

阿里云 数据传输

产品简介

文档版本：20190801

法律声明

阿里云提醒您 在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的”现状“、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含”阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 禁止： 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告： 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明： 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 确定 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
<code>##</code>	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
<code>[]</code> 或者 <code>[a b]</code>	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
<code>{ }</code> 或者 <code>{a b}</code>	表示必选项，至多选择一个。	<code>swich {stand slave}</code>

目录

法律声明.....	I
通用约定.....	I
1 产品优势.....	1
2 功能特性.....	3
2.1 数据同步.....	3
3 规格说明.....	8
3.1 数据迁移链路规格说明.....	8
3.2 数据同步规格说明.....	9

1 产品优势

数据传输服务DTS（Data Transmission Service）支持RDBMS、NoSQL、OLAP等多种数据源之间的数据传输。它提供了数据迁移、实时数据订阅及数据实时同步等多种数据传输方式。相对于第三方数据流工具，DTS提供丰富多样、高性能、高安全可靠的传输链路，同时它提供了诸多便利功能，极大方便了传输链路的创建及管理。

丰富多样

DTS支持多种同构或异构数据源之间的迁移同步，例如Oracle->MySQL、Oracle->Postgres Plus Advanced Server（PPAS）。对于异构数据源之间的迁移，数据传输服务支持结构对象定义的转化，例如将Oracle中的同义词转换为PPAS中对应的同义词定义。

DTS支持多种传输方式，包括数据迁移、实时数据订阅及数据实时同步。其中实时数据订阅及数据实时同步均为实时数据传输方式。

数据实时同步支持两个数据源之间的单向/双向同步，适用于数据异地灾备、异地多活(单元化)、应用就近访问、查询报表分流、实时数据仓库等应用场景。

为降低数据迁移对应用的影响，数据迁移功能支持不停服迁移，可实现在数据迁移过程中，应用停机时间降低到分钟级别。

高性能

DTS使用高规格服务器来保证每条迁移或同步链路都能拥有良好的传输性能。

对于数据迁移，DTS底层采用了多种性能优化措施，全量数据迁移高峰期时性能可以达到70MB/s，20万的TPS。

相对于传统的数据同步工具，DTS的实时同步功能能够将并发粒度缩小到事务级别，能够并发同步同张表的更新数据，从而极大地提升同步性能。高峰期时，数据同步性能可以达到30,000RPS。



说明:

性能指标仅供参考，DTS的数据迁移和数据同步性能受到源端、目标端性能、网络延时、网络带宽、链路规格等多方面因素影响。

安全可靠

DTS底层为服务集群，如果集群内任何一个节点宕机或发生故障，控制中心都能够将这个节点上的所有任务秒级切换到其他节点上，链路稳定性高达99.95%。

DTS内部对部分传输链路提供7×24小时的数据准确性校验，快速发现并纠正传输数据，保障传输数据可靠性。

DTS各模块间采用安全传输协议及安全token认证，并具有自动断点续传机制，有效地保证数据传输的可靠性。

简单易用

DTS提供可视化管理界面，提供向导式的链路创建流程，用户可以在其控制台简单轻松地创建传输链路。

[数据传输 DTS 控制台](#) 展示了链路的传输状态及进度，传输性能等信息，用户可以方便管理自己的传输链路。

DTS提供链路断点续传功能，解决网络或系统异常等导致的链路中断问题，且定期监测所有链路的状态，一旦发现链路异常，先尝试自动修复重启。如果链路需要用户介入修复，用户可以直接在控制台修复并重启链路。

2 功能特性

2.1 数据同步

数据同步功能旨在帮助用户实现两个数据源之间的数据实时同步。数据同步功能可应用于异地多活、数据异地灾备、本地数据灾备、数据异地多活、跨境数据同步、查询与报表分流、云BI及实时数据仓库等多种业务场景。

支持的同步拓扑

数据同步源实例	数据同步目标实例	支持的同步拓扑
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	<ul style="list-style-type: none"> · 单向同步 · 双向同步
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	MaxCompute实例（原ODPS）	单向同步
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	AnalyticDB实例	单向同步
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	Datahub实例	单向同步
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	Elasticsearch实例	单向同步

数据同步源实例	数据同步目标实例	支持的同步拓扑
DRDS实例	DRDS实例	单向同步
DRDS实例	AnalyticDB实例	单向同步
DRDS实例	Datahub实例	单向同步
<ul style="list-style-type: none"> · ECS上的自建Redis · 通过专线/VPN网关/智能网关接入阿里云的自建Redis 	<ul style="list-style-type: none"> · Redis实例 · ECS上的自建Redis · 通过专线/VPN网关/智能网关接入阿里云的自建Redis 	单向同步
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL 	<ul style="list-style-type: none"> · ECS上自建的Kafka · 通过专线/VPN网关/智能网关接入阿里云的自建Kafka 	单向同步
<ul style="list-style-type: none"> · RDS For MySQL实例 · ECS上的自建MySQL · 通过专线/VPN网关/智能网关接入阿里云的自建MySQL · RDS for PostgreSQL 	AnalyticDB for PostgreSQL实例	单向同步

