

Nexus 9500-R, Nexus 3000-R : Risolvere i problemi relativi ai rigetti di input

Sommario

[Introduzione](#)

[Premesse](#)

[ITM \(Ingress Traffic Manager\)](#)

[Buffer VOQ in ingresso](#)

[Programmazione e controllo del flusso](#)

[Cause comuni](#)

[Dispositivi interessati](#)

[Risolvere i problemi relativi ai rigetti di input](#)

[Scenario comune - Flusso del traffico da 10G a 1G - Cadute costanti:](#)

[Passaggio 1. Verificare quale coda è interessata dall'interfaccia per gli scarti di input.](#)

[Passaggio 2. Controllare la rappresentazione grafica dei contatori utilizzati per la diagnostica di Broadcom:](#)

[Passaggio 3. Individuare l'ASIC e la porta Gerico a cui appartiene la porta del pannello anteriore con errori di input:](#)

[Passaggio 4. Informazioni sul connettore VOQ e VOQ della porta in ingresso.](#)

[Passaggio 5. Verificare dalla prospettiva di BCM quale coda specifica non è vuota; Sono congestionato.](#)

[Passaggio 6. Individuare la porta congestionata in uscita dal valore della coda non vuoto:](#)

[Passaggio 7. Verificare quale porta del pannello anteriore si trova in ASIC 1 e mappare alla porta Jericho 9 in base ai risultati precedenti.](#)

[Comandi aggiuntivi](#)

[Test di laboratorio aggiuntivi:](#)

[Passaggio 1. Input non elaborato con più interfacce congestionate in uscita.](#)

[Passaggio 2. Input ignorato a causa di SPAN.](#)

[Passaggio 3. Input scartato a causa di un pin per capelli traffico.](#)

[Passaggio 4. Inviare i pacchetti con un indirizzo IP di destinazione sconosciuto.](#)

[Passaggio 5. Input ignorato durante la transizione di una porta di accesso/trunk nello stato di inoltro STP](#)

[Passaggio 6. Input ignorato a causa di un superamento di Eth1/9 della velocità della linea.](#)

Introduzione

In questo documento vengono descritte le cause e le soluzioni dei rigetti in input per Cisco Nexus 9500-R EoR e Nexus 3000-R ToR. Un rifiuto in input indica il numero di pacchetti scartati nella coda di input a causa di congestione. Questo numero include le cadute causate da tail drop e Weighted Random Early Detection (WRED).

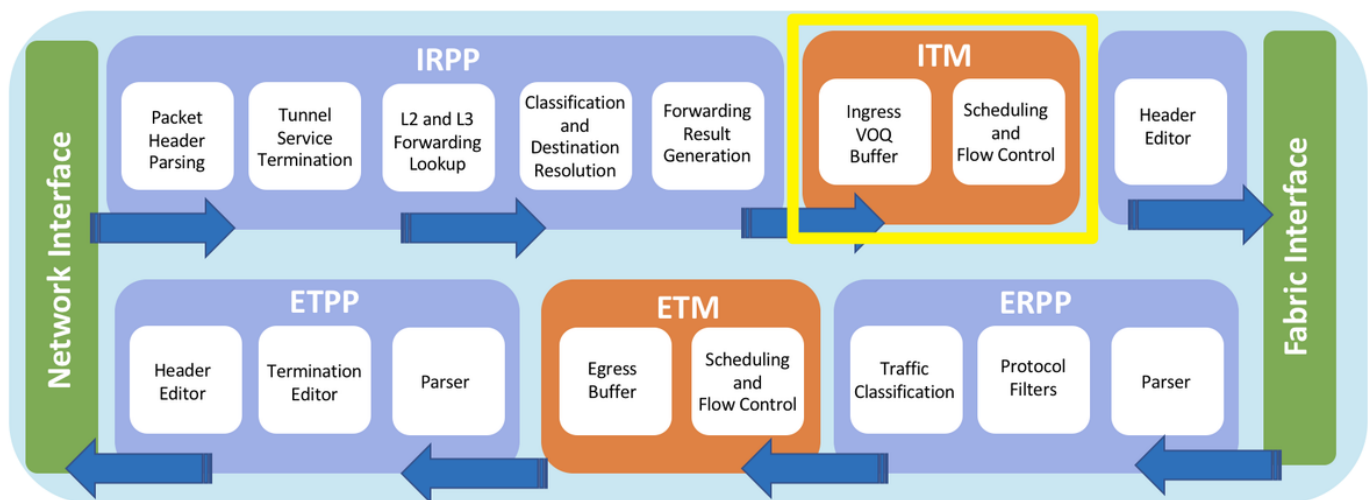
Se si verificano cali casuali/sporadici/storici (che non si verificano più), contattare Cisco TAC per ulteriori indagini. Questa procedura dettagliata è utile quando gli input scartati vengono incrementati di frequente.

