Xen Containers: Better way to run Docker Containers

Sainath Grandhi
sainath.grandhi@intel.com

Contributions: Jun Nakajima

Motivation

- "Containers" being adopted for application development/deploying
- Containers looked upon as lightweight alternative for traditional VMs
- VMs offer stronger application isolation
- Benefits of VMs can be reaped if they are made lightweight and run like containers

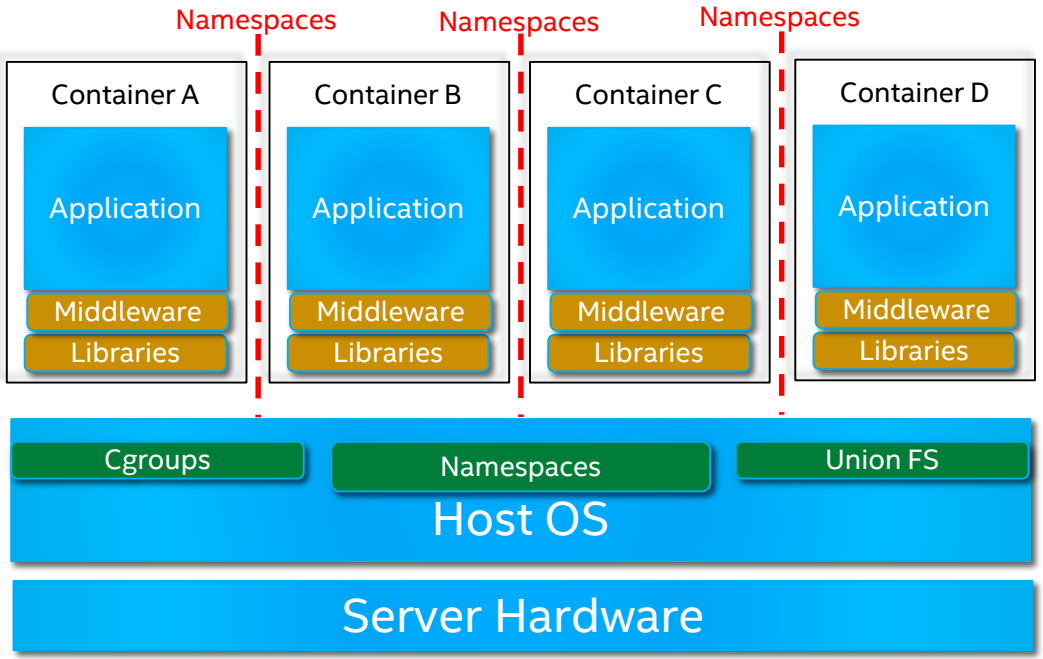
Agenda

Containers

Xen Containers

Numbers

Next Steps


Namespaces
Namespaces
Namespaces
Container A
Container B
Container C
Container D
Application
Application
Application
Application
Middleware
Middleware
Middleware
Middleware
Libraries
Libraries
Libraries
Libraries
Cgroups
Namespaces
Union FS
Host OS
Server Hardware

