关于NE80和NE40 IPC不通问题确认方法及HUB复位的

周灯台 2006-06-12 发表

关于NE80和NE40 IPC不通问题确认方法及HUB复位的
操作指导书

问题背景:
NE80/NE40采用分布式操作系统。主控板（指mpu和sru，下同）通过IPC通道和各单板（指lpu，nat，spu等，下同）进行通信完成协议处理和系统管理的功能。该通道在物理上是以以太HUB的形式实现的：主控板和单板内部各有两个以太端口，分别接在两块主控的hub模块上面。两个以太通道互为备份，确保系统的可靠性。
由于hub模块采用的intel芯片存在bug，使得hub的端口有可能与单板的端口link不上。如果系统是单主控板，主控板和单板之间IPC通道无法建立，主控板和单板之间无法通信；如果系统是双主控，若是hub与单板的两个端口都没有link上，单板将无法和主控板通信，若是hub与单板的一个端口link成功，与另外一个 link不成功，单板和主控之间通信正常，业务也不受影响，但是存在潜在的问题：如果link成功的端口突然因某种原因不通的话，单板将无法与主控通信，会被复位，业务中断。
该问题完全随机出现，而且无法通过硬件措施直接加以解决。软件可以通过一定的措施加以规避，但目前网上运行的大量NE80/NE40设备还没有这种措施。因此发布本指导书以解决现有网上问题。

问题现象:
IPC通道不通问题发生时，表现为如下现象之一：
(1)在系统冷启动（断电重起）或者热启动（命令复位或按键复位等）时有单板注册不上来；或者命令复位单板后单板注册不上。例如：

```
<Quidway>disp dev
NE80's Device status:
Slot #    Type    Online    Register      Status      Primary   NET Port
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
2         LPUB    Present   Registered    Normal      NA        Open
3         LPUB    Present   Registered    Normal      NA        Open
4         LPUB    Present   Registered    Normal      NA        Open
5         UNKNOWN Present   Unregistered  Abnormal    NA        Open
17        MPU     Present   NA            Normal      Master    NA
18        MPU     Present   NA            Normal      Slave     NA
19        NET     Present   Registered    Normal      Master    NA
21        CLK     Present   Registered    Normal      Master    NA
24        DPR     Present   Registered    Abnormal    NA        NA
25        FAN     Present   Registered    Normal      NA        NA
```

(2)所有单板注册成功，业务也正常。但是display communication-channel发现有的单板的某个通道不通。例如：

```
<Quidway>disp dev
NE80's Device status:
Slot #    Type    Online    Register      Status      Primary   NET Port
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
2         LPUB    Present   Registered    Normal      NA        Open
3         LPUB    Present   Registered    Normal      NA        Open
4         LPUB    Present   Registered    Normal      NA        Open
17        MPU     Present   NA            Normal      Master    NA
18        MPU     Present   NA            Normal      Slave     NA
19        NET     Present   Registered    Normal      Master    NA
21        CLK     Present   Registered    Normal      Master    NA
24        DPR     Present   Registered    Normal      NA        NA
25        FAN     Present   Registered    Normal      NA        NA

<Quidway>disp comm
System channel information:
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Slot#     Curr-Chan/Status    Slave-Chan/Status   Switch number
 2         IPC_B/Normal        IPC_A/Normal            3
```

```
 3        IPC_B/Normal        IPC_A/Abnormal          5
 4        IPC_B/Normal        IPC_A/Normal            5
 18        IPC_A/Normal        IPC_B/Normal           0
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

上面环境中显示3号lpu板的IPC_A通道不通。

注意：单主控的系统始终有一个通道不通是正常的，这是因为hub模块在主控板上面，单主控的话就只有单hub了。例如：

```
<Quidway>disp dev
NE80's Device status:
Slot #    Type    Online    Register      Status      Primary   NET Port
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
2        LPUB    Present   Registered    Normal      NA        Open
3        LPUB    Present   Registered    Normal      NA        Open
4        LPUB    Present   Registered    Normal      NA        Open
17        MPU     Present   NA            Normal      Master    NA
19        NET     Present   Registered    Normal      Master    NA
21        CLK     Present   Registered    Normal      Master    NA
24        DPR     Present   Registered    Abnormal    NA        NA
25        FAN     Present   Registered    Normal      NA        NA

