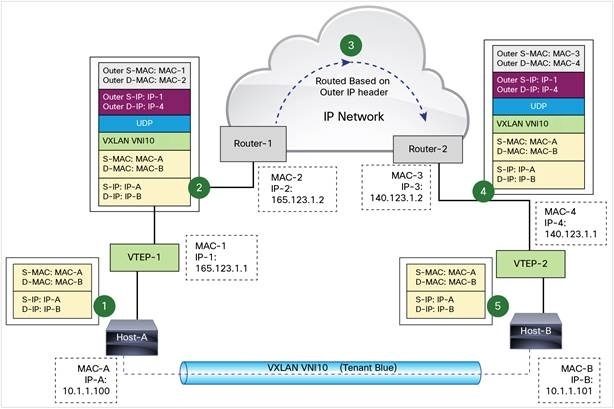


上一节我们介绍了 VXLAN 的封装格式以及 VTEP。  
今天我们将通过例子讨论 VXLAN 封装和转发包的过程，以及 Linux 对 VXLAN 的原生支持。

**VXLAN 包转发流程**

VXLAN 在 VTEP 间建立隧道，通过 Layer 3 网络传输封装后的 Layer 2 数据。  
下面的例子演示了数据如何在 VXLAN 上传输：



图中 Host-A 和 Host-B 位于 VNI 10 的 VXLAN，通过 VTEP-1 和 VTEP-2 之间建立的 VXLAN 隧道通信。  
数据传输过程如下：

1. Host-A 向 Host-B 发送数据时，Host-B 的 MAC 和 IP 作为数据包的目标 MAC 和 IP，Host-A 的 MAC 作为数据包的源 MAC 和 IP，然后通过 VTEP-1 将数据发送出去。
2. VTEP-1 从自己维护的映射表中找到 MAC-B 对应的 VTEP-2，然后执行 VXLAN 封装，加上 VXLAN 头，UDP 头，以及外层 IP 和 MAC 头。此时的外层 IP 头，目标地址为 VTEP-2 的 IP，源地址为 VTEP-1 的 IP。同时由于下一跳是 Router-1，所以外层 MAC 头中目标地址为 Router-1 的 MAC。
3. 数据包从 VTEP-1 发送出去后，外部网络的路由器会依据外层 IP 头进行包路由，最后到达与 VTEP-2 连接的路由器 Router-2。
4. Router-2 将数据包发送给 VTEP-2。VTEP-2 负责解封数据包，依次去掉外层 MAC 头，外层 IP 头，UDP 头 和 VXLAN 头。
5. VTEP-2 依据目标 MAC 地址将数据包发送给 Host-B。

上面的流程我们看到 VTEP 是 VXLAN 的最核心组件，负责数据的封装和解封。

隧道也是建立在 VTEP 之间的，VTEP 负责数据的传送。

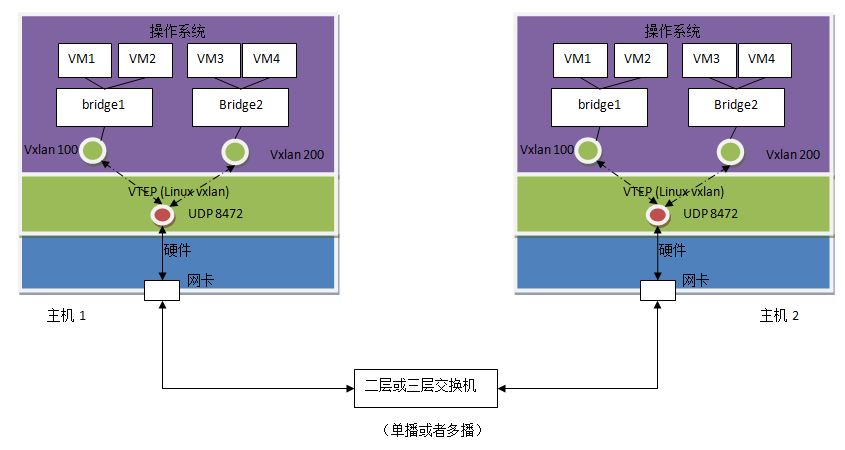
**Linux 对 VXLAN 的支持**

VTEP 可以由专有硬件来实现，也可以使用纯软件实现。

目前比较成熟的 VTEP 软件实现包括：

1. 带 VXLAN 内核模块的 Linux
2. Open vSwitch

我们先来看 Linux 如何支持 VXLAN，Open vSwitch 方式将在后面章节讨论。



实现方式：

1. Linux vxlan 创建一个 UDP Socket，默认在 8472 端口监听。
2. Linux vxlan 在 UDP socket 上接收到 vxlan 包后，解包，然后根据其中的 vxlan ID 将它转给某个 vxlan interface，然后再通过它所连接的 linux bridge 转给虚机。
3. Linux vxlan 在收到虚机发来的数据包后，将其封装为多播 UDP 包，从网卡发出。

到这里，相信大家对 VXLAN 的原理已经有了大致的了解。

下节我们将学习如何在 Neutron 中配置和实施 VXLAN。