阿里云 云数据库 POLARDB

数据迁移/同步

文档版本: 20200103

为了无法计算的价值 | [] 阿里云

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。

- 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。
- 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. 阿里云文档中所有内容,包括但不限于图片、架构设计、页面布局、文字描述,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
0	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。
!	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。
Ê	用于补充说明、最佳实践、窍门 等,不是用户必须了解的内容。	送 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。	执行cd /d C:/window命令,进 入Windows系统文件夹。
##	表示参数、变量。	bae log listinstanceid
		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。	<pre>switch {active stand}</pre>

目录

法律声明	I
通用约定	I
1 POLARDR MySOL	1
11 数据迁移/同步方案概监	1
1.2 数据迁移	
1.2.1 从POLARDB for MySQL迁移至POLARDB for MySQL	2
1.2.2 从ECS上的自建MySQL迁移至POLARDB for MySQL	8
1.2.3 从RDS MySQL迁移至POLARDB MySQL	12
1.2.4 从Amazon Aurora for MySQL迁移至POLARDB for MySQL	17
1.2.5 从本地MySQL迁移至POLARDB for MySQL	22
1.2.6 从POLARDB for MySQL迁移至RDS for MySQL	29
1.2.7 一键升级RDS for MySQL到POLARDB for MySQL	
1.2.8 一键克隆RDS for MySQL到POLARDB for MySQL	
1.3 数据同步	
1.3.1 从RDS MySQL同步至POLARDB MySQL	44
1.3.2 从POLARDB MySQL同步至RDS MySQL	50
1.3.3 从POLARDB MySQL同步至POLARDB MySQL	55
1.3.4 从POLARDB MySQL同步至AnalyticDB for MySQL	61
1.3.5 从POLARDB MySQL同步至AnalyticDB for PostgreSQL	68
1.3.6 从POLARDB MySQL同步至Datahub	75
1.3.7 从POLARDB MySQL同步到Kafka	
1.3.8 从ECS上的自建MySQL同步至POLARDB MySQL	
2 POLARDB PostgreSQL	95
2.1 数据迁移方案概览	
2.2 从自建PostgreSQL迁移至POLARDB for PostgreSQL	
2.3 从RDS for PostgreSQL迁移至POLARDB for PostgreSQL	
3 POLARDB兼容Oracle语法引擎	103
3.1 数据迁移方案概览	103
3.2 从自建Oracle迁移至POLARDB	103

1 POLARDB MySQL

1.1 数据迁移/同步方案概览

云数据库POLARDB提供了多种数据迁移同步方案,可满足不同上云、迁云、同步的业务需求,使 您可以在不影响业务的情况下平滑将数据库迁移、同步至阿里云云数据库POLARDB上面。

通过使用阿里云数据传输服务 (*DTS*) , 您可以实现POLARDB的结构迁移、全量迁移和实时同步。

数据迁移	
------	--

使用场景	文档链接
从RDS迁移至POLARDB	 一键升级RDS for MySQL到POLARDB for MySQL(平滑迁 移,推荐) 一键克隆RDS for MySQL到POLARDB for MySQL 从RDS MySQL迁移至POLARDB MySQL
从POLARDB迁移至RDS	#unique_8
从自建数据库迁移至 POLARDB	 #unique_9 #unique_10
从第三方云数据库迁移至 POLARDB	从Amazon Aurora for MySQL <mark>迁移至</mark> POLARDB for MySQL
POLARDB之间的数据迁移	从POLARDB for MySQL迁移至POLARDB for MySQL

数据同步

使用场景	文档链接
从RDS同步至POLARDB	从RDS MySQL同步至POLARDB MySQL
从POLARDB同步至 POLARDB	从POLARDB MySQL同步至POLARDB MySQL
从自建数据库同步至 POALRDB	从ECS上的自建MySQL同步至POLARDB MySQL
从POLARDB同步至RDS	从POLARDB MySQL同步至RDS MySQL
从POLARDB同步至分析型数 据库	 ・ 从POLARDB MySQL同步至AnalyticDB for MySQL ・ 从POLARDB MySQL同步至AnalyticDB for PostgreSQL

使用场景	文档链接
从POLARDB同步至Datahub	从POLARDB MySQL同步至Datahub
从POLARDB同步至Kafka	从POLARDB MySQL同步到Kafka

1.2 数据迁移

1.2.1 从POLARDB for MySQL迁移至POLARDB for MySQL

本文介绍如何使用数据传输服务(DTS)将一个POLARDB集群的数据迁移到另一个POLARDB集群。

前提条件

- · 已购买源和目标POLARDB for MySQL集群,详情请参见创建POLARDB for MySQL集群。
- ・ 为満足増量数据迁移的要求,源POLARDB for MySQL集群需要开启Binlog,详情请参见如何开启Binlog。
- 注意事项
 - ・如果源数据库没有主键或唯一约束,且所有字段没有唯一性,可能会导致目标数据库中出现重复数据。
 - ・ 对于迁移失败的任务,DTS会触发自动恢复。当您需要将业务切换至目标集群,请务必先停止或
 释放该任务,避免该任务被自动恢复后,导致源端数据覆盖目标集群的数据。

迁移类型介绍

支持结构迁移、全量数据迁移和增量数据迁移,详细介绍请参见#unique_22。



同时使用这三种迁移类型可实现在应用不停服的情况下,平滑地完成数据库迁移。

费用说明

迁移类型	链路配置费用	公网流量费用
结构迁移/全量数据迁 移	不收费。	通过公网将数据迁移出阿里云时将收费,详情请 参见#unique_23。
增量数据迁移	收费,详情请参 见#unique_23。	

增量数据迁移阶段支持同步的SQL操作

- INSERT, UPDATE, DELETE, REPLACE
- ALTER TABLE, ALTER VIEW, ALTER FUNCTION, ALTER PROCEDURE
- CREATE DATABASE、 CREATE SCHEMA、 CREATE INDEX、 CREATE TABLE、 CREATE PROCEDURE、 CREATE FUNCTION、 CREATE TRIGGER、 CREATE VIEW、 CREATE EVENT
- DROP FUNCTION, DROP EVENT, DROP INDEX, DROP PROCEDURE, DROP TABLE, DROP TRIGGER, DROP VIEW
- RENAME TABLE、TRUNCATE TABLE

数据库账号的权限要求

数据库	权限要求
源POLARDB for MySQL	待迁移对象的读权限
目标POLARDB for MySQL	迁移对象的读写权限



关于数据库账号的创建和授权方法,请参见创建POLARDB数据库账号。

操作步骤

- 1. 登录数据传输控制台。
- 2. 在左侧导航栏,单击数据迁移。
- 3. 在迁移任务列表页面顶部,选择迁移的目标集群所属地域。

数据传输	迁移任务列表	华东1(杭州)	华东2(上海)	华北1 (青岛)	华北2(北	凉) 华南1(深	訓)	华北3 (张	家口)	香港	美国(硅谷)	美国 (弗吉尼亚) 新加坡
487105		阿联酋 (迪拜)	德国 (法兰克福]) 马来西亚(吉隆坡)	澳大利亚 (悉尼)	印度	建(孟买)	英国 (伦敦)	日本 (东京)	印度尼西亚(新	珈达)
備范		华北5 (呼和浩特)										
数据迁移											()	DTS常见问题	€刷新
数据订阅											•		

4. 单击页面右上角的创建迁移任务。

5. 配置迁移任务的源库和目标库连接信息。

1.源库及目标库	<u> </u>	2.迁移类型及列表	>		3.映射名称修改	\rangle	4.预检查
* 任务名称: [POLARDB实例间迁移						
源库信息							
* 实例类型:	POLARDB		Ŧ				
* 实例地区:	<u> </u>		v				
* POLARDB空例ID:	nc-hn1w						
* 粉墀庑胀具。	detect						
	utstest			100-01-1-+			
\$K36/FG1/3.	*******		P	测试进安	C MINUELL		
日仁定法自							
日你年间总							
* 实例类型:	POLARDB		•				
* 实例地区:	华东1 (杭州)		•				
* POLARDB实例ID:	pc-bp1l		•				
* 数据库账号:	dtstest						
* 数据库密码:	•••••	٩	Þ	测试连接	⊘ 测试通过		
						取消 上云评估	授权白名单并进入下一步

配置项目	配置选项	配置说明
任务名称	-	DTS会自动生成一个任务名称,建议配置具有业务意义的名称(无唯一性要求),便于后续识别。
源库信息	实例类型	选择POLARDB。
	实例地区	选择源POLARDB集群所属的地域。
	POLARDB实 例ID	选择源POLARDB集群ID。
	数据库账号	填入连接源POLARDB集群的数据库账号。
	数据库密码	填入该账号对应的密码。
		 说明: 源库信息填写完毕后,您可以单击数据库密码后的测试连接来验证填入的信息是否正确。如果填写正确则提示测试通过;如果提示测试失败,单击测试失败后的诊断,根据提示调整填写的源库信息。
目标库信息	实例类型	选择POLARDB。
	实例地区	选择目标POLARDB集群所属的地域。
	POLARDB实 例ID	选择目标POLARDB集群ID。
	数据库账号	填入连接目标POLARDB集群的数据库账号。

配置项目	配置选项	配置说明
	数据库密码	填入该账号对应的密码。
		 说明: 目标库信息填写完毕后,您可以单击数据库密码后的测试连接来验证填入的信息是否正确。如果填写正确则提示测试通过;如果提示测试失败,单击测试失败后的诊断,根据提示调整填写的目标库信息。

6. 完成上述配置后,单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到源和目标POLARDB集群的白名单中,用于保 障DTS服务器能够正常连接源和目标集群。

7. 选择迁移类型和迁移对象。

1.源库及目标库	迁移类型: 🗹 结构迁移	2.迁移类型及列表	量数据迁移	3.映射名称修改	>	4.预检查
	迁移对象 若全局搜索,请先展开树 ● dtstestdata ● Tables ● Views		> <	已选择对象 (急标修到对象行,点击) 情点我 ■ dtstestdata (2个对象) ■ customer ■ order	異個可修改对象名或过違条件) Q) 详
*	全选中 映射名称更改: 0 7	へ进行库表名称批量更改 □)要进行库表名称批量更	全移除 改		
当 1 2	<mark>t意:</mark> 数据迁移只会将源库的数据 数据迁移过程中,不支持DDI	(结构)复制一份到目标数据库 握作,如进行DDL撮作可能导致	并不会对源数据库数据(改迁移失败	结构)造成影响。	取消上一歩	保存 预检查并启动
配置	说明					

配置	说明
迁移类型	・如果只需要进行全量迁移,则同时勾选结构迁移和全量数据迁移。
	间 说明: 为保障数据一致性,全量数据迁移期间请勿在自建MySQL数据库中写入新的 数据。
	 ・如果需要进行不停机迁移,则同时勾选结构迁移、全量数据迁移和增量数据迁 移。
迁移对象	在迁移对象框中单击待迁移的对象,然后单击 ,将其移动至已选择对象框。
	道 说明:
	 · 迁移对象选择的粒度可以为库、表、列三个粒度。 · 默认情况下,迁移完成后,迁移对象的名称保持不变。如果您需要迁移对象在 目标集群中名称不同,那么需要使用DTS提供的对象名映射功能,详情请参 见#unique_24。
	・如果使用了对象名映射功能,可能会导致依赖这个对象的其他对象迁移失败。

8. 单击页面右下角的预检查并启动。

〕 说明:

- · 在迁移任务正式启动之前, 会先进行预检查。只有通过预检查, DTS才能迁移数据。
- ·如果预检查失败,单击具体检查项后的____,查看失败详情。根据提示修复后,重新进行预

检查。

9. 预检查通过后,单击下一步。

10.在弹出的购买配置确认对话框,选择链路规格并勾选数据传输(按量付费)服务条款。

11.单击购买并启动,迁移任务正式开始。

・结构迁移+全量数据迁移

请勿手动结束迁移任务,否则可能会导致数据不完整。您只需等待迁移任务完成即可,迁移 任务会自动结束。

・结构迁移+全量数据迁移+增量数据迁移

迁移任务不会自动结束,您需要手动结束迁移任务。

说明:

请选择合适的时间手动结束迁移任务,例如业务低峰期或准备将业务切换至目标集群时。

- a. 观察迁移任务的进度变更为增量迁移,并显示为无延迟状态时,将源库停写几分钟,此时增量迁移的状态可能会显示延迟的时间。
- b. 等待迁移任务的增量迁移再次进入无延迟状态后,手动结束迁移任务。

概览	迁移任务名 ▼ 请输入迁移任务名进行搜索	援索 排序: 默认排序 ▼ 状态: 全部 ▼
数据迁移		
数据订阅	Ⅱ ID/名称:	状态: 迁移中 暫停住务 重着详情 创建类似住务 升级
数据同步	2018-12-04 16:13:56 创建 全量迁移 100%(已迁移4000000行)	完成 當量迁移 无延迟 1
文件导入导出		
操作日志	□ 启动 暫停 结束 3	共有26条,毎页显示:20条 « < 1 2 > »

12将业务切换至POLARDB集群。

后续步骤

由于用于数据迁移的数据库账号拥有读写权限,为保障数据库安全性,请在数据迁移完成后,删除 相关的数据库账号。

1.2.2 从ECS上的自建MySQL迁移至POLARDB for MySQL

使用阿里云<mark>数据传输服务 (DTS)</mark>,您可以将ECS上的自建MySQL数据库迁移至POLARDB for MySQL实例。使用增量数据迁移,您可以实现源数据库不停机的情况下,平滑地将数据迁移到目 标POLARDB for MySQL实例中。

迁移权限要求

迁移任务配置时,需要提供源ECS上的自建MySQL数据库及目标POLARDB实例的迁移账号。迁 移账号的相关权限详见下表。

📋 说明:

如您尚未创建迁移账号,您可以参考*MySQL*账号创建手册、创建*POLARDB*数据库账号创建迁移账 号并配置所需权限。

迁移数据源	结构迁移	全量迁移	增量迁移
ECS上的自建 MySQL数据库	迁移对象的 SELECT权限	迁移对象的 SELECT权限	迁移对象的SELECT、 REPLICATION CLIENT、 REPLICATION SLAVE权限
目标POLARDB实 例	迁移对象的ALL权 限	迁移对象的ALL权 限	迁移对象的ALL权限

配置迁移任务

- 1. 登录DTS控制台。
- 2. 在左侧导航栏,单击数据迁移。
- 3. 单击数据迁移页面右侧的创建迁移任务。

4. 配置迁移任务的源库及目标库信息。

1.源库及目		2.迁移关型及列表	\rangle	3.预检查
* 174 Att.				
11544:	33版注档FPOLARDB			
源库信息				
* 实例类型:	ECS上的自建数据库			
* 实例地区 :	华东1(杭州)			
* ECS实例ID :	-			
* 数据库类型:	MySQL v			
* 端口:	3306			
* 数据库账号:				
* 数据库密码:	¢	测试连接		
* 连接方式:	● 非加密连接 ● SSL安全连接			
目标库信息				
* 实例类型:	POLARDB			
* 实例地区:	华东1(杭州) 🔻			
* POLARDB实例ID:				
* 数据库账号:	(mailed)			
* 数据库密码:	4>	测试连接		
			取当	トテ運信 授权白名单并进入下一步

配置项目	配置选项	配置说明
任务名称	-	 DTS为每个任务自动生成一个任务名称,任务名称没有唯一性要求。 您可以根据需要修改任务名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。
源库信息	实例类型	选择ECS上的自建数据库。
	实例地区	选择源ECS所属的地域
	ECS实例ID	选择源ECS实例ID。
	数据库类型	选择MySQL
	端口	填入源ECS上的MySQL数据库对外提供服务的端口,默认为 3306。
	数据库账号	填入连接源MySQL数据库的账号。
	数据库密码	填入连接源MySQL数据库账号对应的密码。
	连接方式	根据需求选择非加密连接或SSL安全连接,本案例选择为非加 密连接。
		道 说明: 选择 SSL安全连接时,需要提前准备CA根证书并上传该证 书。

配置项目	配置选项	配置说明
目标实例信息	实例类型	选择POLARDB。
	实例地区	选择目标POLARDB实例所属地域。
	POLARDB 实例ID	选择目标POLARDB实例ID。
	数据库账号	填入连接目标POLARDB数据库的账号。
	数据库密码	填入连接目标POLARDB数据库账号对应的密码。

5. 配置完成后,单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到目标POLARDB实例的白名单中,用于保 障DTS服务器能够正常连接POLARDB实例。迁移完成后如不再需要可手动删除,详情请参 考#unique_27。

6. 选择迁移对象和迁移类型。

创建迁移任务 🔹 返回上级

式移对象 •mydb •yanmen_target825	已选择国际移到对家行,点击编辑可修改目标库对家名或配重过源条件)
🖶 📴 mydb 🗄 🦢 yanmen_target825	amptest1
	3
	>
	< compared with the second sec

配置项目	配置说明
迁移类型	・如果只需要进行全量迁移,那么迁移类型选择结构迁移和全量数据迁移。
	 说明: 为保障数据一致性,全量数据迁移期间请勿在源数据库中写入新的数据。 如果需要进行不停机迁移,那么迁移类型选择结构迁移、全量数据迁移和增量数据迁移。
迁移对象	 在迁移对象框中将想要迁移的数据库选中,单击 移动到已选择对 象框。 迁移对象的选择粒度细化为:库、表、列三个粒度。 默认情况下,对象迁移到目标POLARDB实例后,对象名跟源ECS上的自 建数据库一致。如果您迁移的对象在源实例跟目标实例上名称不同,那么需 要使用DTS提供的对象名映射功能,详细使用方式可以参考库表列映射。

7. 上述配置完成后,单击页面右下角的预检查并启动。



・在迁移任务正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动迁移任务。

·如果预检查失败,单击具体检查项后的

,查看具体的失败详情。根据失败原因修复

后,重新进行预检查。

8. 预检查通过后,单击下一步。

- 在购买配置确认页面,选择链路规格并勾选数据传输(按量付费)服务条款,单击购买并启动,迁移任务正式开始。
 - ・ 全量数据迁移

等待迁移任务完成即可,迁移任务会自动停止。

・増量数据迁移

迁移任务不会自动结束,观察迁移任务的状态显示为增量迁移无延迟的状态时,将源库停写 几分钟,等待增量迁移再次进入增量迁移无延迟状态,手动停止迁移任务。

迁移完成后根据业务需求选择合适的时间,将业务切换至目标POLARDB实例中。

1.2.3 从RDS MySQL迁移至POLARDB MySQL

本文介绍使用DTS将RDS MySQL迁移至POLARDB MySQL的任务配置流程。

背景信息

使用阿里云<mark>数据传输服务₍DTS)</sub>,您可以实现RDS MySQL到POLARDB MySQL的数据迁移。</mark> 通过DTS增量迁移的存储引擎,可以在源RDS实例不停机的情况下,将数据迁移到目标POLARDB MySQL集群。

☴ 说明:

阿里云已支持从RDS MySQL一键迁移到POLARDB MySQL,详情请参见一键升级RDS for MySQL到POLARDB for MySQL。

下面详细介绍如何使用DTS实现RDS MySQL实例与POLARDB集群间的数据迁移。

迁移权限要求

当使用DTS进行RDS实例与POLARDB集群间的数据迁移时,源RDS实例的账号需拥有读写权限、目的POLARDB集群的账号需拥有迁移对象的ALL权限。

创建迁移账号

迁移任务配置时,需要提供源RDS实例及目的POLARDB集群的迁移账号。迁移账号的相 关权限详见上面的 迁移权限要求一节。如果尚未创建迁移账号,您可以参考 RDS实例账号创 建和POLARDB集群账号创建。先在源实例及目的集群中创建迁移账号,并将要迁移的库表的读写权 限授权给上面创建的账号。

