S5800下挂虚拟机迁移慢问题分析案例

一、组网：

无。

二、问题描述：

虚拟机VMX从母机A热迁移到母机B，母机A和母机B在同一个S5800下面，S5800作为虚拟机的网关。挂在另一台S5800下面的第三方服务器以0.01s的间隔ping VMX

母机B端的QEMU在内存，CPU寄存器，VHOST等迁移完成后，向外以子机的源mac向外发RARP报文，然后resume子机

发现如下现象：

热迁移过程中ping中断3.5s，从热迁移日志中可以看到，热迁移整体停机耗时500~600ms，迁移过程中在母机A端和母机B端抓包显示如下：

母机A：

11:01:56 收到母机B发出的RARP报文

11:01:59 收到最后一个ICMP ECHO

可以看到母机A端在收到B母机广播的RARP报文3s后才停止收到ping报文

母机B：

11:15:25 母机B QEMU 往外发RARP包

11:15:28 收到ICMP包

可看到母机B向外发RARP报文3s后才收到ping报文

从抓包中可以明确看到，S5800已经收到了RARP的网络报文，但是3s后才切换mac -> port的转发表转发对应mac的包到正确的母机。S5800交换机上为何会产生这3s的延迟，还需要研究分析一下问题的根本原因。

三、过程分析：

根据客户描述，查看虚拟机对应的ARP表项：

```
display arp 10.221.156.18
```

```
      Type: S-Static  D-Dynamic
```

```
IP Address      MAC Address      VLAN ID  Interface  Aging Type
```

```
10.221.156.18  38ea-a78c-4fb9      2    GE1/0/18   3    D
```

```
display arp 10.221.156.41
```

```
      Type: S-Static  D-Dynamic
```

```
IP Address      MAC Address      VLAN ID  Interface  Aging Type
```

```
10.221.156.41  38ea-a78c-528d      2    GE1/0/16   2    D
```

发现产生延迟问题的根本原因在于虚拟机迁移之后，S5800没有快速更新对应的ARP表项。实验室内根据现场反馈的抓包进行验证，发现：

1、客户发送的RARP报文，在S5800设备上不会直接根据这个报文更新ARP，如果宿主B发送免费ARP，则设备也能很快更新，这个更新时间在100ms以内。所以对于该问题，首先建议让宿主B发送免费ARP来提升虚拟机迁移业务切换时间。

2、如果虚拟机不发送免费ARP，发送的是RARP或者普通的广播报文来更新ARP，基于现场的应用，建议可以配置下面命令，该命令可以提升mac迁移更新ARP的性能。在实验室内经过实际测试也可以在100ms以内刷新ARP：

```
mac-address station-move quick-notify enable
```

此命令用来使能设备快速更新ARP表项功能。

```
undo mac-address station-move quick-notify enable命令用来恢复缺省情况。
```

缺省情况下，设备上并没有使能快速更新ARP表项功能。

四、解决方法：

可以通过下面两种方式来解决此问题：

- 1、调整虚拟机发送免费ARP来实现ARP的快速刷新。
- 2、如果不能调整虚拟机，则在设备上使能快速更新ARP表项功能
mac-address station-move quick-notify enable