

基于 Oracle 的 SSRF 数据存档系统及接口设计

张志远 陈建锋 沈立人

(中国科学院上海应用物理研究所 上海 201800)

摘要 上海光源(SSRF)采用 EPICS 作分布式控制系统的软件平台。Channel Archiver 是 EPICS 自带的 OPI 层数据存档工具集,实现对 EPICS 控制系统运行过程中各种数据的快速存取及检索。本文对 EPICS 数据存档系统进行分析,提出现有的不足,并针对这些不足,结合 Oracle 数据库对存档系统重新设计,实现中央数据库统一存储和管理加速器运行数据,并采用 Web Services 技术提供访问数据库的统一接口。

关键词 SSRF, EPICS, Channel Archiver, Oracle, Web Service

中图分类号 TB533⁺.2

上海光源(SSRF)是高性能的中能第三代同步辐射光源^[1,2],电子束能量 3.5 GeV,其同步辐射光覆盖红外至硬 X 射线,是一座可供数百位科研人员同时开展工作的科学平台。上海光源采用 EPICS^[3]作为分布式控制系统,这是美国洛斯阿拉莫斯国家实验室和阿贡国家实验室联合开发的典型分布式控制系统软件开发环境,广泛应用于许多大型科研装置。为方便用户查询、分析系统参数和加速器运行数据,EPICS 提供了 Channel Archiver 存档工具来采集、存储和检索系统运行过程中的机器参数和历史数据。本文对 EPICS 现有存档系统进行分析,提出其在日常应用中的不足,并针对这些不足,设计了基于 Oracle 数据库的新存档系统,实现稳定、可配置的数据采集及利用。

1 EPICS 存档系统分析

EPICS系统中,Channel Archiver是OPI层工具,其通过基于TCP/IP的通道访问协议(Channel access, CA)访问运行于底层的IOC(Input/Output Controller)服务器。存档引擎采集到IOC回传的数据后,以特定格式将数据保存到本地磁盘;当文件大小达到设定值时,系统会自动建立一个新的数据文件。SSRF控制系统每年存档的数据量可达1 TB以上,可为科研人员提供海量的历史数据。但Channel Archiver原有的存档方式在实际应用中也存在诸多不足^[4],具体表现在:(1) Archiver Engine的数据存储格式固定,用户只能按时间戳等少有的几个条件查询,不能满足用户实际需求;(2) 历史数据分散在众多文件中,不利于各种应用程序访问数据且管理不便。

2 存档系统的设计

为弥补EPICS自带的Channel Archiver工具集的不足,同时用Windows操作系统下丰富的软件和工具以便于程序功能的扩展,本存档系统将工作于Linux平台的Channel Archiver工具移植到Windows平台,并改写存档引擎的数据存储及检索功能,引擎采集到的数据不再按照特定格式以文本形式存入本地磁盘,而是用Oracle作为中央数据库集中存放。存档引擎从IOC采集到的数据经过缓冲区短暂保存后通过网络实时写入中央数据库。中央数据库提供强大的搜索计算功能,实现通用的数据管理,便捷了应用程序对数据的利用,并有利于以后的系统扩展。在中央数据库的基础上,为进一步提高数据通用性,系统采用Web Service的统一数据库接口,极大地方便了以后各种应用程序的开发与功能扩展。

2.1 存档系统的软件结构

改进后的存档系统软件结构如图 1。存档引擎启动后读取 xml 格式的配置信息,根据配置文件中提供的 pv(process variable)信息,与 IOC 建立起基于 CA 的通道连接。连接建立完成后,按照 scan、monitor 和 scan using monitors 三种模式向 IOC 发送数据采集请求。当接收到从 IOC 传回的数据时,存档引擎将数据临时存放于为各个通道配置的 buffer 中,再通过引擎与中央数据库的接口向中央数据库写入存档数据。

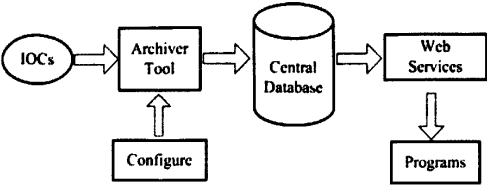


图 1 存档系统结构图
Fig.1 Archiving system.

2.2 存档数据库设计

存档系统中，中央数据库表的设计尤为关键，关系到数据管理效率及应用程序利用数据是否便捷。存档系统中的中央数据库由通道数据表、通道表和通道参数表组成。通道数据表(表1)为存档数据库主表，保存通道连接后采集的通道数据，用常规类型字段结合XMLType字段的方式，将通道类型、通道数据个数、报警状态等主要数据以常规类型字段保存，将不适合常规字段保存的特殊格式数据用XMLType字段保存。通道表是通道名称与通道id的对照表，用整型id代替以字符串形式表示的通道名，可提高存储和查询效率，并能有效节省存储空间。通道参数表保存每次通道连接后的通道参数信息。通道的参数信息通常不变，只在重新连接后变化，因此用一张表单独保存以节省空间。

表 1 通道数据表
Table 1 Channel data table.

