TATA CONSULTANCY SERVICES

Experience certainty.

Confidentiality Statement

Tata Code of Conduct

We, in our dealings, are self-regulated by a Code of Conduct as enshrined in the Tata Code of Conduct. We request your support in helping us adhere to the Code in letter and spirit. We request that any violation or potential violation of the Code by any person be promptly brought to the notice of the Local Ethics Counselor or the Principal Ethics Counselor or the CEO of TCS. All communication received in this regard will be treated and kept as confidential.

TATA CONSULTANCY SERVICES

TATA CONSULTANCY SERVICES
Experience certainty.

vILP – Level 2

Content Manual

Version 1.1

December 2014

(ILP Guwahati)

Table of Content

1. Filters

When a program takes its input from another program, performs some operation on that input, and writes the result to the standard output, it is referred to as a filter. A common use of filters is to modify output. UNIX filters can restructure output.

Some common filters in UNIX are:

- sort – sorts files by line (lexically or numerically)
- uniq – Removes identical adjacent lines

1.5 sort

sort reorders the lines of a file in ascending or descending order.

Syntax:

The default order is ASCII: whitespace, numbers, uppercase, and finally, lowercase letters.

Options:

-k n sort on the nth field of the line
-tchar use char as the field delimiter
-n sort numerically
-r reverse order sort
-u removes repeated lines
-m list merge sorted files in list

Examples:

Below examples will help you to understand sort used with different options:

Consider a file named list has below data

Example:

```
1, Justin Timberlake, Title 545, Price $7.30
2, Lady Gaga, Title 118, Price $7.30
```

```
3, Johnny Cash, Title 482, Price $6.50
4, Elvis Presley, Title 335, Price $7.30
5, John Lennon, Title 271, Price $7.90
```

Sort on the 2nd field of file named "list". File list is comma separated file.

```
$sort –t ,-k 2 list
```

Output:

```
4, Elvis Presley, Title 335, Price $6.30
5, John Lennon, Title 271, Price $7.90
3, Johnny Cash, Title 482, Price $6.50
1, Justin Timberlake, Title 545, Price $6.30
2, Lady Gaga, Title 118, Price $6.30
```

Example: Numerically sorting:

sort –n list, numerically sort file named list

```
$cat list
20
19
5
49
200
```

Output:

```
$sort list
19
20
200
49
5
```

However the below command with the help of option –n will output

```
$sort –n list
5
19
20
```

```
200
```

Example:

```
$cat list1
25
18
5
48
200
```

Sort can sort multiple files also.

```
$sort -n list list1 will produce
5
5
18
19
20
25
48
200
200
```

Example: Numerically sort in reverse order

```
$sort -nr list
200
48
25
20
19
18
 5
```

Syntax:

```
$sort –u list
```

Example: Sort the file list removing the repeated lines.

```
$cat list
Unix
Linux
Solaris
AIX
Linux
HPUX
```

Output: The duplicate 'Linux' record got removed.

```
$sort –u list
 Unix
 Linux
 Solaris
 AIX
 HPUX
```

Sort file1 & file2 and merge

```
$sort –m file1 file2
```

Video Link:

http://www.youtube.com/watch?v=ZLjo_yL6g5s

1.6 uniq

uniq command is used to suppress the duplicate lines from a file. It discards all the successive identical lines except one from the input and writes the output.

Syntax:

uniq [option] filename

Options:

-u lists only the lines that are unique
-d lists only the lines that are duplicates
-c counts the frequency of occurrences

Consider following example.txt file in your unix operating system.

```
$cat example.txt
Unix operating system
unix operating system
unix dedicated server
linux dedicated server
```

Suppress duplicate lines:

The default behaviour of the uniq command is to suppress the duplicate line. Note that, you have to pass sorted input to the uniq, as it compares only successive lines.

```
$uniq example.txt
unix operating system
unix dedicated server
linux dedicated server
```

If the lines in the file are not in sorted order, then use the sort command and then pipe the output to the uniq command.

```
$sort example.txt | uniq
```

Count of lines:

The -c option is used to find how many times each line occurs in the file. It prefixes each line with the count.

```
$uniq -c example.txt
2 unix operating system
1 unix dedicated server
1 linux dedicated server
```

Display only duplicate lines:

You can print only the lines that occur more than once in a file using the -d option. The -D option prints all the duplicate lines.

```
$uniq -d example.txt
unix operating system

$uniq -D example.txt
unix operating system
unix operating system
```

Skip first N fields in comparison

the -f option is used to skip the first N columns in comparison. Here the fields are delimited by the space character.

```
$uniq -f2 example.txt
unix operating system
unix dedicated server
```

In the above example the uniq command, just compares the last fields. For the first two lines, the last field contains the string "system". Uniq prints the first line and skips the second. Similarly it prints the third line and skips the fourth line.

Print only unique lines

You can skip the duplicate lines and print only unique lines using the -u option

```
$uniq -u example.txt
unix dedicated server
linux dedicated server
```