File Sharing, Network Architecture, and Copyright Enforcement: An Overview*

Tilman Klumpp[†]
University of Alberta

May 2013

Abstract

This paper provides an overview of internet file sharing networks and explores the relationship between technological, economic, and legal aspects of file sharing. I chronicle the evolution of content sharing technology since the 1990s and examine the role of network architecture in a copyright holder's choice of enforcement strategy. I also describe how users and developers of file sharing networks have responded to various enforcement tactics. The target audience of this survey consists of economists and legal scholars interested in the technology and economics of file sharing networks and the enforcement of intellectual property rights on these networks.

Keywords: File sharing; peer-to-peer networks; network architecture; intellectual property rights; copyright enforcement.

JEL codes: D85; K11; K42; L82; O33; O34.

---

*I thank Phil Curry, Hugo Mialon, Paul Rubin, and Xuejuan Su for helpful comments. Some of the material contained in this essay has been used in undergraduate game theory courses that I have taught at Indiana University and Emory University. The students in these courses have provided many valuable insights that helped shape the content of this paper. Disclosure: As a graduate student, the author was a moderate user of the original Napster network and, after its shutdown, for a short time also of Gnutella. He since obtained a tenure-track job and has become a (mostly satisfied) paying subscriber of a number of streaming media services.

[†]Department of Economics, University of Alberta. 9-20 Tory Building, Edmonton, AB, Canada T6G 2H4. E-mail: klumpp@ualberta.ca

1 Introduction

File sharing is one of the most popular uses of the internet, accounting for approximately 40% of all data traffic in 2010 (Cisco Systems 2011). As the name indicates, file sharing applications let users make files on their computers available to others via the internet. While anything digital may be distributed in this way, file sharing is popular because it enables consumers to freely obtain content that they would otherwise have to purchase—that is, copyright-protected movies, music, or video games.

This essay chronicles the evolution of file sharing networks since the late 1990s and explores the relationship between technological, economic, and legal aspects of file sharing.[1] My goal is to provide a non-technical, yet concise, overview of the designs of popular file sharing networks and of the implications of network design for copyright enforcement. Economists' understanding of file sharing technology is often of the "black box" type: In goes a music album or movie, out comes a copy on somebody else's computer.[2] A look inside the black box reveals that file sharing networks have interesting economic properties stemming from their internal organization—what I call *network architecture*. I will demonstrate that network architecture impacts users' participation incentives as well as copyright holders' enforcement strategies. Conversely, network architecture has been evolving partly with the objective to escape enforcement, or to make enforcement more difficult.

Any large file sharing network must accomplish two tasks: First, it must provide users with a way of transferring digital files between one another, over the internet. Second, it must provide a directory or indexing service that enables users to search the network for content and determine its location.[3] Each of these functions can be centralized or decentralized to varying degrees, and this is what I mean by a network's architecture. Since both file transfer and directory services are essential for a well-functioning file sharing network, successful litigation against one function will shut the network down, even if the other can still be performed. Thus, whenever an essential network task is centralized, it becomes an obvious litigation target for copyright holders. As a rule, lawsuits against file sharing platforms have

[1] Krishnan *et al.* (2003) provide a short early survey of some of the economic issues related to file sharing networks, many of which are still relevant today. For a general survey of economics of digital piracy, the reader should consult Belleflamme and Peitz (2012).

[2] The black box perspective is nicely displayed in a 2009 *Economist* article on the Pirate Bay trial (The Economist 2009; see also Section 4 of this paper). The article liberally employs the term "illegal file-sharing service" to describe two very different networks (Napster and BitTorrent), two client software programs used to access yet another network (Grokster and Kazaa for the FastTrack network), as well as a directory website that lets users search for content on the BitTorrent network (The Pirate Bay).

[3] Without this second function, users would have to know each other before the file transfer can be executed. This is how individuals have traditionally shared content (e.g., by recording a mix tape for a friend or making a copy of someone else's CD). The appeal of internet file sharing networks lies in the fact that users can copy content from millions of other users without having to know who they are or where they are located.

targeted the most centralized function of a network. This pattern will be demonstrated throughout the survey. A challenge for enforcement arises, however, in fully distributed networks where all essential functions are performed by the peers themselves. In the absence of a single large player that can be sued, the next-best targets for litigation are the individual users. But since only a very small fraction of users can realistically be sued in a network of millions, the litigation risk for any one of them is negligible. Thus, the paper also explores to what extent even a small expected cost of litigation can affect a user's behavior, and what this implies for a copyright holder's enforcement strategy. Finally, the paper documents the legal strategies used against other "enablers" of file sharing (e.g., software developers and internet service providers), as well as various technical attacks that have been launched against file sharing networks.

I will survey several internet content sharing platforms that exist today, or that have existed in the past: Napster, OpenNap, eDonkey, FastTrack, Gnutella, BitTorrent, Megaupload, Rapidshare, and YouTube.[4] For each platform, I will describe its architecture, and discuss its strengths and weaknesses from the perspective of escaping enforcement as well as civil or criminal legal action taken against it. This includes several highly publicized cases, such as the 2000 Napster lawsuit, the 2003 RIAA file sharing lawsuits, the 2009 Pirate Bay trial, and the 2012 shutdown of Megaupload, as well as a number of lesser known cases.

The overview given in this primer is deliberately incomplete in three ways. First, I will, for the most part, steer clear of the broader questions of whether digital piracy is socially harmful and to what extent copyrights should be enforced in the first place. For the purpose of this paper, I will simply assume that copyright holders want to prevent copyright infringement, and focus on the question of *how* this objective can be best achieved.[5] Second, many other file sharing networks exist besides the ones surveyed here. I focus on platforms that are, or were, both widely used and whose architectures give them unique economic properties. In combination, these networks represent the major types of architectures used for file sharing on the internet. Third, while the objective of this essay is to look into the "black box" of file sharing technology, I will not provide the most accurate possible description of each network's organization. What I call network architecture is still a rather coarse summary of a network's inner workings—but one that is economically relevant. Despite this incompleteness, I believe the overview presented herein will be useful to those

[4]The FastTrack and Gnutella file sharing networks are often better known by the names of its third-party client programs (e.g., "Kazaa," "Morpheus," "Grokster," "Limewire"). Similarly, the BitTorrent network may be better known by the names of some of its third-party directory providers (e.g., "The Pirate Bay").

[5]This assumption is reasonable as long as unauthorized content distribution on file sharing networks crowds out legal sales. The literature has arrived at differing, but mostly negative, estimates of the impact of file sharing on sales (see Peitz and Waelbroeck 2004; Michel 2005, 2006; Zentner 2005, 2006; Rob and Waldfogel 2006; Oberholzer-Gee and Strumpf 2007; Hong 2007; Liebowitz 2008; and Danaher *et al.* 2012).

interested in the the intersection of file sharing technology, copyright enforcement, and economic incentives.

The paper is organized along economic characteristics of file sharing networks. This dimension is well aligned with the chronological order in which these networks emerged, the evolution of their technological properties, and copyright holders' choice of enforcement strategy. Section 2 describes the early days of peer-to-peer internet file sharing technology with centralized search—in particular, the original Napster network that existed between 1999 and 2001—and some of its semi-centralized successors. Section 3 focuses on fully distributed peer-to-peer networks, in particular the FastTrack and Gnutella networks. Section 4 describes the BitTorrent network and its reciprocity features. Section 5 examines the role of centralized file hosting platforms as alternatives to peer-to-peer technology, with a focus on Megaupload and YouTube. Section 6 concludes.

2 The early days

2.1 *Napster*

The original *Napster* service, founded in 1999 by Shawn Fanning, John Fanning and Sean Parker (who later co-founded the social networking site *Facebook*), marked the beginning of the internet file sharing era.[6] Napster was the first well-known peer-to-peer (P2P) file sharing network. The P2P design meant that files were transferred directly between two personal computers instead of being uploaded to, and then downloaded from, file servers.

Napster's novelty lay in the fact that the participants in a P2P transfer did not have to know one another. When signing in to a Napster session, a user connected to a directory that listed the shared contents of all users currently signed in to the network, together with the users' IP addresses. To obtain a file listed on the directory, a user's Napster client would contact the corresponding peer and initiate the transfer. By the same token, when a user connected to the service, a listing of the contents of his or her own shared folders was transmitted to the directory. These files became visible to the community and available for retrieval via P2P connections. In essence, Napster's core activity was the operation of this central directory (see Figure 1). The demands on Napster's own infrastructure were modest compared to a traditional file hosting service, as all actual files were stored only on the users' personal computers, not on Napster's servers.

Napster quickly gained popularity as a platform on which MP3 music files could be swapped with ease. Almost all music transferred through Napster was copy-

[6] Digital files have been shared online between personal computers ever since the first telephone modems were marketed in the late 1970s. For the purpose of this essay, I use the term "file sharing" to mean large-scale sharing of digital media files on the internet, which arguably started with Napster coming on line on June 1, 1999.

