

Enhancement the Efficiency of Work with P2P System

Abhilisha Pandurang Bhalke

Post Graduate Scholar, Department of Computer Engineering, Csmss's Chh. Shahu College of Engineering Kanchanwadi, Paithan Road, Aurangabad, Maharashtra, India

To Cite this Article

Abhilisha Pandurang Bhalke, "Enhancement the Efficiency of Work with P2P System", *International Journal for Modern Trends in Science and Technology*, 6(9): 68-72, 2020.

Article Info

Received on 08-August-2020, Revised on 18-August-2020, Accepted on 27-August-2020, Published on 04-September-2020.

ABSTRACT

The P2P system should be used Proximity information to minimize the load of file request and improve the efficiency of the work. Clustering peers for their physical proximity can also rise the performance of the request file. However, very few currently work in a peer group based on demands as peers on physical proximity. Although structured P2P provides more efficient files requests than unstructured P2P, it is difficult to apply because of their strictly defined topology. In this work, we intending to introduce a system for exchange a P2P file for proximity and level of interest based on structured P2P nodes that form physically block in the cluster and other groups physically close and nodes of public interest in sub-cluster based on the hierarchical topology. Querying an effective file is important for the overall P2P file exchange performance. Clustering peers from their common interests can significantly enhance the efficiency of the request file PAIS use an intelligent file replication algorithm to further rise the efficiency of the request file. Create a copy file that is often requested by a group of physically close nodes in their position. In addition, PAIS improves the search for files within the intra-system sub-cluster through various approaches. First, it further classifies interest in the sub-cluster to a number of subsections of interests and groups with common interest-free sub nodes in the group for file sharing. Secondly PAIS creates an over for each group that connects nodes of less node capacity to a higher throughput for the distributed node overload prevention request file. Third, in order to reduce the search for late files, PAIS uses a set of proactive information so that applicant can file knowledge if its requested file is in the neighboring nodes. Fourth, reduce the overhead of collecting information about files using the PAIS, collection of file information based on the Bloom Filter and the corresponding search for files distributed. Fifth, in order to improve the efficiency of file sharing, PAIS ranks the results with a blob of filters in order. Sixth, while the newly visited file is usually re-visited approach, based on the Bloom filter is improved only through the management of new information flowering filter is added to reduce the delay of file search. The experimental result of the Real-world Planet Lab Experiment shows that PAIS significantly reduces overhead and improves the efficiency of scrolling and without sharing files. In addition, the experimental results show high efficiency within the sub-research cluster of file approaches to improve file search efficiency.

KEYWORDS: P2P networks, file sharing system, proximity awareness, file replication, Bloom filter.

I. INTRODUCTION

The last few years, the enormous popularity of the Internet has caused considerable incentive for P2P file-sharing systems. For example, Bit Torrent is about 35 % of all Internet traffic. There are two classes of P2P system unstructured and structured. Unstructured network P2P networks, such as Gnutella and Free net, do not associate responsible for specific node data. Nodes are connected and leave the network according to some loose rules. Currently, unstructured P2P-networks file query method is based on a river when the request applies to all site neighbors, or a random walk, where a request is sent to randomly selected neighbors as long as the file is not found.

However, occasional floods and hikers cannot guarantee the location of the data. Structured P2P network, i.e. Distributed Hash Table (DHT), able to overcome the disadvantages in their features greater efficiency, scalability, and location data determined. They are strictly controlled topology, and their precise data and search algorithms determined precisely based on the DHT data structure and constant hash function. The key responsibilities for the key can always be found, even if the system is in a constant state of change. Most DHT requires $O(\log n)$ transitions for each search query to $O(\log n)$ neighbors by the node, where n - the number of nodes in the system. A key criterion for assessing the P2P file sharing system is the efficiency of its position. Numerous methods have been proposed to increase this efficiency. One method uses supernumerary topology consisting of super nodes with quick links and regular nodes with slower connections. A super node is connected to other super nodes and some regular drives, but a common node is connected to the super node. In this topology, the sites of the super peer topology faster and then create a safer and more stable base.

