Network Design Guidelines

Introduction

Extreme Networks Gigabit Ethernet Layer 3 switches give you the flexibility to solve many problems that plague a wide range of network environments. From enterprise desktops, LAN segments, server groups and the network core, all Extreme switches leverage the same hardware, software and management architecture to deliver a common set of services throughout your infrastructure. In each of these areas, Extreme Networks switches provide smooth and cost-effective system solutions which reduce congestion, increase capacity, eliminate poor routing performance and ease application overload. Ultimately, you'll be able to focus on the applications that run on the network.

Common Network Problems	The Extreme Networks Solution
Slow speed due to software-based routing	Wire-speed routing in ASICs/non-blocking architecture
Low network availability	Redundant links, redundant power, load-shared trunks
Poor application performance	Policy-Based Quality of Service prioritizes applications

Over-Subscription

Over-subscription is one of the most important tools used to design networks. ***Over-subscription*** or ***over-subscription ratios*** deal specifically with points in a network where bottlenecks occur. The impact of improper over-subscription ratios is congestion, which causes packet loss.

However, over-subscription is acceptable in some areas of the network. But what is considered "acceptable" has been fiercely debated in the desktop, segment, server-farm and core application spaces. In general, different aggregation points in the network have unique over-subscription requirements. So the acceptable ratio for desktop connectivity is different from the ratio for server connectivity.

Over-subscription is tolerated at the edge or in feeder branches of a network. In intermediate distribution frame (IDF) or wiring closet layouts, over-subscription defines the traffic model. By understanding over-subscription, you can identify and correct the adverse effects that congestion has on network performance.

Over-subscription ratios are calculated by adding the potential bandwidth requirements of a particular path and dividing the total by the actual bandwidth of the path. Although a ratio larger than 1:0 is considered over-subscribed, it does not necessarily mean congestion will occur.

Allowing a link to have less bandwidth than required to sustain a peak load recognizes that average bandwidth utilization is much lower than the maximum. An acceptable over-subscription ratio is typically about 3:1.

For example, the Summit24 has 24 10/100 Mbps ports for a total of 2.4 gigabits (or 4.8 gigabits full duplex) of potential bandwidth. This potential bandwidth is required to access the network core. Summit24 is equipped with a single Gigabit Ethernet port, which provides 1 gigabit (or 2 gigabits full duplex) of actual bandwidth to the core.

In this case, the Summit24 gigabit uplink is over-subscribed. Its 10/100 Mbps port-count totals 2.4 gigabits of potential bandwidth, but the uplink can support only 1 gigabit of actual bandwidth. While the resulting ratio is 2.4:1, it is well within the acceptable range for over-subscription.

To avoid congestion caused by improper over-subscribed links, the following ratios are advised:

Desktop:	3:1
Segment:	2:1
Server:	1:1
Core:	1:1

Why You Need a Non-Blocking Architecture

Device capacity, which can be limited by hardware architectures, is also crucial to network design. Hardware architectures can be blocking or non-blocking. In terms of capacity, a blocking architecture cannot meet the full-duplex bandwidth requirements of its ports, which results in packet loss. Non-blocking means that a device's internal capacity matches or exceeds the full-duplex bandwidth requirements of its ports, and will not drop packets due to architecture. So, if one link is over-subscribed, it will not affect other links in the same switch.

Consider a device with 1.2 gigabits of backplane capacity. If it has 48 10/100 Mbps Ethernet ports, the backplane could not handle the 9.6 gigabits of full-duplex input capacity because it is a blocking architecture. The backplane can only support six 10/100 Mbps ports, which equal 1.2 gigabits full duplex.

Every switch in the Extreme Networks family has a non-blocking architecture. For example, the non-blocking architecture in the BlackDiamond 6800 has a capacity of 64 gigabits. This architecture allows it to provide 32 full-duplex Gigabit Ethernet ports in a non-blocking configuration. Non-blocking is especially important to core infrastructure designs where the aggregate traffic load of the entire network comes to bear. For a non-blocking architecture at the desktop, the Summit48 switch is well suited for the task. The Summit48 architecture has a capacity of 17.5 gigabits. This allows the Summit48 to provide 48 10/100 Mbps Ethernet ports and 2 full-duplex Gigabit Ethernet ports in a non-blocking configuration.

Software-Based Routing vs. Layer 3 Switching

Layer 3 switching is routing in hardware. Traditional software-based routers forward in the 500,000 packets per second range. Layer 3 switches forward up to 48 million packets per second. Although they both route packets, the integration of forwarding using high-speed ASIC technology differentiates Layer 3 switches from traditional software-based routers.

Several issues should be addressed when designing a routing infrastructure, including where to route and performance. Determining where to route is perhaps most challenging because variations in network design have a direct effect on routing performance. When considering a network design, it is important to remember that the constraints of traditional router performance have driven the model. This is not the case with today's faster Layer 3 switching.

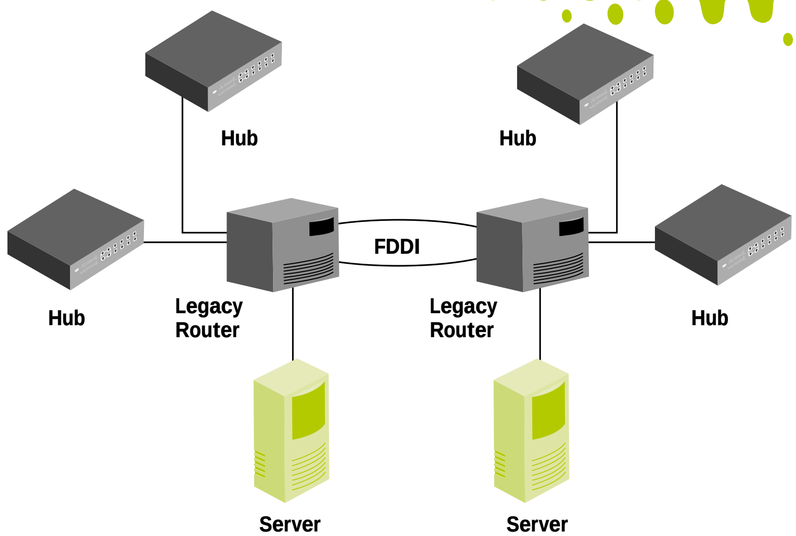

Figure 1

As shown in Figure 1, traditional infrastructures relied on one or two expensive routers in the core. The network was then designed around these core routers. This failed to solve the problem of scalability because software-based routers have limited capacity and port density.

