Mirroring FreeBSD

摘要

這是份介紹如何 mirror FreeBSD,主要是針對網路中心管理者或託管於大型資料中心的管理者.

目⊠

1. 聯繫方式 1
2. FreeBSD mirrors 的需求 1
3. 如何Mirror FreeBSD 站台 3
4. Where to mirror from 4
5. Official Mirrors 6
6. Some statistics from mirror sites 6

我們目前不接受新Mirror站點的申請.

1. 聯繫方式

The Mirror System Coordinators can be reached through email at mirror-admin@FreeBSD.org.
There is also a FreeBSD mirror sites mailing lists.

2. FreeBSD mirrors 的需求

2.1. 磁碟空間

磁碟空間是最為需要. 根據你想要 mirror 的發行版、CPU架構 ,可能會消耗大量的磁碟空間.另外請注意 官方鏡像站需要完整 mirror。網站內容亦需要完整鏡像。且這裡所述的數字是反應目前版本狀態 (如 10.4-RELEASE/11.1-RELEASE)。而不斷的開發與發行將會增加所需空間。並請務必保留一些 (約10-20%) 額外空間。這裡大約估計如下：

- 完整的作業系統套件 FTP 站所需：1.4 TB
- CTM deltas： 10 GB
- 網站： 1GB

目前 FTP Distribution 的磁碟使用可在 ftp://ftp.FreeBSD.org/pub/FreeBSD/dir.sizes 找到。

2.2. 網路連線/頻寬

當然，你一定要能連上 Internet。 頻寬需求多少，這要看你所想要的 mirror 程度而定。 若只想要 mirror 一部份的 FreeBSD 檔案以作為網站或 intranet 的局部用途，
那麼頻寬需求會明顯比成為公共服務用途的小一些。 若想成為 official mirror
之一的話，那麼頻寬就勢必得增加才夠用。以下，我們僅列出一些估計值以做為參考：

- 本地站台，沒有要公共存取： 基本上沒有最低需求，但是 < 2Mbps 同步將會非常緩慢
- 非官方公共站台: 34Mbps 是不錯的開始.
- 官方站台: > 100Mbps 是建議值,並且你的主機必須盡可能連接靠近邊界路由器.

2.3. 系統需求，CPU，RAM

這取決於預期的客戶端數量,這是由伺服器的策略決定的。也會受到您提供的服務類型而影響.普通的 FTP 或 HTTP 服務可能不需要大量的資源。注意如果您提供rsync. 這可能會對 CPU
和記憶體的需求產生巨大的影響，因為會消耗大量記憶體。 以下只是給你一個非常粗略的的例子。

針對一個較常被瀏覽的網站 rsync，您須考量處理器大約 800Mhz 至 1Ghz，並且安裝最少 512MB RAM，這或許是成為一個 官方 站台的最小需求.

為了一個經常使用的網站你絕對需要更多 RAM (2GB是不錯的開始) 並且儘可能有更多 CPU ，
這也表示你需要一個 SMP 系統。

您也會需要考慮有一個較快的磁碟系統。在管理 SVN repository 需要一個快速的磁碟系統 (強烈建議 RAID)。有自己的快取記憶體的 SCSI
控制器也可以加快速度，因為大多數這些服務會對磁碟進行大量的小幅修改。

2.4. 提供的服務

每個鏡像站都需要一有一組可用的核心服務。除了這些所需的服務之外，還有許多伺服器管理員可以選擇提供的選用服務。本節將說明您可以提供哪些服務以及如何實作這些服務。

2.4.1. FTP (需要提供給FTP檔案集)

這是最基本的服務之一。需要為每個鏡像站提供公共的 FTP distributions 。 FTP 存取必須是匿名的，
不允許上傳/下載比率 （這是一件荒謬的事），上傳功能不是必需的 （且必須 絕不允許 FreeBSD 檔案空間）。另外，FreeBSD archive 應該在路徑 /pub/FreeBSD 下。

