Virtualization in the Cloud: Featuring Xen

Lars Kurth
Xen Community Manager
lars.kurth@xen.org

FREENODE: lars_kurth

@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.

Global Public Computing

"*This dissertation proposes a new **distributed computing paradigm**, termed global public computing, which allows*

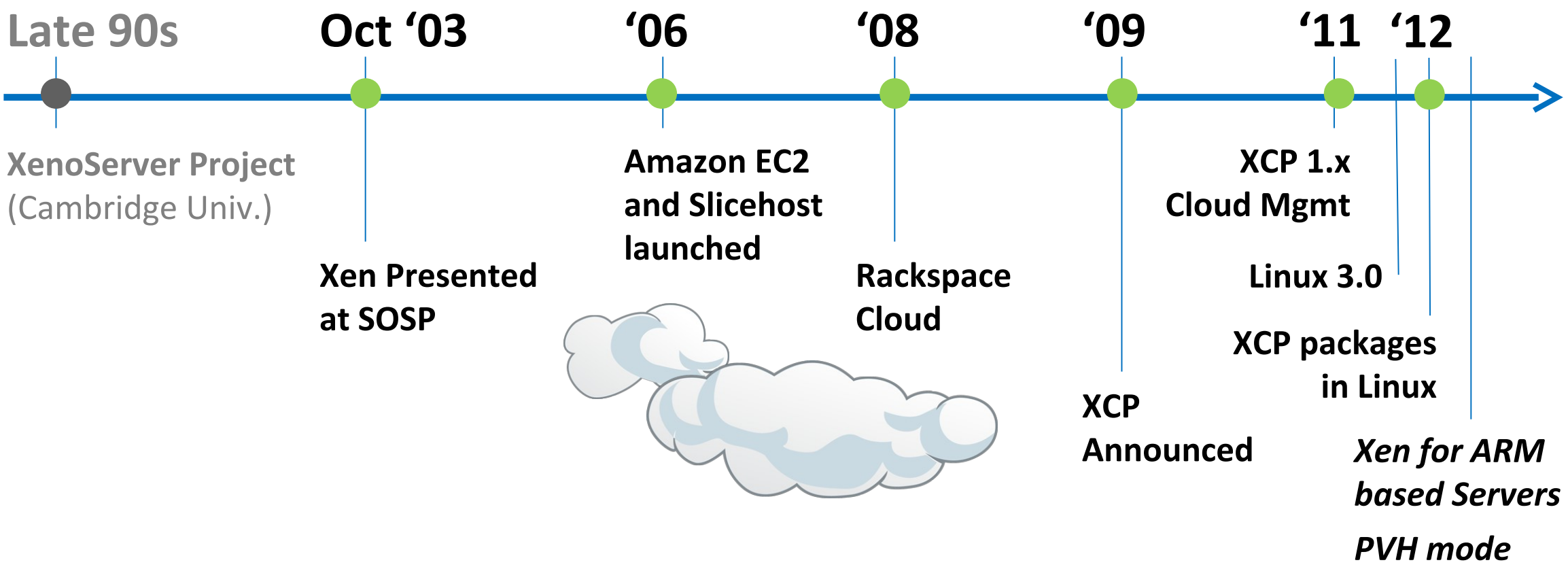
*any user to run any code anywhere. Such platforms **price***

***computing resources**, and ultimately **charge users for resources** consumed.*"

Evangelos Kotsovinos, PhD dissertation, 2004

Xen™

A Brief History of Xen in the Cloud



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

Xen.org

- Guardian of Xen Hypervisor and related OSS Projects
- Xen Governance similar to Linux Kernel
- Projects
  - Xen Hypervisor (led by 5 committers, 2 from 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, …)

