



基于微服务架构的城市交通大数据治理框架

吴丹 李旭芳 刻洁 黄荣裕

摘要: 随着交通行业的快速发展和智能交通系统的广泛应用,大量交通数据不断生成和累积,对数据治理带来了更高的要求。本文以微服务架构为基础,结合中国交通行业的实际情况,探讨交通行业数据治理的关键问题及解决方案。通过分析交通行业大数据应用的现状和典型数据特征,提出了一种基于微服务架构的数据治理框架,涵盖了数据采集、清洗、存储、分析、隐私保护、质量管理、共享与开放以及法律法规合规等方面。本文旨在为致力于交通行业数字化转型的数据治理团队提供有效参考与帮助。

关键词: 交通大数据;数据治理;微服务;数据隐私保护;数据共享

引言

近年来,随着交通行业的发展和智能交通系统的应用,大量交通数据被生成和积累,数据治理的问题逐渐由企业扩展至政府领域,成为学术界研究的热点^[1]。交通行业作为国家经济的重要支柱,其数据治理水平对整个社会的运行和发展具有重要意义。随着大数据、云计算、物联网等技术的发展,交通行业数据呈现出爆炸式增长,数据类型和来源日趋多样化^[2]。这使得交通行业数据治理面临诸多挑战,如数据质量、数据安全、数据隐私、数据共享等。为了应对这些挑战,本文提出了一种基于微服务架构的交通行业数据治理框架。

交通行业大数据特征及应用中的问题

1.数据类型多样、数据量大

交通行业数据包括结构化数据(如数据库中的表数据)、半结构化数据(如XML、JSON等格式的数据)和非结构化数据(如图像、视频、文本等)。数据产生速度快,数据量大,如实时定位数据、报警数据,对数据存储和处理能力提出了较高要求。

2.数据实时性要求高

数据的实时性要求较高,如实时路况信息、交通拥堵预测等,需要快速响应和处理。为了满足这些需求,数据处理系统需要具备低延迟、高吞吐量的特点,以便在有限的时间内分析和处理大量实时数据。

3.数据安全性和隐私保护要求高

交通数据中可能包含个人隐私信息,如位置数据、出行记录等,如何在保障数据安全和隐私的前提下进行数据治理成为一大挑战。此外,交通数据还涉及国家安全和公共安全等方面的问题,因此需要在数据收集、存储和处理过程中确保数据的安全性和合规性。

4.数据地域性强

交通数据具有明显的地域性特征,不同地区的交通状况、道路状况和出行需求差异较大。因此,在进行交通大数据分析时,需要充分考虑地域

特点,以便为各地区提供更有针对性的交通优化建议和措施。

5.数据共享与协同需求

交通行业涉及多个部门和企业,如政府部门、公共交通公司、出租车公司等。为了实现交通行业的高效治理,需要在这些部门和企业之间实现数据共享和协同。这要求数据治理框架具备良好的互操作性和标准化,以便在不同系统和平台之间实现顺畅的数据交换。

基于微服务架构的数据治理框架

一个基于云计算和微服务架构的数据治理框架可以有效解决交通行业大数据应用中的问题。

这个框架包含数据采集与集成、实时数据处理与分析、数据存储与管理、数据安全与隐私保护、数据共享与协同、微服务架构六大方面。

1.数据采集与集成

如图1所示,该框架利用各种数据采集工具和接入程序,实时收集来自不同来源的交通数据,如传感器、监控摄像头、GPS设备、公共交通系统等。数据接入程序负责对原始数据进行清洗、去重、格式转换等操作,以提高数据质量。数据治理服务根据数据标准对数据进行深度融合和加工,数据集成模块将整合不同来源的结构化、半结构化和非结构化数据,为后续分析提供统一的数据视图。



图1 数据采集清洗入库流程

2.实时数据处理与分析

如图1所示,该框架使用流处理技术(Apache Kafka)对实时数据进行处理和分析。实时数据处理模块可以对交通状况、拥堵情况等实时监测和预测,为交通管理部门提供实时决策支持。此外,利用机器学习和深度学习技术可以从海量交通数据中挖掘有价值的信息和模式,为交通优化提供智能建议。

3.数据存储与管理

该框架采用 Clickouse 存储大量交通数据,能够提高存储能力,满足数据量大和地域性强的挑战。同时,通过数据分层存储策略,将热数据存储在高速缓存(Redis)中,以提高数据访问速度。

4.数据安全与隐私保护

为确保数据在传输和存储过程中的安全性,该框架采用加密技术对数据进行加密,支持 SM4、

AES、SM3、FPE、XGS 等加密算法。对于涉及个人隐私和国家安全的敏感数据,使用数据脱敏(如数据掩码、伪造数据等)和访问控制技术,确保只有授权用户可以访问敏感数据。同时,遵守相关法律法规,确保数据治理的合规性。

5.数据共享与协同

该框架构建了数据共享平台,通过 API 和数据交换规范实现跨部门和企业的数据共享。数据共享平台可提供数据质量评估、数据血缘分析等功能,以确保数据的准确性、完整性和一致性。此外,平台可支持数据权限管理,以便在保护隐私的同时实现数据共享。图2是静态数据共享流程,对于动态数据,会根据交通实时数据的数据紧凑性、交换实时性和基于数据流的交换特点,建设有别于基于报文交互式的服务调用通道的交通实时数据交换通道。



