

Normal Checking Rules Prevent Corruption Scenarios

And a set of natural rules (which would probably be part of an implementation's error handling independent of this presentation)

John L. Hufferd

T11/08-055v1

Purpose of slides

This set of slides describes a set of Rules that should be included in Robust installations and can be used to prevent the administrative errors described in T11/07-558v0, entitled “FCoE: Fabric Crosstalk”

These slides are a results of many conversations with David Black who authored T11/07-558v0. There are three types of issues described in that presentation

It has been agreed that with the Discovery Protocol that we agreed upon and with appropriate rules that scenarios 1 & 2 have been addressed. The 3rd scenario, has been in debate. These slides are intended to show all the tools we have to work with independent of ACLs.

However, use of the appropriate “Best Practices” ACLs would prevent many of the scenarios from ever happening.

Overview

The following slides deal with Scenario 1, 2 and 3 (as seen in T11/07-558v0)

Additional Scenarios have been probed to examine some additional effects of duplicate MAC addresses including:

- when the invalid LAN interconnect is created
- what FCoE connections are allowed
- Etc.

A set of rules and assumptions will be asserted

Those rules will be used in case studies that show the response to the Scenarios' errors

- Several examples are shown to demonstrate how the rules prevent the error conditions
- The Rules will probably be implemented in good FCoE products anyway but need to be identified here to ensure that there are appropriate solutions that provide robustness to the Fabric

Note: A simple ACL with "Eth=FCoE Deny", and "Eth=FIP Deny" on non-connected switch ports is a "Best Practice" that would prevent all of the erroneous cable plugging Scenarios; regardless of that products should be constructed with Rules set forth here just in case an ACL error occurs.



Instantiation of a Vx_Port

In this presentation the term “Instantiation” of a Vx_Port will refer to the process that the FCoE controller (see FCoE Model) goes through within the End Node or FCF to enable a path to be created between the MAC and the upper layers. i. e. The “Instantiation” process establishes:

- The FCoE Entity (LEP – Link End Point) for a VN_Port, a VE_Port, or a VF_Port
- The Vx_Port FC Entity, that handles the L2 FC Frames that arrive at the Vx_Port’s FCoE Entity (LEP)

This instantiation process is done by the FCoE Controller when FLOGI Responses are sent or received, or when normal (not FIP) ELS frames are sent to begin a VE_Port to VE_Port interaction

Knowledge of conflicts that are identified in the “Rules” that follow are often based on knowledge local to the FCoE Controller, the FC Entity, or the FCF itself, however, some are based in knowledge within the Fabric.

Rules for FCoE Entity Instantiation (I1 – I2)

- I1. During the instantiation process of any Vx_Port type, if the MAC address of the Remote MAC, which is part of the instantiation, is equal to a Local MAC address (including MAC Address for ENode FCoE Controller, or VN_Port,)
 - It is an Error → Throw an exception
 - Do not instantiate:
 - FCF should reject FLOGIs, and ELS;
 - End Nodes should not issue FLOGIs (based on info found during Discovery)
 - End Nodes should Logout if sent an FLOGI ACC FIP response (from a request with S_ID=0) with a descriptor that contains a MAC Addr that already exists on the End Node (other than the FCoE Controller MAC Addr)
 - Optionally, for VN_Ports all links on the MAC must be shut down until Admin restarts it

- I2. An FCF should NOT share a MAC between two different port types.

(E.g. VF and VE ports should not share the same MAC)

 - If the FCF equipment permits it, Best Practices should dictate that they not be configured in that manner
 - Any FCF equipment that permits the sharing of a port type should (VF & VE), during the instantiation process, check to see if a potential VN_Port to be assigned, is Equal to any other remote Vx_Port previously instantiated to that FCF's MAC
 - If so, it is an Error → Throw an exception
 - Do not instantiate (reject the FLOGI, or ELS)
 - Shut down the MAC until Admin restarts it
 - Likewise no VE port should be instantiated on an FCF if the remote MAC Address is equal to any other instantiated links' remote MAC Address (with same error processing)

Rules for FCoE Entity Instantiation (I3)

13. If during the instantiation process of a VF_Port on an FCF, the Remote End_Node MAC Address, which is part of the instantiation, is equal to the Remote MAC address of any VE_Port on the entire FCF

OR

If during the instantiation process of a VE_Port on the FCF, the Remote MAC Address, which is attempting to obtain instantiation, is equal to the Remote MAC address of any VF_Port on the entire FCF

- It is an Error → Throw an exception
- Do not instantiate the requesting Vx_Port (reject the FLOGI or ELS)
- Optionally, Shut down the instantiated FCF VE_Port or VF_Port MAC
- Do not permit that Remote MAC Address to be instantiated again until Admin permits

Rules for Accepting a Frame at a MAC from a remote Vx_Port (A1 – A2)

A1. Is the SA of the arriving frame equal to the Remote peer of an instantiated Vx_Port using this MAC?

Yes: Accept the Frame and route it to the appropriate Vx_Port

No:

This is an Error → Throw an Exception

Discard the Frame

Optionally, Shut down the NIC until Admin action?

A2. If not instantiated Vx_Port (just a FCoE Controller) and the frame is not part of FIP:

This is an Error → Throw an Exception

Discard the Frame

Optionally, Shut down the NIC until Admin Action?

Rules for Accepting a Frame at a VF_Port FC Entity or VN_Port FC Entity (FC1 – FC2)

FC1. If a VF_Port FC Entity receives a FC Frame with an S_ID to which it is not instantiated:

It has been misrouted and is an Error → Throw an Exception

Discard the Frame

Optionally, Shut down the NIC until Admin intervention

FC2. If a VN_Port FC Entity receives a FC Frame with FC values -- D_ID, S_ID, OX_ID, Sequence#, etc -- that are not correct:

It has been misrouted and is an Error → Throw an Exception

Discard the Frame

Shut down the NIC until Admin intervention

Validation Rules during Discovery & Login (D1 – D2)

D1. When an End Node is selecting an VF_Port after receiving the advertisements from the FCFs

- It must see if any of the advertised FCF MAC addresses are equal to its own, if so:
 - It is an Error → Throw an Exception
 - Optionally, Shut Down the End Node MAC (including existing logins)
 - For HW FCoE NICs it should not be opened again until Admin reset
 - For Software FCoE the MAC address (or the Virtual MAC) should not be reused until the FCoE Software is restarted, or another MAC address is chosen

D2. After receiving a discovery multicast request, the FCF must determine if it is aware of the MAC address, of the SA, as being the MAC address of any MAC on any FCoE controller of which it is aware, if so:

- It is an Error → Throw an Exception
- Do not return an advertisement

Proposed additional Discovery Action and Checking Rule (D3)

D3. End-Nodes Should listen for “ALL_FCFs” Multicasts

- If the SA of any “ALL_FCFs” Multicast is equal to the MAC Address of the End_Node
 - It is an Error → Throw an Exception
 - Optionally, Shut Down the End Node MAC (including existing logins)
 - For HW FCoE NICs it should not be opened again until Admin reset
 - For Software FCoE the MAC address (or the Virtual MAC) should not be reused until the FCoE Software is restarted, or another MAC address is chosen
- Appropriate for all Addressing Methods, to prevent DoS attacks on the Discovery process

Statements regarding Robustness of the FCoE Implementations

Discovery of MAC Addresses duplication

Within a Single Fabric (and a Fabric Wide discovery management Domain)

- End_Node MAC addresses that are a duplication of FCF MAC addresses, will be found during discovery or instantiation and the Admin will be alerted
 - Because of Rule **D1 & D2**
 - + Discovery prevents Duplication of FCF MAC addresses
 - Because of **Rule I1 & I3**
 - + Instantiation Process prevents Duplicates of FCF MAC addr
- End_Node MAC addresses that are duplicates of other End_Nodes will be discovered
 - Because of **Rule I2** (No 2 VN_Ports with duplicate MAC Addresses can be used with the same FCF Port)
 - Because of **Proposed Rule D3** (End_Node detects dubs because of Multicast)



Corruption Prevention

Frames that are misrouted to the wrong End_Node will be discarded, & an Admin Alert will be issued, because of:

- Rule A1 and A2
 - Prevents corruption by eliminating frames without the right SA or if not yet instantiated
- Rule FC2
 - Prevents corruption by eliminating frames that have the wrong values for the FC part of the frame

Probability (P) of error escape is:

- Virtually Zero with only HW assigned MAC addresses
- Probability of a duplicate Host MAC address that causes a corruption problem (refer to T11/07-691V0)
 - With Rule FC2 only: $5.5 \cdot 10^{-14} \leq P$ (error escape) $\leq 8.8 \cdot 10^{-17}$
 - With Rules A1 & A2 -- Virtually 0

Looping prevention

Frames that are intended for an End Node but are misrouted to a VF_Port will be discarded, along with Admin Alert, because of:

- Rule A1, A2, and FC1
 - Prevents looping by eliminating frames without the right SA at VF_Ports
 - Prevents Looping by not accepting frames if FCF Port not instantiated
 - Prevents Looping by discarding frames at the FCF if not the right S_ID at the VF_Ports
- If the Misrouted Frames can not enter the FCF, it can not loop

Loop prevention via ACLs

All sorts of ways that cables can be plugged wrong

All sorts of worst case timing can occur with mistaken cable plugs

A simpler “Best Practices” approach, is the use of ACLs on all Open Ports, which at a minimum says

- Eth=FCoE Deny and
- Eth=FIP Deny

In other words the Admin needs to take special action to permit FCoE frames to enter that port.

This is normal Admin action when properly plugging in new links
It prevents mistaken cabling events from causing Storage Errors
Maybe an automatic ACL for FCoE sensitive Ethernet switches

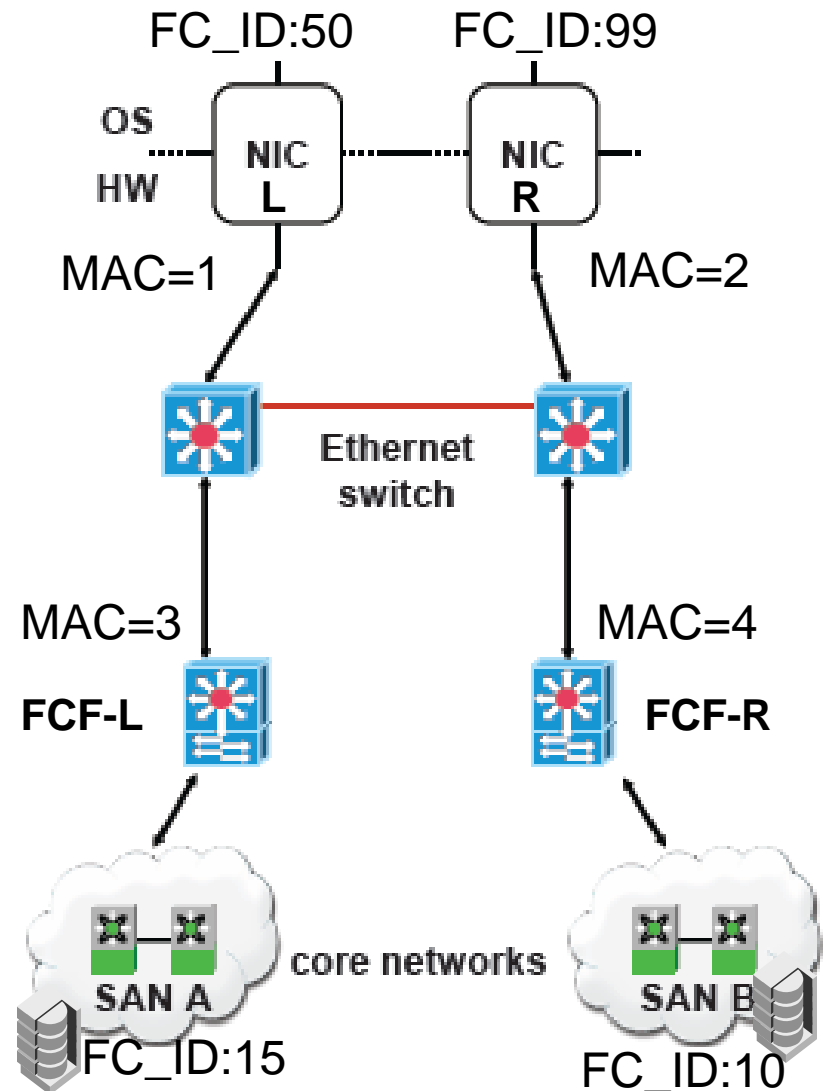


Backup Slides

The details of the Scenarios and Rule Based Solutions

Nomenclature

1. The NIC on the Left has the ID of NIC-L and a MAC Address that may be shown as MAC:1
2. The NIC on the Right has the ID of NIC-R and a MAC Address that may be shown as MAC:2
3. The FCF on the Left has the ID of FCF-L and a MAC Address that may be shown as MAC:3
4. The FCF on the Right has the ID of FCF-R and a MAC Address that may be shown as MAC:4
5. Some other NIC, FCF, or MAC Address will be shown as
 - NIC-x
 - FCF-x
 - MAC:x
6. FC_ID:x will indicate the FC Source or Destination ID (S_ID:x, or D_ID:x)



Problem Scenario 1

Mistake: Cross connect edge switches with VLAN 1 link

- Result: 2 instances of same FCoE MAC on new VLAN 1
 - Spanning tree: new link stays active

Nothing breaks immediately

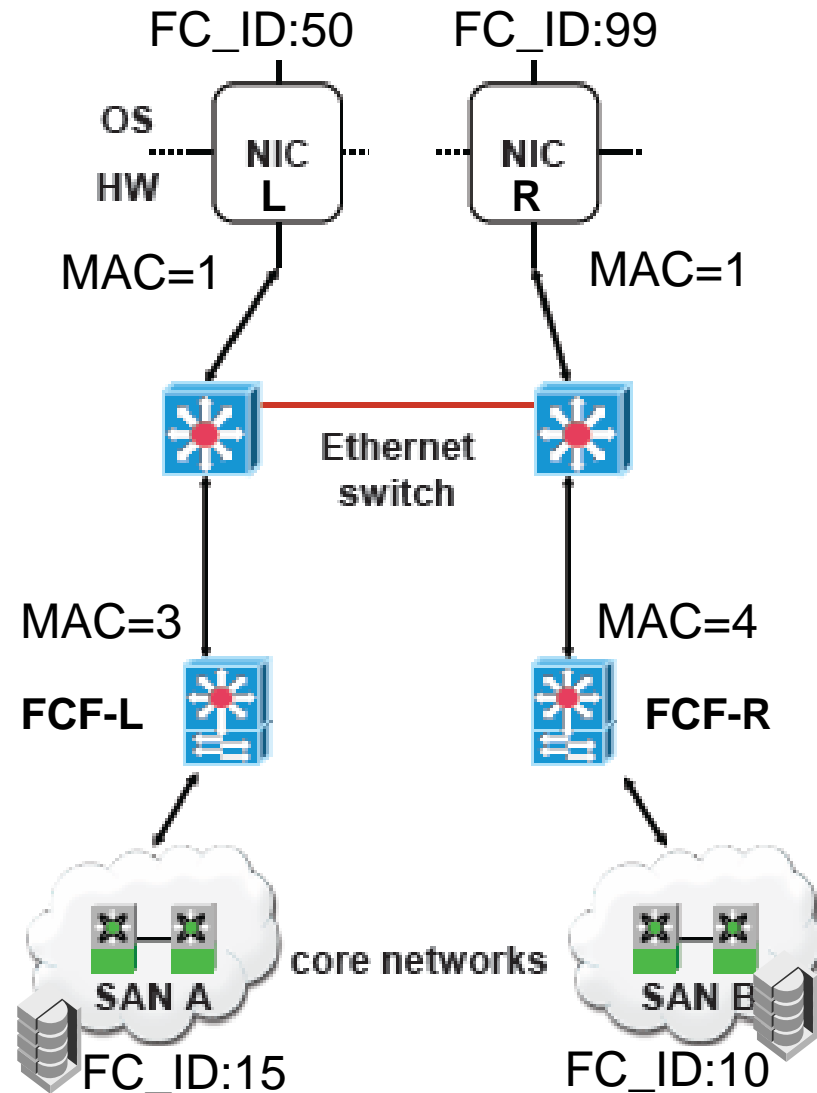
- Both switches learned not to forward MAC-1 traffic to other switch.

