“Київський фаховий коледж зв’язку”

Циклова комісія Комп’ютерної інженерії

**ЗВІТ ПО ВИКОНАННЮ**

**ЛАБОРАТОРНОЇ РОБОТИ №6**

з дисципліни: «Операційні системи»

**Тема: «“Команди Linux для архівування та стиснення даних. Робота з текстом”»**

Виконали:

Cтуденти групи: КСМ-03Б

Команда 2:

Нестолій Н.І

Усенко С. І.

Титов О. О

Перевірив викладач:

Сушанова В.С

Київ 2022

*Робота студентів групи КСМ-03Б Команда 2: Нестолій Н., Усенко С., Титов 0.;*

**Мета роботи:**

1. Отримання практичних навиків роботи з командною оболонкою Bash.
2. Знайомство з базовими командами для архівування та стиснення даних.
3. Знайомство з базовими діями при роботі з текстом у терміналі.

**Матеріальне забезпечення занять**

1. ЕОМ типу IBM PC.

2. ОС сімейства Windows (Windows 7).

3. Віртуальна машина – Virtual Box (Oracle).

4.Операційна система GNU/Linux – CentOS.

5. Сайт мережевої академії Cisco netacad.com та його онлайн курси по Linux

# Завдання для попередньої підготовки:

*Готувала матеріал студентка Усенко С.*

1. **Прочитайте короткі теоретичні відомості до лабораторної роботи та зробіть невеликий словник базових англійських термінів з питань призначення команд та їх параметрів.**

File compression - a data compression method in which the logical size of a file is reduced to save disk space for easier and faster transmission over a network or the Internet.

Lossless compression - a class of data compression that allows the original data to be perfectly reconstructed from the compressed data with no loss of information. Lossless compression is possible because most real-world data exhibits statistical redundancy.

Lossy compression - the class of data compression methods that uses inexact approximations and partial data discarding to represent the content. These techniques are used to reduce data size for storing, handling, and transmitting content.

Archive file - a computer file that is composed of one or more files along with metadata. Archive files are used to collect multiple data files together into a single file for easier portability and storage, or simply to compress files to use less storage space.

Compression tools - a type of utility software. Furthermore, these tools help us to compress and decompress files. An important part of a computer is storage space, it is very important to maintain this storage

**4. На базі розглянутого матеріалу дайте відповіді на наступні питання:**

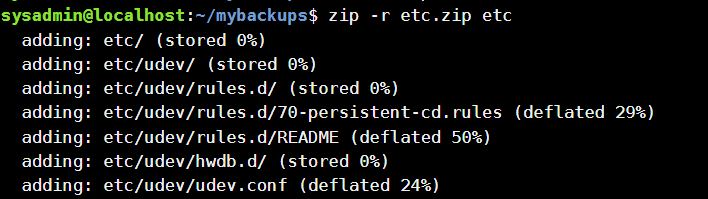
***4.1 Яке призначення команд tar, xz, zip, bzip, gzip? Зробіть короткий опис кожної команди та виділіть їх основні параметри. Яким чином їх можна встановити.***

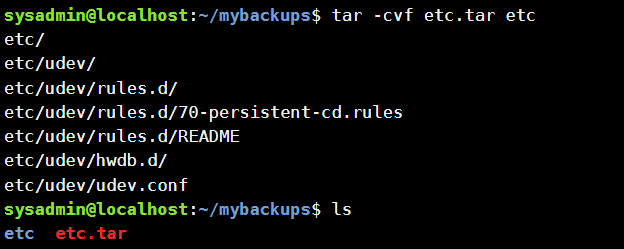
- The Linux ‘tar’ stands for tape archive, is used to create Archive and extract the Archive files. **-c** creates archive, **-x** extracts archive, **-z** zips archive using gzip.

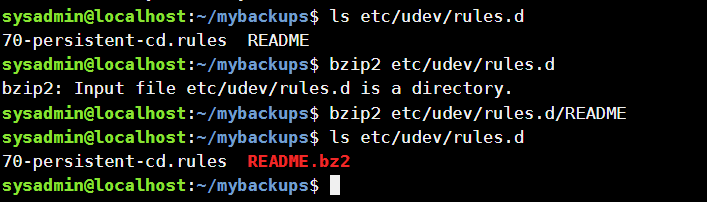
- The xz compression utilities leverage a compression algorithm known as LZMA2. **-l** flag lists statistics about the compression of the file, **-c** flag sends the compressed output to standard out, **-e** flag enables extreme compression mode.  
- The zip command is a command-line tool in Linux that allows us to create an archive of files and directories. **-d** option removes the file from the zip archive, **-u** option updates the file in the zip archive, **-m** option deletes the original files after zipping.

