

MaxCompute

准备工作

准备工作

使用 MaxCompute 服务前，您需要准备一个阿里云账号，具体操作如下：

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](#)，单击右上角的 **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，切勿借给他人使用并定期更新密码。

实名认证阿里云账号

阿里云账号需要进行实名制认证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（为使后续步骤顺利进行，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

为了保证您在数加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。

您在登录阿里云官网后，在右上角用户名下单击 accesskeys 即可进入 [个人管理控制台](#)，在此页面进行创建新的 accesskeys，禁用甚至删除已有 accesskeys 等操作。

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](#)。

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，如下图所示：



单击 **立即开通** 后，出现下图：



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，您首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。

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

登录 大数据开发套件管理控制台，您可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：



填写 **创建项目** 弹出框中的各配置项，单击 **确定**，如下图所示：

创建项目

Maxcompute老用户，填写MaxCompute项目名称，系统为您自动绑定同名的数据开发项目

* 付费方式:

CU预付费

去购买

I/O后付费

* 项目名称:

MaxCompute_demo

显示名:

MaxCompute_demo

项目描述:

MaxCompute项目示例

确定

取消

完成上述操作，即可看到项目列表，如下图所示：



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](#)。

通过客户端可以访问 MaxCompute 的各项功能，本文将介绍如何安装并配置客户端，更多详情请参见：[客户端介绍](#)。

也可以通过大数据开发套件使用MaxCompute,即前面步骤控制台中创建项目后,直接点击项目的“进入工作区” 进行操作,具体使用请参考[大数据开发套件文档](#)。

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。

点击 [此链接](#)，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=*****
# Access 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：

- 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，详细操作请参见：[创建项目空间](#)；
- MaxCompute 提供了公网和私网两个服务地址供用户选择，不同的服务地址对用户的计费结算

有不同影响。详细说明请参见：[Tunnel SDK 简介](#)。

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 概要](#)。

任意非项目空间 Owner 用户必须被加入 MaxCompute 项目空间中，并被授予相对应权限，方能操作 MaxCompute 中的数据、作业、资源及函数。本文将介绍项目空间 Owner 如何将其他用户加入 MaxCompute。

如果您是项目空间 Owner，建议您仔细阅读本文；如果您是普通用户，建议您向 Owner 提出申请，被加入对应的项目空间后再阅读后续章节。

本文的操作均在客户端运行，Linux 系统下运行 ./bin/odpscmd，Windows 下运行 ./bin/odpscmd.bat。

添加用户

添加用户的命令为：

```
ADD USER <full_username> ;
```

示例如下：

假设需要添用户的云账号为 bob@aliyun.com 到某个项目空间，此时该项目空间的 Owner 需要通过客户端执行如下命令：

```
add user ALIYUN$bob@aliyun.com;
```

若是添加RAM子帐号：

```
add user RAM$主帐号:子帐号;
```

如果不确定该用户是否已经存在于此项目空间，请输入如下命令进行查看：

```
LIST USERS;
```

添加用户成功后，此用户仍不能操作 MaxCompute，需要对用户授予一定权限，用户才能在所拥有的权限范围内操作 MaxCompute。关于授权的更多操作，请参见：用户授权及权限查看。

删除用户

删除用户的命令为：

```
REMOVE USER <full_username> ;
```

示例如下：

假设需要在某个项目空间删除用户 bob@aliyun.com，此时该项目空间的 Owner 需要通过客户端执行如下命令：

```
remove user ALIYUN$bob@aliyun.com;
```

若删除RAM子帐号：

```
remove user RAM$主帐号:子帐号;
```

在删除用户前，请确保该用户的所有权限已经被取消。

添加、删除用户的更多详情请参见：项目空间的用户管理。

添加用户后，项目空间 Owner 或者项目空间管理员需要给该用户进行授权，只有用户获得权限后，才能执行操作。所谓授权，即授予用户对 MaxCompute 中的表，任务，资源等客体的某种操作权限，包括：读、写、查看等，详情请参见：项目成员管理，以了解 MaxCompute 角色对应的大数据开发套件中的成员角色及其权限特征。

如果您只是 MaxCompute 的普通用户，请确认已经获取足够权限，快速浏览本章节即可。

MaxCompute 提供了 ACL 授权，跨项目空间数据分享及 项目空间数据保护等多种策略。本文只简单介绍使用最多的 ACL 授权。

ACL 授权

ACL 授权中，MaxCompute 的客体包括：Project，Table，Function，Resource，Instance，Task。每种客体具有不同的操作权限，详细介绍请参见：ACL 授权。

授权的命令为：

```
GRANT privileges ON project_object TO project_subject
```

```
REVOKE privileges ON project_object FROM project_subject

privileges ::= action_item1, action_item2, ...

project_object ::= PROJECT project_name | TABLE schema_name |
INSTANCE inst_name | FUNCTION func_name |
RESOURCE res_name | JOB job_name

project_subject ::= USER full_username | ROLE role_name
```

