Ethernet Switching

A Technology White Paper

Madge Networks, October 1996

Summary

This document is intended to provide a technology backgrounder, description of the LANswitch Backbone Switch and Visage Workgroup Switch architectures and a brief theory of operation. The reader will, upon reviewing this document, have a clear understanding of the operational characteristics and technical/functional advantages of Ethernet switching as implemented on the LANswitch and Visage switches.

Contents

Introduction

The term “LANswitch” is the name given to the switching architecture that was developed as part of the LANswitch family of intelligent switching hub products. It is comprised of both architecture (the modular chassis-based 1.28Gbps high-speed backplane, called the Cellenium bus) and products (LANswitch switching modules).

The initial LANswitch hub, the LET-36, was the first to have, as part of the inherent backplane design, a high-speed bus. It was the first hub in the industry to implement the concept of a “generic” high-speed bus for flexible adaptability to many different network technologies, as well as the use of silicon-based (ASIC) switch processors. Even though the hub began shipping in 1991, no modules (LANswitch switching modules) used this bus until mid 1993, with the exception of some early FPGA (field programmable gate arrays)-based prototype modules used for test and evaluation purposes. The fact that the Cellenium bus will provide the necessary bandwidth for most network user needs into the next millennium is testimony to the foresight of Madge’s design teams and engineers.

There is a large variety of LAN technologies supported by the family of LANswitch modules, including switched Ethernet, Fast Ethernet, FDDI and ATM (Ethernet-to-ATM edge adapter), as well as support for all common media types (10BASE-T, 10BASE-FL, 10BASE-FB, 100BASE-TX, 100BASE-FX). The LANswitch modules are managed using an i960-based management agent module that passively interfaces with the Cellenium bus when inserted into the chassis. The ability to do this comes as a direct result of the unique architectural characteristics of the Madge Ethernet switches.

The Visage family of stackable switches from Madge, introduced in June 1996, was developed to create a new class of scaleable workgroup switches. The goal was to create a flexible, fault-tolerant, high-performance and fully managed workgroup switching architecture, in a stackable form factor, that can accommodate switched Ethernet, Fast Ethernet, and ATM simultaneously, without the “typical” limitations of similar products on the market today, and at a price-per-port that is as close to shared hub per-port prices as possible. The end result is a product that, when combined with the LANswitch backbone switch, provides an end-to-end network with the performance, manageability, features and price that is unmatched in the market.

The Visage product family of workgroup switches consists of five different models, supporting switched Ethernet, switched Ethernet with Full-duplex Fast Ethernet uplinks (copper/100BASE-TX and fiber/100BASE-FX), switched Ethernet with ATM uplinks, and switched, autosensing Ethernet/Fast Ethernet, respectively. Each Visage unit is self-contained, and can be stacked up to 4 units high, interconnected with a high-speed, non-blocking interconnection (Exoplane), in any combination. They are managed via a modular agent in a form factor that allows insertion into any unit, with a design that also allows for redundancy (second agent) within the stack. Similar to the LANswitch management implementation, this agent passively interfaces with the Exoplane high-speed interconnect system. Each stack can accommodate up to 96 fully switched 10BASE-T ports without any performance limitation

Technology Backgrounder

"Switching". This is a networking term that is being heard more and more throughout the networking industry, especially with the advent and subsequent popularity of ATM. In many cases, LAN vendors are using the term in unorthodox ways to stay competitive, but in fact, end up confusing the customer. So what does "switching" really mean? Lets take a closer look.

Dictionaries define switching as: "To connect, disconnect; to shift, transfer, change or divert." In terms of LAN-based switching, this would be the purest, simplest definition. The problem is that there is more than one kind of LAN switch. This is where the LAN vendors typically take liberties in describing their "switching hubs".

The easiest way to understand the differences is to divide the types into two groups: Static Switches and Dynamic Switches. Basically, their names are self-descriptive. Static switches require some kind of manual intervention to function, and dynamic switches function automatically. There is also a less obvious difference in the two switch types, but it is one that is even more important. It has to do with the fact that the two switch types handle LAN traffic in completely different ways. Static switches simply move users (see definition above) from one shared LAN to another shared LAN. This gives LAN administrators the ability to address and handle the adds, moves and changes inherent in today's ever changing corporate environments. This type of switching, however, does not address the issue of bandwidth (performance) that many vendors would like you to believe. The only way a user will get better performance with this kind of switching is if they are moved from a busy shared LAN to a less busy shared LAN.

Dynamic switching, on the other hand, automatically forwards packets between two LAN resources "on-the-fly", giving each user the performance equivalent of having their own private LAN. This is the inherent functionality of ATM. These session-oriented connections are sometimes referred to in the Ethernet world as "private Ethernets". In these types of switches, there is little or no contention for the network. The "collisions" normally associated with Ethernet LANs are substantially reduced or eliminated, giving each user access to the full bandwidth capability of the LAN (e.g., for Ethernet, a full 10Mbps). It is important to realize that in these "segmentless" networks, the concept of segments or rings (shared LANs) is no longer applicable. In dynamic port switches, all communications occur on a point-to-point basis. There is another type of dynamic switch that deals with segment connectivity, referred to as a "segment switch". These were initially designed to replace bridges with a high-speed switching matrix. In these devices, the shared LANs still exists, they are simply interconnected with a dynamic switch.

Within dynamic switch types, there are two functional classes: *dynamic port switches*, and *dynamic segment switches*. The differences between the two types, in most cases, is relatively clear. Dynamic port switches automatically allocate dedicated bandwidth (bandwidth-on-demand) between two individual ports (users in peer-to-peer networks, or a user and server in a client-server environment). Dynamic segment switches are, in essence, a high-speed replacement for bridges (for LAN interconnections), eliminating the bottlenecks inherent to software-based internetworking devices. The bottom line difference between these two kinds of devices is that dynamic port switches are designed to connect users (point-to-point with a source and destination MAC address), and dynamic segment switches are designed to connect networks (with many MAC addresses).

Where the distinction becomes somewhat unclear is when defining the term "segment". Some devices, like the LANswitch modules, are by definition segment switches, with support for more than one MAC address per switched port (8 addresses per port are actually supported). This is due to the fact that they are primarily designed to be port (user)-oriented, as opposed to network oriented. They could support segments of up to 8 users. The difference in how these two kinds of devices are used is ultimately application driven. If an application is bandwidth-intensive, as in client-server, imaging and audio/video applications, the "segments" they reside on must be, by definition, limited to a small number of users in order to maintain an acceptable performance level. As these applications utilize more and more of the network resources, it is not unlikely that they will at some point require ALL the bandwidth that their particular LAN technology can provide, which is where per-port switching is necessary. An example in today's networks would be high-speed workstations doing massive file transfers. Just one peer-to-peer session can use up an entire Ethernet (10 Mbps). It is not impossible to see that as workstations get faster and faster, they will exceed the throughput capabilities of the existing infrastructure, requiring a transition to a faster technology, such as switched FDDI, 100BASE-T, and ultimately ATM that has the ability to provide scaleable performance with the inherent Quality of Service characteristics necessary to support evolving applications, such as video conferencing.

Switch Architectures

From a platform design standpoint, there are basically three architecture types associated with dynamic Ethernet switches. They are described as follows:

A. Standalone Hardware-based "matrix" or "cross-point" switches

These are standalone, hardware-only devices that exist as "black boxes" on the network (Figure 1). The most common implementation is the "cut-through" type of device, where packet processing involves destination address checking only, with no error or format checking. The simplicity enables hardware-based switches the ability to provide high-speed switching with very low latency. These devices typically consist of a "matrix" of point-to-point 10Mbps lines which are switched dynamically.

