

LAMOST 控制系统实时分布式数据库

徐 灵 哲 徐 欣 圻

(中国科学院国家天文台南京天文光学技术研究所 南京 210042)

摘 要

介绍国家“九五”重大科学工程项目“大天区面积多目标光纤光谱天文望远镜 (LAMOST)”中基于 QNX (Quick UNIX) 操作系统的实时分布式数据库的研究情况。此项研制已通过 2002 年 7 月 2 日由 LAMOST 工程指挥部组织的北京、南京和合肥三地专家的评审, 并得到好评。该数据库功能强、工具丰富、操作方便, 在类 UNIX 操作系统下实现了 Windows 的界面风格, 可以满足 LAMOST 控制系统数据在线 (或离线) 处理、存储、程序生成表格和图形等需要, 也可推广应用到其它相关领域。

关 键 词 天文观测设备与观测资料 — 实时分布式数据库 — 控制系统 — LAMOST

分 类 号 P111.2

1 引 言

国家重大科学工程项目“大天区面积多目标光纤光谱天文望远镜 (LAMOST, Large sky Area Multi-Object fiber Spectroscopic Telescope)”^[1]的结构设计别树一帜, 它将成为世界上 4 m 级口径以上的光学望远镜中视场最大、光谱观测效率最高的望远镜。如此高的目标给望远镜控制系统的设计^[2]带来了很大的挑战, 而其中一项挑战就是数据在线处理和实时运作。为此, LAMOST 望远镜总控研制组开发了基于 QNX (Quick UNIX) 操作系统的实时分布式数据库平台。该项研制与当代国际先进大天文望远镜接轨, 在国内大望远镜的使用上尚属首次, 具有开创性意义。

2 LAMOST 望远镜实时分布式数据库软件包设计

国家大科学工程项目 (0983) 资助课题

特约稿 2003-03-31 收到 2003-06-17 收到修定稿

2.1 概念的产生

纵览当前国际上先进的大天文望远镜,人们不难发现,为方便调试和获取高质量的天文目标数据,在线质量分析、望远镜各种状态的历史记录、仪器故障诊断、气象条件和观测室环境参数的长期监测采集等都是提高望远镜的使用率、增加科学产出的必要手段。所有这些数据如何处理是值得研究的问题。一个不争的事实是这些数据的处理,特别是在线质量分析,要求较强的实时性。鉴于天文望远镜观测的高精度和实时性,当代国际上几乎所有的大天文望远镜的控制都建立在网络实时分布式操作系统的环境下。而基于这种操作系统之上的实时分布式数据库则是当代天文光学望远镜控制的又一特色。借鉴国际先进经验,结合实际需要,LAMOST 实时分布式数据库系统的概念便应运而生。

2.2 系统开发平台

1 台 COMPAQ AP500 和 3 台 NEC PIII500 组成一个局域网模拟系统开发平台^[3]。每台计算机上都装有 QNX4.25 实时操作系统^[4]。其中在 1 台 NEC 机上还安装了 EMPRESS v8.60 数据库开发包(测试版)。目前国际软件市场上只有少数几个产品能支持在 QNX4.25 环境下开发分布式关系型数据库,EMPRESS 就是其中的一种。在 QNX 操作系统下 EMPRESS 能利用 QNX 的 Fleet 协议,而对于其它操作系统则提供 ODBC (Open DataBase Connectivity) 接口。LAMOST 实时分布式数据库就是在 QNX 环境下再利用 EMPRESS 软件包研制和开发出来的。

2.3 数据表格

在进行 LAMOST 实时分布式数据库系统设计时,我们考虑了 10 余张数据表格,如:机架指向和跟踪表(内含观测日期、时间、序号、目标信息、机架焦面位置读数和导星星像等重要参数);主动光学控制表(内含施密特校正板的 24 块子镜上的力促动器、位移促动器,以及主镜上的 37 块子镜上的位移促动器在观测过程中不断检测的读数);环境参数表(包括测量日期、时间、地点、所用仪器和环境参数值);联机帮助表(内含望远镜控制系统软件的序列号、升级号、主要应用软件模块的使用说明和演示例子等)等。

