UNIVERSITY OF ONTARIO
INSTITUTE OF TECHNOLOGY

Project LiveSync

Distributed Systems
Prashakar Prabagaran
100489428

[Abstract]

Project LiveSync addresses the issue of file synchronization between platforms. Today, many people use multiple devices to access the files and data they need on a daily basis. We are no longer bound to one device and in-fact shuffle through various devices throughout the day. Why is it that we alternate between devices but the files most important to us generally dont? This is the problem that Project LiveSync hopes to solve.

LiveSync is a service that will run on any given system. It synchronizes files across your various systems. It allows you to make changes to a given set of files and have those changes pushed across all your devices. This process is done without any user intervention except for the initial setup of sync directories. LiveSync is a distributed file system based service that offers an easy way to distribute or synchronize files.

This project addresses the overall problem of distributing useful files across various devices. The LiveSync process synchronizes files that are within a user specified directory. Project objectives include: 1) development of a user-friendly and seamless synchronization service where the user is required to perform a quick set-up and no to very little user intervention should be required post-setup; 2) fast file synchronization between all client devices, speed of file synchronization should be bottlenecked by network throughput as opposed to the processing time of the client services; 3) reliable synchronizations that ensures all files within the specified sync directory is safely synchronized and no issues arise in the sync process; 4) maintain file system integrity as much as possible which entails that when a user requests to keep filesystem directory structures identical, LiveSync ensures that the files found under a given directory tree of one client is also found under the nearly same directory tree of another client.

Once the setup is complete the files that were once available on a remote system will now be available on the localhost under the sync directory that was specified. An option will also exist to preserve directory structure, in which case an attempt will be made to synchronize files in a directory with a filesystem structure identical to the remote system. LiveSync enables users to seamlessly make changes on one device and see those changes applied immediately on another remote device. The user can set up LiveSync to synchronize across more than two devices which allows for even greater accessibility.

LiveSync will use a form of indirect communication. The synchronization service will be time coupled which means after a client is shut down or goes offline at a given time and then re-connects, it will not gain access to the file updates that it missed while offline. The missed file updates will not be stored and sent to the offline client. Instead, the client that goes offline will simply stop receiving any further syncs until the LiveSync service is online. Ideally this service should begin listening for updates immediately after the device is online again.

Unlike a reliable multicast solution, if a given client is not online or is not accepting synchronizations for whatever reason, the sync will still occur for all other clients. The reason for this is because LiveSync is a synchronization services that attempts to provide the most seamless experience possible, in this case live updates for as many clients but is also a best-effort service. It will make its best effort to sync across all the clients however, if a given client is not online, it will still push changes out to all remaining online clients. If this wasn't the case, all the clients will not receive live updates and then all clients will be out-of-date. With this best-effort approach, LiveSync ensures that it is updating as many

clients as possible and therefore provides a more reliable and seamless experience to the remaining online clients.

Synchronizations occur based on last modified time of a file within the specified directory. LiveSync will periodically scan the filesystem for changes to the file. It does this by retrieving the last modified time of each file, notifying the server of an updated file, and the server compares to the last modified times on all other clients (this data is stored on the server). As such, periodic scans will be executed on all clients running LiveSync and when a client detects an update that data will be sent and stored within the server. When the server notices that a specific client has a newer version of the file (based on last modified time), it will pull that file from the client with the newer version and distribute it among the other clients. A restriction here is that files that are intended to be synced cannot exceed a set size. If it exceeds that size, the sync operation will not affect that specific file. This will be made aware to the users upon first synchronization. A notification will outline all files that were exempt from synchronization.

LiveSync requires one server to be running; the LiveSync Server and any number of client devices can be running LiveSync Client. All client devices will communicate with the server. Clients monitor the local filesystem every 5 minutes then detect if any updates have occurred. Instead of the server polling the clients, the clients will invoke a method on the server to check if the local file believed to be the newest is truly the last updated file (check from data stored on the server) and the server will call a pull file method if it is necessary and then distribute file method to sync the new file across all clients. LiveSync will be developed using Python.