



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

« [Connectivity changes for March 2011](#) | [Main](#) | [Connectivity changes for April 2011](#) »

03/27/2011

FCoE Multihop and VE_Ports

In the April EMC Support Matrix, EMC will be posting support for NX-OS 5.0(2)N2(1) for the Nexus 5010, 5020 and 5548. In addition, we will also now support the creation of Virtual E_Ports (VE_Ports) on these products. This post discusses what VE_Ports are, the protocol used to instantiate them and the exact configuration steps required to create them.

VE_Port overview

VE_Ports allow for the formation of FCoE ISLs and the creation of an all FCoE fabric or more precisely a Multihop FCoE fabric. As you will see, once a link between two FCFs has been established using FIP, VE_Ports / FCoE ISLs are initialized using the same protocol that is used to initialize E_Ports / native FC ISLs. Just in case you are interested, the FIP protocol used to establish FCF to FCF connections is described in my previous post, "[FIP, FIP Snooping Bridges, and FCFs \(Part 1 - "FIP", the FCoE Initialization Protocol\)](#)".

From a logical connectivity point of view, an all FCoE fabric is identical to an FC Fabric and it supports all of the same functionality such as zoning, a distributed name server as well as support for RSCNs. As a result, the same types of scalability limits apply to both FC and FCoE Fabrics such as the maximum number of; hops, VN_Ports and Domains. The FC scalability limits have been posted in the EMC Support Matrix for years and are currently no more than 5 hops, no more than 6000 N_Ports and no more than 55 Domains. With an all FCoE fabric, I feel comfortable with 3 hops (because I've tested it), and have no reason to believe that the number of VN_Ports or Domains will need to be limited to less than what is supported in native FC.

From a physical connectivity point of view, the connectivity options currently supported for an all FCoE Fabric are shown in the following diagram. Each of the connectivity options are explained in detail below the diagram.



[Erik Smith](#)

ARCHIVES

[September 2022](#)

[August 2020](#)

[April 2019](#)

[December 2017](#)

[November 2017](#)

[July 2017](#)

[April 2017](#)

[September 2016](#)

[July 2016](#)

[February 2016](#)

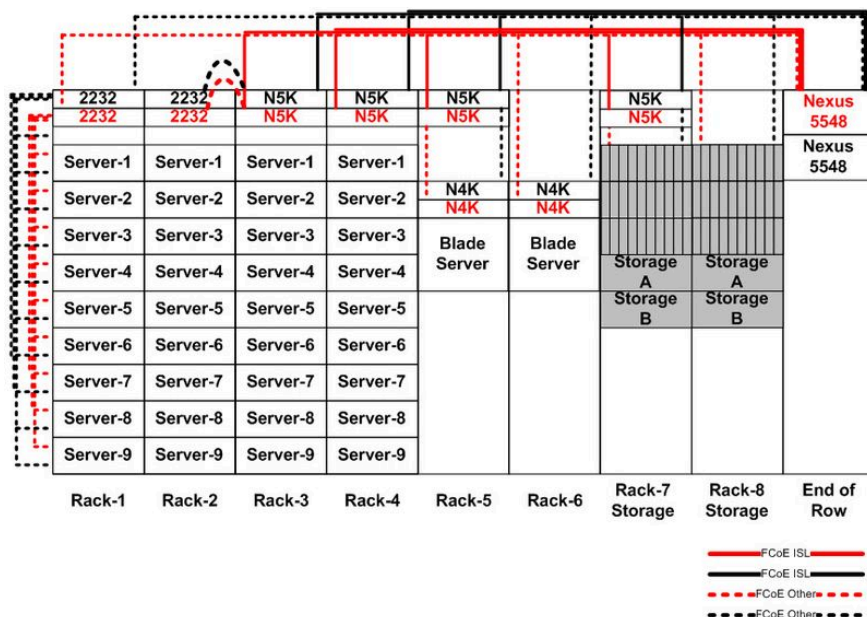
SEARCH

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.



