Cisco Active Network Abstraction Gateway High Availability Solution

White Paper

This white paper describes the Cisco Active Network Abstraction (ANA) Gateway High Availability solution developed and implemented by Cisco Professional Services. High availability protection for the Cisco ANA Unit is not covered here but is fully described in the documentation available online at http://www.cisco.com/go/ana.

Introduction

Cisco[®] Active Network Abstraction (ANA) is a carrier-class network management foundation for service providers providing:

- Element management and mediation
- Network topology discovery and visualization
- Fault management, correlation, and troubleshooting
- Service support

Cisco ANA provides a rich set of applications with an easy-to-use GUI as well as well-defined APIs for operations support systems, helping enable carriers and service providers to respond efficiently to the constant market demand for new, reliable, and more sophisticated services, while hiding the complexity of large, multivendor, multitechnology networks.

Cisco ANA has a multilayer architecture. The first layer is composed of the gateway through which all the OSS/business support system (BSS) applications and clients access the system. The second layer is composed of the interconnected fabric of units, each managing a subset of the network elements in the network.

As the network infrastructure gets larger, the potential impact of network management downtime is greater. A high availability solution minimizes downtime and protects the revenue base.

Included in the product, Cisco ANA provides a clustered N+M high availability mechanism, within the Cisco ANA fabric only. Unit availability is established in the gateway, which runs a protection manager process that continuously monitors all the units in the network. Cisco ANA does not provide a preconfigured solution for high availability for the gateway. This document describes such a solution as implemented and provided by Cisco Professional Services.

Cisco ANA Gateway High Availability Solution

This solution covers the high availability mechanisms designed to handle the failure of the gateway server. Such failure includes hardware failures, operating system failures, and power failures.

The Cisco ANA Gateway High Availability solution is available in two different configurations: local and geographical redundancy.

The local redundancy configuration (Figure 1) uses two gateway servers in a cold standby configuration to provide an automatic failover for local faults without the need to reconfigure IP addresses on the switched/routed network. This solution uses the Veritas Volume Manager for volume management and Veritas Cluster Server (VCS) for clustering management.

Figure 1. Local Redundancy Configuration
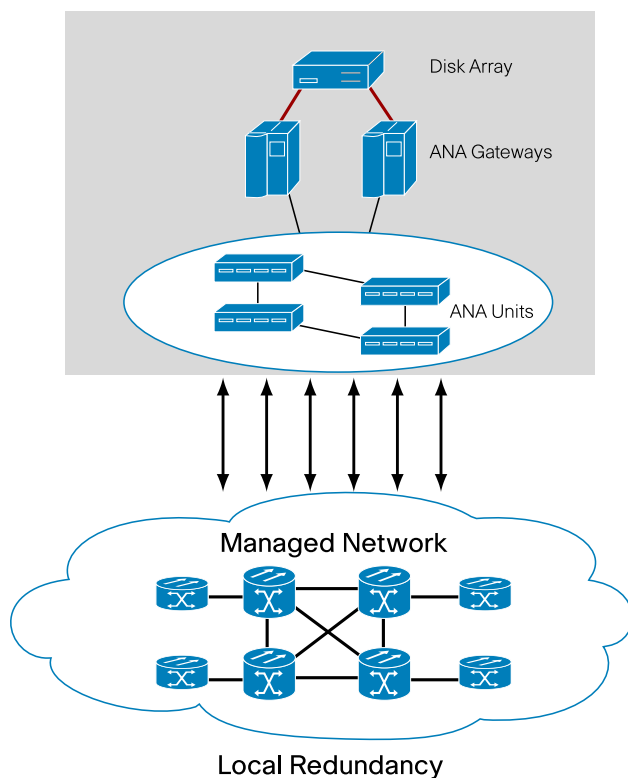
The geographical redundancy configuration (Figure 2) uses one or two gateway servers at each site for a full disaster recovery solution. When the local site is functional, the remote gateway at the secondary site is in standby mode. In the event of a local site failure, an operator must switch manually to the standby server. This requires a change to the gateway IP address, and clients must reconnect using the new address. This solution uses the Veritas Global Cluster option for failover management. Data replication can be done using the Veritas Volume Replicator software or using storage area network (SAN) solutions.

