

Layer 3 Switching Demystified

Introduction

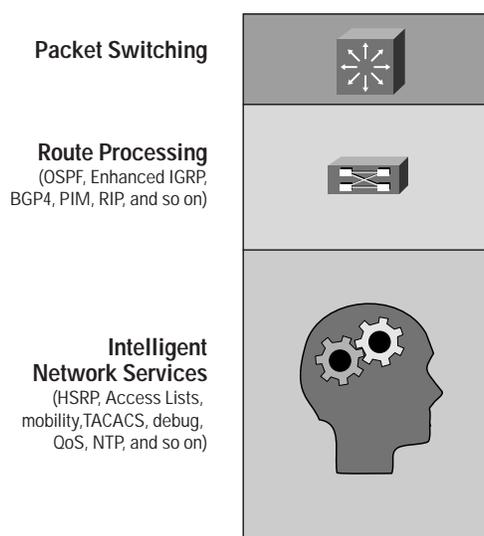
Any time new technology is introduced into the marketplace by networking vendors, there is a period of time when key decision makers at corporations must carefully analyze the onslaught of “techno-jargon” to determine its usefulness and relevance to their business. Recently, the industry has been bombarded with terminology such as Layer 3 switching, Layer 4 switching, multilayer switching, routing switches, switching routers, and gigabit routers. For purposes of this discussion, all these terms essentially represent the same function and as such, the term Layer 3 switching is used to represent them all.

While the performance aspect of Layer 3 switching makes most of the headlines, higher performance in switching packets does not, by itself, promise that all problems are solved in a network. There must be a recognition that application design, mix of network protocols, placement of servers, placement of networking devices, management, as well as the implementation of end-to-end intelligent network services are at least as important—maybe more so—than simply adding more bandwidth and switching capability to the network.

This paper shows that Layer 3 switching refers to a class of high-performance routers optimized for the campus LAN or intranet. Thus the same technology and intelligence that exists in traditional routing platforms need to be incorporated into the Layer 3 switches. This paper explains the major functions of a Layer 3 switch and its importance to

the design and operation of a campus network. Also included is a brief overview of Cisco’s various offerings in the Layer 3 switching marketplace.

Figure 1 Layer 3 Switching Components



A Visit inside a Layer 3 Switch

In discussions about Layer 3 switches, often the prime area of focus is raw performance which refers to the aggregate number of packets that a device can switch in and out over a fixed period of time. Layer 3 switches tend to have packet switching throughputs in the millions of packets per second (pps), while traditional general-purpose routers have evolved from the 100,000 pps range to over a million pps. In essence, aggregate performance is the primary difference between

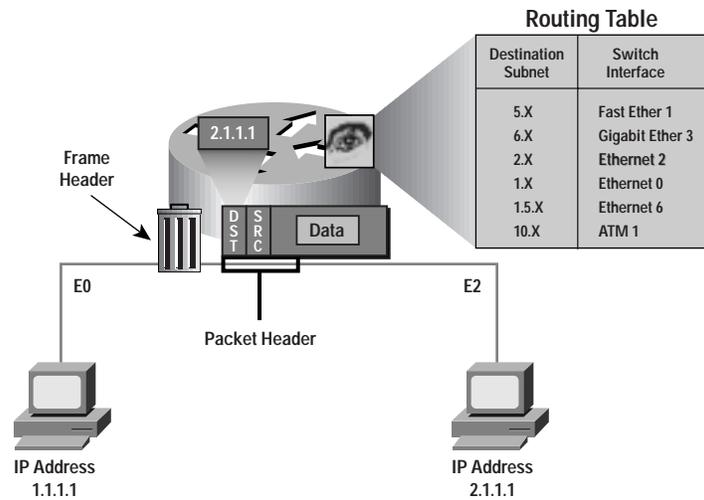
Layer 3 switches and traditional routers. The features and functionality of Layer 3 switches and routers do, however, have numerous similarities. These similarities manifest themselves not only in packet switching methods, but also in the categories of route processing and intelligent network services. All three of these categories are explained further in the remaining sections.