- ADVANTAGES: High performance; simple design; plug-and-play operation; inexpensive.
- DISADVANTAGES: Limited aggregate throughput; limited port density; typically limited to a single technology; difficult to adapt to other technologies (Ethernet - Token Ring, Ethernet - FDDI, Ethernet - 100BASE-T, ATM, etc.); no upgrade path; limited or no management capability (Virtual Networking); no error checking; packet loss under continuous load conditions.

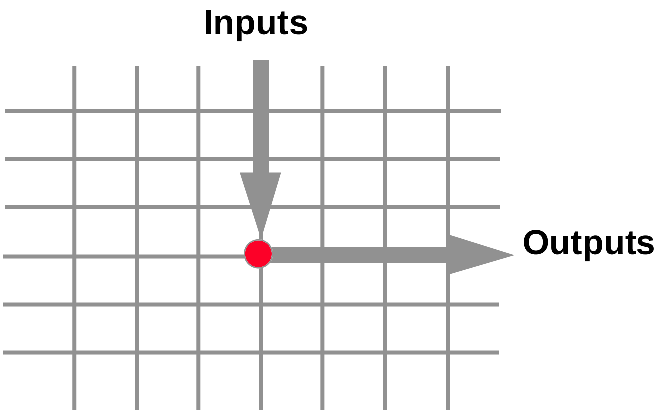

Figure 1 Matrix switch Architecture

B. Centralized software-based (router/bridge) switches

Centralized software-based switches utilize a central processor as a "switching state machine" (Figure 2), enabling packet processing similar to a router, and typically utilize a common inter-module communication technology, such as FDDI. Due to their routing capabilities, these switches can filter various protocols and error packets during transmission.

ADVANTAGES: Modular; intelligence allows filtering capabilities.

DISADVANTAGES: Limited backplane (inter module) speed is a bottleneck; limited port density; slower speed due to software-based design; central point of failure; central processor creates a performance bottleneck; packet loss under heavy loads.

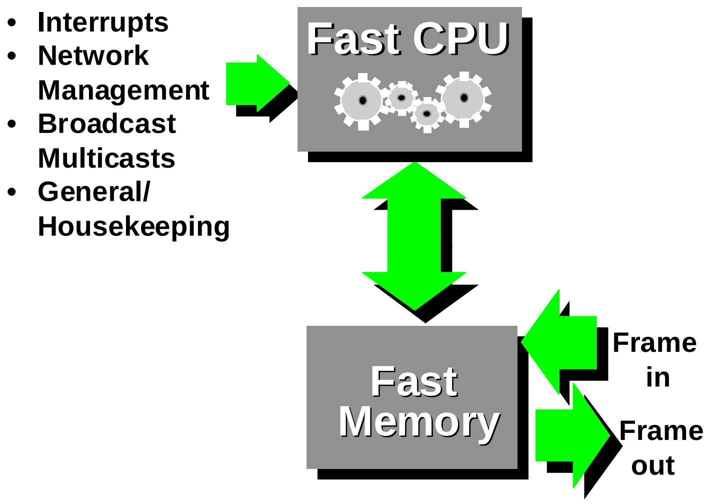

Figure 2 Centralized Software Based Switch Architecture

C. Distributed hardware switches with high-speed backplane

These switches are based on a design of distributed hardware processors (typically custom ASIC's or VLSI's) which dynamically set up point-to-point session connectivity via a common high-speed backplane or bus (figure 3). The distributed processors increase scaleable bandwidth as they are added to the overall system, limited only to the aggregate backplane speed. This packet-based high-speed backplane provides a "common denominator" between protocol types.

- ADVANTAGES: Modular, scaleable architecture offers both flexibility and performance (more modules (switches) = more performance); distributed design provides fault tolerance (no single point of failure); hardware-based switching provides high performance with little or no latency; easy integration of multiple protocol technologies (Ethernet - Token Ring, Ethernet - FDDI, Ethernet - 100BASE-T, ATM, etc.); high port density.
- DISADVANTAGES: None.

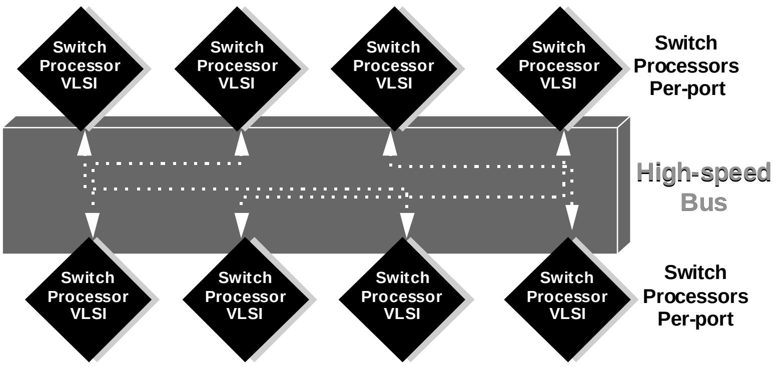

Figure 3 Distributed hardware switches with high-speed backplane

Switch Forwarding Techniques

There are two different data forwarding techniques commonly used today (Figure 4). The first is the cut-through design that only looks at a small portion of the packet. This provides very low latency, but at the expense of little or no error checking , as well as difficulty scaling to high-speed technologies.

The other technique more commonly used is the "store-and-forward" technique. This provides better error checking and scalability, but at the expense of performance degradation due to the delay introduced by storing the packets and having a CPU decide what to do with it.

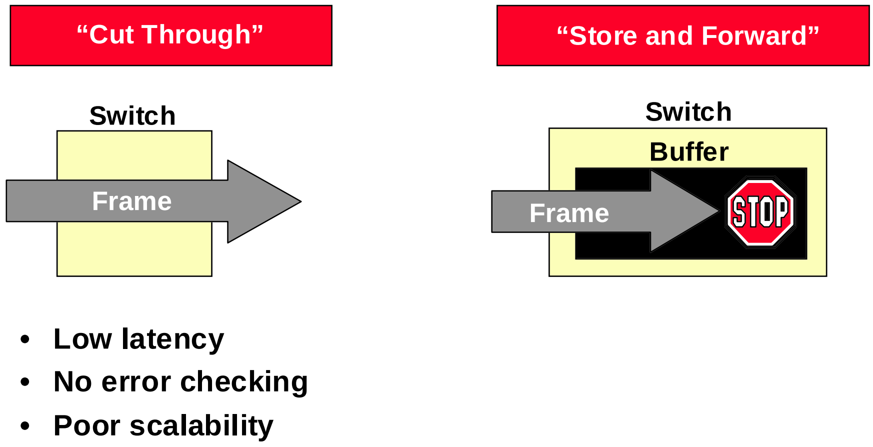

Figure 4 Forwarding Techniques: Cut through and Store and Forward

The LANswitch and Visage switches, on the other hand, uses a technique called "store-zero-forward", or sometimes called modified store-and-forward, only possible with the use of a high-speed backplane and distributed processors (Figure 5). The packet is only stored long enough to determine if the packet is good, and then is forwarded with "zero delay" These forwarding decisions are happening in parallel by the distributed processors so that no performance degradation is encountered. The end results are low latencies of cut-through switching with the error checking and scalability of store-and-forward, realized in a single implementation.

