

# EAST\_VCL 数据服务中心的设计

张明新<sup>1,2</sup>, 罗家融<sup>1</sup>, 陈国龙<sup>2</sup>, 李贵明<sup>1</sup>

(1. 中国科学院等离子体物理研究所, 合肥 230031; 2. 淮北煤炭师范学院, 淮北 235000)

**摘要:** 在 EAST\_VCL 应用系统中, 构建了基于 XML Web Services 的数据服务中心, 实现了多种类型的数据转换功能, 并且组建了实验数据特征抽样、数据压缩等服务。实验证明, 该服务中心不仅具有良好的可扩展性, 而且有效地提高了 EAST\_VCL 应用环境中数据传输性能, 增强了应用的可访问性。

**关键词:** 分布式应用; XML Web services; 数据传输; 可访问性

## Design of Data Service Center for EAST\_VCL

ZHANG Mingxin<sup>1,2</sup>, LUO Jiarong<sup>1</sup>, CHEN Guolong<sup>2</sup>, LI Guiming<sup>1</sup>

(1. Institute of Plasma Physics, Chinese Academy of Sciences, Hefei 230031; 2. Huaibei Coal Industry Teachers College, Huaibei 235000)

**【Abstract】** Data service center is constructed with the XML Web services in ESAT\_VCL, which implements the function of data transformation, and the services that sample the character data from experimental data system and compress the data for transmission. It is proved in the experiments that the data service center not only shows a well expansibility, but also advances the capability of the data transmission and the accessibility in EAST\_VCL.

**【Key words】** Distributed application; XML Web services; Data transfer; Accessibility

EAST(The Experimental Advanced Superconductive Tokamak)是我国自主研发的新一代高性能的超导托卡马克实验装置, 基于该装置的国际合作不断被推向更广泛、更深入的研究领域。传统的实地合作研究模式不仅给合作者带来更大的财力、人力投入, 且更多的合作者仍无法直接参与到实验过程, 显然, 这种合作研究模式愈发不合时宜。利用计算技术、信息技术和网络技术, 构建EAST的虚拟合作研究平台EAST\_VCL (Virtual Collaboratory), 可将更广泛的合作伙伴联系在一起, 实现在虚拟的空间开展项目合作研究<sup>[1]</sup>。

EAST\_VCL是基于异构数据环境的分布式应用系统, 并且, 随着EAST研究的发展, 实验数据量呈几何增长。多种多样的数据格式以及庞大的数据资源, 不仅增加了EAST\_VCL计算的复杂性, 也给应用的数据通信提出了难题。因此, 针对EAST\_VCL分布计算的特点, 构建统一的数据交换中心, 解决数据格式转换和数据传输服务, 能够提高整个应用系统的数据可访问性<sup>[2]</sup>。

### 1 XML Web Services 技术特点分析

XML Web Services是基于HTTP和XML协议的分布式对象技术, 其松散耦合特点是COBRA和DCOM等技术所无法比拟的, 为实现应用系统的可伸缩性提供了新途径。XML Web Services已经得到像Microsoft、IBM、SUN等业界大公司的认可和推崇。作为新一代分布式计算模型, 它是真正将计算设备和资源链接到一起的基础架构。XML Web Services采用多层构架组建分布式计算模型, 它提供一套访问机制, 客户通过配置代理可实现远程访问<sup>[3]</sup>。

### 2 EAST\_VCL 业务逻辑架构

随着 Web 计算架构的优化、技术支撑能力的提升, 不仅越来越多的事务计算转向这种计算结构, 而且, 基于科学计

算的应用也不断采用这样结构。特别是, 插件技术为解决 Web 应用复杂计算提供了支持, 通过 Web 页嵌入, 这种控件能够提高客户端的计算能力。EAST\_VCL 正是基于 Web 模型的分布式应用系统, 它利用.NET 框架的 Windows 控件库模型构建了托管控件, 实现了业务逻辑封装。同时, 该控件在 Web 页中保留着 Windows 应用程序的计算风格和操作特点。

在托卡马克实验中, 随着实验控制技术的不断改进和完善, 放电时间不断增加, 因而, 产生的实验数据量越来越大, 原来实验每炮(炮: 实验组织单位)只有几兆的数据量, 而目前, 达到几百兆、几千兆的大数据, 给 Web 应用的远程计算与数据传输带来更大的困难。

