the Availability Digest

Bank-Verlag – An Update

August 2010

Wolfgang Breidbach and his colleagues may well be the fathers of active/active systems. At least, we don't know of any earlier active/active system. Equally interesting is that the driving motivation for this early active/active system was not only high availability. It was also zero downtime migration.

In an earlier article,[1] we described Bank-Verlag's system as using transaction replication. In a recent, interesting discussion on the Defining Active/Active thread of our LinkedIn Continuous Availability Forum, Wolfgang provided significant additional detail about his approach.

Here is his updated story.

Bank-Verlag and the Debit Card

Bank-Verlag is a subsidiary of an association of over 300 German banks, including Deutsche Bank and all other large banks in Germany. It was established in 1961 to be the publishing arm for the association and published a banking magazine that was distributed to the association's member banks.

In 1985, the association directed Bank-Verlag to start an electronic banking service for its smaller member banks. Wolfgang was hired as Bank-Verlag's first IT employee.

His first task was to provide online banking services and debit-card production for these banks. He was given the use of an IBM System 370 in a data center of one of the association's banks to implement the services.

The debit-card service involved creating cards based on a bank account. Money within a daily limit could be drawn once a day from an ATM. The ATM transaction was booked against the bank account afterwards. Everything needed was encoded in each debit card's magnetic stripe.

As a customer used a card, the date of the last transaction was recorded in the card's magnetic stripe. There was no online tracking of the transactions by a central system. After all, why would this be necessary? The card told all.

[1] Bank Verlag – The Active/Active Pioneer, *Availability Digest*; December 2006. http://www.availabilitydigest.com/private/0103/bank_verlag.pdf

The TV Exposé

Until one fateful day in 1986. Bank-Verlag's management was stunned to learn from a TV investigative report that people were using debit cards fraudulently. The scheme was quite straightforward.

The fraud was initiated by stealing a debit card and by somehow managing to get its PIN. The PIN was obtained perhaps by looking over the owner's shoulder or because the owner had kept the PIN together with his card. At that time, people were not familiar with all the risks.

The thieves simply purchased readily available equipment to read and write card magnetic stripes. They read and wrote down the data in the stripes of the stolen debit cards, including the dates of the last transactions. They then withdrew money from ATMs within the daily limits of the cards and later rewrote the magnetic stripe of each debit card with the old data, especially the date of the last transaction. In this way, they easily could exceed the cards' daily limits. They could use the cards over and over again, and the banks were never the wiser. It took a TV reporter to uncover the fraud.

As a result, Bank-Verlag was ordered by the banking association to build a central authorization system for the banks' debit cards. Wolfgang and his colleges immediately set to work to do just that, again using a data center IBM 370 for the authorization task. The authorization system was front-ended by an IBM Series 1 system as a communication subsystem. They were able to implement this application in a very short time, and the system went into service later that same year.

Using this new system, ATMs reported in real time the contents of the magnetic stripe to the authorization system, which validated the data and returned the data to be recorded by the ATM on the stripe. No longer could one fraudulently modify the data on the stripe.

Of course, if the system were down, no one could withdraw cash using a debit card. However, at that time, ATMs were not so popular. Thus, this was not deemed to be a big problem.

Enter Tandem

Shortly after the new authorization system came online, Deutsche Bank acquired the bank that owned the data center being used by Bank-Verlag. Deutsche Bank closed down that bank's data center, and Bank-Verlag's processing operations were moved to a Deutsche Bank data center. This did not sit well for competitive reasons with the other banks in the association, and they authorized Bank-Verlag to open its own data center.

Bank-Verlag now had the opportunity to revisit its choice of data-processing systems. It investigated several high-availability systems, including those of Nixdorf and Siemans. However, Bank-Verlag had close relations with Banksys, a financial transaction routing service located in Brussels, which was using a little-known system from Tandem Computers.

What impressed Bank-Verlag was that while the providers of the other systems focused on the number of 9s *before* the decimal point to describe their availability. Tandem focused on the number of 9s *after* the decimal point. Other system vendors bragged that they could achieve two 9s before the decimal point (99%). Tandem claimed that they could achieve two 9s after the decimal point (99.99%).

Debit cards were becoming more and more popular, and the inconvenience of not being able to use them because the authorization system was down was rapidly becoming a serious issue.

This was strong motivation for Bank-Verlag to select a highly reliable Tandem system for the new authorization system.

