

Before You Begin

Introduction to Unix for HPC

Agenda – Day 1 – Shell Primer

9:45 Welcome

10:00 – 11:15 Session 1 (Intro & Unit 1)

11:15 – 11:30 Short break

11:30 – 12:30 Session 2 (Unit 2+3)

12:30 – 1:30 Lunch

13:30– 14:45 Session 3 (Unit 4)

14:45 - 15:00 Short Break

15:00 – 16:15 Session 4 (Unit 5)



Agenda – Day 2 – Data & HPC

9:45 - 11:15 Session 1 (Data - Units 1+2+3)

11:15 – 11:30 Short break

11:30 – 12:30 Session 2 (Data – Units 4+5)

12:30 – 1:30 Lunch

13:30 – 15:30 Session 3 (HPC)



General Introduction

- Intersect http://www.intersect.org.au/
 - Who we are?
 - Your Trainer
- Your University IT Contacts
- General Housekeeping
 - Toilets
 - Coffee & Water Facilities
 - Emergency Exits



Course Information

- Who is the course intended for?
- What will be covered during the 2 days?
 - More details in the course outline at <u>http://www.intersect.org.au/course-resources</u>
- What level will you be at upon completion of the course?
- Remember It's YOUR course, so ask questions!



Download training materials

Training Material available from:

http://www.intersect.org.au/course-resources

- Download the Slides and Exercises and save to a local directory, e.g. C:\HPC:
 - Course Outlines: Shell Primer, Data, HPC
 - Slides: Shell Primer, Data, HPC Slides
 - Exercises



Install software

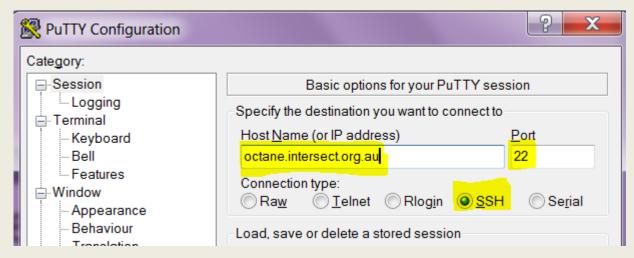
Use your web browser to download Putty and PuttySCP from: http://www.putty.org

- Click on the "Download Putty" link and download:
 - <u>putty.exe</u> (a Telnet and SSH client)
 - <u>pscp.exe</u> (an SCP client, i.e. command-line secure file copy)
- Double click to install on your PC.



Logging On

- Start the putty telnet/ssh client by double clicking on putty.exe and connect to the HPC Machine
 - Host: octane.intersect.org.au
 - Connection Type: ssh
 - Port: 22





Log On to the HPC Machine

- Log into the training HPC system using the Test Account allocated to you, e.g.
 - Account Name: hpc26
 - Password: train26

```
login as: hpc26
Using keyboard-interactive authentication.
Password:
Last login: Fri Feb 15 10:08:28 2013 from 129.94.48.59

- Welcome to Octane@ORANGE
- HPC learning and testing platform
-
Setting up initial SSH keys...
hpc26@octaneorange:~>
```





Shell Primer

Introduction to Unix for HPC

Unit 1: Run Spot, Run

Goals:

- Can run commands, including when there are spaces in arguments
- Can use 'man' to find out more about commands, including arguments and parameters



UNIX Operating System

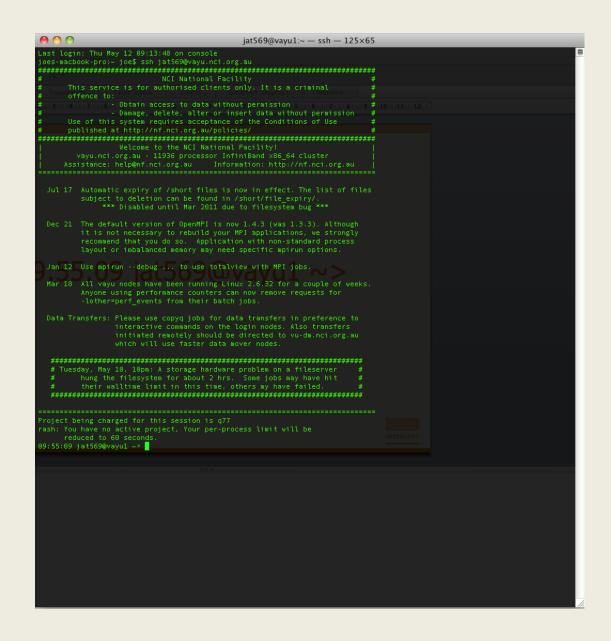
- UNIX was initially developed by AT&T employees at Bell Labs in 1969
- Common Variants today: Linux,
 OpenSolaris, FreeBSD, HP/UX, AIX
- 100's of Major Linux Distributions: Debianbased (e.g. Ubuntu), Red Hat Package Manager (RPM)-based (SUSE, Red Hat & Fedora)
 - Octane uses SUSE LINUX



What is a Shell?

