阿里云
大数据计算服务

新功能发布记录

文档版本：20190906

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
	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。
`courier` 字体	命令。	执行 `cd /d C:/windows` 命令，进入Windows系统文件夹。
*## *	表示参数、变量。	`bae log list --instanceid` *`Instance_ID`*
[]或者[a|b]	表示可选项，至多选择一个。	`ipconfig` *`[-all|-t]`*
{}或者{a|b}	表示必选项，至多选择一个。	`swich` *`{stand | slave}`*

目录

法律声明.....I
通用约定.....I
1 公告..... 1
2 文档最新动态.....14
3 发展历程.....38
4 新功能测试申请.....40

1 公告

本文为您提供关于MaxCompute使用功能的各项更新记录。

2019年9月01日-MaxCompute存储降价

MaxCompute于2019年9月1日开始，对现有的#unique_4进行降价，降价原则：

- 将原来的5个阶梯价调成3个阶梯价，减少存储计费复杂度。
- 将新阶梯单价调低，最终达到降价目的。

2019年9月1日前，原存储阶梯价如下。

存储量	阶梯单价	固定价
大于0小于等于0.5GB	不涉及	0.01元/天
大于0.5GB小于等于100GB	0.0192元/GB/天	不涉及
大于100GB小于等于1TB	0.0096元/GB/天	不涉及
大于1TB小于等于10TB	0.0084元/GB/天	不涉及
大于10TB小于等于100TB	0.0072元/GB/天	不涉及
大于100TB	0.006元/GB/天	不涉及

2019年9月1日起，新存储阶梯价如下。

存储量	阶梯单价	固定价
大于0小于等于0.5GB	不涉及	0.01元/天
大于0.5GB小于等于10TB	0.0072元/GB/天	不涉及
大于10TB小于等于100TB	0.006元/GB/天	不涉及
大于100TB	0.004元/GB/天	不涉及

存储计费统计方式不变，更多信息请参见#unique_4：

- 存储到MaxCompute的数据，包括表（Table）和资源（Resource）等，会按照其数据容量的大小进行阶梯计费，计费周期为天。
- 以小时级别采集您每个项目空间下当前的存储使用情况，并以项目空间为基本单位，计算您当天的存储平均值再乘以单价。由于MaxCompute以项目空间为基本单位计算您当天的存储平均值。因此数据越集中存放在某个项目空间中，存储费用会越低。

假设您的某个项目的每天存储平均值为1PB，原阶梯单价每天收取的费用如下。

```
100GB*0.0192元/GB/天
+ (1024-100) GB*0.0096元/GB/天
```

```
+ (10240-1024) GB*0.0084元/GB/天
+ (102400-10240) GB*0.0072元/GB/天
+ (10240*10240-102400) GB*0.006元/GB/天
=6428.81元/天
```

新阶梯单价每天收取费用如下。

```
10240 GB*0.0072元/GB/天
+ (102400-10240) GB*0.006元/GB/天
+ (10240*10240-102400) GB*0.004元/GB/天
=4411.39元/天
```

原单价与新单价在单个项目空间内对比如下。

数据量(TB)	原价费用(元/天)	新价费用(元/天)	降价比例
10	88.20	73.73	16.41%
100	751.76	626.69	16.64%
1024	6428.81	4411.39	31.38%

从表中可以看出，单个项目空间内存储量越高，在新阶梯单价下，降价率就越高。

2019年8月29日-外表自定义storagehandler实现Outputer接口升级

北京时间2019年8月29日，MaxCompute进行版本升级。期间，您在使用外表自定义storagehandler实现Outputer接口时，如果通过列名而非数字下标获取列数据，可能会引起作业失败。

升级时间：北京时间，2019年8月29日14:00-23:00

升级Region：美国西部1（硅谷）、亚太东南1（新加坡）

2019年8月21日-外表自定义storagehandler实现Outputer接口升级

北京时间2019年8月21日，MaxCompute进行版本升级。期间，您在使用外表自定义storagehandler实现Outputer接口时，如果通过列名而非数字下标获取列数据，可能会引起作业失败。

升级时间：北京时间，2019年8月21日14:00-23:00

升级Region：亚太东北1（东京）、欧洲中部1（法兰克福）、中国（香港）、亚太东南2（悉尼）

影响原因：`Outputer.output(Record record)`中，传入的record为Outputer的上一个Operator产生的记录，列名可能发生变化，系统无法保证固定列名。

例如，表达式`some_function(column_a)`产生的列名是一个临时列名。因此，使用`record.get(列名)`方式来获取列内容的用法都有可能受到影响，建议使用`record.get(index)`方式获取。Outputer里如需获取表的列名，请调用 `DataAttributes.getFullTableColumns()`。

如您遇到相关问题，请提交工单咨询。

2019年6月25日-SQL语言功能升级

- UDT支持Resource

  UDT支持资源（Resource）的访问。您可以在SQL中通过`com.aliyun.odps.udf.impl.UDTExecutionContext.get()`静态方法获取`ExecutionContext`对象，从而访问当前的`ExecutionContext`，进而访问资源（例如文件资源和表格资源）。详情请参见#unique_5。

- UDT实现函数重载

  MaxCompute的UDF使用重载`evaluate`方法的方式来重载函数。这种方式不支持泛型，所以当您需要定义一个可以接受任何数据类型的函数时，必须为每种类型都写一个`evaluate`函数。但是，这种方法依然无法实现个别输入类型（例如ARRAY）的重载函数。在没有提供Resolve注解（Annotation）的情况下，Python UDF或UDTF会根据参数个数决定输入参数，同时支持变长参数，但这种过于灵活的机制也会导致编译器无法静态找到某些错误。通过UDT实现函数重载，可以很好地解决以上问题。UDT支持泛型、类继承、变长参数，为您提供灵活的函数定义方式。详情请参见#unique_5/unique_5_Connect_42_section_uap_og3_qgs。

- UDAF/UDTF支持动态参数

  MaxCompute UDAF和UDTF中的Annotation里，支持在输入参数部分使用`*`、`ANY`，分别表示接受任意长度且任意类型、任意类型的参数。UDTF的返回值使用星号时，表示返回任意个STRING类型，而`ANY`不能用于返回值列表中。详情请参见#unique_6。

- 参数化视图

  MaxCompute传统的视图（VIEW）中，底层可以封装一段逻辑复杂的SQL脚本，调用者可以像读普通表一样调用视图，无需关心底层的实现。传统的视图实现了一定的封装与重用，因此被广泛地使用。但是传统的视图并不接受调用者传递的任何参数，导致代码重用能力低下。MaxCompute当前的SQL引擎支持带参数的视图，支持传入任意表或者其它变量，定制视图的行为。详情请参见#unique_7。

- UDF支持分区裁剪

  UDF已支持分区裁剪，支持的方式是将UDF语句先当做一个小作业执行，再将执行的结果替换到原来UDF出现的位置 。实现的方式有以下两种：

  - 在编写UDF的时候，UDF类上加入Annotation。

    ```
    @com.aliyun.odps.udf.annotation.UdfProperty(isDeterministic=true)
    ```

    说明：

请把引用的*odps-sdk-udf*版本提高到0.30.x或以上。

- 在SQL语句前设置Flag：`set odps.sql.udf.ppr.deterministic = true;`，此时SQL中所有的UDF均被视为`deterministic`。该操作执行的原理是做执行结果回填，但是结果回填存在限制，即最多回填1000个Partition。因此，如果UDF类加入`annotation`，则可能会导致出现超过1000个回填结果的报错。此时您如果需要忽视此错误，可以通过设置Flag：`set odps.sql.udf.ppr.to.subquery = false;`全局关闭此功能。详情请参见#unique_8/unique_8_Connect_42_section_lwx_cv2_ggb。