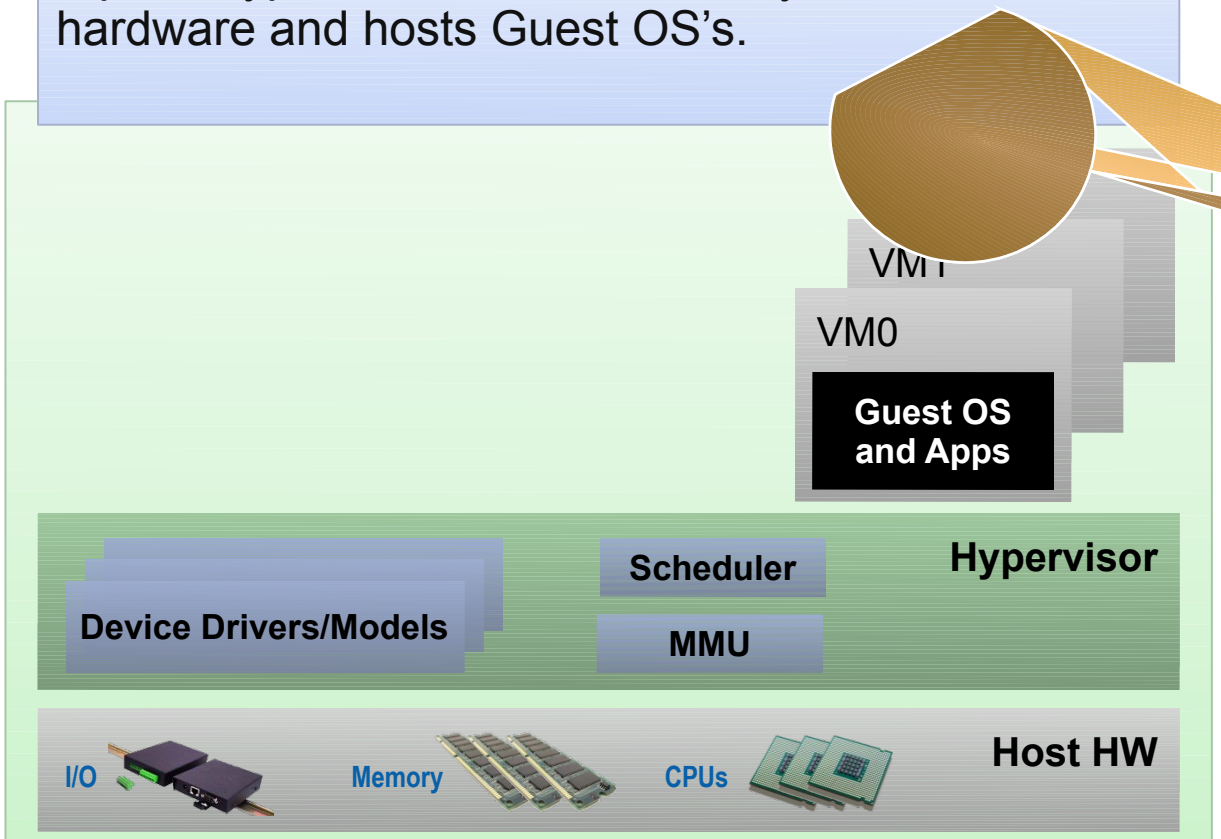
Xen™

Xen Overview

Hypervisor Architectures

Type 1: Bare metal Hypervisor

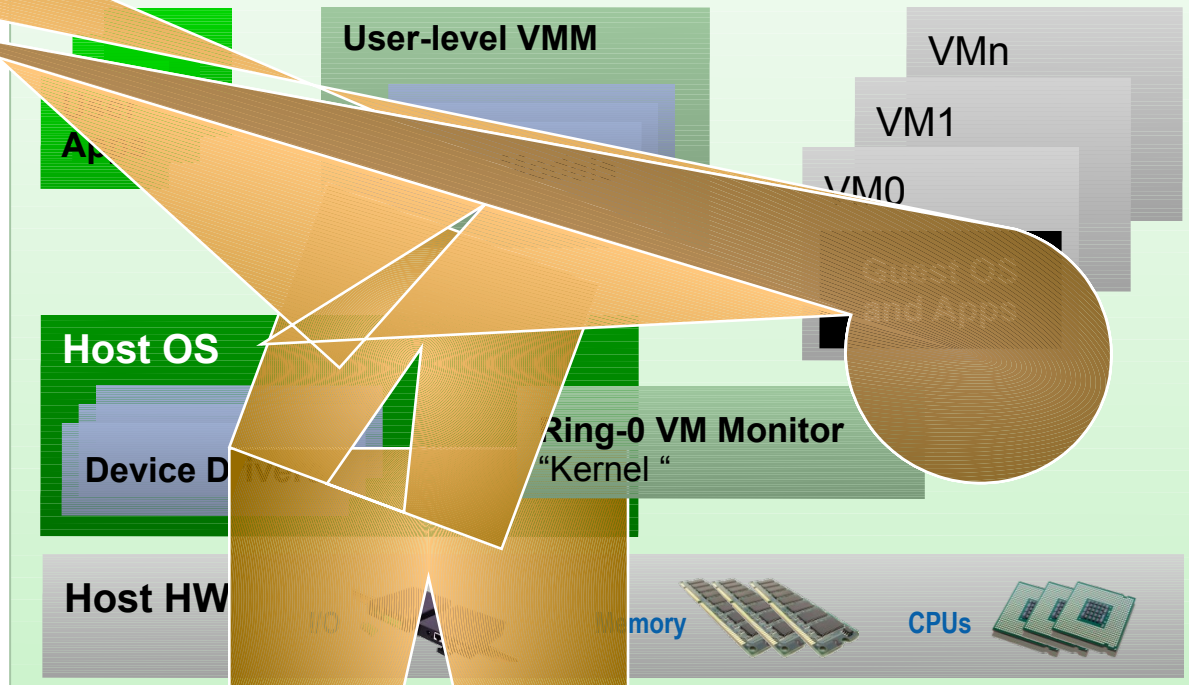
A pure Hypervisor that runs directly on the hardware and hosts Guest OS's.



Provides partition isolation + reliability, higher security

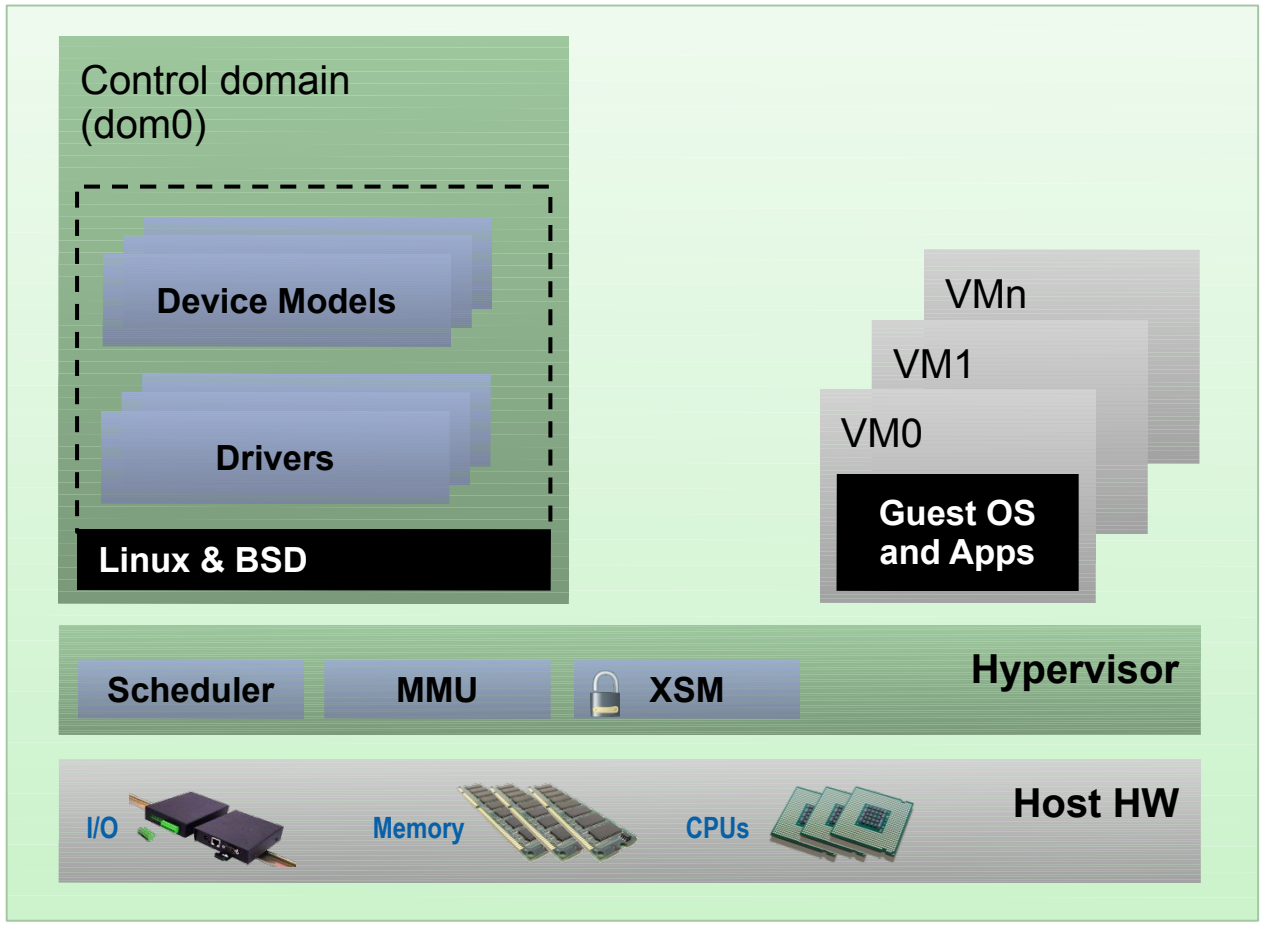
Type 2: OS 'Hosted'