這裡有很多可用的軟體可以架設允許匿名的 FTP 服務 （按字母順序）。

- /usr/libexec/ftpd: FreeBSD 內建的 ftpd 可以使用。請您參閱 ftpd(8)。
- ftp/ncftpd。一個商業軟體套件，免費供教育使用。
- ftp/oftpd：一個以安全性作為主要考量的 ftpd。
- ftp/proftpd：一個模組化且非常有彈性的 ftpd。
- ftp/pure-ftpd: 另一個為安全所設計的 ftpd。
- ftp/twoftpd：如上。
- ftp/vsftpd："非常安全的" ftpd。

FreeBSD 的 ftpd、proftpd 和也許 ncftpd 是最常使用的 FTP
軟體。其他的在鏡像站並沒有大量用戶基礎。需要考慮的一件事情是,您可能需要性地來限制允許同時連線數，從而限制消耗多少網路頻寬和系統資源。

2.4.2. Rsync (給FTP檔案集選用)

rsync 通常是用在存取 FreeBSD
系統中的FTP內容，其他的鏡像站可以使用你的系統當作他們的來源。這個協定和 FTP
有很多不同，它比較不那麼消耗頻寬，只有當比對檔案間有變動才傳輸檔案，而不是整個檔案傳完。rsync 需要較多的記憶體。大小取決於檔案與目錄的數目及同步模組大小。rsync 可以使用 rsh 和 ssh
（現在為預設）來傳輸，
或使用自己的協定單獨存取（這是公共rsync伺服器的首選方法）。可以用認證、連接限制和其他限制。只有一個軟體套件可以用：

- net/rsync

2.4.3. HTTP（網頁需要，FTP 檔案集則是選用）

如果您想提供 FreeBSD 的網頁，您需要安裝一個網頁伺服器。您可以選擇利用 HTTP 提供 FTP
檔案集。網頁伺服器軟體的選擇留給鏡像站管理員選擇。一些最受歡迎的選擇是：

- www/apache22：Apache 是網際網路上最廣泛使用的網頁伺服器。 它被 FreeBSD 計畫廣泛使用。
- www/thttpd：如果您要提供大量的靜態內容，您可能會發現使用諸如 thttpd 之類的應用程式會比 Apache 更有效率。它針對 FreeBSD 的優秀性能進行了最佳化。
- www/boa：Boa 是 thttpd 和 Apache 外的另一個選擇。對於純粹的靜態網頁，它應該會提供比 Apache 更好的性能。在寫這篇文章的時候，它並不包含像在 thttpd 中一樣針對FreeBSD 做最佳化。
- www/nginx：Nginx 是一款高性能的最新網頁服務器，具有低記憶體佔用量和關鍵特色，可以構建現代高效率網頁基礎架構，功能包括 HTTP 伺服器，HTTP 和郵件反向代理，快取，負載平衡，壓縮，請求限制(request throtting)，連接多工與再利用，SSL 卸載和 HTTP 媒體串流。

3. 如何Mirror FreeBSD 站台

好，現在你知道硬體需求和如何提供服務，但不知道如何做。:-) 這節將解釋如何實際 mirror FreeBSD 的不同部分，使用哪些工具以及從哪裡 mirror。

3.1. 鏡像 FTP 站

FTP 部份有最大量的資料需要被 mirror。它包括網路安裝所需的發布集，實際上是原始碼樹快照的分支，可燒錄光碟供安裝系統的ISO映像檔 ，一個可 live 開機的檔案系統，以及一個 port tree 的快照。當然，全都有各種 FreeBSD 版本和各種CPU架構。

The best way to mirror the FTP area is rsync. You can install the port net/rsync and then use rsync to sync with your upstream host. rsync is already mentioned in Rsync (給FTP檔案集選用). Since rsync access is not required, your preferred upstream site may not allow it. You may need to hunt around a little bit to find a site that allows rsync access.

