

Alibaba Cloud

在线迁移服务
AWS S3 迁移教程





文档版本：20220628

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.背景信息	05
2.准备工作	06
3.迁移实施	08
4.后续操作	14
5.地域与域名关系对照表	16

1. 背景信息

本教程讲解如何将AWS S3上的数据迁移到阿里云对象存储OSS。

阿里云在线迁移服务是阿里云提供的存储产品数据通道。使用在线迁移服务，您可以将第三方数据轻松迁移至阿里云对象存储OSS，也可以在对象存储OSS之间进行灵活的数据迁移。

使用在线迁移服务，您只需在控制台填写源数据地址和目标OSS地址信息，并创建迁移任务即可。启动迁移后，您可以通过控制台管理迁移任务，查看迁移进度、流量等信息；也可以生成迁移报告，查看迁移文件列表、错误文件列表。

注意

- 执行在线迁移任务过程中，读取源站数据会产生公网流出流量费，该费用由源站的存储服务提供商在源站收取。
- 在线迁移默认不支持跨境迁移数据（例如，从中国北京的存储空间向美国硅谷的存储空间迁移数据），若您有跨境迁移的需求，您在配置迁移任务前需[提交工单](#)，申请开通创建跨境迁移任务的权限。您需承诺您的业务属于合法经营，数据的流动符合当地法律法规，数据没有违法违规信息。

本教程包含以下内容：

- [准备工作](#)
- [迁移实施](#)
- [后续操作](#)

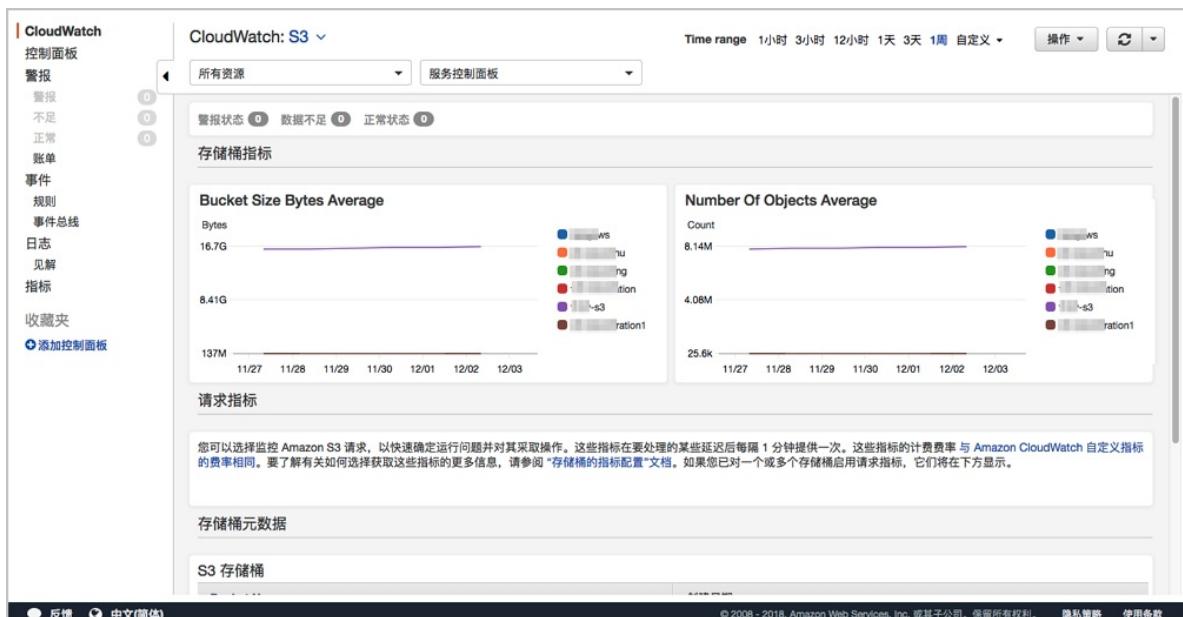
2.准备工作

本文介绍数据迁移之前的准备工作。

AWS S3

- 预估迁移数据

预估需要迁移的数据，包括迁移存储量和迁移文件个数。您可以打开AWS的CloudWatch控制台，选择S3进入，查看您需要迁移的存储桶内文件大小及数量。



说明 为保证您数据的正常迁移，您需要在创建迁移任务时正确填写存储量与对象（文件）数量。

- 解冻源存储桶待迁移数据

对于归档类型的数据，需要您进行解冻操作，待解冻完成后再创建迁移任务。在线迁移服务并不会对源端数据执行解冻操作。若您源端有未解冻、解冻中的文件，则这些文件都会迁移失败。归档类型的数据解冻请参见[归档迁移方案](#)。

