xen.org

Xen Cloud Platform

Lars Kurth

Xen Community Manager

lars.kurth@xen.org

@lars_kurth

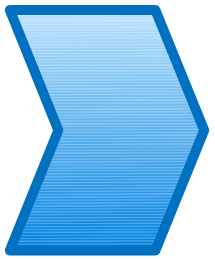
@xen_com_mgr

A Brief History of Xen in the Cloud

Late 90s

XenoServer Project
(Cambridge Univ.)

The **XenoServer project** is building a *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.



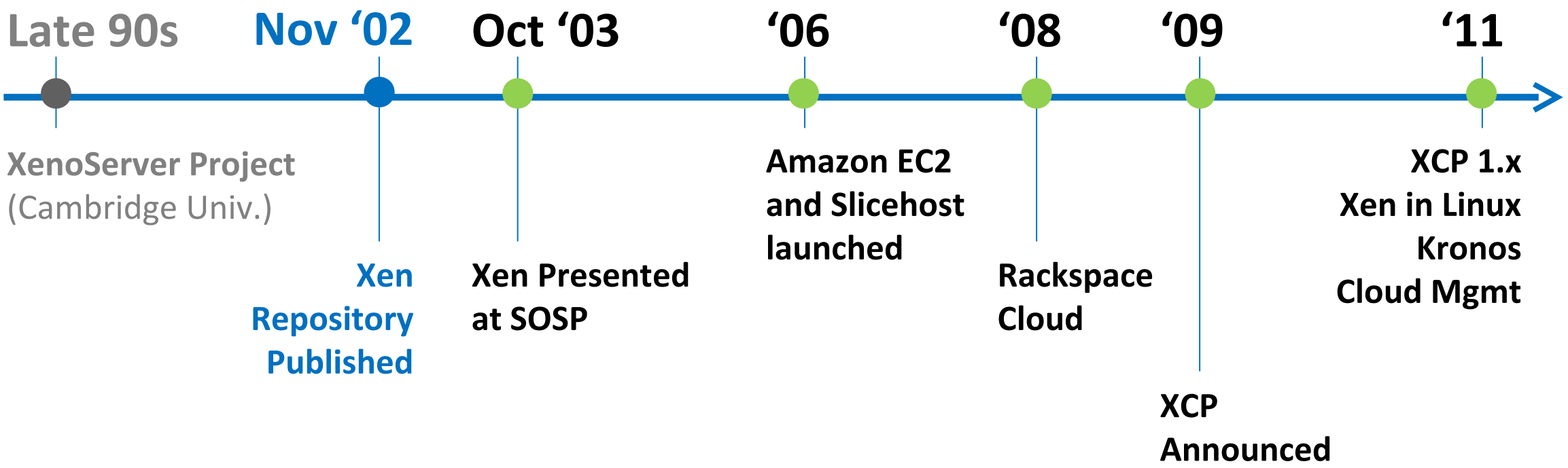
Global Public Computing

"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

A Brief History of Xen in the Cloud



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

xen.org

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 (led by Samsung)
  - PVOPS : Xen components and support in Linux Kernel (led by Oracle)

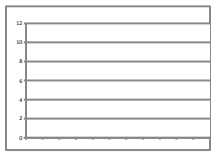
The Xen Community

xen.org

Xen

Xen Contributions & Vendors

By Change Sets *)



2011 Contributions by KLOC **)

*) Does not count activity on XenARM
(as not yet in an official repo)

*) Activity on Development branch (not yet in xen-unstable)
**) Includes PVOPS
***) Figures up to end of Q3 2011

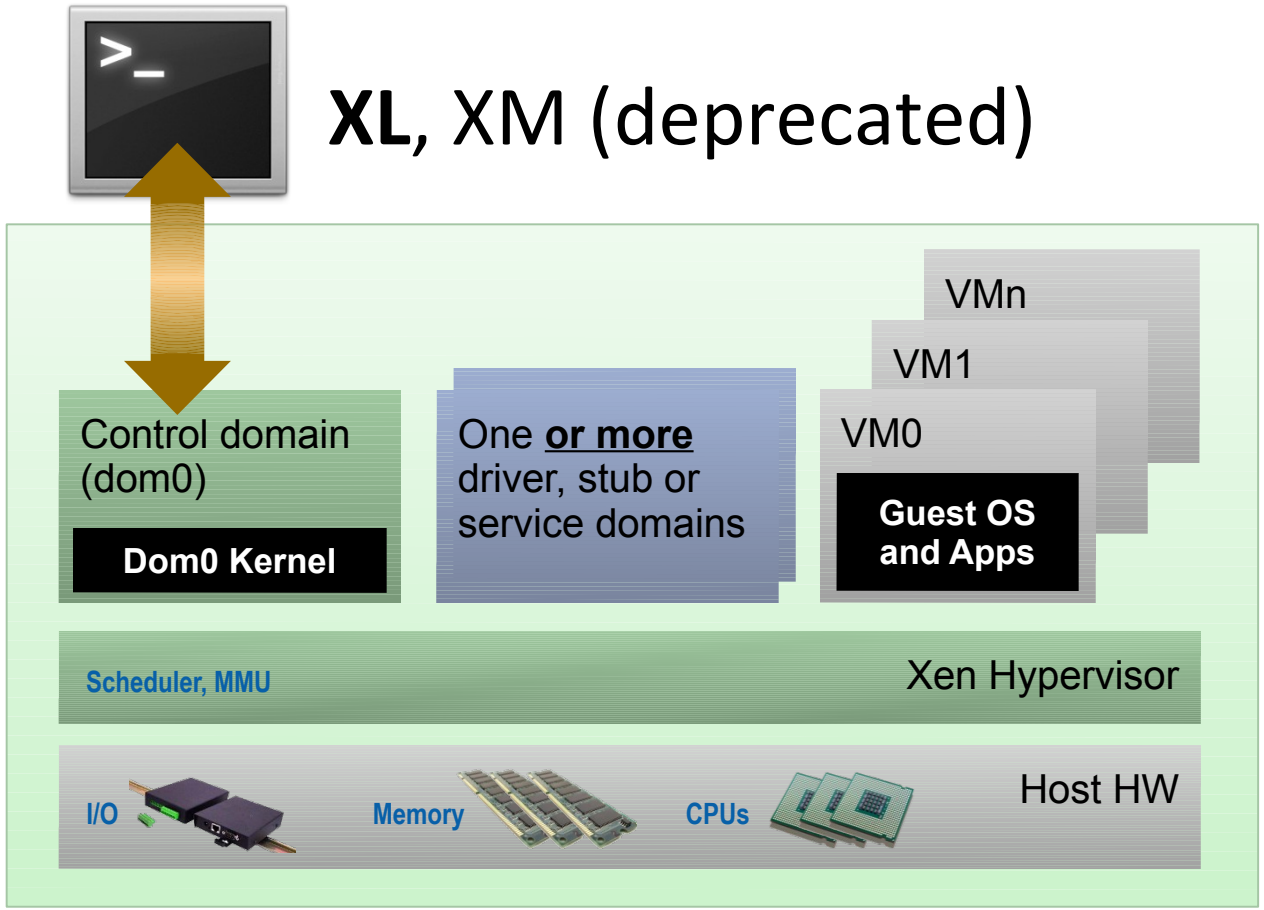
Community & Ecosystem Map

xen.org/community/projects

openSUSE
NetBSD
debian
Eucalyptus
OpenNebula.org
cloudstack
open source cloud computing
openstack
CITRIX
suse
ORACLE
NETLOGIC MICROSYSTEMS
HUAWEI
linode.com
amazon web services

