Contents lists available at www.innovativejournal.in

Asian Journal Of Computer Science And Information Technology

Journal homepage: http://www.innovativejournal.in/index.php/ajcsit

ISSN 2249 - 5126

AN ALGORITHM FOR CONSISTENCY MAINTENANCE IN P2P SYSTEMS

K. Shalini*, Y. Surekha

Department of Computer Science & Engineering, PVP Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

ARTICLE INFO

Corresponding Author:
K. Shalini
Department of Computer Science & Engineering, PVP Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India

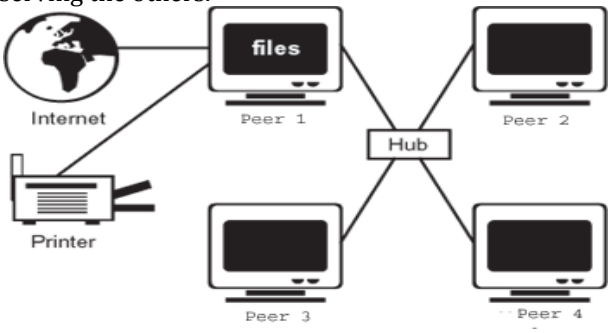
KeyWords: P2P, File replication, consistency maintenance.

ABSTRACT

P2P is a popular technology used for file sharing. File replication and Consistency maintenance are the techniques used in P2P for high system performance. File replication methods specify replica nodes without considering consistency maintenance which may lead to high overhead for unnecessary file replications and consistency maintenance. Consistency maintenance methods update files without considering file replication dynamism which may not give the accuracy of replica consistency. Instead of passively accepting replicas and update messages, we develop a mechanism which combines both file replication and consistency maintenance. This mechanisam is very efficient in file replication and consistency maintenance at low cost. In this mechanism, accepting replicas and updates is based on file query rate and update rate.

INTRODUCTION

The term P2P refers to "peer-to-peer" networking. A peer-to-peer network allows computer hardware and software to function without the need for special server devices. P2P is an alternative to client-server network design. In client-server network , each computer or process on the network is either a client or a server. Servers are powerful computers or processes dedicated to managing disk drives (file servers), printers (print servers), or network traffic (network servers). Clients are PCs or workstations on which users run applications. Clients rely on servers for resources, such as files, devices, and even processing power. In peer-to-peer network, each workstation has equivalent capabilities and responsibilities. This differs from client/server architectures, in which some computers are dedicated to serving the others.



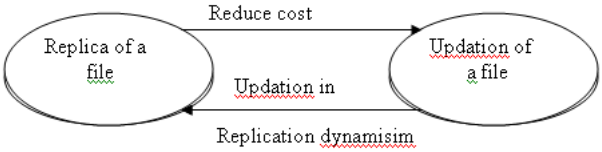
With increase in popularity of Peer to Peer (P2P) networks it has also become one of the medium for spreading of viruses, spywares, ad ware, malware through file sharing applications.

Some of the P2P file sharing programs available on internet are bittorrent, limeware, kazaa, shareaza, imesh, bearshare lite, kceasy, ares galaxy, emule, soulseek, winmx, piolet etc

Most of the people download audio and video files by using P2P file sharing. Whenever a file is requested frequently, the capacity of the node degrades and gives delayed response. File replication is very useful in this situation. In this method , the load is distributed over replica nodes. File consistency maintenance is to maintain consistency between file and its replica nodes.

This paper presents a mechanism which integrates File Replication and Consistency Maintenance to achieve high efficiency in file replication and consistency maintenance at a lower cost.

Fig. 1. Interrelationship between replica of a file and updation of a file



Replication dynamism deals with replica node generation, deletion and failures.

