

阿里云

图数据库 数据导入

文档版本：20220711

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
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

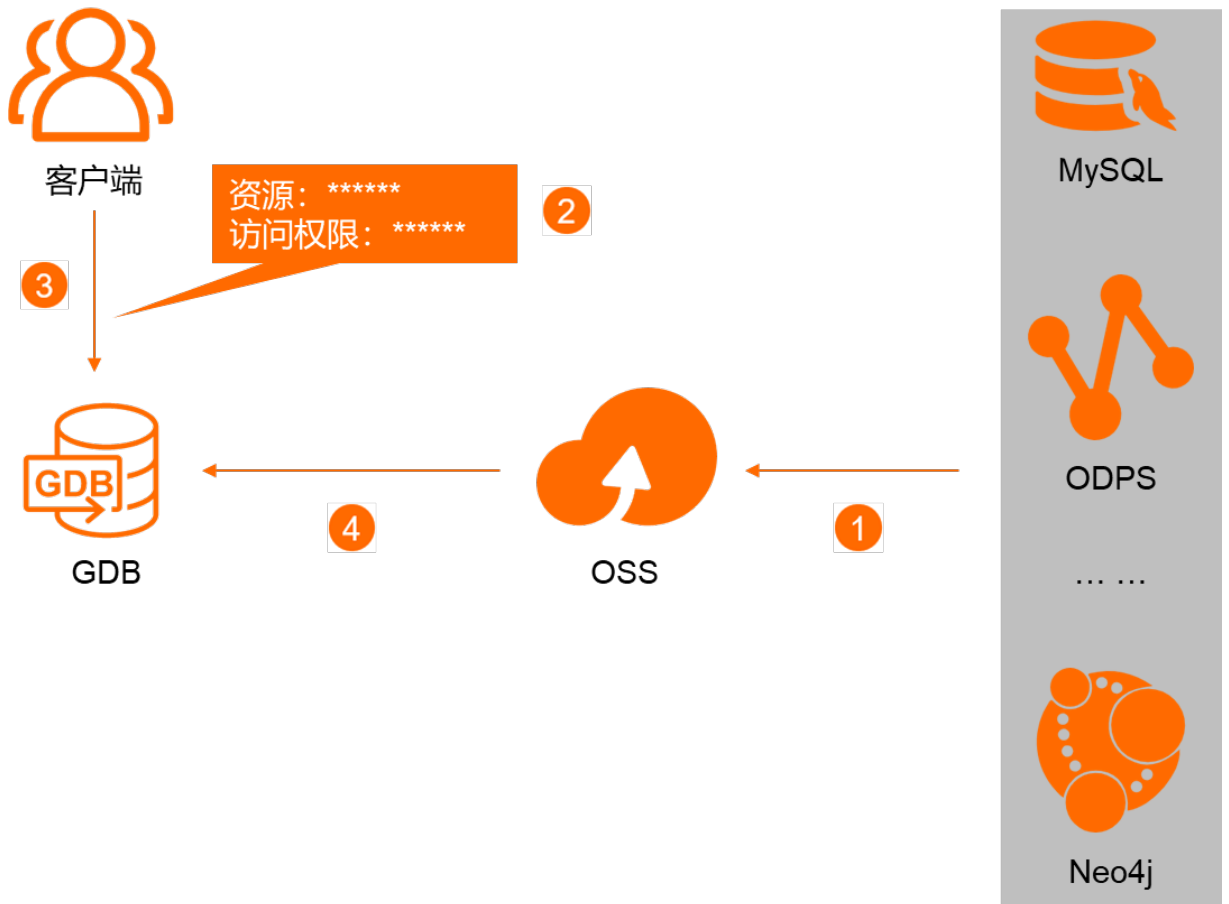
1.数据导入概览	05
2.从阿里云对象存储OSS导入数据至图数据库GDB实例	07
2.1. 从OSS导入数据至图数据库GDB示例	07
2.2. 整体流程介绍	10
2.3. 准备工作	12
2.3.1. 准备工作	12
2.3.2. CSV数据文件格式	14
2.4. OSS授权	20
2.4.1. 通过GDB控制台授权	20
2.4.2. 通过RAM控制台授权	21
2.5. 将数据上传至OSS	23
2.5.1. 使用OSS控制台上传数据文件	23
2.5.2. 使用ossutil工具上传数据文件	25
2.6. 将OSS数据导入至图数据库GDB	26
2.6.1. 导入接口介绍	26
2.6.2. 使用curl命令导入数据文件	32
2.6.3. 使用GdbLoader脚本导入数据文件	35
3.使用DataWorks导入数据至图数据库GDB实例	38
3.1. 配置DataWorks平台GDB Writer	38
3.2. 配置DataWorks平台GDB Reader	44
3.3. 使用DataWorks数据集成从MySQL导入数据到GDB	50

1.数据导入概览

图数据库GDB支持从多种数据源将数据导入至图数据库GDB。本文介绍两种导入方式（数据格式和导入服务）。

从阿里云对象存储OSS导入数据至图数据库GDB实例

图数据库GDB服务端集成有数据导入模块，可以获取您提供的.csv格式的OSS数据文件，并将其解析为图中的点数据和边数据，然后将点数据和边数据导入至图数据库GDB实例。



步骤	步骤说明
①	将需要导入的点数据和边数据分别转换为.csv格式的文件，并上传至OSS（存储对象）的Bucket中。
②	创建对该Bucket有读取和列表访问权限的RAM角色。
③	发送HTTP请求至图数据库GDB实例以启动图数据库GDB加载程序。
④	图数据库GDB加载程序从满足RAM角色的Bucket中读取数据并将数据加载至图数据GDB实例。

使用DataWorks数据集成同步数据到GDB

DataWorks是阿里云上的一站式大数据智能云研发平台，其中数据集成服务支持导出数据到GDB，您也可以使用平台完成数据计算、筛选、转换等处理再同步到GDB。

DataWorks的GDB数据源底层由GdbWriter插件提供服务，提供同等的写入能力，同步配置方法请参见[配置GDB数据源](#)。

2. 从阿里云对象存储OSS导入数据至图数据库GDB实例

2.1. 从OSS导入数据至图数据库GDB示例

本文介绍将数据从OSS导入至图数据库GDB的完整流程。

前提条件

- 已创建图数据库GDB实例，创建方法，请参见[创建主实例](#)。
- 已创建OSS的Bucket，创建方法，请参见[创建存储空间](#)。
- 已获取服务关联角色的ARN（arn:info）信息，获取方法，请参见[授权GDB访问OSS](#)。
- 已获取图数据库GDB的连接地址和端口，获取方法，请参见[获取图数据库GDB实例的信息](#)。

如果您使用本地客户端进行导入操作，连接时需要使用外网地址，申请外网地址的方法，请参见[申请外网地址](#)。

- 已创建数据库账号，创建方法，请参见[创建账号](#)。
- 已将需要访问图数据库GDB的客户端IP地址添加到白名单中，设置白名单的方法，请参见[设置白名单](#)。

准备测试数据

本文将使用air-routes数据集作为测试数据，图数据库GDB的数据分为点文件和边文件，下载链接如下，进入网页另存为即可保存为CSV文件

- [下载点文件](#)。
- [下载边文件](#)。

您也可以使用其它数据进行导入测试，图数据库GDB对CSV数据的文件格式有要求，具体要求，请参见[CSV数据文件格式](#)。

上传数据至OSS

以下内容是通过OSS控制台上传数据。如果您需要通过ossutil工具上传数据，请参见[使用ossutil工具上传数据文件](#)。

1. 登录[OSS管理控制台](#)。
2. 单击左侧导航栏的Bucket列表，然后单击目标Bucket名称。
3. 单击上传文件。
4. 在上传文件面板，配置文件存储目录、读写权限并选择待上传文件。
5. 单击上传文件。

完成上传后，您可以返回文件管理页面查看文件存储路径，在导入数据到图数据库GDB时需要使用该路径。



如上图所示, `gdbOssDemo/node`路径存放了点文件, `gdbOssDemo/edge`路径存放了边文件, 假如Bucket名称为 `gdbbucket`, 则点文件和边文件的OSS路径分别为:

- 点文件

```
oss://gdbbucket/gdbOssDemo/node/
```

- 边文件

```
oss://gdbbucket/gdbOssDemo/edge/
```

导入数据到图数据库GDB

以下内容仅为导入数据到图数据库GDB的示例步骤, 更多参数说明, 请参见[使用curl命令导入数据文件](#)。

1. 在客户端 (例如ECS或本地客户端) 上执行 `curl` 命令, 导入点文件, 示例如下:

```
curl -u gdb_user:PassWord -X POST \
-H 'Content-Type: application/json' \
http://gds-****.graphdb.rds.aliyuncs.com:8182/loader -d '{
  "source" : "oss://gdbbucket/gdbOssDemo/node/",
  "ramRoleArn" : "acs:ram::1*****:role/aliyunserveroleforgdb"
}'
```

说明 请将以上示例中的参数替换为您真实的数据库信息, 使用本地客户端连接时请使用外网地址。

图数据库GDB导入程序检查参数无误后, 将返回导入任务IDloadId, 示例如下:

```
{
  "status" : "200 OK",
  "payload" : {
    "loadId" : "552617AF-4F1E-4CD8-9533-*****"
  }
}
```

2. 在客户端执行 `curl` 命令, 导入边文件, 示例如下:

```
curl -u gdb_user:PassWord -X POST \
-H 'Content-Type: application/json' \
http://gds-****.graphdb.rds.aliyuncs.com:8182/loader -d '
{
  "source" : "oss://gdbbucket/gdbOssDemo/edge/",
  "ramRoleArn" : "acs:ram::1*****:role/aliyunserviceroleforgdb"
}'
```

 **说明** 请将以上示例中的参数替换为您真实的数据库信息，使用本地客户端连接时请使用外网地址。

图数据库GDB导入程序检查参数无误后，将返回导入任务IDloadId，示例如下：

```
{
  "status" : "200 OK",
  "payload" : {
    "loadId" : "552617AF-4F1E-4CD8-9533-*****"
  }
}
```

成功上传后，您可以通过loadId查看导入任务的任务状态，示例如下：

```
curl -u gdb_user:PassWord -X GET 'http://gds-****.graphdb.rds.aliyuncs.com:8182/loader/<loadId>'
```