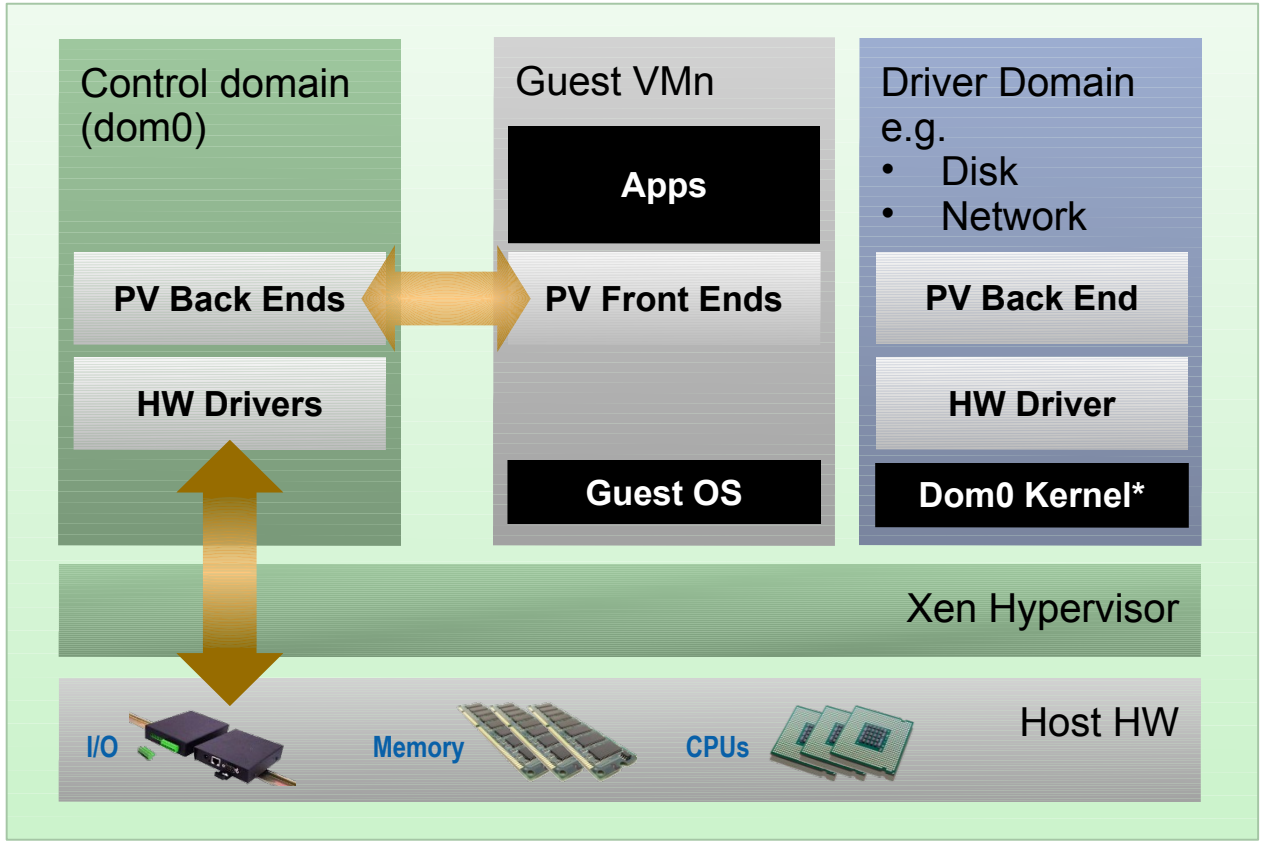
Xen Overview

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
- Driver/Stub/Service Domain(s)
  - A “driver, device model or control service in a box”
  - De-privileged and isolated
  - Lifetime: start, stop, kill

PV Domains & Driver Domains



*) Can be MiniOS

Linux PV guests have limitations:

- limited set of virtual hardware

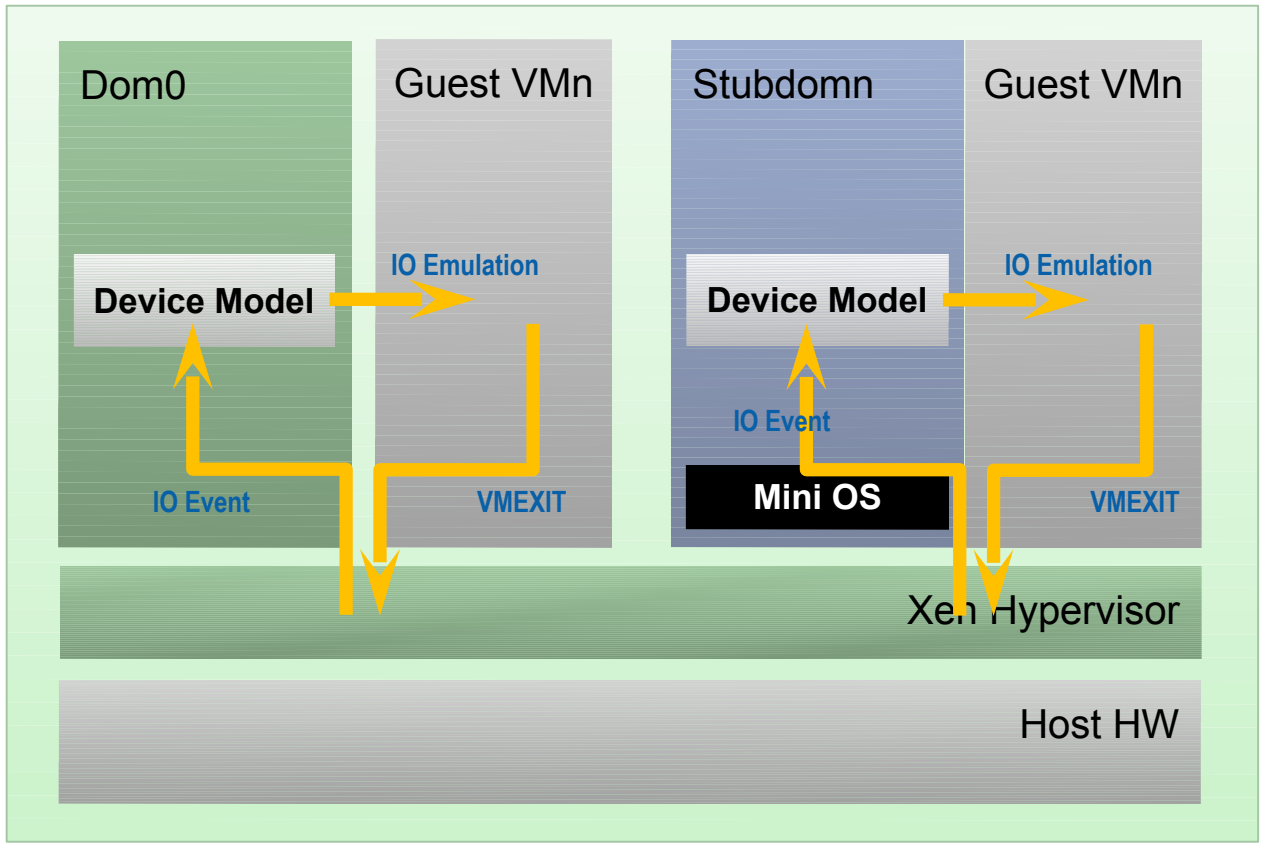
Advantages

- 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

- Install the same way as native Linux

Stub Domains

- Security
- Isolation
- Reliability and Robustness

PV on HVM

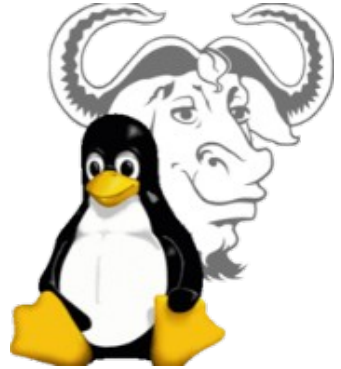
- A mixture of PV and HVM
- Linux enables as many PV interfaces as possible
- This has advantages
  - install the same way as native
  - PC-like hardware
  - access to fast PV devices
  - exploit nested paging
  - Good performance trade-offs
- Drivers in Linux 3.x

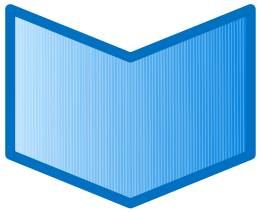
	HVM	PV on HVM	PV
Boot Sequence	Emulated	Emulated	PV
Memory	HW	HW	PV
Interrupts, Timers & Spinlocks	Emulated	PV*	PV
Disk & Network	Emulated	PV	PV
Privileged Operations	HW	HW	PV

*) Emulated for Windows

Xen and the Linux Kernel

Xen was initially a University research project



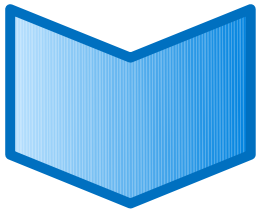


Invasive changes to the kernel to run Linux as a PV guest

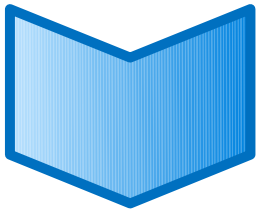
Even more changes to run Linux as dom0

Xen and the Linux Kernel

Xen support in the Linux kernel not upstream



Great maintenance effort on distributions

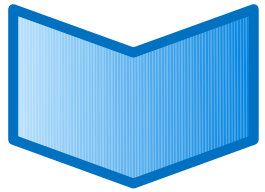


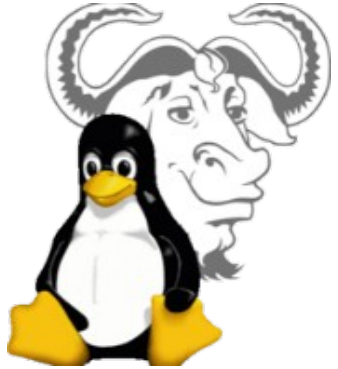
Risk of distributions dropping Xen support

