XML for CIM Model Exchange

A. deVos, Member IEEE Langdale Consultants Narrabeen NSW 2101, Australia S.E. Widergren, Sr. Member, IEEE ALSTOM ESCA Corporation Bellevue, WA 98004 USA J. Zhu, Member IEEE ALSTOM ESCA Corporation Bellevue, WA 98004 USA

Abstract: Organizations responsible for secure power system operations need to model their systems and portions of neighboring systems in support of control and security functions. In the USA, the restructuring of the electric utility industry emphasizes the need to exchange operational system models for Independent System Operators (ISO) and Regional Transmission Organizations (RTO). These models need to be node/breaker oriented in order to meet the needs of control center applications. Unfortunately, the existing model exchange formats derive from planning models that are bus/branch oriented and lack detail required for control center operations. To support these data modeling exchange needs, NERC has adopted an approach that uses the semantic data definitions from the EPRI CIM with the syntax of XML to create XML files containing operational power system models. This paper presents the key aspects of XML and the CIM that make them excellent choices for addressing the operational model exchange needs of our industry.

Keywords: power system modeling, electronic data interchange, data models, power system control, transmission control, data management, data communication, software standards

I. INTRODUCTION

Electric utility organizations have long needed to exchange system modeling information with one another in order to construct simulation environments for power system economics and security analysis. The major motivation for this exchange has been to support system planning functions including transmission planning, maintenance scheduling, and operations planning. For proper analysis, significant portions of neighboring systems must be modeled in addition to an organization's own service territory. As these models tend to be very large, the equipment representations are simplified. For example, complicated substation switching schemes and equipment connections are generally reduced into bus/branch oriented models. Data acquisition and control equipment are also unnecessary to model in this case.

By contrast, the information needed for real-time power system operation requires far greater detail about the field equipment and its connectivity. These models must include the substation bus segments, switches, and measurement details. The resulting model is often referred to as a node/breaker model. Operational models are often initially built from bus/branch oriented planning models. These models are exchanged using planning model formats such as IEEE common format, WSCC format, or vendor formats (e.g., PTI PSS/E). Details of the immediate operating area are then added to these models to meet the energy management needs. Coordination of specific details of neighboring system models has been done on an "as needed" basis, usually with manual model updates.

With the advent of electric power deregulation, transmission operations must be open to promote fair competition among power utilities. To effectively coordinate transmission usage and assure reliable operation in the USA, independent system operators (ISO) and Regional Transmission Organizations These groups must maintain (RTO) are emerging. operational power system models that span multiple service areas to properly oversee safe and reliable operation of the transmission grid. To accomplish this, they must regularly exchange node/breaker detailed models with their member utilities. The planning model formats fall short of meeting this need. The result is that ISO/RTO modelers have resorted to the use of special conversion programs together with manual data manipulations that take node/breaker models from participating operating companies and mend together a super-model for the ISO/RTO's area of jurisdiction. As there are many proprietary formats for these control center models, the costs for initially building an ISO/RTO model are high and the on-going maintenance of model conversion is significant.

In parallel with these developments, the EPRI CCAPI Task Force and its counterpart in the international standards arena (IEC TC57 WG13 on EMS API) have been working on the specification for interfaces to facilitate the interoperation of electric utility software from independent sources. A significant achievement of this effort is the creation of a common information model (CIM) specifically for energy control center systems. This model meets the node/breaker level of detail needs for system operations.

In 1998, NERC began sponsoring what has turned into a series of meetings on Common Power System Modeling (CPSM). The purpose of these meetings is to address the operational model exchange needs of the North American electric utility industry. Early on, the CIM was identified as a good, vendor neutral choice for operations modeling. The problem is that the CIM is an abstract model; it is neither a modeling database specification nor an exchange format.

During this period, XML emerged as the dominant technology for encoding structured documents in new

applications. XML [4] is a markup language developed by the World Wide Web Consortium (W3C) and standardised by a W3C recommendation. It is now the format of choice for document-level data exchange over the public Internet and within many private networks.

As a result of these developments and the work of the authors, a common model exchange format based upon the CIM data definition and XML was proposed to NERC and subsequently adopted by their Data Exchange Working Group. In addition, the CIM XML format is going through the IEC process for standardization. Finally, all major vendors of energy management systems have voiced their support for the format and are presently working to make the promise of open model exchange a reality.