What if one of them forgets?

- Or is “helped” to forget by an administrator or injected frame?

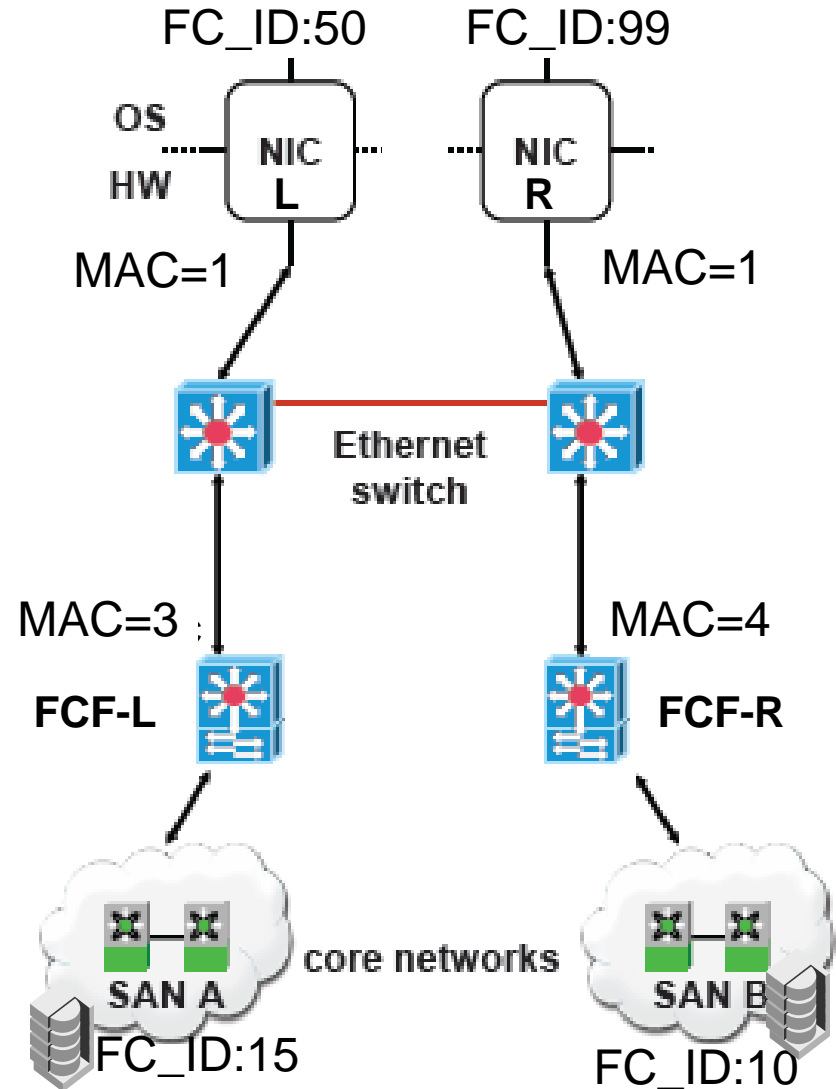
Read data corruption is possible!

- Assume same FC Exchange IDs used with both FCoE N_Ports at same time



Scenario 1: Corruption Avoided with Rule A1

1. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Left FCF (FCF-L) MAC:1 ← → MAC:3
2. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Right FCF (FCF-R) MAC:1 ← → MAC:4
3. I/O is flowing
4. The wire is Plugged wrong and learning is forgotten
5. Outgoing Write data will go to the correct fabric because the DA will be contain the appropriate FCF MAC Addr → no write data corruption
6. Incoming Read data from SAN:A leaves FCF-L with SA=3, DA=1, S_ID:15 & D_ID:50
 - Could be routed to the Right NIC instead of the Left NIC
7. NIC-R will notice that the SA or the D_ID of the incoming frame is not one with which it is instantiated
 - According to [Rule A1](#) the frame must be discarded, an alert issued, (and shutdown the MAC:1?)
8. Data Corruption is avoided for Reads and Writes



What is learned with Scenario 1

Even with duplicate MAC Addresses on the Host systems across different Fabrics and Errant Cabling occurs

It does not matter if the Host MAC addresses are duplicated across different fabrics and then accidentally connected via a cable between the Ethernet Switches --- No corruption will occur if Rule A is used

- I/O headed toward the Storage Controller will not be disrupted by the connection error, because it is still headed toward the MAC address of the appropriate FCF
- I/O headed toward the NIC will be discarded if the frame that is received has the wrong FCF SA – because of [Rule A1](#)

Alerts will be issued to the Admin to request correction, etc.

And the FCoE NIC Optionally, will be shut down until Admin intervention

Following [Rule A1](#) (on End Nodes) will prevent duplicate Host MAC Addresses from causing corruption

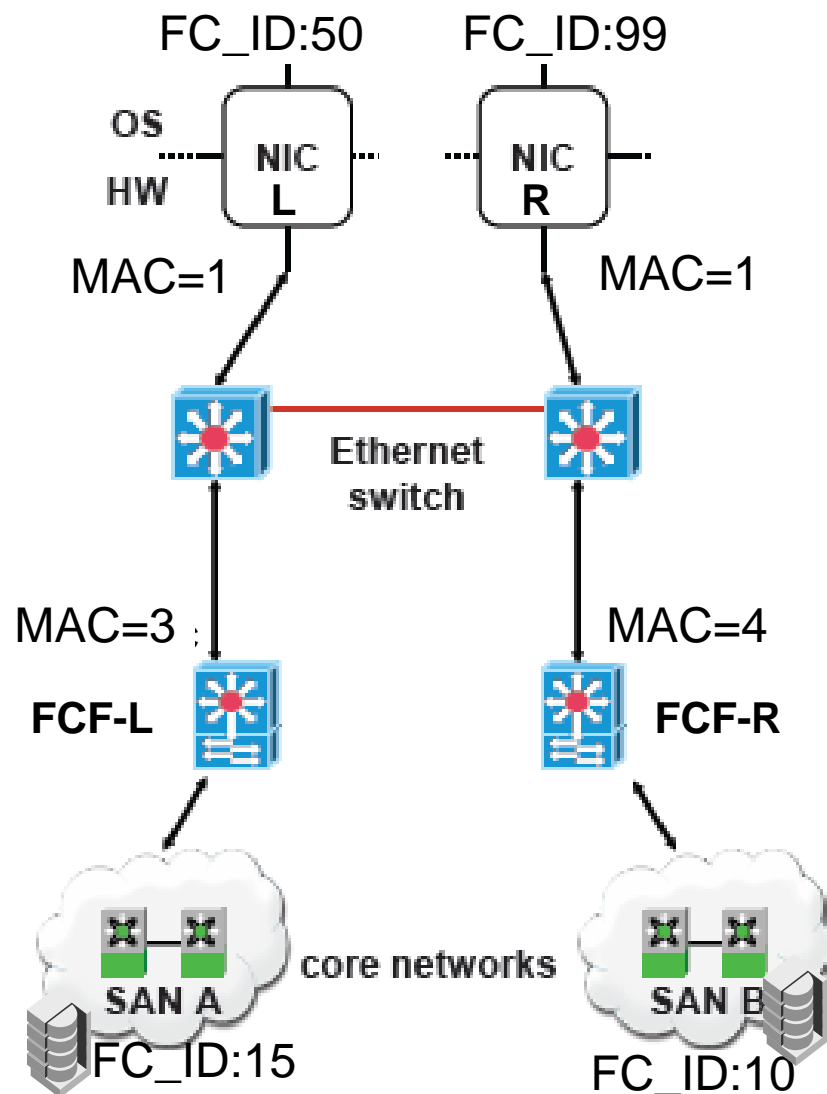
What happens when duplicated NICs occur within the same FCoE fabric?

Need to look at connections to different Ethernet Switches

Need to look at connections to the same Ethernet Switch

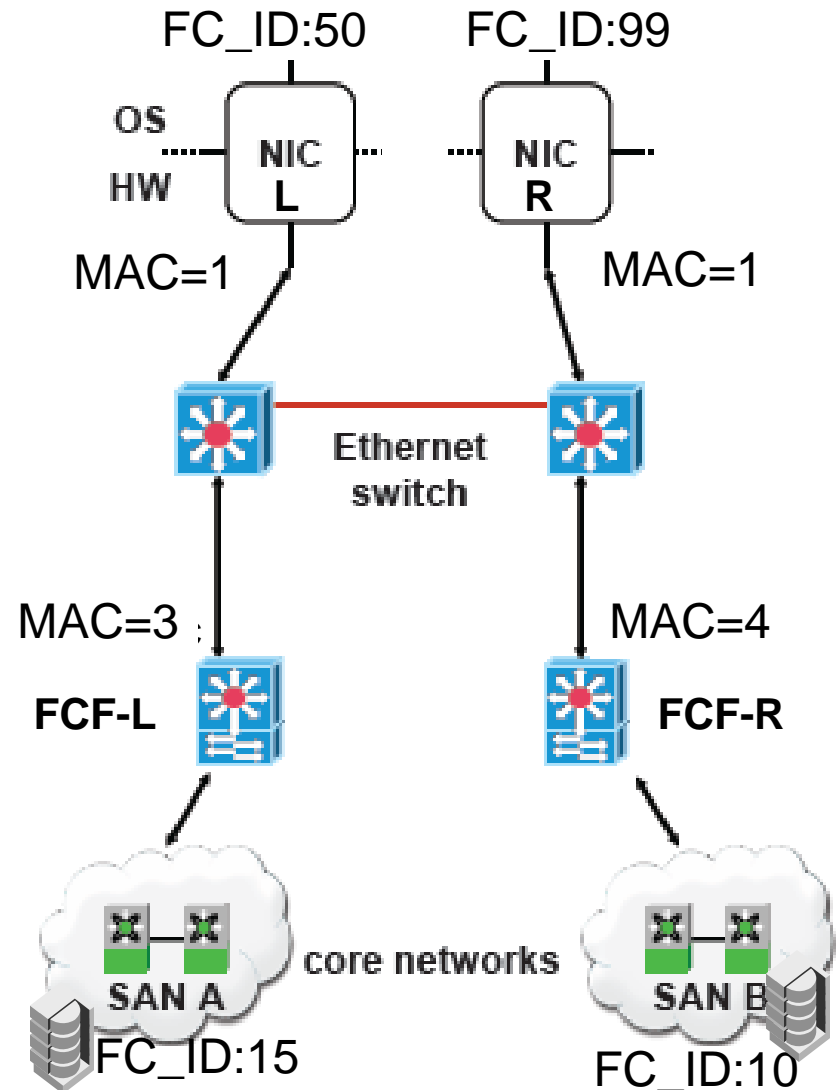
Scenario 1B: Duplicate MACs within the Same Fabric

1. FCF-L discovers FCF-R (and Visa Versa) then instantiates VE_Ports between the FCFs
 - [It is now one Fabric](#)
(There is no cabling error)
2. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Left FCF (FCF-L)
MAC:1 ← → MAC:3
3. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Right FCF (FCF-R)
MAC:1 ← → MAC:4
4. Probably stable but what if one of the switches forgets it learning
5. Data Corruption on Writes or Reads?



