IP-based VLAN switching for Network Services Virtualization
draft-rekesh-ipvlanswitching-01.txt

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on November 18, 2010.

Copyright Notice

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

The objective of this document is to define and formalize the use of IP-based VLAN switching in the context of virtualization of network services. Networking infrastructure appliances may host multiple
virtual instances of network services. IP address inspection on a packet coming from an external network such as the Internet can be used to switch the packet into a VLAN, through use of VLAN tags on Ethernet frames. The tagged frame is subsequently fed to a VLAN-aware networking infrastructure appliance where the packet is directed to a virtualized service instance. This method of IP-based VLAN switching on a packet assists in hosting multiple virtual instances of network services on servers and other appliances.

Table of Contents

1. Introduction ..................................................... 2
2. Conventions used in this document .............................. 3
3. Terminology ...................................................... 4
4. Usage Scenarios ................................................ 4
   4.1. ISP security services ................................... 4
   4.2. Datacenter Security services and Server Load balancing .... 5
5. Protocol Considerations ........................................ 7
   5.1. IP-based VLAN switching for Virtualization Support ........ 7
       5.1.1. 802.3 Ethernet frame ................................ 7
       5.1.2. 802.1q VLAN tagging ................................ 8
       5.1.3. 802.1q-in-q or Stacked VLANs ........................ 9
   5.2. Configuration ............................................... 10
       5.2.1. Configuration for service-facing ports ............. 10
       5.2.2. Configuration for external network facing ports ... 10
   5.3. Functionality ............................................... 11
       5.3.1. Egressing frames on service-facing ports .......... 11
           5.3.1.1. IPv4 unicast payload frames ................. 11
           5.3.1.2. IPv4 broadcast payload frames .............. 11
           5.3.1.3. IPv4 multicast payload frames ............. 11
       5.3.2. Ingressing frames on service-facing ports ......... 11
       5.3.3. Egressing frames on external network facing ports.... 12
       5.3.4. Ingressing frames on external network facing ports... 12
       5.3.5. non-IP unicast payload frames .................... 12
       5.3.6. non-IP broadcast payload frames ................... 12
       5.3.7. Non-IP multicast payload frames .................... 13
6. External Switching Device Considerations ...................... 13
7. Security Considerations ........................................ 13
8. IANA Considerations ........................................... 13
9. References ...................................................... 13

1. Introduction

Ethernet standards, as defined by IEEE 802.3 [1] are currently the most widely deployed wired LAN technology. Ethernet provides virtualization services using the concept of 802.1q VLANs [2].
Virtualization helps provide isolation between IP networks on the same physical LAN, utilize resources more efficiently, reduce clutter of appliances, and offer certain cost benefits.

Ethernet virtualization is often deployed in layer 2 switches, and has steadily gained popularity in supporting network and security services virtualization in network appliances. Such virtualization helps fray costs through optimal use of hardware resources. Examples include virtualization of security services like firewall, ipsec vpn and intrusion prevention at an ISP. The ISP would use some form of service virtualization within a single appliance to cater to multiple customers. Ethernet virtualization can assist in this process, identifying traffic for individual instances through the use of VLAN identifiers. Another example would be large data centers that deploy and load balance server farms for their customers. These deployments too can use ethernet virtualization to identify and segregate network traffic meant for various server farms.

Virtualization may thus rely on VLANs to distinguish traffic to specific service instances. In such cases, VLAN switching needs to be performed with IP address intelligence, since an IP packet entering the premises from outside may not have a VLAN ID already associated. Some element or logic within the networking infrastructure needs to identify the correct service instance that can service the packet, apply a suitable VLAN ID on the packet, and dispatch it to the device handling the service.

The above task of VLAN switching based on IP addresses (a.k.a IP-based VLAN switching) may be implemented as a software module integrated within a target appliance that provides the services, or within a separate switching device external to the target appliance.

This document provides a functional description of the IP-based VLAN switching module.

2. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC-2119 Error! Reference source not found.
3. Terminology

The term "service facing port" refers to a physical port or logical interface that connects to a network or device that hosts virtualized network services.

The term "external network facing port" refers to a physical port or logical interface that connects to a network that does not host virtualized network services.

The term "service domain" refers to a group consisting of a network facing port, all virtualized network services hosted on that network and reachable through the network facing port, along with the VLAN IDs associated with those services.

4. Usage Scenarios

This section describes some deployment scenarios in more detail. Note that these are only representative scenarios and that there can be many others deployments that can benefit from IP-based VLAN switching.