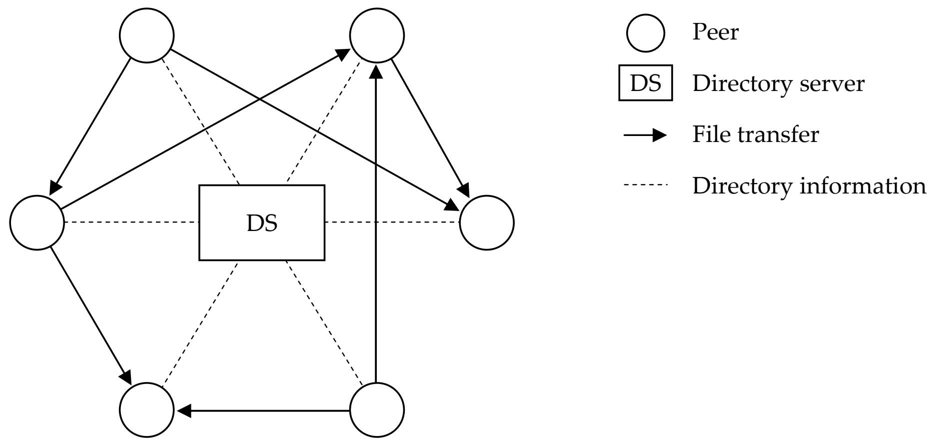

Figure 1: The original *Napster* architecture

right protected. Users would typically share songs that they had ripped from CDs they owned, as well as songs they had previously downloaded from others on the Napster network. At its peak, the Napster directory contained approximately 80 million songs. A large part of the platform's appeal to consumers came from the fact that, at the time, Napster was one of the most convenient ways to experience digital music. Instead of purchasing CDs at the store or waiting for mail-ordered CDs to arrive, Napster users enjoyed instant access to a large online music library. Furthermore, since music is an experience good, many users also valued the fact that Napster allowed them to sample music of unfamiliar artists without committing to a purchase.[7] The later commercial success of legal online music stores such as Apple's *iTunes* or Amazon's MP3 store, and flat-rate streaming media services such as *Rhapsody*, *Spotify*, or *Netflix*, proved that consumers were indeed willing to pay for this convenience without violating copyrights.

Nonetheless, the fact remained that most of the content shared on the Napster network was copyright protected. In 2000, a number of recording companies, represented by the Recording Industry Association of America (RIAA), sued Napster under the U.S. Digital Millennium Copyright Act (DMCA).[8] The RIAA's claim was that the service, while not itself infringing on their intellectual property rights, facilitated copyright infringement by its users.

A few remarks on U.S. copyright law are in order. Under U.S. common law, third parties can be held indirectly liable for copyright infringements under two legal doctrines—contributory infringment, which applies when the third party is aware

[7]There is hence an argument that file sharing can increase sales of copyrighted music by giving consumers a way to discover new music before buying it (e.g., Gopal *et al.* 2006; Peitz and Waelbroeck 2006).

[8]*A&M Records, Inc. v. Napster, Inc.* 239 F.3d 1004 (9th Cir. 2001).

of an infringing activity and capable of stopping it, and vicarious liability, which applies when the third party has control over, and benefits from, the infringing actions of others. There is an obvious efficiency reason for the law to recognize indirect liability: Clearly, it is much less costly to sue one firm whose technology enables the copyright violations of millions of users, instead of suing these (directly liable) users individually (see Landes and Lichtman 2003). Indirect liability, therefore, creates a significant legal risk for any file sharing network in which one or more critical network functions are centralized. This theme will surface on several more occasions throughout the paper.

In case of Napster, plaintiffs argued, the firm knew about the copyright violations of its users and was technically capable of preventing them, for example by shutting down its central directory. Under these conditions, it was liable for contributory infringement. Napster's defense was that its technology also had legitimate uses—a limit to the reach of indirect liability that courts had accepted in the past. In particular, in *Sony v. Universal City* (464 U.S. 417, 1984), popularly known as the "Betamax case," the Supreme Court ruled that VCR manufacturers could not be held liable for copyright violations of their customers, as VCRs had significant non-infringing uses, such as time-shifting (i.e., recording a program to be watched at a later time).

Napster lost the argument in district court and was issued injunction ordering it to shut down. The 9th Circuit U.S. Court of Appeals, while upholding most of the ruling, lifted the injunction based on "commercially significant non-infringing uses" of Napster's technology. Instead, it ordered Napster to block access to infringing material upon notification. Being unable to comply with the court's mandate on the massive scale it had grown to, Napster shut down its service in July of 2001 and began bankruptcy proceedings.[9]

2.2 Toward decentralized search: *OpenNap* and *eDonkey*

While Napster's central directory made it an obvious litigation target, at the time of its demise the transition to P2P networks with decentralized directory services had already begun.

A less centralized version of Napster, called *OpenNap*, had surfaced on the internet soon after the original network's launch in 1999. OpenNap was not related to the original Napster service or its founders even though it relied on Napster's file transfer protocol, which had been reverse engineered by a group of anonymous developers. Instead of utilizing a central directory, however, OpenNap relied on a network of smaller, interconnected directory servers that were independently operated (see Figure 2). The goal was for OpenNap to become resilient to litigation through suffi-

[9]What remained of Napster was acquired by the firm Roxio in 2002 and turned into a legal for-pay streaming service, before being sold to electronics retailer Best Buy in 2008. The ultimate end of the Napster brand came in 2011, when it was merged with the popular music subscription service Rhapsody.

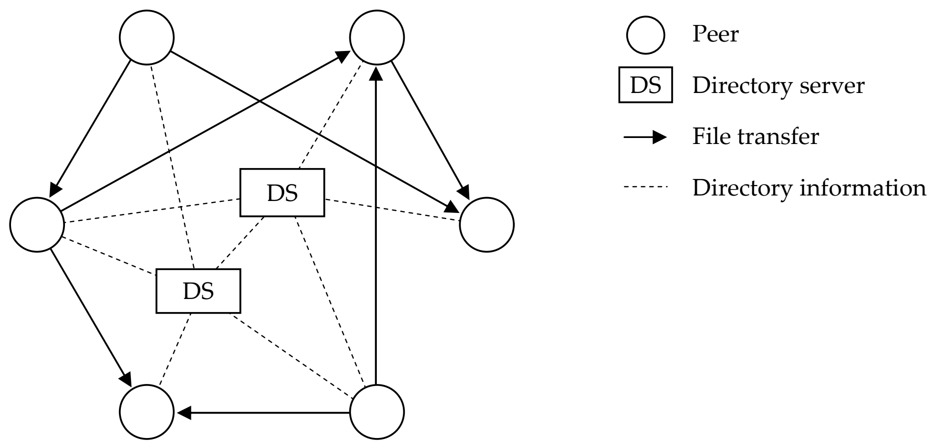

Figure 2: The *OpenNap/eDonkey* architecture

cient decentralization of its critical nodes. However, this approach was unsuccessful. Coming off its recent victory against Napster, the RIAA in 2002 began an aggressive campaign threatening OpenNap directory operators with indirect liability lawsuits. Most directories shut down out of fear of being sued, effectively dismantling the OpenNap architecture. Although a small number of OpenNap servers still operate today, the network is a marginal platform compared to both the original Napster service and later P2P networks.

A similar architecture was employed in the *eDonkey* network, developed by Jed McCaleb in 2000. Similar to OpenNap, search on the eDonkey network was performed by a network of independently operated servers. eDonkey was also the first network that allowed files to be downloaded in fragments from multiple sharing sources, a feature that was later also implemented in the BitTorrent protocol (see Section 4.1). This made it ideal for the sharing of large content files, particularly movies and computer games. The most popular eDonkey server, *Razorback2*, was ordered shut down by a Belgian court in 2006, following a request filed by the Motion Picture Association of America (MPAA).

In addition, copyright holders used two other tactics, which were also employed against the FastTrack and Gnutella networks (see Section 3.3). One was litigation against developers of eDonkey client software: In 2006, MetaMachine, the developer of a popular eDonkey client and main supporter of the eDonkey protocol, agreed to discontinue distribution of its software as part of a $30 million settlement with the RIAA. A second tactic was the technical infiltration of the network. In eDonkey's case, this involved setting up fake eDonkey directory servers that appeared on the network as having a large number of connected users but returned useless search results. Despite these challenges, eDonkey remained a viable P2P file sharing network into the late 2000s, but appears to be declining in popularity (Schulze and Mochalski 2009).

3 Full decentralization

3.1 *FastTrack* and *Gnutella*

The *Gnutella* file sharing protocol was developed in 2000 by Justin Frankel and Tom Pepper. The *FastTrack* protocol was developed around the same time by Jaan Tallin, who later also invented the popular video conferencing software *Skype*. Conceived as alternatives to Napster and OpenNap, which had come under immense legal pressure to shut down, Gnutella and FastTrack were fully decentralized: Not only were file transfers peer-to-peer, but so was the transfer of all directory information.