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

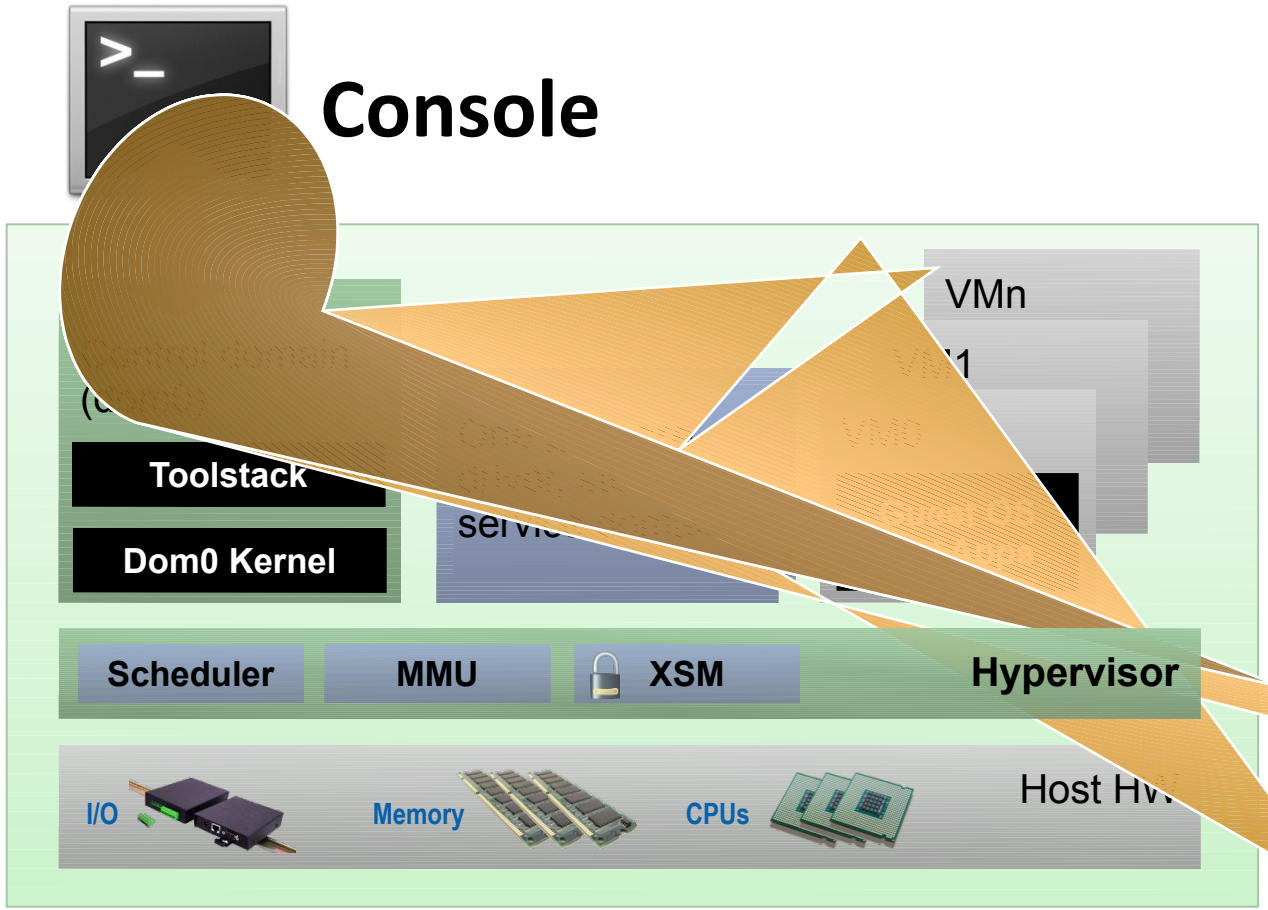
- Re-use of Dom0 kernel components
- Ease of use & Installation
- Isolation & Security

Xen and Linux

- Xen Hypervisor is **not** in the Linux kernel
- **BUT**: everything Xen and Xen Guests need to run is!
- Xen packages are in all Linux distros (except RHEL6)
  - Install Dom0 Linux distro
  - Install Xen package(s) or meta package
  - Reboot
  - Config stuff: set up disks, peripherals, etc.

More info: wiki.xen.org/wiki/Category:Host_Install

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”
- -privileged and isolated
- Lifetime: start, stop, kill

Xen Variants for Server & Cloud

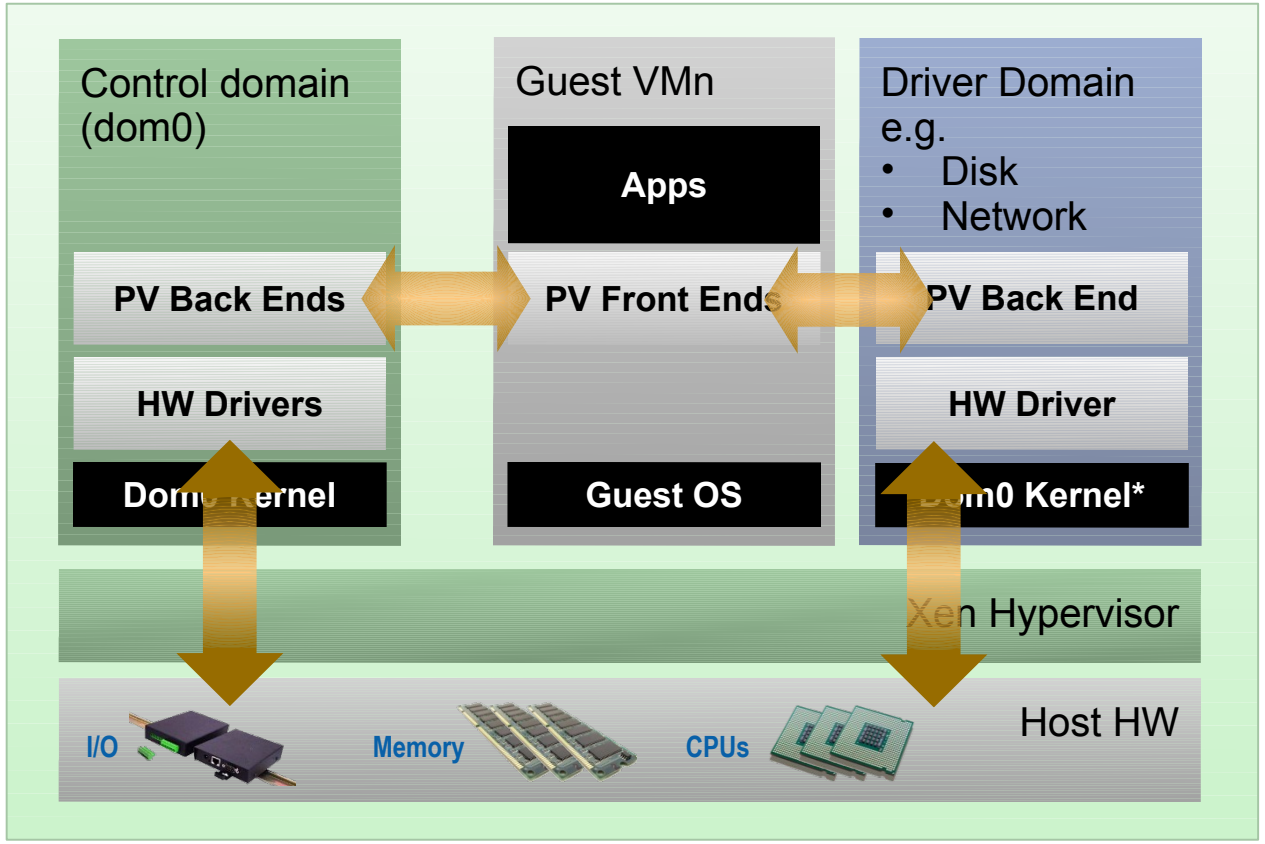
	Xen Hypervisor		XCP
Toolstack / Console	**Default / XL (XM)**	**Libvirt / VIRSH**	**XAPI / XE**
	Increased level of functionality and integration with other components		
Get Binaries from …	**Linux Distros**	**Linux Distros**	**Debian & Ubuntu** **ISO from Xen.org**
Products	**Oracle VM**	**Huawei UVP**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

