

工业物联网设备运行参数数据库的构建与查询优化

滕硕

中国矿业大学徐海学院

DOI:10.32629/acair.v3i4.17910

[摘要] 随着工业物联网(IIoT)的广泛应用,海量设备产生的运行参数数据对存储系统的结构性、实时性与查询效率提出了更高要求。传统数据库在处理高频、非结构化和多源异构数据时存在性能瓶颈与扩展性问题。本文围绕工业物联网场景下设备运行参数的特点,设计并实现了一套高效的运行参数数据库系统。通过构建时间序列优化的数据模型、采用混合存储架构与压缩策略,并引入索引优化与查询重写机制,显著提升了数据插入与查询效率。结合实际工厂运行数据进行测试验证,系统在数据存储性能、查询响应速度与资源利用率方面均表现出良好适应性。

[关键词] 工业物联网; 运行参数; 数据库构建; 查询优化; 时间序列数据

中图分类号: G250.74 文献标识码: A

Construction and Query Optimization of Operational Parameter Databases for Industrial Internet of Things Devices

Shuo Teng

Xu Hai College, China University of Mining and Technology

[Abstract] With the widespread application of the Industrial Internet of Things (IIoT), the massive operational parameter data generated by devices imposes higher demands on the structural integrity, real-time performance, and query efficiency of storage systems. Traditional databases face performance bottlenecks and scalability issues when handling high-frequency, unstructured, and multi-source heterogeneous data. This paper focuses on the characteristics of device operational parameters in industrial IoT scenarios, designing and implementing an efficient operational parameter database system. By constructing a time-series-optimized data model, adopting a hybrid storage architecture with compression strategies, and introducing index optimization and query rewriting mechanisms, data insertion and query efficiency are significantly improved. Testing with actual factory operational data demonstrates that the system exhibits good adaptability in data storage performance, query response speed, and resource utilization.

[Key words] Industrial Internet of Things; operational parameters; database construction; query optimization; time-series data

工业物联网通过传感器、控制器与通信网络,实现了设备状态、运行参数与环境变量的实时采集与远程控制,为工业智能化奠定了基础。然而,IIoT系统中设备运行参数具有数据量大、更新频繁、时序性强等特点,给数据存储与处理带来了巨大挑战。传统关系型数据库难以满足高频写入与低延迟查询的双重需求,而现有的时序数据库在复杂查询支持与系统扩展性方面仍存在不足。为了实现工业设备运行状态的高效监控、故障预警与运行优化,迫切需要构建一套兼顾数据写入效率、查询性能与系统可扩展性的运行参数数据库系统。

1 工业物联网运行参数数据特征分析

1.1 工业设备运行参数的分类与数据特性

工业物联网运行参数需按设备功能与监测目标分类,不同类别数据呈现差异化特性^[1]。从分类来看,可分为电气参数、物理参数、状态参数、环境参数,各类参数对应不同监测需求,共同反映设备运行状态。数据特性方面,首先是强时序性,参数采集按固定或动态周期生成时间序列数据,需保留时间维度关联以分析趋势;其次是高实时性,部分关键参数需实时传输与处理,否则可能延误故障响应;再者是数据价值不均衡,多数时段参数处于正常范围,仅异常数据价值密度高。

1.2 运行参数采集频率与数据吞吐量评估

运行参数采集频率需结合设备类型与监测精度需求动态设定,直接影响数据吞吐量规模^[2]。采集频率划分需遵循“关键参

数高频、常规参数低频”原则:对核心生产设备的安全相关参数(如化工反应釜压力、机床主轴转速),采集频率需达到毫秒级或秒级,确保实时捕捉异常波动;对非关键参数,可采用分钟级或小时级采集,降低数据冗余。数据吞吐量评估需综合采集频率、设备数量与参数维度:单条生产线若包含数百台设备,每台设备监测数十项参数,高频采集下日均数据量可达TB级甚至PB级;且吞吐量存在周期性波动,生产高峰期设备满负荷运行,参数采集频次与数据量显著增加,低谷期则相应减少。

1.3 多源异构数据集成需求与挑战

工业互联网多源异构数据集成需满足“全量汇聚、统一管理”需求,同时面临技术与管理层面挑战^[3]。需求方面,首先是数据格式统一,需将不同传感器、不同设备生成的异构数据(如JSON格式的状态数据、二进制格式的振动波形数据、文本格式的故障日志)转化为标准化格式,消除数据交互壁垒;其次是数据时空对齐,需关联同一设备不同参数的采集时间、同一区域不同设备的空间位置,确保数据关联性以支持跨参数、跨设备分析;再者是数据质量保障,需过滤噪声数据、补全缺失数据,避免劣质数据影响后续分析。挑战层面,一是协议兼容性问题,不同工业设备采用Modbus、Profinet、OPCUA等不同通信协议,数据接入需适配多协议;二是实时集成难度大,高频数据与低频数据同时接入时,易出现数据拥堵或延迟;三是数据主权与安全问题,多源数据可能来自不同部门或供应商,集成过程需保障数据隐私与访问权限控制,避免数据泄露或越权使用。

