

排除12000系列路由器上的“QM_SANITY_WARNING”消息故障

目录

[简介](#)

[先决条件](#)

[要求](#)

[使用的组件](#)

[规则](#)

[问题](#)

[解决方案](#)

[背景](#)

[情形 1：](#)

[方案 2：](#)

[情形 3：](#)

[故障排除命令](#)

[示例：](#)

简介

本文档介绍调试数据包缓冲区耗尽的过程可能发生的消息运行IOS的12000系列思科路由器中的不同线卡。这太常见了 由于缺乏GSR缓冲区管理知识，因此在更换实际正常运行的硬件方面浪费了宝贵的时间和资源。

先决条件

要求

读者应概述Cisco 12000系[列路由器架构](#)。

使用的组件

本文档中的信息基于以下软件和硬件版本：

- Cisco 12000 系列互联网路由器
- 支持千兆位交换机路由器的Cisco IOS®软件版本

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备都以已清除（默认）配置启动。如果您使用的是真实网络，请确保您已经了解所有命令的潜在影响。

规则

有关文档规则的详细信息，请参阅 Cisco 技术提示规则。

问题

GSR或12000系列思科路由器具有真正的分布式架构。这意味着每个LC都运行其自己的Cisco IOS软件映像副本，并且具有自行完成数据包转发决策的智能。每张线卡都有自己的功能

1. 转发查找
2. 数据包缓冲区管理
3. QOS
4. 流量控制

在GSR中，分组交换期间最重要的操作之一是缓冲区管理，由位于线卡中的各种缓冲区管理ASIC(BMA)执行。以下是一些与GSR缓冲区管理相关的消息，这些消息可能在生产过程中显示在路由器日志中。在以下部分中，我们将讨论可能导致这些消息出现在路由器日志中的不同触发器，以及要采取哪些纠正措施来缓解问题。在某些情况下，这还可能导致丢包，可能表现为协议摆动并导致网络影响。

```
%EE48-3-QM_SANITY_WARNING:ToFab FreeQ缓冲区耗尽
```

```
插槽1:9月16日19:06:40.003 UTC:%EE48-3-QM_SANITY_WARNING:ToFab FreeQ池#2中可用的可用缓冲区(1)很少
```

```
插槽8:9月16日19:06:45.943 UTC:%EE48-3-QM_SANITY_WARNING:ToFab FreeQ池#1中可用的可用缓冲区(0)很少
```

```
插槽0:9月16日19:06:46.267 UTC:%EE48-3-QM_SANITY_WARNING:ToFab FreeQ池#2中提供的可用缓冲区(2)很少
```