More info: xen.org/community/presentations.html

Xen™

Xen : Types of Virtualization

PV Domains & Driver Domains



*) Can be MiniOS

Technology:

- Paravirtualization

Linux PV guests have limitations:

- limited set of virtual hardware

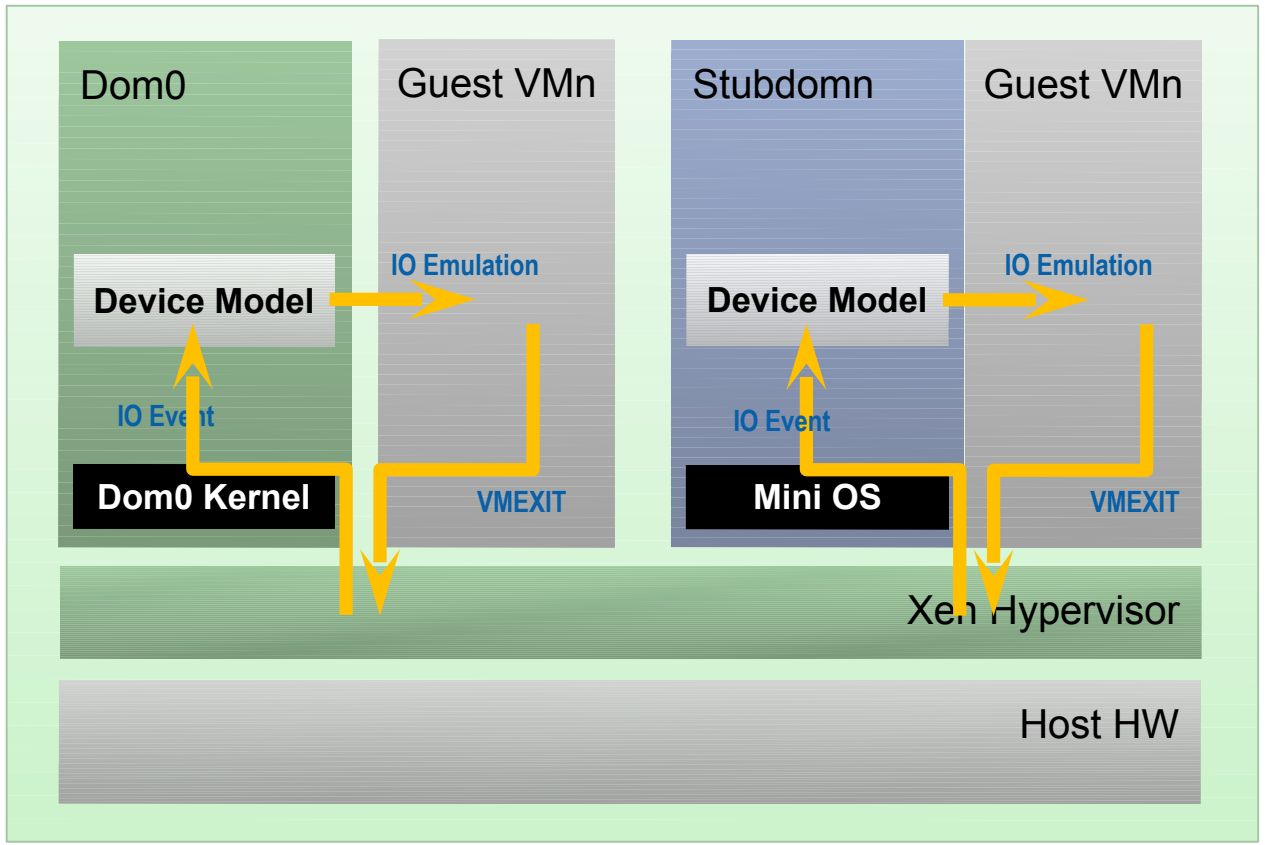
Advantages

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

Driver Domains

- Security
- Isolation
- Reliability and Robustness

HVM & Stub Domains



Technology:

- Shows emulation using QEMU/Device Model (SW Virtualization)
- In other situation HW can be used

Disadvantages

- Emulation slower than PV (mainly I/O devices)

