Supporting the Smart Grid 2.0 using In-Memory Database Systems

Stephen Dillon
Technology Strategy, Schneider Electric
North Andover, MA. U.S.A.
stephen.dillon@schneider-electric.com

Reviewer: John Brodeur
Technology Strategy, Schneider Electric
North Andover, MA. U.S.A.
john.brodeur@schneider-electric.com

***Abstract*— the Smart Grid introduces a paradigm shift in the requirements for processing and storing data. In-Memory Database Management Systems (IMDBS) are designed to support the Big Data requirements of the Smart Grid, real-time data analytics and decision-making systems.**

"In-memory enterprise data management provides the necessary equipment to excel in a future where businesses face ever-growing demands from customers, partners, and shareholders." [3]

In this paper, we will explore the need for the processing of data in real-time to support the Smart Grid, use cases, and the solutions provided by in-memory database systems.

I. Introduction

"The smart grid has come to describe a next-generation electrical power system that is typified by the increased use of communications and information technology in the generation, delivery and consumption of electrical energy." [1]

The Smart Grid is undergoing an evolution, as part of what is referred to as the Smart Grid 2.0, as smart meter providers and utilities attempt to live up to the expectations including real-time dynamic pricing, demand control response, and line loss analysis.

"Ultimately, what makes a grid smart is the technology to monitor in real time the current operational state of the generation, transmission, and distribution network and the ability to automatically respond to those conditions as quickly as possible. The primary goal is to provide greater visibility, transparency, and control over the generation, delivery and use of electric power." [26]

This potential cannot be realized without a new breed of data management tools and systems that can handle Big Data and do so in real-time.

There are five main triggers of the Smart Grid. [2]

- With renewable energy, consumers can produce their own energy and access a 'greener' energy mix
- Flexible distribution enables a more responsive and stable electrical network
- Active energy efficiency and energy management make energy visible and allow individuals to act on their consumption
- Electric vehicles are revolutionizing the perception of mobility and, at the same time, access to energy, its use, and storage
- Real-time grid management enables anticipation of consumption and adaptation of the offer

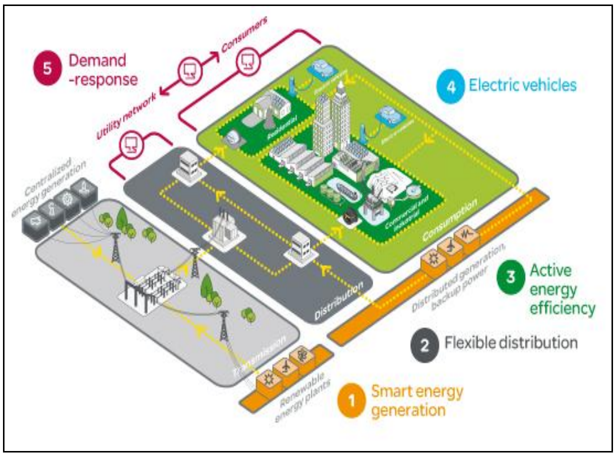

Figure 1 – Five Main Triggers of the Smart Grid

II. What is Real-time?

Real-time - adj. "Of or relating to computer systems that update information at the same rate as they receive data, enabling them to direct or control a process…"

We can define "Real-time analytics" as the process of performing aggregations, identifying patterns, and crunching numbers *at the time of data ingestion*. We can also apply the term "real-time" towards the concept of "On-demand analytics". On-demand analytics, instead of all the processing of the data taking place at the time of data ingestion, can instead take place against data that has already been collected whether it was an hour, a day or a year ago. This is considered real-time within the context of the raw data of a database being processed without the need to pre-aggregate data for analysis. It simply becomes a matter of perspective as to when the data is ingested by the query engine of the IMDBS. This paper will use the term "real-time analytics" to refer to both of these use cases.

There is no question that real-time data processing and analytics is of great importance to the industry, and one mustn't look far for examples of applications that provide real-time metering and monitoring.

For example, Schneider Electric's "Smart Cities" endeavor seeks to implement a "bottom -up, systems-centric approach with a top-down, **data-centric** one" in an effort to enable an intelligent infrastructure. "Smart cities start with **smart systems** … Electric grids, gas distribution systems, water distribution systems, public and private transportation systems, commercial buildings, hospitals, homes. It is the improvement and integration of these critical city systems — done in a step-by-step manner—that become the cornerstones to making a smart city a reality." [4]

"Once the data is collected and available, **real-time systems** can use the data to automate management of city infrastructure, resulting in significant performance and cost advantages." [4]

III. Use Cases

This "new era of intelligent energy management" [2] has given rise to hundreds of user documented use cases for the Smart Grid that includes several high-level categories such as. [11]

- Customer Service
- Distribution Operations
- Market Operations
- Distributed Energy Resources
- Federated System Management Functions
- Transmission Operations

We will briefly reference the following examples, to provide scope for some of the problems we can solve with an In-Memory Database Management System, including VEE, Grid Stability, Real-time Pricing & Billing, and Alarm Management.