The advent of Layer 3 switching dramatically changed the way we design networks. Low-cost Layer 3 switches eliminate our dependence on expensive core routers because they can be deployed throughout the network – from desktops at the edge, to data centers at the core. And they bring with them wire-speed performance, whether switching at Layer 2 (hardware-based bridging) or Layer 3 (hardware-based routing).

Wire-speed performance at Layer 3 means that network architects are no longer bound by the performance penalty usually associated with software-based routers. The increased latency introduced by Layer 3 subnet boundaries do not exist with wire-speed Layer 3 switching.

Layer 3 switches let you design networks that are driven by the physical flow of packets, instead of designing around the limitations of software-based routers and other slow network equipment. Ultimately, this allows you to focus on the logical organization of your network and optimize the performance of applications that run on it.

Quality of Service

Layer 3 switches let you deploy routing throughout a network. Packets are routed in hardware at wire speed. Application performance improves significantly.

However, networks require much more than just "speeds and feeds." Gigabit Ethernet offers increased speed and bandwidth, but you still need to control it. And only quality of service can deliver this control.

QoS gives you the power to predict and control network behavior by prioritizing applications, subnets and end stations, and guaranteeing them a specific amount of bandwidth or giving them priority. So, while a good network design can reduce the negative effects of over-subscription and blocking, QoS can deliver end-to-end control over traffic flows. Controlling the behavior of the network means that network managers can allocate limited resources based on policies rather than congestion.

Signaled QoS vs. Policy-Based QoS

There are two primary ways to deliver QoS today. One uses signaling, while the other is policy based.

For example, the resource reservation protocol (RSVP) uses signaling. Pairs of RSVP-enabled end-stations use signaling across the network to reserve bandwidth for an application session. However, it is not possible to control how much bandwidth is reserved or when it can be reserved. In terms of control, network managers can only restrict the use of RSVP to certain end-stations. It is also costly and time-consuming to deploy because applications and end-stations must be altered to handle RSVP requests.

Extreme Networks provides a powerful alternative to signaled QoS by allowing traffic classification based on a set of rules, or policies, that run in the switch. Utilizing 802.1D-1998 (formerly 802.1p prioritization), there is no active path or communication needed between network devices and end-stations. Instead, information is embedded in each packet using 802.1Q tagging.

Policy-Based QoS involves classifying network traffic into groups and assigning QoS profiles to each group. Traffic groups can be classified by MAC address, physical port, 802.1Q, 802.1D-1998, destination IP address and well-known TCP/UDP port numbers.

The process of identifying network traffic is performed by switches equipped with packet classification features. As packets enter a switch, they are compared against policies that have been set for traffic groups and are placed in a hardware queue. The policies associated with each queue guarantee a specific amount of bandwidth and traffic priority. By classifying traffic within the switch and embedding QoS information within each packet, network managers can set and enforce QoS policies anywhere on the network.

Having Policy-Based QoS within each switch gives network managers a consistent way to predict and control traffic across the entire network. In addition, the deployment and administration of QoS is dramatically improved because end-stations do not require modifications in order to classify traffic.

Hardware-Based QOS

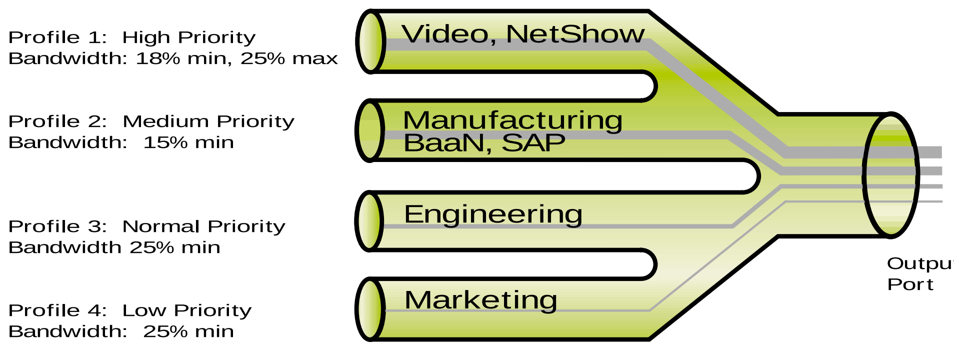

Figure 2

Ports. By applying priorities to individual input ports, 802.1D-1998 priority flags can be used to propagate QoS throughout the entire enterprise network – from the desktop to the core.

VLANs. This is QoS in a simple form. By applying a QoS profile to port-based VLANs, you can quickly control congestion where they are trunked through an uplink.

Subnets. Protocol-based VLANs can tie QoS to specific IP, IPX and AppleTalk subnets. This allows you to easily single out a specific group of users and provide them with explicit QoS.

Applications – TCP/UDP Port Numbers. The key to application-specific QoS is having visibility into the transport layer (Layer 4). This gives you control over the applications using individual port numbers. Applications that use well-known port numbers can have QoS profiles applied to them.

The Application Spaces

Each application space shown in Figure 3 – desktop, segment, server and core – has different requirements when it comes to over-subscription, routing and QoS. Extreme's Summit and BlackDiamond switches are designed with different configurations and capabilities to meet the unique requirements of each space.

