14. VLANs

14.1 Overview

Virtual LANs have been made possible as the switching infrastructure has replaced the traditional shared media LANs. An individual switch port can be assigned to a logical LAN, the next switch port can be easily assigned to a completely different logical LAN. This is made possible by 'tagging' the frames entering the port so that these frames are identified as belonging to a particular logical LAN whilst they travel along the switch fabric of the box. Once these frames are sent out of their logical LAN ports, the tag is removed from the frame. In the past, proprietary frame tagging has been implemented by Cisco (Inter Switch Link or ISL) and Bay Networks (Lattis Span), but the standard is now defined by **802.1q** and may be the one that you go for in order to allow interoperability between different manufacturers boxes.

More recently frame tagging has been applied to trunk ports as links to other switches and routers within the wider network, carry multiple VLANs. This 'tagging' is very important since it enables the VLANs to spread 'Enterprise-wide' as the backbones take the bulk of network traffic and VLAN information. This is why VLANs need to operate across high-bandwidth trunked FDDI (802.10), Fast Ethernet (ISL and 802.1q) and ATM links (ATM LANE).

One key difference between ISL and 802.1q is that ISL allows multiple instances of Spanning Tree (i.e. one per VLAN) to exist on a trunk link whereas 802.1q only allows one instance of Spanning Tree on a trunked link. Which trunking method you use influences the network design somewhat particularly if you are wishing to make the most of the network connections and not have part of the network not being used because Spanning Tree has blocked some ports. ISL allows you to loadshare VLANs across parallel links so you don't have to have network ports not being used. This does not mean that you are completely tied in to using Cisco equipment only since you can 'map' ISL VLANs to 802.1q VLANs before entering a load-sharing section of the network and also other manufacturers such as Lucent support ISL anyway.

The use of switches has meant that microsegmentation has increased the number of Collision Domains thereby minimising the number of collisions that occur across this 'flat' network. This in turn frees up more bandwidth for data but at the cost of increasing the amount of broadcast traffic. VLANs are the next step where multicast and broadcast traffic is restricted to each of the VLANs and reduces the possibility of broadcast storms.

On a large network one Spanning Tree would take a very long time to converge. At the best of times even a RIPv1 network converges more quickly than a Spanning Tree network. For this reason, due to the greater likelihood of changes occurring the bigger the network is, it is a good thing to have one instance of Spanning Tree for each VLAN which decreases the calculation time, improves scalability and accommodates changes more efficiently.

VLANs allow you to connect any user to any logical LAN. The benefit here is that the user could be anywhere in the building or even another building. A particular department does not have to have all its employees physically situated in the same place. In addition, security is easily maintained since the only way to communicate between the virtual LANs is by routing between them, either by way of a router (slow) or via a layer 3 switch. VLANs also simplify moves, adds and changes plus give opportunities to load share traffic. Ideally, you should look to design your logical networks such that at least 80% of LAN traffic is maintained with the LAN, and a maximum of 20% of traffic routed between LANs. Note here that Layer 3 switching makes this less of an imperative. Ultimately, layer 3 switches will completely replace routers as far as intraLAN routing is concerned, the routers will remain as edge devices, doing the job that they are best at.

In addition, multiple logical networks can be multiplexed through one physical connection, provided that the connection is between two boxes that use the same **Multi-Link Trunking (MLT) standard**.

Originally, VLANs were simply based on port ID i.e. different ports being assigned to different VLANs. This is fine if the different groups are local, but not flexible enough to accommodate campus-wide VLANs. 'Port-centric' VLANs require no lookup table and are easy on the processor especially if an ASIC is taking care of the switching, plus there is a high level of security as packets