Premesse

La serie R utilizza l'architettura Ingress VOQ. L'architettura VOQ emula le code in uscita nel buffer in entrata con le code virtuali. Ogni porta di uscita ha otto code per il traffico unicast e otto code per il traffico multicast. Il traffico può essere classificato in classi di traffico basate sul valore CoS (Class-of-Service) o DSCP (Differentiated Services Code Point) nei pacchetti e quindi accodato nella coda virtuale corrispondente per quella classe di traffico.

La serie R utilizza un meccanismo di credito distribuito per trasferire il traffico sulla struttura. Prima che un pacchetto lasci il VOQ, lo scheduler del buffer in entrata richiede un credito per la porta e la priorità specifiche nel buffer in uscita. Il credito viene richiesto a uno scheduler crediti in entrata per la porta e la priorità di destinazione. Se è disponibile spazio di buffer, lo scheduler di uscita concede l'accesso e invia la concessione di credito allo scheduler di buffer in entrata. Se nel buffer in uscita non è disponibile spazio di buffer, la programmazione in uscita non concede un credito e il traffico viene memorizzato nel VOQ finché non sarà disponibile il credito successivo.

Di seguito è riportata la pipeline di inoltro pacchetti per la piattaforma -R. In questo articolo viene esaminato il componente **Gestione traffico in ingresso**. Ulteriori dettagli sull'architettura in questo [link](#)



ITM (Ingress Traffic Manager)

Il gestore del traffico in entrata (ITM) è un blocco nella pipeline in entrata. Esegue le fasi relative al traffico in coda in VOQ, pianifica il traffico per la trasmissione sulla struttura e gestisce i crediti.

Buffer VOQ in ingresso

Il blocco buffer VOQ in entrata gestisce sia il buffer on-chip che il buffer del pacchetto off-chip. Entrambi i buffer utilizzano l'architettura VOQ e il traffico viene accodato in base alle informazioni provenienti dall'IRPP (Ingress Receiver Packet Processor). Sono disponibili 96.000 VOQ per il traffico unicast e multicast.

Programmazione e controllo del flusso

Prima di trasmettere un pacchetto dalla pipeline in entrata, il pacchetto deve essere pianificato per il trasferimento sulla struttura. Lo scheduler in entrata invia una richiesta di credito allo scheduler

in uscita che si trova nel blocco del gestore del traffico in uscita. Quando il gestore del traffico in entrata riceve il credito, inizia a inviare il traffico al processore di pacchetti di trasmissione in entrata. Se il buffer di uscita è pieno, il traffico verrà memorizzato nel buffer nella coda dedicata rappresentata dalla porta di uscita e dalla classe di traffico.

Cause comuni

In genere, i rifiuti di input possono essere rilevati per i motivi seguenti in vari componenti hardware Nexus

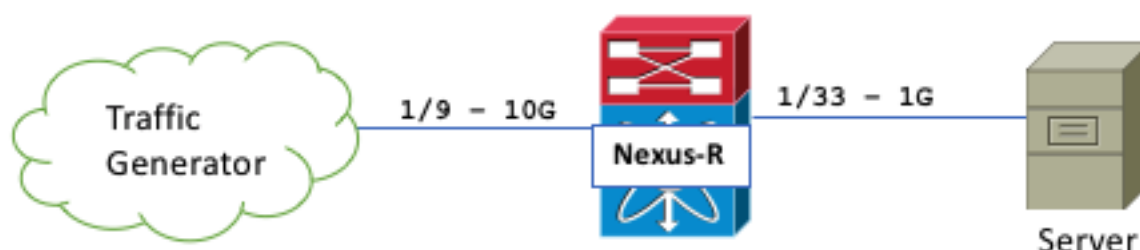
- Flussi di traffico che congestionano le interfacce in uscita (come 10G in entrata e 1G in uscita)
- Porta di destinazione SPAN con sottoscrizione eccessiva: si applica a tipi di hardware specifici.

