«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 КСМ-03Б

Команда: Кучмій-Зікеєв

М.О, Зінченко І.О. та

Звєрьков А.Г.

Перевірив викладач

Сушанова В.С.

Київ 2022

**Завдання для попередньої підготовки.**

***Виконав студент Зінченко І.***

**1.** ***tar*** First of all, tar has been victorious for creating archives on a magnetic stripe, but at the same time, tar has been victorious for collecting more files in the middle of one file, for expanding software security, and also for direct recognition - for creating an archive file system.

***xz*** compresses one file on input, but does not fit a set of files into one archive. In this way, we squeeze a file, which can be an archive, for example, we create such Unix programs, like tar or cpio.

The ***zip*** utility is ideal for coding on a Linux system. It is also necessary to work with the same file in another OS (for example, from Windows or MS DOS), more universal є zip, so it is suitable for exchanging files, solving problems from the summation. Victory zip and for archiving, and for embossing. After the procedure, the files may have a .zip extension.

The ***bzip*** utility is recognized to defeat the compressing function. The main plus is in pairing with other utilities, that bzip2 can be compressed with a great file format, but it also comes in the security of the process of archiving itself, which can be negatively identified when working with bulk files. The very principle of recognition and recognition is similar to gzip, but with a different compression algorithm, which itself implements compression of data at a high level. After the procedure, the files may have a .bz extension.

When squeezing, as an argument, the file names are entered, as it is necessary to compress, so you can additionally add options (keys). Assignments the file is replaced with a version with extensions .gz., Saving the file name, modes, access rights, modification hour and marks of the original file. As a result, gzip shows a smaller amount of compression, lower yogo іnshі twins (for example, bzip2), vikoristovuyuchi dosit the old algorithm, but with tsomu є nevibaglivim to system resources і, as a variant, suitable for machines with small strains.

**2.** A characteristic feature of the greater types of data is their superfluousness. Stupin surplus of data to lie in the type of data. For example, for video data, the levels of excess in dekilka are more low for graphic data, and the levels of excess of graphic data, in their own line, are greater for the level of excess of textual data. The main factor that contributes to the level of oversupply is the adopted coding system. The butt of coding systems can be zvichayny movi splkuvannya, yakі є nі chim іnhim, like coding systems to understand those ideas for the development of thoughts. Thus, it has been established that the coding of textual data for the additional benefit of the Ukrainian language gives an average excess of 20-25% more than the lower coding of similar data for the purposes of the English language. Particularly relevant is the problem of the need for the collection of great amounts of information with insignificant amounts of data. In connection with this, the problem of relieving excess or squeezing data is gradually blamed. If the methods of compressing data reach ready-made files, then often the term “data compressing” is replaced by the term “archiving data”, the compressing version of the data is called an archive, and the programs that implement the compressing methods are called archivers.

**3.** The ***cat*** command in Linux is one of the most useful utilities you should definitely learn. Being an abbreviation of the English word "concatenate" (concatenation), it allows you to create, combine, and also display the contents of files on the command line or in another file. The utility does not require additional installations as it is pre-installed in the coreutils package on any Debian or Red Hat based system.

The ***less*** command renders the text file in such a way that it allows the scroll bar to scroll up and down one row at a time. less also wins over the lock for man-sides, as discussed in previous lessons.

***more*** Victory for displaying together great file expansions. The more command is a simpler command with a small set of options, for example, in the case of a similar command less, but it’s enough to complete its main task.

The ***head*** command allows you to use the utility of the same name to display the first few lines from a text file or the output of another utility.

The ***tail*** command allows the user to read the commands of a file from the end. Also, with the help of this utility, you can track new information updated in real time in a particular file. This makes it very easy to check the latest entries in the system. Also, the tail command in Linux allows you to view multiple

**4.** The command shell in UNIX is a command-line interface in Unix-like operating systems, so that it can pick up commands, as it sends a message, or as it reads from files. Such files with shell commands are called shell scripts (scripts, programs). These scenarios are not compiled, but are interpreted by the shell. Tse means that the shell will read the script from the cob to the end, row by row, jokingly assigned there commands and vykonoyuchi їх; on the input of this code, the compiler transforms the whole program to look, attached to the code by the machine - then the file with such code can be written in the shell script. The peculiarity of the language of the shell is characteristic - rich operations, like in traditional languages, programming and awakening, vykonuyutsya with additional help from old programs.

