Scaling the S in SDN at Azure

Albert Greenberg
Distinguished Engineer & Director of Engineering
Microsoft Azure Networking
ExpressRoute Partners

2013: SDN in the Cloud

2014: Host SDN and Express Route

2015: Controller and SmartNIC

Coming soon: Containers and SONiC
SDN in Windows Server 2016

- Bringing Cloud-scale Technology to Enterprise on Premises with Windows Server 2016

- Create Virtual Networks, Load Balancers, Security Groups, and more from Azure in Private Cloud

- Supports Future Protocols - OVSDB, VxLAN for open integration of control + data plane, using VFP, Azure’s Virtual Switch
What’s New In Data Plane?

Improving Scale
PacketDirect

New Windows direct I/O model for accelerating virtual networking and NFV

Port of VFP (Azure Virtual Switch) to PD:
- Up to 2x improvement in

Supported on major merchant silicon NICs

Initially for public cloud, coming to Windows Server 2016 soon!
HW is needed for scale, perf, and COGS at 40G+
• Multi-year ASIC cycle too slow
• Need agility and serviceability – SDN

Roll out Hardware as we do Software!
What’s New In Control Plane?

SDN for Containers
Containers Improve Density

10X utilization: 100s of VMs per host to 1000s of Containers per VM, each with its own network.
Goals for Container Management

**Speed**
- Instantly provision container network

**Scale**
- 100s of 1000s of containers in same virtual network

**Seamless DevOps**
- Common APIs for VMs and Containers
Open SDN Solution for Containers

Principles
- Consistent solution: Works across VMs and containers
- Cross platform: Works on Linux and

How: New primitives
- Network namespace isolation extended to VFP: tremendous scale without being limited by OS/port limitations
Open SDN for Containers in Public Cloud

NAT-based Container Network

Underlay Container Network

VM

Container 1
172.16.0.2
eth0
veth
172.16.0.1
bridge0
vswitch

Container 2
172.16.0.3
Eth0
veth

Bridge

VM policies

10.0.0.2
eth0
Veth

Container 1
10.0.0.2
eth0
veth

Container 2
10.0.0.3
eth0
veth

VM (Linux / Windows)

NAT

veth

namespace

VFP

namespace
Demo

On Prem

Windows VM (Media Client)

GW

Network Controller

Azure

Ubuntu VM

Container 1 (PHP server)

Ubuntu VM

Container 2 (PHP server)
What’s New In Physical Network?

Software for Open Networking in the Cloud (SONiC)
Goals

Velocity
• Feature agility weekly
• Bug mitigation daily

Serviceability
• Modular, fine-grained hitless upgrades and restarts
• `.deb/,.rpm/container preferred over`

Reliability
• Improvements with zero customer impact.
SAI

- OCP Contribution Dec 2014
- OCP Acceptance July 2015
- Common API and behavior model across multiple ASICs
- Hardware/software disaggregation.
- Independent evolution of hardware and software.

SONiC

- OCP Contribution March 2016
- Provides L2/L3 functionalities ready for cloud
- Loosely-coupled modular design
- Separation of states and logic

Logos:
- Arista
- Broadcom
- Barefoot Networks
- Cavium
- Centec Networks
- Dell
- Mellanox
- Metaswitch Networks
- Microsoft
SONiC Completes the Open-Source Switch Stack

Linux

Network Applications

Object Library

Switch State Service

ASIC SDK

SAI

User Space

Object Library

PAS

SDI

Kernel

PAL - sysfs

HW Peripheral drivers

netdev

Network device drivers

ASIC PCI driver

Various

OCP

Chassis Supplier

ASIC Supplier

Linux

Switch Hardware

Fan/Power/LED

Transceiver

ASIC
SONiC Demo: Servicing BGP

- Sender
- SONiC
- Quagga -> Gobgpd
- Commercial Switch
- Vendor BGP
- Receiver

BGP session
Unlocks Simpler, Best in Breed Technology

Enables Ecosystem
- Old Vs. New techniques
- Monitoring: Push Vs. Pull
- Identity: AD Vs. TACACS

Monitoring example
- SNMP pulling not very efficient
- Mapping counters to MIB back?
Want to be a part of Scaling the S in SDN? We’re Hiring!

albert@microsoft.com