

**A Comparison of Fibre Channel over Ethernet / Converged
Enhanced Ethernet Proposals**

07-336v0

Brocade Engineering Team

1 Introduction

Two proposals have been offered to t11 for consideration to provide for transport of Fibre Channel frames over Ethernet fabrics with certain defined characteristics (e.g. support of jumbograms). One is T11/07-303v0 FCoE - Specification (FCoE), submitted by Cisco et. al. The other is T11/07-292v2 FCoCEE Proposal for FC-BB-5 (FCoCEE), submitted by Brocade et. al. Not surprisingly, these proposals are similar in many respects. However, there are a number of significant differences.

This document attempts to summarize the differences between the two proposals in an objective manner. However, it must be noted that this document was developed by Brocade based on Brocade's best understanding of FCoE. We welcome all input from the developers of FCoE to correct any misunderstandings that might be present in this document.

2 Basic Difference Between the Architectural Models

FCoE proposes two architectural models, Type C and Type D:

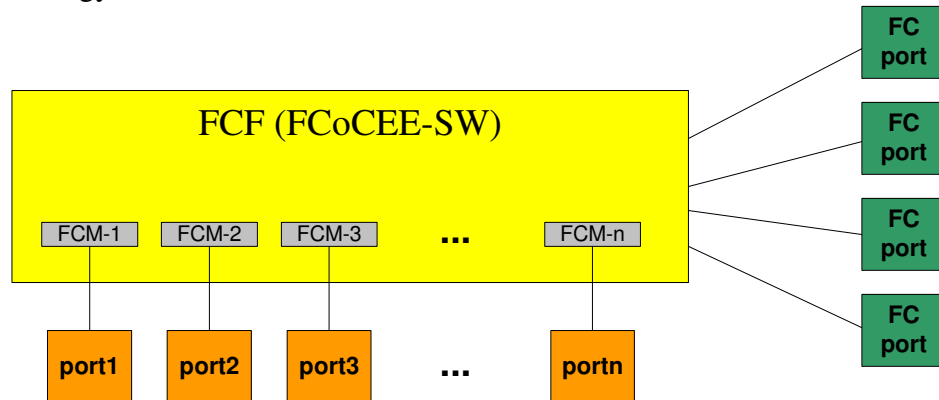
- The Type C model contains a single Ethernet Bridge. FCoE seems to imply that the inclusion of an 802.1Q or 802.1D switch is required.
- The Type D model contains 2 or more Ethernet Bridges. The FCoE proposal seems to discourage the use of this model.

FCoCEE allows, but does not require the presence of an Ethernet Bridge. In a deployment carrying FCoCEE traffic exclusively, there is no need for an Ethernet bridge. In this case, the FCoCEE switch operates exactly like a FC-SW switch, except for the addition of the FCoCEE entity on each FCoCEE port that performs the encapsulation / de-encapsulation function.

Observations / Questions:

- Is there a need for defining two Types of forwarding models, one for directly connected deployment and another for in-directly connected deployment? Although FCoE provides two different models, FCoE seems to discourage usage of the Type D FCF model.
- Does FCoCEE support both the deployments equally well with a single model. If so, does having a single model simplify user deployment?
- FCoCEE doesn't require an Ethernet Bridge inside the FCoCEE switch. The FCoCEE model appears to support the deployments illustrated in the FCoE proposal without requiring an Ethernet Bridge.

- For reference, below is an illustration of the FCoCEE model using FCoE terminology:



