

经济与管理

陈德莉. 上海烟草数据中心卷烟零售数据库的架构优化 [J]. 中国烟草学报, 2019, 25 (5) . CHEN Deli. Architecture optimization of cigarette retail database in Shanghai tobacco data center [J]. Acta Tabacaria Sinica, 2019, 25(5). doi: 10.16472/j.chinatobacco.2018.403

上海烟草数据中心卷烟零售数据库的架构优化

陈德莉

上海烟草集团有限责任公司, 信息中心上海, 上海长阳路717号 200082

摘 要: 为解决烟草零售户订单数据处理的性能瓶颈问题, 本文引入大规模并行处理的思路, 对上海烟草数据中心平台中存储烟草零售订单数据的数据库架构重新进行了规划和设计, 部署了大规模并行处理架构的数据库, 并优化数据处理业务, 提升了数据的加工性能, 由优化前的 10 小时以上缩短到优化后的 2 小时以内。

关键词: 数据库; 并行处理; 性能优化

随着烟草行业信息化发展的不断深入, 上海烟草数据中心的数据仓库逐年积累的数据不断增多, 在国家局烟草专卖局推出下行数据之后, 上海烟草数据中心接入物流扫码数据、零售户订单数据, 社会库存数据等, 其中零售户订单数据的数据量逐年增多, 逐步达到 10TB 以上, 难以保障原有的数据加工、数据更新等数据处理性能。基于上述背景, 研究并提出了零售户数据处理性能优化方法。

目前, 处理海量数据的软件产品分为两类, 一类是基于大规模并行处理 (MPP) 架构的关系型数据库; 另一类是基于 Hadoop 平台的软件产品。MPP (Massively Parallel Processing), 即大规模并行处理, 在数据库非共享集群中, 每个节点都有独立的磁盘存储系统和内存系统, 业务数据根据数据库模型和应用特点划分到各个节点上, 每台数据节点通过专用网络或者商业通用网络互相连接, 彼此协同计算, 作为整体提供数据库服务。非共享数据库集群有完全的可伸缩性、高可用、高性能、优秀的性价比、资源共享等优势。MPP 采用无共享资源结构, 优势体现在大规模存并行计算上^[1-3]。Hadoop 分布式文件系统 (HDFS) 是运行在通用硬件上的分布式文件系统, 它可提供一个高度容错性和高吞吐量的海量数据存储解决方案^[2]。通过对这两种软件产品的评估, 并基于现有关系型数据库的运行环境, 决定引入 MPP 架构的关系型数据库: 一方面可以通过改变

数据库的架构提升数据处理性能; 另一方面, 可以不改变现有数据库产品以及相关技术降低性能优化造成的影响程度, 以最小的投入完成性能提升^[3]。在升级数据库架构之后, 对现有的数据架构做了进一步的优化, 主要包括了数据物理模型和数据处理两方面。

1 零售户订单数据处理现状

1.1 数据库架构

上海烟草数据中心零售户订单数据库架构如图 1 所示: 零售户订单数据存储在一台 DB2 单机数据库

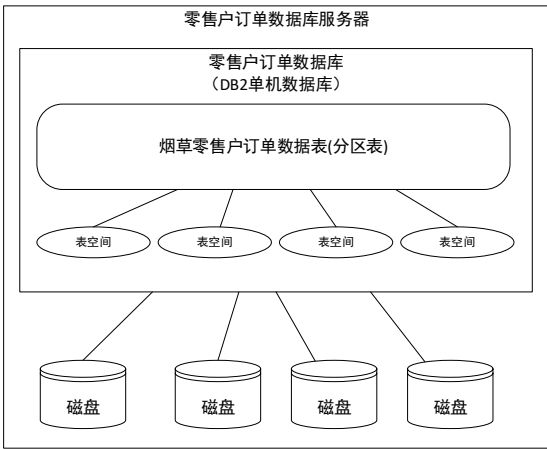


图 1 上海烟草数据中心零售户订单数据架构现状

Fig. 1 Current architecture of retail order database of Shanghai tobacco data center

中,数据以分区表的方式存储,每个月作为一个数据分区,每个分区的数据分别存储在不同的表空间上。

1.2 数据处理业务

零售户订单数据的应用主要是统计若干零售户订单指标用于一些分析型报表,指标包括零售户需求量和订单量、三维五率、零售户个数等。零售户订单数据处理的业务流程如图2所示:

