Delivering information you can trust
December 2006

IBM Information Management software

Is your data protected?

Safeguarding critical business information and customer privacy with data masking and IBM Information Server

Contents

2 *Effective data security: The key to generating trust in consumer relationships*
4 *Data masking protects sensitive information during outsourcing or offshoring*
5 *Considerations in data masking*
8 *The IBM approach to protecting customer privacy through data masking*
9 *Integrated business information forms the foundation for effective data masking*
15 *Next steps toward a secure information management program*

As movie and music studios, retailers, financial institutions and publishing companies have increasingly made their content and services available online, the role of providing and directing access via voice, data, wireless and broadband media has fallen to telecommunications companies. And these companies know convergence means big business. According to the Tower Group, digital content will generate more than US$11 billion per year by 2009. Additionally, a PaymentOne study indicates that about 46 percent of consumers would purchase even more digital content if they could add the purchase to their existing telephone or broadband bills.[1]

What do these statistics mean for telecommunications organizations? Most importantly, companies must develop and maintain high levels of consumer trust. Without physical products or stores to create a branding experience, customer care becomes a key touchpoint and the primary way to create positive customer impressions.

Effective data security: The key to generating trust in consumer relationships

As the importance of consumer trust has increased, cooperation and coordination with strategic partners has also become an integral part of the telecommunications industry. Providing a seamless customer experience often means sharing customer data with business partners and outsourced or offshored contractors, which can make it vulnerable to theft. According to the U.S. Privacy Rights Clearinghouse, as many as 91 million sensitive customer or employee records have been taken without permission since February 2005.[2]

Unfortunately for telecommunications companies, stories of data security breaches are often played out in the headlines as well as in the boardroom. Customers are aware of all the ways their information can be compromised: denial-of-service attacks, pretexting, social engineering, hacking and "bluesnarfing" on Bluetooth devices, to name a few. They are also acutely aware of the risks associated with identity theft. As a result, protection of customer privacy through secure identity management has moved rapidly up the list of concerns for telecommunications providers.

The business consequences of security breaches can be severe—ranging from public relations damage and loss of customer confidence to lost revenue and legal battles. The average financial impact of security breaches has escalated dramatically,[3] and security threats are growing in numbers and sophistication.[4] For example:

- *In January 2006, the Federal Communications Commission proposed fining AT&T Inc. and Alltel Corporation $100,000 each for apparently failing to certify that they have procedures to protect customers' personal phone records. Despite an AT&T spokesman's assertion that the company did have systems in place to protect customer data, the two companies violated commission rules by failing to have a corporate officer file an annual certificate stating the company has procedures to ensure customer confidentiality.*[5]
- *On January 20, 2006, Honeywell discovered that a former employee had posted payroll data, Social Security numbers and other personal information about 19,000 of the company's workers on the Web. However, the man had not hacked into Honeywell's systems—he had simply "exceeded his authorized access," according to a Honeywell spokesman.*[6]

- *In August 2006, up to 19,000 AT&T customers had their credit card information stolen by hackers who breached security at the company's online equipment store. In addition to absorbing the loss of revenue associated with shutting down the store, AT&T was forced to notify all the credit card companies whose cards might have been involved. The company made public statements acknowledging its deep regret over the incident and offered to pay for credit monitoring for consumers whose information had been stolen.*[7]

Data masking protects sensitive information during outsourcing or offshoring

Information is the most valuable resource telecommunications companies have at their disposal, and breaches in customer trust cannot easily be repaired. Thus, it is of paramount importance that companies protect private consumer data under all circumstances—especially when working with contractors and partners.

One method of protecting critical data is called *data masking*. This technique, in which data is protected by modifying it so that no sensitive information remains, can provide a way to safeguard important data without disrupting outsourced customer service or product development. Data masking disguises sensitive information such as Social Security numbers and credit card numbers in production and non-production databases by replacing it with realistic-looking false data. This data can be referred to as masked, jumbled, encrypted, blocked, scrubbed or sanitized. The masked data is still usable for application development, maintenance and testing.