The term ‘service facing port’ is used to distinguish physical ports on the IP-based VLAN switch module that are connected to the device(s) that run virtualized services.

4.1. ISP security services

A network appliance at an ISP provides firewalling, ipsec vpn, intrusion prevention and other security services to the ISP’s customers. For cost reasons, a single appliance supports multiple virtual service instances that may be assigned to individual customers.

Security service instances within the appliance work with virtual interfaces that are mapped to separate VLAN IDs. Packets arriving from the Internet through the IP-based VLAN switch are destined for subscriber or customer networks and have destination IP addresses accordingly.

To distinguish which packet must be sent to which virtual instance of the service, the external IP-based VLAN switch may inspect the destination IP address against a set of preconfigured addresses, address ranges and/or address lists, and insert a distinguishing VLAN tag into the ethernet packet before forwarding it to the appliance. In essence, the packet is switched to one of many defined VLANs based on IP address inspection. Once the packet enters the appliance, this
VLAN id is used to direct the packet to the right virtual interface where the packet is processed by a service instance dedicated to that customer.

```
Customer/Subscriber Side
      ^   ^   ^   ^   ^
      |   |   |   |   |
      +-------+-------+------+
      |Provider Edge Router|
      '-----------------'  

ISP Security Appliance

|Virtual | Virtual | Virtual| ...| Virtual|
|Instance|Instance|Instance|     |Instance|

| 1    | 2    | 3    | 2048|
```

Service-facing Port (VLAN IDs 1,2,3...2048)

```
|IP-based VLAN switch|
'---------------------'  
      ^   ^   ^   ^
      |   |   |   |
     Internet Side
```

Figure 1 An ISP security services deployment

A schematic of the usage scenario is shown in Figure 1.

This method of inspection and switching in the external device requires IP subnets, IP addresses or a range of IP addresses to be mapped to a VLAN ID on the service-facing port(s) of the IP-based VLAN switch.

4.2. Datacenter Security services and Server Load balancing

In this deployment scenario, a network appliance at a data center provides server load balancing services to customers, along with
security services like traditional firewalls, web application firewalls and intrusion prevention. The appliance supports multiple virtual instances, with load balancing public IP addresses assigned for each customer.

Virtual service instances within the appliance operate on VLAN interfaces. Each service instance is assigned a unique public IP
address. A packet arriving from the Internet, destined for a particular load balancing IP address or customer network address, will be VLAN-tagged based on the destination IP address of the packet. This allows the packet can be sent to a VLAN interface and reach a virtual service instance on the appliance. This technique can require an IP subnet, IP address or a range of IP addresses to be mapped to a VLAN ID on the service-facing port(s) of the IP-based VLAN switch.

5. Protocol Considerations

5.1. IP-based VLAN switching for Virtualization Support

IP-based VLAN switching operates on service-facing 802.3 ports. A service-facing port is one that connects to a VLAN-based IP services network as described in section 3.

A service-facing port is part of a 'service domain'. A service domain is a set of IP services in the service network, along with the defined 802.3 ports and VLANs for those services.

Frames entering a service domain through a service-facing port are VLAN tagged as per 802.1q standards. VLAN tagging is configured specifically for the port, with VLAN IDs assigned based on configured IP subnets, IP addresses or ranges. VLAN tagging in this context applies only to service-facing ports, and only to frames egressing those ports. If a frame already has a VLAN tag, an outer VLAN tag is applied resulting in stacked VLANs.

The frames that exit a service domain through a service-facing port have those corresponding VLAN tags stripped. These are ingressing frames on service-facing ports. This stripping ensures that the IP-based VLAN tagging is specific and restricted to a given service domain, and will not conflict with other service domains on other ports, or with the general non-domain network, even if same VLAN IDs are used across them. If a frame has multiple VLAN tags (stacked VLANs), then the outermost VLAN tag is matched and stripped.

The IEEE 802.3 frame format is assumed in this document. This format, with VLAN tagging additions are summarized below.

