A New Xen in SUSE® Linux Enterprise Server 12

Jim Fehlig
Software Engineer
jfehlig@suse.com

Russell Pavlicek
Xen Project Community Evangelist
Russell.Pavlicek@XenProject.org

Agenda

- New Xen in SUSE Linux Enterprise Server 12
  - System changes
  - New Xen features
  - Toolstack changes
- Xen Project Community Update

grub2

- /etc/default/grub
  - GRUB_CMDLINE_XEN_DEFAULT
    - Options for xen.gz, e.g. loglvl, dom0_mem, etc
  - GRUB_CMDLINE_LINUX_DEFAULT
    - Options for kernel-xen
  - Run 'grub2-mkconfig -o /boot/grub2/grub.cfg' after changes
- Upstream grub supports a Xen PV image
  - /usr/lib/grub2/x86_64-xen/grub.xen in grub2-x86_64-xen package
  - Use as 'kernel' for PV domains
    - Runs grub against the domain's image, similar to booting physical machine
    - More secure and cloud-friendly than pygrub

btrfs

- Host
  - Default partitioning scheme
    - '/' gets a small share of the pie
    - VM images created in /var/lib/{kvm,libvirt,xen}/images
  - Snapshots
    - Generally not useful to snapshot VM image files
- Virtual Machine
  - No support in pygrub
    - Use grub-xen or another filesystem supported by pygrub
  - May cause problems in other hypervisors (XenServer) that don't have btrfs support in their tools

systemd

- All services converted to systemd
  - systemctl restart xendomains.service
  - rcxendomain restart for back compat
- journalctl
  - journalctl [--follow] --unit libvirtd.service
  - journalctl /usr/sbin/libvirtd

SLES® 12 Xen Stack

- Xen 4.4.1
- kernel 3.12.x
- libvirt 1.2.5
- virt-install 1.1.x, vm-install 1.x.x
- virt-manager 1.1.x

New Features

- New toolstack!
- NUMA scheduler affinity
- Modern qemu
- Event channel scalability
- Intel Nested Virtualization (tech preview)
- SPICE
- GRUB2
- qdisk disk backend
- Hardware support
  - vAPIC, ACPI v5

kernel-xen

- SLES12 retains the “xenified” kernel
- Changes to kernel-default generally apply to kernel-xen
- pvops kernel for SLES12 SP1?

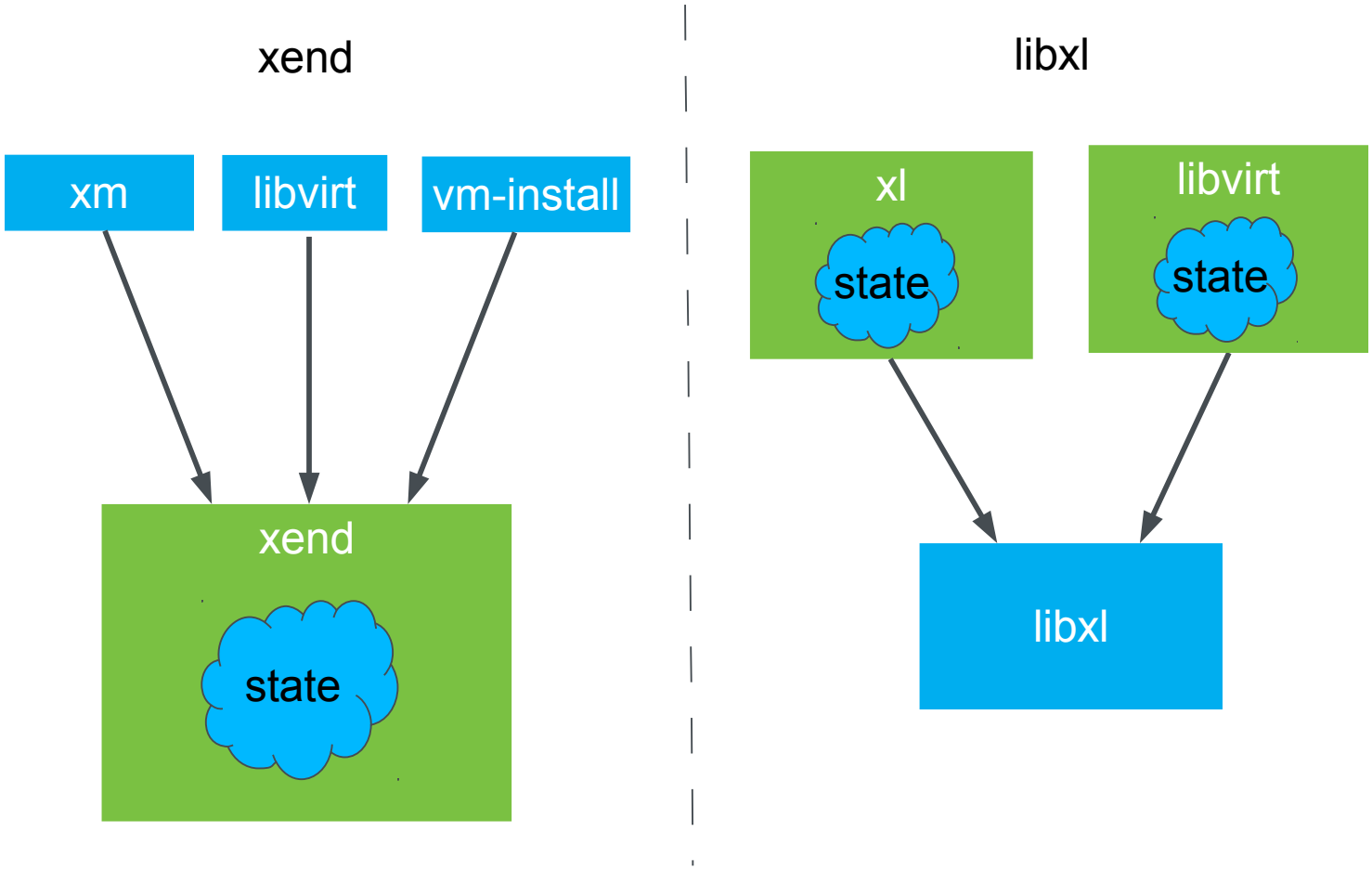
New Toolstack!

