Networking in the New Millennium: Business Applications for Peer-to-Peer and Grid Computing

Richard Taylor  
*University of Houston*

Nellai Subramaniam  
*University of Houston*

Follow this and additional works at: [http://aisel.aisnet.org/amcis2004](http://aisel.aisnet.org/amcis2004)
Networking in the New Millennium: Business Applications for Peer-to-Peer and Grid Computing

Richard G. Taylor  
University of Houston  
rgtaylor@uh.edu

Nellai Subramaniam  
University of Houston  
nsubramaniam@uh.edu

ABSTRACT
With the introduction of Napster in 1999, 18-year-old Shawn Fanning showed the world the power of peer-to-peer (P2P) computing. Even though Napster did not survive, P2P continued to live on and flourish. While the focus has been on the personal use of P2P computing, businesses have started to find ways to take advantage of the P2P model in their organizations. P2P proponents forecast that businesses can save billions of dollars by implementing P2P into their networking environments. An extension of P2P computing is grid computing. While grid computing was initially used for scientific and academic research, the past two years have seen a significant increase in the use of grid computing by commercial institutions. The main objective of this paper is to investigate how peer-to-peer and grid computing systems have been taken out of the scientific, academic, and personal usage environments and successfully applied in business environments.

Keywords: peer-to-peer, P2P, grid computing, distributed computing

INTRODUCTION
Shortly after personal computers were introduced into the business world it became apparent that connecting computers together would bring added benefits to an organization, including information sharing and resource sharing [Fitzgerald and Dennis, 2002]. Thus networking was born. By the early 1990s, more than 60 percent of all microcomputers in American businesses were networked [Fitzgerald and Dennis, 2002]. Communications technology continued to advance resulting in more sophisticated networks. Networking continued to evolve with the introduction of the Internet into the business environment, allowing worldwide interconnectivity. Entering the new millennium, businesses are finding new ways to take advantage of networking. Borrowing from sociology, social network theory tells us that the greater the number of ties in a network, the greater the overall density and power of the network (Turner 2003). Businesses are applying this theory in their environments, creating new types of powerful networks. This paper will look at two look at two them—P2P computing and grid computing.

“P2P is a revolution that will change computing as we know it.”
Andy Grove, chairman of Intel [Rutherford, 2000]

What’s the big deal with P2P? It’s just another way of doing things we have always done using the Internet and other networks—exchanging and finding data, and/or collaborating remotely. However, P2P does these things faster and more dynamically by making use of distributed, interconnected computers. The power of peer-to-peer computing, better known as P2P, was demonstrated in 1999 when an 18-year-old university student, Shawn Fanning, created Napster. Napster permitted users around the world to share music files. Napster was an instant success and in just 18 months had attracted almost 40 million users [Laudon and Laudon, 2002].

Napster opened up a universe of music sharing to millions of people. However, Napster was more than a free source of music. It was “an exploration of the unknown that widened cultural horizons” [Oram, 2001]. At the peak of its success, Napster boasted over 70 million registered users with up to 1.57 million users online simultaneously [Shirky, Turelove, Dornfest, and Gonze, 2001]. Napster, though, was a challenge to intellectual property laws and was shut down by court order; but P2P refused to die. As of June 2002, an estimated 19% of Americans over the age of 12 had downloaded music files.
from various P2P systems. P2P file sharing applications also accounted for five of the top 10 downloads for the download.com Web site in the last week of June 2002, together constituting 4.5 million downloads [Lee, 2003].

Downloading music is not the only practical use of the P2P architecture. The Forrester research organization predicts that by 2005, about 43% of broadband households will be taking advantage of P2P to share such things as digital photos, digital video, addresses, and calendars [Forrester, 2001]. Online consumers already have access to tools which can be used to create personal rich media, and those who have broadband access are already exchanging this personal rich media at least once per week [Forrester, 2001]. But many of these people are already getting tired of continually updating Web sites with this information, or receiving emails with large attachments. It would be much easier for them to allocate a secure area on their PCs to permit access to family and friends.

The initial success of Napster and other file sharing systems convincingly demonstrated that P2P computing can readily accommodate rapid growth and support millions of computers. Technology managers are finding ways to leverage P2P’s strengths, while avoiding or overcoming its management challenges. Joining the IT managers are a number of vendors, including IBM, Intel, Sun Microsystems, and Microsoft who have all embraced P2P computing as demonstrated by several new products and well-funded initiatives [Edwards, 2002].

