

Wget – Download Tool di GNU/Linux

FAJAR YUSRAN ZEBUA

fajarzebua@yahoo.com

fajarzebua@gmail.com

Lisensi Dokumen:

Copyright © 2003-2006 IlmuKomputer.Com

Seluruh dokumen di IlmuKomputer.Com dapat digunakan, dimodifikasi dan disebarkan secara bebas untuk tujuan bukan komersial (nonprofit), dengan syarat tidak menghapus atau merubah atribut penulis dan pernyataan copyright yang disertakan dalam setiap dokumen. Tidak diperbolehkan melakukan penulisan ulang, kecuali mendapatkan ijin terlebih dahulu dari IlmuKomputer.Com.

PENDAHULUAN

Surfing di dunia maya atau dengan kata lain berinternet-ria sudah menjadi kehidupan sehari-hari sebagian orang. Internet digunakan sebagai media untuk refreshing atau untuk keperluan studi. Browser yang paling populer digunakan adalah Opera, Internet Explorer, Mozilla Firefox, Konqueror, dll. Sebagian besar aktifitas di Internet selain *browsing* dan *chatting* adalah mendownload suatu file dari suatu web server. Browser seperti Opera dan Mozilla Firefox secara *default* memiliki *download manager* untuk menangani file-file yang di download di Internet. Kemampuan yang dimiliki oleh download manager bawaan suatu browser tidak secanggih aplikasi *stand-alone* lainnya, seperti Internet download manager, Download Accelerator Plus (DAP), Flashget, dll. Aplikasi *stand-alone* ini mendownload file-file di Internet dengan membagi-baginya dalam beberapa ukuran kecil, kemudian menyatukannya kembali pada sisi *client* (orang yang mendownload).

Aplikasi di atas sebagian besar berjalan pada sistem operasi Windows. Nah, bagaimana sekarang apabila pengguna Linux ingin mendownload suatu file di Internet dan ingin menggunakan aplikasi sekelas aplikasi di Windows? Jangan khawatir, untuk sistem operasi GNU/Linux terdapat tool dengan lisensi GPL (*GNU Public License*) yaitu **wget**. Wget pertama kali dirilis pada tahun 1995 oleh Hrvoje Niksic dirilis dibawah lisensi GNU (general public license). Wget adalah aplikasi freeware yang dibuat dalam bahasa C dan cukup sederhana, namun mempunyai kemampuan yang sangat powerful. Aplikasi ini dipakai untuk melakukan download, baik melalui protokol HTTP, HTTPS, ataupun FTP.

Wget memiliki beberapa kelebihan antara lain :

- a. **Gratis.** Seperti juga software dengan lisensi GPL lainnya, wget dapat didapatkan secara gratis.
- b. **Non interaktif.** Wget merupakan tool non interaktif, berarti anda bisa melakukan browsing tanpa perlu interaksi terus menerus dengan wget.
- c. **Mirroring.** Wget dapat melakukan mirror terhadap situs tertentu.
- d. **Resume.** Wget dapat melakukan resume download, sehingga sangat berguna untuk koneksi internet yang tidak stabil atau lambat.
- e. **Dukungan ekstensi file.** Dapat mendownload file-file jenis tertentu. Contoh file *.jpg.
- f. Dan masih banyak lagi. ☺

OPSI-OPSI WGET

-t (tries)

dengan opsi -t ini wget akan selalu mencoba apabila koneksi anda terputus-putus biasanya di ikuti angka seperti wget -t45 [URL]. Maka wget akan mencoba sebanyak 45 kali.

-c (continue)

Dengan opsi ini jika download anda terputus di tengah-tengah maka wget akan mendownload dari titik putus, wget tidak akan mendownload dari awal lagi.

-r (recursive)

Opsi -r akan mendownload seluruh isi situs, -r akan membuat wget menelusuri seluruh link. Opsi ini bermanfaat apabila anda membuat mirror sebuah site.

