



Linux Essentials

Arquivamento e Compactação

Arquivamento e Compactação

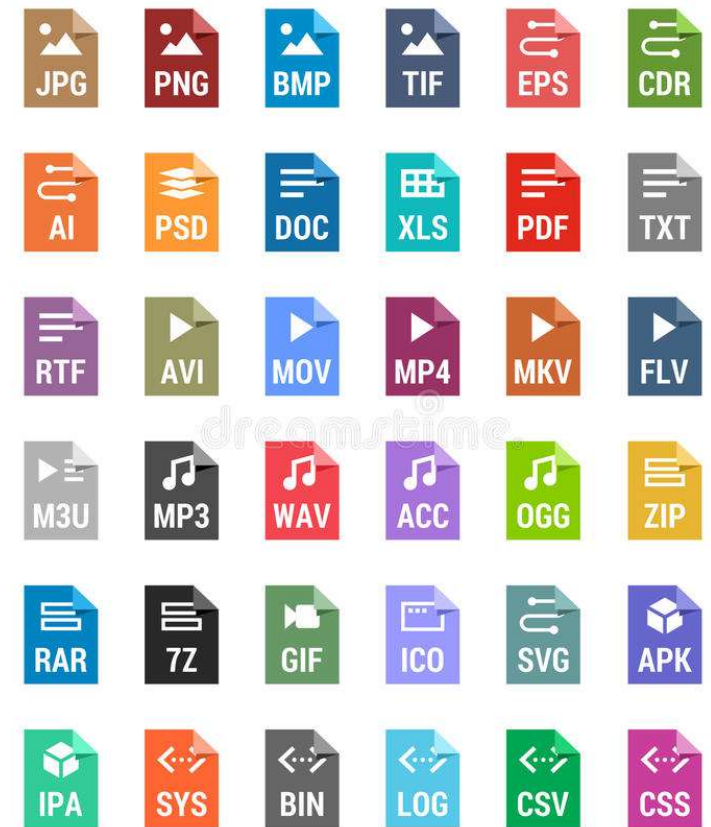
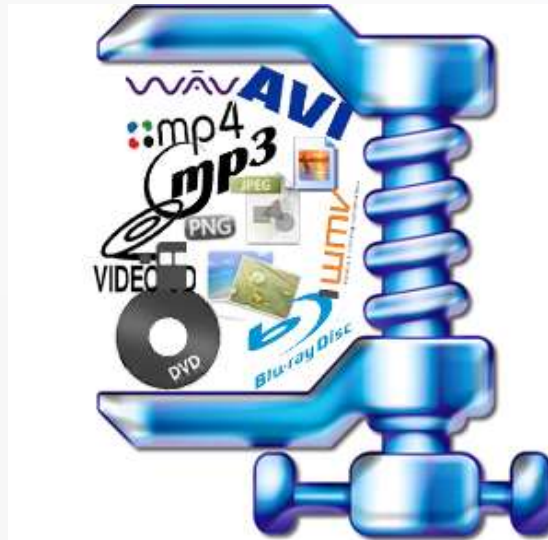
- Iremos abordar como gerenciar arquivos compactados na linha de comando.
- O arquivamento de arquivos é usado quando um ou mais arquivos precisam ser transmitidos ou armazenados da maneira mais eficiente possível.
- Existem dois aspectos fundamentais que este capítulo explora:
 - Arquivamento: combina vários arquivos em um, o que elimina a sobrecarga em arquivos individuais e facilita a transmissão.
 - Compactação: Reduz os arquivos, removendo informações redundantes.