Scenario 1B: Corruption prevented by Rule A1

1. FCF-L discovers and instantiates with FCF-R
 - FCF-L (MAC:3) $\leftarrow \rightarrow$ FCF-R (MAC:4)
2. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Left FCF (FCF-L) MAC:1 $\leftarrow \rightarrow$ MAC:3
 - FCF-L is selected by Admin policy
3. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Right FCF (FCF-R) MAC:1 $\leftarrow \rightarrow$ MAC:4
 - If FCF-R is selected by Admin policy
4. Outgoing Data will always go to FCF-L or FCF-R because it has a target MAC address of either MAC:3 or MAC:4
 - Therefore, no Data Corruption on Writes
5. Incoming Data Frame leaves FCF-L with SA=3, DA=1, S_ID:15 & D_ID:50
 - Could be routed to the Right NIC instead of the Left NIC
6. NIC-R will notice that the SA of the incoming frame is not one with which it is instantiated
 - According to [Rule A1](#), the frame must be discarded, an alert issued, **(and the NIC shut down?)**
7. Data Corruption is avoided for Reads and Writes



What is learned with Scenario 1B

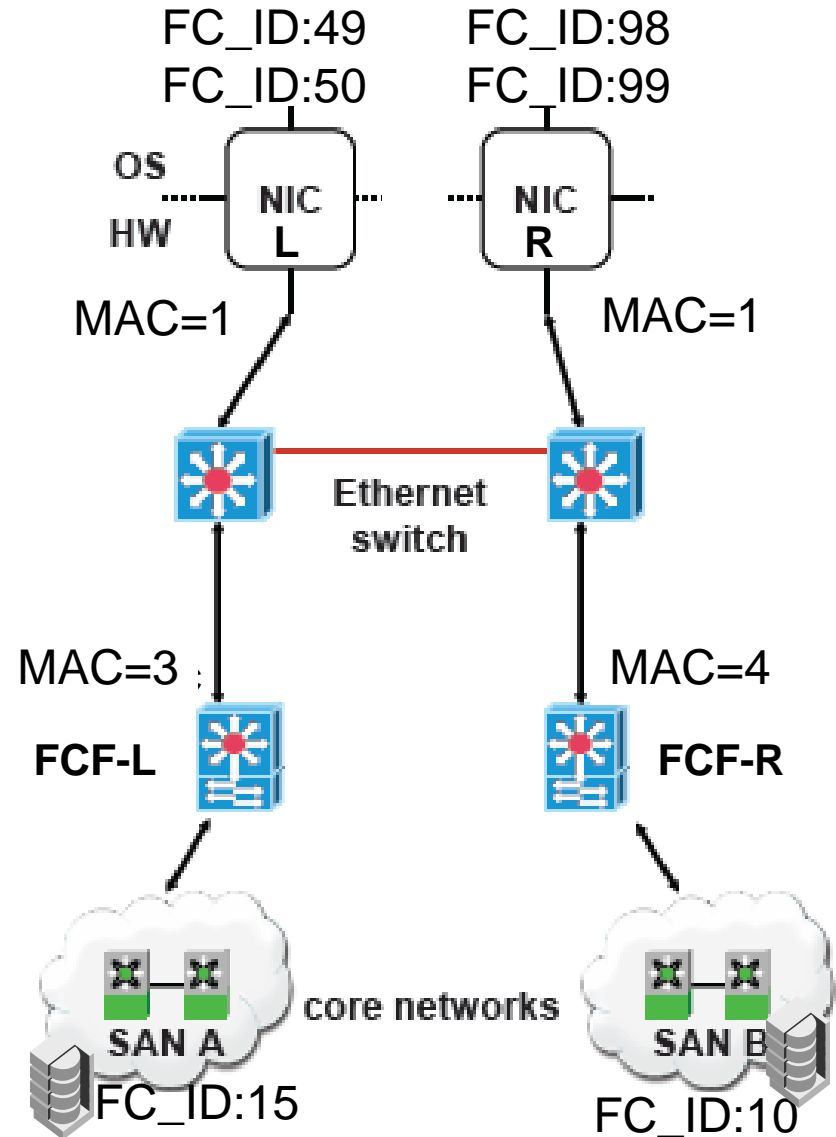
Even with duplicate MAC Addresses on the Host systems across the same Fabric

- No corruption will occur if **Rule A1** is used on the End Node
 - I/O headed toward the Storage Controller will not be disrupted by the connection error, because it is sent toward the MAC address of the appropriate FCF
 - I/O headed toward a Host NIC will be discarded if the wrong frame is received – because of **Rule A1**
- Alerts will be issued to the Admin to request correction
- And the NIC Optionally will be shut down until Admin intervention
- Following **Rule A1** (on the End Nodes) will prevent duplicate Host MAC Addresses from causing corruption



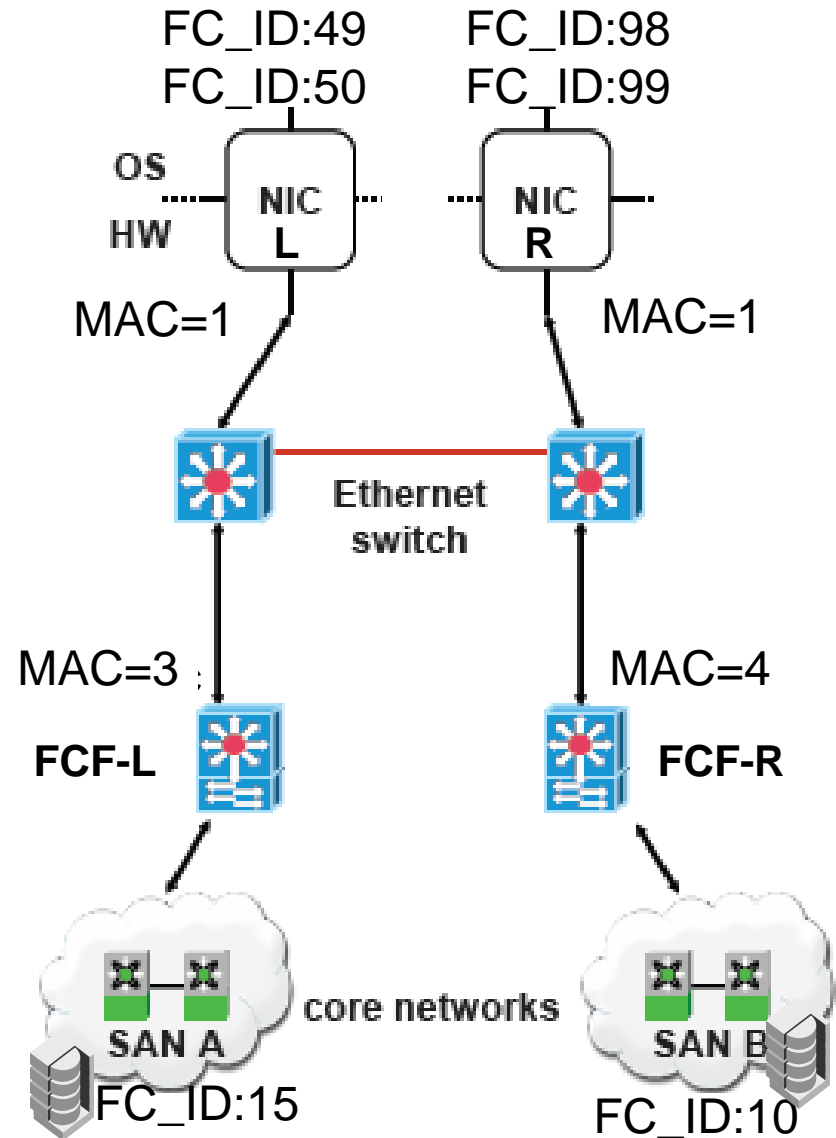
Scenario 1C: Duplicate MACs within the Same Fabric

1. FCF-L discovers FCF-R (and Visa Versa) then instantiates VE_Ports between the FCFs
 - [It is now one Fabric](#)
(There is no cabling error)
2. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with both the Left & Right FCFs
 - NIC-L & FCF-L is MAC:1 ← → MAC:3
 - NIC-L & FCF-R is MAC:1 ← → MAC:4
3. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Right FCF
 - NIC-R & FCF-L is MAC:1 ← → MAC:3
 - NIC-R & FCF-R is MAC:1 ← → MAC:4
4. Probably stable but what if one of the switches forgets it learning
5. Data Corruption on Writes or Reads?



Scenario 1C: Corruption prevented by Rule FC2

1. An VE_Port is established between LCF-L & FCF-R
2. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with both FCFs
 - NIC-L & FCF-L is MAC:1 ← →MAC:3
 - NIC-L & FCF-R is MAC:1 ← →MAC:4
3. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with both FCFs
 - NIC-R & FCF-L is MAC:1 ← →MAC:3
 - NIC-R & FCF-R is MAC:1 ← →MAC:4
4. Outgoing Data will always go to FCF-L or FCF-R because it has a target MAC address of either MAC:3 or MAC:4
 - Therefore, no Data Corruption on Writes
5. Incoming Data Frame leaves FCF-L with SA=3, DA=1, S_ID:15 & D_ID:50 might be routed to the wrong MAC
 - Could be routed to the Right NIC instead of the Left NIC
6. NIC-R's FC Entity will notice that the D_ID of the incoming FC frame is not equal to its own FC_ID
 - According to [Rule FC2](#), the frame must be discarded (and an alert issued and the NIC shut down)
7. Data Corruption is avoided for Reads and Writes



What is learned with Scenario 1C

Even with duplicate MAC Addresses on the Host systems across the same Fabric

NOTE: since it is “one fabric” there is NO duplication of FC_IDs

- No corruption will occur if **Rule FC2** is used by the Host NIC
 - I/O headed toward the Storage Controller will not be disrupted by the connection error, because it is sent toward the MAC address of the appropriate FCF
 - I/O headed toward a Host NIC will be discarded if the wrong frame is received – because of **Rule FC2**
- Alerts can be issued to the Admin to request correction
- The FCoE NIC Optionally will be shut down until Admin intervention
- Following **Rule FC2** (on End Nodes) will prevent duplicate Host MAC Addresses from causing corruption



Scenario 1D: Duplicate MACs within the Same Fabric & duplicate of an FCF MAC Addr

1. FCF-L discovers FCF-R (and Visa Versa) then instantiates VE_Ports between the FCFs

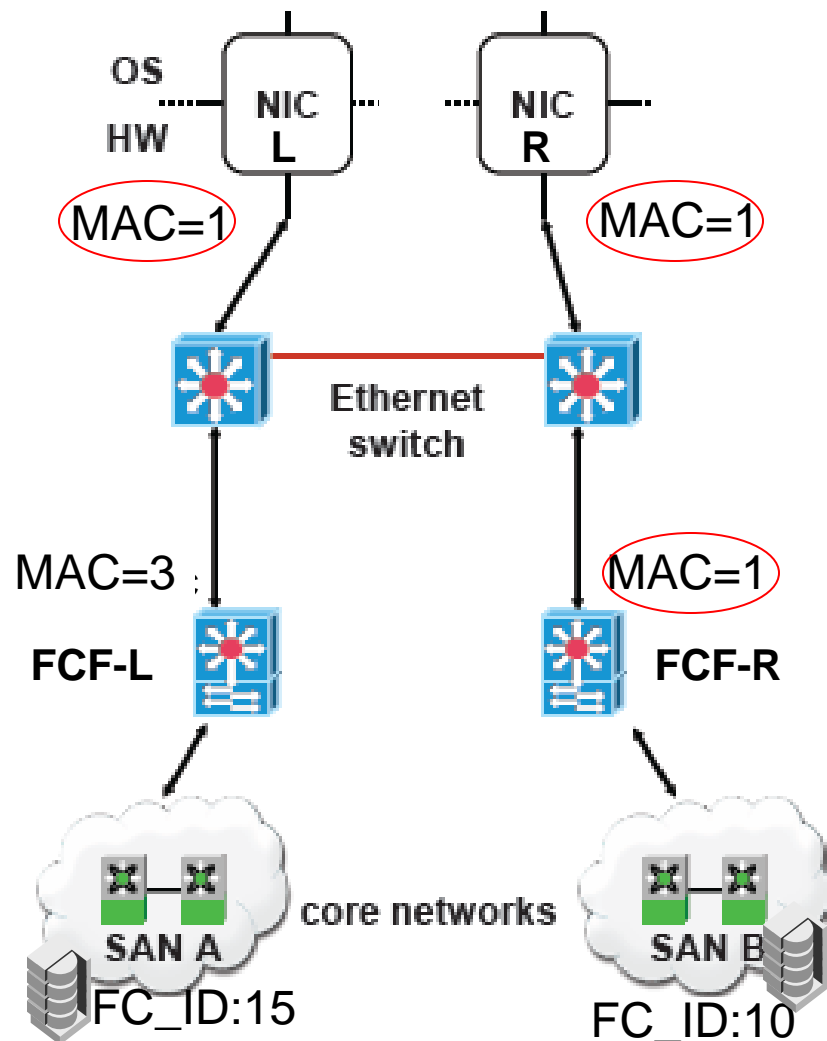
- [It is now one Fabric](#)
(There is no cabling error)

2. Left NIC with MAC:1 discovers and attempts to instantiate a VN_Port to VF_Port connection with the Left FCF (FCF-L) MAC:1 ← → MAC:3

3. Right NIC with MAC:1 discovers and attempts to instantiate a VN_Port to VF_Port connection with the Right FCF (FCF-R) MAC:1 ← → MAC:1

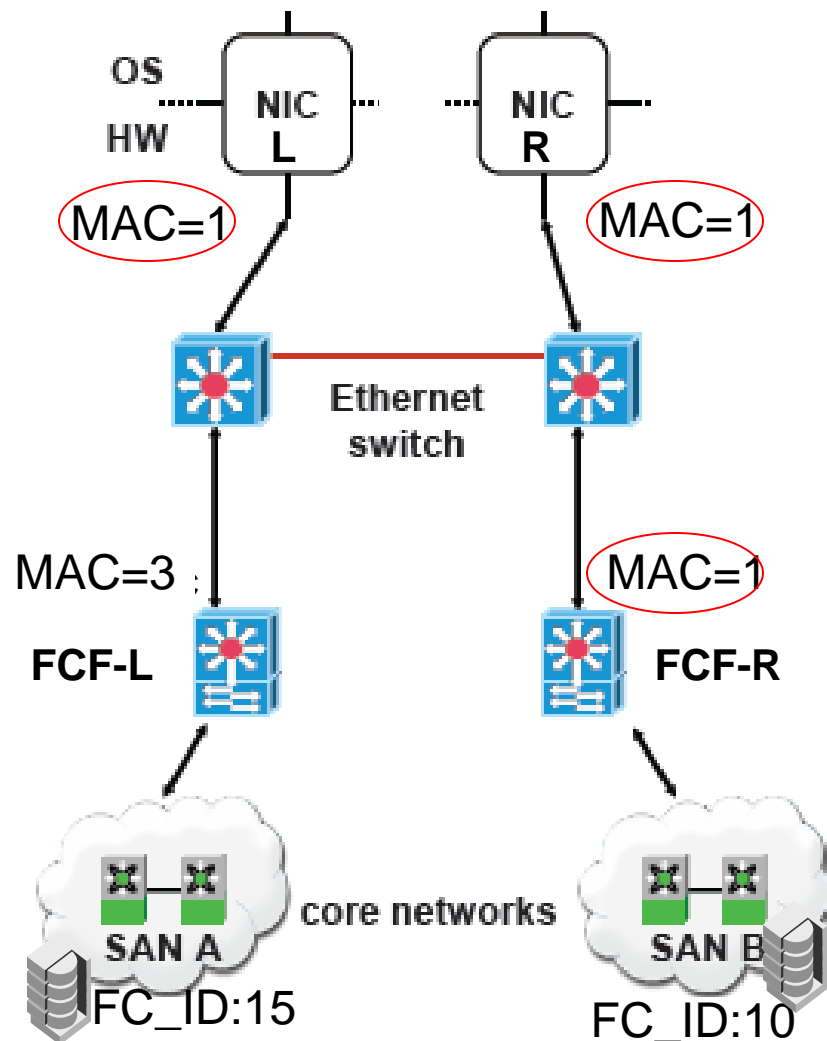
4. Probably stable but what if one of the switches forgets it learning