-p (page requisite)

Opsi ini akan memerintah kan wget untuk mendownload halaman depan sebuah situs lengkap dengan gambar dan semua yang ada pada halaman depan situs tersebut.

-k (converts links)

opsi ini penting digunakan apabila kita membuat mirror sebuah situs agar bisa di browse offline.

-A (accept list)

opsi ini biasanya selalu di ikuti dengan nama file yang ingin kita download contohnya:

```
$ wget -r -A gif,jpg,jpeg [URL]
```

dengan command diatas maka wget akan mendownload semua file gif, jpg dan jpeg yang terdapat pada suatu site.

-R (reject List)

Opsi -R berbeda dengan opsi -r (recursive). -R adalah kebalikan dari -A. contoh :

```
$ wget -r -k -R gif [URL]
```

wget tidak akan mendownload file gif tersebut.

-np(no parent)

dengan opsi ini -r tidak akan menelusuri keatas path. Contoh :

```
$ wget -r -k -np www.fajarzebua.net/tes
```

maka wget akan mengambil seluruh file di www.fajarzebua.net/tes tapi tidak mengambil file di section lain.

-nc (no clobber)

opsi ini merupakan kebalikan dari -c (continue). Jika -c mendownload dari titik putus maka untuk -nc mendownload file tanpa mengganggu file yang sudah setengah download.

-o

Opsi -o berguna untuk mendownload file dan disimpan dalam file tertentu contoh:

```
$ wget -o tes.txt [URL].
```

Masih banyak lagi opsi-opsi lain seperti -I, -b, -S -m, -x, -q dan lain lain. Untuk informasi lebih lanjut silahkan baca di manual wget. atau dengan mengetikkan `wget -help`.

STUDI KASUS : DOWNLOAD SCRIPT DENGAN WGET DAN CRONTAB

Pada studi kasus kali ini, kita akan mencoba mendownload sebuah script dengan wget, dan kemudian script dijalankan oleh **crontab**.

Direktori kerja kita adalah di `~/fajar/tes` atau di **home** direktori (sesuaikan dengan direktori di tempat Anda). Disini kita akan memulai membuat script-nya dulu, kemudian script ini akan di jalankan oleh **crontab**. Buatlah file dengan nama '**start-dl.sh**', gunakan editor '**vi**' atau '**nano**'. Kemudian ketik :

```
#!/bin/sh

case "$1" in
start)
    URLFILE=$(cat url.file)
    wget -c -b -T 0 -t 0 -o log -i $URLFILE &
    ps -auxww | grep $URLFILE | grep -v grep | awk '{ print $2 }'
> wget.pid
    ;;

stop)
    kill `cat ./wget.pid`
    rm -f ./wget.pid
    ;;

*)

    echo "Gunakan Perintah : '$0 {start|stop}' ."

esac
exit 0
```

Penjelasan :

Pada line `wget -c -b -T 0 -t 0 -o log -i $URLFILE` option :

-c adalah '**continueable**' atau sama dengan fungsi resume apabila download terhenti.

-b adalah menjalankan proses '**background**' wget.

-T 0 adalah men-set waktu timeout dalam hitungan detik.

-t 0 adalah **'retry'**, wget akan mencoba mendownload terus-menerus/infinite karena di set dengan value **'0'**, wget akan terus mendownload jika koneksi timeout, apabila value di set **'5'** maka download akan berhenti apabila timeout lebih dari **5** kali.