The following chapters discuss the design of the CIM XML language in relation to other XML-based languages and other CIM-based standards.

II. MOTIVATION FOR AN XML APPROACH

The CIM XML language is one of a large number of XMLbased languages that have been developed for various purposes since XML itself emerged. Applications of XML include languages dedicated to particular software tools, languages for horizontal applications such as graphics, and vertical applications for particular industries. There are now an increasing number of industry groups codifying data exchange formats in terms of XML resulting in standards such as HL7 for the health care industry, CML for the chemical industry, and OFX and OTP for the financial and retail industries.

A pragmatic reason for using XML in all these areas, and for power system model exchange in particular, is the availability of tools and libraries. Moreover, XML is accompanied by an extensive technology infrastructure covering functions such as transformation, presentation, query, schema and exchange protocols.

However, common adoption of XML has other benefits. It reduces the time and effort required to learn different systems. It imparts a degree of compatibility to different industry standards against the day when they come into contact with each other in the enterprise or on the public Internet.

III. DEFINING AN XML-BASED LANGUAGE

Each new application of XML involves a design process in which application-specific vocabulary and syntax are defined.

A. Vocabularies

XML stands for eXtensible Markup Language. Two of its antecedents are the Standard Generalized Markup Language (SGML) and Hypertext Markup Language (HTML).

Because of the web, HTML is the most familiar. In contrast to HTML, XML is generic. While HTML defines a fixed vocabulary of tags with which to create web pages, XML supports flexibly defined vocabularies and is not limited to one application. In this, it resembles SGML. However, XML is a simpler, more lightweight language than SGML.

To illustrate the principle of definable vocabularies and the general appearance of XML, Figure 1 shows an XML document that describes an IEEE PES meeting.

Figure 1. A Sample XML Document

The vocabulary used here includes the words "meeting", "where" and "country". At some level, each XML application needs to introduce a vocabulary like this. The CIM XML language introduces a power system oriented vocabulary that includes "transformer" and "breaker". These vocabulary items are drawn from the CIM schema.

B. Syntax and Semantics

The example document also follows a syntax. At a basic level it follows XML syntax which gives it a regular, hierarchical structure of elements and attributes. The construction beginning with <where> and ending with </where> is an element, and the construction ID="SUM'2000" is an attribute. When a document follows XML syntax, it is said to be *well-formed*.

To be useful, the document must follow an application-level syntax as well. For example, the <country> element must appear within the <where> element and not the other way around. Both must be contained by the <meeting> element.

Finally, the document author and the reader must agree on the meaning of things. They must understand that "meeting" means an IEEE PES meeting, "country" denotes a nation and "where" connects the two in the sense of a subject and predicate. In our example, the semantics seem obvious to a speaker of English and need little attention. However, this is not always true when a large technical vocabulary is deployed as in the CIM XML language.

C. Document Type Definitions and Other Specifications

In most applications it is essential to provide a precise specification of the vocabulary and syntax used. This requirement is most obvious when communication between independent parties is involved, such as the different transmission system operators who will use the CIM XML language. An XML Schema or Document Type Definition (DTD) can often provide the necessary specification, as shown Figure 2.

```
<DOCTYPE pes_meeting [
   <!ELEMENT meeting (where, when)>
   <!ELEMENT where (country, city)>
   <!ELEMENT when (month, date, year)>
   ...
]>
```

Figure 2. A Sample Data Type Definition (DTD)

Using DTD notation, Figure 2 establishes the vocabulary and syntax of a class of documents, of which our specimen in Figure 1 is an example. An XML processor can now read the DTD, together with a PES meeting document, and determine if the latter conforms to the former. A document that conforms to its type definition is said to be *valid*.

However, DTD's are not suitable for all XML applications. As we will see, the CIM XML language requires a more datacentric specification technique, in place of the DTD.

D. A Document Exchange Scenario

Once an XML language is agreed among its users, whether by DTD or some other means, documents can be exchanged. One exchange scenario is pictured in Figure 3. Here it is envisaged that the ultimate sources and destinations of the data are databases. The scenario also shows a web browser for visual presentation.