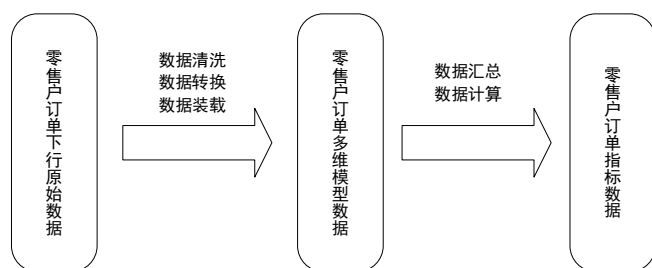


图2 上海烟草数据中心零售户订单数据处理业务流程

Fig. 2 Retail order data processing flow of Shanghai tobacco data center

零售户订单数据分为三层,第一层是零售户订单原始数据,该层数据为零售户订单原始数据,通过国家局零售户订单数据下行到数据中心,该层数据加工不对数据做任何的转换,是国家局零售户订单数据的拷贝副本;第二层是零售户订单多维模型数据,该层数据由原始数据通过数据清洗、数据转换、数据装载等数据处理之后生成。数据清洗是指清洗原始数据中的脏数据,数据转换是指将原始数据转换为星型或者雪花型模型数据;第三层是零售户订单指标数据,该层数据由多维模型数据计算生成,用于各类报表和应用。

2 零售户订单数据处理性能问题分析

根据统计每个数据处理环节的耗时,上海烟草数据中心处理零售户订单数据的性能瓶颈主要在于数据汇总。原始订单数据的时间细粒度为“日”,平均每天的数据量为800万条,通常都需要计算某段周期的

数据,例如周订单量、月订单量、月零售户个数、三个月零售户个数、年零售户个数等,这些数据汇总处理耗用大量的时间。其中,统计零售户个数耗时最长,零售户个数指标的统计口径为:周期内有订单量的零售户数量。计算零售户指标的耗时如下:

表1 计算零售户指标耗时情况

Tab. 1 Time consuming for calculating retailer indexes

统计指标	计算平均耗时(分钟)
周零售户个数	38 分钟
月零售户个数	63 分钟
近3月零售户个数	156 分钟
近1年零售户个数	747 分钟

通过分析上述性能问题,发现统计零售户订单数据的性能瓶颈主要在于数据汇总环节。现有烟草零售户订单数据部署在一台单机DB2数据库上,若要提升数据处理性能,使得计算时间能够减少至60分钟以内,在硬件方面,升级单机服务器硬件资源已经意义不大,需要使用支持并行处理的海量数据处理产品,横向扩展硬件资源,提升数据处理并行度。为了验证其可行性,将若干年份的烟草零售户订单数据迁移到MPP架构的DB2数据库以及Hadoop平台的HIVE中,测试两个平台下汇总1年烟草零售户订单数据的性能,发现两者的数据汇总性能相对于现单机数据库平台均提升,运行耗时如表2所示。

根据上述测试,引入海量数据处理产品提升零售户订单数据的处理性能是可行的。此外,目前的数据处理效率也并非达到最优在数据处理业务方面,从多维模型数据直接计算各周期的指标数据,会存在部分数据重复汇总的情况,例如月零售户指标与近3个月零售户指标,月度数据是重复汇总的。因此还需要优化现有数据架构,包括数据物理模型,数据处理业务流程等,以此提升数据处理效率。

表2 在DB2 MPP数据库、Hadoop平台、DB2单机数据库下汇总1年零售户订单数据

Tab. 2 Performances of DB2 MPP, Hadoop and DB2 for summing 1-year retailer order data

服务配置/性能	DB2 MPP 数据库	Hadoop 平台 HIVE	DB2 单机数据库
服务器配置	CPU: 10 核 2.4G 赫兹 内存: 128G 磁盘: 主节点(1个) 1TB 处理节点(4个) 5TB	CPU: 10 核 2.4G 赫兹 内存: 128G 磁盘: 主节点(1个) 1TB 处理节点(4个) 5TB	CPU: 10 核 2.4G 赫兹 内存: 128G 磁盘: 5TB
运行耗时(分钟)	8.6 分钟	11.3 分钟	32 分钟

