Federation Based Solution for Peer-to-Peer Network Management

Jilong Wang and Jing Zhang

Network Research Center, Tsinghua University, Beijing 100084, China
wjl@cernet.edu.cn, zhj_021@tom.com

Abstract. Recently, Peer-to-Peer (P2P) technology has become one of the hottest topics in the research area of Internet. With a variety of P2P applications, especially those applications sharing large-size file resources among a large scale of user community, P2P brought us an even more exciting Internet life. However, P2P also made lots of troubles to network managers because of consuming too much network bandwidth sometimes. Lacking of effective management solution, some ISPs plan to block all P2P services on their network boundaries. In this paper, we propose a federation-based solution for peer-to-peer network management. By setting up a P2P federation, ISPs and P2P service providers can work together for P2P network management. From P2P federation service, ISPs can get P2P nodes information of their own network to make some control to those that disturb normal network services. At the same time, service providers of P2P may get routing information of specific ISPs to optimize the routing of P2P network itself by joining in federation. Under such scenario, it will save much cost for ISPs to detect and control P2P traffic and give up the idea of killing P2P services. Save and help to the development of P2P are the most important target of this solution.

Keywords: Peer to Peer, Federation, Management.

1 Introduction

Currently, Peer-to-Peer (P2P) technology is widely applied in Internet application systems, especially those applications sharing large-size video files among a large scale of user community.

P2P changes the traditional Client/Server communication model into a point-to-point model. P2P traffic is no longer converged at a few computers that play as server hosts, but is distributed among each node in the network, which makes the distribution of traffic more reasonable and helps avoid congestion of network.

Although P2P technology has lots of advantages, it brings troubles to network operation sometimes. On account of being frequently used to transfer large size files such as audio and video clips, P2P applications may consume much network bandwidth. Furthermore, P2P application is always apt to use the network bandwidth as much as it possible. In a certain study [2], it has estimated that the total P2P volume goes about 80% of the total downloading traffic in the Internet, and the ratio has been constantly on the rise, nearly twice a year.

Y. Shi et al. (Eds.): ICCS 2007, Part IV, LNCS 4490, pp. 765–772, 2007.

Long-time congestion brings the network management great difficulty. Let's take account of the case of BitTorrent (BT) [1], one P2P application for file sharing. Providing there are numerous clients using BT to download and upload files at the same time, this application will take up lots of network bandwidth and influence other services in the Internet [3]. On the other hand, lack of proper management of P2P applications also causes a series of social problems, relating to copyright, privacy, security and so on. Owing to these referring matters, Internet Service Providers (ISPs) urgently desire to find an effective way to monitor and control the P2P applications. But it is not an easy subject. Current P2P applications tend to disguise their protocol specifications, operate on the random port numbers, and even use the custom port number, such as port 80 of HTTP, intentionally. These characteristics of new-style P2P applications make it even more difficult to monitor and control P2P traffic.

In this paper, we propose a federation-based solution for peer-to-peer network management. Not by measurement, but by exchanging information among ISPs and P2P providers, which is named PSPs under federation model. By setting up a P2P federation, ISPs and PSPs can work together for P2P network management. From P2P federation service, ISPs can get P2P nodes information of their own network to make some control to those that disturb normal network services. Meanwhile, PSPs may get routing information of specific ISPs to optimize the routing of P2P network itself by joining in federation. Under such circumstance, it will save much cost for ISPs to detect and control P2P traffic and give up the idea of killing P2P. Save and help to the development of P2P are the most important target of this solution.

The rest of this paper is structured as follows:Section2 analyzes the previous work in the areas of P2P traffic measurement and management, and then indicates some problems. Section 3 presents the framework of the solution. Section 4 describes some details, including considerations of implementation. Finally, we conclude the paper in Section 5.

2 Previous Works

In order to monitor and manage the P2P applications, we need to distinguish P2P traffic from other network load.

At present, there are basically two ways to measure P2P traffic. One is active measurement, and the other is passive measurement which contains two methods: payload analysis and nonpayload analysis.

2.1 The Active Measurement

In this method we collect P2P nodes information by setting some P2P crawlers. P2P crawlers are some special nodes deployed in the P2P network, connected with other live hosts through TCP. Communicating with known hosts, crawler establishes a peer-list and adds hosts that were freshly found into the list. In this way, crawler host acquires the information of other peers in the P2P network, from which we could know the nodes' (peers') actual distribution and operating status.