在EAST\_VCL的应用中, 为了避免像一般Web应用中频繁的访问和数据传输过程, 采用客户端缓存技术, 在计算过程中只进行一次性的数据交换(读/取), 这样, 既有效地保持了计算的连续性, 又维护了系统运行的时效性, 也大大提高了系统的运行效能<sup>[4,5]</sup>。

### 3 EAST\_VCL 数据交换服务

对于 EAST\_VCL 系统, 这些应用所引用的数据类型各不相同, 对于 HT-7 实验数据, 远程控制系统引用的是文本格式数据, 而对 EAST 实验, 则引用的 MDSPulse 格式数据。因此, 构建数据交换中心可以为 EAST\_VCL 各子系统提供一致数据服务。EAST\_VCL 的数据服务中心是采用可跨平台、可

**基金项目:** 国家自然科学基金资助项目(10101008); 安徽省优秀青年基金资助项目(06046009)

**作者简介:** 张明新(1966—), 男, 博士生, 主研方向: 网络工程与虚拟技术; 罗家融, 研究员; 陈国龙, 教授; 李贵明, 博士生

**收稿日期:** 2006-01-09 **E-mail:** mx\_zhang@ipp.ac.cn

伸缩性和松散耦合特点的 XML Web Services 技术为架构。

### 3.1 数据服务

对于目前 Internet 状况,进行大量数据传输势必消耗掉许多带宽资源,严重影响应用系统的性能。对于像 EAST\_VCL 这样的分布式应用系统,在一次长脉冲放电过程中,一个诊断信号的数据量大到几十兆(81 349 炮,放电时间,IPA 信号的数据 25.1 兆,文本格式),在一般的网络带宽下,系统几乎无法正常运行。

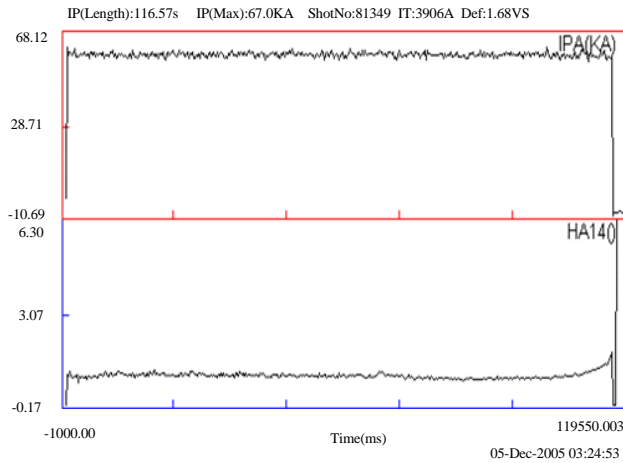


图 1 抽样后 IPA 与 HA14 信号的时间区间对应关系

在以 XML Web Services 架构的数据服务中心,针对托卡马克实验特点,采用以下措施,有效地解决了大容量数据交换的问题。

(1)数据抽样 设计专门的抽样算法,对分析的实验数据进行抽取,可形成轻量的数据集。抽样的关键是基准时间(即信号抽样的时间范围)的计算,(在 HT-7 和 EAST 实验中,所有诊断信号的采集基于同步的触发信号),等离子体电流信号(IPA)是托卡马克实验重要的物理特征信号之一,抽样算法的基准时间是参照此信号的时序关系进行计算的(见图 1)。抽样算法可以按照步长和区间两种方法进行抽样,在步长抽样方法中,对所分析的数据在 IPA 有效区间中按步长进行抽取;在区间抽样方法中,对所分析的数据在按数据段进行抽取。对于远程数据分析系统采用抽样方法,可以实现理想的传输数据量级,以便提高系统运行的性能和效率。

(2)数据压缩 是减少数据量,提高传输效率的另一种有效方法。LZO 数据(免费的软件)就是一种压缩的格式,采用这样方法,需要在客户端也集成解压工具,并且,压缩后采用的是二进制格式,再需要进行格式序列化处理,以及 XML 格式封装<sup>[6]</sup>。在 .NET 框架中,数据序列化可以实现数据格式的转换,在 EAST\_VCL 中,远程控制系统可以采用此方法。

由于控制的信号只有十几道,在 HT-7 实验中,这些信号数据采用文本格式进行存储,具有较高的压缩比(LZO1.0.8 版本的压缩比一般在 40%左右)。

(3)数据加密 在 EAST\_VCL 应用中,一些重要的信息必须进行加密,才可以有效保护数据的安全,在远程控制系统中,涉及到实验装置的运行安全性,为了保护有效的远程控制,对控制流程采用用户信息的跟踪监控措施,在往返的数据流中,需要对用户信息进行加密处理。

