

[19] China National Intellectual Property Administration



[12]Description of Invention Patent Application Publication

[21] Application Number: 200710013549.4

[51] Int. Cl.

H04L 12/28 (January 2006)

H04L 12/56 (January 2006)

H04L 12/02 (January 2006)

[43] Publication Date: August 15, 2007

[11] Publication Number CN 101018172 A

[22] Application Date: February 9, 2007

[21] Application Number: 200710013549.4

[71] Applicant: Inspur Electronic Information Industry Co., Ltd.
Address: No. 224, Shanda Road, Lixia District, Jinan City,
Shandong Province, 250014

[72] Inventors: Xiaofeng LI; Wenjun ZHUANG; Wei CUI; Yonghui
LIU;
Jian LIU; Yawei ZHAO

[74] Patent Agency: Jinan Xinda Patent Agency Co., Ltd.

Agent: Ming JIANG

Claims: 3 pages; Description: 8 pages; Drawings: 3 pages

[54] Title of Invention

METHOD FOR OPTIMIZING P2P TRANSMISSION IN
METROPOLITAN AREA NETWORK

[57] Abstract

The present invention relates to an optimization method for P2P application in a metropolitan area network, and specific optimization steps are as follows: a. a P2P node, when needing to download a certain resource, being referred to as a requesting P2P node hereinafter; b. first, performing querying by means of a P2P index server to obtain a list of P2P nodes having the resource, these nodes being referred to as resource P2P nodes; c. finding, from among the resource P2P nodes, nodes that are the nearest neighbors of the requesting P2P node on the network; and d. downloading the resource from one or more nearest neighbor resource P2P nodes via the network, so that P2P traffic is kept as far away from a core network as possible and transmitted at the edge of the network, thereby reducing the P2P traffic flowing through backbone networks such as a core network and an aggregation network. Compared with the prior art, the method of the present invention, by adding a topology server and the index server, limits P2P traffic to occur at the edge of the network as much as possible, reducing the transmission load on the backbone networks and thereby avoiding network congestion caused by P2P. Therefore, the method of the present invention has great potential for widespread application.

Subnet IP address	Aggregation switch number (s1)	Access switch number (s2)	Cell switch number (s3)	Building switch number (s4)
...	...	...	...	...
10.30.11.65 / 26	2	3	5	4
10.30.11.129 / 26	2	3	5	5
...	...	...	...	...

METHOD FOR OPTIMIZING P2P TRANSMISSION IN METROPOLITAN AREA NETWORK

TECHNICAL FIELD

The present invention relates to a network transmission optimization technology, and relates in particular to a method for reducing network backbone bandwidth occupancy and preventing network conflicts for P2P application in a metropolitan area network.

BACKGROUND OF THE INVENTION

P2P is an important technology in current network transmission, and provides new technical support for network resource sharing. The P2P technology achieves high transmission performance by maximizing the use of the network bandwidth, and therefore, the P2P technology is rapidly popularized in network applications. Compared with a conventional client/server mode, the P2P technology realizes a distributed storage and sharing capability in the peer-to-peer network. Each node computer is both a client that downloads from other nodes and a storage server that can share content stored thereby with other nodes, so that more massive and richer content can be provided, which becomes one of the important attractions of broadband networks.

With the advanced development and application of the P2P technology, P2P-based text, voice and video communications and various multimedia services have begun to emerge and flourish, and people pay more attention to the development and application of the P2P technology.

However, while applying the P2P technology, people found that the P2P technology has the following disadvantage. The P2P technology preemptively applies the network bandwidth, which occupies most of the network bandwidth, generally more than 60% of the total bandwidth, thereby seriously affecting the normal operation of other services. As a result, broadband operators generally resist or even ban P2P technology. Banning P2P will greatly reduce the attractiveness of broadband networks to users. On the one hand, the operators want to provide P2P services on their networks to make up for the insufficient content resources they provide, so as to compete with other broadband operators. On the other hand, the operators do not want the application of P2P to occupy a large amount of bandwidth, which further affects normal operation of other services. Therefore, this puts the operators in a dilemma regarding the application of P2P technology.

In a metropolitan area network of a broadband operator in the prior art, a user computer is generally connected to a building switch, then connected to a cell switch, and then connected to an access layer switch or router (hereinafter referred to as an access switch), further connected to an aggregation layer switch or router (hereinafter referred to as an aggregation switch), and finally connected to a core switch or router, and a core network of the metropolitan area network is generally formed by one or more core switches or routers. FIG. 1 is a model of a metropolitan area network, which forms a tree structure as a whole. In general, one building switch forms one subnet.

For a metropolitan area network formed by using an ADSL or Cable Modem technology, a user generally accesses an access layer device directly through an ADSL access device or a Cable Modem, and the tree structure is still maintained.

Current P2P file transfer software causes most of the P2P traffic to pass through a backbone network. The closer to a core layer, the larger the P2P traffic, the more serious the bandwidth occupation problem, ultimately causing network congestion and affecting normal business operations. The root of this problem is that when P2P software searches for a service node, it considers that the nodes are completely identical, and the hierarchical structure of the metropolitan area network is not considered; and in the metropolitan area network, the service nodes are selected approximately randomly, resulting in that most of the traffic passes through the backbone network.

As shown in FIG. 1, assuming that the metropolitan area network has a total of n_1 aggregation switches, each aggregation switch has n_2 access switches, each access switch has n_3 cell switches, each cell switch has n_4 building switches, and each building switch is connected to n_5 user computers, the number of computers included in the entire metropolitan area network is $N = n_1 * n_2 * n_3 * n_4 * n_5$. Hereinafter, a link that connects a building computer to a cell switch is referred to as a building-to-cell uplink, a link that connects a cell computer to a switch is referred to as a cell-to-access uplink, a link that connects an access computer to an aggregation switch is referred to as an access-to-aggregation uplink, and a link that connects an aggregation computer to a core network is referred to as an aggregation-to-core uplink.

The following analysis is performed on the conventional P2P traffic load:

When a P2P node needs to download a file with m bytes, it is referred to as a requesting P2P node, and assuming that $x\%$ of computers have the resource, they are referred to as resource P2P nodes,

that is, there are $N * x\%$ resource P2P nodes in total. According to statistical rules, these nodes are uniformly distributed in the entire network.

The requesting P2P node randomly selects resource P2P nodes, and obtains $m/(N * x\%)$ bytes of file content from each resource P2P node.

Therefore, a user node downloads, from an adjacent computer on its own building switch,

$\frac{(n5-1)*x\%*m}{N*x\%}$ bytes of file content, which is approximately equal to $m/(n1 * n2 * n3 * n4)$ bytes, only accounting for $1/(n1 * n2 * n3 * n4)$ of the entire file, and the remaining content needs to be downloaded from a cell switch through a building-to-cell uplink, with a ratio of $\frac{n1*n2*n3*n4-1}{n1*n2*n3*n4}$;

$\frac{(n4*n5-1)*x\%*m}{N*x\%}$ a user node can download, from a computer of its own cell, bytes of file content, which is approximately equal to $m/(n1 * n2 * n3)$ bytes, only accounting for $1/(n1 * n2 * n3)$ of the entire file, and the remaining content needs to be downloaded from an access switch through a cell-to-access uplink, with a ratio of $\frac{n1*n2*n3-1}{n1*n2*n3}$;

$\frac{(n3*n4*n5-1)*x\%*m}{N*x\%}$ a user node can download, from a computer within an access network where it is located, bytes of file content, which is approximately equal to $m/(n1 * n2)$ bytes, only accounting for $1/(n1 * n2)$ of the entire file, and the remaining content needs to be downloaded from an aggregation switch through an access-to-aggregation uplink, with a ratio of $\frac{n1*n2-1}{n1*n2}$;

$\frac{(n2*n3*n4*n5-1)*x\%*m}{N*x\%}$

a user node can download, from a computer within an aggregation network where it is located, bytes of file content, which is approximately equal to $m/n1$ bytes, only accounting for $1/n1$ of the entire file, and the remaining content needs to be downloaded from a core network through an

aggregation-to-core uplink, with a ratio of $\frac{n1-1}{n1}$.

When there are y requesting P2P nodes on the network, each node downloads m bytes of content, the traffic generated on the network is:

P2P traffic downloaded from the core network on each aggregation-to-core uplink is $y * m * \frac{n1-1}{n1^2}$,

and the total traffic generated on the core network is $y * m * \frac{n1-1}{n1}$

P2P downlink traffic on each access-to-aggregation uplink is $y * m * \frac{n1 * n2 - 1}{(n1 * n2)^2}$, which is approximately equal to $y * m / (n1 * n2)$

P2P downlink traffic on each cell-to-access uplink is $y * m * \frac{n1 * n2 * n3 - 1}{(n1 * n2 * n3)^2}$, which is approximately equal to $y * m / (n1 * n2 * n3)$

P2P downlink traffic on each building-to-cell uplink is $y * m * \frac{n1 * n2 * n3 * n4 - 1}{(n1 * n2 * n3 * n4)^2}$, which is approximately equal to $y * m / (n1 * n2 * n3 * n4)$

P2P traffic generated between computers in each building is $y * m / (n1 * n2 * n3 * n4)^2$

Therefore, the ratio of the total traffic $y * m$ generated by P2P, the traffic of the core network, the downlink traffic of the aggregation-to-core uplink, the downlink traffic of the access-to-aggregation uplink, the downlink traffic of the cell-to-access uplink, the downlink traffic of the building-to-cell uplink and the P2P traffic between the computers in the building is $1:(n1-1)/n1:(n1-1)/n1^2:1/(n1 * n2):1/(n1 * n2 * n3):1/(n1 * n2 * n3 * n4):1/(n1 * n2 * n3 * n4)^2$. In this metropolitan area network model, $n1 = 8$, $n2 = 8$, $n3 = 8$, $n4 = 8$, and $n5 = 20$, that is, the metropolitan area network has $N = 81920$ computers. The calculated P2P traffic load ratios of the network at all levels are, for example, the "- ♦ -" broken lines in FIGS. 2 and 3, where the vertical coordinate in FIG. 2 is linear, and the vertical coordinate in FIG. 3 is logarithmic. Apparently, most of the network traffic generated by P2P flows through the core network, which is the reason why P2P causes network congestion. The closer to the edge of the network, the lighter the traffic load.

SUMMARY OF THE INVENTION

At present, among the applications of P2P technology, the main applications that occupy bandwidth are file sharing applications. A large number of files are transmitted repeatedly on the network, causing network congestion problems. Because most of the P2P traffic flows through the backbone network, the main congestion occurs on the backbone network. In order to solve the

congestion problem in the application of the P2P technology, the present invention proposes the following technical solutions.

