VIAVI

Accelerating Development and Troubleshooting of Data Center Bridging (DCB) Protocols Using Xgig®

The new Data Center Bridging (DCB) protocols provide important mechanisms for enabling priority and managing bandwidth allocations between different types of traffic. However, given their complexity, it can be difficult to discern from standard network operations whether these mechanisms have been implemented robustly or not. This whitepaper will explore key areas of these protocols which require verification and show developers how to test that network equipment performs as expected.

The new Data Center Bridging (DCB) protocols currently being developed to carry Fibre Channel, TCP/IP, and IPC traffic over a single, converged 10 Gigabit Ethernet network. To date, the DCB Protocols consist of:

- **Priority Flow Control (PFC)**, which defines a new Pause frame format where the pause time is set per priority.
- **Enhanced Transmission Selection (ETS)**, which is not a communication protocol in itself but rather defines a standard way to allocate link bandwidth to each priority.
- **Center Discovery And Exchange Protocol (DCBX)**, which is used by switches and end devices to configure and advertise the PFC and ETS configurations.
- **Congestion Notification (CN)**, which defines the mechanism to manage long-term congestion in a large-scale converged network. As the Congestion Notification specification is still in the early stages of development, this whitepaper will not address it.

White Paper

Using Priority to Allocate Bandwidth

DCB networks are configured to enable VLAN tags on every frame and assign a 3-bit priority to each frame. Since each priority can be paused independently of the others, administrators can configure the network to limit the bandwidth allocation assigned to each type of traffic (i.e., by assigning IP traffic to a different priority than FCoE traffic). Figure 1 shows a partial FCoE frame with a VLAN header and a PCP priority of 0x7.

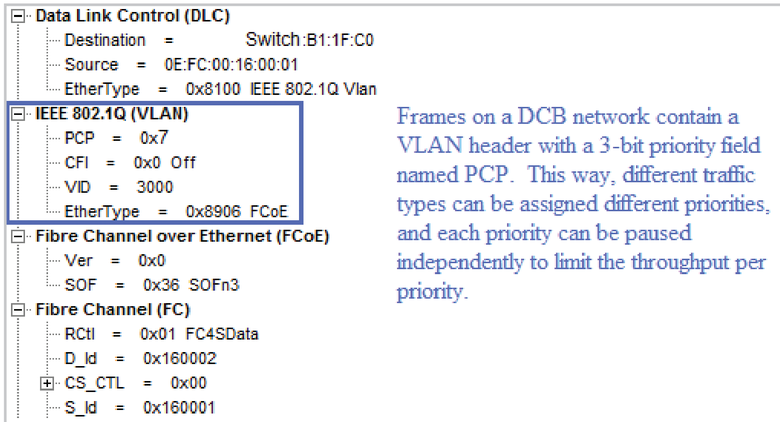


Figure 1: FCoE frame with VLAN header and priority 0x7

Figure 2 shows a Pause frame requesting that the link partner pause all traffic with priority 7 for 1363 µs. Note that all frames with priority other than 7 are still allowed during this period.

Key Verification Points to Consider

PFC, ETS, and DCBX provide important mechanisms for enabling priority and managing bandwidth allocations. However, it can be difficult to determine from network operation alone whether these mechanisms have been implemented properly and robustly, because they are Ethernet link level protocols. A hardware-based protocol analyzer will need to be put in-line to capture 100% of the data on the wire for thorough verifications. In order to have confidence in the implementations of DCB protocols, developers need to analyze performance and responsiveness to specifically verify that each protocol operates as expected. Important points to focus upon include:

- Verifying that the bandwidth assigned to each priority matches the ratio configured using ETS.
- For priorities where PFC is enabled, verifying that the switch/end device actually sends Pause frames on the receive side to limit the throughput as configured using ETS.
- Verifying that Pause frame recipients actually pause frame transmission as requested and for the duration selected.
- For switches, verifying that no frames have been lost on paused priorities by comparing the traffic on both sides of the switch.

