Basic Shell Scripting

Dr. Zach Byerly
HPC User Support Specialist

LSU & LONI HPC
LSU HPC: hpc.lsu.edu
LONI: loni.org
sys-help@loni.org

Louisiana State University
Baton Rouge
February 9, 2022
Outline

- *Introduction to Linux Shell*
- Shell Scripting Basics
  - Variables/Special Characters
  - Arithmetic Operations
- Beyond Basic Shell Scripting
  - Control flow
  - Functions
- Advanced Text Processing Commands (grep, sed, awk)
What is a Linux Shell?

- An application running on top of the kernel and provides a command line interface to the system
- Types of shell with varied features
  - **sh**
    - the original Bourne shell.
  - **ksh**
    - one of the three: Public domain ksh (pdksh), AT&T ksh or mksh
  - **bash**
    - the GNU Bourne-again shell. It is mostly Bourne-compatible, mostly POSIX-compatible, and has other useful extensions. It is the default on most Linux systems.
  - **csh**
    - BSD introduced the C shell, which sometimes resembles slightly the C programming language.
  - **tcsh**
    - csh with more features. csh and tcsh shells are NOT Bourne-compatible.
What can you do with a shell?

- Check the current shell you are using
  - `echo $0`
- List available shells on the system
  - `cat /etc/shells`
- Change to another shell
  - `csh`
- Date
  - `date`
- `wget`: get online files
  - `wget https://website.com/filename.tgz`
- Compile and run applications
  - `gcc hello.c -o hello`
  - `./hello`
- Automate lots of commands using a script
- Use the shell script to run jobs
Shell Scripting

- Script: a program written for a software environment to automate execution of tasks
  - A series of shell commands put together in a file
  - When the script is executed, those commands will be executed one line at a time automatically
  - Shell script is interpreted, not compiled.

- The majority of script programs are “quick and dirty”, where the main goal is to get the program written quickly
  - Often, we write scripts and only use them ourselves
  - Shell scripts can be made robust so that many other people can use them over and over.
When NOT to use Shell Scripting...

- Selected situations:
  - Resource-intensive tasks, especially where speed is a factor (sorting, hashing, recursion [2] ...)
  - Procedures involving heavy-duty math operations, especially floating point arithmetic, arbitrary precision calculations, or complex numbers (use C++ or FORTRAN instead)
  - Complex applications, where structured programming is a necessity (type-checking of variables, function prototypes, etc.)
  - Extensive file operations required (Bash is limited to serial file access, and that only in a particularly clumsy and inefficient line-by-line fashion.)
  - Need native support for multi-dimensional arrays, data structures, such as linked lists or trees
  - Need to use libraries or interface with legacy code
#!/bin/bash

# My first bash script
# by Zach Byerly

echo "Hello World!"

[mtiger@smic training]$: bash hello_world.sh # run using bash
Hello World!
[mtiger@mike2 training]$: chmod +x hello-world.sh # make executable
[mtiger@mike2 training]$: ./hello-world.sh # execute the script
Hello World!
Outline

• Introduction to Linux Shell
• Shell Scripting Basics
  • Variables/Special Characters
  • Arithmetic Operations
• Beyond Basic Shell Scripting
  • Flow Control
  • Functions
• Advanced Text Processing Commands (grep, sed, awk)
**Variables**

- Rules for variable names
  - Must start with a letter or underscore
  - Number can be used anywhere else
  - Do not use special characters such as @,#,%,$
  - Case sensitive
  - Allowed: VARIABLE, VAR1234able, var_name, _VAR
  - *Not* allowed: lvar, %name, $myvar, var@NAME, myvar-1
- To reference a variable, prepend $ to the name of the variable
- Example: $PATH, $LD_LIBRARY_PATH, $myvar etc.
Global and Local Variables

- Two types of variables:
  - Global (Environmental) variables
    - Inherited by subshells (child process, see next slide)
    - Provide a simple way to share configuration settings between multiple applications and processes in Linux
    - Using all uppercase letters by convention
    - Example: `PATH, LD_LIBRARY_PATH, DISPLAY` etc.
    - `printenv/env` list the current environmental variables in your system.
  - Local (shell) variables
    - Only visible to the current shell
    - Not inherited by subshells
- Local (Shell) variables are only valid within the current shell, while environment variables are valid for all subsequently opened shells.
- Example: useful when running a script, where exported variables (global) at the terminal can be inherited within the script.