3 数据库架构优化

3.1 数据库架构优化设计

目前，基于大规模并行处理（MPP）技术处理海量数据的软件产品分为两类，一类是基于大规模并行处理（MPP）架构的关系型数据库；另一类是基于 Hadoop 平台的软件产品。上海烟草集团数据中心列举了若干评估项对两类产品进行评估，评估项列表如下所示：

表 3 海量数据处理产品选型评估项列表
Tab. 3 List of evaluation for massive data processing products

评估项	评估标准	MPP 关系型数据库	Hadoop 平台产品
处理海量数据能力	具备：40 分 不具备：0 分	具备（40 分）	具备（40 分）
可扩展性 （考虑是否可扩展以及扩展便捷性）	强：10 分 中：5 分 弱：0 分	中：5 分	强：10 分
产品供应商服务保障 （考虑产品是否有大型厂商提供技术服务保障）	好：10 分 中：5 分 差：0 分	好：10 分	中：5 分
技术先进程度 （考虑技术的先进性）	高：10 分 一般：5 分 低：0 分	一般：5 分	高：10 分
涉及现有系统变更 （考虑使用产品后，对现有系统改造的工作量大小）	大：0 分 中：10 分 小：20 分	小：20 分	大：0 分
运维人员对产品熟悉程度 （考虑现有运维人员对产品的熟悉程度）	熟悉：10 分 一般：5 分 不熟悉：0 分	熟悉：10 分	一般：5 分
综合得分	—	90 分	70 分

根据评估结果，决定采用基于大规模并行处理（MPP）架构的关系型数据库作为处理的烟草行业订单数据的平台，提升处理性能。针对单机数据库在处理海量数据时的性能问题，重构数据库架构为大规模并行处理 MPP 架构，以并行处理的方式计算烟草零售户订单数据相关计量指标，提升数据处理性能。简单来说，MPP 是将任务并行的分散到多个服务器和节点上，在每个节点上计算完成后，将各自部分的结果汇总在一起得到最终的结果。总体架构如图 3 所示：

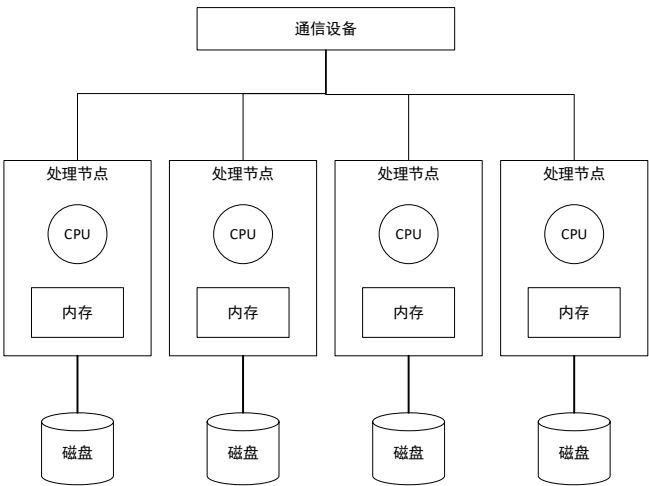


图 3 大规模并行处理总体架构

Fig .3 Massive parallel processing (MPP) architecture

根据上述大规模并行处理 MPP 数据库架构，在数据中心增设以大规模并行处理 MPP 为架构的数据库，用于存储和处理烟草零售户订单数据。MPP 数据库中的订单数据均匀的分布在每个处理节点的磁盘上，独享每个处理节点的硬件资源，在处理数据时，每个节点并行计算，达到提升性能的目的。

3.2 数据库架构优化实现

上海烟草数据中心配置 5 台服务器用于部署 MPP 数据库，数据库平台为 IBM DB2。其中，1 台服务器作为主节点，用于数据库软件安装和处理节点之间的通信，4 台服务器作为处理节点，用于数据存储和计算。其硬件配置如下：

表 4 上海烟草数据中心 MPP 数据库硬件配置
Tab. 4 MPP hardware configuration of Shanghai tobacco data center

服务器	硬件配置
dcmp1 主节点	CPU：10 核 2.4G 赫兹
dcmp2 处理节点	内存：128G
dcmp3 处理节点	磁盘：
dcmp4 处理节点	主节点 1TB
dcmp5 处理节点	处理节点 5TB