5.1.1. 802.3 Ethernet frame

The 802.3 frame has the following format as shown in Figure 3:
### Acronyms

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE</td>
<td>64-bit preamble field</td>
</tr>
<tr>
<td>SFD</td>
<td>Start frame delimiter (8 bits)</td>
</tr>
<tr>
<td>DA</td>
<td>Destination MAC address</td>
</tr>
<tr>
<td>SA</td>
<td>Source MAC address</td>
</tr>
<tr>
<td>LT</td>
<td>Length field (carries the EtherType for Ethernet II frames)</td>
</tr>
<tr>
<td>Payload</td>
<td>802.3 LLC/SNAP headers and data</td>
</tr>
<tr>
<td>FCS</td>
<td>Frame Check Sequence</td>
</tr>
</tbody>
</table>

**Figure 3 802.3 Frame Format**

The 802.3 frame format is compatible in structure with other Ethernet frame formats. In such cases, the 2-byte Length Field (LT) contents are used to distinguish the frame type. A length field value exceeding the maximum frame size for an 802.3 frame indicates a different frame type.

For example Ethernet II frames use a value starting from 0800 hex. For Ethernet II frames, the payload section in the 802.3 format aligns with that of the Ethernet II frame data payload, and lacks 802.3 LLC/SNAP headers.

The reader is referred to IEEE 802.3 standards for frame format details [1] [2].

#### 5.1.2. 802.1q VLAN tagging

VLAN tagging follows the IEEE 802.1q specifications. This standard requires inserting the VLAN tag as an additional 4 bytes between the Source address and the EtherType/Length field. The VLAN tagged frame is formatted as in Figure 4.
802.1q Tagged Frame

<table>
<thead>
<tr>
<th>field</th>
<th>PRE</th>
<th>SFD</th>
<th>DA</th>
<th>SA</th>
<th>Tag</th>
<th>LT</th>
<th>Payload</th>
<th>FCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>bytes</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>46 - 1500</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>field</th>
<th>TPID</th>
<th>PRI</th>
<th>CFI</th>
<th>VID</th>
</tr>
</thead>
<tbody>
<tr>
<td>bits</td>
<td>16</td>
<td>3</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

Figure 4 802.1q VLAN Tagging

The Tag Protocol Identifier (TPID) carries an Ethernet Type value which is used to identify the frame as a VLAN tagged frame.

This document is not specifically concerned about various details of the frame format, other than using the fact that VLAN tagging is supported on Ethernet frames. The reader is referred to 802.3 standards for more details [1] [2].

5.1.3. 802.1q-in-q or Stacked VLANs

When 802.3 frames are already VLAN tagged, stacked VLAN frames may be used to insert the additional VLAN tag. The format of such a doubly-tagged frame is summarized in Figure 5.
Doubly-tagged Frame

<table>
<thead>
<tr>
<th>field</th>
<th>PRE</th>
<th>SFD</th>
<th>DA</th>
<th>SA</th>
<th>Tag2</th>
<th>Tag1</th>
<th>LT</th>
<th>Payload</th>
<th>FCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>bytes</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>46 - 1500</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 5 Stacked VLANs

For IP-based VLAN switching, the tag with the desired VLAN ID is inserted before (i.e. nesting over) any existing tags.

5.2. Configuration

5.2.1. Configuration for service-facing ports

The module MUST provide the ability to configure one or more IPv4 subnet addresses against a VLAN ID for all service-facing ports.

The IP-based VLAN switching module SHOULD provide the ability to configure one or more IPv4 addresses against a VLAN ID on all service-facing ports.

The IP-based VLAN switching module SHOULD provide the ability to configure multiple IPv4 addresses against a VLAN ID on all service-facing ports.

The module SHOULD provide the ability to configure one or more IPv4 address ranges against a VLAN ID for all service-facing ports.

The module SHOULD provide the ability to configure IPv4 broadcast addresses against a VLAN ID for all service-facing ports.

The module SHOULD provide the ability to specify the 802.1p priority level to be used in the priority field of an inserted tag.

The module SHOULD provide the ability to enable or disable IP-based VLAN tagging support for egressing frames on service-facing ports.

The module SHOULD provide the ability to enable or disable IP-based VLAN tagging support for ingressing frames on service-facing ports.

5.2.2. Configuration for external network facing ports

IP-based VLAN switching does not have any specific requirements for these ports, other than following Ethernet standards.
5.3. Functionality

This section discusses the operation of the IP-based VLAN switching module in more detail.

5.3.1. Egressing frames on service-facing ports

Much of IP-based VLAN switching functionality is exercised when frames are sent out on service-facing ports. Various cases are covered below.

5.3.1.1. IPv4 unicast payload frames

If a service-facing port is configured with one or more IPv4 subnets, addresses, or ranges, and an egressing frame on the port contains an IP packet with an IPv4 unicast address that matches the configuration, then a VLAN tag with the corresponding VLAN ID MUST be inserted into the frame before (i.e. nesting over) any existing VLAN tags. If an 802.1p priority has also been configured, it MUST be applied to the frame.

