Table of Contents

A good distributed file system provides the following types of services
1. Storage service
2. True file service
3. Name/Directory service

Desirable features of a good distributed file system
1. Transparency
2. User mobility
3. Performance
4. Simplicity and ease of use
5. Scalability
6. High availability
7. High reliability
8. Security
9. Heterogeneity

Two main purposes of using files:

1. Permanent storage of information on a secondary storage media.
2. Sharing of information between applications.

A file system is a subsystem of the operating system that performs file management activities such as organization, storing, retrieval, naming, sharing, and protection of files.

A good distributed file system provides the following types of services

1. Storage service

Allocation and management of space on a secondary storage device thus providing a logical view of the storage system.

2. True file service

Includes file-sharing semantics, file-caching mechanism, file replication mechanism, concurrency control, multiple copy update protocol etc.

3. Name/Directory service

Responsible for directory related activities such as creation and deletion of directories, adding a new file to a directory, deleting a file from a directory, changing the name of a file, moving a file from one directory to another etc.

Desirable features of a good distributed file system

1. Transparency

- Structure transparency: Clients should not know the number or locations of file servers and the storage devices.
- Access transparency: Both local and remote files should be accessible in the same way.
- Naming transparency: The name of the file should give no hint as to the location of the file. The name of the file must not be changed when moving from one node to another.
- Replication transparency: The clients do not need to know the existence or locations of multiple file copies.

2. User mobility

User should not force to work on a specific node but should have the flexibility to work on different nodes at different times.

3. Performance

The performance of the file system is usually measured as the average amount of time needed to satisfy client requests.

4. Simplicity and ease of use

User interface to the file system be simple and number of commands should be as small as possible.

5. Scalability

A good distributed file system should be designed to easily cope with the growth of nodes and users in the system.

6. High availability

A distributed file system should continue to function in the face of partial failures such as a link failure, a node failure, or a storage device crash. A highly reliable and scalable distributed file system should have multiple and independent file servers controlling multiple and independent storage devices.

7. High reliability

Probability of loss of stored data should be minimized. System should automatically generate backup copies of critical files.

8. Security

Users should be confident of the privacy of their data.

9. Heterogeneity

There should be easy access to shared data on diverse platforms (e.g. Unix workstation, Wintel platform etc).