静态数据共享流程

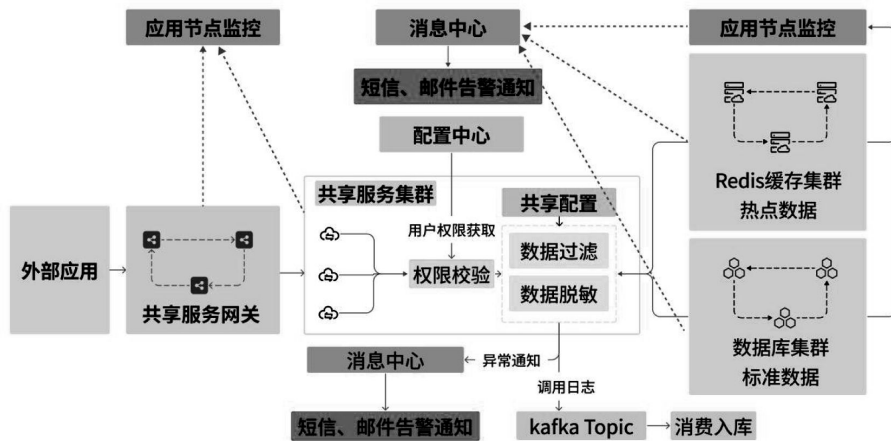


图 2 数据共享流程

6.微服务架构

该框架采用微服务架构 (如 Spring Cloud、Istio) 组织各个数据处理和分析模块, 实现模块化、

解耦合可扩展性。微服务架构有助于快速响应不断变化的业务需求, 提高系统的灵活性和可维护性。技术架构如图 3 所示。



(接上图)

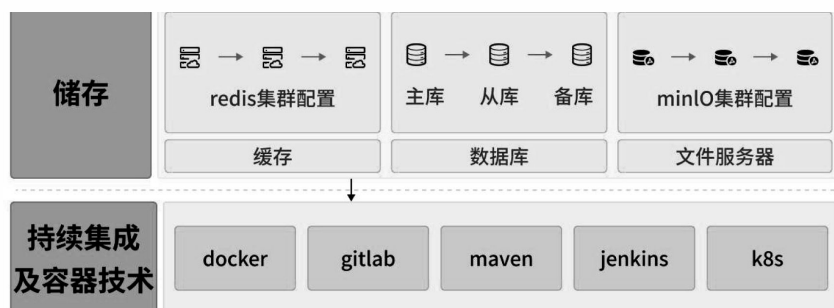


图3 基于微服务技术的数据治理技术架构

结语

本文提出了一种基于微服务架构的交通行业数据治理框架,能够为交通管理部门、企业和公众提供高质量的数据服务,助力实现智能交通和城市可持续发展,已在宝鸡市交通管理局的实际案例中取得了良好效果。然而,仍存在一些挑战和局限性,如非结构化数据处理、数据源变化、数据共享与隐私保护之间的平衡等。为应对这些挑战,采用了自研的数据集成工具处理非结构化数据;引入数据版本控制应对数据源变化;以及采用数据脱敏技术(如数据掩码、伪造数据等)对敏感数据进行处理;采用加密技术(如AES、RSA)对数据进行加密存储和传输;在保护隐私的同时实现数据共享。

未来研究可以从以下几个方向展开:一是研究新兴技术(如区块链、人工智能等)在交通行业数据治理中的应用;二是探讨如何进一步完善和优化现有数据治理框架,以适应不断变化的需求;三是深入研究数据治理过程中的伦理和法律问题,确保数据治理的合规性和道德性。通过优化和完善这一框架,不断为交通行业数据治理领域的发展提供有力支持。

(作者单位:厦门卫星定位应用股份有限公司)

参考文献

- [1] 孙颖.治理主体视角下政府数据治理路径研究[J].超星期刊,2021(1):140.
- [2] 张伟,李婷.大数据时代交通行业数据治理策略研究[J].信息技术,2018,35(12):82-85.

- [3] 周涛,唐小勇.城市交通大数据挖掘与应用实践[M].上海:同济大学出版社,2022.
- [4] 朱小杰.城市交通综合数据中心系统研究和实现[J].交通世界,2017,29:37-38.
- [5] 汪祖云.交通数据中心数据总体架构与数据共享交换平台的设计研究[J].交通运输系统工程与信息,2005(1):74-76.
- [6] 中国信息通信研究院云计算与大数据研究所.数据安全治理实践指南 [EB/OL].(2021-07)[2023-09-27].
<http://www.caict.ac.cn/kxyj/qwfb/ztbg/202107/P020210720377857004616.pdf>.
- [7] 徐国庆.通过数据治理提升数据应用价值思路探析[J].中小企业管理与科技,2023(1):42-24.
- [8] 罗健,欧阳钺.交通数据治理框架及应用分析——以成都为例[J].黑龙江交通科技,2023(5):161-163.
- [9] 王磊,刘洋.基于云计算的交通行业数据治理体系研究[J].计算机工程与应用,2019,55(6):208-212.
- [10] 赵婧,陈晓红.交通行业大数据治理框架研究[J].计算机科学,2020,47(S1):328-332.
- [11] 孙琦,马丽萍.交通行业数据治理中的隐私保护技术研究[J].计算机工程,2021,47(2):248-252.
- [12] 李婧,王瑞.基于区块链的交通行业数据治理体系研究[J].计算机工程与设计,2021,42(10):2468-2472.