Advanced Data Management and Analytics for Automated Demand Response (ADR) based on NoSQL

Bert Taube

Versant Corporation
255 Shoreline Drive, Suite 450

Redwood City, CA 94065

btaube@versant.com

Rolf Bienert

The OpenADR Alliance
275 Tennant Avenue, Suite 202

Morgan Hill, CA 95037

Rolf@openadr.org

Keywords: Automated Demand Response (ADR), NoSQL, Object-Oriented Data Models, Model-Based Data Management, Advanced Data Analytics, Connectivity, Interoperability, Pattern Recognition, Prediction

Abstract

The largest energy producers in the US currently operate Demand Response programs totaling more than 31GW of subscribed resources. Many of these programs are operated manually or with very little automation. The recent trend to move to an interoperable Smart Grid has initiated significant efforts to automate a large percentage of the available resources.

While manual Demand Response already requires good planning, forecasting and data management, automated Demand Response will multiply the need for specific strategy mechanisms based on adequate data management and analytics. Furthermore, automated demand response will provide additional telemetry data which will need to be added to the existing information.

The proposed paper will discuss advanced data management and analytics as it applies to the needs of fully integrated ADR solutions. It will show the different data sources and characteristics within the ADR framework and derive advanced object-oriented data management concepts applicable for data storage on OpenADR servers. It will become evident how object-oriented data models within the NoSQL data management concept support far better performance in the raw data enhancement process as well as the subsequent data analysis. The paper will show how object-oriented data management allows for fast, effective and reliable data access within the ADR framework and thus supports strategic data usage, program execution as well as the necessary accounting (e.g. billing) in an ADR solution. Last but not least, the paper will explore potential applications of the NoSQL based data management and analysis concept within the ADR framework such as automated fault detection in components and the identification of root causes.

1. INTRODUCTION AND BACKGROUND

OpenADR standard development has evolved through research, pilots, and commercialization. The OpenADR 1.0 communication specification by Lawrence Berkeley National Laboratory (LBNL) DR Research Center and the California Energy Commission (CEC) (Piette et al., 2009a) is implemented in California's commercial Automated DR (Auto-DR) programs (Wikler et al., 2008), and is now an accepted standard in the industry. The new standard, called OpenADR 2.0, is a result of contributions from various standards organizations and the OpenADR stakeholders. The OpenADR Alliance (Alliance) is the managing entity for OpenADR 2.0 and is the provider of certification and testing programs for interoperability.

NoSQL – based data management has been has been around for two decades. It is based on an object-oriented database approach which is used in network industries such as energy, telecommunications and transportation to track large number of objects. Unlike relational (SQL) or serialized databases, object-oriented databases offer seamless integration with object-oriented languages. Unlike SQL—which encompasses its own database language apart from the programming language—the object database uses the OO programming language as its data-definition language (DDL) and data-manipulation language (DML). The application objects are the database objects. Query is used for optimization based on use cases, not as the sole means of accessing and manipulating the underlying data. There is no application code needed to manage the connectivity between objects or how they are mapped to the underlying database storage. Object databases use and store object identity directly, bypassing the need for the CPU and memoryexpensive set based JOIN operations using SQL. Object databases exhibit traditional database features, such as queries, transaction handling, backup, and recovery, along with advanced features such as distribution and fault tolerance.

1.1. Background

OpenADR provides non-proprietary, standardized interfaces to enable electricity service providers to communicate DR and Distributed Energy Resource (DER) signals to customers using a common language and existing communications such as the Internet (Piette et al., 2009b). These OpenADR data models facilitate price-responsive and reliability DR. As shown in figure 1 below, this is achieved through open Application Programming Interfaces (APIs) that provide two-way communications between the service provider (Utility/ISO) and customers (Sites) through a logical interface of OpenADR server (called a Demand Response Automation Server).

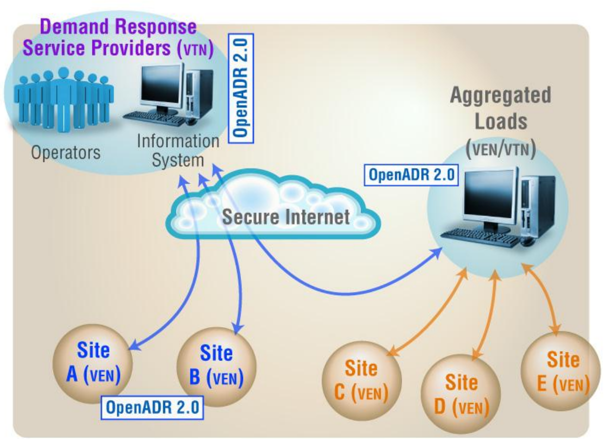

Figure 1: OpenADR Communication Architecture

The communications between the service providers and consumers in OpenADR 2.0 have evolved generically as the Virtual Top Node (VTN) and the Virtual End Node (VEN), respectively. The VTN/VEN pair structure allows a chain of hierarchy from the parent (the one that issues primary DR signal) to the multiple parent/child relationships all the way to the end-use devices (OASIS, 2011).

