



it's about time

Outside the Utility Box

Solving the utility sector's Big Data dilemmas with cross-industry innovation

Author:

Leslie Goldsmith



Introduction

Organizations across all sectors of today's business landscape experience a glut of data from countless traditional streams, such as from customer sales and transactional information as well as more novel ones like machine-to-machine communications, sensors and meters, social networks and mobile devices. This has caused an evolution in the way we store and manage data, as well as in its application and analysis—all generally for the better. But to be useful, that unprecedented amount of data must be turned into actionable information.

The abundance and variety of data, dubbed Big Data, has forced businesses to transition away from the traditional methods of management and storage. In order to gain the business insights needed from these traditional and new types of data, radical approaches to data management and business intelligence must be developed to move beyond the data warehouses of the past two decades. This need for a new approach is sharply evident in sectors like utilities, where collecting enormous amounts of new data from sources, such as smart meters and telemetry devices on the grid, has become the norm, and where timely analysis of that data can yield significant business- transforming insights.

Far from being seen solely as a challenge, businesses are recognizing that great opportunity comes from Big Data. The true challenge, of course, is in how to deal with it successfully and transform it into value. There is no single one-size-fits-all answer to this challenge for all industry sectors.

This paper looks at Big Data's impacts, especially in the utility business—its opportunities and challenges, and how organizations can look to solutions used by other industries to leverage Big Data to make better, more informed and more timely business decisions.

The Big Deal about Big Data

Interest in Big Data is tremendous, and that's because the demand for insight into corporate information has never been higher.

In fact, the value of Big Data-related hardware, software and professional services is projected to continue to grow 17 percent annually to reach \$76 billion by 2020¹. Some industries' Big Data related spending is particularly notable, like financial services firms' whose investments in Big Data are expected to grow 22% a year until 2020².

Why? Not many years ago, only the largest of multinationals considered the value from Big Data (then dubbed VLDB or very large databases), or even had access to it. But new streams of data, much of it from the web, mean that now even small companies witness Big Data opportunities and challenges. Two forces are helping drive the need for systems to better manage and leverage Big Data: The Internet of Things and increased competition.

Put simply, nearly every device or appliance, home, car and worn product is increasingly becoming interconnected—each with identifiable digital pulses, sending data that we can use to help organizations become smarter. This is propelling industries toward greater automation in the same way it furthers the electric smart grid by enabling us to better understand the behaviors of suppliers, consumers and events.

1 "The Big Data Market 2014–2020: Opportunities, Challenges, Strategies, Industry Verticals and Forecasts", Signals and Systems Telecom (June 2014)

2 Bain Insights, "Insights Infographic: The Who, Why and How of Big Data", Forbes (29 October 2013)



At the same time, competition is high, business cycles are short, and business transforming opportunities come and go quickly. Today's competitive forces leave little or no room to waste time in the constant drive to meet more unforgiving customer expectations. Making business decisions under such conditions demands near real-time information that is complete, precise and can be quickly analyzed and acted on. The key here is actionable information, not just data.

But not all data is created equal, and different approaches are often called for when it comes to processing it efficiently.

Today, both structured data—produced at phenomenal rates by machine-generated activities and web-generated events—and unstructured data—like that of images, text, video, etc.—are prevalent, and the need to fuse the two in order to perform analysis poses a challenge for both private and public sector organizations. But value is hidden in both types of data. It would be a mistake to continue to incorporate only traditional low-velocity structured data into business intelligence systems, as this would ultimately limit the value of the decisions that resulted from it.

Compounding this challenge is the assumption of up-to-the-minute data on a 24x7 basis. And add to this the presence of time series information, like that frequently gained through the data collected by utility companies, and the problem of just how to analyze and query this over-abundance of data becomes even greater.

Utility companies have a strong need to analyze sets of time-series information. An unprecedented amount of actionable information can be captured via smart meters, grid telemetry devices and other instrumentation—or can be lost, depending on one's perspective.

Solutions have emerged, but they have often been Band-Aids applied to past technologies to shore them up to present needs. For example, SQL-type databases, even if retrofitted with mechanisms to handle Big Data, aren't always the best way to phrase particular types of queries. For one thing, time-series queries can't easily be expressed using SQL due to its inherent lack of understanding of temporality. As a result, relating events that are in play around a particular time, or computing moving window functions of stored values, requires a lot of fidgeting in SQL.

Hadoop is often seen as a panacea to cure Big Data ailments, but that's overly simplistic. In fact, it's becoming recognized that hybrid approaches involving solutions that may include Hadoop, massively parallel processing (MPP) enterprise data warehouses (EDW), in-memory columnar databases, stream computing, NoSQL, document databases and other technologies that support extreme analytics in the cloud are emerging in the toolkit of methods to manage Big Data. One solution can't fix all Big Data problems, because it is not actually a single problem that must be solved but rather a class of problems.

Looking Outside and Beyond

Sometimes innovation can come from the likeliest sources, but finding it requires looking outside one’s industry sector.

Cross-industry innovation can yield excellent results. To begin with, it immediately cuts down on uncertainty and risk, which ironically is one of the benefits being sought from business intelligence in the first place. And, while there is no one-size-fits-all Big Data solution, some of the same or similar challenges may have been overcome by an outside industry.

Some sectors use cross-industry often. For example, a recent study of health care firms identified the automotive and aviation sectors as lead industries to which they look for technologies, processes or even business models³. The approach can be similarly compelling in the world of application development.

