“Київський фаховий 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”**

Виконала студентка

групи РПЗ-93а

Команда 6: Колосюк Д.С., Мількевич В.П.

Перевірив викладач

Сушанова В.С.

Київ 2022

**Мета роботи:**

1. Отримання практичних навиків роботи з командною оболонкою Bash.

2. Знайомство з базовими командами для архівування та стиснення даних.

**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 Мількевич В.***

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

|  |  |
| --- | --- |
| instance | екземпляр |
| distribution | розповсюдження |
| opportunities | можливості |
| compression techniques | техніки стиснення |
| reconstruct the original file | відновити вихідний файл |
| placeholder | заповнювач |
| implements | знаряддя |
| measurement of memory | вимірювання пам'яті |

***Готувала матеріал студентка Колосюк Д.***

1. На базі розглянутого матеріалу дайте відповіді на наступні питання:

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

Gzip — це утиліта для стиснення та відновлення файлів, яка використовує алгоритм Deflate.

Згідно з традицією програмування UNIX, gzip виконує лише одну функцію: стискаючи і розпаковуючи один файл, він не може упакувати кілька файлів в один архів. У разі стиснення до вихідного розширення файлу додається суфікс .gz. Щоб упакувати кілька файлів, вони зазвичай спочатку архівуються в один файл за допомогою утиліти tar, а потім цей файл стискається за допомогою gzip. Таким чином, стиснені архіви зазвичай мають подвійне розширення .tar.gz.

У найпростішому випадку він викликається у такому форматі:

[користувач] $ gzip файл

Tar є єдиним архівом колекції файлів і папок, але на відміну від zip, tar не стискається.

Створіть файл tar, що містить усі ваші зображення, зберігаючи структуру папок, використовуючи таку команду: tar -cvf photos ~ / photos

Bzip2 — безкоштовна утиліта командного рядка з відкритим кодом (а також алгоритм) для стиснення даних.

bzip2 виконує лише одну функцію: стискає та розпаковує один файл. Розширення .bz2 додається до імені файлу за замовчуванням. Щоб упакувати кілька файлів, вони зазвичай спочатку архівуються в один файл за допомогою утиліти tar, а потім стискаються за допомогою bzip2. Такі архіви зазвичай мають розширення .tar.bz2.

bzip2 стискає більшість файлів ефективніше, але повільніше, ніж традиційні gzip або zip. \

bzip2 можна використовувати як у поєднанні з tar, так і окремо: файл bzip2 для стиснення та bzip2 -d file.bz2 для розпакування

Xz — це комп’ютерна програма стиснення без втрат, яка використовує алгоритм LZMA2, а також відповідний формат файлу.

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

Застібка на блискавці

У Linux легко використовувати архіви Zip та Unzip.

Щоб перевірити встановлення, введіть "sudo apt-get install zip unzip"

Після встановлення ми можемо використовувати zip для створення архівів і розпакувати, щоб розширити їх до оригіналів.

2.2 Наведіть три приклади реалізації архівування та стискання даних різними командами.

bzip2 стискає більшість файлів ефективніше, але повільніше, ніж традиційні gzip або zip

gzip Щоб упакувати декілька файлів, вони зазвичай спочатку архівуються в один файл за допомогою утиліти tar, а потім цей файл стискається gzip

zip Найпростіший спосіб використовувати zip - це вказати йому ім'я zip-архіву, який ви хочете створити, а потім явно назвати кожен файл, який має бути до нього включений.

**Контрольні запитання**

***Готувала матеріал студентка Мількевич В.***

1. Надайте порівняльну характеристику процесам стискання та архівування.

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

Архівування інформації - стиснення одного або кількох файлів з метою зменшення обсягу даних при зберіганні на носіях або при передачі даних по каналах зв'язку, включаючи Інтернет, і розміщення стиснутих файлів в одному архівному файлі.

Коли методи стиснення даних застосовуються до готових файлів, замість «стиснення даних» часто використовується термін «стиснення даних», стиснена версія даних називається архівом, а програмне забезпечення, що реалізує методи стиснення, — архіваторами.

Залежно від об'єкта, в якому знаходяться дані, що підлягають стиснення, розрізняють:

1. Стиснення (архівування) файлів: використовується для зменшення розміру файлів під час підготовки до передачі по каналах зв'язку або для транспортування на зовнішніх носіїв малої ємності;

2. Стиснення (архівування) папок: використовується як засіб зменшення обсягу папок перед тривалим зберіганням, наприклад, при створенні резервних копій;

3. Стиснення (стиск) дисків: використовується для підвищення ефективності використання дискового простору шляхом стиснення даних під час запису їх на носій (зазвичай за допомогою операційної системи).

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.

Ark — це інструмент для архівації KDE, що входить до пакету Kdeutils.

File Roller — це комп'ютерний архіватор для робочого столу GNOME. File Roller — це графічна оболонка, яка забезпечує єдиний графічний інтерфейс для консольних програм архівування. Поширюється під загальною суспільною ліцензією GNU.

