

July 2010

## Distributed Shared Files Management

Saurabh Malgaonkar

*Thadomal Shahani College Of Engineering, Bandra(W), Mumbai University,*  
saurabhmalgaoonkar@gmail.com

Sakshi Geeta Surve

*Thadomal Shahani College Of Engineering, Bandra(W), Mumbai University,* geetams24@rediffmail.com

Manish Bhelande

*KC College of Engineering, Thane(E),* manishbhelande@gmail.com

Follow this and additional works at: <https://www.interscience.in/imr>



Part of the [Business Administration, Management, and Operations Commons](#), and the [Human Resources Management Commons](#)

---

### Recommended Citation

Malgaonkar, Saurabh; Surve, Sakshi Geeta; and Bhelande, Manish (2010) "Distributed Shared Files Management," *Interscience Management Review*. Vol. 3 : Iss. 2 , Article 9.

DOI: 10.47893/IMR.2010.1062

Available at: <https://www.interscience.in/imr/vol3/iss2/9>

This Article is brought to you for free and open access by the Interscience Journals at Interscience Research Network. It has been accepted for inclusion in Interscience Management Review by an authorized editor of Interscience Research Network. For more information, please contact [sritampatnaik@gmail.com](mailto:sritampatnaik@gmail.com).

# Distributed Shared Files Management

Saurabh Malgaonkar<sup>1</sup>, Sakshi “Geeta” Surve<sup>2</sup> & Manish Bhelande<sup>3</sup>

<sup>1&2</sup>Thadomal Shahani College Of Engineering,Bandra(W),Mumbai University,

<sup>3</sup>KC College of Engineering,Thane(E)

E-mail : <sup>1</sup> saurabhmilgaonkar@gmail.com, <sup>2</sup>geetams24@rediffmail.com, <sup>3</sup>manishbhelande@gmail.com

---

**Abstract** – Most often file sharing is the common and basic requirement when users work on a particular domain or area of interest. Users can use software that connects in to a peer-to-peer network to access shared files on the computers of other users (i.e. peers) connected to the network. Files of interest can then be downloaded directly from other users on the network. So this concept is similar to a distributed file system where files are distributed across the network but the users have an illusion of a centralized file system and also avoids its high complexity and cost of implementation. Emphasis is more on location and access transparency.

**Keywords** – P2P File Access; Network Sharing; DistributedShared File System; Large File Sharing.

---

## I. INTRODUCTION

In a computer system a file is a named object that comes into existence by explicit creation, is immune to temporary failures in the system and persists until explicitly destroyed. The two main purposes of using files are as follows:

1. Permanent storage of information.
2. Sharing of information.

A user creates many files on his machine and updates them accordingly if required. Access made to those files depends on the requirements of the user. Most often file sharing is the common and basic requirement when users work on a particular domain or area of interest. Example, in a hospital, the records of the patients are essential in every department so they are shared accordingly on the system. So users need to share the files that are necessary and access those shared files quickly. So file sharing becomes the practice of distributing or providing access to digitally stored information such as computer programs, multimedia (audio, images, and video), documents, or electronic books. It may be implemented through a variety of ways. Common methods of storage, transmission, and distribution[1] used in file sharing include manual sharing using removable media, centralized server on computer networks, World Wide Web-based hyperlinked documents and the use of distributed peer-to-peer networking[1][2].

Users can use software that connects in to a peer-to-peer network to access shared files on the computers of other users (i.e. peers) connected to the network. Files of interest can then be downloaded directly from other users on the network. So this concept is similar to a distributed file system where files are distributed across the network but the users have an illusion of a centralized file system and also avoids its high complexity and cost of implementation.

The most common and feasible approach is to use peer to peer file sharing[2] for implementing a distributed shared files management system. In addition to these advantages it will also enable to support the following:

1. Remote Information Sharing: It will enable to access to information that is being shared by a remote machine.
2. User Mobility: As the system will reflect all the files shared by the nodes present in the system, user can access them from anywhere.
3. Availability: For better fault tolerance, the systems shared file entries are available to the users even in the temporary failure of the main directory controller.

