Demand Response and Grid Management are power utility initiatives driven by Federal and State stimulus programs to make the electric grid more reliable, secure and intelligent. The first step in making a grid “smart” is by replacing existing end user power utility meters with smart meters that support bi-directional communication between utility central control systems, the electric substations, the line, pad and underground transformers and to the power consumer. Utilities are scheduled to spend over $40B in the next five years on smart meters and smart grid infrastructure upgrades.

A key requirement for demand response and grid reliability is managing the acquisition and analysis of energy data from meters, transformers and substations. Smart meter data is used to reduce customer management and billing costs as well as supporting time of day energy pricing. Grid data is used to limit the impact of load balancing and grid failures including the monitoring of performance of various assets such as transformers, switches, substations equipment and monitoring grid zones for power utilization.

To meet this goal, GridPlex is designing a Cloud based platform to deliver to utilities, information on electrical load by different components of the grid, down to the transformer. The platform utilizes a virtual in memory parallelization and transaction analytics engine from μCirrus, coupled with a data aggregation and query solution, either from Query Objects or Netezza. Web centric utility metric dashboards are provided by GridPlex. In parallel, GridPlex is in discussions with VereCloud and a potential Carrier Provider (CSP) to use the Telco’s IT infrastructure for the Cloud service function for participating utilities.
The Emerging Smart Grid Data Problem

Smart Grids, by inference, are electrical distribution systems with an added information component. Data is captured at different time intervals, at multiple points on the grid and at each point of use. The purpose of data capture is to determine the grid’s quality of service, predict potential problem areas and provide information to support pricing models to different classes of end users.

The volume of data collected can be enormous depending on the number of observations per time unit. A recent report from the IQPC Smart Grid Forum\(^1\) estimates that 100 million Smart Meters would produce 1000 petabyte of data per year (a petabyte equals 1 thousand terabytes or 1 million gigabytes). While some of the information is transient, to support longitudinal models these volumes can easily overwhelm data storage requirements, even with cloud computing. A more serious problem is the ability in real time to analyze such data volumes using traditional relational database schemas.

Utilities understand that managing this information architecture may be best served through vendors that provide IT data management services. To this end, all major IT vendors are “re-inventing” their existing solutions to sell a Smart Grid information model. However, such offerings are usually based on large amounts of data storage, thousands of licenses for virtualized cluster servers and the hardware itself.

Current Solutions

Relational databases have evolved into many variants to address specific information problem sets. These variants include star schemas (see image right), bit indexing engines, normalized data, neural schemas and so forth. Almost all are rooted in the original “table” schema of rows (records, instances) and columns (attributes, metrics).

All these approaches are classically defined as “Newtonian” where data is stored in a physical structure and queried through efficient scanning of the physical data.

There is a limit to the efficiencies of such approaches because data storage growth is linear, requiring more hardware, complex partitioning and parallelism of the query engine, and technical support – all adding to system cost.

Alternative approaches abstract the physical data into a non-relational form that still permits standard SQL queries. Many of these solutions rely on very large memory allocations in servers, and complicated parallel query optimization, making the maintenance and operation expensive.

\(^1\) IQPC Forum
Advanced Solutions

The Query Object (QO) method is radically different but at the same time solves the limitations and problems stated earlier. Based on the concept of lossless fractal compression, QO normalizes the data into numerical discrete and continuous formats and creates a multi-dimensional schema that is SQL compliant. This mathematical construct is extremely efficient in managing sparse data (many nulls) and significantly reduces the physical size of the object which grows in size in log-log terms rather than with linear growth using other relational methods. Attribute relationships (referred to as dimensions in a star schema) are expressed in the form of polynomial equations with fractional coefficients.

This approach further reduces the object size. The easiest way to describe the difference between QO and relational methods is that QO stores all possible answers to any query as a single fractal point in a multi-dimensional space. Each dimension represents an attribute. The combination of any subset of dimensions represents an address where the metrics are stored. In other words, the metric stored represents the result of a relational query. If no metric exists for such a point (i.e. a null value), then it is not stored in the object. In this manner further compression is realized. Simply put, if no answer exists to a query, then no space is reserved in the object. In this way, fractal compression is lossless.

QO data structure is further optimal in that queries with different dimensions can be extracted (iterated) from higher and lower dimensions. What this means one can instantly calculate a fractal address in the structure from higher or lower dimensional addresses. A single “get” by address is all that is needed, compared with scanning rows and aggregating across dimensions that a relational model requires.

The key benefit of the fractal data model is that it is scale invariant in storage growth and delivers a relatively constant response time, regardless of the complexity of a SQL query. This alone sets it apart for solving Very Large Database Schemas (VLDB) with low infrastructure costs and IT management. Examples in the telecom market show a compression savings between 5000 and 10000 to 1 compared with Oracle’s solution. Independent studies confirmed that with average query tests, QO outperformed its competition by 50 percent faster query response times and with more complex queries, 95 percent faster.

QO applications have a long history in financial services, telecommunication network analysis, call detail record analysis, medical data, environmental sciences and business marketing solutions. Originally invented in 1987, QO has evolved through several generations of improvements, was ported to a number of hardware platforms and operating systems, and demonstrated its efficiency and cost savings over well established vendors such as Oracle, Microsoft, Informix, Sybase and IBM.

QO Solution for Smart Grid

The energy grid consists of power generation sources, transmission lines, sub stations for transformer voltage step down, distribution lines to transformer vaults and poles and finally, drop lines to the power users (consumers). A typical utility may have tens of thousands of such grid points where data is collected continuously at defined time intervals. Each transformer is considered a data acquisition point on energy flow through the grid. In addition, each smart meter generates additional data at different time intervals. The combined data is used for Smart Grid Demand Response, grid quality of service and maintaining reliable power to end consumers. While this describes a directional flow of data from user and from transformer to a centralized command function...
in the utility, it supports the ability to automatically manage power usage through smart devices that are connected to energy using appliances (pumps, heating and cooling, heavy machinery, and consumer appliances such as washers, dryers and interruptible devices.

