POWER5 Virtualization: How to work with VLANs using the IBM Virtual I/O Server

Nigel Griffiths

November 18, 2008
(First published September 29, 2005)

Learn how to improve security between logical networks. In this article, Nigel Griffiths describes the benefits of working with VLANs using the IBM Virtual I/O Server (VIO Server).

Introduction

This article assumes the following:

- System administration experience in setting up Ethernet networks, but no experience with virtual Local Area Networks (VLANs)
- Experience in pSeries® hardware and, in particular, POWER5™-based machines
- Experience with the Hardware Management Console (HMC) and creating logical partitions (LPARs)
- Experience with the IBM Virtual I/O Server (VIO Server) in simple configuration

Note: The term VIO Server is used in this article to refer to the IBM Virtual I/O Server LPAR running on a pSeries POWER5-based machine. The term VIO client is used to refer to any LPAR making use of the virtual SCSI (small computer system interface) and/or virtual Ethernet services of the VIO Server.

Basic networks and basic VLAN

You can run two different IP address ranges on the same Ethernet cable segment. For example, all cables being plugged into the same hub. This would give you two logical networks on one physical network. While this does work well, there is a large problem — one machine, with the right network setup and software, can capture all the packets for both IP address ranges. So, there is no security between the logical networks. Figure 1 below illustrates this point:
To stop this security hole, Cisco first invented ISL (Inter Switch Link), which has since moved on and become the IEEE-802.1Q standard that covers VLANs. This standard, in its simplest form, requires an Ethernet switch with IEEE-802.1Q VLAN to support ports on the switch that have extra attributes (VLAN tagging), so machines (or their PCI adapters) can use VLAN tagging. So in Figure 1 above, you would want to replace the Ethernet hub with an Ethernet switch -- this means it is intelligent in sending packets. The ports within the switch are capable of belonging to one or more VLANs with a default VLAN for the port -- this is called Port VLAN ID (PVID). Packets that arrive at the switch with a VLAN tag are permitted to pass, depending on whether this matches one of the VLANs that the port is a member of. If a packet arrives without a VLAN tag, the switch then tags it with the PVID of the switch port. The routing of packets within the network will now rely not only on the subnet, but also on the VLAN tags. The VLAN tag separates the ports into what are effectively separate physical networks -- for example, using two VLANs with VLAN IDs of 31 and 45. If you assign the same VLAN ID (31) to ports where machines A, B, and C are plugged in and VLAN ID 45 to the ports where machines 1, 2, and 3 are plugged in, the switch will not transfer packets between the two logical networks -- that is two sets of machines. Most devices are completely unaware of this VLAN tagging at the switch and do not accept packets with VLAN tags, although most devices can be configured to accept and understand the tags. With this in mind, most switches, if they receive a packet with a VLAN tag matching the PVID, strip the tag from the packet before sending it out of the switch to the attached device. If the VLAN tag matches one of the possible additional VLANs that the port belongs to (for example, not the default PVID), the tag will be left intact. This allows devices that are both VLAN aware and VLAN unaware to exist side by side. See Figure 2 below.

See Figure 3 below for a simpler view. By using VLAN IDs and port tagging, the Ethernet switch logically separates the two VLANs so that packets are no longer communicated between the two sets of machines.
It is also worth pointing out that using the switch to add VLAN ID tagging means the network administrator has complete control of where packets go and who can see them. This is how most network administrators like to operate, as it is invisible to the machines and their system administrators. Your network administrators are probably using VLANs without you knowing, as there are many benefits when running a large network. I did not know this until I asked, but every floor in the IBM building in which I work is a separate VLAN. Listed below are some other reasons for using VLANs:

- Limits packets to a subset of the company network backbone
- Limits pointless network flow of packets between non-server machines
- Increases performance
- Provides security -- using VLANs to limit hostile connections (for example, direct Internet connection)
- Allows you to monitor network traffic for planning

If your operating system allows it, you can also add these VLAN tags on the machines themselves. In addition, the Ethernet switch can be told not to automatically add VLAN tags on the port. This decentralizes the point of control away from the switch to many machines, making it much harder to control -- keep in mind that a rogue device could read all Ethernet packets on the network with a simple administration mistake. In general, it is better to control the VLAN tagging at the switches for simplicity.

