CLOUD ARCHITECTURE & PERFORMANCE WORKLOADS

Field Activities
Jeremy Eder
Principal Performance Engineer
Red Hat, Inc
@jeremyeder

Matt Smith
Senior Solution Architect
Red Hat, Inc
@rhmjs
CLOUD ARCHITECTURE
“Cloud” - A Buzzword Review

- On-demand self-service
  - Accelerate Time-to-Value
  - Multi-tenancy

- Broad network access
  - Heterogeneous, Standardized Interfaces

- Resource pooling
  - Shared Compute, Network, Storage

- Rapid elasticity
  - Grow, Shrink, Repurpose

- Measured service
  - Metering and Reporting

NIST SP 800-145: “The NIST Definition of Cloud Computing”

#redhat #rhsummit
Red Hat Cloud Infrastructure
CLOUD ADOPTION
Common Cloud Workloads

- Web Applications
  - Horizontal scalability/resiliency
  - Rebuild, don’t repair

- Development Environments
  - On-demand self-service
  - Templates matching production
  - Automated testing
Growing Cloud Workloads

• Distributed / Grid Compute
  ○ Resource Reallocation

• Big Data / Data Analytics
  ○ Horizontal Scalability/Resiliency
  ○ On-demand self-service “DAaaS”
  ○ Resource Reallocation
Emerging Cloud Workloads

- Telco
  - Network Functions Virtualization (NFV)
  - Packet Processing/Analysis
- Retail / Industrial
  - “Internet of Things” (IoT)
- Financial Services Industry
  - Regulatory Compliance
  - Transactional, Low-Latency
Performance Requirements

- High Compute Throughput
- High Network Throughput
- Low Network Latency
Performance & Cloud ?
TECHNICAL APPROACH
Typical Virtualization Stack

- APPLICATION
- IP STACK
- VIRTUAL NETWORK DRIVER
- HYPERVERSOR
- IP STACK / BRIDGE / ENCAPSULATION
- NETWORK DRIVER
- SERVER HARDWARE
- NETWORK CARD
Typical Virtualization Stack

APPLICATION
IP STACK
VIRTUAL NETWORK DRIVER
HYPervisor
IP STACK / BRIDGE / ENCAPSULATION
NETWORK DRIVER
SERVER HARDWARE
NETWORK CARD

ABSTRACTION !!
Typical KVM Virtualization Stack

Red Hat Enterprise Linux Guest
- Application
- IP Stack
- Virtio
- KVM
- IP Stack / Bridge / Encapsulation
- Network Driver
- Server Hardware
- Network Card

Red Hat Enterprise Linux Host
Red Hat Enterprise Linux

# tuned-adm profile virtual-host

# tuned-adm profile virtual-guest
KVM: Industry Leading Results

SPEC virt_sc® 2013 Benchmark Leadership
Highest Red Hat based result versus highest non Red Hat result by socket
(As of June 18, 2015)

http://www.spec.org/virt_sc2013/results/

(June 2015) SPECvirt_sc2013
OpenStack & NUMA

- Non-Uniform Memory Access
- The NUMA topology defines the locality of CPU and Memory
- Processes can be bound to NUMA nodes for optimal performance
OpenStack & NUMA

- KVM awareness of NUMA aligns the guests virtual NUMA topology with the physical topology of the host

- RHEL-OSP 6 supports specifying NUMA requirements in the “flavor”
OpenStack NUMA Configuration

Instance with 8 vCPUs and 4GB RAM

```
 hw:numa_nodes=2
 hw:numa_cpus.0=0,1,2,3,4,5
 hw:numa_cpus.1=6,7
 hw:numa_mem.0=3
 hw:numa_mem.1=1
```
NUMA: Process == KVM thread

No NUMA scheduling

Node 0
Process 37
Process 29
Process 19
Process 61

Node 1

Node 2

Node 3

Process 61

With NUMA Scheduling

Node 0
Proc 29
Proc 19
Proc 61

Node 1

Node 2
Proc 37

Node 3
Simple OpenStack Deployment

- APPLICATION
- IP STACK
- VIRTIO
- KVM
- IP STACK / BRIDGE / ENCAPSULATION
- NETWORK DRIVER
- SERVER HARDWARE
- NETWORK CARD