The file sharing domain is necessary and distributed thus need further justification on using peer to peer technologies[2] on that domain. Peer to peer file sharing is economically efficient. When the user wants to find specific information, searching for the same would require a lot of human

efforts and time. If the upcoming technologies are clubbed with the existing ones it can help better understand the whole system. Thus extending the idea of peer to peer in the file sharing environment helps better built the whole system. For accomplishing this task a directory server is used to better organize the user shared files information in the related domain.

The main advantages of this approach are:

1. Scalability: It can easily accommodate more users and hence making it more scalable.
2. Bandwidth: It will enable to save the network bandwidth as only the required files when required are transmitted among the users.
3. Distributed control: In this, there is a need for a central point which we name as a controller, which will manage various shared file lists from all the clients and will be handling the shared file lists distribution scenario as per the clients entry or exit in the network.
4. Fault Tolerance: The plan is to make this file sharing scenario fault tolerant so a replica controller will be maintained that will be frequently updated from the primary controller so even if the primary controller fails the system continues to operate.

The goal of this project is not only to achieve a distributed shared files management system that will allow clients that are distributed location wise to share files among themselves but also to give better performance in terms of file access.

A diagrammatic interaction of the user with the system is as follows:

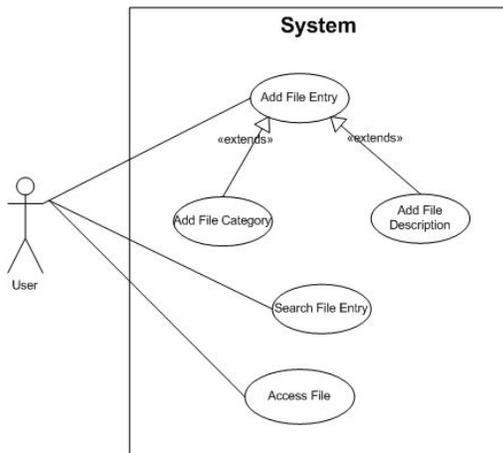


Fig. 1 : User Interaction with the System

When a user interacts with the system when joined to a particular network, the user adds the file entry that needs to be shared among the clients in the network. The user mentions the category of the file and also adds its description so the other users are aware about the contents of the file. A user can also search for a particular file entry from the required parameters (name, category or description) as the users in a network can share hundreds of file entries and it is impossible to look for a particular entry manually. Once the user finds the required file entry, the user with its help can access the file by receiving it from the client who is sharing that particular file. A common interface in each client also lists the overall shared file entries of all the clients in the network for access along with the search mechanism.

## II. TYPICAL P2P DISTRIBUTED FILE SHARING SYSTEM

We present a literature review which includes the basic file sharing details, distributed file system, various technologies, existing communication protocols etc. The development of file sharing system has triggered two to three decades ago.

As the various file applications grew from hundreds to thousands to millions, the interest along with the resources for file sharing. File sharing began in 1999 with the introduction of Napster[3][4], a file sharing program and directory server that linked people who had files with those who requested files. The central index server was meant to index all of the current users and to search their computers. When someone searched for a file, the server would find all of the available copies of that file and present them to the user. The files would be transferred between the two private computers. While Napster connected users through a directory server, these new services connected users remotely to each other. These services also allowed users to download files other than music, such as movies and games. One limitation was that only music files could be shared. After Napster shut down the most popular of these new services was Gnutella[5].

### A. P2P System Advantages

In [7] the key features in P2P file transfer are highlighted. Using P2P technique, execute nodes can share data with each other instead of fetching files only from the central manager or a file server, saving plenty of time. In a typical file sharing environment which follows a centralized approach where the files need to be shared are transferred to the central server and the nodes accessing them from

it has its great disadvantage. In case of the failure of the central server the entire system comes to a standstill and when the system is operating on heavy load the central server becomes a bottleneck. Also it does not reflect the updated file entries, so when a user updates a file, it needs to be uploaded again to the central server. So this paper suggest a peer to peer sharing where all the major disadvantage of a centralized sharing system have been taken care off.

