

Espelhando o FreeBSD

Resumo

Um artigo em andamento sobre como espelhar o FreeBSD, destinado à administradores de hubs.

Índice

1. Informações de contato	1
2. Requisitos para um site espelho do FreeBSD	1
3. Como espelhar o FreeBSD	4
4. De onde espelhar	6
5. Espelhos Oficiais	8
6. Algumas estatísticas dos sites espelho	9



Nós não estamos aceitando novos sites espelho neste momento.

1. Informações de contato

Os Coordenadores do Sistema de Espelhamento podem ser contatados pelo email mirror-admin@FreeBSD.org. Existe também uma [lista de discussão de sites espelho do FreeBSD](#).

2. Requisitos para um site espelho do FreeBSD

2.1. Espaço em disco

O espaço em disco é um dos requisitos mais importantes. Dependendo do conjunto de releases, arquiteturas e grau de cobertura que você deseja espelhar, uma quantidade enorme de espaço em disco pode ser consumida. Também tenha em mente que espelhos *oficiais* provavelmente precisam ser completos. As páginas web devem ser sempre espelhadas completamente. Observe também que os números indicados aqui refletem o estado atual (para 10.4-RELEASE/11.2-RELEASE). Desenvolvimentos adicionais e novas releases só aumentarão a quantidade necessária. Também certifique-se de ter algum espaço extra (cerca de 10-20%) apenas para ter certeza de que não irá faltar espaço. Aqui estão alguns números aproximados:

- Distribuição FTP completa: 1.4 TB
- Deltas do CTM: 10 GB
- Páginas Web: 1GB

O uso atual de disco da Distribuição por FTP pode ser encontrado em <ftp://ftp.FreeBSD.org/pub/FreeBSD/dir.sizes>.

2.2. Conexão de Rede/Largura de Banda

Claro, você precisa estar conectado à Internet. A largura de banda necessária depende do uso pretendido do site espelho. Se você quiser espelhar apenas algumas partes do FreeBSD para uso na sua rede local/intranet, a demanda pode ser muito menor do que se você quiser disponibilizar os arquivos publicamente. Se você pretende se tornar um site espelho oficial, a largura de banda necessária será ainda maior. Podemos apenas dar estimativas aproximadas aqui:

- Site local, sem acesso público: basicamente sem valor mínimo, mas se for menor que < 2 Mbps, pode deixar a sincronização bem lenta.
- Site público não oficial: 34 Mbps é provavelmente um bom começo.
- Site oficial: > 100 Mbps é recomendado, e seu host deve estar conectado o mais próximo possível do seu roteador de borda.

2.3. Requisitos de Sistema, CPU, RAM

Isto depende muito do número esperado de clientes, que é determinado pela política do servidor. O dimensionamento também é afetado pelo tipo de serviços que você deseja oferecer. Serviços FTP ou HTTP simples podem não exigir uma quantidade enorme de recursos. Tenha cuidado se você fornecer o rsync. Isso pode ter um grande impacto nos requisitos de CPU e memória, já que este serviço é considerado um devorador de memória. Os exemplos a seguir, visam lhe dar uma ideia simples deste dimensionamento.

Para um site com visitação moderada o qual ofereça o serviço de rsync, você pode considerar uma CPU entre 800 MHz - 1 GHz, e pelo menos 512 MB de memória RAM. Esta é provavelmente a configuração mínima para um site espelho *oficial*.

Para um site com visitação frequente, você definitivamente vai precisar de mais memória RAM (considere 2 GB como um bom ponto de partida) e possivelmente de mais poder de processamento (CPU), o que pode significar que você precisará ir para um sistema multiprocessado (SMP).

Você também pode querer considerar um subsistema de disco rápido. As operações no repositório SVN requerem um subsistema de disco rápido (o RAID é altamente recomendado). Um controlador SCSI que possua um cache próprio também pode acelerar as coisas, já que a maioria desses serviços incorrem em um grande número de pequenas modificações no disco.

2.4. Serviços para oferecer

Todo site espelho é obrigado a disponibilizar um conjunto de serviços básicos. Em adição a estes serviços obrigatórios, existe um grande número de serviços opcionais aos quais o administrador do servidor pode optar por oferecer. Esta sessão irá detalhar quais serviços você pode oferecer, bem como implementá-los.

