XenSummit Asia

November 2-3, 2011
Seoul, Korea

Xen: the Past, Present and Exciting Future

Ian Pratt

Chairman of Xen.org,
SVP and Founder of Bromium

Sponsored by:

SAMSUNG & KOREA UNIVERSITY

Outline

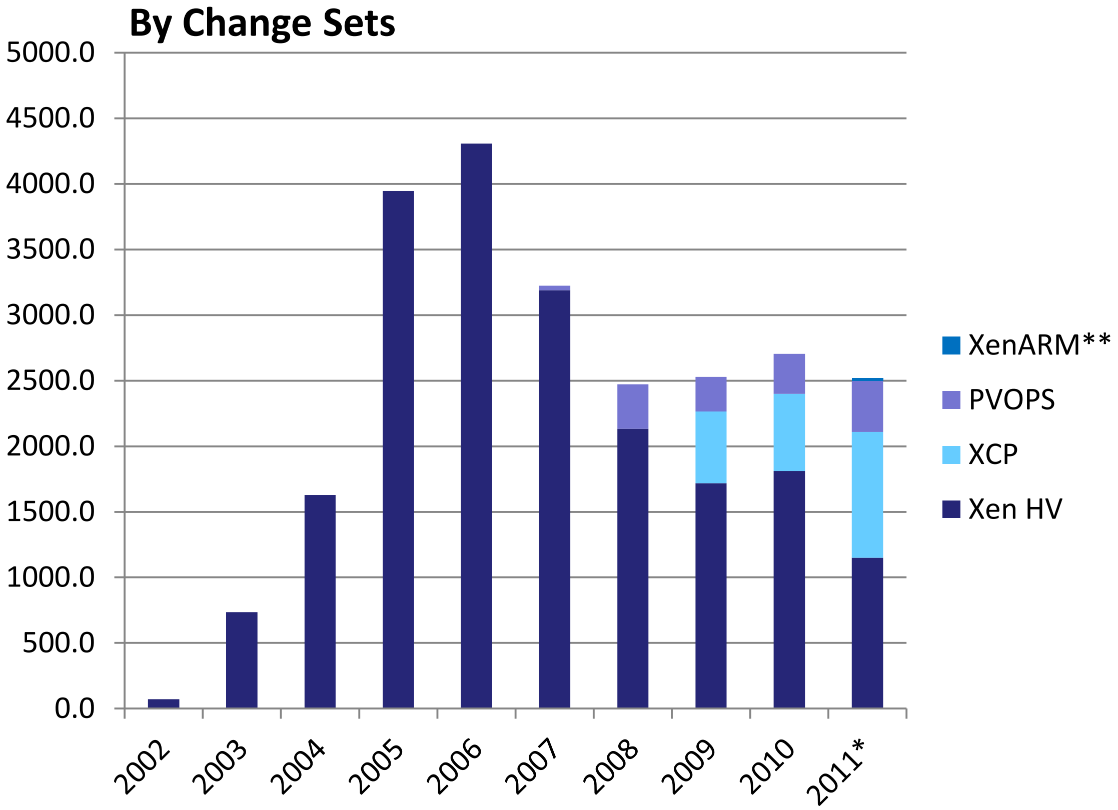
- Community Update
- Xen 4 Review
- Xen and the next wave of virtualization

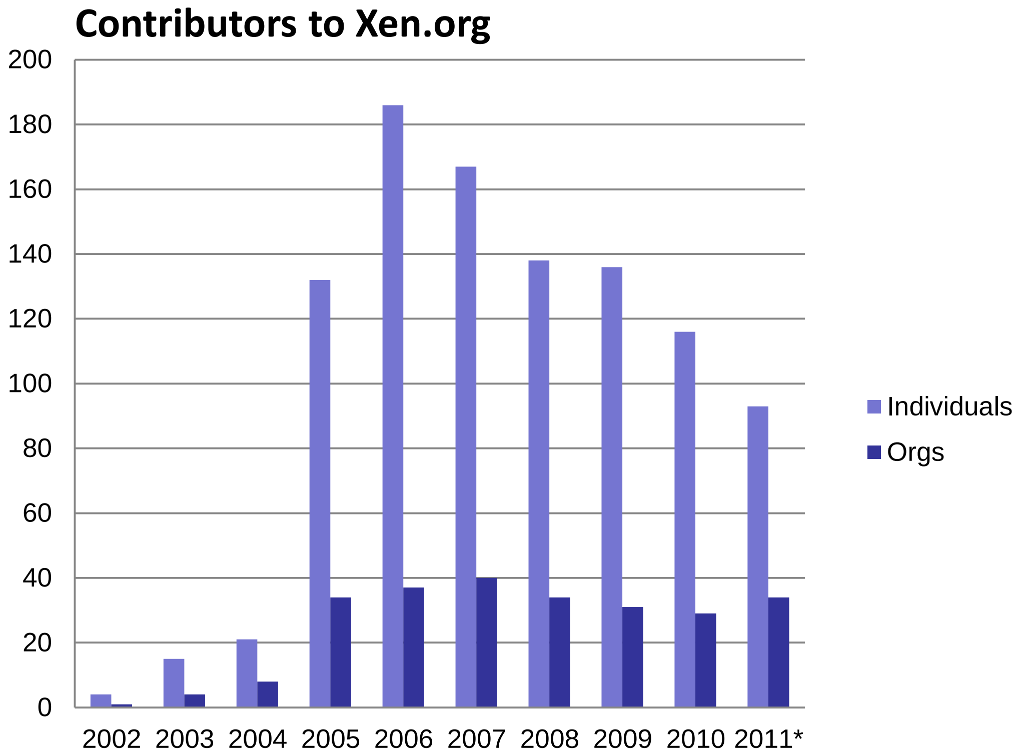
COMMUNITY UPDATE

2011 Highlights

- Inclusion of Xen into Linux 3 (and distros)
- New Initiatives:
  - Project Kronos
  - Xen.org Governance
  - Renewed focus on Xen for ARM
- Successful Community Initiatives
  - Documentation Day
  - Google Summer of Code
  - Hackathons: Cambridge (Citrix) and Munich (Fujitsu)
- Lars Kurth: (not so) new Community Manager

