

Managing big data for smart grids and smart meters

*Meet the challenge posed by the growing volume, velocity
and variety of information in the energy industry*



Executive summary

Evolving technologies in the energy and utilities industry, including smart meters and smart grids, can provide companies with unprecedented capabilities for forecasting demand, shaping customer usage patterns, preventing outages, optimizing unit commitment and more. At the same time, these advances also generate unprecedented data volume, speed and complexity.

To manage and use this information to gain insight, utility companies must be capable of high-volume data management and advanced analytics designed to transform data into actionable insights. For example, designing effective demand response programs requires that utilities execute advanced analytics across a combination of data about customers, consumption, physical grid dynamic behavior, generation capacity, energy commodity markets and weather.

This paper examines the business requirements, technical challenges and IBM solutions for a variety of data-driven decision-making and planning imperatives in the energy and utilities industry.

Leveraging powerful new opportunities in the energy industry

Today, many utilities are moving to smart meters and grids as part of long-range plans to ensure a reliable energy supply, incorporate distributed generation resources, develop innovative storage solutions, reduce the need to build new power plants and enable customers to have more control over their energy use.

Many are deploying smart meter systems as a first step, which means they have an immediate technical challenge on their hands. Going from one meter reading a month to smart meter readings every 15 minutes works out to

96 million reads per day for every million meters. The result is a 3,000-fold increase in data that can be overwhelming if not properly managed.

There is an upside, of course: the additional data generated opens up new opportunities, allowing energy companies to do things they never could before. Data gathered from smart meters can provide better understanding of customer segmentation, behavior and how pricing influences usage—if companies have the capability to use that data. For example, time-of-use pricing encourages cost-savvy retail customers to run their washing machines, dryers and dishwashers at off-peak times. These customers not only save money but also require less generation capacity from their energy providers, which means lower capital outlay for new generation and overall greater operational efficiency for utilities.

But the possibilities don't end there. With the additional information available from smart meters and smart grids, it is possible to transform the network and dramatically improve the efficiency of electrical generation and scheduling. However, the new mix of resources available requires more granular forecasting, load planning and unit commitment analysis than ever before to avoid inefficient energy trading or dispatching too much generation.

The ongoing growth in micro-generation is resulting in more small generators scattered geographically throughout a region, as opposed to the traditional model of a centralized power plant serving a large area. The advent of the “prosumer”—the consumer who also produces electricity for the grid—adds to the resource mix. Planning for a potential surge in electric vehicles and charging stations, which both charge and discharge from the grid, adds to the complexities and opportunities. The intermittency of wind and solar generation must also be factored into utilities' calculations.

Using predictive analytics on their data, companies can make a wide range of forecasts such as:

- How much excess energy will be available, when to sell it and whether the grid can transmit it
- When and where equipment downtime and power failures are most likely to occur
- Which customers are most likely to feed energy back to the grid, and under what circumstances
- Which customers are most likely to respond to energy conservation and demand reduction incentives
- How to manage the commitment of larger, traditional plants in a scenario where peaks from distributed generation are becoming relevant

To capitalize on these possibilities, organizations in the energy and utility industry require solutions that can help them manage high volumes of data and analyze patterns in that data to optimize business outcomes.

Generating big value from big data

Given this changing data environment, utilities are focusing increased attention on business intelligence and advanced analytics to support data-driven decision making and planning. These analytics require an integrated view of company data and alignment of data across disparate operational groups and lines of business. Utilities that integrate and analyze this data can gain insight into their operations and assets, enabling them to take proactive action rather than simply reacting to events after they happen. The results may include increased profitability, a reduced carbon footprint, increased safety, enhanced regulatory interaction and improved customer satisfaction.

Making the most of information from smart meters and smart grids increasingly requires dealing with what is called big data—meaning data that is high in volume, velocity, variety or all three. While data volumes in the power industry may not equal those of traditionally data-intense industries, they are growing larger than many power companies may be prepared to handle.

Velocity refers to the speed requirement for collecting, processing and using the data. This is likely to be relevant for data generated by sensors and new grid instrumentation. Many analytical algorithms can process vast quantities of information if there is time for the job to run overnight. But for real-time tasks such as equipment reliability monitoring, outage prevention or security monitoring, overnight is not good enough. Companies must be prepared for streaming data and relatively large volume data movement.