2.4.1. FTP (necessário para o conjunto de arquivos do FTP)

Este é um dos serviços mais básicos, e ele é obrigatório em todos os sites espelhos que oferecem acesso público às distribuições via FTP. O acesso ao FTP deve ser anônimo, e não é permitido o uso de nenhum controle nas taxas de upload/download (o que seria uma coisa ridícula de qualquer maneira). Não é necessário ter o upload de arquivos habilitado (e isso *nunca* deve ser permitido na área onde os arquivos do FreeBSD são mantidos). Os arquivos do FreeBSD devem ficar disponíveis sob o caminho /pub/FreeBSD.

Existem diversos softwares disponíveis que podem ser configurados para operar como um servidor de FTP anônimo. Por exemplo (em ordem alfabética)

- [/usr/libexec/ftpd](#): o próprio ftpd do FreeBSD pode ser usado. Não deixe de ler o manual do [ftpd\(8\)](#).
- [ftp/ncftpd](#): Um pacote comercial, grátis para uso educacional.
- [ftp/oftpd](#): Um ftpd projetado tendo a segurança como foco principal.
- [ftp/proftpd](#): Um ftpd modular e muito flexível.
- [ftp/pure-ftpd](#): Outro ftpd desenvolvido com segurança em mente.
- [ftp/twoftpd](#): Mais um ftpd desenvolvido com segurança em mente.
- [ftp/vsftpd](#): Um ftpd "muito seguro".

O ftpd nativo do FreeBSD, o proftpd, e talvez o ncftpd são alguns dos servidores de FTP mais utilizados. Os demais não possuem uma grande base de usuários entre os sites espelhos. Um item a ser considerado é que você pode precisar de flexibilidade para controlar quantas conexões simultâneas serão permitidas no servidor, limitando desta forma o consumo do seu link IP e dos demais recursos do sistema.

2.4.2. Rsync (opcional para o conjunto de arquivos do FTP)

O rsync é muitas vezes oferecido para acesso ao conteúdo da área de FTP de um site espelho do FreeBSD, desta forma outros sites espelhos podem utilizar o seu sistema como fonte para se espelhar. O protocolo do rsync é diferente do FTP em muitos aspectos. Ele é muito mais amigável em relação ao consumo de banda IP, uma vez que quando um arquivo é alterado ao invés de transferi-lo por completo novamente, ele transfere apenas as diferenças entre as duas versões do arquivo. O rsync requer uma grande quantidade de memória para cada instância. A quantidade de memória alocada depende do tamanho do modulo sincronizado em termos do número de diretórios e de arquivos. O rsync pode utilizar [rsh](#) e o [ssh](#) (que agora é padrão) para transporte dos dados, ou então utilizar o seu próprio protocolo para acesso stand-alone (este é o método preferido para um servidor público de Rsync). Obrigatoriedade de autenticação, limites ao número de conexões simultâneas e outras restrições podem ser aplicadas. Há apenas um pacote de software disponível para se implementar um servidor de Rsync:

- [net/rsync](#)

2.4.3. HTTP (necessário para as páginas web, opcional para o conjunto de arquivos do FTP)

Se você deseja disponibilizar as páginas web do FreeBSD, você vai precisar instalar um servidor web. Opcionalmente você poderá oferecer acesso a sua árvore de FTP via HTTP. A escolha do software do servidor web é uma escolha do administrador do site espelho. As opções mais populares são:

