

**IJCRR**

Vol 05 issue 23

Section: Technology

Category: Research

Received on: 28/09/13

Revised on: 21/10/13

Accepted on: 12/11/13

UNSTRUCTURED PEER – TO - PEER NETWORKS USING OPTIMIZING OVERLAY TOPOLOGIES

M. Arulraja, K. Umasankar, S. Manikandan

Department of Computer Science and Engineering, Prist University, Kumbakonam, Thanjavur, TN, India

E-mail of Corresponding Author: siskin2008@gmail.com

ABSTRACT

In its simplest form, a peer-to-peer (P2P) network is created when two or more PCs are connected and share resources without going through a separate server computer. A P2P network can be an ad hoc connection—a couple of computers connected via a Universal Serial Bus to transfer files. A P2P network also can be a permanent infrastructure that links a half-dozen computers in a small office over copper wires. Or a P2P network can be a network on a much grander scale in which special protocols and applications set up direct relationships among users over the Internet.

Keywords: Serial bus, protocols, copper wires

INTRODUCTION

The initial use of P2P networks in business followed the deployment in the early 1980s of free-standing PCs. In contrast to the mini mainframes of the day, such as the VS system from Wang Laboratories Inc., which served up word processing and other applications to dumb terminals from a central computer and stored files on a central hard drive, the then-new PCs had self-contained hard drives and built-in CPUs. The smart boxes also had onboard applications, which meant they could be deployed to desktops and be useful without an umbilical cord linking them to a mainframe.

Many workers felt liberated by having dedicated PCs on their desktops. But soon they needed a way to share files and printers. The obvious solution was to save files to a floppy disk and carry the disk to the intended recipient or send it by interoffice mail

MATERIALS AND METHODS

That practice resulted in the term "sneaker net." The most frequent endpoint of a typical sneaker net was the worker who had a printer connected to

his machine. While sneaker nets seemed an odd mix of the newest technology and the oldest form of transportation, the model is really the basis for today's small P2P work groups. Whereas earlier centralized computing models and today's client/server systems are generally considered controlled environments in which individuals use their PCs in ways determined by a higher authority, a classic P2P workgroup network is all about openly sharing files and devices. In general, office and home P2P networks operate over Ethernet (10M bit/sec.) or Fast Ethernet (100M bit/sec.) and employ a hub-and-spoke topology. Category 5 (twisted-pair) copper wire runs among the PCs and an Ethernet hub or switch, enabling users of those networked PCs access to one another's hard drives, printers or perhaps a shared Internet connection.

BOTH CLIENT AND SERVER

In effect, every connected PC is at once a server and a client. There's no special network operating system residing on a robust machine that supports special server-side applications like directory services (specialized databases that control who

has access to what). In a P2P environment, access rights are governed by setting sharing permissions on individual machines.

For example, if User A's PC is connected to a printer that User B wants to access, User A must set his machine to allow (share) access to the printer. Similarly, if User B wants to have access to a folder or file, or even a complete hard drive, on User A's PC, User A must enable file sharing on his PC. Access to folders and printers on an office P2P network can be further controlled by assigning passwords to those resources.

TRENDS AND IMPACT

The first appearance of open source systems such as Napster in 1999 radically changed file-sharing mechanisms. The traditional client-server file sharing and distribution approach using protocols like FTP (File Transfer Protocol) was supplemented with a new alternative — P2P networks. At the time, Napster was used extensively for the sharing of music files. Napster was shut down in mid-20012 due to legal action by the major record labels.

The shutting of Napster did not stop the growth of P2P applications. A number of publicly available P2P systems have appeared in the past few years, including Gnutella, KaZaA, WinMX and BitTorrent, to name but a few. From analysis of P2P traffic in 2007, BitTorrent is still the most popular file sharing protocol, accounting for 50-75% of all P2P traffic and roughly 40% of all Internet traffic³.

