

# HT-7 超导托卡马克总控数据发布系统研究

王 华<sup>1,2</sup>, 罗家融<sup>2</sup>, 瞿连政<sup>2</sup>, 魏沛杰<sup>1,2</sup>

(1. 中国人民解放军炮兵学院二系, 合肥 230031; 2. 中科院等离子体物理研究所, 合肥 230031)

**摘 要:** 为了适应 HT-7 超导托卡马克实验的需要, 有必要研制实验总控数据发布系统, 使实验人员通过该系统获取实验中各控制系统数据, 从而及时了解实验中装置的运行状态。文章介绍了系统的硬件结构和系统实现过程, 并给出了系统测试结果。

**关键词:** HT-7; 数据发布; 等离子体; 托卡马克

## Study of Control Data Publishing System on HT-7 Superconduction Tokamak

WANG Hua<sup>1,2</sup>, LUO Jiarong<sup>2</sup>, QU Lianzheng<sup>2</sup>, WEI Peijie<sup>1,2</sup>

(1. Dept. 2, Artillery Institute of PLA, Hefei 230031; 2. Institute of Plasma Physics, Chinese Academy of Sciences, Hefei 230031)

**【Abstract】** To meet the requirement of experiments on HT-7 Tokamak, it is necessary to develop the experimental control data publishing system. And it is possible for scientists to acquire the control data of devices in experiments and make decision how to do in time. The article introduces the system hardware structure and the system realization process, and gives some test results.

**【Key words】** HT-7; Data publishing; Plasma; Tokamak

### 1 概述

HT-7 是国内第 1 个超导托卡马克受控核聚变实验装置, 中科院等离子体所已经在该装置上进行了多轮放电实验, 获得了大量实验数据及多项重要成果。HT-7 控制系统由总控系统和各个分控系统组成。总控系统是整个 HT-7 控制系统的核心, 负责完成对所有分控系统全部的监视、协调和控制操作, 保证等离子体放电实验的正常进行。HT-7 分控系统包括极向场控制、密度控制、离子回旋(ICRF)控制、低杂波(LHCD)控制等控制系统, 各分控系统在放电过程中共同作用, 控制等离子体的行为和状态。

目前, HT-7 数据采集系统已经实现了对实验中所有诊断信号的数据采集、压缩、存储、数据发布, 实验人员可以很方便地通过 GT7(HT-7 数据浏览软件)和基于 Web 远程数据获取系统浏览数据。随着 HT-7 物理实验的不断深入, 为了更好地控制等离子体, 实验人员希望能及时方便地查看放电过程中各控制子系统的控制数据, 因此需要研制总控数据发布系统。但是相比 HT-7 数据采集系统, 控制数据的收集、传输和发布存在以下两个问题:

(1)控制数据的格式极不统一。由于各控制子系统的采用的操作系统和控制软件的不同, 各控制子系统都规定了自己的控制数据格式, 如密度系统用脉冲时刻、脉冲宽度、脉冲幅度这样的三维数组来记录控制数据, 限制器温度数据则采用 EXCEL 文件格式存储。

(2)控制数据中的放电零时刻不统一。各控制子系统处于实验装置的不同位置, 某些控制子系统得不到中央定时系统的基准零时刻触发信号, 还有某些控制系统是采用软件定时的方式进行采集触发, 所有这些导致各控制数据中的放电零时刻都不统一。

### 2 系统方案设计

HT-7 的网络系统包括 3 种类型的子网: 所内局域网, 控

制子网、采集子网。所内局域网是等离子体所内部的局域网, 它通过中科院的代理服务器接入 Internet。总控系统和各个分控系统都处于控制子网当中, 控制子网通过具有防火墙功能的路由器以光纤方式接入所内局域网, 控制子网在有了防火墙的保护以后, 可以减少计算机病毒的侵袭和来自外部的攻击。如图 1 所示。

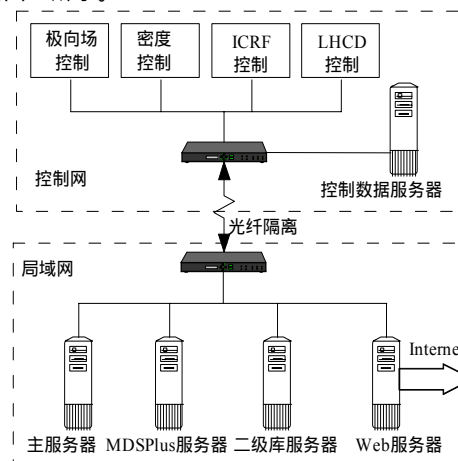


图 1 系统网络结构

HT-7 的数据服务器由主服务器、MDSPlus 服务器和 2 级库服务器组成。主服务器也是原始数据服务器, 由于实验过程中, 数据量巨大, 因此以 LZ0 格式存储实验原始数据, 数据的压缩比约为 1:3。MDSPlus 服务器是专为脉冲实验设计的进行数据采集存储与分析的一种数据库服务器。2 级库

