

*Claudio DeSanti (Cisco), Erik Smith (EMC),
Craig Carlson (QLogic), Ed McGlaughlin (QLogic)
Pat Thaler (Broadcom)*

1 Distributed FCF

1.1 Definitions

1.1.1 Controlling FCF: An FCF able to control a set of FDFs in order to create a Distributed FCF.

1.1.2 Controlling FCF Set: The Switch_Names of the up to two Controlling FCFs that are part of a Distributed FCF.

1.1.3 Distributed FCF: A set of FDFs associated with at least one Controlling FCF, that controls the operations of the set of FDFs. A Distributed Switch is defined by the administrative configuration of the Controlling FCF Set and of the FDF Set.

1.1.4 FDF (FCoE Data-Plane Forwarder): A simplified FCoE switching entity that forwards FC frames among VA_Ports and VF_Ports through a FCDF Switching Element (see 1.4). An FDF shall support at least one VA_Port capable FDF-MAC and may support one or more VF_Port capable FDF-MAC. Each FDF-MAC shall be coupled with an FCoE Controller function and may be coupled with a Lossless Ethernet bridging element (see 1.4).

1.1.5 FDF Set: The Switch_Names of the FDFs that are part of a Distributed FCF.

1.1.6 VA_Port (Virtual A_Port): An instance of the FC-2V sublevel of Fibre Channel that communicates with another VA_Port and that is dynamically instantiated on successful completion of a FIP ELP Exchange. A VA_Port is uniquely identified by an A_Port_Name Name_Identifier and is addressable by the VA_Port connected to it through the A_Port Controller address identifier (i.e., FFFFF9h).

1.1.7 VA_Port/FCoE_LEP pair: A VA_Port and its associated FCoE_LEP.

1.2 Overview

A Distributed FCF is a set of FDFs associated with at least one Controlling FCF, that controls the operations of the set of FDFs. Figure 1 shows an example of Distributed FCF composed of a Controlling FCF and two FDFs.

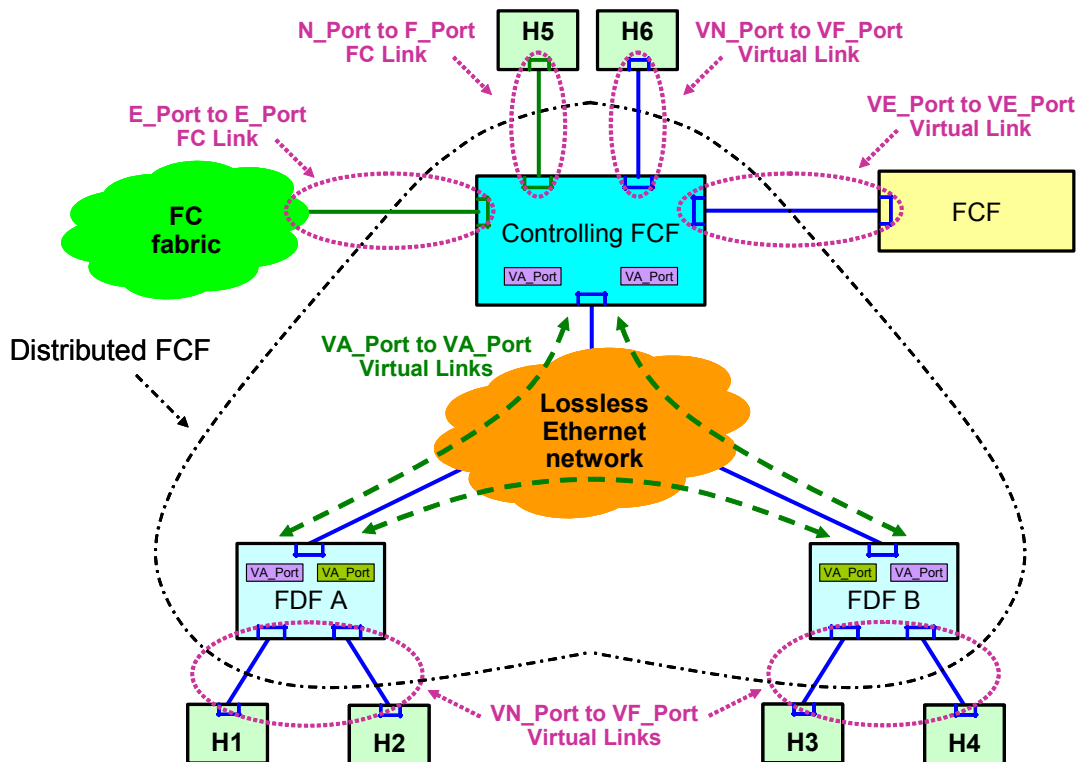


Figure 1 – Example of Distributed FCF

From an external point of view (i.e., outside the dotted and dashed black line in figure 1), a Distributed FCF behaves as an FCF. In particular, a Distributed FCF supports the instantiation of VN_Port to VF_Port Virtual Links with ENode MACs and of VE_Port to VE_Port Virtual Links with FCF-MACs. VN_Port to VF_Port Virtual Links are supported by both FDFs and Controlling FCFs, while VE_Port to VE_Port Virtual Links are supported only by Controlling FCFs. This means that it is possible to connect a Distributed FCF to another FCF or to a native FC fabric only through a Controlling FCF, not through an FDF.