VEE

A precursor to all other use cases, and arguably the most important, is the ability to acquire quality data from smart meters. Validation, Editing, and Estimation engines, commonly referred to as VEE, are a common method used to verify raw meter data and provide gap analysis, spike detection, and fill in estimated data where required. Most VEE engines have historically attempted to perform these activities in batch form, typically during off hours, in order to handle the performance requirements and offset stress placed on the underlying database management systems. The need to gain insight into this data more quickly, than batch processing allows, gave rise to the popularity of Complex Event Processing techniques to act upon data at the time of ingestion.

Applications across the smart grid that perform real-time data analytics demand a real-time capability to perform VEE activities. In-memory technologies are ideally suited to fulfill this requirement.

Grid Stability

Many Energy Service Providers and Market Operators provide customer-side Demand Response and Load Control programs. These programs ensure the stability of the grid and the stable operation during times of peak demand or system emergencies arising from generator outages, transmission and/or distribution constraints. AMI – Advanced Metering Infrastructure supports these types of programs. "Advanced metering systems are comprised of state-of-the-art electronic/digital hardware and software, which combine interval data measurement with continuously available remote communications." [27]

With some programs, the customer…reduces the required load upon instruction from the Energy Service Provider or Market Operator…or a Curtailment Service Provider remotely reduces the load. Some of these programs are conducted on a voluntary basis, where the customer can opt to maintain the level or load, or mandatory, where the customer either will be dropped off the system or will incur significant financial penalties for noncompliance. [13]

The customer's ability to monitor and react to events relies heavily upon applications and services that allow for the up to the minute (or better) processing of data.

Other examples of Grid Stability include precise forecasting similar to Siemens "self-learning software" based on neural networks. This system can forecast the electrical output of renewable energy sources over a 72-hour period with more than 90 percent accuracy. The data helps grid operators calculate power demand in their networks and fairly exactly order the amount of additional electricity required in advance. [5] It is now possible for similar forecasting efforts to be performed by real-time analytics or for such a system to benefit from the real-time processing capabilities of in-memory technology.

Real-Time Pricing & Billing

We understand as customers and employees of the Energy industry that electricity pricing is variable throughout the time of day. "Nearly all retail consumers, however, are charged some average price that does not reflect the wholesale price at the time of consumption. "[[23] Real-time pricing offers the consumer the opportunity for cost savings by reacting to wholesale market pricing.

RTP also enables Demand Response programs by providing the consumer and utilities transparency in actual load, better alignment of rates to the cost of services. [24]

Benefits of RTP include:

- A choice in energy providers per RTP
- Unlock the potential for demand response (DR) to save capital and operational costs and empower consumers
- Better match of retail rates to cost of service creates transparency, unleashes DR (and distributed generation & storage)
- FERC suggests ~20% DR capacity is achievable
- Historically, DR used to manage system peak load

Historically, meters have only reported consumption in 15 minute intervals, submitting the data daily, and third-party systems have been utilized to process billing and consumption as batch jobs. Smart meters instead offer the ability to monitor consumption and billing immediately. These meters can produce a continuous stream of data and as a result the opportunity exists for "…Energy providers to offer new flexibility to their consumers. Rather than getting an invoice once a year with real energy consumption, the consumer is able to track its energy consumption at least on daily basis." [6]

In-memory technology can be utilized to provide pricing & billing details in real-time, i.e. each time the consumer logs in, meter readings are aggregated on-the-fly via sub second queries without the need for pre-aggregated totals. "…only raw meter data and no pre-aggregation totals" need to be stored so "There is no need to store either daily or monthly consumption aggregates to improve performance. This redundancy-free way of storing only raw data minimizes storage demands and eliminates the need for updating totals by the application explicitly." [6]

The need to process data quickly will become paramount as the number of installed smart meters increases, customers will begin to demand access to wholesale pricing and utilities will offer RTP as a common practice.

Alarm Management

Remote Energy Monitoring systems are commonly used to receive remote signals and measurement thresholds which are sent from meters as events to be processed as alarms to be sent to an administrator or operator. [12]

These events are typically cross-referenced against customer defined business rules and provide a decision-making capability. For example, if the battery power of a system falls below 75% contact the operator via email, phone, or perhaps a product recommendation is generated and sent to the customer as part of a marketing campaign.

The advent of the smart meter and support for other use cases such as demand response and RTP will increase the importance of alarming functionality. The necessity for real-time processing of these events, such as the ability to automatically shut down a remote system, especially for critical infrastructure systems will have to be supported by underlying technologies capable of immediate responsiveness utilizing high performance event handling.

IV. Big Data

The Smart Grid generates Big Data.

The term Big Data can be best described as "…data that exceeds the processing capacity of conventional database systems. The data is too big, moves too fast, or doesn't fit the structures of your database architectures. To gain value from this data, you must choose an alternative way to process it." [22]

Typically when discussing Big Data we will hear of the "Three Vs" Volume, Velocity, and Variety. In brief, the Volume required to be categorized as Big Data is approximately 3 TB or more of data circa 2013. Velocity on the other hand is a vague metric and can be summarized as any amount of work that must be done to ingest, store, transform, retrieve, or report data. Finally, Variety can be viewed as the multiple formats data can be received or stored as. An example would be various meters all communicating a kWh in various storage formats such as Xml or Comma Separated Value CSV.