- The shell is a command line interpreter or shell that provides an <u>interface to the</u> <u>UNIX operating system</u>.
- A shell has a single purpose to allow users to enter commands to execute, or create scripts containing commands, to direct the operation of the computer







More on UNIX Shells

- This course uses BASH
 - BASH = the Bourne-Again Shell
 - Default shell for most Linux systems
- Other common Shells include:
 - C Shell (csh)
 - K Shell (ksh)
 - TENEX C Shell (tcsh)



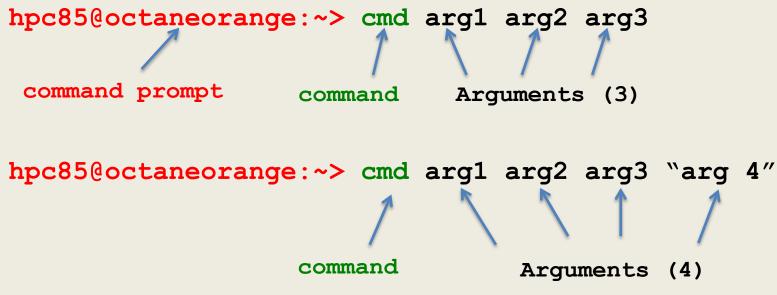
What does BASH do?

- The shell (command line interpreter)
 <u>interprets</u> the commands entered by the
 user and passes those to the UNIX
 operating system.
- When enter a command and hit return
 - BASH <u>parses</u> the command line into <u>tokens</u>
 - 1st token is <u>interpreted</u> as the **command**
 - Remaining tokens are <u>interpreted</u> as arguments



Anatomy of a command

- Commands comprise of:
 - a command that invokes a program
 - arguments to that program





Commands and Shells

- Commands allow users to interact with the operating system via the shell
- Two-way communication is possible between the shell and commands
 - e.g. the ls command will return a list of files in a directory





Exercise 1(a)

Running Commands

Command	Description	
ls	Shows a directory listing	
w	Shows who is logged on to the system and what they are doing	
w hpc22	Shows login, idle time, and what user hpc22 is doing	
finger	Shows login, username & login details of users on the system	
finger hpc22	Shows user hpc22's login details including name, idle time, login time as well as some of their environment settings	
date	Prints the system time and date	
uptime	Tells you how long the system has been running	



Command Line Options (Flags)

- Command Line Options / Flags modify the operation of the command.
 - There are two forms of flags Short Form & Long Form.
 - Flags are case sensitive!
- Short form options start with a single hyphen "-"
 - rm -f <filename> -> force removal of file
- Long form Options start with a double hyphen "--"
 - rm --force <filename> -> force removal of file

Both commands do the exact same thing!

Sample Command & Flags

- -ls-a
 - short form flag "-a"
 - Shows all files incl. files starting with .
- ls --all
 - long form flag "--all"
 - Shows all files incl. files starting with .



More Commands & Flags

```
-ls -1
```

- short form flag "-1"
- Shows the long format listing for all files in the directory



Combining Flags

- Flags can be combined in any order but still mean the same thing
- The following command and flags all show the directory contents and show one file per line:
 - ls -a -1
 - ls -1 -a
 - 1s -la
 - ls -al
 - ls --all -1
 - ls -1 -all



Command Line Arguments

- <u>Command Line Parameters</u> are arguments sent to the program being called.
 - There are two forms of parameters Short Form & Long Form.
 - Parameters are case sensitive!
- Short form options start with a single hyphen "-"
 - tail -n 20 <filename> → tail from the last
 20 lines of the file
- Long form Options start with a double hyphen "--"
 - tail --lines=20 <filename> → tail from the last 20 lines of the file



Exercise 1(b)

Flags and Parameters

Command	Description
ls	Shows a directory listing
ls -l	(long) Lists directory contents of current directory using long listing format
ls -a	(all) Lists directory contents of current
lsall	directory and does not ignore entries starting with .
ls -la ls -all -a	(all+long) Lists directory contents of current directory using long listing format and does not ignore entries starting with .
ls -format=horizontal	Lists directory contents of current directory horizontally
lsformat=single-column	Lists directory contents of current directory in a single column

Using the "man" pages

- All Unix systems come equipped with manual or "man" pages which contain documentation about how each in-built command works
- Man pages can be accessed using the following command format:

```
man program, e.g. man ls
```



The online manual

 man is a command that takes as its argument the name of another command.

