

Carrier-Class Availability for Enterprises

As enterprises go borderless and move towards IP communications and mission-critical applications, there is a strong need for carrier-class availability. In converged environments, this is especially important, as in order to carry voice traffic and real-time traffic, the network infrastructure needs to deliver the same level of availability as the public switched telephone network. While most high availability networks are focused on the reliability of the hardware devices, Alcatel has designed its next generation switching platform – the OmniSwitch 7000 series – on the basis that availability needs to be extended throughout the enterprise to ensure the an uninterrupted flow of information to users and resources.

CARRIER CLASS AVAILABILITY FOR ENTERPRISES

As enterprises go borderless and move towards IP communications and mission-critical applications, there is a strong need for carrier-class availability.

Introduction

The trends towards convergence of voice and data on a single network infrastructure and IP mission-critical networking are gaining momentum, making it necessary for enterprise networks to deliver carrier-class availability. In today's enterprise networks, downtime can be very expensive, with significant costs that need to be evaluated when designing a high availability network. It can also be more than money that is lost; it can be customers, business transactions and time-critical communications. In converged environments, this is especially important, as in order to carry voice and real-time traffic, the network infrastructure must deliver the same level of availability as the public switched telephone network.

While hardware failures tend to capture the most attention, according to analysts as much as 80% of downtime is caused by people, processes and configuration errors. Thus to achieve the highest possible availability, suitable operational processes must be in place for all facets of the network: hardware, software, applications, security, networking, server farms and backup systems. Network management and security must be designed to maintain availability, taking advantage of automated features to reduce human error, prevent problems and minimize the associated downtime.

Alcatel's view on network availability is that it needs to go beyond the devices and basics of hardware redundancy. Availability must be built into the infrastructure and extend throughout the network links and paths to ensure that all computing resources, applications and services are readily available to users at all times. To maximize interoperability across the enterprise, the network must also comply with industry standards to enable businesses to make full use of their investment in existing equipment and applications.

Alcatel has taken the lead in delivering carrier-class availability to the enterprise. The OmniSwitch 7000 incorporates a wide range of features, such as smart continuous switching, wirespeed performance, native server load balancing and firewall clustering, all of which are crucial aspects of availability as they help to ensure that mission-critical applications and network security are "always on".

OmniSwitch 7000

The OmniSwitch 7000 was built from the ground up to facilitate convergence in enterprises and Ethernet Metropolitan Area Network (eMAN) implementations with wirespeed Quality of Service (QoS) so that voice and mission critical traffic could be given priority. Offering high performance and carrier-class features, OmniSwitch 7000 is a next generation multi-gigabit Ethernet switch that also offers enhanced network management with a common interface for managing both voice and data networks. In-line power feeding is built into the switch, eliminating the need to provide power sources for Internet Protocol (IP) phones and devices.

Carrier-Class Availability

The term carrier-class availability, which comes from the telecommunications world, is most often associated with 99.999% availability. When applying the concept of carrier-class availability to a data network, it is important to take a closer look at how it is defined and also at the design of the network.

When designing the network and building the infrastructure, many areas must be addressed to achieve carrier-class availability. At the device level, all system components, including the management processor and fabric, must be redundant, hot swappable and capable of failovers that are fast and transparent to users. At the network level, there is a need for topological redundancy and resilience in the network links to ensure that no single point of failure exists that could disrupt the entire network.

Redundancy is the key in carrier-class networks. It involves having duplicate components, typically running in parallel, that provide backup in the event of a failure. Hot swapping refers to the action of adding, removing or replacing a component while the switch is operating. Load sharing allows multiple components to run at the same time, with the intelligence to determine which components are available and algorithms that determine how the load is to be shared.

Additionally, the device must be designed according to established physical standards for carrier-grade

equipment, which includes being certified for compliance with the New Equipment Building System (NEBS). This is a strict set of standards with extensive testing; it was developed by Bellcore to meet performance, quality, environmental and safety requirements in central offices and carrier environments.

So, how is 99.999% availability measured and what does it really reflect?

99.999% Availability

The multiple “nines” have been used for many years in the telecom industry to describe availability. Carrier-class availability is generally equated with “five nines”, or 99.999% availability, which amounts to about five minutes downtime a year, or just under one second per day. Typically, this measurement does not include any “planned downtime” for scheduled maintenance and system upgrades. *Table 1* shows the multiple “nines” availability and the associated downtimes.