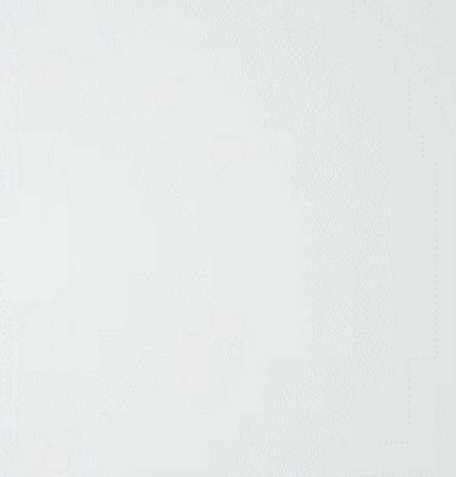
Compactação

- A compactação reduz a quantidade de dados necessários para armazenar ou transmitir um arquivo enquanto o armazena de forma que o arquivo possa ser restaurado.
- O algoritmo de compactação é um procedimento que o computador usa para codificar o arquivo original e, como resultado, torná-lo menor.
- Quando se fala em compressão, existem dois tipos:
 - Sem perdas: nenhuma informação é removida do arquivo.
 - Com perdas: as informações podem ser removidas do arquivo.

Compactação de Dados

- É a transformação de um conjunto de símbolos em outro com um tamanho reduzido.
- Simplificando é diminuir o tamanho de um determinado arquivo.





Dados para o computador

- Para um computador, os Dados ou arquivos, são praticamente a mesma coisa, sequências de números binários.
- O tratamento deles são feitos pelo programa interpretador.
- Exemplo:
 - Um arquivo .doc (arquivo de texto com formatação) por padrão no Windows é interpretado, pelo programa de edição de texto Word.



Exemplo_arquivo_Word.doc



Exemplo_arquivo_Word.docx

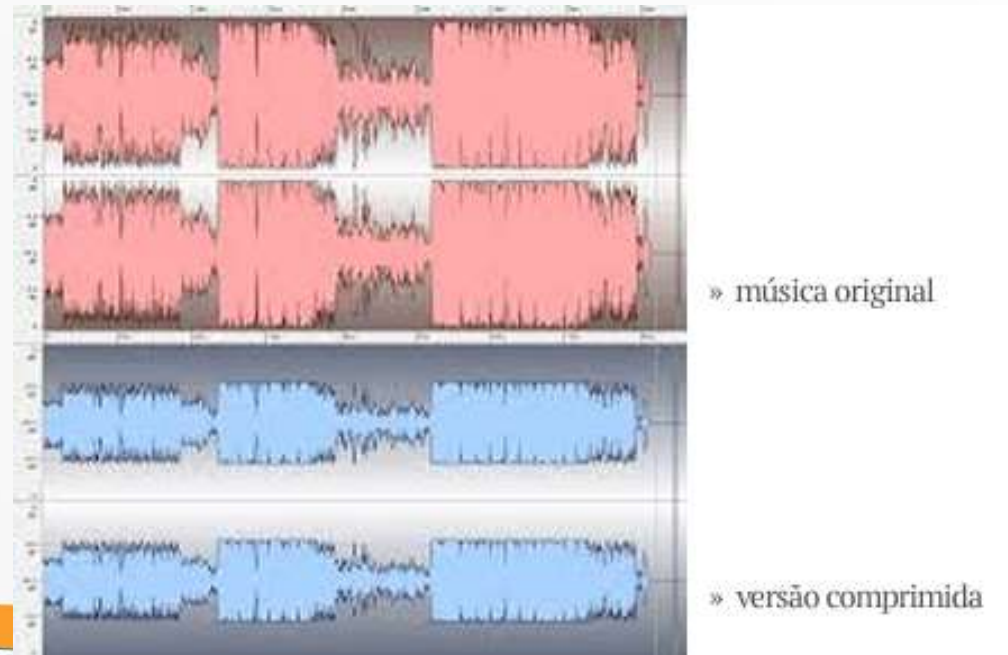
Compactação de Dados

Existem vários tipos de algoritmos para compactação.

- E para cada tipo de arquivo também existe uma técnica.

- Áudio
- Vídeo
- Imagem

- Exemplo de um algoritmo de compactação bem conhecido é MP3 (MPEG3).