Figure 2. Geographical Redundancy Configuration
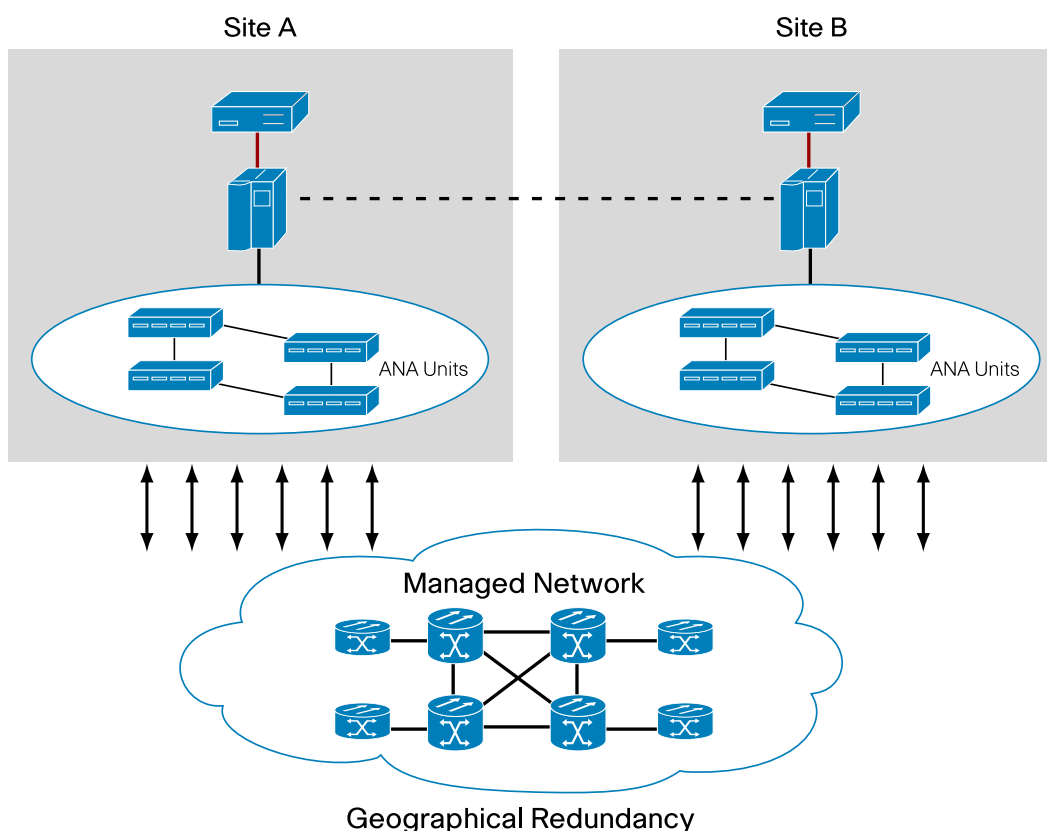
Split-Brain: In a geographical redundancy configuration, it is not recommend (even though it is technically possible) to use automatic failover to the standby gateway in order to avoid a "split-brain" scenario where each node in the cluster may mistakenly decide that the other node has gone down and attempt to start services that the other node is still running. This would lead to data corruption in the Oracle database from which it is difficult to recover.

Even though the above solution uses Veritas Software, Cisco is not an OEM or reseller for Veritas or for Oracle or Sun. The software licenses and support contracts for the Veritas, Oracle, and Sun components of the solution must be purchased separately from the respective vendors. The Cisco ANA product and the Cisco service to implement the high availability solution with all components must be purchased from Cisco Advanced Services.