Perhaps one of the main misconceptions relates to

Availability	Downtime per year
99.9999%	32 sec
99.999%	5 min 15 sec
99.99%	52 min 36 sec
99.9%	8 hr 46 min
99%	3 days 15 hr 40 min

what components are included in the computation of network availability. Surprisingly, in the voice world, 99.999% only applies to a few components in the Private Branch Exchange (PBX); it does not extend to the phones or user devices. For example, in a PBX it typically applies only to the central processor, power supplies and switch matrix; it does not usually include the line cards, electrical power supply, software, operating system or planned downtime (maintenance, upgrades, fixes). The components in a data networking device are not yet as clearly defined, so while the concept is often applied, it is difficult to make any direct comparisons.

The most common way to increase reliability is to introduce redundancy, which involves a cost trade-off. This generally means having redundant components running in parallel, for example, “hot standby” components. Reliability can be enhanced by adding components in parallel. As an example, if one component has an availability of 99% and the aim is to achieve 99.999%, two redundant components need to be added so that there are three operating in parallel.

Serial or sequential components, such as cabling or

network links where only one is operating at a time, are different as adding further serial components actually decreases reliability. For example, if three components each with a 99% reliability are connected in series, the reliability is reduced as the failure of one component will impact the other two. To increase the reliability of serial components, redundancy or parallel paths need to be deployed.

Measuring Availability

While the terms *availability* and *reliability* are often used interchangeably, it is important to note that availability is usually defined as network uptime and the ability of a network to deliver continuous operation without interruption to users. Reliability is a part of availability; it is a measurement based on the life expectancy or failure rate of the components. There are two parts to the measurement of availability: the Mean Time Between Failures (MTBF) and the Mean Time To Repair (MTTR). The equation to measure availability is as follows:

$$\frac{MTBF}{MTBF+MTTR} = (.9xxxx)$$

MTBF refers to reliability and the average time between failures. As the reliability of a component usually decreases over time, replacing components before it is anticipated that they will fail can improve this metric and the overall availability.

MTTR is the average time needed to repair (or restore) the system and bring it back into service or operation. In the case of redundant configurations, this includes how quickly failovers, hot swapping and other fault tolerance mechanisms operate in order to minimize any disruption to operation.

Device Level Availability

The reliability of components at the device level is critical for achieving high availability. Carrier-class switching platforms, such as the OmniSwitch 7000, implement redundancy, hot swapping and load sharing across multiple components to reduce downtime caused by component failures.

Chassis Management Modules

Chassis Management Modules (CMM) are the management / supervisory modules that are critical to the entire operation of the switch. They are typically designed to run in redundant configurations with one CMM having the primary role and the other a secondary role. The primary manages the current switch operations, while the secondary runs in parallel, serving as a “hot standby”. In the event of a failure, the secondary CMM takes over. What is critical to availability is how fast this failover occurs; in many cases it can result in all connections being dropped. In the OmniSwitch 7000, both

CMMs are synchronized at all times, and a technology called “smart continuous switching” is used to ensure that there is no interruption to users or to the flow of data.

Smart Continuous Switching

In the case of smart continuous switching, all source learning, spanning tree functions and established routes are distributed throughout the network interface modules instead of using a central engine. In the event of a management module failure, the system automatically switches over to the hot standby with no loss of connections or switch fabric capacity. Established layer 2/layer 3 traffic, including voice conversations, continues without interruption. Furthermore, smart continuous switching enables new connections to be made, ensuring that all users have access to network resources – an industry first for the enterprise.

Network Interface Modules

Network interface (NI) modules are the interface cards that provide the connectivity for ports and switching intelligence. In the OmniSwitch 7000, these include fast Ethernet NIs (ENI) and Gigabit Ethernet NIs (GNI); each interface has its own CPU for distributing intelligence across the switch. The NI modules are incorporated with additional technologies, including IEEE 802.3ad dynamic link aggregation, server load balancing and IEEE 802.1w rapid reconfiguration, to provide a fail-safe, scalable and flexible environment. These technologies may be combined to maximize throughput and availability for mission-critical applications and server farms.

Network Level Availability

At the network level, redundancy and resilience are important elements in realizing high network availability. The OmniSwitch 7000 offers advanced routing, load balancing and link aggregation, as well as mechanisms for fast rerouting and reconfiguration of links between switches, servers and other network devices.

