bhyve Provisioning and Monitoring

Michael Dexter
bhyvecon Tokyo 2014

bhyve.org | bhyvecon.org
editor@callfortesting.org
An Evolution
2002 – 2005

jail(8) VMs in disk images
(root, usr, var)
managed via rc(8)
Limitations

Explosive FreeBSD 5.0
Image Corruption
jail(8) Limited Scope
(compat_linux?)
Limitations

Installers that are designed to be used every two years with a new system
Limitations

Unpopular official jail management infrastructure
Exploration

OpenBSD/SysJail
NetBSD/mult
NetBSD/Xen
Agreement

“We need a Hypervisor”
Surprise!

BHyVe
(now “bhyve”)
The BSD HyperVisor
Goal

Stress Test bhyve
Expose Bugs
SPARE YOU FROM THEM
Limitations

Installers that are designed to be used every two years with a new system...
Limitations

Limited official bhyve management structure (vmrun.sh)
Limitations

ISO boot only
8G disk image
zvols? iSCSI?
Continued Evolution

bhyve (build) menu
bhyve-scripts
bhyve-script
Lessons

SCRIPT EVERYTHING
SCRIPT YOUR SCRIPTS
TEST ASSUMPTIONS
BUILD FUNCTIONS
Current State

vmrc management framework
mkvm.sh provisioning script
Layout

/usr/local/etc/vm.conf
(host variables)

/usr/local/etc/rc.d/vm

/usr/local/vmrc/*
Layout

/usr/local/etc/vm.conf

host_nic="em0"  # Host's active NIC for bridged networking
host_bridge="0"  # bridge0 (Fine for most configurations)
host_vmroot="/usr/local/vmrc/"  # Directory for all vmrc components
host_vmdir="/usr/local/vmrc/vm/"  # VM disk images and mount points
host_distdir="/usr/local/vmrc/distributions/"  # OS Distributions
host_templates="/usr/local/vmrc/templates/"  # VM Templates
host_zpool="zroot"  # ZFS pool for VM zvols
Layout
/usr/local/vmrc/vm/
vm0 freebsd1 openbsd2
Layout

/usr/local/vmrc/distributions/

10.0-RELEASE/base.txz...
Layout
/usr/local/vmrc/templates/freebsd10

# FreeBSD 10.0-RELEASE Template

# VM operating system variables
vm_os_type="freebsd"        # freebsd, openbsd, or linux (required)
vm_os_ver="10.0-RELEASE"    # Exact OS version if auto-fetching
dist_site="ftp://ftp.freebsd.org/pub/FreeBSD/releases/amd64/amd64/10.0-RELEASE/"
iso_site="ftp://ftp.freebsd.org/pub/FreeBSD/releases/amd64/amd64/ISO-IMAGES/10.0/"
iso_img="FreeBSD-10.0-RELEASE-amd64-disc1.iso"
# VM boot variables
vm_cpus="1" # Number of VM virtual CPUs (max 16) (required)
vm_ram="1024" # VM RAM Allocation in MB (required)
vm_console="nmdm" # stdio, nmdm, tmux or tmux-detached (required)
virtio_type="ahci-hd" # "ahci-hd" or "virtio-blk" (required)
vm_hostbridge="" # "amd_" for the AMD hostbridge
bhyve_flags="" # Additional bhyve(8) flags
grub_boot_cmd="" # grub-bhyve command to boot from IMG
grub_iso_cmd="" # grub-bhyve command to boot from ISO

# Optional existing boot device in /dev/ i.e. "ada2"
vm_device="" # Existing bootable block device (sans /dev/)
# FreeBSD-specific VM provisioning variables

```
vm_dev_type="img"  # "img" for image, "zvol" or blank for other device
vm_dev_util="gpart" # Partitioning tool "fdisk" or "gpart" (FreeBSD only)
vm_dev_layout="gpt" # "gpt" or "mbr" volume layout (FreeBSD only)
vm_dev_fs="ufs"    # "ufs" or "zfs" (FreeBSD only)
vm_dev_size="2G"   # M or G for raw "img" volumes (FreeBSD only)
vm_pool="cesspool" # VM ZFS pool name (FreeBSD only)
vm_timezone="America/Los_Angeles"    # VM timezone (FreeBSD only)
vm_hostname="bhyve" # VM hostname (FreeBSD only)
vm_ipv4="192.168.1.210" # VM IPv4 address (blank for DHCP) (FreeBSD only)
vm_gw="192.168.1.1"  # VM IPv4 gateway (FreeBSD only)
vm_searchdomain="" # VM resolv.conf searchdomain (FreeBSD only)
vm_dns="192.168.1.1" # VM resolv.conf IPv4 DNS server (FreeBSD only)
vm_password="bsd" # VM password (clear text for now) (FreeBSD only)
vm_distsets="base.txz kernel.txz"
#vm_pkgs=""
# Additional NICs and DEVs
# zpool template
```
mkvm.sh script

sh mkvm.sh
sh mkvm.sh <template> <name>

virtual/physical host installer
vm rc script

/usr/local/etc/rc.d/vm start
service vm start
service vm status
service vm attach vm0
service vm mount vm0
service vm iso vm0
service vm jail vm0
service qemu vm0
Console Considerations

stdio
nmdm
tmux, tmux-detached
Backing Storage

Raw images
zvols
malloc images
Existing devices
VM Filesystems

fdisk/MBR/UFS

gpart/MBR/UFS

gpart/GPT/UFS

gpart/MBR/ZFS

gpart/GPT/ZFS
VM Filesystems Cont.

Aligned ZFS pools
RaidZ pools
Mirrored Stripes
Thin Provisioning
Prototyping
v2p, p2v
Goals

Stress Test bhyve... quickly
Test edge cases – iSCSI?
Perhaps something useful
Visual Profiling

Keep it simple
Hand-rolled telemetry send
Generate load
Graphite/Carbon/Whisper
Come to my talk Saturday
Questions? Demo?
Thank you!