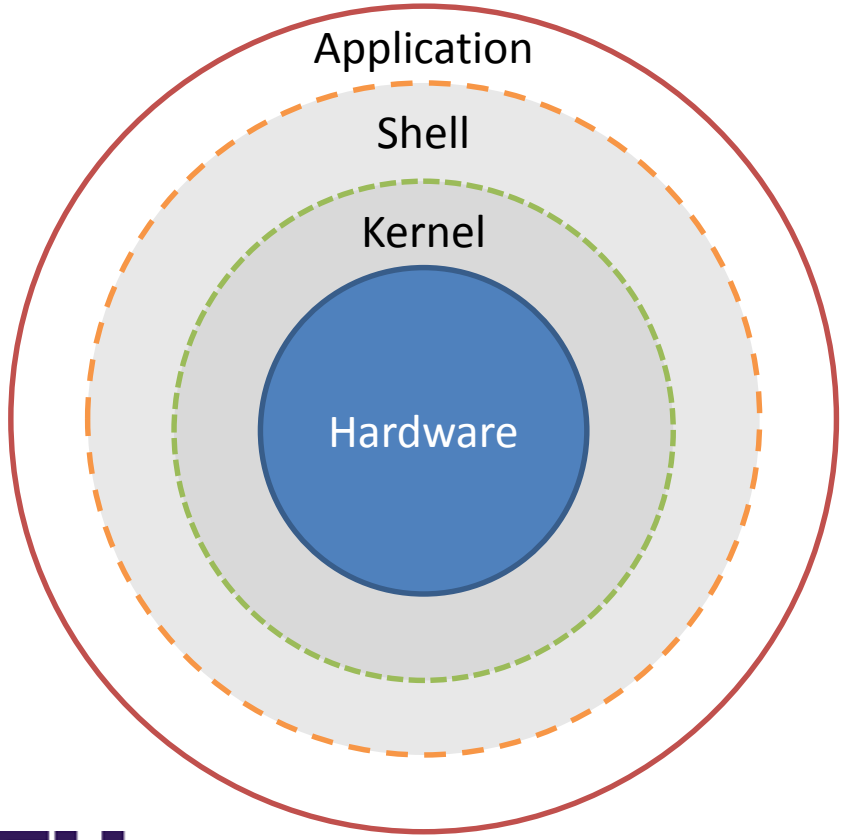
Shell Scripting

Shaohao Chen and Le Yan
HPC User Services @ LSU

Outline

- Introduction to Linux Shell
- Shell Scripting Basics
- Beyond Basic Shell Scripting
  - Arithmetic Operations
  - Arrays
  - Flow Control
  - Command Line Arguments
  - Functions
- Advanced Text Processing Commands (grep, sed, awk)

What Do Operating Systems Do?



- Operating systems work as a bridge between hardware and applications
  - Kernel: hardware drivers etc.
  - Shell: user interface to kernel
  - Some applications (system utilities)

: Operating System

Kernel

- Kernel
  - The kernel is the core component of most operating systems
  - Kernel's responsibilities include managing the system's resources
  - It provides the lowest level abstraction layer for the resources (especially processors and I/O devices) that application software must control to perform its functions
  - It typically makes these facilities available to application processes through inter-process communication mechanisms and system calls

Shell

- Shell
  - The command line interface is the primary user interface to Linux/Unix operating systems.
  - Each shell has varying capabilities and features and the users should choose the shell that best suits their needs
  - The shell can be deemed as an application running on top of the kernel and provides a powerful interface to the system.

Type of Shell

- **sh (Bourne Shell)**
  - Developed by Stephen Bourne at AT&T Bell Labs
- **csh (C Shell)**
  - Developed by Bill Joy at University of California, Berkeley
- **ksh (Korn Shell)**
  - Developed by David Korn at AT&T Bell Labs
  - Backward-compatible with the Bourne shell and includes many features of the C shell
- **bash (Bourne Again Shell)**
  - Developed by Brian Fox for the GNU Project as a free software replacement for the Bourne shell
  - Default Shell on Linux and Mac OSX
  - The name is also descriptive of what it did, bashing together the features of sh, csh and ksh
- **tcsh (TENEX C Shell)**
  - Developed by Ken Greer at Carnegie Mellon University
  - It is essentially the C shell with programmable command line completion, command-line editing, and a few other features.

Shell Comparison

Software	sh	csh	ksh	bash	tcsh
Programming language	y	y	y	y	y
Shell variables	y	y	y	y	y
Command alias	n	y	y	y	y
Command history	n	y	y	y	y
Filename autocompletion	n	y*	y*	y	y
Command line editing	n	n	y*	y	y
Job control	n	y	y	y	y

*: not by default

http://www.cis.rit.edu/class/simg211/unixintro/Shell.html

Linux Shell Variables

- Linux allows the use of variables
  - Similar to programming languages
- A variable is a named object that contains data
  - Number, character or string
- There are two types of variables: ENVIRONMENT and user defined
- Environment variables provide a simple way to share configuration settings between multiple applications and processes in Linux
  - Environment variables are often named using all uppercase letters
  - Example: `PATH, LD_LIBRARY_PATH, SHELL, DISPLAY` etc.
  - `printenv:` list all environment variables
- To reference a variable, prepend $ to the name of the variable, e.g. `$PATH, $LD_LIBRARY_PATH`
  - Example: `$PATH, $LD_LIBRARY_PATH, $DISPLAY` etc.