It is important to note that one does not have to possess all three of the "Vs" in order to have Big Data.

V. Data Storage Trends

There are two data storage & processing trends in Big Data that allow us to support Smart Grid use cases; NoSQL and NewSQL.

Each of these categories seeks to support the handling of vast amounts of data Volume through scale-out architectures that leverage the compute power of a cluster. They each offer different solutions however and as such meet the needs of different use cases within the Smart Grid. *Whereas this paper focuses on the application of In-Memory Database technology and real-time analytics, we will only focus briefly on NoSQL and instead focus on NewSQL in detail.*

NoSQL

*NoSQL supports the Volume and Variety of data. It **does not support** the Velocity and real-time processing or analytical needs of the Smart Grid.*

NoSQL DBMS provide alternative data storage capabilities that Relational Database Management Systems do not. These technologies are not meant to replace RDBMS but instead compliment them with alternative storage solutions. Use cases include the storage of unstructured discussion threads, multi-structured data such as application logs, or meter data stored in industry standard information models such as the 61850 format. In these scenarios, a non-relational data store would be an ideal candidate as opposed to attempting to shoehorn data into a relational model.

In 2013 there are an estimated 75 NoSQL technologies offered across four topologies including Document-stores, Key-Value stores, Graph databases, and Column-stores. All of these have similar characteristics including non-relational data storage, no transactional support, and originally they were designed for off-line or "batch" analytics of data.

NoSQL technologies support the processing of large volumes of data per a scaled-out, shared nothing architecture using an Elastic Load Balanced approach and the MapReduce programming model [20].

Real-time processing has not gone unnoticed by the NoSQL community. There is a growing trend to make Hadoop operate in near real-time using technologies such as Storm. Software provider MapR is one such company now delivering "near real-time" results against Hadoop. These offers definitely provide potential solutions for different use cases. By comparison, an IMDBS will always beat a disk-bound system running the same query against the same data simply due to the very nature of data retrieval via memory versus disk.

NewSQL

*NewSQL **does support** the Volume, Velocity, and real-time analytical needs of the Smart Grid.*

NewSQL is the evolution of the traditional RDBMS. The scope of this paper is not to cover all types of NewSQL offers but instead only focus on IMDBS. Briefly however these technologies provide the capability to:

1. Horizontally scale out databases by sharding data across multiple servers
2. Virtualize these databases as one to our applications
3. In the case of IMDBS, provide in-memory storage of an entire database's transactions for the fast and efficient analysis of data a.k.a. Real-Time Analytics.

VI. Traditional RDBMS

Before discussing in-memory technologies further, it is important to understand its predecessor and why a new solution is needed to support the Smart Grid.

Traditional Relational Database Management Systems, such as Oracle and Microsoft SQL Server, are disk-based storage mediums and as such come with an inherent overhead for processing, storing, and retrieving data. In a breakthrough paper entitled "OLTP Through the Looking Glass and What We Found There" it was demonstrated that approximately 93% of the work a DBMS performs is overhead and only approximately 7% is used to perform useful work when retrieving a record. [10] See Figure 2.

If traditional RDBMS can be viewed as a hammer, NoSQL technologies would be a screw gun, and NewSQL would be the pneumatic hammer.

VII. In-memory Database Management Systems

RAM is the new disk!

An In-Memory Database Management System (IMDBS) stores all of a database's transactional data in main memory. Past efforts in this field only provided Read operation optimization. Other technologies exist that have often been mistaken as in-memory databases, which only utilized cache to store a working set of data rather than an entire database in main memory. Thus I will distinguish between these other technologies by stating that the term IMDBS is synonymous with the term "Modern IMDBS" within the context of this paper.

"In-memory technology is set to revolutionize enterprise applications both in terms of functionality and cost due to a vastly improved performance." [3]

Modern In-Memory database management systems are instead designed to not be disk-bound, to be both Read &

Write optimized, and eliminate the overheads caused by locking contention and latching typically found on a disk-based system.

In-memory technologies, such as **VoltDB** (voltdb.com), provide for Read and Write performance optimization, massive scalability, durability and highly-availability. The traditional big three vendors of RDBMS (Oracle, Microsoft, and MySQL) have all noticed this growing trend and have attempted to add on their own in-memory features or plan to within an upcoming release. For example, Microsoft has announced that SQL Server '14 code named "Hekaton" will include the ability to store some data from "hot tables" in memory while cold or inactive data will remain disk-bound. It is unclear at the time of this paper how "Hekaton" will address the horizontal-scalability benefits modern IMDBS provide but it is clear the days of simply scaling systems up have been surpassed by modern technologies and architectures available in IMDBS.