Done effectively, data masking can provide a variety of benefits. Most importantly, sensitive data is protected from misuse or theft during offshored and outsourced application development, maintenance and testing. In addition, the data masking process has the added benefit of streamlining metadata documentation by cataloguing data definitions for other uses. This occurs because the processes are reusable and are thereby shared for all applications. Companies can use IBM consulting resources to rapidly produce a pilot application and conduct knowledge transfer to the client IT organization for future applications.

Considerations in data masking

Data masking on large and complex legacy systems is not a trivial exercise. Some of the obvious difficulties are described below.

Data documentation: Most legacy systems were developed decades ago and have poor or non-existent data documentation. This makes it difficult to identify the sensitive fields, embedded business rules and redefinitions.

Data relevancy: The masked data is used for testing and application maintenance. It must pass through the range checks and other edits an application may apply. For example, if a Social Security number is replaced with XXX-XX-XXXX, these letters may cause errors in an application screen that is looking for nine digits.

Referential integrity: Most applications maintain referential integrity across tables and subsystems. If a phone number is used as a key, it must be changed in a consistent way across all the tables and across all the subsystems that share the data.

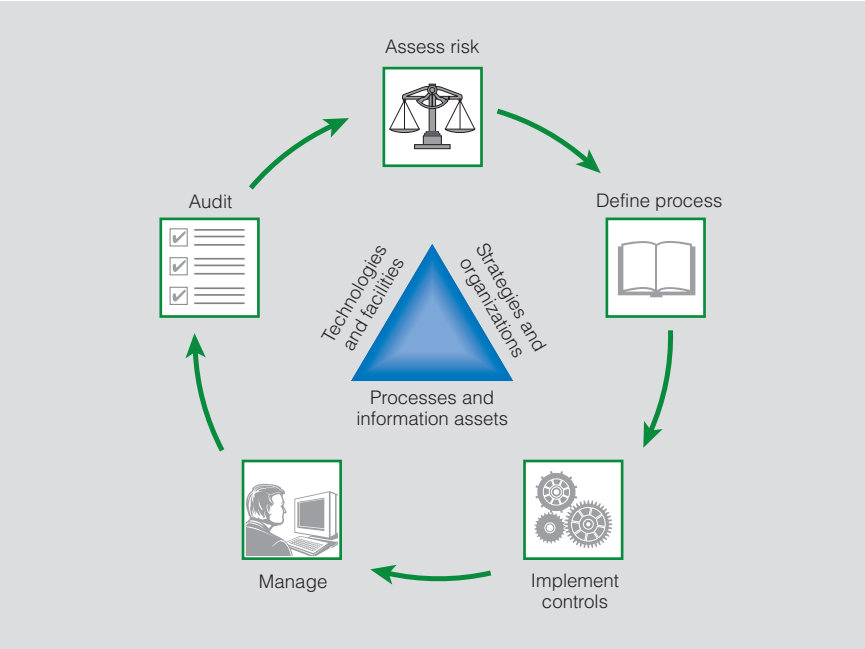
Preservation of data errors: A testing organization would test an application to identify all the potential ways it may break. The realistic production data typically carries data quality problems that must be maintained in the masked test data to check all the error conditions in an application.

Free form data: Critical data may appear anywhere, including in free form text, such as in a notes section.

Data masking is extremely useful for safeguarding sensitive information; however, it can become unwieldy and unmanageable if too much data is masked unnecessarily. Taking into account the organization's strategies, processes, information assets, technologies and facilities, telecommunications companies must continuously examine their risks and capabilities to determine the appropriate masking volume.

Figure 1 shows steps that work together to help achieve an acceptable and manageable level of data masking.

Figure 1: Steps to help achieve effective data masking



The IBM approach to protecting customer privacy through data masking

The IBM approach to data masking brings together IBM® WebSphere® Information Analyzer and IBM WebSphere DataStage® software, as well as data masking utilities developed in IBM research and development labs. Once an integrated information infrastructure is in place, data can be masked in six general steps:

1. ***Extract data from source systems****. After the company has identified the data to be masked, that information must be located and collected. WebSphere DataStage can be used to perform this function.*

2. ***Use data discovery and analysis to understand content and dependencies.*** *With the ability to understand and analyze the meanings, relationships and lineage of information, WebSphere Information Analyzer software can help telecommunications companies discover, define and model information content and structure. By automating data profiling and data quality auditing within systems, organizations can establish an understanding of data sources and relationships, as well as eliminate the risk of utilizing or proliferating "bad" data.*

