Fibre Channel SAN
основы

Сергей Целиков
Системный инженер SAN,
+7 916 860-2579, sergey.tselikov@broadcom.com
Июнь 2020

BROADCOM®

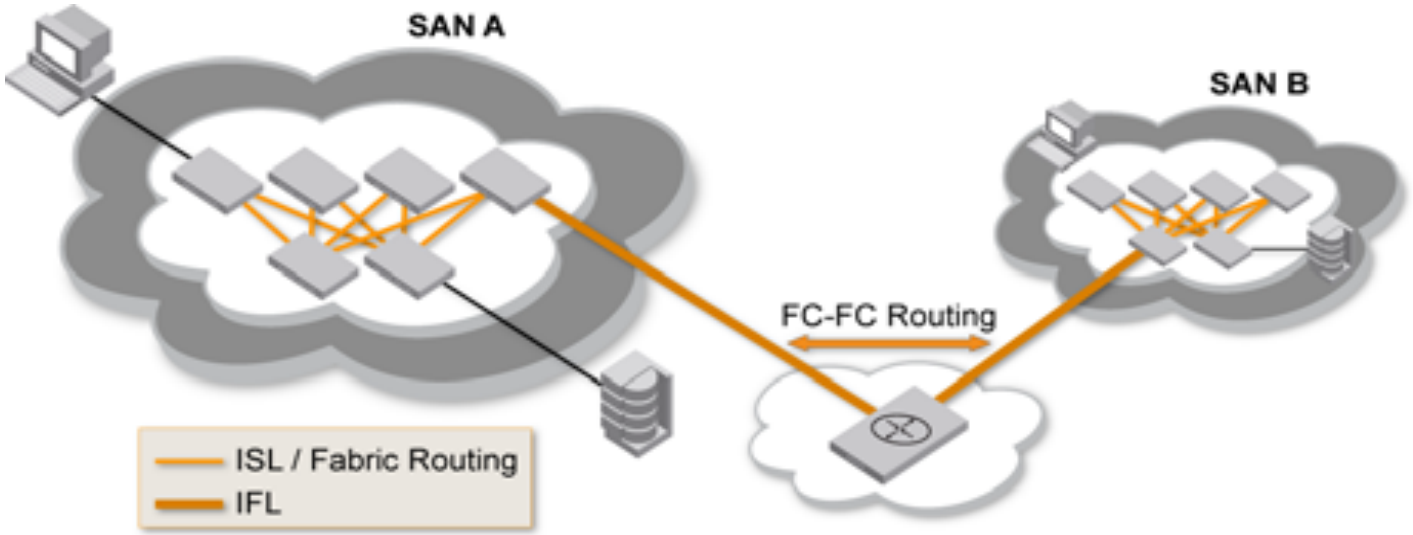
Программа

- FC Routing. Data Flow in FC Fabric.
- ISL Trunking.
- FC-to-FC routing.
- Virtual Fabrics.
- Access Gateway.
- FICON.
- Extension Solutions.
- Licensing.
- SAN Design Best Practices.

FC Routing
Data Flow in FC Fabric

Routing – Overview

- Fabric Routing is logic used by a switch to pass frames from the source domain toward the destination domain
- This module focuses on Fabric Shortest Path First (FSPF) routing
  - Also referred to as ‘Layer 2’ routing
- There is a separate discussion on FC-to-FC routing
  - Also referred to as ‘Layer 3’ routing



SAN Basics

LAN / SAN Comparison

LAN

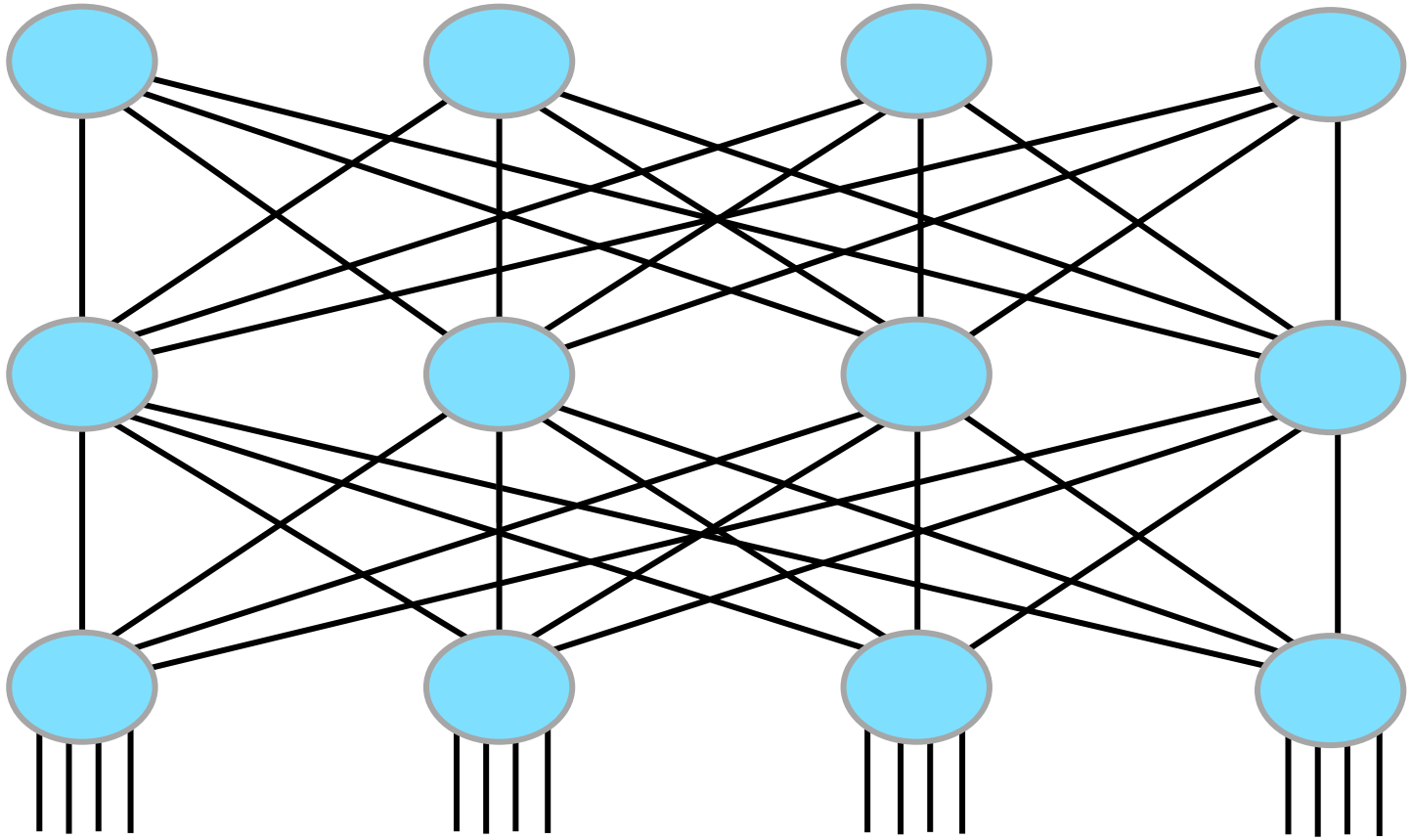
- In a layer 2 LAN, you have to worry about loops.

SAN

- Since its inception 25 years ago, Fibre Channel has incorporated layer 3 style intelligence at layer 2 – ergo, no loop concerns at all.
- A protocol similar to OSPF is used in layer 2 FC called “FSPF”, or “Fabric Shortest Path First” and it is an open standard (ANSI T11 fc-sw-5).
- Effortless L2 equal-cost multi-pathing

