Virtualization in the Cloud: Featuring Xen

Lars Kurth

Xen Community Manager

lars.kurth@xen.org

FREENODE: lars_kurth

Follow me!

@lars_kurth

@xen_com_mgr

A Brief History of Xen in the Cloud

Late 90s

XenoServer Project
(Cambridge Univ.)

The **XenoServer project** is building

> *public infrastructure for wide-area distributed computing*.

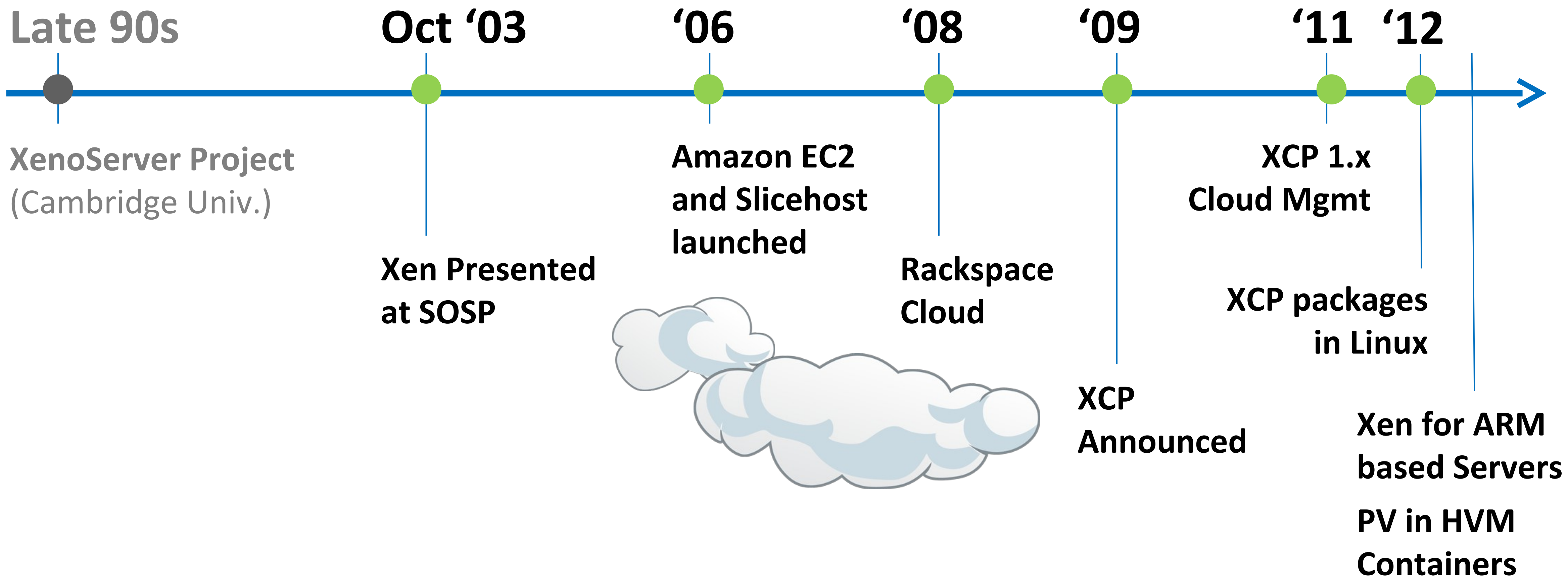
We envisage a world in which **XenoServer** execution platforms will be scattered across the globe and available for any member of the public to submit code for execution.

<u>Global Public Computing</u>

*"This dissertation proposes a new **distributed computing paradigm**, termed global public computing, which allows any user to run any code anywhere. Such platforms **price computing resources**, and ultimately **charge users for resources** consumed."*

Evangelos Kotsovinos, PhD dissertation, 2004

A Brief History of Xen in the Cloud



Late 90s
XenoServer Project
(Cambridge Univ.)

Oct '03
Xen Presented
at SOSP

'06
Amazon EC2
and Slicehost
launched

'08
Rackspace
Cloud

'09
XCP
Announced

'11
XCP 1.x
Cloud Mgmt

'12
XCP packages
in Linux

Xen for ARM
based Servers

PV in HVM
Containers

The Xen Hypervisor was designed for the Cloud straight from the outset!

Xen.org

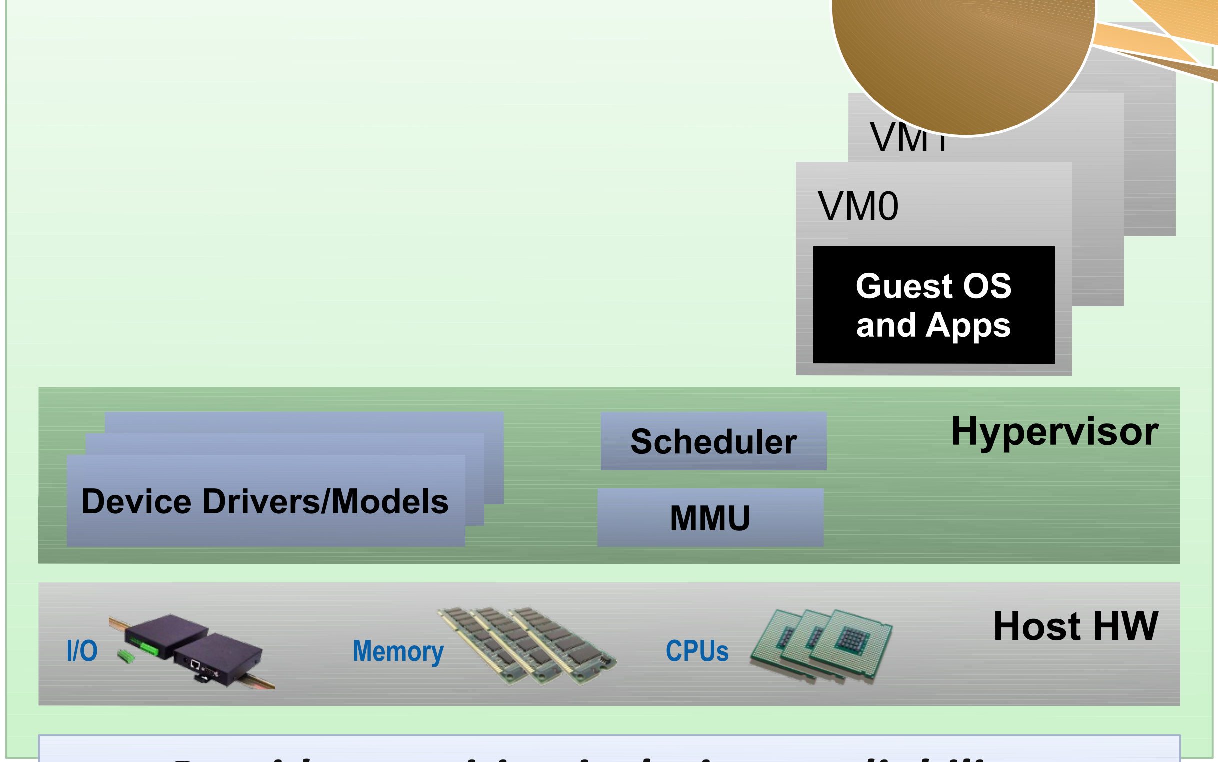
- Guardian of Xen Hypervisor and related OSS Projects
- Xen project Governance similar to Linux Kernel
- Projects
  - Xen Hypervisor (led by Citrix)
  - Xen Cloud Platform aka XCP (led by Citrix)
  - Xen ARM : Xen for mobile devices (led by Samsung)
  - PVOPS : Xen components and support in Linux Kernel (led by Oracle)
- 10+ vendors contributing more than 1% to the project
  (AWS, AMD, Citrix, GridCentric, Fujitsu, Huawei, iWeb, Intel, NSA, Oracle, Samsung, Suse, …)