- bzip2 command in Linux is used to compress and decompress, but has a slower decompression time and higher memory use. **-z** option forces compression, **-k** option does compression but does not delete the original file, **-d** option is used for decompression of compressed files, **-t** option does the integrity check of the file and does not decompresses the file.

- Gzip is one of the most popular compression algorithms that allow you to reduce the size of a file and keep the original file mode, ownership, and timestamp. **-k** option saves original file, **-v** option shows the percentage reduction and the names of the files that are being processed, **-r**  option compresses all files in a given directory, from **-1 to -9** changes compression level. **4.2 *Наведіть три приклади реалізації архівування та стискання даних різними командами.***







***4.3. Яке призначення команд cat, less, more, head and tail? Зробіть короткий опис кожної команди та виділіть їх основні параметри. Яким чином їх можна встановити.***- The cat command is a utility command in Linux. One of its most common usages is to print the content of a file onto the standard output stream. Other than that, the cat command also allows us to write some texts into a file. Use **-A** option to show all, **-b** to number nonempty output lines, **-s** to suppress repeated empty output lines.

**-** The less command is a Linux terminal pager that shows a file's contents one screen at a time. **-E** option adds automatic exit upon reaching the end of file, **-f** option forces less to open non-regular files (a directory or a device-special file), **-i** option allows ignoring of any case sensivity.

**-** More command is used to view the text files in the command prompt, displaying one screen at a time in case the file is large (For example log files). The more command also allows the user do scroll up and down through the page. **-d** command displays “[Press space to continue, ‘q’ to quit.]” and displays “[Press ‘h’ for instructions.]” when wrong key is pressed, **-f** option does not wrap the long lines and displays them as such, **-p** option clears the screen and then displays the text.

- Head command is the complementary of Tail command. The head command, as the name implies, print the top N number of data of the given input. **-n (any number)** Prints the fixed amount of lines instead of first 10 lines. By using **-v** option, data from the specified file is always preceded by its file name.

- Tail is the complementary of head command. The tail command, as the name implies, print the last N number of data of the given input. Options are same as Head command ones.

***4.4.*** ***Поясніть принципи роботи командної оболонки з каналами, потоками та фільтрами***

Assume for a moment that you have a modern machine with 8 or more threads. Can you start seeing how if we would be able to execute code – eight parallel threads all at the same time, each running on a different CPU thread (or shared across all threads) – this way it would execute much faster then a single-threaded process running on a single CPU thread (which may be co-shared with other running processes)? The gains realized will depend a bit on what is being executed, but gains there will be, almost always!

***4.5.*** ***Яке призначення команди grep?***

Grep is an essential Linux and Unix command. It is used to search text and strings in a given file. In other words, grep command searches the given file for lines containing a match to the given strings or words. It is one of the most useful commands on Linux and Unix-like system for developers and sysadmins.

**Хід роботи**

*Готував матеріал студент Нестолій Н.*

**1. Початкова робота в CLI-режимі в Linux ОС сімейства Linux:**

***1.1. Запустіть віртуальну машину VirtualBox, оберіть CentOS та запустіть її. Виконайте вхід в систему під користувачем: CentOS, пароль для входу: reverse (якщо виконуєте ЛР у 401 ауд.) та запустіть термінал.***

***1.2. Запустіть віртуальну машину Ubuntu\_PC (якщо виконуєте завдання ЛР через академію netacad)***

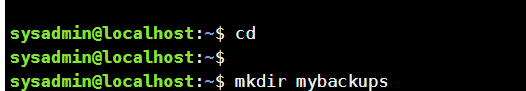
***1.3. Запустіть свою операційну систему сімейства Linux (якщо працюєте на власному ПК та їївстановили) та запустіть термінал.***

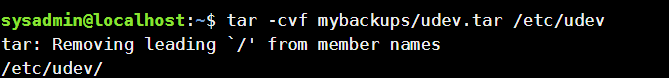
**2. Опрацюйте всі приклади команд, що представлені у лабораторних роботах Lab 9: Archiving and Compression та Lab 10: Working With Text. Створіть таблицю для опису цих команд:**