### Editing Variables

<table>
<thead>
<tr>
<th>Type</th>
<th>sh/ksh/bash</th>
<th>csh/tcsh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shell (local)</td>
<td>name=value</td>
<td>set name=value</td>
</tr>
<tr>
<td>Environment (global)</td>
<td>export name=value</td>
<td>setenv name value</td>
</tr>
</tbody>
</table>

#### With export

```
$ export v1=one
$ bash
$ echo $v1
one
```

#### Without export

```
$ v1=one
$ bash
$ echo $v1
$ echo $v1
```
How to inherit variables in the script?

Using the `source` command, it has a synonym in dot "." (period)

Syntax:

```bash
#!/bin/bash
export myvar="newvalue"

source source_var.sh
```

```
newvalue
```
### List of Some Environment Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PATH</td>
<td>A list of directory paths which will be searched when a command is issued</td>
</tr>
<tr>
<td>LD_LIBRARY_PATH</td>
<td>Colon-separated set of directories where libraries should be searched for first</td>
</tr>
<tr>
<td>HOME</td>
<td>Indicate where a user's home directory is located in the file system.</td>
</tr>
<tr>
<td>PWD</td>
<td>Contains path to current working directory.</td>
</tr>
<tr>
<td>OLDPWD</td>
<td>Contains path to previous working directory.</td>
</tr>
<tr>
<td>TERM</td>
<td>Specifies the type of computer terminal or terminal emulator being used</td>
</tr>
<tr>
<td>SHELL</td>
<td>Contains name of the running, interactive shell.</td>
</tr>
<tr>
<td>PS1</td>
<td>Default command prompt</td>
</tr>
<tr>
<td>PS2</td>
<td>Secondary command prompt</td>
</tr>
<tr>
<td>HOSTNAME</td>
<td>The system's host name</td>
</tr>
<tr>
<td>USER</td>
<td>Current logged in user's name</td>
</tr>
<tr>
<td>DISPLAY</td>
<td>Network name of the X11 display to connect to, if available.</td>
</tr>
</tbody>
</table>
Quotations

• Single quotation
  – Enclosing characters in single quotes (') preserves the literal value of each character within the quotes. A single quote may not occur between single quotes, even when preceded by a backslash.

• Double quotation
  – Enclosing characters in double quotes (" ) preserves the literal value of all characters within the quotes, with the exception of ‘$’, ‘`’, ‘\’

• Back “quotation?”
  – Command substitution (` `) allows the output of a command to replace the command itself, enclosed string is executed as a command, almost the same as $( )
Always use double quotes around variable substitutions and command substitutions: "$foo", "${foo}"
# | Start a comment line.
$ | Indicate the name of a variable.
\ | Escape character to display next character literally
{} | 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)
| | Pipe: use the output of a command as the input of another one
> | Redirections (0<: standard input; 1>: standard out; 2>: standard error)
### Special Characters

<table>
<thead>
<tr>
<th>Character</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>??</code></td>
<td>Exit status for the last command, 0 is success, failure otherwise</td>
</tr>
<tr>
<td><code>$$</code></td>
<td>Process ID variable.</td>
</tr>
<tr>
<td><code>[ ]</code></td>
<td>Test expression, eg. if condition</td>
</tr>
<tr>
<td><code>[[ ]]]</code></td>
<td>Extended test expression, more flexible than <code>[ ]</code></td>
</tr>
<tr>
<td><code>$[ ]</code>, <code>$(( ))</code></td>
<td>Integer expansion</td>
</tr>
<tr>
<td>`</td>
<td></td>
</tr>
</tbody>
</table>
Outline