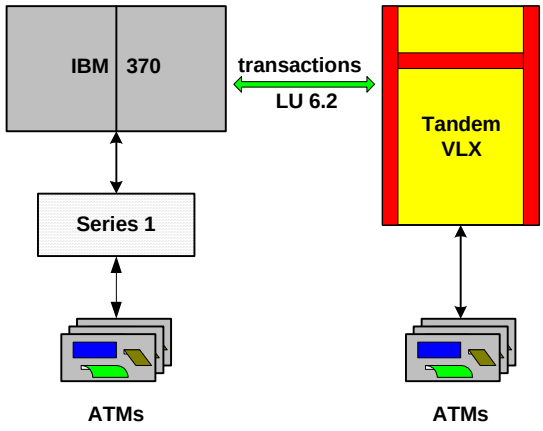
Furthermore, Banksys in Belgium had an application package that ran on Tandem and that did a lot of what Bank-Verlag needed. That clinched the deal. Of course, Bank-Verlag's staff had to write additional applications, which they did in TAL and COBOL. They put their first Tandem VLX into service in 1988.

The Active/Active Inspiration

Now came the tough part. All of the debit-card authorization was still being done on the IBM system, but Bank-Verlag had to move that processing to the Tandem system without denying service to the debit-card holders. It did not want to do this as a "big bang" migration but rather as a controlled, incremental migration. Thus was born active/active.

To accomplish this, Wolfgang's group modified both the IBM and the Tandem applications so that each would send debit transactions that it had processed to the other system. Transactions were interchanged as LU 6.2 messages, a protocol supported by both systems. The strategy was to process each transaction on each system so that the systems would remain synchronized.

The group then copied the debit-card database from the IBM system to the Tandem system and moved a few ATMs to the Tandem system. Data collisions were not a problem, as a debit card could not be at two ATMs at the same time (at least, legally). Now, any debit card transaction executed on one system was also executed on the other system.



Is This the World's First Active/Active System?

As Bank-Verlag became comfortable with the new system, it moved more ATMs to the Tandem system until all had been moved. This entire migration process took only a few days. The IBM system was kept available as a hot backup.

Extending to Disaster Tolerance

In 1989, the growth in the use of debit cards was exploding; and the authorization system needed to be expanded. In addition, Bank-Verlag wanted to configure its system to be geographically distributed in order to achieve a degree of disaster tolerance. Therefore, it purchased an additional VLX and installed it about three kilometers from the original data center (Germany is not plagued by hurricanes or earthquakes, so three kilometers was deemed to be a safe separation distance). After the installation of this system, Bank-Verlag decommissioned the IBM system.

Wolfgang and his group dusted off their active/active facility and brought it up to support the new node. A problem that Bank-Verlag faced was that the LU 6.2 communications software licensed from Tandem was very expensive. Bank-Verlag therefore rewrote the transaction-replication logic to use SNA LU0 over X.21 leased lines instead. This communication technique was faster, imposed less CPU overhead, and required no specially-licensed communication software from Tandem.

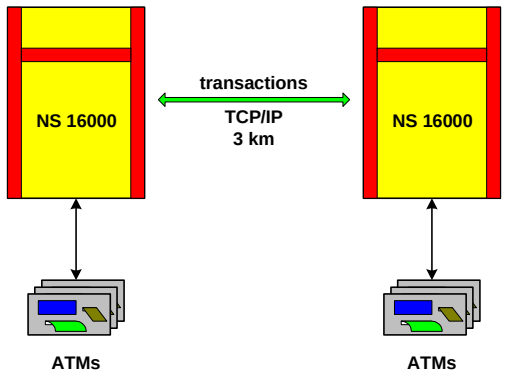
The Growth of the System

As the years went by, the Tandem (then Compaq, then HP NonStop) systems grew to Cyclones, then to K-series systems, then to S-series systems, and today to NonStop Integrity systems.

The conversion to S-series was particularly painful for two reasons. First, the S-series did not support the X.21 communication controllers. The applications had to be rewritten to use the SWAN controllers and TCP/IP, and all of the communication cables had to be changed and rerouted. Secondly, the debit-card transaction volume had outgrown the capacity of the K-series disks; and the transition to S-series disks had to be made.

Because of scheduling pressures, Wolfgang had to put into production in just a few weeks a beta version of the new application. It had not been thoroughly tested but ended up working well. He had scheduled a downtime window of 24 hours to replace a node (he did one node at a time) but was able to cut over in 12 to 15 hours.

Recently, Bank-Verlag upgraded from two six-processor S-series (S86006 and S72006) to a pair of four-processor Integrity NS16000s. This time, no change in communication subsystems or disk systems was required (though Bank-Verlag migrated S-disks to Integrity disks to save maintenance charges); and the conversion went smoothly with no significant problems.



Bank-Verlag Today

As its systems upgraded, so did Bank-Verlag's responsibilities. It now acts as a financial transaction gateway for debit and credit cards issued by any bank. As a user requests services at an ATM or makes a purchase at a retail store, his transaction must be authorized by the bank that issued the card. In some cases, Bank-Verlag acts as the issuer. In most cases, however, the transaction is routed by the Bank-Verlag system to the bank that issued the credit or debit card for authorization. The acceptance or rejection of the transaction is then routed back to the originating ATM or retail point-of-sale (POS) device to conclude the transaction. If the issuing bank's authorization system is down, Bank-Verlag can provide stand-in authorization until the authorization system is returned to service.