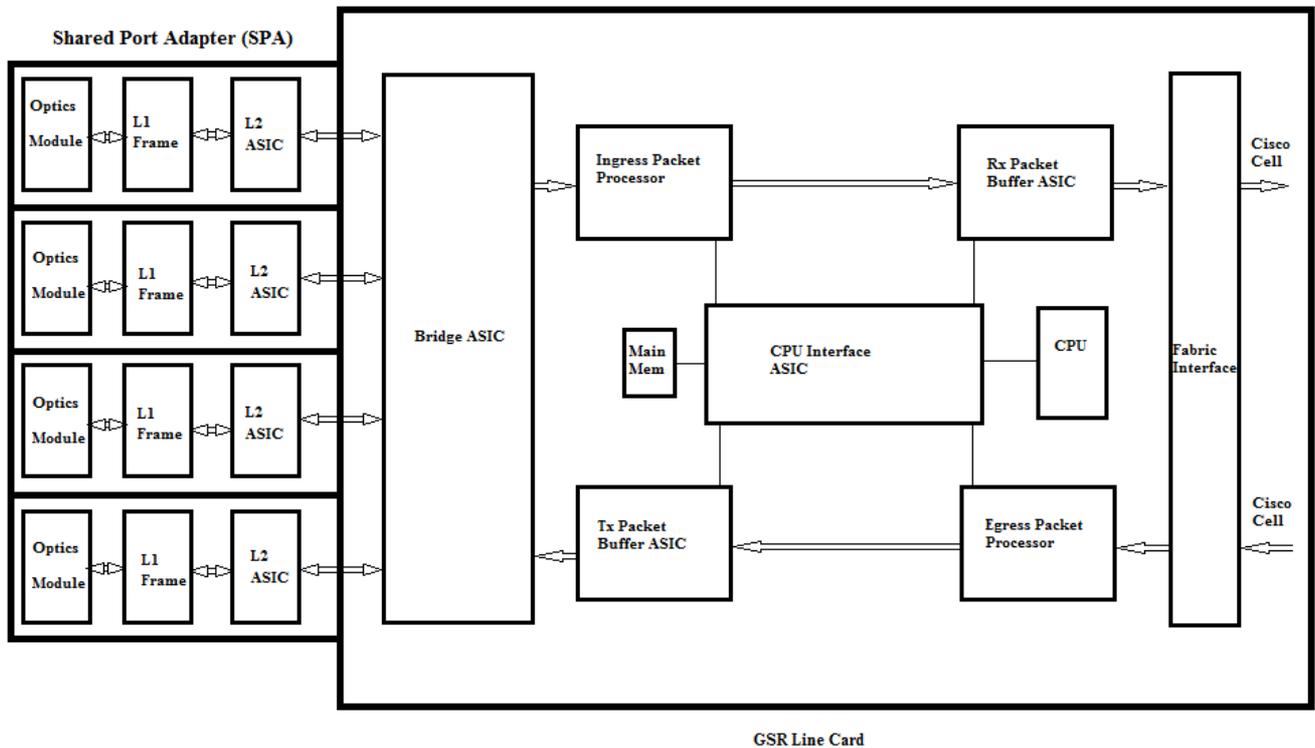
```
插槽8:9月16日19:06:47.455 UTC:%EE48-3-QM_SANITY_WARNING:ToFab FreeQ缓冲区耗尽。恢复ToFab缓冲区
```

```
插槽8:9月16日19:06:47.471 UTC:%EE192-3-BM_QUIESCE:
```

解决方案

背景

要排除QM-SANITY警告错误故障，我们需要了解GSR线卡上的数据包流。下图说明 C12k线卡的主块和数据包流路径。



Cisco 12000系列互联网路由器上的线卡(LC)有两种类型的内存：

- 路由或处理器内存（动态RAM - DRAM）：此内存主要使板载处理器能够运行Cisco IOS软件并存储网络路由表（转发信息库 — FIB、邻接）
- 数据包内存（同步动态RAM - SDRAM）：线卡分组存储器暂时存储等待线卡处理器的交换决定的数据分组。

从上图看，GSR线卡具有专用的数据包缓冲ASIC（专用集成电路），这些ASIC也称为缓冲区管理ASIC(BMA)，在线卡上执行数据包缓冲和缓冲区队列管理功能。为了支持高吞吐量/转发速率，将任一方向上的数据包内存划分到不同大小的内存池中，这些内存池旨在转发不同MTU大小的数据包。

物理层接口模块(PLIM)接收的帧对卡进行第2层处理，并将其DMA传输到PLIM卡中的本地存储器。一旦所接收的数据单元完成，PLIM中的ASIC就与入口BMA联系，并请求适当大小的缓冲区。如果授予缓冲区，则数据包将移至线卡入口数据包内存。如果没有可用缓冲区，则数据包将被丢弃，而忽略的接口计数器将启动。入口分组处理器对分组执行功能处理，做出转发决策，并将分组移至与出口线卡对应的toFab队列。交换矩阵接口ASIC(FIA)将数据包分段到思科信元，并将信元传输到交换矩阵。然后，出口线卡上的FIA从交换机交换矩阵接收数据包，并进入FrFab队列，在该队列中重组数据包，然后再进入出口PLIM，最后在线路上发送数据包。

FrFab BMA从特定缓冲区池中选择缓冲区的决定取决于入口线卡交换引擎的决定。由于整个设备上的所有队列大小和顺序相同，因此交换引擎会通知传输LC将数据包放入其进入路由器的相同编号队列中。

在交换数据包时，入口线卡上用于移动数据包的特定缓冲池的队列大小将减1，直到出口线卡中的BMA返回缓冲区。在此，我们还应注意到，完整的缓冲区管理是由缓冲区管理ASIC在硬件中完成的，为了实现无缺陷操作，BMA必须将缓冲区从源地返回到原始池。