The remainder of this whitepaper will show developers how to specifically test that network equipment performs as expected.

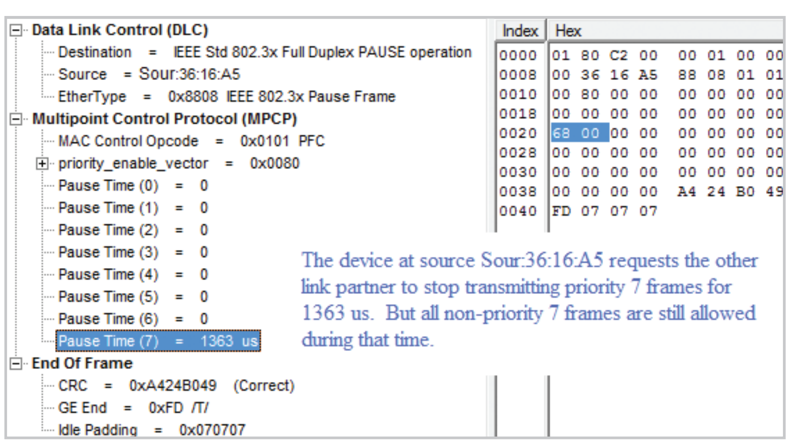


Figure 2: Pause frame stopping all traffic with priority 7 for 1363 µs

Test Configuration

Figure 3 shows a switch with two high-rate traffic sources that are both directed to the same output port. The ETS configuration for FCoE traffic is 40% of the full line rate bandwidth (1200 MB/s) for an output of around 480 MB/s. PFC is enabled for FCoE, which has been assigned to Priority 3. The remaining bandwidth is available for IP traffic, which is not assigned a priority nor has PFC enabled. IP bandwidth will eventually reach 60% of 1200 MB/s, or around 720 MB/s.

Trigger Configuration

To verify that the switch operates efficiently under this configuration, we'll need to capture and save bi-directional traffic on the switch inputs and output using TraceControl in Xgig Analyzer. Since we are interested in when the switch starts pausing FCoE traffic, set the Analyzer to trigger on PFC Pause frames. The configuration shown in Figure 4 will capture 4085 MBs of bi-directional traffic across all inputs/output. 5% of the buffer will hold the traffic sent before the Pause frame while the remaining 95% of the buffer will be used to capture traffic sent after the first Pause frame is detected. In order to enable capture for a longer period of time, conserve buffer space by truncating the frames after 120 bytes in length.

