

Content Distribution: P2P or Multicast

Zhiyuan Zhang

Abstract

Content distribution is the original application of both multicast and peer-to-peer (P2P) technologies. Since they are two fundamentally different approaches, it is a controversy to discuss who plays better. In this position paper, I will argue that although multicast solution conserves some total network bandwidth consuming, P2P technique is better on content distribution applications in terms of feasibility, flexibility, performance, and robustness. In addition, several different scenario cases are also investigated, which demonstrates that P2P is better from views of different network participants, the content distributor, the receiver, and the Internet Service Provider (ISP). Finally, we also discuss the legal consideration and commercial model issues of P2P.

1. Introduction

Because the popularity of content distributions, it has a broad prospect to develop technologies to these applications. Taking File Sharing for example, assuming there are thousands of users want to download a considerable file from a server, how to deliver the file efficiently? In traditional method, the sender should copy the file thousands of times, and then deliver the copy to every demander respectively. It is inefficient because of duplicate information distribution. Multicast is a technique implemented in

network layer, which is mainly proposed to save bandwidth because the original sender can just send one copy of file, and depends on routers to duplicate only necessary copies of contents and deliver them to receivers. On the other hand, rather than touch the underneath network layer, the P2P approach for content distribution is implemented in application level. The basic idea of P2P is to make receivers closer to the original sender get the data first, and after that, they became new senders, the later coming receivers then download the data or data chunks from one or more such senders respectively and simultaneously, no matter if the sender is the original one or any new ones, depending on the connection rate between them. At this point, this "late comer" becomes a sender as well, thus making the number of senders equals to the number of receivers.

Although both multicast and P2P technologies can improve the performance of transmission in theory, we believe that P2P is better than multicast in content distribution applications. In the following sections, we will compare both technologies in terms of implementation feasibility, application flexibility, performance and robustness.

In section 2, we begin to introduce backgrounds of the two totally different technique approaches from model to general process. In section 3, we will

discuss the advantages to apply P2P rather than multicast. Then, we will describe and refute some counterclaims that multicast is better than P2P. To explicit our claims, we also discuss 3 scenarios in section 5 from views of senders, receivers, and ISPs (Internet Services Providers) respectively. After those discussions, we point out that even though P2P applies better in content distribution, we also need consider several other issues like legality and commercial model to guarantee its prevalence.

2. Background

2.1. Multicast Technology

Multicast is a bandwidth-conserving technology specifically designed to reduce traffic by simultaneously delivering a single stream of information to potentially thousands of corporate recipients or homes. [1] The sender sends a single datagram to the multicast address, and the intermediary routers take care of making copies and sending them to all receivers that have registered their interest in data from that sender (see Figure 1). [5]

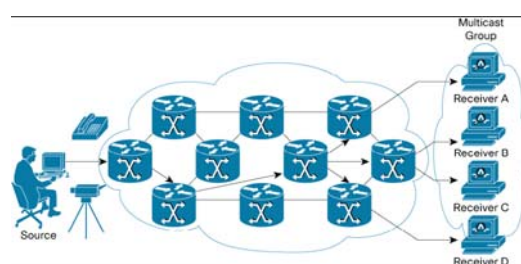


Figure 1

There are some Key words to describe multicast technology.

- Group: In multicast technology, receivers are joined into groups to

require same data. The sender will only send one copy to this group, and router will replicate copies and send them to every grouper.

- Spanning Tree: It is the spanning (distribution) tree to determine which routers should replicate and sent copies to groupers. If groupers join or leave the group, the spanning tree needs to be recomputed.
- IGMP: It is a kind of protocol support multicast technology, be used between routers and end-users.
- PIM: It is a kind of protocol support multicast technology, be used among routers.

The general process of multicast communication occurs as follow. [2] Firstly, receivers send IGMP message to routers to require join into the group. Then the routers will use the multicast protocols to compute the spanning tree, and add the receiver into the group. The receiver can get data from the resource. If a receiver wants to leave the group, it only needs to not reply the multicast routers' periodic queries. If routers find receivers leave out of the group, the spanning tree should be recomputed again.

2.2. peer-to-peer technology

Different from the traditional centralized client/server model, P2P network is organized in a decentralized manner, where all computers (peers) in such network are both clients and servers (Figure 2). Peers are dynamically consisted with the P2P network in ad-hoc manner. Anyone can

join or leave the network freely as well as query/get interested content from peers. Meanwhile, other peers can also get their interested contents from the current receiver.

