
KubeCon
CloudNativeCon
North America 2022
BUILDING FOR THE ROAD AHEAD
DETROIT 2022
TiKV
TiKV
HELM

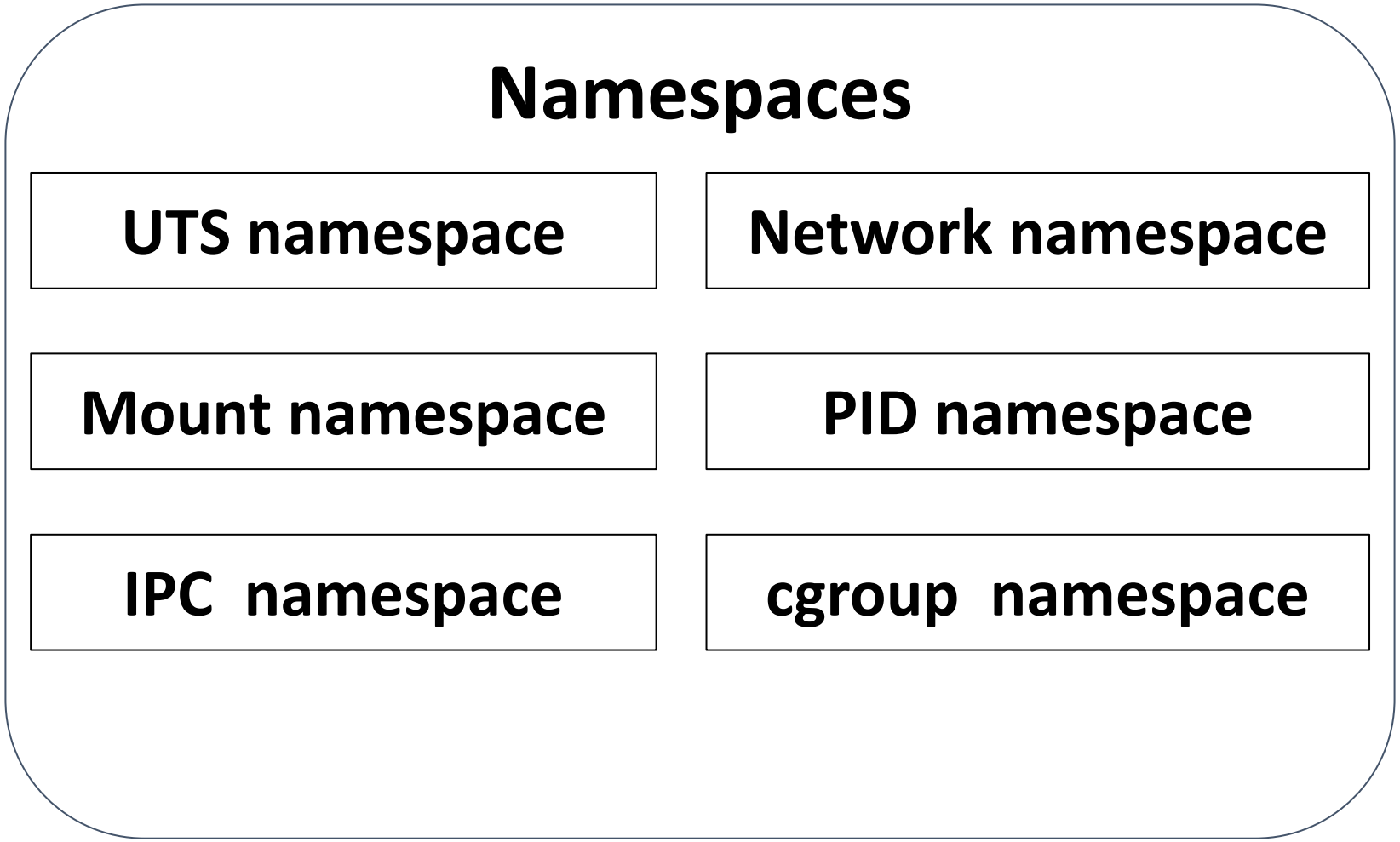
BUILDING FOR THE ROAD AHEAD