Xen Overview

Architecture Considerations

Type 1: Bare metal Hypervisor

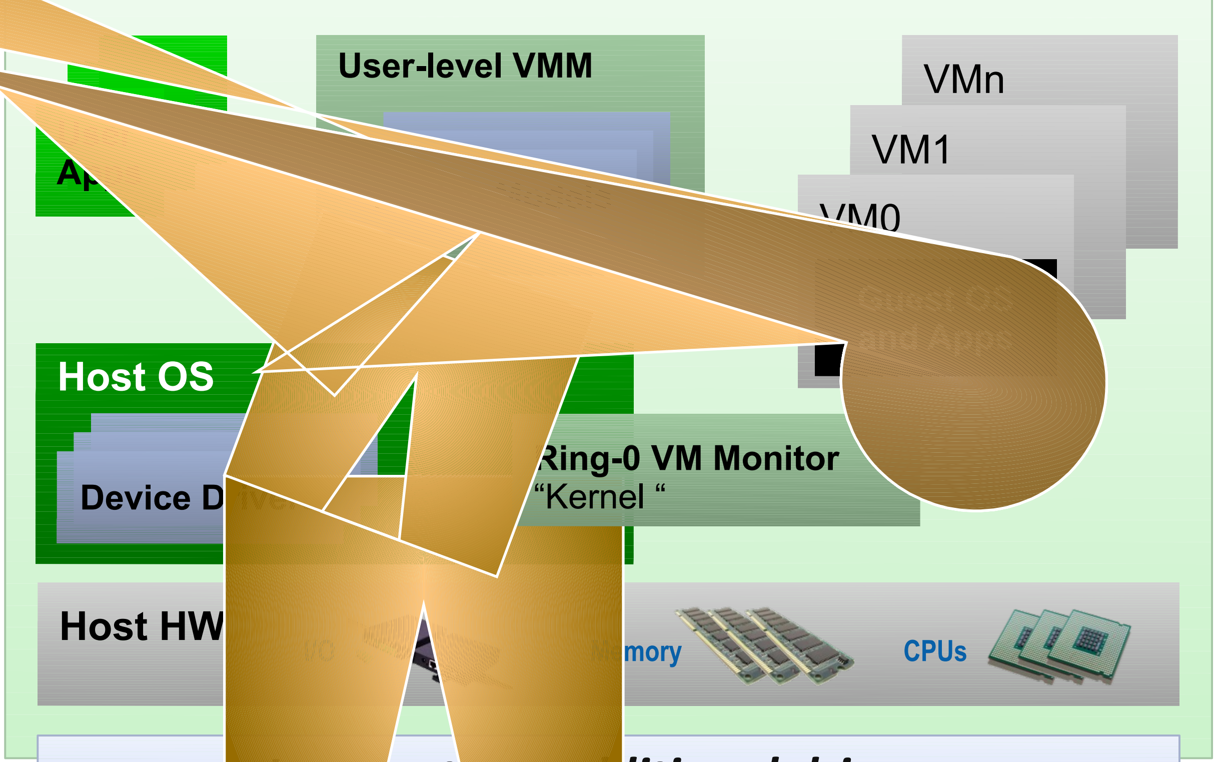
A pure Hypervisor that runs directly on the hardware and hosts Guest OS's.