备注：上述授权命令中有关角色（ROLE）的相关部分用户可以暂时不予关注，待后续章节介绍。user full_username，若为阿里云主帐号，写法为 ALIYUN\$xxxx，若为RAM子帐号，写法为 RAM\$主帐号:子帐号。

示例如下：

假设用户云账号为 bob@aliyun.com 已经被添加到某个项目空间（\$user_project_name），需要给它授予建表、获取表信息和执行的权限。项目空间管理员（admin 角色）或者该项目空间 owner 在客户端可以执行如下命令：

```
grant CreateTable on PROJECT $user_project_name to USER ALIYUN$bob@aliyun.com;
-- 向 bob@aliyun.com 授予名为 "$user_project_name" 的 project 的 CreateTable（创建表）权限

grant Describe on Table $user_table_name to USER ALIYUN$bob@aliyun.com;
-- 向 bob@aliyun.com 授予名为 "$user_table_name" 的 Table 的 Describe（获取表信息）权限

grant Execute on Function $user_function_name to USER ALIYUN$bob@aliyun.com;
-- 向 bob@aliyun.com 授予名为 "$user_function_name" 的 Function 的 Execute（执行）权限
```

查看权限

查看权限的命令为：

```
show grants for $user_name;
```

示例如下：

假设需要查看用户云账号 bob@aliyun.com 在当前项目空间的权限，需要在客户端可以执行如下命令：

```
show grants for ALIYUN$bob@aliyun.com
```

查看RAM子帐号权限：

```
show grants for RAM$主帐号:子帐号;
```

备注：关于查看权限的详细介绍请参见：[查看权限](#)。

目前，MaxCompute 支持阿里云账号和 RAM 账号两种账号体系，本文主要介绍如何添加 RAM 子账号。添加 RAM 子账号有以下两种方式：

通过大数据开发套件进行操作，详情请参见：如何添加成员及授权；

通过 MaxCompute 客户端常用命令进行操作，详情如下：

注意：MaxCompute 只能够识别 RAM 的账号体系，不能识别 RAM 的权限体系。即用户可以将自身的任意 RAM 子账号加入 MaxCompute 的某一个项目中，但 MaxCompute 在对该 RAM 子账号做权限验证时，并不会考虑 RAM 中的权限定义。

默认情况下，MaxCompute 项目只能够识别阿里云账号系统，用户可以通过list accountproviders;命令查看该项目所支持的账号系统，通常情况下仅会看到 ALIYUN 账号，例如：

```
odps@ ****>list accountproviders;
ALIYUN
```

注意：只有项目空间的owner有权限进行accountproviders的相关操作。

由上可见，只能看到ALIYUN账号体系，如果想添加对 RAM 账号的支持，可以执行add accountprovider ram;如下所示：

```
odps@ odps_pd_inter>add accountprovider ram;
OK
```

通过list accountproviders;来查看账号支持情况，如下所示：

```
odps@ ****>list accountproviders;
ALIYUN, RAM
```

由上可见，这个项目空间已经能够支持RAM账号体系，即可以向这个项目空间添加 RAM 子账号并授予某张表的Describe权限，如下所示：

```
odps@ ****>add user ram$bob@aliyun.com:Alice;
OK: DisplayName=RAM$bob@aliyun.com:Alice
odps@ ****>grant Describe on table src to user ram$bob@aliyun.com:Alice;
OK
```

此时，bob@aliyun.com的 RAM 子账号Alice就可以通过自己的AccessKeyID及AccessKeySecret登录MaxCompute 并对表src进行desc操作。

备注：

- 获取 RAM 子账号AccessKeyID及AccessKeySecret的相关操作请参见：[RAM 介绍](#)；
- 更多 MaxCompute 添加/删除用户的操作请参见：[安全介绍之用户管理](#)；
- 更多有关授权的操作请参见：[安全介绍之授权操作](#)。

同样，您可以通过remove user命令删除自身的 RAM 子账号。示例如下：

```
odps@ ****>revoke describe on table src from user ram$bob@aliyun.com:Alice;
OK
-- 回收子账号 Alice 权限

odps@ ****>remove user ram$bob@aliyun.com:Alice;
Confirm to "remove user ram$bob@aliyun.com:Alice;" (yes/no)? yes
OK
-- 删除子账号
```

如果您是项目空间的 owner，也可以通过remove accountprovider将 RAM 账号系统从当前项目中删除，如下所示：

```
odps@ ****>remove accountprovider ram;
Confirm to "remove accountprovider ram;" (yes/no)? yes
OK

odps@ ****>list accountproviders;
ALIYUN
```

