Crear réplicas de FreeBSD

Resumen

Un artículo en progreso sobre cómo crear una replica de FreeBSD, destinado a los administradores de hub.

Tabla de contenidos

1. Información de contacto . . . 1
2. Requisitos para las réplicas de FreeBSD . . . 1
3. Cómo replicar FreeBSD . . . 4
4. De donde replicar . . . 6
5. Replicas oficiales . . . 8
6. Algunas estadísticas de las réplicas . . . 9

No aceptamos nuevas réplicas en este momento.

1. Información de contacto

Se puede contactar con los Coordinadores del Sistema de Réplicas en el email mirror-admin@FreeBSD.org. También existe Lista de correo de las réplicas de FreeBSD.

2. Requisitos para las réplicas de FreeBSD

2.1. Espacio en disco

El espacio en disco es uno de los requisitos más importantes. Dependiendo del conjunto de releases, arquitecturas y cómo de completas quieres que sean tus réplicas, se podría consumir una enorme cantidad de espacio en disco. Ten en cuenta también que para las réplicas *oficiales* se podría requerir que fueran completos. Las páginas web se tienen que duplicar en la réplica de forma completa. También ten en cuenta que los números utilizados aquí reflejan el estado actual (en 12.0-RELEASE/11.3-RELEASE). Desarrollos posteriores y releases sólo pueden incrementar las cantidades requeridas. Asegúrate también de tener algo (un 10-20%) de espacio extra para estar seguro. Aquí hay algunos números aproximados:

- Distribución FTP completa: 1.4 TB
- Deltas de CTM: 10GB
- Páginas web: 1GB

El uso de disco actual de la Distribución FTP se puede encontrar en

ftp://ftp.FreeBSD.org/pub/FreeBSD/dir.sizes.

2.2. Conexión de red/ancho de banda

Por supuesto, necesitas estar conectado a internet. El ancho de banda requerido depende del uso previsto para la replica. Si solo deseas replicar algunas partes de FreeBSD para uso local en tu sitio/intranet, la demanda puede ser mucho menor que si quieres que los archivos estén disponibles públicamente. Si pretendes convertirte en una replica oficial, el ancho de banda requerido será aún mayor. Aquí solo podemos dar estimaciones aproximadas:

- Sitio local, sin acceso público: básicamente, no hay un mínimo, pero con < 2 Mbps puede que la sincronización sea muy lenta.
- Sitio público no oficial: 34 Mbps es probablemente un buen comienzo.
- Sitio oficial: se recomiendan > 100 Mbps, y tu host debe estar conectado lo más cerca posible de tu border router.

2.3. Requisitos del sistema, CPU, RAM

Esto depende mucho del número de clientes esperado, que está determinado por la política del servidor. El tamaño también se ve afectado por el tipo de servicios que quieres ofrecer. Los servicios FTP o HTTP simples pueden no necesitar una gran cantidad de recursos. Ten cuidado si proporcionas rsync. Esto puede tener un gran impacto en los requisitos de CPU y memoria, ya que este servicio consume una gran cantidad de memoria. Los ejemplos siguientes tienen como objetivo darte una idea bastante aproximada.

Para un sitio que ofrece rsync y que sea visitado con frecuencia moderada, podrías considerar una CPU con unos 800MHz - 1 GHz y al menos 512 MB RAM. Este es probablemente el mínimo que quieres para un sitio *oficial*.

Si es un sitio con un número de visitas alto, definitivamente, necesita más RAM (2GB es un buen comienzo) y posiblemente más CPU, esto también significa la necesidad de ir hacia un sistema que disponga de SMP.

También deberías de considerar un sistema rápido de discos. Las operaciones en el repositorio SVN requieren de un sistema de discos que sea rápido (se recomienda RAID). Un controlador SCSI que disponga de una caché propia también puede acelerar las cosas, ya que la mayoría de estos servicios realizan una gran cantidad de pequeñas modificaciones en el disco.

2.4. Servicios a ofrecer

Se require que cada replica tenga un conjunto de servicios centrales disponibles. Además de estos servicios que son requeridos, hay una serie de servicios opcionales que los administradores del servidor pueden escoger ofrecerlos o no. Esta sección explica qué servicios puedes proporcionar y cómo implementarlos.