常见问题

- Q：导入相同的文件后，系统返回 `LOAD_NOT_SUPPORT TO RESUME LOAD` 错误。

A：需要将之前导入数据生成的loadId删除，删除语句如下：

```
curl -u gdb_account:***** -X DELETE 'http://gds-123.graphdb.rds.aliyuncs.com:****/loader/552617AF-4F1E-4CD8-9533-A2EC154688DC'
```

如果有很多loadId，您可以使用以下脚本批量删除：

```
#!/bin/bash
GDB_HOST=YOU_GDB_HOST
GDB_PORT=YOU_GDB_PORT
GDB_USER=YOUR_GDB_USER
GDB_PASSWORD=YOUR_GDB_PASSWORD
curl -u "${GDB_USER}:${GDB_PASSWORD}" "http://${GDB_HOST}:${GDB_PORT}/loader"| python3 -c
'import json,sys;print("\n".join(json.load(sys.stdin)["payload"]["loadIds"]))'| while read
TASK;
do
  echo delete task: $TASK
  curl -u "${GDB_USER}:${GDB_PASSWORD}" -XDELETE "http://${GDB_HOST}:${GDB_PORT}/loader/
${TASK}"
done
curl -u "${GDB_USER}:${GDB_PASSWORD}" "http://${GDB_HOST}:${GDB_PORT}/loader"
```

- Q：导入文件时，提示 `Commit record Error(21-0)` 错误。

A：您需要将导入参数failOnError修改为false，修改操作，请参见[导入接口介绍](#)。

更多错误信息，请参见[错误信息列表](#)。

2.2. 整体流程介绍

本文介绍将数据文件导入至图数据库GDB实例中的前提条件及操作流程。

前提条件

- 已创建图数据库GDB实例，创建方法请参见[创建实例](#)。
- 已创建Bucket，创建方法请参见[创建存储空间](#)。

🔍 说明

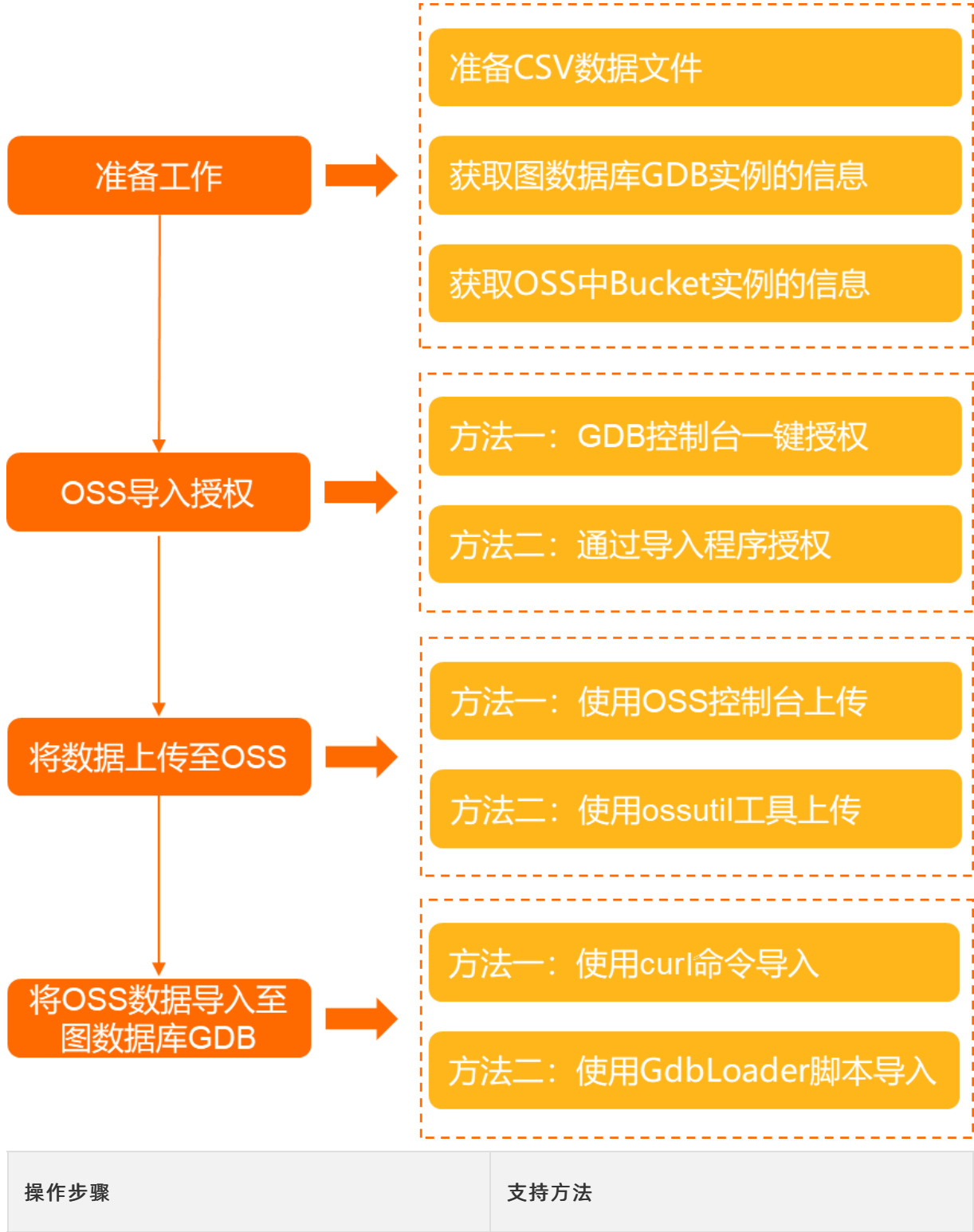
如果是在对象存储OSS控制台创建，必须与已创建的图数据库GDB实例在同一地域。

- 如果应用部署在ECS实例，您需要确保图数据库GDB实例和ECS服务器所在地域、网络类型相同，并建议所在可用区相同，以保证网络的连通性。ECS实例所在地域、可用区和网络类型的查看方法请参见[查看实例信息](#)。

🔍 说明

当ECS服务器的网络类型为经典网络时，您可以将经典网络迁移至专有网络，迁移方法请参见[ECS实例从经典网络迁移到专有网络](#)。

操作流程



操作步骤	支持方法
准备工作	准备工作 包含如下内容： • 准备CSV数据文件 • 获取图数据库GDB实例信息 • 获取OSS中的Bucket信息
OSS导入授权	• 图数据库GDB控制台一键授权 • 通过RAM控制台授权
将数据上传至OSS	• 使用OSS控制台上传数据文件 • 使用ossutil工具上传数据文件
将OSS数据导入至图数据库GDB	• 使用curl命令导入数据文件 • 使用GdbLoader脚本导入数据文件

2.3. 准备工作

2.3.1. 准备工作

将数据文件（点文件和边文件）导入至图数据库GDB实例前，您需要先准备好需要导入的数据文件、获取导入过程中使用到的实例信息等，使导入操作更加流畅。

前提条件

- 已创建图数据库GDB实例，创建方法请参见[创建实例](#)。
- 已创建Bucket，创建方法请参见[创建存储空间](#)。

② 说明

如果是在对象存储OSS控制台创建，必须与已创建的图数据库GDB实例在同一地域。

- 如果应用部署在ECS实例，您需要确保图数据库GDB实例和ECS服务器所在地域、网络类型相同，并建议所在可用区相同，以保证网络的连通性。ECS实例所在地域、可用区和网络类型的查看方法请参见[查看实例信息](#)。

② 说明

当ECS服务器的网络类型为经典网络时，您可以将经典网络迁移至专有网络，迁移方法请参见[ECS实例从经典网络迁移到专有网络](#)。

准备CSV数据文件

- 格式要求：请参见[CSV数据文件格式](#)。
- 构造测试数据文件时，您可以从以下地址下载点文件和边文件。
 - 点文件的下载地址：<https://github.com/krlawrence/graph/blob/master/sample-data/air-routes-latest-nodes.csv>
 - 边文件的下载地址：<https://github.com/krlawrence/graph/blob/master/sample-data/air-routes-latest-edges.csv>

获取图数据库GDB实例的信息

1. 登录[图数据库GDB控制台](#)。
2. 在页面左上角，选择目标实例所在地域。
3. 在实例列表中，单击目标实例ID或目标实例所在行操作列的管理。
4. 根据获取信息，执行对应操作。

- 连接地址和端口

在基本信息页面的基本信息区域，查看目标实例的连接地址（内网地址或外网地址）和端口（内网端口或外网端口）。

基本信息

设置白名单

实例ID	gds-██████████	实例名称	HTTP清除数据 设置
地域可用区	可用区H	实例类型及系列	主实例 高可用版
内网地址	gds-██████████	内网端口	8182
外网地址	gds-██████████	外网端口	3734
VPC网络	vpc-██████████	VPC交换机	vsw-██████████
存储类型	cloud_essd	当前版本号	1.0.27
实例创建时间	2021年11月25日 18:03:54	可维护时间段	02:00-03:00 设置
支持的图查询语言	Gremlin		

- 账号

在左侧导航栏，单击[账号管理](#)。

在账号列表中，查看目标账号信息。

- 如果您忘记目标账号的密码，您可以单击目标账号所在行操作列的重置密码，重新设置密码，设置方法请参见[重置密码](#)。
- 如果您还未创建账号，您可以单击账号列表左上方的创建账号为目标实例创建账号，创建方法请参见[创建账号](#)。

获取OSS中Bucket实例的信息

1. 登录[OSS管理控制台](#)。
2. 在左侧导航栏，单击[Bucket列表](#)。
3. 在Bucket列表中，查看目标Bucket实例的名称和所在地域，并根据所在地域确认该地域Bucket实例的服务接入地址（Endpoint），具体请参见[访问域名和数据中心](#)。

说明

当您在导入数据时 `source` 参数使用 `http://<endpoint-name>/<bucket-name>/<object-key-name>` 格式的URI时，您需要获取Bucket实例的服务接入地址（即替代URI中的 `<endpoint-name>`）。