"Given the dramatic increase in RAM sizes over the past several decades, there is every reason to believe that many OLTP systems already fit or will soon fit into main memory, especially the aggregate main memory of a large cluster." [10] "Modern servers can provide multiple terabytes of main memory and allow the complete transactional data to be kept in main memory." [3] Thus it is commonly said that *RAM is the new disk.*

IMDBS are Persistent!

IMDBS are designed to meet the modern needs of the Smart Grid and Enterprise systems. There is a misconception by some that data will not be persisted in the event a server is powered off. This is not true. It is important to realize that an IMDB is not a cache store. These modern systems are designed to persist to logs and replicated servers. This makes them highly durable and recoverable in the event of a disaster.

ACID Compliancy

NewSQL technologies, and specifically IMDBS, support ACID compliancy.

The ACID properties of traditional RDBMS (Atomicity, Consistency, Isolation, and Durability) have long been a main staple of enterprise systems. This set of properties ensures that transactions are able to be processed reliably. Any modern in-memory database should meet ACID compliancy.

NoSQL technologies on the other hand do not support ACID and are non-transactional. These systems are referred to as "eventually consistent". This means that not all data is guaranteed to be the same on all copies of the database at a given point in time. In some cases it may take 5 minutes for data to synchronize and in other cases data may be lost altogether. While there are certainly use cases for NoSQL databases, some that may belong to the Smart Grid, we will claim for the purposes of this paper and analytics within the Smart Grid that ACID compliancy is a very important requirement within our relational OLTP data stores.

SQL Compliancy

SQL (Structured Query Language) has been the de facto programming language of database engineers since the 1970s and how most people think in database terms. All modern IMDBS should support some level of SQL (Structured Query Language) compliancy. Each vendor may provide a varying degree of SQL commands but unlike NoSQL technologies, a standard mechanism for data operations exists and one that is very familiar to engineers.

In fact many NoSQL technologies are now implementing, or plan to, SQL-like syntax atop their technologies.

Performance

A key tenet of in-memory technology is the ability to process, store, and retrieve data more efficiently than a traditional RDBMS. IMDBS meet this capability first by removing the overhead of a traditional DBMS. Again; only 7% of the work conducted on a traditional OLTP database was used to retrieve a record. [10] That means that 93% of the work done by our traditional systems is spent handling overhead processes instead of performing the work we desire as demonstrated in Figure 2.

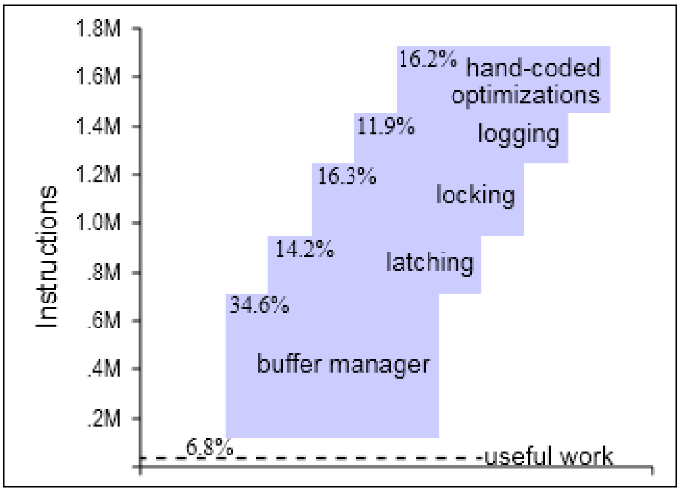

Figure 2

Second, data can be scaled horizontally across nodes in what are referred to as "shards". A shard, as depicted in Figure 3, is similar to a traditional "partition" except that a shard includes an entire copy of the database schema which is populated by a subset of the data. [21] This distribution of data allows database queries to be written that can efficiently query only the necessary data instead of all data. Additionally, the compute power of each node (server) can be harnessed without affecting the other a.k.a. a Shared-nothing architecture.

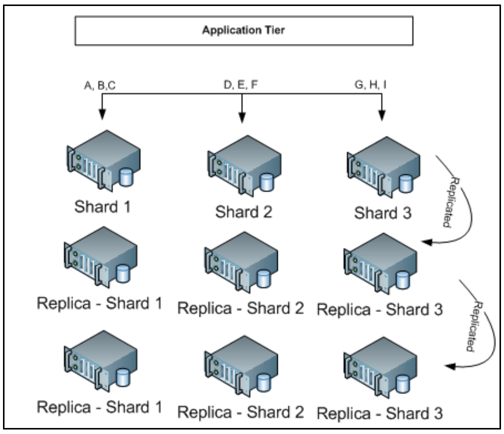

Figure 3

Third, IMDBS support the analytical capabilities and performance of an OLAP system via an OLTP schema. These schemas can include either row storage or as columnar storage. There are multiple in-memory offers on the market that appeal to such use cases as well as a hybrid approach and offer both column- storage as well as row-storage. Pending the volume of data and the type of analytics you wish to perform, each type of storage has its place.

VIII. Real-time Analytics

In-memory databases = real-time analytics.

