

数据传输服务 **UDTS**

产品文档

目录

目录	2
概览	14
什么是数据传输服务	17
数据传输	17
数据集成	17
实例类型	18
旗舰版	18
轻量版	18
基础版	19
使用限制	20
功能限制	20
配额限制	20
计费说明	21
计费模式	21
费用列表	21
回收与删除	23
通知渠道	23
回收	23

删除	24
NAT网关实现数据库公网迁移	25
地域列表	25
操作步骤	25
支持数据类型	30
MySQL 迁移到 MySQL	32
迁移内容	32
前提条件	34
功能限制	38
注意事项	40
SSL 安全连接	40
跨版本迁移	41
MySQL 填写表单	42
MySQL 迁移到 TiDB	46
迁移内容	46
前提条件	48
功能限制	52
注意事项	53
SSL 安全连接	54
MySQL 填写表单	55
MySQL与TiDB兼容性说明	58

MySQL 迁移到 ClickHouse	62
功能限制	62
迁移内容	63
填写表单	63
MySQL 迁移到 Kafka	65
1. 功能限制	65
2. 迁移内容	66
3. 表单填写	66
4. Kafka传输数据格式	67
MySQL 迁移到 PostgreSQL	68
1. 功能限制	68
2. 迁移内容	69
3. 表单填写	69
4. 权限要求	70
5. 数据类型映射	71
MySQL 迁移到 MAXIR	75
1. 功能限制	75
2. 迁移内容	77
3. 表单填写	77
4. 权限要求	79
5. 数据类型映射	79

MySQL双向同步	83
创建双向同步任务	83
注意事项	84
SQL Server 迁移到 SQL Server	85
功能限制	85
迁移内容	86
填写表单	86
SQL Server 迁移到 MySQL	88
功能限制	88
迁移内容	89
填写表单	89
SQL Server 迁移到 Kafka	91
1. 功能限制	91
2. 迁移内容	92
3. 表单填写	92
4. Kafka传输数据格式	93
TiDB	94
TiDB 填写表单	94
预检查项	94
CSV源	96

CSV源填写表单	96
MongoDB	97
功能限制	97
注意事项	98
MongoDB填写表单	98
MongoDB	102
源 DynamoDB 填写表单	102
目标 MongoDB 填写表单	102
备注说明	103
Redis	104
功能限制	104
注意事项	104
Redis填写表单	105
断点续传	105
PostgreSQL	106
前提条件	106
功能限制	106
注意事项	107
PostgreSQL 填写表单	108
PostgreSQL 迁移到MAXIR	111

前提条件	111
功能限制	111
注意事项	112
PostgreSQL 填写表单	113
数据类型映射	115
ElasticSearch	118
ElasticSearch 填写表单	118
功能原理	119
技术优势	121
传输速度	121
创建UDTS任务	123
点击【创建任务】	123
基本设置	124
填写任务名称	124
填写源目详细信息	125
预检查	128
完成创建	128
预检查	130
填写表单后点击【立即购买】	130
启动任务	132
启动UDTS迁移任务	132

获取任务详情	133
查看进度	134
支持数据类型	134
查看任务运行历史	140
停止任务	142
停止UDTS迁移任务	142
修改任务	143
修改UDTS迁移任务	143
更新任务名称和备注	145
基于全量任务创建增量任务	147
步骤一 列出全量任务	147
步骤二 查看任务详情	147
步骤三 创建增量任务	148
删除任务	151
删除UDTS迁移任务	151
数据校验	152
支持数据类型	152
前提条件	152

校验内容	153
使用方法	153
注意事项	155
支持数据类型	157
MySQL	158
迁移内容	158
功能说明	158
功能限制	159
TiDB	160
迁移内容	160
功能说明	160
功能限制	161
创建任务	162
填写任务名称	162
添加数据源	163
添加目标	167
预检查	169
数据源预检查	169
目标预检查	171
预检查失败	173

启动任务	175
启动UDTS数据集成任务	175
任务详情	176
查看进度	179
支持数据类型	179
在详情中查看	179
停止任务	181
停止UDTS数据集成任务	181
修改任务	182
编辑UDTS数据集成任务的备注	182
删除任务	184
删除数据集成任务	184
监控指标	185
监控项	185
监控信息查看	186
适用范围	188
监控指标告警	189
告警配置	189

消息订阅	191
如何设置消息订阅	191
FAQ	193
1 问: 现在支持多少区域间的跨域迁移?	193
2 问: MySQL 全量迁移需要满足哪些条件	193
3 问: MySQL 增量迁移需要满足哪些条件	193
4 问: 增量提示 only support ROW format binlog 该如何操作	193
5 问: ERROR 1292 (22007): Incorrect date value: '0000-00-00' for column	196
6 问: 如何判断MySQL-MySQL增量任务中, 源目数据库已经同步完成	197
7 问: error="Error 1040: Too many connections" 错误处理	197
8 问: row size too large (\u003e 8126). Changing some columns to TEXT or BLOB or using ROW_FORMAT=DYNAMIC or ROW_FORMAT=COMPRESSED may help. In current row format, BLOB prefix of 768 bytes is stored inline	198
9 问: Column count doesn't match value count	198
10 问: sync: Type:ExecSQL msg:"exec jobs failed,err:Error 1205:Lock wait timeout exceeded;try restarting transaction"	199
11 问: MySQL 全量任务运行中会有数据库锁吗? 什么时候释放	200
12 问: Redis 迁移出现 ERR illegal address	200
13 问: Error 3140: Invalid JSON text	200
14 问: Mongodb 迁移出现 error reading collection: cursor id 5707195885304103447 not found	201
15 问: Error 1264: Out of range value for column 'xxx' at row 100	201
16 问: 目标库为TiDB时, 出现 Error 1071: Specified key was too long; max key length is 3072 bytes	202
17 问: 目标库为TiDB时, 出现 Error: incorrect utf8 value f09f8c80 for column xxx	202
18 问: 目标库为TiDB时, 出现 Error 1071: Specified key was too long; max key length is 3072 bytes	203
19 问: Cannot add foreign key constraint	203
20 问: 源库是MySQL, 增量同步时出现 ERROR 1236 (HY000): The slave is connecting using CHANGE MASTER TO MASTER_AUTO_POSITION = 1, but the master	

has purged binary logs containing GTIDs that the slave requires	204
21 问: The table 'xxx' is full 或者 No space left on device	206
22 问: Error 1785: When @@GLOBAL.ENFORCE_GTID_CONSISTENCY = 1, updates to non-transactional tables can only be done in either autocommitted statements or single-statement transactions, and never in the same statement as updates to transactional tables	207
23 问: 迁移报错 Invalid default value for 'xxxx_time'	207
24 问: error creating indexes for xx.xx: createIndex error: WiredTigerIndex::insert: key too large to index, failing 1089	208
25 问: OOM command not allowed when used memory > 'maxmemory' or rss_memory not enough.	208
26 问: ddtLog.v_HBUSer: error creating collection ddtLog.v_HBUSer: error running create command: Cannot create view when the featureCompatibilityVersion is 3.X.	209
27 问: 源库是Redis, 报错信息: [PANIC] read error, please check source redis log or network	209
28 问: 源库是MongoDB, 报错信息: Failed: error creating intents to dump: error creating intents for database config: error getting collections for database config: (Unauthorized) not authorized on config to execute command { listCollections: 1, filter: {}, cursor: {}, lsid: { id: UUID("12d62805-acea-472f-9862-ca27253c107e") }, \$db: "config" }	209
29 问: 源库是 MySQL, 报错信息: Deadlock found when trying to get lock; try restarting transaction	210
30 问: load: return result; END; This function has none of DETERMINISTIC, NO SQL, or READS SQL DATA in its declaration and binary logging is enabled (you might want to use the less safe log_bin_trust_function_creators variable	210
预检查错误信息及解决方法	211
1 MySQL	211
2 TiDB	220
3 MongoDB	222
4 Redis	222
如何从阿里云RDS MySQL迁移到UCloud UDB	224
配置外网地址和白名单	224
创建UDTS任务	225

启动UDTS任务	226
如何从阿里云 云数据库Redis 迁移到 UCloud 云内存Redis	229
配置外网地址和白名单	229
设置账号增量同步权限	230
创建UDTS任务	233
自建IDC如何传输数据到UCloud	235
搭建专线	235
创建UDTS任务	235
运行UDTS任务	237
托管云如何传输数据到公有云区域	239
创建UDTS任务	239
运行UDTS任务	241
跨VPC/跨项目/跨账号数据迁移	243
创建UDTS任务	243
运行UDTS任务	245

概览

- 产品简介
 - 什么是数据传输服务
 - 实例类型
 - 使用限制
- 计费说明
 - 计费指南
 - 回收与删除
- 公网迁移
 - NAT网关实现公网迁移
- 数据传输
 - 支持数据类型
 - 源为MySQL
 - MySQL迁移到MySQL
 - MySQL迁移到TiDB
 - MySQL迁移到ClickHouse
 - MySQL迁移到Kafka
 - MySQL迁移到PostgreSQL
 - MySQL迁移到MAXIR
 - MySQL双向同步
 - 源为SQLServer
 - SQL Server迁移到SQL Server
 - SQL Server迁移到MySQL
 - SQL Server迁移到Kafka
 - 源为TiDB
 - TiDB
 - 源为CSV
 - CSV
 - 源为MongoDB

- MongoDB
- 源为 AWS DynamoDB
 - DynamoDB
- 源为 Redis
 - Redis
- 源为 PostgreSQL
 - PostgreSQL 迁移到 PostgreSQL
 - PostgreSQL 迁移到 MAXIR
- 源为 ElasticSearch
 - ElasticSearch
- 功能原理与传输速度
 - MySQL 迁移
- 操作指南
 - 创建任务
 - 预检查
 - 启动任务
 - 获取任务详情
 - 查看进度
 - 查看运行历史
 - 停止任务
 - 修改任务
 - 基于全量任务创建增量任务
 - 删除任务
 - 数据校验
- 数据集成
 - 支持数据类型
 - 源为 MySQL
 - MySQL 迁移到 MySQL
 - MySQL 迁移到 TiDB
 - 操作指南
 - 创建任务
 - 预检查
 - 启动任务

- 获取任务详情
- 查看进度
- 停止任务
- 删除任务
- 监报告警
 - 监控指标
 - 监控指标告警
 - 消息订阅
- FAQ
- 预检查FAQ
- 最佳实践
 - 如何从阿里云RDS MySQL迁移至UCloud UDB
 - 如何从阿里云 云数据库Redis 迁移至UCloud 云内存Redis
 - 自建IDC如何传输数据至UCloud
 - 托管云如何传输数据至公有云区域
 - 跨VPC/跨项目/跨账号数据迁移

什么是数据传输服务

数据传输服务UDTS(UCloud Data Transmission Service) 支持多种同构,异构数据源之间进行 全量/增量 数据传输。UDTS可以轻松帮助用户调整数据架构,跨机房数据迁移,实时数据同步进行后续数据分析等。UDTS无需停止当前数据库服务使用,极大的减少了数据传输对于业务的影响。

主要包括以下几大功能模块:

数据传输

数据传输主要负责数据的全量迁移与增量同步。适用于业务上云、跨云容灾、跨可用区迁移等场景。

数据集成

数据集成专注于多源数据汇聚,为用户建立多对一的数据同步通道。可通过指定策略自动处理数据冲突,同时可对库表名称进行修改。适用于企业BI库集成、数据库架构调整等场景。

实例类型

2022年1月1日起提供三种实例类型供用户选择。

旗舰版

配置 1T 磁盘，适合大型传输业务使用。

数据类型	全量传输速率上限（单位Mbps）	增量同步ops上限
MySQL	800	20000
TiDB	800	NA
Redis	800	90000
MongoDB	800	50000
PostGre	800	40000

轻量版

配置200G磁盘，适合中型传输业务使用。

数据类型	全量传输速率上限（单位Mbps）	增量同步ops上限
------	------------------	-----------

MySQL	240	4000
TiDB	240	NA
Redis	400	60000
MongoDB	400	20000
PostGre	400	20000

基础版

配置100G磁盘，适合小型传输业务使用。

数据类型	全量传输速率上限（单位Mbps）	增量同步ops上限
MySQL	80	2000
TiDB	80	NA
Redis	80	30000
MongoDB	80	10000
PostGre	80	10000

使用限制

功能限制

当前UDTS的目的端仅支持UCloud公有云。

配额限制

每个用户的默认配额为 15,如果用户要使用超过15个任务,请联系技术支持调整配额。如果需求的配额量较大,请至少提前一个月联系技术支持,以保证有足够的时间进行资源调度。

计费说明

UDTS 从2022年01月01日开始对“增量同步”、“全量+增量”及“双向同步”任务收费，“全量”任务继续免费。

计费模式

采用预付费模式。任务创建后开始收费("全量"任务免费), 删除任务后停止计费。如不长期同步数据, 请及时删除实例以停止计费。

费用列表

单位(元/月)

版本	价格
基础版	350
轻量版	600
旗舰版	1100

对于在2022年1月1日之前创建的未设置版本“增量同步”、“全量+增量”及“双向同步”任务, 按如下方式补充版本并开始计费

任务类型	版本	价格
全量任务	轻量版	免费
增量任务	Redis设置为旗舰版, 其它都设置为轻量版	Redis 1100/其它 600
全量+增量任务	Redis设置为旗舰版, 其它都设置为轻量版	Redis 1100/其它 600
双向同步	轻量版	600

回收与删除

通知渠道

所有通知消息将通过邮件、短信以及站内信的方式通知到您设置的通知接收人。

设置通知人:<https://console.ucloud.cn/umon/contact>

回收

预付费按年付/按月付

- 当您的资源过期前3天:给您设置的通知接收人发送资源即将过期预警;
- 当您的资源过期当天:给您设置的通知接收人发送资源已过期通知;
- 当您的资源过期后1天:给您设置的通知接收人发送UDTS任务即将被停止提醒,当天对UDTS任务执行停止操作,停止后的任务需先完成续费方可启动继续使用,因欠费停止的任务续费后会自动重启;
- 当您的资源过期后2天:给您设置的通知接收人发送UDTS任务即将被回收提醒,当天对UDTS任务执行回收操作(回收后的资源不可找回,还请及时续费)。

预付费按小时付

若您的账户可用余额充足,并开启了自动续费,将自动扣费。若您的账户可用余额不足以支持扣费,将产生欠费订单;

- 欠费订单产生时:发送UDTS任务可能被回收通知;
- 欠费订单+24H:发送UDTS任务停服通知,且当天停止任务。

- 欠费订单+48H:发送UDTS任务删除通知,且当天删除任务。

删除

如果您在资源有效期内删除资源,则系统将根据您的使用时长扣除费用,退还剩余的预付金额至您的账户,具体退费规则见删除资源退费。

NAT网关实现数据库公网迁移

地域列表

- 香港
- 台北
- 新加坡
- 东京
- 首尔
- 曼谷
- 雅加达
- 胡志明市
- 孟买
- 洛杉矶
- 华盛顿
- 圣保罗
- 莫斯科
- 迪拜
- 伦敦
- 拉各斯

以上地域公网迁移需用通过NAT网关实现

操作步骤

1. 创建NAT网关

U 全部产品

NAT网关列表 / 创建NAT网关

网络设置

出外网模式 普通模式 白名单模式

所属VPC DefaultVPC

指定子网 无可用于子网, 请到控制台创建>>

外网弹性IP 现有 新购

线路 BGP 电信

覆盖多家网络运营商线路, 自动选择最优线路访问

计费方式 带宽计费 流量计费 共享带宽 [查看文档](#)

带宽上限 100 M

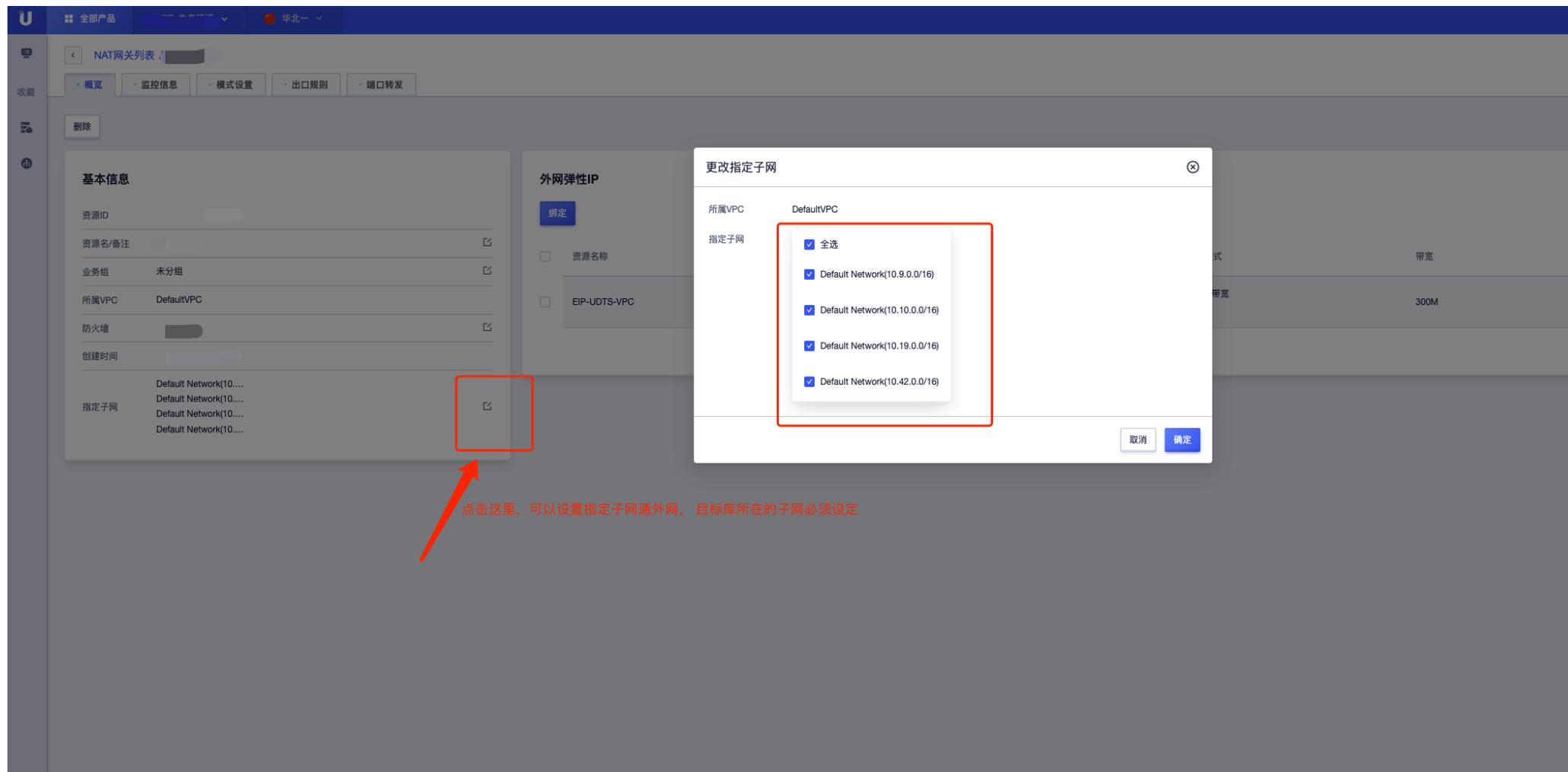
防火墙 非Web服务器推荐(22, 3389)

管理设置

名称 迁移NAT

业务组 未分组

2. 配置NAT网关



3. 创建任务

3.1 源网络类型选择专线

3.2 源填写公网地址

3.3 目标选择要迁移的目标库

U 全部产品 全球服务

数据迁移 / 创建任务

基础设置

资源类型

如不需要长期迁移，建议增量迁移完成后，及时删除实例，从而停止计费。

资源类型	全量	增量
旗舰版	传输速度 800 Mbps 磁盘空间 1000 G	MySQL 20000 QPS MongoDB 50000 QPS Redis 90000 QPS PostgreSQL 40000 QPS
基础版	传输速度 80 Mbps 磁盘空间 100 G	MySQL 2000 QPS MongoDB 10000 QPS Redis 30000 QPS PostgreSQL 10000 QPS

任务设置

任务类型

全量任务：将数据源库中指定库表/整库的全量数据一次性传输到目标端；增量任务：将数据源库中指定库表/整库的增量数据持续传输到目标端，任务持续运行。

全量任务 增量任务 **全量+增量任务** 双向同步

数据源类型

MySQL MongoDB Redis PostgreSQL DynamoDB SQL Server

传输目标类型

MySQL TIDB PostgreSQL Kafka ClickHouse MAXIR

mysql 数据源

您尚未设置数据源 立即设置

更多设置

任务名称

请输入名称

最大重试次数

UDTS为运行失败的任务提供了自动重启任务的功能，重试次数最大限定5次。

3

设置数据源

类型设置

网络类型

用户可以使用内网地址+VPC的方式，或者直接使用公网地址进行数据传输。内网地址当前仅开放右侧可选区域。

内网地址 外网地址 **专线地址**

专线地址

端口

管理设置

用户名

密码

数据库名

多库使用 隔开，也可使用“全库传输”

表名

最大速率

如对速率配置有疑问，可查看 [建议速率配置表](#)

MB/s

NoLocks

默认关闭，对于无法获取super权限的友商RDS服务需要开启，UDB获取super权限详见FAQ

OFF

SSL安全连接

OFF

网络类型选择专线

填写源库公网地址

支持数据类型

数据源	数据目标	全量迁移	增量同步
MySQL	MySQL	支持	支持
MySQL	TiDB	支持	支持
MySQL	PostgreSQL	支持	支持全量+增量
MySQL	Kafka	支持	支持全量+增量
MySQL	ClickHouse	支持	支持
MySQL	MAXIR	支持	支持
TiDB	TiDB	支持	暂不支持
TiDB	MySQL	支持	暂不支持
MongoDB	MongoDB	支持	支持
PostgreSQL	PostgreSQL	支持	支持全量+增量
CSV	MySQL	支持	N/A
Redis	Redis	支持	支持全量+增量
Dynamodb	MongoDB	支持	支持全量+增量
ElasticSearch	ElasticSearch	支持	暂不支持

SQL Server	SQL Server	支持	支持全量+增量
SQL Server	MySQL	支持	支持全量+增量
SQL Server	PostgreSQL	支持	支持全量+增量
SQL Server	Kafka	支持	支持全量+增量

MySQL 迁移到 MySQL

UDTS支持MySQL作为数据传输源/目标,支持版本有 MySQL(包含Percona版)5.5/5.6/5.7/8.0; MariaDB 10.1.2 及以上,以及兼容MySQL的数据库,像PolarDB等。

迁移内容

全量和增量迁移以下内容

- Database、Table 结构及数据
- 视图(View)
- 函数(Function)、存储过程(Procedure)

在全量迁移时,为了保证数据的一致性,防止数据冲突,会在迁移前清理目标库中的数据,清理的内容为本次迁移对应的 Database 和 Table,具体如下

数据库名设置	表名设置	清理内容
*		清理源库中除内置库之外的数据库(参考备注说明 1)
db1,db2,db3		清理任务指定的多个 DB(参考备注说明 2)
db1		清理任务指定的单个 DB
db1	table1,table2,table3	清理任务指定的单个DB 下指定的多个 Table(参考备注说明 3)
db1	table1	清理任务指定的单个DB 下指定的单个 Table

备注说明:

- 1、假设在源库上执行 SHOW DATABASES 的内容如下

```
> SHOW databases;

INFORMATION_SCHEMA
PERFORMANCE_SCHEMA
mysql
sys
mydb1
mydb2
```

当前的内置库为 mysql、sys、INFORMATION_SCHEMA、PERFORMANCE_SCHEMA, 排除这些 DB 后, 剩余 mydb1和mydb2, 则在迁移任务运行时, 会在目标数据库中执行

```
DROP DATABASE IF EXISTS `mydb1`;
DROP DATABASE IF EXISTS `mydb2`;
```

- 2、迁移任务配置的 数据库名为 db1,db2,db3, 则在目标数据库中会执行

```
DROP DATABASE IF EXISTS `db1`;
DROP DATABASE IF EXISTS `db2`;
DROP DATABASE IF EXISTS `db3`;
```

- 3、迁移任务配置的 数据库名为 db1, 迁移的 表名 为 table1,table2,table3, 则在目标数据库中会执行

```
DROP TABLE IF EXISTS `db1`.`table1`;
```

```
DROP TABLE IF EXISTS `db1`.`table2`;  
DROP TABLE IF EXISTS `db1`.`table3`;
```