> 由於 rsync 客戶端的數量將對伺服器主機產生重大影響，因此大多數管理員會對伺服器負荷加以限制。對於 mirror 站台，您應該詢問您要 mirror 站台的管理人員他們的管理政策，也許需要對您的主機開放例外（因為您是一個 mirror 站）。

一個需要mirror FreeBSD官網的指令如下:

```
% rsync -vaHz --delete rsync://ftp4.de.FreeBSD.org/FreeBSD/ /pub/FreeBSD/
```

Consult the documentation for rsync, which is also available at http://rsync.samba.org/, about the various options to be used with rsync. If you sync the whole module (unlike subdirectories), be aware that the module-directory (here "FreeBSD") will not be created, so you cannot omit the target directory. Also you might want to set up a script framework that calls such a command via cron(8).

3.2. Mirroring 網頁

FreeBSD 網站應只能透過 rsync 指令來mirror.

一個 mirror FreeBSD 網站的指令應該看起像這樣：

```
% rsync -vaHz --delete rsync://bit0.us-west.freebsd.org/FreeBSD-www-data/
/usr/local/www/
```

3.3. Mirroring 套件

由於對頻寬，儲存空間和管理的要求非常高，FreeBSD 計畫決定不允許公眾 mirror 套件.
對於擁有大量伺服主機的網站，建議為 pkg(8) 使用 HTTP proxy
快取可能會有所幫助。或者，您可以使用以下指令獲得套件與相依套件：

```
% pkg fetch -d -o /usr/local/mirror vim
```

一旦這些套件包被下載，就必須執行以下命令來產生套件庫數據：

```
% pkg repo /usr/local/mirror
```

一旦套件被下載並且已經生成了套件庫的數據，就可以透過 HTTP
協定將套件提供給客戶端機器。有關更多訊息，請參閱 pkg(8) 的 man pages，特別是 pkg-repo(8) 頁面。

3.4. 我多久應該mirror?

Every mirror should be updated at a minimum of once per day. Certainly a script with locking to prevent multiple runs happening at the same time will be needed to run from cron(8). Since nearly every admin does this in their own way, specific instructions cannot be provided. It could work something like this:

1. Put the command to run your mirroring application in a script. Use of a plain /bin/sh script is recommended.
2. Add some output redirections so diagnostic messages are logged to a file.
3. Test if your script works. Check the logs.
4. Use crontab(1) to add the script to the appropriate user’s crontab(5). This should be a different user than what your FTP daemon runs as so that if file permissions inside your FTP area are not world-readable those files can not be accessed by anonymous FTP. This is used to "stage" releases — making sure all of the official mirror sites have all of the necessary release files on release day.

   Here are some recommended schedules:

   - FTP fileset: daily
   - WWW pages: daily

4. Where to mirror from

This is an important issue. So this section will spend some effort to explain the backgrounds. We will say this several times: under no circumstances should you mirror from ftp.FreeBSD.org.

4.1. A few words about the organization

Mirrors are organized by country. All official mirrors have a DNS entry of the form ftpN.CC.FreeBSD.org. CC (i.e. country code) is the top level domain (TLD) of the country where this mirror is located. N is a number, telling that the host would be the Nth mirror in that country. (Same applies to wwwN.CC.FreeBSD.org, etc.) There are mirrors with no CC part. These are the mirror sites that are very well connected and allow a large number of concurrent users. ftp.FreeBSD.org is actually two machines, one currently located in Denmark and the other in the United States. It is NOT a master site and should never be used to mirror from. Lots of online documentation leads

"interactive"users to ftp.FreeBSD.org so automated mirroring systems should find a different machine to mirror from.

Additionally there exists a hierarchy of mirrors, which is described in terms of tiers. The master sites are not referred to but can be described as Tier-0. Mirrors that mirror from these sites can be considered Tier-1, mirrors of Tier-1-mirrors, are Tier-2, etc. Official sites are encouraged to be of a low tier, but the lower the tier the higher the requirements in terms as described in FreeBSD mirrors 的需求. Also access to low-tier-mirrors may be restricted, and access to master sites is definitely restricted. The tier-hierarchy is not reflected by DNS and generally not documented anywhere except for the master sites. However, official mirrors with low numbers like 1-4, are usually Tier-1 (this is just a rough hint, and there is no rule).