Note: I have only scratched the surface of what VLANs can do. For example, Switch ports can have multiple VLANs, send every packet to a particular port regardless of the VLAN (this is often used for monitoring and security), and much more complex configurations. Let's now move on to VLANs in the Hypervisor.

The virtual Ethernet in the Hypervisor

In the IBM eServer™ p5 pSeries machines have a Hypervisor, which is used to control the LPARs in the machine. This Hypervisor is firmware that all operating systems (AIX® and Linux®) use to interact with the hardware. It also provides a virtual Ethernet service, so that without an Ethernet
adapter, the LPARs can communicate over the Hypervisor (through memory) using TCP/IP (Transmission Control Protocol/Internet Protocol).

The Hypervisor virtual Ethernet is implemented to behave just like a VLAN-aware Ethernet switch -- so every use of the Hypervisor virtual Ethernet is using VLAN technology -- even if you are not aware of it. See Figure 4 below:

**Figure 4. Two logical networks on two VLAN networks using Hypervisor virtual Ethernet**

For two LPARs to communicate using virtual Ethernet, you need to add a virtual Ethernet adapter to each LPAR and set the PVID to the same number -- the two LPARs can use standard TCP/IP tools to communicate like Telnet, FTP, SSH, and so forth. This is exactly the same model as in the Basic networks and basic VLAN section for say machines A and B to communicate, but Machine A and Machine 2 can not communicate.

The PVID is added to outgoing packets from an LPAR. This is called tagging or VLAN tagging. These tags are then used to decide where the packet goes and what is stripped off when the packet gets delivered to the other LPARs with the same PVID. LPARs are unaware of this tagging/untagging. VLAN tagging is used to separate LPARs using different PVIDs. It can also have multiple internal networks where packets are routed between a subset of the LPARs within the pSeries chassis to avoid going to the wrong LPAR.

**Note:** This VLAN security model is used so that you can be sure that potential hackers cannot jump from one internal network to another. An LPAR can only see packets on its VLAN, and not others.

There is another VLAN feature that you need to be aware of. For example, if a packet arrives at a destination virtual Ethernet adapter (the analogue of a Ethernet Switch port) with a matching VLAN tag (that is the packet VLAN ID equals the PVID), the tag is then stripped off (the packet is untagged). This is part of the IEEE 802.1Q standard and happens regardless of how the VLAN tag got there in the first place, either arriving on a tagged port (virtual Ethernet adapter) or added by the originating adapter in a machine or LPAR. Also note that if the packet VLAN ID does not equal the PVID but equals one of the additional VLAN IDs that the port permitted to pass, the VLAN tag is not removed and the packet is allowed to pass with the VLAN tag left intact.
A worked example

The previous sections have been about Ethernet packets within a machine and LPARs communicating with themselves across the internal Hypervisor network. There is, of course, a need to get packets to and from an external Local Area Network (LAN). In a VIO Server environment, it is the VIO Server that provides the link between internal and external LANs. This gets a little more complex, as you might have many VLANs within the machine that need to be connected to multiple VLANs outside the machine in a secure way. Therefore, the VIO Server needs to be connected to all of the VLANs but not let packets move between the VLANs, as this would break the security model.

Figure 5 below illustrates how to set up this type of configuration:

Figure 5. A worked example

There are multiple requirements for this to work:

1. LPARs must be able to communicate with other LPARs on the same VLAN. Requirement -- internal VLAN communications are simple to setup.
2. For many LPARs to talk to a single virtual Ethernet adapter in the VIO server, you have to use IEEE 802.1Q standard on the virtual Ethernet adapter of the VIO Server to allow more than single PVID tagged packets to be accepted at the virtual Ethernet adapter. Requirement -- one virtual Ethernet adapter/port for multiple VLANs on the VIO Server.
3. Also, you do not want the VLAN tags stripped from the arriving packets, as the VIO Server could not forward them on to the right external VLAN. Requirement -- no stripping of VLAN tags on the VIO Server.
4. In addition, you also have to enable the Shared Ethernet Adapter (SEA) to allow packets with many VLANs to pass through it. Requirement -- allow the SEA to transmit packets for many VLANs.

