Chapter I

Core Concepts in Peer-to-Peer Networking

Detlef Schoder
University of Cologne, Germany

Kai Fischbach
University of Cologne, Germany

Christian Schmitt
University of Cologne, Germany

Abstract

This chapter reviews core concepts of peer-to-peer (P2P) networking. It highlights the management of resources, such as bandwidth, storage, information, files, and processor cycles based on P2P networks. A model differentiating P2P infrastructures, P2P applications, and P2P communities is introduced. This model provides a better understanding of the different perspectives of P2P. Key technical and social challenges that still limit the potential of information systems based on P2P architectures are discussed.

Introduction

Peer-to-peer (P2P) has become one of the most widely discussed terms in information technology (Schoder, Fischbach, &Teichmann, 2002; Shirky, Truelove, Dornfest, Gonze, & Dougherty, 2001). The term *peer-to-peer* refers to the concept that in a network of equals (peers) using appropriate information and communication systems, two or more individuals are able to spontaneously collaborate without necessarily needing central coordination (Schoder & Fischbach, 2003). In contrast to client/server networks, P2P networks promise improved scalability, lower cost of ownership, self-organized and decentralized coordination of previously underused or limited resources, greater fault tolerance, and better support for building ad hoc networks. In addition, P2P networks provide opportunities for new user scenarios that could scarcely be implemented using customary approaches.

This chapter is structured as follows: The first paragraph presents an overview of the basic principles of P2P networks. Further on, a framework is introduced which serves to clarify the various perspectives from which P2P networks can be observed: P2P infrastructures, P2P applications, P2P communities. The following paragraphs provide a detailed description of each of the three corresponding levels. First, the main challenges—namely, interoperability and security—of P2P infrastructures, which act as a foundation for the above levels, are discussed. In addition, the most promising projects in that area are highlighted. Second, the fundamental design approaches for implementing P2P applications for the management of resources, such as bandwidth, storage, information, files, and processor cycles, are explained. Finally, socioeconomic phenomena, such as free-riding and trust, which are of importance to P2P communities, are discussed. The chapter concludes with a summary and outlook.

P2P Networks: Characteristics and a Three-Level Model

The shared provision of distributed resources and services, decentralization and autonomy are characteristic of P2P networks (M. Miller, 2001; Barkai, 2001; Aberer & Hauswirth, 2002, Schoder & Fischbach, 2002; Schoder et al., 2002; Schollmeier, 2002):

1. Sharing of distributed resources and services: In a P2P network each node can provide both client and server functionality, that is, it can act as both a

provider and consumer of services or resources, such as information, files, bandwidth, storage and processor cycles. Occasionally, these network nodes are referred to as servents—derived from the terms client and server.

2. Decentralization: There is no central coordinating authority for the organization of the network (setup aspect) or the use of resources and communication between the peers in the network (sequence aspect). This applies in particular to the fact that no node has central control over the other. In this respect, communication between peers takes place directly.

   Frequently, a distinction is made between pure and hybrid P2P networks. Due to the fact that all components share equal rights and equivalent functions, pure P2P networks represent the reference type of P2P design. Within these structures there is no entity that has a global view of the network (Barkai, 2001, p. 15; Yang & Garcia-Molina, 2001). In hybrid P2P networks, selected functions, such as indexing or authentication, are allocated to a subset of nodes that as a result, assume the role of a coordinating entity. This type of network architecture combines P2P and client/server principles (Minar, 2001, 2002).

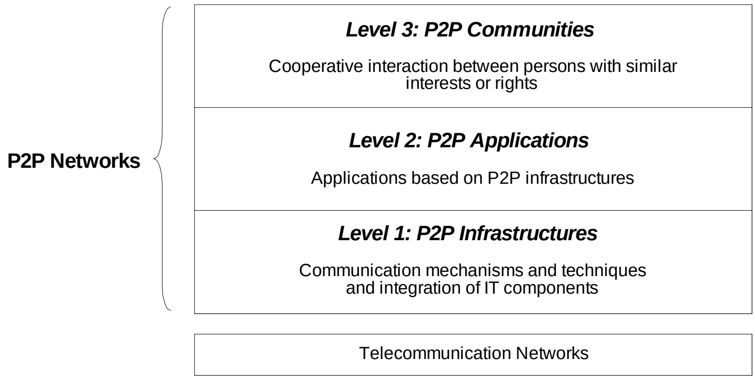
3. Autonomy: Each node in a P2P network can autonomously determine when and to what extent it makes its resources available to other entities.

On the basis of these characteristics, P2P can be understood as one of the oldest architectures in the world of telecommunication (Oram, 2001). In this sense, the Usenet, with its discussion groups, and the early Internet, or ARPANET, can be classified as P2P networks. As a result, there are authors who maintain that P2P will lead the Internet back to its origins—to the days when every computer had equal rights in the network (Minar & Hedlund, 2001).

Decreasing costs for the increasing availability of processor cycles, bandwidth, and storage, accompanied by the growth of the Internet have created new fields of application for P2P networks. In the recent past, this has resulted in a dramatic increase in the number of P2P applications and controversial discussions regarding limits and performance, as well as the economic, social, and legal implications of such applications (Schoder et al., 2002; Smith, Clippinger, & Konsynski, 2003). The three level model presented below, which consists of P2P infrastructures, P2P applications, and P2P communities, resolves the lack of clarity in respect to terminology, which currently exists in both theory and practice.

Level 1 represents P2P infrastructures. P2P infrastructures are positioned above existing telecommunication networks, which act as a foundation for all levels. P2P infrastructures provide communication, integration, and translation functions between IT components. They provide services that assist in locating

Figure 1. Levels of P2P networks



and communicating with peers in the network and identifying, using, and exchanging resources, as well as initiating security processes such as authentication and authorization.

Level 2 consists of P2P applications that use services of P2P infrastructures. They are geared to enable communication and collaboration of entities in the absence of central control.

Level 3 focuses on social interaction phenomena, in particular, the formation of communities and the dynamics within them.

In contrast to Levels 1 and 2, where the term *peer* essentially refers to technical entities, in Level 3 the significance of the term *peer* is interpreted in a nontechnical sense (peer as person).

P2P Infrastructures

The term *P2P infrastructures* refers to mechanisms and techniques that provide communication, integration, and translation functions between IT components in general, and applications, in particular. The core function is the provision of interoperability with the aim of establishing a powerful, integrated P2P infrastructure. This infrastructure acts as a "P2P Service Platform" with standardized APIs and middleware which in principle, can be used by any application (Schoder & Fischbach, 2002; Shirky et al., 2001; Smith et al., 2003).

Among the services that the P2P infrastructure makes available for the respective applications, security has become particularly significant (Barkai, 2001). Security is currently viewed as the central challenge that has to be resolved if P2P networks are to become interesting for business use (Damker, 2002).

