Scientific Journal of Impact Factor (SJIF): 4.72

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

"Emerging Technologies in the Computer World", January -2017

An Efficient and Scalable Framework for Peer-to-Peer Clustering and Firework Query Model

Prof. Shilpa Pimpalkar, Department of Computer Engineering, AISSMS's IOIT

Bhosale Prachi[1], Gadve Vaishnavi[2], Khopade Mayuri[3], Pawar Kajal[4].
Computer Engineering, AISSMS's IOIT

Abstract — *Efficient and trustworthy file querying is very important to the general performance of peer-to-peer (P2P) file sharing systems. rising strategies area unit setting out to address this challenge by exploiting on-line social networks (OSNs). However, current OSN-based strategies merely cluster common-interest nodes for prime potency or limit the interaction between social friends for prime trustiness, that provides restricted sweetening or contradicts the open and free service goal of P2P systems. very little analysis has been undertaken to totally and hand in glove leverage OSNs with integrated thought of proximity and interest. during this work, we tend to analyze a Bit Torrent file sharing trace, that proves the requirement of proximity- and interest-aware bunch. supported the trace study and OSN properties, we tend to propose a social Network integrated P2P file sharing system with increased potency and trustiness (SoNet) to totally and hand in glove leverage the common-interest, proximity-close and trust properties of OSN friends. SoNet uses a ranked distributed hash table (DHT) to cluster common-interest nodes, then any cluster proximity-close nodes into subcluster, and connects the nodes in an exceedingly subcluster with social links. Thus, once queries move trustable social links, they additionally gain higher likelihood of being with success resolved by proximityclose nodes, at the same time enhancing potency and trustiness. The results of trace-driven experiments on the realworld Planet Lab test bed demonstrate the upper potency and trustiness of SoNet compared with alternative systems..*

Keywords-P2P,clustering,proximity,awareness

I. INTRODUCTION

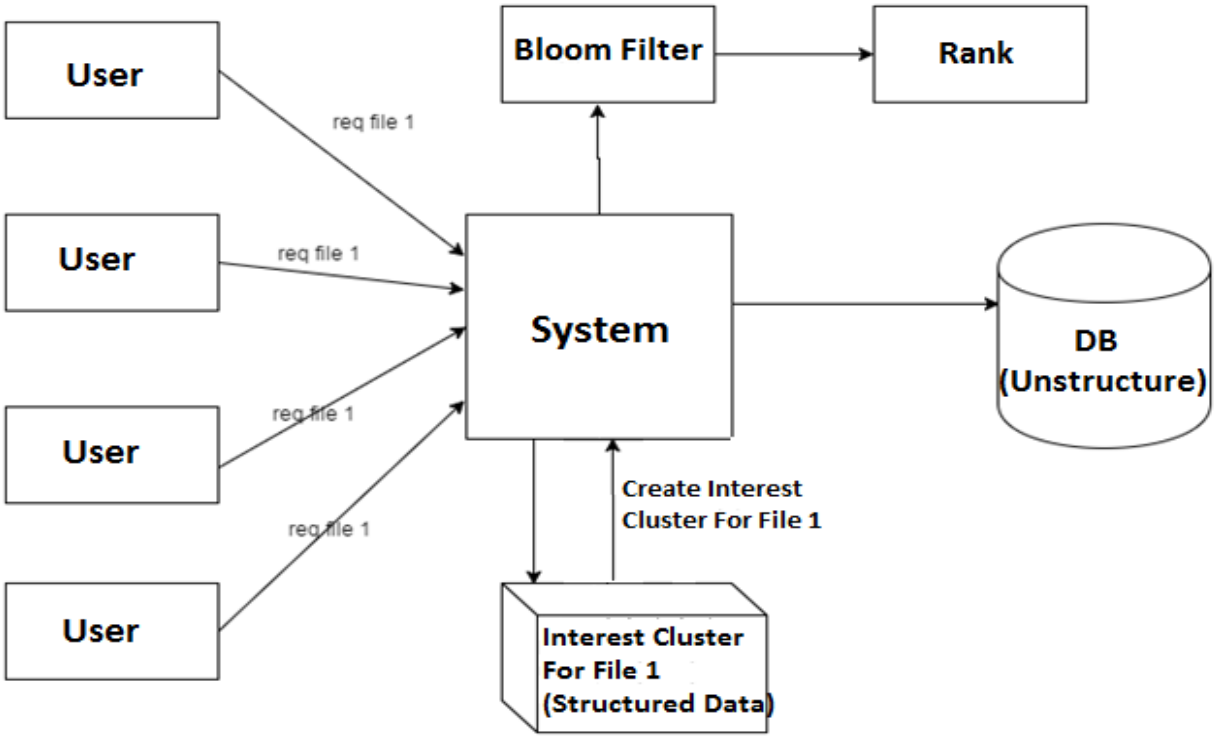
Over the past few years, the Brobdingnagian quality of the web has made a major stimulation to P2P file sharing systems. There square measure 2 categories of P2P systems: unstructured and structured. Unstructured P2P networks like Gnutella and Freenet doesn't assign responsibility for information to specific nodes. Nodes be a part of and leave the network in step with some loose rules. Currently, unstructured P2P networks' file question technique is predicated on either flooding wherever the question is propagated to all or any the node's neighbors, or random-walkers wherever the question is forwarded to haphazardly chosen neighbors till the file is found. However, flooding and random walkers cannot guarantee information location. Structured P2P networks i.e., Distributed Hash Tables (DHTs), will overcome the drawbacks with their options of upper potency, quantifiability, and settled information location. It strictly controlled topologies, and their operation algorithms and information placement square measure exactly outlined supported a DHT organisation and consistent hashing operate. The node is answerable for a key will invariably be found eventhough if the system is in a very continuous state of modification. Most of the DHTs need O (log n) hops per operation request with (O log n) neighbors per node, wherever n is that the variety of nodes within the system.