<Quidway>disp comm
System channel information:
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Slot#     Curr-Chan/Status    Slave-Chan/Status   Switch number
  2        IPC_A/Normal        IPC_B/Abnormal          0
  3        IPC_A/Normal        IPC_B/Abnormal          0
  4        IPC_A/Normal        IPC_B/Abnormal          0
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

问题的确认：

仅依据上面描述的现象就认为IPC通道不通是不够严谨的，尤其是设备不注册的时候。那么如何确认是IPC通道问题呢？下面就各个现象进行逐步排查和问题确认。

a.现象：所有单板都注册不上。

步骤1：确保主控的启动文件正确，如果是硬盘启动，请确保硬盘没有被隔离；确保单板和版本配套关系正确。

步骤2：观察主控板面板的act灯，确定有无主用主控板。如果无主用主控板，则可能主控板未插好，或者硬件存在问题，请参照有关手册排查。

步骤3：用display dev看一下有无主用网板或交换模块。对NE80产品，用disp dev直接看网板的主备用状态，如果无主用网板，则可能网板未插好或者硬件存在问题，请参照有关手册排查。对NE40-8/NE40-4，用disp dev <sru_slot>可以看出sru板上交换模块是否正常，如果主sru板的交换模块不在位或异常，基本确定主sru硬件存在问题，请参照有关手册排查。NE40-2 sru板不存在交换模块。例如：

```
<Quidway>disp dev 9
Switch Module detail information:              //NE40-8 sru交换模块信息
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Net board status                          Normal
 Port loss of signal intermittent errors       0000
 Port crc errors                           0000
 Port yellow cell not received errors          0000
 Port yellow cell transmit time out            0000
 Port flow control violation                   0000
 Port loss of receive synchro error            0000
 CRC error counter from Pz0                    0
 CRC error counter from Pz1                    0
 HSCX master channel test                      Ok
 HSCX channel A errors counter                 0

Clock Module detail information:               //NE40-8 sru时钟模块信息
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 Current status:                       Normal
 Current source:                       0
 Workmode:                             auto
 Output SSM Level:                     lnc
 Module work:                          normal
```

```
SRAM:                          normal
Output 38.88MHz clock:             normal
PLL TRU050:                      normal
PLL 88915:                      normal
OSC:                           normal
I2C bus:                       normal
EPLD:                          normal
HDLC:                          normal
E1 port A:                     normal
E1 port B:                     normal
SSM control:                     off
Current source step into pull-in range
Clock is in hold mode
OSC working normal, warm-up completed

Sru9's detail information:
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Date of slot1 is reset: 2004-12-21
 ……
```

步骤4：display comm看一下IPC通道是不是都通。由于单板在启动过程中存在IPC通道abnormal的时候，所以需要分时观察多次，建议在两分钟内disp comm 十 次。对每块单板而言，IPC_A，IPC_B通道如果都abnormal，单板注册不上，否则就可以注册。如果确认所有单板的双通道一直都abnormal，那么可以基本确认是IPC通道问题。

b.现象：个别单板注册不上

步骤1：确保主控的启动文件正确，如果是硬盘启动，请确保硬盘没有被隔离；确保单板和版本配套关系正确。

步骤2：display dev <主控板槽位>看设备是否因为某种原因复位过。如果复位了表示单板这会儿可能正在启动；如果复位原因是自检失败，表示单板初始化过程中出现错误，此时用display logfile buffer或者more日志文件可以看到更具体的自检失败原因。如果单板三次因为自检不通过，则很可能单板本身发生故障。例如：

```
<Quidway>disp dev 17
Mpu17's detail information:
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Date of slot2 is reset: 2004-12-18
Time of slot2 is reset: 14:56:57
Sequence of event occur: 8
Reason is:Lpu selftest error， and reset lpu!
Date of slot2 is reset: 2004-12-18
Time of slot2 is reset: 14:54:13
Sequence of event occur: 7
Reason is:Lpu selftest error， and reset lpu!
Date of slot2 is reset: 2004-12-18
Time of slot2 is reset: 14:51:28
Sequence of event occur: 6
Reason is:Lpu selftest error， and reset lpu!
Date of slot2 is reset: 2004-12-18
Time of slot2 is reset: 14:48:40
Sequence of event occur: 5
Reason is:Board is pulled in!
<Quidway>disp logf b
…
#Dec18 14:56:56 2004 Quidway
SRM/5/SlfTstErr:OID=1.3.6.1.4.1.2011.2.17.0.157，Lpu2 selftest error!
%Dec 18 14:56:56 2004 Quidway SRM/5/FSUSelftest:LPU2 NP start error!
%Dec 18 14:56:57 2004 Quidway SRM/1/Reset:Slot2 is reset， the reason is:Lpu sel
ftest error， and reset lpu! .
…
```