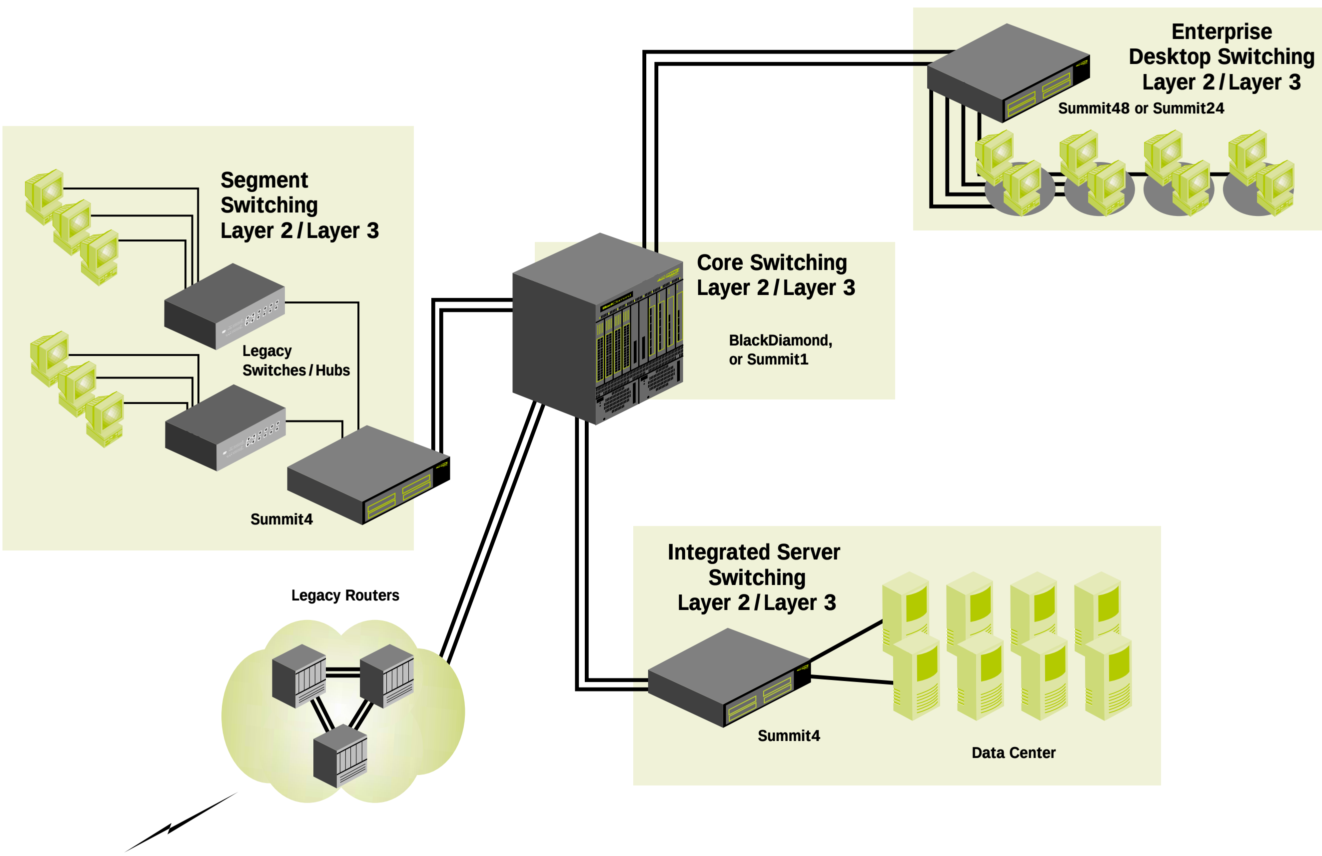

Figure 3: Extreme Networks' complete end-to-end systems solution covers the full spectrum, from redundant, fault-tolerant carrier-class chassis to high-density, low-cost desktop switching.

Product Descriptions

BlackDiamond 6800 – The BlackDiamond 6800 is a 10-slot modular chassis switch with an internal capacity of 64 gigabits. Using four-port Gigabit Ethernet modules or 32-port 10/100 Mbps Ethernet modules, the BlackDiamond 6800 supports a maximum configuration of either 32 Gigabit Ethernet ports or 256 10/100 Mbps Ethernet ports. It is non-blocking and routes 48 million packets per second. The six-port Gigabit Ethernet modules, which also have a non-blocking architecture, can be used to create a 48-port Gigabit Ethernet configuration in the BlackDiamond 6800.

Summit1 – The Summit1 is an 8-port Gigabit Ethernet switch with an internal capacity of 17.5 gigabits. It is non-blocking and routes 11.9 million packets per second.

Summit4 – The Summit4 has 16 10/100 Mbps Ethernet ports, six Gigabit Ethernet ports and an internal capacity of 17.5 gigabits. It is non-blocking and routes 11.3 million packets per second.

Summit24 – The Summit24 has 24 10/100 Mbps Ethernet ports, one Gigabit Ethernet port and an internal capacity of 8.5 gigabits. It is non-blocking and routes 5.1 million packets per second.

Summit48 – The Summit48 Enterprise Desktop Switch has 48 10/100 Mbps Ethernet ports, two Gigabit Ethernet ports and an internal capacity of 17.5 gigabits. It is non-blocking and routes 10.1 million packets per second.

The Enterprise Desktop

Switched connectivity for enterprise desktops is driven by price/performance. The need to aggregate more enterprise desktops to a single switch puts pressure on vendors to sacrifice performance and resource management in favor of reducing the price per port.

This trade off could have a negative impact on performance and overall network health. For example, the uplink on an enterprise desktop switch with high-density 10/100 Mbps Ethernet connections can become congested due to improper over-subscription ratios.

