A Solaris users guide to Linux

Phil Kirk

Introduction

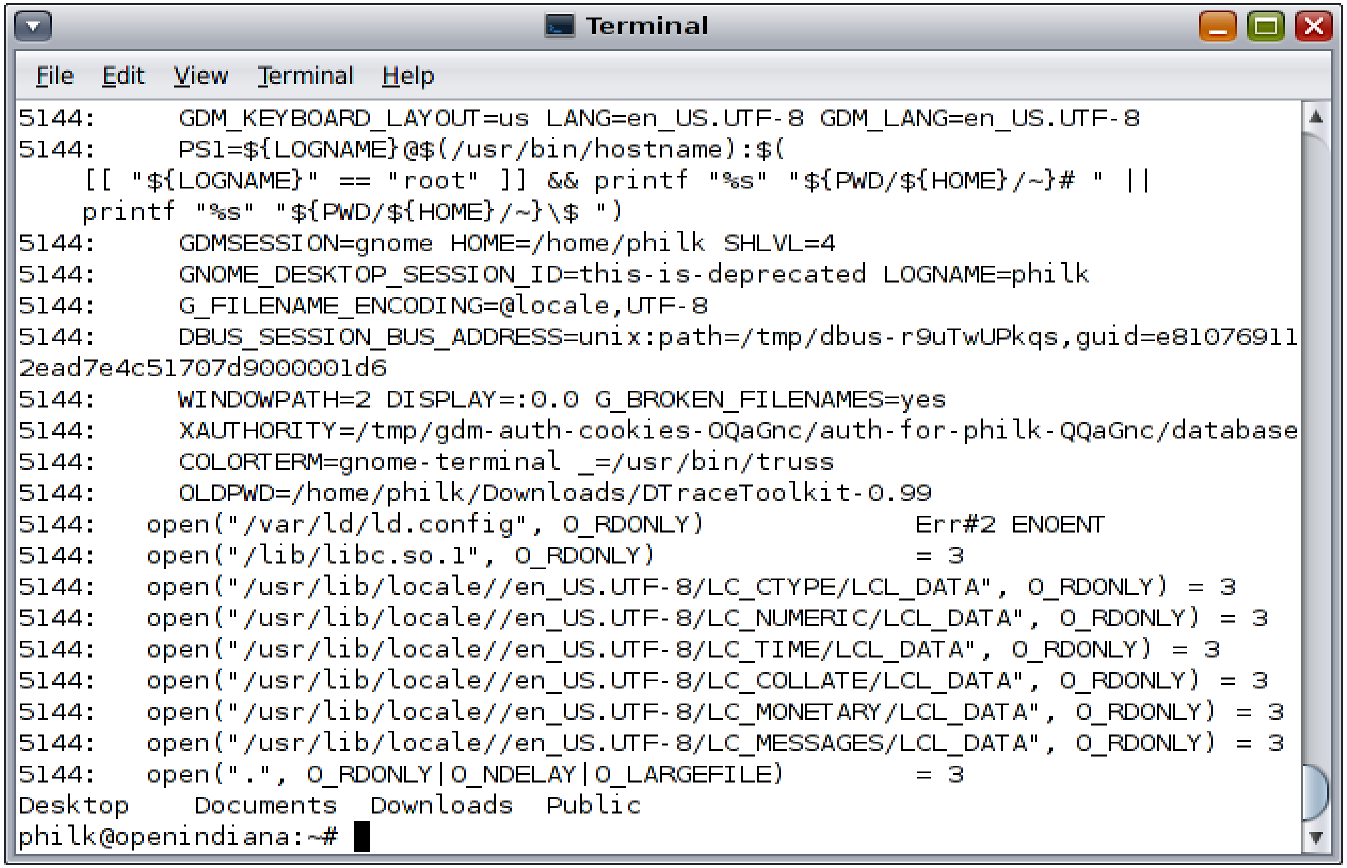
- What this presentation is
  - A look at what equivalent features exist in Linux and how they’re used
  - Features that I use or find interesting
- What this presentation isn’t
  - Which operating system is better than the other
  - How to setup a linux machine
- Environments aren’t just Solaris or Linux, there’s usually a different way to achieve the same thing

Userland tools

- Truss
- Still used a lot even with newer tools like DTrace available
- Common options, -f, -e, -a
- sotruss or truss –u trace library calls

Userland tools cont.

