

ALIBABA CLOUD

阿里云

数据管理
最佳实践

文档版本：20201024

 阿里云

法律声明

阿里云提醒您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
 危险	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。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会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，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	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。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <i>Instance_ID</i>
[] 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>switch {active stand}</code>

目录

1.表结构发布流程化	05
2.跨云账号数据库管理	10
3.在DMS中实现DLA一键建仓与数据开发	13

1.表结构发布流程化

本文档将以设计结构（Develop，后面简称Dev）、测试（Test）、生产发布（Product）的表结构设计步骤向您介绍规范、安全的表结构发布流程。

背景信息

在新项目、新需求或者优化等业务场景中，一般会涉及到新建表、修改表（加减字段、加减索引、调整字段属性、调整索引组成）等结构的变更操作，此时您可以使用DMS的结构设计功能。

DMS推出的结构设计为多套研发环境开发，且在开发过程中自动检测企业需遵循的表结构研发规范，并支持多人在Web页面协同开发的功能。

结构设计功能可保障研发自助变更的同时满足企业的设计规范、不同环境之间相关表结构的一致性，同时大大提升生产发布的成功率。

结构设计的功能说明如下。

功能	说明
保障生产发布的成功率	新的表结构在发布到生产环境前，会依据研发流程在逐个节点中对表结构进行验证，可避免未验证脚本直接发布到生产环境而导致发布失败的问题，保障生产发布的成功率。
保障环境结构的一致性	系统会依据研发流程对逐个环境进行结构变更，可避免因部分环境未执行而导致的测试延误、项目整体发版进展延误等。
保障数据结构的规范性	安全规则 的研发规范包含新建表、修改表字段、修改表索引等多方面的近40条规范，方便研发同学自助设计满足企业研发规范的表结构，可避免因不满足规范的结构发布上线而导致的生产故障。同时，该功能可节省运维人员因人工检查规范而带来的人力投入，并提升研发效能，可以有效支撑快速迭代发版工作。

搭建示例的测试环境

1. 创建3个RDS MySQL数据库实例，分别为：Schema_Design-dev、Schema_Design-test、Schema_Design-product，操作详情请参见[创建RDS MySQL实例](#)。
2. 将3个数据库实例录入至DMS中，将管控模式均设置为安全协同，并分别将数据库环境类型设置成开发（Dev）、测试（Test）、生产（Product）。云数据库录入的操作详情请参见[云数据库录入](#)，数据库实例环境的说明请参见[实例环境说明](#)。
3. 在3个数据库实例中分别创建3个对应的数据库：dev_database、test_database、product_database。

数据库实例与结构设计各节点的对应关系如下。

结构设计节点	设计结构（Dev）	测试（Test）	生产发布（Product）
对应数据库实例名称	Schema_Design-dev	Schema_Design-test	Schema_Design-product
环境类型	dev	test	product
数据库名称	dev_database	test_database	product_database

步骤一：配置研发规则

默认的研发流程为：设计结构（Dev）、生产发布（Product）。以下步骤将介绍如何新建安全规则并配置设计结构（Dev）、测试（Test）、生产发布（Product）的研发流程。

1. 登录DMS控制台。
2. 在页面顶部的菜单栏中，单击系统管理 > 安全管理 > 安全规则。



3. 在安全规则页面中，单击左上角的新增规则集按钮，新增一个规则集。

 说明 规则及名称可自定义，本示例中以SD-DT为例。



4. 在安全规则详情页面左侧的导航栏中，单击结构设计。
5. 在结构设计页签中，单击研发流程右侧的操作列下的编辑。



6. 在编辑配置项对话框中，完成如下配置：
 - i. 将第二条节点名称改为 测试 ，将库环境改为 test 。
 - ii. 新增第三条节点，将第三条节点名称改为 生产发布 ，将库环境改为 product ，同时该节点设置为锚点。

 说明 锚点：为工单编辑的冻结点，即锚点所在步骤成功发布之后，工单将会变成已发布状态，不可编辑。

配置界面如下图所示。



7. 单击提交。

步骤二：设置数据库实例的安全规则

将Schema_Design-dev数据库实例（环境类型为Dev）的管控模式设置为安全协同并选择步骤一创建的SD-DT安全规则。

1. 登录DMS控制台。
2. 在控制台首页左侧的实例列表中，右键单击Schema_Design-dev数据库实例。
3. 在弹出的列表中，选择管控模式 > 安全协同 > SD-DT。