(4)数据接口 实现数据格式的转换。在 EAST\_VCL 应

用的不同子系统中,建立了不同的数据接口,为它们提供一致的数据格式转换机制。如远程控制系统,采用 LZO 格式进行数据传输,而在 HT-7 和 EAST 装置中,控制系统产生的数据格式不同,HT-7 生成文本数据,EAST 则产生 MDSPlus 格式数据,接口就必须实现相应转换功能。

### 3.2 数据服务框架

在 EAST\_VCL 应用中,采用客户层、业务层、数据层 3 层交换模型。XML Web Services 的数据服务框架见图 2。

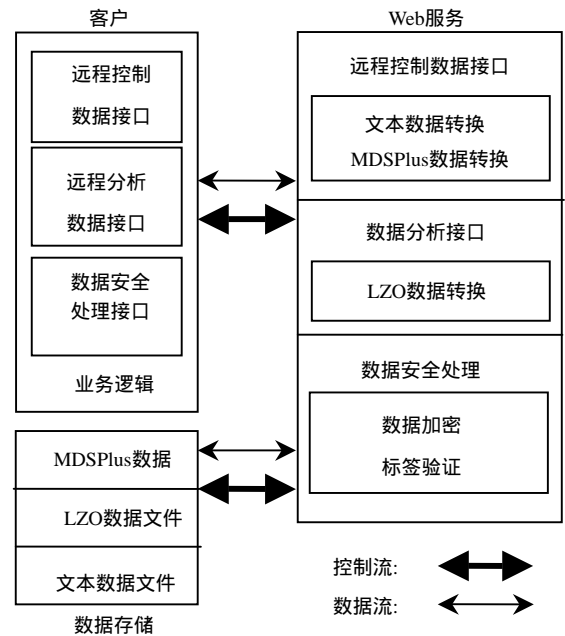


图 2 数据服务中心框架

(1)客户层 实现数据的请求、计算、回写等功能。为实现 LZO 数据交换,将 LZO 程序集嵌入到该层中。

(2)业务层 根据不同的访问请求,启动相应的数据接口,实现对应的数据服务。

(3)数据层 实现底层数据访问。

### 3.3 数据交换的安全控制

数据加密可以增强信息本身的安全性,但是,像远程控制系统,它不仅提供资源访问,而且,用户还能够参与到实验的过程控制,进行协同实验,对这样的访问控制流必须加强安全检查,以保护系统运行的安全性。

在远程控制中,采用了用户标签作为访问流的检查对象,当用户登录时,服务器收集用户的名称、密码以及访问设备的特殊标志,并做成用户标签,在整个访问过程,系统内部一直使用此标签标识该用户,除了数据交换,在设备访问时,服务对每个环节也进行标签验证。

## 4 实验分析

在 2005 年的两轮实验中,EAST\_VCL 得到了比较全面的测试,数据服务中心提供了稳定、可靠的数据通信,友好地配合着相应的远程服务功能。图 3 和图 4 展示了不同计算环境中数据分析的图谱;A 图是远程系统在 Web 页中显示的数据曲线,a、b、c 曲线反映的是经过抽样后的数据。B 图是本地 GT-7 工具中显示数据曲线,a、b、c 曲线反映的是原始数据。

