

Xen Project: Hypervisor for Clouds



Julien Grall
julien.grall@linaro.org

LINUX FOUNDATION
COLLABORATIVE PROJECTS

About the Xen Project Stack

- The main components:
 - Xen Project Hypervisor, the central FOSS project
 - Xen Project API, the Cloud enabled subproject
 - Better known as “XAPI”
 - Xen Project is a Linux Foundation Collaborative Project
 - These are the subjects of this talk
- And then there's:
 - XenServer, a popular Xen Project-based product
 - Was partially closed source; open-sourced by Citrix in 2013



The Cloud “Problem”



IT Before the Cloud

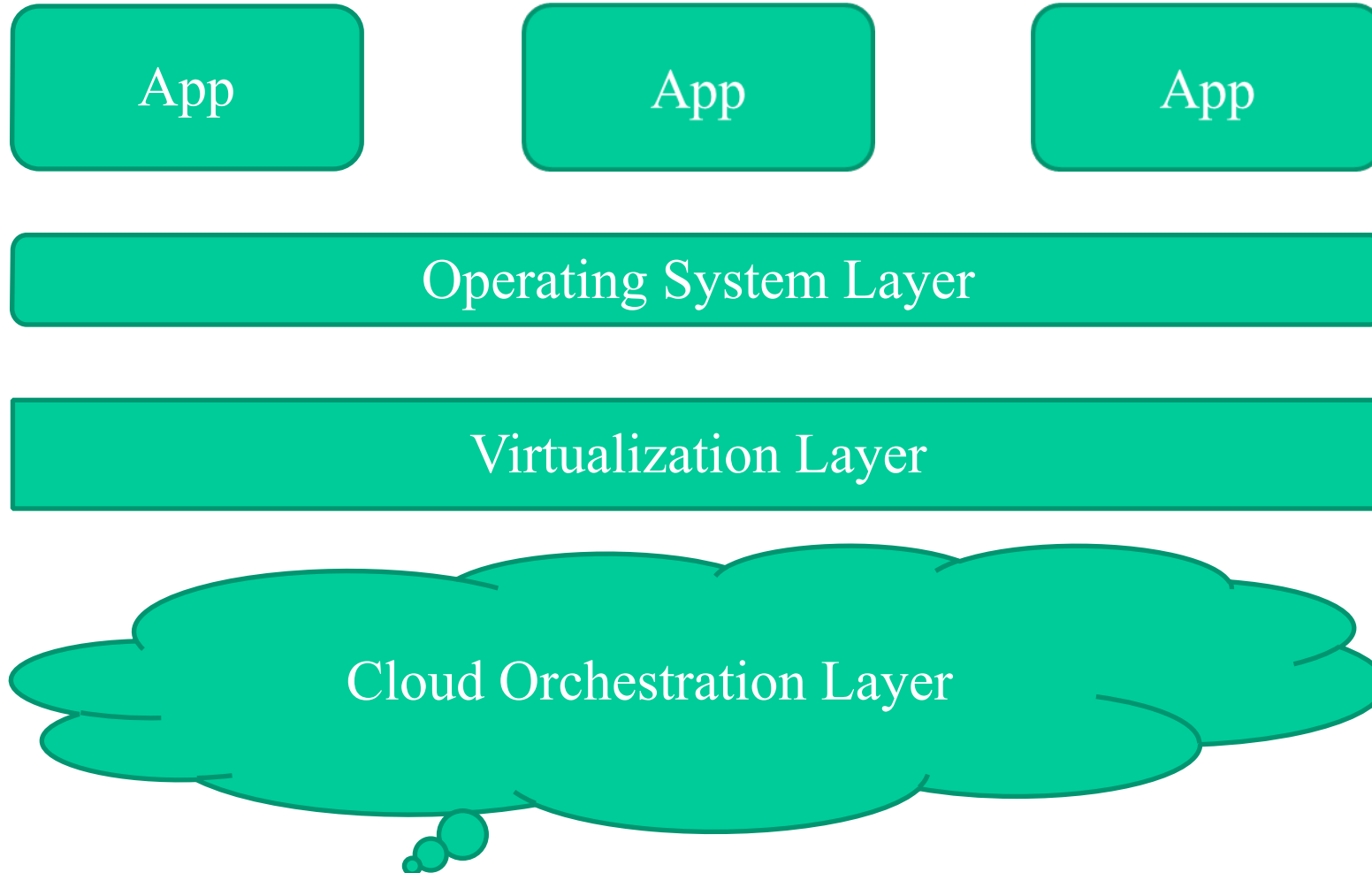
- **Stability is Paramount**
 - The value of IT to the corporation is consistent service availability
 - Service capacity specified a year or more in advance
 - What's up, stays up
- **Change is Bad**
 - Change to status quo is disruptive and dangerous
 - Changes are beaten into submission until they become part of the new status quo - and then they are no longer changes

IT Reinvented in the Cloud

- **Availability of Services is Paramount**
 - The value of IT to the corporation is consistent service availability at levels matching dynamic business demand
 - Service capacity must move with business needs
 - What's up when depends on what's needed when
- **Change is Good**
 - Services must change to cover the needs of the moment
 - Lack of change = lack of value



Cloud 101: Layers of the Cloud



Virtualization in the Cloud

- It must be stable
- It must be secure
- It must be configurable on a large scale
 - The “user at machine” paradigm does not work
 - If it requires a mouse, you’re in trouble
- It must take orchestration (APIs, command line)
- It must be multi-tenant
- It must not lock you into one concept or provider of Cloud



Xen Project: Highly Stable

- Solid track record
 - Amazon's AWS cloud business uses Xen Project
 - Verizon launched a new Xen Project-based cloud
- Linux Foundation Project Partners:
 - Amazon, AMD, ARM, CA, Cisco, Citrix, Google, Intel, NetApp, Oracle, Rackspace, Verizon, and more



Xen Project: Highly Secure

- SELINUX
- FLASK
 - SELINUX capabilities at the VM level by the same team
- Disaggregation
 - Segment device drivers into discrete VMs
- Architectural advantages of a Type-1 Hypervisor
 - See the slides of Advanced Security talk on XenProject.org

