Nithin Raju, Sairam Venugopal
VMware, Inc.
Alin Serdean, Alin Balutoiu
Cloudbase Solutions Srl
OVS on Microsoft Hyper-V
Agenda

• Introduction & Architecture
• Stateful firewall & Demo
• Container support
• Roadmap
• Questions
What is it?

- Full fledged OVS solution for Microsoft Hyper-V
- Windows 8 / Server 2012+
- Collaboration b/n VMware Inc. and Cloudbase Solutions Srl
- Consists of:
  - Ported OVS userspace code
  - Windows datapath as an NDIS driver (OVSEXT)
- Forwarding extension on Microsoft Hyper-V Switch
Architecture

Root Partition (Host)

- `ovs-*ctl`
- `ovs-vswitchd`
- `dpif-netlink`
- `netdev-windows`
- `netlink socket(emulation)`

Child Partitions (Guest)

- `Hyper-V Machine #1`
- `Virtual Machine #2`
- `Hyper-V extensible switch`
- `Physical NIC`

NDIS Stack

Interface device
- Netlink Message Impl.
- Flowtable
- vport table
- Packet Processing

OVS Forwarding Extension
• Registered as a forwarding filter driver
• Registers callbacks for switch management and datapath functions
• Netlink messages over a pseudo device for user-kernel comm.
• Deep/Partial copy between NBLs with ref counting (NDIS limitation)
• Packet flow:
  – Packet from VM enters the switch ingress path and gets sent to OVSEXT driver
  – OVSEXT does flow lookup and sends to userspace upon miss
  – Flow gets added to kernel and packet gets forwarded to the NDIS stack
Features

- Supports VLAN, STT, GRE, VXLAN and Geneve tunnels
- L2/L3 header overwrite
- Offloads: Checksum, TSO, STT (LSO & LRO)
- Sampling
- Recirculation
- Stateful Firewall using an internal connection tracker
- Multiple NICs, VTEPs and Bonding
- All OVS binaries work on Hyper-V
- Named pipes instead of Unix domain sockets
Stateful Firewall

- New Connection Tracker internal to OVS datapath
- Windows Filtering Platform
  - WFP applies the filters for creating the firewall
  - Can be used to audit firewall connections
  - Cannot track the status of the connection
  - Incompatible architecture for OVS (flow based)
- FreeBSD - Packet Filter
  - Didn’t require a full blown packet filter
  - conntrack-tcp parser is ported from FreeBSD
- Port of userspace connection tracker
- Supports a familiar netlink interface
Stateful Firewall

- ovs-dpctl dump-conntrack
- ovs-dpctl flush-conntrack

Netlink Socket (Emulation) - Netfilter like messages

OVS Flow table

Conntrack Module

recirculation

(matched flow)

(new/est/rel)

conntrack action

(-trk)

(new/est/rel)

(-trk)

Userspace (vswitchd) ovs-vsctl cmd
Stateful Firewall

• How does it work?
  – Maintains an expirable internal map to track connections
  – Counters for tracking packet count and size per connection
  – Netlink interface to Dump and Flush Conntrack entries
  – Supports subscribing to Conntrack delete/create events
  – Currently supports IPv4: TCP, UDP, ICMP packets
  – Adding in support for ALGs and fragmented packets
• Stateful Firewall on Openstack with KVM & Hyper-V
  – Install devstack
  – Setup Hyper-V as Compute Node
  – Install neutron-ovs-agent on Hyper-V
  – Create a VM on Hyper-V and KVM hypervisors
  – Setup security groups for both VMs
  – Display firewall policies for ICMP, TCP and UDP traffic
Demo – Stateful firewall with Openstack
Containers on Windows

- Prerequisites: Windows 10 / Server 2016 family

- Container types:
  - Windows Containers – runs as isolated process with shared kernel (does not require Hyper-V feature enabled)
  - Hyper-V Containers – runs in an optimized virtual machine

- Containers connect to a “VM Switch” for networking. NAT service has a port called “HNS Internal NIC”.
Containers on Windows

• Windows Container network adapter
  – Represented as a Hyper-V Virtual Switch internal port
  – Not visible on the host in the UI or ipconfig

• Hyper-V Container network adapter
  – Represented as a Hyper-V Virtual Switch synthetic port

• Requires integration with Docker (demo is part of PoC)
Roadmap

- Auto-addition of Hyper-V VIFs to OVSDB
- ALGs support for Connection Tracking
- IPv6 (tunnels, conntrack)
- Megaflows
- Containers
- Hardware Offloading for tunnels
- QoS
- Investigate OVN support
• Join our Hyper-V IRC meetings at 10 AM PST, Tuesdays at #openvswitch
Multiple Ports

- Multiple NICs
- Bonding
- Multiple VTEPs
- Internal ports (useful for testing)