From an internal point of view (i.e., inside the dotted and dashed black line in figure 1), VA_Port to VA_Port Virtual Links enable FCoE frames forwarding between Controlling FCF and FDFs, as well as between FDFs. In addition, VA_Port to VA_Port Virtual Links are used to exchange control information between Controlling FCF and FDFs (see T11/11-225v0).

FDFs are not able to operate properly without a Controlling FCF, therefore the Controlling FCF is a single point of failure in a Distributed Switch configuration with only one Controlling Switch, as the one shown in figure 1. To avoid this issue, Distributed FCFs may support a redundant configuration of two Controlling FCFs, a Primary one and a Secondary one. The Secondary Controlling FCF keeps its state synchronized with the Primary and is able to take its place in case of failure according to the Controlling Switch Redundancy Protocol (see T11/11-224v0).

Figure 2 shows an example of Distributed FCF including a redundant pair of Controlling FCFs.

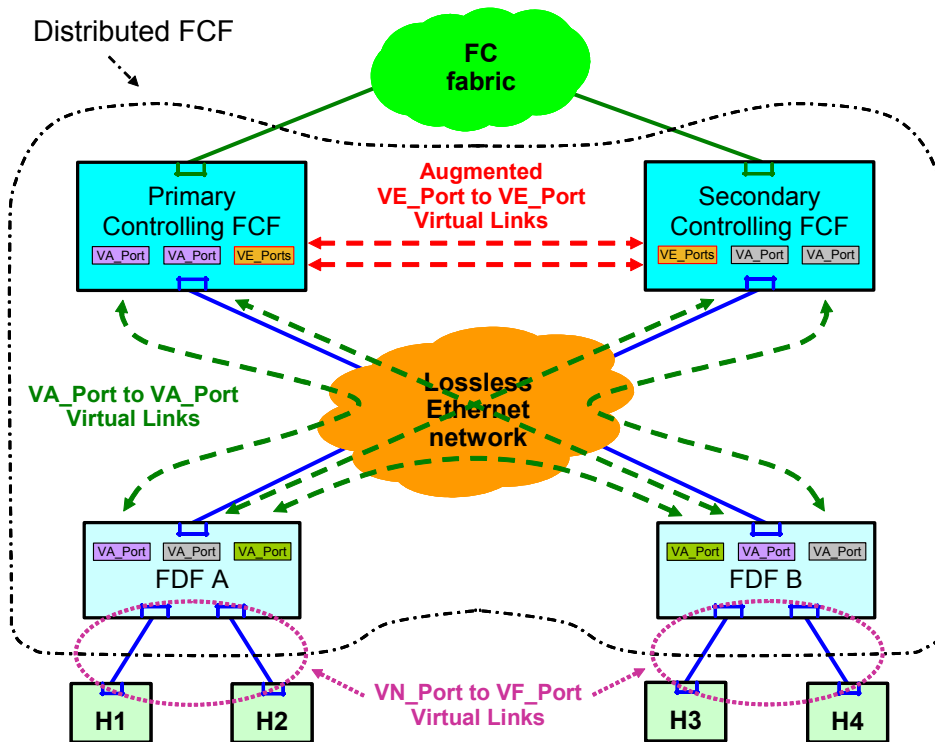


Figure 2 – Example of Redundant Distributed FCF

The two Controlling FCFs in a redundant Distributed FCF instantiate at least two Augmented VE_Port to VE_Port Virtual Link between themselves, where the term ‘augmented’ indicates that Virtual Link is used also for the Redundancy protocol, in addition to normal VE_Port operation (see T11/11-224v0).

The Controlling FCFs use a Virtual Domain_ID to perform N_Port_ID allocations for N_Ports connected to the FDF Set of the Distributed FCF (i.e., the Virtual Domain_ID is used as the most significant byte in all N_Port_IDs allocated to N_Ports that are attached to the FDF Set). Using a Virtual Domain_ID to assign N_Port_IDs enables seamless operation in case of failures of one of the two redundant Controlling FCFs (see T11/11-224v0). As a result, a redundant Distributed FCF typically uses three Domain_IDs: one for each Controlling FCF and one for the Virtual Domain_ID. To properly support the operations of a Virtual Domain_ID, a Controlling FCF shall have a Switch_Name to associate with the Virtual Domain_ID, in addition to its own Switch_Name.

