-v1.29.11-
kubernetes-v1.29.11-
kubernetes-v1.29.12-
kubernetes-v1.29.12-
kubernetes-v1.29.2-
kubernetes-v1.29.2-x86_64.raw
kubernetes-v1.29.3-arm64.raw
```

flatcar / sysext-bakery

[Code](#) [Issues](#) [Pull requests](#) [Actions](#) [Projects](#) [Wiki](#) [Security](#) [Insights](#) [Settings](#)

Releases / latest

latest

github-actions released this Sep 10, 2024

latest

6cb91d8

Release 2024-12-18 08:52

The release adds the following sysexts:

- wasmedge-0.14.1-x86-64.raw
- wasmedge-0.14.1-arm64.raw
- llamaedge-0.14.16-x86-64.raw
- llamaedge-0.14.16-arm64.raw
- k3s-v1.31.3+k3s1-x86-64.raw
- k3s-v1.31.3+k3s1-arm64.raw
- k3s-v1.30.7+k3s1-x86-64.raw
- k3s-v1.30.7+k3s1-arm64.raw
- k3s-v1.29.11+k3s1-x86-64.raw
- k3s-v1.29.11+k3s1-arm64.raw
- crio-v1.31.3-x86-64.raw
- crio-v1.31.3-arm64.raw
- crio-v1.30.8-x86-64.raw
- crio-v1.30.8-arm64.raw
- crio-v1.29.11-x86-64.raw
- crio-v1.29.11-arm64.raw
- rke2-v1.31.3+rke2r1-x86-64.raw
- rke2-v1.31.3+rke2r1-arm64.raw
- rke2-v1.30.7+rke2r1-x86-64.raw
- rke2-v1.30.7+rke2r1-arm64.raw
- rke2-v1.29.11+rke2r1-x86-64.raw
- rke2-v1.29.11+rke2r1-arm64.raw
- kubernetes-v1.32.0-x86-64.raw
- kubernetes-v1.31.4-x86-64.raw
- kubernetes-v1.31.4-arm64.raw
- kubernetes-v1.30.8-x86-64.raw
- kubernetes-v1.30.8-arm64.raw
- kubernetes-v1.29.12-x86-64.raw
- kubernetes-v1.29.12-arm64.raw

sysext-bakery

[Edit Pins](#) [Unwatch](#) 10

main 7 Branches 6 Tags

Go to file

Add file

Code

tormath1 release-build: remove "Skipping \$version"

6cb91d8 · last month

200 Commits

.github/workflows	release.yaml: Split artifact upload and don't delete release	4 months ago
.gitignore	gitignore: Ignore any local tar balls	2 years ago
LICENSE	LICENSE: re-license to Apache 2.0	2 years ago
MAINTAINERS.md	Sync maintainers file from flatcar/flatcar repository	3 years ago
README.md	sysext: add llamaedge recipe	2 months ago
bake.sh	bake: use semver for version check	3 months ago
bake_flatcar_image.sh	bake_flatcar_image.sh: Update list of vendors, improve clean...	2 months ago
code-of-conduct.md	code-of-conduct.md: add CNCF code of conduct.	2 years ago
convert_torcx_image.sh	Default to ID=_any and make use of EXTENSION_RELOAD_M...	11 months ago
create_containerd_sysext.sh	sysext: add containerd recipe	2 months ago
create_crio_sysext.sh	crio: make usage of '\$VERSION' consistent	2 months ago
create_docker_compose_sysext.sh	Use '/usr/bin/env bash' instead of '/bin/bash'	2 years ago
create_docker_sysext.sh	Default to ID=_any and make use of EXTENSION_RELOAD_M...	11 months ago
create_falco_sysext.sh	Bake Sysdig/CNCF Falco (#77)	2 months ago
create_k3s_sysext.sh	create_k3s_sysext.sh: create symlinks for kubectl, ctr and cric...	8 months ago
create_keeplived_sysext.sh	Add keeplived (#65)	9 months ago
create_kubernetes_sysext.sh	Default to ID=_any and make use of EXTENSION_RELOAD_M...	11 months ago
create_llamaedge_sysext.sh	sysext: add llamaedge recipe	2 months ago

Show me!