A method for optimizing P2P transmission in a metropolitan area network of the present invention is as follows:

- a. a P2P node, when needing to download a certain resource, being referred to as a requesting P2P node hereinafter;
- b. first, performing querying by means of a P2P index server to obtain a list of P2P nodes having the resource, these nodes being referred to as resource P2P nodes;
- c. finding, from among the resource P2P nodes, nodes that are the nearest neighbors of the requesting P2P node on the network; and
- d. downloading the resource from one or more nearest neighbor resource P2P nodes via the network, so that P2P traffic is kept as far away from a core network as possible and transmitted at an edge of the network, thereby reducing the P2P traffic flowing through backbone networks such as a core network and an aggregation network.

(1) The step of finding, from among the resource P2P nodes, nodes that are the nearest neighbors of the requested P2P node on the network comprises:

- a. first, finding out whether there is a resource P2P node connected to the same building switch as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;
- b. otherwise, finding out whether there is a resource P2P node connected to the same cell switch or router as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;
- c. otherwise, finding out whether there is a resource P2P node connected to the same access switch or router as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;
- d. otherwise, finding out whether there is a resource P2P node connected to the same aggregation switch or router as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;

e. otherwise, all the resource P2P nodes being considered as nearest neighbor resource P2P nodes.

(2) A step of numbering switches or routers for searching for the nearest neighbor resource P2P node comprises:

- a. numbers s_1 assigned to aggregation switches or routers being from 1 to n_1 ;
- b. numbers s_2 assigned to access switches or routers under each aggregation switch or router being from 1 to n_2 ;
- c. numbers s_3 assigned to cell switches or routers under each access switch or router being from 1 to n_3 ;
- d. numbers s_4 assigned to building switches under each cell switch being from 1 to n_4 ; and
- e. the numbers of the aggregation switches or routers, the access switches and routers, the cell switches and routers and the building switches traversed from a core network to each computer form a location vector $S = (s_1, s_2, s_3, s_4)$ that describes the topological network location of the computer, and the computers under the same building switch have a same location vector.

(3) A calculation method for searching for the nearest neighbor resource P2P node is as follows:

- a. defining a distance vector D of two location vectors $S = (s_1, s_2, s_3, s_4)$ and $S' = (s'_1, s'_2, s'_3, s'_4)$ as

$$D = (d_1, d_2, d_3, d_4) = (s_1, s_2, s_3, s_4) - (s'_1, s'_2, s'_3, s'_4) = (s_1 - s'_1, s_2 - s'_2, s_3 - s'_3, s_4 - s'_4),$$

wherein the subtraction of each component is defined as:

$$d_i = \begin{cases} 0, & s_i = s'_i \\ 1, & s_i \neq s'_i \end{cases}$$

- b. defining a size comparison method for two distance vectors $D = (d_1, d_2, d_3, d_4)$ and $D' = (d'_1, d'_2, d'_3, d'_4)$ as:

$$\left\{ \begin{array}{l} D = D', \text{ When } d1 = d'1, d2 = d'2, d3 = d'3, d4 = d'4 \\ D > D', \text{ When } \left\{ \begin{array}{l} d1 = 1, d'1 = 0 \\ d1 = d'1, d2 = 1, d'2 = 0 \\ d1 = d'1, d2 = d'2, d3 = 1, d'3 = 0 \\ d1 = d'1, d2 = d'2, d3 = d'3, d4 = 1, d'4 = 0 \end{array} \right. \end{array} \right.$$

wherein a small distance vector indicates a nearer neighbor on the network.

(4) The topology server added in the metropolitan area network is used to store a correspondence between subnets and location vectors, stored content includes a subnet address and mask of each building switch, as well as the location vectors, and working steps of the topology server are as follows:

a. according to an IP address, comparing the IP address against stored subnets to find out a subnet to which the IP address belongs, thereby obtaining a location vector of a computer with the IP address;

b. according to an IP address of a requesting P2P node and IP addresses of a plurality of resource P2P nodes, calculating a distance vector between the requesting P2P node and each resource P2P node, and comparing the distance vectors to find a resource P2P node with the smallest distance vector;

c. reflecting, by stored data, a current topology of the metropolitan area network, which is a basis for calculating the nearest neighbor resource nodes.

(5) After the P2P index server obtains resource P2P nodes, the topology server calculates the nearest neighbor nodes, and specific calculation steps are as follows:

a. receiving, by the P2P index server, a resource query request from the requesting P2P node and performing querying to obtain the resource P2P nodes having the resource;

b. sending the IP address of the requesting P2P node and the IP address of each obtained resource P2P node to the topology server; or sending the list of the resource P2P nodes to the requesting P2P node, and sending, by the requesting P2P node, the IP address of the requesting P2P node and the IP address of each resource P2P node to the topology server;

c. calculating, by the topology server, one or more nearest neighbor resource P2P nodes of the requesting P2P node, and then returning the nearest neighbor resource P2P nodes to

the index server, and sending, by the index server, a list of the nearest neighbor resource P2P nodes to the requesting P2P node; and

e. the requesting P2P node connecting to and downloading from the nearest connected resource P2P node.

(6) The list of all resource P2P nodes obtained by the requesting P2P node comprises a location vector of each node, and specific steps of directly calculating the nearest neighbor resource P2P nodes by the requesting P2P node are as follows:

a. when P2P software on a computer is started, first, requesting a location vector of the computer from the topology server;

b. querying, by the topology server, a topology data table for a subnet to which the computer belongs according to an IP address of the computer, thereby obtaining the location vector of the computer, and returning the location vector to the computer;

c. reporting, by each P2P node computer, the location vector of the P2P node computer while reporting resources of the P2P node computer to the P2P index server;

d. when a requesting P2P node computer needs to download a certain resource, first performing querying to obtain a list of nodes having the resource, wherein the list of nodes comprises a location vector of each resource node; and

e. subtracting, by the requesting P2P node, the location vector of each resource node from the location vector of the requesting P2P node to obtain a corresponding distance vector, and selecting one or more resource nodes with the smallest distance from the resource nodes to download the resource.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagram of a metropolitan area network model;

FIG. 2 is a distribution diagram of P2P traffic of a metropolitan area network with a linear coordinate as a vertical ordinate;

FIG. 3 is a distribution diagram of P2P traffic of a metropolitan area network with a logarithmic coordinate;

FIG. 4 is a diagram of a metropolitan area network model after numbers are assigned to switches or routers; and

FIG. 5 is a schematic diagram of a data storage structure of a topology server.

DETAILED DESCRIPTION OF THE EMBODIMENTS

Implementation solution for optimizing P2P

First, numbers are assigned to aggregation switches, access switches, cell switches, and building switches of a metropolitan area network as shown in FIG. 4: Numbers s_1 assigned to the aggregation switches are from 1 to n_1 , numbers s_2 assigned to the access switches under each aggregation switch are from 1 to n_2 , numbers s_3 assigned to the cell switches under each access switch are from 1 to n_3 , and numbers s_4 assigned to the building switches under each cell switch are from 1 to n_4 . In this way, each subnet or building switch can be represented by a vector $S = (s_1, s_2, s_3, s_4)$, which is also a common location vector of the computers under this switch. It is obvious that two computers with the same location vector must access the same building switch.

When the requesting P2P node computer can download from multiple resource computers, first, the location vector $S' = (s'_1, s'_2, s'_3, s'_4)$ of each resource P2P node is subtracted from its own location vector $S = (s_1, s_2, s_3, s_4)$ to obtain a distance vector $D = (d_1, d_2, d_3, d_4)$ from each resource P2P node

$$(d_1, d_2, d_3, d_4) = (s_1, s_2, s_3, s_4) - (s'_1, s'_2, s'_3, s'_4) = (s_1 - s'_1, s_2 - s'_2, s_3 - s'_3, s_4 - s'_4)$$

wherein the subtraction of each component is defined as:

$$d_i = \begin{cases} 0, & s_i = s'_i \\ 1, & s_i \neq s'_i \end{cases}$$

and a size comparison method for two distance vectors $D = (d_1, d_2, d_3, d_4)$ and $D' = (d'_1, d'_2, d'_3, d'_4)$ is defined as:

$$\begin{cases} D = D', & \text{When } d_1 = d'_1, d_2 = d'_2, d_3 = d'_3, d_4 = d'_4 \\ D > D', & \text{When } \begin{cases} d_1 = 1, d'_1 = 0 \\ d_1 = d'_1, d_2 = 1, d'_2 = 0 \\ d_1 = d'_1, d_2 = d'_2, d_3 = 1, d'_3 = 0 \\ d_1 = d'_1, d_2 = d'_2, d_3 = d'_3, d_4 = 1, d'_4 = 0 \end{cases} \end{cases}$$

According to the above algorithm, the distance vectors from the respective resource P2P nodes are compared, and a resource P2P node with the minimum distance is found to download therefrom, so that the result of optimizing the P2P transmission can be achieved.

A server is set up in the metropolitan area network, and stores the topological structure data of the entire metropolitan area network. The data structure is shown in FIG. 5, and the stored content is the correspondence between subnets and location vectors. The server is configured to calculate a location vector according to an IP address of a computer, and is referred to as a topology server hereinafter.

In a specific application, distance comparison calculation may be performed on the P2P node or the topology server.

The working process of performing a distance comparison at a P2P node is as follows:

- A. when P2P software on a computer is started, first, requesting a location vector of the computer from the topology server;
- B. querying, by the topology server, a topology data table for a subnet to which the computer belongs according to the IP address of the computer, thereby obtaining the location vector of the computer, and returning the location vector to the computer;
- C. when the P2P software on the computer is started, first, requesting the location vector of the computer from the topology server;
- D. reporting, by each P2P node computer, a location vector of the P2P node computer while reporting a resource owned by the P2P node computer to a P2P index server;
- E. when the requesting P2P node computer needs to download a certain resource, first performing querying to obtain a list of nodes having the resource, wherein the list of nodes comprises a location vector of each resource node; and
- F. subtracting, by the requesting P2P node, the location vector of each resource node from the location vector of the requesting P2P node to obtain a corresponding distance vector, and selecting one or more resource nodes with the smallest distance from the resource nodes to download the resource.

The working process of performing a distance comparison by the topology server is as follows:

- A. when a requesting P2P node needs to download a certain resource, first sending a query request for the resource to the P2P index server;
- B. receiving, by the P2P index server, a resource query request from the requesting P2P node and performing querying to obtain resource P2P nodes having the resource;
- C. sending the IP address of the requesting P2P node and the IP address of each obtained resource P2P node to the topology server; or sending the list of the resource P2P nodes to the requesting P2P node, and sending, by the requesting P2P node, the IP address of the requesting P2P node and the IP address of each resource P2P node to the topology server;
- D. calculating, by the topology server, one or more nearest neighbor resource P2P nodes of the requesting P2P node, as described in claim 4; and then, returning the nearest neighbor resource P2P nodes to the index server, and sending, by the index server, the list of the nearest neighbor resource P2P nodes to the requesting P2P node; or directly sending the list of the nearest neighbor resource P2P nodes to the requesting P2P node; and
- E. the requesting P2P node connecting to and downloading from the nearest connected resource P2P node.

