SUN BLUEPRINTS APPROVED

Clustering LDAP Directory Servers

By Tom Bialaski - Enterprise Engineering

Sun BluePrints™ OnLine - December 1999

Sun microsystems

ENTERPRISE
ENGINEERING

`http://www.sun.com/blueprints`

Sun Microsystems, Inc.
901 San Antonio Road
Palo Alto, CA 94303 USA
650 960-1300 fax 650 969-9131

Part No.: 806-3968-10
Revision 01, December 1999

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REPONDRE A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.

Clustering LDAP Directory Servers

In my previous article, we looked at several factors that are key to making your directory service highly available. The directory services, NIS, NIS+, and LDAP, which were discussed in that article, all had built-in replication mechanisms that can increase the availability of these services. However, replication provides a read-only duplicate of the directory and cannot be used to directly update directory information.

This article will explore deployment of Sun™ Cluster software to create an environment for LDAP based directory services that are highly available for both read and write access. An overview of the Sun Cluster Data Services framework is presented along with a detailed description of HA LDAP Data Services for the Sun Cluster software.

Why Use Clusters?

The decision whether or not to deploy clustering technology to increase the availability of your Lightweight Directory Access Protocol (LDAP) directory service largely depends on what the impact of losing your master server is. If information is updated infrequently and the ability to update this information in a timely manner is not mission critical, then a temporary outage of the master directory server is probably acceptable.

An example of a non-mission critical deployment of a directory service would be a company employee address book. Since this information only changes when employees join or leave the company or are transferred to another location or department, changes are predictable and do not have to occur in real time. If updates cannot be made for a few hours or even a day, the impact on the company's business is minimum.

However, if the company is an Internet Server Provider (ISP) and is in the business of providing email accounts to thousands of web surfers, an outage of even a few minutes could result in loss of potential customers. If Internet users cannot register for an email service while they are online connected to the ISP's web server, they are likely to go to a competitor, who can create an account for them.

Manual failover can be deployed as an alternative to clusters, but generally requires that a skilled operator be present to execute the failover scripts and take action if anything goes wrong. While this may be an acceptable approach in some environments, the 24x7 nature of some businesses makes this impractical. A safer approach is to take advantage of the automatic failure detection and recovery provided by clustering technology such as Sun™ Cluster software.

Overview of Sun™ Cluster Software

In general, clusters can perform in two roles: clustering for high availability, and clustering for performance. While Sun Cluster software can be used to increase performance of some applications such as database servers, its primary benefit to an LDAP directory service is increased availability. In this role, Sun Cluster software is configured in a *shared nothing* architecture. What this means is that at any time a *resource* is owned by only one of the cluster nodes. If that node fails, then the ownership of its resources is transferred to a working node.

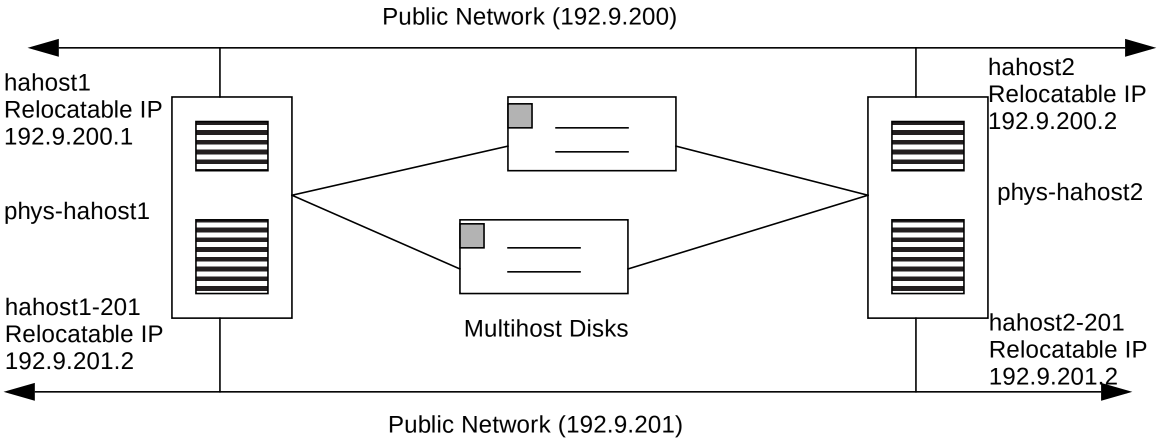
Sophisticated clustering software such as Sun Cluster softwareneeds to perform many housekeeping functions such as cluster creation and cluster reconfiguration. While these are critical functions, they are not unique to LDAP directory services, so they won't be discussed here. Instead, we will focus on what the Sun Cluster framework provides to support highly available LDAP services.