· 支持`show create table`命令

类似Hive中的`show create table`的语法，可以直接查看建表语句便于移植。详情请参见#unique_9/unique_9_Connect_42_section_a8z_v92_gi0。

2019年5月8日-预付费套餐发布

MaxCompute预付费套餐开始售卖，支持的套餐规格为：

· 存储密集型160套餐（含计算资源160CU+存储资源150TB）
· 存储密集型320套餐（含计算资源320CU+存储资源300TB）
· 存储密集型600套餐（含计算资源600CU+存储资源500TB）

购买套餐后，存储超出套餐资源会按量收费，价格比标准版更优惠。具体的套餐说明请参见#unique_10。

2019年3月26日-SQL语言功能升级

· 支持Grouping Set多维聚合分析（Cube，Roll up），适配需要对a列做聚合，也要对b列做聚合，或者同时要按照a、b两列做聚合的场景。具体的使用方法请参见#unique_11。
· 支持INTERSECT/MINUS/EXCEPT，具体使用方法请参见#unique_12。
· 通过外表读取OSS上的ORC和Parquet格式的文件时，支持对文件的列裁剪，有效减少IO量，节省计算资源和成本。
· Java UDX类型系统增强：UDF支持Writable参数，具体使用方法请参见#unique_13。

SQL性能优化

· DynamicDAG：动态优化的必要机制，它可以将一个优化（可能是资源配置或算法选择）延迟到运行时，以便提高优化的精准度，降低产生较差计划的可能性。
· ShuffleRemove Optimization：针对Shuflle的优化。 从本次版本升级开始，MaxCompute将支持对inner join时右表key相同的场景进行Shuffle Remove优化。

2019年3月19日-MaxCompute 支持开发者版

从2019年3月19日开始，MaxCompute 按量后付费购买模式支持开发者版。您过去购买的MaxCompute按量付费模式为标准版。

目前开发者版只支持 北京、上海、杭州、深圳4个Region。一个Region内您只能有一个项目可使用开发者版资源，且仅新创建的项目可选择开发者版资源。

开发者版的SQL计费标准为：

```
一次SQL计算费用=计算输入数据量*SQL价格
```

SQL价格为0.15元/GB。开发者版项目可以转换为标准版，标准版项目不能转换成开发者版。

2019年2月1日-MaxCompute Spark及MaxCompute Lightning开始收费

从2019年2月1日开始，MaxCompute Lightning（交互式分析）服务及MaxCompute Spark服务开始计费。

MaxCompute Lightning查询任务的计费公式为：

`一次Lightning查询费用 = 查询输入量*单价`

MaxCompute Spark采用的计费标准如下：

`Spark任务当日计算费用 = 当日总计算时*0.66元（人民币）`

详情请参见计量计费说明。

2019年1月23日-开始分批升级上海Region的时区定义

MaxCompute将于2019年1月23日开始分批将上海Region的时区定义从tzdata-2011g-1升级到tzdata-2015g-1.1。

升级影响：在升级过程中，如果您对1927年以前的中国时区时间进行操作，执行结果会有漂移，示例如下。

```
select to_date('0001-01-01', 'yyyy-mm-dd');
```

执行结果为`0001-01-01 00:00:09`。

升级完成后，不会有此类影响。

操作建议：如果您在1月23日之后有针对1927年以前的中国时区时间的操作，请注意检查执行结果是否受到影响，并进行相应处理。

2019年1月15日16:00-20:00-中国（香港）Region底层优化

为向您提供更好的产品性能和稳定性，MaxCompute将于北京时间2019年1月15日16:00 - 20:00对中国（香港）Region底层元数据仓库组件进行优化。在优化期间，中国（香港）Region用户的应用可能出现1分钟左右任务无法提交、运行中任务失败的情况。极端情况下，应用不可用时间将延长至30分钟。请您尽量避免在迁移窗口提交作业。其他Region不受影响。如果您有任何问题，可随时通过企业钉钉群或工单反馈。

2018年11月9日-Hash Clustering功能发布

Hash Clustering通过允许在建表时设置表的Shuffle和Sort属性，使得MaxCompute可以根据数据已有的存储特性来优化执行计划，提高效率，节省资源消耗。如果在初始表数据生成时，按照Hash Shuffle和Sort的方式存储，则后续查询中将避免对数据的再次Shuffle和Sort。具体如何在建表时设置、查看、管理Shuffle和Sort属性请参考DDL语句。

说明:

目前Hash Clustering还存在限制：

- 不支持insert into，只能通过insert overwrite来添加数据。
- 不支持Tunnel直接Upload到Range Cluster表，因为Tunnel上传数据是无序的。

2018年11月3日-底层服务软件缺陷修复

2018年11月3日10:00到19:00将对MaxCompute底层服务进行软件缺陷修复。正常情况下，您不会感知到升级过程。极端情况下，可能会产生约20分钟的作业运行失败或作业提交失败。一旦出现作业失败，请您重新提交失败任务或提交工单反馈。

2018年10月31日-SQL外部表功能开启计费

从2018年10月31日开始，MaxCompute SQL外部表功能开始计费。采用的计费标准为一次SQL计算费用=计算输入数据量*SQL复杂度*SQL价格。SQL价格是0.03元/（GB*复杂度），复杂度系数为1。当天的所有计量信息在第二天做一次性汇总收费，并直接体现在您的账户账单中。

2018年10月16日-Python UDF全面开放

从2018年10月16日开始，MaxCompute公共云环境下的Python UDF使用权限将全面开放，新增的Project默认可用Python UDF，具体的Python UDF使用方式请参考文档Python UDF。更多PyODPS的内容介绍请参考相关文档。

2018年10月15日-新增Project的setproject 2.0新类型默认为True

从2018年10月15日开始，MaxCompute创建的新Project，`odps.sql.type.system.odps2`属性默认为True，对setproject的详细说明请参见：#unique_19。

2018年9月29日-SQL语言升级

MaxCompute Sprint29中发布的新功能如下：

- UDJ

  为了满足用户很多跨表操作的需求，MaxCompute S29版本在UDF框架中引进了新的扩展机制：UDJ（User Defined Join）。UDJ提供了让用户自定义Join实现逻辑的接口。如何定义和使用UDJ，详情请参见#unique_20。

- UDT

  为了提升SQL语言编译过程的易用性与语言的表达能力，例如使用select transform可以不创建Function甚至不上传资源的情况下执行其他语言的脚本，MaxCompute2.0基于新一代的SQL引擎增加了UDT（User Defined Type）新功能。

  UDT功能允许用户在SQL中直接引用第三方语言的类或者对象，获取其数据内容或者调用其方法。例如在SQL中直接写Java的表达式，并可以引用JDK中的类。如何使用UDT，详情请参见UDT。

  > 说明：
  > - 目前UDT只支持Java语言。
  > - 所有的运算符都是MaxCompute SQL的语义，不是UDT的语义。
  > - 目前UDT不能用作Shuffle Key：包括join、group by、distribute by、sort by、order by、cluster by等结构的Key。
  > - DDL不支持UDT，也不能将UDT对象insert到表中。屏显的最终结果也不能是UDT类型。

- 支持Outer Lateral View

  MaxCompute Sprint30扩展了Lateral View功能，支持Outer Lateral View，当Table Function不输出任何一行时，对应的输入行在Lateral View结果中依然保留，且所有Table Function输出列为Null。关于Lateral View介绍请参考文档DML语句的#unique_22部分。

