Content Sharing Systems for Digital Media

Jerald Hughes
University of Texas Pan-American, USA

Karl Reiner Lang
Baruch College of the City University of New York, USA

INTRODUCTION

In 1999, exchanges of digital media objects, especially files of music, came to constitute a significant portion of Internet traffic, thanks to a new set of technologies known as peer-to-peer (P2P) file-sharing systems. The networks created by software applications such as Napster and Kazaa have made it possible for millions of users to gain access to an extraordinary range of multimedia files. However, the digital product characteristics of portability and replicability have posed great challenges for businesses that have in the past controlled the markets for image and sound recordings.

'Peer-to-peer' is a type of network architecture in which the various nodes may communicate directly with other nodes, without having to pass messages through any central controlling node (Whinston, Parameswaran, & Susarla, 2001). The basic infrastructure of the Internet relies on this principle for fault tolerance; if any single node ceases to operate, messages can still reach their destination by rerouting through other still-functioning nodes. The Internet today consists of a complex mixture of peer-to-peer and client-server relationships, but P2P file-sharing systems operate as overlay networks (Gummadi, Saroiu, & Gribble, 2002) upon that basic Internet structure.

P2P file-sharing systems are software applications which enable direct communications between nodes in the network. They share this definition with other systems used for purposes other than file sharing, such as instant messaging, distributed computing, and media streaming. What these P2P technologies have in common is the ability to leverage the combined power of many machines in a network to achieve results that are difficult or impossible for single machines to accomplish. However, such networks also open up possibilities for pooling the interests and actions of the users so that effects emerge which were not necessarily anticipated when the network technology was originally created (Castells, 2000). In a narrow sense, P2P file-sharing systems refer to applications that exchange content over computer networks where the nodes act both as client and server machines, requesting and serving files (e.g., Kazaa, BitTorrent). In a wider sense, P2P file-sharing systems also include any application that lets peer users exchange digital content among themselves (e.g., YouTube, Flickr).

TECHNICAL FOUNDATIONS

P2P file-sharing systems function by integrating several digital technologies (see Table 1).

The first digital format for a consumer product was the music CD, introduced in the early 1980s. This format, Redbook Audio, encoded stereo sound files using a sample rate of 44.1 kHz and a sample bit depth of 16 bits. In Redbook Audio, a song 4 minutes long requires 42 megabytes of storage. Even at broadband speeds, downloading files of this size is impractical for many users, so effective compression is a necessary component of efficient file-sharing technology.

The breakthrough for file sharing audio files came from the MPEG (Motion Picture Experts Group) specification for digital video, via the Fraunhofer Institute in Erlangen, Germany, which used the MPEG-1 Layer 3 (thus, 'MP3') specification to develop the first stand-alone encoding algorithms for MP3 files. The MP3 encoding algorithm yields a compression ratio (lossy compression) for standard MP3 files of 11:1 from the original Redbook file. The nominal bitstream rate is 128 kilobits per second, although MP3 encoding tools now allow variable rates, up to 320 kbs. In MP3, a single song of 4 minutes length becomes available in relatively high-quality form in a digital file 'only' 4 megabytes in size.

While MP3s are the most popular for music, many other file types also appear in P2P networks, including .exe (for software), .zip, and many different formats for video, images, and books. Users with fast Internet

Table 1. Enabling technologies for P2P filesharing systems

Encoding for digital media	Music: MP3 (MPEG 1 Layer 3), AAC, WMA, OGG Movies and video: DivX (MPEG 4), Xvid
Multimedia systems and players	Software: Winamp, MusicMatch, RealPlayer, Quicktime Hardware: iPod, Zune, Archos, Media Center PC
Broadband Internet Access	DSL, cable modems, wi-fi, satellite, T1 or T3 lines, Internet 2
P2P filesharing software	Napster, Kazaa, BitTorrent, Grokster, Limewire, Soulseek, Bearshare, eMule

connections may use P2P networks to obtain uncompressed original digital content, stored as '.iso' disk images recorded directly from the source. The swift and nearly universal adoption of MP3 audio format has also driven the development of further supporting hardware in the form of stand-alone players such as the iPod, Zune, and many more. Support for MP3 and other digital media formats is now available in nearly every type of multimedia computing device, including desktops, laptops, PDAs, cellphones, car stereos, handheld gaming platforms, and dedicated 'media player' devices. The MPEG-4 video compression format has had an explosive effect on video file sharing, similar to that previously seen by the MP3 on audio file sharing.

Obtaining digital media files for one's player or viewer means either 'ripping' it oneself from the source, such as a CD or DVD, or finding it on a file-sharing network. P2P file-sharing applications create virtual networks which allow each participant to act as either a servant or a client, able to both receive and distribute files, often simultaneously. P2P software sets up the required communications and protocols for connecting to the virtual network, conducting searches, and managing file exchanges with other machines in the system.

P2P applications use metadata to allow keyword searches. Once a desired file is discovered, the P2P application establishes a direct link between the participating machines so the file can be downloaded. For original Napster, this involved a central index, maintained at Napster's server site, and a single continuous download from a single IP address. P2P users may share all, none, or some of the files on their hard drives. Sharing no files at all, a behavior known as 'free-riding' (Adar & Huberman, 2000), can degrade the performance of the network, but the effect is surprisingly small until a large majority of users are free-riding. For individual song downloads, using one-to-one download protocols works well, but for very large files, such as those used for digital movie files, the download can take hours, and may fail entirely if the servent leaves the network before transfer is complete. To solve this problem, P2P engines like BitTorrent allow users to download a single file from many users at once, thus leveraging not only the storage but also the bandwidth of many machines on the network. A similar technique is used by P2P systems which provide streaming media (not file downloads) in order to avoid the cost and limitations of single-server media delivery. As both the penetration of broadband service and the speed of such service continue to rise, the effectiveness of the P2P approach to large-file distribution also increases.

APPLICATIONS

P2P file-sharing systems have passed through several stages of development. The first generation, the original Napster, was closed by the courts because of copyright infringement. The second generation, widely used in tools such as Kazaa, reworked the architecture to allow for effective file discovery without the use of a central index. However, users themselves were exposed to legal sanctions, as the Recording Industry Association of America (RIAA) filed lawsuits against users who made files of music under copyright available to other network users. The third generation of P2P file-sharing