Xen harder to use

Current State

PVOPS Project





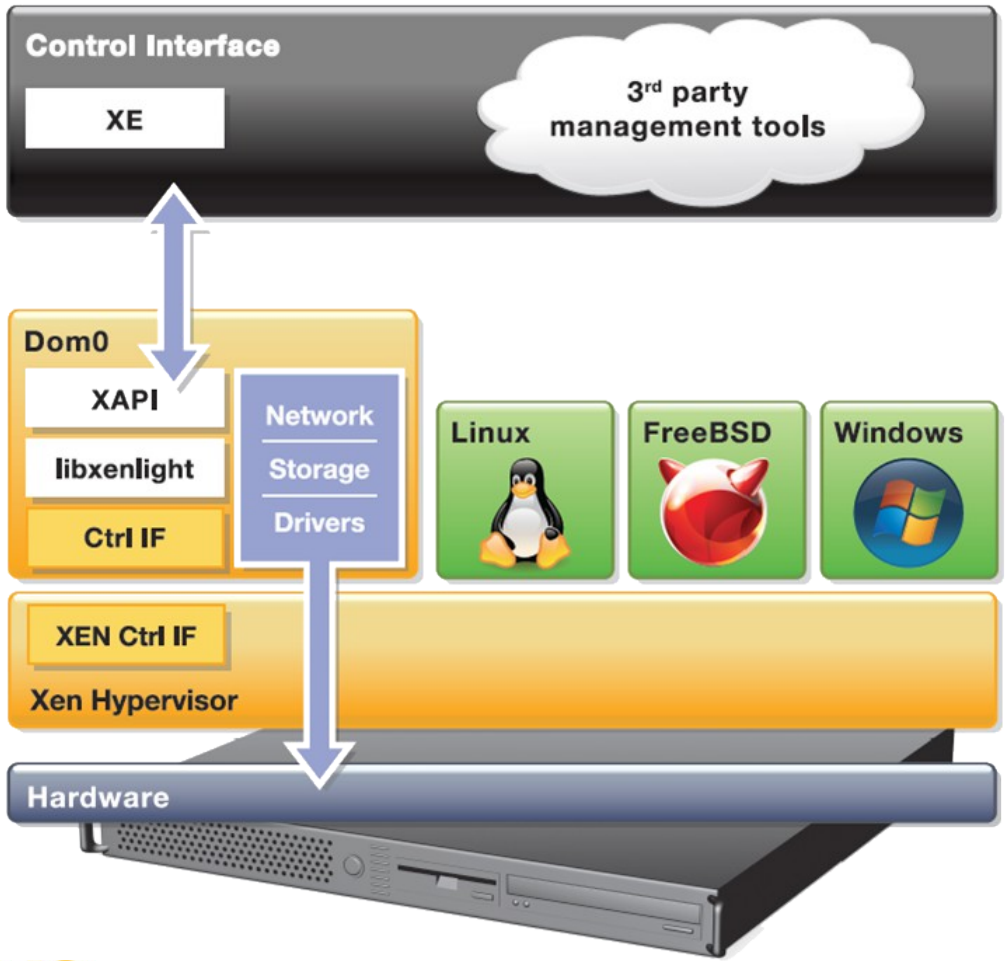
Xen Domain 0 in Linux 3.0+

(it is functional but not yet fully optimized)

On-going work to round out the feature set in Linux 3.2 +

XCP Project

XCP



- Complete vertical stack for server virtualization
- Distributed as a closed appliance (ISO) with CentOS 5.5 Dom0, misc DomU's, network & storage support and Xen API
- Open source distribution of Citrix XenServer

XCP Overview

- Open source version of Citrix XenServer
  - wiki.xen.org/wiki/XCP/XenServer_Feature_Matrix
- Enterprise-ready server virtualization and cloud platform
  - Extends Xen beyond one physical machine and other functionality
  - Lots of other additional functionality compared to Xen
- Built-in support and templates for Windows and Linux guests
- Datacenter and cloud-ready management API
  - XenAPI (XAPI) is fully open source
  - CloudStack and OpenStack integration
- Open vSwitch support built-in

Project “Kronos”: XAPI on Linux

- Make the XAPI toolstack independent of CentOS 5.5
- Extend the delivery model
  - Deliver Xen, XAPI and everything in between (storage manager, network support, OCaml libs, etc.) via your favorite Linux distro
    “apt-get install xcp-xapi” or “yum install xcp-xapi”

- Debian 
- Next: Ubuntu 12.04 LTS 
- Later: other major Linux distro (Fedora, CentOS, etc.) 
  - Volunteers are welcome!

Xen vs. XCP vs. XAPI on Linux

Xen	XCP (up to 1.1)	XAPI on Linux
Hypervisor: latest	lagging	Linux distro
Dom0 OS: CentOS, Debian, Fedora, NetBSD, OpenSuse, RHEL 5.x, Solaris 11, …	CentOS 5.5	Debian, Ubuntu, …
Dom 0: 32 and 64 bits	32 bits	32 and 64 bits
Linux 3 PVOPS Dom0: Yes	No	Yes
Toolstack: XM (deprecated), XL or Libvirt	XAPI + XE (lots of additional functionality to Xen)	Same as XCP
Storage, Network, Drivers: build and get yourself	Integrated with Open vSwitch, multiple storage types & drivers	Get them yourself
Configurations: Everything	constrained by XAPI	Same as XCP
Usage Model: Do it yourself	Shrink wrapped and tested	Do it yourself
Distribution: Source or via Linux\Unix distributions	ISO	Via host Linux distribution

XCP/XAPI Vision & Next Steps

- XCP & XAPI for Linux are the configuration of choice for clouds
  - Optimized for cloud use-cases
  - Optimized for usage patterns in cloud projects
  - XAPI toolstack is more easily consumable
- We are doing this by …
  - XenServer is built from XCP (almost there)
  - Track unstable Xen hypervisor and Linux kernels aggressively (almost there)
  - Deliver into Linux distributions : more flexibility (almost there)
  - Exploit advanced Xen security features
  - Fully open development model (build & test capability)

XCP 1.5 (soon)

- **Architectural Improvements:** Xen 4.1, GPT, smaller Dom0
- **GPU pass through:** for VMs serving high end graphics
- **Performance and Scalability:**
  - 1 TB mem/host
  - 16 VCPUs/VM, 128 GB/VM
- **Networking:** Open vSwitch (default), Active-Backup NIC Bonding
- **Virtual Appliance:** multi-VM and boot sequenced, OVF support
- More guest OS templates

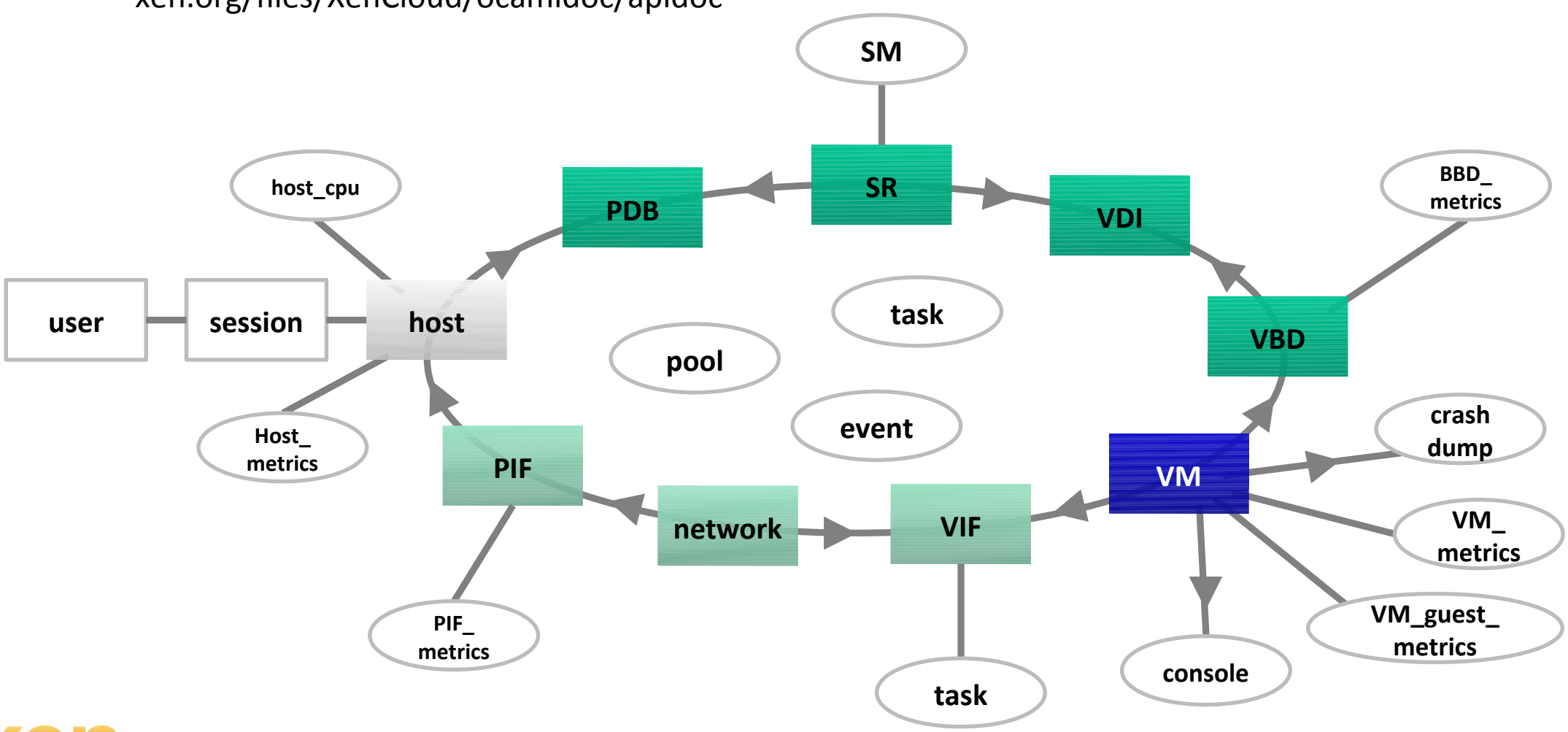
XAPI Overview

XAPI: What is it?