1.2. NoSQL Big Energy Data Management & Analytics

1.2.1 Big Energy Data Features

The problem of "big data" was originally defined as a three-dimensional space of orthogonal variables of volume, velocity, and variety by Laney in 2001 [Doug Laney, 2001]. Laney was operating in a paradigm of a typical business, such as manufacturing, where profitability is often achieved by the minimization of fixed assets, where work in progress (WIP) is measured in days, weeks, or months, and where real-time data collection and analysis are often not critical to ensure the profitability of the organization. The value chain for manufacturing almost always crosses company boundaries. However, in the utility industry; there are vertically integrated and deregulated variants that nevertheless have to act exactly the same.

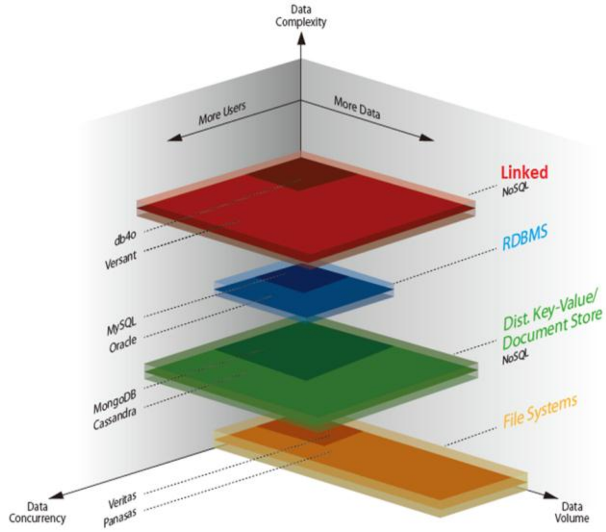

Figure 2: Classification of Data Management Technologies

The utility industry is unique in that the product is consumed virtually simultaneously to its production (but the price may be set years in advance) and the focus is on the utilization of assets (which are often defined by circumstances) rather than the minimization of assets. In this environment, the acquisition of real-time data can be costly and can seriously impact the bottom line. The utility industry must still deal with volume, velocity, and variety, but two new "V's" are introduced: validity and veracity.

Validity is adding a fourth "dimension" to Laney's model, where time is considered. Information in the utility environment often has a "shelf life" and is useful, and therefore needed to be stored, only for a fixed amount of time. After that time, the data may no longer be needed for evaluation. The questions of when to archive or even dispose of data become relevant given the cost of storing large quantities of data.

Veracity is the recognition that the data is not perfect and that achieving "perfect" data has a cost associated with it. The questions become 1) how good must the data be to achieve the necessary level of analysis and 2) at what point does the cost of correcting the data exceed the benefit of obtaining it?

As figure 2 illustrates, it is the linked NoSQL solution that allows solving modern world big data challenges characterized by extreme complexity, concurrency and volume as well as validity and veracity. As a result, there is the need to shift to NoSQL architectures in order to provide the necessary scalability inherent in a big data problem. This will ensure enterprise-linked data management and analytics as required for the real-time utility enterprise.

The following table displays the diversity of challenges inherent in big energy data challenges:

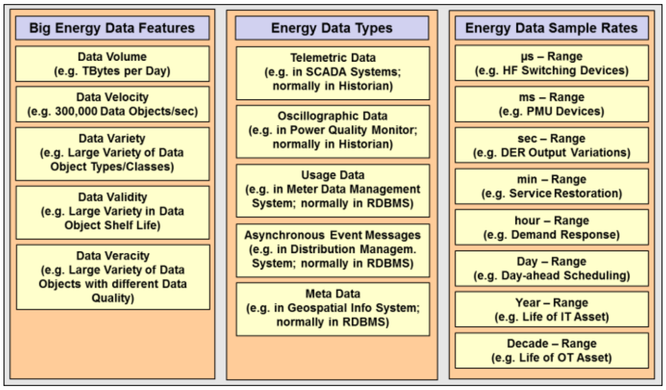

Figure 3: Diversity of Big Energy Data Challenge

As the cost of storage continues to plummet and the bandwidth and I/O speeds of networks and servers continue to increase, it makes less sense to utilize conventional types of data-management technologies (e.g. relational and time-serialized).

1.2.2 Big Energy Data Types

According to Jeffrey Taft, Paul de Martini, and Leonardo von Prellwitz ["Utility Data Management & Intelligence", Cisco White Paper, May 2012], utility data for advanced data management and analytics can be summarized as follows:

1. Telemetric data (e.g. in SCADA systems; conventionally stored in data historians),
2. Oscillographic data (e.g. in power quality monitors; normally stored in data historians),
3. Usage data (e.g. in meter data management systems; frequently stored in relational databases),
4. Asynchronous event message data (e.g. in distribution management systems; often stored in relational databases),
5. Meta data (e.g. in geospatial information systems (GIS); mostly stored in relational databases).

Telemetry and oscillography are often stored in time-serialized database, while usage data, asynchronous event messages, and meta-data are often stored in relational databases. Often meta-data, such as connectivity, is stored as binary large objects (BLOBS) in products such as a geospatial information system (GIS), usage data is stored in a meter data-management system (MDMS), and asynchronous messages are stored in a variety of places, one of which is a distribution management system (DMS).

