# Nexus 9500-R, Nexus 3000-R: Solucionar problemas de descarte de entrada

#### **Contents**

Introduction

Informações de Apoio

Gerenciador de Tráfego de Entrada (ITM)

Buffer de VOQ de entrada

Controle de fluxo e programação

Causas comuns

Hardware aplicável

Solucionar problemas de descarte de entrada

Cenário comum - Fluxo de tráfego de 10G a 1G - Quedas constantes:

Etapa 1. Verifique qual fila está impactada na interface Input Discards.

Etapa 2. Verifique a representação gráfica da Broadcom dos contadores usados para diagnósticos:

Etapa 3. Descubra a que ASIC e que Porta Jericho sua Porta do Painel Frontal com Descartes de Entrada pertencem:

Etapa 4. Entenda qual conector de VOQ e VOQ sua porta de entrada tem.

Etapa 5. Verificar da perspectiva do BCM, qual fila especificamente não está vazia; Ou seja, congestionado.

Etapa 6. Encontre sua porta congestionada de saída do valor da fila não vazio:

Passo 7. Verifique qual porta do painel frontal está em ASIC 1 e Mapas para Jericho Port 9 com base em sua descoberta anterior.

Comandos adicionais

Testes adicionais de laboratório:

Etapa 1. Descartes de entrada com várias interfaces congestionadas de saída.

Etapa 2. Input Discards devido ao SPAN.

Etapa 3. Descarte de entrada devido ao pino do par de tráfego.

Etapa 4. Enviar pacotes com um IP de destino desconhecido.

Etapa 5. A entrada é descartada enquanto uma porta de tronco/acesso passa para o estado de encaminhamento STP

Etapa 6. Input Discards devido ao Eth1/9 exceder a taxa de linha.

## Introduction

Este documento descreve as causas e as soluções de descarte de entrada para o Cisco Nexus 9500-R EoR e o Nexus 3000-R ToR. Um descarte de entrada indica o número de pacotes descartados na fila de entrada devido ao congestionamento. Esse número inclui quedas causadas por queda traseira e Detecção Antecipada Aleatória Ponderada (WRED - Weighted Random Early Detection).

Se você tiver quedas aleatórias/esporádicas/históricas (ou seja, não ocorrendo mais), entre em contato com o TAC da Cisco para obter mais informações. Essa passagem é útil quando as

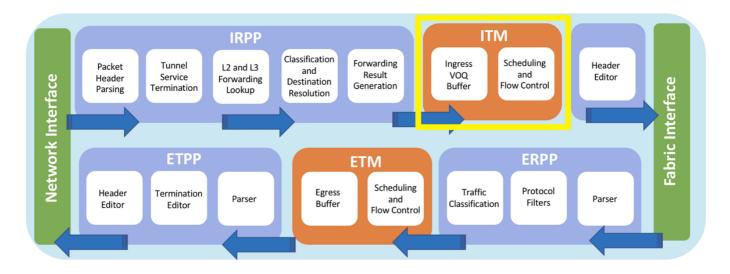
devoluções de entrada são incrementadas com frequência.

## Informações de Apoio

O R-Series usa a arquitetura de ingresso VOQ. A arquitetura VOQ emula filas de saída no buffer de entrada com filas virtuais. Cada porta de saída tem oito filas para tráfego unicast e oito filas para tráfego multicast. O tráfego pode ser classificado em classes de tráfego com base no valor de Classe de Serviço (CoS - Class of Service) ou Ponto de Código de Serviços Diferenciados (DSCP - Differentiated Services Code Point) nos pacotes e enfileirado na fila virtual correspondente para essa classe de tráfego.

O R-Series usa um mecanismo de crédito distribuído para transferir o tráfego pela estrutura. Antes que um pacote seja agendado para sair do VOQ, o agendador de buffer de entrada solicita um crédito para a porta específica e a prioridade no buffer de saída. O crédito é solicitado a um agendador de crédito de entrada para a porta de destino e a prioridade. Se houver espaço de buffer disponível, o agendador de saída concederá acesso e enviará a concessão de crédito ao agendador de buffer de entrada. Se nenhum espaço de buffer estiver disponível no buffer de saída, a agenda de saída não concederá um crédito e o tráfego será colocado em buffer no VOQ até que o próximo crédito esteja disponível.