说明

- 请您务必完成解冻操作后再创建源数据地址和迁移任务。
- 请您根据待迁移的数据总量确保解冻时长，以防迁移期间数据再次变成冻结状态。
- 解冻操作可能会收取一定的费用，计费规则可咨询您的源存储空间服务运营商。

- 创建用于迁移的访问密钥

您可以在AWS控制台的IAM页面创建用户并授予 `AmazonS3ReadOnlyAccess` 权限，然后创建访问密钥。

说明

迁移服务目前暂不支持AWS S3的白名单访问。如果有白名单配置，请取消白名单或[提交工单](#)。

阿里云对象存储OSS

- 创建目标存储空间（Bucket）

创建目标存储空间，用于存放迁移的数据。具体操作，请参见[创建存储空间](#)。

- 创建RAM用户并授予相关权限

- i. 登录[RAM控制台](#)。
- ii. 在左侧导航栏，选择**身份管理 > 用户**。
- iii. 在用户页面，单击**创建用户**。
- iv. 在**创建用户**页面的用户账号信息区域，输入登录名称和显示名称。
- v. 在访问方式区域，选中控制台访问和OpenAPI调用访问。并保存生成的账号、密码、AccessKey ID和AccessKey Secret。
 - a. 控制台访问：设置控制台密码、需要重置密码和MFA多因素认证。
 - b. OpenAPI调用访问：自动为RAM用户生成访问密钥（AccessKey），支持通过API或其他开发工具进行访问。
- vi. 创建RAM用户成功后，在**用户**页面，单击目标RAM用户操作列的**添加权限**，授予RAM用户存储空间读写权限（AliyunOSSFullAccess）和在线迁移管理权限（AliyunMGWFullAccess）。
- vii. 授权完成后，在左侧导航栏，单击**概览**。
- viii. 在账号管理区域中，单击用户登录地址下的链接，使用刚创建的RAM用户的用户名和密码进行登录。

3. 迁移实施

本文介绍数据迁移的注意事项和操作步骤。

注意事项

在线迁移需注意以下问题：

- 在线迁移会占用源地址和目的地址的网络资源。若您的业务比较重要，建议您对迁移任务设置限速，或在空闲时间启动迁移任务。
- 在线迁移前会检查源地址和目的地址的文件，但是若您源和目的地址有相同文件名的文件，且源地址的文件更新时间比目的地址文件更新时间晚，迁移时会直接覆盖目的地址的文件。若两个文件内容不同，必须更改文件名或做好备份。
- 在线迁移目前仅支持单次迁移单个Bucket数据，无法将整个账号的数据一次性迁移。

步骤一：创建源地址

1. 登录[阿里云数据在线迁移控制台](#)。
2. 选择在线迁移服务 > 数据地址，单击创建数据地址。
3. 在创建数据地址页面，配置如下参数，然后单击确认。

参数	是否必需	说明
数据类型	是	选择AWS S3。
数据名称	是	输入3~63位字符。不支持短划线（-）和下划线（_）之外的特殊字符。
Endpoint	是	填写源地址所在的地域。更多信息，请参见 地域与域名对应关系表 。
Bucket	是	AWS S3空间名称（Bucket名称）。 存储空间名称要求前后不带空格、换行、制表符等非法字符。
Prefix	是	<ul style="list-style-type: none">◦ 迁移全部数据：迁移整个Bucket中的数据。 选择迁移全部数据时，无需填写Prefix。◦ 迁移部分数据：迁移指定目录（前缀）的文件。格式要求不能以正斜线（/）开头，要以正斜线（/）结尾，如 <code>docs/</code>。
AccessKey Id和SecretAccess Key	是	输入用于迁移的访问密钥。迁移完成后删除。

4. （可选）如果系统提示该功能在公测中，请单击去[申请提交白名单权限申请](#)。
5. （可选）填写相关信息，提交迁移公测申请。申请通过后，您将收到短信提醒。