- XAPI is the backbone of XCP
  - Provides the glue between all components
  - Is the backend for all management applications
- Call it XAPI or XenAPI
- It's a XML-RPC style API, served via HTTPS
  - Provided by a service on every XCP dom0 host
  - Designed to by highly programmable
  - API bindings for many languages: .NET, Java, C, Powershell, Python
- XAPI is Extensible via plugins
  - E.g. used by OpenStack

XAPI from 30000 Feet

xen.org/files/XenCloud/ocamldoc/apidoc



XAPI Functionality Overview

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

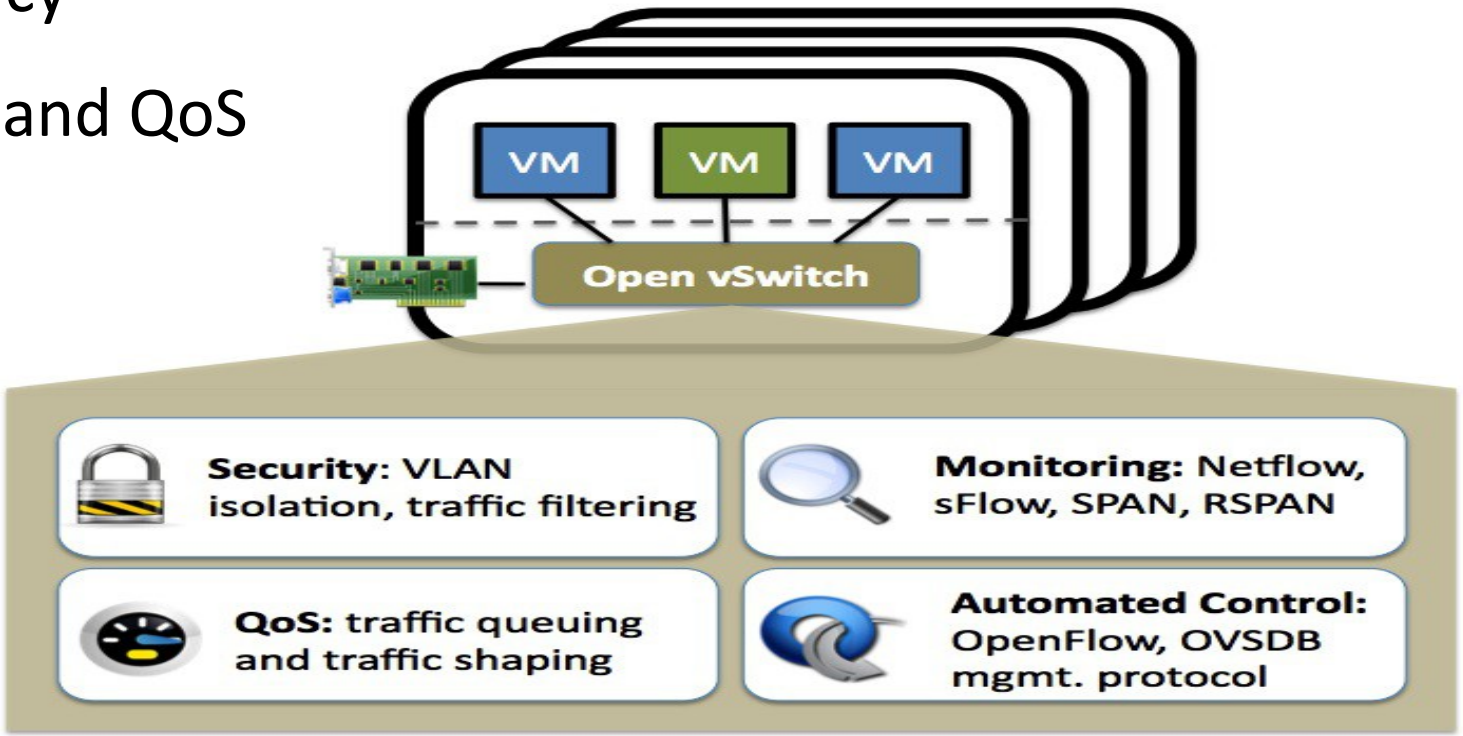
- Full list: wiki.xen.org/wiki/XCP/XenServer_Feature_Matrix

Open vSwitch

- Software switch, similar to:
  - VMware vNetwork Distributed Switch
  - Cisco Nexus 1000V
- Distribution agnostic. Plugs right into Linux kernel.
- Reuses existing Linux kernel networking subsystems.
- Backwards-compatible with traditional userspace tools.
- Free and Open Source http://openvswitch.org/

Why use Open vSwitch with Cloud?

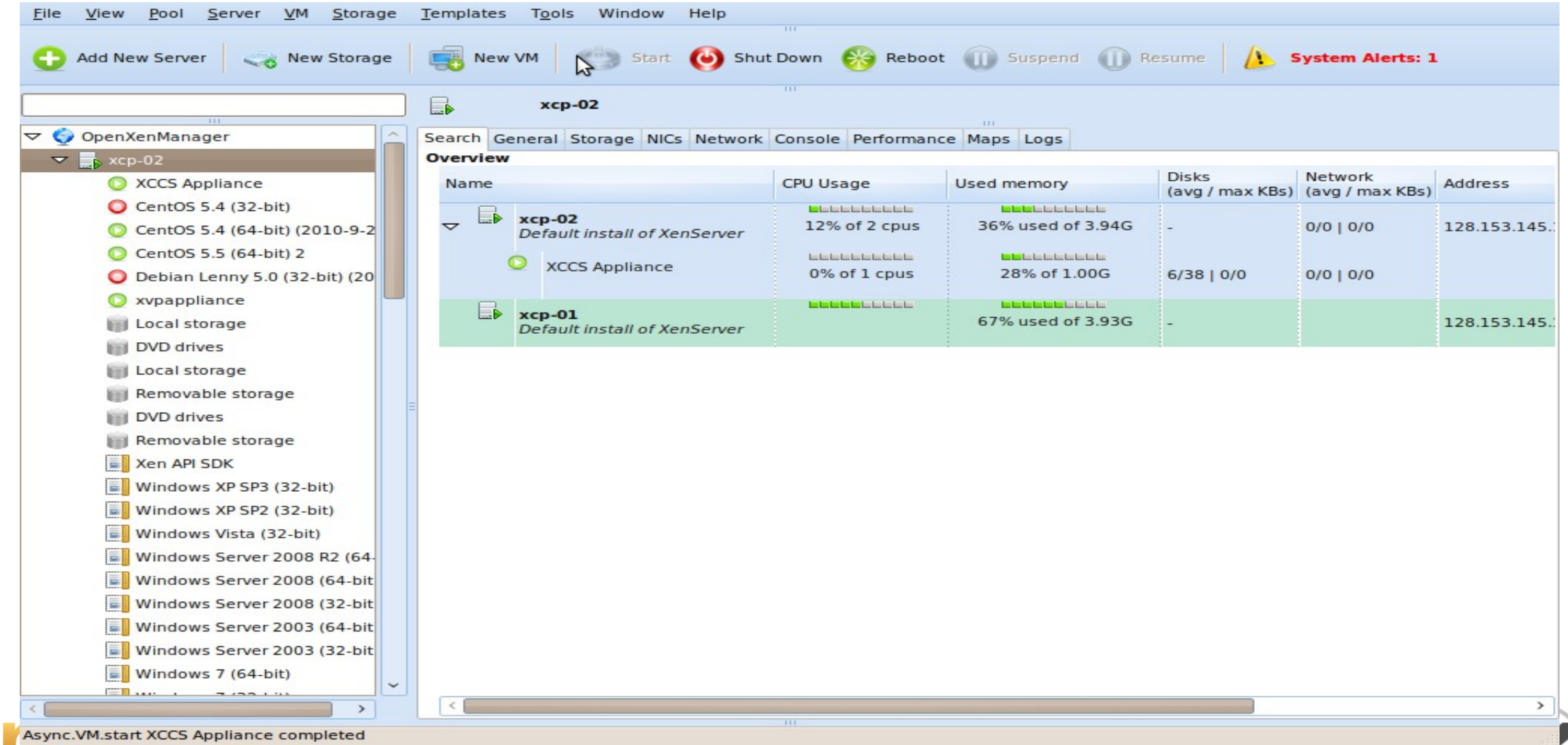
- Automated control: OpenFlow
- Multi-tenancy
- Monitoring and QoS



