

MaxCompute

准备工作

准备工作

准备阿里云账号

使用 MaxCompute 服务前，您需要准备一个阿里云账号。

操作步骤

注册阿里云账号

如果您还没有注册过阿里云账号，请进入 [阿里云官网](#)，单击 **免费注册**，即可进入阿里云账号注册页面创建新的阿里云账号。

注意：

主账号创建成功后，作为阿里云系统识别的资源消费账号，会拥有非常高的权限，请您尽可能保证账号和密码的安全，切勿借给他人使用或定期更新密码。

阿里云账号实名认证

阿里云账号需要进行实名制认证后，才能购买和使用阿里云上的各种产品。如果您还没有实名认证，请进入 [实名认证](#) 页面对账号进行实名认证。**为保证后续操作顺利进行，请务必完成实名认证操作。**

对于企业级用户，最好可以进行企业级认证，以获取更多的便利。更多详情请参见 [会员账号&实名制认证](#)。

创建运行密钥 accesskeys

为了保证您在大数据开发套件中的任务顺利运行，需要为您创建一个运行密钥。该密钥区别于您登录时填写的账号和密码，主要用于在阿里云各产品间互相认证使用权限。运行密钥 AK 包括 Access Key ID 和 Access Key Secret 两部分。具体操作如下：

登录阿里云官网，在右上角的用户名下单击 **accesskeys** 进入 [Access Key 管理](#) 页面。



单击右上角的 **创建 Access Key**，单击弹出框中的 **同意并创建**，即可成功创建。如下图所示：



成功创建 Access Key 后，便自动跳转至 Access Key 管理页面，您可查看相应 AK 的状态并对其进行禁用、删除等操作。如下图所示：

Access Key管理 (2)					刷新	创建Access Key
① Access Key ID和Access Key Secret是您访问阿里云API的密钥，具有该用户完全的权利，请您妥善保管。						
Access Key ID	Access Key Secret	状态	创建时间	操作		
LTAI5C8356N7K8W	显示	启用	2017-07-14 10:32:26	禁用 删除		
LTAI5C8356N7K8W	显示	启用	2017-02-16 17:14:49	禁用 删除		

Accesskeys 一旦被禁用，使用该 Accesskeys 的服务将运行失败并报错，因此如有变更，需要及时关注使用该 Accesskeys 的产品和服务。

注意：

账号的 accesskeys 非常重要，创建成功后，请您尽可能保证 Access Key ID 和 Access Key Secret 的安全，切勿让他人知晓，一旦有泄漏的风险，请及时禁用和更新。

后续步骤

现在，您已经学习了如何准备阿里云账号，您可以继续学习下一个教程。在该教程中您将学习如何开通 MaxCompute 服务。详情请参见 [开通 MaxCompute](#)。

开通MaxCompute

如果您是第一次使用数加产品，请使用阿里云账号登录，并根据 [准备阿里云账号](#) 检查账号是否已可用。如果验证无误，请根据下述操作开通 MaxCompute 服务。

操作步骤

进入 [阿里云 MaxCompute 产品首页](#)，单击 **立即开通**，进入购买页面。如下图所示：

The screenshot shows the MaxCompute purchase interface. At the top, it says '大数据计算服务MaxCompute (包月)'. Below this are tabs for '预付费' (Prepaid) and '按量付费' (Pay-as-you-go). A note indicates that if the user is unsure about the purchase method or pricing, they should click '点击咨询帮你确定'. The main configuration area includes a '区域' (Region) dropdown with options like '华东2 (上海)', '华南1 (深圳)', '香港 (香港)', '亚太东南1 (新加坡)', '亚太东南2 (悉尼)', and '欧洲中部1 (法兰克福)'. The '计算资源单元' (Compute Resource Unit) is set to '10 CU'. A note explains that 1 CU equals 1 core CPU + 4 GB, and for purchases over 500 CU, users should contact sales. The '下载服务' (Download Service) and '存储资源' (Storage Resource) sections show purchase quantities and prices. On the right, the '当前配置' (Current Configuration) summary shows the region as '华南1 (深圳)', 10 CU, 1-month term, and a total price of '¥ 1500.00'. An '立即购买' (Buy Now) button is present. At the bottom, the '购买时长' (Purchase Duration) is set to '1个月' (1 month), with other options like 2, 3, 4, 5, 6, 7, 8, 9 months, 1 year, 2 years, and 3 years available.