- 支持Scalar Subquery

  子查询支持Scalar Subquery，即当Subquery的输出结果为单行单列的时候，可以当做标量来使用。具体使用介绍请参见DML语句#unique_23中的Scalar Subquery部分。

2018年9月26日-元数据仓库组件优化

为了提升您的MaxCompute使用性能，系统即将对元数据仓库组件进行优化，届时您会提前收到以下信息：

尊敬的阿里云用户您好，大数据计算服务MaxCompute为向您提供更好的产品性能和稳定性，将于北京时间2018年9月26日16:00 - 20:00对底层元数据仓库组件进行优化，您的应用可能出现1分钟左右任务无法提交、运行中任务失败的情况。极端情况下，应用不可用时间将延长至30分钟。请您尽量避免在迁移窗口提交作业。感谢您的理解和支持！谢谢。有任何问题，可随时通过企业钉钉群或工单联系反馈。

2018年-MaxCompute例行优化

为了提升您的MaxCompute使用性能， 系统每个月将会对文件数占用较多的project进行小文件合并，届时您将会提前收到以下信息：

尊敬的客户，您好：为了提升您的MaxCompute使用性能， 我们计划于XX年XX月XX日对文件数占用较多的project（只会涉及到5天内未被访问分区）进行小文件合并，此操作是后台自动做，用户无需参与。Merge执行过程中，作业读取正在被Merge的表会有小概率失败，业务上重试即可成功，届时请关注相关业务，如有异常随时通过企业钉钉群进行反馈。

说明：

如果您收到报错：ODPS-0010000:System internal error - PanguOpen4Read4R，说明作业在读取正在被Merge的表，请在业务上重试。

2018年9月17日支持setproject 2.0新类型和兼容hive模式

为了提升MaxCompute 2.0的使用体验，支持对project级别进行新类型和兼容hive模式打开，即对原来session级别的`set odps.sql.type.system.odps2=true|false`和`set odps.sql.hive.compatible=true|false`两个功能，升级为支持setproject级别。project的Owner可根据需要对project进行设置，命令为：

```
setproject odps.sql.type.system.odps2=false|true
setproject odps.sql.hive.compatible=false|true
```

对setproject的详细说明请参见：#unique_19。

说明：

打开project级别的新类型的主要影响：

- 某些隐式类型转换会被禁用，包括string -> bigint, string - > datetime, double->bigint , decimal -> double, decimal -> bigint都是有精度损失或者可能报错的。用户依旧可以用cast来做强制转换。
- 常量类型会变化。单独一个 整形常量，如 123，在旧类型下是bigint类型，在新类型下是int类型。

- udf resolve结果可能变化，比如udf包含bigint和int两个重载。旧类型下一定走bigint的重载，而新类型下可能会被解析到int的重载。
- session级别会覆盖project级别设置。

2018年5月底-NewSQL开启计费

MaxCompute New SQL采用的计费标准为一次SQL计算费用=计算输入数据量*SQL复杂度*SQL价格。当天的所有计量信息在第二天做一次性汇总收费，并直接体现在您的账户账单中。

2018年1月16日流量切换公告

1月16日14~17点对MaxCompute前端、tunnel服务进行流量切换， 如果出现服务连接超时或访问失败，可提交工单进行咨询。

2018年1月15日开放label、Package安全模型

2018年1月15日，MaxCompute的安全模型开发label和Package，即列级别访问控制和跨项目空间的资源分享。

2018年1月10日禁止Full scan功能开放

2018年1月10日 20点新创建的项目（project）默认禁止分区表全表扫描，即对该项目的分区表执行全表扫描语句时会返回失败，必须指定分区条件，若实在需要，可以在对分区表全表扫描的sql语句前加一个set语句`set odps.sql.allow.fullscan=true;`，执行的时候，set语句和SQL语句一起提交执行。

2017年11月28-29日内部维护升级公告

由于MaxCompute内部维护升级出现问题，2017-11-28 22:00到2017-11-29 18:40期间的计算计量和存储计量未及时出账，相应账单会延后推送，但不影响最终计费结果，由此给您造成的不便深感抱歉。

2017年11月1日-华北2（北京）Region开服

2017年11月1日，阿里云数加>MaxCompute 华北2（北京）数据中心将正式开服售卖。

售价与华东2、华南1一致，详情请参见计量计费文档。

开通时，选择什么Region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。

不同区域数据是不能互通，仅针对MaxCompute而言，假如您既购买华东2又购买华北2的服务，那么MaxCompute之间的数据不互通。

如果您的云服务器ECS不在北京Region，若想访问连接北京的MaxCompute，需跨Region进行访问。跨Region访问的各种服务连接请参见访问域名和数据中心。

2017年10月11日-预付费支持自动续费

2017年10月11日，MaxCompute的预付费实例开始支持自动续费，可在阿里云控制台的续费管理中心操作。

MaxCompute续费的操作介绍请参见#unique_28，更多自动续费的介绍请参见自动续费功能。

2017年9月19日-中国（香港）Region开服

2017年9月19日，阿里云数加>MaxCompute中国（香港）数据中心将正式开服售卖。

售价与华东2、华南1一致，详情请参见计量计费文档。

开通时，选择什么Region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。

不同区域数据是不能互通，仅针对MaxCompute而言，假如您既购买华东2又购买中国（香港）的服务，那么MaxCompute之间的数据不互通。

如果您的云服务器ECS不在中国（香港），若想访问连接中国（香港）的MaxCompute，需跨Region进行访问。跨Region访问的各种服务连接请参见访问域名和数据中心。

2017年9月7日-华南1（深圳）Region开服

2017年9月7日，阿里云数加>MaxCompute华南1（深圳）数据中心将正式开服售卖，这是数加>MaxCompute在国内开服的第二个区域。

关于售价

华南1区域价格与华东2一致，主要分三部分进行收费：存储、计算和下载，其中计算（指SQL和MR计算任务）分预付费、按量后付费两种模式，存储和下载都是按量后付费。做预算的具体的售价信息请参见官网定价页或计量计费文档。

华南1区域价格与华东2一致，主要分三部分进行收费：存储、计算和下载，其中计算（指SQL和MR计算任务）分预付费、按量后付费两种模式，存储和下载都是按量后付费。做预算的具体的售价信息请参见计量计费文档。

关于开通

请您确保云账号是实名认证的账号，在开通购买页面进行区域选择时，注意选择华南1。那么什么场景适合选择区域是华南1呢?