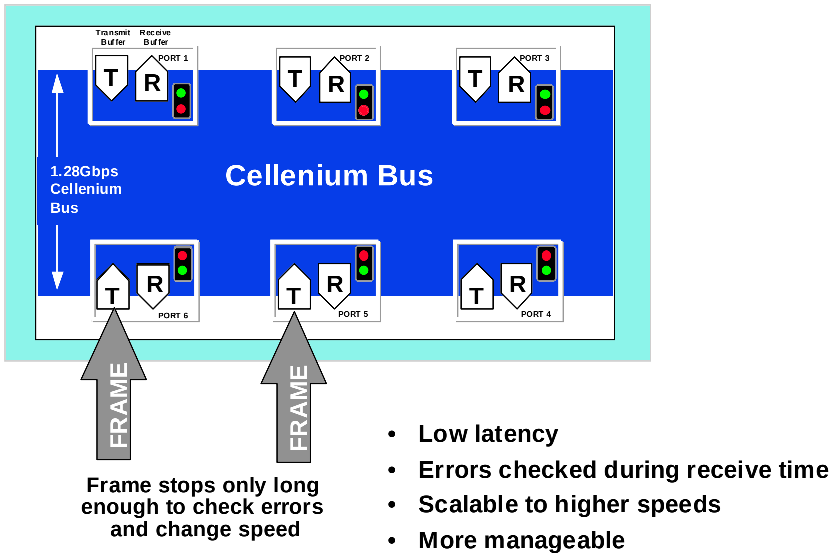

Figure 5 Store-Zero-Forward Forwarding Technique

The appropriate use of a particular forwarding technique is best dictated by the specific advantages and disadvantages of each LAN protocol. In Ethernet switches, the ability to use a store-and-forward technique is important in order to eliminate the propagation of errors. And because the maximum size of an Ethernet packet is not very large, the delay associated with storing the packet is predictable and minimal. Token Ring, on the other hand, is more appropriate for the cut-through technique due to two primary reasons. First, the packets can be quite large, so storing these potentially huge packets can introduce unacceptable delays. And second, the Token Ring protocol has many inherent abilities to handle errors that don't exist in Ethernet, so error checking requirements are minimal, eliminating the need to store the entire packet.

Topologies

One of the major advantages of switching technology is the physical layout of the switched networks themselves. Due to the ongoing "microsegmentation" that today's networks are going through to address the bandwidth requirements of the users, the networks begin to look like mazes, consisting of many internetworking devices (Figure 6). These networks are not only inflexible, but they are extremely difficult to manage and control, and can suffer from delays introduced by these devices. The costs in both man-hours and salaries to support these kinds of networks can be enormous, due to the fact that most of the adds, moves and changes, as well as general administration, must occur after hours, so that disruption to the production network is minimized. After-hours work means overtime in most cases, which only serves to increase the costs. As the networks get larger and more segmented, these "hidden operation costs" increase exponentially.

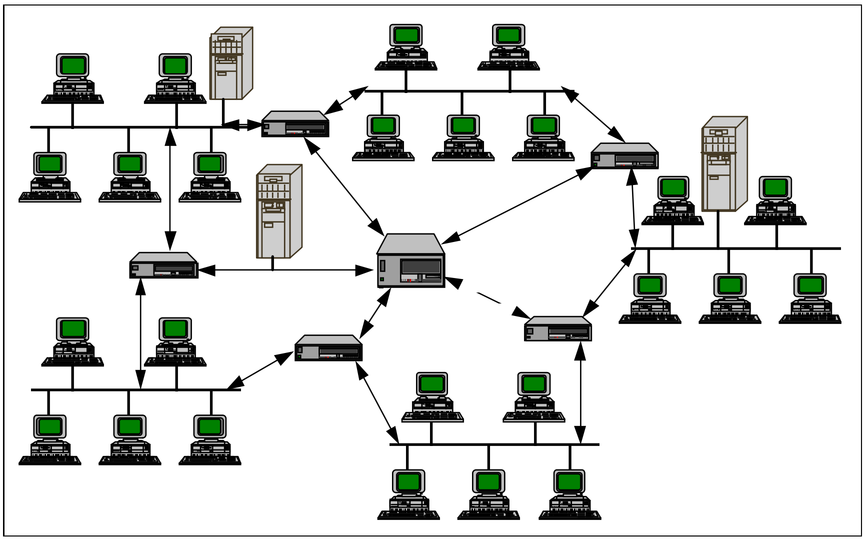

Figure 6 Microsegmented networks become very complex and difficult to manage

In switched networks, the topology, and therefore the design, become much simpler to implement and manage. In the horizontal, everything is "star-wired", and due to the point-to-point nature of switched network sessions, bandwidth becomes a non-issue (Figure 7). Typically, when designing networks, certain assumptions must be made up-front as to the most limiting factors (i.e. delays associated with routers and bridges, potential bottlenecks, etc.). A network is only as fast as its slowest component. With switched networks, most of these limitations are minimized or eliminated altogether. Internetworking devices, especially routers, can be used for the reasons they were initially designed - interconnectivity between disparate LANs and/or WAN connections. With fewer of these internetworking devices to contend with, the networks become much easier to manage, and are infinitely more efficient and flexible.

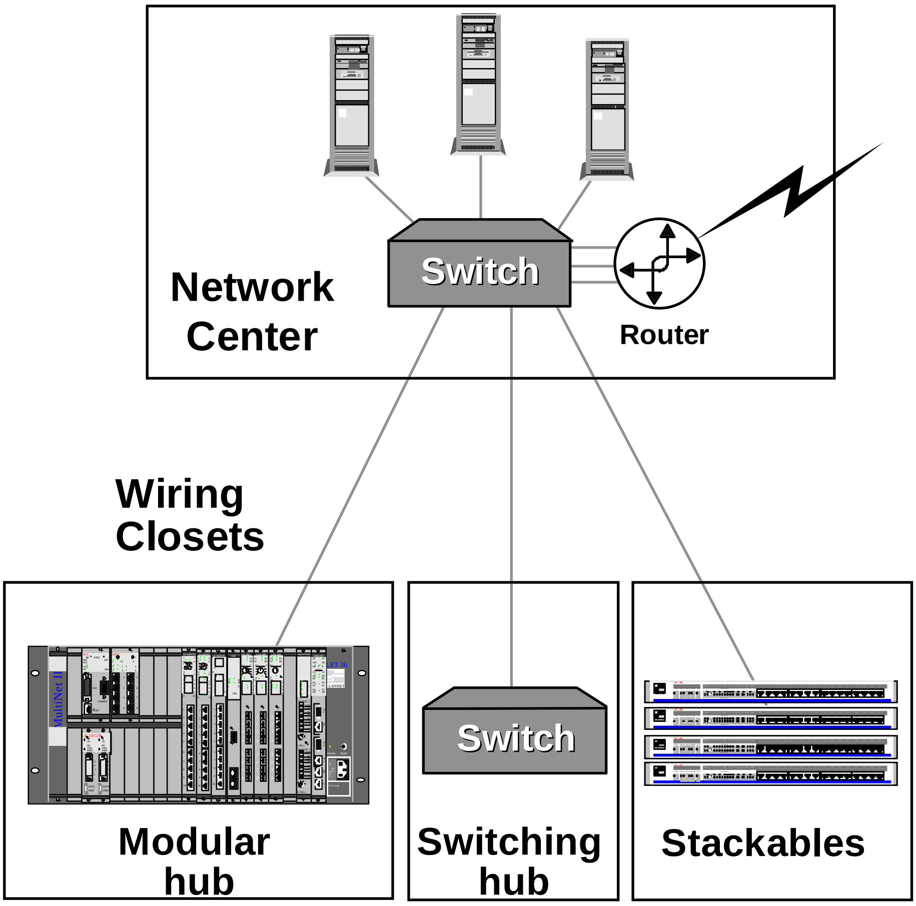

Figure 7 Switched Networks Have Star-Wired Topology

LANswitch Architecture Description

The LANswitch is a packet switch, similar in architecture to a parallel processor. It consists of a high-speed backplane (the Cellenium bus), which is a 64-bit wide parallel bus operating at 20Mhz (yielding the aggregate throughput of 1.28Gbps), and the LANswitch family of switching modules, each with their own switch processor-per-port (Figure 8). This creates a fully distributed switching system that actually scales *up* in performance as more modules/ports are added to the system. What is sometimes describe as "64-bit cells" running on the Cellenium bus are not really "cells" at all, but parallel 64-bit words that make up packets of variable length. This is what gives the LANswitch the flexibility to accommodate any access method that is adapted to the system, regardless of packet size.