Gnutella and FastTrack are not organizations or services, but languages through which computers can communicate with each other.[10] Client programs speaking these languages are supplied by various third parties. Popular client programs include *Kazaa* (the original FastTrack client), *Morpheus* and *LimeWire*. When started, clients begin searching their neighborhood of nearby IP addresses for network peers. Once a peer is identified, its collection of shared files can be searched. Furthermore, by relaying search queries from one peer to the next, large (but generally not exhaustive) portions of the network can be searched for content. Figure 3 illustrates this fully distributed architecture.[11]

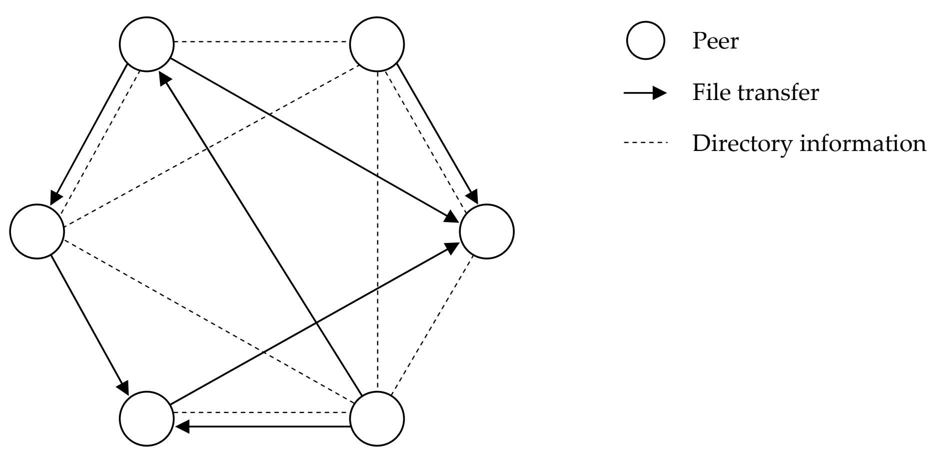

Figure 3: The *Gnutella*/*FastTrack* architecture

[10] Gnutella is an open source protocol. FastTrack is a proprietary protocol and currently owned by the Dutch company Consumer Empowerment.

[11] This is a simplified description of search on fully decentralized P2P networks. FastTrack, as well as later versions of Gnutella, organized peers into hierarchical levels, each with a different role in the search process, in order to optimize the speed with which searches are performed and to relieve the network of search-related traffic. These details are not relevant here, and Figure 3 depicts a flat peer hierarchy.

The decentralized search capabilities of Gnutella and FastTrack proved powerful enough to replicate the functionality of Napster's centralized listing service. After Napster's shutdown in 2001, large-scale unauthorized sharing of content continued on both networks (with FastTrack being the more popular platform initially). But with no single critical node to go after, the recording industry had to adjust its legal strategy when it took on these networks. The next-best targets for litigation were now the individual network users—mostly teenagers and college students—who were directly liable for copyright infringements. Thus, in 2003 the Recording Industry Association of American (RIAA) began suing FastTrack users for copyright infringement.

3.2 The RIAA file sharing lawsuits: A strategic analysis

In principle, two legal strategies against individual network users are conceivable. The first is to sue users for the act of sharing copyrighted content, and the second is to sue users for the act of downloading it.

There is a technological reason, as well as a legal reason, to prefer the former. The technological reason is that the act of a download is difficult to observe. Without monitoring a particular user's internet connection, there is no easy way to observe the occurrence of a download. On the other hand, it is straightforward to observe the incidence of sharing and to document the IP-addresses of sharers: All that is needed is a running Gnutella or FastTrack client whose list of discovered hosts can be scanned for copyright protected content. The legal reason is that, even if a download could be observed, the mere act of downloading copyrighted content is not in itself illegal. What would have to be proven in court is the downloader's knowledge of the fact that the content was protected, as well as his or her intent to keep it after the download. On the other hand, the intent to distribute copyright protected material on the internet is relatively easily established, especially when a user shares a large number of works over an extended period of time.

I will now argue that there is also an economic reason to target sharers instead of downloaders. To this end, I will set up a simple file sharing game between two users. I assume that each user possesses a digital copy of a song that the other wishes to obtain. Both users must decide whether to share their songs, and whether to download a song if shared. Thus, a user has four strategies available:

S,D	Both share and download	(a "typical user")
$S,\sim D$	Share but not download	(a "benevolent user")
$\sim S,D$	Not share but download	(a "freeloader")
$\sim S,\sim D$	Neither share nor download	(a "non-user")

I will also assume that a successful download results in a payoff of $V > 0$, and a zero payoff otherwise. Figure 4 shows the resulting payoff matrix for this game.

	S,D	$S,\sim D$	$\sim S,D$	$\sim S,\sim D$
S,D	V,V	$V,0$	$0,V$	$0,0$
$S,\sim D$	$0,V$	$0,0$	$0,V$	$0,0$
$\sim S,D$	$V,0$	$V,0$	$0,0$	$0,0$
$\sim S,\sim D$	$0,0$	$0,0$	$0,0$	$0,0$

Figure 4: A two-player file sharing game

The game has several pure-strategy Nash equilibria. The payoff-dominant equilibrium is for both players to be "typical users" and play strategy S,D.

Now suppose that users who share face the risk of litigation, and let $r > 0$ represent the expected cost of litigation. The payoff matrix under this new scenario is given in Figure 5.

	S,D	$S,\sim D$	$\sim S,D$	$\sim S,\sim D$
S,D	$V-r, V-r$	$V-r, -r$	$-r, V$	$-r, 0$
$S,\sim D$	$-r, V-r$	$-r, -r$	$-r, V$	$-r, 0$
$\sim S,D$	$V, -r$	$V, -r$	$0,0$	$0,0$
$\sim S,\sim D$	$0, -r$	$0, -r$	$0,0$	$0,0$

Figure 5: The file sharing game when sharers are sued

The previous equilibrium strategy S,D is now dominated by the freeloader strategy $\sim S,D$. Moreover, freeloading is a (weakly) dominant strategy, and this is true for all values of r as long as r is positive. Thus, in order to deter file sharing the perception of even the slightest litigation risk is sufficient.

When the RIAA began filing its lawsuits in 2003, it indeed named only four defendants, with whom it subsequently settled out of court for moderate payments. The defendant base was later expanded, but at no point did it include more than 700 individuals at the same time. In relation to the millions of users who participated in the network, this is a negligible proportion. Suing a small number of individuals for unauthorized sharing is thus a relatively inexpensive way of creating a freeloading incentive.[12]

[12]Gnutella and FastTrack provide several ways of implementing the freeloading strategy. First, most client programs allow users to disable sharing. Doing so may come at a certain penalty in terms of reduced content availability, as many programs also give sharers the option to disable transfers to users who do not themselves share. However, it is relatively easy to circumvent this problem, simply by placing a sufficient number of copyright-free media in the user's shared folder (but not copyright-protected files).

If, instead, users were to face a litigation risk for the act of downloading protected content (assuming that it can be observed and proven in court), the calculus on both the users' and the copyright holders' side changes. The payoff matrix is the now given in Figure 6 (note that a download occurs if and only if one user shares a file and the other downloads it). Observe that strategy S, D is no longer dominated by the freeloading strategy $\sim S,D$, even for large values of r. On the other hand, S,D could be dominated by the "offline" strategy $\sim S,\sim D$ or the "benevolent" strategy $S,\sim D$. For this to happen, however, $r > V$ is required. That is, the expected cost of litigation would have to exceed the value a user would otherwise derive from the file sharing network. In order to achieve the same deterrence effect as before, a much larger perception of litigation risk is required, which means that a much larger number of users would have to be targeted for litigation.[13]

	S,D	$S,\sim D$	$\sim S,D$	$\sim S,\sim D$
S,D	$V-r, V-r$	$V-r, 0$	$0, V-r$	$0,0$
$S,\sim D$	$0, V-r$	$0,0$	$0, V-r$	$0,0$
$\sim S,D$	$V-r, 0$	$V-r, 0$	$0,0$	$0,0$
$\sim S,\sim D$	$0,0$	$0,0$	$0,0$	$0,0$

Figure 6: The file sharing game when downloaders are sued

The RIAA's file sharing lawsuits clearly did not have the objective of recovering meaningful damages from the few individuals that were sued. Instead, the goal was to change user behavior on decentralized file sharing networks. As shown above, there is a strong economic argument in favor suing sharers of protected content, instead downloaders, to achieve the desired deterrence effect. At a modest cost to the copyright holder, this strategy creates an incentive for sharers to become freeloaders, and if sufficiently many sharers become freeloaders, then the file sharing network is, in effect, shut down.