- Example

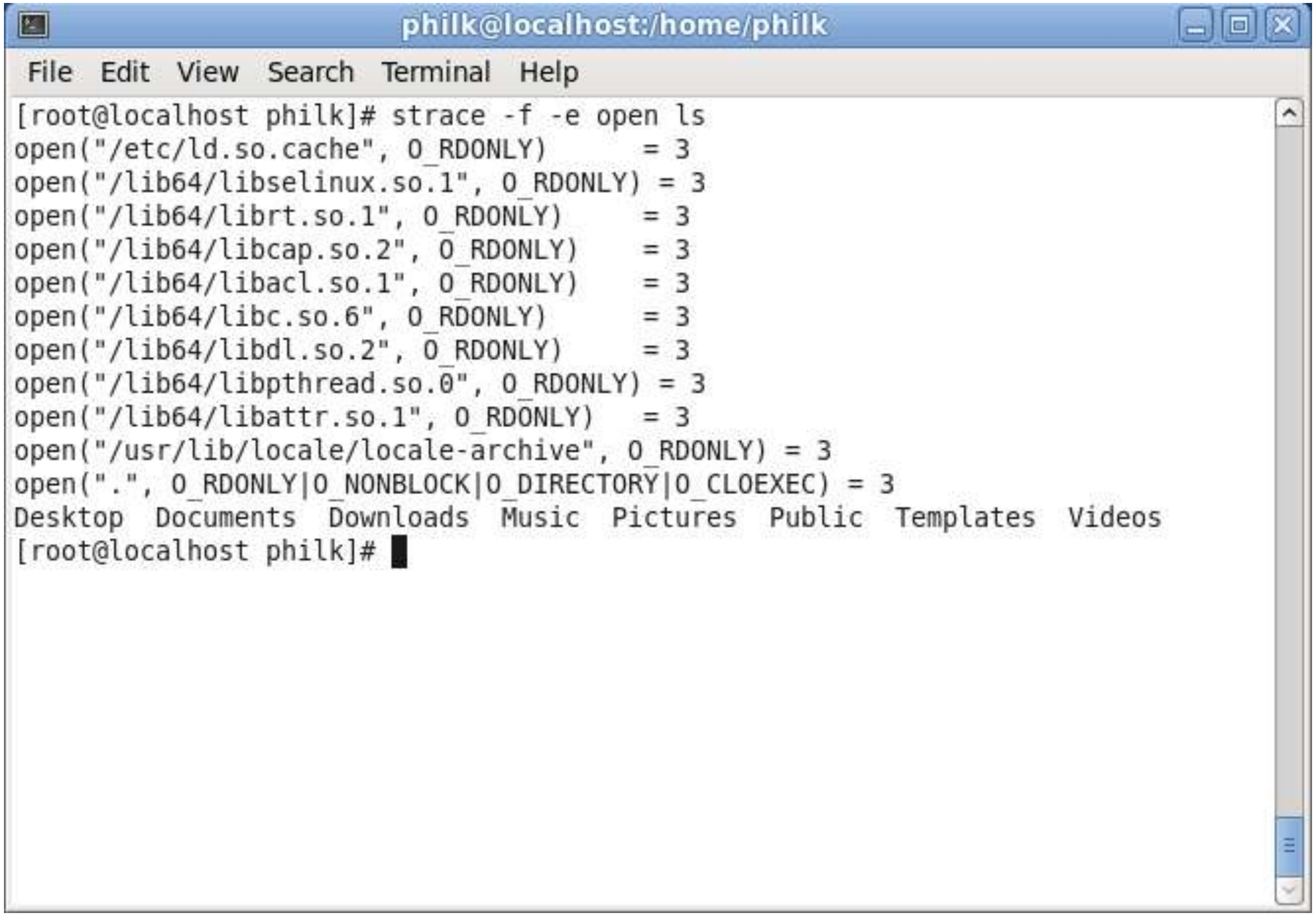


```
5144:      GDM_KEYBOARD_LAYOUT=us LANG=en_US.UTF-8 GDM_LANG=en_US.UTF-8
5144:      PS1=${LOGNAME}@$(/usr/bin/hostname):$(
    [[ "${LOGNAME}" == "root" ]] && printf "%s" "${PWD/${HOME}/~}# " ||
    printf "%s" "${PWD/${HOME}/~}\$ ")
5144:      GDMSESSION=gnome HOME=/home/philk SHLVL=4
5144:      GNOME_DESKTOP_SESSION_ID=this-is-deprecated LOGNAME=philk
5144:      G_FILENAME_ENCODING=@locale,UTF-8
5144:      DBUS_SESSION_BUS_ADDRESS=unix:path=/tmp/dbus-r9uTwUPkqs,guid=e81076911
2ead7e4c51707d9000001d6
5144:      WINDOWPATH=2 DISPLAY=:0.0 G_BROKEN_FILENAMES=yes
5144:      XAUTHORITY=/tmp/gdm-auth-cookies-OQaGnc/auth-for-philk-QQaGnc/database
5144:      COLORTERM=gnome-terminal _=/usr/bin/truss
5144:      OLDPWD=/home/philk/Downloads/DTraceToolkit-0.99
5144:    open("/var/ld/ld.config", O_RDONLY)             Err#2 ENOENT
5144:    open("/lib/libc.so.1", O_RDONLY)                = 3
5144:    open("/usr/lib/locale//en_US.UTF-8/LC_CTYPE/LCL_DATA", O_RDONLY) = 3
5144:    open("/usr/lib/locale//en_US.UTF-8/LC_NUMERIC/LCL_DATA", O_RDONLY) = 3
5144:    open("/usr/lib/locale//en_US.UTF-8/LC_TIME/LCL_DATA", O_RDONLY) = 3
5144:    open("/usr/lib/locale//en_US.UTF-8/LC_COLLATE/LCL_DATA", O_RDONLY) = 3
5144:    open("/usr/lib/locale//en_US.UTF-8/LC_MONETARY/LCL_DATA", O_RDONLY) = 3
5144:    open("/usr/lib/locale//en_US.UTF-8/LC_MESSAGES/LCL_DATA", O_RDONLY) = 3
5144:    open(".", O_RDONLY|O_NDELAY|O_LARGEFILE)        = 3
Desktop    Documents  Downloads  Public
philk@openindiana:~#
```