The two redundant Controlling FCFs instantiate VA_Port to VA_Port Virtual Links to enable the forwarding of FCoE frames and the communication of control information between Controlling FCFs and FDFs. In a redundant configuration, FDFs instantiate VA_Port to VA_Port Virtual Links to each of the Controlling FCFs and between themselves, if they are directly reachable through the Ethernet topology.

A Distributed FCF may have a cascaded FDF configuration when FDFs with at least two VA_Port capable FDF-MACs and independent Lossless Ethernet Bridging Elements are used (see 1.4). Figure 3 shows an example of such a configuration.

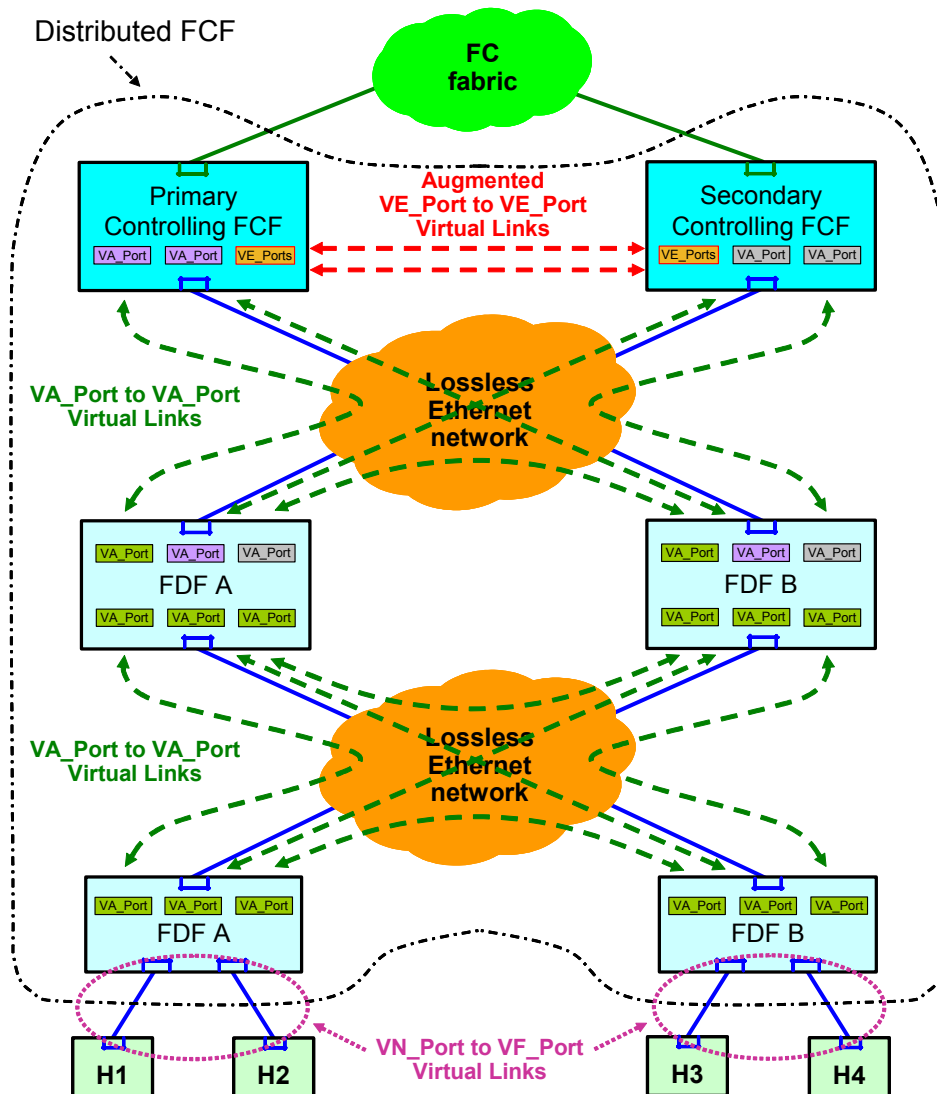


Figure 3 – Example of Distributed FCF with Cascaded FDFs

A Controlling FCF is uniquely identified by its Switch_Name Name_Identifier, as an FCF. An FDF is uniquely identified by its Switch_Name Name_Identifier. A Distributed FCF is defined by an administrative configuration on the Controlling FCFs, listing:

- a) the Switch_Names of the two Controlling FCFs that act as the Primary/Secondary pair for that Distributed FCF (i.e., the Controlling FCF Set); and
- b) the Switch_Names of the FDFs that are part of that Distributed FCF (i.e., the FDF Set).