Abaixo está o Pipeline de Encaminhamento de Pacotes para a plataforma -R. Neste artigo, você se concentra no componente **Ingress Traffic Manager**. Mais detalhes sobre a arquitetura neste **link** 



## Gerenciador de Tráfego de Entrada (ITM)

O gerenciador de tráfego de entrada (ITM) é um bloco no pipeline de entrada. Ele executa etapas relacionadas ao tráfego de fila no VOQ, agenda o tráfego para transmissão pela estrutura e gerencia créditos.

#### Buffer de VOQ de entrada

O bloco de buffer de VOQ de entrada gerencia o buffer no chip e o buffer de pacote fora do chip. Ambos os buffers usam arquitetura VOQ e o tráfego é enfileirado com base nas informações do IRPP (processador de pacote de receptor de entrada). Um total de 96.000 VOQs estão disponíveis para tráfego unicast e multicast.

## Controle de fluxo e programação

Antes de um pacote ser transmitido do pipeline de entrada, o pacote precisa ser programado para transferência pela estrutura. O agendador de ingresso envia uma solicitação de crédito ao agendador de saída localizado no bloco gerenciador de tráfego de saída. Quando o gerenciador de tráfego de entrada recebe o crédito, ele começa a enviar tráfego para o processador de pacote de transmissão de entrada. Se o buffer de saída estiver cheio, o tráfego será colocado em buffer na fila dedicada representada pela porta de saída e pela classe de tráfego.

## Causas comuns

Geralmente, as devoluções de entrada podem ser vistas pelos seguintes motivos em vários hardwares Nexus

- Fluxos de tráfego congestionando interfaces de saída (como entrada de 10G e saída de 1G)
- Porta de destino de SPAN com excesso de assinaturas Aplica-se a tipos de hardware específicos.

## Hardware aplicável

PID

N9K-X9636C-R

N9K-X9636Q-R

N9K-X9636C-RX

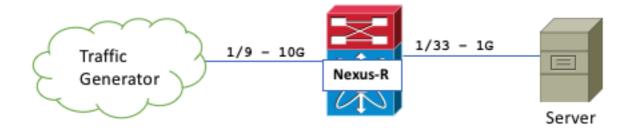
N9K-X96136YC-R

N3K-C36180YC-R

N3K-C3636C-R

## Solucionar problemas de descarte de entrada

Cenário comum - Fluxo de tráfego de 10G a 1G - Quedas constantes:



Em todo este artigo, o valor do contador de "descartes de entrada" e qualquer contador interno de HW que faça referência ao mesmo será alterado à medida que os erros forem aumentando enquanto os testes e os comandos relevantes devem ser capturados ao vivo.

Etapa 1. Verifique qual fila está impactada na interface Input Discards.

Esta etapa será útil mais tarde.

No nosso caso, é Fila 7, a fila padrão - Há 8 filas no total na entrada:

Etapa 2. Verifique a representação gráfica da Broadcom dos contadores usados para diagnósticos:

```
Nexus-R# bcm-shell mod 1 "diag counters g"
/ | \
                                        JERICHO NETWORK INTE
RFACE
                                       \ | /
                                                                                NBI
RX_TOTAL_BYTE_COUNTER
                                            = 10,616,663,796
TX_TOTAL_BYTE_COUNTER
                                          = 41,136
                                            = 10,659,301
RX_TOTAL_PKT_COUNTER
TX_TOTAL_PKT_COUNTER
                                          = 606
RX_TOTAL_DROPPED_EOPS
EPNI
                                         CPU_PACKET_COUNTER
                                             = 606
                                             = 10,659,302
NIF_PACKET_COUNTER
EPE_BYTES_COUNTER
                                          = 41,136
OAMP_PACKET_COUNTER
                                             = 0
EPE_PKT_COUNTER
                                          = 606
OLP_PACKET_COUNTER
EPE_DSCRD_PKT_CNT
                                          = 0
RCY_PACKET_COUNTER
                                             = 0
| IRE_FDT_INTRFACE_CNT
                                       TDR
EGQ
MMU_IDR_PACKET_COUNTER
                                             = 10,659,302
FQP_PACKET_COUNTER
                                          = 606
| IDR_OCB_INTERFACE_COUNTER
```

```
PQP_UNICAST_PKT_CNT
                                          = 606
PQP_DSCRD_UC_PKT_CNT
                                          = 0
PQP_UC_BYTES_CNT
                                          = 48,408
PQP_MC_PKT_CNT
                                       MQI
PQP_DSCRD_MC_PKT_CNT
                                          = 0
PQP_MC_BYTES_CNT
                                          = 0
ENQUEUE_PKT_CNT
                                             = 1,403,078
EHP_UNICAST_PKT_CNT
                                          = 606
DEQUEUE_PKT_CNT
                                             = 1,403,078
EHP_MC_HIGH_PKT_CNT
                                          = 0
DELETED_PKT_CNT
EHP_MC_LOW_PKT_CNT
                                          = 0
ENQ_DISCARDED_PACKET_COUNTER
                                            = 9,256,829
DELETED_PKT_CNT
   Rejects: PORT_AND_PG_STATUS
RQP_PKT_CNT
                                          = 606
RQP_DSCRD_PKT_CNT
                                          = 0
PRP_PKT_DSCRD_TDM_CNT
                                          = 0
PRP_SOP_DSCRD_UC_CNT
                                          = 0
PRP_SOP_DSCRD_MC_CNT
                                          = 0
PRP_SOP_DSCRD_TDM_CNT
                                          = 0
EHP_MC_HIGH_DSCRD_CNT
                                          = 0
EHP_MC_LOW_DSCRD_CNT
ERPP_LAG_PRUNING_DSCRD_CNT
                                          = 0
ERPP_PMF_DISCARDS_CNT
                                          = 0
ERPP_VLAN_MBR_DSCRD_CNT
                                          = 0
FDA
CELLS_IN_CNT_P1 = 0
                                     CELLS_OUT_CNT_P1 = 0
CELLS_IN_CNT_P2 = 0
                                     CELLS_OUT_CNT_P2 = 0
                 = 0
                                     CELLS_OUT_CNT_P3 = 0
CELLS_IN_CNT_P3
                                       IPT
CELLS_IN_TDM_CNT
                    = 0
                                     CELLS_OUT_TDM_CNT
                                                             = 0
CELLS_IN_MESHMC_CNT
                                     CELLS_OUT_MESHMC_CNT
                     = 0
                                                              = 0
EGQ_PKT_CNT
                                                          = 606
CELLS_IN_IPT_CNT
                     = 606
                                      CELLS_OUT_IPT_CNT
ENQ_PKT_CNT
                                             = 1,403,084
EGQ_DROP_CNT
                                          = 0
FDT_PKT_CNT
                                             = 1,402,472
EGQ_MESHMC_DROP_CNT
                                          = 0
CRC_ERROR_CNT
                                             = 0
```

```
EGO TDM OVF DROP CNT
                                           = 606 *
 CFG_EVENT_CNT
                                            = 48,408
  CFG_BYTE_CNT
FDR
                                            = 5,609,892
| IPT_DESC_CELL_COUNTER
                                         = 0
P1_CELL_IN_CNT
IRE_DESC_CELL_COUNTER
                                           = 0
P2_CELL_IN_CNT
                                         = 0
P3_CELL_IN_CNT
TRANSMITTED DATA CELLS COUNTER
                                           = 5,609,892
CELL_IN_CNT_TOTAL
/ \
                                           JERICHO FABRIC INTER
F A C E
                                      \ | /
```

Um QUEUE\_DELETED\_PACKET\_COUNTER maior que zero indica que os pacotes foram EXCLUÍDOS pelo IQM (Ingress Queueing Manager) após a fila. Isso se deve ao fato de uma fila ativa não receber nenhum crédito que sugeriria uma configuração incorreta do esquema de agendamento. Você deve verificar isso através do bcm-shell mod X "getReg IQM\_QUEUE\_DELETED\_PACKET\_COUNTER"

ENQ\_DISCARDED\_PACKET\_COUNTER significa que os pacotes foram descartados ANTES de enfileirar. Você também pode ver este contador definido no BCM (o comando é limpo na leitura):

```
Nexus-R# bcm-shell mod 1 "g iqm_reject_status_bmp" | i i PG|IQM0|IQM1
IQM_REJECT_STATUS_BMP.IQM0[0x1a7]=0x20000000: <VSQF_WRED_STATUS=0,
QNUM_OVF_STATUS=0,PORT_AND_PG_STATUS=1,OCCUPIED_BD_STATUS=0,
IQM_REJECT_STATUS_BMP.IQM1[0x1a7]=0: <VSQF_WRED_STATUS=0,VSQF_MX_SZ_STATUS=0,
PORT_AND_PG_STATUS=0,OCCUPIED_BD_STATUS=0,MULTICAST_ERROR_STATUS=0,
```