以 MPP 架构安装 DB2 数据库的实现过程如下：

（1）数据库安装

在主节点安装数据库软件，设置与各处理节点间的无密码通信，在每个器节点配置通信端口，并为主节点和处理节点配置数据库分区，实现以 MPP 架构安装 DB2 数据库。

(2) 处理节点上的磁盘划分

处理节点用于存储和处理零售户订单数据，可以在每个处理节点中创建若干个容量相同的磁盘分区，提升每个节点的 IO 并行度。

(3) 数据库表空间创建

为数据库设置合理的缓冲池，分区组；可以将磁盘分区作为存储器添加到数据库并创建数据库自动管理的大型表空间和临时表空间，以此提升表空间性能。

(4) 优化数据库参数

调整数据库性能参数优化性能，参数包括：应用堆大小、最大请求 IO 数量、排序堆、索引扫描速度、最大并行度、通信缓冲池大小、最大文件打开数量、应用程序堆大小等。

4 数据架构优化

4.1 数据架构优化设计

在数据库架构升级成大规模并行处理 (MPP) 架构后，原先的数据架构也需要做相应的优化，主要体现在数据物理模型和数据处理流程两个方面。

4.1.1 数据物理模型优化设计

(1) 分区键设计

在大规模并行处理 MPP 架构模式下，任何资源都是非共享的，也包括了磁盘非共享，数据在存储时必须均匀的分布到各处理节点的磁盘上，否则各处理节点处理的数据量不同，导致各节点处理耗时不同。可以使用哈希分布的方式，由数据库将数据键的值映射成哈希值，均匀存储到每个处理节点。设计数据库分区的基本原则是：将数据量大的表分布在所有分区上提高并行处理能力；将数据量小的表放置在单一个分区上；减少分区间的通信^[4]。

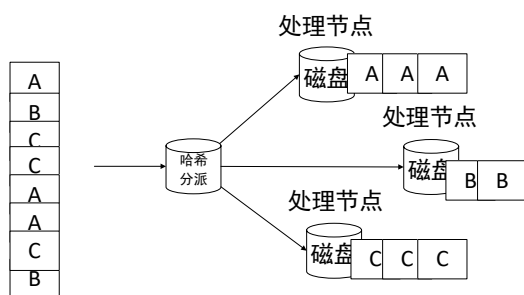


图 4 MPP 数据库均匀分布数据原理

Fig. 4 Data allocation mode of MPP database

数据分区键可以是一个字段，也可以是多个字段，数据分区键的值需要满足数量多，分布广的特点，以确保数据可以均匀分布到每个处理节点。

(2) 数据表分区设计

对于海量结构化数据，可以设置数据分区表提升数据访问性能。表在数据库中是一个逻辑的概念，数据分区表可以视作一个包含了多个子表的表簇^[5]，当存储海量结构化数据时，可以设置数据分布表，数据被拆分为若干个部分，分别存储到数据分区表中的每一个子表，减少不必要的数据库访问。

当设置数据分区表时，需要明确指定表分区键以及键值范围，表分区键以及键值的范围可以根据具体的数据使用情况而定，一般时间序列的数据以日期字段作为数据表分区键，按月、季、年等周期分区。

(3) 数据索引设计

在大规模并行处理 (MPP) 数据库中对表中的字段设置分区索引，提升数据查询性能；除此之外，还可以使用多维聚簇分区索引 (MDC)，MDC 在物理上按照某个或者某几个字段进行群集，采用了“BLOCK”来进行索引的组织，一个 BLOCK 会包含很多条传统索引机制所采用的“行”记录，因而提高了索引的粒度。使得索引的定位变得更快。

4.1.2 数据处理流程优化设计

目前数据处理过程中存在重复汇总数据的情况也影响了性能，MPP 数据库提升数据处理并行度，而优化原有的数据处理业务层次，减少大批量数据重复汇总则可以提升数据处理效率。在多维模型数据层和指标数据层之间添加一个中间层，中间层的数据按时间周期汇总。通过拆解各指标数据的计算过程，提炼其中重复汇总的数据，将这些数据纳入中间层做统一汇总，在指标数据计算时，直接提取已汇总的数据计算，提升计算效率。