Over-subscription Ratios To avoid this, an enterprise desktop switch should support an over-subscription ratio of 3:1 or lower and a non-blocking architecture to switch local traffic without loss. The result is an enterprise desktop switch that is congestion-free under normal conditions and able to leverage QoS to control peak loads.

Layer 3 Switching With Extreme's introduction of the Summit Enterprise Desktop Switch, Layer 3 switching is now an affordable and flexible system solution for the desktop.

Because each Summit switch routes at wire speed, there is no routing performance penalty when deployed at the edge of the network. This increases the flexibility of the network by supporting IP unicast and multicast routing. Multicast and intra-segment traffic are kept on their respective local segments and therefore do not consume core bandwidth.

As a result, you can design a network infrastructure that complements the way your business runs. Another compelling reason to route at the edge is that robustness can be improved by pushing multipath routing protocols, like OSPF, as close to the desktop as possible. The net gain is a more resilient network that can handle change.

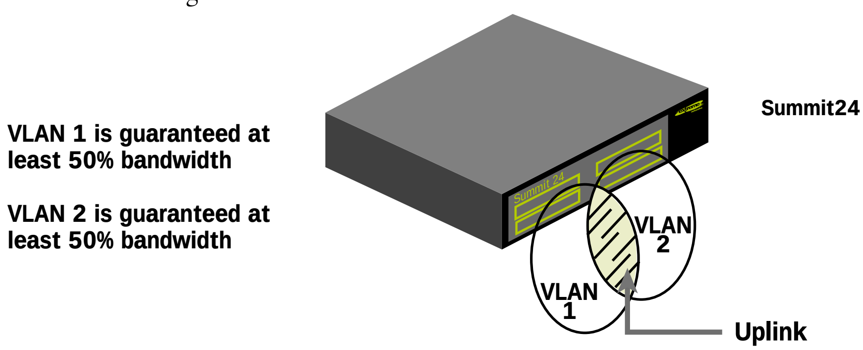

Figure 4

Quality of Service As shown in Figure 4, a Summit24 that supports two VLANs can use QoS to control congestion on the uplink to guarantee fair access to the core network. The solution is as easy as defining two QoS profiles, each with a 50% minimum bandwidth allocation. By guaranteeing bandwidth, both VLANs have a guaranteed level of performance, even when the uplink is over-subscribed.

Port-based priority is a powerful tool for QoS deployment in the desktop space. By assigning a medium or high priority to the input port of a switch, you can provide end-to-end QoS for specific people or equipment on the network.

For example, consider a subnet with a mix of engineering and administrative staff. You can assign high priority to the ports that connect engineering and give them preferential network access. By assigning QoS to individual input ports, you achieve more granular point-deployment of QoS.

Application-specific QoS is another useful traffic control mechanism for desktop switching. You can assign a more granular level of QoS to mission-critical applications to eliminate any chance of congestion. By defining a QoS profile based on the unique IP port number of an application, you can prioritize it and prevent other traffic from interfering with its performance.

Summit48 Enterprise Desktop Switch in the Desktop

Designed to support high-density desktop-to-wiring closet applications, Summit48 is the ultimate Enterprise Desktop Switch with 48 10/100 Mbps ports and two Gigabit Ethernet ports. Summit48 breaks the legacy price/performance barrier by delivering the same non-blocking architecture found in other Summit switches, while offering a considerable increase in port density. As shown in Figure 5, the 2.4:1 over-subscription ratio makes the Summit48 an ideal Enterprise Desktop Switching solution.

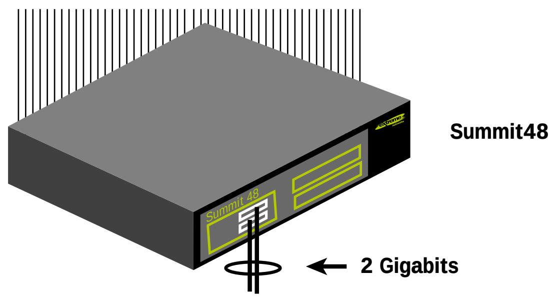

Figure 5

Summit24 Enterprise Desktop Switch in the Desktop

Alternatively, Summit24 is ideally suited for desktop switching when high-density connections are not a critical factor. Like the Summit48, the over-subscription ratio for Summit24 is 2.4:1. Summit24 is an effective aggregation device in this configuration because it offers 24 ports of desktop connectivity with a comfortable over-subscription ratio. In addition, Summit24 has an aggregate forwarding rate of 5.1 million packets per second.

BlackDiamond 6800 as a High-Capacity, Fault-Tolerant Desktop-to-Wiring Closet Solution

Using a BlackDiamond 6800, you can create a wiring closet with 192 10/100 Mbps ports to desktops and aggregate four Gigabit Ethernet ports into a single logical link to the core. BlackDiamond is an ideal desktop aggregation device in environments that require an increased level of redundancy, fault tolerance and load-sharing capabilities. As shown in Figure 6, the over-subscription ratio for BlackDiamond is 2.4:1.

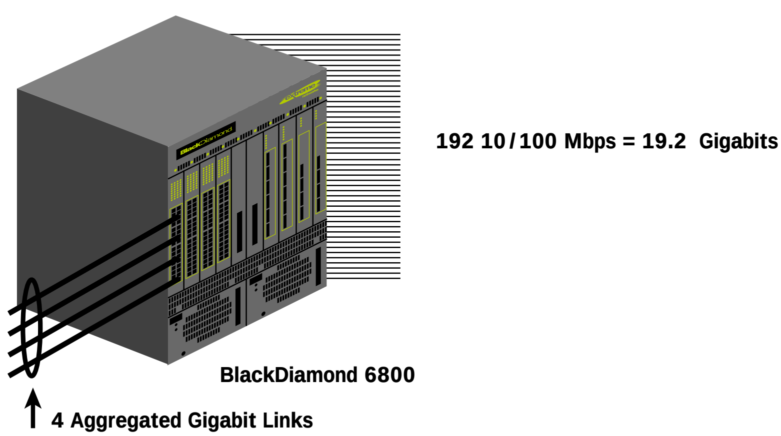

Figure 6

The Segment

Segment switching covers the hierarchy of devices that support desktop switching. This hierarchy of devices typically involves small 10 Mbps Ethernet switches, or feeder hubs, that are attached to higher performing Fast Ethernet switches.