The dilemma is that neither of the prevailing data-management technologies is an ideal way to store, manage, and analyze these data types. This is especially true when one is attempting to analyze across data types.

So the utility industry has come to a point where the data-management technology of the past no longer fits the needs of the industry just at a time when the amounts of data produced are about to increase significantly. What is needed is a data-management technology that is optimized for analysis rather than constraints such as space and speed. Ideally, this database technology would be built much like the grid itself, with classes of assets that have a natural relationship between the classes. This is exactly the capability of a NoSQL-based data management and analytics solution.

1.2.3 Big Energy Data Sampling Rates

But not enough with the challenges implied in the variety of energy data types. Figure 3 also displays the diversity of sample rates used to collect energy data. It illustrates the unique situation of the utility industry, where data time scales vary over 15 orders of magnitude. Traditional methods of data management (relational databases or time-serialized databases) may not have the capability to capture the causal effects that may be on the order of years or decades of events that may occur in the millisecond or microsecond range.

Analyzing huge volumes of data that spans multiple orders of magnitude in time scale is a serious challenge for current data-management technologies prevalent in the utility industry. And again, NoSQL-based data management and analytics can accommodate the variety in time scale.

1.2.4 Compliance with Big Energy Data Standards

In electric power supply (generation, transmission, distribution, consumption) the common information model (CIM) is a standard developed by the electric power industry to fully describe the assets, topology, and processes that make up the grid. The CIM is a set of standards, adopted by the International Electrotechnical Commission (IEC), whose original purpose was to allow application software to exchange information about the configuration and status of the grid.

The CIM is described as a UML model. The central package of the CIM is called the "wires model," which describes the basic components used to transport electricity. The standard that defines the core packages of the CIM is IEC 61970-301, with a focus on the needs of electricity transmission, where related applications include energy-management system, SCADA, planning, and optimization. The IEC 61968 series of standards extend the CIM to meet the needs of electrical distribution, where related applications include distribution-management system, outage-management system, planning, metering, work management, geographic information system, asset management, customer-information systems, and enterprise resource planning.

The CIM UML model, which describes information used by the utility, is an ideal candidate on which to base an object-oriented database. It has several advantages:

- The schema, derived from the standard, would be public and well documented.
- The UML relationships would have been vetted through the standard development process.
- Messages base on the UML model are, in themselves, standards.

By developing a schema in an object-oriented database, based on the CIM that describes the relationship between actual object classes in the utility, it is believed that additional insight into the inner workings of the grid will allow for better, faster, and more insightful and more widely useful analytics.

2. ROLES OF AUTOMATED DEMAND RESPONSE ENTITIES

As discussed above, OpenADR defines two functional entities, the Virtual Top Node (VTN) or server and the Virtual End Node (VEN) client. These functional entities represent the interfaces between the operational entities involved in the execution of the Demand Response programs. The following diagram illustrates some of the possible architectures of an OpenADR Demand Response program.

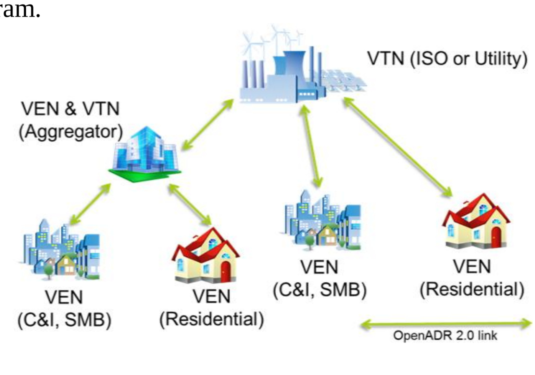

Figure 4: OpenADR Architecture

As shown above, the operational entities can act as either functional entity when it comes to the OpenADR interface. However the roles and responsibilities vary greatly.

2.1. Utility and Independent System Operators (ISO)

The utilities and ISO/RTOs (energy providers) primarily are the responsible entities in Demand Response programs. As such, the burden of calculating the right time, amount and price for a DR event rest with them at all times. Different goals of the DR program can vary the way in which information is being processed as well as the necessary processing speed. The most common goals of DR programs are –

- Peak Load Management: Often events are established hours or even days ahead to soften high demand periods.
- Grid and frequency balancing: In order to cope with fluctuations in generation (e.g. renewable energy), energy providers resort to short term demand side changes to stabilize voltage and frequency which ramping up generation. These modifications of the demand side require very quick data processing as facility response times of 4 seconds are desired.

In order to make appropriate decisions, the energy provider needs to obtain, store, access and process a large variety of information. The most common data sources are shown below.