A number of years ago, my consulting firm was tasked with developing a massive multiplayer online gaming platform for a client. When looking at the technical problems such a system presented, we had an “ah-ha” moment when we recognized the similarities between the broadcast nature of MMOG and the problem of real-time quote distribution for financial markets. Armed with that and the methods we had used for the financial services industry, the challenge was easily overcome— and with certainty of the performance and scalability of the resulting solution.

From Wall Street to Electric Avenue

It’s not surprising that the financial services sector has been dealing for years with many of the Big Data problems utility companies now face. The sheer amount of data that flows through Wall Street is daunting, as is the peak rate at which it flows—over 8 million messages per second (mps). This has increased more than 100-fold over the one minute peak of 56,000 mps reported by the Securities Industry Automation Corp. (SIAC) in 2004. ³ Karoline Bader, Ellen Enkel, Charlott Buchholz, Lorenz Bohn, “A view beyond the horizon: cross-industry innovation in the health care sector”, Performance (May 2013).

But perhaps more importantly, the types of data analysis of value to utility companies— time-series based information incorporating behavior and events affecting consumption and availability—closely parallel the historic information and access patterns commonly found in financial services.

In addition, capital markets and utilities are both industry sectors that need to:

- Perform online data access where sub-second response is important
- Support ad hoc processing of a smaller subset of data within a larger data set
- Enable real-time transactional processing while data is arriving continuously and must be applied to an historical data set
- Work with recent data more often than older data
- Run 24x7 with no scheduled outages



³ Karoline Bader, Ellen Enkel, Charlotte Buchholz, Lorenz Bohn, “A view beyond the horizon: cross-industry innovation in the health care sector”, Performance (May 2013)

Outside the Utility Box

It is because of these parallels that we found for our clients another excellent example of cross-industry innovation based on solutions popular in financial services: using Kx technology. For more than two decades, Wall Street has used Kx's technology, such as kdb+, to handle the demanding high-performance needs and stringent audit requirements of their VLDB and Big Data applications.

With a need to manage time series data and the near-real-time requirements that parallel financial services, Kx technology makes excellent sense for many utility applications, as well as telecommunications, real estate and pharmaceuticals—really any industry with massive amounts of data involving time series events, whether generated by people or machines. Thanks to the database's fundamental design and extended temporal data types, time series events can be manipulated with ease and at a massive scale. Only with a solution like kdb+ can time series analytics be performed innately to the database and not be treated simply as queries involving a time stamp.

Utility Big Data needs more than Patchwork

There is a huge opportunity to leverage the relentless streams of data that organizations of all size are bombarded with today in order to make faster, smarter decisions. In few places can its impact be more influential than in utilities—where the habits and usage patterns of thousands or millions of customers, severe (or minor) weather effects, supplier shortages, and countless other time series events can be game-changing. The kind of game-changing effects that can help utility companies is the constant quest, through smart meters and other instrumentation, to help businesses and consumers make wiser, more educated use of energy. But the deeper one looks at shoring up traditional solutions to handle Big Data challenges with near real-time requirements and time series events, the clearer it becomes that they are inappropriate and even cost-prohibitive.

Thanks to Big Data and wide-scale instrumentation, a revolution around time series analytics is underway, and the technology to help propel it forward has been under our noses in the guise of a financial services solution: kdb+. Optimized for both in-memory and historical data processing, it has set the performance and reliability standard for time series analytics for years.

Even beyond that, the lesson learned from looking to financial services to find the answer to the problems utility companies face around Big Data is an important one: application developers and the businesses they serve must take a less insular look when seeking to solve today's challenges. Often the answers to the questions asked are only an industry away.

About Kx

Founded in 1993, Kx is the provider of kdb+, a time series database for performance-critical environments. Its technology is widely adopted by financial institutions around the world, including RBC, Fidelity, and IEX. The company counts the top ten global investment banks among its customers, many of whom deploy Kx as their enterprise-wide market data server.

In recent years, the company has expanded into new sectors, including utilities, pharmaceuticals, oil and gas, and academic research. Kx is ideally suited to these and other industries that rely heavily on vast data volumes and high-speed analytics.



EMEA

Head Office

3 Canal Quay,
Newry,
BT35 6BP
N. Ireland
Tel: +44 (0)28 3025 2242

Belfast

11-13 Gloucester Street,
BT1 4LS
N. Ireland
Tel: +44 (0)28 9023 3518

Dublin

Fleming Court,
D04 N4X9
Rep. of Ireland
Tel: +353 (0)1 630 7700

London

Cannon Green Building,
1 Suffolk Lane,
EC4R 0AY
United Kingdom
Tel: +44 (0)207 3371210

Americas

New York

45 Broadway,
New York,
NY 10006
USA
Tel: +1 (212) 447 6700

Toronto

31 Lake Shore Road East
Mississauga, On,
L5G 4V5
Canada
Tel: +1 (289) 329 9444

Ottawa

300 Terry Fox Drive,
Kanata, On,
K2K 0E3
Canada
Tel: +1 (613) 216 9095

Palo Alto

#375
555 Bryant Street,
CA 94301
USA
Tel: +1 (650) 798 5155

APAC

Sydney

22 Pitt Street,
Sydney,
NSW 2000
Australia
Tel: +61 (0) 2 9236 5700

Singapore

One Raffles Quay,
North Tower,
#30-03
048583
Singapore
Tel: +65 6592 1960

Hong Kong

Two Exchange Square,
8 Connaught Place,
Central
Tel: +852 2168 0715

Tokyo

Sanno Park Tower,
2-11-1 Nagata-cho,
Chiyoda-ku,
100-6162
Japan
Tel: +81 (0)36 205 3494

