Xen Project 4.4: Features and Futures

Russell Pavlicek

Xen Project Evangelist, Citrix Systems

Xen Project 4.4 Features

Improved Flexibility in Driver Domains

- Linux driver domains used to rely on udev events in order to launch backends for guests
  - Dependency on udev is replaced with a custom daemon built on top of libxl
  - Provides greater flexibility in order to run user-space backends inside of driver domains
  - Example of capability: driver domains can now use Qdisk backends, which was not possible with udev

Improved Event Channel Scalability

- Event channels are para-virtualized interrupts
- Previously limited to either 1024 or 4096 channels per domain
  - Domain 0 needs several event channels for each guest VM (for network/disk backends, qemu etc.)
  - Practical limit of total number of VMs to around 300-500 (depending on VM configuration)

Improved Event Channel Scalability (2)

- New FIFO-based event channel ABI allows for over 100,000 event channels
  - Improve fairness
  - Allows for multiple priorities
  - The increased limit allows for more VMs, which benefits large systems and cloud operating systems such as MirageOS, ErlangOnXen, OSv, HalVM

Experimental PVH Support

- PVH mode combines the best elements of HVM and PV
  - PVH takes advantage of many of the hardware virtualization features that exist in contemporary hardware
- Potential for significantly increased efficiency and performance
- Reduced footprint in Linux and FreeBSD

Tech Preview of Nested Virtualization

- Nested virtualization provides virtualized hardware virtualization extensions to guests
  - Can now run Xen Project, KVM, VMWare or HyperV inside of a guest for debugging or deployment testing
  - Also allows Windows 7 "XP Compatibility mode"
  - Tech Preview not yet ready for production use, but has made significant gains in functionality and reliability
- More information on nested virtualization: see Xen nested

Improved Support for SPICE

- SPICE is a protocol for virtual desktops which allows a much richer connection than display-only protocols like VNC
- Added support for additional SPICE functionality, including:
  - Vdagent
  - clipboard sharing
  - USB redirection

GRUB 2 Support of Xen Project PV Images

- In the past, Xen Project software required a custom implementation of GRUB called pvgrub
- The upstream GRUB 2 project now has a build target which will construct a bootable PV Xen Project image
  - This ensures 100% GRUB 2 compatibility for pvgrub going forward

Indirect Descriptors for Block PV Protocol

- Modern storage devices work much better with larger chunks of data
- Indirect descriptors have allowed the size of each individual request to triple, greatly improving I/O performance when running on fast storage technologies like SSD and RAID
- This support is available in any guest running Linux 3.11 or higher (regardless of Xen Project version)

Improved kexec Support

- kexec allows a running Xen Project host to be replaced with another OS without rebooting
  - Primarily used execute a crash environment to collect information on a Xen Project hypervisor or dom0 crash
- The existing functionality has been extended to:
  - Allow tools to load images without requiring dom0 kernel support (which does not exist in upstream kernels)
  - Improve reliability when used from a 32-bit dom0
  - kexec-tools 2.0.5 or later is required

Improved XAPI and Mirage OS support

- XAPI and Mirage OS are sub-projects within the Xen Project written in OCaml
- Both are also used in XenServer (see http://xenserver.org/) and rely on the Xen Project OCaml language bindings to operate well
- These language bindings have had a major overhaul
  - Produces much better compatibility between XAPI, Mirage OS and Linux distributions going forward

Experimental Support for Guest EFI boot

- EFI is the new booting standard that is replacing BIOS
  - Some operating systems only boot with EFI
  - Some features, like SecureBoot, only work with EFI

Improved Integration With GlusterFS

- You can find a blog post to set up an iSCSI target on the Gluster blog here

Improved ARM Support

- A number of new features have been implemented:
- 64 bit Xen on ARM now supports booting guests
- Physical disk partitions and LVM volumes can now be used to store guest images using xen-blkback (or is PV drivers better in terms of terminology)
- Significant stability improvements across the board
- ARM/multiboot booting protocol design and implementation
- PSCI support

Improved ARM Support (2)

- Same DMA in Dom0 even with no hardware IOMMUs (not sure what the implications of this are)
- ARM and ARM64 ABIs are declared stable and maintained for backwards compatibility
- Significant usability improvements, such as automatic creation of guest device trees and improved handling of host DTBs

Improved ARM Support (3)

- Adding new hardware platforms to Xen Project on ARM has been vastly improved, making it easier for Hardware vendors and embedded vendors to port to their board
- Added support for the Arndale board, Calxeda ECX-2000 (aka Midway), Applied Micro X-Gene Storm, TI OMAP5 and Allwinner A20/A30 boards
- ARM server class hardware (Calxeda Midway) has been introduced in the Xen Project OSSTest automated testing framework

Early Microcode Loading

- The hypervisor can update the microcode in the early phase of boot time
  - The microcode binary blob can be either as a standalone multiboot payload, or part of the initial kernel (dom0) initial ramdisk (initrd)
  - To take advantage of this use latest version of *dracut* with *--early-microcode* parameter and on the Xen Project command line specify: *ucode=scan*.
  - For details see *dracut* manpage and http://xenbits.xen.org/docs/unstable/misc/xen-command-line.html

Xen Project Futures

Still More Fun to Come…

- Xen Automotive
  - Xen Project in the entertainment center of your car?
- XenGT
  - Virtualized GPU support
- Even More ARM Support
  - On your server, in your phone, wherever…
- PVH stability and performance
  - The new hypervisor mode to get harder and faster

Questions?

Russell.Pavlicek@XenProject.org

Twitter: @RCPavlicek