Protect your business and customers from fraud

It’s a nightmare scenario that too many people have experienced: your credit card company calls to ask you about the thousands of dollars you spent in a country you’ve never visited. Or you open your monthly credit card statement only to find dozens of purchases you never made. Someone has stolen your credit card information and accumulated thousands of dollars in charges.

As the world has become more connected, the ease of obtaining and using credit cards has increased dramatically. We’ve also experienced rapid growth of online stores offering convenience and low prices, and that combination of convenience and simplicity, along with increased protections for buyers, has led to an explosion in online credit card purchases over the past decade.

But life has also gotten easier for criminals, who no longer need to take physical risks to steal. In the past, thieves would need to steal merchandise from a retailer, or cash or checks from their victims. Now they can commit criminal acts from the comfort of their own homes by using a mobile device and a broadband connection.

As criminals get more sophisticated with their attacks, the solutions to fight them need to keep up. So how do you stay ahead of these well-organized criminals? What if you could quickly harness the wealth of customer demographic and spending information that is available from credit card customers, without the expense of a supercomputer and team of data scientists? For example, if you could compare typical spending amounts, types of purchases, and spending locations with real-time use data, you could quickly flag spending habits that deviate from normal activities. You could also look at spending patterns that commonly indicate fraud, such as a large purchase at a jewelry or electronics store just minutes after a smaller test purchase. To effectively and affordably track these indicators, you need the ability to process real-time data and perform complex analytics quickly, without spending a fortune on large-scale systems and specialized programmers.

From 2010 to 2011, the percentage of Americans notified that their information was compromised in a data breach in the past 12 months rose from 9% to 15%.1
Fight fraud more easily and cost-effectively

IBM and Intel can help you efficiently process and analyze customer credit card data by using powerful, cost-effective tools that build on your existing infrastructure to help you protect against fraud. By combining commodity storage with the strength of efficient, affordable servers built on Intel® processors and chipsets, you can create an effective, affordable solution for processing and analyzing real-time data that previously could only be accomplished with the largest computer systems in the world.

IBM takes full advantage of this processing power with sophisticated software to help you optimize and manage processing and analysis of your real-time data. And IBM has revolutionized how you interact with all that big data by providing a spreadsheet-like interface that puts enormous capabilities in the hands of non-programmers. You no longer need to hire a dedicated data scientist or team of Java® developers to access and analyze Apache Hadoop® data. This easy-to-use interface lets you quickly visualize big data through graphs and charts at a fraction of the time and cost incurred from traditional methods.

INTEL AND IBM: COLLABORATION AND INNOVATION

For more than 15 years, IBM and Intel have collaborated to ensure optimal levels of performance and power consumption, with an ideal price-to-performance ratio for database systems.

By running IBM® DB2® pureScale® on the Intel® Xeon® processor and IBM® System x®, organizations can experience near-linear scalability for their mission critical applications, while enjoying leading levels of data availability.

Intel provides IBM software with multiple performance advantages through:

- Innovation in security and virtualization technology
- Multi-core optimization
- Storage and networking products
- IBM System x joint platform innovation
Build off your current infrastructure

Your company has probably invested a significant amount of time and capital in its infrastructure. Your IT managers want to create additional value by using the best technology, such as emerging big data solutions, to better protect against credit card fraud, but they probably don’t want to replace their existing relational database systems, which contain massive amounts of valuable information.

At the same time, you don’t want to expand your current infrastructure without ensuring scalability and adaptability for changing business needs. You need your systems to have the flexibility to grow and react as quickly as the online world around you.

The key to moving forward is understanding both the limitations and opportunities of traditional relational database management systems (RDBMS) and how the hardware and software solutions from Intel and IBM can help you gain the insights that let you unlock a wealth of actionable customer information currently just out of reach. With this knowledge, your company’s reputation and customers will be better protected from credit card fraud as you respond to emerging threats in real time and on a consistent basis.

Get more from your data

Traditional RDBMS and data warehouse configurations work well for highly structured data, such as credit card transactions, but their rigid design doesn’t easily accommodate unstructured data, such as e-mail or geolocation data. RDMS databases must be carefully designed with the schemas defined in advance. Big data solutions, on the other hand, can rapidly process structured and unstructured data in real time. This advantage adds tremendous flexibility for extracting and correlating useful information from a wide range of data sources.

