

基于规约编码和缓存的智能用电数据处理技术

陶晓峰

(国网电力科学研究院 / 南京南瑞集团公司, 南京 210061)

摘 要:针对智能用电信息采集系统的数据集中采集、数据项和用户数量大的特点,提出了一种基于编码的算法并结合缓存技术的批量数据入库的新技术。先对采集的数据项进行预编码,运用 P/T 映射技术将数据项和存储字段关联并映射到缓存中,通过内存寻址方式先缓存数据,再定时批量提交到数据库,减少了数据库的操作频率,提高了系统的数据入库效率和稳定性,介绍了其中的主要功能模块以及关键技术的实现。

关键词:编码;缓存;映射技术;智能用电

中图分类号:TP311

文献标志码:B

文章编号:1009-0665(2012)01-0056-02

为了大幅提升用户用电体验,实现电力用户用电信息采集系统建设“全覆盖、全采集、全预付费”的总目标,对于大型智能用电信息集成采集系统来说,其要求采集的数据项和用户数越来越多,采集频率也将越来越快,对采集系统的采集模块提出了更高的要求。系统的数据源主要存在以下 2 种对采集造成影响的情况:一是需要采集的数据项可能会分别存到不同的数据表中,二是这些分项数据项并不是在同一时间片上采集的。如果每采集 1 个数据项都做 1 次数据库操作会增加数据库读写频率,从而造成数据库性能下降。因此设计高效数据处理技术对采集效率和系统性能提升都有重要作用。

1 采集数据项编码体系架构

1.1 采集数据项概述

根据 Q/GDW376.1—2009^[1]的功能规范,智能用电采集系统的采集数据项非常多,规范中定义的一类数据有 22 组共 170 多项,二类数据有 28 组共 200 多项。几乎覆盖了所有低压抄表和负控的所有数据项。为了满足智能电网的应用要求,对于一些重要数据项需要实现高频度的实时采集。

1.2 数据项编码和处理体系架构

现设计一种将采集对应的数据项通过一定编码规则产生对应的编码项方法,通过配置将编码项关联到该数据项需要存储的表和字段,然后将这些配置信息形成相应的缓存结构,采集的数据分别对应到不同的缓存结构,最后由入库模块通过解析将数据批量入库。这样就能大大提高入库效率。具体技术架构如图 1 所示。

1.3 编码机制

为了减少数据解析模块和数据库入库模块之间的交互,系统设计了一种通用的数据项编码方

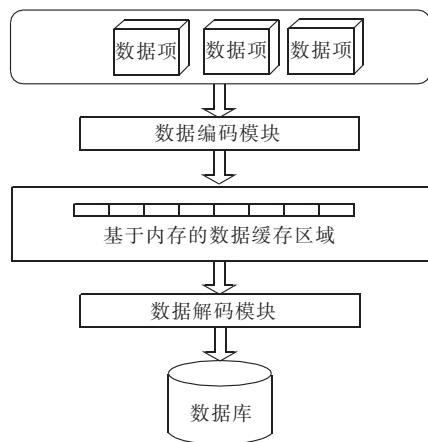


图 1 基于编码和缓存的数据处理架构

法,即将每个数据项根据其对应的规约、数据属性、数据项类型等生成唯一的特征编码,在系统中通过配置将这些数据项编码和数据库中的存储字段相关联,减少数据库表跟程序代码的耦合,提高系统灵活性,各模块之间的数据交互也采用这种编码进行交互,有效提高系统数据的传输效率和处理速度。

2 编码数据缓存方法

2.1 P/T 映射技术

P/T 映射技术是用来解决多数据项和跨表存储数据模型的方案,将数据的编码项和数据库表的具体字段建立相应的对应关系,目的就是将应用层和数据库层进行解耦,达到数据库存储结构有变化可以不用反映到应用层上。使用 P/T 映射技术可以通过在内存中对数据模型的存储和重构,一方面可以实现数据项和数据模型的映射关系,另一方面也解决了数据模型映射的可扩展性,对以后增加的数据项和数据模型可以通过 P/T 映射技术实现自动映射,开发人员不必考虑具体的数据项存在哪张数据表的某个字段,大大简化了程序的编写工作,提高了系统的可维护性、可扩展性和可重用性^[2]。