In this example, you are going to use VLAN IDs of 23, 150, and 160. There is nothing special about these numbers, so simply substitute your own. Note: VLAN IDs 8 and 89 are also used here as otherwise unused VLAN IDs, so be careful of these numbers. Let’s now build up this picture one step at a time in the following sections.
Requirement 1: Internal VLAN communications are simple to set up

Internal VLAN communications are simple set up on the HMC. The LPAR profile is assigned to a virtual Ethernet adapter -- it is assumed you are familiar with the HMC. If you have created an LPAR and an LPAR profile, you should see the virtual resources panels. Each LPAR below has been given a virtual Ethernet adapter with a **PVID**. The **Trunk** and **IEEE802.1Q** are also unselected. As discussed above, the **PVID** will be added to the packets from an LPAR, used to route packets to the right LPARs, and then stripped off before delivery. For example, see **Figure 6** below of four LPARs. LPAR3 and LPAR4 have the same **PVID** (for example, 160) so that they can communicate directly. LPAR1 and LPAR2 have other **PVIDs** so that they cannot talk to LPAR3 or LPAR4, or each other. However, LPAR1 and LPAR2 can communicate with other machines on VLANs external to the machine.

**Figure 6. Requirement 1: Internal VLAN communications are simple to set up**

![Figure 6](image)

**Figure 7** below shows the creation of the virtual Ethernet adapter for LPAR1. The virtual Ethernet adapter settings are **PVID = 23** and the **Trunk** adapter and **IEEE 802.1Q** are unselected. Note that the **Slot** number needs to be unique within an LPAR for each virtual device. Its value is otherwise unimportant for virtual Ethernet devices.

**Figure 7. Virtual Ethernet adapter for LPAR1**

![Figure 7](image)

For LPARs 2, 3, and 4, the same settings are used except for the **PVID** numbers.
Requirement 2: One virtual Ethernet adapter/port for many internal VLANs on the VIO Server

The VIO Server is a single relay point between multiple internal VLANs and the single external physical Ethernet adapter and LAN. To connect up the virtual and physical networks, you need to use the Shared Ethernet Adapter (SEA) device to link one or more virtual adapters to a physical adapter. This single virtual Ethernet adapter connecting to multiple internal VLANs is by far the simplest to set up and covers most cases. Remember that the virtual Ethernet runs over the internal memory of the system -- you cannot gain performance by having extra virtual adapters. In more complex cases, you might have other complete sets -- physical adapter, virtual adapter, and SEA. However, to setup this simple case, when the virtual Ethernet adapter in the VIO Server is created, you need to make sure to name all the internal VLANs that are connected, as shown in Figure 8 below.

Figure 8. One virtual Ethernet adapter/port for many internal VLANs on the VIO Server

Figure 9 below illustrates the purpose of having multiple fields (in the virtual Ethernet adapter creation dialogue) for Additional VLAN IDs in the VIO Server.
Figure 9. Virtual Ethernet adapter properties

Here the Trunk adapter is selected because this VIO Server is going to be using the SEA function to transfer packets to and from the external network. The IEEE 802.1Q is selected so that you can add additional VLAN IDs. The additional VLAN IDs have been typed in the Add field and the Add button clicked to add them to the list of VLAN IDs lower down the panel. This gives you all the VLAN IDs on the internal network that the VIO Server is allowing to arrive at this particular virtual Ethernet adapter/port. Note: These are the same as the PVIDs in the VIO client LPARs. This means packets that arrive with the VLAN tag that the Hypervisor adds can pass over the SEA.