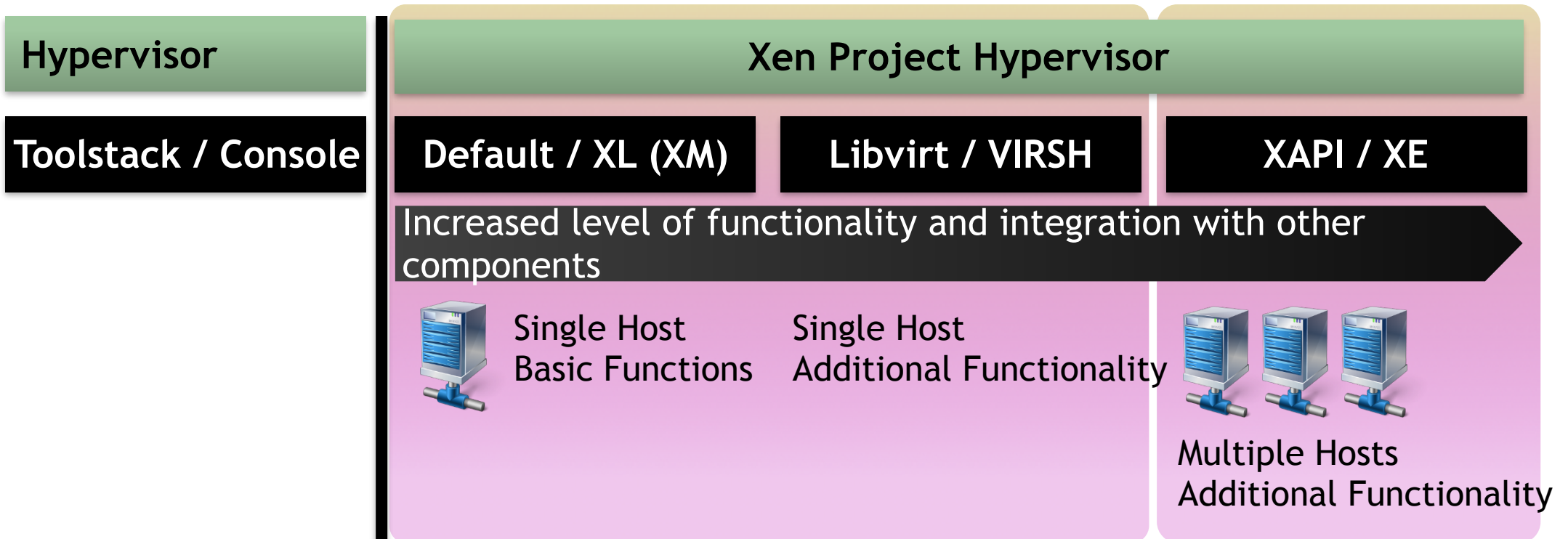


Xen Project: Configurable at Scale

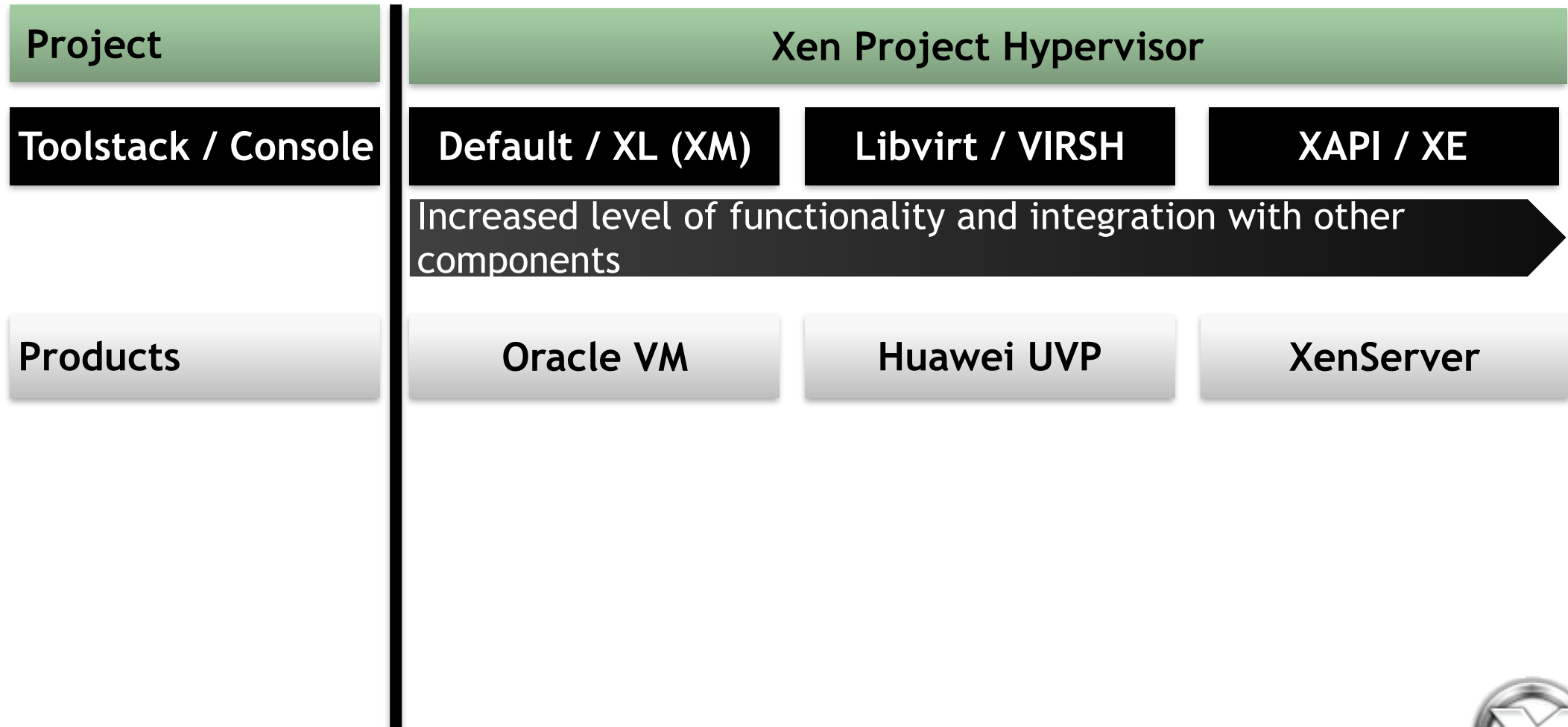
- Toolstacks give rich API and command line capabilities
- Not GUI-centric
- Empowers orchestration via scripting, power tools (*Puppet, Chef, etc.*), GUIs (*XenServer's XenCenter, Xen Orchestra, etc.*), and Cloud layers (*OpenStack, CloudStack, Open Nebula, etc.*)






Xen Project: Rich Toolstacks



Xen Project: Tools for Different Solutions



Xen Project: Tools for Different Clouds

| | | | |
|---------------------|--|--|---|
| Project | Xen Project Hypervisor | | |
| Toolstack / Console | Default / XL (XM) | Libvirt / VIRSH | XAPI / XE |
| | Increased level of functionality and integration with other components | | |
| Products | Oracle VM | Huawei UVP | XenServer |
| Used by ... |  |  |  |



Xen Project: A Multi-tenant Solution

- Multiple groups share common resources securely
 - Clouds require sharing common resources
 - Organizations often need their VMs to be visible to each other, but entirely invisible to all other VMs
 - XAPI makes this happen
 - Critical ability for hosting providers



Xen Project: Doesn't lock you in

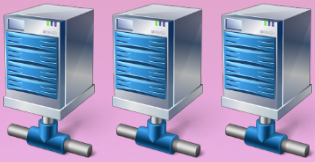
- Xen Project does not force its view of the Cloud on you
- Xen Project does not force you to use a “favored” Cloud solution
- This is one of the reasons why Cloud innovation happens in the world of FOSS: It gives power to the Cloud, but allows Cloud orchestration solutions to innovate
- There is no attempt to bend your efforts to the will of some corporate business plan



XAPI : Orchestration Choices

The Hypervisor

XAPI / XE



Multiple Hosts
Additional Functionality

apache **cloudstack**
open source cloud computing

OpenNebula.org



Xen
Orchestra



Xen Project Healthcheck

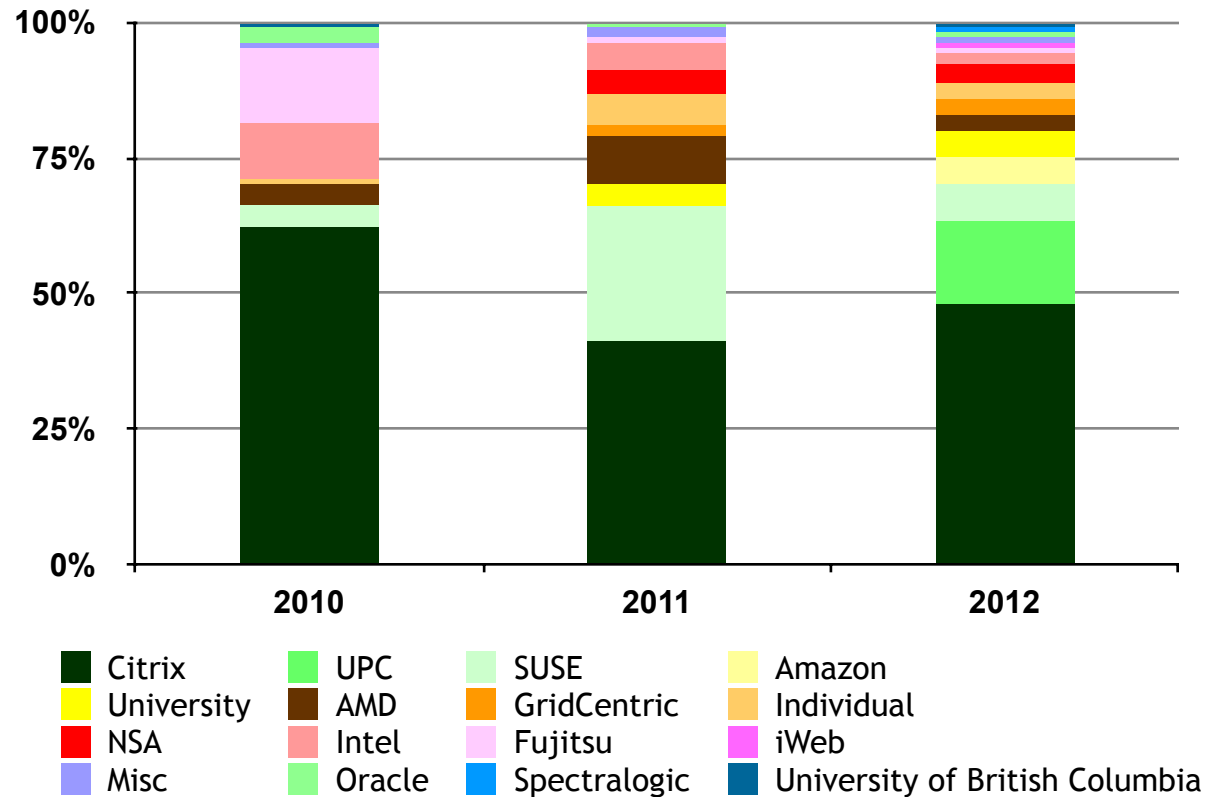


2013: Xen Project Joins Linux Foundation

- See the following teams on the new XenProject.org site:
 - Hypervisor
 - XAPI
 - ARM Hypervisor (for Servers as well as Mobile Devices)
 - Mirage OS
- Governance : mixture between Linux Kernel and Apache
 - Consensus decision making
 - Sub-project life-cycle (aka incubator)
 - PMC style structure for team leadership



Xen Project Contributor Community is Diversifying



- The number of “significant” active vendors is increasing
- New feature development driving new participation



More Xen Project Features...

- Unikernel development and support (Mirage OS, etc.)
- ARM hardware support
- Live Migration of VMs: XenMotion (via XAPI)
- High Availability: Remus (& COLO for non-stop)
- Wide variety of Control Domains supported
- Even wider variety of Guest Domains
- Multiple virtualization modes improve performance