Segment switching shares the same problems as desktop switching, but with added complexity. Feeder hubs that connect multiple desktops to segments will generate more traffic because there are more active devices in the data path. This leads to increased congestion.

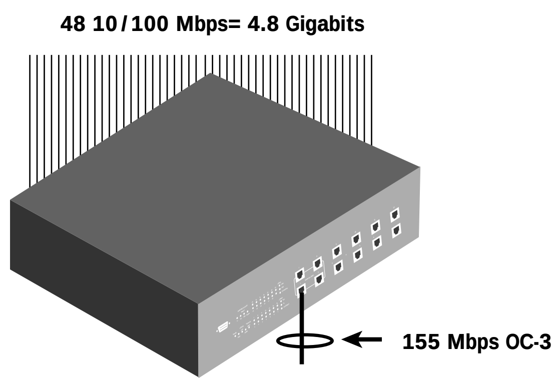

Figure 7: Over-subscription ratio 30:1

Over-subscription Ratios

Figure 7 represents a widely implemented ATM configuration that will unfortunately always be at risk of congestion. The ATM edge device has a single OC-3 connection to the core. This results in a massive over-subscription ratio of 30:1. You can add an OC-12 link, but that would only bring down the ratio to 7:1. This example points out why it's so important to have an over-subscription ratio of 2:1 or lower when segment switching. By keeping the over-subscription ratio to 2:1 or lower, there will always be enough bandwidth to achieve acceptable and consistent performance under maximum traffic loads.

Layer 3 Switching

The segment is perhaps the most optimal place on your network to deploy Layer 3 switching because subnets are typically comprised of multiple segments. By deploying Layer 3 switching where these segments converge, local traffic stays local and traffic is prevented from needlessly entering the core.

Quality of Service

In addition to application-specific QoS, switched segments require traffic prioritization. High-density desktop connections to feeder hubs, and their subsequent connection to 10/100 Mbps Ethernet switches, create a need to assign priorities to different traffic groups. By applying QoS to 10/100 Mbps ports that support feeder hub connections, you can prioritize traffic groups that pass through specific ports and retain better control over network performance.

Summit4 and BlackDiamond 6800 in the Segment

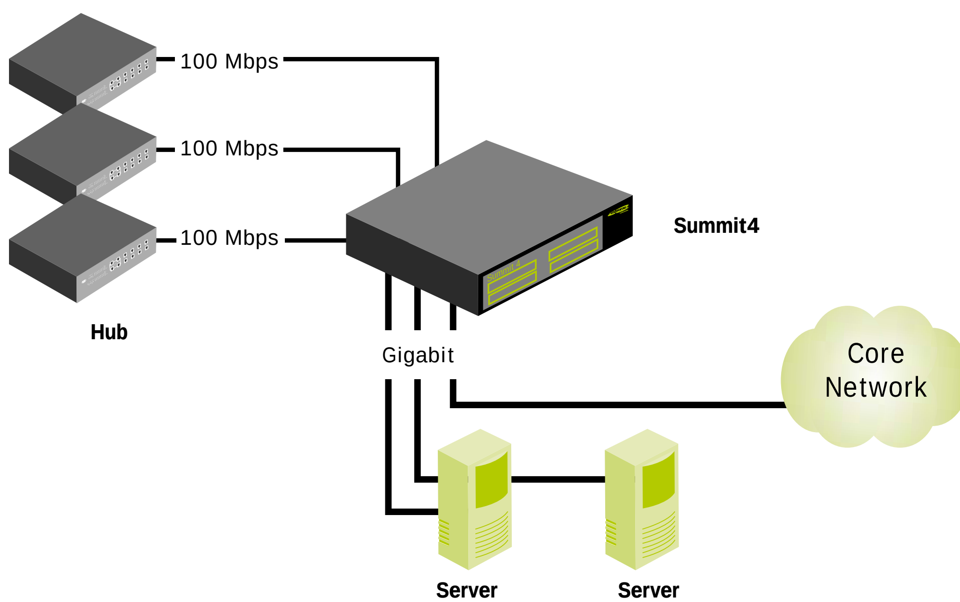

Figure 8

The Summit4 is ideally suited for segment switching. As shown in Figure 8, the Summit4 supports feeder hub segments on its 16 10/100 Mbps Ethernet ports, while the six Gigabit Ethernet ports act as load-shared uplinks and provide server connectivity. This configuration is recommended for smaller networks where Summit4 acts as a collapsed backbone.

As a segment switch, the BlackDiamond 6800 offers a 1:1 over-subscription ratio. The difference is that the BlackDiamond chassis is able to support much higher 10/100 Mbps port densities. The BlackDiamond 6800 supports a maximum of 256 10/100 Mbps feeder hub connections with zero over-subscription.

The Server

Integrated server switching includes server farms, server clusters, database back-ends, web hosting and server backup networks. These topologies require high bandwidth server-to-server connections as well as client-to-server connections from the backbone.

Although the trend has been to centralize servers, recent studies indicate that 66% to 75% of servers are still decentralized. This makes it a challenge to balance client-to-server and server-to-server traffic models, as shown in Figure 9.

