The Data Management Challenge:

Making Extremely Large Amounts of Data Useful and Actionable

Todd Pistorese

OSIsoft
777 Davis Street
San Leandro, CA 94577

Summary

The introduction of AMI has created a data paradox. Traditional AMR solutions built with relational database products are, for the most part, not equipped to perform the real-time processes needed for the future growth of AMI and SmartGrid applications nor are they capable of collection and dissemination of historical data at its original resolution. Yet the promises of AMI will only come from elimination of data latency and the ability to maintain all data at its original resolution. A new approach needs to be taken by viewing the AMI from an operational perspective as well as billing perspective. Products that have supported real-time use of data from electrical networks worldwide are now poised to become the Operational Data Managers in the new AMI and will deliver on the promise of closer ties between market price signals and customer usage of electricity.

Industry History – Supply-Side

Throughout the 100+ year history of the electric power industry distribution systems have suffered from a lack of clear information about the state of the network. As independent utility companies grew, they interconnected with neighboring utilities to provide a modest amount of system stability to their otherwise vertically integrated operations. These connections were initially relatively weak and relied upon primarily for regional power exchanges caused by seasonal variations in demand and availability of generation resources. Since deregulation of the industry and separation of vertically integrated utility business units, the dependence upon a tightly integrated transmission grid has become a major issue. The lack of network capacity in many regions has lead to extreme price fluctuation and rapid rate increases for captive consumers.

Industry Future - Demand-Side

Another alternative to the expensive proposition of supply side expansion and improvement has emerged to complement this strategy. Demand side initiatives coupled with vast improvements in communications capabilities have brought about a significant opportunity in the industry. While transmission engineers and operators have had the ability to view and control the transmission network from end to end for many years, distribution operations have not had this ability. Generally speaking, distribution is a “black box” operation with very little known about the detailed consumption habits of the millions of small residential and commercial users of electricity. Engineers and operators generally know the demand on the circuit breakers at the distribution substations but have no visibility into the minute to minute consumption of electricity beyond the substation fence. Furthermore, the disconnect between price signals and consumption has created an economic crisis for utilities who are limited by regulation on the prices they can charge yet have obligations to serve customers in a volatile market.

Utilities need a more immediate connection between electricity demand and price signals to the end consumers in order to change consumption habits. Traditional utility billing systems rely on highly aggregated data to perform monthly billing. When the industry was composed of vertical utilities which owned their generation sources and sold to captive regional customers in their territory, monthly billing was adequate and easily accommodated the very steady prices for company owned generation. The volatility in electric prices introduced by deregulation and decoupling of generation sources from load serving entities has created a huge need to send immediate price signals to consumers as a means to alter customer usage habits and stabilize electric rates. The uproar created by vast regional service interruptions and volatile electric rates coupled with growth in electric demand has created a tipping point where "business as usual" is not a viable option.

Advanced Metering Infrastructure (AMI)

AMI has emerged as the likely solution for demand-side management and offers a host of new possibilities for appropriate utility rate structures, new consumer services, and price stability. As vendors rush into this new market offering a variety of technologies, utilities are faced with tough decisions about what they need, who should supply them, and whether the solutions they choose can scale and be expanded to support future customer service offerings.

Advancements in communications, availability of broadband, high speed networks and advanced digital metering technologies provide a means to eliminate the price/use disconnect, provide consumers and utilities with a mechanism to intelligently moderate demand during periods of high electric prices and effectively let natural market economics reshape the demand curve for electricity consumption.

The future of electricity consumption will be a smart home that reacts to price without customer intervention once that customer sets their "inconvenience tolerance". Management of consumption even down to individual appliances within the residence will be a common occurrence in the future Smart Grid.

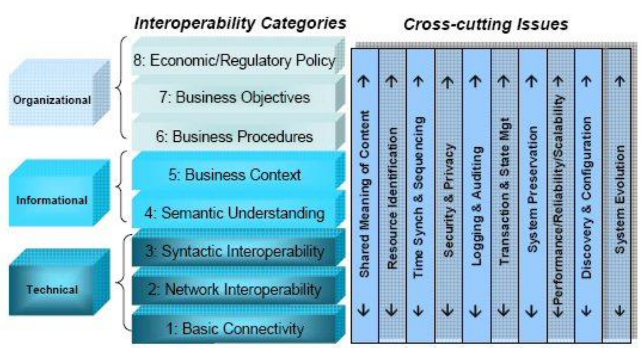
The opportunities (and problems) AMI introduces are complex. Many existing, repackaged AMR solutions cannot scale, suffer from data communication latency, and are not self-configuring or bi-directional. Solutions that have never before had to manage high-speed, massive volumes of distributed data were logically built using relational database designs. These products came to the marketplace from a need for aggregated data used for monthly billing purposes. They never contemplated demand-side management of electricity use, complex billing, operational uses for data or the need for long term retention of high resolution meter data, device control signals, and variable price signals.