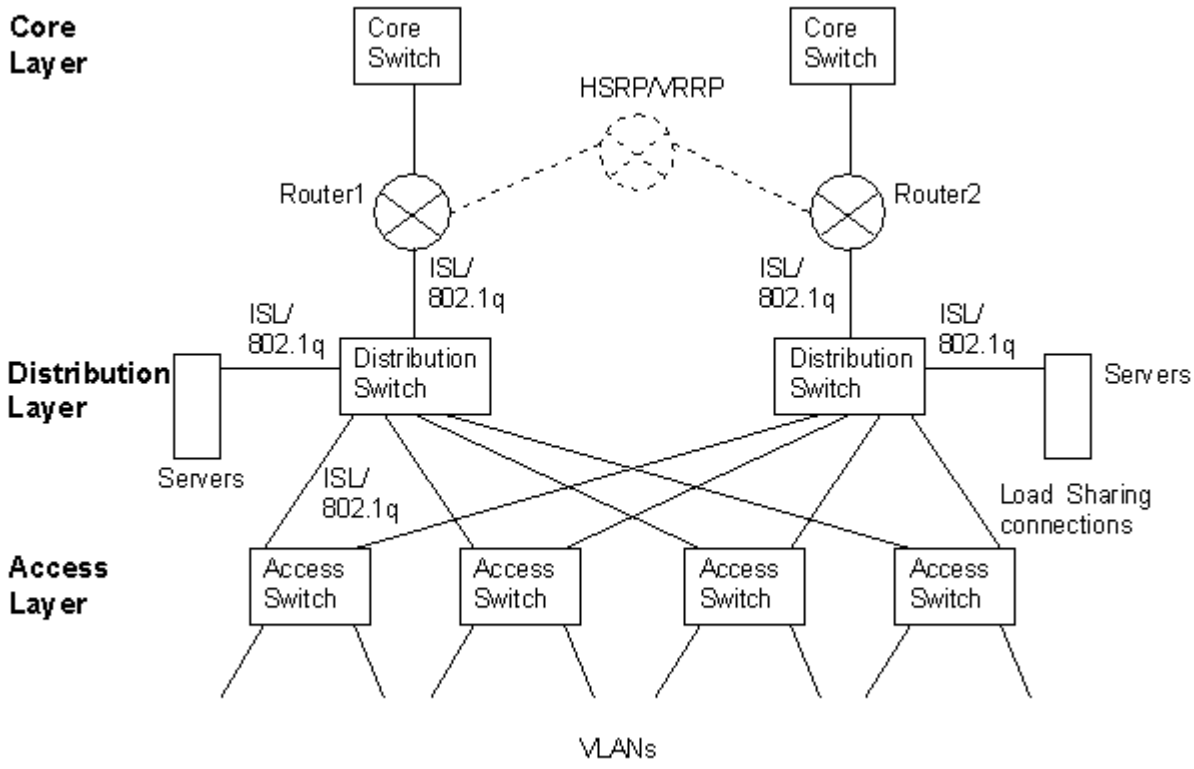
are very unlikely to 'leak' into other VLANs. These type of VLANs are often referred to as **Static VLANs** as they are manually configured.

Membership of a VLAN can be determined by the MAC address. The user can move departments and the MAC address is remembered by the switch and the user remains part of the same logical LAN. VLANs can also be based on protocol or even application. These methods enable Moves and Changes to be implemented with little effort. These types of VLANs are called **Dynamic VLANs** since the ports can automatically determine their VLAN assignments on a per packet basis. The problem with Dynamic VLANs is that a lookup table has to be produced with all the known MAC addresses mapped to the relevant VLANs. Not only is this table likely to be very large, but it is also going to change frequently as new devices are added to the network and old ones are removed. If an organisation has a lot of movement within the building environment, then the Dynamic VLAN may be the best design.

Frame filtering is often used by switches to aid in minimising LAN traffic. Tables are kept and frames are compared to the table entries to varying levels of frame depth. The deeper into the frame the switch has to go, the greater the switch latency. Also, the larger the table; the more the latency and these tables need to be synchronised with other switches.

14.2 Switched Network Design

A good general network design is shown below:



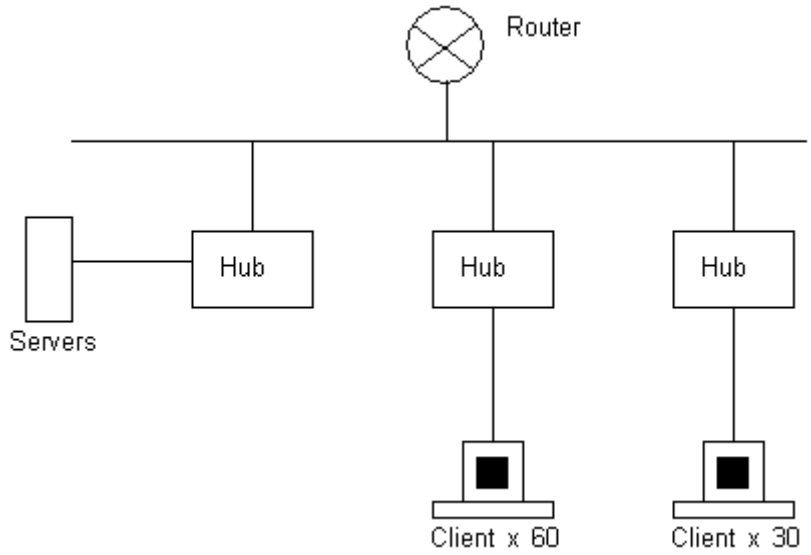
All networks have the Core, Distribution and Access layer functions within them even if two or more of the functions are dealt with by one box. Generally, it is a good idea to have the servers connected to the Distribution switch and let the distribution switch deal with the VLANs. The routers could either be separate 'routers on a stick' or built in to the switch and these should deal with access policies and filtering. The Core switches need to be left purely to switch packets and not worry about policies, routing or server traffic.

One thing to be aware of when upgrading networks is the capability of the servers. It is all very well upgrading the client links and the backbone links, but if the servers can so easily become the

bottleneck in a network. You need to make sure that the server, be it based on NT or Unix, is capable of fast network technology such as Duplex operation and Gigabit Ethernet. Not only do you need to make sure that the server can cope with fibre-based cards for 1000BaseSX or 1000BaseLX, you also need to be sure that the servers internal components such as the hard disk, processor and the bus can cope with a fast network. A duplex gigabit card can completely swamp the whole PCI bus in a standard PC which can cause problems since this PCI bus is meant to be share by the other cards in the machine.

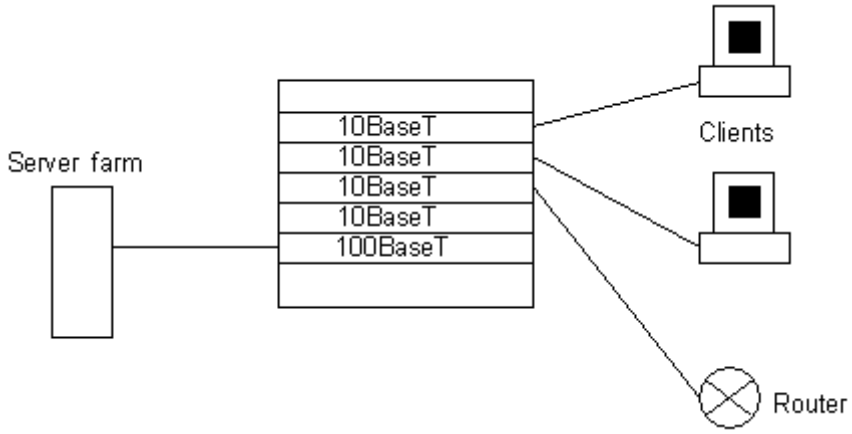
Examples

The following diagram illustrates a typical flat network.



The trouble with this network is that it is just one collision domain and one broadcast domain so it is prone to high collision rates and alot of the bandwidth on the network is going to be given over to broadcasts. The problem with broadcast traffic is that each station on the network be it a server or a client, will have to process the broadcast packets. This processing has to be carried out by the CPU of the computer and can have quite a large effect on the processing capability of the computer.