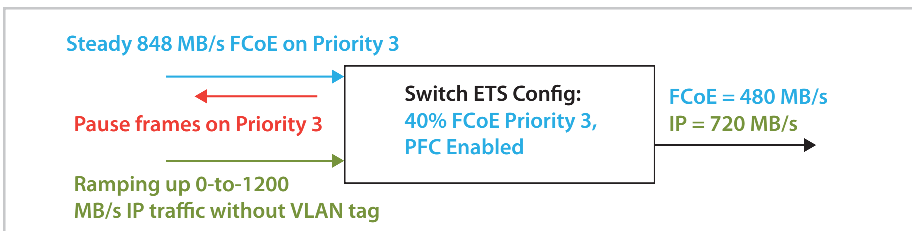


Figure 3: ETS bandwidth allocation configuration

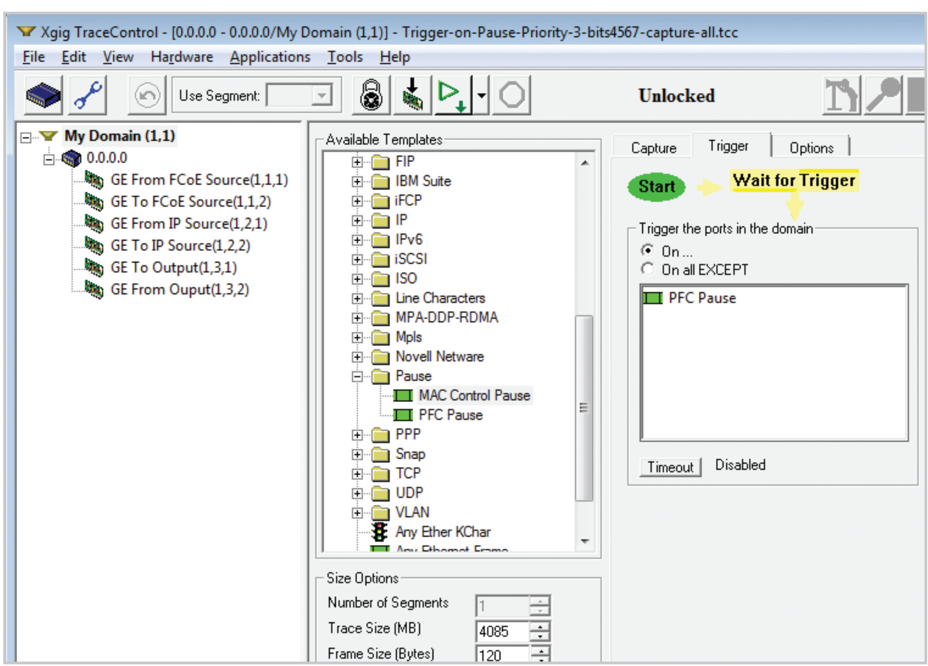


Figure 4: Configuring the Xgig Analyzer to trigger on PFC Pause frames

Displaying Throughput by Priority

Once capture is completed, run Xgig Expert and open both directions of the output. Next, right-click on the graph view and select *Add Counter...* (see Figure 5) to configure a graph showing the MB/s throughput for FCoE (Priority 3) and IP (traffic without a VLAN Tag).

Xgig-Expert then displays both FCoE and IP output traffic (see Figure 6). FCoE traffic started at 848 MB/s, but it quickly decreased to around 465 MB/s when the IP traffic ramped up. This corresponds to 39% of the full line rate. The IP traffic, on the other hand, topped out at around 682 MB/s, or 57% of full line rate. As shown by these percentages, the switch has properly enforced the 40%/60% configured using ETS.

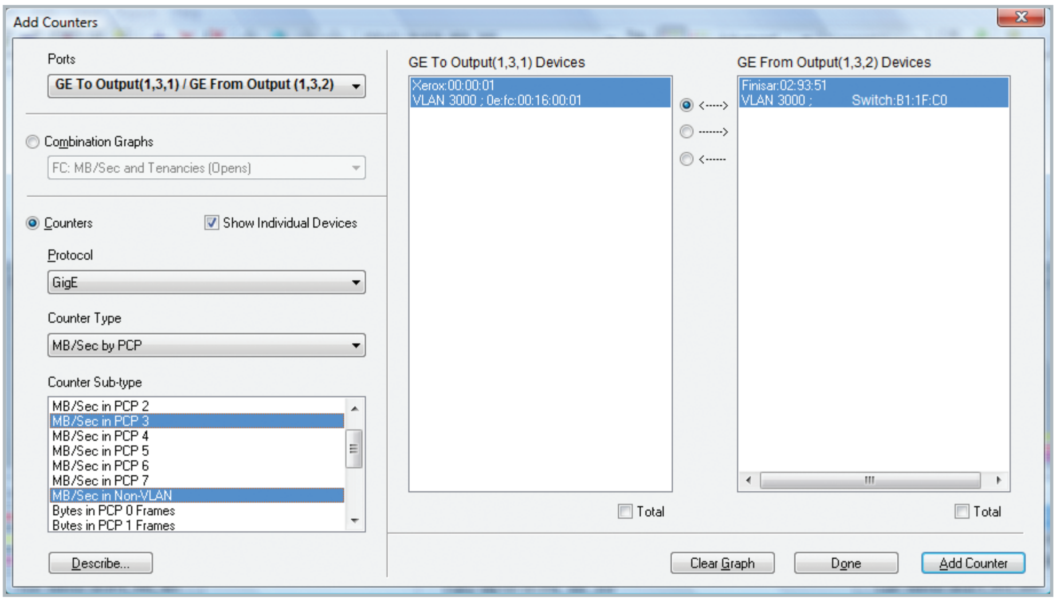


Figure 5: Using Add Counter to configure a graph showing the MB/s throughput for FCoE (Priority 3) and IP (traffic without a VLAN Tag) in Xgig Expert.

