

图引擎服务

常见问题

文档版本 01
发布日期 2024-04-17



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编： 518129

网址： <https://www.huawei.com>

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

安全声明

漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

目录

- 1 客户咨询类..... 1
 - 1.1 关于图引擎服务..... 1
 - 1.2 图引擎服务与图像处理有什么区别? 1
 - 1.3 服务支持使用哪些算法对图进行分析? 1
- 2 API 使用类..... 5
 - 2.1 如何导入数据到图引擎服务? 5
 - 2.2 创建实例或实例已创建, 需绑定 EIP 时, 提示创建委托权限不足怎么处理? 5
 - 2.3 创建实例或实例已创建, 需绑定 EIP 时, 页面提示“委托配额不足”如何处理? 6
 - 2.4 Gremlin/Cypher 查询是否支持几条命令一起执行? 7
 - 2.5 如果点被删除了, 基于该点的边会怎么处理? 7
 - 2.6 点 ID 存在中文时, 修改点属性接口报错“点不存在”怎么处理? 7
- 3 界面问题..... 8
 - 3.1 Gremlin 命令怎么执行和查看运行结果? 8
 - 3.2 为什么有些算法单击右边“运行”按钮没有反应? 8
- 4 其他问题..... 9
 - 4.1 什么是区域与可用区? 9

1 客户咨询类

1.1 关于图引擎服务

图引擎服务（Graph Engine Service，简称GES），是针对以“关系”为基础的“图”结构数据，进行查询、分析的服务。提供分布式、大规模、一体化的图查询和分析能力，内置丰富的图分析算法库，满足多领域应用需求。

- 支持多大规模的图数据
图引擎服务最大支持百亿节点千亿边的超大规模图。
- 性能
通过高性能的内核，满足高并发秒级多跳实时查询。拥有三大核心专利，支持多跳查询秒级响应。
- 应用场景
可以应用于社交网络、精准营销、信贷保险、网络/路径的设计等场景。
- 与Eywa的关系
Eywa是自研的图引擎内核，图引擎服务是基于Eywa的云服务。

1.2 图引擎服务与图像处理有什么区别？

- 图引擎服务
主要用于关系分析，把关系网络抽象成一张形象的图结构数据（例如：假设人是一个一个的点，而人与人之间的关系就是连接各个点的边，图引擎的作用就是分析这张图，挖掘出人与人之间潜在的关系），可以应用于社交网络、精准营销、信贷保险、网络/路径的设计等场景。
- 图像处理
主要用于对图片进行处理。例如，图像识别就是图像处理的功能之一。