Interoperability

Interoperability refers to the ability of any entity (device or application) to speak to, exchange data with, and be understood by any other entity (Loesgen, n.d.). At present, interoperability between various P2P networks scarcely exists. The developers of P2P systems are confronted with heterogeneous software and hardware environments as well as telecommunication infrastructures with varying latency and bandwidth. Efforts are being made, however, to establish a common infrastructure for P2P applications with standardized interfaces. This is also aimed at shortening development times and simplifying the integration of applications in existing systems (Barkai, 2001; Wiley, 2001). In particular, within the World Wide Web Consortium (W3C) (W3C, 2004) and the Global Grid Forum (GGF, n.d.) discussions are taking place about suitable architectures and protocols to achieve this aim. Candidates for a standardized P2P infrastructure designed to ensure interoperability include JXTA, Magi, Web services, Jabber, and Groove (Baker, Buyya, & Laforenza, 2002; Schoder et al., 2002).

- JXTA is an open platform that aims at creating a virtual network of various digital devices that can communicate via heterogeneous P2P networks and communities. The specification includes protocols for locating, coordinating, monitoring and the communication between peers (Gong, 2001; Project JXTA, n.d.).
- Magi is designed to set up secure, platform-independent, collaborative applications based on Web standards. A characteristic of Magi is the shared use of information and the exchange of messages between devices of any type, in particular, handheld devices (Bolcer et al., 2000).
- Web services are frequently classified as an application area in the context of P2P. However, they represent a complementary conceptual design, which can be used as one of the technological foundations for P2P applications. Evidence of this, for example, is the growing change in direction of the middleware initiatives of the National Science Foundation and the Global Grid Forum toward Web services. Both initiatives enjoy wide acceptance in both research and practice and are leading the way in the continuing development of grid computing. With their own initiatives, important players in the market, such as Microsoft with .NET (Dasgupta, 2001), IBM with WebSphere, and Sun with SunONE, are pushing forward the development of Web services. Key technologies of Web services are

the Extensible Markup Language as a file format, the Simple Object Access Protocol for communication, the Web Services Description Language for the description of services, the Web Services Flow Language for the description of work flows, and Universal Description, Discovery, and Integration for the publication and location of services (Baker et al., 2002; Bellwood et al., 2003; Web services activity, 2004; Wojciechowski & Weinhardt, 2002).

- The Jabber Open Source Project (Jabber, 2004) is aimed at providing added value for users of instant messaging systems. Jabber functions as a converter, providing compatibility between the most frequently used and incompatible instant messaging systems of providers, such as Yahoo, AOL, and MSN. This enables users of the Jabber network to exchange messages and present information with other peers, regardless of which proprietary instant messaging network they actually use. Within the framework of the Jabber-as-Middleware-Initiative, Jabber developers are currently working on a protocol that is aimed at extending the existing person-to-person functionality to person-to-machine and machine-to-machine communication (J. Miller, 2001).
- The Groove platform provides system services that are required as a foundation for implementing P2P applications. A well-known sample application that utilizes this platform is the P2P Groupware Groove Virtual Office. The platform provides storage, synchronization, connection, security and awareness services. In addition, it includes a development environment that can be used to create applications or to expand or adapt them. This facilitates the integration of existing infrastructures and applications (such as Web Services or .NET Framework) (Edwards, 2002).

Security

The shared use of resources frequently takes place between peers that do not know each other and, as a result, do not necessarily trust each other. In many cases, the use of P2P applications requires granting third parties access to the resources of an internal system, for example, in order to share files or processor cycles. Opening an information system to communicate with, or grant access to, third parties can have critical side effects. This frequently results in conventional security mechanisms, such as firewall software, being circumvented. A further example is communication via instant messaging software. In this case, communication often takes place without the use of encryption. As a result, the security goal of confidentiality is endangered. Techniques and methods for providing authentication, authorization, availability checks, data integrity, and confidentiality are among the key challenges related to P2P infrastructures (Damker, 2002).

A detailed discussion of the security problems which are specifically related to P2P, as well as prototypical implementations and conceptual designs can be found in Barkai (2001), Damker (2002), Udell, Asthagiri, and Tuvell (2001), Groove Networks (2004), Grid Security (2004), Foster, Kesselman, Tsudic, and Tuecke (1998), and Butler et al. (2000).

P2P Applications: Resource Management Aspects

In the respective literature, P2P applications are often classified according to the categories of instant messaging, file sharing, grid computing and collaboration (Schoder & Fischbach, 2002; Shirky et al., 2001). This form of classification has developed over time and fails to make clear distinctions. Today, in many cases the categories can be seen to be merging. For this reason, the structure of the following sections is organized according to resource aspects, which in our opinion, are better suited to providing an understanding of the basic principles of P2P networks and the way they function. Primary emphasis is placed on providing an overview of possible approaches for coordinating the various types of resources, that is, information, files, bandwidth, storage, and processor cycles in P2P networks.

Information

The following sections explain the deployment of P2P networks using examples of the exchange and shared use of presence information, of document management, and collaboration.

- Presence information: Presence information plays a very important role in respect to P2P applications. It is decisive in the self-organization of P2P networks because it provides information about which peers and which resources are available in the network. It enables peers to establish direct contact to other peers and inquire about resources. A widely distributed example of P2P applications that essentially use presence information are instant messaging systems. These systems offer peers the opportunity to pass on information via the network, such as whether they are available for communication processes. A more detailed description of the underlying architecture of instant messaging system can be found in Hummel (2002).

The use of presence information is interesting for the shared use of processor cycles and in scenarios related to omnipresent computers and information availability (ubiquitous computing). Applications can independently recognize which peers are available to them within a computer grid and determine how intensive computing tasks can be distributed among idle processor cycles of the respective peers. Consequently, in ubiquitous computing environments it is helpful if a mobile device can independently recognize those peers which are available in its environment, for example in order to request Web Services, information, storage or processor cycles. The technological principles of this type of communication are discussed in Wojciechowski and Weinhardt (2002).

- Document management: Customary, Document Management Systems (DMS), which are usually centrally organized, permit shared storage, management, and use of data. However, it is only possible to access data that have been placed in the central repository of the DMS. As a result, additional effort is required to create a centralized index of relevant documents. Experience shows that a large portion of the documents created in a company are distributed among desktop PCs, without a central repository having any knowledge of their existence. In this case, the use of P2P networks can be of assistance. For example, by using the NextPage-NXT 4 platform, it is possible to set up networks that create a connected repository from the local data on the individual peers (Next Page Inc., n.d.). Indexing and categorization of data is accomplished by each peer on the basis of individually selected criteria.