1.3 Controlling FCF Functional Model

Figure 4 shows the functional model of a Controlling FCF, where the bracketed functional components are optional. A Controlling FCF is an FCF that supports the instantiation of VA_Ports over its Lossless Ethernet MACs (FCF-MACs), in addition to VE_Ports and VF_Ports. As for any FCF, each FCF-MAC of a Controlling FCF may be coupled with a Lossless Ethernet bridging element. A Controlling FCF may have a Fibre Channel Fabric interface, providing native E_Port, A_Port, and F_Port connectivity.

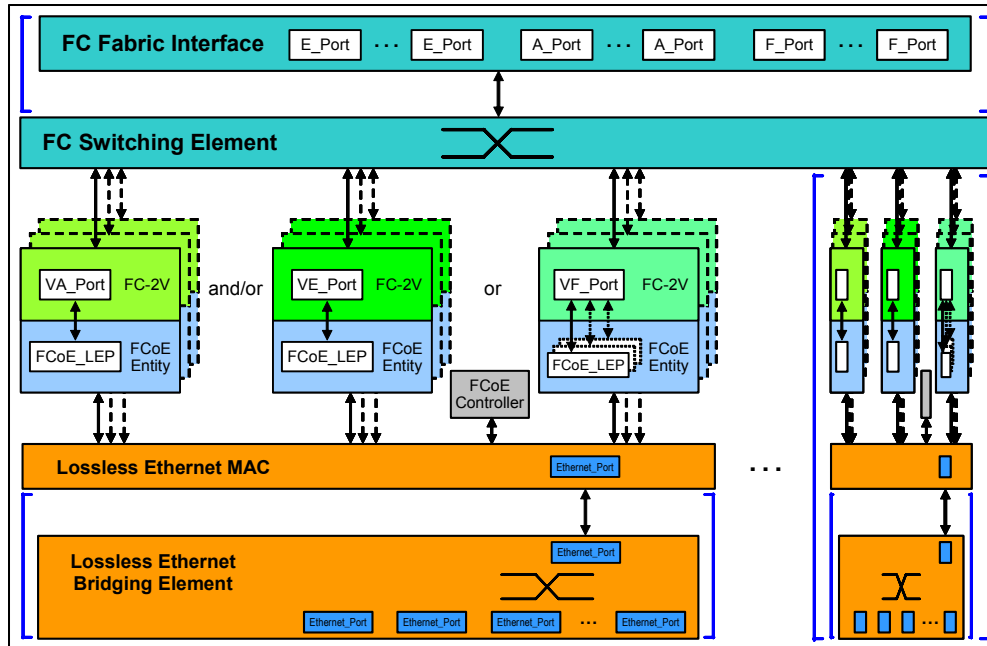


Figure 4 – Controlling FCF Functional Model

NOTE 1 – Other combinations of Lossless Ethernet bridging elements and Lossless Ethernet MACs connections are allowed.

VA_Ports and VE_Ports may coexist over the same FCF-MAC. VF_Ports coexist with neither VA_Ports nor VE_Ports over the same FCF-MAC. An FCF-MAC supporting the instantiation of VA_Ports and VE_Ports is referred to as a VA_Port/VE_port capable FCF-MAC.

The FCoE Controller is the functional entity that performs the FCoE Initialization Protocol (FIP) and instantiates or de-instantiates VA_Ports, VE_Ports, or VF_Ports, as needed.

For a VA_Port/VE_Port capable FCF-MAC, the FCoE Controller:

- a) optionally performs the FIP VLAN discovery protocol to discover FCoE VLANs;
- b) discovers other VA_Port/VE_Port capable FCF-MACs, VE_Port capable FCF-MACs, and VA_Port capable FDF-MACs connected to the same Lossless Ethernet network using the FIP discovery protocol;
- c) instantiates a VE_Port/FCoE_LEP pair on successful completion of each FIP ELP Exchange with a remote VA_Port/VE_Port capable FCF-MAC or a remote VE_Port capable FCF-MAC;
- d) de-instantiates a VE_Port/FCoE_LEP pair on receiving a FIP Clear Virtual Link request;
- e) instantiates a VA_Port/FCoE_LEP pair on successful completion of each FIP ELP Exchange with a remote VA_Port capable FDF-MAC;
- f) de-instantiates a VA_Port/FCoE_LEP pair on receiving a FIP Clear Virtual Link request;

- g) monitors the status of the instantiated VE_Port/FCoE_LEP and VA_Port/FCoE_LEP pairs;
- h) initiates FIP Clear Virtual Link requests as needed to terminate Virtual Links to other VE_Ports or VA_Ports;
- i) transmits periodic FIP Discovery Advertisements to the All-FCF-MACs address every FKA_ADV_PERIOD; and
- j) monitors the status of remote VE_Ports and VA_Ports by maintaining timers and verifying that periodic FIP Discovery Advertisements are received within every FKA_ADV_PERIOD.