Dispositivi interessati

PID
N9K-X9636C-R
N9K-X9636Q-R
N9K-X9636C-RX
N9K-X96136YC-R
N3K-C36180YC-R
N3K-C3636C-R

Risolvere i problemi relativi ai rigetti di input

Scenario comune - Flusso del traffico da 10G a 1G - Cadute costanti:



In questo articolo, il valore del contatore dei "rigetti di input" e di qualsiasi contatore interno HW che fa riferimento allo stesso cambierà con l'aumento degli errori durante il test e i comandi rilevanti devono essere acquisiti in tempo reale.

Passaggio 1. Verificare quale coda è interessata dall'interfaccia per gli scarti di input.

Questo passaggio si rivela utile in seguito.

Nel nostro caso, è la coda 7, la coda predefinita - Ci sono 8 code totali in entrata:

```
Nexus-R# show system internal qos queuing stats interface e1/9 | beg "QUEUE: 7"
QUEUE: 7
=====
ingress dropped packets: 113503981
ingress dropped bytes: 113503981000
enqueued packet count: 74115825
enqueued byte count: 74115825000
```

Passaggio 2. Controllare la rappresentazione grafica dei contatori utilizzati per la diagnostica di Broadcom:

```
Nexus-R# bcm-shell mod 1 "diag counters g"
```

```

/|\
|
| J E R I C H O   N E T W O R K   I N T E
R F A C E   |
|
|\|/
|
+-----+-----+-----+-----+
|
| NBI
|
| RX_TOTAL_BYTE_COUNTER           = 10,616,663,796
TX_TOTAL_BYTE_COUNTER             = 41,136
| RX_TOTAL_PKT_COUNTER           = 10,659,301
TX_TOTAL_PKT_COUNTER              = 606
| RX_TOTAL_DROPPED_EOPS          = 0
|
+-----+-----+-----+-----+
|
| IRE
EPNI
| CPU_PACKET_COUNTER             = 606
|
| NIF_PACKET_COUNTER             = 10,659,302
EPE_BYTES_COUNTER                 = 41,136
| OAMP_PACKET_COUNTER            = 0
EPE_PKT_COUNTER                   = 606
| OLP_PACKET_COUNTER             = 0
EPE_DSCRD_PKT_CNT                 = 0
| RCY_PACKET_COUNTER             = 0
|
| IRE_FDT_INTRFACE_CNT           = 0
|
+-----+-----+-----+-----+
|
| IDR
EGQ
|
| MMU_IDR_PACKET_COUNTER         = 10,659,302
FQP_PACKET_COUNTER                = 606
| IDR_OCB_INTERFACE_COUNTER      = 0
PQP_UNICAST_PKT_CNT               = 606
|
| PQP_DSCRD_UC_PKT_CNT           = 0
|
| PQP_UC_BYTES_CNT               = 48,408

```



```
Nexus-R# show hardware internal errors module 1
```

```
slot 1
```

```
=====
```

```
|-----|
| Device:Forwarding ASIC Role:MAC Mod: 1 |
| Device Statistics Category :: ERROR    |
|-----|
Instance:0
```

```
IQM
```

```
ENQ_DISCARDED_PACKET_COUNTER = 8,233,862
```

```
Rejects: PORT_AND_PG_STATUS
```

```
Instance:1
```

Passaggio 3. Individuare l'ASIC e la porta Gerico a cui appartiene la porta del pannello anteriore con errori di input:

```
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
```

Visualizzazione di Eth1/33 per questo esempio. In una rete reale, non si conosce ancora la porta di uscita congestionata.

