“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 студенти

групи РПЗ-13А та РПЗ-13Б

Eleven Two Zeroes:

Vlad Sapozhnyk

Max Karpenko

Dmytro Onufriiev

Київ 2024

Робота студентів групи РПЗ-13А та РПЗ-13Б Eleven Two Zeroes: Vlad Sapozhnyk, Max Karpenko and Dmytro Onufriiev.

**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 та віртуальна машина Virtual Box (Oracle).
3. ОС GNU/Linux (будь-який дистрибутив).
4. Сайт мережевої академії Cisco netacad.com та його онлайн курси по Linux

**Завдання для попередньої підготовки: *Created by Dmytro Onufriiev***

**Словник термінів**

| **Термін** | **Переклад** |
| --- | --- |
| Compression | Стиснення |
| Decompression | Розпакування |
| Archive | Архів |
| Lossy compression | Стиснення з втратами |
| Lossless compression | Стиснення без втрат |
| Tar file | Файл tar |
| gzip | gzip (інструмент стиснення) |
| bzip2 | bzip2 (інструмент стиснення) |
| xz | xz (інструмент стиснення) |
| DEFLATE algorithm | Алгоритм DEFLATE |
| Burrows-Wheeler algorithm | Алгоритм Burrows-Wheeler |
| LZMA2 algorithm | Алгоритм LZMA2 |
| Block size | Розмір блоку |
| Memory usage | Використання пам'яті |
| Compatibility | Сумісність |
| Verbose output | Детальний вивід |
| Extract | Витягувати |
| Preserve directory structure | Зберігати структуру директорій |
| File extension | Розширення файлу |
| Command line flag | Прапорець командного рядка |
| Standard out | Стандартний вихід |
| Extreme compression variant | Варіант екстремального стиснення |

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

tar: creates and decompresses file archives.

c: creates an archive.

x: unzips the archive.

f: specifies the name of the archive.

v: displays detailed information about the process.

z: compresses the archive using gzip.

j: compresses the archive using bzip2.

J: compresses the archive using xz.

The parameters can be set as follows:

tar -cvf my\_archive.tar my\_folder

xz: compresses and decompresses files.

-c: compresses files.

-d: decompresses files.

-z: sets the compression level (1-9).

-k: Saves file permissions.

You can set the options as follows:

xz -c my\_file.txt > my\_file.xz

zip: creates and decompresses ZIP archives.

-r: creates a recursive archive.

-x: Unzips an archive.

-f: specifies the name of the archive.

-e: enables encryption.

The options can be set as follows:

zip -r my\_archive.zip my\_folder

bzip2: compresses and decompresses files.

-c: Compresses files.

-d: decompresses files.

-z: sets the compression level (1-9).

The options can be set as follows:

bzip2 -c my\_file.txt > my\_file.bz2

gzip: compresses and decompresses files.

-c: Compresses files.

-d: decompresses files.

-z: sets the compression level (1-9).

The options can be set as follows:

gzip -c my\_file.txt > my\_file.gz

2. \*\*Наведіть три приклади реалізації архівування та стискання даних різними командами.

tar -cvzf my\_archive.tar.gz my\_folder

This command creates the archive my\_archive.tar.gz, which contains a compressed copy of the folder my\_folder.

zip -r -e my\_archive.zip my\_folder

This command will create a ZIP archive my\_archive.zip containing a recursive copy of the my\_folder folder. The archive will be encrypted, so you will need a password to unzip it.

bzip2 -c my\_file.txt > my\_file.bz2

This command will compress the file my\_file.txt and save the result to the file my\_file.bz2.

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

cat: displays the contents of files on the console.

-n: number lines.

-E: Show end-of-line characters.

-T: print tabs as spaces.

-s: compress consecutive blank lines.

An example of how to use the options:

cat -n my\_file.txt

less: view files page by page.

/: search for text.

n: Go to the next line.

p: go to the previous line.

h: help.