Local Redundancy

Cisco ANA gateway local redundancy is implemented as a 1+1 cold standby in a dual-node cluster. One server, the primary gateway, hosts the ANA and Oracle processes, while the ANA and Oracle data are located on an external disk array connected to each server using redundant connections. The two servers maintain a heartbeat between them that allows the Veritas Cluster Server application on each server to monitor the health of the other server. VCS can detect faults in the ANA or Oracle processes and all their dependent components, including the operating system, network, and storage resources. When a failure is detected, VCS gracefully shuts down the application, restarts it on the available server, connects it to the appropriate storage device, and resumes normal operations.

Implementation details

Figure 3.   Local Redundancy

Active ANA Gateway   Standby ANA Gateway

External Disk Array

━━ Fiber - Primary Path
━━ Fiber - Redundant Path
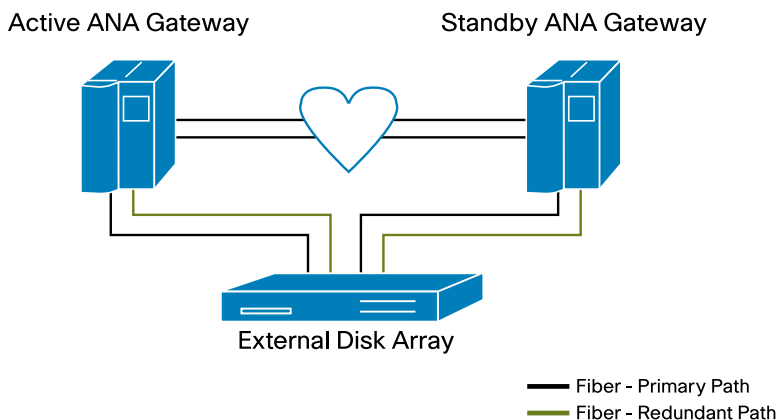
Hardware Requirements

- Two Sun servers. For a list of supported Sun servers for ANA, please refer to the Cisco ANA Installation Guide available at http://www.cisco.com/go/ana. It is recommended that both servers (primary and secondary) should be identical.

- Two power supplies for each Sun server with each power supply plugged in to a separate circuit

- Two internal mirrored boot/root disks in each server

- One Quad Ports FastEthernet or Gigabit Ethernet and one onboard network interface

- Two Fiber Channel (FC) connectors (or other connectors depending on the disk array) for each server

- One external disk array with two RAID controllers. The disk array must be included in the certified hardware compatibility list available from Symantec, the Veritas vendor (http://www.symantec.com).

Software Requirements

- One Cisco Active Network Abstraction Starter Kit Release 3.6

- One Oracle 10g, Enterprise Edition, Release 10.2.0.3, with Partitioning Option

- One Veritas Storage Foundation Standard High Availability, which consists of:

  ◦ Veritas Cluster Server Release 5.0 MP1 (Maintenance Pack 1)

  ◦ Veritas Volume Manager Release 5.0 MP1

  ◦ Veritas Cluster Server Oracle Agent Release 5.0 MP1

- Solaris 10 operating system and all recommended Solaris patches as required in the Cisco ANA Installation Guide

