记一个 IP packet reassembles failure 导致的实例驱逐案例

付思明（Simon Fu）|Senior Support Engineer
Oracle Global Software Support

一般来说，对于 IPC Send timeout，可能的情况有以下几种：

1、节点本地盘 CPU 等待队列超高或 IO 繁忙、空闲物理内存用尽等，这种情况往往是相互伴随发生的，可以从 OSWatcher 的 vmstat 和 iostat 来发现；

2、私网网络发生丢包或异常，从 OSWatcher 的 netstat 和 trace route 输出中可以看到；

3、DRM 或 skgxp 等方面的 Oracle Bug，如（Doc ID 1594578.1）

这个案例属于上述第二种，但由于处理过程比较反复，且最后一次重现时问题指向 IP reassembles failure，而不是 UDP packet drop，所以记录下来以备今后参考。

一、首次发生的症状和分析：

心跳超时导致节点 2 被踢出，然而节点 2 回到集群后，由于 ASM 实例起不来，voting 不可用，再次被踢出。

<<<节点 1 gi alert log

```
2016-07-17 06:16:44.610:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(62109)]CRS-5011:Check of resource "**db" failed: details at "(:CLSN00007:)" in "/opt/rac/11.2.0/grid/log/**db1/agent/crsd/oraagent_oracle/oraagent_oracle.log"
2016-07-17 06:16:44.642:
[crsd(56996)]CRS-2765:Resource 'ora.**db.db' has failed on server '**db1'.
2016-07-17 06:17:01.103:
[cssd(34796)]CRS-1612:Network communication with node **db2 (2) missing for 50% of timeout interval. Removal of this node from cluster in 14.890 seconds《===cssd 网络心跳超时
2016-07-17 06:17:09.105:
[cssd(34796)]CRS-1611:Network communication with node **db2 (2) missing for 75% of timeout interval. Removal of this node from cluster in 6.890 seconds
2016-07-17 06:17:13.107:
[cssd(34796)]CRS-1610:Network communication with node **db2 (2) missing for 90% of timeout interval. Removal of this node from cluster in 2.880 seconds
2016-07-17 06:17:16.000:
[cssd(34796)]CRS-1607:Node **db2 is being evicted in cluster incarnation 363182878; details at (:CSSNM00007:) in /opt/rac/11.2.0/grid/log/**db1/cssd/ocssd.log.
2016-07-17 06:17:44.013:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(62109)]CRS-5818:Aborted command 'check' for resource 'ora.**db.db'. Details at (:CRSAGF00113:) {0:9:14530} in /opt/rac/11.2.0/grid/log/**db1/agent/crsd/oraagent_oracle/oraagent_oracle.log.
2016-07-17 06:17:46.550:
[cssd(34796)]CRS-1601:CSSD Reconfiguration complete. Active nodes are **db1 .
2016-07-17 06:17:46.577:
[crsd(56996)]CRS-5504:Node down event reported for node '**db2'.
2016-07-17 06:17:50.030:
[crsd(56996)]CRS-2878:Failed to restart resource 'ora.**db.db'
2016-07-17 06:17:50.036:
```

