

# 基于纯 XML 数据库的 SCL 文件管理

邹小虎, 何 星

(上海交通大学自动化系, 上海 200240)

**摘 要:** 介绍 IEC61850 系列标准中的变电站配置描述语言的文件结构, 阐述系统的配置流程及各种数据文件的交换过程, 提出在 XML 文件管理维护上采用纯 XML 数据库技术。直接将 XML 文本文件作为存储单元。创建 2 个索引文件, 运用传统的 B+树数据结构建立索引。针对 XML 文件的特性, 采用 XML 路径语言 XPath 表达式的方法进行文件查询。

**关键词:** IEC61850 标准; 变电站配置描述语言; 纯 XML 数据库

## SCL Document Management Based on Native XML Database

ZOU Xiao-hu, HE Xing

(Department of Automation, Shanghai Jiaotong University, Shanghai 200240)

**【Abstract】** This paper introduces the document structure of the Substation Configuration description Language (SCL) in the IEC61850 standard. It expatiates the configuration flow of the system as well as the exchange process of the data document, and brings forward to use the Native XML Database(NXD) technology in maintaining the XML documents. It takes the XML text file as the store unit. The system creates two index file and uses the traditional B+ tree data structure to set up the index. It uses the XPath expression to query in the XML file.

**【Key words】** IEC61850; Substation Configuration description Language(SCL); Native XML Database(NXD)

### 1 概述

IEC61850 是国际电工委员会 TC57 工作组制定的《变电站通信网络和系统》系列标准, 该标准是基于网络通信平台的变电站自动化系统的唯一国际标准。它应用在变电站自动化的设计、开发、工程、维护等各个领域。该标准通过对变电站自动化系统中的对象统一建模, 采用面向对象技术和独立于网络结构的抽象通信服务接口, 增强了设备之间的互操作性, 可以在不同厂家的设备之间实现无缝连接, 实现了智能电子设备的互操作, 大大提高了变电站自动化的技术水平, 以及变电站自动化的安全维护水平, 节约开发验收维护的人力物力, 实现了安全的互操作性。

XML 是 Web 上表示结构化信息的一种标准文本格式, 它没有复杂的语法和包罗万象的数据定义。XML 同 HTML 一样, 都来自标准通用标记语言(SGML)。标记是指计算机所能理解的信息符号, 通过这种标记, 计算机之间可以处理包含各种信息的信息文章

### 2 变电站配置描述语言

变电站配置描述语言(Substation Configuration description Language, SCL)是IEC61850 标准中定义的用来描述智能电子装置(Intelligent Electronic Device, IED)的配置和通信系统。使用该语言, 可规范地描述变电站自动化系统和变电站(开关设备)本身的相对关系。该语言利用了XML的可扩展性, 是根据变电站配置特殊要求定义的一种行业专用描述语言<sup>[1]</sup>, 在语法上遵循XML的语法规则, 在语义上尽可能包含变电站配置所涉及各类对象。XML的跨平台性使得其文件内的数据可以在不同的平台和不同的系统内以某种兼容的方式进行交换。

在 IEC61850-6 标准中, 用 XML Schema 来定义 SCL 文件的结构, 标准中使用了 8 个 Schema 文件, 其中, SCL.xsd

是主文件, 它引用或包含了其他 7 个 Schema 文件。各 Schema 文件通过定义各种简单或复杂类型来定义 SCL 文件中各种元素及其属性值的类别, 保证了配置信息交换格式的一致性。

在用 SCL 进行 IED 远程配置的过程中, SCL 的数据交换有 4 种目的, 因此, 就有 4 种文件在不同工具之间进行数据交换。这 4 种文件以不同的后缀名区分, 但是每个文件的内容必须遵循 SCL 所定义的规则。

具体名称如下:

- (1) 用于描述变电站的配置信息的 .SCD(Substation Configuration Description)文件;
- (2) 用于描述变电站的规格静态信息的 .SSD(System Specification Description)文件;
- (3) 用于描述 IED 配置信息的 .CID(Configured IED Description)文件;
- (4) 用于描述 IED 功能信息的 .ICD(IEE Capability Description)文件。

### 3 系统配置过程

整个系统有 2 种配置器: 系统配置器和 IED 配置器。IED 配置器是在各个 IED 侧生成可以作为输入的 SCL 数据的工具; 系统配置器独立于各 IED, 它可以解析和生成系统配置文件, 描述 IED 和设备之间的关系。上文的 4 种文件, 将在如下的配置过程中介绍<sup>[2]</sup>, 整个配置过程如图 1 所示。