5. Data Corruption on Writes or Reads?



Scenario 1D: Corruption prevented by Rules I1, & I3

1. A VE_Port $\leftarrow \rightarrow$ VE_Port link exist with MAC:3 $\leftarrow \rightarrow$ MAC:1
2. Left NIC with MAC:1 discovers and attempts an instantiation with a VN_Port to VF_Port connection with the Left FCF (FCF-L) MAC:1 $\leftarrow \rightarrow$ MAC:3
 - FCF-L is selected by Admin policy
 - According to [Rule I3](#) the link will NOT be instantiated and an Alert will be issued, and any links with that same remote MAC:1 address (including the VE_Port) will be shut down and prevented from restarting on that FCF until Admin intervention
3. Right NIC with MAC:1 discovers and attempts instantiation of a VN_Port to VF_Port connection with the Right FCF (FCF-R) MAC:1 $\leftarrow \rightarrow$ MAC:1
 - FCF-R is selected by Admin policy
 - According to [Rule I1](#) the instantiation will not occur, and an Alert will be issued and the MAC on that NIC will be shut down until Admin intervention
4. No I/O is flowing
5. Data Corruption is avoided for Reads and Writes since no Data is flowing



What is learned with Scenario 1D

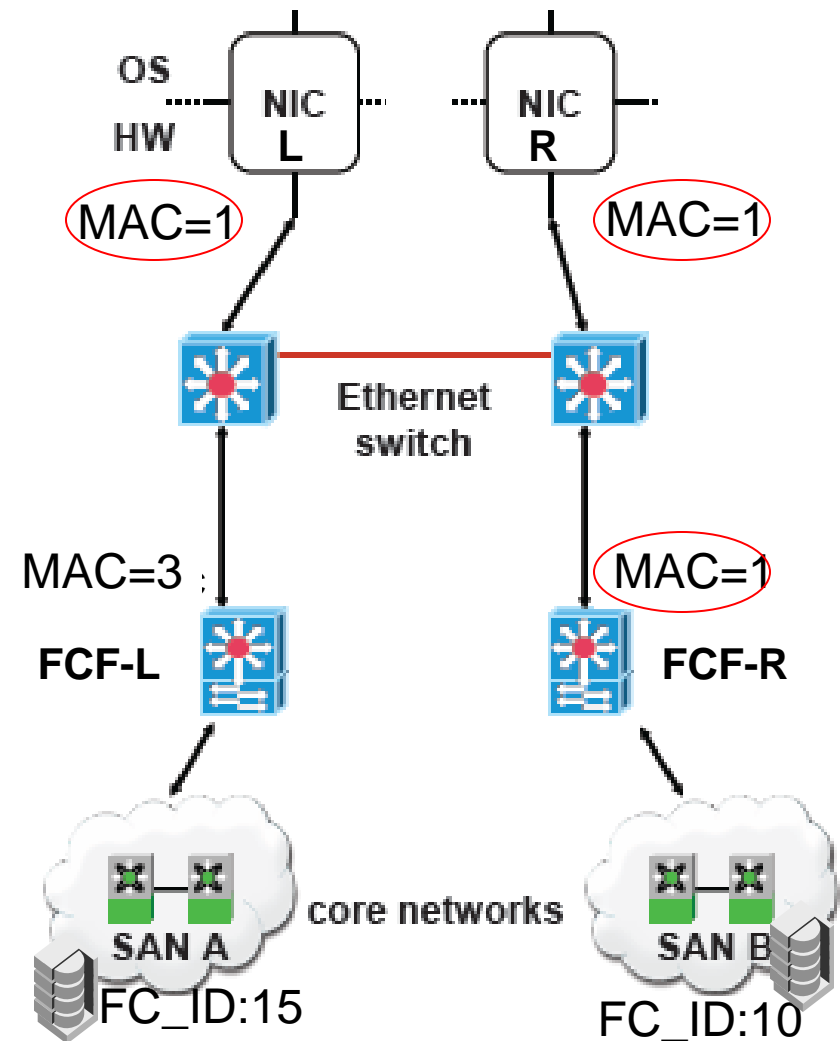
Even with duplicate MAC Addresses on the Host systems across the same Fabric and duplicated with an FCF MAC addresses

- I/O corruption can be avoided in this case because no data will flow
- Alerts will be issued to the Admin to request correction
- End Nodes will shut down the FCoE NIC until Admin intervention
- Following [rules I1 & I3](#) will prevent duplicate Host and FCF MAC Addresses from causing corruption



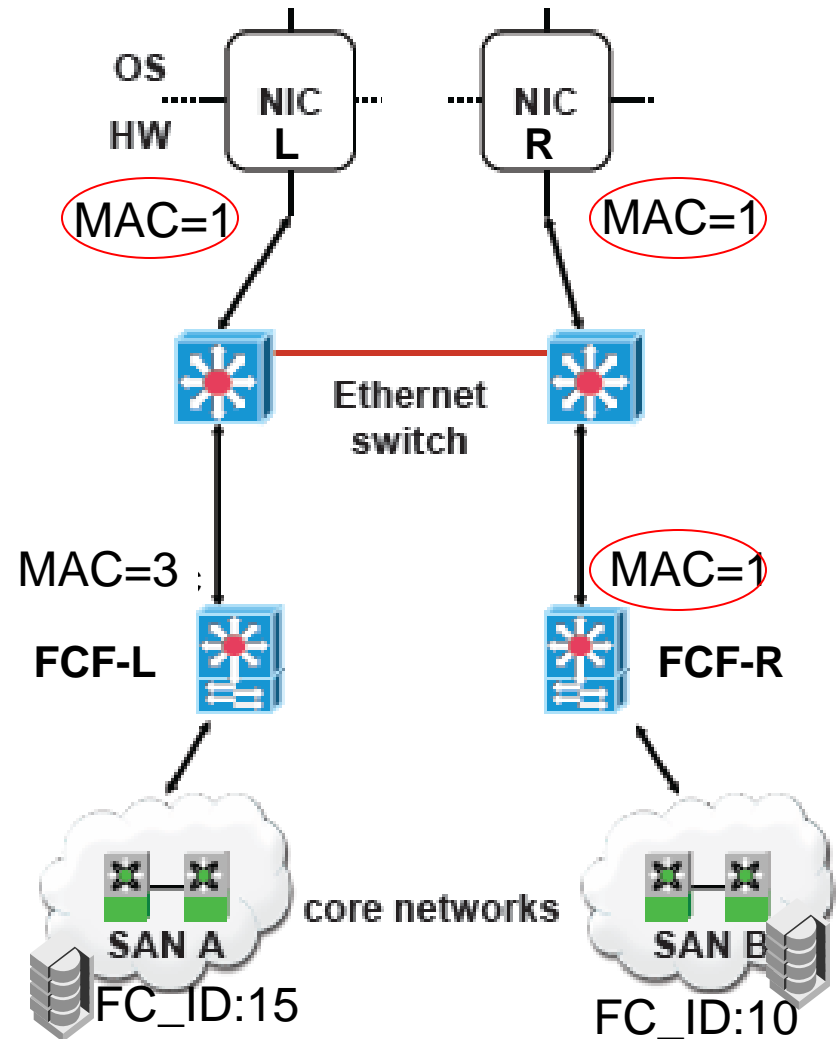
Scenario 1E: Duplicate MACs within the Same Fabric & Duplicate of an FCF MAC Addr

1. FCF-L discovers FCF-R (and Visa Versa) then instantiates VE_Ports between the FCFs
 - [It is now one Fabric](#)
(There is no cabling error)
2. Left NIC with MAC:1 discovers and attempts a VN_Port to VF_Port connection with both the Left & Right FCFs
 - NIC-L & FCF-L is MAC:1 ← → MAC:3
 - NIC-L & FCF-R is MAC:1 ← → MAC:1
3. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the Right FCF
 - NIC-R & FCF-L is MAC:1 ← → MAC:1
 - NIC-R & FCF-R is MAC:1 ← → MAC:4
4. Probably stable but what if one of the switches forgets it learning
5. Data Corruption on Writes or Reads?



Scenario 1E: Corruption prevented by Rules I1, I3

1. An VE_Port is established between LCF-L & FCF-R (MAC:3 \leftarrow \rightarrow MAC:1)
2. Left NIC with MAC:1 discovers and attempts to instantiate a VN_Port to VF_Port connection with both FCFs
 - NIC-L & FCF-R (MAC:1 \leftarrow \rightarrow MAC:1) is denied because of **Rule I1** and an alert will be issued, and the NIC's MAC shut down
 - (The Advertisement might also be prevented because of **Rule D1**)
 - NIC-L & FCF-L is MAC:1 \leftarrow \rightarrow MAC:3 is denied because Host MAC shut down
 - However, if instantiation with FCF-L (MAC:3) was attempted before FCF-R (MAC:1) then it will not be instantiated because of **Rule I3**
 - But if some how instantiated it will be shut down when instantiation is attempted with FCF-R because of **Rule I1**
3. Right NIC with MAC:1 discovers and attempts to instantiate a VN_Port to VF_Port connection with both FCFs
 - All instantiations are denied because of **Rules I1, I3**, and indirectly by **Rule D1**
4. No I/O is flowing
5. Data Corruption is avoided for Reads and Writes since no Data is flowing



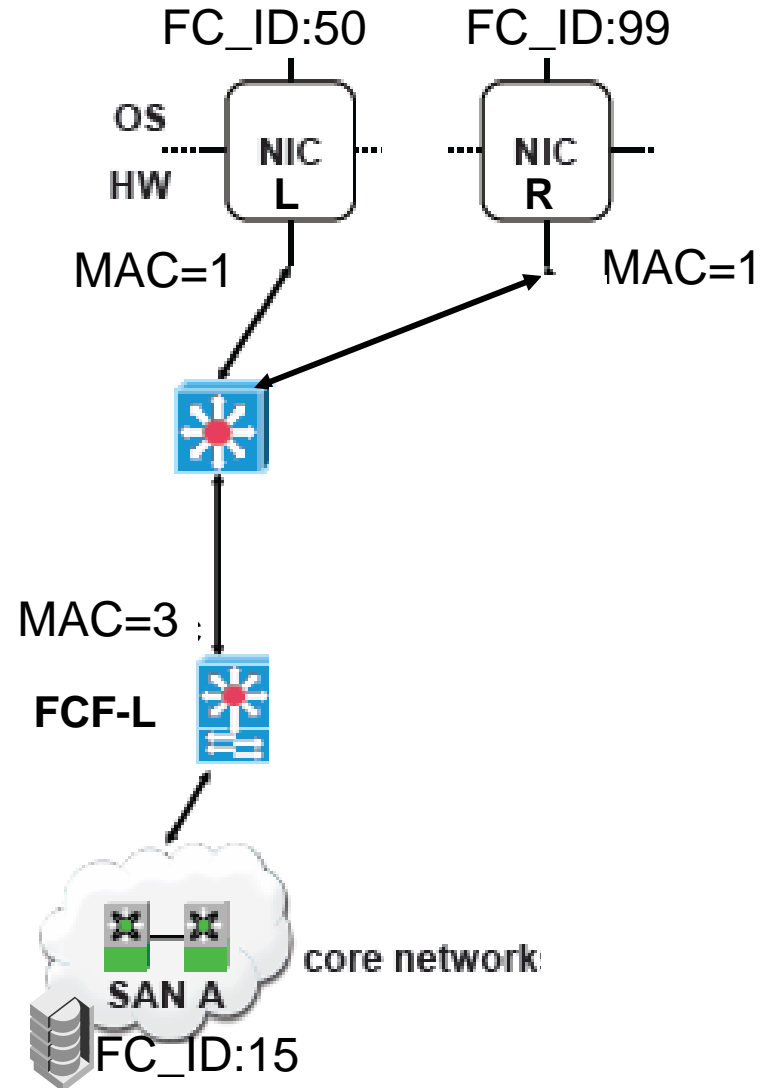
What is learned with Scenario 1E

Even with duplicate MAC Addresses on the Host systems across the same Fabric and duplicated with an FCF MAC addresses

- I/O corruption can be avoid in this case because no data will flow
- Alerts will be issued to the Admin to request correction
- The End Nodes with duplicate MAC Addresses Optionally will be Shut down until Admin intervention
- Following rule 1 & 3 will prevent duplicate Host and FCF MAC Addresses from causing corruption

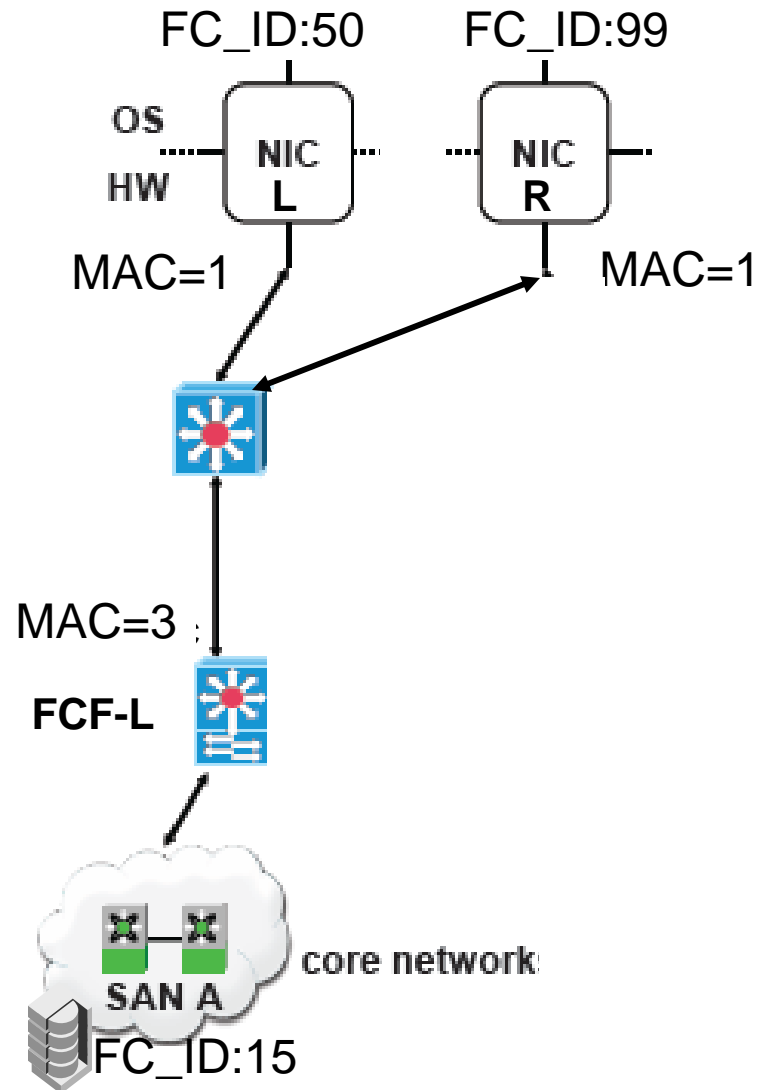
Scenario 1F: Duplicate MACs connected to Same FCF

1. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the FCF (FCF-L) MAC:1 ← → MAC:3
2. Right NIC with MAC:1 also discovers and instantiates a VN_Port to VF_Port connection with the FCF (FCF-L) MAC:3 ← → MAC:4
3. Probably Not stable
4. Data Corruption on Writes or Reads?