[crsd(56996)]CRS-2769:Unable to failover resource 'ora.**db.db'.
2016-07-17 06:17:50.148:
[crsd(56996)]CRS-2773:Server '**db2' has been removed from pool 'Generic'.
2016-07-17 06:17:50.148:
[crsd(56996)]CRS-2773:Server '**db2' has been removed from pool 'ora.**db'.
2016-07-17 06:17:50.149:
[crsd(56996)]CRS-2773:Server '**db2' has been removed from pool 'ora.oggb'.
2016-07-17 06:17:50.157:
[crsd(56996)]CRS-2787:Server pool 'ora.oggb' has fallen below its minimum size. Details at (:CRSPE00140:) {1:8812:33761} in /opt/rac/11.2.0/grid/log/**db1/crsd/crsd.log.
2016-07-17 06:17:52.714:
[cssd(34796)]CRS-1601:CSSD Reconfiguration complete. Active nodes are **db1 **db2 .<===节点 2 重新加入集群
2016-07-17 06:23:26.297:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(34688)]CRS-5011:Check of resource "+ASM" failed: details at "(:CLSN00006:)" in "/opt/rac/11.2.0/grid/log/**db1/agent/ohasd/oraagent_grid/oraagent_grid.log"
2016-07-17 06:23:26.375:
[ohasd(34144)]CRS-2765:Resource 'ora.asm' has failed on server '**db1'.《===节点 1 asm failed
2016-07-17 06:23:26.409:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(34688)]CRS-5011:Check of resource "+ASM" failed: details at "(:CLSN00006:)" in "/opt/rac/11.2.0/grid/log/**db1/agent/ohasd/oraagent_grid/oraagent_grid.log"
2016-07-17 06:23:26.440:
[crsd(56996)]CRS-2765:Resource 'ora.asm' has failed on server '**db1'.
...
2016-07-17 06:30:49.244:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(34688)]CRS-5019:All OCR locations are on ASM disk groups [VOTEDISK], and none of these disk groups are mounted. Details are at "(:CLSN00100:)" in "/opt/rac/11.2.0/grid/log/**db1/agent/ohasd/oraagent_grid/oraagent_grid.log".<===节点 1voting disk 全部不可用
...
2016-07-17 06:31:05.237:
[crsd(56996)]CRS-1013:The OCR location in an ASM disk group is inaccessible. Details in /opt/rac/11.2.0/grid/log/**db1/crsd/crsd.log.
2016-07-17 06:31:05.237:
[crsd(56996)]CRS-1006:The OCR location is inaccessible. Details in /opt/rac/11.2.0/grid/log/**db1/crsd/crsd.log.《===OCR 不可用
...
2016-07-17 06:31:05.607:
[ohasd(34144)]CRS-2765:Resource 'ora.crsd' has failed on server '**db1'.<===所以紧接着 crsd 挂掉
2016-07-17 06:31:21.220:
[cssd(34796)]CRS-1612:Network communication with node **db2 (2) missing for 50% of timeout interval. Removal of this node from cluster in 14.910 seconds《===cssd 网络心跳超时
2016-07-17 06:31:29.222:

[cssd(34796)]CRS-1611:Network communication with node **db2 (2) missing for 75% of timeout interval. Removal of this node from cluster in 6.910 seconds
2016-07-17 06:31:33.224:
[cssd(34796)]CRS-1610:Network communication with node **db2 (2) missing for 90% of timeout interval. Removal of this node from cluster in 2.910 seconds
2016-07-17 06:31:36.137:
[cssd(34796)]CRS-1607:Node **db2 is being evicted in cluster incarnation 363182880; details at (:CSSNM00007:) in /opt/rac/11.2.0/grid/log/**db1/cssd/ocssd.log.
2016-07-17 06:31:38.228:
[ohasd(34144)]CRS-8011:reboot advisory message from host: **db2, component: cssagent, with time stamp: L-2016-07-17-06:31:38.187
[ohasd(34144)]CRS-8013:reboot advisory message text: clsnomon_status: need to reboot, unexpected failure 8 received from CSS
2016-07-17 06:32:08.681:
[cssd(34796)]CRS-1601:CSSD Reconfiguration complete. Active nodes are **db1 <==节点 2 再次被踢出

<<<alert_+ASM1.log
Sun Jul 17 06:23:24 2016
IPC Send timeout detected. Receiver ospid 47129 [《===每次 asm 实例挂掉都是发生在 IPC Send timeout detected 之后。
Sun Jul 17 06:23:24 2016
Errors in file /opt/rac/grid/diag/asm/+asm/+ASM1/trace/+ASM1_lmd0_47129.trc:
Sun Jul 17 06:23:26 2016
Instance termination initiated by instance 2 with reason 1.
Instance 2 received a reconfig event from its cluster manager indicating that this instance is supposed to be down
Please check instance 2's alert log and LMON trace file for more details.
Please also examine the CSS log files.
LMON (ospid: 47127): terminating the instance due to error 481
System state dump requested by (instance=1, osid=47127 (LMON)), summary=[abnormal instance termination].
System State dumped to trace file /opt/rac/grid/diag/asm/+asm/+ASM1/trace/+ASM1_diag_47121_20160717062326.trc
Dumping diagnostic data in directory=[cdmp_20160717062326], requested by (instance=1, osid=47127 (LMON)), summary=[abnormal instance termination].
Instance terminated by LMON, pid = 47127
Sun Jul 17 06:23:30 2016
NOTE: No asm libraries found in the system
MEMORY_TARGET defaulting to 2415919104.
* instance_number obtained from CSS = 1, checking for the existence of node 0...
* node 0 does not exist. instance_number = 1
Starting ORACLE instance (normal)
LICENSE_MAX_SESSION = 0