Single node sysext sysupdate

Self-contained, immutable OS level application images

Easy to generate, easy to compose, easy to manage

Complementary update mechanism

Your Linux distro likely supports it (if it ships systemd). Try it out!



Sysexts and Kubernetes Deployments

Leverage upstream OS maintenance

Take control of Kubernetes



Leverage upstream OS maintenance

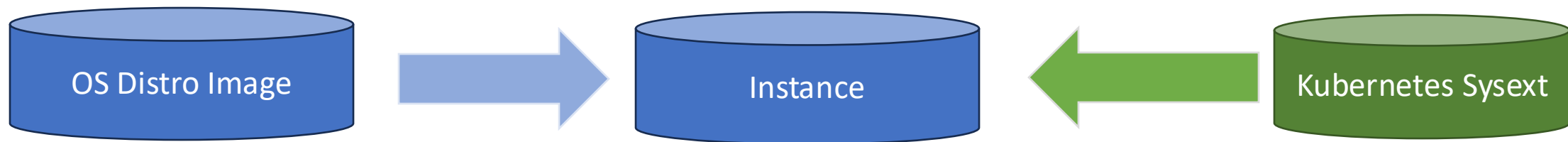
Take control of Kubernetes



Separation of Concerns

Leverage upstream OS maintenance

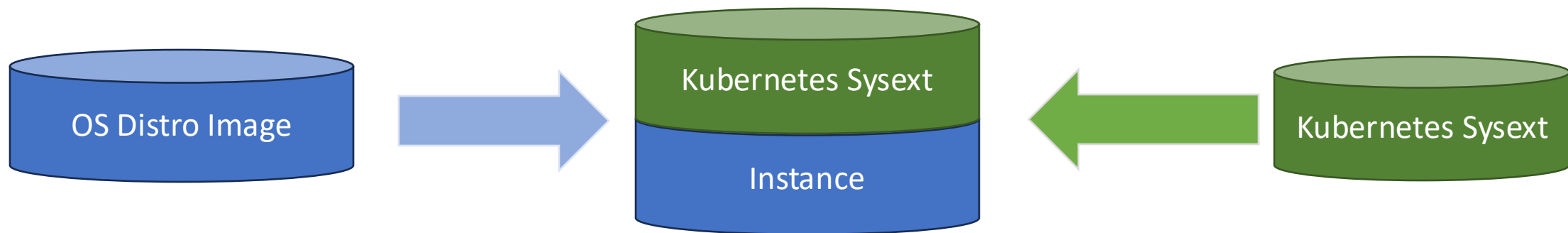
