redhat.

# WHEN CONTAINERS AND VIRTUALIZATION DO - AND DON'T - WORK TOGETHER

JEREMY EDER

redhat.

# Agenda

- Technology Trends
- Container and VM technical Overview
- Performance Data Round-up
- Workload Classification

# Why listen to me...

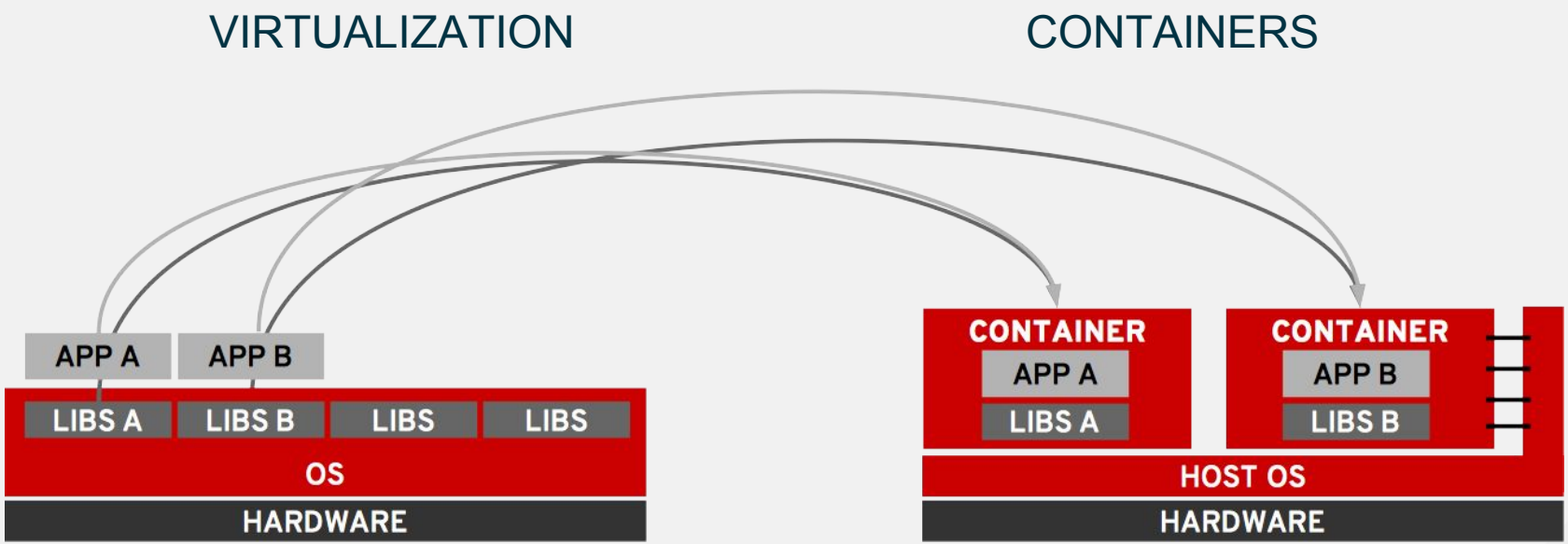
- Co-team lead for container performance and scale team at Red Hat.

- Architect of Red Hat “tuned” project.
- Authored many blogs and whitepapers on container performance, tuning for high frequency trading.

But really, don’t listen to me:

# Listen to your apps.

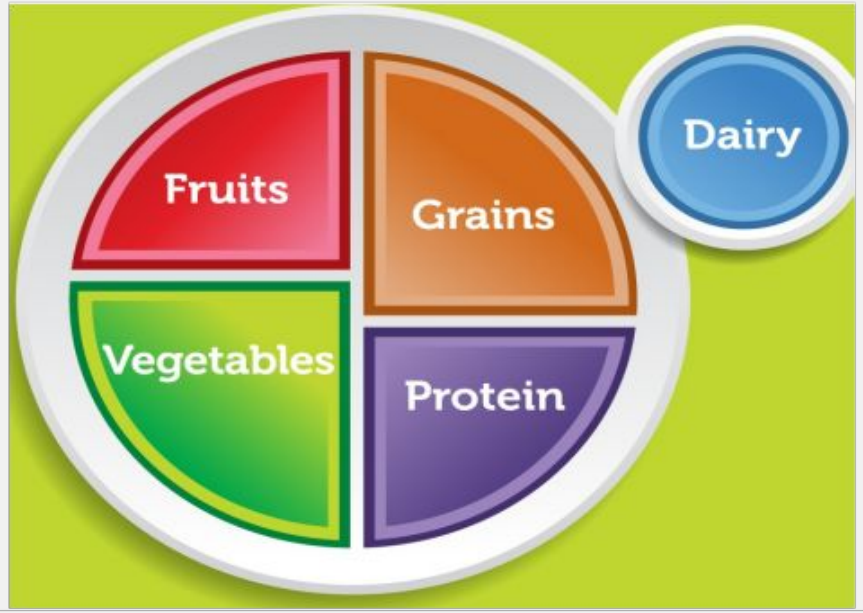
# Containers are an OS Technology

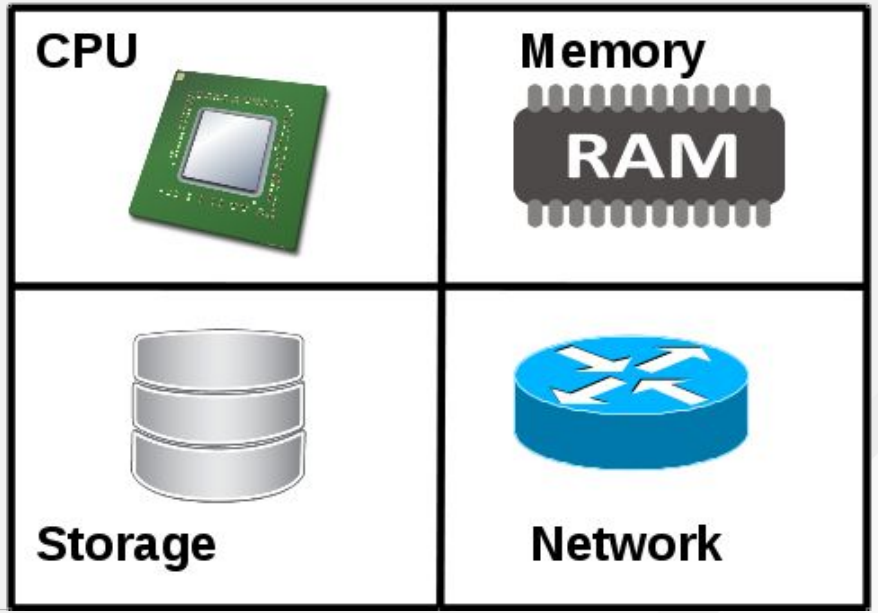


# It’s all about the workloads...

- Some don’t care where they run
  - Batch workloads
- Some care greatly
  - Security, Isolation
  - Uptime
  - Performance
  - Proximity/Locality to data