For the sake of clarity, the diagram shows the physical connectivity options that can be used within a row of equipment racks. The diagram is not intended to indicate a limitation in the types of topologies that are supported; rather it is just intended to help highlight all of the different possibilities. Each Rack is described in detail below:

End of Row- The End of Row (EoR) cabinet contains the aggregation layer switches. In order to maintain the highly available characteristics of FC, two 5548s are shown and they have not been connected together via FCoE ISLs. This allows for the two fabrics to remain logically isolated. The 5548s may be part of the same VPC Domain.

The 5548s in the End of Row cabinet will support:

1. VE_Port (FCoE ISL) connections from other Nexus 5000s,
2. VF_Port (ENode) connections from FCoE initiators and targets,
3. native FC E_Ports (ISLs) to other FC capable Cisco switches
4. F_Port (FC initiator and target) connections from FC initiators and targets.
5. N_Port connections to FC capable Cisco/Brocade/QLogic FC switches
6. uplinks from an FSB, and
7. Fabric Ports to connect to the 2232 FEX module.

Rack 7 & 8 Storage- Storage ports (either FC or FCoE) can be connected to a Top of Rack (ToR) switch or directly back to the 5548 in the End of Row rack (EoR). Connecting a storage port directly to the top of rack switch in a given rack makes sense if the storage port is accessed primarily by servers residing in that rack. If a storage port is going to be accessed by many servers in different racks, the optimal placement of the storage ports may be on the 5548 in the EoR rack.

Rack 5 & 6 FIP Snooping Bridge (FSB)- FIP Snooping Bridges such as the Cisco/IBM 4001i can either be connected to the Nexus 5000 at the top of the rack or connected to the Nexus 5548 at the EoR.

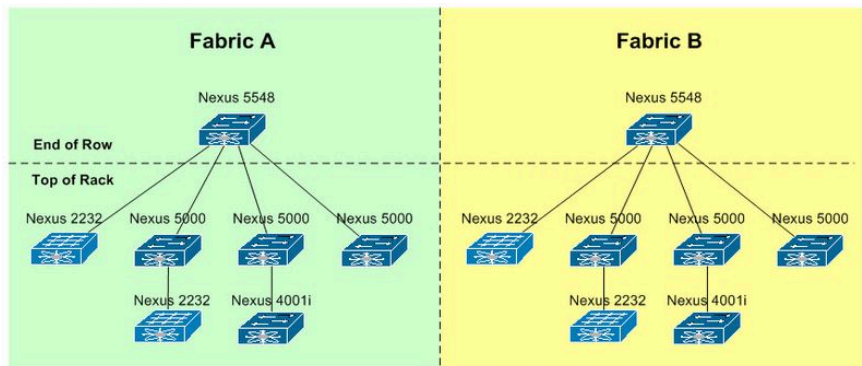
Rack 4 Nexus 5000 (VE_Ports)- The Nexus 5000 can be connected to the 5548 in the EoR via VE_Ports while running in FC-SW mode over FCoE or over FC

while running NPV and FC-SW modes. Currently you cannot utilize FCoE for uplinks to the 5548 at the EoR while running in NPV mode.

Rack 2 & 3 Nexus 5000 (VE_Ports)/Nexus 2232 (Fabric Ports)- The Nexus 2232 can be connected to a Nexus 5000 via Fabric ports and then the Nexus 5000 can be connected to the Nexus 5548 at the EoR via E_Ports.

Rack 1 Nexus 2232 (Fabric ports) - The Nexus 2232 at the ToR can be connected to the Nexus 5548 at the EoR via Fabric ports.

A logical representation of the physical topology is shown below.



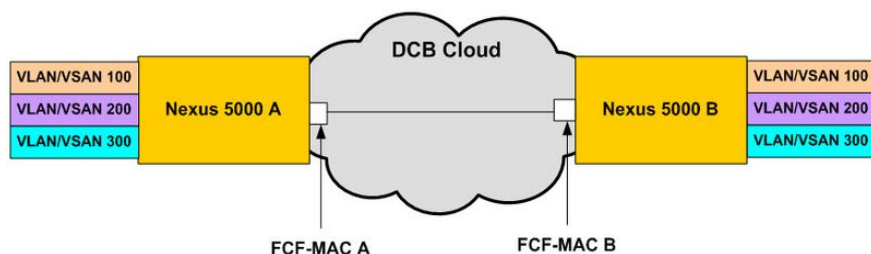
Although it isn't shown in the logical topology above, the FCoE ISLs must be on physically separate links rather than trunked on the ethernet uplinks used to carry non-FCoE traffic. This is not a requirement from an FCoE protocol perspective but it is required when using the current version of NX-OS. This requirement may eventually be removed.