选择地域时，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，示例如下：

- ECS在华南1区域。ECS访问、下载华南1的MaxCompute数据都可以通过内网，快速又避免跨Region走公网下载，从而产生费用。

- 数据在华南1区域。如RDS、OSS、TableStore等在华南1区，那么在和华南1的MaxCompute进行数据传输时可以走内网，避免跨Region配置阿里云内网或VPC网络时不保证连通性的情况。

关于不同区域数据是否互通

仅针对MaxCompute而言，假如您既购买华东2又购买华南1的服务，那么MaxCompute之间的数据不互通。

如果您的云服务器ECS不在华南1，若想访问连接华南1的MaxCompute，需跨Region进行访问。跨Region访问的各种服务连接请参见访问域名和数据中心。

关于数据跨Region迁移

若您原来已经开通购买华东2区域，目前想使用华南1区域，所以希望将华东2的所有数据迁移到华南1。建议您通过以下方式进行迁移：

通过配置数据集成任务进行数据同步，如在华南1区域创建好的项目中，配置数据集成任务，来读取华东2区域的表，写入到华南1区域的表中。如果遇到问题可以提工单进行咨询。

说明：
如果您对新服购买有疑惑，可通过工单咨询或加入MaxCompute购买钉钉群（群号：11782920）进行咨询。

2017年8月16日-MR开启计费

MaxCompute MapReduce任务开启按量计费。计费标准如下：

```
MR任务当日计算费用=当日总计算时 * 0.46元（人民币）
```

一个MR任务一次执行成功的计算时=任务运行时间（小时）*任务调用的core数量。

如一个MR任务一次执行成功是调用了100core并花费0.5小时，那么本次执行计算时为：0.5 小时 × 100core= 50个计算时。

详情请参见计量计费说明。

2017年7月-开始分批升级MaxCompute2.0

说明：
2017年9月底已经全面升级开放MaxCompute2.0，MaxCompute2.0 NewSQL及非结构化（外表）处理能力公测阶段，不收取费用，具体收费时间等待官方通知。

MaxCompute2.0已经正式开始，升级后的MaxCompute2.0完全拥抱开源生态，支持更多的语言功能，带来更快的运行速度，同时新版本会执行更严格的语法检测。升级的具体时间会在分批升级时另行通知。

升级后需要对部分SQL语法做出调整，详情请参见修改不兼容SQL实战进行修改。

本次MaxCompute2.0产品升级以下功能：

- 全新的SQL引擎

  更快的SQL执行引擎：降低企业大数据分析成本，SQL执行效率更高。
- 生态兼容

  MaxCompute显著提升了针对Hive等开源产品的兼容性，直接兼容Hive的UDF/UDAF/UDTF，为Hive开发的大部分UDF/UDAF/UDTF可不经修改，直接在MaxCompute中运行。

  详情请参见INSERT操作、SELECT操作、数据类型、日期函数、数学函数、字符串函数、其他函数和JAVA UDF等相关文档。
- 非结构化处理能力

  通过MaxCompute SQL直接处理OSS/TableStore（OTS）数据，从而处理音频、视频、图像、气象等非结构化数据以及K-V类型的数据。

  详情请参见前言、访问OTS非结构化数据和访问OSS非结构化数据。
- MaxCompute Studio

  MaxCompute编译器基于MaxCompute2.0全新自主研发的SQL引擎，尤其配合使用MaxCompute Studio，提供了丰富的错误提示、警告的功能及作业排队展示。详情请参见巧用MaxCompute编译器的错误和警告。

2017年06月01日

尊敬的阿里云客户：

MaxCompute目前只针对SQL作业进行收费，MapReduce作业近期将开启收费。

MaxCompute近期将进行2.0版本升级，升级过程中会逐批进行，本次版本变更将带来一些新特性：

- SQL执行引擎

  降低大数据分析成本，SQL执行效率更高。
- 非结构化处理能力

  通过MaxCompute SQL直接处理OSS/TableStore（OTS）数据，从而处理音频、视频、图像、气象等非结构化数据以及K-V类型数据。

- 生态兼容

  MaxCompute MapReduce SDK高度兼容了开源Hadoop MapReduce的使用场景。
- 基于IntelijJ的本地开发插件

  本地对MaxCompute作业进行开发，调试。
- 实时数据上传

  新增DataHub模块，帮助企业数据实时上传。
- 新型算法平台

  基于GPU的深度学习算法框架，在线预测服务等。

  如果您在系统升级过程中遇到问题，请及时提交工单联系我们。

2016年09月16日

尊敬的阿里云客户：

MaxCompute在11月1日进行了计费变更。本次变更升级存储及计算两个方面：

- 升级后，修复部分SQL作业无法计费的问题。
- 升级后，解决部分场景下MaxCompute存储费用存在的问题.升级后，存储将按照实际情况收费。

以上两处之前均不存在向您多收取费用的情况，对于历史作业不做费用补收。您可以通过计量计费说明中获取账单详情的相关指引，从阿里云官网获取每天详细的计量计费信息。

目前MaxCompute只有SQL作业收费（不包括UDF），UDF/MapReduce/Graph/PAI等作业仅是公测状态。具体收费计划请关注官网公告。

感谢您长期以来对MaxCompute产品的支持。

2 文档最新动态

本文为您介绍了MaxCompute 2.0版本内容更新的最新动态，基于此您可以了解MaxCompute中增加了哪些新功能、语法新特性和权限变更，为您开发项目提高效率。

2019年更新记录

时间	特性	类别	描述	产品文档
2019.09.02	新增PyODPS的聚合操作示例	新示例	新增PyODPS的聚合操作示例	#unique_43
2019.09.02	新增PyODPS的列运算示例	新示例	新增PyODPS的列运算示例	#unique_44
2019.08.13	新增PyODPS的Sequence及执行操作最佳实践	新示例	新增PyODPS的Sequence及执行操作最佳实践。	#unique_45
2019.08.07	新增MaxCompute数据迁移至OSS最佳实践	新示例	新增MaxCompute数据迁移至OSS最佳实践。	#unique_46
2019.08.01	新增PyODPS读取分区表数据最佳实践	新示例	新增PyODPS读取分区表数据最佳实践。	#unique_47
2019.08.01	新增PyODPS查看一级分区最佳实践	新示例	新增PyODPS查看一级分区最佳实践。	#unique_48
2019.07.29	新增MaxCompute数据迁移至OTS最佳实践	新示例	新增MaxCompute数据迁移至OTS最佳实践。	#unique_49
2019.07.25	新增PyODPS使用第三方包最佳实践	新示例	新增PyODPS使用第三方包最佳实践。	#unique_50
2019.07.22	新增PyODPS节点实现结巴中文分词最佳实践	新示例	新增PyODPS节点实现结巴中文分词最佳实践。	#unique_51
2019.07.22	新增PyODPS参数传递最佳实践	新示例	新增PyODPS参数传递最佳实践。	#unique_52

时间	特性	类别	描述	产品文档
2019.07.12	新增set命令odps.sql.reshuffle.dynamicpt说明	新命令	该命令用于设置动态分区，以避免拆分动态分区时产生过多小文件。	#unique_53
2019.07.11	新增tablestore.columns.mapping说明点	新说明	在指定映射时，提供的属性列必须是Table Store表的列，否则即使外表可以创建成功，查询时也会报错。	#unique_54/unique_54_Connect_42_sec
2019.07.09	新增子账号进行权限管理的案例	新案例	以案例分析的形式为您介绍如何使用子账号进行权限管理。	#unique_55
2019.07.08	新增数据质量保障教程	新教程	新增数据质量保障教程。	#unique_56
2019.07.05	新增UDF支持分区裁剪的说明引导	新说明	新增UDF支持分区裁剪的说明引导。	#unique_57/unique_57_Connect_42_sec
2019.07.05	新增成都地区的外网Endpoint和外网Tunnel Endpoint	新功能	新增成都地区的外网Endpoint和外网Tunnel Endpoint。	#unique_58
2019.07.04	新增开通MaxCompute服务说明	新说明	开通MaxCompute服务后，您可以进入管理控制台创建项目。MaxCompute管理控制台即为DataWorks管理控制台，因此您在创建项目前，请先开通DataWorks服务。	#unique_59