Variable Names

- Rules for variable names
  - Must start with a letter or underscore
  - Number can be used anywhere else
  - Do not use special characters such as @,#,%,$
  - (again) They are case sensitive
  - Example
    - Allowed: `VARIABLE, VAR1234able, var_name, _VAR`
    - Not allowed: `1var, %name, $myvar, var@NAME`

Editing Variables (1)

- How to assign values to variables depends on the shell

Type	sh/ksh/bash	csh/tcsh
Shell	`name=value`	`set name=value`
Environment	`export name=value`	`setenv name=value`

- Shell variables is only valid within the current shell, while environment variables are valid for all subsequently opened shells.

Editing Variables (2)

- Example: to add a directory to the PATH variable

```
sh/ksh/bash: export PATH=/path/to/executable:${PATH}
csh/tcsh: setenv PATH /path/to/executable:${PATH}
```

  - `sh/ksh/bash`: no spaces except between export and PATH
  - `csh/tcsh`: no "=" sign
  - Use colon to separate different paths
  - The order matters: more forward, higher priority.

Basic Linux Commands

Name	Function
`ls`	Lists files and directories
`cd`	Changes the working directory
`mkdir`	Creates new directories
`rm`	Deletes files and directories
`cp`	Copies files and directories
`mv`	Moves or renames files and directories
`pwd`	prints the current working directory
`echo`	prints arguments to standard output
`cat`	Prints file content to standard output

- Use option --help to check usage of commands

File Editing in Linux

- The two most commonly used editors on Linux/Unix systems are:
  - `vi` or `vim` (vi improved)
  - `emacs`
- `vi/vim` is installed by default on Linux/Unix systems and has only a command line interface (CLI).
- `emacs` has both a CLI and a graphical user interface (GUI).
  - if emacs GUI is installed then use `emacs -nw` to open file in console
- Other editors you may come across: `kate, gedit, gvim, pico, nano, kwrite`
- To use `vi` or `emacs` is your choice, but you need to know one of them
- **For this tutorial, we assume that you already know how to edit a file with a command line editor.**

Outline

- Introduction to Linux Shell
- Shell Scripting Basics
- Beyond Basic Shell Scripting
  - Arithmetic Operations
  - Arrays
  - Flow Control
  - Command Line Arguments
  - Functions
- Advanced Text Processing Commands

Scripting Languages

- A script is a program written for a software environment that automate the execution of tasks which could alternatively be executed one-by-one by a human operator.
- Shell scripts are a series of shell commands put together in a file
  - When the script is executed, it is as if someone type those commands on the command line
- The majority of script programs are ``quick and dirty'', where the main goal is to get the program written quickly.
  - Compared to programming languages, scripting languages do not distinguish between data types: integers, real values, strings, etc.
  - Might not be as efficient as programs written in C and Fortran, with which source files need to be compiled to get the executable

Startup Scripts

- When you login to a *NIX computer, shell scripts are automatically loaded depending on your default shell
- `sh/ksh` (in the specified order)
  - `/etc/profile`
  - `$HOME/.profile`
- `bash` (in the specified order)
  - `/etc/profile` (for login shell)
  - `/etc/bashrc or /etc/bash/bashrc`
  - `$HOME/.bash_profile` (for login shell)
  - `$HOME/.bashrc`
- `csh/tcsh` (in the specified order)
  - `/etc/csh.cshrc`
  - `$HOME/.tcshrc`
  - `$HOME/.cshrc` (if `.tcshrc` is not present)
- `.bashrc, .tcshrc, .cshrc, .bash_profile` are script files where users can define their own aliases, environment variables, modify paths etc.

An Example

```
# .bashrc

# Source global definitions
if [ -f /etc/bashrc ]; then
        . /etc/bashrc
fi

# User specific aliases and functions
alias c="clear"
alias rm="/bin/rm -i"
alias psu="ps -u apacheco"
alias em="emacs -nw"
alias ll="ls -lF"
alias la="ls -al"
export PATH=/home/apacheco/bin:${PATH}
export g09root=/home/apacheco/Software/Gaussian09
export GAUSS_SCRDIR=/home/apacheco/Software/scratch
source $g09root/g09/bsd/g09.profile

export TEXINPUTS=.:/usr/share/texmf//:/home/apacheco/LaTeX//:${
    TEXINPUTS}
export BIBINPUTS=.:/home/apacheco/TeX//:${BIBINPUTS}
```

Writing and Executing a Script

- Three steps
  - Create and edit a text file (`hello.sh`)

```
#!/bin/bash
# My First Script
echo "Hello World!"
```

  - Set the appropriate permission

```
~/Tutorials/BASH/scripts> chmod 755 hello.sh
```

  - Execute the script

```
~/Tutorials/BASH/scripts> ./hello.sh
Hello World!
```

Components Explained

```
#!/bin/bash
# My First Script
echo "Hello World!"
```

- The first line is called the "Shebang" line. It tells the OS which interpreter to use. In the current example, bash
  - For tcsh, it would be: `#!/bin/tcsh`
- The second line is a comment. All comments begin with "#".
- The third line tells the OS to print "Hello World!" to the screen.