Contribution Statistics

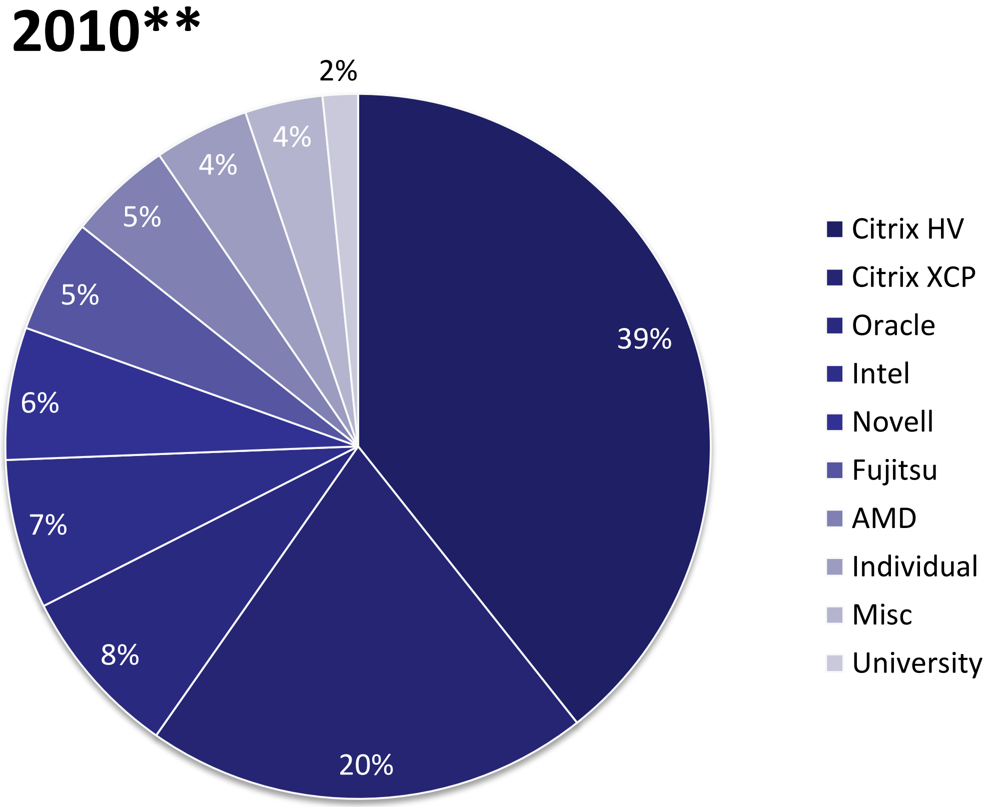

By Change Sets

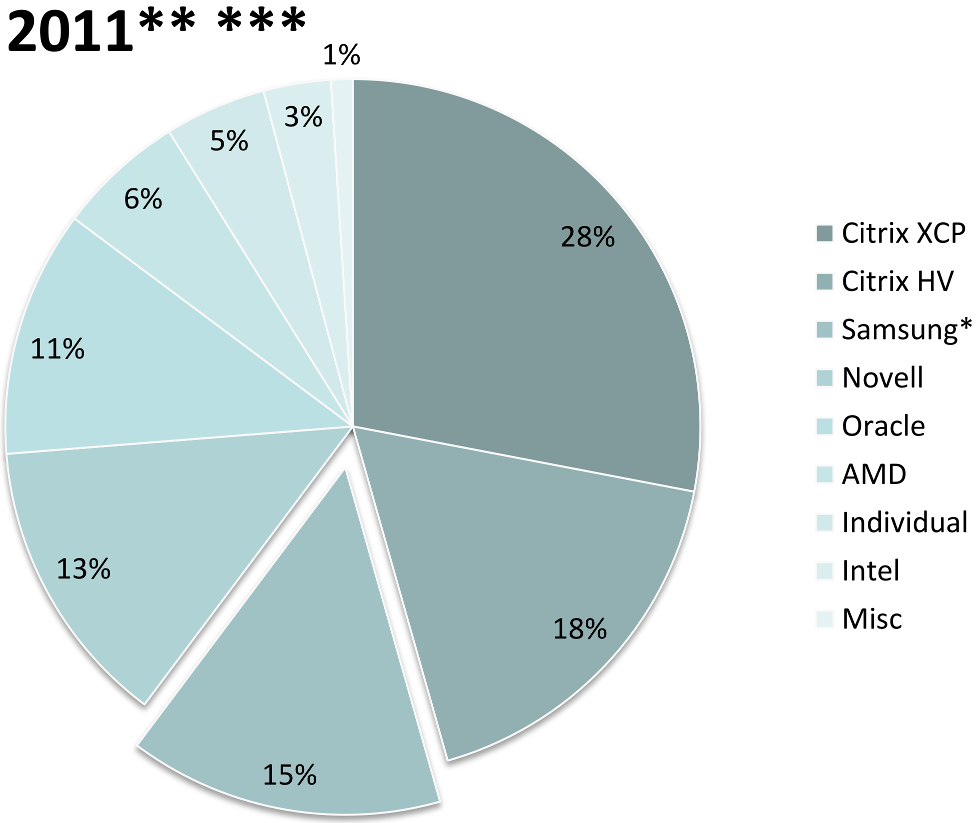

Contributors to Xen.org

*) End of Sept 2011

**) Activity on Development branch (not yet in xen-unstable)

2010 & 2011 Contributors (by KLOC)