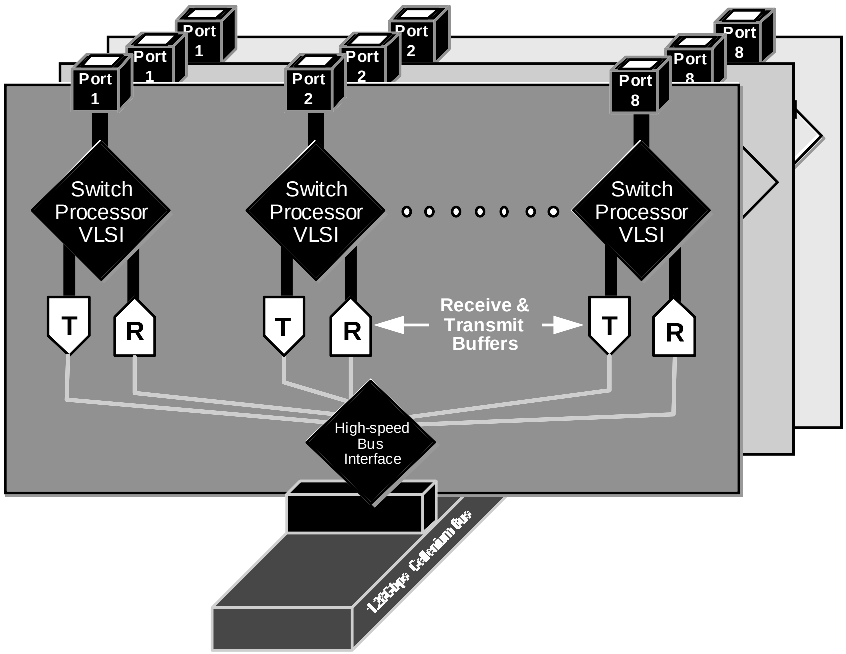

Figure 8 Processor-Per-Port LANswitch Architecture

There are many advantages to this approach that are not inherently obvious. With all the hype about ATM, it seems that unless you have a switch that uses ATM as a backplane data transfer mechanism, you are presumed to have a switch with limited capabilities. This is simply not true. In fact, if the network environment consists of *any* combination of LAN technologies, other than pure end-to-end ATM, an ATM-based backplane is severely limiting for traditional LAN switching due to the translation that must occur.

As an example, switches that use ATM for the backplane must translate every non-ATM packet to ATM and back again. This involves both translation and fragmentation, adding to a degradation in system performance and increases in delays as a result of the unnecessary overhead. Added to this is the issue of LAN Emulation, which adds additional overhead. In ATM networks operating with other LAN types, this software is the mechanism that actually performs the translation from Ethernet/Token Ring-to-ATM and back. It is the necessary integration piece for networks evolving to ATM from legacy (Ethernet or Token Ring) when ATM is implemented in the backbone as a first step towards a total end-to-end ATM network.

On the other hand, the LANswitch uses a more optimized approach, one that is more flexible and adaptable to supporting multiple LAN technologies without compromising performance. Each switch processor with an interface to the Cellenium bus sends data to the bus in the form of parallel 64-bit wide words that make up packets of variable length, regardless of the LAN technology. This means that the Cellenium bus sees and handles packets of any LAN type in the same manner, which gives the LANswitch its unique flexibility, adaptability and manageability. And, because there is no fragmentation, emulation or translation necessary, there is no performance degradation, which gives the LANswitch its robust, scaleable performance, regardless of the mix of LAN types and speeds existing on the Cellenium bus. For networks transitioning to ATM by using ATM in the backbone, the required emulation occurs outside the LANswitch, preserving the performance.

Another aspect of the efficiency of the design is that information regarding virtual LAN and priority are simply added to the packet, or tagged, so there is again no performance degradation. The primary reason the LANswitch is able to accomplish this is that each switched port accesses the Cellenium bus via its own dedicated silicon-based switch processor. This gives the LANswitch distributed, scaleable performance that is unaffected by added ports or with the addition of additional forwarding information (VLAN and priority) that must be processed. All of the forwarding information is being processed in parallel in the distributed switching ASIC’s.

Visage Architecture Description

The Visage family of switches is based on distributed switching VLSIs which provide the ability to yield extremely high performance and functionality levels at very reasonable prices. It is a distributed VLSI-based architecture, similar to the design of the Madge LANswitch, providing inherent fault tolerance characteristics, as well as upward scalability for performance and port density.

Each Visage switch contains an internal high-speed switching fabric made up of distributed switching VLSIs, which are in turn interconnected to other switch units via a high-speed 4 Gbps interconnection called the Exoplane, which together form a single, logical Visage switch. The scaleable switch fabric of the Visage family ensures that there is no performance degradation when adding users to the hub or hubs to the stack. The aggregate throughput capability in a 4 unit Visage “stack” is 4 Gbps, thus yielding an actual throughput capacity of 5.8 million packets-per-second.

One difference in the Visage switch is that there is a memory matrix used to create a shared 32,000 MAC address capability that is used for all ports in a switch, regardless of the number of units in a “stack”. With this implementation, there is no difference between port or segment switching ports...all ports have multiple MAC address support for segment and/or port switching.

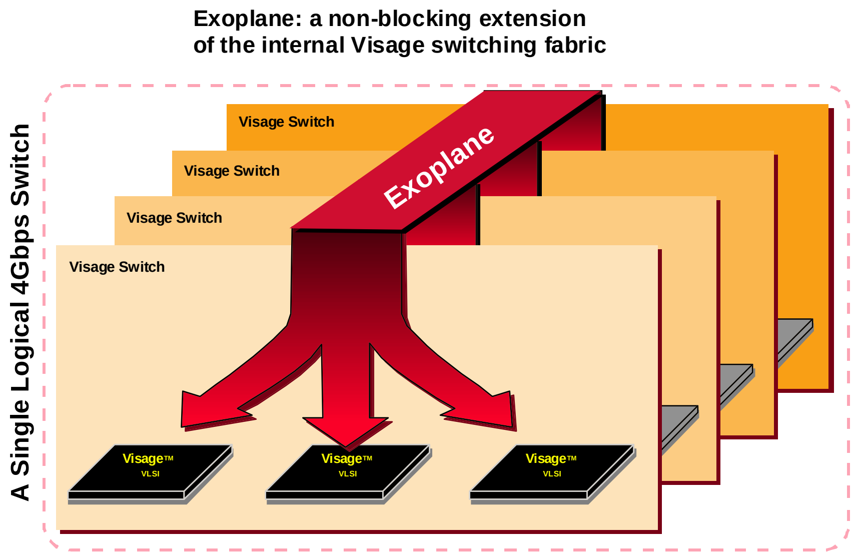

Figure 9 The Visage Distributed Architecture

The Visage VLSI switch on a chip is a technological breakthrough which yields inherent speed and reliability. The 260,000 gates and over a million transistors have a complexity which almost matches that of an I486 processor, and the highest ever density of custom Ethernet switching silicon.