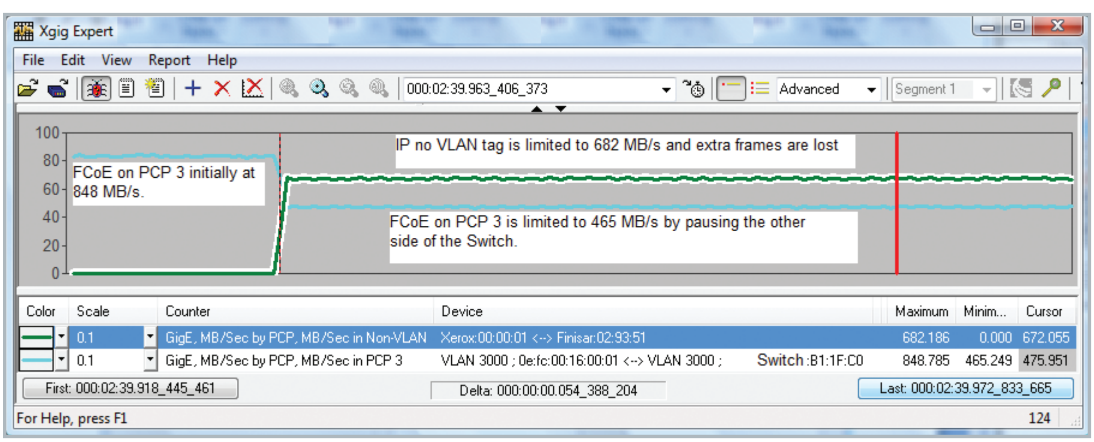


Figure 6: Xgig Expert Graphic View showing ratio of FCoE to IP traffic over time.

Verifying the Sending of Pause Frames

The next step is to verify that the switch actually sent Pause frames to control the rate of Priority 3 FCoE traffic. If the switch failed to send these Pause frames, then it must have dropped FCoE frames to decrease the throughput from 848 MB/s on the input to 465 MB/s on the output. To verify what the switch actually did, close the output ports and re-open both directions of the FCoE input in Xgig-Expert. Xgig-Expert will then analyze all the Pause frames in the capture and report on each one. Figure 7 shows Expert's summary concerning the Pause frames in our example.

Overall, Expert found 596 *PFC Pause Requests* sent to control Priority 3 traffic. For these Pause frames, the recipient was given 1.4 μs to stop transmitting frames upon receiving the request. Row 1 in the figure, *Frame Received while paused during valid overlap,* shows that Expert found 25 frames that were sent during the 1.4 μs period. Row 4, *Frame Received while PFC Class Paused,* shows that Expert reports an error on 565 frames that were sent after the initial 1.4 μs Pause time. This means that the pause recipient was either not pausing or was not responding quickly enough to the Pause frames, i.e. it took longer than 1.4 μs to pause.

Measuring Pause Frame Response Time

Determining how long the recipient took to respond to the pause request is one critical parameter in measuring the PFC performance. To access more detailed pause information, generate a report by clicking on *Report/Create New Report*.... Select all the ports in the *Report Setup* dialog and click *Create Report*. The report will contain all the statistics compiled by Xgig-Expert. The ones of interest in this case are grouped under *GigE\ Gigabit Ethernet - PFC Flow Control Timings* (see Figure 8).

