

能量管理系统的实时数据库技术

陆杏全

(国电南瑞科技股份有限公司, 江苏省南京市 210003)

摘要: 分析了能量管理系统(EMS)中实时数据库技术的现状以及影响实时数据库技术发展的原因,分析了 EMS 的应用环境特点,对设计开发符合 IEC 61970 EMS-API 标准的新一代 EMS 的实时数据库管理系统提出了一些观点,特别是如何提高数据库访问效率,使之满足 EMS 中实时应用的要求。

关键词: 能量管理系统; 数据库; 数据库管理系统; 实时数据库; 面向对象

中图分类号: TM73; TP311.13

0 引言

数据库管理系统(DBMS)是能量管理系统(EMS)支持软件平台的最主要成分,这一观点已得到 EMS 厂家和用户的广泛认同,但关于怎样的 DBMS 才符合 EMS 应用实时性要求,即实时 DBMS 问题却一直没有得到很好的解决。通用的关系型数据库管理系统(RDBMS)产品出现后,其通用、灵活的特点和诸多优良品质赢得了很高声誉,在 EMS 领域之外获得了广泛应用。国内外一些 EMS 厂家也试图直接采用 RDBMS 支持 EMS,但结果并不理想。IEEE 和 CIGRE 早些年成立过有关工作小组,针对此问题在国外 EMS 厂家和用户中进行过大规模的调查研究,结果认为不能用 RDBMS 代替 EMS 中的实时 DBMS,原因在于 RDBMS 操纵数据的速度慢,不能满足大多数 EMS 应用功能(通常被称为实时应用和准实时应用)对信息处理和控制响应速度的要求;同时,客户/服务器(C/S)模式对于分布式 EMS 而言还存在服务器延迟和网络通信延迟问题。尽管这是几年前的调查结果,其后 RDBMS 产品在不断升版,硬件(CPU 和网络)速度也大大提高,但实测表明,RDBMS 每秒的事务处理量仍然不超过 10^3 数量级。服务器(软件)延迟和网络通信延迟问题可能随着计算机速度和网络速度的提高而得到一定程度的缓解,但问题依然存在,甚至有时造成的延缓非常明显,而且没有明确的规律可循。因此用 RDBMS 支持 EMS 可能使许多应用功能的速度指标远不能令人满意,甚至达不到有关的国家/行业标准所规定的指标(实际上其中多数指标也是多年前制定的,现在看来要求很低)。

因此,适合开放、分布式 EMS 的实时数据库技术远未达到成熟的地步,仍值得国内外 EMS 厂家

和用户关注。近年来国际电工委员会(IEC)正在制定的 IEC 61970 EMS-API 系列标准(草案)公布后^[1],关于 EMS 如何满足 61970 标准提出的公共信息模型(CIM)和组件接口规范(CIS)要求,从而实现真正开放的问题引起 EMS 厂家的广泛重视。虽然 IEC 声明,发布 61970 标准的基本目的是为了减少向(投运后的)EMS 中增加新应用所需要的费用和时间,为了保护对(投运后的)EMS 中正在有效工作的应用(软件,可能有硬件)的投资,减少不同厂家生产的系统之间实现互联所需要的费用和时间,并且强调 CIM 只是一个数据模式,不是数据库。也就是说,IEC 61970 标准涉及的只是数据模式和接口问题,没有规定用什么样的 DBMS 去实现这种数据模式。但国内的 EMS 厂家,不论其原有的 EMS 产品如何实现,都受此标准的触动,开始认真考虑改进其实时数据库技术,使其产品在实现 IEC 61970 标准的基础上做到真正的开放。为此,本文根据笔者多年来从事 EMS 应用和实时数据库技术研究的经历,以及对 IEC 61970 EMS-API 标准的认识,在分析影响实时数据库技术发展的因素和 EMS 应用环境特点的基础上,就新一代 EMS 的实时 DBMS 实现方法提出一些观点,重点将放在如何提高效率方面。