  In addition to linking distributed data sources, P2P applicationa can offer services for the aggregation of information and the formation of self-organized P2P knowledge networks. Opencola (Leuf, 2002) was one of the first P2P applications that offer their users the opportunity to gather distributed information in the network from the areas of knowledge that interest them. For this purpose, users create folders on their desktop that are assigned keywords that correspond to their area of interest. Opencola then searches the knowledge network independently and continuously for available peers that have corresponding or similar areas of knowledge without being dependent on centrally administered information. Documents from relevant peers are analyzed, suggested to the user as appropriate, and automatically duplicated in the user's folder. If the user rejects respective suggestions, the search criteria are corrected. The use of Opencola results in a spontaneous networking of users with similar interests without a need for a central control.

- Collaboration: P2P groupware permits document management at the level of closed working groups. As a result, team members can communicate synchronously, conduct joint online meetings, and edit shared documents, either synchronously or asynchronously. In client/server-based groupware a corresponding working area for the management of central data has to be set up and administered on the server for each working group. In order to avoid this additional administration task, P2P networks can be used for collaborative work. The currently best-known application for collaborative work based on the principles of P2P networks is Groove Virtual Office. This system offers similar functions (instant messaging, file sharing, notification, co-browsing, whiteboards, voice conferences, and data bases with real-time synchronization) to those of the widely used client/server-based Lotus products, Notes, Quickplace, and Sametime, but does not require central data management. All of the data created are stored on each peer and are synchronized automatically. If peers cannot reach each other directly, there is the option of asynchronous synchronization via a directory and relay server. Groove Virtual Offce offers users the opportunity to set up so-called shared spaces that provide a shared working environment for virtual teams formed on an ad hoc basis, as well as to invite other users to work in these teams. Groove Virtual Office can be expanded by system developers. A development environment, the Groove Development Kit, is available for this purpose (Edwards, 2002).

Files

File sharing is probably the most widespread P2P application. It is estimated that as much as 70% of the network traffic in the Internet can be attributed to the exchange of files, in particular music files (Stump, 2002) (more than one billion downloads of music files can be listed each week [Oberholzer & Strumpf, 2004]). Characteristic of file sharing is that peers that have downloaded the files in the role of a client subsequently make them available to other peers in the role of a server. A central problem for P2P networks in general, and for file sharing in particular, is locating resources (lookup problem) (Balakrishnan, Kaashoek, Karger, Morris, & Stoica, 2003). In the context of file sharing systems, three different algorithms have developed: the flooded request model, the centralized directory model, and the document routing model (Milojicic et al., 2002). These can be illustrated best by using their prominent implementations—Gnutella, Napster, and Freenet.

P2P networks that are based on the Gnutella protocol function without a central coordination authority. All peers have equal rights within the network. Search requests are routed through the network according to the flooded request model which means that a search request is passed on to a predetermined number of peers. If they cannot answer the request, they pass it on to other nodes until a predetermined search depth (ttl = time-to-live) has been reached or the requested file has been located. Positive search results are then sent to the requesting entity which can then download the desired file directly from the entity that is offering it. A detailed description of searches in Gnutella networks, as well as an analysis of the protocol, can be found in Ripeanu, Foster, and Iamnitchi (2002) and Ripeanu (2001). Due to the fact that the effort for the search, measured in messages, increases exponentially with the depth of the search, the inefficiency of simple implementations of this search principle is obvious. In addition, there is no guarantee that a resource will actually be located. Operating subject to certain prerequisites (such as nonrandomly structured networks), numerous prototypical implementations (for example, Aberer et al., 2003; Crowcroft & Pratt, 2002; Dabek et al., 2001; Druschel & Rowstron, 2001; Nejdl, et al. 2003; Pandurangan & Upfal, 2001; Ratnasamy, Francis, Handley, Karp, & Shenker, 2001; Lv, Cao, Cohen, Li, & Shenker, 2002; Zhao et al., 2004) demonstrate how searches can be effected more "intelligently" (see, in particular, Druschel, Kaashoek, & Rowstron [2002], and also Aberer & Hauswirth [2002] for a brief overview). The FastTrack protocol enjoys widespread use in this respect. It optimizes search requests by means of a combination of central supernodes which form a decentralized network similar to Gnutella.

In respect of its underlying centralized directory model, the early Napster (Napster, 2000) can be viewed as a nearly perfect example of a hybrid P2P system in which a part of the infrastructure functionality, in this case the index service, is provided centrally by a coordinating entity. The moment a peer logs into the Napster network, the files that the peer has available are registered by the Napster server. When a search request is issued, the Napster server delivers a list of peers that have the desired files available for download. The user can obtain the respective files directly from the peer offering them.

Searching for and storing files within the Freenet network (Clarke, Miller, Hong, Sandberg, & Wiley, 2002; Clarke, 2003) takes place via the so-called document routing model (Milojicic et al., 2002). A significant difference to the models that have been introduced so far is that files are not stored on the hard disk of the peers providing them, but are intentionally stored at other locations in the network. The reason behind this is that Freenet was developed with the aim of creating a network in which information can be stored and accessed anonymously. Among other things, this requires that the owner of a network node does not know what documents are stored on his/her local hard disk. For this reason, files and peers are allocated unambiguous identification numbers. When a file is created, it is transmitted, via neighboring peers, to the peer with the identification number that is numerically closest to the identification number of the file and is stored there. The peers that participate in forwarding the file save the identification number of the file and also note the neighboring peer to which they have transferred it in a routing table to be used for subsequent search requests.

The search for files takes place along the lines of the forwarding of search queries on the basis of the information in the routing tables of the individual peers. In contrast to searching networks that operate according to the flooded request model, when a requested file is located, it is transmitted back to the peer requesting it via the same path. In some applications, each node on this route stores a replicate of the file in order to be able to process future search queries more quickly. In this process, the peers only store files up to a maximum capacity. When their storage is exhausted, files are deleted according to the least recently used principle. This results in a correspondingly large number of replicates of popular files being created in the network, whereas, over time, files that are requested less often are removed (Milojicic et al., 2002).

In various studies (Milojicic et al., 2002), the document routing model has been proven suitable for use in large communities. The search process, however, is more complex than, for example, in the flooded request model. In addition, it can result in the formation of islands—that is, a partitioning of the network in which the individual communities no longer have a connection to the entire network (Clarke et al., 2002; Langley, 2001).

Bandwidth

Due to the fact that the demands on the transmission capacities of networks are continuously rising, in particular due to the increase in large-volume multimedia data, effective use of bandwidth is becoming more and more important. Currently, in most cases, centralized approaches in which files are held on the server of an information provider and transferred from there to the requesting client are primarily used. In this case, a problem arises when spontaneous increases in demand exert a negative influence on the availability of the files due to the fact that bottlenecks and queues develop.

Without incurring any significant additional administration, P2P-based approaches achieve increased load balancing by taking advantage of transmission routes that are not being fully exploited. They also facilitate the shared use of the bandwidth provided by the information providers.