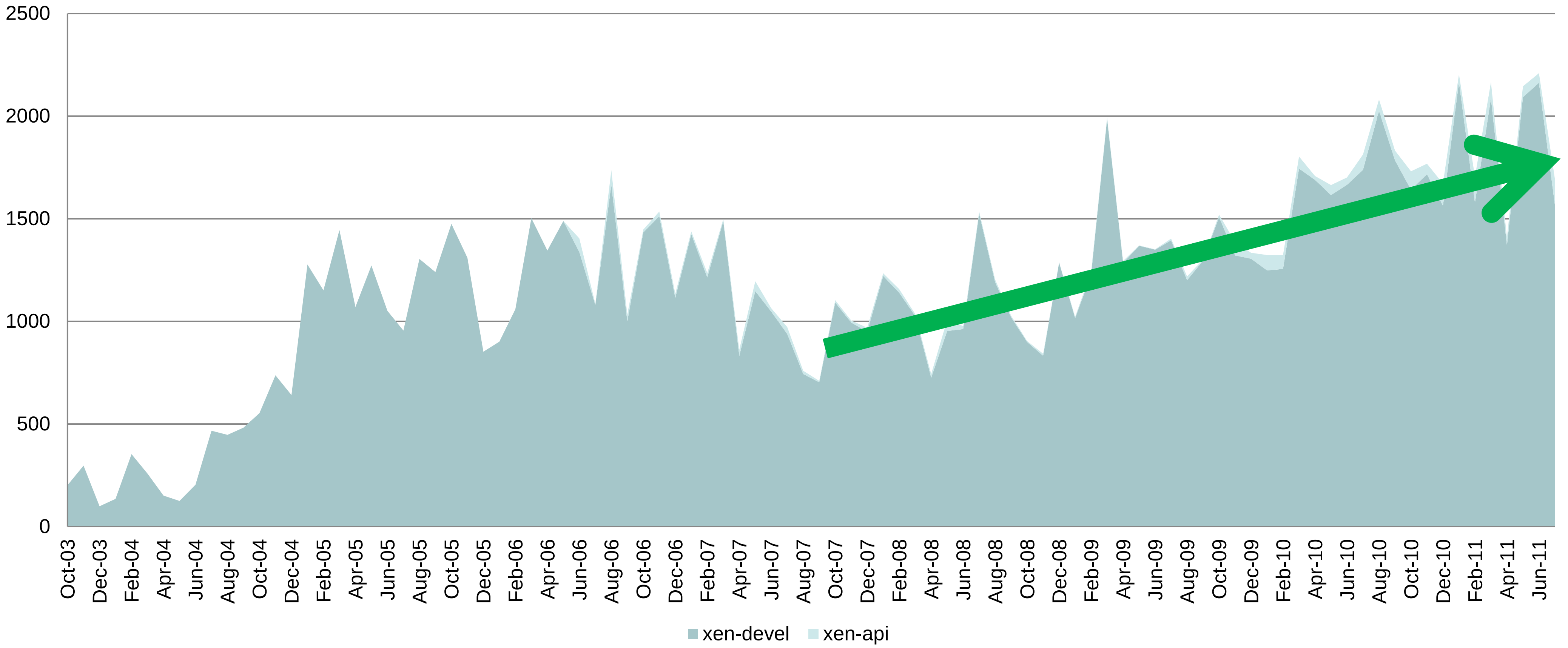




*) Activity on Development branch (not yet in xen-unstable)
**) Includes PVOPS
***) Until Sept 2011

Developer mailing list traffic

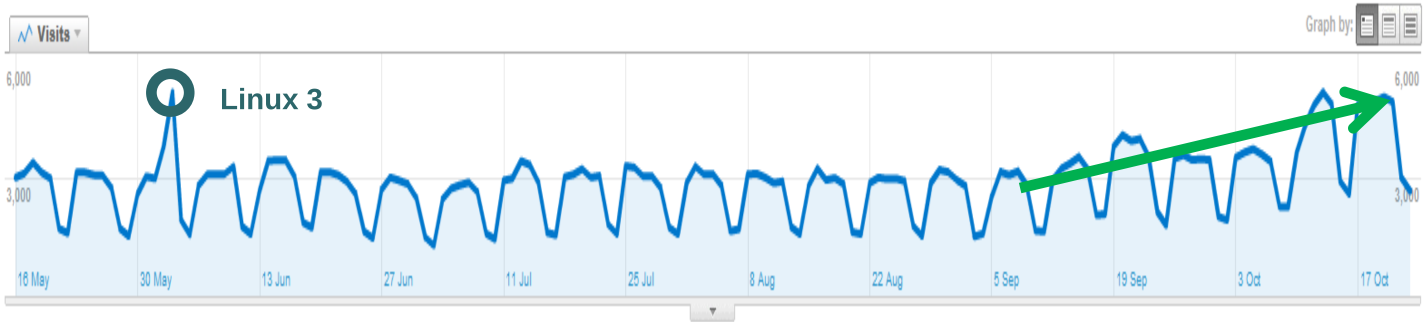
Conversations, excluding patches

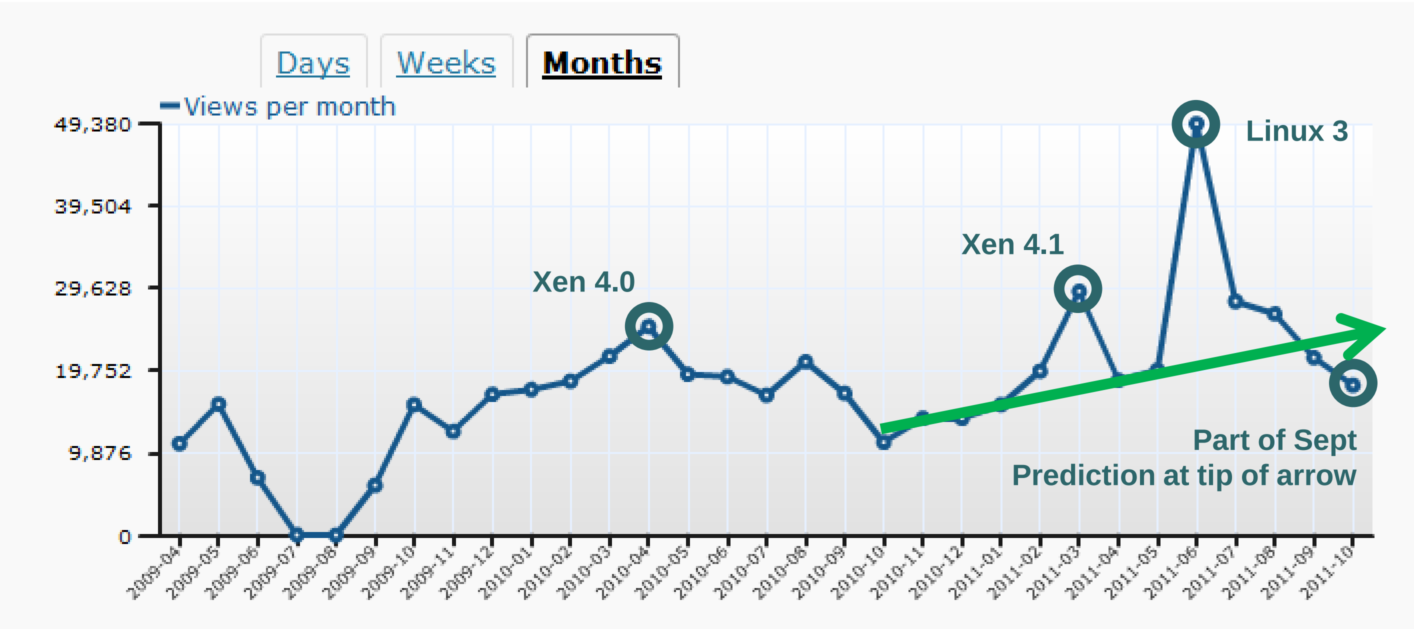


Formalized Governance

- How to contribute
  (had this for a long time, but was poorly documented)
- Election of Maintainers, Committers & Project Leads
  - Committer Election in September
  - Jan Beulich (Novell) : Committer on Xen HV project
    - 2009 : 107 patches changing 11746 lines of code
    - 2010 : 147 patches changing 7613 lines of code
    - 2011 : 130 patches changing 27377 lines of code (as of Sept)
- Project Lifecycle
  - Xen HV & XCP migrated to new lifecycle

Xen.org Web site Activity





Website Traffic/Day:

Notable traffic increase since September (almost double)

Coincides with changes to content and new content!