Example

Optimized P2P access strategy

If the P2P traffic can be limited to the edge of the network as much as possible, the bandwidth pressure of the backbone network can be effectively reduced. Therefore, the P2P access mode is modified, and the following new download strategy is adopted:

When a node needs to download a certain resource, and when other computers under a building switch where the node is located possess the resource, the resource is downloaded only from the computers under the building switch; otherwise, if other computers in a current cell network possess the resource, the resource is downloaded only from the computers in the current cell network; otherwise, if other computers in a current access network possess the resource, the resource is downloaded only from the computers in the current access network; otherwise, if other computers in a current aggregation network possess the resource, the resource is downloaded only from the computers in the current aggregation network; and if none of the computers in the current aggregation network possesses the resource, the resource is downloaded from another aggregation

network via a core network. Such a P2P download strategy can push P2P traffic to the edge of the network as much as possible, thereby reducing the load of the backbone network. The traffic distribution characteristics of the P2P using the strategy are analyzed below under statistical rules:

A requesting P2P node needs to download a file with m bytes, and among N computers on the metropolitan area network, the proportion of resource P2P nodes having the resource is $x\%$.

If there are one or more resource P2P nodes under the current building switch, the download traffic will only occur inside the current building switch, and no downlink traffic will be generated on the building-to-cell uplink. If there is no next resource node under the current building switch, all the m bytes need to be downloaded through the building-to-cell uplink. The probability p_5 that there is no resource node under the current building switch is:

$$p_5 = \left(\frac{N - n_5}{N} \right)^{N \cdot x\%} = \left(\frac{n_1 * n_2 * n_3 * n_4 - 1}{n_1 * n_2 * n_3 * n_4} \right)^{N \cdot x\%}$$

Then, under statistical conditions, the P2P downlink traffic generated on the building-to-cell uplink is $p_5 * m$ bytes, and the traffic downloaded from other computers of the current building is $(1 - p_5) * m$ bytes.

Similarly, the probability p_4 that there is no resource P2P node in the current cell network is:

$$p_4 = \left(\frac{N - n_4 * n_5}{N} \right)^{N \cdot x\%} = \left(\frac{n_1 * n_2 * n_3 - 1}{n_1 * n_2 * n_3} \right)^{N \cdot x\%}$$

Therefore, the P2P download traffic generated on the cell-to-access uplink is $p_4 * m$ bytes.

The probability p_3 that there is no resource P2P node in the current access network is:

$$p_3 = \left(\frac{N - n_3 * n_4 * n_5}{N} \right)^{N \cdot x\%} = \left(\frac{n_1 * n_2 - 1}{n_1 * n_2} \right)^{N \cdot x\%}$$

Therefore, the P2P download traffic generated on the access-to-aggregation uplink is $p_3 * m$ bytes.

The probability p_2 that there is no resource P2P node in the current aggregation network is:

$$p_2 = \left(\frac{N - n_2 * n_3 * n_4 * n_5}{N} \right)^{N \cdot x\%} = \left(\frac{n_1 - 1}{n_1} \right)^{N \cdot x\%}$$

Therefore, the P2P download traffic generated on the aggregation-to-core uplink is $p_2 * m$ bytes.

If there are y requesting P2P nodes uniformly distributed in the metropolitan area network to

download the resource, the P2P traffic inside each building switch is $\frac{y * (1 - p_5) * m}{n_1 * n_2 * n_3 * n_4}$ bytes, the P2P

downlink traffic generated on each building-to-cell uplink is $\frac{y * p_5 * m}{n_1 * n_2 * n_3 * n_4}$ bytes, the P2P

downlink traffic generated on the cell-to-access uplink is $\frac{y * p_4 * m}{n_1 * n_2 * n_3}$ bytes, the P2P downlink traffic

generated on each access-to-aggregation uplink is $\frac{y * p_3 * m}{n_1 * n_2}$ bytes, the P2P downlink traffic

generated on each aggregation-to-core uplink is $\frac{y * p_2 * m}{n_1}$ bytes, and all P2P traffic flowing through the core is $y * p_2 * m$ bytes.

Therefore, under this optimization strategy, the ratio of the total traffic $y * m$ generated by P2P, the traffic of the core network, the downlink traffic of the aggregation-to-core uplink, the downlink traffic of the access-to-aggregation uplink, the downlink traffic of the cell-to-access uplink, the downlink traffic of the building-to-cell uplink and the P2P traffic between the computers inside the building is $1:p_2:p_2/n_1:p_3/(n_1 * n_2):p_4/(n_1 * n_2 * n_3):p_5/(n_1 * n_2 * n_3 * n_4):(1 - p_5)/(n_1 * n_2 * n_3 * n_4)$. According to the above metropolitan area network model parameters, the traffic loads generated by P2P when x is 0.1%, 0.05%, 0.02% and 0.01% are calculated, respectively, and the results are shown in the solid broken lines in FIGS. 2 and 3. Apparently, the P2P load on the core network inside the access switch is significantly reduced by transferring the traffic to the very edge of the network - inside the building switch. For example, when the computers having resources account for 1/10,000 of the entire network ($x = 0.01\%$), the load of the core network is about 1/3 of that in the conventional P2P mode, and the P2P traffic in the building is increased by about one order of magnitude, thereby increasing the transmission speed of users.

CLAIMS

1. A method for optimizing P2P transmission in a metropolitan area network, characterized in that a topology server and a P2P index server are added in the metropolitan area network, and P2P traffic is limited to occur at an edge of the network as much as possible, reducing transmission load of backbone networks and avoiding network congestion caused by P2P; and specific steps are as follows:

- a. a P2P node, when needing to download a certain resource, being referred to as a requesting P2P node hereinafter;
- b. first, performing querying by means of the P2P index server to obtain a list of P2P nodes having the resource, these nodes being referred to as resource P2P nodes;
- c. finding, from among the resource P2P nodes, nodes that are the nearest neighbors of the requesting P2P node on the network; and
- d. downloading the resource from one or more nearest neighbor resource P2P nodes via the network, so that P2P traffic is kept as far away from a core network as possible and transmitted at an edge of the network, thereby reducing the P2P traffic flowing through backbone networks such as a core network and an aggregation network.

2. The method according to claim 1, wherein the step of finding, from among the resource P2P nodes, nodes that are the nearest neighbors of the requesting P2P node on the network comprises:

- a. first, finding out whether there is a resource P2P node connected to the same building switch as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;
- b. otherwise, finding out whether there is a resource P2P node connected to the same cell switch or router as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;

c. otherwise, finding out whether there is a resource P2P node connected to the same access switch or router as the requesting P2P node, and if so, the resource P2P node being a nearest neighbor resource P2P node;

d. otherwise, finding out whether there is a resource node connected to the same aggregation switch or router as the requesting P2P node, and if so, the resource node being a nearest neighbor resource P2P node;

e. otherwise, all the resource P2P nodes being considered as nearest neighbor resource P2P nodes.

3. The method according to claim 1, wherein a step of assigning numbers to switches or routers for searching for a nearest neighbor resource P2P node comprises:

a. numbers s_1 assigned to aggregation switches or routers being from 1 to n_1 ;

b. numbers s_2 assigned to access switches or routers under each aggregation switch or router being from 1 to n_2 ;

c. numbers s_3 assigned to cell switches or routers under each access switch or router being from 1 to n_3 ;

d. numbers s_4 assigned to building switches under each cell switch being from 1 to n_4 ; and

e. the numbers assigned to the aggregation switches or routers, the access switches and routers, the cell switches and routers and the building switches traversed from a core network to each computer form a location vector $S = (s_1, s_2, s_3, s_4)$ that describes the topological network location of the computer, and the computers under the same building switch have a same location vector.

4. The method according to claim 1, wherein a calculation method for searching for the nearest neighbor resource P2P node is as follows:

a. defining a distance vector D of two location vectors $S = (s_1, s_2, s_3, s_4)$ and $S' = (s'_1, s'_2, s'_3, s'_4)$ as $D = (d_1, d_2, d_3, d_4) = (s_1, s_2, s_3, s_4) - (s'_1, s'_2, s'_3, s'_4) = (s_1 - s'_1, s_2 - s'_2, s_3 - s'_3, s_4 - s'_4)$, wherein subtraction of each component is defined as:

$$d_i = \begin{cases} 0, & s_i = s'_i \\ 1, & s_i \neq s'_i \end{cases}; \text{ and}$$

b. defining a size comparison method for two distance vectors $D = (d_1, d_2, d_3, d_4)$ and $D' = (d'_1, d'_2, d'_3, d'_4)$ as:

$$\begin{cases} D = D', & \text{When } d_1 = d'_1, d_2 = d'_2, d_3 = d'_3, d_4 = d'_4 \\ D > D', & \text{When } \begin{cases} d_1 = 1, d'_1 = 0 \\ d_1 = d'_1, d_2 = 1, d'_2 = 0 \\ d_1 = d'_1, d_2 = d'_2, d_3 = 1, d'_3 = 0 \\ d_1 = d'_1, d_2 = d'_2, d_3 = d'_3, d_4 = 1, d'_4 = 0 \end{cases} \end{cases}$$

wherein a small distance vector indicates a nearer neighbor on the network.

5. The method according to claim 1, wherein the topology server added in the metropolitan area network is used to store a correspondence between subnets and location vectors, stored content includes a subnet address and mask of each building switch, as well as the location vectors, and working steps of the topology server are as follows:

a. according to an IP address, comparing the IP address against stored subnets to find out a subnet to which the IP address belongs, thereby obtaining a location vector of a computer with the IP address;

b. according to an IP address of a requesting P2P node and IP addresses of a plurality of resource P2P nodes, calculating a distance vector between the requesting P2P node and each resource P2P node, and comparing the distance vectors to find a resource P2P node with the smallest distance vector; and

c. reflecting, by stored data, a current topology of the metropolitan area network, which is a basis for calculating nearest neighbor resource nodes.