The logical approach of a relational database "billing centric" solution to more data volume has been to introduce a "clearing house" for meter data referred to as a Meter Data Management System (MDMS). Because relational databases are very good at transaction processing (like an ATM machine withdrawal) but inadequate for time-series data management and distributed processing of events the MDMS proposes to summarize incoming data into hourly or even less granular bits of information. This eliminates any possible use of the data for operational analysis of customer usage, creates further data latency that all but makes real-time price signals moot, and creates an inflexible environment for analyzing sophisticated Time of Use (TOU) rate structures. Furthermore, data aggregation at the MDMS is done in a predefined manner which tightly couples the backend systems to whatever aggregation decisions are made today. Rate analysis using different TOU structures against historical data is rendered impossible. And the subtle changes in consumption patterns are lost in data aggregation.

A fundamental shift in the utilities perspective on the problem must occur before a viable solution can be implemented. To complement the new breed of sophisticated metering and device control products being created, utilities need to look at the problem more closely from an operational point of view while retaining the ability to aggregate data for billing purposes and ensure system scalability for more complex, direct associations between customer demand and electric prices. Someday, all new real-time billing solutions will be required but a well designed data collection and dissemination solution today will allow utilities to defer investment in new billing and accounting systems while enabling phased implementation of AMI. What is needed is an Operational Data Manager.

The Case for an Operational Data Manager

The Gridwise Interoperability Framework defines several cross cutting issues that in fact have been topics of conversation and system design in the real-time operations world for several years.

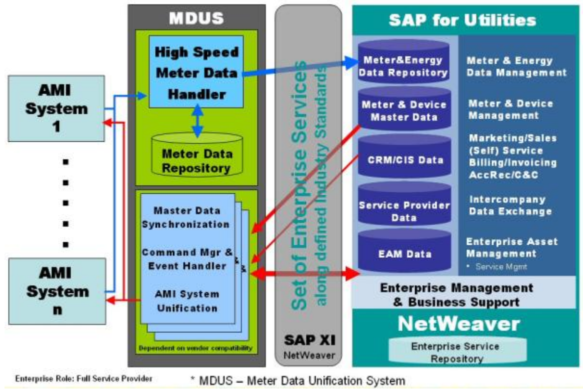


GridWise Interoperability Framework

Most of these issues have been addressed in an environment that is extremely time sensitive and dependent on high reliability, where decision delays or failure to deliver data would result in blackouts and significant loss of revenue. This environment is the realm of real-time operation systems.

Eventually, the validation and aggregation functions of the MDMS will be dissected and moved into more appropriate parts of the meter to utility data flow to eliminate data latency and enhance flexibility and scalability. Head-end systems will become immediate data validators. Operational Data Management systems created using highly efficient, real-time solutions derived from operations oriented products will not only hold all incoming data at original resolution and maintain history on real-time metering and device control signals, but they will be device-aware and automatically accept new data when a meter is added to the AMI network, maintain history on device relationship to the distribution network, and they will tightly integrate with spatial systems (GIS) that manage the distribution assets of the utility.

Back office systems will have the ability to change rate structures beyond the limited, preconfigured boundaries of a MDMS, as the utility receives approval for TOU rates. The Operational Data Manager will provide data in any aggregation the billing and accounting systems demand for any timeframe the utility wants to review. Creation of TOU rates will be based on revenue analysis using historical consumption information available at its original fidelity. Utilities will be able to make very convincing cases before their Public Utility Commissions on the positive aspects of closer coupling of price signals to consumer demand using historical consumption patterns against proposed TOU rates.



SAP's view of AMI

A marriage of technologies used traditionally for network control and real-time data management with sophisticated head-end metering and control devices will enable utilities to implement AMI in a phased in and financially manageable manner without the need for a "big bang" wholesale replacement of all back office accounting and customer billing systems in order to gain the benefits of AMI.