2.4.1. FTP (requerido para el conjunto de archivos FTP)

Este es uno de los servicios más básicos, y es requerido en cada réplica que ofrezca distribuciones FTP públicas. El acceso a FTP debe ser anónimo, y no se permite ningún control sobre las tasas de subida/descarga (de todas formas, es ridículo). La subida de archivos no es necesaria (y *nunca* debe permitirse en el área donde se guardan los archivos de FreeBSD). Además, los archivos de FreeBSD deben estar disponibles en la ruta /pub/FreeBSD.

Existe una gran cantidad de software disponible que se puede configurar para permitir FTP anónimo (en orden alfabético).

- `/usr/libexec/ftpd`: Se puede usar el propio ftpd de FreeBSD. Asegúrate de leer ftpd(8).
- ftp/ncftpd: Un paquete comercial, gratis para uso educativo.
- ftp/oftpd: Un ftpd diseñado con la seguridad como principal objetivo.
- ftp/proftpd: Un ftpd modular y muy flexible.
- ftp/pure-ftpd: Otro ftpd desarrollado con la seguridad como principal objetivo.
- ftp/twoftpd: Lo mismo que el anterior.
- ftp/vsftpd: El ftpd "muy seguro" ("very secure").

`ftpd` de FreeBSD, `proftpd` y quizás `ncftpd` están entre los FTPds más usados. Los otros no tienen una base de usuarios tan grande entre los sitios de réplicas. Algo a considerar es que podrías necesitar flexibilidad para limitar el número de conexiones simultáneas que se permiten, limitando por tanto cuánto ancho de banda y recursos de sistema se consumen.

2.4.2. Rsync (opcional para el conjunto de archivos FTP)

A menudo, se ofrece rsync para acceder a los contenidos del área FTP de FreeBSD, de esta forma, otras réplicas pueden usar tu sistema como su fuente. El protocolo es diferente de FTP en muchos aspectos. Es mucho más amigable con el ancho de banda, ya que solo se transfieren las diferencias entre los archivos en lugar del archivo completo. rsync requiere una cantidad significativa de memoria para cada instancia. El tamaño depende del tamaño del módulo que se vaya a sincronizar, en términos de cantidad de directorios y archivos. rsync puede usar `rsh` y `ssh` (por defecto) como transporte, o usar su propio protocolo para el acceso independiente (este es el método preferido para los servidores de rsync públicos). Se pueden aplicar autenticación, límites de conexión y otras restricciones. Solo hay un paquete de software disponible:

- net/rsync

2.4.3. HTTP (requerido para las páginas web, opcional para el conjunto de archivos FTP)

Si quieres ofrecer las páginas web de FreeBSD, necesitarás instalar un servidor web. Opcionalmente, puedes ofrecer el conjunto de archivos FTP a través de HTTP. La elección del software del servidor web se deja en manos del administrador de la réplica. Algunas de las opciones más populares son:

- www/apache24: es todavía uno de los servidores web con más despliegues en Internet. Se usa

de forma extensiva en el Proyecto FreeBSD.

- www/boa: Boa es un servidor HTTP single-task. A diferencia de los servidores web tradicionales, su proceso no se divide por cada conexión entrante, ni crea varias copias de sí mismo para manejar múltiples conexiones. Aún así, debería proporcionar un rendimiento considerablemente grande para contenido puramente estático.
- www/cherokee: Cherokee es un servidor web muy rápido, flexible y fácil de configurar. Es compatible con las tecnologías más extendidas en la actualidad: FastCGI, SCGI, PHP, CGI, conexiones cifradas SSL/TLS, vhosts, autenticación de usuarios, codificación on the fly y balanceo de carga. También genera archivos logs compatibles con Apache.
- www/lighttpd: es un servidor web seguro, rápido, estándar y muy flexible que ha sido optimizado para entornos de alto rendimiento. Tiene unos requisitos de memoria muy bajos comparados con otros servidores web y se preocupa por la carga de la CPU.
- www/nginx: nginx es un servidor web de alto rendimiento con bajo consumo de memoria y características clave para crear una infraestructura web moderna y eficiente. Las características incluyen, un servidor HTTP, proxy inverso de HTTP y correo electrónico, almacenamiento en caché, balanceo de carga, compresión, limitación de solicitudes, multiplexación y reutilización de conexiones, descarga de SSL y streaming de medios por HTTP.
- www/thttpd: Si vas a ofrecer una gran cantidad de contenido estático, es posible que el uso de una aplicación como thttpd sea más eficiente que otras. También está optimizado para tener un rendimiento excelente en FreeBSD.