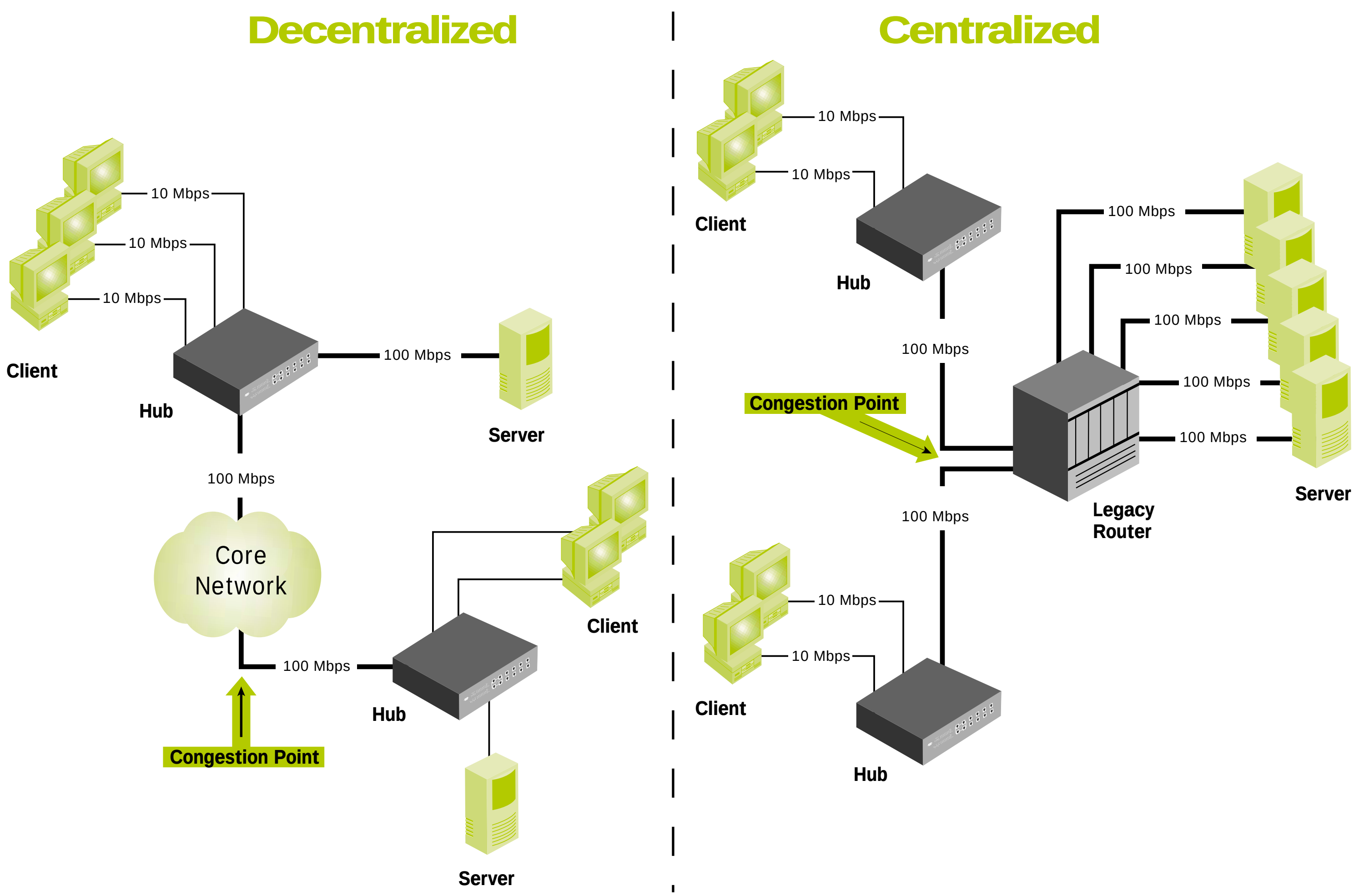

Figure 9

When decentralized server groups are used to optimize client traffic, bottlenecks occur between servers because the core connection becomes improperly over-subscribed. When optimizing traffic between centralized servers, client-to-server connections will suffer due to the performance limitations of legacy routers.

The resulting decrease in server and application performance is caused by limitations in contemporary network architectures. If the proper traffic balance cannot be achieved in both instances, overall network performance degrades.

Non-blocking Extreme believes that non-blocking and no over-subscription is critical in designing a solid server-switching environment. Non-blocking switches are critical to balancing the server-to-server traffic patterns. By having no oversubscribed connections to server groups – whether client-to-server or server-to-server – you won't have to worry about traffic patterns well into the future. This allows you to focus on optimizing applications through quality of service.

Layer 3 Switching Layer 3 switching in server farms produces a more resilient design. By keeping server traffic to their local subnets, multipath routing can be used in high-availability designs. In addition, server-to-server or multicast server traffic can be partitioned from the rest of the network. And finally, there is no longer a need to decentralize servers.

By providing penalty-free Wire-Speed IP Routing, servers can be placed in clusters or farms and still perform as if they were local to the users. The data center infrastructure can be leveraged to its maximum potential.

Link Aggregation Extreme Networks switches allow you to take advantage of network interface card (NIC) link aggregation, which is now available with some server platforms. By grouping four 100 Mbps Ethernet NICs into one logical link, you can create up to 800 Mbps of throughput between the server and the switch.

Network managers can also expand switch-to-switch bandwidth by utilizing Extreme Networks' link aggregation capability. With this feature, network managers can aggregate up to four full-duplex Gigabit Ethernet ports between Extreme switches into a single, 8 Gbps logical link.

Quality of Service Quality of service in the server switching space is best implemented by giving mission-critical applications higher priority and bandwidth. For example, you can define a profile that gives priority to server-to-server traffic. This is vital to database clusters because one server might need access to multiple back-end resources to complete a request. By implementing QoS here, you regulate the traffic flow in favor of server-to-server traffic and increase the database cluster's ability to service requests.

