Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

Linux Containers (LXC)

Linux System Administration Meeting #1

Wednesday, 7th May 2014 15:00 – 17:00

BFH – Room R1 1.21 (Rockhall, Biel)

Daniel Baumann <daniel.baumann@bfh.ch>

IT System Engineer, Infrastructure Team

Overview

Introduction

- ITS Linux System Administration Ressources, Part 1

Linux Containers (LXC)

- Theory
- Practice

Aperitif

ITS Linux System Administration Ressources

ITS Linux System Administration Ressources (Part 1)

https://linux.bfh.ch

- External Content: generic and world readable
- Internal Content: BFH specific and BFH wide readable

- Partially manual content (Software Documentation)
- Partially automatic content (System Documentation)
- Download of Linux Images

https://lists.bfh.ch

- BFH Lists: public and world readable
- ITS Lists: private and Team internal/BFH internal

https://otrs.bfh.ch

- Sub-Queue: Infrastructure Services / Linux

Linux Container (LXC) Theory

Linux Container (LXC) Theory: Leight Weight „Virtualization“

- Nothing excatly new...
  - IBM LPAR
  - FreeBSD Jails
  - Solaris Zones
  - Linux chroot, Linux VServer, Linux OpenVZ, ...
- ...or not really the same...
  - Full/Para-Virtualization: KVM, Xen
  - Proprietary: VMware
- ...now, it is done right this time (aka Generic Ressource Management):
  - Linux Containers (LXC)
  - Docker
  - libvirt-lxc, lmctfy, ...

Linux Container (LXC) Theory: Linux Ressource Management (1)

- Process Attributes
  - Process ID (PID)
  - Parent Process ID (PPID)
  - Nice number
  - Terminal (TTY)
  - Real and Effective User (RUID/EUID)
  - Real and effective Group owner (RGID/EGID)

Linux Container (LXC) Theory: Linux Ressource Management (2)

- Process Permission Checks (aka Capabilities)
  - Good Examples: CAP_MAC_OVERRIDE, CAP_MKNOD, CAP_SYS_TIME
  - Otherwise: CAP HELL :(
  - see capabilities(7)

Linux Container (LXC) Theory: Linux Ressource Management (3)

- Namespaces (NS)
  - providing isolation through new syscall flags (6 for clone) and new 3 syscalls (share/unshare/setns)
  - ips (System V IPC)
  - mnt (Mountpoints, Filesystems)
  - net (Networking)
  - pid (Processes)
  - user (UIDs/GIDs)
  - uts (Hostname)
  - In the Queue: devices, time

Linux Container (LXC) Theory: Linux Ressource Management (4)

- Control Groups (CGroups)
  - providing hierarchical ressource management through grouping of processes (ressource controllers, /sys/fs/cgroup)
  - blkio (block devices)
  - cpu (CPU Access), cpuacct (CPU Accounting), and cpuset (CPU Assignment)
  - devices
  - freezer
  - memory
  - swap
  - net_cls (Network Class identifier)
  - net_prio (Network Priority)
  - perf_event, hugetlb, debug

Linux Container (LXC) Practice

Linux Container (LXC) Practice: Installation

- Simple: sudo apt-get install lxc lxc-stuff debootstrap bridge-utils
- Make sure:
  - most recent Debian version of LXC on the host system
  - most recent version of live-debconfig in /usr/share/lxc/packages (unless building sid)
- Note: Bumpy road ahead (live-debconfig to puppet migration)

Linux Container (LXC) Practice: Configuration (1)

Enable IP Forward

- echo "net.ipv4.ip_forward = 1" > /etc/sysctl.d/ip_foward.conf
- sysctl -p

Optionally add Boot parameter for Memory/Swap Ressource Controller

- cgroup_enable=memory
- swapaccount=1

Mount /sys/fs/cgroups (on non-systemd systems; #601757)

- mount cgroup -t cgroup /sys/fs/cgroup

Linux Container (LXC) Practice: Configuration (2)

Bridge (dhcp)

```
# /etc/network/interfaces

auto lo

iface lo inet loopback

iface eth0 inet manual

auto br0

iface br0 inet dhcp
    bridge_ports    eth0
    bridge_fd   0
    bridge_hello    0
    bridge_maxwait  0
    bridge_stp  0
```

Linux Container (LXC) Practice: Terminology

- 2 Contexts: Host System and Container
- LXC Commands: /usr/bin/lxc-*
- LXC Wrapper: /usr/bin/lxc
- LXC Container: /var/lib/lxc/$container
- Templates: /usr/share/lxc/templates
- Rootfs: /var/lib/lxc/$container/rootfs
- Config: /var/lib/lxc/$container/config
- Privileged and Unprivileged Containers

Linux Container (LXC) Practice: Usage (1)

Create a container

- lxc-create -t debconfig -n example.org

Start a container

- lxc-start -n example.org
- lxc-start -n examle.org -d

Stop a container

- lxc-stop -n example.org
- lxc-kill -n example.org

Linux Container (LXC) Practice: Usage (2)

Remove a container

- lxc-destroy -n example.org

- **Pause/Unpause a container**
- lxc-freeze -n example.org
- lxc-unfreeze -n example.org

Attach a Console to container

- lxc-console -n example.org

List containers

- lxc-list
- lxc-ls –fancy

Linux Container (LXC) Practice: Best Practice

- Always use FQDNs
- Container Preseeding
- Data Separation
- Use as many containers as reasonably needed/possible, only limit is IP address usage
- Combine LXC with your favourite tools (kiss)

Linux Container (LXC) Practice: Future Areas

Future

- Unprivileged containers
- Nested containers
- systemd/cgmanager
- criu
- puppet
- further finetuning and polishing Debian integration (config includes, image tarballs, etc.)

Rant

- libvirt-lxc
- Docker

Thank You for Your Attention.

Linux Container (LXC): Links

Namespaces/CGroups

- LWN Namespaces in Operation: http://lwn.net/Articles/531114/
- Kernel CGroups: Documentation/cgroups/cgroups.txt
- Pro-Linux CGroups Overview:
  http://www.pro-linux.de/artikel/2/1464/ressourcen-verwaltung-mit-control-groups-cgroups.html

LXC

- LXC Homepage: http://linuxcontainers.org
- LXC Networking Explained: http://containerops.org/2013/11/19/lxc-networking/
- LXC 1.0 Blog Post Series: http://www.stgraber.org/2013/12/20/lxc-1-0-blog-post-series/
- BFH: https://linux.bfh.ch/software/lxc/

Further Reading...

- LWN The failure of operating systems and how we can fix it: http://lwn.net/Articles/524952/

Linux Container (LXC): Todays Debian unstable Image for LXC

Component Versions

- Debian unstable (sid) 2014-05-07 with lxc 1.0.3-1
- live-build: 4.0~alpha37-1
- live-images: 4.0~alpha24-1
- live-debconfig: 4.0~alpha32-1

Build

- install live-build and „sudo lb build“ in live-images/images/lxc-server

Download

- https://linux.bfh.ch/files/images/lxc-server/

  Beware: it's Debian unstable (sid), anything except LXC is broken on that image

Future

- Official lxc-server Image for BFH (wheezy/wheezy-backports, and newer) and official for Debian (as of jessie)