

## Real-Time Credit and Debit Card Fraud Detection A Shadowbase® Real-Time Business Intelligence Solution

October, 2009

Much of today's consumer activity depends upon plastic. Credit cards and debit cards are used to purchase products and services as well as to withdraw cash from Automated Teller Machines (ATMs). In order for a card transaction to be approved, the ATM, retailer, or food establishment must submit the request, and the bank that issued the card must authorize the purchase or cash withdrawal. Furthermore, this approval must be provided in real-time since the customer is waiting for it. A key problem is quickly identifying suspicious or fraudulent uses so that those transactions can be rejected and the card suspended.

A card transaction is captured by such devices as a point-of-sale (POS) terminal in a store, a customer's browser communicating with a website, or an ATM. The information concerning the request must be rapidly gathered from the servicing network to which the devices are connected, sent to the issuing bank for authorization/approval, and the response rapidly returned in order for the system to complete the transaction.

A major provider offers interbank transaction-switching services for just this purpose. It uses a redundant network of powerful HP NonStop computers to implement an authorization and message switch that gathers the customer transactions from the servicing network, routes them to the appropriate issuing banks for authorization, and then returns the authorization or rejection responses back to the servicing network for delivery to the origination point.



Thousands of merchants connect hundreds of thousands of POS devices to the provider's network to service their retail counters. When a merchant's customer makes a purchase with a card, that card is read via the merchant's POS device, and the amount is entered. The purchase details are transmitted to the provider's switch, which forwards this information to the bank that issued the credit or debit card. The bank authorizes or rejects the transaction and returns a response through the provider's switch to the POS device. The transaction can be rejected for many reasons, such as exceeding a daily limit, exceeding the account's balance limit, an expired card, or a card with a credit hold.

In addition to brick-and-mortar stores and other physical locations such as mall kiosks, the provider services online merchants that receive credit or debit card information over the internet from a customer's browser. This information is sent from the merchant's website to the switching provider's computers, and authorization proceeds as described above.

The provider's switch also routes ATM transactions for many tens of thousands of ATM devices for authorization. ATMs are controlled by a servicing bank. If a servicing bank receives an ATM request involving a card that the bank did not issue, the request details are sent to the provider's switch, which handles the request in the same way that it handles POS transactions. The data is sent to the issuing bank, and the authorize/reject response is returned via the servicing bank to the ATM.

A critical problem faced by the issuing banks is that of fraudulent transactions. This problem has, of course, exploded with the high speed nature of electronic submission of credit and debit transactions. An ATM or POS transaction may be fraudulent, for instance, if the credit card is stolen. Online purchases may be fraudulent if the credit card number, expiration date, and CID number are copied from the card. This copying can easily be done, for example, by a waiter taking a customer's credit card to pay for a meal or when a phone order is placed and paid for with a credit card. Identifying fraudulent transactions typically takes hours or days, and many such transactions may slip through before a hold can be put on the card.<sup>1</sup> Worse, because the information can be quickly shared with thieves in multiple countries, they can rapidly attack via multiple avenues by submitting many different types of transactions simultaneously, anticipating that the lesser (slower, etc) infrastructure that some of them may take will allow at least some of them to get through successfully.<sup>2</sup>

Recognizing this problem, the switching-service provider decided to provide a capability to the issuing banks to flag in real-time those transactions that appeared to be suspicious, and possibly fraudulent. In this way, at the bank's option, suspicious or fraudulent transactions were identified and rejected much earlier and faster than previously possible.