However, there are two serious limitations using this methodology. Firstly, it does not adapt to the management of large scale P2P network since this method depends on

the number of connections that crawlers keep. And one significant character of P2P applications is just the enormous user community. Secondly, this method is concerned with specific protocols. Therefore, it is not apt to P2P systems having several types of protocols, especially when some protocols are not open.

2.2 The Passive Measurement—Payload Analysis

Payload analysis of P2P traffic is based on identifying characteristic bit strings in packet payload. It analyses the potential character (usually using the IP packet head and the sixteen bits of the data packets) in the packet payload to identify the type of P2P traffic. E.g. the character sign of BitTorrent protocol is "0x13Bit" [1]. We can distinguish BT traffic by analyzing whether the packet payload contains "0x13Bit". In the same way, eDonkey packets can be identified by "0xe319010000" [7].

The disadvantage of this method is the inefficient application layer identification. Due to the demand of dealing all the data packets, the speed of analysis is slow, which means that it is not adapt to the real-time analysis. In addition, this method is not able to deal with the unknown protocol either.

2.3 The Passive Measurement—NonPayload Analysis

Using nonpayload analysis methodology, we can identify the P2P traffic without inspecting the user payload. The analysis needs only some packet header information such as the connection patterns of source and destination IPs and ports.

From some related references, at present there are six out of nine popular P2P applications using both TCP and UDP as the transport layer's protocols. These P2P applications usually use TCP to transfer actual data and use UDP to control traffic [2]. Due to this characteristic, we can identify P2P hosts by looking for pairs of source-destination hosts that use both transport protocols (TCP and UDP). Whereas the usage of both TCP and UDP is not only for the P2P protocols, it is also used for some other application protocols such as DNS to transport. Therefore, it is hard to distinguish the P2P application from other applications which have the same characteristic, especially in the case of some P2P applications intentionally using the custom port number.

For the descriptions presented above, we strongly commit to the notion that although the management of P2P applications is very important, at the present there is still not an adaptive and effective methodology to analyze so many kinds of P2P applications.

3 Federation Based Solution for P2P Network Management

Because of the diversity and variability of P2P protocols, it is quite hard for ISPs to monitor and manage P2P applications just by active and passive measurement. In order to deal with the matters in the network caused by P2P applications, ISPs attempt several means (e.g. ban the port numbers P2P protocols usually using), and do their best to forbid the P2P applications. At the same time, PSPs create a series of improved protocols and renew-edition client software to confront the forbiddance from ISP, which makes the management of P2P applications more difficult. Thus, both sides of ISP and P2P are plunging into a hostile competition.

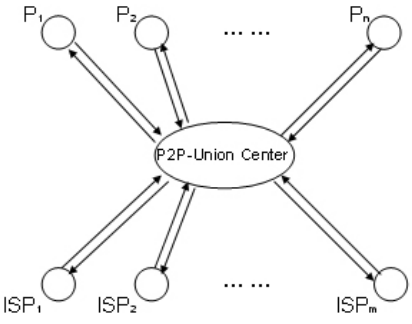

Fig. 1. Architecture design of P2P federation. P1…Pn represent various of P2P service providers, ISP1…ISPn represent Internet Service Providers.

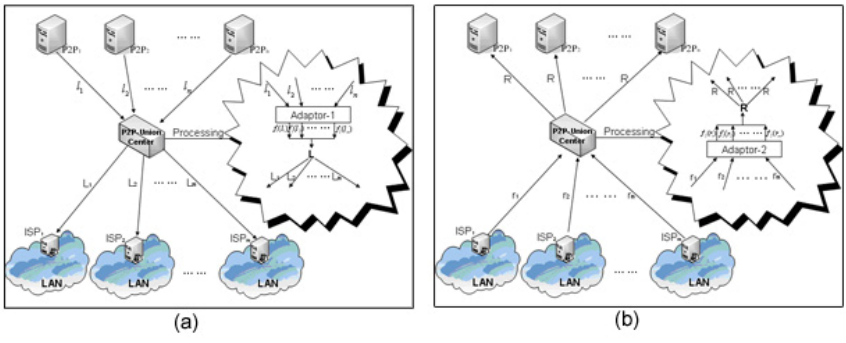

Fig. 2. (a) ISPs learn P2P network information from federation. (b) P2Ps learn ISP's routing information from federation.

To solve above problems, we propose to establish a service called P2P-Federation that helps to exchange P2P client and communication information between ISPs and PSPs. Using this methodology, we can easily deal with the current problems existing in P2P management such as multiple protocols, disguising port numbers and so on.

