

# OSS PIPE

---

OSS PIPE 是一个用于 S3 兼容对象存储间进行文件迁移的工具。

下载地址: [https://github.com/jiashiwen/oss\\_pipe/releases/download/0.1.0/oss\\_pipe\\_x86\\_64-apple-darwin](https://github.com/jiashiwen/oss_pipe/releases/download/0.1.0/oss_pipe_x86_64-apple-darwin)

[https://github.com/jiashiwen/oss\\_pipe/releases/download/0.1.0/oss\\_pipe\\_x86\\_64-linux](https://github.com/jiashiwen/oss_pipe/releases/download/0.1.0/oss_pipe_x86_64-linux)

## 主要功能

- Transfer 对象存储间文件迁移, 比如 AWS 到 阿里云的迁移, 或从阿里云到华为云的迁移
- Download 从对象存储下载到本地
- Upload 本地目录上传到OSS
- LocalToLocal 本地多线程复制
- TruncateBucket 清空Bucket (慎用!)
- OssCompare object 比较

## Getting Stated

How to build

- 安装rust编译环境

```
curl --proto '=https' --tlsv1.2 -sSf https://sh.rustup.rs | sh
```

- 构建

```
apt install -y pkg-config
```

```
git clone https://github.com/jiashiwen/oss_pipe.git
cd oss_pipe
cargo build --release
```

## 基本使用

oss\_pipe 支持命令执行模式和交互式执行模式。您可以通过 `oss_pipe [subcommand]` 执行执行任务, 比如

```
oss_pipe parameters provider
```

也可以通过 `oss_pipe -i` 命里进入交互模式。交互模式支持按 'tab' 键进行子命令提示。

## 定义任务

通过 `oss_pipe template` 命令生成模板

```
oss_pipe template transfer /tmp/transfer.yml
```

transfer.yml 文件内容

```
task_id: '7071651847576096769'
name: transfer task
task_desc: !Transfer
source:
  provider: ALI
  access_key_id: access_key_id
  secret_access_key: secret_access_key
  endpoint: http://oss-cn-beijing.aliyuncs.com
  region: cn-north-1
  bucket: bucket_name
  prefix: test/samples/
target:
  provider: JD
  access_key_id: access_key_id
  secret_access_key: secret_access_key
  endpoint: http://s3.cn-north-1.jdcloud-oss.com
  region: cn-north-1
  bucket: bucket_name
  prefix: test/samples/
bach_size: 100
task_threads: 12
max_errors: 1
meta_dir: /tmp/meta_dir
target_exists_skip: false
start_from_checkpoint: false
large_file_size: 104857600
multi_part_chunk: 10485760
exclude: null
include: null
```

修改 `access_key_id` `secret_access_key` 等参数，适配自己的任务。`template` 命令按照任务类型创建模版, 模板描述请参考[参考手册](#)。`parameters` 支持参数查询，包括支持的provider 以及 任务类型

## 执行任务

`osstask` 子命令用于执行任务

```
oss_pipe osstask filepath/task.yml
```

# 参考手册

---

## 命令详解

oss\_pipe 同时支持命令行模式和交互模式 oss\_pipe -i 进入交互模式。交互模式使用'tab'键进行子命令提示。

- osstask

通过yaml描述文件执行相关任务

- 命令格式

```
osstask <filepath>
```

- 命令行模式示例

```
oss_pipe osstask yourpath/exec.yml
```

- 交互模式示例

```
oss_pipe> osstask yourpath/exec.yml
```

- template

生成任务模板,通过子命令指定任务类型

- 命令格式

```
template [subcommand] [file]
```

- 命令行模式示例

```
template transfer /tmp/transfer.yml
```

- 交互模式示例

```
oss_pipe> template transfer /tmp/transfer.yml
```

- parameters

参数查询, 输出所支持的 oss 供应商, 以及任务类型。

- 命令格式

```
parameters [subcommand]
```

- 命令行模式示例

```
parameters provider
```

- 交互模式示例

```
oss_pipe> template parameters task_type
```

- tree  
显示命令树。

- 命令格式

```
tree
```

- exit  
退出交互模式。

- 命令格式

```
exit
```

# Oss 提供商支持

提供商代码	提供商描述
AWS	Amazon Web Services
ALI	阿里云
JD	京东云
HUAWEI	华为云

# 任务类型

Download

```

# 任务id, 非必填
task_id: '7064075547986497537'
# 任务名称, 非必填
name: download task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !Download
# 源对象存储
source:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 下载目标目录
local_path: /tmp
# 批次大小, 既每批下载文件数据量
batch_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
# 大文件拆分时每块大小, 单位为Bytes
multi_part_chunk: 10485760

