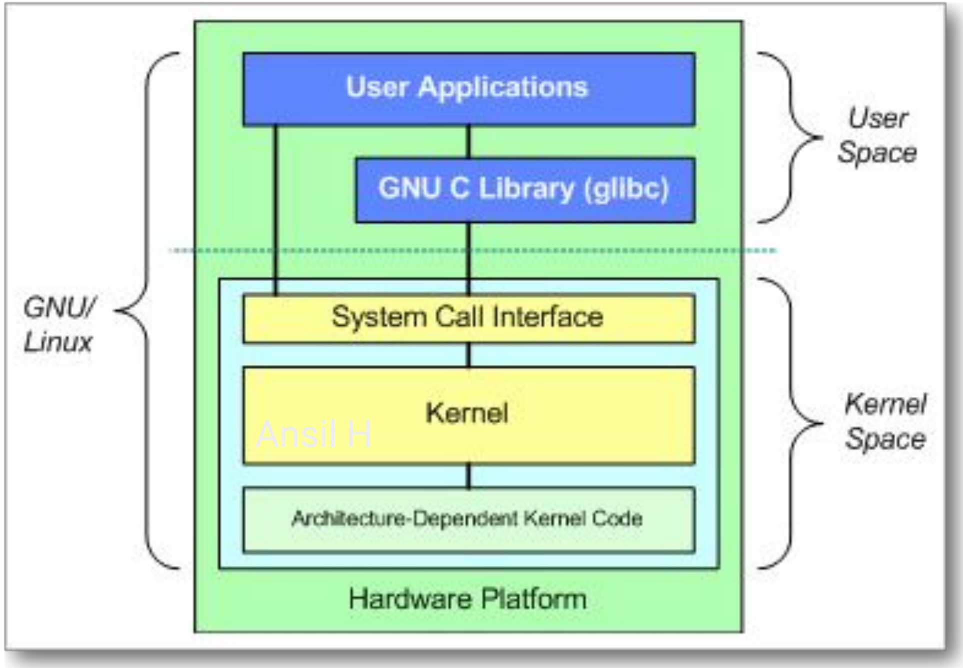
Container Networking Fundamentals - Part 1

Ansil H
Lead SRE@Armorblox

Linux OS fundamentals



Why we need containers?

- Process isolation for security
- Resource usage restriction
- Dependency management
- Lifecycle management

The Building Blocks

- Namespaces
- Cgroups
- Capabilities

Namespaces

Namespace	Constant	Isolates
Cgroup	CLONE_NEWCGROUP	Cgroup root directory
IPC	CLONE_NEWIPC	System V IPC, POSIX message queues
Network	CLONE_NEWNET	Network devices, stacks, ports, etc.
Mount	CLONE_NEWNS	Mount points
PID	CLONE_NEWPID	Process IDs
User	CLONE_NEWUSER	User and group IDs
UTS	CLONE_NEWUTS	Hostname and NIS domain name

CGroups or Control Groups

- Resource Limiting
- Prioritization
- Accounting
- Control/Freeze

Linux Capabilities

Root user can do anything but, can we have an “admin” user with restricted access to some of the privileged operations?

“Split privileged kernel calls and group them into related functionality”

Eg:- The “ping” binary may not work without **`CAP_NET_RAW`** capability.