Advanced

Description	Timestamp	Source	Destination	Ports	Type	Total ...	Value 1	Valu...
Frame Received while paused during valid overlap	02:39.933_899_291	VLAN 3000 ; 0e:fc:00:16:00:04	VLAN 3000 ; Switch:B1:1F:C0	1,2,1 / 1,2,2	GigE	25	Time since Pause 0.315 (us)	PCP 3
PFC Pause Request Extended	02:39.930_944_354	Switch:B1:1F:CA	IEEE Std 802.3x Full Duplex PAUSE operation	1,2,1 / 1,2,2	GigE	22	Pause Time 51.933 (us)	PCP 3
Illogical/useless Pause Release	02:39.930_069_024	Switch:B1:1F:CA	IEEE Std 802.3x Full Duplex PAUSE operation	1,2,1 / 1,2,2	GigE	4914	Pause Time 0 (us)	PCP 0
Frame Received while PFC Class Paused	02:39.930_030_856	VLAN 3000 ; 0e:fc:00:16:00:04	VLAN 3000 ; Switch:B1:1F:C0	1,2,1 / 1,2,2	GigE	565	Time since Pause 1.641 (us)	PCP 3
PFC Pause Released	02:39.930_029_216	Switch:B1:1F:CA	IEEE Std 802.3x Full Duplex PAUSE operation	1,2,1 / 1,2,2	GigE	602	Pause Time 0 (us)	PCP 0
PFC Pause Request	02:39.930_029_216	Switch:B1:1F:CA	IEEE Std 802.3x Full Duplex PAUSE operation	1,2,1 / 1,2,2	GigE	596	Req Time 3253.713 (us)	PCP 3

Figure 7: Expert's Summary of Pause Frames

Gigabit Ethernet - PFC Flow Control Timings

From FCoE Source(1,1,1) - To FCoE Source(1,1,2)	% PFC Pause Time	PFC Pause Time (Avg. - us)	PFC Pause Time (Min - us)	PFC Pause Time (Max - us)	PFC Pause Time (Total - us)	Frame overlap time (Avg. - us)	Frame overlap time (Min - us)	Frame overlap time (Max - us)	Frame overlap time (Total - us)
Switch:B1:1F:CA --> IEEE Std 802.3x Full Duplex PAUSE operation	60.608	43.582	31.831	53.787	25,931.464	0.000	0.000	0.000	0.000
VLAN 3000 ; Switch:B1:1F:C0 --> VLAN 3000 ; 0e:fc:00:16:00:04	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Any --> Any	30.304	21.791	0.000	53.787	25,931.464	0.000	0.000	0.000	0.000
Switch:B1:1F:CA <-- IEEE Std 802.3x Full Duplex PAUSE operation	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
VLAN 3000 ; Switch:B1:1F:C0 <-- VLAN 3000 ; 0e:fc:00:16:00:04	0.000	0.000	0.000	0.000	0.000	1.470	0.013	1.820	865.719
Any <-- Any	0.000	0.000	0.000	0.000	0.000	0.735	0.000	1.820	865.719

Figure 8: Detailed Expert Report Showing Time to Respond to Pause Frames

Overall, the FCoE source side of the Switch was paused 60.608% of the entire trace time. This means that the switch only let traffic through about 40% of the time, which is what Priority 3 FCoE traffic was configured to inside the Switch. Thus, this statistic in itself shows that the Switch behaved perfectly.

However, the report also shows that frames were sent up to 1.8 μs after the Pause request which explains why Xgig Expert reported errors on 565 frames. The average PFC Pause Time was 43.582 μs so the recipient did pause for a significant amount of time but just not as quickly as it should have. Not reacting quickly enough to Pause Requests can potentially lead to frame losses at the Switch, which not only translates to degraded performance but also compromises the lossless quality of the FCoE network.

Detecting Lost Frames

To detect frame losses, enable *Cross Port Analysis* in Xgig Expert's Preferences dialog (see Figure 9).

