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 fazer o espelhamento 6
5. Espelhos Oficiais 8
6. Algumas estatísticas dos sites espelho 9

Não estamos aceitando novos espelhos da comunidade neste momento.

1. Informações de contato

Os coordenadores do Sistema de Espelhamento podem ser contatados por e-mail através de mirror-admin@FreeBSD.org. Também existe a lista de discussão sobre sites espelhos 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 versões, arquiteturas e grau de completude que você deseja espelhar, uma enorme quantidade de espaço em disco pode ser consumida. Além disso, espelhos *oficiais* provavelmente precisam ser completos. As páginas da web devem sempre ser espelhadas completamente. Observe também que os números aqui declarados refletem o estado atual (em 12.0-RELEASE/11.3-RELEASE). O desenvolvimento e os lançamentos futuros só aumentarão a quantidade necessária. Certifique-se também de manter cerca de 10-20% de espaço extra apenas para garantir. 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 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 não há mínimo, mas menos de 2 Mbps pode tornar a sincronização muito lenta.
- Site público não oficial: 34 Mbps provavelmente é um bom começo.
- Site oficial: é recomendado ter mais de 100 Mbps, e seu host deve estar conectado o mais próximo possível do seu roteador de borda.

2.3. Requisitos de Sistema, CPU, RAM

Uma coisa que isso depende é o número esperado de clientes, que é determinado pela política do servidor. Também é afetado pelos tipos de serviços que você deseja oferecer. Serviços FTP ou HTTP simples podem não requerer uma quantidade enorme de recursos. Tenha cuidado se você fornecer o rsync. Isso pode ter um grande impacto nos requisitos de CPU e memória, pois ele é considerado um devorador de memória. Os dados seguintes são apenas exemplos para dar uma ideia aproximada.

Para um site moderadamente visitado que oferece rsync, você pode considerar uma CPU atual com cerca de 800 MHz - 1 GHz e pelo menos 512MB de RAM. Isso provavelmente é o mínimo que você deseja para um site *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 deve considerar um subsistema de disco rápido. As operações no repositório SVN requerem um subsistema de disco rápido (RAID é altamente recomendado). Um controlador SCSI que tenha seu próprio cache também pode acelerar as coisas, uma vez que a maioria desses serviços incorre em um grande número de pequenas modificações no disco.

2.4. Serviços para oferecer

Cada site espelho deve ter um conjunto de serviços básicos disponíveis. Além desses serviços obrigatórios, há vários serviços opcionais que os administradores de servidor podem escolher oferecer. Esta seção explica quais serviços você pode fornecer e como implementá-los.

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

Este é um dos serviços mais básicos e é necessário para cada espelho que oferece distribuições por

meio de FTP publico. O acesso FTP deve ser anônimo e não são permitidas proporções de upload/download (algo ridículo de qualquer maneira). A capacidade de upload não é necessária (e *nunca* deve ser permitida para o espaço de arquivos do FreeBSD). Além disso, os arquivos do FreeBSD devem estar 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. Certifique-se de ler ftpd(8).
- ftp/ncftpd Um pacote comercial, gratuito para uso educacional.
- ftp/oftpd: Um servidor FTP projetado com segurança como foco principal.
- ftp/proftpd: Um ftpd modular e muito flexível.
- ftp/pure-ftpd: Outro servidor FTP desenvolvido com segurança em mente.
- ftp/twoftpd: Como acima.
- ftp/vsftpd: O ftpd "muito seguro".

O `ftpd` do FreeBSD, o `proftpd` e talvez o `ncftpd` estão entre os FTPds mais comumente usados. Os outros não têm uma grande base de usuários entre os sites espelho. Uma coisa a considerar é que você pode precisar de flexibilidade para limitar quantas conexões simultâneas são permitidas, limitando assim quanto de largura de banda de rede e recursos do sistema são consumidos.

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

O Rsync é frequentemente oferecido para acesso ao conteúdo da área FTP do FreeBSD, para que outros sites espelho possam usar seu sistema como fonte. O protocolo é diferente do FTP de muitas maneiras. É muito mais amigável para a largura de banda, pois apenas as diferenças entre os arquivos são transferidas em vez de arquivos inteiros quando são alterados. O Rsync requer uma quantidade significativa de memória para cada instância. O tamanho depende do tamanho do módulo sincronizado em termos do número de diretórios e arquivos. O Rsync pode usar `rsh` e `ssh` (agora padrão) como transporte ou usar seu próprio protocolo para acesso autônomo (este é o método preferido para servidores Rsync públicos). A autenticação, limites de conexão e outras restrições podem ser aplicadas. Há apenas um pacote de software disponível:

- net/rsync

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

Se você deseja oferecer as páginas web do FreeBSD, precisará instalar um servidor web. Você pode opcionalmente oferecer o conjunto de arquivos FTP via HTTP. A escolha do software do servidor web é deixada a critério do administrador do espelho. Algumas das opções mais populares são:

- www/apache24: O Apache ainda é um dos servidores da Web mais amplamente implantados na Internet. Ele é usado extensivamente pelo Projeto FreeBSD.
- www/boa: O Boa é um servidor HTTP de single-task. Ao contrário dos servidores web tradicionais, ele não faz um fork para cada conexão recebida, nem faz muitas cópias de si

mesmo para lidar com várias conexões. contudo, ele deve fornecer um desempenho consideravelmente melhor para conteúdo puramente estático.