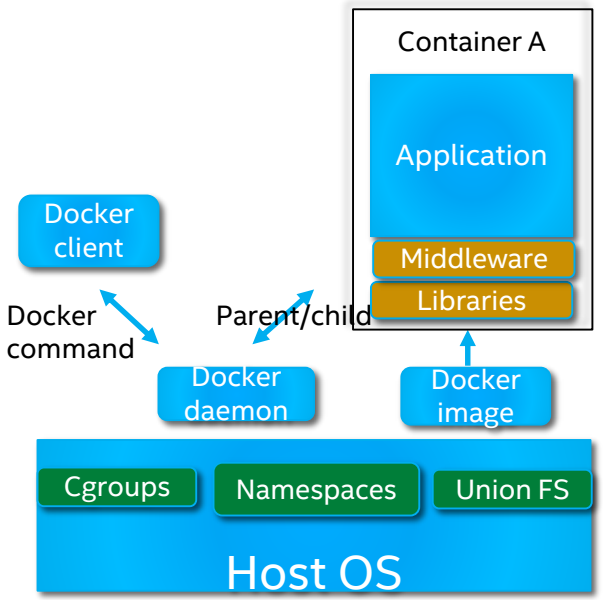
Docker Containers

Docker – a one stop solution for running, building and packaging containers

Running
- docker run/create/stop

Building
- docker build

Packaging
- docker push/pull/commit



Bare metal containers - Isolation

Isolation provided by Host OS

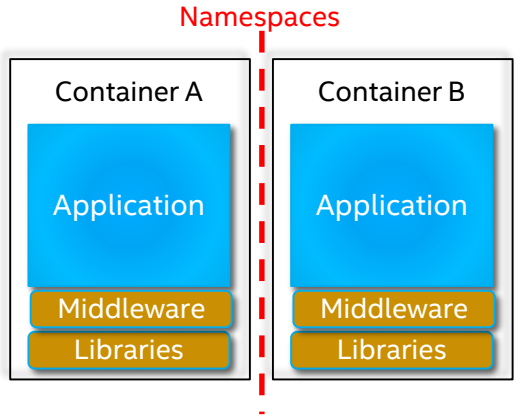
- Security compromised kernel can be exploited by malicious images/applications for namespace intrusion
- Enabling cgroups and namespaces in the kernel increases the kernel attack surface

Malicious public images

- Over 30% of Official Images in Docker Hub Contain High Priority Security Vulnerabilities
  http://www.banyanops.com/blog/analyzing-docker-hub/

Multi-tenant Cloud Providers

- Google: “we see the VM as the only truly safe isolation.... Until we see foolproof security for containers, we will always double-bag our customers' workloads”
  http://www.informationweek.com/cloud/infrastructure-as-a-service/google-docker-does-containers-right/d/d-id/1319146