Big data solutions are designed to address the challenges presented by the three Vs of data management: volume, velocity, and variety. In the context of credit card transaction analysis, volume corresponds to the thousands of credit card transactions that occur every second of every day. Velocity refers to how quickly data can be processed. This factor is particularly important for analyzing transactions for fraud indicators and taking immediate corrective action. Hadoop-based solutions, for example, excel at near real-time modeling and pattern matching of tremendous amounts of data. And finally, variety refers to the types of data being analyzed. In addition to traditional transactional data, granular, real-time credit card analysis also takes other types of data, such as location, into account.
Big solutions for big data

IBM and Intel provide technologies that can help simplify management for IT, make your data work harder for you, and reduce time and complexity for your analysts. Powerful IBM® System x® servers built on Intel processors bring you scalability, efficiency, and high performance that can keep up with the demands of real-time processing and complex analysis.

The following configuration describes a typical use case for handling large volumes of credit card transactions with high performance and real-time analytics for fraud prevention. This configuration, built using IBM® InfoSphere® BigInsights™, is capable of providing continuous, rapid analysis of big data, moving and transforming big data at speeds close to rated network transmission speed (line rate), and processing tens of thousands of credit card transactions each second.

PROCESSING DATA IN REAL TIME

The process begins with real-time processing of credit card transactions by using IBM® InfoSphere® Streams: a high-performance computing product that can rapidly ingest, analyze, and correlate information as it arrives from thousands of real-time sources. In real-world scenarios, IBM InfoSphere Streams processed up to 140 million transactions per hour, with the capability to scale to even higher rates with additional nodes.²

The transaction data is passed on to IBM® DB2® pureScale® for high-performance processing. Built on IBM® DB2®, IBM DB2 pureScale offers clustering technology that helps deliver high availability and scalability to online transaction processing (OLTP) applications. In this scenario, IBM DB2 pureScale performs processing of both customer account and demographics data. In real-world scenarios, IBM DB2 pureScale processed up to 300 million rows per hour.²

TRANSFORMING AND ENRICHING DATA

IBM InfoSphere Information Server, which includes IBM® InfoSphere® DataStage®, unloads the data from IBM DB2 pureScale to transform and enrich the data for analysis. IBM InfoSphere Information Server helps create and maintain trusted information to support strategic business initiatives, including big data, point-of-impact analytics, business intelligence, data warehousing, master data management, and application consolidation and migration. In this configuration, InfoSphere Information Server handles extract, transform, and load (ETL) operations between IBM DB2 pureScale, Apache Hadoop, and IBM® PureData™ System for Analytics. In real-world scenarios, IBM InfoSphere BigInsights was able to perform operations on 4.3 billion rows per hour.²
ANALYZING DATA FOR FRAUDULENT ACTIVITY

The transformed and enriched data is loaded into IBM InfoSphere BigInsights, which forms the foundation for the credit card fraud solution by helping you identify scenarios and patterns for identifying fraudulent activity. InfoSphere BigInsights extends the power of Apache Hadoop with enhanced software features to help you find insights in complex big data. Apache Hadoop is a collection of many different low-level open-source projects and tools. IBM InfoSphere BigInsights organizes and enhances this open source technology to withstand the demands of your enterprise, adding administrative, workflow, provisioning, and security features, along with best-in-class analytical capabilities from IBM Research.

Powerful, scalable System x servers powered by the Intel® Xeon® processor E5 family or Intel Xeon processor E7 family form the hardware foundation for IBM InfoSphere BigInsights.

The transformed and enriched reporting data from the InfoSphere Information Server is loaded into IBM PureData System for Analytics, where it can be accessed quickly to provide insight and actionable information for fraud detection and analysis. Powered by IBM® Netezza® technology, PureData System for Analytics is a data appliance that helps simplify and optimize the performance of data services for analytic applications, enabling very complex algorithms to run in minutes.