- www/cherokee: O Cherokee é um servidor web muito rápido, flexível e fácil de configurar. Ele suporta as tecnologias amplamente utilizadas hoje em dia: FastCGI, SCGI, PHP, CGI, conexões criptografadas 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 borda de alta performance com uma pegada de memória baixa e recursos importantes para construir uma infraestrutura web moderna e eficiente. Os recursos incluem um servidor HTTP, proxy reverso HTTP e de e-mail, cache, balanceamento de carga, compressão, limitação de solicitações, multiplexação e reutilização de conexões, offload SSL e streaming de mídia HTTP.
- www/thttpd: Se você vai servir uma grande quantidade de conteúdo estático, pode descobrir que usar um aplicativo como o thttpd é mais eficiente do que outros. 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 que precisa ser espelhada. Inclui os *conjuntos de distribuição* necessários para instalação via rede, os *branches*, que são na verdade snapshots dos diretórios de código fonte, as *imagens ISO* para gravar CD-ROMs com a distribuição de instalação, um sistema de arquivos live (ao vivo) e um snapshot da árvore de ports. Tudo, é claro, para várias versões do FreeBSD e várias arquiteturas.

A melhor maneira de espelhar a área FTP é com o rsync. Você pode instalar o pacote pelo port: net/rsync[] e, em seguida, 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 obrigatório, seu site upstream preferido pode não permiti-lo. Você pode precisar procurar um pouco para encontrar um site que permita o acesso rsync.

> Como o número de clientes rsync terá um impacto significativo no processamento do servidor, a maioria dos administradores impõe limitações em seu servidor. Para um espelho, você deve perguntar ao mantenedor do site de onde está sincronizando sobre sua política e talvez pedir uma exceção para seu host (já que você é um 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 (ao contrário de subdiretórios), esteja ciente de que o diretório do módulo (aqui "FreeBSD") não será criado, portanto, você não pode omitir o diretório de destino. Além disso, você pode querer configurar um framework de script que chame esse comando via cron(8).

3.2. Espelhando as páginas WWW

Desde a migração da documentação para Hugo/Asciidoctor em 25/01/2021, a sincronização do site com rsync não funciona mais.

Existem estudos em andamento para implementar um espelho do website com infraestrutura oficial.

Para os antigos espelhos do website, uma maneira de obter o espelhamento hoje em dia é construir o site localmente com o endereço correspondente em que será hospedado.

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

Verifique mais detalhes sobre as ferramentas de construção no livro Primer do Projeto de Documentação do FreeBSD para Novos Contribuidores.

Observe que o site foi dividido em www.FreeBSD.org e docs.FreeBSD.org, e existem links entre eles; além disso, no momento, a variável `HUGO_baseURL` não cobrirá todos os links, dessa forma, o espelhamento do site não é encorajado.

3.3. Espelhando os Pacotes

Devido aos requisitos muito altos de largura de banda, armazenamento e administração, o Projeto FreeBSD decidiu não permitir espelhos públicos de pacotes pré compilados. Para sites com muitas máquinas, pode ser vantajoso executar um proxy HTTP com cache habilitado para o processo pkg(8). Alternativamente, pacotes específicos e suas dependências podem ser baixados executando algo como o seguinte:

```
% 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
```

Depois que os pacotes forem baixados e os metadados do repositório forem gerados, sirva os

pacotes para as máquinas clientes via HTTP. Para obter informações adicionais, consulte as páginas man do pkg(8), especificamente a página pkg-repo(8).

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

Todo espelho deve ser atualizado pelo menos uma vez por dia. Certamente, um script com bloqueio para evitar que várias execuções ocorram ao mesmo tempo será necessário para ser executado pelo cron(8). Como cada administrador faz isso de sua própria maneira, instruções específicas não podem ser fornecidas. Este processo poderia funcionar , por exemplo, desta forma:

1. Coloque o comando para executar sua aplicação de espelhamento em um script. O uso de um script `/bin/sh` simples é recomendado.
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 do que seu daemon FTP é executado, para que, se as permissões de arquivo dentro de sua área FTP não forem legíveis por todo o mundo, esses arquivos não possam ser acessados pelo FTP anônimo. Isso é usado para "preparar" as releases - garantindo que todos os sites de espelho oficiais tenham todos os arquivos da release necessários no dia do seu lançamento.

Aqui estão alguns agendamentos recomendados:

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

4. De onde fazer o espelhamento

Esta é uma questão importante. Portanto, esta seção dedicará algum esforço para explicar o os motivos que estão por trás desta orientação. Diremos isso várias vezes: em nenhuma circunstância você deve espelhar a partir de `ftp.FreeBSD.org`.

