NTT

Usernetes Generation 2

Kubernetes in Rootless Docker

https://github.com/rootless-containers/usernetes

Akihiro Suda, NTT

Usernetes github.com/rootless-containers/usernetes

- Rootless Kubernetes
  - Even if an attacker has escaped from a Pod, or gained an access to kubelet API, the attacker still cannot gain the root privilege of the node
- Implemented by putting kubelet, CRI, OCI, CNI, etc. in a user namespace
  - UserNS: Linux kernel's feature that maps a non-root user to a fake root (the root privilege is limited inside the namespace)
- Multi-node networking is possible with VXLAN

History

- Began in 2018
  - As old as Rootless Docker (pre-release at that time) and Rootless Podman
- The changes to Kubernetes was merged in Kubernetes v1.22 (Aug 2021)
  - Feature gate: “KubeletInUserNamespace” (Alpha)
- The feature gate was also adopted by:
  - kind (with Rootless Docker or Rootless Podman)
  - Minikube (with Rootless Docker or Rootless Podman)
  - k3s

Usernetes Gen 1 vs Gen 2

"The hard way"

Similar to `kind` and minikube, but supports real multi-node

	Gen 1 (2018-2023)	Gen 2 (2023-)
Host dependency	RootlessKit	Rootless Docker, Rootless Podman, or Rootless nerdctl (contaiNERD CTL)
Supports kubeadm	No	Yes
Supports multi-node	Yes, but practically No, due to complexity	Yes
Supports hostPath volumes	Yes	Yes, for most paths, but needs an extra config

Usage

```
# Bootstrap the first node
make up
make kubeadm-init
make install-flannel
```

```
# Enable kubectl
make kubeconfig
export KUBECONFIG=$(pwd)/kubeconfig
kubectl get pods -A
```

```
# Multi-node
make join-command
scp join-command another-host:~/usernetes
ssh another-host make -C ~/usernetes up kubeadm-join
```