RELATED WORK

The file replication methods copies files near file owners , file requesters or along a query path from a requester to a owner. PAST [2], CFS [3], and Backslash [4] replicate each file on close nodes near the file's owner. In LAR [5] and Gnutella [6], overloaded nodes replicate a file at requesters. In these methods, file owners rigidly determine replica nodes and nodes accept replicas. They are unable to keep track replica utilization to reduce underutilized replicas and ensure high utilization of existing replicas. In efficient and adaptive decentralized file replication algorithm in P2P file sharing systems called

EAD [8] , traffic hubs that carry more query load are chosen as replica nodes. The nodes continuously check their query load in order to create copy for the file and remove low utilized replicas. Replication in a structured P2P system is to decrease file query time, while replication in an unstructured P2P system is to decrease the search time.

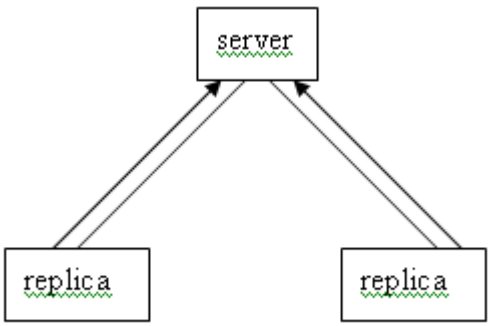
File consistency methods are based on structure [7] and message spreading [9].In structure based methods, stable replica nodes are used but it is not true in practice because of file replication dynamism. In message spreading, unnecessary and redundant messages are generated and is not sure that all replicas receive update messages. Therefore the methods lead to unnecessary file replications and overhead in consistency maintenance.

In file replication and consistency maintenance methods, nodes accept replicas and update messages. They are unable to keep track the utilization of replicas to determine the need of file replicas and replica updates. Minimization of the number of replicas helps to reduce unnecessary updates in consistency maintenance. Here the number of replicas are based on queries.

PROPOSED SYSTEM

Our system is a combination of both file replication and consistency maintenance. Both are dependent on each other. Instead of accepting replicas and update messages, it integrates file replication and consistency maintenance by letting each node autonomously determine the need for file replication and update based on file query rate and update rates. File replication places replicas in frequently visited nodes to guarantee high utilization of replicas, and meanwhile reduce underutilized replicas and overhead of consistency maintenance.

Fig. 2 Combined approach of file replication and consistency maintenance



In the above figure, the straight line represents the link between replica node and server and the arrow mark represents that the replica polls the server for update, to make sure that an update file is available to the client.

Consistency maintenance aims to guarantee file fidelity of consistency at a low cost with file replication dynamism consideration. Using adaptive polling, this ensures timely update operation and avoids unnecessary updates. The basic idea of this approach is to use file query and update rate to direct file replication and consistency maintenance.

Adaptive File Replication

Combined approach of File Replication and Consistency maintenance mechanism is developed by using EAD [8] file replication algorithm. This algorithm achieves an optimized trade-off between query efficiency and overhead in file replication. File replication component addresses two main problems 1) The point at which the replicas should be generated , and are not underutilized. 2) To remove underutilized and unnecessary replicas so that the overhead for consistency maintenance is less.

- The need to determine replica nodes

Whenever a popular file like audio and video is accessed continuously, the server will be degraded and will give delayed response .In that situation placing a replica for that file is suitable.

- Creation of replica for popular file

Creation of a replica is based on file query rate. A replica is created when the requesters request continuously a file. If the query rate is less to that replica., then the replica node is deleted

- Adaptation of Replica node

A file might not be accessed frequently .so the replica node should frequently update their query passing rate, so that underutilized replicas can be removed

File Consistency maintenance

To maintain consistency between frequently and not frequently updated files and its copies is an important aspect in peer-peer file sharing systems.. Here the node might be generated or failed or deleted.. To be aware of this we employ a polling method, in which the replica node itself polls the server for update continuously.

In our proposed system , the replica node frequently polls the file owner for update. It is based on two main problems 1) In what frequency , the replica node polls the sever for update 2) to reduce the polling operations to save cost and to maintain accuracy in consistency maintenance.