时间	特性	类别	描述	产品文档
2019.06.28	新增创建项目的前提条件说明	新说明	创建项目前，您需要先开通DataWorks服务，并确保开通的Region与MaxCompute开通的Region保持一致。	#unique_60
2019.06.25	新增时区配置功能介绍	新功能	新增时区配置功能介绍。	#unique_61
2019.06.25	新增MaxCompute跨项目迁移的最佳实践	新示例	介绍了同区域下不同的MaxCompute项目如何实现配置与数据的迁移。	#unique_62
2019.06.25	SQL语言功能升级	新功能	· UDT支持资源的访问 · UDT实现函数重载 · UDAF/UDTF支持动态参数 · 支持参数化视图 · UDF支持分区裁剪 · 支持直接查看建表语句	#unique_63/unique_63_Connect_42_sec unique_5_Connect_42_sect unique_8_Connect_42_sect unique_9_Connect_42_sect
2019.06.20	新增教程	新教程	新增搭建互联网在线运营分析平台的教程。	搭建互联网在线运营分析平台
2019.06.19	更新not in/in子查询返回数量超1000时的解决方案	新方案	更新not in/in子查询返回数量超1000时的解决方案。	如何使用not in实现上万级别数据量的子查询
2019.06.19	新增PyODPS最佳实践参考示例	新示例	新增PyODPS最佳实践参考示例。	#unique_64

时间	特性	类别	描述	产品文档
2019.06.18	新增Package授权引导说明	新说明	完成Package授权后，使用方在读Package中的表时，表名的写法为：表所属Project名称.表名。	#unique_65/unique_65_Connect_42_sec
2019.06.18	新增预付费资源组监控报警介绍	新功能	阿里云的云监控服务中的监控报警功能可用于监测MaxCompute预付费资源组的使用情况。	#unique_66
2019.06.17	新增VALUES文档引导说明	新说明	介绍了如何在业务测试阶段，创建数据进行简单的测试。	#unique_67
2019.06.05	新增复杂类型数据的上传下载示例	新示例	新增复杂类型数据的上传下载示例。	#unique_68
2019.06.04	新增复杂类型数据的说明	新说明	MaxCompute的复杂数据类型可以被任意嵌套使用。	#unique_69/unique_69_Connect_42_sec
2019.06.03	新增阿里云Elasticsearch数据迁移至MaxCompute最佳实践	新示例	新增阿里云Elasticsearch数据迁移至MaxCompute最佳实践。	#unique_70
2019.06.03	新增服务注销与资源释放介绍	新说明	开通MaxCompute服务后，不支持服务的注销。	#unique_71
2019.06.01	新增教程	新教程	新增构建与优化数据仓库的教程。	构建与优化数据仓库
2019.05.31	新增Java SDK示例	新示例	新增Java SDK示例。	#unique_72、#unique_73、
2019.05.29	新增Kafka数据迁移至MaxCompute最佳实践	新示例	新增Kafka数据迁移至MaxCompute最佳实践。	#unique_77

时间	特性	类别	描述	产品文档
2019.05.28	新增用户自定义获取计量信息说明	新说明	可通过计量API获取MaxCompute使用记录。	#unique_78
2019.05.24	新增MaxCompute的ACID语义说明	新说明	当遇到并发写入时，MaxCompute会根据ACID进行并发写的保障。	#unique_79、#unique_80/unique_80_Connect_42_sec unique_81_Connect_42_sec
2019.05.13	更新Spark开发指南	新案例	新增Java/Scala开发示例和PySpark开发示例。	#unique_82、#unique_83、
2019.05.10	新增欠费预警接收设置步骤说明	新说明	主账号可以设置子账号接收欠费预警消息。	#unique_85/unique_85_Connect_42_sec
2019.05.08	新增预付费套餐版说明	新规格	介绍预付费套餐版的基础资源、售价、注意事项、使用说明。	#unique_86、#unique_10
2019.05.06	新增数仓建设指南	新指南	介绍数仓建设规范。	#unique_87、#unique_88、
2019.05.06	新增表设计指南	新指南	介绍表设计指南。	#unique_94、#unique_95、
2019.05.06	新增数据管理指南	新指南	介绍数据管理指南。	#unique_98、#unique_99、
2019.05.06	新增安全管理案例之行级别权限控制	新案例	介绍如何实现行级别权限控制。	#unique_103
2019.04.29	新增日期函数from_utc_timestamp	新函数	新函数可以将一个UTC时区的时间戳转换成一个指定时区的时间戳。	#unique_104
2019.04.28	新增服务规格类型	新规格	当前售卖的规格类型包括预付费标准版、按量付费标准版、按量付费开发者版。	#unique_105、#unique_106

时间	特性	类别	描述	产品文档
2019.04.26	新增项目空间生命周期操作命令说明；新增IP白名单支持的IPV6格式	新功能	项目空间操作中增加生命周期操作命令说明。IP白名单已支持IPV6格式。	#unique_108
2019.04.18	新增脚本模式SQL	新功能	MaxCompute当前的SQL引擎已支持脚本模式。	#unique_109
2019.04.16	变更Tunnel Endpoint的自动路由规则	新规则	当不配置Tunnel Endpoint时，将自动路由到Service Endpoint所在网络对应的Tunnel Endpoint。	#unique_110
2019.04.15	新增开启数据保护机制后的数据流出方法	新方法	可以通过设置Exception Policy开启数据保护机制后的数据流出方法。	#unique_111
2019.04.11	新增InstanceTunnel介绍及示例	新功能	可以通过InstanceTunnel下载以select关键字开头的用于获取数据的SQL Instance。	#unique_112、#unique_113
2019.03.26	Java UDX类型系统支持UDF Writable参数	新规格	Java UDX类型系统增强，支持UDF Writable参数。	#unique_13
2019.03.26	新增支持INTERSECT/MINUS/EXCEPT集合操作	新规格	支持INTERSECT/MINUS/EXCEPT集合操作。	#unique_12

时间	特性	类别	描述	产品文档
2019.03.26	新增支持Grouping Set多维聚合分析	新规格	支持持Grouping Set多维聚合分析（Cube, Rollup），解决既需要对a列做聚合，也要对b列做聚合，或者同时要按照a，b两列做聚合的，因此需要写很多UNION ALL类场景问题。	#unique_11
2019.02.27	新增获取URL中指定位置的字符示例	体验优化	向您介绍如何应用indexOf()、substring()、lastIndexOf()等字符串处理方法获取URL中指定位置的字符的Java UDF示例。	#unique_114
2019.02.27	新增取余函数示例	体验优化	向您介绍了如何实现不同类型数值相除获取余数的Java UDF示例。	#unique_115
2019.02.26	新增获取指定日期格式的时间示例	体验优化	向您介绍如何通过evaluate()方法获取指定格式时间的Java UDF示例。	#unique_116
2019.02.14	新增获取字符串（不含分隔符）value示例	体验优化	向您介绍如何获取不包含分隔符的字符串中指定key对应的value值的Java UDF示例。	#unique_117
2019.02.13	新增使用正则表达式替换字符串示例	体验优化	向您介绍如何实现使用正则表达式替换字符串的Java UDF示例。	#unique_118
2019.01.31	常见问题目录调整	体验优化	常见问题目录细分，文档归类整理。	常见问题