Agenda

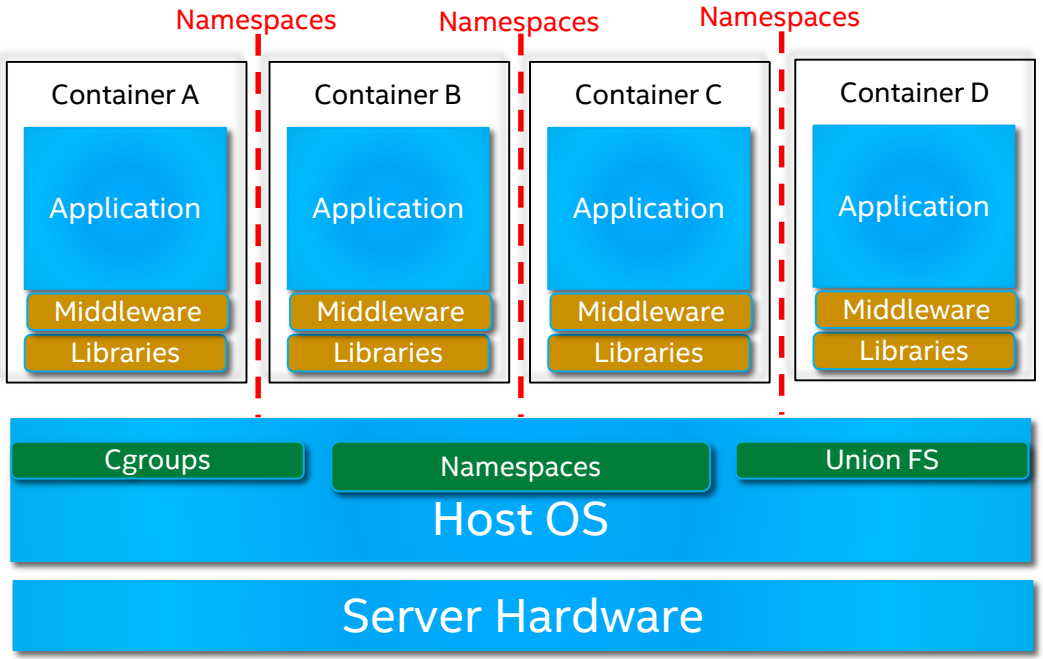
Containers

Xen Containers

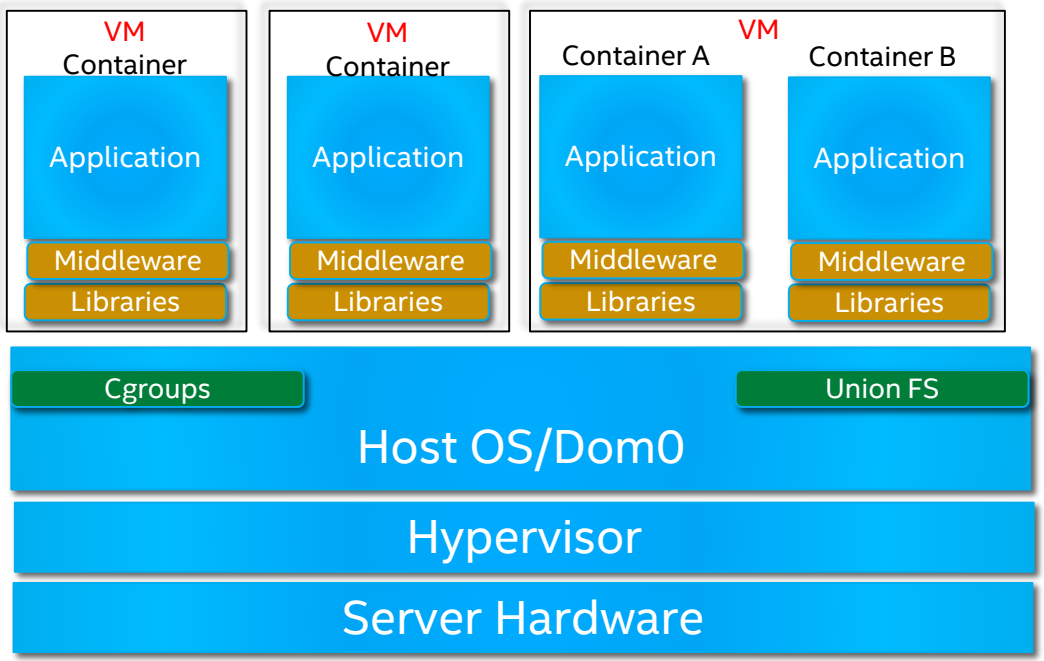
Numbers

Next Steps

Containers


Namespaces
Namespaces
Namespaces
Container A
Container B
Container C
Container D
Application
Application
Application
Application
Middleware
Middleware
Middleware
Middleware
Libraries
Libraries
Libraries
Libraries
Cgroups
Namespaces
Union FS
Host OS
Server Hardware