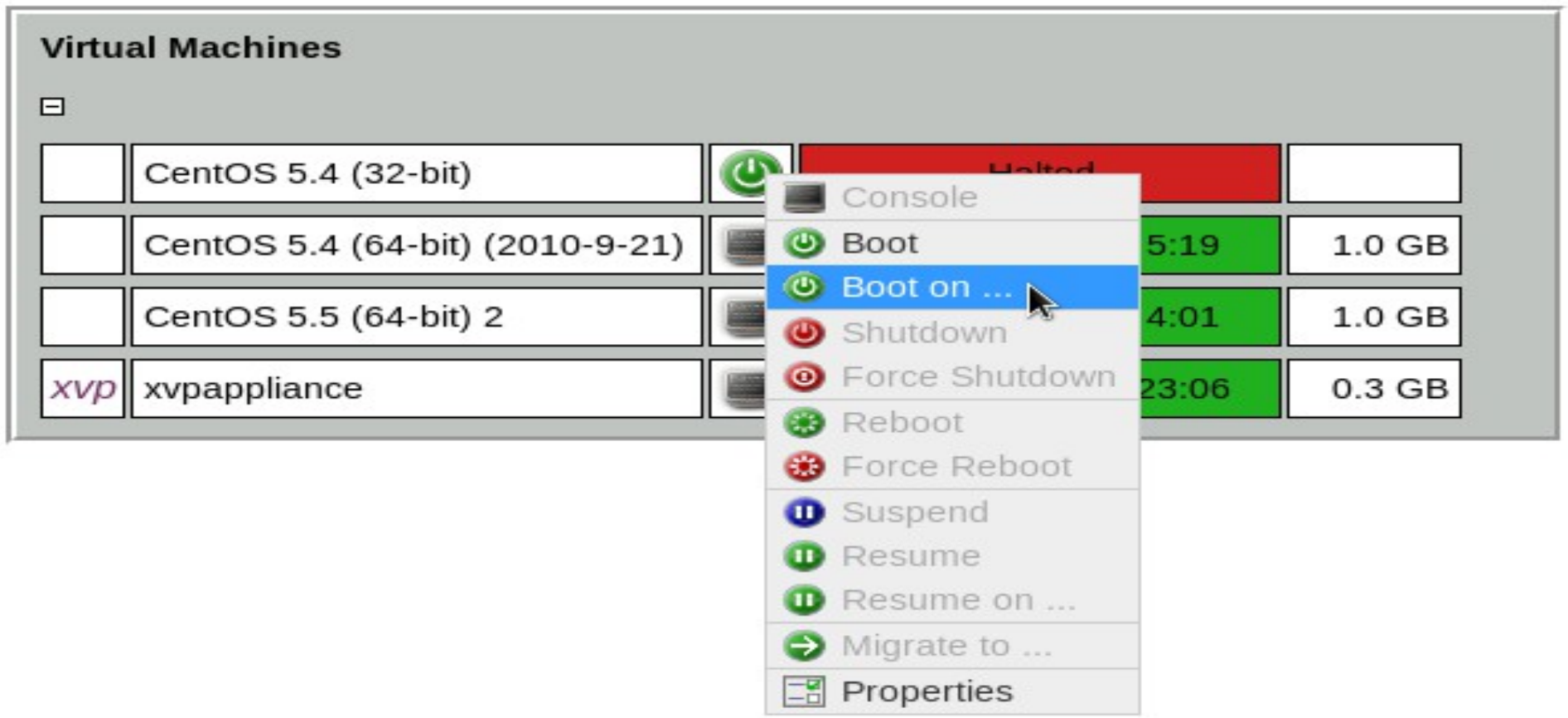
XAPI Management Options

- XAPI frontend command line tool: XE (tab-completable)
- Desktop GUIs
  - Citrix XenCenter (Windows-only)
  - OpenXenManager (open source cross-platform XenCenter clone)
- Web interfaces
  - Xen VNC Proxy (XVP)
    - lightweight VM console only
    - user access control to VMs (multi-tenancy)
  - XenWebManager (web-based clone of OpenXenManager
- XCP Ecosystem:
  - xen.org/community/vendors/XCPProjectsPage.html
  - xen.org/community/vendors/XCPProductsPage.html

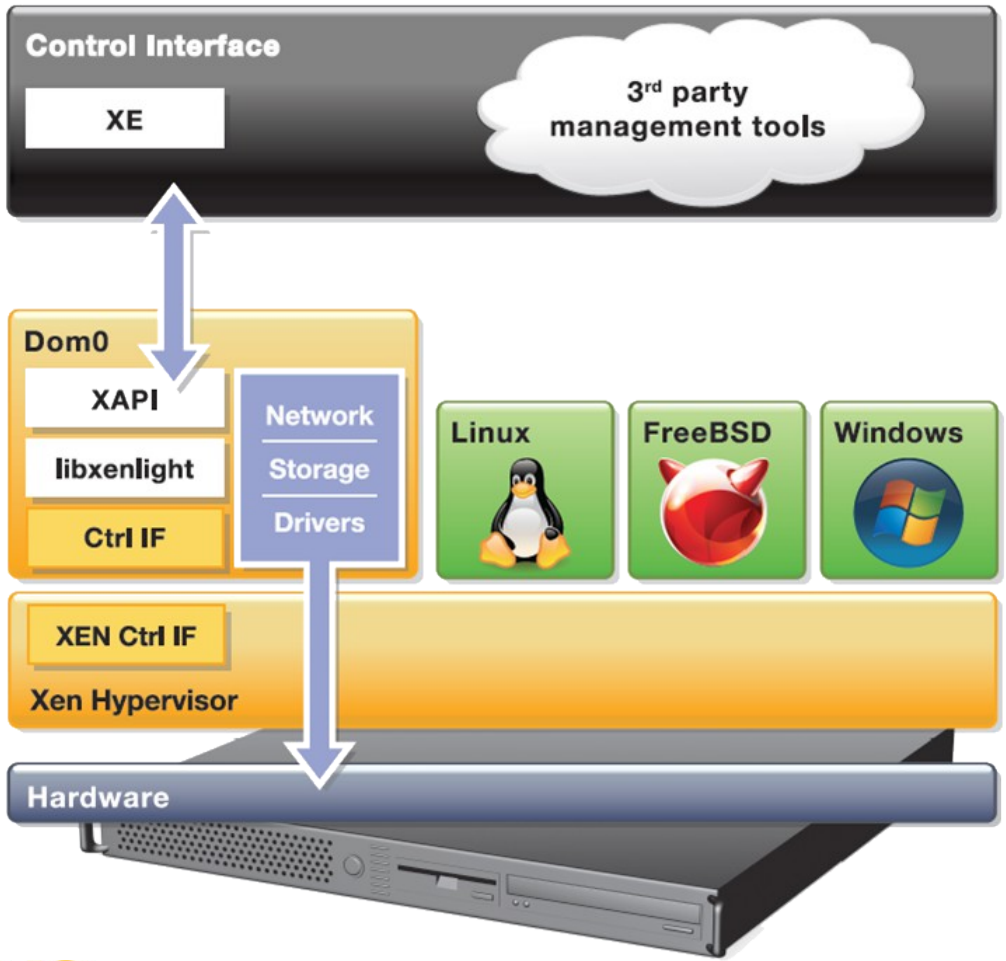
OpenXenManager



Xen VNC Proxy (XVP)



XCP and Cloud Orchestration Stacks



cloudstack
open source cloud computing

OpenNebula.org

openstack™

xen.org

Xen

Cloud VM vs. Cloud Package(s) in Dom0

Cloud VM (DomU)

Pros

- Isolation of cloud VM
- Security properties
- Pre-package + appliance

Cons

- Slightly more complex
- Less flexible

Cloud Package(s) in Dom0

Pros

- Simple install
- Flexibility
- Simpler overall

Cons

- Less isolation
- Cloud node is a potential entry point to compromise Dom0

Xen Hypervisor Project

xen.org

Xen

Xen 4.1 Release: 21 March 2011

- Very large system support
  - 4 TB; >255 CPUs
  - Reliability, Availability, Scalability enhancements
- CPU Pools for system partitioning
- Page sharing enhancements
- Hypervisor emergency paging / compression
- New “xl” lightweight control stack
- Memory Introspection API
- Enhanced SR-IOV support
- Software-implemented Hardware Fault Tolerance

Upcoming Xen 4.2 Release

- **Security:** Intel Supervisor Mode Execution Protection, XSM / Flask improvements
- **Scalability:** increased VM density for VDI use-cases, up to 256 Host CPUs for 64 bit HV , Multiple PCI segment support, prefer oxenstored
- **Performance:** PCI pass-through for Linux Guests, AMD SVM DecodeAssist support, Remus memory image compression
- EFI support
- Libvchan cross domain comms in Xen mainline
- XL improvements, XEND is formally deprecated
- Documentation improvements (e.g. man pages)

Xen, Security, QoS and the Cloud


“Security and QoS/Reliability are amongst the top 3 blockers for cloud adoption”
www.colt.net/cio-research
xen.org
Xen