- [www/apache24](#): O Apache é o servidor web mais amplamente utilizado na internet. Ele é usado extensivamente pelo projeto FreeBSD.
- [www/boa](#): O Boa é um servidor HTTP single-task. Ao contrário dos servidores Web tradicionais, o seu processo não se divide para cada conexão de entrada e nem cria muitas cópias de si mesmo para lidar com várias conexões. Entretanto, ele fornece um desempenho excelente para conteúdo puramente estático.
- [www/ Cherokee](#): O Cherokee é um servidor web muito rápido, flexível e fácil de configurar. Ele suporta as tecnologias difundidas atualmente: FastCGI, SCGI, PHP, CGI, conexões criptografadas por SSL/TLS, vhosts, autenticação de usuários, codificação on the fly e balanceamento de carga. Ele também gera arquivos de log compatíveis com o Apache.
- [www/lighttpd](#): O lighttpd é um servidor web seguro, rápido, compatível com os padrões e muito flexível o qual foi otimizado para ambientes de alto desempenho. Tem um consumo de memória muito baixo em comparação com outros servidores Web, bem como um baixo consumo de CPU.
- [www/nginx](#): O nginx é um servidor web de alto desempenho com baixo consumo de memória e recursos-chave para construir uma infraestrutura web moderna e eficiente. Os recursos incluem um servidor HTTP, proxy reverso de HTTP e email, armazenamento em cache, balanceamento de carga, compactação, limitação de solicitações, multiplexação e reutilização de conexões, descarregamento de SSL e streaming de mídia por HTTP.
- [www/thttpd](#): Se você estiver servindo uma grande quantidade de conteúdo estático, você pode descobrir que usar uma aplicação como o thttpd é mais eficiente do que outros servidores web. Ele também é otimizado para ter um excelente desempenho no FreeBSD.

3. Como espelhar o FreeBSD

Ok, agora você conhece os requisitos e sabe como oferecer os serviços, mas não sabe como obtê-los. :-) Esta seção explica como realmente espelhar as várias partes do FreeBSD, quais ferramentas utilizar e de onde espelhar.

3.1. Espelhando o site FTP

A área FTP possui a maior quantidade de dados a serem espelhados. Ela inclui os *conjuntos de distribuição* necessários para a instalação em rede, os *branches* que são snapshots das árvores de código fonte, as *Imagens ISO* para gravar CD-ROMs com a distribuição de instalação, um sistema de arquivos ativo e um snapshot da árvore de ports. E claro, tudo isso para as várias versões do FreeBSD e diversas arquiteturas.

A melhor maneira de espelhar a área FTP é com o rsync. Você pode instalar o port [net/rsync](#) e então

usar o rsync para sincronizar com seu host upstream. O rsync já foi mencionado em [Rsync \(opcional para o conjunto de arquivos do FTP\)](#). Como o acesso rsync não é necessário, seu site de upstream preferido pode não permitir isso. Talvez você precise procurar um pouco mais para localizar um site que permita acesso por rsync.



Como o número de clientes rsync terá um impacto significativo na performance do servidor, a maioria dos administradores impõe limitações em seus servidores. Para um espelho, você deve perguntar ao mantenedor do site com o qual você está sincronizando sobre sua política, e talvez pedir uma exceção para o seu host (já que você também é um site espelho).

Um exemplo de linha de comando para espelhar o FreeBSD pode ser verificada abaixo:

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

Consulte a documentação do rsync, que também está disponível em <http://rsync.samba.org/>, sobre as várias opções a serem usadas com o rsync. Se você sincronizar o módulo inteiro (diferentemente dos subdiretórios), esteja ciente de que o diretório do módulo (aqui "FreeBSD") não será criado, então você não pode omitir o diretório de destino. Além disso, você pode querer configurar um script que chame tal comando via [cron\(8\)](#).

3.2. Espelhando as páginas WWW

O site do FreeBSD deve ser espelhado apenas via rsync.

Uma linha de comando para espelhar o site do FreeBSD pode parecer com:

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

3.3. Espelhando os Pacotes

Devido a exigências muito altas de largura de banda, armazenamento e administração, o Projeto FreeBSD decidiu não permitir espelhos públicos de pacotes. Para sites com muitas máquinas, pode ser vantajoso executar um proxy HTTP para fazer cache do [pkg\(8\)](#). Alternativamente, pacotes específicos e suas dependências podem ser baixados executando algo assim:

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

Quando esses pacotes forem baixados, os metadados do repositório devem ser gerados executando:

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

Uma vez que os pacotes tenham sido baixados e os metadados para o repositório tenham sido

gerados, sirva os pacotes até as máquinas clientes via HTTP. Para obter informações adicionais, consulte as páginas de manual do [pkg\(8\)](#), mais especificamente a página [pkg-repo\(8\)](#).

