International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 05 | May -2017 www.irjet.net p-ISSN: 2395-0072

Geographically Clustered P2P File Sharing Implementation

Sanoj Saju[1], Praahas Amin[2]

[1]*M. tech student, Sahyadri College of Engineering and Management, Karnataka, India.*
[2]*Assistant professor, Sahyadri College of Engineering and Management, Karnataka, India*

Abstract - *The overall performance of peer to peer file sharing system is determined by the efficient query. The main drawback of P2P file sharing system is traffic generated during the file transfer. An efficient query is required to avoid traffic, to improve the file query performance clustering the peer based on proximity. An approach is made to improve the efficiency of file query. The clustering can be done on their physical juxta position. In this work we deal with the structured P2P file sharing system. It is based on the physical closeness and nodes based on the geographical state. GCFS uses an technique of clustering the node based on proximity interest of the nodes. GCFS uses a novel lookup function include DHT and intelligent replication file algorithm to provide an efficient query. It creates an replicated file to reduces hotspot and thus result in efficiency of file query. GCFS uses an approach to increase the file efficiency, they forms a cluster based on the locations, then a sub cluster are formed based on interest and proximity. To avoid an overlay of nodes the lower capacity nodes connect with higher capacity node.DHT keeps track of nodes an key is mapped on to the nodes. To reduce a file searching delay an proactive file information collection is used, to improve file sharing efficiency it uses a bloom filter. It checks for the newly added bloom filter to reduce file searching delay.*

Key Words: Bloom Filter, Intelligent Replication File, Chord Technique, proximity aware.

1. INTRODUCTION

In the current year's P2P document sharing framework has picked up prominence in the web, for record sharing framework more movement is produced i.e. bit torrent. Efficient and trustworthy file query results in the overall performance of peer to peer file sharing. In P2P sharing system classified into two classes: structured and unstructured P2P network. In an unstructured P2P network depends either on flooding where nodes will be assigned to random selected neighbor node before locating the file.

Despite the fact that the key hub consistently changes DHT is dependable to keep refresh of the records in the hubs. To enhance the record area effectiveness a few strategies have being proposed, time taken to find the document from the hub depends the proficiency of the framework. Strategies have being proposed to build the effectiveness, versatility and deterministic information area.

To accomplish consistency support and load adjusted in organized P2P, it utilizes super companion topology, there are many issue in finding the documents they incorporate customer peer super associate connection, stack adjusting, record area and so on. It is understood utilizing super companion innovation, this system bring about making document questioning more solid.

In super companion topology they are delegated standard hub and super hubs, they are commonly connected with each other, general hub is encompassed to a super hub, which result in quicker network. In this way, these build a more predictable and steady spine. Bunching of the hubs and document replication in the system could bring about effectiveness in the record area.

Bunching procedure includes vicinity mindfulness, where physically close hubs are assembled together, inside this physical close hub are again arranged and gathered in light of intrigue, hub with same intrigue are grouped together. To keep up the consistency in P2P framework because of their successive information refreshes, An overlay system is set up with two layer for every imitation gathering. The two layer are delegated harmony imitation hubs and normal copy hubs, these aides in the consistency support. Sprout Filter with CAM strategy is utilized to perceive the softened words up a nano memory.

2. PROBLEM DEFINITION

The efficiency of the P2P file sharing depends on the ability to locate the file with the least query search. Unstructured P2P doesn't provide an efficient query because it is difficult to realize due to strictly defined topologies. To improve the file querying performance clustering is required, clustering is based on the common interest and closeness of the node. According to the geographical layout and closeness of the node it is clustered. To define both proximity and interest clustering is not possible in super peer technologies.

3. RELATED WORK

Super peer topology is used by fast track, morphens, due to firmly defined topologies it is harder to understand in P2P. Hierarchical structure of structured P2P helps us to group the peer based on interest and proximity. Proximity aware interests clustered have being grouped they are: (a) Super peer topology, (b) Proximity awareness, (c) Interest based

files sharing. It uses Intelligent File Replication algorithm that replicate the requested file near the physical node. In structured P2P system the super peer network is for efficient and scalable file consistency maintains, super peer network is for load balancing. To improve the file location efficiency, they are clustered on the basis of closeness.
Clustering technique in the P2P file sharing system involves proximity awareness. Hierarchical secure load balancing scheme in a P2P cloud system is proposed by liu et al. super nodes are balanced first, and then it depends on super node to balance other nodes /regular nodes. Each super node balances the node /regular node surrounded by it. Each super node caches the files recently requested by the other nodes and send a request to the super node to solve their request, self-organizing super node architecture to facilitate file querying.

