

Last updated Saturday, Nov 16, 2019

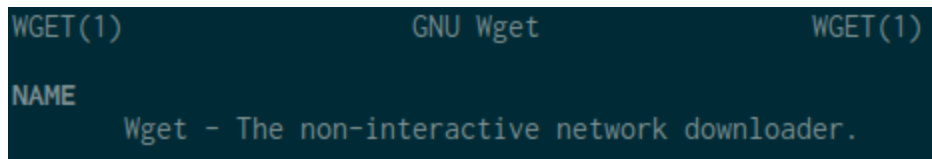
Linux and Unix wget command tutorial with examples

Tutorial on using wget, a Linux and UNIX command for downloading files from the Internet. Examples of downloading a single file, downloading multiple files, resuming downloads, throttling download speeds and mirroring a remote site.

Estimated reading time: 7 minutes

Table of contents

- [What is the wget command?](#)
- [How to download a file](#)
- [How to change the name of a downloaded file](#)
- [How to turn off output](#)
- [How to turn off verbose output](#)
- [How to resume a download](#)
- [How to download multiple URLs](#)
- [How to throttle download speeds](#)
- [How to download in the background](#)
- [How to log output to a file](#)
- [How to mirror a website](#)
- [How to find broken links](#)
- [How to set the user agent](#)
- [How to view response headers](#)
- [Further reading](#)



What is the wget command?

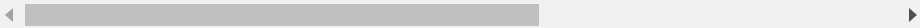
The **wget** command is a command line utility for downloading files from the Internet. It supports downloading multiple files, downloading in the background, resuming downloads, limiting the bandwidth used for downloads and viewing headers. It can also be used for taking a mirror of a site and can be combined with other UNIX tools to find out things like broken links on a site.

How to download a file

To download a file with **wget** pass the resource you would like to download.

```
wget http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016.0
--2016-09-16 11:04:40--  http://archlinux.mirrors.uk2.net/iso/2016.09
Resolving archlinux.mirrors.uk2.net (archlinux.mirrors.uk2.net)... 83
Connecting to archlinux.mirrors.uk2.net (archlinux.mirrors.uk2.net)|8
HTTP request sent, awaiting response... 200 OK
Length: 792723456 (756M) [application/x-iso9660-image]
Saving to: 'archlinux-2016.09.03-dual.iso'
```

```
archlinux-2016.09.03-dual.iso 2%[>
```



The output of the command shows **wget** connecting to the remote server and the HTTP response. In this case we can see that the file is 758M and is a MIME type of **application/x-iso9660-image**. The file will be saved as **archlinux-2016.09.03-dual.iso**. Finally the standard output of **wget** provides a progress bar. This contains the following from left to right.

- The name of the file
- A thermometer style progress bar with the percentage downloaded
- The amount of the file that has been downloaded
- The download speed
- The estimated time to complete the download


How to change the name of a downloaded file

To change the name of the file that is saved locally pass the **-O** option. This can be useful if saving a web page with query parameters. In the following example suppose that the URL **https://petition.parliament.uk/petitions?page=2&state=all** is to be downloaded.

```
wget "https://petition.parliament.uk/petitions?page=2&state=all"
--2016-09-16 11:15:26-- https://petition.parliament.uk/petitions?pag
Resolving petition.parliament.uk (petition.parliament.uk)... 52.48.19
Connecting to petition.parliament.uk (petition.parliament.uk)|52.48.1
HTTP request sent, awaiting response... 200 OK
Length: 25874 (25K) [text/html]
Saving to: 'petitions?page=2&state=all'

petitions?page=2&state=all 100%[=====

2016-09-16 11:15:27 (1.55 MB/s) - 'petitions?page=2&state=all' saved
```




Using the **ls** command shows that the file has been saved as **petitions?page=2&state=all**. To specify a different filename the **-O** option may be used.

```
wget "https://petition.parliament.uk/petitions?page=2&state=all" -O .
--2016-09-16 11:18:04-- https://petition.parliament.uk/petitions?pag
Resolving petition.parliament.uk (petition.parliament.uk)... 52.48.19
Connecting to petition.parliament.uk (petition.parliament.uk)|52.48.1
HTTP request sent, awaiting response... 200 OK
Length: 25874 (25K) [text/html]
Saving to: 'petitions.html'

petitions.html 100%[=====

2016-09-16 11:18:06 (1.38 MB/s) - 'petitions.html' saved [25874/25874]
```