3.4. Com que frequência eu devo atualizar o conteúdo do meu espelho?

Todo site espelho deve ser atualizado no mínimo uma vez por dia. Certamente, um script com bloqueio para impedir que várias execuções ocorram ao mesmo tempo será necessário para executar a partir do [cron\(8\)](#). Como quase todo administrador faz isso à sua maneira, instruções específicas não podem ser fornecidas. Mas poderia ser algo como:

1. Coloque o comando para executar seu aplicativo de espelhamento em um script. Recomenda-se o uso de um script simples `/bin/sh`.
2. Adicione alguns redirecionamentos de saída para que as mensagens de diagnóstico sejam registradas em um arquivo.
3. Teste se o seu script funciona. Verifique os logs.
4. Use o [crontab\(1\)](#) para adicionar o script ao [crontab\(5\)](#) do usuário apropriado. Este deve ser um usuário diferente daquele sob o qual o daemon FTP está sendo executado, de forma que, se as permissões de arquivo dentro de sua área FTP não forem legíveis por todos, esses arquivos não poderão ser acessados por FTP anônimo. Isto é usado para fazer o "stage" de uma release - assegurando que todos os sites espelhos oficiais tenham todos os arquivos necessários da release no dia do lançamento.

Aqui estão alguns agendamentos recomendados:

- Conjunto de arquivos FTP: diariamente
- Páginas WWW: diariamente

4. De onde espelhar

Esta é uma questão importante. Então, esta seção vai gastar algum esforço para explicar mais a fundo. Nós diremos isso várias vezes: sob nenhuma circunstância você deve espelhar a partir do <ftp.FreeBSD.org>.

4.1. Algumas palavras sobre a organização

Os espelhos são organizados por país. Todos os espelhos oficiais possuem uma entrada de DNS no formato <ftpN.CC.FreeBSD.org>. O *CC* (ou seja, o código do país) é o *domínio de nível superior* (TLD) do país onde esse espelho está localizado. *N* é um número, dizendo que o host seria o espelho *Nth* daquele país. (O mesmo se aplica a <wwwN.CC.FreeBSD.org>, etc.) Há espelhos sem partes *CC*. Esses são os sites espelhos que são muito bem conectados e permitem um grande número de usuários simultâneos. O <ftp.FreeBSD.org> é na verdade duas máquinas, uma atualmente localizada na Dinamarca e outra nos Estados Unidos. Este *NÃO* é um site mestre e nunca deve ser usado para espelhar. Muitos documentos on-line conduzem os usuários "interativos" para <ftp.FreeBSD.org>, portanto os sistemas automatizados de espelhamento devem utilizar uma máquina diferente para

se espelhar.

Além disso, existe uma hierarquia de espelhos, descrita em termos de *camadas*. Os sites mestres não são referenciados, mas podem ser descritos como *Tier-0*. Espelhos que espelham desses sites podem ser considerados *Tier-1*, espelhos de *Tier-1*mirrors, são *Tier-2*, etc. Os sites oficiais são encorajados a ter um *nível* baixo, mas quanto mais baixo o nível, maiores os requisitos em termos, conforme descrito em [Requisitos para um site espelho do FreeBSD](#). Também o acesso a espelhos de baixo nível pode ser restrito, e o acesso a sites mestres é definitivamente restrito. A *tier-hierarchy* não é refletida pelo DNS e geralmente não é documentada em nenhum lugar, exceto nos sites mestres. No entanto, os espelhos oficiais com números baixos como 1-4, geralmente são *Tier-1* (isso é apenas uma dica, e não há regra).

4.2. Ok, mas de onde devo baixar os arquivos agora?

Em nenhuma circunstância você deve espelhar a partir de <ftp.FreeBSD.org>. A resposta simples é: do site que está mais próximo de você em termos de Internet ou que lhe ofereça o acesso mais rápido.

4.2.1. Eu só quero espelhar de algum lugar!

Se você não tem nenhuma intenção ou requisito especial, a declaração em [Ok, mas de onde devo baixar os arquivos agora?](#) se aplica. Isso significa:

1. Verifique quais fornecem acesso mais rápido (número de saltos, tempos de ida e volta) e ofereçam os serviços que você pretende usar (como rsync).
2. Entre em contato com os administradores do site escolhido, informando sua solicitação e perguntando sobre seus termos e políticas.
3. Configure o seu espelho conforme descrito acima.