1 影响实时数据库技术发展的因素

1.1 EMS 开发方存在的认识误区

a. 虽然 EMS 厂家意识到实时 DBMS 对其 EMS 产品非常重要,却总认为自己是搞 EMS 集成的,主要力量应放在接口和功能实现上,而实时 DBMS 只是整个 EMS 的一个“专用”的“后台角色”,不可能成为一个单独的商品软件;而商用数据库技术已经非常完美,希望能走 RDBMS 的捷径,因此在这方面投入的开发力量不够。

b. 有的厂家曾经为此付出了不少努力,但投入的开发力量中太偏重于计算机专业,对 EMS 应用数据库体系和应用特点不甚了解(甚至认为只有 SCADA 数据库才是实时数据库),又没有关于实时 DBMS 的资料可参考,其开发结果往往过于受 RDBMS 教材的束缚,太拘泥于经典的“客户/服务器(C/S)”模式。不少 EMS 厂家采用了所谓“基于 RDBMS 的实时数据库技术”,即在 RDBMS 定义和生成的关系数据库基础上,结合一些实现方法不同的加速处理(如基于 cache 技术的内存共享或基于 hash 表的内存共享),作为 EMS 的支持工具。这种做法本身就反映了他们对 EMS 中实时数据库概念的误解,这样的“实时数据库”无论是数据模式、基本管理模式、访问方法还是定义界面,都离不开 RDBMS,难免在一定程度上影响 EMS 应用系统的使用效果,至少对具有实时或准实时处理要求的 EMS 高层应用的开发和维护带来不便,因为支持这些应用的数据库也应该是实时数据库。

1.2 EMS 用户方存在的认识误区

a. 长期以来在其技术规范书中只提到“实时数据库”(并且大多停留在实时数据概念上),到了近年来才提及 DBMS(也仅仅是概念性提及而已),用户更关注的是 EMS 应用功能和人机界面。即使到现在,也往往只认可接口软件和应用软件的价格,很少认可厂家对其实时 DBMS(或者整个支持平台)的单独报价。而且在评价人机界面的特性时,较多的机会是“以貌取人”,很少关注其“后台角色”的质地如何。这些因素有可能影响 EMS 厂家精心开发实时 DBMS 的积极性。

b. 在购买 EMS 时可能要求厂家承诺其提供的产品具有支持“二次应用开发”的能力,但是太笼统,对于其中很关键的实时 DBMS 没有很具体的要求(例如支持到什么程度、灵活性、方便性、技术指标等)。近年来,用户方可能在文字上提及所谓的“应用程序接口”,但他们(绝大多数)并不亲自(或让第三方)尝试。这样,厂家不能获取来自用户(或第三方)的关于其实时数据库技术方面的信息反馈,从而反过来也影响了实时数据库技术的发展。

2 EMS 应用环境的特点

就一般 DBMS 而言,安全性、完整性、可靠性、保密性、灵活性(适应力)、服务功能、用户界面、并发能力、访问效率等都在性能指标之列。但事实上,有些性能本身可能是相互制约的,可能需要根据不同的应用场合,在各项性能要求间做出权衡。在一个 EMS 中,除了一部分历史数据应用功能外,使用人员要求大多数应用软件具有很高的信息处理和控制

响应速度。其中 SCADA 类应用(如数据采集、状态监视、告警处理、事故追忆等)是高频率访问数据库和对电力系统事件快速响应的最典型例子,特别是当电力系统受到突发事件的干扰时,短时间内要接收、记录、处理、报告大量的事故/事件,因此这类应用对快速存取数据的要求是不言而喻的;另一类应用(如在线拓扑分析、实时状态估计、在线潮流、仿真软件等)则访问量较大和计算量大,为了争取处理过程在相当短的时间内完成,要求尽量减少访问数据库所花费的时间。