- RIP xm/xend
  - Officially deprecated upstream in Xen 4.3
  - Removed from sources in upcoming Xen 4.5
  - Managed (or persistent) domains are gone
  - /etc/xen/xend-config.sxp is gone
  - domUloader is gone
  - Support for s-expression config format is gone
- Welcome xl/libxl (aka libxenlight)!

xl/libxl

- Fast, light-weight, improved concurrency
- Technical preview in SUSE Linux Enterprise Server 11 SP3
- Default toolstack in openSUSE13.1
- Only toolstack in SLE12/Factory
- Recommend disabling xend when using xl/libxl

xend vs libxl



xl vs xm

- xl is just another (in-tree) libxl client
  - Don't use with other libxl clients like libvirt!
    - Use 'xl -f …' to force use of xl with “foreign” domains
- man 1 xl
- No managed domain support
- Compatible subcommands
  - Exception: new, start, delete, suspend, resume
- Behavior adjustable via /etc/xen/xl.conf
  - man 5 xl.conf
- 'xl migrate' needs a properly configured ssh setup on both dom0 hosts

RIP /etc/xen/xend-config.sxp

- Replaced with /etc/xen/xl.conf (man 5 xl.conf)

xend-config.sxp	xl.conf
network-script	Use distro tools
vif-script	vif.default.script
dom0-min-mem	No equivalent, use Xen cmdline option
enable-dom0-ballooning	autoballoon
dom0-cpus	No equivalent, use Xen cmdline option
vncpasswd	Set per domain
device-{create,destroy}-timeout	No equivalent
xend-domain-lock	No equivalent

dom0-min-mem

- Set via parameters to xen.gz
  - dom0_mem=min:<size>,max:<size>,<size>
  - dom0_mem=min:512,2G
- Recommend fixed dom0 mem size + autoballooning off
  - dom0_mem=4G
  - autoballoon=”off” in /etc/xen/xl.conf
  - libvirt turns off autoballooning if dom0_mem is set

dom0-cpus

- Set via parameters to xen.gz
  - dom0_max_vcpus=<num>
  - dom0_max_vcpus=<min>-<max>
  - dom0_max_vcpus=4-8

xl vs xm Config Format

- man 5 xl.cfg
- xm config is a subset of xl
  - Existing xm config should just work with xl
  - Exception: embedded python
- xl supports a new disk config sytax
- New features only supported in xl format
  - SPICE, paging options (hap, shadow_memory), device model, ...
- libvirt supports both formats (plus s-expression)

xl Disk Config Syntax

- Positional

  target, format, vdev, access, devtype

  /var/lib/xen/images/dom/disk0.raw, raw, xvda, rw

  /foo.iso, raw, xvdc, ro, cdrom

- Key/value

  format=raw, vdev=xvda, access=rw, target=/var/lib/xen/images/dom/disk0.raw

  format=raw, vdev=xvdc, access=ro, target=/foo.iso, devtype=cdrom

- xm disk config syntax still supported

  file:/var/lib/xen/images/dom/disk0.qcow2,xvda,w'

- /usr/share/doc/packages/xen/misc/xl-disk-configuration.txt

Disk Backends

- qdisk
  - Default backend for raw and qcow2 formats
  - Adjustable cache modes: writethrough, directsync, unsafe
  - Active upstream community
- blkbk
  - Only used for 'phy:'
  - To use with raw image files (like SLE11)

```
losetup /dev/loopx /image/file/path
xl:       phy:/dev/loopx
libvirt: <driver name='phy/><source file='/dev/loopx'/>
```