前提条件

UDS 在迁移时,可以通过预检查完成对需要条件的检查,包括以下内容

- 连接性检查,包括主机地址、端口、用户名和密码等信息的正确性。
- 权限检查,检查迁移时需要的权限。
- 配置检查,如 sql_mode、binlog 格式、唯一键检查等。

连接性检查

连接性检查会对填写的 主机地址,端口,用户名,密码进行检查

- 主机连通性检查:如果填写了错误的主机地址或者端口,会出现连通性检查失败: Project OR IP:10.19.37.212 OR Vpc:uvnet-u0ecvp, Subnet:subnet-2slodr error, please check again, 请确认当前的 主机地址,端口,选择的 VPC ID,所属子网,以及项目是否正确。
- 用户名密码检查:如果填写了错误的用户名或者密码,会出现连通性检查失败: Error 1045: Access denied for user 'root1'@'10.42.255.213' (using password: YES), 请确认用户名和密码正确。
- 防火墙和白名单:如果主机地址,端口,用户名和密码均正确
 - 如果 MySQL 为主机自建,需要排查主机的 iptables 规则
 - 如果设置了白名单,请联系技术支持获取白名单地址

权限检查

默认配置迁移需要源数据库账号拥有 SUPER 权限,如果没有 SUPER 权限,请选择 "NoLocks" 模式。

如果用户需要在全量阶段开启 NoBinlog 模式,那么在目标库还需要有 SUPER 权限。

类型/权限	源库权限(开启 NoLocks)	源库权限(未开启 NoLocks)	目标库权限
全量	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "PROCESS"	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "RELOAD", "LOCK TABLES", "PROCESS"	"SELECT", "INSERT", "UPDATE", "CREATE", "DROP", "ALTER", "DELETE", "INDEX", "CREATE VIEW", "CREATE ROUTINE"
增量	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW"	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW"	"SELECT", "INSERT", "UPDATE", "CREATE", "DROP", "ALTER", "DELETE", "INDEX", "CREATE VIEW", "CREATE ROUTINE", "ALTER ROUTINE"
全+增	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "PROCESS"	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "RELOAD", "LOCK TABLES", "PROCESS"	"SELECT", "INSERT", "UPDATE", "CREATE", "DROP", "ALTER", "DELETE", "INDEX", "CREATE VIEW", "CREATE ROUTINE", "ALTER ROUTINE"

sql_mode 检查

为了保证迁移能正确执行,最好保持 源和目标数据库的 sql_mode 一致,可以通过以下命令查询 sql_mode

```
select @@sql_mode;
```

--- 或者

```
show variables like "sql_mode";
```

如果目标数据库和源数据的 sql_mode 不一致,可以通过下面的命令进行修改

```
SET GLOBAL sql_mode='xxx';
```

其中具体的值,可以通过连接源数据库查询。

binlog 格式检查

如果迁移的任务类型为 增量、全量+增量, 或者全量任务后期需要进行增量迁移,需要源库开启 binlog,且格式设置为ROW, image设置为FULL

```
binlog_format 为 ROW  
binlog_row_image 为 FULL
```

查询方式:

```
show global variables like 'binlog_format';  
show global variables like 'binlog_row_image';
```

设置方式:

```
set global binlog_format = "ROW" ;  
set global binlog_row_image = "FULL" ;
```

备注: 如果是 MySQL 5.5 ,没有 binlog_row_image 这个变量,不需要设置

MyISAM 引擎表检查

如果源库需要迁移的表中包括 MyISAM 引擎表,同时目标库开启了 GTID ,可能导致 MySQL 1785 错误,报错信息如下:

```
When @@GLOBAL.ENFORCE_GTID_CONSISTENCY = 1, updates to non-transactional tables can only be done in either autocommitted statements or single-statement transactions, and never in the same statement as updates to transactional tables
```

建议用户将 MyISAM 引擎表转换为 InnoDB 引擎表,或者关闭目标库的 GTID 模式。查询方式:

```
# 在源库中查询数据库db1中是否存在 MyISAM 表
select table_schema, table_name
from information_schema.tables
where engine = 'MyISAM'
and table_type = 'BASE TABLE'
and table_schema in (db1);

# 在目标库中查询是否开启了 GTID
show global variables like 'gtid_mode';
```

设置方式:

```
# 方案一:修改源库
# 将 MyISAM 表 table1 的引擎修改为 InnoDB
alter table table1 ENGINE = InnoDB;

# 方案二:修改目标库
# 关闭目标库的 GTID 模式
```

```
set global gtid_mode = "ON_PERMISSIVE";
set global gtid_mode = "OFF_PERMISSIVE";
set global gtid_mode = "OFF";
```

功能限制

MyISAM 引擎表

UDTS 支持 MyISAM 引擎表的全量迁移及增量同步,但是有以下限制:

- 导出数据时MyISAM表允许读,不允许写,并且写SQL会阻塞之后的读SQL直到表转储完毕,可能会影响业务。
- 不支持一条事务中同时更新MyISAM引擎表和InnoDB引擎表。
- 对于损坏的表,在迁移之前要进行修复。

建议用户将MyISAM 引擎表转换为InnoDB引擎表以后再使用UDTS迁移。

数据量

建议单个全量迁移任务所迁移的数据量不超过200G,最大支持500G。如果所要迁移的数据量超过了500G,可以将任务拆分为多个任务进行迁移。UDTS 提供按库、按表、按多库、按多表等多维度的迁移方式。

如果迁移任务超过了500G,且无法拆分为多个任务,请联系技术支持。

其它限制

- 源库不能存在同名但大小写不一致的库或表,否则同步可能会异常,建议全部采用小写格式。

- 不迁移 test 以外的内置数据库。
- 在转储过程中,如果待迁移的数据库上有执行 DDL 语句,UDTS 任务会失败,用户可以选择一个不会执行 DDL 语句的时间段重启任务。
- 全量和增量阶段暂时不支持 event 和 trigger。
- 增量同步开启之前,需要关闭 event. `` -- 停止所有 event

```
SET GLOBAL event_scheduler = OFF;
```

```
-- 如果按库迁移,停止指定库的 event 即可
```

```
-- 查找对应的 event 并停止
```

```
SHOW EVENTS;
```

```
ALTER EVENT hello DISABLE;
```

```
``
```

- 源库不能有超过4G的binlog文件,建议大小是1G,否则同步会出错或失败。
- 不支持 MySQL 8.0 的新特性 binlog 事务压缩 Transaction_payload_event。使用 binlog 事务压缩有导致上下游数据不一致的风险。
- 如果源是阿里云 MySQL RDS,所有表必须要有显式主键,否则增量同步功能不支持。
- 数据库名中包含特殊字符'\$'时,只支持单库迁移。
- 表名或者库名不能包含特殊字符'.',否则迁移失败。
- 增量迁移时,DDL 语句仅支持字符集ascii/latin1/binary/utf8/utf8mb4。

注意事项

存储空间

如果迁移目标MySQL数据库开启了Binlog, 目标数据库产生的 Binlog 会占用存储空间。当数据量较大时(超过200G), 建议用户打开 NoBinlog 选项, 这样在全量迁移的过程中目标数据库不会产生Binlog, 减少迁移对磁盘的额外需求, 也可加快全量迁移速度。如果在迁移的过程中一定要开启Binlog, 请为目标数据库创建较大的存储空间或者定时清理不需要的Binlog 文件, 以免存储空间不足导致任务失败(根据经验, 迁移3TB的数据会产生约3TB的Binlog文件, 即源数据库存储空间为3TB, 目标需要6TB存储空间)。

UDS MySQL 支持按保留时长和磁盘使用百分比设置Binlog自动清理策略, 详情参考UDS文档。目标库如果是高可用UDS, 在开启 NoBinlog 后会产生高可用告警, 用户可以忽略此告警, 等待全量任务完成后重做目标库高可用即可。

如果任务失败, 在调整好目标数据库配置之后, 可以重新启动任务, 任务会重新开始。

主从切换

如果源MySQL数据库是主从结构, 当主从发生切换时, 对UDS产生的影响如下:

- 发生在全量迁移阶段, 全量迁移失败, 需要重启任务。
- 发生的增量同步阶段, 如果使用GTID同步, 则无影响; 如果没有使用GTID同步, 则同步失败, 需要重新做全+增。

SSL 安全连接

为了提高链路的安全性, 支持使用 SSL 证书连接数据库。SSL 在传输层对数据进行加密, 提升通信数据的安全性, 但会增加一定的网络连接响应时间。

UDS 当前支持 pem 格式的证书, 如果您使用的是其它格式的证书, 可以先转换为 pem 格式, SSL 安全连接可以在源或者目标中设置。

如何使用

在源或者目标中, 将 SSL 安全连接 选项打开, 同时上传 ca 证书

端口*	<input type="text"/>
用户名*	<input type="text"/>
密码*	<input type="password"/>
最大速率 ①	不填写默认为64 MB/s <input type="button" value="▲"/> <input type="button" value="▼"/>
NoBinlog	<input type="checkbox"/> OFF
SSL安全连接	<input checked="" type="checkbox"/> ON
CA根证书*	<input type="button" value="请选择文件"/> 请选择符合规则的文件
Client证书	<input type="button" value="请选择文件"/> 请选择符合规则的文件
Client私钥	<input type="button" value="请选择文件"/> 请选择符合规则的文件

跨版本迁移

跨版本迁移时, 请注意源库中是否存在与目标库不兼容的特性, 例如:

- MySQL8.0 版本的默认 collation utf8mb4_0900_ai_ci 在 MySQL5.x 版本中不兼容。如需从 MySQL8.0 迁移至 MySQL5.x, 请修改 collation 为 MySQL5.x 支持的类型, 例如 utf8mb4_general_ci
- MySQL5.6 版本支持数据类型 geometry 字段为空字符串, MySQL5.7 及以后版本不支持。如需从 MySQL5.6 迁移至更高版本, 请确保源库中 geometry 字段不为空字符串

MySQL 填写表单

数据源表单

参数名	说明
地址类型	支持内网地址, 外网地址, 专线地址三种方式。内网地址需要填写VPC和子网信息; 外网地址支持IP和域名两种方式; 专线地址既支持IP, 也支持域名, 如果使用域名需要用户网络有外网出口。
端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
数据库名	MySQL数据库名称。所有库传输请填写 *; 指定一个数据库传输, 请填写数据库名; 指定多个数据库传输, 依次输入多个数据库名, 库名之间使用英文逗号隔开。(如果数据库名称中包含空格则无法做增量迁移)
表名	MySQL传输表名。只有当“数据库名”为指定一个数据库时有效。若不填, 默认为迁移指定库中的所有表; 指定一张表传输, 请填写表名; 指定多张表传输, 依次输入多张表名, 表名之间使用英文逗号隔开
最大速率	外网/专线的速率范围为 1-56 MB/s; 内网的速率范围为 1-1024 MB/s

Nolocks	默认关闭,对于无法获取 SUPER 权限的友商RDS服务需要开启,UDB获取 SUPER 权限详见FAQ
SSL 安全连接	默认关闭,当您需要使用证书连接数据库时,可以打开该选项,具体可以参考

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
最大速率	内网的速率范围为 1-1024 MB/s
NoBinlog	当数据源和目标都是MySQL时,在全量阶段可以设置关闭目标端binlog文件的写入,默认不关闭。开启NoBinlog需要当前用户拥有 SUPER 权限。
SSL 安全连接	默认关闭,当您需要使用证书连接数据库时,可以打开该选项,具体可以参考

建议速率配置表

源限速配置参考值

数据库可用内存大小 (G)	内网建议值 (MB/s)	外网/专线建议值 (MB/s)
(0, 2)	10	10
[2, 4)	15	15
[4, 6)	20	20

[6, 8)	30	30
[8, 10)	40	40
[10, 20)	50	50
[20, 40)	60	56
[40, 60)	70	56
[60, 80)	80	56
[80, 100)	90	56
≥ 100	100	56

目标限速配置参考值

数据库可用内存大小 (G)	内网建议值 (MB/s)
(0, 1)	1
[1, 2)	2
[2, 4)	5
[4, 6)	10
[6, 8)	20
[8, 10)	30
[10, 20)	50
[20, 40)	60

[40, 60)	70
[60, 80)	80
[80, 100)	90
≥ 100	100

MySQL 迁移到 TiDB

UDTS支持MySQL作为数据传输源传输到TiDB,支持版本有 MySQL(包含Percona版)5.5/5.6/5.7/8.0; MariaDB 10.1.2 及以上,以及兼容MySQL的数据库,像PolarDB等。

迁移内容

全量和增量迁移以下内容

- Database、Table 结构及数据

在全量迁移时,为了保证数据的一致性,防止数据冲突,会在迁移前清理目标库中的数据,清理的内容为本次迁移对应的 Database 和 Table,具体如下

数据库名设置	表名设置	清理内容
*		清理源库中除内置库之外的数据库(参考备注说明 1)
db1,db2,db3		清理任务指定的多个 DB(参考备注说明 2)
db1		清理任务指定的单个 DB
db1	table1,table2,table3	清理任务指定的单个DB 下指定的多个 Table(参考备注说明 3)
db1	table1	清理任务指定的单个DB 下指定的单个 Table

备注说明:

- 1、假设在源库上执行 SHOW DATABASES 的内容如下

```
> SHOW databases;

INFORMATION_SCHEMA
PERFORMANCE_SCHEMA
mysql
sys
mydb1
mydb2
```

当前的内置库为 `mysql`、`sys`、`INFORMATION_SCHEMA`、`PERFORMANCE_SCHEMA`，排除这些 DB 后，剩余 `mydb1`和`mydb2`，则在迁移任务运行时，会在目标数据库中执行

```
DROP DATABASE IF EXISTS `mydb1`;
DROP DATABASE IF EXISTS `mydb2`;
```

- 2、迁移任务配置的 数据库名为 `db1`、`db2`、`db3`，则在目标数据库中会执行

```
DROP DATABASE IF EXISTS `db1`;
DROP DATABASE IF EXISTS `db2`;
DROP DATABASE IF EXISTS `db3`;
```

- 3、迁移任务配置的 数据库名为 `db1`，迁移的 表名 为 `table1`、`table2`、`table3`，则在目标数据库中会执行

```
DROP TABLE IF EXISTS `db1`.`table1`;
DROP TABLE IF EXISTS `db1`.`table2`;
DROP TABLE IF EXISTS `db1`.`table3`;
```

前提条件

UDS 在迁移时,可以通过预检查完成对需要条件的检查,包括以下内容

- 连接性检查,包括主机地址、端口、用户名和密码等信息的正确性。
- 权限检查,检查迁移时需要的权限。
- 配置检查,如 sql_mode、binlog 格式、唯一键检查等。

连接性检查

连接性检查会对填写的 主机地址,端口,用户名,密码进行检查

- 主机连通性检查:如果填写了错误的主机地址或者端口,会出现连通性检查失败: Project OR IP:10.19.37.212 OR Vpc:uvnet-u0ecvp, Subnet:subnet-2slodr error, please check again, 请确认当前的 主机地址,端口,选择的 VPC ID,所属子网,以及项目是否正确。
- 用户名密码检查:如果填写了错误的用户名或者密码,会出现连通性检查失败: Error 1045: Access denied for user 'root1'@'10.42.255.213' (using password: YES), 请确认用户名和密码正确。
- 防火墙和白名单:如果主机地址,端口,用户名和密码均正确
 - 如果 MySQL 为主机自建,需要排查主机的 iptables 规则
 - 如果设置了白名单,请联系技术支持获取白名单地址

权限检查

默认配置迁移需要源数据库账号拥有 SUPER 权限,如果没有 SUPER 权限,请选择 "NoLocks" 模式。

如果用户需要在海量阶段开启 NoBinlog 模式,那么在目标库还需要有 SUPER 权限。

类型/权限	源库权限(开启 NoLocks)	源库权限(未开启 NoLocks)	目标库权限
全量	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "PROCESS"	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "RELOAD", "LOCK TABLES", "PROCESS"	"SELECT", "INSERT", "UPDATE", "CREATE", "DROP", "ALTER", "DELETE", "INDEX", "CREATE VIEW"
增量	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW"	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW"	"SELECT", "INSERT", "UPDATE", "CREATE", "DROP", "ALTER", "DELETE", "INDEX", "CREATE VIEW"
全+增	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "PROCESS"	"SELECT", "REPLICATION SLAVE", "REPLICATION CLIENT", "SHOW VIEW", "RELOAD", "LOCK TABLES", "PROCESS"	"SELECT", "INSERT", "UPDATE", "CREATE", "DROP", "ALTER", "DELETE", "INDEX", "CREATE VIEW"

sql_mode 检查

为了保证迁移能正确执行,最好保持 源和目标数据库的 sql_mode 一致,可以通过以下命令查询 sql_mode

```
select @@sql_mode;

--- 或者

show variables like "sql_mode";
```

如果目标数据库和源数据的 sql_mode 不一致,可以通过下面的命令进行修改

```
SET GLOBAL sql_mode='xxx';
```

其中具体的值,可以通过连接源数据库查询。

binlog 格式检查

如果迁移的任务类型为 增量、全量+增量, 或者全量任务后期需要进行增量迁移,需要源库开启 binlog,且格式设置为ROW, image设置为FULL

```
binlog_format 为 ROW  
binlog_row_image 为 FULL
```

查询方式:

```
show global variables like 'binlog_format';  
show global variables like 'binlog_row_image';
```

设置方式:

```
set global binlog_format = "ROW" ;  
set global binlog_row_image = "FULL" ;
```

备注: 如果是 MySQL 5.5 ,没有 binlog_row_image 这个变量,不需要设置

从MySQL迁移到TiDB时, 检查源库字符集

TiDB目前支持的字符集包括ascii/latin1/binary/utf8/utf8mb4。从MySQL迁移到TiDB时,如果源库中需要迁移的表或表中某一字段采用的字符集不包含在上述字符集之中,则无法迁移。

查询方式:

```
show create table table1;
```

设置方式:

```
# 将表 table1 的字符集修改为 utf8
alter table task character set utf8;

# 将表 table1 中 column1 字段的字符集修改为 utf8
alter table table1 change column1 column1 varchar(200) character set utf8;
```

lower_case_table_names 检查

TiDB 的lower_case_table_names值为2,并且只能为2,意思是表名和字段名默认大小写不敏感。如果源MySQL lower_case_table_names值为0(大小写敏感),并且存在表名相同但是大小写不同的表,迁移会出错,需要提前在源库修改此类表名。建议库名,表名,字段名等统一采用小写形式。

查询方式:

```
show global variables like 'lower_case_table_names';
```

功能限制

MyISAM 引擎表

UDTS 支持 MyISAM 引擎表的全量迁移及增量同步,但是有以下限制:

- 导出数据时MyISAM表允许读,不允许写,并且写SQL会阻塞之后的读SQL直到表转储完毕,可能会影响业务。
- 不支持一条事务中同时更新MyISAM引擎表和InnoDB引擎表。
- 对于损坏的表,在迁移之前要进行修复。

建议用户将MyISAM 引擎表转换为InnoDB引擎表以后再使用UDTS迁移。

数据量

建议单个全量迁移任务所迁移的数据量不超过200G,最大支持500G。如果所要迁移的数据量超过了500G,可以将任务拆分为多个任务进行迁移。UDTS 提供按库、按表、按多库、按多表等多维度的迁移方式。

如果迁移任务超过了500G,且无法拆分为多个任务,请联系技术支持。

其它限制

- 源库不能存在同名但大小写不一致的库或表,否则同步可能会异常,建议全部采用小写格式。
- 不迁移除 test 以外的内置数据库。
- 在转储过程中,如果待迁移的数据库上有执行 DDL 语句,UDTS 任务会失败,用户可以选择一个不会执行 DDL 语句的时间段重启任务。
- 全量和增量阶段不支持 event, trigger, procedure, function。
- 增量同步开启之前,需要关闭 event. `` -- 停止所有 event

```
SET GLOBAL event_scheduler = OFF;
```

```
-- 如果按库迁移, 停止指定库的 event 即可
```

```
-- 查找对应的 event 并停止
```

```
SHOW EVENTS;
```

```
ALTER EVENT hello DISABLE;
```

```
...
```

- 源库不能有超过4G的binlog文件, 建议大小是1G, 否则同步会出错或失败。
- 不支持 MySQL 8.0 的新特性 binlog 事务压缩 Transaction_payload_event。使用 binlog 事务压缩有导致上下游数据不一致的风险。
- 如果源是阿里云 MySQL RDS, 所有表必须要有显式主键, 否则增量同步功能不支持。
- 数据库名中包含特殊字符'\$'时, 只支持单库迁移。
- 表名或者库名不能包含特殊字符'.', 否则迁移失败。
- 增量迁移时, DDL 语句仅支持字符集ascii/latin1/binary/utf8/utf8mb4。

注意事项

主从切换

如果源MySQL数据库是主从结构, 当主从发生切换时, 对UDTS产生的影响如下:

- 发生在全量迁移阶段,全量迁移失败,需要重启任务。
- 发生的增量同步阶段,如果使用GTID同步,则无影响;如果没有使用GTID同步,则同步失败,需要重新做全+增。

SSL 安全连接

为了提高链路的安全性,支持使用 SSL 证书连接数据库。SSL 在传输层对数据进行加密,提升通信数据的安全性,但会增加一定的网络连接响应时间。

UDTS 当前支持 pem 格式的证书,如果您使用的是其它格式的证书,可以先转换为 pem 格式,SSL 安全连接可以在源或者目标中设置。

如何使用

在源或者目标中,将 SSL 安全连接 选项打开,同时上传 ca 证书

端口*	<input type="text"/>
用户名*	<input type="text"/>
密码*	<input type="password"/>
最大速率 ①	不填写默认为64 MB/s <input type="text"/>
NoBinlog	<input type="checkbox"/> OFF
SSL安全连接	<input checked="" type="checkbox"/> ON
CA根证书*	<input type="button" value="请选择文件"/> 请选择符合规则的文件
Client证书	<input type="button" value="请选择文件"/> 请选择符合规则的文件
Client私钥	<input type="button" value="请选择文件"/> 请选择符合规则的文件

MySQL 填写表单

数据源表单

参数名	说明
地址类型	支持内网地址, 外网地址, 专线地址三种方式。内网地址需要填写VPC和子网信息; 外网地址支持IP和域名两种方式; 专线地址既支持IP, 也支持域名, 如果使用域名需要用户网络有外网出口。
端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
数据库名	MySQL数据库名称。所有库传输请填写 *; 指定一个数据库传输, 请填写数据库名; 指定多个数据库传输, 依次输入多个数据库名, 库名之间使用英文逗号隔开。(如果数据库名称中包含空格则无法做增量迁移)
表名	MySQL传输表名。只有当“数据库名”为指定一个数据库时有效。若不填, 默认为迁移指定库中的所有表; 指定一张表传输, 请填写表名; 指定多张表传输, 依次输入多张表名, 表名之间使用英文逗号隔开
最大速率	外网/专线的速率范围为 1-56 MB/s; 内网的速率范围为 1-1024 MB/s
Nolocks	默认关闭, 对于无法获取 SUPER 权限的友商RDS服务需要开启, UDB获取 SUPER 权限详见FAQ
SSL 安全连接	默认关闭, 当您需要使用证书连接数据库时, 可以打开该选项, 具体可以参考

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	TiDB 连接端口
用户名	TiDB 连接用户名

密码	TiDB 数据库对应用户密码
最大速率	内网的速率范围为 1-1024 MB/s
兼容MySQL自增模式	默认关闭, TiDB默认自增ID和MySQL行为不一样, 开启此选项可以在大多数情况下兼容MySQL自增模式, 具体区别请参考: https://docs.pingcap.com/zh/tidb/stable/auto-increment/#mysql-%E5%85%BC%E5%AE%B9%E6%A8%A1%E5%BC%8F

建议速率配置表

源限速配置参考值

数据库可用内存大小 (G)	内网建议值 (MB/s)	外网/专线建议值 (MB/s)
(0, 2)	10	10
[2, 4)	15	15
[4, 6)	20	20
[6, 8)	30	30
[8, 10)	40	40
[10, 20)	50	50
[20, 40)	60	56
[40, 60)	70	56
[60, 80)	80	56
[80, 100)	90	56