IBM PLATFORM SYMPHONY

IBM® Platform Symphony management software enhances IBM® InfoSphere® BigInsights™ by providing highly efficient job scheduling that performs up to 63 times the raw scheduling performance of Apache Hadoop. Platform Symphony is optimized for reliability, low latency, and resource sharing and, unlike the open-source solution, the Platform Symphony MapReduce implementation can restart failed services automatically, delivering improved reliability.

Platform Symphony can be used to monitor all elements of the hardware and software stack and assign workloads evenly across elements of the infrastructure based on current processor speed, server input/output (I/O) performance, or storage capacity and speed. You can even turn off unused resources when monitoring systems through Intelligent Platform Management Interface (IPMI). The result is a more efficient infrastructure that is easier to manage and is better optimized for resource utilization.
PERFORMING COMPLEX ANALYTICS WITHOUT A DATA SCIENTIST

IBM® BigSheets is the cornerstone for this fraud prevention use case. BigSheets, part of InfoSphere BigInsights, is a spreadsheet-like interface that lets business users explore broad varieties and volumes of data (such as structured and unstructured) and discover new insights without writing any code. By using familiar tools, users can rapidly visualize and analyze data without having to rely on Java developers and data scientists. This do-it-yourself analytics capability even lets users create new tools on the fly, without help from IT. Figure 1 shows the BigSheets user interface for InfoSphere BigInsights.

CREATING A MORE AFFORDABLE HARDWARE FOUNDATION

The following hardware components provide efficient, scalable platforms for the use case components described above. With System x servers using cost-effective Intel Xeon processor E5 and Intel Xeon processor E7 family processors, you can create powerful clusters with capabilities that were previously out of reach for all but the largest organizations, using supercomputers.

Intel and IBM have worked together for more than 10 years to improve software performance on Intel processors, and more than 20 years to improve overall system performance. Each new Intel processor generation brings increased performance and new features, and IBM and Intel work to fine-tune IBM server systems and leading software to take advantage of the latest hardware capabilities.

Figure 1. IBM InfoSphere BigInsights offer a powerful, easy-to-use visual interface for creating complex analytics without writing code. With a few clicks of the mouse, you can create collections from petabytes of data, easily run macros to quickly analyze your data in multiple ways, and then enrich your data set by finding relationships and publishing rich visualizations to reveal practical business insights.
In addition to enhanced performance and efficiency, Intel Xeon processor E5 and Intel Xeon processor E7 family servers include Intel® Advanced Encryption Standard–New Instructions (Intel® AES-NI) for accelerated encryption, and Intel® Virtualization Technology (Intel® VT) for hardware-enhanced protections and efficiencies for virtualized systems.

**IBM System x3750 M4 and x3650 M4 servers powered by the Intel Xeon processor E5 family**

IBM System x3750 M4 and IBM System x3650 M4 servers provide the raw processing power needed to handle high-velocity data streams, along with the reliability, availability, and serviceability (RAS) features required for critical credit card analysis systems.

**IBM System x3630 M4 and IBM System x3550 M4 servers powered by the Intel Xeon processor E5 family**

IBM System x3630 M4 servers serve as the compute nodes for large Apache Hadoop clusters, while IBM System x3550 M4 servers function as management nodes. Both feature enhanced power and cooling efficiencies and scale easily to accommodate your needs as your clusters grow in size.

**IBM RackSwitch™ G8316, RackSwitch G8264, and RackSwitch G8052**

These switches are optimized for big data and cloud applications that require high bandwidth and low latency.

- IBM RackSwitch G8316 supports sixteen 40 Gigabit Ethernet (GbE) QSFP+ ports in a 1U form factor.
- IBM RackSwitch G8264 supports up to sixty-four 10 GbE SFP+ ports in a 1U form factor.
- IBM RackSwitch G8052 supports forty-eight 1 GbE RJ45 ports and four standard 10 GbE SFP+ ports in a 1U form factor.