这次可以从 OSWatcher 的 traceroute 输出中看到很大私网传输时延：

<<<**db1_prvtnet_16.07.17.0600.dat

zzz ***Sun Jul 17 06:17:11 CST 2016

traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets

1 **db1-priv1 (*.*.*.50) 0.087 ms 0.015 ms 0.008 ms

traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets

1 **db2-priv1 (*.*.*.51) 0.176 ms 0.409 ms 0.453 ms

zzz ***Sun Jul 17 06:17:27 CST 2016<==============================在这个心跳超时时间点上，节点 2 的私网 ip traceroute 产生很大时延，第 27 次发包才通。

traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets

1 **db1-priv1 (*.*.*.50) 0.047 ms 0.010 ms 0.007 ms

traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets

1 * * *

2 * * *

3 * * *

4 * * *

5 * * *

6 * * *

7 * * *

8 * * *

9 * * *

10 * * *

11 * * *

12 * * *

13 * * *

14 * * *

15 * * *

16 * * *

17 * * *

18 * * *

19 * * *

20 * * *

21 * * *

22 * * *

23 * * *

24 * * *

25 * * *

26 * * *

27 * * **db2-priv1 (*.*.*.51) 0.155 ms《==========

zzz ***Warning. Traceroute response is spanning snapshot intervals.《====私网中断导致 traceroute 响应跨越了采样间隔。

zzz ***Sun Jul 17 06:17:57 CST 2016

traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets

1 **db1-priv1 (*.*.*.50) 0.088 ms 0.010 ms 0.010 ms

traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets
 1 **db2-priv1 (*.*.*.51) 0.202 ms 0.192 ms 0.193 ms
且节点 2 重启后，仍然有持续的丢包现象：
zzz ***Sun Jul 17 06:31:48 CST 2016
traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 0.033 ms 0.007 ms 0.008 ms
traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets《====traceroute 再次出现持续丢包
 1 * * *
 2 * * *
 3 * * *
 4 * * *
 5 * * *
 6 * * *
 7 * * *
 8 * * *
 9 * * *
10 * * *
11 * * *
12 * * *
13 * * *
14 * * *
15 * * *
16 * * *
17 * * *
18 * * *
19 * * *
20 * * *
21 * * *
22 * * *
23 * * *
24 * * *
25 * * *
26 * * *
27 * * *
28 * * *
29 * * *
30 * * *
zzz ***Warning. Traceroute response is spanning snapshot intervals.
zzz ***Sun Jul 17 06:32:19 CST 2016
traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 0.050 ms 0.012 ms 0.023 ms
traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 460.791 ms !H 460.782 ms !H *
zzz ***Sun Jul 17 06:32:34 CST 2016

traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 0.038 ms 0.009 ms 0.011 ms
traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets《===这次丢包要比上一次更频繁
 1 * * *
 2 * * *
 3 * * *
 4 * * *
 5 * * *
 6 * * *
 7 * * *
 8 * * *
 9 * * *
10 * * *
11 * * *
12 **db1-priv1 (*.*.*.50) 1176.770 ms !H * *
zzz ***Sun Jul 17 06:32:49 CST 2016
traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 0.142 ms 0.010 ms 0.009 ms
traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets
 1 * * *
 2 * * *
 3 * * *
 4 * * *
 5 * * *
 6 * * *
 7 **db1-priv1 (*.*.*.50) 812.835 ms !H * *
zzz ***Sun Jul 17 06:33:04 CST 2016
traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 0.060 ms 0.010 ms 0.017 ms
traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 24.540 ms !H 24.522 ms !H 24.481 ms !H
zzz ***Sun Jul 17 06:33:19 CST 2016
traceroute to **db1-priv1 (*.*.*.50), 30 hops max, 60 byte packets
 1 **db1-priv1 (*.*.*.50) 0.068 ms 0.020 ms 0.026 ms
traceroute to **db2-priv1 (*.*.*.51), 30 hops max, 60 byte packets
 1 * * *
 2 * * *
 3 * * *
 4 * * *
 5 * * *
 6 * * *
 7 **db1-priv1 (*.*.*.50) 197.495 ms !H 197.508 ms !H 197.497 ms !H