An example of how to use the parameters:

less -F my\_file.txt

more: view files page by page.

/: Search for text.

Enter: Go to the next page.

q: exit.

An example of using the parameters:

more +5 my\_file.txt

head: displays the first lines of the file.

-n: specifies the number of lines to print.

-q: do not print the header.

An example of using the parameters:

head -10 my\_file.txt

tail: prints the last lines of the file.

-n: specifies the number of lines to print.

-f: follow file changes.

An example of using the parameters:

tail -f my\_file.txt

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

Channels are virtual pipes that allow you to transfer data between processes. One process can write data to a channel, and another can read from it. Channels are used for communication between teams.

Threads are multiple channels that can be used to process data in parallel. Data from one channel can be split into streams and then processed simultaneously by different commands. Threads are used to improve performance.

Filters are commands that take data as input and transform it before outputting it as output. Filters can be used to format data, sort, search, and other operations. Filters are used for flexible data processing.

5. \*Яке призначення команди grep?

grep is a command used to search for strings matching a specific pattern in one or more files.

It has the following options:

-E: Use an extended regular expression.

-i: Ignore case sensitive characters.

-n: print the line number where a match is found.

-c: Print only the number of matches.

-A: print the context around the match.

-B: print the context after the match.

Example:

grep "hello" my\_file.txt

This command will find all lines in the file my\_file.txt that contain the word "hello".

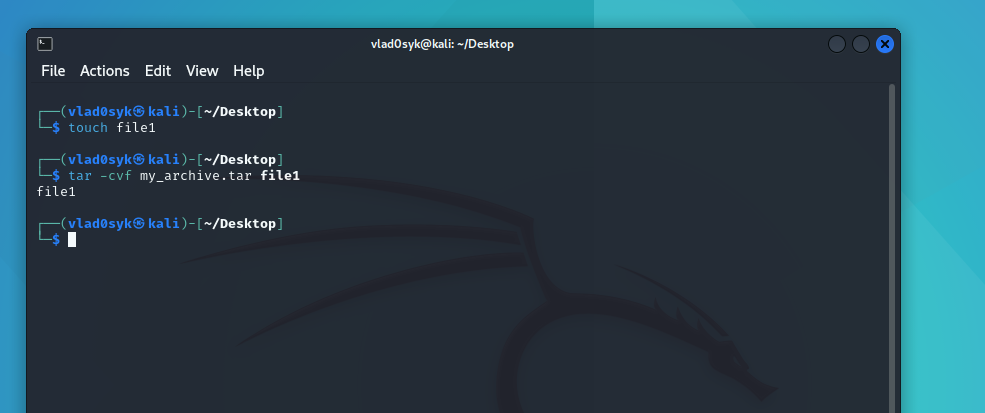
**Хід роботи. *Created by Vlad Sapozhnyk***