Take control of Kubernetes



Separation of Concerns

Leverage upstream OS maintenance

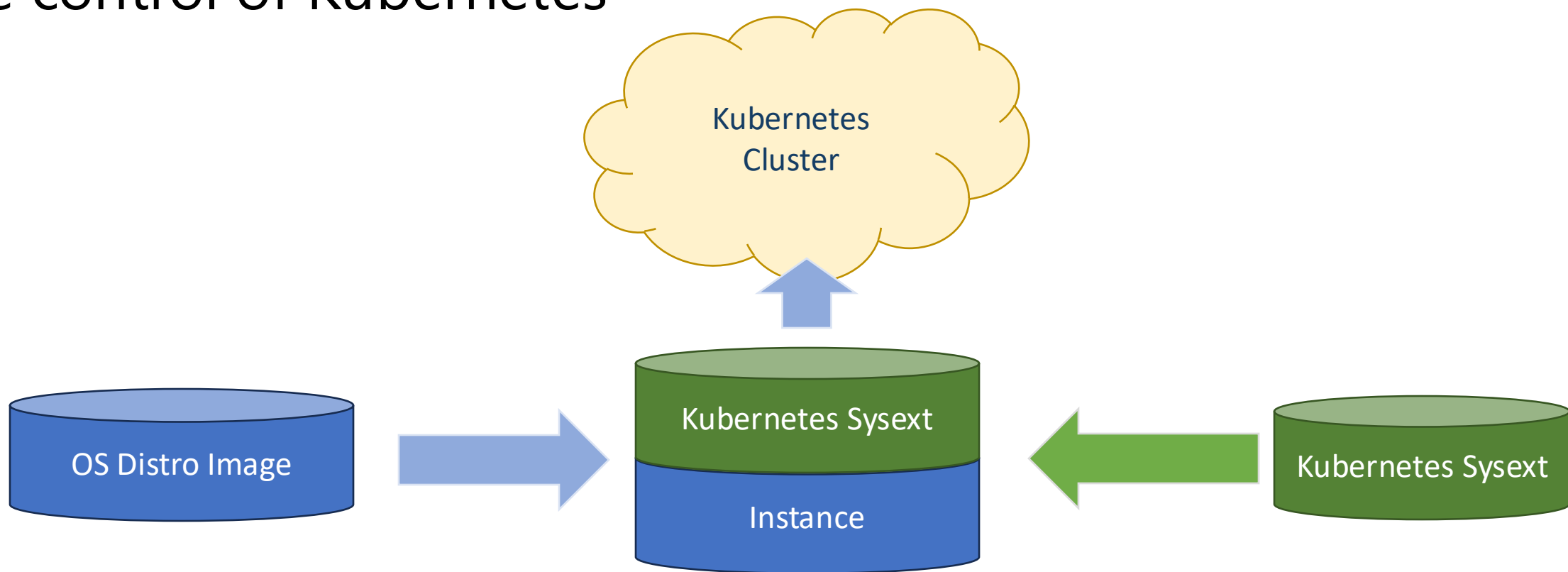
Take control of Kubernetes



Separation of Concerns

Leverage upstream OS maintenance

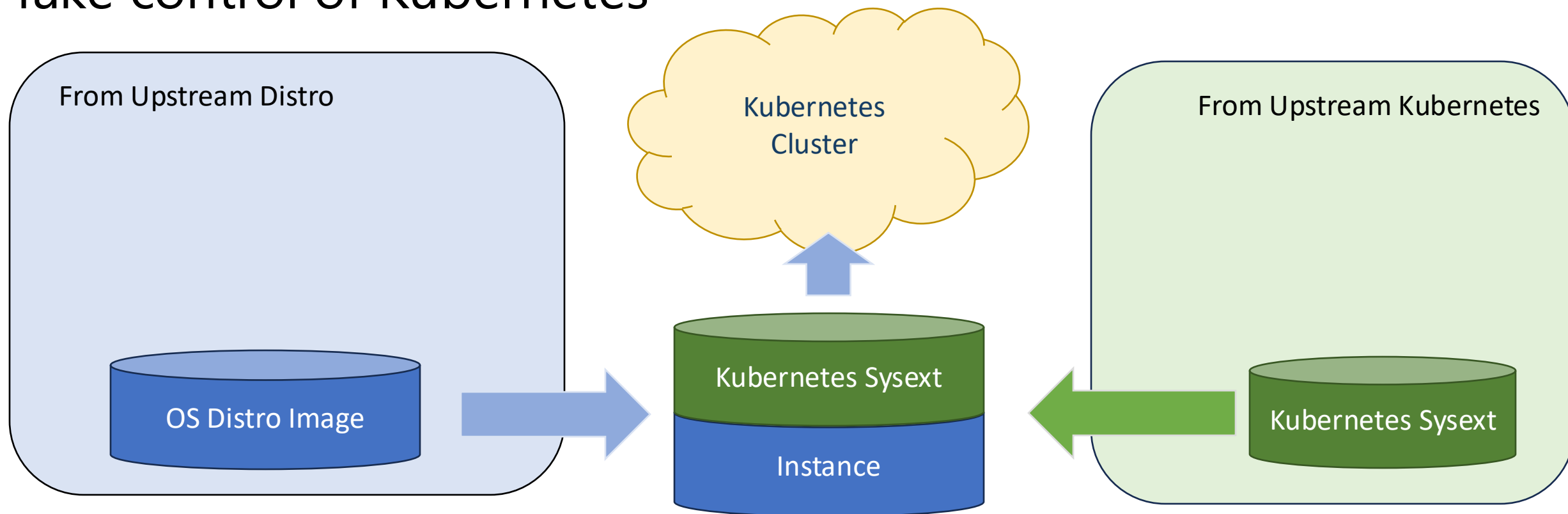
Take control of Kubernetes



Separation of Concerns

Leverage upstream OS maintenance

Take control of Kubernetes



Loose coupling

OS and Kubernetes maintained independently

Separate In-Place Updates

Cut down version / support matrix

Version Matrix? Big issue with pre-built images.



