International Journal of Computer Applications Technology and Research
Volume 2– Issue 3, 374 - 377, 2013

Comparative Analysis of File Transfer Algorithms to Reduce Average Download Time in Peer to Peer Networks

Anjali Malekar
CSE Department
RCET Bhilai

Neelabh Sao
CSE Department
RCET Bhilai

Abstract: File sharing using the peer-to-peer techniques are very effective and popular among Internet users. The peer-to-peer (P2P) network is used for downloading large-volume contents such as movies and software. To address the problem of flexibility and scalability, we propose a distributed peer-to-peer architecture and algorithms for data clustering. In general the file download can take more time depending on the level of network congestion or the service capacity fluctuations. In this paper we proposed distributed random chunk based periodic switching algorithm for file distribution. This algorithm remove the limitation of the existing systems i.e. Parallel downloading, Chunk based switching, periodic switching, dynamically distributed parallel periodic switching algorithm thus reduce the average download time. In this method corresponding file is divided into many small chunks, peers select their targets for downloading in a random fashion, thus increase the average download time in P2P networks and then analyze a simple, distributed algorithm to effectively remove these negative factors, thus minimizing the average download time.

Keywords: Peer-to-Peer Networks, Distributed Content Sharing, average download time.

1. INTRODUCTION

Peer to peer (P2P) network is widely used for content sharing application. Peer to peer networks are Decentralized networks, it improves the capability of resources sharing, and maximizes the utilization of resources. In this paper we analyze various algorithms [1] for average download time in a P2P network with multiple competing downloading peers. We investigate the relationship between the average download time, in peer to peer networks. More recent research in the direction of evaluating P2P systems has focused on performance. Peer selection schemes were evaluated in [2], where measurements are used to optimize the selection of good peers and improve the overall system performance. The work in [3] is also evaluate system performance in peer to peer network, and define Variable Chunk based switching scheme in which the file to be downloaded is divided into many chunks, but in this scheme a user downloads the chunks sequentially one at a time. Early work [5], [6] are study on clustering technique for average download time in peer to peer networks. The approach of using the average service capacity to analyze the average download time has been a common practice in the literature [7],[8],[9],[10].

1.1 Peer to peer network:

A peer to peer (P2P) computer network is type of network in which each workstation has equivalent capabilities and responsibilities. Each computer in this network can act as a client or server. Client means placing a request (i.e.) client is a running application programs on a local site that requests service from a running application program on a remote site. Server means a program that can provide services to others program. Peer to peer network can be set up within the home, a business, or over the Internet. Peer to peer network can be used for sharing content such as audio, video, data, or anything in digital format. P2P is a distributed application architecture that partitions tasks or workloads among peers. Peers are equally privileged when a P2P network is established over the Internet a distributed network can be established where the sharing of files is split between all the users in the network that are storing a given file.

P2P network is different from client-server network. Client-server network involves multiple clients connecting a single, central server. Content distribution is a centralized one, where the content is distributed from the centralized server to all clients requesting the document. Clients send request to the centralized server for downloading the file. Server accepts the request and sends the file as response to the request. In most client-server setups, the server is a dedicated computer whose entire purpose is to distribute files. There are many drawbacks of client-server network such as scalability problem arises when multi requests arises at a single time, servers need heavy processing power, downloading takes hours when clients increases, requires heavy storage in case of multimedia content.

Content providers that use peer to peer network can benefit from a cost-effective distribution of content to thousands of simultaneous users, both Internet-wide and in private networks.P2P content distribution is only viable by satisfying the requirements of both the content providers and the end-users. The advantage of peer-to-peer networking is the easier control concept not requiring any additional coordination entity and not delaying transfers by routing via server entities. In this network, software applications can be installed on the single computer and shared by every computer in the network.

This network allows reducing the traffic at a central server by delegating a fraction of it to the server's clients. Peer-to-peer architectures can be very effective in reducing the server traffic, to an extent that a large user population can be supported while requiring only limited resources at the server side. Unlike traditional distributed computing, P2P networks aggregate large number of computers and possibly mobile or handheld devices, which join and leave the network frequently. Nodes in a P2P network, called peers, play a variety of roles in their interaction with other peers. When accessing information, they are clients. When serving information, they are server.