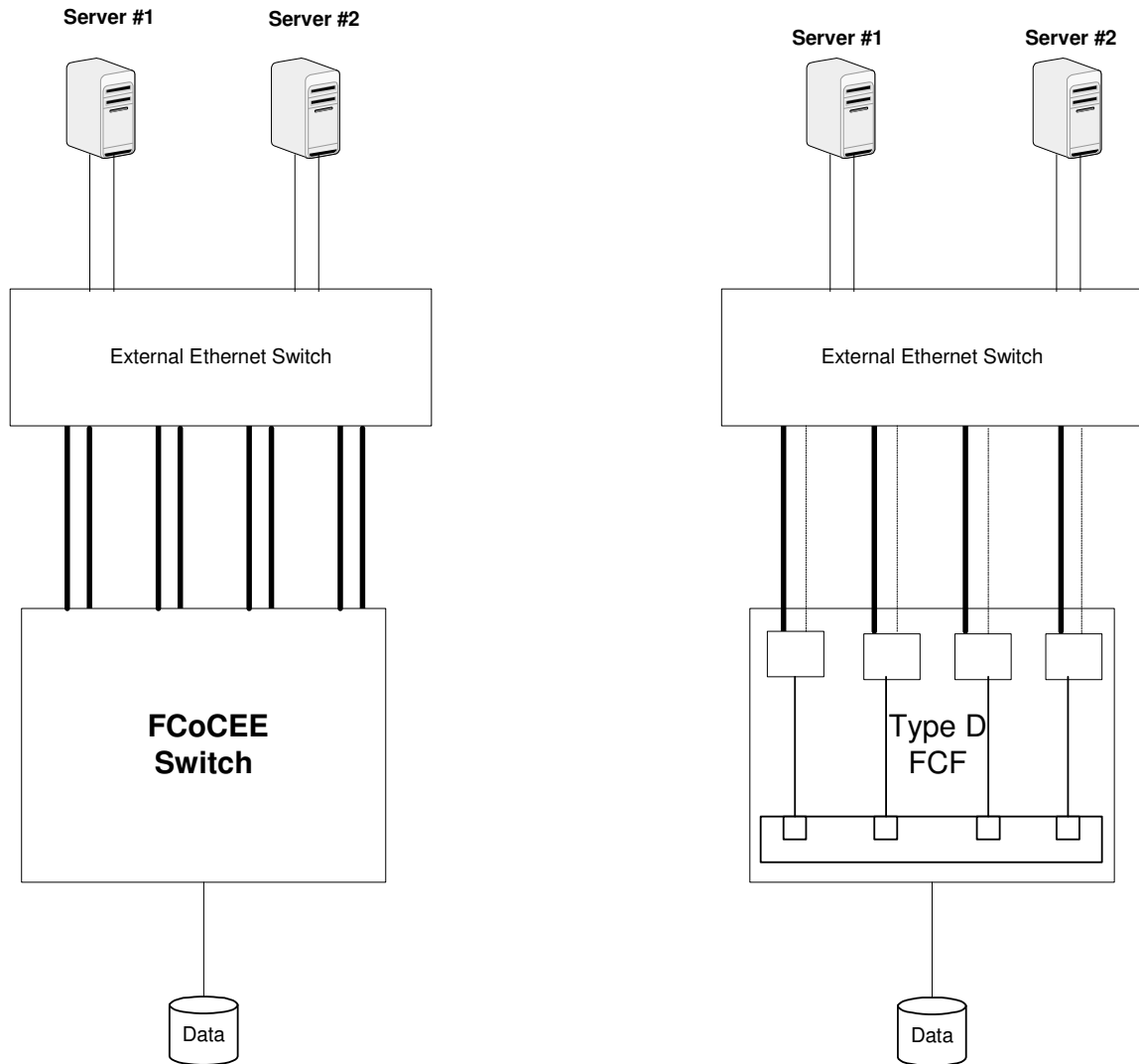
- In this alternative model each FCoCEE port has an FCM and each FCM has a universal MAC. In this way the Ethernet addressing is constrained to one simple form.
- In many deployments one might find it convenient to have a combined FCoCEE switch with a standard 802 switch. The FCoCEE architecture supports this in a straightforward manner. To accomplish this, the two switches are simply logically connected to each Ethernet port in parallel. The FCoCEE switch handles the FCoCEE traffic, and the 802 switches handle the remaining Ethernet traffic. Even in this case, the FCoCEE model does not require multiple architectures. This approach maintains the model the users are accustomed to today; that is, storage switches handle storage traffic (FCoCEE-SW) and Ethernet bridges handle by Ethernet traffic (802).

3 Multipathing and Associating VN_Ports to FCMs

The introduction of Ethernet switches into the network creates an additional complexity that does not exist in standard Fibre Channel. In Fibre Channel, an N_Port is attached to exactly one F_Port. With certain FCoE/FCoCEE topologies, it is possible for an N_Port to reach multiple F_Ports. Therefore, a mechanism must be provided in either proposal to associate an N_Port with one of the possibly multiple reachable ports.

The differences in the architectural models result in differences in the way multipathing operates and the method that N_Ports associate themselves with F_Ports. The figures below are used to illustrate these differences:

Note: The following analysis is based on our current understanding of the FCoE proposal and makes certain assumptions to finishing the analysis. It is perfectly possible for some of the assumptions to be not valid and our understanding of the proposal is not complete.