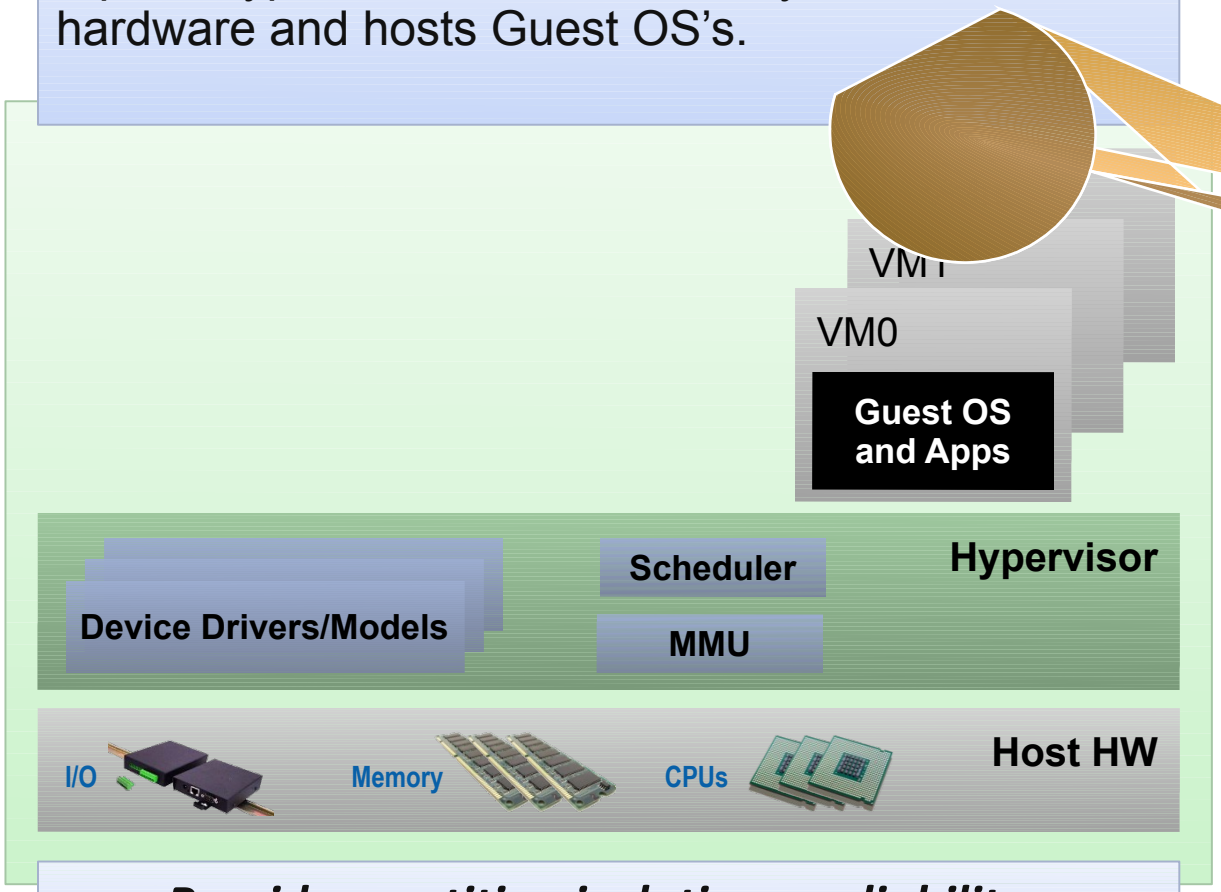
Security and the Next Wave of Virtualization

- Security is key requirement for Cloud
- Security is the primary goal of virtualization on the Client
  - Desktop, Laptops, Tablets & Smart Phones
- Maintaining isolation between VMs is critical
  - Spatial and Temporal isolation
  - Run multiple VMs with policy controlled information flow
    - E.g. Personal VM; Corporate VM; VM for web browsing; VM for banking

Architecture Considerations

Type 1: Bare metal Hypervisor

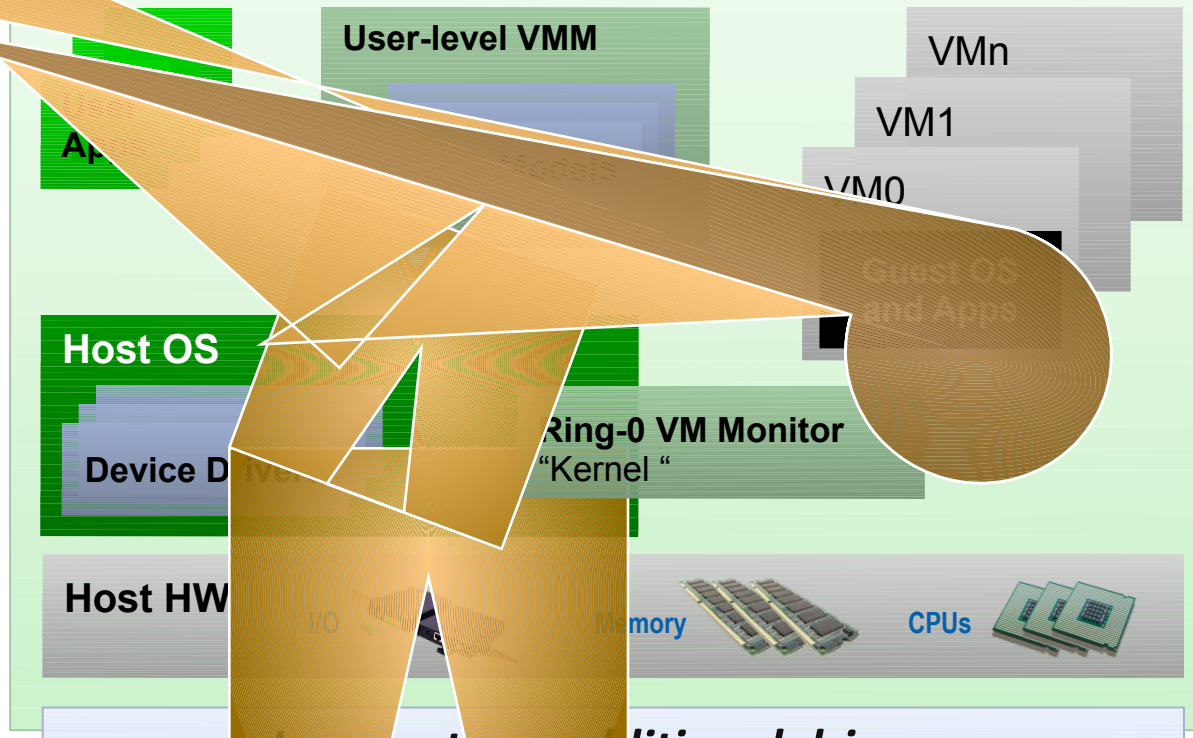
A pure Hypervisor that runs directly on the hardware and hosts Guest OS's.



Provides partition isolation + reliability, higher security

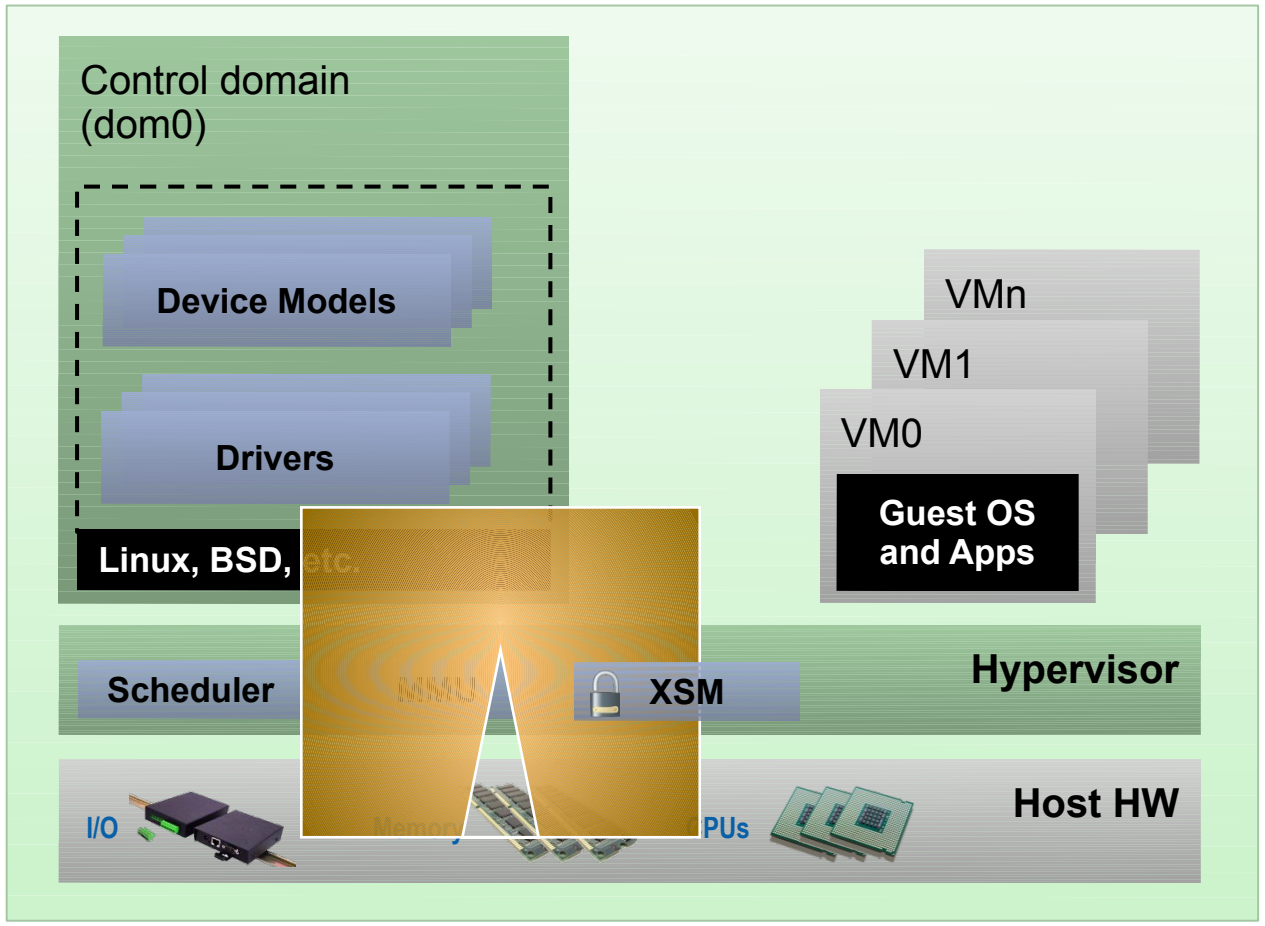
Type 2: OS 'Hosted'

