Peer Reviewed Article

Convergence of on-line community technologies: Internet relay chat (IRC) and peer-to-peer (P2P) file sharing

C.J.P. Niemand
Library and Information Centre
Rand Afrikaans University
cni@rau.ac.za

C.W. Rensleigh
Department of Information Studies
Rand Afrikaans University
cr@rau.ac.za

Contents

1 Introduction

The Internet makes it possible to bring together people with common interests and/or ideas in a virtual community without ever having to meet face-to-face. These virtual communities have existed for over three decades (Cothrel 2000:21). However, the dramatic surge in Internet access in recent years, which has been a result of the availability of greater processing power at lower prices, has expanded and diversified the on-line communities as never before. Client–server technology has been the main technology used by these on-line communities. In recent years, peer-to-peer technology has identified new possibilities for on-line communities.

In this article, on-line communities are discussed, as well as how peer-to-peer technology is used in conjunction with client–server technology to expand the functionality of these on-line communities. In the first part of this article, on-line communities are discussed, in the second part, peer-to-peer computing and in the last part Internet relay chat (IRC) and how peer-to-peer file-sharing technology has been incorporated into a client–server environment.

2 On-line community

2.1 Social communities revisited

Human life exists in a network of relationships. Without the basic concept of relationships, healthy human development cannot take place. Because humans need stability and familiarity within their developmental stage, they must be in the same place for extended periods of time. According to Fernback and Thompson (1995), the structural process that is associated with a community is communication.

The words community and communication both stem from the same Latin root word, communis, which means common (Williams 1983). A community can be defined as follows:

- A combination of place, content and people, and the interactions and relationships between them (Tillman 2000)
- A group of people with a shared interests, purposes or goals, who get to know each other better over time (Kim 2000:28).

The above definitions correspond with the *Random House dictionary of the English language*, which defines a community as follows:

'A social group of any size whose members resides in a specific locality, share government and often have a common cultural and historical heritage. Or, a group sharing common characteristics or interests and perceived or perceiving itself as distinct in some respect from the larger society within which it exists' (Random House 1987:414).

2.2 On-line communities defined

According to Balasubramanian and Mahajan (2000), the origins of the virtual or on-line community can be traced to the Electronic Information Exchange System (EIES) implemented in 1976. The EIES focused on computerized conferencing that would allow users to exercise a collective intelligence capability (Kitchin 1998).

It should be noted that in the literature the terms 'on-line community' and 'virtual community' are used interchangeable. Some definitions of an on-line or virtual community may include the following:

- Groups of people who communicate with each other via electronic media (Romm, Pliskin and Clarke 1997:261)
- A group of people who come together on the World-Wide Web to communicate, bond, and/or perform particular functions. It can operate with or without a membership (Dillard and Hennard 2002:29).

In essence, the collective term or definition for an on-line or virtual community can be summarized as 'the vibrant new village of activity within the larger cultures of computing' (Escobar 2001:64).

2.3 Characteristics of on-line communities

Uzokwelu and Gunnarsdottir (2001) identify seven criteria that characterize on-line or virtual communities. According to them, virtual communities are:

- All kinds of people and all kinds of interests– the concept virtual community refers to an on-line environment where all kinds of people come together to do any, some of or

all of the following activities: exchange information and ideas, share common interests, entertain each other, seek different kinds of help, offer support to one another, trade goods and services, and foster relationships.

- User based – virtual or on-line communities provide an environment for users to connect and communicate with each other. The more users the community has, the more value is added to the community.
- Specific – the distinctive interests of the members of an on-line community determine the scope and specific objectives of that community.
- Organised – on-line communities have organizers or organizing bodies who facilitate the growth and maintenance of the community.
- Integrated – virtual communities make it possible for the integration of content with communication.
- Efficient – on-line communities help companies to save time and money when they are used to build better and watertight business processes and company–customer relationships.
- Interactive – this might be the best aspect of on-line communities. It is possible for community members to communicate one to one, one to many and many to many.