(1) IED 功能描述及出厂信息以及 IED 配置工具都是各个厂商提供的, 这些信息通过配置工具后整理出了 ICD 文件, 可以认为是每一个 IED 配置出一个 ICD 文件, 多个 IED 产生出多个 ICD 文件, 作为 SCL 配置器的输入, 这些文件都可以

**作者简介:** 邹小虎(1979 -), 男, 硕士, 主研方向: 电网软件, 电力数据库; 何 星, 副教授

**收稿日期:** 2007-09-28 **E-mail:** zouxiaohu0528@yahoo.com.cn

看作是静态文件。数据流向为 A1->A2->A。

(2)SSD 文件是描述变电站系统的文件,对于一个投入运行的变电站,站内设施是很少改动的,但是还应考虑到使尽可能多的 IED 设备兼容,包括新增 LN 节点等。因此,在变电站的运行过程中,需要手动增删某些信息。此 SSD 文件也将作为 SCL 配置器的输入,即 A + B->C。

(3)SCL 配置器通过输入的 SSD 文件和各个 IED 的 ICD 文件进行配置,生成了整个系统的 SCD 文件,该文件描述了各孤立的 IED 和变电站自动控制中心之间的关联关系。同时它又作为 IED 配置器的输入,产生一系列真正完整的 IED 功能描述文件——CID 文件。

(4)可以将各个 CID 文件下载到各个独立的 IED 之内,从而完成配置过程并作好通信连接的准备工作。

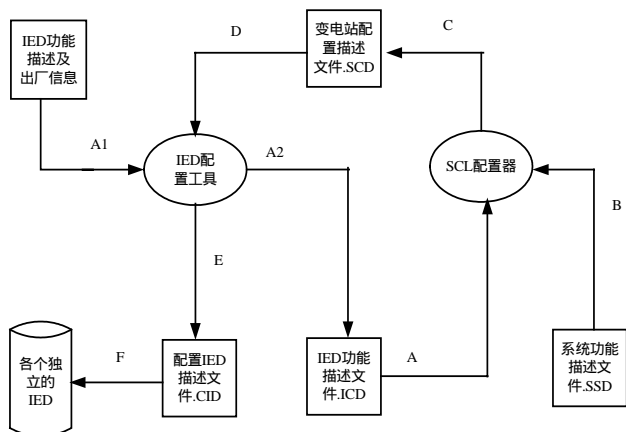


图1 系统配置数据流程

通过以上过程,可以将系统看作为一个 client-server 结构,各个独立的 IED 即为客户端,而配置工具和产生的一系列文件均置于 server 端,完成的功能是进行配置、产生出各种配置文件,以及对这些文件的存储、管理和维护。

## 4 文件管理方法

### 4.1 纯 XML 数据库技术的应用

在 server 端内存储着各个配置器产生的文件,这些文件在语法上实质均是 XML 文件。XML 数据是树型结构,当前有 2 种处理 XML 文件的方法<sup>[3]</sup>:

- (1)XML-Enabled DBMS(XED);
- (2)纯 XML 数据库(Native XML Database, NXD)。

XED 是在原有数据库基础上扩展了 XML 支持模块,完成 XML 数据和数据库之间的格式转换和传输。从存储尺度上,可以把整个 XML 文档作为 RDBMS 表中一行,或把 XML 文档进行解析后存储到相应的表格中。

NXD 出现在 XML 数据处理领域内,一般采用层次数据存储模型,保持 XML 文档的树形结构,省略了 XML 文档和传统数据库的数据转换过程。

一个公认的描述为<sup>[4]</sup>:

- (1)相对于文档的数据,它为 XML 文档定义了一个(逻辑)模型,并根据模型存储文件。
- (2)它以 XML 文件作为其基本(逻辑)存储单元,类似于关系数据库中以表中的记录作为基本(逻辑)存储单元。
- (3)对底层的物理存储模型无特殊要求。

因此, NXD 专门用于存储 XML 文档并且保持了其完整性,存储均以文档为基本单元。NXD 不是一种全新的数据库底层处理模式,它不取代现存的数据库,而仅仅用于帮助人

们更好地处理 XML 文档。

### 4.2 采用纯 XML 数据库技术的原因

分析将要存储的 SCL 文件,文件的特点是:

- (1)文件的语法格式(Schema)已经定出。IEC 已经规定了 SCL 文件的语法格式,这将有助于索引、查询等功能的实现;
- (2)SCL 对象建模采用的是面向对象的方法,继承的层次结构是多重的,难以将其二维化;
- (3)SCL 文件以数据为中心,注重的是文件内的信息而非文件的格式。

采用 XED 技术的优缺点如下:

- (1)需要根据 IEC61850 系统的层次体系结构建表,而其层次关系严谨而复杂,层层嵌套,使得建表的工作十分繁琐;
- (2)由于文件是以数据为中心的,因此将海量的数据导入至传统的关系型数据库的表内,工作量十分巨大;
- (3)由于是基于成熟的关系型数据库技术,因此在查询及索引方面效率较高。

采用纯 XML 数据库技术处理 SCL 文件有如下特点:

- (1)无须对 IEC61850 系统的层次体系结构建表,直接将 SCL 文件作为基本处理单元;
- (2)无须将 SCL 文件内描述的海量数据进行重新整理,直接从文件中获取;
- (3)由于是较新的技术,因此在查询方面还未有成熟的理论。但是针对 SCL 文件而言,已经给出了 Schema,并且所作的查询及索引均是在 SCL 文件基础之上完成的,因此极大地降低了处理该类文件的难度。

由于本文针对 SCL 这个特定的文件进行处理,并且考虑到文件的容量较小,因此采用纯 XML 数据库技术来进行这些文件的管理与维护,对这些文件建立索引,提供出查询及增删改的接口。

### 4.3 文件的索引技术

系统创建了 2 个索引文件如下:

- (1)dom.index: 存放在在页文件中的节点,并且做到将实际节点的唯一识别符与文件相关联;
- (2)element.index: 对 XML 文件中的元素和属性进行索引。

从数据结构来说,静态索引是个多分树(multiway tree);而动态索引是一种平衡多分树(balanced multiway tree),一般称之为 B-树。B-树有不少变种,常用的有 B+树。B+树大量地用在存储索引中,传统的关系型数据库中的动态索引大多都是采用 B+树技术。当前 B+树技术已成为数据库领域中较为成熟的技术。

在本系统中,所有的索引都是基于 B+树技术的。建立以上 2 个索引文件可以减少内部 B+树的页文件的数量,并可以提高整个文件内的查询性能。

XML 数据的存储(dom.index)是整个文件存储的核心部分。它由单页文件组成,在此文件中,所有的文档节点均按照 W3C 的文档对象模型进行存储。数据则来源于同一个文件中的多分 B+树。使用 B+树将给定文档中的上级元素的唯一节点标识符与数据块中的节点存储地址关联起来,具体如图 2 所示。

只有高级的元素使用 B+树进行索引,而属性、文本节点以及文件节点结构中低级的节点仅仅是写入数据页,并没有在 B+树中加入键。但是可以通过在树中寻找其最近祖先的方法,来提供一个查询这些类型节点的接口。采用此种存储技

术,就没有必要用指针去指向节点间的连接。DOM的实现完全依赖于<sup>[5]</sup>编号方法(numbering Schema)用以决定节点间的关系。例如,要获得一个节点的父节点,可以通过计算节点的标识符来获得,而相应的节点可以通过索引的查询获得。这样,在索引文件(dom.index)中的数据存储的大小将小于原先的数据源文件。

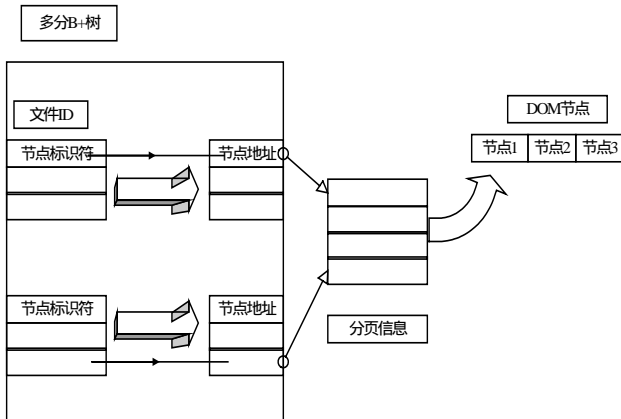


图2 XML数据存储组织

在索引文件 element.index 中,元素和属性名被映射成唯一节点标识符。每个索引中的入口由 1 个键组成,是 1 个二元组:

```
<document-id, name-id>
```