NAME

```
find - search for files in a directory hierarchy
```

SYNOPSIS

```
find [-H] [-L] [-P] [path...] [expression]
```

 The [] indicate an optional argument (you don't type these)



Exercise 1(c)

The Calendar

Command	Description
man <command name=""/>	Print the online manual entry for <command name=""/>
cal	Print a calendar



Unit 2: Where am I?

Goals:

- Can explore navigate a hierarchy of directories
- Can specify files with relative and absolute paths
- Can create a hierarchy of directories
- Can copy and move files between directories



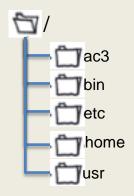
UNIX Directories

- Called 'folders' under windows
 - Exactly the same concept

 You run into them a LOT more under Unix than you do under windows

- A directory is a container
 - It can contain files and other directories

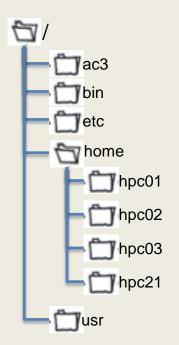




```
What are the names of the contents of the directory called "/" (root folder)?
```

```
/ac3
/bin
/etc
/home
/usr
```

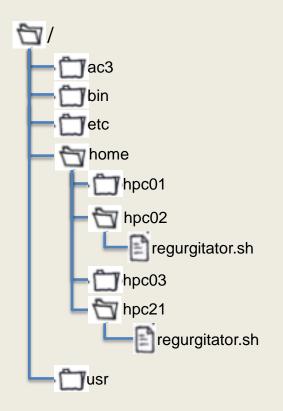




What are the names of the contents of the directory called /home?

```
/home/hpc01
/home/hpc02
/home/hpc03
/home/hpc21
```





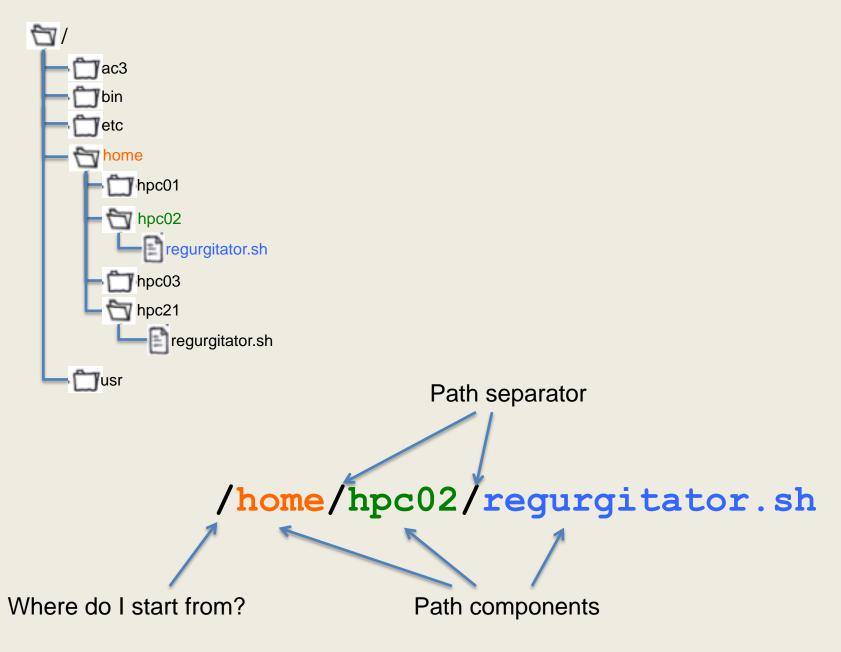
What are the names of the contents of the directory called /home/hpc02?

/home/hpc02/regurgitator.sh

What are the names of the contents of the directory called /home/hpc21?

/home/hpc21/regurgitator.sh





Rules So Far

- There is one root, called "/"
 - This is different from windows, where there is one root for each disk drive C:, D:, etc

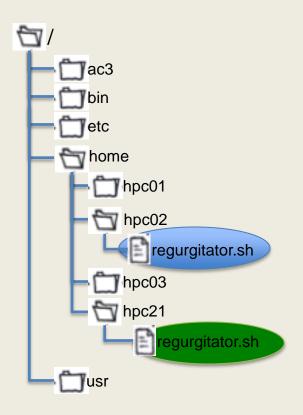
 A path from the root designates exactly one file: such a path is called an "absolute path"



The Home Directory

- Each account (user) has a special directory called his/her home directory
 - That's where you 'get put' when you log in. (More on this later)
- BASH uses the "~" character to indicate "home directory"
- Two forms
 - ~ "my home directory"
 - ~hpc01 "hpc01's home directory"