Server Connectivity

Figure 4.   Connections between the Primary and Secondary
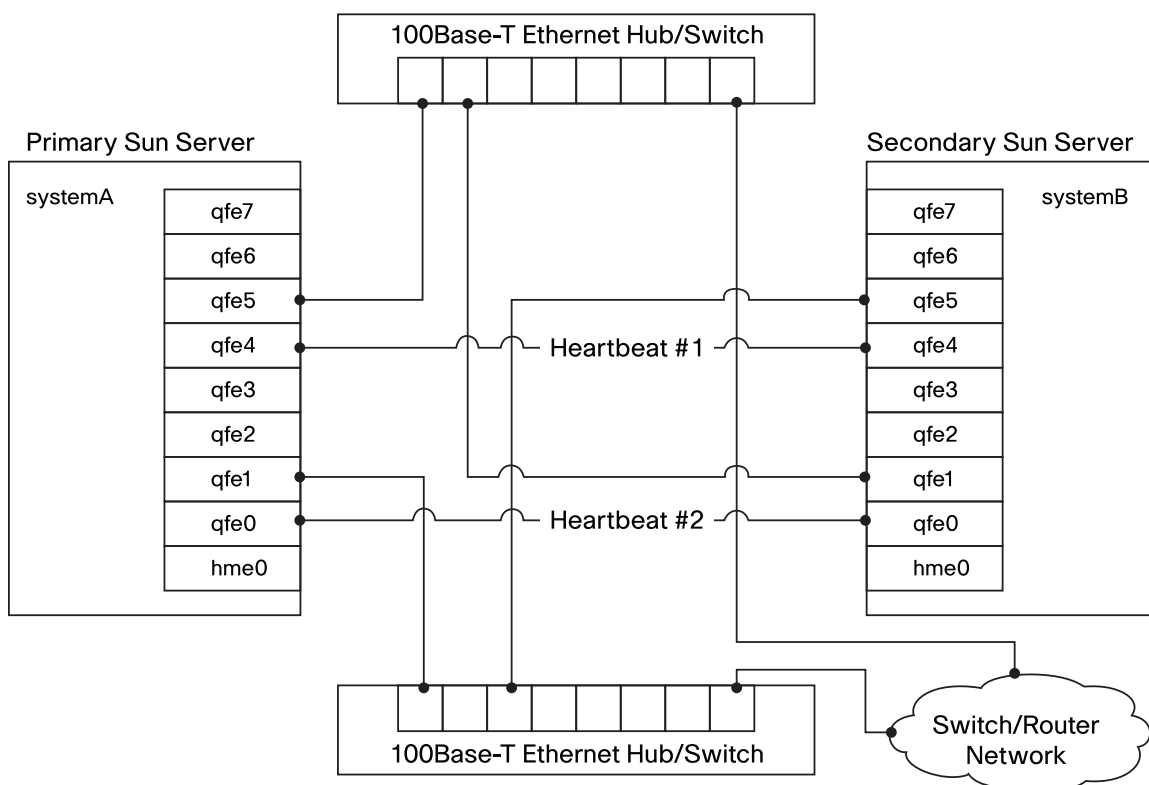
Cluster heartbeat is facilitated using redundant interfaces through direct connect crossover (reversing) Ethernet cables. If the connection between the servers fails, the heartbeat can also be sent across the LAN. Heartbeat messages run across a low latency transport (LLT) protocol, providing the high-speed, low-latency mechanism that Veritas uses for both local and geographical redundancies to help ensure that all nodes get status update information at the same time.

If using IP Multipathing (IPMP), the server's interfaces can be made available and plumbed up through the OS, requiring less VCS administration. VCS is used to assign a logical IP address to the dual-node cluster. All ANA clients (GUI and northbound interface), all units, and the Oracle application will use the logical IP address. In this way, failover can occur rapidly with the server's IP address being totally transparent.

Internal and External Storage

High availability local redundancy requires both internal and external storage for the two gateways. Internally, each gateway must contain at least two disks (the root "/" disks) containing the operating system plus Veritas. Those disks are mirrored (RAID 1) using Veritas Volume Manager. If both the primary and secondary gateways are equipped with four disks instead of two, a more reliable RAID 1+0 configuration can be used. (Note: The 1+0 configuration is mirrored sets in a striped set, different from 0+1, striped sets in a mirrored set.)

Externally, the Cisco ANA High Availability mechanism requires a single external storage unit containing a set of mirrored disks. If the disk array has RAID controllers (usually, it does), Veritas Volume Manager is only used to create one or more logical volumes since mirroring and striping are automatically done by the controller at a lower level. Multiple RAID configurations are possible, depending on the level of availability the customer wants to implement. Two suggested options are: RAID 5+0 (mirroring/parity + striping), using three or five disks, or RAID 1+0 (mirroring + striping), using two or four disks.

The final configuration is transparent to ANA.

As said initially in this section, internal disks only contain the operating system and Veritas. ANA and Oracle applications are installed on the external disk array on shared directories available for mounting on either of the two gateways.

Service Group and Critical Processes

Resources in the cluster are grouped together to form a logical and hierarchical object called a service group. In this proposed solution, we just have one service group (it is also possible to have two service groups, one for ANA and one for Oracle, to provide more flexibility). Its representation, using Veritas Resource Dependency View, would look similar to Figure 5.