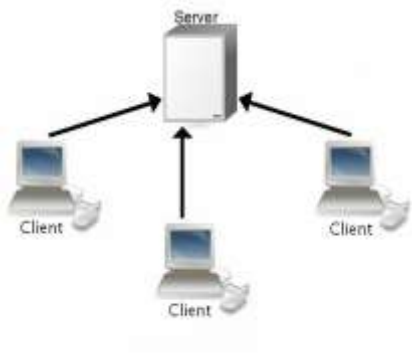

Fig 1 Client Server architecture

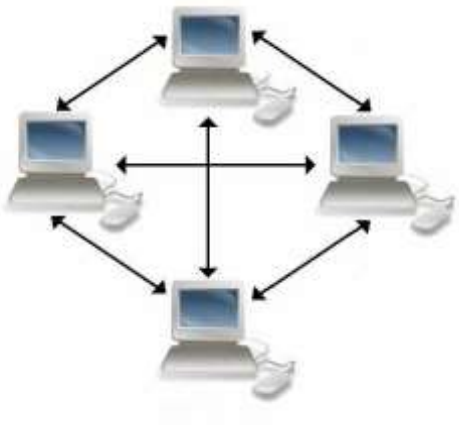

Fig 2 Peer to peer architecture

2. CLUSTERING AND MEDIA STREAMING

In this paper we proposed clustering technique for file splitting in order to reduce transfer delay. In clustered P2P model divides the sharing file into many chunks and Transmits the Chunks to each peer in the cluster. In this paper many algorithms are used for file splitting such as random chunk based switching algorithm, variable chunk based switching algorithm and etc.In chunk based switching algorithm downloaded file is divided in to many chunks either sequentially or randomly.

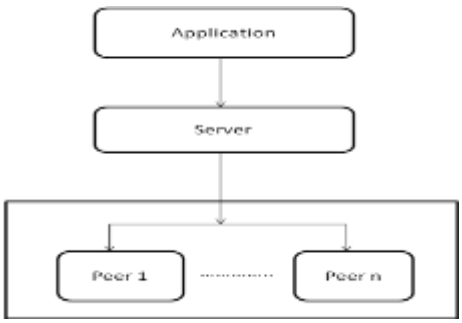

Fig 3- Architecture of server clustering

Streaming is a process of generating and delivering a steady, isochronous flow of data packets over networking medium, e.g., the Internet, from a source to a destination. Streaming media usually denotes digital audio and video data; however an alternative that alleviates the bandwidth cost problem by offering a service to deliver continuous media streams directly between peer nodes. Examples of streaming systems that may be used to distribute stored content are Narada, HMTP, and Pastry. Media streams are generated by live sources (e.g., cameras and microphones) and the data is forwarded to other nodes in real-time. A streaming process can be separated into three stages that overlap in time (Figure 1): data acquisition, data delivery and data presentation. Data acquisition is the stage that determines how the streaming content is acquired, packetized and distributed for streaming. The data presentation stage represents the methods on how to buffer, assemble and render the received data. Data delivery is the process of how the stream data is transported from the source to the destination. The source, the destination and all the intermediate nodes in a streaming system participate in a topology that is constructed based on the specific system's protocol. In a P2P streaming system, this network architecture exhibits peer-to-peer characteristics.

3. TECHNIQUES FOR FILE DOWNLOADING IN PEER TO PEER NETWORKS

In this paper we present various Techniques for content distribution .The goal of these techniques is reduce file downloading time in peer to peer network. Suppose that a downloading peer wants to download a file of size F from N possible source peers. Let ci be the average end-to-end available capacity between the downloading peer and the source peer (i = 1, 2, . . . ,N). The actual value of ci is unknown before the downloading peer actually connects to the source peer i.The average service capacity of the network, A(ĉ). The actual value of ci is unknown before the downloading peer actually connects to the source peer i

3.1 Parallel downloading

Parallel downloading is one of the ways to reduce the download time. In parallel download, a file is divided into k chunks of equal size and single file is allowed to download in parallel with simultaneous connections. Parallel downloading is better than single downloading. In the network with single one user, parallel downloading may not reduce the download time up to the mark. The parallel downloading may perform well if the capacity of each possible source peer is known so as to allocate larger chunks to faster connections and smaller chunks to slower connections.

Downloading a file of size F from the network, average service capacity, the download time T is given by

$$T = F / A(\hat{c})$$

$$= \frac{FN}{\sum_{i=0}^{N} c_i}$$

Where $$\hat{c} = \frac{1}{N} \sum_{i=0}^{N} c_i$$

3.2 Random chunk Based Switching

In the random chunk-based switching scheme, the file to be downloaded is divided into many small chunks. In this

method the user randomly selects a new source peer and connects to it to retrieve a new chunk. The download time for one chunk is independent of that of the previous chunk. The switching of source peers based on chunk can reduce the correlation in service capacity between chunks. A file of size F is divided into m chunks of equal size, and let be the download time for chunk. Then, the total downloads time

$$T_{chunk} = \sum_{j=0}^{m} \frac{1}{N} \sum_{i=0}^{N} \frac{F/m}{c_i}$$

3.3 Random Periodic Switching

In this method the downloader randomly chooses a source peer at each time slot, independently of everything else. It is observed that both the spatial heterogeneity and the temporal correlation in the service capacity can significantly increase the average download time of the users in the network. Random periodic switching strategy will always reduce the average download time and that the average download time under the random periodic switching is given by F/(ĉ).Random periodic switching removes the negative impact of both the heterogeneity and the correlations. Our algorithm is extremely simple and does not require any information about the system.