A significant benefit from moving operational real-time solutions into the AMI space will come from better visibility into the consumption patterns (and customer response to real time price signals). For instance, distribution operators will be able to optimize networks based on data collected by the AMI system and combined with SCADA, distribution automation, capacitor control systems, equipment sensors, etc. to:

- Manage distributed generation resources;
- Maximize feeder efficiency;
- Manage circuit voltage profiles;
- Monitor grid equipment health;
- Optimize circuit loading;
- Reduce outage response times and switching analysis;

- Monitor demand response events.

As such, an Operational Data Manager must be able to manage extremely large volumes of data from a variety of sources with multiple data frequencies and latencies, including real-time. An Operational Data Manager needs to:

- Be head end system vendor and meter or demand device neutral and capable of interfacing to all device management systems
- Be scalable;
- Be highly available;
- Self discover and configure device additions and changes;
- Support aggregations and complex calculations on demand,
- Synch with other systems of record;
- Manage event data;
- Store multiple years of data online;
- Support multiple hierarchies (routes, network, asset types);
- Reconcile data with other source systems;
- Provide security and health monitoring of IT systems;
- Provide easy-to-use data access tools for a variety of users.

By combining the data collected by the Advanced Metering Infrastructure (AMI) system and SCADA, distribution automation, capacitor control systems, and equipment sensors, distribution operators will be able to optimize networks to:

- Manage distributed generation resources
- Maximize feeder efficiency
- Manage circuit voltage profiles
- Monitor grid and distribution device health
- Optimize circuit loading
- Reduce outage response times and switching analysis
- Monitor demand response events.

Requirements of an Operational Data Manager

For the intelligent grid and smart grid, an Operational Data Manager should meet a number of requirements in order to be effective for utilities' needs both today, and in the future with AMI. The features and requirements outlined below will provide a useful checklist that can be used when considering vendors for your Operational Data Manager and compliance with recommendations of the GridWise Interoperability Framework (GWIF).

Scalability
(GWIF – System Evolution)

AMR systems were originally designed to replace the manual meter reading conducted for monthly billing of retail customers. These systems and infrastructure, however, are not capable of handling readings of greater fidelity.

AMI systems, on the other hand, are being designed to gather data for sophisticated new programs outside of meter reading for billing such as demand response and service reliability. These projects will include the data from large numbers of meters as well as many other real-time sensors on distribution devices.

Moving toward the intelligent and smart grid, it is estimated that there will be between six and twenty times the number of meters in terms of points to be measured—current, voltage, status, peak values, external sensors, internal devices, etc. The Operational Data Manager should make these readings and events available and actionable to the operations center in real-time to support grid management.

High Availability
(GWIF – System Preservation)
AMI systems will provide the meter and device data for advanced functionality at utilities to support grid management via demand response and other data intensive applications. Many of these applications will require availability on the same par as mission critical environments like SCADA.

Specifically, the Operational Data Management System will be required to respond to various equipment and telecommunications failures, security patch and operating system upgrades, and back-up of both the data and the system. These and other events will need to be performed on-line with no data loss or loss of functions.

Demonstration of the high availability approach with particular emphasis on no data loss and non-stop function will be an important requirement of the Operational Data Manager.

Smart Connectors
(GWIF - Discovery and Configuration)
AMI systems are extremely large. As such, it is not practical to require manual configuration of these systems, either during the initial build or for updates as changes arise. For this reason, resources within the Operational Data Manager that support all interfaces to AMI data, external data or structure must be built and maintained without requiring manual intervention. To meet these requirements, interfaces must self discover and automatically configure device additions and changes. Smart connectors are an example of this type of interface.

Meta-Data Management
(GWIF – Resource Identification)
A good Operational Data Management System must have the ability to share resource definitions and configuration information with other products such as GIS. Part of the function of a Meta-Data management component of the consists of the ability to maintain procedures, relationships, models, and aliases to the AMI head-end data points. Meta-Data should be manageable from any of the integrated systems and have the ability to identify when an object is introduced to the AMI network that is not part of the network model. This concept advances Smart connectors to smart models.

Analytics
(GWIF - Performance/Reliability/Scalability)
The net result of the AMI system data collection will be a stream of data that contains meter readings, events, and status messages. Extremely high volumes of events will need to be managed and processed in near real-time, with no data loss. This processing may include calculations for validation, data framing, event filtering, demand response results and notifications to support business processes, such as billing, tamper and theft, outage management and grid management. Near real-time analytical capabilities such as support for aggregations and complex calculations will be an

important feature of Operational Data Managers. Descriptions of how events are managed and processed at this scale, as well as the analytical and reporting capabilities are key requirements for an Operational Data Manager.