6. The method according to claim 5, wherein after the P2P index server obtains the resource P2P nodes, the topology server calculates the nearest neighbor nodes, and specific calculation steps are as follows:

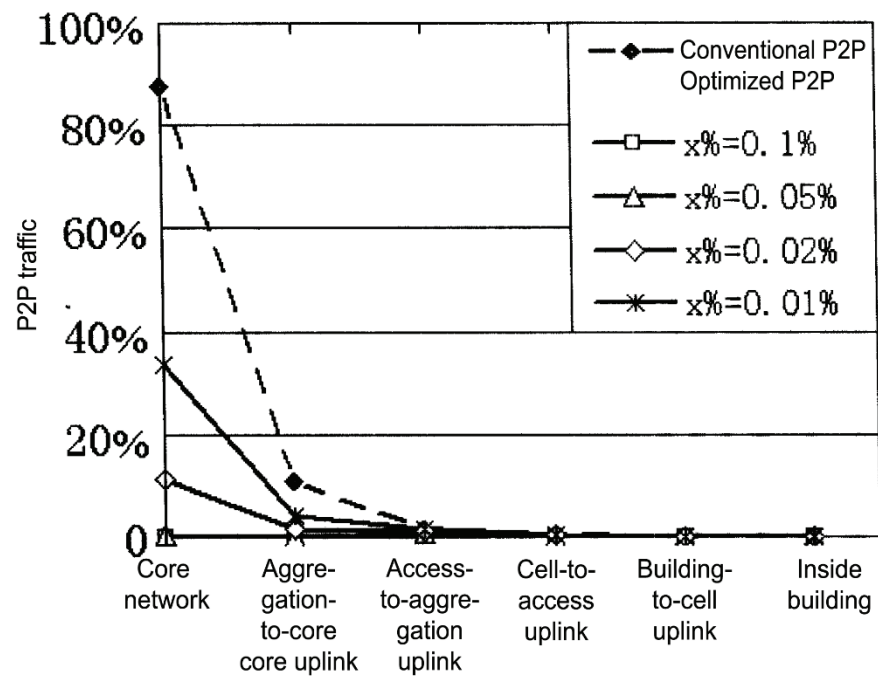
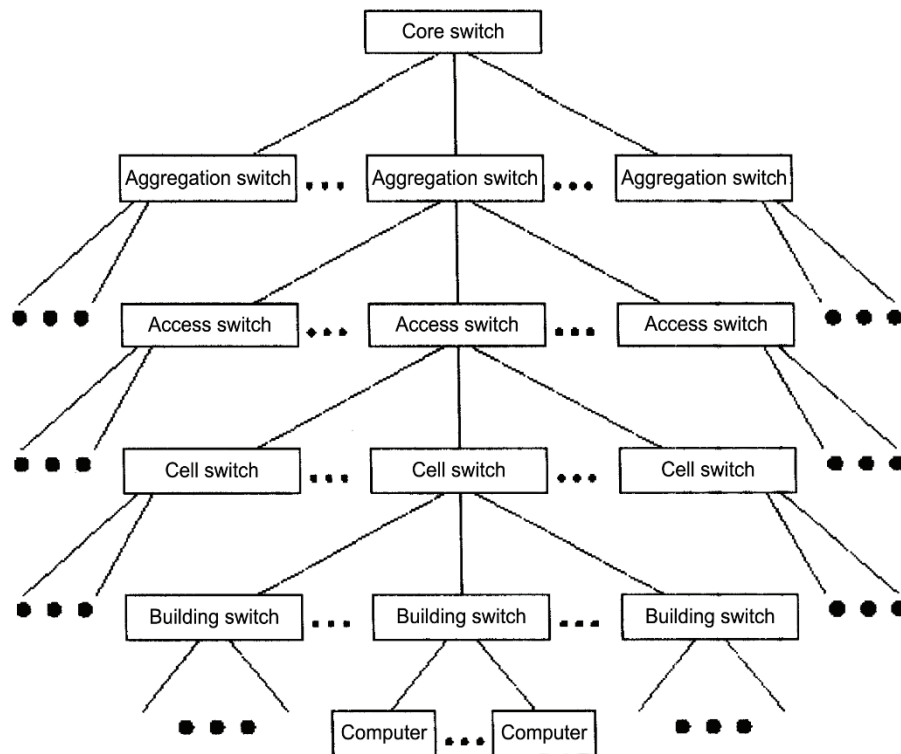
- a. receiving, by the P2P index server, a resource query request from the requesting P2P node and performing querying to obtain the resource P2P nodes having the resource;
- b. sending the IP address of the requesting P2P node and the IP address of each obtained resource P2P node to the topology server; or sending the list of the resource P2P nodes to the requesting P2P node, and sending, by the requesting P2P node, the IP address of the requesting P2P node and the IP address of each resource P2P node to the topology server;
- c. calculating, by the topology server, one or more nearest neighbor resource P2P nodes of the requesting P2P node, and then returning the nearest neighbor resource P2P nodes to the index server, and sending, by the index server, a list of the nearest neighbor resource P2P nodes to the requesting P2P node; and
- d. the requesting P2P node connecting to and downloading from a nearest connected resource P2P node.

7. The method according to claim 1, wherein the list of all resource P2P nodes obtained by the requesting P2P node comprises a location vector of each node, and specific steps of directly calculating the nearest neighbor resource P2P nodes by the requesting P2P node are as follows:

- a. when P2P software on a computer is started, first, requesting a location vector of the computer from the topology server;
- b. querying, by the topology server, a topology data table for a subnet to which the computer belongs according to an IP address of the computer, thereby obtaining the location vector of the computer, and returning the location vector to the computer;
- c. reporting, by each P2P node computer, a location vector of the P2P node computer while reporting a resource owned by the P2P node computer to the P2P index server;

d. when a requesting P2P node computer needs to download a certain resource, first performing querying to obtain a list of nodes having the resource, wherein the list of nodes comprises a location vector of each resource node; and

e. subtracting, by the requesting P2P node, the location vector of each resource node from the location vector of the requesting P2P node to obtain a corresponding distance vector, and selecting, from the resource nodes, one or more resource nodes with the smallest distance to download the resource.



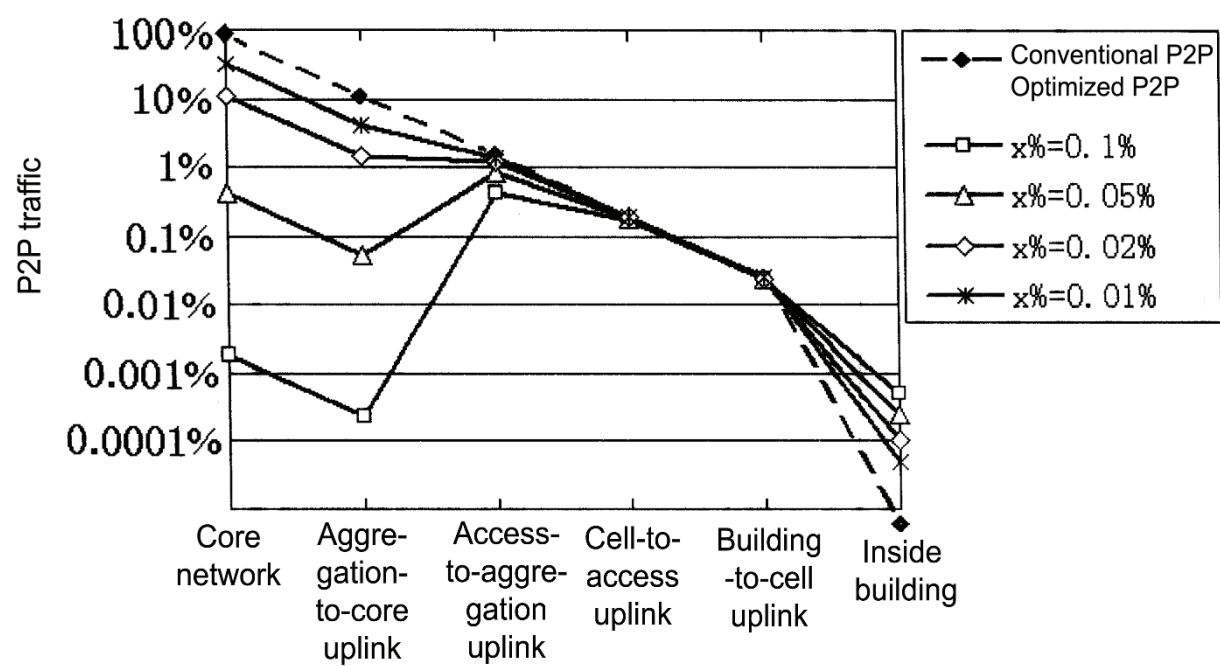


FIG. 3

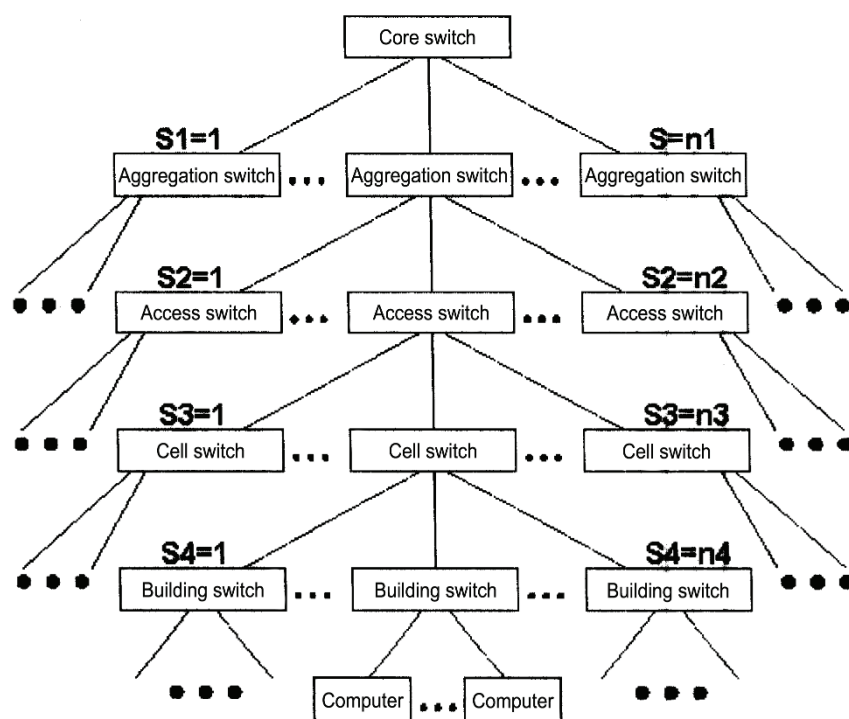


FIG. 4

Subnet IP address	Aggregation switch number (s1)	Access switch number (s2)	Cell switch number (s3)	Building switch number (s4)
...	...	...	...	...
10.30.11.65 / 26	2	3	5	4
10.30.11.129 / 26	2	3	5	5
...	...	...	...	...