4. Limitation of the Existing methods

In Parallel Downloading, if the downloader stuck with any one of the bad source peer over k peers, then it waits for long time until getting the chunk. The download time of this method is the maximum time taken by any of the k peers that take the longest time to complete. The main disadvantage of the Chunk-Based Switching is that if we get stuck in a bad source peer with very low service capacity, downloading a fixed amount of bytes from that source peer may still take a In Random Periodic Switching, the downloader randomly chooses a source peer at each time slot and it may get stuck with bad source peer. So this method cause too much overhead associated with switching to many source peers and integrating those many chunks into a single file

4.1 Dynamically Distributed Parallel Periodic Switching

Dynamically Distributed Parallel Periodic Switching algorithm that effectively removes correlations in the capacity fluctuation and the heterogeneity in space, thus greatly reducing the average download time. Dynamically Distributed Parallel Periodic Switching algorithm has two methods (i) parallel connection (ii) parallel random chunk based switching. In first method the source selection function does not change in fixed time slot t, but instead of choosing a single source peer, here the downloader chooses multiple fixed source peers over possible source peers and it makes permanent connection for the fixed time slot t.In second method the source selection function changes for each randomly selected time slot but instead of choosing a single source peer, here the downloader randomly chooses multiple fixed source peers over possible source peers and it makes parallel connection with that k source peers for each randomly selected time slot.

4.2 Distributed variable chunk based parallel switching

In this method chunk size changes with time and provides no. Of parallel connections which changes with hardware utilization which overcomes of Chunk Based Switching algorithm. This method also removes correlations in the capacity fluctuation and the heterogeneity in space. If bandwidth available is increased then downloading can complete before specified time. If bandwidth available is decreased then downloader will search another peer with good bandwidth and get it replaced. After downloading all chunks from the all sources, the system will check whether the entire file got downloaded or not. But this method is not cost effective.

Another limitation of existing algorithm

1) We find that the existing algorithm has follows limitation-

2) Source selection function randomly selects source peer but that source may be with low bandwidth.

3) A source has more bandwidth but downloader cannot utilize this bandwidth.

4) Random chunk based switching is sequential approach.

5) The main disadvantage of random chunk based switching is that, if we get stuck in a source peer with very low service capacity, downloading a fix amount of bytes from that source peer may take a long time..

5. PREPOSED ALGORITHM

In this paper we propose distributed random periodic switching algorithm. In this method, the source selection function changes for each randomly selected time slot as in simple Random Chunk Based Switching of existing Periodic Switching. But instead of choosing a single source peer, here the downloader randomly chooses multiple peers. In this algorithm, the file to be downloaded is divided into many small chunks. Whenever a user completes a chunk from its current source peer, the user randomly selects a new source peer and connects to it to retrieve a new chunk .The switching of source peers based on chunk can reduce the correlation in service capacity between chunks.

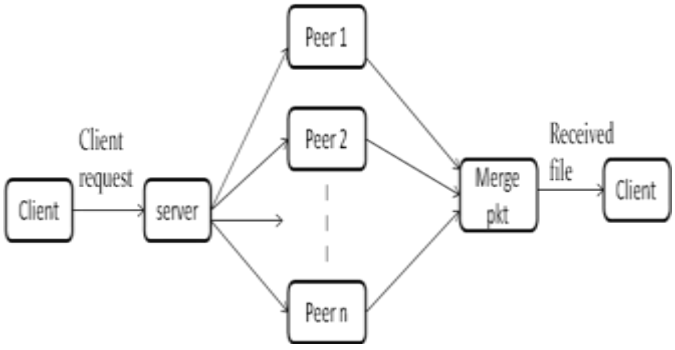

Fig. 4 clustered peer to peer (P2P) architecture

Algorithm

- Input: Node-servers network capacity at node i=[0,1……n]
- Output: X, The total time to complete download of entire k mb file.
- X ← 0
- // file that has been broken into m kb chunks
- For j = 0 to m-1
- int rand =random(0,n)
- int ChunkDownLoadTime = (k MB * 1024KB)/bandwidth
- X = X + ChunkDownLoadTime
- End for
- X = X / 60 \\convert seconds into minutes
- return X
- Exit