Even before the RIAA lawsuits, freeloading and other types of "selfish" behaviors, such as not relaying search queries, occured on the Gnutella network (Adar and Huberman 2000, Shneidman and Parkes 2003, Asvanund *et al.* 2003). However, there is some empirical evidence that the lawsuits induced additional freeloading behavior. In a Pew telephone survey of U.S. internet users, self-reported file sharing

[13]If the copyright holder is constrained in the number of lawsuits that can be litigate at the same time, there may be multiple equilibria: If there are few active users, the risk of being sued for downloading may indeed be a strong enough deterrent. On the other hand, there is strength in numbers, so that with sufficiently many users the risk of being sued becomes negligible and does no longer act as a deterrent. The stylized two-player game examined above does not permit a formal analysis of this interesting aspect of file sharing litigation. The point here is simply that it is easier to deter sharing than downloading.

dropped by about half during the year 2003 (Wingfield 2004). Hughes *et al.* (2005) found that the proportion of freeloading Gnutella users increased significantly between 2000 and 2004, although no causal link is established to the RIAA lawsuits. Bhattacharjee *et al.* (2006) found evidence that that the announcement of legal action by the RIAA, as well as the lawsuits themselves, led a majority of users to decrease the number of shared files; however, there were some differences in the response of substantial sharers compared to users who shared a relatively small number of files.[14]

There may, however, exist a limit to how much sharing can realistically be deterred through the threat of legal action. Krishnan *et al.* (2008) develop a game-theoretic analysis of contributions to file sharing networks with differentiated content. In their model, users who possess content that is of high value to others are also the least likely to free-ride, even when contributions are moderately costly, for example because of legal risks. In an empirical study, Cox *et al.* (2010) find that among users of the BitTorrent network (introduced in the next section) so-called *first seeders*—users who make available the first instance of a file—are those users who hold extremely low estimates about the probability of getting caught for file sharing, or who view file sharing as a "philanthropic" activity. In a survey of European file sharers, Tushar and Rochelandet (2008) find that a desire for "cultural diversity" is associated with a propensity to share content (while perceived legal risks had a neutral effect on sharing).

Indeed, the RIAA file sharing lawsuits did not deter as many sharers as would be needed for an effective network shutdown. While some users' behavior seems to have changed, the number of users as well as files available on Gnutella and FastTrack continued to grow during the early-to-mid 2000s. It is even possible that the publicity generated by the lawsuits helped further popularize file sharing among new users. In late 2008, the RIAA announced that it would abandon its strategy of suing P2P file sharers, and instead enlist the help of internet service providers to combat illegal file sharing (McBride and Smith 2008). This effort will be discussed in more detail in Section 4.2.

3.3 Other enforcement strategies

In addition to its file sharing lawsuits, the RIAA and its members also pursued a number of other tactics to curb online file sharing on FastTrack and Gnutella. This included covert technical attacks on the networks themselves. If a file sharing network is flooded with enough bogus files masquerading as copies of popular songs

[14] I also have anecdotal evidence that the RIAA's approach had the desired effect: An informal poll of my game theory students in 2003 and 2004 at Indiana University indicated that almost all had used file sharing networks in the previous year, and that at least a quarter had become freeloaders after learning of the lawsuits. (Another quarter stated that they were not aware that freeloading was an option, but would now consider changing their strategy.)

or films, it becomes more difficult for users to find specific content. To achieve this goal, a number of copyright holders reportedly procured the services of *Overpeer*, a secretive New York City-based firm that operated an array of virtual peers to infiltrate file sharing networks in the early 2000s (Maguire 2003). Overpeer injected large amounts of corrupted data into the FastTrack network by exploiting a security hole of the FastTrack protocol.[15] The attack was successful and rendered FastTrack virtually unusable by 2005. However, it did not work well against the Gnutella network and BitTorrent network, which employ stronger authentication measures. Overpeer's flooding of the FastTrack network also did little to reduce overall file sharing traffic, as FastTrack users simply migrated to Gnutella and BitTorrent. Ultimately, the firm turned out to be a 'one-trick pony' and ceased operations in 2007.

The recording industry also sued the makers of various FastTrack and Gnutella client programs. In a long-running legal battle, media company *MGM* accused *Grokster* (maker of a FastTrack client by the same name) of facilitating copyright violations and sought damages. In 2005, the case was heard by the United States Supreme Court, which ruled against Grokster on the basis that it had failed to provide sufficient evidence of non-infringing uses of its software.[16] Grokster later paid $50 million in damages to MGM and other recording companies, and shut down. Similar battles involved *StreamCast* (maker of the Gnutella client Morpheus) as well as *Lime Group* (maker of the Gnutella and BitTorrent client LimeWire). StreamCast's legal costs forced it to declare bankruptcy in 2008. Lime Group was issued an injunction to remotely disable its clients in 2010, and agreed to pay $105 million in damages.[17]

Despite the legal defeats of developers of some of the most popular P2P client programs, Gnutella remained an active file sharing community through the 2000s. However, the fully decentralized directory system that allowed Gnutella to escape enforcement created a technical challenge of its own: An exponentially increasing portion of data traffic on Gnutella was devoted to forwarding directory queries between users instead of the actual file transfers (*search overhead*), severely degrading the performance of the network. As BitTorrent and hosting services have gained in popularity among file sharers, Gnutella has now become a ghost network.

[15] All file sharing networks authenticate, to varying degrees, the files being shared. This usually involves *hashing* of files. A hash is a digital signature, shorter than the file itself, that has the property that two files with the same hash are, with probability almost one, identical. The hashing algorithm employed in the FastTrack network was relatively weak, however, and made it easy to deliberately generate files that had the same hash as others but were actually different. This shortcoming allowed Overpeer's staff to generate fake copies of popular content that appeared legitimate to FastTrack users, but were in fact unusable.

[16] *MGM Studios, Inc. v. Grokster, Ltd.* 125 Sup. Ct. 2764 (2005). After *Sony v. Universal City*, the question remained whether, in the absence of evidence for significant non-infringing uses, equipment makers could be sued under indirect liability. The Grokster trial settled this question in the affirmative.

[17] Older versions of the software, that cannot be remotely disabled, are still in use, as is a reverse engineered version called *LimeWire Pirate Edition*.

4 Multiple sourcing and reciprocity

4.1 *BitTorrent*

The *BitTorrent* protocol, developed by Bram Cohen in 2001, represents yet another step in the evolution of file sharing networks. Like Gnutella or FastTrack, BitTorrent utilizes P2P file transfers without having to rely on a single central directory node to let users find content on the network. However, the BitTorrent protocol departs from these networks in two important ways.

First, BitTorrent divides the download process into a large number of different P2P connections over which small fragments of the downloaded file are transferred from multiple sharing sources. These are then stitched together by the BitTorrent client software to reconstitute the actual content file. Multiple sourcing (first used in the eDonkey network) significantly increases the reliability of large downloads. For example, if a sharing source goes offline during a file transfer, the download process continues uninterrupted by automatically utilizing other sources. Second, the BitTorrent protocol requires each user who is on the receiving end of a download to also be a source of the received content, for the duration of the download process. Reciprocity ensures that the supply of content on the network automatically matches its demand, provided it is *seeded* by some users. A seed is a user who shares all fragments that make up a content file. Should file become popular and a larg number of other users simultaneously decide to download it, they become sources of the same file and thus increase its supply on the network in lockstep with demand, while relieving the seed.

Multiple sourcing and reciprocity have necessitated a BitTorrent directory system that is more centralized than the networks examined in the previous section. Finding content on the BitTorrent network involves two steps. First, a user must find a descriptor document, called a *torrent file*, that describes how the desired content file has been broken up into fragments and how these fragments can be identified. The part of BitTorrent's functionality that enables searching for torrent files is called *indexing*. Second, the user's client software must locate peers on the network that can supply the fragments described in the torrent file. This part of BitTorrent's functionality is called *tracking*. Both indexing of torrent files and tracking of peers are needed for a working BitTorrent directory, and both tasks can, in principle, be decentralized. However, when BitTorrent was initially developed no method existed to quickly and reliably distribute the necessary relational data structures over a large network.[18] What evolved instead was an ecosystem of indexing and tracking servers

[18] In particular, a Gnutella/FastTrack-style host discovery process is insufficient in the BitTorrent environment. When downloading a file from multiple sources, the BitTorrent client must ensure that the fragments it receives originated from the same initial copy of the file. For example, several copies of the movie *Titanic* may be available on the network, and these will typically not be identical if they originated from different initial copies of the movie. Thus, in order to download the movie *Titanic* it is not enough to search the network for *Titanic (Part 1)*, *Titanic (Part 2)*, etc., and then patch them

operated by a number of independent third parties. These servers, taken together, constitute BitTorrent's directory system.[19] Figure 7 illustrates the network architecture underlying the BitTorrent file sharing protocol.