Userland tools cont.

- strace
- Much like truss, very similar options
- ltrace trace dynamic libraries as well as system calls
- But there are differences
- No stop on signal, depending on ltrace version problems with threads
- Example

Userland tools cont.



philk@localhost:/home/philk
File Edit View Search Terminal Help

```
[root@localhost philk]# strace -f -e open ls
open("/etc/ld.so.cache", O_RDONLY)      = 3
open("/lib64/libselinux.so.1", O_RDONLY) = 3
open("/lib64/librt.so.1", O_RDONLY)     = 3
open("/lib64/libcap.so.2", O_RDONLY)    = 3
open("/lib64/libacl.so.1", O_RDONLY)    = 3
open("/lib64/libc.so.6", O_RDONLY)      = 3
open("/lib64/libdl.so.2", O_RDONLY)     = 3
open("/lib64/libpthread.so.0", O_RDONLY) = 3
open("/lib64/libattr.so.1", O_RDONLY)   = 3
open("/usr/lib/locale/locale-archive", O_RDONLY) = 3
open(".", O_RDONLY|O_NONBLOCK|O_DIRECTORY|O_CLOEXEC) = 3
Desktop  Documents  Downloads  Music  Pictures  Public  Templates  Videos
[root@localhost philk]# 
```

Linkers and compilers

- Lots of similar linker options
- LD_DEBUG, LD_PRELOAD, LD_LIBRARY_PATH
- man ld.so, ld.so.1
- Examples where the linker can be exploited to help find problems
- LD_DEBUG, which libraries are getting loaded in what order
- LD_PRELOAD, interpose on library calls

Linkers and compilers cont.