Disk Backends

- Tap1..N
  - Slow upstream death
  - Everything except vhd mapped to qdisk

Log Files

- /var/log/xen/xl-<dom-name>.log
- /var/log/xen/qemu-dm-<dom-name>.log
- /var/log/xen/xen-hotplug.log
  - 'set -x' at beginning of /etc/xen/scripts/xen-hotplug-common.sh

Troubleshooting

- Supportconfig for most issues
- Increase log level
  - GRUB_CMDLINE_XEN_DEFAULT=”loglvl=all guest_loglvl=all”
- Serial line
  - GRUB_CMDLINE_XEN_DEFAULT=”loglvl=all guest_loglvl=all console=com1 com1=11520,8n1”
  - GRUB_CMDLINE_LINUX_DEFAULT=”console=ttyS0,11520”
- Debug keys
  - xl debug-keys h; xl dmesg
  - xl debug-keys q; xl dmesg

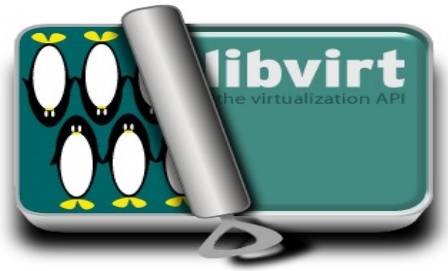
Regressions

- pvusb
- Domain migration from SLE N to SLE N+1 hosts
  - SLE11 SP3 -> SLE12

libvirt Changes

- New, stateful libxl driver
  - Another libxl client
  - (don't mix and match libxl client applications)
- Backwards compatible with legacy Xen driver
  - Existing domXML should just work after switching Xen toolstacks
  - Exception: Currently no dom0!
- Provides managed domain support
- Logs
  - /var/log/libvirt/libvirtd.log
  - /var/log/libvirt/libxl/<dom-name>.log

libvirt Changes



- New package structure
  - Minimal libvirt-daemon
  - Dynamically loaded libvirt-daemon-driver-<driver> modules
  - Allows constructing a custom libvirt daemon

Xen Config Format <-> libvirt domXML

- virsh domxml-{from,to}-native <format> /config/path
  - Formats: xen-sxp, xen-xm, xen-xl (in progress)
- xen2libvirt.py
  - Convert Xen config to libvirt domXML
  - Option to import Xen config to libvirt
  - Option to convert or import a tree of Xen config
  - Supports xen-sxp and xen-xm formats

virt-install vs vm-install

- virt-install
  - virt-manager's default installation tool
  - Requires libvirt
  - Supports remote installations
  - Does not emit a config file
- vm-install
  - Works with or without libvirt
  - '--use-xl' option to force xl
- Logs
  - ~/.cache/virt-manager/virt-manager.log

SLE11 SP3 -> SLE12 Host Upgrade

- Export domain config

  virsh dumpxml > dom.xml

- Upgrade to SLE12
- Import domain config

  virsh define dom.xml

- Or use xen2libvirt utility
  - xen2libvirt –recursive /var/lib/xend/domains
- No live migration of domains from SLE11 SP3 to SLE12 host