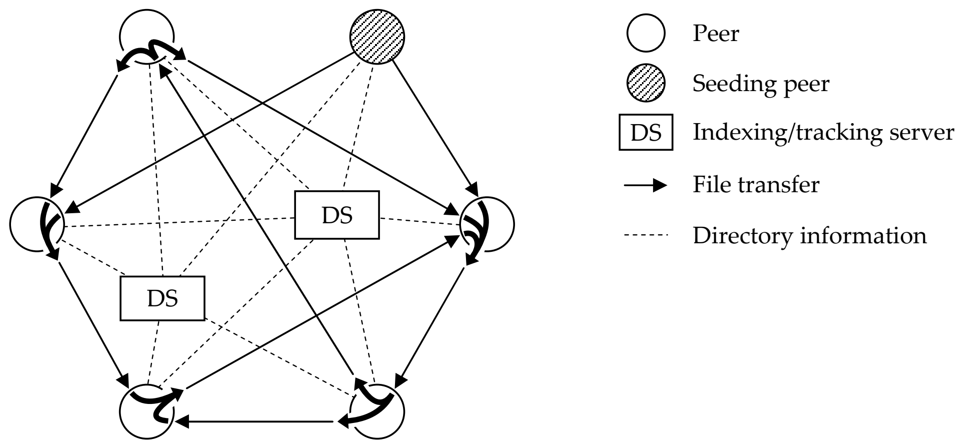

Figure 7: The *BitTorrent* architecture

The directory system's tracking nodes also fulfill a second, equally important role. The reciprocity principle requires monitoring of peers. Without monitoring, the network could be invaded by modified clients that do not share. The tracking servers in the BitTorrent architecture therefore publicize summary statistics of each peer's downloading and sharing history. Peers with poor share-to-download ratios can then be penalized by others through their refusal to share.[20] This "tit-for-tat" approach does not, however, ensure the seeding of content to the network. This does not restrict the availability of popular content that is constantly being sought by many users—the demand from these users ensures adequate supply, as explained earlier. For less popular content, however, a seeding problem may arise. Some users therefore organize themselves into subnetworks, or clubs, which can be joined by in-

together. The process must also ensure that *Titanic (Part 1)* is the first fragment, and *Titanic (Part 2)* the second fragment, of the same original file containing the movie *Titanic*. Thus, the client must obtain information not only about the location of individual file fragments, but also about how each fragment is related to the many others scattered across the network. This is why the BitTorrent protocol needs both indexing and tracking for a functioning directory: Indexing lets users find available instances of the movie *Titanic*, and tracking lets the client software find the available fragments associated with a particular instance.

[19]The same party may provide both tracking and indexing, but this is not necessary. Entry into the BitTorrent indexing and tracking market is easy, and providers fund their operations through advertising on their search pages (in case of indexing) and through donations.

[20]One loophole exists to circumvent the reciprocity rule, as BitTorrent currently exempts new peers, who do not have much to share, from it. *BitThief* (`bitthief.ethz.ch`; accessed May 10, 2012) is an experimental BitTorrent client that continuously pretends to be a new peer in order to take advantage of reciprocity exemptions (see also Locher *et al.* 2006). The client is not widespread, however, and adjustments to the BitTorrent protocol could be made if necessary to close this loophole.

vitation only and that enforce sufficient seeding by their members through the threat of exclusion. Club-operated tracking servers are used to perform the necessary monitoring of club members' seeding activities.

4.2 Copyright enforcement in the BitTorrent network

What does BitTorrent's architecture mean for copyright enforcement? Interestingly, BitTorrent's reciprocity and internal monitoring features imply that copyright holders may not always want to curb unauthorized sharing of content. Minniti and Vergari (2010) develop a theoretical model in which the presence of file sharing clubs can improve the profit firms earn from content sales. Suppose that content is a horizontally differentiated good, and that the legal purchase of some content—a movie on a DVD, say—is required to join a file sharing club. In this case, the DVD essentially becomes a bundle containing the purchased movie itself, as well as access to copies of different movies that are bought and shared by the other members of the club. In analogy to bundling by a monopolist firm, this effect can increase the profits of the (oligopolistic) content industry.[21] As Minniti and Vergari (2010) show, conditions under which this is the case are most likely to arise in small, emerging markets. In mature markets, on the other hand, content owners generally have an interest in preventing file sharing. We focus on the latter case in the discussion below.

As was the case with FastTrack and Gnutella, the absence of any single player central to the network's functionality makes individual users potential targets of copyright litigation. However, enforced reciprocity implies that the option becoming a freeloader is not available (or much more costly). Using BitTorrent is, in a sense, a commitment to not being a freeloader. The implications of this commitment can be seen in Figure 8, which depicts the same game as Figure 5 but with strategy $\sim S, D$ removed. Because it was precisely the freeloading strategy $\sim S, D$ that previously dominated the strategy S, D, the enforcement approach of turning sharers into freeloaders is now impossible.

	S,D	$S,\sim D$	$\sim S,\sim D$
S,D	$V-r, V-r$	$V-r, -r$	$-r, 0$
$S,\sim D$	$-r, V-r$	$-r, -r$	$-r, 0$
$\sim S,\sim D$	$0, -r$	$0, -r$	$0, 0$

Figure 8: Suing *BitTorrent* users

[21] This mechanism is different from, and independent of, the sampling/discovery effects for experience goods discussed in Section 2.1.

In order to prevent file sharing, copyright holders would have to drive users off the network entirely, by choosing r large enough to make $\sim S, \sim D$ an attractive strategy relative to S, D. If $r > V$, strategy $\sim S, \sim D$ becomes a strictly dominant strategy. A very high probability of getting caught, or a very high penalty if caught, could hence undermine file sharing on BitTorrent.

However, $r > V$ is not necesssary for litigation to be successful. Note that, for all positive r, the strategy pair ($\sim S, \sim D$; $\sim S, \sim D$) is a strict Nash equilibrium of the game in Figure 8, and for $r > V/2$, this pair becomes the risk-dominant equilibrium: If player 1 was unsure about the strategy used by player 2, strategy $\sim S, \sim D$ would be a best response for player 1 against a larger set of mixed strategies of player 2 than S, D, and vice versa (see Harsanyi and Selten 1988). Thus, if network users coordinate on the risk-dominant equilibrium, sharing can be deterred as long as the expected penalty from litigation is more than half of the value a user attaches to the content that can be obtained over the network.[22] Yet, even this penalty would be non-negligible in relation to V. Suing a small number of users is hence unlikely to create a sufficiently strong risk perception for it to be an effective deterrent.

Copyright holders have responded to this challenge by returning to an indirect liability approach. Enforcement strategy has targeted two entities, in particular.

First, the industry decided to go after the third-party indexing and tracking operators that comprise the BitTorrent directory system, and whose servers represent the most centralized nodes of the network. The largest directory provider in the BitTorrent ecosystem is the Swedish website *The Pirate Bay (TPB)*. TPB bills itself as a "performance art project" and operates as a non-profit organization under Swedish tax law. Its 5.7 million users are required to register with the site that they can then search for content (more precisely, for torrent files). In May of 2006, TPB's offices were searched and its servers were confiscated by Swedish police on suspicion of copyright infringements by TPB users.[23] Based on the evidence discovered in the search, Swedish prosecutors filed criminal charges against TPB in a Stockholm court in 2008. The case was joined by the *International Federation of the Phonographic Industry (IFPI)*, an association of copyright holders, that sued for civil damages. In April of 2009, the court found four individuals associated with TPB guilty of copyright violations and sentenced each to a prison term of one year, as well as payment of fines and damages totaling about \$3.5 million.[24] The verdict was later upheld by an appeals court, which increased the damages but reduced the prison sentences. The

[22]For experimental evidence concerning the predictive power of the risk-dominance criterion, see van Huyck *et al.* (1990) and Cabrales *et al.* (2000).

[23]The raid caused the website to go offline for several days. When it went back online, its registered user base more than doubled due to the publicity the events had received in the media.

[24]Case B 13301-06 (District Court of Stockholm, Sweden 2009).

Swedish supreme court refused to hear TPB's subsequent appeal in 2012. Currently, the website operates under a different domain name.[25]

Second, copyright holders have increasingly focused their enforcement efforts on internet service providers (ISPs). Being "bottlenecks" between individual users and the file sharing community, ISPs can assist in copyright enforcement in three principal ways. First, they can block subscribers' access to certain offending sites, such as The Pirate Bay. While straightforward to implement, this measure can be circumvented relatively easily by users, for example by connecting to proxy servers or virtual private networks. Second, ISPs could potentially monitor their subscribers' data traffic for protected content and selectively block transmission of infringing data, using *deep packet inspection*. This measure is more complex to implement, and also raises obvious privacy concerns. Third, ISPs can implement a range of deterrents against subscribers who engage in copyright violations. These include the display of warning messages, a reduction of connection speeds, as well as suspension or termination of a subscriber's service. Using the notation of Figure 8, these measures substantially increase r, at little to no cost for rights owners (the costs are mostly borne by ISPs).

ISPs have come under mounting pressure to do all of these things in recent years. Since 2008, several national and international copyright groups have been engaged in civil lawsuits against European ISPs, alleging that the latter facilitate copyright infringements by their subscribers. European courts have consistently ruled against ISPs in this matter and required ISPs to block subscribers' access to BitTorrent sites, in particular to The Pirate Bay. However, European courts have so far been unwilling to require ISPs to indiscriminately monitor their subscribers' traffic.[26] On the legislative side, France adopted a graduated "three strikes" penalty scheme for copyright violations in 2009. Under the so-called HADOPI law (named after the government agency created to administer it), the first identified copyright violation by an individual results in an e-mail warning and initiates monitoring of the offender's internet connection by his or her ISP. The second violation results in a notice delivered by registered mail; the third may result in fines and suspension of internet privileges.[27]

[25] In February 2012, TPB changed its domain name from `thepiratebay.org` to `thepiratebay.se`. It is likely that the change was made to prevent a possible domain seizure by U.S. authorities. (U.S. law enforcement agancies can seize `.org`-domains, as the registrars of these domains are located on American soil.)