## The Service Provider's Switching System

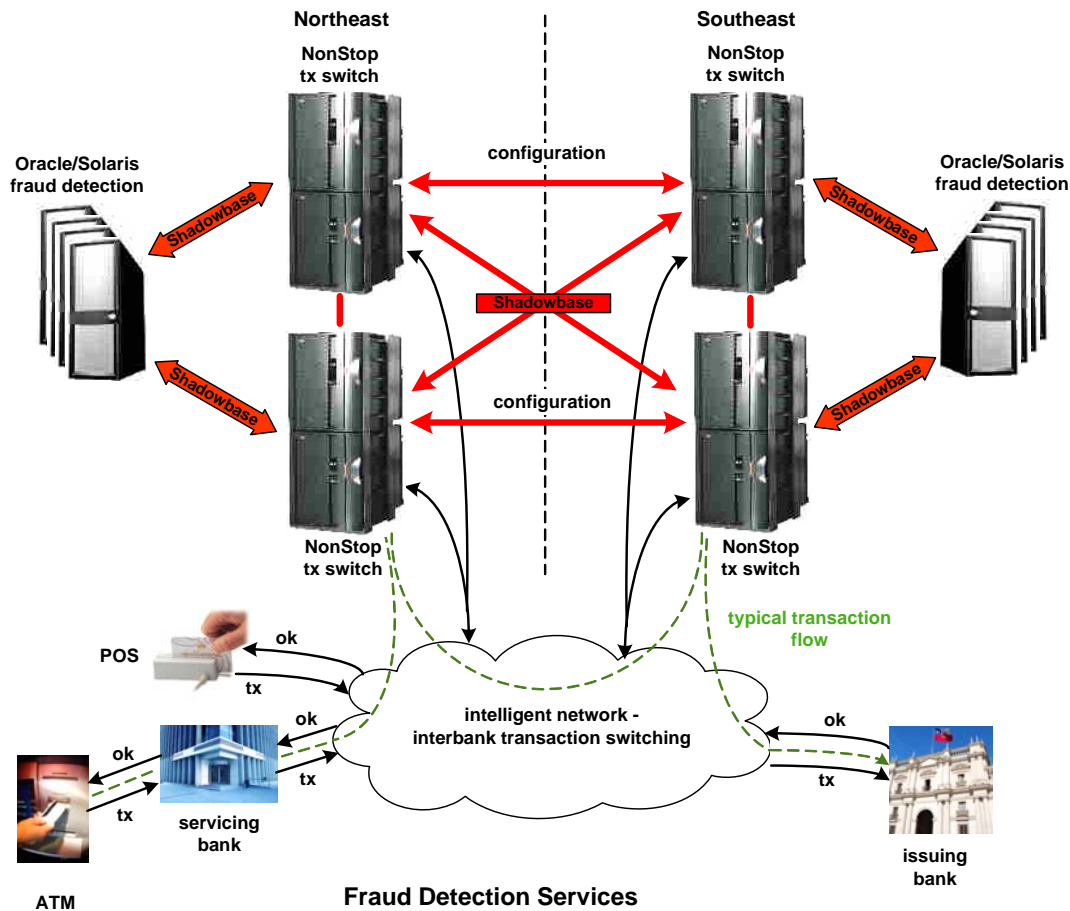
Before delving into the provider's fraud detection service, it is important to understand its transaction switch that provides the backbone for this service. As shown in Figure 1, the service provider maintains a four-node NonStop active/active system<sup>3</sup> to supply transaction-switching services. Each node is a multi-processor NonStop Integrity Server. Two nodes of the active/active system are located in the provider's northeastern U.S. data center, and the two other nodes are located in the provider's southeastern U.S. data center. The geographical separation of the two data centers guarantees survivability in the event of either a localized or regional disaster that takes down one of the centers.

---

<sup>1</sup> ATM and debit-card transactions require a PIN and are better protected against fraud.

<sup>2</sup> For example, in some cases, a store merchant may decide to allow a POS transaction to go through, even when it cannot get it submitted for authorization to the issuing bank (due to local network issues, etc), so as not to inconvenience the customer.

<sup>3</sup> W. H. Highleyman, P. J. Holenstein, B. D. Holenstein, *Breaking the Availability Barrier: Survivable Systems for Enterprise Computing*, AuthorHouse; 2004.



**Fraud Detection Services  
Figure 1**

The nodes are interconnected in a mesh network by Shadowbase bi-directional data replication engines. In this type of architecture, Shadowbase handles extracting the application's database changes, and replicating and applying them to all of the other database copies to keep them properly synchronized. Therefore, any database changes made at a node by the resident applications are subsequently made to all of the other nodes.