```

## Transfer

```

# 任务id, 非必填
task_id: '7064079488245698561'
# 任务名称, 非必填
name: transfer task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !Transfer
# 源对象存储

```

```

source:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: Ali
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://oss-cn-beijing.aliyuncs.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 目标对象存储
target:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 批次大小, 既每批下载文件数据量
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
# 大文件拆分时每块大小, 单位为Bytes
multi_part_chunk: 10485760

```

## Upload

```

# 任务id, 非必填
task_id: '7064079488245698561'
# 任务名称, 非必填
name: upload task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !Upload
# 上传目标目录
local_path: /tmp
# 目标对象存储
target:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 批次大小, 既每批上传文件数据量
batch_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
# 大文件拆分时每块大小, 单位为Bytes
multi_part_chunk: 10485760

```

## LocalToLocal

```

# 任务id, 非必填
task_id: '7064087275835101185'
# 任务名称, 非必填
name: local to local task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !LocalToLocal
# 源目录

```

```

source_path: /tmp/source
# 目标目录
target_path: /tmp/target
# 批次大小, 既每批复制文件数据量
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
# 大文件拆分时每块大小, 单位为Bytes
multi_part_chunk: 104857

```

## TruncateBucket

```

# 任务id, 非必填
task_id: '7064088180835880961'
# 任务名称, 非必填
name: truncate bucket task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !TruncateBucket
# 要清理的bucket描述
oss:
  # 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
  provider: JD
  # 对象存储 access_key
  access_key_id: access_key_id
  # 对象存储 secret_access
  secret_access_key: secret_access_key
  # 对象存储endpoint, 请参考供应商相关文档
  endpoint: http://s3.cn-north-1.jdcloud-oss.com
  # 对象存储区域, 请参考供应商相关文档
  region: cn-north-1
  # 对象存储 bucket
  bucket: bucket_name
  # 对象存储prefix
  prefix: test/samples/
# 批次大小, 既每批删除文件数据量
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1

```



```
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
```

## OssCompare

```
# 任务id, 非必填
task_id: '7064090414587973633'
# 任务名称, 非必填
name: oss compare task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !OssCompare
# 源对象存储
source:
  # 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 查找
  provider: Ali
  # 对象存储 access_key
  access_key_id: access_key_id
  # 对象存储 secret_access
  secret_access_key: secret_access_key
  # 对象存储endpoint, 请参考供应商相关文档
  endpoint: http://oss-cn-beijing.aliyuncs.com
  # 对象存储区域, 请参考供应商相关文档
  region: cn-north-1
  # 对象存储 bucket
  bucket: bucket_name
  # 对象存储prefix
  prefix: test/samples/
# 目标对象存储
target:
  # 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 查找
  provider: JD
  # 对象存储 access_key
  access_key_id: access_key_id
  # 对象存储 secret_access
  secret_access_key: secret_access_key
  # 对象存储endpoint, 请参考供应商相关文档
  endpoint: http://s3.cn-north-1.jdcloud-oss.com
  # 对象存储区域, 请参考供应商相关文档
  region: cn-north-1
  # 对象存储 bucket
  bucket: bucket_name
  # 对象存储prefix
  prefix: test/samples/
# 批次大小, 既每批下载文件数据量
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
```

```
# 对象过期时间差，当差值小于指定值时被视为过期时间一至  
exprirs_diff_scope: 10  
# 从checkpoint开始执行  
start_from_checkpoint: false
```

## 参考手册

### 命令详解

oss\_pipe 同时支持命令行模式和交互模式 oss\_pipe -i 进入交互模式。交互模式使用'tab'键进行子命令提示。

- osstask  
通过yaml描述文件执行相关任务

- 命令格式

```
osstask <filepath>
```

- 命令行模式示例

```
oss_pipe osstask yourpath/exec.yml
```

- 交互模式示例

```
oss_pipe> osstask yourpath/exec.yml
```

- template  
生成任务模板,通过子命令指定任务类型

- 命令格式

```
template [subcommand] [file]
```

- 命令行模式示例

```
template transfer /tmp/transfer.yml
```

- 交互模式示例

```
oss_pipe> template transfer /tmp/transfer.yml
```

- parameters  
参数查询，输出所支持的 oss 供应商，以及任务类型。
  - 命令格式

```
parameters [subcommand]
```