Provides partition isolation + reliability, higher security

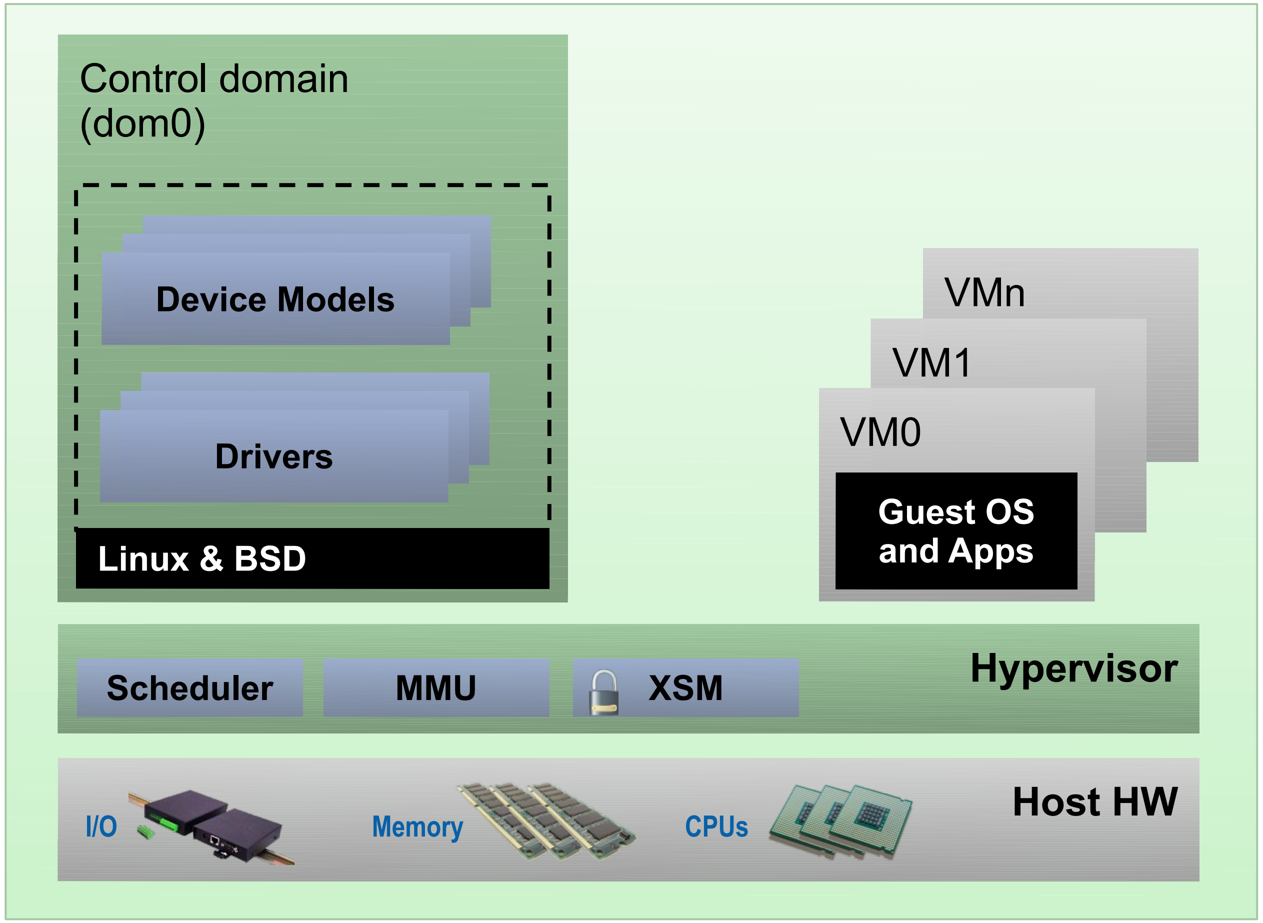
Type 2: OS 'Hosted'

A Hypervisor that runs within a Host OS and hosts Guest OS's inside of it, using the host OS services to provide the virtual environment.



Low cost, no additional drivers
Ease of use & installation

Xen: Type 1 with a Twist



Thinner hypervisor

- Functionality moved to Dom0

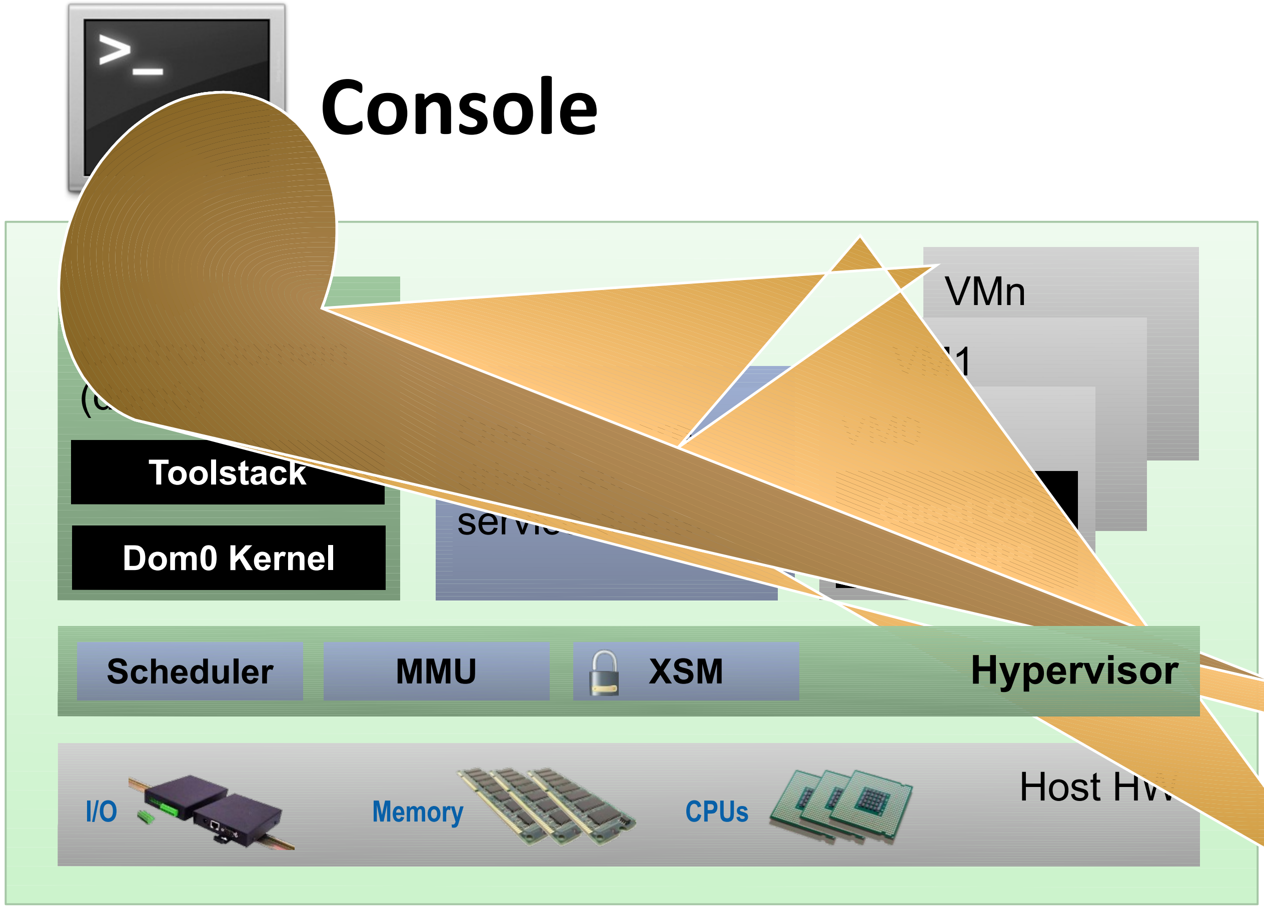
Using Linux PV OPS

- Using Linux Device Drivers
- PV, PV on HVM and PVH modes
- Sharing components with KVM

