

GPUs on OpenShift

A PATTERN TO ENABLE ACCELERATOR CARDS

Eduardo Arango

OpenShift Software engineer



1
hpckp'20
Barcelona



Kubernetes operators

Put the System-admin
knowledge into
Kubernetes

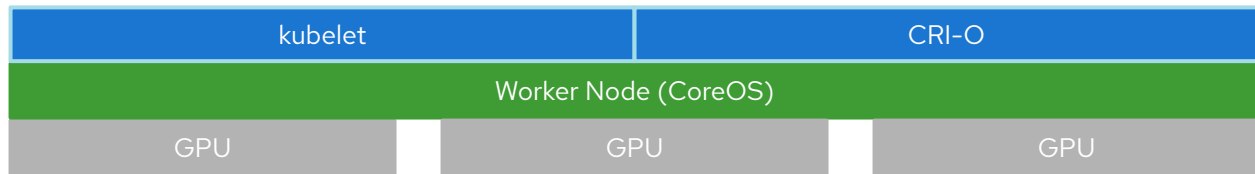
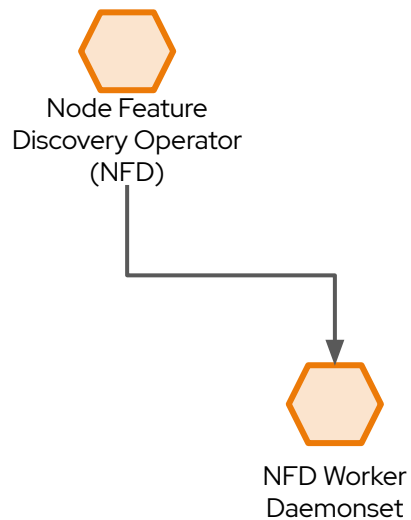
What is an Operator?

An Operator is a method of packaging, deploying and managing a Kubernetes-native application. A Kubernetes-native application is an application that is both deployed on Kubernetes and managed using the Kubernetes APIs and kubectl tooling.