The FCoE_LEP is the functional entity performing the encapsulation of FC frames into FCoE frames in transmission and the decapsulation of FCoE frames into FC frames in reception. An FCoE_LEP operates according to the MAC address of the local link end-point and the MAC address of the remote link end-point. When encapsulating FC frames into FCoE frames, the MAC address of the local link end-point shall be used as source address and the MAC address of the remote link end-point shall be used as destination address of the generated FCoE frame. When decapsulating FC frames from FCoE frames, the FCoE_LEP shall verify that the destination address of the received FCoE frame is equal to the MAC address of the local link end-point and shall verify that the source address of the received FCoE frame is equal to the MAC address of the remote link end-point. If either check fails the FCoE frame shall be discarded.

For a VA_Port/VE_Port capable FCF-MAC, the MAC address of the local link end-point is the FCF-MAC address and the MAC address of the remote link end-point is the MAC address of the remote FCF-MAC or FDF-MAC with which a FIP ELP Exchange has been successfully completed.

A VA_Port is an instance of the FC-2V sublevel of Fibre Channel that is dynamically instantiated together with its FCoE_LEP on successful completion of a FIP ELP Exchange, according to the rules specified in 1.6.2. A VA_Port receives FC frames from the FC Switching Element and sends them to its FCoE_LEP for encapsulation and transmission over the Lossless Ethernet network. In a similar way, a VA_Port sends FC frames received from its FCoE_LEP to the FC Switching element. A VA_Port is uniquely identified by an A_Port_Name Name_Identifier and is addressed by the A_Port Controller address identifier (i.e., FFFFF9h).

1.4 FDF Functional Model

Figure 5 shows the functional model of an FDF, where the bracketed functional components are optional. An FDF is functionally composed of an FCDF Switching Element (see FC-SW-6) with at least one Lossless Ethernet MAC (FDF-MAC). Each FDF-MAC shall be coupled with an FCoE Controller function. Each FDF-MAC may be coupled with a Lossless Ethernet bridging element. An FDF supports the instantiation of VA_Ports or VF_Ports over its FDF-MACs. The FCDF Switching Element may be coupled with a Fibre Channel Fabric interface, providing native A_Port and F_Port connectiv-

ity. An FDF forwards FCoE frames addressed to one of its FDF-MACs based on the D_ID of the encapsulated FC frames.

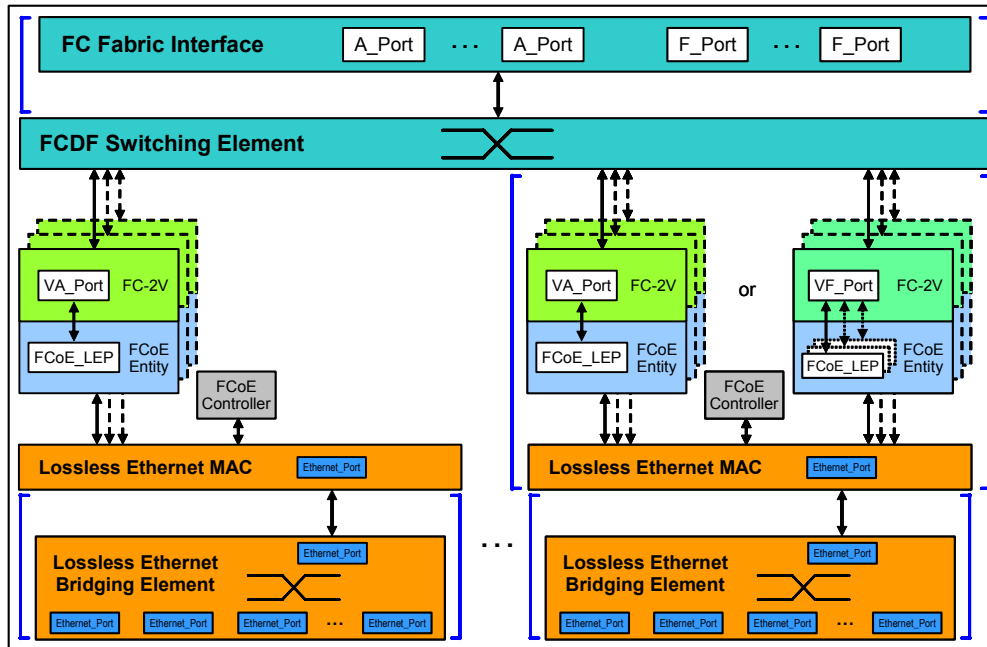


Figure 5 – FDF Functional Model

NOTE 2 – Other combinations of Lossless Ethernet bridging elements and Lossless Ethernet MACs connections are allowed.