2.4 On-line community activities

According to Agre (1998:82), the life of every community includes some form of activity within a particular institutional logic. This includes the activities through which particular kinds of media are used. Community-building tools or activities may include the following (Boettcher 1999):

- E-mail – the e-mail can be seen as one of the simplest and widely used computer-mediated communications (CMC). Not only was e-mail one of the first Internet applications, it was also one of the reasons why many people were originally attracted to the Internet (Schneider 2002:49). Within a community, an e-mail list (or listserv) can be seen as a tool that connects people of the community via e-mail messages. There is one central address to which a member of the community sends messages for the group. A person who receives mail has the choice to respond either to the sender or to the whole list of members within the community.
- Newsgroups – mewsgroups are to mailing lists what bulletin boards are to newsletters. Rather than cluttering a mailbox with messages from mailing lists, the user can follow certain topics at his or her convenience. The messages are kept in files that can be read using a newsreader program.
- Chat – chat can be defined as communication by community members who are on-line at the same time, typing messages to each other. Chat can be performed in one of two situations, namely public chat rooms (open to anyone) and private chat rooms (where only those of the community can enter). Chat can be seen as a many-to-many communication tool (there are a number of people in a chat room conversing at the same time), but it can also be used for one-on-one meetings and other work-oriented applications.
- Message boards – the on-line version of a message board is much like a message board in an office environment. The on-line message board communication can be seen as asynchronous (the message board participants do not have to be on-line at the same time). According to Boettcher (1999), there are two ways to organize messages in a message board system, namely threaded and linear. With a threaded system, messages are arranged into 'threads' or topics. A message is attached to the message to which it is replying, whether or not it appears in chronological order. With a linear system, each posting in a given topic arrives in chronological order. The result is more like a real conversation. Linear message boards are sometimes called conferencing.

3 Peer-to-peer computing

The concept of to peer-to-peer (P2P) computing is not new. The basic form of P2P computing has existed since the Internet took form during the 1970s. Today's P2P infrastructure is not just the collection of what was done during the 1970s, but is also the sum of the evolutions of earlier experiments (Barkai 2002:39).

3.1 Defining P2P computing

Depending on the definition, P2P computing can be seen as one of the oldest architectures in the world of telecommunication. The telephone system, the discussion forums of Usenet, or the early form of the Internet can be classified as P2P systems (Schoder and Fischbach 2003:27).

What is P2P? Is it a set of protocols, an IT architecture, a design philosophy stressing decentralization, a business model or merely a fad? In answering the above question, a simple definition of P2P computing is that 'it refers to the topology and architecture of the computers in a system' (Bricklin 2000) 'in which each computer has equivalent capabilities and responsibilities' (Miller 2001:19). As can be seen from the above definition, P2P is in complete contrast to the traditional client–server network architecture, 'for each participating computer or node in a P2P system is called a "peer" meaning that the participants interact as equals' (Kubiatowicsz 2003:33). This means that participants act both as a client and a server, as illustrated in Figure 1.