Two transformation steps are shown as part of the document transfer sequence (labelled XSL trans). XSL is a language for specifying transformations of XML documents [7] and an XSL processor would be one way to bridge any differences between the agreed language for exchanging information and the language expected by the end systems.



Figure 3. XML-Based Data Exchange Mechanism

E. Large Schema in XML

Compared to many applications of XML, the CIM XML language deals with a very extensive data schema. Translating the CIM schema directly into a DTD presents several problems.

- 1. The CIM assumes an entity-relationship view of data in which instances are nodes in a directed, labelled graph (DLG). However, XML provides a hierarchical structure. There are no general facilities in DTD notation to define a DLG.
- 2. The CIM is subject to change, in part because of its size and scope. However, most of the model remains unchanged between versions and there is a need for applications to recognise the compatible subset between versions. A DTD does not provide the necessary version control.
- 3. A CIM XML document may need to carry extended data, beyond that standardized by the CIM. However, the extended data must not interfere with recognition of the other content. Moreover, the extended data should be recognizable at the level of entities and relationships, even if the relationship names and entity types are unknown. Again, there is no general way to specify this with a DTD.

Similar comments apply to the successor of the DTD, XML Schema.

F. RDF – An XML Knowledge Representation Language

The W3C recommendation entitled Resource Description Framework or RDF has addressed the general problem of representing entities and relationships, that is a Directed Labeled Graph (DLG), in XML. The RDF specifications contain three components: data model, syntax, and schema [5,6].

The RDF data model is drawn from Knowledge Representation. It is a simple and general view of information and therefore relatively easy to project onto other models. Once that is done, RDF syntax can be used to encode the information and RDF schema can be used to describe or constrain it.

In the RDF model, a *resource* is anything that can be identified. A Uniform Resource Identifier (URI) is used to designate a resource. A *property* is any characteristic of a resource that can be described with a *value*. The triple: (resource, property, value) is the atomic unit of information in RDF and is called a *statement*. The value in a statement can be a literal, such as a string. It can also be another resource, thus statements form the arcs in a graph.

Figure 4 shows a diagram of a simple RDF statement. In this diagram, the oval node represents the resource, the arc represents the named property, and the rectangle represents the value of the property.



Figure 4. A Sample of RDF Statement

RDF uses XML as its syntax. Figure 5 shows a sample of an RDF XML document used to describe the above RDF data model. As shown in this figure, each tag has a namespace as its prefix. For example, tag <Author> has namespace des as its Namespace des is specified with an URI at the prefix. beginning of the document. This allows different vocabularies to be combined in the same document. In the case of CIM XML, there might be extensions and multiple versions of the CIM vocabulary. In this example, for instance, the prefix des implies that the property Author is defined in a particular schema. Other tags, such as <Description>, are defined in the rdf namespace.

Figure 5. A Sample of RDF/XML Schema Definition

IV. CIM: COMMON INFORMATION MODEL

As an initial step to create an open data exchange environment in the EMS industry, parties exchanging information must agree on the definition of common power system entities and their relationships. The Common Information Model (CIM) is a cornerstone of the EPRI CCAPI effort to facilitate the integration of independently developed software components into energy management systems (EMS). The CIM specifies common semantics for power system resources, their attributes and relationships. As a result of several years of effort with contributions from industry experts across the world, the CIM has matured to the point where it is gaining wide recognition in the EMS arena.

The CIM [1] provides a comprehensive, logical view of EMS information for transmission network analysis, generation control, SCADA, and operator training simulation. The CIM is documented as a set of class diagrams using the Unified Modeling Language (UML). UML specifies the CIM in an abstract manner that allows for open implementations (i.e., there is no restriction to relational or object oriented or other modeling technologies). Figure 6 shows a fragment of the CIM class diagram in UML notation.

As shown in Figure 6, the base class of the CIM is the PowerSystemResource class, which is defined to represent a generic power system component. Derived from this abstract class are a variety of subclasses representing various power system equipment entities, such as lines, capacitors, breakers, transformers, and substations. Relationships between classes are also represented including resource ownership, groupings into substations, etc. The CIM systemically names each class, its attributes and relationships, thus creating a common data dictionary that facilitates system and application integration in the EMS industry.