ARJ — архіватор файлів. Розроблено Робертом К. Юнгом. (Походження назви ARJ: архіватор Роберт Юнг). Версія ARJ 1.00 була випущена в лютому 1991 року за умовно-безкоштовною ліцензією. Стиснення ARJ подібне до PKZIP 1.02. Існує також версія ARJ з відкритим кодом, доступна в більш ніж десяти операційних системах, включаючи різні версії UNIX і Linux. Також існує версія російського NLV, що дозволяє захищати архіви за допомогою алгоритму шифрування ГОСТ.

bzip2 — безкоштовна утиліта командного рядка з відкритим кодом (а також алгоритм) для стиснення даних. Розроблено та вперше опубліковано Джуліаном Сьюардом у липні 1996 року (версія 0.15). Стабільність і популярність компресора з роками зростали, і версія 1.0 була випущена наприкінці 2000 року. Згідно з традиціями програмування UNIX, bzip2 виконує лише одну функцію: стискає та розпаковує один файл.

gnochive - графічний інтерфейс для архіваторів GNOME.

Easibox - консольна утиліта для створення різноманітних архівів.

Lha - File Archiver - використовує метод стиснення LZW (файли .lzh).

kArchiver - Розширений архіватор, розуміє tar, tgz, tar.gz, bz2, zip, rar та інші.

mediArchiver - медіаархіватор для зберігання всього багатства медіафайлів.

Порівняйте алгоритми стиснення, які використовуються в командах (програмах), що використовуються в Linux.

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

Використовуються два типи методів – алгоритми стиснення без втрат і без втрат. Перший дозволяє відновити файл до початкового стану без втрати жодного біта інформації в стиснутому файлі. Другий — це типовий підхід до виконуваних файлів, тексту та електронних таблиць, де втрата слів або цифр змінить інформацію.

Алгоритм Д. Хаффмана та алгоритм Шеннона-Фано передбачають перекодування даних. Символи, які частіше зустрічаються в повідомленні, перекодуються меншою кількістю символів і рідше більшою кількістю символів. Наприклад, щоб закодувати в Windows-1251 попереднє речення, яке містить 114 символів, включаючи пробіли, потрібно використовувати 1 байт даних для кожного символу, або 114 байт. Якщо порахувати, скільки разів символ зустрічається в реченні, то вийде, що загальна кількість символів у реченні – 28, найпоширеніший символ «пробіл» – 14 разів, символ «і» – 10, «о» - 9, "в" - 7, "в" - 6 і т. д. Принаймні один раз є символи "щ", "б", "в", "-", ".". Якщо ви використовуєте алгоритм стиснення даних для двійкового коду (0 і 1) і позначаєте найпоширеніші символи одним символом (1 біт): символ «пробіл» - 0, а символ «і» - 1, за яким слідує частота повторення - два символи. 2 біти): символ "o" - 00, символ "c" - 01 і т. д. Для менш поширених символів будуть використані чотири двійкові цифри або 4 біти. В результаті загальна довжина коду закодованого таким чином пропозиції становитиме 90 біт або 11 байт і 2 біти. Як бачимо, загальна кількість коду речення зменшиться в 10 разів. Звичайно, це досить спрощена схема кодування, не забувайте, що стиснутий таким чином файл повинен містити таблицю кодування та додаткову службову інформацію. Тому при кодуванні дуже маленьких файлів може бути не зменшення, а, навпаки, збільшення розміру файлу. Для файлів з великою кількістю текстових або цифрових даних цей алгоритм забезпечує високий ступінь стиснення.

Алгоритм кодування Run-Length (RLE) частіше використовується для кодування графічних даних. Цей метод передбачає заміну послідовності подібних даних її скороченим описом. Наприклад, зображення логотипу Apple (рис. 2.49) використовує лише два кольори — чорний і білий. Якщо позначити чорний цифрою «1», а білий — «0», то код першого рядка зображення повинен складатися з 48 цифр «0» (48 біт).

Але їх можна замінити описом із зазначенням коду символу та кількості повторів - 0 48. Коду опису буде виділено 9 біт: 1 біт до цифри "0" і чотири біти до двох цифр кількості повторів. . Другий ряд пікселів буде кодуватися таким же чином, а третій: 0 29 1 4 0 15-23 біти. Для кольорових зображень з великою кількістю монохромних фрагментів цей алгоритм забезпечить високу ефективність стиснення.

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

Сьогодні ми дуже часто зустрічаємо архівні файли на всіх платформах: Windows, Mac або Linux

Найпопулярнішою утилітою для архівування Linux є tar. Використовується практично скрізь, для архівування вихідного коду, пакування пакетів. Для стиснення використовується багато інших програм, залежно від алгоритму стиснення, наприклад zip, bz, xz, lzma тощо.

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

Існує два типи архіваторів:

Архіватори, що працюють у режимі командного рядка. Суть роботи з цими архіваторами полягає у введенні команд, які відповідають імені виконуваного файлу програми. Дуже важливо не забути про налаштування відповідних параметрів.

Архіватори оболонки – це програми зі зручним інтерфейсом, який полегшує будь-якому користувачеві виконання операцій з архівуванням.

Стиснення даних – це процес перекодування даних для зменшення розміру файлу.

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

Використовується для більш ефективного використання пристроїв зберігання та передачі даних. Стиснення засноване на усуненні надлишку, що міститься в вихідних даних. Найпростішим прикладом надмірностей є повторення фрагментів у тексті. Вирішуються такі проблеми, як дублювання даних, передача даних та робота з загальними документами.

**Висновки**

***Готувала матеріал студентка Колосюк Д***

Під час виконання лабораторних робіт ми отримали практичні навички роботи з командною оболонкою Bash. А також вивчили основні команди для архівування та стиснення даних.