Figure 1 Simplified configuration of P2P and client–server network
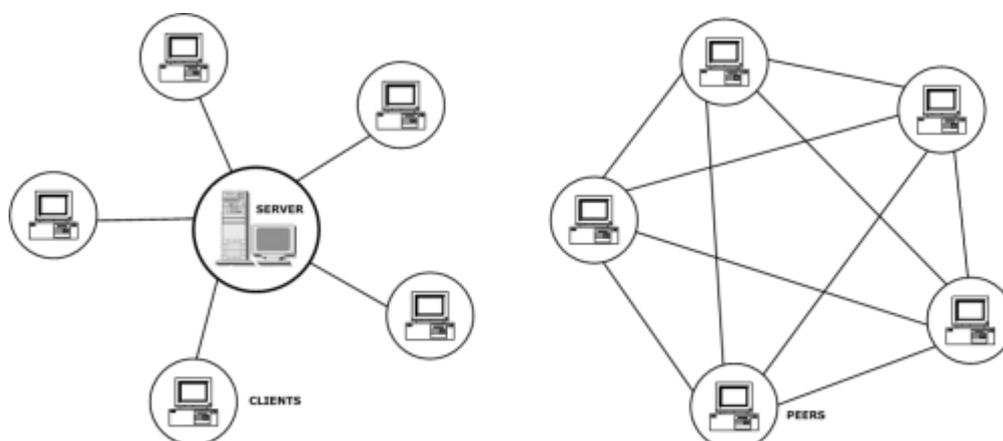
Another way to characterize P2P systems is not by a common technology but rather by what it does (Lee 2002). With P2P, computers share data and resources, such as spare computing cycles and storage capacity via the Internet or private networks. Machines can also communicate directly and manage computing tasks without using central servers (Clark 2001:18). In essence, there is a blur in the distinction between server, client and router because individual computers fulfil these roles, communicating and sharing resources without dedicated servers (Lee 2003:49).

Building on the above definitions, Miller (2001:19) identified the following five key characteristics within a P2P system:

- The network facilitates real-time transmission of data or messages between the peers
- Peers can function as both client and server

- The primary content of the network is provided by peers
- The network gives control and autonomy to the peer
- The network accommodates peers who are not always connected and who may not have permanent Internet Protocol (IP) addresses.

Related terms that are sometimes incorrectly used as interchangeable with P2P computing are P2P networking and P2P communications. A P2P network allows every computer in the network to act as a server to every other user on the network. Dedicated servers may or may not be present. P2P communications enable both participants to initiate, manage and terminate the session (Barkai 2002:19). It can therefore be said that a P2P network implies the P2P communication between the computers in the network. However, it should be noted that P2P communication could also occur between computers in a network that is not a P2P network.

In conclusion, P2P networks and communications refer to physical networks and the modes of communications. In contrast, P2P computing refers to an end-user application level environment.

3.2 Various P2P models

All computer systems can be classified into either a distributed or a centralized system, as depicted in Figure 2. Within the taxonomy of computer systems, the distributed systems can then be sub-divided into the client–server model and the P2P model (Milojicic, Kalogeraki, Lukose, Nagaraja, Pruyne, Richard, Rollins and Xu 2002). The P2P model can further be sub-divided into the following categories:

- Pure P2P
- P2P with a simple discovery server
- P2P with discovery and lookup servers
- P2P with discovery, lookup and content servers (Dreamtech 2002:2–5).

Figure 2 Taxonomy of computer systems
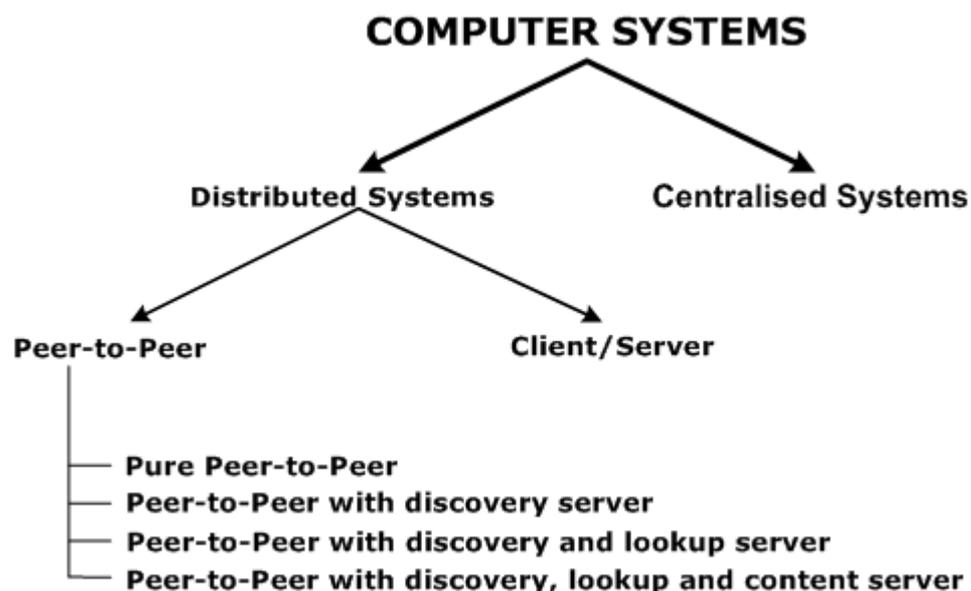
3.2.1 Pure P2P
The pure P2P model consists of and is entirely dependent on personal computers. This architecture is different from the traditional networking model, consisting of personal