Ubuntu-22 + Kubernetes 1.29 + AWS

Version Matrix? Big issue with pre-built images.



Ubuntu-22 + Kubernetes 1.29 + AWS

Ubuntu-22 + Kubernetes 1.29 + Azure

Version Matrix? Big issue with pre-built images.



Ubuntu-22 + Kubernetes 1.29 + AWS

Ubuntu-22 + Kubernetes 1.29 + Azure

Ubuntu-22 + Kubernetes 1.29 + GCP

Version Matrix? Big issue with pre-built images.



Ubuntu-22 + Kubernetes 1.29 + AWS

Ubuntu-22 + Kubernetes 1.29 + Azure

Ubuntu-22 + Kubernetes 1.29 + GCP

Ubuntu-22 + Kubernetes 1.30 + AWS

Ubuntu-22 + Kubernetes 1.30 + Azure

Ubuntu-22 + Kubernetes 1.30 + GCP

Version Matrix? Big issue with pre-built images.



Ubuntu-22 + Kubernetes 1.29 + AWS

Ubuntu-22 + Kubernetes 1.29 + Azure

Ubuntu-22 + Kubernetes 1.29 + GCP

Ubuntu-22 + Kubernetes 1.30 + AWS

Ubuntu-22 + Kubernetes 1.30 + Azure

Ubuntu-22 + Kubernetes 1.30 + GCP

Ubuntu-22 + Kubernetes 1.31 + AWS

Ubuntu-22 + Kubernetes 1.31 + Azure

Ubuntu-22 + Kubernetes 1.31 + GCP

Version Matrix? Big issue with self-hosted images.



Ubuntu-24 + Kubernetes 1.29 + AWS

Ubuntu-24 + Kubernetes 1.29 + Azure

Ubuntu-24 + Kubernetes 1.29 + GCP

Ubuntu-24 + Kubernetes 1.30 + AWS

Ubuntu-24 + Kubernetes 1.30 + Azure

Ubuntu-24 + Kubernetes 1.30 + GCP

Ubuntu-24 + Kubernetes 1.31 + AWS

Ubuntu-24 + Kubernetes 1.31 + Azure

Ubuntu-24 + Kubernetes 1.31 + GCP

Version Matrix? Big issue with self-hosted images.



RHEL9 + Kubernetes 1.29 + AWS

RHEL9 + Kubernetes 1.29 + Azure

RHEL9 + Kubernetes 1.29 + GCP

RHEL9 + Kubernetes 1.30 + AWS

RHEL9 + Kubernetes 1.30 + Azure

RHEL9 + Kubernetes 1.30 + GCP

RHEL9 + Kubernetes 1.31 + AWS

RHEL9 + Kubernetes 1.31 + Azure

RHEL9 + Kubernetes 1.31 + GCP

Version Matrix? Big issue with self-hosted images.