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

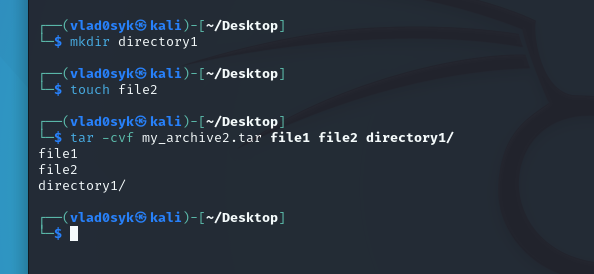
| Name of the command | Its purpose and functionality |
| --- | --- |
| mkdir mybackups | Create a new mybackups directory in the user's home directory. |
| tar -cvf mybackups/udev.tar /etc/udev | The tar command is used to combine several files into a single file. In this case, the contents of the /etc/udev directory will be saved to the udev.tar archive in the mybackups directory. The -c option tells the tar command to create a tar file. The -v option stands for "verbose," which tells the tar command to show what it is doing. The -f option is used to specify the name of the tar file. |
| tar –tvf mybackups/udev.tar | Display the contents of a tar file by using the available options (t = list contents, v = verbose, f = filename). |
| tar –zcvf mybackups/udev.tar.gz /etc/udev | To create a tar file that is compressed use -z option: The -z option makes use of the gzip utility to perform compression. |
| xz words | Using **xz** and **unxz** to compress and uncompress a file is also very similar to using **gzip** and **gunzip**. The compressed file is created with a **.xz** extension. The extension is removed when uncompressed. Execute the following commands to compress a copy of the words file. |
| unxz words.xz | Execute the following commands to uncompress the words.xz file. |
| zip words.zip words | Use the zip command to compress the words file. |
| zip -r udev.zip /etc/udev | Compress the /etc/udev directory and its contents with zip compression. |
| unzip -l udev.zip | To view the contents of a zip archive, use with the -l option with the unzip command. |
| rm -r etc  unzip udev.zip | To extract the zip archive, use the unzip command without any options. In this example we first need to delete the files that were created in the earlier tar example. |
| echo "Hello World"  echo "Hello World" > mymessage  cat mymessage | Use the redirection symbol > along with the echo command to redirect the output from the normal output of stdout (to the terminal) to a file. The cat command can be used to display file contents and will be used in this example to verify redirected output to the file. Type the following. |
| echo "Greetings" > mymessage  cat mymessage | When you use the > symbol to redirect stdout, the contents of the file are first destroyed. Type the following commands to see a demonstration. |
| cat mymessage  echo "How are you?" >> mymessage  cat mymessage | You can avoid clobbering a file by using >> instead of >. By using >> you append to a file. Execute the following commands to see a demonstration of this. |
| find ~ -name "\*bash\*" | The find command is a good command to demonstrate how stderr works. This very flexible command allows searching with a host of options such as filename, size, date, type and permission.  The find command will begin the search in the directory specified and recursively search all of the subdirectories. For example, to search for files beginning in your home directory containing the name bash. |
| find /etc -name hosts 2> err.txt  cat err.txt | To redirect stderr (error messages) to a file, issue the following command. |
| find /etc -name hosts > std.out 2> std.err  cat std.err  cat std.out | You can also redirect stdout and stderr into two separate files. |
| find /etc -name hosts > find.out 2>&1  cat find.out | To redirect both standard output (stdout) and standard error (stderr) to one file, first redirect stdout to a file and then redirect stderr to that same file by using the notation 2>&1. |
| tr a-z A-Z  this is interesting  how do I stop this?  ^D | Standard input (stdin) can also be redirected. Normally stdin comes from the keyboard, but sometimes you want it to come from a file instead. For example, the tr command translates characters, but it only accepts data from stdin, never from a file name given as an argument. This is great when you want to do something like capitalize data that is inputted from the keyboard (Note: Press **Control**+**d**, to signal the tr command to stop processing standard input). |
| tr A-Z a-z > myfile  Wow, I SEE NOW  This WORKS! | The tr command accepts keyboard input (stdin), translates the characters and then redirects the output to stdout. To create a file of all lower-case characters, execute the following. |
| cat myfile  tr a-z A-Z < myfile | Execute the following commands to use the tr command by redirecting stdin from a file. |
| ls -l /etc | more | Another popular form of redirection is to take the output of one command and send it into another command as input. For example, the output of some commands can be massive, resulting in the output scrolling off the screen too quickly to read. Execute the following command to take the output of the ls command and send it into the more command, which displays one page of data at a time. |
