Open questions around vitualized IPsec

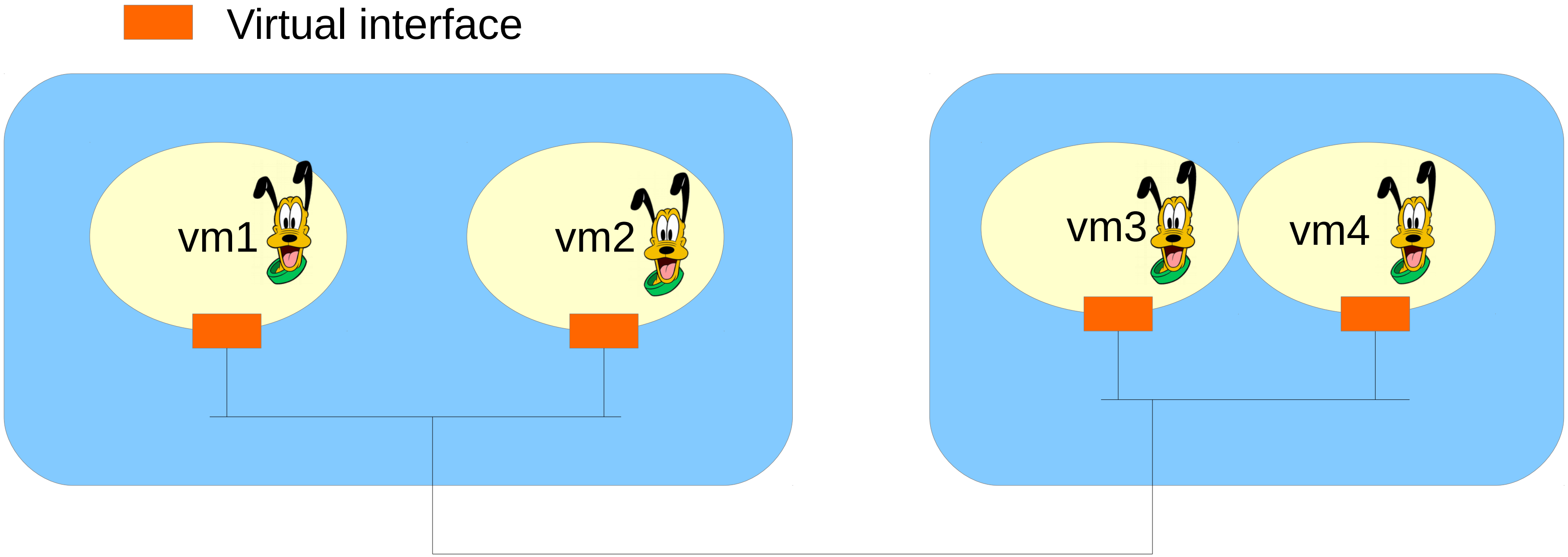
Sowmini Varadhan (sowmini.varadhan@oracle.com)

Linux IPsec workshop, March 2018, Dresden, Germany

Definitions

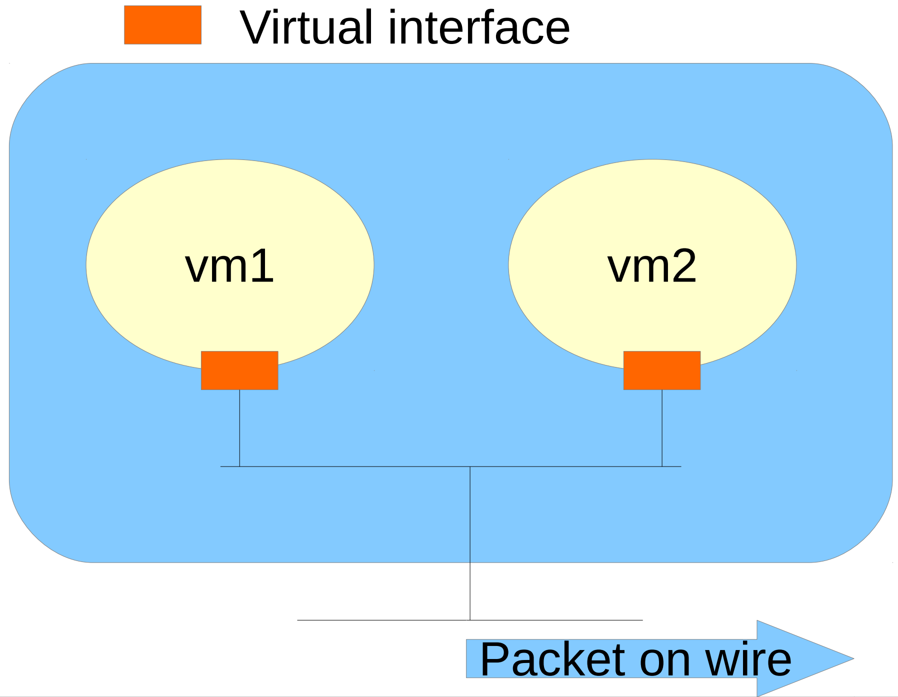
- Hypervisor (aka Virtual Machine Monitor): software, firmware or hardware that creates and runs virtual machines
- Without loss of generality, we use “pluto” (libreswan IKE daemon) as a typical example of a standards-conformant IKE implementation in the slides that follow.
- Host-terminated IPsec: IKE daemon like pluto runs inside the virtual machine. The VM is fully aware of the IPsec SADB and SPD
- Device-terminated IPsec: IPsec transforms are done outside the VM, in the hypervisor.
- Virtual Interface: network device assigned to the Virtual Machine- can be an SRIOV VF, Xen netback driver, member of veth pair, macvlan, 802.1q interface …