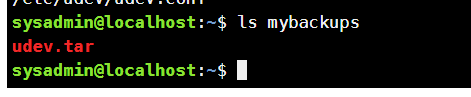
|  |  |
| --- | --- |
| **gzip** | Gzip is one of the most popular compression algorithms that allow you to reduce the size of a file and keep the original file mode, ownership, and timestamp. |
| **tar** | The Linux 'tar' stands for tape archive, is used to create Archive and extract the Archive files. |
| **bzip2** | bzip2 command in Linux is used to compress and decompress the files i.e. it helps in binding the files into a single file which takes less storage space as the original file use to take. It has a slower decompression time and higher memory use. |
| **zip** | bzip2 command in Linux is used to compress and decompress the files i.e. it helps in binding the files into a single file which takes less storage space as the original file use to take. It has a slower decompression time and higher memory use. |
| **cat** | The cat command is a utility command in Linux. One of its most common usages is to print the content of a file onto the standard output stream. Other than that, the cat command also allows us to write some texts into a file. |
| **less** | The less command is a Linux terminal pager that shows a file's contents one screen at a time. |
| **more** | more command is used to view the text files in the command prompt, displaying one screen at a time in case the file is large (For example log files). |
| **head** | more command is used to view the text files in the command prompt, displaying one screen at a time in case the file is large (For example log files). |
| **tail** | more command is used to view the text files in the command prompt, displaying one screen at a time in case the file is large (For example log files). |
| **sort** | SORT command is used to sort a file, arranging the records in a particular order. By default, the sort command sorts file assuming the contents are ASCII. |
| **wc** | wc stands for word count. As the name implies, it is mainly used for counting purpose. It is used to find out number of lines, word count, byte and characters count in the files specified in the file arguments. |
| **cut** | wc stands for word count. As the name implies, it is mainly used for counting purpose. It is used to find out number of lines, word count, byte and characters count in the files specified in the file arguments. |
| **grep** | In Linux and Unix Systems Grep, short for “global regular expression print”, is a command used in searching and matching text files contained in the regular expressions. |
| **egrep** | The egrep command belongs to the family of the grep command which is used for pattern searching in Linux. If you have used the grep command, egrep works the same as grep -E (grep Extended regex') does. |
| **find** | The find command in UNIX is a command line utility for walking a file hierarchy. It can be used to find files and directories and perform subsequent operations on them. |
| **tr** | The Linux ‘tar’ stands for tape archive, is used to create Archive and extract the Archive files. tar command in Linux is one of the important command which provides archiving functionality in Linux. |

**3.Ознайомтесь з командою tar та за її допомогою виконати у терміналі наступні дії:**

* ***створити файл з розширенням .tar;***

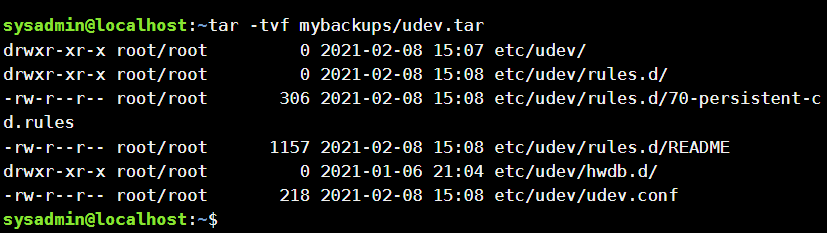




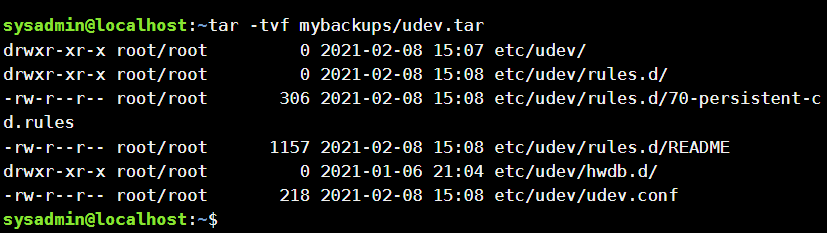


* ***створити файл з розширенням .tar, що складається з декількох файлів і каталогів одночасно;***

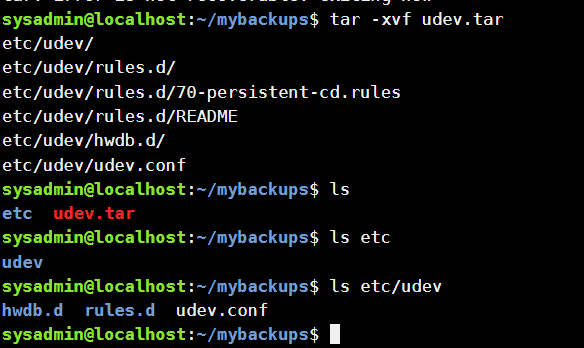
Already did that with first command, contents of archive:



* ***перегляду вмісту файлу;***



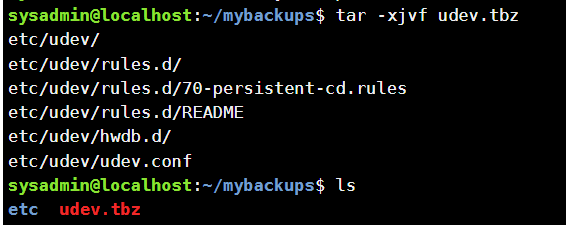
* ***витягти вміст файлу tar;***



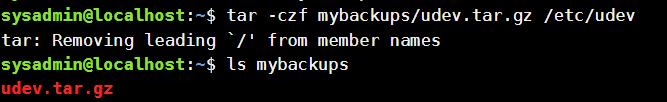
* ***створити архівний файл tar, стиснений за допомогою bzip;***



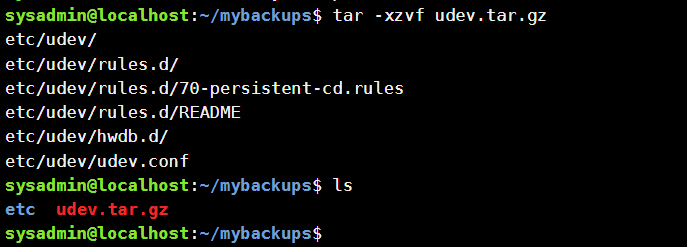
* ***витягти вміст файлу tar bzip;***



* ***створити архівний tar файл, стисненого за допомогою gzip;***



* ***витягти вміст файлу tar gzip.***



1. ***Як буде відбуватись перенаправлення потоків виведення в bash для наступних дій з командами (позначено як cmd) та файлами (позначено як file):***

|  |  |
| --- | --- |
| **Команда** | **Що виконує команда?** |
| cmd 1> file | Redirects input into file with overwriting its contents |
| cmd > file | Same as first one, same commands |
| cmd 2> file | Redirects input into file without overwriting its contents, adding information to the file. |
| cmd >> file | Same as «**cmd 2> file**» |
| cmd &> file | redirects both stdout (>) and stderr (>>) to the result of the expansion of file |
| cmd > file 2>&1 | Redirects input into file with overwriting its contents, after that 2>&1 redirects stderr (>>) to the current value of stdout (>) |
| cmd >> file 2>&1 | Redirects input into file without overwriting its contents, after that 2>&1 redirects stderr (>>) to the current value of stdout (>) |
| cmd 2> &1 > /dev/null | 2>&1 redirects stderr (>>) to the current value of stdout (>), and being redirected into /dev/null, which redirects output into the void |
| cmd 2> /dev/null | Redirects stderr (2>) to the void. |
| cmd 1 | cmd2 | Reads all output from first command and redirects it into input of another |
| cmd 1 2>&1 | cmd2 | Redirects stderr (>>) to the current value of stdout (>), and uses that as input for cmd2 |

1. ***Розгляньте наведені нижче приклади та поясніть, що виконують дані команди та який тип перенаправлення потоків вони використовують:***

|  |  |  |
| --- | --- | --- |
| **Команда (контейнер команд)** | **Що виконує команда?** | **Який потік перенаправлення?** |
| $echo "It is a new story." > story | Inputs «It is a new story» into story file | Stdout (>) |
| $ date > date.txt | Redirects output of date command to date.txt | Stdout (>) |
| $ cat file1 file2 file3 > bigfile | Redirects contents of all three files to the bigfile | Stdout (>) |
| $ls -l >> directory | Redirects information from ls -l command to directory file, without overwriting it | Stderr (>>) |
| $ sort < file1\_unsorted > file2\_sorted | Sorts contents of file1\_unsorted and redirects output to file2\_sorted | Stdout (< >) |
| $ find -name '\*.txt' > file.txt 2> /dev/null | Find all files specified by .txt, output their contents to file.txt and output all to void. | Stdout (>) |
| $ cat file1\_unsorted | sort > file2\_sorted | Read all output from first command and redirect it into input of second one, as other file | Stoud (>) |
| $ cat myfile | grep student | wc -l | Read all output from first command and redirect it into input of second one, find student in it, count amount of «student» words in it | Stdout (>) |

**Контрольні запитання**

*Готував матеріал студент Титов О.*

1. **Надайте порівняльну характеристику процесам стискання та архівування.**

gzip and bzip2 are for compressing single file. bzip2 is more efficient than gzip to compress the file but takes more time to do the compression.

zip and tar are for compressing and archiving file. tar can use gzip or bzip2 to do the compress and then archive.

tar, compressing and archiving files. It supports archive compressing through gzip and bzip2. If you are compressing more than 2 files, tar is recommended instead of gzip or bzip2

bzip2 has notably better compression ratio than gzip, which has to be the reason for the popularity of bzip2; it is slower than gzip especially in decompression and uses more memory. However the memory requirements of bzip2 should be nowadays no problem even on older hardware.