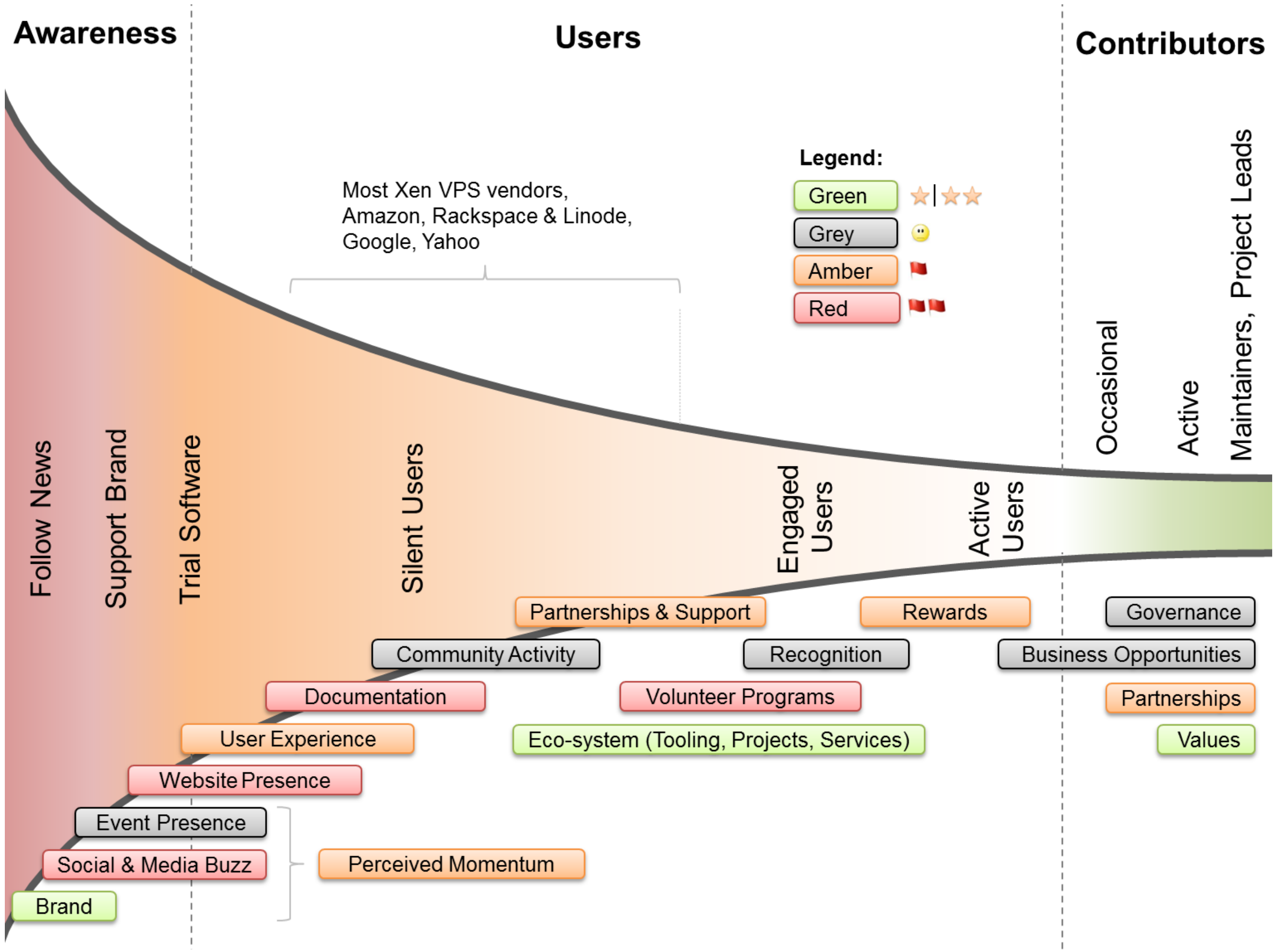
Blog Traffic/Month:

100% average increase of traffic compared to one year ago

Product and project news roundup

- Xen support in Linux 3
  - More in Linux 3.1 (and subsequent releases)
- Xen support (DomU and Dom0) back in Linux distros
  - Debian
  - Ubuntu 11.10
  - Fedora 16
- Recent product releases that distribute Xen
  - Oracle VM 3.0
  - XenServer 6.0
  - XenClient 2.0
  - Beta’s of QubesOS and RC’s of openSuse 12.1

Where do we need to improve?



Awareness

Users

Contributors

Legend:

Green	★|★★
Grey	😐
Amber	🚩
Red	🚩🚩

Most Xen VPS vendors, Amazon, Rackspace & Linode, Google, Yahoo

Follow News

Support Brand

Trial Software

Silent Users

Engaged Users

Active Users

Occasional

Active

Maintainers, Project Leads

Partnerships & Support

Rewards

Governance

Community Activity

Recognition

Business Opportunities

Documentation

Volunteer Programs

Partnerships

User Experience

Eco-system (Tooling, Projects, Services)

Values

Website Presence

Event Presence

Social & Media Buzz

Perceived Momentum

Brand

Focus for 2012

- New website
- Better media presence
- More focus on users (individual & commercial)
- More and better documentation
- New benchmarks, feature comparisons, etc.
- Formalize volunteer activities such as “documentation day”

2011 & 2012 Event Calendar

- LinuxCon Brazil, Sao Paulo, Nov 17-18
- USENIX Lisa, Boston, Dec 4-9
- Planning to co-locate XenSummit NA with LinuxCon (LinuxCon NA, San Diego, Aug 27-28)
  - Not yet finalized, but should be soon
- OSCON, Portland, July 16-20

Community Summary

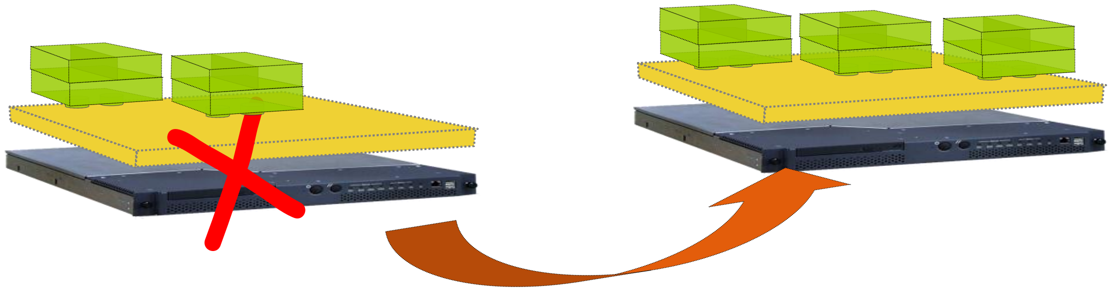
- The Xen Developer community is healthy for a 10 year old project
- The inclusion of Xen support into Linux 3.x has made a big impact
  - Getting questions by many new users
  - Building new and productive relationships with many people in the Linux and BSD communities
- Up to now Xen.org was almost exclusively looking after developers
  - Successful open source communities bring their users and developers together
  - Xen.org needs to focus on
    - Reconnecting to its users
    - On making it easy to get started with Xen and engage new users
      - Many opportunities in Cloud Projects
    - On better communicating the Xen advantages
- Everybody can help with this!