配置迁移任务

当上面的所有前置条件都配置完成后,可以开始正式的数据迁移。下面详细介绍迁移任务配置流 程。

- 1. 登录DTS控制台。
- 2. 在页面左侧,选择数据迁移。
- 3. 单击右上角创建迁移任务。

数据传输	迁移任务列表 华东1(记州) 华东2(上海) 华北1(育岛) 华北2(北京) 华南1(深圳) 华	21は3(3%を口) 骨相 美国(桂谷) 美国(弗吉尼亚) 新加坡 阿联酋(迪年) 20月時 文件号	入创建迁移任务
概范	德国(法兰克福) 马米西亚(吉隆坡) 澳大利亚(悉尼) 印度(孟买) 日本(5	に死) 印度に四亚(釉加広) 終は5(呼和治特)	
数据迁移	迁移任务名 ▼	▼ 状态: 全部 ▼	
数据订阅			
数据同步	□ ID/名称: / dtsd9tcbni6 状态: 未启动	修改	任务配置 创建类似任务
· 文件导入导出	2018-08-18 12:48:04 创建		完成
Sarayeash			

4. (可选) 填写任务名称。

DTS 为每个任务自动生成一个任务名称,任务名称没有唯一性要求。您可以根据需要修改任务 名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。

5. 填写源实例信息。

创建迁移任务 1200上级

	1.源库及目标库	2.迁移	5类型及列表	> 3.预检查
	*任务名称: h	angzhou-small		
源库信息				
	* 实例类型:	RDS实例	\$	
	*实例地区:	9.81 (854)	\$	
	* RDS实例ID:	m-m-mail-mail-	•	其他阿里云账号下的RDS实例
	*数据库账号:	jærghrinst		
	* 数据库密码:		4>	测试连接
	* 连接方式:	●非加密连接 ○SSL安全连接		

- · 实例类型:选择RDS实例。
- · 实例地区:选择RDS实例所在的地域。
- · RDS实例ID:选择配置迁移的源RDS实例的实例ID。
- ·数据库账号:连接RDS实例的账号。
- ・数据库密码:以上数据库账号的密码。
- 6. 单击测试连接,确认DTS可以连接到源RDS实例。

7. 填写目标POLARDB集群信息。

目标库信息						
•实例类型:	POLARDB \$					
• 实例地区:	1833 (896) ÷					
◆ POLARDB实例ID:	pc-in-finite and v					
*数据库账号:	langi.ted					
*数据库密码:	هې	测试连接				
			Ę	仅消	上云评估	授权白名单并进入下一步

- · 实例类型:选择POLARDB。
- · 实例地区:选择POLARDB集群所在的地域。
- · POLARDB实例ID:选择配置迁移的目标POLARDB集群的集群ID。
- ·数据库账号:连接POLARDB集群的账号。
- ・数据库密码:以上数据库账号的密码。
- 8. 单击测试连接,确认DTS可以连接到目标POLARDB集群。
- 9. 单击右下角的授权白名单并进入下一步。这个步骤DTS会将DTS服务器的IP地址添加到 源RDS实例及目标POLARDB的白名单中,避免因为RDS实例及POLARDB设置了白名 单,DTS服务器连接不上实例导致迁移失败。

10选择迁移类型和迁移对象。

1.源库及目标库		2.迁移类型及列表		3.预检查
	★ 迁移类型: ☑ 结构迁移 ☑ 全量数调迁移 □ 增量数调迁移 全量数据迁移过程中,如果源库有数据更新,这部分更新数据不保证被迁移 为保证迁移数据的一致性,建议选择结构迁移+全量数据迁移+增量数据迁移	到目标实例 多		
	迁移对象		已选择刘象(黑标移到刘象行,点击编辑可修改刘象名或立遗条件)详情点我	
		`		
		>		
	全选中		全移除	

- ・迁移类型:
 - 结构迁移

DTS 会将迁移对象的结构定义迁移到目标集群。目前 DTS 支持结构迁移的对象包括:表。其他对象如视图、同义词、触发器、存储过程、存储函数、包、自定义类型等暂 不支持。

- 全量数据迁移

DTS 会将源数据库迁移对象的存量数据全部迁移到目标集群。

- 増量迁移

DTS 会将迁移过程中源实例的数据变更同步到目标集群。如果迁移期间进行了 DDL 操 作,那么这些结构变更不会迁移到目标集群。

如果只需要进行全量迁移,那么迁移类型选择:结构迁移+全量数据迁移。

如果需要进行不停机迁移,那么迁移类型选择:结构迁移+全量数据迁移+增量数据迁移。

送 说明:

结构迁移和全量数据迁移均不收费,增量迁移需要收费。

·迁移对象:选择要迁移的对象,单击向右的箭头,将选中的对象添加到右侧。

迁移对象的选择粒度细化为:库、表、列三个粒度。默认情况下,对象迁移到目标POLARDB集群后,对象名跟源RDS实例一致。如果您迁移的对象在源实例跟目标集群上名称不同,那么需要使用DTS提供的对象名映射功能,详细使用方式可以参考<u>库表列映射</u>。

📃 说明:

- 暂时不支持对系统表的迁移。
- 目标集群中不能有和迁移对象同名的对象。将鼠标移至右侧框中的对象,单击编辑,即可 修改迁移后的对象名。

11.单击右下角的预检查并启动,成功后单击下一步。

如果预检查失败,可以单击具体检查项后的检查结果,查看具体的失败详情,并根据失败原因修 复后,重新进行预检查。

管理控制台	产品与服务	§ -		预检查			X 备家	案 帮助与文档 ▼	amp****@aliyun-inner.com ▼
数据传输		迁和	8任务列表		预检	₫查失败91%		こ別新	文件导入 创建迁移任务
数据迁移		迁移	任务名 🛊	检测项 日印//中K/K112旦	检测内容 12旦日19355年1935万代校正古 <i>两</i> 比让19支不	检测结果			
数据订阅				同名对象存在性检查	检查目的库是否存在跟待迁移对象同名的结构对象	(失败))			
数据同步			ID/名称: /	源端同名对象存在性检查	检查源端不同库是否存在同名的结构对象	成功			
产品文档			2016-08-05	约束完整性检查	检查迁移表依赖的外键父表是否迁移	成功			
	Ę		101534	存储引擎检查	检查迁移表是否有不支持的存储引擎	成功			可爱洋教红条
			ID/名称: c	连接数检查	检查源库、目标库的连接数是否满足要求	成功			HUJELLIY IL 75
			ID/名称: c	SQL_MODE 检查	检查源库和目标库中SQL_MODE是否合法	成功			查看迁移详情 升级
			2016-07-27						
			Laward			取消	i		
			ID/名称: d	itssoko4mx9tt2e / dts8jxyxv6c	状态: 过移失败	,立即修复			查看迁移详情 升级
					共有10433	条, 每页显示: 20务		< 1 2 3	3 39 GO

12确认DTS购买信息,阅读并勾选服务条款,单击立即购买并启动。

当预检查通过后,可以启动迁移任务,任务启动成功后,可以在任务列表中查看迁移的具体状态 及迁移进度。

如果选择了增量迁移,那么进入增量迁移阶段后,源库的更新写入都会被DTS同步到目标 POLARDB集群。迁移任务不会自动结束。如果用户只是为了迁移,那么建议在增量迁移无延迟 的状态时,源库停写几分钟,等待增量迁移再次进入无延迟状态后,停止掉迁移任务,直接将业 务切换到目标POLARDB集群上即可。 13.单击目标地域,查看迁移状态。迁移完成时,状态为已完成。

至此,完成RDS实例到POLARDB数据迁移任务配置。

1.2.4 从Amazon Aurora for MySQL迁移至POLARDB for MySQL

背景信息

本文介绍使用阿里云数据传输服务 (*DTS*) ,将 Amazon Aurora for MySQL数据库迁移到阿里 云POLARDB for MySQL的步骤及注意事项。

前提条件

- ·迁移的源数据库实例支持公网连接。
- · 已创建支持MySQL 5.6 版本的Amazon Aurora 实例。
- ・已经创建阿里云POLARDB集群。
- ・已经创建拥有读写权限的账号。

迁移限制

- ・ 迁移仅支持 MySQL 5.6 的Amazon Aurora。
- ·结构迁移不支持 event 的迁移。
- · 对于MySQL的浮点型float/double, DTS通过round(column, precision)来读取该列的 值,若列类型没有明确定义其精度,对于float,精度为38位,对于double类型,精度为308
 、请先确认DTS的迁移精度是否符合业务预期。
- ·如果使用了对象名映射功能,依赖这个对象的其他对象可能迁移失败。
- · 当选择增量迁移时,源端的 MySQL 实例需要开启 binlog。
- · 当选择增量迁移时, 源库的 binlog_format 要为 row。
- · 当选择增量迁移且源 MySQL 为 5.6 版本时,它的 binlog_row_image 必须为 full。
- ・当选择增量迁移时,增量迁移过程中如果源MySQL实例出现因实例跨机迁移或跨机重建等导致
 的binlog 文件ID乱序,可能导致增量迁移数据丢失。

注意事项

- ・如果源数据库没有主键或唯一约束,且所有字段没有唯一性,可能会导致目标数据库中出现重复数据。
- · 对于七天之内的异常任务,DTS会尝试自动恢复,可能会导致迁移任务的源端数据库数据覆盖 目标实例数据库中写入的业务数据,迁移任务结束后务必将DTS访问目标实例账号的写权限用 revoke命令回收掉。

操作步骤

1. 登录Amazon Aurora数据库实例,单击数据库名称,在连接页面查看终端节点和端口。

RDS 〉 数据库 〉 cluster1-1 〉 t1 t1						(8)	文 操作	•
相关 Q、 /#:::: 此前中								0
- 数据库名称 ▲	角色 ▼ 引擎 ▼	区域和可用区 🔻	大小 ▼	状态 ▼	CPU	当前活动	维护 ▼	VP
C E cluster1-1	区域性的 Aurora MySQL	us-east-2	1 实例	❷ 可用				-
• t1	写入器 Aurora MySQL	us-east-2c	db.t2.small	⊘ 可用	7.33%	2选择/秒		vpc
<mark>连接</mark> 盖控 日志和事件 配置 标签								
连接								
终端节点和端口	联网			安全的	生			
终端节点	可用区			VPC 安	全组			
t1. rds.amazonaws.com	us-east-2c			defaul (活跃)	t (sg-35839559))			
第二	VPC							

- 2. 登录DTS控制台。
- 3. 在左侧菜单栏单击数据迁移,单击右上角创建迁移任务。
- 4. (可选)填写任务名称。

DTS 为每个任务自动生成一个任务名称,任务名称没有唯一性要求。您可以根据需要修改任务 名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。

5. 填写源库和目标库信息,具体参数配置说明如下:

库类别	参数	说明
源库信息	实例类型	源库实例类型,这里选择有公网IP的自建数据库。
	实例地区	这里选择源库所在的地区,如果您的实例进行了访问限 制,请先放开对应地区公网IP段的访问权限后,再配置数据 迁移任务。
		道 说明: 可以单击右侧获取DTS IP段查看、复制对应地区的IP段。
	数据库类型	源数据库类型,这里选择MySQL。
	主机名或IP地址	Amazon Aurora数据库的终端节点。
	端口	Amazon Aurora数据库的端口。
	数据库账号	Amazon Aurora数据库具有读写权限的账号。
	数据库密码	Amazon Aurora数据库账号对应的密码。

库类别	参数	说明
目标库信息	实例类型	这里选择POLARDB。
	实例地区	目标实例所在的地区。
	POLARDB实例ID	对应地区下的实例ID,这里选择要迁移到的目标实例的ID。
数据库账号 目标实例的拥有读写权限的账号		目标实例的拥有读写权限的账号。
	数据库密码	目标实例的对应账号的密码。

* 任务名称: d		
源库信息		
* 实例类型:	有公网IP的自建数据库 ▼	
* 实例地区:	美国(硅谷)	获取DTS IP段
* 数据库美型:	MySQL •	
* 主机名或IP地址:	.rds.amazonaws.com	
* 流口:	3306	
* 数据库账号:		
* 数据库密码:	••••••	测试连接 🛇 测试通过
目标库信息		
* 实例类型:	POLARDB	
* 实例地区:	华东1(杭州)	
* POLARDB实例ID:	pc-	
* 数据库账号:		
* 数据库密码:	••••••• Ø	测试连接 📀 测试通过

- 6. 填写完成后单击测试连接,确定源库和目标库都测试通过。
- 7. 单击右下角授权白名单并进入下一步。
- 8. 勾选对应的迁移类型, 在迁移对象框中将要迁移的数据库选中, 单击

	移动到已选择对
/	

象框。

说明:为保证迁移数据的一致性,建议选择结构迁移+全量数据迁移+增量数据迁移。

结构迁移和全量数据迁移暂不收费,增量数据迁移根据链路规格按小时收费。

1 创建迁移任务 * 返回上级				
1.源库及目标库		2.迁移类型及列	表	3.预检查
	■ 过移完型: ○ 總和這移 ○ 金融数编述移 □ 滑稽数编述移 全品放振计校订程中、如果源与有效重要。 法部分要称素不保证统计经 为保证计经数据的一边性。 建议透明和记述书: 全面数新正书 常面数描述	·到目标实例 多		
	迁鄉对象		已透掉对象(就标移到对象行,点击编辑可修改对象名或过续条件)详情点我	
	e 🔤 sys		📔 db1	
		>		
		<		
	全选中		全核除	
	注意: 1. 数据迁移日会将遗库的数据(结构)复制一份到目标数据库,并不会对遗数或 2. 数据迁移过程中,不支持DDL操作,如进行DDL操作可能导致迁移失败	黑库数据(结构) 近	立位影响 。	
				取消 上一步 保存 预检查并启动

9. 单击预检查并启动,等待预检查结束。



如果预检查失败,可以根据错误项的提示进行修复,然后重新启动任务。

预检查		×
		预 <u>检查通过</u> 100%
检测项	检测内容	检测结果
源库连接性检查	检查数据传输服务器是否能连通源数据库	成功
源库权限检查	检查源数据库的账号权限是否满足迁移要求	成功
目的库连接性检查	检查数据传输服务器是否能连通目的数据库	成功
目的库连接性检查	检查数据传输服务器是否能连通目的数据库	成功
目的库权限检查	检查目的数据库的账号权限是否满足迁移要求	成功
存储引擎检查	检查迁移表是否有不支持的存储引擎	成功
		下一步

10.单击下一步,在购买配置确认对话框中,勾选《数据传输(按量付费)服务条款》并单击立即购 买并启动。

如果选择了增量迁移,那么进入增量迁移阶段后,源库的更新写入都会被DTS同步到目标 POLARDB集群。迁移任务不会自动结束。如果您只是为了迁移,那么建议在增量迁移无延迟的 状态时,源库停写几分钟,等待增量迁移再次进入无延迟状态后,停止掉迁移任务,直接将业务 切换到目标POLARDB集群上即可。

11.单击目标地域,查看迁移状态。迁移完成时,状态为已完成。

ID/名称: / dtsvzev3df8 🖌	秋态:已完成		查看详情	创建类似任务
2019-01-15 13:58:52 创建			2019-01-15	14:10:03 完成
结构迁移 100%		全量迁移 100%(已迁移0行)		

至此,完成 Amazon Aurora 数据库到阿里云 POLARDB 数据迁移任务配置。

1.2.5 从本地MySQL迁移至POLARDB for MySQL

使用阿里云数据传输服务_(DTS),您可以实现本地MySQL到POLARDB for MySQL集群的数 据迁移。通过DTS增量迁移的存储引擎,可以实现在本地应用不停服的情况下,将数据迁移到目 标POLARDB for MySQL集群。

本文介绍使用DTS进行本地MySQL迁移至POLARDB for MySQL的任务配置流程。

支持同步的SQL操作

对于本地 MySQL > POLARDB 数据迁移,DTS支持同步的SQL操作如下:

INSERT, UPDATE, DELETE, REPLACE

ALTER TABLE, ALTER VIEW, ALTER FUNCTION, ALTER PROCEDURE

CREATE DATABASE, CREATE SCHEMA, CREATE INDEX, CREATE TABLE, CREATE PROCEDURE, CREATE

```
FUNCTION, CREATE TRIGGER, CREATE VIEW, CREATE EVENT
```

DROP FUNCTION、 DROP EVENT、 DROP INDEX、 DROP PROCEDURE、 DROP TABLE 、 DROP TRIGGER、 DROP

VIEW

RENAME TABLE, TRUNCATE TABLE

前提条件

- ・已经创建阿里云POLARDB集群。
- · 已经创建拥有读写权限的账号。
- 已经开通本地MySQL的远程访问权限。开通命令为grant all privileges on *.* to < username>@'<ipaddress>' identified by "<password>";

📕 说明:

- <username>: 本地MySQL数据库的用户名;
- <ipaddress>: 允许登录数据库的ip地址, localhost 表示只能本地登录数据库, %表示任何ip地址都能登录数据库;
- <password>:数据库的用户名对应的密码。

注意事项

・执行迁移任务前建议提前做好数据备份。

・ 对于七天之内的异常任务,DTS会尝试自动恢复,可能会导致迁移任务的源端数据库数据覆盖
 目标实例数据库中写入的业务数据,迁移任务结束后务必将DTS访问目标实例账号的写权限用
 revoke命令回收掉。

迁移限制

- ・迁移仅支持 MySQL 5.6版本。
- · 结构迁移不支持 event 的迁移。
- · 对于MySQL的浮点型float/double, DTS通过round(column,precision)来读取该列的 值,若列类型没有明确定义其精度,对于float,精度为38位,对于double类型,精度为308
 ,请先确认DTS的迁移精度是否符合业务预期。
- ·如果使用了对象名映射功能,依赖这个对象的其他对象可能迁移失败。
- · 当选择增量迁移时,源端的 MySQL 实例需要开启 binlog。
- · 当选择增量迁移时, 源库的 binlog_format 要为 row。
- · 当选择增量迁移且源 MySQL 为 5.6 版本时,它的 binlog_row_image 必须为 full。
- · 当选择增量迁移时,增量迁移过程中如果源MySQL实例出现因实例跨机迁移或跨机重建等导致的binlog 文件ID乱序,可能导致增量迁移数据丢失。

迁移权限要求

当使用 DTS 进行本地 MySQL > POLARDB 迁移时,不同的迁移类型,对源端和目标端 MySQL 实例的迁移账号权限要求如下:

迁移类型	结构迁移	全量数据迁移	增量数据迁移
本地 MySQL 实例	select	select	super
			select
			replication slave
			replication client
目标端 POLARDB 集 群	读写权限	读写权限	读写权限

迁移流程

DTS 在进行本地 MySQL > POLARDB 数据迁移时,为了解决对象间的依赖关系,提高迁移成功率。结构对象及数据的迁移顺序如下:

1. 结构对象:表、视图的迁移。

2. 全量数据迁移。

3. 结构对象:存储过程、函数、触发器、外键的迁移。

4. 增量数据迁移。

📃 说明:

如果没有选择增量数据迁移,那么当全量数据迁移完成后,任务列表中的迁移进度为:结构迁移 100%,全量迁移100%,迁移状态为迁移中。此时迁移任务正在进行步骤3中的对象的迁移。请 勿手动结束任务,否则会造成迁移数据丢失。

操作步骤

- 1. 登录DTS控制台。
- 2. 在左侧菜单栏单击数据迁移,单击右上角创建迁移任务。
- 3. (可选)填写任务名称。

DTS 为每个任务自动生成一个任务名称,任务名称没有唯一性要求。您可以根据需要修改任务 名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。

4. 填写源库和目标库信息,具体参数配置说明如下:

库类别	参数	说明
源库信息	实例类型	源数据库实例类型,选择有公网IP的自建数据库。
	实例地区	选择源数据库所在的地区。
	数据库类型	源数据库类型,选择MySQL。
	主机名或IP地址	源数据库的公网IP地址。
	端口	源数据库的端口。
	数据库账号	源数据库具有读写权限的账号。
	数据库密码	源数据库账号对应的密码。
目标库信息	实例类型	选择POLARDB。
	实例地区	目标实例所在的地区。
	POLARDB实例ID	对应地区下的实例ID,选择要迁移到的目标实例的ID。
	数据库账号	目标实例的拥有读写权限的账号。

