“Київський коледж зв’язку”

Циклова комісія Комп’ютерної інженерії

**ЗВІТ ПО ВИКОНАННЮ**

**ЛАБОРАТОРНОЇ РОБОТИ №5**

з дисципліни: «Операційні системи»

**Тема: «Команди Linux для архівування та стиснення даних»**

Виконали студенти

групи КСМ-93б

Скидан Д.О.

Логвинеко

Р.О.

Галка Р.О.\_\_\_\_\_\_\_\_

Перевірив викладач

Повхліб В.С. \_\_\_\_\_\_\_

Київ 2021

**Мета роботи:**

* Отримання практичних навиків роботи з командною оболонкою Bash.
* Знайомство з базовими діями при роботі з довідкою.
* Знайомство з базовими діями при роботі з файлами та каталогами.

**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:**

1.Прочитайте короткі теоретичні відомості до лабораторної роботи та зробіть невеличкий словник

базових англійських термінів з питань призначення команд та їх параметрів.

2. **tar** (англ. *tape archive* — стрічковий архів) — формат бітового потоку або файлу архіву, а також назва традиційної для Unix програми для роботи з такими архівами. Програма tar була стандартизована в POSIX.1-1998, а також пізніше в POSIX.1-2001. Спочатку програма tar використовувалася для створення архівів на магнітній стрічці, зараз же tar використовується для зберігання кількох файлів усередині одного файлу, для розповсюдження програмного забезпечення, а також за прямим призначенням — для створення архіву файлової системи. Однією з переваг формату tar при створенні архівів є те, що в архів записується інформація про структуру каталогів, про власника й групу окремих файлів, а також мітки часу файлів.



**xz** — комп'ютерна програма стиснення без втрат, що використовує алгоритм LZMA2, а також відповідний формат файлу

Архіватор xz міг бути задуманий як обрізана версія програми 7-Zip, який має свій власний формат, а не використовує її формат *7z*. Бо хоча 7-Zip, застосовуючи компресію LZMA2, і здатна стискати щільно (за рахунок швидкодії), та її формат архіву, зроблений переважно для Windows, не підтримує функціональності Unix (цьому формату не вистачає підтримки метаданих Unix-подібних файлових систем).

xz стискає один файл на вході, а не вміщає набір файлів в один архів. Таким чином, він стискає файл, який може бути й архівом, наприклад створеним такими Unix-програмами, як tar або cpio.

**Zip** - це часто використовувана функція стиснення. Вона портативна і проста у використанні. Ви навіть можете розпакувати в Windows файли, створені в Linux!

Утиліта Unzip за замовчуванням недоступна в більшості дистрибутивів Linux, але її легко встановити. З .zip можна домогтися рівня стиснення, аналогічного .tar.gz.

Нижче наведено кілька сценаріїв, в яких ви можете використовувати zip-файли:

Коли ви працюєте з системами на базі Windows і Unix і часто перемикаєтеся між ними. Архіватор не тільки стискає файли, але також є службовим програмою для створення пакетів файлів. Працює на декількох операційних системах;

Щоб заощадити трафік. Якщо у вас обмежена пропускна здатність, zip може використовуватися для передачі файлів між двома серверами;

Для швидкого перенесення файлів. Утиліта Zip зменшує розмір файлу, що скорочує час його перенесення;

Більш швидке завантаження і скачування директорій;

Економія місця на диску;

Розпакування захищених паролем архівів .zip

Коли потрібна хороша ступінь стиснення

Пам'ятайте, для того, щоб використовувати unzip в Linux, спочатку вам необхідно підключитися до сервера VPS по SSH.

**Утиліта bzip** призначена виконувати функцію стиснення. Головний плюс в порівнянні з іншими утилітами, що bzip надає можливість стиснення з великим відсотком, але поступається в швидкості виконання процесу самої архівації, що може негативно позначатися при роботі з об'ємними файлами. Сам принцип дії і призначення схожий на gzip, але використовує інший алгоритм стиснення, тим самим реалізує стиснення даних з високим ступенем.

Права доступу, власника, як і оригінальний файл, що дає в подальшому можливість зберегти ці атрибути при розпакуванні файлів. Також, якщо ви стискаєте однією командою кілька файлів не буде створюватися спільний архів (кожен файл буде окремо). Для виконання архівування можна спочатку через tar створити загальний архів і потім вже його відразу стиснути.

Основні опції схожі з gzip, але є і ряд відмінностей, наприклад відсутній ключ -r, або наявність параметрів:

-k - залишає оригінальний вихідний файл разом зі стисненим, без заміни;

-s - скоротити вживані ресурси процесами архівації, при цьому збільшується час виконання стиснення.

