Connectrix Brocade: How to troubleshoot a slow drain device in a SAN

Article Number:000015254 Version:2

Key Information

Audience: Level 30 = Customers	**Original Create Date:** Wed Nov 23 17:05:17 GMT 2011	**Article Type:** How To
Channels: Customer , Internal App	**First Published:** Fri May 31 02:58:29 GMT 2013	**Validation Status:** Final Approved
Originally Created By: William Curren	**Last Modified:** Mon Aug 12 08:57:45 GMT 2013	**Last Published:** Mon Aug 12 08:57:45 GMT 2013

Summary:

Article Content

Goal: How to troubleshoot a slow drain device in a SAN

Issue: Slow performance on multiple hosts in the fabric

I/O errors on multiple hosts in the fabric

SCSI errors

Slow drain

Performance issue

Congestion

Environment: Product: Connectrix Brocade Switches All

Brocade SW: Fabric OS 6.3.1x or later

Cause: A slow drain device is an end devices(can be host/storage/Inter-switch link (ISL)) that does not return buffer credits quickly enough to the switch causing frames to back up through the fabric and thus causing fabric congestion.

Resolution:
1. The first step in troubleshooting any performance related issues is to validate and correct any physical layer issues in the fabric. The **porterrshow™** output will provide an overview of the error counters on each switch.

   Note: These counters are historical and are based on the delta of when the last time they were cleared out. In order to get a valid base line you must clear the counters out by using the below commands and take periodic snap shots of the output of the porterrshow™ command.

   Commands to clear the counters:

   #>**statsclear**
   #>**slotstatsclear**

2. If the physical layer is clean and you suspect a slow drain device in the fabric, the next step in isolating the device is to determine in which direction the I/O is being congested. (We will assume a multi-switch fabric). To do this we check the specific counter details for the Inter-Switch Links (ISLs) using the **portstatsshow <slot/port>™**

   Note: If you are looking at the Supportsaves/Supportshow from the switch, the portstatsshow™ command will refer to the port based on its port index instead of slot/port. If you are live on the switch depending on what Fabric O.S. you are you can reference by port index by using portstatsshow i <port index>.

Example from Supportsave:

a. Find the port index of the ISL port(s) on the switch using the islshow™ command

```
islshow:
  1:271-> 20 10:00:00:00:00:00:00:02  2 Brocade   sp:  8.000G bw: 8.000G
```

b. Check the counters on ISL port in question.

```
portstatsshow 271   (output  is truncated to the fields we are interested in)

tim_txcrd_z        2883985979  Time TX Credit Zero (2.5Us ticks)
tim_txcrd_z_vc  0- 3:  217142      0        1562012    1218509250
tim_txcrd_z_vc  4- 7:  58123331   1605574244  0        0
tim_txcrd_z_vc  8-11:  0        0       0      0
tim_txcrd_z_vc 12-15:  0         0       0       0 0       0

er_rx_c3_timeout    0        Class 3 receive frames discarded due to timeout
er_tx_c3_timeout    2118        Class 3 transmit frames discarded due to timeout
```

Time at Zero Credit:

In a high performance environment it is normal for buffer credits to reach zero at times and buffer credits reaching zero does not mean there is a congestion problem. If however, the buffer credits reaching zero is a very high value in regard to total frames sent it could indicate a congestion problem. It is normal for this counter to show a value of 10 30 % of the total frames sent (there are many caveats to this counter, which explains the wide range of percentages). A percentage much higher may indicate a congestion problem.

3.
1. The class 3 discard counters along with the time at zero credit counter will determine which direction the I/O is being congested. On this port, there are transmitting timeouts (er_tx_c3_timeout), but no receive timeouts.
2. This indicates the I/O being transmitted from port 271 of this switch is being held up, but we are not having issues receiving frames.
3. This indicates that the slow drain device is downstream of this switch

   Note: The er_tx_c3_timeout and er_rx_c3_timeout are only available after Fabric O.S. 6.3.1x. Prior versions will have a similar counter er_c3_timeout. This however does not give you the direction of the discard so it makes troubleshooting where the congestion is coming from difficult.

