

Big data, big potential

*Leverage the right database foundation
to supercharge your big data journey*



1

**The big
potential
for big data**

2

**Delivering
answers at
the speed
of business**

3

**Addressing
the unique
challenges of
sensor data**

4

**Your next step
on the big data
journey begins
with IBM**

5

Resources

The big potential for big data

The era of big data is here. From retail businesses and telecommunications companies to financial services firms and government agencies, organizations are collecting a tremendous volume and variety of data, ranging from online clicks, transactions and machine-generated sensor data to social media posts, emails and videos.

Executives around the globe are developing new plans to capitalize on this wealth of data. In fact, 83 percent of CIOs surveyed by IBM intend to use business intelligence and analytics to boost the competitiveness

of their organizations.¹ In a separate survey, IBM found that 63 percent of organizations reported using information (including big data) and analytics to create a competitive advantage.² And according to a study by *InformationWeek*, 62 percent of survey participants are using advanced analytics to optimize business operations, 44 percent are identifying business risks and 44 percent hope to predict new business opportunities.³

Retailers want to analyze purchasing trends and customer sentiment to enhance the customer experience, deliver real-time



promotions and improve the effectiveness of marketing campaigns. Telecommunications companies want to scan millions of call detail records to predict customer churn. Financial services firms plan to conduct real-time analysis of streaming market data to drive investment decisions. And insurance firms want to analyze vast stores of customer information to assess fraud risks before signing new policies.

For many organizations, realizing the full potential of big data will require a journey—one that begins with an assessment of their current database solutions.

**What do you need from a database?
Performance and flexibility are key.**

A database must deliver rapid access to data and facilitate real-time analytics to provide answers to questions at the speed of business. Flexibility is also essential; a database should enable your organization to handle the wide variety of data available today—including structured, semi-structured and unstructured information.

IBM® database solutions provide the foundation you need for capitalizing on the big potential of big data. With capabilities designed to deliver fast analytics results, enhance availability, efficiently store large data volumes, accommodate a broad variety of data and control costs, these solutions provide a robust foundation for your big data journey.



Delivering answers at the speed of business

Performance

IBM DB2® is designed to supercharge big data analytics, providing robust data warehousing capabilities to deliver the performance you need to generate new insights fast. DB2 speed of thought analytics with new BLU Acceleration uses dynamic in-memory technology along with innovations such as parallel vector processing, actionable compression and data skipping to dramatically accelerate reporting and analytics.

Dynamic in-memory technology offers the benefits of in-memory processing without the limitations or costs of in-memory-only systems. For analytical queries that need to scan through large sets of data, organizing

data in columns provides a big speed advantage. DB2 with BLU Acceleration enables column processing of terabytes of data that are compressed to fit in memory on the database server. It uses advanced encoding to maximize compression and preserves the order of this encoding so the compressed data can be quickly analyzed.

What can performance gains mean for your business? The results depend on the use case. For example, instead of waiting for a weekly or monthly sales report, a retailer could run ad hoc queries to assess sales performance in real time and then fine-tune in-store promotions or online advertising campaigns. A telecommunications company could leverage big data to



predict the lowest-cost routing of calls and anticipate customer churn. Financial services companies could analyze big data to reduce investment risks, and banks could use it to avoid losses from fraud.

Availability

While speed and efficiency are key, availability is critical too. Online purchases, web page lookups, queries, analytics tasks and other transactions must continue uninterrupted even during planned updates, unplanned demand spikes and other occurrences. Any availability or performance problems can quickly drive away your customers, jeopardize revenues, slow the productivity of business users and prevent timely decision making.

DB2 always-available transactions with enhanced IBM DB2 pureScale® reliability help maintain availability during planned updates and unplanned disasters. With DB2 pureScale, you can scale your environment easily—without interruption—so you can accommodate short-term sales promotions, seasonal demand spikes and longer-term growth. IBM DB2 High Availability Disaster Recovery (HADR) capabilities help ensure that transactions will keep flowing even after large-scale failures.