UNIX lets you choose a range of different skins:

sh, or Bourne Shell one of the first shells, like a bula vikoristana in UNIX-like environments. Tse base shell, small and with a small set of capabilities. Vaughn is the de facto standard shell, which is present on the skin system from UNIX. On Linux, /bin/sh can be a symbolic link to bash. This is designed to keep the confusion with UNIX programs. bash, or Bourne Again Shell Linux standard shell. In most of the Linux boxers, bash itself is installed as a standard shell. In deaky sensi bash - tse superbudova over sh, nabir additional and additional modules. In this way, the new Bourne shell is summed up with the great Bourne shell: commands that work with sh will also work with bash, but not the same language.csh, or a C-oriented shell (C Shell) the syntax of the shell is similar to my C programming. tcsh, or the C-oriented shell has been reduced (Turbo C Shell) nadbudova over csh. ksh, or Korn Shell The bula was written on the cob of the 1980s and took away the best at that time the possibility of Born's shell and csh. Є nadbudovaya over the shell of Bourne, that guarantees the writing of scripts for the shell of Bourne. The POSIX standard is the Korn shell with sub-options. zsh is a modern shell, a great sum from bash.

**5.** Grep is one of the most useful system administrator tools, a kind of “Swiss knife” that helps a specialist solve many everyday routine tasks. It is most commonly used when searching for strings and patterns in a group of files or subfolders. We'll look at this command in more detail later in this article. Grep quickly and efficiently finds a line in a given file using some pretty clever and incredibly flexible criteria. Although in real life most of the tasks solved with the command in question are simple, there are many more complex uses of it that most people do not know about. This utility is based on the extended regular expression syntax added to UNIX right after the original implementation of this concept in 1981.

*There are two ways to use Grep, each with its own characteristics:*

Can be used to search for a given file or files in the system (including recursive subfolder searches).

Grep can also run in filter pipeline mode, taking input (usually via a pipe) from another command (or a series of preceding commands).

Thus, by combining these approaches, the application point of Grep becomes universal: you can search for files in directories according to complex criteria, filter the necessary processes in memory, search for substrings in the contents of one (or several) files, somehow process the output (results of work) of others commands through pipes, parse log files according to certain criteria, etc.

**Хід роботи.**

**2. *Виконав студент Звєрьков А.***

|  |  |
| --- | --- |
| Назва команди | Її призначення та функціональність |
| mkdir mybackups | Create a new mybackups directory in your home directory |
| tar -cvf mybackups/udev.tar /etc/udev | The tar command is used to combine multiple files into a one file. In this case, the contents of the /etc/udev directory will be saved in the udev.tar archive in the mybackups directory. The -c option tells the tar command to create a tar file. The -v option means ”verbose”, which tells the tar command to demonstrate that it does The -f option is used to specify the name of the tar file. |
| tar -cf archive.tar file1 file2 file3 | Creates an archive named archive.tar from the files named file1, file2, file3 |
| tar -cjf archive.tar.bz2 file1 file2 | To compress the archive with the bzip2 algorithm, invoke tar with the -j option. The following command creates a tar.bz2 archive from the given files: |
| tar -czf archive.tar.gz file1 file2 | The -z option tells tar to compress the archive using the gzip algorithm as it is created. For example, to create a tar.gz archive from given files, you would run the following command: |

**3. *Виконав студент Звєрьков А.***

1-3) 

4) 

5) 

6) 

7) 

8) 

**4. *Виконав студент Кучмій-Зікеєв М.***

|  |  |
| --- | --- |
| **Команда** | **Що виконує команда?** |
| cmd 1> file |  |
| cmd > file |  |
| cmd 2> file |  |
| cmd >> file |  |
| cmd &> file |  |
| cmd > file 2>&1 |  |
| cmd >> file 2>&1 |  |
| cmd 2>&1 > /dev/null |  |
| cmd 2> /dev/null |  |
| cmd1 | cmd2 |  |
| cmd1 2>&1 | cmd2 |  |

**5. *Виконав студент Кучмій-Зікеєв М.***

|  |  |
| --- | --- |
| **Команда**  **(контейнер команд)** | **Що виконує команда?** |
| $echo "It is a new story." > story | display the text "It is a new story." when opening the story file |
| $ date > date.txt | the current time and date will be written to the file |
| $ cat file1 file2 file3 > bigfile | merging multiple files into one |
| $ls -l >> directory | output the entire contents of a directory in single line format |
| $ sort < file1\_unsorted > file2\_sorted | sorts the contents of one file into another |
| $ find -name '\*.txt' > file.txt 2> /dev/null | Find all files with the extension .txt write the result to file.txt and do not save it in the system |
| $ cat file1\_unsorted | sort > file2\_sorted | move the contents of file1 sorted to file2 sorted after sorting the contents |
| $ cat myfile | grep student | wc -l | move contents of myfile to student find newlines and output them |

**Відповіді на контрольні запитання**

***Виконав студент Зінченко І.***

***1.*** Squeeze data

To change the size data use squeezing data for special algorithms (methods) of squeezing.

