UNIX AND LINIX
IN INFOCOMMUNICATION
Week 7

O. Sadov

UNIX/Linux administration

By installing the system on your computer, you become more or less an administrator and you need to have some basic administration skills. The most important tasks:

- Users and groups management;
- Working with repositories and packages;
- Devices and drivers handling;
- File systems configuring;
- Archiving and backups;
- Network administration.

Typically, system administration in different UNIX-like systems is the most different part of the system, although the general approaches to administration are more or less the same everywhere. On some systems, you have tools that can help you perform some of the *adminisconsolehelpertrative* tasks. For example:

- `gnome-control-center` in systems with GNOME UI
- RHEL: simple text config — `setup`, GUI-configs — `system-config-*`
- commercial systems provide their own more or less administrator-friendly tools

As we understand it, we need superuser rights to perform such tasks. Some systems may require stricter restrictions where system administration tasks can be decoupled from those of a security officer using mandatory access control (MAC) systems, such as those developed by the National Security Agency (NSA) SELinux subsystem in the Linux kernel.

Let's take a look at the RH '`setup`' tool:

```
$ setup
You are attempting to run "setup" which requires administrative
privileges, but more information is needed in order to do so.
Authenticating as "root"
Password:
```

We have to enter the root password and after that we can do some settings:

- Authentication configuration
- Keyboard configuration
- System services

But when we run '`system-config-date`', the system asks for the user's password. This is because these programs use different machinery for increasing privileges:

```
$ ls -l /usr/bin/setup
lrwxrwxrwx. 1 root root 13 Nov 9 2019 /usr/bin/setup
                                                -> consolehelper
```

The `setup` program is just a symbolic link to '`consolehelper`', a tool that allows console users to easily run system programs. And the `pkexec` runner is used to execute '`system-config-date`':

```
$ cat /usr/bin/system-config-date
#!/bin/sh

exec /usr/bin/pkexec \
           /usr/share/system-config-date/system-config-date.py
```

A more general way is to just switch to the 'root' superuser, and the first way to do this is with the `su` command:

```
man su
```

`su` — run a command with substitute user and group ID, by default — to 'root' superuser. For such a switch, we need to say the password of this user.

When called without arguments ‘su’ defaults to running an interactive shell as ‘root’. A very important option is just a ‘dash’, it’s mean — starts the shell as login shell with an environment similar to a real login.

After switching to superuser “root” your prompt will change from a dollar sign to a hash sign:

```
$ id
...
$ su -
Password:
# id
...
# logname
...
```

On BSD systems, for security reasons, only users in the ‘wheel’ group (group 0) can use ‘su’ as ‘root’, even with the ‘root’ password. In many UNIXes and Linux the Plugin Authentication Module (PAM) is now being used to fine tune the privilege change. The settings for this subsystem are located in the /etc/pam.d/ directory.

```
$ ls /etc/pam.d/
```

And one of the applications whose config files we can find in this directory is the ‘sudo’ command. The default PAM security policy allows users configured appropriately in ‘/etc/sudoers’ to run commands with ‘root privileges. And you don’t need to know the password of ‘root’ user to do this.

Also, by default only one command is executed with ‘sudo’, instead of ‘su’ where we have to use the ‘-c’ option to run one command. This reduces the chances of an unexpected error for an inexperienced user. And this is, for example, the default policy for Ubuntu systems. When Ubuntu is installed, a standard root account is created, but no password is assigned to it. You cannot log in as root until you assign a password for the root account. Only ‘sudo’ may be used with such default settings.

To allow a regular user to run ‘sudo’ this way on RH based systems such as Fedora, RHEL, CentOS, our NauLinux, you must add this user to the ‘wheel’ group (as in BSD). And the easiest way to get a ‘root’ shell session like in ‘su’ with ‘sudo’ in Ubuntu is to just run it ‘sudo -i’ (interactive).

Users and groups

As we remember, users are one of the three pillars on which the UNIX world stands.

You can use some graphical interfaces to manage users and groups, but simple CLI utilities are often more convenient. There are:

- `adduser`, `useradd` — create a new user or update default new user information
- `groupadd` — create a new group
- `passwd` — update user's authentication tokens

To create a new user, we (as 'root') simply have to run the program '`adduser`' and set a password with '`passwd`'. But it's not over yet!