Host terminated IPsec



- Pluto runs inside each virtual machine
- IPsec association is between IP addresses owned by the VM (e.g., vm1’s ipaddr and vm3’s ipaddr).
- There may not even be an underlay encapsulation
- The scope of the SPI numbering space is within the VM, so we may well have an IPsec tunnel using SPI “X” between vm1 ↔ vm3, as well as vm2 ↔ vm4
- Offload needs to track both the SPI **and** a unique identifier for the vm (vlan and mac address)

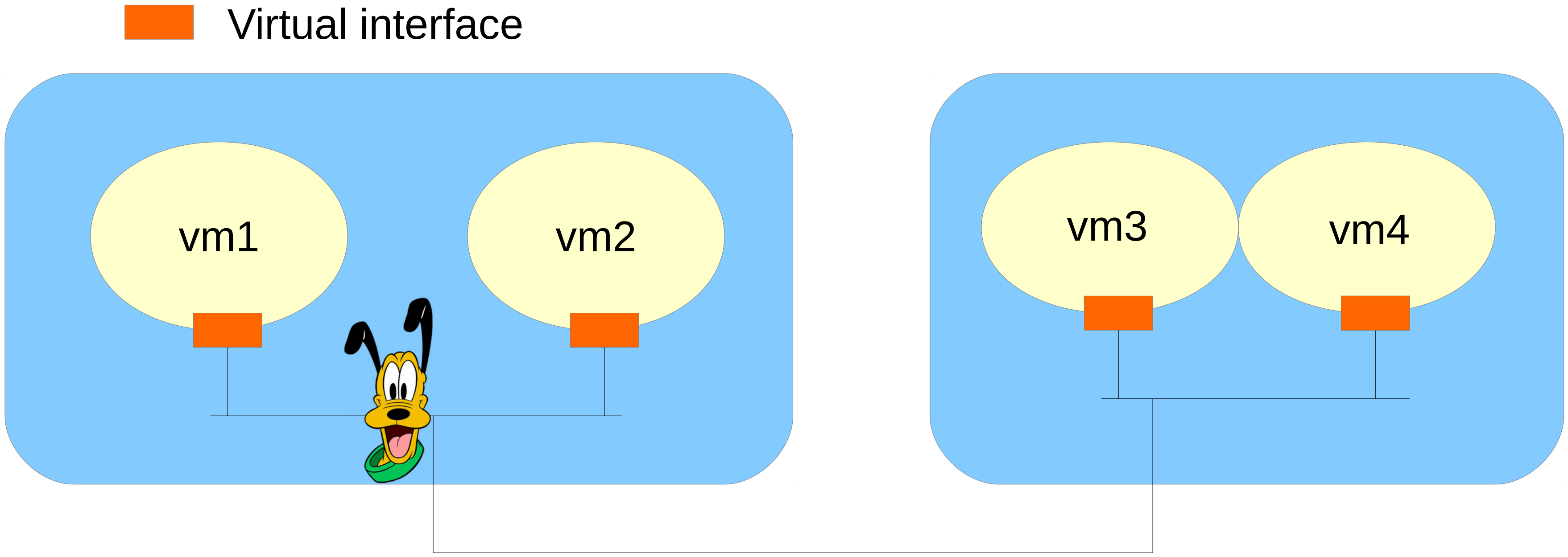
Interaction with other encapsulations



Underlay IP And ether header	UDP hdr	Vxlan hdr	vm1 tenant frame with overlay IP header and TCP payload

- e.g VXLAN: For clear traffic, hypervisor will select UDP src port based on fields in tenant frame; tenant frame may be encrypted (in the host-terminated Ipsec model), so UDP src port selection needs to make sure we have the desired entropy in the SPI
- What if vm1 wants TSO offload of the (overlay) TCP packet, and hypervisor needs to enforce IPsec offload of the (underlay) UDP/VXLAN packet?

Device terminated IPsec



- Pluto runs on the hypervisor
- If there is some type of underlay, and the IPsec association is between underlay IP address and remote node’s IP address, pluto/IKE config is straight-forward, can be done with existing support for VTI/VPN etc.
- How will control plane work if there is no underlay (e.g., if the above is a flat L2 subnet) and the hypervisor does not really “own” the outer-most IP address of the outgoing packet?