Variety signifies the increasing array of data types, which are collected not only from traditional sources like industrial control systems but also from security cameras; weather forecasting systems, maps, drawings and pictures; and the web. The variety of data is likely to become increasingly pertinent to utilities as they begin to analyze social media and call center dialogues as part of their decision-making and planning processes.

To obtain the most value from their data, many IT organizations focus on leveraging real-time data sources and analytics, bringing together multiple data sources. They are turning to techniques such as filtering and analyzing data on the fly, using tailored analytics tools to process a variety of data in its native format, and bringing arrays of parallel processors to bear on incoming data.

IBM offers an integrated suite of products designed to enable IT to leverage big data in a variety of ways that can contribute to the success of energy companies. Capabilities include time-series data flow, streaming data analysis, data security, data warehousing archiving, data mining and reporting. Each of the following power and utility industry use cases presents technical challenges that IBM products and capabilities are designed to address:

- Managing smart meter data
- Monitoring the distribution grid
- Optimizing unit commitment
- Optimizing energy trading
- Forecasting and scheduling loads

Managing smart meter data

Managing the large volume and velocity of information generated by short-interval reads of smart meter data can overwhelm existing IT resources. Ensuring the privacy of sensitive customer meter data is also a major issue in smart meter deployments. Additionally, meter data must typically be retained for many years to satisfy emerging regulatory requirements.

Technical challenges

Data storage costs can explode due to increased data volumes and retention requirements if organizations are using traditional, relational database technologies. Additionally, report generation and analytics can become painfully slow due to high volumes—companies may not be able to load and analyze all of the information fast enough to support decision making. Applications begin to drag and IT may struggle to meet service-level agreements.

IBM big data solutions

- Use IBM® Informix® TimeSeries to capture and load meter data as part of a meter data management system.
- Deploy IBM InfoSphere® Optim™ Data Management solutions for archiving meter data to comply with mandated retention periods.
- Use IBM InfoSphere Guardium® software to help ensure customer information privacy.
- Employ IBM InfoSphere Streams to process and analyze data in motion.
- Leverage IBM Netezza® data warehousing appliances and IBM InfoSphere BigInsights solutions as the stores for deep analytics and customer behavior analysis.

IBM big data platform capabilities at a glance

Integrated data management solutions and analytics engines from IBM enable energy and utility companies to:

Analyze a variety of information

- Apply novel analytics to a broad set of mixed information
- Analyze mixed data sets that could not be analyzed before

Analyze information in motion

- Perform streaming data analysis
- Easily handle large volume data bursts

Analyze high volumes of information

- Cost-efficiently process and analyze petabytes of information
- Manage and analyze structured, relational data to extract value

Discover and experiment

- Perform ad hoc analysis
- Enable data discovery and experimentation
- Gain insight and value from large volumes of low-economic-value data

Manage and plan

- Enforce data structure, integrity and control
- Ensure consistency for repeatable queries

IBM software supports the full range of information lifecycle requirements for smart meter and smart grid applications.

Meter data management benchmark results show speed of Informix TimeSeries

In addition to proof of concept tests, IBM and AMT-SYBEX developed a benchmark to demonstrate the capability of the Informix TimeSeries software to enable the Affinity Meterflow™ application to offer exceptional linear scalability. The benchmark was conceived and conducted to illustrate these benefits in tangible ways. It tested the preparation, loading and validation, estimation and editing (VEE) of meter data for a 10 million-meter utility, as well as a “day-in-the-life” scenario for a 100 million-meter utility. The results showed that daily end-to-end processing time remained constant with increased historical data (see Table 1), and storage use remained linear with increasing data. Over the 31-day test period, total storage required for interval and register data for 100 million meters was less than 4 terabytes—a fraction of the amount often required for a meter group this size.

Process	Average elapsed time	Average throughput rate (records/sec)
Preparation and technical verification	2 hrs. 10 min.	628,205
Data load	3 hrs. 14 min.	420,962
Validation, estimation and editing (VEE)	2 hrs. 11 min.	623,409

Table 1: Scalability benchmark results

To learn more about the results, please visit ibm.com/software/data/informix/smart-meter

Monitoring the distribution grid

Utilities need to proactively identify abnormal conditions and take action to both prevent power delivery disruptions and optimize overall grid reliability. Distribution companies can improve both customer satisfaction and regulatory compliance by reducing the number and duration of power outages.

Causes of the outages must be identified quickly and prioritized appropriately so crews can be efficiently dispatched with the correct solutions.

