The Smart Grid as a Semantically Enabled Internet of Things

Andrew Crapo
GE Global Research
1 Research Circle
Niskayuna, NY 12309
crapo@ge.com

Ray Piasecki
GE Energy
500 Davis Street
San Francisco, CA 94111
Ray.Piasecki@ge.com

Xiaofeng Wang
GE Energy
1990 West Nasa Blvd
Melbourne, FL 32904
wang.xiaofeng@ge.com

Keywords: semantic modeling, interoperability, Internet of things

Abstract

Most visions of the smart grid see a multitude of devices, large and small, networked together in what might be described as, to use a term familiar to some, an Internet of Things. These things will include sensors, controllers, and the onboard-computers which will increasingly be associated with grid components. It is natural, given our history both with electrical transmission and distribution networks and with telecommunications/computing networks, to see these devices as connected to one another to form a network with information flowing between the nodes. However, it is also possible to envision the communication network of the smart grid as the infrastructure which supports a semantic model, or a federation of semantic models, that resides on the network and captures all relevant aspects of the smart grid as it was, as it is, and as it is expected to be. Each device, enabled with the knowledge of its own capabilities and purpose expressed with the semantics of shared ontologies, will be able to "plug-and-play" at the semantic level. In other words, once connected at the communication level, the device will know how to draw the information that it needs from the semantic model in which it exists and/or to provide the information that it senses or computes back to the semantic model for use by others – human or device. This will give us true interoperability; not interoperability based on the syntax of a message or the definition of an API, but interoperability because devices inherently can "talk" to one another about the things they care about.

1. INTRODUCTION

Significant advances often occur at the intersection of fields of endeavor where synergies emerge between the thinking of researchers with different views who are approaching a common or related problem from different directions. The problems around smart grid security and interoperability have important commonalities with two other technology thrusts emerging in different fields. From the field of networked communication comes the Internet of Things (IOT). The number of devices "on the Internet" has surpassed the number of people with Internet access and 24 billion devices are expected to be connected by 2020 [1, 2].

From the creators of the World Wide Web comes Linked Data, also known as the Semantic Web [3, 4]. The Linked Data initiative sees the Web, including its numerous devices, as suppliers and users of self-describing data—self-describing because the metadata is drawn from shared ontologies. Such data is inherently interoperable at the semantic level once interoperability is achieved at the communication or syntactic level.

The objectives and challenges of these two emerging technologies have much in common with some of the challenges of the smart grid. As a smart grid community we will be wise to examine carefully the opportunities to leverage research in these related fields. In this paper we will set the stage by summarizing some key smart grid IT challenges, briefly describe IOT and Semantic Web technologies, give a summary and update of our work in semantic modeling for smart grid, share our evolving vision of the future, and ground that vision with a use case around interoperability of two smart grid standards.

2. THE SMART GRID INFORMATION CHALLENGE

The information side of smart grid poses a number of well-known challenges.

Reliability

The Internet does not currently come with any guarantee of quality of service. Will the Internet provide a sufficient level of reliability for those smart grid functions that will require wide area networks? Will grid operators and service providers be forced to create their own network infrastructure for critical applications? And even if they do, will these really be more reliable than the public network?

Security

Information security seems a game in which the good guys at best stay a half-step ahead of the bad guys. From Stuxnet to WikiLeaks, there are frequent reminders of the difficulties of keeping information systems and data secure. Smart grid will likely be seen, at least in some of its operations, as a high-stakes target for hackers. Getting my energy usage data may not attract much talent, but blacking out New York City or Los Angeles might. Security is a very complicated business. Even if devices and networks have sufficient security capability, can the work force be trained to properly configure and maintain them to achieve secure operations?

Scalability and performance

While some functions, such as reading a meter, may not require a rapid response, others, such as load balancing and VAR control, will require near-real time communication. From streaming video to ubiquitous sensors, the data demands on the Internet continue to grow rapidly. Smart grid will add its own demands for bandwidth with a bi-directional flow of information between grid components. Will the bandwidth be available when it's needed?