FIG. 5



TRANSLATOR'S CERTIFICATE OF TRANSLATION

Questel-Orbit, Incorporated
2331 Mill Road, Suite 600
Alexandria, VA 22314, U.S.A.
www.questel.com

Questel Project Number: FRPC2520013UT-LZ_Q
Date: April 23, 2025

To whom it may concern:

I, Roger Chao, a translator fluent in the Chinese and English languages, on behalf of Questel, do solemnly and sincerely declare that the following is, to the best of my knowledge and belief, a true and correct translation of the document(s) listed below in a form that best reflects the intention and meaning of the original text.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true. I further declare that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Original Document Identifier: Application CN101018172A.pdf
Translated Document Identifier: CN101018172A_en-US.docx

Signed this 23rd day of April, 2025

Signature

Roger Chao

Name

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

H04L 12/28 (2006.01)

H04L 12/56 (2006.01)

H04L 12/02 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200710013549.4

[43] 公开日 2007 年 8 月 15 日

[11] 公开号 CN 101018172A

[22] 申请日 2007.2.9

[21] 申请号 200710013549.4

[71] 申请人 浪潮电子信息产业股份有限公司
地址 250014 山东省济南市历下区山大路 224 号

[72] 发明人 李晓峰 庄文君 崔 卫 刘永辉
刘 剑 赵亚伟

[74] 专利代理机构 济南信达专利事务所有限公司
代理人 姜 明

权利要求书 3 页 说明书 8 页 附图 3 页

[54] 发明名称

一种优化城域网内 P2P 传输的方法

[57] 摘要

本发明是一种针对城域网的 P2P 应用的优化方法，具体优化步骤是：a. 当一个 P2P 节点需要下载某资源时，以下称之为请求 P2P 节点；b. 首先通过 P2P 索引服务器查询到拥有该资源的 P2P 节点列表，这些节点称之为资源 P2P 节点；c. 在资源 P2P 节点中寻找出与请求 P2P 节点在网络上最近邻的节点；d. 资源通过网络上从最近邻的一个或多个资源 P2P 节点下载，使 P2P 流量尽量远离核心网，在网络边缘传输，从而降低流经核心网和汇聚网等骨干网络的 P2P 流量。本发明的方法和现有技术相比，通过增加拓扑服务器和索引服务器，把 P2P 流量尽量限制在网络的边缘发生，降低骨干网的传输负载，以此避免 P2P 造成网络堵塞。因而具有很好的推广使用价值。

子网 IP 地址	汇聚交换机 编码(s1)	接入交换机 编码(s2)	小区交换机 编码(s3)	楼宇交换机 编码(s4)
...	...	...	...	...
10.30.11.65 / 26	2	3	5	4
10.30.11.129 / 26	2	3	5	5
...	...	...	...	...

1、一种优化城域网内 P2P 传输的方法，其特征在于在城域网内增设拓扑服务器和 P2P 索引服务器，把 P2P 流量尽量限制在网络的边缘发生，降低骨干网的传输负载，避免 P2P 造成网络堵塞，具体步骤如下：

- a、当一个 P2P 节点需要下载某资源时，以下称之为请求 P2P 节点；
- b、首先通过 P2P 索引服务器查询到拥有该资源的 P2P 节点列表，这些节点称之为资源 P2P 节点；
- c、在资源 P2P 节点中寻找出与请求 P2P 节点在网络上最近邻的节点；
- d、资源通过网络上从最近邻的一个或多个资源 P2P 节点下载，使 P2P 流量尽量远离核心网，在网络边缘传输，从而降低流经核心网和汇聚网等骨干网络的 P2P 流量。

2、根据权利要求 1 所述的方法，其特征在于，在资源 P2P 节点中寻找与请求 P2P 节点在网络上最近邻的节点的步骤如下：

- a、首先寻找是否有与请求 P2P 节点接入同一个楼宇交换机的资源 P2P 节点，如果有则是最近邻的资源 P2P 节点；
- b、否则寻找是否有与请求 P2P 节点接入同一个小区交换机或路由器的资源 P2P 节点，如果有则是最近邻的资源 P2P 节点；
- c、否则寻找是否有与请求 P2P 节点接入同一个接入交换机或路由器的资源 P2P 节点，如果有则是最近邻的资源 P2P 节点；
- d、否则寻找是否有与请求 P2P 节点接入同一个汇聚交换机或路由器的资源 P2P 节点，如果有则是最近邻的资源 P2P 节点；
- e、否则所有的资源 P2P 节点都认为是最近邻的资源 P2P 节点。

3、根据权利要求 1 所述的方法，其特征在于，查找最近邻资源 P2P 节点所进行的交换机或路由器编码的步骤如下：

- a、汇聚交换机或路由器的编码 s_1 为从 1 到 n_1 ；
- b、每个汇聚交换机或路由器下的接入交换机或路由器的编码 s_2 为从 1 到 n_2 ；
- c、每个接入交换机或路由器下的小区交换机或路由器的编码 s_3 为从 1 到 n_3 ；
- d、每个小区交换机下的楼宇交换机的编码 s_4 为从 1 到 n_4 ；
- e、从核心网到每个计算机，所经过的汇聚交换机或路由器、接入交换机和路由器、小区交换机和路由器和楼宇交换机的编码构成了描述这台计算机拓扑网络位置的位置向量 $S=(s_1,s_2,s_3,s_4)$ ，同一个楼宇交换机下的计算机具有相同的位置向量。

4、根据权利要求 1 所述的方法，其特征在于，查找最近邻资源 P2P 节点的计算方法如

下:

a、定义 2 个位置向量 $S=(s1,s2,s3,s4)$ 和 $S'=(s'1,s'2,s'3,s'4)$ 的距离向量 D 为

$$D=(d1,d2,d3,d4)=(s1,s2,s3,s4)-(s'1,s'2,s'3,s'4)=(s1-s'1,s2-s'2,s3-s'3,s4-s'4)$$

其中各分量的减法定义为:

$$d_i = \begin{cases} 0, & si = s'i \\ 1, & si \neq s'i \end{cases}$$

b、定义 2 个距离向量 $D=(d1,d2,d3,d4)$ 和 $D'=(d'1,d'2,d'3,d'4)$ 的大小比较方法为

$$\begin{cases} D = D', & \text{当 } d1 = d'1, d2 = d'2, d3 = d'3, d4 = d'4 \\ D > D', & \text{当 } \begin{cases} d1 = 1, d'1 = 0 \\ d1 = d'1, d2 = 1, d'2 = 0 \\ d1 = d'1, d2 = d'2, d3 = 1, d'3 = 0 \\ d1 = d'1, d2 = d'2, d3 = d'3, d4 = 1, d'4 = 0 \end{cases} \end{cases}$$

距离向量小, 表示在网络上更近邻。

5、 根据权利要求 1 所述的方法, 其特征在于, 在城域网中增设的拓扑服务器用于存储各子网与位置向量的对应关系, 存储内容为各楼宇交换机的子网地址和掩码, 以及位置向量, 拓扑服务器的工作步骤如下:

a、根据 IP 地址, 与存储的各子网进行比对, 寻找出所属子网, 从而得到该 IP 地址计算机的位置向量;

b、 根据一个请求 P2P 节点的 IP 地址和多个资源 P2P 节点 IP 地址, 计算出请求 P2P 节点与各资源 P2P 节点的距离向量, 并进行距离向量的比较, 找出距离向量最小的资源 P2P 节点。

c、所存储的数据反映城域网当前的拓扑结构, 是计算最近邻资源节点的判断依据。

6、 根据权利要求 5 所述的方法, 其特征在于, P2P 索引服务器查询出资源 P2P 节点后, 由拓扑服务器计算出最近邻节点, 具体计算步骤如下:

a、P2P 索引服务器接收请求 P2P 节点的资源查询请求, 查询出拥有该资源的资源 P2P 节点;

b、把请求 P2P 节点的 IP 地址和查询出的各资源 P2P 节点 IP 地址发送给拓扑服务器; 或者把资源 P2P 节点列表发送给请求 P2P 节点, 请求 P2P 节点把自己的 IP 地址和各资源 P2P 节点的 IP 地址发送给拓扑服务器;

c、拓扑服务器计算出与请求 P2P 节点最近邻的一个或多个资源 P2P 节点, 然后返回给索引服务器, 索引服务器把最近邻的资源 P2P 节点列表发送给请求 P2P 节点;

d、该请求 P2P 节点与最近连的资源 P2P 节点进行连接和下载。

7、 根据权利要求 1 所述方法, 其特征在于, 请求 P2P 节点获得的所有资源 P2P 节点列

表中包含各节点的位置向量，请求 P2P 节点直接计算最近邻的资源 P2P 节点的具体步骤如下

- a、计算机上的 P2P 软件启动时，首先向拓扑服务器请求自己的位置向量；
- b、拓扑服务器根据计算机的 IP 地址，在拓扑数据表中查询所属的子网，从而得到该计算机的位置向量，并返回给该计算机；
- c、各 P2P 节点计算机向 P2P 的索引服务器上报告自己所拥有的资源时，同时上报自己的位置向量；
- d、当请求 P2P 节点计算机要下载某资源时，首先查询得到拥有该资源的节点列表，其中包含各资源节点的位置向量；
- e、请求 P2P 节点分别把自己的位置向量与各资源节点的位置向量相减得到相应的距离向量，从中选择距离最小的一个或多个资源节点，进行资源的下载。

一种优化城域网内 P2P 传输的方法

技术领域

本发明涉及一种网络传输优化技术，尤其是一种在城域网的 P2P 应用中降低对网络主干带宽占用，防止网络冲突的方法。

技术背景

P2P 是目前网络传输中的一个重要技术，为实现网络资源共享提供了新的技术支持。P2P 技术是通过尽其所能利用的网络带宽，获得较高的传输性能，因此在网络应用中得到迅速的普及。P2P 技术和传统的客户机/服务器模式相比，实现了对等网下的分布式存储和共享能力，每个节点计算机既是向其它节点进行下载的客户端，又是一个存储服务器，能与其它节点共享其存储内容，因此能够提供更海量和更丰富的内容，成为宽带网的重要吸引力之一。

随着 P2P 技术的发展和应用的深入，基于 P2P 的文字、语音和视频通信，以及多种多媒体服务开始出现和蓬勃发展，人们更加重视 P2P 技术的开发与应用。

但是人们在 P2P 技术应用的同时，发现其存在的不足是，P2P 技术对网络带宽抢占式应用，其占用了网络的大部分带宽，一般能够占到总带宽的 60%以上，从而严重影响了其它业务的正常运行，这种结果导致宽带运营商对 P2P 技术一般采用抵制甚至遭到封杀，封杀 P2P 会大幅度降低宽带网对用户的吸引力，运营商一方面希望能够在其网络上提供 P2P 服务，弥补自己提供的内容资源的不足，以与其它的宽带运营商进行竞争；另一方面，运营商又不希望 P2P 的应用占用大量的带宽，进而影响其它业务的正常进行。因而，使得运营商对 P2P 技术的应用处于进退两难的地步。

