N7K替换65引发的意外

目录

硬件平台

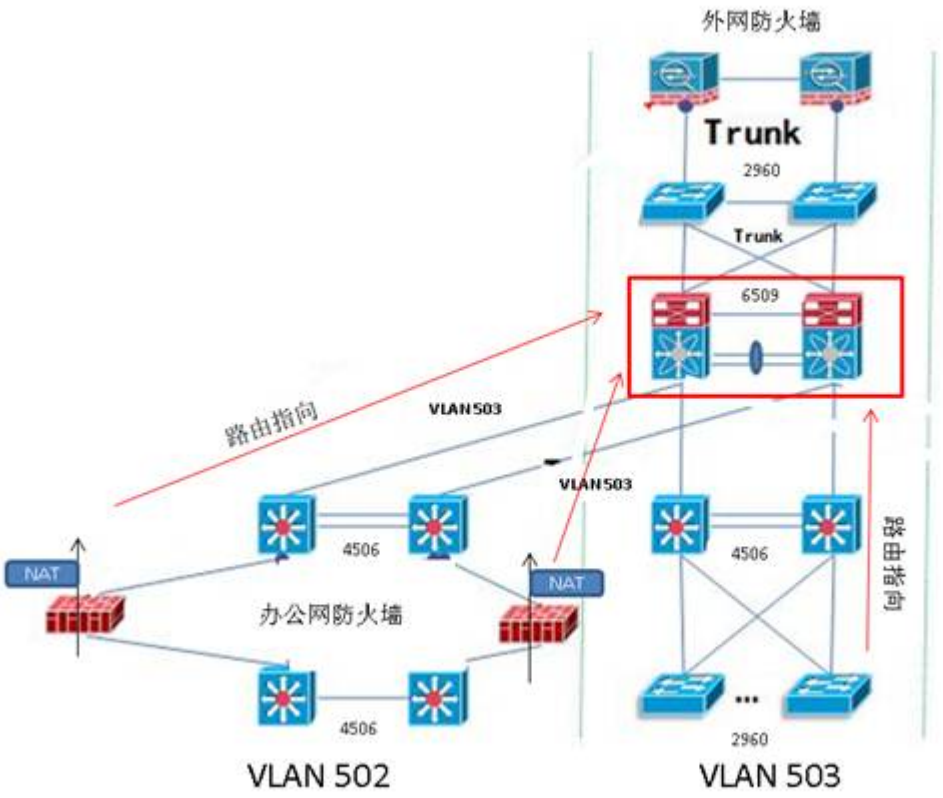
N7K， 6509

软件版本

N7K: 5.2(5)

6509：122-33.SXH5

故障发生前的网络状况：



1. 内部网络分为2个VLAN，分别是办公网（VLAN502）和业务网（VLAN503）；
2. 网络中所有的4506和2960交换机只提供2层连接）；
3. 办公网（VLAN502）的网关在办公网防火墙上，由办公网防火墙做NAT地址转换后转为业务网地址（VLAN503）传出，办公网防火墙的网关指向6509(升级后为N7K)的VLAN 503；
4. 业务网（VLAN503）的主机网关直接指向6509(升级后为N7K)的VLAN503；
5. 6509(升级后为N7K)上通过默认路由将流量由VLAN503引向外网防火墙。

6509和N7K上interface vlan 503的配置如下：

N7k:

```
Ip route 0.0.0.0 0.0.0.0 211.151.203.222()
interface Vlan503
  description business
  no shutdown
  ip address 211.151.203.194/27
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
  hsrp version 2
  hsrp 503
    preempt
    ip 211.151.203.193
```

6509：

（接口相关的OSPF的配置与N7K上完全一致，此处省略）

```
interface Vlan503
  description business
  ip address 211.151.203.194 255.255.255.224
  standby  ip 211.151.203.193
```

```
Ip route 0.0.0.0 0.0.0.0 211.151.203.222()
```