Privacy

I sat down recently with my grandson and looked at an inside temperature profile of our home. A lot could be inferred from just this simple time series—when we were home, when we were gone on vacation to visit his family. A home full of smart appliances and a home energy management system will provide sufficient data to not only minimize energy consumption but to identify our private behaviors. For example, the data might notify burglars of when we are away and what to do to disarm our homes. How much data leaves the home? Who has access to it? How will it be protected?

Complexity

The electrical grid is a complex analog system and has not previously been under centralized control other than its macro structure. In the past reliability has largely been achieved by making the grid respond automatically to changing conditions. The Internet and networked communication systems in general are complex digital systems. Complex systems with non-linear interactions are subject to cascading failures and the integration of two systems as complex as the electrical and communication grids could expose unanticipated interactions leading to unexpected failures [5].

These challenges are not unique to smart grid. They are inherent in the rapidly evolving interconnected world of people and devices in which we live. Two thrusts in information technology seem particularly relevant to smart grid developers. The first is the Internet of Things (IOT), which envisions people far outnumbered by connected devices, of varying degrees of "smartness", in the Internet of the future. The other is the semantic enablement of the Web: the Semantic Web and its Linked Data. This paper will briefly discuss these research directions and describe ways in which they are relevant to smart grid development.

3. THE INTERNET OF THINGS

The "Internet of Things" is a term that describes a world in which physical and virtual objects are networked together, enabling them to "talk" to each other to exchange data and services. Is this something that exists today or simply a vision for the future? The answer is both. There are currently an estimated 9 billion devices connected to the Internet, of which 6 billion are mobile. While that number is expected to more than double by 2020, perhaps the more relevant projection is the anticipated revenue for mobile device operators—nearly $2.1 trillion [6]. There is an Internet of Things today, just as there is an electricity grid today, but both the grid and the IOT are expected to be very much different in the future—and in similar ways. Some observers see the smart grid as just one example of the way the IOT will change the world [7]. In other words, the evolution of the IOT may well define and drive the way that the smart grid's networks of smart devices are connected and interact, not the other way around. That will put smart grid developers in the role of wisely choosing from rapidly evolving technologies rather than needing to invent them.

The IOT has its share of problems to overcome and naysayers to convince [8]. Can the infrastructure support such a huge expansion of the Internet? Is the Internet reliable enough to trust our safety and security to its erratic performance? What about privacy—if your wine glass is tracking how much you drink and telling you when to stop, who else now has that data? Why would anyone go to the trouble and expense of making umbrella handles smart enough to tell you when it is going to rain when you can already get that information from a variety of sources? Not surprisingly, these questions somewhat parallel those often asked about the smart grid. In fact, there is one more set of analogous questions asked. Does the evolution of the IOT mean that China, with a top-down view of control, leap-frogs the west? Can't the objectives of the smart grid be more efficiently and reliably achieved by coupling distributed storage and control in every home and business, creating a distributed intelligence much more like that found in biological ecosystems? And much more like the capitalistic structures of Western economies?

4. LINKED DATA AND THE SEMANTIC WEB

The Semantic Web envisions the World Wide Web of the future as one of self-describing data where the content of the Web is "understandable" to both people and computers. Data that is self-describing in a way that is understandable by other people and computers (not just the data's author) is achieved by tagging the data using metadata that is shared across communities of interest. This structured metadata is called a semantic model or an ontology. The structured metadata forms a formal, logical model and this allows the model to be checked for consistency, completeness, etc. It also allows instance data described using the model to be validated and reasoned over to infer logical conclusions not explicitly in the original data. Data on the Semantic Web can be linked or connected in the sense that the data makes sense when it is put together with other data from the same or related domains [9]. Linked Data is a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using standards such as Uniform Resource Identifiers (URIs), Resource Description Framework (RDF), RDF Schema (RDFS), and the Web Ontology Language (OWL) [3].

The tools of the Semantic Web are not the most sophisticated but they provide practical solutions to important problems. First of all, the identity problem—the problem of identifying things in information systems and knowing when two identifiers refer to the same thing and when they refer to different things—is flexibly solved through the use of URIs to identify all resources, which includes devices as well as the classes to which they belong and the [types of] relationships between them. The XML namespace standard underlying URIs makes the use of the same local name (URI fragment) for different concepts a non-issue. OWL nicely supports the interfacing of disparate models in which the same concept is given different names in the two models through the OWL equivalence tags for classes, properties, and instances. This is identification at Web scale.