在现有技术的宽带运营商的城域网中，用户计算机一般接入楼宇交换机，再连接到小区交换机，然后连接到接入层交换机或路由器，以下称为接入交换机，进而接入汇聚层交换机或路由器，以下称为汇聚交换机，最后连接到核心交换机或路由器，一般由一台到多台核心交换机或路由器构成城域网的核心网。图 1 是一个城域网的模型，整个构成一个树形结构。通常情况一个楼宇交换机构成一个子网。

对于采用 ADSL 或者 Cable Modem 技术构成的城域网，则用户通过 ADSL 接入设备或者 Cable Modem 一般直接接入到接入层设备，仍然保持树形结构。

目前的 P2P 文件传输软件会造成 P2P 流量大部分经过主干网络，越接近核心层，P2P 流量越大，占用带宽问题越严重，最终造成网络阻塞，影响了正常业务的开展。这种问题出现的根

源在于 P2P 软件在寻找服务节点时,认为各节点完全相同,没有考虑城域网的层次结构,在城域网内,近似随机地选择服务节点,导致大部分的流量经过主干网络。

如图 1 所示,假定城域网共有 n_1 个汇聚交换机,每个汇聚交换机下有 n_2 个接入交换机,每个接入交换机下有 n_3 个小区交换机,每个小区交换机下有 n_4 个楼宇交换机,每个楼宇交换机连接 n_5 台用户计算机,则整个城域网拥有的计算机台数为 $N=n_1*n_2*n_3*n_4*n_5$ 。以下把楼宇计算机接入小区交换机的链路简称楼宇上联小区链路,把小区计算机接入交换机的链路简称小区上联接入链路,把接入计算机接入汇聚交换机的链路简称接入上联汇聚链路,把汇聚计算机接入核心网的链路简称汇聚上联核心链路。

我们对常规 P2P 流量负载作以下分析:

当 P2P 节点要下载 m 字节的文件时,称之为请求 P2P 节点,假定有 $x\%$ 的计算机上有该资源,称之为资源 P2P 节点,即共有 $N*x\%$ 个资源 P2P 节点。按照统计规律,这些节点会均匀分布在整个网络中。

请求 P2P 节点随机选择资源 P2P 节点,将会从每个资源 P2P 节点获得 $m/(N*x\%)$ 字节的文件内容。

因此用户节点从自己楼宇交换机上相邻的计算机下载 $\frac{(n_5-1)*x\%*m}{N*x\%}$ 字节的文件内容,约等于 $m/(n_1*n_2*n_3*n_4)$ 字节,仅占整个文件的 $1/(n_1*n_2*n_3*n_4)$,剩余的内容需要经过楼宇上联小区链路从小区交换机下载,比例为 $\frac{n_1*n_2*n_3*n_4-1}{n_1*n_2*n_3*n_4}$;

一个用户节点能够从自己小区的计算机下载 $\frac{(n_4*n_5-1)*x\%*m}{N*x\%}$ 字节的文件内容,约等于 $m/(n_1*n_2*n_3)$ 字节,仅占整个文件的 $1/(n_1*n_2*n_3)$,剩余的内容需要经过小区上联接入链路从接入交换机下载,比例为 $\frac{n_1*n_2*n_3-1}{n_1*n_2*n_3}$;

一个用户节点能够从自己所在接入网内计算机下载 $\frac{(n_3*n_4*n_5-1)*x\%*m}{N*x\%}$ 字节的文件内容,约等于 $m/(n_1*n_2)$ 字节,仅占整个文件的 $1/(n_1*n_2)$,剩余的内容需要经过接入上联汇聚链路从汇聚交换机下载,比例为 $\frac{n_1*n_2-1}{n_1*n_2}$;

一个用户节点能够从自己所在汇聚网内计算机下载 $\frac{(n_2*n_3*n_4*n_5-1)*x\%*m}{N*x\%}$ 字节的文件内容,约等于 m/n_1 字节,仅占整个文件的 $1/n_1$,剩余的内容需要经过汇聚上联核心链路从核心网络下载,比例为 $\frac{n_1-1}{n_1}$ 。

当网络上有 y 台请求 P2P 节点,每个都下载 m 字节的内容,在网络上产生的流量情况为:

每个汇聚上联核心链路上从核心网下载的 P2P 流量为 $y*m*\frac{n_1-1}{n_1^2}$,在核心网络上总共产生

的流量为 $y * m * \frac{n1-1}{n1}$

每个接入上联汇聚链路路上的 P2P 下行流量为 $y * m * \frac{n1 * n2 - 1}{(n1 * n2)^2}$, 约等于 $y * m / (n1 * n2)$

每个小区上联接入链路路上的 P2P 下行流量为 $y * m * \frac{n1 * n2 * n3 - 1}{(n1 * n2 * n3)^2}$, 约等于 $y * m / (n1 * n2 * n3)$

每个楼宇上联小区链路路上的 P2P 下行流量为 $y * m * \frac{n1 * n2 * n3 * n4 - 1}{(n1 * n2 * n3 * n4)^2}$, 约等于 $y * m / (n1 * n2 * n3 * n4)$

在每个楼宇内计算机间产生的 P2P 流量为 $y * m / (n1 * n2 * n3 * n4)^2$

因此 P2P 产生的总流量 $y * m$ 、核心网流量、汇聚上联核心链路下行流量、接入上联汇聚链路下行流量、小区上联接入链路下行流量、楼宇上联小区链路下行流量和楼宇内计算机间的 P2P 流量的比例为 $1 : (n1-1)/n1 : (n1-1)/n1^2 : 1/(n1 * n2) : 1/(n1 * n2 * n3) : 1/(n1 * n2 * n3 * n4) : 1/(n1 * n2 * n3 * n4)^2$ 。在这个城域网模型中, 取 $n1=8, n2=8, n3=8, n4=8, n5=20$, 即城域网拥有 $N=81920$ 个计算机, 计算得到的各级网络 P2P 流量负载比例如图 2 和图 3 的“-◆-”折线, 其中图 2 纵坐标采用线性坐标, 图 3 采用对数坐标。显然 P2P 产生的网络流量绝大部分流经核心网, 这就是 P2P 造成网络阻塞的原因。而越接近网络的边缘, 流量负载越轻。

发明内容

目前在 P2P 技术的应用中, 占用带宽的主要是文件共享应用, 大量的文件在网络上反复传递, 造成了网络堵塞问题, 因为绝大部分的 P2P 流量会流经骨干网, 所以主要的堵塞是发生在骨干网上。为了解决 P2P 技术的应用中存在的阻塞问题, 本发明提出以下解决的技术方案。

本发明的优化城域网内 P2P 传输的方法是:

- a、当一个 P2P 节点需要下载某资源时, 以下称之为请求 P2P 节点;
- b、首先通过 P2P 索引服务器查询到拥有该资源的 P2P 节点列表, 这些节点称之为资源 P2P 节点;
- c、在资源 P2P 节点中寻找出与请求 P2P 节点在网络上最近邻的节点;
- d、资源通过网络上从最近邻的一个或多个资源 P2P 节点下载, 使 P2P 流量尽量远离核心网, 在网络边缘传输, 从而降低流经核心网和汇聚网等骨干网络的 P2P 流量。

(1) 在资源 P2P 节点中寻找与请求 P2P 节点在网络上最近邻的节点的步骤如下:

- a、首先寻找是否有与请求 P2P 节点接入同一个楼宇交换机的资源 P2P 节点, 如果有则是最近邻的资源 P2P 节点;
- b、否则寻找是否有与请求 P2P 节点接入同一个小区交换机或路由器的资源 P2P 节点, 如

果有则是最近邻的资源 P2P 节点；

c、否则寻找是否有与请求 P2P 节点接入同一个接入交换机或路由器的资源 P2P 节点，如果有则是最近邻的资源 P2P 节点；

d、否则寻找是否有与请求 P2P 节点接入同一个汇聚交换机或路由器的资源 P2P 节点，如果有则是最近邻的资源 P2P 节点；

e、否则所有的资源 P2P 节点都认为是最近邻的资源 P2P 节点。

(2) 查找最近邻资源 P2P 节点所进行的交换机或路由器编码的步骤如下：

a、汇聚交换机或路由器的编码 s1 为从 1 到 n1；

b、每个汇聚交换机或路由器下的接入交换机或路由器的编码 s2 为从 1 到 n2；

c、每个接入交换机或路由器下的小区交换机或路由器的编码 s3 为从 1 到 n3；

d、每个小区交换机下的楼宇交换机的编码 s4 为从 1 到 n4；

e、从核心网到每个计算机，所经过的汇聚交换机或路由器、接入交换机和路由器、小区交换机和路由器和楼宇交换机的编码构成了描述这台计算机拓扑网络位置的位置向量 $S=(s1,s2,s3,s4)$ ，同一个楼宇交换机下的计算机具有相同的位置向量。

(3) 查找最近邻资源 P2P 节点的计算方法如下：

a、定义 2 个位置向量 $S=(s1,s2,s3,s4)$ 和 $S'=(s'1,s'2,s'3,s'4)$ 的距离向量 D 为

$$D=(d1,d2,d3,d4)=(s1,s2,s3,s4)-(s'1,s'2,s'3,s'4)=(s1-s'1,s2-s'2,s3-s'3,s4-s'4)$$

其中各分量的减法定义为：

$$d_i = \begin{cases} 0, & s_i = s'_i \\ 1, & s_i \neq s'_i \end{cases}$$

b、定义 2 个距离向量 $D=(d1,d2,d3,d4)$ 和 $D'=(d'1,d'2,d'3,d'4)$ 的大小比较方法为

$$\begin{cases} D = D', & \text{当 } d1 = d'1, d2 = d'2, d3 = d'3, d4 = d'4 \\ D > D', & \text{当 } \begin{cases} d1 = 1, d'1 = 0 \\ d1 = d'1, d2 = 1, d'2 = 0 \\ d1 = d'1, d2 = d'2, d3 = 1, d'3 = 0 \\ d1 = d'1, d2 = d'2, d3 = d'3, d4 = 1, d'4 = 0 \end{cases} \end{cases}$$