Figure 5.   Solution with One Service Group
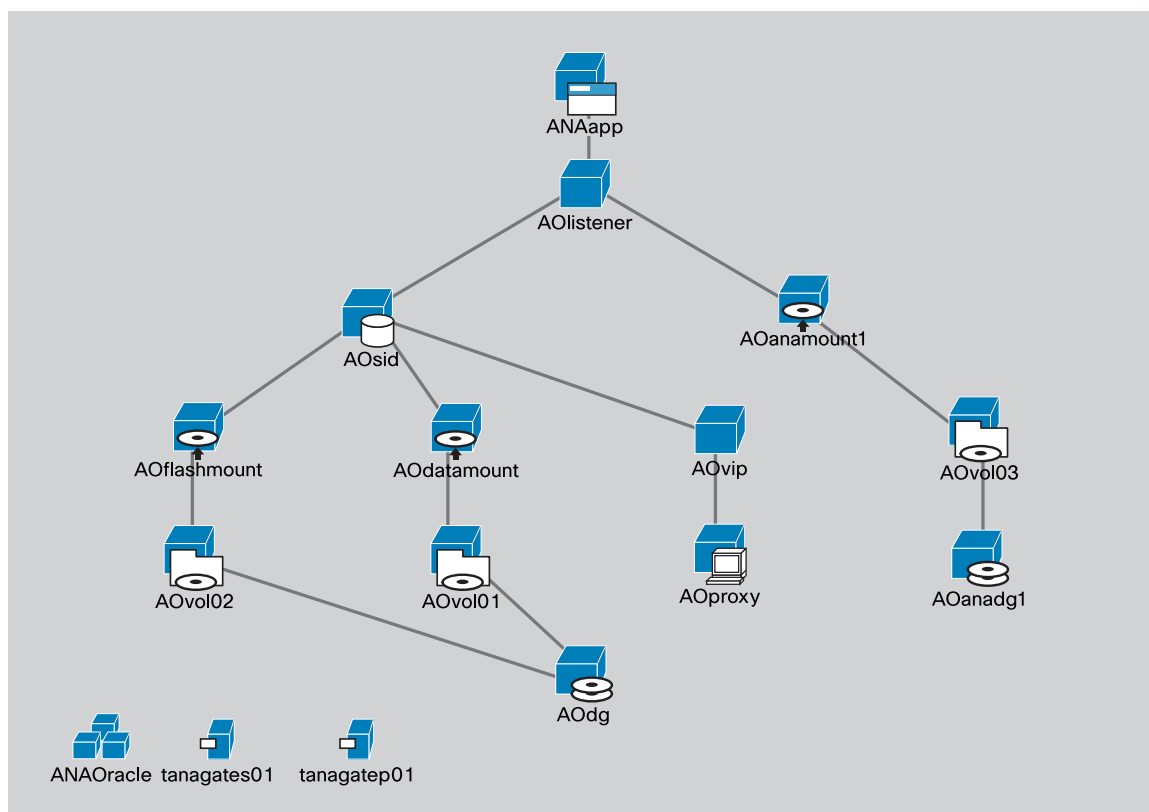
The lowest levels are more related to hardware such as network interface cards and disks; the higher levels are more relevant to ANA resources. In between, we have Oracle resources. When the system fails over (or the administrator performs a manual switch), VCS will start/stop processes using this structure. Since the ANA application depends directly on all the resources in its service group, all resources are designated critical and monitored by Veritas agents.

Note: While VCS is monitoring the ANA and Oracle applications, those applications should only be started and stopped using the Veritas Cluster Server GUI or Veritas command-line interface (CLI) commands, since using native commands would trigger an automatic failover.

Automatic Failover

The local cluster is configured to fail over locally automatically in the event of a failure in one of the critical resources. Thus, a failure in any of these resources (not including allowances for restarts) will cause the ANA service group to fail over. The failover process includes the shutdown (from the top of the tree down) of all the resources in the

service group on the current server and the startup (from the bottom of the tree up) of all the resources on the other server. As mentioned above, all ANA clients, all units, and the Oracle application will keep using the logical IP address. The Veritas agent (IPMultiNICB) will take care of the physical IP failover.