根据自身需求进行选择、购买，MaxCompute 有按 CU 预付费和按 I/O 后付费两种付费方式，详细介绍请参见 [计量计费说明](#) 中的计算计费模块。

关于区域的选择，您需要考虑的最主要因素是 MaxCompute 与其他阿里云产品之间的关系，如：
：ECS 在什么区域，数据在什么区域。

单击 **立即购买** 确认订单后，单击 **去开通**，如下图所示：



跳转至服务开通成功的页面，如下图所示：



后续步骤

现在，您已经学习了如何开通 MaxCompute 服务，您可以继续学习下一个教程。在该教程中您将学习如何创建项目空间。详情请参见 [创建项目](#)。

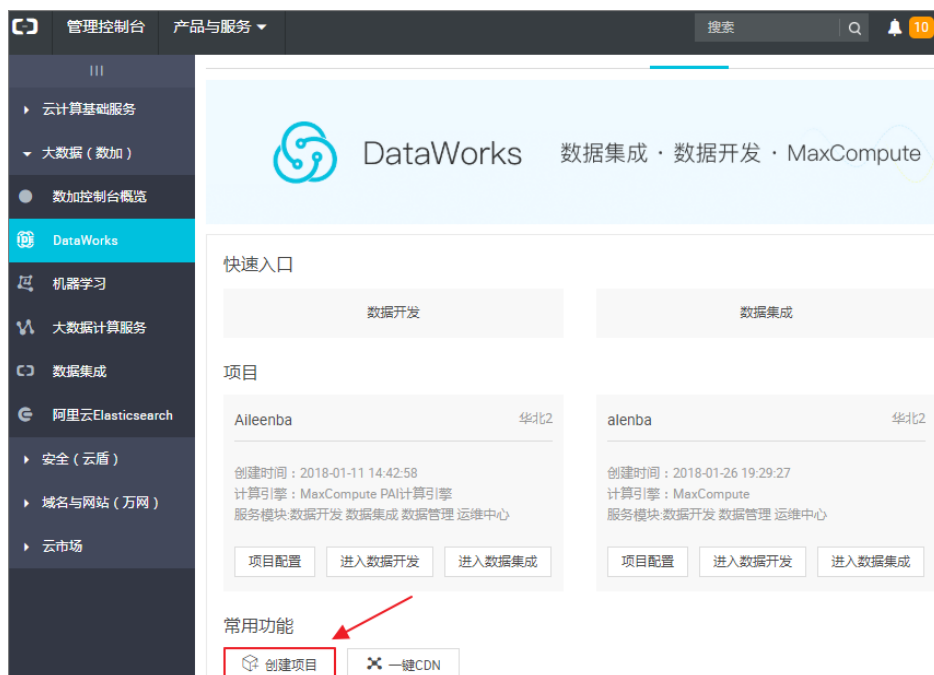
创建项目

您成功开通MaxCompute服务后，即可根据下文的操作创建项目空间。

操作步骤

主账号登录DataWorks管理控制台创建MaxCompute项目，您可以通过以下两种方式进行创建。

单击控制台概览页面**常用功能**下的**创建项目**，如下图所示：



导航至控制台项目列表页面，单击**创建项目**，如下图所示：



填写**创建项目**弹出框中的各配置项，选择Region及服务，如Region没有购买相关的服务，会直接显示**该Region下暂无可用服务**，数据开发、运维中心、数据管理默认是选中的。如下图所示：

创建项目

选择region

华东2 华北1 华北2 华东1 华南1 香港 美西1 亚太东南 1 华北3
欧洲中部 1 亚太东南 2 亚太东南 3 亚太东北 1

选择计算引擎服务

☐  MaxCompute ☐ 按量付费 ☐ 包年包月 [去购买](#)
开通后，您可在DataWorks里进行ODPS SQL，ODPS MR任务的开发。

☐  机器学习PAI ☐ 按量付费 [去购买](#)
开通后，您可使用机器学习算法、深度学习框架及在线预测服务。使用机器学习PAI，需要使用MaxCompute

选择DataWorks服务

☐  数据集成
开通后，您可在DataWorks里进行数据集成任务的开发，快捷实现二十多种数据源之间的数据同步。

☒  数据开发、运维中心、数据管理
您可以进行工作流编排、周期调度任务、查询所有表的信息和权限，相关服务目前处于公测阶段。

取消

下一步

选择计算引擎服务：

MaxCompute：大数据计算服务（MaxCompute，原名ODPS）是一种快速、完全托管的TB/PB级数据仓库解决方案，能够更快速为您解决海量数据计算问题，有效降低企业成本，并保障数据安全。详情请参见MaxCompute文档。

