引文格式：刘荣虎,邵隆,樊文有. 云环境下空间数据异步分析架构设计与应用[J]. 地理空间信息, 2025, 23(4):63-66. doi:10. 3969/j.issn. 1672-4623.2025.04.013

云环境下空间数据异步分析架构设计与应用

刘荣虎[1]，邵　隆[1]，樊文有[2]
（1. 北京天耀宏图科技有限公司，湖北 武汉 430076；2. 中国地质大学（武汉）地理与信息工程学院，湖北 武汉 430074）

摘　要：针对云环境下空间数据分析耗时过长导致HTTP通信超时的问题，采用消息中间件异步通信技术、地图服务自动发布技术以及地图可视化技术，设计了云环境下空间数据异步分析架构，实现了空间数据异步分析、系统即时响应的能力。该架构已在国家重点研发计划课题系统研发中得到应用，为云环境下顾及用户交互体验，开展大规模空间数据分析、场景复现、结果查看提供了一种可行方案。该架构同样也适用于GIS应用系统资源不足、性能低下影响用户交互体验的场景，具有一定的实用价值。

关键词：云环境；空间数据；空间分析；异步通信
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2025）04-0063-04

Design and Application of Spatial Data Asynchronous Analysis Architecture in Cloud Environment

LIU Ronghu[1], SHAO Long[1], FAN Wenyou[2]
(1. Beijing Tianyao Hongtu Technology Co., Ltd., Wuhan 430076, China;
2. School of Geography and Information Engineering, China University of Geosciences (Wuhan), Wuhan 430074, China)

Abstract: In order to solve the problem of HTTP communication timeout caused by spatial data analysis too long time consuming in cloud environment, we used the messaging middleware asynchronous communication technology, map service automatic publishing technology and map visualization technology to design spatial data asynchronous analysis architecture in cloud environment, and realized the capability of spatial data asynchronous analysis and instantaneous response of system. The architecture has been applied in the research and development of national key research and development projects, offering a feasible solution for considering user interaction experience in cloud environments, conducting large-scale spatial data analysis, scenario reproduction, and result viewing. This architecture is also applicable to the scenarios where insufficient GIS application system resource and poor performance affecting user interaction experience, which has certain practical value.

Key words: cloud environment, spatial data, spatial analysis, asynchronous communication

随着测绘、遥感技术的发展进步，空间数据种类不断增多，数据类型不断丰富，数据量呈几何级增长[1-2]。空间数据分析是挖掘这些数据价值的重要途径，因此大规模空间数据分析成为了研究热点。

传统的大规模空间数据分析处理的方法主要有多进程多线程计算[3-4]、多核众核计算[5-6]等，其核心思想是通过提高单机硬件的利用率来提升计算效率。然而随着芯片生产工艺遇到瓶颈，依靠单机开展大规模空间数据分析已不能满足需求[7]。

近年来云计算、大数据技术兴起，大规模空间数据分析处理方法逐渐向更具优势的分布式、并行化方向转变。随着SpatialHadoop、GeoSpark、GS-Hadoop等地理大数据框架的出现，众多学者将其用于实践。Karthi[8]、Li等[9]在云环境下利用Spatial Hadoop机制存储大规模空间数据，优化时空查询分析算法，显著提高了查询分析的性能。彭旭[10]使用GeoSpark结合SparkRDD和JTS工具，实现了空间相交分析近10倍的性能提升。Abdul等[11]定义了扩展形状文件格式（.shpx），利用GS-Hadoop框架，通过MapReduce内存映射访问形状文件组件文件，使得空间数据处理取得了8.3倍的加速。

总体来说，上述方法在实践中均具有较好的表现，在性能上均有提升，但在与云计算技术整合应用场景下，由于采用HTTP协议进行通信，在数10 s未能完成分析任务，易造成HTTP连接超时，影响用户体验[12]。

项目来源：基于全域时空信息的智能城市混合现实及推演技术（2019YFB2103203）。
第一作者简介：刘荣虎（1986—），硕士，工程师，从事大数据GIS、智慧城市软件研发工作，E-mail：liuronghu@greatmap.com.cn。

为此，本研究立足于云环境下大规模空间数据分析处理的需要，在不改变原有空间分析技术和算法的基础上，从避免HTTP通信超时、提升用户交互体验的角度开展研究实践，设计了基于消息中间件通信的空间数据异步分析框架，可作为分布式、并行化空间分析技术架构的补充。