库类别	参数	说明			
	数据库密码	目标实例的对应账号的密码。			
	*任务名称: dt	szuak08it			
源库信息					
	* 实例类型:	有公网IP	的自建数据库 ▼		
	* 实例地区:	华东1(札	沆州) ▼	获取DTS IP段	
	*数据库类型:	MySQL	۲		
	* 主机名或IP地址:				
	* 端□:	3306			
	* 数据库账号:]	
	* 数据库密码:	••••••		测试连接	⊘ 测试通过
目标库信息					
	* 实例类型:	POLARD	POLARDB V		
	* 实例地区:	华东1(杭州)]	
	* POLARDB实例ID:	· · · · · · · · · · · · · · · · · · ·]	
	* 数据库账号:]	
	* 数据库密码:	•••••	4 >	测试连接	⊘ 测试通过

5. 填写完成后单击测试连接,确定源库和目标库都测试通过。

6. 单击右下角授权白名单并进入下一步。

- 7. 选择迁移类型和迁移对象。
 - ・迁移类型:
 - 结构迁移

DTS 会将迁移对象的结构定义迁移到目标集群。目前 DTS 支持结构迁移的对象有:表、 视图、触发器、存储过程、存储函数。

- 全量数据迁移

DTS 会将源数据库迁移对象的存量数据全部迁移到目标集群。由于全量迁移并发Insert 导致目标实例的表存在碎片,全量迁移完成后目标实例的表空间会比源实例大。

如果用户只进行全量数据迁移,那么迁移过程中本地 MySQL 实例新增的业务写入不会被同步到目标 RDS for MySQL 实例。

- 增量数据迁移

DTS 会将迁移过程中源实例的数据变更同步到目标集群。如果迁移期间进行了 DDL 操 作,那么这些结构变更不会迁移到目标集群。

如果只需要进行全量迁移,建议迁移类型选择:结构迁移+全量数据迁移。

如果需要进行不停机迁移,建议迁移类型选择:结构迁移+全量数据迁移+增量数据迁移。



结构迁移和全量数据迁移均不收费,增量迁移需要收费。

·迁移对象:选择要迁移的对象,单击向右的箭头,将选中的对象添加到右侧。

迁移对象的选择粒度细化为:库、表、列三个粒度。默认情况下,对象迁移到目标POLARDB集群后,对象名跟源RDS实例一致。如果您迁移的对象在源实例跟目标集群上名称不同,那么需要使用DTS提供的对象名映射功能,详细使用方式可以参考库表列映射。

📕 说明:

- 暂时不支持对系统表的迁移。

目标集群中不能有和迁移对象同名的对象。将鼠标移至右侧框中的对象,单击编辑,即可 修改迁移后的对象名。

创建迁移任务 \$200上级				
1.源库及目标库		2.迁移类型及列:	R 🔶	3.预检查
	 • 这些出版: ○ 结构过程 · 这些比較正確 · 」 · 」 · 」	¥到目标实例 够		
	迁移对象		已选择对象(就标移到对象行。点击编辑可称改对象名或过读条件)详情自我	
	t 🖕 test2		test1	
		>		
		<		
	全选中		全部除	
	注意: 1. 数据迁移只会将源库的数据(结构) 图制一份到目标数据库,并不会对源数据 2. 数据迁移过程中,不支持ODL操作,如进行ODL操作可能导致迁移失败	医卑激症(结构)道	成影响,	
				取消 上一步 保存 预检查并启动

8. 单击右下角的预检查并启动,成功后单击下一步。



如果预检查失败,可以单击具体检查项后的检查结果,查看具体的失败详情,并根据失败原因 修复后,重新进行预检查。

预检查			\times
		预检查通过10	0%
检测项	检测内容	检测结果	
源库连接性检查	检查数据传输服务器是否能连通源数据库	成功	-
源库连接性检查	检查数据传输服务器是否能连通源数据库	成功	
源库权限检查	检查源数据库的账号权限是否满足迁移要求	成功	
目的库连接性检查	检查数据传输服务器是否能连通目的数据库	成功	
目的库连接性检查	检查数据传输服务器是否能连通目的数据库	成功	I.
目的库权限检查	检查目的数据库的账号权限是否满足迁移要求	成功	
存储引擎检查	检查迁移表是否有不支持的存储引擎	成功	*
		下一	步

9. 确认DTS购买信息,阅读并勾选服务条款,单击立即购买并启动。

如果选择了增量迁移,那么进入增量迁移阶段后,源库的更新写入都会被DTS同步到目标 POLARDB集群。迁移任务不会自动结束。如果您只是为了迁移,那么建议在增量迁移无延迟的 状态时,源库停写几分钟,等待增量迁移再次进入无延迟状态后,停止掉迁移任务,直接将业务 切换到目标POLARDB集群上即可。

10.单击目标地域,查看迁移状态。迁移完成时,状态为已完成。

迁移	任务名 ▼ 请编入迁移任务名进行搜索	搬業 排序: 数以排序 v 状态: 全部 v	
	ID/名称:	状态: 已完成	查看洋语 创建类似任务
	2019-01-25 14:22:59 创建		2019-01-25 14:26:03 完成
	结构迁移 100%		全量迁移 100%(已迁移0行)

至此,完成本地 MySQL 数据库到阿里云 POLARDB 的数据迁移任务。

1.2.6 从POLARDB for MySQL迁移至RDS for MySQL

本文介绍使用阿里云<mark>数据传输服务 (DTS</mark>) ,将POLARDB for MySQL迁移至RDS for MySQL的 操作步骤及注意事项。

迁移前准备工作

·为源集群设置IP白名单

迁移前,您需要为源POLARDB集群设置白名单,在白名单中添加DTS的IP段。

▋ 说明:

您只需放开目标数据库所在区域对应的*DTS IP*段。本案例中,目标数据库地区为杭州,您只需 要添加杭州地区的DTS IP地址段。

・ 创建数据库账号

迁移任务配置时,需要提供源POLARDB集群及目的RDS实例的迁移账号。如果尚未创建迁移 账号,您可以参考*POLARDB*创建集群账号和*RDS for MySQL*创建账号。先在源集群及目的实例中 创建迁移账号,并将要迁移的库表的读写权限授权给上面创建的账号。

迁移权限要求

当使用 DTS 进行POLARDB > RDS for MySQL 迁移时,不同的迁移类型,对源端和目标端数据 库的迁移账号权限要求如下:

数据库类型	结构迁移	全量数据迁移
源 POLARDB 集群	只读权限	只读权限
目的RDS for MySQL实例	读写权限	读写权限

注意事项

- · POLARDB暂不支持增量迁移;
- ·保证迁移数据一致性,在开始迁移前,需停止写入数据到源POLARDB集群;
- ・为保证迁移成功,目标实例的存储空间应大于源POLARDB集群已使用空间。

迁移步骤

- 1. 登录DTS控制台。
- 2. 在左侧导航栏选择数据迁移,然后在右上角单击创建迁移任务。
- 3. 填写源库和目标库信息,具体参数配置说明如下。

库类别	参数	说明
源库	实例类型	源库实例类型,选择有公网IP的自建数据库。

库类别	参数	说明
	实例地区	源POLARDB集群所在的地域。
	数据库类型	源数据库类型,选择MySQL。
	主机名或IP地址	POLARDB集群的公网集群地址。您可以参考 <u>查看连接地址</u> 。
	端口	默认的3306端口。
	数据库账号	POLARDB集群的账号。
	数据库密码	POLARDB集群账号的密码。
目标库	实例类型	目标实例的类型,选择RDS实例。
	实例地区	目标实例的地域。
	RDS实例ID	目标实例的ID。
	数据库账号	目标实例的拥有读写权限的账号。
	数据库密码	目标实例的对应账号的密码。

库类别	参数	说明
	连接方式	有非加密传输和SSL安全连接两种连接方式,选择SSL安全加 密连接会显著增加CPU消耗。
		道 说明: 只有支持并开启了 <i>SSL</i> 安全连接的实例才需要选择SSL安全连接。

*任务名称:		
源库信息		
* 实例类型:	有公网IP的自建数据库 ▼	
* 实例地区:	华东1(杭州)	获取DTS IP段
*数据库类型:	MySQL	
* 主机名或IP地址:	rwlb.rds.aliyuncs.com	
* 端口:	3306	
* 数据库账号:		
* 数据库密码:	••••••• Ø	测试连接 🛛 测试通过
* 数据库密码:	••••••• Ø	测试连接 🛛 测试通过
* 数据库密码: 目标库信息	••••••• Ø	测试连接 ⊘ 测试通过
* 数据库密码: 目标库信息 * 实例类型:	••••••• 《 RDS实例	测试连接 ⊘ 测试通过
* 数据库密码: 目标库信息 * 实例类型: * 实例地区:	••••••• ♥ RDS实例 ¥东1(杭州)	测试连接 ⊘ 测试通过
* 数据库密码: 目标库信息 * 实例类型: * 实例地区: * RDS实例ID:	•••••••••••••••••••••••••••••••••••••	测试连接 ⊘ 测试通过
* 数据库密码: 目标库信息 * 实例类型: * 实例地区: * RDS实例ID: * 数据库账号:	•••••••● RDS实例 ▼ 华东1(杭州) ▼ 「m 'td	测试连接 ⊘ 测试通过
* 数据库密码:	••••••••● RDS实例 ▼ 华东1(杭州) ▼ rm *td ● ●	测试连接 ⊘ 测试通过

4. 填写完毕后单击测试连接,确定源库和目标库都测试通过。

5. 单击授权白名单并进入下一步。

- 6. 选择迁移类型和迁移对象。
 - · 迁移类型:选择结构迁移和全量数据迁移。(暂不支持增量迁移。)为保证迁移数据一致
 性,在开始迁移前,需停止写入数据到源集群。
 - 结构迁移:

DTS 会将迁移对象的结构定义迁移到目标实例。目前 DTS 支持结构迁移的对象包括:表。其他对象如视图、同义词、触发器、存储过程、存储函数、包、自定义类型等暂 不支持。

- 全量数据迁移:

DTS 会将源数据库迁移对象的存量数据全部迁移到目标实例。

·迁移对象:选择要迁移的对象,单击向右的箭头,将选中的对象添加到右侧。



- 暂时不支持对系统表的迁移。

目标集群中不能有和迁移对象同名的对象。将鼠标移至右侧框中的对象,单击编辑按
 钮,即可修改迁移后的对象名。

* 迁移类型: ☑ 结构迁移 ☑ 全量数据迁移 □ 增量数据迁移 全量数据迁移过程中、如果源库有数据更新,这部分更新数据不保证被迁移 为保证迁移数据的一致性,建议选择结构迁移+全量数据迁移+增量数据迁移	到目标实例 3	
迁移対象	> <	已选择对象(急标移到对象行,点击编辑可修改对象名或过滤条件)详情点我 ■ ■ ■ ■ ■ ■ ■
全选中 注意: 1. 数据迁移只会将原库的数据(结构)复制一份到目标数据库,并不会对源数据 2. 数据迁移过程中,不支持DDL操作,如进行DDL操作可能导致迁移失败	库数层(结构))	全移除 202影响。 取消 上一步 保存 预检查并启动
7. 单击预检查并启动,等待预检查结束。

- i台 flH・
ルック・

如果检查失败,可以根据错误项的提示进行修复,然后重新启动任务。

	<u> 预检查通过</u> 100%
检测内容	检测结果
检查数据传输服务器是否能连通源数据库	成功
检查数据传输服务器是否能连通源数据库	成功
检查源数据库的账号权限是否满足迁移要求	成功
检查数据传输服务器是否能连通目的数据库	成功
检查数据传输服务器是否能连通目的数据库	成功
检查目的数据库的账号权限是否满足迁移要求	成功
於杏洋铊末旦丕右太支持的左腔引数	яðтн
	下一步

 8. 单击下一步,在购买配置确认对话框中,勾选《数据传输(按量付费)服务条款》并单击立即购 买并启动。



迁移任务列	则表 华东1(杭州)	华东2(上海)	华北1 (青岛)	华北2(北京)	华南1 (深圳)	华北3(张家口)	香港	美国(硅谷)	美国 (弗吉尼亚)	新加坡	阿联酋 (迪拜)	德国(法兰克福)	马来西亚(吉隆坡)			
	演大利亚 (悉尼)	印度(孟买)	英国 (伦敦)	日本 (东京)	印度尼西亚(雅加	达) 华北5(1)	附和浩特)							C 1005	文件导入	ANTELEFAL
													0.046001938	- 1000	2002/	Chapter of LENG
迁移任务名	▼ 请输入迁移任务3	名进行搜索		搜索 排序	: 默认排序	▼ 状态:	全部	•								
□ ID/名4	弥:dts	18		状态:	已完成										查看》	情 创建类似任务
2019-0	03-08 14:38:17 创建 汗移 100%								全量迁移 100%(已迁移200	000行)				2019-	03-08 14:44:22 完成

1.2.7 一键升级RDS for MySQL到POLARDB for MySQL POLARDB支持将RDS for MySQL一键升级为POLARDB for MySQL。

前提条件

- ・源RDS实例版本为RDS for MySQL 5.6高可用版。
- ・源RDS实例未开启TDE和SSL。
- ・源RDS实例的表存储引擎为InnoDB。
- ·如果RDS处于高安全模式(数据库代理模式),需要创建有高权限账号(请参见#unique_33/ unique_33_Connect_42_section_wkq_j35_q2b),或者切换到高性能模式(参见#unique_34),才 能进行一键升级。

<		(运行中) 1 返回失例》末	「平海江修賀 9名同語 操作指引 登录数据库 迁移数据库 重启实例 G 刷新
基本信息	-	数据库连接	
账号管理 备份恢复	Ŀ	实闭连接 读写分离	
数据库管理	ŧ.	数据库连接	如何连接RDS 为什么连接不上 初换为专有网络 初换访问模式 修改连接地址 申请外网地址
数据库连接	J.	网络类型: 经典网络 ❷	数据库代理状态(原高安全模式): 已开通 ❷
数据库代理		内网地址: aliyuncs.com	内网端□: 3306
监控与报警	-	温馨提示:请使用以上访问连接串进行实例连接,VIP在业务维护中可能会变化。	
数据安全性		\	
服务可用性		切时便和活的图	查看其他組合切问機式路构图

背景信息

POLARDB是阿里云自研的下一代关系型云数据库,主要优势如下:

- ・存储容量高:最高可达100TB。
- ・性能高:最高可以提升至MySQL的6倍。
- · Serverless存储:存储容量无需提前购买,自动扩缩容,按使用量计费。
- · 临时升配:临时升级规格,轻松应对短期的业务高峰。

详情请参见#unique_35。

一键升级功能可以将RDS for MySQL一键升级为POLARDB for MySQL,升级后POLARDB集 群包含源RDS实例的账号、数据库、IP白名单和必要的参数。

一键升级的功能亮点

- ・迁移完全免费。
- ・迁移过程数据0丢失。
- ・支持増量迁移,停机时间小于10分钟。
- ・支持回滚, 迁移失败可以在10分钟内恢复。

迁移流程

1. 参见从*RDS*迁移的说明,创建一个与源RDS实例数据相同的POLARDB集群,源RDS实例的增量数据会实时同步到该POLARDB集群。



需要在7天内修改应用端的数据库地址为POLARDB地址,确认业务正常,以及单击完成迁移。 单击完成迁移会中断RDS和POLARDB之间的数据同步。

2. 单击迁移切换。该操作将源RDS实例修改为只读,将POLARDB集群修改为可读可

写,POLARDB的增量数据会实时同步到RDS。修改数据库连接地址。具体操作请参见迁移切换。

门 说明:

迁移切换后,也可以选择迁移回滚。

3. 完成迁移。

注意事项

- ・迁移只能在相同地域内进行。
- ・源RDS实例在迁移时不能修改参数。

从RDS迁移

本操作将创建一个与源RDS实例数据相同的POLARDB集群,源RDS实例的增量数据会实时同步 到该POLARDB集群。

- 1. 登录POLARDB控制台。
- 2. 单击创建新集群。
- 3. 选择包年包月或按量付费页签。
- 4. 设置以下参数。

参数	说明
地域	源RDS for MySQL实例所在地域。
	道 说明: 新建的POLARDB集群也在此地域。
创建方式	选择从RDS迁移。即从RDS实例克隆一个POLARDB集群,同时保持 数据同步。默认开启新集群的Binlog。
源RDS引擎	源RDS实例的引擎类型,不可变更。

参数	说明			
源RDS版本	源RDS实例的版本,不可变更。			
源RDS实例	可选的源RDS实例,不包括只读实例。			
可用区	可用区是地域中的一个独立物理区域,不同可用区之间没有实质性区 别。 您可以选择将POLARDB集群与ECS创建在同一可用区或不同的可用 区。			
网络类型	POLARDB集群的网络类型,不可变更。			
VPC网络 VPC交换机	POLARDB集群所属的VPC和虚拟交换机。请确保POLARDB集群与 需要连接的ECS创建于同一个VPC,否则它们无法通过内网互通,无 法发挥最佳性能。			
数据库引擎	POLARDB集群的数据库引擎,不可变更。			
节点规格	按需选择,建议不低于源RDS实例规格。所有POLARDB节点均为独 享型,性能稳定可靠。详情请参见#unique_40。			
节点个数	无需选择。系统将自动创建一个与主节点规格相同的只读节点。			
存储费用	无需选择容量,根据实际数据使用量按小时计费。详情请参 见#unique_40。			
集群名称	填写集群名称用于区分业务用途。如果留空,系统将自动生成一个集 群名称。创建集群后还可以修改。			

5. 设置购买时长(仅针对包年包月集群),然后单击右侧的立即购买。

6. 确认订单信息,阅读和勾选服务协议,单击去开通,完成支付。

7. 进入POLARDB控制台,查看新建的POLARDB集群的状态。

🗾 说明:

- 集群创建后开始从RDS实例同步数据,7天内需要完成修改数据库连接地址以及完成迁移操
 作。超过7天将自动关闭迁移功能。
- ·您可以在此步骤选择取消迁移,相关影响请参见迁移常见问题。

迁移切换

满足以下条件后,您可以进行迁移切换,然后修改应用里的数据库连接地址。

・已完成<u>从RDS</u>迁移的操作。

・复制延迟小于60秒。

<	集群 pc-	• 运行中		登录数据库
基本信息	基本信息			
▼ 配置与管理	集群ID	pc-	集群名称	pc 编辑
账号管理	地域	华东1(杭州)	数据分布的可用区	杭州 可用区G (主), 杭州 可用区I
数据库管理	兼容数据库	MySQL 5.6	状态	• 运行中
备份恢复	VPC 可维护窗口	vpc 02:00-03:00 修改	交换机	VSW-
参数配置				
▼ 诊断与优化	从RDS迁移			
集群总览	迁移切换 取消迁移			
性能监控	源RDS	m	状态	● 数据同步中
问题分析	源RDS读写状态	读写	POLARDB读写状态	只读
慢 SQL	复制关系复制延迟	从RDS (m) 到POLARDB (pc)同步		
▼ 日志与审计	复制到期时间	 2019年07月21日 12:39:14 (剩余时长 6天 23小时 47分 22秒)		
SQL 洞察		到期后,将关闭迁移功能,源和目标数据库均恢复【读写】状态,脱离复制关系。		

- 1. 进入POLARDB控制台。
- 2. 找到目标集群,单击集群的ID。
- 在基本信息页面单击迁移切换,在弹出的对话框中单击确定。本操作将源RDS实例修改为只 读,将POLARDB集群修改为可读可写,同时会将POLARDB集群的新增数据同步到RDS实 例。

基本信息	基本信息	
▼ 配置与管理	集群ID	pc-
账号管理	地域	华东1 (杭州)
数据库管理	兼容数据库	MySQL 5.6
备份恢复	VPC	
参数配置		02.00-03.00 (State
▼ 诊断与优化	从RDS迁移	
集群总览	迁移切换 取消迁移	8
性能监控	源RDS	rm-
问题分析	源RDS读写状态	读写
	复制关系	从RDS (m-
慢 SQL	复制延迟	0
▼ 日志与审计	复制到期时间	2019年07月15日 15:27:49 (剩余时长 6天 23小时 59分 5秒)
SQL 洞察		到期后,将关闭迁移功能,源和目标数据库均恢复【读写】状态,脱离复制关系。
道 说明:		