Você pode sempre observar isso rapidamente com o **show hardware internal errors module X** (comando limpa na leitura):

Etapa 3. Descubra a que ASIC e que Porta Jericho sua Porta do Painel Frontal com Descartes de Entrada pertencem:

```
Nexus-R# show interface hardware-mappings | i i Eth1/9|--|Name|Eth1/33
HName - Hardware port name. None means N/A

Name Ifindex Smod Unit HPort HName FPort NPort VPort SrcId

Eth1/9 1a001000 0 0 9 xe9 255 8 -1 0 << ASIC 0, Jericho Port 9
Eth1/33 1a004000 2 1 9 xe9 32 -1 0 << ASIC 1, Jericho Port 9
```

Exibindo Eth1/33 para este exemplo. Em uma rede real, você ainda não conhecerá a porta de saída congestionada.

Etapa 4. Entenda qual conector de VOQ e VOQ sua porta de entrada tem.

Esse comando mostra detalhes de nós para o fluxo de entrada de VoQ para uma porta específica. Além disso, ele nos mostra o saldo de crédito atual do VoQ.

O VOQ da porta é derivado desta maneira:

LCs são baseados em 0 - Módulo 1 é 0, Módulo 2 é 1, etc

Há 256 IDs de porta de sistema por LC

```
ID = (LC * ID da porta do sistema) + número FP
Eth1/9 = (0 * 256) + 9 = 9
ID VOQ = 32 + (ID da porta do sistema * 8)
Eth1/9 = 32 + (9 * 8) = 104
```

Nosso VOQ para Eth1/9 será, portanto, 104 que corresponde à saída previamente coletada

```
module-1# show hardware internal jer-usd ingress-vsq buffer-occupancy front-port 9
             VSQF BUFFER OCCUPANCY
| Front port 9
|max global shared |
                                         157286
|max ocb buffer occupancy |
                                          0 |
+----
              COSQ 0
                                          4 |
3280 |
rate class
granted buffers per port
shared buffers occupied
                                          127792 <<<<
granted buffers occupied
                                           3280
|shared buffer max occupancy |
                                           127792 | <<<<
```

Etapa 5. Verificar da perspectiva do BCM, qual fila especificamente não está vazia; Ou seja, congestionado.

Etapa 6. Encontre sua porta congestionada de saída do valor da fila não vazio:

Se a Fila for 303, lembre-se de que essas filas são na verdade um intervalo para que possam ser 303 + 7 ou 303-7 - A questão é: qual porta tem um VOQ que corresponda em um intervalo de 296-303 ou, alternativamente, 303-310?

Sabe-se que a Fila 7 em Eth1/9 está congestionada, então 303 é realmente a mais alta em seu intervalo, de modo que a faixa de 296-303 é uma estimativa bem instruída.

```
module-1# show hardware internal jer-usd info voq asic 1
|Unit|JerPort| Voq| VoqConn| SE| HR|CreditBal|
 1 1 232 56 81957 8 3ffff
           72 81989
     2 240
                    16 3fffff
 24 3ffff
                    32 3ffff
     5 264 120 82085 40 3ffff
                    48 3ffff
                    56 3ffff
     8 288 168 82181 64 3ffff
                    72 3a5 <<< 296 +7 would give us 303
 <snip>
```

Exiba o mesmo para o ásico 0 - Não mostrado aqui para brevidade; você observaria na coluna Voq que seu intervalo de interesse não está nesse ASIC

Observe algumas coisas na saída acima:

- Nossa porta congestionada de saída está no ASIC 1.
- Nossa porta congestionada de saída tem um VOQ de 296 e 303 equivaleria à Fila 7 nessa porta.
- Observe a coluna Saldo de crédito Há muito poucos créditos restantes nesta interface a serem concedidos e é por isso que nosso Eth1/9 de entrada inicia o buffer.

Passo 7. Verifique qual porta do painel frontal está em ASIC 1 e Mapas para Jericho Port 9 com base em sua descoberta anterior.