- gcc and cc have the same set of options they just have different names
- Same for gdb and dbx
- Of course gcc and gdb run on Solaris as well
- Useful gdb things, gdb scripts and gdb tty mode , though with ddd or other graphical debuggers perhaps you don’t need to

PTools

- pstack, pgrep, pldd, etc.
- Exist on both Solaris and Linux but some things aren’t quite the same and Linux doesn’t have all the same tools
- As an example pstack under linux doesn’t work on core files
- Of course it’s it’s easy enough to use gdb but...

Solaris DTrace

- Everyone’s familar with Dtrace now but just incase
- Gives observability into both the kernel and userland
- Scriptable via D, an awk like language minus looping and conditional structures to prevent unsafe operations
- Production safe

Solaris Dtrace cont.

- Things like the Dtrace tool kit make it very easy to answer questions about what’s happening on the system
- Examples

System tap

- Not always part of the installed distribution
- Gives observability into both the kernel and userland though for userland you may need debuginfo packages
- Stap scripts are more C like and allow looping constructs, safety is provided inside the SystemTap framework
- Production system safe?

SystemTap cont.

- Lots of examples much like the Dtrace tool kit to help answer common questions quickly
- As a comparison I noticed there wasn’t an opensnoop.stp in the examples so wondered how hard it would be to write
- Examples

Perf

- Originally developed as a userland program to provide access to the prerformance counters
- Extended to work with kernel tracepoints
- Can be used to look at userland too
- http://www.linux-kongress.org/2010/slides/lk2010-perf-acme.pdf
- Examples

LTT

- Linux Trace Toolkit
- Aims to produce an effecient full system tracing acility
- Allow tracing userland, kernel and also provides tools for viewing and analysing trace output
- http://lttng.org/

DTrace

- Paul Fox’s port
- OLE

Kernel debugging

- On Solaris mdb, kmdb
- Crash dump facilities are managed through `dumpadm(1M)`
- All installed as part of core system

Linux tools

- kgdb
- Kdb
- crash
- Crash dump facilities managed via kexec/kdump
- On SMP machines extra config needed
- Tools not part of many standard distributions and additional packages need to be installed

Kernel debugging cont.

- Different approaches to debugging tools
- Linux tools tend to aim and kernel developers, source code debugging via kgdb
- Solaris tools developed to root cause system problems the first time they occur

Kernel debugging cont.

- What sort of problems do you want to solve?
- As an example consider a system falling over due to a read/write lock not been released
- On Solaris it’s fairly straightforward to start trying to find the culprit
- First find the lock we’re interested in
  - mdb>::walk thread | ::findstack
- Then find who owns the lock

Kernel debugging cont

  - mdb><addr>::kgrep | ::whatis
- On Linux it wouldn’t be quite so straightforward, crash doesn’t have the same feature set as Solaris.

Virtualization

- Solaris zones
- Open Solaris, KVM
- LDOMS

Solaris zones

- Resource containers
- Lightweight virtualization
- Configured using zoneadm(1M) zonecfg(1M)
- Rich feature set
- Example

Solaris Zones cont.