What are the names of the contents of the directory called ~

It depends who asks: If the user "hpc02" asks, then the answer is /home/hpc02/regurgitator.sh

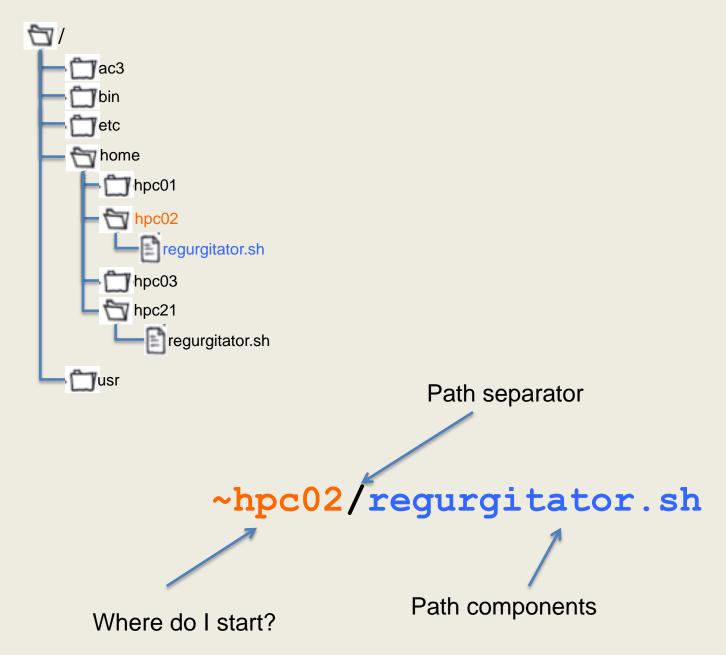
What are the names of the contents of the directory called ~hpc21

/home/hpc21/regurgitator.sh

What is the absolute path of the file ~hpc02/regurgitator.sh

/home/hpc02/regurgitator.sh







Rules So Far

- There is one root, called "/"
- A path starting with "/" means "from the root"
- A path starting with "~" means "from the home"



Exercise 2(a)

Finding your way around

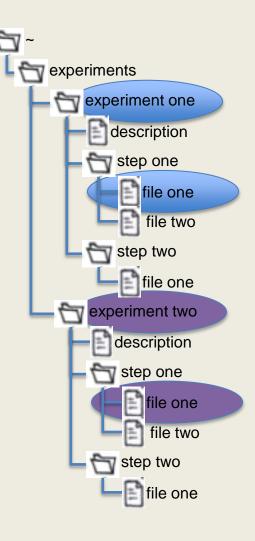
Command	Description
cd	Change the working directory to your home directory
cd <path></path>	Change directory to <path></path>
ls	List the contents of the working directory
ls <path></path>	List the contents of <path></path>



The Working Directory

- BASH maintains a "working" or "current" directory. That
 is "the directory that you are currently in"
 - Your prompt will show you your current working directory
- Linux uses the "." character to denote the working directory. You can use this when running commands
 - ls -la \cdot \rightarrow show the listing for the current dir
- The command "pwd" will list the current working directory





What is the absolute path of "step one/file one"

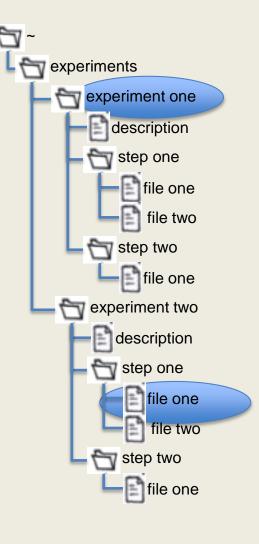
What is the working directory?



Looking upwards

- UNIX uses ".." to denote the parent of a directory. You can use this when running commands, e.g.
 - $cd .. \rightarrow change up 1 directory$
 - -cd ../hpc01 → change up 1 directory, then down 1 directory to hpc01
- UNIX understands ".." as a "normal directory" and it can appear in a path