Special Characters (1)

#	Starts a comment line.
$	Indicates the name of a variable.
\	Escape character to display next character literally
{ }	Used to enclose name of variable
;	Command separator. Permits putting two or more commands on the same line.
;;	Terminator in a case option
.	"dot" command. Equivalent to `source` (for bash only)

Special Characters (2)

`$?`	Exit status variable.
`$$`	Process ID variable.
`[]`	Test expression.
`[[ ]]`	Test expression, more flexible than []
`$[], $(())`	Integer expansion
`||, &&, !`	Logical OR, AND and NOT

Quotation

- Single quotation
  - Enclosed string is read literally
- Double quotation
  - Enclosed string is expanded
- Back quotation
  - Enclose string is executed as a command

Quotation - Examples

```
[shaohao@mike1 bash_scripts]$ str1='echo $USER'
[shaohao@mike1 bash_scripts]$ echo $str1
echo $USER
[shaohao@mike1 bash_scripts]$ str2="echo $USER"
[shaohao@mike1 bash_scripts]$ echo $str2
echo shaohao
[shaohao@mike1 bash_scripts]$ str3=`echo $USER`
[shaohao@mike1 bash_scripts]$ echo $str3
shaohao
```

Quotation – More Examples

```
#!/bin/bash

HI=Hello

echo HI           # displays HI
echo $HI          # displays Hello
echo \$HI         # displays $HI
echo "$HI"        # displays Hello
echo '$HI'        # displays $HI
echo "$HIAlex"    # displays nothing
echo "${HI}Alex"  # displays HelloAlex
echo `pwd`        # displays working directory
echo $(pwd)       # displays working directory


~/Tutorials/BASH/scripts/day1/examples> ./quotes.sh
HI
Hello
$HI
Hello
$HI

HelloAlex
/home/apacheco/Tutorials/BASH/scripts/day1/examples
/home/apacheco/Tutorials/BASH/scripts/day1/examples
~/Tutorials/BASH/scripts/day1/examples>
```

Outline

- Introduction to Linux Shell
- Shell Scripting Basics
- Beyond Basic Shell Scripting
  - Arithmetic Operations
  - Arrays
  - Flow Control
  - Command Line Arguments
  - Functions
- Advanced Text Processing Commands

Arithmetic Operations (1)

- You can carry out numeric operations on integer variables

Operation	Operator
Addition	+
Subtraction	–
Multiplication	*
Division	/
Exponentiation	** (bash only)
Modulo	%

Arithmetic Operations (2)

- bash
  - `$((...))` or `$[...]` commands
    - Addition: `$((1+2))`
    - Multiplication: `$[$a*$b]`
  - Or use the `let` command: `let c=$a-$b`
  - Or use the `expr` command: `c=`expr $a - $b``
  - You can also use C-style increment operators:

    `let c+=1` or `let c--`

Arithmetic Operations (3)

- tcsh
  - Add two numbers: `@ x = 1 + 2`
  - Divide two numbers: `@ x = $a / $b`
  - You can also use the expr command: `set c = 'expr $a % $b'`
  - You can also use C-style increment operators:
    `@ x -= 1` or `@ x++`
- Note the use of space
  - bash: space required around operator in the `expr` command
  - tcsh: space required between @ and variable, around = and numeric operators.

Arithmetic Operations (4)

- For floating numbers
  - You would need an external calculator like the GNU basic calculator (`bc`)
    - Add two numbers

```
echo "3.8 + 4.2" | bc
```

    - Divide two numbers and print result with a precision of 5 digits:

```
echo "scale=5; 2/5" | bc
```

    - Call `bc` directly:

```
bc <<< "scale=5; 2/5"
```

    - Use `bc -l` to see result in floating point at max scale:

```
bc -l <<< "2/5"
```

Arrays (1)

- bash and tcsh supports one-dimensional arrays
- Array elements may be initialized with the `variable[i]` notation: `variable[i]=1`
- Initialize an array during declaration
  - **bash**: `name=(firstname 'last name')`
  - **tcsh**: `set name = (firstname 'last name')`
- Reference an element `i` of an array `name`: `${name[i]}`
- Print the whole array
  - **bash**: `${name[@]}`
  - **tcsh**: `${name}`
- Print length of array
  - **bash**: `${#name[@]}`
  - **tcsh**: `${#name}`

Arrays (2)

- Print length of element `i` of array `name`: `${#name[i]}`
  - Note: In **bash** `${#name}` prints the length of the first element of the array
- Add an element to an existing array
  - **bash** `name=(title ${name[@]})`
  - **tcsh** `set name = ( title "${name}")`
  - In the above **tcsh** example, title is first element of new array while the second element is the old array name
- Copy an array name to an array user
  - **bash** `user=(${name[@]})`
  - **tcsh** `set user = ( ${name} )`

Arrays (3)

- Concatenate two arrays
  - **bash** `1`
  - **tcsh** `set nameuser=( "${name}" "${user}" )`
- Delete an entire array: `unset name`
- Remove an element i from an array
  - **bash**
```
unset name[i]
```
  - **tcsh**
```
@ j = $i - 1
@ k = $i + 1
set name = ( "${name[1-$j]}" "${name[$k-]}" )
```
- Note
  - **bash**: array index starts from 0
  - **tcsh**: array index starts from 1

Arrays (4)

name.sh

```
#!/bin/bash

echo "Print your first and last name"
read firstname lastname

name=($firstname $lastname)

echo "Hello " ${name[@]}

echo "Enter your salutation"
read title

echo "Enter your suffix"
read suffix

name=($title "${name[@]}" $suffix)
echo "Hello " ${name[@]}

unset name[2]
echo "Hello " ${name[@]}
```

```
~/Tutorials/BASH/scripts/day1/examples> ./name.sh
Print your first and last name
Alex Pacheco
Hello  Alex Pacheco
Enter your salutation
Dr.
Enter your suffix
the first
Hello  Dr. Alex Pacheco the first
Hello  Dr. Alex the first
```

name.csh

```
#!/bin/tcsh

echo "Print your first name"
set firstname = $<
echo "Print your last name"
set lastname = $<

set name = ( $firstname $lastname)
echo "Hello " ${name}

echo "Enter your salutation"
set title = $<

echo "Enter your suffix"
set suffix = "$<"

set name = ($title $name $suffix )
echo "Hello " ${name}

@ i = $#name
set name = ( $name[1-2] $name[4-$i] )
echo "Hello " ${name}
```

```
~/Tutorials/BASH/scripts/day1/examples> ./name.csh
Print your first name
Alex
Print your last name
Pacheco
Hello  Alex Pacheco
Enter your salutation
Dr.
Enter your suffix
the first
Hello  Dr. Alex Pacheco the first
Hello  Dr. Alex the first
```