By default, all resources are polled every 60 seconds, so fault detection can take from 0–60 seconds. From the time a fault is noticed and the failover triggered, it takes on the order of 2 to 3 minutes until the ANA application begins to come up on the new active gateway. At that point, complete ANA gateway startup time for the Agent Virtual Machine (AVM) processes will vary depending on the network size. During this time, no alarms will be recorded.

If the two servers are identical, there would be no performance difference for ANA. The administrator will have to repair the fault that generated the failover and only then revert back manually to the original gateway, since the solution is not designed to do it automatically.

Warm Standby Option

In a standard cold standby high availability configuration, the secondary server would potentially wait a long time before it starts playing an active role. As the reader may remember, ANA comes with an option where the Oracle application (and data) can be installed on a separate server. By combining these two, we can have a warm standby option for local redundancy high availability where both servers are active at the same time. One is running only ANA processes, the other one only the Oracle application. Also in this case, both ANA and Oracle data are kept in the external disk array. The difference in this solution is that two volumes are used, one for Oracle and one for ANA. Each volume can be mounted independently on the server running the corresponding application.

When ANA or Oracle fails, VCS gracefully shuts down the application and starts it on the other server. The only difference is when Oracle is failing. In this case Veritas Cluster Server will also trigger an ANA restart after Oracle is up and running.

Using the second server to run either Oracle or ANA is also a good resource allocation benefit that makes business sense. Distributing the load across two servers not only is a good example of system resources allocation that improves overall solution performance, but also represents an immediate return on investment to those customers who decided to implement the high availability solution.

Geographical Redundancy

Cisco ANA gateway geographical redundancy extends the local cluster capability across two remote sites. On each site we have a cluster made up of one or optionally two servers (dual-node geographical redundancy). The VCS Global Cluster Option, an additional Veritas component required for this specific solution, combines those two clusters to form a global cluster and provide a seamlessly integrated disaster recovery solution. As opposed to local redundancy, in case of fault of critical resources, gateway failover isn't automatic and requires manual intervention from the administrator. On one hand, automatic failover would appear more attractive and would reduce system downtime, but on the other hand, manual intervention would avoid the risk of incurring split-brain conditions and consequent data corruption in the Oracle database. For complete geographical redundancy, redundant units are located at the remote site to be used in case the units at the local site are unavailable. A more cost-effective solution can be implemented using a single unit fabric.

Implementation Details

Figure 6. Geographically Redundant Implementation
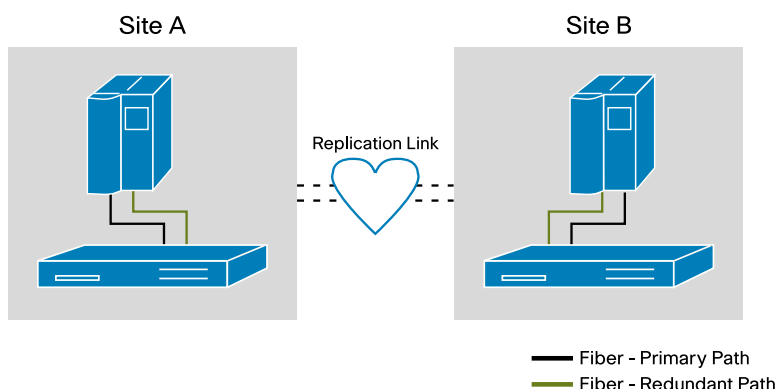
Hardware Requirements

- Two Sun servers for a single node cluster and four for a dual node cluster. An additional option can be implemented using 2+1 servers (described at the end of this section). For a list of supported Sun server for ANA, please refer to the Cisco ANA Installation Guide document available at http://www.cisco.com/go/ana. It is recommended that all servers should be identical
- Two power supplies for each Sun server with each power supply plugged in to a separate circuit
- Two internal mirrored boot/root disks in "each" server
- One Quad Ports FastEthernet or GigEthernet (or equivalent card with at least two FastEthernet or GigEthernet ports), plus one onboard network interface
- Two Fiber Channel (FC) connectors (or other connectors depending on the disk array) for each server
- Two external disk arrays with two RAID controllers. The disk arrays must be from the certified list available from Veritas.

