- K8s makes opinionated choices about how pods are networked, to be specific the following requirements should be fulfilled by a networking implementation
  - all Pods can communicate with all other Pods without using network address translation.
  - All Nodes can communicate with all Pods without NAT
  - the IP that Pod sees itself is the same IP that others see it as.
- Given these constraints we are left with four distinct networking problems to solve
  - Container-to-Container Network
  - Pod-to-Pod networking
  - Pod-to-Service Networking
 - Internet-to-Service Networking
- NAT Overview
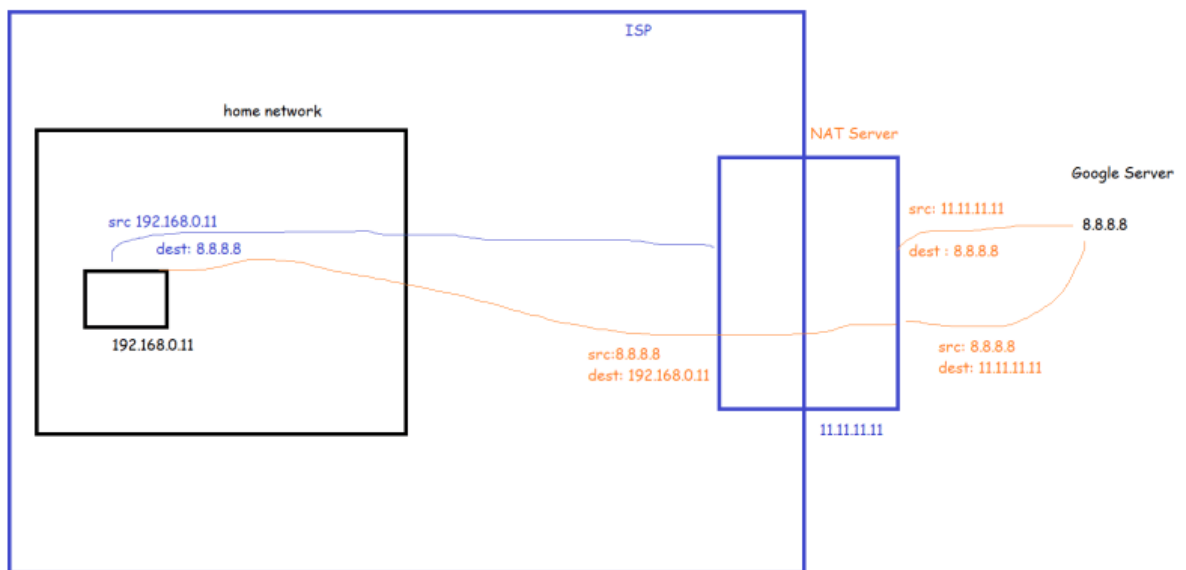
Container to Container Networking

- In Linux, each running process communicates with *network namespace* that provides a networking stack with its own routes, firewall rules and network devices
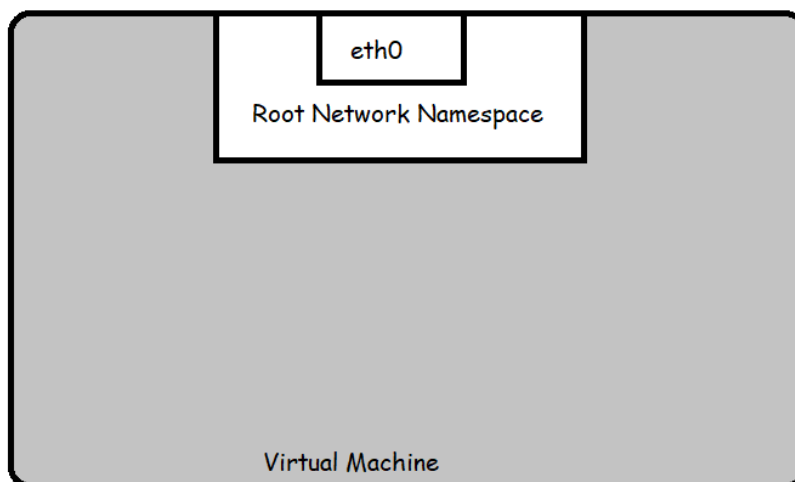

- A Pod is modelled as group of Docker Containers that share same network name space.