Advantages

- No kernel support needed

Stub Domains

- Security
- Isolation
- Reliability and Robustness

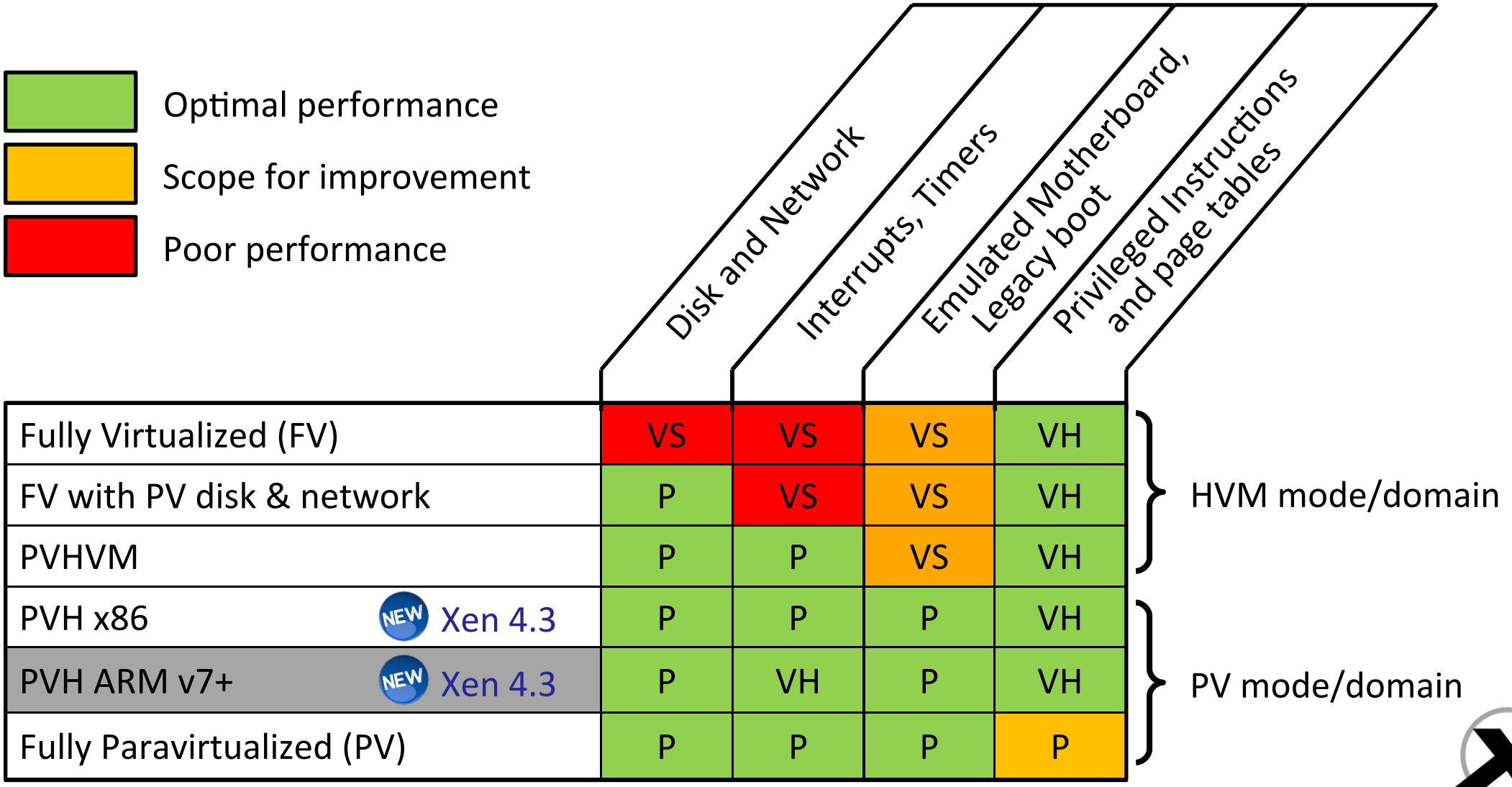
The Virtualization Spectrum

VS	Virtualized (SW)
VH	Virtualized (HW)
P	Paravirtualized

	Disk and Network	Interrupts, Timers	Emulated Motherboard, Legacy boot	Privileged Instructions and page tables	
Fully Virtualized (FV)	VS	VS	VS	VH	HVM mode/domain
FV with PV disk & network	P	VS	VS	VH	
PVHVM	P	P	VS	VH	
PVH x86 NEW Xen 4.3	P	P	P	VH	PV mode/domain
PVH ARM v7+ NEW Xen 4.3	P	VH	P	VH	
Fully Paravirtualized (PV)	P	P	P	P	

Xen™

The Virtualization Spectrum



PVH Benefits

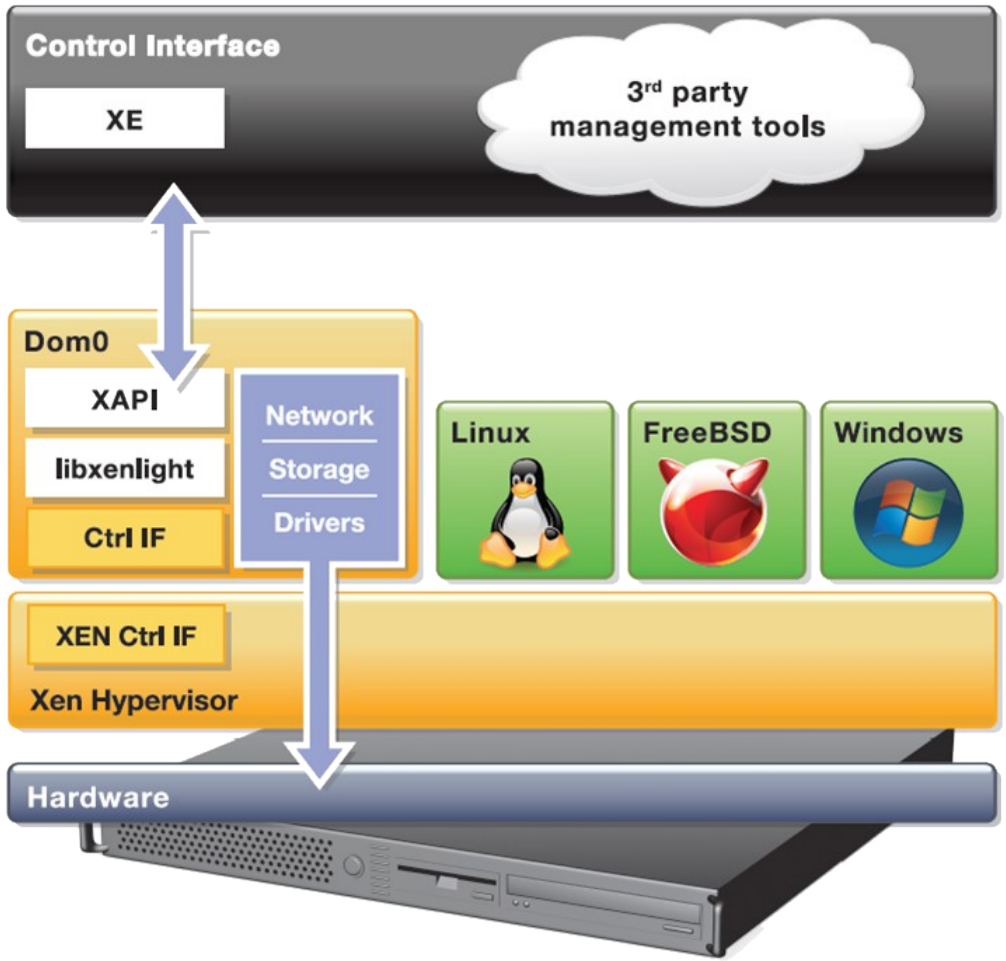
- Solves a number of historical problems with PV and HVM
  - AMD 64 bit and x86-64 architecture is not a good match for PV for Privileged Instructions and Page Tables
  - Will allow to simplify the Xen and PVOPS architecture in the longer term
- Fastest of PV and HVM on all architectures
  - No need for emulation
  - Uses HW virtualization where it is fastest
  - Uses PV where PV is fastest
  - Should provide the best trade-offs for most work-loads

More info: wiki.xen.org/wiki/Virtualization_Spectrum &
xen.org/xensummit/xs12na_talks/M9.html

Xen™

XCP Project

XCP – Xen Cloud Platform



Complete stack for server virtualization

- Extends Xen to cover multiple hosts
- Adds further functionality and integrations for cloud, storage and networking to Xen HV
- GPLv2
- XenServer is a commercial XCP distro