步骤二：创建目的地址

1. 选择在线迁移服务 > 数据地址，然后单击创建数据地址。
2. 在创建数据地址页面，配置相关参数，然后单击确定。

参数	是否必需	说明
数据类型	是	选择OSS。
数据名称	是	输入3~63位字符。不支持短划线 (-) 和下划线 (_) 之外的特殊字符。
数据所在区域	是	选择目的地址所在的地域。
开通并使用传输加速	否	<p>在线迁移服务使用OSS的传输加速服务，需要开通Bucket的传输加速服务。开启传输加速后，会在30分钟内生效，请在30分钟后再创建迁移任务。</p> <p> 注意 开启了传输加速的Bucket会收取传输加速费用。关于传输加速的更多信息，请参见传输加速。</p>
OSS Endpoint	是	<p>根据您的数据所在区域，选择一个Endpoint。更多信息，请参见访问域名</p> <ul style="list-style-type: none"> ◦ http://oss-cn-endpoint.aliyuncs.com：表示使用HTTP协议的外网Endpoint访问OSS。 ◦ https://oss-cn-endpoint.aliyuncs.com：表示使用HTTPS协议的外网Endpoint访问OSS。 <p> 说明 将第三方数据迁移到OSS时，只能选择外网Endpoint访问OSS。</p>
AccessKey Id和AccessKey Secret	是	输入用于迁移的AccessKey。更多信息，请参见 准备工作 。
OSS Bucket	是	选择一个存储空间，用于存储迁移数据。
OSS Prefix	否	<p>格式要求不能以正斜线 (/) 开头，必须以正斜线 (/) 结尾，例如 <code>data/to/oss/</code>。如果为空表示迁移至所选存储空间的根目录。</p> <p> 注意 若您迁移的源地址文件中有以正斜线 (/) 开头的文件名，配置目的地址的时候需要添加一个OSS Prefix，否则会导致迁移失败。例如：需要迁移的文件中包含 <code>/test/test.png</code> 这个文件，您需要添加一个OSS Prefix，如：<code>oss/</code>。当迁移完成后，<code>/test/test.png</code> 的OSS文件名变为 <code>oss//test/test.png</code>。</p>

步骤三：创建迁移任务


1. 选择**在线迁移服务 > 迁移任务**，单击**创建迁移任务**。
2. 在**创建迁移任务**面板，阅读迁移服务条款协议，选中**我理解如上条款，并开通数据迁移服务**，单击**下一步**。
3. 在弹出的**费用提示**对话框，单击**确认**，继续创建。
4. 在**配置任务**页签，配置如下参数，单击**下一步**。

参数	是否必选	说明
任务名称	是	输入3~63位小写字母、数字、短划线(-)，且不能以短划线(-)开头或结尾。
源地址	是	选择已创建的源地址。
目的地址	是	选择已创建的目的地址。
指定目录	否	<p>设置迁移时包含或排除指定目录下的文件和子目录。</p> <ul style="list-style-type: none"> 不过滤：不过滤迁移目录。 排除：迁移时，不迁移排除目录下的文件和子目录。 包含：迁移时，只迁移包含目录下的文件和子目录。 <div> <p> 说明</p> <ul style="list-style-type: none"> 此功能需要开通白名单。更多信息，请参见高级功能使用说明。 目录不能以正斜线(/)或者反斜线(\)开头，并且目录中不能出现两个正斜线(/)，两个半角句号(..)和半角双引号(")，提交的所有目录总字符长度不能超过10 KB。 目录要以正斜线(/)结尾，例如docs/。 最多可设置20个排除目录或者包含目录。 </div>

参数	是否必选	说明
迁移方式	是	<p>选择迁移数据的方式。</p> <ul style="list-style-type: none"> ◦ 全量迁移：根据迁移起点时间迁移一次指定迁移起点时间之后的全量数据，数据迁移完成后任务结束。 <p>如果迁移完成后源数据有变化，您可以再次提交全量迁移任务，系统将仅迁移变化的数据。</p> <ul style="list-style-type: none"> ◦ 增量迁移：按设定的增量迁移间隔和增量迁移次数执行迁移任务。 <ul style="list-style-type: none"> ■ 首次根据迁移起点时间迁移指定迁移起点时间之后的全量数据。首次迁移完成后，按照增量迁移时间间隔执行增量迁移任务，将源地址从前次迁移任务开始后到下次迁移开始前新增或修改的增量数据迁移至目的地址。 ■ 如果配置的增量迁移次数是N，则执行1次全量迁移，之后执行N-1次增量迁移。 <p>例如：设置迁移间隔1小时，迁移次数5次，迁移起点时间为2019-03-05 08:00，当前时间为2019-03-10 08:00。则首次迁移最后修改时间在2019-03-05 08:00~2019-03-10 08:00之间的文件。假设迁移任务1小时完成，第二次迁移则从2019-03-10 10:00（迁移1小时，迁移间隔1小时）开始，迁移最后修改时间在2019-03-10 08:00~2019-03-10 10:00之间的文件，共进行1次全量迁移和4次增量迁移。</p> <div>  注意 全量迁移和增量迁移均会在每次迁移开始前，对源地址和目的地址的文件进行对比。如果迁移同名文件，则以下三种情况的目标地址文件会被覆盖。 <ul style="list-style-type: none"> ◦ 源地址文件与目标地址文件的Content-Type不一致，目标地址文件会被覆盖。 ◦ 源地址文件的最后修改时间晚于目标地址文件的最后修改时间，目标地址文件会被覆盖。 ◦ 源地址文件与目标地址文件的大小不一致，目标地址文件会被覆盖。 </div>