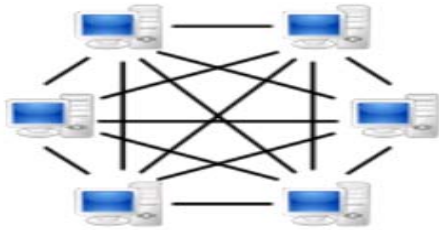


Figure 2

In a P2P content distribution network, the receivers will automatically find the "best" distributor or distributors to get the contents, in terms of the transmission latency and bandwidth between itself and distributors. Therefore, the receiver can get the content in almost the possible fastest rate. Once the receiver gets the content, it becomes another distributor, thus make other peers (receivers) be able to get the content data from it as well, by the same way as mentioned above. Another issue in P2P content distribution network is how a peer can know where to get its interested data. Normally, there are mainly three approaches: the first technique is making all peers sharing a centralized global directory, which is used to track the position of each content object. Peers directly query the directory to get the information and then get data from relative peers found in the directory; another way to do this is that the peers just flood their queries into the network, in which there is no centralized mechanism; the third approach is the "middle" one, where there are some "super nodes" due to their better connectivity, bandwidth or stronger

CPU, and every other nodes try to be within 1-hop within super nodes, and then query with their "super delegates".

3. My claims

As section 2 mentioned, we find both multicast and P2P can improve performance of network in different aspects. However, multicast is much more difficult to implement because it needs to change the current simple, unreliable, unicast internet model. In this section, we will compare both methods in feasibility, application range, performance and robustness.

3.1. Implementation Feasibility

Multicast technology needs routers support both unicast and multicast transmissions, apply IGMP between edge routers and end users, PIM among routers, which most applied routers can not achieve. Moreover, if we want to extend multicast to WAN, not only few but all routers should support these requirement which is a huge amount of investment. Who are the potential investors? It is clearly that end-users do not care whether he/she receives files from multicast transmission or unicast transmission, so they will not intend to pay for it. It is only Internet Service Provider (ISP) will invest in multicast in order to reduce bandwidth conserving. However, is it a good deal? In order to support multicast content distribution, ISPs should replace all of their routers into a multicast style. And even it is not enough if only a single ISP deciding to replace its devices, every ISP in the world should make a agreement, which makes the Internet multicast almost a mission impossible.

Even though every router is eligible to multicast, the cost of implementation, maintenance and overall management far exceeds bandwidth savings. [2] On the other hand, P2P is implemented on the application layer, which requires nothing to do with the well developed TCP/UDP/IP protocols at all. End users only have to download the P2P software, and install it onto the operating system of their own computers, and then they can start to enjoy the benefits. The ISP may be even unaware of the whole process. Therefore it is much easier than multicast, which will touch and require modifications of the network layer and replacement of existing routers, to implement on current existing network infrastructure.

3.2. Application Flexibility

Besides it is the P2P solution more feasible on the Internet for content distribution, other than to multicast, P2P is also better in terms of its flexibility of supporting various applications. Due to inherent technique limits of multicast, all receivers in such kind of network have to be synchronized to receive multicast data from the sender, at a particular point of time; in other words, the receivers cannot "randomly" get part of data chunks they want, but wait for the sender to "broadcast" the whole. Although this kind of characteristic is OK for some "one to many" content distribution applications, such as live audio or video broadcasting, online video conference, distance learning, it is lacking or inefficient on supporting other "many to many", "many to one", and even "one to one" applications. However, P2P can support them well.

Taking file sharing for consideration, in a P2P file sharing system, a user can download the file any time as he/she like, not from the far far away original distributor, but just from his/her neighbor peer, who is even in exactly the same intranet domain of the user. So the user can get access to the file immediately and download it with little latency. In addition, if the file is chunked, even the user can get it much faster by downloading those chunks from multiple peers (many to one), simultaneously. Obviously, multicast cannot support such kind of feature. It even cannot support file sharing well. Some other content distribution applications only P2P can support or support better include VoIP, software distribution, online streaming, video on demand (VOD), personal data synchronization and many more broader usages.