XEN 4.1 REVIEW

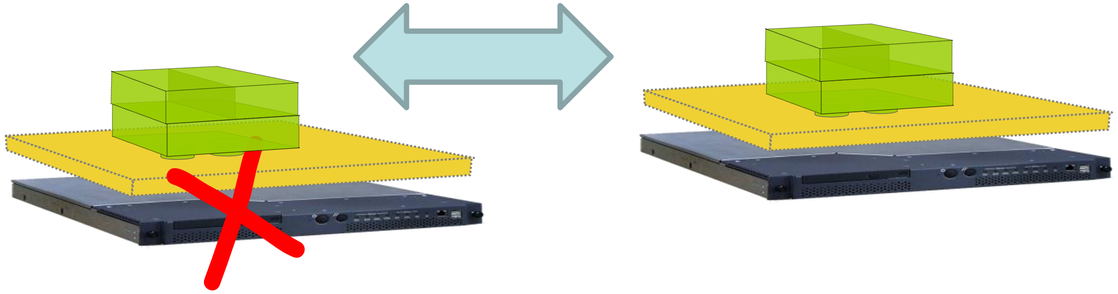
Xen 4.1 Release – 21 March 2011

- 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

Hardware Fault Tolerance



- *Restart-HA* monitors hosts and VMs to keep apps running



- *Hardware Fault Tolerance* with deterministic replay or checkpointing

Xen’s Software-Implemented Hardware Fault Tolerance enables true High Availability for unmodified applications and operating systems

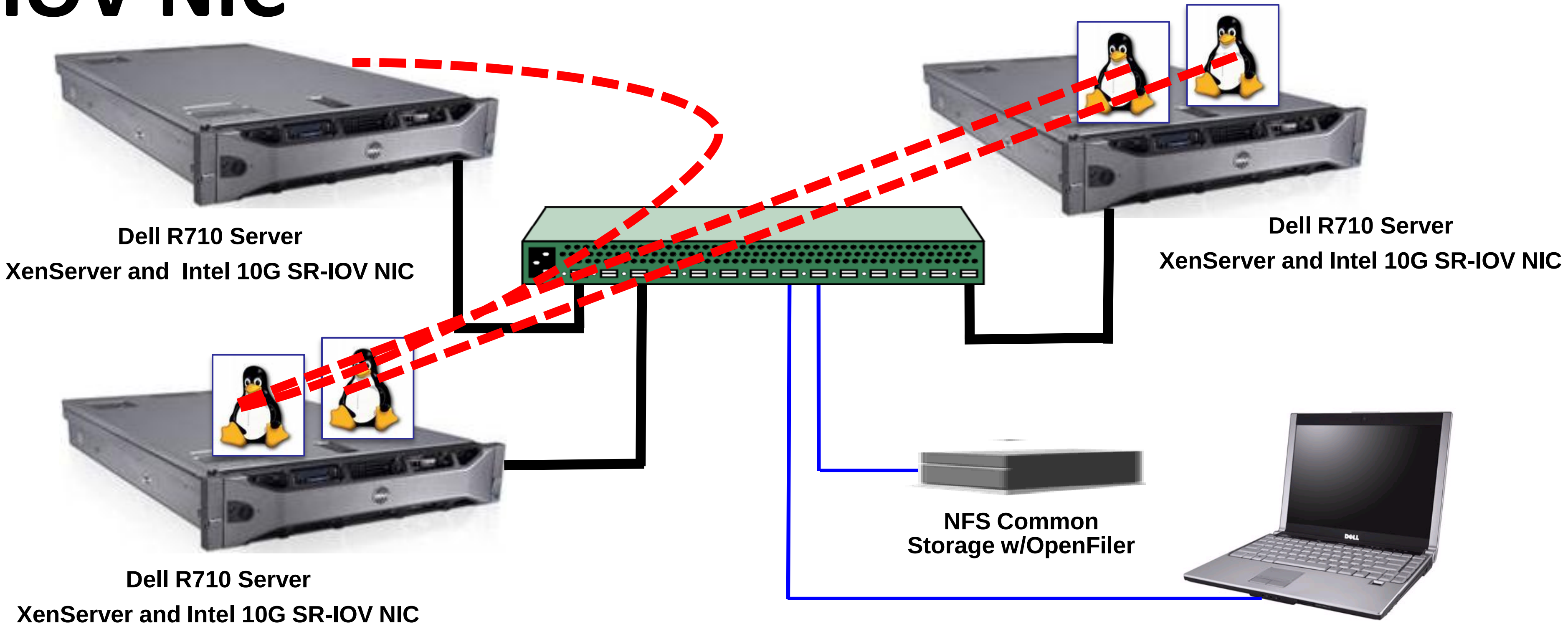
Hardware Fault Tolerance

- University of British Columbia’s “Remus” project is now in Xen 4
- Smart checkpointing approach yields excellent performance
  - VM executes in parallel with checkpoint transmission, with all externally visible state changes suppressed until checkpoint receipt acknowledged
  - Checkpoints delta compressed
- Checkpointing possible across wide-area, even for multi-vCPU guests