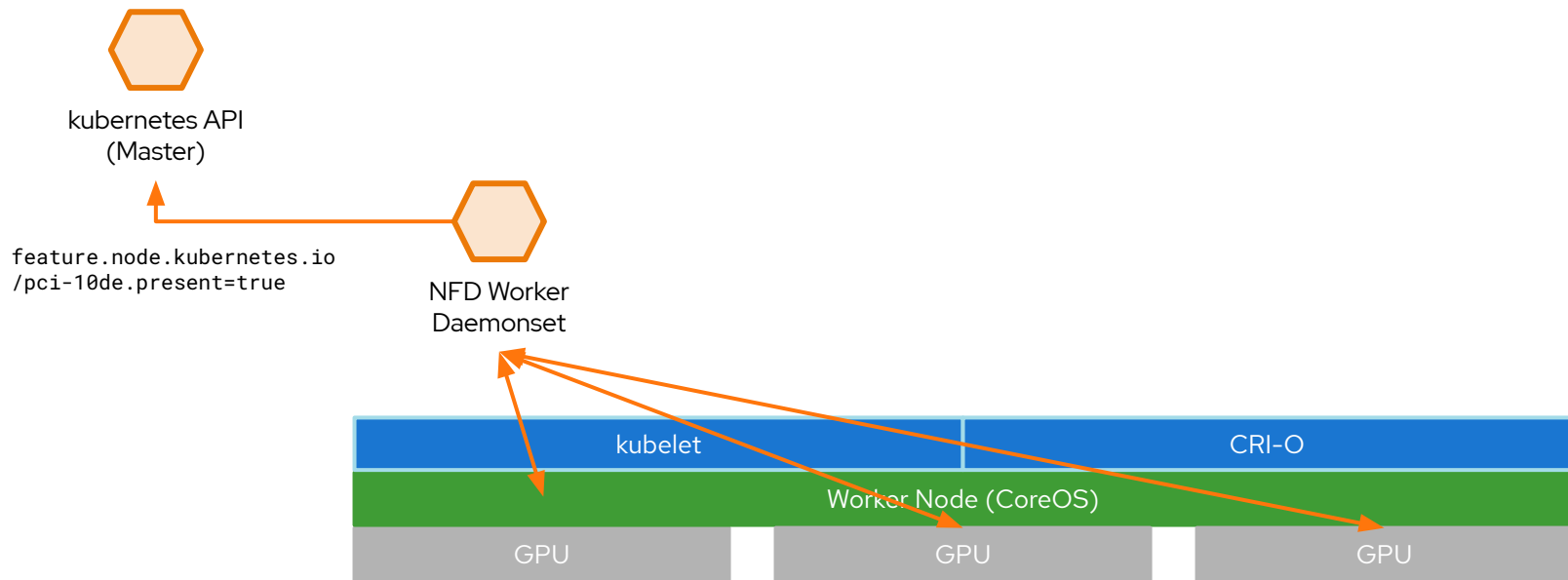
Special Resources and Devices

Enabling GPU, Network,
and other special
resources for workloads

NFD finds certain resources

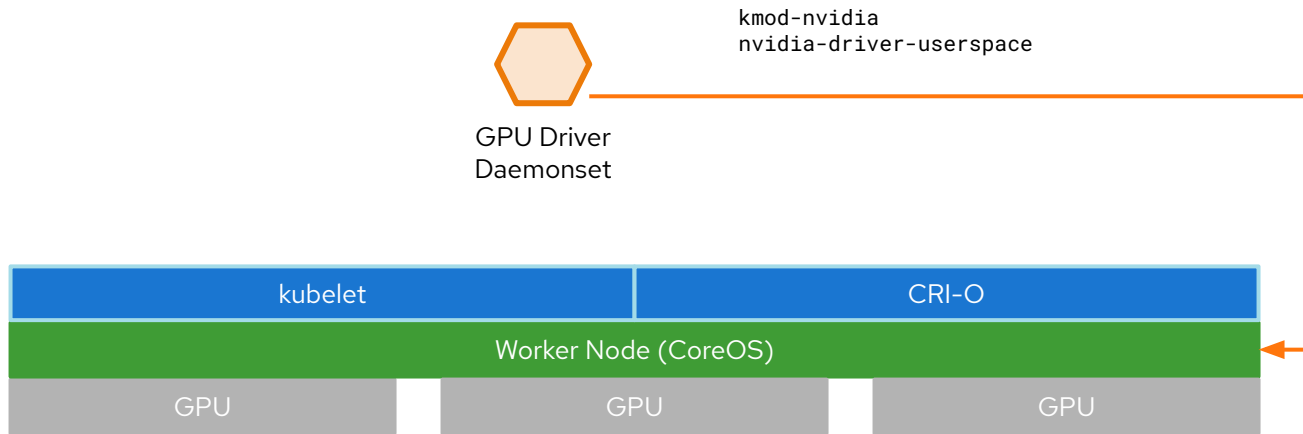


NFD labels nodes



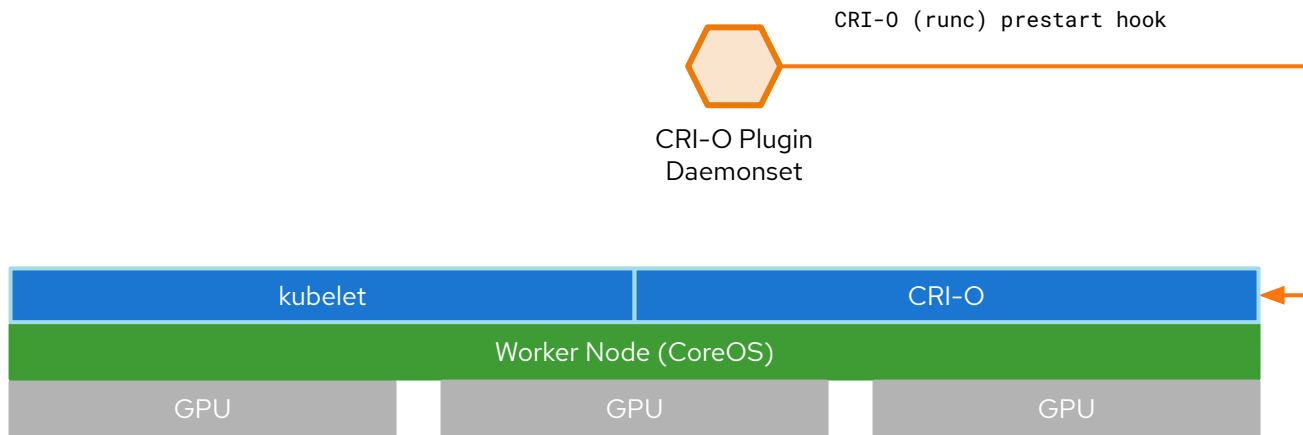
GPU Driver installs kmod and userspace drivers

feature.node.kubernetes.io
/pci-10de.present=true
...



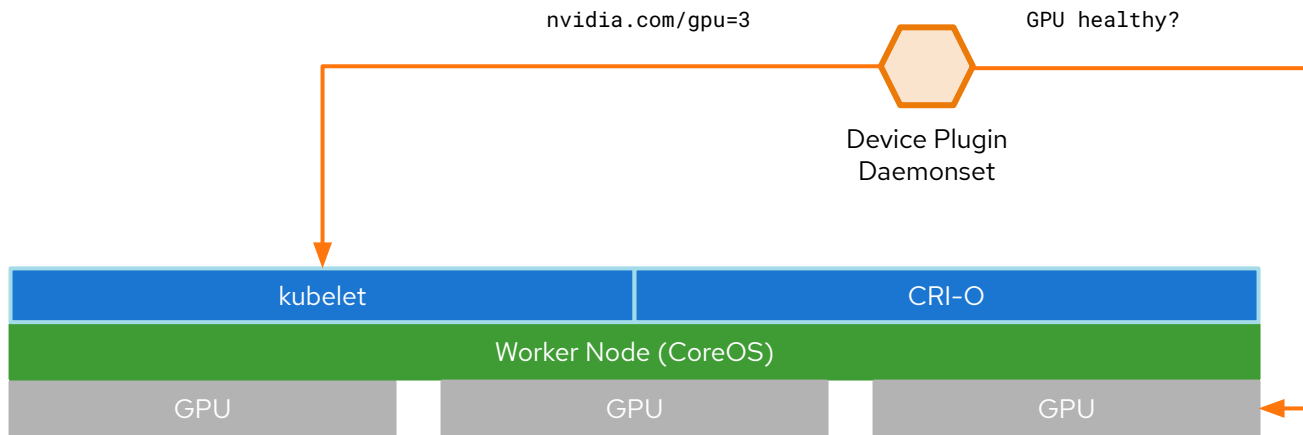
CRI-O Plugin installs prestart hook

feature.node.kubernetes.io
/pci-10de.present=true
...