### B. A Basic P2P File Sharing System

In [8] a reliable and simple P2P file sharing system is described which avoids unnecessary data redundancy and connectivity issues among peers by maintaining an adapter which optimizes the working of the entire file sharing system. This approach also makes it highly scalable in nature. Since a file resides at local node and is shared only when required there is no need for separate update policy, the shared file itself reflects the updates version of the file all the time as the user updates it. The adapter maintains a list of all the files that are shared by the users and updates it accordingly then sends it to all the users. When this list is imported by the users they can perform file sharing operations among themselves.

### C. Reliability and fault tolerance

The disadvantages of a client server file system which do not scale with respect to the number of users and exhibit a single point failure are further highlighted in [6]. So the focus is more on the distributed peer to peer aspect rather than a centralized one. The important aspect of this paper is fault tolerance achieved by replicating data, hence the data being available even in case of temporary failures. The other aspect to be considered is user mobility.

## III. DSFM DESIGN

It is necessary to keep the system in a constant flow and achieve the targeted goals of the proposed system at the same time. So the after studying carefully all the literatures [6][7][8] and highlighting the key features and drawbacks from them, the overall workflow of the proposed system will be as follows:

The three most important components of the system are:

1. Client: The client system will allow the clients to share the files that are required as well as provide an interface that will allow the client to access the files

shared by all the nodes in the system(centralized view of shared files). When a client joins the network its shared files will be added to the system and when the client leaves the network all his shared files entries are discarded from the system. After a client receives the updated shared files list, it can access the file from the respective clients.

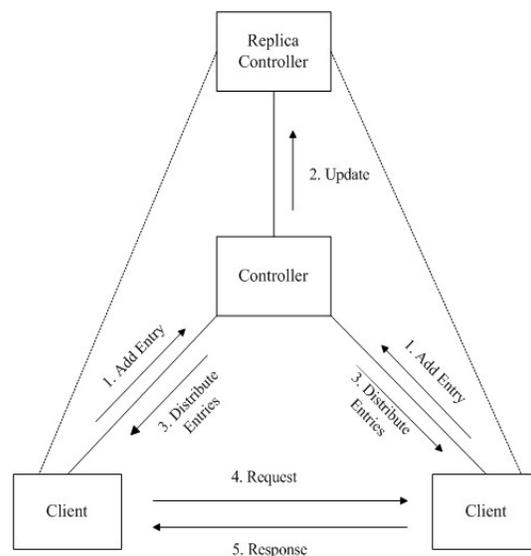


Fig. 2: Overall System Workflow

2. Controller: The main controller will store the address and shared files information of the client peers and will be responsible for distributing them to all the clients in the network. It's basic task is just to index the file entries from all the clients and distribute them accordingly. Also when a client joins or leaves the network it will update its shared files entries accordingly and inform the remaining clients. So through the controller we will be able to achieve the scenario of overall file sharing.

3. Replica Controller: The project plans to replicate the main controller so even when the main module fails temporary the system does not comes to a standstill, the requests can be handled from the replica controller and the system continues to operate. So our system will be fault tolerant.

### A. General Basic Functionality

The following diagram denotes the basic interaction among the modules and functionality of each module. This is the normal scenario highlighted when the system is working with the

primary controller when fully functional.

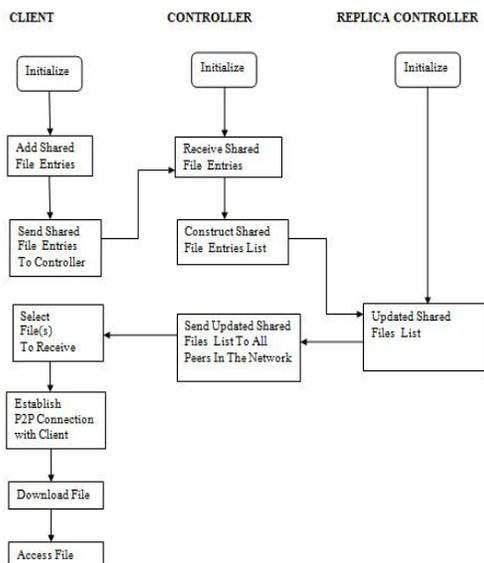


Fig. 3 : General Flow Chart

**B. File Sharing Process**

The following diagram illustrates the scenario that enables to achieve the basic file sharing process among the various clients in the network.