- Increased load balancing: In contrast to client/server architectures, hybrid P2P networks can achieve a better load balancing. Only initial requests for files have to be served by a central server. Further requests can be automatically forwarded to peers within the network, which have already received and replicated these files. This concept is most frequently applied in the areas of streaming (for example, PeerCast [“PeerCast”, n.d.], P2P-

Radio ["P2P-Radio", 2004], SCVI.net ["SCVI.NET", n.d.]) and video on demand. The P2P-based Kontiki network (Kontiki, n.d.) is pursuing an additional design that will enable improved load balancing. Users can subscribe to information channels or software providers from which they wish to obtain information or software updates. When new information is available the respective information providers forward it to the peers that have subscribed. After receiving the information, each peer instantaneously acts as a provider and forwards the information to other peers. Application areas in which such designs can be implemented are the distribution of eLearning courseware in an intranet (Damker, 2002, p. 218), the distribution of antivirus and firewall configuration updates (for example, Rumor [McAfee, n.d.]) and also updating computer games on peer computers (for example, Descent (Planet DESCENT, n.d.) and Cybiko [Milojicic et al., 2002]).

- Shared use of bandwidth: In contrast to client/server approaches, the use of P2P designs can accelerate the downloading and transport of big files that are simultaneously requested by different entities. Generally, these files are split into smaller blocks. Single blocks are then downloaded by the requesting peers. In the first instance each peer receives only a part of the entire file. Subsequently, the single file parts are exchanged by the peers without a need for further requests to the original source. Eventually, the peers reconstruct the single parts to form an exact copy of the original file. An implementation utilizing this principle can be found in BitTorrent (Cohen, 2003).

Storage

Nowadays, Direct Attached Storage (DAS), Network Attached Storage (NAS), or Storage Area Networks (SAN) are the main design concepts used to store data in a company. These solutions have disadvantages such as inefficient use of the available storage, additional load on the company network, or the necessity for specially trained personnel and additional backup solutions.

However, increased connectivity and increased availability of bandwidth enable alternative forms of managing storage which resolve these problems and require less administration effort. With P2P storage networks, it is generally assumed that only a portion of the disk space available on a desktop PC will be used. A P2P storage network is a cluster of computers, formed on the basis of existing networks, that share all storage available in the network. Well-known approaches to this type of system are PAST (Rowstron & Druschel, 2001), Pasta (Moreton, Pratt, & Harris, 2002), OceanStore (Kubiatowicz et al., 2000), CFS (Dabek, Kasshoek, Karger, Morris, & Stoica, 2001), Farsite (Adya et al., 2002), and Intermemory (Goldberg & Yianilos, 1998). Systems that are particularly suitable for explaining the way in which P2P storage networks work are PAST, Pasta, and OceanStore. They have basic similarities in the way they are constructed and organized. In order to participate in a P2P storage network, each peer receives a public/private key pair. With the aid of a hash function, the public key is used to create an unambiguous identification number for each peer. In order to gain access to storage on another computer, the peer has to either make available some of its own storage, or pay a fee. Corresponding to its contribution, each peer is assigned a maximum volume of data that it can add to the network. When a file is to be stored in the network, it is assigned an unambiguous identification number, created with a hash function from the name or the content of the respective file, as well as the public key of the owner. Storing of the file and searching for it in the network take place in the manner described as the document routing model before. In addition, a freely determined number of file replicates are also stored. Each peer retrieves its own current version of the routing table which is used for storage and searches. The peer checks the availability of its neighbors at set intervals in order to establish which peers have left the network. In this way, new peers that have joined the network are also included in the table.

To coordinate P2P storage networks, key pairs have to be generated and distributed to the respective peers and the use of storage has to be monitored. OceanStore expands the administrative tasks to include version and transaction management. As a rule, these tasks are handled by a certain number of particularly high-performance peers that are also distinguished by a high degree of availability in the network. In order to ensure that a lack of availability on the part of one of these selected peers does not affect the functional efficiency of the entire network, the peers are coordinated via a Byzantine agreement protocol (Castro, 2001). Requests are handled by all available selected peers. Each sends a result to the party that has issued the request. This party waits until a certain number of identical results are received from these peers before accepting the result as correct.

By means of file replication and random distribution of identification numbers to peers using a hash function, the P2P storage network automatically ensures that various copies of the same file are stored at different geographical locations. No additional administration or additional backup solution is required to achieve protection against a local incident or loss of data. This procedure also reduces the significance of a problem which is characteristic of P2P networks: in P2P networks there is no guarantee that a particular peer will be available in the network at a particular point in time (availability problem). In case of P2P storage networks, this could result in settings where no peer is available in the network that stores the file being requested. Increasing the number of replicates stored at various geographical locations can, however, enhance the probability that at least one such peer will be available in the network.

The low administration costs, which result from the self-organized character of P2P storage networks, and the fact that additional backup solutions are seldom required are among the advantages these new systems offer for providing and efficiently managing storage.

Processor Cycles

Recognition that the available computing power of the networked entities was often unused was an early incentive for using P2P applications to bundle computing power. At the same time, the requirement for high-performance computing, that is, computing operations in the field of bio-informatics, logistics, or the financial sector, has been increasing. By using P2P applications to bundle processor cycles, it is possible to achieve computing power that even the most expensive supercomputers can scarcely provide. This is effected by forming a cluster of independent, networked computers in which a single computer is transparent and all networked nodes are combined into a single logical computer. The respective approaches to the coordinated release and shared used of distributed computing resources in dynamic, virtual organizations that extend beyond any single institution, currently fall under the term *grid computing* (Baker et al., 2002; Foster, 2002; Foster & Kesselman, 2004; Foster, Kesselman, & Tuecke, 2002; GGF, n.d.). The term *grid computing* is an analogy to customary power grids. The greatest possible amount of resources, in particular computing power, should be available to the user, ideally unrestricted and not bound to any location—similar to the way in which power is drawn from an electricity socket. The collected works of Bal, Löhr, and Reinefeld (2002) provide an overview of diverse aspects of grid computing.

One of the most widely cited projects in the context of P2P, which, however, is only an initial approximation of the goal of grid computing, is SETI@home (Search for Extraterrestrial Intelligence) (Anderson, 2001). SETI@home is a scientific initiative launched by the University of California, Berkeley, with the goal of discovering radio signals from extraterrestrial intelligence. For this purpose, a radio telescope in Puerto Rico records a portion of the electromagnetic spectrum from outer space. This data is sent to the central SETI@home server in California. There, they take advantage of the fact that the greater part of processor cycles on private and business computers remains idle. Rather than analyzing the data in a costly supercomputer, the SETI server divides the data into smaller units and sends these units to the several million computers made available by the volunteers who have registered to participate in this project. The SETI client carries out the calculations during the idle processor cycles of the participants' computers and then sends back the results. In the related literature, SETI@home is consistently referred to as a perfect example of a P2P application in general, and more specifically, a perfect example of grid computing (Oram, 2001; M. Miller, 2001). This evaluation, however, is not completely accurate, as the core of SETI@home is a classical client/server application, due to the fact that a central server coordinates the tasks of the nodes and sends them task packets. The peers process the tasks they have been assigned and return the results. In this system there is no communication between the individual nodes. SETI@home does have, however, P2P characteristics (Milojicic et al., 2002). The nodes form a virtual community and make resources available in the form of idle processor cycles. The peers are to a large extent autonomous, as they determine if and when the SETI@home software is allowed to conduct computing tasks (Anderson, 2001; Anderson, Cobb, Korpela, Lebofsky, & Werthimer, 2002). The shared accomplishment of these types of distributed computing tasks, however, is only possible if the analytic steps can be separated and divided into individual data packets.