有三种情况下，GSR数据包缓冲区管理可能会遇到压力或故障，导致数据包丢失。以下是三个场景。

情形 1：

硬件队列管理失败。当出口BMA无法返回数据包缓冲区或将数据包缓冲区返回到错误的缓冲池时,会发生此情况。如果缓冲区返回到不正确的池,我们将看到一些缓冲池在增长,而一些缓冲池在一段时间内耗尽,并最终影响具有耗尽缓冲池大小的数据包。我们开始将QM-Sanity警告视为数据包缓冲区耗尽并超过警告阈值。

使用QM健全性调试和 `show controllers tofab queues` 命令,检查您是否受此条件影响。请参阅故障排除部分,了解如何启用QM健全性阈值。

此情况通常由硬件故障引起。检查路由器上的以下输出并查找奇偶校验错误或线卡崩溃。解决方法是更换线路卡。

```
show controllers fia
```

```
show context all
```

```
show log
```

示例:

从QM健全性调试和`show controller tofab queue`可以看到Pool 2正在增大,而Pool 4正在运行低速。这表示池4正在丢失缓冲区,并且它正返回到池2。

QM健全性调试:

```
插槽5:10月25日04:41:03.286 UTC:池 1:雕刻尺寸102001:当前大小73078
```

```
插槽5:10月25日04:41:03.286 UTC:池 2:切割大小78462:当前大小181569
```

```
插槽5:10月25日04:41:03.286 UTC:池 3:雕刻尺寸57539:当前大小6160
```

```
插槽5:10月25日04:41:03.286 UTC:池 4:雕刻尺寸22870:当前大小67
```

```
插槽5:10月25日04:41:03.286 UTC:IPC FreeQ:切割尺寸600:当前大小600
```

show controllers tofab queues:

```
<snip>
```

```
Qnum正尾#Qelem LenThresh
```

```
— —
```

4个非IPC空闲队列:

```
102001/102001 (指定/划分缓冲区),39.1%,80字节数据大小
```

```
1 13542 13448 73078 262143
```

```
78462/78462 (指定/划分的缓冲区),30.0%,608字节数据大小
```

```
2 131784 131833 181569 262143
```

```
57539/57539 (指定/划分的缓冲区),22.0%,1616字节数据大小
```

3 184620 182591 6160 262143

23538/22870 (指定/划分的缓冲区) ,8.74%,4592字节数据大小

4 239113 238805 67 262143

<snip>

方案 2 :

下一跳设备或转发路径上的流量拥塞。在此场景中，GSR向其发送流量的设备无法以GSR的速度处理，因此，下一跳设备向GSR发送暂停帧，要求其降速。如果在GSR PLIM卡上启用流量控制，路由器将执行暂停帧并开始缓冲数据包。最终，路由器将耗尽缓冲区，导致QM健全性错误消息和数据包丢弃。随着数据包缓冲区耗尽并超过警告阈值，我们将开始看到QM-Sanity警告。有关如何查找QM健全性阈值的故障排除部分。

使用出口接口上的**show interface**输出检查路由器是否受此场景的影响。下面的捕获提供了接口接收暂停帧的示例。该行动计划将研究下一跳设备拥塞的原因。

GigabitEthernet6/2为up，线路协议为up

小型可插拔光纤正常

硬件为GigMac 4端口千兆以太网，地址为000b.455d.ee02(bia 000b.455d.ee02)

描述:思科悉尼实验室

Internet address is 219.158.33.86/30

MTU 1500 bytes, BW 500000 Kbit, DLY 10 usec, rely 255/255, load 154/255

Encapsulation ARPA, loopback not set

Keepalive set (10 sec)

全双工，1000Mbps，链路类型为强制打开，介质类型为LX

输出流控制打开，输入流控制打开

ARP type:ARPA, ARP Timeout 04:00:00

Last input 00:00:02, output 00:00:02, output hang never

上次清除“show interface”计数器7w1d

排队策略随机早期检测(WRED)