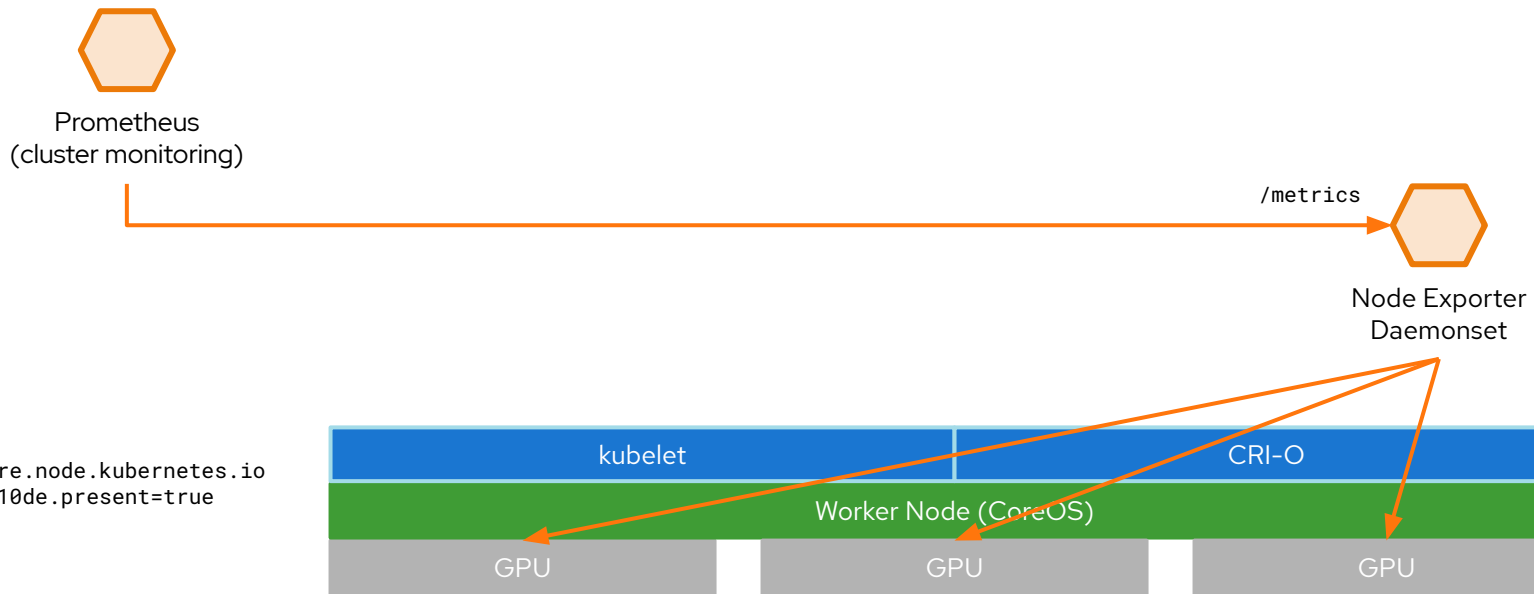


Device Plugin informs kubelet of resource details

feature.node.kubernetes.io
/pci-10de.present=true
...