Shrinking data is the process of recoding data with the method of changing files.

Distinguish algorithms for squeezing, which secure squeezing without wasting data, and algorithms that transfer private waste of data.

Data embossing methods:

D. Huffman's algorithm and the Shannon-Fano algorithm transfer data recoding. Symbols, which are more often used in repetitions, are recoded with a smaller number of symbols, and those that are more common - with a larger one. For a file compressed in this way, the table of coding and additional service information is to be misplaced.

The RLE (Run-Length encoding) algorithm is often used to encode graphic data. This method of transferring the replacement of the sequence of data of the same type with a short description.

Archive data

The loss of tribute can lead to even more serious consequences, this is due to the need for the creation of copies of the tribute. The most important data are duplicated, recorded on other hard disks, but only on optical disks. Sound, for the sake of transparency and change of data, files and folders for the hour of creating backup copies are packed into one file. Such copies of data are called archives, and files, which are packed like stinks, are archive files, or, for simplicity, archives.

Archiving is the creation of copies of data from the archives of special programs that can be used to extract data.

Most operating systems have their own warehouse to create backup copies of data. These programs, call, copy data for updating the operating system on the same or on a different disk, and also create archives of files of the user.

2. Ashampoo Zip Free

Ashampoo ZIP Free is easy to use the ZIP compression tool. This software supports creating and extracting with 7-ZIP, LHA, TAR, CAB, and many other formats. It has an integrated preview facility to effortlessly view files. Feature:

***Feature:***

* It provides a user-friendly graphical interface.
* Offers integration with Windows Explorer.
* It allows you to create the unlimited size of the archive.
* You can repair broken zip files.
* The tool helps you to create SFX (Special Effects) archive.
* You can integrate with the Windows operating system and windows server.
* Support for RAR format.

***WinRAR***

WinRAR is a strong competitor of Winzip. It allows you to create ZIP and RAR archives. The tool is available in more than 50 languages. It automatically recognizes and selects the ideal compression method.

***Feature:***

* It enables you to compress files quickly and easily.
* You can split the archive file into a separate volume.
* This tool uses 256-bit password encryption for security.
* It uses a compression algorithm that helps you to compress multimedia files and executable files.
* WinRAR helps you to save your save disk space and transmission costs.
* The software has many integrated functions to manage your compress archive.

***7zip***

7-zip is a popular open-source file compression tool. This software uses strong AES-256 encryption technique for security. It is one of the best free WinZip alternatives that has a self-extracting facility for 7z file format. The app can be used for business and commercial sectors. Supported formats by 7-zip are 7z, GZIP, XZ, and BZIP2.

***Feature:***

Integration with Windows shell.

It has a file manager to effectively manage your files.

This tool can be used from the command line.

Offers plugin for a FAR manager (A software that manages archive and files)

It can be Localized for more than 87 languages.

7-zip uses LZMA (Lempel Ziv Markov Chain Algorithm) and LZMA2 for the compression.

***Bandzip***

Bandizip is a free archiver for MS windows. It supports archive formats like ZIP, RAR, 7Z, and RAR5. This tool allows users to extract multiple files. The software helps you to easily open the extracted folder.

***Feature:***

This application supports Windows operating system.

It is free software.

Extract files in more than 30 formats.

You can compress archive with multi-volume and password.

It enables you to compress fast with multi-core.

Bandizip supports compression of a file having more than 4 GB.

This tool offers security through an AES256 encryption algorithm.

**Express Zip**

Express Zip is easy to use and comprehensive file archive and compression tools available. This tool enables you to unzip zip archives and create zip compatible folders and files. It supports a wide range of popular archive formats, including ROpen RAR, 7Z, TAR, CAB, and more.

**Feature:**

* Create compatible zip files.
* It enables you to unzip archives.
* This tool has improved compression ratio.
* Email ZIP files in two easy clicks.
* You can open and extract data from ISO (International Organization for standardization) and RAR files.

***3.*** Image embossing is the use of algorithms for compressing data to images that are captured in digital form. As a result of compression, the size of the image changes, which changes the hour of transmission of the image as much as possible and saves space for saving. The embossing of the image is divided into embossing with embossing and embossing without embossing. Embossing without loss is more suitable for custom-made images, such as graphics, program icons, or for special features, for example, as an image is recognized for further processing by image recognition algorithms. Algorithms for squeezing with costs with an increased degree of squeezing, as a rule, generate artefacts that are good for the human eye.