Spanning Tree / Rapid Reconfiguration

The Spanning Tree algorithm and Protocol (STP) is a self-configuring algorithm that provides data path redundancy, while ensuring there is only one data path between any two switches. The STP allows IEEE 802.1Q tagged ports and link aggregate logical ports to participate in the calculation of the STP topology. Multiple paths can be provided between two or more switches, but only one path can be active at any one time. The remaining paths are placed in a “blocking” mode. If a primary path fails, an alternative data path is brought out of the blocking mode into a forwarding state, thereby re-establishing the connection between switches. The transition takes between 30 and 50 seconds for the IEEE 802.1d STP.

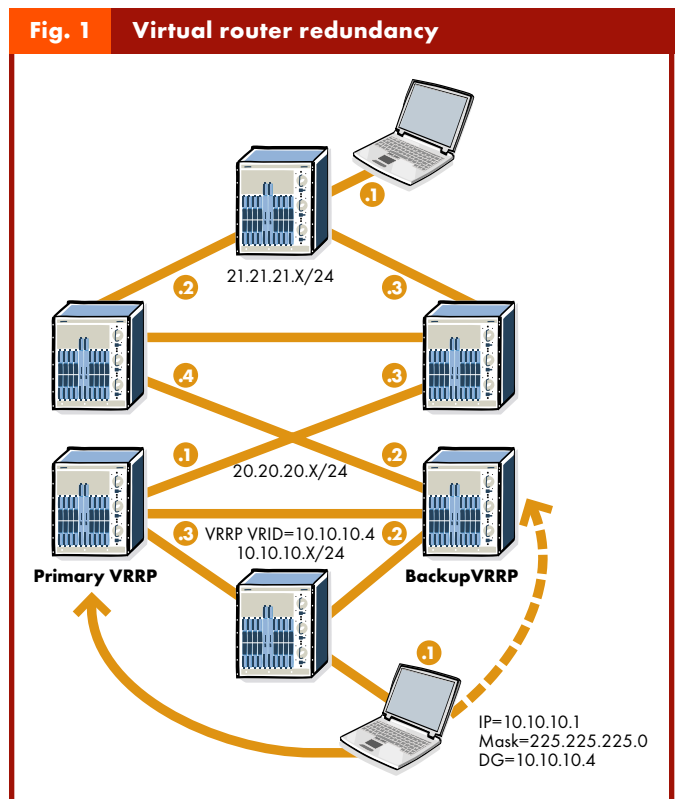
The Rapid Reconfiguration Protocol (RRP) integrated

into the OmniSwitch 7000 provides even faster recovery from a link outage. The ability of the RRP to make a quick transition from discarding to forwarding is based on calculating alternate and backup paths during spanning tree convergence. The port transition has been enhanced with the IEEE 802.1w amendment to the 802.1d STP, enabling the RRP to transition from a blocking (discarding) state to a forwarding state in under one second.

Virtual Router Redundancy

The Virtual Router Redundancy Protocol (VRRP) is a standard protocol that provides topological redundancy and resilience by eliminating the single point of failure inherent in static route environments. VRRP dynamically assigns the responsibility for a virtual router to a physical router, allowing several routers within a network to use the same virtual IP address, which is referred to as the Virtual Router IDentification (VRID). If the physical router becomes unavailable, the highest priority backup router switches over to the master state.

As shown in *Figure 1*, a VRRP router is a physical router, such as an OmniSwitch 7000. Both gateway switches are configured for VRRP functionality. The client sets gateways to the Virtual Router IDentification (VRID) and forwards all packets to that IP address. In the case of an active failure, the backup VRRP switch will immediately begin forwarding packets as the active



VRRP switch. The implementation of VRRP provides node redundancy for packets exiting a broadcast domain without the need to configure dynamic routing or router discovery protocols. Alcatel's implementation can also load share traffic when both routers are running.

OSPF Equal Cost MultiPath

The Open Shortest Path First Equal Cost MultiPath (OSPF ECMP) routing technique is implemented for routing packets along multiple paths in a load sharing configuration. The cost is a value or metric assigned to a path; it is typically based on the number of routing hops, distance or link speed to a specific destination. When implemented in the OmniSwitch 7000, the forwarding engine identifies the available paths and the next hop by using OSPF ECMP and the equal cost metrics that have been established for a specific destination. *Figure 2* shows a typical OSPF ECMP scenario. Nodes with equal cost paths have the ability to load share their traffic across multiple paths using round robin, which can enhance performance in addition to providing redundancy in the event of a network path or device failure.