Figure 10 The Visage VLSI, Containing On-Board Management

Advanced Switch Features of LANswitch and Visage

Fairness

In Figure 11 you will see a diagram that describes the reason the LANswitch and Visage have such robust operational characteristics, that is, one that does not break down under heavy network load and provides the same robust, fair operation as Ethernet. In this case you see six different ports, each consisting of a buffer for transmitting data and one for receiving data. In addition, there is a bus arbitration or access mechanism, indicated with a “stoplight” symbol, that allows each port that needs the use of the bus equal and fair access to it. The overall effect is a powerful switch that stays balanced so that no port is ignored and no port can monopolize the bus.

It also works much like a statistical TDM, in that ports that do not need access to the bus are simply bypassed so that no unnecessary delay is introduced. This makes for very streamlined, efficient switch operation.

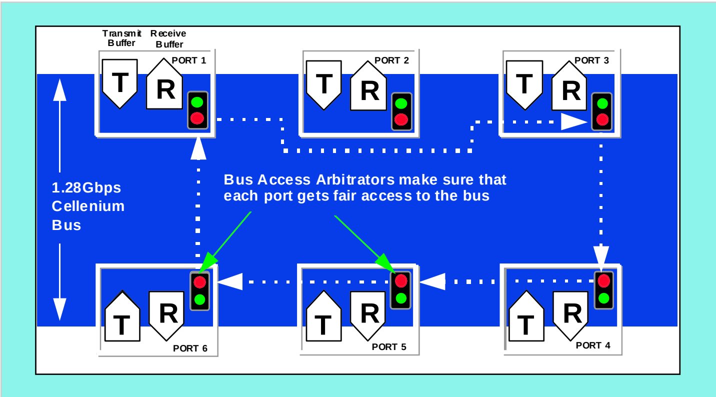

Figure 11 LANswitch and Visage Fairness Mechanism

Priority

In Figure 12, we have the same basic diagram, but we have now introduced the concept of a port that has been put into a higher priority mode to deliver the required higher Quality of Service (QoS). The basic mechanism is one that provides priority data delivery, without affecting the overall fairness of the system. If this mechanism were not able to maintain fairness, the system would break down, losing data and dropping network connections that are sensitive to delays.

As we see, the feature essentially changes the rate at which the prioritized port has access to the high-speed bus, and in addition, only effects the ports that are communicating with the prioritized port. Other ports on the switch would act as in the previous diagram. In this simplified diagram, the prioritized port transmits to the bus, then port 2, then the prioritized port, then port 6, and so on. The end result is that ports receiving data from the prioritized port will receive more of the higher priority data then of lower priority data, maintaining the necessary quality of service for the prioritized data without blocking the other ports. This rate of data prioritization can be controlled or adjusted via network management software by selecting of any of 4 (for LANswitch) or 2 (for Visage) available priority levels.

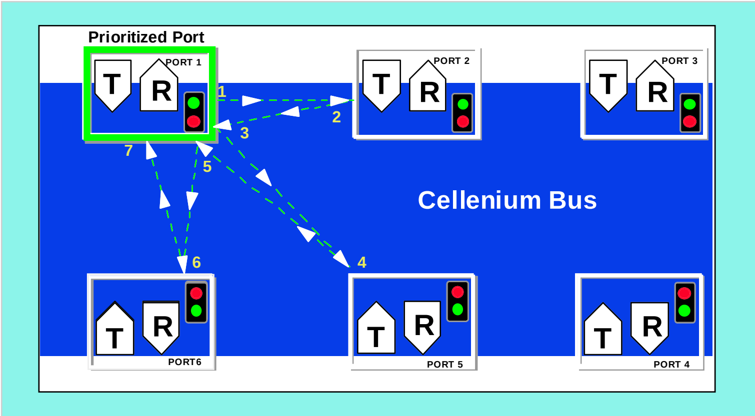

Figure 12 Priority Data Delivery

Congestion Management

Most real-world networks are based on client/server architectures. Putting a switch in the network to provide dedicated bandwidth to each client will only go so far when you consider that all the clients are communicating with a single server. At this point, it becomes clear why an effective method for handling this congestion problem is necessary. The LANswitch and Visage switches do this in a very efficient way, using a two-step process.

In Figure 13, you see the basic diagram again. You will notice that the fairness mechanism is still there, always working in the background. Ports 2, 3 and 6 are clients communicating with a server that is attached to port 4. In the first step, there is communication occurring between the associated receive and transmit buffers of the ports that updates the status of each buffer. Transmit buffers are filling and transmitting data to receive buffers that are filling and emptying. Remember, this is happening at 1.28Gbps, resulting in a “buffer depletion rate” that is over 10 times faster than Ethernet. When this is considered, along with the fairness mechanism working in the background, the chance for a congestion situation that cannot be accommodated by the system is extremely rare. It should also be noted that this capability is completely independent of the technology outside of the bus, as it only dealing with data in the form of “words” on the high-speed bus. Another aspect of this mechanism is that in the high-speed (100Mbps) switch ports, the buffers are considerable larger, and additionally enhance this capability.

But there are cases where this congestion problem can exceed the capabilities of the smart buffering system working in concert with the high-speed bus, especially with many very fast workstations communicating with a single server.

When this happens, the 2nd step is initiated. In this case, both the transmit and receive buffers have filled and cannot accept additional data. In other switches, this scenario would cause a significant packet loss, mostly from the resultant collisions and retransmissions. In the LANswitch, a throttling mechanism is employed, consisting of a standard Ethernet carrier sense signal sent to the sending station. The end effect is that the station, for a very brief time, senses that the network is busy, and only transmits when the busy signal is gone, which means no collisions and therefore no retransmissions. This can cause severe performance degradation due to the fact that packets are typically retransmitted in “blocks” of 8 packets or more, not just single packets. The result is that no data is lost under any congestion situation, as has been successfully tested by several third-party, unbiased external test facilities under very rigorous conditions.

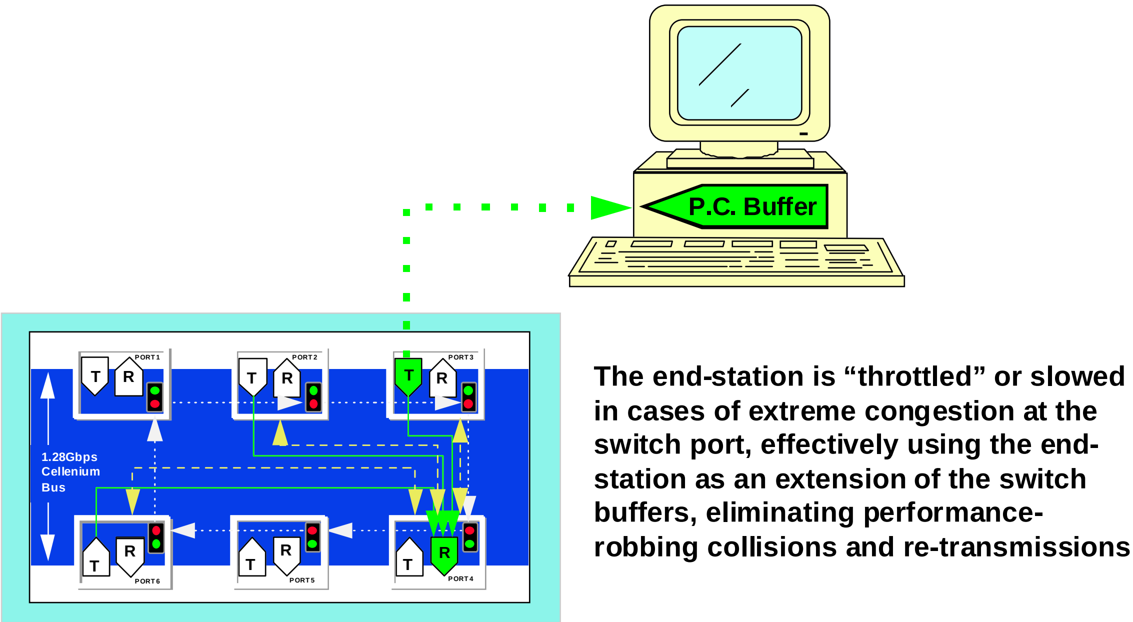

Figure 13 LANswitch and Visage Congestion Management

Layer 3 Switching

Switches place users who are connected to different physical segments on one virtual segment. The main advantage of Virtual segmentation is that Network design becomes independant of physical location. When high speed communication between different logical groups is required, Layer 3 switches allow you to go one step further. They give the freedom to place users who are defined on different logical subnets or virtual LANs on one virtual segment, thereby freeing the network from logical constraints imposed by network layer architecture. Layer 3 switches expand the capabilities of hardware based switches to encompass functions that are currently handled by slower, software based routers.