Flow Control

- Shell scripting languages execute commands in sequence similar to programming languages such as C and Fortran
  - Control constructs can change the order of command execution
- Control constructs in bash and tcsh are
  - Conditionals: `if`
  - Loops: `for, while, until`
  - Switches: `case, switch`

if statement

- An if/then construct tests whether the exit status of a list of commands is 0, and if so, execute one or more commands

bash

```
if [ condition1 ]; then
  some commands
elif [ condition2 ]; then
  some commands
else
  some commands
fi
```

tcsh

```
if ( condition1 ) then
  some commands
else if ( condition2 ) then
  some commands
else
  some commands
endif
```

- Note the space between condition and the brackets
  - bash is very strict about spaces.
  - tcsh commands are not so strict about spaces
  - tcsh uses the `if-then-else` `if-else-endif` similar to Fortran

File Tests

Operation	bash	tcsh
File exists	`if [ -e .bashrc ]`	`if ( -e .tcshrc )`
File is a regular file	`if [ -f .bashrc ]`	
File is a directory	`if [ -d /home ]`	`if ( -d /home )`
File is not zero size	`if [ -s .bashrc ]`	`if ( ! -z .tcshrc )`
File has read permission	`if [ -r .bashrc ]`	`if ( -r .tcshrc )`
File has write permission	`if [ -w .bashrc ]`	`if ( -w .tcshrc )`
File has execute permission	`if [ -x .bashrc ]`	`if ( -x .tcshrc )`

Integer Comparisons

Operation	bash	tcsh
Equal to	`if [ 1 -eq 2]`	`if (1 == 2)`
Not equal to	`if [ $a -ne $b ]`	`if ($a != $b)`
Greater than	`if [ $a -gt $b ]`	`if ($a > $b)`
Greater than or equal to	`if [ 1 -ge $b ]`	`if (1 >= $b)`
Less than	`if [ $a -lt 2 ]`	`if ($a < 2)`
Less than or equal to	`if [ $a -le $b ]`	`if ($a <= $b)`

String Comparisons

Operation	bash	tcsh
Equal to	`if [ $a == $b ]`	`if ($a == $b)`
Not equal to	`if [ $a != $b ]`	`if ($a != $b)`
Zero length or null	`if [ -z $a ]`	`if ($%a == 0)`
Non zero length	`if [ -n $a ]`	`if ($%a > 0)`

- One might think that these "[" and "]" belong to the syntax of Bash's if-clause: No they don't! It's a simple, ordinary command, still!

if [*expression*]	if test *expression*
`if [ ! -e .bashrc ]`	`if test ! -e .bashrc`

Logical Operators

Operation	Example
! (NOT)	`if [ ! -e .bashrc ]`

&& (AND)	`if [ -f .bashrc ] && [ -s .bashrc ]`
	`if [[ -f .bashrc && -s .bashrc ]]`
	`if ( -e .tcshrc && ! -z .tcshrc )`

| (OR)	`if [ -f .bashrc ] || [ -f .bash_profile ]` `if [[ -f .bashrc || -f .bash_profile ]]`

Examples

```
read a
if [[ "$a" -gt 0 &&  "$a" -lt 5 ]]; then
  echo "The value of $a lies somewhere between 0 and 5"
fi
OR
if [ "$a" -gt 0 ] && [ "$a" -lt 5 ]; then
  echo "The value of $a lies somewhere between 0 and 5"
fi
```

```
set a = $<
if ( "$a" > 0  &&  "$a" < 5 ) then
  echo "The value of $a lies somewhere between 0 and 5"
endif
```

Loop Constructs

- A loop is a block of code that iterates a list of commands as long as the loop control condition is evaluated to true
- Loop constructs
  - bash: `for`, `while` and `until`
  - tcsh: `foreach` and `while`

For Loop - bash

- The `for` loop is the basic looping construct in **bash**

```
for arg in list
do
  some commands
done
```

- The `for` and `do` lines can be written on the same line:

  `for` *arg* `in` *list*`;` `do`

- `for` loops can also use C style syntax

```
for i in $(seq 1 10)
do
  touch file${i}.dat
done
```

```
for i in $(seq 1 10); do
  touch file${i}.dat
done
```

```
for ((i=1;i<=10;i++))
do
  touch file${i}.dat
done
```

For Loop - tcsh

- The `foreach` loop is the basic looping construct in **tcsh**

```
foreach i (`seq 1 10`)
  touch file$i.dat
end
```

While Loop

- The `while` construct tests for a condition at the top of a loop and keeps going as long as that condition is true.
- In contrast to a `for` loop, a `while` loop finds use in situations where the number of loop repetitions is not known beforehand.
- bash

```
while [ condition ]
do
  some commands
done
```

