

eSubnet Fragment Article

High Availability: Technologies

Network availability directly impacts productivity. This article covers a range of topics focused on improved uptime: the process, of selecting the right technology, and the tools to bring your network to a practical state of always up. This is the first in a series of articles on High Availability (HA).

Understanding Uptime

Sometimes management requires that services and data are always available no matter what. They are looking at the cost of downtime without considering the cost of uptime. The pinnacle of uptime, referred to as the “five-nines” is represented mathematically as 99.999% or making your network unavailable only 0.0001% of the time. This should be quantified: does value measure workdays, a 24 hour day, and is it monthly or annual. The table below shows how much time this is, in minutes.

Table 1: Uptime percentage in time

Uptime Percentage	Availability				Down Time
	24 hour day	Weekly	Monthly (4 weeks)	Annually (12 months)	
99	1425.6	9,979.2000	39,916.8000	479,001.6000	3.5 days
99.9	1438.56	10,069.9200	40,279.6800	483,356.1600	8 hrs
99.99	1439.856	10,078.9920	40,315.9680	483,791.6160	under 1 hour
99.999	1439.9856	10,079.8992	40,319.5968	483,835.1616	under 5 min
Total minutes	1,440	10,080	40,320	483,840	

According the table above, delivering five-nines over the course of a year equates to 4.8384 minutes, just less than five minutes of downtime, where as three-nines equates to a day.

The HA Principle: N+1

In a high availability environment, widgets will break. You need to have a spare. Count the number of widgets you need and buy one more widget. This is the N+1 principle of high availability; it applies to all parts of the organization’s IT infrastructure, and perhaps most often seen with servers.

Most parts of the infrastructure can be made redundant as well. Hard drives are given redundancy through RAID controllers, power supplies are removable and servers can typically use more than one. Having more than one source of power is also important, as when the server has the option for two power supplies, plugging them into the same power source defeats the purpose. I would strongly suggest having two Uninterruptable Power Supplies (UPS) connected to two separately supplied power sources; never plug the server’s two power supplies into the same phase, or power source. True Enterprise servers further allow CPUs to be swapped out without shutting down the server.

Following this line with your network means you have to look at all its parts. You need two of everything; core routers, firewalls, links to distribution switches, load balancers, links to Internet providers, remote locations, and so on.

The only area where it is nearly impossible to have an active spare is with your distribution switches unless you have two network drops and two NICs in every PC. For distribution switches it is more cost-effective have a spare switch on hand. Just be sure to backup the configuration files regularly.

Redundant Components (enterprise versus residential)

I often hear SMB organizations say that they can’t afford to be like larger companies and therefore buy and use residential-grade equipment. But they can’t afford to do this, and need to plan like enterprises do. SMBs still look to deliver products and services competitively with the larger players, and should act accordingly, even in IT infrastructure. IT vendors know that the SMB market is huge overall, and supply the support and products needed.

Using residential-grade products has significant drawbacks: there is less memory, CPUs are slower and redundancy capacity is usually lacking. These lacks make the devices affordable for home users. But when the business’ competitive success is on the line, you can’t rely on those the same devices.

There are some concessions to be had by smaller companies. For example, large companies need larger internet pipes. Back-up solution and storage capacity needs are also greater. Smaller company shouldn’t relax when it comes to availability, reliability and redundancy. But they don’t need to purchase the same capacity.

Redundant Paths in Layer-1 through Layer-3

For maximum network redundancy the lower 3 layers of the OSI model must have duplication.

Layer-1

The physical connections in Layer-1 between all network components should be installed and wherever possible not in parallel. Often referred to as diverse path, this applies to all links and devices. When selecting your multiple Internet providers, ensure that they use a different means of access from each other on to the premises. For distribution switches, having diverse path is a simple way to safe guard connectivity in the event someone accidentally breaks the primary path. The second link can also save from failed hardware.

Layer-2

Layer-2 communication has traditionally been managed by spanning tree protocol (IEEE 802.1D), or STP. STP provides loop free connectivity in a meshed switched networking environment by blocking ports which have a longer path to the root bridge. STP will automatically calculate the number of hops to the root bridge. For STP the root bridge is considered the centre of the network. Using its own algorithm, STP will calculate the root bridge by looking first for the switch with the lowest bridge priority, where the default value is 32,768. In the event that more then one switch has the lowest priority, the switch with the lowest MAC address is assigned as the root bridge. To ensure proper functionality within the network it is essential that network administrators manually identify the appropriate switch as the root bridge.

Other variations on spanning tree have been developed to make improvements to the protocol since its release in 1985. These include Rapid-STP (IEEE 802.1w) introduced in 1998, and Cisco Per VLAN STP (PVST) which provides granular control over Cisco's proprietary ISL VLANs and PVST+ for use with the standards based IEEE 802.1q VLANs. Regardless of which spanning tree protocol is used, any delays in failover, coupled with the fact that one link is rendered useless, are unacceptable for certain environments and an alternative solution is required.