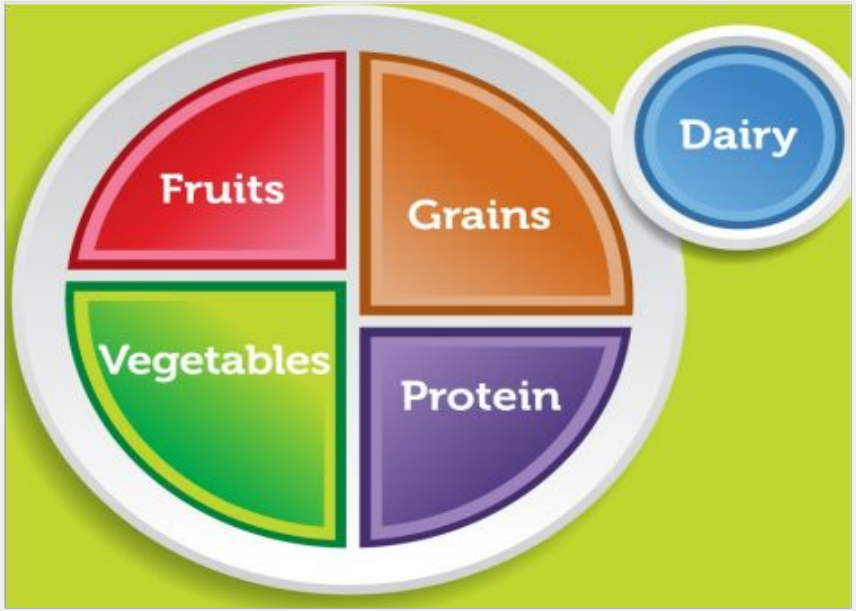
# What is a workload? Subsystems

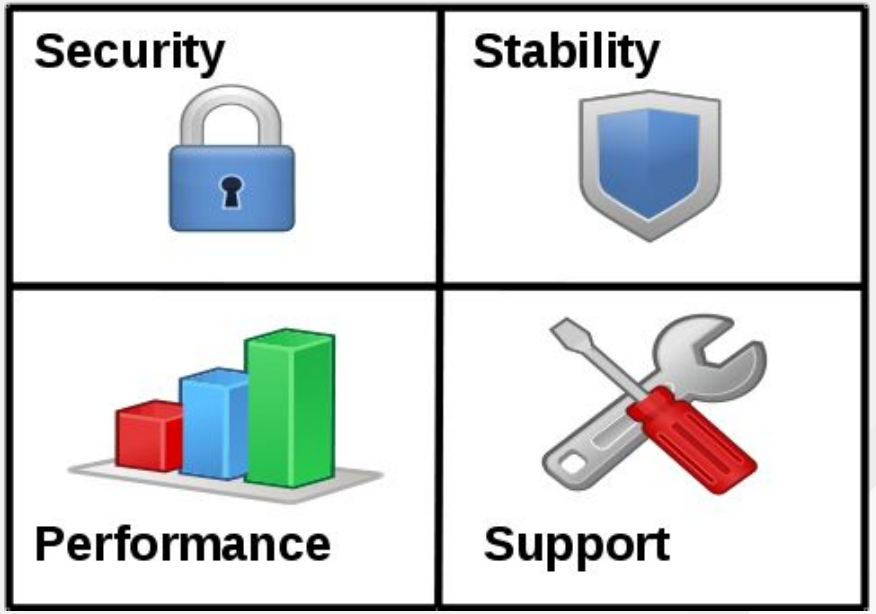


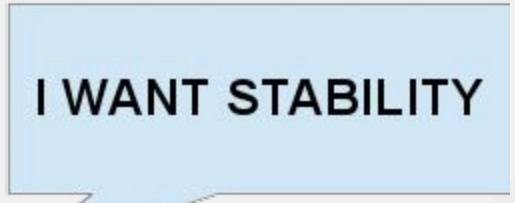
# What is a workload? Business Requirements

.