- ·数据同步的延迟超过60秒时无法进行迁移切换。
- ・切换过程一般小于5分钟。

4. 刷新页面,当POLARDB读写状态显示为读写后,尽快修改应用里的数据库连接地址。

从RDS迁移								
迁移回滚 完成记	E8							
源RDS	rm-	状态	● 数据库切换完成					
源RDS读写状态	只读	POLARDB读写状态	读写					
复制关系	从POLARDB (pc-I) 到RDS (rm) 同步							
复制延迟	0							
复制到期时间	2019年07月15日 15:27:49 (剩余时长 6天 23小时 46分 25秒)							
	到期后,将关闭迁移功能,源和目标数据库均恢复【读写】状态,脱离复制关系。							



迁移切换完成后,也可以选择迁移回滚。

完成迁移

从*RDS*迁移后,需要在7天内修改数据库连接地址以及单击完成迁移。该操作将中断POLARDB集 群和RDS实例间的数据同步。



由于本操作将中断POLARDB集群和RDS实例间的数据同步,不再提供迁移回滚功能,建议您使 用一段时间POLARDB集群,确认正常后再执行本操作。

- 1. 进入POLARDB控制台。
- 2. 找到目标集群,单击集群的ID。

3. 在基本信息页面,单击完成迁移,在弹出的对话框中单击确定。

基本信息			
集群ID	pc-	集群名称	从RDS迁移编辑
地域	华东1(杭州)	兼容数据库	MySQL 5.6
状态	 运行中 	可维护窗口	02:00-03:00 修改
VPC	Vpc	交换机	VSW-
从RDS迁移			
迁移回滚 完成迁移			
源RDS	rm-	状态	● 数据库切换完成
源RDS读写状态	只读	POLARDB读写状态	读写
复制关系	从POLARDB (pc)) 到RDS (rm) 同步		
复制延迟	0		
复制到期时间	2019年06月17日 15:42:10 (剩余时长 6天 23小时 22分 12秒)		
	到期后,将关闭迁移功能,源和目标数据库均恢复【读写】状态,脱离复制关系。		

完成迁移	×
完成本次迁移任务。	
将中断RDS MYSQL实例 rm	
□ 关闭 POLARDB 集群的Binlog	
关闭Binlog会带来少量的写入性能提升,但需要重启POLARDB。您也可以稍后在 【参数配置】中关闭。	
确定取消	

- ・単击确定后,系统将在约2分钟内中断同步关系,期间完成迁移按钮不会消失,请勿重复单 击。
- ·您可以选择是否关闭POLARDB集群的Binlog。关闭Binlog会带来少量的写入性能提升,但需要重启POLARDB。
- 4. 如果不再需要源RDS实例,可以释放实例。

迁移回滚

在完成迁移前,如果您发现数据存在异常等问题,可以进行回滚操作,快速恢复至迁移前的状态(RDS实例为可读可写,POLARDB集群为只读,同时会将RDS实例的数据同步到POLARDB集 群)。

- 1. 进入POLARDB控制台。
- 2. 找到目标集群,单击集群的ID。
- 3. 在基本信息页面单击迁移回滚,在弹出的对话框中单击确定。

基本信息			
集群ID	pc-	集群名称	从RDS迁移编辑
地域	华东1(杭州)	兼容数据库	MySQL 5.6
状态	• 运行中	可维护窗口	02:00-03:00 修改
VPC	vpc	交换机	VSW-
从RDS迁移			
迁移回滚 完成迁移			
源RDS	rm-	状态	● 数据库切换完成
源RDS读写状态	只读	POLARDB读写状态	读写
复制关系	从POLARDB (pc)) 到RDS (rm) 同步		
复制延迟	0		
复制到期时间	2019年06月17日 15:42:10 (剩余时长 6天 23小时 22分 12秒)		
	到期后,将关闭迁移功能,源和目标数据库均恢复【读写】状态,脱离复制关系。		

说明:

单击确定后RDS实例为可读可写,POLARDB集群为只读,同时会将RDS实例的数据同步 到POLARDB集群。当源RDS读写状态显示为读写后,请尽快修改应用里的数据库连接地址 为RDS连接地址。

迁移常见问题

・从RDS迁移会影响源RDS实例吗?

答:不会影响源RDS实例的正常运行。

· 平滑迁移对业务有影响吗?

答:平滑迁移能够保证迁移过程不丢失数据,停机时间小于10分钟,如果有需要还可以进行回 滚。

· 取消迁移会有什么影响?

答:取消迁移后,源RDS实例可以修改参数;POLARDB集群恢复可读可写,且数据不会释 放。手动取消时可以选择是否关闭POLARDB集群的Binlog,自动取消时不会关闭。

相关API

API	描述
#unique_42	创建POLARDB集群。
	道 说明: 一键升级时,参数CreationOption取值需要 为MigrationFromRDS。
#unique_43	查询POLARDB集群的迁移状态。
#unique_44	修改迁移任务,进行任务的切换或回滚。
#unique_45	取消或完成迁移。

1.2.8 一键克隆RDS for MySQL到POLARDB for MySQL

POLARDB支持从RDS for MySQL一键克隆数据到新的POLARDB for MySQL集群。

前提条件

- ・源RDS实例版本为RDS for MySQL 5.6高可用版。
- ・源RDS实例未开启TDE和SSL。
- ・源RDS实例的表存储引擎为InnoDB。
- ・如果RDS处于高安全模式(数据库代理模式),需要创建有高权限账号(请参见#unique_33/

unique_33_Connect_42_section_wkq_j35_q2b),或者切换到高性能模式(参见#*unique_34*),才能进行一键克隆。

<		(运行中) 本运阅实例列表	网络平滑迁移	到专有网络	操作指引	登录数据库	迁移数据库	重启实例	备份实例	C 刷新
基本信息 账号管理		数据库连接								
备份恢复		实例距接 读写分离								
数据库管理		数据库连接	如何這	接RDS 📀	为什么连接不上	切换为专有网络	切换访问模式	修改连接加	地址 申请外	小网地址
数据库连接		网络类型: 经共网络 🞯		数据库代理	■状态(原高安全	注模式): 已开通 🌘	>			
敗据库代理		内网地址: allyuncs.com		内网端口:	3306					
监控与报警	-	温馨提示:请使用以上访问连接串进行实例连接,VIP在业务维护中可能会变化。								
数据安全性										
服务可用性		访问模式站科图						查看3	其他組合访问模式	尤结构图

背景信息

POLARDB是阿里云自研的下一代关系型云数据库,主要优势如下:

- ・存储容量高:最高可达100TB。
- ・性能高:最高可以提升至MySQL的6倍。
- · Serverless存储:存储容量无需提前购买,自动扩缩容,按使用量计费。
- ·临时升配:临时升级规格,轻松应对短期的业务高峰。

详情请参见#unique_35。

一键克隆功能将会新建一个与源RDS实例的数据相同的POLARDB集群,POLARDB集群 包含源RDS实例的账号、数据库、IP白名单和必要的参数。源RDS实例的增量数据不会同步 到POLARDB集群。

| ■ 说明:

如果需要在新建POLARDB集群的同时,使源RDS实例的增量数据实时同步到POLARDB集群,即实现平滑迁移(不停机迁移),请参见#unique_46。

一键克隆的功能亮点

- ・免费
- ・克隆过程数据0丢失

一键克隆的操作步骤

- 1. 登录POLARDB控制台。
- 2. 单击创建新集群。
- 3. 选择包年包月或按量付费页签。
- 4. 设置以下参数。

参数	说明
地域	源RDS for MySQL实例所在地域。
	道 说明: 克隆的POLARDB集群也在此地域。
创建方式	集群的创建方式:
	 · 默认创建: 创建一个全新的POLARDB集群。 · 从RDS克隆: 基于所选的RDS实例,克隆一个数据完全一样的POLARDB集群。 · 从RDS迁移: 先从RDS实例克隆一个POLARDB集群,同时保持同步。默认开启新集群的Binlog。 这里选择从RDS克隆。
源RDS引擎	源RDS实例的引擎类型,不可变更。
源RDS版本	源RDS实例的版本,不可变更。
源RDS实例	可选的源RDS实例,不包括只读实例。

参数	说明
可用区	可用区是地域中的一个独立物理区域,不同可用区之间没有实质性区 别。 您可以选择将POLARDB集群与ECS创建在同一可用区或不同的可用 区。
网络类型	POLARDB集群的网络类型,不可变更。
VPC网络 VPC交换机	POLARDB集群所属的VPC和虚拟交换机。请确保POLARDB集群与 需要连接的ECS创建于同一个VPC,否则它们无法通过内网互通,无 法发挥最佳性能。
数据库引擎	POLARDB集群的数据库引擎,不可变更。
节点规格	按需选择,建议不低于源RDS实例规格。所有POLARDB节点均为独 享型,性能稳定可靠。详情请参见#unique_40。
节点个数	无需选择。系统将自动创建一个与主节点规格相同的只读节点。
存储费用	无需选择容量,根据实际数据使用量按小时计费。详情请参 见#unique_40。
集群名称	填写集群名称用于区分业务用途。如果留空,系统将自动生成一个集 群名称。创建集群后还可以修改。

5. 设置购买时长(仅针对包年包月集群),然后单击右侧的立即购买。

6. 确认订单信息,阅读和勾选服务协议,单击去开通。

常见问题

从RDS克隆会影响源RDS实例吗?

答:不会影响源RDS实例的正常运行。

相关API

API	描述
#unique_42	创建POLARDB集群。
	送 说明: 一键克隆时,参数CreationOption取值需要 为CloneFromRDS。

后续步骤

请尽快将应用的数据库连接地址修改为POLARDB的地址,详情请参见#unique_32。

1.3 数据同步

1.3.1 从RDS MySQL同步至POLARDB MySQL

本小节介绍如何使用数据传输服务DTS快速创建RDS MySQL实例同POLARDB MySQL集群间的 实时同步作业,实现RDS MySQL到POLARDB MySQL增量数据的实时同步。

前提条件

已购买目标POLARDB MySQL集群,详情请参见创建POLARDB MySQL集群。

注意事项

- ·如果同步对象为单个或多个表(非整库),那么在数据同步时,请勿对源库的同步对象使用ghost或pt-online-schema-change等类似工具执行在线DDL变更,否则会导致同步失败。
- ・ 全量初始化过程中,并发INSERT会导致目标集群的表碎片,全量初始化完成后,目标集群的表 空间比源集群的表空间大。
- ・如果数据同步的源集群没有主键或唯一约束,且记录的全字段没有唯一性,可能会出现重复数 据。

支持同步的SQL操作

- DML: INSERT, UPDATE, DELETE
- DDL: CREATE TABLE、ALTER TABLE、RENAME TABLE、TRUNCATE TABLE、 DROP TABLE

支持的同步架构

- ・一对一单向同步
- ・一对多单向同步
- ・级联单向同步
- 多对一单向同步

关于各类同步架构的介绍及注意事项,请参见#unique_48。

功能限制

・不兼容触发器

同步对象为整个库且这个库中包含了会更新同步表内容的触发器,那么可能导致同步数据不一 致。例如数据库中存在了两个表A和B。表A上有一个触发器,触发器内容为在INSERT一条 数据到表A之后,在表B中插入一条数据。这种情况在同步过程中,如果源集群表A上进行了 INSERT操作,则会导致表B在源集群跟目标集群数据不一致。

此类情况须要将目标集群中的对应触发器删除掉,表B的数据由源集群同步过去,详情请参见<u>触</u> 发器存在情况下如何配置同步作业。

・ RENAME TABLE限制

RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A,如果同步过程中源 集群将表A重命名为表B,那么表B将不会被同步到目标库。为避免该问题,您可以在数据同步 配置时,选择同步表A和表B所在的整个数据库作为同步对象。

操作步骤

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdlo

说明:

购买时,选择源实例为MySQL、目标实例为POLARDB,并选择同步拓扑为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。
- 4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

数据传输	同步作业列表 华东1 (杭州	华东2 (上海) 华北1	(青岛) 华北2 (北京)	华南1 (深圳) 华北3 (张	家口) 香港		TRAUXIL A
477.14	美国 (硅谷)	美国 (弗吉尼亚) 新加	1坡 阿联酋 (迪拜) 領	唐国 (法兰克福) 马来西亚 (1	吉隆坡)	(所选地域为同步作业目标实例	所在的地域)
概党	演大利亚 (恩	尼) 印度 (孟买) 英国	(伦敦) 日本 (东京)	印度尼西亚(雅加达) 华	比5 (呼和浩特)		
数据迁移							2 刷新
数据订阅						Ŭ	
数据同步	同步作业名称 ▼		搜索	排序: 默认排序 🔻	状态: 全部	Ŧ	

5. 定位至已购买的数据同步实例,单击配置同步链路。

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实	例 2.选择同步对象	>	3.高级设置	<u> </u>		4.预检查
同步作业名称:	RDS_TO_POLARDB					
源实例信息						
实例类型:	RDS空例	*				
实例地区:	华东1 (杭州)					
* 实例ID:	rm-bn		其他阿里云账号下的RDS实例			
		-				
90旗件项号:	dtstest					
* 数据库密码:	*************	Ø >				
* 连接方式:	● 非加密连接 ○ SSL安全连接					
目标实例信息						
实例类型:	POLARDB					
实例地区:	华东1 (杭州)					
* POLARDB实例ID:	pc-bp	•				
* 数据库账号:	dtstest					
* 数据库密码:		«				
					HINCLE	低现内存并并进入工 正
					取消	授权白名单并进入下一步

配置项目	配置选项	配置说明
任务名称	-	 DTS为每个任务自动生成一个任务名称,任务名称没有唯一性要求。 您可以根据需要修改任务名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。
源实例信息	实例类型	选择RDS实例。
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。
	数据库账号	填入连接RDS实例的数据库账号。
		 说明: 该账号需具备待同步对象的SELECT、REPLICATION CLIENT、REPLICATION SLAVE权限。 当源RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时,无需配置数据库账号和数据库密码。
	数据库密码	填入数据库账号对应的密码。

配置项目	配置选项	配置说明					
	连接方式	根据需求选择非加密连接或SSL安全连接,本案例选择为非加密 连接。					
		道 说明: 选择SSL安全连接时,需要提前开启RDS实例的SSL加密功 能,详情请参见 <mark>设置</mark> SSL加密。					
目标实例信	实例类型	固定为POLARDB,不可变更。					
息	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。					
	POLARDB 实例ID	选择目标POLARDB集群ID。					
	数据库账号	填入连接POLARDB集群的数据库账号。					
		前明: 用于数据同步的数据库账号需具备目标同步对象的ALL权限。					
	数据库密码	填入数据库账号对应的密码。					

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到源RDS实例和目标POLARDB集群的白名单中,用 于保障DTS服务器能够正常连接源和目标实例。 8. 配置目标已存在表的处理模式和同步对象。

1.选择同步通道的源及目标实例 2.选择	洞步对象	3.高级设置	\rightarrow	4.预检查
同步架构: 单向同步				
目标已存在表的处理模式: 💿 预检查并报错拦截 🔘 🤊	石操作			
源库对象		已选择对象(鼠标移到对象行,点击编	辑可修改对象名或过法	1条件)详情 点我
若全局搜索,请先展开树 Q			Q	
🖃 늘 dtstestdata		🦢 dtstestdata (1个对象)		
Tables		i order		
📰 vip 관 🚰 polardbdatabase				
	>			
	<			
全选		全选		
		13		
*映射名称更改:	認 ◎ 要进行库表名称批量更改			
				取消 上一步 下一步

配置项目	配置说明
目标已存在表的处理 模式	 预检查并报错拦截:检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表,则通过该检查项目;如果目标数据库中有同名的表,则在预检查阶段提示错误,数据同步作业不会被启动。
	 说明: 如果目标库中同名的表不方便删除或重命名,您可以#unique_50来 避免表名冲突。 无操作:跳过目标数据库中是否有同名表的检查项。
	 警告: 选择为无操作,可能导致数据不一致,给业务带来风险,例如: 表结构一致的情况下,如果在目标库遇到与源库主键的值相同的记录,在初始化阶段会保留目标库中的该条记录;在增量同步阶段则会覆盖目标库的该条记录。 表结构不一致的情况下,可能会导致无法初始化数据、只能同步部分列的数据或同步失败。

配置项目	配置说明
选择同步对象	在源库对象框中单击待同步的对象,然后单击 > 将其移动至已选择
	对象框。
	同步对象的选择粒度为库、表。
	道 说明:
	 · 如果选择整个库作为同步对象,那么该库中所有对象的结构变更操作都会同步至目标库。 · 默认情况下,同步对象的名称保持不变。如果您需要改变同步对象在目标集群中的名称,那么需要使用DTS提供的对象名映射功能,详情请参见#unique 50。

9. 上述配置完成后,单击页面右下角的下一步。

10 配置同步初始化的高级配置信息。

创建同步作业	€ 返回数据同步列表					
	1.选择同步通道的源及目标实例	\rightarrow	2.选择同步对象	3.高级设置		4.预检查
	同步初始化: 🗹 结构初始化	☑ 全量数据初始化				
					取消 上一步	保存预检查并启动

同步初始化类型细分为:结构初始化,全量数据初始化。选择结构初始化和全量数据初始 化后,DTS会在增量数据同步之前,将源数据库中待同步对象的结构和存量数据,同步到目标 数据库。

11.上述配置完成后,单击页面右下角的预检查并启动。



- ・ 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同 步作业。
- ・如果预检查失败,单击具体检查项后的,查看失败详情。根据提示修复后,重新进行预

检查。

12.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

13等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面,查看数据同步作业的状态。

同步作业名称 🔻	搜索	排序: 默认排序 ▼ 状态	: 全部 ▼			
□ 实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼		操作
hangzhou-hangzhou-small	同步中	延时:565	按量付费	单向同步	暂停同步 转包年包月	升级 更多
1 11日 11日 11日 11日 11日 11日 11日 11日 11日 1				共有1条 ,每页显示	示:20条 《 < 1 >	> >>

1.3.2 从POLARDB MySQL同步至RDS MySQL

本文介绍如何使用数据传输服务DTS快速创建POLARDB MySQL集群和RDS MySQL实例间的实 时同步作业,实现POLARDB MySQL到RDS MySQL的增量数据实时同步。

前提条件

- · 已购买POLARDB MySQL实例,详情请参见创建POLARDB MySQL实例。
- · POLARDB MySQL实例已开启Binlog,详情请参见如何开启Binlog。

注意事项

- ・如果同步对象为单个或多个表(非整库),那么在数据同步时,请勿对源库的同步对象使用ghost或pt-online-schema-change等类似工具执行在线DDL变更,否则会导致同步失败。
- ・全量初始化过程中,并发INSERT会导致目标集群的表碎片,全量初始化完成后,目标集群的表 空间比源集群的表空间大。
- ・如果数据同步的源集群没有主键或唯一约束,且记录的全字段没有唯一性,可能会出现重复数 据。

支持同步的SQL操作

- DML: INSERT, UPDATE, DELETE
- DDL: CREATE TABLE, ALTER TABLE, RENAME TABLE, TRUNCATE TABLE, DROP TABLE

操作步骤

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdlo

说明:

购买时,选择源实例为POLARDB,选择目标实例为MySQL,并选择同步拓扑为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。

4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

乍业目标实例所在的地域)
常见问题 刷新
f

5. 定位至已购买的数据同步实例,单击配置同步链路。

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实例		2.选择同步对象	>	3.高级设置	>	4.预检查
同步作业名称:	POLARDB源					
源实例信息						
实例类型:	POLARDB					
实例地区:	华东1 (杭州)					
* POLARDB实例ID:	pc-		•			
* 数据库账号:	dtstest					
* 数据库密码:	•••••		₫>			
目标实例信息						
实例类型:	RDS实例		v			
实例地区:	华东1 (杭州)					
* 实例ID:	rm-		•			
* 数据库账号:	dtstest					
* 数据库密码:	•••••		₫>			
* 连接方式:	● 非加密连接 ○ SSL安全	2连接				
					取消	授权白名单并进入下一步

配置项目	配置选项	配置说明
任务名称	-	 DTS为每个任务自动生成一个任务名称,任务名称没有唯一性要求。 您可以根据需要修改任务名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。
源实例信息	实例类型	选择POLARDB。
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。
	实例ID	选择作为数据同步源的POLARDB实例ID。
	数据库账号	填入源POLARDB实例的数据库账号。
	数据库密码	填入数据库账号对应的密码。
目标实例信息	实例类型	选择为RDS实例。