This allows you to have several messages about the slow spine and, therefore, provides higher scalability. In super-peer-to-peer networks, they occupy the second place between centralized P2P networks and are completely symmetric and have the potential to combine the benefits of centralized and distributed research. Another class of methods to improve the efficiency of file allocation is through the structure of information proximity. A contactless logical abstraction derived from a P2P system does not necessarily correspond to physical proximity information in reality. The

shortest path in accordance with the routing protocol is not necessarily the shortest physical path. This disagreement becomes a serious obstacle to the implementation and optimization of the performance of P2P file-sharing system. The P2P system should use Proximity information to reduce the load of file requests and improve the efficiency of the work. In other words, assigning or replicating a file to a node physically closer to the applicant can greatly help the applicant to recover the file so efficiently. Clustering is known about proximity can be used for physical groups of immediate physicians to effectively improve work efficiency.

The third class of methods for improve the efficiency of file allocation for a group of nodes with similar interests which reduce the latency of the file location. Although there are many super-peer topology proposals based on proximity and based on common interests with different possibilities, several methods can group a peer on the basis of proximity and interest. In addition, most of these methods are in unstructured P2P systems that do not have a strict policy for building topology. They cannot be directly applied to common DHT, despite their higher efficiency of file allocation.

II. LITERATURE WORK

We will discuss the most important works related to PAIS in three groups: super-peers of the topology of awareness of proximity and sharing of files of interests. Super-peers topology. Interactive (10) and Morpheus (20) using the topology of super-peers. Super-peer network (8) for efficient and scalable file consistency assurance in structured P2P systems. Our previous work built a network of super-peers for load balancing. (9) Garbacki et al. (21) proposed super-peer architecture for a self-organizing network that solves four tasks completely decentralized: such as peer clients are associated with super-peers such as super-peers finding files since the load is balanced between super -supers, and how the system copes with node failures. Methods of informing proximity to use topology information in the routing of a P2P overlay include the geographical layout, routing, and proximity of the selected neighbor series. The geographical method of linking the trace space is the logical overlap identifier in the physical network so that the neighboring nodes in the ID space are also closest to the physical network. It is used in CAN topologically aware (11). In the proximity routing to

the method, a logical overlay is constructed without taking into account the underlying physical topology.

The category of interest-based file-sharing networks, called schema-based networks. They use explicit schema to describe content of a pair, based on semantic descriptions, and allow aggregate and integrate data from distributed data source.

Liu et al. Online storage system offered with peer support .Work on using Bloom filter technology to search for files. Despite the efforts to effectively position files in P2P systems, there is a several works that combine root topology using clustering methods with interest and intimacy. Moreover, it is difficult to implement in DHT due to topology policy and well-defined data distribution. This document describes how to resolve the PAIS problem, using the DHT hierarchical structure.

III. PROBLEM STATEMENT

3.1 Existing Model -The basic test for the P2P file sharing system is its file location efficiency. In order to rise these numerous methods, efficiency has been proposed. One method uses the super-peer topology, which consists of super nodes with faster connections and regular nodes with slow connections. Super node connects to other nodes and the super nodes some regular and regular node connects to the super node.

In this super-peer topology, the nodes in the center of the networks are faster and give a more reliable and stable cork. This allows you to have several messages about the slow spine and, therefore, provides higher scalability. In super-P2P networks, they occupy the second place between centralized P2P networks and are complete symmetric and have the potential to combine the benefits of centralize and distributed research.

Another class of methods to improve the efficiency of file allocation is through the structure of information proximity. The third class of methods for improves the efficiency of the location of files for a group of nodes with similar interests, which reduced the delay in the location of files.