An important point to note in the above diagram is that from an FC logical point of view, a Fabric A / Fabric B topology has been maintained. If the Customer desires, the ToR switches could be connected together to allow for the use of vPC and active/active NIC Teaming /Bonding. Although this breaks the "Air gap" separation requirement (more on this later), the FC portion of the network will remain logically isolated.

FCoE VE Port Virtual Link Instantiation

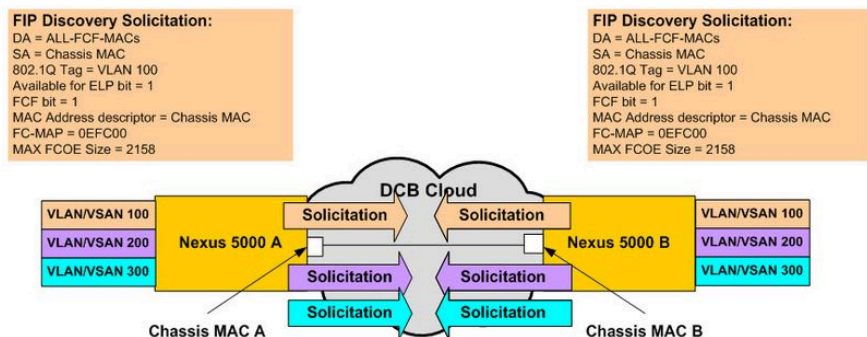
Note: This section describes the FCoE VE_Port Virtual Link Instantiation process currently being used between Cisco Nexus products. The reason for this is that Cisco is the only FCF vendor currently providing this functionality. When other vendors provide VE_Port functionality, this section will be updated to show the differences (if any).

The Virtual Links used to support the instantiation of Virtual E_Ports (VE_Ports) are created in a manner that is similar to the process used to instantiate a virtual link that supports the instantiation of a Virtual F_Port (VF_Port) as described in my previous post "[FIP, FIP Snooping Bridges, and FCFs \(Part 1 - "FIP", the FCoE Initialization Protocol\)](#)". One major difference is that FIP VLAN Discovery is not used. For the sake of this example, the following topology will be used.



FIP Discovery Solicitation

The process starts with both sides of the link transmitting DCBX frames. Once the DCBX parameters have been exchanged, both sides will transmit Solicitations on every VLAN that FCoE has been allowed on. In this example, we will assume that the Ethernet / vFC interface on the Nexus has been configured to allow all VLANs / VSANs. In the diagram below, only the details for the Solicitation frames for VLAN/VSAN 100 are shown. The Solicitations for VLAN/VSAN 200 and 300 would contain similar information.



The Solicitations are transmitted to the multicast Destination Address (DA) of ALL-FCF-MACs. The SA of these frames is the Chassis MAC of the Nexus. The 802.1Q tag will contain the VLAN that the Solicitation is being performed on.

The Available for ELP bit will need to be set to one in order for FIP to proceed to the next phase (FIP Advertisement).

The FCF bit indicates that the Frame was transmitted by an FCF.

The FC-MAP is checked by both sides to ensure that this value matches. If it does not, FIP will not proceed to the next phase and the virtual link will not be instantiated. The FC-MAP value prevents unintentional FC Fabric merges and should be administratively set to a value other than the default if you have multiple FCoE Fabrics in the same data center and you do not want an accidental connection between two FCFs to result in a fabric merge.

The Max FCoE Size field is set to the maximum size FCoE Frame supported by each side.