Figure 6. A Fragment of the CIM Wires Model

To better manage this large model, the CIM is divided into "packages". The Core package describes classes common to nearly all parts of the model. The Wires package contains information on the electrical characteristics of transmission and distribution networks. The Topology package describes connectivity information. Together, these packages substantially represent the modeling information needed for network security analysis, and are relevant for ISO/RTO model exchange.

V. CIM XML LANGUAGE

The CIM XML language is an application of RDF to CIM. It is defined by a confluence of the CIM, RDF schema, and RDF syntax specifications.

A. CIM RDF Schema

An RDF schema [2] has been defined by codifying the CIM's abstract model with the RDF schema vocabulary. Since RDF is general enough to describe UML concepts, see [10], the conversion is straightforward. Figure 7 illustrates some of the correspondences.



Figure 7. Converting CIM Object Model to a CIM RDF Data Model

Resources correspond to objects, properties correspond to object attributes, and relationships such as inheritance are represented by RDF schema properties such as subClassOf.

Figure 8 shows a fragment of the CIM RDF schema, in RDF syntax, corresponding to figure 7.

```
<rdfs:Class rdf:ID="Switch">
    <rdfs:label>Switch</rdfs:label>
    <rdfs:subClassOf
rdf:resource="#ConductingEquipment"/>
</rdfs:Class>
<rdfs:Class rdf:ID="Breaker">
    <rdfs:label>Breaker</rdfs:label>
    <rdfs:subClassOf rdf:resource="#Switch"/>
</rdfs:Class>
 <rdf:Property rdf:ID="Switch.NormalOpen">
   <rdfs:label>NormalOpen</rdfs:label>
<rdfs:domain resource="#Switch"/>
    <rdfs:range rdf:resource="#Boolean"/>
</rdf:Property>
<rdf:Property rdf:ID="Breaker.AmpRating">
   <rdfs:label>AmpRating</rdfs:label>
<rdfs:label>AmpRating</rdfs:label>
<rdfs:domain resource="#Breaker"/>
<rdfs:range rdf:resource="#Real"/>
</rdf:Property>
```

Figure 8. A Sample of CIM RDF Schema Definition

This fragment uses standard RDF schema vocabulary such as "domain" and "range" to define CIM items such as "Breaker.AmpRating". In the full version [2], the standard RDF schema vocabulary is extended to represent additional UML association concepts such as inverse roles and multiplicity. A virtue of the RDF schema vocabulary is that it is designed for this type of extension. This enables the CIM to be translated from UML to RDF schema with sufficient fidelity.

The final result is a concrete schema, encoded in RDF syntax and employing RDF concepts that software tools can readily interpret. Moreover, instances of the classes in this schema have a well-defined representation in RDF syntax.

B. CIM XML Document



Figure 9. XML-Based EMS Data Exchange Mechanism

With an agreed upon CIM RDF schema, an EMS power system model can be converted for export as an XML document, see Figure 9. This document is referred to as a CIM XML document. All of the tags (resource descriptions) used in the CIM XML document are supplied by the CIM RDF schema. The resulting CIM XML model exchange document can be parsed and the information imported into a foreign system. By choosing XML, implementers are able to make use of a growing set of development tools to facilitate the creation of import and export software.

C. A CIM XML Prototype

To verify the XML-based model exchange proposal, a case study was conducted to import and export EMS data. A software program exported a 60-bus power system model from a vendor's database system to create a CIM XML document. This document was then successfully parsed and the information browsed by a program developed by an independent party. The speed of the development effort that supported this simple test of interoperability demonstrated the effectiveness of the approach.

Figure 10 shows a section of a prototype CIM XML document. It describes a substation, called "East", owned by a company called "BPA" that contains a normally-open circuit breaker made by "Admirable Electric". Each resource has a unique identifier (ID) associated with it. The circuit breaker description is embedded in the substation element, while the company is described as a separate element and referenced by the substation with its ID. As we can see, tags in this document are always prefixed by the cim namespace, indicating that these tags are defined in CIM RDF schema.

Figure 10. A CIM XML Document Prototype

D. Simplified Syntax

RDF syntax provides many ways to represent the same set of data. For example, an association between two resources can be written with a resource attribute (as per the cim:MemberOfCompany element in figure 10) or by nesting one element within another (as per the cim:Contain element). This could make it difficult to use some XML tools, such as XSLT processors, with the CIM XML document.