P2P technology is not just used for media file sharing. For example, in the bioinformatics research community, a P2P service called Chinook⁴ has been developed to facilitate exchange of analysis techniques. The technology is also used in other areas including IP-based telephone networks, such as Skype⁵, and television networks, such as PPLive⁶. Skype allows people to chat, make phone calls or make video calls. When launched, each Skype client acts as a peer, building and refreshing a table of

reachable nodes ⁷ in order to communicate for chat, making phone calls or video calls. PPLive shares live television content. Each peer downloads and redistributes live television content from and to other peers⁸.

GOVERNANCE AND REGULATIONS

In the U.S., a number of politicians have raised concerns about possible threats to national security due to P2P network technology. The possibility of accidental leaks of classified information by government officers to foreign governments, terrorists or organized crime via P2P file sharing programs has prompted a view that “new laws and rules should be enacted to protect personal information held by federal agencies and other organizations”. The proposal does not restrict P2P networks as a whole, but attempts to strike “a balance that protects sensitive government, personal and corporate information and copyright laws”⁹.

A P2P network itself is only a form of technology, and is not related to disputes over content and intellectual property rights. However, there have been court cases in Hong Kong against illegal P2P activities. In 2005, a Hong Kong resident was convicted of Peer-to-peer Network Page 7 of 14 breaching the Copyright Ordinance by uploading illegal copies of copyrighted works to the Internet using the BitTorrent peer-to-peer file sharing program, and making files available for download by other Internet users¹⁰.

SECURITY THREATS

A P2P network treats every user as a peer. In file sharing protocols such as BT, each peer contributes to service performance by uploading files to other peers while downloading. This opens a channel for files stored in the user machine to be uploaded to other foreign peers. The potential security risks include:

1. TCP ports issues

Usually, P2P applications need the firewall to open a number of ports in order to function

properly. BitTorrent, for example, will use TCP ports 6881-6889 (prior to version 3.2). The range of TCP ports has been extended to 6881-6999 as of 3.2 and later¹⁶. Each open port in the firewall is a potential avenue that attackers might use to exploit the network. It is not a good idea to open a large number of ports in order to allow for P2P networks.

2. Propagation of malicious code such as viruses

As P2P networks facilitate file transfer and sharing, malicious code can exploit this channel to propagate to other peers. For example, a worm called VBS. Gnutella was detected in 2000 which propagated across the Gnutella file *Peer-to-peer Network* Page 10 of 14

sharing network by making and sharing a copy of itself in the Gnutella program directory¹⁷.

Algorithm 1: Building Hierarchical Summaries

1. for each peer
2. for each document
3. Generate its vector vd by VSM
4. Generate peer weighted term dictionary vp
5. for each document vector vd
6. transform it into $D(vp)$ dimensionality
7. generate high-dimensional point for vd by SVD
8. Pass vp to its super peer
9. for each super peer
10. Generate group weighted term dictionary vs
11. for each vp
12. transform it into $D(vs)$ dimensionality
13. generate high-dimensional point for vp by SVD
14. Pass vs to other super peers
15. Generate global weighted term dictionary vn
16. for each vs
17. Transform it into $D(vn)$ dimensionality
18. Generate high-dimensional point for vs

RESULT AND DISCUSSION

While P2P networks open a new channel for efficient downloading and sharing of files and data, users need to be fully aware of the security threats associated with this technology. Security

measures and adequate prevention should be implemented to avoid any potential leakage of sensitive and/or personal information, and other security breaches. Before deciding to open firewall ports to allow for peer-to-peer traffic, system administrators should ensure that each request complies with the corporate security policy and should only open a minimal set of firewall ports needed to fulfil P2P needs. For end-users, including home users, care must be taken to avoid any possible spread of viruses over the peer-to-peer network.

CONCLUSION

The P2P paradigm is becoming increasingly popular for developing internet-scale applications. P2P content sharing systems, as popularized by the initial endeavors of Napster, Gnutella, etc., are receiving increasing attention from academics and industry as an important class of internet data management applications.

In this paper we have presented the problem of managing content and resources in P2P sharing systems so to ensure the efficiency of operation. Efficiency is viewed both from the point of view of the system, in the sense of ensuring globally fair load distribution among all Peers, and from the point of view of the users, in the sense of facilitating low user-request response times.

