

¿Requieres de una instalación o configuración de Linux o sus servicios?

¿Un desarrollo WEB empresarial a la medida?

¿Un curso o capacitación a la medida?

Revisa el sitio de **SERVICIOS** ([index.php?cont=servicios](http://www.linuxtotal.com.mx/index.php?cont=servicios)) de LinuxTotal

LINUXTOTAL.COM.MX - Información y servicios en Linux y Open Source

URL: http://www.linuxtotal.com.mx/index.php?cont=info_admon_017

Aprende wget con ejemplos

Copyright © 2005-2017 LinuxTotal.com.mx

Se concede permiso para copiar, distribuir y/o modificar este documento siempre y cuando se cite al autor y la fuente de [linuxtotal.com.mx](http://www.linuxtotal.com.mx) y según los términos de la GNU Free Documentation License (<http://www.gnu.org/licenses/translations.html>), Versión 1.2 o cualquiera posterior publicada por la Free Software Foundation.

Autor: Sergio González D. (sergio.gonzalez.duran@gmail.com)

Hay decenas de aplicaciones para descargar archivos, la mayoría basadas en interfaces Web y de escritorio, y para todos los sistemas operativos. Sin embargo en Linux (y también hay versión para Windows) existe el poderoso descargador de archivos **wget**. Esta considerado como el descargador (downloader) más potente que existe, soporta http, https y ftp. En este artículo, a través de varios ejemplos aprenderás sus usos y opciones más comunes.

Ejemplos básicos

```
wget http://ejemplo.com/programa.tar.gz
```

La manera más simple de usarlo, descarga el archivo indicado.

```
wget http://ejemplo.com/programa.tar.gz ftp://otrositio.com/descargas/video.mpg
```

Es posible indicar más de una descarga a la vez, incluso con distintos protocolos.

```
wget http://ejemplo.com/*.pdf
```

Otra manera de descargar varios archivos, con extensión similar.

```
(creamos una lista en archivos.txt que serán descargados)
http://ejemplo.com/programa.tar.gz
http://ejemplo.com/rpm/paquete.rpm
ftp://otrositio.com/descargas/distro.iso

(descargamos todos indicando el archivo)
wget -i archivos.txt
```

Si vas a descargar varios archivos, indícalos a través de una lista en un archivo.

```
wget -c http://ejemplo.com/distro.iso
wget -i -c archivos.txt
```

Si la descarga se interrumpió por algún motivo, continuamos la descarga desde donde se haya quedado con la opción `c`.

```
wget -o reporte.txt http://ejemplo.com/programa.tar.gz
```

Proporciona un reporte (log) sobre la descarga.

```
$ wget -o /reporte.log --limit-rate=50k ftp://ftp.centos.org/download/centos5-dvd.iso
```

En descargas muy largas como el iso de alguna distro de Linux, puedes limitar el ancho de banda de la descarga en específico, ya que podría ser que la descarga se llevará todo el ancho de banda durante varias horas.

```
wget --http-user=admin --http-password=secreto http://ejemplo.com/archivo.mp3
```

Si es un sitio donde se requiere de usuario/contraseña usa estas opciones.

```
wget -t 50 http://ejemplo.com/pelicula.mpg
```

Por defecto, **wget** realiza 20 intentos de establecer la conexión e iniciar la descarga, en sitios muy saturados es posible que ni con 20 intentos se logré, con la opción `t` (tries) aumenta a más intentos.

```
wget -t inf http://ejemplo.com/pelicula.mpg
```

¿Quieres que intente hasta que logré la conexión?, usa la opción `'inf'` de intentos infinitos.

Ejemplos avanzados

```
wget www.linuxtotal.com.mx
```

Puedes descargar toda una página completa, **wget** no está limitado a solo descargas de archivos. Sitios completos pueden descargarse también.

```
wget -p www.linuxtotal.com.mx
```

Lo mismo que lo anterior pero con esta opción *p* descarga además todos los elementos extras necesarios de la página como hojas de estilos, imágenes en línea, etc.

```
wget -r www.ejemplo.com -o reporte.log  
wget -r -l10 www.ejemplo.com -o reporte.log
```

Con la opción *r* se descarga recursivamente hasta 5 niveles del sitio.