3. ***Create metadata to support data masking.*** *Telecommunications companies must identify any potential problems or special considerations for the data to be masked, and then decide how the masking will be performed and who will have access to the information. Companies must also develop validation procedures during this step.*

4. ***Use data masking utilities and metadata for data masking.*** *WebSphere DataStage complements the functions of WebSphere Information Analyzer by allowing telecommunications companies to transform and aggregate any volume of information in batch or real time through visually designed logic.*

5. ***Validate masked data to make sure masking has been completed.*** *WebSphere DataStage can be used to verify proper completion of the masking process.*

6. ***Load masked data to the target platform.*** *WebSphere DataStage also provides the ability to deliver information to the people, processes or applications that need it. Data can be moved in large bulk volumes from location to location or accessed in place when it cannot be consolidated, all reusing the same core logic. Information delivery enables telecommunications organizations to help ensure that information is always available, when and where it is needed, while also protecting sensitive customer data.*

Integrated business information forms the foundation for effective data masking

For data masking to be effective at an acceptable and manageable level, information must be used and made available consistently across the enterprise. Addressing this issue hinges on effective, flexible integration of key business data.

Every telecommunications company runs multiple operating environments, each of which creates and stores data to satisfy the unique business requirements at that business level. Convergence of technologies can make it difficult for telecommunications providers to keep track of all the sensitive data housed in various systems across the enterprise. Customer data often resides on siloed infrastructures, making information integration a prerequisite for data masking. To compete effectively in the industry, telecommunications organizations must make sure all data related to operational risk and financial performance is exposed and available—internally and externally—to the right people at the right time, while blocking access to entities that should not have access.

Furthermore, preventing privacy breaches is only part of winning customer loyalty. Increased attention to the customer experience has put pressure on telecommunications providers to explore new revenue streams. Often, this means integrating content and sharing revenues with third parties—a task that can prove impossible without real-time access to information from multiple sources both within and beyond the enterprise. Scattered, inconsistent account information can destroy customer loyalty by causing a frustrating customer care experience—but a single view of the customer can facilitate smooth service and provide insightful analysis for development of future services.

With integrated data, companies can:

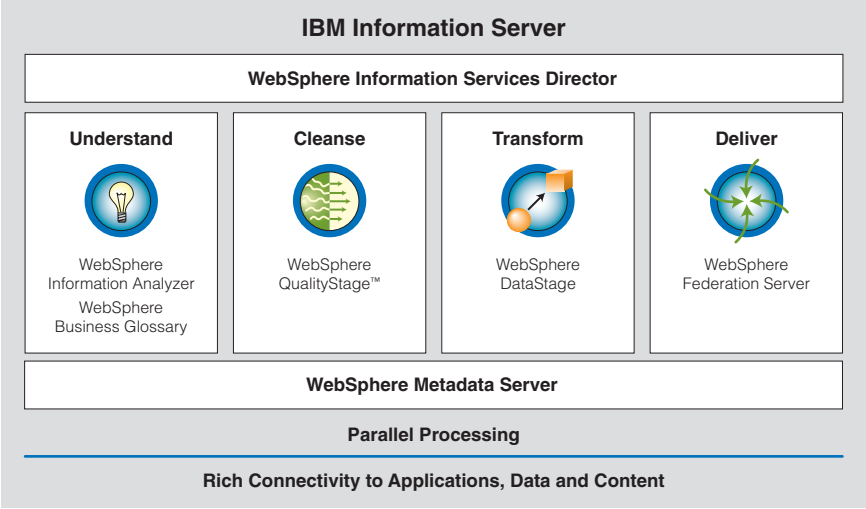
- *Synchronize data to deliver a single view of the enterprise*
- *Integrate transactional data to enable real-time analytics*
- *Consolidate redundant applications and systems*
- *Address regulatory requirements with reliable, accurate and timely information*
- *Reduce integration costs to 20 percent to 50 percent less than manual coding*
- *Empower the enterprise with authoritative, timely and consistent information*

IBM provides the foundation for effective data masking with a unique portfolio of solutions designed to enable the delivery of Information on Demand. IBM Information Server is designed to help telecommunications companies manage convergence and consolidation by consolidating customer, product, network and billing information—helping to reduce complexities and inconsistencies from data silos and unsynchronized databases.

