

WWWOFFLE – Web proxy with features for dial–up users

Source: <http://linux.derkeiler.com/Newsgroups/comp.os.linux.announce/2003-12/0005.html>

From: Andrew M. Bishop (*amb_at_gedanken.demon.co.uk*)

Date: 12/02/03

Date: Tue, 2 Dec 2003 13:38:14 CST

WWWOFFLE – World Wide Web Offline Explorer – Version 2.8a
=====

The WWWOFFLE programs simplify World Wide Web browsing from computers that use intermittent (dial–up) connections to the internet.

Description

The WWWOFFLE server is a proxy web server with special features for use with dial–up internet links. This means that it is possible to browse web pages and read them without having to remain connected.

Basic Features

- Caching of HTTP, FTP and finger protocols.
- Allows the 'GET', 'HEAD', 'POST' and 'PUT' HTTP methods.
- Interactive or command line control of online/offline/autodial status.
- Highly configurable.
- Low maintenance, start/stop and online/offline status can be automated.

While Online

- Caching of pages that are viewed for later review.
- Conditional fetching to only get pages that have changed.
 - Based on expiration date, time since last fetched or once per session.
- Non cached support for SSL (Secure Socket Layer e.g. https).
- Can be used with one or more external proxies based on web page.
- Control which pages cannot be accessed.
 - Allow replacement of blocked pages.
- Control which pages are not to be stored in the cache.
- Requests compressed pages from web servers (compile time option).
- Requests chunked transfer encoding from web servers.

While Offline

- Can be configured to use dial–on–demand for pages that are not cached.
- Selection of pages to download next time online

comp.os.linux.announce: WWWOFFLE – Web proxy with features for dial-up users

- Using normal browser to follow links.
- Command line interface to select pages for downloading.
- Control which pages can be requested when offline.
- Provides non-cached access to intranet servers.

Automated Download

- Downloading of specified pages non-interactively.
- Options to automatically fetch objects in requested pages
 - Understands various types of pages
 - HTML 4.0, Java classes, VRML (partial), XML (partial).
 - Options to fetch different classes of objects
 - Images, Stylesheets, Frames, Scripts, Java or other objects.
 - Option to not fetch webbug images (images of 1 pixel square).
- Automatically follows links for pages that have been moved.
- Can monitor pages at regular intervals to fetch those that have changed.
- Recursive fetching
 - To specified depth.
 - On any host or limited to same server or same directory.
 - Chosen from command line or from browser.
 - Control over which links can be fetched recursively.

Convenience

- Optional information footer on HTML pages showing date cached and options.
- Options to modify HTML pages
 - Remove scripts.
 - Remove Java applets.
 - Remove stylesheets.
 - Remove shockwave flash animations.
 - Indicate cached and uncached links.
 - Remove the blink tag.
 - Remove refresh tags.
 - Remove links to pages that are in the DontGet list.
 - Remove inline frames (iframes) that are in the DontGet list.
 - Replace images that are in the DontGet list.
 - Replace webbug images (images of 1 pixel square).
 - Demoronise HTML character sets.
 - Stop animated GIFs.
 - Remove Cookies in meta tags.
- Provides information about cached pages
 - Headers, raw and modified.
 - Contents, images, links etc.
 - Source code unmodified by WWWOFFLE.
- Automatic proxy configuration for Netscape.
- Searchable cache with the addition of the ht://Dig, mnoGoSearch (UdmSearch) or Namazu programs.
- Built in simple web-server for local pages.
 - Allows CGI scripts
- Timeouts to stop proxy lockups
 - DNS name lookups.
 - Remote server connection.
 - Data transfer.

comp.os.linux.announce: WWWOFFLE – Web proxy with features for dial-up users

- Continue or stop downloads interrupted by client.
 - Based on file size or fraction downloaded.
- Purging of pages from cache
 - Based on URL matching.
 - To keep the cache size below a specified limit.
 - To keep the free disk space above a specified limit.
 - Interactive or command line control.
 - Compression of cached pages based on age.
- Provides compressed pages to web browser (compile time option).
- Use chunked transfer-encoding to web browser.

Indexes

- Multiple indexes of pages stored in cache
 - Servers for each protocol (http, ftp ...).
 - Pages on each server.
 - Pages waiting to be fetched.
 - Pages requested last time offline.
 - Pages fetched last time online.
 - Pages monitored on a regular basis.
- Configurable indexes
 - Sorted by name, date, server domain name, type of file.
 - Options to delete, refresh or monitor pages.
 - Selection of complete list of pages or hide un-interesting pages.

Security

- Works with pages that require basic username/password authentication.
- Automates proxy authentication for external proxies that require it.
- Control over access to the proxy
 - Defaults to local host access only.
 - Host access configured by hostname or IP address.
 - Optional proxy authentication for user level access control.
- Optional password control for proxy management functions.
- Can censor incoming and outgoing HTTP headers to maintain user privacy.

Configuration

- All options controlled using a configuration file.
- Interactive web page to allow editing of the configuration file.
- User customisable error and information pages.

Changes

Since version 2.8:

Bug Fixes:

Fix compilation problems on Win32 (Cygwin). Fixed upgrade-config.pl script error with WWWOFFLE User-Agent. Fixed bugs with detecting type of compressed data. Fixed compilation problems with dietlibc. Fix installation permission problems. Added additional paths to search scripts. Perform a cleaner socket shutdown. Fixed CSS in wwwoffle.css file. Fix potential and actual problems with io functions.

comp.os.linux.announce: WWWOFFLE – Web proxy with features for dial-up users

New Features:

Modify response headers before performing HTML modifications.

Documentation:

Added FAQ question about Content-Length header.

Availability

Version 2.8a uploaded, but may not be available yet

FTP server: <ftp://ftp.ibiblio.org/pub/Linux/apps/www/servers/wwwoffle-2.8a.tgz>

FTP server: <ftp://ftp.demon.co.uk/pub/unix/httpd/wwwoffle-2.8a.tgz>

Web page: <http://www.gedanken.demon.co.uk/wwwoffle/>

Author & Copyright

This program is copyright Andrew M. Bishop 1996,97,98,99,2000,01,02,03
(amb@gedanken.demon.co.uk) and distributed under GPL.

email: amb@gedanken.demon.co.uk

[Please put wwwoffle in the subject line]

--

Andrew.

Andrew M. Bishop

amb@gedanken.demon.co.uk

<http://www.gedanken.demon.co.uk/>

Send submissions for comp.os.linux.announce to: cola@stump.algebra.com #
PLEASE remember a short description of the software and the LOCATION. #
This group is archived at <http://stump.algebra.com/~cola/> #
#####