距离向量小，表示在网络上更近邻。

(4) 在城域网中增设的拓扑服务器用于存储各子网与位置向量的对应关系，存储内容为各楼宇交换机的子网地址和掩码，以及位置向量，拓扑服务器的工作步骤如下：

a、根据 IP 地址，与存储的各子网进行比对，寻找出所属子网，从而得到该 IP 地址计算机的位置向量；

b、根据一个请求 P2P 节点的 IP 地址和多个资源 P2P 节点 IP 地址，计算出请求 P2P 节点与各资源 P2P 节点的距离向量，并进行距离向量的比较，找出距离向量最小的资源 P2P 节点。

c、所存储的数据反映城域网当前的拓扑结构，是计算最近邻资源节点的判断依据。

(5) P2P 索引服务器查询出资源 P2P 节点后，由拓扑服务器计算出最近邻节点，具体计算步骤如下：

a、P2P 索引服务器接收请求 P2P 节点的资源查询请求，查询出拥有该资源的资源 P2P 节点；

b、把请求 P2P 节点的 IP 地址和查询出的各资源 P2P 节点 IP 地址发送给拓扑服务器；或者把资源 P2P 节点列表发送给请求 P2P 节点，请求 P2P 节点把自己的 IP 地址和各资源 P2P 节点的 IP 地址发送给拓扑服务器；

c、拓扑服务器计算出与请求 P2P 节点最近邻的一个或多个资源 P2P 节点，然后返回给索引服务器，索引服务器把最近邻的资源 P2P 节点列表发送给请求 P2P 节点；

e、该请求 P2P 节点与最近连的资源 P2P 节点进行连接和下载。

(6) 请求 P2P 节点获得的所有资源 P2P 节点列表中包含各节点的位置向量，请求 P2P 节点直接计算最近邻的资源 P2P 节点的具体步骤如下

a、计算机上的 P2P 软件启动时，首先向拓扑服务器请求自己的位置向量；

b、拓扑服务器根据计算机的 IP 地址，在拓扑数据表中查询所属的子网，从而得到该计算机的位置向量，并返回给该计算机；

c、各 P2P 节点计算机向 P2P 的索引服务器上报告自己所拥有的资源时，同时上报自己的位置向量；

d、当请求 P2P 节点计算机要下载某资源时，首先查询得到拥有该资源的节点列表，其中包含各资源节点的位置向量；

e、请求 P2P 节点分别把自己的位置向量与各资源节点的位置向量相减得到相应的距离向量，从中选择距离最小的一个或多个资源节点，进行资源的下载。

附图说明

图 1 是城域网模型图；

图 2 是纵坐标采用线性坐标的城域网 P2P 流量分布图；

图 3 是采用对数坐标的城域网 P2P 流量分布图；

图 4 是给交换机或路由器编码后的城域网模型图；

图 5 是拓扑服务器的数据存储结构示意图。

实施方式

优化 P2P 的实现方案

首先对城域网的汇聚交换机、接入交换机、小区交换机和楼宇交换机进行编码,如图 4: 汇聚交换机的编码 s_1 为从 1 到 n_1 , 每个汇聚交换机下的接入交换机的编码 s_2 为从 1 到 n_2 , 每个接入交换机下的小区交换机的编码 s_3 为从 1 到 n_3 , 每个小区交换机下的楼宇交换机的编码 s_4 为从 1 到 n_4 。这样每个子网或楼宇交换机都可以用一个向量表示 $S=(s_1,s_2,s_3,s_4)$, 也是这个交换机下的各计算机共同的位置向量。显然位置向量相同的 2 台计算机必定接入同一台楼宇交换机。

当请求 P2P 节点计算机可以从多个资源计算机下载时, 首先用自己的位置向量 $S=(s_1,s_2,s_3,s_4)$ 分别减去各资源 P2P 节点的位置向量 $S'=(s'_1,s'_2,s'_3,s'_4)$, 得到与各资源 P2P 节点的距离向量 $D=(d_1,d_2,d_3,d_4)$

$$(d_1,d_2,d_3,d_4)=(s_1,s_2,s_3,s_4)-(s'_1,s'_2,s'_3,s'_4)=(s_1-s'_1,s_2-s'_2,s_3-s'_3,s_4-s'_4)$$

其中各分量的减法定义为:

$$d_i = \begin{cases} 0, & s_i = s'_i \\ 1, & s_i \neq s'_i \end{cases}$$

并且定义距离向量 $D=(d_1,d_2,d_3,d_4)$ 和 $D'=(d'_1,d'_2,d'_3,d'_4)$ 比较大小的方法

$$\begin{cases} D = D', & \text{当 } d_1 = d'_1, d_2 = d'_2, d_3 = d'_3, d_4 = d'_4 \\ D > D', & \text{当 } \begin{cases} d_1 = 1, d'_1 = 0 \\ d_1 = d'_1, d_2 = 1, d'_2 = 0 \\ d_1 = d'_1, d_2 = d'_2, d_3 = 1, d'_3 = 0 \\ d_1 = d'_1, d_2 = d'_2, d_3 = d'_3, d_4 = 1, d'_4 = 0 \end{cases} \end{cases}$$

按照上述算法, 比较与各资源 P2P 节点的距离向量, 找出距离最小的资源 P2P 节点, 从其进行下载, 就可以达到优化 P2P 传输的结果。

在城域网中设置一台服务器, 存储着整个城域网拓扑结构数据, 数据结构如图 5, 存储内容是各子网与位置向量的对应。该服务器用于根据计算机的 IP 地址计算出位置向量, 以下称之为拓扑服务器。

具体应用中, 进行距离比较计算可以在 P2P 节点或者拓扑服务器上进行。

在 P2P 节点进行距离比较的工作过程:

- A、计算机上的 P2P 软件启动时, 首先向拓扑服务器请求自己的位置向量;
- B、拓扑服务器根据计算机的 IP 地址, 在拓扑数据表中查询所属的子网, 从而得到该计算机的位置向量, 并返回给该计算机;
- C、计算机上的 P2P 软件启动时, 首先向拓扑服务器请求自己的位置向量;
- D、各 P2P 节点计算机向 P2P 的索引服务器上报告自己所拥有的资源时, 同时上报自己的位置向量;

E、当请求 P2P 节点计算机要下载某资源时，首先查询得到拥有该资源的节点列表，其中包含各资源节点的位置向量；

F、请求 P2P 节点分别把自己的位置向量与各资源节点的位置向量相减得到相应的距离向量，从中选择距离最小的一个或多个资源节点，进行资源的下载。

拓扑服务器进行距离比较的工作过程：

A、当请求 P2P 节点要下载某资源时，首先向 P2P 索引服务器发送该资源的查询请求；

B、P2P 索引服务器接收请求 P2P 节点的资源查询请求，查询出拥有该资源的资源 P2P 节点；

C、把请求 P2P 节点的 IP 地址和查询出的各资源 P2P 节点 IP 地址发送给拓扑服务器；或者把资源 P2P 节点列表发送给请求 P2P 节点，请求 P2P 节点把自己的 IP 地址和各资源 P2P 节点的 IP 地址发送给拓扑服务器；

D、按照权利要求 4 所述，拓扑服务器计算出与请求 P2P 节点最近邻的一个或多个资源 P2P 节点。然后返回给索引服务器，索引服务器把最近邻的资源 P2P 节点列表发送给请求 P2P 节点；或者直接发送给请求 P2P 节点；

E、该请求 P2P 节点与最近连的资源 P2P 节点进行连接和下载。

实施例

优化的 P2P 访问策略

如果能够将 P2P 的流量尽量限制在网络边缘，则能够有效降低主干网络的带宽压力。因此对 P2P 访问模式进行改造，采用以下新的下载策略：

当一个节点要下载某一资源时，当所在楼宇交换机下的其它计算机拥有该资源时，只从本楼宇交换机下的计算机下载；否则，如果本小区网络内的其它计算机拥有该资源时，只从本小区网络内的计算机下载；否则，如果本接入网络内的其它计算机拥有该资源时，只从本接入网络内的计算机下载；否则，如果本汇聚网络内的其它计算机拥有该资源时，只从本汇聚网络内的计算机下载；如果本汇聚网络内的计算机都没有该资源，则通过核心网络，到别的汇聚网络进行下载。这样的 P2P 下载策略能够尽量把 P2P 流量推向网络边缘，降低主干网络的负载。下面在统计规律下分析采用策略的 P2P 的流量分布特性：

一个请求 P2P 节点要下载 m 字节的文件，在城域网上的 N 台计算机中，其中拥有该资源的资源 P2P 节点比例为 $x\%$ 。

如果在本楼宇交换机下存在一个或者更多的资源 P2P 节点，下载流量将只在本楼宇交换机内部发生，在楼宇上联小区链路上不产生下行流量；如果本楼宇交换机下一个资源节点也没有，全部 m 字节都要通过楼宇上联小区链路下载。本楼宇交换机下不存在资源节点的几率 p_5 为

$$p5 = \left(\frac{N - n5}{N} \right)^{N \cdot x\%} = \left(\frac{n1 \cdot n2 \cdot n3 \cdot n4 - 1}{n1 \cdot n2 \cdot n3 \cdot n4} \right)^{N \cdot x\%}$$

则在统计情况下在楼宇上联小区链路上产生的 P2P 下行流量为 $p5 \cdot m$ 字节，而从本楼宇其它计算机下载的流量为 $(1-p5) \cdot m$ 字节。

同样在本小区网络内不存在资源 P2P 节点的几率 $p4$ 为

$$p4 = \left(\frac{N - n4 \cdot n5}{N} \right)^{N \cdot x\%} = \left(\frac{n1 \cdot n2 \cdot n3 - 1}{n1 \cdot n2 \cdot n3} \right)^{N \cdot x\%}$$

从而在与小区上联接入链路上产生的 P2P 下载流量为 $p4 \cdot m$ 字节。

在本接入网络内不存在资源 P2P 节点的几率 $p3$ 为

$$p3 = \left(\frac{N - n3 \cdot n4 \cdot n5}{N} \right)^{N \cdot x\%} = \left(\frac{n1 \cdot n2 - 1}{n1 \cdot n2} \right)^{N \cdot x\%}$$

从而接入上联汇聚链路上产生的 P2P 下载流量为 $p3 \cdot m$ 字节。

在本汇聚网络内不存在资源 P2P 节点的几率 $p2$ 为

$$p2 = \left(\frac{N - n2 \cdot n3 \cdot n4 \cdot n5}{N} \right)^{N \cdot x\%} = \left(\frac{n1 - 1}{n1} \right)^{N \cdot x\%}$$

