Berner Fachhochschule
Haute école spécialisée bernoise
Bern University of Applied Sciences

Linux Containers (LXC)

Thursday, 23th January 2015 15:00 – 16:00

PH Bern, Zentrale Dienste, Fabrikstrasse 2, Bern

Daniel Baumann <daniel.baumann@bfh.ch>

IT System Engineer, Team Infrastructure Services

Overview

Linux Containers (LXC)

- Theory: Ressource Management
- Practice: Installation, Configuration, Usage, etc.

BFH ITS

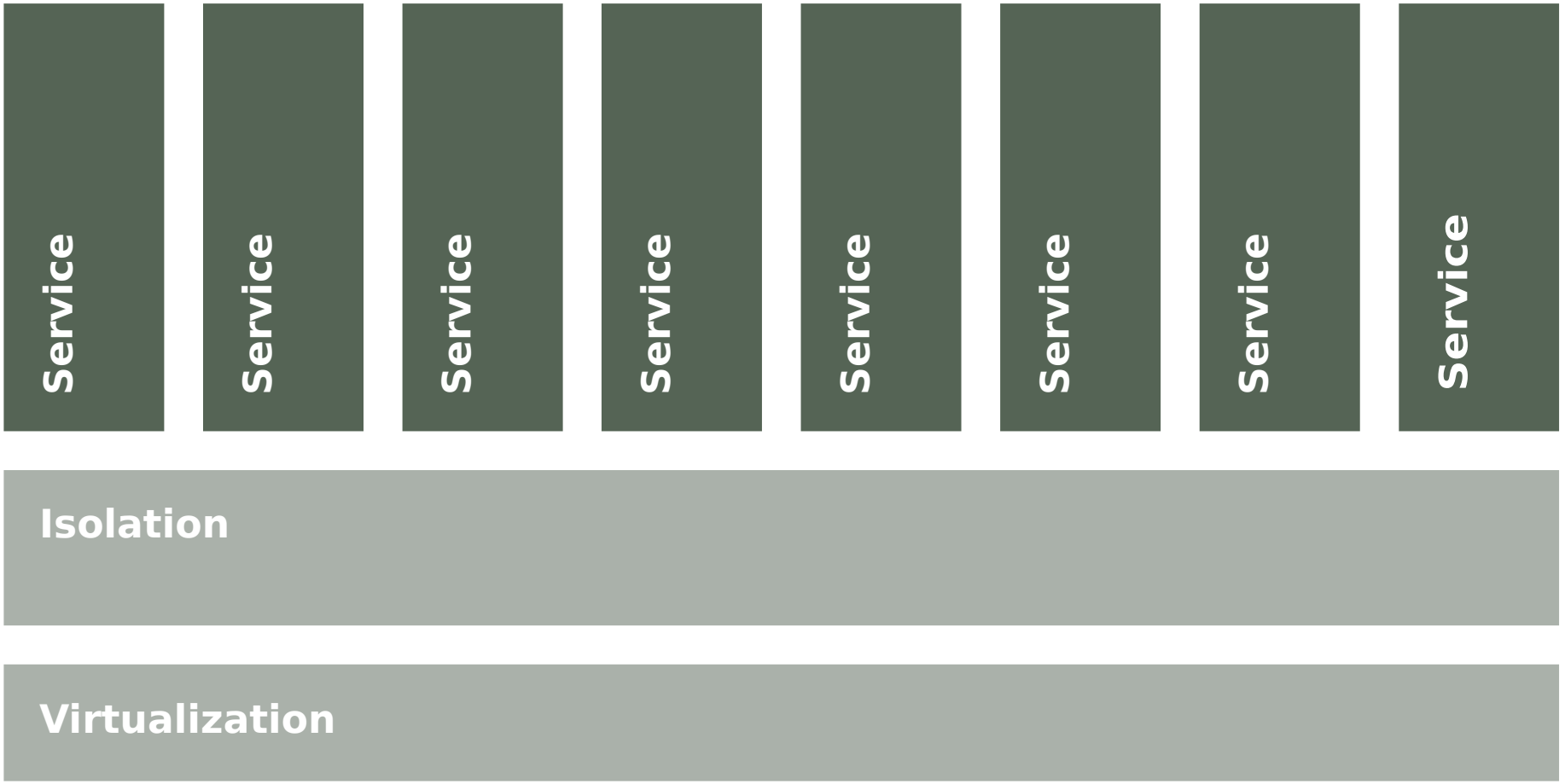