SLES + Kubernetes 1.29 + AWS

SLES + Kubernetes 1.29 + Azure

SLES + Kubernetes 1.29 + GCP

SLES + Kubernetes 1.30 + AWS

SLES + Kubernetes 1.30 + Azure

SLES + Kubernetes 1.30 + GCP

SLES + Kubernetes 1.31 + AWS

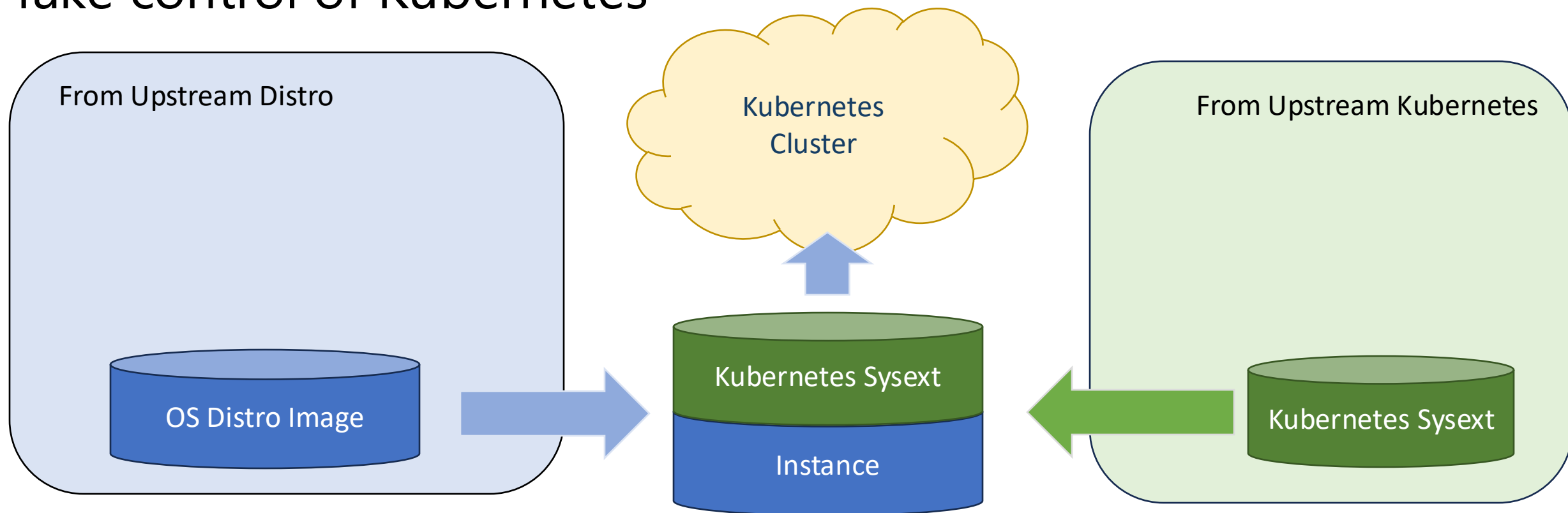
SLES + Kubernetes 1.31 + Azure

