Bash Shell Scripting

Advance Shell Scripting - I

Persistent University

Contents

Advance Shell Scripting - I

1. At the end of this module , you will be able to understand:

2. Filters:

3. Commonly used Filters:

   sort

   uniq

   head

   tail

   wc

   Tr,Pr, Split, cut

Filters

- **Filters** : If a Linux command accepts its input data from the standard input and produces its output (result) on standard output is known as a filter.

- Filters usually works with pipes.

- **Syntax:**
  - command1 | command2
  - command1 file.txt | command2
  - command1 args < input.txt | command2

- Where command2 is a filter command.

- **Example:-**
  - cut -d: -f1 /etc/passwd | sort | uniq | grep vivek

- In this example grep command acts as a filter

Persistent

Filters contd…

- Filters
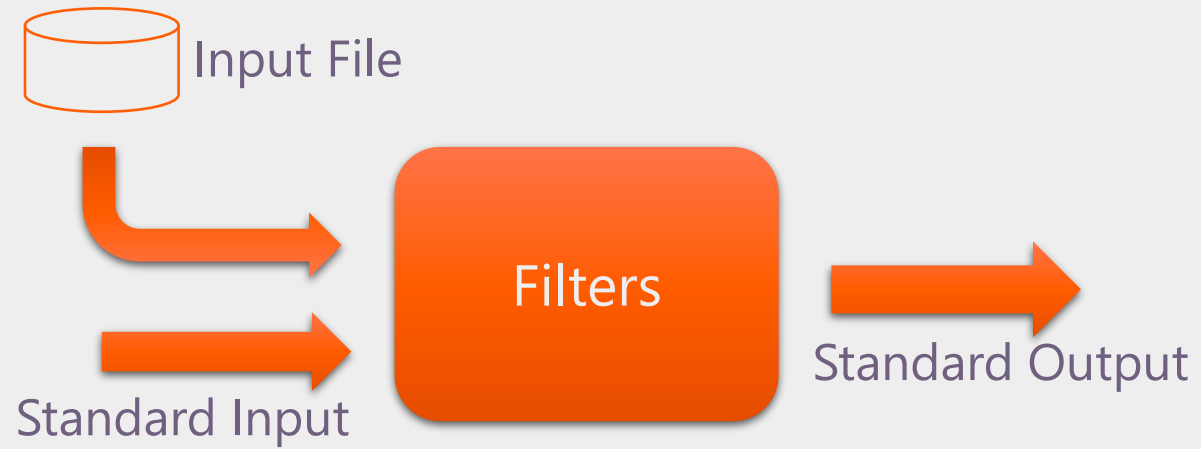
- sort
- tail
- wc

- uniq
- pr
- cut

- head
- tr
- split

Commonly Used Filters

Filter	What it does
sort	Sorts input lines.
uniq	Discards all but one of successive identical lines.
head	Shows beginning of a file (default 10 lines).
tail	Shows end of a file (default 10 lines).
wc	Counts characters, words and lines.
tr	Translates or deletes specified character sets.
pr	Pr command is used for printing files
split	Split is used to split large file into two or more files
cut	Cut command is used for text processing

Commonly Used Filters

Filter	What it does
sort	Sorts input lines.
uniq	Discards all but one of successive identical lines.
head	Shows beginning of a file (default 10 lines).
tail	Shows end of a file (default 10 lines).
wc	Counts characters, words and lines.
tr	Translates or deletes specified character sets.
pr	Pr command is used for printing files
split	Split is used to split large file into two or more files
cut	Cut command is used for text processing

Sort

- Sorts standard input then outputs the sorted result on standard output.

- Sort is a simple and very useful command which will rearrange lines in a text file so that they are sorted numerically and alphabetically

- Example:

$**cat input_file**
Unix
Linux
Solaris
Linux
AIX

$**sort input_file**
AIX
Linux
Linux
Solaris
Unix

Uniq

- Discards identical lines in a file.

- Example:

$**cat sample.txt**

unix operating system

unix operating system

unix dedicated server

linux dedicated server

$**uniq sample.txt**

unix operating system

unix dedicated server

linux dedicated server

$**uniq -d sample.txt**

unix operating system

$**uniq -u sample.txt**

unix dedicated server

linux dedicated server

$**uniq –c sample.txt**

2 unix operating system

1 unix dedicated server

1 linux dedicated server

Persistent

Uniq

- Please find below details of the uniq command options:

  c   : Count of occurrence of each line.

  d   : Prints only duplicate lines.

  D   : Print all duplicate lines

  f   : Avoid comparing first N fields.

  i   : Ignore case when comparing.

  s   : Avoid comparing first N characters.

  U   : Prints only unique lines.

  w   : Compare no more than N characters in lines

Head

- Prints the first N lines from the file.