支持的数据源版本

源数据库	目标数据库	源库版本	目标库版本
MySQL	MySQL	5.1、5.5、5.6、5.7、8.0	5.1、5.5、5.6、5.7、8.0
MySQL	Elasticsearch	5.1、5.5、5.6、5.7、8.0	所有版本的Elasticsearch
MySQL	ADS	5.1、5.5、5.6、5.7、8.0	所有版本的ADS
MySQL	MaxCompute	5.1、5.5、5.6、5.7、8.0	所有版本的MaxCompute
MySQL	Datahub	5.1、5.5、5.6、5.7、8.0	所有版本的数据hub
DRDS	DRDS	所有版本的DRDS	所有版本的DRDS
DRDS	ADS	所有版本的DRDS	所有版本的ADS
DRDS	Datahub	所有版本的DRDS	所有版本的数据hub

源数据库	目标数据库	源库版本	目标库版本
Redis	Redis	2.8、3.0、3.2、4.0  说明： 当源库为阿里云Redis实例时，数据库版本需为4.0版本。	2.8、3.0、3.2、4.0
MySQL	Kafka	5.1、5.5、5.6、5.7、8.0	0.10、1.0
RDS for PostgreSQL	AnalyticDB for PostgreSQL	9.4、10	不涉及

数据同步对象

同步对象的选择粒度可以为：库、表、列，您可以根据需要选择同步的对象。

数据同步支持库、表、列名映射，您可以进行不同库名之间的数据，或两个不同表名之间的数据同步。

数据同步支持列选择，您可以根据业务需求，只同步表中的某几列数据。

同步作业状态说明

同步作业是数据实时同步的基本单元。如果要进行两个实例间的数据同步，必须在数据传输控制台创建同步作业。

同步作业在创建及运行过程中，不同阶段会处于不同的状态，具体如下表所示。

作业状态	状态说明	可进行操作
预检中	同步作业正在进行启动前的预检查。	<ul style="list-style-type: none"> 查看同步配置 删除同步 复制同步配置 配置监控报警
预检查失败	同步作业预检查没有通过。	<ul style="list-style-type: none"> 预检查 查看同步配置 修改同步对象 修改同步速度 删除同步 复制同步配置 配置监控报警

作业状态	状态说明	可进行操作
未启动	迁移任务已经通过迁移之前的预检查，但是尚未启动。	<ul style="list-style-type: none"> · 预检查 · 开始同步 · 修改同步对象 · 修改同步速度 · 删除同步 · 复制同步配置 · 配置监控报警
同步初始化中	同步作业正在进行同步初始化。	<ul style="list-style-type: none"> · 查看同步配置 · 删除同步 · 复制同步配置 · 配置监控报警
同步初始化失败	同步作业在初始化过程中，同步失败。	<ul style="list-style-type: none"> · 查看同步配置 · 修改同步对象 · 修改同步速度 · 删除同步 · 复制同步配置 · 配置监控报警
同步中	同步作业正常同步中。	<ul style="list-style-type: none"> · 查看同步配置 · 修改同步对象 · 修改同步速度 · 删除同步 · 复制同步配置 · 配置监控报警
同步失败	同步作业同步异常。	<ul style="list-style-type: none"> · 查看同步配置 · 修改同步对象 · 修改同步速度 · 删除同步 · 复制同步配置 · 配置监控报警
暂停中	同步作业执行了暂停，处于暂停状态。	<ul style="list-style-type: none"> · 查看同步配置 · 修改同步对象 · 修改同步速度 · 删除同步 · 复制同步配置 · 配置监控报警

当任务处于同步失败或同步初始化失败的状态时，7天之内这个同步实例会被后台启动恢复。如要停止同步，您需要手动暂停或结束处于同步失败或同步初始化失败的任务。

高级特性

数据同步支持以下高级特性，简化产品的使用难度。

- 动态增减同步对象。

在数据同步过程中，用户可以随时增加或减少需要同步的对象。

- 完善性能查询体系。

数据同步提供同步延迟、同步性能(RPS、流量)趋势图，用户可以方便查看同步链路的性能趋势。

- 完善监控体系。

数据同步提供同步作业状态、同步延迟的报警监控功能。用户可以根据业务敏感度，自定义同步延迟报警阈值。

- 支持多种ETL特性。

- 支持库表列三级对象名映射，可以实现对源跟目标实例的库名、表名或列名不同的两个对象之间进行数据同步。
- 支持迁移数据过滤，可以对待同步的表设置某种SQL条件过滤要所需同步的数据。