Programas Compactadores

- Mais conhecidos são:
 - Winzip
 - Winrar
 - Cdex
 - Format Factory
 - 7zip
 - Gzip
 - Xz
 - bzip
- O resultado é o arquivo compactado com uma extensão diferente.
 - Zip
 - Rar
 - Mp3
 - Jpg
 - Rmvb
 - Mp4

GZip

- A ferramenta Gzip é a mais popular e “clássica” para compressão de arquivos. O Gzip mantém o nome do arquivo original gerando uma extensão “.gz” do arquivo compactado. É uma ferramenta GNU, presente nativamente, na maioria das distribuições Linux.
- O Gzip usa um algoritmo de compressão conhecido como “DEFLATE, muito usado, também, em outras tecnologias imagem PNG, protocolo web HTTP e o protocolo de conexão remota segura SSH.

Compactando com Gzip

- O Linux oferece diversas ferramentas para compactar arquivos, a mais comum é o gzip. No exemplo abaixo antes e depois da compactação:

```
sysadmin@localhost:~/Documents$ ls -l longfile*  
-rw-r--r-- 1 sysadmin sysadmin 66540 Dec 20 2017 longfile.txt  
sysadmin@localhost:~/Documents$ gzip longfile.txt  
sysadmin@localhost:~/Documents$ ls -l longfile*  
-rw-r--r-- 1 sysadmin sysadmin 341 Dec 20 2017 longfile.txt.gz
```

- O tamanho original do arquivo chamado longfile.txt é 66.540 bytes.
- O arquivo é compactado o comando **gzip** com o nome do arquivo como argumento.
- Após a conclusão do comando, o arquivo original desaparece e uma versão compactada com uma extensão de arquivo .gz é deixada em seu lugar, o tamanho agora é de 341 bytes.

Descompactando com Gzip

- O comando **gzip** fornecerá essas informações usando a opção **-l**.

```
sysadmin@localhost:~/Documents$ gzip -l longfile.txt.gz
      compressed      uncompressed   ratio uncompressed_name
          341             66540   99.5% longfile.txt
```

- Os arquivos compactados podem ser restaurados à sua forma original (descompactação) usando o comando **gunzip** ou o comando **gzip -d**.
- Depois que o **gunzip** funcionar, o arquivo *longfile.txt* será restaurado ao tamanho e nome originais:

```
sysadmin@localhost:~/Documents$ gunzip longfile.txt.gz
sysadmin@localhost:~/Documents$ ls -l longfile*
-rw-r--r-- 1 sysadmin sysadmin 66540 Dec 20 2017 longfile.txt
```

Outros Compactadores

- Existem outros comandos que operam virtualmente de forma idêntica ao gzip e gunzip. Estes incluem **bzip2** e **bunzip2**, bem como **xz** e **unxz**.
- O comando gzip usa o algoritmo de compactação de dados Lempel-Ziv, enquanto os utilitários bzip usam um algoritmo de compactação diferente chamado classificação de blocos Burrows-Wheeler, que pode compactar arquivos menores que gzip às custas de mais tempo de CPU. Esses arquivos podem ser reconhecidos porque possuem uma extensão .bz ou .bz2 em vez de uma extensão .gz.
- As ferramentas xz e unzip são funcionalmente semelhantes ao gzip e gunzip, pois usam o algoritmo de cadeia Lempel-Ziv-Markov (LZMA), que pode resultar em tempos de CPU de descompactação mais baixos que estão no mesmo nível do gzip, ao mesmo tempo que fornece as melhores taxas de compactação normalmente associadas. com as ferramentas bzip2. Os arquivos compactados com o comando xz usam a extensão .xz.

Arquivando com o TAR

- Arquivar é quando você empacota muitos arquivos ou diretórios em um arquivo.
- O utilitário tradicional do UNIX para arquivar arquivos é chamado tar, que é uma forma abreviada de *TApe aRchive*.
- O **Tar** tem três modos com os quais é útil se familiarizar:
 - **Criar:** Crie um novo arquivo a partir de uma série de arquivos.
 - **Extrair:** Extraia um ou mais arquivos de um arquivo.
 - **Lista:** Mostra o conteúdo do arquivo sem extrair.