SLES + Kubernetes 1.31 + GCP

Separation of Concerns

Leverage upstream OS maintenance

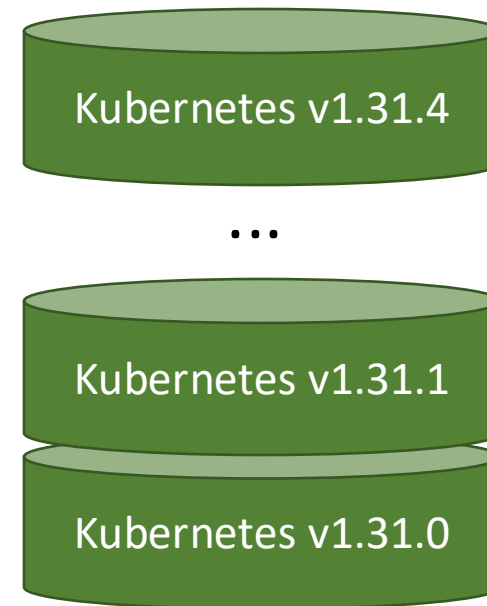
Take control of Kubernetes



In-Place Updates

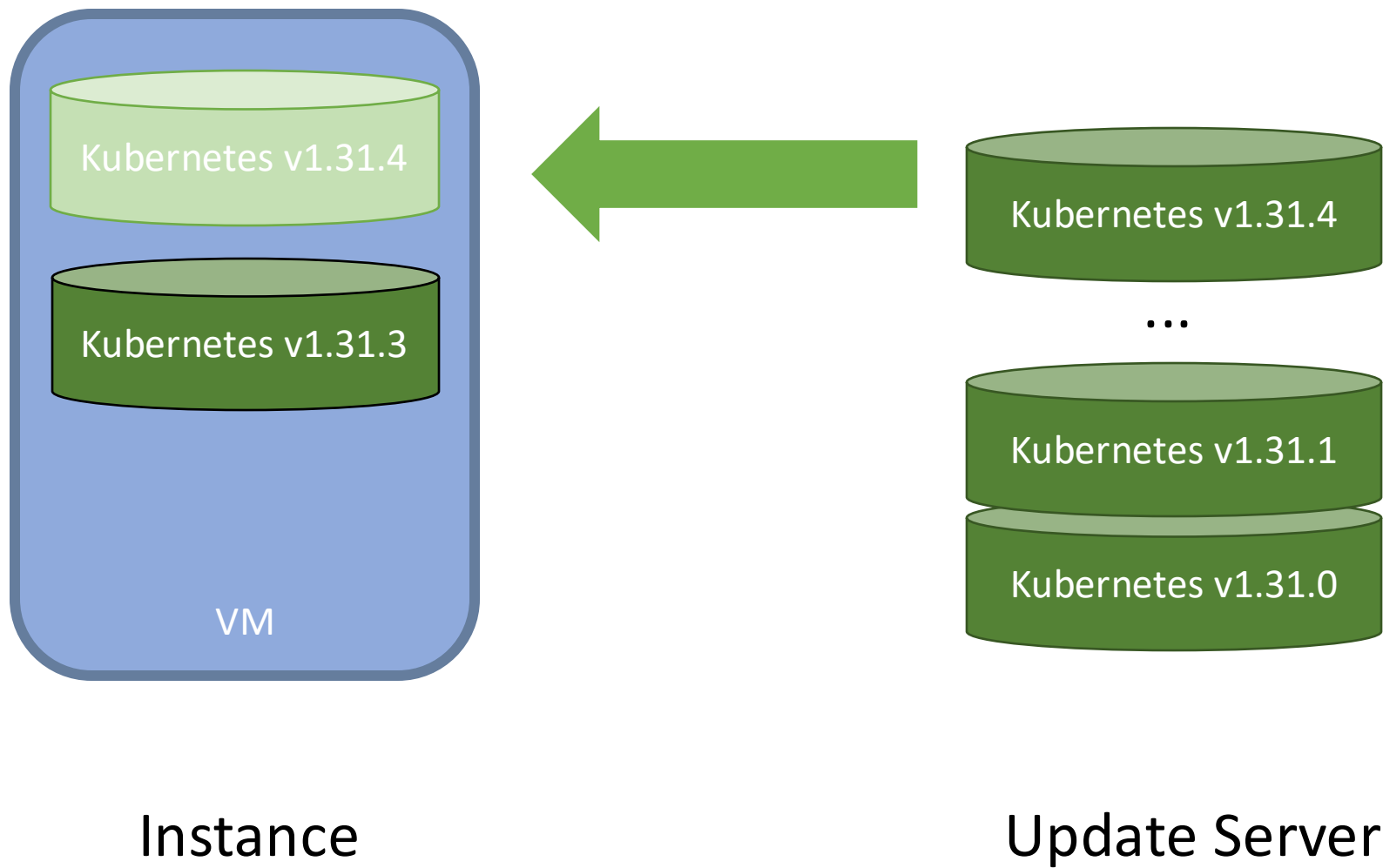


Instance

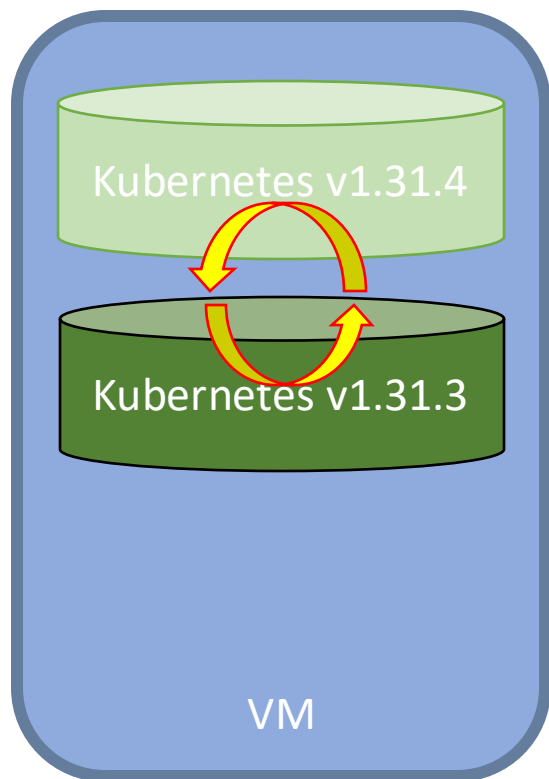