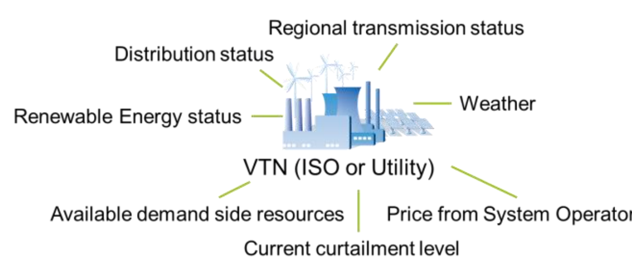

Figure 5: Select Data Sources Energy Provider

2.2. Aggregators

Demand Response Aggregators receive requests to provide demand side load reduction (curtailment) from the energy providers. Typically this also includes a time frame, energy amount and pricing information. The aggregator represents a VEN (client) to the requester. The aggregator then applies its own business logic to calculate which resources to deploy to satisfy the request of the energy provider. This can be geographically based but also more generally focused on the energy amount.

While the information provided to the aggregator is not very data intensive, the aggregator still has to process a lot of information it receives from the connected resources. In particular at the time of an event, the aggregator is under pressure to produce the contracted curtailment and has to process the feedback from the resources (facilities).

2.3. Facilities

As show in the NIST diagram below, there is a great variety of interfaces with the customer domain. Each and any of these systems can contribute to the vast amount of data that has to be processed for Demand Response. In particular Fast Demand Response will require real time access to a variety of data from this domain.

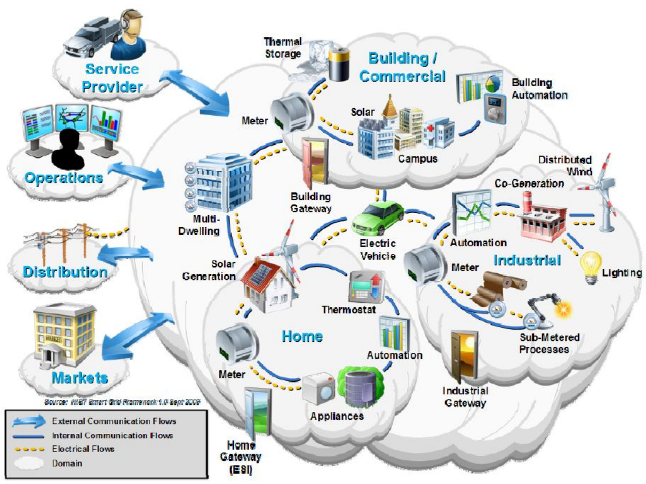

Figure 6: NIST Smart Grid Framework – Customer Domain

3. INTERFACES BETWEEN ENTITIES

OpenADR 2.0 uses a server to client communication architecture between one client and one server. It does not network between clients or servers and does not create large networks of OpenADR signals. Instead, after an OpenADR 2.0 message has been created and distributed to a client (or clients), the client systems can decide which strategies can be applied. Subsequently, if these systems also act as servers to other participants, they can create new OpenADR 2.0 messages that serve the selected strategy.

During the registration process, servers and clients exchange additional data out-of-band in order to establish the required contractual and program related rules and regulations.

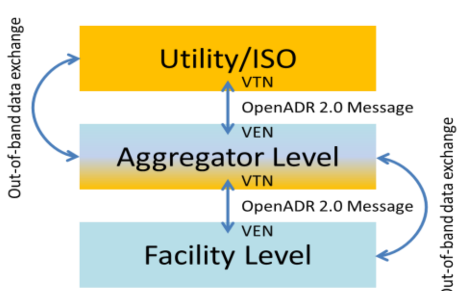

Figure 7: Data Interfaces

4. DATA SOURCES AND DATA ACCESS

4.1 Data Management and Analytics Problem in ADR

The following table displays an overview of data layers, stakeholders, levels, and typical data type examples as found in a fully-automated demand response schema.

More specifically, we have the following data layer categories and data type examples:

1. **Utility enterprise layer**: energy savings, pricing, reporting/monitoring, scheduling,
2. **Campus/District layer**: alarms, monitoring, scheduling, energy data,
3. **Premises/System layer**: energy mode, ADR signals, alarming, scheduling,
4. **Zone layer**: occupied mode, load shed mode, lighting scene,
5. **Room layer**: occupied mode, load shed mode, lighting scene,
6. **Device layer**: temperature, pressure, status, set points, mode, scene.

Layer	Stakeholder	Data level	Data Types Examples
Enterprise	Owner/Master integrator/Facility Staff/Application Developer/Aggregator	5, 6 - Aggregate, Monitor, Report	Energy savings, pricing, reporting/monitoring, scheduling
Campus/ District	Owner/Master integrator/Facility Staff/Application Developer/Aggregator	4, 5 - Schedule, Report, Monitor	Alarms, Monitoring, Scheduling, Energy Data,
Premises/ System	Owner/Integrator/Facility Staff/Application Developer	3, 4 - DR, Load shed, control, monitor, schedule	Energy mode, ADR Signals, Alarming, Scheduling
Zone	User/Occupant/ Manufacturer/Vendors/ Integrators	2, 3 - Status/Mode/Scene, schedule	Occupied mode, Load Shed mode, Lighting scene
Room	User/Occupant/ Manufacturer/Vendors/ Integrators	2 – Status Mode Scene	Occupied mode, Load Shed mode, Lighting scene
Device	Manufacturer/Vendors Integrators	1 - on/off/control, low level data	Temp, pressure, status, set points, mode, scene

Figure 8: Energy Data Types and Access in ADR

The data implied in the above six categories can be classified into the five energy data types listed in section 1.2.2. As a result, a comprehensive automated demand response configuration represents a specific case of the general description of a big energy data management and analytics problem as outlined in chapter 1.