3.1.1 Disadvantages Of Existing System - Despite the fact that various types of avant-garde topologies have been proposed based on proximity and interest; different techniques can be grouped peers both in proximity and interest. Also, most of these methods are in unstructured P2P systems, which do not have a strict policy for topology

building. They cannot be applied directly to the general DHT, despite their file allocation being more efficient.

3.2 Proposed System - This article is system for exchange a P2P file with the dimension of proximity and a cluster of interests (PAIS) over a structured P2P system. Form physically close nodes in the cluster and more physically closed groups and nodes of common interest in the subgroup. Also, it places files with the same interests together and makes them available use the DHT routing lookup () function. More importantly, it retains all the advantages of P2P DHT unstructured. Based on the DHT research policy, rather than on the transfer, the construction of PAIS consumes much less cost to display the node of the cluster and display the cluster to influence subgroups. PAIS uses an intelligent file replication algorithm to further improve the efficiency of file search. Create a copy file that is often request by a group of physically close nodes in their position. In addition, PAIS improves sub-cluster files in various approaches.

First, it further classifies interest in the sub-cluster to a number of subsections of interests and groups with common interest-free sub nodes in the group for file sharing. Secondly, PAIS creates an overlay for each group that connects nodes of less node capacity to a higher throughput for the distributed node overload prevention request file. Third, in order to reduce the search for late files, PAIS uses a set of proactive information so that the applicant can file knowledge if its requested file is in the neighboring nodes. Fourth, to reduce the overhead of collecting information about the file, PAIS uses the collection of information about files based on the flowering of filters and the corresponding search for files distributed. Fifth, in order to improve the efficiency of file sharing, PAIS ranks the results with a blob of filters in order. Sixth, while the newly visited file is usually re-visited approach, based on the Bloom filter is improved only through the management of new information flowering filter is added to reduce the delay of file search.

3.2.1 Advantages of Proposed System - The methods proposed in this articles can benefit many modern applications, such as content delivery networks, video system for requesting P2P and exchanging data in online social networks. Let's introduce a detailed PAIS drawing. And suitable for a file sharing system where files can be divided into

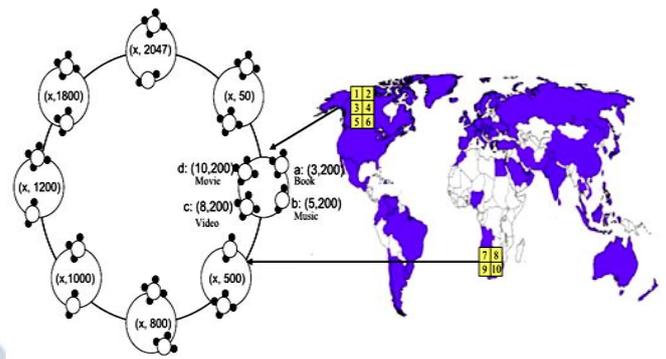
different interests and any interest can be divided into several subgroups of interests. It brings together peers based on interest and close, using the hierarchic structure of structured P2P. PAIS uses the algorithm for intelligent file replication, which reproduces the file, often requires physically close nodes near their physical location, and improve the efficiency of file search. PAIS improve the efficiency of file search and the proximity between nodes of common interest from variety of approaches.

IV. METHODOLOGY

4.1 PAIS: A proximity-aware interest-clustered p2p file sharing system. - In our last work, we studied the traces of the Bit Torrent user's activity for analyzing the behavior of the file-sharing user. We found that the restoration of the long-distance file exists. Thus, we can group node physically close to the cluster in order to improve the efficiency of file share. In addition, colleagues tend to see files in a few percent. Thus, we can group the nodes that share an interest in the subgroup. Finally, the popular file in each interest are distributed among peers globally distributed.

Thus, we can use file replication between sites for popular file files and use the system at the level of researching unpopular files. We introduce the detailed design of PAIS below. It is suitable for a file sharing system in which files can be classified into several percentages and any percentages can be classified into several sub-interest.