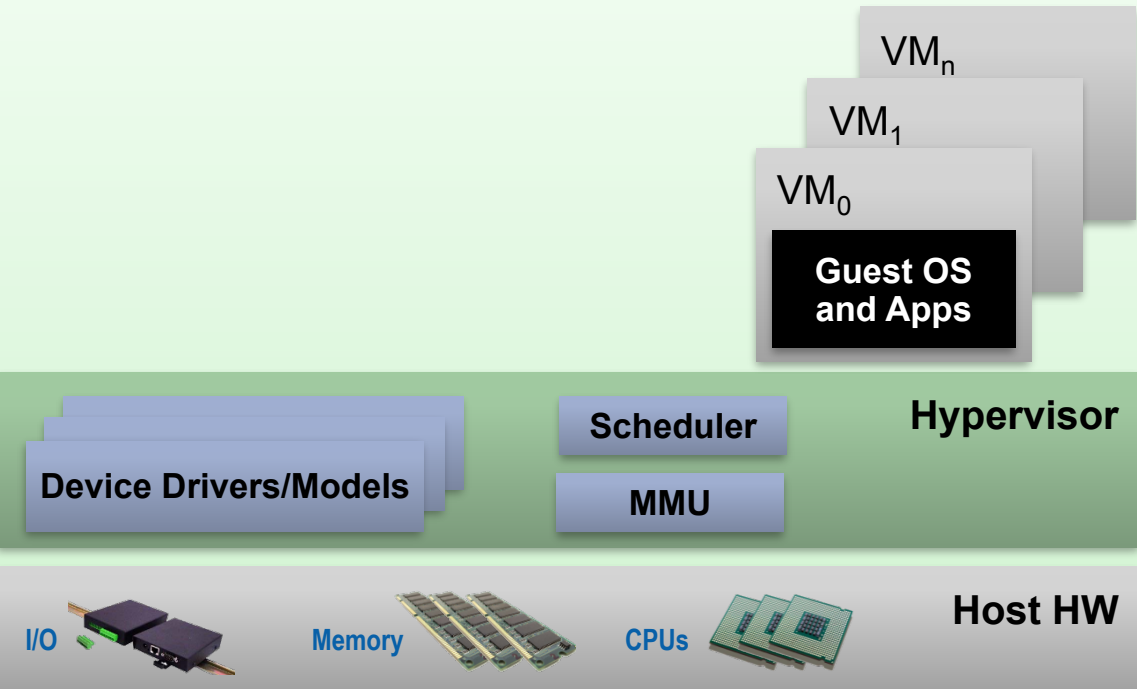
Hypervisor Architecture



Hypervisor Architectures

Type 1: Bare metal Hypervisor

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

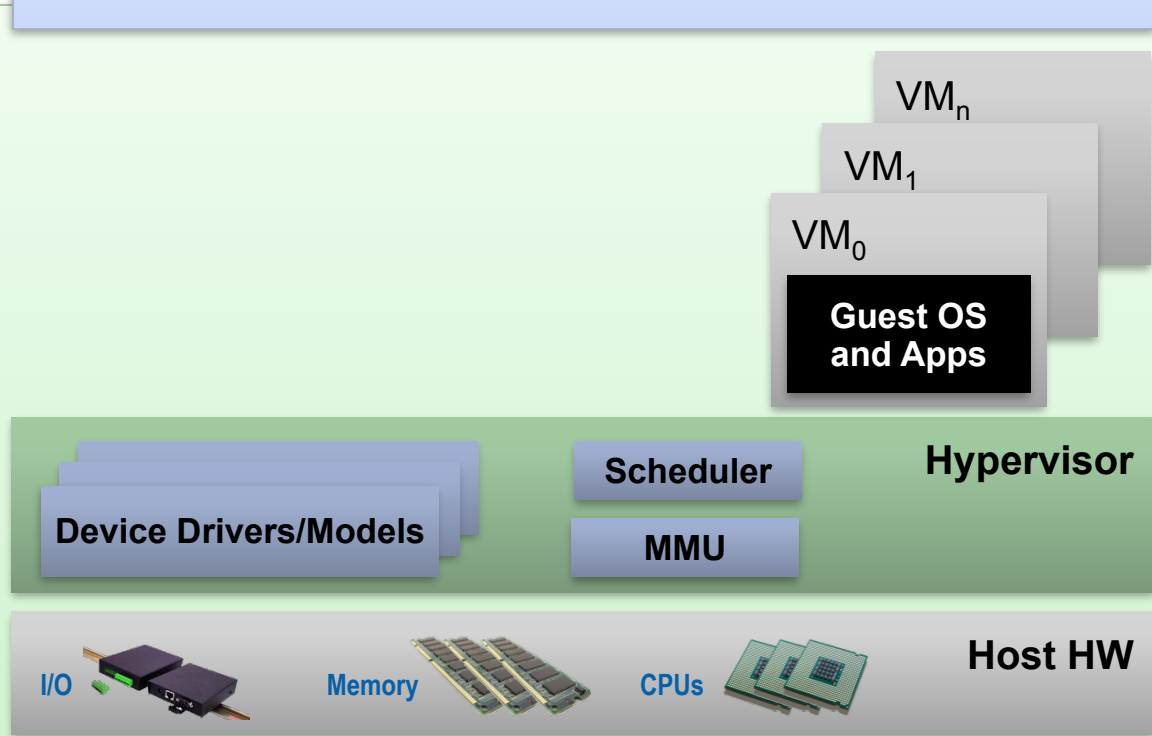


*Provides partition isolation + reliability,
higher security*

Hypervisor Architectures

Type 1: Bare metal Hypervisor

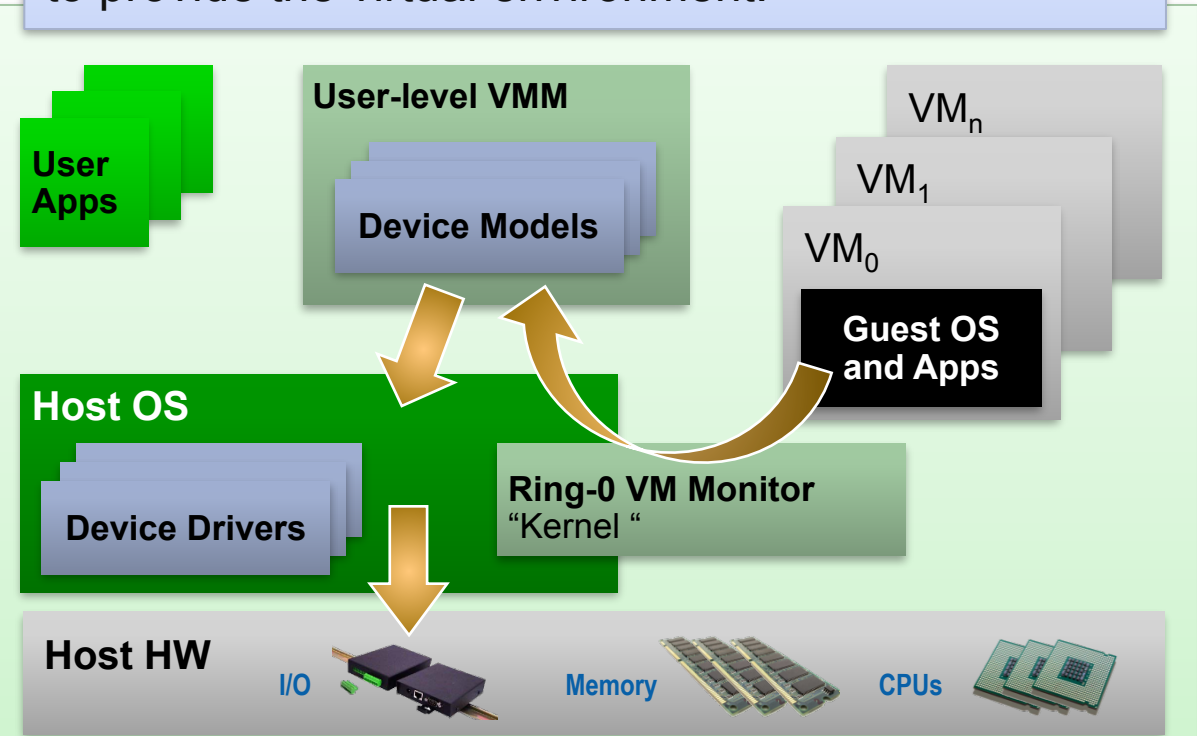
A pure Hypervisor that runs directly on the hardware and hosts Guest OS's.



Provides partition isolation + reliability, higher security

Type 2: OS 'Hosted'

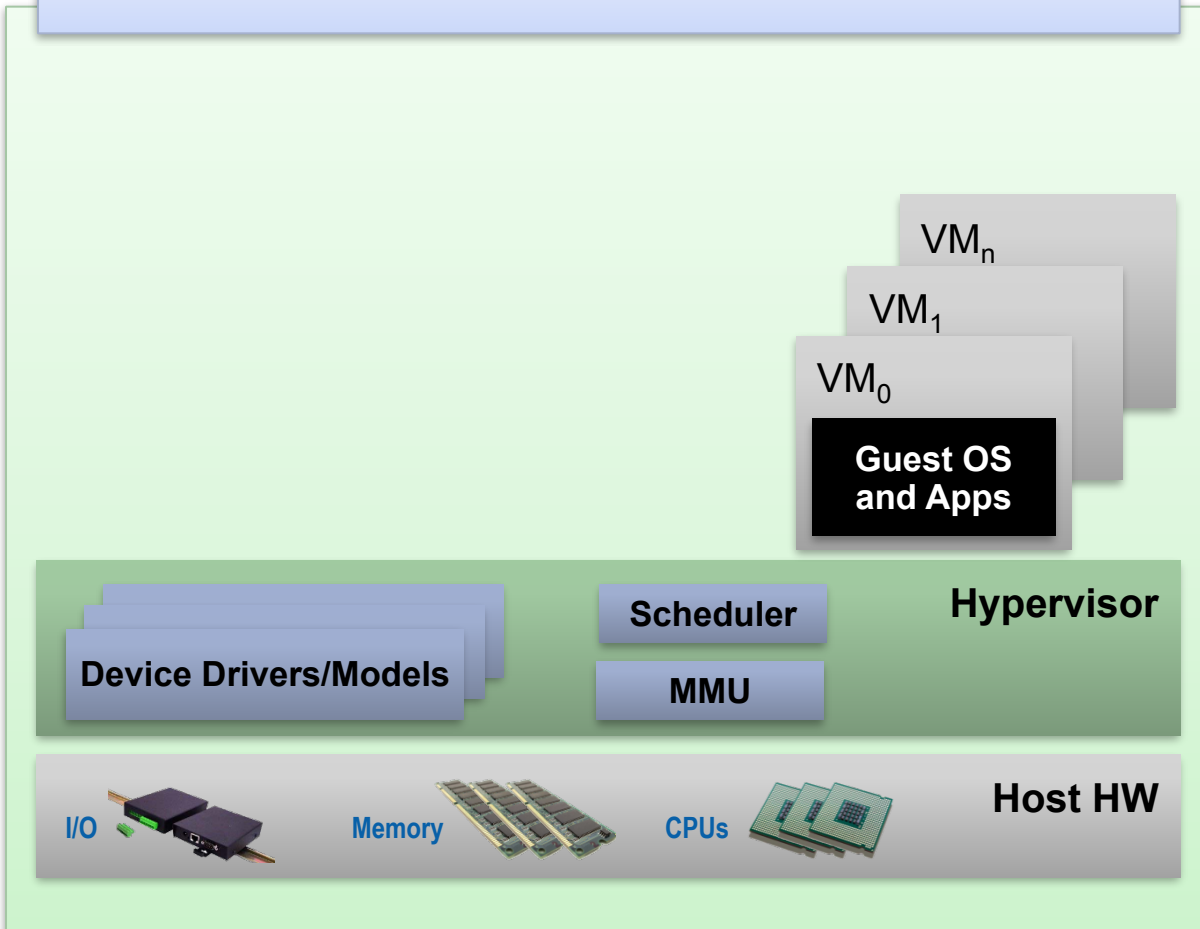
A Hypervisor that runs within a Host OS and hosts Guest OS's inside of it, using the host OS services to provide the virtual environment.