从 netstat 输出来看，一直存在 packets to unknown port received 增长，在 6:17 增速变大，但差异不是很明显。

zzz ***Sun Jul 17 06:14:55 CST 2016
    4061139294 total packets received
Udp:
    2971717805 packets received
    2157844 packets to unknown port received.<===2157844-2157812=32
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:15:10 CST 2016
    4061141080 total packets received
Udp:
    2971718916 packets received
    2157876 packets to unknown port received.<=== 2157876-2157844=32
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:15:26 CST 2016
    4061142796 total packets received
Udp:
    2971719982 packets received
    2157908 packets to unknown port received.<=== 2157908-2157876=32
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:15:41 CST 2016
    4061144499 total packets received
Udp:
    2971721039 packets received
    2157940 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:15:56 CST 2016
    4061146297 total packets received
Udp:
    2971722135 packets received
    2157972 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:16:11 CST 2016
    4061148021 total packets received
Udp:
    2971723201 packets received
    2158004 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:16:26 CST 2016
    4061149204 total packets received

Udp:
    2971723920 packets received
    2158036 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:16:41 CST 2016
    4061150165 total packets received
Udp:
    2971724507 packets received
    2158068 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:16:56 CST 2016
    4061151213 total packets received
Udp:
    2971725024 packets received
    2158183 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:17:11 CST 2016
    4061151985 total packets received
Udp:
    2971725503 packets received
    2158228 packets to unknown port received.
    31 packet receive errors
    SndbufErrors: 162
zzz ***Sun Jul 17 06:17:27 CST 2016
    4061152501 total packets received
Udp:
    2971725791 packets received
    2158265 packets to unknown port received.《===2158265-2158228=37
    31 packet receive errors
    SndbufErrors: 162

因此，建议客户：

1、检查私网稳定性；

2、请提供下两节点的 udp buffer 配置，因为从 netstat 输出来看，udp packets to unknown port received 一直在增长:

#cat /etc/sysctl.conf

并在之后建议将 udp buffer 调大为如下值：

net.core.rmem_default = 4194304
net.core.rmem_max = 4194304
net.core.wmem_default = 4194304
net.core.wmem_max = 4194304

二、第二次重现的症状和分析

1、症状：

与上次有所差异，NHB 丢包情况有所缓解，所以这次直接就是 asm 实例 ipc send timeout 导致 asm 挂掉，然后是 voting 引起的 node eviction，类似于上次节点 2 由于网络心跳超时被驱逐后的情况：