4. PROPOSED SYSTEM

The time taken by the system /software to search /locate a file in a network determines the efficiency of the system. For an efficient and scalable P2P network sharing an efficient file retrieval technique is needed. It proposes a technique based on the geographical area and the closeness of the node based on the interest and clustering. It is based on cycloid structure of P2P network. It is a hierarchical structure, where they are clustered based on physical closeness of the node, then they are sub clustered based on their interest. The clustered nodes are again grouped on the basic of geographical location, to maintain the consistency for the peer to peer system due to their frequent data update, replica of the file is created.

An overlay network is established with two layers, the upper layer is based on the Distributed Hash Table and it is a powerful and consists of a stable replica nodes called Chord Replica Node (CRN) and the lower layer consist of an Ordinary Replica Nodes. Ordinary Replica Node is connected to the Chord Replica Node. When a file is to be replicated a message is initialized on the upper layer and the upper layer initialize a tree called Update Message Propagate Tree (UMPT).

4.1 ALGORITHM:

4.1.1 NODE N JOINING IN PAIS ALGORITHM:

Step1: Each server tests its coordinating table areas and Antecedent once in a while to guarantee they are correct.
Step2: If one of its neighbors neglects to respond in the midst of a particular day and age T, the server finds and partners with another neighbor.
Step3: In a sub-bunch, a server picks a discretionary server from its fortifications that will supplant it upon its departure or disappointment.
Step4: It moreover informs all clients regarding the discretionary server. Preceding a server abandons, it requests the assistant server to be the new server and tells all clients.
Step5: The clients then connect with the new server. To deal with the effect of a server disappointment on its clients.
Step6: Each client tests its server periodically. On the remote possibility that a client c does not get an reply from its server s in the midst of T, c acknowledge that s falls
Level, and connects with the auxiliary server.

4.1.2 NODE N LEAVING IN PAIS ALGORITHM:

Step1: in the event that it is the server in the sub-group of intrigue i at that point
Step2: in the event that it has a super node(s) in its reinforcement list then
Step3: find super node from its backup list to supplant itself
Step4: inform its customers about the server change
Step5: else
Step6: tell its customers to rejoin in the framework
Step7: end if
Step8: execute leaving capacity in the Cycloid DHT
Step9: else
Step10: tell its server about its takeoff
Step11: end if
Step12: end for

4.1.3 LOOKING UP FILE IN PAIS ALGORITHM

Step1: When hub I need to recoup a record, if the Archive's key is one of the requester's leeway properties; it uses the intra-sub cluster looking for.
Step2: Node I sends the requesting to its server in the Sub-group of the intrigue.
Step3: Every time a server gets a sale, it checks in the event that its sub-group has the requested record
Step4: If yes, the server sends the record zone to the Requester particularly.

Step5: If the record's key is not one of the requester's advantage attributes, hub i checks the nearness of the record or a propagation of the report in its group.
Step6: If there is a proliferation of the archive, it should be secured in a sub-group closest to ID
Step7: The requesting is sent along the servers in each sub-bunch in the requester's group.
Step8: If there is no requested archive or proliferation of the requested record, the record request coordinating is performed.
Step9: center point i figures the ID of the record and Passes on a message of Lookup (fileID)

4.1.4 DISTRIBUTED INTRA SUB-CLUSTER PROTOCOL

Step1: hub sends demand to parent with time to live.
Step2: higher limit hub gives records to lower limit hubs.
Step3: if hub has higher limit then

Step4: straightforwardly associated with kids hubs
Step5: load is circulated to higher limit hubs.