4.2 Data Management and Analytics Solution for ADR

The following diagram (figure 5) represents a new and fully integrated software-development solution that has been applied in network industries such as telecommunications, transportation, defense, or financial services.

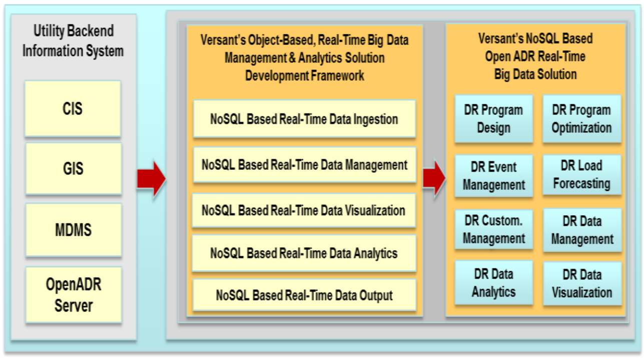

Figure 9: NoSQL Data Management & Analytics Framework

This platform allows for an automated means to ingest, manage, and analyze big and complex data volumes in real time. The approach is based on object-oriented programming, allowing data to be modeled as objects and classified into object classes. This produces the right match with the nature of data generated by network topologies. The platform's ability to model data entities as schema according to UML also corresponds with the UML reference model for the electric utility established in the IEC CIM standard, where interoperability between all network devices used in a Smart Grid is also specified. It will allow for a stochastic topological model to be established between the devices via network configuration models and associate real-time data with those topologies.

Given the objectives and the solution outlined above it provides the following integrated data management and analytics components for the ADR schema:

1. Developing, storing, retrieving, and managing of network configuration models such as the Smart Meter measurement network in the NoSQL database technology (leverages a relational database's industry standard API where needed).

2. Analyzing the runtime for optimal model connectivity to reflect network characteristics in a NoSQL implementation.

3. Evolving the afore-mentioned models for suitability in analytical methods for optimal situational awareness in an ADR schema.

4. Real-time data ingestion with big data Ingestion Framework. It embraces two options regarding the data ingestion API (Versant JPA API and Versant Ingestion API). Both interfaces can be fed with data coming from 1) Hadoop/MapReduce (discrete data extraction based on data transformation, data aggregation, and data summation), 2) real-time data monitors (streaming data extraction based on IP content data, such as twitter and the Web), data from sensor networks (such as traffic and energy), and transactional data (such as ticker), and 3) data virtualization (enterprise data extraction such as classic ETL system data). The ingested data goes through context discovery to enable semantic enrichment. External data can include weather data (lightning, wind, and precipitation), existing asset data, and performance of the communication network.

5. Build real-time model. Develop model of connectivity while ingesting data, providing a linked NoSQL data format that leverages emerging NoSQL data-management technology.

6. Object-oriented data-management/storage framework: Implemented architectural scale patterns commonly found in NoSQL technology. Deployment of both scale-up on modern n-core process architectures and horizontal scale-out patterns in partitioned systems.

7. Object data modeling to support NoSQL architectural patterns for optimal methods for situational-awareness analysis. This modeling includes the data management/storage regarding all of the time-series data generated by the network devices. This part leverages the huge experience in real-time data management with a variety of challenges to big-data management addressed for mission-critical applications in energy, transportation, defense, research, and media.

8. In-Database Analytic Framework: The In-Database Analytic Framework embraces a number of analytic features such as 1) graph closures, 2) sub-graph traversal queries, 3) discovery of unstructured data relations, 4) descriptive statistics, 5) inferential statistics, and 6) forecasting and prediction. The In-Database Analytics are empowered by 1) graph analysis (hyper-graph model analysis to enable path determination, pattern matching, and forensic data analysis), 2) complex event analysis (correlation of streaming data to enable moving averages, packet reassembly, or complex rule triggers), and 3) statistical/quantitative analysis (R/S+ to do statistics and machine learning). This capability includes applying a combination of the above techniques to identify patterns, revealing critical actors in system function necessary for controlling network stability under multiple simultaneous events. It will also address the identification of system steady-state basing and divergence-detection models.

Based on the availability of the network configuration data as well as real-time data from smart meters, PMUs, or SCADA systems it models and manages the device configuration as well as ingest and analyze the real-time data. It also provides the necessary platform and interface to allow for numerical methods necessary for large-scale simulation and optimization.

5. NOSQL DEMAND RESPONSE AUTOMATION

Demand response (DR) represents a concept where energy seller and buyer negotiate and agree on a commodity contract (energy represents the commodity) which rules how much energy can be saved on the energy procurement side and, therefore, does not need to be provided by the energy seller. Depending on the time of the day or the season of the year energy demand is adjusted (reduced) based on the contractual rules. In such a way, the prevention of load shedding scenarios is supported and, hopefully, not an issue at all moving forward.