上面2号板连续3次np启动失败。复位3次后，主控板不再复位2号板。2号板就一直注册不上来。

步骤3：display comm看IPC通道是不是通的。单板在启动加载阶段有时与主控不通，需要多看几次，建议在两分钟内十次。如果通了，则不是IPC不通问题。如果一直不通，则说明可能有问题。

步骤4：将单板拔出再插入，多次执行display comm发现看IPC通道仍是不通。则可能单板发生故障，如bootroom发生改写等问题，也可能是IPC通道不通。在使用其他手

段排除了单板故障（比如单板换槽位就注册上），则说明存在IPC通道不通问题。

c.现象：所有单板注册正常，业务也正常

步骤1：**display communication-channel**发现有的单板IPC通道不通。这种情况可以确定是IPC通道不通。这种情况只有双主控才可能发生。

NE80 hub复位操作指导:

由于NE80和NE40在硬件设计上有差异，所以分开介绍IPC通道不通问题的规避措施。操作在单主控和双主控的系统有不同，也分开说明。

注意：规避措施的核心是复位hub模块。由于问题的复杂性，下面的操作都在主用主控板的串口上操作，不能使用远程操作。此规定适合NE80和NE40.

1 单主控

单主控环境IPC通道不通的问题只能表现为所有或其中几块单板都注册不上。不可能存在所有单板注册正常，业务也正常的情况。在确认是IPC通道不通后，按如下步骤进行：

1)主控在17槽位：

步骤1：用户模式下执行reset slot hub_a；

步骤2：display comm看单板 IPC通道通不通，两分钟内执行十次。如果IPC通道通则意味着单板可能会注册上。IPC通道link成功时有dem channel 好的打印，如：

%Dec 18 15:53:32 2004 Quidway SRM/5/DEMInure:Slot 2 DEM channel 1 inures!

步骤3：步骤2步执行完后如果在十分钟以内所有单板都注册成功，本次操作有效，本次操作完成；如果十分钟后现象依旧，本次操作无效，可重复执行步骤1和2，每次间隔不低于10分钟。最多三次，本次操作完成。

2) 主控在18槽位：

步骤1：在诊断模式下执行set mpu mem *18* fe00a000 1 df，然后执行set mpu mem *18* fe00a000 1 ff。这一步可以事先写好脚本，然后拷贝到终端上面执行()，脚本如下：

sys

en_

set mpu mem *18* fe00a000 1 df

y

set mpu mem *18* fe00a000 1 ff

y

注意： 在两条set mpu mem之间要有200ms的时间间隔。如果使用北研vtp测试平台，可以设置命令行延时来保证。如果使用其他telnet平台，请分批拷贝两条命令到终端执行以保证间隔。由于执行set mpu mem *18* fe00a000 1 df后hub不工作了，需要尽快执行set mpu mem *18* fe00a000 1 ff以使hub恢复工作。所以间隔时间不宜太大。此外在y后面不要忘了敲“回车”键。在后面提到的在诊断模式下执行命令也要遵从此约束。

步骤2：display comm看单板 IPC通道通不通，两分钟内执行十次。如果IPC通道通则意味着单板可能会注册上。IPC通道link成功时有dem channel 好的打印，如：

%Dec 18 15:53:32 2004 Quidway SRM/5/DEMInure:Slot 2 DEM channel 1 inures!