II.PROPOSED SYSTEM

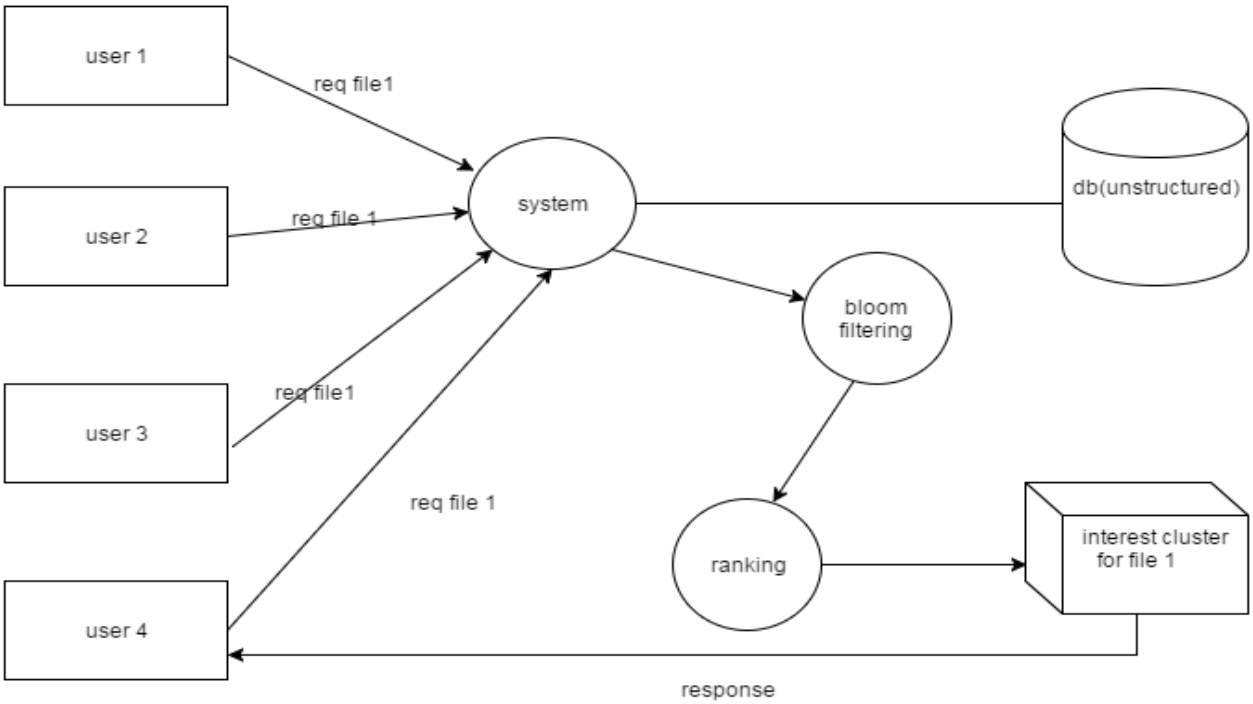
In this paper represents a proximity-aware and interest-clustered P2P file sharing System (PAIS) on a structured P2P system. It forms physically-close nodes into a cluster and additional teams physically-close and common-interest nodes into a sub-cluster. It additionally places files with identical interests along and create them accessible through the DHT Lookup() routing operate. a lot of significantly, it keeps all benefits of DHTs over unstructured P2Ps. wishing on DHT search policy instead of broadcasting, the PAIS construction consumes abundant less value in mapping nodes to clusters and mapping clusters to interest sub-clusters. PAIS uses associate intelligent file replication formula to additional enhance file search potency. It creates replicas of files that area unit oft requested by a gaggle of physically shut nodes in their location. Moreover, PAIS enhances the intra sub-cluster file looking through many approaches 1st, it additional classifies the interest of a sub-cluster to variety of sub-interests, and clusters common-sub-interest nodes into a gaggle for file sharing. Second, PAIS builds associate overlay for every cluster that connects lower capability nodes to higher capability nodes for distributed file querying whereas avoiding node overload. Third, to cut back file looking out delay, PAIS uses proactive file info assortment so a file requester will understand if its requested file is in its near nodes.Fourth, to cut back the overhead of the file info assortment, PAIS uses bloom filter based mostly file info assortment and corresponding distributed file looking out. Fifth, to boost the file sharing potency, PAIS ranks the bloom filter leads to