By virtue of being able to process an entire database's transactions in main memory, IMDBS are able to perform real-time analytics, both as data is ingested as well on-demand against an OLTP schema. Traditional OLTP based RDBMS could not perform these analytical operations because of the significant overhead of the disk-based systems. This meant an OLAP schema had to be implemented and the management overhead of designing data cubes, and tables to store the facts and dimensions to perform the pre-aggregation of data for pre-selected queries.

Traditional OLAP systems are not able to support the dynamic processing and performance requirements of Big Data or it is too costly to do so. This lack of flexibility led to the advent of analytical systems such as NoSQL's Hadoop which is designed to perform off-line batch processing and analytics.

IX. Leveraging the Cloud

Time, Effort, Cost

The emergence of the Cloud, as the modern data center, has given rise to newer and more powerful system architectures that allow systems to leverage multiple compute nodes operating in parallel. These architectures support the ubiquitous nature of data in the Smart Grid. No longer should data merely be viewed as belonging to only one system hidden from other systems in an enterprise.

Modern IMDBS leverage the capabilities of the Cloud and base their system architectures on the power of horizontal-scalability and a shared-nothing architecture. This is not to say that these technologies cannot be deployed on physical hardware, or virtual machines in a private data center, but when considering Time, Effort, and Costs the choice to deploy these systems on a Cloud provider's IaaS offer is attractive; especially at the point of initial deployment. The Cloud can become costly if not managed correctly in the long-term.

Traditional RDBMS by comparison could not achieve the same architectures, scalability, or performance of the Cloud. Those that could come remotely close could not do so within the same **Time, Effort, or Costs.** Thus it is my contention that the Cloud has become the new data center; especially for Big Data solutions.

X. Decision-making & the OODA Loop

Observe, Orient, Decide, Act

The Smart Grid is non-linear. One can think of the decision-making and processing requirements of the Smart Grid in similar terms to a driver on the highway; constantly processing information about the road conditions, weather, vehicle warning lights, and other vehicles. The Smart grid also requires the processing of continuous streams of data, many of which can be processed and analyzed immediately, to provide continuous insight that will allow a company to action.

"There are two main aspects of a database systems, namely database management (data storage, transaction processing, and querying), and data mining (analysis of data to gain certain knowledge or facilitate certain decision making)." [7]

The concepts of Real-time analytics and Decision-making systems are not proprietary to any one industry. It is first a concept of *time-scale* and *customer maneuverability* that allows one to become oriented in a competitive situation. Second; this is a time-based strategy supported by analytical systems that are capable of processing large volumes of data. I will further summarize real-time analytics as the ability to quickly perform decision-making against current environmental conditions and events.