In the current, non-automated demand response approach the two parties negotiate the necessary DR conditions within the contractual framework on a daily, weekly or even monthly basis. The means of communications to exchange the necessary information are phone, email and alike. It is not at all real-time or standards based as it could be if the interface between the contractual parties was properly defined and automated given that the necessary technology is available. In order to create an effective execution of the DR process governed by the commodity contract, it is simply required to layout an approach which is based on as much automation as possible. Only in such a way, the necessary broad participation in DR programs will be realized through productivity which enables the economics expected. The OpenADR industry consortium provides the correct response to the need for a non-proptietary, open-standards based DR interface between the contractual parties (energy seller and buyer) involved. It allows electricity sellers and buyers to communicate DR signals that use a common language and existing means of communications such as the Internet.

The DR scenario in a power delivery process represents a 3-stage process with two parties (energy seller and buyer) each involved in every stage of DR. More specifically, we find three contractual scenarios to describe a fully deployed DR solution starting with the power producer and ending with the power consumer:

1. First commodity contract between power producer and power transmission organization (energy seller: power producer, energy buyer: power transmission organization),
2. Second commodity contract between power transmission organization and power distribution organization (energy seller: ISO/RTO, energy buyer: utility),
3. Third commodity contract between power distribution organization and power end-consumers (energy seller: utility, energy buyer: power end-consumer).

Each of the three scenarios will be analyzed in the following to describe the DR process, the closed-loop control problem implied to automate it, and the data management and analytics challenges involved.

5.1. Automation of Demand Response between Power Generation and Transmission

The demand response scenario between power generation and transmission is based on a commodity contract agreed upon between the power producer and the power transmission organization (see figure 6). As an automated concept it represents a multi-loop control problem driven by the process of finding an optimum between the time-varying energy price negotiated and the corresponding energy amount provided and consumed. It is impacted by the time of the day, the season of the year, and geospatial locations of the end-consumers (constraints imposed).

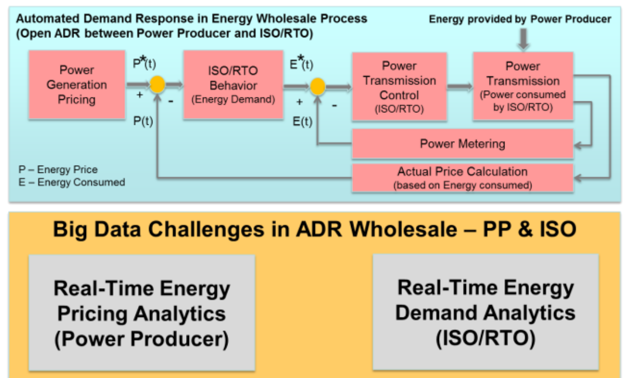

Figure 10: OpenADR between Power Producer and ISO/RTO

In order to build an automated DR solution the following challenges must be addressed from a data management and analytics perspective:

- **Finding/Predicting the optimal energy load profile (reference value) for every point in time of the year**: The transmission organization has the challenge to express a well-defined energy demand (reference value) at all times to every power producer contractually involved in the ADR solution (energy trading process). To do so, it needs to understand the required energy supply demanded by every connected utility (load profile) at all times.
- **Finding/Predicting the optimal energy price profile (reference value) for every point in time of the year**: The power producer has the challenge to calculate a well justified energy price (reference value) at all times to the transmission organization contractually bound to within the ADR solution (energy trading process). To do so, the power

producer needs to be provided with the total energy demand required by the transmission company (load profile) at all times.

A digital, real-time, automated DR approach, therefore, requires a big data prediction engine which can ingest, manage and analyze highly resolved big data streams in real time. This analytics engine must be capable of correlating the energy and pricing data ingested with the constraints (e.g. time of day, season, geospatial location) that define this nonlinear optimization problem to be solved. Only in such a way, meaningful reference values based on load profiles can be determined and used to identify the necessary energy to be supplied as well as the optimum price to be charged by the power producer.

5.2 Automation of Demand Response between Power Transmission and Distribution

The demand response scenario between power transmission and distribution is based on a commodity contract agreed upon between the power transmission company and the power distribution organization (utility). As an automated concept it represents a multi-loop control problem driven by the process of finding an optimum between the time-varying energy price negotiated and the corresponding energy amount provided and consumed (see figure 7). It is impacted by the time of the day, the season of the year, and geospatial locations of the end-consumers (constraints imposed).