*Low cost, no additional drivers
Ease of use & installation*

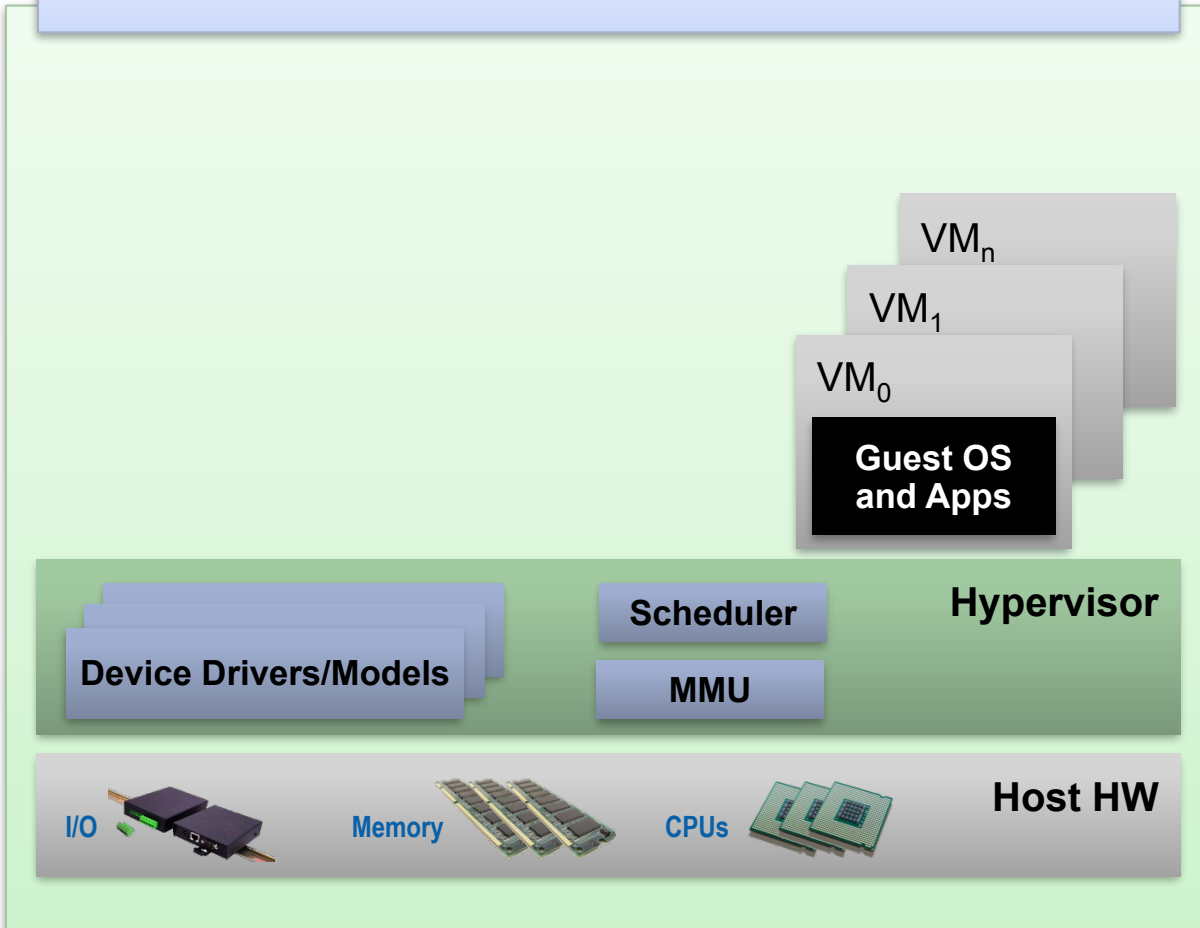
Xen Project: Type 1 with a Twist

Type 1: Bare metal Hypervisor

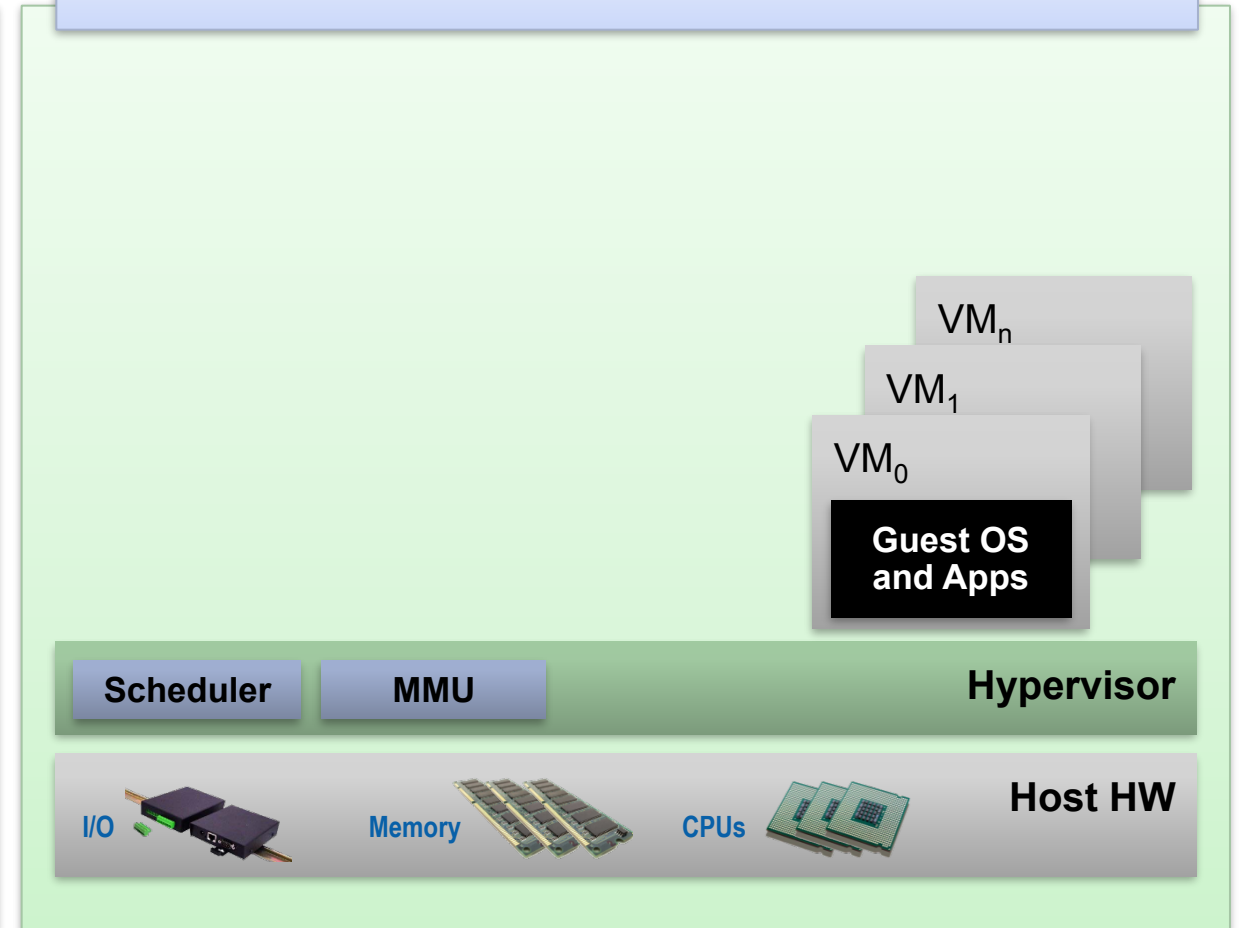


Xen Project: Type 1 with a Twist

Type 1: Bare metal Hypervisor

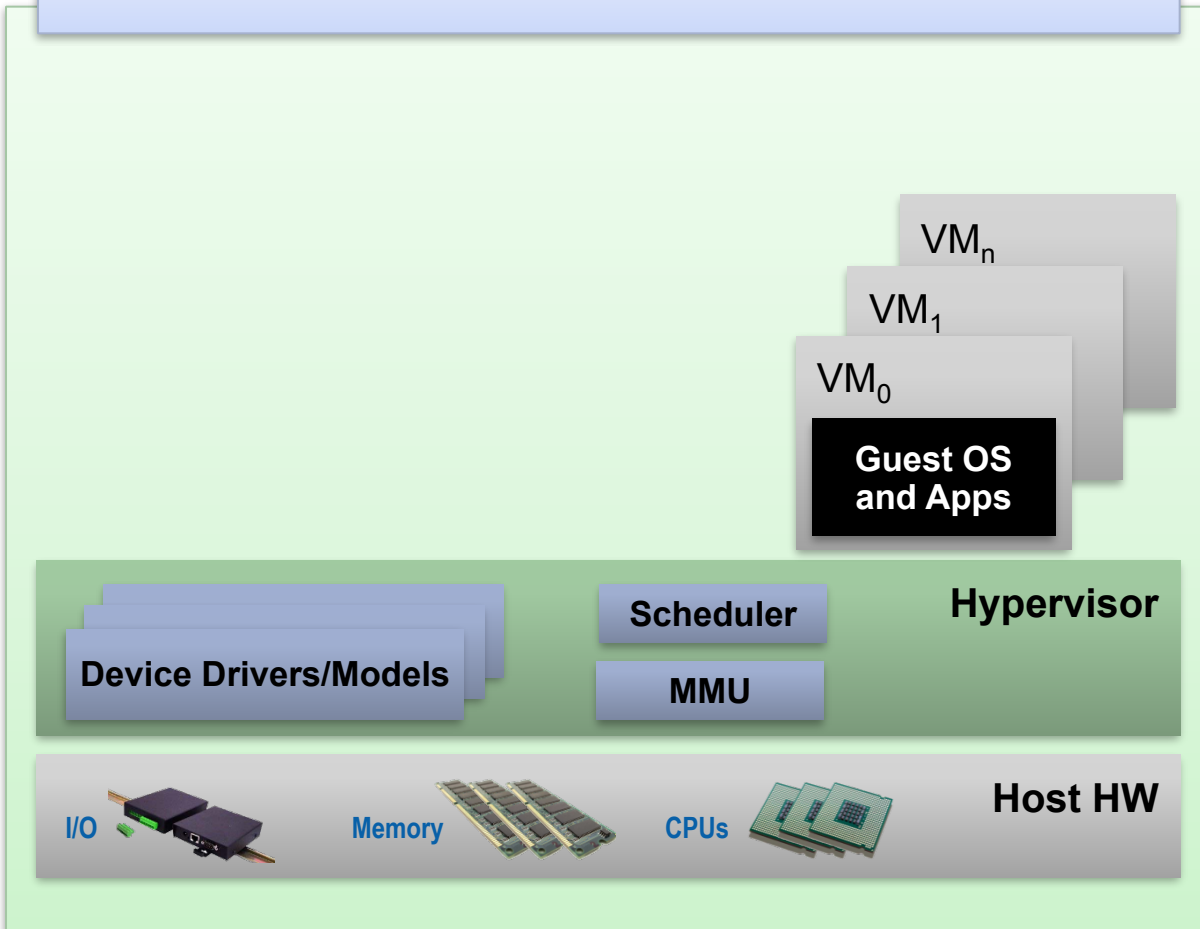


Xen Project Architecture

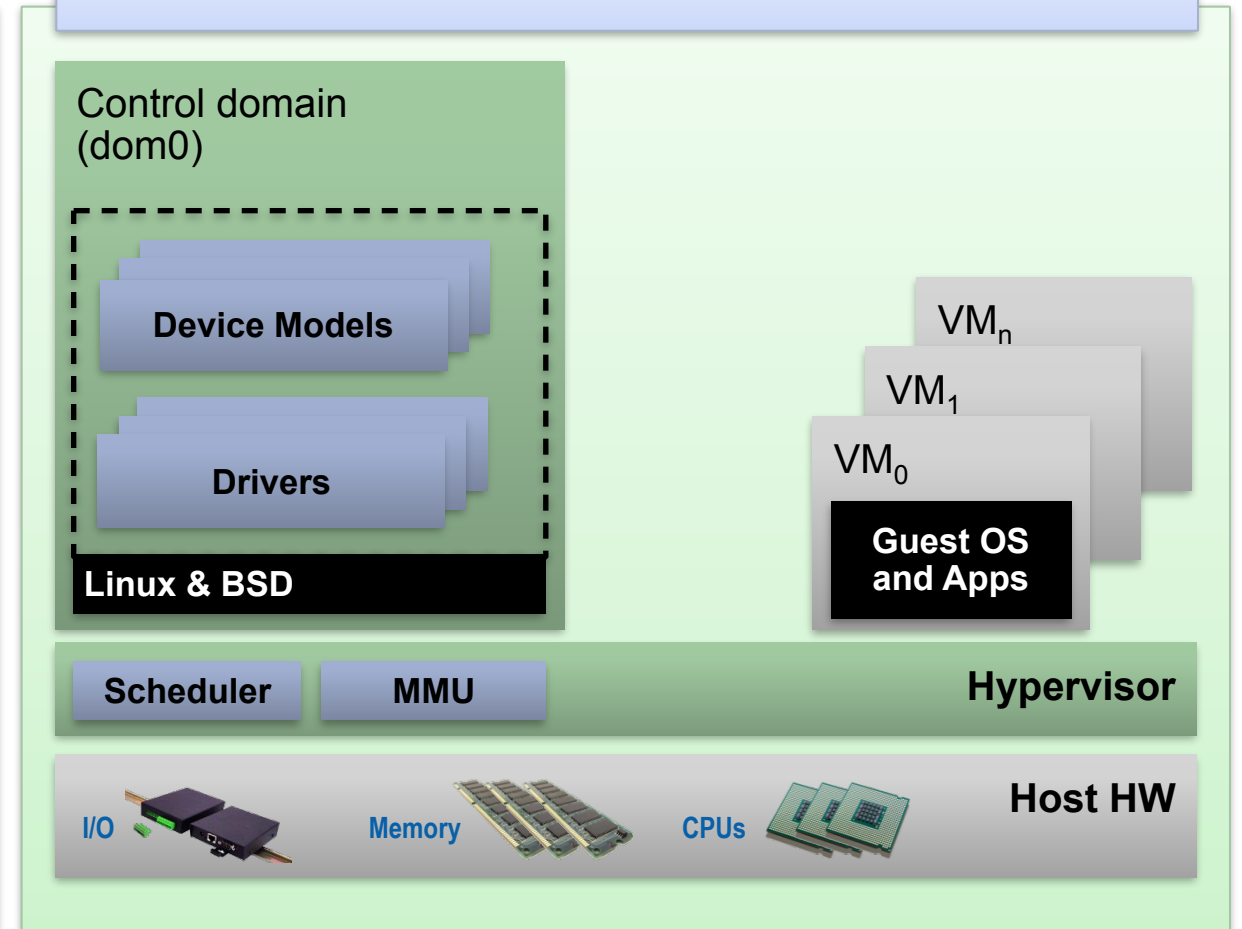


Xen Project: Type 1 with a Twist

Type 1: Bare metal Hypervisor