4.2. Ok, but where should I get the stuff now?

Under no circumstances should you mirror from ftp.FreeBSD.org. The short answer is: from the site that is closest to you in Internet terms, or gives you the fastest access.

4.2.1. I just want to mirror from somewhere!

If you have no special intentions or requirements, the statement in Ok, but where should I get the stuff now? applies. This means:

1. Check for those which provide fastest access (number of hops, round-trip-times) and offer the services you intend to use (like rsync).
2. Contact the administrators of your chosen site stating your request, and asking about their terms and policies.
3. Set up your mirror as described above.

4.2.2. I am an official mirror, what is the right site for me?

In general the description in I just want to mirror from somewhere! still applies. Of course you may want to put some weight on the fact that your upstream should be of a low tier. There are some other considerations about official mirrors that are described in Official Mirrors.

4.2.3. I want to access the master sites!

If you have good reasons and good prerequisites, you may want and get access to one of the master sites. Access to these sites is generally restricted, and there are special policies for access. If you are already an official mirror, this certainly helps you getting access. In any other case make sure your country really needs another mirror. If it already has three or more, ask the "zone administrator" (hostmaster@CC.FreeBSD.org) or FreeBSD mirror sites mailing lists first.

Whoever helped you become, an official should have helped you gain access to an appropriate upstream host, either one of the master sites or a suitable Tier-1 site. If not, you can send email to mirror-admin@FreeBSD.org to request help with that.

There is one master site for the FTP fileset.

ftp-master.FreeBSD.org

This is the master site for the FTP fileset.

ftp-master.FreeBSD.org provides rsync access, in addition to FTP. Refer to 鏡像 FTP 站.

Mirrors are also encouraged to allow rsync access for the FTP contents, since they are Tier-1 -mirrors.

5. Official Mirrors

Official mirrors are mirrors that

- a) have a FreeBSD.org DNS entry (usually a CNAME).
- b) are listed as an official mirror in the FreeBSD documentation (like handbook).

So far to distinguish official mirrors. Official mirrors are not necessarily Tier-1-mirrors. However you probably will not find a Tier-1-mirror, that is not also official.

5.1. Special Requirements for official (tier-1) mirrors

It is not so easy to state requirements for all official mirrors, since the project is sort of tolerant here. It is more easy to say, what official tier-1 mirrors are required to. All other official mirrors can consider this a big should.

Tier-1 mirrors are required to:

- carry the complete fileset
- allow access to other mirror sites
- provide ftp and rsync access

Furthermore, admins should be subscribed to the FreeBSD mirror sites mailing lists. See this link for details, how to subscribe.

> It is very important for a hub administrator, especially Tier-1 hub admins, to check the release schedule for the next FreeBSD release. This is important because it will tell you when the next release is scheduled to come out, and thus giving you time to prepare for the big spike of traffic which follows it.
>
> It is also important that hub administrators try to keep their mirrors as up-to-date as possible (again, even more crucial for Tier-1 mirrors). If Mirror1 does not update for a while, lower tier mirrors will begin to mirror old data from Mirror1 and thus begins a downward spiral… Keep your mirrors up to date!

5.2. How to become official then?

We are not accepting any new mirrors at this time.

6. Some statistics from mirror sites

Here are links to the stat pages of your favorite mirrors (a.k.a. the only ones who feel like providing stats).

6.1. FTP site statistics

- ftp.is.FreeBSD.org - hostmaster@is.FreeBSD.org - (Bandwidth) (FTP processes) (HTTP processes)
- ftp2.ru.FreeBSD.org - mirror@macomnet.ru - (Bandwidth) (HTTP and FTP users)