-o log adalah membuat output log dengan nama **log**

-i \$URLFILE adalah membaca url dari file, yang dimana file tersebut berisi url-url yang akan didownload. **\$URLFILE** adalah variable dari line **'URLFILE=\$(cat url.file)'**

Kemudian, save dan **'chmod u+x start-dl.sh'** agar bisa di eksekusi. Buatlah sebuah file biasa dengan nama **'tes.txt'**, kemudian ketikkan url file yang akan di download, misalnya :

'http://fajarzebua.net/music/indonesia/kenangan.mp3'. Save. Buatlah sebuah file biasa lagi dengan nama **'url.file'**, kemudian ketikkan nama file yang menyimpan url yang tertulis diatas seperti **'tes.txt'**. Save. Adapun fungsi dari file **'url.file'** adalah menyimpan variable yang fungsinya akan dipanggil untuk membuat sebuah pid file. Adapun script yang memanggil *variable* tersebut seperti yang telah di tuliskan, yaitu :

```
#-snip-#
URLFILE=$(cat url.file)
wget -c -b -T 0 -t 0 -o log -i $URLFILE &
ps -auxww | grep $URLFILE | grep -v grep | awk '{ print $2 }' >
wget.pid
#-snip- #
```

dan pada fungsi **awk '{print \$2}' > wget.pid** adalah mengambil output pada layar yang dimana **print \$2** adalah mencetak pada kolom ke 2, yaitu urutan process id yang sedang running, kemudian di simpan ke sebuah file yang bernama **wget.pid**, yang dimana file ini akan digunakan oleh crontab.

Kemudian , buat sebuah file dengan nama **'cronwget.sh'**, ketikkan :

```
#!/bin/sh

if test -r ./wget.pid; then
    WGETPID=$(cat ./wget.pid)
    if $(kill -CHLD $WGETPID >/dev/null 2>&1)
    then
        exit 0
    fi
fi
./start-dl.sh start &>/dev/null
```

Chmod file **'cronwget.sh'** menjadi **'chmod u+x cronwget.sh'** agar bisa di eksekusi.

Kemudian, atur waktu crontab, pada saat 5 menit akan men-check proses wget tadi, jika proses masih running, lanjutkan download. Tetapi, jika proses download berhenti, jalankan kembali. Proses terhenti bisa karena **'request time out'** atau komputer shutdown atau restart. Disini peran crontab di perlukan.

Sekarang, edit crontab dengan command : '**crontab -e**', maka crontab akan mengedit dengan editor 'vi'. Kemudian isikan dengan :

```
5 * * * * cd ~/download/mp3; ./cronwget.sh >/dev/null 2>&1
```

kemudian, save!

Sampai disini, selesai☺. Tinggal tunggu 5 menit, crontab akan mengeksekusi **cronwget.sh** tadi. Jika sudah lebih dari 5 menit, ketikkan :

```
[fajar@zebua tes]$ ps x
  PID TTY          STAT       TIME COMMAND
 1331 ?            S          0:01 klogd -c 1 -2
 5547 ?            R          0:00 /usr/sbin/sshd
 5548 pts/0        R          0:00 -tcsh
 5591 ?            S          0:00 wget -c -b -T 0 -t 0 -o log -i tes.txt
 9648 pts/0        R          0:00 ps x
```

Untuk memastikan apakah proses sudah running, output menampilkan '**wget -c -b -T 0 -t 0 -o log -i tes.txt**' berarti proses download sudah running.

CATATAN :

Untuk nama file url seperti '**tes.txt**', jika anda melakukan/menjalankan beberapa download, sangat di sarankan/harus untuk pemberian nama file url berbeda.

Biografi Penulis



Fajar Yusran Zebua. Dilahirkan di Nias pada tanggal 25 Juni 1984. Lulus dari SMU Negeri 1 Gunungsitoli Kabupaten NIAS pada tahun 2002. Sebentar lagi akan menyelesaikan kuliah di Institut Sains dan Teknologi AKPRIND (ISTA) Yogyakarta. Mulai mengenal komputer pada umur 10 tahun dan pada saat itu tertarik belajar Sistem Operasi UNIX. Saat ini sedang mendalami database (sql server), jaringan komputer (khususnya *wireless connection*) dan juga pemrograman Java, Visual Basic.

Informasi lebih lanjut mengenai penulis ini bisa didapat melalui :

fajarzebua@yahoo.com

fajarzebua@gmail.com