P2P COMPUTING

P2P has actually been around for a long time, including the connections made with original computer modems [Fox, 2001]. Another early, but less sophisticated, version of a P2P network was the “sneaker net”. Users had personal computers on their desks, but they were not interconnected. To transfer information, data was copied to a disk and carried to another computer. Typically, this was done to find a computer with a printer attached [Cope, 2002]. The telephone network, discussion forums of Usernet, and the earliest form of the Internet can also be classified as P2P systems [Minar and Hedlund, 2001]. The earliest form of the Internet communicated from computer to computer, without a server in between to facilitate the connection. One of the breakthroughs of ARPANet was its concept of connecting computers as equal peers [Spangler, 2001].

The development of free-standing PCs in the early 1980’s, lead to the initial use of P2P networks in businesses.

The true nature of P2P however is still uncertain. Is it an IT architecture, a business model, a set of protocols, a design philosophy stressing decentralization, or merely a fad? The answer may be “all of the above.” P2P is defined as a “technology that enables two or more peers to collaborate spontaneously in a network of equals (peers) by using appropriate information and communication systems without the necessity for central coordination” [Schoder and Fischbach, 2003: 27]. P2P is a form of distributed processing linking computers through the Internet or private networks for the purpose of sharing processing tasks [Laudon and Laudon, 2002]. This allows users to share files via interconnected virtual private or public networks. P2P networks can handle huge numbers of users simultaneously and rarely crash because of their decentralization. The P2P architecture also makes them more scaleable and much less vulnerable to distributed DOS (denial of service) attacks [Krebs, 2001].

The main concept behind P2P is decentralization. Without a centralized server for connection, users are connected through other users. This allows users to communicate directly with one another. The PCs in this interconnected network are referred to as “peers”, because they offer their resources to other computers on the network [Fox, 2001]. Resources can include both disk space and processing power. The term “peer” indicates that each node is treated as an equal. In reality, however, this can be misleading. Computers have different CPUs, storage capacity, memory and network connectivity. Some computers are managed professionally while some are not and some computers reside at network hubs where others are at the edges or the network [Kubiatowicz, 2003]. However, when accessing information they are all clients, when providing information to other peers, they are servers. P2P generally assumes that each peer is acquainted with only a small number of other peers with which it can exchange information and services. Acquaintances tend to change constantly, there is no central control, and peers remain totally autonomous throughout their participation in a P2P network [Penserini, Liu, Mylopoulos, Panti, and Spalazzi, 2003].

In a pure P2P environment, every client must know where the other clients are, and be able to connect to them. Another variation of the P2P model, called hybrid P2P, introduces a server that can be used for managing the various peer devices or to store information, including replicated data, for disconnected peers. In hybrid P2P, the server only plays a supporting role, as opposed to the leading role it plays in client/server applications. The hybrid model resembles a ring network topology, where segments of the traditional application and directory, along with the message storage and files, reside on the hard drives or the networked PCs. Unlike client/server applications, though, hybrid P2P applications only use server resources.
when a segment of the ring is no longer available, such as when the network is experiencing heavy traffic, a user is no longer logged onto the network, or when a connection is severed [Edwards, 2001]. With the hybrid architecture, the central server can store content replications for disconnected peers, as well as other information to manage peer devices and user profiles. With larger systems, multiple servers may be used to store the directory information needed to find other systems and identify services of facilities available on them. Servers are also used to connect to systems for replication, content exchange, or management purposes [Edwards, 2001].

Hybrid P2P networks are appealing to businesses, because they still afford a level of control that is lost with pure P2P. Napster is an example of a hybrid P2P network, an architecture which proved to be its downfall. The Recording Industry Association of America was able to totally shut down Napster by a court order shutting down the primary server that was used to direct request for all of Napster’s millions of users. Other file sharing sites, such as KaZaa, have been able to stay in operation because they are pure P2P networks. There is no central server that could be shut down to stop the transfer of files. Each peer is connected directly to a group of peers. Legal action must now be, and has recently been, directed at individual users.

Whether pure or hybrid, P2P is divided into three basic divisions: instant messaging, affinity communities, and distributed computing [Rutherford, 2000].

**Instant Messaging**

Instant messaging (IM) allows interconnected users to exchange text messages and files synchronously. It is a tool that is already being utilized by many organizations, and serves as a good example of a P2P application that has already been adopted by the business world as a central communications tool. It has been predicted that the number of business IM users will grow to over 180 million users by 2004 [Spangler, 2001]. IM allows more synchronous communication than email, allowing remote workers to collaborating on a project via a richer media offering instantaneous chat [Rutherford, 2000].