The combined approach has a time-to-refresh (TTR) value with each replica node of a file. It denotes at what time the replica should poll the file owner to keep its replica updated. a node should poll the owner to keep its replica updated. The TTR value is changed frequently based on the results of each polling. It takes file query rate for poll time determination. TTRquery and TTRpoll denotes the next time, where the file is updated.

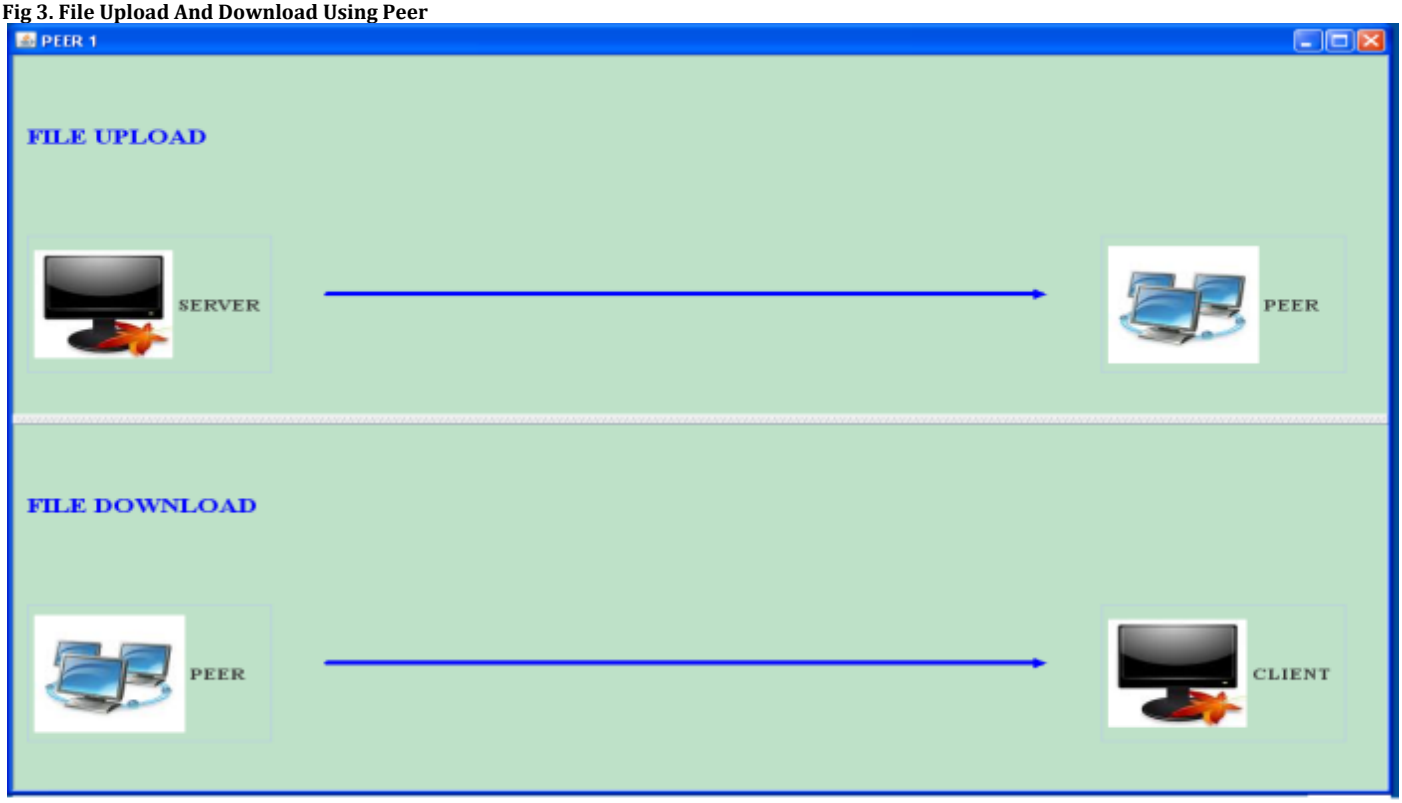
//Algorithm for file consistency maintenance

