

Anatomia do BitTorrent

a Ciência da Computação por trás do protocolo

Paulo Cheadi Haddad Filho
Orientador: José Coelho de Pina

Trabalho de Conclusão de Curso



IME-USP

Universidade de São Paulo
São Paulo, 2013

Sumário

Capa	i
Sumário	i
Lista de Figuras	iii
Glossário	iv
Lista de tarefas pendentes	v
1 Introdução	1
2 Napster, Gnutella e BitTorrent	2
2.1 Período pré-torrent	2
2.2 Nascimento do BitTorrent	4
2.3 Mundo pós-torrent	4
3 Anatomia do BitTorrent	5
3.1 Busca por informações	5
3.2 Fontes de arquivos	5
3.3 Jogo da troca de arquivos	5
4 Conceitos de Computação no BitTorrent	6
4.1 Estruturas de dados, listas ligadas e árvores	7
4.2 Funções de hash	7
4.3 Criptografia	7
4.4 Bitfields	7
4.5 Protocolos de redes	7
4.6 Multicast	7
4.7 Roteamento de pacotes	7
4.8 Retomada de downloads	7
4.9 Conexão com a Internet	7

4.10	Threads	7
4.11	Engenharia de Software	7
5	Comentários Finais	8
6	Bibliografia	9
7	Visão Pessoal	10

Lista de Figuras

Glossário

ISP

do inglês *Internet Service Provider*; fornecedores de acesso a Internet, que são empresas que vendem serviço e equipamento que permitem o acesso de um computador pessoal acessar a Internet. [2](#)

MP3

do inglês *MPEG-1/2 Audio Layer 3*; formato patenteado de compressão de dados de áudio digital que usa um método de compressão de dados com perdas. [2](#)

Lista de tarefas pendentes

■ Como faz referência a uma coisa deste tipo?	2
■ Fazer pequena introdução aqui.	6

Capítulo 1

Introdução

Aqui vou explicar o objetivo do trabalho e o que será mostrado ao longo dele.

Capítulo 2

Napster, Gnutella e BitTorrent

Para entendermos como e por que o BitTorrent se tornou o que é hoje, devemos voltar um pouco no tempo e rever a história que precedeu à sua criação, que é o fim da década dos anos 1990.

2.1 Período pré-torrent

[1] [4]

Entre o final dos anos 80 e o início dos 90, a Internet deixou de ser uma rede de computadores usada somente por entidades governamentais, laboratórios de pesquisa e universidades, passando a ter seu acesso comercializado para o público em geral pelos [fornecedores de acesso a Internet \(ISP\)](#) [2]. Com o advento do [formato de áudio MP3 \(MP3\)](#) no final de 1991 e do seu primeiro reprodutor de áudio [MP3](#) [3] Winamp, o tráfego da Internet aumentou devido ao aumento da troca direta de arquivos.

Em 1998, dois sites de compartilhamento de músicas foram criados: o AudioGalaxy.com e o MP3.com, ambos sites de busca de MP3 sobre arquivos que seus usuários faziam upload. Ambos os sites ajudaram na popularização do MP3 como forma de se escutar músicas no dia a dia, que foi catalisado quando indústrias de equipamentos eletrônicos anunciaram seus sucessores do toca-fitas: os tocadores de mídia portáteis, popularmente conhecidos como MP3 players.

Não demorou muito tempo para a indústria da música, representada pela RIAA (Recording

Como
faz re-
ferência
a uma
coisa
deste
tipo?

Industry Association of America, Associação da Indústria de Gravação da América) começar a enxergar essa popularização do formato MP3 como um perigo real de diminuição de seus lucros com as vendas de discos, e então passou a encabeçar a frente de combate contra a troca de arquivos protegidos por direitos autorais pela Internet. Pouco tempo depois, durante o governo Clinton, o Senado americano aprova uma lei de proteção a conteúdos com direitos autorais, estabelecendo regras mais específicas sobre esses direitos autorais e se tornando obstáculo àqueles sites de compartilhamento por penalizar os usuários por infringir esses direitos.

No segundo semestre de 1999 surgiu o Napster, um serviço de compartilhamento de MP3 que inovou por desfigurar o modelo cliente-servidor ao possuir conexões entre usuários e não a servidores, criando assim a primeira rede P2P.