The vision of grid computing described earlier, however, extends far beyond projects such as SETI@home. At an advanced stage of development, it should not only be possible for each network node to offer its own resources, but it should also be possible for it to take advantage of the resources available in the P2P network. The first implementations suitable for industrial use are already being announced by the big players in the market. It will most probably be quite some time before a generally available, open grid platform is created, due to the fact that suitable middleware architectures and APIs are still in the development phase (Baker et al., 2002). A currently influential initiative is the Globus Project (The Globus Alliance, 2004), which is working on a standardized middleware for grid application and has been greeted with wide acceptance throughout the grid community. The project is being supported by important market players such as IBM, Microsoft, Sun, HP, and NEC.

P2P Communities

The term *virtual community* was introduced by Licklider and Taylor (1968): "[I]n most fields they will consist of geographically separated members, sometimes grouped in small clusters and sometimes working individually. They will be communities not of common location but of common interest." Today, numerous variations and extensions of the original definition can be found in the related literature. Schoberth and Schrott (2001) have identified a minimum consensus among the various definitions according to which common interest, common norms and values, as well as a common interaction platform are the elements that constitute a virtual community. Emotional links, continuity, alternating relationships, and self-determination represent additional qualifying elements. Based on this, we define P2P communities as virtual communities that use P2P applications as communication and interaction platforms to support the communication, collaboration, and coordination of working groups or groups of persons. Virtually no research has been conducted on the extent to which, or if in fact, all of the elements cited are evident or fulfilled in concrete P2P communities. In contrast, it is relatively easy to identify common interests and common infrastructures in the context of P2P communities. Users of file sharing networks, for example, wish to exchange music and operate an interaction platform to do so. This platform is established by the networked entities and the protocols such as FastTrack or Gnutella. In the anything@home networks, users are linked by interests, for example, for the search for extraterrestrial life forms (SETI@home) or for a cure for AIDS (fightaids@home). The existence or the pronounced development and efficiency of common norms and values in P2P networks can scarcely be determined from the currently available research. In this respect, it can at least be assumed that the establishment of common norms and values will increase as the availability of sanctions and reputation mechanisms grows. The early Napster clients, for example, enable users to deny access to their own resources for selected persons who have no resources or only undesired resources to offer, so that free-riding persons can be barred from the community.

It remains open as to whether qualifying elements such as emotional attachment or a feeling of belonging or continual membership in a community exist—or for that matter, whether they are even relevant. On the other hand, the criterion for self-determination is directly fulfilled by the autonomy of peers described as a characteristic of P2P networks at the onset.

Initial steps toward studying P2P communities are currently being undertaken in the areas of the restructuring of value chains by P2P communities, P2P communities as a business model and trust, as well as free riding and accountability. These will be described briefly in the following paragraphs.

- Restructuring of value chains: Hummel and Lechner (2001) use the music industry as an example of the fact that P2P communities can achieve scales that result in changes to the configuration of the value chain, as well as having a tendency to transfer control over individual value creation steps from the central market players to consumers and, in some cases, to new intermediaries.
- Communities as business models: Viewing communities as business models originates from Hagel and Armstrong (Armstrong & Hagel, 1996; Hagel & Armstrong, 1997). These authors broke away from the idea of viewing communities as purely sociological phenomena and observed virtual communities as a business strategy for achieving economic goals. Their research interest in this respect focuses on the question regarding the

extent to which virtual communities can actually be used by individuals with commercial interests (Hummel & Lechner, 2002). It is particularly unclear which monetary and nonmonetary attractions motivate potential members to participate.

- Trust: There are limitations on virtual collaboration structures. This is of particular importance for P2P applications, as trust is an essential prerequisite for opening up an internal system to access by others. Limitations on trust also constitute limitations on collaboration in P2P networks.

  As a result, it is necessary to have designs that allow trust to be built up between communication partners. Reputation can prove to be a very important element for the creation of trust in P2P networks (Lethin, 2001). Reputation is aggregated information regarding the previous behavior of an actor. In P2P networks, it is possible to take advantage of a wide spectrum of central and decentralized reputation mechanisms that have already been discussed and, in part, implemented in the area of electronic commerce (Eggs, 2001; discussed with particular reference to P2P in Eggs et al. [2002]).

- Free riding and accountability: The success of the file sharing communities established by early Napster and other similar applications has a remarkable origin. Individual maximization of usage in P2P communities leads to—with regard to the community—collectively desirable results. This is due to the fact that when a file is downloaded, a replicate of the music file is added to the database of the file sharing community. These dynamics are threatened by free riders by denying access to the downloaded file or moving the file immediately after downloading so that the collective database doesn't increase. Free-riding peers use the resources available in the P2P network, but do not make any resources available (Adar & Hubermann, 2000). For most P2P applications in general, and for file sharing systems in particular, this can create a significant problem and prevent a network from developing its full potential. Free riding reduces the availability of information as well as the level of network performance (Golle, Leyton-Brown, & Mironov, 2001; Ramaswamy & Liu, 2003). A possible solution for overcoming this problem is accountability (Lui, Lang, & Kwok, 2002). It consists of the protocolling and assignment of used or provided resources and the implementation of negative (debits) or positive (credits in the form of money or user rights) incentives. In view of the absence of central control, however, difficult questions arise regarding the acceptance, enforceability, privacy of user data, and so forth, and as a result, the practicality of such measures.

Schoberth and Schrott (2001) have noted a scarcity of empirical research in the area of virtual communities (and consequently, in the area of P2P communities) as well as a lack of models for illustrating the development, interaction, and disintegration processes of virtual communities. There is a clear need for research regarding the development, motivation, stabilization, and control of P2P communities.

Conclusion

The introduction of a three-level model comprising P2P infrastructures, P2P applications, and P2P communities helps to differentiate the discussion regarding P2P networks. The management of significant resources such as information, files, bandwidth, storage, and processor cycles can benefit from the existence of P2P networks.

Initial experience with protocols, applications, and application areas is available and provides evidence of both the weaknesses and the potential of P2P-based resource management. The use of P2P networks for resource management promises advantages such as reduced cost of ownership, good scalability, and support for ad hoc networks.