参数	是否必选	说明
迁移起点时间	是	<p>选择迁移数据的起点时间。</p> <ul style="list-style-type: none"> 迁移全部：迁移所有时间的文件。 指定时间：只迁移指定时间之后创建或修改的文件。 <p>例如指定时间设置为2018-11-01 08:00:00，则只迁移2018年11月01日8点之后创建或修改的文件，在这个时间之前创建或修改的文件被忽略。</p>
增量迁移间隔	是（针对增量迁移）	默认值1小时，最大值24小时。
增量迁移次数	是（针对增量迁移）	默认值1次，最大值30次。
文件覆盖方式	是	<p>源地址中文件和目的地址中文件同名时，迁移过程中执行的覆盖方式。包括如下选项：</p> <ul style="list-style-type: none"> 最后修改时间优先：对于同名文件，判断两个文件的LastModified，即最后修改时间。 <ul style="list-style-type: none"> 如果源地址中文件的LastModified晚于目的地址中文件的LastModified，则执行覆盖。 如果源地址中文件的LastModified早于目的地址中文件的LastModified，则执行跳过。 如果两个文件的LastModified相同，则继续判断两个文件的Size和Content-Type是否均相同。如果两个文件的Size和Content-Type均相同，则执行跳过；如果两个文件的Size或者Content-Type中存在至少一个不同，则执行覆盖。 条件覆盖：对于同名文件，判断两个文件的LastModified、Size和Content-Type是否相同。 <ul style="list-style-type: none"> 如果两个文件的LastModified、Size和Content-Type均相同，则执行跳过。 如果两个文件的LastModified、Size和Content-Type中存在至少一个不同，则执行覆盖。 全覆盖：对于同名文件，不进行任何判断，直接执行覆盖。 不覆盖：对于同名文件，不进行任何判断，直接执行跳过。

5. 在性能调优页签的数据预估区域，填写待迁移存储量和待迁移文件个数。

 **说明** 为了迁移任务的顺利进行，请尽量准确进行数据预估。更多信息，请参见[预估迁移数据](#)。

6. （可选）在性能调优页签的流量控制区域，设置限流时间段和最大流量，然后单击添加。

❓ 说明 为了不影响您在线业务的访问，建议您根据业务访问的波峰和波谷来设置迁移时的限流时间段和最大流量。

7. 单击**创建**，等待迁移任务完成。

4. 后续操作

本文介绍迁移任务创建后的一些后续操作。


查看迁移任务状态

迁移任务创建后，有以下四种状态：

- 迁移中：数据正在迁移中，请您耐心等待。
- 创建失败：迁移任务创建失败，您可以查看失败原因，重新创建迁移任务。
- 已完成：迁移任务完成，您可以查看迁移报告。
- 失败：迁移任务失败，您可以生成并查看迁移报告，之后重新迁移失败的文件。

修改限流

数据迁移过程中，您可以根据您的实际情况随时修改限流参数。

1. 登录[数据在线迁移控制台](#)。
2. 选择[在线迁移服务](#) > [迁移任务](#)。在迁移任务列表中，单击对应任务的管理。
3. 单击停止，并确认停止。
4. 单击流量时间规划图的重置。
 - 如需增加限流，选择合适的限流时间段及最大流量，单击添加。
 - 如需删除限流，请单击对应限流条目后的按钮。
 - 如需修改限流，请先删除对应的限流条目后再添加。
5. 单击提交，然后单击启动，重新启动迁移任务。

查看迁移报告

1. 打开迁移任务列表，单击对应任务的管理。
2. 单击生成迁移报表。待报告生成后，单击导出，导出迁移报告。

迁移报告中，文件列表一栏包含三个文件名：

- 以 `_total_list` 结尾的文件名代表总迁移文件列表。
 - 以 `_completed_list` 结尾的文件名代表已迁移完成文件列表。
 - 以 `_error_list` 结尾的文件代表迁移失败文件列表。
3. 在[对象存储OSS控制台](#)，找到自动生成的文件夹 `aliyun_mgw_import_report/`，其中包含迁移报告中列出的三个文件。您可以下载这些文件，查看详细的文件列表，推荐使用[ossbrowser](#)工具查看。