Update Server

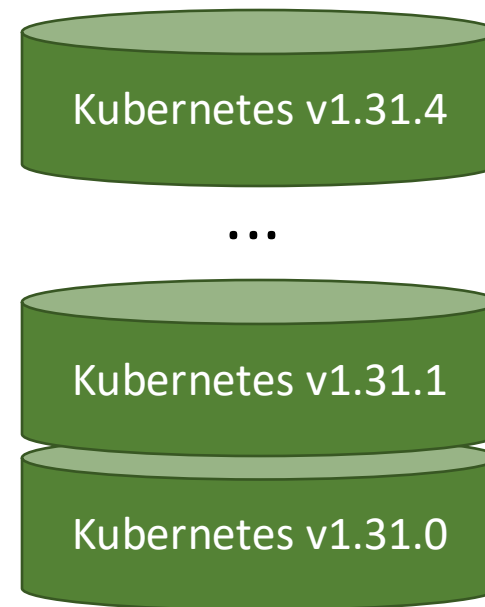
In-Place Updates



In-Place Updates

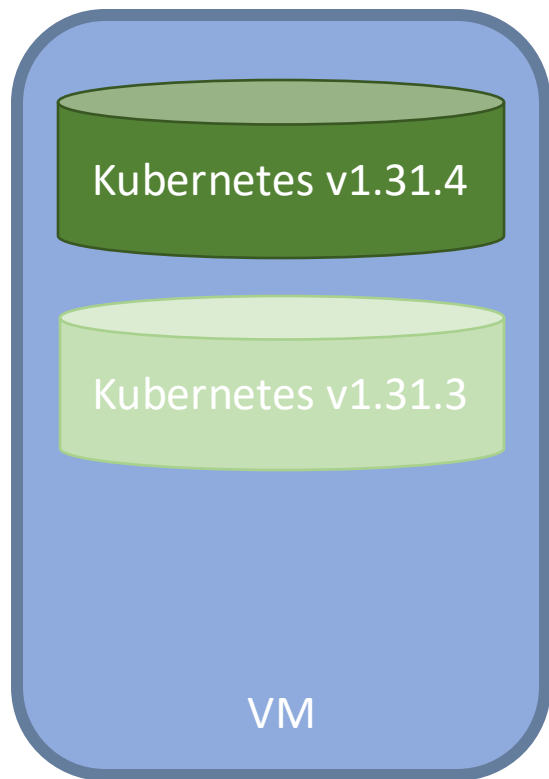


Instance

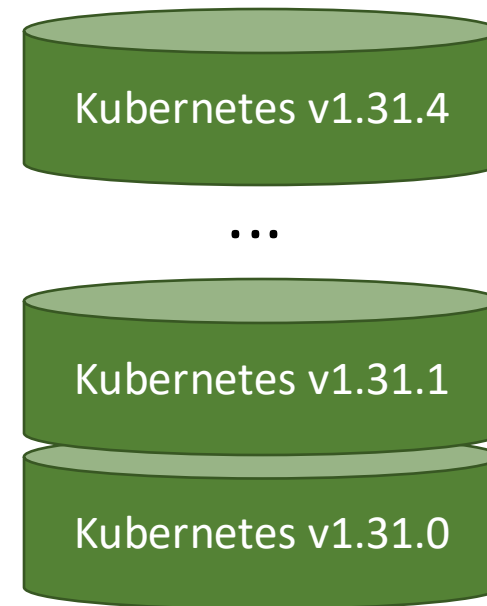


Update Server

In-Place Updates



Instance



Update Server

Show me!

