Chapter 2e. basic Unix tools

This chapter introduces commands to **find** or **locate** files and to **compress** files, together with other common tools that were not discussed before. While the tools discussed hereare technically not considered **filters**, they can be used in **pipes**.

2e.1. find

The **find** command can be very useful at the start of a pipe to search for files. Here are some examples. You might want to add **2>/dev/null** to the command lines to avoid cluttering your screen with error messages.

Find all files in **/etc** and put the list in etcfiles.txt

```
find /etc > etcfiles.txt
```

Find all files of the entire system and put the list in allfiles.txt

```
find / > allfiles.txt
```

Find files that end in .conf in the current directory (and all subdirs).

```
find . -name "*.conf"
```

Find files of type file (not directory, pipe or etc.) that end in .conf.

```
find . -type f -name "*.conf"
```

Find files of type directory that end in .bak .

```
find /data -type d -name "*.bak"
```

Find files that are newer than file42.txt

```
find . -newer file42.txt
```

Find can also execute another command on every file found. This example will look for *.odf files and copy them to /backup/.

```
find /data -name "*.odf" -exec cp {} /backup/ \;
```

Find can also execute, after your confirmation, another command on every file found. This example will remove *.odf files if you approve of it for every file found.

```
find /data -name "*.odf" -ok rm {} \;
```

2e.2. locate

The **locate** tool is very different from **find** in that it uses an index to locate files. This is a lot faster than traversing all the directories, but it also means that it is always outdated. If the index does not exist yet, then you have to create it (as root on Red Hat Enterprise Linux) with the **updatedb** command.

```
[clim@sop ~]$ locate Samba

warning: locate: could not open database: /var/lib/slocate/slocate.db:...
warning: You need to run the 'updatedb' command (as root) to create th...
Please have a look at /etc/updatedb.conf to enable the daily cron job.
[clim@sop ~]$ updatedb

fatal error: updatedb: You are not authorized to create a default sloc...
[clim@sop ~]$ su -

Password:
```

Most Linux distributions will schedule the **updatedb** to run once every day.

2e.3. date

The **date** command can display the date, time, time zone and more.

```
clim@sop ~$ date

Sat Apr 17 12:44:30 CEST 2010
```

A date string can be customised to display the format of your choice. Check the man page for more options.

```
clim@sop ~$ date +'%A %d-%m-%Y'
Saturday 17-04-2010
```

Time on any Unix is calculated in number of seconds since 1969 (the first second being the first second of the first of January 1970). Use **date +%s** to display Unix time in seconds.

```
clim@sop ~$ date +%s
1271501080
```

When will this seconds counter reach two thousand million ?

```
clim@sop ~$ date -d '1970-01-01 + 2000000000 seconds'
Wed May 18 04:33:20 CEST 2033
```

2e.4. cal

The **cal** command displays the current month, with the current day highlighted.

```
clim@sop ~$ cal
     April 2010
Su Mo Tu We Th Fr Sa
             1  2  3
 4  5  6  7  8  9 10
11 12 13 14 15 16 17
18 19 20 21 22 23 24
25 26 27 28 29 30
```

You can select any month in the past or the future.

```
clim@sop ~$ cal 2 1970
   February 1970
Su Mo Tu We Th Fr Sa
 1  2  3  4  5  6  7
 8  9 10 11 12 13 14
15 16 17 18 19 20 21
22 23 24 25 26 27 28
```

2e.5. sleep

The **sleep** command is sometimes used in scripts to wait a number of seconds. This example shows a five second **sleep**.

```
clim@sop ~$ sleep 5
clim@sop ~$
```

2e.6. time

The **time** command can display how long it takes to execute a command. The **date** command takes only a little time.

```
clim@sop ~$ time date
Sat Apr 17 13:08:27 CEST 2010


real    0m0.014s
user
sys     0m0.008s

        0m0.006s
```

The **sleep 5** command takes five **real** seconds to execute, but consumes little **cpu time**.

```
clim@sop ~$ time sleep 5

real    0m5.018s
user
sys     0m0.005s
```

This **bzip2** command compresses a file and uses a lot of **cpu time**.

```
clim@sop ~$ time bzip2 text.txt

real    0m2.368s
user
sys     0m0.847s
```

2e.7. gzip - gunzip

Users never have enough disk space, so compression comes in handy. The **gzip** command can make files take up less space.

```
clim@sop ~$ ls -lh text.txt

-rw-rw-r-- 1 clim clim 6.4M Apr 17 13:11 text.txt
clim@sop ~$ gzip text.txt

clim@sop ~$ ls -lh text.txt.gz
```

You can get the original back with **gunzip**.

```
clim@sop ~$ gunzip text.txt.gz
clim@sop ~$ ls -lh text.txt

-rw-rw-r-- 1 clim clim 6.4M Apr 17 13:11 text.txt
```

2e.8. zcat - zmore

Text files that are compressed with **gzip** can be viewed with **zcat** and **zmore**.

```
clim@sop ~$ head -4 text.txt

/

/opt

/opt/VBoxGuestAdditions-3.1.6

/opt/VBoxGuestAdditions-3.1.6/routines.sh
clim@sop ~$ gzip text.txt

clim@sop ~$ zcat text.txt.gz | head -4
```

2e.9. bzip2 - bunzip2

Files can also be compressed with **bzip2** which takes a little more time than **gzip**, but compresses better.

```
clim@sop ~$ bzip2 text.txt
clim@sop ~$ ls -lh text.txt.bz2
-rw-r--r-- 1 clim clim 569K Apr 17 10:11 text.txt.bz2
```

Files can be uncompressed again with **bunzip2**.

```
clim@sop ~$ bunzip2 text.txt.bz2
clim@sop ~$ ls -lh text.txt
-rw-r--r-- 1 clim clim 6.4M Apr 17 10:11 text.txt
```

2e.10. bzcat - bzmore

And in the same way **bzcat** and **bzmore** can display files compressed with **bzip2**.

```
clim@sop ~$ bzip2 text.txt

clim@sop ~$ bzcat text.txt.bz2 | head -4

/

/opt
```

2e.11. practice: basic Unix tools

2e.1.1 Explain the difference between these two commands. This question is very important. If you don't know the answer, then look back at the **shell** chapter.

```
find /data -name "*.txt"
```

```
find /data -name *.txt
```

2e.2.1 Explain the difference between these two statements. Will they both work when there are 200 **.odf** files in **/data** ? How about when there are 2 million .odf files ?

```
find /data -name "*.odf" > data_odf.txt
```

```
find /data/*.odf > data_odf.txt
```

2e.3.1 Write a find command that finds all files created after January 30th 2010.

2e.4.1 Write a find command that finds all *.odf files created in September 2009.

2e.5.1 Count the number of *.conf files in /etc and all its subdirs.

2e.6.1 Here are two commands that do the same thing: copy *.odf files to /backup/ . What would be a reason to replace the first command with the second ? Again, this is an important question.

```
cp -r /data/*.odf /backup/
```

```
find /data -name "*.odf" -exec cp {} /backup/ \;
```

2e.7.1 Create a file called **loctest.txt**. Can you find this file with **locate** ? Why not ? How do you make locate find this file ?

2e.8.1 Use find and -exec to rename all .htm files to .html.

2e.9.1 Issue the **date** command. Now display the date in YYYY/MM/DD format.

2e.10.1 Issue the **cal** command. Display a calendar of 1582 and 1752. Notice anything special ?