<<<GI alert log

2016-08-09 13:17:34.560:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(58754)]CRS-5818:Aborted command 'start' for resource 'ora.asm'. Details at (:CRSAGF00113:) {0:15:1653} in /opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log.
2016-08-09 13:17:36.566:
[ohasd(58104)]CRS-2757:Command 'Start' timed out waiting for response from the resource 'ora.asm'. Details at (:CRSPE00111:) {0:15:1653} in /opt/rac/11.2.0/grid/log/**db2/ohasd/ohasd.log.
2016-08-09 13:18:06.569:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(58754)]CRS-5818:Aborted command 'check' for resource 'ora.asm'. Details at (:CRSAGF00113:) {0:15:1653} in /opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log.
2016-08-09 13:18:27.476:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(58754)]CRS-5019:All OCR locations are on ASM disk groups [], and none of these disk groups are mounted. Details are at "(:CLSN00100:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log".<===All OCR disks were offlined.
2016-08-09 13:18:27.496:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(58754)]CRS-5019:All OCR locations are on ASM disk groups [VOTEDISK], and none of these disk groups are mounted. Details are at "(:CLSN00100:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log".
2016-08-09 13:18:30.590:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(58754)]CRS-5019:All OCR locations are on ASM disk groups [VOTEDISK], and none of these disk groups are mounted. Details are at "(:CLSN00100:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log".
2016-08-09 13:25:03.140:
[cssd(61077)]CRS-1662:Member kill requested by node **db1 for member number 2, group ocr_**db-cluster
2016-08-09 13:25:03.163:
[ohasd(58104)]CRS-2765:Resource 'ora.crsd' has failed on server '**db2'.
2016-08-09 13:26:27.761:
[cssd(61077)]CRS-1608:This node was evicted by node 1, **db1; details at (:CSSNM00005:) in /opt/rac/11.2.0/grid/log/**db2/cssd/ocssd.log.
2016-08-09 13:26:27.761:
[cssd(61077)]CRS-1656:The CSS daemon is terminating due to a fatal error; Details at (:CSSSC00012:) in /opt/rac/11.2.0/grid/log/**db2/cssd/ocssd.log
2016-08-09 13:26:27.762:
[cssd(61077)]CRS-1652:Starting clean up of CRSD resources.
2016-08-09 13:26:27.892:
[cssd(61077)]CRS-1608:This node was evicted by node 1, **db1; details at (:CSSNM00005:) in /opt/rac/11.2.0/grid/log/**db2/cssd/ocssd.log

<<<asm alert
Tue Aug 09 12:37:35 2016
IPC Send timeout detected. Sender: ospid 95611 [oracle@**db2 (PING)]
Receiver: inst 1 binc 454362809 ospid 216469
Tue Aug 09 12:41:39 2016
Restarting dead background process DIA0
Tue Aug 09 12:41:39 2016
DIA0 started with pid=8, OS id=172270
Tue Aug 09 12:44:02 2016
IPC Send timeout detected. Sender: ospid 95611 [oracle@**db2 (PING)]
Receiver: inst 1 binc 454362809 ospid 216469
Tue Aug 09 12:50:04 2016
IPC Send timeout detected. Sender: ospid 95611 [oracle@**db2 (PING)]
Receiver: inst 1 binc 454362809 ospid 216469
Tue Aug 09 12:52:04 2016
Restarting dead background process DIA0
Tue Aug 09 12:52:04 2016
DIA0 started with pid=8, OS id=209238
Tue Aug 09 12:56:02 2016
IPC Send timeout detected. Sender: ospid 95611 [oracle@**db2 (PING)]
Receiver: inst 1 binc 454362809 ospid 216469
Tue Aug 09 13:02:22 2016
IPC Send timeout detected. Sender: ospid 95611 [oracle@**db2 (PING)]
Receiver: inst 1 binc 454362809 ospid 216469
Tue Aug 09 13:02:28 2016
Restarting dead background process DIA0
Tue Aug 09 13:02:28 2016
DIA0 started with pid=8, OS id=230220
Tue Aug 09 13:06:09 2016
NOTE: ASM client **rac2:**rac disconnected unexpectedly.
NOTE: check client alert log.
NOTE: Trace records dumped in trace file /opt/rac/grid/diag/asm/+asm/+ASM2/trace/+ASM2_ora_169679.trc
Tue Aug 09 13:06:46 2016
NOTE: ASMB process exiting, either shutdown is in progress
Tue Aug 09 13:06:46 2016
NOTE: client exited [113293]
NOTE: force a map free for map id 2
Tue Aug 09 13:06:46 2016
PMON (ospid: 95586): terminating the instance due to error 481<===
Instance terminated by PMON, pid = 95586
...
Tue Aug 09 13:12:58 2016
IPC Send timeout detected. Sender: ospid 139823 [oracle@**db2 (LMD0)]
Receiver: inst 1 binc 454362816 ospid 216475