Arquivamento de arquivos – Modo criar

```
tar -c [-f ARCHIVE] [OPTIONS] [FILE...]
```

- Criar um arquivo com o comando **tar** requer duas opções nomeadas:

-c	Criar um arquivo.
-f ARQUIVO	Usar arquivo compactado. O argumento ARQUIVO será o nome do arquivo resultante.

- O exemplo a seguir mostra um arquivo tar, também chamado de tarball, sendo criado a partir de vários arquivos:

```
sysadmin@localhost:~/Documents$ tar -cf alpha_files.tar alpha*
sysadmin@localhost:~/Documents$ ls -l alpha_files.tar
-rw-rw-r-- 1 sysadmin sysadmin 10240 Oct 31 17:07 alpha_files.tar
```

Arquivamento de arquivos – Modo listar

```
tar -t [-f ARCHIVE] [OPTIONS]
```

- Dado um arquivo **tar**, compactado ou não, você pode ver o que ele contém usando a opção **-t**. O próximo exemplo usa três opções:

-t	Lista os arquivos dentro do arquivo
-j	Descompacta o arquivo
-f ARQUIVO	Executa no arquivo determinado

- O exemplo a seguir lista o conteúdo do arquivo : alpha_files.tar.

```
sysadmin@localhost:~/Documents$ tar -t -f alpha_files.tar
```

Arquivamento de arquivos – Modo extrair

```
tar -x [-f ARCHIVE] [OPTIONS]
```

- Você pode extrair o arquivo com a opção **-x**. O exemplo a seguir usa o padrão semelhante aos outros modos:

-x	Extraí os arquivos dentro do arquivo
-j	Descompacta o arquivo
-f ARQUIVO	Executa no arquivo determinado

- O exemplo a seguir lista o conteúdo do arquivo : alpha_files.tar.

```
sysadmin@localhost:~/Documents$ cp alpha_files.tar School
sysadmin@localhost:~/Documents$ tar -x -f alpha_files.tar
```

Utilitário ZIP

- É o utilitário de arquivamento padrão da Microsoft. ZIP não é tão predominante no Linux, mas é bem suportado pelos comandos **zip** e **unzip**.
- O modo padrão do **zip** é adicionar arquivos a um arquivo e compactá-lo.

```
zip [OPTIONS] [zipfile [file...]]
```

- O exemplo a seguir mostra um arquivo compactado chamado alpha_files.zip sendo criado:

```
sysadmin@localhost:~/Documents$ zip alpha_files.zip alpha*
```

Zipando um Diretório

- O comando zip não será recursivo em subdiretórios por padrão (o tar faz isso), então você deve usar a opção **-r** para indicar que a recursão deve ser usada.

```
sysadmin@localhost:~/Documents$ zip -r School.zip School
```

Descompactando arquivo Zip

- Descompactar o arquivo school.zip.
- A opção **-l** *list* do comando **unzip** lista arquivos em arquivos .zip:

```
sysadmin@localhost:~/Documents$ unzip -l School.zip
```


Atividade Prática

- Na máquina virtual Linux Ubuntu.
- Baixar o arquivo do site <https://br.wordpress.org/download/>
- `# wget https://br.wordpress.org/latest-pt_BR.tar.gz`
- Após download, descompactar com o comando:
- `#tar xzvf latest-pt_BR.tar.gz`

Resumo

- **bunzip2**
- Descomprime um arquivo comprimido bzip2.
- **bzcat**
- Exibe o conteúdo de um arquivo comprimido bzip.
- **bzip2**
- Comprime arquivos usando o algoritmo e o formato bzip2.
- **gunzip**
- Descomprime um arquivo comprimido gzip.
- **gzip**
- Comprime arquivos usando o algoritmo e o formato gzip.

- **unzip**
- Descomprime e extrai conteúdo de um arquivo ZIP.
- **xz**
- Comprime arquivos usando o algoritmo e o formato xz.
- **zcat**
- Exibe o conteúdo de um arquivo comprimido gzip.
- **zip**
- Cria e comprime pacotes ZIP.
- **tar**
- Cria, atualiza, lista e extrai pacotes tar.