In other words

- Driver re-use
- Ease of use & Installation
- Isolation & Security

Basic Xen Concepts



Control Domain aka Dom0

Dom0 kernel with drivers

Xen Management Toolstack

Trusted Computing Base 🟩

Guest Domains

Your apps

E.g. your cloud management stack

er/Stub/Service Domain(s)

river, device model or control

ce in a box”

e-privileged and isolated

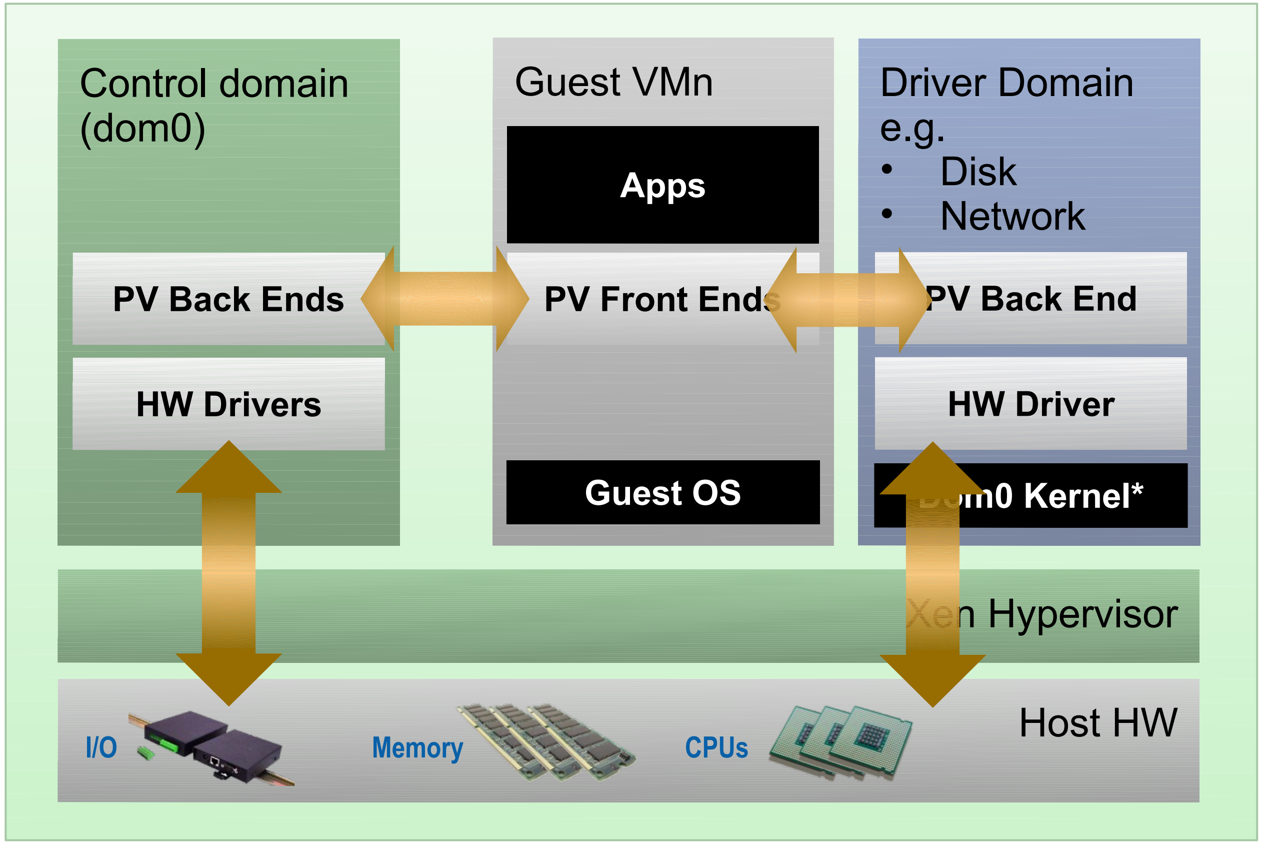
Lifetime: start, stop, kill

Xen Variants for Server & Cloud

	Xen Hypervisor		XCP
Toolstack / Console	Default / XL (XM)	Libvirt / VIRSH	XAPI / XE
	Increased level of functionality and integration with other components		
Get Binaries from …	Linux Distros	Linux Distros	Debian & Ubuntu ISO from Xen.org
Products	Oracle VM	Huawei UVP	Citrix XenServer
Used by …	amazon web services™ More info …	Many Others	CLOUD SERVERS™ Custom server instances on demand More info …

Xen : Types of Virtualization

PV Domains & Driver Domains



*) Can be MiniOS

Linux PV guests have limitations:

- limited set of virtual hardware

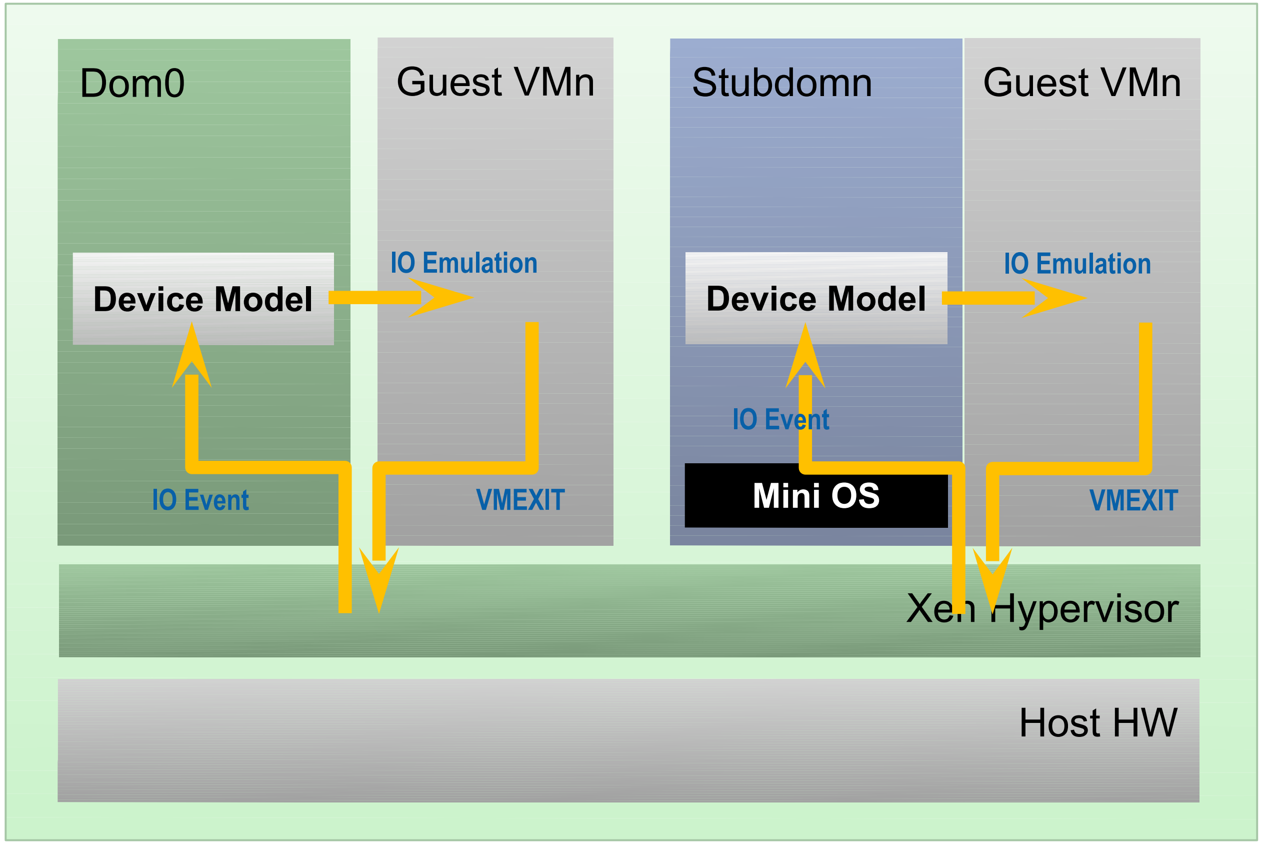
Advantages

- Fast
- Works on any system (even without virt extensions)

Driver Domains

- Security
- Isolation
- Reliability and Robustness

HVM & Stub Domains



Disadvantages

- Slower than PV due to Emulation (mainly I/O devices)

Advantages

- No kernel support needed

Stub Domains

- Security
- Isolation
- Reliability and Robustness

PV on HVM

- HVM guest with PV elements
- Linux enables as many PV interfaces as possible
- This has advantages
  - Install the same way as native
  - PC-like hardware
  - Access to fast PV devices
  - Exploit nested paging
  - Good performance trade-offs
- Drivers in Linux 3.x

	HVM	PV on HVM	PV
Boot Sequence	Emulated	Emulated	PV
Memory	HW	HW	PV
Interrupts, Timers & Spinlocks	Emulated	PV*	PV
Disk & Network	Emulated	PV	PV
Privileged Operations	HW	HW	PV

*) Emulated for Windows

PV in HVM Containers: Xen 4.3

- Salient Features
  - Dom0 runs in ring0
  - Event channel (no APIC)
  - Native page tables
  - Native IDT
- Fastest of PV and HVM
  - No need for emulation
  - Uses HW, where PV is slower than HVM
- Being up streamed now

More info …

NEW

	HVM	PV on HVM	PVH	PV
Boot Sequence	Emulated	Emulated	PV	PV
Memory	HW	HW	HW	PV
Interrupts, Timers & Spinlocks	Emulated	PV*	PV	PV
Disk & Network	Emulated	PV	PV	PV
Privileged Operations	HW	HW	HW	PV

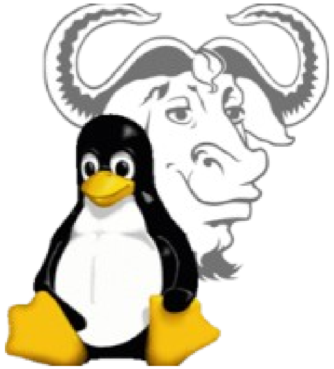
*) Emulated for Windows

Xen and Linux

Xen and the Linux Kernel

Xen was initially a University research project

Invasive changes to the kernel to run Linux as a PV guest and Dom0



Current State

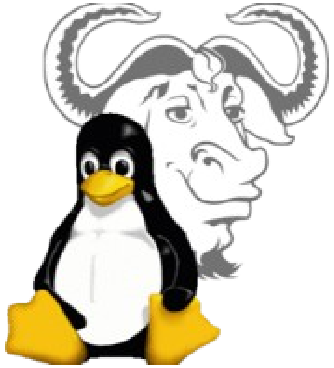
PVOPS Project

Xen support in Linux 3.0+

(it is functional – some optimizations missing)

On-going optimization work in Linux 3.6 +

Supporting new Xen 4.3 functionality (e.g. PVH, ARM)