Actually, adduser is also black magic — in fact, all data related to users and groups is placed in common text files that can only be modified with ordinary text editors:

```
less /etc/passwd
```

The file format is quite simple — one line per user with colon-separated fields:

```
$ man 5 passwd
```

The fields, in order from left to right, are:

1. User name: the string a user would type in when logging into the operating system: the logname. Must be unique across users listed in the file.
2. Information used to validate a user's password; And at the very beginning, the password data was actually placed in this field. But we can read this file as a regular user, this is a design requirement. Did users have the ability to read passwords of other users at this time? Not. In Robert Morris and Ken Thompson's classic article "Password Security: A Case History" about the UNIX password system, Morris described a real-life incident he himself saw:

> Perhaps the most memorable such occasion occurred in the early 1960s when a system administrator on the CTSS system at MIT was editing the password file and another system administrator was editing the daily message that is printed on everyone's terminal on login. Due to a software design error, the temporary editor files of the two users were interchanged and thus, for a time, the password file was printed on every terminal when it was logged in.

And the main idea of UNIX passwords is not to believe that you can simply hide them. Better not to save passwords in the system at all. Actually, when creating a password, a random code was simply generated (the so-called SALT code), and then from this code and password by means of a one-way 'crypt' procedure with the DES algorithm:

```
man crypt
```

And the result of this operation cannot be decrypted (actually, we received some kind of hash) — when entering the system, the system receives SALT from the password field, encrypts it with the entered password, and simply compares it with the contents of the password field.

In most modern uses, this field is usually set to "x" (or "*", or some other indicator) with the actual password information being stored in a separate shadow password file. On Linux systems, setting this field to an asterisk ("*") is a common way to disable direct logins to an account while still preserving its name, while another possible value is "*NP*" which indicates to use an NIS server to obtain the password.[2] Without password shadowing in effect, this field would typically contain a cryptographic hash of the user's password (in combination with a salt).

3. User identifier number, used by the operating system for internal purposes. It need not be unique. Moreover, a superuser is simply a user with a zero user ID, and you can have multiple superusers in addition to the traditional "root" superuser. For example, you can create some superuser with UID 0 and a name like "halt" and with the command "shutdown" as a shell for the user, and provide a password for that user to anyone who needs to shutdown the system at night.

4. Group identifier number, which identifies the primary group of the user; all files that are created by this user may initially be accessible to this group. You can change this default during the current session with the command ‘`newgrp`’.

5. Gecos field, commentary that describes the person or account. Some early Unix systems at Bell Labs used GE/Honeywell mainframe computers with General Comprehensive Operating System (GCOS) for print spooling and various other services, so this field was added to carry information on a user’s GECOS identity.

   Typically, now this is a set of comma-separated values including the user’s full name and contact details which may be used by some commands for example by mail user agent.

6. Path to the user’s home directory.

7. Program that is started every time the user logs into the system. For an interactive user, this is usually one of the system’s command line interpreters (shells). For example, for pseudo-users who do not need interactive sessions, this could be ‘`nologin`’ or just ‘`false`’ executables, which will exit immediately upon startup.

The description of the groups is also placed in a text file:

```
less /etc/group
```

In this file, we see a similar format:

```
man 5 group
```

1. group_name — the name of the group.

2. password — Password field that has never been used

3. GID — the numeric group ID.

4. user_list — a list of the usernames that are members of this group, separated by commas.

Finally, a file with real data regarding passwords:

```
ls -l /etc/shadow
```

As we can see, only the superuser has access to this file. The transfer of password hashes from '/etc/passwd' to this file was carried out to prevent brute-force attacks using modern computing equipment, which is now becoming cheaper and cheaper. And we can see the hashes of the passwords in the second field of the records for each user:

```
man 5 shadow
```

A password field which starts with an exclamation mark means that the password is locked. The rest of the characters in the string represent the password field before the password was locked, and you can simply remove the exclamation mark to unlock it.

On a multiuser system with many administrators, it is advisable to use the 'vipw' and 'vigr' commands to avoid conflicts when multiple administrators are editing the same file at the same time:

```
man vipw
```

This file-based machinery for handling of user accounts is not hardcoded. You can switch to network authentication services such as LDAP or Winbind using the setup utility:

```
setup
```

or simply by editing the text configuration file '/etc/nsswitch.conf'

```
less /etc/nsswitch.conf
```

Other security related settings on Linux systems can be done in the '/etc/security' and PAM configuration directories:

```
ls /etc/security/
ls /etc/pam.d/
```

As we can see, the UNIX system administration paradigm does not hide the details from the user, everything can be configured by hands or scripts. For beginners, such systems simply provide more or less user-firendly tools and wizards to lower the barrier to entry.