2.4 数据库分布设计

LAMOST 望远镜控制系统 TCS (Telescope Control System) 包括一个总控平台^[2]和 8 个子系统。这些子系统分别是方位驱动、高度驱动、校准镜主动光学控制、主镜主动光学控制、焦面控制、导星控制、圆顶控制和环境参数控制。总控平台和 8 个子系统分布运行在 4 台计算机上,互相之间采用进程通讯。为了减少网络间数据的流量,数据库的分布式设计采用了数据分布优化策略^[7]。对于机架指向和跟踪表、主动光学控制表和环境参数表,由于它们要存储大量的实时观测数据,所以分别把这 3 张表建立在与其相关的子系统节点上,这样可以减少大量的网络通讯量。对于联机帮助表则采用复制的方法在每个节点上复制一份,用户在调用帮助时只需使用本机的资源。至于其它表因为数据量小,修改不频繁,而且主要在工程维护时使用,因此和总控平台都放在一个节点上。这就是数据库的底层安排。应该指出,上述数据分布优化处理增强了系统的坚固性即对故障的免疫力,不至于“牵一发而动全身”,有利于系统的维护和扩充。至于对数据库的安全保护、数据库完整性保护和维护,则采用一定的配置,充分利用 EMPRESS 提供的功能自动完成。对于数据库中的数据目前我们采用增量转储的方法,这样万一数据被破坏也可以恢复。为了进一步确保将来 LAMOST 望远镜实时分布式数据

库的可靠性, 整个节点的冗余也在考虑之中。但由于目前平台上的计算机数量有限, 故暂时没有实施。

2.5 程序风格、结构和功能

考虑使用方便以及操作灵活, 鉴于 Windows 操作系统已被大部分人所熟悉, 本程序在类 UNIX 系统的 QNX 操作系统下, 通过大约 13000 多行 C 语言编程实现了标准的 Windows 风格。如支持鼠标、菜单和快捷键操作, 支持拖曳操作和支持多种分辨率等功能便可见一斑。

由于 EMPRESS 的测试版没有图形开发工具, QNX 操作系统的图形开发工具 PHAB (PHoton Application Build) 也无法直接调用 EMPRESS 测试版提供的库, 所以本程序采用了如图 1 所示的结构。其中界面进程的开发是借助 PHAB, 用 C 语言编制的。服务进程是用 QNX shell 下的 WATCOM C 开发的。它们之间的通讯则通过共享内存来实现。

LAMOST 望远镜实时分布式数据库软件包具有一系列功能, 如建表、修改表、删除表、建立索引、删除索引、运行 SQL (Structured Query Language) 语句、选择列条件产生表格、调整表格的列、调整表格的字体、调整表格的颜色、调整表格的大小、调整表格的宽度、选择查询条件、选择查询顺序、选择列条件产生图形、设置图形的横坐标和纵坐标、设置图形的题头、设置图形的缩放、插入数据和删除数据。这些功能基本上实现了常用的 DBMS (DataBase Management System) 功能。用户可以利用程序提供的 GUI (Graphical User Interfaces) 接口, 方便地建立自己的表格、图形, 并加以保存。