2.3.2. CSV数据文件格式

本文介绍需要导入至图数据库GDB实例的CSV数据文件格式。

背景信息

CSV数据文件遵循RFC 4180 CSV规范。更多信息，请参见[CSV文件的一般格式和MIME类型](#)。

注意

所有文件必须采用UTF-8格式编码。

注意事项

- 必须将点和边分别放到单独的文件。

说明

加载程序可以在单个加载任务中加载多个点文件和多个边文件。

- 每一个加载任务的加载文件集必须位于Bucket的同一文件夹中，不能使用文件名和文件扩展名来区分点文件和边文件。
- 每个文件都包含一个逗号分隔的标题行，此标题行由系统列标题和属性列标题组成。

标题行是数据文件的第一行，由 `\n` 或 `\r\n` 结束，行内不支持使用双引号、换行符等特殊字符，除分隔符逗号（,）外的其余字符均被解析为标题内容。

系统列标题

标题类型	参数/语法	是否必需	说明
点标题	<code>~id</code>	是	点的ID。
	<code>~label</code>	是	点的标签，只支持单个标签。
	<code>~id</code>	是	边的ID。

标题类型	参数 / 语法	是否必需	说明
边标题	<code>~from</code>	是	源点的点ID。
	<code>~to</code>	是	目标点的点ID。
	<code>~label</code>	是	边的标签，只支持单个标签。

属性列标题

- 属性列为普通属性时，您可以通过以下语法指定属性列，类型名称不区分大小写。

```
propertyname:type
```

 注意

- 列标题中不允许使用空格，因此，属性名称也不能包含空格。
- 属性不支持List数据类型，只能有一个值。

示例：指定名为 `age`、类型为 `Int` 的属性的列标题。

文件中的每行都需要在该位置具有整数或保留为空。

```
age:Int
```

- 属性列为SET类型数据时，您可以使用以下几种语法指定属性列。

 注意

图数据库GDB中仅点支持SET属性，边不支持SET属性，边中设置了SET属性均会导致错误。

```
#兼容默认格式，普通单值属性。
propertyname:type

#普通单值属性
propertyname:type(single)

#SET属性，数据行中包含单值。
propertyname:type(set)

#SET属性，数据行中包含多值。
propertyname:type[]

#SET属性，数据行中包含多值。
propertyname:type(set) []
```

行格式

项目	说明
分隔符	行中字段是用逗号分隔的，行间是用换行符（ <code>\n</code> ）或回车换行符（ <code>\r\n</code> ）分隔的。
空白字段	非必需列（如用户定义的属性） 允许使用空白字段。空白字段仍需要逗号分隔符。
顶点ID	每个顶点文件中所有顶点的 <code>~id</code> 值都必须是唯一的。 <code>~id</code> 值相同的多个顶点行适用于图形中的单个顶点。
边ID	每个边文件中所有边的 <code>~id</code> 值都必须是唯一的。 <code>~id</code> 值相同的多个边行适用于图形中的单个边。
标签	标签是字符串类型，区分大小写。
字符串值	引号是可选的，当包含逗号等特殊字符时需要使用双引号（ <code>"</code> ）括起来。
<code>propertyname: type(set)</code>	SET属性字段，表示数据字段是属性 <code>propertyname</code> 的一个属性值，类型是 <code>type</code> 。  注意 覆盖写入一个SET属性时，相当于追加一个属性值。如果原来是一个普通属性，更新后是一个SET属性。

项目	说明
<code>propertyname: type(set) []</code>	SET属性字段，表示数据字段是属性 <code>propertyname</code> 的属性值数组，数组元素类型是 <code>type</code> 。数组元素之间使用分号（;）分隔。

属性覆盖逻辑

原属性逻辑	新属性逻辑	结果
普通属性	普通属性	原属性值被覆盖，保留最后一次的写入。
普通属性	SET属性	更新后是SET属性，原属性值会合并成SET中的一个属性值。
SET属性	SET属性	SET属性追加，属性值是并集。
SET属性	普通属性	更新后是一个普通属性，属性值是最后一次的写入。

CSV文件格式规范

与SET属性相关的格式规范具体如下：

- SET属性单值字段：保持与原解析格式一致，解析后的值是SET属性的一个属性值。
- SET属性数组字段：使用分号（;）分隔数组元素，例如整型数组字段 `32;45;12` 。
- SET属性数组字段：可以使用双引号（"）将整个字段括起来，解析时会去掉字段前后的双引号，例如整型数组字段 `"32;45;12"` 。
- SET属性数组字段类型是字符串：单个数组元素不能包含有分号（;）字符，会被解析成元素分隔符。

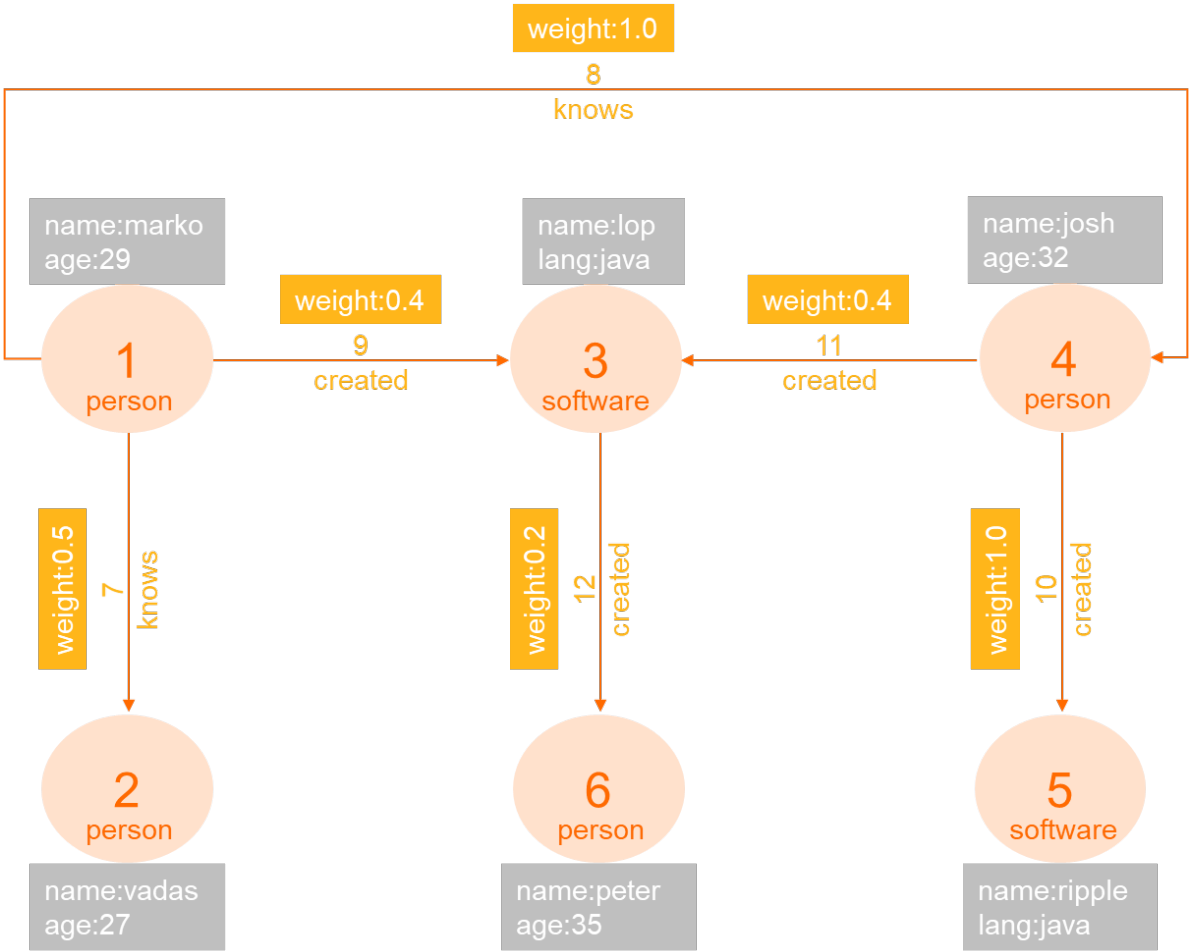
数据类型

属性值类型	字段类型	取值范围
Bool或Boolean	Boolean	• FALSE或false • TRUE或true

属性值类型	字段类型	取值范围
整数	char	-128 ~ 127
	short	-32768 ~ 32767
	int	-2 ³¹ ~ 2 ³¹ - 1
	long	-2 ⁶³ ~ 2 ⁶³ - 1
小数	float	32位IEEE 754浮点
	double	64位IEEE 754浮点
字符串	string	• 双引号是可选的。 • 字符串中如果包含逗号(,)、换行(<code>\r\n</code> 或 <code>\n</code>), 双引号(")等特殊字符, 需要使用双引号(")括起来。示例: "Hello,World" • 如果用双引号括起来的字符串中包含有双引号, 可根据CSV规则使用两个双引号转义。示例: "Hello""World"" • 字符串中不包含特殊字符, 可以不使用双引号括起来。 • 支持长度为零的空字符串, 表示为: " ", 不使用双引号的空白域表示不存在, 为 <code>nullptr</code> 。

普通示例

- 图数据库GDB数据文件（以TinkerPop modern图为例）



- CSV数据文件如下所示：

点文件

~id	~label	name:string	lang:string	age:int
1	person	marko		29
2	person	vadas		27
3	software	lop	java	
4	person	josh		32
5	software	ripple	java	
6	person	peter		35

边文件

~id	~from	~to	~label	weight:double
7	1	2	knows	0.5
8	1	4	knows	1.0
9	1	3	created	0.4
10	4	5	created	1.0
11	4	3	created	0.4
12	6	3	created	0.2

SET属性示例

CSV数据文件

~id	~label	fruits:string(set)[]	codes:long(set)[]
1	person	apple;pear	22
2	person	banana;bitterorange	25;81
3	person	"cherry;blackberry;grape"	3;12
4	person		"54;11"

2.4. OSS授权

2.4.1. 通过GDB控制台授权

从对象存储OSS的Bucket中导入数据时，您需要具有访问该Bucket的权限。本文介绍通过图数据库GDB控制台授权并获取该角色对应的ARN（Alibaba Cloud Resource Name）信息的方法。