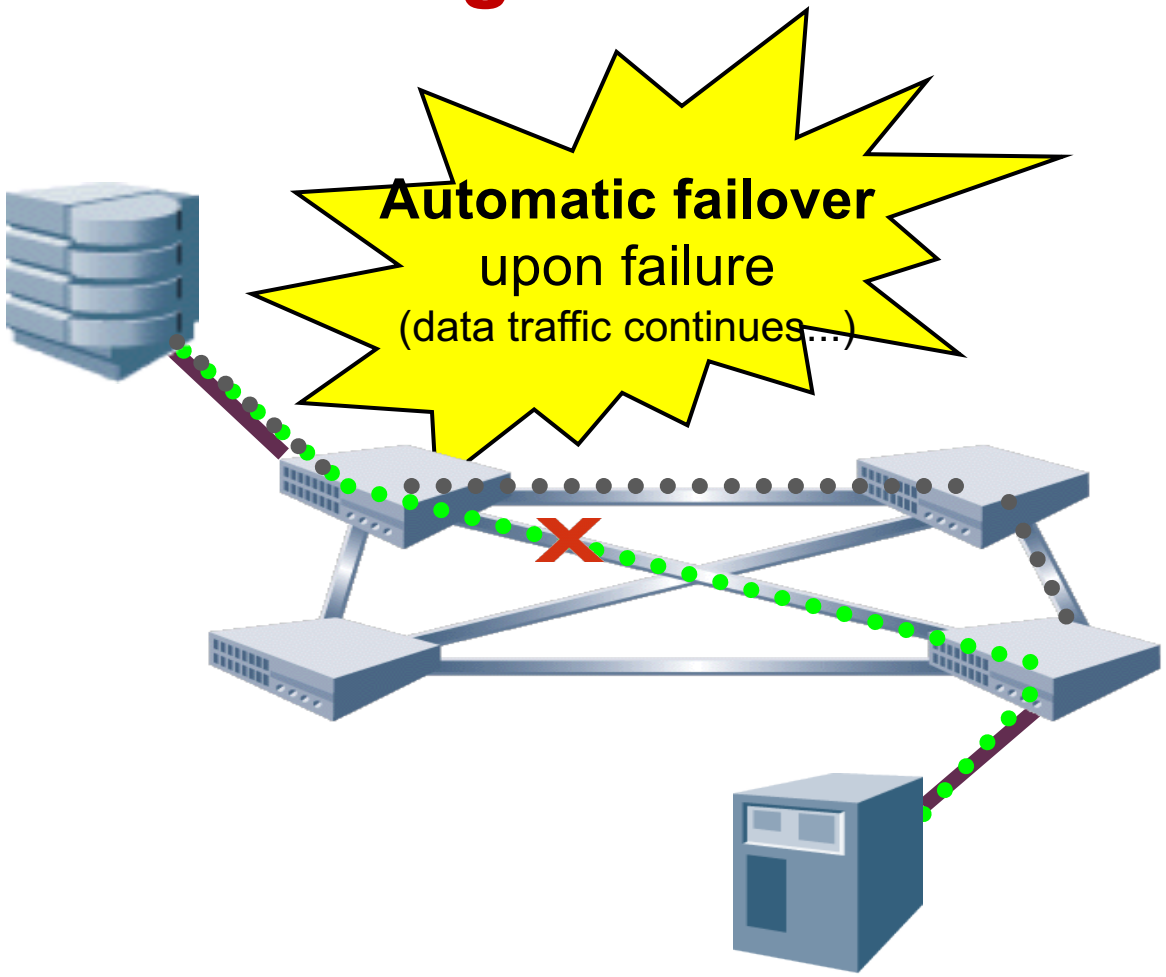
BROADCOM

FC SAN = All Links Active And Forwarding

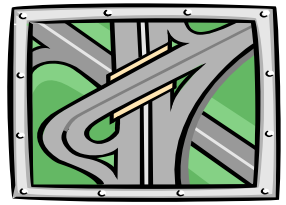


BROADCOM

Self-Healing SAN Fabric



SAN fabric remains operational when ISL or switch fails (or reboots)



Fabric Shortest Path First

- FSPF is the magic that glues switches together and forms a protected network.
- Automatic fabric topology **discovery** and path selection.
- Dynamically configures routes when SAN fabric configurations change (**self-learning**).
- Can apply static routing.
- Automatically identifies the **next best** path upon failure.
- FSPF protects the network only; node protection requires dual paths into the SAN fabric.

BROADCOM

Fabric Terminology

- Inter-Switch Links (ISLs)
  - E_Port-to-E_Port links
  - Communicate using Class F service
- Principal switch
  - Selected when the fabric initializes, before routing is established
  - Manages the assignment of unique domain IDs
  - Provides time synchronization to all other switches in the fabric
- Principal ISL
  - ISL used to communicate between the principal switch and other switches in the fabric

Principal Switch; Principal ISL's

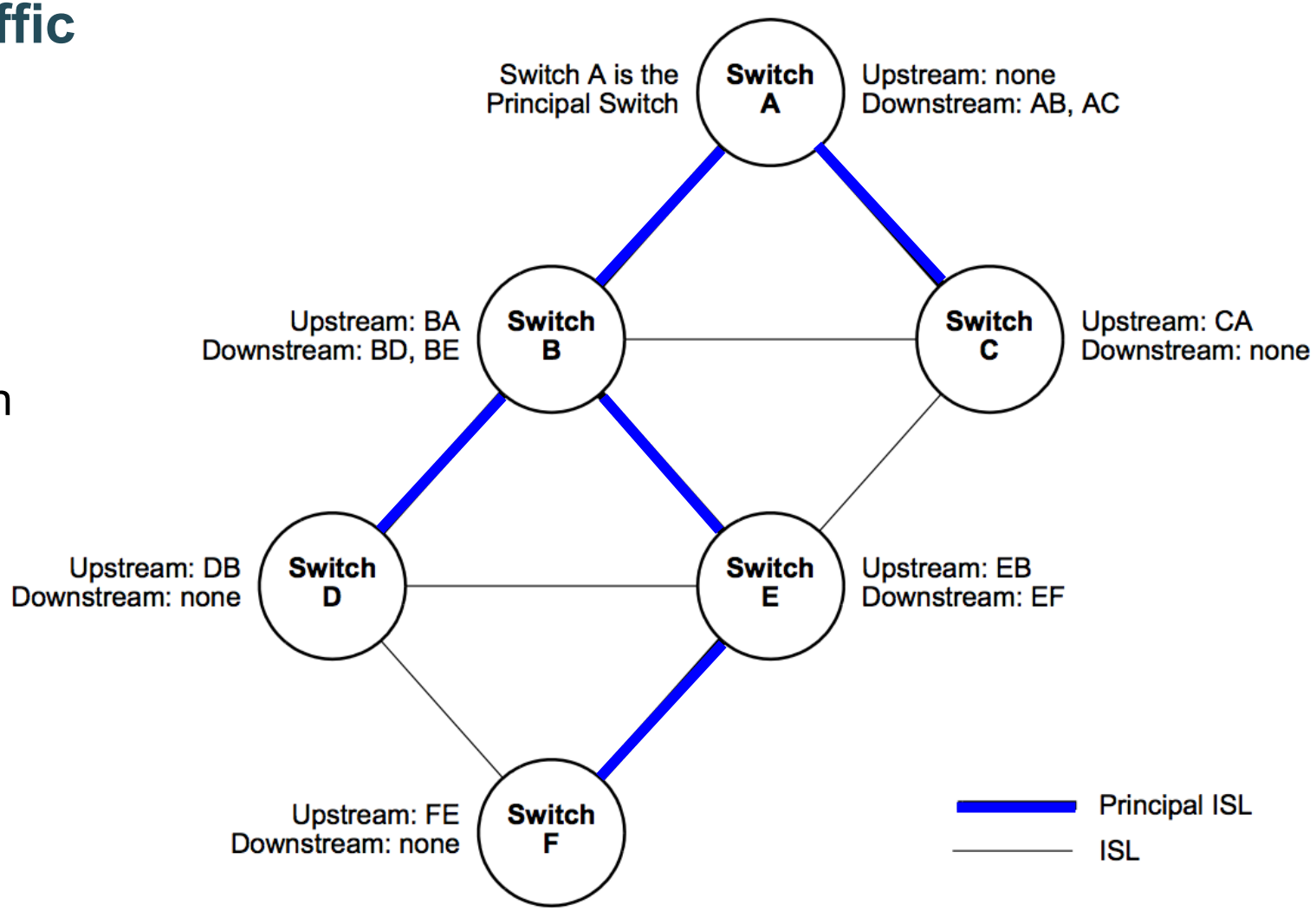
Special Paths for Class F Traffic

Principal Switch:

- Ensures No Domain ID Conflicts
- Time Sync, etc.