Packet Switching

In many respects, the switching of packets is the simplest part of the operation of a Layer 3 switch. The only major difference between the packet switching operation of a router and a Layer 3 switch is the physical implementation. In general-purpose routers, packet switching takes place using microprocessor-based engines, whereas a Layer 3 switch performs this using application specific integrated circuit (ASIC) hardware.

Irrespective of the implementation, the job of the packet switching engine is to examine incoming packets, determine the destination address, (IP address 2.1.1.1 in Figure 2) compare this address with the contents of the routing table (created by the dynamic routing protocol, described in the next section) and forward the frame out the appropriate interface (Ethernet 2 in this case). The switching engine also does some packet manipulation by creating an Ethernet frame with its own Ethernet Media Access Control (MAC) address as the source, decrementing the Time-To-Live (TTL) field in the IP header and recalculating the frame check sequence (FCS). This switching technique is sometimes referred to as “packet by packet”, and is precisely what today’s routers do in the networks all over the world.

Figure 2 Packet Switching Operation



Another important function of packet switching is the ability to perform longest-match lookups. For example, a destination address of 1.5.6.6 has two possible entries in the routing table. The first possibility is interface Ethernet 0 because there is a match on the first digit, 1. There is also a match on Ethernet 6 on the first two digits, 1.5. Clearly 1.5.X is a more specific match, and therefore, the packet should go out the Ethernet 6 interface. This fundamental capability to perform longest match lookups will be exploited when looking at some intelligent network services related to mobility.

While these scenarios may seem complex, Layer 3 packet manipulation operations are well documented and understood. The real issue that people must face related to performance is when to take advantage of higher-performing technology. Many people confuse aggregate performance with speed of applications, a flawed assumption. If network demand is in the order of 10,000 pps (for example) and current Layer 3 devices process packets at 200,000 pps, then replacing them with devices that perform even at a trillion pps will not, in fact, speed up the applications. This setup is comparable to adding lanes on a freeway where traffic is already running at the speed limit—it simply won't get you there any faster.

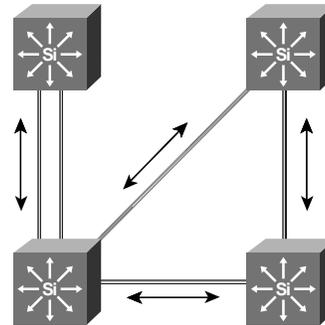
The most important criteria for the selection of a Layer 3 switch is not raw performance, but is both route processing and intelligent network services. The next two sections discuss the other requirements of Layer 3 switches that serve as a deployment prerequisite.

Route Processing

One of the fundamental capabilities of routers and Layer 3 switches is the creation of routing tables that automatically adjust themselves to the ever-changing network topologies caused by link failures, device failures, and additions and deletions to the network. Dynamic routing protocol technology is used in Layer 3 switches and routers to populate these Layer 3 routing tables. The Internet is the prime example of a network that stresses routing protocols to their fullest, and it has been an innovation center in routing protocol development. It is important to note here that routing

protocol processing is a software-based activity that operates independently from packet switching discussed in the previous section.

Figure 3 Routing Protocol Exchanges



From a historical perspective in the Internet, the first dynamic routing protocol deployed was the Routing Information Protocol (RIP). RIP is classified as a “distance vector” routing protocol and as such it exhibits characteristics of periodic updates (even when there are no changes in the topology), slow convergence around failed links (in the order of minutes), and limited scalability because of finite hop count restrictions. RIP has been obsoleted by newer generations of routing protocols that solve its inherent scalability and overhead limitations. The relatively new RIP Version 2 does not overcome any of the limitations of RIP but it does allow for an IP addressing technique known as variable length subnet masks (VLSM), which is sometimes used to conserve IP address space.