授权GDB访问OSS

② 说明 请使用阿里云账号（主账号）或拥有AliyunGDBFullAccess权限的RAM用户进行以下操作。

1. 登录图数据库GDB控制台。
2. 单击数据导入-OSS授权。
3. 根据返回结果判断是否授权成功。
 - 授权成功：系统提示关联角色创建成功，GDB可以正常访问您的OSS。
 - 授权失败：系统提示授权失败。请使用阿里云账号或拥有AliyunGDBFullAccess权限的RAM用户进行操作。

获取ARN信息

授权GDB访问OSS后，系统会自动创建服务关联角色AliyunServiceRoleForGDB，后续使用OSS导入数据时需要使用该角色的ARN（arn:info）信息，获取ARN信息方式如下：

- 登录访问控制控制台查看。



- RAM用户如果无法登录访问控制控制台，可以将以下示例中的 <Aliuid> 替换为阿里云账号UID，即为ARN信息：

```
acs:ram::<Aliuid>:role/aliyunserviceforGDB
```

例如阿里云账号UID为 123456789，那么ARN信息为 acs:ram::123456789:role/aliyunserviceforGDB。

② 说明 您可以使用阿里云账号（主账号）登录阿里云控制台，将鼠标移动到右上角头像处获取阿里云账号UID。



2.4.2. 通过RAM控制台授权

从对象存储OSS的Bucket中导入数据时，您需要具有访问该Bucket的权限。本文介绍通过在RAM（访问控制）控制台创建策略进行授权并创建AccessKey的方法。

操作步骤

- 1. 登录RAM访问控制控制台。
- 2. 创建权限策略。
 - i. 在左侧导航栏，单击权限管理 > 权限策略。
 - ii. 在权限策略页面，单击左上角的创建权限策略。
 - iii. 在创建权限策略页面，单击脚本编辑。
 - iv. 在策略文档输入框中输入以下内容。

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": [
        "oss:GetObject",
        "oss:GetBucket",
        "oss:GetObjectMeta",
        "oss:ListObjects"
      ],
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

 **说明** `Resource`：请您替换为导入数据文件的具体路径。如果当前未规划好导入路径，您可以使用星号（*）代替。

- v. 单击下一步。
- vi. 在基本信息区域，设置以下参数。

参数	参数说明
名称	按照以下规则自定义权限策略的名称。 ■ 包含英文字母、数字和下划线（_）。 ■ 长度不超过128个字符。
备注	建议自定义为具有业务意义的名称（无唯一性要求），便于后续识别。不超过1024个字符。

- vii. 单击确定。
- 3. 为您的阿里云账号创建AccessKey。

- i. 单击页面右上角的，并选择AccessKey管理。



- ii. 在AccessKey页面，单击页面左上角的创建AccessKey。
- iii. 在弹出的查看Secret对话框中，单击下载CSV文件或复制，记录AccessKey ID和AccessKey Secret。
- iv. 单击确定。

2.5. 将数据上传至OSS

2.5.1. 使用OSS控制台上传数据文件

对象存储OSS控制台支持上传大小不超过5 GB的文件，如果您需要导入图数据库GDB实例的数据文件不超过5 GB，您可以将数据文件上传至对象存储OSS的Bucket中，然后从Bucket中将数据文件导入至图数据库GDB实例。本文介绍将数据文件上传至对象存储OSS的Bucket的方法。

前提条件

- 已创建与目标图数据库GDB实例在同一地域、可用区的Bucket实例，创建方法请参见[创建存储空间](#)。
- 已准备好需要导入的CSV数据文件（边文件和点文件大小均不超过5 GB），数据文件具体要求请参见[准备工作](#)。

操作步骤

1. 登录[OSS管理控制台](#)。
2. 单击左侧导航栏的Bucket列表，然后单击目标Bucket名称。
3. 在文件管理页签，单击上传文件。
4. 在上传文件面板，按以下说明完成基础配置项。

参数	说明
上传到	设置文件上传到OSS后的存储路径。 ◦ 当前目录：将文件上传到当前目录。 ◦ 指定目录：将文件上传到指定目录，您需要输入目录名称。若输入的目录不存在，OSS将自动创建对应的文件夹并将文件上传到该文件夹中。
文件ACL	选择文件的读写权限。 ◦ 继承Bucket：以Bucket读写权限为准。 ◦ 私有（推荐）：只有文件Owner拥有该文件的读写权限，其他用户没有权限操作该文件。 ◦ 公共读：文件Owner拥有该文件的读写权限，其他用户（包括匿名访问者）都可以对文件进行访问，这有可能造成您数据的外泄以及费用激增，请谨慎操作。 ◦ 公共读写：任何用户（包括匿名访问者）都可以对文件进行访问，并且向该文件写入数据。这有可能造成您数据的外泄以及费用激增，若被人恶意写入违法信息还可能会侵害您的合法权益。除特殊场景外，不建议您配置公共读写权限。 有关文件ACL的更多信息，请参见Object ACL。
待上传文件	选择您需要上传的文件或文件夹。 您可以单击扫描文件或扫描文件夹选择本地文件或文件夹，或者直接拖拽目标文件或文件夹到待上传文件区域。 如果上传文件夹中包含了无需上传的文件，请单击目标文件右侧的移除将其移出文件列表。  注意 ◦ 如果上传的文件与存储空间中已有的文件重名，则会覆盖已有文件。 ◦ 使用拖拽方式上传文件夹时，OSS会保留文件夹内的所有文件和子文件夹。 ◦ 文件上传过程中，请勿刷新或关闭页面，否则上传任务会被中断且列表会被清空。

如果在上传文件时需要配置高级选项，例如设置自定义元数据，服务端加密方式等，请参见[上传文件](#)。

5. 单击上传文件。

此时，您可以在上传列表页签查看各个文件的上传进度。上传完成后，您可以在目标路径下查看上传文件的文件名、文件大小以及存储类型等信息。

2.5.2. 使用ossutil工具上传数据文件

ossutil支持以命令行形式管理OSS数据，您可以使用ossutil工具将数据文件上传至对象存储OSS的Bucket中。

前提条件

已准备好需要导入的CSV数据文件，数据文件具体要求请参见[准备工作](#)。

操作步骤

1. 在客户端下载并安装ossutil工具，具体请参见[下载和安装](#)。
2. 执行如下命令，创建Bucket。

```
./ossutil64 mb oss://<bucket>
```

? 说明

- `<bucket>`：Bucket的名称，建议根据业务意义自定义，方便后续识别。
- 关于创建Bucket的更多信息，请参见[mb（创建存储空间）](#)。

示例：

```
./ossutil64 mb oss://examplebucket
```

3. 执行如下命令，上传点文件和边文件。

- 上传点文件

```
./ossutil cp air-routes-latest-nodes.csv oss://<bucket>/air-routes/<nodes.csv>
```

? 说明

- `<bucket>`：请替换为创建的Bucket的名称。
- `<nodes.csv>`：请替换为点文件的名称。

示例：

```
./ossutil cp air-routes-latest-nodes.csv oss://examplebucket/air-routes/air-routes-latest-nodes.csv
```

- 上传边文件

```
./ossutil cp air-routes-latest-nodes.csv oss://<bucket>/air-routes/<edges.csv>
```

? 说明

- `<bucket>`：请替换为创建的Bucket的名称。
- `<edges.csv>`：请替换为边文件的名称。

示例：

```
./ossutil cp air-routes-latest-nodes.csv oss://examplebucket/air-routes/air-routes-latest-edges.csv
```

2.6. 将OSS数据导入至图数据库GDB

2.6.1. 导入接口介绍

您可以发送请求至 `<gdb-endpoint>:<port>/loader` 终端，实现将数据文件从OSS的Bucket中导入至图数据库GDB实例，还可以查看任务状态和删除导入任务。本文介绍添加导入任务（POST）、查看导入任务状态（GET）和删除导入任务（DELETE）的应用示例。

前提条件

- 已创建与目标数据库GDB实例，创建方法请参见[创建实例](#)。

说明

如果Bucket是在对象存储控制台创建，请确保图数据库GDB实例和Bucket在同一地域。

- 已将数据文件上传至Bucket中，上传方法请参见[使用OSS控制台上传数据文件](#)或[使用ossutil工具上传](#)。

添加导入任务

使用HTTP POST请求发送至 `<gdb-endpoint>:<port>/loader` 终端，将数据从OSS的Bucket导入至图数据库GDB实例中。

注意

MIME类型必须为application/json。

- 请求语法

```
{
  "source" : "${OSS-Path}",
  "format" : "csv",
  "ramRoleArn" : "string",
  "accessKey" : "string",
  "secretKey" : "string",
  "mode": "NEW|RESUME|AUTO",
  "failOnError" : "TRUE/FALSE",
  "parallelism" : "HIGH/MEDIUM/LOW"
}
```

- 请求参数

参数	类型	说明
<code>source</code>	string	OSS的Bucket中存储数据文件的文件夹URI。 ◦ 可指向单个文件或文件夹的URI。如果指定文件夹，图数据库GDB实例将加载文件夹中的每个数据文件，但不包含子文件夹下的文件。 ◦ 文件夹中可以包含多个点文件和多个边文件。 ◦ 加载过程中先加载点文件，且会自动跳过边文件和非点文件。 URI可以采用以下两种格式： ◦ （推荐）<code>oss://<bucket-name>/<object-key-name></code> ◦ <code>http://<endpoint-name>/<bucket-name>/<object-key-name></code> ❓ 说明 ◦ <code><bucket-name></code>：请替换为Bucket的名称。 ◦ <code><object-key-name></code>：请替换为数据文件名称。 ◦ <code><endpoint-name></code>：请替换为Bucket所在地域对应的服务接入地址。您需要确保图数据库GDB实例与Bucket所属同一地域。
<code>format</code>	string	固定为csv。
<code>ramRoleArn</code>	string	请替换为将数据文件导入Bucket时进行的授权ARN信息，查看方法请参见图数据库GDB控制台一键授权。 ❓ 说明 使用 <code>ramRoleArn</code> 方法进行数据导入时，导入任务受临时安全令牌（STS）有效时间的限制，不能超过10小时，超过会导致任务中断（即数据不能全部导入）。如果您需要导入的数据量较大，建议您采用 <code>accessKey</code> 或 <code>secretKey</code> 的方法导入。
<code>accessKey</code>	string	请替换为您的阿里云账号AccessKey信息的AccessKey ID，查看方法请参见通过访问控制台授权。 ❓ 说明 配置 <code>accessKey</code> 和 <code>secretKey</code> 中的任意一个参数即可。