故障的触发原因：

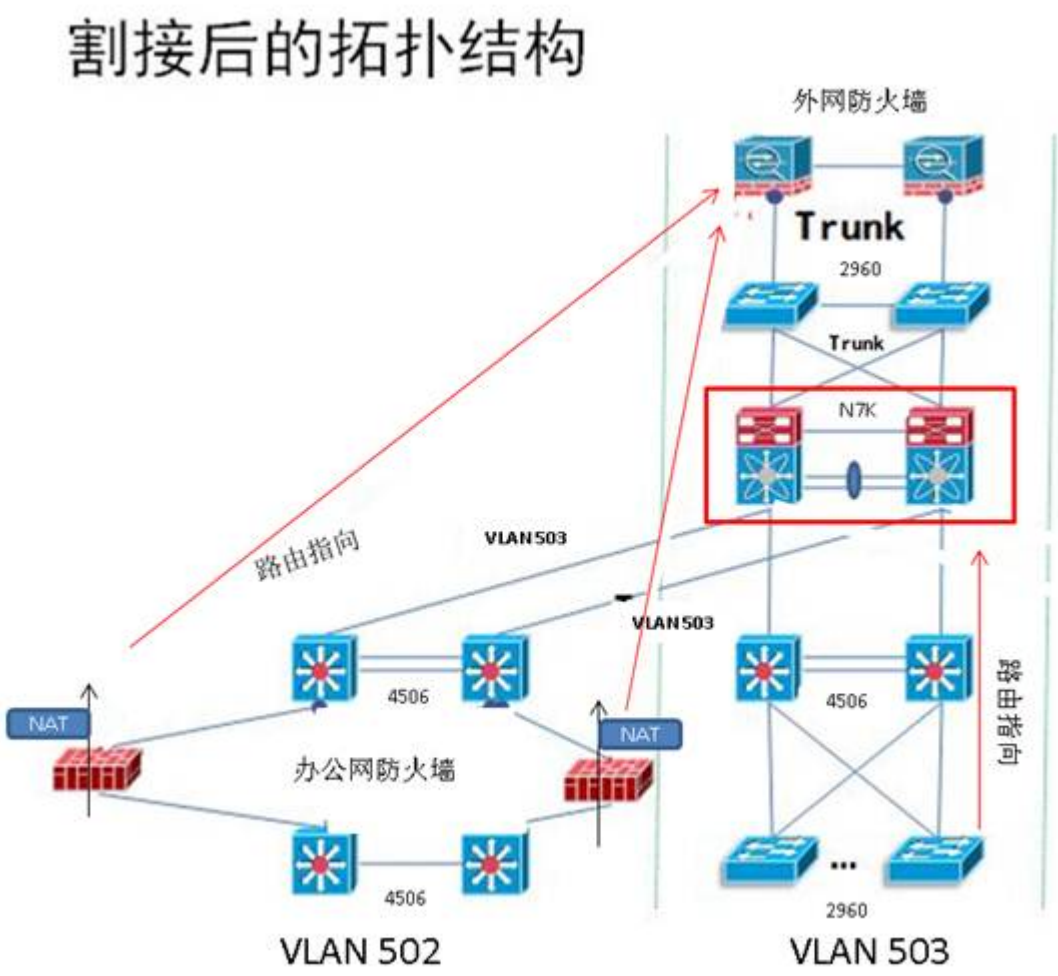
将6509升级为N7K，所有配置做相应命令格式的转换后从6509完整移植到N7K。

设备升级后的故障现象：

1. VLAN502上的主机到外网数据传输丢包严重；
2. VLAN 503上所有主机到外网数据完全中断。

临时解决方案：

1. 将内网防火墙的网关从N7K转到外网防火墙（如下图）上以后，VLAN502的丢包情况消除。



2. 检查VLAN503内服务器的配置，发现所有服务器的IP地址都是24位掩码而N7K上VLAN503的地址为27位掩码，将server上的地址掩码改为27位后，VLAN503的数据传输恢复正常。

问题：

1. 为什么办公网防火墙的网关指向N7K时办公网VLAN502的数据会丢包而业务网VLAN503没有这个问题？
2. 如果VLAN503里的主机因为某些历史原因IP地址的掩码与网关不一致，为什么用6509做网关就没有问题但用N7K就不行？