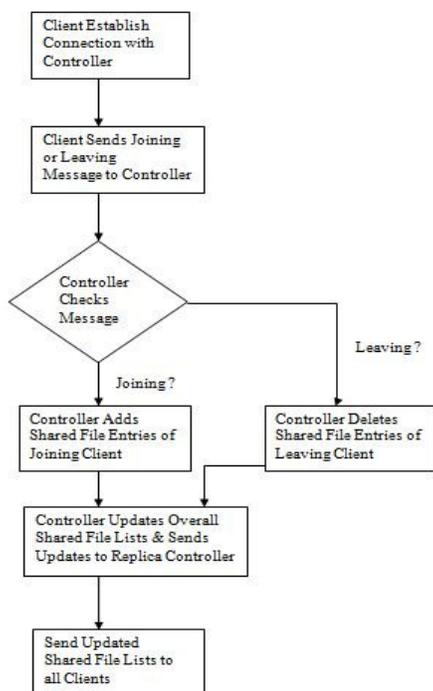


Fig. 4 : Basic File Sharing Process

**C. Fault Tolerance Scenario**

The following diagram illustrates the scenario about how the clients detect the temporary no response or failure of the primary controller and redirect to the replica controller.

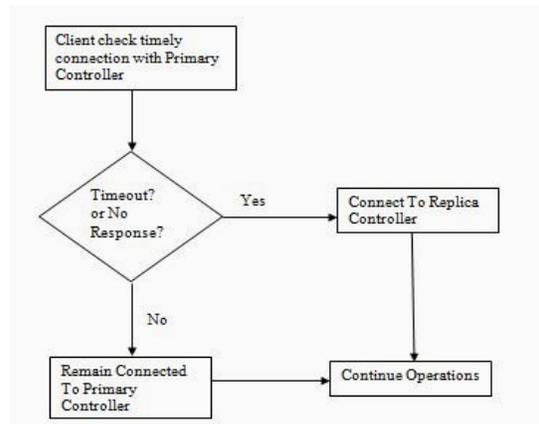


Fig. 5 : Handling Primary Controller Failure

**REFERENCES**

- [1] Kenneth P. Birman (2005), "Reliable distributed systems".
- [2] Rüdiger Schollmeier, "A Definition of Peer-to-Peer Networking for the Classification of Peer-to-Peer Architectures and Applications", Proceedings of the First International Conference on Peer-to-Peer Computing IEEE, pp.149-160, 2002.
- [3] Menta, Richard, "Napster Clones Crush Napster. Take 6 out of the Top 10 Downloads CNet", MP3Newswire, <http://www.mp3newswire.net/stories/2001/topclones.html>, July 20, 2011.
- [4] Ante, Spencer, "Inside Napster". Business Week. Retrieved 2011-04-10, [http://www.businessweek.com/2000/00\\_33/b3694001.htm](http://www.businessweek.com/2000/00_33/b3694001.htm), June 7, 2011.
- [5] "Gnutella File Sharing", <http://slashdot.org/story/00/03/14/0949234/Open-Source-Napster-Gnutella>, 15 August 2011.
- [6] Sunil Chakravarthy and Chitranjan Hota, "Secure Resilient High Performance File System for Distributed Systems", IEEE International Conference on Computer & Communication Technology, pp. 87-92, 2010.
- [7] Gao Ying, Liu Guan Yao and Huang JianCong, "A New Method of File Transfer In Computational Grid Using P2P Technique", IEEE International Conference on Network Computing and Information Security, pp. 332-336, 2011.
- [8] Rui Zhao, Ruhua Liu and Guangxuan Fu, " P2P File Sharing Software in IPv4/IPv6 IEEE International Conference on Software and networks, pp. 367-370, 2011.