ATOL: Kernel Services

Marek Grác
xgrac@fi.muni.cz

Red Hat Czech s.r.o. / Faculty of Informatics, Masaryk University

Advanced Topics of Linux Administration

Kernel Images and Variants

- Architectures supported: x86, x86-64, IA64, PowerPC64, s390x
- Three kernel versions available for x86:
  - Regular: one or more processsors but 4GB of RAM or less
  - PAE (big-iron): multiple processors and up to 64GB of RAM
  - Xen: needed for virtualization
- Kernel always installed under `/boot/vmlinuz-*`

Kernel Modules

- Modules are small kernel extensions that may be loaded and unloaded at will
- Can implement drivers, filesystems, firewall, and more
- Are located under `/lib/modules/(uname -r)/`
- Compiled for a specific kernel version and area provided with the kernel package
- Third party modules may be added
- Tainted modules – 3rd party non-GPL modules

Kernel Modules Utilities

- *lsmod* provides a list of loaded modules
- *modprobe* can load and unload modules
- `/etc/modprobe.conf` used for module configuration:
  - Parameters to pass to a module whenever it is loaded
  - Aliases to represent a module name
  - Commands to execute when a module is loaded or unloaded

/proc Filesystem

- `/proc` used to get or set kernel configuration
- Virtual filesystem: files not stored on hard disk
- Entries not persistent: modifications get reinitialized after a reboot
- Used to display process information, memory resources, hardware devices, kernel memory, etc.
- Can be used to modify network and memory subsystems or modify kernel features
- Modifications apply immediately

/proc Examples

- Read-only files
  - `/proc/cpuinfo`
  - `/proc/1/*`
  - `/proc/partitions`
  - `/proc/meminfo`
- Read-write files
  - `/proc/sys/kernel/hostname`
  - `/proc/sys/net/ipv4/ip_forward`
  - `/proc/sys/vm/drop_caches`
  - `/proc/sys/vm/swappiness`

sysctl : Persistent Kernel Configuration

- *sysctl* adds persistence to `/proc/sys` settings
- Statements added to `/etc/sysctl.conf` automatically reflected under `/proc` after a reboot
- Configuration maintained or monitored using the *sysctl* commands:
  - List all current settings: *sysctl -a*
  - Reload settings from `sysctl.conf`: *sysctl -p*
  - Set a `/proc` value dynamically: *sysctl -w net.ipv4.ip_forward=1*

Accessing Drivers Through /dev

- Files under `/dev` used to access drivers
- Reading from and writing to those files are valid operations:
  - Read from serial port: *cat /dev/ttyS0*
  - Write to serial port: *echo Message > /dev/ttyS0*
- Three files attributes determine which driver to access:
  - Device type (character or block)
  - Major number
  - Minor number

Device Node Examples

- Block Devices
  - `/dev/hda`, `/dev/hdc` – IDE hard disk, CDROM
  - `/dev/sda`, `/dev/sdc` – SCSI, SATA, or USB Storage
  - `/dev/md0`, `/dev/md1` – Software RAID
- Character Devices
  - `/dev/tty[0-6]` – virtual Consoles
  - `/dev/null`, `/dev/zero` – software Devices
  - `/dev/random`, `/dev/urandom` – random Numbers

Managing /dev with udev

- udev manages files stored under `/dev/`
- Files only created if corresponding device is plugged in
- Files are automatically removed when the device is disconnected
- udev statements under `/etc/udev/rules.d` determine:
  - Filenames
  - Permissions, Owners and groups
  - Commands to execute when a new device shows up

Adding Files Under /dev

- The right way to add a `/dev` entry involes udev:
  - Create a new file under `/etc/udev/rules.d`
  - Insert a statement such as:
    `KERNEL==sda, NAME=usbkey, SYMLINK=usbstorage`
  - This creates a device file named `usbkey` and a symlink named `usbstorage` next time `/dev/sda` gets plugged.
- Files can be added manually with *mknod*:
  - *mknod /dev/usbdevice b 8 0*
  - *mknod* not persistent **!**

Exploring Hardware Devices

- A snapshot of all connected devices is maintained by HAL
- HAL is acronym for Hardware Abstraction Layer
- *hal-device* lists all devices in text mode.
- *hal-device-manager* displays all devices on a graphical window.
- *lspci* and *lsusb* list devices connected to the PCI and USB buses, respectively.
- The `/proc` and `/sys` filesystems also contains bus and device specific information.

Lab: Installation

- Goals:
  - A system that does not respond to *ping*
  - A system that provides `/dev/myusbdisk` automatically after a reboot.

Lab: Prepare a paper

- Themes:
  - Describe features of HAL
  - Kernel modules: tuning and performance
- Format:
  - Short presentation (15–20 minutes; 5-7 slides)
  - Paper containing comparision (500 words)