1. If a query is requested for a file then
2. include an update request within query of the file
3. else send the update request
4. if the aecceptance reply is given from the owner of the file
   / /check the conditions
5. If the file is a valid one then
   TTR = TTR_{old} + α. // α is a constant
6. If the file is a stale one then
7. TTR = TTR_{old} / β. // β is a constant. we need to update replica of a file.
8. if Time to refresh rate(TTR) is greater then maximum or less then minimum TTR then
9. TTR = max(TTR_{min}, min(TTR_{max}, TTR))
10. if time to refresh rate is less than or equal to query then TTR_{poll} = T_{query}
11. else
12. TTR_{poll} = TTR

when TTR > T_{query}, that is, the file is queried at a higher rate than change rate, then the file should be updated timely based on TTR. As a result, TTR_{poll} should be calculated based on the following formula [1]

$$TTR_{poll} = \begin{cases} T_{query} & TTR \le T_{query}, \\ TTR & TTR > T_{query}. \end{cases}$$

Fig 3. File Upload And Download Using Peer



CONCLUSION

This paper proposes the combined approach of file replication and consistency maintenance which is highly efficient at low cost. Instead of accepting replicas and updates, nodes determine the need for replicas based on file query rate and update rate. This approach guarantees high utilization of replicas, high query efficiency and accurate consistency maintenance. It reduces redundant file replicas, consistency maintenance overhead, and unnecessary file updates.

Replica node polls the file owner , this might not be true that all the requesters of the file can have up-to-date files. although its performance is better when compared to other file consistency maintenance algorithms..

REFERENCES

[1]. Haiying (Helen) Shen,"IRM: Integrated File Replication and Consistency Maintenance in P2P Systems", Parallel and Distributed Systems, IEEE Transactions on Jan 2010

A. Rowstron and P. Druschel, "Storage Management and Caching in PAST, a Large-Scale, Persistent Peer-to-Peer Storage Utility," Proc. ACM Symp. Operating Systems Principles (SOSP), 2001.

[2]. F. Dabek, M.F. Kaashoek, D. Karger, R. Morris, and I. Stocia,"Wide Area Cooperative Storage with CFS," Proc. ACM Symp. Operating Systems Principles (SOSP), 2001.

[3]. T. Stading, P. Maniatis, and M. Baker, "Peer-to-Peer Caching Schemes to Address Flash Crowds," Proc. First Int'l Workshop Peerto- Peer Systems (IPTPS), 2002.

[4]. V. Gopalakrishnan, B. Silaghi, B. Bhattacharjee, and P. Keleher, "Adaptive Replication in Peer-to-Peer Systems," Proc. 24th Int'l Conf. Distributed Computing Systems (ICDCS), 2004.

[5]. Gnutella Home Page, http://www.gnutella.com, 2008.

[6]. S. Tewari and L. Kleinrock, "Analysis of Search and Replication in Unstructured Peer-to-Peer Networks," Proc. ACM SIGMETRICS, 2005.

[7]. H. Shen, "EAD: An Efficient and Adaptive Decentralized File Replication Algorithm in P2P File Sharing Systems," Proc. Eighth Int'l Conf. Peer-to-Peer Computing (P2P '08), 2008.

[8]. G. Xie, Z. Li, and Z. Li, "Efficient and Scalable Consistency Maintenance for Heterogeneous Peer-to-Peer Systems," IEEE Trans. Parallel and Distributed Systems, vol. 19, no. 12, pp. 1695-1708, Dec. 2008.

[9]. G. Xie, Z. Li, and Z. Li, "Efficient and Scalable Consistency Maintenance for Heterogeneous Peer-to-Peer Systems," IEEE Trans. Parallel and Distributed Systems, vol. 19, no. 12, pp. 1695-1708, Dec. 2008.