Passaggio 4. Informazioni sul connettore VOQ e VOQ della porta in ingresso.

```
Nexus-R# attach module 1
```

```
module-1# show hardware internal jer-usd info voq asic 0 port 9
```

```
+-----+
|Unit|JerPort| Voq| VoqConn| SE   |   HR   |CreditBal|
+-----+
| 0  |    9  | 104|    176| 82213|    72   | 16a .  |
+-----+
```

Questo comando mostra i dettagli del flusso per il VoQ in entrata per una porta specifica. Inoltre, mostra il saldo attivo corrente del VoQ.

Il VOQ della porta è derivato nel modo seguente:

I moduli LC sono basati su 0: il modulo 1 è 0, il modulo 2 è 1, ecc

Sono disponibili 256 ID porte di sistema per LC

ID = (LC * ID porta di sistema) + numero FP

Eth1/9 = (0 * 256) + 9 = 9

VOQ ID = 32 + (ID porta sistema * 8)

Eth1/9 = 32 + (9 * 8) = 104

Il VOQ per Eth1/9 sarà quindi 104, che corrisponde all'output precedentemente raccolto

```
module-1# show hardware internal jer-usd ingress-vsqs buffer-occupancy front-port 9
```

VSQF BUFFER OCCUPANCY		
Front port 9		
max global shared		157286
max ocb buffer occupancy		0
COSQ 0		
rate class		4
granted buffers per port		3280
shared buffers occupied		127792 <<<<
granted buffers occupied		3280
shared buffer max occupancy		127792 <<<<

Passaggio 5. Verificare dalla prospettiva di BCM quale coda specifica non è vuota; Sono congestionato.

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

Core 0:
Ingress VOQs Sizes (format: [queue_id(queue_size)]):
[303(191338496B)] << the Queue ID belongs to your Egress CONGESTED port!

Core 1:
<empty>

Passaggio 6. Individuare la porta congestionata in uscita dal valore della coda non vuoto:

Se la coda è 303, ricordare che queste code sono in realtà un intervallo, quindi possono essere 303 + 7 o 303-7. La domanda è: quale porta ha un VOQ che corrisponde a un intervallo di 296-303 o, in alternativa, 303-310?

È noto che la Coda 7 su Eth1/9 è congestionata, quindi 303 è in realtà il più alto nel suo range, quindi l'intervallo di 296-303 è una stima ben istruita.


```

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|
| 1| 2| 240| 72| 81989| 16| 3ffff|
| 1| 3| 248| 88| 82021| 24| 3ffff|
| 1| 4| 256| 104| 82053| 32| 3ffff|
| 1| 5| 264| 120| 82085| 40| 3ffff|
| 1| 6| 272| 136| 82117| 48| 3ffff|
| 1| 7| 280| 152| 82149| 56| 3ffff|
| 1| 8| 288| 168| 82181| 64| 3ffff|
| 1| 9| 296| 184| 82213| 72| 3a5| <<< 296 +7 would give us 303
| 1| 10| 304| 200| 82245| 80| 3ffff| << It cannot be this one as 303 is not included
| 1| 11| 312| 216| 82277| 88| 3ffff|
<snip>

```

Visualizzare lo stesso per l'elemento di base 0 - Non qui per brevità; nella colonna Voq si noterà che l'intervallo di interesse non è incluso nell'ASIC

Di seguito sono riportati alcuni elementi dell'output precedente:

- La porta congestionata in uscita è su ASIC 1.
- La nostra porta congestionata in uscita ha un VOQ di 296 e 303 equivarrebbe alla coda 7 su quella porta.
- Notare la colonna Saldo credito - Ci sono pochissimi crediti rimanenti su questa interfaccia da concedere, ed è per questo che il nostro ingresso Eth1/9 inizia la memorizzazione nel buffer.

Passaggio 7. Verificare quale porta del pannello anteriore si trova in ASIC 1 e mappare alla porta Jericho 9 in base ai risultati precedenti.

```

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

```

A questo punto, è stata rilevata la porta congestionata in uscita - Determinare se si è verificato un errore di frammentazione nella rete, se è stato configurato SPAN e la porta di destinazione è 1G durante l'origine di una o più interfacce 10G o se si tratta di un problema di collo di bottiglia/progettazione.

