

从 Internet 上获取资源的有效方法

高玉邦 (中科院网络信息中心)

耿磊 (北京经济学院信息系)

摘要:本文根据 UNIX 系统多任务、多用户的特性,将用 FTP 获取 Internet 资源的工作放在后台或通过限定执行时间,避开网络的高峰期来完成,既提高了效率,减少用户等待的时间,又缓解网络的拥挤。

一、问题的提出

Internet 是全球最大的国际网路,目前,它拥有 5 万个网络,500 万台主机,并以每月 4 万台主机的速度增加。Internet 拥有丰富的资源,很多主机都提供了“匿名”帐号,使网上用户可以很容易地利用文件传输命令:ftp(该命令遵从:File Transfer Protocol)以“anonymous”为用户名,Email 地址为通行字,来访问并共享该主机的资源,极大地方便了用户。然而,有了这些资源,当我们去取的时候,会发现我们遇到的一个极大的问题便是传输速度。尤其在我国,中国国家计算与网络设施 NCFC(The National Computing and Networking Facility of China)工程刚投入运行一年多,租用连至美国进入 Internet 的专线只有 64kbps,在同一时间内正在使用的用户更无法计数。因此,网上的传输速度便成为一个瓶颈。要从远地主机,尤其是从国外的主机上获取资源,对用户来说,将需一个漫长的等待时间。

本文基于 UNIX 系统多任务、多用户的特性,利用 UNIX 系统自身提供的基本命令,把获取资源的作业(JOBS)提交给后台来完成,从而免除了用户等待的痛苦,且能获取所需的资源。

二、实现的方法

1. 初始化设定文件.netrc

在 UNIX 系统中,许多工具程序都可籍由其所属的初始化设定文件作一些参数的设定工作,如:sh 有 profile, csh 有 cshrc, vi 有 exrc 等等,同样,ftp 也有它的初始化设定文件 netrc。netrc 文件是一个隐含文件,它包

含了 ftp 命令自动登录所需的一些信息,它存放在用户的根目录(HOME)下,每当用 ftp 命令登录时,ftp 命令将首先在根目录下查找该文件。如果该文件存在,则根据该文件的设定实现自动登录;否则将以交互式的方式进行登录。

下面是 netrc 文件包含的自动登录的一些主要信息,它们之间以空格、TAB(S)或空行分开:

(1) machine host. 这里,host 是一个欲登录的远程主机的名字。关键字“machine”是 ftp 进入自动登录到特定主机的定义的开始,并直到遇到下一个“machine”或文件的结束为止。

default 可以以缺省的方式来定义,其设定同“machine”一样,不同的是,它可以匹配任何的主机名,所以,在文件.netrc 中,只能有一个“default”,而且它必须出现在所有的“machine”之后。它的一般用法为:

default login anonymous password username@

(2) login user. user 是在远程主机上的用户名。如果在.netrc 文件中有 login 参数的话,则在自动登录进程中将采用给定的“user”作为用户名进行登录。如果此用户不存在,则登录将失败。

(3) password password. password 是自动登录时用到的通行字,自动登录进程将该 password 提供给远程主机。该 password 一定要在远程主机上建立,并在.netrc 文件中给出,否则,自动登录将提示用户手工输入通行字。需要注意的是,如果在.netrc 文件中包含有用户的通行字,系统为保证用户的安全,要求.netrc 文件的模式一定为 600,即只有用户本身才能读写,而对其它用户进行限制,以免用户通行字的泄漏,否则的话,将会提示:

Error—.netrc incorrect mode

Remove password or correct mode

ftp>

自动登录失败。

(4)macdef macro。这里,macro 是 ftp 子命令的宏定义。宏定义包含了所有要执行的 ftp 子命令,以一个空行作为一个宏定义的结束。定义的方法是:

ftp> macdef macro <—宏定义的名字

子命令 1

子命令 2

:

子命令 n

<—一个宏定义的结束

ftp> macro <—执行该宏定义

如果"macro"为"init",则自动登录进程将执行"int"宏定义所包含的所有的 ftp 子命令。如果没有宏定义,或宏定义的名字不是"init",将无法执行 ftp 子命令,并可能会提示出错。

根据以上的分析,我们就可以得到一个用于 ftp 自动登录的初始化设定文件.netrc。

例:

¥ cat .netrc

machine ns.cnc.ac.cn login user password bldadk

machine osf.cnc.ac.cn login user passowrd abcdef

machinegingko.ict.ac.cn loginuser password 123456

default login anonymous password user@

macdef init

pwd

dir

bye

¥

若要登录到"ns.cnc.ac.cn"主机上,则输入:¥ ftp ns.cnc.ac.cn
将根据第一行,自动完成远程主机的登录过程。

若要登录到"nctuccca.edu.tw"主机上,则输入:

¥ ftp nctuccca.edu.tw

将根据第三行,自动登录到台湾的"nctuccca"主机上。

这样,通过.netrc 文件,便实现了 ftp 的自动登录及对登录主机的访问。

2.后台处理的两种方法

了解了 ftp 自动登录的实现之后,便可以采取以下两种方法到远程主机上获取资源,而无需一直占用终端。

(1)直接放入后台处理。第一种方法非常简单,例如

到"ftp.cnc.ac.cn"主机上获取资源,首先建立好.netrc 文件:

¥ cat .netrc

default login anonymous password user@

macdef init

cd / pub, RFC, rfc14xx

prompt

mget rfc1432.txt

bye

¥

然后,在系统提示符下输入:

¥ ftp ftp.cnc.ac.cn

系统回显:

[1]10522

它表示一个系统进程 PID 之后,便可做其它事或关掉终端,而主机则会替您完成所有的工作。

这种方式是一旦用户提交,主机马上同远程主机进行通信,并将通信过程在您的终端上显示出来,无论您在做什么工作。另外,这种方式即使在网络的高峰期,仍然执行,但传输的速度是非常慢的,有时甚至会失败的。

(2)第二种方式是用 UNIX 系统的"at"命令.UNIX 系统的"at"命令从标准输入中读取或接收一个参数作为以后执行的命令,它允许限定命令执行的时间,格式为:

¥ at time

命令

ctrl-d <—按"ctrl"键再按"d"键

同样,在系统回显一个进程号之后,用户就不需关心任务的执行了。当任务完成之后,便自动发送一个 Email,汇报任务实现的全过程,包括任务的成功或失败。

因此,我们要以通过限定命令报告的时间,避开网络的高峰期,在网络空闲的时候,来获取远程主机的资源。这不但会提高传输的速度,而且也会保证传输的成功率。

如上例,我们不妨将获取资源的任务放在第二天凌晨 1 点执行,输入如下:

¥ at 01:00

ftp ftp.cnc.ac.cn

ctrl-d

fob USER.7849592.a at MON Apr 12:00 1995

(任务提交时间)

第二天,您就会发现,那么多的资料已在您休息的时间传了过来。