步骤3：步骤2步执行完后如果在十分钟以内所有单板都注册成功，表示本次操作有效，本次操作完成；如果十分钟后现象依旧，本次操作无效，可重复执行步骤1和2，每次间隔不低于10分钟。最多三次，本次操作完成。

2 双主控

双主控环境IPC通道不通的问题表现为所有或其中几块单板都注册不上或者所有单板注册正常，业务也正常的情况。在确认是IPC通道不通后，按如下步骤进行：

步骤1：disp communication-channel查看故障通道哪个hub上。有两种情况：

1）故障分布在一块hub上面：所有abnormal都集中在ipc_a或者ipc_b上面。

2）故障分布在两块hub上面：有的abnormal在ipc_a，有的abnormal在ipc_b上面。

步骤2：如果故障都在一块HUB上面，则在用户配置模式下复位该hub。分两种情况：

1）17槽位为主18槽为备：

a）如果所有abnormal都集中在ipc_a，则执行reset slot hub_a命令；

b）如果所有abnormal都集中在ipc_b，则执行reset slot hub_b命令；

2）18槽位为主17槽为备：

a）如果所有abnormal都集中在ipc_a，则执行reset slot hub_b命令；

b）如果所有abnormal都集中在ipc_b，则执行reset slot hub_a命令；

步骤3：如果故障分布在两块hub上面：

1）把较多故障的hub上的业务往把较少故障的hub上业务切换，命令：

switch communication-channel IPC *slotid*;

2）复位较多故障的hub。同样：

a）17槽位为主18槽为备：

i 如果abnormal在ipc_a通道更多，则执行reset slot hub_a命令；
ii 如果abnormal在ipc_b通道更多，则执行reset slot hub_b命令；
b）18槽位为主17槽为备：
i 如果abnormal在ipc_a通道更多，则执行reset slot hub_b命令；
ii 如果abnormal在ipc_b通道更多，则执行reset slot hub_a命令；
步骤4：前面三步执行完后如果所有单板的通道都通表示本次操作有效，本次操作完成；如果现象依旧或者通道通断情况发生变化，可重复步骤1 - 3。每次间隔不低于10分钟。最多三次，本次操作完成。
例如：

```
<Quidway>disp comm
System channel information:
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Slot#     Curr-Chan/Status    Slave-Chan/Status   Switch number
  2         IPC_A/Normal        IPC_B/Abnormal        29
  3         IPC_A/Normal        IPC_B/Normal           6
  4         IPC_A/Normal        IPC_B/Normal           6
  5         IPC_B/Normal        IPC_A/Abnormal         6
  18        IPC_B/Normal        IPC_A/Abnormal         3
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

比较abnormal的情况，hub_a故障更多，我们复位hub_a。但这时2，3，4单板与通过ipc_a与17槽位通信，我们把3，4槽位的主通道切换到ipc_b上面。2号槽因ipc_b异常无法切换。

```
<Quidway>switch communication-channel ipc  3 b
Switch Succeeded!
<Quidway>switch communication-channel ipc  4 b
Switch Succeeded!
<Quidway>disp comm
System channel information:
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
  Slot#     Curr-Chan/Status    Slave-Chan/Status   Switch number
  2         IPC_A/Normal        IPC_B/Abnormal        29
  3         IPC_B/Normal        IPC_A/Normal           6
  4         IPC_B/Normal        IPC_A/Normal           6
  5         IPC_B/Normal        IPC_A/Abnormal         6
  18        IPC_B/Normal        IPC_A/Abnormal         3
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

最后我们复位hub_a，因为本例17槽位是master，所以直接复位hub_a:

```
<Quidway>reset slot hub_a
```

复位完后我们disp comm发现所有ipc_a通道都正常。只有2号槽ipc_b通道不正常，复位hub_b:

```
<Quidway>reset slot hub_b
```

复位完后我们disp comm发现所有通道都正常。本次操作完成。

五、NE40 hub复位操作指导

1 单主控

单主控环境IPC通道不通的问题只能表现为所有或其中几块单板都注册不上。不可能存在所有单板注册正常，业务也正常的情况。在确认是IPC通道不通后，按如下步骤进行：

步骤1：在诊断模式下执行s*et sru mem sruslotid fe00a000 1 ef*，然后执行*set sru mem sruslotid fe00a000 1 ff*。