Xen Project Architecture



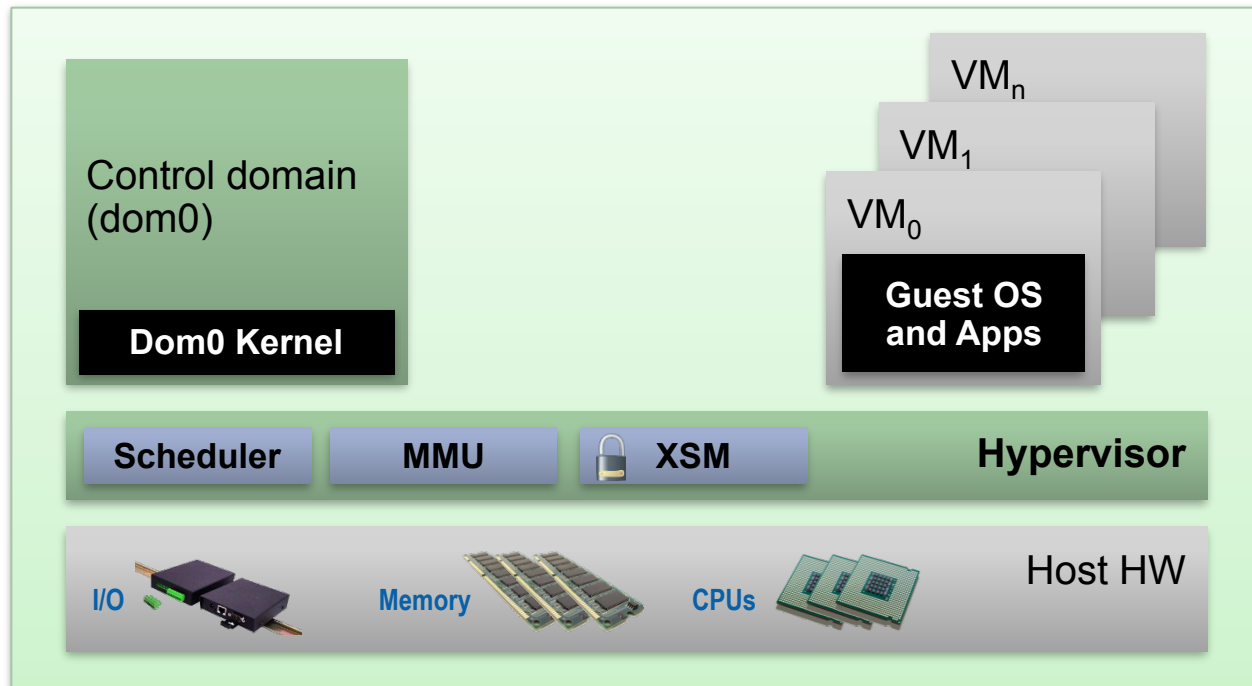
Xen Project and Linux

- Xen Project Hypervisor is not in the Linux kernel
- **BUT**: everything needed to run the hypervisor is
- Xen Project packages in all distributions (not in RHEL6, but CentOS 6 via Xen4CentOS)
 - Install Control Domain (Dom0) Linux distribution
 - Install Xen Project package(s) or meta package
 - Reboot
 - Configure stuff: set up disks, peripherals, etc.