In response to RIP's shortcomings, Cisco was an early pioneer in the development of alternative routing protocols. Interior Gateway Routing Protocol (IGRP) was developed by Cisco to address the limitations of RIP, as well as provide for greater flexibility in determining the path that a packet takes through the network by considering link speeds, link quality, delay, and other metrics into the intelligent forwarding decision. Consequently, IGRP was quickly adopted as the accepted standard routing protocol in numerous enterprise networks because of its flexibility, robustness and ease-of-use.

Further developments in a class of routing protocols designated as “link-state” took place in the early 1990s with the publishing of the Open Shortest Path First (OSPF) routing

protocol specification. Around failures, OSPF has the properties of very fast convergence with fewer routing protocol updates in stable topologies. Link-state routing protocols exchange link status information with neighbors and flood this information around the network such that each device has an identical map of the network topology. While the fast-convergence benefits of OSPF are great, they do come with some complexity. Network designs using OSPF must consider an area structure to enable the network to scale, and this scenario can add artificial constraints on the physical network design topology and complexity in some cases.

OSPF remains a complicated protocol and as such, the stability of implementations vary considerably. With Cisco IOS software, years of OSPF field experience in the Internet and largest Internets have been integrated into all of Cisco's Layer 3 switch offerings.

In order to take advantage of faster converging routing protocols, network designers must change from one routing protocol to another. This migration is a very time-consuming and complex process. To address the needs of the huge installed base of customers using IGRP as their routing protocol, Cisco developed Enhanced IGRP to bring the benefits of fast convergence that OSPF exhibits, while maintaining the simplicity of the original IGRP protocol. Enhanced IGRP has also been leveraged to support other routed protocols such as AppleTalk and Novell IPX. Cisco customers have a choice of fast-converging routing protocols to suit their business needs with standards-based OSPF or the simplicity of Enhanced IGRP. Both OSPF and Enhanced IGRP use the VLSM techniques mentioned earlier to permit IP address conservation.

In most real-world networks, numerous routing protocols are used across the enterprise to optimize network connectivity. In this "multilingual" environment, there is a need for sophisticated translation between these routing protocols so that they all work in harmony. Cisco's leading-edge routing protocol redistribution technology in Cisco IOS software has again been developed over the years, with Internet engineers who have built and debugged the most complex network in the world. These routing protocol redistribution features allow network designers to tailor their network designs and operational characteristics to suit individual needs.

Regardless of the routing protocol used, networks are typically designed with redundancy to load share packets across parallel links. In Figure 3, between the two Layer 3 switches with parallel links, the routing protocol and corresponding routing table would create two entries for packets traversing those two switches and hence would set up load sharing. This fundamental capability allows for flexible, fault-tolerant design topologies, as well as incremental bandwidth between switches.

Another key development in routing protocol technology that is critical to the successful deployment of network applications is in the area of IP multicast routing. Numerous applications use this technology, such as Microsoft Netshow, Precept IP/TV, TIBCO stock market feeds, Intel Proshare, data replication applications, and many others that are available in the commercial marketplace. It is indeed a prudent move for network designers to consider IP multicast, as it is likely that these applications will take corporations by storm much the same way that a little-known application called Mosaic (soon to be known by the more popular term, World Wide Web) did just a few short years ago.

To support these applications, Layer 3 switches also need to support multicast routing protocols. With the extensive work done in the Internet multicast backbone (MBONE), it is clear that the multicast routing protocol of choice is Protocol Independent Multicast (PIM). PIM has the property of operating independent of the underlying unicast routing protocol (Enhanced IGRP, OSPF, RIP, static, and so on) and supporting applications that have both servers sending to multiple destinations (dense mode) or numerous small workgroups operating in different multicast groups (sparse mode). Again, PIM is a complex protocol that takes years of field experience in a challenging environment to ensure a robust implementation. Fortunately, Cisco has been at the forefront of both the creation of PIM and the rollout of the Internet MBONE to a point where PIM is deployed in over 50 percent (and growing) of its 1000+ routers.