Principal ISL's

- Establish a Path to Principal Switch (Principal Links)
- Used for FSPF link-state updates, etc.
  - Note: SAN traffic can use all links, these links are only special because they are designated for fabric-stabilization traffic



BROADCOM

Principal Switch Commands

- Set the preferred principal switch priority using:

  `SW1:admin>` **`fabricprincipal --enable -p 0x01 -f`**

  `-p` Sets the principal selection priority for the switch

  `-f` Forces a fabric rebuild immediately after enabling on a switch

- Selection process:
  - If none of the switches has a priority setting, switch with the lowest WWN becomes principal
  - If switches have a priority setting, then only those switches participate in the selection
  - A switch with the lowest priority becomes the principal switch
  - If more than one switch has lowest priority, the switch with the lowest WWN becomes the principal switch

```
RSL_SWT121:admin> fabricshow
Switch ID   Worldwide Name           Enet IP Addr    FC IP Addr      Name
-------------------------------------------------------------------------
  2: fffc02 10:00:00:60:69:80:04:5e  10.255.255.121  0.0.0.0         "RSL_SWT121"
129: fffc81 10:00:00:60:69:80:05:1c  10.255.255.129  0.0.0.0         "RSL_SWT129"
153: fffc99 10:00:00:60:69:50:0d:d6  10.255.255.153  0.0.0.0       >"RSL_SWT153"
157: fffc9d 10:00:00:60:69:51:2d:57  10.255.255.157  0.0.0.0         "RSL_SWT157"
```

Indicates the Principal Switch

BROADCOM

Fabric Shortest Path First (FSPF)

- Calculates the minimum cost path from switch-to-switch
- Downloads the routing tables to the ASICs
- The ASIC routes frames using the lowest cumulative costs of all available links
- Path vs. Route
  - Paths are all the possible ways to get from one switch to another
    - Each ISL has a metric cost
    - Cumulative cost is based on the sum of the costs of all traversed ISLs
  - Routes are the least cost paths, and get put into the switch's routing table
    - Dynamic Load Sharing (DLS) will assign data across all equal cost routes in relation to the ratio of the ISL's link capacity
    - In-order Delivery (I0D) ensures ordered frame delivery when ingress ports are re-routed due to topology changes

FSPF

Determining Paths

Four Main Components of FSPF:

#1. Hello Protocol

- Establish connectivity with a neighbor Switch
- Establish the identity of the neighbor Switch
- Exchange FSPF parameters and capabilities;

#2. Replicated Link State Database

- Has protocols and mechanisms to keep the databases synchronized across the Fabric;

#3. Path Computation Algorithm

#4. Routing Table Update

BROADCOM

FSPF
Basics

The Link State Database is central to the operation of FSPF.

- It is a replicated database where all Switches in the Fabric have the same exact copy of database at all times
- The database consists of a collection of Link State Records (LSRs).

Path computation is local

- The results of the computation are not distributed to other Switches, only topology information is distributed. This is a characteristic of link-state path selection protocols.

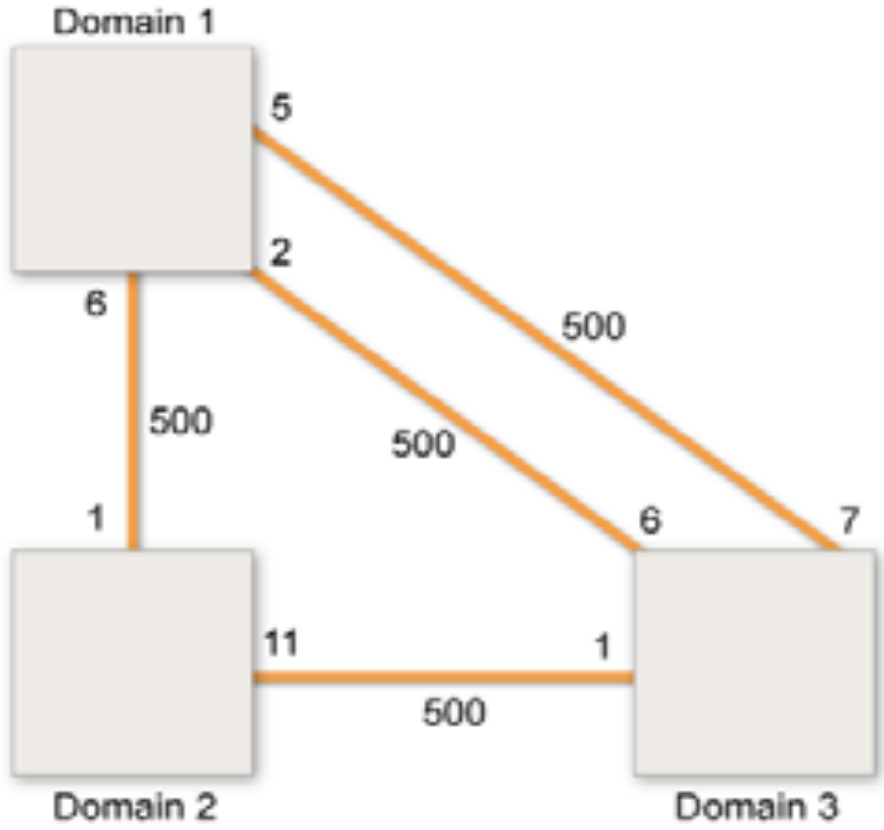
FSPF

Determining Paths

Operation	Starting Condition	Process	Ending Condition
1. Perform Initial HELLO Exchange	The Switch originating the HELLO has a valid Domain_ID.	HLO SW_ISL frames are exchanged on the link until each Switch has received a HELLO with a valid neighbor Domain field.	Two way communication has been established
2. Perform Initial Database Exchange	Two way communication has been established.	LSU SW_ISL frames are exchanged containing the Initial database.	Link State Databases have been exchanged.
3. Running State	Initial Database Exchange completed.	Routes are calculated and set up within each Switch. Links are maintained by sending HELLOs every Hello_Interval. Link databases are maintained by flooding link updates as appropriate.	FSPF routes are fully functional.

FSPF Link Cost

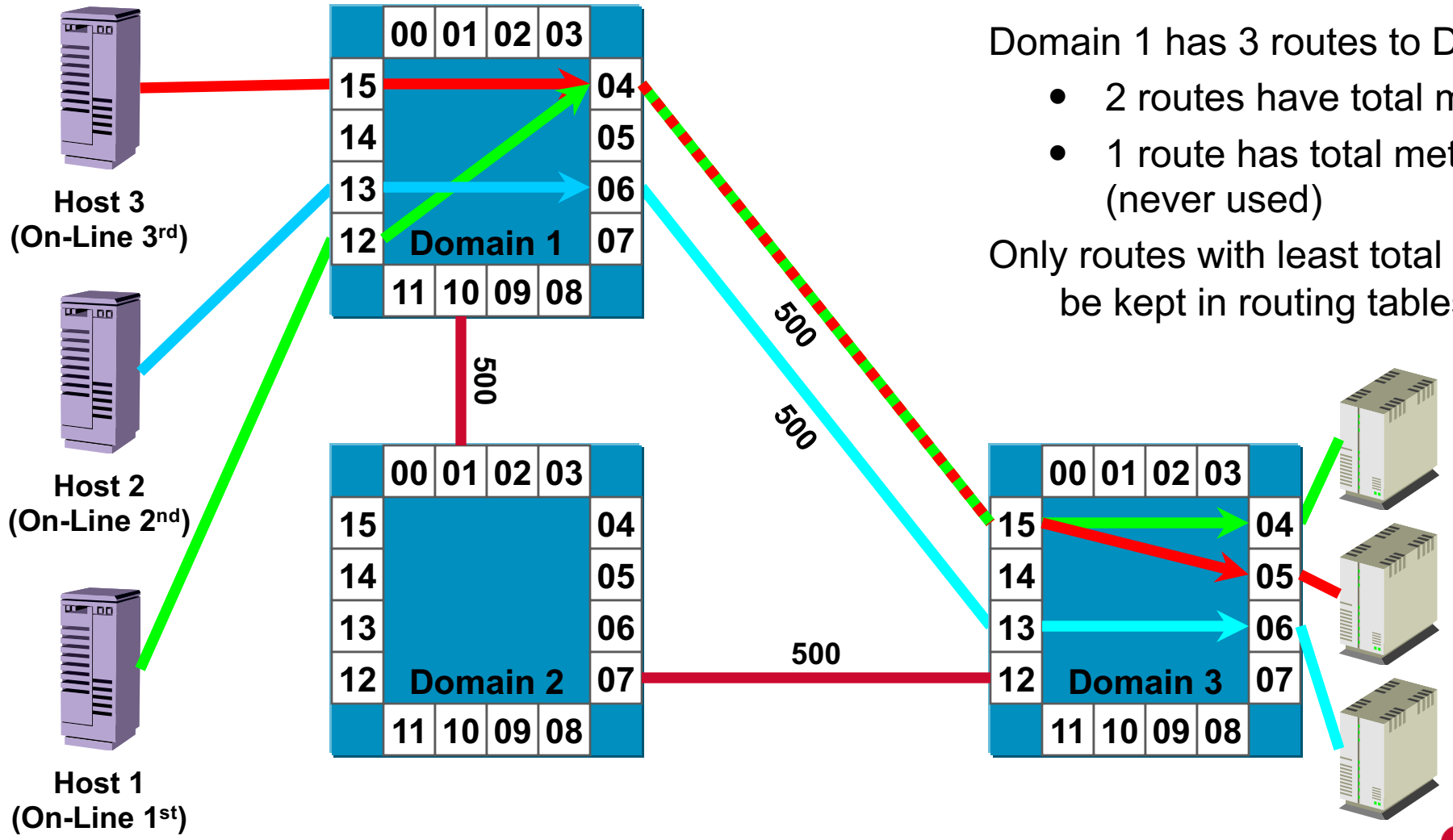
- Metric value assigned to the Tx sides on each ISL
- The Brocade default Link Cost (metric) value for 2, 4, 8, 10, 16 and 32 Gbit/sec links is 500
- Also known as a **Hop**
- Ports 2 and 5 have a cost of 500 from Domain 1 to Domain 3
- Port 6 has a cost of 1000 from D1 to D3
- Lowest cost paths routes are Ports 2 and 5
- FSPF configures the routing table in Domain 1 to only use these the routes on ports 2 and 5 for frames with a destination of Domain 3