**基金项目:** 国家发展计划委员会“计投资(1998)1303 号项目”(子项目)

**作者简介:** 王 华(1976 - ), 男, 博士生, 主研方向: 数据采集与控制系统; 罗家融, 研究员; 瞿连政、魏沛杰, 博士生

**收稿日期:** 2006-04-15 **E-mail:** wanghua\_qq@ipp.ac.cn

服务器存储初步处理过的数据,也以 LZO 格式存储。Web 服务器将实验数据发布到 Internet,供用户远程访问。

在控制子网内设置一台控制数据服务器,它的主要功能是收集各控制子系统的控制数据,并将其转化为标准格式发送到 2 级库服务器中。

总控数据发布系统的主要工作流程如下:

(1)当一次放电结束后,各控制子系统将控制数据传送给控制数据服务器,服务器和各控制系统之间通过基于 TCP/IP 协议的 Socket 方式进行通信。服务器上运行一个服务程序用来接收各控制系统的控制数据。

(2)控制数据服务器对接收的控制数据调用相应的格式转换模块进行数据转换,将不同格式的控制数据转换为 GT7 可识别的标准格式,并发送到二级库服务器上。

(3)二级库服务器以数据采集系统中等离子体电流(IPA)信号为基准,对各控制数据的放电零时刻进行校准。

(4)控制数据用 LZO 压缩算法进行压缩,存储到服务器中。

(5)通过 Web 服务器将控制数据进行远程发布。

### 3 系统实现

#### 3.1 控制数据格式的转换

为了和数据采集系统的数据格式统一,采用 GT7 规定的 txt 格式为标准格式,文件格式由数据头信息和数据体两部分组成。数据头信息记录了本道信号的相关信息,包括信号名、信号单位、触发时刻等信息。数据体全部按照浮点数以文本的形式进行保存,一行保存一个数据点。这样存储方式虽然有一定的冗余,但是方便了后续的数据处理过程。

#### 3.2 放电零时刻的校准

各控制子系统的触发不统一,导致相应控制数据中的放电零时刻都不一致,而且这种定时误差并不是固定的,不能通过定标的方式来固定的扣除,必须是通过软件的方式自动计算出本次放电控制数据零时刻和标准零时刻之间的差距,并在触发时刻这个参数项中进行扣除。

HT-7 的各控制子系统一般都需要采集等离子体电流信号(IP),就以数据采集系统的电流信号(IPA)为基准,通过比较控制数据中的电流信号和基准信号之间的差异,计算出相应的定时误差,进行放电零时刻的校准。如图 2 所示。

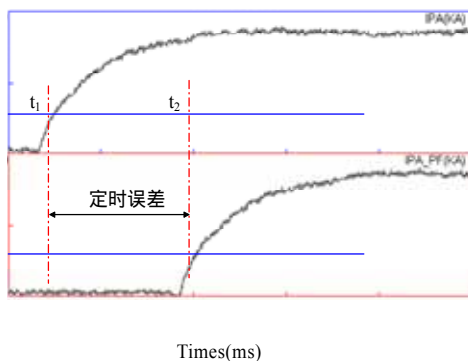


图 2 零时刻校准示意图

图中显示两道信号,IPA是数据采集系统采集的等离子体电流信号,IPA\_PF是极向场控制系统的采集的电流信号,这两道信号从形状和大小看极其相似,仅仅是触发时刻上有差异。在这两道信号上各作一条IP=35KA的水平线,信号IPA和IPA\_PF和水平线各有一个交点,它们的X坐标(时间坐标)分别

记为 $t_1$ 和 $t_2$ ,因为这两道数据其实都是表示的是同一信号,它们到达IP=35KA的时刻之间的差距其实就是两个系统触发之间的定时误差。选取IP=35KA 作为参考线,是因为在这一阶段IP上升的速度较快,可以减小校准时的计算误差。