Scenario 1F: Corruption prevented by Rule FC2

1. Left NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the FCF (FCF-L) MAC:1 ← → MAC:3
2. Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with the FCF (FCF-L) MAC:1 ← → MAC:3
3. Outgoing Data will always go from NIC-L or NIC-R to FCF-L because NIC-L and NIC-R have a DA MAC address of MAC:3
 - Therefore, no Data Corruption on Writes
4. Incoming Data Frame intended for NIC-L leaves FCF-L with SA=3 & DA=1 & FC S_ID:15 & D_ID:50
5. Frame is routed to NIC-R instead of NIC-L
6. Frame will be received by the MAC on NIC-R since it has an instantiation with MAC:3 as an N-Port to F-Port link
7. The Frame will be stripped of its Ethernet and FCoE headers and given to the VN_Port FC Entity
8. The VN_Port FC Entity will notice that the D_ID is not its own FC_ID
 - According to [Rule FC2](#), discard the frame, throw an alert, and shut down the NIC (NIC-R in this case)
9. NIC-L may continue to operate after a SCSI I/O recovery, unless an residual Frame intended for NIC-R is set to NIC-L
 - In which case, it too, according to [Rule FC2](#), will also discard the frame, issue an alert, and shut down NIC (NIC-L in this case)
10. Data Corruption is avoided on Reads and Writes



What is learned with Scenario 1F

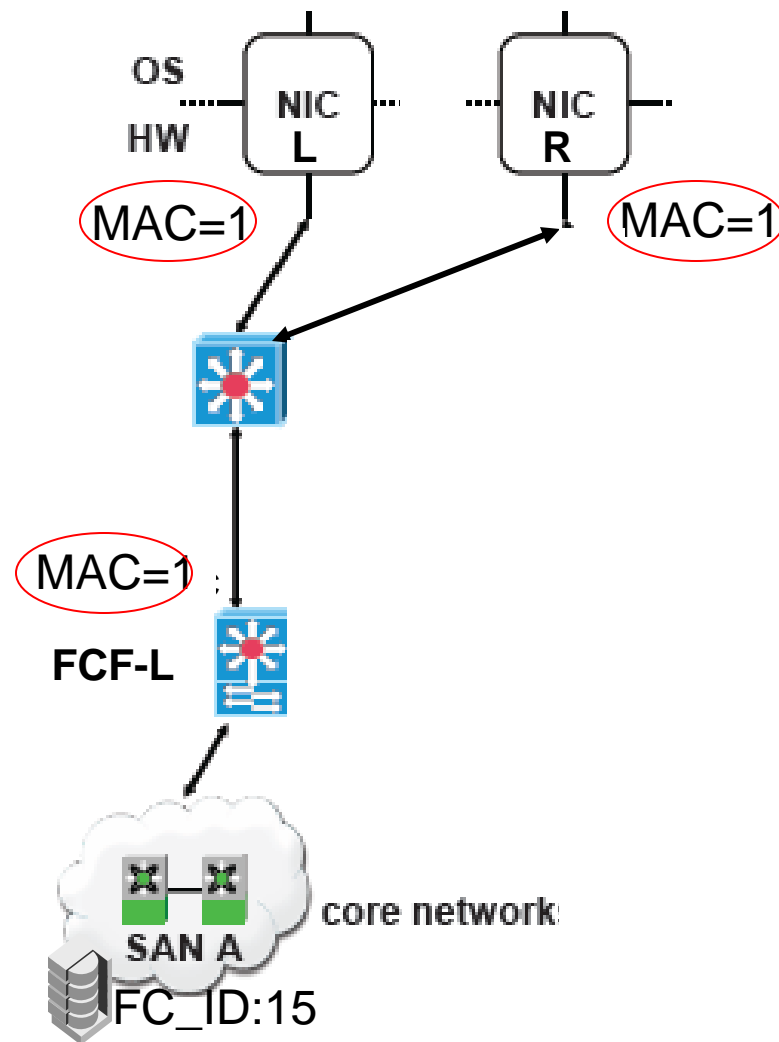
Even with duplicate MAC Addresses on the Host systems accessing a single fabric

- No corruption will occur if **Rule FC2** is used
 - I/O headed toward the Storage Controller will not be disrupted by the connection error, because it is headed toward the MAC address of the appropriate FCF
 - I/O from the storage device will be discarded if the wrong NIC receives the frame and an Alert will be issued – because of **Rule FC2**
- Alerts will be issued to the Admin to request corrections
- The End node(s) with duplicate MAC addresses will be shut down until Admin intervention
- Following **Rule FC2** will prevent duplicate Host MAC Addresses from causing corruption



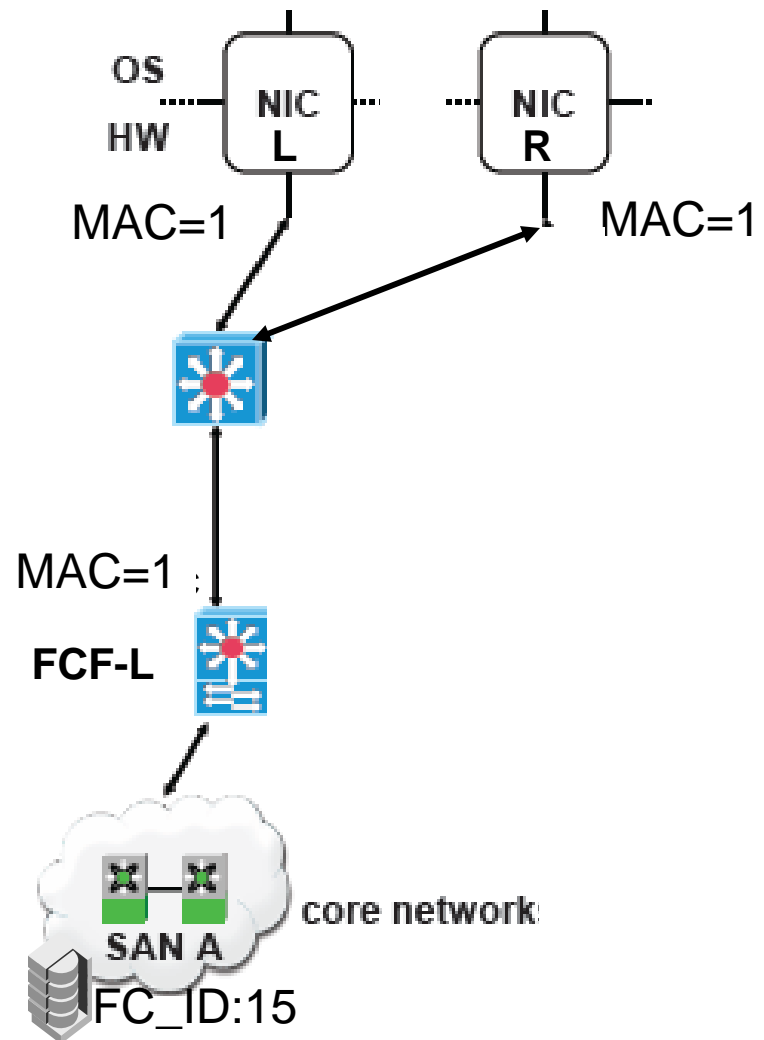
Scenario 1G: Duplicate MACs connected to Same FCF with the same MAC address

1. Left NIC with MAC:1 discovers and attempts instantiation as VN_Port to VF_Port connection with FCF (FCF-L) MAC:1 ← →MAC:1
2. Right NIC with MAC:1 also discovers and attempts instantiation as VN_Port to VF_Port connection with the FCF (FCF-L) MAC:1 ← →MAC:1
3. Probably Not stable
4. Data Corruption on Writes or Reads?



Scenario 1G: Corruption prevented by Rule I1

1. Left NIC with MAC:1 discovers and unsuccessfully attempts instantiation as a VN_Port to VF_Port connection with the FCF (FCF-L) MAC:1 ← → MAC:1
 - Discovery, according to **Rule D2**, will probably not advertise this FCF-L, and will issue Alerts
 - Instantiation with FCF-L must therefore be directed by an Admin
 - According to **Rule I1**, the instantiation will not occur and an Error will be thrown, and the NIC will be shut down until Admin intervention
2. Right NIC (NIC-R) will also be shut down
 - According to **Rule I1**, the instantiation will not occur and an Error will be thrown
 - And the MAC will be shut down until Admin intervention
3. Neither outgoing Data nor incoming data will flow since no NIC is instated as a VN_Port
4. Data Corruption is avoided for Reads and Writes since no Data is flowing



What is learned with Scenario 1G

Even with duplicate MAC Addresses on the Host systems and accessing a single fabric with the same MAC Address on the FCF

- I/O corruption is avoided in this case because no data will flow
- Alerts will be issued to the Admin to request correction
- The End Node NICs with duplicate MAC Addresses will be shutdown until Admin intervention
- Following [Rule I1](#) will prevent duplicate Host MAC Addresses from causing corruption even if the same as the MAC address on the FCF

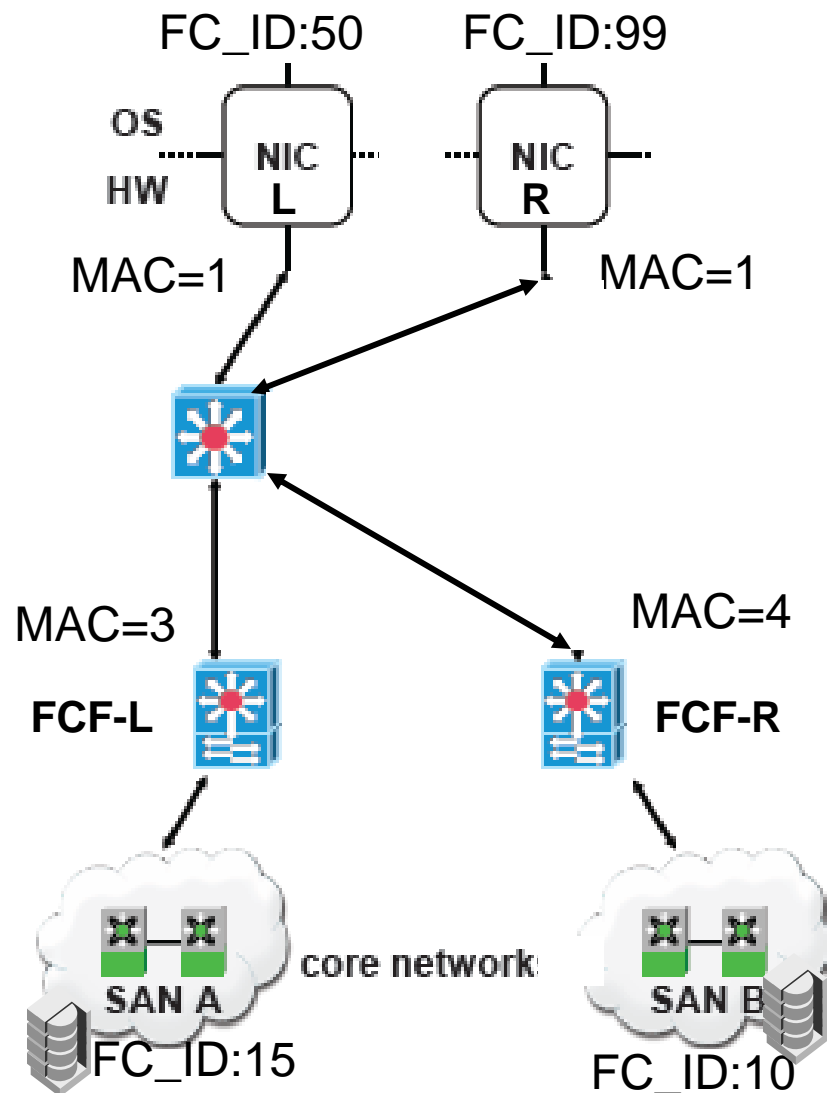


Scenario 1G: Duplicate MACs connected to Different FCFs through a common Ethernet Switch

1. Left NIC & Right NIC with MAC:1 discovers and attempt to instantiate a VN_Port to VF_Port connection with both the Left FCF (FCF-L) MAC:1 ← → MAC:3 and the Right FCF (FCF-R) MAC:1 ← → MAC:4

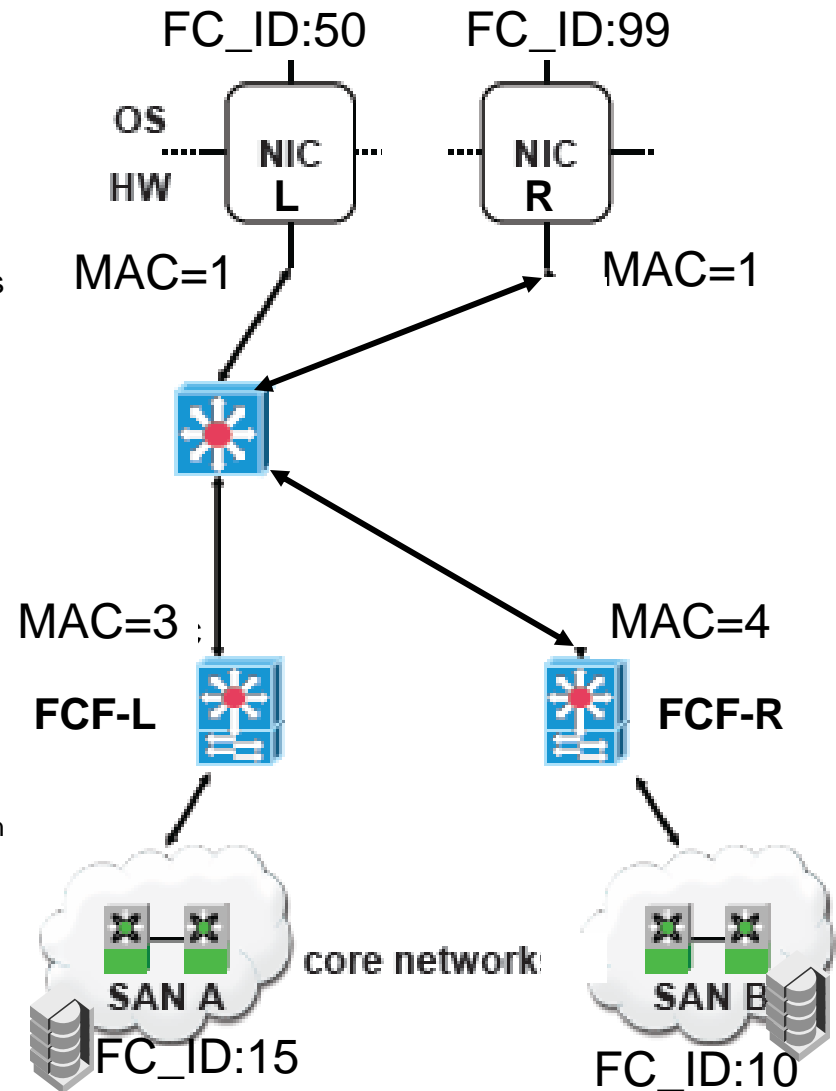
2. What will happen?