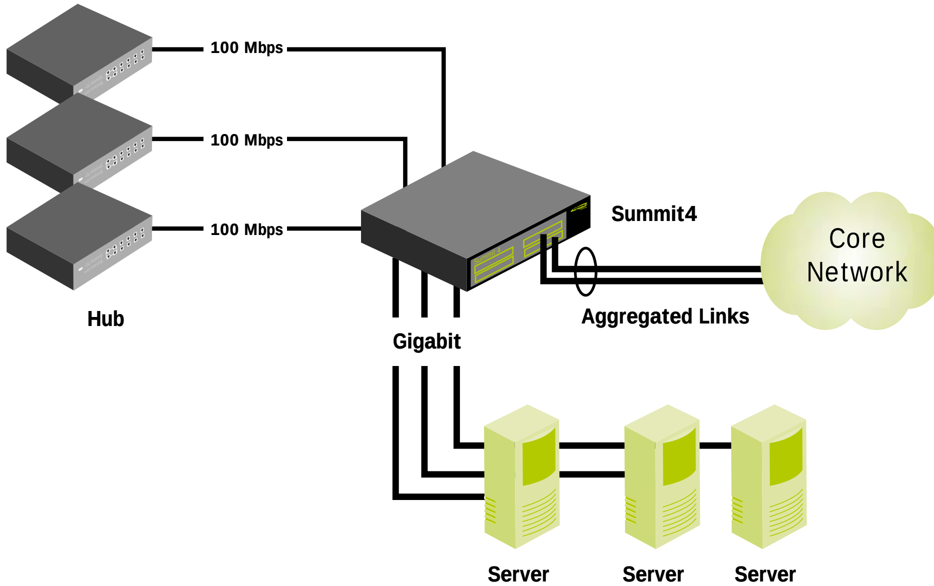

Figure 10

Summit4 and BlackDiamond 6800 for Server Switching

The Summit4 is designed to meet integrated server switching requirements. As shown in Figure 10, the Summit4 supports switched segments on its 10/100 Mbps ports while providing Gigabit Ethernet links to the servers. This gives the servers plenty of bandwidth locally and to the attached segments. However, Figure 11 shows a reversal of the previous Summit4 configuration. By using the multiple Gigabit Ethernet links for attached segments and risers, and 10/100 Mbps links for servers, you can create a configuration that is bandwidth-optimized for client-to-server traffic.

As a chassis-based switch, the BlackDiamond 6800 can be configured to perform the same server switching capabilities as a Summit4, but in environments that require higher port density and modular configuration flexibility.

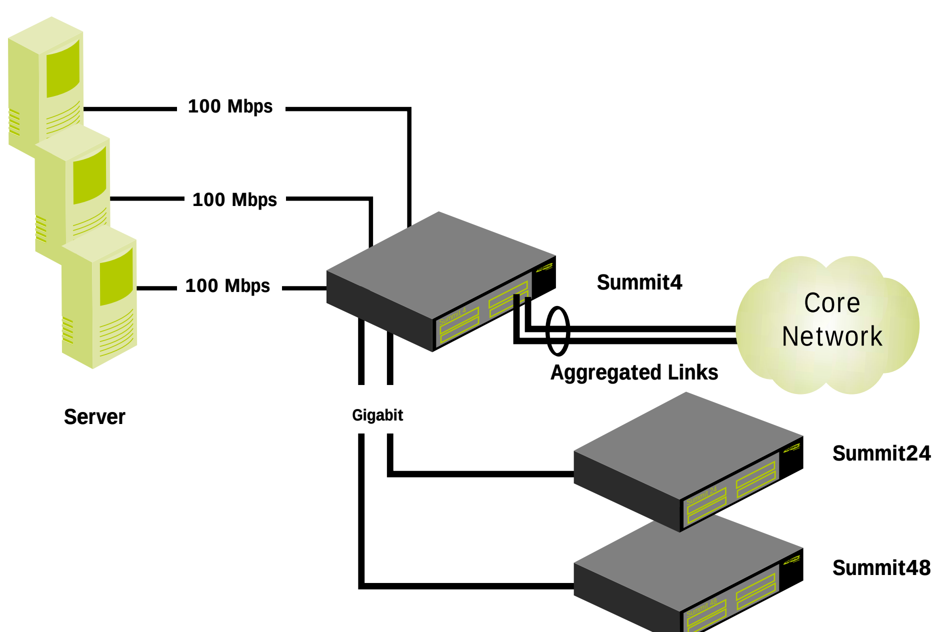

Figure 11

The Core

Core switching is the most critical point in network design. The core supports the aggregate traffic load produced by the network. This is where many networks fail. Wire-speed routing and switching, massive capacity and scalability characterize requirements throughout the core. As a result, core devices should be non-blocking and have no over-subscription. There's no point in surrounding the core with high-capacity feeders if the core cannot handle the traffic load.

Performance

The biggest challenge in the network core is routing performance. The core is where the most intelligent devices reside and is where most traffic converges. However, software-based core routers impede performance due to their blocking architectures and limited packet-per-second forwarding rates.

In the Extreme view, the network core design must meet some key requirements – increase routing performance to wire speed, have a non-blocking architecture to support it, have zero over-subscription and preserve proper over-subscription ratios from the edge into the core.

Layer 3 Switching

Layer 3 switching in the core should be properly balanced with Layer 3 switching at the edge. In the core, Layer 3 switching lets you optimize your network's logical structure. At the edge, Layer 3 switching lets you efficiently organize end-user communities. In most networks, the core is still the best place to route. Routing in the core with Layer 3 switches preserves legacy IP address structures and increases performance by an order of magnitude, thanks to hardware-based routing.

Quality of Service

The core is the easiest place to implement quality of service because the widest range of traffic groups is flowing through it. Application prioritization is the primary reason you need quality of service in the core. For example, a delay-sensitive manufacturing application that runs through the core can be given a high priority and 10% guaranteed bandwidth at all times. This type of traffic segregation and prioritization ensures your mission-critical applications perform optimally and with minimum delay.

