International Journal of Science Engineering and Advance Technology

Enhanced Document Search and Sharing Tool in Organized P2p Frameworks

[1]Mohammad Aslamsultana,[2]SSSN Usha Devi N
[1,2]Dept. of CSE, University College of Engineering
JNTU Kakinada (AUTONOMOUS), AP, India

ABSTRACT:

In internet p2p file sharing system generating more traffic. in this system file querying is important functionality which indicates the performance of p2p system .To improve file query performance cluster the common interested peers based on physical proximity .Existing methods are dedicated to only unstructured p2p systems and they don't have strict policy for topology construction which decreases the file location efficiency. In this project proposing a proximity aware interest –clustered p2p file sharing system implemented in structured p2p file system. It forms a cluster based on node proximity as well as groups the nodes which having common interest into sub-cluster. A novel lookup function named as DHT and file replication algorithm which supports efficient file lookup and access. To reduce overhead and file searching delay the file querying may become inefficient due to the sub-interest supernode overload or failure. Thus, though the sub-interest based file querying improves querying efficiency, it is still not sufficiently scalable when there are a very large number of nodes in a sub-interest group. We then propose a distributed intra-sub-cluster file querying method in order to further improve the file querying efficiency.

KEYWORDS:Proximity awareness, file replication, Bloom filter

1 INTRODUCTION:

A proximity-aware and interest-clustered P2P file sharing System (PAIS) on an organized P2P framework. It frames physically-close nodes into a cluster and further gatherings physically-close and normal interest nodes into a sub-cluster[3]. It likewise puts documents with the same interests together and make them available through the DHT Lookup() directing capacity. All the more importantly, it keeps all preferences of DHTs over unstructured P2Ps. Depending on DHT lookup arrangement as opposed to TV, the PAIS development consumes considerably less cost in mapping nodes to clusters and mapping groups to intrigue sub-bunches. PAIS utilizes a wise document replication algorithm to further upgrade record lookup productivity. It makes copies of documents that are as often as possible asked for by a group of physically close nodes in their area[13][14].

2 LITERATURE SURVEY:

2.1 the lack of global knowledge makes scheduling difficult in P2P systems. Present the new developments concerning locality awareness as well as co-allocation strategies available in the latest release of P2P-MPI. i) The spread strategy tries to map processes on hosts so as to maximize the total amount of available memory while maintaining locality of processes as a secondary objective. ii) The concentrate strategy tries to maximize locality between processes by using as many cores as hosts offer

2.2 Existing data storage systems based on the hierarchical directory-tree organization do not meet the scalability and functionality requirements for exponentially growing data sets and increasingly complex metadata queries in large scale. The decentralized design of SmartStore can improve system scalability and reduce query latency for complex queries.

3 PROBLEM DEFINITION

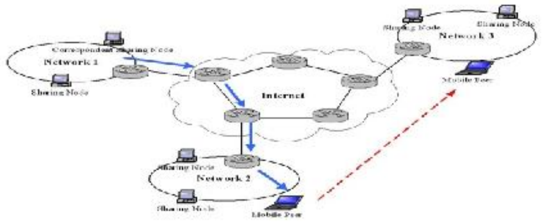
Flooding and irregular walkers can't promise information location. Proximity based and premium based super-peer topologies techniques are on unstructured P2P frameworks that have no strict approach for topology development. They can't be straightforwardly connected to general DHTs notwithstanding their higher document area effectiveness.

4 PROPOSED APPROACH

More altogether, it keeps all inclinations of DHTs over unstructured P2Ps. Contingent upon DHT lookup approach rather than TV, the PAIS advancement exhausts significantly less cost in mapping nodes to groups and mapping clusters to interest sub-clusters.

PAIS uses a wise document replication algorithm to advance enhance record lookup capability.

5 SYSTEM ARCHITECTURE:



6 PROPOSED METHODOLOGY:

6.1 PAIS STRUCTURE:

We can cluster physically close nodes into a group to upgrade record sharing productivity. Likewise, peers tend to visit records in a couple interests. Subsequently, we can encourage group nodes that share an enthusiasm into a sub-cluster. At long last, prominent records in every interest are shared among associates that are all around appropriated. We can utilize document replication between areas for famous records, and utilize framework wide document looking for disagreeable documents.

6.2FILE DISTRIBUTION:

Physically close and normal interest nodes shape a subcluster, they can share records between each other so that a hub can recover its asked for document to its greatest advantage from a physically close hub. For this reason, the sub-cluster server keeps up the record of all documents in its sub-group for record sharing among nodes in its sub-bunch. A node's asked for document may not exist in its sub-cluster. To help nodes discover documents not existing in their sub-clusters, as in conventional DHT systems, PAIS re-disseminates all records among nodes in the system for proficient worldwide search

6.3 FILE QUERYING:

Intra-cluster looking means comprising of intra-sub-cluster seeking and between sub-group looking and between group seeking implies DHT steering. On the off chance that the intra-sub-cluster seeking comes up short, PAIS depends on between sub-group looking. On the off chance that the between sub-cluster seeking fizzles, it will rely on upon DHT routing for document searching

7 ALGORITHM:

7.1 NODE N JOINING IN PAIS ALGORITHM:

Step1: Each server tests its directing table sections and predecessor occasionally to ensure they are right.