- tcsh

```
while ( condition )
  some commands
end
```

While Loop - Example

factorial.sh

```
#!/bin/bash

read counter
factorial=1
while [ $counter -gt 0 ]
do
   factorial=$(( $factorial * $counter ))
   counter=$(( $counter - 1 ))
done
echo $factorial
```

factorial.csh

```
#!/bin/tcsh

set counter = $<
set factorial = 1
while ( $counter > 0 )
   @ factorial = $factorial * $counter
   @ counter -= 1
end
echo $factorial
```

Until Loop

- The `until` construct tests for a condition at the top of a loop, and keeps looping as long as that condition is false (opposite of `while` loop)

```
until [ condition is true ]
do
   some commands
done
```

factorial2.sh

```
#!/bin/bash

read counter
factorial=1
until [ $counter -le 1 ]; do
  factorial=$[ $factorial * $counter ]
  if [ $counter -eq 2 ]; then
    break
  else
    let counter-=2
  fi
done
echo $factorial
```

Switching Constructs - bash

- The `case` and `select` constructs are technically not loops since they do not iterate the execution of a code block
- Like loops, however, they direct program flow according to conditions at the top or bottom of the block

`case` construct

```
case variable in
  "condition1")
    some command
    ;;
  "condition2")
    some other command
    ;;
esac
```

`select` construct

```
select variable [ list ]
do
  command
  break
done
```

Switching Constructs - tcsh

- tcsh has the `switch` constructs

switch construct

```
switch (arg list)
  case "variable"
    some command
    breaksw
endsw
```

dooper.sh

```
#!/bin/bash


echo "Print two numbers"
read num1 num2
echo "What operation do you want to do?"

operations='add subtract multiply divide
        exponentiate modulo all quit'
select oper in $operations ; do
  case $oper in
    "add")
      echo "$num1 + $num2 =" $[$num1 + $num2]
      ;;
    "subtract")
      echo "$num1 - $num2 =" $[$num1 - $num2]
      ;;
    "multiply")
      echo "$num1 * $num2 =" $[$num1 * $num2]
      ;;
    "exponentiate")
      echo "$num1 ** $num2 =" $[$num1 ** $num2]
      ;;
    "divide")
      echo "$num1 / $num2 =" $[$num1 / $num2]
      ;;
    "modulo")
      echo "$num1 % $num2 =" $[$num1 % $num2]
      ;;
    "all")
      echo "$num1 + $num2 =" $[$num1 + $num2]
      echo "$num1 - $num2 =" $[$num1 - $num2]
      echo "$num1 * $num2 =" $[$num1 * $num2]
      echo "$num1 ** $num2 =" $[$num1 ** $num2]
      echo "$num1 / $num2 =" $[$num1 / $num2]
      echo "$num1 % $num2 =" $[$num1 % $num2]
    ;;
    *)
      exit
      ;;
  esac
done
```

dooper.csh

```
#!/bin/tcsh


echo "Print two numbers one at a time"
set num1 = $<
set num2 = $<
echo "What operation do you want to do?"
echo "Enter +, -, x, /, % or all"
set oper = $<

switch ( $oper )
  case "x"
     @ prod = $num1 * $num2
     echo "$num1 * $num2 = $prod"
     breaksw
  case "all"
     @ sum = $num1 + $num2
     echo "$num1 + $num2 = $sum"
     @ diff = $num1 - $num2
     echo "$num1 - $num2 = $diff"
     @ prod = $num1 * $num2
     echo "$num1 * $num2 = $prod"
     @ ratio = $num1 / $num2
     echo "$num1 / $num2 = $ratio"
     @ remain = $num1 % $num2
     echo "$num1 % $num2 = $remain"
     breaksw
  case "*"
     @ result = $num1 $oper $num2
     echo "$num1 $oper $num2 = $result"
     breaksw
endsw
```

```
~/Tutorials/BASH/scripts> ./day1/examples/dooper.sh
Print two numbers
1 4
What operation do you want to do?
1) add 3) multiply  5) exponentiate  7) all
2) subtract 4) divide  6) modulo   8) quit
#? 7
1 + 4 = 5
1 - 4 = -3
1 * 4 = 4
1 ** 4 = 1
1 / 4 = 0
1 % 4 = 1
#? 8
```

```
~/Tutorials/BASH/scripts> ./day1/examples/dooper.csh
Print two numbers one at a time
1
5
What operation do you want to do?
Enter +, -, x, /, % or all
all
1 + 5 = 6
1 - 5 = -4
1 * 5 = 5
1 / 5 = 0
1 % 5 = 1
```

Command Line Arguments (1)

- Similar to programming languages, bash and other shell scripting languages can also take command line arguments
  - Execute: `./myscript arg1 arg2 arg3`
  - Within the script, the positional parameters `$0, $1, $2, $3` correspond to `./myscript, arg1, arg2`, and `arg3`, respectively.
  - `$#`: number of command line arguments
  - `$*`: all of the positional parameters, seen as a single word
  - `$@`: same as `$*` but each parameter is a quoted string.
  - `shift N`: shift positional parameters from `N+1` to `$#` are renamed to variable names from `$1` to `$# - N + 1`
- In csh and tcsh
  - An array `argv` contains the list of arguments with `argv[0]` set to the name of the script
  - `#argv` is the number of arguments, i.e. length of `argv` array