Next close and then re-open the capture in Expert. This time, however, open the input and output ports at the same time. This triggers Xgig Expert to confirm that every frame sent over the input port was received on the output port in the same order. In this example, opening the FCoE input port at the same time as the output port in Xgig Expert does not result in any frame losses reported for FCoE. However, re-opening Xgig Expert with the IP input port and the output port at the same time and performing cross port analysis does reveal that frames were lost (see Figure 10).

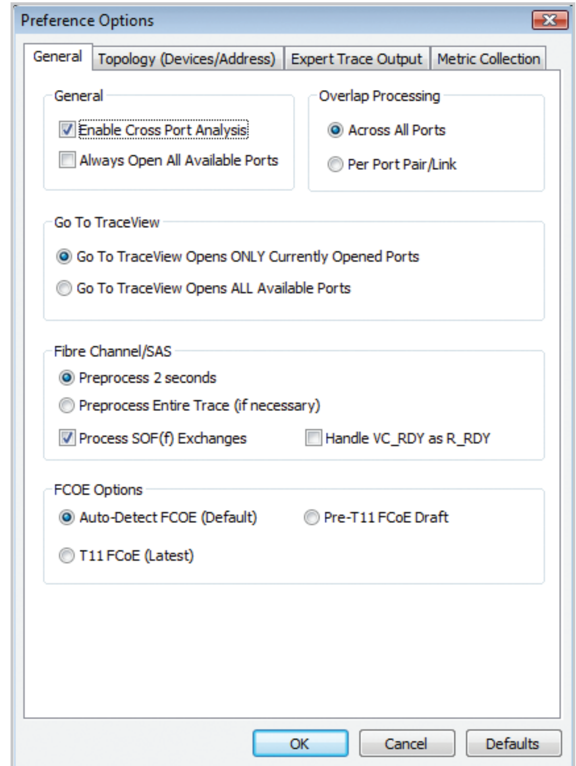
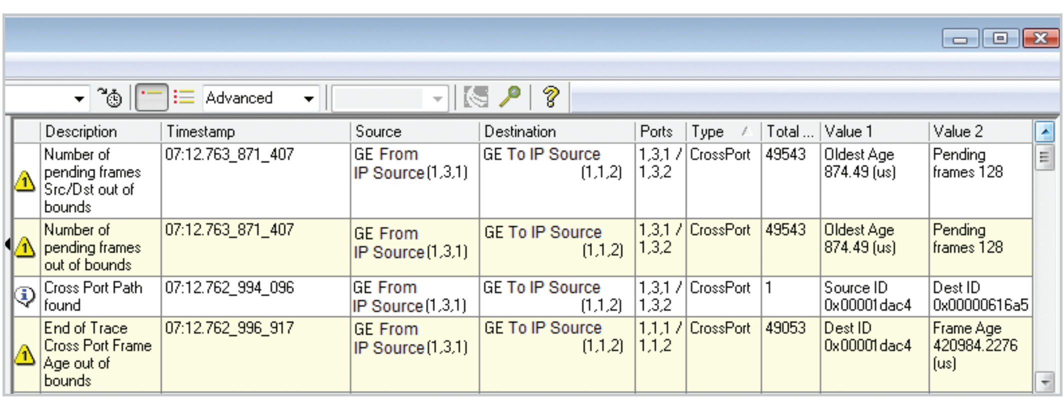


Figure 9: Xgig Expert enabling Cross Port Analysis to Detect Frame Losses

The statistic *Cross Port Path* found serves as an indication that cross port analysis was performed on the trace and that several frames of the input stream were found on the output. The important statistic in this report is *End of Trace Cross Port Frame Age* out of bounds which indicates that 49053 frames were found on the input that were never found on the output. These frames, therefore, were lost at the switch. Note that is consistent with the ETS configuration: since PFC was disabled for IP traffic, the switch had no choice but to drop about 40% of the input IP frames.



