**Laboratory work №5**

**Topic: "Linux commands for archiving and compressing data".**

Performed by students RPZ-93B group

Team: Бушовська О.В, Білобровенко О.С., Скворцов Д.Є.

The teacher checked: Сушанова В.С.

**The goal of the work:**

1. Gaining practical skills in working with the Bash command shell.

2. Familiarity with basic commands for archiving and compressing data.

**Material support of classes**

1. Computer type IBM PC.

2. Windows family of operating systems (Windows 7).

3. Virtual machine - Virtual Box (Oracle).

4. GNU / Linux operating system - CentOS.

5. Cisco Network Academy website netacad.com and its online Linux courses

***Готувала матеріал студентка Білобровенко Олександра***

**Tasks for preliminary preparation.**

1. Read brief theoretical information for laboratory work and make a small dictionary

basic English terms for assigning teams and their parameters.

|  |  |
| --- | --- |
| Англійська | Українська |
| Lossy and Lossless Compression | Стиснення з втратами та без втрат |
| Disk space | Дисковий простір |
| Archiving | Архівування |
| Unpacking | Розархівування |

2. On the basis of the considered material give answers to the following questions:

2.1. What is the purpose of the commands tar, xz, zip, bzip, gzip? Make a brief description of each team and highlight it their main parameters. How to install them.

.tar is an archive file without compression. Tar was originally used to archive data on tape devices. But it also allows you to write the output to a file, and this method has become widely used in Linux for its intended purpose. The tar command is the most commonly used command for archiving multiple files in tarball - a common Linux file format similar to the zip format, with compression being optional. This is a rather complex command with a long list of functions, such as adding new files to an existing archive, listing the contents of the archive, removing content from the archive and much more.

.xz is a lossless computer compression program that uses the LZMA2 algorithm and the corresponding file format. The xz archiver could be thought of as a truncated version of 7-Zip, which has its own format, rather than using its 7z format. Although 7-Zip, using LZMA2 compression, is able to compress tightly (due to speed), its archive format, made mainly for Windows, does not support Unix functionality (this format lacks support for metadata Unix-like file systems). xz compresses a single file on the input, rather than fitting a set of files into a single archive. Thus, it compresses a file, which can also be an archive, such as those created by Unix programs such as tar or cpio.

.zip - (usually) a compressed archive file. The gzip utility is ideal for users running Linux. If you need to work with the same file in another OS (such as Windows or MS DOS), zip is more versatile, as it is suitable for file sharing, avoiding compatibility issues. Zip is used for both archiving and compression. After the procedure, the files have a .zip extension.

. bzip According to the tradition of UNIX programming, bzip2 performs only one function: compressing and unpacking a single file. The .bz2 extension is added to the default file name. To package multiple files, they are usually first archived into one file with the tar utility, and then compressed with bzip2.

.gzip When compressing, the names of the files to be compressed are entered as an argument, and options (keys) can also be used. The specified file is replaced with a compressed version with the .gz extension, retaining the owner, modes, access rights, modification time, and labels of the original source file. As a result, gzip produces a lower percentage of compression than its other counterparts (for example, bzip2), using a fairly old algorithm, but is unpretentious to system resources and, as an option, suitable for machines with low power.

2.2. Here are three examples of how to archive and compress data with different commands.

Archiving:

$ Tar -xvf archive.tar.gz

$ ./filename.shar

zip -r directory\_name1

Compression:

$ Tar -zcvf home.tar.gz ~ /

gunzip archive.tar.gz

zip archive\_name.zip filename1

3. Study Cisco Academy Online Course Materials:

- NDG Linux Unhatched (Chapter 19 - 22 all Topics)

- NDG Linux Essentials (Chapter 9 all Topics)

4. Take the NDG Linux Essentials course on the following topics:

- Chapter 09 Exam

- Midterm Exam (Modules 1 - 9)

5. Take the NDG Linux Unhatched course:

- Assessment

- End of Course FeedbackExternal tool

6. Prepare in electronic form the initial version of the report:

- Title page, topic and purpose of the work

- Glossary of terms

- Answers to item 2.1 and item 2.4 of the tasks for preliminary training.

**Progress.**

1. Initial work in CLI mode in Linux Linux family:

1.1. Start the VirtualBox virtual machine, select CentOS, and start it. Log in

under the user: CentOS, password for login: reverse (if you perform LR in 401 aud.) and lower terminal.

1.2. Start the Ubuntu PC virtual machine (if you are performing LR tasks through the netacad academy)

1.3. Start your Linux operating system (if you are running on your own PC and

installed) and start the terminal.

