

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
Glossário	iii
Lista de Figuras	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	12

Glossário

Linux is a generic term referring to the family of Unix-like computer operating systems that use the Linux kernel. 1

Lista de Figuras

Lista de tarefas pendentes

 Fazer pequena introdução aqui.	6
--------------------------------------------------------------------------------------------------------------------------	---

Capítulo 1

Introdução

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

Linux

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 do compartilhamento de arquivos pela Internet.

2.1 Período pré-torrent

Os primeiros métodos de acesso remoto a arquivos vieram nos anos 1970, onde era-se feito diretamente ao arquivo FTP (1985) e o IRC (1988). O Hotline (1997) ainda permitiam que seus usuários se comunicassem e enviassem arquivos. Até o fim dos anos 1990, esses métodos usados para transferir arquivos seguiam o modelo cliente-servidor. Neste modelo, o cliente é um computador que faz uma requisição de serviço ou recurso para um servidor, que se dedica a aguardar pedidos de clientes.

Com o advento da codificação MP3, os arquivos de música em formato digital passaram a ser largamente utilizados na segunda metade dos anos 1990, devido à grande redução de tamanho de arquivos de áudio nesse formato quando comparado aos outros formatos contemporâneos. Então, 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

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] Vishal Misra Alix L. H. Chow Leana Golubchik. *Improving BitTorrent: A Simple Approach*.
- [2] Walid Dabbous Ashwin Rao Arnaud Legout. *Can Realistic BitTorrent Experiments Be Performed on Clusters?*
- [3] Venkata N. Padmanabhan Ashwin R. Bharambe Cormac Herley. *Analyzing and Improving BitTorrent Performance*. Rel. téc. Microsoft Research, fev. de 2005.
- [4] *BitTorrent and fountain codes: friends or foes?*
- [5] Bo Li Cameron Dale Jiangchuan Liu. *On the Network Topologies Formed in BitTorrent Downloads*.
- [6] *Category:Technical Information* — Vuze Wiki, [Online; acessado em maio de 2013]. 2013. URL: http://wiki.vuze.com/w/Category:Technical_Information.
- [7] *Characterization of BitTorrent swarms and their distribution in the Internet*.
- [8] Jan-Ming Ho Chi-Jen Wu Cheng-Ying Li. *Improving the Download Time of BitTorrent-like Systems*.
- [9] Bram Cohen. *Incentives Build Robustness in BitTorrent*.
- [10] *Comparing BitTorrent Clients in the Wild: The Case of Download Speed*.
- [11] Adrian Popescu David Erman Dragos Ilie. *BitTorrent Session Characteristics and Models*.
- [12] *Do incentives build robustness in BitTorrent?*
- [13] *Evolution and Enhancement of BitTorrent Network Topologies*.
- [14] *Experimental Study of Broadcatching in BitTorrent*.
- [15] Aaron Grunthal. *Efficient Indexing of the BitTorrent Distributed Hash Table*.
- [16] Niklas Carlsson György Dán. *Dynamic Swarm Management for Improved BitTorrent Performance*.

- [17] *Improving BitTorrent Traffic Performance by Exploiting Geographic Locality.*
- [18] *Improving Traffic Locality in BitTorrent via Biased Neighbor Selection.*
- [19] *Index of BitTorrent Enhancement Proposals.* [Online; acessado em maio de 2013]. 2012. URL: http://www.bittorrent.org/beps/bep_0000.html.
- [20] Tomas Isdal. *Using BitTorrent for Measuring End-To-End Internet Path Characteristics.*
- [21] K. Y. Lin. *Survey on BitTorrent.* Mar. de 2006.
- [22] *Modelling BitTorrent-like systems with many classes of users.*
- [23] *On the Efficiency and Cost of Introducing QoS in BitTorrent.*
- [24] *On the Robustness of BitTorrent Swarms to Greedy Peers.*
- [25] Stefano Cucchiella Paolo Di Francesco. *Selfish strategies affecting the BitTorrent protocol.*
- [26] Nicolae Tapus Razvan Deaconescu Razvan Rughinis. *A BitTorrent Performance Evaluation Framework.*
- [27] Dan S. Wallach Scott A. Crosby. *An Analysis of BitTorrent's Two Kademlia-Based DHTs.*
- [28] Alex Sherman. *Guaranteeing Performance through Fairness in Peer-to-Peer File-Sharing and Streaming Systems.*
- [29] Anirban Basu Soumya V. L. "Modified BitTorrent Protocol and its application in Cloud Computing Environment". Em: *International Journal of Computer Science, Engineering and Applications (IJCSEA)* 2.5 (Outubro de 2012), pp. 23–31.
- [30] *Swarming: como BitTorrent Revolucionou a Internet.*
- [31] *The BitTorrent P2P File-Sharing System: Measurements and Analysis.*
- [32] Richard Thommes e Mark Coates. *BitTorrent Fairness: Analysis and Improvements.*
- [33] Umar Saif Umair Waheed Khan. *BitTorrent for the Less Privileged.*
- [34] Konstantinos Psounis Wei-Cherng Liao Fragkiskos Papadopoulos. *Performance Analysis of BitTorrent-like Systems in Heterogeneous Networks.*
- [35] Wikipedia. *File sharing — Wikipedia, The Free Encyclopedia.* [Online; acessado em maio de 2013]. 2013. URL: http://en.wikipedia.org/w/index.php?title=File_sharing&oldid=556034682.
- [36] Wikipedia. *Peer-to-peer file sharing — Wikipedia, The Free Encyclopedia.* [Online; acessado em maio de 2013]. 2013. URL: http://en.wikipedia.org/w/index.php?title=Peer-to-peer_file_sharing&oldid=558141797.
- [37] Wikipedia. *Peer-to-peer — Wikipedia, The Free Encyclopedia.* [Online; acessado em maio de 2013]. 2013. URL: <http://en.wikipedia.org/w/index.php?title=Peer-to-peer&oldid=556683045>.

- [38] Theory.org Wiki. *Main Page — Theory.org Wiki*, [Online; acessado em maio de 2013]. 2013. URL: https://wiki.theory.org/index.php?title=Main_Page&oldid=216.
- [39] Scott Wolchok. *Crawling BitTorrent DHTs for Fun and Profit*.

Capítulo 7

Visão Pessoal