1.3 服务支持使用哪些算法对图进行分析？

为满足用户各种场景需求，图引擎服务提供了丰富的基础图算法、图挖掘算法和图指标算法。

支持在查询编辑器页面使用算法对图进行分析，当前支持的算法列表如下图所示。

表 1-1 算法一览表

算法	介绍
PageRank算法	又称网页排名，是一种由搜索引擎根据网页（节点）之间相互的超链接计算的技术，用来体现网页（节点）的相关性和重要性。
PersonalRank算法	PersonalRank算法又称Personalized PageRank算法。该算法继承了经典PageRank算法的思想，利用图链接结构来递归地计算各节点的重要性。与PageRank算法不同的是，为了保证随机行走中各节点的访问概率能够反映出用户的偏好，PersonalRank算法在随机行走中的每次跳转会以（1-alpha）的概率返回到source节点，因此可以基于source节点个性化地计算网络节点的相关性和重要性（PersonalRank值越高，对source节点的相关性/重要性越高）。
k核算法（k-core）	k-core是图算法中的一个经典算法，用以计算每个节点的核数。其计算结果是判断节点重要性最常用的参考值之一，较好的刻画了节点的传播能力。
k跳算法（k-hop）	从起点出发，通过宽度优先搜索（BFS），找出k层与之关联的所有节点。找到的子图称为起点的ego-net。k跳算法会返回ego-net中节点的个数。
最短路径（Shortest Path）	用于解决图论研究中的一个经典算法问题，旨在寻找图中两节点之间的最短路径。
全最短路（All Shortest Paths）	用于解决图论研究中的一个经典算法问题，旨在寻找图中两节点之间的所有最短路径。
带一般过滤条件最短路径算法（Filtered Shortest Path）	寻找两点间满足过滤条件的最短路径，如有多条，返回任意一条最短路径。
单源最短路（SSSP）	图论中的经典问题，给定一个节点（称为源），该算法给出从该源节点出发到其余各节点的最短路径长度。
点集最短路（Shortest Path of Vertex Sets）	用于发现两个点集之间的最短路径。适用于互联网社交、金融风控、路网交通、物流配送等场景下的区块之间关系分析。
关联路径（n-Paths）	该算法用于寻找图中两节点之间在k层关系内的n条路径。适用于关系分析、路径设计、网络规划等场景。
紧密中心度（Closeness Centrality）	紧密中心度是一个节点到所有其他可达节点的最短距离的平均，该指标可以用来衡量信息从该节点传输到其他节点的时间长短。节点的“Closeness Centrality”越小，其所在图中的位置越中心。

算法	介绍
标签传播 (Label Propagation)	一种基于图的半监督学习方法，其基本思路是用已标记节点的标签信息去预测未标记节点的标签信息。利用样本间的关系建图，节点包括已标注和未标注数据，其边表示两个节点的相似度，节点的标签按相似度传递给其他节点。标签数据就像是一个源头，可以对无标签数据进行标注，节点的相似度越大，标签越容易传播。
Louvain算法	基于模块度的社区发现算法，该算法在效率和效果上都表现较好，并且能够发现层次性的社区结构，其优化目标是最大化整个社区网络的模块度。
关联预测 (Link Prediction)	给定两个节点，根据Jaccard度量方法计算两个节点的相似程度，预测节点之间的紧密关系。
Node2vec算法	通过调用word2vec算法，把网络中的节点映射到欧式空间，用向量表示节点的特征。Node2vec通过回退参数P和前进参数Q来生成从每个节点出发的随机步，它带有BFS和DFS的混合，回退概率正比于 $1/P$ ，前进概率正比于 $1/Q$ ，每个节点出发生成多个随机步，反映出网络的结构信息。
实时推荐 (Real-time Recommendation)	一种基于随机游走模型的实时推荐算法，能够推荐与输入节点相近程度高、关系或喜好相近的节点。该算法可以基于历史购买或浏览数据进行相近商品推荐，也可以针对人进行相近喜好的潜在好友推荐。
共同邻居 (Common Neighbors)	是一种常用的基本图分析算法，可以得到两个节点所共有的邻居节点，直观地发现社交场合中的共同好友、消费领域共同感兴趣的商品，进一步推测两个节点之间的潜在关系和相近程度。
连通分量 (Connected Component)	连通分量代表图中的一个子图，当中所有节点都相互连接。考虑路径方向的为强连通分量 (strongly connected component)，不考虑路径方向的为弱连通分量 (weakly connected component)。 说明 本算法计算得到的是弱连通分量。
度数关联度 (Degree Correlation)	度数关联度算法计算所有边上起点和终点度数之间的Pearson关联系数，常用来表征图中高度数节点是否和高度数节点相连。
三角计数 (Triangle Count)	不考虑边的方向，统计图中三角形个数。三角形越多，代表图中节点关联程度越高，组织关系越严密。
聚类系数 (Cluster Coefficient)	聚类系数是表示一个图中节点聚集程度的系数，证据显示，在现实的网络中，尤其是在特定的网络中，由于相对高密度连接点的关系，节点总是趋向于建立一组严密的组织关系。

