

# 一种大规模分布式实验数据管理系统

## A Large Scale Distributing Experiment Data Management System

16-18

TL631.24

TP274.2

(合肥)中国科学院等离子体物理研究所(230031)

罗家融 季振山 熊斌 谭立湘  
徐浩 田一海 朱琳 周志英

**【摘要】**HT-7 超导托卡马克核聚变装置主要研究和解决非感应等离子体平衡、稳定、输运、波与等离子体的相互作用,以及等离子体与壁的相互作用等问题,以期达到准稳态等离子体参数,为核聚变堆的商业应用作出先期预演。HT-7 实验数据管理系统主要任务是服务于这些装置诊断仪器的数据采集、存档、分析处理和检索。由于 HT-7 超导托卡马克是个准稳态聚变实验装置,其数据采集量达数百兆,又因为这些数据要在短时间(5 min)之内,提供给分布在所内各大楼内的广大物理实验工作者分析使用,因此设计了基于交换机的高速以太网的分布式 HT-7 实验数据管理系统,本文较为详细地介绍了该系统的结构组成。

**关键词:** 分布式数据管理系统, 数据采集, 网络通信, 数据检索

**Abstract:** The main goals for the HT-7 superconducting Tokamak are to study the plasma equilibrium; instability; transport; the reciprocities between the waves and the plasma and the characteristics between the wall and the plasma, under the non-induction current. The tasks of the management system for the experiment data in HT-7U Tokamak are to collect; to archive; to analyze and to search the data. In order to deal with the huge amount of the experiment data of 200 ~ 400 MB per experiment, and they should be deliver to the experimenters who locate in the distributing buildings less than 5 minutes, the management system is designed as a distributing system based on the local area net. The system is described in details.

**Key words:** distributing experiment data management system, data acquisition, communication with the net, data search system

大规模分布式数据管理系统被广泛应用于核实验、国防科研、航空航天及核电站和核聚变实验装置<sup>[1]</sup>等领域。在建立大型实验数据库管理系统时一般应遵循下列原则<sup>[2]</sup>:①数据库管理系统对用户是透明的,即

北京长城电子衡器寺冈电子 电话:(010)65685281

用户无须了解数据库结构及复杂烦琐程序就能方便地使用数据库数据;②数据库应根据实验诊断通道的变化进行扩充,使用灵活;③数据库应便于检索;④数据库应有用户接口、绘图软件包接口等,接口应简单好用。为核聚变实验装置配置的实验数据库管理系统,除应遵循以上原则以外,必须再满足以下几个要求:①有一定的实时性,即在实验间隔的时间内(一般在 10 ~ 20 min 之间),比较快速地把实验数据传输到数据服务器,供广大实验工作者分析,物理人员可以在计算机上做些必要的物理工作,利用本次实验参数指导下次实验的运行;②高速的数据传送,现在一般聚变装置的放电时间长达数秒,因此实验数据量达数百兆字节。这些物理数据必须在较短时间内传至数据服务器供物理人员处理和分析。基于以上的考虑,我们建立大型的分布式 HT-7 实验数据管理系统,该系统由 HT-7 实验数据采集系统 (HT7DAS)、HT-7 实验数据服务器 (HT7DBS)、HT-7 实验数据检索系统 (HT7DSS) 等组成。图 1 是该系统的结构框图。

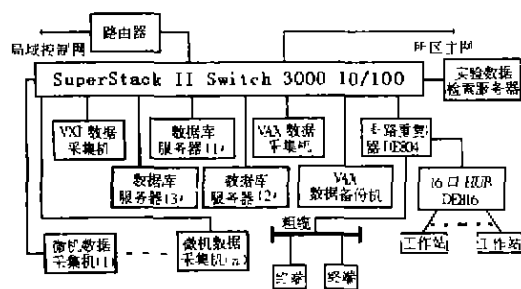


图 1 分布式 HT-7 实验数据管理系统的结构框图

### 1 HT-7 实验数据采集系统 (HT7DAS)

HT-7 数据采集系统主要任务是服务于装置诊断仪器的数据采集和存档。由于 HT-7 超导托卡马克是个准稳态聚变实验装置,其数据采集量达数百兆,因此本系统设计为一种分布式的采集系统,主要包括三种基于不同总线类型的采集模式:CAMAC 总线、PCI 总线和 VXI 总线。