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



Thin hypervisor

- Functionality moved to Dom0

Using Linux PVOPS

- Take full advantage of PV
- PV on HVM
- No additional device drivers (Linux 3.x dom0)

In other words

- low cost (drivers)
- Ease of use & Installation
- Isolation & Security

Xen Security & Robustness Advantages

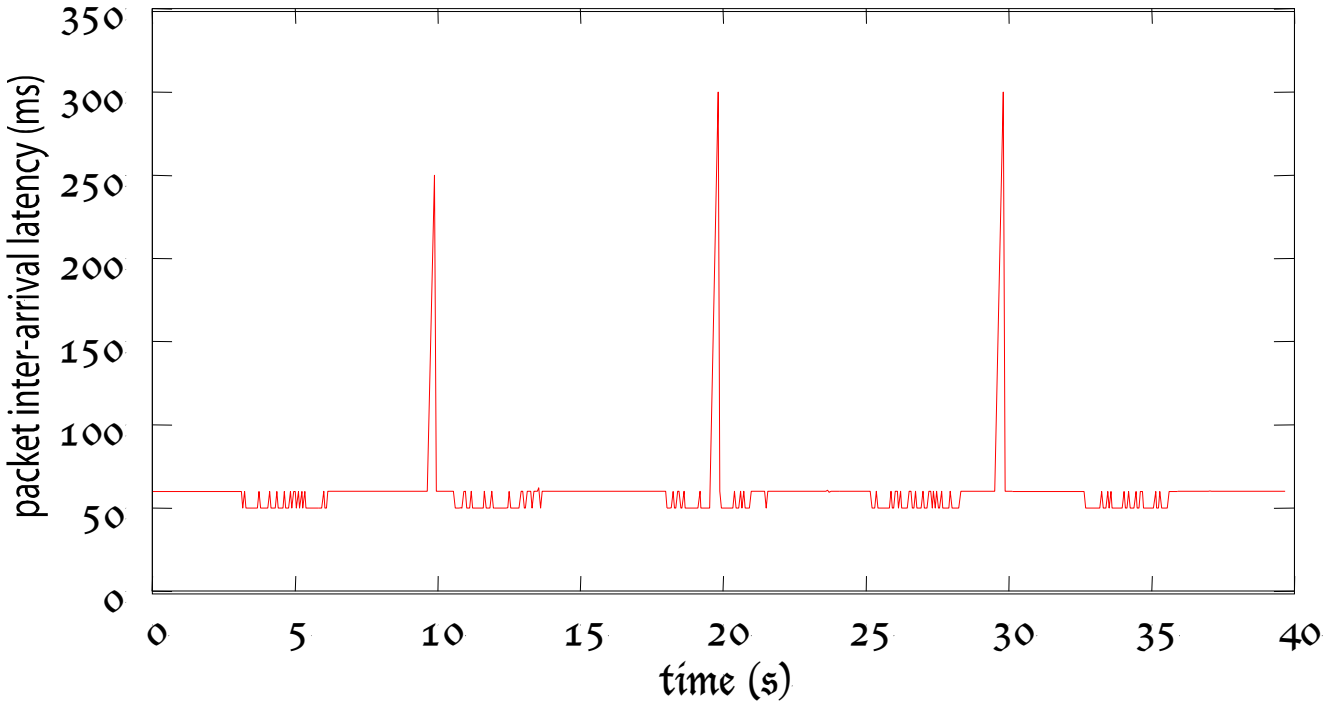
- Even without Advanced Security Features
  - Well-defined trusted computing base (much smaller than on type-2 hypervisor)
  - No extra services in hypervisor layer
- **More Robustness:** Mature, Tried & Tested, Architecture
- Xen Security Modules (or XSM)
  - Developed and contributed to Xen by NSA
  - Generalized Security Framework for Xen
  - The Xen equivalent of SELinux

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

Example: Network Driver Domain for HA

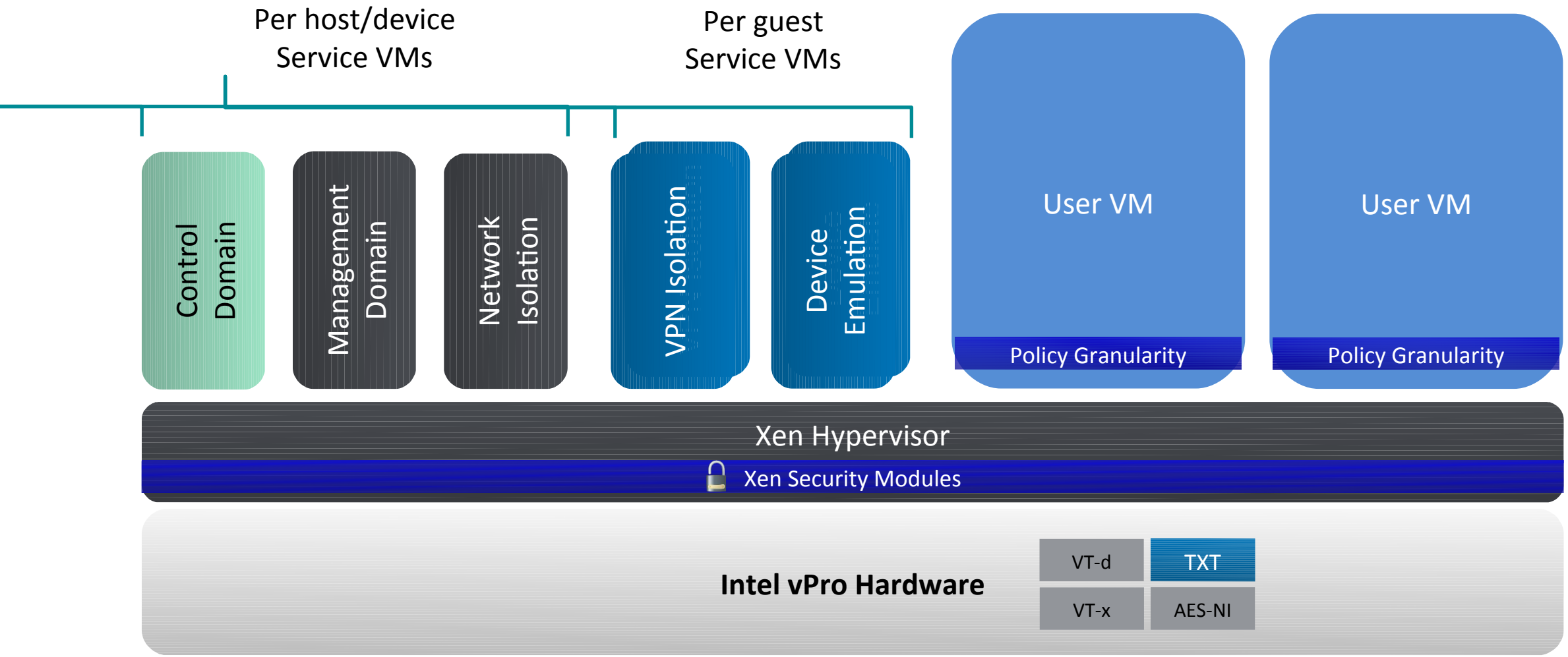
- Detect failure e.g.
  - Illegal access
  - Timeout
- Kill domain, restart
  - E.g. Just 275ms outage from failed Ethernet driver
- Auto-restarts to enhance security



Qubes OS / XenClient XT

- First products configured to take advantage of the security benefits of Xen's architecture
- Isolated Driver Domains
- Virtual hardware Emulation Domains
- Service VMs (global and per-guest)
- Xen Security Modules

Advanced XenClient Architecture



BUT…

- Today, XCP and commercial Xen based Server products
  - Do not make use of XSM
  - Do not make use of Advanced Security Features (Disaggregation)
- Most of these features are poorly documented on xen wiki
- In XCP, work has started to add these features
  - Various articles of how this may be done on the xen wiki
  - Hopefully more information soon
- Commitment on improving docs for Security, Reliability & Tuning

PVOPS : Xen in Linux 3.x

New in Linux 3.1 & 3.2

- Xen-pciback module
- Usability improvements
  - Auto loading of backend modules
  - Helps distros to package / deploy
- Memory Hotplug
- Bug fixes
  - e.g. VGA text console for dom0 fixed
- Many bug fixes: **THANK YOU!**
- Support for more than 256 PCI devices
- Kexec support for PV on HVM
- Laid foundations for HVM Driver Domains
- Blkback/front: added

Planned for 3.3 and beyond

- Documentation improvements
- Continue to round out the feature set, usability, rough edges
- Graphics improvements
- More Blkback and Netback optimisations
- New driver for doing ioctl
- ACPI power management
- Make Netback work much much better than it does now!
- Allow backends and xenstore to run in guests
- Completing work for Device Driver Domains

See full list at PVOPS Wiki

OK, so Upstream has stuff!