**Intel® Ethernet Converged Network Adapter X540 and X520**

These adapters offer high-speed, low-latency 10 GbE performance and are available with support for both SFP+ and Base-T networks.
Putting it all together: components and data flow

A. IBM® InfoSphere® Streams processes credit card transactions.
B. IBM InfoSphere Streams adds approved and rejected transactions into IBM® DB2® pureScale®.
C. IBM InfoSphere Information Server unloads transaction data from IBM DB2 pureScale.
D. IBM InfoSphere Information Server loads transformed and enriched transaction data into IBM® InfoSphere® BigInsights™.
E. IBM InfoSphere Information Server unloads customer credit risk ratings from IBM InfoSphere BigInsights.
F. IBM InfoSphere Information Server updates IBM DB2 pureScale with customer credit risk ratings.
G. IBM InfoSphere Information Server loads transformed and enriched reporting data to IBM® PureData™ System for Analytics data warehouse appliance.
H. IBM® BigSheets provides visualization for analytic applications access and processing of customer credit card transaction history.

Figure 2. Big data components and data flow
Real-time data, powerful analytics

With powerful software from IBM and scalable, efficient hardware from Intel, you can take advantage of transactional data in real time to provide improved business insight and lower risk to your company and customers by more accurately and quickly identifying potential fraudulent activity. This innovative big data solution puts valuable transaction, demographics, and customer usage data at your fingertips by combining affordable storage with powerful, efficient servers built on Intel processors, and easy-to-configure software and management tools from IBM.

IBM BigInsights expands on the power of Apache Hadoop by adding features to improve performance, scalability, and reliability and let your infrastructure grow and change with your workloads and business requirements. With the simplicity of IBM BigSheets, you can easily harness and visualize your big data to create powerful new insights by using a familiar spreadsheet user interface, with no programming required.

What used to only be possible with supercomputers and dedicated data scientists is now achievable in a cost-effective solution that builds on your existing infrastructure using technologies from Intel and IBM.

Learn more

For more information about the Intel Xeon processor platform: www.intel.com/itcenter/products/xeon

For more information about IBM InfoSphere BigInsights: http://ibm.com/software/data/infosphere/biginsights

For more information about IBM InfoSphere Streams: http://ibm.com/software/data/infosphere/streams/

For more information about IBM InfoSphere Information Server: http://ibm.com/software/data/infosphere/info-server/overview

For more information about IBM System x servers: http://ibm.com/systems/x
Sources


3 Intel® Advanced Encryption Standard–New Instructions (Intel® AES-NI) requires a computer system with an AES-NI-enabled processor, as well as non-Intel software to execute the instructions in the correct sequence. AES-NI is available on select Intel® Core™ processors. For availability, consult your system manufacturer. For more information, see http://www.intel.com/content/www/us/en/architecture-and-technology/advanced-encryption-standard--aes-/data-protection-aes-general-technology.html.

4 Intel® Virtualization Technology requires a computer system with an enabled Intel® processor, BIOS, and virtual machine monitor (VMM). Functionality, performance, or other benefits will vary depending on hardware and software configurations. Software applications may not be compatible with all operating systems. Consult your PC manufacturer. For more information, visit http://www.intel.com/content/www/us/en/virtualization/virtualization-technology/hardware-assist-virtualization-technology.html.

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL’S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

UNLESS OTHERWISE AGREED IN WRITING BY INTEL, THE INTEL PRODUCTS ARE NOT DESIGNED NOR INTENDED FOR ANY APPLICATION IN WHICH THE FAILURE OF THE INTEL PRODUCT COULD CREATE A SITUATION WHERE PERSONAL INJURY OR DEATH MAY OCCUR.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked “reserved” or “undefined.” Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

Copies of documents which have an order number and are referenced in this document, or other Intel literature, may be obtained by calling 1-800-548-4725, or go to: http://www.intel.com/#/en_US_01

Copyright © 2013, Intel Corporation. All rights reserved.

Intel, the Intel logo, Xeon, and the Xeon badge are trademarks of Intel Corporation in the U.S. and other countries.

IBM, ibm.com, BigInsights, DataStage, DB2, InfoSphere, Netezza, Platform Symphony, PureData, pureScale, RackSwitch, and System x are trademarks or registered trademarks of IBM in the United States, other countries, or both.

*Other names and brands may be claimed as the property of others.