步骤三：进行结构设计

以下步骤以新建数据库的表结构变更为例，向您介绍设计结构（Dev）、测试（Test）、生产发布（Product）三节点的结构设计项目。

1. 登录DMS控制台。
2. 在顶部菜单栏，选择库表结构 > 结构设计。



3. 单击页面右上角的结构设计。



4. 配置结构设计工单信息，并单击创建工单。

配置	说明
项目名称	填写项目名称便于后期管理，本示例中以结构设计测试为例。
业务背景	描述项目的业务背景信息，例如变更目的、预期目标等，减少沟通成本。
变更基准库	选择dev_database数据库，因为在步骤二时，已将该实例的安全规则设置为ST-DT，系统会根据该安全规则的研发流程创建出设计结构（Dev）、测试（Test）、生产发布（Product）的结构设计流程。 <div style="border: 1px solid #add8e6; padding: 5px; margin-top: 10px;"> ? 说明 系统会根据第一个变更基准库的安全规则创建结构设计的研发流程。 </div>
安全规则	DMS会自动选择，无需配置。
变更相关人	单击新增设置相关人员，被设置的人员都可查看工单并协同工作，非相关人员则不能查看工单（管理员、DBA除外）。

创建工单后，系统会自动跳转到结构设计项目首页，该项目共有5个节点，如下图所示。

5. 在设计结构节点进行配置：

- i. 单击导入建表语句。
- ii. 在文本框中输入以下语句，并单击确定。

```
CREATE TABLE `teacher` (
  `teacher_id` varchar(20) NOT NULL COMMENT '教师编号',
  `teacher_name` varchar(20) NULL COMMENT '教师名字',
  `phonenumner` varchar(20) NULL COMMENT '教师电话号码',
  PRIMARY KEY (`teacher_id`)
) DEFAULT CHARACTER SET=utf8 COMMENT='教师表';
```

iii. 在控制台界面确认基本信息、字段、索引等信息。

■ 基本信息

■ 字段

■ 索引

iv. 单击保存，并在弹出的预检测结果窗口中单击确认无误，提交保存。

- v. 单击执行变更到基准库。
- vi. 在弹出的对话框中，选择执行策略为立即执行，并在确认最终的SQL语句无误后单击提交执行。
- vii. 提交执行后请等待审批完成，待审批完成后，系统会将新的表结构写入到dev_database数据库中。
- viii. 您可以前往SQL Console插入数据或提交数据变更工单对dev_database数据库进行测试。节点验证通过时，即可将该变更结构发布到下一节点中。

示例代码如下。

```
insert into teacher
(teacher_id, teacher_name, phonenumber)
values
('t001', 'teacher001', '1380000001'),
('t002', 'teacher002', '1380000002'),
('t003', 'teacher003', '1380000003');
```

- ix. 在结构设计工单页面中，单击进入下一节点。

 说明 在进行该项操作前，您可以与您的同事在该节点对dev_database数据库进行多次变更设计与验证，系统最终会产生一个 create 脚本，发布到下个节点。

- x. 在弹出的对话框中，再次单击进入下一节点。
6. 在测试节点进行配置：
- i. 单击执行变更到目标库。
 - ii. 选择test_database数据库，并将执行策略设置为立刻执行，单击提交执行。

 说明 提交执行后请等待审批完成，待审批完成后，系统会将变更的结构写入到test_database数据库中。

- iii. 您可以前往SQL Console插入数据或提交数据变更工单对test_database数据库进行测试。

 说明 当该节点验证通过时，即可将该变更结构发布到下一节点中。

- iv. 在结构设计工单页面中，单击进入下一节点。

- v. 在弹出的对话框中，再次单击进入下一节点。

7. 在生产发布节点进行配置：

- i. 单击执行变更到目标库。

- ii. 选择product_database数据库，并将执行策略设置为立刻执行，单击提交执行。

提交执行后请等待审批完成。待审批完成后，系统会将变更的结构写入到product_database数据库中，您也可以将数据写入生产环境中（product_database数据库）。

- iii. 在结构设计工单页面中，单击进入下一节点。

- iv. 在弹出的对话框中，再次单击进入下一节点。

整个流程结束，工单关闭。

2. 跨云账号数据库管理

本文以在DMS中接入多个账号，并录入RDS MySQL数据库为例介绍如何录入跨云账号数据库。

背景信息

当您的企业中存在多个主账号，并且希望跨账号统一管理企业内所有的数据库资源，您可以在您的DMS中添加多个用户，然后按需录入数据库实例。录入后您的账户下所用用户均可访问该数据库。

注意事项

- 当前跨云账号录入仅支持RDS MySQL、RDS SQL Server、RDS PPAS、RDS PostgreSQL类型的数据库。
- 当前仅支持以连接串地址的方式录入跨云账号数据库，暂不支持以实例ID的方式录入。
- 您只能添加您当前账户（含主账号）下的用户的数据库实例，如下图所示：



