

图引擎服务 GES 使用——创建图

1. 任务介绍

图引擎服务（Graph Engine Service），是针对以“关系”为基础的“图”结构数据，进行查询、分析的服务。广泛应用于社交关系分析、营销推荐、舆情分析、路径规划、知识图谱、金融风控等具有丰富关系数据的场景。本次任务将介绍如何使用图引擎服务创建图，这里以电影推荐场景为例（如下图）：

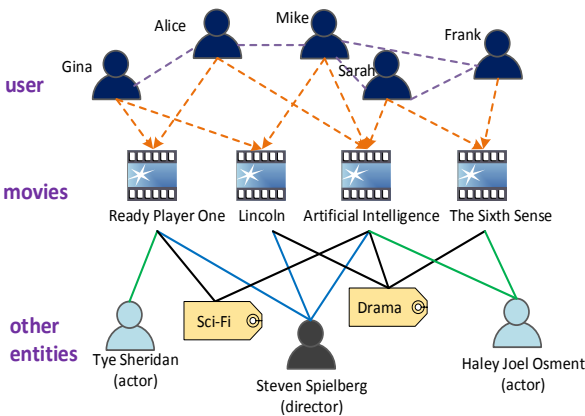


图 1

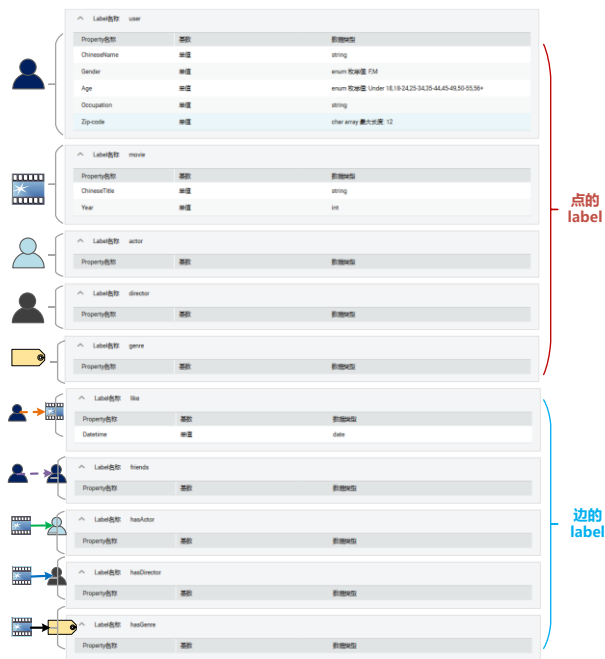
2. 任务执行

如图 1 所示，这张图中包含了点（实体）、边（关系）、以及点和边上面的属性描述，图引擎服务支持的图数据格式包含点文件、边文件以及元数据三部分。点文件和边文件分别存放了点数据和边数据；元数据包含了点和边上的属性描述信息，可以用于描述点文件和边文件中的数据格式。

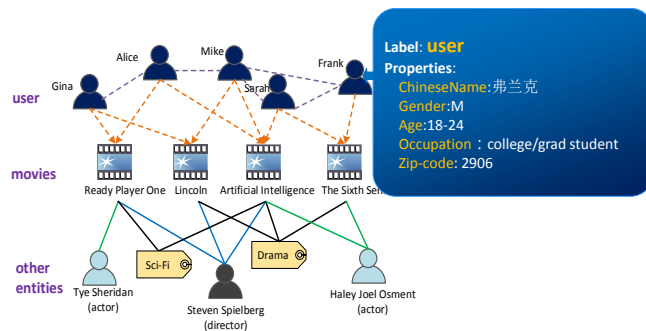
使用图引擎服务时包含四个操作步骤：（1）创建元数据；（2）准备点\边数据文件；（3）创建图；（4）访问图（进行查询和分析）。

