Containers versus Virtualization

Jeremy Eder, Principal Performance Engineer
Scott Herold, Principal Product Manager, RHEV
June 2015
Agenda

- Venting
- Tech Overview
- Workload Classification
- Cold War?
- Performance Data Roundup
TOP 5 MISCONCEPTIONS ABOUT CONTAINERS

1. Containers are new.
2. Containers equal virtualization.
3. Containers are universally portable.
4. Containers are secure by default.
5. Containers are not enterprise-ready.
VIRTUALIZATION AND CONTAINERS

VIRTUALIZATION

CONTAINERIZATION

HARDWARE

HOST OS

HYPERVERSOR

GUEST OS

BINS/ LIBS

APP A

APP A'

APP B

GUEST OS

BINS/ LIBS

GUEST OS

BINS/ LIBS

APP A

APP A'

APP B

APP B'

HOST OS, SHARED SERVICES

BINS/LIBS

BINS/LIBS

HARDWARE
RED HAT ENTERPRISE VIRTUALIZATION

- Centralized Management of KVM Hypervisor
- Self-Service User Portal
- VM Workload Management
- Differentiating Features
WHAT ABOUT DENSITY?

“For every VM, you can run 10 billion containers.” -- Internet
How many containers will you run on one OS instance?

- 1
- 10
- 50
- 100
- > 100
WHAT ABOUT DENSITY?

“This may be the most misleading stat ever.” -- Me
Get off my lawn!
Because it’s ALL about the workloads

• Some don’t care where they run
  – Batch workloads
• Some care greatly
  – Security/Isolation
  – Uptime
  – Performance
What is a workload? Subsystems

- CPU
- Memory
- Storage
- Network
What is a workload? Requirements

- Security
- Stability
- Performance
- Support

Fruits
Grains
Vegetables
Protein
Dairy
Culture, Control

Code Down (Dev) versus Infra Up (Ops)

I WANT CHANGE

I WANT STABILITY
WHEN WILL YOU MAKE SOMETHING THAT MATTERS?

WHEN WILL YOU MAKE SOMETHING COOL?
CONTAINERS VERSUS VIRTUALIZATION:

NEW COLD WAR ?
Minimum Overheads

Sample Virtualization Overhead

Sample Container Overhead
Maximum Overheads

Sample Virtualization Overhead

- Overhead: 15
- Capacity: 85

Sample Container Overhead

- Overhead: 3
- Capacity: 97
# Reducing Overhead in VMs

<table>
<thead>
<tr>
<th>Workload</th>
<th>Mitigation</th>
</tr>
</thead>
</table>
| CPU-intensive          | • CPU Pinning  
                         • Avoid syscalls  
                         • Setup NUMA topology in-Guest                                          |
| Memory-heavy           | • Use hugepages  
                         • NUMA Pinning  
                         • Setup Hugepages in-Guest                                              |
| Network (Latency)      | • SR-IOV  
                         • PCI Passthrough  
                         • Busy Poll                                                          |
| Network (Throughput)   | • Not normally an issue                                                   |
| Storage (Latency)      | • Increase threads  
                         • virtio-blk-dataplane coming soon                                     |
| Storage (Throughput)   | • Not normally an issue                                                   |
CONTAINERS VERSUS VIRTUALIZATION:

PERFORMANCE DATA ROUND-UP
CPU Tests: MPI LINPACK

MPI Linpack % diff vs Bare Metal
RHEL7.1, KVM, Docker

<table>
<thead>
<tr>
<th></th>
<th>% diff vs Bare Metal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare Metal</td>
<td>100</td>
</tr>
<tr>
<td>4 Containers</td>
<td>100</td>
</tr>
<tr>
<td>KVM, SR-IOV</td>
<td>90</td>
</tr>
<tr>
<td>KVM, vhost_net</td>
<td>80</td>
</tr>
</tbody>
</table>
I/O Tests: fio ... Bare Metal, KVM, Atomic, Docker

fio IOPS % diff vs baseline (geomean)

<table>
<thead>
<tr>
<th></th>
<th>%diff vs baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare Metal</td>
<td>100</td>
</tr>
<tr>
<td>Bare Metal + Container</td>
<td>100</td>
</tr>
<tr>
<td>Bare Metal + Atomic</td>
<td>100</td>
</tr>
<tr>
<td>KVM</td>
<td>100</td>
</tr>
<tr>
<td>KVM + Container</td>
<td>100</td>
</tr>
<tr>
<td>KVM + Atomic</td>
<td>100</td>
</tr>
</tbody>
</table>
Application Tests: Business Analytics

Business Analytics
Lower is Better

Time to Complete (seconds)

- Bare Metal
- RHEL6 Container
- RHEL7 Container
- RHEL6 KVM
- RHEL7 KVM
Network Latency and Throughput

netperf Latency and Throughput
Higher is Better

% diff vs Bare Metal

TCP_RR  UDP_RR  TCP_STREAM  UDP_STREAM

Bare Metal  Container  KVM+SR-IOV  KVM
Large OLTP Database, BM vs Container vs KVM

Large OLTP Database (3 instances of 100 Users) Higher is Better

- Bare Metal
- Containers
- KVM

%diff vs baseline
Network Function Virtualization (NFV) Throughput and Packets/sec (RHEL7.x+DPDK)

NFV: Millions of Packets Per Second

RHEL7.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>
RHEL7.1 + Solarflare OpenOnload Bare Metal / KVM / Containers

- Lower is better
- Alternative kernel-bypass mechanism to DPDK
## Workload Classification

<table>
<thead>
<tr>
<th>Workload</th>
<th>Bare Metal</th>
<th>Containers</th>
<th>KVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU-bound</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Memory Intensive</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Disk Latency</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Disk Throughput</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Network Latency</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Network Throughput</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Security</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
<tr>
<td>Uptime (Live Migration)</td>
<td>⚫</td>
<td>⚫</td>
<td></td>
</tr>
<tr>
<td>Deployment Speed</td>
<td>⚫</td>
<td>⚫</td>
<td>⚫</td>
</tr>
<tr>
<td>Alternative OS</td>
<td></td>
<td></td>
<td>⚫</td>
</tr>
</tbody>
</table>
LEARN. NETWORK.
EXPERIENCE OPEN SOURCE.