4. Move to the other side of the ISL on the downstream switch, and check the same counters. Using the ISL show information from step 2 above, the ISL port index 271 connects to ISL port index 20 on the peer switch.

```
portstatsshow 20
tim_txcrd_z        263556      Time TX Credit Zero (2.5Us ticks)
tim_txcrd_z_vc  0- 3:  0        0        263556    0
tim_txcrd_z_vc  4- 7:  0        0       0       0
tim_txcrd_z_vc  8-11:  0        0       0       0
tim_txcrd_z_vc 12-15:  0         0       0       0
er_rx_c3_timeout    983        Class 3 receive frames discarded due to timeout
er_tx_c3_timeout    0        Class 3 transmit frames discarded due to timeout
```

The time at zero transmit credit is fairly low, and have no transmit timeouts, which indicates that there is no congestion transmitting frames to port 271 on the other side of the ISL.

There are receive frames discarded due to timeout, which means it is not able to forward frames transmitted from port 271 on the other side of the the ISl and we are reaching the 500ms timeout and discarding the frames.

This indicates the slow drain device is either connected to this switch, or connected to a downstream switch. If there are one or more downstream switches in the fabric, use the same steps to check the counters on the ISLs to isolate which switch the slow drain device is connected to.

5. Once you have determined which switch the slow drain device is connected to, you will need to isolate the device.

    a. Check the **porterrshow** for physical layer issues.
    b. Check the **switchshow** output for any bad port states. Some examples are below. If you see any ports in this state the port should be disabled and the customer should check the end device attached to the port. You may need to replace the SFP/cable.

    In_sync
    No_sync
    G-port
    Mod_val
    Mod_inv
    Port_flt
    Laser_ftl

    c. Check the **fabriclog --show output**. Look for ports issuing link resets. This can be an indication of the link going through Credit recovery.

    Example:

    #>fabriclog show

```
[output truncated]
01:42:44.383448 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
02:45:04.053762 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
05:58:20.691534 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
23:04:38.181541 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
Switch 0; Sat Oct  1 00:22:10 2011 GMT (GMT+0:00)
00:22:10.471710 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
00:51:57.661666 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
01:55:52.954734 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
03:21:15.018572 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
04:55:47.686715 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
08:46:52.934251 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
Switch 0; Sun Oct  2 06:14:30 2011 GMT (GMT+0:00)
06:14:30.896030 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
20:39:00.733818 SCN LR_PORT (0);g=0x3328          A2,P0 A2,P0 222 NA
[output truncated]
```

    Once you have a list of misbehaving ports you can check the **portstatsshow <slot/port> output** for each port individually.

    If the above steps do not help isolate the problem, you will have to check **portstatsshow** output for all of the end devices.

    From the example above notice that port index 222 is constantly issuing link resets. So now, lets drill into the specific counters on the port index 222 using the 'portstatsshow' command.

    Example from logs:

```
portstatsshow 222
tim_txcrd_z        1448624856  Time TX Credit Zero (2.5Us ticks)
er_rx_c3_timeout    0        Class 3 receive frames discarded due to timeout
er_tx_c3_timeout   32481        Class 3 transmit frames discarded due to timeout
```

    In this case port index 222 was spending an excessive time at 0 buffer credit, discarding frames due to timeout and going through credit recovery. The er_tx_c3_timeout counter tells you that the switch port had data it needed to send to the end device, in this case a Host HBA and the host did not respond within 500ms (a very long time in fibre channel terms) and the switch had to drop the frame.

    This can be caused by bad physical media SFP/cable/end device or because of a load balancing issue on the end device. Faulty physical media can cause r_rdy to be lost. These are use as flow control to replenish buffer to buffer credits indicating the device is ready to accept more data. As a best practice it

is always best to start at the physical layer.

In this case there were no physical layer errors on this port, but this turned out to be a faulty HBA.

Notes: It is not uncommon for there to be multiple slow drain devices connected to multiple switches existing in a fabric. The customer will usually not notice until there is enough backpressure to impact the ISLs. You may have to go through multiple iterations of the above procedure to root out all of the misbehaving devices. The port counters are static. It is important to get periodic baselines on the counters by running the commands statsclear™ and slotstatsclear™ to clear the counters. Allow some run time, preferably a couple of hours, before you collect each set of supportsave outputs. Proactive measures can also be taken by the customer by using the Brocade tools Bottleneckmon, Fabric Watch and/or SAN Health