Flexibility

In the era of big data, you need a database that has the flexibility to support a wide variety of data types and the agility to accommodate change. With its integrated enterprise NoSQL databases (RDF Graph and XML), you can apply the power of DB2 to a full range of data, including structured, semi-structured and unstructured data. NoSQL support can also help you capitalize on emerging opportunities by enabling software developers to accelerate the creation of new applications. These capabilities allow you to fully utilize the data you are collecting and swiftly adapt to changing conditions.

IBM DB2 can deliver extreme availability, helping you keep your customers happy and your business moving.

Leveraging streaming data to provide proactive patient care

The Hospital for Sick Children (SickKids) in Toronto, Canada, wanted to take advantage of new medical monitoring technology to better detect subtle warning signs of complications in pediatric patients. Doctors decided to collaborate with researchers at the IBM T.J. Watson Research Center's Industry Solutions Lab, who were extending a new stream-computing platform to support healthcare analytics.

The collaborative project was Project Artemis—a project designed to help physicians make better, faster decisions regarding patient care. The project was part of the IBM First-of-a-Kind program, which pairs IBM scientists with clients to explore how emerging technologies

can solve real-world business problems. In the earliest iteration of Project Artemis, doctors and scientists focused on early detection of a particular type of infection by monitoring heart rate variability and other indicators.

Project Artemis used IBM InfoSphere® Streams to enable near real-time decision support through the continuous analysis of streaming data. IBM DB2 provided the data management required to support future retrospective analysis of the collected data.

The solution collected and analyzed real-time patient data using algorithms developed collaboratively by clinicians and programmers. By integrating the knowledge and experience of healthcare providers, the solution delivered more robust results than monitoring alone.



Initial successes with Project Artemis led the hospital staff to imagine new ways to implement the solution in a wide range of healthcare settings. For example, healthcare providers could use the solution in conjunction with remote sensors and wireless connectivity to help monitor patients wherever they are while providing life-saving alerts in near-real time.

Addressing the unique challenges of sensor data

Machine-generated sensor data can present particular data management and analysis challenges for organizations in their big data journey. Smart meters, RFID tags, healthcare patient monitoring devices and other sensor sources produce time-stamped data that arrives in frequent, uniform time intervals. Traditional relational databases cannot handle the special internal structure of time-stamped, time-series data.

Organizations need a database with the flexibility to handle this special data type while also enabling interoperation with relational data. The database must deliver outstanding performance and help optimize the storage of the tremendous volume of time-series data generated by sensors (see Figure 1).

IBM Informix® TimeSeries, a built-in feature of Informix software, is designed to address the particular requirements of collecting, storing and analyzing time-series data. With Informix TimeSeries, you gain a relational database that can handle the time-series data type and deliver the performance required for analyzing that data.

Informix TimeSeries stores time-stamped values in a unique way to drastically reduce storage requirements. Efficient data storage helps deliver dramatic performance gains compared with traditional relational databases. In fact, in numerous proofs of concept, Informix TimeSeries has processed smart meter data 50 to 70 times faster than organizations' existing databases while requiring as little as 30 percent of the storage.⁴



Figure 1: Utilities can use IBM Informix TimeSeries to efficiently collect, store and analyze a large volume of time-stamped data from smart meters.

A real-time loader contributes to impressive performance. By dramatically reducing load times, Informix TimeSeries makes data quickly available for processing or querying.

Utilities can use Informix TimeSeries to load data from millions of smart meters in a matter of minutes, instead of the hours it might take other databases. In a recent benchmark, Informix TimeSeries enabled data from 10 million meters to be loaded and processed in less than 36 minutes.⁵ This performance facilitates faster business processes throughout the value chain.

For example, utilities gain the ability to provide rapid feedback to customers to help promote energy conservation.

For financial services firms, performance of Informix TimeSeries assists with fraud prediction and prevention. Firms can rapidly analyze data from biometric sensors that capture and store physical traits (such as fingerprints and retina patterns) to track customer locations and confirm authentication before processing transactions (see Figure 2).