The basic idea of P2P federation service is cooperation. Instead of the traditional method that made a contrastive relationship between ISPs and PSPs, P2P federation solution aims to found collaboration between ISPs and PSPs.

Commonly, a P2P federation system is composed of three parts. They are ISP units (a series of ISPs), P2P units (a series of P2P providers) and a P2P-Federation Center. Figure 1 shows the architecture design of the P2P federation. Figure 2(a) and (b) show the scenarios of exchanging nodes and routing information between ISPs and PSPs.

The P2P-Federation Center works as a backbone to connect ISPs and P2Ps together, and provides services of information exchange. The federation center gets nodes lists from P2P service providers and provides nodes information to ISPs. On the other hand, the federation center may also gains the IP routing information from

ISPs and provides such information to P2P service providers. The communication among ISP, federation center and PSP should be based on the reliable and secure protocols.

4 The Working Procedure and the Experiment

This section describes how the P2P Federation works in detail. The basic Working Procedure of P2P Federation Service is shown as Figure 3.

Step 0: Register. The P2P-Federation Center provides the register service. The acceptable users contain all kinds of ISPs and P2P service providers. For users whose identity is ISP, the personal information they should submit to register includes: userName, userPwd (that is user's password), userIP (that is the IP address used by ISP's server), IPSection (that is the IP sections that is managed by this ISP), routingFormat (that is the format of the routing table this ISP uses). In the same way, as to users whose identity is PSP (P2P service provider), they should also submit their personal information including userName, userPwd, userIP (that is the IP address that the server providing P2P service uses), p2pType (that is the type of the P2P application, such as Bittorrent, eMule, and so on), p2pFormat (that is the format of the P2P's routing information). Then, the P2P-Federation Center provides each user a unique userID to identify. The P2P-Federation Center maintains a database for preserving the users' identity information.

Step 1: P2P servers that have registered in the P2P-Federation Center will transfer Peers lists (l_i) they maintained to the Center in real-time. The content of Peers list (l_i) is the routing information of all the P2P clients which use the service provided by the corresponding P2P server. The formats of these Peers lists ($l_1, l_2 ....., l_n$) from different P2P servers may also be different. It lies on whether the p2pIDs of the different P2P application are the same.

Step 2: The P2P-Federation Center formats the Peers lists (l_i, i=1,…,n) by an adaptor (signed as Adaptor-1). In this way, all the various Peers lists have been formatted in the same mode, that is (IP, p2pID). We can sign the new lists as $f_1(l_1), f_1(l_2), ..., f_1(l_n)$. The function f_1 () represented the process by Adaptor-1.

Step 3: The P2P-Federation Center then adds up all the new lists ($f_1(l_1), f_1(l_2), ..., f_1(l_n)$) and acquires an entire Peers list which is called L. $L = f_1(l_1) + f_1(l_2) + ... + f_1(l_n)$.

Step 4: According to the IPSections given by ISP users who have registered in the Center, the P2P-Union Center decomposes the L into several child lists which is noted as $L_1, L_2 ....., L_m$.

Step 5: The P2P-Federation Center transfers the child lists L_j (j=1,…,m) to corresponding ISP users in real-time.

As shown in Figure 4, we can easily and directly comprehend the process of how to transfer the P2P routing information to ISPs.

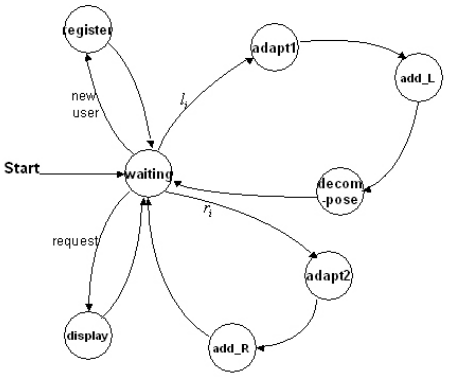

Fig. 3. The Working Procedure of P2P Federation Service

Step6: ISP users that have registered in the P2P-Federation Center will transfer routing tables (r_i) maintained of their own to the Center in time. The content of routing table (r_i) is the routing information of the IPSection managed by the corresponding ISP. The formats of these routing tables ($r_1, r_2 ..... , r_m$) from different ISPs may be different too.

Step 7: The P2P-Federation Center formats the routing tables (r_i, i=1,…,m) by an adaptor (signed as Adaptor-2). In this way, all the various routing tables have been formatted in the same mode, that is (sourceIP, nextRoute, target). We can sign the new lists as $f_2(r_1), f_2(r_2), ..., f_2(r_m)$. The function f_2 () represented the process by Adaptor-2.