Partitions

As storage systems grow, they need to be separated to store different data. And for this the partitioning was invented. There are different partition schemes developed by different vendors like IBM, Apple, Microsoft, etc.

In common PCs the Master Boot Record (MBR) partitioning scheme, widely used since the early 1980s, imposed limitations for use of modern hardware. A major deficiency is the limited size of 32 bits for block addresses and related information. For hard disks with 512-byte sectors, the MBR partition table entries allow a maximum size of 2 TiB. Also, the standard partitioning scheme only supports 4 primary partitions, and as the disk space increases, it will become necessary to implement such complex solutions as extended and logical partitions.

In the late 1990s, Intel developed a new partition table format as part of what eventually became the Unified Extensible Firmware Interface (UEFI). As of 2010, the GUID Partition Table (GPT) forms a subset of the UEFI specification. GPT uses 64 bits for logical block addresses, allowing a ZB disk size. Number of partitions — Depends on the space allocated for the partition table. By default, the GPT contains space to define 128 partitions.

Different systems use different naming schemes for devices and partitions. Modern Linux, for example, has special `/dev/sd` files for SCSI or SATA devices with a naming schema like this:

```
ls /dev/sd*
/dev/sda /dev/sda2 /dev/sda4 /dev/sda6 /dev/sdb /dev/sdb2
/dev/sda1 /dev/sda3 /dev/sda5 /dev/sda7 /dev/sdb1 /dev/sdb5
```

The letter “a” stands for the first device on the bus, and the numbers are the partitions. We can also access disk devices by disk labels:

```
ls -l /dev/disk/by-label
```

BSD disklabels, which also used on many commercial UNIXes, traditionally contain 8 entries for describing partitions. These are, by convention, labeled alphabetically, ’a’ through to ’h’. Some BSD variants have since increased this to 16 partitions, labeled ’a’ through to ’p’.

Also by convention, partitions ’a’, ’b’, and ’c’ have fixed meanings:

- ‘a’ is the “root” partition, the volume from which the operating system is bootstrapped. The boot code in the Volume Boot Record containing the disklabel is thus simplified, as it need only look in one fixed location to find the location of the boot volume;
- ‘b’ is the “swap” partition;
- ‘c’ overlaps all of the other partitions and describes the entire disk. Its start and length are fixed. On systems where the disklabel co-exists with another partitioning scheme (such as on PC hardware), partition ’c’ may actually only extend to an area of disk allocated to the BSD operating system, and partition ’d’ is used to cover the whole physical disk.

On Linux, MBR related tools are:

- ‘fdisk’ — is a simple text-based tool:

```
# fdisk /dev/sda
Command (m for help): m
Command action
   d delete a partition
   g create a new empty GPT partition table
   G create an IRIX (SGI) partition table
   l list known partition types
   m print this menu
   n add a new partition
   o create a new empty DOS partition table
   p print the partition table
   q quit without saving changes
   s create a new empty Sun disklabel
   t change a partition’s system id
   v verify the partition table
   w write table to disk and exit
   x extra functionality (experts only)

Command (m for help): p
...
```

  The most useful operations: m, p, n, t, d, q, w.

- ‘cfdisk’ — is a fullscreen program text-based variant of ‘fdisk’.
- ‘sfdisk’ — non-interactive variant of ‘fdisk‘, it’s useful for scripting.

Partition management programs that support GPT:

- ‘parted’ — GNU Parted.

```
# parted /dev/sda
GNU Parted 3.1
Using /dev/sda
Welcome to GNU Parted! Type ’help’ to view a list of commands.
(parted) help
  align-check TYPE N check partition N for TYPE(min|opt) alignment
  help [COMMAND] print general help, or help on COMMAND
  mklabel,mktable LABEL-TYPE create a new disklabel (partition table)
  mkpart PART-TYPE [FS-TYPE] START END make a partition
  name NUMBER NAME name partition NUMBER as NAME
  print [devices|free|list,all|NUMBER] display the partition table, available
        partitions, or a particular partition
  quit exit program
  rescue START END rescue a lost partition near START and END

  resizepart NUMBER END resize partition NUMBER
  rm NUMBER delete partition NUMBER
  select DEVICE choose the device to edit
  disk_set FLAG STATE change the FLAG on selected device
  disk_toggle [FLAG] toggle the state of FLAG on selected device
  set NUMBER FLAG STATE change the FLAG on partition NUMBER
  toggle [NUMBER [FLAG]] toggle the state of FLAG on partition NUMBER
  unit UNIT set the default unit to UNIT
  version display the version number and copyright information of GNU Parted
(parted)
```

- ‘gparted’ — GUI variant of ‘parted’.