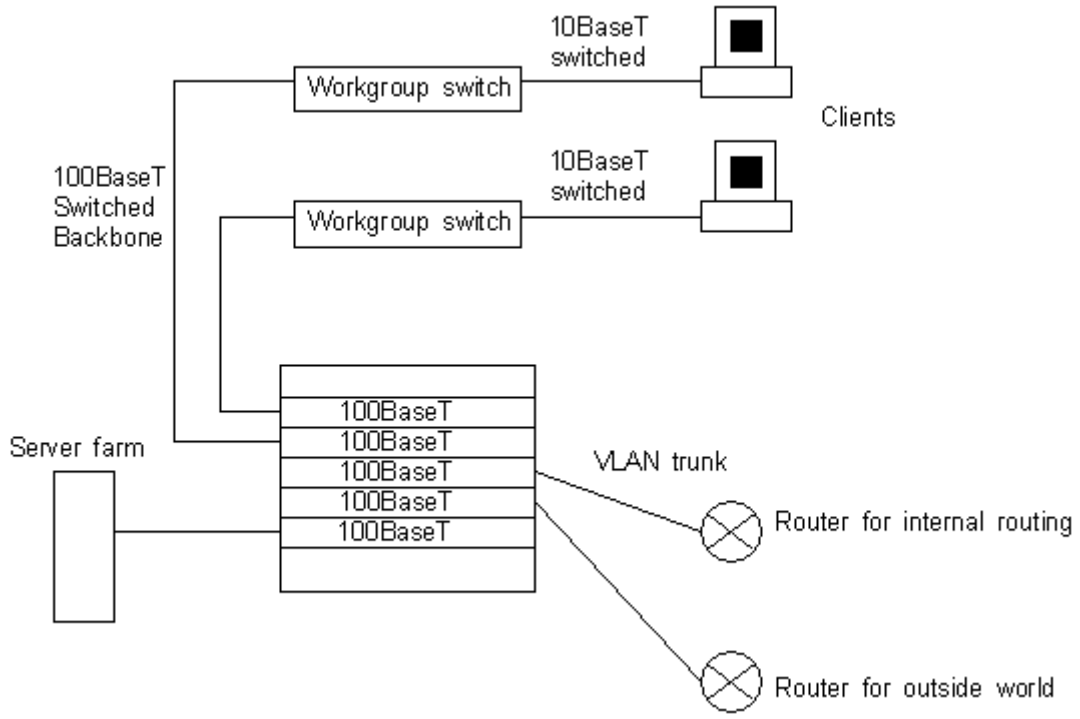
A chassis-based switch can replace the hubs and provide a far more efficient network:



Using the switch, not only is more bandwidth available to each client because each client has it's own collision domain, but also one can configure VLANs to separate certain groups within the organisation and thereby reduce broadcast traffic freeing up even more bandwidth. The router can either remain a separate device or become integrated into the electronics of the switch. You will

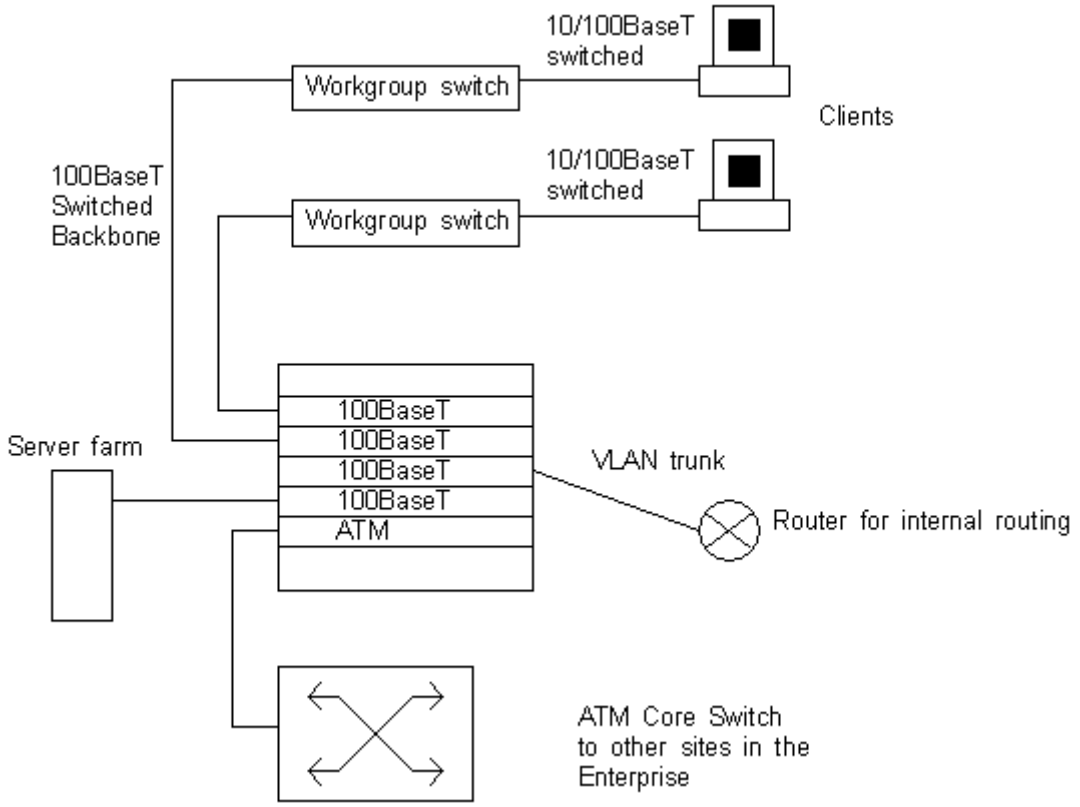
notice that in this small network, the core, distribution and access layers can be satisfied by one box even though the functions are still distinct.