IPC Send timeout to 1.0 inc 12 for msg type 65521 from opid 10
Tue Aug 09 13:13:00 2016
Communications reconfiguration: instance_number 1
Tue Aug 09 13:13:00 2016
Dumping diagnostic data in directory=[cdmp_20160809131300], requested by (instance=1, osid=216473 (LMON)), summary=[abnormal instance termination].
Reconfiguration started (old inc 12, new inc 16)
List of instances:
 2 (myinst: 2)
 Nested reconfiguration detected.
...
Tue Aug 09 13:18:24 2016
IPC Send timeout detected. Sender: ospid 139823 [oracle@**db2 (LMD0)]
Receiver: inst 1 binc 464337747 ospid 100573
IPC Send timeout to 1.0 inc 18 for msg type 53 from opid 10
Tue Aug 09 13:18:26 2016
Communications reconfiguration: instance_number 1
Evicting instance 1 from cluster
Waiting for instances to leave: 1
Tue Aug 09 13:18:27 2016
Dumping diagnostic data in directory=[cdmp_20160809131827], requested by (instance=1, osid=100575 (LMS0)), summary=[abnormal instance termination].
Reconfiguration started (old inc 18, new inc 22)
List of instances:
 2 (myinst: 2)
 Nested reconfiguration detected.
 Global Resource Directory frozen
 Communication channels reestablished
 Master broadcasted resource hash value bitmaps
 Non-local Process blocks cleaned out
Tue Aug 09 13:18:27 2016
 LMS 0: 0 GCS shadows cancelled, 0 closed, 0 Xw survived
 Set master node info
 Submitted all remote-enqueue requests
 Dwn-cvts replayed, VALBLKs dubious
 All grantable enqueues granted
 Post SMON to start 1st pass IR
 Submitted all GCS remote-cache requests
 Post SMON to start 1st pass IR
Tue Aug 09 13:18:27 2016
NOTE: GMON heartbeating for grp 1
 Fix write in gcs resources
Reconfiguration complete

这次情况并不顺利，因为客户在上次处理过程中怀疑私网交换机也有问题，从而自行将私网互联从

交换机改成了直连。但 Oracle 并不支持私网交换机使用直连的情况，这点在官方文档和 RAC FAQ note 上都有说明：

1、官方文档

https://docs.oracle.com/cd/B28359_01/rac.111/b28254/admcon.htm

The interconnect network is a private network that connects all of the servers in the cluster. The interconnect network uses a switch (or multiple switches) that only the nodes in the cluster can access. Configure User Datagram Protocol (UDP) on a Gigabit Ethernet for your cluster interconnect. On Linux and Unix systems, you can configure Oracle Clusterware to use either the UDP or Reliable Data Socket (RDS) protocols. Windows clusters use the TCP protocol. Crossover cables are not supported for use with Oracle Clusterware interconnects.

2、RAC: Frequently Asked Questions [ID 220970.1]中的描述

Is crossover cable supported as an interconnect with RAC on any platform ?

NO. CROSS OVER CABLES ARE NOT SUPPORTED. The requirement is to use a switch:

Detailed Reasons:

1) cross-cabling limits the expansion of RAC to two nodes

2) cross-cabling is unstable:

a) Some NIC cards do not work properly with it. They are not able to negotiate the DTE/DCE clocking, and will thus not function. These NICS were made cheaper by assuming that the switch was going to have the clock. Unfortunately there is no way to know which NICs do not have that clock.

b) Media sense behaviour on various OS's (most notably Windows) will bring a NIC down when a cable is disconnected. Either of these issues can lead to cluster instability and lead to ORA-29740 errors (node evictions).

Due to the benefits and stability provided by a switch, and their afforability ($200 for a simple 16 port GigE switch), and the expense and time related to dealing with issues when one does not exist, this is the only supported configuration.

From a purely technology point of view Oracle does not care if the customer uses cross over cable or router or switches to deliver a message. However, we know from experience that a lot of adapters misbehave when used in a crossover configuration and cause a lot of problems for RAC. Hence we have stated on certify that we do not support crossover cables to avoid false bugs and finger pointing amongst the various parties: Oracle, Hardware vendors, Os vendors etc...

