

P2P Networks vs Multicasting for Content Distribution

Omer Boyaci
Rutgers University
Department of Computer Science
boyaci@cs.rutgers.edu

Abstract

Multicast seemed the only possible technique for one/many-to-many content distribution over many years. However we never had a world-wide content distribution network based on multicasting excluding some small and preconfigured networks such as mbone. Fortunately, today we have lots of content distribution networks such as gnutella, BitTorrent, eDonkey and Napster. All of these networks are built on peer-to-peer networks. We are going to discuss why peer-to-peer networks are better than multicasting for content distribution applications.

1. Introduction

Over the past twenty years Internet mostly carried unicast traffic and the most used architecture on it, was the client/server model. The main reasons of this situation are:

- Almost all of the content is created by companies, libraries, institutions and universities. And the connected end-users are just accessing these information.
- The access speed of the end-users are very low due to expensiveness of broadband links and most of them have non-static IP addresses so that they can not run servers to distribute content to others.
- The end-users did not have enough equipment, processing power and bandwidth to distribute content so they are clients to servers which have enough resources to do.
- Although content distributors want to use their bandwidth/processing power efficiently, there is no way to multicast traffic to masses due to problems in multicasting architecture and implementation.

Besides these technical difficulties the value of the Internet is low on those days compared to today's Internet because according to the Metcalfe's law the value of a network increases exponentially with the network size.

Over the years both the number of users of Internet and the processing power/bandwidth of computers are increased dramatically, so the Internet users want to share content among each other and also the producers of content want to distribute them efficiently to this huge user group. Multicast and P2P networks are two possible solutions for content delivery. P2P networks are more suitable for this environment than multicasting because of the following reasons:

- In order to use multicasting all the underlying Internet structure has to support it, however currently there is no world-wide support for multicasting. But because P2P communication occurs at application layer the interesting users can interact with each other by just installing a software.
- It is very hard to achieve reliable delivery of content with multicasting, however it is very easy with P2P networks.

In the following sections, we will give brief information about both solutions namely multicast and P2P networks; and then we will expand the reasons why P2P better for content distributions.

2. Multicast

Multicast is a proposed solution for efficiently distributing the same packet to more than one host. The packet will not replicated until it is necessary.

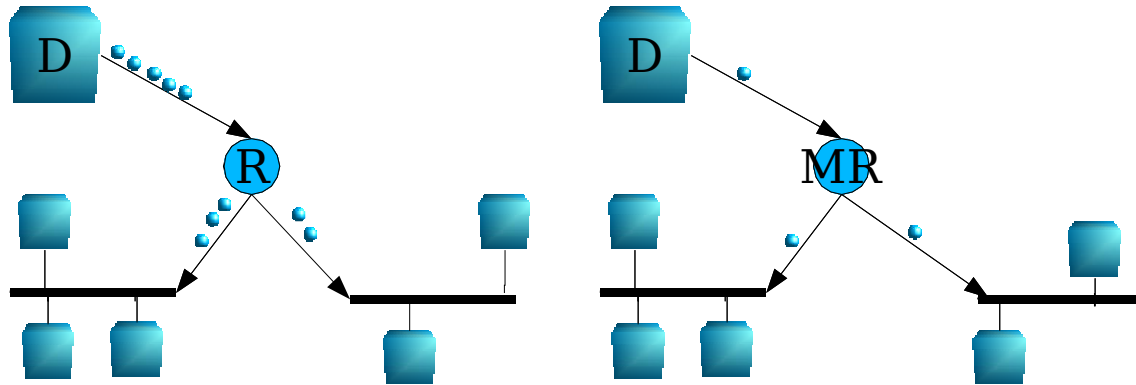


Figure 1.a Unicast

Figure 1.b Multicast

We can illustrate the usefulness and efficiency of multicast with a small example. In Figure 1.a, a content distributor sends 5 copy of the same packet to its five interesting users via unicast. In Figure 1.b, the producer sends only one copy of packet via multicast and then the packet is replicated when it is necessary. So the content distributor uses less bandwidth and also less processing power with multicasting.