An important OWL functionality essential to merging models from different domains is its support for multiple-inheritance. There is no difficulty identifying a device as an instance of a communicating device on the one hand and an instance of an electrical device on the other hand. It is entirely conceivable that the two meta-models, one of communicating devices and one of electrical devices, are quite independent at the meta-level but that a single actual device can be an instance of classes in both models and be fully described and reasoned about using both.

The implications of using semantic models and linked data are significant. Semantic models and linked data enable distributed, autonomous, and interoperable devices and subsystems to connect semantically, share information, and be smart. They are widely seen as the key to interoperability of people and smart devices in the networks of tomorrow.

5. SEMANTICS FOR SMART GRID: EXPERIENCE AND VISION

5.1. What We've Learned

At Grid Interop 2009 we shared our preliminary vision of how semantic technology can address the challenges smart grid poses for interoperability [10]. In 2010 we reported on our evaluation of the CIM as a semantic model and some of the challenges we encountered while using it as the basis of modeling [11]. During the past year we have focused our efforts on making our modeling environment, the Open Source, Eclipse-based Semantic Application Design Language (SADL), which provides a controlled-English interface to OWL and domain rules, more complete and user-friendly and able to accommodate a variety of reasoners via a reasoner/translator plug-in architecture [12] .

On the application side we have gained experience modeling cyber security problems as reported in another presentation at this conference [13]. Semantic models that capture domain concepts in a formal, modular, extensible way provide a foundation for classification and logical reasoning. Rules that capture expertise in the very complex cyber security domain are built on this foundation and expressed in terms of the foundational concepts. Overall our experience reinforces our belief that semantic modeling is a valuable tool for addressing security needs.

Modularity continues to be a substantial concern. The architecting of models so that their core is optimally reusable by extension for different purposes is essential to achieving interoperability across even adjacent spaces. While we do not have a final solution to this issue, we have found the modularity, extensibility, and immediate computability of models represented in the Web Ontology Language (OWL), combined with the version control and validation/regression testing capability of the SADL environment, to be beneficial in building and refactoring models as they evolve.

5.2. Our Vision Going Forward

The smart grid will include many devices that are simultaneously part of a communication network and part of the monitoring/control mechanism of the electrical grid. These will include sensors, remotely controlled switches, and other smarter devices. Smart grid interoperability will be enhanced considerably if the modeling paradigm is able to merge the two models (the communication network model and the electrical network model) in ways that allow integrated model usage. As described above, OWL provides several capabilities that facilitate this kind of merging, including URIs and support for multiple-inheritance.

5.2.1. A Hypothetical Example of Data Rollup

The semantics of the smart grid domain – generation, transmission, distribution – can be captured in a set of modularized ontologies or semantic models. Some of these models already exist (e.g., the CIM) and some are yet to be developed. The simplest of devices, a temperature sensor for example, may be given an understanding of its capability and purpose using concepts (metadata tags) drawn from standardized semantic models. The device's understanding may include concepts for geospatial location and time. The content of messages to and from the device is in the form of self-describing data—models composed of instance data and the appropriate metadata tags from the shared semantic models.

Suppose the sensor is on a turbine generator. The control system of the turbine generator may receive a report of temperature from our lowly sensor, which is immediately understood in the context of a broader semantic model than the one possessed by the sensor. This pattern repeats itself both at higher levels and at adjacent levels as needed. Much like a multidisciplinary team, the members of any segment of the smart grid have knowledge of themselves and, when needed, of their neighbors, superiors, and subordinates. As soon as the team members are able to communicate they become interoperable not just at the communication level but at the semantic level.

The communication network of the smart grid constitutes the infrastructure on which semantic models reside. Devices, once they can plug-and-play at the communication level, are able to plug-and-play at the semantic level. Each network node, with a sufficient semantic understanding of its capability and purpose, knows how to obtain information that it needs from the semantic ecosystem in which it exists. Information that it senses or computes is added to the semantic ecosystem for use by others – artifact or human. Furthermore, because a semantic model is a formal representation with model theory and proof, new and pre-existing information is constantly being considered by reasoners , which infer logical entailments that are added to the model. This reasoning can occur at many levels with the scope at each level determined by the breadth of semantic understanding available to the reasoner.

