DPDK on Microsoft Azure

Daniel Firestone
Madhan Sivakumar

DPDK Summit - San Jose – 2017
Microsoft Azure

The need for DPDK in the Cloud – NFV

Accelerated Networking in Azure

Enhancements to DPDK for Cloud usage
  - Host Serviceability and VM Migration support

Public Preview of DPDK on Azure
As customers migrate on-prem networks and apps to the cloud, NFV has become wildly popular.

Popular appliances in Azure Marketplace: Load Balancing, L7 filtering, web application firewalls, application gateways, DDoS protection, SD-WAN, and more.

DPDK can improve performance, throughput, latency, and reliability for this important class of workloads.

Goal: How can we make DPDK appliances work well in the public cloud?
Azure Accelerated Networking:
Fastest Cloud Network!

- Highest bandwidth VMs of any cloud
  - Up to 30Gbps for full-sized regular compute VMs
  - Standard Linux VM with CUBIC gets 30+Gbps on a single connection

- Consistent low latency network performance
  - Provides SR-IOV to the VM
  - Up to 10x latency improvement – as low as sub-10us within tenants
  - Increased packets per second (PPS)
  - Reduced jitter means more consistency in workloads

- Enables workloads requiring native performance to run in cloud VMs
  - >2x improvement for many DB and OLTP applications
Accelerated Networking Internals: SR-IOV

SDN/Networking policy applied in software in the host

Without accelerated networking:

- VM 1
- Virtual switch
- Physical server 1
- Network card
- Physical switch
- Azure infrastructure services
- Azure hosting infrastructure

With accelerated networking:

- VM 1
- Virtual switch
- Physical server 1
- Network card
- Physical switch
- FPGA-based SmartNIC acceleration used to apply all policies

SDN/Networking policy applied in software in the host

FPGA-based SmartNIC acceleration used to apply all policies
The Need for Serviceability

- Support migration of VMs
- Host updates including PF driver updates
- Completely transparent to the VM network stack
- Completely transparent to the DPDK applications
Accelerated Networking with serviceability & migration support

Diagram:
- User
- Kernel
  - TCP/IP
  - NetVSC
  - VF net dev
  - VMBUS to host partition
  - Virtual Function
- HW

#DPDKSummit
DPDK with serviceability & migration support

- DPDK application binds over failsafe PMD
- Transparent failsafe works over VF PMD and TAP PMD
- TAP PMD will create TC rules with action “redirect” to forward traffic from NetVSC to TAP net_dev and from TAP to NetVSC

#DPDKSummit
DPDK drivers

- TAP PMD
  - custom scan
  - probe
  - notification

- VF PMD
  - probe
  - notification

DPDK buses

- VDEV
  - scan
  - devargs
  - at init or hotplug

- PCI
  - scan
  - id
  - at init or hotplug

Linux drivers

- NetVSC netdev
- TAP netdev
- VF netdev

Transparent failsafe

#DPDKSummit
Launched production DPDK access to select registered developers at Ignite (September 2017) based on 17.08

Got lots of great feedback from the community

Demo: A10 vThunder appliance running over DPDK on Azure at 30Gbps line-rate
New Announcement - DPDK Public Preview!

- DPDK capable VMs are now available in Canada East!
- All Accelerated Networking VM sizes supported
  - 4+ core VMs
- Look for additional region availability and GA in the coming months!
- We want to work with the community to find ways to make servicing and migration completely transparent to DPDK applications
- Developers - Please try DPDK on Azure and give us feedback!

For any questions regarding the preview, please email AzureDPDK@Microsoft.com
Questions?