配置项目	配置选项	配置说明
	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。
	数据库账号	填入目标RDS实例的数据库账号。
	数据库密码	填入数据库账号对应的密码。
) 说明: 当目标RDS实例的数据库类型为MySQL 5.5或MySQL 5.6时,无需配置数据库账号和数据库密码。
连接方式	连接方式	根据需求选择为非加密连接或SSL安全连接,本案例选择为非 加密连接。
		道 说明: 选择 SSL安全连接时,需要提前开启RDS实例的SSL加密功 能,详情请参见 <mark>设置</mark> SSL加密。

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到源POLARDB实例和目标RDS实例的白名单中,用 于保障DTS服务器能够正常连接源和目标实例。 8. 配置目标已存在表的处理模式和同步对象。

1.选择同步通道的源及目标实例 2.选择同步对象		3.高级设置	\rightarrow	4.预检查
同步架构: 单向同步				
目标已存在表的处理模式: 💿 预检查并报错拦截 🔍 无操作				
源库对象		已选择对象(鼠标移到对象行,点击领	扁嶺可修改对象名或过》	<mark>条件)详情</mark> 点我
若全局搜索,请先履开附 Q			Q	
🖃 🧧 dtstestdata		🍙 dtstestdata (1个对象)		
□ <mark>●</mark> Tables		i order		
🗄 🥁 polardbdatabase				
	>			
	<			
全选		全选		
*1111111111111111111111111111111111111	主义分批导面功			
	~~ 中小山車丈以			
				取消 上一步 下一步

配置项目	配置说明
目标已存在表的处理 模式	 预检查并报错拦截:检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表,则通过该检查项目;如果目标数据库中有同名的表,则在预检查阶段提示错误,数据同步作业不会被启动。
	 说明: 如果目标库中同名的表不方便删除或重命名,您可以#unique_50来 避免表名冲突。
	· 无操作:跳过目标数据库中是否有同名表的检查项。
	 警告: 选择为无操作,可能导致数据不一致,给业务带来风险,例如: 表结构一致的情况下,如果在目标库遇到与源库主键的值相同的记录,在初始化阶段会保留目标库中的该条记录;在增量同步阶段则会覆盖目标库的该条记录。 表结构不一致的情况下,可能会导致无法初始化数据、只能同步部分列的数据或同步失败。

配置项目	配置说明
选择同步对象	在源库对象框中单击待同步的对象,然后单击 > 将其移动至已选择
	对象框。
	同步对象的选择粒度为库、表。
	道 说明:
	·如果选择整个库作为同步对象,那么该库中所有对象的结构变更 操作都会同步至目标库。
	 . 默认情况下,同步对象的名称保持不变。如果您需要改变同步对 象在目标集群中的名称,那么需要使用DTS提供的对象名映射功 能,详情请参见#unique_50。

9. 上述配置完成后,单击页面右下角的下一步。

10 配置同步初始化的高级配置信息。

创建同步作业	t 返回数据同步列表					
	1.选择同步通道的源及目标实例	>	2.选择同步对象	3.高级设置		4.预检查
	同步初始化: 🗹 集构初始化	☑ 全量数据初始化				
					取消 上一步	保存预检查并启动

同步初始化类型细分为:结构初始化,全量数据初始化。选择结构初始化和全量数据初始 化后,DTS会在增量数据同步之前,将源数据库中待同步对象的结构和存量数据,同步到目标 数据库。

11.上述配置完成后,单击页面右下角的预检查并启动。

▋ 说明:

- · 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同步作业。
- ・如果预检查失败,单击具体检查项后的,查看失败详情。根据提示修复后,重新进行预

检查。

12.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

13等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在数据同步页面,查看数据同步作业的状态。

同步作业名称 ▼	搜索	排序: 默认排序 V 状	态: 全部 ▼		
□ 实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼	操作
hangzhou-hangzhou-small	同步中	延时:565 室秒 速度:0TPS(0.00MB/s)	按量付费	单向同步 暂停同步	转包年包月 日 升级 更多
■ 暂停同步 释放同步				共有1条 ,每页显示:20条 «	< 1 > »

1.3.3 从POLARDB MySQL同步至POLARDB MySQL

本文介绍如何使用数据传输服务DTS快速创建POLARDB MySQL集群和POLARDB MySQL集群 间的实时同步作业,实现POLARDB MySQL到POLARDB MySQL的增量数据实时同步。

前提条件

- · 已购买源和目标POLARDB MySQL集群,详情请参见创建POLARDB MySQL集群。
- ・源POLARDB MySQL集群已开启Binlog,详情请参见如何开启Binlog。

注意事项

- ・如果同步对象为单个或多个表(非整库),那么在数据同步时,请勿对源库的同步对象使用ghost或pt-online-schema-change等类似工具执行在线DDL变更,否则会导致同步失败。
- ・ 全量初始化过程中,并发INSERT会导致目标集群的表碎片,全量初始化完成后,目标集群的表 空间比源集群的表空间大。
- ・如果数据同步的源集群没有主键或唯一约束,且记录的全字段没有唯一性,可能会出现重复数 据。

支持同步的SQL操作

- DML: INSERT, UPDATE, DELETE
- DDL: CREATE TABLE、ALTER TABLE、RENAME TABLE、TRUNCATE TABLE、 DROP TABLE

功能限制

・不兼容触发器

同步对象为整个库且这个库中包含了会更新同步表内容的触发器,那么可能导致同步数据不一致。例如数据库中存在了两个表A和B。表A上有一个触发器,触发器内容为在INSERT一条

数据到表A之后,在表B中插入一条数据。这种情况在同步过程中,如果源集群表A上进行了 INSERT操作,则会导致表B在源集群跟目标集群数据不一致。

此类情况须要将目标集群中的对应触发器删除掉,表B的数据由源集群同步过去,详情请参见<u>触</u> 发器存在情况下如何配置同步作业。

・ RENAME TABLE限制

RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A,如果同步过程中源 集群将表A重命名为表B,那么表B将不会被同步到目标库。为避免该问题,您可以在数据同步 配置时,选择同步表A和表B所在的整个数据库作为同步对象。

操作步骤

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdlo

Ð	说明
---	----

购买时,选择源实例和目标实例均为POLARDB,并选择同步拓扑为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。
- 4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

数据传输	同步作业列表 华东1(杭州) 华东2 (上海)	华北1 (青岛)	华北2 (北京)	华南1 (深圳)	华北3 (张家口)	香港		
	美国 (硅	谷) 美国 (弗吉尼亚) 新加坡 阿爾	送酋 (迪拜) · 德	国 (法兰克福) 马	马来西亚 (吉隆坡)		(所选地域为同步作业目标实例	所在的地域)
概览	海十年山田	(※2) 印度(子菜) 苗園 ((公敦)	日本 (左南)	印度尼西亚 (強加)	()() ()()()()()()()()()()()()()()()()()	19年4年)		
*****==?~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~		(2005) PDSE (mm./	.) Secimi (16-6X)	Live (2000)	NUBERCHART (SEALH)		(H19)		
级以简定作为								(?) DTS常见问题	€ 刷新
数据订阅									
数据同步	同步作业名称 🔻			搜索排	·序: 默认排序	▼ 状态:	全部	Ŧ	

5. 定位至已购买的数据同步实例,单击配置同步链路。

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实	例	2.选择同步对象	\geq	3.高级设置	>	4.预检查
日半作业々争。						
PD221F3上中小:	POLARDB_TO_POLARDB					
源实例信息						
实例类型:	POLARDB					
实例地区:	华东1 (杭州)					
* POLARDB实例ID:	pc-bp		•			
* 数据库账号:	dtstest					
* 数据库密码:	•••••		\$ >			
目标实例信息						
实例类型:	POLARDB					
实例地区:	华东1 (杭州)					
* POLARDB实例ID:	pc-bp:		•			
* 数据库账号:	dtstest					
* 数据库密码:	•••••		()			
					取消	授权白名单并进入下——步

配置项目	配置选项	配置说明				
同步作业名 称	-	 DTS为每个任务自动生成一个同步作业名称,该名称没有唯一 性要求。 您可以根据需要修改同步作业名称,建议配置具有业务意义的 名称,便于后续识别。 				
源实例信息	实例类型	固定为POLARDB,不可变更。				
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。				
	POLARDB 实例ID	选择源POLARDB集群ID。				
数据库账号		填入连接POLARDB集群的数据库账号。				
		道 说明: 该账号需具备待同步对象的SELECT、REPLICATION CLIENT、REPLICATION SLAVE权限。				
	数据库密码	填入数据库账号对应的密码。				
目标实例信	实例类型	固定为POLARDB,不可变更。				
息	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。				
	POLARDB 实例ID	选择目标POLARDB集群ID。				

配置项目	配置选项	配置说明
	数据库账号	填入连接POLARDB集群的数据库账号。
		间 说明: 用于数据同步的数据库账号需具备目标同步对象的ALL权限。
	数据库密码	填入数据库账号对应的密码。

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到源和目标POLARDB集群的白名单中,用于保 障DTS服务器能够正常连接源和目标集群。 8. 配置目标已存在表的处理模式和同步对象。

1.选择同步通道的源及目标实例 2.选择	洞步对象	3.高级设置	\rightarrow	4.预检查
同步架构: 单向同步				
目标已存在表的处理模式: 💿 预检查并报错拦截 🔘 🤊	石操作			
源库对象		已选择对象(鼠标移到对象行,点击编	辑可修改对象名或过法	1条件)详情 点我
若全局搜索,请先展开树 Q			Q	
🖃 늘 dtstestdata		🦢 dtstestdata (1个对象)		
Tables		i order		
📰 vip 관 🚰 polardbdatabase				
	>			
	<			
全选		全选		
		13		
*映射名称更改:	認 ◎ 要进行库表名称批量更改			
				取消 上一步 下一步

配置项目	配置说明
目标已存在表的处理 模式	 预检查并报错拦截:检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表,则通过该检查项目;如果目标数据库中有同名的表,则在预检查阶段提示错误,数据同步作业不会被启动。
	说明: 如果目标库中同名的表不方便删除或重命名,您可以#unique_50来 避免表名冲突。
	・无操作:跳过目标数据库中是否有同名表的检查项。
	
	步阶段则会覆盖目标库的该条记录。 - 表结构不一致的情况下,可能会导致无法初始化数据、只能同 步部分列的数据或同步失败。

配置项目	配置说明
选择同步对象	在源库对象框中单击待同步的对象,然后单击 > 将其移动至已选择
	对象框。
	同步对象的选择粒度为库、表。
	道 说明:
	·如果选择整个库作为同步对象,那么该库中所有对象的结构变更 操作都会同步至目标库。
	 . 默认情况下,同步对象的名称保持不变。如果您需要改变同步对 象在目标集群中的名称,那么需要使用DTS提供的对象名映射功 能,详情请参见#unique_50。

9. 上述配置完成后,单击页面右下角的下一步。

10 配置同步初始化的高级配置信息。

创建同步作业	t 返回数据同步列表					
	1.选择同步通道的源及目标实例	>	2.选择同步对象	3.高级设置		4.预检查
	同步初始化: 🗹 集构初始化	☑ 全量数据初始化				
					取消 上一步	保存预检查并启动

- 说明:

同步初始化类型细分为:结构初始化,全量数据初始化。选择结构初始化和全量数据初始 化后,DTS会在增量数据同步之前,将源数据库中待同步对象的结构和存量数据,同步到目标 数据库。

11.上述配置完成后,单击页面右下角的预检查并启动。

▋ 说明:

- · 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同步作业。
- ・如果预检查失败,单击具体检查项后的,查看失败详情。根据提示修复后,重新进行预

检查。

12.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

13等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面,查看数据同步作业的状态。

同步作业名称 ▼	搜索	排序: 默认排序 V 状:	态: 全部 ▼		
□ 实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼	操作
hangzhou-hangzhou-small	同步中	延时:565	按量付费	单向同步 暫停同步	转包年包月 升级 更多
1				共有1条 ,每页显示:20条 。	$\langle 1 \rangle \gg$

1.3.4 从POLARDB MySQL同步至AnalyticDB for MySQL

分析型数据库MySQL版(AnalyticDB for MySQL),是阿里巴巴自主研发的海量数据实时高 并发在线分析(Realtime OLAP)云计算服务,可以对千亿级数据进行毫秒级的即时多维分析透 视和业务探索。通过数据传输服务DTS(Data Transmission Service),您可以将POLARDB MySQL同步到AnalyticDB for MySQL,帮助您快速构建企业内部BI、交互查询、实时报表等系 统。

前提条件

・已创建目标AnalyticDB MySQL实例,详情请参见创建AnalyticDB MySQL (2.0) 或创建 AnalyticDB MySQL (3.0)。



不支持将华北1(青岛)、美国东部1(弗吉尼亚)、英国(伦敦)地域的AnalyticDB MySQL(2.0)实例作为同步的目标实例;不支持将美国西部1(硅谷)地域的AnalyticDB MySQL(3.0)实例作为同步的目标实例。

- ・确保目标AnalyticDB MySQL具备充足的存储空间。
- · POLARDB MySQL已开启Binlog,详情请参见如何开启Binlog。
- 如果同步的目标为AnalyticDB for MySQL(2.0),那么源RDS for MySQL待同步的对象不能包含AnalyticDB for MySQL(2.0)保留的库名和列名,否则将造成数据同步失败或DDL操作同步失败。

注意事项

- · 请勿在数据同步时,对源库的同步对象使用gh-ost或pt-online-schema-change等类似工具 执行在线DDL变更,否则会导致同步失败。
- 由于AnalyticDB MySQL(3.0)本身的使用限制,当AnalyticDB MySQL(3.0)实例中的
 节点磁盘空间使用量超过80%,该实例将被锁定。请提前根据待同步的对象预估所需空间,确
 保目标实例具备充足的存储空间。

术语/概念对应关系

POLARDB MySQL	AnalyticDB for MySQL
数据库	 AnalyticDB for MySQL (2.0) : 表组 AnalyticDB for MySQL (3.0) : 数据库
表	 AnalyticDB for MySQL (2.0) : 表 AnalyticDB for MySQL (3.0) : 表

支持同步的SQL操作

目标数据库版本	支持的SQL操作
AnalyticDB MySQL 2.0	・ DDL操作:ADD COLUMN ・ DML操作:INSERT、UPDATE、DELETE
AnalyticDB MySQL 3.0	・ DDL操作:CREATE TABLE、DROP TABLE、RENAME TABLE 、TRUNCATE TABLE、ADD COLUMN、DROP COLUMN ・ DML操作:INSERT、UPDATE、DELETE
	 说明: 如果在数据同步的过程中变更了源表的字段类型,同步作业将报错并 中断。您可以提交工单处理或参照文末的方法来手动修复,详情请参见#unique_17/unique_17_Connect_42_d7e829。

数据库账号的权限要求

数据库	所需权限
POLARDB MySQL	待同步对象的读权限。
AnalyticDB for MySQL (2.0)	无需填写数据库账号信息,DTS会自动创建账 号并授权。
AnalyticDB for MySQL (3.0)	读写权限。

数据库账号的创建和授权方法,请参见创建POLARDB MySQL数据库账号或创建AnalyticDB for

MySQL数据库账号。

数据类型映射关系

详情请参见#unique_51。

操作步骤

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdlo

📋 说明:

购买时,选择源实例为POLARDB、目标实例为分析型数据库AnalyticDB,并选择同步拓扑 为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。
- 4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

数据传输	□ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □ □
	美国(桂谷) 美国(弗吉尼亚) 新加坡 阿联酋(迪拜) 德国(法兰克福) 马来西亚(吉隆坡) (所选地域为同步作业目标实例所在的地域
概览	演大利亚(悉尼) 印度(孟买) 英国(伦敦) 日本(东京) 印度尼西亚(雅加达) 华北5(呼和浩特)
数据迁移	⑦ DTS常见问题
数据订阅	
数据同步	同步作业名称 ▼

5. 定位至已购买的数据同步实例,单击配置同步链路。

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实	例 2	ADS账号授权	3.选择同步对象	>	4.预检查
同步作业名称:	POLARDB_TO_AnalyticDB for My	SQL			
源实例信息					
500米刑,					
文 内决至,					
头彻地区:	华东1(杭州)				
* POLARDB实例ID:	pc-bp	•			
* 数据库账号:	dtstest				
* 数据库密码:	******	4>			
目标实例信息					
实例类型:	ADS				
实例地区:	华东1 (杭州)				
*版本:	◎ 2.0 ◎ 3.0				
* 数据库:	am-bp:	•			
* 数据库账号:	dtstest				
* 数据库密码:	•••••	4>			
				F	(1) (1) (1) (1) (1) (1) (1) (1) (1) (1)

配置项目	配置选项	配置说明
同步作业名称	-	 DTS为每个同步作业自动生成一个名称,该名称没有唯一 性要求。 您可以根据需要修改同步作业的名称,建议配置具有业务 意义的名称,便于后续的识别。
源实例信息	实例类型	固定为POLARDB,不可变更。
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。
	POLARDB实 例ID	选择源POLARDB集群ID。
	数据库账号	填入POLARDB集群的数据库账号。
	数据库密码	填入数据库账号对应的密码。
目标实例信息	实例类型	固定为ADS,不可变更。
	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。

配置项目	配置选项	配置说明
	版本	根据目标AnalyticDB for MySQL实例的版本,选 择2.0或3.0。
		道 说明:
		 ・选择为2.0后,DTS将自动创建数据库账号并进行授 权,您无需配置数据库账号和数据库密码。
		・选择为3.0后,您还需要配置数据库账号和数据库密码。
	数据库	选择作为数据同步目标的AnalyticDB for MySQL实例ID。
	数据库账号	填入AnalyticDB for MySQL的数据库账号。
	数据库密码	填入数据库账号对应的密码。

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到POLARDB MySQL和AnalyticDB for MySQL的 白名单中,用于保障DTS服务器能够正常连接源和目标实例。

8. 配置同步策略及对象信息。

1.选择同步通道的源及目标实例 > 2.ADS账号授权		3.选择同步对象		4.预检查
同歩初始化: ● 结构初始化 ● 全量数据 日际已存在表的处理概 正 ● 預除室井报错拦截 ● 无缘作 予 潮明并: ● 夏 ● 否 通 か聞作姓型: ● Insert ● Update ● ● Drop Tat 源库対象 ■ atsestdata ● Tables ■ atsestdata ● Tables ■ name ■ vip ● dtstestdata ● Tables ■ name ■ vip ● dtstestdata ● Tables ■ name ■ vip	初始化 Delete 《Alter Table le 《Create Database	 ♥ Truncate Table ♥ Drop Database 已选择对象(鼠标移到对象行,点击都情点我 □ dtstestdata (1个对象) □ order 	- 現在 日本 日本 日本 日本 日本 日本 日本 日本 日本 日本	
		全选		
*映射名称更改: 不进行库表名称批量更改 	◎ 要进行库表名称批量更	έ¢		
			取消 上-	-步 下步

配置项目	配置说明
同步初始化	默认情况下,您需要同时勾选结构初始化和全量数据初始化。预检查完成 后,DTS会将源实例中待同步对象的结构及数据在目标实例中初始化,作 为后续增量同步数据的基线数据。
目标已存在表的处 理模式	 · 预检查并报错拦截:检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表,则通过该检查项目;如果目标数据库中有同名的表,则在预检查阶段提示错误,数据同步作业不会被启动。
	 说明: 如果目标库中同名的表不方便删除或重命名,您可以#unique_50来避免表名冲突。
	・无操作:跳过目标数据库中是否有同名表的检查项。
	▲ 警告: 选择为无操作,可能导致数据不一致,给业务带来风险,例如:
	 表结构一致的情况下,在目标库遇到与源库主键的值相同的记录,则会保留目标实例中的该条记录,即源库中的该条记录不会同步至目标数据库中。
	 表结构不一致的情况下,可能会导致无法初始化数据、只能同步部 分列的数据或同步失败。