BROADCOM

Fibre Channel Routing

- Route Selection Default Behavior



linkCost Default = “500”

Domain 1 has 3 routes to Domain 3

- 2 routes have total metric = 500
- 1 route has total metric = 1000 (never used)

Only routes with least total metric will be kept in routing tables

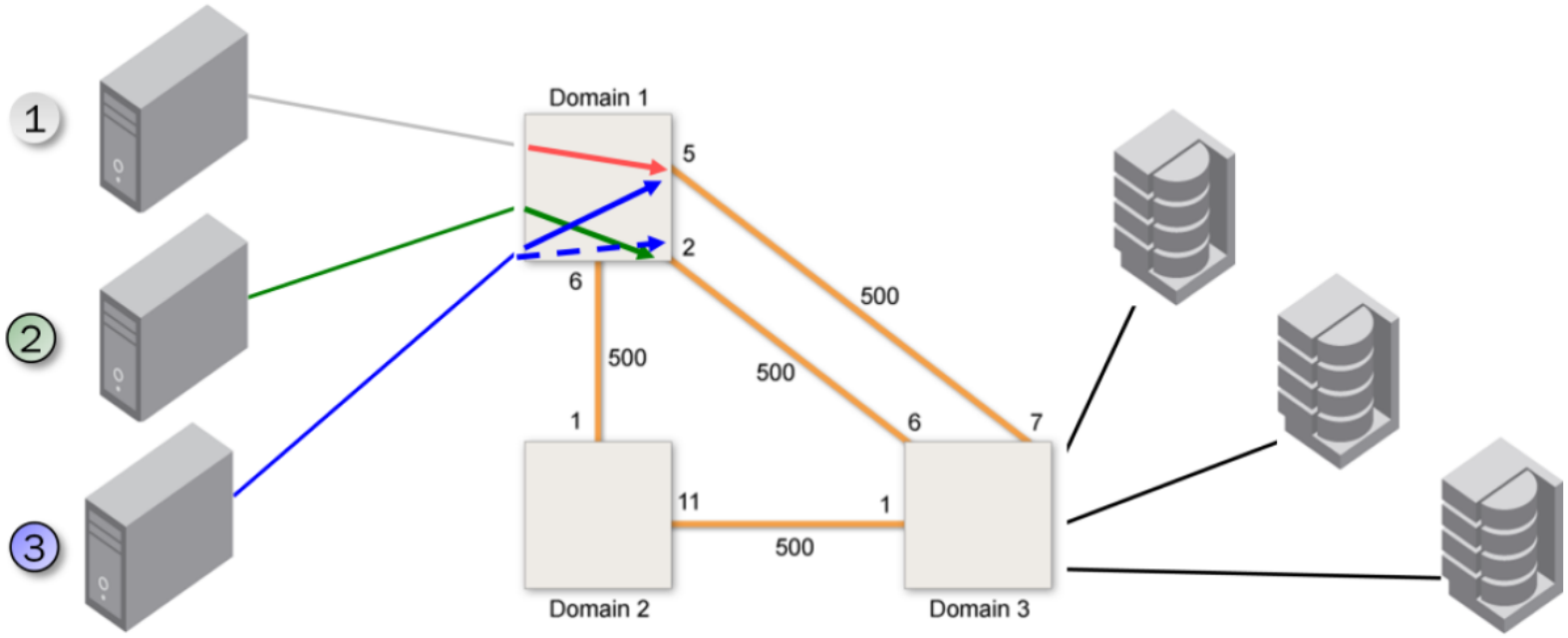
BROADCOM

Routing Policies

- The routing policy is *unidirectional* and responsible for selecting a route based on one of two user-configurable routing policies:
  - Port-based routing
  - Exchange-based routing
- Each switch has its own routing policy
  - Different policies can exist in the same fabric
- 4/8/16/32 Gbps ASICs use the FSPF protocol and either Port-based routing or Exchange-based routing
  - Exchange-based routing is Brocade's factory default setting

Port-based Routing

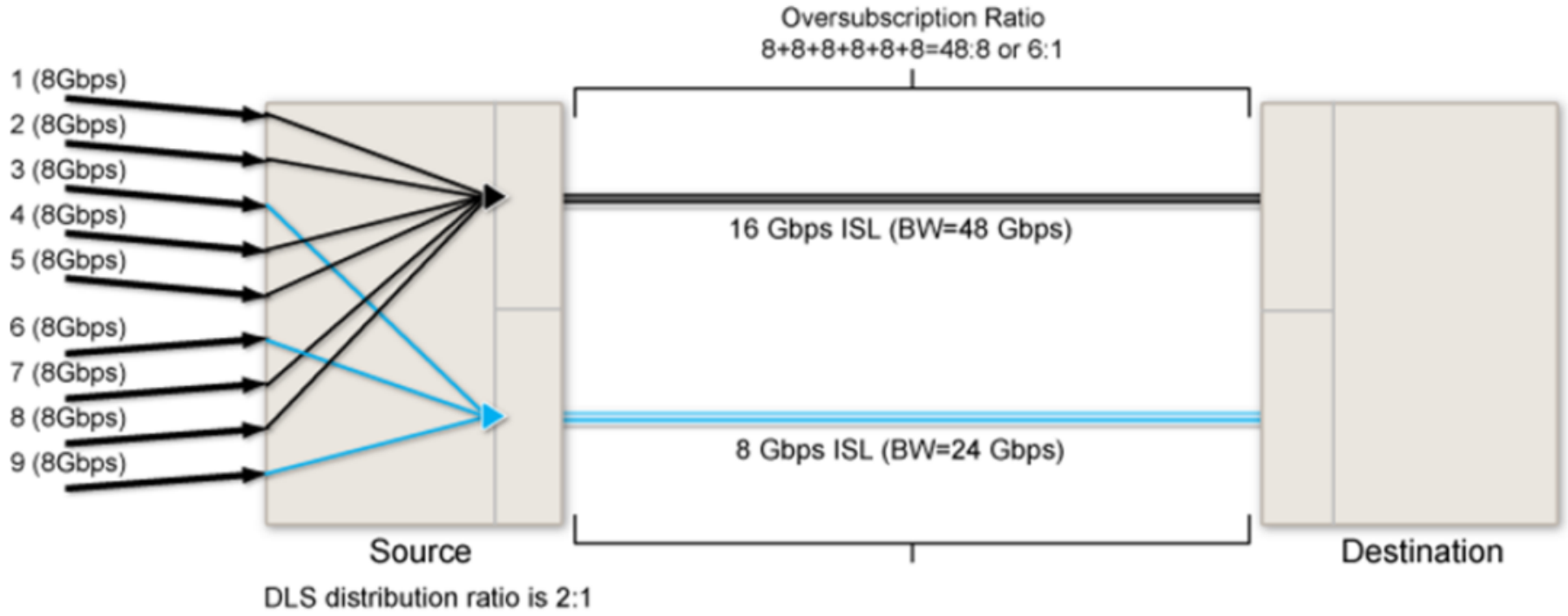
- The Link State algorithm calculates the cost of each link and determines the lowest cost path within each switch
- Input port from the source is assigned to an output port toward the destination domain, known as a route
- Routes are allocated via round-robin assignment
- Chosen routes are used until one of the devices in the fabric goes offline or the fabric changes