- Easy to create a new zone

  -
    ```
    #zonecfg -z zone
     testzone: No such zone configured
     Use 'create' to begin configuring a new zone.
     zonecfg:testzone> create
    ```
- Easy to add network interface using vnics
  - `#dladm create-vnic -l e1000g0 vnic0`
- Easy to administer via zoneadm

Linux OpenVZ

- Not part of the core distribution
- Mature product used as the base for Parallels Virtuozzo Container
- Looks fairly straightforward to setup containers
- http://wiki.openvz.org/Main_Page

Linux Vserver

- Not part of the standard distribution
- Need to install separate packages and kernel updates
- http://linux-vserver.org/Welcome_to_Linux-VServer.org

Linux containers lxc

- http://lxc.sourceforge.net/
- “LXC is the userspace control package for Linux Containers, a lightweight virtual system mechanism sometimes described as “chroot on steroids”. “
- Mainstream, integrated from 2.6.29
- Seems very similar to zones in terms of configuration
- Not yet as feature rich

Linux containers lxc cont

- Uses libvirt to provide virtualized networking support
- Example of creating a lxc container
  - Various templates available to create linux guests
  - Creating a new fedora guest is as easy as running lxc-fedora though you may want to do more
  - Running the new guest lxc-start
  - Similar commands to zones

Linux containers lxc cont

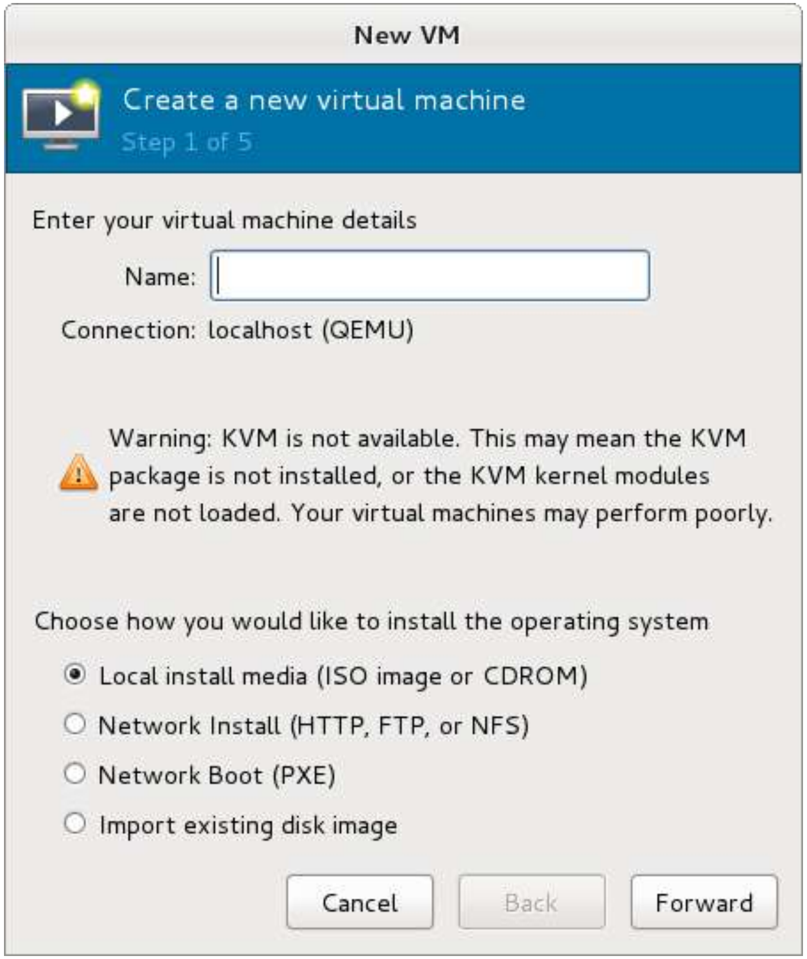
  - lxc-info, lxc-console, lxc-stop

Linux KVM

- Type 2 hypervisor requires Intel VT or AMD V cpu support
- In the linux kernel from 2.6.25
- Possibly require additional packages depending on distribution
- Using virt-manager makes setup and management of virtual machines easy

Linux KVM cont.

- Example



New VM

Create a new virtual machine

Step 1 of 5

Enter your virtual machine details

Name:

Connection: localhost (QEMU)

Warning: KVM is not available. This may mean the KVM package is not installed, or the KVM kernel modules are not loaded. Your virtual machines may perform poorly.