3. Cómo replicar FreeBSD

Ok, ahora ya conoces los requisitos y cómo ofrecer los servicios, pero no cómo obtenerlos. :-) Esta sección explica cómo replicar las distintas partes de FreeBSD, qué herramientas usar y de dónde replicar.

3.1. Replicar el sitio FTP

El área FTP tiene la mayor cantidad de datos que deben ser replicados. Incluye los *conjuntos de la distribución* (distribution sets) necesarios para la instalación en red, las *ramas* (branches) que son snapshots de los árboles del código fuente, las *imagenes ISO* para grabar CD-ROMs con la distribución de instalación, un sistema de archivos live y un snapshot del árbol de ports. Y claro, todo esto para las diversas versiones y arquitecturas de FreeBSD.

La mejor forma de replicar el área de FTP es con rsync. Puedes instalar el port net/rsync y después utilizar rsync para sincronizar con el servidor origen. rsync ya se menciona en Rsync (opcional para el conjunto de archivos FTP). Puesto que el acceso por rsync no es un requisito el sitio desde el que réplicas podrías no permitirlo. Podrías necesitar buscar un poco para encontrar un sitio que te permita acceso por rsync.

Dado que el número de clientes de rsync tendrá un impacto significativo en el rendimiento del servidor, la mayoría de los administradores imponen limitaciones en sus servidores. Para una réplica, debes preguntar al mantenedor del sitio con el que se está sincronizando sobre su política, y tal vez, pedir una excepción para tu

host (ya que tú eres una réplica).

Un ejemplo en la línea de comandos para replicar FreeBSD podría ser de la siguiente forma:

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

Consulta la documentación de rsync, que también está disponible en http://rsync.samba.org/, para conocer las distintas opciones que se pueden utilizar. Si sincronizas un módulo completo (a diferencia de subdirectorios), ten en cuenta que el directorio del módulo (aquí "FreeBSD") no se creará, por lo que no puedes omitir el directorio de destino. También podrías querer tener un script que llame a esos comandos via cron(8).

3.2. Replicando las páginas WWW

Desde la migración de documentos a Hugo/Asciidoctor en 2021-01-25, la replicación del sitio web mediante rsync ya no funciona.

Hay estudios en curso para implementar una réplica del sitio web con la infraestructura oficial.

Para las réplicas antiguas del sitio web, una forma de conseguir la replicación del sitio web a día de hoy es construir el sitio de forma local con la correspondiente dirección que lo albergará.

```
% cd website && env HUGO_baseURL="https://www.XX.freebsd.org/" make
```

Visita el libro FreeBSD Documentation Project Primer for New Contributors para más detalles acerca de las herramientas de construcción.

Ten en cuenta que el sitio web se ha dividido en www.FreeBSD.org y docs.FreeBSD.org y que hay enlaces entre ellos; además, en este momento la variable `HUGO_baseURL` no cubre todos los enlaces, desaconsejando por lo tanto la replicación del sitio web de esta forma.

3.3. Replicando los paquetes

El Proyecto FreeBSD ha decidido no permitir la replicación pública de paquetes debido a los altos requisitos de ancho de banda, almacenamiento y administración. Para sitios con muchas máquinas, podría ser una ventaja tener un proxy HTTP para el proceso de pkg(8). De forma alternativa, se pueden descargar paquetes específicos junto con sus dependencias ejecutando algo como lo siguiente:

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

Cuando estos paquetes se descarguen, los metadatos del repositorio se deben generar ejecutando:

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

Una vez que los paquetes se han descargado y que se ha generado la metadata del repositorio, puedes servir los paquetes hasta las máquinas cliente mediante HTTP. Para obtener información adicional, lee las páginas de manual para pkg(8), en concreto la página pkg-repo(8).

