

Replicated Document Management System

Himanshu Saraiya - 18111023

I G Prasad - 18111025

September 27, 2018

1 Introduction

The basic idea behind Replicated Document Management System is to provide a group communication system that is used by people to share textual and numeric information. In this project we will create a system that operates on personal computers in local-area and wide-area networks, and provides end-users the ability to design and create document databases for specific applications. The system supports group of people working on a set of documents and all of them are not connected to the system all the time. Each person may either work in isolation (offline) on a subset of documents or may be online and modifying/updating some documents(However assumption is that files are not frequently modified and communication usually takes place by adding new files rather than modifying the existing one). The proposed system is based on replication and decentralization of the documents , which improves efficiency and reliability.

2 Application characteristics

The primary Goal of the System is to provide storage and retrieval of the databases of documents with high availability using replication.

Interfaces Provided

1. Databases can be created, opened, closed,and deleted.
2. Documents can be created, opened, updated, closed, and deleted
3. Retrieval of Documents

Replication Design

Each node will have a replica of database and will run the replication algorithm independently. Replication algorithm assures Eventual consistency, which is a form of weak consistency.

1. **Background Replication :** The replication is background activity that doesn't interfere with normal database access.
2. **Works For Rarely Connection Network:** Supports replication of documents on servers that are only rarely in communication with each other. Immediate update/communication between server to server is not required.
3. **Convergence:** The databases will be converged over all the replicas over time.
4. No master copy and no database replica directory is maintained.

References

- [1] Leonard Kawell, Jr., Steven Beckhardt, Timothy Halvorsen, Raymond Ozzie, and Irene Greif. 1988. Replicated document management in a group communication system. In Proceedings of the 1988 ACM conference on Computer-supported cooperative work (CSCW 88). ACM, New York, NY, USA, 395.