[More info: wiki.xenproject.org/wiki/Category:Host_Install](http://wiki.xenproject.org/wiki/Category:Host_Install)



Basic Xen Project Concepts



■ Trusted Computing Base

Control Domain aka [Dom0](#)

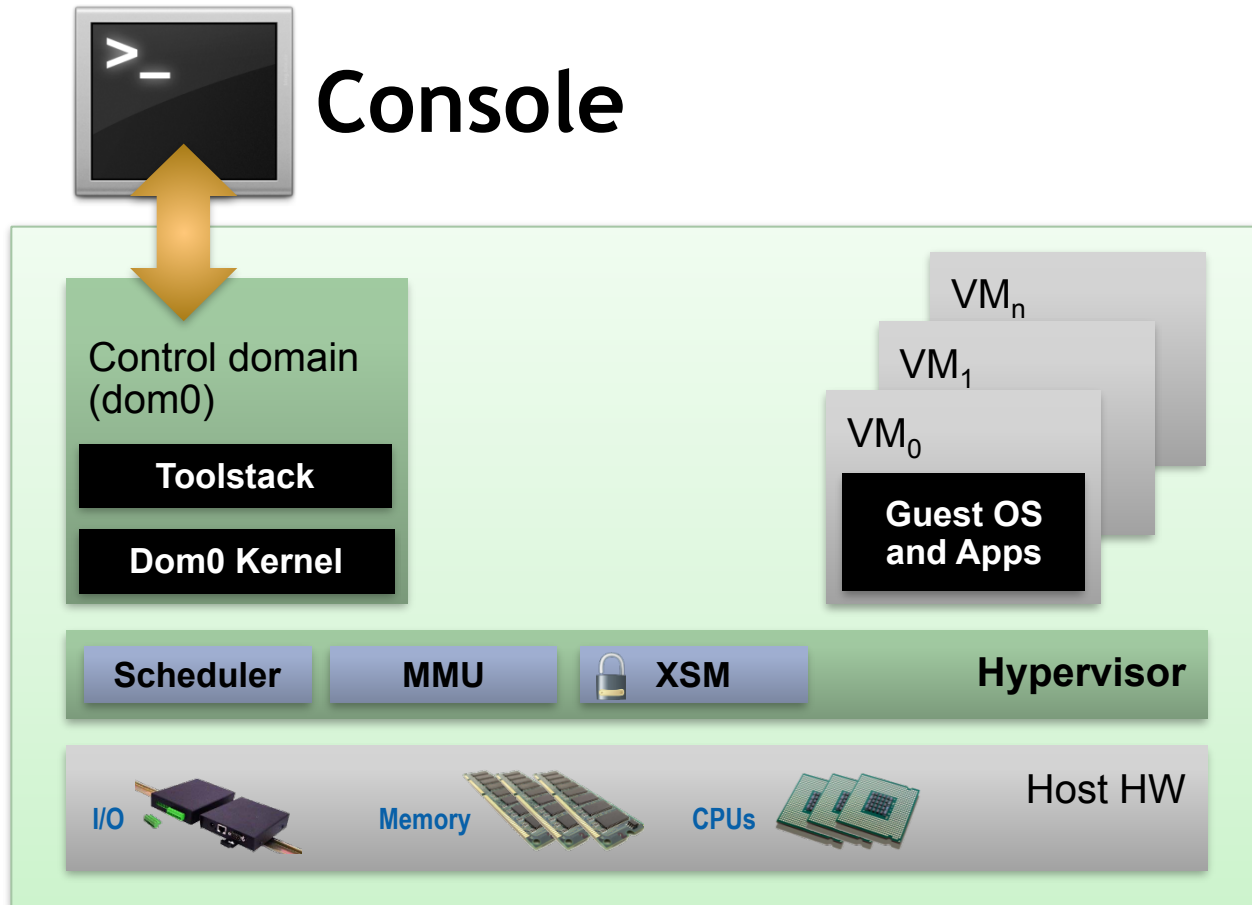
- Dom0 kernel with drivers

Guest Domains

- Your apps



Basic Xen Project Concepts: Toolstack+



Console

- Interface to the outside world

Control Domain aka [Dom0](#)

- Dom0 kernel with drivers
- Xen Project Management Toolstack

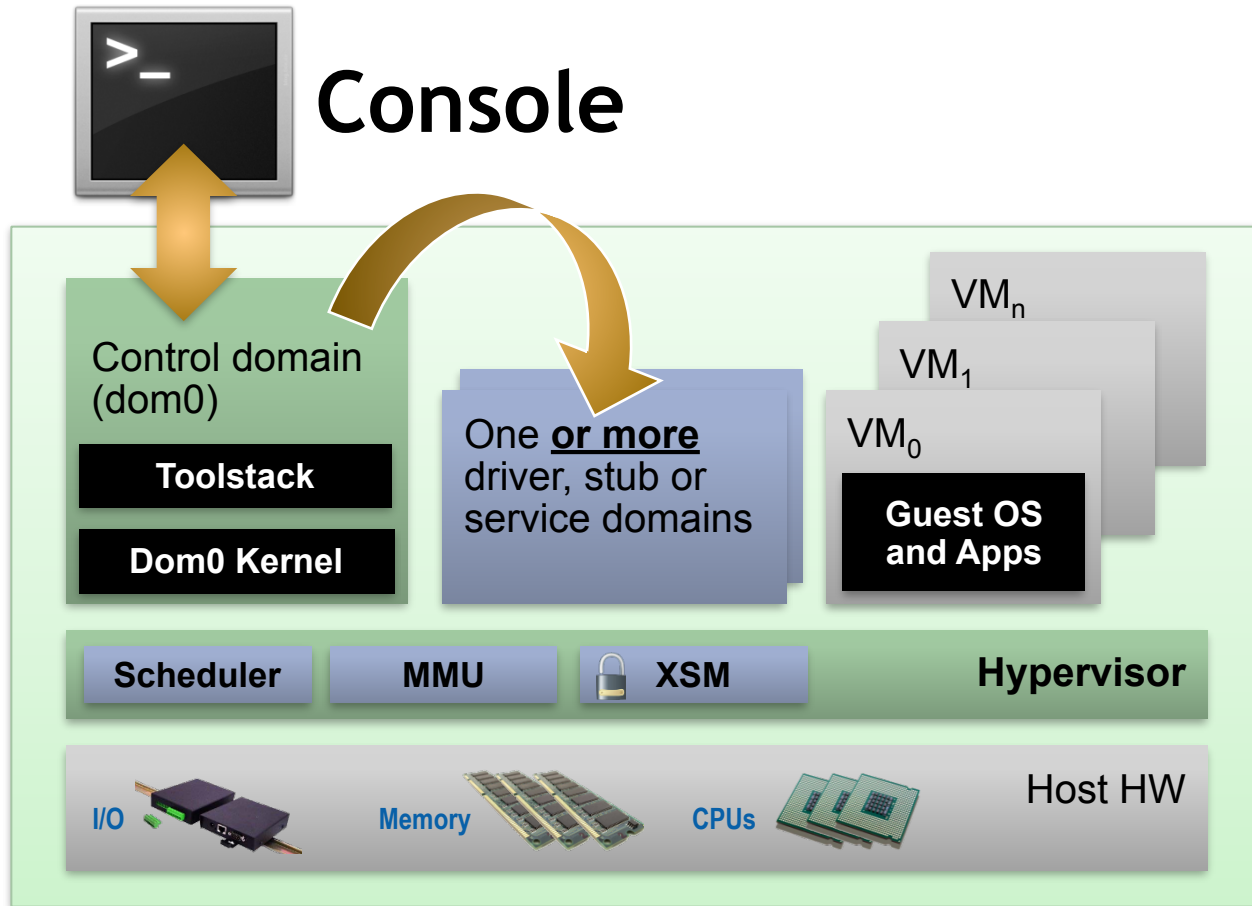
Guest Domains

- Your apps

■ Trusted Computing Base



Basic Xen Project Concepts: Disaggregation



■ Trusted Computing Base

Console

- Interface to the outside world

Control Domain aka [Dom0](#)

- Dom0 kernel with drivers
- Xen Project Management Toolstack

Guest Domains

- Your apps

Driver/Stub/Service Domain(s)

- A “driver, device model or control service in a box”
- De-privileged and isolated
- Lifetime: start, stop, kill



Xen Project: Types of Virtualization



Xen Project Virtualization Vocabulary

- **PV - Paravirtualization**
 - Hypervisor provides API used by the OS of the Guest VM
 - Guest OS needs to be modified to provide the API
- **HVM - Hardware-assisted Virtual Machine**
 - Uses CPU VM extensions to handle Guest requests
 - No modification to Guest OS
 - But CPU must provide VM extensions
- **FV - Full Virtualization (another name for HVM)**



Xen Project Virtualization Vocabulary

- **PVHVM - PV drivers on HVM**
 - Allows HVM guests to use PV disk and I/O drivers
 - No modifications to guest OS
 - Better performance than straight HVM
- **PVH - PV in HVM Container (new in 4.4)**
 - Almost fully PV
 - Uses HW extensions to eliminate PV MMU
 - Eventually best mode for CPUs with virtual H/W extensions

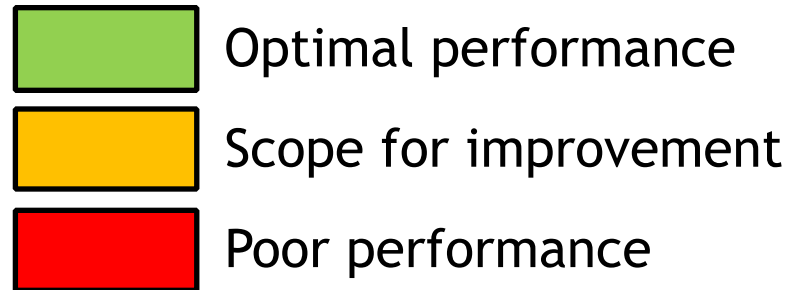


The Virtualization Spectrum

| | |
|----|------------------|
| 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 for disk & network | P | VS | VS | VH | |
| PVHVM | P | P | VS | VH | |
| PVH NEW Xen Project 4.4 | P | P | P | VH | } PV mode/domain |
| Fully Paravirtualized (PV) | P | P | P | P | |