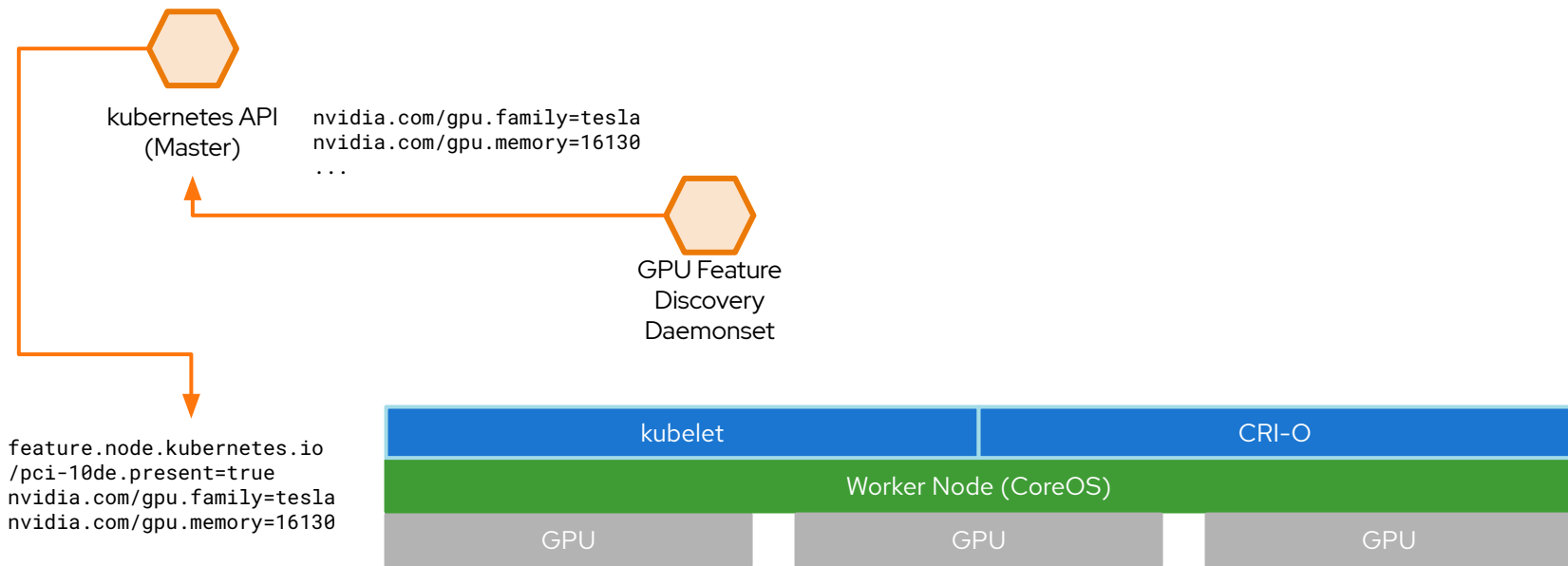


Node Exporter provides metrics on GPU

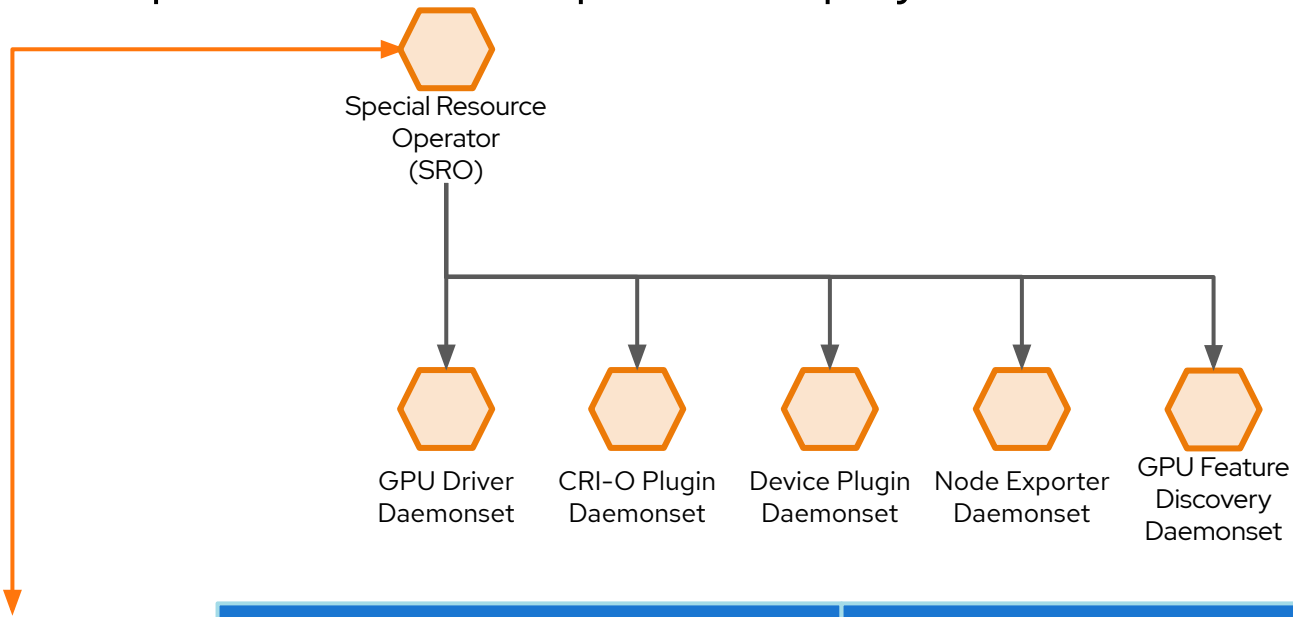


feature.node.kubernetes.io
/pci-10de.present=true
...

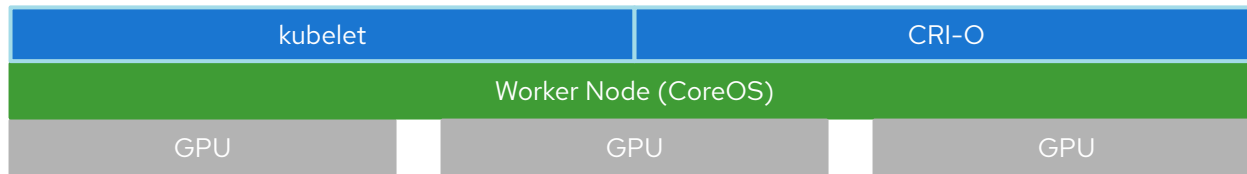
GPU Feature Discovery reports additional capabilities



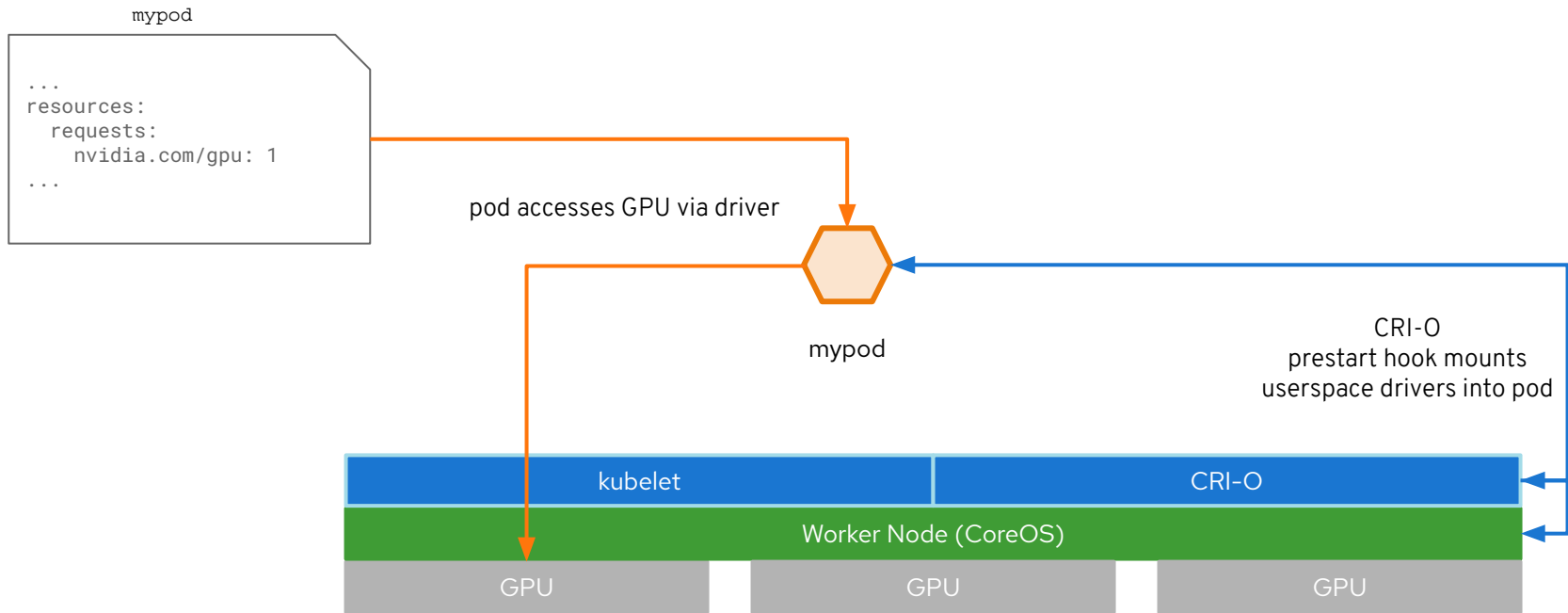
Special Resource Operator deploys to relevant nodes