字段名	字段类型	字段约束
Field name	Field type	Field constraints
pv_id	Number(5,0)	Primary key(1)
status	Number(2,0)	not null
dbr_type	Number(2,0)	not null
dbr_count	Number(2,0)	not null
svr_time	timestamp	Primary key(2)
severity	Number(2,0)	not null
stamp_sec	Number(10,0)	not null
stamp_nsec	Number(10,0)	not null
pv_value	Number(10,4)	
pv_xml	XMLType	

对一条通道数据表中保存的数据，可通过select语句查询pv_id为1021的通道数据而获得，各字段数值间以顿号隔开，分别对应表1中pv_id之后的字段：5、6、1、21-12月-09 03:46:24.942000 下午、2、349598253、630189716、2、<data>null</data>。

2.3 数据存档与数据库接口

存档系统中，中央数据库采用 Oracle 10g Enterprise Edition Release 10.2.0.4.0，运行平台为惠

普公司的ProLiant DL580 G3服务器，服务器操作系统为Centos 4.5。本存档引擎在Windows XP Professional操作系统下运行。EPICS本身没有Oracle数据库接口，本文用OTL作为存档程序与Oracle操控接口。OTL(Oracle, Odbc and DB2-CLI Template Library)是一个C++编译中操控关系数据库的模板库，其通过Oracle提供的OCI接口直接操作Oracle数据库，有运行效率极高、代码简洁、部署容易、无需.net framework支持等优点。OTL框架的实现是对流观念的实现，系统主要通过otl_stream操作数据，主要操作语句如下：

```
otl_stream os(1100, "insert into pvdata values(
    to_timestamp(:f1<char[24]>, 'yyyy-mm-dd hh24:
    mi:ss.ff6' ),
    :f2<int>;:f3<int>;:f4<int>;:f5<int>;:f6<int>;:f7<i
    nt>,
    :f8<int>;:f9<double>, sys.xmlType.createXML
    (:f10<char[101]> )))",
    db); //其中，f1-f10均为占位符
```

当OTL缓冲区被SQL填满时，OTL流将自动刷新缓冲区，批量执行缓冲区中所有的SQL语句。存档系统中的SQL操作主要为INSERT语句，缓冲区的大小将影响其执行时机及服务器磁盘I/O性能，合理设置缓冲区大小对存档系统有重要意义。

在配置为CPU主频2.4 GHz、内存2 G的机器测试插入10⁵条数据响应耗时，当缓冲区大小为1100条记录时，响应耗时5385 ms，当缓冲区大小分别为1150和1050时，响应耗时为5423 ms和5437 ms，继续增大或减小缓冲区，响应耗时会继续增长。可见，该机器上运行存档引擎的最佳缓冲区大小为1100条。

2.4 数据库接口设计

Web应用程序访问数据库的主流接口实现技术主要有CGI(Common Gateway Interface)和Web Service技术。CGI是Web服务器与CGI应用程序间信息传递的一种标准，几乎所有Web服务器都支持CGI标准。但由于CGI应用程序一般都是一个独立的可执行实体，与Web服务器各自占据不同进程，且一个CGI程序只能处理一个用户请求，大量的用户请求会导致极其可观的内存和CPU占用。Web Service是分布式计算领域发展的前沿技术，是一种可接收从Internet或其它系统中传递请求的轻量级、独立的通讯技术。Web Service采取简单易理解的标准Web协议作组件界面描述和协同描述规范，通过XML、

SOAP、WSDL和UDDI等技术^[5]和规则屏蔽不同软件平台的差异^[6], 有高度通用性和易用性、高度集成性及容易发布和部署等优点。Web Service能向外界提供一个通过网络调用的 API(Application Programming Interface, 应用程序编程接口)。Web Service发布后, 用编程的方法通过web调用这个应用程序, 而调用服务的应用程序就是Web Service的客户端。

Web Service 用面向服务的体系架构(SOA)^[7]。SOA 结构中共有三种角色: 服务提供者(Services provider)发布自己的服务, 并对使用自身服务的请求进行响应; 服务注册者(Services broker)注册已发布的服务提供者, 对其分类, 并提供搜索服务; 服务请求者(Services requester)用服务注册者查找所需的服务, 然后使用该服务。这些角色间通过发布、查找和绑定等操作相联系。