时间	特性	类别	描述	产品文档
2019.01.29	MaxCompute账单分析最佳实践	体验优化	向您介绍如何获取MaxCompute账单信息，并通过SQL分析账单数据的最佳实践	#unique_119
2019.01.28	新增使用Java SDK运行安全命令最佳实践	体验优化	向您介绍在MaxCompute Console上运行安全相关的命令的最佳实践	#unique_72
2019.01.24	新增计量计费中对Lightning按量后付费的修订	新规格	增加MaxCompute计量计费中对Lightning按量后付费的修订	#unique_14
2019.01.24	新增DataWorks实现指定UDF被指定用户访问最佳实践	体验优化	向您介绍如何在DataWorks上实现将具体的某个资源（表、UDF等）设置为仅能被指定的用户访问。	DataWorks实现指定UDF被指定用户访问
2019.01.21	安全指南目录调整、增加基础篇和案例篇文档	体验优化	安全指南目录增调整为一级目录，并增加基础篇和案例篇文档	#unique_120
2019.01.18	新增JSON字符串增加键值示例	体验优化	向您介绍如何实现向基础JSON字符串中增加（覆盖）一个或多个键值对（支持嵌套JSON），且接受奇数个参数的Java UDF示例。	#unique_121
2019.01.17	新增不同类型数据转换JSON String示例	体验优化	向您介绍如何实现将不同类型的数据转换为JSON String的Java UDF示例。	#unique_122

时间	特性	类别	描述	产品文档
2019.01.17	新增JSON字符串获取UDF示例程序	体验优化	向您介绍如何实现根据path路径获取JSON字符串并对指定的值进行判断是否存在的Java UDF示例。	#unique_123
2019.01.09	修订Lightning文档	体验优化	标题“Lightning”更改为“交互式分析（Lightning）”。下线服务定价文档。	交互式分析（Lightning）概述
2019.01.09	修订计量计费文档，增加Lightning计费和Spark按量后付费说明	新规格	从2019年2月1日开始，MaxCompute 正式对Lightning服务、Spark任务开始计费。	#unique_14
2019.01.09	增加Spark文档	体验优化	Spark是MaxCompute提供的兼容开源的Spark计算服务，它在统一的计算资源和数据集权限体系之上，提供Spark计算框架，支持用户以熟悉的开发使用方式提交运行Spark作业，以满足更丰富的数据处理分析场景。	#unique_125
2019.01.04	新增UDF示例程序目录和JSON字符串获取示例	体验优化	向您详细介绍如何实现根据path路径获取JSON字符串的Java UDF示例。	#unique_126

2018年更新记录

时间	特性	类别	描述	产品文档
2018.12.20	新增使用Intellij IDEA开发工具进行Java UDF开发最佳实践	体验优化	向您详细介绍如何使用Intellij IDEA开发工具进行Java UDF开发。	#unique_127
2018.12.18	新增解决DataWorks 10M文件限制问题最佳实践	体验优化	向您介绍如何解决用户在DataWorks上执行MapReduce作业的时候，文件大于10M的JAR和资源文件不能上传到Dataworks的问题。	#unique_128
2018.12.17	新增使用Eclipse开发工具进行Java UDF开发最佳实践	体验优化	向您详细介绍如何使用Eclipse开发工具进行Java UDF开发。	#unique_129
2018.12.04	大数据计算服务MaxCompute欧洲西部1（伦敦）节点正式开服售卖	新地域/可用区	大数据计算服务MaxCompute 欧洲西部1（伦敦）节点开服售卖。 可在国际站、中文站、日本站购买节点资源，在控制台按需求开通project。	#unique_130
2018.11.28	新增在MaxCompute上分析IP来源的方法最佳实践	体验优化	向您介绍如何在MaxCompute上分析IP来源的方法，包括下载、上传淘宝IP地址库数据及编写UDF函数、编写SQL四个步骤。	#unique_131

时间	特性	类别	描述	产品文档
2018.11.26	新增从MongoDB提取JSON字段到MaxCompute最佳实践	体验优化	向您介绍如何利用DataWorks数据集成直接从MongoDB提取JSON字段到MaxCompute。	#unique_132
2018.11.15	处理非结构化数据目录更新为外部表	体验优化	处理非结构化数据目录更新为外部表。	外部表
2018.11.13	新增JSON数据从OSS迁移到MaxCompute最佳实践	体验优化	向您介绍如何利用DataWorks数据集成将JSON数据从OSS迁移到MaxCompute，并使用MaxCompute内置字符串函数GET_JSON_OBJECT提取JSON信息。	#unique_133
2018.11.08	新增与MySQL、Oracle内建函数对照表	体验优化	帮助您根据MySQL内建函数查找对应的MaxCompute内建函数。	#unique_134
2018.11.08	标题“导读”更新为“用户必读”，添加MaxCompute使用限制汇总	体验优化	标题“导读”更改为“用户必读”，增加MaxCompute SQL、数据上传下载、MapReduce、安全配置、Lighting、PyODPS等使用限制汇总。	#unique_135
2018.11.05	MaxCompute SQL新窗口函数cume_dist	新功能	该函数用于取得累计分布，相当于求分组中值小于等于当前值的行数占分组总行数的比例。	#unique_136

时间	特性	类别	描述	产品文档
2018.11.03	将Java SDK、Python SDK和PyODPS DataFrame自定义函数中使用第三方包三篇文档移动到SDK参考目录	体验优化	将用户指南中的Java SDK、Python SDK和PyODPS DataFrame自定义函数中使用第三方包三篇文档统一移动到SDK参考目录。	#unique_137、#unique_138
2018.11.02	新增RDS迁移到MaxCompute实现动态分区最佳实践	体验优化	向您介绍如何将RDS数据迁移到MaxCompute表实现动态分区。	#unique_139
2018.10.31	MaxCompute SQL外部表功能计费	新规格	从2018年10月31日开始，MaxCompute SQL外表功能开始计费。MaxCompute下创建和OTS表存储关联的表格，使用SQL查询时，无论通过外网还是经典/VPC网络访问OTS（除非使用STS方式访问），OTS都会收取费用。当您需要内外表混合作业的时候，会分别计费。	#unique_14
2018.10.16	新增数据迁移文档	体验优化	整理将数据移迁移到MaxCompute的最佳实践相关文档。	#unique_140
2018.9.29	MaxCompute SQL引擎新功能User Defined Type（简称UDT）	新功能	允许在SQL中直接引用第三方语言的类或者对象，获取其数据内容或者调用其方法。	#unique_141

时间	特性	类别	描述	产品文档
2018.9.29	MaxCompute UDF框架中新近引入的新扩展机制	新功能	满足了MaxCompute内置Join操作无法实现的灵活的跨表、多表自定义操作，并且减少了对分布式系统底层细节的操作，提高了执行效率。	#unique_142
2018.9.27	PyODPS已全面对外开放	新规格	PyODPS是MaxCompute Python版本的SDK，提供了DataFrame框架，现已全面对外开放。DataWorks也推出了PYODPS任务类型，可在DataWorks的PYODPS节点上直接编辑Python代码操作Maxcompute。	PyODPS
2018.7.30	新增一些典型的长尾问题的场景及其解决方案最佳实践	体验优化	向您介绍平时工作中遇到的一些典型的长尾问题的场景及其解决方案。	#unique_144
2018.5.25	Maxcompute支持OSS压缩格式gzip	新功能	MaxCompute目前只支持通过内置extractor读取OSS上gzip压缩的CSV/TSV数据。与读取非压缩文件相比，主要区别是SERDEPROPERTIES指定的属性。	#unique_145