1 关键技术研究

1.1 基于消息中间件的异步通信技术

消息中间件具有异步处理、异构集成、业务解耦的用途，是微服务和分布式系统中重要的通信组件[13-14]。基于消息中间件的异步架构，能够有效解决系统间耦合的问题，保证系统间通信的可靠性和系统反馈的实时性，给解决云环境下HTTP通信超时问题提供了思路。

目前，各商业软件公司和开源组织也纷纷推出了各自的消息中间件产品，其中阿里巴巴公司推出的RocketMQ具有高性能、高可靠、高实时等特点[15]。实验表明，综合吞吐量、延迟、消息处理能力、系统可靠性、国产化等因素，RocketMQ为最佳选择[16]。

借助RocketMQ异步通信的能力，可将分析任务作为消息进行传输，空间分析程序以异步的方式运行。这样业务程序可在不干扰其他操作的情况下完成耗时的空间分析过程，避免HTTP通信超时的发生。

在RocketMQ中主要有4种角色：①NameServer：注册中心，用于Broker的动态注册与发现；②BrokerServer：主要负责消息的存储、投递和查询；③Producer：负责消息发布，选择相应的Broker集群队列进行消息投递；④Consumer：消息消费的角色[12]。消息传递主要有2种模式，分别是点对点模式（Point to Point，Queue）和发布/订阅模式（Publish/Subscribe，Topic）[17]。本研究选择点对点模式，其大致执行流程为：①Producer向Broker发送消息，同时指定消息发送成功及发送异常的回调方法；②调用API后立即返回，Producer发送消息线程不阻塞；③消息发送成功或失败，回调任务执行回调方法。在该模式下，一条消息可以被多个Consumer随机接收，但最终只能被一个Consumer获取[13]。因此系统支持多个业务程序发送消息和多个分析程序接收消息，能够保证多个消息及时发送、及时处理，从而避免消息堆积影响系统整体性能。

1.2 基于地图可视化的场景复现技术

在云环境下，需要通过浏览器与云端进行交互，空间数据的浏览、分析场景的复现、分析结果的查看，均需要在浏览器端进行。

OpenLayers是一个用于WebGIS客户端开发的JavaScript类库包，它实现了标准化的地理空间数据访问方法，支持由OpenGIS协会制定的标准OGC网络服务规范，可以通过远程调用服务的方式在浏览器端加载地图数据，进行显示和交互操作[18]。GeoServer是OpenGIS Web服务器规范的J2EE实现，全面遵循OGC开放标准的网络地图服务和网络要素服务规范[19]。同时提供了REST风格的服务接口，支持用户通过调用接口的方式实现地图服务的发布与管理[20]。

本研究使用GeoServer提供空间数据可视化相关的服务，借助其REST接口，自动化地将分析结果发布成地图服务；在浏览器端使用OpenLayers，用于空间数据可视化及地图交互，满足空间分析原始数据与分析结果查看的需求。

2 空间数据异步分析架构设计

2.1 总体框架设计

根据云环境下HTTP通信协议的特点，兼顾系统的扩展性，总体框架采用SOA（service-oriented architecture）的思想进行设计。SOA是一种基于组件模型的面向服务的体系结构，它能基于HTTP协议，通过接口和契约的方式，将不同的功能服务单元联系起来[21]，总体框架如图1所示。

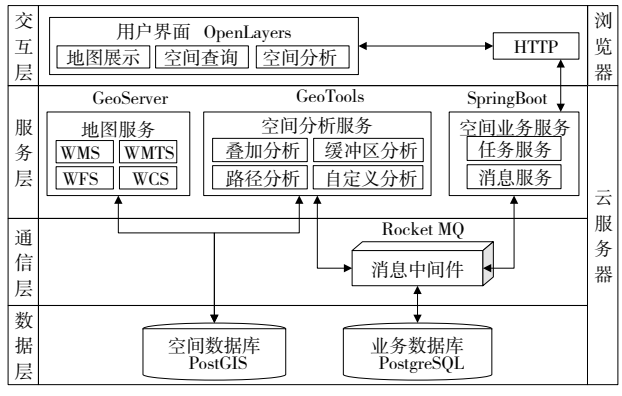


图1 总体框架图

1）交互层：即浏览器，为用户提供地图可视化界面和操作界面；与用户交互，使用HTTP协议向服务层发送请求并接收响应信息。

2）服务层：采用微服务架构，提供业务服务、地图服务、空间分析服务三类服务。业务服务用于任务与消息的管理；地图服务是由空间数据及空间分析结果发布成的OGC标准服务组成；空间分析服务用于接收并执行空间分析任务、调用数据层存储分析结果、

