CTS - CIM Topology Store

Nikolai Dahlem
University of Oldenburg
Business Engineering
Oldenburg, Germany
dahlem@wi-ol.de

Mathias Uslar
OFFIS
Business Information Management
Oldenburg, Germany
uslar@offis.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 and based on them implements a version management system, which facilitates easy exchange of topologies and cooperation between utility companies.

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 dissemble their monopolistic functions like transport und distribution of electricity and reorganize them competetively. Another new development is the success and increased usage of renewable energy sources, which leads to decentralized generation of power, which in turn poses new challenges to the power grid and supporting energy management systems (EMS).

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) [7] establishes a common language and domain model for energy management systems and related data structures [17]. Our solution adopted the CIM as a common language.

Our contribution is structured as follows. Section 2 will provide a short overview on the Common Information Model CIM providing the needed basics for an understanding of the domain ontology for utilities. Section three is going to define the four main use cases we identified in the requirements elicitation phase of the project. To implement those use cases, several existing approaches can be taken into account. The section four deals with the regional model authorities approach by Britton. Our main contribution can be seen in the sections five and six where we describe the newly developed CIM difference format and the versioning approach for the CIM. Section seven and eight deal with the implementation and evaluation of our implemented solution while section nine concludes the paper with future work and general remarks on the approach.

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. Nowadays, it seems as the Common Information Model (CIM) is going to be the most common data exchange format for energy markets in the future [12].

The CIM originates from a project by the American Electric Power Research Institute EPRI starting in the mid 1990s. It evolved to an International Standards being under maintenance by the International Electrotechnical Commission IEC. It is mainly standardized within the IEC TC 57 framework. Core elements of CIM have been adopted in International Standard IEC 61970-301, 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 [8][9][10]. The model has recently also been released in Web Ontology Language OWL [4]. More information on the CIM can be found in [17] and [16]. The semantic dimensions of the CIM and its implications on other standards are discussed in [14]. Other information on the use of CIM for EAI-based messaging can be found at [15]. This contribution is going to focus on the use of the CIM in order to facilitate the exchange of power grid data. This is being standardized as the international standard IEC 61970-552-4.

Deploying CIM leads to a lot of fundamental decisions, for instance if it is advantageous to sustain the currently used database or to implement a new database schema. 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. However, it is quite unlikely that it is possible to introduce a completely new data management in a company; instead, it is necessary to map CIM/XML data from and to the existing database. Those discussions will be out of scope for this paper, you can find more in [12]. This contribution is going to discuss the CIM and its special application from the perspective of RDF-based systems.

In the next section, we are going to present the main use cases we have identified when dealing with a topology exchange system for a utility.

3. Use Cases

We identified four 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 spend 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 accessable.

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.
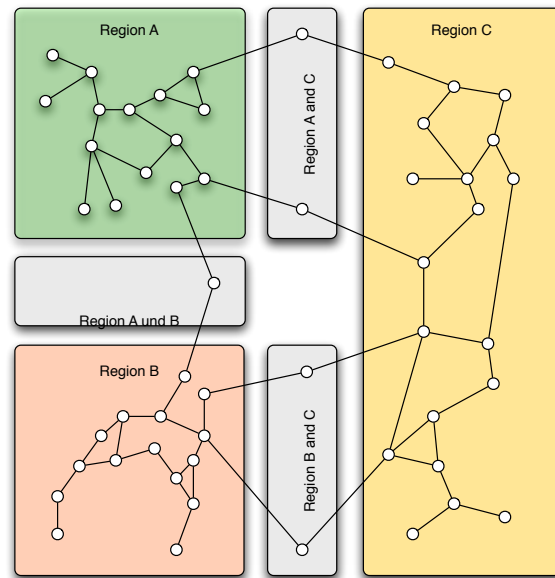
4. Regional Model Authorities

Figure 1. Regional Model Authorities according to [3]

The regional model for dealing with CIM topologies was originally developed by Britton as a theoretical model [3]. 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. Arnold deVos proposed a difference format for the CIM [5]. Unfortunately, this format has some drawbacks:

- It doesn't adhere to the RDF standard. DeVos extends the standards with a new 'parseType'.

- 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 [2].

As a replacement we developed two formats, one using insertions and deletions and one using replacements (cf. [2]). 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 a 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', using the replacement syntax proved to be superior to the format proposed by deVos and or first format.