DETROIT 2022

Run as “root”, not root: user namespaces in Kubernetes

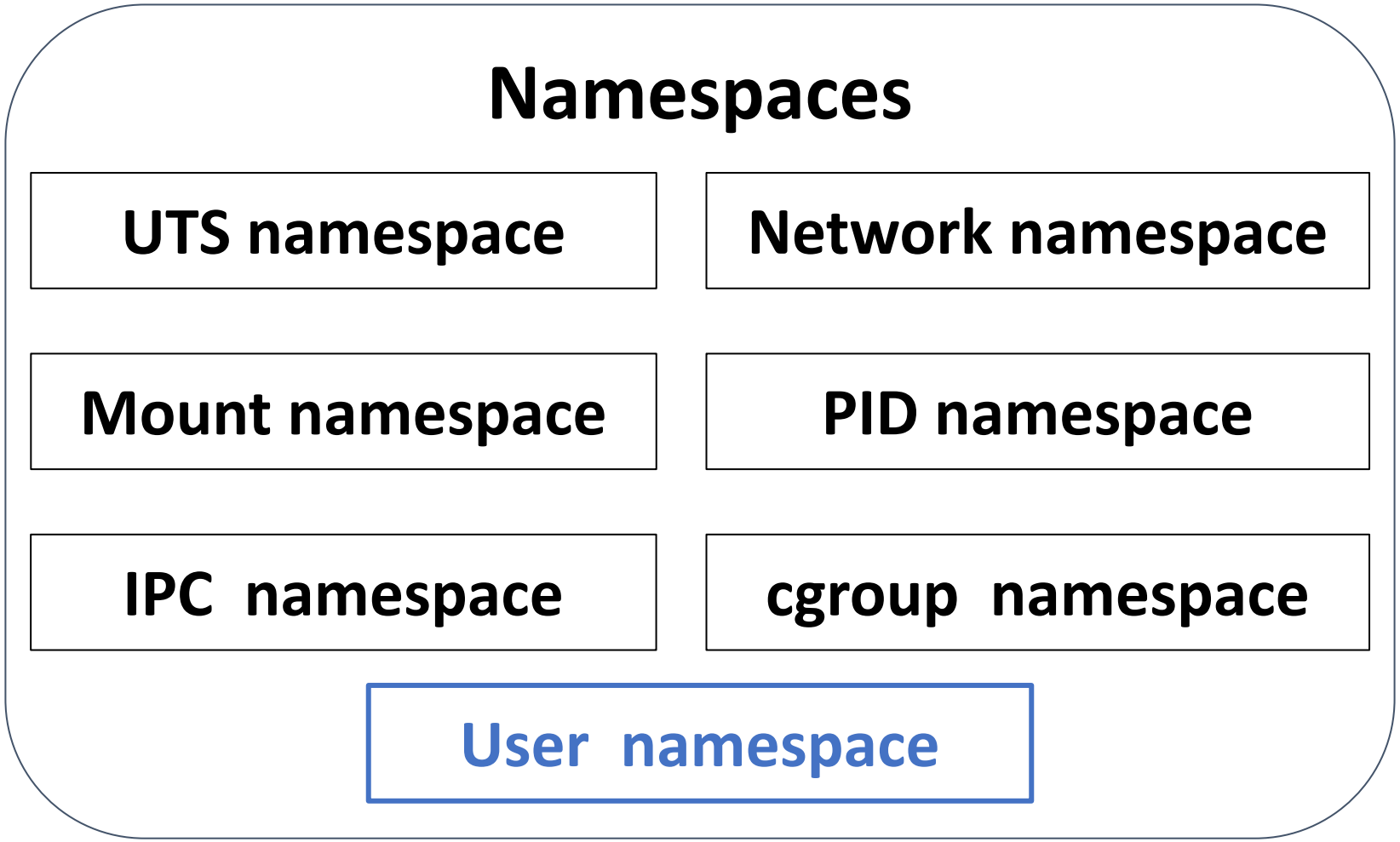
Rodrigo Campos Catelin
Margarita Manterola