配置项目	配置说明
多表归并	 ・ 选择为是:DTS将在每个表中增加dts_data_source列来存储数 据来源,且不再支持DDL同步。 ・ 选择为否:默认选项,支持DDL同步。
	 说明: 多表归并功能基于任务级别,即不支持基于表级别执行多表归并。如果 需要让部分表执行多表归并,另一部分不执行多表归并,您可以创建两 个数据同步作业。
同步操作类型	根据业务勾选需要同步的操作类型,默认情况下都处于勾选状态。
	道 说明: 目前仅支持INSERT、UPDATE、DELETE、ADD COLUMN。
选择同步对象	在源库对象框中单击待迁移的对象,然后单击 > 将其移动至已选择对
	象框。
	同步对象的选择粒度为库、表。
	道 说明:
	 ・如果选择整个库作为同步对象,那么该库中所有对象的结构变更操作 都会同步至目标库。
	・如果选择某个表作为同步对象,那么只有这个表的ADD COLUMN操 作会同步至目标库。
	 默认情况下,同步对象的名称保持不变。如果您需要同步对象在目标 实例上名称不同,那么需要使用DTS提供的对象名映射功能,详情请 参见#unique_50。

9. 上述配置完成后,单击页面右下角的下一步。

10.设置待同步的表在目标库中类型。

- 1.选择同步)	通道的源及目标实例	2.	ADS账号授权	3.选择同步对象	4.预检查		
ADB表组	ADB表名	类型(全部) ▼	主鏈列	分布列	定义状态(全部) 🕶		
dtstestdata	order	分区表 ▼	orderid	orderid v	已定义		
全部设置为分区表	全部设置为维度表	请输入ADB表名	搜索	共	有1条, 每页显示: 20条 《 〈 1 〉 》		
				取消	当 上一步 保存 预检查并启动		
j ii	〕 说明:						

选择了结构初始化后,您需要定义待同步的表在AnalyticDB MySQL中的类型、主键列、分区 列等信息,详情请参见ADB 2.0 SQL手册和ADB 3.0 SQL手册。

11.上述配置完成后,单击页面右下角的预检查并启动。



- ・ 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同 步作业。
- ・如果预检查失败,单击具体检查项后的,查看失败详情。根据提示修复后,重新进行预

检查。

12.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

13.等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面,查看数据同步作业的状态。

同步	作业名称 🔻	搜索	排序: 默认排序 ▼ 状态:	全部 ▼		
	实例ID/作业名称	状态	同步概况	付鶈方式	同步架构(全部) ▼	操作
	nangzhou-hangzhou-small	同步中	延时:565	按量付费	单向同步 暂停同步 转包年包月	升级 更多
	暂停同步 释放同步				共有1条 , 每页显示 : 20条 🛛 « 🧹 1	>

1.3.5 从POLARDB MySQL同步至AnalyticDB for PostgreSQL

分析型数据库PostgreSQL版(原HybridDB for PostgreSQL)为您提供简单、快速、经济高效的PB级云端数据仓库解决方案。通过数据传输服务DTS(Data Transmission Service),您可以将POLARDB MySQL同步至AnalyticDB for PostgreSQL,帮助您快速实现对海量数据的即席查询分析、ETL处理和可视化探索。

前提条件

- ・源POLARDB MySQL实例已开启Binlog,详情请参见如何开启Binlog。
- ・ 已购买目标AnalyticDB for PostgreSQL实例,详情请参见创建AnalyticDB for PostgreSQL实
 例。

注意事项

- ・全量初始化过程中,并发INSERT会导致目标实例的表碎片,全量初始化完成后,目标实例的表 空间比源集群的表空间大。
- ・如果数据同步的源实例没有主键或唯一约束,且记录的全字段没有唯一性,可能会出现重复数据。
・ 在同步的过程中,如果在源库中创建一个表,要将其作为同步对象,那么您需要为数据同步作 业#unique_52。

功能限制

- · 仅支持表级别的数据同步。
- ・不支持结构同步,详情请参见#unique_22。
- ・不支持同歩JSON、GEOMETRY、CURVE、SURFACE、MULTIPOINT、MULTILINES TRING、MULTIPOLYGON、GEOMETRYCOLLECTION类型的数据。

支持同步的SQL操作

- ・ DML操作: INSERT、UPDATE、DELETE。
- ・ DDL操作: ADD COLUMN、RENAME COLUMN。

॑ 説明:

不支持CREATE TABLE操作,如果您需要将新增的表作为同步对象,则需要执行#unique_52操作。

支持的同步架构

- ・1对1单向同步。
- ・1对多单向同步。
- ・多对1单向同步。

关于各类同步架构的介绍及注意事项,请参见#unique_48。

术语/概念对应关系

POLARDB MySQL	AnalyticDB for PostgreSQL
Database	Schema
Table	Table

操作步骤一 在目标实例中创建对应的数据结构

根据源RDS实例中待同步表的数据结构,在目标AnalyticDB for PostgreSQL实例中创建数据 库、Schema及数据表,详情请参考AnalyticDB for PostgreSQL基础操作。

操作步骤二 配置数据同步

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdlo



购买时,选择源实例为MySQL、目标实例为AnalyticDB for PostgreSQL,并选择同步拓扑 为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。
- 4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

数据传输	同步作业列表 华东1 (杭州)	华东2 (上海) 华北1 (青	33) 华北2 (北京)	华南1 (深圳) 华北	は3 (张家口) 香港		
4871145	美国 (硅谷)	美国 (弗吉尼亚) 新加坡	阿联酋 (迪拜) 領	専国 (法兰克福) 马来西	西亚 (吉隆坡)	(所选地域为同步作业目标实例	所在的地域)
191.03	演大利亚 (悉尼) 印度 (孟买) 英国 (论敦) 日本 (东京)	印度尼西亚 (雅加达)	华北5 (呼和浩特)		
数据迁移							€ 刷新
数据订阅						<u> </u>	
数据同步	同步作业名称 🔻		搜索	排序: 默认排序	▼ 状态: 全部	Ŧ	

- 5. 定位至已购买的数据同步实例,单击配置同步链路。
- 6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道	的源及目标实例	2.选择同步对象	>	3.预检查
同步作业?	称: POLARDB数据同步至AnalyticDB for Postgres	5QL		
源实例信息				
实例到	型: 通过专线/VPN网关/智能网关接入的自建数据	ka v		
实例均	区: 华东1(杭州)			
* 对端专有网	络:	-		
数据库纳	型: MySQL			
* IP#	址: 172. 36			
* 10	□: 3306			
* 数据库则	号:			
* 数据库容	码:	<i>ه</i> >		
目标实例信息				
实例类	型: AnalyticDB for PostgreSQL			
实例均	区: 华东1(杭州)			
* 实例	ID :	-		
* 数据库名	称: polardbdatabase			
* 数据库则	·号:			
* 数据库图	码: ••••••	4>		

类别	配置	说明
任务名称	-	 DTS为每个任务自动生成一个任务名称,任务名称没有唯一性要求。 您可以根据需要修改任务名称,建议为任务配置具有业务意义的名称,便于后续的任务识别。

类别	配置	说明					
源实例信息	实例类型	选择通过专线/VPN网关/智能网关接入的自建数据库。					
	实例地区	购买数据同步实例时选择的源POLARDB实例的地域信息,不可 变更。					
	对端专有网 络	选择POLARDB实例所属的VPCID。 您可以登录POLARDB控制合,单击目标实例ID,进入该实例 的基本信息页面来获取。					
	数据库类型	固定为MySQL,不可变更。					
	IP地址	配置POLARDB主实例的私网IP地址,本案例中填 入172.16.20.36。 您可以在电脑中ping目标POLARDB实例的主地址(私网)来获 取私网IP地址。					
	端口	填入POLARDB实例的服务端口,默认为3306。					
	数据库账号	填入连接POLARDB实例的数据库账号。 说明: 该账号需具备待同步对象的SELECT、REPLICATION CLIENT、REPLICATION SLAVE权限。					
	数据库密码	填入POLARDB数据库账号对应的密码。					
目标实例信	实例类型	固定为AnalyticDB for PostgreSQL,无需设置。					
尽 	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。					
	实例ID	选择作为数据同步目标的AnalyticDB for PostgreSQL实例ID。					

类别	配置	说明
	数据库名称	填入目标AnalyticDB for PostgreSQL实例中,待同步的目标 表所属的数据库名称。
	数据库账号	填入目标AnalyticDB for PostgreSQL实例的数据库账号。
		自己 说明: 用于数据同步的数据库账号需具备目标同步对象的ALL权限。
	数据库密码	填入数据库账号对应的密码。

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加POLARDB MySQL和AnalyticDB for PostgreSQL的白名单中,用于保障DTS服务器能够正常连接源和目标实例。

8. 配置同步策略及对象信息。

1.选择同步通道的源及目标实例	2.选择同步对象	3.预益	
1.送評局法施設的原及目标实例 同步初始化: ☑ 全量数据初始化 目标已存在表的处理模式: ◎ 預检童并报错拦截 ◎ 漂: 同步操作类型: ☑ Insert ☑ Update 源库对象 ② ◎ polardbdatabase ③ Tables	2.选择同步对象 空目标表数据 ◎ 无操作 I Delete I Alter Table	3.预检查 已选择对象 (魚标移到对象行,点击编辑可修改对象名或过谋条件) 详 情点我	鼠标移到对象上,点击 编辑入口,即可配置源 跟目标实例的对象名缺 射及迁移列选择
全选			
*缺射名称更改: 不进行库表名称批量更 	攻 ◎ 要进行库表名称批量更		
		取消上一	步 预检查并启动

类别项目	选项	说明
同步策略配 置	同步初始化	选择全量数据初始化。 道 说明: 将源实例中已经存在同步对象的数据在目标实例中初 始化,作为后续增量同步数据的基线数据。

类别项目	选项	说明
	目标已存在表的处理模式	 预检查检测并拦截(默认勾选) 在预检查阶段执行目标表是否为空的检查项目,如 果有数据直接在预检查的目标表是否为空的检查项 中检测并拦截报错。 清空目标表的数据 在预检查阶段跳过目标表是否为空的检查项目。全 量初始化之前将目标表的数据清空。适用于完成同 步任务测试后的正式同步场景。 不做任何操作 在预检查阶段跳过目标表是否为空的检查项目。全 量初始化时直接追加同步数据。适用于多张表同步
	同步操作类型	到一张表的汇息同步场景。 根据业务需求选择需要同步的操作类型: · Insert · Update · Delete · Alter Table
选择同步对 象	-	在源库对象框中单击待同步的表,然后单击 将其 移动至已选择对象框。 ① 说明: · 同步对象的选择粒度为表。 · 如果需要目标表中列信息与源表不同,则需要使 用DTS的字段映射功能,详情请参见#unique_50。

9. 上述配置完成后,单击页面右下角的预检查并启动。

・ 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同步作业。

·如果预检查失败,单击具体检查项后的____,查看失败详情。根据提示修复后,重新进行预

检查。

10.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

11.等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面,查看数据同步作业的状态。

同步作业名称 ▼	搜索	排序: 默认排序 🔻 状态	≅: 全部 ▼		
□ 实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼	操作
hangzhou-hangzhou-small	同步中	延时:565 毫秒 速度:0TPS(0.00MB/s)	按量付费	单向同步 暂停同步 转	抱年包月 升级 更多
■ 暂停同步 释放同步				共有1条 ,每页显示:20条 《 《	: 1 → ≫

1.3.6 从POLARDB MySQL同步至Datahub

阿里云流式数据服务DataHub是流式数据(Streaming Data)的处理平台,提供对流式数据的 发布、订阅和分发功能,让您可以轻松构建基于流式数据的分析和应用。通过数据传输服务(Data Transmission Service,简称DTS),您可以将POLARDB MySQL同步至DataHub,帮助您 快速实现使用流计算等大数据产品对数据实时分析。

前提条件

- · DataHub实例的地域为华东1、华东2、华北2或华南1。
- · DataHub实例中,已创建用作接收同步数据的Project,详情请参见创建Project。
- · 已开启POLARDB MySQL的Binlog,详情请参见如何开启Binlog。
- · POLARDB MySQL中待同步的表需具备主键或唯一约束。

功能限制

- ・不支持全量数据初始化、即DTS不会将源RDS实例中同步对象的存量数据同步至目标DataHub 实例中。
- ・仅支持表级别的数据同步。
- ・不支持新增列的数据同步,即源数据表新增了某个列,该列的数据不会同步至目标DataHub实 例中。
- ・数据同步的过程中,请勿对源库中待同步的表执行DDL变更,否则会导致同步失败。

支持同步的SQL操作

INSERT, UPDATE, DELETE.

操作步骤

1. 购买数据同步作业,详情请参见#unique_49。

说明:

购买时,选择源实例为POLARDB、目标实例为DataHub,并选择同步拓扑为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。
- 4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

数据传输	同步作业列表	华东1 (杭州)	华东2 (上海)	华北1 (青岛)	华北2 (北京)	华南1 (深圳)	华北3 (张家口)	香港		
477114-		美国 (硅谷)	美国 (弗吉尼亚)	新加坡 阿耳	联酋 (迪拜) 領	萬国 (法兰克福)	马来西亚 (吉隆坡)		(所选地域为同步作业目标实例	所在的地域)
備記		演大利亚 (悉尼	(孟买)	英国 (伦敦)	日本 (东京)	印度尼西亚 (雅)	加达) 华北5 (呼利	口浩特)		
数据迁移										€ 刷新
数据订阅										
数据同步	同步作业名称 🔻				搜索	非序: 默认排序	▼ 状态:	全部	v	

- 5. 定位至已购买的数据同步实例,单击配置同步链路。
- 6. 配置同步作业的源实例及目标实例信息。

1.选择同步通道的源	双目标实例 人名法尔尔 人名法尔尔 人名法尔尔 人名法尔尔 人名法尔尔 化合成合金 化合成合金 人名法尔尔 化合成合金 人名法尔尔尔 人名法尔尔尔 化合成合金 人名法尔尔尔 化合成合金 人名法尔尔尔 化合成合金 人名法尔尔尔 化合成合金 人名法尔尔尔尔 化合成合金 人名法尔尔尔尔 化合成合金 人名法尔尔尔尔 化合成合金 人名法尔尔尔尔尔 化合成合金 人名法尔尔尔尔尔尔 化合成合金 人名法尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔尔	2.选择同步对象	>	3.预检查
同步作业名称	POLARDB for MySQL_TO_DataHub			
酒空周注自				
「「赤矢の」」日本				
实例类型	: POLARDB			
实例地区	: 华东1 (杭州)			
* POLARDB实例IE	pc-bp	•		
* 数据库账号	dtstest			
* 数据库密码	:	<i>٩</i> >		
目标实例信息				
实例类型	: DataHub			
实例地区	: 华东1 (杭州)			
* Projec	dtstestdata	•		
				取消 授权白名单并进入下一步

配置项目	配置选项	配置说明
同步作业名称	-	DTS会自动生成一个同步作业名称,建议配置具有业务意义 的名称(无唯一性要求),便于后续识别。
源实例信息	实例类型	固定为POLARDB,不可变更。
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。

配置项目	配置选项	配置说明
	POLARDB实 例ID	选择源POLARDB集群ID。
	数据库账号	填入POLARDB集群的数据库账号。
	数据库密码	填入数据库账号对应的密码。
目标实例信息	实例类型	固定为DataHub,不可变更。
	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。
	Project	选择DataHub实例的Project。

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到POLARDB集群的白名单中,用于保障DTS服务器 能够正常连接POLARDB集群。

8. 配置同步策略和同步对象。

1.选择同步通道的源及目标实例	X	2.选择同步对象		3.预检查
同步初始化: 🗹 结构初	刀始化			
源库对象			已选择对象(鼠标移到对象行,点击编辑可修改对象名或运信点我)	赵滤条件)详
若全局搜索,请先展开树 ♥ ∰ sys ■ dtstestdata ■ Mables		> <	Q dtstestdata (2个对象) customer order	
全选			全洗	
* 映射名称更改: *是否启用新的附加列规则	 不进行库表名称批量更改 是 ● 否 	要进行库:	表名称批量更改	
			取消	上一步 预检查并启动

配置	说明
同步初始化	勾选结构初始化。
	说明: 勾选结构初始化后,在数据同步作业的初始化阶段,DTS会将同步对象的结构信息(例如表结构)同步至目标DataHub实例。
选择同步对象	在源库对象框中单击待迁移的对象,然后单击 > 将其移动至已选择对
	象框。
	 送明: · 同步对象的选择粒度为表。 · 默认情况下,同步对象的名称保持不变。如果您需要改变同步对象在目标实例中的名称,需要使用DTS提供的对象名映射功能,详情请参
	见 #unique_50。

配置	说明
选择附加列规则	DTS在将数据同步到DataHub时,会在同步的目标Topic中添加一些附 加列。如果附加列和目标Topic中已有的列出现名称冲突将会导致数据同 步失败。您需要根据业务需求选择是否启用新的附加列规则为是或否。
	▲ 警告: 在选择附加列规则前,您需要评估附加列和目标表中已有的列是否 会出现名称冲突。关于附加列的规则和定义说明,请参见#unique_53/ unique_53_Connect_42_table_zyo_yb6_p5y。

9. (可选)将鼠标指针放置在已选择对象框中待同步的表名上,单击对象后出现的编辑,然后在 弹出的对话框中设置Shardkey(即用于分区的key)。

	Q 搜索文档、控制台、API、解决方案和资	源 费 用	工単	备案 企业	支持与服务	官网 🖸	Ċ.	Ä	?	简体中文	
1.选择同	步通道的源及目标实例	2.选择同步对象				3	3.预检查				
	同步初始化: ✓ 结构初始化 源库对象		已选择对象 dtstee custor order	<mark>(鼠标移到对象行</mark> stdata (2个对象) mer	,点击编辑可修改对	\$2ىتىنىچ م	件)详情点	親	鼠标移至1, 場入口, 小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小小	对象上, 点击 即可配置源跟 対象名映射万	编目
	<pre></pre>	> <					6				
	全选		全选								
	缺射名称更改:	r批量更改 💿 要进行库表名 请选择附加列规则	称批量更改								

10.上述配置完成后,单击页面右下角的预检查并启动。

说明:

- ・ 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同 步作业。
- ·如果预检查失败,单击具体检查项后的____,查看失败详情。根据提示修复后,重新进行预

检查。

11.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

12.等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面,查看数据同步作业的状态。

同步作业名称 ▼	搜索	排序: 默认排序 ▼ 状:	∑: 全部 ▼		
□ 实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼	操作
hangzhou-hangzhou-small	同步中	延时:565 毫秒 速度:0TPS(0.00MB/s)	按量付费	单向同步 暫停同步	转包年包月 升级 更多
■ 暫停同步 释放同步				共有1条 ,每页显示:20条 «	< 1 > »

Topic结构定义说明

DTS在将数据变更同步至DataHub实例的Topic时,目标Topic中除了存储变更数据外,还会新增一些附加列用于存储元信息,示例如下。



本案例中的业务字段为id、name、address,由于在配置数据同步时选用的是旧版附加列规

则, DTS会为业务字段添加dts_的前缀。

dts_id	dts_name	dts_address	dts_record_id	dts_operation_flag	dts_instance_id	dts_db_name	dts_table_name	dts_utc_timestamp	dts_before_flag	dts_after_flag
10006	- 1810 -	曲靖市	1574832130000000000	U		dtstestdata	customer	1574832130	Y	N
10006	- 19 10	杭州市	1574832130000000000	U		dtstestdata	customer	1574832130	N	Y
10009	382	马鞍山市	1574832919000000000	D		dtstestdata	customer	1574832919	Y	N
10112	100	北京市	1574832919000000000	I		dtstestdata	customer	1574832919	N	Y

结构定义说明:

旧版附加列名称	新版附加列名称	说明
dts_record_id	new_dts_sync_dts_rec	增量日志的记录id,为该日志唯一标识。
	ord_id	 说明: · id的值唯一且递增。 · 如果增量日志的操作类型 为UPDATE,那么增量更新会被拆 分成两条记录,一条为DELETE,一 条为INSERT,并且这两条记录的 dts_record_id的值相同。
dts operat	new dts sync dts ope	操作类型、取值:
ion_flag	ration_flag	・ I:INSERT操作。
		・ D:DELETE操作。
		・ U: UPDATE操作。
dts_instan	new_dts_sync_dts_ins	数据库的server id。
ce_id	tance_id	

