项目：Maxcompute的管理单元；

表：Maxcompute的存储单元，MaxCompute中不同类型计算任务的操作对象（输入、输出）都是表。MaxCompute的表格有两种类型：内部表和外部表。①对于内部表，所有的数据都被存储在MaxCompute中，表中列的数据类型可以是MaxCompute支持的任意一种数据类型版本说明；②外部表：对于外部表，MaxCompute并不真正持有数据，表格的数据可以存放在OSS或OTS中 。MaxCompute仅会记录表格的Meta信息，您可以通过MaxCompute的外部表机制处理OSS或OTS上的非结构化数据，例如，视频、音频、基因、气象、地理信息等。

分区：分区可以理解为分类，通过分类把不同类型的数据放到不同的目录下。分类的标准就是分区字段，可以是一个，也可以是多个。

分区表：分区表是指拥有分区空间的表，即在创建表时指定表内的一个或者某几个字段作为分区列。分区表实际就是对应分布式文件系统上的独立的文件夹，一个分区对应一个文件夹，文件夹下是对应分区所有的数据文件。分区表的意义在于优化查询。查询表时通过WHERE子句查询指定所需查询的分区，避免全表扫描，提高处理效率，降低计算费用。使用数据时，如果指定需要访问的分区名称，则只会读取相应的分区。

生命周期：MaxCompute表的生命周期（Lifecycle），指表（分区）数据从最后一次更新的时间算起，在经过指定的时间后没有变动，则此表（分区）将被MaxCompute自动回收。这个指定的时间就是生命周期。

函数：MaxCompute为您提供了SQL计算功能，您可以在MaxCompute SQL中使用系统的[内建函数](https://help.aliyun.com/document_detail/27864.html)完成一定的计算和计数功能。但当内建函数无法满足要求时，您可以使用MaxCompute提供的Java编程接口开发自定义函数（User Defined Function，以下简称UDF）

自定义函数：自定义函数（UDF）可以进一步分为标量值函数（UDF），自定义聚合函数（UDAF）和自定义表值函数（UDTF）三种类型。

A

AccessKey

AccessKey（简称AK，包括AccessKey ID和AccessKey Secret），是访问阿里云API的密钥。在阿里云官网注册云账号后，可在AccessKey管理页面生成，用于标识用户，为访问MaxCompute或者其他云产品做签名验证。AccessKey Secret必须保密。

安全

MaxCompute多租户数据安全体系，主要包括用户认证、项目的用户与授权管理、跨项目的资源分享以及项目的数据保护。关于MaxCompute安全操作的更多详情请参见安全指南。

C

Console

MaxCompute Console是运行在Window/Linux下的客户端工具，通过Console可以提交命令完成项目管理、DDL、DML等操作。对应的工具安装和常用参数请参见客户端。

D

Data Type

MaxCompute表中所有列对应的数据类型。目前支持的数据类型详情请参见数据类型版本说明。

DDL

数据定义语言（Data Definition Language）。例如创建表、创建视图等操作，详情请参见DDL语句。

DML

数据操作语言（Data Manipulation Language）。例如INSERT操作，MaxCompute DML语法请参见INSERT语句。

F

fuxi

伏羲（fuxi）是飞天平台内核中负责资源管理和任务调度的模块，同时也为应用开发提供了一套编程基础框架。MaxCompute底层任务调度模块为fuxi的调度模块。

I

Instance（实例）

作业的一个具体实例，表示实际运行的Job，类同Hadoop中Job的概念。详情请参见任务实例。

M

MapReduce

MapReduce是处理数据的一种编程模型，通常用于大规模数据集的并行运算。您可以使用MapReduce提供的接口（Java API）编写MapReduce程序，来处理MaxCompute中的数据。编程思想是将数据的处理方式分为Map（映射）和Reduce（规约）。

在正式执行Map前，需要将输入的数据进行分片。所谓分片，就是将输入数据切分为大小相等的数据块，每一块作为单个Map Worker的输入被处理，以便于多个Map Worker同时工作。每个Map Worker在读入各自的数据后，进行计算处理，最终通过Reduce函数整合中间结果，从而得到最终计算结果。详情请参见MapReduce。

O

ODPS

ODPS是MaxCompute的原名。

P

Partition（分区）

分区Partition是指一张表下，根据分区字段（一个或多个字段的组合）对数据存储进行划分。也就是说，如果表没有分区，数据是直接放在表所在的目录下。如果表有分区，每个分区对应表下的一个目录，数据是分别存储在不同的分区目录下。关于分区的更多介绍请参见分区。

Project（项目）

项目（Project）是MaxCompute的基本组织单元，它类似于传统数据库的Database或Schema的概念，是进行多用户隔离和访问控制的主要边界。详情请参见项目。

R

Role（角色）

角色是MaxCompute安全功能里使用的概念，可以看成是拥有相同权限的用户的集合。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。关于角色管理的更多介绍请参见角色管理。

Resource（资源）

资源（Resource）是MaxCompute中特有的概念。如果您想使用MaxCompute的自定义函数（UDF）或MapReduce功能，则需要依赖资源来完成。详情请参见资源。

S

SDK

Software Development Kit软件开发工具包。一般都是一些被软件工程师用于为特定的软件包、软件实例、软件框架、硬件平台、操作系统、文档包等建立应用软件的开发工具的集合。MaxCompute目前支持Java SDK介绍和Python SDK。

授权

项目管理员或者项目 Owner授予您对MaxCompute中的Object（或称之为对象，例如表、任务、资源等）进行某种操作的权限，包括读、写、查看等。授权的具体操作请参见用户管理。

沙箱（Sandboxie）

沙箱是一种按照安全策略限制程序行为的执行环境。沙箱机制是一种安全机制，将Java代码限定在特定的运行范围中，并且严格限制代码对本地系统资源访问，通过这样的措施来保证对代码的有效隔离，防止对本地系统造成破坏。MaxCompute MapReduce及UDF程序在分布式环境中运行时受到Java沙箱的限制。

T

Table（表）

表是MaxCompute的数据存储单元，详情请参见表。

Tunnel

MaxCompute的数据通道，提供高并发的离线数据上传下载服务。您可以使用Tunnel服务向MaxCompute批量上传数据或者向本地进行批量数据下载。相关命令请参见Tunnel命令参考或批量数据通道SDK。

U

UDF

广义的UDF（User Defined Function），代表了自定义标量函数、自定义聚合函数及自定义表函数三种类型的自定义函数的集合。MaxCompute提供的Java编程接口开发自定义函数，详情请参见概述。

狭义的UDF指用户自定义标量值函数（User Defined Scalar Function），它的输入与输出是一对一的关系，即读入一行数据，写出一条输出值。

UDAF

自定义聚合函数（User Defined Aggregation Function），它的输入与输出是多对一的关系， 即将多条输入记录聚合成一条输出值。可以与SQL中的GROUP BY语句联用。详情请参见UDAF。

UDTF

自定义表值函数（User Defined Table Valued Function），用来解决一次函数调用输出多行数据的场景。它是唯一能返回多个字段的自定义函数，而UDF只能一次计算输出一条返回值。详情请参见UDTF。