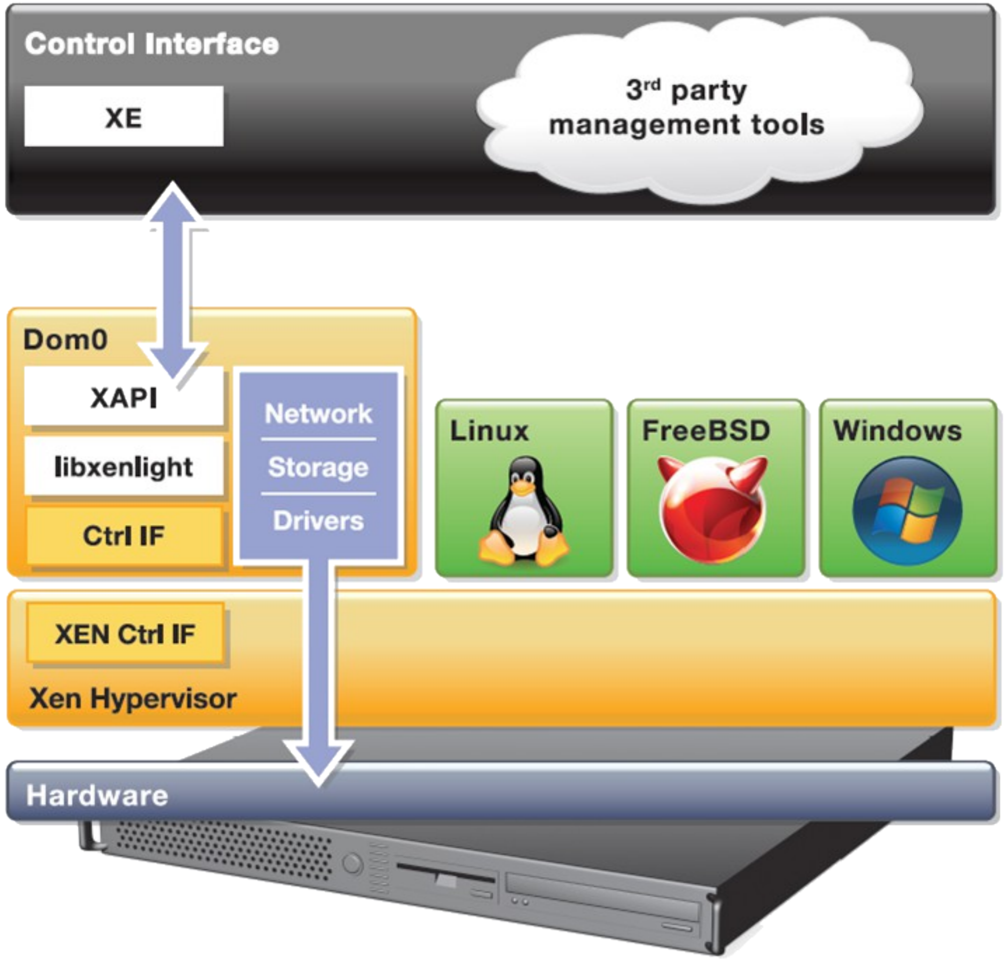
What does this mean?

- Xen Hypervisor is **not** in the Linux kernel
- **BUT**: everything Xen needs to run is!
- Xen packages are mostly in Linux distros
  - Install Dom0 Linux distro
  - Install Xen package(s) or meta package
  - Reboot
  - Config stuff: set up disks, peripherals, etc.

More info …

XCP Project

XCP – Xen Cloud Platform



- GPLv2
- XenServer is a commercial distro
- Complete vertical stack for server virtualization
- Distributed as
  - Appliance (ISO)
  - Packages in Debian & Ubuntu (more distros to come)

Major XCP Features

- VM lifecycle: live snapshots, checkpoint, migration
- Resource pools: flexible storage and networking
- Event tracking: progress, notification
- Upgrade and patching capabilities
- Real-time performance monitoring and alerting

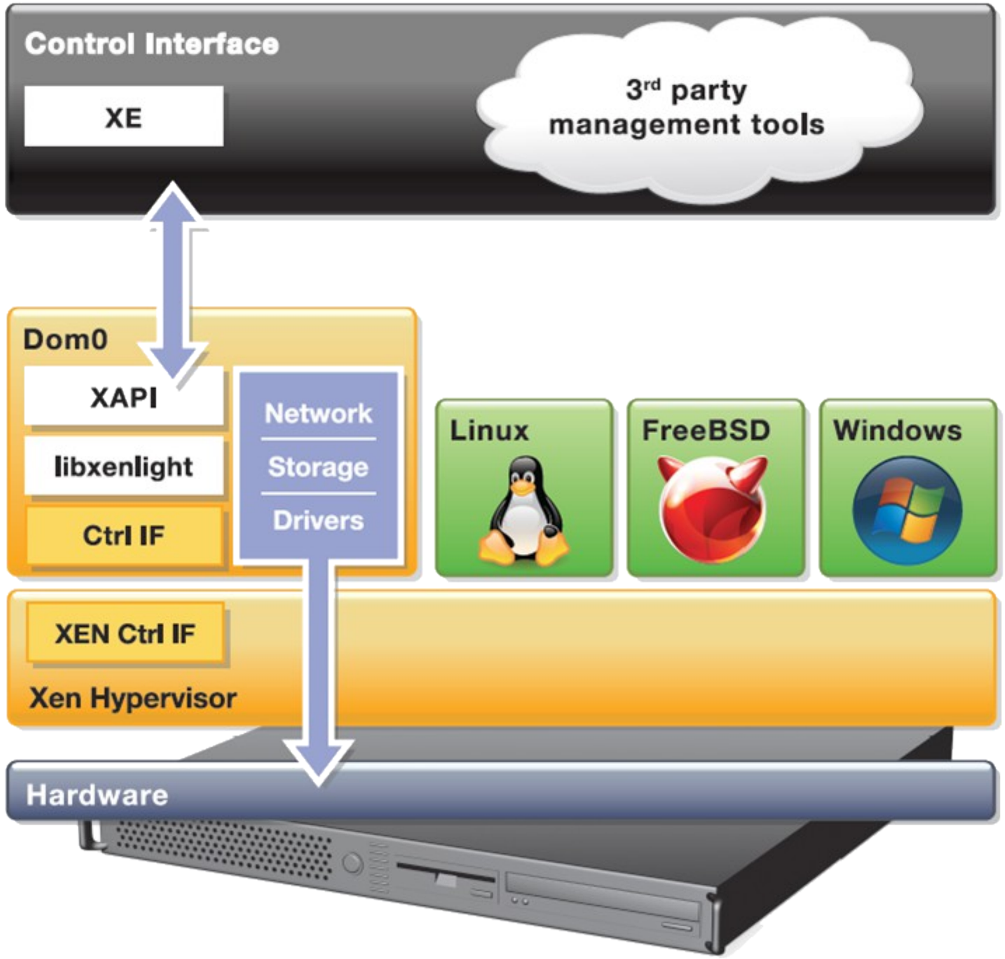
- Built-in support and templates for Windows and Linux guests
- Open vSwitch support built-in (default)