3. Will data corruption occur on Writes or Reads?



Scenario 1G: Corruption prevented by Rule FC2

1. FCF-L and FCF-R discover each other and instantiate a VE_Port to VE_Port connection FCF-L (MAC:3) ← → FCF-R (MAC:4)
2. Left NIC & Right NIC with MAC:1 discovers and instantiates a VN_Port to VF_Port connection with both the Left FCF (FCF-L) MAC:1 ← → MAC:3 and the Right FCF (FCF-R) MAC:1 ← → MAC:4
3. Outgoing Data will always go to the appropriate FCF because it has a destination MAC address of either MAC:3 or MAC:4 as appropriate
 - Therefore, no Data Corruption on Writes
4. Incoming Data Frame intended for NIC-L leaves FCF-L with SA=3 & DA=1 & FC S_ID:15 & D_ID:50
5. Frame is routed to NIC-R instead of NIC-L
6. Frame will be received by the MAC on NIC-R since it has an instantiation with MAC:3 as an N-Port to F-Port link
7. The Frame will be stripped of its Ethernet and FCoE headers and given to the VN_Port FC Entity
8. The VN_Port FC Entity will notice that the D_ID is not its own FC_ID
 - According to [Rule FC2](#), discard the frame, throw an alert, and shut down the NIC (NIC-R in this case)
9. NIC-L may continue to operate after a SCSI I/O recovery, unless an residual Frame intended for NIC-R is set to NIC-L
 - In which case, it too, according to [Rule FC2](#), will also discard the frame, issue an alert, and shut down NIC (NIC-L in this case)
10. Data Corruption is avoided on Reads and Writes



What is learned with Scenario 1G

Even with duplicate MAC Addresses on the Host systems connecting to different Fabrics through a common Ethernet switch:

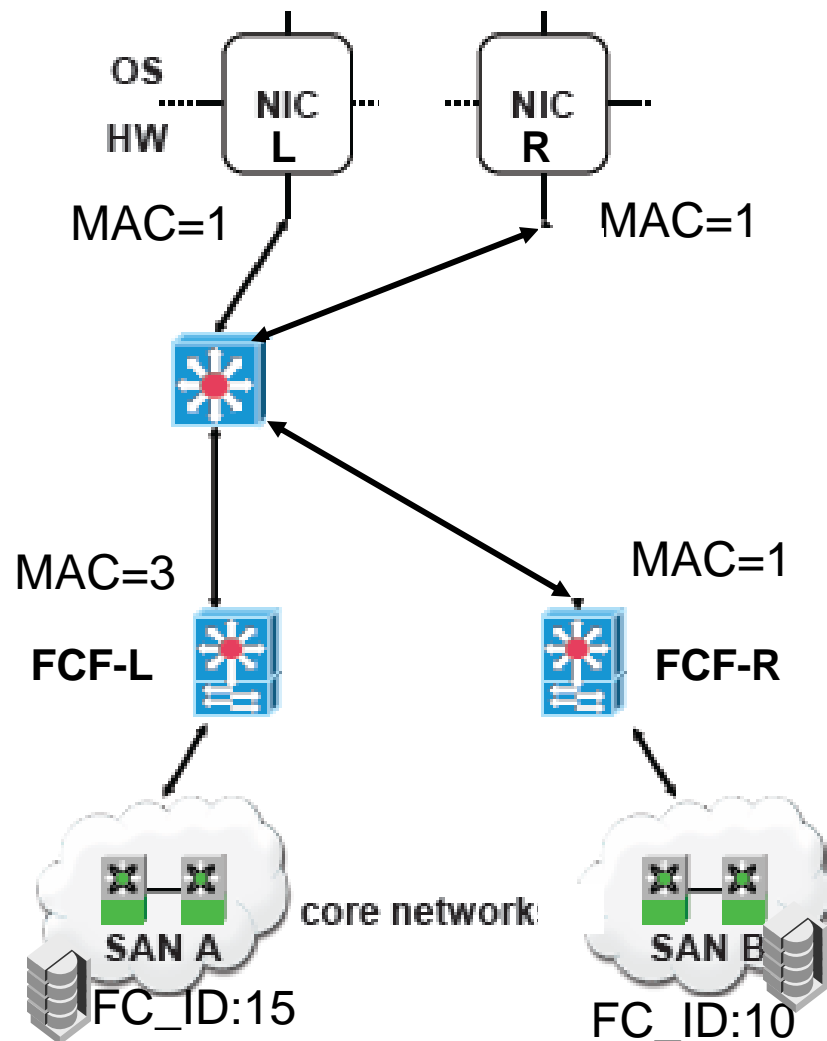
- Alerts will be issued to the Admin to request corrections for unsuccessful instantiations and miss-routed I/O
- No corruption will occur if [Rule FC2](#) is used

Scenario 1H: Duplicate MACs connected to Different FCFs through a common Ethernet Switch

1. Left NIC & Right NIC with MAC:1 discovers and attempt to instantiate a VN_Port to VF_Port connection with both the Left FCF (FCF-L) MAC:1 ← → MAC:3 and the Right FCF (FCF-R) MAC:1 ← → MAC:4

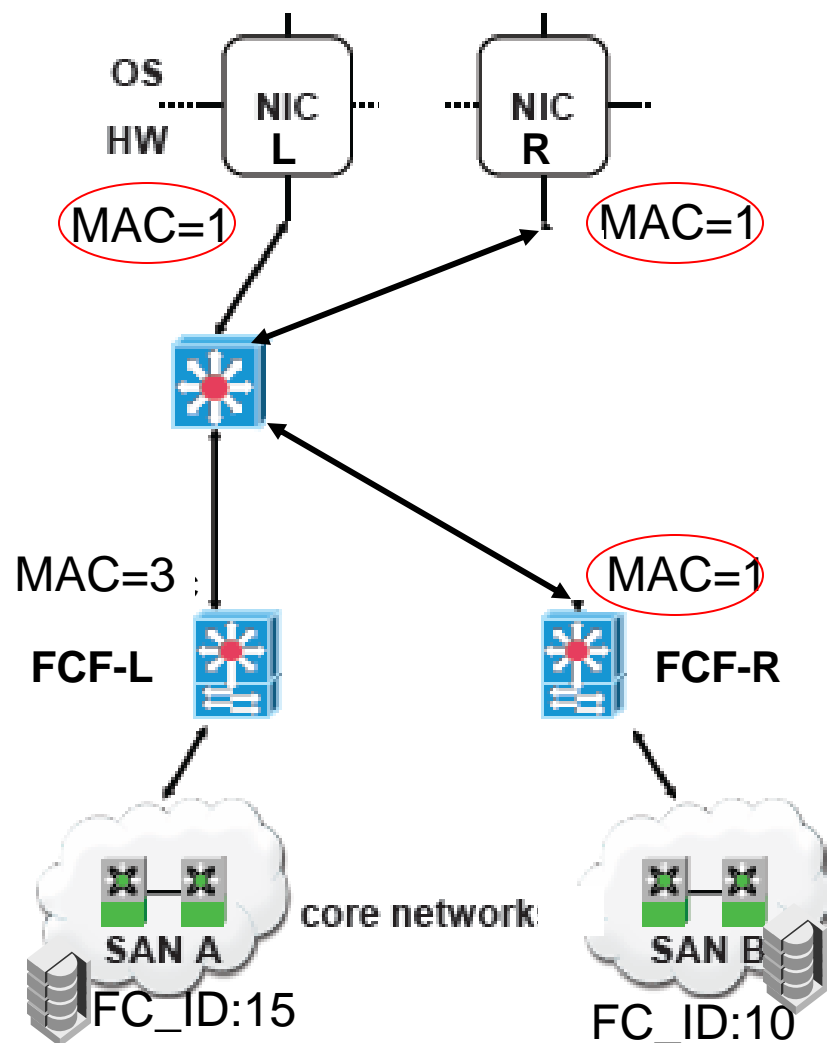
2. What will happen?

3. Will data corruption occur on Writes or Reads?



Scenario 1H: Corruption prevented by Rules I1, I3

1. FCF-R and FCF-L discover each other and instantiate a VE_Port to VE_Port link FCF-L (MAC:3) ← → FCF-R (MAC:1)
2. NIC-L and NIC-R attempt to discover FCFs
3. FCF-R will not advertise to NIC-L or NIC-R, and will issue an Alert based on **Rule D2**
4. NIC-L will attempt to instantiate with FCF-L
 - This will fail according to **Rule I3**, alerts will be issued, and the MAC address of MAC:1 will not be permitted instantiation until Admin action and the VE_Port with FCF-R (MAC:1) will be shut down
5. NIC-R will attempt to instantiate with FCF-L
 - Because of the above, FCF-L will not permit the instantiation with a MAC:1 from the NIC-R until Admin action has occurred
6. Even if directed by Admin, neither NIC-L nor NIC-R will be able to instantiate with FCF-R
 - Instantiation is denied and Alerts issued based on **Rule I1**, and the End Node NICs will be shut down
7. No I/O will flow
8. Data Corruption is avoided on Reads and Writes because no I/O is flowing



What is learned with Scenario 1H

Even with duplicate MAC Addresses on the Host systems across the same Fabric and duplicated with an FCF MAC addresses

- I/O corruption can be avoid in this case because no data will flow
- Alerts will be issued to the Admin to request correction
- FCoE NICs with duplicate MAC Addresses will be shut down until Admin intervention
- Following [Rules I1 & I3](#) will prevent duplicate Host and FCF MAC Addresses from causing corruption

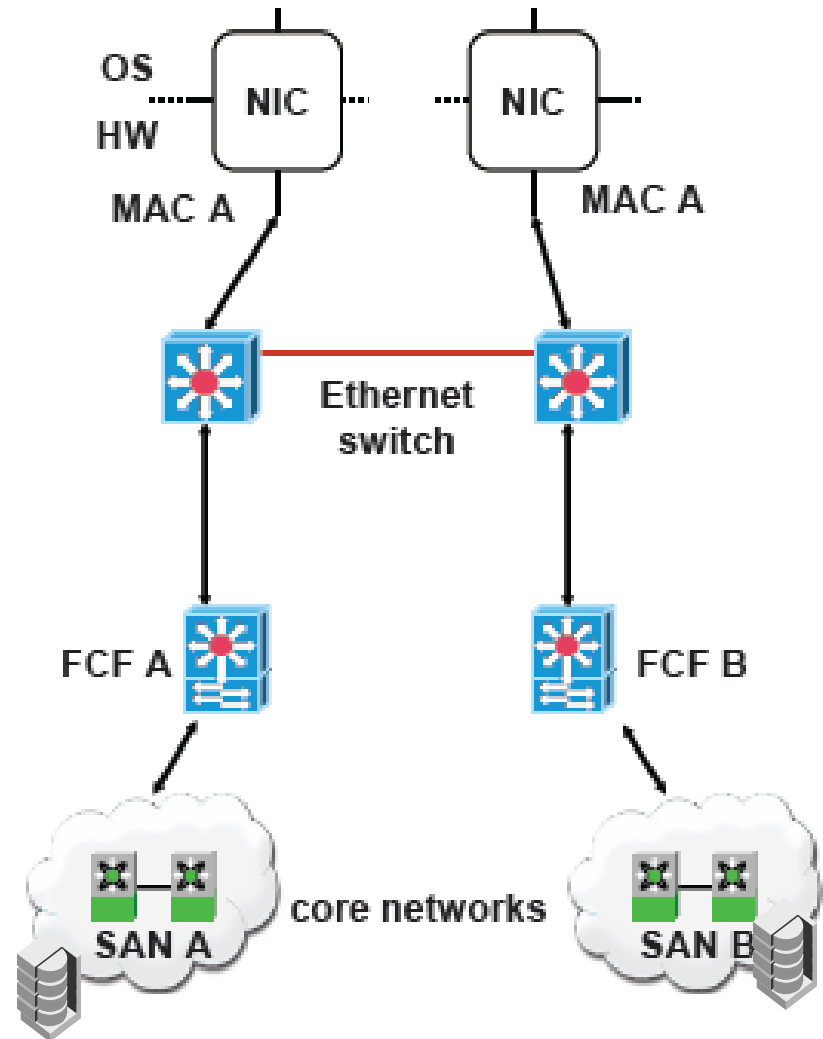
Scenario #2

Same Mistake of: Cross connect edge switches

What happens if one of the top FCoE N_Ports resets and sends FLOGI?

