OSS_PIPE: Rust编写的大规模文件迁移工具

背景

随着业务的发展,文件数量和文件大小会急剧增加,文件迁移的数量和难度不断攀升。oss_pipe 是rust编写的文件迁移工具,旨在支撑大规模的文件迁移场景。

编写 oss_pipe 的初衷

- 同类产品面临的问题
- rust 语言带来的技术红利
- oss_pipe 的基本功能

常见的 oss 迁移工具

- ossimport 阿里出品,java语言编写,支持存量增量同步,支持大部分云场的oss数据源
- ossutil 阿里出品,go语言编写,迁移只是辅助功能,主要是阿里对象存储的管理客户端
- COS Migration 腾讯出品,java语言编写 文档中未见增量迁移部分

	出品方	开发语言	可执行文件size
ossimport	阿里	java	单机版16MB
COS Migration	腾讯	java	31MB
Kodoimport	七牛	golang	31MB
oss pipe	 京东	rust	12MB

同类产品面临的问题

- 编程语言方面,java golang 这些带vm的语言容易产生OOM
- 功能相对齐全的工具,如ossimport需要部署java runtime环境,安装步骤繁琐
- 运行期间不仅需要条件本身参数还需要、复杂场景下还需要对jvm进行调优、对现场工程师要求较高

rust 语言带来的技术红利

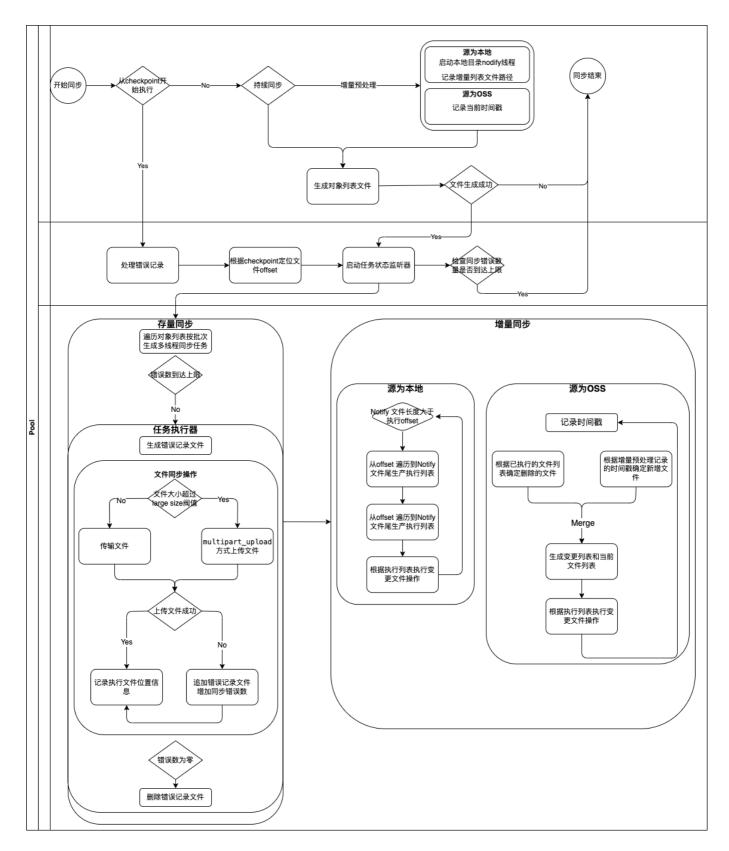
- 内存安全性:通过所有权、借用和生命周期机制,Rust 在编译时确保内存安全,无需垃圾回收。
- 性能:接近 C/C++ 的性能,无额外运行时开销,适合系统级编程。
- 并发处理: 所有权模型简化并发编程, 减少数据共享和复杂性。
- 可靠性和生产力: 强类型系统、模式匹配和丰富的工具链提高代码质量和开发效率。
- 多平台支持:支持多种操作系统和硬件平台,便于跨环境部署。
- 现代语言特性: 结合函数式和泛型编程, 同时保持与传统系统编程语言的兼容性。
- 社区和生态系统: 活跃的社区和成熟的生态系统, 支持快速的语言发展和项目构建。

oss_pipe 的基本功能

• 主要功能

- 。 全量迁移
- 。 存量迁移
- 。 增量迁移
- 。 断点续传
- 。 大文件拆分上传
- 。 正则表达式过滤
- 。 线程数与上传快大小组合控制带宽
- 存储适配及支持列表
 - 。 京东云对象存储
 - 。 阿里云对象存储
 - 。 腾讯云对象存储
 - 。 华为云对象存储
 - 。 AWS对象存储
 - o Minio
 - 。 本地

实现机制



演示

- 文件生成
- 文件上传
- oss 间同步

oss 工具及形态的理解

- 管理客户端TUI
- 迁移工具