Comandi aggiuntivi

Si tratta di funzionalità più avanzate. Non è necessario trovare una porta congestionata in uscita in scenari normali.

```

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-vsqs 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> ] [ ]

```

```

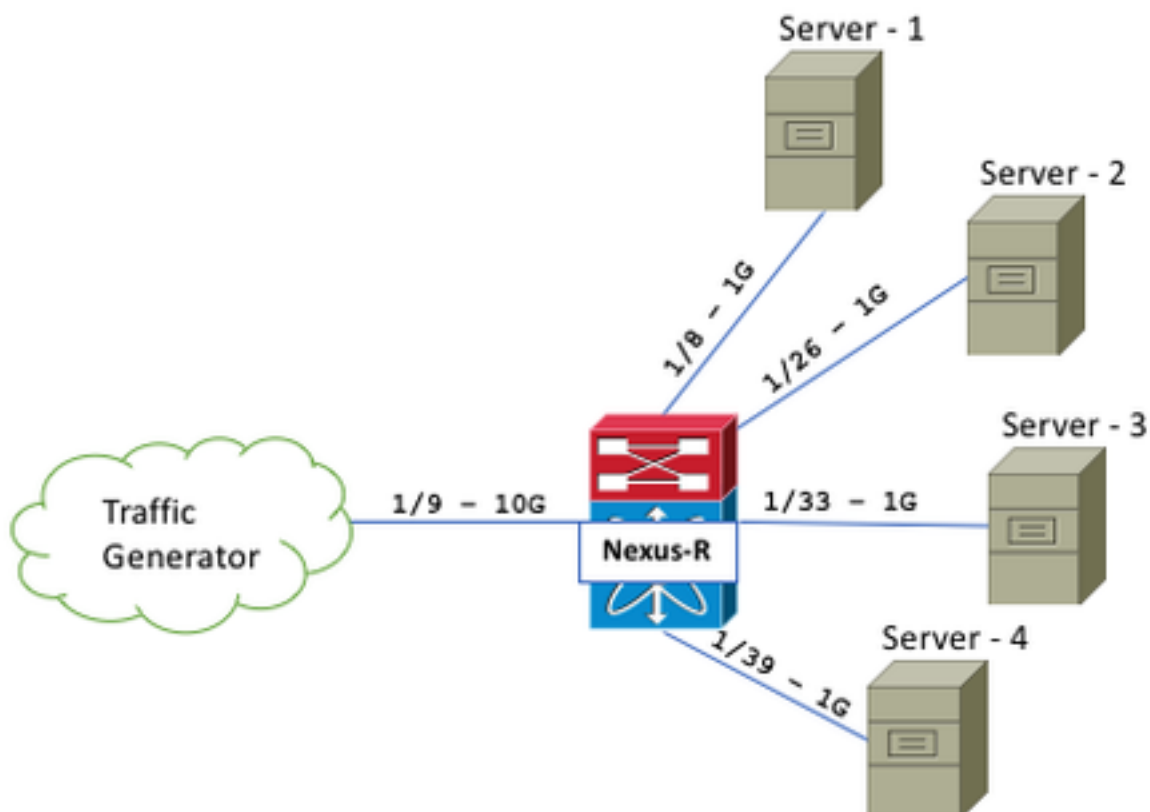
bcm-shell mod X
diag cosq print_flow_and_up dest_id=<flow_id>
diag cosq voq id=<voqid> detailed=1
diag cosq qpair e2e ps=<id>
cosq conn ing
cosq conn egr
dump IPS_CR_BAL_TABLE <voqID>
getReg IQM_QUEUE_MAXIMUM_OCCUPANCY_QUEUE_SIZE

```

Test di laboratorio aggiuntivi:

Passaggio 1. Input non elaborato con più interfacce congestionate in uscita.

Considerare questa topologia in cui il Generatore traffico sta inviando 2G di traffico verso ciascun Server:



Verificare rapidamente quali code non sono vuote. Notare che sono presenti 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)]
```

Determinare le interfacce a cui appartengono queste code. Controllare prima ASIC 0 (viene dimostrato con una sola interfaccia):

```
module-1# show hardware internal jer-usd info voq asic 0
```

```
+-----+
|Unit|JerPort| Voq| VoqConn| SE| HR|CreditBal|
+-----+
| 0| 1| 40| 48| 81957| 8| 3ffff|
| 0| 2| 48| 64| 81989| 16| 3ffff|
| 0| 3| 56| 80| 82021| 24| 3ffff|
...
| 0| 8| 96| 160| 82181| 64| 7b| << 96 + 7 = 103, this is port Eth1/8
<snip>
```