order. Sixth, considering that a recently visited file tends to be visited once more, the bloom filter based mostly approach is increased by solely checking the fresh added bloom filter info to cut back file looking out delay.

III.SYSTEM ARCHITECTURE



IV FLOWCHART



V MATHEMATICAL MODEL

Let W be the whole system which consists:

W = {input, process, output}.

Input: {p, N, F, i}.

Where,

1. p probability of Request.
2. N is number of samples taken for estimation.
3. F is the frequency of number of events.

Process:

We implement the system that periodically transmits raw data samples at the rate of 920Hz. System observes the transmission result of each data sample and estimates the request probability p` by

$$P` = \frac{1}{N} \sum_{i=1}^{N} 1_{F_i}$$

Where N is the number of observations clusters in the network, and Fi denotes the event that the i-th transmission fails.

VI. ADVANTAGES

1. It enhances communication and availability of information.

P2P, particularly with full access to the online, permits ways in which of communication that might merely be not possible before it had been developed. Instant electronic messaging will currently permit users to speak in real time and send files to people where they're within the world, that could be a vast boon for businesses. Also, it permits access to a huge quantity of helpful data, together with ancient reference materials and timely facts, like news and current events.

2. It allows for more convenient resource sharing.

This profit is extremely vital, notably for larger firms that basically have to be compelled to manufacture immense numbers of resources to be shared to any or all the individuals. Since the technology involves computer-based work, it's assured that the resources they needed to urge across would be utterly shared by connecting to a network that their audience is additionally victimisation.

3. It makes file sharing easier.

P2P allows easier accessibility for people to share their files, which greatly helps them with saving more time and effort, since they could do file sharing more accordingly and effectively.

4. It is highly flexible.

This technology is known to be very flexible, as it gives users the opportunity to explore everything about essential

hings, such as software without affecting their functionality. Plus, people will have the accessibility to all information they need to get and share.

5. It is an inexpensive system.

Installing networking software on your device would not cost too much, as you are assured that it lasts and can effectively share information to your peers. Also, there is no need to change the software regularly, as mostly it is not required to do so.

6. IT INCREASES COST EFFICIENCY.

With computer networking, you can use a lot of software products available on the market which can just be stored or installed in your system or server, and can then be used by various workstations.

V CONCLUSION AND FUTURE SCOPE

In recent years, to boost file location potency in P2P systems, interest-clustered super-peer networks and proximity-clustered super-peer networks are projected. though each ways improve the performance of P2P systems, few works cluster peers supported each peer interest and physical proximity at the same time. Moreover, it's more durable to comprehend it in structured P2P systems because of their strictly outlined topologies, though they need high potency of file location than unstructured P2Ps. during this paper, we have a tendency to introduce a proximity-aware and interest-clustered P2P file sharing system supported a structured P2P. It teams peers supported each interest and proximity by taking advantage of a hierarchical data structure of a structured P2P.

We don't would like the additional proxy server to store the information backup,if one in every of node is get unsuccessful then our system not be collapse. thus in future we are going to improve the node capability or network affiliation.

VI APPLICATIONS

1. University and College Level
2. Business Aplication
3. Government Aplication
4. Military Aplication

VII ACKNOWLEDGMENT

We want to thank analysts and distributers for making their assets accessible. We additionally appreciative to Commentator for their significant recommendations.

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-topee 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

