



[Home](#) [Archives](#) [Profile](#) [Subscribe](#)

[Main](#)

02/16/2012

Does FCoE require DCB Ethernet?

I received a question earlier today that asked the following:

"I have heard some folks argue that ETS (Enhanced Transmission Selection) is NOT required to make FCoE work. I'm thinking specifically of the case in which a CNA is connected to a DCB-capable Ethernet switch that is doing FIP Snooping, which is then uplinked to, say, a Cisco Nexus. The argument goes that the DCB-capable switch does NOT require ETS support to make this design work and be stable - supporting PFC would be enough.

What do you think of that?"

I have two different answers; a practical one and then a theoretical one.

Practical

From a practical point of view, I wouldn't even consider using a FIP Snooping Bridge, ENode or FCF that does not support PFC, ETS and DCBX. To figure out why, you'll need to read the theoretical answer...

Theoretical

From a theoretical point of view, the FCoE protocol can actually function without any DCB features (i.e., DCBX, PFC, ETS), it just won't allow the ULP (SCSI-FCP) to work very efficiently. By this I mean the driver supporting the ULP would spend a significant portion of its time in error recovery (because of frame loss due to normal periodic congestion in the Ethernet network) rather than doing useful work. This is specifically a problem for SCSI-FCP, because it "evolved" in a lossless environment (FC / BB_Credit) and as a result no significant amount of work was done on a retransmit like mechanism. In other words, since frame loss happens so infrequently in FC, the easiest (and cheapest) way to recover from a frame loss is to retry the entire exchange.

When attempting to run FC / SCSI-FCP directly over Ethernet, as is done with FCoE, the solution to the frame loss problem is to make the Ethernet link



[Erik Smith](#)

[ARCHIVES](#)

[September 2022](#)

[August 2020](#)

[April 2019](#)

[December 2017](#)

[November 2017](#)

[July 2017](#)

[April 2017](#)

[September 2016](#)

[July 2016](#)

[February 2016](#)

SEARCH

[CATEGORIES](#)

[Subscribe to this blog's feed](#)

DISCLAIMER

This is not an official EMC blog. The opinions expressed here are my personal opinions. Content published here is not read or approved in advance by EMC and does not necessarily reflect the views and opinions of EMC nor does it constitute any official communication of EMC.

lossless, which means that some form of Pause will need to be used. You could use either Pause (802.3x - a.k.a Link Pause) or Priority Flow Control (PFC - 802.1Qbb). The problem with using "Link Pause" is that it pauses the entire link. Since one of the main selling points of FCoE is to converge the network, this means that multiple traffic types will usually be present when FCoE is in use. When Pause is used in this type of environment, all traffic using the link will also be paused. This behavior presents a significant problem because it can cause congestion spreading and I believe it could get in the way of TCPs flow control mechanism because no frames will be dropped. The solution is to use PFC. When using PFC, FCoE traffic can be lossless and other types of traffic can remain lossy.

With the above in mind, ETS is needed to ensure that lossy and lossless traffic are each guaranteed a certain amount of bandwidth. Is it strictly required to enable FCoE to function? No!

Then there is DCBX which isn't required to allow FCoE to function properly either, but it does allow for the ENodes (CNAs) to be properly configured (automatically) to use the DCB settings defined on the switch. This prevents misconfiguration of the ENodes and it makes life easier for the end user.

All of this is interesting from a theoretical point of view, but as I said previously in practice, I wouldn't even think about using an ENode or an FCF that didn't support DCBX, PFC, ETS. These DCB extensions should be considered required in order to make FCoE reliable and easy to use. As a side note, these extensions are also supported by all of the FCoE implementations I am aware of, so we really don't have a choice anyway...

Thanks for reading!

Posted at 09:14 AM | [Permalink](#)

 Reblog (0)

Comments

 You can follow this conversation by subscribing to the [comment feed](#) for this post.



Erik, thank you for that thorough answer. Brilliant, as usual. If it's OK, I would like to add another component to the answer: the popular view. From talking to industry data center architects, in particular those who manage high-performance environments, I think the prevailing view is that FCoE requires DCB to provide a lossless environment and any switch that does not offer ETS will be summarily rejected as an inadequate solution that can place the environment at risk.