>=100	100	56
-------	-----	----

MySQL与TiDB兼容性说明

TiDB 高度兼容MySQL协议、MySQL 常用的功能及语法。MySQL 生态中的系统工具 (PHPMyAdmin、Navicat、MySQL Workbench、mysqldump、Mydumper/Myloader)、客户端等均适用于 TiDB。

但 TiDB 尚未支持一些 MySQL 功能,可能的原因如下:

有更好的解决方案,例如 JSON 取代 XML 函数。目前对这些功能的需求度不高,例如存储过程和函数。一些功能在分布式系统上的实现难度较大。

不支持的功能特性

<https://docs.pingcap.com/zh/tidb/stable/mysql-compatibility/>

- 存储过程与函数
- 触发器
- 事件
- 自定义函数
- 全文语法与索引
- 空间类型的函数 (即 GIS/GEOMETRY)、数据类型和索引
- 非 ascii、latin1、binary、utf8、utf8mb4、gbk 的字符集
- SYS schema

- MySQL 追踪优化器
- XML 函数
- X-Protocol
- 列级权限
- XA 语法 (TiDB 内部使用两阶段提交, 但并没有通过 SQL 接口公开)
- CREATE TABLE tblName AS SELECT stmt 语法
- CHECK TABLE 语法
- CHECKSUM TABLE 语法
- REPAIR TABLE 语法
- OPTIMIZE TABLE 语法
- HANDLER 语句
- CREATE TABLESPACE 语句
- "Session Tracker: 将 GTID 上下文信息添加到 OK 包中"
- JOIN 的 ON 子句的子查询

不兼容的特性

- TiDB 的自增列可以保证唯一, 但只能保证在单个 TiDB server 中自增, 在使用UDTS迁移时打开兼容MySQL自增模式选项, 能保证在多个 TiDB server 中自增, 但不保证自动分配的值的连续性。

在后续TiDB使用时建表语句需要添加 AUTO_ID_CACHE 1, 如:

```
CREATE TABLE `test` (  
  `id` bigint(20) unsigned NOT NULL AUTO_INCREMENT,  
  `name` varchar(20) NOT NULL,  
  `age` int(11) NOT NULL,  
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 AUTO_ID_CACHE 1;
```

- TiDB 可通过 `tidb_allow_remove_auto_inc` 系统变量开启或者关闭允许移除列的 `AUTO_INCREMENT` 属性。删除列属性的语法是：`ALTER TABLE MODIFY` 或 `ALTER TABLE CHANGE`。
- TiDB 不支持添加列的 `AUTO_INCREMENT` 属性，移除该属性后不可恢复。
- TiDB 中 `utf8` 与 `utf8mb4` 的默认排序规则分别为 `utf8_bin` 和 `utf8mb4_bin`，与 MySQL 的默认排序规则 `utf8_general_ci` 和 `utf8mb4_general_ci` 不同，MySQL 源为 5.7 时字符集为 `utf8` 排序规则为 `utf8_general_ci` 的表迁移到 TiDB 时，默认排序规则为 `utf8_bin`，字符集为 `utf8mb4` 排序规则为 `utf8_general_ci` 的表迁移到 TiDB 时，默认排序规则为 `utf8mb4_bin`，如果业务强制要求大小写不敏感，请参照以下步骤修改：

1. mysql `utf8mb4` 完全兼容 `utf8`，可以统一修改源库的表字符集为 `utf8mb4`，此操作会锁表不影响读，会阻塞写操作：

```
alter table table_name convert to character set utf8mb4 collate utf8mb4_general_ci;
```

1. 源库如果单独指定字段的字符集为 `utf8`，也需要单独修改字段的字符集为 `utf8mb4` 排序规则为 `utf8mb4_general_ci`：

```
alter table table_name modify column_name varchar(200) character set utf8mb4 collate utf8mb4_general_ci;
```

1. 设置目标 TiDB `default_collation_for_utf8mb4` 参数值为 `utf8mb4_general_ci`；

TiDB serverless 版本执行 `set global default_collation_for_utf8mb4="utf8mb4_general_ci"`；

TiDB 固定规格版本在参数管理中修改 `default_collation_for_utf8mb4` 参数值为 `utf8mb4_general_ci`；

- TiDB 默认单条 SQL 使用内存上限最大为 1G，可以通过修改 `tidb_mem_quota_query` 调整，默认当 SQL 语句超过 1G 时，会报错：

```
INSERT INTO db1.tbl_test01 SELECT * FROM db2.tbl_test02;
```

数据量较小的情况下 TiDB 可以执行，但是当事务占用内存超过 1G，TiDB 会取消 SQL 执行并报错：

Your query has been cancelled due to exceeding the allowed memory limit for a single SQL query. Please try narrowing your query scope or increase the `tidb_mem_quota_query` limit and try again.

导入较大数据量时,请使用UDTS或者执行脚本分批次导入。

- TiDB 的lower_case_table_names值为2,并且只能为2,意思是表名和字段名默认大小写不敏感。如果源MySQL lower_case_table_names值为0(大小写敏感),并且存在表名相同但是大小写不同的表,迁移会出错,需要提前在源库修改此类表名。

MySQL 迁移到 ClickHouse

UDTS 支持从 MySQL 迁移到 ClickHouse。MySQL 支持版本有 MySQL(包含Percona版)5.5/5.6/5.7/8.0; MariaDB 10.1.2 及以上,以及PolarDB(MySQL兼容版本)。ClickHouse 支持版本 21.3 至 22.8。

功能限制

1. 支持单库迁移,可迁移整库或指定表,不支持迁移存储过程、触发器、视图等。
2. 不支持迁移 DDL 语句,不支持 delete 语句。
3. 增量/全+增迁移时,源库需要开启 binlog,且格式设置为ROW, image设置为FULL。

查询方式:

```
show global variables like 'binlog_format';  
show global variables like 'binlog_row_image';
```

设置方式:

```
set global binlog_format = "ROW" ;  
set global binlog_row_image = "FULL" ;
```

4. 目标库仅支持集群版的 ClickHouse,迁移库表也仅支持集群表。
5. 选择“手动创建库表”时,UDTS 将不会在目标库创建库表,请您在目标库中提前创建待迁移库表,再启动迁移任务。
6. 选择“自动创建库表”时,如果待迁移的库表在目标库中不存在,UDTS 将在目标库中自动创建待迁移的库表;如果待迁移的库表在目标库中已经存在,UDTS 不会重新创建表结构。

7. 选择“自动创建库表”时,要求待迁移的库表必须有主键或唯一索引,如果不满足要求,请选择“手动创建库表”,并在目标库中提前创建待迁移库表。
8. 增量任务只能选择“手动创建库表”。
9. 源库不能存在同名但大小写不一致的库或表,否则同步可能会异常,建议全部采用小写格式。
10. 源库不能有超过4G的binlog文件,建议大小是1G,否则同步会出错或失败。
11. 不支持 MySQL 8.0 的新特性 binlog 事务压缩 Transaction_payload_event。使用 binlog 事务压缩有导致上下游数据不一致的风险。

迁移内容

迁移内容	说明
迁移结构	1. Database、Table 结构及数据 2. 迁移开始前不会清理目标库已存在的库表,如果自动建表不满足要求,客户可以提前在目标库手动建表
迁移范围	仅迁移创建任务时可以查到的库表,任务运行中新增加的表暂时不会自动迁移
DDL	不支持
DML	insert/update

填写表单

数据源表单

参数名	说明
地址类型	支持内网地址,外网地址,专线地址三种方式。内网地址需要填写VPC和子网信息;外网地址支持IP和域名两种方式;专线地址既支持IP,也支持域名,如果使用域名需要用户网络有外网出口。

端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
数据库名	MySQL数据库名称。所有库传输请填写 *；指定一个数据库传输，请填写数据库名；指定多个数据库传输，依次输入多个数据库名，库名之间使用英文逗号隔开。(如果数据库名称中包含空格则无法做增量迁移)
表名	MySQL传输表名。只有当“数据库名”为指定一个数据库时有效。若不填，默认为迁移指定库中的所有表；指定一张表传输，请填写表名；指定多张表传输，依次输入多张表名，表名之间使用英文逗号隔开
最大速率	外网/专线的速率范围为 1-56 MB/s；内网的速率范围为 1-1024 MB/s
Nolocks	默认关闭,对于无法获取 SUPER 权限的友商RDS服务需要开启,UDB获取 SUPER 权限详见FAQ

传输目标表单

参数名	说明
TCP 地址	ClickHouse 集群的 TCP 连接地址,默认端口为 9000 ,多个地址用","相连,示例:192.168.1.100:9000,192.168.1.120:9000
HTTP 地址	ClickHouse 集群的 HTTP 连接地址,默认端口为 8123 ,多个地址用","相连,示例:192.168.1.100:8123,192.168.1.120:8123
用户名	ClickHouse 连接用户名
密码	ClickHouse 数据库对应用户密码

MySQL 迁移到 Kafka

UDTS 支持从 MySQL 迁移到 Kafka。MySQL 支持版本有 MySQL(包含Percona版)5.6/5.7/8.x。Kafka 支持版本 2.x:包括 2.0、2.1、2.2、2.3、2.4、2.5、2.6、2.7、2.8 等。

1. 功能限制

1.1 源MySQL限制

1. 增量/全+增迁移时,源库需要开启binlog,且格式设置为ROW, image设置为FULL。

查询方式:

```
show global variables like 'binlog_format';
```

```
show global variables like 'binlog_row_image';
```

设置方式:

```
set global binlog_format = "ROW" ;
```

```
set global binlog_row_image = "FULL" ;
```

2. 源库不能存在同名但大小写不一致的库或表,否则同步可能会异常,建议全部采用小写格式。
3. 源库不能有超过4G的binlog文件,建议大小是1G,否则同步会出错或失败。
4. 不支持 MySQL 8.0 的新特性 binlog 事务压缩 Transaction_payload_event。使用 binlog 事务压缩有导致上下游数据不一致的风险。

1.2 目标Kafka限制

1. 设置 `auto.create.topics.enable` 为true。
2. 设置 `delete.topic.enable` 为true

2. 迁移内容

迁移内容	说明
迁移结构	Database、Table 结构及数据
迁移范围	仅迁移创建任务时可以查到的库表，任务运行中新增的表暂时不会自动迁移
DDL	CREATE、ALTER、DROP 等语句
DML	Snapshot/insert/update/delete

3. 表单填写

数据源表单

参数名	说明
地址类型	支持内网地址, 外网地址, 专线地址三种方式。内网地址需要填写VPC和子网信息; 外网地址支持IP和域名两种方式; 专线地址既支持IP, 也支持域名, 如果使用域名需要用户网络有外网出口。

端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
数据库名	MySQL数据库名称。所有库传输请填写 *；指定一个数据库传输，请填写数据库名；指定多个数据库传输，依次输入多个数据库名，库名之间使用英文逗号隔开。(如果数据库名称中包含空格则无法做增量迁移)
表名	MySQL传输表名。只有当“数据库名”为指定一个数据库时有效。若不填，默认为迁移指定库中的所有表；指定一张表传输，请填写表名；指定多张表传输，依次输入多张表名，表名之间使用英文逗号隔开
Nolocks	默认关闭,对于无法获取 SUPER 权限的友商RDS服务需要开启,UDS获取 SUPER 权限详见FAQ

传输目标表单

参数名	说明
内网地址	Kafka集群的连接地址,示例:92.168.1.10:9092,192.168.1.11:9092,192.168.1.12:9093
最大速率	Kafka传输限速值,调整传输速率
Topic前缀	UDTS迁移MySQL到Kafka会在目标Kafka上创建对应的Topic,规则是每张表创建一个Topic,每个Topic都会以此参数作为前缀
默认分区数量	UDTS迁移MySQL到Kafka会在目标Kafka上创建对应Topic的默认分区数量

4. Kafka传输数据格式

UDTS 迁移到Kafka的内容,格式为 Debezium JSON 格式,下游同步可使用Flink CDC ,数据格式指定为 debezium-json。

MySQL 迁移到 PostgreSQL

UDTS 支持从 MySQL 迁移到 MAXIR。MySQL 支持版本有 MySQL(包含Percona版) 5.5/5.6/5.7/8.x。PostgreSQL 支持版本 9.4到16.x。

1. 功能限制

1.1 源MySQL限制

1. 增量/全+增迁移时,源库需要开启binlog,且格式设置为ROW, image设置为FULL。

查询方式:

```
show global variables like 'binlog_format';
```

```
show global variables like 'binlog_row_image';
```

设置方式:

```
set global binlog_format = "ROW" ;
```

```
set global binlog_row_image = "FULL" ;
```

1.1.2 源待迁移表必须要有主键。

1.1.3 源库不能存在同名但大小写不一致的库或表,否则同步可能会异常,建议全部采用小写格式。

1.1.4 源库不能有超过**4G**的binlog文件，建议大小是**1G**，否则同步会出错或失败。

1.1.5 不支持 **MySQL 8.0** 的新特性 **binlog** 事务压缩 **Transaction_payload_event**。使用 **binlog** 事务压缩有导致上下游数据不一致的风险。

2. 迁移内容

迁移内容	说明
迁移结构	1. Database、Table 结构及数据 2. 迁移开始前不会清理目标库已存在的库表, 如果自动建表不满足要求, 客户可以提前在目标库手动建表
迁移范围	仅迁移创建任务时可以查到的库表, 任务运行中新增的表暂时不会自动迁移
DDL	不支持
DML	insert/update/delete

3. 表单填写

数据源表单

参数名	说明
地址类型	支持内网地址, 外网地址, 专线地址三种方式。内网地址需要填写VPC和子网信息; 外网地址支持IP和域名两种方式; 专线地址既支持IP, 也支持域名, 如果使用域名需要用户网络有外网出口。

端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
数据库名	MySQL数据库名称。仅支持单库迁移
表名	MySQL传输表名。只有当“数据库名”为指定一个数据库时有效。若不填,默认为迁移指定库中的所有表;指定一张表传输,请填写表名;指定多张表传输,依次输入多张表名,表名之间使用英文逗号隔开
最大速率	外网/专线的速率范围为 1-56 MB/s; 内网的速率范围为 1-1024 MB/s

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	PostgreSQL 连接端口
用户名	PostgreSQL 连接用户名
密码	PostgreSQL 数据库对应用户密码

4. 权限要求

类别	权限说明
----	------

源MySQL	SELECT, RELOAD, SHOW DATABASES, REPLICATION SLAVE, REPLICATION CLIENT
目标PostgreSQL	必须拥有待迁移库的 owner 权限, 或 rolcreatedb 权限

5. 数据类型映射

MySQL数据类型	PostgreSQL数据类型
BIT(1)	BOOL
TINYINT(1)	
TINYINT	INT2
TINYINT UNSIGNED	INT2
SMALLINT	
SMALLINT UNSIGNED	
MEDIUMINT	
MEDIUMINT UNSIGNED	INT4
INT	
INTEGER	
YEAR	

INT UNSIGNED	
INTEGER UNSIGNED	INT8
BIGINT	
BIGINT UNSIGNED	NUMERIC(20,0)
DECIMAL(p, s)	
DECIMAL(p, s) UNSIGNED	NUMERIC(p,s)
NUMERIC(p, s)	
NUMERIC(p, s) UNSIGNED	
FLOAT	
FLOAT UNSIGNED	FLOAT4
DOUBLE	
DOUBLE UNSIGNED	FLOAT8
REAL	
REAL UNSIGNED	

CHAR	
VARCHAR	
TINYTEXT	
MEDIUMTEXT	
TEXT	VARCHAR/TEXT
LONGTEXT	
ENUM	
JSON	
ENUM	
DATE	DATE
TIME(s)	TIME(s)
DATETIME	
TIMESTAMP(s)	TIMESTAMP(s)

BINARY	
VARBINARY	
BIT(p)	
TINYBLOB	
MEDIUMBLOB	BYTEA
BLOB	
LOB	
LONGLOB	
GEOMETRY	

MySQL 迁移到 MAXIR

UDTS 支持从 MySQL 迁移到 MAXIR。MySQL 支持版本有 MySQL(包含Percona版) 5.5/5.6/5.7/8.x。

1. 功能限制

1.1 源MySQL限制

1.1.1 增量/全+增迁移时，源库需要开启binlog，且格式设置为ROW，image设置为FULL。

查询方式:

```
show global variables like 'binlog_format';  
show global variables like 'binlog_row_image';
```

设置方式:

```
set global binlog_format = "ROW" ;  
set global binlog_row_image = "FULL" ;
```

1.1.2 源待迁移表必须要有主键。

1.1.3 源库不能存在同名但大小写不一致的库或表，否则同步可能会异常，建议全部采用小写格式。

1.1.4 源库不能有超过**4G**的binlog文件，建议大小是**1G**，否则同步会出错或失败。

1.1.5 不支持 **MySQL 8.0** 的新特性 **binlog** 事务压缩 **Transaction_payload_event**。使用 **binlog** 事务压缩有导致上下游数据不一致的风险。

1.1.6 源库表名或者库名不能包含特殊字符'**'**，否则迁移失败。

1.2 目标MAXIR限制

1.2.1 MAXIR建表时需要指定**DISTRIBUTED BY** 和**CLUSTER BY** 列属性，例如：

```
CREATE TABLE trade (  
  id bigint,  
  "date" date,  
  shopid int NOT NULL,  
  sku varchar NOT NULL,  
  price decimal(18, 4),  
  primary key(date, id)  
) DISTRIBUTED BY (id) CLUSTER BY (date, id);
```

a. **CLUSTER BY** 列建议需要跟primary key 相同,或者是primary key的一部分,例如是 date 或者是 {date, id}

b. **DISTRIBUTED BY** 的列需要是primary key中一列,并且需要是 cardinality 大的列,在本例中只能是id,不能是date

1.2.2 UDS迁移到MAXIR，如果目标库中表不存在则会自动建表，建表规则如下，客户需要提前评估，如果不满足业务要求客户可以在目标库中提前建表：

1. CLUSTER BY 会和表的主键保持一致。
2. DISTRIBUTED BY 会使用主键，如果是联合主键则会使用第一个字段。
3. 自动创建的表结构会去掉字段的 NOT NULL 属性。

2. 迁移内容

2.1 全量迁移

1. 全量迁移会迁移任务指定的源库中的库表结构和数据到目标库。
2. 表结构中索引仅迁移主键信息(MAXIR仅支持主键)。
3. 全量迁移会清理指定库表在目标库中的数据,但不会删除表结构,如果需要重建表结构需要客户自己在目标库删除或重建。

2.2 增量迁移

1. 增量迁移会实时解析源库binlog,将增量数据同步到目标库。
2. DML操作支持INSERT、UPDATE、DELETE。
3. DDL操作支持 1. CREATE/DROP DATABASE; 2. CREATE/DROP/RENAME TABLE; 3. ALTER TABLE ADD/DROP/MODIFY/CHANGE COLUMN。ALTER TABLE时,不支持同时修改数据类型和字段名。
4. 增量迁移会将源库的binlog位点信息记录在UDTS的任务中,下次增量迁移时会从上次的binlog位点开始解析。

3. 表单填写

数据源表单

参数名	说明
地址类型	支持内网地址, 外网地址, 专线地址三种方式。内网地址需要填写VPC和子网信息; 外网地址支持IP和域名两种方式; 专线地址既支持IP, 也支持域名, 如果使用域名需要用户网络有外网出口。
端口	MySQL连接端口
用户名	MySQL连接用户名
密码	MySQL数据库对应用户密码
数据库名	MySQL数据库名称。所有库传输请填写 *; 指定一个数据库传输, 请填写数据库名; 指定多个数据库传输, 依次输入多个数据库名, 库名之间使用英文逗号隔开。(如果数据库名称中包含空格则无法做增量迁移)
表名	MySQL传输表名。只有当“数据库名”为指定一个数据库时有效。若不填, 默认为迁移指定库中的所有表; 指定一张表传输, 请填写表名; 指定多张表传输, 依次输入多张表名, 表名之间使用英文逗号隔开
最大速率	外网/专线的速率范围为 1-56 MB/s; 内网的速率范围为 1-1024 MB/s

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	MAXIR 连接端口
用户名	MAXIR 连接用户名
密码	MAXIR 数据库对应用户密码

数据库名	<p>迁移到MAXIR 目标库中的 Database名称。</p> <p>源MySQL的多个Database名称,会分别创建为目标MAXIR指定Database下的Schema。例如: 源MySQL迁移 A.t1,A.t2,B.t1,B.t2 到目标MAXIR的 target数据库下,则在MAXIR中看到的迁移结果为: target.A.t1,target.A.t2,target.B.t1,target.B.t2</p>
------	--

4. 权限要求

类别	权限说明
源MySQL	SELECT, RELOAD, SHOW DATABASES, REPLICATION SLAVE, REPLICATION CLIENT
目标MAXIR	必须拥有待迁移库的 owner 权限,或 rolcreatedb 权限

5. 数据类型映射

MySQL数据类型	MAXIR数据类型
BIT(1)	BOOL
TINYINT SMALLINT	INT2

MEDIUMINT INT INTEGER YEAR	INT4
TINYINT UNSIGNED SMALLINT UNSIGNED MEDIUMINT UNSIGNED INT UNSIGNED INTEGER UNSIGNED BIGINT UNSIGNED	NUMERIC(20, 0)
BIGINT	INT8
DECIMAL(p, s) DECIMAL(p, s) UNSIGNED NUMERIC(p, s) NUMERIC(p, s) UNSIGNED	NUMERIC(p,s)
FLOAT FLOAT UNSIGNED	FLOAT4

DOUBLE	
DOUBLE UNSIGNED	FLOAT8
REAL	
REAL UNSIGNED	
CHAR	VARCHAR
VARCHAR	
TINYTEXT	
MEDIUMTEXT	
TEXT	TEXT
LONGTEXT	
ENUM	
JSON	
DATE	DATE
TIME	TIME
DATETIME	TIMESTAMP
TIMESTAMP	TIMESTAMPZ

BINARY	
VARBINAR	
BIT(p)	
TINYBLOB	
MEDIUMBLOB	BYTEA(但不支持写入数据)
BLOB	
LOB	
LONGBLOB	
GEOMETRY	

MySQL双向同步

为满足用户多地域甚至跨云容灾等需求,UDTS推出了MySQL-MySQL双向同步的功能。

当前双向同步相当于是增量同步的模式,并不处理全量数据传输,适用于双库初始数据一致的情况。

创建双向同步任务

选择“双向同步”任务类型

填写双库信息(当前不支持2个数据库都是外网数据库的场景)

[<](#) [数据传输服务](#) / [创建任务](#)

基本设置

任务类型 [?](#)

全量任务

增量任务

全量+增量任务

双向同步

数据源类型

MySQL

目标类型

MySQL

最大重试次数 [?](#)

0

注意事项

- 1 双库需要开启binlog, 格式为ROW, MODE 为FULL
- 2 库内的表必须都要有主键。
- 3 数据库需要开启GTID模式, 但非必填。如果不填写会从当前位置开始同步。
- 4 双向同步的两个表如果是主键自增, 需要两边设置对应的自增步进值, 比如一边使用奇数 ID, 一边使用偶数 ID, 这样数据记录创建时, 不会出现同样 ID 的数据, 或者使用UUID/snowflake, 来确保两表的ID不会重复。
- 5 因为同步可能会受到网络、锁等因素造成延迟, 为保障同步数据的一致性, 您需要确保同一个主键的记录只在双向同步的一个节点进行更新。
- 6 数据同步时, 请勿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。
- 7 为保障双向同步链路的稳定性, 您需要确保DDL更新只能在其中一个同步方向执行。
- 8 目前双向同步版本仅支持MySQL5.6/5.7, 且需要源库和目标库版本一致。

SQL Server 迁移到 SQL Server

UDTS 支持 SQL Server 2008及以后版本之间的全量与全+增迁移任务。

功能限制

1. 支持单库迁移,可迁移整库或指定表,不支持迁移存储过程、触发器、视图等。
2. 全+增迁移不支持 DDL。
3. 全+增迁移时,源库和待迁移的表需要开启cdc功能。

```
-- 指定库名
use dbname
-- 开启cdc功能
exec sys.sp_cdc_enable_db
-- 查看 dbname 库是否开启cdc,返回值为 1 表示已开启
select is_cdc_enabled from sys.databases where name = "dbname";
-- 给 dbo.tablename 表开启cdc
exec sys.sp_cdc_enable_table @source_schema = 'dbo',@source_name = 'tablename',@role_name = null;
-- 查看 dbo.tablename 表是否开启cdc,返回值为 1 表示已开启
select is_tracked_by_cdc from sys.tables where name = 'tablename'
-- 修改 cdc 数据保留时间,至少修改为1440分钟(1天)或以上,建议14400分钟(7天)。
EXECUTE sys.sp_cdc_change_job @job_type = N'cleanup', @retention = 14400;
```