• Introduction to Linux Shell
• Shell Scripting Basics
  • Variables/Special Characters
  • Arithmetic Operations
• Beyond Basic Shell Scripting
  • Flow Control
  • Functions
• Advanced Text Processing Commands (grep, sed, awk)
## Integer Arithmetic Operations

<table>
<thead>
<tr>
<th>Operation</th>
<th>Operator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Addition</td>
<td>+</td>
</tr>
<tr>
<td>Subtraction</td>
<td>-</td>
</tr>
<tr>
<td>Multiplication</td>
<td>*</td>
</tr>
<tr>
<td>Division</td>
<td>/</td>
</tr>
<tr>
<td>Exponentiation</td>
<td>** (bash only)</td>
</tr>
<tr>
<td>Modulo</td>
<td>%</td>
</tr>
</tbody>
</table>
Integer Arithmetic Operations

- `$(...)` or `[$...]` commands
  - `x=$(1+2)` # Addition, suggested
  - `echo ${x}*${x}` # Multiplication, deprecated

- let command:
  - `let c=\$x+\$x` # no space
  - `let c=x+x` # you can omit the \$ sign
  - `let c="x + x"` # can have space
  - `let c+=1` or `let --c` # C-style increment operator

- expr command:
  - `expr 10 / 2` (space required)

Note: Bash is picky about spaces!
Floating-Point

GNU basic calculator (bc) external calculator

- Add two numbers
  \[ \text{echo } "3.8 + 4.2" \mid \text{bc} \]
- Divide two numbers and print result with a precision of 5 digits:
  \[ \text{echo } "\text{scale}=5; 2/5" \mid \text{bc} \]
- Convert between decimal and binary numbers
  \[ \text{echo } "\text{ibase}=10; \text{obase}=2; 10" \mid \text{bc} \]
- Call bc directly:
  \[ \text{bc} \ll "\text{scale}=5; \text{sqrt}(2)" \]
Outline

• Introduction to Linux Shell
• Shell Scripting Basics
  • Variables/Special Characters
  • Arithmetic Operations
• Beyond Basic Shell Scripting
  • *Flow Control*
  • Functions
• Advanced Text Processing Commands (grep, sed, awk)
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
  - Conditionals:
    - if-then-else
    - Switches: case
  - Loops: for, while, until
If Statement

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

```bash
if [ condition ]; then
    Do something
elif [ condition 2 ]; then
    Do something
else
    Do something else
fi
```

- Strict spaces between condition and the brackets (bash)
- `[[ condition ]] extended test` construct is the more versatile Bash version of `[ condition ]`, generally safer to use.
### File Operations

<table>
<thead>
<tr>
<th>Operation</th>
<th>bash</th>
</tr>
</thead>
<tbody>
<tr>
<td>File exists</td>
<td><code>if [ -e test ]</code></td>
</tr>
<tr>
<td>File is a regular file</td>
<td><code>if [ -f test]</code></td>
</tr>
<tr>
<td>File is a directory</td>
<td><code>if [ -d /home ]</code></td>
</tr>
<tr>
<td>File is not zero size</td>
<td><code>if [ -s test ]</code></td>
</tr>
<tr>
<td>File has read permission</td>
<td><code>if [ -r test ]</code></td>
</tr>
<tr>
<td>File has write permission</td>
<td><code>if [ -w test ]</code></td>
</tr>
<tr>
<td>File has execute permission</td>
<td><code>if [ -x test ]</code></td>
</tr>
</tbody>
</table>
## Integer Comparisons