computers and servers. After installing the P2P application on the computer, a dynamic search for other peers is done and a connection is established. The entire communication process occurs without any assistance from servers. In essence, the pure P2P model allows users or clients to set up and determine their own unique networking environment.

3.2.2 P2P with discovery server

Within this environment, as the name suggests, a server provides and functions as a basic administrator. The primary role of the server is to provide the names of already connected peers to the new peer. A new peer notifies the server of his or her existence by logging in and then receives a list of names. It should, however, be noted that the server only provides the names of peers who are connected. The establishment of a connection and communication remain the responsibility of the peers. For resource sharing, a peer must approach each peer within a specified environment until the resource is found. This is in contrast to the traditional client–server architecture, where all content resources are located on the server. The process of searching for resources on the P2P with a simple discovery server is relatively time consuming.

3.2.3 P2P with discovery and lookup server

Within this model, the features of the pure P2P and P2P with a simple discovery server are enhanced with another feature of a server, namely looking at the resources available within the P2P network. This architecture is an enhancement in comparison to the P2P with a basic discovery server, for it reduces the amount of time wasted by visiting each individual peer to obtain shared resource information. In essence, the server initiates the communication process between two peers.

3.2.4 P2P with discovery, lookup and a content server

In typical client–server architecture, all the facets of requesting and requiring information and resources reside with the server. Furthermore the individual peers are not allowed to connect to each other directly. If a peer requires a specific resource, he or she communicates with the server. In turn, the server processes the request and displays the information. A possible disadvantage of this architecture is slower processing of requests by the server, as the volume of requests increases. Another is higher costs and the possibility of single point failure, as the server is responsible for handling all processes.

3.3 Applications of P2P computing

A typical P2P application can be based on any of the models discussed in the previous section, but some prerequisites exist for an application to be an ideal P2P application. Some of the key features that every model of a P2P application should have are the following:

- Tracing other peers: finding other peers connected to the network
- Querying peers for required content: getting lists of shared resources from other peers
- Sharing content or resources with other peers: learning how content are shared among the peer network (Dreamtech 2002:8).

There are at least five distinct types of P2P applications in use today (Miller 2001:30). They are file sharing, instant messaging, distributed search, group collaboration and distributed computing. In this article, only file sharing and instant messaging (IM) are discussed.

3.3.1 P2P file sharing

Mention the word P2P and most people reply 'Napster'. Napster is possibly the best-known type of P2P application. Fortunately, file sharing is much more than Napster or MPEG Audio Layer 3 (MP3) music files. A file can contain any type of information, such as a research report, a case study or a travel journal. Files may also contain multiple formats, for example

a text document may include embedded graphics (Moore and Hebeler 2002:37). The traditional definition of file sharing in a typical shared distributed client–server file system is where two or more different users access the same file at the same time. File sharing in the P2P sense is the transfer of a copy of the file (Barkai 2002:299–300). Two computers connected to a network can copy files from one to the other. Since the Internet is nothing more than a giant global network, there is no reason why any two computers connected to the Internet cannot share files. The process whereby two computers share and copy files over the Internet is what P2P file sharing is all about (Miller 2001:97).