#### 1.1 基于 CAMAC 总线的采集子系统

该数据采集管理子系统基于 DEC/VAX 系列计算

机和 VMS 操作系统<sup>[3]</sup>。由微型 VAX 计算机 VAX4200 承担数据采集,这些数据临时存储在本地硬盘上,同时把数据通过网络实时传送到数据库服务器。大量的数据来源于瞬态记录器模块(ADC 加上缓冲存储器),通过 CAMAC 字节串行分支采集数据,两条字节串行分支使用 KINETIC SYSTEMS 2060 分支驱动器。

该系统软件基于 VAX/VMS 和 MDS 软件系统基础上<sup>[4]</sup>,软件开发使用 VAX 软件工程工具集、VAX 公共数据字典扩充和系统实用程序等。数据采集软件的执行过程是一个表驱动过程,在建立扫描表之前必须由用户创建一个数据库,该数据库是一组索引文件,由数据库定义正文文件描述。一旦数据库建立,利用 CAMAC 服务程序创立扫描表,扫描器程序读表中信息并执行 CAMAC 模块指定的功能,获取 CAMAC 模块数据,而这些数据立刻写入磁盘存入数据库中。

### 1.2 基于 PCI 总线的采集子系统

在本子系统中采用了国内外公司生产的多种不同数据采集卡,以适应不同目标的采集。采用 Windows 或 DOS 操作系统的大量微机承担数据采集,这些数据临时存储在本地硬盘上,同时把数据通过网络实时传送到数据库服务器。该系统软件采用不同的语言编制,主要包括 C、VC++、LabVIEW 和 Delphi。所有的子系统全部采用标准 Socket 网络编程进行数据传送。该子系统的文件以统一的格式传至数据服务器,以适应分布式采集系统的要求。一般对于采集频率要求较低的信号(采样周期大于 10 ms),运用 DOS 操作系统<sup>[5]</sup>。

### 1.3 基于 VXI 总线的采集子系统

采用了航天测控公司军品事业部最新研制的基于 VXI 总线的 AMC3320 高速 4 通道并行模块为核心,组建了本系统。12 块该模块与 HP E8491A 零槽控制器(负责与微机通信)和一台 PIII-500 微机组成了本系统。该系统的采样频率达 1 MHz,其深度达 1M,基本保证了长时间、高精度的物理数据的要求。采用 Windows 操作系统,利用 LabWindow 和 VC++ 的编程进行数据采集,这些数据临时存储在本地硬盘上,同时把数据通过网络实时传送到数据库服务器。

## 2 HT-7 实验数据服务器(HT7DBS)

建立 HT-7 分布式实验数据服务器的目的是为了利用当今先进的数据库管理软件,更有效、更方便地管理实验数据,大大提高实验数据的完整性、一致性、可靠性、共享性;同时为所有的实验工作者提供快捷、友好、简便的交互式数据分析软件。本系统的使用者为一般实验工作者,因此要求系统功能灵活、强大,数据传输和分析准确可靠。

该数据服务器建立了一级数据库(原始数据)、二级数据库(处理过数据)和三级数据库(物理数据)。该

数据库对用户是透明的,用户无须了解数据结构即可方便地获取数据,并通过该系统所提供的绘图软件包(该软件包以 Delphi 为基础开发而成),方便地在自己的微机上作出有关的图形。同时该系统还提供一般常用的数学处理工具,使用户能方便地利用该数据库,对物理数据进行初步的处理。