FCoE Discovery Advertisement

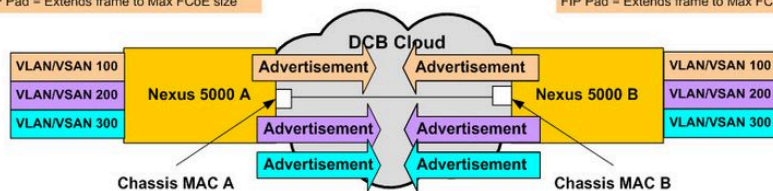
If the Available for ELP bit and FC-MAP values match on both sides of the link, the FIP process will continue with both sides transmitting a Discovery Advertisement.

FIP Discovery Advertisement:

DA = ALL-FCF-MACs
 SA = Chassis MAC
 802.1Q Tag = VLAN 100
 Available for ELP bit = 1
 FCF bit = 1
 MAC Address descriptor = Chassis MAC
 FC-MAP = 0EFC00
 FIP Pad = Extends frame to Max FCoE size

FIP Discovery Advertisement:

DA = ALL-FCF-MACs
 SA = Chassis MAC
 802.1Q Tag = VLAN 100
 Available for ELP bit = 1
 FCF bit = 1
 MAC Address descriptor = Chassis MAC
 FC-MAP = 0EFC00
 FIP Pad = Extends frame to Max FCoE size

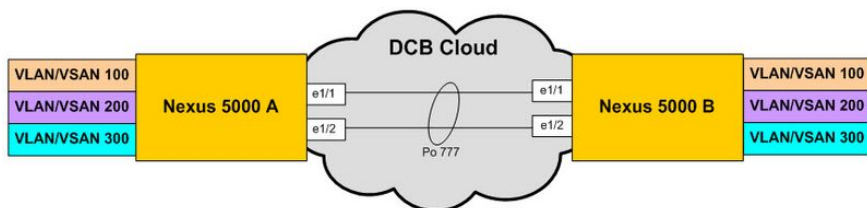


The information contained within the FIP Discovery Advertisement is similar to what is contained in the Solicitation. One difference worth mentioning is that the Advertisement is padded so that the Frame is equal to Max FCoE size. The presence of this padding allows both ends to validate that the network infrastructure between the two VE_Ports is capable of supporting mini Jumbo frames.

Once both sides have received and validated the information contained within the Advertisements, the Virtual link can be instantiated by both sides transmitting ELP on each VLAN/VSAN that supports FCoE. The frames exchanged to complete the VE_Port instantiation are practically identical to those frames used to initialize an FC E_Port and will not be repeated here. For more information, refer to the [Networked Storage Concepts and Protocols TechBook](#) for additional information.

Configuring Cisco VE Ports

VE_Ports allow for two Cisco Nexus 5000 products to be connected together and merged into the same fabric using FCoE links. VE_Ports can be created as either individual links or as shown in the following example configured to use a Port channel. The topology used in this example is shown below.



Setup steps:

On Nexus 5000 A:

1. Create the appropriate VSANs

```
NEXUS5000A# config t
```

Enter configuration commands, one per line. End with CNTL/Z.

```
NEXUS5000A(config)# vsan database
```

```
NEXUS5000A(config-vsan-db)# vsan 200
```

```
NEXUS5000A(config-vsan-db)# vsan 300
```

```
NEXUS5000A(config-vsan-db)# vsan 400
```

```
NEXUS5000A(config-vsan-db)# exit
```

```
NEXUS5000A(config)#
```

2. Create the appropriate VLANs and associate them with the appropriate VSAN.

```
NEXUS5000A(config)# vlan 200
NEXUS5000A(config-vlan)# fcoe vsan 200
NEXUS5000A(config-vlan)# vlan 300
NEXUS5000A(config-vlan)# fcoe vsan 300
NEXUS5000A(config-vlan)# vlan 400
NEXUS5000A(config-vlan)# fcoe vsan 400
NEXUS5000A(config-vlan)# exit
NEXUS5000A(config)#
```

3. Create the port-channel that will be used to connect the two switches.

```
NEXUS5000A(config)# interface port-channel 777
NEXUS5000A(config-if)#switchport mode trunk
NEXUS5000A(config-if)# exit
NEXUS5000A(config)#
```