时间	特性	类别	描述	产品文档
2018.5.25	Maxcompute支持OSS上的Hive文件格式	新规格	存储在OSS上的各种流行的开源数据格式（ORC、PARQUET、SEQUENCEFILE、RCFILE、AVRO和TEXTFILE）可通过非结构化框架在MaxCompute进行处理。	#unique_146
2018.5.25	Maxcompute SQL新语法Select Transform	新功能	Select Transform功能明显简化了对脚本代码对引用，支持Java、Python、Shell、Perl等语言，且编写过程简单，适合adhoc功能的实现。	#unique_147
2018.5.25	Maxcompute SQL新语法SEMI JOIN	新功能	MaxCompute支持SEMI JOIN（半连接）。SEMI JOIN中，右表只用来过滤左表的数据而不出现在结果集中。支持LEFT SEMI JOIN和LEFT ANTI JOIN两种语法。	#unique_148
2018.5.25	Maxcompute SQL支持CASE WHEN的隐式类型转换	新规格	Maxcompute SQL支持CASE WHEN的隐式类型转换。	#unique_149/unique_149_Connect_42_se
2018.5.16	MaxCompute提供全表扫描的设置操作	新功能	MaxCompute提供全表扫描的设置操作。	#unique_19

时间	特性	类别	描述	产品文档
2018.4.27	Dataworks支持MaxCompute PYODPS任务	新规格	DataWorks推出了PYODPS任务类型，集成了Maxcompute的Python SDK，可在DataWorks的PYODPS节点上直接编辑Python代码操作Maxcompute，也可以设置调度任务来处理数据，提高数据开发效率。	PYODPS节点
2018.4.14	MaxCompute自动续费支持过期实例续费	新规格	原先预付费用户实例过期，在续费中心找不到过期实例的续费方式，耽误项目正常使用，新功能可以帮助用户通过续费中心很方便的完成手动续费，保障项目的正常使用，操作步骤为登陆后选择左侧大数据 > 数加控制台 > MaxCompute续费管理台。	无
2018.4.10	大数据计算服务MaxCompute 美国东部1（弗吉尼亚）节点正式开服售卖	新地域/可用区	大数据计算服务MaxCompute 美国东部1（弗吉尼亚）节点开服售卖。 可在国际站、中文站、日本站购买美东节点资源，在控制台按需求开通Project。	#unique_130

时间	特性	类别	描述	产品文档
2018.4.2	MaxCompute Studio 2.9.0 新版本发布	新功能	支持代码检查和快速修复 2.支持graph开发及调试 3.更多改进及Bug Fix，请更新后查看Release Notes：MaxCompute Studio 2.9.0。	#unique_150
2018.3.28	MaxCompute CU管家支持华东1杭州Region	新规格	本次MaxCompute为华东1预付费用户带来了计算资源管理功能；之前，MaxCompute 华东1预付费用户购买了CU计算资源，但并不清楚资源消耗的情况。现在，CU管家不仅可以帮助企业客户了解计算资源的消耗情况，还可以帮助企业客户排查计算任务排队情况，并根据任务重要程度，分组管理计算任务。	#unique_151
2018.3.23	MaxCompute Studio支持查询结果自定义Tunnel endpoint	修复问题	修复功能： 使用Instance Tunnel下载 SQL 结果时，用户可以指定Tunnel endpoint。	配置Endpoint

时间	特性	类别	描述	产品文档
2018.3.22	大数据计算服务MaxCompute 华东1（杭州）节点正式开服售卖	新地域/可用区	大数据计算服务MaxCompute 华东1（杭州）节点开服售卖。 可在中文站、国际站、日本站购买杭州节点资源，在控制台按需求开通Project。	#unique_130
2018.3.22	增强SQL 编辑器功能	体验优化	DataWorks通过在前端实现MaxCompute SQL和编辑器参数等扩展语法的AST解析，来提升脚本开发效率和开发体验，就像本地编辑器一样，提供智能代码提示，语法高亮，语法检查，风险及错误提示，类型检查等一系列功能。	DataWorks 增强SQL编辑器功能

时间	特性	类别	描述	产品文档
2018.2.9	PyODPS 0.7.15 版本发布	新版本/新规格	PyODPS 0.7.15 较与之前版本新增、优化和修复了相关功能，并提升了开发者的体验与效率： 1、在 Python 3.6版本上，开发者可以直接使用 DataFrame 自定义函数了； 2、修复 Python 3.5 版本上, DataFrame 自定义函数可能会导致错误的问题; 3、支持Xflow 运行时可以设置运行参数; 4、从 pandas DataFrame 创建 PyODPS DataFrame 时, as_type 参数可以指定组合类型（如 list 或者 dict 类型）； 5、修复其他BUG。	#unique_138
2018.2.9	大数据计算服务 MaxCompute 印尼开服，亚太东南5（雅加达）节点正式售卖	新地域/可用区	大数据计算服务 MaxCompute 亚太东南5（雅加达）节点正式开服售卖。 可在国际站、中文站、日本站购买雅加达节点资源，在控制台上按需求开通 Project。	#unique_130

时间	特性	类别	描述	产品文档
2018.2.8	MaxCompute服务等级协议正式发布	体验优化	MaxCompute服务已对接服务等级监控（SLA）平台，用户可以直接在官网控制台可用性中心查看每月可用性数据。 服务承诺可实现99.9%的正常运行时间和可用性，如果未能实现承诺，客户可以提出赔偿申请。	MaxCompute服务等级协议
2018.2.5	新增函数“TRANS_ARRAY”说明	新功能	该函数用于将一行数据转为多行的UDTF，将列中存储的以固定分隔符格式分隔的数组转为多行。	其他函数

时间	特性	类别	描述	产品文档
2018.2.3	MaxCompute Studio 2.8.2.3版本发布	体验优化	1.MaxCompute Studio Editor的侧边拦针对完整sql语句增加执行功能。原本当Editor工作区中有多条sql语句而仅需要执行其中某条sql语句时，需要先选中对应语句，再点击工具栏的执行按钮，才会只执行该条语句。本次优化后，可直接在左侧栏对对应的sql语句操作执行，符合开发者习惯，减少操作步骤，提升使用体验。 2.MaxCompute Studio Project Explorer中按类型划分内置函数。原本的函数列表不分类型，一个大列表显示所有函数，查找不方便，同时也无法区分名称相同类型和使用方式不同的函数。本次优化后内置函数列表按类型划分，与官方文档对应，方便开发者查找并使用对应的函数。	无
2018.2.3	添加 top instance命令说明	新功能	添加top instance命令说明。	#unique_152

时间	特性	类别	描述	产品文档
2018.2.3	添加如何查看外部表的”external info”	新功能	添加如何查看外部表的External Info”（包括LOCATION、StorageHandler等）信息的介绍。	#unique_153、#unique_154
2018.2.2	添加MaxCompute管家介绍入口	体验优化	MaxCompute管家就是解决这个预付费计算资源监控管理的问题。目前MaxCompute管家主要提供三个功能，系统状态、资源组分配、任务监控。添加MaxCompute管家介绍入口。	MaxCompute管家
2018.2.1	MaxCompute Console支持top instance命令	新规格	MaxCompute Console 0.29.0版本推出top instance命令，支持开发者通过该命令查询当前项目空间内正在执行或排队的作业具体信息，方便开发者、管理员查看与管理当前项目作业情况。	#unique_152

时间	特性	类别	描述	产品文档
2018.1.31	MaxCompute CU管家	新功能	MaxCompute为预付费用户带来了计算资源管理功能；之前，MaxCompute 预付费用户购买了CU计算资源，但并不清楚资源消耗的情况，现在，CU管家不仅可以帮助企业客户了解计算资源的消耗情况，还可以帮助企业客户排查计算任务排队情况，并根据任务重要程度，分组管理计算任务。	#unique_155
2018.1.22	DataHub 新增四大Region	新地域/可用区	DataHub新增开通华东2（上海）、华北2（北京）、华南1（深圳）以及亚太东南1（新加坡），可以在管控台上根据业务需求开通服务。	无
2018.1.18	MaxCompute Studio 2.8.2.2版本发布	新版本/新规格	MaxCompute Studio 2.8.2.2 主要发布的功能有四点： 功能一： 支持python udf开发。 功能二： 支持非结构化unit test编写。 功能三：支持UDF local run时对复杂数据类型进行操作。 功能四：支持自定义sql语句格式化规则。	#unique_150