输出队列0/40、22713601丢弃；输入队列0/75, 736369丢弃

Available Bandwidth 224992 kilobits/sec

30秒输入速率309068000位/秒，49414个数据包/秒

30秒输出速率3034000000位/秒、73826个数据包/秒

143009959974 packets input, 88976134206186 bytes, 0 no buffer

Received 7352 broadcasts, 0 runts, 0 giants, 0 throttles

0 input errors, 0 CRC, 0 frame, 0 overrun, 0 ignored

0监视器，7352组播，45个暂停输入

234821393504 packets output, 119276570730993 bytes, 0 underruns

传输的73201广播

0 output errors, 0 collisions, 0 interface resets

0 babbles, 0 late collision, 0 deferred

0丢失载波，0无载波，0暂停输出

0 output buffer failures, 0 output buffers swapped out

情形 3：

在因网络设计不佳/流量突发/DOS攻击而超订用时。如果持续存在高流量情况，则可能会发生QM健全性警告其中更多流量被定向到路由器，而线卡可以处理。

要引起此问题，请检查路由器中所有接口上的流量速率。这将揭示出是否有高速链路阻塞了慢速链路。

使用“show interface output”命令。

故障排除命令

- 检查LC的当前QM健全性级别 连接到LC转到启用模式运行**test fab**命令收集“qm_sanity_info”的输出选项q退出**test fab**命令行退出LC
- 配置QM健全性参数 更改配置模式运行**hw-module slot <slot#> qm-sanity tofab warning freq <>**
- 启用/禁用QM健全性调试 连接到LC转到启用模式**Run test fab**命令运行“qm_sanity_debug”。再次运行，它将停止调试选项q以**扩展fab**命令行退出LC
- 要检查GSR交换矩阵接口ASIC统计信息，请执行以下操作：**show controller fia**
- 检查Tofab队列 **show controllers tofab queues**
- 要检查Frfab查询 **show controller frfab queues**

示例：

以下输出来自运行中的实验路由器以演示命令输出。

GSR-1-PE-5#show controller fia

交换矩阵配置：10Gbps带宽（2.4Gbps可用），冗余交换矩阵

主计划程序：插槽17备份调度程序：插槽 16

Fab Epoch no 0停止计数0

从交换矩阵FIA错误

—

redund overflow 0信元丢弃0

单元奇偶0

交换机卡提供0x001F 插槽16 17 18 19 20

监控的交换机卡 0x001F 插槽16 17 18 19 20

插槽： 16 17 18 19 20

名称： csc0 csc1 sfc0 sfc1 sfc2

— — — — —

los 0 0 0 0 0

状态关闭关闭关闭

crc16 0 0 0 0 0 0

到交换矩阵FIA错误

—

sca不按0请求错误0 uni fifo overflow 0

授予奇偶校验0多请求0 uni fifo undrflow 0

cntrl parity 0 uni req 0

multi fifo 0 空dst req 0 握手错误0

单元奇偶校验0

GSR-1-PE-5#attach 1

进入插槽中模块化SPA接口卡的控制台：1

键入“exit”结束此会话

Press RETURN to get started!

LC-Slot1>en

LC-Slot1#test fab

BFLC诊断控制台程序

BFLC(?)[?]:qm_sanity_debug

已启用QM健全性调试

BFLC(?帮助)[qm_sanity_debug]:

插槽1:02:54:33: ToFAB BMA信息

插槽1:02:54:33:FreeQs雕刻数4

插槽1:02:54:33:池 1:雕刻尺寸102001:当前大小102001

插槽1:02:54:33:池 2:切割尺寸78462:当前大小78462

插槽1:02:54:33:池 3:雕刻尺寸57539:当前大小57539

插槽1:02:54:33:池 4:雕刻尺寸22870:当前大小22870

插槽1:02:54:33:IPC FreeQ:切割尺寸600:当前大小600

插槽1:02:54:33:启用的LOQ数768

插槽1:02:54:33:禁用的LOQ数1280

插槽1:02:54:33: ToFAB BMA信息

插槽1:02:54:33:FreeQs雕刻数4

插槽1:02:54:33:池 1:雕刻尺寸102001:当前大小102001

插槽1:02:54:33:池 2:切割尺寸78462:当前大小78462

插槽1:02:54:33:池 3:雕刻尺寸57539:当前大小57539

插槽1:02:54:33:池 4:雕刻尺寸22870:当前大小22870

插槽1:02:54:33:IPC FreeQ:切割尺寸600:当前大小600

插槽1:02:54:33:启用的LOQ数768

插槽1:02:54:33:禁用的LOQ数1280

禁用QM健全性调试

BFLC(?帮助)[qm_sanity_debug]:qm_sanity_info

ToFab QM健全性级别警告

FrFab QM健全性级别无

每20秒触发一次健全性检查

最小值.缓冲区阈值百分比5

BFLC(?)[qm_sanity_info]:q

LC-Slot1#exi

从插槽1断开连接。

连接持续时间 : 00:01:09

GSR-1-PE-5#config t

输入配置命令，每行一条。以 CNTL/Z 结束。

GSR-1-PE-5(config)#hw-module slot 1 qm-sanity tofab warning freq 10

GSR-1-PE-5(config)#end

GSR-1-PE-5#attach 1

02:57:25:%SYS-5-CONFIG_I:Configured from console by console

GSR-1-PE-5#attach 1

进入插槽中模块化SPA接口卡的控制台 : 1

键入“exit”结束此会话

Press RETURN to get started!

LC-Slot1>en

LC-Slot1#test fab

BFLC诊断控制台程序

BFLC(?)[?]:qm_sanity_info

ToFab QM健全性级别警告

FrFab QM健全性级别无

每10秒触发一次健全性检查

最小值.缓冲区阈值百分比5

BFLC(?)[qm_sanity_info]:q

LC-Slot1#exit

从插槽1断开连接。

连接持续时间 : 00:00:27

GSR-1-PE-5#execute-on all show controllers tofab queues

=====线卡 (插槽0) =====

为ToFab缓冲区创建信息

SDRAM大小 : 268435456 bytes, address:E0000000 , 刻制基础 : E0018000

268337152字节刻划大小、4个SDRAM库、16384字节SDRAM页面大小、2个刻划

最大缓冲区数据大小4592字节 , 最小缓冲区数据大小80字节

262141/262141指定/划分的缓冲区

265028848/265028848字节总和指定/划分的缓冲区大小

Qnum正尾#Qelem LenThresh

— —

4个非IPC空闲队列 :

107232/107232 (指定/划分缓冲区) ,40.90%,80字节数据大小

601 107832 107232 262143

73232/73232 (指定/划分的缓冲区) ,27.93%,608字节数据大小

107833 181064 73232 262143

57539/57539 (指定/划分的缓冲区) ,21.94%,1616字节数据大小

181065 238603 57539 262143

23538/23538 (指定/划分的缓冲区) ,8.97%,4592字节数据大小

238604 262141 23538 262143

IPC队列 :

600/600 (指定/划分的缓冲区) ,0.22%,4112字节数据大小

155 154 600 262143

原始队列 (高优先级) :

0 0 0 65535

原始队列 (中等优先级) :

0 0 0 32767

原始队列 (低优先级) :

0 0 0 16383

ToFab队列 :

目标插槽队列#头尾长度阈值

pkts pkts

=====

0 0 0 0 262143

15 2191(hpr)0 0 0 0

组播2048 0 0 0 262143

2049 0 0 0 262143

=====线卡 (插槽1) =====

为ToFab缓冲区创建信息

SDRAM大小 : 268435456 bytes, address:26000000, carve base:26010000

268369920字节刻花大小、4个SDRAM内存、32768字节SDRAM页面大小、2个刻花

最大缓冲区数据大小4592字节 , 最小缓冲区数据大小80字节

262140/261472指定/划分的缓冲区

267790176/264701344字节总和指定/划分的缓冲区大小

Qnum正尾#Qelem LenThresh

— —

4个非IPC空闲队列 :

102001/102001 (指定/划分缓冲区) ,39.1%,80字节数据大小

1 601 102601 102001 262143

78462/78462 (指定/划分的缓冲区) ,30.0%,608字节数据大小

2 102602 181063 78462 262143

57539/57539 (指定/划分的缓冲区) ,22.0%,1616字节数据大小

3 181064 238602 57539 262143

23538/22870 (指定/划分的缓冲区) ,8.74%,4592字节数据大小

4 238603 261472 22870 262143

IPC队列 :

600/600 (指定/划分的缓冲区) ,0.22%,4112字节数据大小

30 85 84 600 262143

原始队列 (高优先级) :

27 0 0 0 65368

原始队列 (中等优先级) :

28 0 0 0 32684

原始队列 (低优先级) :

31 0 0 0 16342

ToFab队列 :

目标插槽队列#头尾长度阈值

pkts pkts

=====

.....