- „Standing on the shoulders of giants“

Linux Container (LXC): Theory

Conventional

Virtualization

Isolation


Service
Service
Service
Service
Service
Service
Service
Service
Isolation
Virtualization

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

- 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-user-chroot, systemd-nspawn ...

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)

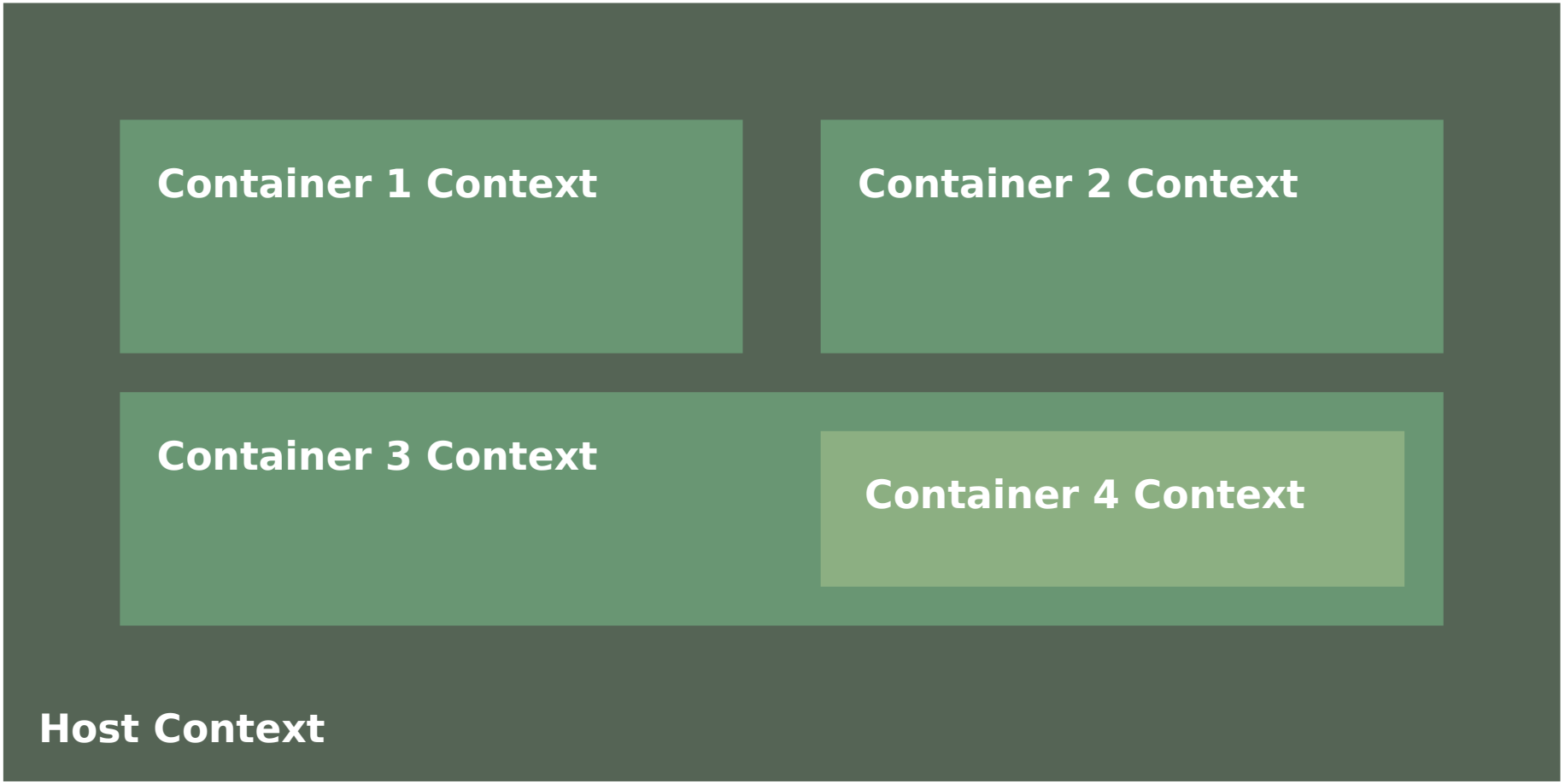
Linux Container (LXC) Theory: Linux Ressource Management (3)

