Survey of Peer to Peer technologies

Milić Bratislav, Krunić Jelena, and Milutinović Veljko

***Abstract*—For several years, P2P is one of technologies that brought lots of attention to it. Many people have been using some of P2P tools, those tools have created small fortunes to their creators, and some of tools and services were legally banned. Yet, only small number of people can claim that understand basics of P2P technology. Idea of this paper is to introduce P2P and its applications, without details of protocols, or advanced techniques**

***Index Terms*—distributed computing, file sharing, p2p, peer**

I. Introduction

As noted in the abstract, goal of this paper is to give you a short introduction to P2P networks, systems and applications. If you are interested in more detailed explanations of P2P, you should read some of the referenced papers, or some of excellent books that were published on this subject.

As with most of new technologies, there is no formal definition of P2P (although, the idea of P2P is not a new one – first telephones and even Arpanet can be considered as P2P). Instead of trying to put all that P2P is in one long and confusing sentence, we will give a list of its features:

1) All peers in P2P network are the same. This is quite different compared to "classical" client-server systems, where each peer (client or server) has its strict responsibilities, and applications that run on them are completely different. Peer in P2P network works *both* as a client and a server.
2) Data and computation is decentralized. All information and resources are placed in peers. This can be considered both appealing and problematic. Appealing sides are *high fault tolerance* (since no peer in the network is more important when compared to other peers, if some of the peers are not functional, that does not influence the network as a whole), and potentially *good scalability*. Scalability is only potentially good side, since it heavily depends on protocol of the network. Bad protocol waists resources and limits efficiency of peers; this can reduce scalability. Bad side is that is very hard *to guarantee QoS* in P2P networks.
3) Search for information in P2P networks is more relevant if compared to *static* searches (such as Google or Yahoo). In case of static search, before some information can be found it first has to be collected by crawler. And between two visits of crawler to certain site, months can pass. In the meantime, information can be put and removed from sites, so search results can be inaccurate. P2P networks give possibility to search for data in *real time*. However, there is again a drawback: depending on protocol, some data may not be found even if some peer in the network has it. Also, there is no guarantee for how long the query will take. So this is again problem of QoS in P2P networks.
4) Peers and their connections are volatile. Algorithms for any kind of improvement (QoS, security, etc.) that are using some topology assumption cannot be used in P2P networks. The only topology assumption that is valid in strict P2P network is volatility.

Interesting property of P2P networks is their ability and need to work outside of the DNS system because DNS is made for half-static environments, where new nodes are placed and removed relatively rarely and typical P2P network gains and looses hundreds of users each second.

II. P2P applications

All mentioned features of P2P make it suitable for different classes of applications. Here, we will present only some of the most known.

Typical applications that are used now include file-sharing systems (Napster, Gnutella, Morpheus, Freenet, Kazaa...) that are used for file sharing between peers of network, distributed calculations (FightAIDS@home SETI@home, Folding@home) where calculations are done on peers that volunteer for that and communication (ICQ, AOL Instant Messenger, some implementations of Internet radio broadcasts).

One of the first P2P applications that was adopted in Internet community was Napster. Napster was used mostly for sharing of .MP3 files, although it offered some other possibilities such as chat rooms or "buddy lists". Strictly speaking, Napster was not completely P2P application. All searches were done on single server (more precisely, on single place – as the number of Napster users was growing, they needed more servers to provide search results). Each user had to register on Napster server and to declare which files is sharing. When some registered user places query, the server gives him respond as a list of peers that have desired file. Then, connection between peers is established and file is transferred. That was the P2P part of the Napster service. Since Napster's main purpose was sharing of .mp3 music, recording companies sued him and stopped its service.

One of many Napster's "descendants" is Gnutella (some fo

others are Freenet, Morpheus, Kazaa,...). Similar to Napster, it was widely used for sharing of .mp3 music. Instead of central server (that can be sued), the Gnutella protocol used distributed search mechanism, which uses peers to search the network. This decentralization removes possibility for any kind of censorship; increases fault tolerance (bottleneck for Napster network was central query server – if it falls, everything falls). But generally, decentralization increases traffic in the network especially for queries.

Interesting aspect of Gnutella network is that since there is no central server, new user has to know address of at least one Gnutella peer in order to connect to the network. Off course that is no problem, since there are web sites that offer lists of reliable peers.

One of large problems that were noticed in file sharing networks are *free riders*. Free riders are peers that do not share files (do not offer files for others) and serve only as clients in network. They create lots of problems in network searching. Since they do not offer no resources (files), there is *absolutely no chance* that search will end in their node. So they are propagating search queries through the network and significantly increase network traffic.

Distributed computing is another interesting aspect of P2P networks. They are especially convenient for calculations that have coarse-grained parallelism. That means that large calculations can be easily divided into completely independent parts, which are calculated on separate nodes. Examples of the most popular P2P calculations are SETI@home that is used in search for extraterrestrial intelligence in the space, FightAIDS@home that is used for AIDS fighting and Folding@home that is used for protein folding analysis. All peers in the network voluntary give part of their computational resources to mentioned organizations. This kind of computation is called P2P because all the computation is done on peers although there is no direct communication between peers.

III. Future

All mentioned systems and networks use their individual protocols. Even some of networks that are used for similar purposes use different and incompatible protocols. Some of protocol implementations do not exist on many platforms.

Sun's JXTA platform tries to solve those problems. It offers open protocol specification (in contrast to Napster and Gnutella protocols that had to be reverse engineered), platform and implementation independence. It is still in research phase, but promises a lot. One of its strong points is that it has Java implementation that can allow it to be used on variety of platforms, from cell phones to large enterprise servers.

Still a lot of work has to be done before JXTA establish itself as widely accepted standard.

IV. Conclusion

At the moment, P2P is used mostly by home and individual users. It still has to find its way to enterprise users. To become accepted in those new environments, P2P standards and protocols has to improve QoS, security, scalability. Some of research has improved (and in some cases solved) some of those issues, but more work has to be done.

References

[1] K. Aberer, M. Hauswirth, "Peer to peer information systems: concepts and models, state of the art, and future systems"

[2] S. McCarthy, B. Hore, I. Issenin, S. Tauro, S. Han, "Survey on P2P File Sharing System"

[3] "Project JXTA: An Open, Innovative Collaboration", Sun Microsystems, www.jxta.org