BlackDiamond 6800 in the Core The BlackDiamond 6800, which routes 48 million packets per second, is designed specifically for high-density, high-reliability core applications. With up to 48 Gigabit Ethernet ports, or up to 256 10/100 Mbps Ethernet ports, it acts as a collapsed backbone that aggregates multiple segments and risers into a central star, as shown in Figure 12.

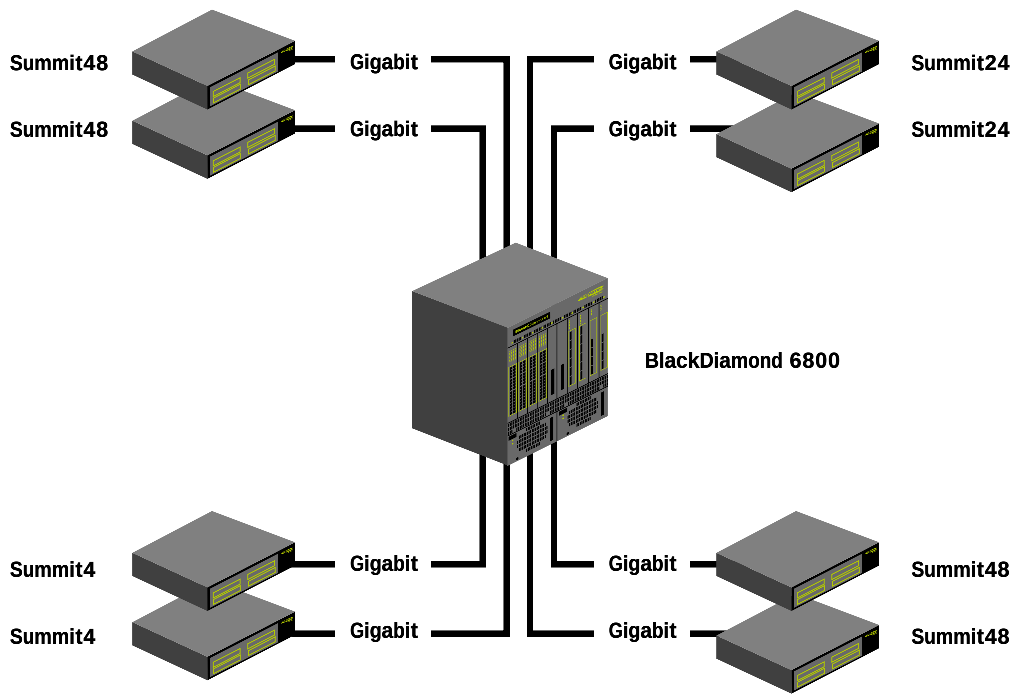

Figure 12

Summit1 in the Core The Summit1, which routes 11.9 million packets per second, is designed for smaller or lower density core applications or for distributed core designs. With eight Gigabit Ethernet ports, it acts as a collapsed backbone that aggregates multiple segments and risers into a central star, as shown in Figure 13.

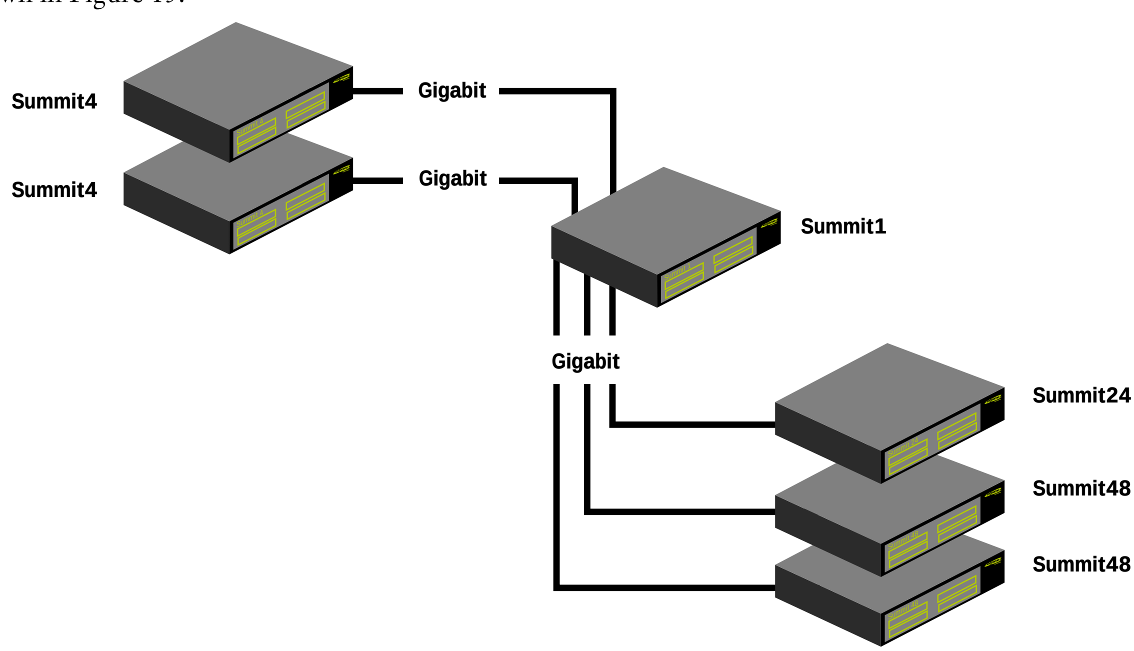

Figure 13

3585 Monroe Street Santa Clara, CA 95051-1450 Phone 408.579.2800 Fax 408.579.3000 Email info@extremenetworks.com
Web www.extremenetworks.com

Rev 2/00   EN-212 Network Design