- 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

Control Groups



Linux Container (LXC): Practice

Linux Container (LXC): Installation

- Simple: sudo apt install lxc debootstrap bridge-utils

- **Notes:**
  - use lxc-tools and lxc-support from BFH (open-infrastructure.net) for better integrated lxc debian template
  - expect minor troubles with vanilla Debian (including jessie), fixes available (more about that later)

Linux Container (LXC): Configuration (1)

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): Configuration (2)

Choose LXC Networking Types

- empty (= no network)
- veth (**prefered**; bridge, with or without NAT)
- macvlan (private, vepa, or bridge)
- vlan
- phys

Enable IP Forward

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

Linux Container (LXC): Terminology

- Contexts: Host System and Container
- LXC Commands: /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): Usage (1)

Create a container

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

Start a container

- lxc-start -n example.org
- lxc-start -n examle.org -d|--daemon
- lxc-start -n example.org -F|--foreground

Stop a container

- Shutdown: lxc-stop -n example.org
- Kill: lxc-stop -n example.org -k|-kill

Linux Container (LXC): 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): Preseeding

lxc-tools

- Uses debconf, therefore User and Admin friendly
- Two Stages: Bootstrapping and Postinst
- Preseed Files: /etc/lxc/debconfig
- Preseed Examples: /usr/share/lxc/preseed

