

阿里云 MaxCompute

准备工作

文档版本：20180806

法律声明

阿里云提醒您在使用或阅读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
	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。	 禁止： 重置操作将丢失用户配置数据。
	该类警示信息可能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，恢复业务所需时间约10分钟。
	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。	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。	单击 确定 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid 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>swich {stand slave}</code>

目录

法律声明.....	I
通用约定.....	I
1 准备阿里云账号.....	1
2 开通MaxCompute.....	4
3 创建项目.....	6
5 用户及角色管理.....	10
6 访问域名和数据中心.....	11

1 准备阿里云账号

背景信息

使用MaxCompute服务前，您需要准备一个阿里云账号。

操作步骤

1. 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。



说明：

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。

2. 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&实名认证](#)。

3. 创建运行密钥accesskeys

为了保证您在DataWorks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：

a) 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。



b) 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：



c) 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 ID	Access Key Secret	状态	创建时间	操作
LTAI5O9MFK2H9P	显示	启用	2017-07-14 10:32:26	禁用 删除
LTAI5O9MFK2H9P	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。

 说明：

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。

2 开通MaxCompute

背景信息

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

1. 进入[阿里云MaxCompute产品首页](#)，单击立即购买，进入购买页面。如下图所示：



您可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，详情请参见[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。

2. 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：



3 创建项目

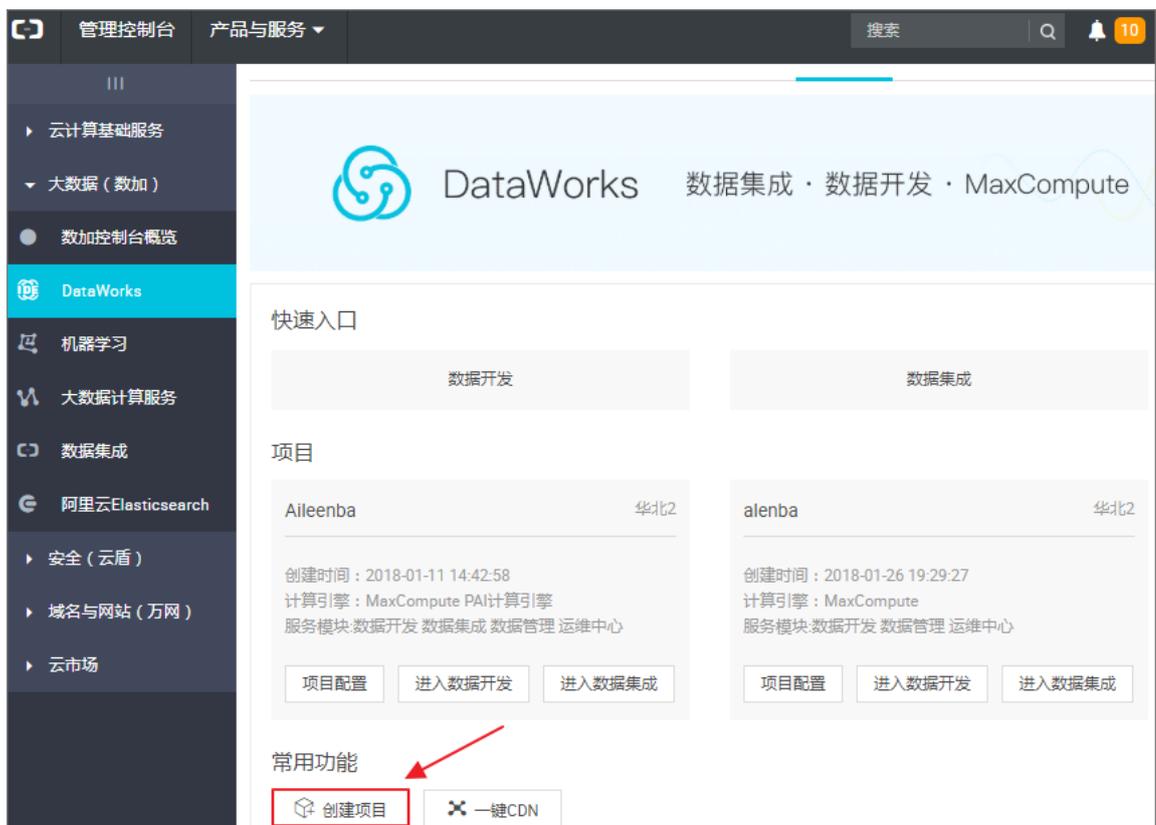
背景信息

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

1. 登录>DataWorks管理控制台，您可以通过以下两种方式创建MaxCompute项目。

- 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：



2. 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MaxCompute SQL，MaxCompute 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 按量付费 [去购买](#)

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 workflow 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。

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](#)。

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](#)。
3. 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：

显示名：

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 ?

MaxCompute项目名称： ?

* MaxCompute访问身份： 个人账号 系统账号 ?

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。

4. 项目创建成功后，您可进入项目列表页面查看相应内容。

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	longgall@163	正常		项目配置 进入数据开发 修改服务 更多



说明：

- 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。

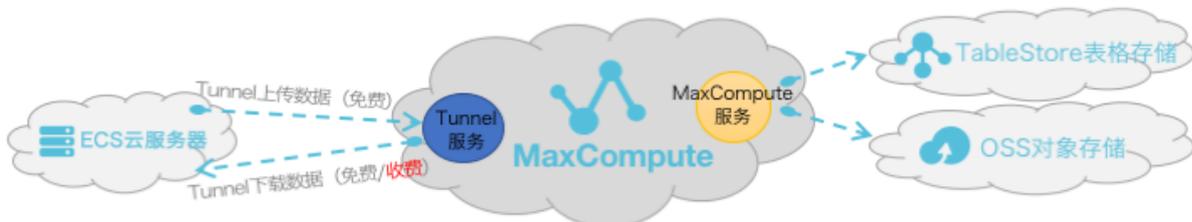
5 用户及角色管理

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](#)。

6 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为您提供了两大类服务连接地址，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。



说明：

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。



说明：

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。



说明：

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。



说明：

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

Region中文名称	所在城市	开服状态	外网Endpoint	外网Tunnel Endpoint
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.aliyun.com/api	http://dt.ap-southeast-5.maxcompute.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

Region中文名称	所在城市	开服状态	外网Endpoint	外网Tunnel Endpoint
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

Region中文名称	所在城市	开服状态	经典网络 Endpoint	经典网络Tunnel Endpoint
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. aliyun-inc.com/ api	http://dt.eu -central-1. maxcompute. aliyun-inc.com

Region中文名称	所在城市	开服状态	经典网络Endpoint	经典网络Tunnel Endpoint
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Endpoint	VPC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

Region中文名称	所在城市	开服状态	VPC网络Endpoint	VPC网络Tunnel Endpoint
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.aliyun-inc.com/api	http://dt.eu-central-1.maxcompute.aliyun-inc.com

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