We have a lot of available tools, tools with a graphical interface for compressing, and we just happen to have a mother on the right with forgiveness and intuitively understand a graphical interface for compressing and decompressing, such as PeaZip or 7zip, ... just . In particular, the first building is practiced with different formats, zokrema needs over 180 of them. But if you are one of the quiet ones, who still loves working with the terminal, then you will have a large number of tools, as you know melodiously:

last but ta unpack: this is a good option, if you need files, how to transfer to other operating systems, but you will know tools for working with files in Microsoft Windows systems, as well as in macOS, and also in others. For example, to squeeze a file or a directory with the name test.

gzip: Better if you just want to port between Unix / Linux operating systems. Possibly, the step of squeezing may be identical to zip, perhaps, troch is smaller, but you will not know the great difference in expanding the file for the help of zip or gzip. For squeezing that unpacking for an additional tool, we can pick up two options for unpacking, and the second option is -do, without the middle picking alias gunzip

bzip2: As before, this algorithm is already present in Unix / Linux operating systems, although in the processes of compressing and decompressing it takes more than an hour, less than gzip. In this way, I didn’t make the jam to a greater degree of tightness, like in the case of xz, shards of files, squeezed under bzip2, take up three more, lower gzip. Axis why it is recommended to unify bzip2 and change which one to choose xz or gzip. If you want everything to be loose in a file type, which you can try to squeeze.

xz: This is the best format for large file expansions, shards of wines show the best compression indicators, but it's also true that it takes more than an hour to complete compression or decompression. It's more recent, less older, so you can rely on more primitive distributions and older Unix systems, as you can't for that tool.

unrar and rar: We can also work with RAR formats in Linux for these tools, although it's not as popular with other \* nix systems, as in the past.

compress and decompress: and although the use of compress is lost and not as popular as the previous ones, I would not want to miss this tool either. It is used to compress files with the .Z extension and does so using a modified Lempel-Ziv algorithm.

***4.***

***RAR***

RAR is a very popular and powerful yet easy-to-use data compression tool. It can create and open archives in all common formats like RAR, ZIP, TAR, GZ, BZ2, 7z, etc. It even functions as a basic file manager. A file can be quickly compressed or unzipped in a few simple steps.

***WinZip - Zip UnZip tool***

WinZip is another very popular program from the PC era. Now also available as an Android app, it supports the commonly used major compression formats. WinZip can create Zip and Zipx files, encrypt files, and open and extract compressed files. It can be used to send compressed large files via e-mail and share them in the cloud.

***ZA Archiver***

ZArchiver is a popular archive manager for Android. It has a simple, convenient, but functional interface. According to the creators, the program does not have permission to access the Internet, and therefore cannot transmit or receive data without the user's knowledge. ZArchiver can help you create compressed files with common file formats such as 7zip, zip, bz2, gz, XZ, lz4, tar and zstd. Similarly, it also supports opening or extracting several commonly used file types.

***Archiver B1***

When using B1 Archiver, you can use B1 Archiver's own file format or the more common zip format to compress data. It is a very powerful and comprehensive archiver program with support for decompression of more than 34 different formats. It can be used in more than 30 languages.

***ALZip***

ALZip is essentially a file manager with support for data compression and decompression. Thus, it can be used for all file manager functions such as open, copy, move, delete, rename files, search files and drag and drop files. Also, it supports compression/decompression in all commonly used major formats such as rar, 7z, egg, alz, tar, tbz, tbz2, tgz, lzh, jar, gz, bz, bz2, lha and alz split archive , egg and rar.

***Simple Unrar***

If you only work with RAR file formats, then this is the program for you. Simple Unrar has a minimalistic and easy-to-use interface. But it is still a powerful tool for extracting RAR archives with faster decompression. At that time, it does not support formats other than RAR, such as zip. It has been translated into several languages such as Chinese (Simplified), Dutch, French, Italian, German, Greek and many others.

***5.*** There is often a need to reduce the size of data stored in computer memory. For this, special methods of data compression are used, which are called algorithms (methods) of data compression. Data compression is used when creating certain types of files, such as graphics such as TIFF, JPEG, PNG or audio such as MPEG3, WMA, for transferring files over a network, etc. Data compression is the process of recoding data to reduce file sizes. There are compression algorithms that provide compression without data loss and algorithms that involve partial data loss. Algorithms with partial data loss (another name "with adjustable data loss") are used when data integrity is not very important. For example, when compressing graphic, video, sound files, since the human senses are not always able to notice a slight difference in shades of colors in a photo, in the reproduction of sound or video data, etc. Or the loss of quality is not as important as the need to reduce file sizes for recording to an optical disc or for use on the Internet. However, these algorithms cannot be applied when compressing text and numeric data.