Xen Project: A Vibrant, Innovative Community

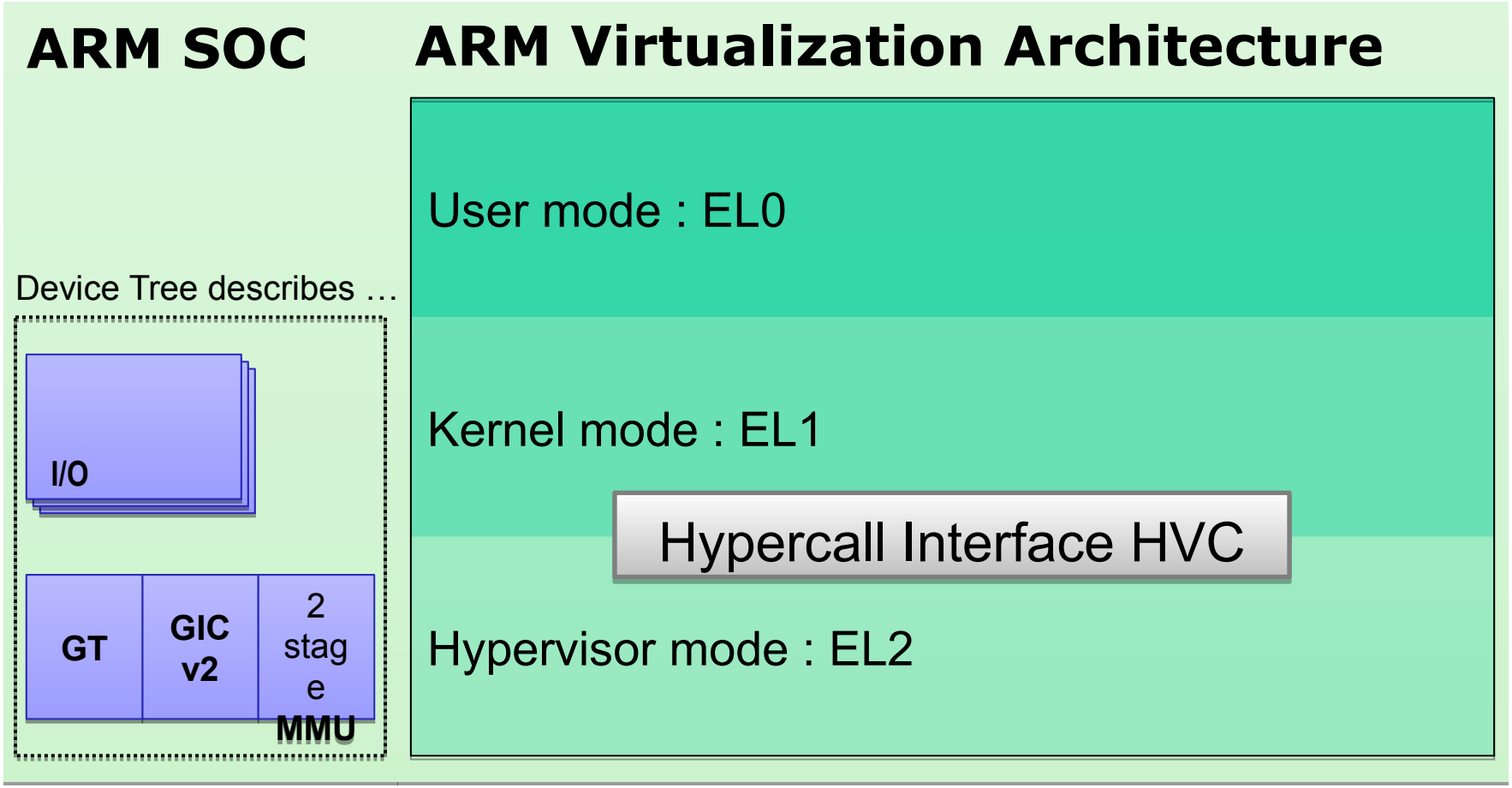
Xen Project: Continuing Innovation

- Began its life over a decade ago with the invention of paravirtualization
- Moving forward in new directions:
  - ARM servers
  - Handheld devices
  - Embedded applications
  - Unikernels/Cloud Operating Systems
  - Better integration with clouds and orchestration technologies