向通信层发送分析结果消息。

3）通信层：采用分布式消息中间件集群，用于空间分析任务消息的传输；同时起到控制流量、降低系统耦合性、提高系统响应速度等作用。

4）数据层：采用分布式数据库集群，用于存储业务数据和空间数据、其主要职责是为空间分析服务提供数据支持，为分析结果提供存储空间。

2.2　消息结构设计

空间分析任务消息定义了业务与空间分析服务之间传递的信息模型，消息结构如图2所示，括号中为示意值，各字段说明如下：

1）采用UUID作为任务的唯一标识。

2）采用时间戳记录任务创建时间和完成时间。

3）任务内容由分析类型和数据列表组成。分析类型主要有叠置分析、缓冲分析以及自定义的分析；数据列表采用数组类型，存储空间分析使用的数据ID。

4）任务指纹：采用Hash算法，将数据列表和分析类型作为参数，计算出任务的Hash值。当有新的分析任务时，只需判断2次任务指纹是否一致，如果一致可直接从数据库中取回分析结果。

5）任务状态：设计了排队中、分析中、成功、失败4种状态。

6）任务结果：空间分析结果的唯一标识码，通过此标识可从空间数据库中获取分析结果数据。

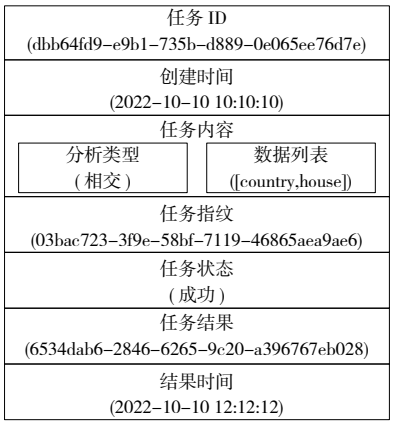


图2　空间分析任务消息结构

2.3　流程设计

总体流程大致分为3个环节，分别是：空间分析消息构建与提交、空间分析消息接收与执行、空间分析结果反馈与查看。

1）空间分析消息构建与提交。客户端采用HTTP协议发送空间分析请求，业务模块根据分析请求各参数构建分析任务；将分析任务提交至消息中间件后，立即结束请求，释放资源，避免操作阻塞。

2）空间分析消息接收与执行。空间分析服务器从消息中间件接收空间分析任务，对任务进行解析，从空间数据服务器中调取相应的空间数据进行空间分析；待分析完成后，将分析结果写入空间数据库，同时自动发布地图服务，修改任务状态信息并提交至消息中间件。

3）空间分析结果反馈与查看。根据反馈信息，业务系统解析任务信息，如果消息状态为“成功”，则可以场景复现的方式查看空间分析结果。客户端调用数据列表信息实现场景的复现，调用分析结果对应的地图服务实现分析结果的查看。

3　应　用

为验证本文所提出的云环境下异步空间分析架构的可行性，以国家重点研发计划课题《基于全域时空信息的智能城市混合现实及推演技术：多规合一空间规划优化分析系统》为应用案例进行试验验证。

3.1　应用背景

依据课题要求，研究面向智能城市的空间规划优化分析技术，实现城市规划和城市建设以及空间治理体系的辅助决策。为了满足空间规划优化分析的需求，系统设计了多个空间分析场景。由于部署在云环境下，且空间数据量庞大、分析计算复杂，易造成HTTP通信超时。

3.2　应用环境与架构

系统部署在雄安云环境中，由雄安云提供存储资源和计算资源。依据空间数据异步分析架构，搭建了Web服务器集群、地图服务器、数据库服务器、空间分析服务器、消息中间件服务器。其中Web服务器集群作为客户端交互入口，接收并响应HTTP请求。

3.3　应用效果

系统以分屏浏览的方式对空间分析原始数据进行展示。在进行空间分析操作时，系统构建分析任务，并将任务作为消息提交到消息中间件，此操作几乎瞬间完成，用户无需处于等待状态；分析任务提交后，用户可随时通过任务中心查看已提交的任务信息；分析成功之后，用户可即时接收到消息提示。系统以任务中心功能模块为入口，通过解析任务信息，以分屏查看的方式展示空间分析原始数据和分析结果。

