QEMU 2.0 and Beyond

CloudOpen 2013

Anthony Liguori <anthony@codemonkey.ws>

About

- QEMU is a fast full system simulator and virtualization engine

- QEMU is Open Source hardware emulation
  - KVM
  - Xen
  - Android SDK (fork)
  - VirtualBox (fork)
  - Just about every embedded SDK out there

Quick History

- Started in 2003 by Fabrice Bellard
  - Author of FFMPEG, JSLinux, and lots of other cool things
  - Portable Just In Time (JIT) translation engine for cross architecture emulation
  - Quickly grew system emulation
    - Starting with PC hardware
- Has been a grass roots, quiet community

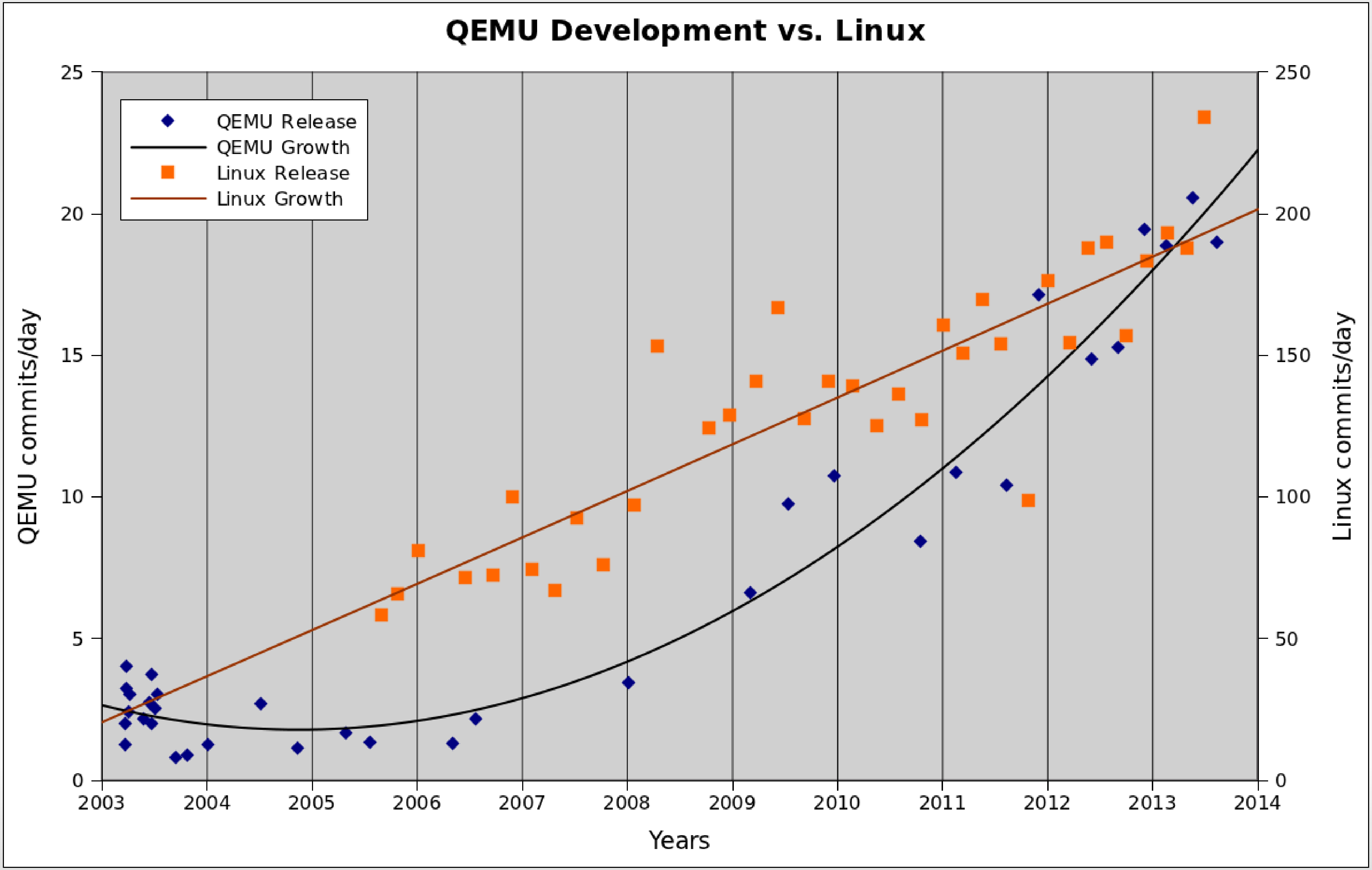
Evolution of QEMU

- Linux user emulation
- System emulation
- Replace dyngen → TCG
- Virtualization support
- Management API
- Block layer
- ...

Growth of the Community

- 10+ years of community building
- Roller coaster ride
- Inclusiveness
  - Wildly different features and missions
  - Rich community
  - Extremely complex command line
  - There be dragons

Growth of the Community



Forks and Merges

- Not always positive
  - Minor forks like qemu-kvm and qemu-dm
  - A few cases where major forks almost happened
- Tremendous effort merging forks back
  - Forks proved importance of compromise

Development Process

- Hierarchical maintainership
  - 40+ submaintainers
  - 250+ contributors
- Two month development cycle, one month stabilization period
- Major releases every 2 years, minor releases every 3 months

Features

- QEMU is the front line for Cloud
  - Xen HVM and all KVM guests
  - Primary interface that the guests communicate with is hardware
- The Linux Kernel unites all distributions
  - QEMU unites the Open Cloud

Features – Virtual I/O

- VirtIO
  - High speed paravirtual I/O framework
  - Designed like hardware
  - Network, disk, serial, hwrng, balloon, …
  - Undergoing standardization via OASIS
- Emulated I/O still improving
  - Improving support for VMware devices and more

Features - Graphics

- VNC and Spice support for remoting
  - Javascript clients available
  - Native WebSockets support
- Virgl
  - 3D graphics for guests based on VirtIO
  - Still a research project
  - Very promising

Features - Storage