An FDF does not support VE_Ports. VA_Ports and VF_Ports do not coexist over the same FDF-MAC. An FDF-MAC supporting the instantiation of VA_Ports is referred to as a VA_Port capable FDF-MAC. An FDF-MAC supporting the instantiation of VF_Ports is referred to as a VF_Port capable FDF-MAC. A VF_Port capable FDF-MAC behaves as a VF_Port capable FCF-MAC. An FDF shall support at least one VA_Port capable FDF-MAC.

The FCoE Controller is the functional entity that performs the FCoE Initialization Protocol (FIP) and instantiates or de-instantiates VA_Ports or VF_Ports, as needed.

For a VA_Port capable FDF-MAC, the FCoE Controller:

- a) optionally performs the FIP VLAN discovery protocol to discover FCoE VLANs;
- b) discovers other VA_Port capable FDF-MACs and VA_Port/VE_Port capable FCF-MACs connected to the same Lossless Ethernet network using the FIP discovery protocol;
- c) instantiates a VA_Port/FCoE_LEP pair on successful completion of each FIP ELP Exchange with a remote VA_Port capable FDF-MAC or VA_Port/VE_Port capable FCF-MAC;
- d) de-instantiates a VA_Port/FCoE_LEP pair on receiving a FIP Clear Virtual Link request;
- e) monitors the status of the instantiated VA_Port/FCoE_LEP pairs;
- f) initiates FIP Clear Virtual Link requests as needed to terminate Virtual Links to other VA_Ports;
- g) transmits periodic FIP Discovery Advertisements to the All-FCF-MACs address every FKA_ADV_PERIOD; and
- h) monitors the status of remote VA_Ports by maintaining timers and verifying that periodic FIP Discovery Advertisements are received within every FKA_ADV_PERIOD.

A VA_Port/VE_Port capable FCF-MAC shall instantiate VE_Port to VE_Port Virtual Links with other VA_Port/VE_Port capable FCF-MACs and VE_Port capable FCF-MACs discovered by FIP discovery on the Lossless Ethernet network. VE_Port to VE_Port Virtual Links with VA_Port/VE_Port capable FCF-MACs belonging to the other Controlling FCF of the Primary/Secondary pair defining the Distributed FCF are used for the Redundancy protocol of the Distributed FCF (see T11/11-224v0). VE_Port to VE_Port Virtual Links are established according to the normal ELP rules (see FC-SW-5).

A VA_Port/VE_Port capable FCF-MAC instantiates VA_Port to VA_Port Virtual Links with VA_Port capable FDF-MACs through a FIP ELP Exchange. These VA_Port to VA_Port Virtual Links are instantiated when the FDF/FCDF flag is set to one in the FIP ELP SW_ACC. A VA_Port/VE_Port capable FCF-MAC shall reject a received FIP ELP Request with the FDF/FCDF flag set to one with Reason Code 'Protocol Error' and Reason Code Explanation 'Invalid Request'.

When operational (i.e., when the Controlling FCF is in state P2 or S2 of T11/11-224v0), a VA_Port/VE_Port capable FCF-MAC shall instantiate VA_Port to VA_Port Virtual Links with VA_Port capable FDF-MACs belonging to FDFs that are part of the Distributed FCF's FDF Set and discovered by FIP discovery on the Lossless Ethernet network. In this case, the FIP ELP Exchange shall be initiated by the VA_Port/VE_Port capable FCF-MAC and the VA_Port capable FDF-MAC shall reply, irrespective of the value of the Switch_Name field in the ELP payload. A VA_Port/VE_Port capable FCF-MAC shall not establish VA_Port to VA_Port Virtual Links with VA_Port capable FDF-MACs belonging to FDFs that are not part of the Distributed FCF's FDF Set. A VA_Port capable FDF-MAC shall not initiate a FIP ELP Exchange with a VA_Port/VE_Port capable FCF-MAC.

A VA_Port capable FDF-MAC instantiates VA_Port to VA_Port Virtual Links with other FDF-MACs through a FIP ELP Exchange. These VA_Port to VA_Port Virtual Links are instantiated when the FDF/FCDF flag is set to one in both FIP ELP Request and FIP ELP SW_ACC.