Different strategies have different impact on the average download time of each peer, which may result in different system dynamics as well, e.g., how fast a downloader can start to contribute (become a source peer) or how fast a peer leaves the system after finishing download. The switching of source peers based on chunk can reduce the correlation in service capacity between chunks. The benefit of chunk-based switching conversion from downloading peers to uploading Peers and thus indirectly affect the average download time

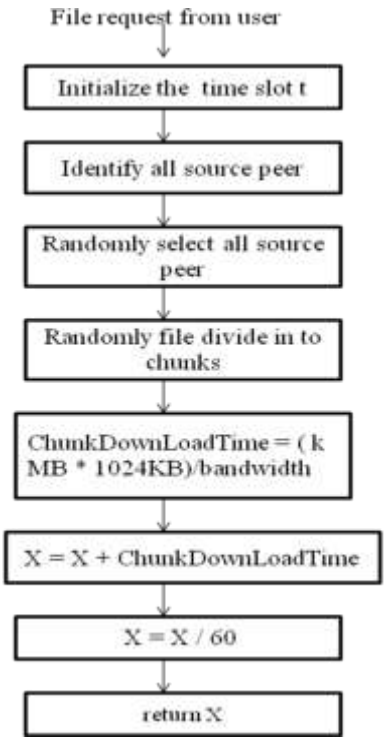

Fig 5- Process of File Splitting

6. CONCLUSION

In this paper, we discussed peer-to-peer systems that have been deployed in file sharing and media streaming. Services to content providers and consumers but also need we investigated existing P2P based file sharing and media streaming applications, and discussed the limitations of their implementations. We studied various techniques that might help improveP2P performance. We have compared several schemes widely used in practice, including chunk-based file transfer, parallel downloading and have shown that all those schemes are not so effective for reducing average download time In this paper we have designed a clustering technique that effectively reduces the average download time. We observed that our new proposed distributed random periodic switching algorithm is performing better than previous technique. From both experimental and analytical viewpoints, we have concluded that proposed model improved the system performance and reduced the network traffic.

7. REFERENCES

[1] P. Satheesh, B. Srinivas, M. V. S. Narayana, "Analysis of Distributed algorithms to Remove Correlations for Reducing Average Download Time in Peer-to-Peer Networks", International Journal of Computer Applications Volume 51– No.10, August 2012.

[2] Kolan Helini, P. Prashanthi and G. Radha Devi, "Analysis of Peer-to-Peer Networks and Study of File Download Minimization", International Journal of Multidisciplinary Science and Engineering, Vol.3, No.1, January2012

[3] Amit Chougule1, Shambhuraj Deshmukh ,"Variable Chunk Based Parallel Switching To Minimizing File Download Time in P2P Network", IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 4, No 2, July 2011

[4] Michael J. Neely , Leana Golubchik,"Utility Optimization for Dynamic Peer-to-PeerNetworks with Tit-For-Tat Constraint", PROC. IEEE INFOCOM, 2011.

[5] Guiyi Wei, Yun Ling, Ye Gu, Yujia Ge," A Dependable Cluster-based Topology in P2P Networks", Journal of communication, vol. 5, no. 1, JANUARY 2010 57

[6] S.Ayyasamy1 and S.N. Sivanandam " Cluster Based Replication Architecture for Load Balancing in Peer-to-Peer Content Distribution", International Journal of Computer Networks & Communications (IJCNC) Vol.2, No.5, September 2010.

[7] Prof. Rakesh Mohanty , Prof. H. S. Behera, Khusbu Patwari,Monisha DashM. Lakshmi Prasanna ,"Priority Based Dynamic Round Robin (PBDRR) Algorithm with Intelligent Time Slice for Soft Real Time Systems", International Journal of Advanced Computer Science and Applications,Vol. 2, No.2, February 2011.

[8] Peer selection Algorithm in Stochastic Content Delivery Networks to Reduce File Download Time by Michael Lehrfeld A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Computer Information Systems Graduate School of Computer and Information Sciences Nova Southeastern, university 2009.

[9] Yuh-Ming Chiu and Do Young Eun," Minimizing File Download Time over Stochastic Channels in Peer-to-Peer Networks,",2008.

[10] [10] Opportunities and Challenges of Peer-to-Peer Internet Video Broadcast by Jiangchuan Liu, Member IEEE, Sanjay G. Rao, Bo Li, Senior Member IEEE, and Hui Zhang , Vol. 96, No. 1, January 2008.