BROADCOM

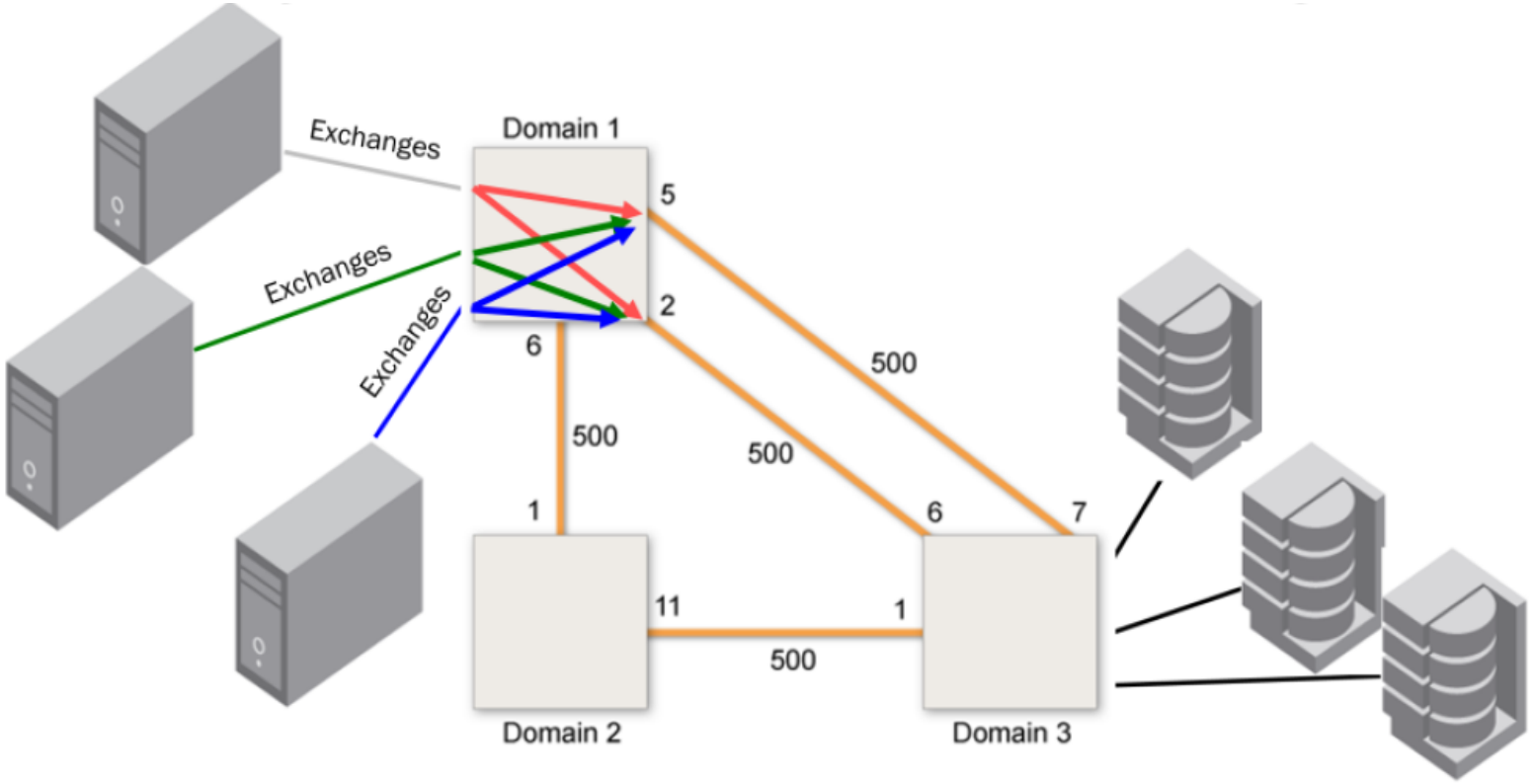
Port-based Routing and DLS

- Nine devices need to share two routes established through the Fabric:
  - Black route is a 16 Gbit/sec ISL with 500 metric cost
  - Blue route is a 8 Gbit/sec ISL with 500 metric cost
  - Ratio for DLS Distribution is 2:1



Exchange-based Routing

- Based on the Source ID (SID), Destination ID (DID), and Fibre Channel originator exchange ID (OXID)
  - Hash algorithm
- Optimizes route utilization for the best performance