A VA_Port capable FDF-MAC shall initiate a FIP ELP Exchange with a discovered VA_Port capable FDF-MAC only if:

- a) it has already at least a VA_Port to VA_Port Virtual Link with the Primary Controlling FCF or another FDF;
- b) it has received from the Primary Controlling FCF the Distributed FCF's FDF Set (see T11/11-225v0); and
- c) the discovered FDF-MAC belongs to an FDF in the Distributed FCF's FDF Set.

A VA_Port capable FDF-MAC shall not initiate a FIP ELP Exchange with a discovered VA_Port capable FDF-MAC if the discovered FDF-MAC does not belong to an FDF in the FDF Set. A VA_Port capable FDF-MAC that has not initiated a FIP ELP Exchange shall reply to a received FIP ELP Request, irrespective of the value of the Switch_Name field in the ELP payload. A VA_Port capable FDF-MAC that has initiated a FIP ELP Exchange (i.e., sent a FIP ELP Request) shall reply to a received FIP ELP Request according to the normal ELP rules (see FC-SW-5).

NOTE 3 – These rules enable an ordered establishments of VA_Port to VA_Port Virtual Links from the Controlling FCF(s) to the peripheral FDFs in a Distributed FCF with cascaded FDFs.

An FDF does not establish VE_Port to VE_Port Virtual Links, therefore an FDF-MAC shall reject a received FIP ELP Request with both Controlling FCF/Switch flag and FDF/FCDF flag set to zero (i.e., a FIP ELP Request coming from an FCF that is not a Controlling FCF) with Reason Code 'Protocol Error' and Reason Code Explanation 'Invalid Request'. An FDF-MAC shall also reject a received FIP ELP Request coming from a Controlling FCF other than the Controlling FCFs that define its Distributed FCF, with Reason Code 'Logical Error' and Reason Code Explanation 'Not Authorized'.

1.6.3 FCoE Virtual Link Maintenance

VA_Port to VA_Port Virtual Link maintenance is performed as for VE_Port to VE_Port Virtual Links.

1.6.4 Distributed FCF Operations

In a Distributed FCF, the Primary Controlling FCF defines the routes for the FDF topology and performs N_Port_ID allocations and deallocations for all its controlled FDFs. The two Controlling FCFs keep their state synchronized.

When becoming operational (i.e., when in state P2 or S2 of T11/11-224v0), a Controlling FCF performs FIP discovery as described in 1.6.1. A Controlling FCF instantiates VA_Port to VA_Port Virtual Links with the discovered FDFs that are directly reachable and are part of its FDF Set by initiating FIP ELP Exchanges, as described in 1.6.2.

Upon instantiating a VA_Port to VA_Port Virtual Link with an FDF, the Primary Controlling FCF shall initiate an FDRN Exchange (see T11/11-225v0) describing that Virtual Link with the Secondary Controlling FCF to keep the state synchronized. Upon completion of this FDRN Exchange, the Primary Controlling FCF shall provide to that FDF the Distributed FCF Membership information through a DFMD Exchange (see T11/11-225v0). At this point the instantiated VA_Port to VA_Port Virtual Link becomes part of the Distributed FCF internal topology (i.e., the set of VA_Port to VA_Port Virtual Links internal to the Distributed FCF). The Primary Controlling FCF shall recompute the N_Port_ID routes and distribute them to each FDF belonging to the Distributed FCF through NPRD Exchanges (see T11/11-225v0).

Upon deinstantiating a VA_Port to VA_Port Virtual Link with an FDF, the Primary Controlling FCF shall initiate an FDUN Exchange (see T11/11-225v0) describing that Virtual Link with the Secondary Controlling FCF to keep the state synchronized. Upon completion of this FDUN Exchange, the Primary Controlling FCF shall recompute the N_Port_ID routes and distribute them to each FDF belonging to the Distributed FCF through NPRD Exchanges.

When becoming operational, an FDF participates in FIP discovery as described in 1.6.1. It then waits for a Controlling FCF or another FDF to initiate a FIP ELP Exchange with it, in order to set up a VA_Port to VA_Port Virtual Link. Upon completing the DFMD Exchange with the Primary Controlling FCF, the FDF becomes able to initiate FIP ELP Requests to instantiate other VA_Port to VA_Port Virtual Links with other FDFs. Upon completing the NPRD Exchange with the Primary Controlling FCF, an FDF becomes able to set up proper forwarding tables to forward FCoE frames inside and outside the Distributed FCF. At this point the FDF enables its VF_Port capable FCF-MACs for FIP discovery and FIP logins from ENodes.