2.2 数据缓存原理

针对智能用电采集系统的数据采集特点:数据项多、用户数多、采集时间比较离散等。系统的批量数据入库处理模块,采用了基于数据项编码的数据缓存机制,其原理类似于内存寻址方式,数据入库模块将数据表结构映射到内存中,保存相应的数据索引和数据类型等基本信息。由数据编码模块将采集来的数据转化成编码项和数值,根据编码项通过索引算法算出偏移量,找到该数值存储的数据区,最后通过写数据操作把数值存到对应的数据区。数据入库模块可以根据索引和数据类型等基本信息迅速地通过读数据操作将数据值转化为对应的存储结构,从而提高数据库入库效率和减少数据库的IO操作次数。具体原理如图2所示。

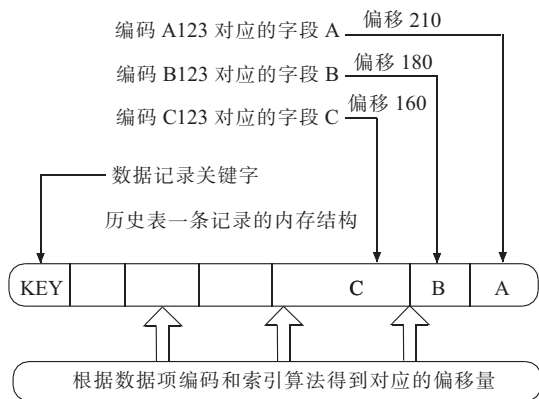


图2 基于编码的数据缓存原理

3 数据入库优化

数据库是信息系统的核心部分,高效、安全和可扩展的数据入库模块是系统的关键。由于每个 SQL 语句都要执行很多次,尽量让该 SQL 的散列值在内存块中存在,数据入库模块先对数据库的数据表的信息预先进行 SQL 的初始化拼装,避免 SQL 的硬解析,这样有利于提高数据库的 SQL 语句的执行效率^[3]。由入库线程定时通过数据解码模块,解析出不同数据表的批量数据记录,先将数据记录插入到数据库的临时表中,再通过 MERGE 语句进行批量入库,其比传统的先判断再选择插入或更新的方式要快很多,而且易于使用,提高了数据入库效率。

4 数据入库容错机制

为了提高系统运行的稳定性,数据入库模块提供了一个基于 LOCAL 缓存的容错机制,主要针对以下2个方面:一是针对用电采集系统数据流量突发性增长情况,即在某个很短的时间段可能会有超过系统设计正常处理能力的情景;二是如果数据库临时出现故障时,系统能尽量减少停运时间。为此,

设计了 LOCAL 缓存机制,先把数据通过文件先缓存到本地,作为虚拟内存的延伸。定时由轮询处理线程对数据库状态进行监测,如果出现数据库发生故障或者缓存数据库记录太多的情况,数据入库容错模块把缓存数据按一定的格式进行序列化,序列化后的数据按时间和数据表分类生成数据文档存储在本采集服务器上^[4]。一旦数据库恢复后数据文档处理模块将数据文档进行反序列化,反序列化后的数据提交给数据入库线程批量入库^[5]。这样出现上述情况时,就能尽可能地避免系统停机,保证采集服务器正常运行;系统还能正常进行数据采集,不会出现一些实时数据丢失的情况,提高了系统的可用率。处理流程如图3所示。

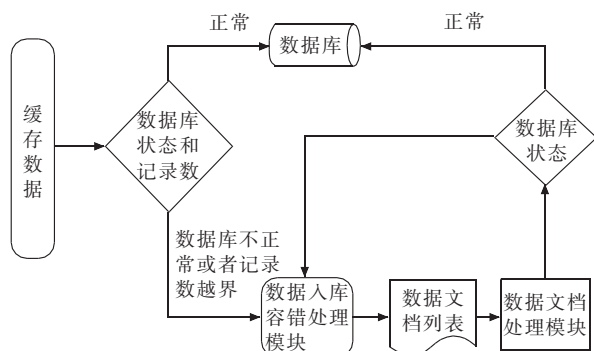


图3 数据入库容错机制流程

5 结束语

该技术针对智能用电信息采集系统的通信规约数据项多、用户数特别庞大的特点,提出了数据项编码化的思想,并运用数据缓存预处理技术,大大提高了数据处理效率。现场实际投运后发现处理能力有了明显提高,对整个智能用电信息采集系统的数据稳定采集和入库起到了关键作用。

参考文献:

- [1] Q/GDW376.1—2009, 电力用户用电信息采集系统通信协议:主站与采集终端通信协议[S].
- [2] 曹晋彰,朱传柏,刘波,等.基于公共信息模型的电网智能调度编码体系研究与实现[J].电力系统自动化,2011(2):43-47.
- [3] SHASHA D,BONNET P.数据库性能调优:原理与技术[M].孟小峰,李战怀,等译.北京:电子工业出版社,2004.
- [4] 周功业,吴伟杰,陈进才.一种基于对象存储系统的元数据缓存实现方法[J].计算机科学,2007(10):146-148.
- [5] 卢成均.缓存机制及其在数据存取层中的应用模型研究[J].计算机应用与软件,2008(12):172-174.