以 EMS 中最基本的高级应用“实时状态估计”软件为例,对于 300 个计算节点的电网(相当于一个中等规模地区电网的 10 kV 以上部分)而言,按照现在的计算机计算速度,状态估计所需计算时间(不含数据库访问时间)应小于 0.5 s,如果数据库访问速度能达到 500 万次/s 以上,那么完成状态估计全过程的时间将接近 1 s(即一半时间花在数据库访问上);如果数据库访问速度只有 10 万次/s,则完成状态估计全过程的时间将超过 5 s(即 90% 以上的时间花在数据库访问上),已经使这个软件丧失了实时性。由此可见,访问效率这一性能指标对 EMS 的 DBMS 而言显得特别突出。

另一方面,如果把一个 EMS 中的每个实时/准实时应用看成一个客户,那么总的客户数量不多,而且总体上是相当稳定的;每个客户要完成的工作也是特定、有规律的;客户除了对公共数据库有读共享要求外,几乎没有对同一数据单元的并发写操作(让各个应用软件操作各自的应用数据库)。也就是说,EMS 的实时应用环境远没有银行、电信、证券、票务等信息处理自动化系统那样复杂(事实上 EMS 还不向外界开放);实时应用数据库也远没有那些系统的数据库那样庞大;而且一个实时应用数据库很可能只是为某一个应用功能“专门”设计和使用的,其作用范围小,初始化过程相当简单。相反,电力系统是不断变化的,用户还可能对已有实时应用提出新的数据处理要求(新功能),因此可能经常在线改变应用数据库结构。此外,EMS 应用软件操作数据库的过程更多的是“零售方式”(频繁交互)而不是“批发方式”(事务处理);用户对 EMS 的图形化界面(实际上也可以看成是应用数据库的表达和操作界面)的功能要求又往往比那些系统更高。所以,对实时 DBMS 各方面性能的具体要求与 RDBMS 产品相比,不应该“一视同仁”。

3 新一代 EMS 的实时 DBMS

3.1 面向对象的数据模式

IEC 61970 EMS-API 标准的公共信息模型

(CIM)是采用面向对象的形式描述的,又规定了组件接口规范(CIS)。结合面向对象技术的发展现状,使人很自然地设想采用面向对象技术去解决问题,实时数据库也采用对象模式,这样有利于 EMS 更加开放,实现 EMS 应用软件“即插即用”。由于各种各样的原因,迄今国内外 EMS 厂家还没有成功使用面向对象 DBMS 支持 EMS 的例子。有些 EMS 公司的实时 DBMS 可以支持混合型数据结构,在某些方面带有面向对象性质,效率很高,应用程序接口和其他工具功能也很强,但本质上只能算是扩展的层次型而不是对象型 DBMS。因此所有的 EMS 厂家可能都面临这样的问题,即在不长的时间里开发新的或改造其原有的实时 DBMS,使之支持对象模式,同时不失高效率和其他必备的性能^[2]。

这里所说的“支持对象模式”并不等于建议使用商业化的面向对象 DBMS,也不等于建议厂家完全按商业化面向对象 DBMS 教科书进行开发。开发者可能设计一个支持纯对象模式的产品,也可能做成一个支持异构模式的 DBMS(包括支持对象模式)。

3.2 数据库的物理存储模式

对于 EMS 实时数据库(规模不大、变动频繁)而言,实现那些理论上优化的模式所花的开发代价与由此带来的实际效果相比是得不偿失的,更何况还有效率降低的负面影响。除非已有现成的程序,否则开发者选用简单的类文件模式(最简单的做法是一个类对应一个二进制文件)比较妥当,正因为采用类文件方法简单,所以易于实现、管理,并且可靠、高效。

3.3 提高效率的途径

一般的数据库系统中,客户程序通过接口调用访问数据库,不必关心数据库的逻辑结构,甚至不关心数据库在什么地方,服务软件通过搜索向客户程序提供它需要的信息。设计 DBMS 时对数据库的物理存储(数据体)、逻辑结构描述以及接口(服务)可能会考虑许多优化因素,如存储空间利用、各种加快(相对而言)查询的措施等。但这种方法通常是低效率的,因为它基于查询服务,只适合于“批发”式应用场合。笔者认为可以采用以下几种方法提高数据库访问效率使之达到“实时”的要求。

a. 指针的使用^[3,4]