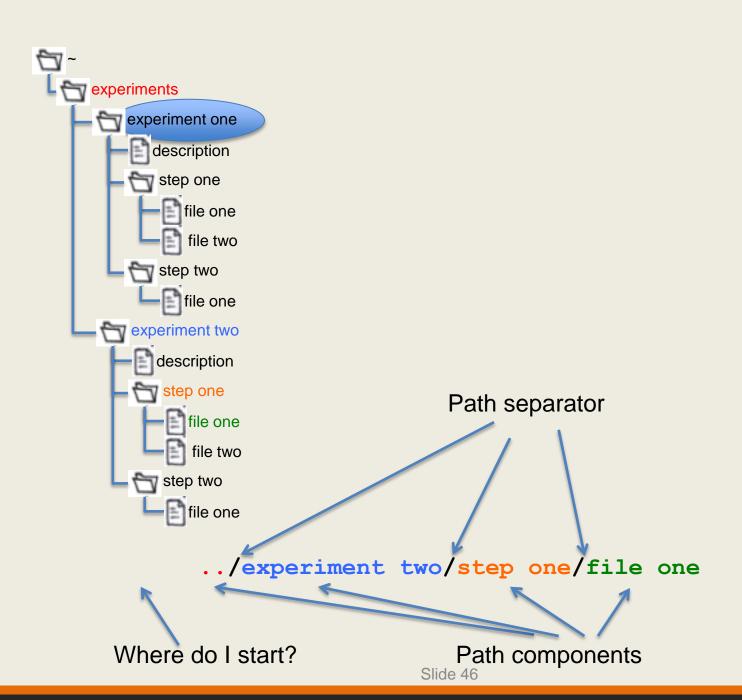




What is the absolute path of .../experiment two/step one/file one

What is the working directory?

INTERSECT





Everything about directories

- There is one root, called "/"
- A path starting with "/" means "from the root"
- A path starting with "~" means "from the home"
- A path starting with anything else means "from the working directory"
- Note the last and second last rules assume a hidden context (current user and current directory)



Exercise 2(b)

Finding your way around

Command	Description
pwd	Print the name of the current working directory
cd	Change directory to the parent directory
ls	List the contents of the working directory
ls .	List the contents of the working directory
ls	List the contents of the parent directory

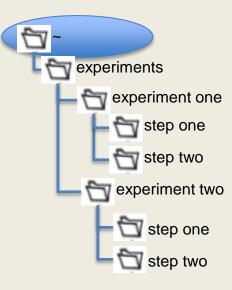


Making New Directories

 You can create a directory, but only one at a time (by default)

```
mkdir </path/new_directory>
mkdir </path/to/new/directory>
```





What happens when you execute the commands?

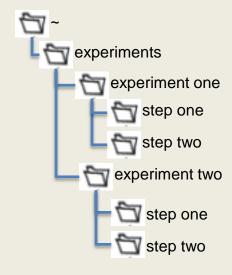
```
mkdir ~/experiments
cd ~/experiments
mkdir "experiment one"
mkdir "experiment one/step one"
mkdir "experiment one/step two"
mkdir "experiment two"
cd "experiment two"
mkdir "step one"
mkdir "step two"
```



Exercise 2(c)

Making Directories

Command	Description
mkdir <path></path>	Make a directory at <path></path>
pwd	Print the name of the current working directory
cd <path></path>	Change directory to <path></path>
ls <path></path>	List the contents of <path></path>





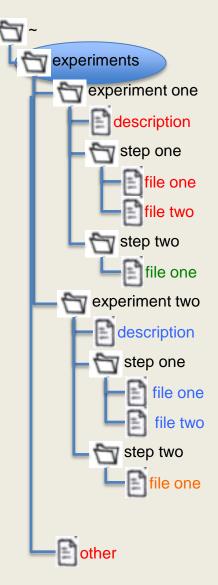
Moving/Copying Files Around

Files live in exactly one location

Files can be copied and moved between directories

```
cp <from path/file> <to path/file>
mv <from path/file> <to path/file>
```





What happens when you execute the commands?

```
cp "experiment one/description" "experiment two"
cd "experiment one"
cp "step one/file one" "../experiment two/step two"
cp "../experiment two/description" ../other
cd "../experiment two/step one"
mv "file one" "../step two"
```



Exercise 2(d)

Moving and Copying Files

Command	Description					
cp <from> <to></to></from>	Copy a file from <from> to <to></to></from>					
mv <from> <to></to></from>	Copy a file from <from> to <to> then remove <from></from></to></from>					
man cp man mv	Print the online manual entry for <command name=""/> . What happens when you provide more than two arguments to cp and mv?					



Unit 3: It's what's inside that counts

Goals:

- Can look inside files (even really big ones)
- Can obtain rudimentary information about files



What's in a file?

my dame has a lame tame crane my dame has a crane that is lame pray, gentle jane, that my dame's lame tame crane feed and come home again

0000000: 0000006: 000000c: 0000012: 0000018: 000001e: 0000024: 000002a: 0000030: 0000036: 000003c:

my dam
e has
a lame
tame
crane.
my dam
e has
a cran
e that
is la
me.pra



What's in a file?