The Virtualization Spectrum



| | 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 for disk & network | P | VS | VS | VH | |
| PVHVM | P | P | VS | VH | |
| PVH NEW Xen Project 4.4 | P | P | P | VH | } PV mode/domain |
| Fully Paravirtualized (PV) | P | P | P | P | |

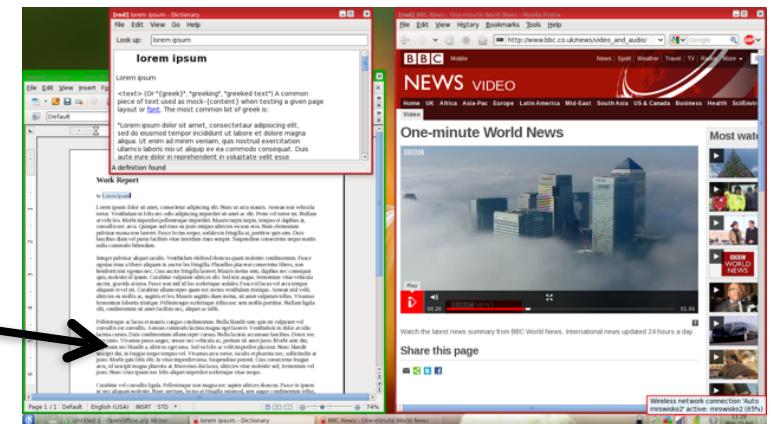
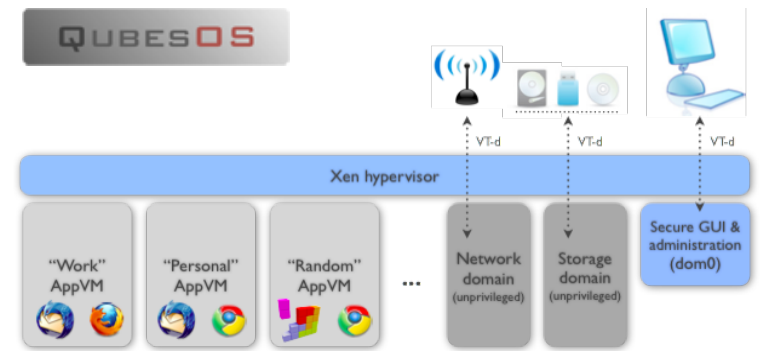
Disaggregation

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](#)"

Used today by [Qubes OS](#) and Citrix XenClient XT

Prototypes for XAPI



See qubes-os.org

Different windows run in different VMs

Benefits of Disaggregation

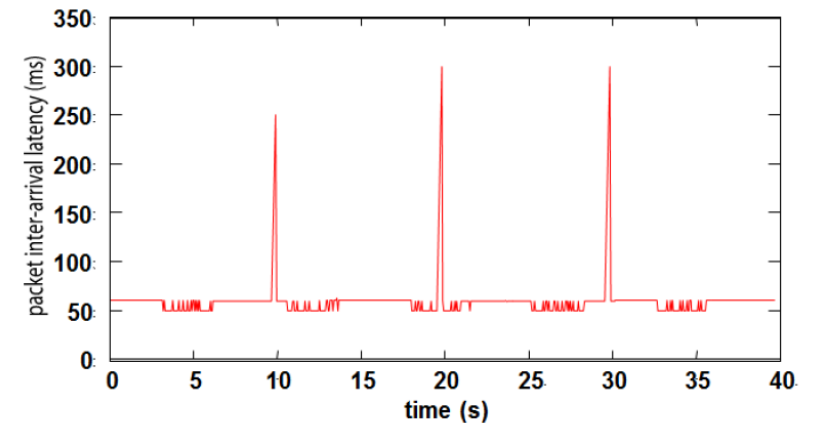
More Security

Increased serviceability and flexibility

Better Robustness

Better Performance

Better Scalability



Ability to safely restart parts of the system
(e.g. just 275ms outage from failed Ethernet driver)



Xen Project 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 Project Security Modules (or XSM) and FLASK
 - XSM is Xen Project equivalent of LSM (Linux Security Modules)
 - FLASK is Xen Project equivalent of SELinux
 - Developed, maintained and contributed to Xen Project by NSA
 - Compatible with [SELinux](#) (tools, architecture)
 - XSM object classes maps onto Xen Project features

[More info: http://www.xenproject.org/component/allvideoshare/video/latest/lfnw2014-advanced-security-features-of-xen-project-hypervisor.html](http://www.xenproject.org/component/allvideoshare/video/latest/lfnw2014-advanced-security-features-of-xen-project-hypervisor.html)



Xen Project Security Modules: FLASK

- What does FLASK provide?
 - Granular security
 - Can a guest domain talk with other guest domains?
 - Can a guest domain only communicate with the Control Domain?
 - Can a Guest domain have memory which cannot be read by the Control Domain?
 - What type of device model is used in this domain?
 - The ability to define multiple security roles on the domain level
 - User types can be defined and assign roles
 - Policy constraint logic

[More info: http://wiki.xenproject.org/wiki/Xen_Security_Modules:_XSM-FLASK](http://wiki.xenproject.org/wiki/Xen_Security_Modules:_XSM-FLASK)



ARM Hypervisor



Xen Project for ARM Servers

- Fully functional for ARM v7 & v8
- ARM v7:
 - Versatile Express, Arndale, Samsung Chromebook, Cortex A15, Allwinner A20/A31
- ARM v8: Fast Model, APM X-Gene “Mustang”

http://wiki.xenproject.org/wiki/Xen_ARM_with_Virtualization_Extensions



Xen Project + ARM = A Perfect Match

ARM SOC

Device Tree describes ...



I/O

GT

GIC
v2

2 stage
MMU

ARM Architecture Features for Virtualization

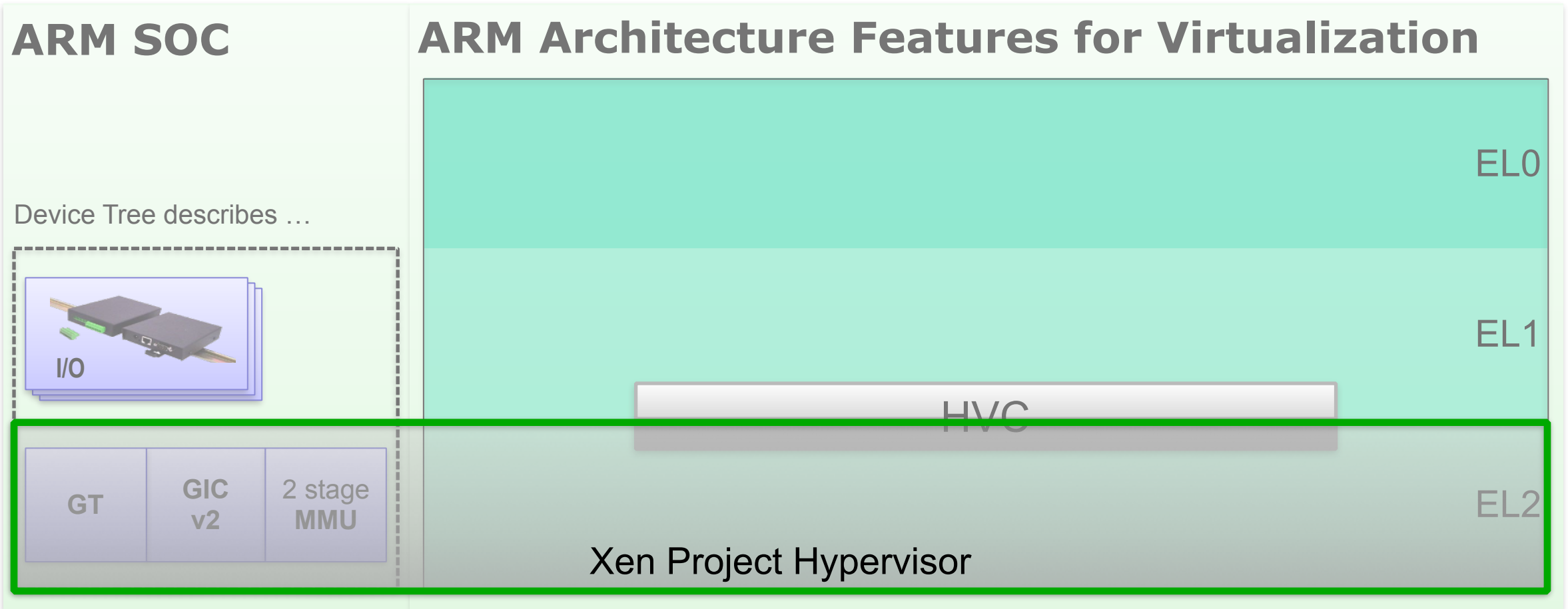
User mode : EL0

Kernel mode : EL1

Hypercall Interface HVC

Hypervisor mode : EL2

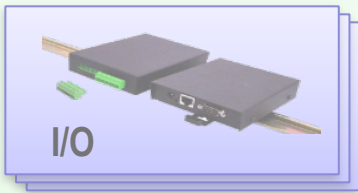
Xen Project + ARM = A Perfect Match



Xen Project + ARM = A Perfect Match

ARM SOC

Device Tree describes ...