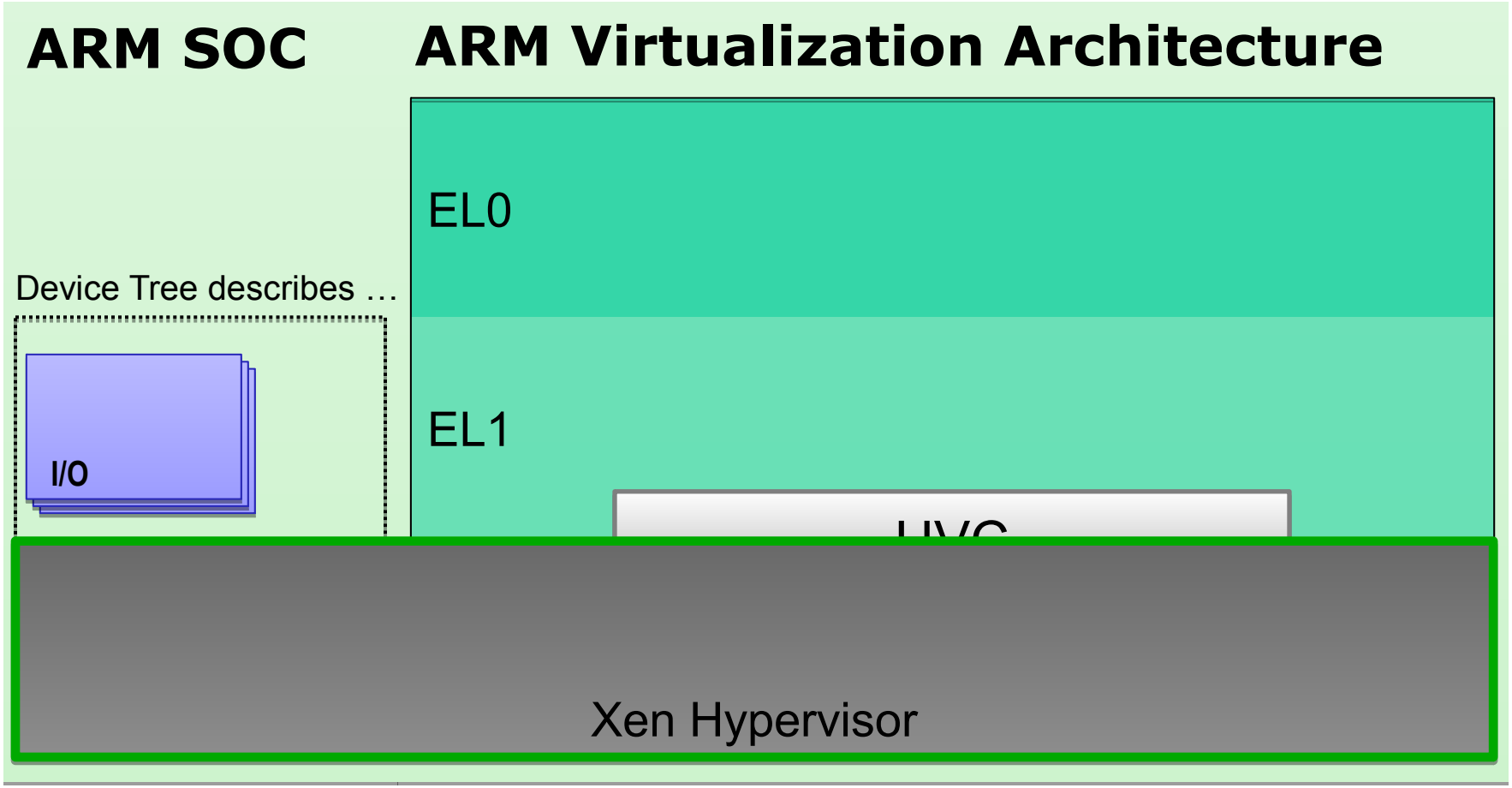
Optimal ARM Hypervisor Design

- Xen Project's design maps directly to the ARM virtualization extensions added in ARM v7
- Type 1 architecture works directly in the ARM Hypervisor Mode
- No need to switch into privileged Kernel Mode
  - Increases security
  - Improves performance
  - Efficient codebase; a fraction of the code required for x86

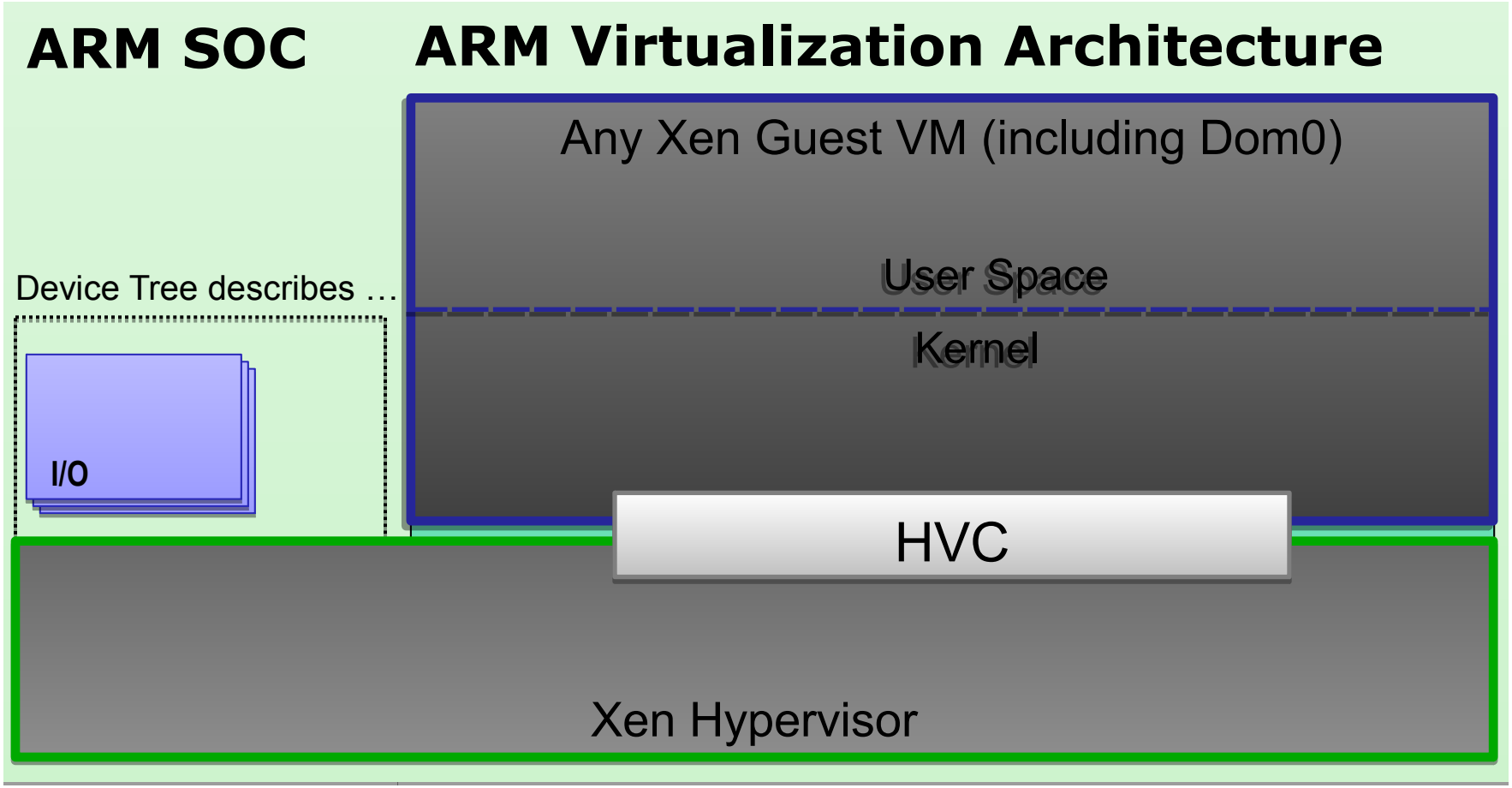
Xen Project on ARM: A Perfect Fit!