Figure 3 illustrates the structure of our second format called 'dahlem2'.

Figure 4 show an excerpt of a actual difference between two version of the PsyCor SmallModel (as used in [1] and depicted in figure 2) in format 'dahlem2'.
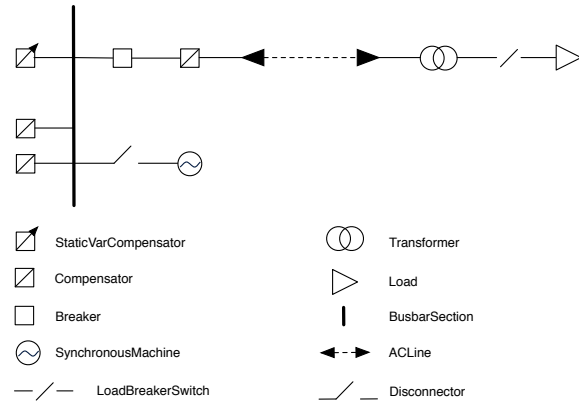

Figure 2. PsyCor SmallModel

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 possibilty to use our new, standards-compliant formats on the other.

Additionally the system supports queries using SPARQL [11]. 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.

The next sections are going to present how the CTS system was implemented and evaluated.

7. Implementation

We have implemented the solution using a distributed client/server architecture using web services. The main CTS system has been implemented in the following manner. 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

```
 1 <rdf:RDF
 2    xmlns:dm="http://cim.dahlem.eu.org/CIM_difference_model#"
 3    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
 4    xmlns:cim="http://iec.ch/TC57/2000/CIM-schema-cimu09b#">
 5  <dm:DifferenceModel rdf:about="http://somewhere.org/somenet">
 6  <rdf:Description>
 7     <dm:preconditions>
 8     <!-- Vorbedingungen -->
 9     </dm:preconditions>
10     <dm:postconditions>
11     <!-- Nachbedingungen -->
12     </dm:postconditions>
13  </rdf:Description>
14  <!-- Einfügeoperation: Aussage 1 wird eingefügt ->
15  <rdf:Statement>
16     <!-- Aussage 1 -->
17     <dm:replaces rdf:parseType="Resource">
18     </dm:replaces>
19  </rdf:Statement>
20  <!-- Löschoperation: Aussage 1 wird gelöscht -->
21  <rdf:Description>
22     <dm:replaces>
23        <rdf:Statement>
24        <!-- Aussage 1 -->
25        </rdf:Statement>
26     </dm:replaces>
27  </rdf:Description>
28  <!-- Ersetzung: Aussage 1 ersetzt Aussage 2 -->
29  <rdf:Statement>
30     <!-- Aussage 1 -->
31     <dm:replaces>
32        <rdf:Statement>
33        <-- Aussage 2 -->
34        </rdf:Statement>
35     </dm:replaces>
36  </rdf:Statement>
37 </rdf:RDF>
```

Figure 3. Structure of difference format dahlem2

distribution.

8. Evaluation

The EPRI sponsors the so called interoperability test [1]. By participating in those tests, the vendors of CIM based systems ensure that their solutions have proper interchangeability and interoperability. Currently, there have been eight major Interop tests. There is a special suite for testing the exchange of topology data and the needed difference models. Overall, our solution has successfully imported the three main topologies, the SMALL Model by PsyCor (depicted in figure 2), the Siemens PTI 100 and the Areva 60 Bus files. There are twelve tests dealing with the difference models which had to be translated into our syntax and then been applied to the CTS system. They passed all three implemented difference models. Furthermore, there are functionalities to validate the models against existing schemes like the Common Power System Model (CPSM [6]) using the HP Jena Eyeball inspectors. Overall, the system fulfils the following requirements elicitated from utilities and vendors:

- Import and export of topologies
- Creating topology parts
- Merging of topologies

```
 1 <rdf:Statement>
 2    <rdf:subject rdf:resource="http://cim.dahlem.eu.org/#
          _242C8FDFE7334E9BA603449C87D89903"/>
 3    <rdf:predicate rdf:resource="http://iec.ch/TC57/2003/CIM-schema-cim10#
          EnergyConsumer.pfixed"/>
 4    <rdf:object>30</rdf:object>
 5    <dm:replaces>
 6       <rdf:Statement>
 7          <rdf:subject rdf:resource="http://cim.dahlem.eu.org/#
                _242C8FDFE7334E9BA603449C87D89903"/>
 8          <rdf:predicate rdf:resource="http://iec.ch/TC57/2003/CIM-schema-
                cim10#EnergyConsumer.pfixed"/>
 9          <rdf:object>0.2</rdf:object>
10       </rdf:Statement>
11    </dm:replaces>
12 </rdf:Statement>
```