[26] A recent advisory opinion issued by the European Court of Justice held that, under current European Union law, ISPs could not be compelled by a court to indiscrimately monitor their subscribers' internet connection in order to enforce the intellectual property rights of others. (European Court of Justice case C-70/10: *Scarlet Extended SA v. Société belge des auteurs compositeurs et éditeurs (SABAM)*, 2011.)

[27] According to the French government, HADOPI costs €12 million annually to run (Reitman 2012). In September 2012, almost three years after its inception, the 60-employee agency called strike three for the first time. A French court subsequently sentenced the offender—a man who claimed that his ex-wife had downloaded songs by the artist Rihanna on BitTorrent—to a fine of €150 (Reitman 2012).

In the United States, two pieces of proposed anti-piracy legislation could impose similar obligations on American ISPs: *SOPA*, introduced in the House of Representatives in 2011; and *PIPA*, introduced in the Senate the same year. Both proposals are strongly supported by U.S. and international copyright groups, and opposed civil rights organizations, consumer advocacy groups, and all major internet companies.[28] After a considerable public outcry over the proposed legislation in early 2012, votes on both bills have been postponed indefinitely. Meanwhile, an agreement to establish a joint copyright enforcement initiative, the *Center for Copyright Information (CCI)*, was signed in 2011 by the RIAA, MPAA, their member companies, and most major U.S. internet service providers.[29] Under the agreement, content owners will scan BitTorrent and other P2P networks for protected content. Participating ISPs have agreed to receive notices from content owners when their subscribers (identified by their IP addresses) share protected materials, and to implement a three-phase schedule of countermeasures of increasing severity.[30] CCI, which began operations in 2012, insists that its activities are educational and not punitive, and that their primary goal is to prevent casual file sharing. The agreement does not specify what ISPs receive in return for their participation, but it is likely that ISPs view CCI as a self-regulatory effort to prevent legal action by content owners and to soften any obligations imposed by potential future legislation.

4.3 Some recent developments

The BitTorrent network is still evolving, as are the ways the network is being used. I will review three recent developments in this regard below.

First, a number of technological advances have diminished BitTorrent's dependence on the indexing and tracking providers. In 2009, The Pirate Bay abandoned its own centralized tracker when decentralized tracking became available (TBD remains an indexing site). In addition, techniques for distributed BitTorrent indexing

---

Legal sales of digital content in France are estimated to have increased by up to 25%, or €13.8 million annually, as a result of HADOPI (Danaher *et al.* 2012).

[28]SOPA (H.R. 3261) is short for *Stop Online Piracy Act*. PIPA (S. 968) is short for *Protect IP Act*, which is itself an acronym for *Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act*. SOPA reaches farther than PIPA in a number of aspects. For a list of organizations supporting and proposing SOPA, see `en.wikipedia.org/wiki/List_of_organizations_with_ official_stances_on_the_Stop_Online_Piracy_Act`(accessed May 14, 2012).

[29]Center for Copyright Information, Memorandum of Understanding. Final version: July 6, 2011. `www.copyrightinformation.org/wp-content/uploads/2013/02/Memorandum-of-Understanding.pdf` (accessed April 25, 2013).

[30]In the first phase of the so-called *Copyright Alert System* a suscriber is warned that his or her IP address has been associated with a copyright violation. The second phase requires the subscribers to acknowledge receipt of the warning. In the third phase the ISP can block popular websites or throttle a user's connection speed (Mullin 2012).

now exist as well.[31] The resulting network approaches Gnutella's degree of decentralization, thus eliminating BitTorrent's main legal vulnerability.

Second, fears of increased monitoring of P2P networks by content owners or ISPs have prompted more users to obfuscate their data traffic on these networks. A technology that is particularly interesting in the context of the BitTorrent network is the *seedbox*, a rented cloud computer that acts as an around-the-clock BitTorrent peer. Seedboxes are usually located in jurisdictions outside of North America or Europe and are accessed by file sharers only to transfer content between the user and the seedbox. File sharers can thus more easily fulfill seeding obligations, resulting in improved content availability and transfer speeds on the BitTorrent network. At the same time, seedboxes remove users one step from the actual file sharing network, thereby enhancing sharers' privacy.

Finally, the number of non-infringing uses of the BitTorrent network is growing. BitTorrent's speed and scalability make it a well-suited platform for the delivery of legal content without the need for expensive backbone infrastructure or large data centers (see Krishnan *et al.* 2007). Computer game developers, for example, routinely dispatch software updates to the gaming community via BitTorrent. A second example is *BitTorrent Sync*, a P2P file synchronization application based on the BitTorrent protocol.[32] Unlike many other popular synchronization services, such as *Dropbox*, BitTorrent Sync does not depend on a central server operated by a third party and thus offers enhanced privacy and a reduced risk of service interruption due to central-node failures.

5 Hosting platforms

Hosting platforms are based on the traditional client-server model of computer networks, instead of the peer-to-peer model. That is, users connect to a central file server to which they upload, and from which they download, files. Compared to P2P networks, central file hosting is more demanding in terms of the necessary server-side infrastructure, but usually delivers faster and and more reliable transfers. Starting in the late 2000s, hosting platforms appear to have replaced peer-to-peer networks as the predominant internet file sharing tool.

5.1 *Megaupload*

Megaupload was a Hong-Kong based file hosting service founded in 2005 by the Finnish-German hacker and internet entrepreneur Kim "Dotcom" Schmitz and shut down by U.S. authorities in 2012. Until its shutdown, Megaupload was by far the

[31] For example, distributed indexing is implemented as a feature ("Torrent Exchange") in the *BitComet* client program (`wiki.bitcomet.com/Torrent_Exchange`; accessed May 14, 2012).

[32] `labs.bittorrent.com/experiments/sync.html` (accessed April 29, 2013).

most popular hosting service, accounting for approximately one third of downloads traffic to and from hosting sites (Labovitz 2012). Megaupload—along with its many sister sites, including *Megavideo*, *Megabox*, and *Megaporn*—allowed users to store digital content on its file servers that could then be downloaded by anyone who knew the content's location. Megaupload did not publish a directory of the content it hosted, which meant that users could not search the platform for content they wished to download from others. Instead, users either shared download links directly among themselves or relied on searchable third-party directories that hosted user-submitted download links (see Figure 9).

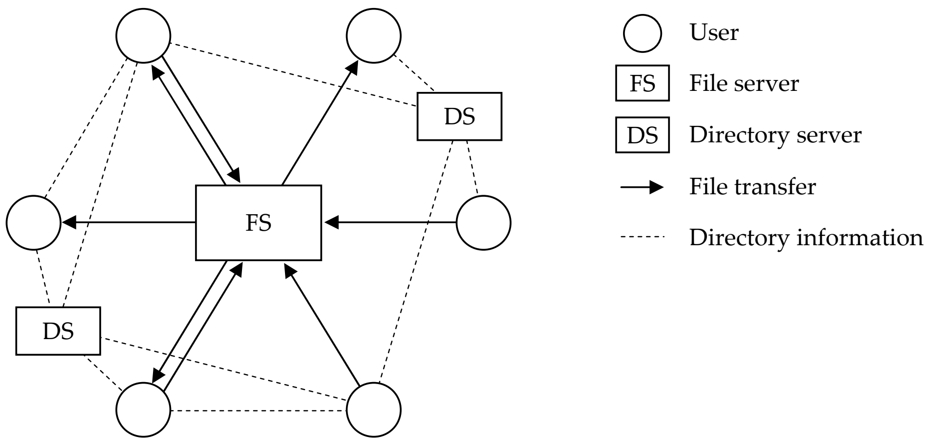

Figure 9: The *Megaupload*/*Rapidshare* architecture

What purpose did this architecture serve? In the 2000 Napster lawsuit, the 9th Circuit U.S. Court of Appeals affirmed a user's right to store multiple copies of digital content on more than one computer. As long as these copies were not made available to others, the transfer of legally acquired content between computers on the internet was permitted for space-shifting purposes. Megaupload's ostensive claim was that it provided a service that let users do precisely this—store personal copies of albums or movies remotely, so that they could be consumed in different locales or on different devices. The lack of a public file directory was meant to underscore the claim that the site was a "file locker" and not a file sharing service.

In reality, Megaupload was anything but a legitimate file hosting service. Most of the platform's revenues came from selling premium access to the site, including unrestricted downloads, large storage capacity, and fast transfer speeds. A smaller fraction came from advertisements served to users during downloads. To maximize the number of users and the number of downloads from its servers, the site made no effort to discourage third parties from providing directory services to its users. It also paid financial rewards to users who uploaded popular content and removed unpopular content to free up storage space. Megaupload profited handsomely from this business model: In the less than seven years that the site was in operation, the U.S. Department of Justice estimates that it earned a profit of approximately 175 million U.S. dollars.[33]

The central infrastructure that allowed Megaupload to extract these profits also placed it at a high risk of legal action. On January 19, 2012, U.S. Department of Justice officials seized several Megaupload servers in Virginia. Kim Schmitz and several of his associates were arrested by New Zealand authorities on the same day while celebrating Schmitz' 38th birthday. Schmitz is currently awaiting extradition to the United States and a federal criminal trial for copyright infringements under the DMCA. Megaupload has been unavailable since the 2012 raid. (Schmitz launched a new service, *Mega*, under a New Zealand domain in 2013.)