5.3.1.2. IPv4 broadcast payload frames

If an egressing frame on a service-facing port contains an IP packet with a configured IPv4 broadcast address that matches subnets, IP addresses or ranges in the configuration, then a VLAN tag with the corresponding VLAN ID MUST be inserted into the frame, before (i.e. nesting) any existing VLAN tags. If there are multiple matches, then the frame MUST be duplicated for each such match and the corresponding VLAN tag inserted into each frame.

5.3.1.3. IPv4 multicast payload frames

If an egressing frame on a service-facing port contains an IP packet with an IPv4 multicast address, then the frame SHOULD be duplicated for each VLAN ID defined in the IP-based VLAN switching configuration, and the respective VLAN ID inserted into each tag.

5.3.2. Ingressing frames on service-facing ports

If a service-facing port has an IP-based VLAN switching configuration that is enabled, then an ingress frame on the port carrying a VLAN tag that matches one of the configured VLAN IDs, MUST have that VLAN tag stripped from the frame before forwarding. If there are multiple
VLAN tags in the frame, the matching should be done against the outermost VLAN tag.

5.3.3. Egressing frames on external network facing ports

IP-based VLAN switching has no additional impact on these frames, which are handled as per Ethernet standards.

5.3.4. Ingressing frames on external network facing ports

The IP-based VLAN switching module may conform to any layer 2, 3 or higher layer processing functionality as required, with no additional requirements beyond imposing Ethernet standards on the processing of these frames.

For example, the frames may be processed as in a layer 2 switch using mac address learning and forwarding, or they may be processed as in a layer 3 switch or router, with routing tables determining the egress port. Egress ports may also be determined based on inspecting a frame’s IP payload destination address, and matching with IP-based VLAN configuration for service-facing ports.

See also section 6 "External Switching Device Considerations".

5.3.5. non-IP unicast payload frames

If a frame egressing a service-facing port carries a non-IP packet, and the frame is a unicast frame, a configuration SHOULD be provided offering one of the following actions based on the EtherType field (DIX format) or the Link Service Access Point (LSAP) Identifiers (802.3 LLC) field.

a) Duplicate the frame for each IP-based VLAN ID associated with the port, and insert the appropriate VLAN tag

b) Send the frame unmodified

c) Drop the frame

This covers various scenarios that may be required for non-IP frames.

5.3.6. non-IP broadcast payload frames

Egressing non-IP broadcast frames on a service-facing port SHOULD be duplicated per IP-based VLAN ID configured on the port. For protocols like ARP to function properly, this is a requirement.
5.3.7. Non-IP multicast payload frames

Egressing non-IP multicast frames on a service-facing port SHOULD be duplicated per IP-based VLAN ID configured on the port.

6. External Switching Device Considerations

IP-based VLAN switching module functions may be integrated into layer 2 switches or bridges, layer 3 switches, routers, load balancers etc. The core functionality as defined by the module is performed in the context of a service-facing port, and is not expected to depend on the packet or frame forwarding mechanisms used within the overall device. For example, a layer 2 switch may use an fdb to forward frames. A layer 3 switch or router may use a routing table. As ethernet frames are received or sent out on service-facing ports, the IP-based VLAN switching logic is applied.

7. Security Considerations

Unless carefully implemented, the requirement to duplicate frames per VLAN ID in certain scenarios described above can create a denial-of-service vulnerability. The module should minimize the amount of resources utilized in this case, such as by avoiding creation of all duplicate frames at once. Frame duplication may be delayed until needed, just before actual transmission. Rate-limiting controls may also be used.

Beyond this, no additional security concerns are envisaged.

8. IANA Considerations

None

9. References

[1] "IEEE Standard for information technology telecommunications and information exchange between systems Local and metropolitan area networks" IEEE 802.3-2005/Cor 1-2006


Authors’ Addresses

John Rekesh
Freescale Semiconductor, Inc.
890 N McCarthy blvd, Ste 120
Milpitas, CA 95035

Phone: (408) 904-2756
Email: john.rekesh@freescale.com

Srinivasa Addepalli
Freescale Semiconductor, Inc.
890 N McCarthy blvd, Ste 120
Milpitas, CA 95035

Phone: (408) 904-2761
Email: saddepalli@freescale.com