File systems

And now that our partitioning is complete, it's time to see how we can use our disk space. As mentioned earlier, UNIX-like systems can treat disks or disk partitions as files. And some databases, for example, can use raw disk partitions to store data with higher performance.

But most of the time, disk space is used in file systems. UNIX-like systems and especially Linux support many different file systems. All the details of their implementation are reduced to a common denominator — the VFS abstraction. Then we can mount them in a single directory tree, navigate through the hierarchy, read, write, work with owners and permissions, according to the restrictions imposed by the original file systems.

The standard tool for creating a new filesystem is '`mkfs`':

```
man mkfs
```

We can choose the type of filesystem and some parameters to create. But it's actually just a wrapper around the real mkfs tools for different types of filesystems:

```
# ls /sbin/mkfs*
/sbin/mkfs /sbin/mkfs.cramfs /sbin/mkfs.ext4 /sbin/mkfs.xfs
/sbin/mkfs.btrfs /sbin/mkfs.ext3 /sbin/mkfs.vfat
```

and they all support their own set of options:

```
# man mkfs.ext4
```

The most commonly used file systems in Linux right now are:

- EXT4 is the Linux journaling file system, or the Fourth Extended File System, which is the successor to the extended file system line originally created in 1992 by Rémy Card to overcome certain limitations of the MINIX file system. The ext4 filesystem can support files up to 16TB and volumes with sizes up to 1 exbibyte (EiB), but this may be limited for certain system versions. For example for RHEL 7/8 — 50TB.

- XFS is a high performance 64-bit journaling file system created by Silicon Graphics, Inc (SGI) in 1993 for their UNIX called IRIX. Although

> XFS scales to exabytes, the host operating system limits can reduce this limit. For example — 500 TB for the maximum file size and file system size for RHEL7 and 1PB/8EB for RHEL8.

Typically ext4 provides better performance for small filesystems on machines with limited I/O capabilities, while XFS provides better performance for large filesystems on machines with high-performance parallel I/O. Also in XFS it is more difficult to reduce the size of the filesystem.

With the '`mkfs`' options, you can set various parameters for creating the filesystem, for example, optimize it for storing large files or for more smaller files.

Once the filesystem is created, we can "`mount`" it. In most cases, this happens automatically when you insert a flash drive or SD card into your computer. But in some cases it needs to be done manually, and you can do it by running the '`mount`' command.

```
man mount
```

You just need to specify the device and directory — mount point, and after mounting you will see the contents of the file system from this device or pseudo device in this directory. And also you can choose the "`mount`" options. For example, we can mount the ISO image with the '`loop`' option:

```
ls /mnt
mount -o loop ...iso /mnt
ls /mnt
```

And then we can '`unmount`' it:

```
umount /mnt
ls /mnt
```

by setting the device or the mount directory as an argument:

```
man umount
```

But Linux/UNIX will not allow you to unmount a device that is busy. There are many reasons for this (such as program accessing partition or open file), but the most important one is to prevent the data loss. You can use the 'fuser' and 'lsof' commands to find the processes that are loading your filesystems:

```
man fuser
...
-k, --kill
...
```

Finally, we can check our filesystem. For journaled file systems, recovering from a power outage is not as relevant, but in some cases it may be useful. In a difficult situation, such as a damaged hard disk, during system boot, you may receive an error message recommending that you run the 'fsck' command to check the file system:

```
man fsck - check and repair a Linux file system
```

As we can see, we have many options for the 'fsck' command, but the main one is 'y' — 'yes'. This means — always try to fix any file system corruption you find automatically, otherwise you could get a zillion troubleshooting questions during the fixup.

After completing the repair, you can find some lost data in a special directory "/lost+found" in the root of the damaged filesystem, which consists of many directories and files whose names contain only numbers — so called 'i-nodes'.