The Megaupload shutdown in 2012 and the Napster shutdown in 2001 demonstrates a key vulnerability of file sharing networks that rely on central critical infrastructure. In order to survive, file sharing networks that depend on a central node must carefully manage their legal risks. For example, the Swiss hosting platform Rapidshare—which has a network architecture similar to that of Megaupload—has cooperated with authorities and copyright holders in individual cases involving copyright violations, and has voluntarily terminated its own upload reward program. The site has so far managed to stay in business despite a number of civil lawsuits and criminal complaints against it.

5.2 *YouTube*

A different case is presented by the video sharing site *YouTube*, which allows users to upload videos that can then be streamed from the site (but not directly downloaded) to its users' computers. Founded in 2005 by Chad Hurley, Steven Chen, and Jawed Karim, and sold to Google in 2006, YouTube is today the largest video sharing community on the internet. It is visited by 800 million unique users every month and serves three billion views per day.[34] YouTube generates revenue from advertising on its website and in videos, but does not publish revenue or profit figures. YouTube's architecture differs from Megaupload in that both file hosting *and* directory functions are fully centralized (see Figure 10).

YouTube also differs from Megaupload in that the content hosted on YouTube is predominantly either copyright free, posted by the copyright holder (or with permission of the copyright holder), or falls under the fair use doctrine. Nevertheless,

---

[33] *United States v. Kim Dotcom*. Indictment, No. 1:12CR3 (U.S. District Court for the Eastern District of Virginia 2012). Megaupload's profligate founder spent his platform's earnings as soon as the cash rolled in. The purchases Kim Schmitz was reported to have made while running Megaupload include an extensive collection of luxury cars (listed in the indictment) and a $500,000 fireworks show that he watched from his private helicopter on New Year's Eve of 2011. Schmitz also rented an outsized New Zealand home that he named "Dotcom Mansion" but that he was unable to purchase after the country's authorities questioned his "good character" (Gallagher 2012).

[34] For a history of YouTube, see Seabrock (2012).

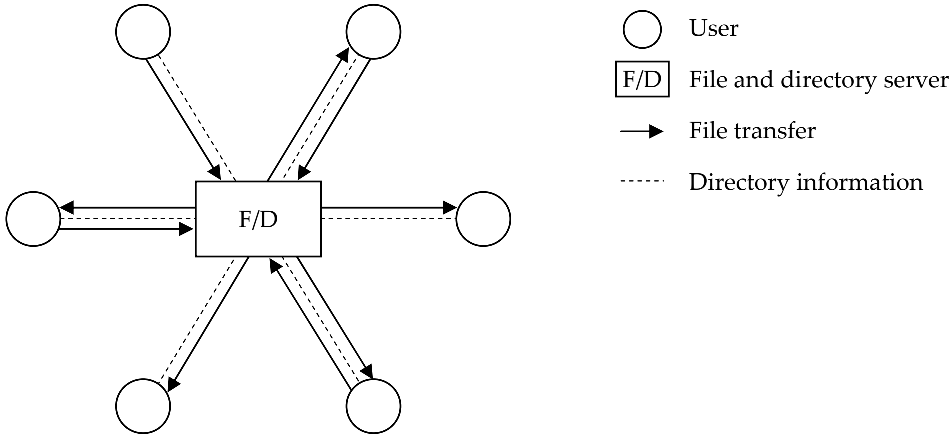

Figure 10: The *YouTube* architecture

YouTube users do upload unauthorized content frequently, and the site has come under attack for facilitating copyright violations by its users. Most prominently, in 2007 cable network operator *Viacom*, along with four other plaintiffs, sued YouTube for one billion dollars in damages, claiming that 160,000 pieces of Viacom's content had been uploaded by YouTube users without its permission.[35]

To navigate these legal challenges, YouTube has adopted an approach based on cooperation with authorities, courts, copyright holders, and its own users. YouTube proactively tries to foster a community based on legal content sharing by educating its users about copyright issues and fair use limitations when uploading new content. In addition, the platform offers content owners tools that assists in identifying infringing content and generating automatic takedown notices, and it terminates a user's account after the third successful takedown action.[36] A particularly interesting feature is YouTube's *Content ID* technology, which is has offered since 2007 and describes on its website as follows:

> "Rights holders deliver YouTube reference files [...] of content they own, metadata describing that content, and policies on what they want YouTube to do when we find a match. We compare videos uploaded to YouTube against those reference files. Our technology automatically

[35] *Viacom International, Inc. v. YouTube, Inc.* No. 07 Civ. 2103 (S.D.N.Y 2010). The district court granted summary judgement for YouTube/Google in 2010, which Viacom appealed. The case has been reinstated by the 2nd Circuit U.S. Court of Appeals in 2012.

[36] YouTube is sometimes overly aggressive when responding to takedown notices. In an ironic twist, shortly before Megaupload's shutdown a pro-Megaupload video was posted to YouTube that featured several artists objecting to their labels' campaign against the service. One of the labels, Universal Music, sent several takedown notices to YouTube requesting that the video be removed. YouTube's automated system promptly complied, despite the fact that Megaupload (and not Universal) actually owned the copyright to the video (Doctorow 2011).

> identifies your content and applies your preferred policy: monetize, track, or block."[37]

Thus, the copyright holder can allow unauthorized content to reside on YouTube's servers in exchange for benefits such as viewership analytics as well as a share of YouTube's advertising revenue generated by the content. Since many views on YouTube originate from users "following" other users, the Content ID monetization option can be regarded as a marketing tool that allows copyright holders to tap not only into YouTube's technical, but also its social network. According to YouTube, Content ID is now used by over $3,000$ copyright holders, including every major television network and movie studio, and accounts for one third of YouTube's monetized views.[38]

6 Conclusion

In the preceding sections, I surveyed a number of different file sharing network architectures and described their economic properties as well as the enforcement tactics used by copyright holders in each instance. Table 1 summarizes these cases.

It is evident that enforcement through litigation has always targeted the most centralized function of each network. When all essential network functions are fully distributed, litigation has targeted individual file sharers, developers of client software, as well as internet service providers. Legal action has resulted in a complete network shutdown in two of the portrayed cases (Napster and Megaupload, both of which featured one fully centralized network function) and an effective shutdown in at least one other case (OpenNap). These legal tactics were flanked by technical attacks, in particular the flooding of networks with junk content.

On the other hand, as the case of YouTube demonstrates, it is possible for commercial operators of large content sharing platforms—even highly centralized ones—to stay "one step ahead" of these challenges by taking proactive measures toward countering infringing uses. In fact, YouTube's high degree of centrality may even help in this regard, as it allows for equally high degree of control over the platform's operations, including operations that address infringing uses.

I conclude this overview with a brief glimpse into the future of online file sharing. Current forecasts predict that, as a proportion of total internet traffic, file sharing is expected to decline from about 40% in 2010 to 20% by 2015 (Cisco Systems 2011). The relative decline is due to other data-intensive internet applications that are expected to grow at a very fast rate. The most important among them is media streaming. Movie and music subscription services, such as Netflix, Rhapsody, or Spotify, are convenient legal alternatives to online piracy. The emergence of these services and

[37] `www.youtube.com/t/contentid` (accessed May 8, 2012).

[38] `www.youtube.com/t/press_statistics` (accessed May 9, 2012).

Network	File hosting/ File transfer	Directory	Litigation targets	Technical attacks
Napster [a]	P2P	Central	Directory	
OpenNap [c]/ eDonkey	P2P	3^{rd}-party	Directory, clients	Yes
Gnutella [c]/ FastTrack [c]	P2P	P2P	Sharers, clients	Yes
BitTorrent	P2P	3^{rd}-party, P2P	Directory, clients, ISPs	Yes
Megaupload [b]/ Rapidshare	Central	3^{rd}-party, P2P	Hosting	
YouTube	Central	Central	Hosting/directory	

[a] Shut down in 2001; [b] shut down in 2012; [c] no longer significant.

Table 1: Network architecture and enforcement strategy

their relatively attractive pricing reflect a significant shift in the business model of the entertainment industry, which was at least partly driven by the threat posed by illegal file sharing (Peitz and Waelbroeck 2005). Despite the fact that file sharing traffic is expected to decline relative to all internet traffic, it is still expected to grow in absolute terms, by about 23% per year over the next several years (Cisco Systems 2011). At the same time, as a fraction of file sharing traffic P2P networks have lost some importance as fast and reliable file hosting services have been gaining in popularity since the late 2000s. However, this trend may be reversing for two reasons. First, as pointed out earlier, the number of non-infringing uses of P2P networks is increasing. Second, the case of Megaupload demonstrates that file hosting services are not immune to prosecution and sudden shut downs. The decentralized architecture and stronger privacy offered by P2P networks may once again make them the preferred alternative for the unauthorized distribution of content.

References

[1] Adar, Eytan and Bernardo Huberman. "Free Riding on Gnutella." *First Monday* 5, Issue 10, 2000.