Container isolation in k8s

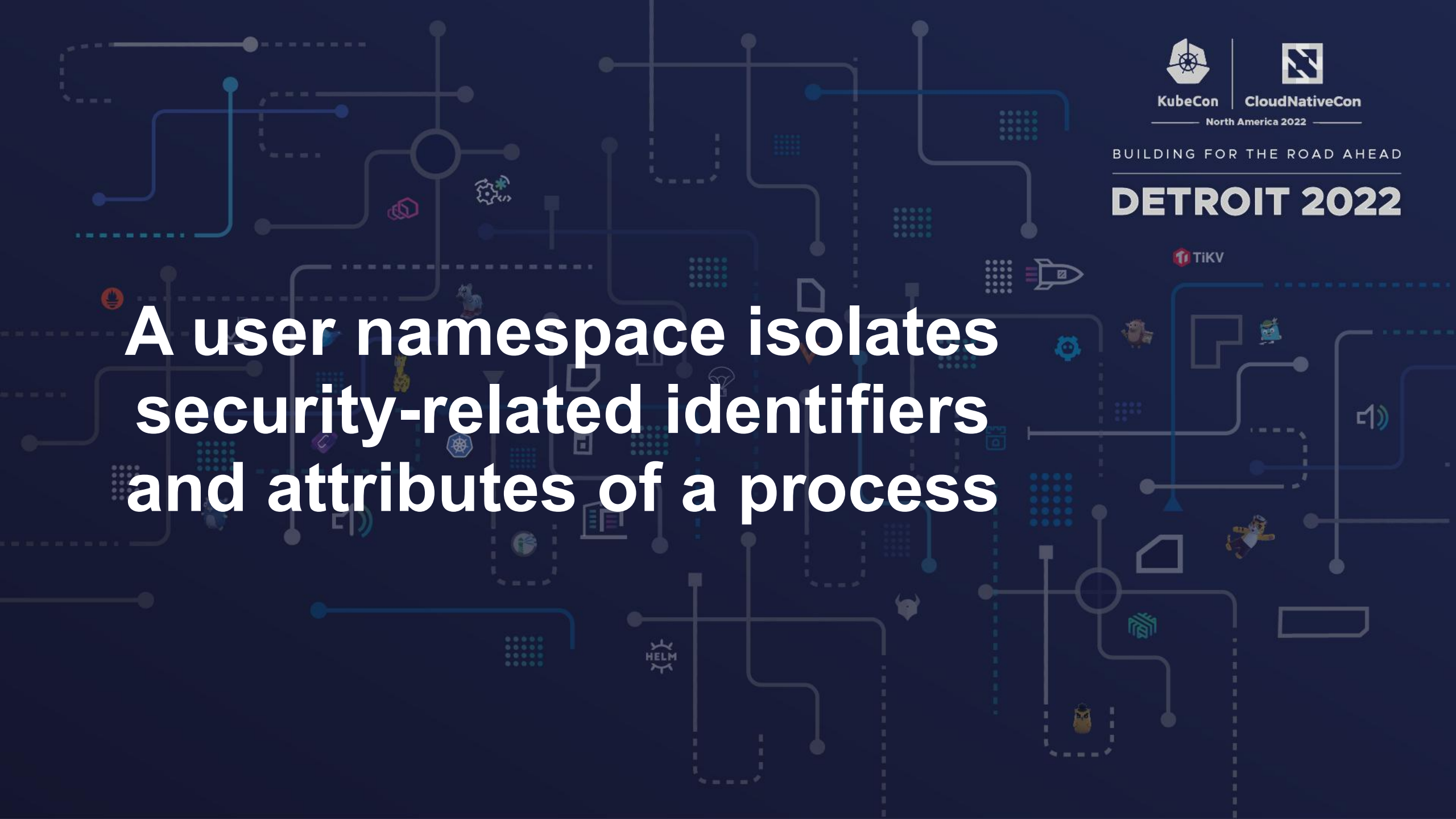


(up to k8s 1.24)