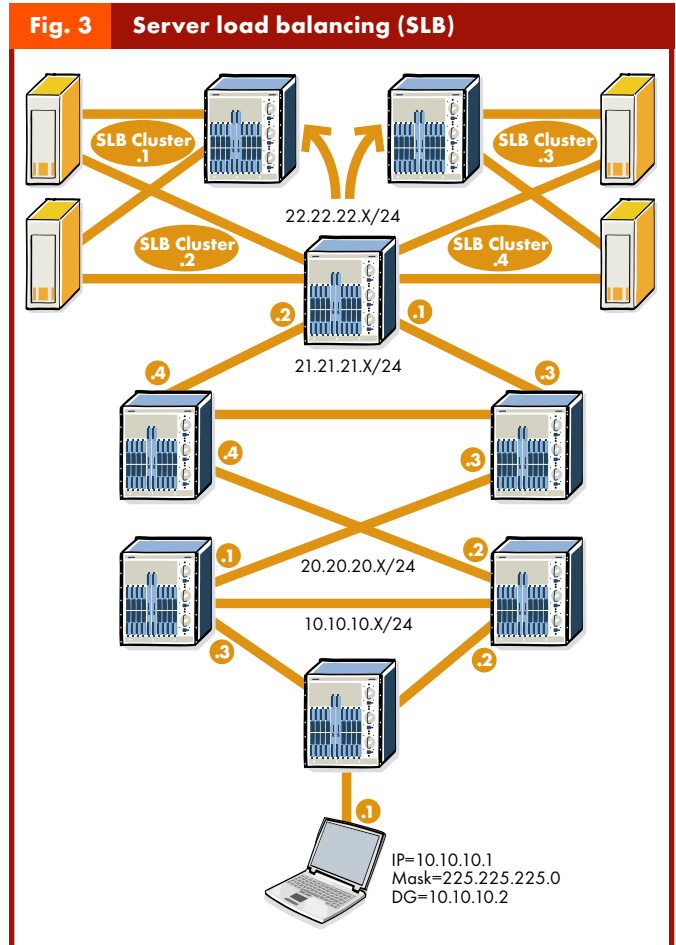
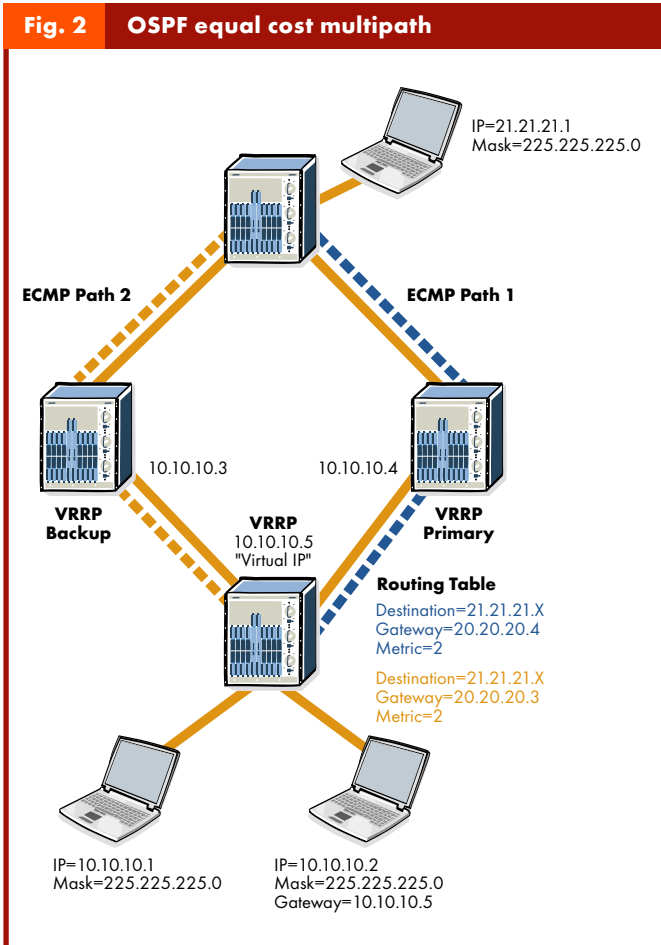
Link Aggregation

Dynamic link aggregation provides redundancy with a resilient uplink capability, allowing multiple virtual links to be established between two switches. If one link in an aggregate fails, all traffic is routed through the remaining links in that aggregate. Dynamic aggregate groups can be created using the standard IEEE 802.3ad Link Aggregation Control Protocol (LACP) for standardized dynamic link aggregation in multi-vendor environments. Alcatel also supports static link aggregation using OmniChannel, a technology designed specifically for linking Alcatel legacy enterprise products.