Choose how you would like to install the operating system

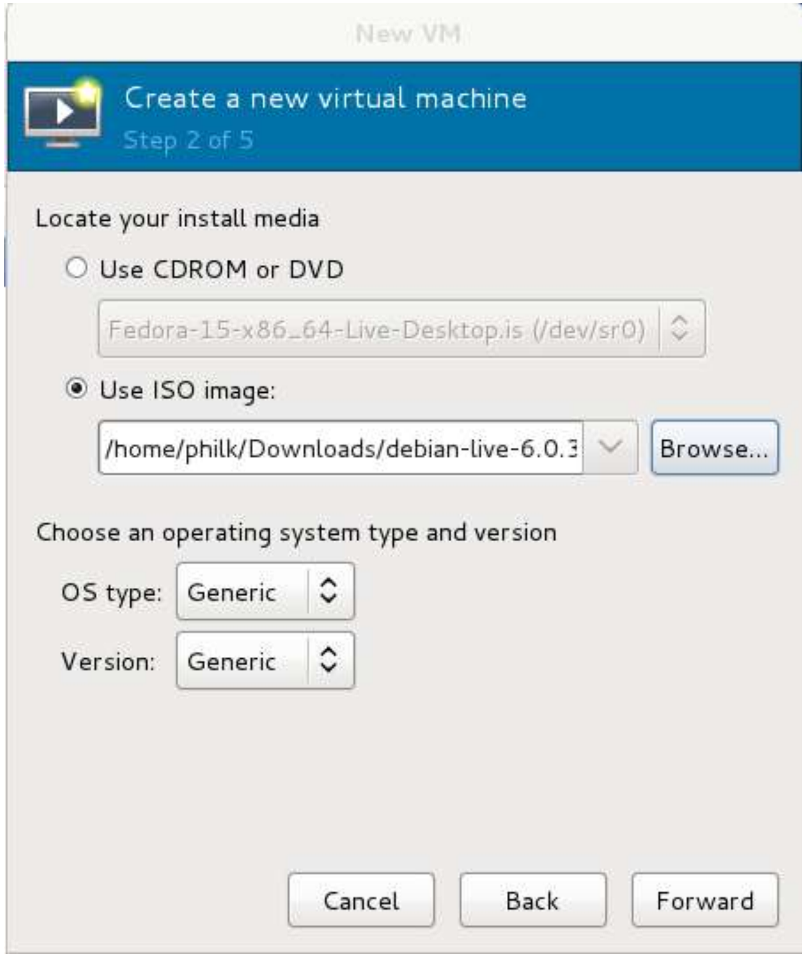
● Local install media (ISO image or CDROM)

○ Network Install (HTTP, FTP, or NFS)

○ Network Boot (PXE)

○ Import existing disk image

Linux KVM cont



Create a new virtual machine

Step 2 of 5

Locate your install media

○ Use CDROM or DVD

Fedora-15-x86_64-Live-Desktop.is (/dev/sr0)

● Use ISO image:

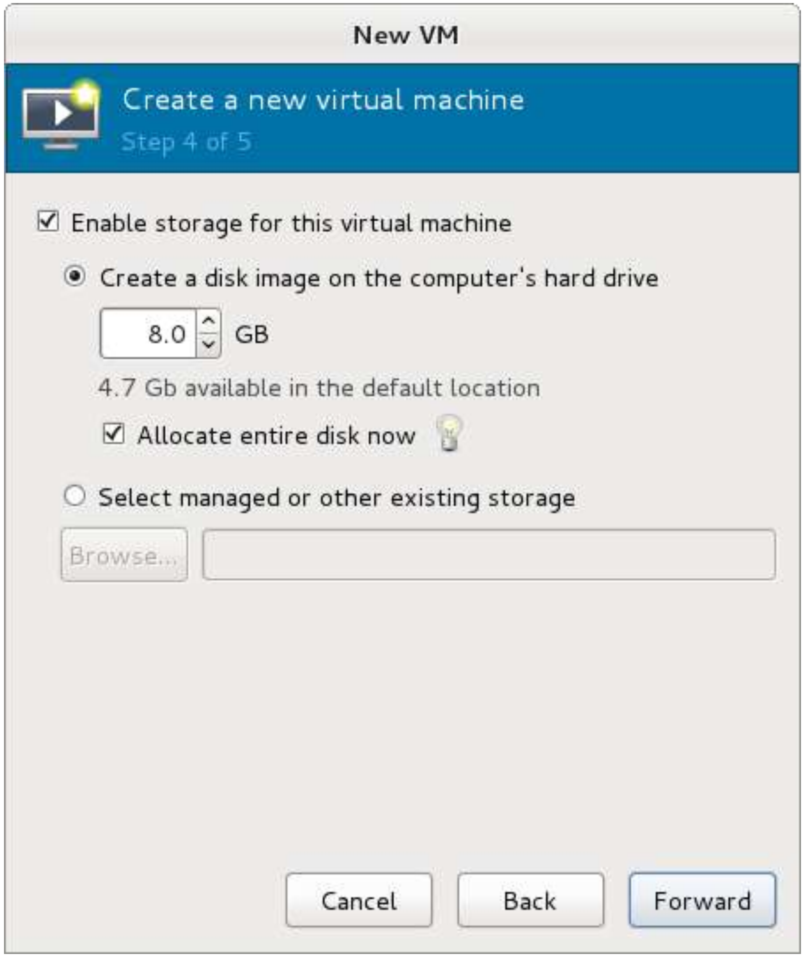
/home/philk/Downloads/debian-live-6.0.3

Choose an operating system type and version

OS type: Generic

Version: Generic

Linux KVM cont



Create a new virtual machine

Step 4 of 5

☑ Enable storage for this virtual machine

● Create a disk image on the computer's hard drive

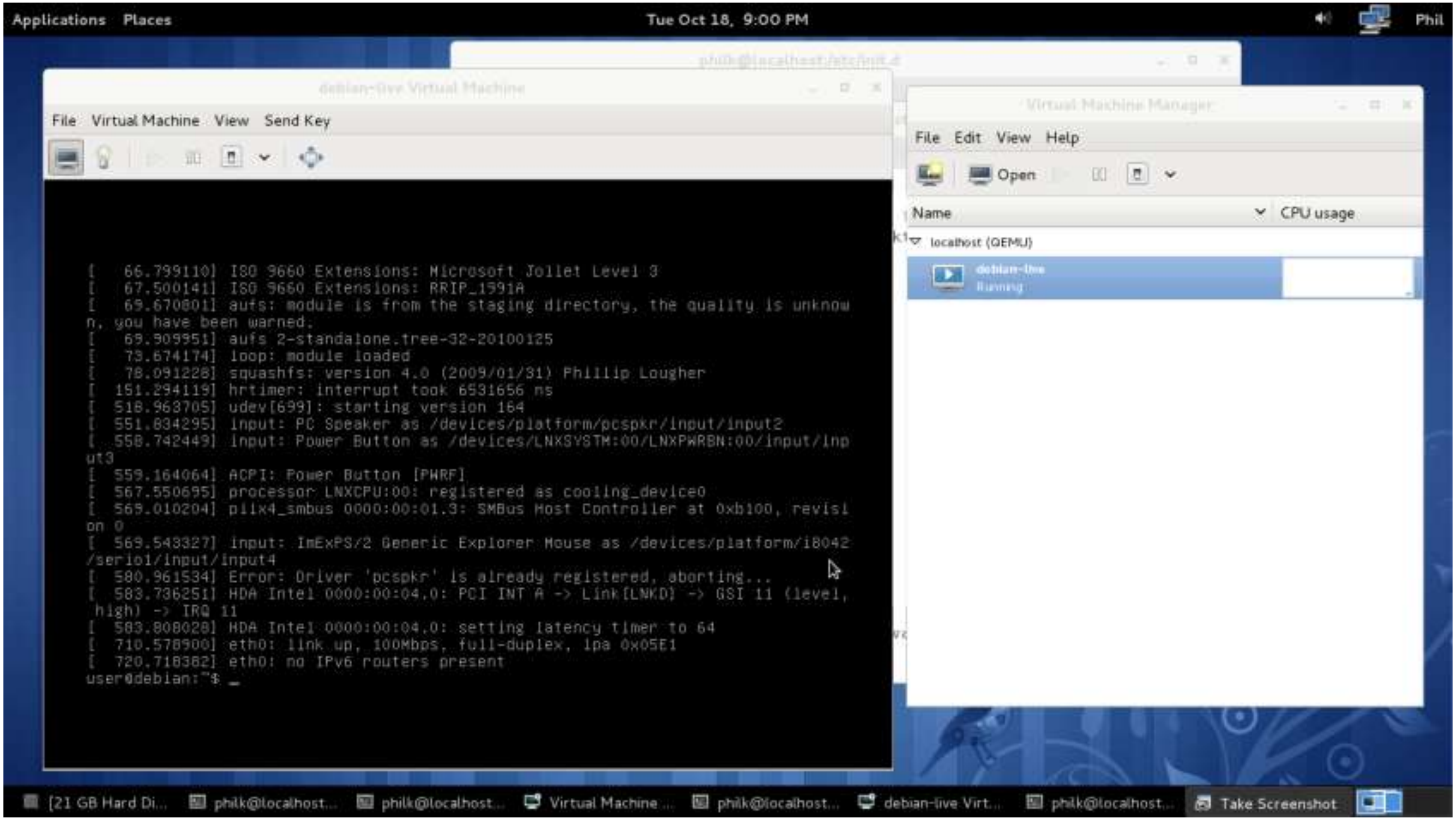
8.0 GB

4.7 Gb available in the default location

☑ Allocate entire disk now

○ Select managed or other existing storage

Linux KVM cont