GT

GIC
v2

2 stage
MMU

ARM Architecture Features for Virtualization

Any Xen Project Guest VM (including Dom0)

EL0

User Space

Kernel

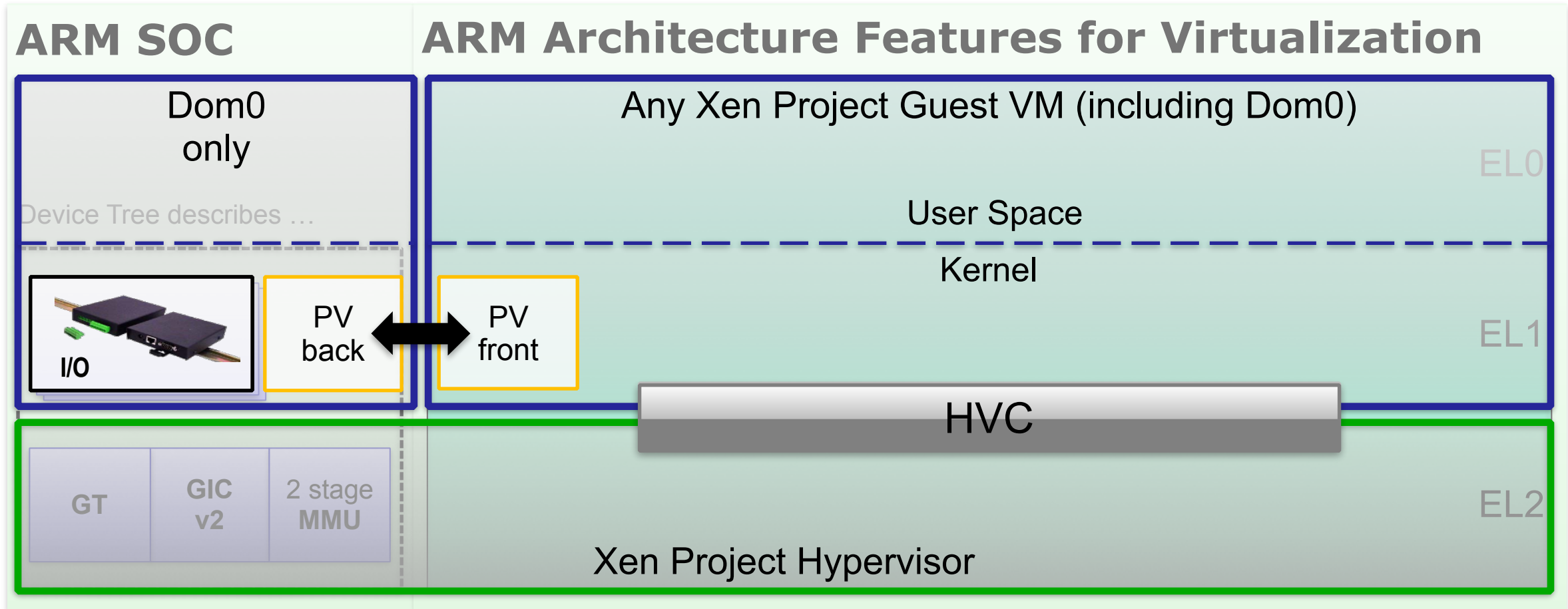
EL1

HVC



EL2

Xen Project Hypervisor

Xen Project + ARM = A Perfect Match



One mode to rule them all

-  Optimal performance
-  Scope for improvement

| | Disk and Network | Interrupts, Timers | Emulated Motherboard Legacy boot | Privileged Instructions and page tables | |
|-------------|------------------|--------------------|----------------------------------|---|-----------------|
| x86: PVHVM | P | P | VS | VH | HVM mode/domain |
| x86: PVH | P | P | P | VH | PV mode/domain |
| ARM v7 & v8 | P | VH | VH | VH | |



Code Size of x86 and ARM Hypervisors

| | | |
|--|----------------|--|
| X86 Hypervisor | 100K -120K LOC | Any x86 CPU |
| <i>ARM Hypervisor for mobile Devices</i> | <i>60K LOC</i> | <i>ARM v5 - v7 (no virtual extensions)</i> |
| <i>ARM Hypervisor for Servers</i> | <i>17K LOC</i> | <i>ARM v7+ (w/ virtual extensions)</i> |



Mirage OS



Library Operating Systems

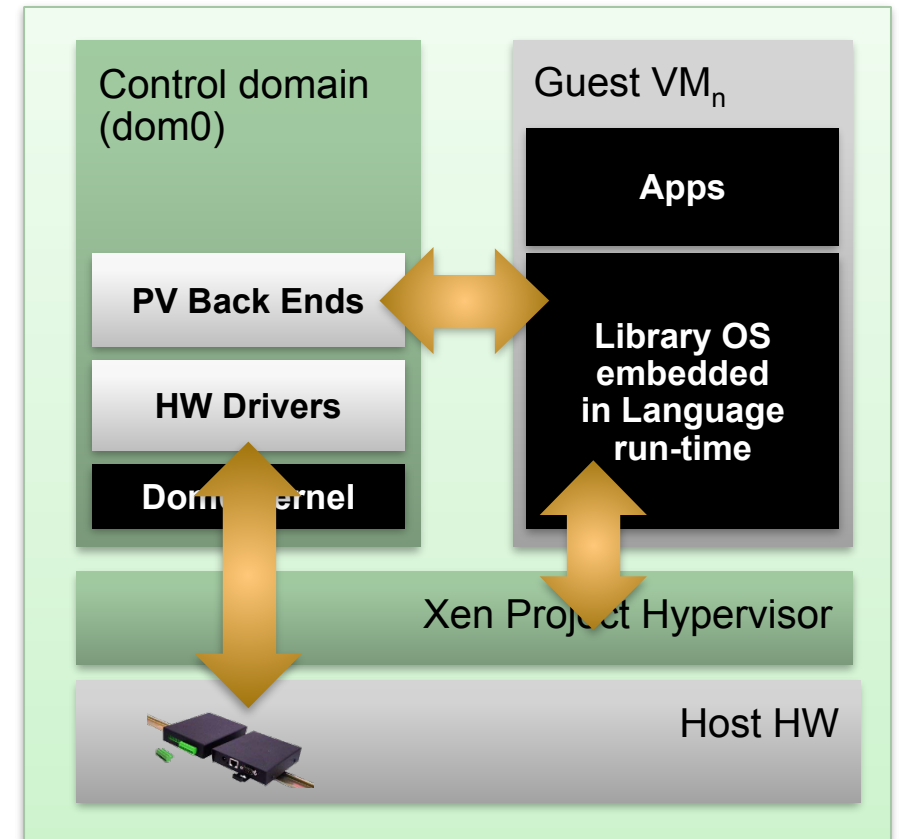
Application stacks only running on Xen Project APIs
Works on any Xen Project cloud or hosting service

Examples

- ErlangOnXen.org : Erlang
- HalVM : Haskell
- Mirage OS : Ocaml
- Osv: Java, C

Benefits:

- Small footprint
- Low startup latency
- Extremely fast migration of VMs



Mirage OS

- Part of the Xen Project incubator
- V2.0 Released July 2014
- Light and small like Docker, but with the full security of the Xen Project Hypervisor
- Clean-slate protocols implementations, e.g.
 - TCP/IP, DNS, SSH, Openflow (switch/controller), HTTP, XMPP

More info: <http://www.xenproject.org/developers/teams/mirage-os.html>



What's Next?



New in Xen Project 4.4 (April 2014)

- PVH mode is here!
- Updated and improved libvirt support
- Xen4CentOS: Xen Project for CentOS 6
- Experimental EFI support & nested virtualization
- Improved ARM, SPICE, GlusterFS support



See slides: <http://www.xenproject.org/component/allvideoshare/video/latest/lf-collaboration-summit-xen-project-4-4-features-and-futures.html>

Coming in Xen Project 4.5 (Dec 2014)

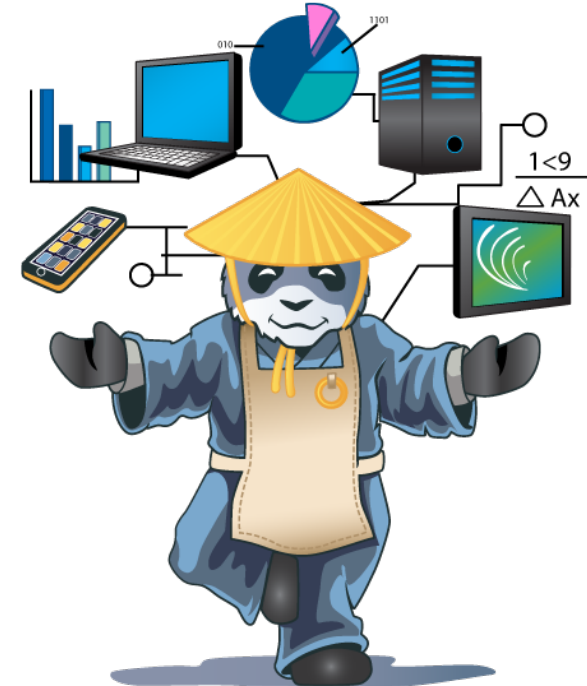
- PVH mode performance improvements
- More Mirage OS and unikernel support
- Even more ARM, libvirt improvements
- REMUS reworked (COLO still in development)
- And much, much more...



See status: http://wiki.xenproject.org/wiki/Xen_Project_Hypervisor_Roadmap/4.5

What's next (and already happening)

- Establish a shared test infrastructure
 - Most major contributors are duplicating effort
- Usability and better distribution integration
- More focus on downstreams
 - Examples: CloudStack and Xen Orchestra
- Xen Automotive
- XenGT (GPU Passthrough)
- Better Libvirt and virt-manager integration
 - Embed Xen Project more into the Linux ecosystem and provide benefits for the wider Linux community



Getting Started with Xen Project

- Document Days (monthly)
- Test Days (prior to release)
- Mailing Lists , IRC, Newsletter
- XenProject.org (sign up, it's free!)



**BE ZEN
HACK **

Hackathon: Expected Spring 2015

Developer Summit: Expected Summer 2015

User Summit: Expected Summer 2015





- **News:** blog.XenProject.org
- **Web:** XenProject.org
 - Help for IRC, Mailing Lists, ...
 - Stackoverflow-like Q&A
- **Wiki:** wiki.XenProject.org
- **Presentations & Videos:** see XenProject.org

Thank You!



Slides available under CC-BY-SA 3.0
From www.slideshare.net/xen_com_mgr