However in order to use multicasting the routers on the way have to understand and behave accordingly to multicast packets. And also we need some protocols to manage joining and leavings and these protocols have to be used by all participating users. Today almost all LANs support multicasting and also some of the autonomous systems support multicasting internally, so in these areas applications can benefit from multicasting. For example, in NYSE and NASDAQ multicasting is used internally to handle the huge amount of traffic, and also lots of the multi-player games such as Quake 3 and Unreal Tournament benefits from multicasting in LANs.

3. P2P Networks

P2P Networks have appeared almost ten years ago and at that time they were used mostly for file sharing especially copyrighted materials such as music and movie files. The motivation behind P2P networks are very strong which is utilizing unused resources. Today most of us have computers with several gigahertz of speed and several gigabytes of hard drives and several kilobits of bandwidth so why not to use all of these resources instead of wasting them by not using.

In P2P networks each host behaves both as a client and server, each computer serves its files to the others and download files from others. The first generation of P2P systems are semi-centralized such as Napster. The central servers just indexes file locations and names and provides search functionality to the peers, the actual files are kept at the peers not in the servers [Figure 2.a]. In terms of functionality of the network there is nothing wrong with these semi-centralized P2P networks, however because users are sharing copyrighted materials these servers became an attacking point for media companies and copyright enforcement institutions such as RIAA.

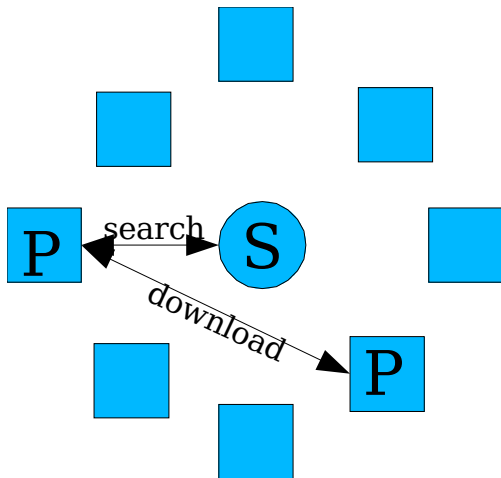


Figure 2.a Semi-centralized

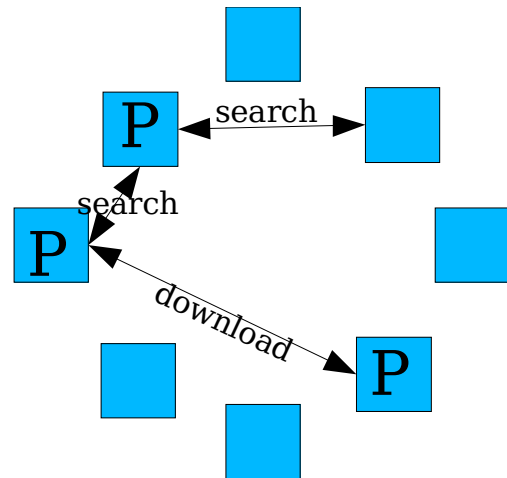


Figure 2.b Decentralized

So, the next generation of peer-to-peer networks are appeared such that there is no central server in the P2P network. Gnutella and Fasttrack are two examples of these decentralized P2P networks. In these systems, commonly users have a list of peers and they are forwarding their search queries to them [Figure 2.b]. If these peers have the requested file they can send it to the requesting peer. If they do not have the requested file, they forward the request to their list of pairs. These next level of peers can communicate with the initial peer directly if they have this file, because most of the time IP address of the requester is inside the search query. This searching process can take more time compared to the centralized approach. However due to its decentralized nature it is very hard to suspend the system completely, because if there is two peers who knows each other the network is exist.

4. The reasons why P2P is better than multicasting for Content Distribution Applications

4.1. Multicasting requires a common infrastructure

Multicasting requires a common infrastructure in terms of both hardware and software. First of all, the routers must support multicasting and they must be configured by system administrators such that multicasting is enabled. Secondly, all of the Autonomous Systems have to agree on a common multicasting protocol.