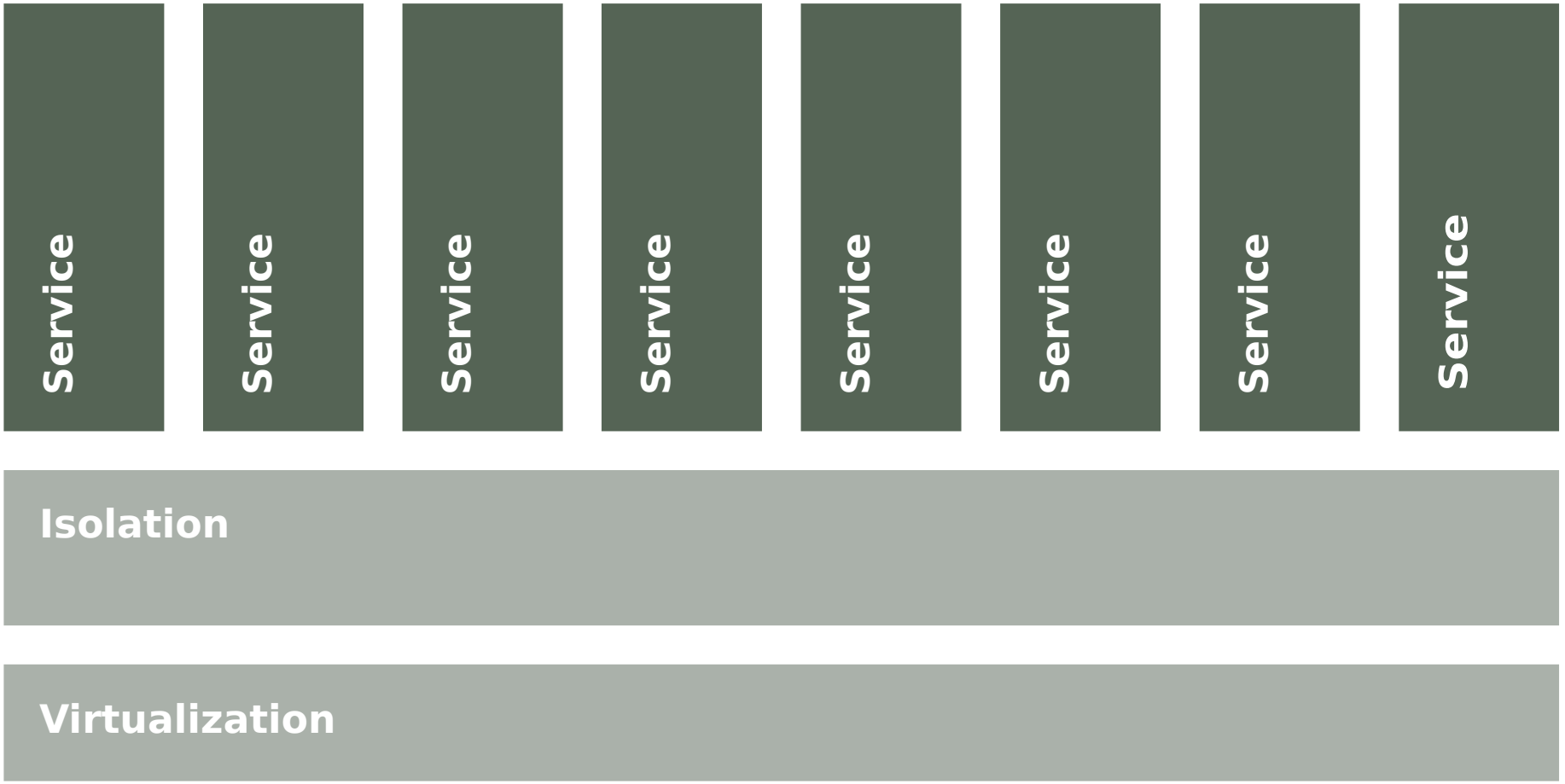
lxc-support

- Uses debconf, therefore User and Admin friendly (dejavu :)
- Current backend shell