2.1 创建元数据

元数据由多个属性（property）集合组成，每个属性集合拥有一个标签（label）名称。例如图 1 所示的图，具有的元数据可以如下所示：



当某个点或边拥有一个 label 时，对应会拥有这个 label 下面的属性，例如：Frank，它的 label 为“user”，那么他也拥有 user 这个 label 下面的属性描述信息：



2.1.1~2.1.4 章节将详细介绍如何创建元数据，您也可以参考《用户指南》中的“元数据管理”章节

2.1.1 登录[华为云官网](#)，选择“EI 企业智能 > 图引擎服务”



2.1.2 选择“进入控制台”



2.1.3 确认左上角的区域选择为“华北-北京一”，然后进行元数据创建：

创建方法：

目前图引擎服务支持两种元数据创建方式：一种是直接创建，一种是导入事先定义好的 xml 文件（本次培训课程采用的元数据 xml 文件可以从第 3 章节的“打卡任务”中下载获得）。

方法一：直接创建方式。点击“元数据管理 > 创建”，添加元数据中的 label 以及 property，点击“确认”完成创建：



注：OBS 创建桶的方法请参考 OBS《控制台指南》。

方法二：导入 xml 文件创建方式。点击“元数据管理 > 导入”，通过导入定义好的 xml 文件，创建元数据，元数据 xml 文件的格式可参考：《用户指南》中的“图数据的格式”->schema 章节。

2.1.4 创建完成后，返回到元数据管理页面，可以点击已经建好的元数据进行查看/编辑，例如下图

图引擎服务

总览

图管理

备份管理

元数据管理

元数据管理 > schema_demo

名称

schema_demo

存储路径

data-ges/demo_schema/schema_demo.xml

定义

点的Label

Label名称

user

Property名称	基数	数据类型
ChineseName	单值	string
Gender	单值	enum 枚举值: F,M
Age	单值	enum 枚举值: Under 18,18-14,25-34,35-44,45-49,50-5...
Occupation	单值	string
Zip-code	单值	char array 最大长度: 12

Label名称

movie

Property名称	基数	数据类型
ChineseTitle	单值	string
year	单值	int

Label名称

actor

Property名称	基数	数据类型
------------	----	------

Label名称

director

Property名称	基数	数据类型
------------	----	------

Label名称

genre

Property名称	基数	数据类型
------------	----	------

边的Label

Label名称

like

Property名称	基数	数据类型
Datetime	单值	date

Label名称

friends

Property名称	基数	数据类型
------------	----	------

Label名称

hasActor

Property名称	基数	数据类型
------------	----	------

Label名称

hasDirector

Property名称	基数	数据类型
------------	----	------

Label名称

hasGenre

Property名称	基数	数据类型
------------	----	------

编辑

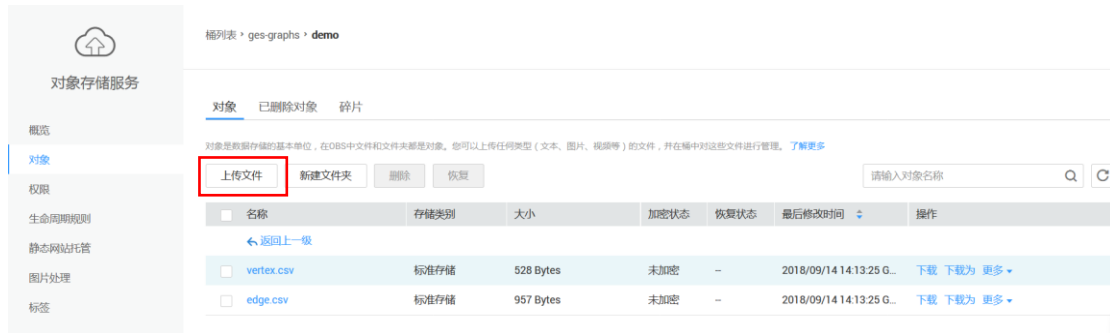
复制

关闭

2.2 准备点边数据

创建图时，用户需要根据定义好的元数据，准备点和边数据集，上传至 [OBS](#)。

注：本次培训课程的涉及到的点边数据集可以从第 3 章节的“打卡任务”中下载获得。



下面对图数据的点边数据集做详细介绍。

2.2.1 点文件

点文件即罗列了各个点的数据信息。每一行为一个点的数据。格式如下：

```
id 1, label, property 1, property 2, ...
```