The D. Huffman algorithm and the Shannon-Fano algorithm involve data recoding. Symbols that occur more often in the message are recoded with a smaller number of symbols, and those that are less frequent - with a larger number. So, for example, to encode in Windows-1251 the previous sentence, which contains 114 characters, including spaces, it is necessary to use 1 byte of data for each character, or 114 bytes. If you count how many times this or that symbol occurs in a sentence, it will turn out that there are 28 total symbols in the sentence, the symbol "space" occurs more often - 14 times, the symbol "i" - 10, "o" - 9, "c" - 7, "v" - 6, etc. The symbols "sh", "b", "z", "-", "." occur at least once. If you apply the data compression algorithm for the binary code (0 and 1) and mark the symbols that occur more often with one symbol (1 bit): the symbol "space" is 0, and the symbol "and" is 1, followed by the frequency of repetition - with two symbols ( 2 bits): character "o" is 00, character "c" is 01, etc. For less common characters, four binary digits or 4 bits will be used. As a result, the total code length of a sentence encoded in this way will be 90 bits or 11 bytes and 2 bits. As we can see, the total amount of sentence code will decrease by 10 times.

Despite the increased reliability of computers and computer data carriers, they still do not provide a full guarantee of data preservation. Data loss can lead to very serious consequences. Thus, the deletion or damage of data on deposits and transfers of clients' funds can lead to the collapse of the bank, the loss of data on ticket sales will complicate the transportation of passengers, the destruction of the results of experiments can nullify many years of scientific research. Even losing a notebook with your friends' phone numbers will cause you trouble. Therefore, there is a need to create copies of data. The most important data is duplicated by writing to other hard drives, optical drives, etc. Usually, for ease of use and reduction of data volumes, files and folders are packed into a single file when creating backup copies. Such copies of data are called archives, and the files into which they are packed are called archive files, or for simplicity - archives. Creating copies of data using special programs that can use data compression is called archiving. Most operating systems have built-in tools for creating backup copies of data. These programs usually copy data to restore the operating system on the same or another disk, and also create archives of user files.

***6.***

Backup

Backup - the process of creating a copy of data on a medium designed to restore data to its original location in the event of damage or destruction by appropriate programs - backup data duplicators.

Backup is necessary for the possibility of quick and inexpensive recovery of information (documents, programs, settings, etc.) in case of loss of a working copy of information for any reason.

In addition, related problems are solved:

duplication of data

data transfer and work with common documents

Backup system requirements:

Reliability of information storage – is ensured by the use of fail-safe equipment of storage systems, duplication of information and replacement of a lost copy with another in case of destruction of one of the copies (including as part of fail-safe). Ease of operation - automation (minimize human involvement if possible: both the user and the administrator). Fast implementation – simple installation and configuration of programs, quick training of users.

Acronis True Image backup program

Backup storage:

"Cloud" backup - recording backup data using "cloud" technology through the online services of special providers;

DVD or CD – backup data recording on CDs;

HDD - recording of backup data on the hard disk of the computer;

LAN – recording of backup data on any machine within the local network;

FTP – recording of backup data on FTP servers;

USB - Write backup data to any USB-compatible device (such as a flash card or external hard drive).

***7.*** Technically, /dev/null in Linux is a virtual device file. Programs and utilities can interact with it as with a real file: request data, as well as send any information to it. But anything you write to /dev/null is sent to the void and forgotten (in other words, sent to a black hole). To understand why this is useful, you must first have a basic understanding of standard output (stdout) and standard error (stderr).

For practical purposes, think of a garbage can, a bottomless pit, or outer space, where you can drop anything without being able to recover it (no matter how hard the guys at NASA try). But if I already have commands like rm, why would I want something new that I'm removing? Since the way the two black holes work is completely different: how would you replace the standard error output in a command in a shell script at runtime? This is where /dev/null occurs.

/dev/null is a virtual device file that, when written to it, sends everything to the void, and when reading from it, it reads zero. The true potential of /dev/null is shown very well in various bash scripts.

**Висновок**

Виконуючи лабораторну роботу №6, ми отримали практичні навички роботи з командною оболонкою Bash. Ознайомилися з базовими командами для архівування та стиснення даних. Ознайомилися з базовими діями при роботі з текстом у терміналі.