While routing requires the division of a network into multiple logical subnets, a Layer 3 switch based network can be much flatter and consequently faster. Logical groupings can be set up as needed based on purely organizational requirements, in the form of virtual LANs. A well designed switched network minimizes the time sensitive traffic that needs to be routed. When routing is needed, Layer 3 functionality provides fast and virtually delay-free connectivity between Virtual LANs, keeping the separate broadcast domains

Layer 3 Architecture Description

A Layer 3 switch should forward packets very quickly, and also have routing intelligence. In the LANswitch Layer 3 switch this is accomplished by splitting the Layer 3 Switch into two distinct functions: fast forwarding and routing intelligence. One half performs all the traditional routing functions. The other half comprises a fast forwarding switching engine which performs the minimal set of routines required to accurately switch the data to its destination. These routines are implemented in hardware (VLSI) so that switching is performed transparently, while the latency added by the routing function is minimized.

The LANswitch Layer 3 switch has a distributed switch architecture based on Layer 2 switching, combined with a very fast backplane operating independently from the switch. It is this that enable it to provide instantaneous communication between source and destination ports.

Flatter network design coupled with efficient broadcast handling is the right approach for better network services. Layer 3 switching is required for switched networks which require inter-Virtual LAN communication, and for organizations who choose to keep their large numbers of desktops unchanged and don’t want to modify their layer three structure today. As inter-subnet traffic levels grow so that they can no longer be efficiently handled by a router, Layer 3 switching should be considered. One of the notable advantages of integrated fast-forwarding (switching) based on Layer 3 protocols is the switching between Virtual LANs. Layer 3 switching provides a very efficient solution that includes those router functions that are still necessary for the network without investing the funds and paying performance penalties for unnecessary features.

The LANswitch Layer 3 switch architecture has a number of advantages. Firstly, it eliminate all dependency on a central processing point (i.e., a central router). The routing and switching functions are completely integrated into a unified product for better performance and reliability. Being an integrated unit, it is not dependent on a central router, nor is it affected by the speed and status of the links connecting the router to the Layer 3 switch. Because the switch incorporates both routing and Layer 3 switching, it is better able to maintain an accurate, real-time database. Another advantage of the LANswitch Layer 3 architecture is that both the switching and the routing functionality enjoy the full internal bus capacity. An external router is limited to the speed of the physical interface. The Layer 3 switch guarantees that only one fast-forward Layer 3 process is required when communication is either local or spans multiple hubs.

By integrating Layer 3 functionality, the Layer 3 distributed switch architecture guarantees redundancy in case of failure and allows for load sharing between multiple Layer 3 switches. The design provides fault-tolerance options, and distributes routing functions amongst the switches. Fault-tolerance ensure no single-point of failure, while distributed routing allows for better performance by better utilizing available resources.

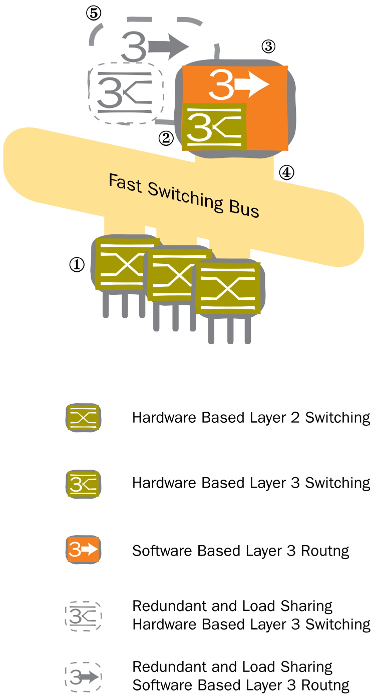

Figure 2 The Madge LANswitch Layer 2 and 3 Switching Integration

Figure 2 explains how the Madge LANswitch Layer 2 and 3 Switching Integration works:

1. Distributed Layer 2 switching is the Core of the Madge switching architecture. MAC layer switching will always be the first choice.

2. Layer 3 switching is the next level of the architecture. Packet switching based on Layer 3 addressing is achieved with a hardware based Layer 3 switch which complements the MAC layer switching when wire-speed communication between Virtual LANs or subnets is required.

3. Layer 3 routing is the highest architectural level. It supplements the Layer 3 switching as well as providing Layer 3 functionality.

4. A unique part of the architecture is the direct attachment of each and every component of the architecture directly to the core of the switch- the Fast Switching Bus. Access to the switching fabric is therefore real-time, unrestricted and highly reliable.

5. The switching architecture is fully distributed and redundant. Figure 2 shows the redundant and load-sharing aspect of the Layer 3 component. The architecture allows for multiple Layer 3 switches as well as Layer 3 routing with load sharing.

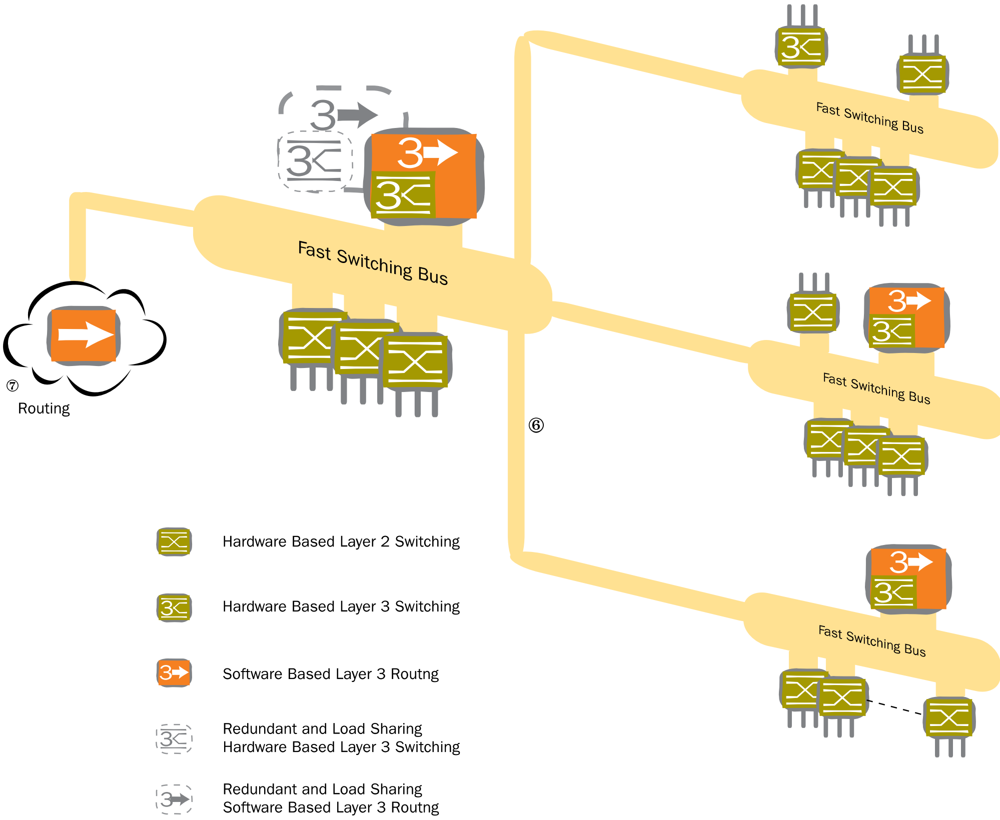

Figure 3 The Madge LANswitch Switching Architecture

Figure 3 shows the Madge LANswitch Layer 2 and 3 switching architecture in a wider context. (The numbers below refer to figure 3)

6. The Layer 3 intelligence guarantees one hop Layer 3 switching. The packet enters the Layer 3 logic only once, and is then forwarded directly to the end stations.