3.4. ¿Con qué frecuencia debo actualizar el contenido de mi replica?

Cada réplica se debería actualizar como mínimo una vez al día. Seguramente para ejecutar un script desde cron(8) el script necesite bloqueos para evitar varias instancias ejecutándose al mismo tiempo. Puesto que prácticamente cada administrador hace esto de una forma distinta, no se pueden proporcionar instrucciones específicas. Podría funcionar algo como esto:

1. Escribe el comando para ejecutar su réplica en un script. Se recomienda el uso de un script simple de `/bin/sh`.
2. Agrega algunas redirecciones de salida para que los mensajes de diagnóstico se registren en un archivo.
3. Comprueba si el script funciona. Comprueba los logs.
4. Usa crontab(1) para añadir el script al crontab(5) del usuario correspondiente. Este debería ser un usuario diferente del que utiliza tu demonio de FTP de forma que si los permisos de los ficheros dentro de tu área de FTP no son legibles para todo el mundo esos ficheros no pueden ser accedidos mediante un FTP anónimo. Esto es lo que se usa para "preparar" las releases - asegurarse de que todas las réplicas oficiales tienen todos los ficheros de release el día de la liberación de la misma.

Aquí hay algunas programaciones recomendadas:

- Conjunto de archivos FTP: diario
- Páginas WWW: diario

4. De donde replicar

Este es un punto importante. Así que en esta sección vamos a esforzarnos un poco para explicar los antecedentes. Lo diremos varias veces: bajo ninguna circunstancia deberías replicar desde `ftp.FreeBSD.org`.

4.1. Algunas palabras sobre la organización

Las réplicas se organizan por país. Todas las réplicas oficiales tienen una entrada DNS de la forma `ftpN.CC.FreeBSD.org`. *CC* (es decir, código de país o country code) es el *top level domain* (TLD) del país donde se encuentra la réplica. *N* es un número que indica que el host debería ser la réplica

número *N* en ese país. (Lo mismo se aplica a `wwwN.CC.FreeBSD.org`, etc.) Hay réplicas sin la parte *CC*. Estas son las réplicas que están bien conectadas y que soportan un gran número concurrente de usuarios. `ftp.FreeBSD.org` son en realidad dos máquinas, una situada actualmente en Dinamarca y la otra en los Estados Unidos. *NO* es un sitio maestro y no debería utilizarse para replicar desde él. Muchos documentos online guían a los usuarios "interactivos" a `ftp.FreeBSD.org` por lo que lo sistemas automáticos de replicado deberían encontrar una máquina diferente desde la que replicar.

Existe además una jerarquía de réplicas que se describe en base a niveles (*tiers*). Los sitios maestros no están referenciados pero se pueden describir como *Tier-0*. Los mirrors que replican de estos sitios se consideran *Tier-1*, réplicas de los *Tier-1* son *Tier-2*, etc. Se recomienda que los sitios oficiales estén en un *tier* bajo, pero cuanto más bajo el tier mayores son los requisitos en los términos descritos en Requisitos para las réplicas de FreeBSD. Además el acceso a las réplicas de tier bajos podría estar restringido y seguro que el acceso a los sitios maestros está restringido. La jerarquía basada en niveles no tiene relación con DNS y en general no está documentada excepto para los sitios maestros. Sin embargo, las réplicas oficiales con números bajos como 1-4 son normalmente *Tier-1* (esto es sólo una aproximación, no hay una regla).

4.2. Ok, pero, ¿de dónde debo descargar los archivos?

Bajo ninguna circunstancia deberías replicar desde `ftp.FreeBSD.org`. La respuesta corta es: desde el sitio que está más próximo a ti en términos de Internet, o el que te proporciona el acceso más rápido.

4.2.1. ¡Solo quiero crear una réplica de algún sitio!

Si no tienes ninguna intención o requisito especial, se aplica lo que se dice en Ok, pero, ¿de dónde debo descargar los archivos?. Esto significa:

1. Comprueba aquellos que proporcionan un acceso rápido (número de saltos, tiempos de ida y vuelta) y que ofrecen servicios que tienes pensado usar (como rsync).
2. Ponte en contacto con los administradores del sitio elegido, indicando tu solicitud y preguntando acerca de sus términos y políticas.
3. Configura tu replica como se describe anteriormente.