In todays Internet the Tier 1 ISPs such as Sprint and MCI experiencing huge amount of traffic on their backbones, so almost all of their routers are configured or built for high throughput. Supporting multicasting on these routers will decrease their throughput so most of these backbone routers are not supporting multicasting. Even one of the new routers did not check IP checksum in order to achieve high speed(50Gb/s) although it is necessary by the IP Procotol [1].

So the situation for multicasting is not better than 20 years ago. However, the situation for P2P networks is far better than 10 years ago. The main reason behind these development is P2P networks are open, such that anyone can design and build a P2P network and starts using it immediately. Because everyone can participate to the development of P2P networks new ideas and techniques can appear easily and even competition can occur among them so any of these networks tries to be the best. Actually today we have different P2P networks specialized on different problems, for example BitTorrent is designed to allow fast propagation of a single big file.

4.2. Achieving Reliable Delivery with Multicast is very hard

Achieving reliable distribution of content is very hard with multicasting. It is true that not all types of content requires reliability, however most of the content such as music or movie files, applications and documents requires reliability.

The main reasons behind the hardness of reliable multicasting are:

- The bandwidths of the participants can vary dramatically, so the distributor has to adjust its speed to some level which will permit most of the users to receive it.
- There has to be a way to retransmit missing packets of each listener, but it is hard to achieve this.

There are some proposed solutions to achieve reliability in multicast. One example is NACK based approaches and another example is circulating the same data more than one times. A list of reliable multicast protocols can be found at [2]. A pure NACK based approach is impractical with huge number of receivers. And also circulating the same data more than on times can be inadequate for low-bandwidth users.

In P2P networks reliability is almost default, because all the transmissions between peers are via unicast. So a reliable transmission protocol such as TCP can be used at no cost.

5. Conclusion

Multicasting and peer-to-peer networking are two different content distribution techniques. Multicasting has a long history, it is started in 1980s however today we have almost no system which uses multicasting for world-wide content distribution.

P2P networks borned in mid 1990s and they became very popular and useful. Today we are using them to distribute content to the masses. For example, the Fedora Linux is distributed via BitTorrent so that everybody gets the ISOs as soon as they released [4]. And 3D Gamers which host more than 200 Gigabytes of 3D game demos, screen shots and movies is also using BitTorrent to distribute these files[5]. Microsoft Research is also looking to benefit from peer-to-peer networks for streaming media content distribution and Web flash crowd alleviation with CoopNet Project[6]. Nokia is also researching peer-to-peer technology for content distribution [3]. So in the near future with the advances in peer-to-peer content delivery we may video conference with our friends via our mobile phone, although its low bandwidth and processing power. The tools needed to develop P2P applications are also released, for example jxta.org has java and C libraries for peer-to-peer application development [7]. Many companies such as Nokia and Siemens and many universities such as Rutgers University, The Applied Software Systems Laboratory and Stanford University, California are using and contributing to JXTA [8][9].

Due to its open and free nature P2P technologies seems to improve and we are going to see more P2P applications for content delivery in the future. Multicasting can be used for the same purpose in LANs and on some special networks such as mbone, however for general porpose content distribution applications peer-to-peer networks are the best choice.

6. References

- [1] J. Turner and N. Yamanaka, "Architectural Choices in Large Scale ATM Switches". Washington University Technical Report. 1997
- [2] http://www-net.cs.umass.edu/sigcomm_mcast/talk1.html
- [3] Peer-to-peer technology for content distribution, <http://www.nokia.com/nokia/0,,54018,00.html>
- [4] Fedora Project, <http://fedora.redhat.com/>
- [5] 3D Gamers, <http://www.3dgamers.com/>
- [6] CoopNet Project at Microsoft Research, <http://research.microsoft.com/~padmanab/projects/coopnet/>
- [7] jxta.org, www.jxta.org
- [8] University Spotlight Archive, <http://www.jxta.org/universities/universityarchive.html>
- [9] Company Spotlight Archive, <http://www.jxta.org/project/www/companies/companyarchive.html>
- [10] Wikipedia, <http://en.wikipedia.org>