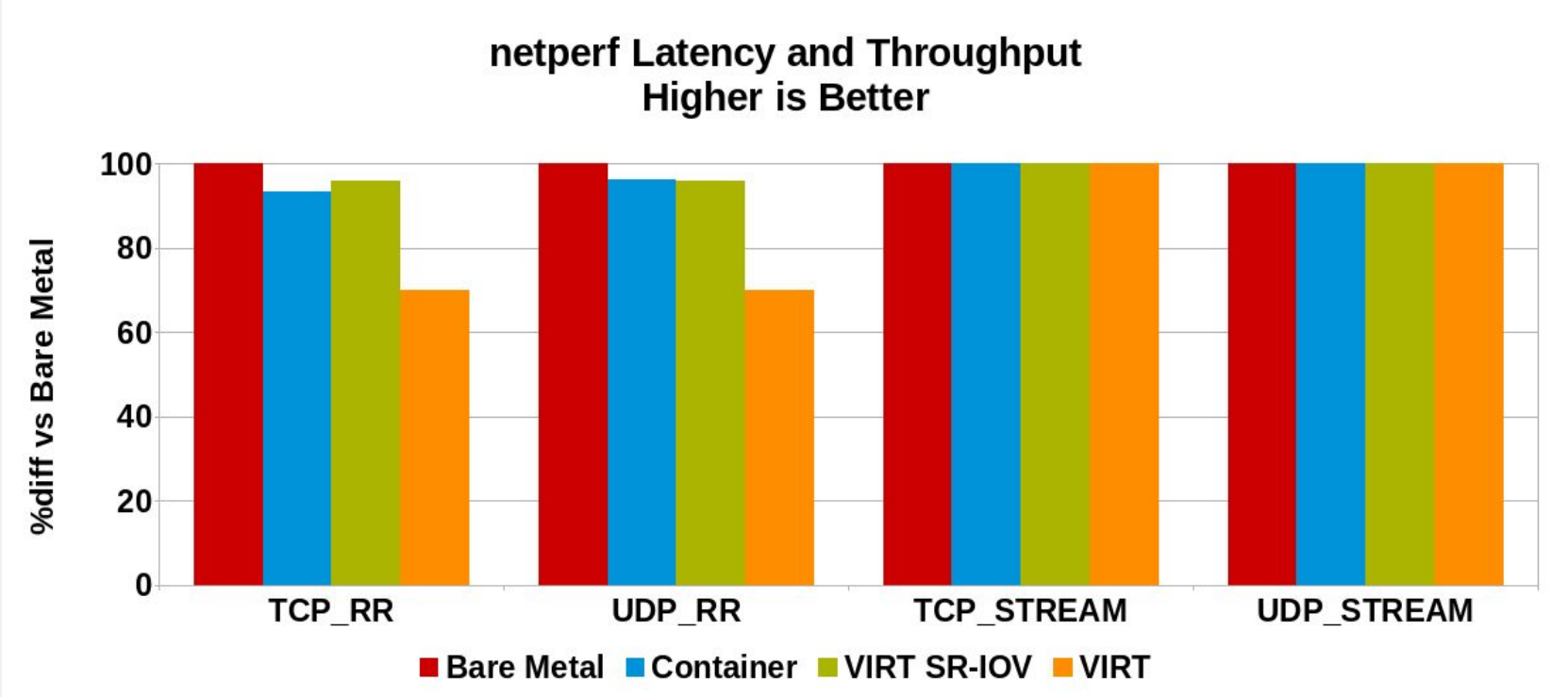




# Code Down (Dev) versus Infra Up (Ops)

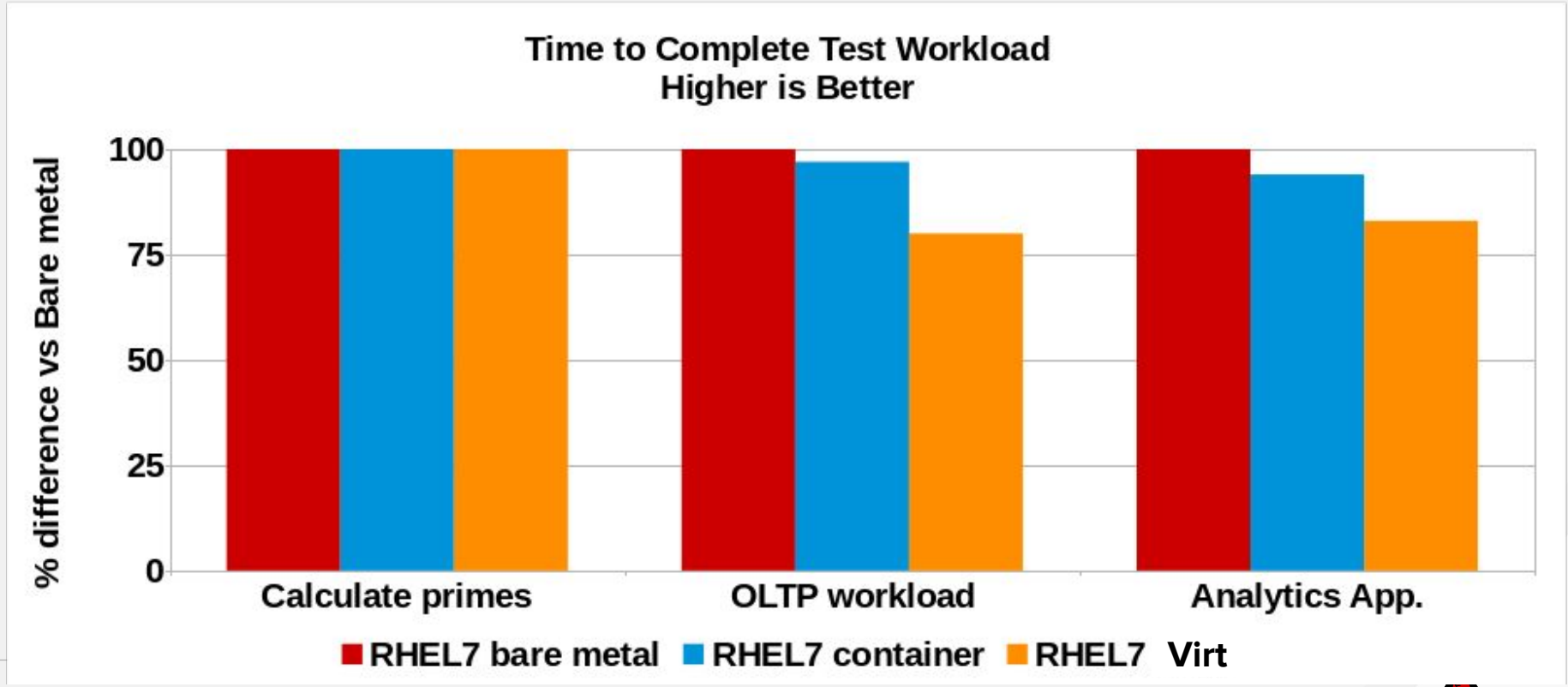


# CONTAINERS AND VIRTUALIZATION: PERFORMANCE DATA ROUND-UP

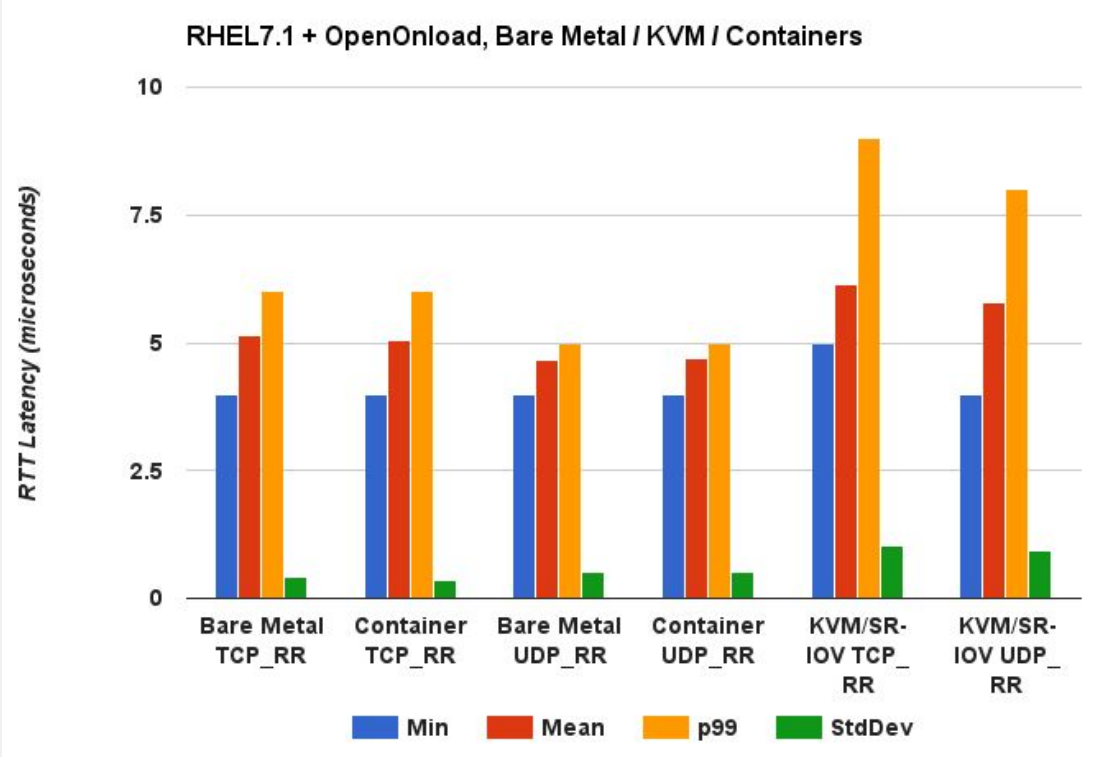