| cut -d: -f1 /etc/passwd | In the following example, you will use a command called cut to extract all of the usernames from a database called /etc/passwd (a file that contains user account information). First, try running the cut command by itself. |
| cut -d: -f1 /etc/passwd | sort | The output in the previous example was unordered and scrolled off the screen. In the next step you are going to take the output of the cut command and send it into the sort command to provide some order to the output. |
| cut -d: -f1 /etc/passwd | sort | more | Now the output is sorted, but it still scrolls off the screen. Send the output of the sort command to the more command to solve this problem. |
| cat /etc/passwd | The /etc/passwd is likely too large to be displayed on the screen without scrolling the screen. To see a demonstration of this, use the cat command to display the entire contents of the /etc/passwdfile. |
| more /etc/passwd | Use the more command to display the entire contents of the /etc/passwd file. |
| h | While you are in the more command, you can view the help screen by pressing the **h** key. |
| <SPACE> | Press the **Spacebar** to view the rest of the document. |
| less /etc/passwd  /bin  nnnNNNq | Use the less command to display the entire contents of the /etc/passwd file. Then search for the word bin, use **n** to move forward, and **N** to move backwards. Finally, quit the less pager by typing the letter **q.** |
| head /etc/passwd | You can use the head command to display the top part of a file. By default, the head command will display the first ten lines of the file. |
| tail /etc/passwd | Use the tail command to display the last ten lines of the /etc/passwd file. |
| head -2 /etc/passwd | Use the head command to display the first two lines of the /etc/passwd file. |
| ls /etc | tail -5 | Execute the following command line to pipe the output of the ls command to the tail command, displaying the last five file names in the /etc directory. |
| head -n -20 /etc/passwd | Another way to specify how many lines to output with the head command is to use the option -n -#, where # is the number of lines counted from the bottom of the output to exclude. Notice the minus symbol - in front of the #. For example, if the /etc/passwd contains 27 lines, the following command will display lines 1-7, excluding the last twenty lines. |
| cd /etc  grep sshd passwd | The use of grep in its simplest form is to search for a given string of characters, such as sshd in the /etc/passwd file. The grep command will print the entire line containing the match. |
| grep root passwd | Regular expressions are "greedy" in the sense that they will match every single instance of the specified pattern. |
| grep '^root' passwd | To limit the output, you can use regular expressions to specify a more precise pattern. For example, the caret ^ character can be used to match a pattern at the beginning of a line; so, when you execute the following command line, only lines that begin with root should be matched and displayed. |
| grep 'sync' passwd | Match the pattern sync anywhere on a line. |
| grep 'sync$' passwd | Use the $ symbol to match the pattern sync at the end of a line. |
| grep '.y' passwd | Use the period character . to match any single character. For example, execute the following command to match any character followed by a 'y'. |
| grep 'sshd|root|operator' passwd | The pipe character, |, or "alternation operator", acts as an "or" operator. For example, execute the following to attempt to match either sshd, root or operator. |
| grep -E 'sshd|root|operator' passwd | Use the -E switch to allow grep to operate in extended mode in order to recognize the alternation operator. |
| egrep 'no(b|n)' passwd | Use another extended regular expression, this time with egrep with alternation in a group to match a pattern. The strings nob and non will match. |
| head passwd | grep '[0-9]' | The [ ] characters can also be used to match a single character. However, unlike the period character ., the [ ] characters are used to specify exactly what character you want to match. For example, if you want to match a numeric character, you can specify [0-9]. Execute the following command for a demonstration. |
| grep -E '[0-9]{3}' passwd | Suppose you want to search for a pattern containing a sequence of three digits. You can use { } characters with a number to express that you want to repeat a pattern a specific number of times; for example: {3}. The use of the numeric qualifier requires the extended mode of grep. |