Enhanced SR-IOV

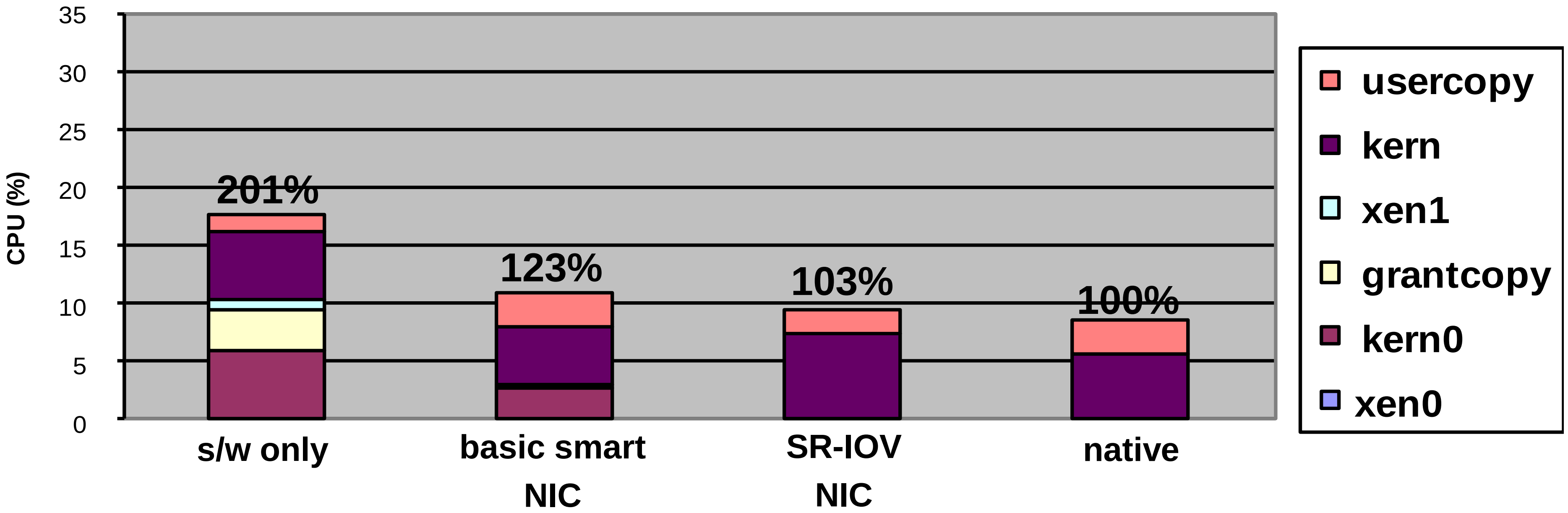
- SR-IOV: Single Root IO Virtualization
  - Virtualization friendly IO devices
- High performance, high efficiency, low latency
- Enables even the most demanding applications to now be virtualized
- Compatible with live relocation via hotplug of acceleration driver “plugin” module
  - Retain primary benefit of physical hardware abstraction

SR-IOV NIC


Dell R710 Server
XenServer and Intel 10G SR-IOV NIC
Dell R710 Server
XenServer and Intel 10G SR-IOV NIC
Dell R710 Server
XenServer and Intel 10G SR-IOV NIC
NFS Common
Storage w/OpenFiler

- 80 Gb/s bi-directional aggregate throughput between 4 VM pairs
- Low latency, High CPU efficiency
- Live relocation between hosts - Even hosts with different NICs

Network Performance



- New Smart NICs reduce CPU overhead substantially
- Care must be taken with SR-IOV NICs to ensure benefits of VM portability and live relocation are not lost
- Need for an industry standard for “driver plugins”

THE NEXT VIRTUALIZATION WAVE

Security will drive the Next Wave of Virtualization

- Security is key requirement for Cloud
- Security is the primary goal of virtualization on the Client
  - Desktop, Laptops, Smart Phones, etc
- 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
- Enables “out-of-band” management and policy enforcement
  - Malware detection, remote access, image update, backup, VPN, etc.

Xen Introspection API

- Allows a suitably privileged VM to monitor and control the execution of another VM
  - Interpose on disk and network IO path
  - Mark VM memory as immutable, no-execute etc
  - Inspect/modify CPU and memory state
- Enables robust anti-malware, anti-root kit
  - Cannot be disabled/bypassed by guest VM

⇨ Virtualized can be more secure than physical!

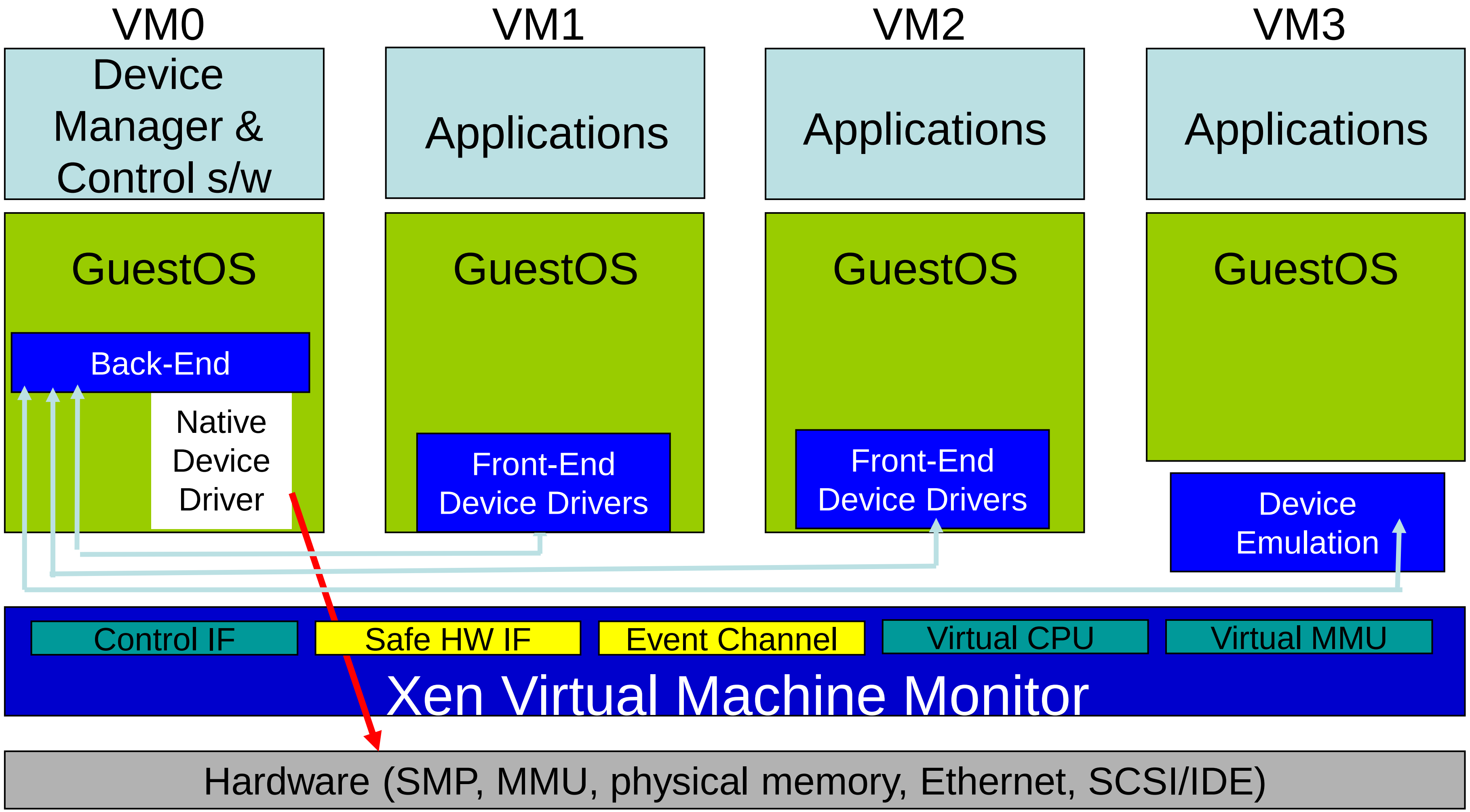
Secure Isolation

- Use good software engineering practice
  - Thin hypervisor: minimize code running with privilege
  - Disaggregate and de-privilege functionality into dedicated Service VMs
  - Narrow interfaces between components
  - Hypervisors are simpler than OSes, simpler than OS kernels
  - Use modern high-level languages where possible
- New hardware technologies help
  - VT-x, VT-d, EPT: reduce software complexity, enhanced protection
  - TPM/TXT: Enable Dynamic Root of Trust