- Example:

$**cat days**
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday

$**head –n3 days**
Sunday
Monday
Tuesday

$**head –c5 days**
Sunda

$**head –n-3 days**
Sunday
Monday
Tuesday
Wednesday

Tail

- Prints the last N lines from the file.

- Example:

$**cat days**
Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday

$**tail –n3 days**
Thursday
Friday
Saturday

$**tail –c5 days**
rday

$**head -5 days | tail -3**
Tuesday
Wednseday
Thurseday

$**tail –n-3 days**
Tuesday
Wednesday
Thursday
Friday
Saturday

WC (word count)

- WC or word count command is used to display newline count, word count, byte and characters count in a file

- Finds out newline count, word count, byte and characters count in a file

    wc -l : Prints the number of lines in a file.

    wc -w : prints the number of words in a file.

    wc -c : Displays the count of bytes in a file.

    wc -m : prints the count of characters from a file.

    wc -L : prints only the length of the longest line in a file.

Persistent

Wc (word count)

- Lets see :

examples of wc command

$ **cat file.txt**
Red Hat
CentOS
Fedora
Debian
Scientific Linux
OpenSuse
Ubuntu
Xubuntu
Linux Mint
Pearl Linux
Slackware
Mandriva

$ **wc file.txt**
 12 16 112 os.txt

$ **wc –l file.txt**
 12 os.txt

$ **wc –w file.txt**
 16 os.txt

$ **wc –c file.txt**
 112 os.txt

Pr and Tr

- pr - Paginate or columnates file for printing.
    - pr files.txt | lpr

- tr – translates, deletes or squeeze repeated characters.
    - echo "hello world" | tr "[a-z]" "[A-Z]"
    - echo "hello world" | tr "[:lower:]" "[:upper:]"

Cut

- Cut command is used for text processing. This command is used to extract portion of text from a file by selecting columns. Lets perform operations on test.txt file

- Contents of test.txt file are as:

- $ Cat test.txt

  - cat command for file oriented operations.

  - cp command for copy files or directories.

  - ls command to list out files and directories with its attributes.

- 1) Cut command is used to select column of characters.

  - To extract only a desired column from a file use –c option

  - The following example displays 2nd character from each line of a file test.txt

  - $ cut -c2 test.txt

  - A

  - p

  - s

Cut Command Contd…

- In the similar way **Cut** command is used for following things:
    - Cut is used to select column of characters using range.
    - Select Column of Characters using either Start or End Position.
    - Cut is used to Select a Specific Field from a File.
    - Cut is used to Select Multiple Fields from a File.
    - Cut is used to Select Fields Only When a Line Contains the Delimiter.
    - Cut is used to Select All Fields Except the Specified Fields.

Persistent

Split

- Splits file into multiple files.

- Example:

$ **cat file**	$ **split file**	$ **split -l 3 file**	$ **split -l 3 file F**	$ **split -l 3 -d file F**
Unix	$ **cat xaa**	$ **cat xaa**	$ **cat Faa**	$ **cat F00**
Linux	Unix	Unix	Unix	Unix
AIX	Linux	Linux	Linux	Linux
Solaris	AIX	AIX	AIX	AIX
HPUX	Solaris	$ **cat xab**	$ **cat Fab**	$ **cat F01**
Ubuntu	HPUX	Solaris	Solaris	Solaris
Cygwin	Ubuntu	HPUX	HPUX	HPUX
	Cygwin	Ubuntu	Ubuntu	Ubuntu
		$ **cat xac**	$ **cat Fac**	$ **cat F02**
		Cygwin	Cygwin	Cygwin

Persistent

Quiz

- By specifying option r to sort command, all records in the file are sorted in _____ alphabetical order.

- You can combine the _____ and _____ command to print lines between the file.

- To print the length of longest line in a file using wc command, use option _____.

- The command cut –d':' –f2 filename, splits line by _____ character.

- To specify suffix in split command, use option ____.

Persistent

Quiz Solution

- By specifying option r to sort command, all records in the file are sorted in **descending** alphabetical order.

- You can combine the **head** and **tail** command to print lines between the file.

- To print the length of longest line in a file using wc command, use option **–L.**

- The command cut –d':' –f2 filename, splits line by **colon(:)** character.

- To specify suffix in split command, use option **–a.**

Persistent

Summary

- In this module, we have learnt about filters used in shell.

- Now, you should be able to answer following questions:
    - What are filters?
    - How to use sort and uniq?
    - What is usage of head and tail filters?
    - How to split lines in a file using cut and split?

Persistent

Reference material

- http://bash.cyberciti.biz/guide/Filters

- http://www.tldp.org/LDP/abs/html/textproc.html

- http://www.tuxradar.com/content/exploring-filters-and-pipes

Key contacts

- **Persistent Interactive**

  Bhimashankar Gavkare

  bhimashankar_gavkare@persistent.com

Thank You !!!

Persistent University