REFERENCES

1. S. Agrawal, S. Chaudhuri, and G. Das. Dbxplorer: A system for keyword-based search over relational databases. In ICDE, 2002.
2. S. Berchtold and D. A. Keim. Indexing high-dimensional spaces: Database support for next decade's applications. ACM Computing Surveys, 33(3):322–373, 2001.
3. G. Bhalotia, C. Nakhe, A. Hulgeri, S. Chakrabarti, and S. Sudarshan. Keyword

- searching and browsing in databases using banks. In ICDE, 2002.
4. C. Buckley, A. Singhal, M. Mitra, and G. Salton. New retrieval approaches using smart. In TREC 4., pages 25–48., 1995.
 5. A. Crespo and H. Garcia-Molina. Routing indices for peer-to-peer systems. In 28th Intl. Conf. on Distributed Computing Systems, 2002.
 6. F. M. Cuenca-Acuna and T. D. Nguyen. Text-based content search and retrieval in ad hoc p2p communities. In International Workshop on Peer-to-Peer Computing, 2002.
 7. Freenet. <http://freenet.sourceforge.com/>
 8. Gnutella. <http://gnutella.wego.com/>.
 9. V. Hristidis and Y. Papakonstantinou. Discover: Keyword search in relational databases. In VLDB, 2002.
 10. Napster. <http://www.napster.com/>.
 11. SVD Package. <http://www.netlib.org/svdpack/>.
 12. C. Palmer and J. Steffan. Generating network topologies that obey power law. In GLOBECOM, 2000.
 13. C.H. Papadimitriou, H. Tamaki, P. Raghavan, and S. Vempala. Latent semantic indexing: A probabilistic analysis. In PODS, 1998.
 14. S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker. A scalable content-addressable network. In SIGCOMM, 2001.
 15. Stoica, R. Morris, D. Karger, F. Kaashoek, and H. Balakrishnan. Chord: A scalable peer-to-peer lookup service for internet applications. In SIGCOMM, 2001.
 16. P. Triantafillou, C. Xiruhaki, M. Koubarakis, and N. Ntarmos. Towards high performance peer-to-peer content and resource sharing systems. In CIDR, 2003.
 17. R. Weber, H. Schek, and S. Blott. A quantitative analysis and performance study for similarity search methods in high dimensional spaces. In VLDB, pages 194–205, 1998.
 18. S. K.M. Wong, W. Ziarko, V. V. Raghavan, and P. C.N. Wong. On modeling of information retrieval concepts in vector spaces. In TODS, 1987.
 19. B. Yang and H. Garcia-Molina. Comparing hybrid peer-to-peer systems. In VLDB'2001, 2001.
 20. B. Yang and H. Garcia-Molina. Improving efficiency of peer-to-peer search. In 28th Intl. Conf. on Distributed Computing Systems, 2002.
 21. B. Yang and H. Garcia-Molina. Designing a super-peer network. In ICDE, 2003.
 22. Sen, S., And Wang, J. Analyzing peer-to-peer traffic across large networks. In Proc. of ACM SIGCOMM Internet Measurement Conference (2002).
 23. Sripanidkulchai, K. The popularity of Gnutella queries and its implications on scalability. In O'Reilly's OpenP2P (2001).
 24. Stoica, I., Morris, R., Karger, D., Kaashoek, F., And Balakrishnan, H. Chord: A scalable peer-to-peer lookup service for internet applications. In Proc. of SIGCOMM (2001).
 25. Tang, C., Xu, Z., And Dwarkadas, S. Peer-to-peer information retrieval using self-organizing semantic overlay networks. In Proc. of SIGCOMM (2003).
 26. Yang, B., And Garcia-Molina, H. Efficient search in peerto-peer networks. In Proc. of the International Conference on Distributed Computing Systems (ICDCS) (2002).
 27. Zhao, B. Y., Kubiawicz, J., And Joseph, A. D. Tapestry: An infrastructure for fault-tolerant wide-area location and routing. Tech. Report UCV/CSD-01-1141, Computer Science Division, UC, Berkeley, 2001.