

Semantic Web Technologies for Power Grid Management

Mathias Uslar

OFFIS – Institut für Informatik
Escherweg 2
26121 Oldenburg
uslar@offis.de

Nikolai Dahlem

Universität Oldenburg
Abteilung Wirtschaftsinformatik
26111 Oldenburg
dahlem@wi-ol.de

Abstract: Utility companies worldwide are facing a multitude of new challenges, which can not be met with the historically grown, monolithic IT systems currently in use. Service oriented architectures are heavily used to integrate heterogeneous systems. One key aspect of this integration is the use of a common language, such as the Common Information Model (CIM). Important use cases are exchange of topologies and storage of topology histories. This paper describes the development and implementation of CIM Topology Store (CTS). CTS introduces new difference models for CIM/XML-serialized topologies.

1 Interoperability in the energy utilities domain

Various reasons result in a growing need for energy distribution companies to exchange data between them. One reason is the close coupling and the physical interdependencies between networks operated by different companies. Other reasons are challenges posed by new developments like the legal unbundling laws passed in 1998 by the EU. These laws will force energy utility companies to disassemble their monopolistic functions like transport und distribution of electricity and reorganize them competitively. EMS systems often use proprietary data exchange formats and must be connected to other systems. Building specialized adaptors for interconnection between the systems is the most common and time-consuming task for IT-departments at utilities. The Common Information Model (CIM) [IEC03] establishes a common language and domain model for energy management systems and related data structures. Our solution adopted the CIM as a common language.

2 Common Information Model IEC 61970

A data exchange format should be suitable for the data interchange with other companies. Therefore, it is beneficial to use an approved standard instead of proprietary formats used previously. It seems as the Common Information Model (CIM) is going to be the most common data exchange format for energy markets in the future. The CIM originates from a project by the EPRI Institute starting in the mid 1990s. It evolved to an International Standards being under maintenance by the International Electrotechnical Commission IEC, standardized within the TC 57 framework. Core elements of CIM have been adopted in International Standard IEC 61970, other elements are currently drafts or committee drafts for vote and will be standardized as IEC 61970-302 and -303. The CIM is an extensive ontology for the domain of electric utilities and is available in different formats, e.g., in Rational Rose UML, XMI and RDF Schema. The model has recently also been released in Web Ontology Language OWL. The semantic dimensions of the CIM and its implications on other standards are discussed in [Us105]. This contribution is going to focus on the use of the CIM in order to facilitate the exchange of power grid data. As CIM does not offer any official database model and data exchange is based on the RDF representation of CIM (called CIM/XML), it seems consequential to establish an RDF-based database instead of a relational database schema. This contribution is going to discuss the CIM and its special application from the perspective of RDF-based systems. Next, we are going to present the main use cases we have identified when dealing with a topology exchange.

3 Use Cases

We identified several use cases for the exchange of topology information for power grids between energy distribution companies:

Exchange of complete topologies at regular intervals: Hitherto, data is exchanged using proprietary formats, which leads to high complexity. Adapters have to be implemented and maintained. Due to this, it is only feasible to exchange data on a quarterly basis. Additional time and effort is spent on validation of complete topologies by experts, where only the changes from the last data exchange are of interest.

Exchange of parts of topologies: In order to be able exchange parts of topologies, i.e. voltage levels, substations, a flexible mechanism is needed to extract required information from topologies.

Merging of independently created topology parts: The merging of multiple, independently created topology parts into one consistent topology is a major issue for the collaboration between utilities. Imagine a scenario where multiple regional utilities cooperate with a grid provider. The grid provider needs to merge the topologies provided by the regional utilities into one consistent topology in order to evaluate it. This process raises questions regarding responsibility for topology parts and consistency.

Traceability of topology evolutions: Power grids are constantly evolving and changing their state. These changes must be saved and made accessible.

In order to implement these use cases, we have to take several existing contributions to the CIM into account, mainly the theoretical model developed by Britton on regional model authorities which will be presented in the next section.

4 Regional Model Authorities

The regional model for dealing with CIM topologies was originally developed by Britton as a theoretical model [Brit06]. It is used to evaluate the responsibilities and interfaces between different utilities having to interconnect their grids while still adhering to requirements from their own infrastructures. We have to distinguish between three types of regions. The first region, the base region, is directly under control and responsibility of utility. The so called boundary region, the second region, forms a border between different base regions. All changes within a boundary region have to be validated by the dispatchers in charge of the associated base regions. Resources from one region may only have associations to resources belonging to the same region or to associated base regions. Aggregated regions, our last type of region, are defined being as the aggregation of a number of base regions with their respective boundary regions. We use this model in the CTS system for our versioning and authorisation system for topologies. The topologies are stored and responsibilities for certain base regions are assigned to different users. They are held responsible for their individual base region and have to communicate with the other operators when they have to commit a change either to a boundary region or another base region. This model is going to be combined in CTS with a reasonable model for both difference format and versioning for the CIM described in the next sections.

5 CIM-Difference Format

It is inefficient to transfer complete topologies. To overcome this inefficiency a difference format for CIM-topologies needs to be developed. DeVos proposed a difference format for the CIM [deV02]. Unfortunately, this format has some drawbacks. It doesn't adhere to the RDF standard and extends the standard with a new 'parseType'. Furthermore, it uses a syntax with insertions and deletions which is unintuitive for users, because most changes to topologies are modifications, resulting in one insertion and one deletion. Additionally deVos only specifies a format, but no method to compute the difference between two RDF-graphs. To compute the differences between two RDF-graphs we used the method described in [BLC06]. As a replacement we developed two formats, one using insertions and deletions and one using replacements (cf. [BLC06]). Furthermore, we included pre- and postconditions as an enhancement in our formats. Both formats are compliant to the RDF standard and use reification viz. make statements on statements, i.e. marking a statement as an insertion or precondition.