Upon instantiating a VA_Port to VA_Port Virtual Link with another FDF or with the Secondary Controlling FCF, an FDF shall perform a FDRN Exchange with the Primary Controlling FCF to inform it of the new Virtual Link. Upon completing a FDRN Exchange with an FDF, the Primary Controlling FCF shall initiate another FDRN Exchange with the same parameters with the Secondary Controlling FCF to keep the state synchronized. Upon completion of this FDRN Exchange, the Primary Controlling FCF shall recompute the N_Port_ID routes and distribute them to each FDF belonging to the Distributed FCF through NPRD Exchanges.

NOTE 4 – A VA_Port to VA_Port Virtual Link with the Secondary Controlling FCF may be instantiated before the VA_Port to VA_Port Virtual Link with the Primary Controlling FCF. The FDF recognizes the Primary Controlling FCF because it is the one from which it receives the DFMD Request. In this case, the FDF initiates with the Primary Controlling FCF the FDRN Exchange describing the Virtual Link with the Secondary Controlling FCF upon completing the DFMD Exchange.

Upon deinstantiating a VA_Port to VA_Port Virtual Link with another FDF or with the Secondary Controlling FCF, an FDF shall perform a FDUN Exchange with the Primary Controlling FCF to inform it of the disappeared Virtual Link. Upon completing a FDUN Exchange with an FDF, the Primary Controlling FCF shall initiate another FDUN Exchange with the same parameters with the Secondary Controlling FCF to keep the state synchronized. Upon completion of this FDUN Exchange, the Primary Controlling FCF shall recompute the N_Port_ID routes and distribute them to each FDF belonging to the Distributed FCF through NPRD Exchanges.

Upon receiving on a VF_Port capable FDF-MAC a FIP FLOGI Request or a FIP NPIV FDISC Request from an ENode MAC, an FDF shall send a VNRN Request (see T11/11-225v0) to the Primary Controlling FCF to inform it of the newly reachable VN_Port. If the Primary Controlling FCF rejects the VNRN Request, the FDF shall also reject the FIP FLOGI Request or FIP NPIV FDISC Request. If the Primary Controlling FCF accepts the VNRN Request, it shall allocate to the newly reachable VN_Port an N_Port_ID from the Virtual Domain_ID, generate appropriate RSCN(s), and update the Fibre Channel Name Server. The Primary Controlling FCF shall also recompute the Zoning ACLs for the affected N_Port_IDs. The Primary Controlling FCF shall then distribute this information to the Secondary Controlling FCF and to each FDF belonging to the Distributed FCF through an NPZD Exchange indicating an N_Port_ID allocation (see T11/11-225v0). The NPZD Requests sent to the Secondary Controlling FCF and to the FDF that sent the VNRN Request shall include the FLOGI / NPIV FDISC LS_ACC Parameters (see T11/11-225v0); the NPZD Requests sent to the other FDFs shall not include them. The NPZD Requests sent to the Secondary Controlling FCF shall carry zero Peering Entries. Upon receiving the NPZD Request including the FLOGI / NPIV FDISC LS_ACC Parameters, the FDF shall accept the FIP FLOGI Request or a FIP NPIV FDISC Request and complete the instantiation of the VN_Port to VF_Port Virtual Link.

When a VN_Port to VF_Port Virtual Link is deinstantiated, an FDF shall perform a VNUN Exchange (see T11/11-225v0) with the Primary Controlling FCF to inform it that the VN_Port is now unreachable. Upon completing a VNUN Exchange, the Primary Controlling FCF shall deallocate the N_Port_ID previously assigned to that VN_Port, generate appropriate RSCN(s), and update the Fibre Channel Name Server. The Primary Controlling FCF shall also recompute the Zoning ACLs for the affected N_Port_IDs. The Primary Controlling FCF shall then distribute this information to the Secondary Controlling FCF and to each FDF belonging to the Distributed FCF through NPZD Exchanges indicating an N_Port_ID deallocation (see T11/11-225v0). These NPZD Requests shall not include the FLOGI / NPIV FDISC LS_ACC Parameters (see T11/11-225v0).

When a new Zone Set is activated in the Fabric, the Primary Controlling FCF shall recompute the Zoning ACLs for all N_Port_IDs allocated in the Virtual Domain_ID and distribute them to the FDFs of the Distributed FCF through AZAD Exchanges (see T11/11-225v0).

Upon receiving on a VF_Port capable FCF-MAC a FIP FLOGI Request or a FIP NPIV FDISC Request from an ENode MAC, a Controlling FCF shall allocate to the newly reachable VN_Port an N_Port_ID from its Domain_ID (i.e., not from the Virtual Domain_ID) if it accepts the received FIP FLOGI or FIP NPIV FDISC Request.