4.2 PAIS Structure -PAIS was designed based on the Cycloid-structured P2P network. Cycloid - an effective taxation of a constant precision level with $n = d \cdot 2d$, where d - its size. It reaches the temporal complexity of $O(d)$ for each search query with $O(1)$ neighbors per node. Each Cycloid node is represented by a pair of indices $(k, a_{d-1}a_{d-2}...a_0)$, where k - cyclic code, and $(a_{d-1}a_{d-2}...a_0)$ are cubic index. The cyclic index is an integer from 0 to $d - 1$, and the cubic index - a binary number between 0 and $2d - 1$. Knots of the same cubic index at the end of their cyclic index $d \bmod a$ short loop, which we call a cluster.



4.3 PAIS Construction and Maintenance -The analysis method can be used to denote the proximity of the nodes in the network used by the indices. Landmark clustering was rarely adopted to create information proximity. It is based on the intuition that nodes close to each other probably have similar distances to some nodes of selected reference points. Suppose that there are nodes of m points that are randomly distributed on the Internet.

V. EXPERIMENTAL RESULTS

We implemented a PAIS Planet Lab prototype, current test bench distribution, to measure PAIS performance over other P2P file sharing systems. We set the experimental environment in accordance with results of Bit Torrent tracking studies. We randomly selected 350 Planet Lab nodes worldwide. Among these sites, we randomly selected thirty sites as backward nodes to calculate the number of Hilbert nodes. We united all the nodes in 169 different places, in accordance with the similarity of their Hilbert numbers.

We used 56,076 Bit Torrent track files. The amount of interest for the system was set to 20, so we also set the Cycloid DHT size to 20. In the experiments, we simulated 100,000 peers by default. Each peer was randomly assigned to cluster positions of all clusters 169 and the further arbitrarily assigned Planet Lab site to this location. According to the peers' requests, focus on about 20 percent of all their interests. So, we chose randomly four of interest (20 percent to 20 percent) for each partner, his interests.

The files are sub clustered randomly with the interests of the most common files 160 places, and then randomly assigned to nodes in the subclass. Eighty percent of all requests addressed to the applicant file with the owners in one place, including 70 percent of his requests in the interests of the applicant. According to research, 80 percent of all requests from a partner focused on his

interests, and each of the other requests is in an arbitrary interest outside his own interests. Percent Query means a request for a randomly selected file in this regard. We also allow each file copied to another peer node in a different location, to test the performance of a file search, proximity support.

VI. CONCLUSION

In recent years, to improve efficiency of file allocation in P2P systems, we have proposed super-P2P networks and clusters of interest in super-peer-to-peer cluster networks. While both strategies will improve the performance of P2P systems, some work together with peers relying both on mutual interest and physical proximity. In addition, it is more difficult to implement in structured P2P systems because of their strictly defined topologies, even if they have a high efficiency of the location of the files compared to unstructured P2P networks.

In this article, we introduce a system for exchanging a P2P file, known and compatible with a series, based on structured P2P. Peers group is based on the interest of both on the proximity of the advantage of the hierarchical structure of structured P2P. PAIS uses intelligent file replication algorithm that reproduces a file often requested by physically close nodes near their physical location to improving the efficiency of file search. Finally, PAIS improves the efficiency of finding files between proximity nodes and common interests through a series of approaches. Tracks on the experimental results of Planet Lab demonstrate the effectiveness of PAIS than other P2P file sharing systems. Dramatically reduces overhead and leads to significant improvement in the efficiency of the location of files also in node dynamism In addition, experimental results show the effectiveness of approaches to improving the efficiency of file search in between proximity nodes and common interests.