3 规格说明

3.1 数据迁移链路规格说明

注意事项

本文中提供的性能指标仅用于提供测试数据，并不作为产品SLA的评判标准。

名词解释

- 规格: 数据传输为用户提供的不同性能的链路规格，以增量迁移性能为衡量标准。
- 表数量: 该测试模型下表的总数。
- 记录大小: 增量迁移时每条记录的大小。
- RPS: 表示在增量数据迁移的源端，每秒由于数据增删改（包括INSERT、UPDATE、DELETE等）产生的数据记录变化的数量。



说明:

- 如果一条SQL语句中包含对多行数据的操作，则计为多条数据记录变化；对同一条记录反复增删改则计为多条数据记录变化。
- 每次commit操作也会被计为一次数据记录变化的次数。

测试模型

测试流程：在两个RDS for MySQL实例之间创建增量迁移任务，在源RDS for MySQL上进行压测，检测增量迁移同步到目标RDS for MySQL实例的性能表现。

表 3-1: 测试环境配置

实例	RDS 实例配置	参考极限性能
源实例	<ul style="list-style-type: none"> · 实例规格: rds.mys2.8xlarge · 内存: 48000MB · 最大连接数: 2000 	<ul style="list-style-type: none"> · 最大QPS: 18000 · 最大IOPS: 14000
目标实例	<ul style="list-style-type: none"> · 实例规格: rds.mys2.8xlarge · 内存: 48000MB · 最大连接数: 2000 	<ul style="list-style-type: none"> · 最大QPS: 18000 · 最大IOPS: 14000

测试模型:

- 测试表数量为20。
- 测试表均有主键。
- 记录大小为1KB。
- 每个事务平均包含两条DML操作，一条commit。其中INSERT、UPDATE、DELETE的比例为3:1:2。

测试结果

源实例地区	目标实例地区	实例间网络延迟(单位:毫秒)	规格	TPS	QPS
华东1(杭州)	华东1(杭州)	0.26	small	2566	8981
华东1(杭州)	华东1(杭州)	0.26	medium	4726	16541
华东1(杭州)	华东1(杭州)	0.26	large	6378	23204
华东1(杭州)	华北1(青岛)	26	small	2469	8641
华东1(杭州)	华北1(青岛)	26	medium	4856	16996
华东1(杭州)	华北1(青岛)	26	large	5439	20400
华东1(杭州)	华北2(北京)	26	small	2533	8866
华东1(杭州)	华北2(北京)	26	medium	5038	17633
华东1(杭州)	华北2(北京)	26	large	6829	26100
华东1(杭州)	美国西部 1(硅谷)	175	small	1753	6135
华东1(杭州)	美国西部 1(硅谷)	175	medium	2837	9929
华东1(杭州)	美国西部 1(硅谷)	175	large	3884	15500
亚太东南 1(新加坡)	美国西部 1(硅谷)	198	small	1104	4000
亚太东南 1(新加坡)	美国西部 1(硅谷)	198	medium	1724	6334
亚太东南 1(新加坡)	美国西部 1(硅谷)	198	large	2256	8300



说明:

上述测试模型的数据是各种数据迁移链路规格的极限性能，如果数据迁移的表没有主键、网络延迟大、存在更新热点或者源实例与目标实例存在性能瓶颈时，无法达到上表中的增量迁移性能。

3.2 数据同步规格说明

注意事项

该性能规格参考的是测试数据，不作为产品SLA的评判标准。

名词解释

- 规格：数据传输服务为用户提供了不同性能的链路规格，以同步的记录数为衡量标准。
- RPS: 表示每秒同步的SQL语句的次数，包括begin、commit、DML语句（insert、delete、update）及DDL语句。

数据同步规格说明

数据同步根据同步链路的同步性能上限，定义了四种规格：

micro、small、medium、large，当满足以下条件时，各个规格的性能上限才能达到下表中的性能上限。

- 源实例的压力要大于等于各规格对应的性能上限。
- 目标实例的写入性能不是瓶颈，能够支持各规格对应的性能压力。
- 源实例/目标实例与DTS服务器的网络延迟小于等于2ms。

规格定义	性能上限(RPS)
micro	200以内。
small	200~2000。
medium	2000~5000。
large	无限制。 <div style="background-color: #f0f0f0; padding: 5px; margin-top: 10px;">  说明： large规格的线上运行性能依赖于网络环境、源实例和目标实例的性能。 </div>