在联系客户之后，客户表示可以切换回私网交换机，但认为该问题与是否直连无关，因为：

1、该问题在使用私网交换机时也频繁发生

2、改为直连之后，私网 traceroute 时延消失了，且现在的情况变为 ASM 先出问题，然后是 voting 导致的节点驱逐

之后客户提交了更换回私网交换机之后重现的 log 和 osw 监控数据：

症状：

<<<gi alert log

2016-08-12 19:38:07.680:

[/opt/rac/11.2.0/grid/bin/oraagent.bin(58737)]CRS-5011:Check of resource "+ASM" failed: details at "(:CLSN00006:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log"

2016-08-12 19:38:13.054:

[/opt/rac/11.2.0/grid/bin/oraagent.bin(58737)]CRS-5011:Check of resource "+ASM" failed: details at "(:CLSN00006:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/ohasd/oraagent_grid/oraagent_grid.log"

2016-08-12 19:38:41.950:
[crsd(113932)]CRS-1012:The OCR service started on node **db2.
2016-08-12 19:38:42.046:
[evmd(80121)]CRS-1401:EVMD started on node **db2.
2016-08-12 19:38:43.486:
[crsd(113932)]CRS-1201:CRSD started on node **db2.
2016-08-12 19:38:44.476:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(127772)]CRS-5011:Check of resource "**db" failed: details at "(:CLSN00007:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/crsd/oraagent_oracle/oraagent_oracle.log"
2016-08-12 19:38:44.476:
[/opt/rac/11.2.0/grid/bin/oraagent.bin(127772)]CRS-5011:Check of resource "**rac" failed: details at "(:CLSN00007:)" in "/opt/rac/11.2.0/grid/log/**db2/agent/crsd/oraagent_oracle/oraagent_oracle.log"

我对比了下 IPC send timeout detected 五分钟（IPC send timeout 阈值是五分钟）前 IP 层 vmstat 信息的变化，比较奇怪的是这次从 UDP 和 traceroute 上均看不出问题，traceroute 时延也不大。于是我进一步对比了下 IPC send timeout detected 前五分钟（IPC 超时阈值为 300s）的 IP 层的指标变化：

1、IP packet reassembles failed 变化

$ egrep "zzz|packet reassembles failed" **cpdb2_netstat_16.08.12.1800.dat
zzz ***Fri Aug 12 18:01:17 CST 2016
    3529 packet reassembles failed
zzz ***Fri Aug 12 18:01:32 CST 2016
    3529 packet reassembles failed
zzz ***Fri Aug 12 18:01:48 CST 2016
    3529 packet reassembles failed
zzz ***Fri Aug 12 18:02:03 CST 2016
    3529 packet reassembles failed
zzz ***Fri Aug 12 18:02:18 CST 2016
    3531 packet reassembles failed《===在这个时间点，IP packet reassembles failed 增长 2
zzz ***Fri Aug 12 18:02:33 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:02:48 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:03:03 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:03:19 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:03:34 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:03:49 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:04:04 CST 2016
    3531 packet reassembles failed
zzz ***Fri Aug 12 18:04:19 CST 2016
    3531 packet reassembles failed

zzz ***Fri Aug 12 18:04:34 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:04:49 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:05:04 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:05:20 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:05:35 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:05:50 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:06:05 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:06:20 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:06:35 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:06:50 CST 2016《===在这个时间点，IP packet reassembles failed 增长 12
3543 packet reassembles failed
zzz ***Fri Aug 12 18:07:06 CST 2016
3543 packet reassembles failed
zzz ***Fri Aug 12 18:07:21 CST 2016《===在这个时间点，IP packet reassembles failed 增长 8
3551 packet reassembles failed
zzz ***Fri Aug 12 18:07:36 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:07:51 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:06 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:21 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:36 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:51 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:07 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:22 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:37 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:52 CST 2016
3551 packet reassembles failed

zzz ***Fri Aug 12 18:10:07 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:22 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:37 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:52 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:11:08 CST 2016
3552 packet reassembles failed
zzz ***Fri Aug 12 18:11:23 CST 2016
3552 packet reassembles failed
zzz ***Fri Aug 12 18:11:38 CST 2016
3553 packet reassembles failed
zzz ***Fri Aug 12 18:11:53 CST 2016
3557 packet reassembles failed
zzz ***Fri Aug 12 18:12:08 CST 2016
3557 packet reassembles failed
zzz ***Fri Aug 12 18:12:23 CST 2016《====在这个时间点，IP packet reassembles failed 增长 15
3568 packet reassembles failed
zzz ***Fri Aug 12 18:12:38 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:12:53 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:13:09 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:13:24 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:13:39 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:13:54 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:14:09 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:14:24 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:14:39 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:14:54 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:15:09 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:15:25 CST 2016
3568 packet reassembles failed