- 待迁移的表必须存在主键,建议采用 int 类型作为主键。主键为 string 类型且数据量较大时,迁移速度会受到影响。
- 待迁移的表中主键的自增属性无法迁移
- 待迁移的表中 rowversion 字段无法迁移

迁移内容

迁移内容	说明
迁移结构	1. Database、Table 结构(索引信息目前只迁移主键)及数据 2. 迁移开始前不会清理目标库已存在的库表,如果自动建表不满足要求,客户可以提前在目标库手动建表
迁移范围	仅迁移创建任务时可以查到的库表,任务运行中新增加的表暂时不会自动迁移
DDL	不支持
DML	insert/update/delete

填写表单

数据源表单

参数名	说明
地址	支持内网地址,专线地址两种方式。内网地址需要填写VPC和子网信息;专线地址既支持IP,也支持域名,如果使用域名需要用户网络有外网出口
端口	SQL Server 连接端口
用户名	SQL Server 连接用户名

密码	SQL Server 数据库对应用户密码
数据库名	SQL Server 数据库名, 仅支持单库迁移
表名	SQL Server 传输表名, 表名之间使用英文逗号隔开。示例: dbo.tablename1, dbo.tablename2, schema 为 dbo 时可以省略, 只填写 tablename, 示例: tablename1, tablename2

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	SQL Server 连接端口
用户名	SQL Server 连接用户名
密码	SQL Server 数据库对应用户密码

SQL Server 迁移到 MySQL

UDTS 支持 SQL Server 2008及以后版本到 MySQL 5.5及以后各版本的全量与全+增迁移任务。

功能限制

1. 支持单库迁移,可迁移整库或指定表,不支持迁移存储过程、触发器、视图等。
2. 全+增迁移不支持 DDL。
3. 全+增迁移时,源库和待迁移的表需要开启cdc功能。

```
-- 指定库名
use dbname
-- 开启cdc功能
exec sys.sp_cdc_enable_db
-- 查看 dbname 库是否开启cdc,返回值为 1 表示已开启
select is_cdc_enabled from sys.databases where name = "dbname";
-- 给 dbo.tablename 表开启cdc
exec sys.sp_cdc_enable_table @source_schema = 'dbo',@source_name = 'tablename',@role_name = null;
-- 查看 dbo.tablename 表是否开启cdc,返回值为 1 表示已开启
select is_tracked_by_cdc from sys.tables where name = 'tablename'
-- 修改 cdc 数据保留时间,至少修改为1440分钟(1天)或以上,建议14400分钟(7天)。
EXECUTE sys.sp_cdc_change_job @job_type = N'cleanup', @retention = 14400;
```

- 待迁移的表必须存在主键或唯一索引,建议采用 int 类型作为主键。主键为 string 类型且数据量较大时,迁移速度会受到影响。
- 待迁移的表中主键的自增属性无法迁移
- 待迁移的表中 rowversion 字段无法迁移
- 目标库为 MySQL 时,不迁移源库的 schema 名,如果源库中存在不同 schema 中的同名表,只会迁移其中一张表。

迁移内容

迁移内容	说明
迁移结构	1. Database、Table 结构及数据 2. 迁移开始前不会清理目标库已存在的库表,如果自动建表不满足要求,客户可以提前在目标库手动建表
迁移范围	仅迁移创建任务时可以查到的库表,任务运行中新增的表暂时不会自动迁移
DDL	不支持
DML	insert/update/delete

填写表单

数据源表单

参数名	说明
地址	支持内网地址,专线地址两种方式。内网地址需要填写VPC和子网信息;专线地址既支持IP,也支持域名,如果使用域名需要用户网络有外网出口
端口	SQL Server 连接端口

用户名	SQL Server 连接用户名
密码	SQL Server 数据库对应用户密码
数据库名	SQL Server 数据库名, 仅支持单库迁移
表名	SQL Server 传输表名, 表名之间使用英文逗号隔开。示例: dbo.tablename1, dbo.tablename2, schema 为 dbo 时可以省略, 只填写 tablename, 示例: tablename1, tablename2

传输目标MySQL表单

参数名	说明
地址类型	目标暂时只支持内网
端口	MySQL 连接端口
用户名	MySQL 连接用户名
密码	MySQL 数据库对应用户密码

SQL Server 迁移到 Kafka

UDTS 支持从 SQL Server 迁移到 Kafka。SQL Server 支持版本为 SQL Server 2008 R2及以上。Kafka 支持版本 2.x:包括 2.0、2.1、2.2、2.3、2.4、2.5、2.6、2.7、2.8 等。

1. 功能限制

1.1 源SQL Server限制

1. 源库和待迁移的表需要开启cdc功能。

```
-- 指定库名
use dbname

-- 开启cdc功能
exec sys.sp_cdc_enable_db

-- 查看 dbname 库是否开启cdc,返回值为 1 表示已开启
select is_cdc_enabled from sys.databases where name = "dbname";

-- 给 dbo.tablename 表开启cdc
exec sys.sp_cdc_enable_table @source_schema = 'dbo',@source_name = 'tablename',@role_name = null;

-- 查看 dbo.tablename 表是否开启cdc,返回值为 1 表示已开启
select is_tracked_by_cdc from sys.tables where name = 'tablename'

-- 修改 cdc 数据保留时间,至少修改为1440分钟(1天)或以上,建议14400分钟(7天)。
EXECUTE sys.sp_cdc_change_job @job_type = N'cleanup', @retention = 14400;
```

1.2 目标Kafka限制

1. 设置 `auto.create.topics.enable` 为true。
2. 设置 `delete.topic.enable` 为true

2. 迁移内容

迁移内容	说明
迁移结构	Database、Table 结构及数据
迁移范围	仅迁移创建任务时可以查到的库表，任务运行中新增的表暂时不会自动迁移
DDL	CREATE、ALTER、DROP 等语句
DML	Snapshot/insert/update/delete

3. 表单填写

数据源表单

参数名	说明
地址	支持内网地址, 专线地址两种方式。内网地址需要填写VPC和子网信息; 专线地址既支持IP, 也支持域名, 如果使用域名需要用户网络有外网出口

端口	SQL Server 连接端口
用户名	SQL Server 连接用户名
密码	SQL Server 数据库对应用户密码
数据库名	SQL Server 数据库名, 仅支持单库迁移
表名	SQL Server 传输表名, 表名之间使用英文逗号隔开。示例: dbo.tablename1, dbo.tablename2, schema 为 dbo 时可以省略, 只填写 tablename, 示例: tablename1, tablename2

传输目标表单

参数名	说明
内网地址	Kafka集群的连接地址, 示例: 92.168.1.10:9092, 192.168.1.11:9092, 192.168.1.12:9093
最大速率	Kafka传输限速值, 调整传输速率
Topic前缀	UDTS迁移MySQL到Kafka会在目标Kafka上创建对应的Topic, 规则是每张表创建一个Topic, 每个Topic都会以此参数作为前缀
默认分区数量	UDTS迁移MySQL到Kafka会在目标Kafka上创建对应Topic的默认分区数量

4. Kafka传输数据格式

UDTS 迁移到Kafka的内容, 格式为 Debezium JSON 格式, 下游同步可使用Flink CDC, 数据格式指定为 debezium-json。

TiDB

UDTS支持 TiDB 作为数据传输源/目标,支持版本 3.* 以上

TiDB 填写表单

参数名	说明
IP	内网地址需要填写VPC和子网信息。
端口	连接端口
用户名	连接用户名
密码	用户名对应密码
数据库名	如果是单个数据库,请填写数据库名;如果是多个数据库,填写以逗号分割的数据库列表;如果是所有数据库,请填写*
表名	传输表名,只有迁移单个数据库时有效。如果是单个表,请填写表名;如果是多个表,填写以逗号分割的列表;若不填或者填写*,则整库迁移

预检查项

以TiDB为源库时,TiDB参数tikv_gc_life_time建议大于1小时,以避免迁移过程中发生 GC 导致迁移失败。

查询方式:

```
select VARIABLE_NAME, VARIABLE_VALUE from mysql.tidb where VARIABLE_NAME = "tikv_gc_life_time"
```

设置方式:

```
set global tidb_gc_life_time = '24h';
```

CSV源

UDTS可以将US3上保存的CSV文件,传输至MySQL的数据类型。

CSV源填写表单

参数名	说明
保存存量数据	开启保存存量数据,服务不会清空目标库存量数据。如果没开启,服务开始前会清空目标库的存量数据。
重复数据处理类型	对于数据目标中,主键数据重复的处理方式。"忽略",不会进行覆盖;"替换",导入的数据或后入库的数据会覆盖先前的数据;"更新",对于数据进行赋值操作。
URL	CSV所在的公网链接
列顺序调整	可依次填写数据目标的字段名,UDTS会根据CSV的列顺序依次匹配字段顺序用以调整传输的列顺序。
更新策略	当重复数据处理类型选择 "更新",Type:"CSVData"CSV文件列的索引号,对应的Data为整型,从0开始;"Function" SQL支持的函数如now() ; "Fixed" 固定值如"people"。
赋值策略	当重复数据处理类型选择 "更新",Type:"CSVData"CSV文件列的索引号,对应的Data为整型,从0开始;"Function" SQL支持的函数如now() ; "Fixed" 固定值如"people"。

MongoDB

UDTS 支持 MongoDB (自建/UDB MongoDB/其它云厂商提供的标准MongoDB服务) 单节点、副本集、分片集之间的相互迁移。

源目数据库版本支持 3.0 至 6.0。由于跨版本特性兼容问题,不支持跨大版本迁移。例如,从3.x迁移到5.x时,需要创建4.x版本的中转库,先从3.x迁移到4.x,再从4.x迁移到5.x。

功能限制

1. 不迁移系统内置库, config/local/admin
2. 当源或者目标MongoDB版本低于3.6时,增量迁移不支持DDL
3. 当源库为分片集群时,增量迁移不支持DDL
4. 增量迁移源库必须开启oplog

副本集的迁移

1. 当源库地址是单地址时,源库集群发生主从切换之后,任务暂时可以运行,但任务失败/停止后重启时无法启动
2. 当目标库地址是单地址时,目标库集群发生主从切换之后,任务可能运行异常,任务失败/停止后重启时无法启动
3. 源库/目标库是副本集时,推荐在地址处同时填写主备地址,如 primary:port,secondary:port。这样源库/目标库发生主从切换之后,任务会自动切换到新的primary节点继续运行
4. 源库/目标库为 UMongoDB NVMe型 时,填写访问地址中所有地址。

分片集群的迁移

1. 源库为分片集群时需要填写 mongos地址,所有分片地址,以及Config Server节点地址。
2. 目标为分片集时,迁移地址需要填写mongos路由地址
3. 在迁移前需要关闭源库Balancer
4. 清理源库中的孤儿文档,防止迁移时遇到id冲突的问题

5. 当源库和目标库都是分片集群时,那么也将默认同时迁移源库中的片键

注意事项

1. MongoOplgTs需填写UTC时间(与CST时间差8小时),例如:期望同步点为北京时间2021-03-01 20:10:10时,应填写2021-03-01T12:10:10Z

2. UDB MongoDB 副本集填写示例,填写所有地址

MongoDB

UMongoDB管理

参数模板管理

创建集群 删除 更改业务组

名称	资源ID	可用区	版本	磁盘类型	访问地址	创建时间	类型
	umongodb-rs-mgpsu8oyiz0	华北一可用区B	MongoDB 4.2	RSSD云盘	mongodb://root:***@10.9.173.190:27017,10.9.159:27017,10.9.33.150:27017?replicaSet=umongodb-rs-mgpsu8oyiz0	2023-08-11	副本集
luna5	umongodb-rs-mmfxmmuc2v	华北一可用区B	MongoDB 5.0	RSSD云盘	mongodb://root:***@10.9.36.111:2...	2023-08-15	副本集
修改名称及备注	umongodb-rs-mpf6bw5jgcd	华北一可用区B	MongoDB 3.4	RSSD云盘	mongodb://root:***@10.9.159.112:...	2023-08-17	副本集
修改名称及备注	umongodb-sc-mpj36rswx4	华北一可用区B	MongoDB 3.6	RSSD云盘	mongodb://root:***@10.9.85.40:27...	2023-08-17	分片集
修改名称及备注	umongodb-sc-mpj4mhv8aqd	华北一可用区B	MongoDB 4.2	RSSD云盘	mongodb://root:***@10.9.83.18:27...	2023-08-17	分片集
修改名称及备注	umongodb-sc-mpj5gl36xxj	华北一可用区B	MongoDB 5.0	RSSD云盘	mongodb://root:***@10.9.123.112:...	2023-08-17	分片集

3. 跨版本迁移时,请注意源库中是否存在与目标库不兼容的特性

MongoDB填写表单

源为副本集或者单节点表单填写说明

参数名	说明
地址	提供内网地址, 外网地址, 专线地址三种方式, 内网地址需要填写VPC和子网信息, 外网地址支持ip和域名两种。 副本集填写所有节点地址, 如: 192.168.1.100:27017,192.168.1.120:27017,192.168.1.130:27017
授权DB	授权数据库名, UDB MongoDB默认为admin
Collection	集合, MongoDB 文档组, 若不填, 默认为迁移指定库中的所有集合; 指定一个集合传输, 请填写集合名; 指定多个集合传输, 依次输入多个集合名, 集合名之间使用英文逗号隔开
数据库名	MongoDB数据库名称。传输全库(系统库除外) 请填写 * ; 指定一个数据库传输, 请填写数据库名; 指定多个数据库传输, 依次输入多个数据库名, 库名之间使用英文逗号隔开。
用户名	MongoDB连接用户名
密码	MongoDB数据库对应用户密码
MongoOplogTs	(仅增量任务需填写) 增量开始的oplog位置, 即增量同步点, 格式为1970-01-01T00:00:00Z (UTC time)

源为分片集表单填写说明

参数名	说明
-----	----

数据节点IP与端口	分片集源库填写所有分片的地址,每个分片填写一行,如:192.168.1.100:27017,192.168.1.120:27017,192.168.1.130:27017
<input type="checkbox"/>	多分片请换行填写,如三分片填写三行: 192.168.1.100:27017,192.168.1.120:27017,192.168.1.130:27017 192.168.1.101:27017,192.168.1.121:27017,192.168.1.131:27017 192.168.1.102:27017,192.168.1.122:27017,192.168.1.132:27017
路由节点地址	分片集群的mongos 地址, 如: 192.168.2.126:27017,192.168.2.136:27017
配置节点地址	分片集群的Config server 地址, 如: 192.168.1.200:27017,192.168.1.220:27017,192.168.1.230:27017
授权DB	授权数据库名 ,UDB MongoDB默认为admin
Collection	集合, MongoDB 文档组,若不填,默认为迁移指定库中的所有集合;指定一个集合传输,请填写集合名;指定多个集合传输,依次输入多个集合名,集合名之间使用英文逗号隔开
数据库名	MongoDB数据库名称。传输全库(系统库除外)请填写 * ;指定一个数据库传输,请填写数据库名;指定多个数据库传输,依次输入多个数据库名,库名之间使用英文逗号隔开。
用户名	MongoDB连接用户名 (相同用户名密码填写此项,各角色不同请分别填写)
密码	MongoDB数据库对应用户密码 (相同用户名密码填写此项,各角色不同请分别填写)

目标为副本集或分片集表单填写

参数名	说明
地址	提供内网地址。 副本集填写所有节点地址,如: 192.168.1.100:27017,192.168.1.120:27017,192.168.1.130:27017 分片集填写mongos路由地址, 如:192.168.1.100:27017,192.168.1.120:27017

授权DB	授权数据库名 ,UDB MongoDB默认为admin
用户名	MongoDB连接用户名
密码	MongoDB数据库对应用户密码

MongoDB

UDTS 支持将 AWS DynamoDB 迁移到 MongoDB

源 DynamoDB 填写表单

参数名	说明	示例
所在地域	DynamoDB 所在地域	us-east-2
AWS access key	AWS 授权 access key	AKIA5XXIWQJLFJSLW5DJNMF
AWS secret key	AWS 授权 secret key	wIHcUhT/ysfl8lxaw7nx/UDcbwnB0t
AWS token	AWS 授权 token	
自建时 DynamoDB 地址	自建 DynamoDB 的地址(参考备注1)	http://192.168.100.103:8000

目标 MongoDB 填写表单

参数名	说明	示例
地址	10.9.43.27	us-east-2

所在地域/VPC/子网	在控制台页面选择即可	
端口	MongoDB 连接端口	27017
用户名	MongoDB 用户名	root
密码	MongoDB 密码	123456
授权DB	MongoDB 的授权 DB, UDB 默认为 admin	admin
数据库名	DynamoDB 迁移到 MongoDB 中对应的名称(具体请参考备注2)	mydb

备注说明

1. 当源为自建时, 自建时 DynamoDB 地址 为必填, 否则 AWS access key 和 AWS secret key, 所在地域 为必填
2. 在 DynamoDB 中只有 Table, 没有 Database, 迁移到 MongoDB 时需要指定一个 DB 名, 迁移后的数据会存储在该 DB 中。

Redis

UDTS 可以将原生 Redis 或使用代理的分布式版本(包括 UCloud Redis 分布式版、Codis、Twemproxy)作为数据源/目标 进行全量+增量任务的传输。当前支持的版本为 3.0 至 6.0 各版本与 7.0 主备版。跨大版本迁移可能存在兼容性问题,建议通过中转库迁移。例如,从3.x/4.x迁移到7.x时,需要创建5.x/6.x版本的中转库,先从3.x/4.x迁移到5.x/6.x,再从5.x/6.x迁移到7.x。

功能限制

- 如果源是读写分离版,需要使用 rump 模式迁移,迁移只能做全量,不支持增量。7.0版本不支持 rump 模式

注意事项

- 对于源和目的集群节点数目不需要一致
- 需要确保源库的 repl-diskless-sync 配置为 NO
- 如果源库为分布式版(包括 UCloud Redis 分布式版、Codis、Twemproxy),需要创建多个任务,每个任务的源库填写一个分布式节点的地址,类型选择主备版。
- 目标库为 UCloud Redis 分布式版时,目标库类型选择分布式版,地址为代理地址。
- 目标库为 UCloud Redis 分布式 Cluster 版本且提供了代理地址时,目标库类型选择分布式版,地址为代理地址。
- 目标库为 UCloud Redis 分布式 Cluster 版本且没有提供代理地址时,目标库类型选择集群版,地址为分片地址,多个地址之间用";"相连。
- 目标库为集群版时,只支持迁移db0。
- UDS不会自动清理目标库中的数据,目标库中与源库中 Key 相同的数据会被覆盖,Key 不同的数据保留。
- Redis的过期key在到期的时候不会自动删除,一般通过惰性删除(访问该key的时候判断过期,过期就删除,否则不删除)和定期删除(定期迭代删除一定量的key),所以源库中会存在一定量的过期key,同步到目标库的时候是不会同步过期key的,所以会有数量不一致

Redis填写表单

参数名	说明
地址	数据库地址,只填写主(master)地址。集群模式下,多个地址用";"相连,例如:192.168.1.100:6379;192.168.1.120:6379;192.168.1.130:26379
密码	Redis密码,非必填
Redis类型	集群模式/主备模式/分布式版
VPC ID	VPC ID,redis 所属 vpc
所属子网	VPC下所属子网, redis 所属子网
是否开启Rump同步	当用户无 psync 权限时需要通过Rump同步。

断点续传

满足以下条件将自动开启断点续传功能

- 任务已运行到增量阶段
- 源目类型一致,比如主备模式传输到主备模式,集群模式传输到集群模式
- 源目均为集群模式时,源目集群节点数一致
- 源库版本为 3.0 至 6.0

PostgreSQL

UDTS 支持 PostgreSQL 作为数据传输源/目标, 支持版本9.4到16.x。

前提条件

- 增量同步时用户需要开启数据日志, wal_level需要设置为logical, max_replication_slots需要大于1
- 待迁移的表需具备主键或着唯一索引, 否则可能会产生重复数据。具体参考增量同步处理表没有主键或者没有唯一索引的情况

所需权限

类型	源库	目标库
全量	SELECT	待迁移库的 owner 权限, 或 rolcreatedb 权限
全量+增量	SELECT, REPLICATION	待迁移库的 owner 权限, 或 rolcreatedb 权限

功能限制

- 一个“全量+增量”任务只能同步一个数据库, 如果有多个数据库需要同步, 则需要为每个数据库创建任务。

- 同步对象仅支持数据表,暂时不支持DDL语句的同步。如果需要同步DDL语句,需要手动在目标库中执行对应的DDL操作,然后重启数据同步任务。
- 增量同步不支持GEOMETRY、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。
- 由于数据库自身的特性,不支持从高版本迁移到低版本。
- 增量阶段不支持处理generated列。

注意事项

增量同步处理表没有主键或者没有唯一索引的情况

- 如果待迁移的表即没有主键也没有唯一索引,需要执行 ALTER TABLE tb_xxx REPLICA IDENTITY FULL, 否则无法进行增量迁移。
- 如果待迁移的表没有主键但是有唯一索引比如idx_unq_name, 这种情况有两种处理方式:

1. 与没有主键的处理方式一样,执行 ALTER TABLE tb_xxx REPLICA IDENTITY FULL, 这种方式复制效率比较低。
2. 手动指定复制的唯一索引, ALTER TABLE tb_xxx REPLICA IDENTITY USING INDEX idx_unq_name, 这种方式复制效率较高。

- 如果待迁移的表没有主键但是有唯一索引,并且在任务的运行过程中更换了用于复制的唯一索引,需要重启任务否则任务可能会失败。

增量同步进度保存与复制槽 slot 说明

- 增量同步期间,UDTS会在目标库中创建public.udts_pgsync_progress的数据表,用于记录同步的进度等信息,同步过程中请勿删除,否则会导致任务异常。
- 增量同步期间,UDTS会在源库中创建前缀为udts_的replication slot用于复制数据。
- 客户暂停同步任务后,由于源库中存在前缀为udts_的replication slot用于复制数据,会导致wal清理不掉持续占用磁盘,如果需要长时间停止任务建议删除此slot,并删除任务。
- 客户删除任务后,需要手动在源库删除对应 slot,操作步骤如下:

操作步骤:

1. 例如使用UDTS增量迁移的数据库为 db_service_car。
2. 执行 `select * from pg_replication_slots` ,返回结果如下:

```
-----  
| slot_name | plugin | slot_type | datoid | database | active |  
-----
```

```
| udts_dd27ef9195294b49a5d424eda8f399f7 | test_decoding | logical | 4839920 | db_service_car | f |  
| udts_050a1f4b1cc84b4bb6460e5477ec3d79 | test_decoding | logical | 2695822 | db_service_xxx | t |  
-----
```

3. 找到 database 为 db_service_car 并且前缀为 udts 的 slot_name。
4. 执行 `select pg_drop_replication_slot('udts_dd27ef9195294b49a5d424eda8f399f7');`
5. 再执行一次 `select * from pg_replication_slots`, 确认已删除。

PostgreSQL 填写表单

数据源表单

参数名	说明
地址类型	提供内网地址,外网地址,专线地址三种方式,内网地址需要填写VPC和子网信息,外网地址支持ip和域名两种
端口	PostgreSQL 连接端口
用户名	PostgreSQL 连接用户名
密码	PostgreSQL 数据库对应用户密码
数据库名	PostgreSQL 待迁移数据库名称
表名	PostgreSQL 传输表名,可选项。若不填,整库迁移;若填写,迁移指定的表。具体参考 表单表名填写规则
最大速率	外网/专线的速率范围为 1-256 MB/s,默认40 MBps (即 320 Mbps);内网的速率范围为 1-1024 MB/s,默认 80 MBps(即 640 Mbps)

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	PostgreSQL连接端口
用户名	PostgreSQL连接用户名
密码	PostgreSQL数据库对应用户密码
最大速率	内网的速率范围为 1-1024 MB/s,默认 80 MBps(即 640 Mbps)

表单表名填写规则

- 1. 表名仅支持[a-z][A-Z][0-9]以及_等字符。
- 2. 多个表名以,隔开
- 3. 表名没有添加schema的默认追加 public. schema。
- 4. 表名填写 public.tb_test01,schema01.tb_test02,表示只迁移 public.tb_test01,schema01.tb_test02 这两张表。