机器学习PAI：机器学习指的是机器通过统计学算法，对大量的历史数据进行学习从而生成经验模型，利用经验模型指导业务。详情请参见机器学习PAI。

选择DataWorks服务：

数据集成：数据集成是阿里集团对外提供的稳定高效、弹性伸缩的数据同步平台。致力于提供复杂网络环境下、丰富的异构数据源之间数据高速稳定的数据移动及同步能力。详情请参见数据集成概述。

数据开发：该页面是您根据业务需求，设计数据计算流程，并实现为多个相互依赖的任务，供调度系统自动执行的主要操作页面。详情请参见数据开发概述。

运维中心：该页面可对任务和实例进行展示和操作，您可以在这里查看您所有任务的实例。详情请参见运维中心概述。

数据管理：数据管理模块中可以看到组织内全局数据视图、分权管理、元数据信息详情、数据生命周期管理、数据表/资源/函数权限管理审批等操作。详情请参见数据管理概述。

配置新建项目的基本信息和高级设置。

创建项目

基本信息

* 项目名称：

haha01

显示名：

haha01|

项目描述：

高级设置

* 启动调度周期：

开

?

* 本项目中能下载select结果：

开

?

面向MaxCompute

MaxCompute项目名称：

haha01

?

* MaxCompute访问身份：

☒ 个人账号

☐ 系统账号

?

* Quota组切换：

默认后付费资源组

▼

上一步

创建项目

基本配置：

项目名：项目名长度需要在3到27个字符之间。

显示名：显示名不能超过27个字符。

高级配置：

启用调度周期：控制当前项目是否启用调度系统，如果关闭则无法周期性调度任务。

允许在本项目中直接编辑任务和代码：当前项目成员在本项目中新建/编辑代码文件的权限，如果关闭则无法新建/编辑代码文件。

本项目中能下载select结果：控制数据开发中select出的数据结果是否能够下载，如果关闭无法下载select的数据查询结果。

MaxCompute项目名称：创建项目过程中可以创建MaxCompute相同的名称项目。

MaxCompute访问身份：个人账号，系统账号。推荐安全要求高的企业用户使用个人账号。

Quota组切换：Quota用来实现磁盘配额。

项目创建成功后，项目列表即会显示相应内容，如下图所示：

项目列表									
华北2 华东1 华东2 华南1 香港 美国1 亚太东南1 美东1 欧洲中部1 亚太东南2 亚太东南3 亚太东北1 亚太南部1 亚太东南5									
搜索									
项目/显示名	创建时间	管理员	状态	开通服务	操作				
Aileen Aileen	2018-01-26 19:29:27	longg@aliyun.com	正常		项目配置 进入数据开发 修改服务 更多				

注意：

您一旦成为项目空间Owner，意味着该项目内**所有东西**都是您的，在给别人赋权之前，任何人无权限访问您的空间。

对于一般用户，并非一定要创建项目空间，只要被加入到某个项目便可使用MaxCompute。

后续步骤

现在，您已经学习了如何创建项目，您可以继续学习下一个教程。在该教程中您将学习如何安装配置客户端并通过客户端访问MaxCompute的各项功能。详情请参见安装配置客户端。

安装并配置客户端

通过客户端可以访问MaxCompute的各项功能，本文将为您介绍如何安装并配置客户端，更多详情请参见客户端介绍。

注意：

推荐您使用MaxCompute Studio工具，该工具集成java环境，可进行MaxCompute SQ脚本开发/执行、数据管理、可视化分析日志、JAVA（UDF、MR）开发，还可以将console集成到这个工具中进行使用。详情请参见Studio文档。