Step2: If one of its neighbors fails to react amid a specific day and age T, the server finds and associates with another neighbor.

Step3: In a sub-cluster, a server chooses an optional server from its reinforcements that will supplant it upon its takeoff or failure.

Step4: It additionally tells all customers about the optional server. Prior to a server leaves, it asks for the auxiliary server to be the new server and tells all customers.

Step5: The customers then associate with the new server. To handle the impact of a server failure on its customers.

Step6: Each customer tests its server occasionally. On the off chance that a customer c does not get an answer from its server s amid T, c accept that s falls flat, and associates with the secondary server.

7.2 NODE N LEAVING IN PAIS ALGORITHM:

Step1: if it is the server in the sub-cluster of interest i then

Step2: if it has a supernode(s) in its backup list then

Step3: find supernode from its backuplist to replace itself

Step4: notify its clients about the server change

Step5: else

Step6: notify its clients to rejoin in the system

Step10: end if

Step11: execute leaving function in the Cycloid DHT

Step12: else

Step13: notify its server about its departure

Step14: end if

Step15: end for

7.3 LOOKING UP FILE IN PAIS ALGORITHM:

Step1: When node i needs to recover a record, if the document's key is one of the requester's advantage properties, it utilizes the intra-subcluster seeking.

Step2: Node i sends the solicitation to its server in the sub-cluster of the interest.

Step3: Every time a server gets a solicitation, it checks if its sub-cluster has the asked for record

Step4: If yes, the server sends the record area to the requester specifically.

Step5: If the record's key is not one of the requester's advantage characteristics, node i checks the presence of the document or a reproduction of the document in its cluster.

Step6: If there is a reproduction of the document, It ought to be put away in a sub-cluster nearest to ID

Step7: The solicitation is sent along the servers in every sub-cluster in the requester's cluster.

Step8: If there is no asked for document or reproduction of the asked for record, the record demand directing is performed.

Step9: hub i computes the ID of the document and conveys a message of Lookup(fileID)

7.4 DISTRIBUTED INTRA SUB-CLUSTER PROTOCOL:

Step1: node sends request to parent with time to live.

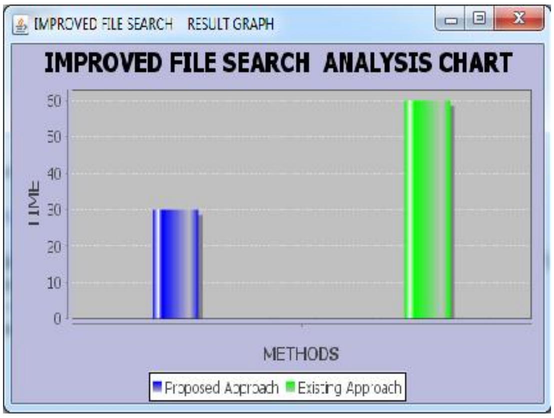
Step2: higher capacity node provides files to lower capacity nodes.

Step3: if node has higher capacity then

Step4: directly connected to children nodes

Step5: load is distributed to higher capacity nodes

8 RESULTS:



At long last proposed DHT based Intra subcluster document questioning convention sets aside less time for record questioning in PAIS structure

9 CONCLUSION&FUTURE WORK:

Present work in p2p systems concentrates on intrest cluster super peer networks as well as proximity clustered super peer networks. In this project introducing distributed intra subcluster querying which improves the file query efficiency and reduces file available latency also. Future research direction on improve the proposed techniques according to future need means if number of nodes increases in pais structure it is very difficult to improve the file search delay.

10 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 Unobservability, 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-topeerlookup 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- basedpublish/subscribe systems," Peer-to-Per 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 peer-topeer systems," in Proc. 22nd Int. Conf. Distrib. Comput. Syst., 2002, pp. 23–32.

MOHAMMAD ASLAMSULTANA received B.Tech degree in computer science and engineering from Jawaharlal Nehru Technological University,Kakinada (JNTU-K). She is currently pursuing M.Tech degree in Computer Science and Engineering from Jawaharlal Nehru Technology University, Kakinada (JNTU-K).Her research areas include peer to peer networking and online social networks.

Smt SSSN Usha Devi N received B.Tech & M.Tech degree in Computer Science and Engineering from Jawaharlal Nehru Technological University. She is currently an Assistant Professor in department of CSE at University of Engineering, JNTU Kakinada. She is pursuing her PhD from National Institute of Technology, Trichy(NIT-T) in the area of Artificial Intelligence. She has published 30 technical papers in various international journals and conferences. She has authored three text books on Artificial Intelligence. She has 6 years of experience in teaching to engineering students. Her bibliography was listed in Marquis Who is Who in World, USA. She is the reviewer of some international journals and conferences. Her research areas include deep learning, cellular automata, data mining and big data analytics.