Requirement 3: No stripping of VLAN tags on the VIO Server

You do not want the VLAN ID added to the packets as it leaves VIO client LPARs (set using the PVID number) to be removed on entering this VIO Server LPAR. Note that in Figure 10 below, the PVID is set to a seemingly weird number, 88 for example. This must be an otherwise unused VLAN ID (clearly it must not be in a PVID either, as it is the same thing). If a packet arrived with this VLAN ID (for example, 88), then the VLAN ID tag would be stripped off (untagged) and it could not be forwarded on to the right external VLAN. For example, if a packet with VLAN ID 88 arrived, then 88 would be removed. If this was sent through the SEA to the external physical network, it would arrive at the Ethernet switch untagged. What happens next is purely up to the settings on the Ethernet switch. For example, it could be discarded or sent on to a default VLAN, but the chances of it going to VLAN ID 88 are remote unless your network administrator has explicitly set this up (for example, the Ethernet switch port has a PVID of 88). Untagged packets in a VLAN environment are generally a mistake or assumed to be on some unimportant or legacy network, and best avoided.
Figure 10. Requirement 3: No stripping of VLAN tags on the VIO Server

![Diagram showing Requirement 3: No stripping of VLAN tags on the VIO Server]

Requirement 4: Allow the SEA to transmit packets for many VLANs

If you are familiar with creating a SEA in a non-VLAN environment, then you will be familiar with the VIO Server command: `mkvdev -sea`. This creates a low-level "bridge" of the internal virtual Ethernet to the external physical Ethernet. It is not obvious, but it does this at a low level in the TCP/IP stack and is efficient because packets avoid rising through the TCP/IP stack before being forwarded -- that would take POWER5 CPU cycles. The command takes the virtual Ethernet adapter, physical Ethernet adapter, a couple of other parameters, and creates a higher level adapter. In the example below, virtual adapter ent2 and physical adapter ent0 creates the new super adapter ent3 -- it is this adapter (actually en3) on which you might add an IP address, but this is covered in more detail later. The example command is:

```
mkvdev -sea ent0 -vadapter ent2 -default ent2 -defaultid 89
```

This VIO Server command displays the new super adapter name in example en3, or you can find this by using the VIO Server command:

```
lsdev -virtual
```

Look for the SEA entry in the list.

Note 1: The "-default ent2" is the default internal virtual VLAN to send the packet on if the it is untagged. In this example, there is only one virtual Ethernet adapter, but this would be used in more complex cases of multiple virtual Ethernet adapters specified in the command.

Note 2: The "-defaultid 89" is the VLAN ID to use for untagged packets. In effect, this is the PVID of the SEA. See Figure 11 below. In this diagram, you are not expecting untagged packets. By using an unused number like this, these packets are dropped, as no client will accept VLAN 89 packets.
This newly created SEA would only forward untagged packets using the PVID 89; however, the Virtual Ethernet adapter behind it will not accept the VLAN ID, so all untagged packets would be dropped. You have to tell the VIO Server SEA that the other VLAN IDs are allowed. This is done with the VIO Server command:

\[ \text{mkvdev –vlan} \]

In this example, you want these three VLANs to have access to the external Ethernet. To do this, run the following commands on the VIO Server:

- \[ \text{mkvdev –vlan ent3 –tagid 23} \]
- \[ \text{mkvdev –vlan ent3 –tagid 150} \]
- \[ \text{mkvdev –vlan ent3 –tagid 160} \]

This creates yet more Ethernet adapter names -- one for each of the previous commands.

If you have an internal virtual Ethernet that you don't what to be visible for security reasons, do not use the previous command on that VLAN. For example, there is no need for the network between the Application Server LPAR and the relational database management system (RDBMS) LPAR to appear externally to the physical Ethernet.

VIO Server, to IP address or not!

Here are some considerations when putting the VIO Server on the network:

1. VIO Server network access is optional.
You can access the VIO Server from a HMC using a secure and private HMC to service processor network. This makes having a network address and access to the VIO Server optional. If the VIO Server does not appear on any network at all, it makes the VIO Server very secure -- no network = no possible security risk.

