the Availability Digest

www.availabilitydigest.com

European Bank's Active/Active ATM Network

June 2009

A large multinational European bank uses Base24-atm from ACI (http://www.aciworldwide.com/) to run its ATM network. For over fifteen years, this network has been managed by an active/active configuration[1] comprising Tandem systems (now HP NonStop servers). During this time, the bank has experienced no major outages, either planned or unplanned, attributable to the NonStop system, not even during three system upgrades.

The ATM Network

Among its retail services, the bank provides ATM services. The ATM machines offer a variety of features. In addition to providing cash, these features include reviewing account balances, transferring funds between accounts, printing statements, paying bills, PIN management, and even topping up mobile phones. The system, in fact, provides all of the bank's retail customer banking services except for Internet banking.

In addition to its network of ATMs, the bank also services point-of-sale (POS) devices for retail merchants. So far as the network is concerned, POS devices are made to look like ATMs; and communication with these devices is handled in the same manner as it is with ATMs.

Between the 4,000 ATM and POS devices, over 1.5 million transactions per day are processed. The peak time, interestingly, is seven minutes past 1 PM on the last Friday of each month at lunch time, just after paychecks are issued.

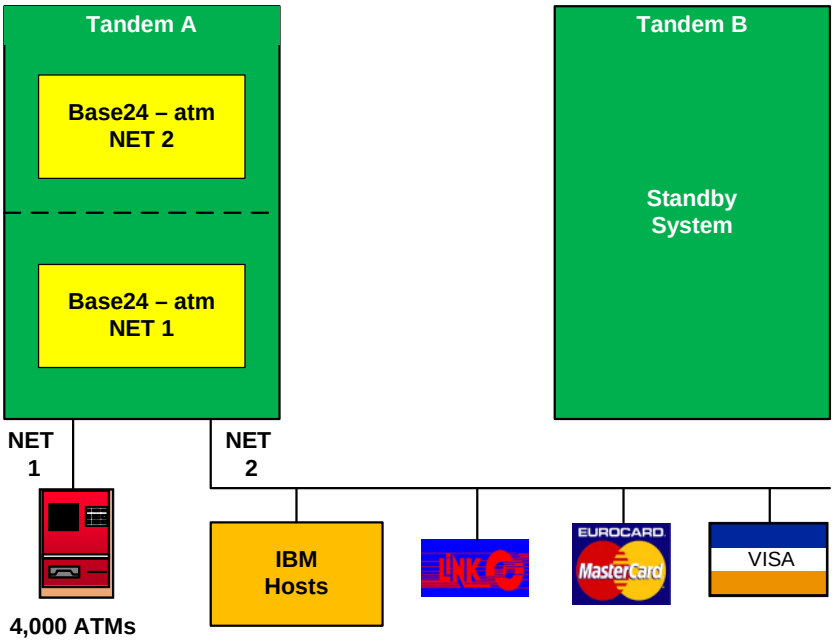
The Original ATM Network

The original implementation for the ATM network used an active/standby Tandem configuration. Transaction traffic was carried over two logical networks, NET 1 and NET 2. One logical network handled the ATMs. The other network carried POS traffic from VISA, MasterCard, and LINK and connected the ATM system with the bank's IBM hosts.

All point-of-sale traffic comes from VISA, with some outgoing traffic being routed to MasterCard. LINK is a financial clearing house that routes POS transactions between the POS devices and the bank issuing the credit card being used at that POS device.

The primary Tandem system (Tandem A) normally handled all traffic. Had it experienced a problem, transaction traffic was rerouted to the standby system, Tandem B, which provided full functionality to the network.

[1] What is Active/Active?, *Availability Digest*; October, 2006.



Active/Standby ATM System

A recurring problem with this system was communication faults. Communication links used the X25 and LU 6.2 protocols and proved somewhat unreliable. When a network went down, the ATMs or POS devices suffered an outage. Furthermore, transient errors on the network caused the ATMs or POS devices to time out, requiring that the user reenter the transaction details.

The Move to Active/Active

The ATM System's Active/Active Configuration

To alleviate the communication problems described above, the bank reconfigured its active/standby system into an active/active pair of Tandem nodes. In this configuration, both nodes are actively processing transactions against a common application database. Each node has its own copy of the database, and these copies are kept in synchronism via data replication.