Therefore, a subset of the RDF Syntax has been proposed [3] for use in CIM XML documents. The aim of this syntax is to make it easier for implementers to construct serializers and deserializers, as well as to improve the effectiveness of general XML tools when used with CIM XML documents. The proposed syntax is a proper subset of the standard RDF syntax. Thus, it can be read by existing RDF deserializers. In this, it differs from some other proposals for a simplified RDF syntax, such as 8], [9].

VI. STANDARDS FOR MODEL EXCHANGE

To benefit the electric utility industry, the CIM XML model exchange format needs to be recognized and widely supported. To generate acceptance, the CIM XML format was formally proposed to and adopted by NERC as a recommendation for model exchange between transmission security coordinators. In support of this decision related standards initiatives are underway.

The Object Management Group's (OMG) has adopted an interface standard [11] to query power system model information, called the Data Access Facility (DAF). Like, the CIM XML language, the DAF is based on the RDF data model and shares the same CIM schema codified in RDF. While CIM XML enables a model to be exchanged as a document, DAF enables an application to navigate among its resources and properties. Because the two standards have a common basis, implementations should benefit by sharing large parts of their data handling logic.

Initiatives have also been taken in the IEC international standards body. In TC57 WG13 on EMS API, the CIM RDF schema definition and the CIM XML exchange format are going through the standardization process. This work is progressing in concert with EPRI sponsored CCAPI Task Force activities. A group of interested parties, including major EMS vendors and modeling tool suppliers, have been meeting to resolve the details of the format and CIM interpretation issues. The objective is to hold interoperability tests, that will demonstrate successful interpretation of CIM XML documents imported into and exported from independent parties.

VII. CIM VERSIONS & EXTENSIONS

Though the CIM has reached a level of maturity, it is still a work in progress. At this writing, CIM version u08b is undergoing revisions. The CIM will continue to be revised over time to correct errors, incorporate improvements, and include extensions desirable to standardize. In fact, nearly every implementation will require a means to address customized modeling not available in the CIM.

A. CIM Versioning

To unambiguously specify the CIM XML version, a Unique Resource Identifier (URI) is used for the cim namespace. A CIM XML document declares the cim namespace with the version being used in a statement such as the following:

xmlns:cim="http://www.iec.ch/tc57/schema/cimu08b"

Software reading the CIM XML document can then detect the CIM XML version and respond accordingly.

B. Model Extensions

The X in XML stands for eXtensible. This means that CIM XML documents can be extended to model vendor or utility

special needs. The CIM RDF schema can be extended with new classes and attributes by providing a separate namespace. Because a separate namespace is used, the customized CIM XML documents clearly delineate what is CIM standard and what is custom. Several different custom extensions can exist and be clearly identified within the same XML document. When these customized documents are imported to information systems that know nothing about the extensions, the elements with the unknown tags can be simply ignored. The following declaration identifies an extended namespace, bpa.

In our prototype CIM XML document, we can add a non-CIM attribute, OriginalPO, to the breaker class, as shown below. These customized tags for BPA can be simply ignored if a system import program is not interested in such extensions.

```
<cim:Contain>
<cim:Breaker ID="2"
cim:PowerSystemResourceName="11023"
cim:Manufacturer="Admirable Electric"
bpa:OriginalPO="PO123123123"
cim:NormalOpen="true"/>
</cim:Contain>
```

The RDF schema corresponding to this extension can be added to a separate RDF schema document thereby keeping the CIM RDF schema clearly separate and allowing each to evolve independently.

VIII. CONCLUSIONS

XML has gained broad acceptance as a language to facilitate enterprise information exchange, and its usage is expanding. With its extensibility and flexibility, the exchange of XML documents provides a lose coupling approach for integration across disparate systems. The fact that it is a standard embraced by the Internet and the software industry means that it has strong development tool support, familiarity with programmers, and an attractive future. The CIM is also gaining acceptance as a standard in the electric utility industry. The combination of these two standards yields a powerful approach for meeting the demands of regional transmission organizations in the restructured energy environment. Besides addressing these needs, the CIM XML approach holds promise for meeting other information exchange needs such as communicating power transactions or transmission reservations.