第一列为点的 id，是点数据的唯一标识；第二列为点所具有的 label；第三列开始为该 label 下的个属性取值，属性的顺序与元数据——对应，示例：

```
Frank,user,弗兰克,M,18-24,college/grad student,2906
```

```
Gina,user,吉娜,F,35-44,sales/marketing,60202
```

```
Ready Player One,movie,头号玩家,2011
```

```
Haley Joel Osment,actor
```

```
Steven Spielberg,director
```

```
Drama,genre
```

```
...
```

2.2.2 边文件

边文件罗列了各个边的数据信息，一行为一条边的数据。格式如下所示，id 1、id 2 表示一条边的两个端点的 id：

```
id 1, id 2, label, property 1, property 2, ...
```

示例：

```
Gina,Lincoln,like,2017 /12/22 16:30
```

```
Sarah,The Sixth Sense,like,2017/12/22 22:00
```

```
Frank,Sarah,friends
```

```
The Sixth Sense,Haley Joel Osment,hasActor
```

```
The Sixth Sense,Drama,hasGenre
```

```
...
```

2.3 创建图

2.3.1 点击左侧“图管理>创建图”



2.3.2 配置参数和规格

- (1) 在创建图页面，选择“区域”和“可用区”。

“区域”：集群工作区域，可在左上方菜单栏选择，目前工作区域仅支持华北-北京一。

“可用区”：可用区是使用独立电源和网络资源的物理区域。通过内部网络互联，再以物理方式进行隔离，提高了应用程序的可用性。建议您在不同的可用区下创建集群。选择集群工作区域下关联的可用区。

- (2) 填写“图名称”。例如，ges_demo。

图名称的长度需满足如下条件。

- 4~64 个字符，且以字母开头。
- 字母不区分大小写。
- 只能包含字母、数字或下划线，不能包含其他特殊字符。

- (3) “GES 软件版本”：选择当前的版本号。

- (4) 选择网络信息，包含“虚拟私有云”、“子网”和“安全组”。

“虚拟私有云”：VPC 即虚拟私有云，是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。选择需要创建集群的 VPC，单击“查看虚拟私有云”，可进入 VPC 服务查看已创建的 VPC 名称和 ID。若没有 VPC，请先创建一个新的 VPC。

“子网”：通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。子网在可用区内才会生效。

选择需要创建集群的子网，可进入 VPC 服务查看 VPC 下已创建的子网名称和 ID。

“安全组”：安全组是一个逻辑上的分组，为同一个 VPC 内具有相同安全保护需求并相互信任的弹性云服务器提供访问策略。单击“查看安全组”可了解安全组详情。4~64 个字符，且以字母开头。字母不区分大小写。只能包含字母、数字或下划线，不能包含其他特殊字符。