1. **Які програми, окрім наведених в роботі, можуть використовуватись для стискання та архівування файлів та каталогів в ОС Linux? Наведіть приклади та їх короткий опис.**

xz is a general-purpose data compression and decompression command-line tool similar to gzip and bzip2. It can be used to compress and decompress the files. The native file format of xz is .xz. But it can also support other various formats to compress or decompress files.xz gives us complete control over the compression and decompression of files. In Linux, xz tool is come by default with the system you don’t need to install it. Now let’s see how to use the xz tool to compress the files.

**xz filename**

ar command is used to create, modify and extract the files from the archives. An archive is a collection of other files having a particular structure from which the individual files can be extracted. Individual files are termed as the members of the archive.

**ar [OPTIONS] archive\_name member\_files**

1. **Порівняйте алгоритми стискання, що використовуються в командах (програмах), використовуваних в Linux. Які з алгоритмів можна вважати найшвидшим та найефективнішим?**

bzip2 has notably better compression ratio than gzip, which has to be the reason for the popularity of bzip2; it is slower than gzip especially in decompression and uses more memory. However the memory requirements of bzip2 should be nowadays no problem even on older hardware.

1. **Опишіть програмні засоби для стискання та архівування, що можуть бути використані у вашому мобільному телефоні.**

Android has built-in file manager, which support archiving and extraction of zip-files, but nothing more. You can download third-party programs from play market or internet, which can help you with more in-depth archive management.

1. **Опишіть та порівняйте програмні засоби для стискання та (де)архівування даних у ОС сімейства Windows.**

Common archiving tools for the Windows are 7zip and WinRAR:

WinRAR is a trialware file archiver utility for Windows, developed by Eugene Roshal of win.rar GmbH. It can create and view archives in RAR or ZIP file formats, and unpack numerous archive file formats. To enable the user to test the integrity of archives, WinRAR embeds CRC32 or BLAKE2 checksums for each file in each archive. WinRAR supports creating encrypted, multi-part and self-extracting archives.

WinRAR is a Windows-only program.

7-Zip is an open source file archiver designed originally for Microsoft Windows. 7-Zip primarily uses the 7z archive format, and can also read and write to several other archive formats. The program can be used from either a command-line interface, graphical user interface, or Windows shell integration. 7-Zip began in 1999 and is actively developed by Igor Pavlov. It is related to a cross-platform port, p7zip.

1. **Поясніть яким чином стиснення та архівування даних може бути використано для резервування даних. В яких ще задачах системного адміністрування воно може бути використано.**

The process of reducing the size of a data file is often referred to as data compression. In the context of data transmission, it is called source coding; encoding done at the source of the data before it is stored or transmitted. Source coding should not be confused with channel coding, for error detection and correction or line coding, the means for mapping data onto a signal.

Compression is useful because it reduces the resources required to store and transmit data. Computational resources are consumed in the compression and decompression processes. Data compression is subject to a space–time complexity trade-off. For instance, a compression scheme for video may require expensive hardware for the video to be decompressed fast enough to be viewed as it is being decompressed, and the option to decompress the video in full before watching it may be inconvenient or require additional storage. The design of data compression schemes involves trade-offs among various factors, including the degree of compression, the amount of distortion introduced (when using lossy data compression), and the computational resources required to compress and decompress the data.

1. **Яке призначення директорії файлу /dev/null?**

/dev/null in Linux is a null device file. This will discard anything written to it, and will return EOF on reading. Whatever you write to /dev/null will be forgotten into the void. It’s known as the null device in a UNIX system. In certain situations, the output may not be useful at all.

Since it discards anything written to it, you can move files to /dev/null to delete them. But, this is not a widely used use case. It is mainly used to discard standard output and standard error from an output.

*Робота студентів групи КСМ-03Б Команда 2: Нестолій Н., Усенко С, Титов І.*

**Висновок:**

Під час виконання лабораторної роботи ми отримали практичні навички роботи з оболонкою Bash та ознайомилися з базовими командами для архівування та стиснення даних та навчилися базовим діям при роботі з текстом у терміналі.