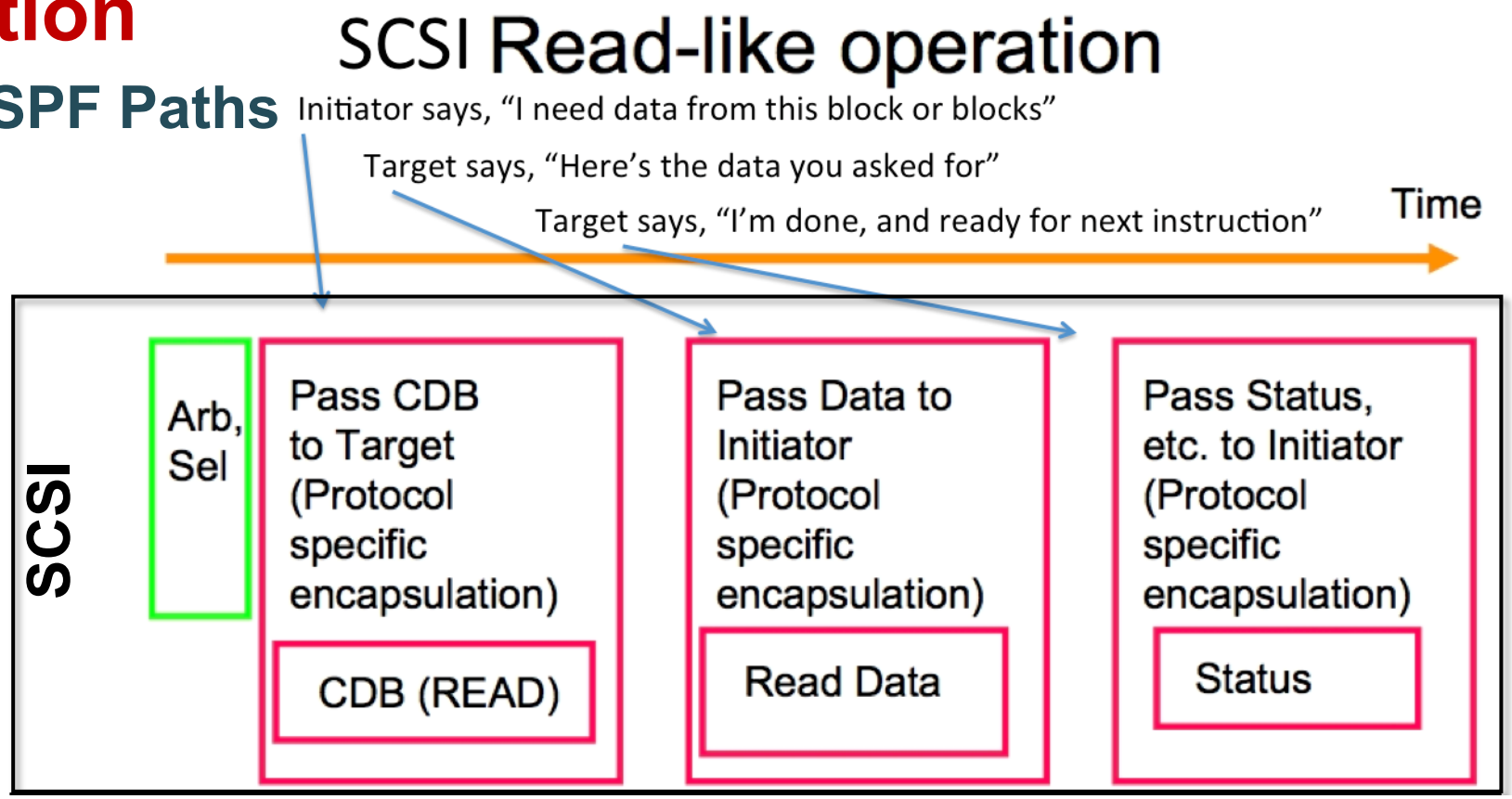


BROADCOM

Dynamic Path Selection

Sharing the Load across FSPF Paths

SCSI Commands are split into sequences of FC frames



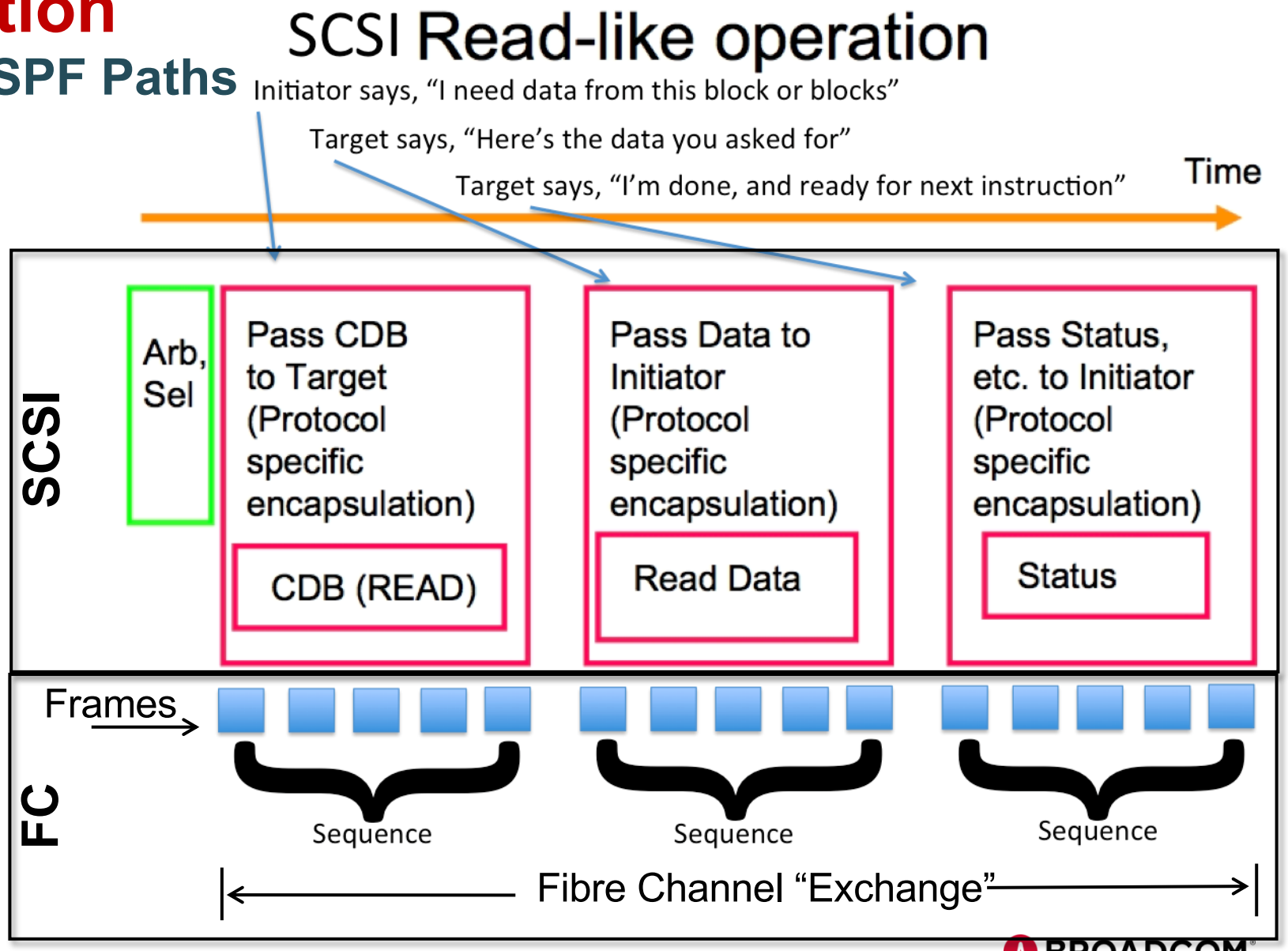
Dynamic Path Selection

Sharing the Load across FSPF Paths

SCSI Commands are split into sequences of FC frames

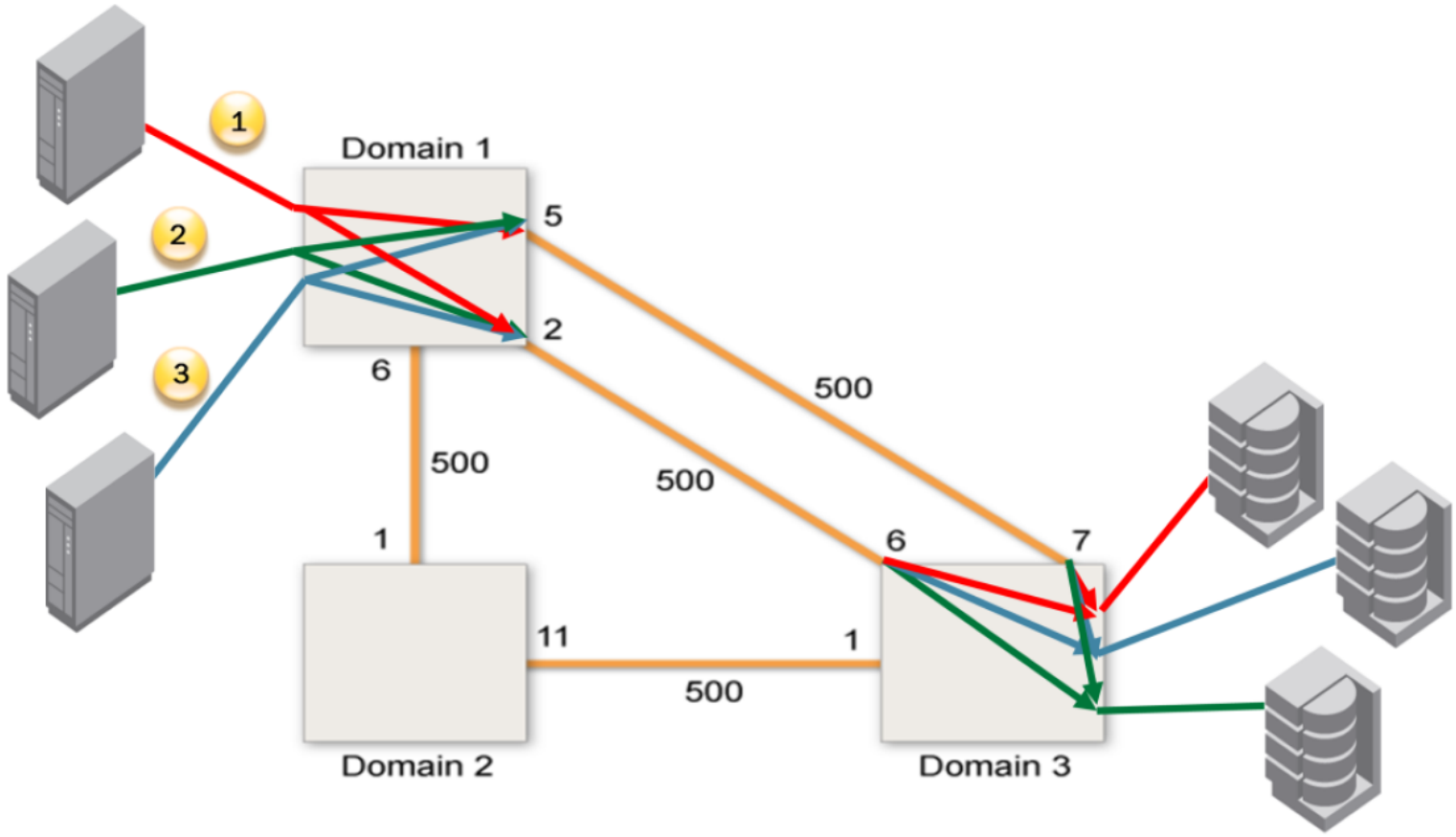
A complete SCSI command maps to a FC "Exchange"

FC Fabric load balances by hashing on these "Exchange ID's" (OXID) and spraying exchanges across equal-cost paths