shift.sh

```
#!/bin/bash

USAGE="USAGE: $0 <at least 1 argument>"

if [[ "$#" -lt 1 ]]; then
   echo $USAGE
   exit
fi

echo "Number of Arguments: " $#
echo "List of Arguments: " $@
echo "Name of script that you are running: " $0
echo "Command You Entered:" $0 $*

while [ "$#"  -gt 0 ]; do
  echo "Argument List is: " $@
  echo "Number of Arguments: " $#
  shift
done
```

```
-/Tutorials/BASH/scripts/day1/examples> ./shift.sh $(seq 1 5)
Number of Arguments:  5
List of Arguments:  1 2 3 4 5
Name of script that you are running:  ./shift.sh
Command You Entered: ./shift.sh 1 2 3 4 5
Argument List is:  1 2 3 4 5
Number of Arguments:  5
Argument List is:  2 3 4 5
Number of Arguments:  4
Argument List is:  3 4 5
Number of Arguments:  3
Argument List is:  4 5
Number of Arguments:  2
Argument List is:  5
Number of Arguments:  1
```

shift.csh

```
#!/bin/tcsh

set USAGE="USAGE: $0 <at least 1 argument>"

if ( "$#argv" < 1 ) then
   echo $USAGE
   exit
endif

echo "Number of Arguments: " $#argv
echo "List of Arguments: " ${argv}
echo "Name of script that you are running: " $0
echo "Command You Entered:" $0 ${argv}

while ( "$#argv"  > 0 )
  echo "Argument List is: " $*
  echo "Number of Arguments: " $#argv
  shift
end
```

```
-/Tutorials/BASH/scripts/day1/examples> ./shift.csh $(seq 1 5)
Number of Arguments:  5
List of Arguments:  1 2 3 4 5
Name of script that you are running:  ./shift.csh
Command You Entered: ./shift.csh 1 2 3 4 5
Argument List is:  1 2 3 4 5
Number of Arguments:  5
Argument List is:  2 3 4 5
Number of Arguments:  4
Argument List is:  3 4 5
Number of Arguments:  3
Argument List is:  4 5
Number of Arguments:  2
Argument List is:  5
Number of Arguments:  1
```

Declare command

- Use the `declare` command to set variable and functions attributes
- Create a constant variable, i.e. read-only
  - `declare -r var`
  - `declare -r varName=value`
- Create an integer variable
  - `declare -i var`
  - `declare -i varName=value`
- You can carry out arithmetic operations on variables declared as integers

```
~/Tutorials/BASH> j=10/5 ; echo $j
10/5
~/Tutorials/BASH> declare -i j; j=10/5 ; echo $j
2
```

Functions (1)

- Like “real” programming languages, bash has functions.
- A function is a code block that implements a set of operations, a “black box” that performs a specified task.
- Wherever there is repetitive code, when a task repeats with only slight variations in procedure, then consider using a function.

```
function function_name {
  command
}
OR
function_name () {
  command
}
```

shift10.sh

```
#!/bin/bash

usage () {
  echo "USAGE: $0 [atleast 11 arguments]"
  exit
}

[[ "$#" -lt 11 ]] && usage

echo "Number of Arguments: " $#
echo "List of Arguments: " $@
echo "Name of script that you are running: " $0
echo "Command You Entered:" $0 $*
echo "First Argument" $1
echo "Tenth and Eleventh argument" $10 $11 ${10} ${11}

echo "Argument List is: " $@
echo "Number of Arguments: " $#
shift 9
echo "Argument List is: " $@
echo "Number of Arguments: " $#
```

```
~/Tutorials/BASH/scripts/day1/examples> ./shift10.sh `seq 1 2 22`
Number of Arguments:  11
List of Arguments:  1 3 5 7 9 11 13 15 17 19 21
Name of script that you are running:  ./shift10.sh
Command You Entered: ./shift10.sh 1 3 5 7 9 11 13 15 17 19 21
First Argument 1
Tenth and Eleventh argument 10 11 19 21
Argument List is:  1 3 5 7 9 11 13 15 17 19 21
Number of Arguments:  11
Argument List is:  19 21
Number of Arguments:  2
```

Functions (2)

- You can also pass arguments to a function
- All function parameters can be accessed via `$1, $2, $3…`
- `$0` always point to the shell script name
- `$*` or `$@` holds all parameters passed to a function
- `$#` holds the number of positional parameters passed to the function

Functions (3)

- Array variable called `FUNCNAME` contains the names of all shell functions currently in the execution call stack.
- By default all variables are global.
- Modifying a variable in a function changes it in the whole script.
- You can create a local variables using the local command

```
local var=value
local varName
```

- A function may recursively call itself even without use of local variables.

factorial3.sh

```
#!/bin/bash

usage () {
  echo "USAGE: $0 <integer>"
  exit
}

factorial() {
  local i=$1
  local f

  declare -i i
  declare -i f

  if [[ "$i" -le 2 && "$i" -ne 0 ]]; then
    echo $i
  elif [[ "$i" -eq 0 ]]; then
    echo 1
  else
    f=$(( $i - 1 ))
    f=$( factorial $f )
    f=$(( $f * $i ))
    echo $f
  fi
}

if [[ "$#" -eq 0 ]]; then
  usage
else
  for i in $@ ; do
    x=$( factorial $i )
    echo "Factorial of $i is $x"
  done
fi
```

```
~/Tutorials/BASH/scripts/day1/examples>./factorial3.sh 1 3 5 7 9 15
Factorial of 1 is 1
Factorial of 3 is 6
Factorial of 5 is 120
Factorial of 7 is 5040
Factorial of 9 is 362880
Factorial of 15 is 1307674368000
```

Outline

- Introduction to Linux Shell
- Shell Scripting Basics
- Beyond Basic Shell Scripting
  - Arithmetic Operations
  - Arrays
  - Flow Control
  - Command Line Arguments
  - Functions
- Advanced Text Processing Commands