PostgreSQL 迁移到MAXIR

UDTS 支持 PostgreSQL 作为数据传输源传输到MAXIR,支持版本9.4到16.x。

前提条件

- 增量同步时用户需要开启数据日志, wal_level需要设置为logical, max_replication_slots需要大于1
- 待迁移的表需具备主键或着唯一索引,否则可能会产生重复数据。具体参考增量同步处理表没有主键或者没有唯一索引的情况

所需权限

类型	源库	目标库
全量	SELECT	待迁移库的 owner 权限,或 rolcreatedb 权限
全量+增量	SELECT,REPLICATION	待迁移库的 owner 权限,或 rolcreatedb 权限

功能限制

- 源待迁移表必须要有主键。

- 一个“全量+增量”任务只能同步一个数据库,如果有多个数据库需要同步,则需要为每个数据库创建任务。
- 同步对象仅支持数据表,暂时不支持DDL语句的同步。如果需要同步DDL语句,需要手动在目标库中执行对应的DDL操作,然后重启数据同步任务。
- 增量同步不支持GEOMETRY、TSQUERY、TSVECTOR、TXID_SNAPSHOT类型的数据同步。
- 由于数据库自身的特性,不支持从高版本迁移到低版本。
- 增量阶段不支持处理generated列。

注意事项

增量同步进度保存与复制槽 slot 说明

- 增量同步期间,UDTS会在目标库中创建public.udts_pgsync_progress的数据表,用于记录同步的进度等信息,同步过程中请勿删除,否则会导致任务异常。
- 增量同步期间,UDTS会在源库中创建前缀为udts_的replication slot用于复制数据。
- 客户暂停同步任务后,由于源库中存在前缀为udts_的replication slot用于复制数据,会导致wal清理不掉持续占用磁盘,如果需要长时间停止任务建议删除此slot,并删除任务。
- 客户删除任务后,需要手动在源库删除对应 slot,操作步骤如下:

操作步骤:

1. 例如使用UDTS增量迁移的数据库为 db_service_car。
2. 执行 `select * from pg_replication_slots` ,返回结果如下:

```
-----  
| slot_name | plugin | slot_type | datoid | database | active |  
-----
```

```
| udts_dd27ef9195294b49a5d424eda8f399f7 | test_decoding | logical | 4839920 | db_service_car | f |
```

```
| udts_050a1f4b1cc84b4bb6460e5477ec3d79 | test_decoding | logical | 2695822 | db_service_xxx | t |
```

3. 找到 database 为 db_service_car 并且前缀为 udts 的 slot_name。
4. 执行 `select pg_drop_replication_slot('udts_dd27ef9195294b49a5d424eda8f399f7');`
5. 再执行一次 `select * from pg_replication_slots`, 确认已删除。

PostgreSQL 填写表单

数据源表单

参数名	说明
地址类型	提供内网地址, 外网地址, 专线地址三种方式, 内网地址需要填写VPC和子网信息, 外网地址支持ip和域名两种
端口	PostgreSQL 连接端口
用户名	PostgreSQL 连接用户名
密码	PostgreSQL 数据库对应用户密码
数据库名	PostgreSQL 待迁移数据库名称
表名	PostgreSQL 传输表名, 可选项。若不填, 整库迁移; 若填写, 迁移指定的表。具体参考 表单表名填写规则

最大速率	外网/专线的速率范围为 1-256 MB/s, 默认40 MBps (即 320 Mbps); 内网的速率范围为 1-1024 MB/s, 默认 80 MBps(即 640 Mbps)
------	---

传输目标表单

参数名	说明
地址类型	目标暂时只支持内网
端口	MAXIR 连接端口
用户名	MAXIR 连接用户名
密码	MAXIR 数据库对应用户密码
数据库名	迁移到MAXIR 目标库中的 Database名称。

表单表名填写规则

- 1. 表名仅支持[a-z][A-Z][0-9]以及_等字符。
- 2. 多个表名以,隔开
- 3. 表名没有添加schema的默认追加 public. schema。
- 4. 表名填写 public.tb_test01,schema01.tb_test02,表示只迁移 public.tb_test01,schema01.tb_test02 这两张表。

数据类型映射

PostgreSQL数据类型	MAXIR数据类型
bigint	bigint
integer	integer
smallint	smallint
date	date
double precision	double precision
real	real
boolean	boolean
text	text
character character varying	varchar
numeric	numeric/numeric(38,10)

money	
cidr	
inet	
macaddr	text
uuid	
xml	
json	
jsonb	
BIGINT	INT8
bit	
bit varying	
bytea	
interval	text
tsquery	
tsvector	
txid_snapshot	
data	date
time	text
timestamp	timestamp

timestampz	timestampz
------------	------------

ElasticSearch

UDTS 支持 ElasticSearch 作为数据传输源/目标,支持版本 6.5.4 到 7.10.2。

ElasticSearch 填写表单

数据源表单

参数名	说明
地址类型	提供内网地址,外网地址,专线地址三种方式,示例:http://10.9.1.1:9200
用户名	ElasticSearch 连接用户名
密码	ElasticSearch 用户名对应密码
索引	需要迁移的 ElasticSearch 索引名称,必填项,多个索引之间用分割,暂不支持*迁移所有索引,示例:testdata1,testdata2

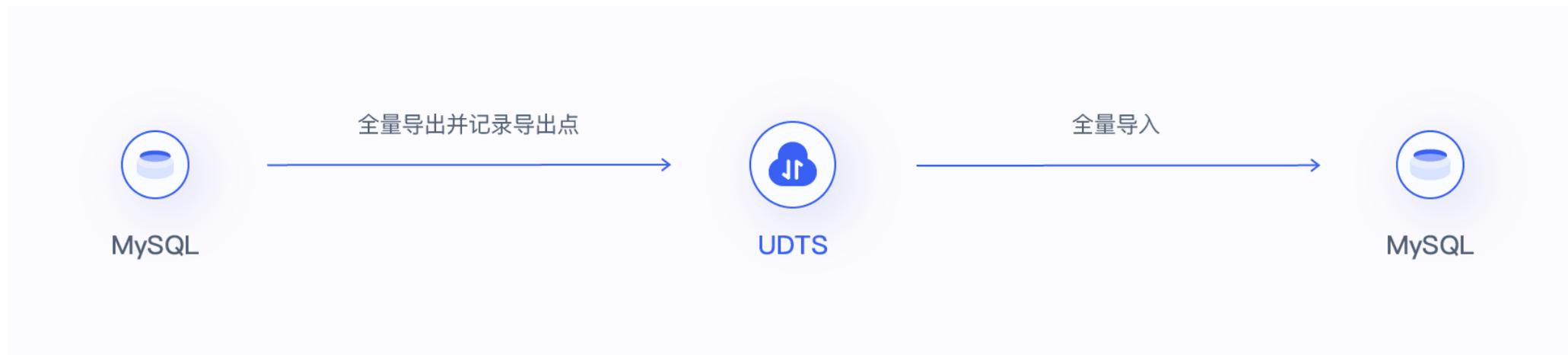
传输目标表单

参数名	说明
地址类型	目标暂时只支持内网,包括 UES 与云主机自建 ElasticSearch

用户名	ElasticSearch 连接用户名
密码	ElasticSearch 用户名对应密码

功能原理

全量迁移



UDS在全球很多个地域部署了服务,会根据源目位置来就近选择运行位置来保证迁移效率。

单向增量同步



UDTs 模拟一个Slave节点,通过Binlog从全量导出点开始获取增量数据。

双向增量同步



与单向增量同步类似，模拟Slave来获取增量数据。同时UDTS对写下去的数据做标记，当有新的Binlog Event的时候，会先检查是否有标记。如果有标记则说明是循环数据，直接丢弃，如果没有标记则加上标记写到对端。

技术优势

- 以服务的形式运行，用户无需考虑物理资源
- 支持多种类型的迁移
- MySQL到MySQL支持不同维度的迁移，可以按表，按库，多库，全库来迁移。
- MySQL到MySQL 支持 ETL
- MySQL到MySQL 支持单向增量同步，也支持双向增量同步。既支持通过GTID同步，也支持通过binlog file/pos同步。支持断点续传。
- 从AWS 迁移MySQL 可以从主库直接进行迁移，不需要搭从库
- 支持通过内网、外网、及专线迁移。
- 提供迁移限速功能。

传输速度

不同场景下，单个UDTS 任务的运行效率。UDTS 可以并发运行，同时提供灵活的迁移方法，既可以按库迁移，又可以按表迁移。迁移速率一般受限于迁移带宽，或者源、目的数据库的性能。通过多任务并发，一般场景都可以100% 利用迁移带宽。

源	目的	类型	带宽(Mbps)	数据量 (GB)	速率 (Mb/s)
AWS RDS for MySQL南美	UCloud 巴西 MySQL	全量导出	200	13	200
AWS RDS for MySQL南美	UCloud 巴西 MySQL	全量导入	200	13	40
新加坡 MySQL	越南MySQL	全量导出	800	10	800
新加坡MySQL	越南 MySQL	全量导入	800	10	80

UCloud 华北一 UDB 8GB HA	UCloud 华北一 UDB 8GB HA	全量导出	200	19	200
UCloud 华北一 UDB 8GB HA	UCloud 华北一 UDB 8GB HA	全量导入	200	19	64
UCloud 华北一 UDB 8GB HA	UCloud 华北一 UDB 8GB HA	全量导出	内网	19	2104
UCloud 华北一 UDB 8GB HA	UCloud 华北一 UDB 8GB HA	全量导入	内网	19	578

单向增量同步峰值支持2W OPS (insert, update, delete)

创建UDTS任务

UDTS为全球服务，不区分地域及可用区。

点击【创建任务】

数据传输服务

数据传输

数据集成

创建任务

停止

启动

删除

🔍

🔄

👉

<input type="checkbox"/>	任务名	任务ID	任务类型▼	创建时间↑	任务状态▼	进度	操作
--------------------------	-----	------	-------	-------	-------	----	----

数据传输(UCloud Data Transmission Service) UDTS支持多种异构数据源之间进行 全量/增量 数据传输。UDTS可以轻松帮助用户调整数据架构，实时数据同步进行后续数据分析。[创建任务](#)

< 1 > 10 条/页 /1

基本设置

选择任务类型

UDTS支持全量迁移任务,单向增量同步任务,全量+增量任务,双向增量同步4种

选择资源类型

UDTS 资源类型分为旗舰版、轻量版和基础版,详细区别请参考[这里](#)

填写任务名称

- 支持设置任务备注,最大长度不超过255
- 支持设置任务重试次数,任务失败后自动重试,最大重试次数为5次

基本设置

任务类型 ?

资源类型

	全量	增量
 旗舰版	传输速度 800 Mbps 磁盘空间 500 G	MySQL  20000 QPS MongoDB  50000 QPS Redis  90000 QPS PostgreSQL  40000 QPS
 轻量版	传输速度 400 Mbps 磁盘空间 500 G	MySQL  4000 QPS MongoDB  20000 QPS

磁盘空间 200 G

Redis 60000 QPS

PostgreSQL 20000 QPS

基础版

传输速度 80 Mbps

磁盘空间 100 G

增量

MySQL 2000 QPS

Redis 30000 QPS

MongoDB 10000 QPS

PostgreSQL 10000 QPS

最大重试次数

任务名称*

备注

数据源* 设置

未设置数据源

传输目标* 设置

未设置传输目标

填写源目详细信息

- 数据源

设置数据源

类型设置

数据源类型

MySQL

TIDB

CSV

MongoDB

US3

Redis

PostgreSQL

DynamoDB

ES

地址类型 [?](#)

内网地址 外网地址 专线地址

内网地址*

所在地域 拉斯各 圣保罗 华北一 广州 上海金融云 上海二 华北二 法兰克福 香港 雅加达 东京

首尔 莫斯科 新加坡 台北 伦敦 洛杉矶 华盛顿 胡志明市

VPC ID DefaultVPC [↻](#)

所属子网 Default Network(10.52.0.0/16) [↻](#)

管理设置

端口*

用户名*

密码*

数据库名*

表名

最大速率 [?](#) MB/s

NoLocks [?](#) OFF

SSL安全连接 OFF

- 数据目标

设置传输目标

类型设置

目标类型

MySQL

TiDB

地址类型 [?](#)

内网地址

内网地址 *

所在地域

拉各斯

圣保罗

华北一

广州

上海金融云

上海二

华北二

法兰克福

香港

雅加达

东京

首尔

莫斯科

新加坡

台北

伦敦

洛杉矶

华盛顿

胡志明市

VPC ID

DefaultVPC [?](#)

所属子网

Default Network(10.52.0.0/16) [?](#)

管理设置

端口 *

用户名 *

密码 *

最大速率 [?](#)

取值范围1-1024 MB/s

NoBinlog

 OFF

SSL安全连接

 OFF

预检查

- 填写任务信息后, 点击【立即购买】, 对任务进行预检查

付费信息

<input checked="" type="radio"/> 按时	0 元
<input type="radio"/> 月付	0 元
<input type="radio"/> 年付	0 元

合计费用 0 元

[立即购买](#)

预检查 详细信息。

完成创建

- 预检查通过后,点击【购买】->【立即支付】完成任务创建

检查结果 ✕

检查类别	结果	信息
数据源连通性检查	● 成功	
数据源权限检查	● 成功	
数据源配置检查	● 成功	
传输目标连通性检查	● 成功	
传输目标权限检查	● 成功	
传输目标配置检查	● 成功	

合计费用 **0.00** 元

您如果购买了 **预付费产品** 将开启 **自动续费** (以最终支付价格为准)

[如何开具发票?](#)

预检查

为了避免任务启动时,因为密码错误、缺少权限、无法连通等问题,导致任务失败。UDS提供预检查功能。

填写表单后点击【立即购买】

付费信息

<input checked="" type="radio"/> 按时	0 元
<input type="radio"/> 月付	0 元
<input type="text" value="1个月"/>	0
<input type="radio"/> 年付	0 元
<input type="text" value="1年"/>	0

合计费用 **0** 元

[立即购买](#)

检查结果 ✕

检查类别	结果	信息
数据源连通性检查	● 成功	
数据源权限检查	● 成功	
数据源配置检查	● 成功	
传输目标连通性检查	● 成功	
传输目标权限检查	● 成功	
传输目标配置检查	● 成功	

- 出现连通性检查失败,请检查端口密码是否正确。
- 出现权限错误,请对应FAQ,查询该任务类型所需权限。
- 出现配置检查失败,请参考预检查FAQ。

启动任务

创建完一个UDTS任务为“已创建”

启动UDTS迁移任务

数据传输服务

数据传输服务

创建任务

停止

启动

删除

<input type="checkbox"/>	任务名	任务ID	任务类型▼	数据量 (条) ↓	剩余时间/总时间 (min) ⌚	任务状态▼	操作
<input type="checkbox"/>	test-1	udts-ehyugq5	全量任务	0	0/0	● 已创建	详情 启动 ...

获取任务详情

- 找到对应的任务, 点击【操作】列中的【详情】按钮

数据传输服务

创建任务 停止 启动 删除

任务名	任务ID	任务类型	数据量 (条)	剩余时间/总时间 (min)	任务状态	操作
test-1	udts-ehyugqj5	全量任务	0	0/0	已创建	详情 启动 ...

- 进入【详情】页面

数据传输服务 / mao-gtid-test-09-03

停止 启动 删除

基本信息

任务名	mao-gtid-test-09-03
任务ID	udts-bn4ftsrc
任务类型	全量任务
数据量 (条)	0
任务状态	完成
Sync	BinlogName: mysql-bin.000003 BinlogPos: 3427 BinlogGTID: 2fd0e29-c957-11e9-bb72-0242ac160002:1-16

数据源信息

数据库名	mao1
IP	113.31.111.251
端口	33306
用户名	root

目标信息

所在地域	上海二
数据库名	mao1
IP	10.23.31.140
端口	3306
用户名	root
VPC ID	uvnet-btr1fd

查看进度

UDTS 任务分为全量任务、增量任务、全+增任务和双向同步任务,每种任务有不同的任务阶段:

- 全量任务:转储中 -> 加载中 -> 完成
- 增量任务:同步中 <-> 已同步
- 全+增任务:转储中 -> 加载中 -> 同步中 <-> 已同步
- 双向同步任务:同步中

各阶段具体情况如下:

- 转储中,表示将全量数据从源库转储至中转机,具体进度以百分比显示
- 加载中,表示将全量数据从中转机加载到目标库,具体进度以百分比显示
- 同步中,表示任务处于增量同步阶段,源库的新增数据会及时同步到目标库,增量同步没有完成状态,如需停止任务请在控制台操作。
- 已同步,表示当前源库新增数据已全部同步至目标库,由于源库随时可能新增数据,同步中 和 已同步的状态会随时转换。

支持数据类型

数据源	数据目标	全量迁移	增量同步
MySQL	MySQL	支持	支持
MySQL	TiDB	支持	支持
TiDB	MySQL	支持	不支持

TiDB	TiDB	支持	不支持
Redis	Redis	支持	支持
MongoDB	MongoDB	支持	支持

全量任务进度

以百分比的形式显示当前任务进度。

方式一：在列表中查看

数据传输

数据传输

数据集成

创建任务

停止

启动

删除



<input type="checkbox"/>	任务名	任务ID	任务类型▼	创建时间⌵	任务状态▼	进度	操作
<input type="checkbox"/>	fh-ws	udts-xf0m0td1	全量+增量任务	2020-11-26 21:18:20	● 已同步		详情 启动 ...
<input type="checkbox"/>	fh-test-json-3	udts-fbxdc4g	增量任务	2020-11-20 17:34:20	● 已同步		详情 启动 ...
<input type="checkbox"/>	mao-progress	udts-30j2sfqz	全量任务	2020-11-18 18:52:57	● 完成	100%	详情 启动 ...

方式二：在详情中查看

[<](#) 数据传输 / mao-progress

概览

运行历史

停止

启动

删除

修改

创建增量任务

基本信息

任务名 mao-progress

任务ID udts-30j2sfqz

任务类型 全量任务

任务状态 ● 完成

进度 100%

Sync BinlogName: mysql-bin.000048
BinlogPos: 15962
BinlogGTID: 38a9d892-b9b4-11ea-942a-52540053d389:1-76967394,40e2643c-b9b4-11ea-93fb-525400e06097:1-4144290

增量任务进度

对于 MySQL/TiDB/Redis 任务,显示剩余数据量大小与预计完成时间。

对于 MongoDB 任务,显示 Oplog 时间戳与同步延迟时间。

[在详情中查看](#)

[<](#) 数据传输 / test-mysql

概览

运行历史

停止

启动

删除

修改

创建增量任务

基本信息

任务名 test-mysql

任务ID udts-41f4fuzm

任务类型 全量+增量任务

任务状态 ● 已同步

剩余binlog大小 0 字节

预计完成时间 0 秒

Sync BinlogName: mysql-bin.000049
BinlogPos: 203892245
BinlogGTID: 40e2643c-b9b4-11ea-93fb-525400e06097:1-6532277,38a9d892-b9b4-11ea-942a-52540053d389:1-76967394

查看任务运行历史

通过运行历史，可以查看一个任务在各阶段所用时间等信息。

- 找到对应的任务，点击【操作】列中的【详情】按钮

数据传输服务

创建任务 停止 启动 删除

任务名	任务ID	任务类型	数据量 (条)	剩余时间/总时间 (min)	任务状态	操作
test-1	udts-ehyugqj5	全量任务	0	0/0	已创建	详情 启动 ...

- 进入【详情】页面

数据传输服务 / mao-gtid-test-09-03

停止 启动 删除

基本信息

任务名	mao-gtid-test-09-03
任务ID	udts-bn4ftsrc
任务类型	全量任务
数据量 (条)	0
任务状态	完成
Sync	BinlogName: mysql-bin.000003 BinlogPos: 3427 BinlogGTID: 2fd10e29-c957-11e9-bb72-0242ac160002:1-16

数据源信息

数据库名	mao1
IP	113.31.111.251
端口	33306
用户名	root

目标信息

所在地域	上海二
数据库名	mao1
IP	10.23.31.140
端口	3306
用户名	root
VPC ID	uvnet-btr1fd

- 切换到【运行历史】标签页

< 数据传输 / mao-progress

概览

运行历史

状态变更时间↑

状态

2020-11-18 18:52:57

● 已创建

2020-11-18 18:53:04

● 已启动

2020-11-18 18:53:04

● 检查中

2020-11-18 18:53:28

● 转储中

2020-11-18 18:55:30

● 加载中

2020-11-18 18:58:10

● 完成

停止任务

运行中的全量任务和增量任务都可以进行停止

停止UDTS迁移任务

<input type="checkbox"/>	任务名	任务ID	任务类型▼	数据量 (条) ⌵	剩余时间/总时间 (min) ⌵	任务状态▼	操作
<input type="checkbox"/>	test-1	udts-ehyugqj5	全量任务	0	0/0	● 已创建	详情 启动 ...
<input type="checkbox"/>	test-aaa	udts-nimz1ub5	全量任务	0	0/0	● 已创建	详情 启动 停止 删除

修改任务

当用户出现任务信息填错等需要修改的时候,UDTS提供了修改任务的功能

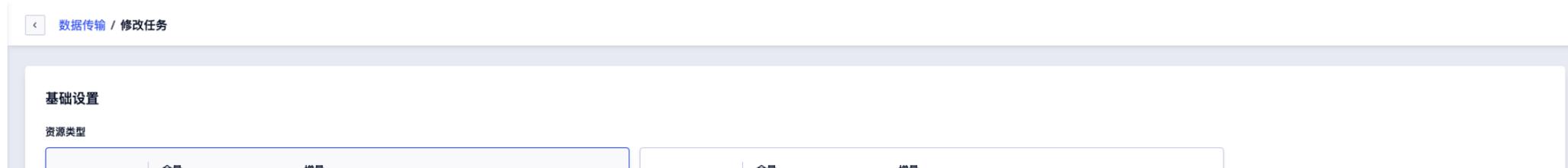
修改UDTS迁移任务

用户无法对同步中的任务进行修改参数。目前开放的修改状态和类型为, 全量任务/全量+增量任务, 状态为“已创建”, “停止”, “失败”, “成功”的任务。



任务状态 ▼	操作
● 已创建	详情 启动 ...
● 已创建	详情 启动 停止 删除
● 停止	详情 启动 修改
● 已创建	详情 启动 ...

进入任务表单对参数进行修改,可修改资源类型、任务名称、数据源/目标的信息等。



< 数据传输 / 修改任务

基础设置

资源类型

全量 增量

旗舰版

传输速度 800 Mbps
磁盘空间 1000 G

MySQL 20000 QPS
MongoDB 50000 QPS
Redis 90000 QPS
PostgreSQL 40000 QPS

基础版

传输速度 80 Mbps
磁盘空间 100 G

MySQL 2000 QPS
MongoDB 10000 QPS
Redis 30000 QPS
PostgreSQL 10000 QPS

轻量版

传输速度 400 Mbps
磁盘空间 200 G

MySQL 4000 QPS
MongoDB 20000 QPS
Redis 60000 QPS
PostgreSQL 20000 QPS

任务设置

任务类型

全量任务：将数据源库中指定库表/整库的全量数据一次性传输到目标端；增量任务：将数据源库中指定库表/整库的增量数据持续传输到目标端，任务持续运行。

[全量+增量任务](#)

数据源类型 *

[MySQL](#)

传输目标类型 *

[MySQL](#)

mysql
数据源

[编辑](#)

所在地域	华北一
内网地址	10.9.39.137
VPC ID	uvnet-u0ecvp
所属子网	subnet-2slodr
端口	3306
用户名	root
数据库名	luna 📄
NoLocks	否
连接方式	非加密连接

mysql
传输目标

[编辑](#)

所在地域	华北一
内网地址	10.9.121.6
VPC ID	uvnet-u0ecvp
所属子网	subnet-2slodr
端口	3306
用户名	root
数据库名	luna 📄
NoBinlog	否
连接方式	非加密连接

被修改的任务状态,会初始化为“已创建”,全+增任务修改后重新启动时,会从全量阶段开始。

备注:目前不提供修改 数据源目类型 和 传输任务类型。

更新任务名称和备注

在任务列表页面编辑UDTS迁移任务的名称和备注

数据传输服务

• 数据传输 • 数据集成

创建任务 停止 启动 删除

<input type="checkbox"/>	任务名 <small>🔍</small>	任务ID	资源类型 <small>▼</small>	任务类型 <small>▼</small>	创建时间 <small>🕒</small>	任务状态	进度	操作
<input type="checkbox"/>	udts-test 修改备注	udts-hv5v1ba1u11	基础版	全量+增量任务	2023-04-17 10:25:35	◦ 停止		详情 启动 ...