Подивитися всі опції і короткий їх опис можна стандартно через параметр -h.

**Утиліта gzip**

У командному рядку можна вказати відразу кілька імен файлів або шаблони імен файлів, просто відокремивши їх пробілами. Але в цьому випадку кожний із зазначених файлів буде заархівований окремо (загальний архів не створюється), як на прикладі нижче. Тут команда ls використовується для виведення списку файлів.

Основні опції утиліти gzip:

-d - для відновлення оригіналу файлу зі стислої версії;

-h - для отримання допомоги про спосіб використання утиліти, ключах та інше;

-N - залишає оригінальний вихідний файл разом зі стисненим (без заміни вихідного файла стислим), зберігає його часові мітки;

-l - отримання інформації про стиснутий файл, його новий розмір, відсоток стиснення, назву вихідного файлу і т.д.;

-r - виконати архівування рекурсивно повністю всієї директорії (всі підкаталоги і всі файли в цих підкаталогах);

-t - перевірка на коректність проходження процедури стиснення файлів, якщо файли пошкодилися, з’явиться повідомлення з помилкою (використовувати зв'язкою з ключем -v для виведення повідомлення з результатом проведення перевірки);

-v - показ додаткових відомостей про процес стиснення.

При стисненні, в якості аргументу, вводяться імена файлів, які необхідно стиснути, також додатково можна застосовувати опції (ключі). Зазначений файл замінюється стислою версією з розширенням .gz., Зберігаючи власника, режими, права доступу, час модифікації і мітки вихідного оригінального файлу. В результаті gzip видає менший відсоток стиснення, ніж його інші побратими (наприклад, bzip2), використовуючи досить старий алгоритм, але при цьому є невибагливим до ресурсів системи і, як варіант, підходить для машин з невеликими потужностями.

2.Алгоритм Д. Хаффмана та алгоритм Шеннона-Фано,Алгоритм RLE.

3.Створіть таблицю команд вивчених у п.2 ходу роботи у наступному вигляді:

|  |  |
| --- | --- |
| Назва команди | Її призначення та функціональність |
| tar | Створює архівні файли за допомогою стиснення та без стиснення |
| xz | Стискає та розпаковує файли в архівний файл |
| zip i unzip | Використовується для стиснення та розпакування архівних файлів |
| bzip2 і gzip | Стискає та розпаковує файли в архівний файл |

**1.СТИСНЕННЯ ДАНИХ**

Стиснення даних — це процедура перекодування даних з метою зменшення їх обсягу.

Для цього використовують різні методи, зокрема заміну часто повторюваних послідовностей коротшим значенням (кодом). Такими послідовностями моясуть бути масиви пікселів однакового кольору в растровому зображенні, слова в тексті тощо. Наприклад, будь-яка тисяча символів текст}' українською мовою містить літер «о» чи «і» на порядок більше, ніж літер «ф». Припустимо, що кожний символ текст}' кодується двійковою послідовністю певної довжини. Зрозуміло, що використовуючи коротші коди для поширених символів, загальна довжина закодованого тексту буде меншою.

Буває стиснення без втрат, коли вихідні дані можна відновити без викривлень, і з втратами, коли дані відновлюються лише з викривленнями, які проте моясуть бути непомітними для людини. Стиснення без втрат застосовують до текстової інформації, а стиснення із втратами — до звукової, графічної та відеоіп-формації, і саме воно дає змогу значно зменшити обсяги даних. В останньому разі па основі спеціальних досліджень визначають, якою інформацією можна поясертвувати. Відомо, наприклад, що

зір людини дуже чутливий до зміни яскравості й значно менше — до колірного відтінку, тому під час стиснення фотографій інформацію про колір частини пікселів можна видалити. їхній колір під час розпакування зображення визначатиметься за кольором сусідніх пікселів.

**АРХІВУВАННЯ**

Для стиснення даних можна використовз'вати спеціальні програми — архіватори. Файли, створені цими програмами, називають архівами.

Архіватор записує вміст файлів у архів і долучає до нього інформацію про імена файлів та обсяги їхніх оригіналів. Архіватор може створити один або кілька архівів з одного чи кількох файлів.

Архіватор — це програма, що стискає та об’єднує файли.

Процедура стиснення потребує певного часу. Що більший ступінь стиснення, то менша швидкість цього процесу, і навпаки.