这一步最好事先写好脚本，然后拷贝到终端上面，脚本如下：

```
sys
en_
set sru mem sruslotid fe00a000 1 ef
y
set sru mem sruslotid fe00a000 1 ff
y
```

注意： 在两条set sru mem之间要有200ms的时间间隔。如果使用北研vtp测试平台，可以设置命令行延时来保证。如果使用其他telnet平台，请分批拷贝两条命令到终端执行以保证间隔。由于执行set sru mem *sruslotid* fe00a000 1 df后hub不工作了，需要尽快执行set sru mem *sruslotid* fe00a000 1 ff以使hub恢复工作。所以间隔时间不宜太大。此外在y后面不要忘了敲“回车”键。在后面提到的在诊断模式下执行命令也要遵从此约束。

步骤2：display comm看单板 IPC通道通不通，两分钟内执行多次。如果IPC通道通则意味着单板可能会注册上。IPC通道link成功时有dem channel 好的打印，如：
%Dec 18 15:53:32 2004 Quidway SRM/5/DEMInure:Slot 2 DEM channel 1 inures!
步骤3：步骤2步执行完后如果在十分钟以内所有单板都注册成功，表示本次操作有效， 本次操作完成；如果十分钟后现象依旧，本次操作无效，可重复执行步骤1和2，每次间隔不低于10分钟。最多三次，本次操作完成。
2 双主控
双主控环境IPC通道不通的问题表现为所有或其中几块单板都注册不上或者所有单板注册正常，业务也正常的情况。在确认是IPC通道不通后，按如下步骤进行：
步骤1：disp communication-channel查看故障通道哪个hub上。有两种情况：
1）故障分布在一块hub上面：所有abnormal都集中在ipc_a或者ipc_b上面；
2）故障分布在两块hub上面：有的abnormal在ipc_a，有的abnormal在ipc_b上面。
步骤2：如果故障都在一个hub上面：
1）把故障的hub上的业务往把较少故障的hub上业务切换，命令：
switch communication-channel ipc *slotid*;
2）复位hub，命令：
sys
en_
set sru mem *sruslotid* fe00a000 1 ef
y
set sru mem *sruslotid* fe00a000 1 ff
y
*sruslotid*为该hub所在槽位。Ipc_a对应本板hub，ipc_b对应对等板 hub。
此处注意项同前。
步骤3：如果故障分布在两块hub上面并且主备主控板之间ipc_a通道不通：
1）先把单板主通道往ipc_b上面切换，命令：
switch communication-channel ipc *slotid*
2）复位ipc_a所在hub：
sys
en_
set sru mem *sruslotid* fe00a000 1 ef
y
set sru mem *sruslotid* fe00a000 1 ff
y
*sruslotid*为该hub_a所在槽位(即主用主控板槽位)。
此处注意项同前。
步骤4：如果故障分布在两块hub上面并且主备主控板之间ipc_a通道正常：
1）把较多故障的hub上的业务往把较少故障的hub上业务切换，命令：
switch communication-channel ipc *slotid*;
2）复位较多故障的hub。同样：
sys
en_
set sru mem *sruslotid* fe00a000 1 ef
y
set sru mem *sruslotid* fe00a000 1 ff
y
*sruslotid*为故障hub所在的主控板槽位。Ipc_a对应主用主控板hub，ipc_b对应备用主控板hub。
此处注意项同前。
步骤5：上述步骤执行完后如果所有通道恢复正常，本次操作完成。如果现象依旧或者通道状态发生变化，可重复执行步骤1 - 4。每次间隔不低于10分钟。最多三次，本次操作完成。

其他:
1 复位hub会造成主控和所有单板的一路通道瞬时断开。如果是单主控，意味着主控和单板通信中断。该过程可能持续2到5秒。这对于多数协议来说是可以容忍的，但对于rstp，
vrrp等部分二层协议可能产生影响，也许会导致瞬时的路径切换或回路。所以建议要在系统流量很小的时候进行操作。
2 版本说明：此指导书适用于VRP3.10-2321及以前版本。
3 如果在S8016遇到类似问题，处理方法参照NE80处理方法，但个别命令可能略有不同。
4 遇到其他问题请及时通知研发。