<table>
<thead>
<tr>
<th>Operation</th>
<th>bash</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal to</td>
<td>if [ 1 –eq 2 ]</td>
</tr>
<tr>
<td>Not equal to</td>
<td>if [ $a –ne $b ]</td>
</tr>
<tr>
<td>Greater than</td>
<td>if [ $a –gt $b ]</td>
</tr>
<tr>
<td>Greater than or equal to</td>
<td>if [ 1 –ge $b ]</td>
</tr>
<tr>
<td>Less than</td>
<td>if [ $a –lt 2 ]</td>
</tr>
<tr>
<td>Less than or equal to</td>
<td>if [ $a –le $b ]</td>
</tr>
</tbody>
</table>
## String Comparisons

<table>
<thead>
<tr>
<th>Operation</th>
<th>bash</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equal to</td>
<td><code>if [ $a == $b ]</code></td>
</tr>
<tr>
<td>Not equal to</td>
<td><code>if [ $a != $b ]</code></td>
</tr>
<tr>
<td>Zero length or null</td>
<td><code>if [ -z $a ]</code></td>
</tr>
<tr>
<td>Non zero length</td>
<td><code>if [ -n $a ]</code></td>
</tr>
</tbody>
</table>
## Logical Operators

<table>
<thead>
<tr>
<th>Operation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>! (NOT)</td>
<td>if [ ! -e test ]</td>
</tr>
<tr>
<td>&amp;&amp; (AND)</td>
<td>if [ -f test ] &amp;&amp; [ -s test ]</td>
</tr>
<tr>
<td></td>
<td>if [[ -f test &amp;&amp; -s test ]]</td>
</tr>
<tr>
<td></td>
<td>if ( -e test &amp;&amp; ! -z test )</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>if [[ -f test1</td>
</tr>
</tbody>
</table>
If Condition Examples

Example 1:
read input
if [ $input == "hello" ]; then
  echo hello;
else echo wrong;
fi

Example 2

touch test.txt
if [ -e test.txt ]; then
  echo "file exist"
elif [ ! -s test.txt ]; then
  echo "file empty"
fi

What happens after

echo "hello world" >> test.txt
Loop Constructs

• A loop is a block of code that iterates a list of commands as long as the loop control condition stays true
• Loop constructs
  for, while and until
For Loop Examples

Example1:
for arg in `seq 1 4`
do
echo $arg;
touch test.$arg
done

How to delete test files using a loop?
rm test. [1-4]

Example 2:
for file in `ls /home/$USER`
do
cat $file
done
While Loop

- The `while` construct test 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` is used when loop repetitions is not known beforehand.

```bash
read counter
while [ $counter -ge 0 ]
  do let counter--
     echo $counter
  done
```
Outline

• Introduction to Linux Shell
• Shell Scripting Basics
  • Variables/Special Characters
  • Arithmetic Operations
• Beyond Basic Shell Scripting
  • Flow Control
  • *Functions*
• Advanced Text Processing Commands (grep, sed, awk)
Functions

• A function is a code block that implements a set of operations. Code reuse by passing parameters,
  • Syntax:
    
    ```
    function_name () {
        command...
    }
    ```
  • By default all variables are global.
  • Modifying a variable in a function changes it in the whole script.
  • Create a local variables using the `local` command, which is invisible outside the function
    ```
    local var=value
    local varName
    ```
Pass Arguments to Bash Scripts

- Note the difference between the arguments passed to the script and the function.
- All parameters can be passed at runtime and accessed via $1, $2, $3..., add {} when $>=10
- $0: the shell script name
- Array variable called `FUNCNAME` contains the names of all shell functions currently in the execution call stack.
- `$*` or `@`: all parameters passed to a function
- `#$`: number of positional parameters passed to the function
- `$?`: exist code of last command
- `$$`: PID of current process
Function example

#!/bin/bash

func_add () # define a simple function
{
    local x=$1 # 1st argument to the function
    local y=$2 # 2nd argument to the function
    result=$((x + y))
    # echo "result is: " $result
}

a=3; b=4
echo "a= $a, b= $b"
result="nothing"

echo "result before calling the function is: " $result
func_add $a $b # note this is arguments to the function

echo "result by passing function arguments is: " $result
func_add $1 $2 # note this is command line arguments

echo "result by passing command line arguments is: " $result
Outline

• Introduction to Linux Shell
• Shell Scripting Basics
  • Variables/Special Characters
  • Arithmetic Operations
• Beyond Basic Shell Scripting
  • Flow Control
  • Functions
• Advanced Text Processing Commands (grep, sed, awk)
Advanced Text Processing Commands

- grep
- sed
- awk
grep & egrep

• **grep**: Unix utility that searches a pattern through either information piped to it or files.
• **egrep**: extended grep, same as `grep -E`
• **zgrep**: compressed files.