# Network Latency and Throughput



# Performance of Large “Expensive” Apps :-)

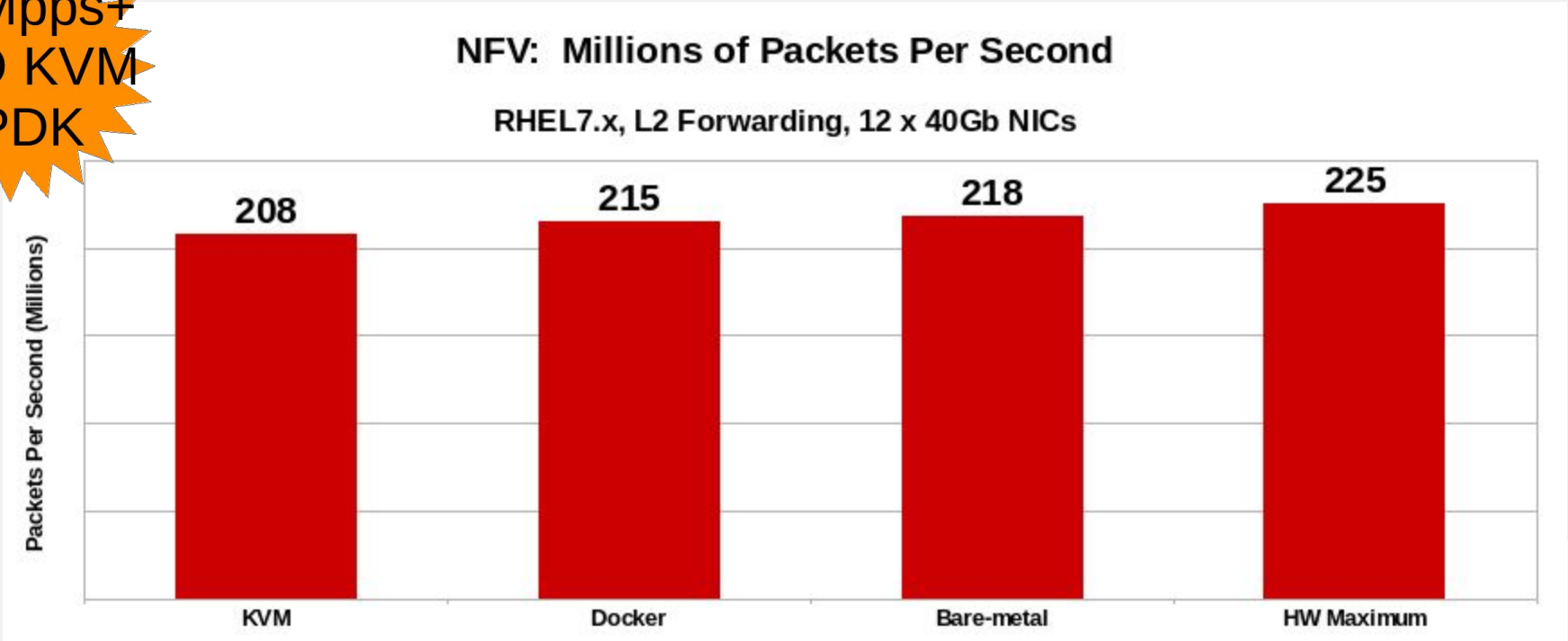


# RHEL7 + Containerized Solarflare OpenOnload



# Network Function Virtualization (NFV) Throughput and Packets/sec (RHEL7.x+DPDK)

208Mpps+ INTO KVM DPDK



# Speedups for Virtual Machines

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

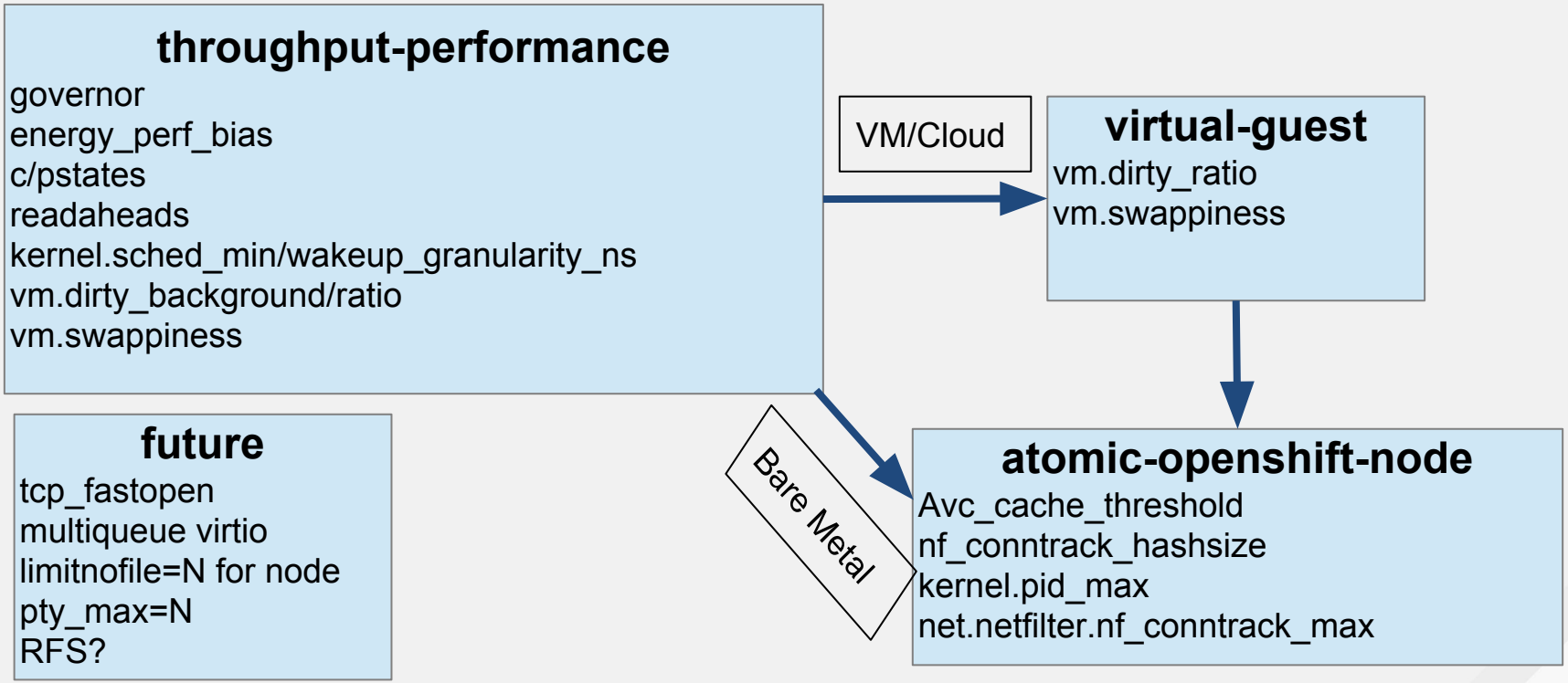
# What is “tuned”?