Local Cluster from Scratch



Operating on Scale

AUTOMATE
ALL
THE UPDATES!



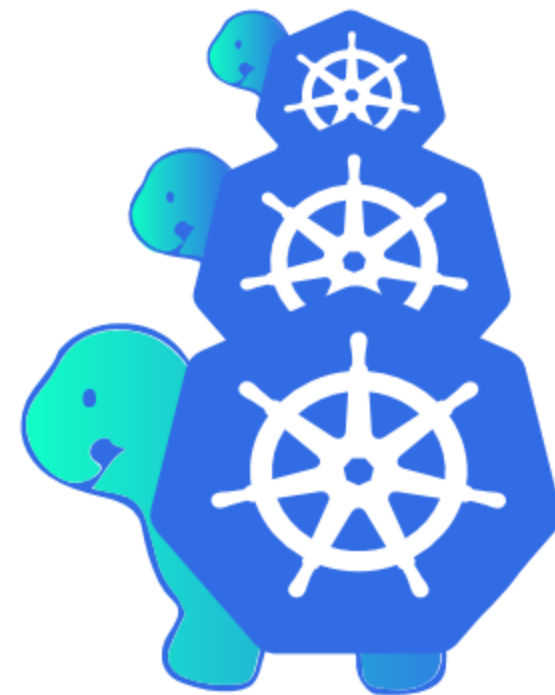
Scenario: large ClusterAPI deployment



Cluster Management System (not an API)

Provision / operate workload clusters from
management cluster

Think hosted Kubernetes offers

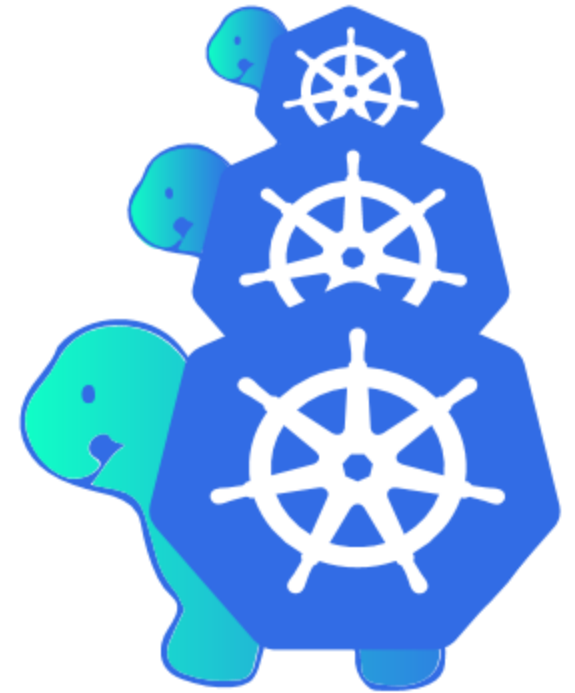


Prepared, but deactivated by default

CAPI workload cluster provisioning:

1. Install provider to management cluster
2. Generate workload cluster configuration
3. Apply config to management cluster

=> Patch workload cluster config after
2., before 3.



Kubernetes Reboot Daemon, CNCF project

<https://github.com/kubereboot/kured>

Drain, Reboot, Un-cordon, Done

Flexible configuration to cover special needs / corner cases



Show me!

Introspect CAPZ cluster, use Kured to in-place update



Thank you