Two Flavours

- Appliance (ISO using CentOS Dom0)
- 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)

More info: wiki.xen.org/wiki/XCP_Release_Features

Xen™

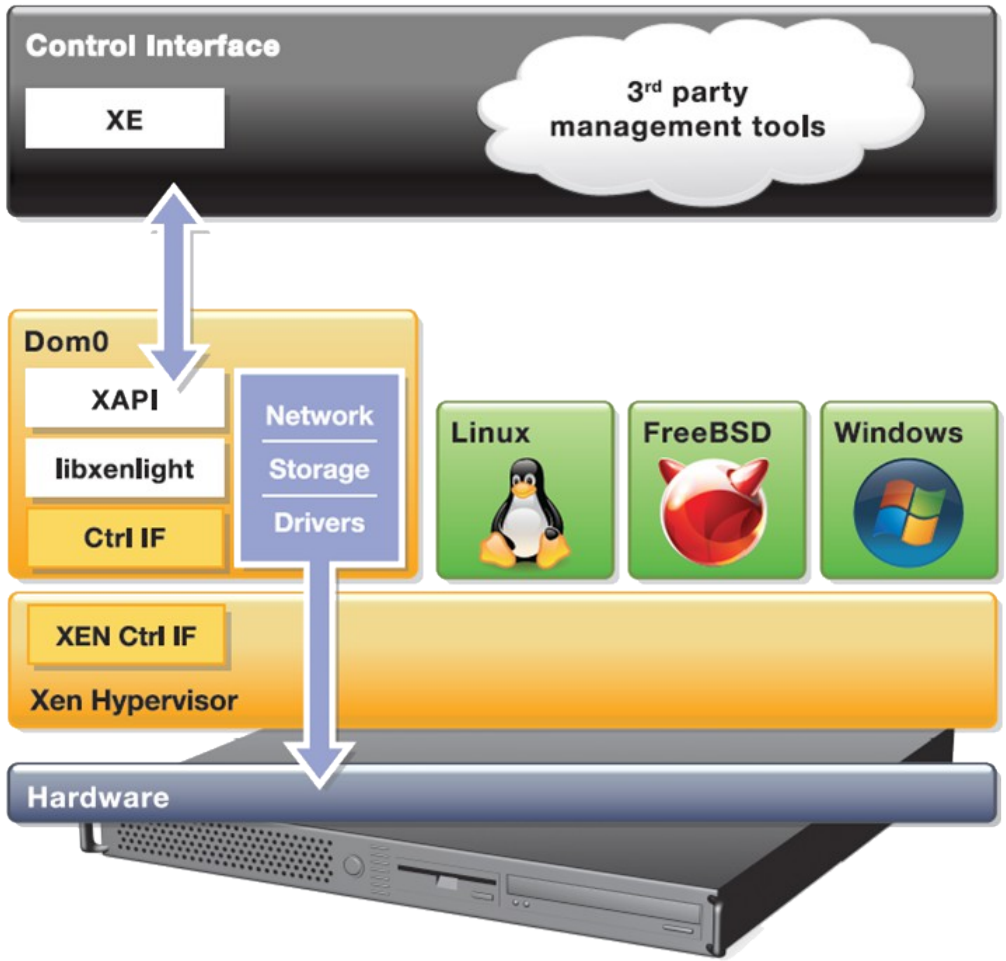
XCP 1.6 Beta

- **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: xen.org/download/xcp/releasenotes_1.6.0.html &
xen.org/download/xcp/index_1.6.0.html

Xen™

XCP and Cloud Orchestration Stacks



apache cloudstack
open source cloud computing

OpenNebula.org

openstack™

Xen™

Challenges for FOSS hypervisors

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

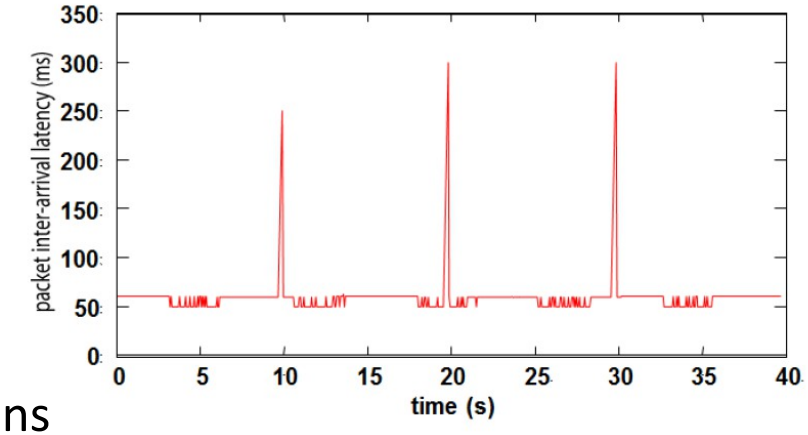
www.colt.net/cio-research

System characteristics cloud users care about: “Robustness, Performance, Scalability & Security”

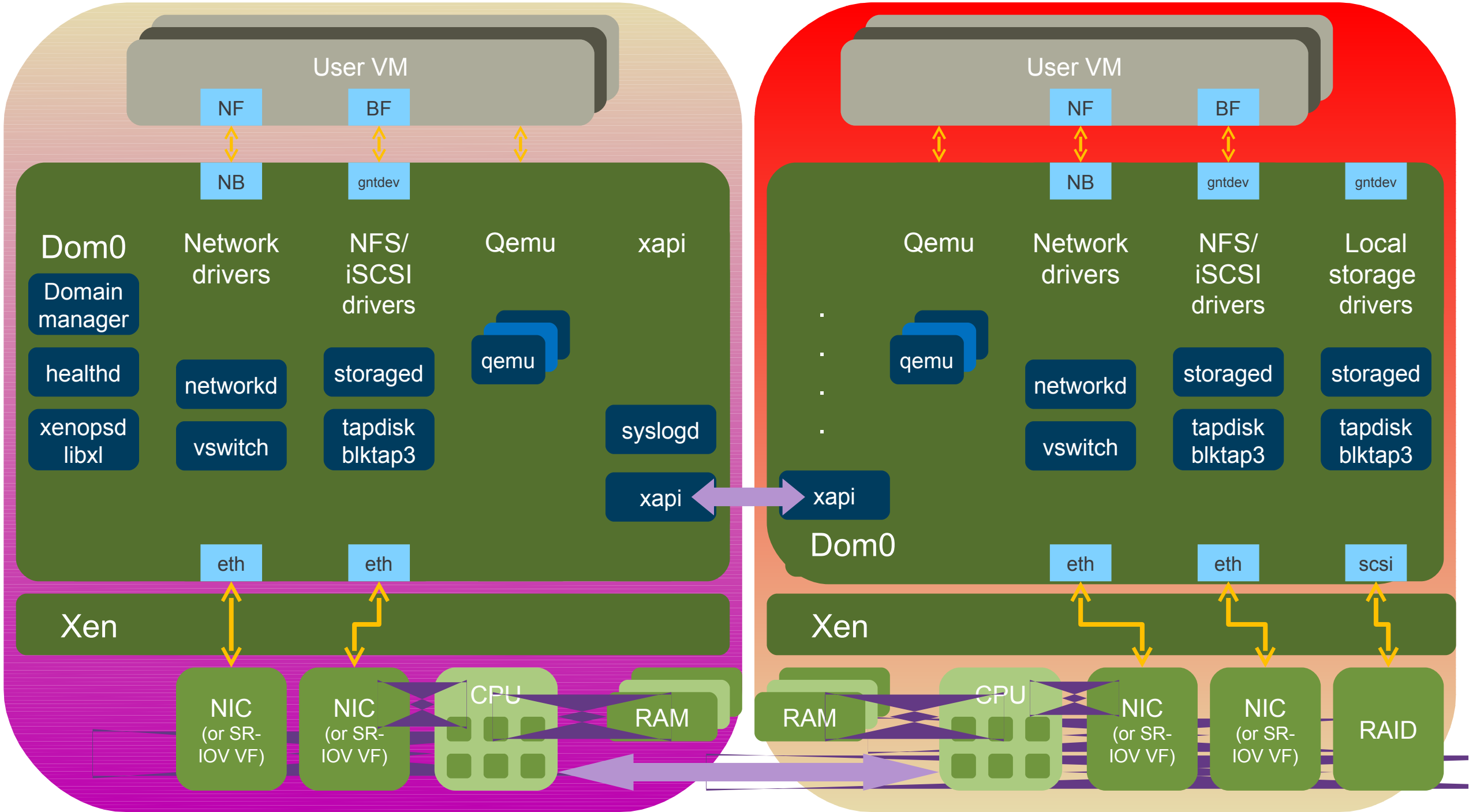
Results XCP User Survey 2012 – 90% of users quoted these as most important attributes