添加用户

1. 登录[DMS控制台](#)。
2. 在顶部导航栏，单击系统管理 > 用户管理。
3. 单击页面左上角的新增按钮。



4. 在阿里云账号文本框中，输入需要添加用户的账号。



 **说明** 账号获取方法：账号所有人访问[阿里云账号uid](#)，查看账号ID。



5. 为待添加用户选择一个角色。

 **说明**

- 普通用户：不具备额外权限。仅允许登录数据管理DMS服务，服务内的所有操作按需申请权限后方可处理。更多关于普通用户的介绍，请参见[我是普通用户](#)。
- DBA：具备所有数据库表的直接查询权限，具备系统管理-实例管理、任务管理、安全规则、配置管理这几项系统管理入口的操作，但不具备提交非权限申请以外其他类型工单的权限（需要显示开通权限才可提交）。更多关于DBA的介绍，请参见[我是DBA](#)。
- 管理员：具备所有数据库表的直接查询权限，具备系统管理-实例管理、任务管理、安全规则、配置管理、用户管理、操作日志、访问IP白名单等这几项系统管理入口的操作，但不具备提交非权限申请以外其他类型工单的权限（需要显示开通权限才可提交）。更多关于管理员的介绍，请参见[我是管理员](#)。
- 安全管理员：具备所有数据库表的直接查询权限，具备系统管理-智能化运维、操作日志、数据保护伞、敏感数据管理、权限管理、库表结构、数据方案等这几项系统管理入口的操作，但不具备提交非权限申请以外其他类型工单的权限（需要显示开通权限才可提交）。更多关于安全管理员的介绍，请参见[我是安全管理员](#)。

6. 单击确认即可。

按需录入数据库

1. 登录DMS控制台。
2. 单击页面左上角的新增实例 / 批量录入，在弹框中单击新增实例按钮。



3. 单击云数据库标签页。
4. 在云数据库区域中选择MySQL。
5. 在弹出的新增实例对话框中，按照下表录入实例信息。



标签页名称	参数名称	说明
基本信息	数据库来源	选择数据库实例的来源。此处选择云数据库。
	数据库类型	选择数据库实例的类型。
	实例地区	选择数据库实例所在的地区。
	录入方式	当前仅支持通过连接串地址的方式录入跨账号资源。
	连接串地址	输入实例的连接地址。 <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;"> ? 说明 数据库接入DMS时，系统会自动尝试往云数据库的白名单中添加DMS的服务器访问地址，若自动添加失败请手动添加，详情请查看RDS设置IP白名单，DMS白名单列表。 </div>
	数据库账号	输入数据库的登录账号。
	数据库密码	输入数据库账号对应的密码。
	管控模式	选择数据库的管控模式，详情请参见 管控模式
高级信息	环境类型	选择数据库环境的类型。
	实例名称	自定义实例的名称。
	开启跨库查询	选择是否开启跨库查询功能，详情请参见 跨库查询
	不锁表结构变更	选择是否开启不锁表结构变更功能，详情请参见 不锁表结构变更

高级信息 标签页名称	参数名称	说明
	实例DBA	选择一个DBA角色进行后期权限申请等流程。
	查询超时时间 (s)	设定安全策略, 当达到设定的时间后, SQL窗口执行的查询语句会中断, 以保护数据库安全。
	导出超时时间 (s)	设定安全策略, 当达到设定的时间后, SQL窗口执行的导出语句会中断, 以保护数据库安全。

6. 完成以上信息填写, 在基本信息标签页中, 单击左下角的测试连接, 等待测试通过。

 说明 如果测试连接失败, 请按照报错提示检查您录入的实例信息。

7. 连接成功, 单击提交。

3.在DMS中实现DLA一键建仓与数据开发

本文将介绍在DMS中接入DLA数据湖，利用DLA一键建仓功能将RDS MySQL数据库接入DLA数据库，并在DMS中对DLA数据库进行任务流开发，简要介绍在DMS接入DLA的开发与应用。

背景信息

数据湖分析（Data Lake Analytics，简称DLA）是无服务器（Serverless）化的云上交互式查询分析服务。DLA完全采用SQL与服务端进行交互，兼容标准的SQL，提供丰富的内置函数支持。DLA支持多路数据源接入分析，提供了多样化、异构的数据源分析能力。您不仅能够对阿里云OSS中的数据、Tablestore中的数据进行分析，还能将这两者之间的数据进行关联性分析，可以帮助您解决将不同种类的数据进行联合分析的问题。同时DLA全面融合MPP（Massively Parallel Processing，即大规模并行处理）和DAG（Directed Acyclic Graph）技术，超强的横向分析扩展能力、向量化执行优化、算子流水线优化，提供良好的资源隔离和优先级调度。使用DMS任务流在DLA上做周期性数据分析可获得如下好处：

- 任务流只需一次定义，即可周期性地自动被调度执行，大大减少人工操作的成本。
- 周期执行任务并提前产生出分析结果，业务应用可直接查询结果获得快速响应。
- 任务流执行时产生的中间结果可被复用，用于其他的数据分析任务，从而最大化DLA的资源利用。