正如 [用户授权及权限查看](#) 中所说，所有被加入 MaxCompute 的用户只有被授予相应的权限后才能操作 MaxCompute。当项目空间下存在大量用户时，这个授权动作会显得非常繁琐。在这种情况下，项目空间管理员可以对拥有相同权限的用户进行分类，每类用户有相同的权限，我们称之为 **角色**。多个用户可以同时存在于一个角色下，一个用户也可以隶属于多个角色。给角色授权后，该角色下的所有用户拥有相同的权限。

创建角色

创建角色的命令格式为：

```
CREATE ROLE <roleName>;
```

示例如下：

假设要创建一个 player 角色，需在客户端输入如下命令：

```
create role player;
```

添加用户到角色

添加用户到角色的命令格式为：

```
GRANT <roleName> TO <full_username>;
```

示例如下：

假设要将用户 bob@aliyun.com 加入 player 角色中，需在客户端输入如下命令：

```
grant player to bob@aliyun.com;
```

给角色授权的语句与给用户授权相似，您可参见：[用户授权示例](#) [用户授权及权限查看](#) 及[授权的详细介绍](#) 对用户或角色进行授权。

删除角色中的用户

删除角色中的用户的命令格式为：

```
REVOKE <roleName> FROM <full_username>;
```

示例如下：

假设将用户 bob@aliyun.com 从 player 角色中删除，需在客户端输入如下命令：

```
revoke player from bob@aliyun.com;
```

删除角色

删除角色的命令格式为：

```
DROP ROLE <roleName>;
```

示例如下：

假设要删除 player 角色，需在客户端输入如下命令：

```
drop role player;
```

注意：删除角色前，请确认 player 下的所有用户都已经被移出该角色。

更多有关角色及其授权的介绍请参见：[项目空间的角色管理](#) 及 [对用户或角色进行授权](#)。

在本章节中，着重介绍MaxCompute region开通情况和连接方式，解答大家在与其它云产品（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 向用户提供了两大类服务连接地址：

(1) MaxCompute服务本身的连接地址：用户可以向 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、创建一个作业等等。

(2) 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。当用户想通过 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：

- ① 从外网访问MaxCompute服务和Tunnel服务；
- ② 从阿里云经典网络访问MaxCompute服务和Tunnel服务；
- ③ 从阿里云VPC网络访问MaxCompute服务和Tunnel服务。

数据上传

需要明确的是Tunnel数据上传无论走哪一种网络形态都为免费，如上示意图所示。

数据下载

用户无论是在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：

- ① 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**；
- ② 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**。

此处需要注意的是，由于阿里云数据中心各个Region部署和网络情况不一致，若用户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

MaxCompute 2.0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，具体可以详见：访问OSS数据 和 访问TableStore数据。网络连通性配置说明，如下：

- ① MaxCompute和TableStore/OSS在同一Region情况下，建议配置阿里云经典网络或VPC网络链接方式，其外网也可以进行连通。
- ② MaxCompute和TableStore/OSS不在同一Region情况下，配置外网访问方式进行连通。若跨Region情况下，用户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 目前仅在华东2(上海)区域开服，用户可以申请使用华东2（上海）区域的MaxCompute。用户的数据存储和计算消耗均发生在华东2（上海）区域。

Region 中文名称	所在城市	开服状态	外网 Endpoint	外网 Tunnel Endpoint	经典网络 Endpoint	经典网络 Tunnel Endpoint
华东2	上海	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun.com	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun-inc.com
华北2	北京	未开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-beijing.maxcompute.aliyun.com	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-beijing.maxcompute.aliyun-inc.com
华北3	张家口	未开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun.com	http://service.cn.maxcompute.aliyun.com/api	

华南1	深圳	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun.com	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shenzhen.maxcompute.aliyun-inc.com	
华东1	杭州	未开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun.com	http://service.cn.maxcompute.aliyun.com/api		
香港	香港	已开服	http://service.cn-hongkong.maxcompute.aliyun.com/api	http://dt.cn-hongkong.maxcompute.aliyun.com	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.com/api	http://dt.ap-southeast-1.maxcompute.aliyun.com	http://service.ap-southeast-1.maxcompute.aliyun-inc.com/api	http://dt.ap-southeast-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2	上海	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shanghai.maxcompute.aliyun-inc.com	
华北2	北京	未开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-beijing.maxcompute.aliyun-inc.com	

华北3	张家口	未开服	http://service.cn.maxcompute.aliyun.com/api		
华南1	深圳	已开服	http://service.cn.maxcompute.aliyun.com/api	http://dt.cn-shenzhen.maxcompute.aliyun-inc.com	
华东1	杭州	未开服	http://service.cn.maxcompute.aliyun.com/api		
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	

访问原则

对于已开服的Region，或者有接入点的，用户可以通过公网、经典网络、VPC网络方式连接MaxCompute服务；

对于未开服的Region，并且没有接入点的，用户只能通过外网连接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。