How to turn off output

To turn off output use the **-q** option. This prevents wget from writing to standard output and makes it totally silent.

```
wget -q http://www.bbc.co.uk
ls
index.html
```

How to turn off verbose output

To turn off verbose output use the `-nv` option. This limits the output of `wget` but provides some useful information.


```
wget -nv http://www.bbc.co.uk
2016-09-16 11:23:31 URL:http://www.bbc.co.uk/ [172348/172348] -> "ind

```

How to resume a download

To resume a download use the `-c` option. This makes `wget` for a file in the folder that the command was run from of the same name as the remote file. If there is a file then `wget` will start the download from the end of the local file. This can be useful if a remote server dropped a connection in the middle of a download or if your network dropped during a download. In the following example the download of the Arch Linux iso is resumed.

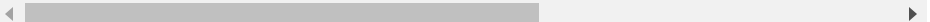
```
wget -c http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016-09-16 11:30:15-- http://archlinux.mirrors.uk2.net/iso/2016.09
Resolving archlinux.mirrors.uk2.net (archlinux.mirrors.uk2.net)... 83
Connecting to archlinux.mirrors.uk2.net (archlinux.mirrors.uk2.net)|8
HTTP request sent, awaiting response... 206 Partial Content
Length: 792723456 (756M), 734083359 (700M) remaining [application/x-i
Saving to: 'archlinux-2016.09.03-dual.iso'

archlinux-2016.09.03-dual.iso  7%[+++

```

The download is resumed from the end of the local file.

How to download multiple URLs

To download multiple files at once pass the `-i` option and a file with a list of the URLs to be downloaded. URLs should be on a separate line. In the following example a list of Linux ISOs is saved in a file called `isos.txt`.

```
http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016.09.03-dual.iso
https://www.mirrorservice.org/sites/cdimage.ubuntu.com/cdimage/releases
http://mirror.ox.ac.uk/sites/mirror.centos.org/6.8/isos/x86_64/CentOS

```

To download these files in sequence pass the name of the file to the `-i` option.

```
wget -i isos.txt
```

How to throttle download speeds

To throttle download speeds use the `--limit-rate` option. This will limit the amount of bandwidth available to `wget` and can be useful to prevent `wget` consuming all the available bandwidth. In the following example the download is limited to 400k.

```
wget --limit-rate 400k http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016-09-16 12:05:12-- http://archlinux.mirrors.uk2.net/iso/2016.09
Resolving archlinux.mirrors.uk2.net (archlinux.mirrors.uk2.net)... 83
Connecting to archlinux.mirrors.uk2.net (archlinux.mirrors.uk2.net)|8
HTTP request sent, awaiting response... 200 OK
Length: 792723456 (756M) [application/x-iso9660-image]
```

```
Saving to: 'archlinux-2016.09.03-dual.iso.6'  
archlinux-2016.09.03-dual.iso    0%[
```

It is possible to see in the output the download is limited to 400KB/s.

How to download in the background

To download in the background use the background use the `-b` option. This can be useful in that no terminal is needed to run the download. The `wget` command will return the pid number of the process and also the name of the file that output will be logged to.

```
wget -b http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016.09.03-dual.iso  
Continuing in background, pid 4135.  
Output will be written to 'wget-log'.
```

How to log output to a file

To direct `wget` output to a log file use the `-o` option and pass the name of a file.

```
wget -o iso.log http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016.09.03-dual.iso
```

To append output to a file use the `-a` option. If no file is present it will be created.

```
wget -a iso.log http://archlinux.mirrors.uk2.net/iso/2016.09.03/archlinux-2016.09.03-dual.iso
```

How to mirror a website

The `wget` command can mirror a remote website for local, offline browsing. It has many options for converting links and limiting downloads of certain file types.

```
wget -mk -w 20 http://www.example.com/
```

The options here are as follows

- `-m` turn on mirroring
- `-k` make links suitable for local browsing
- `-w` wait 20 seconds between each download. This is to avoid placing extra strain on the remote server.

Once the command has finished a local copy of the remote site will be available.

How to find broken links

To find broken links on a site `wget` can spider the site and present a log file that can be used to identify broken links.

```
wget -o wget.log -r -l 10 --spider http://example.com
```