算法	介绍
点集共同邻居 (Common Neighbors of Vertex Sets)	可以得到两个点集合（群体集合）所共有的邻居（即两个群体临域的交集），直观的发现与两个群体共同联系的对象，如发现社交场合中的共同好友、消费领域共同感兴趣的商品、社区群体共同接触过的人，进一步推测两点集合之间的潜在关系和联系程度。
点集全最短路 (All Shortest Paths of Vertex Sets)	点集最短路算法用于发现两个点集之间的所有最短路径，可应用于互联网社交、金融风控、路网交通、物流配送等场景下的区块之间关系的分析。
带一般过滤条件环路检测 (Filtered Circle Detection)	目的是寻找图中所有满足过滤条件的环路。适用于金融风控中循环转账检测、反洗钱，网络路由中异常链接检测，企业担保圈贷款风险识别等场景。
子图匹配 (Subgraph Matching)	子图匹配 (subgraph matching) 算法的目的是在一个给定的大图里面找到与一个给定小图同构的子图，这是一种基本的图查询操作，意在发掘图重要的子结构。
带过滤全对最短路径 (Filtered All Pairs Shortest Paths)	带过滤全对最短路径 (Filtered All Pairs Shortest Paths) 是寻找图中任意两点之间满足条件的最短路径。当前，考虑到实际应用场景，此算法需要用户指定起点集 (sources) 和终点集 (targets)，本算法将返回起点集合到终点集合之间满足条件的两两全最短路径。
带过滤全最短路径 (Filtered All Shortest Paths)	带过滤全最短路径 (Filtered All Shortest Paths) 是在最短路径算法 (Shortest Path) 基础上支持条件过滤，寻找图中两节点之间满足条件的全最短路径。
TopicRank算法	TopicRank算法12345热线多维度话题排序算法之一，适用于政务12345热线投诉话题排序。
带过滤的n_paths算法 (2.2.22)	带过滤的n_paths算法是给定起始点source、目的点target、跳数k、路径数n、过滤条件filters，找出source和target间不多于n条的k跳无环路径。

2 API 使用类

2.1 如何导入数据到图引擎服务？

图引擎服务在创建图完成后，您可以导入图数据，或者当您需要新增图数据时，可以利用“导入”功能进行增量导入。

具体操作请参考[导入数据](#)。

说明

- 当前仅支持 1.1.8 以上版本的图的增量导入功能。
- 为防止系统重启时，不能正常恢复导入图数据，建议在使用图期间，不要删除存储在OBS中的数据。
- 数据列的分隔符默认为逗号，暂不支持自定义。

2.2 创建实例或实例已创建，需绑定 EIP 时，提示创建委托权限不足怎么处理？

给用户所属的用户组赋予安全管理员角色。

步骤1 单击“服务列表”，选择“统一身份认证服务 IAM”。

步骤2 在“用户”页面，单击用户列表中该用户名前面的 ，查看当前用户所属的用户组名称。

步骤3 在“用户组”页面，单击该用户组所在行右侧的“操作>权限配置”，进入“用户组权限”页面。

步骤4 在“用户组权限”列表中，单击“全局服务”所在行右侧的“操作>设置策略”，进入“设置策略”页面。

步骤5 在左侧的“可选策略”列表中，在“Security Administrator”前打勾，单击“确定”，完成操作。

----结束

2.3 创建实例或实例已创建，需绑定 EIP 时，页面提示“委托配额不足”如何处理？

用户在第一次使用GES服务的时候需要授权，授权过程会在IAM（统一身份认证）页面创建例如名为“ges_admin_trust”的委托。而一个用户最多只能创建10个委托，委托个数超限之后无法创建委托，所以页面会出现“委托配额不足”的提示。