Figure 4. Example difference format dahlem2

- Saving of topology histories
- Tracking meta data like creator of a region
- Search functionality for topologies

9. Conclusions

As of today, there did not exist any 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. Our approach is more RDF-compliant and des also adhere to the CIM Users group interoperability test suite, we used the CIMXML Interop 6 difference files for a check the overall systems passed in all cases. 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 furthermore added a system for version control and regional modelling authorities to the difference formats. Further functionality completes the CTS system. It is possible to validate the CIM topologies against various schemes like CPSM and provides a suitable notification for system changes via Jabber or email. 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. Some performance improvements lie with using a different database which is better supported by Jena, e.g. postgreSQL. Furthermore, the current RDF triple stores are expected to gain more overall performance.

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. These could eventually be resolved by integrating SPARUL, the SPARQL Update Language [13]. More work can be done on providing sound formal foundation for SPARQL equal to the current relational algebra. On the other hand, we are going to improve the user interface for the provided server based on Web Services.

References

[1] Report on the Sixth Control Center Application Program Interface (CCAPI) Interoperability Test: The Power of the Common Information Model (CIM) and Generic Interface Definition (GID) to Exchange Power System Data, EPRI, Palo Alto, CA: 2004.[Report Number 1010117].

[2] 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.

[3] 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.

[4] J. Britton and A. deVos. CIM-based standards and CIM evolution. *Power Systems, IEEE Transactions on*, 20:758–764, 2005.

[5] A. deVos. RDF Difference Models - Representing the Difference between two RDF Models. http://www.langdale.com.au/CIMXML/DifferenceModels-R05.pdf, 04 2002.

[6] J. Evans and K. Hunter. CPSM Minimum Data Requirements in Terms of the EPRI CIM Version 1.9, 2004.

[7] IEC. 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.

[8] IEC. Draft IEC 61970: Energy Management System Application Program Interface (EMS-API) - Part 503: CIM XML Model Exchange Format - Draft 3b. Technical report, IEC - International Electrotechnical Commission, 2004.

[9] IEC. IEC 61970-501: Energy management system application program interface (EMS-API) - Part 501: CIM RDF Schema - Revision 4. Technical report, IEC - International Electrotechnical Commission, 2004.

[10] IEC. Draft IEC 61970-452-rev3: CIM Network Applications Model Exchange Specification. Technical report, IEC - International Electrotechnical Commission, 2006.

[11] E. Prud'hommeaux and A. Seaborne. SPARQL Query Language for RDF - W3C Candidate Recommendation 6 April 2006. http://www.w3.org/TR/2006/CR-rdf-sparql-query-20060406/, 2006.

[12] S. Schulte, R. Berbner, R. Steinmetz, and M. Uslar. Implementing and Evaluating the Common Information Model in a Relational and RDF-based Database. In *Information Technologies in Environmental Engineering - ITEE 2007 - Third International ICSC Symposium*, 2007.

[13] A. Seaborne and G. Manjunath. SPARQL/Update - A language for updating RDF graphs. http://jena.hpl.hp.com/ afs/SPARQL-Update.html, 2007.

[14] M. Uslar. Semantic Interoperability within the Power Systems Domain. In *Proceedings of the First International Workshop on Interoperability of Heterogeneous Information Systems in conjunction with the ACM Conference on Information and Knowledge Management CIKM 2005 Bremen*, 2005.

[15] M. Uslar. Das Common Information Model CIM und seine Nutzung für EAI in Energieversorgungsunternehmen. In *Proceedings der XML-Tage 2006 in Berlin*, 2006.

[16] M. Uslar. The Common Information model for utilities: An Introduction and Outlook on Future Applications. In *Proceedings der XML-Tage 2006 in Berlin*, 2006.

[17] M. Uslar, T. Schmedes, A. Lucks, T. Luhmann, L. Winkels, and H.-J. Appelrath. Interaction of EMS related systems by using the CIM standard. In *Proceedings of the ITEE 2005 Magdeburg*, 2005.