The fundamental challenge with P2P file-sharing networks is that it operates outside the Internet domain name server (DNS) addressing system. According to Miller (2001:98), this enables each peer on the network to connect regardless of which Internet service provider (ISP) is used and, in some cases, ensures anonymity for the individual personal computer (PC).

After a connection to a specific P2P network is established, the next step in file sharing is the search for a specific file. The specific P2P model being used determines how this task takes place. A pure P2P model does not need an index server and files are found in real-time. In contrast, the 'brokered' models use the creation of an index of useable or shared files on servers. Once the specific file is located, the transfer is done via the hypertext transfer protocol (HTTP).

The following are some examples of P2P Internet file-sharing systems:

- KaZaA (www.kazaa.com)
- Freenet (www.freenetproject.org)
- Gnutella (gnutella.wego.com)
- iMesh (www.imesh.com)
- eDonkey2000 (www.edonkey2000.com).

The commonly shared file types on P2P Internet file-sharing systems include the following:

- Audio and/or video Files, for example MP3 and Audio Video Interleaved (AVI) files
- Images
- Documents
- Software.

It should be noted that successful P2P computing is a two-way street of uploading and downloading files. A serious concern is the problem of 'freeloaders'. A freeloader can be defined as an individual peer within the P2P computing environment who downloads more resources then he or she uploads within the system. These freeloaders or 'Net leeches' defeat the goal of a fully decentralized system and can be viewed as a potential cause of the downfall of the P2P system (Lee 2002). P2P systems can combat this effect by incorporating coercive measures into their design and deployment, ensuring that users actually donate resources (Dingledine and Freedman 2001:292). This sentiment is supported by Lee (2002), who adds that systems should be equipped with anti-freeloading mechanisms, such as tying the number of downloads to the number of uploads or using a point system that rewards the contributors.

3.3.2 P2P communication

With more than 200 million users, instant messaging (IM) is by far the largest current P2P application – even if most IM users do not realise that they are engaging in P2P computing (Miller 2001:30). This is supported by Moore and Hebeler (2002:20) who state that all IM platforms are P2P and that they operate in a private space that is created among the peers

talking to one another.

All that is required for IM is software on the users side, and Internet protocol addresses of the two peers who want to communicate. The IM network operates outside the DNS, allowing users to connect in real-time to perform a text-based conversation. Within this network, the messages flow from one user to another without the intervention of a server. It should be noted that each of the major IM environments operate on its own proprietary network, making communications between them virtually impossible.

The following are some of the available IM environments

- American online (AOL) Instant Messenger (www.aim.com)
- ICQ (Web.icq.com)
- Microsoft (MSN) Messenger (messenger.msn.com)
- Yahoo! Messenger (messenger.yahoo.com) (Miller 2001:31–32).

4 Internet relay chat and P2P

Internet relay chat (IRC) is a virtual meeting place where people from all over the world can meet and communicate (text based) on a whole diversity of interests, ideas and issues, independent of where they are in the world. IRC allows people to participate in group discussions on one of the many thousands of IRC channels on hundreds of IRC networks, or they can talk in private to family and friends (Vonck 2003). It should be noted that IRC is not the only form of Internet-based chat. A substantial amount of Internet users use Web-based chat forums instead of IRC. These chat forums are accessible via a Web browser, whereas client software must be downloaded and installed for IRC.

IRC is one of the most popular and most interactive services on the Internet. Using the IRC client software to connect to an IRC network (server), a user can

- Exchange text messages with other users
- Engage in real-time communications
- Hold chats with multiple simultaneous users
- Send files back and forth between users (Miller 2001:215).

The IRC networks are constructed via traditional client–server technology that requires users (clients) to connect to an IRC server. All chat channels as well as group discussions are hosted on the server. In addition to using client–server technology, an IRC user can connect and communicate outside of the IRC network (client and server) directly with another user who uses the direct client-to-client (DCC) facility, which can be seen as P2P technology.