参数	类型	说明
<code>secretKey</code>	string	请替换为您的阿里云账号AccessKey信息的AccessKey Secret，查看方法请参见通过访问控制台授权。  说明 配置 <code>accessKey</code> 和 <code>secretKey</code> 中的任意一个参数即可。
<code>mode</code>	string	加载任务模式。取值如下： - NEW：将创建新的加载任务，可用于图数据库GDB清除旧数据后重新加载数据。 - RESUME：检查系统是否已经存在当前加载任务（通过 <code>source</code> 判断）。如果当前加载任务存在，则当该加载任务出现终止现象时，系统会自动恢复加载。如果不存在则会停止加载。  注意 - 加载程序会避免重新加载已经成功完成的任务，并且只会尝试处理失败的文件。 - 目前不支持RESUME模式，不能恢复失败的加载任务。 - AUTO：判断是否存在相同的加载任务，并在发现一个时恢复加载，与RESUME模式相似。如果不存在相同的加载任务，则创建一个新的加载任务，与NEW模式相似。 默认值：AUTO
<code>failOnError</code>	string	加载任务异常时是否完全停止。取值如下： - TRUE：会完全停止，但错误点之前的数据仍然存在。 - FALSE：不会完全停止，加载程序尝试加载所有数据并跳过出错的条目。  注意 当设置为FALSE时，会尝试加载解析完成的每条记录，如果记录 <code>id</code> 对应的点或边存在，会尝试更新属性。 默认值：FALSE

参数	类型	说明
<code>parallelism</code>	string	加载任务的并发数。取值如下： ? 说明 - 您需要升级图数据库GDB实例至v1.0.20或以上版本才能使用该选项。 - 限制加载任务占用的系统资源，可降低加载数据时对线上请求的影响。 - HIGH：加载程序的并发数是实例CPU核数量的2倍。 - MEDIUM：加载程序的并发数与实例CPU核数量相等。 - LOW：加载程序的并发数是实例CPU核数量的1/2。 默认值：HIGH

● 响应语法

```
{
  "status" : "200 OK",
  "payload" : {
    "loadId" : "<loaderId>"
  }
}
```

查询成功将返回“200 OK”，其他错误返回可以查看[错误信息列表](#)。

? 说明

- `<loaderId>`：当前导入任务的ID。
- 当错误出现后，响应体中将返回JSON对象。其中，message对象包含对错误的描述。

查看导入任务

使用HTTP GET请求发送到 `<gdb-endpoint>:<port>/loader` 终端，查看导入任务的状态。

? 说明

查看指定导入任务的状态，必须包含 `<loaderId>` 作为URL参数，也可将 `<loaderId>` 追加到URL路径。

● 请求语法

```
#查看导入任务
GET <gdb-endpoint>:<port>/loader

#查看指定的导入任务
#方法一：
GET <gdb-endpoint>:<port>/loader/<loaderId>
#方法二：
GET <gdb-endpoint>:<port>/loader?loaderId=<loaderId>
```

- ❓ 说明
- `<gdb-endpoint>`：请替换为目标数据库GDB实例的连接地址（内网地址或外网地址）。
 - `<port>`：请替换为目标数据库GDB实例连接地址对应的端口（内网端口或外网端口）。
 - `<loaderId>`：请替换为指定任务的ID。如果不指定，则返回所有加载任务的ID列表。

● 返回语法

```
{
  "payload":{
    "datatypeMismatchErrors": long,
    "errors":[],
    "fullUri": "${OSS-Path}",
    "insertErrors": long,
    "loaderId": "${loaderId}",
    "parsingErrors": long,
    "retryNumber": long,
    "runNumber": long,
    "status": "LOAD_COMPLETED",
    "totalDuplicates": long,
    "totalRecords": long,
    "totalTimeSpent": long
  },
  "status": "200 OK"
}
```

● 返回参数

参数	类型	说明
<code>fullUri</code>	string	加载文件的URI。
<code>runNumber</code>	long	加载任务运行编号，重新启动时增加。

参数	类型	说明
<code>retryNumber</code>	long	加载任务重试编号，运行期间自动重试时增加。
<code>status</code>	string	加载任务当前状态。 LOAD_COMPLETED: 加载成功。
<code>totalTimeSpent</code>	long	加载任务耗费的时间（ms），包括提取文件列表、解析文件、下载文件和加载文件到数据库实例的时间。
<code>totalRecords</code>	long	已经加载的记录总数。
<code>totalDuplicates</code>	long	重复记录的数量。
<code>parsingErrors</code>	long	解析错误的数量。
<code>datatypeMismatchErrors</code>	long	数据类型与指定数据类型不匹配的记录数量。
<code>insertErrors</code>	long	由于错误而无法插入的记录数量。
<code>loaderId</code>	string	加载任务的ID。
<code>errors</code>	List	错误日志列表。 最多保留1000条记录，读取后清空。

查询成功将返回“200 OK”，其他错误返回可以查看[错误信息列表](#)。

取消导入任务

使用HTTP DELETE请求发送到 `<gdb-endpoint>:<port>/loader` 终端，取消导入任务。

② 说明

取消导入任务，必须包含 `<loaderId>` 作为URL参数，也可将 `<loaderId>` 追加到URL路径。

● 请求语法

```
DELETE <gdb-endpoint>:<port>/loader?loaderId=<loaderId>
```

```
DELETE <gdb-endpoint>:<port>/loader/<loaderId>
```

② 说明

- `<gdb-endpoint>`：请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>`：请替换为图数据库GDB实例连接地址对应的端口（内网端口或外网端口）。
- `<loaderId>`：请替换为需要取消的导入任务ID。

取消任务请求完成时，服务端会清理该任务的记录信息。对于取消失败的任务，您可以在查明原因并进行修复后重新进行取消。

● 响应语法

```
no response body
```

通过HTTP状态码反馈请求执行结果。查询成功将返回“200 OK”，其他错误返回可以查看[错误信息列表](#)。

2.6.2. 使用curl命令导入数据文件

本文介绍使用curl命令将数据文件从对象存储OSS的Bucket中导入至图数据库GDB实例、查看导入任务和取消导入任务的方法。

前提条件

- 已创建目标图数据库GDB实例，创建方法请参见[创建主实例](#)。

② 说明 如果Bucket是在对象存储控制台创建，请确保图数据库GDB实例和Bucket在同一地域。

- 已将数据文件上传至Bucket中，上传方法请参见[使用OSS控制台上传数据文件](#)或[使用ossutil工具上传数据文件](#)。

操作步骤

1. 在客户端执行如下命令，将存储至对象存储OSS的Bucket中的数据文件导入至图数据库GDB实例。

```
curl -i -u <username>:<password> -X POST \
-H 'Content-Type: application/json' \
http://<mygdb-endpoint>:<port>/loader -d '
{
  "source" : "oss://<bucket>/air-routes/<filename>",
  "ramRoleArn" : "<arninfo>",
}
```

? 说明

- `<username>` : 请替换为图数据库GDB实例的账号名称。
- `<password>` : 请替换为图数据库GDB实例的账号密码。
- `<mygdb-endpoint>` : 请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>` : 请替换为图数据库GDB实例连接地址的端口号（内网端口或外网端口）。
- `<bucket>` : 请替换为存储数据文件的Bucket的名称。
- `<filename>` : 请替换为需要导入的数据文件的名称。
- `<arninfo>` : 请替换为您的阿里云账号中AliyunServiceRoleForGDB角色的ARN信息，具体请参见[获取ARN信息](#)。
- 复制时不要在 \ 后加空格

示例：

◦ 导入点文件

```
curl -i -u gdb_account:***** -X POST \
-H 'Content-Type: application/json' \
http://gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com:****/loader -d '
{
  "source" : "oss://examplebucket/air-routes/air-routes-latest-nodes.csv",
  "ramRoleArn" : "acs:ram::140692647406****:role/aliyunserviceroleforgdb",
}
```

◦ 导入边文件

```
curl -i -u gdb_account:***** -X POST \
-H 'Content-Type: application/json' \
http://gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com:****/loader -d '
{
  "source" : "oss://examplebucket/air-routes/air-routes-latest-edges.csv",
  "ramRoleArn" : "acs:ram::140692647406****:role/aliyunserviceroleforgdb",
}
```

回显如下类似信息，说明图数据库GDB导入程序检查参数正确，并返回导入任务ID（`loadId`）。

```
{
  "status" : "200 OK",
  "payload" : {
    "loadId" : "552617AF-4F1E-4CD8-9533-A2EC154688DC"
  }
}
```

2. 根据导入任务ID，执行对应命令，查看导入任务状态或取消导入任务。

○ 查看导入任务状态

```
curl -u <username>:<password> -X GET 'http://<mygdb-endpoint>:<port>/loader/<loadId>'
```

🔍 说明

- `<username>`：请替换为图数据库GDB实例的账号名称。
- `<password>`：请替换为图数据库GDB实例的账号密码。
- `<mygdb-endpoint>`：请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>`：请替换为图数据库GDB实例连接地址的端口号（内网端口或外网端口）。
- `<loadId>`：请替换为中返回的导入任务ID。

示例：

```
curl -u gdb_account:***** -X GET 'http://gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com:****/loader/552617AF-4F1E-4CD8-9533-A2EC154688DC'
```

如果回显信息中包含错误信息，具体请参见[错误信息列表](#)。

○ 取消导入任务

🔔 注意

- 取消正在后台执行的任务会阻塞等待任务停止再返回。
- 任务中已经导入至图数据库GDB实例的点数据或边数据不会回滚，将保留在图数据库GDB实例中。

```
curl -u <username>:<password> -X DELETE 'http://<mygdb-endpoint>:<port>/loader/<loadId>'
```

🔍 说明

- `<username>`：请替换为图数据库GDB实例的账号名称。
- `<password>`：请替换为图数据库GDB实例的账号密码。
- `<mygdb-endpoint>`：请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>`：请替换为图数据库GDB实例连接地址的端口号（内网端口或外网端口）。
- `<loadId>`：请替换为中返回的导入任务ID。

示例：

```
curl -u gdb_account:***** -X DELETE 'http://gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com:****/loader/552617AF-4F1E-4CD8-9533-A2EC154688DC'
```