During our work we compared the two difference formats and found that for the major percentage of real-world examples our second format, called 'dahlem2' (depicted in figure 1), using the replacement syntax proved to be superior to the format proposed by deVos and our first format.

```

<rdf:Statement>
  <rdf:subject rdf:resource="http://cim.dahlem.eu.org/#
    _242C8FDFE7334E9BA603449C87D89903"/>
  <rdf:predicate rdf:resource="http://iec.ch/TC57/2003/CIM-schema-cim10#
    EnergyConsumer.pfixed"/>
  <rdf:object>30</rdf:object>
</dm:replaces>
  <rdf:Statement>
    <rdf:subject rdf:resource="http://cim.dahlem.eu.org/#
      _242C8FDFE7334E9BA603449C87D89903"/>
    <rdf:predicate rdf:resource="http://iec.ch/TC57/2003/CIM-schema-
      cim10#EnergyConsumer.pfixed"/>
    <rdf:object>0.2</rdf:object>
  </rdf:Statement>
</dm:replaces>
</rdf:Statement>

```

Figure 1: Format *dahlem2* instance excerpt

6 CIM-Versioning

Based on the difference formats we developed a version management system, which features regional model authorities as described in section 4 and combines them with user and role management. The version management system supports the difference format as proposed by deVos as well as the two formats described in section 5. This ensures compatibility with existing systems on the one hand and offers the possibility to use our new, standards-compliant formats on the other. Additionally the system supports queries using SPARQL. This enables the user to select topology parts using the full power of a RDF query language. The query component is integrated with the RMAs and queries can be executed on all three types of regions.

7 Implementation and Evaluation

We have implemented the solution using a distributed client/server architecture using web services. It is a classic back end system. The overall solution can be integrated into the utilities service-oriented architecture using web services. This was done using Apache Tomcat and Apache Axis. The overall data for both the topologies and meta-data is kept in a MySQL database. For saving RDF, we used the HP Jena framework. HP Jena is a good API to get access to RDF graphs. It provides options to save the data as a flatfile, in memory or using different relational database systems. SPARQL has been implemented using a command line client for testing purposes. A complete installation and implementation was done using a VMWare server as a developing system and reference installation for distribution. The EPRI sponsors the so called interoperability test. By participating in those tests, the vendors of CIM based systems ensure that their solutions have proper interoperability.

There is a special suite for testing the exchange of topology data and the needed difference models. Overall, our solution has successfully imported three main topologies, the SmallModel by PsyCor, the Siemens PTI 100 and the Areva 60 Bus files and passed for the implemented difference models. Furthermore, there are functionalities to validate the models against existing schemes like the Common Power System Model (CPSM) using the HP Jena Eyeball inspectors. Overall, the system fulfils the following requirements elicited from utilities and vendors like import and export of topologies, creating topology parts, merging of topologies, saving of topology histories, tracking meta-data like creator of a region and search functionality for topologies

8 Conclusions

As of today, there is no other solution addressing the particular problem. One solution was the first to combine the both different requirements imposed by the users and providing a RDF-compliant way of a difference model format. Our solution is mainly based on a pure RDF-syntax which has proven to be more robust and understandable even by humans than the existing approach by deVos [deV02]. Our approach is more RDF-compliant and does also adhere to the CIM Users group interoperability test suite. The difference model has been improved in various ways; it is more robust by using plain RDF-syntax adding pre and post conditions and more information than the existing format. For the CTS System, we have added a system for version control and regional modelling authorities to the difference formats. Additionally, it is possible to validate the CIM topologies against various schemes like CPSM. The data is stored using RDF and can be queried using SPARQL, different versions of a single region and different regions can be combined within one single query. Currently, much of the overall performance is influenced by the general speed of HP Jena, its JDBC driver and the MySQL database. Currently, SPARQL does not provide a solution for changing the data, being a query language and no full data manipulation language. We currently have some kind of technology breaches because of using standard SQL for the relational-based versioning data. On the other hand, we are going to improve the user interface for the provided server based on Web Services.

References

- [BLC06] T. Berners-Lee and D. Connolly: Delta: An Ontology for the Distribution of Differences between RDF Graphs, <http://www.w3.org/DesignIssues/Diff.html>, 2006
- [Brit06] J. Britton: Designing Model Exchange Processes with CIM and 'RMA Sets', In: Proceedings of the IEEE PSCE 2006 meeting, Panel-19: Progress with the CIM Standard, 2006
- [deV02] A. deVos: RDF Difference Models - Representing the Difference between two RDF Models, <http://www.langdale.com.au/CIMXML/DifferenceModelsR05.pdf>, 04 2002
- [IEC03] IEC 61970-301: Energy management system application program interface (EMS-API) - Part 301: Common Information Model (CIM) Base, Technical report, IEC - International Electrotechnical Commission, 2003
- [Usl05] M. Uslar: Semantic Interoperability within the Power Systems Domain, In: Proceedings of the ACM Conference on Information and Knowledge Management CIKM 2005 Bremen, 2005