- (5) 选择公网访问的方式，包含“暂不使用”、“现在购买”和“使用已有”，用户根据自身需求选择。

- (6) 填写图相关参数。

选择 Schema (元数据)，边数据集、点数据集等数据只支持从 OBS 中选择，建议您提前将数据文件

存储至“[OBS\(对象存储服务\)](#)”中，导入 OBS 的操作请参见 OBS《[控制台指南](#)》。

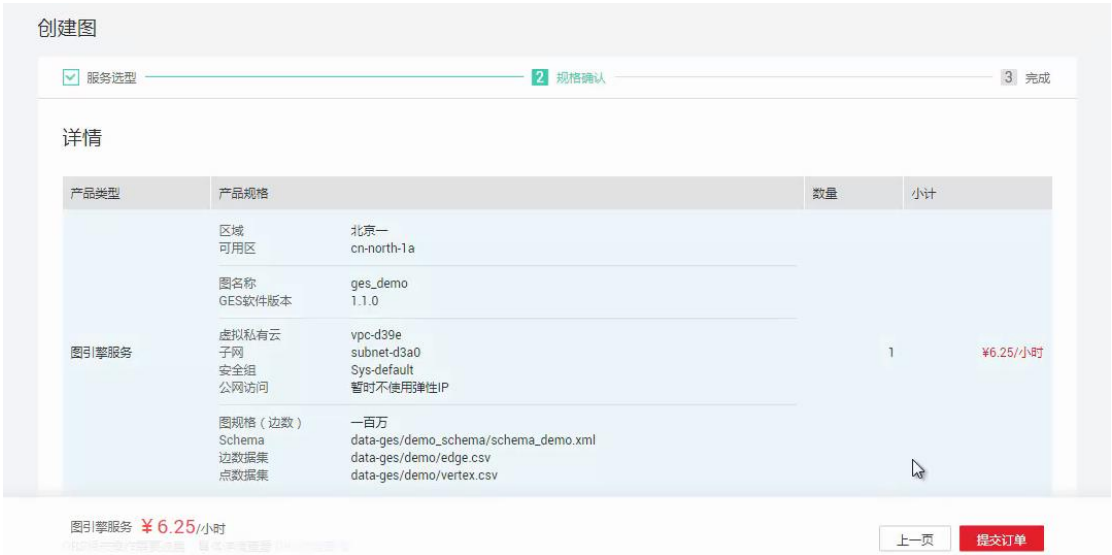
图数据（包含 Schema、边数据集、点数据集）需按照模板中的格式导入（模板中已包含一份电影数据的数据，可下载后直接导入）

- o 用途：“企业生产”提供更高的规格、更好的可靠性和更快的性能；“开发学习”支持小规模完整功能体验，主要用作学习和开发调试，不适合于生产部署。
- o 图规格（边数）：以边为单位，可选择一万、一百万、一千万、一亿、十亿和千亿。根据用户当前的配额，系统会提示可创建的图的数量以及边数。
- o Schema：选择图的元数据。如果没有元数据，可单击“创建新的 Schema”创建，创建 Schema 的操作指导请参考[元数据管理](#)。
- o 边数据集：所有描述边的结构、标签和属性的信息构成了边数据集。
- o 点数据集：所有描述点的 ID、标签和属性的信息构成了点数据集。