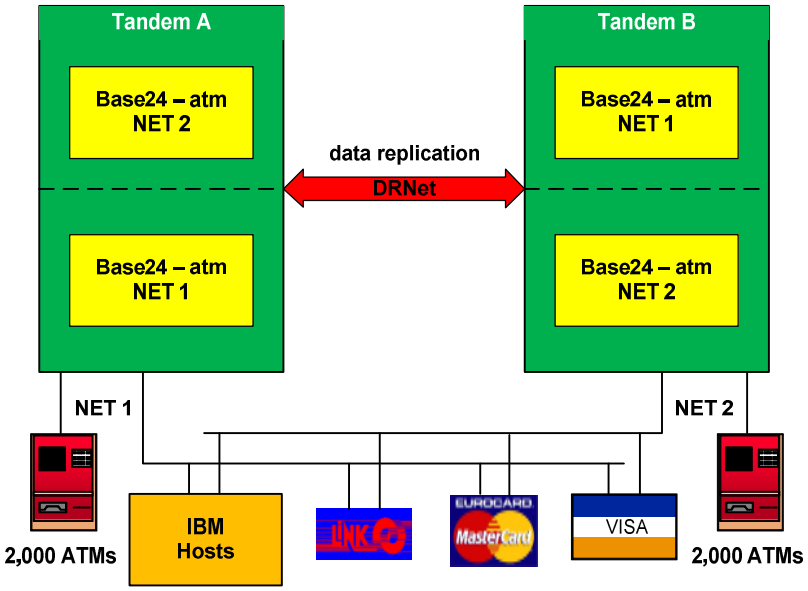
Each node is sized so that it can handle the entire transaction load should its companion node fail.

To provide disaster tolerance, the nodes are located at two different sites ten kilometers apart. In this way, an incident that might take down one site will not cause a total system failure. The other node at the remote site will take over the transaction load.

The concept of dual logical networks is maintained. The 4,000 ATMs are split between the two nodes, with 2,000 ATMs connected to each site. The POS network connecting with VISA, MasterCard, LINK, and the bank's IBM hosts is replicated so that there are separate connections to each node.

The bank refers to this configuration as AB (logical network 1 on node A and logical network 2 on node B). If one node is down, the system is running in AA or BB mode.

As time has passed, communication-line quality has improved; and the bank has moved away from the old X25 and LU 6.2 networks to a TCP/IP network.



ATM Network Active/Active System

The system is designed to avoid data collisions. Each ATM card has a unique number. Though multiple cards issued to the same family may carry the same number imprinted on the card, the magnetic stripe carries additional identifying numbers. Should two transactions come in simultaneously with the same card number, thus creating a data collision, the transactions are rejected as fraudulent.

The splitting of the ATMs between the two nodes ensures that ATM traffic is load-balanced. POS traffic is also load-balanced. VISA traffic is split between the two nodes using the first numbers on the cards. LINK uses its own algorithm to balance its transaction load. IBM host traffic comprises periodic file refreshes that do not have a real-time urgency.

Should a node or network fail, all POS traffic is immediately routed to the surviving node so that there is no interruption in POS transaction processing. Only those ATMs that are connected to the failed node or network are down. If multiple ATMs are located at a single location, care is taken to distribute them between the two nodes. Therefore, a node failure will still leave some ATMs operational at each multiple-ATM site. The downed ATMs are rerouted to the surviving site by making updates to the DNS servers, following which the ATMs are restored to service.

As of this writing, each node is a ten-CPU NonStop 76000 server. The bank may upgrade the nodes to NonStop Integrity servers in the future.

Maintaining Database Synchronization

Each of the ATM network nodes is executing transactions locally. This means that they both must have a local copy of the application database, and these copies must be synchronized with each other. When a change is made to one of the databases, this change must be immediately reflected in the other database.

Database synchronization is accomplished by asynchronously replicating data bidirectionally between the Node A and B databases. The bank uses the DRNet data-replication engine from Network Technologies International (www.network-tech.com) to accomplish this.

A variety of files are replicated. They include:

- *the Card Authorization File (CAF)*, which contains the details for each ATM card, including card transactions and card status (card lost, card blocked, etc.). The CAF contains 35 million records and is updated on each card transaction. In addition, the CAF file is rebuilt from the IBM hosts every six months and is partially refreshed every month. A full refresh is made to both Nodes A and B and does not require replication. A partial refresh is sent to one side and is replicated to the other side. In both cases, refreshing happens in parallel with ongoing transaction replication.

- *the M(TLF) file,* which is a subset of the Transaction Log File (TLF). The TLF file is the Base24-atm real-time event log that records all events occurring within the system. The M(TLF) file contains only those entries that are financial transactions.

- a variety of fairly static files such as those that provide information concerning institutions, access authorization, and encryption keys.

Node or Network Failure and Recovery

Should a node or a communication link fail, all transactions can be routed to the surviving node since it has an up-to-date copy of the application database. As long as the other node is down or is not connected, the surviving node will process all transactions and will queue the changes that it makes. Once the downed node is ready to be returned to service, the surviving node will drain its queue of changes to the downed node, thus synchronizing the downed node's database. At this time, the downed node can be returned to service; and the transaction load can be split once again between the two functioning nodes.

This same capability is used to roll upgrades through the system. First, one node is taken down and upgraded. It is then returned to service, and the other node is taken down and upgraded. Service is never lost to the users of the system.

Testing Failover

A major benefit of active/active systems over active/standby systems is that it is known that the contingency system is up and running. Since the standby system in an active/standby configuration is idle (or is at least not running the application that it is backing up), there is a possibility that the failover to the standby system will fail. Testing the standby system is a costly and risky operation and is often not thoroughly done by many organizations.