Software Requirements

- One Cisco Active Network Abstraction Starter Kit Release 3.6
- One Oracle 10g, Enterprise Edition, Release 10.2.0.3, with Partitioning Option
- Two Veritas Storage Foundation Enterprise HA/DR, which consists of:
  - Veritas Cluster Server Release 5.0 MP1 (Maintenance Pack 1)
  - Veritas Volume Manager Release 5.0 MP1
  - Veritas Cluster Server Oracle Agent Release 5.0 MP1
- Two Veritas Volume Replicator 5.0 MP1
- Two Veritas Global Cluster Option 5.0 MP1
- Solaris 10 operating system and all recommended Solaris patches as required in the Cisco ANA Installation Guide

Data Replication

Data replication between the two sites can be implemented either through software using Veritas Volume Replicator (VVR) or by using a SAN architecture, which requires specific hardware supporting storage-based replication. We will discuss the first option only. The initial data synchronization can take a long time, since this consists of copying the entire contents of the two volumes. At run time, data replication should be done asynchronously using VVR over IP networks. This software includes controls to reduce the impact that replication can have on scarce network

resources. Through efficient volume-level replication based on actual application writes, VVR keeps WAN traffic to a minimum by replicating only the data that actually changes (incremental replication). VVR increases existing bandwidth efficiency through asynchronous replication. Data replication requires that the volumes being replicated be mounted only at one site at a time.

This white paper doesn't provide bandwidth requirement specifications since this changes based on the actual implementation. In any case, we suggest using VVR robust logging capabilities to model bandwidth requirements based on average application activity (rather than peak activity).

Manual Failover

Local and remote clusters constantly exchange LLT heartbeats over an IP network. In case of loss of heartbeat, Veritas Cluster Server registers the fault and awaits manual intervention for failover. In the event of a failure at the local site, the operator can verify that the local site is indeed down or that loss of heartbeat is due or a loss of network connectivity between the two sites before manually failing over to the remote site. Manual failover will also switch all the unit servers, having all AVMs loaded on the units on the other site. Each cluster is assigned with a different logical IP address; thus, any client application that uses an IP address will have to log in again using the new IP address. In case the hostname alias is used, Veritas Cluster Server can update a Domain Name System (DNS) server with the new address of the gateway, allowing client applications to reconnect without knowing the new IP address.

2+1 Gateway Option

Similarly to the warm standby option described in the local redundancy section, also in the geographical we can have a dual-node cluster on the primary site (while still using a single-node cluster on the secondary) with ANA and Oracle running on two servers to increase overall system performance. In this case, since the local cluster is a dual-node cluster, a local cluster failure occurs when a critical resource has failed on both servers.

2+2 Symmetrical Option

To complete the list of proposed options, more from an educational slant rather than suggesting it as an option, is the more expensive 2+2 gateway configuration where, on each site, two servers are deployed. This is indeed a possible and viable option that can implement a complete and symmetrical disaster recovery solution.

Conclusion

The Cisco ANA gateway high availability solution addresses the business needs for those customers who are interested in minimizing downtime, protect their revenue base, provide carrier-class reliability and reduce risk of lost management information. With the proposed scenarios above, customers have a better understanding on what is delivered with each option and can choose which one best fits their network needs. More technical implementation details can be found in the Cisco ANA Gateway high availability solution installation guide which can be found at http://www.cisco.com/en/US/products/ps6776/prod_installation_guides_list.html.

For More Information

For more information about Cisco Active Network Abstraction, visit http://www.cisco.com/go/ana, contact your local Cisco account representative, or send an e-mail to ask_ana_pm@external.cisco.com.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV
Amsterdam, The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

Printed in USA

C11-539603-00   06/09