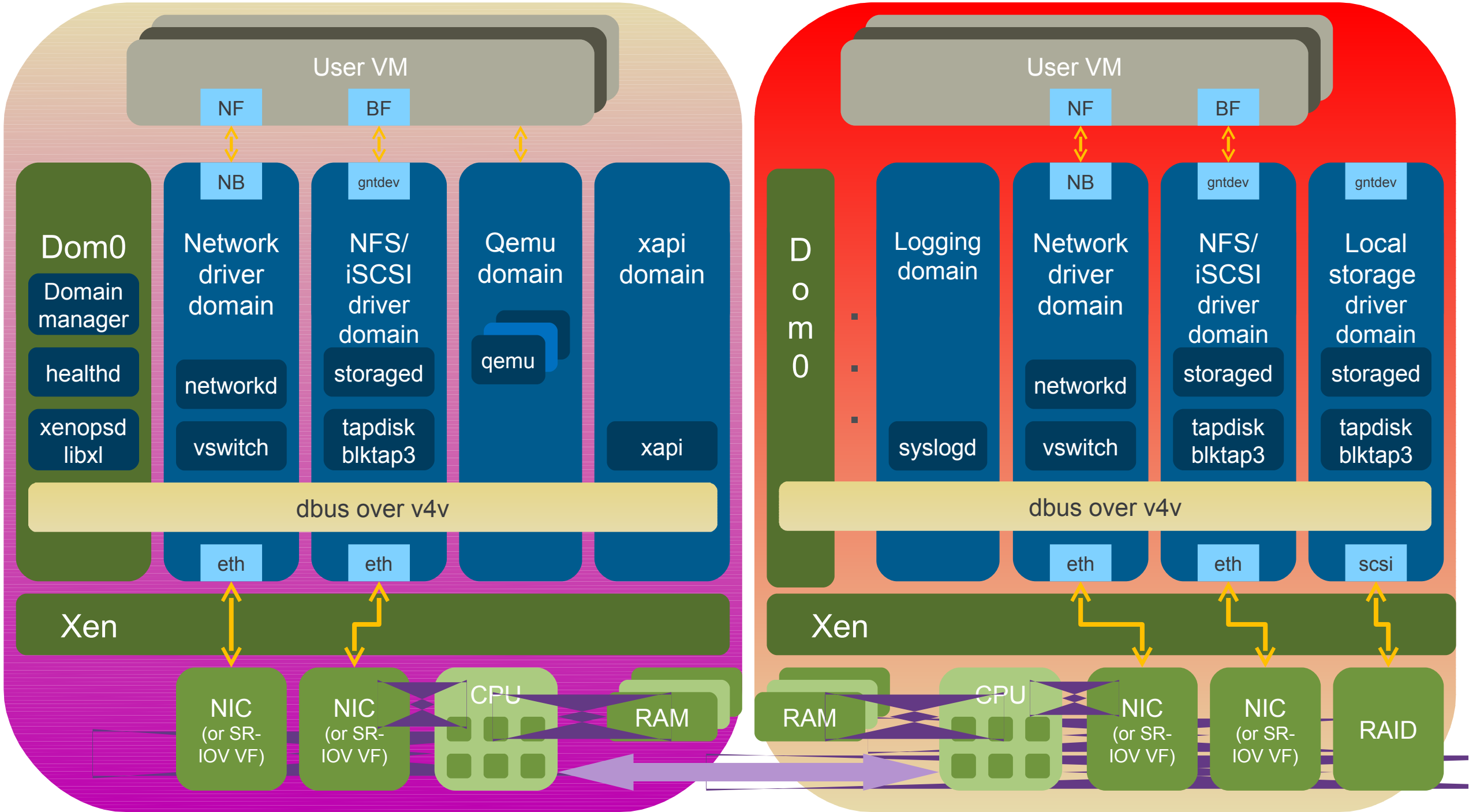
Disaggregation

- Split Control Domain into Driver, Stub and Service Domains
  - See: ”Breaking up is hard to do” @ Xen Papers
  - See: “Domain 0 Disaggregation for XCP and XenServer”
- Unique benefit of the Xen architecture
  - **Robustness:** ability to safely restart parts of the system (e.g. just 275ms outage from failed Ethernet driver)
  - **Performance:** lightweight, Xen scheduler
  - **Scalability:** more distributed system (less reliable on Dom0)
  - **Security**: Minimum privilege; Narrow interfaces; Restart domains