本系统采用 Web Service 提供应用程序访问中央数据库的统一接口, 存档系统采用的数据库接口中间件的 Web 服务体系结构如图 2。为实现 Web Service 中事务处理、安全性和身份识别等功能, 系统中用 XFire 技术开发 Web Services 应用程序。XFire^[8]是 codehaus 推出的设计上面面向文档的 Java SOAP 框架, 是新一代 Web Services 引擎。XFire 通过提供简单的 API 及支持标准协议, 方便地开发面向服务的应用程序。XFire 内建在 STAX 基于流的 XML 解析引擎基础上, 有很高的性能。XFire 框架是一种基于 Servlet 技术的 SOA 应用开发框架, 要在基于 XFire 应用框架开发的企业中正常运行, 需 JDK 和 Servlet 容器的支持。本系统采用 JDK 版本 5.0, 用 Tomcat(版本为 5.5.9)作 XFire 的运行容器, XFire 为 1.2.6 版本。

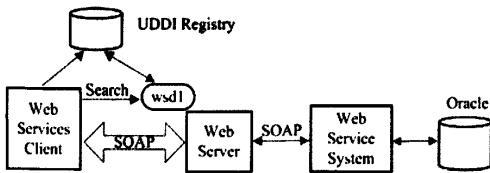


图2 Web 服务体系结构图
Fig.2 Web services architecture.

系统中用于生成 web services 应用 java 文件的主体为通过 JDBC 连接到 Oracle 的链接类, 其主要伪代码如下:

```

public class SSRFDB{
    private static Connection con = null;
    private static String DBdriver = "oracle.jdbc.
  
```

```

driver.OracleDriver";
    private static String SSRFurl = "jdbc:oracle:
thin:@10.30.1.16:1521:SSRFORCL";
    private static String SSRFuser = "chengh";
    private static String SSRFpassword = "chengh";
    public SSRFDB(){
    public int initDB(String sdriver,String url,String
suser,String spassword){...} //初始化
    public int Close(){...} //关闭数据库
    public String[][] execQuery(String sSQL){...} //将
查询的结果按照二维数组返回
    public int execUpdate(String sSQL){...} //执行没
有返回值的 SQL 语句
    }
  
```

3 总结

上海光源(SSRF)采用EPICS控制系统作软件平台, Channel Archiver作EPICS自带的数据库存档工具。由于SSRF数据量庞大, Channel Archiver采用的文本文件形式的本地存储方案及数据格式不能完全满足日常应用需求。新设计的存档系统采用Oracle作为中央数据库实现数据的实时存档及管理, 并用Web Service技术提供不同系统环境及编程语言下应用程序访问中央数据库的统一接口, 增强了系统的功能性与扩展性。存档系统中引擎的设计根据Channel Archiver源代码改写。经测试表明, 系统可长时间稳定运行。

参考文献

- 1 国家上海同步辐射装置中心(筹). 上海同步辐射装置工程初步设计, 2001
National Center for Shanghai Synchrotron Radiation Facility (plan). Shanghai Synchrotron Radiation Facility project preliminary design, 2001
- 2 EPICS-Experimental Physics and Industrial Control System. <http://www.aps.anl.gov/epics/>
- 3 Shanghai Synchrotron Radiation Facility. <http://ssrf.sinap.ac.cn/english/1/Introduction.htm>
- 4 祝 晴, 蒋 舸 扬, 李 林, 等. 核电子学与探测技术, 2007, 5(3): 521-522
ZHU Qing, JIANG Geyang, LI Lin, *et al.* Nucl Electron Detect Technol, 2007, 5(3): 521-522
- 5 W3C.org. Web Services Architecture. <http://www.w3.org/TR/ws-arch/>
- 6 郭海波, 吴 保 国. 电脑知识与技术, 2009, 11(31):

8628–8631

WU Haibo, WU Baoguo. Comput Knowl Technol, 2009,

11(31): 8628–8631

7 Kuassi Mensah. Oracle database programming using Java

and Web Services. USA: Elsevier Digital Press, 2006

8 Codehaus XFire. <http://xfire.codehaus.org/>

Oracle-based data archiver system and interface design for SSRF

ZHANG Zhiyuan CHEN Jianfeng SHEN Liren

(Shanghai Institute of Applied Physics, Chinese Academy of Sciences, Shanghai 201800, China)

Abstract SSRF, as a national scientific base in China to provide synchrotron radiations from IR to hard X-rays for hundreds of scientists working simultaneously on the experimental stations, needs an efficient software platform for the facility control and user service. It uses a distributed control system of EPICS, which is commonly used home and abroad in large scientific facilities. Channel Archiver is an archiving toolset for EPICS, which is used to store and retrieve different types of data generated by EPICS. Based on analysis of Archive System, we have found some ways to overcome the defects in existing, and designed a new archiver system using Oracle database. This offers a unified interface for varieties of programs.

Key words SSRF, EPICS, Channel Archiver, Oracle, Web Service

CLC TB533⁺.2