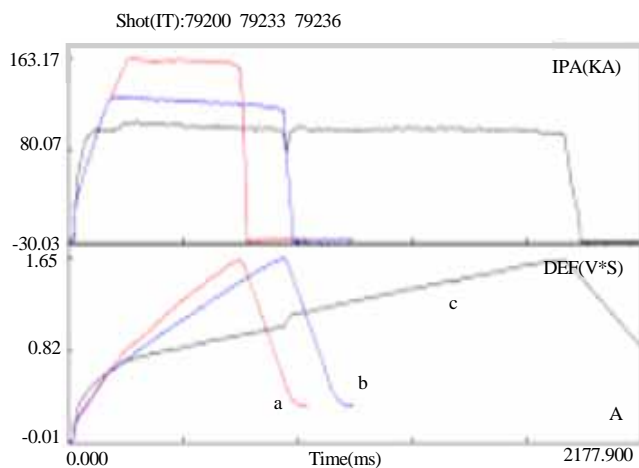
在 IPA 有效时间范围,DEF 数据在两个系统下表现的物

理特征是一致的，因此，研究人员利用远程分析系统可以获得与本地工具软件一样的数据分析结果。

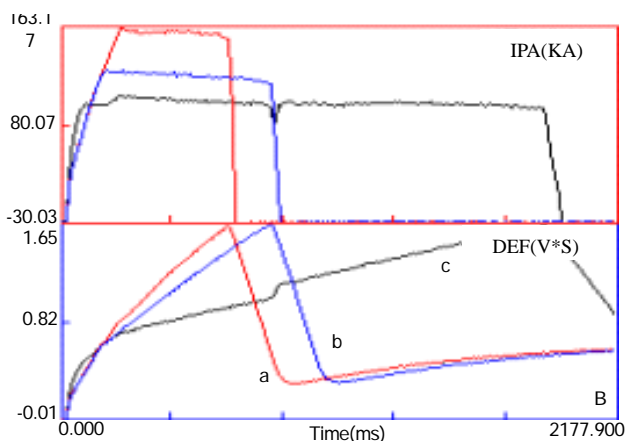
远程控制系统采用压缩数据方法来降低数据量，以提高数据的传输性能，此方法在客户端要额外增加解压的系统时间。通过进一步对数据服务中心的传输能力测试，反映了该服务机制能够满足目前实验的应用要求<sup>[7]</sup>，见表 1。

表 1 数据传输性能测试平均值(单位: s)

数据量	Intranet (100M)	Internet (ADSL 4M)
77k B	0.137	0.678
869k B	1.634	5.564
2843k B	4.788	12.046

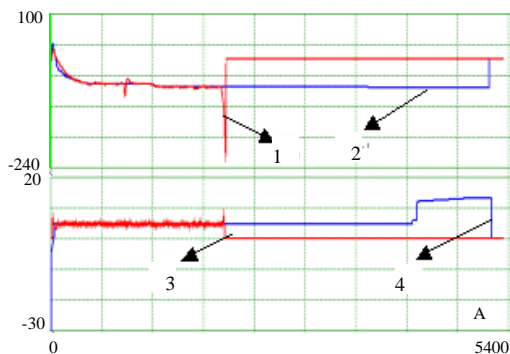


(a) 远程数据分析系统绘制的曲线



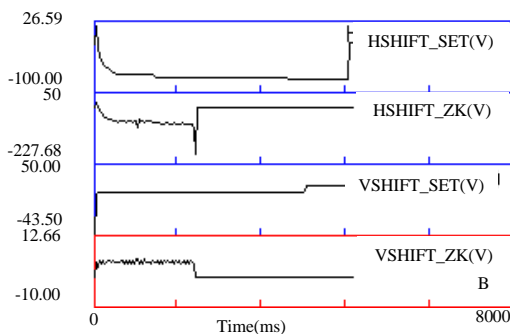
(b) 本地运行系统 (GT-7) 绘制的曲线

图 3 两种曲线



(a) 远程控制系统绘制的曲线

1:HSHIFT\_ZK, 2:HSHIFT\_SET 3:VSHIFT\_ZK, 4:VSHIFT\_SET



(b) 本地运行系统 (GT-7) 绘制的曲线

图 4 79200 炮的垂直位移和水平场位移控制曲线

## 5 结束语

针对科学实验系统，在构建 Web 模型的分布式应用时，往往因为计算的复杂性、数据信息量的因素，制约着应用系统的性能，EAST\_VCL 通过 .NET 框架与 XML Web Services 技术不仅实现了远程科学计算，而且构建了伸缩性能好、跨平台能力强的数据服务中心，在目前 Internet 环境中较好地解决了应用制约因素。

## 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 祝翠玲, 蒋志方, 王 强等. 基于 Internet 的城市烟气实时监控系  
统[J]. 计算机工程, 2005, 31(10): 200-202.
- 3 赵培英, 段会川, 李 欣. 基于 Web Services 的数字仪表监视系  
统设计[J]. 计算机应用, 2005,27(10): 69-72.
- 4 黄陈蓉, 吴慧中. 基于嵌入式的制造系统的远程监测系统[J]. 计  
算机工程, 2005, 31(11): 179-181.
- 5 王喜鹏, 陈灼民, 钱 静等. EAST 装置技术诊断分布式数据采集  
系统的研制[J]. 微计算机信息, 2005,21(19): 78-80.
- 6 李 梅, 毛善君, 卞泉洲. 基于 Web 的煤矿三维可视化控件设计  
与实现[J]. 计算机工程与应用, 2004,40(28): 8-11.
- 7 刘艳芳. 远程控制系统互联网传输时延预测算法的研究[J]. 微计  
算机信息, 2005, 21(19): 70-72.