3.3. Performance

In this part, we will argue that the P2P solution overwhelms the multicast on content distribution not only on its feasible implementation, flexibility to support various of applications, but also its performance is better than multicast. In a P2P network, peers are always trying to get the data from their closes peers, from which they can download data at the highest rate. Moreover they can download chunks of contents from several peers simultaneously. It is highly possible that their neighbors are located in the same intranet domain or local network, then it is reasonable to get a much higher download rates rather than directly to retrieve the contents from the original distributor, who may evenly locate at the opposite

side of the world. While if applied multicast solution, the transmitting performance depends on the connectivity between the sender and the end receiver, according to the latency and bandwidth between them. Obviously, it will not as good as directly get data from the computer next to you. The basic reason is that in multicast solution, only the original sender contributes the upload bandwidth, while in P2P network all the peers do the contribution. Admittedly, the multicast solution will save some network bandwidth in one aspect, however, it also cannot be neglected that in order to support multicast, all the routers have to distinguish whether a packet is for unicast or multicast, thus increasing the routing overhead. In addition, to maintain the local and global multicast routing tree on the Internet is another considerable overhead. All in all, even if the multicast technology really decreases the network bandwidth usage, with little overhead, it is obvious that the sender or receivers care little about it.

3.4. Robustness

Robustness is also an important advantage of P2P over multicast. In model of multicast, it depends on the resource (sender) to distribute contents to groupers. From the spanning tree of multicast, we can conclude that if the root (sender), crash, no other nodes can take charge of the sender, so the whole mechanism will collapse. However, the distributed nature of peer-to-peer networks increases robustness in case of failures. In P2P system, receivers communicate with their peers, so if one peer crash, it just

means the peer leaves the peer group. [4] It is also feasible to find other close peers to join into peer group to keep sources abundant.

4. Counterclaims

Although we consider P2P is better than multicast in content distribution application, some people have opposite views. In this section, we will suggest opposite views firstly and then refute them.

4.1. Multicast has been applied in LANs

Some counterparts declare that since multicast has been well developed in several LANs, it will definitely apply successfully in WAN. At first, we should admit that some applications of multicast perform well in LAN, such as the Seismic Research Symposium [2] [6]. However, we should realize that be successful in little range does not mean to do the same or even better in WAN. First of all, we need to consider the multicast backbone. If routers of backbones can support the multicast, the data complexity they need to consider will be much more increased. Routers need to discern from multicast to unicast, apply particular protocols. Moreover, it should consider the computing of spanning tree. When a receiver joins into the group, it needs to compute the spanning tree which may also affect performance.

4.2. Multicast can release the bandwidth

It is the multicast rather than P2P can use bandwidth more efficiently. Since the resource just sends out one copy of file, it will release the volume of traffic

in the internet whereas P2P will increase it because there are much more chunks are transmitted in the internet at the same time. However, we should consider much more factors rather than only bandwidth-conserving to evaluate a technology. Besides bandwidth, we also need to consider the rate of transmit, the feasibility of implementation and so on. As section 3 mentioned, P2P has better implementation feasibility, broader application range, be faster in most cases, better in robustness. Therefore P2P is better on overall performance than multicast.

5. Scenarios Study

After comparisons from deployment, application range, performance and robustness, in this section, we will discuss three detailed scenarios to strengthen the conclusion.

5.1. From Distributors' View: Beta Version Game Distribution

If a company wants to distribute a Beta Version of its new game, which technology it will prefer? Firstly, the company wants as more as possible players to download the game, because the more people download the game, the more information will feedback to help improve the game, and means more people be attract to this game and probably be potential consumers. Secondly, they hope every player feel satisfied through downloading, because if they feel convenient, they will be more likely to purchase the game. Thirdly, the capacity of its own server is too limited, and the connection rates and quality are also not good enough to support downloading from tens of

thousands of players all around the world. If they prefer multicast, this technology will release their congestion on some degree, however, the ISP of potential consumers may not support multicast transmission, so the consumers do not adapt the technology at all. Even though consumers use multicast to download the game, they will not feel any difference through the downloading process. However if it choose P2P, people can enjoy this technology by just install a software. Moreover, potential consumers may feel convenient by faster downloading through closer peers rather than just one resource. The company's network will also be not so congested because peers share the downloading pressure. Therefore, after comparing the investment cost, degree of users' satisfaction, and performance, the company will prefer P2P technology rather than multicast.

5.2. From Receivers' View : File Sharing

Assuming a CS student Peter wants to get an Open Source Software, he can download it by multicast or P2P. If he prefer multicast, he need to join a group (automatically) and wait for the resource distributes the software. If it exists high latency between Peter and resource, as a long distance, Peter need to wait a long time to download the software. However, if he applies P2P, it is easy to find closer peers who can also store the file and can sent it to him at a higher speed. Moreover, the file can be divided into several chunks, so Peter can receive different chunks from different peers simultaneously, which will also speed up the download rate.