高优先级

0 2176(hpr)0 0 0

1 2177(hpr)0 0 0

2 2178(hpr)0 0 0

3 2179(hpr)0 0 0

4 2180(hpr)553 552 0

5 2181(hpr)0 0 0

6 2182(hpr)0 0 0

7 2183(hpr)0 0 0

8 2184(hpr)0 0 0

9 2185(hpr)0 0 0

10 2186(hpr)0 0 0

11 2187(hpr)0 0 0

12 2188(hpr)0 0 0

13 2189(hpr)0 0 0

14 2190(hpr)0 0 0

15 2191(hpr)0 0 0

组播

2048 0 0 0

2049 0 0 0

2050 0 0 0

2051 0 0 0

2052 0 0 0

2053 0 0 0

2054 0 0 0

2055 0 0 0

=====线卡 (插槽3) =====

为ToFab缓冲区创建信息

SDRAM大小 : 268435456 bytes, address:E0000000 , 刻制基础 : E0018000

268337152字节刻划大小、4个SDRAM库、16384字节SDRAM页面大小、2个刻划

最大缓冲区数据大小4112字节 , 最小缓冲区数据大小80字节

262142/262142指定/划分的缓冲区

230886224/230886224字节总和指定/划分的缓冲区大小

Qnum正尾#Qelem LenThresh

— —

3个非IPC空闲队列 :

94155/94155 (指定/划分的缓冲区) ,35.91%,80字节数据大小

601 94755 94155 262143

57539/57539 (指定/划分的缓冲区) ,21.94%,608字节数据大小

94756 152294 57539 262143

109848/109848 (指定/划分缓冲区) ,41.90%,1616字节数据大小

152295 262142 109848 262143

IPC队列 :

600/600 (指定/划分的缓冲区) ,0.22%,4112字节数据大小

207 206 600 262143

原始队列 (高优先级) :

0 0 0 65535

原始队列 (中等优先级) :

0 0 0 32767

原始队列 (低优先级) :

0 0 0 16383

ToFab队列 :

目标插槽队列#头尾长度阈值

pkts pkts

=====

0 0 0 0 262143

1 0 0 0 262143

2 0 0 0 262143

3 0 0 0 262143

.....

2049 0 0 0 262143

2050 0 0 0 262143

2051 0 0 0 262143

2052 0 0 0 262143

2053 0 0 0 262143

2054 0 0 0 262143

2055 0 0 0 262143

GSR-1-PE-5#execute-on slot 2 show controller frfab queues

=====线卡 (插槽2) =====

为FrFab缓冲区创建信息

SDRAM大小 : 268435456 bytes, address:D0000000 , 刻花底座 : D241D100

230567680字节刻花大小、4个SDRAM内存、16384字节SDRAM页面大小、2个刻花

最大缓冲区数据大小4592字节 , 最小缓冲区数据大小80字节

235926/235926指定/划分的缓冲区

226853664/226853664字节总和指定/划分的缓冲区大小

Qnum正尾#Qelem LenThresh

— —

4个非IPC空闲队列 :

96484/96484 (指定/划分的缓冲区) ,40.89%,80字节数据大小

11598 11597 96484 262143

77658/77658 (指定/划分的缓冲区) ,32.91%,608字节数据大小

103116 103115 77658 262143

40005/40005 (指定/划分缓冲区) ,16.95%,1616字节数据大小

178588 178587 40005 262143

21179/21179 (指定/划分缓冲区) ,8.97%,4592字节数据大小

214748 235926 21179 262143

IPC队列 :

600/600 (指定/划分的缓冲区) ,0.25%,4112字节数据大小

66 65 600 262143

组播原始队列 :

0 0 0 58981

组播复制空闲队列 :

235930 262143 26214 262143

原始队列 (高优先级) :

78 77 0 235927

原始队列 (中等优先级) :

11596 11595 0 58981

原始队列 (低优先级) :

0 0 0 23592

接口队列 :

接口队列#头尾长度阈值

pkts pkts

=====

0 0 103107 103106 0 32768

3 178588 178587 0 32768

1 4 103110 103109 0 32768

7 11586 11585 0 32768

28 0 0 0 32768

11 0 0 0 32768

312 0 0 0 32768

15 0 0 0 32768

GSR-1-PE-5#