## Tuning profile delivery mechanism

# Tuned Profiles throughout Red Hat Products

| | |
|---|---|
| RHEL7 Desktop/Workstation<br>**balanced** | RHEL7 Server/HPC<br>**throughput-performance** |
| RHEL6/7 KVM Host, Guest<br>**Virtual-host/guest** | RHEV<br>**virtual-host** |
| Red Hat Storage<br>**rhs-high-throughput, virt** | RHEL OSP (compute node)<br>**virtual-host** |
| RHEL Atomic<br>**atomic-host, atomic-guest** | OpenShift<br>**openshift-master,node** |

# Tuned Profiles



# KVM vs Container Performance (HP results)

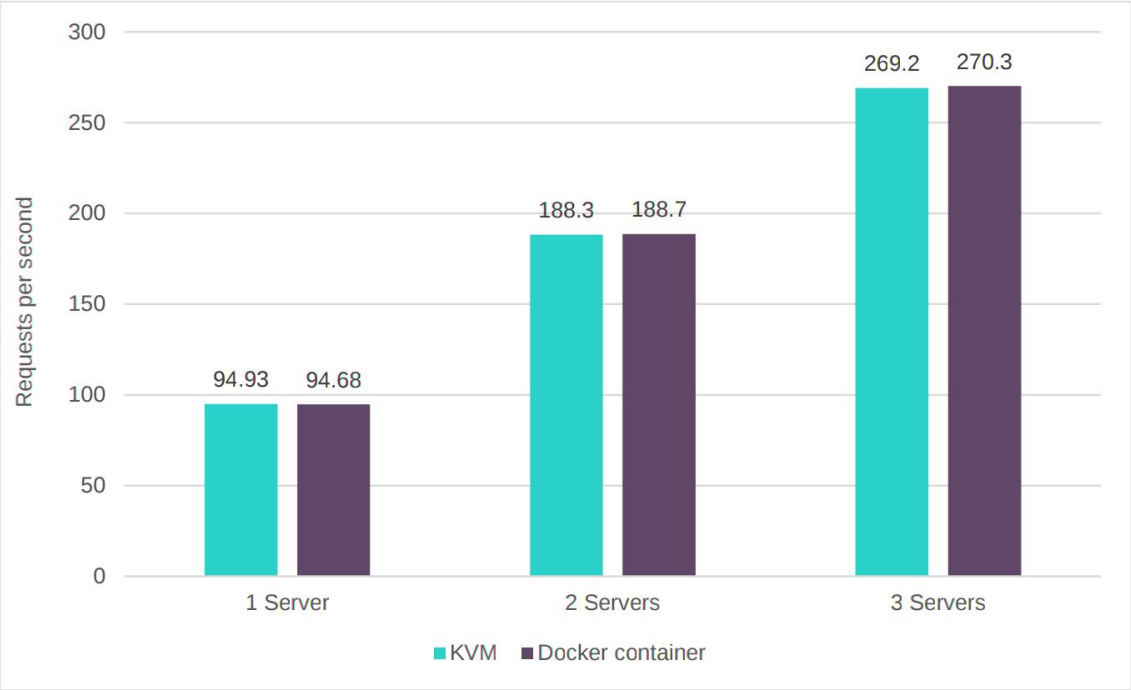



**Figure 10.** Comparison of Docker containers and KVMs (ProLiant DL580 Gen8 with RHEL 7.1)

## Distributed Environment

- Java application server
- Internet Message Access Protocol (IMAP) server
- Batch server