Achieving Availability and Scalability

The Bank-Verlag active/active infrastructure was developed long before the advent of commercially available replication engines. Bank-Verlag therefore had to develop its own data-replication facility. Rather than replicate each change as it is made to the database, a transaction is totally processed by one of the nodes in the application network. The resulting changes for that transaction are then sent as a group to the other node.

Transaction Replication

A transaction can be sent to either node in the active/active network. That node, which we will call the primary node for the transaction, processes the entire transaction. In some cases, an ATM or POS device will always send its transactions to the same node (if that node is up, otherwise to

the other node). In other cases, an ATM or POS device will round-robin its transactions to both nodes, skipping a node if it is down.

Typically, processing by the primary node includes:

1. Converting the transaction to the internal application message structure, including decryption.
2. Preparing the transaction for internal processing, including PIN verification (if Bank-Verlag is the card issuer) or for routing.
3. If the card was issued by another bank, re-encrypting the transaction and sending it to the issuing bank.
4. Receiving and decrypting the response from the issuing bank.
5. Processing the transaction request according to the authorization response.
6. Encrypting the response and returning it to the ATM or POS device.

At this point, the internal application message is sent to the other node, which we will call the secondary node for the transaction. The secondary node performs only Step 5 above. Based on the authorization response, the transaction message is processed; and the secondary database is updated.

Thus, each node has a current copy of the application database. A transaction can be sent to either node, and it will be properly processed. Should a node fail, all that is required to maintain service is to route all transactions to the surviving node. In over two decades of operation, Bank-Verlag has never seen an outage in this system.

Data-Collision Detection and Resolution

It is possible, though highly unlikely, that a data collision can occur. After all, there is only one copy of a debit or credit card; and it cannot be used in two places at the same time (though this is possible with online transactions). Therefore, the application provides data-collision detection. If a collision is detected, the result of the transaction's primary node is accepted; and the results of the secondary node are rejected. Fortunately, this situation has proven to be extremely rare; and such transactions are reported and scrutinized very carefully by Bank-Verlag staff.

Zero-Downtime Upgrades

Furthermore, since the nodes are loosely coupled by the replication process, they do not have to be the same. They can use different hardware and software. Keep in mind that the application was originally developed with one node being an IBM system and the other node being a Tandem system.

This capability has proven to be important during the many system upgrades that Bank-Verlag has performed. It can upgrade one node to new hardware and software and then can run on both the old and the new nodes for a while to ensure that the new node is operating properly before upgrading the second node. If it experiences a problem with the upgrade, it can fall back to the old node until the problem is corrected. Years ago, for instance, Bank-Verlag moved from key-sequenced Enscribe files to SQL. It upgraded one node to SQL and ran with a mixed system for six weeks as a trial with no problems before upgrading the other node to SQL.

Scalability

This infrastructure is not limited to two nodes. The secondary node does not have to do any of the encryption or decryption, any of the communications with external authorization systems, or much of the processing. It is estimated that secondary processing only amounts to about 20% to 30% of

the primary processing load. Therefore, the system is inherently scalable. Nodes can be added to carry additional load without having to increase the size of any of the nodes. The only change that needs to be made to the application when a node is added is to include it in the replication of transactions.

Bank-Verlag has recently completed the move of one of its data centers. To ensure continuous availability during the move, it brought up a third node so that there would always be two operating nodes as one node was moved to its new site. The move was eminently successful, with no downtime experienced by the bank's users.

Postscript

Bank-Verlag has since been reorganized. It is now a holding company managing two operating companies:

- BV Media continues to be the publishing arm of the association banks.
- BV Payment Systems provides banking services, card authorization, and secure PIN letters for the association banks (a secure PIN letter is a tamper-proof letter used to notify customers of PINs for their debit cards).

BV Payment Systems runs NonStop servers as well as several UNIX-based IBM systems. Wolfgang is now the Director of NonStop Systems for BV Payment Systems. His new applications have left TAL and COBOL far behind and are written in C, C++ and Java.

Looking at the history of Bank-Verlag, one has to be impressed with the speed and efficiency that characterized many of its migrations. Wolfgang attributes this in great part to the fact that Bank-Verlag remains a fairly small company. This gives it an agility difficult to find in larger companies.

Are Wolfgang and his colleagues really the fathers of active/active systems? They certainly are pioneers in these architectures. There is a quaint American saying that "You can always tell a pioneer – he is the one with the arrows in his back." Wolfgang and his colleagues have escaped the arrows, but they are still the earliest active/active pioneers of whom we know. If anyone reading this knows of an earlier pioneer, please inform us at the Availability Digest so that we can chronicle that experience.