指针概念不是层次型数据库独有的,面向对象型数据库中也大量存在(实际上面向对象数据库也包含层次型机制)。指针可以用来表达类间(具体讲是对象实例间)的关联(关系),例如表达不同类对象之间 1:1 关联、1:n 关联、m:n 关联的指针和同类对象之间链接的指针等。指针充分利用了静态查询的结果(例如数据库“装入-验证”阶段),提高了

应用程序访问数据库的效率。合理定义和使用指针,不仅使对象(属性数据)之间的相互联系表达得更清楚,而且可大大加快查询速度;但指针应该由 DBMS 自动生成和动态维护,而且不妨将指针设计成对象的一种“隐式属性”,这将对应用软件的功能实现和日常维护提供很大的便利。

b. “避免”磁盘访问^[3~5]

磁盘 I/O 相对于 CPU 处理速度而言很慢,所以在使用驻磁盘的数据库时应该映射到内存使之变为“驻内存”数据库(不锁定,可以是共享方式)。目前有的 EMS 产品的实时 DBMS 采用这种方法,效果很好。这里所说的“驻内存”数据库与早期 SCADA 系统采用过的常驻内存型数据库完全不同,那种数据库在系统引导阶段就装入内存,没有可编辑能力,不能在线改变结构(任何数据库结构改变都将导致系统重新引导),无法适应现代的 EMS 应用(特别是高级应用)。同时,这里所说的“驻内存”数据库与现有的“基于 RDBMS 的实时数据库”也完全不同,因为后者的本质是常驻内存的,也不能在线改变结构。

c. “避免”网络通信,即改“异地访问”为“当地访问”

由于异地访问时服务器延迟和通信延迟造成效率降低,有些厂家的公共实时数据库采用了多副本方法,使应用软件所在地驻有一份副本,并在应用软件所在地建立该应用“专有”的应用数据库,实现应用软件当地访问公共实时数据库和自身应用数据库。此时 DBMS 的最高管理层的主要精力是如何保持公共数据库副本与其“原本”之间的同步,而不是响应客户,因此网络通信工作量体现在服务器层而不是应用层。这样处理可以使数据库访问效率提高至少 1 个数量级。

d. “越过”服务器(软件)

即使实现了数据库“驻内存”和当地访问,如果仍然一律采用 C/S 模式,两阶段提交,由于还存在(当地)服务器延迟问题,所以提高效率有限;假如采用不经过服务器的函数接口方式,则可以使访问效率再提高约 1 个数量级(即在当代计算机上,访问速度达到 10 万次/s 以上)。

e. “越过”逻辑层^[3,4]

为了实现更高的速度,可以提供特别的接口(函数),使应用软件越过数据库的逻辑层直达物理层,或者说让应用软件的数据空间与数据库空间实现规范化的地址映射,这样可以使效率再提高 2 个数量级,即访问速度达到 1 000 万次/s 以上。

表面上看这样似乎使应用程序与数据库数据结构之间的联系越来越紧密,但这种紧密不等于简单

地“捆绑”。将应用程序与数据库数据结构“捆绑”在一起带来的弊端是当数据结构(包括公共数据库和专用数据库)改变后,应用程序需要重新编译连接,因此极不灵活,当应用软件模块很多时还带来维护方面的困难,特别在应用软件开发阶段,数据库结构变更较频繁,矛盾尤其突出。我们需要的是一种“自适应的紧密联系”,使应用软件效率高的同时,又不会被数据库数据结构束缚手脚,这也是使用“越过”逻辑层方法时实时 DBMS 必须解决的问题。这一点与 EMS 应该是“松耦合”系统^[2]的观点并不矛盾,所谓松耦合主要是指应用与应用之间的关系,按面向对象的概念,每个应用作为一个对象具有自治性,使 EMS 具有良好的可扩展性,应用对象“插入”、“卸出”很方便,系统剪裁很灵活。如果把应用数据库和应用程序都看成是同一个应用对象的一部分(一个看成属性,一个看成行为),那么应用程序与应用数据库的联系适当地紧密一些是合理的,更何况除了效率提高外没有什么“毒副作用”。