This power-point presentation.

```
0000000:
           PK . . .
0000006:
 000000c:
 0000012:
 0000018:
 000001e:
 [Conte
0000024:
 nt Typ
           es].xm
000002a:
 0000030:
 0000036:
 000003c:
```

What is a 'text file'?

109 104 10

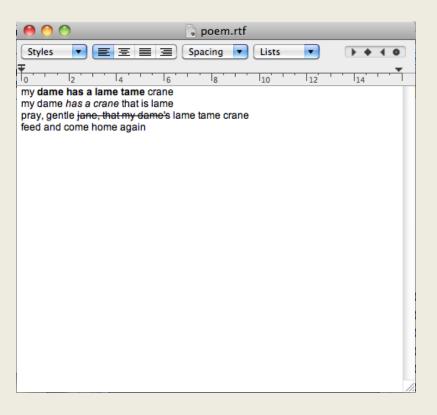
```
0000000:
 0000006:
000000c:
 0000012:
 0000018:
 000001e:
0000024:
 000002a:
 0000030:
0000036:
 000003c:
```

my dam
e has
a lame
tame
crane.
my dam
e has
a cran
e that
is la
me.pra



Dec	H	Oct	Char		Dec	Нх	Oct	Html	Chr	Dec	Нх	Oct	Html	Chr	Dec	: Нх	Oct	Html Ch	<u>ır</u>
0	0	000	NUL	(null)	32	20	040	@#32;	Space	64	40	100	a#64;	0	96	60	140	`	`
1				(start of heading)				6#33 ;	_	1	41	101	A	A				a#97;	a
2	2	002	STX	(start of text)	34	22	042	@#3 4 ;	rr	66	42	102	B	В	98	62	142	4#98;	b
3				(end of text)				a#35;					C					c	
4				(end of transmission)				\$	-				D					d	
5				(enquiry)				a#37;					E					e	
6				(acknowledge)	I			a#38;		I . –			F					f	
7				(bell)				'					G			7 1		g	
8		010		(backspace)				a#40;					H					a#104;	
9		011		(horizontal tab)	ı			a#41;					I					i	
10		012		(NL line feed, new line)				6#42;					a#74;					j	
11		013		(vertical tab)				6#43;					K					k	
12		014		(NP form feed, new page)				6#44;					L					l	
13		015		(carriage return)				6# 4 5;					M					m	
14		016		(shift out)				a#46;					N					n	
15		017		(shift in)				6#47;					O					o	
		020		(data link escape)				6#48;					P					p	
		021		(device control 1)				6# 4 9;					Q	_				q	
		022		(device control 2)				2					R					r	
								3					S					s	
				(device control 4)				4					T					t	
				(negative acknowledge)	I			6#53;					U					u	
				(synchronous idle)				6#5 4 ;					V					v	
		027		(end of trans. block)	I			6#55;		I			W					w	
				(cancel)				8					X					x	
		031		(end of medium)	I			6#57;		I			Y					y	
		032		(substitute)				6#58;					Z					z	
								6#59;					[_				{	
		034		(file separator)				4#60;					& # 92;	-				4 ;	
		035		(group separator)				=					& # 93;	_				}	
		036		(record separator)				a#62;					^					~	
31	1F	037	US	(unit separator)	63	ЗF	077	4#63;	2	95	5F	137	_	_	127	7F	177		DEL
													S	ourc	e: w	ww.	.Look	upTables	.com

Two sides of the same coin



{\rtf1\ansi\ansicpg1252\cocoartf1038\cocoa subrtf350 {\fonttbl\f0\fswiss\fcharset0 ArialMT;} {\colortbl;\red255\green255\blue255;} \paperw11900\paperh16840\margl1440\ma rgr1440\vieww9000\viewh8400\viewkind0 \deftab720 \pard\pardeftab720\ql\qnatural

\fo\fs24 \cf0 my
\b dame has a lame tame
\b0 crane\
my dame
\i has a crane
\i0 that is lame\
pray, gentle \strike \strikec0 jane, that my
dame's\strike0\striked0 lame tame crane\
feed and come home again\

Slide 60

Why do we care?