In an active/active configuration, it is known that both nodes are working since they are both actively processing transactions. Should one node fail, all that needs to be done is to reroute transactions to the surviving node.

However, active/active failover should still be tested periodically. For instance, there may be problems in the mechanism for rerouting transactions. The bank tests failover of the ATM system twice yearly, failing over to Node A as well as to Node B. This is often done in conjunction with hardware and software upgrades.

Switchover is managed via an operator console developed by the bank. The console keeps track of the status of all ATMs in the network and also manages the failover process.

Following the failover to a single node, the operational node continues to replicate data changes to the other node if it is up; or it queues changes for later recovery if the other node is down.

The ATM Failover Problem

A major advantage of active/active systems is that recovery from a node or network failure can be accomplished in seconds. However, in the bank's ATM system, it takes about 45 minutes to restore service to the affected ATMs. Why is this?

The current ATMs are not network-intelligent. They connect only to a single IP address. Therefore, in order to switch the ATMs to their alternate node, the network has to be changed. This is done by modifying the routing entries in the DNS server, a process that takes some time.

This switchover time has been deemed acceptable to the bank since to the customer standing at the ATM, it is no more inconvenient than when the ATM runs out of cash. If the customer is at a site with multiple ATMs, he simply has to move to an ATM that is connected to the surviving node and that is therefore still functional.

The bank may modify the ATMs so that they can switch IP addresses if they get no response. Each node will have a different IP address for the ATM connections. If an ATM determines that it is getting no response from the node to which it is connected, it will simply switch to the IP address of the other node and will resubmit the transaction. With this technique, ATM connections can be rerouted in seconds.

The Bank's Experience with Active/Active

After fifteen or more years of using an active/active configuration for its ATM network, the bank has accumulated a wealth of experience with this technology.

Major Benefits

- The problems with the contingency system are minimized. It is known that both nodes are operational since they are both actively processing transactions. System availability is improved since there are no failover faults.

- In addition, all communication links are being actively used. The bank knows when it has a communication-link problem and does not have to wait until it tries to use a backup link to find this out.

- Each node must be configured to handle the entire load should one node fail. Therefore, during normal operation, each node is handling only half the load, leading to better performance due to reduced CPU and disk loading. In short, all available capacity is being actively used.

- Likewise, each communication link is sized to handle the entire transaction load. Since transaction load is distributed over two links, during normal operation each link is running at half capacity, thus improving performance.

Issues

- System management is more complex. System-management tools have to run on one node and be able to access and change parameters on the other node. These tools must run in AB, AA, and BB modes as nodes fail and are restored.

- The 45-minute ATM downtime during a contingency switch is too long. This downtime should be measured in seconds – short enough so that a customer standing at the ATM is not inconvenienced.

- System sizing is complex. Performance data is taken on a single node on a quiet day (to minimize the inconvenience of ATM failover downtime). Performance data is also taken on peak days. It is desired to configure the nodes so that a single node can handle the peak-day load. The question that arises is how much load can a multi-CPU NonStop server handle and still perform well. Is it 60%? 80%? Is it dependent upon the NonStop server model? On the application architecture? This question has not yet been answered to the bank's satisfaction.

Future Plans

This ATM system is hardly static. The bank is considering several modifications and upgrades to the system. They include the following:

- The time to reconfigure the system from dual-node to single-node operation and back must be reduced. The correction of this problem requires modifications to the existing ATMs so that they can detect a fault and automatically connect to the other node by switching IP addresses. In this way, they can recover from a fault quickly. To return to dual-node operation, all that is required is for the current host node to drop the connection being used by the ATMs that are to be switched back to their home node.

- The bank feels that it must remove its dependence upon certain in-house file transfer and replication utilities. They are costly to maintain, and support staff tend to forget how they work. They plan to move these functions to DRNet.

- The bank is considering upgrading to Integrity NonStop servers in the next several years. When it does, it may reconsider its use of active/active. The bank's original motivation for active/active was to survive the high incidence of communication-line failures that were being experienced in the late 1980s/early 1990s. This is no longer a problem with today's IP networking infrastructure.

  The bank is considering moving to a "sizzling hot standby" configuration in which the nodes are running active/active with bidirectional replication; but all transaction activity is routed to only one node, with the other node acting as a backup. This configuration, the bank feels, will be much easier to manage while still providing the continuous availability and fast failover of active/active systems.

Summary

In today's financial environment, the availability of a bank's ATM network takes on new meaning. The public is so wary about the well-being of banks that the failure of an ATM network might be perceived as a warning that the bank may be about to fail. The result could well be a run on the bank, which might be a self-fulfilling prophecy – it could cause the bank to fail.

This bank moved to a highly-available active/active ATM and POS network many years ago; and it is well-experienced in the operation, benefits, and issues of the technology. Its active/active configuration has proven itself in the high availability that it has achieved. The bank has experienced no planned or unplanned major outages in its ATM network in over fifteen years of operation. This has been true even in the face of many upgrades, including several major hardware upgrades.

The bank has been a pioneer and a leader in the use of active/active technology and is well positioned to reap future benefits of this important experience.