2. Study all the examples of commands presented in the laboratory work of the NDG Linux course Essentials:

- Lab 9: Archiving and Compression

***Готувала матеріал студентка Білобровенко Олександра***

3. Create a table of commands studied in paragraph 2 of the course of work as follows:

|  |  |
| --- | --- |
| Command name | Its purpose and functionality |
| tar | The tar command is used to create archive files, merge multiple files into a single file. By default, it does not compress the data. |
| gzip | The **gzip** command is intended for lossless data compression using the utility of the same name, using the Lempel-Ziv (LZ77) algorithm with Huffman coding. |
| bzip2 | The bzip2 command is for lossless data compression using a suitable utility using the Burrows-Wheeler algorithm |
| zip | The [zip](https://linux.die.net/man/1/zip) command is a command-line tool in Linux that **allows us to create an archive of files and directories**. Besides that, it also provides a multitude of functionalities for manipulating an archive. |
| unzip | unzipping files zip |
| xz | xz is a lossless compression program and file format that includes the LZMA/LZMA2 compression algorithms |
| unxz | XZ-format compression utilities |

4. Familiarize yourself with the tar command and use it to:

- create a file with the extension .tar;

# tar cvf sedicomm-15-12-20.tar /home/OS\_works /

- create a file with the extension .tar, consisting of several files and directories at once;

# tar rvf OS\_works.tar Chapter\_03\_Exam.png

# tar rvf OS\_works.tar Operation\_System\_Folder

- view the contents of the file;

# tar tvf OS\_works.tar.bz2

- extract the contents of the tar file;

tar xvf OS\_works.tar.bz2

or

tar xvf OS\_works.tar.bz2 -C /home/Dmitriy/Documents/

- create a tar archive file compressed with bzip;

# tar cvfj OS\_works.tar.bz2 /home/Dmitriy/Documents

- extract the contents of the tar bzip file;

tar xvf OS\_works.tar.bz2

- create an archive tar file compressed with gzip;

# tar cvfj OS\_works.tar.tb2 /home/Dmitriy/Documents

- extract the contents of the tar gzip file.

To unarchive the compressed tar.gz file, simply use this the same command for the tar.bz2 command.

tar xvf OS\_works.tar.bz2

**Test questions**

1. Give a comparative description of compression and archiving processes.

Archiving means that you take 10 files and combine them into one file, with no difference in size. If you start with 10 100KB files and archive them, the resulting single file is 1000KB. On the other hand, if you compress those 10 files, you might find that the resulting files range from only a few kilobytes to close to the original size of 100KB, depending upon the original file type.

Compression is a process of taking some input data, and by using some sophisticated algorithm, compressing it (transform the bits, effectively), in order to have the same entity that weighs less size.

This is useful if you want to keep more data in a less space (space is always limited resource), or if you just want to have a faster file-transfer throughout networks.

Popular compression utility programs, on Linux distributions, are:

gzip (frequently used);

bzip2 (less frequently used, yet produces smaller output file than gzip);

xz (most space-efficient tool, in Linux, so far)

zip (often used for decompressing data, that was compressed on other systems using zip, like Windows OS).

Note, that generally, more efficient compression method is, more time it takes.

Archiving, on the other hand, can be thought of like putting some different files into one box. If you have 5 files, each of a size of 10kb, archiving those will give you 5 x 10 = 50kb, and that is it.

Note, that on Linux, we have a very good program tar, which, when given an input, does both:

archives the input (first step);

and then compresses that archive.

2. What programs, other than those listed in the paper, can be used for compression and archiving files and directories in Linux? Give examples and a brief description.

I described all possible ways to implement file compression and archiving in the previous task.

***Готувала матеріал студентка Бушовська Ольга***

3. Compare the compression algorithms used in the commands (programs) used in

Linux. Which of the algorithms can be considered the fastest and most efficient?

The gzip command uses the Lempel-Ziv data compression algorithm, while the bzip utilities use a different compression algorithm called Burrows-Wheeler block sorting, which can compress files smaller than gzip at the expense of more CPU time. These files can be recognized because they have a .bz or .bz2 extension instead of a .gz extension.

The xz and unxz tools are functionally similar to gzip and gunzip in that they use the Lempel-Ziv-Markov (LZMA) chain algorithm, which can result in lower decompression CPU times that are on par with gzip while providing the better compression ratios typically associated with the bzip2 tools. Files compressed with the xz command use the .xz extension.

4. Describe the compression and archiving software that can be used in yours

mobile phone.

Program SD Maid

5. Describe and compare software tools for compressing and (de) archiving data in the family OS Windows.

Examples of archivers are WinZIP, WinRAR, 7-zip, Power-Archiver, WinAce, Ark, Btar, AndroZip, FreeArc and others. The result of these programs is an archive file that contains compressed or uncompressed files and folders. Additional measures can be used in the archiving process to protect data from unauthorized access, such as setting a password to access data in the archive.

Depending on the algorithms used to compress and archive data, there are the following archive file formats: ZIP, RAR, TER, ARJ, CAB, LZH, ACE, 7z and others. Most often, especially on the Internet, use archive files in ZIP format.

6. Explain how data compression and archiving can be used to back up data.

The computer has limited memory, so it is extremely important for a programmer or anyone working in the field of computer technology to use this resource efficiently and rationally. When we have data that takes up a certain amount of memory, but we do not need it at the moment, we can compress or archive it, thus freeing up space (memory). The advantage of this option over the deletion option is that we can return this data at any time if necessary, so it turns out that we have reserved it.

In general, the computer OS itself can do regular archiving of information. This is required in case of system failure or complete failure.

***Conclusion:***

During this lab, we learned how to archive and compress files using the command line.