所有的实验数据在实验过程中,准实时地从 VAX 微型计算机和所有分布式微机采集机传输到用作数据库服务器的 Workstation 上(该服务器具有海量存储能力),并把实验数据按照统一的格式存入数据库中,建立一级数据库。同时由另一台服务器对数据进行初步处理,生成二级数据库。由于生成三级数据库一般均需较长时间,因此三级数据库在实验结束之后生成。所有的用户都可以在本地机上利用本系统的画图软件包,自动地从该服务器上选取所需数据(一级数据库、二级数据库或三级数据库),作出相应的物理图像。

由于 HT7DBS 采用的是分布式采集形式,各采集子系统具有不同的操作系统,其存储在本地的数据格式不完全相同,因此通过网络通信技术,将分布式数据以统一的格式传输到服务器是必须的。为了保证网络的传输速度,设置了 3 台双 CPU 服务器,在有大量用户使用服务器的数据时,120MB 的数据量在 3 min 内可全部传输完毕。满足了物理人员在较短时间内处理和分析数据的需要。

在 HT7DBS 通信系统里,主要使用了两种数据通信方法:

① 共享。在硬盘上建立一个文件,一个应用程序往该文件里写数据(可以不关闭文件,但必须刷新缓冲区),另一个应用程序以共享方式打开这个文件并读取其中的内容,这便是最简单的数据交换方式。对于网络用户而言,只要两台终端上安装的都是 WIN 3.11 或 Windows 9.X(或 NT),则只要设置一下目录共享,映射成网络驱动器,同样可以简单地实现数据交换,采用这种方法不仅避免了通信双方握手信号的规定,而且可以一次进行多个文件的存取,与多台机器交换数据,提高了通信的速度。但缺点也是显而易见的:只能适合于可采用轮询方式获取数据的程序(效率低下),网络映射的驱动器绝对不能变动或取消(可靠性差)。

② 标准的 Sockets 和 WINSOCK 编程。Sockets 的内核简单,可将其看成是一个双向的节点,一个应用程序可以通过它先向另一个程序建立连接(建立在一个双方都认可的端上,以便于区分同时运行的几个通信程序),然后就可以彼此交换数据了。

在本系统中此方法主要用于不同操作系统之间的通信,如 VAX 的 VMS 系统、路由器的 Linux 系统与数据服务器的 Windows NT 操作系统之间的通信,应用不

同的编程语言:ANSI C、Visual C++、Delphi 编写而成的,它们不仅解决了系统中的一些网络通信问题,取得了良好的实际应用效果,而且为用不同语言在不同操作系统下进行 SOCKETS 程序开发提供了极好的模型样本。图2是本系统的结构和数据流程图。

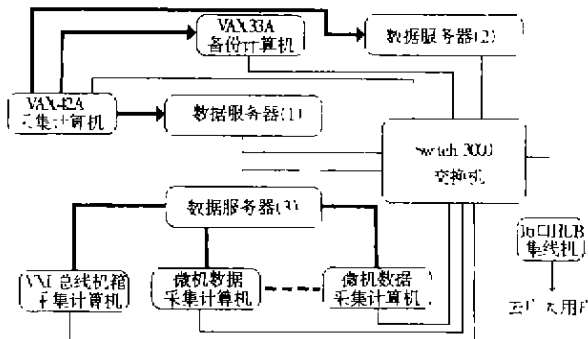


图2 实验数据服务器的结构和数据流程图

### 3 基于 WWW 的 HT-7 实验数据检索库

随着 Internet 的迅猛发展,基于 WWW 的信息发布和数据共享技术已成为新的热点。WWW 是目前 Internet 上最为成熟的技术,其最突出的优点在于它的简单性、开放性和良好的用户界面,但它又是一种检索机制不够完善的分布式信息系统,考虑到以往的各种数据、信息大多存储在各种数据库系统上,因此将 WWW 与数据库系统结合起来,相互取长补短,形成集数据管理、分布式网络功能和支持超文本、超媒体于一体的、具有实时性和交互性的分布式信息系统,已成为当前 Web 业界应用的潮流。