控制数据的定时误差  $\Delta t = t_2 - t_1$ ,将 IPA\_PF 信号中的数据头中触发时刻减去  $\Delta t$ ,得到就真正的放电零时刻,实现了放电零时刻的校准。实验结果表明,用这种方法进行零时刻校准的计算误差小于 1ms。

#### 3.3 数据远程发布

Web 服务器负责对控制数据的远程发布。它主要分为两个部分:Web Services 和应用服务程序,如图 3 所示。

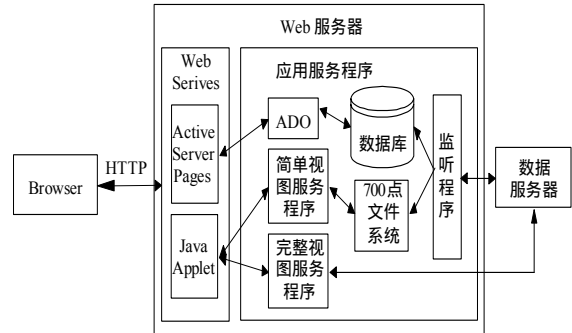


图 3 Web 服务器结构

其中 Web Services 又包含动态网页编程和 Java Applet 的编写。传统网页编程生成的文档类型主要为静态文档和动态文档。浏览器在取得文档后,文档内容不会再改变。对于数据波形图而言,由于要求响应用户鼠标动作,传统的网页编程已经不能满足设计要求。采用 Java Applet 显示实验数据波形图可以满足系统要求。Java Applet 是一类活动文档,它是一个包括计算并显示值的小应用程序。当浏览器访问实验数据波形图时,服务器返回一个浏览器可以本地执行的 Java Applet 程序。当 Applet 程序运行时,它可以和用户交互执行并不停地改变显示。

应用层的视图服务程序用来对 Java Applet 提供数据支持,并隔离 Java Applet 对数据服务器的直接访问。简单视图服务程序的数据来自 700 点文件系统,是用 Java 编写的应用程序。为了减轻主数据服务器的负担,避免和实验人员争用资源,完整视图和监听程序的数据均取自于 MDSPlus 数据服务器。监听程序用于在实验过程中,侦听实验最新进展,一旦一次新的放电完成,程序就对数据进行处理,计算出最新放电情况的相关特征值,并产生 700 点的数据。监听程序保证了系统的实时性。

### 4 系统测试结果

对系统的测试主要包括:(1)从放电结束开始控制数据经过格式转换、零时刻校准、数据压缩等处理,最终存储在服务器所需要的时间;(2)用户浏览数据需要的时间。

#### 4.1 数据发布速度测试

每次放电长度不同,控制数据的大小也不同,表 1 显示不同放电长度、不同数据大小的条件下数据发布时间的测试结果。测试表明:在短脉冲放电的条件下,系统可以在较短的时间内将控制数据发布到数据服务器上(通常<10s),随着数据量的增加,系统的发布时间也会相应增加。在长脉冲放电(放电长度为 320s)的条件下,数据发布的时间也小于 1 分钟,达到系统设计的要求。

**表 1 数据发布时间测试**

格式 炮号	放电长度 (s)	控制数据量 (MB)	发布时间(s)
82000	10	11.5	8
83256	15	16.3	12
82975	74	44.6	20
82990	252	103.0	31
83026	320	137.1	48

**4.2 远程浏览数据测试**

用户的浏览速度是评价这个系统是否满足要求的一个重要指标。通过完整视图，局域网内用户可以迅速获取全部实验数据。通过简单视图，Internet 用户即使使用 56Kbps 电话线拨号上网，也可以较快地获取数据。表 2 和表 3 分别为通过完整视图和简单视图浏览数据的测试速度。

**表 2 完整视图测试速度**

	单信号(数据量≈4MB)	10 道信号(数据量≈50MB)
浏览速度	3s	8s

**表 3 简单视图测试速度**

	局域网	56Kbps 电话线	宽带
单信号(数据量≈8kB)	<1s	3~4s	<1s
10 道信号(数据量≈80kB)	<2s	6~7s	<2s

在系统中，Applet 和服务程序间采用 Socket 通信进行数据传输，相对于在服务器中，直接将信号波形以图片格式传送给客户端来说，这种方法的速度要慢一些。创建 Socket 连接需要一定的时间，同时，网络的情况也会影响数据传输的时间。服务器通过网络将数据发送到客户端，需要经过多个路由转发，若其中某个路由屏蔽了当前的端口，数据将无