2. Hostile network concerns
In a hostile network environment, making the VIO Server accessible on the network can be dangerous; if the VIO Server is compromised (hacked), then all networks are vulnerable, as the VIO Server can access them all. The hacker would have to assign an IP address to each VLAN adapter and then get access to them. If one of the VLANs is directly connected to the Internet network, you might have a problem. To avoid this, it might be best to use a physical adapter that is directly connected to the LPAR for vulnerable external networks. This also gives you the means to physically disconnect the cable to stop an attack.

3. Dynamic logical partition changes
If you want to perform dynamic logical partition changes to the VIOS LPAR (for example, changing the number of CPUs or the entitlement) it will need to be on a network that can provide communication between HMC and the VIOS OS.

4. Upgrading the VIO Server
There are two methods to upgrade the VIO server code -- from CD or using FTP. Make sure that you do give yourself the option to add an IP address to the VIO Server if you are going to use the FTP option.

For non-hostile networks, it is perfectly normal for the VIO Server to go on the network to make management a little more flexible. To do this, use a second virtual Ethernet adapter, add an IP address, and run the VIO Server command, `mktcpip`, on that second virtual Ethernet adapter and not the SEA itself. If you are operating VLANs for LPARs, you will want to put the VIO Server on a specific VLAN. In Figure 12 below, the VIO Server is being placed on VLAN 23.

**Figure 12. Putting the VIO Server on the network**

On the VIO Server, use the `mktcpip` command, for example:
mktcpip -hostname op34 -inetaddr 9.137.62.34 -interface en3 -netmask 255.255.255.0 -gateway 9.137.62.1

Notes:

- This extra virtual Ethernet adapter would make the SEA adapter name en4 -- one higher.
- The VIO Server command sometimes requires "ent" and "en" interfaces. This can get confusing. For example, if you use "ent" when "en" should have been used, then you will receive an error message about "entt" (note the extra "t"), because the command added the "t" to convert "en" to "ent".

Why not put the IP address of the VIO Server on the SEA?

As an alternative, you could use a second physical Ethernet in the VIO Server for network traffic directly to the VIO Server. For example, let's assume you want to back up over the network from the VIO Server. A second physical Ethernet adapter would be a good choice for this scenario. The reason for this is because the second physical Ethernet adapter can separate the client LPAR network traffic from large volume backups that might benefit from using large IP packets. In the example, you have a second adapter called en1 in the VIO Server that can be used.

The unused physical Ethernet port in the example was just a coincidence. It was included in the example because many machines have dual Ethernet port adapters or dual Ethernet ports on the motherboard. By having this in the example, it might warn you to be on the lookout for the case of more than one physical Ethernet port per adapter.

Summary

In this article, you have learned some of the basics for how a VLAN works and that the virtual Ethernet of the Hypervisor behaves just like a VLAN switch.

You have set up a VIO Server to allow access from many client LPARs using different VLAN IDs and how to set up the SEA to transfer Ethernet packets to and from external VLANs.

You have also seen the best way to place the VIO Server on a VLAN or a direct network for remote systems management or backup access.

Acknowledgements

Special thanks to Dave Williams and Chris Milsted for their input and review.
Related topics

• Look at these:
  • Advanced POWER Virtualization on IBM eServer p5 Servers: Introduction and Basic Configuration provides an introduction to the hardware and software that supports and manages the virtual I/O environment on POWER5 systems.
  • Learn how to install SUSE LINUX Enterprise Server 9 on POWER5.
  • Get more information on the virtualization capabilities of the IBM POWER5 servers.
  • Visit the IBM Hardware Information Center, your source for technical information about IBM servers.
  • Get your hands on application development tools and middleware products from DB2®, Lotus®, Rational®, Tivoli®, and WebSphere®. You can download evaluation versions of the products at no charge, or select the Linux or Windows® version of developerWorks’ Software Evaluation Kit.
  • Want more? The developerWorks eServer zone hosts hundreds of informative articles and introductory, intermediate, and advanced tutorials on the eServer brand.