REFERENCES

- [1] BitTorrent. (2013) [Online]. Available: <http://www.bittorrent.com/>
- [2] Gnutella home page. (2003) [Online]. Available: <http://www.gnutella.com>
- [3] I. Clarke, O. Sandberg, B. Wiley, and T. W. Hong, "Freenet: A distributed anonymous information storage and retrieval system," in Proc. Int. Workshop Des. Issues Anonymity Observability, 2001, pp. 46–66.
- [4] I. Stoica, R. Morris, D. Liben-Nowell, D. R. Karger, M. F. Kaashoek, F. Dabek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup protocol for internet applications," *IEEE/ACM Trans. Netw.*, vol. 11, no. 1, pp. 17–32, Feb. 2003.
- [5] A. Rowstron and P. Druschel, "Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems," in Proc. IFIP/ACM Int. Conf. Distrib. Syst. Platforms Heidelberg, 2001, pp. 329–350.
- [6] B. Y. Zhao, L. Huang, J. Stribling, S. C. Rhea, A. D. Joseph, and J. Kubiatowicz, "Tapestry: A resilient global-scale overlay for service deployment," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 1, pp. 41–53, 2004.
- [7] H. Shen, C. Xu, and G. Chen, "Cycloid: A scalable constant-degree P2P overlay network," *Perform. Eval.*, vol. 63, pp. 195–216, 2006.
- [8] Z. Li, G. Xie, and Z. Li, "Efficient and scalable consistency maintenance for heterogeneous peer-to-peer systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 12, pp. 1695–1708, Dec. 2008.
- [9] H. Shen and C.-Z. Xu, "Hash-based proximity clustering for efficient load balancing in heterogeneous DHT networks," *J. Parallel Distrib. Comput.*, vol. 68, pp. 686–702, 2008.
- [10] FastTrack. (2003) [Online]. Available: http://www.fasttrack.nu/index_int.html
- [11] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker, "Topologically-aware overlay construction and server selection," in Proc. IEEE INFOCOM, 2002, pp. 1190–1199.
- [12] M. Waldvogel and R. Rinaldi, "Efficient topology-aware overlay network," in Proc. ACM Workshop Hot Topics Netw., 2002, pp. 101–106.
- [13] Y. Zhu and H. Shen, "An efficient and scalable framework for content-based publish/subscribe systems," *Peer-to-Peer Netw. Appl.*, vol. 1, pp. 3–17, 2008.
- [14] C. Hang and K. C. Sia, "Peer clustering and firework query model," in Proc. Int. World Wide Web Conf., 2002.
- [15] A. Crespo and H. Garcia-Molina, "Routing indices for P2P systems," in Proc. 22nd Int. Conf. Distrib. Comput. Syst., 2002, pp. 23–32.
- [16] W. Nejdl, M. Wolpers, W. Siberski, C. Schmitz, M. Schlosser, I. Brunkhorst, and A. Loeser, "Super-peer-based routing and clustering strategies for RDF-based peer-to-peer networks," *J. Web Semantics*, vol. 1, no. 2, pp. 137–240, 2004.
- [17] P. A. Bernstein, F. Giunchiglia, A. Kementsietsidis, J. Mylopoulos, L. Serafini, and I. Zaihrayeu, "Data management for peer-to-peer computing: A vision," in Proc. 5th Int. Workshop World Wide Web Databases, 2002, pp. 89–94.
- [18] A. Y. Halevy, Z. G. Ives, P. Mork, and I. Tatarinov, "Piazza: Data management infrastructure for semantic web applications," in Proc. 12th Int. Conf. World Wide Web, 2003, pp. 556–567.
- [19] K. Aberer, P. Cudr_e-Mauroux, and M. Hauswirth, "The chatty web: Emergent semantics through gossiping," in Proc. 12th Int. Conf. World Wide Web, 2003, pp. 197–206.
- [20] P. Garbacki, D. H. J. Epema, and M. van Steen, "Optimizing peer relationships in a super-peer network," in Proc. Int. Conf. Distrib. Comput. Syst., 2007, p. 31