Dynamic aggregate groups can be created between an OmniSwitch 7000 and another vendor's switch if that vendor supports the LACP standard. Link aggregation can be made more resilient by utilizing multiple GNI or ENI modules to guarantee maximum availability in the event of a link or interface failure.

Server Load Balancing and Redundancy

When planning for high availability throughout the network, server load balancing with fully redundant servers, dual homing (use of two redundant connections),



clustering and backup servers are all recommended so that applications and information are always accessible. Redundant servers must have identical content so that there is no loss of mission-critical data or information. *Figure 3* shows an example of a typical server load balancing configuration. Critical servers are dual homed and clustered for node and link redundancy to provide maximum performance and throughput as well as high availability.

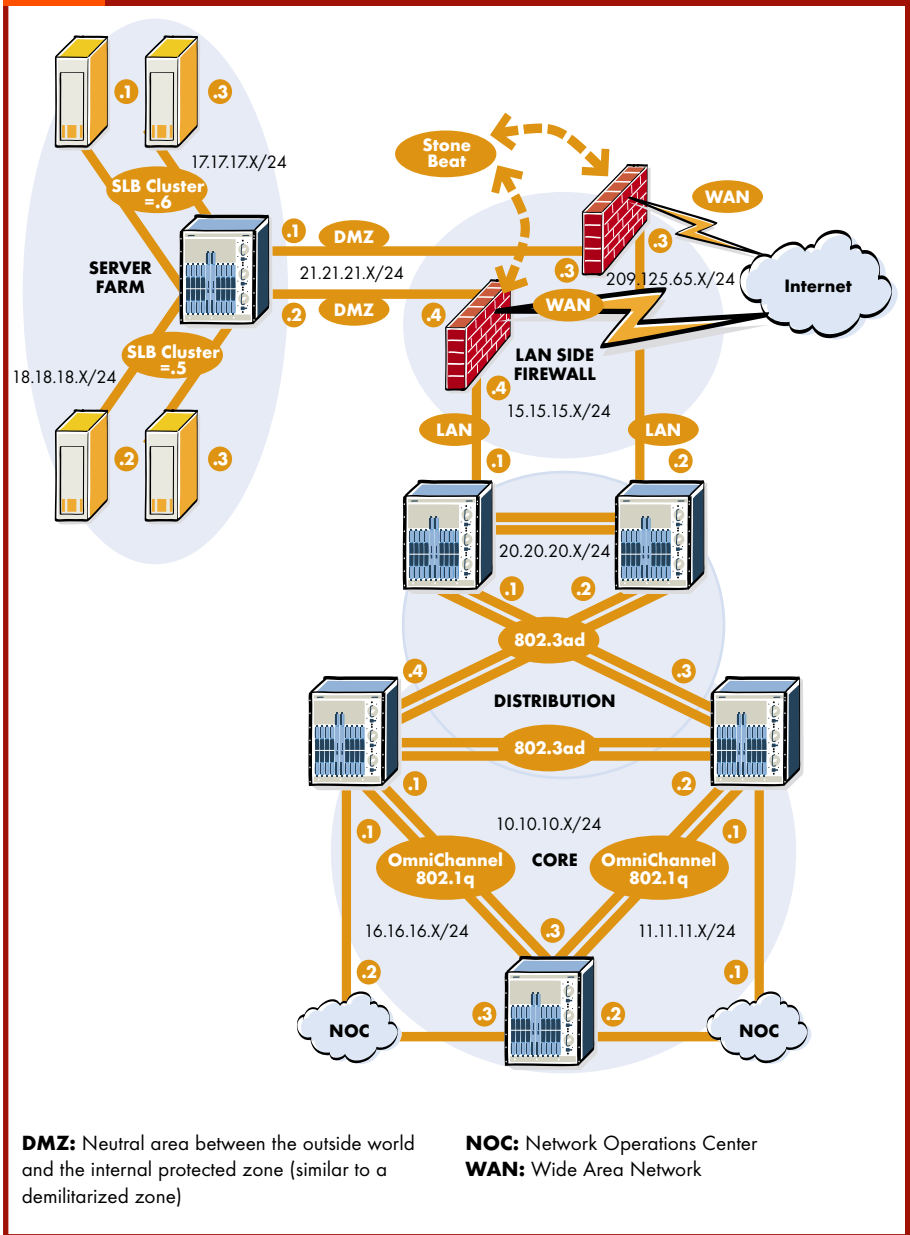
The OmniSwitch 7000 implementation is an integral part of the enterprise infrastructure, providing wirespeed server load balancing as a native part of the switch architecture. This switch-based server load balancing offers several advantages. It works across all the network interface modules, is wirespeed across the switch, does not require the purchase and configuration of an additional device or module, and allows any mix of network interfaces to be used. Up to 75 servers and fifteen server load balancing clusters are supported on the OmniSwitch 7000, offering a wide range of server combinations.