- So I can just install <favorite distro> and use Xen?
  - Yes!
    - But, check whether your distributions has 3.0+ kernel
  - For details visit Dom 0 Kernels for Xen Wiki
  - Some distros don't enable all backends – please open distro bugs (and let xen-devel know)
- Or you can build a v3.x Linux kernel with Xen 4.1.2 on existing distro.
  - Details, explanations, etc: XenParavirtOps Wiki

How you can help

- Take Linux 3.2 or 3.3 RCs (soon) for a spin with Xen 4.1.2
- Run it first without Xen to establish a baseline
- Then run it under Xen and see what happens
- Please send e-mail to xen-devel with what works and with what does not.

Xen ARM Project

xen.org

Xen

Xen ARM History

x86 Xen Hypervisor Release (Cambridge University)

Xen ARM 1st Release: ARM9 Xen Hypervisor, Mini-OS (Samsung)

Xen ARM 2nd Release: Paravirtualized Linux kernel (v2.6.24), Xen tool (Samsung)

Xen ARM 3rd Release: ARM11MPCore Support (Samsung)

Xen ARM 4th Release: Performance Optimization (Samsung)

Xen ARM 5th Release: Cortex-A9 MPCore Support (Samsung)

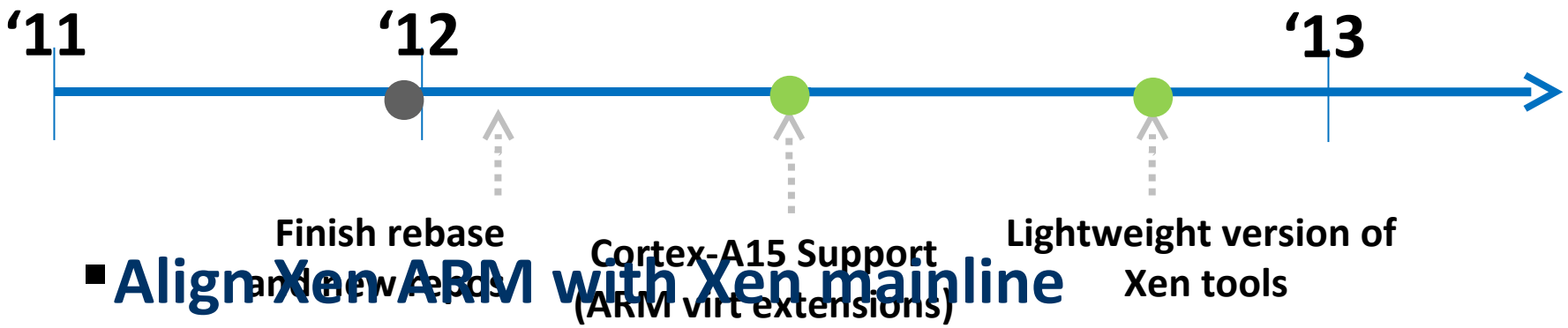
More information:

- wiki.xen.org/wiki/Xen ARM (PV) & xen-arm mailing list
  - Good overview in slides and papers links section
- wiki.xen.org/wiki/Xen_ARMv7_with_Virtualization_Extensions

From Mobiles to Laptops to Servers

- Smart Phones
  - **HW Consolidation:** AP(Application Processor) and BP(Baseband Processor) can share multicore ARM CPU SoC in order to run both Linux and Real-time OS efficiently
  - **OS Isolation:** important call services can be effectively separated from downloaded third party applications by Xen ARM combined with access control
  - **Rich User Experience:** multiple OS domains can run concurrently on a single smartphone
- Client Virtualization: Qubes OS / XenClient / XenClient XT
- ARM based Servers: ARM v7 & v8

Current Developments



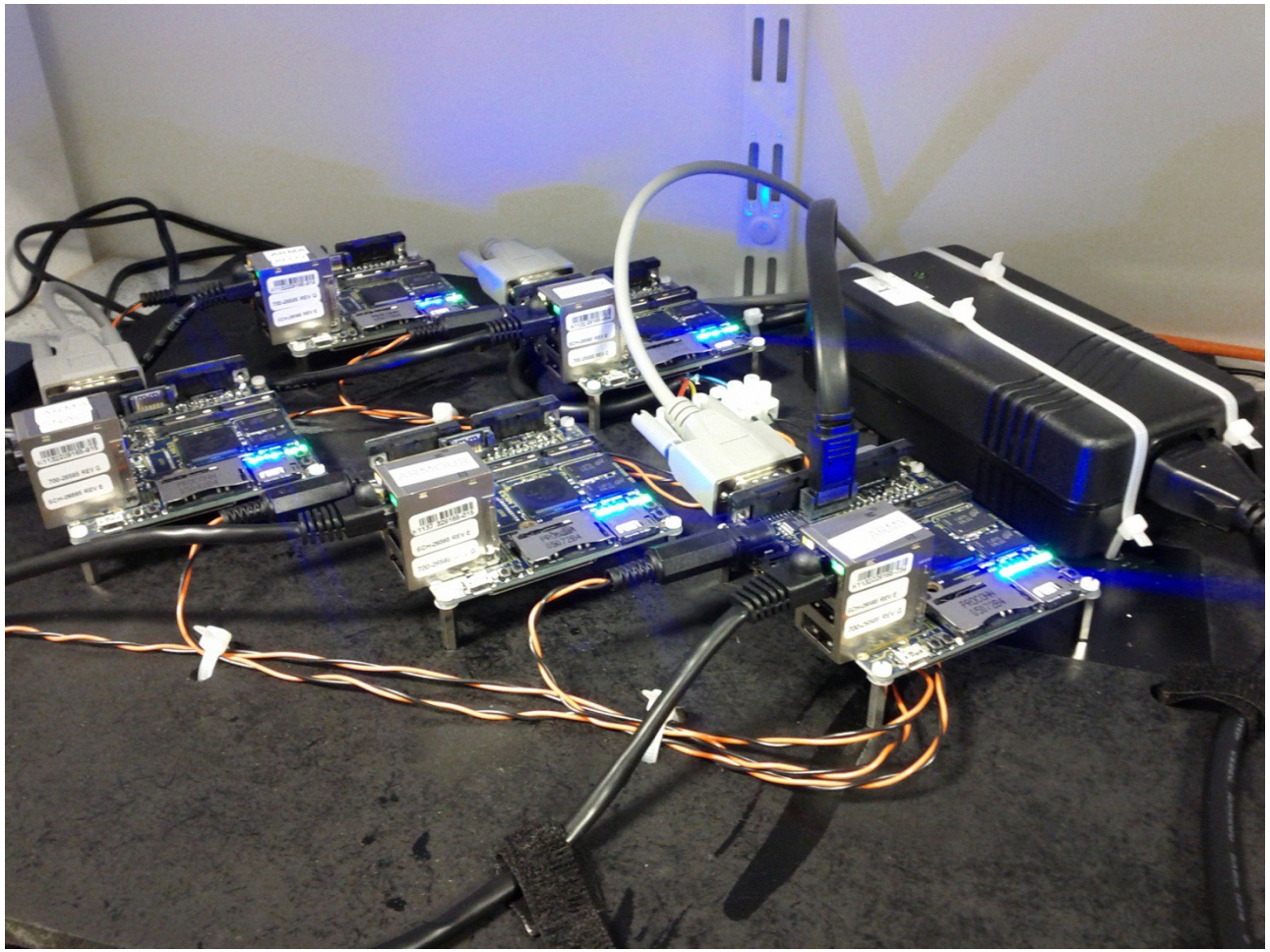
- **Align Xen ARM with Xen mainline**

Key Activities

- Publish xen-unstable.hg: public repo for Xen ARM that is routinely synced with xen-unstable.hg
  - Many parts of the Xen ARM has been rewritten for the integration
  - Publish source for PV port of ARM Linux Kernel

- **Prototyping of Cortex A15 support using ARM virtualization extensions**
  - First patches have made it into xen-unstable.hg

A bit of fun: our ARM Build Farm



10 Freescale i.MX53 Loco Quickstart boards

Running Debian "armhf" with a mainline 3.2.0 kernel

Speed up development of Xen for Cortex A15
(avoid cross compilation)

Summary: Why Xen?

xen.org

- 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
- XCP & XAPI
  - Ready for use with cloud orchestration stacks
  - XCP and XAPI on Linux: flexibility and choice
  - Lots of additional improvements for cloud coming in 2012
- Flexibility and choice of Usage Models
  - Also one of the challenges for Xen
- Catching up on “Ease of deployment and getting started”
- Open Source with a large community and eco-system

Resources

xen.org

Xen

Xen Resources

- **IRC:** ##xen @ FREENODE
- **Mailing List:** xen-users & xen-api
- **Wiki:** wiki.xen.org
  - Beginners & User Categories
- **Excellent XCP Tutorials**
  - A day worth of material @ xen.org/community/xenday11

How to Contribute

- Same process as for Linux Kernel
  - Same license: GPLv2
  - Same roles: Developers, Maintainers, Committers
  - Contributions by patches + sign-off
    (Developer Certificate of Origin)
  - Details @ xen.org/projects/governance.html

Shameless Marketing

Vendors in the Xen community are hiring!

Vendors in the Xen community are hiring!

Vendors in the Xen community are hiring!

xen.org/community/jobs.html

Questions …