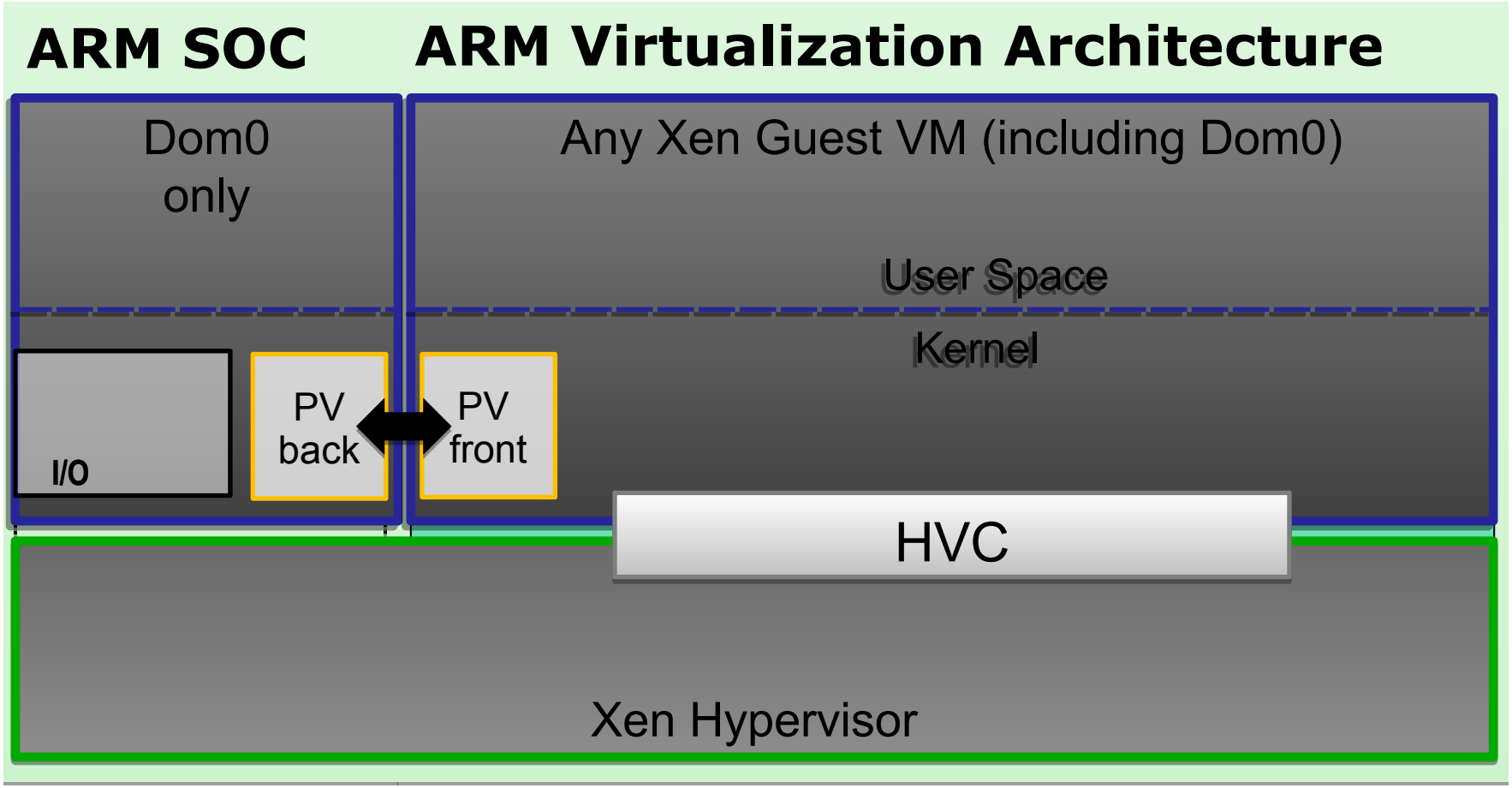
Xen Project on ARM: A Perfect Fit!



Xen Project on ARM: A Perfect Fit!