Has that been your experience, too?

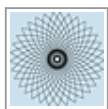
Digging a bit on the technical side of ETS, my understanding is that ETS does not only provide guaranteed BW, which one can configure today with legacy QoS tools, like Cisco's MQC, but that it will allow a statistical multiplexing type model on top of that. In that model, the traffic types that fall under the ETS Traffic Class (TC-2) and are scheduled according to the ETA algorithm (which I know is not specified in the standard - usually its WRR) will not only be

guaranteed a configured BW percentage of the link, but will also be able to have more BW DYNAMICALLY assigned to them as the offered load of the other traffic type drops/fluctuates.

For example, in a scenario in which I have LAN and SAN traffic under the ETS Traffic Class, and the LAN traffic is assigned 60% of the BW and the SAN 40%, that means on a 10G link, the LAN gets 6G and the SAN gets 4G guaranteed. HOWEVER, as I understand the technology, each traffic type will be able to exceed that guaranteed amount of 60 and 40 IF the other traffic class's offered load does not meet its minimum. So, if at time t1, the LAN traffic's offered load is only 2G, not 6G, and the SAN traffic needs to burst, ETS will schedule the SAN traffic such that it can use the 4G that the LAN traffic is not using at that particular time. In short, each traffic type can use the other's BW if they're not using it at a particular type. This is done dynamically.

Is this your understanding of what ETS is and what it does?

Posted by: Victor | [02/16/2012 at 09:49 AM](#)



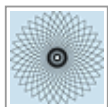
Hi Victor, thanks! Your description matches my understanding of ETS. There's a little bit of additional information in the EMC FCoE TechBook if you're interested.

Posted by: [Erik Smith](#) | [02/16/2012 at 10:10 AM](#)



Thanks, my good man.

Posted by: Victor | [02/16/2012 at 10:18 AM](#)



Hi Eric,

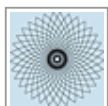
I agree. The resilience and reliability of FC has caused a lack of interest (or shortage of development resources, depends how you look at it) in the ULP side of the protocols like SCSI. The networking peeps have eliminated this by building a very resilient routable framework on top of that called TCP/IP. :-)

For this reason SCSI works on TCP/IP but not ethernet. Hence the invention of FCoE. :-)

If you leave out DCB, ETS and PFC I would argue you will find yourself on very thin ice when trying to push channel protocols through a CNA.

Regards,
Erwin

Posted by: [Erwin van Londen](#) | [03/15/2012 at 02:12 AM](#)



Hi Erwin, I agree. TCP/IP provides a reliable transport and that block I/O protocols (perhaps most protocols that utilize TCP) require a lossless or seemingly lossless transport in order to function properly.

Regards, Erik

Posted by: [Erik Smith](#) | [03/15/2012 at 08:32 AM](#)



As you say all devices in practice require DCBX and pfc to get running by default, some have manual override to allow configuration override something that was in the draft dcb standards but not in the final version. Some but not all devices require ets to be configured as well. Practically the same qos can be achieved without ets through normal l2 cos at the priority level (ets being at the group of priority levels - something many people forget). However in reality I 100% agree with your position that the best practice is clearly DCBX PFC and ETS (and FIP Sooping on any L2 switch).

Posted by: Simon Gordon | [05/19/2012 at 02:28 AM](#)



Hi, Does VN2VN FCoE require DCB Ethernet? and how to config it?

Posted by: jason | [04/28/2014 at 11:48 PM](#)



Hi Jason, all (FC-BB-5 and FC-BB-6) variations of FCoE require a lossless transport, so yes DCB Ethernet is a requirement with VN2VN.

How you configure support for VN2VN is platform dependent. Which CNA, DCB capable switch and Storage platform are you planning to use?

Posted by: [Erik Smith](#) | [04/29/2014 at 08:52 AM](#)

The comments to this entry are closed.

[Brass Tacks](#) [Powered by TypePad](#)