名称只能包含中英文、数字以及-_.[:];最大支持长度128字节;备注最大支持长度255字节

修改名称及备注 ✕

名称

备注

基于全量任务创建增量任务

UDTS 对以下表格中所列出的全量传输任务，提供基于已完成的全量任务快速创建增量任务的功能。

源类型	目标类型
MySQL	MySQL
MySQL	TiDB

步骤一 列出全量任务

在UDTS产品首页默认会列出所有任务，从中找到已完成的全量任务”

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	数据量 (条) ⬇	剩余时间/总时间 (min) ⓘ ⬇	创建时间 ⬇	任务状态 ▼	操作
<input type="checkbox"/>	vpc-test	udts-hw5mcj4a	全量任务	0	1/1	2019-12-23 14:39:46	● 完成	详情 启动 ...

[<](#) [1](#) [>](#) 10 条/页 [▾](#) /1

点击“详情”按钮进入下一步。

步骤二 查看任务详情

确认 sync 栏有 binlog 数据。如果 sync 栏没有数据说明源数据库没有开启binlog, 无法创建增量任务进行数据同步。

< 数据传输服务 / mao-gtid-test-09-03

停止 启动 删除 **创建增量任务**

基本信息	
任务名	mao-gtid-test-09-03
任务ID	udts-bn4ftsrc
任务类型	全量任务
数据量 (条)	0
任务状态	● 完成
Sync	BinlogName: mysql-bin.000003 BinlogPos: 3427 BinlogGTID: 2fd0e29-c957-11e9-bbf2-0242ac160002:1-16

数据源信息	
数据库名	mao1
IP	113.31.111.251
端口	33306
用户名	root

目标信息	
所在地域	上海二
数据库名	mao1
IP	10.23.31.140
端口	3306
用户名	root
VPC ID	uvnet-btr1fd

点击“创建增量任务”按钮进入下一步。

步骤三 创建增量任务

填写增量任务的任务名称, 并点击编辑进行补充源目数据库的密码

基础版

全量

传输速度 80 Mbps

磁盘空间 100 G

增量

MySQL 2000 QPS

Redis 30000 QPS

MongoDB 10000 QPS

PostgreSQL 10000 QPS

最大重试次数

任务名称

备注

1年 折合: 291.67元 / 月

合计费用 350 元

立即购买

数据源 编辑

mysql

外网地址	192.168.0.1
端口	3307
用户名	root
数据库名	testd 📄
SyncData	BinlogName: binlog.000023 BinlogPos: 164160150
连接方式	非加密连接

传输目标 编辑

mysql

所在地域	华北一
内网地址	10.9.5.138
VPC ID	uvnet-u0ecvp
所属子网	subnet-2slodr
端口	3308
用户名	root

3, 点击编辑, 补充目标库密码等信息

填写数据库密码, 以及任务名称来完成创建。

设置数据源

类型设置

数据源类型

MySQL

MongoDB

地址类型 

内网地址

外网地址

专线地址

外网地址 *

106.75.90.95

管理设置

端口 *

3307

用户名 *

root

密码 *

长度范围 1 ~ 32

源目的数据库密码需要重新填写

数据库名 *

test

表名

Sync Data

BinlogGTID

删除任务

非运行中的任务可以进行删除

删除UDTS迁移任务

<input type="checkbox"/>	任务名	任务ID	任务类型▼	数据量 (条) 1k	剩余时间/总时间 (min) 0k	任务状态▼	操作
<input type="checkbox"/>	test-1	udts-ehyugqj5	全量任务	0	0/0	● 已创建	详情 启动 ...
<input type="checkbox"/>	test-aaa	udts-nimz1ub5	全量任务	0	0/0	● 已创建	详情 启动 停止 删除

数据校验

UDTS为数据传输任务提供数据校验功能,用于校验源库和目标库的数据一致性。支持行数校验与完整数据校验两种模式。

支持数据类型

数据源	数据目标	行数校验	完整数据校验
MySQL	MySQL	支持	支持
MySQL	TiDB	支持	支持
TiDB	MySQL	支持	支持
TiDB	TiDB	支持	支持
MongoDB	MongoDB	支持	支持

前提条件

启动数据校验任务需满足以下条件:

1. 全量任务需运行到“完成”状态
2. 增量/全+增任务需运行到“同步中”/“已同步”状态,运行中的任务可以直接启动数据校验,也可以停止任务后再进行校验
3. 任务状态为失败时不支持数据校验

4. 如果当前任务有运行中的数据校验任务,不支持重复启动数据校验

校验内容

数据校验功能仅支持校验通过UDTS任务迁移的库表,不支持校验未迁移库表。

1. 行数校验:仅对比源库和目标库中每张表的行数,不比较具体数据内容。
2. 完整数据校验:对比源库和目标库中具体的数据内容。
 - MySQL数据校验时,校验源库和目标库中所有数据,只在源库中存在的数据和只在目标库中存在的数据都会被标记为不一致。
 - MongoDB数据校验时,以源库为准,只校验源库中的数据在目标库是否一致。

使用方法

1. 全量任务可以在创建任务时选择“全量完成时进行数据校验”,可选择两种迁移模式,全量任务完成后会自动启动数据校验任务。

更多设置

自动数据校验
开启后，当全量任务完成会自动进行数据校验。

ON OFF

任务名称*

最大重试次数
UDTS为运行失败的任务提供了自动重启任务的功能，重试次数最大限定5次。

2. 满足启动条件的任务均可在任务列表页或任务详情页点击按钮启动数据校验任务。

任务状态	进度	操作
● 增量-已同步		<input type="button" value="详情"/> <input type="button" value="启动"/> <input type="button" value="⋮"/>
● 增量-已同步		<input type="button" value="详情"/> <input type="button" value="启动"/> <input type="button" value="⋮"/>
● 增量-已同步		<input type="button" value="详情"/> <input type="button" value="启动"/> <input type="button" value="⋮"/>
● 增量-已同步		<input type="button" value="详情"/> <input type="button" value="启动"/> <input type="button" value="⋮"/>

暂停

修改资源类型

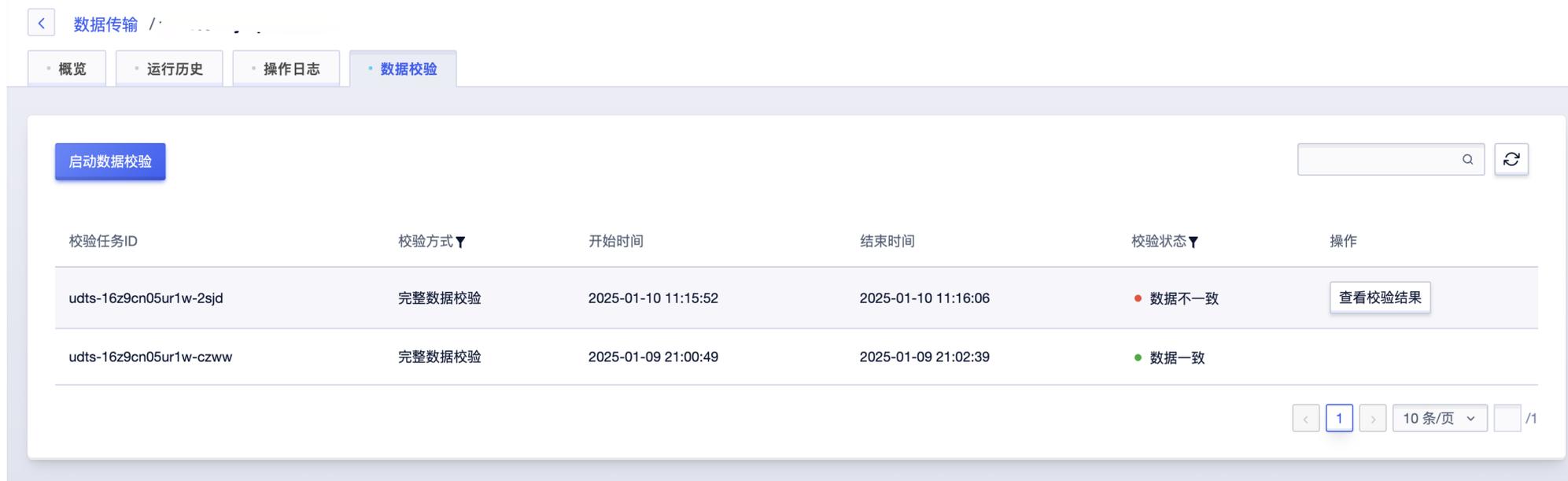
数据校验

删除

修改



3. 数据校验的详细信息可在 任务详情页 -> 数据校验页面 查看。



注意事项

1. MySQL 完整数据校验需要库表中存在主键或唯一索引。无唯一索引的表只校验行数，不校验具体数据内容。
2. 数据校验任务会占用源库和目标库的读资源，可能导致数据库负载上升。建议在业务低峰期执行数据校验任务。
3. 如果在源库有持续数据写入的情况下开启数据校验，最近写入数据的校验结果可能不准确，建议在源库停写的情况下执行数据校验任务。

4. 数据不一致时,前端页面最多展示前10条差异数据,详情请下载完整数据校验报告。报告文件保留3天,请及时下载。
5. 完整数据校验中,如果某张表的`不一致数据`超过10000行,则停止对这张表的校验。请在修复数据后重新校验。

支持数据类型

UDTS有奖用户调研,期待您的加入!

数据源	数据目标	全量迁移	增量同步
MySQL	MySQL	支持	支持
MySQL	TiDB	支持	支持
TiDB	TiDB	支持	不支持

MySQL

数据集成的基本要求与数据迁移完全相同，具体信息请参考数据迁移说明

迁移内容

数据集成当前只迁移 Database、Table 及其数据，不迁移视图(View)、函数(Function)、存储过程(Procedure)等其它内容。

功能说明

执行顺序

- 数据源按添加显示顺序排序，前面的任务优先执行。
- “增量任务”与“全量+增量任务”均为并行执行。
- 当某一数据源的任务类型为全量任务时，只有该全量任务完成后才启动后续任务。

是否保留原数据

数据源配置中提供【是否保留目标库的原数据】选择项，默认为“是”。

- 如果保留默认“是”选择，新数据只会不断的添加到目标数据库中。如果遇到冲突则按设定的冲突解决策略执行。
- 如果选择否，在任务开始之前会删除ETL设置中目标中对应的库表。

数据冲突解决策略

数据源配置中提供【数据冲突解决方式】选择项，用于选择数据发生冲突时所采取的处理方式。

- 如果选择替换,则会使用replace into对存在的数据进行替换
- 如果选择保留,则保留原有数据,忽略新数据。

数据过滤策略

数据源配置中提供【数据过滤策略】选择项，用于在增量同步中过滤某些类型的 DML 语句。

- 可以选择过滤 insert/update/delete 中的任意一项或多项。
- 默认不选择,增量同步所有类型的 DML 语句。

ETL

- 填写ETL信息时,如果勾选了数据库或者表,则必须填写映射名,如果不需要进行映射,默认填写与数据源相同的库名/表名
- 填写ETL信息时,如果勾选了库,未选择表,则迁移该库下的所有表

功能限制

- 数据集成任务暂不支持DDL语句,因为DDL语句可能会改变目标数据库的表结构,破坏ETL功能。

TiDB

数据集成的基本要求与数据迁移完全相同,具体信息请参考数据迁移说明

迁移内容

数据集成当前只迁移 Database、Table 及其数据,不迁移视图(View)等其它内容。

功能说明

执行顺序

- 数据源按添加显示顺序排序,前面的任务优先执行。
- 当某一数据源的任务类型为全量任务时,只有该全量任务完成后才启动后续任务。

是否保留原数据

数据源配置中提供【是否保留目标库的原数据】选择项,默认为“是”。

- 如果保留默认“是”选择,新数据只会不断的添加到目标数据库中。如果遇到冲突则按设定的冲突解决策略执行。
- 如果选择否,在任务开始之前会删除ETL设置中目标中对应的库表。

数据冲突解决策略

数据源配置中提供【数据冲突解决方式】选择项，用于选择数据发生冲突时所采取的处理方式。

- 如果选择替换，则会使用replace into对存在的数据进行替换
- 如果选择保留，则保留原有数据，忽略新数据。

ETL

- 填写ETL信息时，如果勾选了数据库或者表，则必须填写映射名，如果不需要进行映射，默认填写与数据源相同的库名/表名
- 填写ETL信息时，如果勾选了库，未选择表，则迁移该库下的所有表

功能限制

- 数据集成任务暂不支持DDL语句，因为DDL语句可能会改变目标数据库的表结构，破坏ETL功能。

创建任务

选择"数据集成"标签页,点击"创建任务"按钮



填写任务名称

- 支持设置任务备注,最大长度不超过255

基本设置

任务名称*

备注

数据源*

未设置数据源

+ 添加数据源

传输目标*

未设置传输目标

设置

添加数据源

点击添加数据源, 填写源库信息

添加数据源

类型设置

资源类型

	全量	增量	
🚩 旗舰版	传输速度 800 Mbps	MySQL 20000 QPS	MongoDB 50000 QPS
	磁盘空间 500 G	Redis 90000 QPS	PostgreSQL 40000 QPS

轻量版

传输速度 400 Mbps
磁盘空间 200 G

MySQL 4000 QPS
Redis 60000 QPS
MongoDB 20000 QPS
PostgreSQL 20000 QPS

基础版

全量
传输速度 80 Mbps
磁盘空间 100 G

增量
MySQL 2000 QPS
MongoDB 10000 QPS
Redis 30000 QPS
PostgreSQL 10000 QPS

数据源类型

MySQL

任务类型

全量任务 增量任务 全量+增量任务

地址类型

内网地址 外网地址 专线地址

内网地址*

所在地域

拉各斯	圣保罗	华北一	广州	上海金融云	上海二	华北二	法兰克福
香港	雅加达	东京	首尔	莫斯科	新加坡	台北	伦敦
洛杉矶	华盛顿	胡志明市					

VPC ID

所属子网

管理设置

端口*

用户名*

密码*

最大速率 MB/s

最大重试次数

NoLocks OFF

是否保留目标库的原数据 是 否

数据冲突解决方式 替换 保留

数据过滤策略 过滤insert 过滤update 过滤delete

进行预检查,成功通过后点击【确定】

可按需添加多个数据源

数据源*

 **mysql**

地址	端口	用户名	
10.9.47.226	3306	root	 ETL 

+ 添加数据源

点击【ETL】,选择需要迁移的库与表,将自动填写映射值,若有需要,可更改目标库的映射库名与表名。

ETL设置

映射设置

全选 →

- > mao1
- > fh3
- > fh4
- > fh2
- > a.b
- > fh
- > mao2
- ▼ integration2
 - t_csv1
 - t_csv
 - t_csv2
- > integration1

取消 确认

添加目标

填写目标库信息,完成预检查后,点击确认

设置传输目标

类型设置

目标类型

MySQL

TIDB

地址类型 [?](#)

内网地址

内网地址 *

所在地域

拉各斯

圣保罗

华北一

广州

上海金融云

上海二

华北二

法兰克福

香港

雅加达

东京

首尔

莫斯科

新加坡

台北

伦敦

洛杉矶

华盛顿

胡志明市

VPC ID

DefaultVPC



所属子网

Default Network(10.52.0.0/16)



管理设置

端口*

用户名*

密码*

预检查

取消

确定

预检查

为了避免任务启动时,因为密码错误、缺少权限、无法连通等问题,导致任务失败。UDTS提供预检查功能。

数据源预检查

添加数据源

类型设置

资源类型

 旗舰版	全量 传输速度 800 Mbps 磁盘空间 500 G	增量 MySQL  20000 QPS MongoDB  50000 QPS Redis  90000 QPS PostgreSQL  40000 QPS
 轻量版	全量 传输速度 400 Mbps 磁盘空间 200 G	增量 MySQL  4000 QPS MongoDB  20000 QPS Redis  60000 QPS PostgreSQL  20000 QPS
 基础版	全量 传输速度 80 Mbps 磁盘空间 100 G	增量 MySQL  2000 QPS MongoDB  10000 QPS Redis  30000 QPS PostgreSQL  10000 QPS

数据源类型

 MySQL

任务类型

 全量任务 增量任务 全量+增量任务

地址类型

 内网地址 外网地址 专线地址

内网地址*	<input type="text"/>																								
所在地域	<table border="1"><tr><td>拉各斯</td><td>圣保罗</td><td>华北一</td><td>广州</td><td>上海金融云</td><td>上海二</td><td>华北二</td><td>法兰克福</td></tr><tr><td>香港</td><td>雅加达</td><td>东京</td><td>首尔</td><td>莫斯科</td><td>新加坡</td><td>台北</td><td>伦敦</td></tr><tr><td>洛杉矶</td><td>华盛顿</td><td>胡志明市</td><td></td><td></td><td></td><td></td><td></td></tr></table>	拉各斯	圣保罗	华北一	广州	上海金融云	上海二	华北二	法兰克福	香港	雅加达	东京	首尔	莫斯科	新加坡	台北	伦敦	洛杉矶	华盛顿	胡志明市					
拉各斯	圣保罗	华北一	广州	上海金融云	上海二	华北二	法兰克福																		
香港	雅加达	东京	首尔	莫斯科	新加坡	台北	伦敦																		
洛杉矶	华盛顿	胡志明市																							
VPC ID	DefaultVPC <input type="button" value="刷新"/>																								
所属子网	Default Network(10.52.0.0/16) <input type="button" value="刷新"/>																								

管理设置

端口*	<input type="text"/>
用户名*	<input type="text"/>
密码*	<input type="password"/> <input type="button" value="预检查"/>
最大速率 <input type="button" value="帮助"/>	取值范围1-1024 MB/s <input type="text"/>
最大重试次数 <input type="button" value="帮助"/>	0 <input type="text"/>
NoLocks <input type="button" value="帮助"/>	<input type="checkbox"/> OFF
是否保留目标库的原数据	<input checked="" type="radio"/> 是 <input type="radio"/> 否
数据冲突解决方式	<input type="radio"/> 替换 <input checked="" type="radio"/> 保留
数据过滤策略	<input type="checkbox"/> 过滤insert <input type="checkbox"/> 过滤update <input type="checkbox"/> 过滤delete

取消

确定

检查结果

设置数据源

设置

任务组

地址

端口

用户名

密码

测试连通性

最大速率 ①

最大重试次数 ①

NoLocks ①

是否保留目标库的原数据

检查结果

检查类别	结果	信息
数据源连通性检查	● 成功	
数据源权限检查	● 成功	
数据源配置检查	● 成功	

取消 确定

不填写默认为峰值 MB/s

0

OFF

是 否

目标预检查

设置传输目标

类型设置

目标类型

MySQL

TiDB

地址类型 ?

内网地址

内网地址*

所在地域

拉各斯

圣保罗

华北一

广州

上海金融云

上海二

华北二

法兰克福

香港

雅加达

东京

首尔

莫斯科

新加坡

台北

伦敦

洛杉矶

华盛顿

胡志明市

VPC ID

DefaultVPC



所属子网

Default Network(10.52.0.0/16)



管理设置

端口*

用户名*

密码*

预检查

取消

确定

检查结果

检查类别	结果	信息
传输目标连通性检查	● 成功	
传输目标权限检查	● 成功	
传输目标配置检查	● 成功	

取消 确定

预检查失败

连通性检查失败

- 请检查端口密码是否正确。
- 如果有白名单, 请检查白名单设置。
- 如果有防火墙, 请检查防火墙设置。

权限检查失败

- 出现权限错误, 请对应FAQ, 查询该任务类型所需权限。

配置检查失败

出现配置检查失败, 请参考预检查FAQ。

启动任务

创建数据集成任务后, 该任务状态为【已创建】

启动UDTS数据集成任务

🔔 数据传输

数据传输 数据集成

创建任务 停止 启动 删除

搜索 🔍 刷新 🔄 重置 ⌂

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	创建时间 ↑	任务状态 ▼	操作
<input type="checkbox"/>	fn2020	udts-qb0nwe4j	数据集成	2020-11-18 17:44:07	● 已创建	详情 启动

任务详情

数据传输

数据传输 数据集成

创建任务 停止 启动 删除

任务名	任务ID	任务类型	创建时间	任务状态	操作
fh2020	udts-qb0nwe4j	数据集成	2020-11-18 17:44:07	已创建	详情 启动 ...

点击【详情】，查看任务基本信息、数据源信息以及目标信息

[<](#) 数据集成 / fh2020

停止 启动 删除

基本信息

任务名	fh2020
任务ID	udts-qb0nwe4j
任务类型	数据集成
任务状态	● 完成

目标信息

数据类型	mysql
所在地域	上海二
内网地址	10.25.213.150
VPC ID	uvnet-jcdfqp
所属子网	subnet-f4zkdb
端口	3306
用户名	root

数据源信息

任务类型	所在地域	地址	VPC ID	所属子网	端口	用户名	ETL设置	进度	任务状态
全量任务		106.75.240.103			63306	root	查看	100%	● 完成
全量任务	cn-sh2	10.25.164.159	uvnet-jcdfqp	subnet-f4zkdb	3306	root	查看	100%	● 完成

[<](#) 1 [>](#) 10条/页 /1

在数据源信息中, 点击【ETL设置】列中的【查看】, 可查看库表映射详情

ETL设置

映射设置

▼ integration1

integration1_a

任务类型	所在地域	地
全量任务		10
全量任务	cn-sh2	10

取消 确认

查看进度

以百分比的形式显示各数据源任务进度。

支持数据类型

数据源	数据目标	全量迁移
MySQL	MySQL	支持

在详情中查看

数据源信息

任务类型	所在地域	地址	VPC ID	所属子网	端口	用户名	ETL设置	进度	任务状态
全量任务	cn-sh2	10.25.164.159	uvnet-jcdfqp	subnet-f4zkdb	3306	root	查看 	100%	● 完成

停止任务

任务正在运行时,可点击【停止】按钮停止任务

停止UDTS数据集成任务

创建任务 停止 启动 删除

搜索 刷新 帮助

<input type="checkbox"/>	任务名	任务ID	任务类型▼	创建时间↑	任务状态▼	操作
<input type="checkbox"/>	fh2020	udts-l0wwdguh	数据集成	2020-11-26 15:54:41	● 已启动	详情 启动 ...
<input type="checkbox"/>	mao-si-test-stable-2	udts-i2h31gw4	数据集成	2020-11-18 17:03:58	● 完成	详情 启动 停止 删除

修改任务

对于数据集成类型,UDTS提供了修改任务备注的功能

编辑UDTS数据集成任务的备注

<input type="checkbox"/>	mao-t2t-0527 修改备注  <input type="button" value="修改"/>	udts-mxumtpf4	全量任务	2021-05-27 20:43:52	● 完成	100%	<input type="button" value="详情"/> <input type="button" value="启动"/> <input type="button" value="⋮"/>
--------------------------	--	---------------	------	---------------------	---	------	--

最大支持备注长度255字节

修改备注

备注

演示样例

取消 确定

删除

任

uc

udts-salbize4

数据集成

2021-04-07 13:57:27

停止

删除任务

非运行中的任务可以进行删除, 对于正在运行中的任务需要先停止。

删除数据集成任务

数据传输

数据传输

数据集成

[创建任务](#) [停止](#) [启动](#) [删除](#) [🔍](#) [🔄](#) [🔔](#)

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	创建时间 ⌵	任务状态 ▼	操作
<input type="checkbox"/>	fh2020	udts-qb0nwe4j	数据集成	2020-11-18 17:44:07	● 已创建	详情 启动 ⋮
<input type="checkbox"/>	mao-si-test-stable-2	udts-l2h31gw4	数据集成	2020-11-18 17:03:58	● 完成	详情 启动 停止 删除