The following design extends the previous one for a larger network:



Here, the clients are fed from 'access' or 'workgroup' switches perhaps located in different closets elsewhere in the building. In this case, they are linked via 100BaseT switched links which could be copper (100BaseT) or fibre (100BaseFX). Alternatively, they could be ATM or even FDDI links depending on the capabilitites of the main distribution and access switches. This is because ATM LAN Emulation or FDDI 802.10 can be used to maintain the VLANs across the trunk. The trunk to the local router could be ISL (which would allow load sharing of VLANs) or 802.1q, this local router could be contained within the switch, making the switch a layer 3 switch.

In the diagram below, the network has been extended further to an Enterprise level:



The ATM switch takes on the role of the core switch and links to other sites which also have core switches. The backbone within in site could be Gigabit Ethernet, again depending on the capabilities of the access and distribution switches. The server farm and the routing is carried out by the distribution switch, it is best to leave the core switches to high speed switching and, as described before, the distribution switches should take on the burden of access policies. If 100BaseT is to be fed to the desktop it becomes more important to upgrade the bandwidth between the access and distribution switches to either Gigabit Ethernet or a multiple link such as 4 x 100BaseT links forming a 400BaseT channel (when operating in duplex mode this can effectively allow a maximum throughput of 800Mb/s).

In a large campus environment, multiple switch/routers at the distribution level can provide the opportunity to form resilient links to the core via the use of HSRP or VRRP. These protocols can provide alternative paths for specified VLANs.

The Virtual Router redundancy protocol gives the possibility of load sharing traffic across the routers by setting up separate groups with the aim of roughly half the local traffic going through one router and half through another by way of two or more virtual default gateways.

14.3 Trunking Overview

The trunking protocol used can be important in deciding the structure of the network. Although 802.1q is the industry standard it has a limitation in that it only allows one instance of Spanning Tree in a trunk link even if there are a number of VLANs in the trunk. Cisco's ISL, however, allows an instance of Spanning Tree per VLAN and the advantage of this is that trunk ports can remain available for some VLANs if blocked for other VLANs and hardware need not remain dormant as it would if 802.1q was implemented.

Trunking protocols such as Cisco's ISL could be used in the link to the servers. The advantage of this is that Intel manufacture ISL capable NICs and these NICs allow the server to have a number of different VLANs, and therefore completely different networks, to be served by the server as if the server was made up of completely different devices. The NIC cards themselves take care of the processing of frames so the CPU of the server does not have to take the load! Another advantage of using ISL-aware NICs is that the traffic does not have to be dealt with by a router and this therefore aids to keeping the network fast.

14.4 Cisco's Inter-Switch Link (ISL)

Cisco use a proprietary tagging method called **Inter-Switch Link (ISL)** which takes a different approach to tagging the Ethernet frame. Instead of increasing the frame size by inserting fields, ISL encapsulates the Ethernet frame.

Cisco's **Inter-Switch Link (ISL)** allows Per VLAN Spanning Tree (PVST) so multiple VLANs can exist across a trunk link. Multiple Spanning Trees allow load sharing to occur at layer 2 by assigning different port priorities per VLAN. 802.1q only allows Mono Spanning Tree (MST) i.e. one instance of Spanning Tree trunk.