旧版附加列名称	新版附加列名称	说明
dts_db_name	new_dts_sync_dts_db_ name	数据库名称。
dts_table_ name	new_dts_sync_dts_tab le_name	表名。
dts_utc_ti mestamp	new_dts_sync_dts_utc _timestamp	操作时间戳,即binlog的时间戳(UTC 时 间)。
dts_before _flag	new_dts_sync_dts_bef ore_flag	所有列的值是否更新前的值,取值:Y或N。
dts_after_ flag	new_dts_sync_dts_aft er_flag	所有列的值是否更新后的值,取值:Y或N。

关于dts_before_flag和dts_after_flag的补充说明

对于不同的操作类型,增量日志中的dts_before_flag和dts_after_flag定义如下:

• INSERT

当操作类型为INSERT时,所有列的值为新插入的值,即为更新后的值,所以dts_before _flag取值为N, dts_after_flag取值为Y, 示例如下。

dts_id dts_name dts_address dts_record_id dts_operation_flag dts_instance_id dts_db_name dts_table_name dts_utc_timestamp dts_before_flag dts_after_flag
10112 北京市 1574832919000000000 I dtstestdata customer 1574832919 N Y

• UPDATE

当操作类型为UPDATE时,DTS会将UPDATE操作拆为两条增量日志。这两条增量日志的dts_record_id、dts_operation_flag及dts_utc_timestamp对应的值相同。

第一条增量日志记录了更新前的值,所以dts_before_flag取值为Y,dts_after_flag取 值为N。第二条增量日志记录了更新后的值,所以dts_before_flag取值为N,dts_after_ flag取值为Y,示例如下。



• DELETE

当操作类型为DELETE时,增量日志中所有的列值为被删除的值,即列值不变,所以 dts_before_flag取值为Y, dts_after_flag取值为N,示例如下。

 dts_iname
 dts_address
 dts_record_id
 dts_operation_flag
 dts_instance_id
 dts_dts_table_name
 dts_utc_timestamp
 dts_before_flag
 dts_after_flag

 10009
 马鞍山市
 1574832919000000000
 D
 dtstestdata
 customer
 1574832919
 Y
 N

后续操作

配置完数据同步作业后,您可以对同步到DataHub实例中的数据执行计算分析。更多详情,请参见阿里云实时计算。

1.3.7 从POLARDB MySQL同步到Kafka

Kafka是应用较为广泛的分布式、高吞吐量、高可扩展性消息队列服务,普遍用于日志收集、监控数据聚合、流式数据处理、在线和离线分析等大数据领域,是大数据生态中不可或缺的产品之一。 通过数据传输服务DTS(Data Transmission Service),您可以将POLARDB MySQL同步至自建Kafka集群,扩展消息处理能力。

前提条件

- ・Kafka集群的版本为0.10或1.0版本。
- · Kafka集群已创建用于存储待同步数据的Topic。
- · POLARDB MySQL已开启Binlog,详情请参见如何开启Binlog。

注意事项

如果源数据库没有主键或唯一约束,且所有字段没有唯一性,可能会导致目标数据库中出现重复数 据。

功能限制

- ・仅支持表粒度的数据同步。
- ・不支持DDL操作的数据同步。
- ・不支持自动调整同步对象。

如果在同步的过程中,对源库中待同步的表执行了重命名操作,且重命名后的名称不在同步对象中,那么该表将不再被同步到目标Kafka集群中。如果该表还需要同步,那么您需要#unique_52。

支持同步的SQL操作

DML操作: INSERT、UPDATE、DELETE、REPLACE。

消息格式

同步到Kafka集群中的数据以avro格式存储, schema定义详情请参见DTS avro schema定义。

在数据同步到Kafka集群后,您需要根据avro schema定义进行数据解析。

费用说明

详情请参见#unique_23。

操作步骤

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdlo



购买时,选择源实例为POLARDB、目标实例为Kafka,并选择同步拓扑为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。
- 4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

数据传输	同步作业列表	华东1 (杭州)	华东2 (上海)	华北1 (青岛)	华北2 (北京)	华南1 (深圳)	华北3 (张家口)	香港		1942.110
100110		美国 (硅谷)	美国 (弗吉尼亚)	新加坡阿	联酋 (迪拜) 待	国 (法兰克福)	马来西亚 (吉隆坡)		(所选地域为同步作业目标实例	所在的地域)
概党		演大利亚 (悉尼)	印度 (孟买)	英国 (伦敦)	日本 (东京)	印度尼西亚 (雅	加达) 华北5 (呼和]浩特)		
数据迁移									⑦ DTS常见问题	€ 刷新
数据订阅									.	
数据同步	同步作业名称 🔻				搜索 抖	i序: 默认排序	, ▼ 状态:	全部	Ŧ	

5. 定位至已购买的数据同步实例,单击配置同步链路。

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实	例 >	2.选择同步对象	3.	高级设置	\rangle	4.预检查
问步作业者称:	POLARDB_TO_Kafka					
源实例信息						
实例类型:	POLARDB					
实例地区:	华东1 (杭州)					
* POLARDB实例ID:	pc-bp	-				
* 数据库账号:	dtstest					
* 数据库密码:	•••••	4 >				
目标实例信息						
实例类型:	ECS上的自建数据库	Ŧ				
实例地区:	华东1 (杭州)					
* ECS实例ID:	i-bp	•				
数据库类型:	Kafka	Ŧ				
* 端口:	9092					
数据库账号:	dtstest		非必填项			
数据库密码:	•••••	\$ >	非必填项			
* Topic:	customer	v	获取Topic列表			
	请先点击右侧按钮,获取Top	pic列表后选择具体的Topic				
* Kafka版本	1.0	T				
					取消	授权白名单并进入下一步

配置项目	配置选项	配置说明
同步作业名称	-	 DTS为每个数据同步作业自动生成一个名称,该名称没有 唯一性要求。 您可以根据需要修改同步作业名称,建议配置具有业务意 义的名称,便于后续的任务识别。
源实例信息	实例类型	选择POLARDB。
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。
	POLARDB 实例ID	选择POLARDB集群ID。
	数据库账号	填入POLARDB集群的数据库账号,需要具备Replication slave、Replication client及所有待同步对象的Select权 限。
	数据库密码	填入该账号对应的密码。

配置项目	配置选项	配置说明
目标实例信息	实例类型	 ・Kafka集群部署在ECS上时,选择ECS上的自建数据库 ・Kafka集群部署在本地服务器时,选择通过专线/VPN网 关/智能网关接入的自建数据库。
		 说明: 选择通过专线/VPN网关/智能网关接入的自建数据 库时,您需要配置VPC ID并填写IP地址和端口信息。
	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。
	ECS实例ID	选择部署了Kafka集群的ECS实例ID。
	数据库类型	选择为Kafka。
	端口	Kafka集群对外提供服务的端口,默认为9092。
	数据库账号	填入Kafka集群的用户名,如Kafka集群未开启验证可不填 写。
	数据库密码	填入Kafka集群用户名对应的密码,如Kafka集群未开启验证 可不填写。
	Торіс	a. 单击击右侧的获取Topic列表。 b. 下拉选择具体的Topic名称。
	Kafka版本	根据目标Kafka集群版本,选择对应的版本信息。

7. 单击页面右下角的授权白名单并进入下一步。

📕 说明:

此步骤会将DTS服务器的IP地址自动添加到源POLARDB集群的白名单和目标ECS实例的内网 入方向安全组规则中,用于保障DTS服务器能够正常连接源和目标实例。

8. 配置目标已存在表的处理模式和同步对象。

1.选择同步通道的源及目标实例 2.选择同步对象		3.高级设置	>	4.预检查
同步架构: 单向同步				
目标已存在表的处理模 式: ● 预检查并报错拦截 ◎ 无爆作				
源库对象		已选择对象 (鼠标移到对象行,点击) 信点我	编辑可修改对象名或过	过滤条件)详
若全局搜索,请先展开树 Q ♥ sys ● dtstestdata ♥ Tables ● order ♥ dtstestdatanew ♥ dtstestdatanew	> <	■ customer 源库名:dtstestdata ■ customer	Q a (1个对象)	
全选		全洗		
*缺射名称更改: 不进行库表名称批量更改 	要进行库表名称批量更	22		
				取消 上一步 下一步

配置项目	配置说明
目标已存在表的处理 模式	 预检查并报错拦截:检查目标数据库中是否有同名的表。如果目标 数据库中没有同名的表,则通过该检查项目;如果目标数据库中有 同名的表,则在预检查阶段提示错误,数据同步作业不会被启动。
	 说明: 如果目标库中同名的表不方便删除或重命名,您可以#unique_50来 避免表名冲突。
	・无操作:跳过目标数据库中是否有同名表的检查项。
	 ▲ 警告: 选择为无操作,可能导致数据不一致,给业务带来风险,例如: - 表结构一致的情况下,如果在目标库遇到与源库主键的值相同
	 的记录,在初始化阶段会保留目标库中的该条记录;在增量同步阶段则会覆盖目标库的该条记录。 表结构不一致的情况下,可能会导致无法初始化数据、只能同步部分列的数据或同步失败。

配置项目	配置说明
选择同步对象	在源库对象框中单击待同步的对象,然后单击 > 将其移动至已选择
	对象框。
	道 说明:
	 · 仅支持表粒度的数据同步。 · 默认情况下,同步对象的名称保持不变。如果您需要同步对象在目标实例上名称不同,那么需要使用DTS提供的对象名映射功能,详情请参见#unique_50。

9. 上述配置完成后,单击页面右下角的下一步。

10 配置同步初始化的高级配置信息。

创建同步作业	返回数据同步列表					
1.选	择同步通道的源及目标实例	\rangle	2.选择同步对象	3.高级设置		4.预检查
	同步初始化: 🗹 结构初始化	☑ 全量数据初始化				
					取満 上一步	保存 预检查并启动



同步初始化类型细分为:结构初始化,全量数据初始化。选择结构初始化和全量数据初始 化后,DTS会在增量数据同步之前,将源数据库中待同步对象的结构和存量数据,同步到目标 数据库。

11上述配置完成后,单击页面右下角的预检查并启动。

	〕 说明:	
•	・在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同	IJ
	步作业。	
•	・如果预检查失败,单击具体检查项后的,查看失败详情。根据提示修复后,重新进行	颀

检查。

12 在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

13等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面, 查看数据同步作业的状态。

同步作业名称 ▼	搜索	排序: 默认排序 ▼ 状	态: 全部 ▼		
□ 实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼	操作
hangzhou-hangzhou-small	同步中	延时:565	按量付费	单向同步 暂停同步	转包年包月 / 升级 更多
■ 暂停同步 释放同步				共有1条 ,每页显示:20条 《	$\langle 1 \rangle \gg$

1.3.8 从ECS上的自建MySQL同步至POLARDB MySQL

POLARDB是阿里巴巴自主研发的下一代关系型分布式云原生数据库,可完全兼容MySQL,具 备简单易用、高性能、高可靠、高可用等优势。通过数据传输服务DTS(Data Transmission Service),您可以将自建的MySQL数据库同步至POLARDB MySQL,本文以ECS上的自 建MySQL为例介绍配置流程。

前提条件

已购买POLARDB MySQL集群,详情请参见创建POLARDB MySQL集群。

注意事项

- ・如果同步对象为单个或多个表(非整库),那么在数据同步时,请勿对源库的同步对象使用ghost或pt-online-schema-change等类似工具执行在线DDL变更,否则会导致同步失败。
- · 全量初始化过程中,并发INSERT会导致目标集群的表碎片,全量初始化完成后,目标集群的表 空间比源集群的表空间大。
- ・如果数据同步的源集群没有主键或唯一约束,且记录的全字段没有唯一性,可能会出现重复数 据。

支持同步的SQL操作

- DML: INSERT, UPDATE, DELETE
- DDL: CREATE TABLE、ALTER TABLE、RENAME TABLE、TRUNCATE TABLE、 DROP TABLE

功能限制

・不兼容触发器

同步对象为整个库且这个库中包含了会更新同步表内容的触发器,那么可能导致同步数据不一致。例如数据库中存在了两个表A和B。表A上有一个触发器,触发器内容为在INSERT一条

数据到表A之后,在表B中插入一条数据。这种情况在同步过程中,如果源集群表A上进行了 INSERT操作,则会导致表B在源集群跟目标集群数据不一致。

此类情况须要将目标集群中的对应触发器删除掉,表B的数据由源集群同步过去,详情请参见<u>触</u> 发器存在情况下如何配置同步作业。

・ RENAME TABLE限制

RENAME TABLE操作可能导致同步数据不一致。例如同步对象只包含表A,如果同步过程中源 集群将表A重命名为表B,那么表B将不会被同步到目标库。为避免该问题,您可以在数据同步 配置时,选择同步表A和表B所在的整个数据库作为同步对象。

同步前准备工作

#unique_54°



用于数据同步的数据库账号需具备待同步对象的SELECT、REPLICATION CLIENT、REPLICATION SLAVE权限。

支持的同步架构

- ・一对一单向同步
- ・一对多单向同步
- ・级联单向同步
- ・多对一单向同步
- 关于各类同步架构的介绍及注意事项,请参见#unique_48。

操作步骤

1. #unique_49/unique_49_Connect_42_section_39h_fto_gdl。



购买时、选择源实例为MvSQL、目标实例为POLARDB、并选择同步拓扑为单向同步。

- 2. 登录数据传输控制台。
- 3. 在左侧导航栏,单击数据同步。

4. 在同步作业列表页面顶部,选择数据同步实例所属地域。

										15 HUX 13 A
数据传输	同步作业列表	华东1 (杭州)	华东2 (上海)	华北1 (青岛)	华北2 (北京)	华南1 (深圳)	华北3 (张家口)	香港		
		美国 (硅谷)	美国 (弗吉尼亚)	新加坡 阿爾	联酋 (迪拜)	德国 (法兰克福)	马来西亚 (吉隆坡)		(所选地域为同步作业目标实例	所在的地域)
概览		· 津大利亚 (悉尼)) 印度 (孟买)	英国 (伦敦)	日本 (东京)	印度尼西亚 (雅	加达) 华北5 (呼和	□浩特)		
数据迁移										
									(?) DTS常见问题	€ 刷新
数据订阅										
数据同步	同步作业名称 •				搜索	排序: 默认排序	; v 状态:	全部	v	
数据同步	同步作业名称 🔻				搜索	排序: 默认排序	, ▼ 状态:	全部	Ŧ	

5. 定位至已购买的数据同步实例,单击配置同步链路。

6. 配置同步通道的源实例及目标实例信息。

1.选择同步通道的源及目标实	例	2.选择同步对象	3.高级设置	>	4.预检查
同步作业名称:	MySQL_TO_POLARDB				
源实例信息					
实例类型:	ECS上的自建数据库]		
实例地区:	华东1 (杭州)				
* ECS实例ID:	i-bp	-			
数据库类型:	MySQL				
* 端口:	3306]		
* 数据库账号:	dtstest]		
* 数据库密码:	•••••	4]		
目标实例信息					
实例类型:	POLARDB				
实例地区:	华东1 (杭州)				
实例地区: * POLARDB实例ID:	华东1 (杭州) pc-bp	-			
实例地区: * POLARDB实例ID: * 数编库账号:	华东1(杭州) pc-bp. dtstest	-]		
 	华东1 (杭州) pc-bp dtstest	-]]		
	华东1 (杭州) pc-bp dtstest	- 4]]		

配置项目	配置选项	配置说明
同步作业名 称	-	DTS会自动生成一个同步作业名称,建议配置具有业务意义的名称(无唯一性要求),便于后续识别。
源实例信息	实例类型	选择ECS上的自建数据库。
	实例地区	购买数据同步实例时选择的源实例地域信息,不可变更。
	ECS实例ID	选择自建MySQL数据库所属的ECS实例ID。
	数据库类型	固定为MySQL,不可变更。
	端口	填入自建MySQL的数据库服务端口。

配置项目	配置选项	配置说明
	数据库账号	填入连接自建MySQL的数据库账号。
		道 说明: 用于数据同步的数据库账号需具备待同步对象 的SELECT、REPLICATION CLIENT、REPLICATION SLAVE权限。
	数据库密码	填入数据库账号对应的密码。
目标实例信	实例类型	固定为POLARDB,不可变更。
息	实例地区	购买数据同步实例时选择的目标实例地域信息,不可变更。
	POLARDB 实例ID	选择目标POLARDB集群ID。
	数据库账号	填入连接POLARDB集群的数据库账号。
		间 说明: 用于数据同步的数据库账号需具备目标同步对象的ALL权限。
	数据库密码	填入数据库账号对应的密码。

7. 单击页面右下角的授权白名单并进入下一步。



此步骤会将DTS服务器的IP地址自动添加到ECS实例的内网入方向安全组规则和目标POLARDB集群的白名单中,用于保障DTS服务器能够正常连接源和目标实例。

8. 配置目标已存在表的处理模式和同步对象。

1.选择同步通道的源及目标实例 2.选择同步对象		3.高级设置	>	4.预检查
同步架构: 单向同步				
目标已存在责的处理模式: ⑧ 预给查并报错拦截 ◎ 无握作				
源库对象		已选择对象(鼠标移到对象行,点击	编辑可修改对象名或过泸	徐 件)详情点我
若全局搜索,请先展开树 Q			Q	
🖃 🧧 dtstestdata		👕 dtstestdata (1个对象)		
🖂 📴 Tables		i order		
🖃 🗤 polardbdatabase				
	>			
	<			
全选		全选		
	** 存在出口下:			
~ 瞬刻名称更成: ● 小进行库表名称抗量更改 ◎ 要进行。	军 表冶称批量更改			
				取消 上一步 下一步

配置项目	配置说明
目标已存在表的处理 模式	 预检查并报错拦截:检查目标数据库中是否有同名的表。如果目标数据库中没有同名的表,则通过该检查项目;如果目标数据库中有同名的表,则在预检查阶段提示错误,数据同步作业不会被启动。
	说明: 如果目标库中同名的表不方便删除或重命名,您可以#unique_50来 避免表名冲突。
	・无操作:跳过目标数据库中是否有同名表的检查项。
	▲ 警告: 选择为无操作,可能导致数据不一致,给业务带来风险,例如:
	 表结构一致的情况下,如果在目标库遇到与源库主键的值相同的记录,在初始化阶段会保留目标库中的该条记录;在增量同步阶段则会覆盖目标库的该条记录。 表结构不一致的情况下,可能会导致无法初始化数据、只能同步部分列的数据或同步失败。

配置项目	配置说明
选择同步对象	在源库对象框中单击待同步的对象,然后单击 > 将其移动至已选择
	对象框。
	同步对象的选择粒度为库、表。
	道 说明:
	 如果选择整个库作为同步对象,那么该库中所有对象的结构变更 操作都会同步至目标库。
	 . 默认情况下,同步对象的名称保持不变。如果您需要改变同步对 象在目标集群中的名称,那么需要使用DTS提供的对象名映射功 能,详情请参见#unique_50。

9. 上述配置完成后,单击页面右下角的下一步。

10 配置同步初始化的高级配置信息。

创建同步作业	€ 返回数据同步列表						
	1.选择同步通道的源及目标	示实例	\rangle	2.选择同步对象	3.高级设置		4.预检查
	同步初始化:	☑ 结构初始化	☑ 全量数据初始化				
						取消 上一步	保存 预检查并启动

同步初始化类型细分为:结构初始化,全量数据初始化。选择结构初始化和全量数据初始 化后,DTS会在增量数据同步之前,将源数据库中待同步对象的结构和存量数据,同步到目标 数据库。

11.上述配置完成后,单击页面右下角的预检查并启动。



- ・ 在数据同步作业正式启动之前,会先进行预检查。只有预检查通过后,才能成功启动数据同 步作业。
- ・如果预检查失败,单击具体检查项后的,查看失败详情。根据提示修复后,重新进行预

检查。

12.在预检查对话框中显示预检查通过后,关闭预检查对话框,同步作业将正式开始。

13等待同步作业的链路初始化完成,直至处于同步中状态。

您可以在 数据同步页面,查看数据同步作业的状态。

同步	作业名称▼	搜索	排序: 默认排序 ▼ 状态:	全部 •		
	实例ID/作业名称	状态	同步概况	付费方式	同步架构(全部) ▼	操作
	nangzhou-hangzhou-small	同步中	延时:565	按量付费	单向同步 暂停同步 转包年包月	升级 更多
	暂停同步 释放同步				共有1条 , 每页显示 : 20条 🛛 « 🧹 1 🔅	> >>