By partitioning the DHT identifier space, after the update message is passed to all the node regarding the new entry/deletion. Thus CRN nodes are stable nodes thus helps in maintaining the stable cluster after each update. The update entry as well as entry of each node is maintained in the server, when a request is made to search, the server check its index file and perform its search operation. The search is made in two stages they are inter cluster and intra cluster. If the search found success, the node sends the location of the file to the requested client, if not found it will make an inter cluster search. The efficiency of file searching depends on the DHT lookup among the nodes, cluster are sub clustered based on the key. The File Replication helps to improve the file location efficiency, Depending on the hash function such as SHA1 is used in the DHT lookup because it has a collision resistant nature. Computation is infeasible to find different message that facilitate the same message digest. Hash function is efficient to cluster message based on message dissimilarity. File is searched in a distributed manner, It forwards to the interest super node. It check for time to live to find out how long the message can survive, every message is associated with TTL, its value decreased by one at each time when a message passes through peer, if TTL is zero message will be dropped and no longer it will be forwarded, thus we can conclude that file querying is efficient in interest clustered and DHT lookup method can enhance file searching, for the lookup efficiency files are replicated to neighbor physical node if it is frequently requested.
Bloom Filter allows to check an element set and the reduant representation of the set. Redundancy is used for error detection and correction. Bloom Filter is used to correct and detect the error in the element set. CBF can be used to correct the error, it enable a cost efficient solution. To improve the file sharing efficiency the Bloom Filter result content are ranked to reduce the delay in file searching it check only the new added Bloom Filter information.

5. CONCLUSION

Due to firmly defined topologies it is harder to understand in structured P2P. In recent year several techniques used to improve the file sharing efficiency. The interest clustering and proximity aware clustering results in perk up performance of P2P system. In this paper we introduce a clustering based on structured P2P. Hierarchical structure helps in grouping the structured P2P based on interest and proximity of the client. A group of client based on their interest are grouped when they want to download a same file each of the client downloading will help the other client by sharing the folder /data, which results in faster downloads, thus result in decrease conjunction in the central server. Bloom filter using CAM based structure helps in resuming the broken downloads thus clustering based on geographic location results in lookup efficiency in node dynamism, it also results in reduces overhead

REFERENCES

[1] Haiying Shen Senior Memeber,IEEE, Guoxin Liu , Student Member,IEEE, Lee Ward“A Proximity Aware Interest Clustered P2P File Sharing System” ,Vol 26, No 6 ,Jun 2015.
[2] G.Rajashekaram, Nimmakayala Madhavi, Chittala Kamala, Kotha Sumanth, “A New Approach to Intra Sub Cluster File Searching Implemented in P2P File System”.IJSEAT,Vol 4,Issue 4,April 2016.
[3] Anju S Kumar, Ratheesh S, Manoj M, “ A Review on Efficient File Sharing in Clustered P2P System",IJIRCE,Vol 3,Issue 11,Nov2015 .
[4] Suravarapu Anusha,M.Srihari, M.Nagendra Kumar, “Implementation of Bloom Filter Using Cam Based Structure",IJSEAT,Vol.4,Issue 11,Nov 2016.
[5] K.Bhargavi, B.Soumya“An Innovative Peer to Peer Information Querying Structure to Decrease Overheads",IJITR Vol,4,Issue 4,June 2016.
[6] Talari Meena,V, Sundar Ratnam, “A Profitless Data Utilization Theory for Hierarchical Structures ",IJITR Vol 4,Issue 5,Sept 2016.
[7] Karthi Naveen Abyudav, V Krishna Chaitanya “A closeness Alert Interest Gathered P2P Folder Sharing System*",IJITR Vol 4*, Issue 6,Nov 2016.
[8] M Satya Priyanka, V.Ravi Kishore “Improved File Search and Sharing Mechanismin Structured P2P System",IJRCCT,Vol 5,Issue 8,Aug 2016.
[9] Varalakshmi .T, Arul Murugan .R“A Survey on P2P File Sharing Systems using Proximity Aware Interest Clustering ", IJRASET ,Vol 3 ,Issue 12, Dec 2015.
[10] Sanoj Saju, Praahas Amin “Geographically Clustered P2P File Sharing”,IRJET,Vol 4, Issue 4,April 2017.