Observations / Questions

- FCoE requires spanning tree protocol to be enabled on the FCF to eliminate the loops between the internal Ethernet Bridges and the external Ethernet switch. It appears that FCoE restricts each Internal Bridge to be in a single VLAN. It will be assumed that each Ethernet Bridge can only participate in a single VLAN. This effectively makes only one external facing port to be in forwarding state on each internal Bridge. All the other ports are in blocked mode. Ensuring multipathing in this configuration will require extensive VLAN and MSTP configuration as described below:

Configuration steps on each Internal Ethernet Bridge (this step does not apply to FCoCEE since FCoCEE does not contain any internal bridges):

- Enable multiple spanning tree (MSTP).
- Configure one spanning tree per link between each internal bridge and the external switch (so there will be four spanning tree instances). One

would have to configure more if more ISLs are desired between the switches.

- c. Configure 4 VLANs. Configure each port on each bridge to be in a separate VLAN. Effectively there are four separate non-overlapping forwarding topologies between the two switches.
- d. Now, configure each FCM attached to each internal Bridge to be in separate VLANs. **Note: It's not clear in the FCoE proposal that a single FCM behind an Ethernet Bridge in a Type D FCF can participate in multiple VLANs. It will be assumed that it does not.**

Configuration steps on the External Switch (common to both proposals):

- e. Enable multiple spanning tree (MSTP).
 - f. Configure one spanning tree per inter switch link. (so there will be four spanning tree instances enabled on both sides since there are a total of four links).
 - g. Configure 4 VLANs. Configure each ISL to be in separate VLANs.
 - h. Configure servers facing ports to be in their respective VLANs. Server1 gets VLAN1 and VLAN2. Server2 gets VLAN3 and VLAN3.
- One could argue that link aggregation could be used to solve the above problem. It is not clear that this would work. Note that between a given HBA and FCM, all frames use the same source / destination MAC addresses, and there is no IP content. Therefore, there is no indication for the external switch to allow it to identify separable flows to split between the links. Consequently, even if link aggregation is enabled, it is likely that all of the traffic will flow on a single link.
 - With FCoCEE, there are no loops with respect to the FCoCEE traffic and therefore all of the links remain available. Recall that the FCoCEE model does not include an Ethernet bridge. Consequently, from the point of view of the external bridge, it is the only bridge in the network.
 - With FCoCEE, each link between the external switch and the FCoCEE switch is terminated by an FCoCEE entity. Therefore, traffic may be shared among all available links.
 - If the FCoCEE switch includes a parallel Ethernet switch, the above observations remain true for the FCoCEE traffic. Of course, normal spanning tree or link aggregation techniques would be required to support the non-FCoCEE traffic.
 - With FCoE, it is possible that the external switch will flood the FLOGI to multiple FCMs. This problem is exacerbated if one imagines multiple FCoE switches connected to the external switch. FCoE is incomplete in its specification of how this is handled.
 - FCoCEE Switches are configured to associate each XN_Port to an XF_Port. This solves the problem of the possible flooding of FLOGI's. Enabling MMRP/GMRP on

the External switch provides multicast suppression if desired. No server configuration is required.

- FCoCEE does not require any configuration to keep all the provisioned links between the external switch and the FCoCEE switch to be active. Multipathing works as it would in a standard Fibre Channel fabric. Any lossless, jumbogram capable, but otherwise no frills external Ethernet switch will work just fine.

4 MAC Address Generation

FCoCEE uses conventional MAC addresses, i.e., the globally unique MAC address assigned to the NIC (or the locally administered address assigned to the NIC via a virtualization layer) is used for FCoCEE traffic as well as other non-FCoCEE traffic.

FCoE generates a locally administered address by concatenating an OUI with the corresponding (source or destination) Fibre Channel ID of the encapsulated Fibre Channel frame.

Observations / Questions:

- FCoCEE reduces the total number of MAC addresses that need to be handled by the edge Ethernet switches compared to FCoE. This results in reduced L2 forwarding tables and associated learning/aging/flooding overheads in the edge switches, especially when large IO consolidation is deployed through the network.
- FCoE requires IVL deployment to avoid MAC address collisions resulting from same FC-ID assigned in different virtual fabrics. Since this entire set of MAC addresses are indeed processed by the FCF, this stresses the scalability of the internal Ethernet Bridge. This also stresses the HBA implementations in a server consolidation environment, where each bladed server chassis may host 128-256 virtual machines going forward.
- FCoE appears to specify a one-to-one mapping between VLANs and the Virtual Fabrics. FCoCEE has no such restriction. FCoCEE architecturally decouples VLANs from Virtual Fabrics. This allows for Fibre Channel standard based VF tagging, negotiation and discovery by XN_Ports independent of VLANs. FCoE mandates an identity mapping between VID and VFID, thereby restricting the VFs to the VLANs negotiated or configured already. Additionally, this one-to-one mapping requires deployment of Independent VLAN Learning (IVL) in the CEE cloud. Support for IVL is optional in IEEE 802.1. Although IVL are widely available and deployed, they impose scalability limits. Shared VLAN Learning, which is inherently more scalable, does not appear to work with FCoE.
- FCoE utilizes locally administered addresses. Other applications, such as certain HPC and virtualized operating systems also generate locally administered addresses.