2.6.3. 使用GdbLoader脚本导入数据文件

本文介绍使用GdbLoader脚本将数据文件从对象存储OSS的Bucket中导入至图数据库GDB实例、查看导入任务和取消导入任务的方法。

前提条件

- 已创建目标图数据库GDB实例，创建方法请参见[创建主实例](#)。

 **说明** 如果Bucket是在对象存储控制台创建，请确保图数据库GDB实例和Bucket在同一地域。

- 已将数据文件上传至Bucket中，上传方法请参见[使用OSS控制台上传数据文件](#)或[使用ossutil工具上传数据文件](#)。
- 已在客户端安装Python，具体请参见[安装Python](#)。

操作步骤

- 在客户端下载并安装GdbLoader.py工具。

- 执行如下命令，下载图数据库GDB工具包。

```
wget https://github.com/aliyun/alibabacloud-gdb-tools/archive/refs/heads/master.zip
```

- 执行如下命令，解压图数据库GDB工具包。

```
unzip master.zip
```

- 执行如下命令，添加导入任务。

- 添加导入点文件任务

```
python GdbLoader.py --host <mygdb-endpoint> --port <port> --username <username> --password <password> --todo add_task --source oss://<bucket>/air-routes/<nodes.csv> --arn <arninfo>
```

说明

- <mygdb-endpoint>**：请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- <port>**：请替换为图数据库GDB实例连接地址对应的端口号（内网端口或外网端口）。
- <username>**：请替换为图数据库GDB实例的账号名称。
- <password>**：请替换为图数据库GDB实例的账号密码。
- <bucket>**：请替换为存储数据文件的Bucket的名称。
- <nodes.csv>**：请替换为需要导入的点文件的名称。
- <arninfo>**：请替换为您的阿里云账号中AliyunServiceRoleForGDB角色的ARN信息。

示例：

```
python GdbLoader.py --host gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com --port
**** --username gdb_account --password **** --todo add_task --source oss://exampl
ebucket/air-routes/air-routes-latest-nodes.csv --arn acs:ram::140692647406****:role/a
liyunserveroleforgdb
```

◦ 添加导入边文件任务

```
python GdbLoader.py --host <mygdb-endpoint> --port <port> --username <username> --pas
sword <password> --todo add_task --source oss://<bucket>/air-routes/<edges.csv> --arn
<arninfo>
```

❓ 说明

- `<mygdb-endpoint>` : 请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>` : 请替换为图数据库GDB实例连接地址对应的端口号（内网端口或外网端口）
- `<username>` : 请替换为图数据库GDB实例的账号名称。
- `<password>` : 请替换为图数据库GDB实例的账号密码。
- `<bucket>` : 请替换为存储数据文件的Bucket的名称。
- `<edges.csv>` : 请替换为需要导入的边文件的名称。
- `<arninfo>` : 请替换为您的阿里云账号中AliyunServiceRoleForGDB角色的ARN信息。

示例：

```
python GdbLoader.py --host gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com --port
**** --username gdb_account --password **** --todo add_task --source oss://exampl
ebucket/air-routes/air-routes-latest-edges.csv --arn acs:ram::140692647406****:role/a
liyunserveroleforgdb
```

回显如下类似信息，说明图数据库GDB导入程序检查参数正确，并返回导入任务ID（ `loadId` ）。

```
{
  "status" : "200 OK",
  "payload" : {
    "loadId" : "552617AF-4F1E-4CD8-9533-A2EC154688DC"
  }
}
```

3. 根据导入任务ID，执行对应命令，查看导入任务状态或取消导入任务。

◦ 查看导入任务状态

```
python GdbLoader.py --host <mygdb-endpoint> --port <port> --username <username> --pas
sword <password> --todo get_task --loaderId <loadId>
```

说明

- `<mygdb-endpoint>`：请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>`：请替换为图数据库GDB实例连接地址对应的端口号（内网端口或外网端口）。
- `<username>`：请替换为图数据库GDB实例的账号名称。
- `<password>`：请替换为图数据库GDB实例的账号密码。
- `<loadId>`：请替换为中返回的导入任务ID。

示例：

```
python GdbLoader.py --host gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com --port
**** --username gdb_account --password **** --todo get_task --loaderId 552617AF-4F
1E-4CD8-9533-A2EC154688DC
```

取消导入任务

注意

- 取消正在后台执行的任务会阻塞等待任务停止再返回。
- 任务中已经导入至图数据库GDB实例的点数据或边数据不会回滚，将保留在图数据库GDB实例中。

```
python GdbLoader.py --host <mygdb-endpoint> --port <port> --username <username> --pas
sword <password> --todo delete_task --loaderId <loadId>
```

说明

- `<mygdb-endpoint>`：请替换为图数据库GDB实例的连接地址（内网地址或外网地址）。
- `<port>`：请替换为图数据库GDB实例连接地址对应的端口号（内网端口或外网端口）。
- `<username>`：请替换为图数据库GDB实例的账号名称。
- `<password>`：请替换为图数据库GDB实例的账号密码。
- `<loadId>`：请替换为中返回的导入任务ID。

示例：

```
python GdbLoader.py --host gds-bp1t4x427r15331915****.graphdb.rds.aliyuncs.com --port
**** --username gdb_account --password **** --todo delete_task --loaderId 552617AF
-4F1E-4CD8-9533-A2EC154688DC
```

3.使用DataWorks导入数据至图数据库GDB实例

3.1. 配置DataWorks平台GDB Writer

本文为您介绍GDB Writer支持的数据类型、字段映射和数据源等参数及配置示例。

说明

- GDB Writer仅支持使用[独享数据集成资源组](#)，不支持使用默认资源组和[自定义资源组](#)。
- 开始配置GDB Writer插件前，请首先配置好数据源，详情请参见[配置Graph Database数据源](#)。
- 由于点和边的数据集成任务的配置不同，请您分别配置点和边的数据集成任务。

使用须知

- 必须先运行点的同步任务，运行成功后，才可以运行边的同步任务。
- 点有以下约束规则：
 - 点必须具备类型名（即点名称，对应label）。
 - 点的主键ID为必选，必须保证在点范围内唯一，且类型必须是string（如果不是string类型，GDB Writer插件会强制转换）。
 - 请谨慎选择点的主键映射规则idTransRule。如果选择None，需要保证点的ID在全局点的范围内唯一。
- 边有以下约束规则：
 - 边必须具备类型名（即边名称，对应label）。
 - 边的主键ID为可选：
 - 如果填写，则需要保证在全局边范围内唯一。
 - 如果不填写，则GDB服务端默认生成一个UUID，类型必须是string（如果不是string类型，GDB Writer插件会强制转换）。
 - 请谨慎选择边的主键映射规则idTransRule。如果选择None，需要保证边的ID在全局点边的范围内唯一。
 - 边必须选择srcIdTransRule和dstIdTransRule，且必须和导入点时选择的idTransRule一致。
- 示例的字段名或枚举值，如果没有特殊说明，均为大小写敏感。
- 目前GDB服务端仅支持UTF-8编码格式，要求来源数据均为UTF-8编码格式。
- 由于网络限制，运行数据集成任务时，只能使用[独享资源模式](#)。

参数说明

参数	描述	是否必选	默认值
datasource	数据源名称，脚本模式支持添加数据源，此配置项填写的内容必须与添加的数据源名称保持一致。	是	无

参数	描述	是否必选	默认值
labelType	写入数据的类型： - 枚举值 <i>VERTEX</i> 表示点。 - 枚举值 <i>EDGE</i> 表示边。	是	无
writeMode	导入ID重复时的处理模式： - 枚举值 <i>INSERT</i> 表示写入重复ID会报错，错误记录数加1。 - 枚举值 <i>MERGE</i> 表示写入重复ID时，属性用新值覆盖旧值，不计入错误。 - 枚举值 <i>SKIP</i> 表示写入重复ID时直接跳过，也不计入错误。	是	INSERT
label	类型名，即点或边名称。 label支持从源列中读取，例如 <code>#{0}</code> ，表示取第1列字段作为类型名，源列索引从0开始。	是	无
srcLabel	表示起点的点名称： - 当写入类型为点时，不需要填写。 - 当写入类型为边，并且srcIdTransRule为 <i>none</i>，也不需要填写，否则为必填项。	否	无
dstLabel	表示终点的点名称： - 当写入类型为点时，不需要填写。 - 当写入类型为边，并且srcIdTransRule为 <i>none</i>，也不需要填写，否则为必填项。	否	无
idTransRule	主键ID的转换规则： - 枚举值 <i>none</i> 表示映射的值不进行转换，保持主键源字段。 - 枚举值 <i>labelPrefix</i> 表示将映射的值转换为 <code>{label名}{源字段}</code>	是	none
srcIdTransRule	当写入边类型时，表示起点主键ID的转换规则： - 枚举值 <i>none</i> 表示映射的值不进行转换，保持主键源字段，此时可以不填写srcLabel - 枚举值 <i>labelPrefix</i> 表示将映射的值转换为 <code>{label名}{源字段}</code>	写入边类型时必选	none
dstIdTransRule	当写入边类型时，表示终点主键ID的转换规则： - 枚举值 <i>none</i> 表示映射的值不进行转换，保持主键源字段，此时可以不填写dstLabel。 - 枚举值 <i>labelPrefix</i> 表示将映射的值转换为 <code>{label名}{源字段}</code>	写入边类型时必选	none
column::name	点或边字段映射关系配置name字段。 点或边的字段名，当映射关系为属性时表示属性名。	是	无