说明：

点数据集和边数据集当前仅支持英文路径和文件夹格式。

2.3.3 单击“立即购买”，确认信息后，单击“提交订单”开始创建图。



2.3.4 系统将自动跳转到“图管理”页面，您创建的图将展现在图列表中。

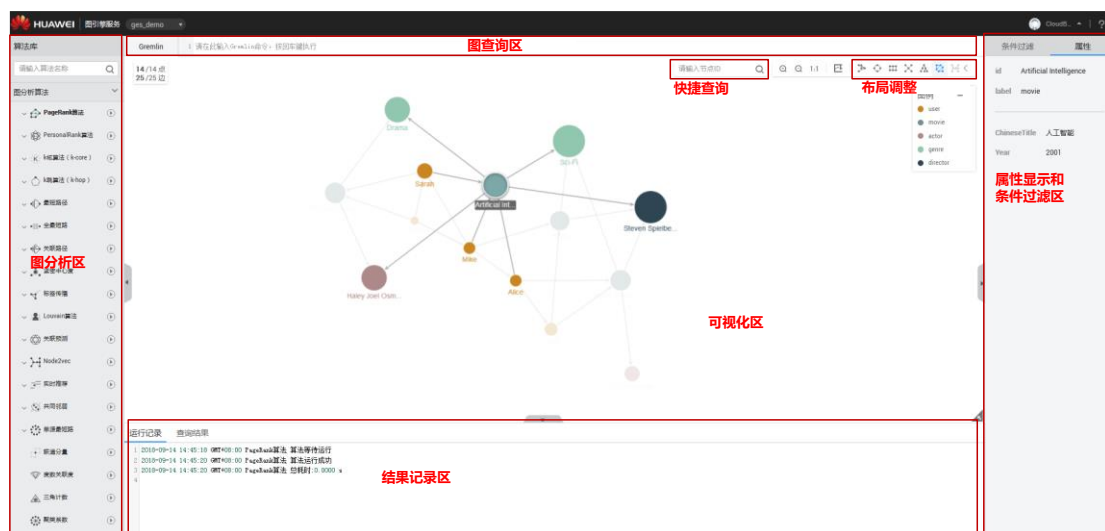


2.4 访问图

2.4.1 图运行状态变为运行中时，点击“访问”，访问对应的图界面。



图界面各功能区如图所示：



图查询区：支持标准图查询语言 Gremlin，兼容您的使用习惯

图分析区：提供丰富的图分析算法，简单易用，满足社交关系分析、营销推荐、舆情分析、路径规划、知识图谱、金融风控等具有丰富关系数据的各类应用场景分析需求

可视化区：支持，所见即所得可视化展示

结果记录区：操作记录可见，支持 JSON 格式结果导出，轻松获取

3. 打卡任务

熟悉 GES 各部分功能，在 GES 动手实践：

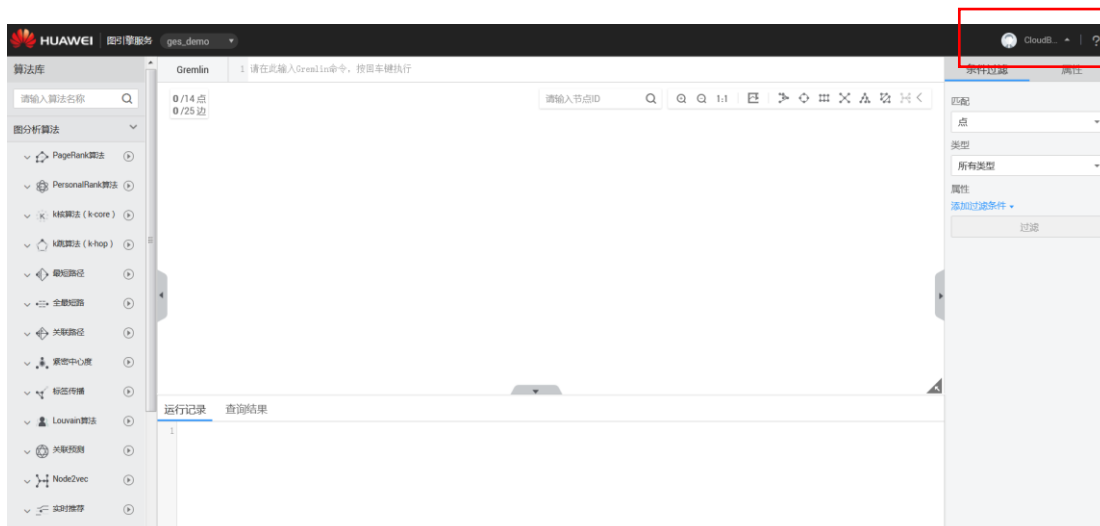
创建图 → 访问图 → 熟悉图界面各功能区（如 2.4.1 章节）→ 删除图。

注：创建图时，图规格参数中的“用途”选项请选择为“开发学习”；

样例数据：电影推荐场景的样例数据：

<https://ges-training-data.obs.myhwclouds.com/huawei-ges-training-data-v1.0.zip>

截图反馈：反馈访问图之后的图界面，例如下：



注意：完成打卡后，请及时到控制台删除闲置的图，以免造成欠费。

删除图方法：[“图管理”](#) → 点击“更多” → 删除



附录：GES 服务用户指南

https://support.huaweicloud.com/usermanual-ges/ges_01_0001.html#