按照以下步骤删除不用的委托，然后重新授权即可。

步骤1 登录华为云官网或图引擎服务产品页，在右上角用户名下拉框中，单击进入“账号中心”页面。

图 2-1 官网界面



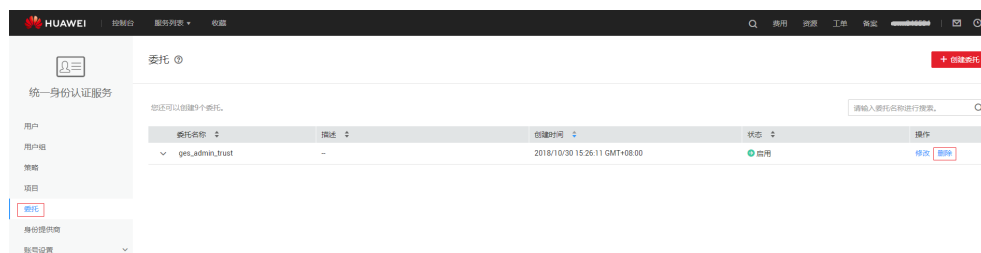
步骤2 在“账号中心”页面右上角用户名下拉框中，单击进入“统一身份认证”页面。

图 2-2 账号中心界面



步骤3 在“统一身份认证”服务页面左侧页签栏中选择“委托”选项，删除页面中不使用的委托，保证GES有创建的委托配额。

图 2-3 统一认证服务界面



步骤4 删除委托之后，重新授权进入GES页面创建实例即可。

----结束

2.4 Gremlin/Cypher 查询是否支持几条命令一起执行？

- Gremlin支持多条命令一起执行，命令之间用分号隔开。

例如：

```
graph = EywaGraph.open('ges_6715');g = graph.traversal();g.V().limit(1)
```

- Cypher暂时不支持多条命令一起执行。

Cypher查询的语法请参考[Cypher查询](#)。Gremlin查询的语法请参考[Gremlin查询](#)。

2.5 如果点被删除了，基于该点的边会怎么处理？

GES基于属性图（Property graph）模型导入图数据，一个属性图是由点、边、标签（Label）和属性（Property）组成的有向图。

点又称作节点（Node），边又称作关系（Relationship），点和关系是最重要的实体。

- 图数据模型中的点代表实体，如交通网络中的车辆、通信网络中的站点、电商交易网络中的用户和商品、互联网中的网页等。
- 图数据模型中的边代表关系，如社交网络中的好友关系、电商交易网络中用户评分和购买行为、论文中作者之间的合作关系、文章之间的索引关系等。

如果点被删除了，基于该点的边会自动删除。

2.6 点 ID 存在中文时，修改点属性接口报错“点不存在”怎么处理？

点ID有中文的情况下，修改点属性接口报错“点不存在”时，您需要进行如下两步操作：

1. 在headers中添加utf-8支持：“Content-Type”:”application/json;charset=UTF-8”。
2. 在url中使用url编码格式来编码中文。

3 界面问题

3.1 Gremlin 命令怎么执行和查看运行结果？

- 执行

在图编辑器页面，您可以在此页面对当前图进行查询分析，在页面下方的Gremlin输入框中，输入一行Gremlin命令后，按“回车”键执行。

- 查看结果

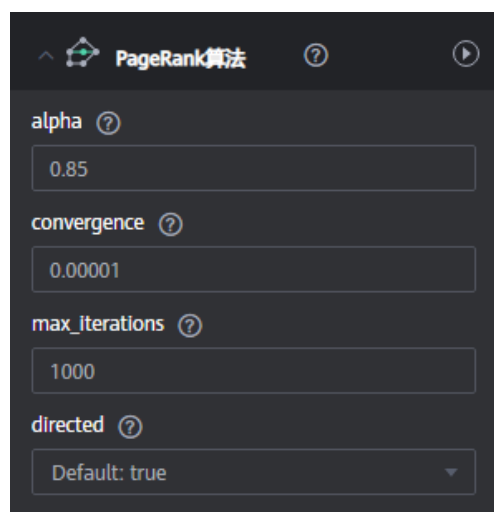
执行Gremlin命令后在“运行记录”可以看到命令运行情况，在“查询结果”可以看到命令执行结果，如果是查点和边的命令在中间画布可以看到点和边的图形化展示。

