InterMezzo File System: Synchronizing Folder Collections
Stelias Computing Inc

Author: Peter J. Braam PhD, braam@stelias.com, http://www.stelias.com.

Introduction

Today's computer users are increasingly facing a daily problem keeping collections of documents "in sync" between their computers. The most well known example is the laptop/desktop problem where users have to determine which of the two computers contains the latest version of a collection of documents.

During manual synchronization of collections of folders and files many hazards face the user. Documents might easily get overwritten, the latest versions of documents may accidentally fail to arrive on a mobile computer before disconnection or many duplicates might be created, which gradually lead to chaotic situations in folder collections.

Manual synchronization or the lack thereof is very costly to organizations.

The InterMezzo file system has been designed to address synchronization problems in a scalable, efficient and secure manner.

InterMezzo's mechanisms

InterMezzo is a client-server file system capable of **journaling versions and updates** to folder collections, both while network connectivity is present and during **disconnected operation**. Unlike other network file systems, InterMezzo wraps around an existing disk file system, with a kernel level module named **Presto** (Linux) or **Vivace** (Windows).
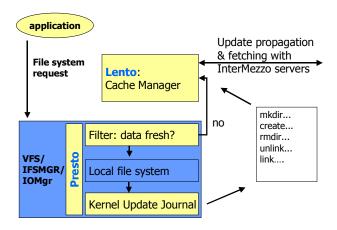

Figure 1: InterMezzo filtering

Presto intercepts updates made to folders or files, and writes journal records that detail both what was done and which version of the folder or file was affected by the update.   InterMezzo can also intercept accesses to data and fetch newer versions of folders and documents.

When more than one client is *connected* to the server, InterMezzo will arbitrate updates through the acquisition of **write tokens.**   At most one system can hold a write token for a given folder collection, and only this system can make updates to the collection. This is done in order to avoid **conflicting updates.** Updates will be propagated immediately to all connected systems that are authorized and registered for update forwarding.

The result is that folder collections on several computers are kept in sync during periods of connectivity.   This results in collections of files and folders, just as if these were shared and updated from a single storage location.

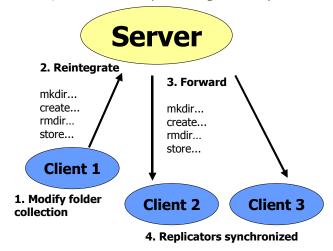However, InterMezzo's capabilities go well beyond this.

Figure 2: Update forwarding

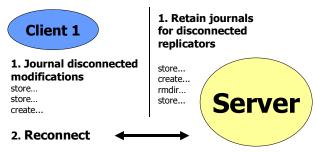## Disconnected operation & reintegration

When a client computer stops being connected to the server holding the authoritative version of the folder collection, InterMezzo switches to **disconnected operation.** This can happen, for example, when a mobile client is taken offline, or when a network or server fails. In disconnected mode, InterMezzo allows the client full access to the folder collection and continues to journal updates and versions. This provides **high availability** to the data. Of course, the disconnected system can no longer be sure that its updates of documents are unique, since the

activities of other clients are not available during disconnection.

When connectivity becomes available again, InterMezzo synchronizes changes stored in journals among client and servers as follows. InterMezzo will start **update propagation** to forward changes already stored on servers but not yet available on the re-connected client. Subsequently, InterMezzo will **reintegrate** updates made on the client to the servers. When update propagation and reintegration have completed the folder collections will be synchronized.

During update propagation it is possible that a **conflicting update** is detected. This happens when the same object (folder or file) was modified on a disconnected client and on a client that has already reintegrated an update to the server. When a conflict is detected, InterMezzo interrupts update propagation to the newly connected client until the **conflict resolution policy** has completed. During conflict resolution, only the re-connecting client is affected.

InterMezzo does not enforce a particular conflict resolution policy. Conflict handlers can use the InterMezzo conflict API to enable a graphical interface for conflict resolution, or alternatively activate a totally automatic tool to deal with the conflict, depending on requirements.

Client 1

1. Journal disconnected modifications
store…
store…
create…

1. Retain journals for disconnected replicators

store…
create…
rmdir…
store…

Server

2. Reconnect

    a. Server forwards modification journals
    b. Handle conflicts
    c. Reintegrate client journals

3. Client and server synchronized

Figure 3: InterMezzo reconnection sequence

Related Technology

InterMezzo was started at Carnegie Mellon University as part of the Coda project, see http://www.coda.cs.cmu.edu. Coda heavily influenced our design decisions. InterMezzo does not offer all features of Coda but is much more modular and easier to integrate with existing file systems.

Microsoft is offering IntelliMirror with Windows 2000. The first release of this technology has features substantially similar to those of InterMezzo. Its design is completely different. InterMezzo can inter-operate smoothly with IntelliMirror.

Solutions based on running synchronization tools, http or ftp downloads tend to be less scalable, involve user intervention and have no or limited concurrent access capabilities due to lack of lock and version management.

Summary

InterMezzo is a file system providing synchronization and high availability of folder collections.

InterMezzo's features comprise:

- **Update reintegration and forwarding:** when updates to files and folders are made a journal is automatically maintained which describes these updates in sufficient detail.
- During **connectivity** the updates in the journal are reintegrated to a server in the background immediately after they have been made. They are then propagated to individual clients authorized and registered for replicating certain folder collections as soon as these clients are available on the network.
- InterMezzo supports **disconnected operation** during periods of failed networks or voluntary disconnection of mobile client computers.
- **Servers** in InterMezzo engage in **update forwarding** to clients that have registered with the servers as **replicators** of folder collections.
- InterMezzo can handle **conflicting updates** and provides an API through which custom graphical user interfaces can extend InterMezzo's core technology.
- **Secure TCP services**, such as ssl, firewalls and proxies, can be integrated with InterMezzo's networking infrastructure to provide a highly secure file sharing environment.

Licensing & availability

Linux versions of InterMezzo will be licensed under the GNU General Public License. InterMezzo can easily be ported to other operating systems, including Windows 9x, NT and 2000.

See http://www.inter-mezzo.org for further information, binaries, sources and documentation for InterMezzo. Red Hat Software and Los Alamos National Laboratory are contributing to InterMezzo's development.

Contact Information

Stelias Computing Inc., was founded in 1995, and engages in consulting and software development in the area of file and storage systems, with an emphasis on Linux. We can be reached at:

filesystems@stelias.com
Phone (403) 678 6910, Fax (403) 678 6922
8580 Magnolia Trail, # 315
Minneapolis, MN 55344