作者简介:

陶晓峰(1977),男,江苏张家港人,工程师,从事智能用电采集系统的设计和开发工作。

(下转第60页)

库户端一般都用 Windows 系统,需安装 Java1.4 版本、Tomca 和用 JAVA 控件新增告警窗程序。

4.1.1 Windows 下安装 Java1.4 版本

Windows 操作系统一般都没有安装 java, 或安装的版本不是 1.4, 推荐全新安装 Java1.4。安装步骤如下: 从版本库中 open2000e_tool\web_software\java\windows 处获取安装程序 j2sdk1_4_2_08-windows-i586-p.exe; 双击安装程序弹出安装界面, 安装过程比较简单, 基本选择默认设置, 一路 next 下去即可完成安装, 默认安装在 c:\j2sdk1.4.2_08 下, 也可以改变安装路径; 安装完成后需要新建 JAVA_HOME 环境变量并修改 Path 变量。

4.1.2 Windows 下安装 Tomcat

从常州调控系统 open3000 版本库 open2000e_tool\web_software\tomcat 下获取 tomcat.tar 安装包。把 tomcat.tar 解压到 D:\ 目录下, tar xvf tomcat.tar (PC 上 open2000e\bin 目录下也有 tar.exe 命令, 不建议使用 WinRAR 等解压) 新建 CATALINA_HOME 环境变量并修改 Path 变量。

4.1.3 JAVA 控件新增告警窗程序

借助公司原有 III 区调度 SCADA 网页浏览系统, 用 JAVA 控件新增告警窗程序。用户只需重新打开一次调度 SCADA 网页, 让页面缓存重新读取新的控件, 将设置文件拷贝至本机指定目录即可正常浏览。

4.2 按责任区划分

根据需求侧重点的不同, 将告警窗显示内容按行政区域(监控中心及各个操作班、金坛、溧阳)、告警来源、告警类型等进行分类。通过责任区划分的方式, 让每个部门进入相应的责任区进行操作, 告警窗中只显示该部门所关心的相关信号告警动作。

5 结束语

2010 年 8 月 18 日, 公司原属于变电工区的监控中心正式并入调度中心, 因此操作班和监控中心已不属于同一个部门。变电运行人员想获知现场告警信息, 本需要在变电运行操作班增加 OPEN3000 工作站。电网跨区实时告警功能的设计与实现后, 变电所各个操作班要了解变电所运行情况时, 只需在普通 PC 机上进行插件, 就可以在 III 区进行浏览。至此, 变电所的投运调试工作也变得更为便捷, 而且 I 区的网络安全也有了重要的保障。

参考文献:

- [1] 程汉湘. 电力电子技术[M]. 北京: 科学出版社, 2007.
- [2] 李天阳, 郭剑虹. 集中中心告警数据处理技术和方法[J]. 电力系统自动化, 2010(22): 113-116.

作者简介:

杨 菲(1980), 女, 江苏常州人, 工程师, 从事电力自动化运行与维护工作。

Design and Realization of Trans-regional Real-time Alarm Function of Power Grid

YANG Fei

(Changzhou Power Supply Company, Changzhou 213003, China)

Abstract: Trans-regional real-time alarm function of power grid achieves the application expansion of alarm service of dispatching & monitoring system in zone III. Based on the research of the functions of balance and synchronization of trans-regional real-time alarm and the client machine installation, real-time alarm data of substations can be released to the Internet by Webpage upload method, and which can be browsed and inquired by legal LAN clients. Above of these make substation's operation and debugging easier, and ensure the network safety of zone I.

Key words: alarm function; webpage; OPEN3000; trans-region

(上接第 57 页)

Smart Electricity Consumption Data Processing Technology Based on Protocol of Encoding and Cache

TAO Xiao-feng

(State Grid Electric Power Research Institute, Nanjing 210061, China)

Abstract: According to the characteristics of smart electricity consumption data acquisition system: centralized data acquisition, large quantity of data items and customers, a new batch data loader technology is proposed which is based on coding algorithm and cache technology. The acquisition data items are precoded and then mapped into the cache data area using P / T data mapping technology, and batch submitted to the database at scheduled time to reduce the operation frequency of the database. The algorithm greatly enhances the system's data storage efficiency and stability. Finally, the main function modules and the realization of key technologies are introduced.

Key words: encode; cache; mapping technology; smart electricity consumption