2 运行参数数据库的系统架构设计

2.1 数据模型设计:结构化与时序数据融合建模

工业互联网运行参数数据库模型需实现结构化数据与时序数据的深度融合,兼顾查询灵活性与存储效率。结构化数据建模针对设备基础信息、参数定义等静态数据,采用关系模型设计,通过表结构关联设备与参数属性,支持多维度筛选查询。时序数据建模需突出时间维度,采用时序模型设计,将“设备ID-参数ID-时间戳”作为联合主键,存储对应参数值,同时保留数据采集时的附加信息(如采集传感器编号、数据质量标记),确保时序数据与设备、参数属性的关联可追溯。融合建模关键在于建立结构化数据与时序数据的关联机制:通过设备ID、参数ID实现静态属性与动态时序数据的跨模型关联,既支持基于结构化属性筛选时序数据,也支持基于时序数据反查设备静态信息,避免两种数据割裂存储导致的查询效率低下。

2.2 存储架构:冷热数据分层与压缩存储策略

存储架构设计需采用“冷热数据分层+压缩存储”组合策略,平衡存储成本与访问效率。冷热数据分层依据数据访问频率与时效性划分:热数据指近7天或30天内的实时数据、高频访问数据,需存储于高性能存储介质,确保毫秒级查询响应;温数据指30天至1年内的历史数据,访问频率中等,可存储于普通机械硬盘或分布式存储系统;冷数据指1年以上的归档数据,仅用于合规审计或长期趋势分析,访问频率极低,可存储于低成本归档存储。

2.3 数据一致性与高可用性机制设计

数据一致性与高可用性是工业物联网数据库的核心保障,需通过多层机制确保数据可靠与服务不中断^[4]。数据一致性机制需解决“分布式存储与并发访问”引发的一致性问题:首先采用分布式事务协议,确保多节点数据写入时要么全部成功、要么全部回滚,避免部分节点写入失败导致的数据不一致;其次实现数据版本控制,对更新频繁的参数数据保留历史版本,防止并发修改覆盖;再者通过定时数据校验,及时发现并修复不一致数据。

3 查询优化策略与实现方法

3.1 多维索引机制与查询路径选择

多维索引机制需适配工业物联网数据的多维度查询需求,同时通过智能路径选择提升查询效率。索引设计需覆盖高频查询维度:针对设备维度,建立设备ID索引,支持快速查询单台或多台设备的所有参数数据;针对参数维度,建立参数ID索引,支持按参数类型筛选跨设备数据;针对时间维度,建立时间戳索引,优化时序范围查询;针对状态维度,建立设备故障状态索引,加速异常数据检索。

3.2 查询重写与执行计划优化算法

查询重写与执行计划优化需通过语法与逻辑层面优化,减少无效计算,提升查询执行效率。查询重写针对用户输入的原始查询语句,通过等价变换简化逻辑:例如将嵌套子查询改写为连接查询,减少表扫描次数;将重复条件合并,降低过滤复杂度;对聚合查询,若存在预计算结果,则直接调用预计算数据,避免重复计算。执行计划优化算法需从物理执行层面规划最优步骤:采用动态规划或贪心算法生成候选执行计划,评估各计划的资源消耗,选择代价最低的方案。例如对多表连接查询,优化表连接顺序,将数据量小的表作为驱动表,减少连接过程中的数据传输量;对大数据量排序查询,采用外部排序结合分区技术,避免内存溢出;对时序数据的滑动窗口查询,利用数据时序连续性优化窗口计算,减少数据重复读取,通过多维度优化确保复杂查询仍能高效执行。

3.3 面向复杂分析的SQL与时序查询融合

面向工业场景的复杂分析需求,需实现标准SQL与时序查询语言的深度融合,兼顾通用性与专业时序分析能力。融合设计需保留SQL的通用性,支持传统关系型查询,同时扩展时序专属语法,满足时序分析需求。例如支持“SELECT设备ID,AVG(温度)OVER(PARTITION BY设备ID ORDER BY时间戳ROWS BETWEEN 10 PRECEDING AND CURRENT ROW)AS近10次温度均值FROM时序表”此类融合查询,既利用SQL的聚合与分区语法,又通过窗口函数实现时序滑动分析。