7. The design is fully compliant with all Layer 3 routing standards, so that compatibility with standard routers is assured.

Switch Management

The ability to manage a network is critical, especially a networks become large and/or companies become dependent on them. This is no less true for switched networks. As a matter of fact, this capability is even more important in switched networks where the switches are placed in the center of the network where the bandwidth problems are usually first encountered. The problem arises because a ll of the traditional tools and protocols available to manage networks (SNMP, RMON, sniffers etc.) were not designed to manage switched networks and are therefore only minimally useful. These traditional tools were design to manage networks based on the concept of “broadcast” concepts, where all traffic is sent to a common media, and can therefore be controlled and monitored.

In the case of switched networks, all connections are “point-to-point”, and even more important, these connections happen dynamically. These on-the-fly point-to-point connections cannot be easily managed on most of the switches on the market today, because there is no common area in the switch to accomplish this task without effecting the performance of the switch. The standard network management protocols do not address these connection-oriented networks, either.

As a means of defining what network management does, it can be said that there are basically two tasks to accomplish: *control* of the network (assignment of virtual networks and priority, turning ports on and off. etc.) and *monitoring* (looking at the network and being able to interpret statistics, performance, etc.).

MADGE has taken a unique approach with a new product and concept for managing the LANswitch. It is comprised of three elements: standard SNMP for *control*, and a unique combination of an enhanced RMON application (SMON) and powerful RISC-based agents (NMA-RS and Visage NMA) for processing and *monitoring*. Although a detailed description of this new network management solution is beyond the scope of this document, we will briefly describe it as it pertains to the overall switching architecture.

Control of Switched Networks

Virtual networks are logical groups of users, segmented because of security, organizational or performance reasons. This function has become necessary as switches have been implemented in networks, due to the nature of switched network access. Switches are inherently secure, due to the point-to-point session-oriented connections, as opposed to the broadcast techniques used in shared networks (Figure 14). There is also the inherent capability for any resource to have access to any other resource in the switched network, which may be an issue if there are legitimate reasons to segment users on the network. Virtual networking allows the ability to group appropriate users or resources into "logical workgroups", segmenting them from other users in the switched network. This segmentation is not implemented for bandwidth reasons: the switch takes care of that naturally. The users are grouped for reasons of security or convenience only.

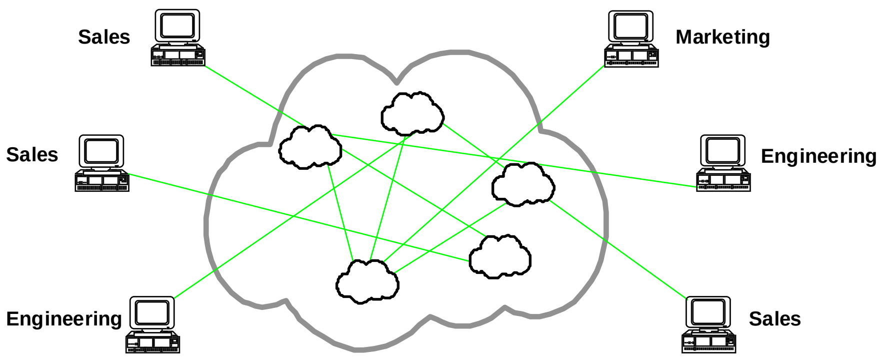

Figure 14 Point-to-Point Communication in Switched Networks

In the LANswitch and Visage switches, this segmentation is accomplished on a per-port basis, creating workgroups consisting of groups of ports and the associated users and resources attached to them (Figure 15). The LANswitch and Visage switches have an additional unique characteristic for virtual networking in that the virtual networks can be created "hierarchically", creating a layered virtual network environment. This can explained by defining the two "classes" of virtual networks: standard and global.

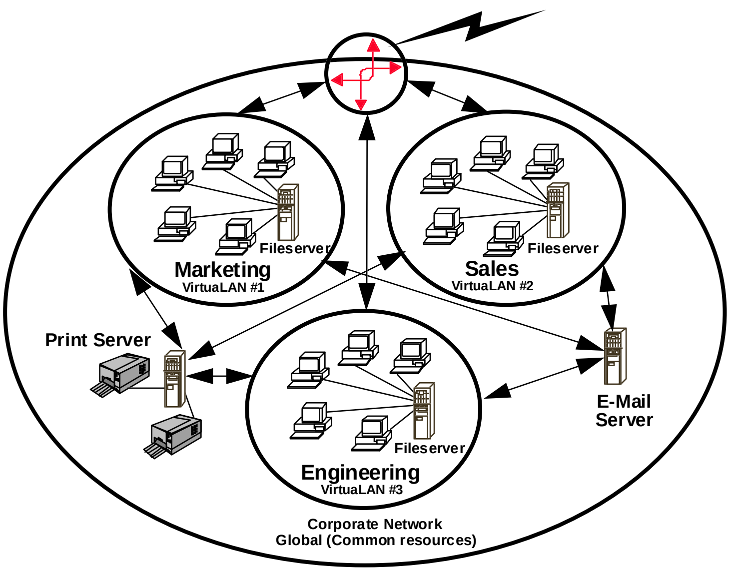

Figure 15 Virtual LAN Segmentation

Standard virtual networks are the segmented groups of ports/users, and by definition, have no access to other virtual workgroups except through a bridge or router used to interconnect them. Traffic is contained within each workgroup. The global virtual network is a “higher layer” network that allows access to all standard virtual networks within it. This can be used as a “common” area where resources or users that require direct switched access to all virtual networks are placed. These can be devices such as common servers or gateways, or common users such as group leaders responsible for multiple workgroups. This global layer allows direct switched access without the performance degradation of bridge/router-based connections, as well as freeing up the router CPU-cycles for true internetworking requirements.

Another function that comes under the heading of network management control is the assignment of the 4 levels of priority (on a per-port basis), discussed previously, as well as being able to disable/enable ports, set redundant ports, and many other functions typical of SNMP-based management.

The major architectural advantage of the Madge Ethernet switch implementations of network management is that they are completely distributed systems, with distributed switch processors as opposed to a central processor. In addition, the information regarding virtual networks and priority are in the form of “tagged” headers appended to the packets as part of the switch forwarding mechanism. This combination of silicon-based forwarding and distributed processors provides very predictable switch performance that is unaffected by the addition of VLAN or priority information.

Monitoring Switched Networks

Monitoring switched networks presents some unique problems that simply cannot be addressed by the conventional, traditional tools and protocols in use today. As stated previously, SNMP and RMON management protocols were designed to monitor “shared” networks, where transmitted data is “broadcast” to the media, making it relatively simple to “see” what is happening on the network. These protocols do not know what a switch is. The disadvantage to this approach is that a monitoring device (agent, probe, etc.) must be on each LAN to monitor the traffic. The problem has been getting worse due to the constant “microsegmentation” that has been going on in order to keep network performance at acceptable levels, requiring more and more of these devices to be installed. This increases both the cost and complexity of these shared LANs.