时间	特性	类别	描述	产品文档
2018.1.18	大数据计算服务MaxCompute 印度开服，亚太南部1（孟买）节点正式售卖	新地域/可用区	大数据计算服务MaxCompute 亚太南部1（孟买）节点正式开服售卖。 可在国际站、中文站、日本站购买孟买节点资源，在控制台上按需求开通project。	#unique_130
2018.1.15	基于Package的跨项目空间的资源分享功能发布	新功能	Package是一种跨项目空间共享机制，主要用于解决跨项目空间授权问题。	#unique_156
2018.1.15	基于LabelSecurity的列级别安全访问控制功能发布	新功能	MaxCompute推出基于表字段级别的安全访问控制，更细粒度的管控。	#unique_157
2018.1.10	MaxCompute禁止Full Scan功能开放	新功能	禁止全表扫描分区表，避免资源浪费，同时防止全表扫描带来的不必要费用。	#unique_158
2018.01	新增项目空间的数据保护文档	体验优化	介绍项目空间的数据保护背景动机，介绍保护机制以及在开启保护机制后如何让数据流出的方法。	#unique_111
2018.01	新增“Hive UDF兼容示例”示例	体验优化	以一个示例介绍Hive的UDF如何在MaxCompute上使用。	#unique_159
2018.01	添加TableTunnel接口说明以及底层实现和使用限制	体验优化	添加TableTunnel接口说明以及底层实现和使用限制。	#unique_160

时间	特性	类别	描述	产品文档
2018.01	添加公开数据集介绍入口，给试用产品的用户提供测试数据	体验优化	添加公开数据集介绍入口，给试用产品的用户提供测试数据。	公开数据集介绍
2018.01	新增Full scan功能开放的通告	新功能	由于Full Scan功能发布，在这些文档中添加相关说明。	通告、常用命令-其他操作SQL常见错误

3 发展历程

本文为您介绍了MaxCompute从诞生到成熟的发展历程。了解更多MaxCompute的最新动态请参见#unique_162。

关键性里程碑

- 2009年9月，MaxCompute大数据平台飞天项目正式启动。
- 2010年10月，阿里巴巴集团自主研发的第一代云计算平台稳定运行。
- 2013年8月，平台的单集群规模已达到5000台。
- 2014年7月，平台开始对外提供服务，完全替换开源体系。
- 2015~2018年，平台开始日趋成熟，单集群已过万台，性能再优化，实现了全球部署。

产品荣誉

- 2018年11月， MaxCompute，DataWorks和AnalyticDB代表阿里云入选Forrester Wave™ Q4 2018云数据仓库报告。
- 2018年9月， 基于公共云的Bigbench在100TB规模上，MaxCompute的性能指标较2017年10月提升了一倍，达到18176.71QPM。此外，在超小型10TB规模的指标上，MaxCompute的性能是其它开源竞品性能的3倍。
- 2018年4月，MaxCompute的多个客户案例荣获“2017大数据优秀产品和应用解决方案案例”奖。
- 2018年3月，MaxCompute登上Forrester《2018年一季度云端数据仓库》大数据服务榜单。
- 2018年3月，Gartner发布了《2017年分析型数据管理解决方案（DMSA）魔力象限》报告，阿里云作为云服务商成功冲进Gartner魔力象限。
- 2017年10月，TPC的benchmark适配MaxCompute，进行了全球首次基于公共云的bigbench大数据基准测试，数据规模拓展到100TB，成为首个突破7000分的引擎，性能达到7830QPM。
- 2017年6月， MaxCompute获得中国国际软件博览会金奖。
- 2016年11月，在CloudSort竞赛中，MaxCompute以$0.82/TB的成绩获得Indy（专用目的排序）和Daytona（通用目的排序） 两个子项的世界冠军，打破了AWS在2014年保持的纪录 $ 4.51/TB。
- 2015年10月，在GraySort竞赛中，MaxCompute用377秒完成了100TB的数据排序，打破了此前Apache Spark创造的1406秒纪录。

产品认证

- **国内首家工信部单集群万台扩展能力认证**
- **工信部信通院和中电标准化研究院认证**

深度参与和推动全球大数据领域标准化建设

- **MaxCompute代表阿里巴巴计算平台，成为国际TPC委员会大数据评测标准BigBench的专委，是国内唯一的一家入选企业。**
- **全球两大热门计算存储标准化开源体系ORC社区的PMC，MaxCompute成为近两年贡献代码量最多的贡献者，引导存储标准化。**
- **MaxCompute积极投入全球最热门优化器项目Calcite，拥有一个专委席位，是国内前两家具备该领域影响力的公司。**

4 新功能测试申请

本文为您介绍如何进行新功能测试申请以及每月最新的可测试功能。

参与新功能测试申请方法

1. 进入提交工单页面，在选择问题所属产品时单击MaxCompute后的提问。

2. 单击MaxCompute新功能测试申请。

3. 在右侧弹窗中单击创建工单。

4. 请详细填写表单中的内容，尤其是框中的项目，以便判断新功能与您业务场景的匹配度，提高申请通过率。

5. 工单提交后，我们会尽快审核您的测试申请，根据您的业务需求评估与新功能的匹配度。如果通过审核，我们会通过工单、电话或邮件的形式联系您。

2019年8月新功能测试申请

- 权限模型2.0

  MaxCompute权限模型2.0是对过去权限系统重构和升级，实现更精准、细颗粒度的授权机制。权限模型2.0包括：支持表列级别权限控制、Label授权到Role、Policy支持GRANT方式授权、Package资源支持细粒度授权等。权限模型2.0具备更细颗粒度的安全管理能力，支持更多数据安全管理和授权场景。

- 数据脱敏

  MaxCompute基于数据保护伞共同支持数据脱敏功能，关于数据保护伞详情请参见#unique_164。数据脱敏支持默认脱敏规则和自定义脱敏规则，您可以通过Flag关闭或打开数据脱敏功能。您还可以对某些敏感信息通过脱敏规则进行数据变形，保护您的敏感隐私数据。对您的商业敏感数据，在不违反系统规则的前提下，可以对您的真实数据（身份证号、手机号、卡号、客户号等个人信息）进行改造并测试。

- 存储加密

  MaxCompute提供数据存储加密功能，最大程度保障您的数据安全。即使您的数据遗失或被盗，也无法被解读出有意义的内容。MaxCompute以项目为单位支持表的加密存储，目前仅支持全表加密。存储支持两种加密算法：AES CTR或RC4算法。MaxCompute接入KMS密钥管理系统，以保障密钥的安全性。

- Information Schema元数据服务

  MaxCompute Information_Schema以视图的形式提供准实时的元数据，包括Tables、Partitions、Columns、UDFs、Resources、Users、Roles、Packages、数据对象的授权信息、作业运行信息等。您可以基于准实时的Information Schema数据，实现对数据的权限管理及行为审计。

2019年9月新功能测试申请

VPC网络IP白名单设置

MaxCompute支持设置VPC内网络IP以及VPC ID作为白名单设置，对VPC环境的访问控制将更加灵活和精细化。过去仅支持设置VPC的出口IP作为白名单设置，且为非region化的域名，加入VPC的内部IP无效果。