Divide and Conquer Scheme for Efficient Load Balancing in P2P network

Mohanjeet Singh
Research Scholar
IKGPTU Jalandhar, Punjab, India

Anuj Kumar Gupta
Professor, Chandigarh Engineering College
Mohali, Punjab, India

Abstract- A consequential piece of the network traffic is produced by P2P applications. To provide quality of services, P2P needs contribution of varied intermediate peers. P2P network allocate the structural opportunity of reciprocally sharing and dispenser user-created contents inside diversity of user society environments. These networks are the collections of huge quantity of diverse nodes recognized as peers. These peers contain various exciting features like self-configuration, adaptation and organization. One problem of peers is that they are scattered. To make the searching and fetching fast these peers should be well organized. In this paper we have proposed multicasting technique which can reduce the communication distance between peers, resources will be better utilized, minimal wastage of resources will be there and contents will be fetched properly.

Keywords-P2P, Self-adaption, structured, unstructured, load balance

I. INTRODUCTION

P2P is a system which allows building a resource sharing system between many users. As pointed out by Shirky [16], "Peer-to-Peer is a class of utilizations that take gain of resources- stockpiling, CPU cycles, content, human presence- available at the edges of the Internet". One event that has marked the Internet growth in the current decade is the birth of the P2P network. P2P technology company; Cache Logic has expected flux proportion in related to internet communication traffic for the year 2005 to be between 64% and 84 %. P2P network make an overlay network on the upper of physical network.This network is a collection of large amount of heterogeneous nodes called as peers. These peers have some exciting features like self-configuration, adaptation and organization. P2Pis a distributed structure where software running on individual node offerssame functions. There is no need of centralized server because peer can act as a server or client, which stays away from the centralized server problems and solo point of failure [1]. P2P network offers the opportunity of reciprocally sharing and dispensing resources like video and music. P2P networks were delineate and were contemplate appropriate for massive contents dispensation across networks. Dispersed, scalable, economical, cooperative resource sharing, self-organizing encourage service providers to install real-time applications above heterogeneous network. Peer-to-Peer networks are scattered schemes in which software running on every node offers same functions. P2P networks try to offer extensive list of features like choice of close peers, excessive storage, successful found of items, data constancy, categorized naming andvalidation.

We can broadly divided P2P in Unstructured and Structured network. In Unstructured as the name demonstrates they do not keep up any particular structure on the network. There is very little deliberate data to help list objects crosswise nodes. There fore nodes frequently locally keep up objects they share. Search in unstructured network is pretty much a visually impairedactivity, transient queries to identify whether they have wanted objects. Since search is completed locally. Complicated queries like wildcards can be carrying as they were in a central system. Search space risesin extent to recall rate on the grounds that search is ultimately a visually impaired procedure. In Structured networks it keeps a structure to retain a network. Utmost Structured peer to peer work by distributed hash tables (DHT) where a key is given, they can proficiently detect an object containing the key by keeping neighboring tables entry. Nodes sustain routing information regarding nodes which helpful to reach all nodes in the overlay network [2]. Structured overlays offer a boundary on the amount of messages desired to discover some object in the overlay. It is mainly essential when searching for occasionally happening on less famous objects. Routing table (Local) which is maintained by every peer is used by way of the sending algorithm. As soon as the peer connectsin overlay using a particular bootstrap technique, routing table is initialized. Peers occasionally

interchange routing table modifications as part of overlay repairs. Peers who are nearby in the overlay can be far in the underlying network as the address space is virtualized but peer addresses are normally randomly allocated. The process of any peer-to-peer delivery system depends on a network of peer as well as connections among them. This network is planned on top of some independently from the primary physical computer network there for it is stated as an "overlay" network.Each peer keeps a small routing table containing NodeIDs and IP addresses of its adjacent peer. Queries are transmitted through overlay paths to peers through the NodeIDs that are nearer to the key. A basic feature of these procedures is to locate a suitable node in the overlay. Routing table which comprises the explanation of another node (NodeID and IP) is used by node to discover an suitable node. Each DHT-based system keep sits own organization strategy for the content objects as well as its key space and routing policies.The underlying network path among two peers may not be same as the path on the DHT base overlay network.