4.2.2. Eu sou um espelho oficial, qual é o site certo para mim?

Em geral, a descrição em [Eu só quero espelhar de algum lugar!](#) ainda se aplica. É claro que você pode querer colocar algum peso no fato de que seu upstream deve ser de nível baixo. Existem algumas outras considerações sobre os espelhos *oficiais* que são descritos em [Espelhos Oficiais](#).

4.2.3. Eu quero acessar os sites principais!

Se você tiver boas razões e pré-requisitos, poderá querer e obter acesso a um dos sites mestres. O acesso a esses sites é geralmente restrito e existem políticas especiais para acesso. Se você já é um espelho *oficial*, isso certamente irá ajudar você a obter acesso. Em qualquer outro caso, certifique-se de que seu país realmente precisa de outro espelho. Se já tiver três ou mais, pergunte ao "administrador de região" (hostmaster@CC.FreeBSD.org) ou [Listas de discussão de sites espelho do FreeBSD](#) primeiro.

Quem ajudou você a se tornar um espelho *oficial* deve ajudar você a obter acesso a um host de upstream apropriado, seja um dos sites mestres ou um site Tier-1 adequado. Caso contrário, você pode enviar um email para mirror-admin@FreeBSD.org para solicitar ajuda com isso.

Existe um site principal para o conjunto de arquivos FTP.

4.2.3.1. ftp-master.FreeBSD.org

Este é o site principal do conjunto de arquivos FTP.

O ftp-master.FreeBSD.org fornece acesso por rsync, além do acesso normal por FTP. Consulte [Espelhando o site FTP](#).

Os espelhos também são encorajados a permitir acesso por rsync para o conteúdo FTP, já que eles são espelhos de *Tier-1*.

5. Espelhos Oficiais

Espelhos oficiais são os espelhos que

- a) tem uma entrada de DNS FreeBSD.org (geralmente um CNAME).
- b) são listados como um espelho oficial na documentação do FreeBSD (como no Handbook).

Até agora, para distinguir espelhos oficiais. Espelhos oficiais não são necessariamente espelhos *Tier-1*. No entanto, você provavelmente não encontrará um espelho *Tier-1* que também não é oficial.

5.1. Requisitos especiais para sites espelhos oficiais (Tier-1)

Não é tão fácil declarar os requisitos para todos os sites espelhos oficiais, uma vez que o projeto é um pouco tolerante aqui. É mais fácil dizer o que um *site espelho oficial Tier-1* precisa ter. Todos os outros sites espelhos oficiais podem considerar isto como um grande *deve*.

Os sites espelhos Tier-1 precisam:

- ter o conjunto completo de arquivos
- permitir acesso a outros sites espelho
- fornecer acesso por FTP e rsync

Além disso, os administradores devem estar inscritos nas [listas de discussão de sites espelho do FreeBSD](#). Consulte [este link](#) para obter detalhes em como se inscrever.

É *muito* importante para um administrador de hub, especialmente administradores de hub de Tier-1, verificar o [calendário de lançamentos](#) para a próxima versão do FreeBSD. Isto é importante porque lhe dirá quando o próximo lançamento está programado para sair, e assim lhe dará tempo para se preparar para o grande pico de tráfego que virá.



Também é importante que os administradores do hub tentem manter seus espelhos o mais atualizados possível (novamente, ainda mais importante para os espelhos Tier-1). Se o Mirror1 não for atualizado por um tempo, os espelhos de camada inferior começarão a espelhar os dados antigos do Mirror1 e, portanto, iniciará uma espiral descendente... Mantenha seus espelhos atualizados!

5.2. Como se tornar um site espelho oficial?

Não estamos aceitando novos sites espelhos neste momento.

6. Algumas estatísticas dos sites espelho

Aqui estão os links para as páginas de estatísticas dos seus sites espelho favoritos (também conhecidos como os únicos que têm a boa vontade de fornecer as estatísticas).

6.1. Estatísticas do site FTP

- <ftp.is.FreeBSD.org> - hostmaster@is.FreeBSD.org - (Largura de Banda) (Processos FTP) (Processos HTTP)
- <ftp2.ru.FreeBSD.org> - mirror@macomnet.ru - (Largura de Banda) (Usuários HTTP e FTP)