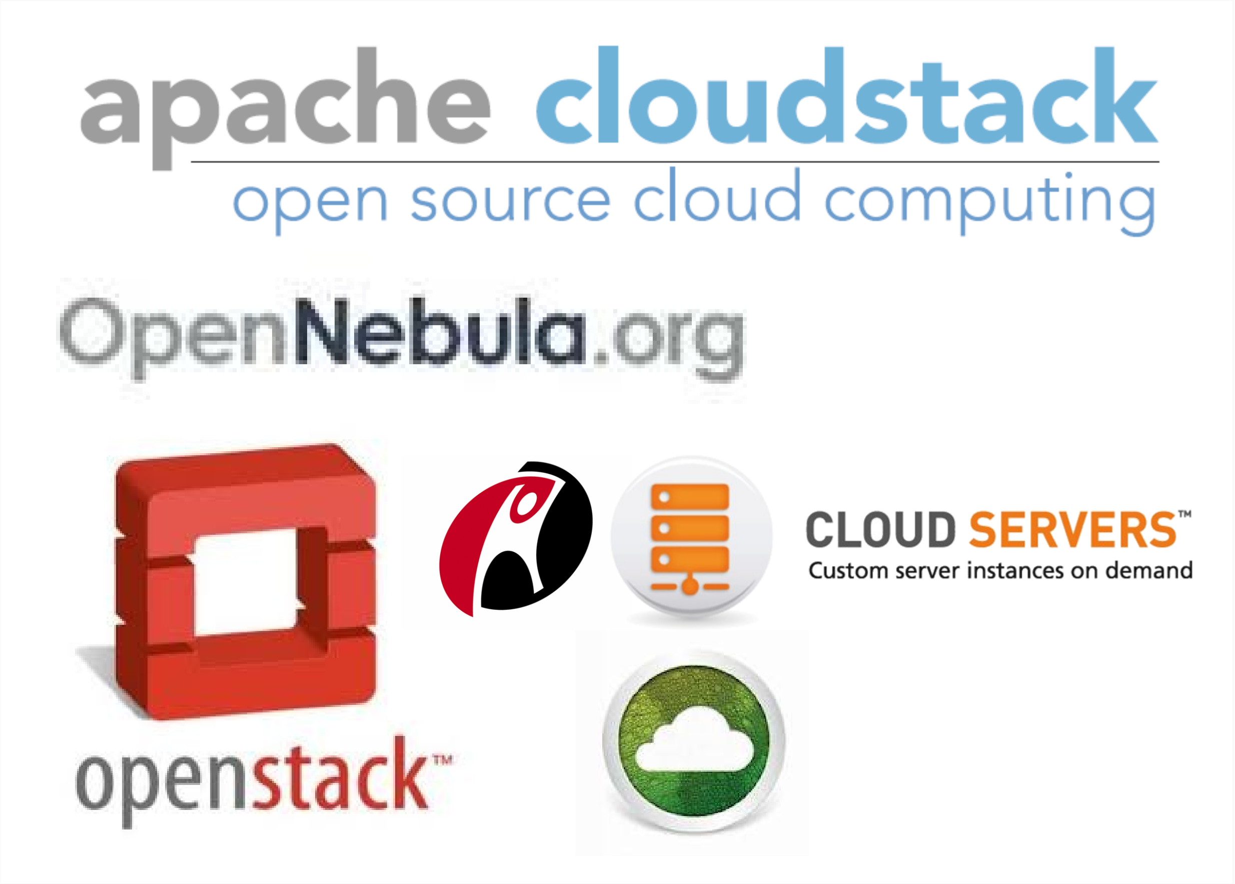
XCP 1.6 – to ship in Sep/Oct 12

- **Internal Improvements:**
  Xen 4.1.2, CentOS 5.7 with kernel 2.6.32.43, Open vSwitch 1.4.1
- **New format Windows drivers:** installable by Windows Update Service
- **Networking:** Better VLAN scalability, LACP bonding, IPv6
- **More guest OS templates:** Ubuntu Precise 12.04, RHEL/CentOS, Oracle Enterprise Linux 6.1 & 6.2, Windows 8
- **Storage XenMotion:**
  - Migrate VMs between hosts or pools without shared storage
  - Move a VM's disks between storage repositories while the VM is running

More Info …

XCP and Cloud Orchestration Stacks




apache cloudstack
open source cloud computing
OpenNebula.org
openstack™
CLOUD SERVERS™
Custom server instances on demand

Challenges for FOSS hypervisors

“Security and QoS/Reliability are amongst the top 3 blockers for cloud adoption”

www.colt.net/cio-research

Security and the Next Wave of Virtualization

- Security is a key requirement for Cloud
- Security is the primary goal of virtualization on the Client
  - Xen's advanced security features were developed for security sensitive Desktop use-cases (NSA)
- Maintaining isolation between VMs is critical (multi-tenancy)