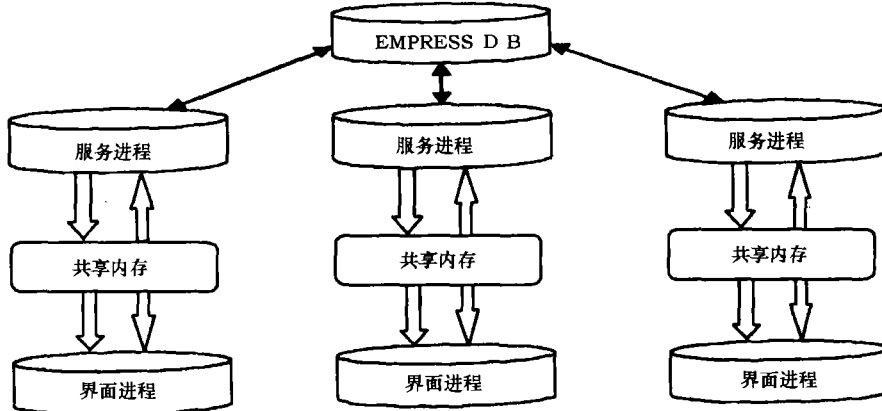


图 1 LAMOST 实时分布式数据库平台的系统结构

2.6 技术实现

(1) 数据库服务程序的实现

服务层负责接收界面进程发来的消息, 并对其进行处理, 然后同数据库交互。服务程序读取消息中的 SQL 指令, 根据指令中的关键字来判断 SQL 指令的类型。对于执行类型的 SQL 语句, 例如插入、删除、修改等, 由服务程序直接运行并得到状态量返回给界面进程; 对于查询类型的 SQL 指令, 则采用 SQL 语句中的动态 SQL 技术来实现。由于在查询之前无法确切知道所要查询的列的数量和列的类型, 只有通过动态 SQL 语句来实现。

(2) 进程间的通讯

界面进程和服务进程的交互采用 QNX 实时操作系统环境下的消息通讯方式,借助于 Send 和 Receive 函数通过共享内存实现传输。共享内存区提供了最快的进程间通讯的手段。一旦这样的内存区映射到共享它的进程空间,这些进程间数据的传递就不再涉及内核。然而往该共享内存区存放信息或从中取走信息的进程间通常需要某种形式的同步,例如互斥锁、条件变量、读写锁、记录锁、信号灯等。在本系统开发过程中,我们采用的是信号灯方式。

(3) 界面程序的编写

应用界面是利用 PHAB 来完成的。编程用的是 C 语言。QNX4.25 系统中的 PHAB 编程环境功能比 Windows 系统下的图形开发工具功能简单了不少,为此在开发过程中,我们使用了一系列的相关数据库开发技术。例如用动态控件来组合成较复杂的控件;用创建控件仓库的方式提供一次生成、多次调用的功能;用链表的方式对动态控件的资源进行管理;用流式文件的存储方式对用户完成的报表或图表进行保存等。

3 LAMOST 望远镜实时分布式数据库实时性能

2.4 节已经介绍了如何将不同的数据库表格尽可能地分布在不同的本地节点上,以节省数据通讯流量和加快传输速度。说到底,实时分布式数据库的实时性能在很大程度上取决于它所基于的实时分布式操作系统。我们配置的操作系统是 QNX,它是目前国际上公认的最优秀的实时分布式操作系统之一^[4],尤其适合于 x86 下的 PC 网络结构。其内部使用的 Fleet 通讯协议使得各节点上的资源完全透明,形成一个大容量高性能的虚拟机,且数据传输速度大大超过常用的 TCP/IP 协议。实时系统最基本的要求有两个:一是在确定的时间内执行计算或进行处理并对外部的异步事件作出响应;二是用户能对服务进程优先级进行设置。强实时系统响应时间的期望值通常大约从 μs 到 ms 级。中断响应指标是操作系统实时性能的一个重要方面。对于 166 MHz 奔腾处理器来说, QNX4.25 系统提供的典型中断延迟时间为 $3.3 \mu\text{s}$,典型调度延迟时间为 $4.7 \mu\text{s}$ ^[5]。不言而喻,若处理器速度更快的话(目前我们使用的是 500 MHz 的处理器),中断响应时间就更短。由此可见,我们的软硬件配置从平台级提供了一个强实时系统。