In addition to communicating in a P2P fashion, users can share files that can be downloaded by other users in the system. The shared files are not loaded to a central server and are only available when the user who is sharing the files is connected to the Internet (system). This is a form of P2P with a discovery and lookup server. Search software using Web-based technology, such as xDCCSPY and IRCSPY, can be used to search for shared resources that are available on the participating users' machines. In essence, both client–server and P2P technologies are being used from within the same system, creating a 'best of both worlds' scenario as depicted in Figure 3.

Figure 3 Simplified view of Internet relay chat (IRC) using both client–server and P2P technologies
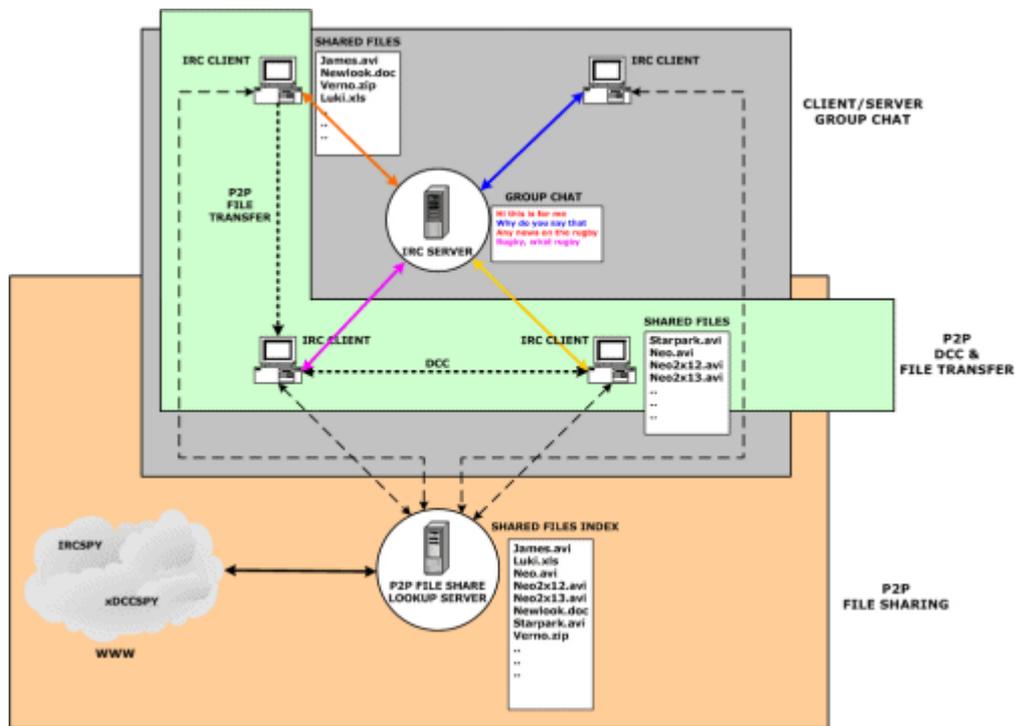
5 Conclusion

On-line communities have existed for almost three decades and it seems that their user base continues to grow. When the Internet became part of our everyday business, client–server technologies were favoured as a platform for new applications. As of late, the use of P2P technology in on-line applications has been expanding rapidly. This is even more so in on-line communities where files are shared. It seems that future developments in this field will favour the P2P technologies.

6 References

Agre, P.E. 1998. Designing genres for new media: social, economic and political contexts. In C*ybersociety 2.0: Revisiting computer-mediated communication and community*. Edited by S.G. Jones. London: Sage.

Balasubramanian, S. and Mahajan, V. 2000. *The economic leverage of the virtual community.* [Online]. Available WWW: http://cci.bus.utexas.edu/research/white/leveraging.pdf (Accessed 28 May 2003).

Barkai, D. 2002. *Peer-to-peer computing: technologies for sharing and collaborating on the Net.* Hillsboro, Oregon: Intel Press.