The IBM Information on Demand approach also can help telecommunications companies better understand their customers. A single view of the customer across multiple channels, systems and business units helps enable accurate customer analysis, which reduces customer churn and provides more highly targeted marketing. Additionally, by allowing companies to provide differentiated services based on customer needs, the customer experience is enhanced.

Furthermore, IBM Information Server can help businesses introduce new products and services. By facilitating the introduction of new applications, products, services and third-party content, the platform can play a role in providing services in real time while integrating content and revenue sharing with third-party vendors (see Figure 2).

Figure 2: IBM Information Server helps provide a single view of the customer



Based on an Information on Demand approach, IBM Information Server offers five fundamental capabilities:

1. *The ability to connect to all relevant sources of information wherever they reside, whether structured or unstructured, mainframe or distributed, internal or external*
2. *The ability to understand the content, quality and structure of data sources (including their meanings, relationships and lineage) prior to integration*
3. *The ability to standardize and cleanse data to provide a consistent view of any element of product or pricing information, as well as assure data quality and consistency*
4. *The ability to effectively and efficiently collect, transform and enrich high volumes of data from source to target in a timely manner*
5. *The ability to federate information and make it accessible to people, processes and applications as if it were a single source—without actually moving or copying the source data*

IBM Information Server is designed to provide access to a broad range of information sources, a wide range of integration functionality and outstanding flexibility for Service Oriented Architecture (SOA), event-driven processing, scheduled batch processing and standard application programming interfaces (APIs) such as SQL and Java™. By using SOA to publish consistent, reusable services for information, telecommunications companies can make it easier for processes to get the information they need from across a heterogeneous IT environment.

The IBM Information Server advantage

- *A comprehensive, unified foundation for enterprise information architectures, scalable to any volume and processing requirement*
- *Auditable data quality as a foundation for trusted information across the enterprise*
- *Metadata-driven integration, providing breakthrough productivity and flexibility for integrating and enriching information*
- *Consistent, reusable information services—along with application services and process services*
- *Accelerated time to value with proven, industry-aligned solutions and expertise*
- *Broadest and deepest connectivity to information across diverse sources—structured, unstructured, mainframe and applications*

Next steps toward a secure information management program

Before your company embarks on a data masking and information management program, you need to take stock of your current data security situation. Consider these questions:

- *How safe and secure is your company's supply chain, and how much safe-time do you need? What dependencies does your business have with suppliers from a safety and security perspective?*
- *How does your company compare on safety, security and privacy with industry peers?*
- *Is your investment being spent in the right areas? Relative to this comparison, how efficient is your organization?*
- *How does your company increase awareness and create a security program?*
- *How do you keep your security program up to date?*
- *How secure are the assets under your management and/or with your business solutions?*
- *How do you keep your organization's security skills current?*
- *Are you getting the most value for your organization's safety and security tools (for example, currency, availability or cost)?*

IBM can help with these evaluations, as well as help you design an effective data security program. The combination of IBM Information Server and data masking processes can help telecommunications companies safeguard business-critical information that is paramount to customer privacy and consumer trust.

For more information

To learn more about data masking and information integration solutions from IBM for the telecommunications industry, please contact an IBM representative or visit **ibm.com**/software/data/integration

IBM Software Group
Route 100
Somers, NY 10589

Printed in the United States of America
December 2006

[1] Shwartz, Susana. "Identity Management and Protection: A Key Differentiator." *Billing World & OSS Today,* November 2006.

[2] http://www.out-law.com/page-7243.

[3] CSI/FBI Computer Crime and Security Survey. 2003.

[4] Carnegie Mellon Software Engineering Institute. CERT Coordination Center. 2006. www.cert.org/stats/cert_stats.html.

[5] Associated Press via MSNBC. January 31, 2006. www.msnbc.msn.com/id/11118732/.

[6] McMillan, Robert. IDG News Service. February 6, 2006. www.computerworld.com/securitytopics/security/story/0,10801,108434,00.html.

[7] Out-Law News. August 30, 2006. www.out-law.com/page-7243.

TAKE BACK CONTROL WITH **Information Management**

IMW11801-USEN-00