- BASH is more or less committed to ASCII text
- In particular
 - your keystrokes result in ASCII codes being sent to the shell
 - when a program generates output, it is interpreted by the shell as ASCII codes and displayed on the screen (even some nonprintables)



Exercise 3(a)

Looking into files

Command	Description				
ls	You've met before				
<pre>cat <filename></filename></pre>	Print a file to the terminal (catenate)				
less <filename></filename>	Like cat, but less at a time				
head <filename></filename>	Like cat, but just the top of the file				
tail <filename></filename>	Like cat, but just the end of the file				
wc <filename></filename>	Word count – count the words in a file				
du	D isk u sage – how much space is used				



Unit 4: It's all just files

Goals:

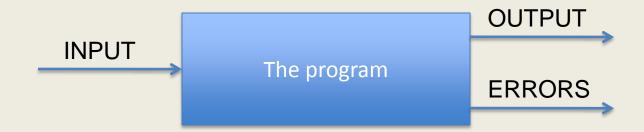
- Can save the output of programs into files
- Can chain commands together



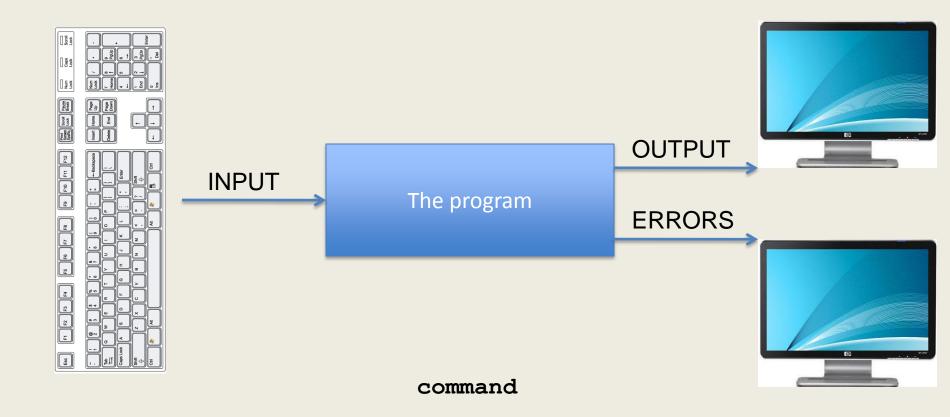
The core magic of the Linux command line

- The next bit is the core magic of Linux, as far as this module goes.
 - The keyboard is just a file
 - The console is just a file
- Define What is input?
- Define What is output?



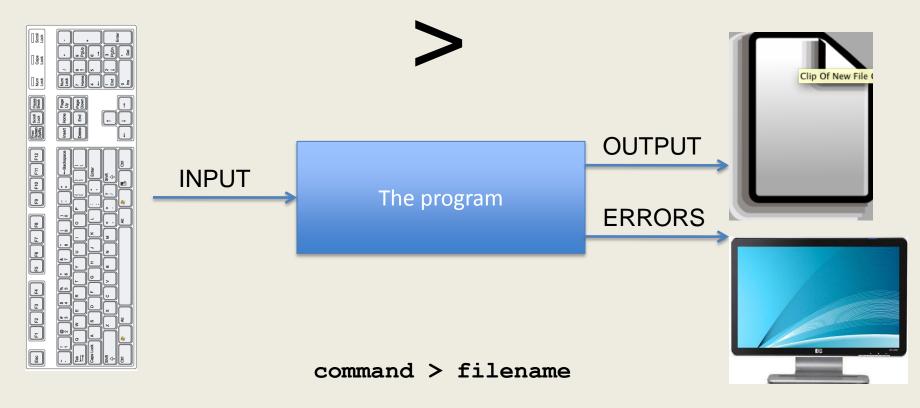






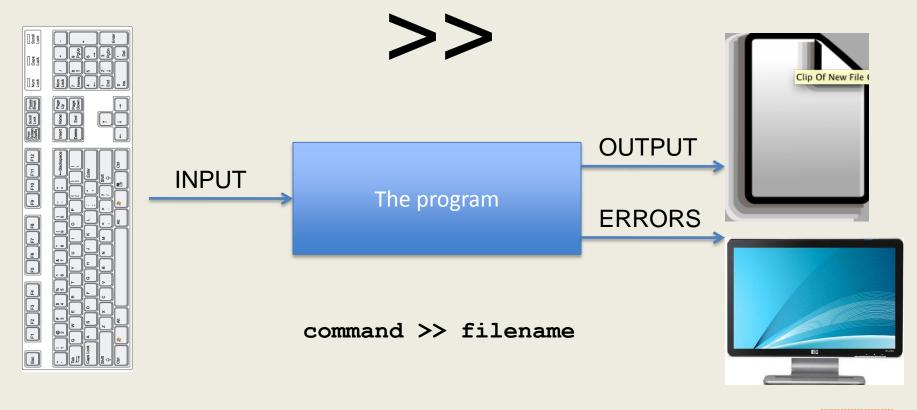


Redirect OUTPUT elsewhere...