Neste ponto, você encontrou a porta congestionada de saída - Determine se há algo que esteja se interrompendo incorretamente na rede, você configurou o SPAN e sua porta de destino é 1G enquanto está fornecendo uma ou mais interfaces 10G ou se esse é um problema de gargalo/projeto.

## Comandos adicionais

Eles são mais avançados - não são necessários para encontrar a porta do Congresso de Saída em cenários normais.

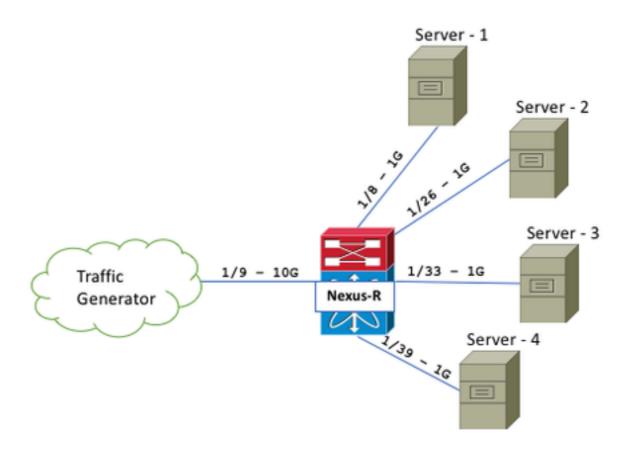
```
attach module X
show hardware internal jer-usd tm_debug asic <slot> module <module>
show hardware internal jer-usd info voq [ asic <instance> ] [ port <port> ] [ ]
show hardware internal jer-usd info non-empty voq asic [ <instance> ] [ ]
show hardware internal jer-usd info voq-profile { QueueThreshold drop_p <dp> | OCBThreshold } [
```

```
asic <instance> ] [ port<port> ] [ ]
show hardware internal jer-usd info voq-connector front-port <port> [ ]
show hardware internal jer-usd stats vsq { front-port <port> | inband asic <slot> | recycle-port
<port> asic <slot> }
show hardware internal jer-usd ingress-vsq buffer-occupancy front-port <port>
show hardware internal jer-usd info IQM { counter | rate } asic <instance> dst-port <port> [
interval <int> ] [ ]
show hardware internal jer-usd info SCH { counter | rate } asic <instance> dst-port <port> [
interval <int> ] [ ]
bcm-shell mod X
diag cosq print_flow_and_up dest_id=<flow_id>
diag cosq voq id=<voqid> detailed=1
diag cosq qpair e2e ps=<id>
cosq conn ing
cosq conn egr
dump IPS_CR_BAL_TABLE <voqID>
getReg IQM_QUEUE_MAXIMUM_OCCUPANCY_QUEUE_SIZE
```

## Testes adicionais de laboratório:

#### Etapa 1. Descartes de entrada com várias interfaces congestionadas de saída.

Considere esta topologia em que o Gerador de Tráfego está enviando 2G de tráfego para cada Servidor:



Verifique rapidamente quais filas não estão vazias - Observe que há 4:

```
Nexus-R# bcm-shell mod 1 "diag cosq non_empty_queue"

Core 0:
Ingress VOQs Sizes (format: [queue_id(queue_size)]):
[103(29475840B)] [247(29379584B)] [303(56452096B)] [351(76020736B)]
```

Determine a que interfaces essas filas pertencem - Verifique primeiro o ASIC 0 (ele demonstra apenas com uma interface):

Repita o mesmo processo para os outros três valores da Fila: 247, 303 e 351.

#### Etapa 2. Input Discards devido ao SPAN.

Definindo Eth1/33 como uma porta de destino de SPAN ao definir Eth1/9 como uma porta de origem de SPAN na direção RX

```
Nexus-R# show run mon

monitor session 1

description SPAN TEST INPUT DISCARDS
source interface Ethernet1/9 rx
destination interface Ethernet1/33
no shut

Nexus-R# show int e1/9 | i i input.disc
0 input with dribble 9314306 input discard
```

Etapa 3. Descarte de entrada devido ao pino do par de tráfego.

Envio de pacotes com SRC 10.10.10.10 e DEST 192.168.10.10 em que Eth1/9 está em 10.10.10.1/24 - Isso não resulta em um Input Discard; no entanto, você vê este contador:

```
Nexus-R# bcm-shell mod 1 "diag counters g"
| / |\
| JERICHONETWORKINTERFACE |
```



Etapa 4. Enviar pacotes com um IP de destino desconhecido.

