OLTP: 行存储

OLAP:列存储

工作负载通常包括: 1. 事务数据库需要提供及时的商业决策统计信息 2. 数据仓库需要满足实时分析的事务需求

SAP HANA: 同时满足OLTP和OLAP高效处理 通过复杂的多步骤记录生命周期管理方法实现

## 1. 记录的生命周期管理

三个物理记录表示过程: L1, L2, Main

• L1:接收传入数据请求,行存储(快速插入、删除、更新、投影),无压缩,每个节点保存10000-100000条数据

L2:列存储,为内存使用提供字典化的编码,不排序,需要使用二级索引支持点查询优化,适合千万条数据

Main:列存储,高压缩比(字典排序,在字典中的位置以比特打包方式存储),同时字典也是压缩的

统一表访问:访问不同存储的通用抽象接口,记录传播是异步的(运行时互不干扰)

记录在生命周期中以不同方式传播(转换?)

行格式到列格式的转换: 行被每个列值分割, 列接列的方式插入到L2阶段

具体步骤: 1. 添加新条目附加到字典中(并行) 2. 列值以字典编码存储 (并行) 3. 删除L1中的条目