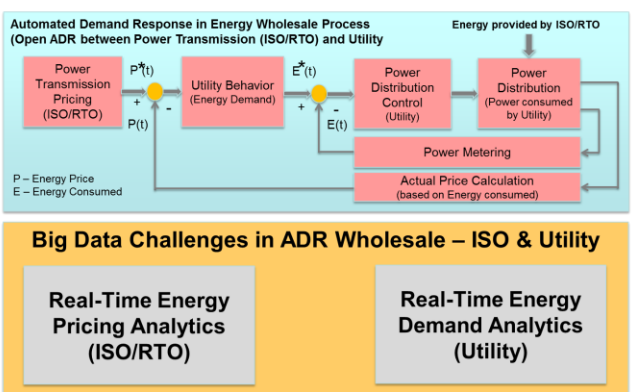

Figure 11: OpenADR between ISO/RTO and Utility

In order to build an automated DR solution the following challenges must be addressed from a data management and analytics perspective:

- **Finding/Predicting the optimal energy load profile (reference value) for every point in time of the year**: The distribution organization (utility) has the challenge to express a well-defined energy demand (reference value) at all times to every power transmission organization (ISO/RTO) contractually involved in the ADR solution (energy trading process). To do so, it needs to understand the required energy supply demanded by every connected power consumer (load profile) at all times.
- **Finding/Predicting the optimal energy price profile (reference value) for every point in time of the year**: The power transmission company has the challenge to calculate a well justified energy price (reference value) at all times for the power distribution organization contractually bound to within the ADR solution (energy trading process). To do so, the power transmission company needs to be provided with the total energy demand required by the utility (load profile) at all times.

A digital, real-time, automated DR approach, therefore, requires a big data prediction engine which can ingest, manage and analyze highly resolved big data streams in real time. This analytics engine must be capable of correlating the energy and pricing data ingested with the constraints (e.g. time of day, season, geospatial location) that define this nonlinear optimization problem to be solved. Only in such a way, meaningful reference values based on load profiles can be determined and used to identify the necessary energy to be supplied as well as the optimum price to be charged by the power transmission organization.

5.3 Automation of Demand Response between Power Distribution and Consumption

The demand response scenario between power distribution (utility) and the end-consumer is based on a commodity contract agreed upon between the utility and the power end-consumer. As an automated concept it represents a multi-loop control problem driven by the process of finding an optimum between the time-varying energy price negotiated and the corresponding energy amount provided and consumed (see figure 8). It is impacted by the time of the day, the season of the year, and geospatial locations of the end-consumers (constraints imposed).

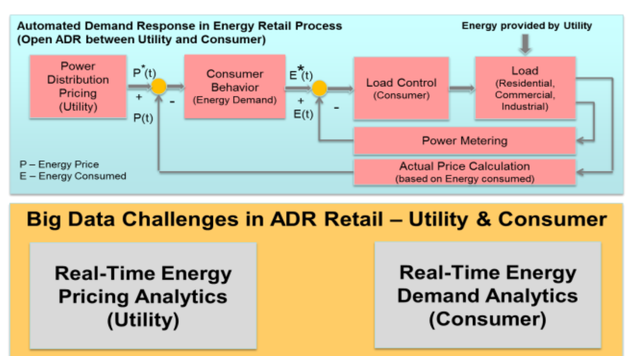

Figure 12: OpenADR between Utility and End-Consumer

In order to build an automated DR solution the following challenges must be addressed from a data management and analytics perspective:

- **Finding/Predicting the optimal energy load profile (reference value) for every point in time of the year**: The power end-consumer has the challenge to express a well-defined energy demand (reference value) at all times to the power distribution company (utility) contractually involved in the ADR solution (energy trading process). To do so, the consumer needs to understand the required amount of energy to be supplied (load profile) at all times.
- **Finding/Predicting the optimal energy price profile (reference value) for every point in time of the year**: The power distribution company has the challenge to calculate a well justified energy price (reference value) at all times to be charged to the end-consumer contractually bound to within the ADR solution (energy trading process). To do so, the power distribution company needs to be provided with the total energy demand required by the end-customer (load profile) at all times.

A digital, real-time, automated DR approach, therefore, requires a big data prediction engine which can ingest, manage and analyze highly resolved big data streams in real time. This analytics engine must be capable of correlating the energy and pricing data ingested with the constraints (e.g. time of day, season, geospatial location) that define this nonlinear optimization problem to be solved. Only in such a way, meaningful reference values based on load profiles can be determined and used to identify the necessary energy to be supplied as well as the optimum price to be charged by the power distribution organization (utility).

6. CONCLUSION

In order to create a fully automated demand response solution, a variety of real-time, big data challenges need to be addressed in a complex commodity-type contractual framework with power producers, transmission companies (ISO/RTO), power distribution organizations (utilities) and power end-consumers as the main stakeholders involved. Big data analytics engines need to solve nonlinear optimization problems to determine optimal pricing and predict load profiles under the constraints of changing times of the day, seasonal differences and different environmental conditions depending on the locations of the end-consumers. A suitable real-time big data management solution needs to be integrated with the analytic solution to support the necessary data formats suitable for the analytics to be performed. Only a linked object-based big data management and analytics solution framework can address the large data diversity (5Vs, 5 different data types and corresponding variances in sample rates) found in demand response schemes as well as the connectivity models (e.g. UML CIM) to provide an integrated big data solution for running interoperable automated demand response (ADR) in real-time which supports the OpenADR standards to ensure non-proprietary solutions.