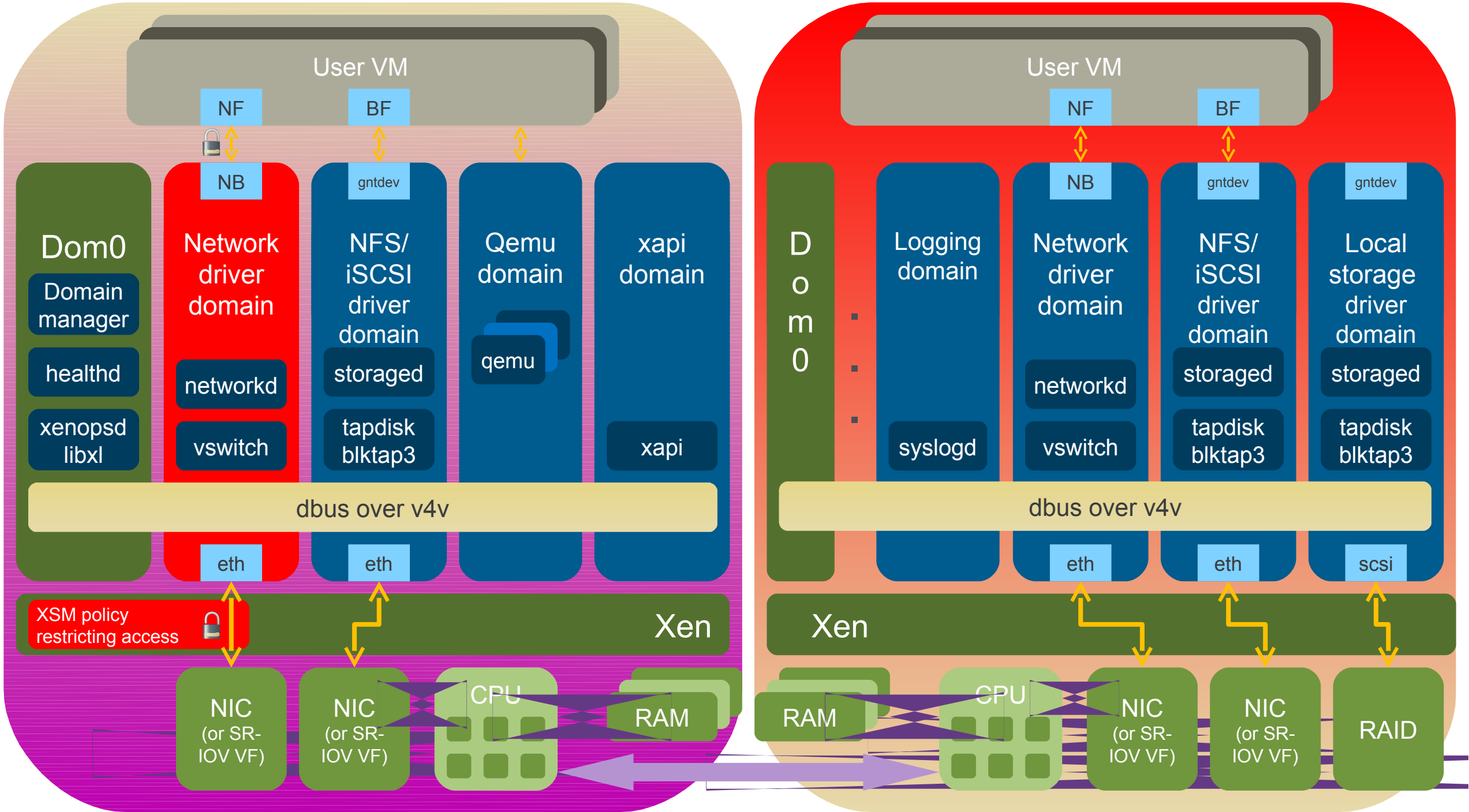
- Used today by Qubes OS and Citrix XenClient XT
- Prototypes for XCP and XenServer


User VM
NF
BF
NB
gntdev
Dom0
Network drivers
NFS/ iSCSI drivers
Qemu
xapi
Domain manager
healthd
xenopsd libxl
networkd
vswitch
storaged
tapdisk blktap3
qemu
syslogd
xapi
eth
eth
Xen
NIC (or SR-IOV VF)
NIC (or SR-IOV VF)
CPU
RAM
User VM
NF
BF
NB
gntdev
gntdev
Qemu
Network drivers
NFS/ iSCSI drivers
Local storage drivers
qemu
networkd
vswitch
storaged
tapdisk blktap3
storaged
tapdisk blktap3
xapi
Dom0
eth
eth
scsi
Xen
RAM
CPU
NIC (or SR-IOV VF)
NIC (or SR-IOV VF)
RAID


User VM
NF
BF
NB
gntdev
Dom0
Domain manager
healthd
xenopsd libxl
Network driver domain
networkd
vswitch
NFS/ iSCSI driver domain
storaged
tapdisk blktap3
Qemu domain
qemu
xapi domain
xapi
dbus over v4v
eth
Xen
NIC (or SR-IOV VF)
CPU
RAM
User VM
NF
BF
NB
gntdev
gntdev
Dom0
Logging domain
syslogd
Network driver domain
networkd
vswitch
NFS/ iSCSI driver domain
storaged
tapdisk blktap3
Local storage driver domain
storaged
tapdisk blktap3
dbus over v4v
eth
eth
scsi
Xen
RAM
CPU
NIC (or SR-IOV VF)
NIC (or SR-IOV VF)
RAID

Xen Security Advantages

- Even without Advanced Security Features
  - Well-defined trusted computing base (much smaller than on type-2 HV)
  - Minimal services in hypervisor layer
- 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
- XSM together with Disaggregation
  - Security sensitive Desktop use-cases developed for the NSA


User VM
NF
BF
NB
gntdev
Dom0
Domain manager
healthd
xenopsd libxl
Network driver domain
networkd
vswitch
NFS/ iSCSI driver domain
storaged
tapdisk blktap3
Qemu domain
qemu
xapi domain
xapi
dbus over v4v
eth
XSM policy restricting access
Xen
NIC (or SR-IOV VF)
CPU
RAM
User VM
NF
BF
NB
gntdev
gntdev
Dom0
Logging domain
syslogd
Network driver domain
networkd
vswitch
NFS/ iSCSI driver domain
storaged
tapdisk blktap3
Local storage driver domain
storaged
tapdisk blktap3
dbus over v4v
eth
eth
scsi
Xen
RAM
CPU
NIC (or SR-IOV VF)
NIC (or SR-IOV VF)
RAID

News from the Xen Community

Cool new functionality & initiatives

- New PVH virtualization mode (Oracle)
  - Patches being up-streamed to Xen and Linux (3.7 & 3.8) as we speak
  - Sweet spot for performance
- Xen for ARM servers (using new PVH mode)
  - Cortex A15/ ARM v7: can start guests on Versatile Express; Samsung Chromebook next
  - ARM v8: porting work started on simulator and patches being up-streamed
- New Xen ports
  - FreeBSD Xen port (SpectraLogic & HP)
  - Xen MIPS port (by BroadCom)
- Language run-times running on bare-metal Xen
  - ErlangOnXen.org , Openmirage.org

More info: wiki.xen.org/wiki/Xen_Roadmap/4.3 & wiki.xen.org/wiki/XCP_Roadmap

Xen™

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 (lists.xen.org)
- **Wiki:** wiki.xen.org
- **Ecosystem pages:** xen.org/community/ecosystem.html
- **Presentations & Videos:** xen.org/community/presentations.html

Questions …

@lars_kurth
@xen_com_mgr

FREENODE: lars_kurth