Technical challenges

Monitoring grid operations in real time involves large volumes of high-velocity data. Companies must also be able to make correlations between network events and network failures, understand which patterns indicate network problems, pinpoint fault locations and identify solutions to prevent service disruptions.

IBM solutions

- Use InfoSphere Streams to detect anomalies and correlate network events and failures in real time.
- Analyze historical data at rest in data warehousing appliances to determine patterns that result in major outages, and embed that pattern analysis into InfoSphere Streams to detect patterns and take corrective action to help prevent future outages.
- Integrate the IBM solutions with existing outage and distribution management systems and enterprise asset management systems.

Optimizing unit commitment

Companies must optimize the scheduling of their generation assets, taking into account a broad range of constraints to generate an optimal solution. These considerations include cost, emissions, ability to use existing delivery infrastructure and other factors—for example, wind and solar energy sources may be heavily weather-dependent and intermittent, requiring analysis of large weather data sets to forecast output. Based on emerging events, workforce members and other key assets must be dynamically re-prioritized to focus attention on the highest priority resources.

Technical challenges

Analysis must accurately predict which units need to be operational to meet but not exceed demand. This ability allows an energy company to optimize its energy source mix and avoid

both unanticipated excess capacity and costly spot market purchasing. Planners need detailed and accurate demand and supply forecasts by time and location. Large volumes of diverse information are needed to dynamically run the necessary algorithms to reflect unforeseen factors.

IBM solutions

- Employ IBM WebSphere® ILOG® optimization software to execute unit commitment algorithms.
- Run predictive analysis using IBM SPSS® software.
- Use IBM Netezza data warehousing appliances to manage data and accelerate business intelligence queries.
- Employ IBM Cognos® reporting software for presentation of results.

Forecasting and scheduling loads

Accurate demand forecasting is essential to energy planning and trading. Companies must be able to predict when they can profitably sell excess power and when they need to hedge supply.

They must determine when it is economically advantageous to buy, sell or trade power on the open market. By anticipating purchases well in advance, organizations are better able to obtain favorable prices.

Accurate load forecasting is critical for scheduling generation operations. Increasingly, energy companies need to incorporate various renewable sources and electric vehicles into the generation mix, and determine where to spend on new capabilities to address changing demand patterns of the future. The introduction of distributed and micro-generation significantly extends the complexity of load and capacity forecasting. Load forecasting models must also consider weather and energy trading conditions.

Technical challenges

Companies must understand which parameters—temperature, weather, day of the week or month, holidays, prior usage, price incentives and others—actually drive demand. To achieve this insight, large volumes of granular historical information must be analyzed and correlations identified.

Case in point: A power delivery company conquers big data

Managing the data involved in electronically reading millions of smart meters is the first dimension of the big data challenge for many utilities. Smart meter data management based on Informix TimeSeries enables an electricity utility in the US to support time-of-use pricing and pinpoint outages while reducing operations and maintenance costs.

Table 2 shows some of the performance and storage efficiency improvements the utility realized with Informix TimeSeries, including dramatically reducing the amount of time required to perform critical data analysis and storage tasks.

Challenges	Other relational database	Informix TimeSeries
Load time for 1 million meters	7 hours	18 minutes
Time to run required reports	2 to 7 hours	25 seconds to 6 minutes
Storage required for 1 million meters	1.3 TB	350 GB

Table 2: Performance and storage efficiency improvements with Informix TimeSeries

IBM solutions

- Leverage IBM Netezza for management of historical conditions and demand patterns.
- Use IBM SPSS predictive analytics solutions to uncover relevant patterns and drivers of demand and mine future demographic projections to determine where to build new plants and transmission facilities.
- Run WebSphere ILOG software to optimize generation scheduling and dispatch.
- Employ Cognos reporting software to present results.

Defining data strategy according to business imperatives

As companies deploy smart meter systems and plan for smart grid initiatives, it is essential to consider which data management and analytics capabilities are the most appropriate for the particular utility. IBM solutions are designed to drive optimum business outcomes and can span the spectrum from the utility's network and generation performance to customer operations and regulatory compliance (see Figure 1).

Each company's data management and analytics roadmap should be determined by the organization's overarching business imperatives and the prevailing regulatory and market model. In a highly regulated energy market, for example, the focus may be on grid reliability and customer satisfaction. In a deregulated and competitive market, retailers will be interested in customer acquisition and retention. Improving security of supply and the cost and quality of service may be a better strategy in such a case.

