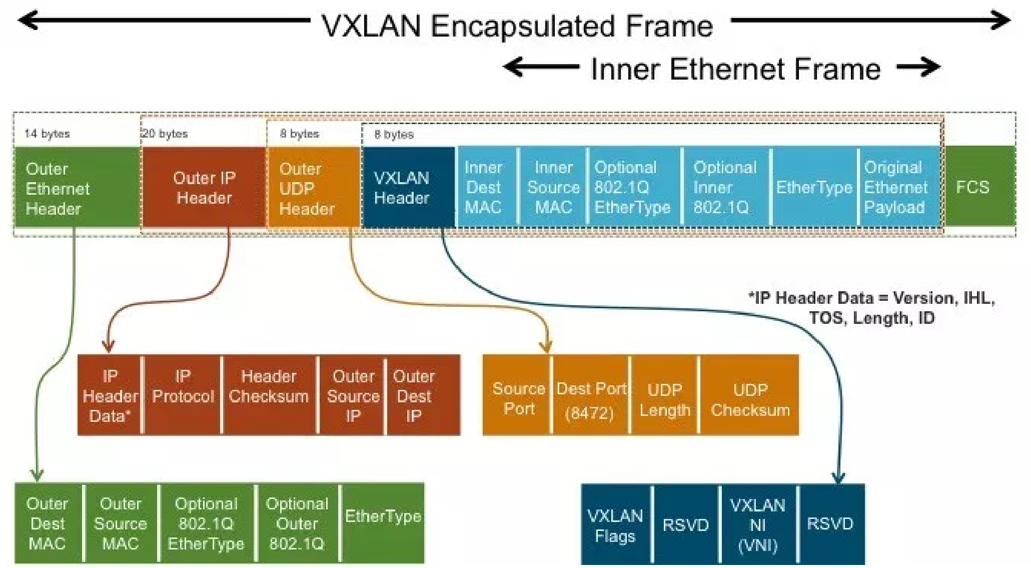
NSX for vSphere and VXLAN UDP Port Number

In this blog post I thought it might be a good idea to share with you the current state of the VXLAN default port number and how this is implemented in the current NSX for vSphere platform.

So, just in case you were not aware and this question get's raised from time to time, and especially with Service Providers who are for example integrating and using VXLAN with say the Cisco N1KV, and which raises a very good point on the issue that the default port number used by VMware for VXLAN is port 8472, however in April 2013, IANA reserved UDP port number 4789 for VXLAN. This is also documented in the IETF RFC7348 "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Network".

The following diagram depicts the VXLAN Frame Format



To explain the above VXLAN Frame Format, let us start with the Inner frame which means the original frame sent out by a VM before VXLAN encapsulation. Then during encapsulation, a VXLAN header including VNI and a few flag bits (a single instance bit in the official draft to IETF), an UDP header, an IP header, and last an ethernet header are added to the orignal frame. The encapsualted frame is sent out to the physical network. In the NSX for vSphere release, the outer UDP header is using destination port 8472. The source port is calculated based on flows of the inner frame. This is mainly for ECMP to have physical switch/router support ECMP be able to distribute different flows to different paths for load balancing. The outer IP header uses the destination VTEP's IP as the destination IP, and the source VTEP IP as the source IP. If the frame is a BUM packet (broadcast, unkown unicast (destination VTEP unknown), or multicast) a per VNI multicast group is used for the outer destination IP. The outer ethernet header may include a VLAN tag (the physical VLAN used for VXLAN transport).

So, does this mean that VMware still use the UDP port 8472 for VXLAN? The answer is yes, we currently still use UDP port 8472 for VXLAN. So, why was this done I hear you ask? It was mainly due to compatibility considerations when upgrading why we still use the VMware port number. But what are the concern's with changing the port number and how will this affect customers upgrading from previous versions? All great questions I feel. Let me see if I can shed some light on this.

1. When we set the port number (via the REST API, or if we where to perform it automatically during an upgrade process) there will be a dataplane outage. We, VMware, can not atomically update the port number on all ESXi hosts in the environment, so in the time between the first and last ESXi host receiving the port number update message (plus

some time for any packets already in transit) some packets will be mis-parsed and dropped. Not desirable I am sure you would agree.

2. Now it a customer has a firewall(s) or other similar devices on the physical network, they could be set as default to blocking UDP port 4789, so if we automatically change the port number it could completely break VXLAN traffic, and we would have no way of knowing before hand. In a greenfield installation, you would have flushed out such issues before deploying any production workloads, but if this happens to be on an upgrade, it could be catastrophic.

So, for these reasons, and at this time we still need to stay with using the UDP port number of 8472. The good news is that we do have the ability to change the UDP port number using the REST API, but ONLY the using the REST API for those customers who want to change that. The plan moving forward is we, VMware have discussed the need for using the IANA port number to interoperate with HW VTEPs. Nothing is committed at this time, but one plan is to prompt the user, with a warning about the above potential issues to change the UDP port when they try to enable the HW VTEPs in the UI and the port number is not 4789, and in the case it would nit be possible to enable HW VTEPs until this is completed, but we would still require explicit user action to change the port number so they are fully aware of any potential outages.

This being the case, and as mentioned above there is a way to change the UDP port number of 8472 for VXLAN using the REST API. What is the REST API call? The REST API call to change the UDP VXLAN port number is as follows:

```
METHOD : PUT
URL : /2.0/vdn/config/vxlan/udp/port/<port_number>
```