4.2.2. Soy una replica oficial, ¿cuál es el sitio correcto para mí?

En general la descripción en ¡Solo quiero crear una réplica de algún sitio! todavía es válida. Por supuesto que podrías darle importancia al hecho de que el servidor desde el que réplicas debería ser de un tier bajo. Hay otras consideraciones acerca de las réplicas *oficiales* que se describen en Replicas oficiales.

4.2.3. ¡Quiero acceder a los sitios maestros!

Si tienes buenos motivos y prerequisitos podrías necesitar acceso a alguno de los sitios maestros. El acceso a estos sitios está restringido por lo general y hay ciertas políticas de acceso especiales. Si ya

eres una réplica *oficial* eso facilita que obtengas acceso. En cualquier otro caso asegúrate de que tu país realmente necesita otra réplica. Si ya tiene tres o más pregunta primero al "administrador de zona" (hostmaster@CC.FreeBSD.org) o Lista de correo de las réplicas de FreeBSD.

Quienquiera que te haya ayudado a convertirte en una replica, un *oficial* (official) debería de haberte ayudado a obtener acceso a un host upstream adecuado, ya sea uno de los sitios maestros o a un Tier-1 adecuado. De lo contrario, puedes enviar un correo electrónico solicitando ayuda a mirror-admin@FreeBSD.org.

Hay un sitio maestro para el conjunto de archivos FTP.

4.2.3.1. ftp-master.FreeBSD.org

Este es el sitio maestro para el conjunto de archivos FTP.

`ftp-master.FreeBSD.org` proporciona acceso rsync, además de por FTP. Lee Replicar el sitio FTP.

Se alienta a las réplicas a ofrecer acceso rsync para el contenido de FTP puesto que son réplicas *Tier-1*.

5. Replicas oficiales

Las réplicas oficiales, son réplicas que

- a) tienen una entrada DNS tipo `FreeBSD.org` (normalmente un CNAME).
- b) son listados como una replica oficial en la documentación de FreeBSD (como el handbook).

Hasta aquí cómo se distinguen las réplicas oficiales. Las réplicas oficiales no son necesariamente *Tier-1*. Sin embargo probablemente no encuentres una réplica *Tier-1* que no sea también oficial.

5.1. Requisitos específicos para las réplicas oficiales (tier-1)

No es tan sencillo establecer los requisitos para todas las réplicas oficiales, ya que el proyecto es bastante tolerante al respecto. Es más sencillo decir qué se requiere de una *replica oficial tier-1*. El resto de réplicas oficiales deberían de considerar esto como algo que se *debe de cumplir*.

Las réplicas tier-1 están obligadas a:

- disponer del conjunto de archivos completo
- permitir el acceso a otras réplicas
- proporcionar acceso FTP y rsync

Es más, los administradores deberían suscribirse a Lista de correo de las réplicas de FreeBSD. Visita See este enlace para los detalles sobre cómo suscribirse.

Para un administrador de hub, especialmente de hubs Tier-1, es *muy* importante

> comprobar la release schedule para la siguiente release. Es importante porque te dirá cuándo está planificada para salir la nueva versión y por lo tanto te da tiempo para prepararte para el pico de tráfico que le sigue.
>
> También es importante que los administradores de las réplicas intenten mantenerlas lo más actualizas posible (esto es crucial para las réplicas Tier-1). Si el Mirror1 no se actualiza durante un tiempo, las réplicas de tiers menores comenzarán a replicar los datos antiguos del Mirror1 y, por lo tanto, empezará una espiral descencente... ¡Mantén tus réplicas actualizadas!

5.2. ¿Entonces, cómo convertirse en una replica oficial?

No estamos aceptando nuevas réplicas en este momento.

6. Algunas estadísticas de las réplicas

Aquí dispone de enlaces a las páginas de estadísticas de sus réplicas favoritas (también conocidas como las únicas que desean proporcionarlas).

6.1. Estadísticas de los sitios FTP

- 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)