IX. ACKNOWLEDGEMENTS

The authors wish to recognize all of those involved in the CCAPI Task Force and IEC TC57 WG13 for their considerable efforts in the creation of the CIM. In particular, we applaud Leila Schneberger for her contribution in representing the CIM in UML and creating software that mechanically generates CIM RDF schema. We also want to acknowledge the role of Margaret Goodrich in working with

the appropriate committees of NERC on the CIM XML format to recommend this approach.

X. REFERENCES

- "Common Information Model (CIM) for the Control Center Application Program Interface", EPRI, Sep., 1998
- [2] "CIM RDF Schema" Exported from Ccapi.mdl Version: CIMU08b, Leila Schneberger, http://www.cim-logic.com/cim-rdf/CIM-schemacimu08b.xml
- "Simplified RDF Syntax for Power System Model Exchange", Arnold deVos, October 2000, http://www.langdale.com.au/DAF/ PSModelExchange.pdf
- [4] "Extensible Markup Language (XML) 1.0", W3C Recommendation, Feb.1998, http://www.w3.org/TR/REC-xml
- [5] "Resource Description Framework (RDF) Model and Syntax Specification", W3C Recommendation, February 1999, http://www.w3.org/TR/REC-rdf-syntax/
- [6] "Resource Description Framework (RDF) Schema Specification", W3C Candidate Recommendation, March 2000, http://www.w3.org/TR/rdf-schema/
- [7] "XSL Transformations (XSLT)Version 1.0", W3C Recommendation November 1999, http://www.w3.org/TR/xslt/
- [8] "A Strawman Unstriped Syntax for RDF in XML", Tim Berners-Lee, November 1999, http://www.w3.org/DesignIssues/Syntax
- [9] "Simplified Syntax for RDF", Sergey Melnik, November 1999, http://www-db.stanford.edu/~melnik/rdf/syntax.html
- [10] "A Discussion of the Relationship Between RDF-Schema and UML", Walter W. Chang, August 1998, http://www.w3.org/TR/1998/NOTErdf-uml-19980804
- [11] "Utility Management System Data Access Facility, Revised Submission", ALSTOM ESCA, OMG TC Document utility/99-11-02, November 1999

XI. BIOGRAPHIES

Arnold deVos (M' 96) received his BE(Elec.) from the NSW Institute of Technology, Sydney, Australia in 1981. He worked in system operations at the NSW Electricity Commission and developed software for load forecasting and operational data handling. He also developed inter-utility data exchange standards for the S.E. Australian interconnection. Subsequently he joined Megadata P/L (now part of Logica, UK) where he was the architect of the MOSAIC SCADA system, and developed its distributed database and user interface. In 1995 he joined ALSTOM ESCA where he developed EMS power system modelling tools. He is now a partner in Langdale Consultants where he consults on system integration issues and is the author of the CORBA standard interface for EPRI CIM models.

Steve Widergren (M' 78, SM' 92) received his BSEE and MSEE from the University of California, Berkeley in 1975, '78 respectively. He developed and maintained programs for system planning and control center operations at Pacific Gas & Electric, and American Electric Power. For the past 18 years he has been associated with ALSTOM ESCA contributing to the areas of network security analysis, project delivery, and productization of the firm's EMS product line. He is involved in developing tools to address power system modeling issues, and in the design of software solutions for operation in the restructured energy environment. He is a vice-chair of the

PES Energy Control Center Subcommittee, a member of the EPRI CCAPI Task Force, and a USA technical expert to IEC TC57 WG13 on EMS API.

Jun Zhu (M' 94) received his B.S. degree from Huazhong University of Science & Technology, Wuhan, China, in 1986, his M.E. degree from Nanjing Automation Research Inst. (NARI), Nanjing, China, in 1989, and his Ph.D. degree from Clemson University, Clemson, SC, in 1994, all in Electrical Engineering. Presently, he is working with ALSTOM ESCA in the areas of power system modeling and distribution management (DMS). Previously, he developed distribution analysis software for Power Technologies, Inc. (PTI). He is a member of IEC TC57 WG 14 on DMS Interfaces.