- Reduced cost of ownership: The expenditure for acquiring and operating infrastructure and communication systems can be lowered by using existing infrastructures and reducing the administration and user costs. For example, P2P storage networks avoid the necessity for operating a central server for the purpose of storing the entire data volume for a network. By utilizing existing resources, the relative costs for each peer are reduced in respect of data storage or distributed computing. In addition, the operation of central or, as the case may be, additional storage servers or high performance computers, is no longer necessary to produce the same total performance within the network. In the context of collaboration, P2P groupware applications, such as Groove Virtual Office, frequently make the central administration of a server or the central distribution of rights when forming working groups obsolete.
- Scalability: Implementations have shown that P2P networks reduce the dependency on focal points and thereby reduce the potential for bottlenecks by spatially distributing information and creating replicates. This enhances their scalability. Hybrid file sharing networks, such as early Napster, have scalability advantages in comparison to client/server approaches. This results from the direct exchange of documents between peers without the

assistance of a server. With this method, early Napster was in a position to serve more than six million users simultaneously. New approaches such as OceanStore and PAST are aimed at providing their services for several billion users with a volume of files greater than 10^{14} (Milojicic et al., 2002, p. 13).

- Ad hoc networks: P2P networks are ideal for the ad hoc networking of peers because they tolerate intermittent connectivity. As a result, it is conceivable that the underlying philosophy of P2P networks will be increasingly utilized by approaches such as grid computing, mobile business, and ubiquitous computing—in particular when the task at hand is to establish communication between spontaneously networked peers or entities (PDAs, mobile telephones, computers, smart devices, or things in general) in the absence of coordinating, central authorities (Chtcherbina & Wieland, 2002).

These advantages are currently counteracted by some disadvantages. Security mechanisms, such as authentication and authorization as well as accountability, are easier to implement in networks with a central server. Equally, the availability of resources and services cannot always be guaranteed in networks with a small number of participants due to their intermittent connectivity. In file sharing networks, for example, in order to provide the desired level of availability, a correspondingly large number of replicates have to be created. This, however, results in increased use of storage and can have a negative effect on the resource advantages that could be otherwise be achieved.

At all three levels, two points become clear: the importance of P2P networks as information system architectures and the pressing need for research and development in this field. To the best knowledge of the authors, there has been no quantitative comparative study to date, which takes P2P networks into account when investigating the advantages of various information system architectures. However, it should be noted that the potential of P2P networks lies in the establishment of new application scenarios and a "simple" comparison of performance would probably be inadequate. According to this potential, the authors expect that knowledge management will especially benefit from P2P-based resource management. When thinking of knowledge as a further resource, we have to distinguish between implicit and explicit knowledge. Implicit knowledge is tied to individuals and therefore is shared face-to-face. Face-to-face communication can be supported very well by P2P knowledge management systems (Tiwana, 2003). P2P groupware, for example, facilitates self-organized and ad hoc networking of knowledge workers and therefore offers different communication channels. Explicit knowledge, however, is represented in documents. In order to enable access to this knowledge, a centralized knowledge management system has to create a repository of all available documents. This organizational effort can be omitted by using P2P knowledge management systems. Due to this, decentralized knowledge repositories enable access to up-to-date and as yet unrecognized knowledge better than centralized approaches can do (Susarla, Liu, & Whinston, 2003).

It remains to be seen which further application scenarios can be established and how effectively and how efficiently decentralized, in principle, self-controlled, P2P networks can fulfill the requirements for fair cost distribution, trustworthiness, and security (Schoder & Fischbach, 2003). In any case, alongside client/server structures, P2P networks offer a self-contained approach for the organization and coordination of resources. They have both advantages and disadvantages that will have to be evaluated in the relevant context.

References

Aberer, K., & Hauswirth, M. (2002). An overview on peer-to-peer information systems. Retrieved November 15, 2004, from http://lsirpeople.epfl.ch/hauswirth/papers/WDAS2002.pdf

Adar, E., & Hubermann, B.A. (2000). Free riding on Gnutella. *First Monday*, *5*(10). Retrieved November 15, 2004, from http://www.firstmonday.dk/issues/issue5_10/adar/index.html

Adya, A., Bolosky, W.J., Castro, M., Cermak, G., Chaiken, R., Douceur, J.R., et al. (2002). FARSITE: Federated, Available, and Reliable Storage for an Incompletely Trusted Environment. Retrieved November 15, 2004, from http://research.microsoft.com/sn/Farsite/OSDI2002.pdf

Anderson, D. (2001). SETI@home. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 67–76). Sebastopol, CA: O'Reilly.

Anderson, D.P., Cobb, J., Korpela, E., Lebofsky, M., & Werthimer, D. (2002). SETI@home: An experiment in public-resource computing. *Communications of the ACM, 45*(11), 56–61.

Armstrong, A., & Hagel, J. (1996). The real value of on-line communities. *Harvard Business Review, 74*, 134–141.

Baker, M., Buyya, R., & Laforenza, D. (2002). Grids and grid technologies for wide-area distributed computing. *International Journal on Software: Practice & Experience (SPE), 32*(15), 1437–1466.

Bal, H.E., Löhr, K.-P., & Reinefeld, A. (Eds.) (2002). *Proceedings of the Second IEEE/ACM International Symposium on Cluster Computing and the Grid*. Washington, DC.

Balakrishnan, H., Kaashoek M.F., Karger, D., Morris, R., & Stoica, I. (2003). Looking up data in P2P systems. *Communications of the ACM, 46*(2), 43–48.

Barkai, D. (2001). *Peer-to-peer computing. Technologies for sharing and collaboration on the net*. Hillsboro, OR: Intel Press.

Bellwood, T., Clément, L., & von Riegen, C. (2003). UDDI Version 3.0.1 - UDDI Spec Technical Committee Specification. Retrieved November 15, 2004, from http://uddi.org/pubs/uddi_v3.htm

Bolcer, G.A., Gorlick, M., Hitomi, P., Kammer, A.S., Morrow, B., Oreizy, P., et al. (2000). Peer-to-peer architectures and the Magi™ open-source infrastructure. Retrieved November 15, 2004, from http://www.endeavors.com/pdfs/ETI%20P2P%20white%20paper.pdf

Butler, R., Welch, V., Engert, D., Foster, I., Tuecke, S., Volmer, J., et al. (2000). A national-scale authentication infrastructure. *IEEE Computer, 33*(12), 60–66.

Castro, M. (2001). Practical Byzantine fault tolerance. Retrieved November 15, 2004, from http://www.lcs.mit.edu/publications/pubs/pdf/MIT-LCS-TR-817.pdf

Chtcherbina, E., & Wieland, T. (2002). In the beginning was the peer-to-peer. Retrieved November 15, 2004, from http://www.drwieland.de/beginning-p2p.pdf

Clarke, I. (2003). Freenet's Next Generation Routing Protocol. Retrieved November 15, 2004, from: http://freenet.sourceforge.net/index.php?page=ngrouting

Clarke, I., Miller, S.G., Hong, T.W., Sandberg, O., & Wiley, B. (2002). Protecting free expression online with Freenet. *IEEE Internet Computing, 6*(1), 40–49.