**Usage:**
```bash
grep <options> <search pattern> <files>
```

**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 containing the word `bash` in current directory
  
  ```bash
grep bash *
  ```

- Search files NOT containing the word `bash` in current directory
  
  ```bash
grep -v bash *
  ```

- Repeat above search using a case insensitive pattern match and print line number that matches the search pattern
  
  ```bash
grep -iN bash *
  ```

- Search files not matching certain name pattern
  
  ```bash
ls | grep -vi fun
  ```
grep Examples

100 Thomas Manager Sales $5,000
200 Jason Developer Technology $5,500
300 Raj Sysadmin Technology $7,000
500 Randy Manager Sales $6,000

• grep OR

grep ‘Man\|Sales’ employee.txt
-> 100 Thomas Manager Sales $5,000
  300 Raj Sysadmin Technology $7,000
  500 Randy Manager Sales $6,000

• grep AND

grep –i ‘sys.*Tech’ employee.txt
-> 100300 Raj Sysadmin Technology $7,000
**sed**

- "stream editor" to parse and transform information
  - information piped to it or from files
- line-oriented, operate one line at a time and allow regular expression matching and substitution.
- \$ substitution command
## sed Commands and Flags

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

#!/bin/bash

# My First Script

echo "Hello World!"
• Delete blank lines from a file

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

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

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

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

```bash
#!/bin/bash
echo "Hello World!"
```
• Add flag `-e` to carry out multiple matches.

```bash
# Example of using `sed` with `-e`
cat hello.sh | sed -e 's/bash/tcsh/g' -e 's/First/Second/g'
```

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

• Alternate form

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

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

• The default delimiter is slash (`/`), can be changed

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

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

- Replace-in-place with a backup file
  
  ```bash
  sed -i.bak '/First/Second/i' hello.sh
  ```

- echo with sed
  
  ```bash
  $ echo "shell scripting" | sed "s/[si]/?/g"
  $$hell $$cr$$pt$$ng
  
  $ echo "shell scripting 101" | sed "s/[0-9]/#/g"
  $ shell scripting ###
  ```
awk

- The *awk* text-processing language is useful for tasks such 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.
How Does awk Work

- `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.
- `$0` Print the entire line
- `$1, $2, $3, ...` for each column (if exists)
- `NR` number of records (lines)
- `NF` 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 $0}'
11:18am up 14 days 0:40, 5 users, load average: 0.15, 0.11, 0.17

uptime | awk '{print $1,NF}'
11:18am 12

uptime | awk '{print NR}'
1

uptime | awk -F, '{print $1}'
11:18am up 14 days 0:40

for i in $(seq 1 3); do touch file${i}.dat ; done
for i in file* ; do
  > prefix=$(echo $i | awk -F. '{print $1}')
  > suffix=$(echo $i | awk -F. '{print $NF}')
  > echo $prefix $suffix $i; done

file1 dat file1.dat
file2 dat file2.dat
file3 dat file3.dat
Getting Help

- User Guides
  - LSU HPC: [http://www.hpc.lsu.edu/docs/guides.php#hpc](http://www.hpc.lsu.edu/docs/guides.php#hpc)
  - LONI: [http://www.hpc.lsu.edu/docs/guides.php#loni](http://www.hpc.lsu.edu/docs/guides.php#loni)
- Documentation: [http://www.hpc.lsu.edu/docs](http://www.hpc.lsu.edu/docs)