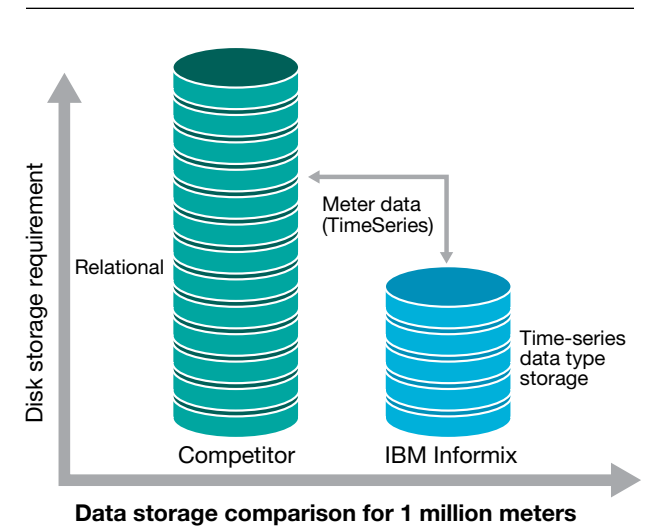


Figure 2: IBM Informix TimeSeries can substantially reduce the storage capacity required for time-series data compared to traditional relational databases.

Reducing heating bills by 30 percent with IBM Informix TimeSeries

In housing developments that use district heating, residents often pay a fixed monthly heating fee. Camden Council, one of 32 borough councils in London, launched a first-of-its-kind heat metering program to provide a cost-effective approach that rewards residents for energy efficiency.

Camden Council turned to Hildebrand Technology, a London-based technology incubator and energy consulting firm, to help introduce an individual metering program that would give each household accountability for its energy usage and encourage energy-saving

practices. Hildebrand, which provided the solution as a service, selected IBM Informix TimeSeries to handle the large volume of time-series data coming in from meters.

“[W]e benchmarked Informix TimeSeries against other databases for capturing meter data,” says Joshua Cooper, CEO of Hildebrand. “Informix TimeSeries was the only one able to handle 50,000 data points per second, scalable to three million homes and beyond....By using Informix TimeSeries in our system, we can give organizations the ability to analyze data on a continuous basis so they can immediately see the impact of weather changes or changing fuel prices.”



The project has highlighted the impressive savings that residents can achieve with greater energy information at their disposal. Some residents could save up to 30 percent on their bills—more than GBP500 over the first two years. At the same time, Camden Council can reduce its carbon footprint.

Your next step on the big data journey begins with IBM

IBM DB2 and IBM Informix include capabilities designed to economically store and process the large volumes and variety of data available to organizations today. With DB2, you gain the performance for real-time decision making, the availability for mission-critical application support and the flexibility to handle a diverse array of data types, all in a cost-effective solution. Informix TimeSeries provides the unique capabilities needed for efficiently and effectively handling large-scale time-series data—an increasingly prevalent type of data in our instrumented world.



Big data holds tremendous potential for organizations in a range of fields. As you begin to deploy new solutions to explore and capitalize on this data, DB2 and

Informix TimeSeries can provide the robust data management foundation required for your journey.

Resources



To learn more about IBM DB2 and IBM Informix TimeSeries, visit:

- [IBM DB2 database software](#)
- [IBM Informix TimeSeries](#)



© Copyright IBM Corporation 2013

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
May 2013

IBM, the IBM logo, ibm.com, DB2, InfoSphere, Informix and pureScale 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

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 and client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. 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.

¹ 2011 IBM Global CIO Study, The Essential CIO, ibm.com/ciostudy

² The IBM Institute for Business Value and the Saïd Business School at the University of Oxford, “Analytics: The real-world use of big data: How innovative enterprises extract value from uncertain data,” 2012, ibm.com/services/us/gbs/thoughtleadership/ibv-big-data-at-work.html

³ *InformationWeek*, “2013 Analytics & Info Management Trends,” November 2012, <http://reports.informationweek.com>

⁴ Based on multiple proofs of concept with IBM customers. See ibm.com/software/data/informix/smart-meter/whytimeseries2.html

⁵ IBM and AMT-SYBEX demonstrate the capability of IBM Informix TimeSeries software to enable the Affinity Meterflow application to offer linear scalability up to 100 million meters to load and process meter data at 30-minute intervals in less than 8 hours. For more information about these benchmarking results, visit ibm.com/software/data/informix/smart-meter/



Please Recycle