`feature.node.kubernetes.io
/pci-10de.present=true
nvidia.com/gpu.family=tesla
nvidia.com/gpu.memory=16130`



GPU workload deployment



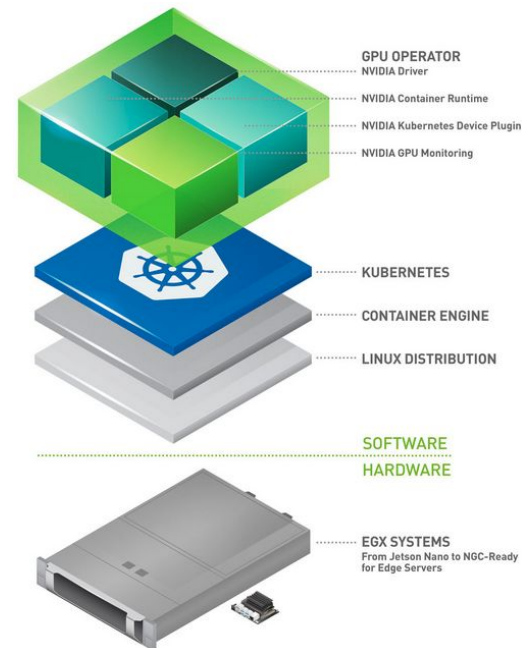
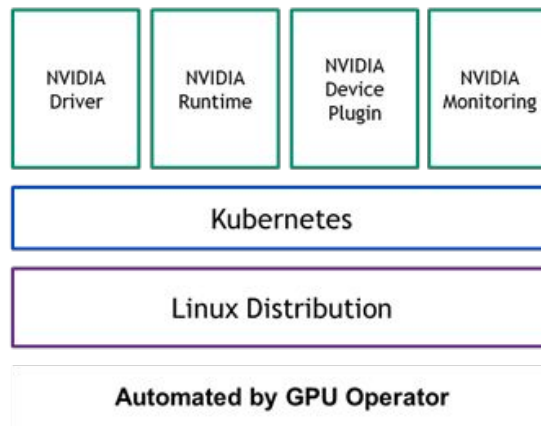
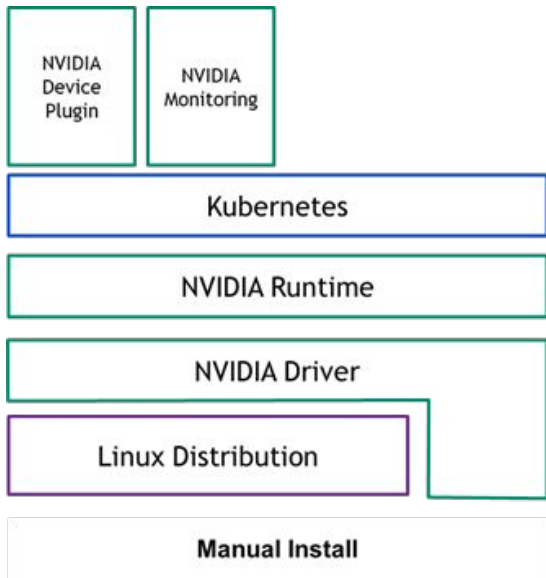
NVidia GPU operator GA