POS devices within a store are generally connected to a common controller that is assigned to a particular active/active node (typically, the one that is geographically closest), and all of their transactions are routed to the assigned node via an intelligent network provided by the switch provider. Servicing banks are similarly assigned to nodes in the active/active network. Thus, the load is balanced across the application network.

By contrast, issuing banks are typically connected to one or more of the active/active nodes. Smaller banks connect to one or more nodes via an issuing-bank interconnection network. Large banks may have direct connections to all four of the provider's switch nodes.

Given this architecture, any node can receive a transaction for any card as shown in Figure 1. If the node receiving the transaction is not a node that has a connection to the issuing bank, it will route the transaction to the appropriate node via the provider's intelligent network. Routing is done via the first six digits of the card number, which identifies the issuing bank. The authorization or rejection response flows in the opposite direction, from the issuing bank to its connected node to the node to which the POS device or servicing bank is assigned and then to the source of the transaction. However, if the node

receiving the transaction does have a connection to the issuing bank, as would be the case for most large banks, then only that one node is involved in the routing of the transaction and its response.

The active/active system configuration is used advantageously to eliminate both planned and unplanned downtime. For example, to eliminate planned downtime, if a node needs to be upgraded, its load is transferred to the other node in the same data center. The node set for upgrade is downed, upgraded, tested, and returned to service. The transaction flow for which it is responsible is then restored. By using the companion node in the same data center, the upgrade process is totally within the control of that data center's IT staff.

The intelligence of the interfacing network can automatically detect a node failure. Should it detect such a failure, the transaction flow of the failed node is rerouted to a companion node in the other data center. In this way, application services are restored whether the fault was only a node failure or was a disaster taking down the entire data center.

Consequently, the intelligent network can route traffic to any of the four nodes in the active/active system depending upon the transaction source, the card number, and the status of the nodes. This routing includes incoming transactions from the POS and ATM devices, outgoing transactions to an issuing bank, and the returned responses. In order to accomplish this routing, every node must know the system configuration – that is, the details of the devices and banks, the primary nodes to which they are assigned, and the nodes that are currently handling them. The configuration information in the four nodes within the active/active system are kept in synchronization by replicating changes made to the configuration of one node to the other nodes via the Shadowbase bi-directional replication network.

If a node is taken down for maintenance or if it fails, Shadowbase replication network queues the configuration changes made by the other nodes. Upon restoration, Shadowbase network resynchronizes the downed node's database by draining and applying these queued changes.

The transaction-switching system provides services for hundreds of thousands of devices. The provider's switching services survive any fault in the system, whether it is a network device, a processing node, or an entire data center. Since the network is automatically switching transactions from a failed or downed node to a surviving node, faults are transparent to the provider's customers.

## **Fraud Detection – The Old Way**

Issuing banks typically provide their own fraud detection. A typical fraud detection process proceeds as follows.

The important information associated with a transaction is written to a transaction log by the bank's authorization system. This information includes the card identification, the transaction amount, the time of the transaction, and the location of the transaction. The log is sent to a separate fraud detection system every few hours. The fraud detection system is optimized to perform complex analyses on the transaction log to look for fraudulent activity against any particular card or account. In this way, a series of transactions made over time against a card or an account can be analyzed.

The fraud detection system flags a suspicious transaction with a severity flag and writes this information to another log. Periodically, the log is sent to the bank's authorization system, which takes appropriate action on the card. A credit hold might be placed on the card so that all further transactions will be rejected until the issue is researched or until the problem is resolved. Alternatively, upon the next attempted transaction, the merchant might be informed to ask the customer to call the bank in order to authorize the transaction.

This method is still generally the primary fraud detection procedure in use today. The problem with this method is that it typically takes hours or even days to flag a card that is perhaps being used fraudulently.

During this time, the bank can experience significant losses as additional fraudulent transactions are made. In general, the bank is responsible for such transactions.

## **Fraud Detection in Real-Time**

The transaction-switching service provider realized that there was an opportunity to provide a unique and important service to the issuing banks. If it could detect suspicious or fraudulent activity in real-time, it could stop fraudulent transactions at the retail counter or at the ATM much sooner, or in some cases, even before they were authorized. This service would be a value-added service that would distinguish it from other ATM/POS switching networks.