中间层可以视为指标数据的预处理层，由于在计算指标数据时，通常的统计周期为周、月、季、年等。中间层可以充分利用 MPP 数据库大规模并行处理的性能优势，预先按时间周期快速汇总数据，缩小了在指标计算时的数据规模，提升处理性能；对于个别无

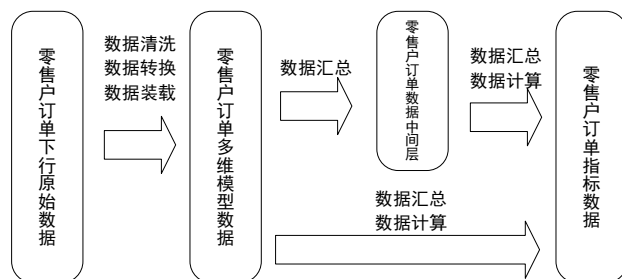


图 5 数据处理流程优化设计

Fig. 5 Optimization of data processing flow

法使用中间层数据计算的指标，也可直接使用多维模型数据层的数据计算。

4.2 数据架构优化实现

4.2.1 数据物理模型优化实现

零售户订单数据中涉及的业务数据主要为零售户信息和卷烟订单信息。其中零售户信息包括零售户号、地区、城乡分类、经营业态、经营规模；订单信息包括：订单号、卖出方、卷烟条码、需求量、订单量、订单金额。

(1) 分区键设定

在 MPP 数据库中零售户订单数据的分区键设为 [订单号, 卷烟条码], 订单号和卷烟条码其两者的组合满足数量多, 分布广的特点, 是合适的分区键。

(2) 数据表分区设定

以数据的业务日期为表分区键按月分区数据。大部分零售户订单数据涉及的统计指标以月为周期, 按月分区数据, 可以有效提升数据访问性能。

(3) 数据索引设定

设置[卖出方]、[零售户号]、[卷烟条码]、[地区]字段的分区索引, 提升这些字段与维度表关联的性能; 设置[业务日期, 地区编码]字段为多维聚簇分区索引 (MDC), 使相同键值的数据在物理上存储在相邻的数据块, 提升数据读取性能。

4.2.2 数据处理流程优化实现

在统计烟草零售户订单数据相关的指标前, 预先以周、月等时间周期维度以及组织机构、地区等维度汇总订单量、需求量、零售户数等度量作为烟草零售户订单中间层数据, 部署在 MPP 数据库上, 缩小统计最终指标的数据源的数据量, 以此提升数据处理性能。计算后的指标数据其数据量已经减少, 以单机数据库的方式存储, 提供各数据集市和应用系统使用。数据量对比如下所示:

表 5 1 年烟草零售户订单数据在预处理前与预处理后的数据量
Tab. 5 Amount of 1-year retail order data before and after pretreatment

预处理时间周期	预处理前	预处理后	指标数据
周数据	约 2 TB	约 300 G	小于 50 G
月数据	约 2 TB	约 130 G	小于 50 G

5 结论

在完成上述优化后, 将单机数据库中的零售户订单多维模型数据同步到 MPP 架构数据库, 并初始化中间层汇总数据, 统计从零售户订单多维模型数据处

理成为指标数据的运行耗时, 验证零售户数指标计算性能。分别统计优化前在单机数据库上使用原有业务处理流程的运行耗时和在 MPP 数据库上使用优化后业务处理流程的运行耗时 (两者均计算 1 年的零售户订单量数据), 其计算耗时如下:

表 6 优化前单机数据库与优化后 MPP 数据库
计算零售户指标耗时
Tab. 6 Time consuming of calculating retailer indexes before and after optimization

统计指标	优化前耗时	优化后耗时
中间层数据	—	27 分钟
周零售户数	38 分钟	2 分钟
月零售户数	63 分钟	6 分钟
近 3 月零售户数	156 分钟	12 分钟
近 1 年零售户数	747 分钟	23 分钟
合计	1004 分钟	70 分钟

对比优化前后的数据同时耗时, 优化后的性能提升 90% 以上, 优化效果显著。

本文针对上海烟草数据中心零售户订单数据处理性能问题进行了全面的分析和研究, 给出了优化方法, 主要在数据库架构和数据处理业务两个方面, 从实施过程和验证的结果来看, 解决了上海烟草数据中心零售户订单数据处理性能问题。