- Containers within a Pod all have same IP address and port space assigned through network namespace assigned to the pod and can find each other via localhost
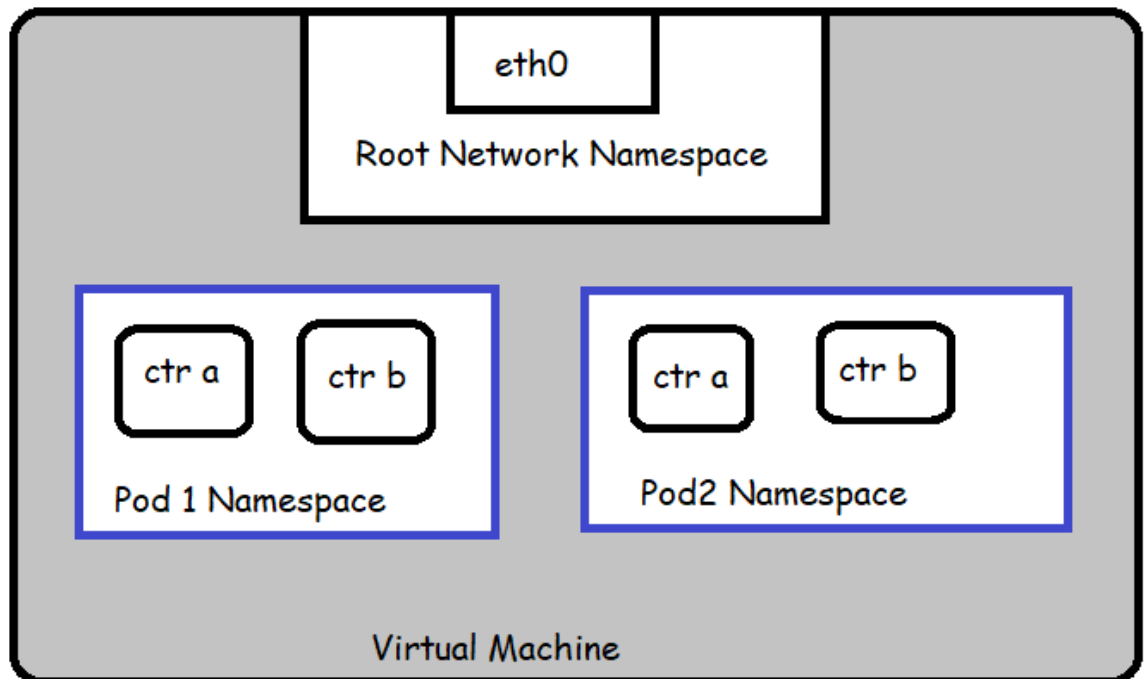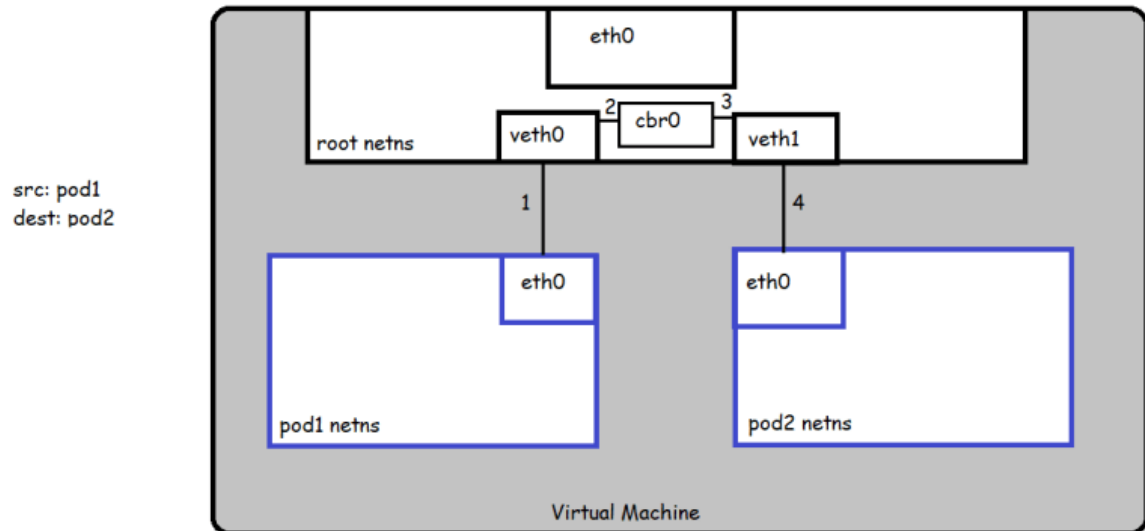