Demo - Container from scratch

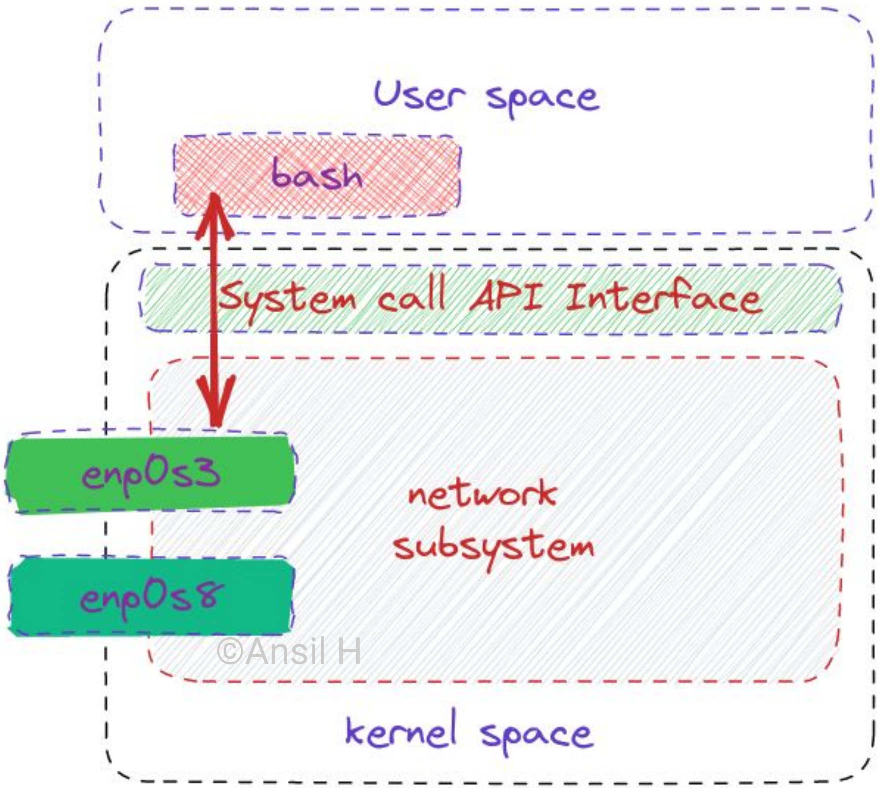
Let’s create a container from scratch using the “unshare” command.

In this demo, we will use “busybox” to simulate the needed binaries for the container.