Boettcher, S. 1999. *What types of virtual communities can I build and what tools are available?* [Online]. Available WWW: http://www.fullcirc.com/community/communitytypes.htm (Accessed 28 May 2003).

Bricklin, D. 2000. *Thoughts on peer-to-peer.* [Online]. Available WWW: http://www.bricklin.com/p2p.htm (Accessed 15 August 2003).

Clark, D. 2001. Face-to-face with peer-to-peer networking. *Computer* (January):18–21.

Cothrel, J.P. 2000. Measuring the success of an online community. *Strategy and Leadership* 28(2):18–21.

Dillard, M. and Hennard, J. 2002. The approaching age of virtual nations. *Futurist* 36(4):24–29.

Dingledine, R. and Freedman, M.J. 2001. Accountability. In: *Peer-to-peer: harnessing the benefits of a disruptive technology*. Edited by A. Oram. Sebastopol, California: O'Reilly.

Dreamtech Software Team. 2002. *Cracking the code: peer-to-peer application development*. New York: Hungry Minds.

Escobar, A. 2001. Welcome to Cyberia. Notes on the anthropology of cyber culture. In: *The Cybercultures Reader*. Edited by D. Bell. New York: Routledge.

Fernback, J. and Thompson, B. 1995. V*irtual communities: abort, retry, failure?* [Online]. Available WWW: http://www.well.com/user/hlr/texts/VCcivil.html (Accessed 30 May 2003).

Kim, A.J. 2000. *Community building. secret strategies for successful online communities*. New York: Peachpit Press.

Kitchin, R. 1998. *Cyberspace*. West Sussex: John Wiley.

Kubiatowicsz, J. 2003. Extracting guarantees from chaos. *Communications of the ACM* 46 (2):33–38.

Lee, J. 2003. An end user perspective on file sharing systems. *Communications of the ACM* 46(2):49–53.

Lee, J. 2002. *Peer-to-peer file sharing systems: what matters to the end user*. [Online]. Available WWW: http://ccs.mit.edu/papers/pdf/wp217.pdf (Accessed 15 June 2003).

Miller, M. 2001. *Discovering P2P*. San Francisco: Sybex.

Milojicic, D.S., Kalogeraki, V., Lukose, R., Nagaraja, K., Pryune, J., Richard, B., Rollins, S. and Xu, Z. 2002. *Peer-to-peer computing*. [Online]. Available WWW: http://www.hpl.hp.com/techreports/2002/HPL-2002-57.pdf (Accessed 15/06/2003).

Moore, D. and Hebeler, J. 2002. *Peer-to-peer: building secure, scalable and manageable networks*. Berkley, California: McGraw-Hill.

Random House dictionary of the English language. 2nd ed. 1987. New York: Random House.

Romm, C., Pliskin, N. and Clarke, R.J. 1997. Virtual communities and society: toward an integrative three-phase model. *International Journal of Information Management* 17(4):261–270.

Schoder, D. and Fischbach, K. 2003. Peer-to-peer prospects. *Communications of the ACM* 46 (2):27–29.

Schneider, G.P. 2002. *Electronic commerce*. 3rd ed. Boston, Massachusetts: Thomson.

Tillman, H.N. 2000. *Virtual community building using Internet tools.* [Online]. Available WWW: http://www.hopetillman.com/il00/vc.html (Accessed 2 June 2003).

Uzokwelu, C.J. and Gunnarsdottir, L.A. 2001. *What roles do virtual communities play in the new economy?* [Online]. Available WWW: http://www.it-c.dk/~nice/Virtual_Communities_site/printversion.pdf (Accessed 2 June 2003).

Vonck, T. 2003. *Welcome to the MIRC homepage!* [Online]. Available WWW: http://www.mirc.com/index.html (Accessed 4 August 2003).

Williams, R. 1983. *Keywords: a vocabulary of culture and society*. New York: Oxford University Press.

Disclaimer

top