```
[   66.799110] ISO 9660 Extensions: Microsoft Joliet Level 3
[   67.500141] ISO 9660 Extensions: RRIP_1991A
[   69.670801] aufs: module is from the staging directory, the quality is unknow
n, you have been warned.
[   69.909951] aufs 2-standalone.tree-32-20100125
[   73.674174] loop: module loaded
[   78.091228] squashfs: version 4.0 (2009/01/31) Phillip Lougher
[  151.294119] hrtimer: interrupt took 6531656 ns
[  518.963705] udev[699]: starting version 164
[  551.834295] input: PC Speaker as /devices/platform/pcspkr/input/input2
[  558.742449] input: Power Button as /devices/LNXSYSTM:00/LNXPWRBN:00/input/inp
ut3
[  559.164064] ACPI: Power Button [PWRF]
[  567.550695] processor LNXCPU:00: registered as cooling_device0
[  569.010204] piix4_smbus 0000:00:01.3: SMBus Host Controller at 0xb100, revisi
on 0
[  569.543327] input: ImExPS/2 Generic Explorer Mouse as /devices/platform/i8042
/serio1/input/input4
[  580.961534] Error: Driver 'pcspkr' is already registered, aborting...
[  583.736251] HDA Intel 0000:00:04.0: PCI INT A -> Link[LNKD] -> GSI 11 (level,
 high) -> IRQ 11
[  583.808028] HDA Intel 0000:00:04.0: setting latency timer to 64
[  710.578900] eth0: link up, 100Mbps, full-duplex, lpa 0x05E1
[  720.718382] eth0: no IPv6 routers present
user@debian:~$ _
```

Thankyou

Any questions