[2] Asvanund, Atip, Sarvesh Bagla, Munjal Kapadia, Ramayya Krishnan, Michael Smith, and Rahul Telang. "Intelligent Club Management in Peer-to-Peer Networks." Working Paper: Carnegie Mellon University, Heinz School of Public Policy and Management, 2003.

[3] Belleflamme, Paul and Martin Peitz. "Digital Piracy: Theory." In: Martin Peitz and Joel Waldfogel (eds.), *The Oxford Handbook of the Digital Economy*, Oxford University Press, 489–530, 2012.

[4] Bhattacharjee, Sudip, Ram Gopal, Kaveepan Lertwachara, and James Marsden. "Impact of Legal Threats on Online Music Sharing Activity: An Analysis of Music Industry Legal Actions." *Journal of Law and Economics* 49, 91–114, 2006.

[5] Cabrales, Antonio, Walter Garcia-Fontes, and Massimo Motta. "Risk Dominance Selects the Leader: An Experimental Analysis." *International Journal of Industrial Organization* 18, 137–162, 2000.

[6] Cisco Systems. "Cisco Visual Networking Index: Forecast and Methodology, 2010–2015." White Paper, 2011. (`www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf`; accessed May 12, 2012.)

[7] Cox, Joe, Alan Collins, and Stephen Drinkwater. "Seeders, Leechers and Social Norms: Evidence From the Market for Illicit Digital Downloading." *Information Economics and Policy* 22, 299–205, 2010.

[8] Doctorow, Cory. "Universal Music Files Fraudulent Copyright Complaints with YouTube, Censors Pro-Megaupload Song." *BoingBoing*, December 10, 2011. (`boingboing.net/2011/12/10/universal-music-files-fraudule.html`; accessed May 10, 2012.)

[9] Danaher, Brett, Michael Smith, Rahul Telang, and Siwen Chen. "The Effect of Graduated Response Anti-Piracy Laws on Music Sales: Evidence from an Event Study in France." Working Paper: Wellesley College, 2012.

[10] Gallagher, Sean. "The Fast, Fabuluous, Allegedly Fraudulent Life of Megaupload's Kim Dotcom." *Wired*, January 26, 2012. (`www.wired.com/threatlevel/2012/01/kim-dotcom`; accessed May 8, 2012.)

[11] Gopal, Ram, Sudip Bhattacharjee, and Lawrence Sanders. "Do Artists Benefit from Online Music Sharing?" *Journal of Business* 79, 1503–1533, 2006.

[12] Harsanyi, John and Reinhard Selten. *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: MIT Press.

[13] Hong, Seung-Hyun. "Measuring the Effect of Digital Technology on the Sales of Copyrighted Goods: Evidence from Napster." Working Paper: University of Illinois, Department of Economics, 2007.

[14] Hughes, Daniel, Geoff Coulson, and James Walkerdine. "Free Riding on Gnutella Revisited: The Bell Tolls?" *IEEE Distributed Systems Online* 6, Issue 6, 2005.

[15] Krishnan, Ramayya, Michael Smith, and Rahul Telang. "The Economics of Peer-to-Peer Networks." *Journal of Information Technology Theory and Application* 5, 31–44, 2003.

[16] Krishnan, Ramayya, Michael Smith, Zhulei Tang, and Rahul Telang. "Digital Business Models for Peer-to-Peer Networks: Analysis and Economic Issue." *Review of Network Economics* 6, 194–213, 2007.

[17] Krishnan, Ramayya, Michael Smith, Zhulei Tang, and Rahul Telang. "The Virtual Commons: Why Free-Riding Can Be Tolerated in File Sharing Networks." Carnegie Mellon University, Heinz Research Paper 59, 2008.

[18] Labovitz, Craig. "File Sharing in the Post MegaUpload Era." *DeepField Networks Blog*, February 7, 2012. (`blog.deepfield.net/2012/02/07/file-sharing-in-the-post-megaupload-era`; accessed May 11, 2012.)

[19] Landes, William and Douglas Lichtman. "Indirect Liability for Copyright Infringement: Napster and Beyond." *The Journal of Economic Perspectives* 17 (2), 113–124, Spring 2003.

[20] Stan Liebowitz. "Testing File-sharing's Impact by Examining Record Sales in Cities." *Management Science* 54, 852–859, 2008.

[21] Locher, Thomas, Patrick Moor, Stefan Schmid, and Roger Wattenhofer. "Free Riding in BitTorrent is Cheap." Conference proceedings, *HotNets 2006*.

[22] Maguire, James. "Hitting P2P Users Where It Hurts." *Wired*, January 13, 2004. (`www.wired.com/entertainment/music/news/2003/01/57112`; accessed May 9, 2012.)

[23] McBride, Sarah and Ethan Smith. "Music Industry to Abandon Mass Suits." *The Wall Street Journal*, December 19, 2008. (`online.wsj.com/article/SB122966038836021137.html`; accessed May 10, 2012.)

[24] Michel, Norbert. "Digital File Sharing and the Music Industry: Was There a Substitution Effect?" *Review of Economic Research on Copyright Issues* 2, 42–52, 2005.

[25] Michel, Norbert. "The Impact of Digital File Sharing on the Music Industry: An Empirical Analysis." *Topics in Economic Analysis & Policy* 6, Issue 1.

[26] Minniti, Antonio and Cecilia Vergari. "Turning Piracy into Profits: A Theoretical Investigation." *Information Economics and Policy* 22, 379–390, 2010.

[27] Mullin, Joe. "How ISPs will do 'six strikes': Throttled speeds, blocked sites." *Ars Technica*, November 16, 2012. (`arstechnica.com/tech-policy/2012/11/how-`

isps-will-do-six-strikes-throttled-speeds-blocked-sites; accessed April 25, 2013.)

[28] Nandi, Tushar and Fabrice Rochelandet. "The Incentives for Contributing Digital Contents over P2P Networks: An Empirical Investigation." *Review of Economic Research on Copyright Issues* 5, 19–35, 2008.

[29] Oberholzer-Gee, Felix and Koleman Strumpf. "The Effect of File Sharing on Record Sales: An Empirical Analysis." *Journal of Political Economy* 115, 1–42, 2007.

[30] Peitz, Martin and Patrick Waelbroeck. "The Effect of Internet Piracy on CD Sales: Cross Section Evidence." *Review of Economic Research on Copyright Issues* 1, 71–79, 2004.

[31] Peitz, Martin and Patrick Waelbroeck. "An Economist's Guide to Digital Music." *CESifo Economic Studies* 51, 359–428, 2005.

[32] Peitz, Martin and Patrick Waelbroeck. "Why the Music Industry May Gain from Free Downloading: The Role of Sampling." *International Journal of Industrial Organization* 24, 907–913, 2006.

[33] Reitman, Rainey. "French Anti-Piracy Law Claims First Victim, Convicted of Failing to Secure His Internet Connection." *Deeplinks Blog*, Electronic Frontier Foundation, September 13, 2012. (www.eff.org/deeplinks/2012/09/french-anti-piracy-law-claims-first-victim-convicted-failing-secure-his-internet; accessed April 28, 2013.)

[34] Rob, Rafael and Joel Waldfogel. "Piracy and the High C's: Music Downloading, Sales Displacement, and Social Welfare in a Sample of College Students." *Journal of Law and Economics* 49, 29–62, 2006.

[35] Schulze, Hendrik and Klaus Mochalski. "Internet Study 2008/2009." ipoque GmbH, 2009. (www.ipoque.com/en/resources/internet-studies; accessed May 12, 2012.)

[36] Seabrock, John. "Streaming Dreams." *The New Yorker*, January 16, 2012.

[37] Shneidman, Jeffrey and David Parkes. "Rationality and Self-Interest in Peer to Peer Networks." In: M. Kaashoek and Ion Stoica (eds.), *Peer-to-Peer Systems II (Lecture Notes in Computer Science)*, Springer, 2003.

[38] Sohn, David and Andrew McDiarmid. "Dangerous Bill Would Threaten Legitimate Websites." *The Atlantic*, November 17, 2011. (www.theatlantic.com/technology/archive/2011/11/dangerous-bill-would-threaten-legitimate-websites/248619; accessed May 14, 2012.)

[39] *The Economist*. "Online Pirates at Bay." April 17, 2009. (www.economist.com/node/13518830?story_id=13518830; accessed May 9, 2012.)

[40] van Huyck, John, Raymond Battalio, and Richard Beil. "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *American Economic Review* 80, 2234–2249, 1990.

[41] Wingfield, Nick. "Online Swapping of Music Declines in Wake of Suits." *The Wall Street Journal*, January 5, 2004. online.wsj.com/article/0,,SB107326093879736500,00.html; accessed May 10, 2012.)

[42] Zentner, Alejandro. "File Sharing and International Sales of Copyrighted Music: An Empirical Analysis With a Panel of Countries." *Topics in Economic Analysis & Policy* 5, Issue 1, Article 21.

[43] Zentner, Alejandro. "Measuring the Effect of File Sharing on Music Purchases." *The Journal of Law and Economics* 49, 63–90, 2006.