4. Create the vfc interface that will be associated with the port-channel, specify it as an E_Port and allow the appropriate VSANs.

```
NEXUS5000A(config)# interface vfc777
NEXUS5000A(config-if)# bind interface port-channel777
NEXUS5000A(config-if)# switchport mode E
NEXUS5000A(config-if)# no shutdown
NEXUS5000A(config-if)# exit
NEXUS5000A(config)#
```

5. Add the ethernet interfaces to the port-channel

```
NEXUS5000A(config-if)# int e1/1-2
NEXUS5000A(config-if-range)# switchport mode trunk
NEXUS5000A(config-if-range)# channel-group 777 mode active
NEXUS5000A(config-if-range)# exit
NEXUS5000A(config)#
```

6. Repeat steps 1 - 5 on Nexus 5000 B.

7. At this point the VE_Port port-channel should initialize and can be viewed using the **show fcoe database** command.

```
NEXUS5000A(config)# show fcoe database
```

```
-----
INTERFACE    FCID      PORT NAME      MAC ADDRESS
-----
...

```

VE Ports:

```

-----
INTERFACE      MAC ADDRESS      VSAN
-----
vfc777          00:05:73:af:13:68      200
vfc777          00:05:73:af:13:68      300
vfc777          00:05:73:af:13:68      400

```

NEXUS5000A(config)#

Afterword

Last Sunday, while I was in the process of writing the technical content of this post, a friend of mine from Cisco sent me an asynchronous notification that J Metz had just posted a blog regarding [FCoE Multihop](#). My initial response was “Darn!” (or some variant thereof), followed by “he beat me to it!” Well, here we are a week later and I’m finally getting around to posting my FCoE multihop post. For the sake of transparency, let me state that because I wanted to avoid contaminating my thought process, this morning was the first time I actually read through the entirety of J’s post and the comments that follow. For the most part, I found myself agreeing with J and Brad, but I understand where Brook is coming from and I think he raises some valid points that need to be seriously considered. One point that I want to stress is this, we don’t care which protocol or topology you use to connect to our storage! We care very much that whatever protocol or topology you do use, you are able to design, implement and maintain it with as little effort (or drama) as possible. As long as this is possible we will strive to support whatever protocols and topology our customers are asking for...

Thanks for reading!

Posted at 10:16 PM | [Permalink](#) Reblog (0)

Comments

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



Hi Eric,

Thank you for such a detailed description, but I am in a fix here. I followed the exact steps given by you for VE port configuration on the cisco nexus. Although on both nexus switches, my vfc's say it is bound to port-channel 777, port is E and mode is TE.

They are just not able to talk with each other. What should be the priority-flow-control settings on the port-channel interface?? I have set them to auto/off/on but no luck. The pre-requisite of identical FC-MAPS on both switches is also followed. Following is the VFC output. Note that on both the switches, I added vfc777 only under vsan 2 in the vsan database.

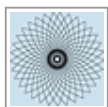
```

switch(config-if)# show interface vfc 777
vfc777 is trunking (Not all VSANs UP on the trunk)
Bound interface is port-channel777
Hardware is Virtual Fibre Channel
Port WWN is 23:08:00:0d:ec:b1:d6:ff
Admin port mode is E, trunk mode is on
snmp link state traps are enabled
Port mode is TE
Port vsan is 2
Trunk vsans (admin allowed and active) (2,10,20,200)
Trunk vsans (up) ( )
Trunk vsans (isolated) ( )
Trunk vsans (initializing) (2,10,20,200)
1 minute input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
1 minute output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
0 frames input, 0 bytes
0 discards, 0 errors
0 frames output, 0 bytes
0 discards, 0 errors
last clearing of "show interface" counters never
Interface last changed at Mon May 9 11:12:26 2011

```

Thanks,
Somu

Posted by: [Account Deleted](#) | [05/09/2011 at 08:23 AM](#)



Hi Somu, sorry to hear you're having a tough time. If you could send me the following output from both switches that you are trying to connect together via the VE_Port, I'll take a look and see if I can figure out where the problem is.

```

show running-config
show int brief
show fcoe database
show int port-channel 777

```

You can email them to me at [\[email protected\]](#)

Posted by: [Erik Smith](#) | [05/09/2011 at 12:48 PM](#)

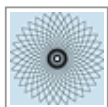


Erik,

Do I understand this correctly that the FCoE traffic is separated from the IP traffic on the ISL link? In short, this is NOT a converged network (FCP + IP over Ethernet), but a separate network physical path for FCoE and a second separate network physical path for IP traffic?