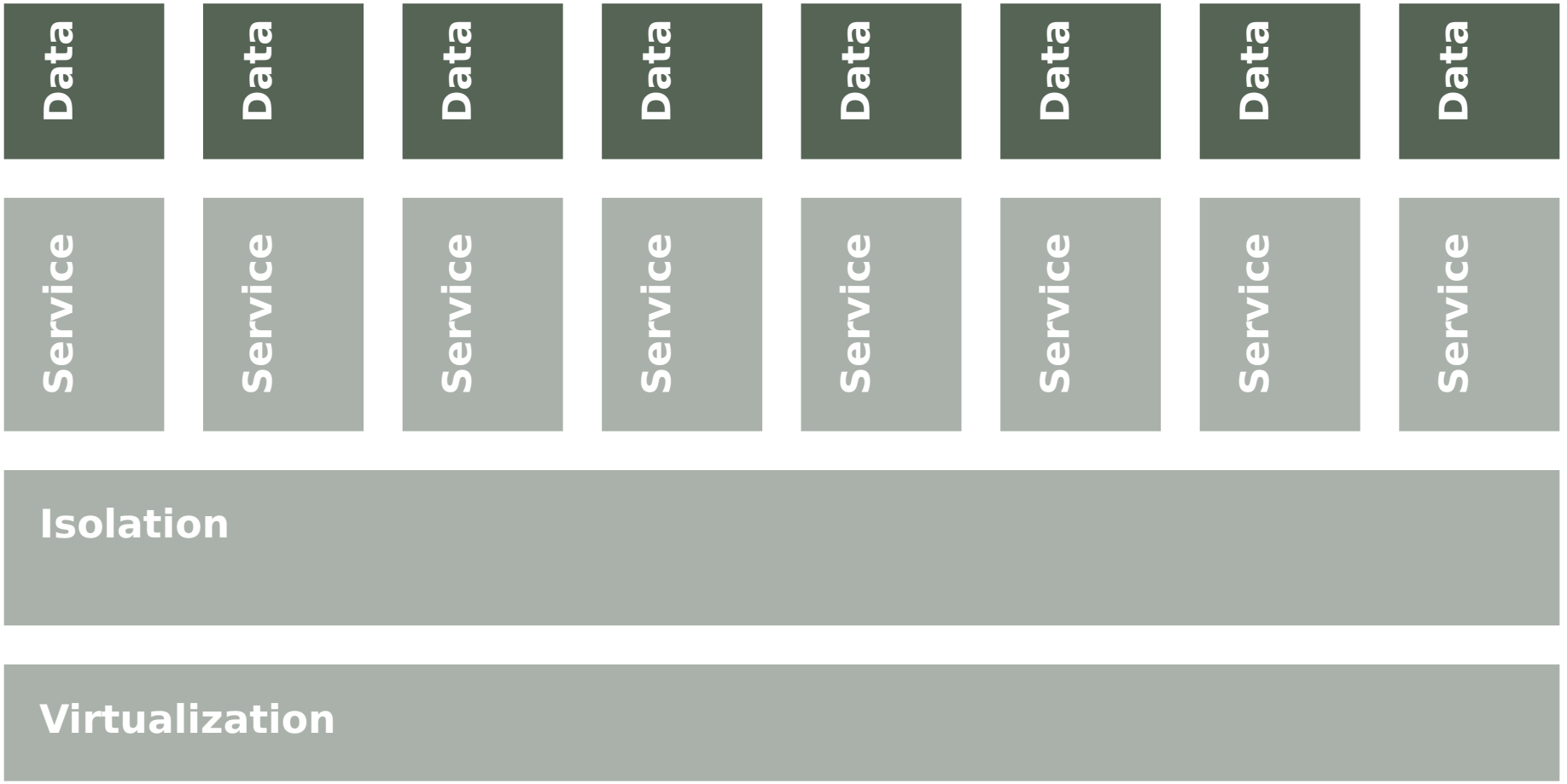
Linux Container (LXC): Best Practice

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

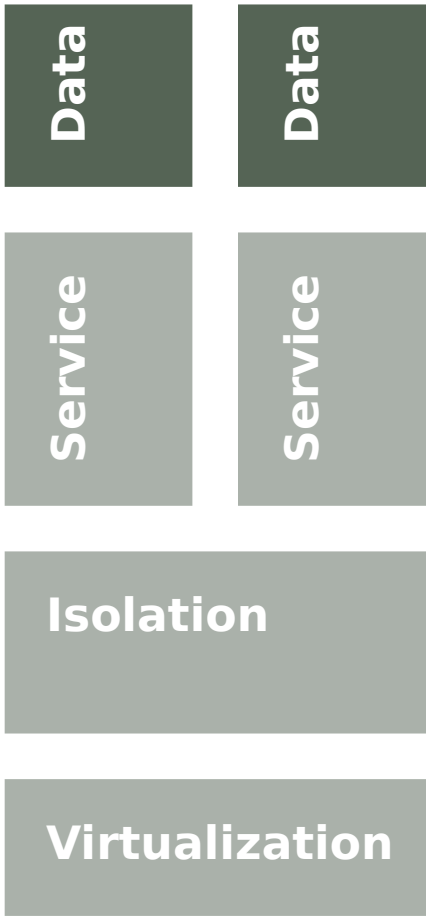
Isolation


Service
Service
Service
Service
Service
Service
Service
Service
Isolation
Virtualization

Separation


Data
Data
Data
Data
Data
Data
Data
Data
Service
Service
Service
Service
Service
Service
Service
Service
Isolation
Virtualization

Linux Container (1)