Cohen, B. (2003). Incentives Build Robustness in BitTorrent. Retrieved Nevember 15, 2004, from: http://www.bitconjurer.org/BitTorrent/bittorrentecon.pdf

Crowcroft, J., & Pratt, I. (2002). Peer to peer: Peering into the future. *Proceedings of the IFIP-TC6 Networks 2002 Conference,* 1–19.

Dabek, F., Brunskill, E., Kaashoek, M.F., Karger, D., Morris, R., Stoica, I., et al. (2001). Building peer-to-peer systems with Chord. A distributed lookup service. *Proceedings of the 8th Workshop on Hot Topics in Operating Systems*, 81–86.

Dabek, F., Kasshoek, M.F., Karger, D., Morris, R., & Stoica, I. (2001). Wide-area cooperative storage with CFS. *Proceedings of the 18th ACM Symposium on Operating Systems Principles*, 202–215.

Damker, H. (2002). Sicherheitsaspekte von P2P-Anwendungen in Unternehmen. In D. Schoder, K. Fischbach, & R. Teichmann (Eds.), *Peer-to-peer – Ökonomische, technologische und juristische Perspektiven* (pp. 209–228). Berlin: Springer.

Dasgupta, R. (2001). *Microsoft .NET as a development platform for peer-to-peer applications*. Intel Corporation.

Druschel, P., & Rowstron, A. (2001). Scalable, distributed object location and routing for large-scale peer-to-peer systems. *Proceedings of IFIP/ACM International Conference on Distributed Systems Platforms (Middleware), Heidelberg, Germany,* 329–350.

Druschel, P., Kaashoek, F., & Rowstron, A. (Eds.). *Peer-to-peer systems, First International Workshop, IPTPS 2002, Cambridge, MA, USA, March 7–8, 2002, Revised Papers.* Springer.

Edwards, J. (2002). *Peer-to-peer programming on Groove*. Indianapolis, IN: Addison-Wesley.

Eggs, H. (2001). *Vertrauen im Electronic Commerce. Herausforderungen und Lösungsansätze*. Wiesbaden, Germany: Deutscher Universitätsverlag.

Eggs, H., Sackmann, S., Eymann, T., & Müller, G. (2002). Vertrauengenerierende Institutionen für P2P-Netzwerke. In C. Weinhardt & C. Holtmann (Eds.), *E-Commerce – Netze, Märkte, Technologien* (pp. 43–54). Heidelberg, Germany: Physica.

Foster, I. (2002). The grid: A new infrastructure for 21st century science. *Physics Today, 55*(2). Retrieved November 15, 2004, from http://www.aip.org/pt/vol-55/iss-2/p42.html

Foster, I., & Kesselman, C. (Eds.) (2004). *The grid: Blueprint for a new computing infrastructure* (2nd edition). San Francisco: Morgan Kaufmann.

Foster, I., Kesselman, C., & Tuecke, S. (2002). The anatomy of the grid. In D. Schoder, K. Fischbach, & R. Teichmann (Eds.), *Peer-to-peer – Ökonomische, technologische und juristische Perspektiven* (pp. 119–152). Berlin: Springer.

Foster, I., Kesselman, C., Tsudic, G., & Tuecke, S. (1998). A security architecture for computational grids. *Proceedings of the 5th ACM Conference on Computer and Communication Security,* 83–92.

Global Grid Forum (GGF) (n.d). Retrieved November 15, 2004, from http://www.gridforum.org/

The Globus Alliance (2004). Retrieved November 15, 2004, from http://www.globus.org/

Golle, P., Leyton-Brown, K., & Mironov, I. (2001). Incentives for sharing in peer-to-peer networks. *Proceedings of the ACM Conference on Electronic Commerce, Tampa, FL*, 264–267.

Goldberg, A., & Yianilos, P. (1998). Towards an archival intermemory. In *Proccedings of the IEEE International Forum on Research and Technology Advances in Digital Libraries, Santa Barbara*, CA, 147–156.

Gong, L. (2001). JXTA: A network programming environment. *IEEE Internet Computing, 5*(3), 88–95.

Grid Security Infrastructure (2004). Retrieved November 15, 2004, from http://www-unix.globus.org/toolkit/docs/3.2/gsi/index.html

Groove networks (2003). Retrieved October 28, 2003, from http://groove.net/

Hagel, J., & Armstrong, A. (1997). *Net gain. Expanding markets through virtual communities*. Boston: Harvard Business School Press.

Hummel, J., & Lechner, U. (2001). The community model of content management, a case study of the music industry. *Journal of Media Management, 3*(1), 4–14.

Hummel, J., & Lechner, U. (2002). Business models and system architectures of virtual communities. From a sociological phenomenon to peer-to-peer architectures. *International Journal of Electronic Commerce, 6*(3), 41–53.

Hummel, T. (2002). Instant messaging – Nutzenpotenziale und Herausforderungen. In D. Schoder, K. Fischbach, & R. Teichmann (Eds.), *Peer-to-peer – Ökonomische, technologische und juristische Perspektiven* (pp. 59–70). Berlin: Springer.

Jabber (2004). Retrieved November 15, 2004, from http://www.jabber.org/about/overview.php

Kontiki (n.d.). Retrieved November 15, 2004, from http://kontiki.com/

Kubiatowicz, J., Bindel, D., Chen, Y., Czerwinski, S., Eaton, P., Geels D., et al. (2000). OceanStore: An architecture for global-scale persistent storage. *Proceedings of Architectural Support for Programming Languages and Operating Systems, New York*, 190–201.

Langley, A. (2001). Freenet. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 123–132). Sebastopol, CA: O'Reilly.

Lethin, R. (2001). Reputation. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 341–353). Sebastopol, CA: O'Reilly.

Leuf, B. (2002). *Peer-to-peer. Collaboration and sharing over the Internet*. Boston: Addison-Wesley.

Licklider, J.C.R., & Taylor, W. (1968). The computer as a communication device. In *Science and Technology* (pp. 21–40).

Loesgen, B. (n.d.). XML interoperability. Retrieved November 15, 2004, from http://www.topxml.com/conference/wrox/2000_vegas/Powerpoints/brianl_xml.pdf

Lui, S.M., Lang, K.R., & Kwok, S.H. (2002). Participation incentive mechanisms in peer-to-peer subscription systems. *Proceedings of the 35th Hawaii International Conference on System Sciences*, 302b.

Lv, Q., Cao, P., Cohen, E., Li, K., & Shenker, S. (2002). Search and replication in unstructured peer-to-peer networks. *Proceedings of the 16th ACM International Conference on Supercomputing*, 84–95.

