

Research Proposal

Title: Multi User File Management System Based on Apples Time Machine Using FSEvents

Author: Tracy Footitt

Supervisor: Dr Chris McDonald

Background

With the release of Mac OS X Leopard, comes Time Machine, a breakthrough in backup and recovery systems [1]. Time Machine is an automatic backup system, which was introduced by Apple in an attempt to encourage more people to backup their data [5]. After an initial full backup to a specified device, Time Machine will automatically make hourly incremental backups. These backups are consolidated into daily backups kept for a month and as many weekly backups as will fit onto the drive. When the drive begins to fill up, the backups are analysed to delete selected files in order to preserve as wide a range of backups as possible. The only way a file will not be included in a backup is if it was both created and destroyed in between hourly backups.

One benefit of Time Machine is that if a hard drive dies or is replaced, a full restore is possible from the latest Time Machine backup. The other main benefit is the ability to view the contents of a folder at any point in the past for which files in that folder have existed. This enables files to be recovered from the past to either replace or coexist with the current file. Other benefits associated with an automatic backup system, are that backups are not reliant on users specifying when backups should be performed.

Time Machine, like most incremental backup systems, only creates backups of files that have been modified since the last backup was performed [2]. In order to enable browsing of the entire backup, Time Machine relies on hard-links to the unchanged files [5]. However Time Machine gains a significant speed advantage of other backup systems through its use of FSEvents. This means that instead of having to traverse the entire file system and compare the files with the latest backup, it can simply read from a list all of the directories where changes have been made since the last backup.

Time Machine is a simple and easy to use as a tool for users who generally dont bother with a backup strategy. However one feature that is not included in Time

Machine is the ability to perform a backup at a time that is nominated by the user as opposed to only scheduled backups [3]. Whilst a backup system needs to be automatic to ensure that it is implemented, a voluntary component would enable a user to perform a backup in between those scheduled.

Aim

The aim of this project is to develop a backup and recovery system similar to Time Machine, that is suitable for the needs of multiple users, in order to improve the management of the UWA CSSE file system. Whilst Time Machine is ideal for single users, it does not scale well for multiple users. In environments such as UWA, there is often the need for multiple users to have a copy of the same file. E.g. all students enrolled in a particular course could download and save a copy of the lecture notes, resulting in multiple copies of the same file being stored on the system. There may also never be a need for these files to be modified resulting in them only ever being read from. Therefore there is the need for a system that reduces/removes the redundancy of storing multiple copies of the same file. In a University environment especially, there are numerous files which are owned by multiple users. The goal is to create a management system where each file is stored once for multiple users rather than multiple times.

It will need to be determined which users have ownership or access rights to a file to make sure that only users with the correct permissions are given access to a file. If a user modifies a file, then a new file is created which only that user has access to. There is also a need within a University environment to enable groups of students to be able to work on a set of files, whilst each knowing which files they have modified.

The main benefit of incremental backups to students, is that more backups are available for the current day (when problems are most likely to be discovered) in addition to backups for previous days.

It is hoped that a working system can put into place in the UWA CSSE Undergraduate labs to run alongside the current system. This will allow for testing to ensure that files are only accessible by users with the correct permissions, and that sufficient backups are taken and maintained.

Method

1. Research and Familiarisation. It is envisaged that the initial stage in the project will be research and familiarisation with past work in the area along with ideas that will be beneficial in completing the project. Ideally this research will be conducted within the first couple of weeks so that an understanding of further requirements of the project can be developed.
 - (a) Research of and Familiarisation with Time Machine. Because the proposed system incorporates several features from Time Machine, it is important to understand the features offered by Time Machine, the features

that could have been included, and the techniques that are employed within Time Machine. This includes research into FSEvents, aliases, and Mac OS X [4].

- (b) Research of other backup systems past and present. In order to gain an understanding of the requirements of the proposed system, it is required to investigate previous and current systems. The purpose of this research is to gain an understanding of features that can be included in the proposed system and features that may not be necessary.
 - (c) Familiarisation with Time Capsule.
2. Implementation of backup system. This stage corresponds to the actual construction of the proposed system. This will require determining and implementing the required features of the system.
 3. Installation of backup system in CSSE Undergrad labs alongside existing system. This refers to the actual installation of the proposed system, whilst maintaining the current system so that both are able to run in parallel. The aim for this stage is to have the installation in place for the start of the second semester.
 4. Evaluation of backup system. Once the system has been installed, it needs to be evaluated to determine if it is performing correctly and also to determine how well it is working.
 - (a) Ease of use. Test the ease of use of the system and how intuitive it is.
 - (b) Robustness and Effectiveness. Determine the effectiveness of the system and compare it to the current system. In order to be effective, the system needs to be able to provide sufficient backups to cover most requirements, whilst only permitting file access to those users who have the required permissions and minimising redundant storage of files. Test the Robustness of the system at different times and under different conditions.

Software and Hardware Requirements

This project will require the following items, all of which will be supplied by the school.

- Apple Quad Xeon 64-bit server
- Apple iMac desktop machine
- Time Capsule NAS

References

- [1] V. Vellanki & Z. Kurmas. A. L. Chervenak. Protecting file systems: A survey of backup techniques.

- [2] Apple. Time machine: a giant leap backward.
- [3] J. Fingas. Setting up time machine: as simple as possible. 2007.
- [4] A. Singh. Mac os x internals: A systems approach. 2001.
- [5] J. Siracusa. Mac os x 10.5 leopard: the ars technica review. page 14, 2008.