Advanced Text Processing Commands

- `grep` & `egrep`
- `sed`
- `awk`

grep & egrep

- `grep` is a Unix utility that searches through either information piped to it or files.
- `egrep` is extended grep (extended regular expressions), same as `grep -E`
- Use `zgrep` for compressed files.
- Usage: `grep <options> <search pattern> <files>`
- Commonly used options

`-i`	ignore case during search
`-r,-R`	search recursively
`-v`	invert match i.e. match everything except *`pattern`*
`-l`	list files that match *`pattern`*
`-L`	list files that do not match *`pattern`*
`-n`	prefix each line of output with the line number within its input file.
`-A num`	print `num` lines of trailing context after matching lines.
`-B num`	print `num` lines of leading context before matching lines.

grep Examples

- Search files that contain the word `node` in the examples directory

```
egrep node *

checknodes.pbs:#PBS -o nodetest.out
checknodes.pbs:#PBS -e nodetest.err
checknodes.pbs:for nodes in "${NODES[@]}"; do
checknodes.pbs: ssh -n $nodes 'echo $HOSTNAME '$i' ' &
checknodes.pbs:echo "Get Hostnames for all unique nodes"
```

- Repeat above search using a case insensitive pattern match and print line number that matches the search pattern

```
egrep -in node *

checknodes.pbs:20:NODES=(`cat "$PBS_NODEFILE"` )
checknodes.pbs:21:UNODES=(`uniq "$PBS_NODEFILE"` )
checknodes.pbs:23:echo "Nodes Available: " ${NODES[@]}
checknodes.pbs:24:echo "Unique Nodes Available: " ${UNODES[@]}
checknodes.pbs:28:for nodes in "${NODES[@]}"; do
checknodes.pbs:29: ssh -n $nodes 'echo $HOSTNAME '$i' ' &
checknodes.pbs:34:echo "Get Hostnames for all unique nodes"
checknodes.pbs:39: ssh -n ${UNODES[$i]} 'echo $HOSTNAME '$i' '
```

sed

- `sed` ("stream editor") is Unix utility for parsing and transforming text files.
  - Also works for either information piped to it or files
- `sed` is line-oriented - it operates one line at a time and allows regular expression matching and substitution.
- `sed` has several commands, the most commonly used command and sometime the only one learned is the substitution command, *s*

```
> echo day | sed 's/day/night/'
> night
```

List of sed commands and flags

Flags	Operation	Command	Operation
`-e`	combine multiple commands	`s`	substitution
`-f`	read commands from file	`g`	global replacement
`-h`	print help info	`p`	print
`-n`	disable print	`i`	ignore case
`-V`	print version info	`d`	delete
`-r`	use extended regex	`G`	add newline
		`w`	write to file
		`x`	exchange pattern with hold buffer
		`h`	copy pattern to hold buffer
		`;`	separate commands

sed Examples (1)

- Add the -e to carry out multiple matches.

```
cat hello.sh | sed -e 's/bash/tcsh/g' -e 's/First/First tcsh/g'

#!/bin/tcsh
# My First tcsh Script
echo "Hello World!"
```

- Alternate form

```
sed 's/bash/tcsh/g; s/First/First tcsh/g' hello.sh

#!/bin/tcsh
# My First tcsh Script
echo "Hello World!"
```

- The default delimiter is slash (/), but you can change it to whatever you want which is useful when you want to replace path names

```
sed 's:/bin/bash:/bin/tcsh:g' hello.sh

#!/bin/tcsh
# My First Script
echo "Hello World!"
```

sed Examples (2)

- `sed` can also delete blank lines from a file

```
sed '/^$/d' hello.sh

#!/bin/bash
# My First Script
echo "Hello World!"
```

- Delete line `n` through `m` in a file

```
sed '2,4d' hello.sh

#!/bin/bash
echo "Hello World!"
```

- Insert a blank line above every line which matches *`pattern`*

```
sed '/First/{x;p;x}' hello.sh

#!/bin/bash


# My First Script

echo "Hello World!"
```

sed Examples (3)

- Insert a blank line below every line which matches *pattern*

```
sed '/First/G' hello.sh

#!/bin/bash

# My First Script


echo "Hello World!"
```

- Insert a blank line above and below every line which matches *pattern*

```
sed '/First/{x;p;x;G}' hello.sh

#!/bin/bash


# My First Script


echo "Hello World!"
```

sed Examples (4)

- Print only lines which match *pattern* (emulates `grep`)

```
sed -n '/echo/p' hello.sh

echo "Hello World!"
```

- Print only lines which do NOT match *pattern* (emulates `grep -v`)

```
sed -n '/echo/!p' hello.sh

#!/bin/bash
# My First Script
```

- Print current line number to standard output

```
sed -n '/echo/ =' quotes.sh

5
6
7
8
9
10
11
12
13
```

awk

- The `awk` text-processing language is useful for such tasks as:
  - Tallying information from text files and creating reports from the results.
  - Adding additional functions to text editors like "vi".
  - Translating files from one format to another.
  - Creating small databases.
  - Performing mathematical operations on files of numeric data.
- `awk` has two faces:
  - It is a utility for performing simple text-processing tasks, and
  - It is a programming language for performing complex text-processing tasks.
- `awk` comes in three variations
  - `awk` : Original AWK by A. Aho, B. W. Kernighnan and P. Weinberger from AT&T
  - `nawk` : New AWK, also from AT&T
  - `gawk` : GNU AWK, all Linux distributions come with `gawk`. In some distros, `awk` is a symbolic link to `gawk`.