基于 WWW 的物理实验数据检索系统是一项将目前广为流行的 Internet/Intranet 技术应用于 HT-7 数据采集系统,采用当今最为流行的浏览器/服务器模式建立的、基于 WWW 的多用户分布式物理实验数据检索系统。用户可以在自己熟悉的 Windows 环境下,利用通用浏览器软件(Netscape Navigator 和 Internet Explorer)随时查询 HT-7 物理实验中自己所关心的关键数据信息。同时系统建立了一个二级图库,在检索的同时提供给用户浏览实验中关键信号的波形图。该系统为用户提供了方便快捷的查询方式,数据安全机制和应用程序接口,它将有助于在实验中不断总结规律,不断创新突破<sup>[6]</sup>。

基于 WWW 的物理实验数据检索系统是建立在浏览器/服务器体系结构上的系统,这是目前广为流行、较为先进的软件开发模式。浏览器/服务器模式是从客户机/服务器模式发展而来的,它的本质仍然是“请求驱动”,但是它在综合多项技术的同时,克服了客户机/服务器结构的一些不足。因此该检索系统具有如下特色:①解决了客户机/服务器应用中存在的客户端

跨多平台的问题。浏览器/服务器模式以 Web 技术为核心,客户端具有统一的用户界面,即通过浏览器可以访问多个应用服务器。而基于客户机/服务器模式的软件系统则往往要求一个客户端,一个服务器成对出现,如果一个客户机要访问多个服务器,就必须配置多种客户软件。②客户端真正实现了零维护。浏览器/服务器结构的系统维护费用相当低廉,软件版本的更新不涉及用户,只需将服务器端的软件更新,具有极强的扩展性。③客户端的软硬件配置要求不高。基于浏览器/服务器结构的检索系统是真正意义上的瘦客户机/胖服务器模式,客户机的软硬件配置要求降到最低限度。④简单灵活的用户界面。以浏览器作为统一的用户界面,使用起来非常简单、易操作,免去了对用户的培训。系统应用环境为基于浏览器/服务器体系结构,网络操作系统选用 Windows NT 4.0 Server,数据库管理系统选用 Microsoft SQL Server 6.5,客户端开发平台选用面向对象的编程语言 Delphi 4.0 来进行设计和实施。图3是本系统的逻辑结构流程图。

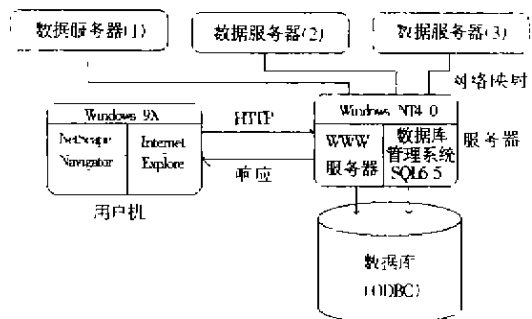


图3 实验数据检索库的逻辑结构流程图

### 4 结论

以微机为主建立的大规模分布式 HT-7 实验数据管理系统,具有成本低、系统更新换代、技术改进及维护修理容易等优点。同时,该系统的各项技术性能及可靠性,均已满足了大型核聚变实验装置物理实验的需要,系统的操作管理简单容易,完全适合中国的应用环境,在我国具有一定的推广价值。

#### 参考文献

- 1 Van H, et al. CODAS: THE JET CONTROL AND DATA ACQUISITION SYSTEM. Fusion Technology, 1997, 11: 120
- 2 陈燎原等. 分布式微机大规模 DAS 实现技术. 计算机自动测量与控制, 1996(2): 42
- 3 中国科学院等离子体物理研究所计算机应用研究室. IPP 数据采集系统. 计算机自动测量与控制, 1994(4): 15
- 4 中国科学院等离子体物理研究所计算机应用研究室. IPP-DAS 数据采集软件. 计算机自动测量与控制, 1994(4): 18

作者简介:罗家融,男,1948年生,研究员,博士生导师,现在中国科学院等离子体物理研究所从事分布控制、数据采集和网络通信系统的研究。

(收稿日期:2000-07)□