法传递给客户端。如果采用图片格式传输，由于图片属于静态文档，则无法响应用户鼠标操作。通过 Java Applet 实现此系统，虽然在一定程度上降低了用户的浏览速度，但从实际测试速度来看，已能基本达到用户需求。

**5 结束语**

HT-7 总控数据发布系统在 2005 年冬季 HT-7 物理实验中进行了首次应用。系统将来自极向场、密度等 4 个分控系统的控制数据(共计 60 道信号)经过格式转换、零时刻校准、数据压缩等处理后进行发布。经过测试，系统稳定可靠，数据发布速度和数据远程访问速度都达到设计要求。对今后 EAST(中科院等离子体所新一代的全超导 Tokamak 装置)控制数据的发布具有借鉴意义。

**参考文献**

- 1 Luo J R, Wang H Z, Ji Z S, et al. The Distributed Control and Data System in HT-7 Tokamak[J]. IEEE Transactions on Nuclear Science, 2002, 49(2): 496-500.
- 2 Cheng Ting, Luo Jiarong, Meng Yuedong, et al. A Web-based System for Remote Data Browsing in HT-7 Tokamak[J]. Plasma Science & Technology, 2005, 7(4): 2926-2928.
- 3 Wan Yuanxi, HT-7 Team, HT-7U Team. Overview of Steady State Operation of HT-7 and Present Status of the HT-7U Project[J]. Nuclear Fusion, 2000, 40(6): 1057.
- 4 谢向阳. HT-7U 采集和控制系统设计[D]. 合肥: 中科院等离子体所, 2000.
- 5 朱琳. HT-7 核聚变实验数据系统的研究[D]. 合肥: 中科院等离子体所, 2004.

(上接第 215 页)

这种方法在保证生成速度的前提下取得了不错的效果。但是我们也发现算法中存在很多值得进一步研究的地方，可以归纳为以下几个方面。

(1)基于滑动窗口的动态摘要属于基于统计方法的摘要，由文章中的一二个片断组成。因为没有对句子进行重组润色，所以这种方法生成的摘要有时也存在不够连贯的缺点。如果能在截断的位置对文本结构进行分析，那么将能提高摘要的连贯性。

(2)单个检索词的片断不存在检索词距离、覆盖检索词的比例等概念。那么，计算出来的片断分数则会依赖于某个单一条件，生成的摘要就不会很理想。所以针对单个检索词应该有特殊的处理方法，可以运用命名实体识别、文本聚类 and 用户日志分析等多种技术，这将是我们的未来的工作。

(3)歧义问题。由于考虑到速度因素，算法没有使用中文分词技术，因此生成的摘要不可避免会出现歧义问题。一种可以改进的方法是保存搜索引擎创建索引时的分词结果，在生成摘要的时候重复利用，从而避免了再次分词。

**参考文献**

- 1 刘挺, 吴岩, 王开铸. 自动文摘综述[J]. 情报学报, 1998, 16(1): 63-69.
- 2 Tombros A, Sanderson M. Advantages of Query Biased Summaries in

Information Retrieval[C]//Proceedings of ACM SIGIR Conference. 1998: 2-10.

- 3 黄纯敏, 吴郁莹. 网路文件自动摘要[C]//台湾区网际网路研讨会 TANET. 中国: 台湾. 1999.
- 4 White R, Ruthven I. Web Document Summarization: A Task-oriented Evaluation[C]//Proceedings of the 1<sup>st</sup> International Workshop on Digital Libraries. 2001: 3-7.
- 5 Ruthven I, Tombros A. A Study on the Use of Summaries and Summary-based Query Expansion for a Question-answering Task[C]//Proc. of the 23<sup>rd</sup> BCS European Annual Colloquium on Information Retrieval Research. 2001: 41-53.
- 6 Callan J P. Passage-level Evidence in Document Retrieval[C]//Proceedings of the 7<sup>th</sup> Annual ACM SIGIR Conference on Research and Development in Information Retrieval. ACM Press, 1994: 302-310.
- 7 Luhn H P. The Automatic Creation of Literature Abstracts[J]. IBM Journal of Research and Development, 1958, 2(2): 159-165.
- 8 孙春葵, 李蕾, 杨晓兰, 等. 基于知识的文本摘要系统研究与实现[J]. 计算机研究与发展, 2000, 37(7): 874-881.
- 9 Hand T, Sundheim B. The Tipster Summac Text Summarization Evaluation[R]. Working Papers from the Summac Conference, NIST, Fairfax, VA, 1998.