Miller, J. (2001). Jabber – Conversational technologies. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 77–88). Sebastopol, CA: O'Reilly.

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

Milojicic, D.S., Kalogeraki, V., Lukose, R., Nagaraja, K., Pruyne, J., Richard, B., et al. (2002). Peer-to-peer computing. Retrieved November 15, 2004, from http://www.hpl.hp.com/techreports/2002/HPL-2002-57.pdf

Minar, N. (2001). Distributed systems topologies: Part 1. Retrieved November 15, 2004, from http://www.openp2p.com/pub/a/p2p/2001/12/14/topologies_one.html

Minar, N. (2002). Distributed systems topologies: Part 2. Retrieved November 15, 2004, from http://www.openp2p.com/pub/a/p2p/2002/01/08/p2p_topologies_pt2.html

Minar, N., & Hedlund, M. (2001). A network of peers – Peer-to-peer models through the history of the Internet. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 9–20). Sebastopol, CA: O'Reilly.

Moreton, T., Pratt, I., & Harris, T. (2002). Storage, mutability and naming in Pasta. Retrieved November 15, 2004, from http://www.cl.cam.ac.uk/users/tlh20/papers/mph-pasta.pdf

Napster protocol specification (2000). Retrieved November 15, 2004, from http://opennap.sourceforge.net/napster.txt

Nejdl, W., Wolpers, M., Siberski, W., Schmitz, C., Schlosser, M., Brunkhorst, I., & Löser, A. (2003). *Super-Peer-Based Routing and Clustering Strategies for RDF-Based Peer-to-Peer Networks*. Retrieved November 15, 2004, from: http://www.kbs.uni-hannover.de/Arbeiten/Publikationen/2002/www2003_superpeer.pdf

Next Page Inc. (n.d.). NXT 4 build and deployment white paper. Retrieved November 15, 2004, from http://www.nextpage.com/pdfs/collateral/wp/nxt4_build_wp090903lo.pdf

Oberholzer, F., & Strumpf, K.(2004). The Effect of File Sharing on Record Sales - An Empirical Analysis. Retrieved November 15, 2004, from: http://www.unc.edu/~cigar/papers/FileSharing_March2004.pdf

Oram, A. (Ed.) (2001). *Peer-to-peer: Harnessing the benefits of a disruptive technology*. Sebastopol, CA: O'Reilly.

P2P-Radio (2004). Retrieved November 15, 2004, from: http://p2p-radio.sourceforge.net/

Pandurangan, P.R.G., & Upfal, E. (2001). Building low-diameter p2p networks. *Proceedings of the 42nd annual IEEE Symposium on the Foundations of Computer Science*, 1–8.

Planet DESCENT (n.d.). Retrieved October 28, 2003, from http://www.planetdescent.com/

Project JXTA (n.d). Retrieved November 15, 2004, from http://www.jxta.org/

Ramaswamy, L., & Liu, L. (2003). Free riding: A new challenge to peer-to-peer file sharing systems. *Proceedings of the 36th Hawaii International Conference on System Sciences,* 220.

Ratnasamy, S., Francis, P., Handley, H., Karp, R., & Shenker, S. (2001). A scalable content-addressable network. *Proceedings of the ACM SIGCOMM'01 Conference,* 161–172.

Ripeanu, M. (2001). Peer-to-peer architecture case study: Gnutella network. *Proceedings of IEEE 1st International Conference on Peer-to-Peer Computing*.

Ripeanu, M., Foster, I., & Iamnitchi, A. (2002). Mapping the Gnutella network: Properties of large-scale peer-to-peer systems and implications for system design. *IEEE Internet Computing Journal*, *6*(1), 51–57.

Rowstron, A., & Druschel, P. (2001). Storage management and caching in PAST, a large-scale, persistent peer-to-peer storage utility. *Proceedings of ACM SOSP'01*, 188–201.

Schoberth, T., & Schrott, G. (2001). Virtual communities. *Wirtschaftsinformatik, 43*(5), 517–519.

Schoder, D., & Fischbach, K. (2002). Peer-to-peer. *Wirtschaftsinformatik, 44*(6), 587–589.

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

Schoder, D., Fischbach, K., & Teichmann, R. (Eds.) (2002). *Peer-to-peer – Ökonomische, technologische und juristische Perspektiven*. Berlin: Springer.

Schollmeier, R. (2002). A definition of peer-to-peer networking for the classification of peer-to-peer architectures and applications. *Proceedings of the First International Conference on Peer-to-Peer Computing*, 27–29.

SCVI.NET(n.d.). Retrieved November 15, 2004, from: http://www.scvi.net/.

Shirky, C., Truelove, K., Dornfest, R., Gonze, L., & Dougherty, D. (Eds.) (2001). *2001 P2P networking overview*. Sebastopol, CA: O'Reilly.

Smith, H., Clippinger, J., & Konsynski, B. (2003). Riding the wave: Discovering the value of P2P technologies. *Communications of the Association for Information Systems, 11*, 94–107.

Stump, M. (2002). Peer-to-peer tracking can save cash: Ellacoya. Retrieved November 15, 2004, from http://www.ellacoya.com/news/pdf/10_07_02_mcn.pdf

Susarla, A., Liu, D., & Whinston, A.B. (2003). Peer-to-peer enterprise knowledge management. In C.W. Holsapple (Ed.), *Handbook on knowledge management 2 – Knowledge directions* (pp. 129–139). Berlin: Springer.

Tiwana, A. (2003). Affinity to infinity in peer-to-peer knowledge platforms. *Communications of the ACM*, *46*(5), 77–80.

Udell, J., Asthagiri, N., & Tuvell, W. (2001). Security. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 354–380). Sebastopol, CA: O'Reilly.

W3C World Wide Web Consortium (2004). Retrieved November 15, 2004, from *http://www.w3.org/*

Web services activity (2004). Retrieved November 15, 2004, from http://www.w3.org/2002/ws/

Wiley, B. (2001). Interoperability through gateways. In A. Oram (Ed.), *Peer-to-peer: Harnessing the benefits of a disruptive technology* (pp. 381–392). Sebastopol, CA: O'Reilly.

Wojciechowski, R., & Weinhardt, C. (2002). Web services und Peer-to-peer-Netzwerke. In D. Schoder, K. Fischbach, & R. Teichmann (Eds.), *Peer-to-peer – Ökonomische, technologische und juristische Perspektiven* (pp. 99–118). Berlin: Springer.

Yang, B., & Garcia-Molina, H. (2001). *Comparing hybrid peer-to-peer systems*. Retrieved November 15, 2004, from http://www-db.stanford.edu/~byang/pubs/hybridp2p_med.pdf

Zhao, B., Huang, L., Stribling, J., Rhea, S., Joseph, A., & Kubiatowicz, J. (2004). Tapestry: A Resilient Global-scale Overlay for Service Deployment. In IEEE Journal on Selected Areas in Communications, Vol 22, No. 1.