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

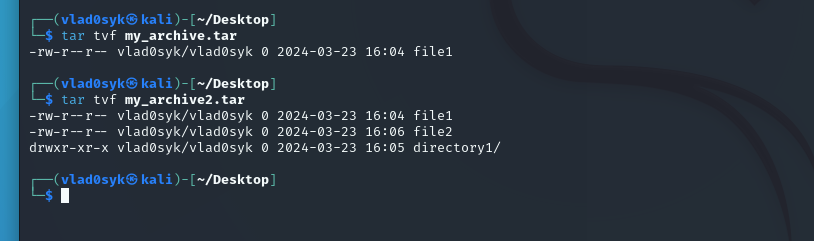
* створити файл з розширенням .tar;



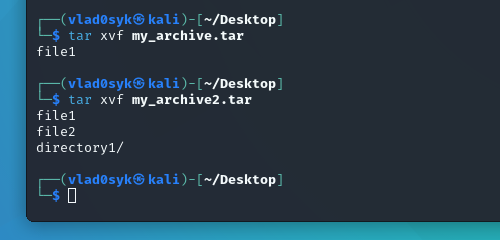
* створити файл з розширенням .tar, що складається з декількох файлів і каталогів одночасно;



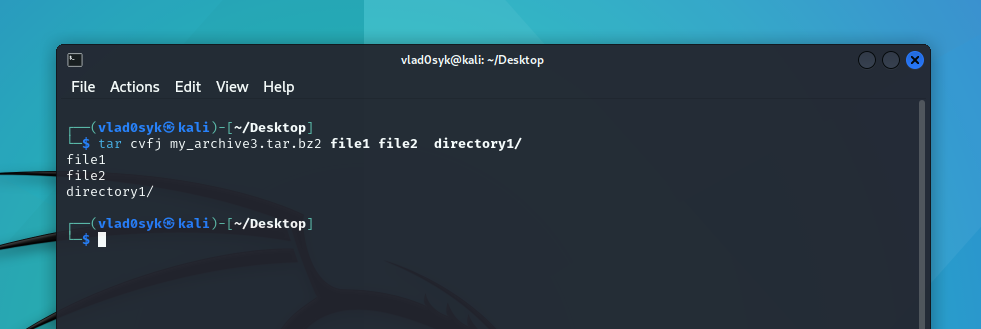
* перегляду вмісту файлу;



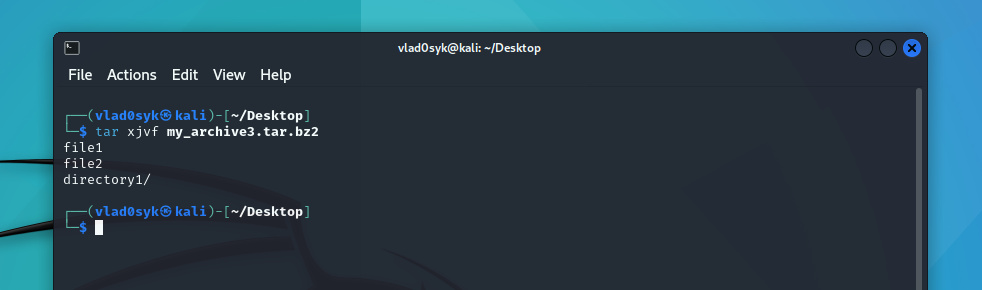
* витягти вміст файлу tar;



* 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):

| Command | What does command do? |
| --- | --- |
| cmd 1> file | The command cmd 1> file attempts to redirect the output of the command cmd to a file named file. |
| cmd > file | The command cmd > file attempts to redirect the standard output of the command cmd to a file named file. |
| cmd 2> file | The command cmd 2> file attempts to redirect the standard error of the command cmd to a file named file. |
| cmd >> file | The command cmd >> file attempts to append the standard output of the command cmd to a file named file. |
| cmd &> file | The command cmd &> file redirects both the standard output **and** standard error of the command cmd to a file named file. |
| cmd > file 2>&1 | The command cmd > file 2>&1 combines two redirection operators to control the output of the command cmd (assuming cmd is a valid command) |
| cmd >> file 2>&1 | The command cmd >> file 2>&1 combines redirection techniques to achieve a specific outcome. |
| cmd 2>&1 > /dev/null | The command cmd 2>&1 > /dev/null silences the output of the command cmd, regardless of whether it's successful or produces errors. |
| cmd 2> /dev/null | The command cmd 2> /dev/null aims to suppress any error messages that might be generated during the execution of the command cmd. |
| cmd1 | cmd2 | The command cmd1 | cmd2 uses a concept called piping (|) to connect the output of one command (cmd1) to the input of another command (cmd2). |
| cmd1 2>&1 | cmd2 | The command cmd1 2>&1 | cmd2 combines redirection and piping to achieve a specific outcome. |