awk Syntax

- Simplest form of using `awk`
  - `awk pattern {action}`
    - `pattern` decides when `action` is performed
  - Most common action: `print`
  - Print file dosum.sh: `awk '{print $0}' dosum.sh`
  - Print line matching bash in all `.sh` files in current directory: `awk '/bash/{print $0}' *.sh`

Awk Examples

- Print list of files that are csh script files

```
awk '/^#\!\/bin\/tcsh/{print FILENAME}' *

dooper.csh
factorial.csh
hello1.sh
name.csh
nestedloops.csh
quotes.csh
shift.csh
```

- Print contents of hello.sh that lie between two patterns

```
awk '/^#\!\/bin\/bash/,/echo/{print $0}' hello.sh

#!/bin/bash
# My First Script
echo "Hello World!"
```

How awk Works

- `awk` reads the file being processed line by line.
- The entire content of each line is split into columns with space or tab as the delimiter. The delimiter can be changed as will be seen in the next few slides.
- To print the entire line, use `$0`.
- The intrinsic variable `NR` contains the number of records (lines) read.
- The intrinsic variable `NF` contains the number of fields or columns in the current line.
- By default the field delimiter is space or tab. To change the field delimiter use the `-F<delimiter>` command.

```
uptime

11:18am up 14 days 0:40, 5 users, load average: 0.15, 0.11, 0.17

uptime | awk '{print $1,NF}'

11:19am 0.17

uptime | awk -F: '{print $1,NF}'

11 0.12, 0.10, 0.16

for i in $(seq 1 10); do touch file${i}.dat ; done
ls file*

file10.dat file2.dat file4.dat file6.dat file8.dat
file1.dat file3.dat file5.dat file7.dat file9.dat

for i in file* ; do
> prefix=$(echo $i | awk -F. '{print $1}')
> suffix=$(echo $i | awk -F. '{print NF}')
> echo $prefix $suffix $i
> done

file10 dat file10.dat
file1 dat file1.dat
file2 dat file2.dat
file3 dat file3.dat
file4 dat file4.dat
file5 dat file5.dat
file6 dat file6.dat
file7 dat file7.dat
file8 dat file8.dat
file9 dat file9.dat
```

Arithmetic Operations (1)

- `awk` has in-built support for arithmetic operations

Operator	Operation	Operator	Operation
+	Addition	++	Autoincrement
-	Subtraction	--	Autodecrement
*	Multiplication	+=	Add to
/	Division	-=	Subtract from
**	Exponentiation	*=	Multiple with
%	Modulo	/=	Divide by

```
echo | awk '{print 10%3}'
1
echo | awk '{a=10;print a/=5}'
2
```

Conditionals and Loops (1)

- `awk` supports
  - `if ... else if .. else` conditionals.
  - `while` and `for` loops
- They work similar to that in C-programming
- Supported operators: ==, !=, >, >=, <, <=, ~ (string matches), !~ (string does not match)

```
awk '{if (NR > 0 ){print NR,":", $0}}' hello.sh

1 : #!/bin/bash
2 :
3 : # My First Script
4 :
5 : echo "Hello World!"
```

Conditionals and Loops (2)

- The `for` command can be used for processing the various columns of each line

```
cat << EOF | awk '{for (i=1;i<=NF;i++){if (i==1){a=$i}else if (i==NF){print a}else{a+=$i}}}'

1 2 3 4 5 6
7 8 9 10
EOF
15
24

echo $(seq 1 10) | awk 'BEGIN{a=6}{for (i=1;i<=NF;i++){a+=$i}}END {print a}'

61
```

Further Reading

- BASH Programming http://tldp.org/HOWTO/Bash-Prog-Intro-HOWTO.html
- Advanced Bash-Scripting Guide http://tldp.org/LDP/abs/html/
- Regular Expressions http://www.grymoire.com/Unix/Regular.html
- AWK Programming http://www.grymoire.com/Unix/Awk.html
- awk one-liners: http://www.pement.org/awk/awk1line.txt
- sed http://www.grymoire.com/Unix/Sed.html
- sed one-liners: http://sed.sourceforge.net/sed1line.txt
- CSH Programming http://www.grymoire.com/Unix/Csh.html
- csh Programming Considered Harmful
- http://www.faqs.org/faqs/unix-faq/shell/csh-whynot/
- Wiki Books http://en.wikibooks.org/wiki/Subject:Computing

Exercises

1. Write a shell script to
   - Print “Hello world!” to the screen
   - Use a variable to store the greeting
2. Write a shell script to
   - Take two integers on the command line as arguments
   - Print the sum, different, product of those two integers
   - Think: what if there are too few or too many arguments? How can you check that?
3. Write a shell script to read your first and last name to an array
   - Add your salutation and suffix to the array
   - Drop either the salutation or suffix
   - Print the array after each of the three steps above
4. Write a shell script to calculate the factorial and double factorial of an integer or list of integers

Next Tutorial – Distributed Job Execution

- If any of the following fits you, then you might want come
  - I have to run more than one serial job.
  - I don't want to submit multiple job using the serial queue
  - How do I submit one job which can run multiple serial jobs?
- Date: Sept 30th, 2015

Getting Help

- User Guides
  - LSU HPC: http://www.hpc.lsu.edu/docs/guides.php#hpc
  - LONI:http://www.hpc.lsu.edu/docs/guides.php#loni
- Documentation: http://www.hpc.lsu.edu/docs
- Online courses: http://moodle.hpc.lsu.edu
- Contact us
  - Email ticket system: sys-help@loni.org
  - Telephone Help Desk: 225-578-0900
  - Instant Messenger (AIM, Yahoo Messenger, Google Talk)
    - Add “lsuhpchelp”