ISL only runs on point-to-point links on Fast Ethernet (copper or fibre) and Token Ring (ISL+). Although ISL will operate over 10Mbps links it is not recommended! ISL runs between switches, from switches to routers and from switches to Intel and Xpoint Technologies NICs which understand ISL, thereby allowing servers to distinguish between VLANs.

With ISL the data frame is not touched but is encapsulated according to the following process:

- The frame enters the switch and is stored in the port's buffer.
- The SAINT/SAGE encapsulates the ISL on a trunk port.
- The encapsulation has 30 bytes of information, 26 bytes for the header (VLAN ID and port number) and 4 bytes for the FCS.
- The frame is switched to the destination port(s).
- The SAINT/SAGE dencapsulates the frame before it is sent out of a normal port, or leaves it alone if the port is a trunk port.

The following diagram details the ISL frame tagging format:

Bits	40	4	4	48	16	24	24	15	1	16	16	Variable	32
	ISL Multicast Address	Type field	User field	Source port address	Length	AAAA03	OUI	VLAN ID	B	Index	R	Original Frame	FCS

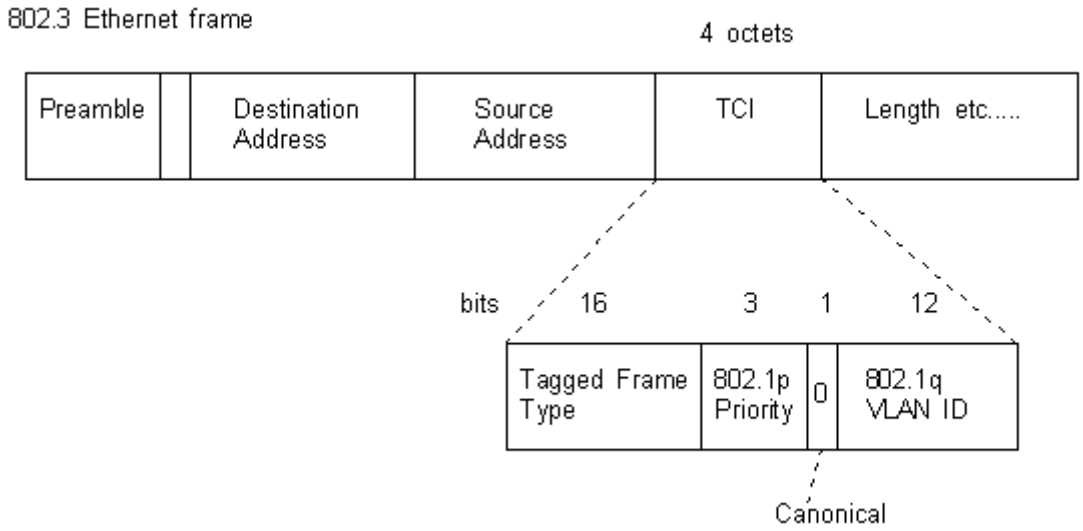
- **ISL Multicast Address** - this reserved address is **0x01000C0000** (40 bits).
- **Type Field** - this identifies the type of frame that is encapsulated, **0000** is Ethernet, **0001** is Token Ring, **0010** is FDDI, **0011** is ATM.
- **User Field** - **0000** means Normal Priority, **0001** means Priority 1, **0010** means Priority 2 and **0011** means High Priority. Similar to the Class of Service field in 802.1p.
- **Source address** - this is the MAC address of the frames source.
- **Length** - this is the length of the frame excluding the fields up to AAAA03 and also excluding the FCS.
- **AAAA03** - indicates that the ISL frames use SNAP LLC.
- **OUI** - this is the Organisational Unique Identifier of the source of the frame i.e. the first three bytes of the Source Address.
- **VLAN ID** - Cisco use the lowest 10 bits of the 15 to give a possible 1024 different VLANs although only 250 VLANs can be active at any one time. The Catalyst 3000 supports 64 VLANs and the 7000 series routers support 255 VLANs.
- **B** - when set to '1' this bit indicates whether the frame is a BPDU, CDP or VTP frame and the frame is sent straight to the NMP for processing.
- **Index Field** - indicates the port number of the source port.
- **R** - this is set to **0x0000** for Ethernet frames but for Token Ring or FDDI frames the AC or FC fields are placed here e.g. for FDDI an FC of 0x12 would mean 0x0012 is placed in the R field.
- **Original frame** - can be up to 24575 bytes in length.
- **FCS** - This is the extra FCS added by ISL.

