Layer 2 switching, VLAN tagging, and CompactPCI

By Curtis A. Schwaderer

CompactPCI Systems

Layer 2 switching and VLAN tagging are being increasingly useful for a variety of communications applications. These uses go beyond standard Internet definitions and intended uses for the technology, wireless infrastructure, data centers, and voice over IP networks are all starting to utilize the capabilities specified by the Internet Engineering Task Force (IETF) request for comments (RFC) specifications that govern layer 2 switching and VLAN processing. CompactPCI manufacturers are beginning to be required to understand and apply this technology if they expect to sell their products into convergence networks in the future.

In this month's column, we'll look at the technology and specifications behind layer 2 switching and VLAN processing. Then we'll run through a typical data flow and application of the technology and look at some products that implement layer 2 switching and VLAN processing that can be used in various communications system applications.

Technology overview

What is *layer 2 switching* and *VLAN tagging*? These terms refer to making packet forwarding decisions using the data link layer (i.e. Ethernet, token ring, or FDDI) header of a packet. The term *routing* refers to processing a packet using the the network layer (IP, IPX, Appletalk) header of the packet. The most common application of layer 2 switching involves the Ethernet header at layer 2.

The main benefit of layer 2 switching is to make efficient use of network bandwidth. Historically, as networks grew, the first switches were nothing more than *port fan-out connections*, that is, all traffic arriving on all ports would be sent out all ports without any processing. As networks become more complex, the situation arises where two or more ports on the same switch may be directly or indirectly connected on the same local network (LAN). Figure 1 shows an example of this scenario where there are two potential looping problems within the network. Switch 1 can access network C directly through port 1 or indirectly through switch 2. Two ports of switch 2 are also directly connected to network C. In cases where a packet comes in on one port and leaves on multiple ports that may be connected to the same LAN, duplicate packets are created on the wire as the switch fans the incoming packets out all the ports. Other switches in the network end up sending the same packet back to the originating switch through the other port(s). The same packet ends up looping throughout these redundant connections indefinitely. Many of the first network attacks made use of this property to overload the network by causing congestion by creating traffic loops.

Layer 2 switching eliminates looping traffic by defining something called a *spanning tree* and spanning tree protocol that configures the spanning tree. This makes equipment more intelligent about what their ports are connected to. The spanning tree protocol identifies ports that are connected to the same LAN and configures the switch to send any given packet out only one port that is connected to a given LAN. This eliminates the problems associated with loops. Spanning tree is also smart enough to keep track of these redundant connections and if the *primary* goes down for some reason, the spanning tree will begin to send packets on the "secondary" port on the same LAN within the switch.

VLAN processing is an extension to the concept of layer 2 switching. VLAN stands for *Virtual LAN* and adds a four-byte tag field between the link layer header (i.e. Ethernet) and the network layer header (i.e. IP). This tag contains, among other things, a VLAN Identifier (VID) and associated user priority field.

VLAN aware switches are smart enough to look for this tag and make a switching decision based on tag information that determines which port(s) to send the incoming packet out on. The VLAN standard defines three kinds of traffic: untagged, priority tagged, and VLAN tagged. Untagged packets are packets without any VLAN tag. Priority tagged packets are packets with a VLAN tag, but a VID of zero (the NULL VID) and a valid priority field within the VLAN tag. VLAN tagged packets contain a VLAN tag with a valid VID field (non-zero). VLAN-aware switches must be able to classify and forward packets of all three types in order to work with legacy equipment as well as other VLAN-aware equipment.

Incoming packets may be untagged, priority tagged, or VLAN tagged. Depending on how the switch is configured, an untagged packet may leave priority-or-VLAN-tagged. Incoming VLAN tagged packets may leave untagged or even tagged with a different VLAN ID (VID).

VLAN processing doesn't stop there. There are also multiple ways of configuring what constitutes a VLAN. The VLAN specifications define a port-based approach where each port is a member of a particular VLAN. All traffic coming or going on this port would be a member of the configured

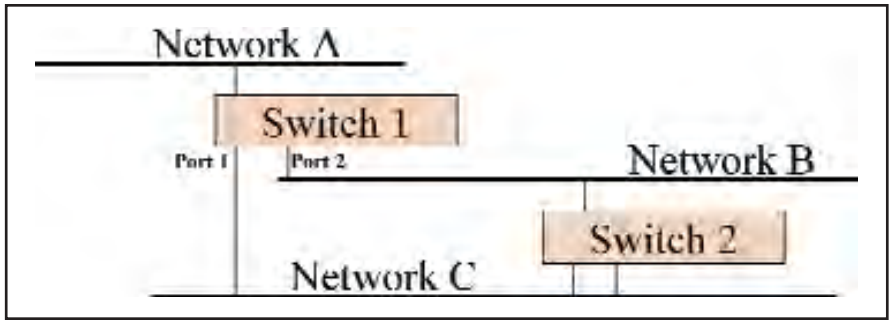

Figure 1. Example network configuration with loops

VLAN. There is also a MAC-based approach to VLAN processing where membership on a VLAN is defined by the source MAC address of a remote host. This is a common approach for remote mobile user access. A VLAN can be defined for a particular mobile host and wherever the traffic from this source comes from, it's identified as belonging to a particular VLAN. Protocol based VLANs key off of the protocol field of the Ethernet header. Application based VLANs key off of the TCP/IP header information to map to a particular VLAN. Since only port-based VLAN processing is defined by the specifications, the other methods are not as interoperable on a global scale.