您也可以通过DataWorks使用MaxCompute，即前面步骤控制台中创建项目后，直接单击项目列表下对应项目的**进入工作区**进行操作，详情请参见DataWorks文档。

安装客户端

注意：

因为客户端是用Java开发的，所以在安装前，请确保您的机器上有JRE 1.7或以上版本（建议优先使用JRE 1.7/1.8，其中JRE 1.9已经支持，JRE 1.10暂时还不支持）。

单击[此处](#)，下载MaxCompute客户端。

解压下载的文件，解压后可以看到如下4个文件夹：

```
bin/  conf/  lib/  plugins/
```

配置客户端

编辑conf文件夹中的odps_config.ini文件，对客户端进行配置，如下所示：

```
access_id=*****
access_key=*****
# Accesss ID及Access Key是用户的云账号信息，可登录阿里云官网，进入管理控制台accesskeys页面进行查看。
project_name=my_project
# 指定用户想进入的项目空间。
end_point=https://service.odps.aliyun.com/api
# MaxCompute服务的访问链接
tunnel_endpoint=https://dt.odps.aliyun.com
# MaxCompute Tunnel服务的访问链接
log_view_host=http://logview.odps.aliyun.com
# 当用户执行一个作业后，客户端会返回该作业的LogView地址。打开该地址将会看到作业执行的详细信息。
https_check=true
#决定是否开启HTTPS访问
```

注意：

建议您根据自己的Region配置客户端，否则会出现无法访问等错误。

odps_config.ini文件中使用#作为注释，MaxCompute客户端内使用两个减号，即--作为注释。

您要提前创建好项目空间方可在配置文件中进行指定，详情请参见[创建项目空间](#)。

MaxCompute提供了公网和私网两个服务地址供您选择，不同的服务地址对您的下载计费结算有不同影响。详情请参见[访问域名和数据中心](#)。

运行客户端

修改好配置文件后，便可运行bin目录下的MaxCompute（Linux系统下运行./bin/odpscmd，Windows下运行./bin/odpscmd.bat），示例如下：

```
create table tbl1(id bigint);
insert overwrite table tbl1 select count(*) from tbl1;
select 'welcome to MaxCompute!' from tbl1;
```

更多SQL语句的介绍请参见[SQL概述](#)。

用户及角色管理

项目空间（Project）是MaxCompute实现多租户体系的基础，是用户管理数据和计算的基本单位，也是计量和计费的主体。当用户申请创建一个项目空间之后，该用户就是这个空间的所有者（Owner），即这个项目空间内的所有对象（表，实例，资源，UDF等）都属于该用户。这就是说，除了Owner之外，任何人都无权访问此项目空间内的对象，除非有Owner的授权许可。

如果您通过DataWorks进行添加用户及授权等操作，请参见[添加项目成员和角色](#)。

如果您通过MaxCompute安全管理命令进行用户管理，请参见[用户管理](#)，以了解如何添加/删除用户、给用户授权（包括RAM子账号管理）。

如果您通过MaxCompute安全管理命令进行角色管理，请参见[角色管理](#)，以了解如何创建/删除角色，如何给角色授权。

授权及权限查看的更多详情，请参见[授权和查看权限](#)。

访问域名和数据中心

本文将为您介绍MaxCompute Region的开通情况和连接方式，解答您在与其他云产品（ECS、TableStore、OSS）互访场景中遇到的网络连通性和下载数据收费等问题。



由上图可知，从服务层面来看，MaxCompute为您提供了两类服务连接地址，如下所示：

MaxCompute服务本身的连接地址：您可以向MaxCompute发出除数据上传、下载外的所有请求，例如：创建表、删除某个函数、创建一个作业等。

MaxCompute Tunnel服务的连接地址：上传、下载数据的能力是通过MaxCompute Tunnel服务提供的。当您想通过Tunnel上传、下载数据时，可以通过Tunnel提供的链接地址发起请求。