Therefore, there is a potential that conflicts may occur between the independently generated locally administered addresses. FCoE attempts to avoid this issue by using an Organizationally Unique Identifier (OUI) as part of its address. However, OUIs have no significance in locally administered addresses and therefore no assurance against address collisions is provided.

- FCoE attempts to make the conversion from Fibre Channel to FCoE stateless. It does not appear that it succeeds in this. When the next hop of the frame is another FCoE switch, then the MAC address of the FCM must be used. This implies state in the conversion process. Edge devices are handled differently using the mapping function. It is not clear that the state saved justifies the increase in FCM complexity.
- FCoE approach appears to increase the complexity on NICs since it requires the XN_Port to remember and listen to multiple locally administered MAC addresses, one per FC-ID assigned.

5 Security

Differences in the way that FLOGIs are processed between the two proposals provide contrast in the level of inherent security provided.

- In FCoE, the CEE cloud never really learns the F_Port Controller's unique MAC address; therefore all initial FLOGI frames are flooded in the network. It's not clear how it will be ensured that only the FCFs process these frames. It's possible that a rogue station could connect itself into the network and keep sending this FLOGI and create a DOS attack on non-FCF entities, because non-FCF entities have to at least pick up these frames to drop them.
- FCoCEE ports are configured to receive and process the specific multicast group addresses. These FCoCEE ports can use standard IEEE GMRP/MMRP (a standard protocol to manage multicast groups in Ethernet networks) to inform the Ethernet switches to forward all initial FLOGI frames "only" to these FCoCEE ports. Non-FCoCEE Ethernet ports in the network are not even aware of these frames.
- FCoE includes a method for direct communication between VN_Ports when there is an Ethernet-only path. In this case, how is access control enforced in the data-path with no FC Zoning involved?

6 Data Integrity

Proper operation of the Fibre Channel Protocol requires that frames be delivered through the fabric within $(R_A_TOV - E_D_TOV) / 2$ or that the frame not be delivered.

Violating this requirement in certain circumstances can result in undetected data corruption.

In conventional Fibre Channel fabrics, this requirement is met by enforcing a maximum time that a frame can live in a single switch (500ms is typical). However, FCoE fabrics become more complex. It is not hard to imagine a fabric that consists of a FCoE Server, an edge Ethernet Switch, a core Ethernet switch, an FCoE switch providing conversion to Fiber Channel, a Fibre Channel core, and maybe even an additional Fibre Channel edge switch. The number of hops therefore becomes considerably more than what one typically sees in pure Fibre Channel. Additionally, it should be noted that the default frame lifetime specification for Ethernet is one second, and may be configured to up to four seconds (2x to 8x the typical Fibre Channel value).

FCoE appears to ignore this issue.

FCoCEE addresses this issue in the same manner it has been addressed in FC-IFR, FCIP, and iFCP. A timestamp is added to each frame allowing an end-to-end (within the FCoCEE domain) timeout to be enforced. Therefore, the frame lifetime settings of individual Ethernet switches becomes irrelevant.

7 Cut Through and Minimum Frame Length

Ethernet requires a minimum frame length of 64 bytes. The minimum frame length for Fibre Channel is 28 bytes.

FCoE requires that padding be added to the Ethernet frame for short Fibre Channel frames to bring the total frame size up to 64 bytes. As a result, a length field is required in the header to indicate the amount of padding present. This has the unfortunate consequence of disabling cut-through operation when transferring a frame from Fibre Channel to Ethernet. This is due to the fact that the entire Fibre Channel frame must be received in order to set the value in the length field.

FCoCEE does not encounter this problem. The headers in FCoCEE are designed to ensure that the minimum frame size is always at least 64 bytes. Thus cut through operation is not precluded.

8 VE_Port Discovery

FCoE provides an automated VE_Port discovery mechanism. FCoCEE provides no such mechanism and could well benefit from this. Additional study is underway.