O sucesso foi rápido, e em dezembro de 1999 já enfrentava sua primeira ação judicial, vinda de várias grandes gravadoras. Sua popularidade foi crescendo, atingindo o auge de 13.6 milhões de usuários em fevereiro de 2001. A sua vida começou a complicar em julho de 2001, quando foi iniciada uma ação judicial por responsabilidade na distribuição de conteúdo protegido. O caso foi finalizado em setembro de 2001, obrigando a desligar o serviço de indexação, que tornou a rede P2P indisponível, e ainda a pagar 26 milhões de dólares por danos já causados e também pelos futuros, fazendo com que pedisse falência.

O sucesso do Napster, mesmo que por curto período tempo, mostrou a eficiência que redes P2P poderiam ter, e com isso novos softwares e modelos de redes foram sendo lançados, porém tentando contornar o ponto fraco do antecessor a fim de não serem novos alvos de medidas judiciais. A solução para isso foi tentar descentralizar o mecanismo de indexação e busca, que foi o calcanhar de Aquiles do Napster.

A pioneira nessa tentativa foi a rede Gnutella, que foi lançada em 2003 mas que em sua primeira versão não conseguiu manter o bom desempenho do Napster. A busca era demorada e inconsistente, pois era repassada aleatória e finitamente de peers para seus vizinhos, o que podia terminar em buscas sem resultados mesmo quando um arquivo estava sendo compartilhado por alguém conectado à rede.

continua...

2.2 Nascimento do BitTorrent

Aqui contarei a história do surgimento do protocolo.

2.3 Mundo pós-torrent

Aqui discorrerei sobre as consequências do uso do BitTorrent, desde as influências acadêmico-tecnológicas da sua difusão, bem como explicarei as questões jurídicas envolvidas no seu uso.

Capítulo 3

Anatomia do BitTorrent

Aqui mostrarei como funciona o BitTorrent de forma linear, como se fosse um humano o utilizando.

3.1 Busca por informações

Falarei quais informações um arquivo torrent pode conter e como o Transmission as utiliza, tentando mostrar detalhes da troca de mensagens dele com servidores de rastreamento.

3.2 Fontes de arquivos

Mostrarei o processamento dos dados adquiridos na seção anterior e como ele organiza a lista das fontes de arquivos usando a tabela hash DHT Kademlia.

3.3 Jogo da troca de arquivos

Explicarei o algoritmo tit-for-tat padrão do protocolo BitTorrent, que vem da Teoria dos Jogos, e como o Transmission o implementa.

Capítulo 4

Conceitos de Computação no BitTorrent

Fazer pequena introdução aqui.

Aqui mostrarei detalhes técnicos sobre as partes coadjuvantes do BitTorrent e do Transmission.

4.1 Estruturas de dados, listas ligadas e árvores

4.2 Funções de hash

4.3 Criptografia

4.4 Bitfields

4.5 Protocolos de redes

4.6 Multicast

4.7 Roteamento de pacotes

4.8 Retomada de downloads

4.9 Conexão com a Internet

4.10 Threads

4.11 Engenharia de Software

Capítulo 5

Comentários Finais

Capítulo 6

Bibliografia

- [1] Wikipedia. *File sharing* — *Wikipedia, The Free Encyclopedia*. [misc; acessado em maio de 2013]. 2013. URL: http://en.wikipedia.org/w/index.php?title=File_sharing&oldid=556034682.
- [2] Wikipedia. *Internet service provider* — *Wikipedia, The Free Encyclopedia*. [misc; accessed 29-September-2013]. 2013. URL: http://en.wikipedia.org/w/index.php?title=Internet_service_provider&oldid=573549991.
- [3] Wikipedia. *MP3* — *Wikipedia, The Free Encyclopedia*. [misc; accessed 29-September-2013]. 2013. URL: <http://en.wikipedia.org/w/index.php?title=MP3&oldid=574123988>.
- [4] Wikipedia. *Timeline of file sharing* — *Wikipedia, The Free Encyclopedia*. [misc; acessado em 28 de setembro de 2013]. 2013. URL: http://en.wikipedia.org/w/index.php?title=Timeline_of_file_sharing&oldid=571061187.

Capítulo 7

Visão Pessoal