Host:
`/srv/lxc/data/$fqdn`

Container:
`/srv/$fqdn`

Host:
`/srv/lxc/containers/$fqdn`

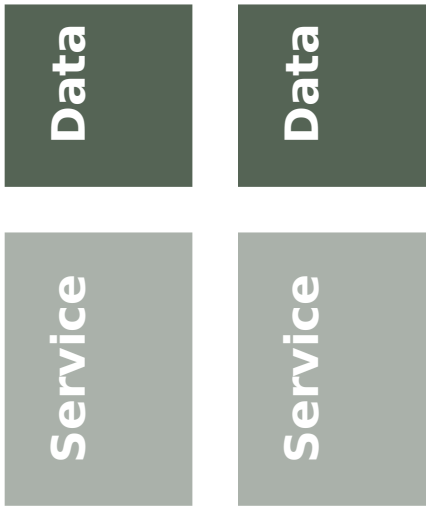
Container:
`/`

Host:
/

Container:
n/a

Host:
n/a

Linux Container (2)



Storage:
`bfhfilerbe01.bfh.ch:/vol/lx_data_01_v0_sata_enge`

Storage:
`bfhfilerbe02.bfh.ch:/vol/lx_containers_01_sas_enge`

Storage:
becenter

Storage:
becenter

Linux Container (LXC): Future Areas

Future

- Unprivileged containers
- Nested containers
- Systemd Integration
- criu
- openvswitch
- LSM (SELinux)
- LXD
- ansible
- rebase debconf based template on top of live-build

Rant

- libvirt-lxc
- Docker
- Security

BFH ITS

Projects (1/3)

Debian (debian.org)

- The **Universal** Operating System, since 1993
- Project with ~1'250 Developers/Maintainers
- ~50'000 Packages per Release

Debian Live (live.debian.net)

- Official Debian Subproject, since 2006
- Handful of People
- Defacto Industry Standard

Projects (2/3)

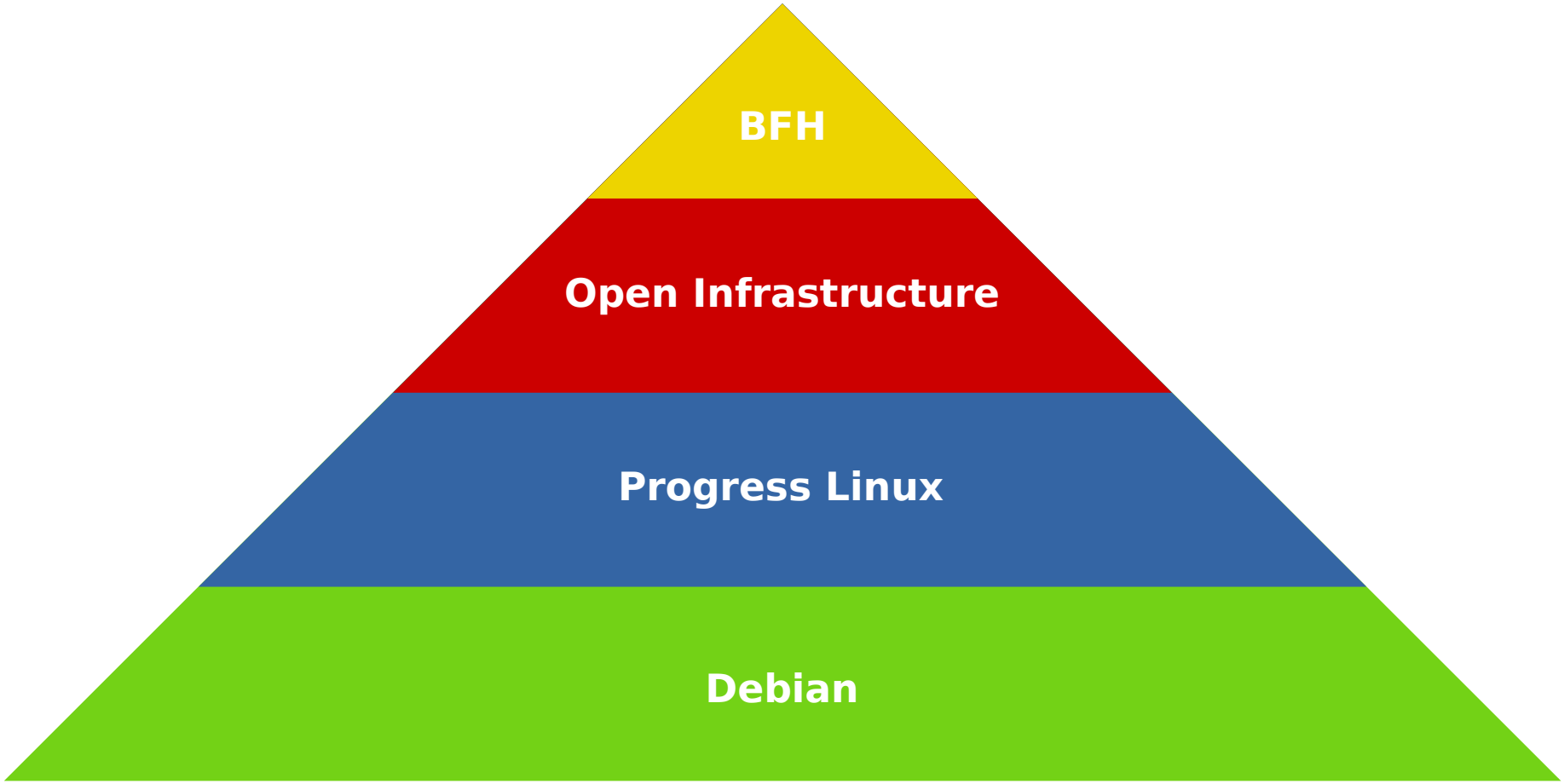
Progress Linux (progress-linux.org)