5.2.2. The Configuration Management Problem

An area of significant concern to the smart grid community is the ease with which large numbers of devices can be configured and managed. Concepts drawn from IOT and Linked Data again come to our aid. Both devices and the data which they generate can be self-describing. For example, a new meter attached to the communication grid can already know its identity. It can query the grid to find out the new context in which it has been placed. As it begins to generate data, the data can be self-describing both in terms of what the data is (voltage, power, energy consumed, etc.) and in terms of where it comes from (what it's related to and how). Similarly, information sent to the meter, e.g., real-time pricing, can be self-describing and easily used by the meter to communicate with an energy management system. Installing, configuring, and managing a myriad of devices and integrating the data that they will produce to enable the higher-level computation necessary to higher-level benefits will require Web-scale approaches.

5.2.3. A More Concrete Interoperability Example

As another and more concrete example of how semantic technology can address interoperability, consider the problem of harmonizing the IEC CIM and the IEC 61850 models. The CIM is "a common semantic model to unify and integrate the data from a myriad of systems involved in support of real-time electric utility operations" while 61850 is a "series of standards for substation automation" [14]. These two IEC standards have many structural differences even though they sometimes model the same actual equipment. Harmonization of the two models could proceed in one of two basic directions.

One approach is to redefine the semantic models explicit in each standard so that there is a single, shared semantic core. This is the approach described by Becker [14]. This is a daunting task for several reasons. First of all, the two standards do not even use the same meta-language to capture their models. The CIM uses UML as its meta-model while IEC 61850 uses its own custom meta-model consisting of Logical Node, Common Data Class, etc. Secondly, standards bodies would have to accept the new, harmonized model and be willing to cope with issues such as backward compatibility. Finally, one might reasonably suppose that at least some differences are due to differences in the point-of-view and needs of the two communities and that a common, harmonized model can only go so far as there will be special and even possibly incompatible needs that will exist.

An alternate approach is to leave both standards as they are and define the mappings between concepts in one model and concepts in the other model. A portion of this mapping can be achieved axiomatically using, for example, the OWL tags for instances (owl:sameAs, owl:differnentFrom, owl:allDifferent), for classes (owl:equivalentClass, owl:disjointWith), and for properties (owl:equivalentProperty). This works when the concepts of the two models are well-aligned but use different names for the same concept. The remainder of the mapping can be achieved, for example, by composing rules on top of the shared semantics that define mappings between structures in one model and the inferred structures in the other model.

The two approaches are not mutually exclusive and many possible combinations might be used. For example, the community might collaboratively build a smart grid "upper ontology" that expresses the fundamental concepts common to CIM and 61850 and possibly Smart Energy Profile (SEP). Leveraging the modularity and extensibility of OWL, this upper level ontology could be extended to provide the unique modeling concepts required by each of the standards, with rule-based translation as needed. Or perhaps the extensions of the common upper-level ontology could be aligned with the seven smart grid domains as defined in the NIST smart grid interoperability framework.

Topology, measurements, and configuration are areas of overlap in the 61850 and CIM models. These areas contain important information elements that must be shared between 61850 based field controllers and enterprise level DMS, GIS and SCADA/EMS applications found in utility control centers. Mapping 61850 and CIM concepts in these categories will be a key enabler to the integration of field devices and utility back office systems.

For example, a Multi-Function Bay Controller/RTU IED residing in a substation would potentially be using 61850-specified Logical Node (LN), Common Data Class (CDC) and Common Data Attribute (CDA) constructs to express device functions and measurements to external applications. It is desired to interface the 61850 based IED controller (located in the field) with CIM based DMS and GIS energy management applications (located in the utility control center). Figure 1 is a class diagram snippet which describes how the IEC and CIM models express a circuit breaker object. CIM includes the notion of a circuit breaker class directly, but 61850 uses the generic notion of a ConductingEquipment class which would be cast as type "circuit breaker" using the CBR enumerated type. From the example it can be seen that the two standards use very different class constructs as well as have differences in data attribution. Rather than change the models, or the applications, a semantic mediation service can be used to perform the translation externally. The semantic mediation service would use class equivalence axioms to map and translate instances of the 61850 logical node data classes (output by the IED) to instances of CIM classes (used by DMS and GIS). In addition to the fundamental class mapping the semantic mediation service would also apply constraints and business rules during the translation processing. The rules and constraints would enable tuning the translation to application-specific requirements and would add/remove attributes, translate values, and validate the types and ranges.