Envie pacotes com SRC 10.10.10.10 e DEST 192.168.10.10 em que Eth1/9 está em 10.10.10.1/24 e Eth1/33 é uma porta L3 na sub-rede 172.16.0.1/30 - Sem contador de queda, sem descartes de entrada mesmo quando o destino é **desconhecido**.

# Etapa 5. A entrada é descartada enquanto uma porta de tronco/acesso passa para o estado de encaminhamento STP

Enviar pacotes em que Eth1/9 é apenas um tronco amplo (ou porta de acesso) - é registrado como um Descarte de entrada enquanto a porta passa para um estado de encaminhamento STP.

```
Nexus-R(config)# int e1/9
Nexus-R(config-if)# switchport mode trunk
Nexus-R# bcm-shell mod 1 "diag counters g" | i i -- | IQM | ENQ_DISCARD | Rejects
PQP_MC_PKT_CNT
                                              = 1,678,949
                                           = 11,369,033
PQP_DSCRD_MC_PKT_CNT
ENQ_DISCARDED_PACKET_COUNTER
                                               = 1,289,182
                                             = 11,369,081
DELETED_PKT_CNT
   Rejects: QUEUE_NOT_VALID_STATUS
Discards: SRC_EQUAL_DEST
Nexus-R# show span int e1/9
Vlan Role Sts Cost Prio.Nbr Type
VLAN0001 Desg BLK 2 128.9 P2p
VLAN0010 Desg BLK 2 128.9 P2p
```

QUEUE\_NOT\_VALID\_STATUS é uma queda devido à decisão do Packet Processor (PP) de descartar ou um destino inválido recebido dos blocos do Packet Processor (PP).

## Etapa 6. Input Discards devido ao Eth1/9 exceder a taxa de linha.

Envia 10G+ para Eth1/9 resultaria em um tipo diferente de queda, já que você está maximizando Eth1/9 no primeiro lugar - Ainda conta como um Input Discard:

```
/ | \
                                                     JERICHO NETWORK INTE
RFACE
                                         \ | /
                                                                                   NBT
                                              = 53,913,106,009
RX_TOTAL_BYTE_COUNTER
                                           = 1,164,231
{\tt TX\_TOTAL\_BYTE\_COUNTER}
RX_TOTAL_PKT_COUNTER
                                              = 54,145,395
TX_TOTAL_PKT_COUNTER
                                            = 17,029
RX_TOTAL_DROPPED_EOPS
                                         TRE
EPNI
CPU_PACKET_COUNTER
                                              = 17,010
NIF_PACKET_COUNTER
                                              = 54,145,476
EPE_BYTES_COUNTER
                                            = 5,721,307
OAMP_PACKET_COUNTER
                                              = 0
EPE_PKT_COUNTER
                                            = 50,703
OLP_PACKET_COUNTER
                                              = 0
EPE_DSCRD_PKT_CNT
                                            = 0
RCY_PACKET_COUNTER
                                              = 16,837
IRE_FDT_INTRFACE_CNT
                                               = 0
                                         IDR
EGQ
                                          | MMU_IDR_PACKET_COUNTER
                                              = 54,128,577
                                           = 50,703
FQP_PACKET_COUNTER
                                              = 0
| IDR_OCB_INTERFACE_COUNTER
                                            = 50,683
PQP_UNICAST_PKT_CNT
PQP_DSCRD_UC_PKT_CNT
                                            = 0
PQP_UC_BYTES_CNT
                                           = 5,216,716
                                           = 20
PQP_MC_PKT_CNT
                                         IQM
PQP_DSCRD_MC_PKT_CNT
                                           = 20
PQP_MC_BYTES_CNT
                                           = 2,079
ENQUEUE_PKT_CNT
                                              = 5,463,323
EHP_UNICAST_PKT_CNT
                                            = 50,683
DEQUEUE_PKT_CNT
                                              = 5,594,400
                                            = 20
EHP_MC_HIGH_PKT_CNT
DELETED_PKT_CNT
EHP_MC_LOW_PKT_CNT
                                              = 48,716,055
ENQ_DISCARDED_PACKET_COUNTER
DELETED_PKT_CNT
                                            = 40
  Rejects: VOQ_MX_QSZ_STATUS
<snip>
```