As such, an additional information layer is supported to model how smart devices manage power consumption during peak demand periods. This is of value to utilities in that peak demand power requirements are reduced, providing significant savings in spot market energy purchasing and decreased need for capital requirements for creating new energy. Finally, the emergence of renewable energy production at point of use from solar, wind, geothermal and carbon neutral sources requires knowledge on how to manage excess power produced and fed back into the grid. Understanding how powering electric vehicles from the grid will impact the overall life of existing transformers is also a concern.

Grid data is primarily about electrical conditions: voltage, phase, current, temperature, and asset attributes such as equipment location and type, asset life, maintenance schedules and so forth. An important measure using SCADA phasors provides grid power state estimation. With Smart Meters there are additional data attributes, customer usage, pricing and smart device readings that drive time of day pricing and a snapshot of how these smart devices are controlled. For both data models, the frequency of measurement is the most important attribute, establishing a time view of both the grid reliability and power consumption by end users.

**QO Approach to the Smart Grid Data Problem**

QO is the most cost effective solution to the Smart Grid Data problem. Information gathered from grid monitoring and smart meter readings is perfectly suited to using QO for data analytics and business intelligence for energy utilities.

The QO grid data model consists of capturing grid point data at constant time intervals. For example, measuring attributes at each transformer at one second intervals creates a state map of multi-dimensional transformer data that is aggregated to a one minute object. Each one minute object is aggregated to a higher time order object (hours, days, weeks, months, years). QO replaces the one-minute state map every minute but aggregates the measurements into the higher time order persistent state map of the grid. Likewise, the Smart Meter points of measurement are captured at one to five minute intervals and follow the same aggregation method to create a historical object.

Transformers serve several end use power points, each of which is monitored using the Smart Meter. A typical transformer may serve five to seven residential homes or a single commercial user. Regardless, there is a physical plant mapping of transformers to smart meters that supports the joining of the QO grid object with the QO energy consumer object. This approach gives a complete, holistic view of energy flow and use over the grid, in real time. As an example, the table below shows the types of data collected using the above method. It is not exhaustive, simply as a reference. The second table links the transformer metrics with the Smart Meter metrics to create a single virtual QO from separate objects.

![Grid Data Analytics for Electric Utilities](image)

2 Synchrophasor Measurements Over SCADA

---

**Author:** Andre Szykier CEO/CTO GridPlex Networks LLC

**Copyright © 2010**
Smart Grid Data Volumes

Notwithstanding claims that the Smart Grid will annually generate more than a 1,000 Petabytes of data, simulations for a typical utility with 100,000 transformers and 1 million smart meters indicate that data volumes are not trivial. Using the grid metrics identified in the table above shows that for an average utility, transformer readings at 1 minute intervals can create 4 terabytes of data while smart meter readings at 5 minute intervals produce 35 terabytes. Increasing the frequency of measures will negatively impact the problem of grid data management and timely response.

Smart Devices

The emerging Home Energy Management (HEM) market is dependent on communication standards, smart device and smart appliance manufacturers. The ability to remotely control power consumption during defined time periods of specific appliances is a key goal of Smart Grid energy management. While the Smart Meter can report power usage and other metrics over time, the lack of interoperable standards limits the ability of smart devices in communicating with the meter.

One solution is to provide a sub-meter device that handles this problem of talking to smart devices and managing their operation as well as capturing data on use over time. This unit acts as an intelligent computer and gives consumers a measure of control over power use on their premise and behind the meter.

The value of such an approach is that the controller can handle communication protocols that can talk to smart meters as well as talk to pole mounted transformer monitoring units. This takes away the complexity of each smart device having to be communication compatible with various Smart meter manufacturer products. It also works at premises with standard meters, because the controller can communicate to the outside through the power line, wireless, through wireline, by using allocated emergency band radio spectrum or through IP addressable Internet modems or routers.

One advantage is an on premise intelligent controller can take the measured attributes and convert them using first order normalization, passing the result to the data acquisition function of the QO engine through various communication methods described earlier. Likewise, the transformer monitor can take the normalized data from the on premise controller and add this to its payload of transformer measurements that were normalized at a different time event.

This approach simplifies the data architecture transmitted to the central processor and reduces the time of loading into the one minute state map. By distributing first order normalization across thousands of grid points, a real time analysis is possible using the QO system.

---

Footnote:

3 Data payload estimate for a transformer reading in this example uses 80 bytes. For the smart meter and sub meter the data payload can be as high as 870 when taking into account readings from up to 20 smart devices such as thermostats, pumps, HVAC and large energy consuming appliances.
Conclusion

Smart Grid energy management will produce a significant and continuous volume of information that utilities need to analyze and store. Traditional solutions relying on large storage and clustered processors will not be able to keep up with the problem of data staging, let alone support fast response times without significant investment in IT systems. This cost will be difficult to pass on to power customers and will not be subsidized by federal and state governments. Since utilities will gravitate to outsourcing the IT function to vendors rather than manage the process internally, the information model will be expensive and difficult to manage, without any benefits to the consumer.

The Query Object model allows utilities to easily deliver Smart Grid energy data at a fraction of the costs using alternative IT methods. The design is proven in different industries, handles unlimited volumes and is capable of loading billions of records into any business intelligence and analytics product. As one of the only solutions that can manage and analyze the call detail records in telecom carriers, QO can provide the equivalent solution and value to energy utilities and their customers.