数组值包含了 document 的 id 和节点的 id,这与元素和属性匹配键中的标示名相对应。例如,要找出所有的在 Substation 内的 Device 元素,搜索引擎将会要求提供一个单一的索引查找以获得指向 Device 元素的所有的节点标识符。

#### 4.4 文件的查询技术

基于以上提及的索引方法,在查询的过程中可以采用 XPath 表达式进行查询。XPath 是一种用于对 XML 文档的元件寻址的语言,设计为 XSLT 和 XPointer。XPath 将一个 XML 文档建模成一棵节点树,有不同类型的节点,包括元素节点、属性节点和正文节点。由于所查询文件的 XML Schema 都已经确定,因此查询的效率很高。

一段示例 SCL 文件片断如下,它描述了变电站内 VoltageLevel 的信息。

```
<?xml version="1.0"?>
<SCL>
  <Header Ref="sclT1-1" Revision="A" NameStructure="
  IEDName"/>
  <Substation Ref="pingyaobian">
    <VoltageLevel Ref="D1" Voltage="220KV">
      <Bay Name="Q1">
        <LNode Ref="ln1" LNClass="PDIS" IEDRef="
        D1Q1KA3"/>
        <Device Ref="I1" type="CTR">
          <LNode Ref="1" LNClass="TCTR" LDInst="LD1"
          IEDRef="D1Q1KA1"/>
        </Device>
      </Bay>
    </VoltageLevel>
  </Substation>
</SCL>
```

```
<Connection CNodeRef="L1"/>
</Device>
...
</Bay>
</VoltageLevel>
</Substation>
...
</SCL>
```

要查找变电站内设备的 XPath 查询非常简单,即为 //LNode[@LNClass='TCTR'],意思为选择上下文节点的所有具有 LNClass 属性且值为 TCTR 的 LNode 节点。在输入 //LNode[@LNClass='TCTR']后,得到的查询结果如下:

```
LNode Ref="1" LNClass="TCTR" LDInst="LD1" IEDRef="
D1Q1KA1"
```

运用 XPath 表达式基本可以满足用户所提出的查询的需求,但如何将用户的查询自动转换成程序内可用的标准的 XPath 表达式仍有待研究。

XPath与SQL类似,都是声明性语言编写查询,但其编写复杂逻辑能力较弱<sup>[6]</sup>。可以通过 javax.xml.xpath 把 XPath 结合到 java 里去,这样就发挥了两者的优势,java 做处理逻辑,XPath 做声明查询。随着越来越多的数据转向 XML, javax.xml.xpath 与 java.sql 一样变得越来越重要。本系统做了试验性的管理方法,在 java 代码中直接嵌入 XPath 的表达式,之后还需提供规范的查询接口以供用户进行查询。

#### 5 结束语

本文根据 IEC61850 的数据模型定义,介绍了变电站内 IED 系统配置的过程,指明 XML 文件在整个协议中的作用。根据 XML 跨平台的特性,将其作为数据存储与交换的载体,使用纯 XML 数据库技术,根据变电站系统的特性,对系统内的 XML 文件进行维护和管理,做了索引和查询。

但本文对 XML 文件的处理仅是文件管理的层面,还未上升至 XML 数据库层面,对于数据库中的事务管理以及完整性约束等方面的问题未加考虑。如何规范文件的查询及增加删改是下一步的研究方向。

#### 参考文献

- [1] 邹晓玉,王浩. IEC61850 标准中 SCL 语言的几个实践应用问题探讨[J]. 电力系统自动化, 2006, 30(15): 77-80.
- [2] 李蓓,沐连顺. 基于关系模型映射的 IEC61850 SCL 配置器建模[J]. 电网技术, 2006, 30(10): 94-99.
- [3] 朱亮. Native XML 数据库技术详解[Z]. IBM DeveloperWorks, 2003-03-27.
- [4] 李骥,陈福生. Native-XML 数据库综述[J]. 计算机工程与设计, 2004, 25(6): 932-934.
- [5] Wolfgang Meier. eXist: An Open Source Native XML Database[C]//Proc. of Web and Database-related Workshops on Web, Web-services, and Database Systems. Berlin, Germany: Springer-Verlag, 2003.
- [6] Harold E. Java 语言的 XPath API[Z]. IBM DeveloperWorks, 2006-09-05.