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

| Command  (command container) | What does the command do? | What is the redirection flow? |
| --- | --- | --- |
| $echo "It is a new story." > story | Creates a file named "story" and writes "It is a new story." into it. | Standard output (echo) is redirected (>) to create a new file named "story" and write the message "It is a new story." into it. |
| $ date > date.txt | Stores the current date and time in a file named "date.txt". | Standard output (date) is redirected (>) to create a new file named "date.txt" and store the current date and time in it. |
| $ cat file1 file2 file3 > bigfile | Combines contents of "file1", "file2", and "file3" into a new file named "bigfile". | Standard output (cat) of concatenated files ("file1", "file2", and "file3") is redirected (>) to create a new file named "bigfile". |
| $ls -l >> directory | Appends a list of files and their details to a file named "directory". | Standard output (ls -l) containing the directory listing is appended (>>) to a file named "directory". If "directory" doesn't exist, it's created. |
| $ sort < file1\_unsorted > file2\_sorted | Sorts the contents of "file1\_unsorted" alphabetically and saves them in "file2\_sorted". | Input (<) for sort comes from the contents of "file1\_unsorted".  Standard output (sort) of the sorted lines is redirected (>) to create a new file named "file2\_sorted". |
| $ find -name '\*.txt' > file.txt 2> /dev/null | Lists paths of .txt files, saving them in "file.txt", and hides any errors. | Standard output (find) containing filenames is redirected (>) to create a file named "file.txt".  Standard error (2>) is redirected to /dev/null, which discards any error messages from find. |
| $ cat file1\_unsorted | sort > file2\_sorted | Similar to previous sort command, but uses a pipe to connect commands for sorting. | Standard output (cat) of "file1\_unsorted" is sent through a pipe (|) as input for sort.  Standard output (sort) of the sorted lines is redirected (>) to create a new file named "file2\_sorted". |
| $ cat myfile | grep student | wc -l | Counts lines containing "student" in "myfile". | Standard output (cat) of "myfile" is sent through a pipe (|) as input for grep.  Standard output (grep) of lines containing "student" is sent through another pipe (|) as input for wc -l.  Standard output (wc -l) containing the line count is displayed on the console (not redirected). |

**Контрольнi запитання: *Created by Max Karpenko***

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

Архівування та стискання даних – це різні процеси, які виконують відмінні, але доповнюючі один одного завдання. Архівування зосереджено на об'єднанні множини файлів або папок у єдиний архівний файл, що полегшує їх зберігання, управління та передачу, в той час як стискання зменшує обсяг даних, використовуючи алгоритми для економії місця на диску або під час передачі даних через мережу. Хоча обидва процеси можуть використовуватися окремо, вони часто застосовуються разом для оптимізації зберігання та передачі даних. Наприклад, після архівування групи файлів у один архів, цей архів можна додатково стиснути, щоб зменшити його розмір, що є особливо корисним для великих наборів даних або для ефективного розподілу програмного забезпечення.

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

В ОС Linux, крім загальновідомих інструментів для стискання та архівування файлів, таких як tar, gzip, bzip2, існують і інші програми, які можуть бути використані для подібних завдань. Наприклад, програма xz використовує алгоритм **LZMA** для стискання, який забезпечує високу ступінь компресії, але може вимагати більше часу та обчислювальних ресурсів. Інший інструмент, **lrzip** (Long Range ZIP), оптимізований для роботи з дуже великими файлами, використовуючи різні алгоритми стискання для досягнення кращої ефективності. **Zstd** (Zstandard), розроблений **Facebook**, є відносно новим алгоритмом стискання, який забезпечує дуже швидке стискання та розпакування при збереженні високого рівня компресії. Ці інструменти доповнюють стандартний набір програм Linux для роботи з архівами і стисненими файлами, пропонуючи додаткові опції для оптимізації зберігання даних та їх передачі.