Standards

There are three key standards for layer 2 switching and VLAN processing. These standards are governed by the Institute for Electrical and Electronics Engineers (IEEE). These standards are IEEE-802.1D, IEEE-802.1P, and IEEE-802.1Q:

- The IEEE-802.1D standard defines the notion of a bridge and how the spanning tree protocol works to define bridges and bridge membership.
- The IEEE-802.1P adds the notion of queue prioritization that defines up to eight prioritized traffic classes. This enables traffic to be tagged for faster or prioritized delivery.
- Finally, IEEE-802.1Q defines the VLAN tag and associated processing. This specification also functionally specifies VLAN Registration among VLAN-aware switches (Generic VLAN Registration Protocol, or GVRP). IEEE-802.1Q also talks about multicast registration (GMRP). Since one of the first target uses of a VLAN was to support videoconferencing and voice over IP networks, GMRP provides an automated way to setup and tear down these kinds of conference groups.

Processing

Now that you've been exposed to some of the features and functionality, let's take a closer look at how this technology supports an application like voice over IP conferencing. Let's say we're setting up a conference group with three remote members with associated MAC addresses A, B, and C. A quick review through the previous technology discussion tells us these members wouldn't be configured using a port based VLAN because every packet coming into that port would be included in the VLAN (not the desired result). Since the members would be associated with MAC addresses and the ports they may come in on may be indeterminate, MAC based VLAN processing is in order. Further, what if MAC member A is running multiple network applications using the same network interface? MAC based VLAN processing would include all traffic coming from MAC address A into the VLAN. For more complex situations like this, application based VLAN processing could be used (i.e. if the videoconference group uses port 100, application-based VLAN tagging would apply the videoconference VLAN based on the TCP port in the TCP header).

A video-conferencing software application may configure the MAC VLAN tables to enter MAC addresses B and C with VID 10 (for instance). To accommodate the MAC address A member, a protocol based rule would be provisioned by the software to say any packets with MAC A and accessing port 100 (the conferencing port) is also part of VID 10. This way packets from MAC A not destined for TCP port 100 would pass through untagged (or perhaps out other ports on other VLANs if configured).

Once the tables are configured, it's up to the microcode. On header reception, the microcode applies application-based VLAN rules to determine if it's a match. If a packet comes in on TCP port 100 with source MAC A, an entry in the VLAN table would be found for that packet and the VLAN tagging result in the table for that entry would be applied. The packet would then be forwarded out the member ports for VLAN 10.

Products

Doing a quick search, I was able to find three software products that could be used to implement layer 2 switching and VLAN processing:

- LVL7 layer 2 product
- RadiSys Microcode Solutions Library (Ethernet) product
- Wind River's Tornado for Managed Switches (TMS) product

Tornado for Managed Switches (TMS) from Wind River is a combination of the VxWorks kernel, Tornado development environment, and layer 2 and 3 IP switching and routing C code that Wind River acquired as a result of the purchase of a company called Routerware. This software product provides a layer 2 and 3 IP switching/routing solution complying to the pertinent IEEE specifications and IETF RFCs. The product is intended to run on a processor like a PowerPC, Pentium, MIPS, or ARM. Wind River did announce support for the IBM, Motorola, and Intel network processors. However, the software is all C code and the bridging and routing functionality runs on the core CPU of these network processors, not on the microengines. However, integrating this product with RadiSys Microcode Solutions Library (described below), the communications product manufacturer could create an optimized network processor based solution and achieve much higher performance.

Layer 2 bridging and VLAN processing is an ideal application for network processor technology. As described above, the rules based and processing are ideal for implementing in microcode. The LVL7 and RadiSys products are specifically engineered for network processor-based operation.

LVL7 also does software products and services for layer 2 bridging and layer 3 routing. LVL7 has done some coding with network processor architectures such as Vitesse, but it's unclear how much of this code is C and how much runs in the microengine environment. The layer 2 bridging support implements the three IEEE specifications described previously. The implementation does not implement MAC-based or application-based VLAN processing algorithms.

The RadiSys Microcode Solutions Library (Ethernet) product also contains layer 2 bridging and layer 3 routing implemented in microcode for the Intel IXP family of network processors. The implementation conforms to the three IEEE specifications described above, but is only suitable for network processor based communications devices since it's a microcode implementation. RadiSys does supply C code for the building of the bridging, routing, and VLAN tables. There is also C code that performs spanning tree protocols and other configuration and control aspects of layer 2 bridging and layer 3 routing. The microcode uses macros to abstract the specifics of table access for bridge, VLAN, or routing table processing. This makes the microcode implementation more easily integrated with products like Tornado for Managed Switches or the Linux layer 2/3 open source implementation. One unique aspect to the RadiSys implementation is that the 802.1Q implementation includes port-

based, MAC-based, and application-based VLAN processing. While MAC and application based VLAN processing is not explicitly specified in IEEE-802.1Q, the product provides defined and documented APIs to support the VLAN processing scenarios at all these levels.

Summary

The virtual LAN concept and specifications enable the ability to create logical LANs that overlay physical LANs to create a domain of access for many classes of users. This concept is being applied to many kinds of applications in wireless and wire-line equipment. As IP continues to find its way into multiple communications topologies, CompactPCI manufacturers must understand this technology and how to incorporate it into their products in order to be successful in the future.