- Debian Derivative Project focused on System Integration, since 2010
- Not a Debian Fork (like e.g. Ubuntu) but a Repository on top of Debian (Debian+ Model)
- Handful of People
- 100 to 200 Packages per Release
- Aims to be **generic**
- Fixing paper cut bugs, optimizes for the majority/modern systems, integrating backports properly

Open Infrastructure (open-infrastructure.org)

- Infrastructure Project, since 2014.. (association founding meeting in 2 weeks! :)
- Handful of People
- Specific ready-made out-of-the-box Solutions (e.g. GNOME Desktop, LXC Server, ...)

Standing on the shoulders of giants



Projects (3/3)

Management Summary

- Debian: Original Equipment Manufacturer (OEM)
- Progress Linux: Value-Added Reseller (VAR)
- Open Infrastructure: Solution Provider
- BFH: Local Instance

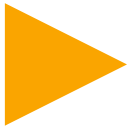
Techie Summary

- Debian: The best a Geek can get
- Progress Linux: PPA with bugfixes and missing packages for Debian
- Open Infrastructure: Upstream Code, providing missing glue for System Administration
- BFH: Branding

Common Principles

- no NIH Syndrom; DRY instead of WET; don't believe in Competitive Advantage wrt/ Infrastructure

...or even shorter:



https://linux.bfh.ch/

Thank You for Your Attention.

❤ Source Code is freely available

git clone git://git.bfh.ch/git/staff/bad9/other/talks.git

Further Information

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/, https://linux.bfh.ch/services/lxc, and https://linux.bfh.ch/servers/

Contact Information

BFH

- Homepage: https://linux.bfh.ch
- Mailing List: bfh-linux-users@lists.bfh.ch
- IRC: irc://irc.oftc.net #bfh-linux-users

Open Infrastructure

- Homepage: https://open-infrastructure.net
- Mailing List: project@lists.open-infrastructure.net
- IRC: irc://irc.oftc.net #open-infrastructure

Progress Linux

- Homepage: https://progress-linux.org
- Mailing List: project@lists.progress-linux.org
- IRC: irc://irc.oftc.net #progress-linux