- FLOGI multicast could go to both FCFs
- Both FCoE N_Ports may wind up in the same fabric with different FCIDs

Loss of multipath redundancy!



Solution to Scenario #2

The inclusion of the Discovery protocol will:

- Deal with fabric separation
- The ENodes learn the Unicast addresses of the FCFs and the Fabric_Name
 - The Fabric_Name allows ENodes to detect merged fabrics configurations
- Not send FLOGI in Broadcast/Multicast, but only in Unicast
 - The Fabric_Name may be compared to the one returned in the FLOGI Accept for consistency

Scenario #3

Mistake: Cross connect edge switches

Send FC frames to left NIC's FCID

- Left FCF thinks the right FCF's MAC is for an FCoE N_Port

Nothing breaks immediately, but

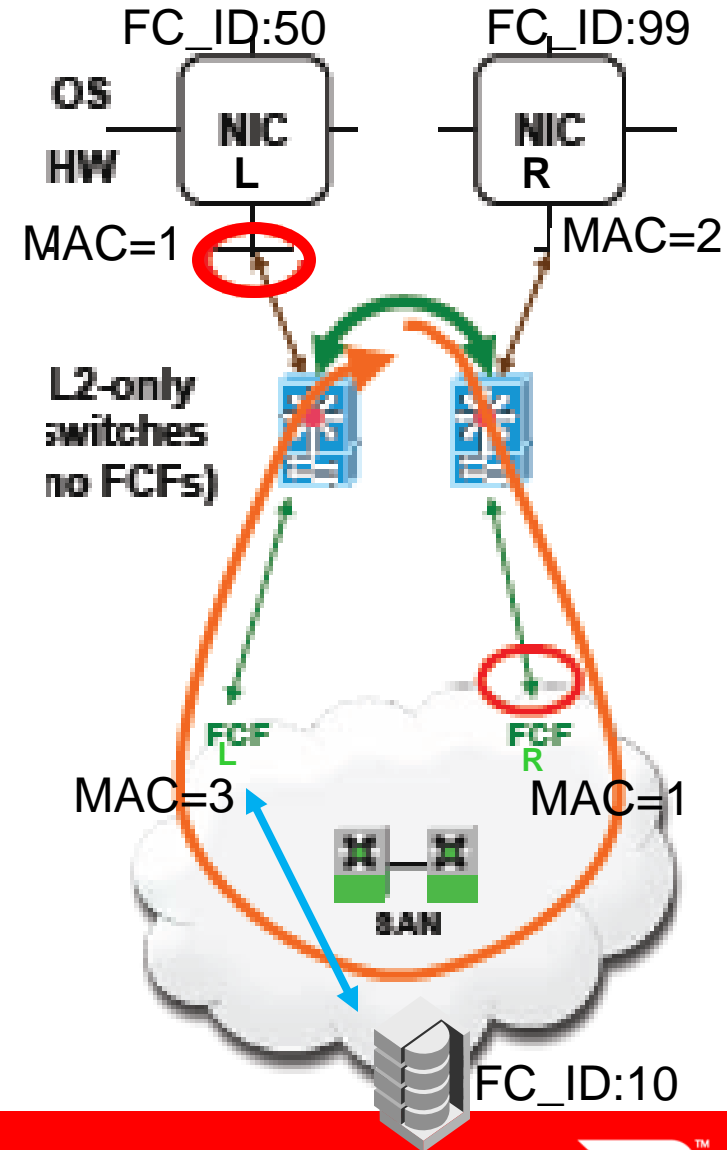
- Suppose left switch forgets where that MAC is located
- Right FCF may send frame that "helps" that switch forget

Left NIC frames now loop forever:

- Sent to left FCF for NIC N_Port
- Left FCF uses N_Port's MAC
- Frame arrives at right FCF

Destination FCID hasn't changed

- Fabric forwards to left FCF

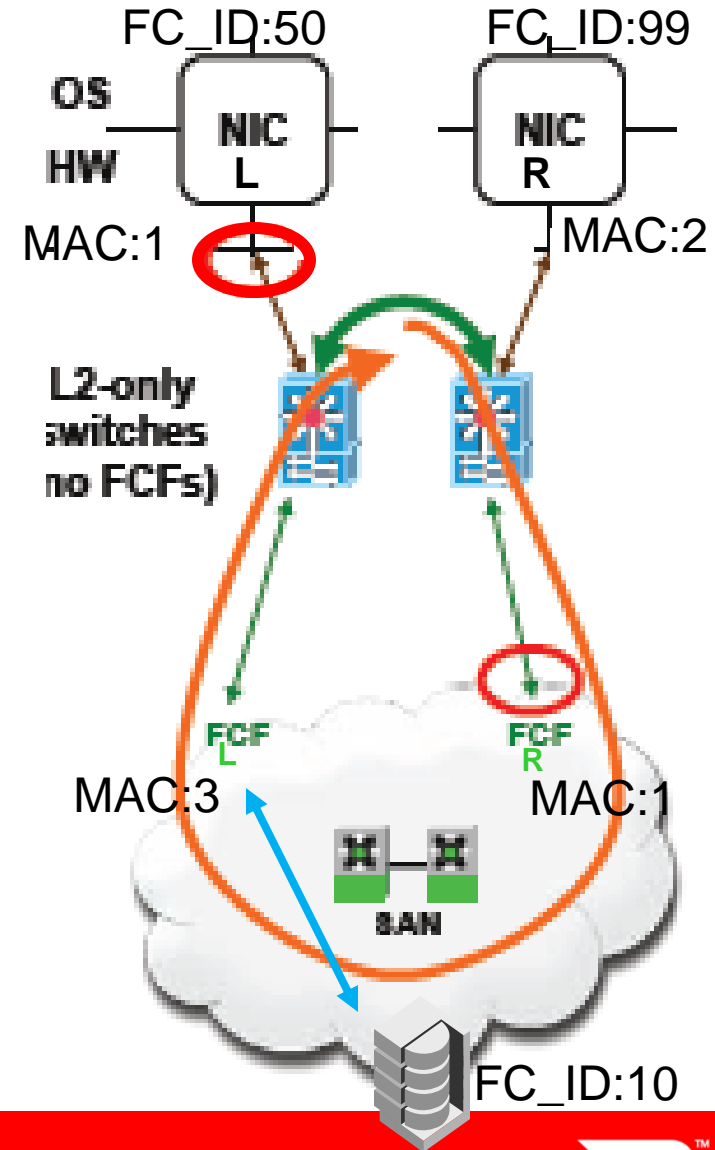


Case 1: Looping prevented by Rule A1 or FC1

- NIC-R instantiates a VN_Port \leftarrow \rightarrow VF_Port relationship with FCF-R
 - MAC:2 \leftarrow \rightarrow MAC:1
- NIC-L instantiates a VN_Port \leftarrow \rightarrow VF_Port relationship with FCF-L
 - MAC:1 \leftarrow \rightarrow MAC:3
- The wire is Plugged wrong
- A Frame leaves FCF-L with SA=3 & DA=1, S_ID:10, and D_ID:50
- Frame will be received by the MAC on FCF-R since it has an instantiation of MAC:1 (as an F-Port)
- Chose 1 of the following implementations
 - There is no instantiated Vx_Port on FCF-R (MAC:1) with a Remote Peer of MAC:3
 - According to [Rule A1](#), issue an alert, discard the frame, **shut down the MAC:1 on FCF-R?**
 - The Frame will be stripped of its Ethernet and FCoE headers and given to the VF_Port FC Entity

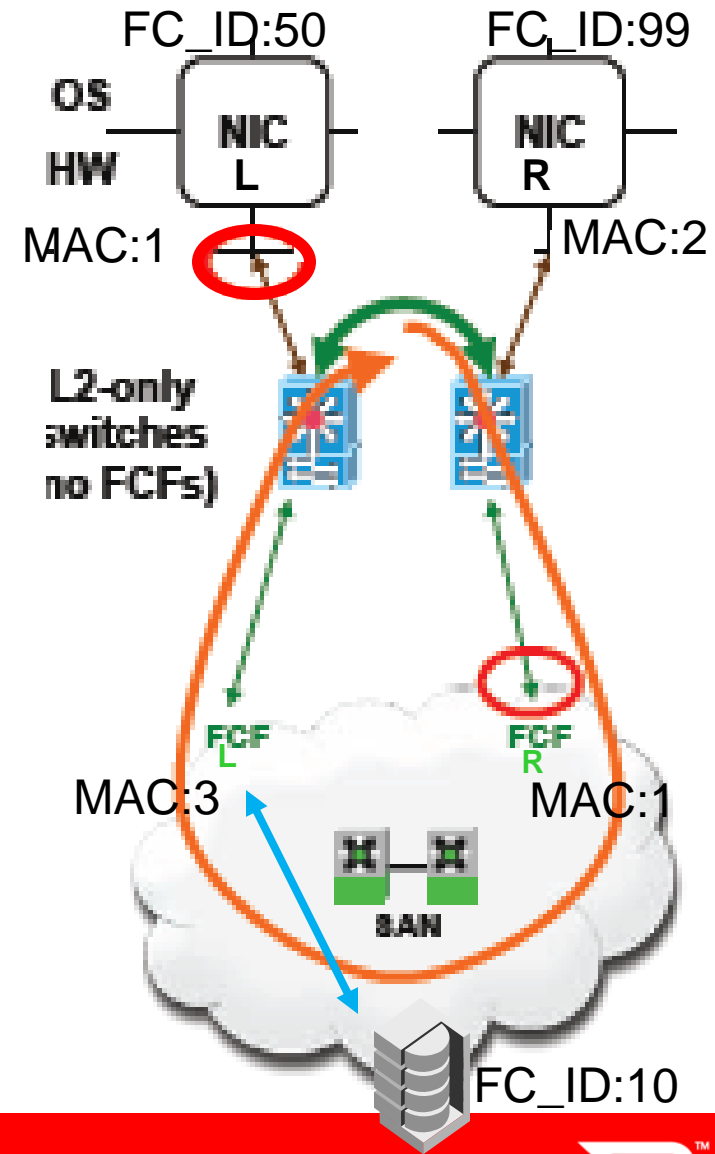
And
The VF_Port FC Entity will not have been instantiated with the N_Port that is represented by the S_ID of the FC Frame that it receives

 - According to [Rule FC1](#), discard the frame, send an Alert, **shut down the MAC:1 on the FCF-R?**
- Looping is prevented



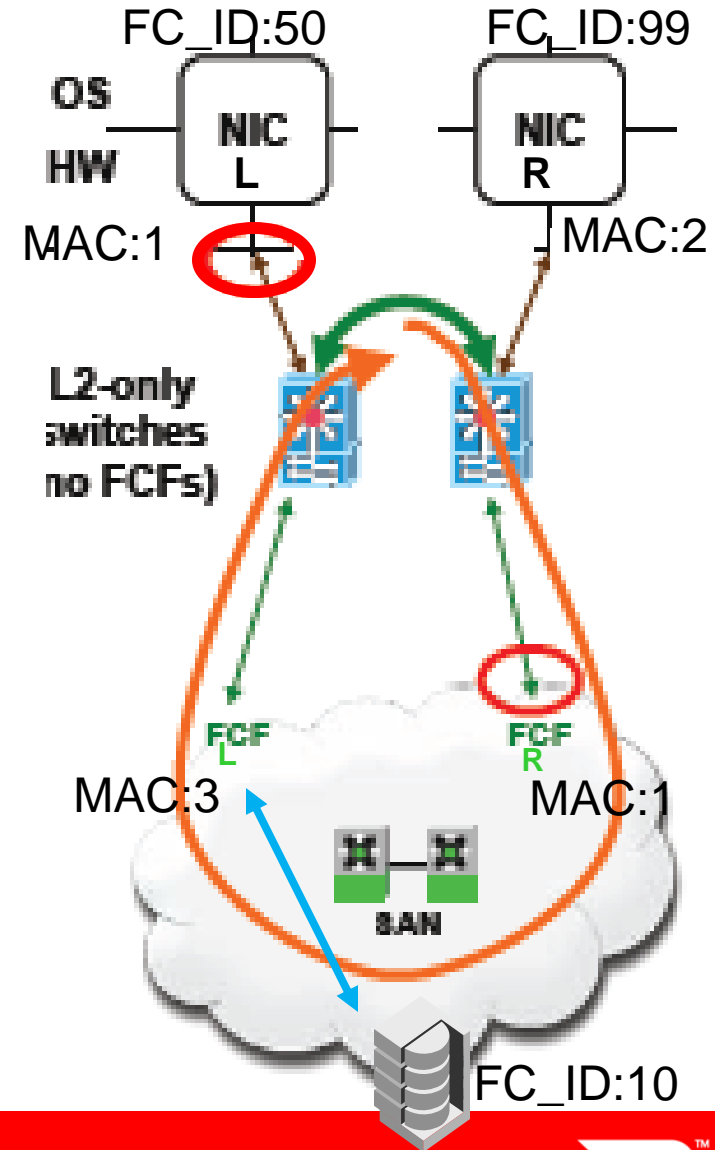
Case 3: Looping prevented by Rules I2/I3

1. NIC-L instantiates as a VN_Port $\leftarrow \rightarrow$ VF_Port relationship with FCF-L
 - MAC:1 $\leftarrow \rightarrow$ MAC:3 and NIC-R instantiates with FCF-R MAC:2 $\leftarrow \rightarrow$ MAC:1
2. The wire is Plugged wrong
3. FCF-L discovers FCF-R
4. FCF-L with MAC:3 attempts to instantiate a VE_Port $\leftarrow \rightarrow$ VE_Port relationship with FCF-R with MAC:1
 - According to [Rule I2/I3](#), deny the instantiation, issue an alert, and shut down the MAC:3 or the other Vx_Ports, and do not permit an instantiation with any MAC:1 until Admin permits
5. No Data will flow to/from NIC-L since the FCF-L port has been shut down.
6. Looping or Data Corruption is prevented