3.2 为什么有些算法单击右边“运行”按钮没有反应？

这种算法需要设置参数才能运行。

您需要单击该算法左侧  图标，输入正确的参数后，单击  执行算法。

图 3-1 设置算法参数



4 其他问题

4.1 什么是区域与可用区？

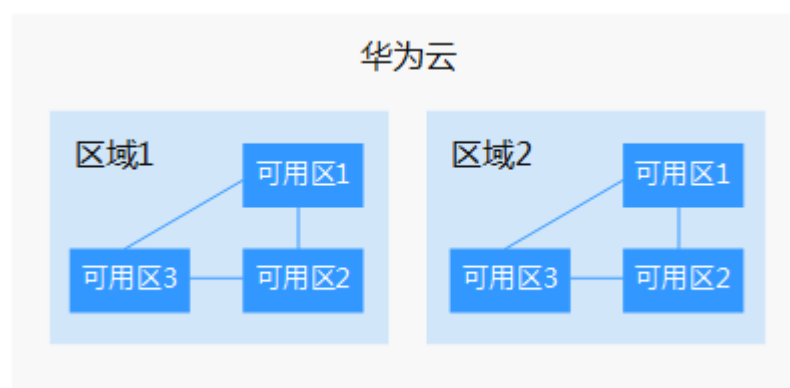
什么是区域、可用区

我们用区域和可用区来描述数据中心的位置，您可以在特定的区域、可用区创建资源。

- 区域（Region）：从地理位置和网络时延维度划分，同一个Region内共享弹性计算、块存储、对象存储、VPC网络、弹性公网IP、镜像等公共服务。Region分为通用Region和专属Region，通用Region指面向公共租户提供通用云服务的Region；专属Region指只承载同一类业务或只面向特定租户提供业务服务的专用Region。
- 可用区（AZ，Availability Zone）：一个AZ是一个或多个物理数据中心的集合，有独立的风火水电，AZ内逻辑上再将计算、网络、存储等资源划分成多个集群。一个Region中的多个AZ间通过高速光纤相连，以满足用户跨AZ构建高可用性系统的需求。

图4-1阐明了区域和可用区之间的关系。

图 4-1 区域和可用区



目前，华为云已在全球多个地域开放云服务，您可以根据需求选择适合自己的区域和可用区。

如何选择区域？

选择区域时，您需要考虑以下几个因素：

- 地理位置

一般情况下，建议就近选择靠近您或者您的目标用户的区域，这样可以减少网络时延，提高访问速度。不过，在基础设施、BGP网络品质、资源的操作与配置等方面，中国大陆各个区域间区别不大，如果您或者您的目标用户在中国大陆，可以不用考虑不同区域造成的网络时延问题。

- 在除中国大陆以外的亚太地区有业务的用户，可以选择“中国-香港”、“亚太-曼谷”或“亚太-新加坡”区域。
- 在非洲地区有业务的用户，可以选择“南非-约翰内斯堡”区域。
- 在欧洲地区有业务的用户，可以选择“欧洲-巴黎”区域。

- 云服务之间的关系

如果多个云服务一起搭配使用，需要注意：

- 不同区域的弹性云服务器、关系型数据库、对象存储服务内网不互通。
- 不同区域的弹性云服务器不支持跨区域部署在同一负载均衡器下。

- 资源的价格

不同区域的资源价格可能有差异，请参见[华为云服务价格详情](#)。

如何选择可用区

是否将资源放在同一可用区内，主要取决于您对容灾能力和网络时延的要求。

- 如果您的应用需要较高的容灾能力，建议您将资源部署在同一区域的不同可用区内。
- 如果您的应用要求实例之间的网络延时较低，则建议您将资源创建在同一可用区内。

区域和终端节点

当您通过API使用资源时，您必须指定其区域终端节点。有关区域和终端节点的更多信息，请参阅图引擎服务的[地区和终端节点](#)。