Xen Project on ARM: A Perfect Fit!



What Good is ARM?

- ARM servers on the horizon
  - Low power consumption
  - Low cooling needed
  - Higher density
  - Potential for higher density of servers at lower operational cost
  - Perfect for the forecast of increasing cloudiness of datacenters

ARM in Embedded and Handhelds

- Embedded systems
  - New solutions: Hypervisors in embedded devices
  - Already have Xen Project in x86 network security devices
  - Coming soon: Xen Project hypervisor in your *car*?
    - <Wiki page reference here>
- Handheld devices
  - New thinking: Hypervisor on your phone
  - Why wipe out your life when you can simply erase one of many identities?

Enabling the Future, Not Dictating It

- We want to enable *your* vision, not force you into ours
  - Huge difference between Open Source and closed source solutions in this area
  - Many key innovations come from the desire of the community to meet their own needs
  - SUSE Linux + Xen Project = the power and flexibility to enable your vision, regardless what it looks like
    - Regardless of cloud: SUSE Cloud, OpenStack, CloudStack, OpenNebula
    - Regardless of orchestration: Xen Orchestra, ConVirt, virt-manager

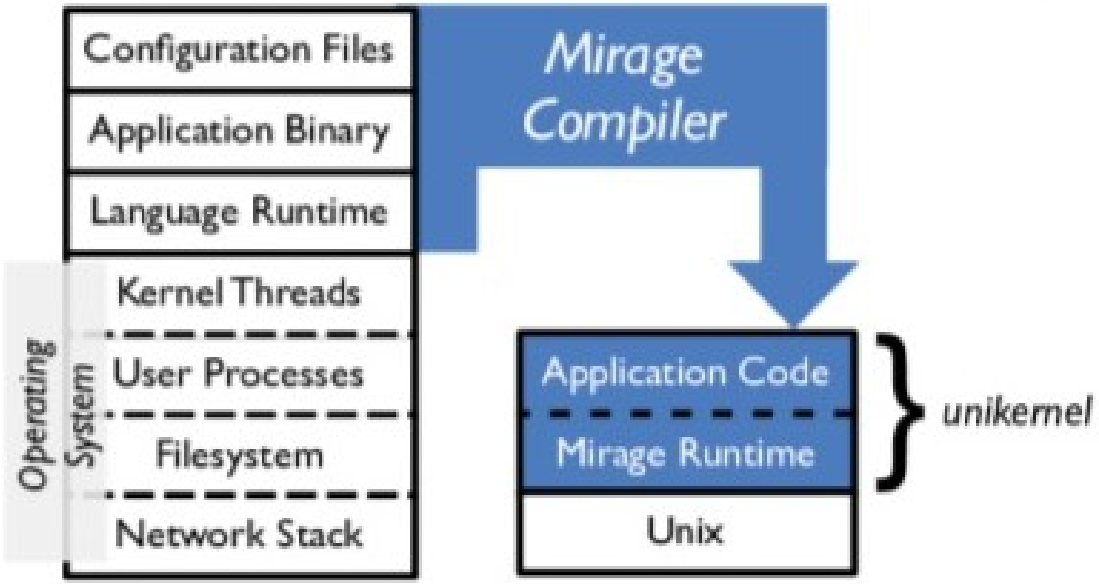
Unikernel: Powering Next Gen Clouds

- Xen Project leading the way with Mirage OS
- Unikernels employ small, secure, packaged VMs
  - Many advantages of Docker but with smaller instance sizes and very high security
  - Potential to power much higher density clouds with improved security
  - Not just Mirage OS from Xen Project, but support for many others
    - OSv, HaLVM, LING (formerly Erlang-on-Xen), ClickOS, and more

Unikernel Approach: Mirage OS

Swap system libraries to target different platforms:

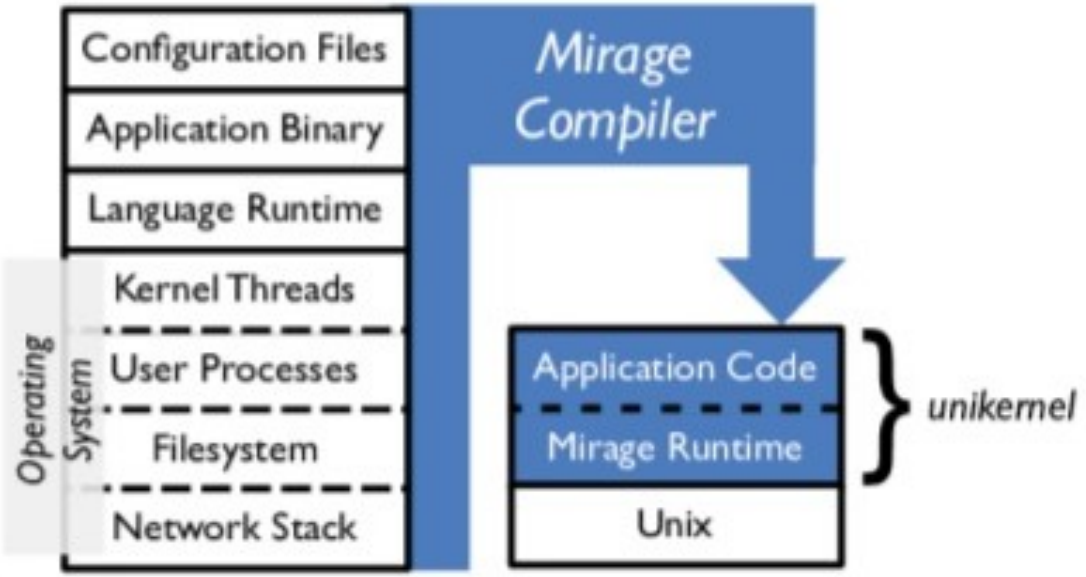
develop application logic using native Unix.