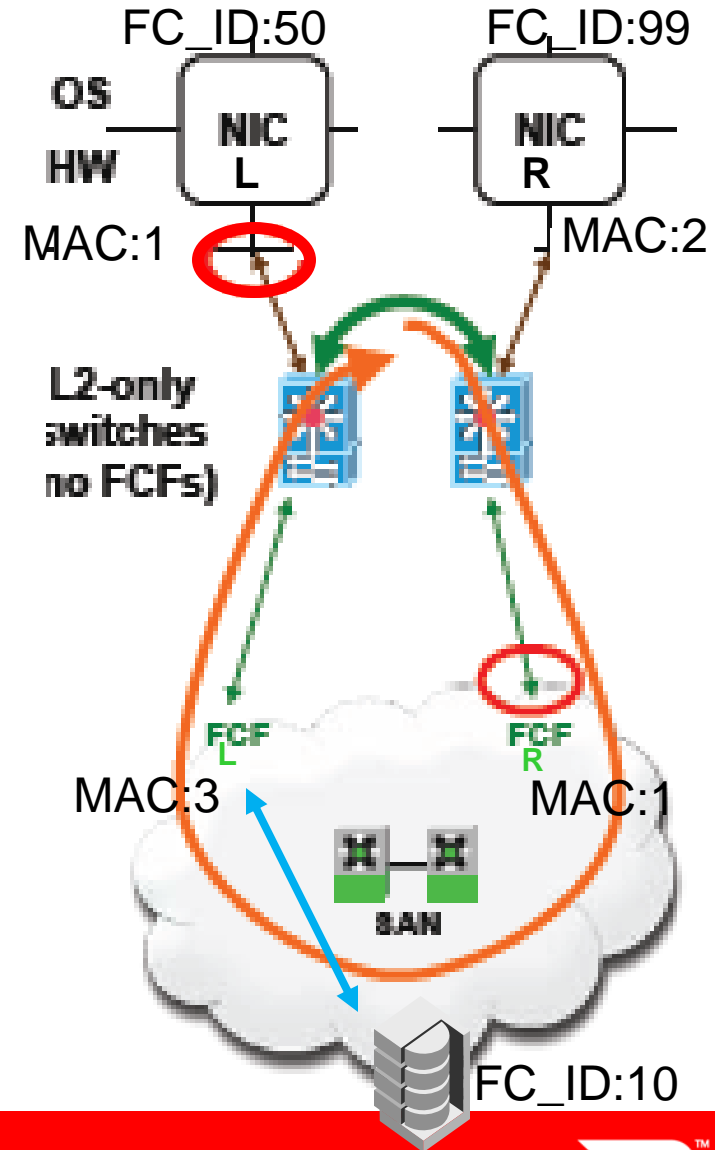
Case 4: Looping prevented by Rule I2/I3

1. The wire is Plugged wrong
2. FCF-L with MAC:3 instantiates a VE_Port $\leftarrow \rightarrow$ VE_Port relationship with FCF-R with MAC:1
3. NIC-L attempts to instantiate a VN_Port $\leftarrow \rightarrow$ VF_Port relationship with FCF-L
 - MAC:1 $\leftarrow \rightarrow$ MAC:3
 - According to [Rule I2/I3](#), deny the instantiation, issues an alert, and shut down the MAC:3 or the other Vx_Ports, and do not permit an instantiation with MAC:1 until Admin permits
4. NIC-R with MAC:2 instantiates with FCF-R
 - MAC:2 $\leftarrow \rightarrow$ MAC:1
5. A Frame leaves NIC-R (MAC:2) headed for FCF-R (MAC:1)
6. Frame gets sent to NIC-L
7. NIC-L does not have any instantiation with MAC:2
 - According to Rule A/B discard the frame, issue an Alert, **(and shut down the NIC?)**
 - **No Corruption on Data writes**
8. Then Assume that a a Read gets sent from NIC-R and it goes to FCF-R correctly
9. A returning frame from FCF-R to NIC-R
10. There is no Looping or corruption since it is prevented because no I/O is flowing from NIC-L (MAC:1)



Case 5: Looping prevented by Rules I1 & I2/I3

1. The wire is Plugged wrong
2. FCF-L with MAC:3 instantiates a VE_Port \leftarrow \rightarrow VE_Port relationship with FCF-R with MAC:1
3. NIC-L attempts to instantiate a VN_Port \leftarrow \rightarrow VF_Port relationship with FCF-L and FCF-R
 - MAC:1 \leftarrow \rightarrow MAC:3 & MAC:1 \leftarrow \rightarrow MAC:1
 - According to [Rules I1, & I2/I3](#), deny the instantiation, issue an alert, and shut down the MAC or the other Vx_Ports, and do not permit an instantiation with MAC:1 until Admin permits
4. NIC-R instantiates VN_Port \leftarrow \rightarrow VF_Port relationships with FCF-R and FCF-L
 - MAC:2 \leftarrow \rightarrow MAC:3 & MAC:2 \leftarrow \rightarrow MAC:1
5. Read Data Frame will be set to NIC-R without issue
6. Looping is prevented but I/O is flowing only to NIC-R



Now lets make it harder

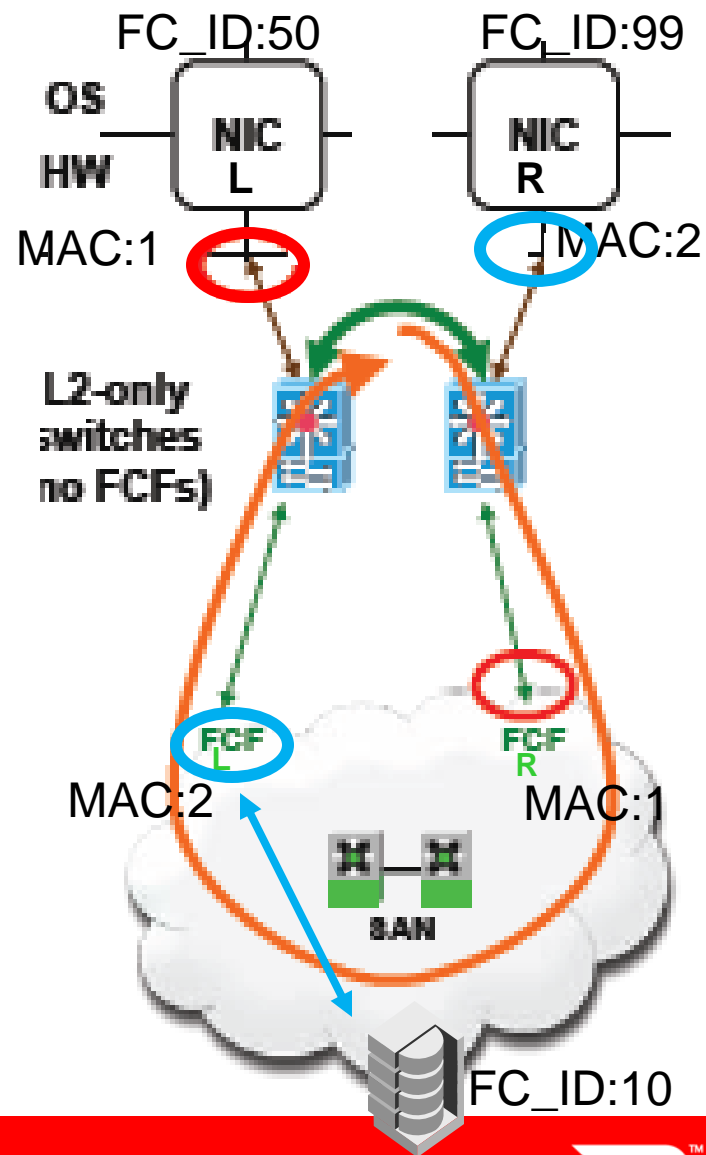
Two pairs of NICs have duplicate MAC addresses

1. NIC-L has a duplicated MAC address with FCF-R
2. NIC-R has a duplicated MAC address with FCF-L

Now lets have a cabling mistake

Case 6: Looping prevented by Rule FC1 & FC2

1. NIC-R with MAC:2 also Equals FCF-L with MAC:2
2. FCF-R instantiates with NIC-R as a VN_Port \leftarrow \rightarrow VF_Port
 - MAC:2 \leftarrow \rightarrow MAC:1
3. NIC-L instantiates as a VN_Port \leftarrow \rightarrow VF_Port relationship with FCF-L
 - MAC:1 \leftarrow \rightarrow MAC:2
4. The wire is Plugged wrong
5. Commands & Data may not even reach the FCFs (will be reflected to the other NIC and then Rule FC2 will apply and reject the frame)
6. A Frame leaves FCF-L with SA=2 & DA=1 & FC S_ID:10 & D_ID:50
7. Frame gets transferred to FCF-R (because it has a MAC:1)
8. Frame will be received by the MAC on FCF-R since it has an instantiation with MAC:2 as an N-Port to F-Port.
9. The Frame will be stripped of its Ethernet and FCoE headers and given to the VF_Port FC Entity
10. The VF_Port FC Entity will not have been instantiated with the N_Port that is represented by the S_ID of the FC Frame that it receives
 - According to Rule FC1, discard the frame, send an Alert, etc.
11. Looping is prevented



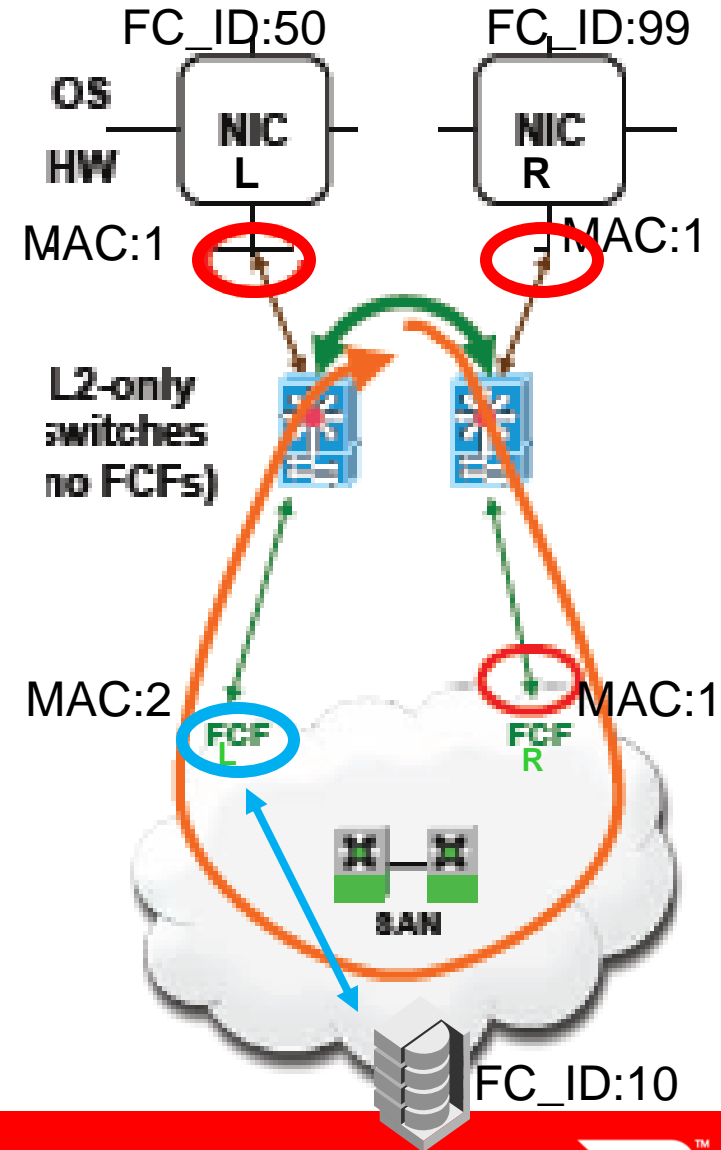
Now lets make it even harder

1. Two NICs have duplicate MAC addresses
2. And they are duplicated MAC address with FCF-R

Now lets have a cabling mistake

Case 7: Looping prevented by Rules I1, A1, & A2

1. NIC-R with MAC:1 has same MAC Addr as NIC-L and FCF-R
2. NIC-R instantiates with FCF-R
 - Instantiate blocked because of [Rule I1](#), & an Alert is issued, etc
3. NIC-L instantiates as a VN_Port ← → VF_Port relationship with FCF-L
 - MAC:1 ← → MAC:2
3. The wire is Plugged wrong
4. A Frame leaves FCF-L with SA=2 & DA=1 & FC S_ID:10 & D_ID:50
5. Frame gets transferred to FCF-R (because it has a MAC:1)
6. FCF-R will not have been instantiated with any NIC or FCF with a MAC:2
 - According to [Rule A2](#) Discard the frame, and issue an Alert, etc
 - Looping will be prevented
7. Or if Frame gets transferred to NIC-R (because it has a MAC:1)
8. NIC-R will not have been instantiated with FCF-L
 - According to [Rule A1](#) Discard the frame, and Issue an Alert, etc
9. Looping is prevented



Value of Unused Port ACLs the FCoE Header

The Looping errors described above can also be avoided by Preventing FCoE related Frames from erroneously being sent from one Fabric to another

- ACLs with
 - Ethertype=FCoE Deny and
 - Ethertype=FIP Deny
 - Will prevent Data and Discovery Frames from entering the wrong fabric
- Can be added to top of any existing ACLs, and removed simply
- Can be dynamically updated when appropriate

Summary

If vendors following normal error checking rules, that almost all vendors will need to implement, they will be able to prevent Data corruption issues on either incoming data or outgoing data.

These same normal error checking rules will prevent Data Loops from occurring when things are erroneously cabled

The Rules are natural and are probably nothing that would not be implemented anyway

Even with duplicate Host NIC MAC Addresses, data corruption can be avoided and Error alerts issued as appropriate, along with MAC shut down that requires Admin intervention

If open Ethernet Switch Ports have an ACLs that just Denys all FCoE related traffic, then all the erroneously wiring Crosstalk problems can be avoided, independent of these checking rules

Probably having the implementations follow the example Rules, and recommending “Best Practices” ACLs with the “Eth=FCoE Deny” and Eth=FIP Deny” action for open ports, is a good “Belt & Suspenders” type of approach (also applies to any switch interconnections that are not intended to pass FCoE related frames – e.g. ports used for IP interconnects)

Additional Backup Slides

What about the Software FCoE in a VM

Though this is a problematical approach this will probably occur

- Probably used for testing etc
- Unlikely to be used for production
- It is incredibly rare that VMware produces duplicate MAC addresses
- Address duplication can be avoided by Administrative overrides
- If administrative action is not taken, the following may occur
 - IP Messages will be hosed up
 - The FCoE FIP discovery will be hosed up
 - FCoE data flow can be disrupted, unless the Checking Rules are used
- Probably only other Test Software-FCoE will be impacted
 - Number is small enough that it is easy for Admin to handle
- But even then, Data Corruption is prevented by the checking techniques shown in these slides