II. NETWORK STRUCTURE

By structure, we state whether the overlay network is made ad hoc as nodes are added, otherwise it is created on particular rules. According to the structure, we classify peer-to-peer networks as follows:

Unstructured -The location of content (files) is not related to the overlay topology. In an unstructured network, datanormally needs to be found. Searching is completed by the help of brute force techniques, like flooding the network by dispersal queries in a breadth-first or depth-first manner till the desired content is found.

Structured - These have appeared mostly in a try to address the scalability problems that unstructured systems were handle with. The overlay topology is strongly organized and files are positioned at exactly specified localities. These systems fundamentally deliver a mapping among content as well as position in the form of dispersed routing table. As a result queries can be forward to the node with preferred content. Structured systems propose a scalable result for accurate match queries, with the intention of queries where the accurate identifier of the needed data object is known. A drawback of structured network is that it is difficult to retain the structure necessary for effective routing messages.

III. SUPPER NODE

Supper node is a reference node who knows the network address. It response to the request send by participants of peer-to-peer system. Supernodes must be distributed in peer-to-peer network so that following can be achieved.

Right to use: Nodes which are not working as supernodes have low latency access to single supernode or more.Access can be measured in hop counts or delay.

Dispersion: Supernodes must be evenly spreadall over the overlay network; they should not be clustered within only a few sub-regions of the overlay.

Proportional: A pre-described ratio of supernodes toward non-supernodes should be maintained to full fillapplication-based performanceneeds.

Load equilibrium: Supernodes should not provide services to more than *n*non-supernodes, where *n* can be organized locally founded on the resource ability of every supernode.

The design founded on superpeers has following main benefits. Primary, some system tasks, like assigning hosting data or services, high-performance peers, i.e., super-peers, to enhance the overall system reliability and performance. Second, the use of super-peers permit the system to bound the number of nodes in distributed algorithms, such as [25, 26, 27, 28], which scale down the performance and become very costly as soon as the system size is big. Thus, the super-peer design can improve the scalability of a P2P system. Though apply of super-peers initiate some new problems that need to be discussed. To select super-peers, the system wishes to determine the numbers of super-peers are desired as well as which peers are utmost suitable to work as super-peers.

Additionally, the system desires to sustain and non-stop accommodate the super-peer set in reply to peer entrances and leavings, modifications in the existing load and changes in peer capabilities. The system also requests to allocate clients among superpeers. It migrate them after super-peers quit the network or fail. Preferably, the load among super-peers must be equal to guarantee the system's scalability as well as fault-tolerance. This must all be performed in a dynamic and decentralized manner, with no central coordination or authority. In P2P a supernode is a node which also functions as a relayer as well as proxy servers, control data flow and links for other users. A supernode usually needs extra network bandwidth plus CPU time. The supernodes essential be well dispersed through the P2P overlay network, also satisfy additional needs like load balance, resource requirements, adaptability to churn, and heterogeneity. To select Supernode is very complicated task, for the reason that it must reply to joins and leaves the nodes and working an atmosphere that is extremely heterogeneous.

A node cannot work as supernode without meet following qualifications like:

Resources: like CPU ability, memory storage;
Resilience: As soon as supernode leaves the network, new supernodes would immediately take over its charge.
Safety: Supernodes may be unsafe to denial of service affect. Effected supernodes can damage the system by failed to ahead the messages or by forwarding incorrect information.
Heterogeneity: Supernodes must have good capability and network connectivity.