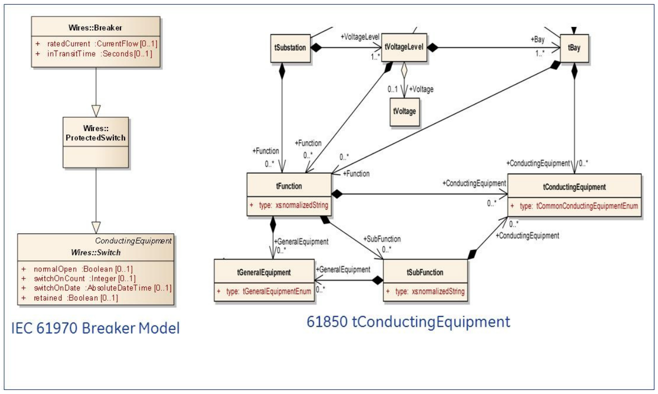

Figure 1: Circuit Breaker in CIM and IEC 61850

5.3. Semantic Modeling Challenges

Semantic reasoning is memory intensive. It does not lend itself well to parallelization in the map reduce fashion unless the instance data and reasoning problem of a particular scenario is itself quite small but there are many parallel scenarios to be evaluated. For larger, tightly coupled scenarios the problem is not easily divisible. A solution to this problem may be emerging in technologies such as massive memory machines and RAM clouds. Cray's XMT supercomputing system is "a scalable massively multithreaded platform with a shared memory architecture" designed for graph analytics, of which semantic reasoning is a subset [15].

In some sense this view of ubiquitous semantic models is the opposite of that of object-oriented programming with its encapsulation, information hiding, and polymorphism. A semantic model makes everything available. On the other hand, semantic models are arguably object-oriented and at least some semantic languages are very strong on inheritance, abstraction, and modularity. Ubiquitous access to information by semantic query achieves many of the same objectives as object-oriented programming but in ways that are more conducive to interoperability of relatively autonomous devices in an Internet of things. A node on this IOT can ask for any information for which it can formulate a semantic query—it can ask for anything it can understand.

Consequently, information access control is very important to protect the quality of the data and prevent it from being accessed inappropriately. Graph query languages like SPARQL allow semantic patterns to be defined and matched in the semantic models much like SQL allows relational patterns to be defined and matched in relational tables. Thus semantic model access control is more like access control in a relational database than it is like access control in processes executing procedural code. Information about who should be allowed to create, read, update, and delete information is not addressed by semantics implicitly but can certainly be modeled using semantics if that proves to be the best approach.

Data provenance is also an important topic for the semantically enabled IOT. In today's information systems the provenance of data is often implied by its location. Provenance of tomorrow's data will need to be much more explicit and a part of the self-describing data content. Authentication of the provenance of semantic information is not unlike today's authentication challenges and similar solutions can be applied.

6. CONCLUSION

Linked data that draws its metadata from a semantic model can enable distributed, autonomous and interoperable devices and subsystems to connect semantically, share information, and be smart. The technologies to enable an Internet of Things that is highly configurable and interoperable are the goal of both the IOT and the Semantic Web communities. Smart grid can be the poster child for both of these research groups if we seek the synergies between the practical problems we face and the solutions they offer.