文件格式如下：

- 总迁移文件列表：源地址+文件名+源版本ID+目的版本ID+源是否最新版本+源是否删除版本+文件大小（Byte）+最后修改时间。源地址的格式为：`<厂商>://<bucketName>/<prefix>/<objectName>`，例如：`oss://bucket-test1022/myprefix/testfile.txt`。
- 已迁移完成文件列表：文件名+源版本ID+目的版本ID+源是否最新版本+源是否删除版本+文件大小（Byte）+校验值（CRC64）+迁移完成时间+是否跳过迁移。

② 说明 是否跳过迁移：对于源、目的端都存在的同名文件，在线迁移服务是否跳过迁移。可能的值为：1（跳过）/0（未跳过）。此值受客户创建任务时文件覆盖方式策略的影响。

- 迁移失败文件列表：文件名+源版本ID+目的版本ID+源是否最新版本+源是否删除版本+迁移开始时间+迁移结束时间+错误描述。

② 说明 如下4个字段在启用多版本迁移任务时才具有意义：

- 源版本ID：此版本在源端的版本ID。
- 目的版本ID：此版本在目的端的版本ID。
- 源是否最新版本：此版本在源端是否是该文件的最新版本。可能的值为：true/false。
- 源是否删除版本：此版本在源端是否是一个删除版本。可能的值为：true/false。

迁移失败后重试

如果迁移任务失败，可查看迁移失败文件列表，找出失败原因并排除。之后，在迁移任务列表中，单击失败任务的管理，然后单击重试，重新迁移失败的文件。

5.地域与域名关系对照表

本文提供AWS S3的地域与域名关系对照表。

RegionName	Region	Endpoint	Protocol
US East (Ohio)	us-east-2	http://s3.us-east-2.amazonaws.com http://s3-us-east-2.amazonaws.com	HTTP/HTTPS
US East (N. Virginia)	us-east-1	http://s3.us-east-1.amazonaws.com http://s3-us-east-1.amazonaws.com	HTTP/HTTPS
US West (N. California)	us-west-1	http://s3.us-west-1.amazonaws.com http://s3-us-west-1.amazonaws.com	HTTP/HTTPS
US West (Oregon)	us-west-2	http://s3.us-west-2.amazonaws.com http://s3-us-west-2.amazonaws.com	HTTP/HTTPS
Canada (Central)	ca-central-1	http://s3.ca-central-1.amazonaws.com http://s3-ca-central-1.amazonaws.com	HTTP/HTTPS
Asia Pacific (Seoul)	ap-northeast-2	http://s3.ap-northeast-2.amazonaws.com http://s3-ap-northeast-2.amazonaws.com	HTTP/HTTPS
Asia Pacific (Osaka-Local)	ap-northeast-3	http://s3.ap-northeast-3.amazonaws.com http://s3-ap-northeast-3.amazonaws.com	HTTP/HTTPS

RegionName	Region	Endpoint	Protocol
Asia Pacific (Singapore)	ap-southeast-1	http://s3.ap-southeast-1.amazonaws.com http://s3-ap-southeast-1.amazonaws.com	HTTP/HTTPS
Asia Pacific (Sydney)	ap-southeast-2	http://s3.ap-southeast-2.amazonaws.com http://s3-ap-southeast-2.amazonaws.com	HTTP/HTTPS
Asia Pacific (Tokyo)	ap-northeast-1	http://s3.ap-northeast-1.amazonaws.com http://s3-ap-northeast-1.amazonaws.com	HTTP/HTTPS
China (Beijing)	cn-north-1	http://s3.cn-north-1.amazonaws.com.cn	HTTP/HTTPS
China (Ningxia)	cn-northwest-1	http://s3.cn-northwest-1.amazonaws.com.cn	HTTP/HTTPS
EU (Frankfurt)	eu-central-1	http://s3.eu-central-1.amazonaws.com http://s3-eu-central-1.amazonaws.com	HTTP/HTTPS
EU (Ireland)	eu-west-1	http://s3.eu-west-1.amazonaws.com http://s3-eu-west-1.amazonaws.com	HTTP/HTTPS
EU (London)	eu-west-2	http://s3.eu-west-2.amazonaws.com http://s3-eu-west-2.amazonaws.com	HTTP/HTTPS

RegionName	Region	Endpoint	Protocol
EU (Paris)	eu-west-3	http://s3.eu-west-3.amazonaws.com http://s3-eu-west-3.amazonaws.com	HTTP/HTTPS
South America (São Paulo)	sa-east-1	http://s3.sa-east-1.amazonaws.com http://s3-sa-east-1.amazonaws.com	HTTP/HTTPS