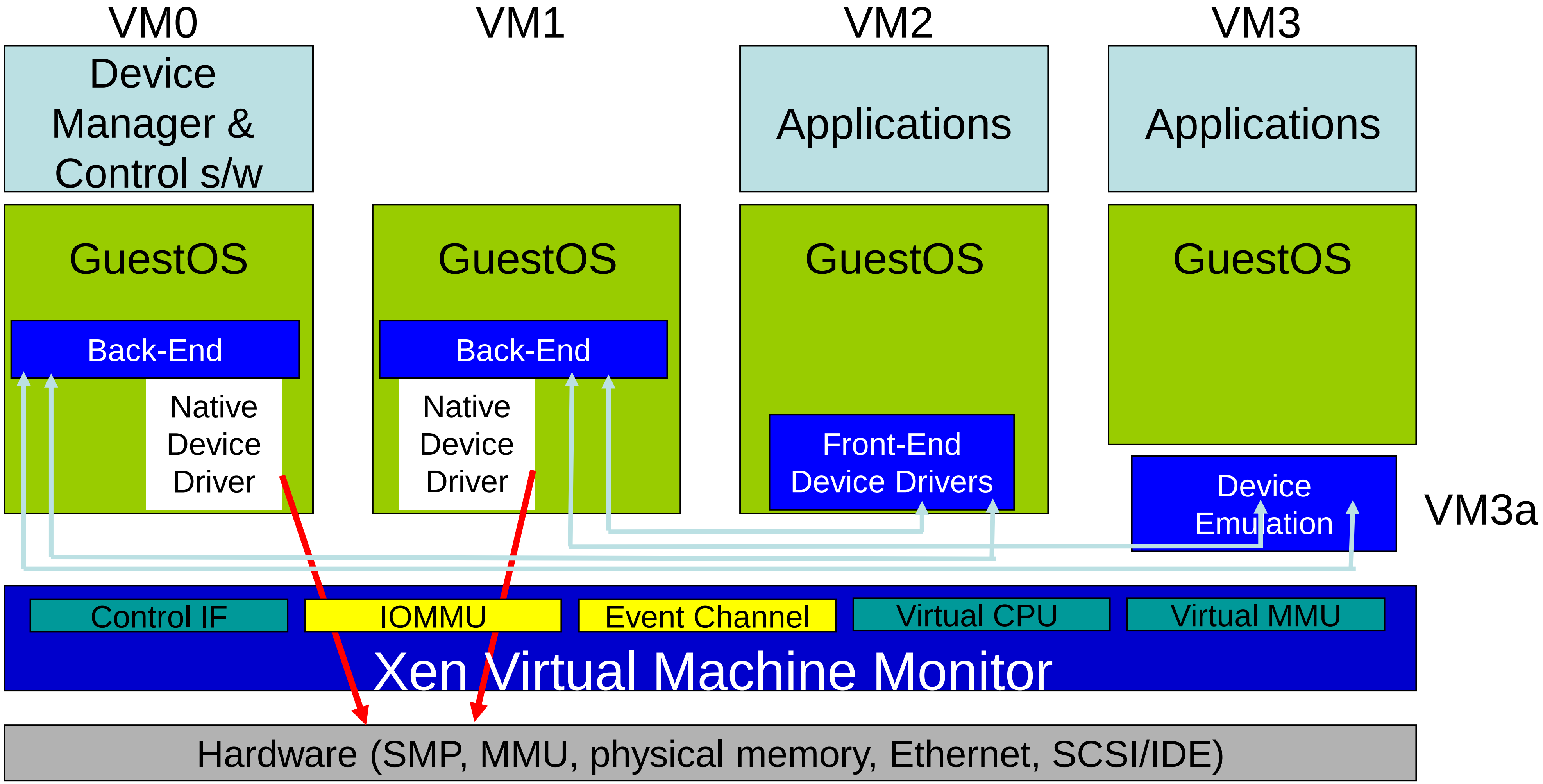
XenClient XT / Qubes OS

- 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 / SElinux
- Measured Launch (TXT)