Firewall Clustering

Firewalls are important tools for securing the network from outside intruders. Consequently, they must be configured for high availability to ensure that all traffic is properly handled and securely managed. While server load balancing can help balance traffic at layers 3 and 4, firewalls operate at the Medium Access Control (MAC) layer so special clustering mechanisms are needed. The OmniSwitch 7000 employs industry leading Stone Beat firewall clustering for enhanced protection; intrusion protection is built into the device to proactively defend against denial-of-service, spoofing and other attacks. Firewall clustering technology provides automatic failover within a firewall cluster, eliminating a possible single point of network failure and ensuring that mission-critical Internet / intranet connections are secure and available.

Figure 4 is an example of a high availability security network solution. Mission-critical servers and firewalls are

Fig. 4 High availability security network



clustered using Stone Beat firewall clustering technology for maximum availability and accessibility. A multilayer hierarchical design ensures the optimal placement of security technologies within a high availability network solution.

Metropolitan Networks

There are increasing opportunities for enterprises to connect their LANs to Metropolitan Area Networks (MAN), and take advantage of new high availability offerings from service providers. This is accelerating the

move towards borderless computing for enterprises by providing transparent connectivity with Synchronous Optical Network / Synchronous Digital Hierarchy (SONET/SDH) in metropolitan networks. Ethernet-based MANs (eMAN) are gaining ground and offering distinct advantages to both enterprises and service providers with their ease of deploying provisioned services, such as managed security services, bandwidth allocation and service level agreements. eMANs will allow many enterprises to seamlessly extend their boundaries and take advantage of borderless computing and services for carrier-class availability.