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

Алгоритми стискання, які використовуються в програмах для Linux, різняться за швидкістю та ефективністю. gzip використовує алгоритм DEFLATE, який забезпечує хороший баланс між швидкістю та ступенем стискання. bzip2 використовує алгоритм Burrows–Wheeler, що забезпечує кращу компресію, ніж gzip, але працює повільніше. xz використовує алгоритм LZMA, який пропонує ще вищий рівень компресії за рахунок швидкості. Zstd (Zstandard) є відносно новим і вирізняється своєю здатністю швидко стискати та розпаковувати дані, при цьому забезпечуючи високу ступінь компресії. Таким чином, за швидкістю Zstd можна вважати одним з найшвидших, тоді як за ефективністю стискання (ступенем компресії) xz з алгоритмом LZMA є одним із найефективніших.

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

На мобільних телефонах для стискання та архівування файлів можуть використовуватися різноманітні додатки. Для Android і iOS популярні такі програми, як WinZip і RAR, що дозволяють створювати, відкривати та розпаковувати архіви форматів ZIP і RAR. Також існують додатки типу 7-Zipper для Android, які підтримують широкий спектр форматів, включаючи 7z, і пропонують функції шифрування. Ці програми надають зручний інтерфейс для роботи з архівами безпосередньо з мобільного пристрою, дозволяючи керувати файлами на ходу.

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

В ОС сімейства Windows для стискання та деархівування даних існує кілька популярних програмних засобів. Вбудований інструмент Windows дозволяє створювати та розпаковувати ZIP-архіви без додаткового програмного забезпечення. WinRAR є широко відомим за підтримку великої кількості форматів, включаючи RAR і ZIP, та пропонує опції шифрування та розподілу архівів на частини. 7-Zip виділяється підтримкою формату 7z, який забезпечує високий рівень компресії, та можливістю працювати з багатьма іншими форматами. WinZip пропонує інтуїтивно зрозумілий інтерфейс та інтеграцію з хмарними сховищами. В порівнянні, 7-Zip найбільш відомий своєю ефективністю стискання та відкритим кодом, WinRAR славиться своєю універсальністю та налаштуваннями безпеки, а WinZip – зручністю у використанні та додатковими функціями для роботи з хмарними сервісами.

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

Стиснення та архівування даних для резервного копіювання дозволяє зменшити обсяг зберігання, спрощуючи управління резервними копіями та знижуючи витрати на зберігання. Це також сприяє швидшій передачі даних через мережу при віддаленому резервному копіюванні. Крім резервного копіювання, стиснення та архівування корисні в інших задачах системного адміністрування, таких як оптимізація передачі файлів між серверами, архівування старих даних для економії місця, створення єдиних пакетів дистрибутивів програмного забезпечення для спрощення розгортання, а також шифрування важливих даних для забезпечення безпеки перед архівуванням.

**7. \*\*Яке призначення директорії файлу /dev/null?**

Директорія файлу **/dev/null** у Unix-подібних операційних системах слугує як "**чорна діра**" для даних. Все, що направляється до /dev/null, **ігнорується** або "**знищується**", тобто дані не зберігаються і не обробляються. Це може використовуватися для відкидання небажаного виводу програм або процесів.

**Висновок:** Пiд час виконання лабораторної роботи, ми отримали практичнi навички роботи з командною оболонкою Bash; Ознайомились з базовими командами для архівування та стиснення даних; Ознайомились з базовими діями при роботі з текстом у терміналі.