监控指标

监控项

指标名	单位	适用数据库	说明
剩余数据量大小	Byte	MySQL/Redis	Mysql: 表示目标库还未与源库同步的binlog数据量大小 Redis: 表示源库与目标库之间复制偏移量的差距
预计剩余时间	s	MySQL	表示目标库与源库达成同步的预计剩余时间
剩余数据变化量	Byte	MySQL	连续两次剩余数据量大小之间的差值, 计算方式: 剩余数据变化量 = 上一次剩余数据量大小 - 当前剩余数据量大小。剩余数据变化量大于等于 0, 说明目标库写入速度大于源库新增数据速度, 任务状态正常; 剩余数据变化量小于 0, 说明目标库写入速度小于源库新增数据速度, 需要检查数据库状态
同步延迟时间	s	Mysql/Mongo/Redis	Mysql: 表示同步到目标库的最后一条binlog执行时间与当前时间的差值 Mongo: 表示同步到目标库的最后一条oplog执行时间与当前时间的差值 Redis: 目标库接收到最后一次复制更新的时间与当前时间的差值
源库拉取数据量	Bps	所有数据库	表示 UDS 每秒从源库拉取的数据量
源库拉取数据行数	个/s	Mysql 增量阶段	表示 UDS 每秒从源库拉取的数据行数

UDTS与源库的网络延迟	ms	所有数据库(内网或专线任务)	表示 UDTS 与源库之间的网络延迟时间
目标库写入数据量	Bps	所有数据库	表示 UDTS 每秒写入目标库的数据量
目标库写入数据行数	个/s	Mysql 增量阶段	表示 UDTS 每秒写入目标库的数据行数
UDTS与目标库的网络延迟	ms	所有数据库	表示 UDTS 与目标库之间的网络延迟时间

监控信息查看

在任务详情界面查看监控信息。





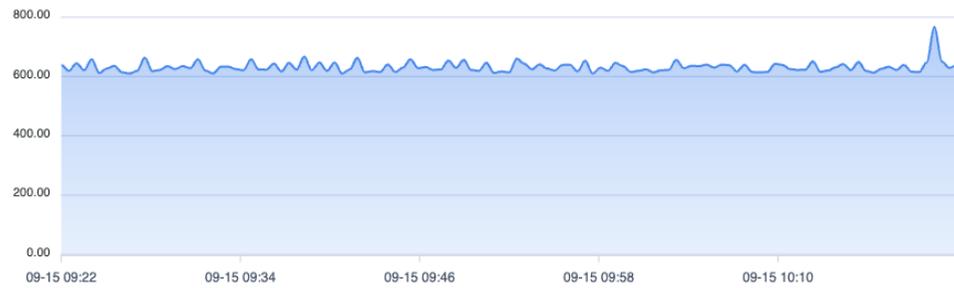
剩余数据变化量

● 剩余数据变化量



源库拉取数据量(Bps)

● 源库拉取数据量



同步延迟时间(s)

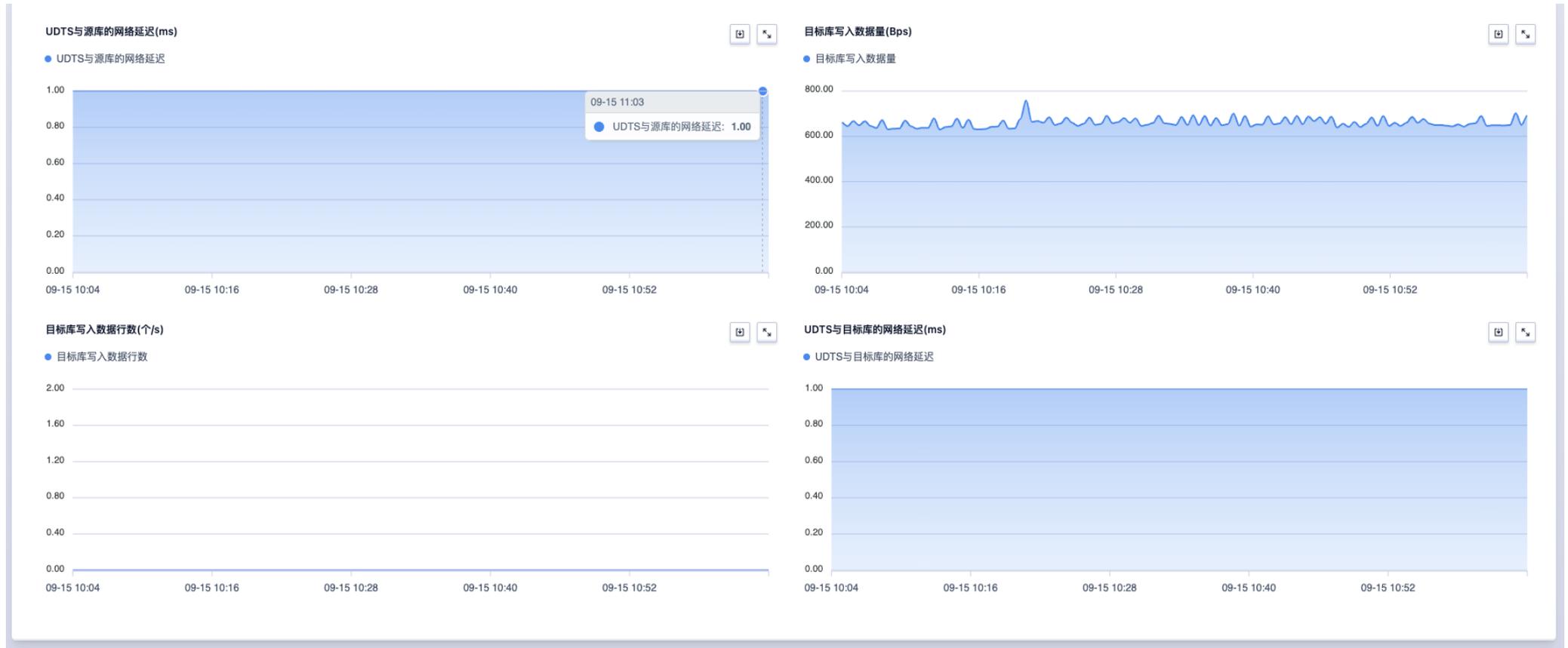
● 同步延迟时间



源库拉取数据行数(个/s)

● 源库拉取数据行数





适用范围

所有的数据传输任务,以及2023年6月1日以后创建的数据集成任务均支持监控。

2023年6月1日以前创建的数据集成任务不支持监控,如有需要,可以重新创建数据集成任务。

2023年7月20日以前创建的任务对部分监控指标不支持,如有需要请联系技术支持。

监控指标告警

告警配置

在任务详情界面点击告警模版, 选择默认告警模版或自定义模版。

停止
启动
删除
修改
创建增量任务

基本信息

任务名	test
任务ID	udts-tkpsxawn
资源类型	基础版
任务类型	全量+增量任务
任务状态	● 已创建
告警模板	默认数据传输告警模版 ✎

数据源信息



mysql

所在地域: 华北一

告警管理

通过模板您可以进行更方便的告警设置。我们提供了一些默认告警模板，您可以直接应用到您的主机或其他云资源。您也可以自定义模板，创建或编辑模板请移步至[告警模板列表](#)

默认数据传输告警模版 D

无告警

监控对象	告警阈值	通知对象
剩余数据变化量	小于: 0	默认组

取消
确认

默认模版监控对象为剩余数据变化量,告警方式为小于 0 时告警,为避免数据波动带来的误告警,设置触发周期为 3 次,即剩余数据变化量连续 3 次小于 0 时触发告警。

消息订阅

利用 消息订阅,UDTS可以在传输任务失败时,通知用户。

如何设置消息订阅

进入 资源监控UMon - 消息订阅: <https://console.ucloud.cn/umon/umsg>

打开任务失败通知

资源监控	消息订阅			
资源监控	保存			
告警模板	您可以进入 通知人管理 添加或修改联系人信息，操作成功后返回本页面进行关联通知组选择。 ×			
告警记录	消息类型	通知对象	消息订阅通知方式	是否开启
通知人管理	UEDN网络告警通知 [?]	默认组	<input checked="" type="checkbox"/> 邮件通知 <input checked="" type="checkbox"/> 短信通知	<input checked="" type="checkbox"/>
监控代理	USSSL证书过期通知 [?]	默认组	<input checked="" type="checkbox"/> 邮件通知 <input checked="" type="checkbox"/> 短信通知	<input checked="" type="checkbox"/>
消息订阅	UFileToken过期告警通知 [?]	默认组	<input checked="" type="checkbox"/> 邮件通知 <input checked="" type="checkbox"/> 短信通知	<input checked="" type="checkbox"/>
	物理网络变更通知 [?]	默认组	<input checked="" type="checkbox"/> 邮件通知 <input checked="" type="checkbox"/> 短信通知	<input checked="" type="checkbox"/>
	UCDN证书过期告警通知 [?]	默认组	<input checked="" type="checkbox"/> 邮件通知 <input checked="" type="checkbox"/> 短信通知	<input checked="" type="checkbox"/>
	UDTS 任务失败告警通知 [?]	默认组	<input checked="" type="checkbox"/> 邮件通知 <input checked="" type="checkbox"/> 短信通知	<input checked="" type="checkbox"/>

FAQ

1 问：现在支持多少区域间的跨域迁移？

UDTS 跨域迁移利用到了 UDPN或者专线, 已支持区域且有UDPN/专线传输线路的路径皆可支持。已支持区域详见控制台。

2 问：MySQL 全量迁移需要满足哪些条件

请参考

3 问：MySQL 增量迁移需要满足哪些条件

请参考

4 问：增量提示 **only support ROW format binlog** 该如何操作

增量迁移要求 MySQL(包括UDB MySQL) binlog_format 值为 ROW, 可以根据情况选择下面的方式之一来完成 binlog_format 的变更; 变更完成后, 需要重新执行一次 UDS 的 **全量 + 增量**

1. 修改配置文件（默认为 `my.cnf`），重启 MySQL

```
[mysqld]
...
binlog_format = ROW
binlog_row_image = FULL
...
```

备注：如果是 MySQL 5.5，没有 `binlog_row_image` 这个变量，不需要设置

2. 通过 MySQL 命令设置

需要特别注意的是，如果通过 MySQL 命令设置 `binlog_format`，当 MySQL 存在连接往数据库中写入数据时，写入的 `binlog_format` 还是原值，需要将连接断开后才会生效。

```
FLUSH TABLES WITH READ LOCK;
FLUSH LOGS;
SET GLOBAL binlog_format = 'ROW';
-- 同样的，如果是 MySQL 5.5，不需要设置 binlog_row_image
SET GLOBAL binlog_row_image = 'FULL';
FLUSH LOGS;
UNLOCK TABLES;
```

变更完成后，可以通过以下命令断开现有的连接

```
-- 查看当前所有连接
```

```
> show processlist;
+----+-----+-----+-----+-----+-----+-----+-----+
| Id | User | Host | db | Command | Time | State | Info |
+----+-----+-----+-----+-----+-----+-----+-----+
| 495 | root | 10.20.5.1:56820 | <null> | Query | 0 | starting | show processlist |
| 497 | root | 10.20.5.1:56828 | <null> | Sleep | 3 | | <null> |
+----+-----+-----+-----+-----+-----+-----+-----+

-- 通过 kill 断开所有 session,如果能确认哪些连接有写操作,可以只 kill 这些连接
> kill 497

FLUSH LOGS;
```

如果你使用的是 master-slave 模式,需要执行以下命令

在 slave 节点上执行

```
stop slave;
```

在 master 上执行

```
FLUSH TABLES WITH READ LOCK;
FLUSH LOGS;
SET GLOBAL binlog_format = 'ROW';
FLUSH LOGS;
UNLOCK TABLES;
```

在 slave 上执行

```
start slave;
```

备注:

通过 `SET GLOBAL binlog_format = 'ROW'`; 设置参数,

再次通过 `show global variables like 'binlog_format'`; 查询到的值还是原值, 需要断开连接后重新连接才会显示变更后的值。

3. 云数据

如果你使用的是云数据, 可以通过复制当前使用的配置文件, 修改对应的 `binlog_format` 和 `binlog_row_image` 配置项, 然后使用该配置重新应用数据库。

5 问: ERROR 1292 (22007): Incorrect date value: '0000-00-00' for column

这是由于迁移的 MySQL(UDB-MYSQL) 的 `sql_mode` 不一致导致的, 目标库的日期不允许 '0000-00-00' 这种无效的日期。

可以通过下面的命令查看 `sql_mode`

```
select @@sql_mode;
```

--- 或者

```
show variables like "sql_mode";
```

如果当前 `sql_mode` 中包括 `NO_ZERO_IN_DATE`、`NO_ZERO_DATE`, 将其删除, 并在查询出来的 `sql_mode` 值的后面, 增加 `ALLOW_INVALID_DATES` 来允许插入这种无效的日期

通过以下命令对 sql_mode 进行修改

```
SET GLOBAL sql_mode='xxxx,ALLOW_INVALID_DATES';
```

这里的 xxxx 是指原有查询出来的 sql_mode 值(需要删去 NO_ZERO_IN_DATE 和 NO_ZERO_DATE)

6 问： 如何判断MySQL-MySQL增量任务中，源目数据库已经同步完成

当增量任务中目标库的数据和源库达到一致时,增量任务的状态会由 同步中 变成 已同步。

已同步 状态说明:

指迁移任务中指定的 db 和 table 对应的目标数据库数据和源数据库达到一致。比如:当迁移任务为 * 时,已同步 是指目标数据库内置库(sys, test, INFORMATION_SCHEMA, PERFORMANCE_SCHEMA 等)之外的所有 db 达到和源库数据一致。当迁移任务为多 db 时,比如 db1,db2 ,是指目标数据的 db1,db2 和源数据库数据达到一致。当源数据库数据发生任何变化(即使发生变化的 db 不在迁移任务中),已同步 状态也会再次变化为 同步中,直到数据再次达到一致。

7 问： error="Error 1040: Too many connections" 错误处理

数据库服务器连接达到上限。

```
show variables like '%max_connections%';
+-----+-----+
| Variable_name | Value |
+-----+-----+
```

```
| max_connections | 151 |  
+-----+-----+  
1 row in set (0.01 sec)
```

解决方法:

1. 增加 max_connections: set GLOBAL max_connections=1000;

8 问: row size too large (\u003e 8126). Changing some columns to TEXT or BLOB or using ROW_FORMAT=DYNAMIC or ROW_FORMAT=COMPRESSED may help. In current row format, BLOB prefix of 768 bytes is stored inline

解决方法:

在目标库执行

```
set GLOBAL innodb_file_format = 'Barracuda';  
set GLOBAL innodb_strict_mode = OFF;
```

9 问: Column count doesn't match value count

阿里云迁移到 UCloud 的增量任务出现类似的错误, Column count doesn't match value count: 2 (columns) vs 3 (values), 是因为源库中存在隐藏主键, 导致数据的 columns

和 values 不匹配。

解决方法:查找没有主键的表,并增加主键

- 查找没有主键的表

```
SELECT
table_schema,
table_name
FROM
information_schema.tables
WHERE (table_schema, table_name)
NOT in( SELECT DISTINCT
table_schema, table_name FROM information_schema.columns
WHERE
COLUMN_KEY = 'PRI')
AND table_schema NOT in('sys', 'mysql', 'INFORMATION_SCHEMA', 'PERFORMANCE_SCHEMA');
```

- 对这些没有主键的表增加主键

10 问: sync: Type:ExecSQL msg:"exec jobs failed,err:Error 1205:Lock wait timeout exceeded;try restarting transaction"

当目标数据库配置较低时可能会遇到这个问题,用户重新启动任务即可。任务会自动从上次同步完成的位置继续执行(即支持断点续传功能)

11 问：MySQL 全量任务运行中会有数据库锁吗？什么时候释放

默认模式下：全量任务运行中，为了保证数据的一致性，会执行 FTWRL (FLUSH TABLES WITH READ LOCK)，如果迁移的数据库中存在 MyISAM 等非事务表时，锁会在这些非事务表数据备份完成之后释放（锁释放的时间与表数据大小有关）；对于 InnoDB 引擎表，会立即释放锁，并通过 START TRANSACTION WITH CONSISTENT SNAPSHOT 开启读一致的事务。

- 对于 MyISAM表，正在转储中的表允许读，不允许写，直到表转储完毕
- 对于 InnoDB表，正在转储中的表允许读写。

Nolock 模式下：不会执行FTWRL (FLUSH TABLES WITH READ LOCK)，对于 MyISAM表，正在转储中的表允许读，不允许写，并且写SQL会阻塞之后的读SQL直到表转储完毕。

12 问：Redis 迁移出现 ERR illegal address

原因可能是用户开启了白名单设置，但是UDTS机器的 IP 不在白名单内。如需要白名单地址，请联系技术支持。

13 问：Error 3140: Invalid JSON text

在MySQL 同步过程中出现 Error 3140: Invalid JSON text: "The document is empty." at position 0 in value for column，原因是源库校验不严格。数据库中的字段要求为 NOT NULL，但是数据中存在值为 NULL 的数据。

有两个解决方法，根据需要处理：

- 对源库中的数据进行修复，将所有值 NULL 的数据修正为正确的值（这也符合业务逻辑需要）
- 或者对目标库中的表进行修改，将字段修改为允许为 NULL。例如表为 period_progress，字段为 total

```
ALTER TABLE `period_progress` CHANGE `total` `total` JSON NULL;
```

14 问: Mongodb 迁移出现 error reading collection: cursor id 5707195885304103447 not found

mongodb 的cursor 默认有效时间是10分钟, 如果数据无法在10分钟内处理完成,cursor将会过期, 导致无法获取剩下的数据 解决方法

- 登录mongodb 的shell 切换至admin库 将cursor超时时间设置成1天

```
db.runCommand( { setParameter:1 , "cursorTimeoutMillis":86400000} )
```

15 问: Error 1264: Out of range value for column 'xxx' at row 100

在 MySQL2TiDB 的数据导入过程中,可能会出现 Error 1264: Out of range value for column 'xxx' at row 100, 这里的原因是 数据在源数据库中没有被强制校验,即非正常的数
据也可以正常存放,但导入到目标时目标数据库要求数据必须满足指定的约束。

比如:

```
CREATE TABLE `aaa` (  
  `id` bigint(20) NOT NULL AUTO_INCREMENT,  
  `lot` float(10,7) NOT NULL DEFAULT '0.0000000',  
  PRIMARY KEY (`id`)  
) ENGINE=InnoDB DEFAULT CHARSET=utf8 COLLATE=utf8_bin;
```

假如源数据库中存在一张结构如上的表,表中要求字段 lot 的数据类型为 float, 长度为 10, 小数点部分为 7 位, 如果数据中存在一条数据为 -1000 的数据, 则导入时可能会失败, 而出现
Error 1264: Out of range value for column 'lot' 的情况。

解决方案有两个, 需要对源数据库进行调整:

- 调整表结构,加大字段长度

```
select max(`lot`) from `aaa`;  
select min(`lot`) from `aaa`;
```

获得最大值和最小值,然后通过 alter 修改表结构,使其符合具体的条件。比如将上面的表 aaa 的 lot 修改为 float(11,7)

```
ALTER TABLE `aaa` MODIFY `lot` float(11,7) NOT NULL DEFAULT '0.0000000',
```

- 修正数据,使其符合表结构的限制

```
select `id` where `lot` > 999 OR `lot` < -999;
```

对查找出来的数据进行修正。

16 问: 目标库为TiDB时, 出现 **Error 1071: Specified key was too long; max key length is 3072 bytes**

需要增大目标库TiDB文件中的配置项max-index-length

17 问: 目标库为TiDB时, 出现 **Error: incorrect utf8 value f09f8c80 for column xxx**

需要设置目标库tidb_skip_utf8_check

```
set global tidb_skip_utf8_check=1;
```

18 问：目标库为TiDB时，出现 **Error 1071: Specified key was too long; max key length is 3072 bytes**

可以通过将TiDB配置文件中 max-index-length 项的值改成 12288 来解决

19 问： **Cannot add foreign key constraint**

在 MySQL 迁移到 MySQL/TiDB 时,如果出现 Cannot add foreign key constraint,需要检查表结构的 CHARSET,当前表结构及其依赖的外键的表结构 CHARSET 要一致。

以下为示例,当 group 和 user 表的 CHARSET 不同时,可能会出现错误: Cannot add foreign key constraint

```
CREATE TABLE `group` (  
  `id` int(11) NOT NULL AUTO_INCREMENT,  
  `name` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NOT NULL,  
  PRIMARY KEY (`id`)  
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4;
```

```
CREATE TABLE `user` (  
  `id` int(11) NOT NULL AUTO_INCREMENT,  
  `name` varchar(255) CHARACTER SET utf8mb4 COLLATE utf8mb4_bin NOT NULL,  
  PRIMARY KEY (`id`)  
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4;
```

```
`id` int(11) NOT NULL AUTO_INCREMENT,  
`username` varchar(255) NOT NULL,  
`password` varchar(255) DEFAULT NULL,  
`g_id` int(11) NOT NULL,  
PRIMARY KEY (`id`),  
FOREIGN KEY (`g_id`) REFERENCES `group` (`id`)  
) ENGINE=InnoDB DEFAULT CHARSET=utf8;
```

出现该问题时如何处理：

1. 在源数据库中,执行以下 SQL,修改表的 CHARSET,推荐使用兼容的编码替换老的编码(使用 utf8mb4 替换 utf8),使 user 和 group 的 CHARSET 一致。

```
ALTER TABLE `user` CONVERT TO CHARACTER SET utf8mb4 COLLATE utf8mb4_bin;
```

2. 重新运行 UDS 任务

20 问：源库是MySQL，增量同步时出现 **ERROR 1236 (HY000): The slave is connecting using CHANGE MASTER TO MASTER_AUTO_POSITION = 1, but the master has purged binary logs containing GTIDs that the slave requires**

这是因为增量同步需要的binlog同步点,无法在源库中找到。可能的原因有以下四种：

- 1 用户是全+增任务,全量的时间超过了binlog的最大过期时间,导致全量结束时,记录的binlog同步点被过期清除了。这种情况需要调大源库binlog的过期时间。
- 2 用户源库的binlog被手动清除。

- 3 用户的任务中途暂停或者失败过,然后间隔一段时间后再启动的任务,在此期间binlog过期被清除了。
- 4 用户增量同步没有使用GTID进行同步,并且发生主从切换。

解决方法:以上情况都需要重新执行一次 UDS 的全量 + 增量,直接重启任务无法解决该问题。步骤:

- 对于全量 + 增量任务,在任务详情页点击编辑任务,填入相关任务信息,然后选择保存任务。当前任务的进度会被清除,启动任务后会重新从全量开始。
- 对于直接创建增量任务的,需要重新做一次全量迁移,然后重新填写正确的BinlogName、BinlogPos、GTID,启动任务之后会从指定的 binlog 位置开始同步。

关于如何修改binlog日志过期时间有以下两种方式:

- 1 临时生效,重启数据库后失效

查看默认设置的过期时间

```
show variables like "%expire_logs%";
```

设置保留 15 天

```
set global expire_logs_days=15
```

刷新日志

```
flush logs;
```

查看新生成的 binlog 日志

```
show master status\G;
```

注意:以上命令在数据库执行会立即生效,请确定设置数据的保留日期,以免误删

- 2 修改配置文件,永久生效

MySQL 配置文件的位置请根据实际情况进行修改

```
vim /etc/my.cnf
```

修改有关binlog过期相关的配置

```
[mysqld]
expire_logs_days=15
```

注意:在配置文件修改后,需要重启才能永久生效。设置为 0 表示永不过期,单位是天。

21 问: The table 'xxx' is full 或者 No space left on device

一般是目标磁盘空间不足导致的,需要升级目标磁盘空间,或者清理目标数据库中的数据。

The table 'xxx' is full 还有一种特殊的情况是 table 使用了 memory 存储引擎

可以通过修改 MySQL 的配置文件的下面两个值来解决(或者将 table 修改为其它存储引擎),该值默认为 16M

```
tmp_table_size = 256M
max_heap_table_size = 256M
```

这个值设置成多少合适呢?这个需要根据实际情况处理

- 当前 MySQL 实例拥有的内存大小是多少,则该值为上限
- memory 引擎表期望存储的数据量大小是多少

22 问: Error 1785: When @@GLOBAL.ENFORCE_GTID_CONSISTENCY = 1, updates to non-transactional tables can only be done in either autocommitted statements or single-statement transactions, and never in the same statement as updates to transactional tables

在 MySQL 的同一个事务中,如果既操作了支持事务的表(比如 innodb),又操作了不支持事务的表(比如 myisam),在数据库开启 gtid 的情况下,会出现该错误。

解决方法,下面选择其一即可

- 在源数据库中,将非事务引擎的表(一般情况下是 myisam)修改为事务引擎的表(一般情况下是 innodb)
- 在目标数据库中,关闭 gtid

23 问: 迁移报错 Invalid default value for 'xxxx_time'

在迁移的过程中,如果目标数据库的 sql_mode 存在 NO_ZERO_DATE 或者 NO_ZERO_IN_DATE,并且源数据库中存在 0000-00-00 这样的日期值和 00:00:00 这样的时间值,如果存在这样的时间值,在加载数据的时候会出现错误。

解决方法:

- 修改目标数据库中的 sql_mode,去掉 NO_ZERO_DATE 和 NO_ZERO_IN_DATE,增加 ALLOW_INVALID_DATES 操作步骤可以参考 FAQ 第 5 条
- 如果目标库 sql_mode 中有 STRICT_TRANS_TABLES 避免产生其他问题也可以先去掉,等迁移完成后再添加回来

24 问: error creating indexes for xx.xx: createIndex error: WiredTigerIndex::insert: key too large to index, failing 1089

MongoDB 在2.6 版本引入 failIndexKeyTooLong 参数,在4.2版本之后废弃该参数,用于限制 index 的长度

解决方法:在目标数据库执行

```
db.runCommand({
  "setParameter": 1,
  "failIndexKeyTooLong": false}
);
```

25 问: OOM command not allowed when used memory > 'maxmemory' or rss_memory not enough.