为测试异步分析架构的效果，利用多边形图层的求交分析进行实验，实验结果如表1所示。实验结果对比分析如下：

1）在数据量较小的情况下，同步模式和异步模

式应用效果相差不大，用户体验均良好。

2）在中等数据规模的情况下，同步模式需要等待数10 s，此时已严重影响用户体验。

3）在较大数据规模的情况下，由于HTTP超时，同步模式此时已无法返回结果，系统不可用。

4）异步模式在任何数据量情况下，均可做到即时响应，无需等待，用户体验良好。

5）在各数据量情况下，异步模式与同步模式总体耗时接近。由此可见，异步通信架构不会对系统整体性能造成影响。

表1 实验结果对比

分析类型	数据量	同步模式总耗时	异步模式总耗时	同步模式等待	异步模式等待/ms
相交分析	80条土地利用50条不规则面	6 s	7 s	6 s	1
相交分析	80条土地利用50条不规则面	73 s	75 s	73 s	1
相交分析	80条土地利用50条不规则面	12 min	12 min 21s	超时	1
相交分析	80条土地利用50条不规则面	6 h 37 min	6 h 42 min	超时	1

4 结 语

云GIS时代已来临，在云环境下大规模空间数据分析处理等耗时场景中，HTTP请求超时已成为常态化问题。本文设计的空间数据异步分析架构能够做到在云环境下即时响应用户请求，有效避免了HTTP超时问题的发生。目前，本文的研究成果已经成功应用于“雄安多规合一空间规划优化分析系统”中。应用效果表明，该架构能够在大规模空间数据分析场景下，具备即时响应、场景复现、结果查看等能力。该架构同样适用于单节点系统中资源不足、性能低下、影响用户交互体验的场景，因此本文的研究成果具有良好的应用前景。

参考文献

[1] 夏伟,李志坚,蔡文婷,等. 云GIS架构下的电网数据并行计算及编辑技术研究[J]. 测绘地理信息,2022,47(6):101-105

[2] 王凯,曹建成,王乃生,等. Hadoop支持下的地理信息大数据处理技术初探[J]. 测绘通报,2015(10):114-117

[3] 陶伟东. 多核/众核环境下栅格数据分析并行化方法研究[D]. 南京:南京大学,2017

[4] 陈能成,龚健雅,朱欣焰,等. 多级缓冲提升海量影像数据在线服务质量[J]. 测绘通报,2007(6):19-22

[5] 左尧,王少华,钟耳顺,等. 高性能GIS研究进展及评述[J]. 地球信息科学学报,2017,19(4):437-446

[6] 关雪峰,曾宇媚. 时空大数据背景下并行数据处理分析挖掘的进展及趋势[J]. 地理科学进展,2018,37(10):1 314-1 327

[7] Shekhar S,Xiong H,Zhou X. Encyclopedia of GIS(Springer Reference) 2nd ed[M]. Springer,2017

[8] Karthi P. Secure Geospatial Data Storage Using SpatialHadoop Framework in Cloud Environment[J]. Int J of Internet Technology and Secured Transactions,2018,8(4):516-527

[9] Li X,Yu H Y,Yuan L,et al. Query Optimization for Distributed Spatio-temporal Sensing Data Processing[J]. Sensors,2022,22(5):1 748

[10] 彭旭. 基于GeoSpark的空间大数据计算优化[J].电子技术与软件工程,2019 (11):172

[11] Abdul J, Alkathiri M, Potdar M B. Geospatial Hadoop (GS-Hadoop) an Efficient Mapreduce Based Engine for Distributed Processing of Shapefiles[C]//2016 2nd International Conference on Advances in Computing, New York: IEEE,2016

[12] David G,Brian T,Marjorie S,et al. HTTP权威指南[M]. 北京:人民邮电出版社,2020

[13] 金凤华. 分布式消息中间件应用场景及建设设想[J]. 中国金融电脑,2020 (9):64-66

[14] 王善勤,王立辉,颜洲,等. 面向高并发的消息中间件设计与实现[J]. 海南师范大学学报(自然科学版),2023,36(1):29-37

[15] 丁威,梁勇. RocketMQ实战[M]. 北京:人民邮电出版社,2022

[16] 石建良. 基于RocketMQ框架的消息通信与监控系统[D]. 武汉:华中科技大学,2019

[17] 李璐,张广泉. 消息中间件的体系结构研究[J]. 苏州大学学报(工科版),2007(3):10-14

[18] 郭明强,黄颖. WebGIS之OpenLayers全面解析(第2版)[M]. 北京:电子工业出版社,2021

[19] 袁长征,林江伟,汪剑云. 轨道交通控制保护区空间数据分析与可视化系统设计与实现[J]. 测绘通报,2019(增刊):210-213

[20] 吴瑞龙,何华贵,张鹏程,等. 基于GeoServer REST的矢量地图服务自动发布研究与应用[J]. 测绘地理信息,2021,46(增刊):32-35

[21] 肖建华,王厚之,彭清山,等. 地理时空大数据管理与应用云平台建设[J]. 测绘通报,2016(4):38-42