注意：

由于各Region部署和网络连接状况不一致，在Tunnel数据的下载计费规则上也不统一。

访问来源及下载数据收费规则说明

从阿里云各Region部署及网络情况来看，通常情况下分为三种连接方式，如下所示：

从外网访问MaxCompute服务和Tunnel服务。

从阿里云经典网络访问MaxCompute服务和Tunnel服务。

从阿里云VPC网络访问MaxCompute服务和Tunnel服务。

注意：

创建MaxCompute的项目时无需指定网络，只在需要连接项目时才需要指定通过什么网络进行连接。

数据上传

Tunnel数据上传无论走哪一种网络形态都免费，如上示意图所示。

数据下载

您无论是在ECS云服务器的哪一个Region上进行Tunnel服务请求进行下载数据，网络连通性设置都需满足如下形态定义：

两者在同一Region内，Tunnel下载请求走阿里云经典网络/VPC网络都**免费**。

两者不在同一Region内或没有条件满足同Region访问，则需走外网跨Region访问请求，此条件下的数据下载将会进行**计费**。

注意：

由于阿里云数据中心各个Region部署和网络情况不一致，若您选择**通过阿里云经典网络/VPC网络**进行跨Region的访问，则MaxCompute产品方**不承诺、不保证**其永久连通性。

MaxCompute访问外部表的连通性

MaxCompute2.0支持读写OSS对象存储数据，同时也支持读写TableStore表格存储数据，详情请参见访问OSS非结构化数据和访问OTS非结构化数据。

网络连通性的配置说明，如下所示：

MaxCompute和TableStore/OSS在同一Region情况下，建议配置阿里云经典网络或VPC网络连接方式，其外网也可以进行连通。

MaxCompute和TableStore/OSS不在同一Region情况下，配置外网访问方式进行连通。在跨Region的情况下，您可选择配置阿里云经典网络或VPC网络则不保证其连通性。

MaxCompute开通Region和服务连接对照表

从Region部署情况来看，MaxCompute目前国内国外陆续开服，您可以申请使用对应区域的MaxCompute，您的数据存储和计算消耗均发生在开通使用的区域。

注意：所有Endpoint都支持http和https，若需要请求上加密，请用https。

外网网络下Region和服务连接对照表

Region中文名称	所在城市	开服状态	外网Endpoint	外网Tunnel Endpoint
华东1	杭州	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-hangzhou.maxcompute.aliyun.com
华东2	上海	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun.com
华北2	北京	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-beijing.maxcompute.aliyun.com
华南1	深圳	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shenzhen.maxcompute.aliyun.com
香港	香港	已开服	http://service.cn-hongkong.maxcompute.aliyun.com/api	http://dt.cn-hongkong.maxcompute.aliyun.com
亚太东南1	新加坡	已开服	http://service.ap-southeast-1.maxcompute.aliyun.com/api	http://dt.ap-southeast-1.maxcompute.aliyun.com
亚太东南2	澳洲	已开服	http://service.ap-southeast-2.maxcompute.aliyun.com/api	http://dt.ap-southeast-2.maxcompute.aliyun.com
亚太东南3	吉隆坡	已开服	http://service.ap-southeast-3.maxcompute.aliyun.com/api	http://dt.ap-southeast-3.maxcompute.aliyun.com
亚太东南5	雅加达	已开服	http://service.ap-southeast-5.maxcompute.	http://dt.ap-southeast-5.maxcompute.

			aliyun.com/api	aliyun.com
亚太东北1	日本	已开服	http://service.ap-northeast-1.maxcompute.aliyun.com/api	http://dt.ap-northeast-1.maxcompute.aliyun.com
欧洲中部1	法兰克福	已开服	http://service.eu-central-1.maxcompute.aliyun.com/api	http://dt.eu-central-1.maxcompute.aliyun.com
美国西部1	硅谷	已开服	http://service.us-west-1.maxcompute.aliyun.com/api	http://dt.us-west-1.maxcompute.aliyun.com
美国东部1	弗吉尼亚	已开服	http://service.us-east-1.maxcompute.aliyun.com/api	http://dt.us-east-1.maxcompute.aliyun.com
亚太南部1	孟买	已开服	http://service.ap-south-1.maxcompute.aliyun.com/api	http://dt.ap-south-1.maxcompute.aliyun.com