Port Channel Groups are the alternative to spanning tree. On Cisco based equipment a port channel group is a collection of interfaces which are bundled together and behave as a single interface. The connecting cables can take diverse path between switches. In the event that one of the links suffers a failure, connectivity continues on without delay, with only a reduction in the available bandwidth.

Technologies are now available which link two physical switches into one logical switch. This logical device, or Virtual Switching System (VSS), allows two or more ports on two such linked switches to act as a single logical port with a total throughput rate of all ports combined. In the event that one of the linked switches suffers a failure or if a physical connection fails, connectivity continues on without delay, resulting in only a reduction in the available bandwidth.

Layer-3

Layer-3 redundancy requires a look at the routers and connections in use along the network path.

Hot Standby Routing Protocol (HSRP) can be configured on two devices to ensure that the Layer-3 address is always available. Each device is given a unique Layer-3 address and a third 'floater' address is shared between the devices. At any given moment, the floater address is owned by one router. One of the routers is configured to be the primary. If the primary device fails, the floater address is assumed by the secondary device and traffic continues to flow.

Layer-3 redundancy in the network path is significant for operations with remote locations. These are extended or wide area networks (WANs) such as campus wide networks, city wide networks and larger. A detailed explanation of this subject is enough to fill a book and so I'm going write about generalities.

To achieve Layer-3 resiliency in a WAN, a dynamic routing protocol is required. Dynamic routing protocols learn about the network from neighbor routers and select the best path for traffic. There are two main flavors of dynamic routing protocols, hop-count and link-state. They are both well named. The simpler is hop-count routing, the classic of which is Routing Information Protocol (RIP). With RIP enabled, the routers make routing decisions by selecting the path with the least number of routers between point A and B. This determination is done with complete disregard for link speed or congestion. This could result, for instance, in the traffic flowing through two 10 Mbps links instead of through three 100 Mbps links. The alternative is to use a link-state protocol. These assign a value to each link based on speed, where higher speed equates to a lower value. They route through the path with the lower cost or metric. In the previous example, the two 10 Mbps links are used for failover only, while the primary traffic flows through the three 100 Mbps links.

Beyond Layer-3

There are a number of ways to ensure such demanded services are available when accessed. For instance, we could deploy more servers. The trick is to distribute the users' access on to those servers in a way that makes sense.

One of the earliest means was by using the domain name system (DNS). By listing the same server name, but with different IP addresses in the DNS zone file (see example below), users are sent to more than one server. If we had three such entries in our zone file then the first request is responded to with the first address, the second request with the second address, and so on until it wrapped back to the first. This is round-robin DNSing. However, the DNS server could still send users to a server that was down. In this case, every third request would break.

Sample zone file entries

```
www A 192.168.100.101
www A 192.168.100.102
www A 192.168.100.103
```

A better means of controlling connectivity to applications is to deploy a load balancer. A load balancer is part server, part client, part router and optionally part firewall. The load balancer accepts the request from the end user and passes it on to a server. The load balancer keeps track of connections to the servers and will take a server out of the pool if it stops responding. Load balancers with advanced features can be configured to do much more than just pass traffic, such as:

- They can act as an SSL front end, thus reducing the load on the servers allowing them to serve more connections;
- They can share TCP sessions for multiple external clients to a single server. Without the need to repeatedly perform the TCP handshake the servers are freed to serve more pages.

Truly advanced models such as the Citrix NETScaler – Enterprise Edition can also inspect HTTP/HTTPS traffic to validate content, thereby increasing security.

As more of our daily application-based tasks move online and with our ever-increasing reliance on them, the deployment of load balancers will also increase. Such high availability improvements will become part of the base cost of doing business competitively.

Conclusion

There is no doubt that more and better uptime is good business. To achieve this, you aren't only adding more components to your network; you're also adding complexity; which requires more care, more advanced skill sets and typically better documentation. A full migration to the coveted five-nines might be within your budget when your business requires that level of high availability.

Originally published Jan, 2010

About the Author

Richard Danielli is the founder and President of eSubnet Enterprises. He has broad expertise in the fields of networking and data security. Based in Toronto, eSubnet provides superior customized solutions for networking and data security. Additional articles are available at <http://www.esubnet.com/fragment.html>.

Reprints

1. Commercial use: Commercial newspapers, magazines and websites that wish to reprint an article should contact eSubnet for permission.
2. Non-Commercial use: Any reproduction of content (excerpt or full text) by nonprofit organizations or companies wishing to share information with their members or clients requires that you notify us, whether for print or online usage.

Regarding reprints, contact Mr. Danielli at [rdanielli\(at\)esubnet.com](mailto:rdanielli@esubnet.com) or at (416) 203-1223.