7. ACKNOWLEDGEMENT

The authors would like to thank Barry Haaser from the OpenADR Alliance, Dave Hardin from EnerNoc, Mark Knight from Logica, and Girish Ghatikar from Lawrence Berkeley National Laboratories for their invaluable assistance and contributions to various aspects of this paper's knowledge base.

References

ANSI/ASHRAE. (2001). "BACnet®: A Data Communication Protocol for Building Automation and Control Networks." ASHRAE. REPLACED by ANSI/ASHRAE 135-2004.

Beecher, P., J. Lin. (2010). "Guiding Principles Document." Utilities Communications Architecture Open Smart Grid Conformance Task Group.

Doug Laney, 3-D Data Management: Controlling Data Volume, Velocity and Variety, February 2001

Ghatikar, G., M. A. Piette, S. Fujita, A. McKane, J. H. Dudley, A. Radspieler Jr., K. C. Mares, and D. Shroyer. (2010b). "Demand Response and Open Automated Demand Response Opportunities for Data Centers." LBNL-3047E

Jeffery Taft, Paul De Martini, and Leonardo von Prellwitz, "Utility Data Management & Intelligence", Cisco White Paper, May 2012.

Kiliccote S., M. A. Piette, G. Ghatikar, E. Koch, D. Hennage, J. Hernandez, A. Chiu, O. Sezgen, and J. Goodin. (2009). "Open Automated Demand Response Communications in Demand Response for Wholesale Ancillary Services." In Proceedings of Grid-Interop Forum 2009, November 17–19, 2009. LBNL-2945E.

Motegi, N., M. A. Piette, D. Watson, S. Kiliccote, and P. Xu. (2007). "Introduction to Commercial Building Control Strategies and Techniques for Demand Response." LBNL-59975

NIST. (2010). "NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 1.0." NIST Special Publication 1108

NIST. (2011). "Draft NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 2.0."

OASIS Energy Market Information Exchange Technical Committee (EMIX TC). (2011). http://www.oasis-open.org/committees/emix

OASIS Energy Interoperation Technical Committee (EI TC). (2011). http://www.oasis-open.org/committees/energyinterop

OASIS Web Services Calendar Technical Committee (WS-Cal TC). (2011). http://www.oasis-open.org/committees/ws-calendar

The OpenADR Alliance (Alliance). (2011a). "The OpenADR Primer: An Introduction to Automated Demand Response and the OpenADR Standard."

The OpenADR Alliance (Alliance). (2011b). "OpenADR Message Transport Mechanism White Paper." (DRAFT)

Piette M. A, G. Ghatikar, S. Kiliccote, E. Koch, D. Hennage, P. Palensky, and C. McParland. (2009a). "Open Automated Demand Response Communications Specification (Version 1.0)." California Energy Commission, PIER. CEC-500-2009-063.

Piette, M. A., G. Ghatikar, S. Kiliccote, D. Watson, E. Koch, and D. Hennage. (2009b). "Design and Operation of an Open, Interoperable Automated Demand Response Infrastructure for Commercial Buildings", *Journal of Computing Science and Information Engineering*, vol. 9, issue 2, no. 2.

Piette, M. A., D. Watson, N. Motegi, and S. Kiliccote (2007a). Automated Critical Peak Pricing Field Tests: 2006 Pilot Program Description and Results. LBNL-62218

OASIS Energy Interoperation Technical Committee. (2011). "Energy Interoperation Version 1.0: Committee Specification Draft 02/ Public Review Draft 02)

Wikler, G, I. Bran, J. Prijyanonda, S. Yoshida, K. Smith, M.A Piette, S. Kiliccote, G. Ghatikar, D. Hennage, and C. Thomas. (2008). Pacific Gas & Electric Company 2007 Auto-DR Program: Task 13 Deliverable: Auto-DR Assessment Study.

Biography

Bert Taube leads Versant's efforts to build advanced NoSQL data management and analytics solutions applicable to large and complex data challenges in the energy and smart grid sector. Prior to Versant, Mr. Taube held a variety of senior management positions in energy & automation as well as in semiconductors for companies such as Siemens or Supertex. He developed intelligent project solutions (FACTS, HVDC) integrated in high-voltage power transmission networks. He also served as an advisor, investment & technology consultant for several energy and clean technology startups in Silicon Valley. Mr. Taube holds a Ph.D. in electrical engineering from the University of Rostock and an MBA from the Kellogg School of Management at Northwestern University.

Rolf Bienert is the Technical Director of the OpenADR Alliance. In this capacity, Rolf oversees the technical developments and the currently forming certification program of the Alliance. Most recently Rolf was the Technical Manager for Telecom and the Global Competence Center for emerging communication protocols at a global testing firm. In this role, Rolf was involved in many international standardization efforts and guided the company's efforts in these areas. Rolf has been an active member of the NIST SGTCC, OpenADR Alliance, USNAP Alliance, SunSpec Alliance, ZigBee Alliance, and other organizations driving the development of new technologies with a specific focus on certification and interoperability. Rolf holds a Masters Degree in EE from the TH Saarbruecken, Germany.