得益于这样一个实时平台,加上在数据库设计方面如上所述的技术处理,已研制成的 LAMOST 望远镜实时分布式数据库的实时性能的实测情况又如何呢? LAMOST 望远镜实时分布式数据库所面临的实时性要求最强的任务是,记录机架跟踪时的高度和方位信息。当代国际上大天文光学望远镜跟踪时的位置数字伺服环的校准时间大约是从每 20 ms 到每 110 ms 一次。LAMOST 望远镜与此相仿,取最苛刻的情况即每 20 ms 一次。我们做了如下测试:分别读取高度和方位信息,产生一条记录,再加上时间信息和其它相关信息,这条记录的大小不超过 100 个字节。实测时以 266 字节为一记录长度,每次向数据库中插入 1000 个记录,共插入 10 次,并计算插入每个记录的平均时间。这样的测试分别在两台具有不同主振频率处理器的奔腾机上进行,以便比较。测试结果列于表 1。由表 1 可知,机器的配置对插入速度影响较大,但是即使是较慢的一种配置,也可以满足目前的应用需求。况且将来 LAMOST 真正使用的机器性能肯定要比现在的配置好许多。

表 1 两台不同奔腾机插入一条记录的平均时间的比较

序号	μs	
	奔腾 512 , 512 M 内存的 compaq 工作站	奔腾 233 , 192 M 内存的兼容机
1	12139	18158
2	12129	18178
3	12129	18148
4	12149	18158
5	12139	18258
6	12129	18118
7	12109	18138
8	12129	18128
9	12119	18128
10	12119	18108

为了进一步提高实时性, 建立基于内存而不是磁盘的实时分布式数据库将是我们今后研究的内容. 目前国际上也有这样的动向^[6].

4 结 束 语

LAMOST 实时分布式数据库软件包只是 LAMOST 望远镜控制系统中的一个子课题. 由于我国以往的大天文望远镜的研制中没有这方面的经验, 所以我们在开发过程中遇到了不少困难, 成果来之不易. 这一成果对我国今后大天文望远镜控制系统发展和进一步研制一定会有裨益, 对横向相关领域也可提供一些借鉴.

参 考 文 献

- 1 苏洪钧, 崔向群. 见: 崔向群编. 97 天文望远镜及仪器新技术学术交流会议论文集, 庐山: 中国天文学会天文仪器与技术专业委员会, 1997: 83
- 2 徐欣圻. 天体物理学报, 2000, 20(增刊): 43
- 3 Xu X Q, Zhou J, Xu L Z. In: Andersen T ed. Integrated Modeling of Telescopes, Lund: SPIE, 2002: 145
- 4 Kolnick F. QNX 4 Real-time Operating System, Canada: Basis Computer Systems, 1998: 27
- 5 徐欣圻. 天文仪器与技术, 1999, 1: 71
- 6 <http://www.commsdesign.com/>, 2003
- 7 徐洁磐. 数据库系统原理, 上海: 上海科学技术文献出版社, 1997: 193

The Real Time Distributed Database for LAMOST Control System

Xu Lingzhe Xu Xinqi

(National Astronomical Observatories/Nanjing Institute of Astronomical Optics & Technology, Chinese Academy of Sciences, Nanjing 210042)

Abstract

A brief research summary of a real time distributed database package on the basis of QNX OS is presented. This package is developed for the Large sky Area Multi-Object fiber Spectroscopic Telescope (LAMOST), which is one of the national large scientific projects in the Ninth Five-Year Plan, and has passed a critical review and won positive response among the experts from Beijing, Nanjing and Hefei at the evaluation meeting on July 2, 2002, organized by the LAMOST Engineering Headquarters. The database package is powerful in functions, rich in development tools and easy to operate with a Windows style under UNIX-like operating system. It satisfies the requirement for online/offline data process, storage, programmable creation of tables and graphs for LAMOST control system. The database package can be adapted to the applications in other related areas.

Key words astronomical facilities and observational data—real time distributed database—control system—LAMOST