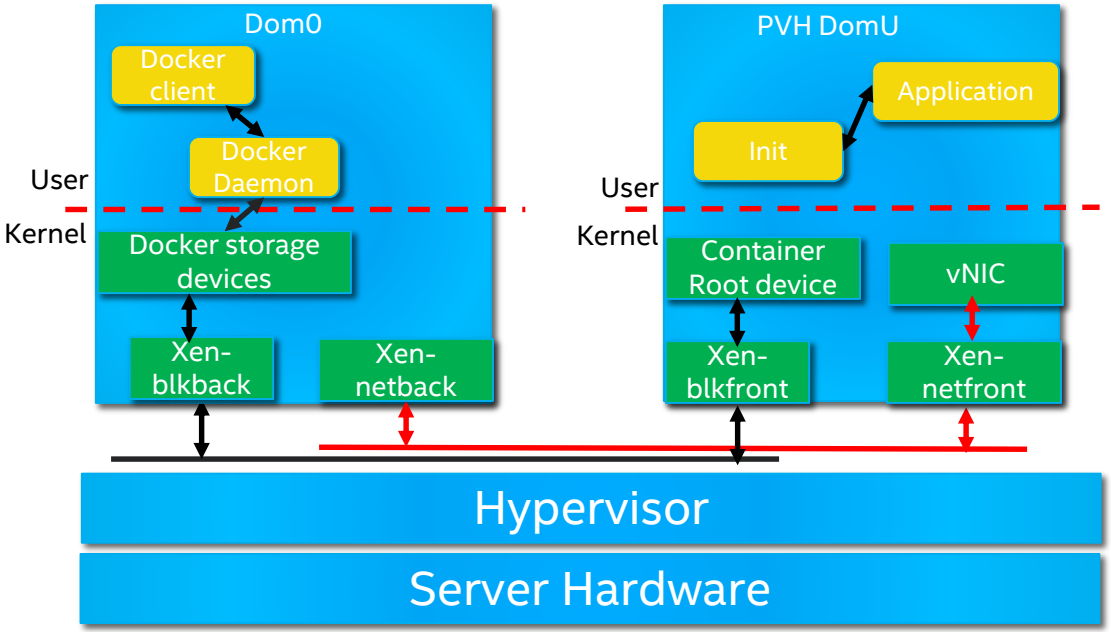
VM Containers



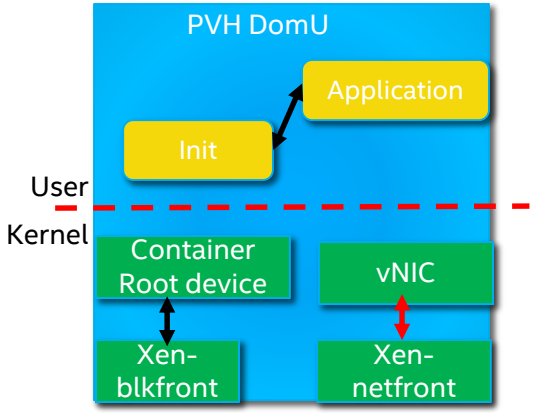
Xen PVH Containers

- VM containers good for multi-tenant cloud providers
  - Group containers from a tenant onto a VM
- Great infrastructure in place for guest isolation
- PVH for app containers
  - Boot to guest kernel in protected mode
  - PV performance for disk and network
  - Hardware virtualized performance for CPU and memory
- Why PVH (vs. HVM)
  - No dependence on QEMU
  - No BIOS
  - Faster Boot time

Xen Containers with Docker



Xen Containers with Docker – Guest Anatomy



Minimal Kernel

- Minimally configured kernel

Init

- Init service to mount application rootfs and configure network

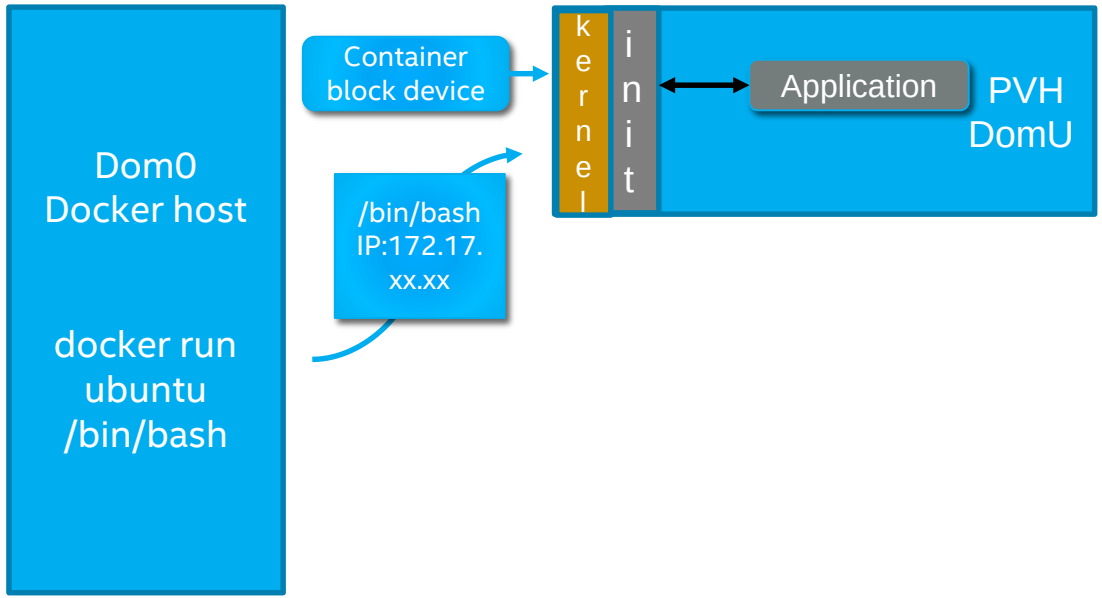
Storage

- Docker container volume as rootfs

Networking

- Docker subnet IP and docker bridge gateway

Xen Containers with Docker – Guest Configuration



Storage

- ❑ Docker devicemapper storage backend – container volume

Application path

- ❑ Application path from docker run/exec command

Network IP

- ❑ DHCP/docker subnet for interoperability with docker containers

Agenda

Containers

Xen Containers

Numbers

Next Steps

Numbers

	PVH	HVM	Comments
Domain Creation	224	184	Time spent by xl toolstack to setup domain
To drop into container shell	1380	**2503**	Time taken to boot the minimal kernel and drop into shell from container rootfs

Guest Memory Used – 16MB

Config:

Host
Xeon® CPU E5-2699 v3
Memory – 60GB
Dom0 Memory – 4GB
Dom0 vCPUs – 8

Guest
Memory – 128MB
vCPU - 1

Agenda

Containers

Xen Containers

Numbers

Next Steps

Next Steps

Docker Volumes

- PV VirtFS for supporting docker volumes

Pods (Multiple applications in a VM)

- Leverage systemd as the init service inside VM to resource control multiple applications

Q & A