Xen Security & Robustness Advantages

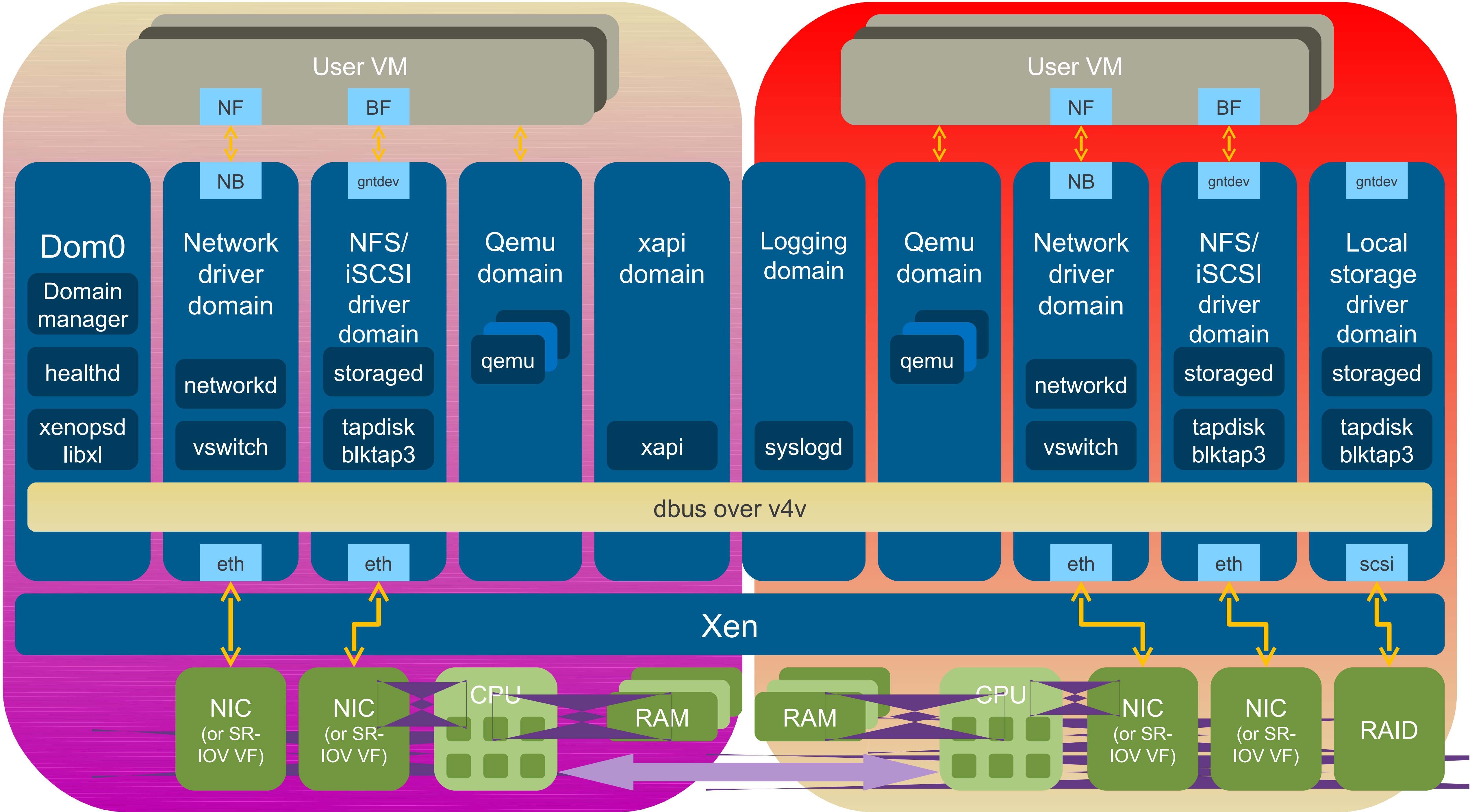
- Even without Advanced Security Features
  - Well-defined trusted computing base (much smaller than on type-2 HV)
  - Minimal services in hypervisor layer
- **More Robustness:** Mature, Tried & Tested, Architecture
- Xen Security Modules (or XSM)
  - Developed, maintained and contributed to Xen by NSA
  - Generalized Security Framework for Xen
  - Compatible with SELinux (tools, architecture)
  - XSM object classes maps onto Xen features

Advanced Security: Disaggregation

- Split Control Domain into Driver, Stub and Service Domains
  - Each contains a specific set of control logic
  - See: ”Breaking up is hard to do” @ Xen Papers
  - See: “Domain 0 Disaggregation for XCP and XenServer”

- Unique benefit of the Xen architecture
  - **Security**: Minimum privilege; Narrow interfaces
  - **Robustness:** ability to safely restart parts of the system
  - **Performance:** lightweight, e.g. Mini OS directly on hypervisor
  - **Scalability:** more distributed system (less reliable on Dom0)

- Used today by Qubes OS and Citrix XenClient XT
- Soon for XCP and XenServer



User VM

NF

BF

NB

gntdev

Dom0

Domain manager

healthd

xenopsd libxl

Network driver domain

networkd

vswitch

eth

NFS/ iSCSI driver domain

storaged

tapdisk blktap3

eth

Qemu domain

qemu

xapi domain

xapi

Logging domain

syslogd

Qemu domain

qemu

Network driver domain

networkd

vswitch

eth

NFS/ iSCSI driver domain

storaged

tapdisk blktap3

eth

Local storage driver domain

storaged

tapdisk blktap3

scsi

dbus over v4v

Xen

NIC (or SR-IOV VF)

NIC (or SR-IOV VF)

CPU

RAM

RAM

CPU

NIC (or SR-IOV VF)

NIC (or SR-IOV VF)

RAID

News from the Xen Community

Cool new functionality & initiatives

- Xen for ARM using HW virt (using new PVH mode)
  - Started our first guest domain, including PV console disk and network devices!
  - No emulation (QEMU is needed)
- New PVH virtualization mode (Oracle)
- FreeBSD Xen port (SpectraLogic & HP)
- Xen MIPS port (by BroadCom)
- Language run-times running on bare-metal Xen
  - Openmirage.org, ErlangOnXen.org
- Disaggregation is moving from Client into Server and Cloud
- Portable Service VMs
  - Agree interface and mechanism to allow service VMs across products and hosting services

Summary: Why Xen?

- Designed for the Cloud : many advantages for cloud use!
  - Resilience, Robustness & Scalability
  - Security: Small surface of attack, Isolation & Advanced Security Features
- Widely used by Cloud Providers and Vendors
- XCP
  - Ready for use with cloud orchestration stacks
  - Packages in Linux distros: flexibility and choice
- Open Source with a large community and eco-system
  - Exciting new developments in the pipeline

- **IRC:** ##xen @ FREENODE
- **Mailing List:** xen-users & xen-api
- **Wiki:** wiki.xen.org
- **Excellent XCP Tutorials**
  - A day worth of material @ xen.org/community/xenday11
- **Ecosystem pages**
- **Presentations:** slideshare.net/xen_com_mgr
- **Videos:** vimeo.com/channels/xen

Questions …

Followme!

@lars_kurth
@xen_com_mgr

FREENODE: lars_kurth