Exchange-based Routing and DLS

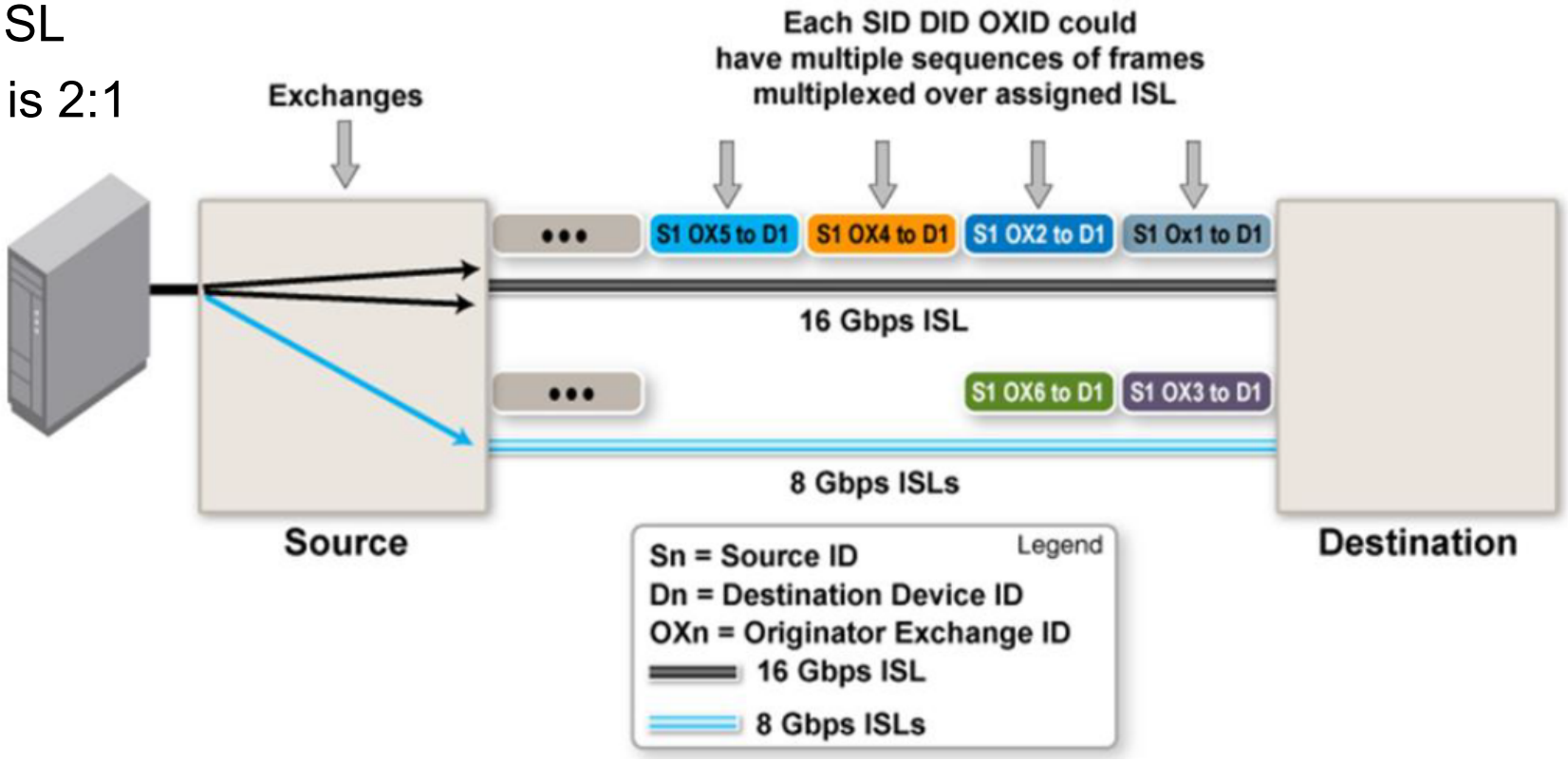
- DLS allocates a route from shortest equivalent paths based on **link cost** and **bandwidth**



BROADCOM

Exchange-based Routing and DLS (cont.)

- A single device shares two equal cost Paths established through the Fabric:
  - Black Path is a 16 Gbit/sec ISL
  - Blue Path is a 8 Gbit/sec ISL
  - Ratio for DLS Distribution is 2:1



Routing Policy Selection

- The aptpolicy command is used to change routing policies

```
R11-ST10-B51:admin> aptpolicy
   Current Policy: 3 0(ap)
   3: Default Policy
   1: Port Based Routing Policy
   3: Exchange Based Routing Policy
         0: AP Shared Link Policy
         1: AP Dedicated Link Policy
```

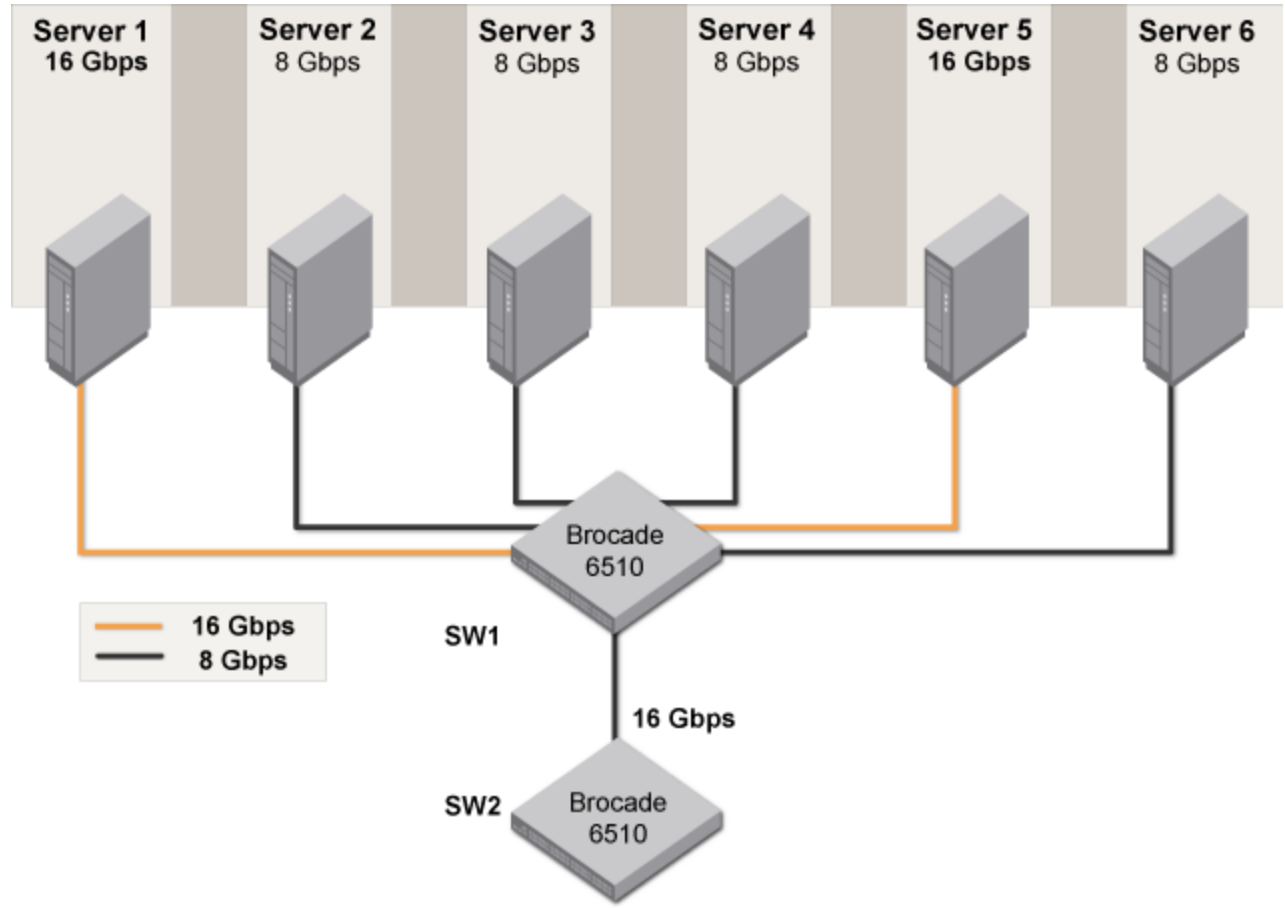
- The switch must be disabled before changing the Routing policies

Routing Terminology

- **Oversubscription**
  - The *possible* contention for bandwidth by devices through an ISL
- **Congestion**
  - The *actual* contention for bandwidth by devices through an ISL
- ISL oversubscription ratio calculation methods:
  - By port count
    - (# of Device ports) to the (# of E_Ports) expressed as 15:1, 7:1, 3:1
    - Easier to calculate, less accurate
  - By bandwidth
    - (Sum of Ingress ports bandwidth) to the (Sum of ISL Egress route ports bandwidth) expressed as 7:1, 3:1
    - More difficult to calculate, but more accurate than by port count

ISL Over-subscription

- What is the over-subscription ratio by port count? 6:1
- What is the over-subscription ratio by bandwidth? 4:1