IV. LITERATURE WORK

The objective of efficient overlay network construction is to organize participating peers into a logical topology that must infer the original topology. Actually , topology which is not appropriate can result in extra overhead and can reduce the system performance drastically. The overlay construction should be scalable. ChristakidisAthanasios [3] has proposed a locality aware and balanced overlay for p2p live streaming which can adapt to the dynamic behavior of the participating peers and the underlying network. KalmanGraffi [4] has proposed a load-balanced architecture for P2P-based multimedia streaming and a stream provider selection mechanism, which can be applied on any distributed hash table (DHT). MubasharMushtaq [5] has proposed a hybrid overlay network mechanism which combines semantic aware peers and location aware peer organization in which nodes are arranged based on the two factors like semantics and its location. FanChoa [6] has proposed Improvement of Structured P2P Routing Algorithm Based on NN-CHORD which can solve the problem of resource location in structured P2P networks. However, because the routing table in Chord suffers serious information redundancy and the lookup direction of Chord is only clockwise, it is not very effectual. Therefore, a routing algorithm BNN-Chord (Bidirectional Neighbor'sNeighbor Chord) based on NN-Chord (Neighbor's Neighbor Chord) is proposed. Zhao-Kui Li [7] has proposed P2P algorithm using supernodes and self-organizing clustering. This algorithm chose a sequence of robust nodes as super-nodes, which work as cluster head of self organizing structure. The outcome proves that the algorithm can clearly,enhance routing functioning as well as solve topology conflict.OtsoKassinen [8] has proposed Analysis of Messaging Load in a P2PP Overlay Network under Churn. They analyze the performance of a DHT-based P2P overlay network in resource access and overlay maintenance activities. They use a cross platform implementation of the protocol known as P2PP. The performance metric is the request accomplishment success ratio in the overlay network. The appraisals are conducted in the presence of churn i.e. the peers' continuous joining and leaving the overlay, which is vitalfor realistic simulations. The time-periods of churning peers both online and offline follow an exponential distribution. In addition to churn, they use overlay size and resource lookup activity as mutable parameters. Measurements of request success ratio can be used to approximate the performance of a P2P overlay network and the load inflicted on it, once its variety of peers and activity level area unit noted. Bow-Nan Cheng et al [9] propose a brand new methodology of building a virtual structure associated introduce a way to route packets through an unstructured overlay network. They introduce Virtual Direction Routing (VDR). VDR may be a light-weight and scalable overlay network routing protocol that uses the thought of virtual directions to perform node data seeding and search. State data is simulated at nodes on virtual orthogonal lines originating from every node and sporadically updated. When a path search is started, rather than flooding the network, query packets also are forwarded on virtual orthogonal lines till associate intersection with the seeded state happens. VDR scales well while not imposing DHT-like graph structures (e.g., trees, rings) and the path stretch compared to random-walk protocols is terribly sensible. The exchange is supplementary latency by choosing suboptimal methods.

V. PROBLEM FORMULATION

Mainly present P2P systems assist the basic object lookup by identifier. A few P2P systems can control extra complicated keyword queries, which determine documents containing keywords in queries. The majority searching approaches are forwarding-based. Beginning by client node, a query is transmitted node to node until the node which has the desired data is reached. To forward query messages, each node must keep information about some other nodes called neighbors. The information of these neighbors constitutes the routing table of a node. Behind the success of searching information, node organization play significant role.

Content fetching and delivery has become one of the most popular P2P applications because of high scalability and low cost implementation. There are many techniques for fetching the content but still there are many problems. For fetching we have got to search out the data sources. Because Peers are distributed in P2P network there for contents are scattered and duplicated in a distributed fashion. To find the object, P2P networks use broadcasting. Hyper Cup [19] which uses broadcasting to search, a large volume of traffic is created during the search process. Some of the problems in P2P network that use broadcasting are their scalability limitation and unnecessary traffic [29]. Another weakness of data broadcast is that a client has to listen to the entire broadcast cycle to retrieve all data items it wants. Not only the access time which depends on the length of the broadcast cycle could

be quite long it also means that the client has to be active in the entire cycle to get perhaps just a few data items from of the broadcast [30]. Multicast has been proposed as a key alternative for large-scale group communication. Multicast technology is directed towards distributed applications. Distance learning and video conference multimedia applications can be used in the network in a measurable and effective way. The effective use of the network and the reduction of the load in traffic sources permit services and applications to be accessed by a great number of participants. The economy of network resources associated to the reduction of the load in the applications and servers makes the network less prone tojams, and, thus, more available to be used. There we need best multicasting approach. For multicasting approach we need a construction of multicasting tree in structured format.

VI. PROPOSED METHODOLOGY

As we have seen that broadcasting is not efficient. It has scalability limitation and unnecessary traffic. To get a few data items from the broadcasting the client has to be active in the entire cycle and has to check the entire broadcast cycle to retrieve all data items it wants. Access time could be quite long because it depends on the length of the broadcast cycle. Multicasting approach can be used in a measurable and effective way mainly on those areas where there is a need of huge data transfer in multimedia applications.