4.1. Algumas palavras sobre a organização

Os espelhos são organizados por país. Todos os espelhos oficiais têm uma entrada DNS com a forma `ftpN.CC.FreeBSD.org`. *CC* (ou seja, código do país) é o *domínio de nível superior* (TLD) do país onde esse espelho está localizado. *N* é um número, indicando que o host seria o *N-ésimo* espelho nesse país. (O mesmo se aplica a `wwwN.CC.FreeBSD.org`, etc.) Existem espelhos sem a parte *CC*. Esses são os sites de espelho que estão muito bem conectados e permitem um grande número de usuários simultâneos. O `ftp.FreeBSD.org` na verdade são duas máquinas, uma atualmente localizada na Dinamarca e outra nos Estados Unidos. Não é um site principal e nunca deve ser usado para se

espelhar a partir dele. Muita documentação online direciona usuários "interativos" para `ftp.FreeBSD.org`, portanto, sistemas automatizados de espelhamento devem encontrar um servidor diferente para espelhar.

Além disso, existe uma hierarquia de espelhos, que é descrita em termos de *níveis*. Os sites principais não são mencionados, mas podem ser descritos como *Nível-0*. Espelhos que espelham a partir desses sites podem ser considerados *Nível-1*, espelhos de espelhos *Nível-1* são *Nível-2*, etc. Sites oficiais são incentivados a ter um nível baixo, mas quanto mais baixo o nível, maiores os requisitos como descrito em Requisitos para um site espelho do FreeBSD. O acesso a espelhos de baixo nível pode ser restrito e o acesso a sites principais é definitivamente restrito. A hierarquia de *nível* não é refletida pelo DNS e geralmente não é documentada em nenhum lugar, exceto para os sites principais. No entanto, espelhos oficiais com números baixos como 1-4, geralmente são *Nível-1* (isso é apenas uma dica aproximada 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 curta é: a partir do site que estiver mais próximo de você em termos de Internet ou que oferecer o acesso mais rápido.

4.2.1. Eu só quero espelhar de algum lugar!

Se você não tem intenções ou requisitos especiais, a afirmação em Ok, mas de onde devo baixar os arquivos agora? se aplica. Isso significa:

1. Verifique quais sites oferecem o acesso mais rápido (número de saltos, tempos de ida e volta) e oferecem 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 site espelho conforme descrito acima.

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

Em geral, a descrição em Eu só quero espelhar de algum lugar! ainda se aplica. Claro, você pode querer dar mais importância ao fato de que o site upstream deve ter um nível baixo. Existem algumas outras considerações sobre espelhos *oficiais* que são descritas em Espelhos Oficiais.

4.2.3. Eu quero acessar os sites principais!

Se você tiver bons motivos e bons pré-requisitos, poderá querer e obter acesso a um dos sites principais. O acesso a esses sites geralmente é restrito e existem políticas especiais para acesso. Se você já é um espelho *oficial*, certamente isso ajuda 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 zona" (hostmaster@CC.FreeBSD.org) ou na lista de discussão sobre sites espelhos do FreeBSD primeiro.

Quem o ajudou a se tornar um espelho *oficial* deve ter lhe ajudado a obter acesso a um host upstream apropriado, seja um dos sites principais ou um site Tier-1 adequado. Se não, você pode enviar um e-mail 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` disponibiliza acesso via rsync, além do FTP. Consulte Espelhando o site FTP para obter mais informações.

Também é encorajado que os espelhos permitam o acesso ao rsync para o conteúdo FTP, já que eles são espelhos *Tier-1*.

5. Espelhos Oficiais

Espelhos oficiais são os espelhos que

- a) ter uma entrada 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 necessariamente são espelhos *Tier-1*. No entanto, provavelmente você não encontrará um espelho *Tier-1* que não seja oficial.

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

Não é tão fácil estabelecer requisitos para todos os espelhos oficiais, já que o projeto é bastante tolerante nesse aspecto. É mais fácil dizer o que é exigido dos *espelhos oficiais de nível 1*. Todos os outros espelhos oficiais devem considerar isso uma grande *recomendação*.

Os sites espelhos Tier-1 precisam:

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

Além disso, os administradores devem se inscrever no lista de discussão sobre sites espelhos do FreeBSD. Consulte este link para obter detalhes sobre como se inscrever.

> É *muito* importante para um administrador de um site espelho, especialmente administradores de espelhos Tier-1, verificar o calendário de lançamento para o próximo release do FreeBSD. Isso é importante porque informará quando o próximo release está programado para sair, dando tempo para se preparar para o grande aumento de tráfego que o segue.

> É importante também que os administradores dos hubs tentem manter seus espelhos atualizados o máximo possível (novamente, ainda mais crucial para os espelhos Tier-1). Se o Mirror1 não atualizar por um tempo, os espelhos de nível inferior começarão a espelhar dados antigos do Mirror1 e, assim, começa uma espiral descendente... Mantenha seus espelhos atualizados!

5.2. Como se tornar um site espelho oficial?

Por favor, entre em contato com os Administradores do Cluster conforme documentado em https://www.FreeBSD.org/administration/#t-clusteradm.

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 de HTTP e FTP)