添加RDS实例

1. 登录**DMS控制台**。
2. 单击页面左上角的新增实例 / 批量录入，在弹框中单击新增实例按钮。
3. 单击云数据库标签页。
4. 在云数据库区域中选择MySQL。
5. 在弹出的新增实例对话框中按照下表录入实例信息。

标签页名称	名称	说明
基本信息	数据库来源	选择数据库实例的来源。此处选择云数据库。
	数据库类型	选择数据库实例的类型。
	实例地区	选择数据库实例所在的地区。
	录入方式	选择数据库实例的录入方式。支持实例ID和连接串地址两种录入方式。
	实例ID	选择数据库实例的实例ID。该选项仅在录入方式选择实例ID时可见。
	连接串地址	输入实例的连接地址。该选项仅在录入方式选择连接串地址时可见。
	数据库账号	输入数据库的登录账号。
	数据库密码	输入数据库账号对应的密码。
	管控模式	选择数据库的管控模式，详情请参见 管控模式 。

标签页名称	名称	说明
高级信息	环境类型	选择数据库环境的类型。
	实例名称	自定义实例的名称。
	开启跨库查询	选择是否开启跨库查询功能，详情请参见 跨库查询 。
	不锁表结构变更	选择是否开启不锁表结构变更功能，详情请参见 不锁表结构变更 。
	实例DBA	选择一个DBA角色进行后期权限申请等流程。
	查询超时时间 (s)	设定安全策略，当达到设定的时间后，SQL窗口执行的查询语句会中断，以保护数据库安全。
	导出超时时间 (s)	设定安全策略，当达到设定的时间后，SQL窗口执行的导出语句会中断，以保护数据库安全。

6. 完成以上信息填写，在基本信息标签页中，单击左下角的测试连接，等待测试通过。

 **说明** 如果测试连接失败，请按照报错提示检查您录入的实例信息。

7. 连接成功，单击提交。

添加DLA实例

1. 登录[DMS控制台](#)。
2. 单击页面左上角的新增实例 / 批量录入，在弹框中单击新增实例按钮。
□
3. 单击云数据库标签页。
4. 在云数据库区域中选择DLA-数据湖分析。
□
5. 后续步骤与添加RDS MySQL实例一致，详情可参见[添加RDS实例](#)。

任务开发

1. 登录[DMS控制台](#)。
2. 在顶部菜单栏中，单击数据工厂 > 任务编排。
3. 在任务编排页面的自由编排任务区域，单击新建任务流。
□
4. 在新建任务流对话框，自定义填写任务流名称和描述后，单击确认。
□
5. 将左侧任务类型中DLA一键建仓拖拽到页面中空白区域。
6. 配置任务节点。
 - i. 单击页面中新建的任务节点，在右侧属性面板中单击内容设置页签，进入到任务节点的SQL编辑页面。

ii. 在内容设置页中配置以下参数，并单击保存。

配置	说明
数据来源	在已接入DMS的实例中选择需要入DLA数据湖的数据库源，本案例中为RDS MySQL数据库。
Schema名称	填入需要在DLA中创建的Schema名称。
OSS数据位置	填入目标OSS数据位置。
高级选项	通过一键建仓向导创建Schema时，您可以在高级选项中自定义一些设置，例如过滤字段或者表等，详细请参见 高级选项功能 。

 **说明** 更多关于DLA一键建仓的介绍请参见[一键建仓](#)。

iii. 单击变量设置页签，设置变量并单击保存。

 **说明** 更多DMS支持的时间格式请参见[任务编排介绍](#)。

7. 将左侧任务类型中DLA-SQL拖拽到页面中空白区域。

8. 配置任务节点。

- i. 单击页面中的DLA-SQL任务节点，在右侧属性面板中单击节点名称，在任务名文本框输入 `DLA-SQL-day`，单击保存。
- ii. 单击右侧属性面板中的内容设置页签，在该页面中编辑报表生成语句。
- iii. 单击变量设置，参考步骤6设置变量并单击保存。

 **说明** 完成SQL、变量配置之后，单击SQL预览确认是否满足预期。确认无误后，单击属性面板右下角保存。

9. 返回到DAG图中，拖动鼠标连接2个任务节点。

10. 完成DAG图的流程走向编辑后，在右侧属性面板中开启调度，完成配置后单击保存即可。

 **说明** 任务流调度周期支持每月、每周、每天多维度配置，您可以按需求选择调度周期，本案例为每个工作日的早上5点执行任务。

后续步骤

- DLA一键建仓生成数据库后，若子账号没有查询权限，需要先申请权限，详情请参见[权限管理](#)。
- 您可以通过DMS提供的数据保护伞功能进行数据安全管控，详情请参见[数据保护伞](#)。