Buying a Layer 3 switch without the richness and depth of routing protocols is somewhat akin to a driverless car. The car can certainly travel very fast in the direction that it is pointed, but the intelligence lies in the driver, who needs to make all the decisions about where it should go and when to stop and turn. It is clear then, that while packet switching



performance makes the headlines of trade magazines, it is the routing protocols, the nerve center of a Layer 3 switch, that makes the Internet and intranets function in a way that ensures stable, resilient networks.

Intelligent Network Services

While packet switching performance is a given, and stable and proven routing protocols are a deployment prerequisite for Layer 3 switches, it is the intelligent network services that permit applications to run on the network as well as enable the creation of a cost-effective, operational environment to support day-to-day operations and management of the enterprise intranet. Once again, the foundation behind these intelligent network services is the Cisco IOS software that anchors all of Cisco's Layer 3 switching solutions.

Numerous intelligent network services enable the deployment of mission-critical applications such as SAP, Peoplesoft, BAAN, Microsoft Backoffice, Oracle, and others. Deployment of mission-critical applications demands an end-to-end network design where intelligent networking elements work together to solve problems. One of the tools available is Hot Standby Router Protocol (HSRP), which transparently allows end users to route around failed Layer 3 switches with minimal application downtime and no impact on the end-user station.

Virtually all network managers consider the devices that comprise the core of their network to be a key strategic asset and as such, protect access to it to eliminate the possibility of unauthorized users crippling the network. Security features that are required to protect the network include encrypted passwords and configuration files, access control lists (which allow only certain users to access a device), and authentication systems such as TACACS+, which authenticates user logons and records login attempts. Virtually all these features are used in operational networks today, and it is key that for easy migration to Layer 3 switching technology, this operational infrastructure remain intact.

To solve or detect a network problem, it might be necessary to access a device and make some configuration changes. Remote access is available by the Telnet protocol and in fact, multiple inbound and outbound Telnet sessions are often used to reach the device that needs the modification. In making a configuration change, it is imperative that the

change take effect immediately on the Layer 3 switch in question, without having to reboot or partially reboot the device and hence take users off line.

Other services impact the bottom line by enabling cost-effective personnel adds, moves, and changes. From a user mobility perspective, it costs corporations tremendous amounts of money to move users around the network, and Cisco IOS software offers mobility solutions to satisfy all types of networking environments. Tools such as Dynamic Host Configuration Protocol (DHCP) relay support in Layer 3 switches, along with the Local Area Mobility IOS feature, permit mobility of users with modern dynamically addressable IP stacks, as well as those bound to a static IP address scheme. Local Area Mobility leverages the flexibility of routing protocol redistribution and longest-match lookups described earlier, to provide a transparent IP mobility solution.

Based on Cisco's experience building the Internet and the largest campus intranets, years of operational practice have been integrated into the Cisco IOS software in the form of troubleshooting and network debugging tools. Every Layer 3 switch has an extended suite of online network debugging capability that allows network managers to remotely troubleshoot network problems. Without this debug functionality, it would be necessary to deploy multiple network analyzers at different points in the network and waste precious networking staff to physically install and remove these probes as problems arise. Clearly, with these Layer 3 switches anchoring the core of mission-critical networks, problems must be quickly identified and corrected to minimize application downtime.

Numerous other troubleshooting and debug tools are integrated into Cisco IOS software. Examples include the syslog feature that allows network managers to centrally collect warning, error, and network debug information from multiple devices for later analysis. Having this capability, however, is useful only if the collected information is time-stamped and the system time-of-day clocks are synchronized in all the Layer 3 switches. The Network time Protocol (NTP) component of IOS software is used to perform this time-of-day synchronization between Layer 3 switches and wiring closet LAN switches as well.