Description	Timestamp	Source	Destination	Ports	Type	Total ...	Value 1	Value 2
Number of pending frames Src/Dst out of bounds	07:12.763_871_407	GE From IP Source(1,3,1)	GE To IP Source (1,1,2)	1,3,1 / 1,3,2	CrossPort	49543	Oldest Age 874.49 (us)	Pending frames 128
Number of pending frames out of bounds	07:12.763_871_407	GE From IP Source(1,3,1)	GE To IP Source (1,1,2)	1,3,1 / 1,3,2	CrossPort	49543	Oldest Age 874.49 (us)	Pending frames 128
Cross Port Path found	07:12.762_994_096	GE From IP Source(1,3,1)	GE To IP Source (1,1,2)	1,3,1 / 1,3,2	CrossPort	1	Source ID 0x00001dac4	Dest ID 0x00000616a5
End of Trace Cross Port Frame Age out of bounds	07:12.762_996_917	GE From IP Source(1,3,1)	GE To IP Source (1,1,2)	1,1,1 / 1,1,2	CrossPort	49053	Dest ID 0x00001dac4	Frame Age 420984.2276 (us)

Figure 10: Xgig Expert Cross Port Analysis of IP Input and Output Ports

Summary

This example shows users how to verify that DCB protocols are not only operating as expected but also how to confirm that the protocols have been implemented properly. In addition to verifying that actual bandwidth utilization matches assigned bandwidth allocations, users need to verify that Pause Frames were actually responsible for managing bandwidth. This is achieved by confirming that Pause Frames were sent as quickly as they should have been, ensuring that Pause Frames were received and acted upon as expected, and verifying through cross port analysis of input and output ports that no frames were lost on paused priorities.

To assist developers in comprehensively verifying compliance, measuring performance, and guaranteeing device robustness, Xgig Analyzer 4.5 provides a full complement of metrics, reports, and functions specifically designed to analyze DCB protocols:

Xgig-Expert Metrics

- MB/Sec, %Utilization, Frames/Sec, Bytes/Frame for each priority (PCP) and for non-VLAN traffic
- % PFC Pause Time
- PFC Pause Time (Avg. – us)
- PFC Pause Time (Min – us)
- PFC Pause Time (Max – us)
- PFC Pause Time (Total – us)
- PFC Pause Request Frames
- PFC Pause Release Frames
- PFC Expired Pause Frames
- PFC Extended Pause Frames
- PFC Extraneous Release Frames
- Frame overlap time (Avg. – us)
- Frame overlap time (Min – us)
- Frame overlap time (Max – us)
- Frame overlap time (Total – us)

Xgig-Expert Experts

- PFC Pause Request
- PFC Pause Time out of bounds
- PFC Pause Request Extended
- PFC Pause Released
- PFC Pause Expired
- Illegal/Illogical usage of both MAC Pause and PFC frames
- Frame Received while PFC Class Paused
- PFC Frame has pause value but not enabled
- Illogical/useless Pause Release
- Frame Received while paused during valid overlap

Xgig-Expert Cross Port Analysis

- Out of Order Frame in Network
- Out of Order Frame by Src/Dst
- Cross Port latency out of bounds
- End of Trace Cross Port Frame Age out of bounds
- Number of pending frames Src/Dst out of bounds
- Number of pending frames out of bound
- EOF modified in network
- Cleared Cross Port Frame Age out of bounds

Xgig-TraceView

- New Priority and VLAN/VSAN columns in the default configurations, allowing users to quickly differentiate the frames by their priority
- Double-clicking on the Priority column brings up the Quick Find/Filter/Hide dialog. This allows users to quickly isolate a single priority in the entire trace, hiding everything else.
- Decodes 3 versions of the DCBX protocol

Xgig-TraceControl

Provides templates to trigger and filter:

- The PFC Pause frames
- 3 versions of the DCBX protocol
- Any priority inside each frame

Contact Us **+1 844 GO VIAVI**
(+1 844 468 4284)

To reach the Viavi office nearest you, visit viavisolutions.com/contacts.

dcbprotocolsxgig-wp-san-tm-ae
30162785 500 0117