

THINGS YOU SHOULD KNOW ABOUT...

P2P

Scenario

For his senior project in computer science, Zach wrote a program that pulls weekly regional economic data—including statistics such as jobless claims, housing starts, changes in average wages, and other financial metrics—from a variety of government, industry, and nonprofit sources. Users of the program can represent these economic factors on a map of the United States, showing regional differences and trends over time. The program allows users to correlate economic conditions with a wide range of social and demographic data and build digital maps that represent these relationships. For example, a user can combine wage data with gate receipts at sports parks to investigate whether people whose real wages are declining go to more or fewer ball games.

Zach released his application as an open-source program, available from a server in the computer science department. Because researchers from across a wide spectrum of disciplines found the application useful, however, the server quickly bogged down due to the number of people trying to download the program. The department didn't have a budget for more servers to meet the demand, so Zach used BitTorrent to allow people to share the application through a P2P network. With this approach, anyone who wants the program can get it from anyone else in the P2P network who has a copy of it.

Over the course of the semester, Zach made continual improvements to the application, and each new version was distributed through the P2P network, allowing many users from around the world the ability to access the updated version quickly and easily, without having to rely on a server at Zach's university. To remain current, the application uses large data files that are updated weekly, including economic and demographic data from a variety of sources. These frequent and large updates to the data files are also distributed by the P2P network. In this way, academics around the globe are easily able to conduct research based on up-to-date versions of both the application and the data files. For his part, Zach can spend his time improving the program and looking for ways to enhance its functionality, rather than managing an overtaxed server or trying to find the money to buy more servers.

1 What is it?

Peer-to-peer (P2P) is an approach to content distribution in which digital files are transferred between “peer” computers over the Internet. In a client/server structure, servers store content and, when requested, send that content to client machines—a one-to-many distribution model. With P2P, the client machines are peers and communicate with one another. A computer in the network can submit a request for specific content, and any peer in the network that has a copy of the file can send it, resulting in a many-to-many model that does not rely on a central repository. Because they do not rely on a central server to deliver content, P2P networks tend to be faster and much more reliable—as long as at least one other computer in the network has a file, others can access it. Any user of a P2P network is as likely to be a contributor as a consumer, and, in this sense, P2P approximates the original conception of the Internet as a network of connections among individuals and organizations that give and take information.

2 Who's doing it?

P2P is used as an efficient, robust means to distribute a wide range of content including software, games, books, movies, and music. Many distributors of Linux, for example, use P2P networks to distribute the latest versions of the open-source operating system. Similarly, OpenOffice can be downloaded through any of several P2P networks, and corrections and updates for computer games such as *World of Warcraft*, which are popular with students, are frequently distributed through P2P. Wikipedia compiled a CD with more than 2,000 articles selected specifically for schools and children, and this resource can be downloaded from a P2P network. Researchers routinely share data and computer code using P2P technology, and scientists at NASA use the technology to distribute satellite imagery. Commercial providers of news, music, and video have begun to use P2P methods as well. In January, CNN used P2P to stream the presidential inauguration to more than 650,000 simultaneous users. Amazon and Google have experimented with consumer services based on this technology, and media distributors such as Time Warner have used it to relieve the traffic pressure on their download servers.

3 How does it work?

P2P networks require installing an application on Internet-connected computers, which become nodes, or peers, that communicate with one another. The nodes create large, interconnected networks through which files are transferred. P2P users can see lists of files that are available on other nodes in the network and can submit a request to the network for any available file. In some services, the entire file is transferred from one peer computer to

more >>

the requestor; with other protocols, such as BitTorrent, different parts of a file are downloaded from multiple nodes on the network and reassembled on the requestor's computer, further improving the efficiency of file transfers by reacting to moment-by-moment fluctuations in network traffic. In both models, files are exchanged between peers without access controls or oversight of transfers. Any node can share or request any file on the network.

4 Why is it significant?

By some estimates, a quarter of all adults—including half of all college students—have used P2P networks, and, according to the Pew Internet & American Life Project, users are not significantly differentiated by income, gender, or race. By lodging copies of files in numerous locations, P2P balances traffic loads, which might otherwise overwhelm servers, and minimizes the chance of service breakdown due to localized server or communication outages. P2P helps optimize network bandwidth and the storage and processing resources of connected computers by building a relationship of trust among peers, creating new opportunities for the efficient transfer of files—particularly very large files and those that are in extremely high demand. As a new channel for content distribution, P2P changes the conventional hierarchy of information. The roles of producer, consumer, and gatekeeper of digital content blur, and the structure that gives content owners control over who can have what is replaced by one in which information and resources are available much more widely than before.

At the same time, however, removing the central authority over content distribution has resulted in copyright infringement on some P2P networks. Music and other media files can easily be reproduced and exchanged outside the distribution channels of the individuals and organizations that hold the copyright for those materials. The music and movie industries, in particular, contend that file sharing has cost them enormous amounts of money in lost sales, and representatives of media companies have gone to great lengths to try to mitigate this activity, including prosecuting alleged violators. Because file traders are frequently only identified by a network address, complaints of copyright infringement generally go through Internet service providers (ISPs) to identify individual users. Because colleges and universities often function as the ISPs for campus users, higher education has found itself in the middle of this controversy.

5 What are the downsides?

Special care may be necessary to ensure the integrity of shared files because of the multiple, dispersed sources in play. Some programs based on P2P expose considerable vulnerabilities. P2P networks have been used for transmission of viruses, malware, and corrupted or mislabeled files. Individuals who inadvertently place sensitive files in network-shared directories have been victims of identity theft when those files were downloaded by others. Moreover, some networks have seen their bandwidth sapped by users trading music and movies.

The advantages that P2P offers have largely been overshadowed by controversy. The recording and motion picture industries, the U.S. Congress, and some colleges and universities tend to use P2P as a synonym for illegal transfers of files, and some have sought to ban the practice altogether. Rather than looking for ways to put P2P to use in the service of teaching or research, colleges and universities instead find themselves wrangling with resource usage and legal issues. Students are frequent targets of copyright-infringement legal actions, and legislation on higher education sometimes includes requirements that colleges and universities step up efforts to enforce copyright.

6 Where is it going?

P2P technology has the potential to play an important, positive role in the fulfillment of institutional missions of teaching, research, and the dissemination of knowledge. For the moment, however, P2P has become a highly visible and rancorous front in the battle over copyright and sustainable business models in the age of digital reproduction and electronic distribution. P2P is one of a range of technologies that have upset the traditional balance of control over the creation, reproduction, dissemination, and consumption of knowledge and information. New expectations and relationships must develop to support intellectual property considerations while expanding access to digital content and taking advantage of new technological capabilities.

7 What are the implications for higher education?

P2P has the potential to bring more information and resources to more people and applications than otherwise would be possible. In some fields, P2P already facilitates an efficient, easy exchange of files between researchers around the world without relying on a single location to meet all of the demand. As rising numbers of academics in a widening array of disciplines use technology in their scholarship, a higher premium is placed on the ability to access the disparate data that will drive research and teaching in the years to come, and P2P can be an important tool in this effort. The conflict over copyright, however, continues to color policy discussions about its use on campus.

EDUCAUSE

EDUCAUSE is a nonprofit membership association created to support those who lead, manage, and use information technology to benefit higher education. A comprehensive range of resources and activities is available to all EDUCAUSE members. The association's strategic directions include focus in four areas: Teaching and Learning; Managing the Enterprise; E-Research and E-Scholarship; and the Evolving Role of IT and Leadership. For more information, visit educause.edu.