目标数据库是MongoDB分片集群时,目标分片的容量可能出现内存不够的情况

解决方法:升级目标集群的分片容量

26 问： ddtLog.v_HBUSer: error creating collection ddtLog.v_HBUSer: error running create command: Cannot create view when the featureCompatibilityVersion is 3.X.

MongoDB 迁移时,以上报错说明源库与目标库的 featureCompatibilityVersion 取值不一致,且源库中存在与目标库不兼容的特性,例如视图、排序规则等

解决方法:检查并调整源库与目标库不兼容的内容

27 问： 源库是Redis，报错信息：[PANIC] read error, please check source redis log or network

Redis 迁移时,以上报错说明源库写入数据量大于缓冲区大小

解决方法:执行以下命令修改源库 client-output-buffer-limit 参数取值

```
config set client-output-buffer-limit 'slave 536870912 0 0'
```

28 问： 源库是MongoDB，报错信息：Failed: error creating intents to dump: error creating intents for database config: error getting collections for database config: (Unauthorized) not authorized on config to execute command { listCollections: 1, filter: {}, cursor: {}, lsid: { id: UUID("12d62805-acea-472f-9862-ca27253c107e") }, \$db: "config" }

MongoDB 全量迁移时, 以上报错说明源库用户缺少备份权限。

解决方法: 执行以下命令修改源库用户权限

```
db.grantRolesToUser("root",[{"role":"backup",db:"admin"}])
```

29 问: 源库是 MySQL, 报错信息: **Deadlock found when trying to get lock; try restarting transaction**

MySQL 全量迁移, load 阶段出现以上报错, 说明全量数据加载到目标库时, 多线程并发运行导致死锁。

解决方法: 将目标库限速值修改为 1

30 问: **load: return result; END; This function has none of DETERMINISTIC, NO SQL, or READS SQL DATA in its declaration and binary logging is enabled (you *might* want to use the less safe log_bin_trust_function_creators variable**

目标库执行以下命令后重试:

```
set global log_bin_trust_function_creators="ON";
```

预检查错误信息及解决方法

1 MySQL

1.1

错误信息:

log_bin is xxx, and should be ON

binlog_format is xxx, and should be ROW

binlog_row_image is xxx, and should be FULL

log_bin: xxx, 需要修改为 ON

binlog_format: xxx, 需要修改为 ROW

binlog_row_image: xxx, 需要修改为 FULL

解决方法:

如果涉及增量同步, 包括 增量任务、全量+增量任务、双向同步以及全量任务之后再增量任务的场景, 要求源数据库开启 binlog 功能, 且

- binlog_format 为 ROW
- binlog_row_image 为 FULL

备注: 如果只进行全量迁移, 可以忽略这个问题

1.1.1 源库未开启过 binlog

修改配置文件(默认为 my.cnf),重启 MySQL

```
[mysqld]
...
log-bin = /data/mysql/logs
binlog_format = ROW
binlog_row_image = FULL
...
```

如果您使用的是云厂商的数据库服务,需要修改对应的配置文件,使用修改后的配置重新启动数据库。

备注: MySQL 5.5 没有 binlog_row_image 这个变量,不需要设置

1.1.2 源库开启过 binlog, 但是 binlog_format 或 binlog_row_image 值不对

需要特别注意的是,如果通过 MySQL 命令设置 binlog_format,当 MySQL 存在连接往数据库中写入数据时,写入的 binlog_format 还是原值,需要将连接断开后才会生效。

```
FLUSH TABLES WITH READ LOCK;
FLUSH LOGS;
SET GLOBAL binlog_format = 'ROW';
-- 同样的,MySQL 5.5 不需要设置 binlog_row_image
SET GLOBAL binlog_row_image = 'FULL';
FLUSH LOGS;
```

```
UNLOCK TABLES;
```

变更完成后,可以通过以下命令断开现有的连接

```
-- 查看当前所有连接
> show processlist;
+----+-----+-----+-----+-----+-----+-----+-----+
| Id | User | Host | db | Command | Time | State | Info |
+----+-----+-----+-----+-----+-----+-----+-----+
| 495 | root | 10.20.5.1:56820 | <null> | Query | 0 | starting | show processlist |
| 497 | root | 10.20.5.1:56828 | <null> | Sleep | 3 | | <null> |
+----+-----+-----+-----+-----+-----+-----+-----+

-- 通过 kill 断开所有 session,如果能确认哪些连接有写操作,可以只 kill 这些连接
> kill 497

-- kill 掉原来的 session 之后,flush logs 保证新的 binlog 格式为 ROW
> FLUSH LOGS;
```

如果你使用的是 master-slave 模式,需要执行以下命令

在 slave 节点上执行

```
stop slave;
```

在 master 上执行

```
FLUSH TABLES WITH READ LOCK;  
FLUSH LOGS;  
SET GLOBAL binlog_format = 'ROW';  
FLUSH LOGS;  
UNLOCK TABLES;
```

在 slave 上执行

```
start slave;
```

备注:

通过 SET GLOBAL binlog_format = 'ROW'; 设置参数,

再次通过 show global variables like 'binlog_format'; 查询到的值还是原值, 需要断开连接后重新连接才会显示变更后的值。

1.2

错误信息:

The tables that need to be migrated from the source database include MyISAM engine tables, while the target database has GTID enabled, which may result in migration task failure.

源库需要迁移的表中包括 MyISAM 引擎表, 同时目标库开启了 GTID , 可能导致迁移失败

解决方法:

如果源库需要迁移的表中包括 MyISAM 引擎表, 同时目标库开启了 GTID , 可能导致 MySQL 1785 错误, 报错信息如下:

```
When @@GLOBAL.ENFORCE_GTID_CONSISTENCY = 1, updates to non-transactional tables can only be done in either autocommitted statements or single-
```

```
statement transactions, and never in the same statement as updates to transactional tables
```

建议用户将 MyISAM 引擎表转换为 InnoDB 引擎表,或者关闭目标库的 GTID 模式。

查询方式:

```
# 在源数据库(db1)中查询是否存在 MyISAM 表
select table_schema, table_name
from information_schema.tables
where engine = 'MyISAM'
and table_type = 'BASE TABLE'
and table_schema in (db1);

# 在目标库中查询是否开启了 GTID
show global variables like 'gtid_mode';
```

设置方式:

```
# 方案一:修改源库
# 将 MyISAM 表 table1 的引擎修改为 InnoDB
alter table table1 ENGINE = InnoDB;

# 方案二:修改目标库
# 关闭目标库的 GTID 模式
set global gtid_mode = "ON_PERMISSIVE";
set global gtid_mode = "OFF_PERMISSIVE";
```

```
set global gtid_mode = "OFF";
```

1.3

错误信息:

max_allowed_packet of the source is xxx, which is larger than max_allowed_packet of the target yyy

源库 max_allowed_packet 值 xxx 大于目标库 max_allowed_packet 值 yyy

解决方法:

源库的 max_allowed_packet 值大于目标库的 max_allowed_packet 时,可能导致目标库写入数据失败,建议调整目标数据库的 max_allowed_packet 值,使源目保持一致。

在源库中执行

```
> show global variables like "max_allowed_packet";
+-----+-----+
| Variable_name | Value |
+-----+-----+
| max_allowed_packet | 4194304 |
+-----+-----+
```

然后使用获取的 max_allowed_packet 值,在目标数据库中执行,比如

```
set global max_allowed_packet = 4194304;
```

1.4

错误信息:

```
table xxx have no primary key or at least a unique key
```

表:xxx 需要主键或唯一键

解决方法: 如果迁移涉及到增量同步,包括 增量任务、全量+增量任务、双向同步以及全量任务之后再增量任务的场景,需要为每张表设置主键或唯一键,否则在增量同步阶段可能出现重复数据。如果只进行全量迁移,可以忽略这个问题。

```
alter table xxx add primary key(xxxx);
```

1.5

错误信息:

```
sql_mode may cause error. Please check sql_mode NO_ZERO_DATE/NO_ZERO_IN_DATE in target db
```

sql_mode 可能导致报错,检查目标库的 sql_mode NO_ZERO_DATE/NO_ZERO_IN_DATE。

解决方法: 如果源库与目标库的 sql_mode 不一致,可能会导致部分数据无法迁移到目标库。

```
# 在源库中查询 sql_mode
show variables like "sql_mode";

# 在目标库中修改 sql_mode,使其与源库保持一致
SET GLOBAL sql_mode='xxxx';
```

1.6

错误信息:

log_slave_updates should be ON

log_slave_updates 需要修改为 ON

解决方法: 如果您使用从库作为迁移的源,那么要求从库开启 log_slave_updates,否则从库中没有 binlog 日志,导致无法迁移

修改配置文件(默认为 my.cnf),重启 MySQL

```
[mysqld]
...
log_slave_updates = 1
...
```

1.7

错误信息:

please stop event: 'db1':event1

增量同步前,需要关闭 event: 'db1':event1

解决方法:

增量同步开启之前,需要关闭 event

```
# 停止所有 event
```

```
SET GLOBAL event_scheduler = OFF;
```

```
# 如果按库迁移,停止指定库的 event 即可
```

```
# 查找对应的 event 并停止
```

```
USE db1;
```

```
SHOW EVENTS;
```

```
ALTER EVENT event1 DISABLE;
```

1.8

错误信息:

The variable innodb_XXX has different values in source and target. Please modify the variables to ensure consistency.

源库与目标库的参数 XXX 取值不同,建议修改参数保持一致

解决方法:

源库与目标库的 innodb 相关参数不一致,可能导致数据迁移报错,建议修改目标库参数,与源库保持一致

```
# 在目标库中修改不一致的参数取值
```

```
set GLOBAL innodb_file_format = 'Barracuda';
```

```
set GLOBAL innodb_strict_mode = OFF;
```

1.9

错误信息:

The variable `lower_case_table_names` has different values in source and target. Please modify the variables to ensure consistency.

源库与目标库的参数 `lower_case_table_names` 取值不同, 建议修改参数保持一致

解决方法:

源库与目标库的 `lower_case_table_names` 参数不一致, 可能导致数据迁移报错, 建议修改目标库参数, 与源库保持一致

修改配置文件(默认为 `my.cnf`), 重启 MySQL

```
[mysqld]
...
lower_case_table_names = 0
...
```

如果您使用的是云厂商的数据库服务, 需要修改对应的配置文件, 使用修改后的配置重新启动数据库。

2 TiDB

2.1

错误信息:

`tikv_gc_life_time` is xxx, and should be great than 1h

`tikv_gc_life_time`: xxx, 需要不小于 1h

解决方法:

如果迁移的任务类型为 全量任务、全量+增量任务, 且迁移的数据量较大时, 需要 tikv_gc_life_time 的值大于 转存 时间。一般情况下 2T 的数据, 大约需要 45min, 我们建议将该值设置为 1h 以上

在 TiDB 中执行以下语句

```
update mysql.tidb set VARIABLE_VALUE="1h" where VARIABLE_NAME="tikv_gc_life_time";
```

2.2

错误信息:

TiDB dose not support charset in table xxx. Please change charset to any one of 'ascii/latin1/binary/utf8/utf8mb4'.

TiDB 不支持此表采用的字符集 xxx , 请修改字符集为 'ascii/latin1/binary/utf8/utf8mb4' 中任意一种

解决方法:

TiDB目前支持的字符集包括ascii/latin1/binary/utf8/utf8mb4。

从MySQL迁移到TiDB时, 如果源库中需要迁移的表或表中某一字段采用的字符集不包含在上述字符集之中, 则无法迁移。

查询方式:

```
show create table table1;
```

设置方式:

```
# 将表 table1 的字符集修改为 utf8  
alter table task character set utf8;
```

```
# 将表 table1 中 column1 字段的字符集修改为 utf8  
alter table table1 change column1 column1 varchar(200) character set utf8;
```

3 MongoDB

3.1

错误信息:

Source and target version do not match, source version is 3.0, and target version is 5.0

源库与目标库版本不匹配,源库: 3.6.23 ,目标库: 5.0.14

解决方法:

Mongo 暂不支持跨大版本迁移,从3.x迁移到5.x时,需要创建4.x版本的中转库,先从3.x迁移到4.x,再从4.x迁移到5.x。

4 Redis

4.1

错误信息:

Source and target version do not match, source version is 4.0, and target version is 7.0

源库与目标库版本不匹配, 源库: 4.0 , 目标库: 7.0

解决方法:

Redis 跨大版本迁移可能存在兼容性问题, 建议通过中转库迁移, 例如, 从3.x/4.x迁移到7.x时, 需要创建5.x/6.x版本的中转库, 先从3.x/4.x迁移到5.x/6.x, 再从5.x/6.x迁移到7.x。

4.2

错误信息:

The source database version is 7.0, and does not support rump mode

源库版本 7.0 不支持 rump 模式

解决方法:

Redis 7.0 版本不支持 rump 模式, 建议源库开启 psync 权限再进行迁移。

如何从阿里云RDS MySQL迁移到UCloud UDB

适用场景:通过UDTS将阿里云 RDS MySQL数据库迁移到UCloud UDB MySQL。

配置外网地址和白名单

白名单地址请联系技术支持获取。

确保UDTS可以从外部访问阿里云的数据库,如图:

基本信息		设置白名单	^
实例ID: rm-bp187f0877	名称: rm-bp187f0877		
地域可用区: 华东1 (杭州) 可用区B	类型及系列: 常规实例 (基础版)		
内网地址:	内网端口: 3306		
外网地址:	外网端口: 3306		
存储类型: SSD云盘			
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。			

准备好UDB目标端资源

[<](#) MySQL / 创建数据库

地域

地域

北京二

上海二

广州

香港

洛杉矶

华盛顿

法兰克福

曼谷

首尔

新加坡

高雄

莫斯科

东京

台北

迪拜

雅加达

孟买

圣保罗

伦敦

拉各斯

胡志明市

可用区

可用区B

可用区C

可用区D

可用区E

基础配置

数据库类型 [?](#)

高可用版

容灾方案 [?](#)

北京二可用区D



创建UDTS任务

基本设置

任务类型 ?

全量任务

增量任务

全量+增量任务

双向同步

数据源类型

MySQL

TiDB

CSV

MongoDB

UFile

目标类型

MySQL

TiDB

最大重试次数 ?

0

设置数据源

地址类型 ?

内网地址

外网地址

专线地址

外网地址*

填写外网地址

如果是搭建的专线也可以使用专线地址。

由于阿里云RDS提供的高级账号并非root账号,需要在UDTS任务中开启 NoLocks。

启动UDTS任务

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	数据量 (条) ↓	剩余时间/总时间 (min) ⌚ ↓	创建时间 ↓	任务状态 ▼	操作
<input type="checkbox"/>	alirds-udb	udts-zarwqatu	全量任务	0	0/0	2019-12-17 15:13:57	● 已创建	详情 启动 ...

执行后可以观测任务状态

基本信息

任务名 alirds-udb

任务ID udts-zarwqatu

任务类型 全量任务

数据量 (条) 0

任务状态 ● 完成

Sync BinlogName: mysql-bin.000004
BinlogPos: 3578
BinlogGTID: 364e64f9-2094-11ea-bc3d-00163e0a3a43:1-11

如何从阿里云 云数据库Redis 迁移到 UCloud 云内存Redis

适用场景:通过UDTS将阿里云 云数据库Redis迁移到云内存Redis上。

注:阿里云 Redis 2.8如果要增量同步需要升级到4.0版本才可执行。

配置外网地址和白名单

白名单地址请联系技术支持获取

确保UDTS可以从外部访问阿里云的数据库,如图:

<  **test** (运行中) 登录数据库 ? 迁移Redis数据库 ? 刷新 修改密码 ? 清除数据 变更配置 ?

实例信息 | 实例信息

性能监控

报警设置

白名单设置

SSL设置

参数设置

账号管理

备份与恢复

连接管理

▶ 日志管理

▶ CloudDBA

基本信息

[迁移可用区 ?](#) [大版本升级 ?](#) [小版本升级](#) ▼

实例ID: r-bp1mgo74 复制	实例名称: test ✎	状态: 运行中
可用区: 杭州 可用区H	网络类型: VPC网络 vpc-bp18mxqtfov 复制	可维护时间段: 02:00-06:00 设置
实例规格: 4G集群版 (2节点)	交换机信息: vsw-bp1viojm24m 复制	最大私网带宽: 192 MB/s 临时调整带宽 ?
版本: Redis 4.0 社区版	Lua状态:	最大连接数: 20,000
实例释放保护: 关闭 设置		

连接信息

[Redis连接问题排查文档](#) [修改连接地址](#) [设置免密访问 ?](#)

私网连接地址 (Host): 复制	端口号 (Port): 6379
公网连接地址 (Host): 复制 释放公网连接	端口号 (Port): 6379
连接密码: ?	

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。

设置账号增量同步权限

Reids 2.8 版本或仅使用全量迁移请忽略此步骤

此步骤适用于阿里云 Redis 4.0+ 标准版。

进入侧边栏“账号管理”

 **test** (运行中)

实例信息

实例信息

基本信息

实例ID: r-bp1mgo7-  复制	实
可用区: 杭州 可用区H	网
实例规格: 4G集群版 (2节点)	交
版本: Redis 4.0 社区版	Lu
实例释放保护: 关闭 设置	

- 实例信息
- 性能监控
- 报警设置
- 白名单设置
- SSL设置
- 参数设置
- 账号管理**
- 备份与恢复

连接管理

- ▶ 日志管理
- ▶ CloudDBA

连接信息

私网连接地址 (Host) : r-bp1mgo74[redacted]dis.rds.aliyuncs.com [复制](#)

公网连接地址 (Host) : r-bp1mgo74wgx[redacted]dis.rds.aliyuncs.com [复制](#)

创建一个有“复制”权限的账号。

创建账号 ×

*账号:

以字母开头，由小写字母、数字、下划线 () 组成。长度不超过16个字符。

*权限设置: 只读 读写 复制

*密码:

密码:

大写字母、小写字母、数字、特殊字符占三种，长度为8~32位；特殊字符为!@#\$%^&*()_+ -=

*确认密码:

备注说明:

创建UDTS任务

根据需要,选择 全量 或 全量+增量 任务

基本设置

任务类型 ?

全量任务	增量任务	全量+增量任务	双向同步
------	------	---------	------

数据源类型

MySQL	TiDB	CSV	MongoDB	UFile	Redis	PostgreSQL
-------	------	-----	---------	-------	-------	------------

目标类型

Redis

最大重试次数 ?

0	▲	▼
---	---	---

注:增量时 密码 需要填写具有“复制”权限的 账号:密码 来作为密码。

例如: 用户名 ucloud 密码 password UDS任务中 数据源 密码请填写 ucloud:password

对于阿里云 Redis 4.0+ 集群版/读写分离版, 需要开启Rump选项迁移。

自建IDC如何传输数据到UCloud

当用户需要将自建IDC数据迁移到UCloud,可以通过数据库的外网地址或者搭建专线进行传输。

搭建专线

如用专线地址,用户需要提前搭建自建IDC与UCloud之间的专线。

创建UDTS任务

地址类型选择专线地址/外网迁移请选择外网地址。

设置数据源

地址类型 [?](#)

内网地址 外网地址 专线地址

专线地址*

端口*

用户名*

密码*

数据库名*

表名

最大速率 [?](#)

Sync Data ?"/>

ETL设置 [?](#) OFF

目标端填写打通的UCloud端内网地址和对应VPC

数据源信息

专线地址 192.168.120.98

端口 3306

用户名 root

数据库名 *

目标信息

所在地域 北京二

IP 172.16.120.45

VPC ID uvnet-05vasnpr

所属子网 subnet-l1oabcd2

端口 3306

用户名 root

数据库名 *

图中示例源端地址为自建IDC内网地址,和目标端地址所处UCloud内网,原本无法通信。通过专线实现内网传输。

如果对于专线的带宽有要求防止被UDTS的传输任务大量使用。可以对任务的传输进行限速,目前仅支持MySQL-MySQL的全量阶段。

运行UDTS任务

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	数据量 (条) 1k	剩余时间/总时间 (min) ① 1k	创建时间 1k	任务状态 ▼	操作
<input type="checkbox"/>	vpc-test	udts-hw5mcj4a	全量任务	0	1/1	2019-12-23 14:39:46	● 完成	详情 启动 ...

托管云如何传输数据到公有云区域

当用户需要将UCloud托管云数据迁移到公有云区域,可以直接通过UDTS专线地址的传输模式进行传输。

创建UDTS任务

地址类型选择专线地址

设置数据源

地址类型 ? 内网地址 外网地址 专线地址

专线地址*

端口*

用户名*

密码*

数据库名* 多库使用","隔开, 也可使用""全库传输

表名

最大速率 ? 不填写默认为峰值 MB/s

Sync Data ServerID* ?

ETL设置 ? OFF

目标端填写公有云区域内网地址和对应VPC

数据源信息

专线地址 192.168.120.98

端口 3306

用户名 root

数据库名 *

目标信息

所在地域 北京二

IP 172.16.120.45

VPC ID uvnet-05vasnpr

所属子网 subnet-l1oabcd2

端口 3306

用户名 root

数据库名 *

图中示例源端地址为托管云数据库内网地址,通过专线实现内网传输。

运行UDTS任务

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	数据量 (条) 1k	剩余时间/总时间 (min) ① 1k	创建时间 1k	任务状态 ▼	操作
<input type="checkbox"/>	vpc-test	udts-hw5mcj4a	全量任务	0	1/1	2019-12-23 14:39:46	● 完成	详情 启动 ...

跨VPC/跨项目/跨账号数据迁移

当用户因为一些特殊需求,已经将不同的vpc,或者不同的项目,甚至不同的账号进行内网打通后,需要用到数据迁移的情况,也可以使用UDTS来完成。

创建UDTS任务

由于UDTS当前版本目标端仅支持内网的地址,会对当前账号项目下的资源进行校验,所以创建任务需要创建在目标库所在的账号,项目下。

即:如果用户要将A项目下的数据库,迁移至B项目的数据库。需要在B项目下创建任务。

地址类型选择专线地址(需要提前保证两边内网可通信),仅填写源端内网地址即可。

设置数据源

地址类型 [?](#) 内网地址 外网地址 专线地址

专线地址*

端口*

用户名*

密码*

数据库名*

表名

最大速率 [?](#)

Sync Data ?"/>

ETL设置 [?](#) OFF

目标端填写公有云区域内网地址和对应VPC

数据源信息

专线地址 192.168.120.98

端口 3306

用户名 root

数据库名 *

目标信息

所在地域 北京二

IP 172.16.120.45

VPC ID uvnet-05vasnpr

所属子网 subnet-l1oabcd2

端口 3306

用户名 root

数据库名 *

运行UDTS任务

<input type="checkbox"/>	任务名	任务ID	任务类型 ▼	数据量 (条) ⌵	剩余时间/总时间 (min) ⌵	创建时间 ⌵	任务状态 ▼	操作
<input type="checkbox"/>	vpc-test	udts-hw5mcj4a	全量任务	0	1/1	2019-12-23 14:39:46	● 完成	详情 启动 ...