http://h20195.www2.hpe.com/V2/getpdf.aspx/4AA6-2761ENW.pdf

# Workload → Infrastructure Mapping

| Color | Meaning |
|---|---|
| | Mature<br>No Perf Concerns |
| | Immature<br>Limited Perf Concerns |
| | Difficult/Impossible<br>(currently) |

| Workload | Bare Metal | Containers | Virt |
|---|---|---|---|
| CPU Intensive | | | |
| Memory Intensive | | | |
| Disk I/O Latency | | | |
| Disk I/O Throughput | | | |
| Network Latency | | | |
| Network Throughput | | | |
| Security | | | |
| Uptime (Live Migration) | | | |
| Deployment Speed | | | |
| Alternative OS | | | |

# Workload ➔ Infrastructure Mapping: Build Farm

| Icon | Meaning |
|---|---|
|  | Mature and/or No Perf Concerns |
|  | Immature and/or Limited Perf Concerns |
|  | Mixed Concerns |
|  | Not Applicable |

| Attribute | Build Farm |
|---|---|
| CPU Intensive | High |
| Memory Intensive | High |
| Disk I/O Latency | Low |
| Disk I/O Throughput | High |
| Network Latency | Low |
| Network Throughput | High |
| Security | Low |
| Uptime (Live Migration) | N/A |
| Deployment Speed | High |
| Alternative OS | N/A |

# Workload ➔ Infrastructure Mapping: memcached

| Icon | Meaning |
|---|---|
|  | Mature and/or No Perf Concerns |
|  | Immature and/or Limited Perf Concerns |
|  | Mixed Concerns |
|  | Not Applicable |

| Attribute | memcached |
|---|---|
| CPU Intensive | Medium |
| Memory Intensive | Medium |
| Disk I/O Latency | Low |
| Disk I/O Throughput | Low |
| Network Latency | High |
| Network Throughput | High |
| Security | N/A |
| Uptime (Live Migration) | N/A |
| Deployment Speed | Low |
| Alternative OS | N/A |

# Workload ➔ Infrastructure Mapping: Stock Trading

| Icon | Meaning |
|---|---|
|  | Mature and/or No Perf Concerns |
|  | Immature and/or Limited Perf Concerns |
|  | Mixed Concerns |
|  | Not Applicable |

| Attribute | Stock Trading |
|---|---|
| CPU Intensive | High |
| Memory Intensive | High |
| Disk I/O Latency | Low |
| Disk I/O Throughput | Low |
| Network Latency | High |
| Network Throughput | High |
| Security | Low |
| Uptime (Live Migration) | N/A |
| Deployment Speed | N/A |
| Alternative OS | N/A |

# Workload ➔ Infrastructure Mapping: Gluster

| Icon | Meaning |
|---|---|
|  | Mature and/or No Perf Concerns |
|  | Immature and/or Limited Perf Concerns |
|  | Mixed Concerns |
|  | Not Applicable |

| Attribute | Gluster |
|---|---|
| CPU Intensive | Low |
| Memory Intensive | Low |
| Disk I/O Latency | High |
| Disk I/O Throughput | High |
| Network Latency | High |
| Network Throughput | High |
| Security | N/A |
| Uptime (Live Migration) | N/A |
| Deployment Speed | Low |
| Alternative OS | N/A |

# Workload ➔ Infrastructure Mapping: Animation Render Farm

| Icon | Meaning |
|---|---|
|  | Mature and/or<br>No Perf Concerns |
|  | Immature and/or<br>Limited Perf Concerns |
|  | Mixed Concerns |
|  | Not Applicable |

| Attribute | Animation |
|---|---|
| CPU Intensive | High |
| Memory Intensive | Medium |
| Disk I/O Latency | Medium |
| Disk I/O Throughput | High |
| Network Latency | Medium |
| Network Throughput | High |
| Security | Low |
| Uptime (Live Migration) | Low |
| Deployment Speed | High |
| Alternative OS | N/A |

# It’s all about the workloads.

# FUTURE OPPORTUNITIES

- Red Hat Performance Engineering Blogs

  https://developers.redhat.com/blog/tag/performance/
- Red Hat Performance Tuning Guide
  https://access.redhat.com/documentation/en-US/Red_Hat_Enterprise_Linux/7/html-single/Performance_Tuning_Guide/
- Learn more about Red Hat Summit at redhat.com/summit

redhat.

# THANK YOU