Unikernel Approach: Mirage OS

Swap system libraries to target different platforms:

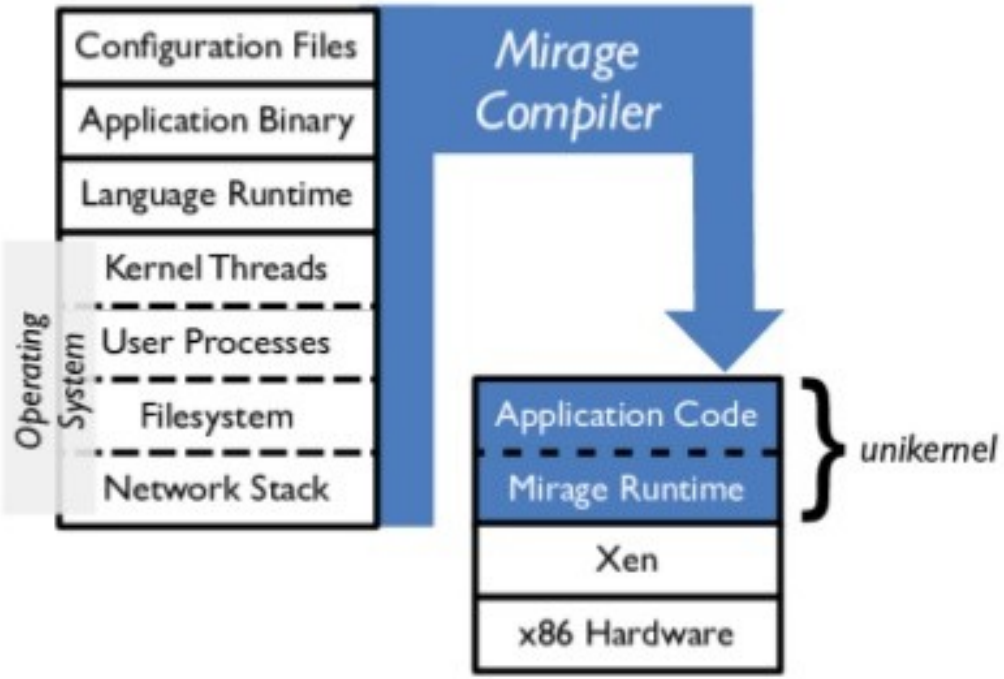
test unikernel using Mirage system libraries.



Unikernel Approach: Mirage OS

Swap system libraries to target different platforms:

deploy by specialising unikernel to Xen.



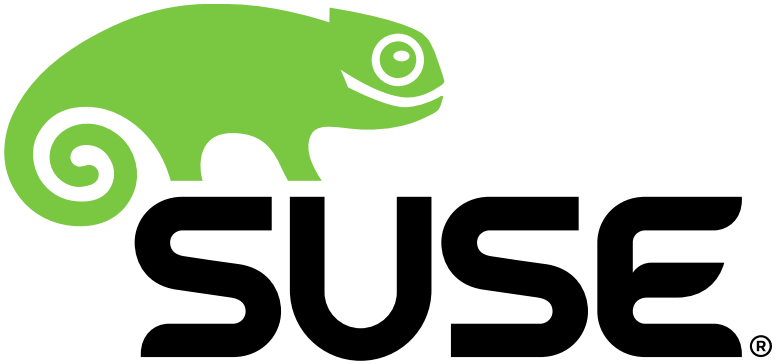
Unikernel Results from Mirage OS

- **Contained:** easy deployment and management
- **Compact:** smaller attack surface, sub-second boot
- **Efficient:** Scales to 1000s of VMs per host
- *Amazingly* small footprints for appliances!
  - DNS server: *less than 500 kilobytes*
  - Webserver: *less than 700 kilobytes*
  - Openflow switch: *less than 400 kilobytes*
- More information at: http://www.openmirage.org/

Questions?

- XenProject.org
- suse.com/documentation/sles12/

Thank you.


SUSE®
We adapt. You succeed.