In our proposal we separate whole network in little logical regions (clusters). Nodes in every region pick a SuperNode. The separation of the regions will be done in such a way that it decreases the communication gap between nodes toward SuperNode. The Nodes which exist in the specific region, communicates straight forwardly by that specific SuperNode. Choice of SuperNode in the rest of regions is in light of mention position in all regions; mention position is the center position of every region. Node nearest toward reference point is chosen as SuperNode first and foremost, then next nearest node thus on till least nearest node. In every round just one SuperNode is chosen in every region moreover, we uses multi-hop method for inter region communication to reduce communication span.

We will follow the steps to develop the regions:

In primary step network is split into m equal distant concentric squares. For simplicity, we take $n = 3$. Accordingly network is split into three identical distance concentric squares: inner square (ls), center square (Cs) and external square (Es). In next step we split the area among two squares called curve Regions (CR) and Non curve Regions (NCR). As cluster are stationary so one SuperNode is chosen in every Non curve Regions.

Middle spot of every NCR is treated as reference spot for choice of SuperNode in that area, nearby node from middle reference point is elected as SuperNode. After that, next nearby node from the reference position is elected as SuperNode for further rounds.

VII. CONCLUSION

In P2P network peers are distributed there for contents are also scattered. To find the required content broadcasting or multicasting techniques may be applied. Since we have observed that there are several confines in broadcasting and the result is delay in searching. In this paper we have proposed multicasting technique for divide the P2P network. This technique divides the network field into sub regions. It is a hybrid approach of static clustering and dynamic cluster head selection. For cluster communication it uses two level hierarchies. It make square and rectangular regions which further divide the network into small regions to reduce the communication distance for intra cluster and inter cluster. We propose that if we divide the network in this manner than the resources will be better utilized, minimal wastage of resources will be there and contents will be fetched properly.

ACKNOWLEDGMENT

I would like to thanks I. K. Gujral Punjab Technical University Jalandhar, for providing me opportunity and technical support to complete this research work.

REFERENCES

[1] Sungsu Kim, “Semantic Overlay Network for Peer-to-Peer Hybrid Information Search and Retrieval”, In 12th IFIP/IEEE 1M 2011.

[2] J. Buford, H. Yu and E. K. Lua, “P2P Networking and Applications”, Morgan Kaufmann Publishers (2008)

[3] ChristakidisAthanasios, “Locality aware structured overlay for P2P live streaming”, Springer Berlin Heidelberg- 2006

[4] KalmanGraffi, “Load Balancing for Multimedia Streaming in Heterogeneous Peer-to-Peer Systems”, ACM New York, NY, USA ©2008

[5] MubasharMushtaq, Toufik Ahmed, “Hybrid Overlay Networks Management for Real-Time Multimedia Streaming over P2P Networks”, IEEE International Conference on Management of Multimedia and Mobile Networks and Services: Real-Time Mobile Multimedia Services 2008.

[6] Fan Choa, “Improvement of Structured P2P Routing Algorithm Based on NNCHORD”, IEEE 7th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM), 2011.

[7] Zhao-Kui Li; Yan Wang, "P2P routing advanced algorithm using super-nodes and self-organizing clustering", IEEE 2nd International Conference on Future Computer and Communication (ICFCC), 2010.

[8] OtsoKassinen, ErkkiHarjula and Mika Ylianttila, "Analysis of Messaging Load in a P2PP Overlay Network under Churn", Proceedings of the Global Communications Conference
Miami, Florida, USA, 6-10 December 2010, pages 1-5, IEEE.

[9] Bow-Nan Cheng, "Virtual Direction Routing for Overlay
Networks", IEEE Ninth International Conference on Peer-to-Peer Computing, 2009.

[10] H.Balakrishnan, M. F .Kaashoek, D.Karger, R.Morris, and I.Stoica, "Looking up data in P2P systems", Communications of ACM, Vol.46, No.2, 2003.

[11] N.Daswani, H.Garcia -Molina, and B.Yang, "Open problems in data -sharing peer-to-peer systems", Proc. of the 9th International Conference on Database Theory (ICDT'0 3), 2003.