经典网络下Region和服务连接对照表

Region中文名称	所在城市	开服状态	经典网络Endpoint	经典网络Tunnel Endpoint
华东1	杭州	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-hangzhou.maxcompute.aliyun-inc.com
华东2	上海	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-shanghai.maxcompute.aliyun-inc.com
华北2	北京	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-beijing.maxcompute.aliyun-inc.com
华南1	深圳	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-shenzhen.maxcompute.aliyun-inc.com
香港	香港	已开服	http://service.cn-hongkong.maxcompute.aliyun-inc.com/api	http://dt.cn-hongkong.maxcompute.aliyun-inc.com
亚太东南1	新加坡	已开服	http://service.a	http://dt.ap-

			p-southeast-1.maxcompute.aliyun-inc.com/api	southeast-1.maxcompute.aliyun-inc.com
亚太东南2	澳洲	已开服	http://service.ap-southeast-2.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-2.maxcompute.aliyun-inc.com
亚太东南3	吉隆坡	已开服	http://service.ap-southeast-3.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-3.maxcompute.aliyun-inc.com
亚太东南5	雅加达	已开服	http://service.ap-southeast-5.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-5.maxcompute.aliyun-inc.com
亚太东北1	日本	已开服	http://service.ap-northeast-1.maxcompute.aliyun-inc.com/api	http://dt.ap-northeast-1.maxcompute.aliyun-inc.com
欧洲中部1	法兰克福	已开服	http://service.eu-central-1.maxcompute.aliyun-inc.com/api	http://dt.eu-central-1.maxcompute.aliyun-inc.com
美国西部1	硅谷	已开服	http://service.us-west-1.maxcompute.aliyun-inc.com/api	http://dt.us-west-1.maxcompute.aliyun-inc.com
美国东部1	弗吉尼亚	已开服	http://service.us-east-1.maxcompute.aliyun-inc.com/api	http://dt.us-east-1.maxcompute.aliyun-inc.com
亚太南部1	孟买	已开服	http://service.ap-south-1.maxcompute.aliyun-inc.com/api	http://dt.ap-south-1.maxcompute.aliyun-inc.com

VPC网络下Region和服务连接对照表

在VPC网络下访问MaxCompute，只能使用如下Endpoint和Tunnel Endpoint。

Region中文名称	所在城市	开服状态	VPC网络	VPC网络Tunnel
------------	------	------	-------	-------------

			Endpoint	Endpoint
华东1	杭州	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-hangzhou.maxcompute.aliyun-inc.com
华东2	上海	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-shanghai.maxcompute.aliyun-inc.com
华北2	北京	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-beijing.maxcompute.aliyun-inc.com
华南1	深圳	已开服	http://service.cn.maxcompute.aliyun-inc.com/api	http://dt.cn-shenzhen.maxcompute.aliyun-inc.com
香港	香港	已开服	http://service.cn-hongkong.maxcompute.aliyun-inc.com/api	http://dt.cn-hongkong.maxcompute.aliyun-inc.com
亚太东南1	新加坡	已开服	http://service.ap-southeast-1.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-1.maxcompute.aliyun-inc.com
亚太东南2	澳洲	已开服	http://service.ap-southeast-2.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-2.maxcompute.aliyun-inc.com
亚太东南3	吉隆坡	已开服	http://service.ap-southeast-3.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-3.maxcompute.aliyun-inc.com
亚太东南5	雅加达	已开服	http://service.ap-southeast-5.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-5.maxcompute.aliyun-inc.com
亚太东北1	日本	已开服	http://service.ap-northeast-1.maxcompute.aliyun-inc.com/api	http://dt.ap-northeast-1.maxcompute.aliyun-inc.com
欧洲中部1	法兰克福	已开服	http://service.eu-central-1.maxcompute.	http://dt.eu-central-1.maxcompute.

			aliyun-inc.com/api	aliyun-inc.com
美国西部1	硅谷	已开服	http://service.us-west-1.maxcompute.aliyun-inc.com/api	http://dt.us-west-1.maxcompute.aliyun-inc.com
美国东部1	弗吉尼亚	已开服	http://service.us-east-1.maxcompute.aliyun-inc.com/api	http://dt.us-east-1.maxcompute.aliyun-inc.com
亚太南部1	孟买	已开服	http://service.ap-south-1.maxcompute.aliyun-inc.com/api	http://dt.ap-south-1.maxcompute.aliyun-inc.com

访问原则

对于已开服的Region，您可以通过公网、经典网络、VPC网络方式连接MaxCompute服务。

通过配置外网Tunnel Endpoint地址下载数据进行收费，价格为0.8元/GB。