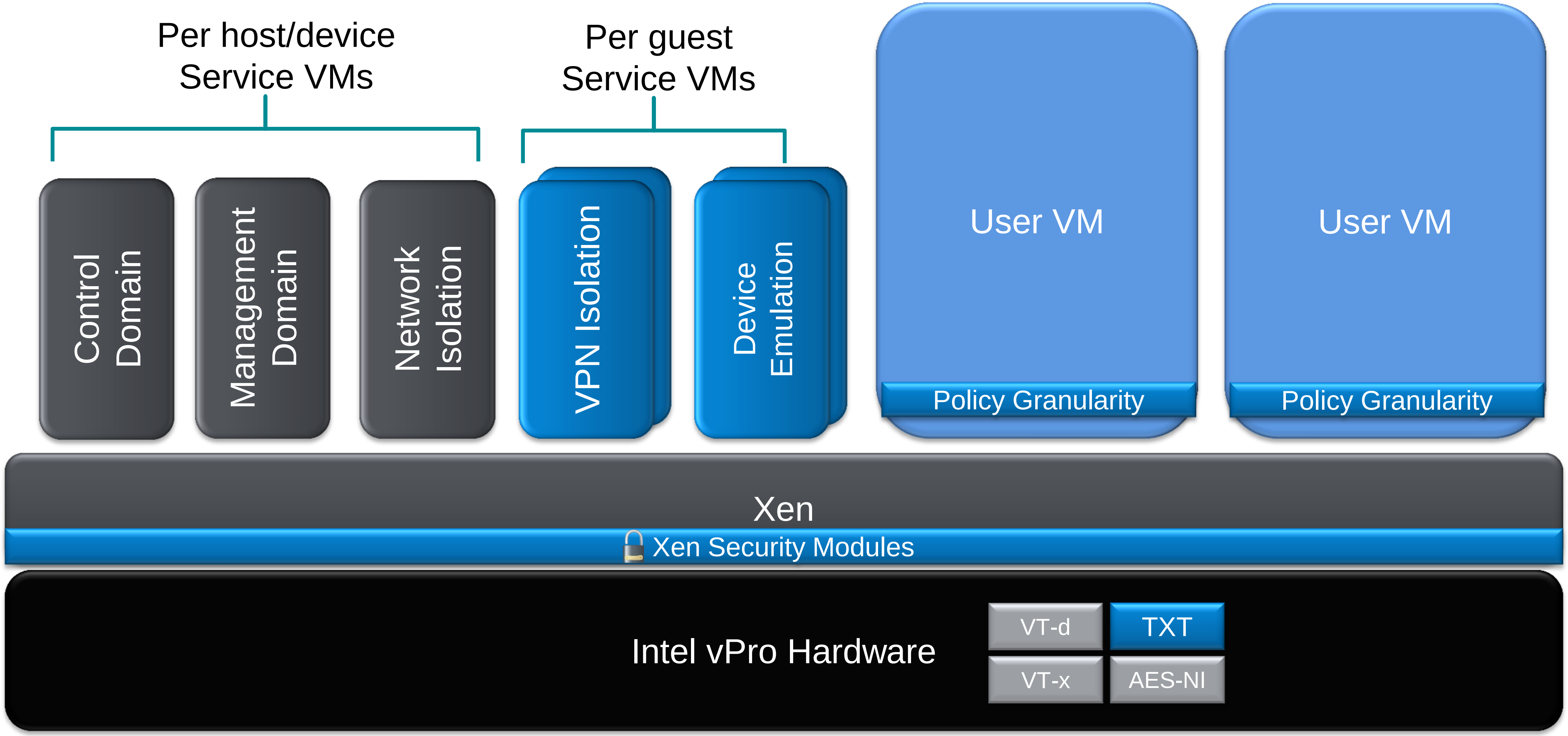
Typical Xen Configuration



Xen Driver Domains



Advanced XenClient Architecture

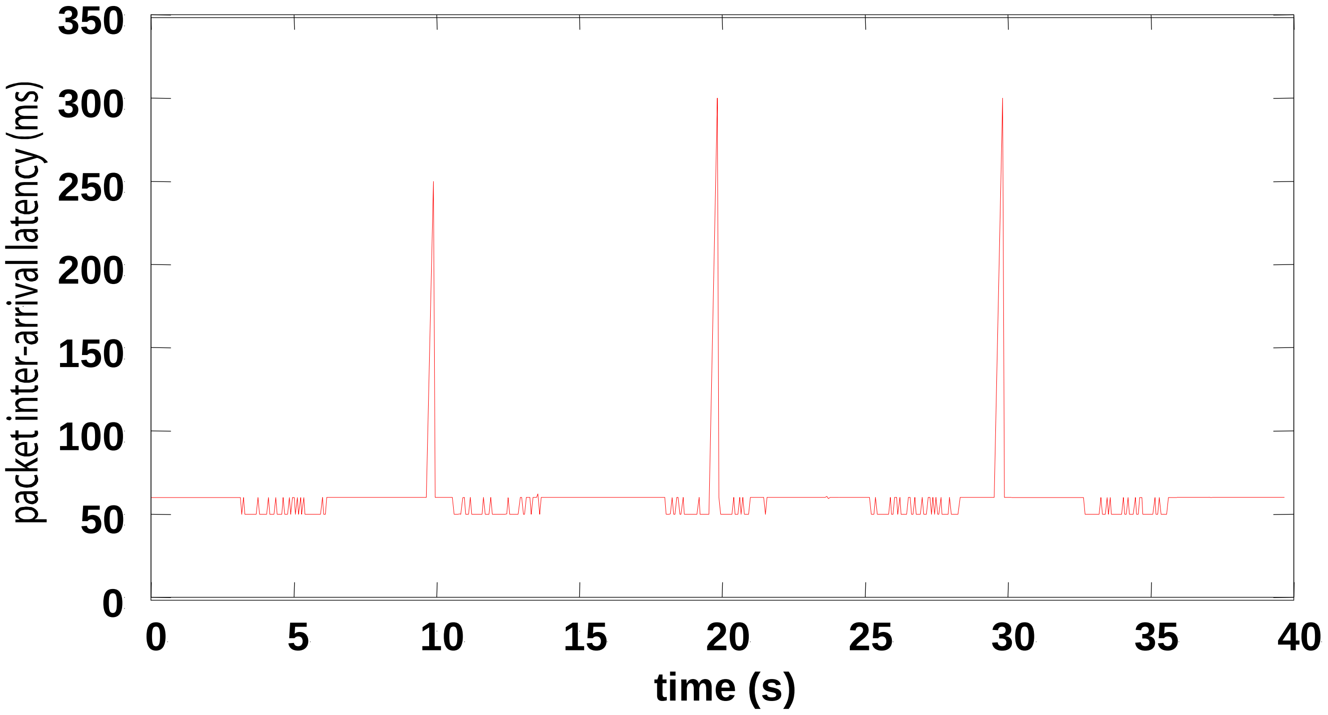


Disaggregation

- Unique benefit of the Xen architecture:
- Security
  - Minimum privilege; Narrow interfaces
- Performance
  - Lightweight e.g. minios directly on hypervisor
  - Exploit locality – service VMs see a subset of the machine, run close to resources with which they interact
- Reliability
  - Able to be safely restarted

Isolated Driver VMs for High Availability

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



Proposal

- We should strive to get all Xen products and deployments to take full advantage of the Xen architecture
- We need to make this much easier!
- Proposal: define and maintain a reference architecture and implementation that embodies best practice recommendations

Reference Architecture

- Define using new technologies
  - Latest stable Xen
  - Linux 3.x pvops
    - Optimization effort required
  - Libxl control stack
    - For easy consumption by other vendor tool stacks

Target Features

- Network restart-able driver domains
  - Integrated OpenFlow vswitch
- Storage restart-able driver domains
  - Also allows easier deployment of new storage options e.g. vastsky, ZFS
- Qemu emulation domains
- Xen Security Modules
- Measured Launch via TXT
- Roadmap for enhanced security and performance features
  - E.g. the SR-IOV network plugin / vswitch architecture

Implementation

- Need an initial reference implementation
  - Easily consumable by users
- XCP could fulfil this role
  - Showcase latest Xen technologies
  - Optimized for OpenStack
- Aim to be as kernel/toolstack etc agnostic to allow easy adoption by all vendors

Summary

- Xen project continues to thrive!
  - Great success in Cloud and Client
- Key architectural security, reliability and performance benefits that are unique to Xen
  - We need to do a better job of getting the message out!
  - We need to do a better job of actually taking advantage of the benefits in all Xen products