Con la opción *l* se aumenta el nivel de recursividad hasta el número indicado, también es posible usar 'inf' como valor de recursividad infinito.

```
wget --convert-links -r http://www.sitio.com/  
(o también)  
wget -k -r http://www.sitio.com/
```

Por defecto, los enlaces dentro del sitio apuntan a la dirección del dominio completo. Si deseas descargar el sitio recursivamente y después estudiarlo fuera de línea (off line) usa la opción *convert-links* que los convertirá en enlaces locales, apuntando a las páginas dentro del sitio descargado.

```
wget -r -l3 -P/tmp/sitio ftp://www.unsitio.com/
```

Por defecto, el sitio a descargar se guarda en el mismo directorio donde estés ejecutando **wget**, con la opción *P* puedes indicar un directorio distinto.

```
wget --mirror http://www.linuxtotal.com.mx/  
(o también)  
wget -m http://www.linuxtotal.com.mx/
```

De este modo obtienes una copia (un espejo) completa del sitio. La opción *--mirror* es igual a usar las opciones *-r -l inf -N* que indica recursividad a nivel infinito y obteniendo la marca de tiempo original de cada archivo descargado (opción *-N*).

```
wget --mirror --convert-links --html-extension http://www.linuxtotal.com.mx  
(o también)  
wget -m -k -E http://www.linuxtotal.com.mx
```

Si descargas el sitio completo para verlo fuera de línea (off line) es posible que varios archivos descargados no se abran, debido a extensiones como .cgi, .asp o .php, es entonces posible indicarle a **wget** con la opción *E* o *--html-extension* que convierta todos los archivos a extensión .html.

```
wget -H -r -l3 -k -E -p http://miblog.sitiodeblogs.com
```

La opción *H* (span hosts) expande la recursividad a los sitios desde donde se enlace el original. Este comando descargará el sitio con una recursividad de 3 niveles (-r -l3), convirtiendo los enlaces a locales para examinarlo off line (-k), convierte todo a extensiones .html (-E), descarga completamente imágenes y demás elementos de cada página (-p) y además descarga la página externa de los enlaces que apunten fuera del sitio (-H).

Este tipo de descargas, dependiendo del sitio, puede llenar todo un disco duro!!!!, así que úsalo a discreción.

Un ejemplo extremo de la potencia de wget

El siguiente ejemplo viene muy bien detallado en este sitio (<http://www.veen.com/jeff/archives/000573.html>), es una línea que busca archivos de música mp3 desde una lista de sitios que previamente hayas definido, recuerda, un renglón por cada sitio.

```
wget -r -l1 -H -t1 -nd -N -np -A.mp3 -erobots=off -i sitiosmp3.txt
```

Veámoslo por partes:

- [-r -l1] recursivo en solo un nivel, solo buscará un nivel en cada subdirectorio que encuentre por sitio.
- [-H] si dentro del sitio encuentra un enlace externo, que lo siga, ya que posiblemente lleve a otro sitio de descarga de música.
- [-t1] -t es igual --tries, solo hará un intento de conexión, si no lo logra continua al siguiente sitio.
- [-nd] indica no crear subdirectorios para cada archivo descargado, es decir, todo nos lo deja en el mismo lugar.
- [-N] conserva la fecha y hora (timestamp) del archivo original que se descargue, de otro modo pondría la fecha y hora de nuestro sistema.
- [-np] no parent, no seguir enlaces a directorios superiores, solo del actual y uno hacía abajo, por -r -l1
- [-A.mp3] -A indica a wget el tipo de archivo a descargar solamente, en este caso solo "mp3".
- [-erobots=off] esto evita que wget ignore los archivos 'robots.txt' que pudiera haber, ya que puede ser que dentro de estos archivos haya indicaciones de que subdirectorios los buscadores no deben hurgar (incluyendo a wget). Con esto lo evitamos y buscamos en todo el nivel.
- [-i sitiosmp3.txt] el archivo con la lista (hecha a tu gusto, con unos cuantos o decenas de sitios de música) de los sitios desde donde buscar "mp3" y descargarlos.

Si ejecutas este comando de manera diaria o a través de un cron ([index.php?cont=info_admon_006](#)) entonces se descargará la música, y gracias a la marca de tiempo solo descargará los que sean más nuevos que el original.

Los anteriores ejemplos demuestran el enorme poder de "wget" y el porque es considerado el rey de los "downloaders".

¿Requieres de una instalación o configuración de Linux o sus servicios?

¿Un desarrollo WEB empresarial a la medida?

¿Un curso o capacitación a la medida?

Revisa el sitio de SERVICIOS ([index.php?cont=servicios](https://www.linuxtotal.com.mx/index.php?cont=servicios)) de LinuxTotal

Copyright © LinuxTotal.com.mx 2006-2017
info@linuxtotal.com.mx · linuxtotal.com.mx@gmail.com