Simplifying GPU

Management in

Kubernetes

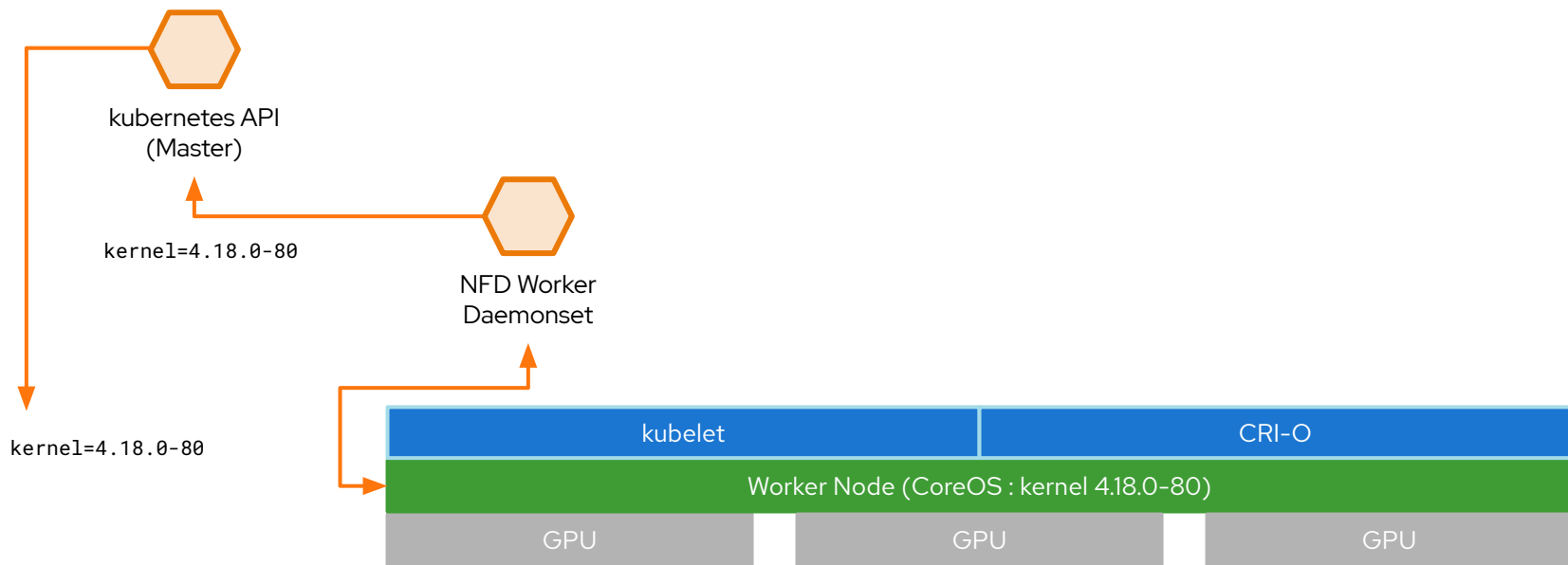
NVIDIA GPU Operator



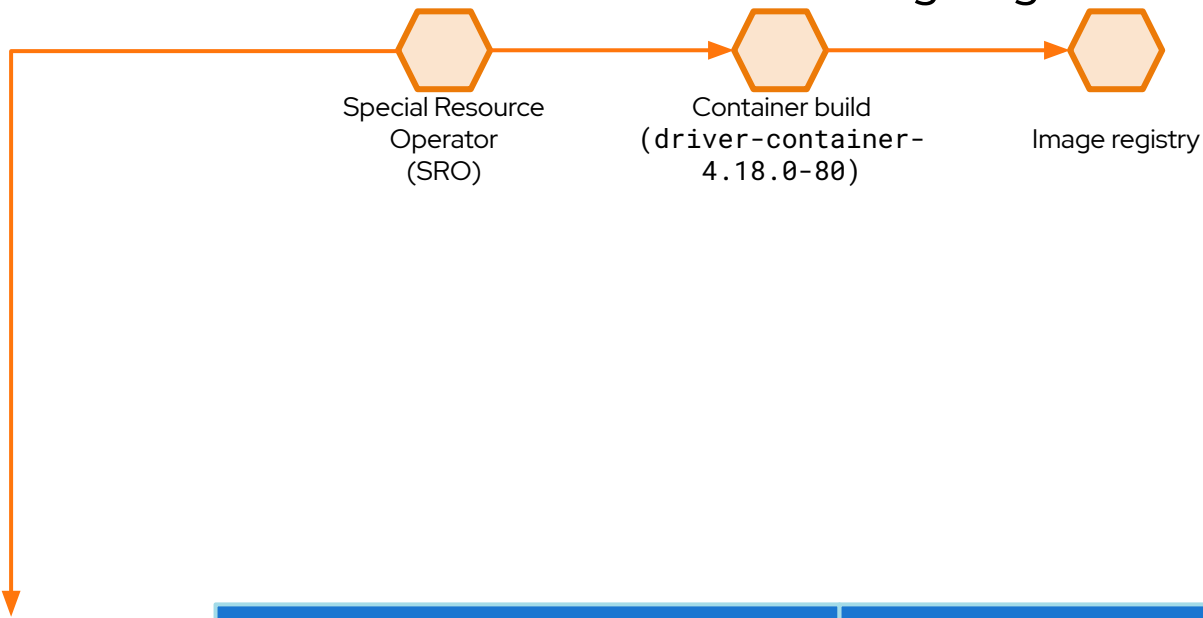
Automated Kernel Module Matching

Rebuilding kernel
modules on the fly to
correspond to system
updates

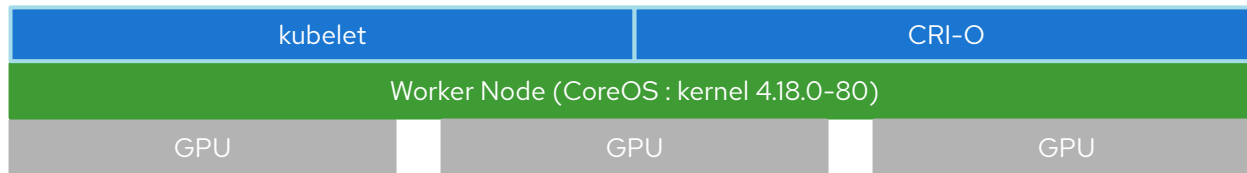
NFD detects kernel version and labels node



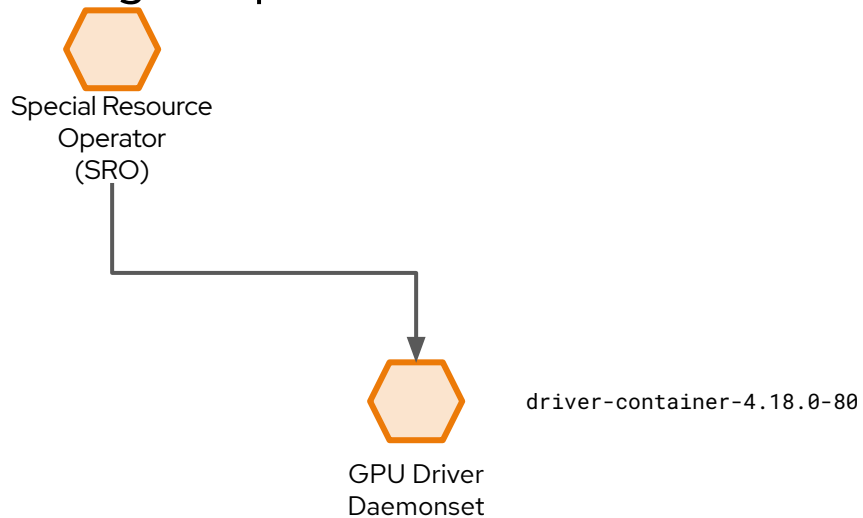
SRO builds driver container image against kernel