3.4 对其他性能的考虑

如上所述,对于实时 DBMS 的各方面性能,应该考虑 EMS 的环境特点。例如关于安全性,假如像上面所说的,应用数据库运行在非常单纯的环境里(出于 EMS 本身安全的要求,它并不向外界开放),一个数据库只是某个应用的“专用品”,能造成其破坏的主要还是该应用本身(并且可能只在该应用的调试阶段),该数据库即使被破坏也只是暂时影响该应用软件,或者可以用很简单的方法加以恢复(例如有一个备份即可),那么在开发实时 DBMS 的初期可以不必在各种防止被破坏的手段或者被破坏后的恢复手段等方面花太大的工夫;同样,对于并发访问问题,假如应用系统中不存在或很少存在这类问题,或者即使发生也不造成严重后果,那么在开发的初期也可大大简化关于锁的机制(涉及锁的粒度、上锁策略、优先级、队列管理、如何解死锁等一系列问题)。笔者用这个例子并不表明对实时数据库安全性的轻视,只是借此说明一种策略,关于各种性能要求都应区别轻重缓急,权衡处置。总之,实时 DBMS 使

用场合不同于 RDBMS,不要按与 RDBMS 类似的思路去设计开发实时 DBMS,否则还不如买一个 RDBMS 合算,因为毕竟没有哪个 EMS 厂家在这个领域的技术力量足以跟商用数据库公司抗衡。

4 结语

EMS 的实时数据库技术还不很完善,IEC 61970 标准(草案)发布后,许多 EMS 厂家把这个已被“搁置”(因为他们认为已经解决了)的问题重新提到议事日程上来。本文在分析 EMS 应用环境特点的基础上,对设计开发实时 DBMS 的一些技术问题提出了自己的看法,重点介绍了如何提高数据库效率,因为笔者认为这是实时 DBMS 区别于其他 DBMS 的根本点。

参考文献

- 1 IEC 61970 EMS-API系列标准(IEC 61970 EMS-API Standard). 2002
- 2 陆杏全(Lu Xingquan). 电力调度集成化信息系统(Integrated System for Power Dispatch Automation). 见:中国电机工程学会 2001 电力信息技术应用与发展学术研讨会(In: CSEE Conference on Power System Information Technology Application and Development). 杭州(Hangzhou); 2001
- 3 陆杏全,任洛卿(Lu Xingquan, Ren Luoqing). EMS 应用数据库管理软件包(EMS Application Database Manager). 电力系统自动化(Automation of Electric Power Systems), 1992, 16(2)
- 4 陆杏全,任群,鲍伟峰(Lu Xingquan, Ren Qun, Bao Weifeng). EMS 应用数据库管理系统 Corona (Corona, A Database Management System for EMS Applications). 见:中国电机工程学会电力系统与电网技术学术年会论文集(In: CSEE Conference on Power System and Power Technology). 宜昌(Yichang); 1999
- 5 张慎明,卜凡强,姚建国,等(Zhang Shenming, Bu Fanqiang, Yao Jianguo, et al). 遵循 IEC 61970 标准的实时数据库管理系统(Real-time Database Management System That Conforms To IEC 61970 Standard). 电力系统自动化(Automation of Electric Power Systems), 2002, 24(24): 26~30

陆杏全(1945—),男,硕士,高级工程师,主要研究方向为 EMS 高级应用软件和数据库管理系统。E-mail: lxq@nari-china.com

REAL-TIME DATABASE TECHNIQUE IN ENERGY MANAGEMENT SYSTEM

Lu Xingquan

(NARI Technology Development Co, Nanjing 210003, China)

Abstract: The current situation of real-time database technique in energy management systems (EMSs) and the factors which have been affecting the development of this technique are analyzed. The characteristics of application environment in an EMS are described. Some key points of view in the development of real-time database management system for the new generation EMS which conforms to the IEC 61970 EMS-API standard are proposed, especially for the technique relating to upgrading of database access efficiency.

Key words: energy management system; database; database management system; real-time database; object-oriented