Data Synchronization (GWIF – Time Synch and Sequencing)

Data synchronization across systems of record is an important component of the Operational Data Manager and an important requirement for an intelligent grid. AMI data will need to be reconciled and combined with SCADA, Distribution Management Systems (DMS) and other operational systems complete with audit trails.

This data needs to be synchronized with systems that contain connectivity and asset information such as GIS, OMS and EAM systems. In addition, analysis of multiple versions of connectivity will be important. For example, operations may need analysis and knowledge of the system from last year's peak day as the system was switched on that day where the planning department may need to analyze the steady state system for the same time period. Operational Data Managers need to be able to provide multiple models for multiple audiences in the utility.

Data Presentation and Access (GWIF – Shared Meaning of Content and Resource Identification)

The Operational Data Manager is primarily responsible for the data stream of readings and notifications that come from meters and devices and the proper analysis and interpretation. Internally, this is a combination of the data stream traffic, results of the analytics, and any context needed to make sense out of the information. Many other utility systems will need this information in an open environment. The Operational Data Manager will need to provide data access via a Services Oriented Architecture to these other applications and support data standards such as CIM, IEC61850 and OPC.

Event Management (GWIF - Performance/Reliability/Scalability and Transaction State Management)

Many applications need to be informed by exception of changes or specific events. It is not practical for these systems to actually process the raw data stream at the rate managed by the Operational Data Manager.

For example, an outage notification or "last gasp" of a meter can indicate an outage or a meter change or theft. Event management must not only identify the events for each application but also ensure that the event was actually transmitted and that the event was "real."

The event management function of the Operational Data Manager must be scalable both for the number of events and the numbers of consumers of that event. This feature is critical to making real-time data useful and actionable for operations today and ultimately the management of the intelligent grid.

Historization of Data (GWIF – System Preservation)

Utilities will have higher fidelity information than has been available in the past and, in particular, data that will support operations, planning, scheduling, cap and trade certification, and other real-time functions. An often overlooked component of the Operational Data Manager is the historization of this high fidelity information. Multiple years of on-line storage and fast retrieval of this data is critical to intelligent analysis and action.

Network Support
(GWIF – Security and Privacy)
Experience with AMR projects has demonstrated the importance of the health of the network and communication infrastructure. For AMI systems this importance will be compounded and with a much higher data throughput. For this reason, the Operational Data Manager should support state-of-the-art forensics and monitoring for a large scale self-healing network.

Reconciliation and Validation
(GWIF – Logging & Auditing)
In addition to the real time analytics that work directly on the data stream, meter and device data must be subjected to significant *reconciliation* and *validation* to ensure what was sent out was received and to ensure readings are correct. The Operational Data Manager must have reconciliation and validation processes that are fast, reliable and flag suspect data. Tracking of any edits, changes, deletions or alteration of data must be logged in an auditable database.

High Security Installation
(GWIF – Security and Privacy)
AMI systems with Home Area Network (HAN) capability will extend the reach of the utility to users (residential, commercial and industrial) and possibly make these systems a potential target for cyber terrorism security breaches. High security installations must include a system for all online application of updates from the software vendors involved (Operating System, EMS, SCADA, AMI systems, etc.). This feature should be tightly integrated with high availability.

Health Monitoring and Reporting
(GWIF – Security and Privacy and Discovery & Configuration)
AMI systems may be distributed throughout a service territory which may or may not be contiguous and may or may not have dedicated, secure communications. As such, it will be necessary to provide security and health monitoring of IT systems. The Operational Data Manager should support this feature.

An Intelligent Grid and Smart Grid

It is critical to make data useable, actionable, and accessible to multiple entities internal and external to the utility when creating your path to the intelligent grid.

The Operational Data Manager can serve as your one tool that adds value by combining all operational data sources together, along with meter data, even when the meter data is daily or monthly. Providing all data in a single Operational Data Manager and seamlessly accessing other systems of record provides users with a complete operational view enabling short term goals such as effective asset management and long term goals such as the intelligent grid. If implemented effectively, the Operational Data Manager can provide utilities a robust infrastructure that:

- Offers extensive and flexible data collection capabilities

- Provides important data management requirements such as data archiving, fast data retrieval, event management, alarming, real time SQC, scheduling, and advanced calculation, and
- Enables intelligent initiatives such as the smart grid and smart substation

Conclusion

Data management is a critical ingredient in creating the intelligent grid. The Operational Data Manager is also an important component to maximizing the benefit of AMI systems by providing timely reconciliation of AMI data with all critical operational data sources and making it actionable to multiple audiences internal and external to the utility.