- Convergence around qcow2
  - New modes and extension mechanism
- Improved support for snapshots
- virtio-blk dataplane
  - 95% of bare metal performance on large storage array

Features - Migration

- Convergance algorithm
  - Must race guest to complete migration
- New techniques to win against guest
  - XBLRE – Compression
  - RDMA – Raw performance
  - Guest delay – Cheat

Features - Migration

- Live block copy
  - Cloud loves local storage
  - Migration traditionally requires shared storage
  - Live block copy allows movement of local storage
- Live update
  - Reduce scheduled downtime by efficiently performing localhost migration
  - Potential to combine with kexec for full system update

Features - Managability

- QEMU Monitor Protocol (QMP)
  - JSON based RPC
- Formally specified in a schema language
- Support for commands and notification
- Rigid compatibility guarantees

Features - Security

- Virtualized hwrng
  - Provide better entropy to guests
- Layered security model
  - Unprivileged
  - Mandatory Access Control via SELinux
  - Sandboxed using seccomp mode 2

Features - Core

- QEMU Big Lock
  - Introduction of VSMP mirrors Linux kernel
- Systematic break up of big lock to enable better scalability
- Have used many tricks to avoid it this long
- Unlike Linux, skipping ahead to RCU

Features – Predictions

- Command line interface will be overhauled
  - Git style CLI
- GTK GUI will be expanded for desktop usage
- Storage layer will add RAID and rely less on Linux kernel
- We will solve migration backwards compatibility

QEMU in the Cloud - Consumer

- Should we care about the virtualization layer in the Cloud?
- Open Virtualization prevents the Cloud from becoming a Walled Garden
- Guests created on QEMU are portable across virtualization implementations
  - Not true of proprietary hypervisors

QEMU in the Cloud - Vendors

- Cloud is about much more than virtualization
  - Why invest is reinventing the wheel?
- Open Source is auditable
  - Recent news events make this even more critical
- Ability to contribute to direction of technology

Questions

- Questions

Get Involved

- http://wiki.qemu.org/Contribute/SubmitAPatch
- qemu-devel@nongnu.org