

前面我们学习了 VXLAN，今天讨论跟 VXLAN 紧密相关的 L2 Population。

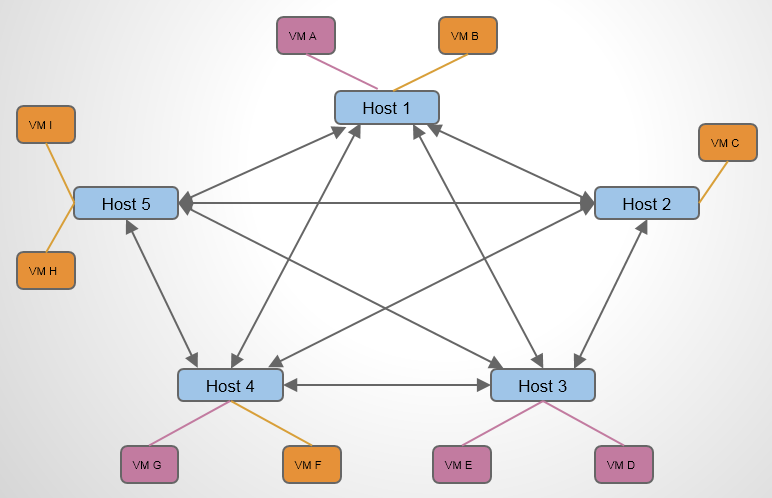
**L2 Population 是用来提高 VXLAN 网络 Scalability 的**。

通常我们说某个系统的 Scalability 好，其意思是：

当系统的规模变大时，仍然能够高效地工作。

L2 Population 到底解决了怎样的 Scalability 问题？

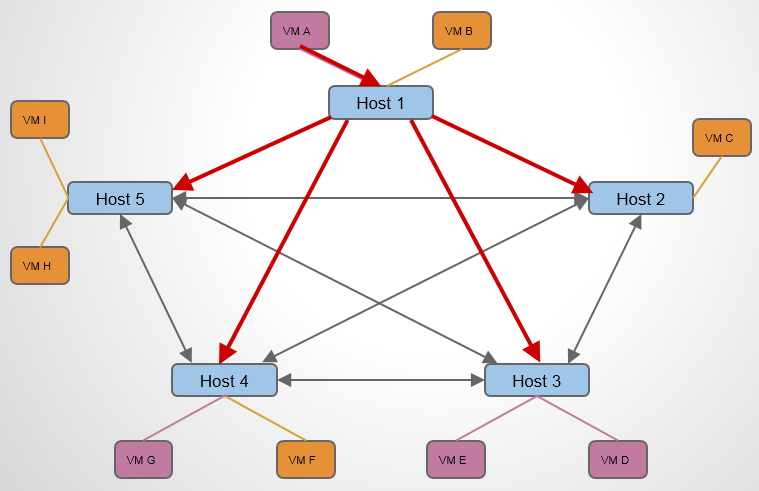
请看下图：



这是一个包含 5 个节点的 VXLAN 网络，每个节点上运行了若干 VM。

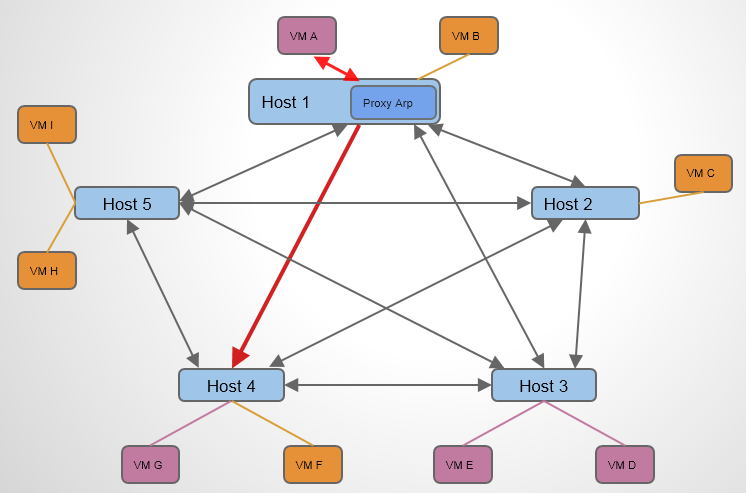
现在假设 Host 1 上的 VM A 想与 Host 4 上的 VM G 通信。  
VM A 要做的第一步是获知 VM G 的 MAC 地址。

于是 VM A 需要在整个 VXLAN 网络中广播 APR 报文：“VM G 的 MAC 地址是多少？”



如果 VXLAN 网络的节点很多，广播的成本会很大，这样 Scalability 就成问题了。

幸好 L2 Population 出现了。



L2 Population 的作用是在 VTEP 上提供 Porxy ARP 功能，使得 VTEP 能够预先获知 VXLAN 网络中如下信息：

1. VM IP -- MAC 对应关系

2. VM -- VTEP 的对应关系

当 VM A 需要与 VM G 通信时：

1. Host 1 上的 VTEP 直接响应 VM A 的 APR 请求，告之 VM G 的 MAC 地址。

2. 因为 Host 1 上的 VTEP 知道 VM G 位于 Host 4，会将封装好的 VXLAN 数据包直接发送给 Host 4 的 VTEP。

这样就解决了 MAC 地址学习和 APR 广播的问题，从而保证了 VXLAN 的 Scalability。

那么下一个关键问题是：  
**VTEP 是如何提前获知 IP -- MAC -- VTEP 相关信息的呢**？  
  
答案是：

1. Neutron 知道每一个 port 的状态和信息； port 保存了 IP，MAC 相关数据。
2. instance 启动时，其 port 状态变化过程为：down -> build -> active。
3. 每当 port 状态发生变化时，Neutron 都会通过 RPC 消息通知各节点上的 Neutron agent，使得 VTEP 能够更新 VM 和 port 的相关信息。
4. VTEP 可以根据这些信息判断出其他 Host 上都有哪些 VM，以及它们的 MAC 地址，这样就能直接与之通信，从而避免了不必要的隧道连接和广播。

理解了工作原理，下节我们学习如何在 Neutorn 中配置 L2 Population。