2 POLARDB PostgreSQL

2.1 数据迁移方案概览

云数据库POLARDB提供了多种数据迁移方案,可满足不同上云、迁云的业务需求,使您可以在不 影响业务的情况下平滑将数据库迁移至阿里云云数据库POLARDB上面。

通过使用阿里云数据传输服务 (DTS) ,您可以实现POLARDB的结构迁移、全量迁移等。

数据迁移

使用场景	文档链接
从RDS迁移至POLARDB	从RDS for PostgreSQL迁移至POLARDB for PostgreSQL
从自建数据库迁移至 POLARDB	从自建PostgreSQL迁移至POLARDB for PostgreSQL

2.2 从自建PostgreSQL迁移至POLARDB for PostgreSQL

本文介绍通过pg_dumpall、pg_dump和pg_restore命令将自建PostgreSQL数据库迁移 至POLARDB for PostgreSQL中。

迁移的源库为RDS for PostgreSQL实例时,请参考从RDS for PostgreSQL迁移至POLARDB for PostgreSQL。

前提条件

POLARDB for PostgreSQL实例的存储空间应大于自建PostgreSQL数据库的存储空间。

注意事项

该操作为全量数据迁移。为避免迁移前后数据不一致,迁移操作开始前请停止自建数据库的相关业

务,并停止数据写入。

准备工作

1. 创建一个Linux操作系统的ECS实例,本案例使用的ECS为Ubuntu 16.04 64位操作系统。详 情请参考创建ECS实例。

📋 说明:

·要求ECS实例和迁移的目标POLARDB for PostgreSQL实例处于同一个专有网络。

· 可创建一个按量付费的ECS实例,迁移完成后释放实例。

2. 在ECS实例中安装PostgreSQL,以便执行数据恢复的命令。详情请参考PostgreSQL官方文档。

📕 说明:

请确保安装的PostgreSQL数据库版本与自建PostgreSQL数据库版本一致。

操作步骤一 备份自建数据库

该操作为全量数据迁移。为避免迁移前后数据不一致,迁移操作开始前请停止自建数据库的相关业 务,并停止数据写入。

1. 在自建PostgreSQL数据库服务器上执行以下命令,备份数据库中的所有角色信息。

pg_dumpall -U <username> -h <hostname> -p <port> -r -f <filename>

参数说明:

- · <username>: 登录自建PostgreSQL数据库的账号。
- · <hostname>: 自建PostgreSQL数据库的连接地址,本机可使用localhost。
- ・ <port>: 数据库服务的端口号。
- · <filename>: 生成的备份文件名称。

示例:

pg_dumpall -U postgres -h localhost -p 5432 -r -f roleinfo.sql

- 2. 命令行提示Password:时,输入数据库账号对应的密码,开始备份数据库中的所有角色信息。
- 3. 使用vim命令将角色信息备份文件中的SUPERUSER替换为polar_superuser。

说明: 如果角色信息备份文件中没有SUPERUSER信息,可跳过本步骤。
-- PostgreSQL database cluster dump
-- Set default_transaction_read_only = off;
Set Geautt_transaction_read_only = off;
Set client_encoding = 'UTF8';
Set standard_conforming_strings = on;
-- Roles
-- CREATE ROLE data1;
ALTER ROLE data1 WITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE manisha WITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE manisha WITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN REPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE manisha WITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN REPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE postgres' MITH SUPERUSER INHERIT CREATEROLE CREATEDB LOGIN REPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE postgres' MITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN REPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE postgres' MITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE postgres' MITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE postgres' MITH NOSUPERUSER INHERIT CREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE feastuser' WITH NOSUPERUSER INHERIT NOCREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS PASSWORD 'md'
-- CREATE ROLE feastuser' WITH NOSUPERUSER INHERIT NOCREATEROLE CREATEDB LOGIN NOREPLICATION NOBYPASSRLS;

-- PostgreSQL database cluster dump complete 4. 在自建PostgreSQL数据库服务器上执行以下命令,备份数据库中的数据。

```
pg_dump -U <username> -h <hostname> -p <port> <dbname> -Fd -j <njobs
> -f <dumpdir>
```

参数说明:

- · <username>: 登录自建PostgreSQL数据库的账号。
- · <hostname>: 自建PostgreSQL数据库的连接地址,本机可使用localhost。
- ・ <port>: 数据库服务的端口号。
- · <dbname>: 要备份的数据库名。
- ・ <njobs>: 同时执行备份作业的并发数。

🧾 说明:

- 参数<njobs>可减少转储的时间,但也会增加数据库服务器的负载。
- **如果您的自建PostgreSQL数据库是9.2以前的版本,您还需要指定**--no-synchroniz ed-snapshots参数。
- · <dumpdir>: 生成的备份文件所属目录。

示例:

```
pg_dump -U postgres -h localhost -p 5432 mytestdata -Fd -j 5 -f postgresdump
```

- 5. 命令行提示Password:时,输入数据库账号对应的密码,数据库开始备份。
- 6. 等待备份完成, PostgreSQL数据库数据将备份至指定的目录中, 本案例为postgresdump。

操作步骤二数据迁移至POLARDB for PostgreSQL

1. 将备份文件所属的目录上传至ECS实例中。

包含角色信息备份文件和数据库备份文件。

2. 在ECS上执行以下命令,将角色信息备份文件中的角色信息迁移至POLARDB for

```
PostgreSQL实例中。
```

```
psql -U <username> -h <hostname> -p <port> -d <dbname> -f <
filename>
```

参数说明:

- · <username>: 登录POLARDB for PostgreSQL数据库的账号。
- · <hostname>: POLARDB for PostgreSQL实例的主地址(私网)。
- ・ <port>: 数据库服务的端口号, 默认为1921。
- · <dbname>: 连接的数据库的名称,默认为postgres。
- · <filename>: 角色信息备份文件名。

```
psql -U gctest -h pc-xxxxxx.pg.polardb.cn-qd-pldb1.rds.aliyuncs.
com -d postgres -p 1921 -f roleinfo.sql
```

- 3. 命令行提示Password:时,输入数据库账号对应的密码,角色信息开始导入。
- 4. 在ECS上执行以下命令,将数据库数据恢复至POLARDB for PostgreSQL实例中。

```
pg_restore -U <username> -h <hostname> -p <port> -d <dbname> -j <
njobs> <dumpdir>
```

参数说明:

- · <username>: 登录POLARDB for PostgreSQL数据库的账号。
- <hostname>: POLARDB for PostgreSQL实例的主地址(私网),详情请参考#unique_59。
- ・ <port>: 数据库服务的端口号, 默认为1921。
- ・ <dbname>: 连接并直接恢复到的目标数据库名。

道 说明:

目标数据库需已存在,如不存在请在目标实例中创建该数据库。

· <njobs>: 同时执行数据恢复作业的并发数。



此选项可减少数据恢复的时间,但也会增加数据库服务器的负载。

· <dumpdir>: 备份文件所在目录。

示例:

```
pg_restore -U gctest -h pc-mxxxxxx.pg.polardb.cn-qd-pldb1.rds.
aliyuncs.com -p 1921 -d mytestdata -j 6 postgresdump
```

5. 命令行提示 Password:时,输入数据库账号对应的密码,数据开始迁移。

苣 说明:

如果忘记密码,请参考#unique_60/unique_60_Connect_42_section_ckb_hpq_tdb。

等待数据迁移完成即可。

2.3 从RDS for PostgreSQL迁移至POLARDB for PostgreSQL

本文介绍通过pg_dump和pg_restore命令将自建PostgreSQL数据库迁移至POLARDB for PostgreSQL中。

迁移的源库为自建PostgreSQL数据库时,请参考从自建PostgreSQL迁移至POLARDB for PostgreSQL。

前提条件

POLARDB for PostgreSQL实例的存储空间应大于RDS for PostgreSQL实例的存储空间。

注意事项

```
该操作为全量数据迁移。为避免迁移前后数据不一致,迁移操作开始前请停止自建数据库的相关业
务,并停止数据写入。
```

准备工作

1. 创建一个Linux操作系统的ECS实例,本案例使用的ECS为Ubuntu 16.04 64位操作系统。详 情请参考创建*ECS*实例。

▋ 说明:

- ·要求ECS实例和迁移的目标POLARDB for PostgreSQL实例处于同一个专有网络。
- ·可创建一个按量付费的ECS实例,迁移完成后释放实例。
- 2. 在ECS实例中安装PostgreSQL,以便执行数据恢复的命令。详情请参考PostgreSQL官方文档。

🗾 说明:

请确保安装的PostgreSQL数据库版本与自建PostgreSQL数据库版本一致。

操作步骤一备份RDS for PostgreSQL数据库

该操作为全量数据迁移。为避免迁移前后数据不一致,迁移操作开始前请停止自建数据库的相关业务,并停止数据写入。

1. 在ECS上执行以下命令,备份数据库中的数据。

```
pg_dump -U <username> -h <hostname> -p <port> <dbname> -Fd -j <njobs
> -f <dumpdir>
```

参数说明:

- · <username>: 登录自建PostgreSQL数据库的账号。
- · <hostname>: 自建PostgreSQL数据库的连接地址,本机可使用localhost。
- ・ <port>: 数据库服务的端口号。
- · <dbname>: 指定要连接的数据库的名称,默认为postgres。
- ・ <njobs>: 同时执行备份作业的并发数。

- 说明:

- 参数<njobs>可减少转储的时间,但也会增加数据库服务器的负载。
- 如果您的自建PostgreSQL数据库是9.2以前的版本,您还需要指定--no-synchroniz ed-snapshots参数。
- · <dumpdir>: 生成的备份文件所属目录。

示例:

```
pg_dump -U postgres -h localhost -p 5432 postgres -Fd -j 5 -f postgresdump
```

- 2. 命令行提示Password:时,输入数据库账号对应的密码,数据库开始备份。
- 3. 等待备份完成, PostgreSQL数据库数据将备份至指定的目录中, 本案例为postgresdump。

操作步骤二数据迁移至POLARDB for PostgreSQL

1. 在ECS上连接POLARDB for PostgreSQL数据库。

```
psql -U <username> -h <hostname> -p <port> -d <dbname>
```

参数说明:

- ・ <username>: 登录POLARDB for PostgreSQL数据库的账号。
- <hostname>: POLARDB for PostgreSQL实例的主地址(私网),详情请参 考#unique_59。
- ・ <port>: 数据库服务的端口号, 默认为1921。
- · <dbname>: 要连接的数据库名称。

示例:

```
psql -h pc-mxxxxxx.pg.polardb.cn-qd-pldb1.rds.aliyuncs.com -p 3433
  -d postgres -U gctest
```

- 根据源RDS for PostgreSQL实例数据库中的角色信息,在目标POLARDB for PostgreSQL实例中创建角色信息,并对数据恢复的目标数据库进行授权,详情请参考官方文档 CREATE ROLE和GRANT。
- 3. 在ECS上执行以下命令,将数据库数据迁移至POLARDB for PostgreSQL实例中。

```
pg_restore -U <username> -h <hostname> -p <port> -d <dbname> -j <
njobs> <dumpdir>
```

参数说明:

- · <username>: 登录POLARDB for PostgreSQL数据库的账号。
- <hostname>: POLARDB for PostgreSQL实例的主地址(私网)。
- ・ <port>: 数据库服务的端口号, 默认为1921。
- ・ <dbname>: 连接并直接恢复到的目标数据库名。

📃 说明:

目标数据库需已存在,如不存在请在目标实例中创建该数据库。

· <njobs>: 同时执行数据恢复作业的并发数。

🧾 说明:

此选项可减少数据恢复的时间,但也会增加数据库服务器的负载。

・<dumpdir>: 备份文件所在目录。

示例:

pg_restore -U gctest -h pc-mxxxxxx.pg.polardb.cn-qd-pldb1.rds. aliyuncs.com -p 1921 -d postgres -j 6 postgresdump

4. 命令行提示Password:时,输入数据库账号对应的密码,数据开始迁移。

📋 说明:

如果忘记密码,请参考#unique_61/unique_61_Connect_42_section_ckb_hpq_tdb。

等待数据迁移完成即可。

3 POLARDB兼容Oracle语法引擎

3.1 数据迁移方案概览

云数据库POLARDB提供了多种数据迁移方案,可满足不同上云、迁云的业务需求,使您可以在不 影响业务的情况下平滑将数据库迁移至云数据库POLARDB。

使用场景	文档链接
从自建数据库迁移至 POLARDB	从自建Oracle迁移至POLARDB

3.2 从自建Oracle迁移至POLARDB

本文介绍使用数据库和应用迁移服务ADAM(Advanced Database & Application Migration)帮助您完成数据库的结构迁移与订正,然后使用数据传输服务DTS(Data Transmission Service)将自建Oracle的数据迁移至POLARDB兼容Oracle语法引擎的集群。

前提条件

· POLARDB兼容Oracle语法引擎的集群(以下简称为POLARDB集群)的polar_comp _redwood_raw_names参数需设置为off,详情请参见设置集群参数。

背景信息

ADAM是阿里云结合阿里巴巴多年的内部业务系统和数据库异构迁移经验(例如Oracle迁移 至MySQL/PPAS/PostgreSQL/POLARDB)研发的,可帮助企业数据库和应用业务系统高效、 轻松上云的产品,能够为迁移Oracle数据库到阿里云相关数据库提供全生命周期支持。

采用ADAM+DTS执行数据迁移的方案优势如下:

- 在迁移前采集待迁移的Oracle数据库信息,然后对采集到的数据进行评估分析,最后得出分析 报告集(包含专家意见)。报告集的内容包括目标数据库方案、源库不同类型对象的兼容度和不 兼容原因、迁移后风险对象及SQL、应用相关的改造建议、迁移方案的成本等。此外,分析结果 还包括目标数据库迁移计划。
- 使用ADAM将源库的Schema对象迁移至目标库,其中完全兼容和改造后兼容的对象在迁移过 程中自动完成改造。对于不兼容对象,您可以根据ADAM STUDIO的提示来做改造并验证正确 性。
- ・使用DTS执行全量数据迁移+增量数据迁移,实现平滑迁移上云。

费用说明

 ADAM:使用ADAM STUDIO工具需要申请License授权(目前免费),您可以联系客户经理 或提交工单申请。



您也可以使用数据库专家服务,由阿里云数据库专家为您提供应用架构和数据库架构治理、评 估选型、制定数据库平滑迁云迁移方案,费用说明请参见<u>计费方式与价格</u>。

· DTS: 详情请参见#unique_23。

数据库账号的权限要求

- 步骤一 使用ADAM执行结构迁移
 - 1. 下载ADAM产品采集客户端。
 - 2. 采集Oracle数据库相关信息。
 - 3. 上传采集数据库包。
 - 4. 在阿里云ADAM SaaS分析平台完成源库数据分析和解决方案分析。
 - 5. 使用ADAM STUDIO工具,将自建Oracle数据库的结构信息迁移至POLARDB集群。请下 载阿里云_ADAM_Studio用户手册,然后按照手册内容执行结构迁移。

🛕 警告:

仅使用ADAM STUDIO执行结构迁移,迁移过程中可评估ADAM STUDIO工具给出的提示,然后执行结构订正以获得更好的兼容性。

步骤二 使用DTS执行数据迁移

🗾 说明:

DTS在执行全量数据迁移时将占用源库和目标库一定的读写资源,可能会导致数据库的负载上 升,在数据库性能较差、规格较低或业务量较大的情况下(例如源库有大量慢SQL、存在无主键表 或目标库存在死锁等),可能会加重数据库压力,甚至导致数据库服务不可用。因此您需要在执行 数据迁移前评估源库和目标库的性能,同时建议您在业务低峰期执行数据迁移(例如源库和目标库 的CPU负载在30%以下)。
1. 配置迁移任务的源库和目标库连接信息。

1.源库及目标库		2.迁移类型及列表	\rightarrow	3.映	时名称修改	\rightarrow	4.预检查
* 17 kg kg skg.							
11穷有标:	Oracle_TO_POLARDB						
源库信息							
* 实例类型:	有公网IP的自建数据库		•	DTS支持链路类型	L		
* 实例地区:	华东1 (杭州)		٣	获取DTS IP段			
* 数据库类型:	Oracle						
* 主机名或IP地址:							
* 端口:	1521						
* 实例类型:	● 非RAC实例 ○ RAC实例						
* SID:	testsid						
* 数据库账号:	dtstest						
* 数据库密码:	*******		\$ ≻	测试连接	⊘ 测试通过		
目标库信息. 							
* 实例类型:	POLARDB		Ŧ				
* 实例地区:	华东1 (杭州)		•				
* POLARDB实例ID:	pc-bp		-				
* 数据库名称:	dtstestdata						
* 数据库账号:	dtstest						
* 数据库密码:	•••••		<⊅	测试连接	⊘ 测试通过		
						取消上	云评估 授权白名单并进入下一步

类别	配置	说明
任务名称	-	DTS会自动生成一个任务名称,建议配置具有业务意义的名称(无唯一性要求),便于后续识别。
源库信息	实例类型	您可以根据源库部署位置,选择有公网IP的自建数据库、ECS上 的自建数据库或通过专线/VPN网关/智能网关接入的自建数据 库。 本文以有公网IP的自建数据库为例介绍配置流程,当自
		建MySQL数据库为其他实例类型时,配置流程与该案例类似。
	实例地区	当实例类型选择为有公网IP的自建数据库时,实例地区无需设置。
		前明: 如果您的自建Oracle数据库进行了白名单安全设置,您需要 在实例地区配置项后,单击获取DTS IP段来获取到DTS服务器 的IP地址,并将获取到的IP地址加入自建Oracle数据库的白 名单安全设置中。

类别	配置	说明			
	数据库类型	选择Oracle。			
	主机名或IP地	填入自建Oracle数据库的访问地址。			
	址	说明: 本案例的自建Oracle数据库的服务端口已开放至公网,此处填 入公网IP地址。			
	端口	填入自建Oracle数据库的服务端口,默认为1521。			
	实例类型	・ 非RAC实例:选择该项后,您需要填写SID信息。 ・ RAC实例:选择该项后,您需要填写ServiceName信息。			
	数据库账号	填入自建Oracle的数据库账号,权限要求请参见数据库账号的 权限要求。			
	数据库密码	填入该数据库账号对应的密码。			
目标库信息	实例类型	选择POLARDB。			
	实例地区	选择目标POLARDB集群所属地域。			
	POLARDB实 例ID	选择目标POLARDB集群ID。			
	数据库名称	填入待迁入数据的数据库名称。			
	数据库账号	填入目标POLARDB集群的数据库账号,权限要求请参见数据库 账号的权限要求。			
	数据库密码	填入该数据库账号对应的密码。			

2. 选择迁移类型和迁移对象。

配置	说明
迁移类型	由于已经在 步骤 一中完成了结构迁移,此处无需勾选结构迁移。
	・如果只需要全量迁移,仅勾选全量数据迁移。
	·如果需要不停机迁移,同时勾选全量数据迁移和增量数据迁移。
	道 说明: 增量数据迁移阶段支持同步的SQL操作为: INSERT、UPDATE、DELETE、CREATE TABLE(不支持分区表和包含函 数的表)、ALTER TABLE、DROP TABLE、RENAME TABLE、CREATE INDEX、ADD INDEX。
	▲ 警告: 为保障数据一致性,全量数据迁移期间请勿在自建Oracle数据库中写入新的数据。

配置	说明
迁移对象	在迁移对象框中选择待迁移的对象(对象的结构须已通过ADAM迁移 至POLARDB集群),然后单击 将其移动至已选择对象框。
	▲ 警告: 由于已使用ADAM完成了结构迁移,请勿执行对象名映射操作,否则将导致迁移 失败。

结束迁移任务



为尽可能地减少数据迁移对业务的影响,建议参考#unique_65文档中介绍的流程执行业务切换并 建立回退方案(将POLARDB集群的增量数据实时迁移回源Oracle数据库)。如果无需切换业 务,则可按照下述步骤结束迁移任务。