zzz ***Fri Aug 12 18:15:40 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:15:55 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:16:10 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:16:25 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:16:40 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:16:55 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:17:10 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:17:26 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:17:41 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:17:56 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:18:11 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:18:26 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:18:41 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:18:56 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:30:57 CST 2016《===节点重启后，IP packet reassembles failed 消失
zzz ***Fri Aug 12 18:31:12 CST 2016
zzz ***Fri Aug 12 18:31:27 CST 2016
zzz ***Fri Aug 12 18:31:42 CST 2016
zzz ***Fri Aug 12 18:31:57 CST 2016

2、IP packet reassembles failed 的变化：

zzz ***Fri Aug 12 18:06:35 CST 2016
3531 packet reassembles failed
zzz ***Fri Aug 12 18:06:50 CST 2016《===在这个时间点，IP packet reassembles failed 增长 12
3543 packet reassembles failed
zzz ***Fri Aug 12 18:07:06 CST 2016
3543 packet reassembles failed
zzz ***Fri Aug 12 18:07:21 CST 2016《===在这个时间点，IP packet reassembles failed 增长 8
3551 packet reassembles failed

zzz ***Fri Aug 12 18:07:36 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:07:51 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:06 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:21 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:36 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:08:51 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:07 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:22 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:37 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:09:52 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:07 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:22 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:37 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:10:52 CST 2016
3551 packet reassembles failed
zzz ***Fri Aug 12 18:11:08 CST 2016
3552 packet reassembles failed
zzz ***Fri Aug 12 18:11:23 CST 2016
3552 packet reassembles failed
zzz ***Fri Aug 12 18:11:38 CST 2016
3553 packet reassembles failed
zzz ***Fri Aug 12 18:11:53 CST 2016
3557 packet reassembles failed
zzz ***Fri Aug 12 18:12:08 CST 2016
3557 packet reassembles failed
zzz ***Fri Aug 12 18:12:23 CST 2016《====在这个时间点，IP packet reassembles failed 增长 15，五分钟后就报了 IPC send timeout detected。
3568 packet reassembles failed
zzz ***Fri Aug 12 18:12:38 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:12:53 CST 2016

3568 packet reassembles failed
zzz ***Fri Aug 12 18:13:09 CST 2016
3568 packet reassembles failed
zzz ***Fri Aug 12 18:13:24 CST 2016
3568 packet reassembles failed

由于客户不便向 redhat support 寻求支持去检查网卡的 IP 封装问题，进一步查找与 IP packet reassembles failed 相关的 node eviction，发现如下：

RHEL 6.6: IPC Send timeout/node eviction etc with high packet reassembles failure (Doc ID 2008933.1)

从文档上看，该问题在 RHEL6.6 和 OEL6.6 上可能发生，且在 OEL6.6 上开过 bug：

BUG 21036841

但由于此问题可以通过调整如下 kernel 参数解决，所以被 OEL 方面认为是 Not a bug：

net.ipv4.ipfrag_high_thresh = 16M

net.ipv4.ipfrag_low_thresh = 15M

另一种解决办法是使用 jumbo frame 巨帧传输，但这需要交换机支持。如想配置巨帧传输，可参考：

Recommendation for the Real Application Cluster Interconnect and Jumbo Frames (Doc ID 341788.1)

三、总结：

这个案例我认为无论是客户还是我自身，都有值得反思的地方，对于最终操作的主角（客户）而言，尝试解决问题时应尽量避免一次做多项变更，特别是不应该尝试 Oracle 并不支持的私网直连。

而对于我自身，在最开始就怀疑过调整 UDP buffer 并不能真正解决这个问题，因为节点重启后的 ASM 实例反复重启不像是短期的私网传输延迟行为。这个案例提醒我们查看网络层传输问题时，不能只看 udp 的指标，udp 正常而 IP 封包不正常仍然可能导致 IPC Send timeout。