|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 简单文件同步器 V1.0   * 设计说明书  |  |  | | --- | --- | | 版本 | 1.0 | | 文档状态: | 完成 | | 作者: | 刘奕鑫 | | 负责人: | 刘奕鑫 | | 创建日期: | 2021年5月23日 | | 更新日期: | 2021年5月23日 | |

目录

[1 简介 2](#_Toc12674)

[1.1 开发目的 2](#_Toc8291)

[1.2 产品功能 2](#_Toc11383)

[1.3 模块描述 2](#_Toc29442)

[2 设计说明 3](#_Toc16447)

[2.1 设计思路 3](#_Toc4706)

[2.2 实现思路 4](#_Toc22806)

# 简介

## 开发目的

课程作业，实现本地目录与S3远程桶的文件同步

## 产品功能

简单的文件同步功能：

1. 使本地某个目录与S3某个Bucket同步；
2. 程序启动时把Bucket的文件同步到本地，需要处理文件冲突；
3. 本地添加、修改了文件，要传到S3；本地删除了文件，也要删除S3的文件；
4. 对于超过20MB的文件，需要使用分块上传下来，需要支持中断传输、恢复传输进度的功能。

## 模块描述

软件主要包括如下模块：

1. 文件管理模块：

负责从远程服务器下载文件到本地、从本地上传文件到远程，需要支持大文件（超过20MB）时使用分块上传下来，需要支持中断传输、恢复传输进度的功能。支持本地和远程的文件删除。

1. 文件同步模块：

启动时下拉远程服务器最新内容，需要解决冲突处理；程序运行中间，需要定期地将更新的文件传到远程。

# 设计说明

## 设计思路

软件的处理流程主要分为两部分：**启动时的下拉同步、运行过程中的实时同步更新**，下面分别展开叙述。

**启动时的下拉同步**

在启动时，软件会检查远程指定桶的所有文件，并将它们与本地指定目录下的文件进行对比。当远程存在本地没有的文件时，下拉到本地。当存在本地有但远程无的文件时， 删除本地的文件；当存在两边都有的同名文件，根据以下方式对比更新，对比逻辑为：根据最后修改时间，来判别远程与本地文件的新旧关系。将远程较为新的文件下拉到本地，本地较为新的文件需要进行文件冲突管理，在程序中我们采用了暂停进行人工查验的方式。

**运行过程中的实时同步**

下拉完成后，程序进入周期性循环，通过对比远程和本地的文件列表的区别，来执行更新操作。若本地添加或者修改了文件，则需要更新上传到远程。若本地删除了，也需要在本地删除。具体而言，按照以下方式实现监控：对比本地和远程文件的列表，添加的情况会对应到本地列表多了一些文件，修改的情况会对应到本地列表某些文件的最后修改时间相比于远程的时间有所变晚，删除的情况会对应本地列表少了一些文件。

## 实现思路

* 使本地某个目录与S3某个Bucket同步；

- 配置好参数文件

* 程序动时把Bucket的文件同步到本地，需要处理文件冲突；

- 启动时，通过getObjectSummaries接口读取远程Bucket的”远程文件列表”L1，然后对比本地目录的”本地文件列表”L2，执行以下的操作：

创建临时文件列表L3，遍历L1的每个文件F1i，若L2中存在同名文件F2i，则对比两者：

1. 若F1i的更新时间早于F2i，说明本地文件更新，需要进行冲突处理，需要手工进行处理
2. 若F1i的更新时间晚于F2i，说明远程文件更新，直接将F1i放到L3
3. 若F1i的更新时间等于F2i，说明远程文件和本地版本一致，不处理

处理完成后，得到了最新文件列表L3， 此时将其与L2对比，然后将文件更新至本地。

* 本地添加、修改了文件，要传到S3；本地删除了文件，也要删除S3的文件；

- 下拉完成后，程序进入周期监控循环，对远程Bucket的”远程文件列表”L1和本地目录的”本地文件列表”L2每个监控循环执行以下的条件判断：

1. 若本地L2中多了远程L1没有的文件，则说明添加了文件，传到S3上；
2. 若本地L2中同名文件的最后修改时间变得更晚，则说明修改了文件，覆盖传到S3上。
3. 若本地L2中少了一些在L1中有的文件，则说明删除了文件，删除S3上的特定文件。

* 对于超过20MB的文件，需要使用分块上传下来，需要支持中断传输、恢复传输进度的功能。

- 对比超过20MB的文件，我们编写了MultipartDownload、MultipartUpload。