4 系统实现与性能测试

4.1 系统部署架构与关键技术实现

工业互联网运行参数数据库需采用分布式部署架构,结合多技术栈实现核心功能。部署架构采用“边缘-云端”协同模式:边缘节点部署在工业现场,负责实时数据采集、预处理与本地缓

存,减少数据传输带宽占用,同时支持边缘侧快速查询;云端节点采用集群部署,由数据接入层、存储层、计算层、应用层组成,接入层适配多协议数据接收,存储层实现冷热数据分层存储,计算层负责查询处理与复杂分析,应用层提供可视化界面与API接口。关键技术实现方面,数据接入层采用消息队列实现高并发数据接收与缓冲,避免数据丢失;存储层结合时序数据库与关系数据库,时序库存储高频时序数据,关系库存储结构化静态数据;计算层引入分布式计算框架,支持批处理与流处理结合,满足实时查询与离线分析需求;同时采用容器化技术实现集群弹性伸缩,根据数据量变化动态调整节点数量,确保系统扩展性与资源利用率平衡。

4.2 实验设计: 数据写入性能与查询效率测试

实验设计需模拟工业实际场景,从数据写入与查询两维度测试系统性能,验证系统适配能力^[5]。数据写入性能测试需构建多场景压力模型:首先是常态写入测试,按常规采集频率向数据库写入多设备多参数数据,监测写入吞吐量、延迟时间与成功率,评估系统常态负载下的稳定性;其次是峰值写入测试,模拟生产高峰期或设备异常时的高频数据写入,测试系统是否能应对瞬时数据量激增,是否出现写入阻塞或数据丢失;再者是长时间连续写入测试,持续数天或数周写入数据,监测存储容量增长、写入性能衰减情况,评估系统长期运行稳定性。查询效率测试需覆盖典型查询场景:包括单设备单参数的时序范围查询、多设备多参数的聚合查询、复杂关联查询、时序分析查询,对每种查询场景记录响应时间、CPU与内存占用率,分析不同查询类型下的系统性能表现,为后续优化提供方向。

4.3 与主流数据库系统性能对比分析

性能对比分析需选择工业物联网领域常用的主流数据库系统,从多维度评估目标系统的优势与不足。对比对象包括专用时序数据库(如InfluxDB、Prometheus)、传统关系数据库(如MySQL、PostgreSQL)、分布式数据库(如HBase、Cassandra)。对比维度需覆盖核心性能指标:数据写入方面,对比不同采集频率

下的写入吞吐量与延迟,评估目标系统在高频写入场景的优势;数据存储方面,对比相同数据量下的存储占用(结合压缩策略),分析目标系统的存储成本优势;查询效率方面,针对不同查询场景(时序查询、关联查询、聚合查询)对比响应时间,突出目标系统在融合查询场景的性能特点;扩展性方面,对比集群节点扩容时的性能线性增长情况,评估目标系统的横向扩展能力;可靠性方面,对比故障恢复时间、数据丢失率,验证目标系统的高可用保障能力。

5 结语

本文针对工业物联网设备运行参数数据高频、异构、时序性强的特点,提出了一种集数据建模、存储优化与查询加速于一体的数据库系统设计方案。通过在数据结构、索引体系与查询优化方面的创新,有效提升了系统在实际工业环境中的响应速度与处理能力。实验结果表明,该系统在数据写入吞吐量、查询延迟与资源利用率方面均优于传统方案。未来将进一步研究系统的智能调度能力与边缘计算集成机制,为工业智能化发展提供更强支撑。

[参考文献]

- [1]吴文鹏.5G赋能下工业物联网设备互联及协同制造模式探索[J].科技资讯,2025,23(19):57-59.
- [2]杨桃李,李欣.基于工业物联网平台的设备自动化管理系统优化设计与实践[J].中国集成电路,2025,34(9):58-63.
- [3]邵汉舒.基于设备管理工业物联网的设备异常监测与故障诊断方法研究[J].物联网技术,2025,15(16):87-90+94.
- [4]李若杼,王嘉欣.工业物联网与电子设备远程检测[J].机械工程与自动化,2025,54(04):216-218.
- [5]陈珑.基于5G通信的工业物联网设备数据低延时传输方法[J].电工技术,2025,(14):237-239+242.

作者简介:

滕硕(2004--),男,汉族,江苏省徐州市人,全日制本科在读,单位:中国矿业大学徐海学院,研究方向:物联网工程。