Програми-архіватори можуть працювати з архівними файлами різних типів. Найлонулярпіші типи архівів — ARJ, ZIP, BAR, 7-Z. Найчастіше застосовують ZIP і RAR. Цю практику запроваджено завдяки програмам WinZip та WinRAR, які мають зручний інтерфейс і багато додаткових функцій, а також забезпечз'їоть досить високий ступінь стиснення.

2. WinZIP, WinRAR, 7-zip, Power-Archiver, WinAce, Ark, Btar, AndroZip, FreeArc

Дуже зручні архіваториЮ які працюють методом архівування дадаючи додаткову безпеку у вигляді пароля для файлів. Моїм фаворотим завжи був та залишається WinRAR, так як він найзручнішим та найпростішим на мою думку.

3. Найкращим на мою думку є утілита tar. **Tar** - це стандартна утиліта, за допомогою якої виконується архівування файлів Linux. Поступово з невеликої програми архівації вона перетворилася на потужний інструмент, що підтримує роботу з багатьма варіантами архівів і алгоритмами стиснення. Програма підтримує велику кількість параметрів.

Ще дуже цікавим є shar. **Shar** дозволяє створювати архіви. По суті, це shell, і для розпакування йому необхідна оболонка Bash або інша сумісна з Bourne Shell. У Shar є кілька переваг, але також він потенційно небезпечний, так як архів являє собою виконуваний файл.

**ar** - утиліта для створення і управління архівами. В основному використовується для архівації статичних бібліотек, але може бути використана для створення будь-яких архівів. Раніше використовувалася досить часто але була витіснена утилітою tar. Зараз використовується тільки для створення та оновлення файлів статичних бібліотек.

**cpio** - означає Copy in and out (скопіювати введення і виведення). Це ще один стандартний архіватор для Linux. Активно використовується в менеджері пакетів Red Hat, а також для створення initramfs. Архівація в Linux для звичайних файлів за допомогою цієї програми не застосовується.

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

**мобільному телефоні.**

**Архівування** - це процес стиснення файлів з метою зменшення їх розміру без втрати вмісту або будь-якого погіршення його якості. На сьогодні,архівування і стискання файлів є дуже актуальними. Існує безліч програм-архіваторів на комп’ютері та телефоні,але найпопулярнішими є: WinRAR та 7-Zip.

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

**Стиснення даних** — це процедура перекодування даних, яка проводиться з метою зменшення їхнього обсягу, розміру, об'єму. Стиснення базується на усуненні надлишку інформації, яка міститься у вихідних даних. Стиснення даних, які не мають властивості надлишку (наприклад випадковий сигнал чи шум), неможливе. Також неможливо стиснути зашифровану інформацію.

**Стиснення без втрат**— метод стиснення даних, при використанні якого закодована інформація може бути відновлена з точністю до біта. Для кожного з типів цифрової інформації, як правило, існують свої алгоритми стиску без втрат. Стиснення без втрат використовується при обробці та збереженні комп'ютерних програм і даних.

**Стиснення з втратами** **інформації** — метод стиснення даних, при якому розпакований файл відрізняється від оригіналу, проте може бути корисним для використання. Стиснення із втратами найчастіше використовується для мультимедіа даних (аудіо, відео, зображення), особливо для потокової передачі даних та телефонії. В цьому контексті такі методи часто називаються кодеками.

До алгоритмів стиснення з втратою інформації відносяться такі алгоритми як **JPEG** (використовуються при стисненні фотозображень) і **MPEG** (використовуються при стисненні відео і аудіо).

**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 даних*** – алгоритмічне перетворення даних, вироблене з метою зменшення їх обсягу. Застосовується для більш раціонального використання пристроїв зберігання і передачі даних. Зворотна процедура називається ***відновленням даних*** (розпакуванням, декомпресією).

Стиснення засноване на усуненні ***надлишків***, що міститься у вихідних даних. Найпростішим прикладом надлишків є ***повторення в тексті фрагментів*** (наприклад, слів природної або машинної мови). Подібний надлишок зазвичай усувається заміною повторюваної послідовності посиланням на вже закодований фрагмент із зазначенням його довжини. Інший вид надшку пов’язаний з тим, що деякі значення в даних, що стискуються зустрічаються частіше інших. Скорочення обсягу даних досягається за рахунок заміни часто повторюваних даних короткими кодовими словами, а рідких – довгими (ентропійне кодування). Стиснення даних, що не володіють властивістю надлишку (наприклад, випадковий сигнал чи шум, зашифровані повідомлення), ***принципово неможливе без втрат***.

В залежності від того, в ***якому об’єкті розміщені дані***, що піддаються стисненню, розрізнюють:

* архівацію файлів;
* архівацію папок;
* архівацію дисків.