14.4 Class of Service and VLANs (802.1p & 802.1q)

Quality of Service (QoS) is becoming more important as data networks begin to carry more time sensitive traffic such as real time voice and video. At layer 2 this is referred to as **Class of Service (CoS)**.

The 802.1 group have been working on an extension to the MAC layer that takes into account CoS. 802.1p is a standard for traffic prioritisation where network frames are tagged with one of eight priority levels, where **7** is high and **0** is low. Switches and routers that are 802.1p compliant can give traffic that is time-sensitive such as voice traffic, preferential treatment if the priority tag has been set to a higher value than other traffic.

In order to accommodate tagging an Ethernet frame a new field has been introduced called the **Tag Control Info (TCI)** field between the Source MAC address and the Length field of the Ethernet frame. This is illustrated below:



- **Tagged Frame Type** - this indicates the type of tag, for Ethernet frames this is currently always **0x8100**.
- **802.1p Priority** - this ranges from binary 000 (0) for low priority to binary 111 (7) for high priority. This maps to the **Quality of Service (QoS)** values used in the TOS field IP precedence values.
    - **000** (0) - Routine
    - **001** (1) - Priority
    - **010** (2) - Immediate
    - **011** (3) - Flash
    - **100** (4) - Flash Override
    - **101** (5) - Critical
    - **110** (6) - Internetwork Control
    - **111** (7) - Network Control
- **Canonical** - this is always **0**.
- **802.1q VLAN ID** - this identifies the VLAN number when trunking VLANs.

Although the frame illustrated is an 802.3 frame, 802.1p/802.1q can also be applied to the Ethernet frame where the TCI is inserted just before the Type field and just after the Source MAC Address.

You will note the similarity between the 802.1p priority field and the Precedence field in the Diff Serv Code Point of the IP datagram. This makes mapping between IP layer 3 and MAC layer priorities much easier.

You will note that the Ethernet frame becomes 'oversized' i.e. grows from the standard maximum size of 1518 bytes to 1522 bytes. This sometimes called a **Baby Giant**. Consequently these frames may be dropped by some network equipment, although most vendors now support 802.1p and 802.1q.

When applying Layer 2 Priority Queueing within a trunk, commonly two priority levels (low and high) are implemented, although as we have seen there is scope though to increase to eight. This

is because each priority has to have its own queue. This is implemented in hardware and is therefore expensive so most manufacturers currently build in two queues per port, a low priority queue for priority levels 0 to 3 and a high priority queue for priority levels 4 to 7. Prioritisation is determined on the outbound packets from a switch, therefore they are already ordered on the inbound ports of the next switch so that prioritisation need not be implemented on the inbound ports, unless the ports are using buffering. Low priority frames or frames without an 802.1p tag are treated with 'best effort' delivery. As time goes on more manufacturers will include separate queues for each priority level to give more granularity and as the applications begin to demand it.

On a 802.1q trunk, one VLAN is NOT tagged. This VLAN is called the **Native VLAN**, and must be configured the same on each side of the trunk. This way, we can deduce to which VLAN a frame belongs when we receive a frame with no Tag, otherwise the frame will remain in whatever tagged VLAN it arrived on even if it is the wrong one. When a switch trunks a frame, it inserts the Tag and then recomputes the FCS.

The 802.1q standard implements Spanning Tree on the Native VLAN, and this applies to all the trunked VLANs, this is called **Mono Spanning Tree (MST)**. Cisco have adapted 802.1q and use a tunnelling mechanism to provide what is called **Per VLAN Spanning Tree Plus (PVST+)** with VLAN numbers up to 1005. This gives the same benefits that ISL gives.

The native VLAN configured on each end of a 802.1q trunk must be the same. A switch receiving a non-tagged frame will assign it to the native VLAN of the trunk. If one end is configured for Native VLAN 1 and the other to Native VLAN 2, a frame sent in VLAN 1 on one side will be received on VLAN 2 on the other. You are then merging VLAN 1 and 2.