Additionally, Peter just needs to upload some sources as a pay back. Therefore, in a receiver's view, Peter will choose P2P rather than multicast.

5.3. From ISPs' View: Costs and Benefit

If an ISP wants to invest a novel technology to improve its performance and competitiveness, it will be more likely to choose P2P rather than multicast. Firstly, if it prefer multicast, it need to purchase a lot of money to buy eligible routers. However, after the investment, it also requires other ISPs' cooperation, so it is uncertain of the performance of bandwidth-conserving. Moreover, the ISP needs cost-recovery and may be price pushing. However, consumers care little about this technology (they don't care their receive file transmit through multicast or unicast) and may not purchase more money on it, thus price pushing may destroy ISPs' competitiveness among other ISPs. Thirdly, it is also complex to compute the volume of traffic of consumers since the ISP needs to distinguish the true or false receivers. On the other hand, compared to multicast, ISP should consider the congestion problem caused by P2P, but it is still more advisable to choose P2P rather than multicast.

6. Extension

Even though P2P is better than multicast in terms of content distribution application, it does not mean P2P will develop successfully in the real commercial world. We also need to consider other issues such as copyright, ISP's support.

Due to P2P's decentralized characteristic, it's more difficult to control the contents' been distributed on the Internet by P2P, compared to the client/server model or multicast, thus making illegal P2P traffic the largest bandwidth consumer on the internet[3]. Some P2P system uses a centralized global directory to track contents (files), such as Napster. This kind of system is vulnerable to copyright's law, if most of contents distributed on it are illegal. Napster is forced to close in 2002. However, in decentralized approach, it's hard to tell what kinds of contents are distributed in the system, who are senders, and who are receivers. Since the P2P network is organized in an ad-hoc manner, it's hard to track if the contents distributed in such a system are legal or illegal, if they are illegal, who are distributing them. Therefore, P2P technology rendered a great challenge to the digital contents copyrights protection. However, I think it's not the fault of the technology, but it's all about how people use the technology.

Another important issue is that the current P2P model does not fit ISPs' economics models well. With the increased P2P usage for content (even legal content) distribution, the traffic volume within a single access ISP or among different ISPs will increase dramatically, even saturating ISP's bandwidth capacity. However, in current P2P model and popular internet accessing plan, e.g. monthly plan, ISP's revenue cannot increase accordingly, and even it will cost ISPs more on those huge volume of incoming/ongoing P2P traffic from/to other ISPs or the ISPs.

Actually, content distributors using P2P to publish their copyright content will indeed save a large amount of money on distribution cost, e.g. they do not need to rent a lot of servers around the world as their download sites to make end users achieve an acceptable download rate. However, the cost and tension is actually pushed to the ISPs' side. If this issue is not addressed, the P2P traffic caused by such kind of content distribution has a high risk to be banned by individual ISP. Some possible solutions are the content distributor pays some money to ISP, or the ISP builds its own P2P content distribution solutions, and charges some money on those who want to publish their product.

7. Conclusion

In this paper, we have compared the both advantages and disadvantages of multicast and P2P technology in the application of content distribution, and get a conclusion that P2P suits better than multicast in content distribution applications. At beginning, we briefly introduce both technologies from models, processes and applications. Then we declared our argument that P2P is more suitable in content distribution than multicast, and pay much attention on attesting our view. Lately we introduce counterparts' view, and refute it. After that we also provide three scenarios to explain and support our view. In the last part, we extend that even though P2P is better than multicast, we also need consider other factors such as legal consideration and commercial model to make sure it development.

References

- [1] "IP Multicast Technical Overview". http://www.cisco.com/en/US/prod/collateral/iosswrel/ps6537/ps6552/prod_white_paper0900aecd804d5fe6.pdf
- [2] "Multicast as a Viable Commercial Enterprise". <http://www.cs.rutgers.edu/~rmartin/teaching/fall08/cs552/position.html>
- [3] " New Generation Content Delivery Network". <http://www.cachelogic.com>
- [4] " Advantages of peer-to-peer networks". <http://www.solyrich.com/P2P-pros-cons.asp>
- [5] " Introduction of Multicast ". <http://www.firewall.cx/multicast-intro.php>
- [6] Agarwal, D. and Lawrence, O. "Using Multicast in the Global Communication Infrastructure for Group Communications ". 21st Seismic Research Symposium.