These examples of intelligent network services represent only a small fraction of what is available in Cisco IOS. Features like quality of service (QoS), advanced queuing mechanisms, multi-protocol routing capability, and others, add to the type of functionality that has been used to build the Internet and intranets and now acts as a fundamental building block of Cisco's Layer 3 switching products.

Overview of Cisco's Layer 3 Switching Solutions

Cisco has several Layer 3 switching solutions to offer its customers. For the large installed base of Cisco switches and routers, there are solutions that leverage existing hardware and software platforms such as the Catalyst® 5000 series, and the Cisco 7500 router series. New Layer 3 switching platforms in the Catalyst 8500 family also offer high-performance, sophisticated QoS features for the campus LAN backbone. All the Cisco Layer 3 switching solutions have the foundation of the IOS software to enable the reliable routing protocol implementations and Intelligent network services described earlier.

With the Catalyst 5000 series multilayer switches, Layer 3 switching functionality is delivered with the Route Switch Module (which handles routing protocols and intelligent network services) and the NetFlow Feature Card that does the high speed Layer 3 packet switching. The Catalyst 5000 series offers Layer 3 switching performance on par with Layer 2 performance, as well as full multiprotocol routing, NetFlow data export, and wire-speed packet filtering. This setup offers a simple migration path for existing Catalyst 5000 customers to deploy a common platform in multiple layers of the network with both Layer 2 and Layer 3 switching.

In the distribution layer and core of the network, the new Catalyst 8500 series Layer 3 switches offer wire-speed Layer 3 switching performance for IP, IPX, IP multicast, and bridging. The QoS functionality inherent in the Catalyst 8500 allows for QoS policy implementations and high-performance network core design. The Catalyst 8500 functionality is available as a standalone platform or integrated into Catalyst 5500 chassis.

The Cisco 7500 family of multiprotocol routers continues to evolve with advanced router system (ARS) enhancements targeted at all three components of Layer 3 switch operations mentioned earlier. For customers using Cisco 7500 routers in the campus backbone, new line card modules called VIP2-50 enhance system performance, whereas higher-performance processor technology in the Route Switch Processor 4 (RSP4) accelerates intelligent services while scaling routing protocol performance.

Conclusions

It should be clear that Layer 3 switching doesn't represent a revolution in technology, but more of an optimization of Internet class routers for the campus intranet. Technologies such as Gigabit Ethernet offer a new link-layer technology, and should not be confused with Layer 3 switching functionality. Learning from the Internet experience, Cisco and its customers realize that it is the combination of performance, robust routing protocol experience and intelligent network services that constitutes the formula for success in new and involving campus networks.



Corporate Headquarters

Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 526-4100

European Headquarters

Cisco Systems Europe s.a.r.l.
Parc Evolic, Batiment L1/L2
16 Avenue du Quebec
Villebon, BP 706
91961 Courtaboeuf Cedex
France
<http://www-europe.cisco.com>
Tel: 33 1 6918 61 00
Fax: 33 1 6928 83 26

Americas

Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
<http://www.cisco.com>
Tel: 408 526-7660
Fax: 408 527-0883

Asia Headquarters

Nihon Cisco Systems K.K.
Fuji Building, 9th Floor
3-2-3 Marunouchi
Chiyoda-ku, Tokyo 100
Japan
<http://www.cisco.com>
Tel: 81 3 5219 6250
Fax: 81 3 5219 6001

Cisco Systems has more than 200 offices in the following countries. Addresses, phone numbers, and fax numbers are listed on the Cisco Connection Online Web site at <http://www.cisco.com>.

Argentina • Australia • Austria • Belgium • Brazil • Canada • Chile • China (PRC) • Colombia • Costa Rica • Czech Republic • Denmark
England • France • Germany • Greece • Hungary • India • Indonesia • Ireland • Israel • Italy • Japan • Korea • Luxembourg • Malaysia
Mexico • The Netherlands • New Zealand • Norway • Peru • Philippines • Poland • Portugal • Russia • Saudi Arabia • Scotland • Singapore