And then you can rename them manually — for example, you found some directory with the files:

- 'passwd', 'group' and 'shadow' — this means that this is '/etc'
- or 'sh', 'ls' and 'cp' — this means '/bin'

and so on...

Swapping

And finally, a few words about swapping. Swapping or paging is a memory management scheme by which a computer stores and retrieves data from secondary storage for use in main memory. It is an MMU-driven virtual memory mechanism that is used in modern operating systems to use secondary storage in order for programs to exceed the amount of available physical memory.

Under the Hood — Virtual Memory

This means you can run applications with a total memory usage that exceeds the physical RAM on your system. The scheduler sends inactive processes to disk swap and loads active tasks from disk into memory. This can reduce overall system performance, but it will increase the ability to run applications.

The main program for creating of swapping area is ‘mkswap‘:

```
man mkswap
```

You just need to specify the device as an argument. Or a file if you need temporary swap space like Microsoft does on Windows.

Then you can enable swap space with the command ‘`swapon`’:

```
man swapon
```

After that, you will see additional swap space using the ‘`free`’ command or in the pseudo file ‘`/proc/meminfo`’. You can also turn off the swap area with the ‘`swapoff`’ command.

OK. But all these mounts and swaps will be connected to our system only until the reboot. To automatically mount them at boot time, we must write them to the filesystems table in the file ‘`/etc/fstab`’:

```
cat /etc/fstab
```

In this text file, we can place static information about connecting file systems and enabling swapping areas:

```
man fstab
```

Each filesystem is described on a separate line; fields on each line are separated by tabs or spaces.

- The first field (fs_spec) describes the block special device or remote filesystem to be mounted.
- The second field (fs_file) describes the mount point for the filesystem. For swap partitions, this field should be specified as ‘none’.

- The third field (fs_vfstype) describes the type of the filesystem. An entry 'swap' denotes a file or partition to be used for swapping.
- The fourth field (fs_mntops) describes the mount options associated with the filesystem.
- The fifth field (fs_freq) is used for these filesystems by the dump command to determine which filesystems need to be backuped.
- The sixth field (fs_passno) is used by the fsck(8) program to determine the order in which filesystem checks are done at reboot time.

After putting some entry in the fstab file, you can run the '`mount`' command with only one of them: device or mount point.

Disk space

Another important task with data in your file system is archiving and backing up. It's wise to look into your filesystems first to analyze disk usage. In case your file system is full on many systems, some graphical disk analysis program will run and you can detect problems visually. But we can also do this job using command line tools that can help you automate some of the admin tasks.

The main tool for reporting file system disk space usage is the '`df`' utility:

```
man df
```

The most useful option is "`-h, -human-readable`" for human readable print sizes in kilobytes, megabytes, gigabytes.

For a more accurate analysis, you can use the '`du`' utility to estimate the file space usage of directories and files:

```
man du
```

So, we can get the size of some directory:

```
# du -hs /tmp/
136K /tmp/
```

And the most commonly used options are “`-k`”, which displays sizes in kilobytes, and “`-x`”, which means that it will scan only this file system and skip directories on other file systems. Let’s take a look at an example of using the command line tools to find the largest directories and files:

```
$ du -kx /tmp | sort -rn | less
```

We examine the directory ‘`/tmp`’, perform a numeric sorting of the sizes of directories and files, and redirect the result to the viewer ‘`less`’ for analysis.

And after finding the largest files and directories, we can clean up our file system and before this archive and back up some data. The easiest way is to simply copy using the ‘`cp -a`’ command to some external drive, or using ‘`scp -rC`’ or ‘`rsync -avz`’ to a remote host.

Also, using the ‘`cp`’ or ‘`scp`’ commands, you can copy any partition or the entire disk, because for us they are just files. But a more efficient way to do this is with the ‘`dd`’ command:

```
man dd
```

By default, it just copies stdin to stdout, perhaps with some re-coding. But the most interesting options for us are: ‘`if`’, ‘`of`’, ‘`bs`’, ‘`count`’, ‘`seek`’ and ‘`skip`’. With a combination of these options, we can select the input and output files, choose the block size to increase speed of I&O, the number of blocks we want to copy, and seek/skip on output/input. Thus, we can cut and paste any fragment from one device or file to another.

We can also use the ‘`od`’ command to low-level view of a file or device in different formats:

```
man od
```

For example — to our hard drive:

```
# od -bc /dev/sda1 | less
```