Virtual IP Addresses

The basic concept behind clustering for High Availability (HA) is the notion of a floating Internet Protocol (IP) address which is temporarily assigned to a server providing a service such as an LDAP directory server. Clients of this service use the virtual IP address to access the server, rather than the server's actual IP address. In the event of a cluster node failure, the virtual IP address associated with the LDAP data service is transferred to a working node.

The nice thing about referencing services by their virtual IP address rather than the actual one, is that the client doesn't need to care about what physical server is actually providing the service. For example, an LDAP client, such as a messaging application would always reference the LDAP server by its virtual IP address and would be unaware if this IP address was transferred to another cluster node.

The following diagram shows what a typical Sun Cluster HA configuration might look like. Two public networks are used for redundancy and relocatable, or virtual addresses, are used by clients accessing services on the servers.



Data Services for Sun™ Cluster Software

To run a particular application in a Sun Cluster environment, two software packages are required:

- Sun Cluster core components
- HA Data Services (for that application)

The Sun Cluster software core components provide the necessary framework for running the application-specific software, or *data services*. The core components are responsible for maintaining a cluster configuration database and other services such as a heartbeat signal between cluster nodes. These core components must be configured before the data services can be installed and configured.

The data services provides functions specific to a particular service or application. For example, the Sun Cluster HA for Netscape LDAP data service is used to manage the failover of the iPlanet™ (formerly Netscape) Directory Server. This service is added as a Solaris™ Operating Environment package after the core Sun Cluster software components are installed.

The HA Data Service for LDAP performs the following functions:

- Monitors the health of the directory service
- Controls the stopping and starting of the directory service
- Attempts to restart a failed directory server on the same node
- Stops the directory server on one cluster node, restarts it on another

Note – Once the HA Netscape LDAP data service is configured with the Sun Cluster software, the directory server is no longer started or stopped manually, because the data service will provide those functions.

Building a Sun™ Cluster with HA LDAP Data Services

While it is possible to build and configure a Sun Cluster software without professional assistance, I strongly recommend contracting the services of a consultant with Sun Cluster software experience. Since the reason you are deploying a Sun Cluster is to increase service availability it is well worth the investment to have an expert perform the initial set up. The following paragraphs outline some of the basic steps required to give you an idea of what is involved.

Building a Sun Cluster software starts with identifying a Sun supported hardware configuration. This configuration usually consists of two (or more) similar servers of which both are physically connected to one or more *multihomed* disk storage devices. The storage devices are used to contain information such as the directory tree and log files. Since the log files are usually kept on a different disk drive than the directory tree, several physical disk drives are usually deployed.

The multihomed storage devices are physically accessible by all the cluster nodes, but only one cluster node has control over it at any time. If the LDAP directory service fails, the control of the multihomed storage devices is transferred to another cluster node.

Another important consideration is the communication channels between the nodes. Separate channels are typically used for cluster communication, the *heartbeat* signal, and client access to the directory service itself.

The LDAP Fault Monitor

The HA LDAP data service contains a fault probe that periodically checks to see if the directory service is functioning properly. The fault probe is run on both the active node and the standby or remote node. On the active node, an `ldapsearch` command is executed on the node running the LDAP service, then the probe waits for the successful completion. On the standby or remote node, the probe attempts to *telnet* into the port (default is 389) on the node that the LDAP server is running on. If the attempted telnet session times out, then the remote node is assumed to have failed.

iPlanet™ Directory Server 4.1 Installation

No special version of the iPlanet Directory Server software is required to run in the Sun Cluster environment. However, you must enter specific configuration parameters and install components in the proper order. The following steps summarize the process.