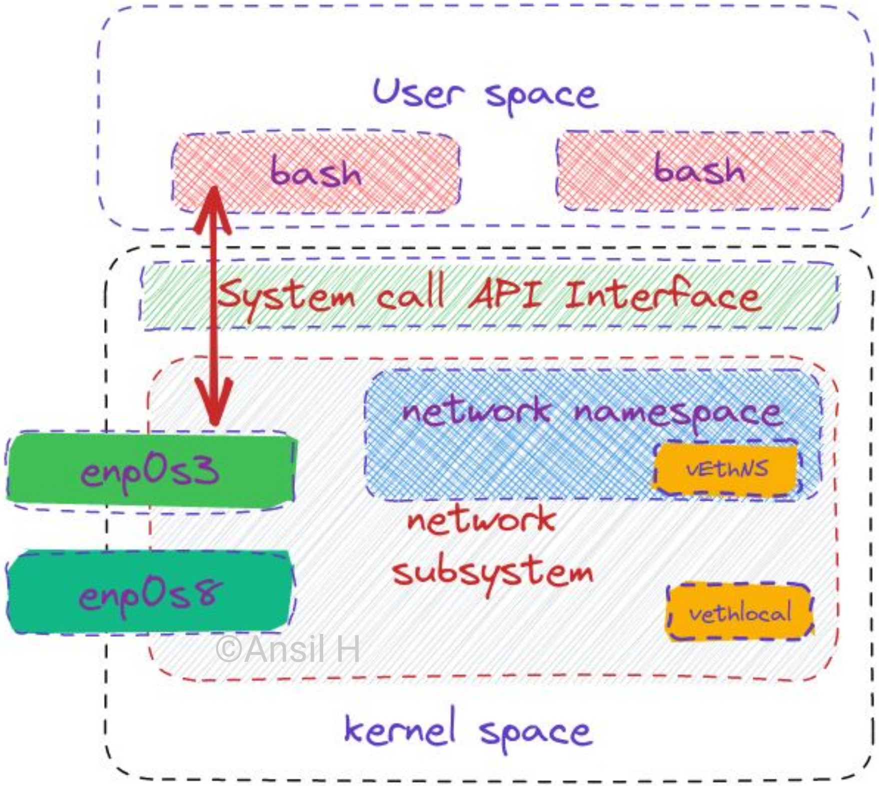
Container networking

vEth Pair

Networking - Host view

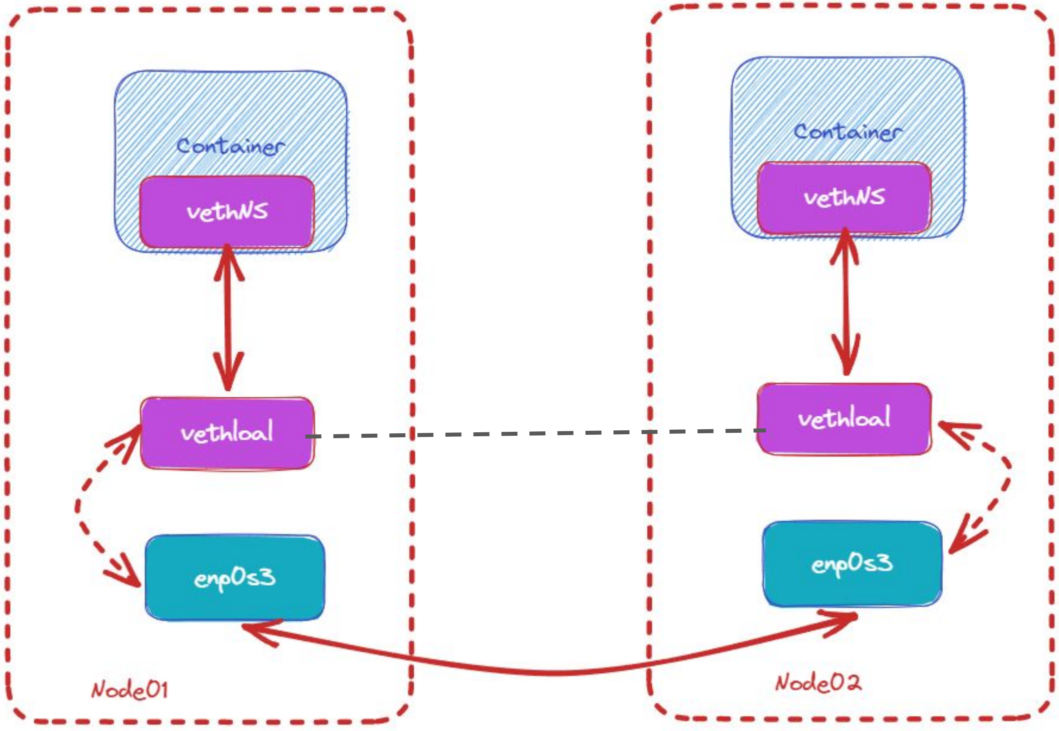


Network namespace and vETH pair



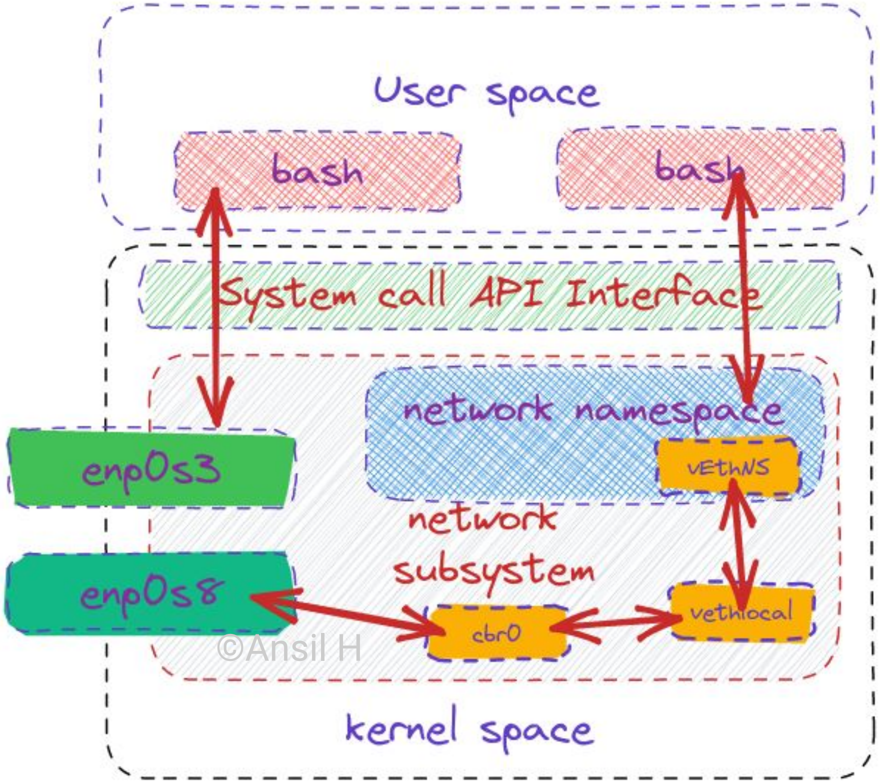
Routing

Routing

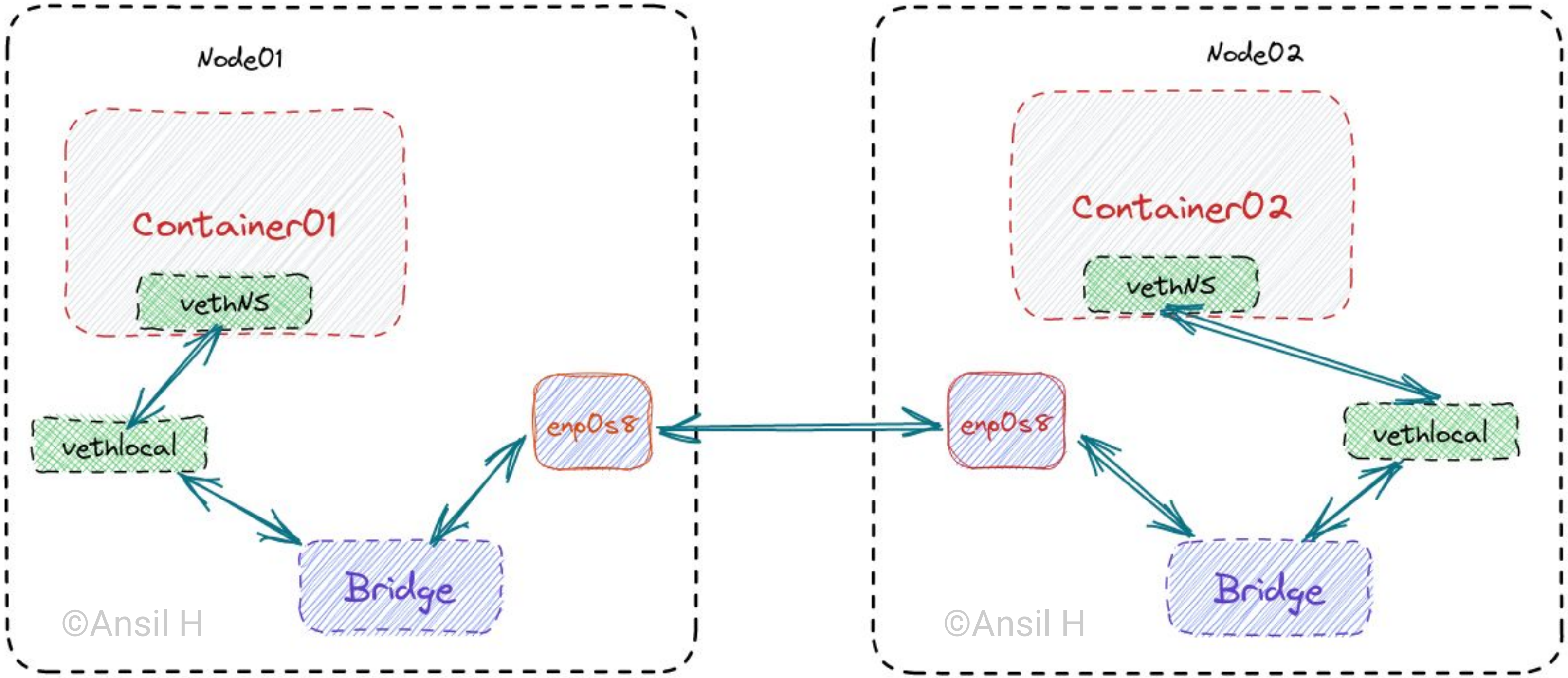

Container
vethNS
vethloal
enp0s3
Node01
Container
vethNS
vethloal
enp0s3
Node02

Bridging

Bridging


User space
bash
bash
System call API Interface
network namespace
vEthNS
enp0s3
network
subsystem
enp0s8
cbr0
vethlocal
©Ansil H
kernel space

Bridging - Demo



Bridging - Demo

In this demo, we will use two hosts called node01 and node02

A container will be started on both nodes using “busybox” and “unshare” commands.

A vEth pair will be created on both nodes

A virtual bridge device will be created on both nodes and will be connected to one of the ethernet device.

Finally one end of the vEth will be connected to the container and other end will be connected to the bridge.

Q&A