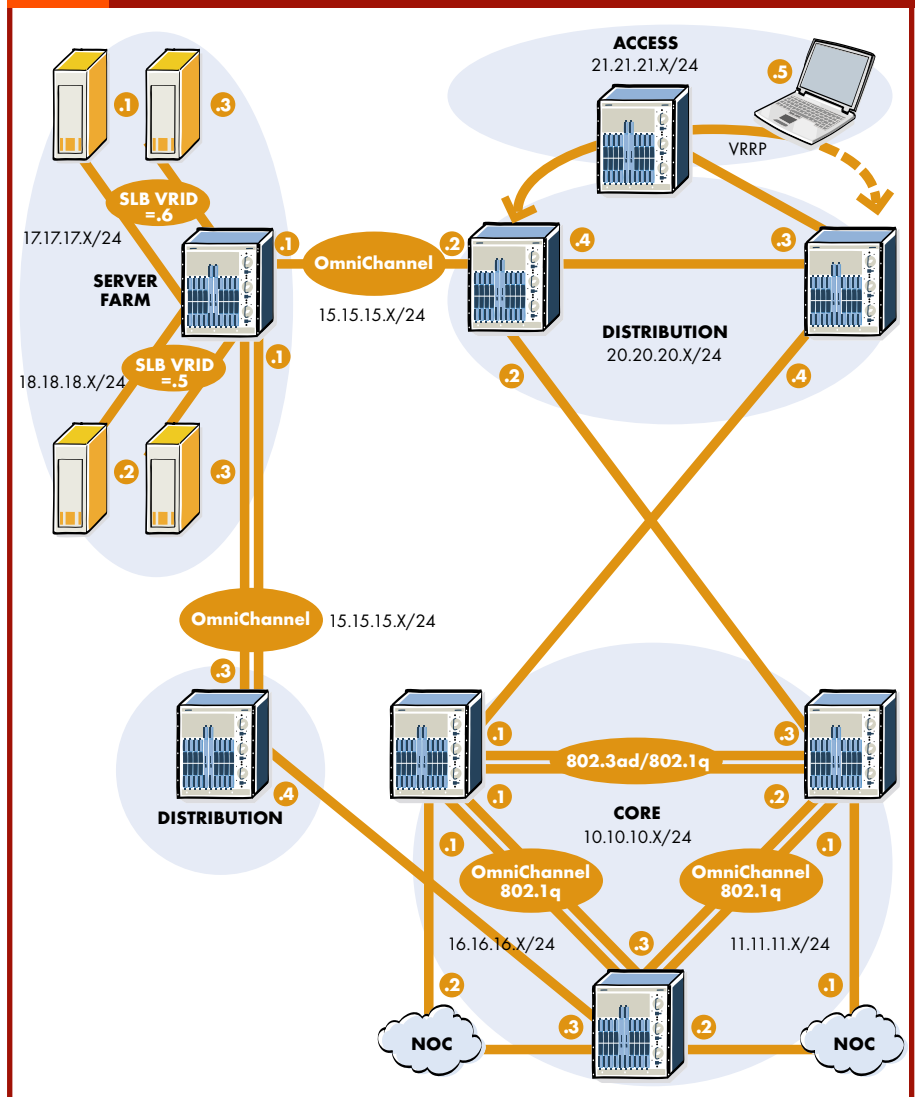
Conclusion

There are many factors and considerations involved in designing a network. Figure 5 shows a typical high availability enterprise network that incorporates many of these features, including server load balancing for mission critical application servers and link aggregation using IEEE 802.3ad and OmniChannel to maximize network level availability.

With the increasing importance of enterprise networks carrying voice and real-time applications, there is a need to deliver carrier-class availability. The Alcatel approach to achieving carrier-class availability in the enterprise is to extend the principles of high availability beyond individual devices to the entire network.

Alcatel's OmniSwitch 7000 switching platforms support IP communications and mission-critical applications, as well as enabling carrier-class availability to be achieved throughout an enterprise.

Fig. 5 High availability enterprise network



Bibliography

- [1] "Reality Check on Five-Nines", Business Communications Review, pp 22-27, May 2002.
- [2] "NSM: Often the Weakest Link in Business Availability", Gartner Group, 3rd July 2001.



Sheri Determan is Senior Director of Marketing for Alcatel Internetworking in Calabasas, California, USA.



Chris Arthmann is a Project Manager in Professional Services for Alcatel Internetworking in Calabasas, California, USA.

Abbreviations

- CMM** Chassis Management Modules
- DMZ** Neutral area between the outside world and the internal protected zone (similar to a demilitarized zone)
- eMAN** Ethernet Metropolitan Area Network
- eMAN** Ethernet-based MANs
 - ENI** Ethernet NI
 - GNI** Gigabit Ethernet NI
 - IP** Internet Protocol
- LACP** Link Aggregation Control Protocol
- MAC** Medium Access Control
- MAN** Metropolitan Area Network
- MTBF** Mean Time Between Failures
- MTR** Mean Time To Repair
- NEBS** New Equipment Building System
 - NI** Network interface
- NOC** Network Operations Center
- OSPF** ECMP Open Shortest Path First Equal Cost MultiPath
- PBX** Private Branch Exchange
- QoS** Quality of Service
- RRP** Rapid Reconfiguration Protocol
- SDH** Synchronous Digital Hierarchy
- SLB** Server Load Balancing
- SONET** Synchronous Optical NETWORK
 - STP** Spanning Tree algorithm and Protocol
- VRID** Virtual Router IDentification
- VRRP** Virtual Router Redundancy Protocol
- WAN** Wide Area Network

ARCHITECTS OF AN INTERNET WORLD



Alcatel and the Alcatel logo are registered trademarks of Alcatel. All other trademarks are the property of their respective owners. Alcatel assumes no responsibility for the accuracy of the information presented, which is subject to change without notice.
© 01 2003 Alcatel. All rights reserved. 3GQ 00002 0010 TQZZA Ed.01