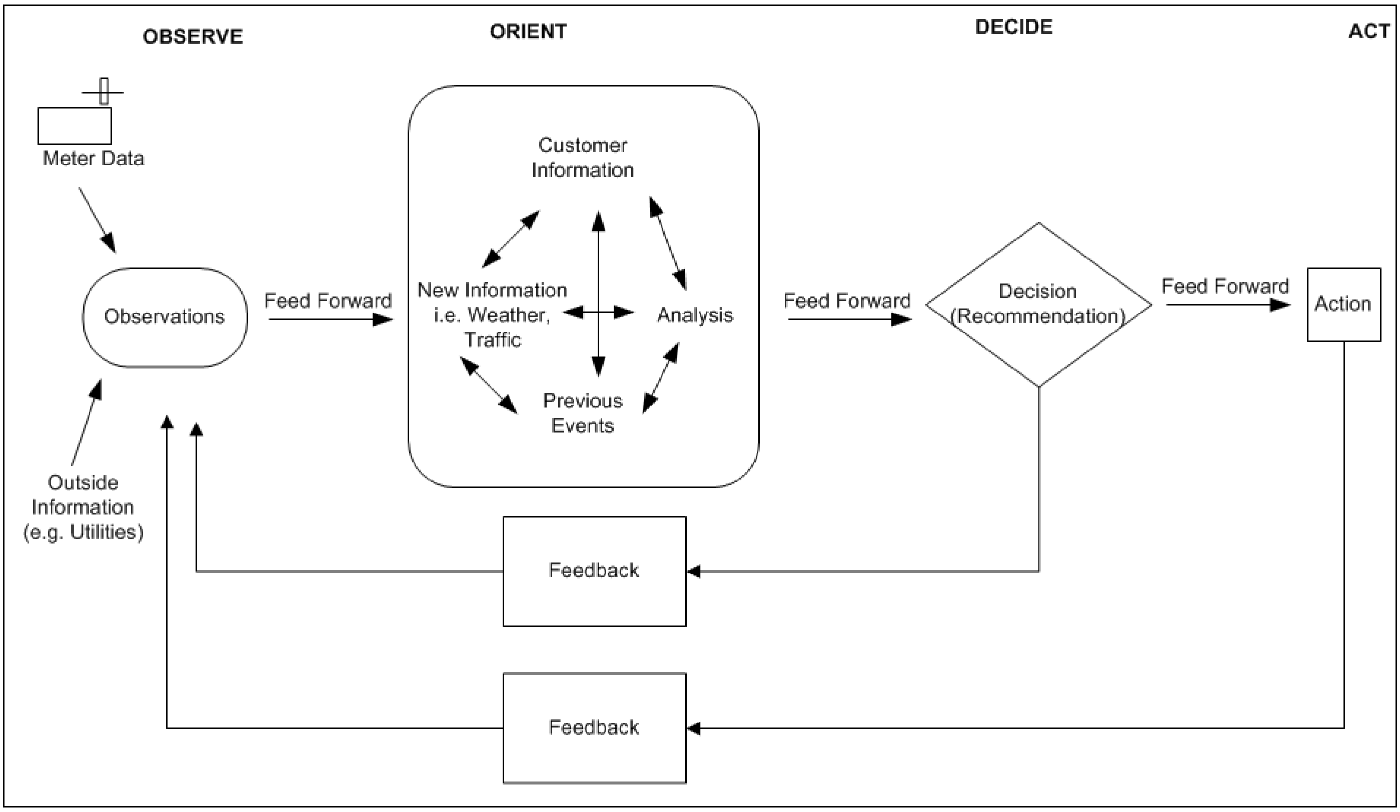

Figure 4

This analytics is typically incorporated with business rules and decision making algorithms that support the needs of a particular industry and enterprise. Unlike other fields of study we have been familiar with in the past decade, such as Complex Event Processing and Batch Data Analytics, IMDBS provide real-time analytics with both the power of complex event processing (CEP) as well as very large volumes of persisted data (Batch Analytics).

The OODA Loop is a complex concept, created by Colonel John Boyd-USAF, that for the purposes of this paper we can paraphrase as the ability to observe and react to unfolding events rapidly and gain an advantage. [14]

The "O-O-D-A Loop" was originally designed for military aerial combat but has been commonly implemented and studied within the private business sector for decades in line with Sun Tzu [18] and Karl Von Clausewitz [19]. This strategic, time-scaled, non-linear, decision-making concept is one that can be applied to the Smart Grid and demonstrates the need for real-time analytical capabilities of IMDBS.

If we look closely enough at the Smart Grid, what the industry really wants to do is answer the question *how do we identify the immediate situation (our situational awareness of the power infrastructure), process the parameters of an event, decide on what actions to take and execute the actions?* This is all with the ultimate goal of energy efficiency and sustainability in mind.

Figure 4 depicts an adaptation of the OODA Loop diagram applied to a meter in the Smart Grid. We can see how the data from the power infrastructure is fed forward through the cycle. "The loop is actually a set of interacting loops that are to be kept in continuous operation". [14]

If we walk through the diagram and apply it to the Smart Grid, the Observation phase is the data collection or ingestion from a smart meter. The application is able to make an observation i.e. the data. The Orientation phase is of the most importance to the business whereas this is where the business value is delivered via knowledge of your customer, a system, the market, energy practices, and predictive analysis. "The single most important part of the cycle is the orientation phase." [17] The Decision phase is where the business rules are applied. If event A happens, B should happen. Again this is not a linear process. Upon a decision being made this is fed back into the loop and becomes a part of the Observation phase. Finally, an action is taken such as a sub station operator must shed load and power down systems. The Action itself is also fed back into the loop and becomes part of the Observation phase.

We can envision that in order maintain these loops of situational awareness and decision making, we must process and analyze data in real-time. From an Energy and Smart Grid perspective, the batch processing of data, *even at the two* *minute mark*, is not conducive to the types of analysis the industry is demanding today.

The concept of the OODA Loop supports the need to perform real-time analysis of data within the Smart Grid.

XI. Interoperability

Interoperability is "the ability of a system or a product to work with other systems or products without special effort from the user is a key issue in manufacturing and industrial enterprise generally." [8]

When we discuss the interoperability of a DBMS, we need to consider multiple factors. First, how easy is it to store data without a lot of overhead required to get data in and out; such as special ETL processes? Second, how easily can we integrate the DBMS with other systems? Is there an API available? Third, how will the data be physically stored; via row or by column? Fourth, is the system scalable and durable? Then also need to consider interoperability at a standards level; for example the Smart Grid Interoperability Panel (SGIP) is an organization that seeks to "align technologies that drive opportunities for creating a more efficient and reliable energy grid". [15]

The interoperability of NewSQL offers, such as IMDBS, will vary by vendor. One such vendor, VoltDB, has taken interoperability into special consideration by providing the following capabilities and features:

- SQL Compliancy
- Standard APIs (Java, .Net)
- Relational data storage
- Highly efficient retrieval of data
- Scale-out architecture
- High-availability\durability
- Special storage of unstructured data as JSON objects

These features lend this technology to integrate with systems more easily than others.

XII. VoltDB

VoltDB is a best in breed, mature, in-memory database management system. It solves both the volume and the critical velocity problem of the Smart Grid 2.0.

