Xen and enclosures

Stephen Childs
Dept. of Computer Science
Trinity College Dublin

October 2007

Overview

- Status of ncm-xen
- Grid-Ireland Xen configuration
- Enclosures summary
- Discussion

ncm-xen status

- For a description of Quattor Xen tools look at this PPT from last workshop
- Updates since March:
  - Some CERN changes (mostly reverted)
  - 0.1.1: namespacing
  - 0.1.2: use getTree, change vifs to list
  - 0.1.3: support for autostart VMs, partition-backed VMs
  - 0.1.4: boot VMs if not running
  - 0.1.5: bugfix for autostart

ncm-xen plans and possibilities

- Delegate filesystem management (e.g. LVM creation) to ncm-filesystems
- Support for other use cases
  - Downloading images
  - Use of copy-on-write
  - ?
- Other VMMs (ncm-virt?)

Grid-Ireland usage

- Used for automatic reinstallation of all VM-based Grid gateways
  - PXE/KS install of host + guests
  - Fully automatic with small Xen patch
- Basic template support developed
  - `install_server` machine type for host
  - Pulls in info from guest profiles
  - Configures grub, filesystems based on variables

Host configuration variables

XEN_GUESTS	List of VMs (corresponding to profiles)
XEN_VG	Default volume group
XEN_BOOTLOADER	pypxeboot, pygrub, etc.
XEN_RAM	Default memory for VM guest if not defined
XEN_BOOTARGS	e.g. vif string needed for pypxeboot
XEN_BOOT_DEVICE	For grub (e.g. /dev/sda2)

Guest configuration

- /software/components/xen/domains/[i] populated from guest profiles
- “Hardware” templates for guests can define RAM size, disk size, etc.
- Needs to be made more flexible and extracted into generic function

Xen templates

rpms/xen/host	Host RPMs (Xen, kernel, glibc, etc.)
rpms/xen/guest_sl3	SL3 guest RPMs
xen/grub	Sets up grub

Enclosures

- Intended to address more complex system configurations
  - Multiple motherboards per case
  - Blades
  - Virtual machines
- Use cases
  - Managing interventions: all nodes within an enclosure must be powered down
  - Generation of Kickstart files: for all nodes within an enclosure

Who's the daddy?

- Two options:
  - parent-references-children
    - Already in use for Xen VMs: "/sw/comps/xen/domains/0/name"="gridgate";
  - children-reference-parent
    - "/hardware/enclosure/name" = "enc0001";
- Choice is important as it makes the other option difficult
- Does a child need to find out its parent?
- Do all use cases fit the same model? (Hardware enclosures vs. software VMs)

Xen VM perspective

- Parent **has** to know about children
  - Generate Xen config files
  - Set up local backing store for VM FS
- Child **shouldn't** know about parent
  - Child profile is not very VM-aware
  - Downtime for child doesn't affect parent
  - Child doesn't need to know when it's migrated

To discuss

- I favour p-r-c approach
  - Works for AII
  - Works for VMs
- What is the use case for c-r-p?
- Should Xen approach be generalised?
- Is “/hardware/enclosure” right?