EXTERNAL

- OPENVSWITCH
- IP STACK
- NETWORK DRIVER
- SERVER HARDWARE
- NETWORK CARD

VXLAN TENANT

EXTERNAL
Typical OpenStack Deployment

- **APPLICATION**
  - IP STACK
  - VIRTIO
  - KVM
  - IP STACK / BRIDGE / ENCAPSULATION
  - NETWORK DRIVER
  - SERVER HARDWARE
  - NETWORK CARD

- **APPLICATION**
  - IP STACK
  - VIRTIO
  - KVM
  - IP STACK / BRIDGE / ENCAPSULATION
  - NETWORK DRIVER
  - SERVER HARDWARE
  - NETWORK CARD

- **OPENVSWITCH**
  - IP STACK
  - NETWORK DRIVER
  - SERVER HARDWARE
  - NETWORK CARD

- VXLAN TENANT
- EXTERNAL
OpenStack Provider Network

**IP STACK**
- NETWORK DRIVER
- SERVER HARDWARE
- NETWORK CARD

**APPLICATION**
- IP STACK
- VIRTIO
- KVM
- IP STACK / BRIDGE / ENCAPSULATION
- NETWORK DRIVER
- SERVER HARDWARE
- NETWORK CARD

**OPENVSWITCH**
- IP STACK
- NETWORK DRIVER
- SERVER HARDWARE
- NETWORK CARD

**EXTERNAL**

**VXLAN TENANT**

**FLAT/VLAN PROVIDER**
Single Root I/O Virtualization (SR-IOV)

- Physical Functions (PFs): full PCIe devices that include the SR-IOV capabilities

- Virtual Functions (VFs): simple PCIe functions, derived from PFs, that only process I/O

- KVM Passthrough of VFs to Guests
OpenStack + SR-IOV

- Ensure that Intel VT-d or AMD IOMMU are enabled in the BIOS and Operating System
- Load the driver with VF configuration
- Enable the `sriovnicsswitch` driver in Neutron
- Enable SR-IOV on the Controller and Compute nodes
- Create an SR-IOV port on a Neutron network
- Launch an instance configured to use the port

Full Detail in Red Hat Customer Portal at https://access.redhat.com
OpenStack with SR-IOV

VXLAN TENANT

FLAT/VLAN PROVIDER

EXTERNAL
Data-plane Development Kit (DPDK)

- Primary development by Intel and 6WIND
- Set of libraries and drivers for fast packet processing
- OpenVSwitch with DPDK for accelerated packet processing with a focus on NFV use-cases, under active development
OpenStack with OVS+DPDK
DPDK config, bare-metal, container, kvm

bare-metal

PF assigned to container

PF assigned to VM
Network Function Virtualization (NFV) Throughput and Packets/sec (RHEL7.x+DPDK)

NFV: Millions of Packets Per Second

RHEL 7.x, L2 Forwarding, 12 x 40Gb NICs

<table>
<thead>
<tr>
<th></th>
<th>Packets Per Second (Millions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KVM</td>
<td>208</td>
</tr>
<tr>
<td>Docker</td>
<td>215</td>
</tr>
<tr>
<td>Bare-metal</td>
<td>218</td>
</tr>
<tr>
<td>HW Maximum</td>
<td>225</td>
</tr>
</tbody>
</table>

208Mpps+ INTO KVM DPDK
Technology Review

- Red Hat Enterprise Linux
- Kernel-based Virtual Machine (KVM)
- Non-Uniform Memory Access (NUMA)
- OpenStack Flat/VXLAN Network Topologies
- Single Root I/O Virtualization (SR-IOV)
- OpenVSwitch (OVS) + Data-plane Development Kit (DPDK)
Other Cloud/Performance Sessions

- **Performance analysis & tuning of Red Hat Enterprise Linux**
  - Wednesday, June 24 1:20 pm - 3:20 pm

- **Performance of OpenStack Cinder on Ceph**
  - Thursday, June 25 4:50 pm - 5:50 pm

- **Containers versus virtualization**
  - Friday 9:45 am - 10:45 am

- **Open source & network functions virtualization**
  - Friday 11:00 am - 12:00 pm
Matt Smith
@rhmjs
verticalindustriesblog.redhat.com

Jeremy Eder
@jeremyeder
developerblog.redhat.com
www.breakage.org
LEARN. NETWORK.
EXPERIENCE OPEN SOURCE.