Thanks.

Posted by: [BRCDbrems](#) | [11/09/2011 at 09:05 PM](#)



Hi Brook, you're right. There are good reasons for this but they're hard to articulate without sounding like I've drowned in the Cisco cool-aid! :-)

At first I really didn't like the idea because it's not necessary for any protocol related reasons and it messed with the pretty converged network picture I had in my head at the time. That having been said, the Customers that I've spoken with about this actually like the idea of separating FCoE from non-FCoE at the first FCF. The reason is this approach allows the storage guys to mostly retain control of the SAN while allowing them to take advantage of the economic advantages that FCoE to the ToR allows for.

BTW, I'm looking forward to working on VCS, that should be very interesting as well.

Regards, Erik

Posted by: [Erik Smith](#) | [11/10/2011 at 03:46 PM](#)



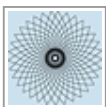
G'day Erik,

Can you confirm if you have tested or implemented 2 x Nexus 5548UP Switches that are connected over distance (40Km) using ISL/E-port's? Let's say 1 N5K is in 1 datacentre and the other is in another datentre. We are also going through a CWDM.

Regards,

Bruce

Posted by: Bruce | [01/31/2012 at 10:39 PM](#)



Hi Bruce, the maximum distance I've tested between 2 5500s was 300m.

That having been said, I believe what you are attempting to do won't work with the 5500s as they don't have enough buffering to allow PFC to function properly and prevent frame loss due to buffer overrun. Basically, their high water mark threshold (the point at which they will transmit PFC frame due to a buffer full condition) does not take into account the extra "buffering" taking place on the link. As a result, they don't reserve enough space for all of the frames in transit and these frames can be dropped under congestion conditions.

I believe the Nexus 7k and the FCoE Blade for the MDS will support up to 40km.

A workaround would be to use FC for the connectivity between the 5548's.

Posted by: [Erik Smith](#) | [02/01/2012 at 08:10 AM](#)



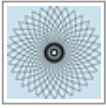
I know this is an old post, but maybe i get lucky....

I have one question: suppose i have a server and a storage fcoe connected to separate 5ks. The 5ks are vonnected by ve ports. The fcoe data packets between the servers would use source and destination mac as the fcmep, wright? Or both servers would use destination mac addrss as the fcf mac? For me the second one seems

more complicated but i can't find an answer...

Thanks for the nice post!

Posted by: Sorin | [06/11/2015 at 01:35 AM](#)



Hi Sorin, to understand MAC Address use with FCoE, it's easiest to keep in mind that FCoE is really just FC operating at the L3 Network layer and using Ethernet at L2. This means that a connection between a server and a storage that crosses two N5ks connected by VE_Ports actually consists of 3 L2 links; one between the Server and the first N5k, the second between the two N5ks and the final one between the N5k and the storage. As with other protocols (e.g., IP) that use Ethernet as a transport, the L2 Addresses (MACs) only have meaning on each L2 link. As a result, they are stripped off by the L3 entities (e.g., FCoE FCF, IP Routers) and new ones are used to transport the L3 protocol to the next L3 entity. All of this means that between the Server and the first N5k, you'll use the Server's FPMA and first FCFs FCF-MAC. Between the two N5ks, you'll use FCF-MACs. And between the N5k and the storage you'll use another FCF-MAC and the Storage ports FPMA. I hope this helps!

Posted by: [Erik Smith](#) | [06/11/2015 at 08:28 AM](#)

The comments to this entry are closed.

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