通过上述的优化实施和验证, 在烟草零售户订单数据处理方面取得了一定效果并积累的较多的实施经验, 后续将在以下几个方面继续优化和完善:

(1) 基于已有的 MPP 数据库架构, 增加若干数据库节点, 测试性能提升与数据库节点的数量关系, 计算最优节点数。

(2) 进一步研究和优化数据处理业务, 提升零售户订单处理性能。

(3) 基于上述研究成果, 在上海烟草数据中心内形成统一的海量数据处理技术标准。

参考文献

[1] 张雨, 蔡鑫, 李爱民, 等. 分布式文件系统与 MPP 数据库的混搭架构在电信大数据平台中的应用 [J] 电信科学, 2013(11): 12-16.
ZHANG Yu, CAI Xin, LI Aimin, et al. Application of distribute file system & MPP database mashup architecture in telecom big data platform[J], Telecommunications Science, 2013(11):12-16.
[2] 孙淳晔, 梁杨. 探秘 MPP 数据库 [J] 互联网天地, 2015(7):73-76.
SUN Chunye, LIANG Yang. Research on massively parallel processor database[J].China Internet, 2015(7):73-76.

- [3] 王格芳. 某金融系统海量数据并行处理架构优化设计与实现 [D]. 中国科学院大学: 工程管理与信息技术学院, 2014.
- [4] 张岩, 刘晓芸, 马玉洁. DB2 数据库分区在商业银行数据仓库中的应用分析 [J]. 信息技术与信息化, 2017(3):55-58.
ZHANG Yan, LIU Xiaoyun, MA Yujie. Application analysis of DB2 database partition feature in commercial bank data warehouse[J]. Information Technology and Informatization, 2017(3):55-58.
- [5] 刘姝. DB2 数据库设计及优化技术研究 [J]. 信息安全与技术, 2011(11):38-40.
LIU Shu. The Design and optimization of DB2 database[J]. Information Security and Technology, 2011,11:38-40.

Architecture optimization of cigarette retail database in Shanghai tobacco data center

CHEN Deli

Information Center, Shanghai Tobacco Group Co., Ltd., Shanghai 200082, China

Abstract: To break through the bottleneck in processing tobacco retailers order data, this paper introduces the idea of massive parallel processing(MPP). The architecture of database for storing tobacco retail data in Shanghai tobacco data center is redesigned. Through deploying database with MPP architecture, optimizing data processing, and improving data processing performance, the data processing efficiency is improved, from more than 10 hours before optimization to less than 2 hours after optimization.

Keywords: database, massive parallel processing, performance optimization

*Corresponding author. Email: chendl@sh.tobacco.com.cn

(上接第 69 页)

Relationships of SPAD value with tissue structure, physiological index and internal quality of flue-cured tobacco upper leaves

SUN Guangwei¹, CHEN Zhenguo¹, WANG Yujun², SUN Jingguo¹, FENG Ji¹,

YU Tao², LI Jianping¹, QIN Guangjiong¹

¹ Hubei Provincial Tobacco Research Institute, Wuhan 430030, China;

² Shandong Agriculture University, Tai'an Shandong 271018, China

Abstract: In order to further quantify the relationship between maturity and quality of flue-cured tobacco upper leaves. Yunyan 87 was used in this experiment. Maturity of upper leaves of Yunyan 87 was classified by SPAD value, and tissue structure, physiological indexes of fresh tobacco and quality of flue-cured tobacco were analyzed. Results showed that the average SPAD value of upper leaves showed a positive linear correlation with free water/bound water and protein content, a negative linear correlation with spongy tissue/palisade tissue ratio, a positive linear correlation with TSNA content present, and a negative linear correlation with polyphenol content and nicotine content. When SPAD values were lower than 17, leaf senescence was obvious, carbon and nitrogen metabolism significantly decreased, antioxidant enzyme activities such as POD, SOD, CAT significantly increased, content of MDA and H₂O₂ significantly increased. The optimum SPAD of upper leaves in southwest Hubei province was 13 ± 1.5 .

Keywords: flue-cured tobacco; upper leaves; SPAD value; maturity; polyphenol; nicotine; TSNA

*Corresponding author. Email: hbskys1@163.com