- 命令行模式示例

```
parameters provider
```

- 交互模式示例

```
oss_pipe> template parameters task_type
```

- tree  
显示命令树。

- 命令格式

```
tree
```

- exit  
退出交互模式。

- 命令格式

```
exit
```

Oss 提供商支持

提供商代码	提供商描述
AWS	Amazon Web Services
ALI	阿里云
JD	京东云

## 提供商代码    提供商描述

HUAWEI	华为云
--------	-----

## 任务类型

### Download

```
# 任务id, 非必填
task_id: '7064075547986497537'
# 任务名称, 非必填
name: download task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !Download
# 源对象存储
source:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 下载目标目录
local_path: /tmp
# 批次大小, 既每批下载文件数据量
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
# 大文件拆分时每块大小, 单位为Bytes
multi_part_chunk: 10485760
```

### Transfer

```
# 任务id, 非必填
task_id: '7064079488245698561'
# 任务名称, 非必填
name: transfer task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !Transfer
# 源对象存储
source:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: Ali
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://oss-cn-beijing.aliyuncs.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 目标对象存储
target:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 批次大小, 既每批下载文件数据量
batch_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
```

```
# 大文件拆分时每块大小, 单位为Bytes  
multi_part_chunk: 10485760
```

## Upload

```
# 任务id, 非必填  
task_id: '7064079488245698561'  
# 任务名称, 非必填  
name: upload task  
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找  
task_desc: !Upload  
# 上传目标目录  
local_path: /tmp  
# 目标对象存储  
target:  
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找  
provider: JD  
# 对象存储 access_key  
access_key_id: access_key_id  
# 对象存储 secret_access  
secret_access_key: secret_access_key  
# 对象存储endpoint, 请参考供应商相关文档  
endpoint: http://s3.cn-north-1.jdcloud-oss.com  
# 对象存储区域, 请参考供应商相关文档  
region: cn-north-1  
# 对象存储 bucket  
bucket: bucket_name  
# 对象存储prefix  
prefix: test/samples/  
# 批次大小, 既每批上传文件数据量  
batch_size: 100  
# 并发数量  
task_threads: 12  
# 任务允许的最大错误数, 达到最大错误数量则任务停止  
max_errors: 1  
# 任务元数据存储位置  
meta_dir: /tmp/meta_dir  
# 当目标文件已存在时是否下载  
target_exists_skip: false  
# 从checkpoint开始执行  
start_from_checkpoint: false  
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes  
large_file_size: 104857600  
# 大文件拆分时每块大小, 单位为Bytes  
multi_part_chunk: 10485760
```

## LocalToLocal

```

# 任务id, 非必填
task_id: '7064087275835101185'
# 任务名称, 非必填
name: local to local task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !LocalToLocal
# 源目录
source_path: /tmp/source
# 目标目录
target_path: /tmp/target
# 批次大小, 既每批复制文件数据量
batch_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 当目标文件已存在时是否下载
target_exists_skip: false
# 从checkpoint开始执行
start_from_checkpoint: false
# 大文件定义, 当文件超过该值时分段下载, 单位为Bytes
large_file_size: 104857600
# 大文件拆分时每块大小, 单位为Bytes
multi_part_chunk: 104857

```

## TruncateBucket

```

# 任务id, 非必填
task_id: '7064088180835880961'
# 任务名称, 非必填
name: truncate bucket task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !TruncateBucket
# 要清理的bucket描述
oss:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix

```

```
    prefix: test/samples/
# 批次大小, 既每批删除文件数据量
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数, 达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
```

## OssCompare

```
# 任务id, 非必填
task_id: '7064090414587973633'
# 任务名称, 非必填
name: oss compare task
# 任务描述, 支持的任务可以通过 oss_pipe parameters task_type 查找
task_desc: !OssCompare
# 源对象存储
source:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 查找
provider: Ali
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://oss-cn-beijing.aliyuncs.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 目标对象存储
target:
# 对象存储供应商, 支持的供应商可以通过 oss_pipe parameters provider 查找
provider: JD
# 对象存储 access_key
access_key_id: access_key_id
# 对象存储 secret_access
secret_access_key: secret_access_key
# 对象存储endpoint, 请参考供应商相关文档
endpoint: http://s3.cn-north-1.jdcloud-oss.com
# 对象存储区域, 请参考供应商相关文档
region: cn-north-1
# 对象存储 bucket
bucket: bucket_name
# 对象存储prefix
prefix: test/samples/
# 批次大小, 既每批下载文件数据量
```



```
bach_size: 100
# 并发数量
task_threads: 12
# 任务允许的最大错误数，达到最大错误数量则任务停止
max_errors: 1
# 任务元数据存储位置
meta_dir: /tmp/meta_dir
# 对象过期时间差，当差值小于指定值时被视为过期时间一至
exprirs_diff_scope: 10
# 从checkpoint开始执行
start_from_checkpoint: false
```