The primary business imperatives of utilities that are considering smart grids and smart meters include customer optimization, regulatory compliance, demand response optimization and operational efficiency. Each of these imperatives is driven by data.

Transforming customer operations

A successful customer engagement program requires a full 360-degree view of the customer. This view is established by integrating customer usage data from smart meters, customers' response to changes in pricing, and other operational and business systems with additional relevant data such as credit and geo-demographic data from external agencies.

Customer insight capabilities allow the utility to better understand the past and predict future utilization patterns. It also enables credit risk management and an understanding of what new offers and services may be most appropriate to specific customers. Furthermore, it can be used to determine optimal tariff strategies, detect energy theft or fraud and design demand response programs.

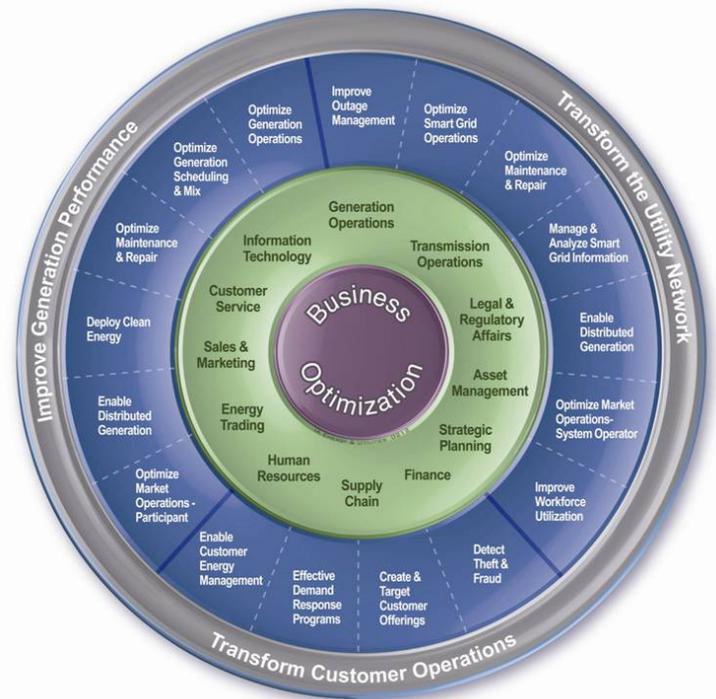


Figure 1: Business optimization map for the energy and utility industry.

Demand response offerings

Demand response programs are widely recognized as one of the essential tools that utility companies need to embrace. Key benefits include peak load shifting and potential elimination of costly spot market energy purchases or capital investment in additional generation capacity. Historically, consumption was calculated at an aggregated level and could not be easily apportioned across the customer base, whereas smart meter data provides granular consumption data for the whole customer base. This data will help determine expected load shedding when demand response events are declared.

Information security and privacy

Information privacy concerns and the high data volumes that the future “smart” world will produce preclude the movement of detailed usage and event data to various operational and business applications. Future operational efficiency may best be achieved by bringing various analytical processes to the data rather than piping the data into different systems.

While it is essential for a utility to be competent in multiple areas—transforming customer operations, optimizing demand response, data security and others—most organizations will prioritize a particular area or two. IBM is ready to help you identify the most valuable areas of focus for your company and create a roadmap from where you are to where you need to be in order to maximize the success of your smart meter or smart grid implementations.

For more information

To learn more about IBM solutions and expertise in the energy and utilities industry, please contact your IBM sales representative, or visit ibm.com/software/data/industry/energy.html

For additional information about IBM solutions for the energy and utility industry, download “The Information Agenda guide for the energy and utilities industry” white paper at: ibm.co/JjRm72



© Copyright IBM Corporation 2012

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
May 2012

IBM, the IBM logo, ibm.com, Cognos, Guardium, ILOG, Informix, InfoSphere, Optim, SPSS and WebSphere are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at ibm.com/legal/copytrade.shtml

Netezza is a trademark or registered trademark of IBM International Group B.V., an IBM Company.

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

The performance data discussed herein is presented as derived under specific operating conditions. Actual results may vary. THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation. Statements regarding IBM’s future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Actual available storage capacity may be reported for both uncompressed and compressed data and will vary and may be less than stated.



Please Recycle