"It is a fully durable in-memory relational database. It combines high-velocity data ingestion, massive scalability, and real-time analytics and decisioning to enable organizations to unleash a new generation of applications that act on data at its point of maximum value." [9]

As depicted in Figure 5, VoltDB outperforms other OLTP traditional relational database management systems by a factor of 45x on a single server. [9]

	Nodes	VoltDB	DBMSx	VoltDB Advantage
"TPC-C-like" workload (VoltDB lab)	1	53,000 TPS	1,555 TPS	45x better throughput
	12	560,000 TPS	N/A	Near-linear (.9) scaling

Figure 5

When a scaled out architecture is implemented, it can reach magnitudes of scalability and performance unmet by traditional relational database technology. Even with a significant investment in hardware, upwards of millions of dollars, traditional systems cannot meet the throughput and transactions per second of an IMDBS as shown in Figure 6.

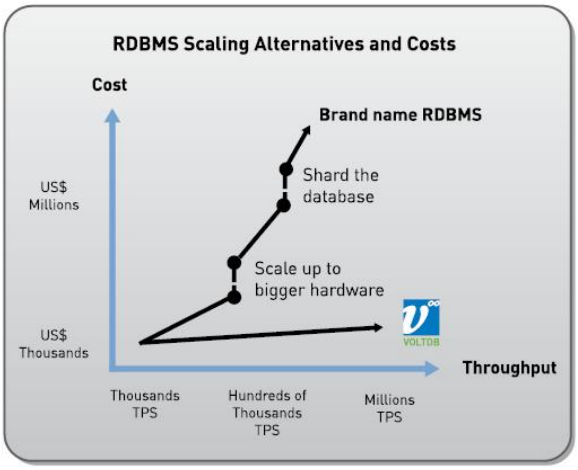

Figure 6

"In-memory architectures are ideal for running data-centric enterprise applications in the cloud. In fact, replication across two nodes where data is kept in main memory both offers higher throughput and is more predictable than writing to disk on a single node." [3]

A recent proof of concept, conducted by Schneider Electric's Technology Strategy team using VoltDB, demonstrated that a Billing analytics query previously executing in upwards of 20+ minutes on SQL Server Enterprise 2008 R2 was able to execute in 452 ms on a 3 node cluster of VoltDB *with only moderate tuning efforts*. As a testament to this technology, another leading vendor in the IMDBS industry claimed per a discussion of these results and the use case that they would not be able to beat this performance.

This technology offers further interoperability, and in some cases advantages over its competitors, by its ability to be deployed to the Cloud, physical servers, *as well as embedded instances*.

XIII. Conclusion

It is clear that the Energy Industry and Smart Grid customers are demanding faster, modern, more efficient data storage and real-time processing and analytical systems.

"We are convinced that in-memory technology is a catalyst for innovation, and the enabler for a level of information quality that has not been possible until now." [3]

Those not indoctrinated in the Smart Grid or Big Data can easily become overwhelmed by the plethora of use cases and technologies. However, the Energy industry is now approaching a period during which competitors will begin to outpace those who do not seize upon these opportunities.

It is important for us to realize that IMDBS are only one piece, albeit a very important one, in Big Data. Any Enterprise level Big Data initiative, including Smart Grid efforts, should utilize multiple technologies in the correct areas to support various data processing and storage requirements. As depicted in Figure 7, these technologies will span the "Database Universe" [28] to provide real-time, interactive capabilities as well as traditional and downstream analytical needs at the batch processing level.

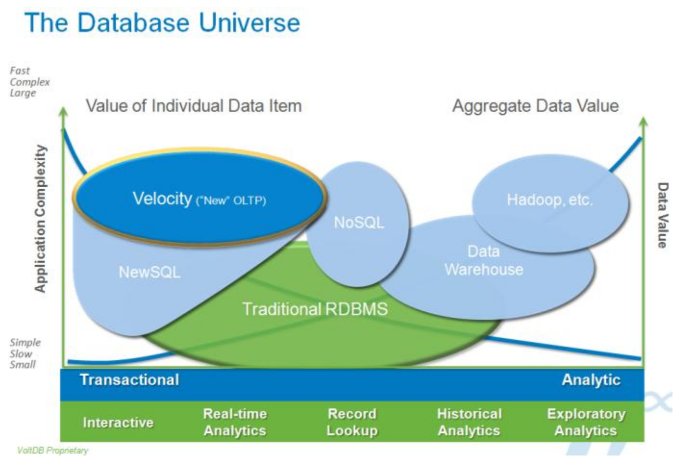

Figure 7

In closing I would leave the reader with the following message. "Big Data will deliver transformational benefits to enterprises …and by 2015 will enable enterprises adopting this technology to outperform competitors by 20% in every available financial metric." [16] It is the challenge of everyone working in the Energy industry and participating in the Smart Grid to understand the opportunities available for sustainability and efficiency.

XIV. Author Bio

Based in Massachusetts, Stephen Dillon is a Data Architect for Schneider Electric's Technology Strategy team. He has over 15 years of experience working with database technologies including Microsoft SQL Server architecture, database design and performance tuning. He has additionally been working with Big Data technologies within NoSQL, NewSQL, and Cloud technologies such as SQL Azure, MongoDB, Amazon EC2, and Windows Azure since 2010.