Internet clothing retailer Lands’ End has capitalized on this tool by allowing on-line shoppers to utilize “Lands End Live” and “Shop with a Friend” [Ives and Piccoli, 2003]. “Lands’ End Live” allows customers to chat online directly with a customer service representative. “Shop with a Friend” allows two shoppers to browse the site together, communicate with each other and add items to a single shopping cart. Title Data, Inc., located in Houston, Texas, recently released an IM product called Sonork, which enables customers to link directly to needed files, thereby reducing support-related calls by as much as $10,000 per month [Kontzer, 2003].

**Affinity Communities**

Affinity communities are groups, such as Napster and KaZaa, formed specifically for file sharing. These P2P environments hold potential for improved collaboration both within and outside the organization. They also facilitate software interaction, which allows programs to send both data inputs and outputs from application to application. Affinity communities “emphasize knowledge integration over acquisition and learning” [Tiwana, 2003: 80]. Vendors who utilize these communities include Groove Networks, Pointera, Kalepa Networks Roku, uRoam, Hilgraeve, FlyCode, GoneSilent, Hotline Communications, Centrata, and the infamous Napster. Groove Networks was founded in 1997 with the mission of “allowing business teams to collaborate on a broad range of activities within secure, shared virtual spaces, in real time or offline” [Rutherford, 2000]. Their systems combine the file-sharing capabilities of Napster with instant messaging and other features to allow groups to share information and communicate in a secure environment [Edwards, 2002]. In 2001, Microsoft invested $51 million in Groove Networks. Since then they have been working together to unleash the power of collaboration in Windows XP and Microsoft SharePoint Technology. This type of P2P advances the premise that “new value comes from sharing information and building on it” [Oram, 2001].

Collaboration is one of the applications identified by Intel as being conducive to P2P networking. P2P collaboration will enable organizations to tap into intellectual property that is locked in their organizational boundaries such as information located in separate offices and on servers with restricted access. This can help organizations unleash the knowledge locked away on employees’ PCs [Cortese, 2001]. P2P connections allow employees to communicate, collaborate and create. This collaboration now involves “connecting people to people, people to machines, and enterprises to enterprises”, increasing the overall knowledge management within organizations [Krill, 2001]. P2P even allows organizations to leverage individual expertise at the periphery of their networks, extending their organizational reach [Tiwana, 2003].

**Distributed Computing**
Distributed computing, which involves pooling the processing power of many computers, is another effective business use of P2P computing. The basic premise is that there are PCs around the organization that are underutilized. Researchers at the University of Wisconsin estimate that organizations use less than 25% of their computing and storage capacities that they have already paid for [Fox, 2001]. Distributed computing allows organizations to harness this unused processing power in desktop PCs. Companies like aerospace giant Boeing, Intel, and oil company Amerada Hess report that they have been able to reduce the need to purchase high-end computer systems, including mainframes, by utilizing P2P networks that tap into processing power already available on their desktop PCs. P2P networks can also lessen bandwidth requirements, which is a key benefit for organizations with networks jammed to capacity [McDougall, 2000]. Distributed computing can have a direct impact on an organization’s bottom lines by defraying the cost of expensive hardware, reducing administrative overhead of maintaining and managing long-running applications, and by increasing business opportunities by accelerating the decision-making process [Lee, 2001].

Distributed computing can also increase an organization’s storage capacity, therefore reducing the need to purchase high-priced server disk drives. Desktop computers are being purchased with larger and larger hard drives. Today’s desktops typically arrive with hard drives of 80 gigabytes or more. For every 100 desktops that can translate into six untapped terabytes of storage. Companies can use P2P computing to increase efficiencies in their distributed storage networks to utilize this unused space. Software that will determine the optimal distribution patterns and the paths for storage loads is already available. This software can also leave a digital blueprint on every client PC to identify where to find the remaining data. Distributed computing can also help smaller businesses utilize their limited bandwidth more efficiently. By sharing processing tasks at the LAN level, distributing storage loads, and limiting downloads from Web servers distributed computing allows traffic to move away from the Internet to the internal corporate LAN, where bandwidth is much more plentiful and easily managed [Edwards, 2001].

GRID COMPUTING

Grid computing is an extension of the P2P distributed computing concept and in essence, grids are P2P systems [Talia and Trunfio, 2003]. Though the terms “peer-to-peer computing” and “grid computing” have been used interchangeably [Chetty and Buyya, 2001], a distinction can be made between them: P2P computing is concerned with sharing low end systems such as PCs connected to the Internet for amassing computing power (e.g., the SETI@home project1), or sharing content as done in Napster and Gnutella. On the other hand grid computing is concerned with aggregating the distributed clusters of computers using well defined protocols and standards. The objective of grid computing is to harness the resources of unused CPU cycles of computers to perform computational tasks at a faster rate. The idea of using the unused CPU cycles originated in the early 1970s with the linking of computers by networks. In 1973, scientists at the XEROX Palo Alto Research center developed a worm program that replicated itself in the memory of each of the 100 Ethernet connected computers. The worms used the idle resources to perform shared computations for rendering realistic computer graphics [Chetty and Buyya, 2001].