Container isolation in k8s

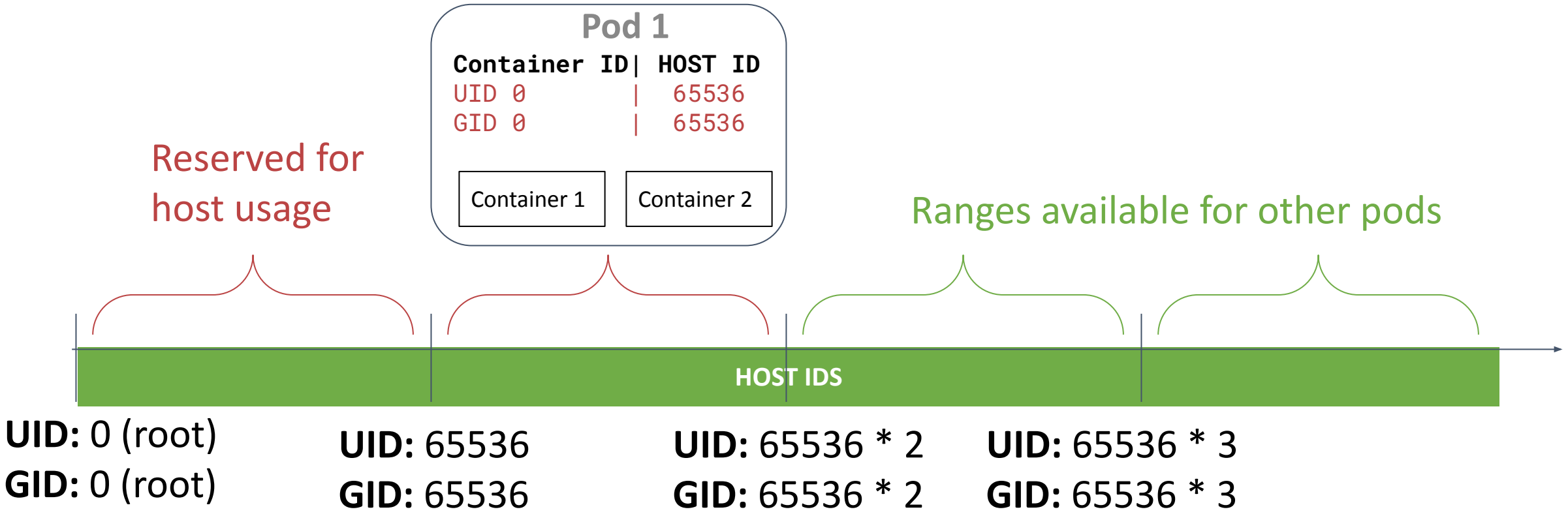


(starting from k8s 1.25)



A user namespace isolates security-related identifiers and attributes of a process

UID Distribution



YAML with user namespaces


```yaml
apiVersion: v1
kind: Pod
metadata:
 name: userns
spec:
 hostUsers: false
 containers:
 - name: shell
   command: ["sleep", "infinity"]
   image: debian
```

Enables user namespaces

Most workloads can enable this without doing any changes in the app

Mixing restrictions


```yaml
apiVersion: v1
kind: Pod
metadata:
 name: userns
spec:
 hostUsers: false
 containers:
 - name: shell
   command: ["sleep", "infinity"]
   image: debian
```

Can't be mixed with:

hostNetwork: true

hostIPC: true

hostPID: true

Mitigated Vulnerabilities

- CVE-2019-5736: Host runc binary can be overwritten from container
  - Score: 8.6 (HIGH)
- CVE-2022-0492: can containers escape? Arbitrary execution as root
  - Score: 7.8 (HIGH)
- Azurescape: This is the **first cross-account container takeover in the public cloud**.

Completely mitigated

- CVE-2017-1002101: subpath volume allows arbitrary file access in host filesystem
  - Score: 9.6 (CRITICAL) / 8.8 (HIGH)
- CVE-2021-30465: mount destinations can be swapped to cause mounts outside the rootfs
  - Score: 8.5 (HIGH)
- …

Mitigated

Root in the container is not root on the host

References

- **KEP 127**: support for user namespaces in pods
  - Long history, the efforts started in 2016!
  - https://kep.k8s.io/127
- Containerd user namespaces support (**#7063**):
  - https://github.com/containerd/containerd/issues/7063
- cri-dockerd support issue (**#74**):
  - https://github.com/Mirantis/cri-dockerd/issues/74
- CAP_SYS_ADMIN: the new root: https://lwn.net/Articles/486306/
- `man 7 user_namespaces`