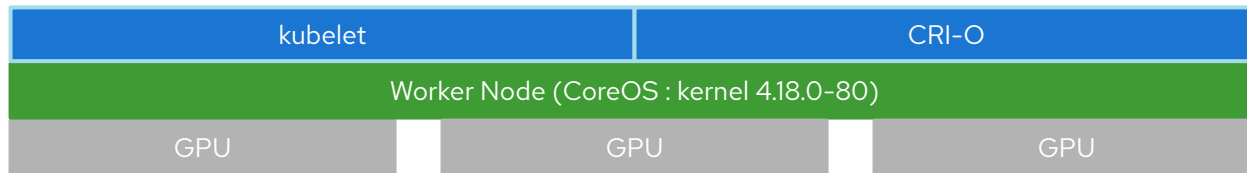
kernel=4.18.0-80



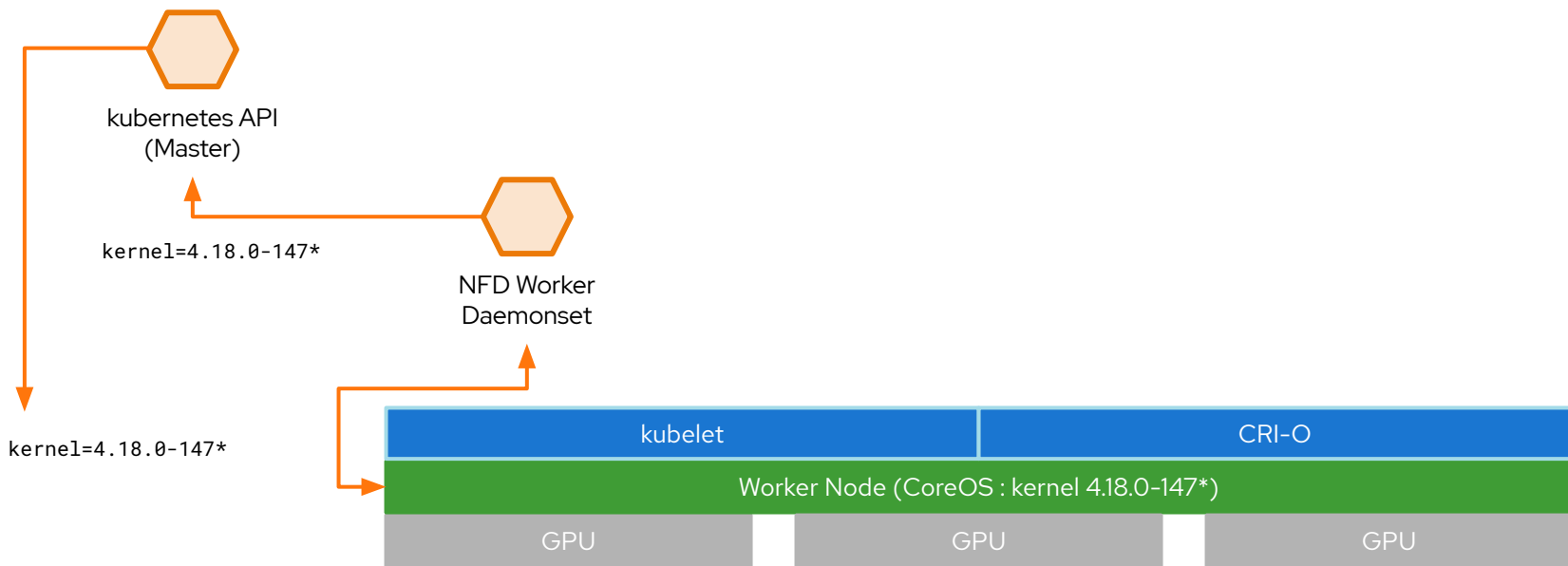
SRO targets specific kernel version hosts