Pod to Pod Networking

- From the pod's perspective, it exists in its own network namespace. Linux Network namespaces can be connected using Linux Virtual Ethernet Device Pair or veth pair
- We want Pods to talk with each other through the root namespace and for this we use a network bridge.
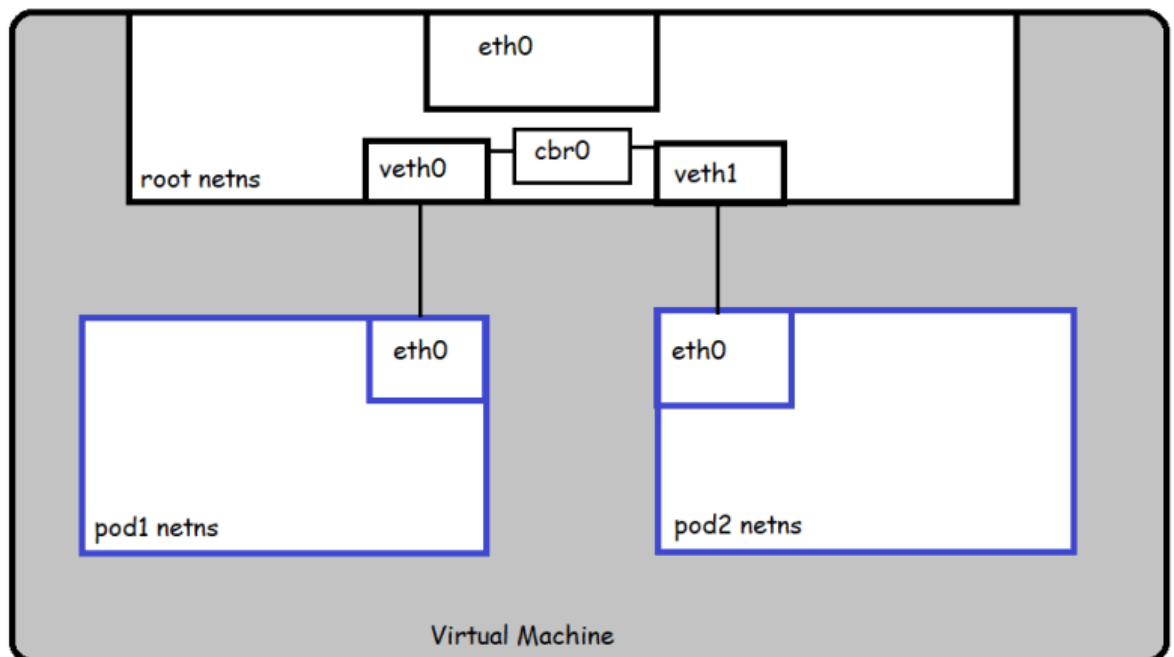
- Bridge implements ARP protocol
- Pod to Pod Networking within Same Node:

  Packets moving b/w pods on same node

  eth0

  root netns 2 veth0 cbr0 3 veth1

  src: pod1
  dest: pod2

  1 eth0 4 eth0

  pod1 netns pod2 netns

  Virtual Machine

- Let's understand the flow
  1. Pod1 sends a packet to its own ethernet device eth0 which is connected via virtual ethernet device to root namespace
  2. The bridge cbr0 is configured with veth0 a network segment attached to it.
  3. Once the packet reaches the bridge, the bridge resolves the network segment to send the packet to veth1 using ARP Protocol
  4. After resolution and the packet once reaches the virtual device veth1 will be forwarded to Pod2's namespace and eth0 device of pod2
- Pod to Pod networking across Nodes
  o Generally, every Node in our cluster is assigned a CIDR block specifying the IP addresses available to Pods running on that node.
  o To be continued

a system for connecting a number of computer systems to form a local area network, with protocols to control the passing of information and to avoid simultaneous transmission by two or more systems.

ARP protocol:

ARP stand for address resolution protocol, it maintains table entry with member. For any new incoming request, it forward packet to all members, once it receive response than it respond to ARP protocol.