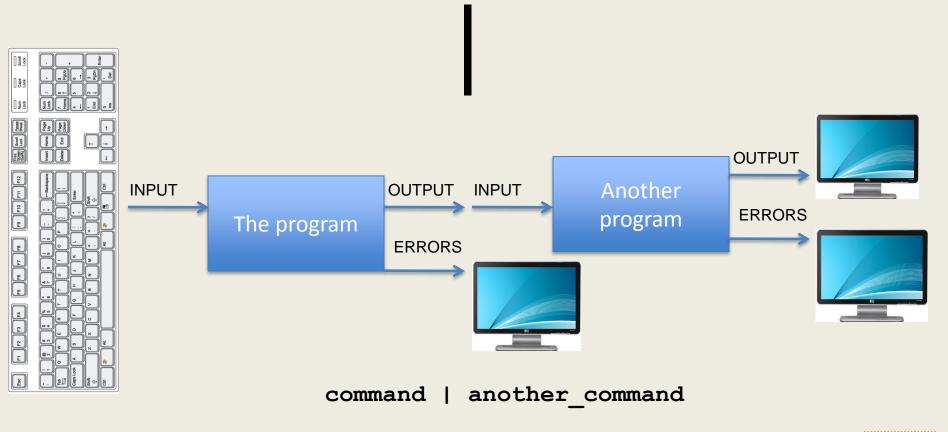


Append OUTPUT elsewhere...





Pipe OUTPUT to a command...





Exercise 4(a)

Looking into files

Command	Description
fortune	Tells your fortune – really! Try it!
<pre>sort <file></file></pre>	Sorts the contents of $\langle file \rangle$, or input if no file is specified.
uniq <file></file>	Removes any redundant lines in <file>, or input if no file is specified. Redundant lines have to be adjacent to be considered redundant.</file>
<pre>cat <file></file></pre>	Prints the contents of <file></file>
<command/> > <file></file>	Run <command/> but rather than print the output to the console, redirect the output to <file>.</file>
<command/> >> <file></file>	Run <command/> but rather than print the output to the console, append the output to <file>.</file>
<command/> <command/>	Pipe one <command/> to another <command/>



Unit 5: Products of our environment

Goals:

- Can set and read environment variables
- Can add programs to the PATH
- Can query the environment for which variables exist



Case study – finding programs

- Recall from previous units
 - Sometimes you have to specify the path of a program

```
regurgitator.sh
./regurgitator.sh
../regurgitator.sh
/bin/regurgitator.sh
```



PATH

- There is a 'special' thing called the PATH
 - Part of what makes it special is encoded
 - Part of what makes it special is convention
- There is an environment full of variables
 - One of those variables is called PATH
 - Use `\$' to tell BASH you're talking about a variable
- To check what an environment variable is set to, use \$< ENVIRONMENT_VARIABLE>
 - e.g. \$PATH



A sample path

```
/bin:/usr/bin:/usr/local/bin:.
```



What does BASH do?

- When you hit return
 - BASH parses the command line into tokens
 - Any 'special' tokens are expanded
 - The first token is treated as the command
 - The remaining tokens are passed to the command as arguments



A simple example

 Your home directory is stored in the environment, you can be taken there directly

cd \$HOME



Exercise 5(a)

Realistic Evaluation of the Environment

Command	Description
\$ <variable></variable>	Evaluates to the contents of the variable <variable>.</variable>
env	Prints all environment variables and their values.
echo <arg1> <arg2></arg2></arg1>	Prints out its arguments, with no additional information.
<pre>grep <pattern> <file></file></pattern></pre>	Prints out all occurrences of <pattern> in <file>. If no file is specified, processes standard input. For example</file></pattern>
	env grep PA
	will print all variables containing 'PA' in them or their value.



Manipulating the Environment

- You can use the environment to your advantage.
 - Update the path
 - Change your prompt
 - Use the environment to share information between your scripts



The export command

The **export** command is used to set the value of an environment variable:



Exercise 5(b)

Manipulating the Environment

Command	Description
export <var>=<val></val></var>	Set a variable called VAR and give it contents VAL . The variable will be available to programs invoked from the shell.



Summing up – Your Skills

 You have now got the rudimentary skills to move around, make and view files, chain commands together and manipulate the environment



Summing up – Hard Lessons

- BASH is wildly picky
- Silent success
- Command-line work is flexible but error prone
- There's a lot of hidden context
 - ENVIRONMENT, WORKING DIRECTORY, STDOUT/STDIN, COMMAND LINE EXPANSION
- BASH is wildly picky



Summing up – the positives

- Command line gives you a more <u>powerful</u> means to control a program
- Allows <u>direct interaction</u> with shell commands (no GUI needed)
- Allow for <u>ease of automation</u> via scripting
- UNIX is more <u>flexible</u> and can be installed on many different types of machines (e.g. mainframes/supercomputers)
- Unix possesses much greater <u>processing</u> <u>power</u> than Windows.