参数	描述	是否必选	默认值
column::type	点或边字段映射关系配置type字段。 点或边字段映射值的类型： - 主键ID仅支持 <i>string</i> 类型，您需要保证ID可以转换为 <i>string</i> 类型，插件会强制转换非 <i>string</i> 类型到 <i>string</i> 类型。 - 普通属性支持的类型：<i>int</i>, <i>long</i>, <i>float</i>, <i>double</i>, <i>boolean</i>, <i>string</i>	是	无
column::value	点或边字段映射关系配置value字段。 点或边字段映射的值，仅字符串类型的值支持自定义拼接： <code>#N</code> 表示直接映射源端值，N为源端column索引，从0开始。 <code>test-#{0}</code> 表示对源端值进行拼接转换，<code>#{0}</code> 值前或后都可以添加固定字符串。 <code>#{0}-#{1}</code> 表示进行多字段拼接，也可以在任意位置添加固定字符串，例如 <code>test-#{0}-test1-#{1}-test2</code>	是	无
column::columnType	点或边字段映射关系配置columnType字段。 点或边字段映射字段的分类，支持以下枚举值： <i>primaryKey</i>表示该字段映射主键ID。 <i>vertexProperty</i>写入类型为点时可选，表示该字段是点的普通属性。 <i>vertexSetProperty</i>写入类型为点时可选，表示该字段是点的SET属性，多个属性值分开多个column，属性值要求基础类型。 <i>vertexJsonProperty</i>写入类型为点时可选，表示该字段是点属性的JSON格式，value结构请参见properties示例。 <i>srcPrimaryKey</i>写入类型为边时必选，表示该字段映射起点主键ID。 <i>dstPrimaryKey</i>写入类型为边时必选，表示该字段映射终点主键ID。 <i>edgeProperty</i>写入类型为边时可选，表示该字段是边的普通属性。 <i>edgeJsonProperty</i>写入类型为边时可选，表示该字段是边属性的JSON格式，value结构请参见properties示例。	是	无

properties示例如下：

```
{ "properties": [
  { "k": "name", "t": "string", "v": "tom" },
  { "k": "age", "t": "int", "v": "20" },
  { "k": "sex", "t": "string", "v": "male" }
]}
```

向导开发介绍

暂不支持向导模式开发。

脚本开发介绍

配置一个写入GDB的数据同步作业，通过脚本模式开发请参见[脚本模式配置任务](#)。

- 点配置示例

```
{
  "order":{
    "hops":[
      {
        "from":"Reader",
        "to":"Writer"
      }
    ]
  },
  "setting":{
    "errorLimit":{
      "record":"100" //错误记录数，表示脏数据最大容忍条数。
    },
    "jvmOption":"","
    "speed":{
      "concurrent":3,
      "throttle":false
    }
  },
  "steps":[
    {
      "category":"reader",
      "name":"Reader",
      "parameter":{
        "column":[
          "*"
        ],
        "datasource":"_ODPS",
        "emptyAsNull":true,
        "guid":"","
        "isCompress":false,
        "partition":[],
        "table":""
      },
      "stepType":"odps"
    },
    {
      "category":"writer",
      "name":"Writer",
      "parameter": {
        "datasource": "testGDB", //数据源名称。
        "label": "person", //label名，即点名称。
        "srcLabel": "", //点类型时此字段无需关注。
        "dstLabel": "", //点类型时此字段无需关注。
        "labelType": "VERTEX", //label类型，"VERTEX"表示点。
        "writeMode": "INSERT", //导入ID重复时处理方式。
        "idTransRule": "labelPrefix", //点的主键转换规则。
        "srcIdTransRule": "none", //点类型时此字段无需关注。
        "dstIdTransRule": "none", //点类型时此字段无需关注。
      }
    }
  ]
}
```

```

        "column": [
            {
                "name": "id",           //字段名。
                "value": "#{0}",       //#{0}表示取源端第1个字段值，支持拼接，0是源
端column索引号。
                "type": "string",      //字段类型。
                "columnType": "primaryKey" //字段分类，primaryKey表示是主键。
            },                       //点的主键，字段名必须是ID且类型是string，该
记录必须存在。
            {
                "name": "person_age",
                "value": "#{1}",       //#{1}表示取源端第2个字段值，同上支持拼接。
                "type": "int",
                "columnType": "vertexProperty" //字段分类，vertexProperty表示是点
的属性。
            },                       //点的属性，支持int、long、float、double、b
oolean和string类型。
            {
                "name": "person_credit",
                "value": "#{2}",       //#{2}表示取源端第3个字段值，同上支持拼接。
                "type": "string",
                "columnType": "vertexProperty"
            },                       //点的属性。
        ]
    },
    "stepType": "gdb"
}
],
"type": "job",
"version": "2.0"
}

```

● 边配置示例

```

{
    "order": {
        "hops": [
            {
                "from": "Reader",
                "to": "Writer"
            }
        ]
    },
    "setting": {
        "errorLimit": {
            "record": "100"           //错误记录数，表示脏数据的最大容忍条数。
        },
        "jvmOption": "",
        "speed": {
            "concurrent": 3,
            "throttle": false
        }
    },
    "steps": [
        {

```

```

        "category": "reader",
        "name": "Reader",
        "parameter": {
            "column": [
                "*"
            ],
            "datasource": "_ODPS",
            "emptyAsNull": true,
            "guid": "",
            "isCompress": false,
            "partition": [],
            "table": ""
        },
        "stepType": "odps"
    },
    {
        "category": "writer",
        "name": "Writer",
        "parameter": {
            "datasource": "testGDB", //数据源名称。
            "label": "use", //label名，即边名称。
            "labelType": "EDGE", //label类型，EDGE表示边。
            "srcLabel": "person", //起点的点名称。
            "dstLabel": "software", //终点的点名称。
            "writeMode": "INSERT", //导入ID重复时的处理方式。
            "idTransRule": "labelPrefix", //边的主键转换规则。
            "srcIdTransRule": "labelPrefix", //起点的主键转换规则。
            "dstIdTransRule": "labelPrefix", //终点的主键转换规则。
            "column": [
                {
                    "name": "id", //字段名。
                    "value": "#{0}", //#{0}表示取源端第1个字段值，支持拼接。
                    "type": "string", //字段类型。
                    "columnType": "primaryKey" //字段分类，primaryKey表示该字段是主键
                },
                {
                    "name": "id", //支持拼接，注意映射规则要与录入点时一致
                    "value": "#{1}",
                    "type": "string",
                    "columnType": "srcPrimaryKey" //字段分类，srcPrimaryKey表示是起点主
                },
                {
                    "name": "id",
                    "value": "#{2}", //支持拼接，注意映射规则要与录入点时一致
                    "type": "string",
                    "columnType": "dstPrimaryKey" //字段分类，dstPrimaryKey表示是终点主
                }
            ]
        }
    }
}

```

g, 该记录选填。

键。

ing, 该记录必须存在。

键。

键。

```
ing, 该记录必须存在。
    {
      "name": "person_use_software_time",
      "value": "#{3}",           //支持拼接。
      "type": "long",
      "columnType": "edgeProperty" //字段分类, edgeProperty表示边的属性。
    },                               //边的属性, 支持int、long、float、double、boolean和string类型。
    {
      "name": "person_regist_software_name",
      "value": "#{4}",           //支持拼接。
      "type": "string",
      "columnType": "edgeProperty"
    },                               //边属性
    {
      "name": "id",
      "value": "#{5}",           //支持拼接。
      "type": "long",
      "columnType": "edgeProperty"
    },                               //边的属性, 字段名是ID。与主键ID不同,
    //该字段为普通属性, 可选。
  ]
  "stepType": "gdb"
}
},
"type": "job",
"version": "2.0"
}
```

② 说明

GDB需要升级到1.0.20或更高版本才支持SET属性, 使用SET属性时请确认实例版本。

3.2. 配置DataWorks平台GDB Reader

本文为您介绍GDB Reader支持的数据类型、字段映射和数据源等参数及配置示例。

② 说明

- GDB Reader仅支持使用[独享数据集成资源组](#)和[自定义资源组](#)。
- 开始配置GDB Reader插件前, 请首先配置好数据源, 详情请参见[配置Graph Database数据源](#)。
- 由于点和边的数据集成任务的配置不同, 请您分别配置点和边的数据集成任务。

使用须知

- 需要分开配置点任务和边任务来导出GDB的点数据和边数据。
- 导出任务按点或边的类型名(即名称, 对应label)遍历数据, 因此您需要正确提供待导出数据类型名。
- GDB中点和边的主键ID字段是字符串类型, 导出应该配置成string类型, 如果配置long等数值类型, 插件会转换成配置的目标类型, 但转换失败时会丢失该记录。

- GDB导出的属性值需要配置成存储类型，如果存储类型与配置类型不一致，插件会转换成目标类型，但可能会转换失败导致该记录丢失。
- 导出点SET属性的一个值时，不能保证多次导出都是同一个值。
- 导出所有属性使用JSON格式输出时，只包含一个属性值的SET属性会被当成普通属性输出。
- 示例的字段名或枚举值，如果没有特殊说明，均为大小写敏感。
- 目前GDB服务端仅支持UTF-8编码格式，导出的数据均为UTF-8编码格式。
- 由于网络限制，运行数据集成任务时，只能使用[独享资源模式](#)。

参数说明

参数	描述	是否必选	默认值
datasource	数据源名称，脚本模式支持添加数据源，此配置项填写的内容必须与添加的数据源名称保持一致。	是	无
labelType	写入数据的类型 ● 枚举值 <i>VERTEX</i> 表示点。 ● 枚举值 <i>EDGE</i> 表示边。	是	无
labels	类型名，即点/边名称；支持读取多个名称数据，用数组表示。	是	无
column::name	点/边映射关系配置 name 字段。 读取属性时必选，提供属性名。	是	无
column::type	点/边映射关系配置 type 字段。 ● 主键ID、类型名label在GDB中都是 <i>string</i> 类型，配置非 <i>string</i> 类型时可能会转换失败。 ● 普通属性支持类型：<i>int</i>, <i>long</i>, <i>float</i>, <i>double</i>, <i>boolean</i>, <i>string</i> ● GDB Reader 尽量将读取到的数据转换成配置要求的类型，但转换失败会导致该条记录错误。	是	无