诊断：

针对问题1：

1. 在N7K上用ethanalyzer抓包，发现VLAN503中有大量数据上到CPU，如：

```
Frame 2 (70 bytes on wire, 70 bytes captured)
  Arrival Time: Jul 19, 2012 02:14:39.862858000
  [Time delta from previous captured frame: 0.000011000 seconds]
  [Time delta from previous displayed frame: 0.000011000 seconds]
  [Time since reference or first frame: 0.000011000 seconds]
  Frame Number: 2
  Frame Length: 70 bytes
  Capture Length: 70 bytes
  [Frame is marked: False]
  [Protocols in frame: eth:ip:icmp:ip:tcp]
Ethernet II, Src: 64:a0:e7:3f:60:c1 (64:a0:e7:3f:60:c1), Dst: 00:10:db:ff:40:90
(00:10:db:ff:40:90)
  Destination: 00:10:db:ff:40:90 (00:10:db:ff:40:90)
    Address: 00:10:db:ff:40:90 (00:10:db:ff:40:90)
    .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
    .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
  Source: 64:a0:e7:3f:60:c1 (64:a0:e7:3f:60:c1)
    Address: 64:a0:e7:3f:60:c1 (64:a0:e7:3f:60:c1)
    .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
    .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
  Type: IP (0x0800)
Internet Protocol, Src: 211.151.230.194 (211.151.230.194), Dst: 211.151.230.201
(211.151.230.201)
  Version: 4
  Header length: 20 bytes
  Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
    0000 00.. = Differentiated Services Codepoint: Default (0x00)
    .... ..0. = ECN-Capable Transport (ECT): 0
    .... ...0 = ECN-CE: 0
  Total Length: 56
  Identification: 0x31a4 (12708)
  Flags: 0x02 (Don't Fragment)
    0.. = Reserved bit: Not Set
    .1. = Don't fragment: Set
    ..0 = More fragments: Not Set
  Fragment offset: 0
  Time to live: 255
  Protocol: ICMP (0x01)
  Header checksum: 0xd564 [correct]
    [Good: True]
    [Bad : False]
  Source: 211.151.230.194 (211.151.230.194)
  Destination: 211.151.230.201 (211.151.230.201)
Internet Control Message Protocol
  Type: 5 (Redirect)
  Code: 1 (Redirect for host)
  Checksum: 0x1261 [correct]
  Gateway address: 211.151.230.222 (211.151.230.222)
  Internet Protocol, Src: 211.151.230.201 (211.151.230.201), Dst: 23.48.17.134 (23.48.17.134)
    Version: 4
    Header length: 20 bytes
    Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
        0000 00.. = Differentiated Services Codepoint: Default (0x00)
        .... ..0. = ECN-Capable Transport (ECT): 0
        .... ...0 = ECN-CE: 0
    Total Length: 40
    Identification: 0x31a4 (12708)
    Flags: 0x02 (Don't Fragment)
```

```
        0.. = Reserved bit: Not Set
        .1. = Don't fragment: Set
        ..0 = More fragments: Not Set
    Fragment offset: 0
    Time to live: 125
    Protocol: TCP (0x06)
    Header checksum: 0xe914 [correct]
        [Good: True]
        [Bad : False]
    Source: 211.151.230.201 (211.151.230.201)
    Destination: 23.48.17.134 (23.48.17.134)
Transmission Control Protocol, Src Port: 3916 (3916), Dst Port: www (80)
    Source port: 3916 (3916)
    Destination port: www (80)
    Sequence number: 383780779
```

2. 查看这些数据包，发现它们上CPU的原因是为了在CPU上触发ICMP的重定向消息；

3. 这些数据包都是从内网防火墙上送过来的，VLAN503里的server没有送数据包到N7K的CPU。

以上三点解释了VLAN502中数据丢包的原因，因为N7K上的COPP会丢弃超限的CPU流量，而这些流量全部都是VLAN502做地址转换后发出的数据包。

*** 为什么VLAN503的数据不需要上CPU触发重定向？**

到VLAN503的主机（windows系统）上查看相关路由表项，发现这些主机上不仅有默认路由，还有很多到外网的32位主机路由，下一跳指向外网防火墙。很显然，这些32位的路由是windows系统的主机依据N7K发出的ICMP 重定向消息添加的。这些主机在第一次将数据发往自己的网关N7K并收到N7K的重定向消息后就修改了自己的路由表，以后的数据直接交给外网防火墙处理而不再以N7K为网关。

所以问题1的答案是：windows主机会依据ICMP重定向消息修改自己的路由表项而防火墙不会这么做，所以VLAN503的数据只需由N7K的CPU处理一次，而VLAN502的数据则被内网防火墙源源不断地送往N7K的CPU去触发重定向，这些包有很大一部分被N7K依据COPP策略无情地丢弃了。

针对问题2：

问题2的答案其实非常明显，就是代理ARP。ARP-proxy在65的端口上的默认启动而N7K端口上默认关闭，所以即便65上interface VLAN503的IP地址和VLAN503里主机的IP地址掩码不同，它仍然会不辞劳苦地默默为主机提供服务。但在N7K上，在默认关闭proxy-ARP的前提下，N7K因为大家掩码不同而对所有VLAN503内的主机的请求都只是冷眼旁观，失去网关支持的VLAN503中的数据万念俱灰，于是。。。