Providing systems fast and powerful enough to perform this service would be quite expensive and complex – it would require integrating the disparate applications of transaction authorization and fraud detection in new and complex ways – a hindrance to any issuing bank that might want to build its own real-time fraud detection system. However, by building such a system that could be shared by many issuing banks, the cost could be justified.

To implement this system, the switching provider installed multiple high-performance servers that could quickly analyze transactions on-the-fly to determine if they were suspicious. The selected servers were large Sun Solaris servers running Oracle databases. Each server comprised eight quad-core CPUs.

Each data center is provided with its own fraud detection complex, comprising multiple Sun Solaris/Oracle servers (the cards/accounts are assigned to particular Sun Solaris/Oracle servers at a site in order to partition the work load). The fraud detection complex is easily scalable to handle additional load by adding additional servers and reassigning the accounts/cards accordingly.

In this new approach, when a transaction is received by a switch node, it is sent not only to the issuing bank for authorization, but it is also replicated in real-time to a fraud detection server via a Shadowbase replication engine. Shadowbase engine routes the transaction to the particular fraud detection server that is monitoring that card or account. Transaction distribution by card number or account is accomplished via routing rules configured into the Shadowbase replication engine.

The powerful fraud detection system rapidly analyzes the transaction on-the-fly and, if suspicious, notifies the switch node via reverse replication using the Shadowbase replication engine. This notification contains a severity flag indicating the degree of suspicion. The goal is to be fast enough to beat the issuing bank's response so that the transaction can be flagged before it is returned to the servicing bank or to the merchant's POS device. In any case, the issuing bank is notified of the fraud finding. If the flagged message is received by the switch node after the authorization response has been returned to the merchant or to the issuing bank, at least the issuing bank (and the provider's message switch) has been notified and can take appropriate action on the next transaction, far sooner than it would otherwise be notified under the old fraud detection approach.

Of course, transactions for a given card may come into both data centers as the card is used at different stores or ATMs. Therefore, in order for fully effective fraud detection to work, both data centers must know all of the transactions for all cards. This is accomplished by the switch nodes, which block transactions as they are received and send them in near-real-time to the opposite data center, where they are recorded in the fraud detection system at that data center. In this way, both fraud detection complexes know about all transactions for each card and can effectively monitor each card transaction for fraudulent use no matter which switching node receives the transaction. Additionally, this provides a full disaster tolerant back up of the fraud detection services at each node, should one of the data centers be lost in a disaster.

The action taken by the switch node for a suspicious transaction can be configured to correspond to the desires of the issuing bank and the merchant. Different levels of actions can be specified for different levels of suspicion severity. In some cases, it may be desirable to reject the transaction. In others, it may

be desirable to request that the customer call the bank before this transaction or additional transactions can be authorized. In still other situations, the issuing bank may want to allow the transaction but leave a voice or e-mail message for the customer notifying him of a potentially suspicious transaction.

## Summary

This real-time fraud detection system is an excellent example of real-time business intelligence (RTBI), in which events as they happen can control the operational actions of an enterprise. Consequently, real-time business intelligence is often referred to as event-driven business intelligence. A fundamental benefit of RTBI systems is that they can integrate in real-time the independent results of diverse heterogeneous systems and consequently affect the actions of an operational system.

In the example described above, a complex suspicious or fraudulent activity determination is made and action taken while a transaction is in the process of being gathered, routed, authorized, and returned to the origination point, or shortly thereafter, typically far sooner than otherwise achievable.

In the transaction switch described above, RTBI is made possible by the high-performance, heterogeneous Shadowbase bi-directional data replication engine. Shadowbase technology can replicate data between a wide variety of databases and platforms, changing the data as it is replicated to meet the needs of the target application or of the target database's schema. The Shadowbase engine is a high-performance, low-latency replication engine that can typically replicate between platforms in tens to hundreds of milliseconds. It is easily scalable to match any needed replication load and is configurable so that capabilities such as, in this case, routing transactions to the proper fraud detection server are simply added.

Real-time business intelligence will provide the competitive edge to companies in the future. Shadowbase solutions are positioned to help your company achieve this edge.