References

[1] Malik, O. (2011, October 13). Internet of things will have 24 billion devices by 2020. Retrieved October 25, 2011, from Structure: http://gigaom.com/cloud/internet-of-things-will-have-24-billion-devices-by-2020/

[2] *Internet Users in the Word Distribution by World Regions-2011*. (2011, March 31). Retrieved October 25, 2011, from Internet World Stats: Usage and Population Statistics: http://www.internetworldstats.com/stats.htm

[3] Linked Data. (n.d.). Retrieved October 27, 2011, from Linked Data - Connect Distributed Data across the Web: http://linkeddata.org/

[4] Semantic Web. (n.d.). Retrieved November 15, 2011, from W3C: Linked Data. (n.d.). Retrieved October 27, 2011, from Linked Data - Connect Distributed Data across the Web: http://linkeddata.org/

[5] Danforth, C. M., Hines, P. D., O'Hara, B., Cotilla-Sanchez, E., & Danforth, C. M. (2011, April). *Cascading Failures: Extreme Properties of Large Blackouts in the Electric Grid.* Retrieved October 26, 2011, from Mathematics Awareness Month: http://mathaware.org/mam/2011/essays/complexsystemsHines.pdf

[6] Lipman, Rebecca. (2011, October 18). *The Internet of Things: The Opportunity of a Lifetime?* Retrieved October 25, 2011, from The Motley Fool: http://www.fool.com/investing/general/2011/10/18/the-internet-of-things-the-opportunity-of-a-lifeti.aspx

[7] Hertzog, C. (2011, October 18). *The Smart Grid Offers a Glimpse into the Internet of Things*. Retrieved October 24, 2011, from the energy collective: http://theenergycollective.com/christine-hertzog/67281/smart-grid-offers-glimpse-internet-things

[8] (2010, December 9). *The "internet of things", The internet of hype*. Retrieved October 25, 2011, from The Economist: http://www.economist.com/blogs/schumpeter/2010/12/internet_thing

[9] Allemang, D., & Hendler, J. (2011). Semantic Web for the Working Ontologist: Effective Modeling in RDFS and OWL, 2nd Edition. Waltham, MA, USA: Morgan Kaufmann.

[10] Crapo, A., Wang, X., Lizzi, J., & Larson, R. (2009, November). *The Semantically Enabled Smart Grid.* Retrieved October 26, 2011, from Gridwise Architecture Council: http://www.gridwiseac.org/pdfs/forum_papers09/crapo.pdf

[11] Crapo, A., Griffith, K., Khandelwal, A., Lizzi, J., Moitra, A., & Wang, X. (2010, December). *Overcoming Challenges Using the CIM as a Semantic Model for Energy.* Retrieved October 26, 2011, from Grid Archecture Council: http://www.pointview.com/data/files/3/2433/1730.pdf

[12] Crapo, A. (2011). *Semantic Application Design Langague (SADL).* Retrieved October 27, 2011, from SourceForge: http://sadl.sourceforge.net

[13] Barnett, Bruce, & Crapo, Andrew. (2011, December). *A Semantic Model for Cyber Security*, forthcoming at Grid Interop 2011.

[14] Becker, D. (. (2010). *Harmonizing the IEC CIM and 61850 Standards via a Unified Model: Key to Achieving Smart Grid Interoperability Objectives (draft).* Palo Alto, CA: Electric Power Research Institute.

[15] *Cray XMT.* (2011). Retrieved 11 15, 2011, from CRAY: The Supercomputer Company: http://www.cray.com/products/XMT.aspx

Biography

Andrew Crapo received a B.S. in Physics from Brigham Young University, an M.S. in Energy Systems from the University of Central Florida, and a Ph.D. in Decision Sciences and Engineering Systems from Rensselaer Polytechnic Institute. He is a senior information scientist at the GE Global Research Center where he has worked since 1980. His work has focused on applications of information science to engineering problems including applied artificial intelligence, human-computer interactions, and information system architectures. More recently he has focused on modeling and the application of Semantic Web technologies to engineering and business problems.

Ray Piasecki received B.S. and M.S. degrees in electrical engineering from San Diego State University in San Diego CA. He is a principal technologist at GE Energy's Smart Grid Center of Excellence in San Francisco CA. He serves as a technical specialist consulting on smart grid systems design, microgrid and distributed energy resource control systems, utility IT modernization, and information modeling and analytics.

Xiaofeng Wang received B.S. and M.S. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree from the Electrical and Computer Engineering Department of Michigan Technological University, Houghton, in 2001. Currently, he is a System Engineer with GE Energy. His interests include power system modeling, enterprise integration, Smart Grid, and Semantic Web Technology.