从而在汇聚上联核心链路上产生的 P2P 下载流量为 $p2 \cdot m$ 字节。

如果城域网中有均匀分布的 y 个请求 P2P 节点要下载该资源，则在每个楼宇交换机内部的 P2P 流量为 $\frac{y \cdot (1-p5) \cdot m}{n1 \cdot n2 \cdot n3 \cdot n4}$ 字节，在每个楼宇上联小区链路上产生的 P2P 下行流量为 $\frac{y \cdot p5 \cdot m}{n1 \cdot n2 \cdot n3 \cdot n4}$ 字节，在每个小区上联接入链路上产生的 P2P 下行流量为 $\frac{y \cdot p4 \cdot m}{n1 \cdot n2 \cdot n3}$ 字节，在每个接入上联汇聚链路上产生的 P2P 下行流量为 $\frac{y \cdot p3 \cdot m}{n1 \cdot n2}$ 字节，在每个汇聚上联核心链路上产生的 P2P 下行流量为 $\frac{y \cdot p2 \cdot m}{n1}$ 字节，所有流经核心的 P2P 流量为 $y \cdot p2 \cdot m$ 字节。

因此在这种优化策略下，因此 P2P 产生的总流量 $y \cdot m$ 、核心网流量、汇聚上联核心链路下行流量、接入上联汇聚链路下行流量、小区上联接入链路下行流量、楼宇上联小区链路下行流量和楼宇内计算机间的 P2P 流量的比例为 $1 : p2 : p2/n1 : p3/(n1 \cdot n2) : p4/(n1 \cdot n2 \cdot n3) : p5/(n1 \cdot n2 \cdot n3 \cdot n4) : (1-p5)/(n1 \cdot n2 \cdot n3 \cdot n4)$ 。按照前面的城域网模型参数，分别计算在 x 为 0.1%、0.05%、0.02%和 0.01%时 P2P 产生的流量负载，结果如图 2 和图 3 的实折线。显然在接入交换机以内的核心网络上的 P2P 负载明显降低，实现方法是把这些流量转移到了网络的最边缘——楼宇交换机内部。例如在拥有资源的计算机占整个网络的万分之一($x=0.01\%$)时，核心网负载约有常规 P2P 模式下的 1/3，而楼宇内的 P2P 流量约提高了一个数量级，提高了用户的传输速度。

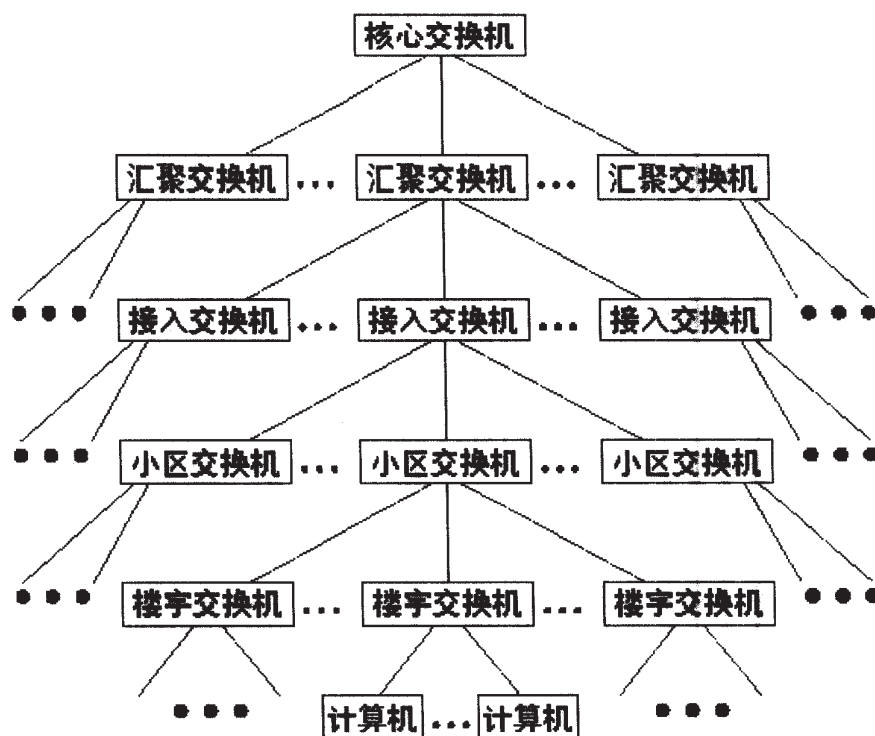


图 1

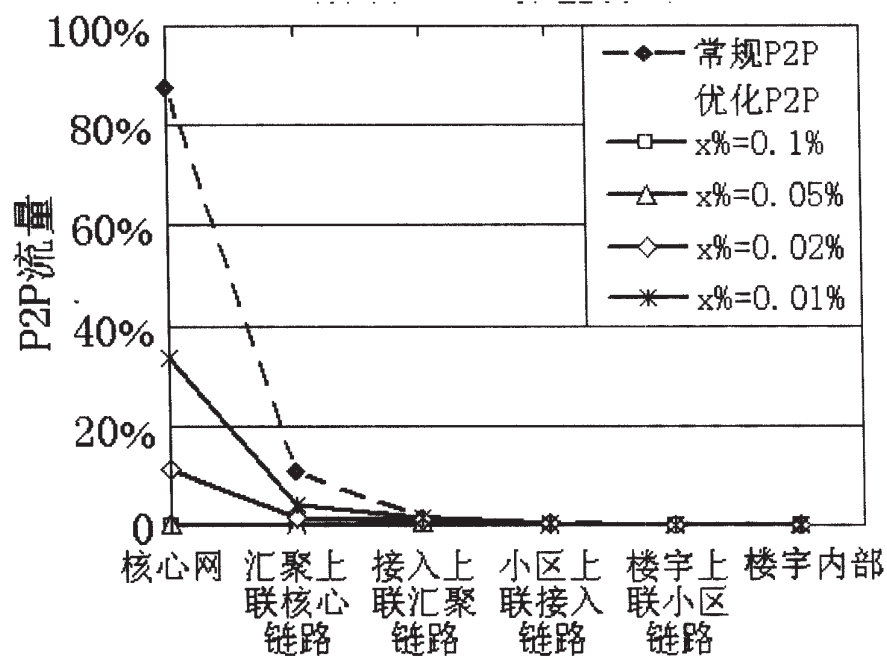


图 2

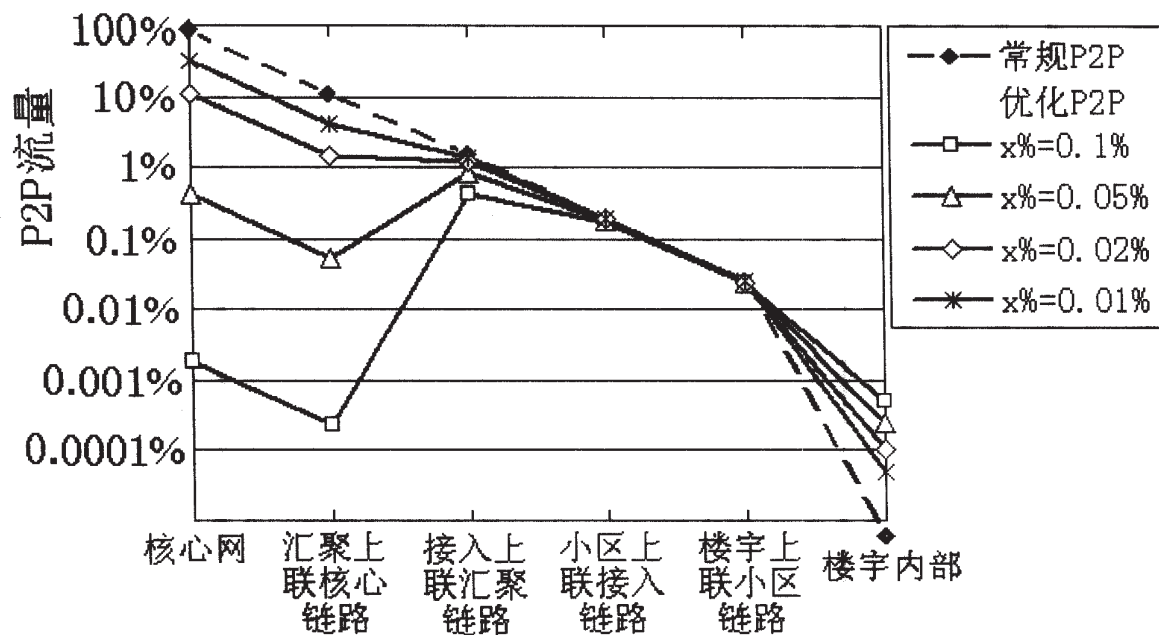


图 3

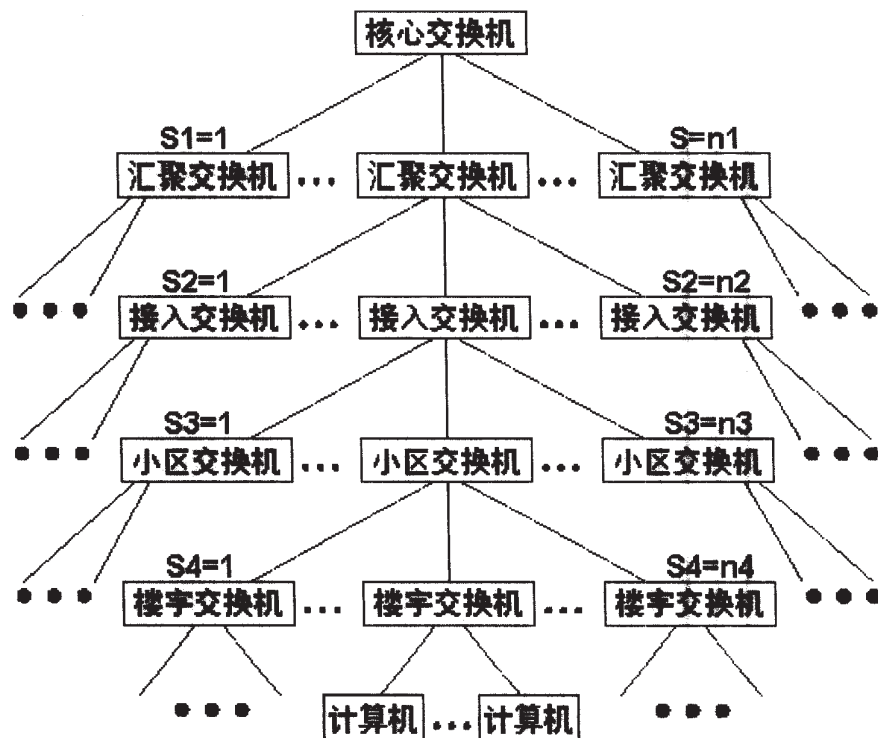


图 4

子网 IP 地址	汇聚交换机 编码(s1)	接入交换机 编码(s2)	小区交换机 编码(s3)	楼宇交换机 编码(s4)
...	...	...	...	...
10.30.11.65 / 26	2	3	5	4
10.30.11.129 / 26	2	3	5	5
...	...	...	...	...

图 5