

# 参考答案

1. 请简述时序数据管理的关键技术和挑战。

答：

时序数据管理的关键技术包括高效存储结构(如时序数据库、列式存储)、数据压缩算法(如 Delta 编码、字典压缩)、索引优化(如时间范围索引)以及流处理框架(如 Flink、Spark Streaming)。这些技术旨在解决时序数据的高写入吞吐、低存储成本及快速查询需求。挑战主要体现在数据量的爆炸性增长、高并发读写压力、实时性与准确性的平衡,以及多维度查询的优化。此外,时序数据常伴随噪声和非平稳性,对数据清洗和异常检测提出了更高要求。

2. 请分析多源数据融合的基本原理和方法。

答：

多源数据融合通过整合不同来源、格式的数据,提升信息的完整性与可靠性。其基本原理包括数据对齐(时间、空间同步)、关联分析(如实体链接)和不确定性建模(如贝叶斯方法)。常用方法可分为三类:传感器级融合(直接合并原始数据)、特征级融合(提取共性特征)与决策级融合(集成局部推理结果)。例如,卡尔曼滤波可用于实时传感器融合,而深度学习模型(如注意力机制)能够自适应加权多源输入,提高融合精度。

3. 请说明模型与数据融合的优势和应用场景。

答：

模型与数据融合结合物理/机理模型与数据驱动方法,兼具理论严谨性和适应性优势。例如,数字孪生通过实时数据校准仿真模型,实现动态预测与优化。应用场景包括工业智能制造(设备故障诊断)、气象预报(数据同化修正模型参数)和自动驾驶(多传感器融合感知)。这种融合能减少纯数据模型对样本量的依赖,提升泛化能力,同时在数据缺失时依靠模型补充信息,增强系统鲁棒性。

4. 请论述实时数据监控的架构和关键技术。

答：

实时数据监控通常采用分层架构:数据采集层(Agent、传感器)、传输层(Kafka、MQTT 消息队列)、处理层(流计算引擎如 Flink)以及可视化层(Grafana、ELK)。关键技术包括流式数据处理(窗口计算、状态管理)、复杂事件处理(CEP 引擎)、低延迟存储(时序数据库)和告警规则引擎。此外,水平扩展的微服务架构和容错机制(如 Checkpoint)保障了系统在高负载下的稳定性与实时性。

5. 请比较不同数据可视化方法的特点和适用场景。

答:不同数据可视化方法的特点和适用场景如下。

①折线图/面积图:适合时序数据趋势分析,如股票走势监控。

②柱状图/热力图:用于分类对比与密度分布,如销售统计或地理数据可视化。

③散点图/气泡图:揭示变量关联与聚类,常见于多维数据分析。

④网络图:展示实体关系,如社交网络分析。

⑤仪表盘:直观显示关键指标状态,适用于实时监控场景。选择时需考虑数据维度、交互需求及受众认知门槛,平衡信息密度与可读性。