Step 8: The P2P-Federation Center then adds up all the new lists ($f_2(r_1), f_2(r_2), ..., f_2(r_m)$) and acquires an entire Peers list which is called R. $R = f_2(r_1) + f_2(r_2) + ... + f_2(r_m)$.

Step 9: The P2P-Federation Center transfer the child lists R to each PSPs in real-time.

We can see the process of how the ISPs transfer the IP routing information to PSPs in the Figure 5.

Step 10: ISP uses the Peers lists to manage P2P traffic. Taking into account of ISP_i , it gets the Peers list L_i from the P2P-Federation Center, in which there is all the P2P clients' routing information (including IP and p2pType). Then ISP_i can confine the hosts whose IP is in the list L_i to a low speed network through infusing specific routing information into the network. Notes that this is just one of possible ways to use the Peers lists.

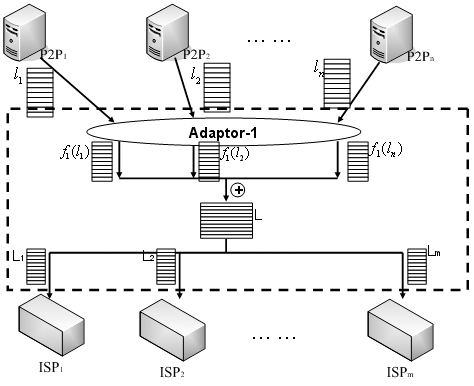

Fig. 4. The Scenario of Step 1 to Step 5

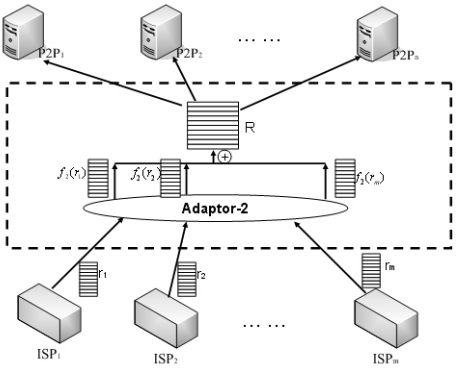

Fig. 5. The Scenario of Step 6 to Step 9

We set up a web server to simulate the P2P-Federation Center which provides some web pages used for registering, logging in and showing the information and maintains a database keeping the peers information. Then we set up two BitTorrent's tracker server to simulate the P2P units, and we use two users registered in the center as ISP units. In this way, the ISP users can acquire the current P2P users' IP and port.

5 Conclusions

This paper discusses P2P management in a new vision. Different from traditional measurement based method, we propose a federation-based solution for P2P management. By exchanging information from a trustable federation center, ISPs can easily acquire P2P peers' information. In this way, we avoid measuring the P2P traffic accurately which is difficult and inextricable to some extent.

At the same time, the P2P service providers who have joined in the federation may get the IP routing information of specific ISPs from the P2P-Federation Center, which help much to improve the quality and efficiency of P2P services.

Therefore, the P2P federation solution described in this paper can achieve a win-win consequence to both ISPs and PSPs. It will advance the P2P technology and improve the management of P2P network.

References

1. Bram Cohen. BitTorrent Protocol Specification v1.0 Identification. http://www.bitconjurer.org/BitTorrent/protocol.html
2. Subhabrata Sen, Jia Wang. Analyzing Peer-to-Peer Traffic Across Large Networks. In: IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 12, NO.2, (APRIL 2004)
3. Bram Cohen. Incentives Build Robustness in BitTorrent. http://bittorrent.com/bittorrentecon.pdf
4. T. Karagiannis, A. Broido, M. Faloutsos, and K. Claffy. Transport Layer Identification of P2P Traffic. In: Proceedings of 2004 ACM SIGCOMM Internet Measurement Conference, IMC 2004, Taormina, Italy, 2004.
5. T. Hamada, K. Chujo, T. Chujo, and X. Yang. Peer-to-Peer Traffic in Metro Networks: Analysis, Modeling, and Policies. In: Proceedings of IEEE Symposium Record on Network Operations and Management Symposium, Seoul, South Korea, (2004).
6. Ashwin R. Bharambe, Cormac Herley, Venkata N. Padmanabhan. Some Observations on BitTorrent Performance. Proceedings of the 2005 ACM SIGMETRICS international conference on Measurement and modeling of computer systems, (Jun. 2005)
7. eDonkey2000. http://www.edonkey2000.com