On switched networks, where every connection is both dynamic and temporary, the problem is significantly exacerbated. If traditional tools were employed, a monitoring device would have to be resident on every port in order to accomplish real-time monitoring. Switches that have CPU-based architectures can rely on the central CPU as the common point to monitor the network, however, it is the CPU itself that must make both forwarding decisions as well as monitor the network connections. This additional processing requirement will have a severe negative effect on the switch performance.

On the LANswitch and Visage switches, the architectures consist of distributed processors on each port, interconnected with a high-speed bus (the Cellenium Bus or Exoplane, respectively) where all switch traffic can be monitored passively without the use of any switching processors. All the power in the monitoring system is in the special agent (NMA-RS or Visage NMA), which is comprised of a fast i960 RISC processor and several custom switching ASIC’s that make up a parallel processing system providing complete analysis of the Cellenium or Exoplane Buses. This management hardware is combined with a unique monitoring application called SMON that runs on top of Madge’s MultiMan management system, which allows continuous monitoring of the switched network on a per-port basis.

The Visage switch management architecture, although similar to the LANswitch, has some different characteristics. The Visage, like the LANswitch, consists of distributed switching silicon (VLSI), interconnected with a high-speed bus (the Exoplane). The management agent, or Visage NMA, consists of an I960 processor and flash memory, but many of the elements that exist on the NMA-RS for the LANswitch are embedded in the silicon of the Visage switch. These elements include RMON counters. The over all effect is that much of the actual monitor processing occurs in the silicon, which optimizes the agent performance and reduces costs at the same time.

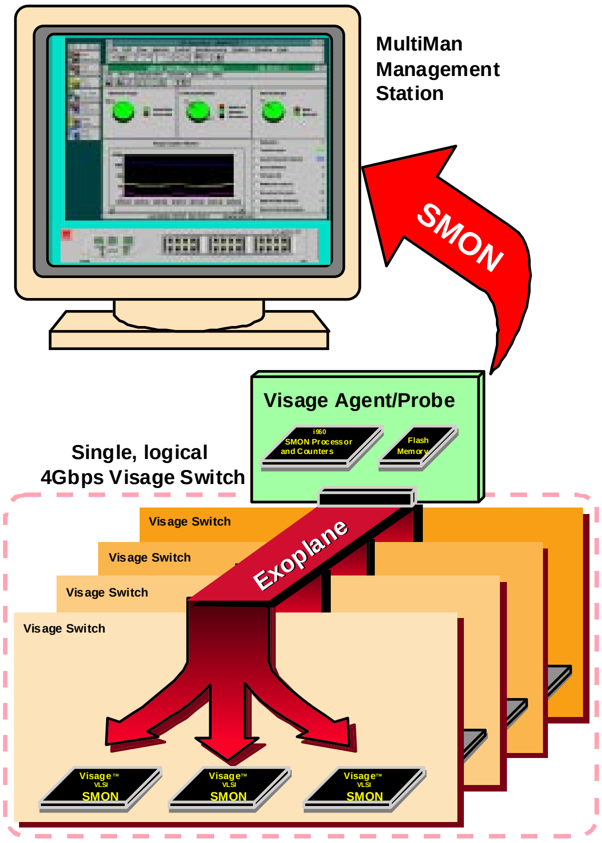

Figure 16 The Visage Distributed VLSI Based Management Architecture

The actual application is based on the RMON standard, with many significant enhancements that allow for the monitoring of switched, rather than shared networks. It possesses a graphic user interface that makes interpretation of collected comprehensive information useful. (Figure 17)

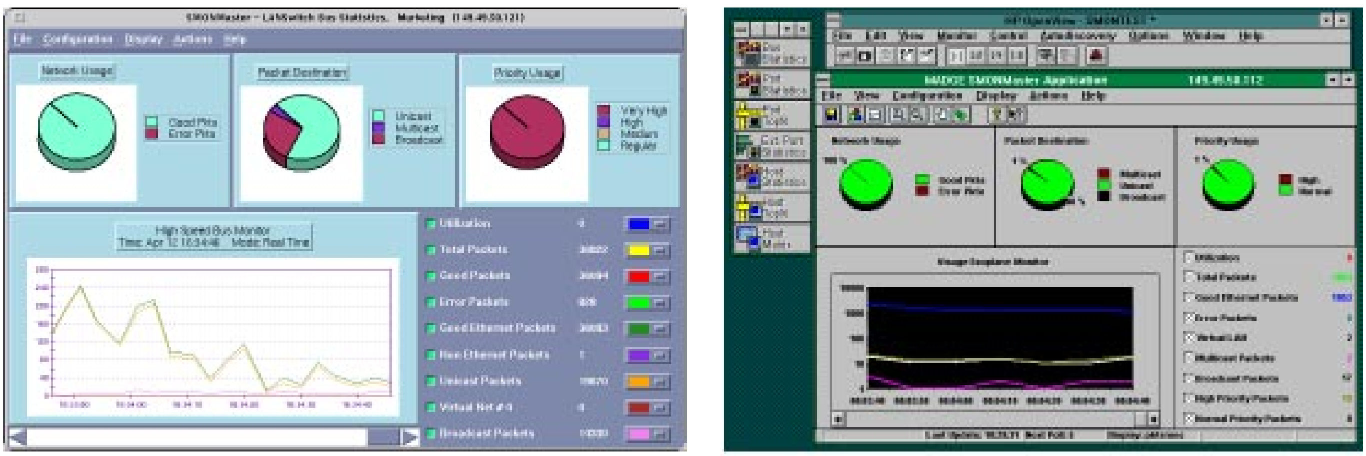

Figure 17 Typical SMONMaster Screens

Conclusion

When comparing switch architectures, a rule of thumb is that all switches are definitely NOT created equal. Some switches are nothing more than fast bridges. Others more closely resemble poorly designed routers. Many have extremely limited functionality beyond plug-and-play operation, lack the ability to grow with the network, lack any fault-tolerance characteristics, or are completely unmanageable, making them inappropriate for use in large enterprise networks. Some switches break down under intense, but common, data traffic patterns (e.g. broadcast traffic), or offer inconsistent performance for the isochronous data types (digitized, network-based video and audio) that are now being implemented on many networks.

The LANswitch was designed with the idea that these worst-case applications would be implemented on large, mission-critical enterprise networks. That is why the LANswitch was designed to inherently possess all of the characteristics required to support these next-generation networks. That is why the LANswitch continues to best all of the competition in world-wide evaluations. And, that is why the LANswitch backbone switch and the Visage workgroup switch will provide the necessary performance required by these next-generation networks well into the next millennium.

Madge Networks

Americas
2310 North First Street
San Jose,
CA 95131-1011
United States
Tel +1 408 955 0700
Fax +1 408 955 0970
http://www.madge.com

Asia, Australia & New Zealand
12/F Li Po Chun Chambers
189 Des Voeux Road,
Central,
Hong Kong
Tel +852 2593 9888
Fax +852 2519 8022
http://www.madge.com

Europe, Middle East & Africa
Knaves Beech Business Park
Loudwater, High Wycombe
Bucks HP10 9Qz
England
Tel +44 1628 858000
Fax +44 1628 858011
http://www.madge.com

Japan
Mita NN Building
1-23, Shiba 4-chome
Minato-ku, Tokyo 108
Japan
Tel +81 3 5232 3281
Fax +81 3 5232 3208
http://www.madge.com

142751 10/96