```
`show interface hardware-mappings`
```

```
-----
Name      Ifindex  Smod Unit HPort HName NPort VPort SrcId
-----
Eth1/8    1a000e00 0 0 8 xe8 7 -1 0
```

Ripetere la stessa procedura per gli altri tre valori di Coda: 247, 303 e 351.

Passaggio 2. Input ignorato a causa di SPAN.

Impostazione di Eth1/33 come porta di destinazione SPAN mentre impostazione di Eth1/9 come porta di origine SPAN in direzione 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
```

Passaggio 3. Input scartato a causa di un pin per capelli traffico.

Invio di pacchetti con SRC 10.10.10.10 e DEST 192.168.10.10 dove Eth1/9 è in 10.10.10.1/24 - Non si verifica un'eliminazione degli input; tuttavia, viene visualizzato questo contatore:

```
Nexus-R# bcm-shell mod 1 "diag counters g"
```

```
| /|\
| J E R I C H O N E T W O R K I N T E R F A C E |
```

```
\|/ |
```

```
+-----+
+-----+
_PACKET_COUNTER = 0 | DELETED_PKT_CNT = 12,027,201 |
| | Discards: INVALID_OTM SRC_EQUAL_DEST
+-----+
+-----+
```

Passaggio 4. Inviare i pacchetti con un indirizzo IP di destinazione sconosciuto.

Inviare pacchetti con SRC 10.10.10.10 e DEST 192.168.10.10 dove Eth1/9 è in 10.10.10.1/24 e Eth1/33 è una porta L3 nella subnet 172.16.0.1/30 - Nessun contatore di rilascio, nessun rifiuto di input anche quando la destinazione è sconosciuta.

Passaggio 5. Input ignorato durante la transizione di una porta di accesso/trunk nello stato di inoltro STP

Invia pacchetti dove Eth1/9 è solo un wide trunk (o porta di accesso) - È registrato come Input Discard mentre la porta passa a uno stato di inoltro 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 |
| IQM |
PQP_DSCRD_MC_PKT_CNT = 11,369,033 |
| ENQ_DISCARDED_PACKET_COUNTER = 1,289,182 |
DELETED_PKT_CNT = 11,369,081 |
| 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
<snip>
```

QUEUE_NOT_VALID_STATUS è una perdita di dati causata dalla decisione del processore pacchetti (PP) di eliminare o da una destinazione non valida ricevuta dai blocchi del processore pacchetti (PP).

Passaggio 6. Input ignorato a causa di un superamento di Eth1/9 della velocità della linea.

Invia 10G+ in Eth1/9 causerebbe un tipo diverso di caduta mentre si sta massimizzando Eth1/9 in prima posizione - Conta ancora come Input Discard:

bcm-shell.0> **diag counters g**

```

/|\
R F A C E |
|
+-----+
|
|
| RX_TOTAL_BYTE_COUNTER = 53,913,106,009 |
TX_TOTAL_BYTE_COUNTER = 1,164,231 |
| RX_TOTAL_PKT_COUNTER = 54,145,395 |
TX_TOTAL_PKT_COUNTER = 17,029 |
| RX_TOTAL_DROPPED_EOPS = 0 |
|
+-----+
|
| IRE
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 |
FQP_PACKET_COUNTER = 50,703 |
| IDR_OCB_INTERFACE_COUNTER = 0 |
PQP_UNICAST_PKT_CNT = 50,683 |
|
PQP_DSCRD_UC_PKT_CNT = 0 |
|
PQP_UC_BYTES_CNT = 5,216,716 |
+-----+
PQP_MC_PKT_CNT = 20 |
|
| 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 |
EHP_MC_HIGH_PKT_CNT = 20 |
| DELETED_PKT_CNT = 0 |
EHP_MC_LOW_PKT_CNT = 0 |
| ENQ_DISCARDED_PACKET_COUNTER = 48,716,055 |
DELETED_PKT_CNT = 40 |
| Rejects: VOQ_MX_QSZ_STATUS |
|

```