1. Load the core Sun Cluster components
2. Install the HA Data Services for LDAP package **`SUNWscns1`**
3. Install the HA Data Services Update, Patch-ID# **108109-01**
4. Install the iPlanet Directory Server 4.1 software on each cluster node.
5. Run the Sun Cluster configuration command: `hadsconfig(1M).`

When installing the iPlanet directory Server 4.1 software as noted in Step 4, you must change the following default parameters:

1. Specify the *logical* host name instead of the *physical* host name of the server where the software is being installed.
2. Change the default server root directory: `/usr/netscape/server4,` should be changed to a directory which resides on the multihost disk.
3. Change the base install directory pathname should be the location of where the start and start scripts, `start-slapd` and `stop-slapd`, reside.

These parameters are also used for input to the `hadsconfig` command. In addition to these parameters, there are some that pertain to how you want the HA LDAP Data Service to operate. These parameters include the following:

- Name of the instance—Multiple instances of the iPlanet Directory Server can be run simultaneously in the Sun Cluster environment. For administrative purposes, assign each instance a unique name tag.

- Takover flag—This parameter specifies whether you want to failover this instance of the directory service to another cluster node. In most cases you want this set to `y` for `yes`.
- Probe interval—This is the time between fault probes. Since the fault probe executes a `ldapsearch` command, an additional load is placed on the directory server and network. The trade-off is between adding the additional load versus the time it takes for the Sun Cluster software to recognize there is a problem. The default here is 60 seconds.
- Probe time-out—This is the time after which the fault probe will time out. Setting this value too low may cause a false failover trigger, so you must take care to set this value correctly. The default here is 30 seconds.

Obviously there is more to managing a Sun Cluster environment than just the initial set up as described here. When a cluster node fails it must be removed from the cluster, repaired, then will need to rejoin the cluster. These mechanisms are defined in the Sun Cluster documentation and instructions specific to your deployment are specified in your data center *run* book.

LDAP Cluster Deployment Options

Once you have made the decision to deploy your LDAP directory server in a Sun Cluster environment, you will need to decide what deployment model is right for you. You can either use the hot standby model or choose to run services on all cluster nodes.

Hot Standby Model

In this configuration, two identical servers are configured and sized with enough capacity to effectively handle the load. One server in the cluster is active all the time while the other server is idle. If the primary server fails, then the secondary server takes over.

The advantage of this model is that there is no degradation in server performance in the event of a failure. The disadvantage, of course, is that the resources on the standby server are not performing productive work. However, a manual switch over can be initiated so that periodical maintenance can be performed on a server without affecting service.

Active Server Model

In this model all the cluster nodes have active services running on them. However, because you cannot have two master LDAP directory servers active at the same time, you either have to run two different instances of the LDAP server or run another type of service.

For example, you could have two separate directory trees, each with its own master server. One of the master servers could be active on each cluster node. In the event of a cluster node failure, the LDAP server on that node would failover to the healthy node. Similarly, you could have an LDAP server active on one node and a web server active on another node. If either service failed then it would fail over to the other node.

The problem with this model is that running the additional service or services on the second node requires additional system resources. If sufficient resources are not available on the node where the services fail over to, then the performance of all services will be degraded.

More predictable results can be obtained on the surviving cluster node by running Solaris Resource Manager™ (SRM) software. While this won’t increase the amount of available resources, a minimum threshold of performance can be maintained. SRM software assigns shares of system resources to different applications. If applications are not currently using their allotted shares, then they can be used by the active applications. This prevents resources from being wasted if they are not currently being used.

Another factor to consider when deploying this model is the effect of reduced performance and the amount of time it would take to repair the failed server.

Conclusion

Deciding whether to deploy your LDAP server in a cluster environment depends on what it is being used for. If updates are performed frequently in an *ad hoc* fashion and the ability to service updates quickly is critical, then a cluster environment would make sense.

LDAP replication is not a replacement for clustering, but can be used in conjunction with clustered servers. Replication provides read only access, but is very effective in providing load balancing of read requests. By deploying a combination of both clustered LDAP master servers and replication servers, both fast read access and high availability for updates can be achieved.

Author’s Bio: Tom Bialaski

Tom has nearly 20 years of experience with the UNIX operating system and has been a Sun engineer since 1984. He is currently a staff engineer on the Sun Blueprints team and is the author of "Solaris Guide for Windows NT Administrators".