The simplest grid may consist of a few machines having the same hardware architecture and operating system and connected via a local area network. In other cases however, a grid may include machines from different departments and consist of dissimilar hardware architecture and operating systems. An intragrid usually refers to machines from different departments within a single organization forming the grid, while an intergrid refers to a grid that crosses organization boundaries [Berstis, 2002]. There are three different types of grids—computational, scavenging, and data. In a computational grid, resources, usually high performance servers, are set aside exclusively for computing power. A scavenging grid uses the unused CPU cycles and resources from desktop machines throughout an organization. In a data grid, the main focus is to provide access to data through distributed storage [Jacob, 2003].

Every grid system must have some kind of management software. This software is used to keep track of resources available to the grid and also to keep track of the different users on the grid. Other jobs of this software include measurement of the following: capacities of the resources, utilization rate of the resources, traffic congestion and bottlenecks, overall usage patterns of resources. Such software also generates statistics, provides for recovery from grid failures, schedules jobs and finds alternate ways to perform tasks such as coordinating parallel execution of jobs.

---

1 SETI@home is a scientific experiment that uses Internet-connected computers in the Search for Extraterrestrial Intelligence (SETI). Complete information about SETI@home can be found at http://setiathome.ssl.berkeley.edu/
Grid computing has made its primary impact on research in the scientific and academic arena, such as the recent research endeavor at Virginia Tech University where researchers have created one of the world’s most powerful supercomputer by linking together 1100 Apple Macintosh computers. This supercomputer is able to compute at 10.28 TFlop/s (trillion operations per second), making it the third fastest supercomputer in the world [Top500, 2003]. The performance of this system is obviously impressive, but the real eye opener is its cost. While typical supercomputers cost in the $100 million to $200 million range, the Virginia Tech system only cost about $5 million [Markoff, 2003]. Other research oriented organizations have also created grid computing systems. In fact, seven of the top ten supercomputers in the world are grid computing systems, as are 208 of the top 500 supercomputers [Top500, 2003].

While grid computing was initially used for scientific and academic research, the past two years have seen a significant increase in the use of grid computing by commercial institutions. With costs significantly lower than those of supercomputer, businesses are realizing they can get supercomputer processing speed for a fraction of the cost of a supercomputer. According to Gartner, by 2006, five percent of organizations that typically use supercomputing cycles will instead turn to cheaper grid computing, a number they expect to see increase to 25% by 2011[Gartner, 2001]. As a result, vendors such as IBM and Oracle are positioning grid computing as an important cog in their computing strategies. Grid computing is driving a new evolution in industries such as bio-medical, financial, oil exploration, motion picture, aerospace and many others. Businesses jumping on the grid computing bandwagon include Shell, BankOne, J.P. Morgan Chase Investment Bank, Charles Schwab, and Novartis A.G.

Shell uses grid-enabled infrastructure for applications that involve interpretation of seismic data [Case Study-Shell, 2003]. The grid is expected to cut seismic data processing time while improving the output quality. Further, because of the open standards used, the grid is expected to provide easy integration of existing software.

“Grid Computing is important to Shell because it offers the potential to create a truly unlimited resource, with a uniform interface to a variety of services. This is a significant opportunity for Shell to engage its independent companies in closer cooperation”.

J.N. Buur, Principal Research Physicist, Shell International Exploration and Production

BankOne is using grid computing to boost performance of analytics for its Chicago based trading floor. Instead of using a supercomputer to process its massive risk analytics, BankOne uses an array of small processors with the computing responsibilities distributed across them. Bank One expects to make significant cost savings on hardware and also utilize the unused processing cycles of the distributed resources [Schmerken, 2003].

JP Morgan Chase Investment Bank (JPMCIB) is offering its business units online access to raw computing power via an enterprise technology infrastructure, which pools the financial firm’s flexible processing resources for computational intensive applications. The project code-named “Compute Backbone” uses grid computing technology to improve the process service levels. The business units will be charged for their processing (peak and off-peak pricing) and thereby reducing costs [Davidson, 2002].

For others in the financial industry, grid computing enables the migration to cheaper computing solutions. Charles Schwab converted an existing wealth-management application and grid enabled it using the Globus Toolkit with IBM eServer running RedHat Linux [Schmerken, 2003]. By running Linux, a free open-source operating system, Charles Schwab expects to save significant capital expenditure on hardware, such as expensive UNIX servers. The project has helped Charles Schwab reduce the processing time of certain transactions from four minutes to fifteen seconds [Case Study-Charles Schwab, 2003].