Stephen holds a Masters in Computer Information Systems from Boston University.

References

[1.] IEEE, "IEEE & Smart Grid, What is the Smart Grid?", available online via http://Smart Grid.ieee.org/ieee-smart-grid

[2.] Schneider Electric, "Smart Grid, Schneider Electric vision", Available online via http://www.schneider-electric.com/sites/corporate/en/group/energy-challenge/smart-grid.page

[3.] Plattner Hasso, Zeier Alexander, "In-Memory Data Management Technology and Applications", 2nd Edition, ISBN 978-3-642-29574-1, PP. xxxii

[4.] Aoun, Charbel, "The Smart City Cornerstone: Urban Efficiency", pps.2-10, Available online via http://www.schneider-electric.com/documents/support/white-papers/smart-cities/998-1185469_smart-city-cornerstone-urban-efficiency.pdf

[5.] Siemens, "Grid Stability Thanks to Precise Forecasting", Available online via http://www.siemens.com/innovation/en/news/2012/e_inno_1217_1.htm

[6.] Schapranow, M., Kuhne, R., Zeier, A., Plattner, H., "Enabling Real-Time Charging for Smart Grid Infrastructures using In-Memory Databases", 2010, 1st IEEE Workshop on Smart Grid Networking Infrastructure

[7.] Aung, Zeyar, "Database Systems for the Smart Grid", pp.1, available online via http://www.aungz.com/PDF/Aung_Chapter.pdf

[8.] Konstantas, D.; Bourrières, J.-P.; Léonard, M.; Boudjlida, N. (Eds.) , "Interoperability of Enterprise Software and Applications:, available online via http://www.springer.com/computer/swe/book/978-1-84628-151-8

[9.] VoltDB, "Technical Overview" white paper available online via http://voltdb.com/downloads/datasheets_collateral/technical_overview.pdf

[10.] Harizopoulos S., Abadi D, Madden S., Stonebraker M, "OLTP Through the Looking Glass and What We Found There", available online via http://hstore.cs.brown.edu/papers/hstore-lookingglass.pdf.

[11.] EPRI, Smart grid Use Case Repository, available online via http://www.Smart Grid.epri.com/repository/repository.aspx

[12.] EPRI-2, "Smart grid Use Case Repository, UC8 Load Shedding (by Order)", Available online via http://smartgrid.epri.com/UseCases/LoadShedding-ByOrder.pdf

[13.] EPRI-4,"Smart Grid Use Case Repository, Demand Response – Utility Commanded Load Control", Available online via http://smartgrid.epri.com/UseCases/DemandResponse-UtilityCommandedLoadControl.pdf

[14.] Boyd, J. (1987). A discourse on winning and losing. Maxwell Air Force Base, AL: Air University Library Document No. M-U 43947 (Briefing slides)

[15.] SGIP, "Smart Grid Interoperability Panel", Available online via http://sgip.org/about_us/#SGIP-alignment-of-technologies

[16.] Gartner, “Hype Cycle for Cloud Computing Shows Enterprises Finding Value in Big Data, Virtualization”, August 2012, Available online via http://www.forbes.com/sites/louiscolumbus/2012/08/04/hype-cycle-for-cloud-computing-shows-enterprises-finding-value-in-big-data-virtualization/
[17.] Coram, R., “Boyd, the Fighter Pilot Who Changed the Art of War”, pp.235, 2004, Back Bay Books
[18.] Sun Tzu, The Art of War
[19.] Von Clausewitz, Karl, On War
[20.] Rhoton, J., Jaukioja, R., “Cloud Computing Architected”, p.69, 2013, British Library Cataloging-In-Publication-Data
[21.] Wilder, B., “Cloud Architecture Patterns”, O’Reilley, 2012
[22.] Dumbhill, Ed, “What is Big Data?”, January 2012, Available online via http://strata.oreilly.com/2012/01/what-is-big-data.html
[23.] Allcot, H., “Rethinking Real-time Electricity Pricing”, 2009, Available online via http://web.mit.edu/ceepr/www/publications/workingpapers/2009-015.pdf .
[24.] Pratt, R., Real-Time Pricing and Demand Response”, pp. 4, July 2011, Available online via http://www.smartgrid.gov/sites/default/files/doc/files/RealTime_Pricing_Demand_Response_201106.pdf
[25.] “Real-time”, The Free Dictionary, Available online via http://www.thefreedictionary.com/real-time
[26.] Flood, Joshua, Lucero, Sam, “Overview: The Evolution of Smart Grid 1.0 to 2.0”, pp.2, 2011, Available online via http://youratt.com/smartgrid/pdf/ABI_ATT_Smart_Grid_Evolution_102611Final.pdf
[27.] EPRI, “Advanced Metering Infrastructure (AMI)”, p.1, Available online via http://www.ferc.gov/eventcalendar/Files/20070423091846-EPRI%20-%20Advanced%20Metering.pdf
[28.] Stonebraker, M., Jarr, S., “Navigating the Database Universe”, Slide 23, Available online via http://www.slideshare.net/lisapaglia/navigating-the-database-universe