参数	描述	是否必选	默认值
column::columnType	点/边字段映射关系配置 columnType 字段。 点/边字段映射字段的分类，支持以下枚举值。 • <i>primaryKey</i> 表示该字段映射主键ID。 • <i>primaryLabel</i> 表示该字段映射点/边名称。 • <i>vertexProperty</i> 导出类型为点时可选，表示该字段是点的属性。 • <i>vertexJsonProperty</i> 导出类型为点时可选，表示该字段是点的属性集合，用JSON格式封装。当配置该类型时，所有属性都会打包到该列，column里面不能再包含其他属性类型。 • <i>srcPrimaryKey</i> 导出类型为边时可选，表示该字段是起点主键ID。 • <i>dstPrimaryKey</i> 导出类型为边时可选，表示该字段是终点主键ID。 • <i>srcPrimaryLabel</i> 导出类型为边时可选，表示该字段是起点的类型名称label。 • <i>dstPrimaryLabel</i> 导出类型为边时可选，表示该字段是终点的类型名称label。 • <i>edgeProperty</i> 导出类型为边时可选，表示该字段是边的属性。 • <i>edgeJsonProperty</i> 导出类型为边时可选，表示该字段是边的属性集合，用JSON格式封装。当配置该类型时，所有属性都会打包到该列，column里面不能再包含其他属性类型。	是	无

vertexJsonProperty格式

```
{ "properties": [
  { "k": "name", "t": "string", "v": "tom", "c": "set" },
  { "k": "name", "t": "string", "v": "jack", "c": "set" },
  { "k": "sex", "t": "string", "v": "male", "c": "single" }
]}
```

edgeJsonProperty格式

```
{ "properties": [
  { "k": "name", "t": "string", "v": "tom" },
  { "k": "sex", "t": "string", "v": "male" }
]}
```

注意

- 边不支持多值属性，封装成JSON格式时，没有 `c` 字段。
- 如果点的多值属性只包含一个属性值时，导出会被当成单值属性处理。

向导开发介绍

暂不支持向导模式开发。

脚本开发介绍

配置一个导出GDB的数据同步作业，通过脚本模式开发请参见[脚本模式配置任务](#)。

- 点配置示例

```
{
  "order":{
    "hops":[
      {
        "from":"Reader",
        "to":"Writer"
      }
    ],
  },
  "setting":{
    "errorLimit":{
      "record":"100" //错误记录数，表示脏数据最大容忍条数。
    },
    "jvmOption":"","
    "speed":{
      "concurrent":3,
      "throttle":false
    }
  },
  "steps":[
    {
      "category":"reader",
      "name":"Reader",
      "parameter":{
        "datasource": "testGDB", //数据源名称。
        "labelType": "VERTEX", // label类型，"VERTEX"表示点
        "labels": ["label1", "label2"], // 类型名称label列表，为空表示导出所有的点
        "column": [
          {
            "name": "id", // 字段名
            "type": "string", // 字段类型
            "columnType": "primaryKey" // 字段分类，表示点的主键id，GDB中是string类型
          },
          {
            "name": "label", // 字段名
            "type": "string", // 字段类型
            "columnType": "primaryLabel" // 字段分类，表示点的label名，GDB中是string类型
          },
          {
            "name": "age", // 属性字段名
            "type": "int", // 属性字段类型
            "columnType": "vertexProperty" // 字段分类，表示点的属性，GDB中基础类型属性
          }
        ]
      },
      "stepType":"gdb"
    },
    {
```

```

        "category": "writer",
        "name": "Writer",
        "parameter": {
            "print": false,          //是否向屏幕打印输出。
            "fieldDelimiter": ",",  //列分隔符。
        }
        "stepType": "stream"
    }
]
"type": "job",
"version": "2.0"
}

```

● 边配置示例

```

{
    "order": {
        "hops": [
            {
                "from": "Reader",
                "to": "Writer"
            }
        ]
    },
    "setting": {
        "errorLimit": {
            "record": "100"  //错误记录数，表示脏数据最大容忍条数。
        },
        "jvmOption": "",
        "speed": {
            "concurrent": 3,
            "throttle": false
        }
    },
    "steps": [
        {
            "category": "reader",
            "name": "Reader",
            "parameter": {
                "datasource": "testGDB", //数据源名称
                "labelType": "EDGE", // label类型, "EDGE"表示边
                "labels": ["label1", "label2"], // 类型名称label列表，为空表示导出所有的边
                "column": [
                    {
                        "name": "id",          // 字段名
                        "type": "string",      // 字段类型
                        "columnType": "primaryKey" // 字段分类，表示边的主键id，GDB中是string类型
                    },
                    {
                        "name": "label",        // 字段名
                        "type": "string",      // 字段类型
                        "columnType": "primaryLabel" // 字段分类，表示边的label名，GDB中是string类型
                    }
                ]
            }
        }
    ]
}

```

```

    },
    {
      "name": "srcId",          // 字段名
      "type": "string",        // 字段类型
      "columnType": "srcPrimarykey" // 字段分类，表示边关联点中起点的id，G
DB中是string类型
    },
    {
      "name": "srcLabel",      // 字段名
      "type": "string",        // 字段类型
      "columnType": "srcPrimaryLabel" // 字段分类，表示边关联点中起点的la
bel名，GDB中是string类型
    },
    {
      "name": "dstId",          // 字段名
      "type": "string",        // 字段类型
      "columnType": "dstPrimarykey" // 字段分类，表示边关联点中终点的
id，GDB中是string类型
    },
    {
      "name": "dstLabel",      // 字段名
      "type": "string",        // 字段类型
      "columnType": "dstPrimaryLabel" // 字段分类，表示边关联点中终点的
label名，GDB中是string类型
    },
    {
      "name": "weight",        // 属性字段名
      "type": "double",        // 属性字段类型
      "columnType": "edgeProperty" // 字段分类，表示边的属性
    }
  ]
}
"stepType": "gdb"
},
{
  "category": "writer",
  "name": "Writer",
  "parameter": {
    "print": false,          //是否向屏幕打印输出。
    "fieldDelimiter": ",",  //列分隔符。
  }
  "stepType": "stream"
}
]
"type": "job",
"version": "2.0"
}

```

② 说明 GDB需要升级到 1.0.20 或更高版本才支持SET属性，使用SET属性时请确认实例版本。

3.3. 使用DataWorks数据集成从MySQL导入数据到GDB

本文为您介绍通过DataWorks的数据集成功能，从RDS MySQL导入数据至图数据库GDB。

前提条件

- 已创建RDS MySQL实例，创建方法请参见[创建RDS MySQL实例](#)。
- 已创建图数据库GDB实例，创建方法请参见[创建主实例](#)。
- 已完成独享资源组配置，配置方法请参见[独享资源组模式](#)。
- 已完成RDS MySQL数据源配置，配置方法请参见[配置MySQL数据源](#)。
- 已完成图数据库GDB数据源配置，配置方法请参见[配置Graph Database数据源](#)。

操作步骤

1. 配置同步任务脚本

- i. 单击目标DataWorks工作空间所在行操作列的进入数据开发。
- ii. 创建离线同步节点，具体请参见[离线同步节点](#)。
- iii. 在选择数据源向导下，执行以下操作。
 - 在数据来源区域，选择数据源类型为MySQL，数据源为已配置完成的MySQL数据源。
 - 在数据去向区域，选择数据源类型为GDB，数据源为已配置完成的GDB数据源。

 **说明** 图数据库GDB当前仅支持通过脚本模式配置数据去向，具体请参见[通过脚本模式配置离线同步任务](#)。

- a. 单击工具栏提示信息中的单击转换为脚本。
- b. 在配置脚本页面，编辑同步任务参数。

以下示例配置是同步数据源 `mysql_loader` 中数据表 `comment` 到GDB数据源 `gdb_loader` 中，将表数据转换成GDB的点，表中部分列映射为点的属性。

```
{
  "type": "job",
  "steps": [
    {
      "stepType": "mysql", // 数据来源类型
      "parameter": {
        "datasource": "mysql_loader", // mysql数据源
        "column": [ // 读取的数据表列名
          "node_id",
          "node_label",
          "m_id",
          "imagefile",
          "creationDate"
        ],
        "connection": [ // 连接信息
          {
            "datasource": "mysql_loader",
            "table": [ // 读取的数据库表
```

```
        "post"
    ]
}
],
"where": "",
"splitPk": "",
"encoding": "UTF-8"
},
"name": "Reader",
"category": "reader"
},
{
    "stepType": "gdb",
    "parameter": {
        "datasource": "gdb_loader",
        "session": true,
        "column": [
            {
                "columnType": "primaryKey",
                "name": "id",
                "type": "string",
                "value": "#{0}"
            },
            {
                "columnType": "vertexProperty",
                "name": "m_id",
                "type": "string",
                "value": "#{2}"
            },
            {
                "columnType": "vertexProperty",
                "name": "imagefile",
                "type": "string",
                "value": "#{3}"
            },
            {
                "columnType": "vertexProperty",
                "name": "creationDate",
                "type": "long",
                "value": "#{4}"
            }
        ],
        "labelType": "VERTEX",
        "maxRecordsInBatch": 32,
        "label": "#{1}",
        "writeMode": "INSERT",
        "idTransRule": "none"
    }
}
```

式

string类型会导致异常。

选VERTEX和EDGE。

新的记录数，在打开批量事务时有效，建议使用默认值。

，可直接提供常量字符串如果本次导入的点名称相同。

T模式时重复写入会报错。

置none表示保持原配置的映射值。

// 数据去向类型

// gdb数据源

// 使用GDB的批量事务模

// 到GDB数据的映射规则

// 主键ID映射

// 主键值类型，配置为非

// 主键映射列序号

// 点属性映射

// 属性名

// 属性值类型

// 属性值映射列序号

// 点属性映射

// 属性名

// 属性值类型

// 属性值映射列序号

// 导入GDB数据类型，可

// 批量事务中一次事务更

// 导入数据点名称label

// 写入数据类型，INSERT

// 主键ID映射规则，配

```
    },
    "name": "Writer",
    "category": "writer"
  }
],
"version": "2.0",
"order": {
  "hops": [
    {
      "from": "Reader",
      "to": "Writer"
    }
  ]
},
"setting": {
  "speed": {
    "record": "3000"
  },
  "route": "zhangbei" //zhangbei为地域名称，您需要将其修改为图数据库GDB实例
                        所在地域名称。
}
```

iv. 单击，并单击.

2. 执行同步任务脚本

- i. 单击脚本窗口右侧导航栏的**基本配置**，在弹出窗口选择**资源组**。
- ii. 单击，开始导入RDS MySQL数据至图数据库GDB。

② 说明

- 如果配置以及网络环境正确，您将看到周期性的同步信息输出，最后同步完成时会输出任务的统计信息。
- 如果您只需要导入少量测试数据到GDB，也可以使用GDB的外网地址创建数据源，使用默认资源组运行导入任务。但由于外网环境复杂，且延时较高，不能保障服务质量，仅限于测试使用。