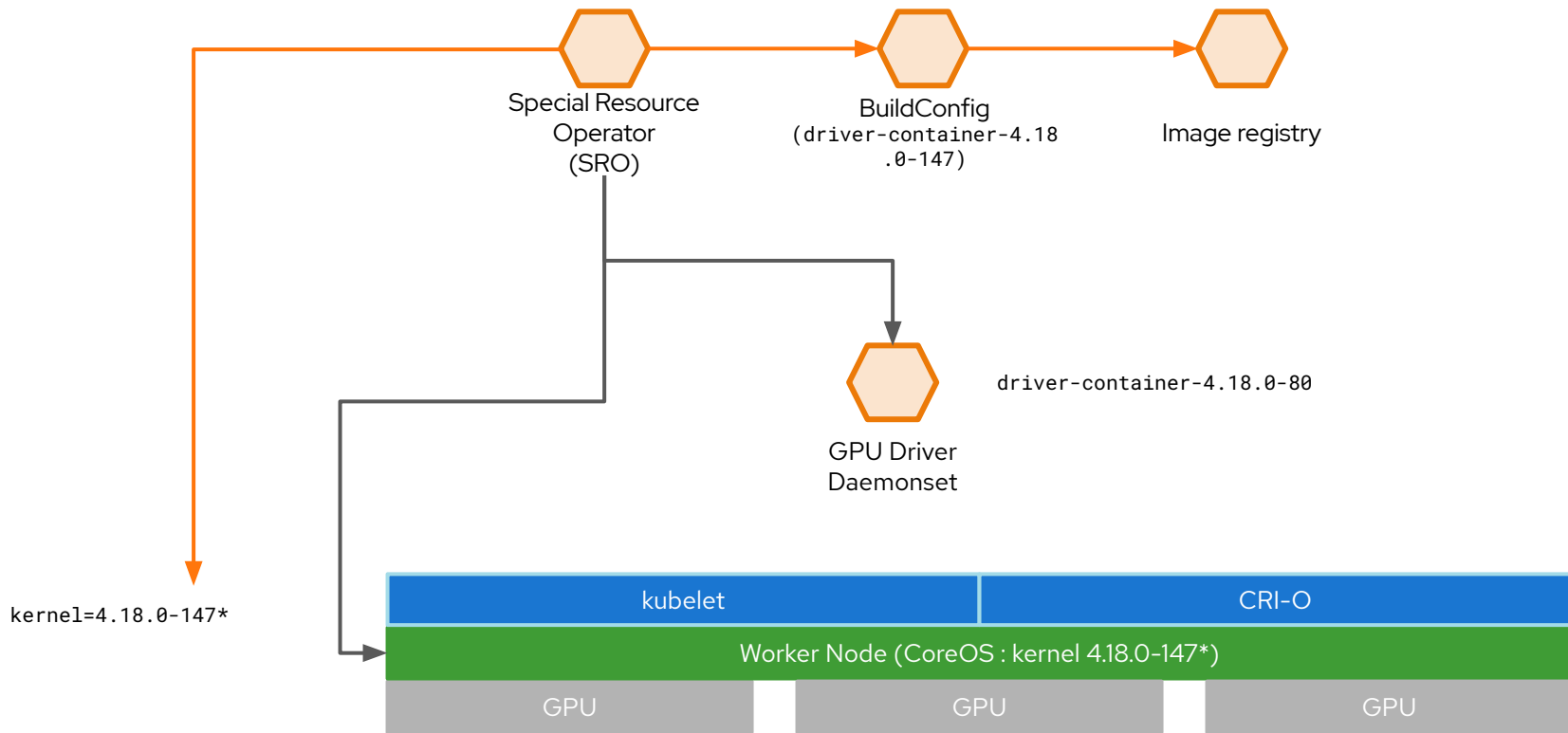
kernel=4.18.0-80



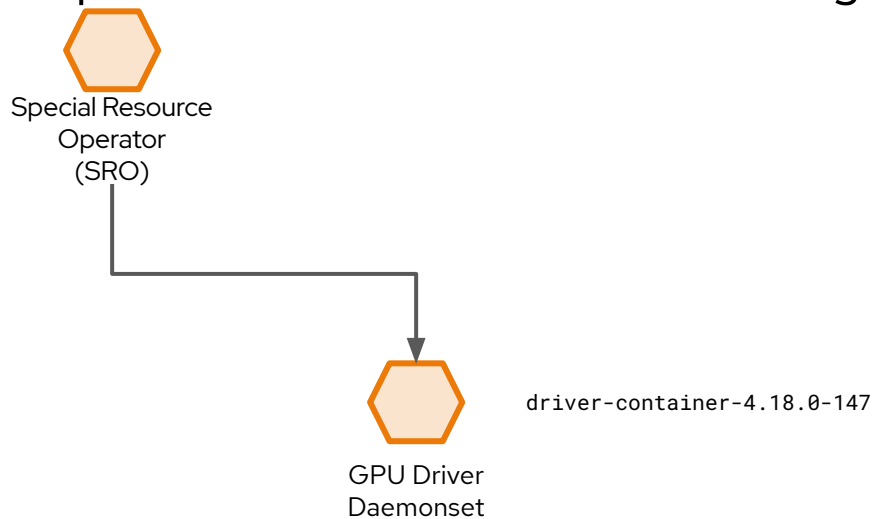
NFD detects updated kernel and relabels node



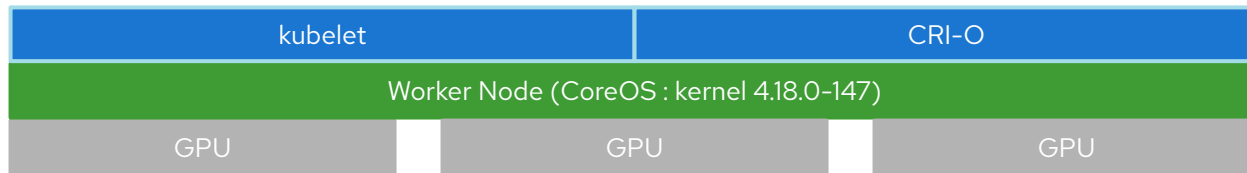
SRO detects mismatch and rebuilds driver container



SRO updates daemonset with new image



kernel=4.18.0-147



Some links:

NFD Operator

- **Building Multiarch ImageStreams with the NFD Operator and OpenShift 4:**
<https://www.openshift.com/blog/building-multiarch-imagestream-with-the-nfd-operator-and-openshift-4>

SRO Operator:

- **Part 1: How to Enable Hardware Accelerators on OpenShift:**
<https://www.openshift.com/blog/part-1-how-to-enable-hardware-accelerators-on-openshift>
- **Part 2: How to enable Hardware Accelerators on OpenShift, SRO Building Blocks:**
<https://www.openshift.com/blog/part-2-how-to-enable-hardware-accelerators-on-openshift-sro-building-blocks>

GPU Operator:

- **Simplifying deployments of accelerated AI workloads on Red Hat OpenShift with NVIDIA GPU Operator:**
<https://www.openshift.com/blog/simplifying-deployments-of-accelerated-ai-workloads-on-red-hat-openshift-with-nvidia-gpu-operator>

Thank you

Red Hat is the world's leading provider of enterprise open source software solutions. Award-winning support, training, and consulting services make Red Hat a trusted adviser to the Fortune 500.

 linkedin.com/company/red-hat

 youtube.com/user/RedHatVideos

 facebook.com/redhatinc

 twitter.com/RedHat