[12] Ahmed Moujane, DalilaChiadmi, Laila Benhlima, FaouziaWadjinny, "Managing Network Dynamicity in a Vector Space Model for Semantic P2P Data Integration", In Communications in Computer and Information Science Springer Volume 241 pp 186 -200 2011.

[13] Zhi Yang, "Exploring peer heterogeneity: Towards understanding and application", In IEEE International Conference on Peer-to-Peer Computing (P2P), 2011.

[14] Fiorese, A," An approach to peer selection in service overlays", In IEEE 7th International Conference on Network and Service Management (CNSM) 2011.

[15] Eytan Adar and Bernardo Huberman, "Free Riding on Gnutella, First Monday". October 2000.

[16] Shirky, C., "What is p2p and what isn't. Network", O'Reilly (2000)

[17] D. P. Anderson, J. Cobb, E. Korpela, M. Lebofsky, and D. Werthimer, "Seti@home: an experiment in public-resource computing", Communications of the ACM, 45(11):56-61, 2002.

[18] R. Bhagwan, S. Savage, and G. M. Voelker, "Understanding availability", In Proceedings of the 2nd International Workshop on Peer-to-Peer Systems, volume 2735 of Lecture Notes in Computer Science, pages 256267. Springer, 2003.

[19] D. Werthimer, J. Cobb, M. Lebofsky, D. Anderson, and E. Korpela, "SETI@HOMEmassivelydistributed computing for seti", Computing in Science and Engineering, 3:78-83, January 2001

[20] Q. Lv, S. Ratnasamy, and S. Shenker,"Can heterogeneity make gnutella scalable?", In Proceedings of the 1st International Workshop on Peer-to-Peer Systems, volume 2429 of Lecture Notes in Computer Science, pages 94103. Springer, 2002

[21] S. Saroiu, P. K. Gummadi, and S. D. Gribble, "Measuring and analyzing the characteristics of napster and gnutella hosts", Multimedia Systems, 9(1):170184, July 2003.

[22] A. I. T. Rowstron and P. Druschel, "Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems", In Proceedings of the 18th IFIP/ACM InternationalConference on Distributed Systems Platforms, pages 329-350. Springer-Verlag, 2001.

[23] A. Singh and M. Haahr, "Creating an adaptive network of hubs using schelling's model", Communications of the ACM, 49(3):6973, 2006.

[24] A. Crespo and H. Garcia-Molina, "Routing indices for peer-to-peer systems", In Proceedings of the 22nd International Conference on Distributed Computing Systems, pages 23-34, Washington, DC, USA, 2002. IEEE Computer Society.

[25] Q. Lv, P. Cao, E. Cohen, K. Li, and S. Shenker,"Search and replication in unstructured peer-to-peer networks", In Proceedings of the 16th International Conference on super
computing, page 84-95. ACM, 2002.

[26] B. Yang and H. Garcia-Molina,"Improving search in peer-to-peer networks", In Proceedings of the 22nd International Conference on Distributed Computing Systems, pages 5-14. IEEE, 2002.

[27] D.Tsoumakos and N. Roussopoulos., "A comparison of peer-to-peer search methods", In Proceedings of the 6th International Workshop on the Web and Databases, pages 61-66, 2003.

[28] J. Li, B. T. Loo, J. M. Hellerstein, M. F. Kaashoek, D. R. Karger, and R. Morris,"On the feasibility of peer-to-peer web indexing and search", In Proceedings of the 2nd International Workshop on Peer-to-Peer Systems, number 2735 in LNCS, pages 207-215. Springer, 2003.

[29] M. Schlosser, M. Sintek, S. Decker, and W. Nejdl. Hypercup shaping up peer-to-peer networks. Technical report, Stanford University, 2001.

[30] Rostami, H. , Habibi, J, Abolhassani, H. ; Amirkhani, M. ; Rahnama, A., "An Ontology Based Local Index in P2P Networks", Second International Conference on Semantics Knowledge and Grid-Cover,Nov 2006.

[31] DikLun Lee, Qinglong Hu, "Indexing techniques for data broadcast on wireless channels", "The 5thInternational Conference on Foundations of Data Organization, Japan November 1998.