Novartis A.G., the Switzerland based drug and pharmaceutical giant, uses grid computing technology to speed up its drug research in a cost-effective way. The grid links about 2700 PCs delivering five TFlops/s of computing power [Case Study-Novartis, 2003]. Manuel Peitsch, head of Novartis A.G.’s department of Informatics and Knowledge Management explains,

“If you look at the desktop PCs in a typical corporation, probably 90 percent of computing cycles are unused. Just by capturing unused cycles on the PCs we have already got, we have created a 5 teraflops supercomputer. We have avoided the

---

1 IBM eServer represents a class of IBM servers with on demand business capabilities. These servers are scalable, follow open standards and have features to manage server workload.
expenses of buying an HPC system, building another computer center, and taking on the people to support it. We invested roughly $400,000 in grid software licensing and figure we have saved at least $2M based on the 2700 seats we have currently. We expect to realize more savings of this nature in the future as our grid expands.”

Though grid computing has opened up a number of opportunities for many organizations to better manage their resources, it is not a panacea for all computational problems nor appropriate for every organization. The configuration of a grid can influence the benefits that accrue to an organization. Most current business applications are not written for parallel processing. Parallelizing these applications can involve a major investment. Standards for ensuring security in grids, monitoring user access and charging for user access are still evolving. Hence it is imperative for organizations to understand what grid computing can and cannot do as well as the associated risks.

Even through there are still issues involving the integration of grid computing into the business world, almost all major IT providers are involved in promoting grid computing in one form or another, and for good reason. According to a study by The Insight Research Corporation [Insight, 2003], total worldwide grid spending is expected to grow from $250 million in 2003 to approximately $4.9 billion in 2008, with financial applications and professional business services accounting for the highest grid expenditures. With more and more companies embracing grid technology, it is slowly moving from research labs to commercially viable solutions. Acknowledging this scenario, Larry Ellison, CEO of Oracle, who has been championing his company’s grid enabled database Oracle 10g states that [Krill, 2003; Songini, 2003].

“It’s not going to happen overnight, but there will be an inexorable move [to grid computing]. The economics are compelling; the reliability is compelling.”

CONCLUSION

The Gartner Group predicts that P2P will "radically change business models" [Rutherford, 2000]. P2P proponents predict that businesses can save billions of dollars by using distributed computing that takes advantage of unused/underutilized bandwidth and resources. Instant messaging and affinity communities can open up data and intellectual property that are otherwise hidden in multiple departmental offices and servers [Rutherford, 2000]. Many also see P2P computing as a solution that will reduce network bottlenecks, enhance collaboration within internal and external workgroups, and unleash computing power from underutilized processors throughout organizations [Edwards, 2001]. However, with the tarnished image of P2P from legal actions taken against Napster and other companies, many IT executives’ feelings about P2P still range from “mild skepticism to pure paranoia” [Scannell, 2001]. This paranoia, along with IT executives reliance on the centralized server to control and secure their organizations’ mission critical data to carry out important functions such as backups, and to host their e-commerce applications has led to their reluctance to deploy P2P technologies.

P2P and grid computing have already proven themselves to be useful business architectures. However, to become standard business architectures, many of the concerns surrounding P2P and grid computing will have to be addressed. P2P should continue to grow within the general public and grid computing will continue to play a significant role in academic and scientific research. However, expect to see more business uses emerge in the near future. It is also likely that the design philosophy underlying P2P will gain importance in the development of mobile and ubiquitous computing, especially when the goal is to establish communication links between mobile network peers, such as PDAs, laptops, and mobile telephones [Schoder and Fischbach, 2003].

“P2P is probably the only architecture in the world that can allow billions of these devices to be interconnected. Today’s client-server architecture can not handle such a large universe of devices. We are not even beginning to use all the power of the machines and people at the edge of our networks, and the intelligence and creativity at the edges will provide some of the greatest value as we move forward” [Gillmore, 2001].

P2P and grid computing provides an architecture which allows businesses to tap this wealth of knowledge and power. In fact, both Microsoft Chairman Bill Gates and Intel Corp. Chairman Andrew Grove both have said they believe P2P is going to be very important [Disabatino, 2000]. Yet, today P2P remains an architecture still waiting for a killer application [Agree, 2003]. Grid computing may just be that killer application. Only time will tell.
REFERENCES

http://www.computerworld.com/networkingtopics/networking/story/0,10801,69883,00.html