

Segundo Proyecto

El objetivo de este proyecto consiste en programar un empacador de archivos. Este es el tipo de funcionalidad que provee el comando *tar* en ambientes UNIX.

El programa *tar*, es usado para almacenar múltiples archivos en un solo archivo. Dentro de los entornos Unix *tar* aparece como un comando que puede ser ejecutada desde la línea de comandos de una consola de texto o desde un simple terminal. El formato del comando *tar* es, comúnmente

```
tar <opciones> <archivoSalida> <archivo1> <archivo2> ... <archivoN>
```

donde *<archivoSalida>* es el archivo resultado y *<archivo1>*, *<archivo2>*, etc; son los diferentes archivos que serán "empaquetados" en *<archivoSalida>*.

Las opciones más comunes son las siguientes:

- -c, --create : crea un nuevo archivo
- -x, --extract : extraer de un archivo
- -t, --list: listar los contenidos de un archivo
- --delete: borrar desde un archivo
- -u, --update: actualiza el contenido del archivo
- -v, --verbose: ver un reporte de las acciones a medida que se van realizando
- -f, --file: empaclar contenidos de archivo, si no está presente asume la entrada estándar.
- -r, --append: agrega contenido a un archivo
- -p, --pack: desfragmenta el contenido del archivo (no presente en tar)

Ejemplos

1. Si queremos empaclar un archivo llamado "index.html" y guardar los datos en "html-paq.tar", lo haríamos con la instrucción

```
tar -cvf html-paq.tar index.html
```

2. Si queremos desempaquetar todo el contenido de un archivo llamado xxx.tar podemos utilizar un comando como este

```
tar -xvf xxx.tar
```

3. Para archivar el contenido de tres archivos doc1.txt, doc2.txt y data.dat

```
tar -cvf foo.tar doc1.txt doc2.txt data.dat
```

4. Si ahora se desea eliminar el contenido del archivo data.dat se ejecutaría

```
tar --delete -vf foo.tar data.dat
```

5. Para agregar ahora un nuevo archivo test.doc a foo.tar se ejecutaría

```
tar -rvf foo.tar test.doc
```

Los contenidos se desempacarán en el directorio actual.

Programación

Se deberá programar el comando *star* ("simple tar", NO "estrella"), de tal forma que acepte los comandos básicos mostrados anteriormente. Para desarrollar su programa usted debe tomar en cuenta los siguientes aspectos:

- El archivo debe organizarse internamente utilizando "memoria contigua". Esto significa que los archivos a agregar se almacenan (al principio) seguidos en el archivo empacado. Note que debe llevar control del byte en donde empieza cada archivo y el byte en donde termina.
- Al crear un archivo empacado este se crea del tamaño necesario para almacenar los archivos agregados. Cuando se borra algún contenido, el archivo empacado no cambia de tamaño sino que se lleva registro de los espacios liberados. Si posteriormente se agrega un nuevo contenido entonces se reutiliza el espacio libre. Si aún así el nuevo contenido no cabe, se hace crecer el archivo empacado. No debe utilizar ningún archivo auxiliar para hacer crecer el archivo.
- Tome en cuenta que un archivo que se agrega puede ya existir en el archivo empacado. Es decir, lo que se desea hacer es actualizar su contenido. Para esto existe la opción update (-u) que sobrescribirá el contenido de un archivo.
- Se debe llevar un control de los espacios libres, para asignar los huecos libres se utilizará la técnica del primer ajuste o el siguiente ajuste.
- Este programa no utilizará los derechos de acceso, que normalmente almacenaría un archivo empacado *tar* en ambiente UNIX.
- La opción de desfragmentación (-p) no es estándar (no está presente en tar) y lo que hace es desfragmentar el contenido almacenado en el archivo empacado y liberar cualquier espacio sin utilizar. Es decir, con este comando se liberarán todos los bloques libres y el tamaño del archivo empacado se ajustará al contenido real existente. Note que no se debe utilizar un archivo temporal para realizar esta función, toda la desfragmentación se debe realizar sobre el contenido del mismo archivo.
- La opción -v muestra información sobre la operación ejecutada. Se puede aplicar dos veces -vv para ver información adicional.

```
tar -rvf foo.tar test.doc
```

Documentación

Se deberá generar una documentación formal, en formato pdf, en donde se describan las diferentes etapas del desarrollo del proyecto, las decisiones de diseño que se tomaron, los mecanismos de programación utilizados, y los resultados de las diferentes pruebas al programa.

Dicha documentación deberá incluir al menos las siguientes secciones:

- Introducción
- Descripción del problema (este enunciado)
- Definición de estructuras de datos
- Descripción detallada y explicación de los componentes principales del programa:
 - Mecanismo de acceso a archivos
 - Estructura de directorios internos
 - Estrategia de administración de espacios libres
 - Procedimiento de desfragmentación del archivo
- Análisis de resultados de pruebas (funcionamiento)
- Conclusiones

Puede agregar diagramas a la documentación que le ayuden a ilustrar las diferentes estructuras de datos y algoritmos principales.

Consideraciones generales

- El proyecto debe ser ejecutado en grupos de a lo más dos personas. Se puede desarrollar en forma individual pero bajo ninguna circunstancia se aceptarán grupos de tres o más personas.
- No se permite la copia de código entre grupos de estudiantes, o entregar código realizado por estudiantes los semestres anteriores, tampoco es permitido utilizar clases o librerías adicionales (desarrolladas por terceros) para simplificar el manejo de los bloques en disco, bloques libres, empaquetar datos, o desfragmentación.
- El proyecto debe ser resuelto mediante lenguaje C y ambiente Unix.
- Puede buscar en Internet mayor información sobre el uso de las diferentes opciones del comando tar, sin embargo, note que existen muchas otras opciones que no serán implementadas en este proyecto.