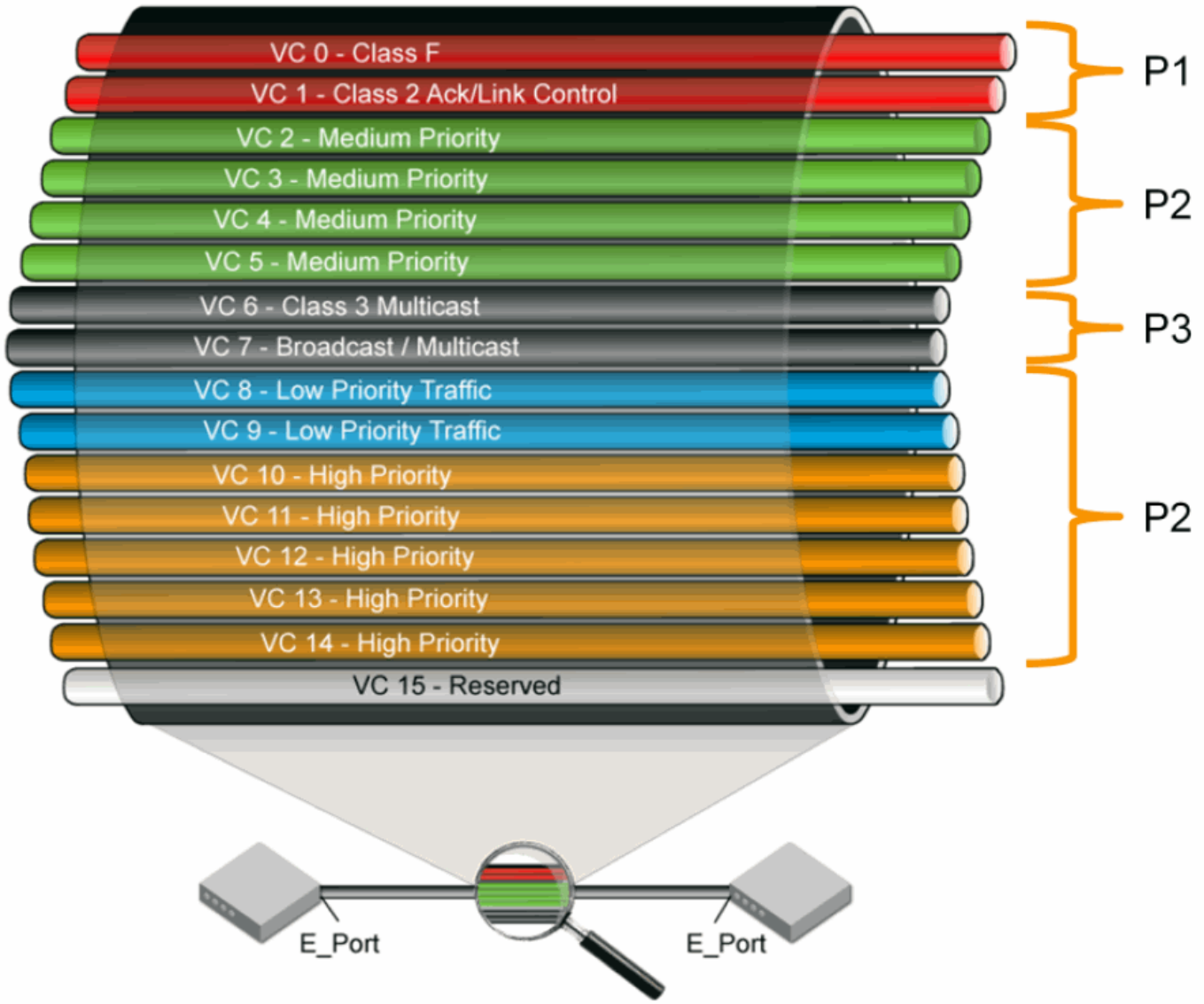
BROADCOM

Virtual Channels

- Virtual Channels (VCs) are unique to Brocade
  - Virtual Channels are transmitting queues used within a single ISL
  - Virtual channels allow the interleaving of frames on inter-switch links for non-blocking routes
- Virtual Channels are divided into 3 priority groups:
  - P1 is the highest which is used for Class F; F_RJT and ACK
  - P2 is the next highest which is used for data frames
    - The data VC channels can be further prioritized to provide higher levels of Quality of Service
  - P3 is the lowest which is used for broadcast and multicast traffic
- Condor3 support
  - VC-level credit recovery, stuck VC detection

Virtual Channels

- Virtual channels (VC) create multiple logical data paths across a single physical link or connection.
- Each VC has its own network resources such as queues and buffer-to-buffer credits assigned.
- Virtual channel technology is the fundamental building block used to construct Adaptive Networking services.

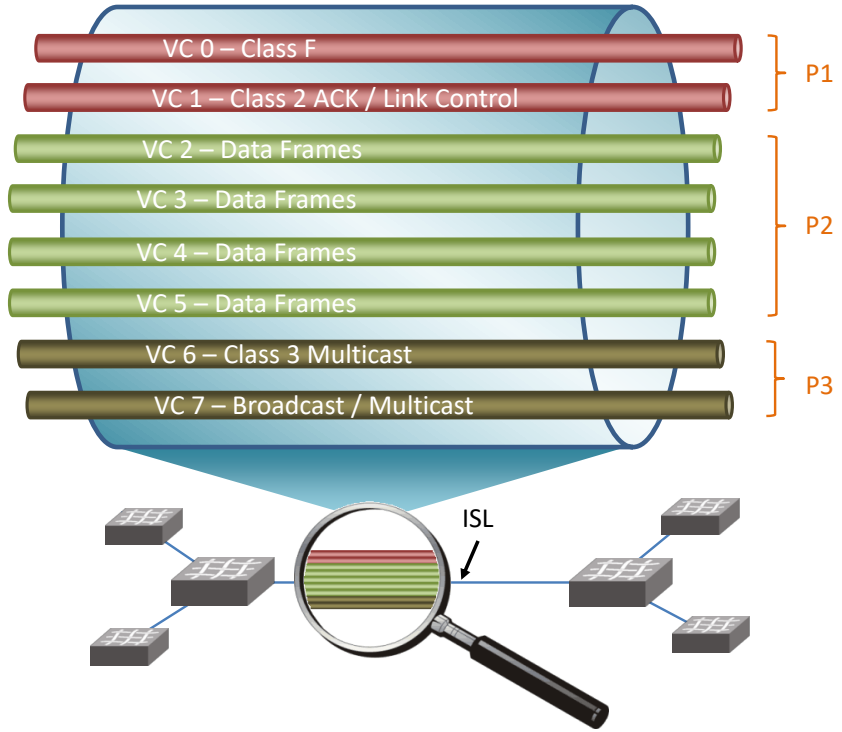


BROADCOM

Virtual Channels

- Virtual Channels are used for ISL traffic optimization
  - Each physical ISL link is partitioned into up to 16 virtual channels.
  - Priority levels enable Quality of Service (QoS) delivery.
  - Each virtual channel (VC) has an independent flow control.
  - Bottlenecks on one VC do not impact other VCs.
- Benefits
  - Reduces congestion and ensures efficient (balanced) flow of data
  - Prioritizes different types of network traffic

Medium priority traffic uses the VCs 2-5. The VCs are selected according the least significant nibble of the 2nd byte of the destination PID (010100)



VC #	Assigned to	VC Selection
VC 0	Class F	
VC 1	Class 2	
VC 2	Medium Priority QoS	Based on PID of destination (0, 4, 8, C)
VC 3	Medium Priority QoS	Based on PID of destination (1, 5, 9, D)
VC 4	Medium Priority QoS	Based on PID of destination (2, 6, A, E)
VC 5	Medium Priority QoS	Based on PID of destination (3, 7, B, F)
VC 6	Class 3 Multicast	
VC 7	Broadcast/Multicast	

BROADCOM

Virtual Channels

Basics – QoS enabled

If QoS is enabled, a different distribution will be used for QoS High and QoS Low VCs.

The medium VC distribution will be unchanged.

*Remark: QoS % distribution will be considered, if a link is fully utilized.

If there is no high/low traffic to be delivered, then all bandwidth will go to the medium VCs.

Default ISL				
VC	PID Used (DDAAPP)	Assigned to	Credits	Bandwidth
0	-	Class F	4	
1	-	Class 2 Ack/Link Control	-	
2	0, 4, 8, C	Data	5	100%
3	1, 5, 9, D	Data	5	
4	2, 6, A, E	Data	5	
5	3, 7, B, F	Data	5	
6	-	Class 3 Broadcast	1	
7	-	Broadcast/Multicast	1	

QoS Enabled ISL				
VC	PID Used (DDAAPP)	Assigned to	Credits	Bandwidth
0	-	Class F	3	
1	-	Class 2 Ack/Link Control	-	
2	0, 4, 8, C	Data - Medium	2	30%
3	1, 5, 9, D	Data - Medium	2	
4	2, 6, A, E	Data - Medium	2	
5	3, 7, B, F	Data - Medium	2	
6	-	Class 3 Broadcast	1	
7	-	Broadcast/Multicast	1	
8	Round Robin*	Data - Low	2	10%
9	Round Robin*	Data - Low	2	
10	Round Robin*	Data - High	2	60%
11	Round Robin*	Data - High	2	
12	Round Robin*	Data - High	2	
13	Round Robin*	Data - High	2	
14	Round Robin*	Data - High	2	
15	-	-	-	

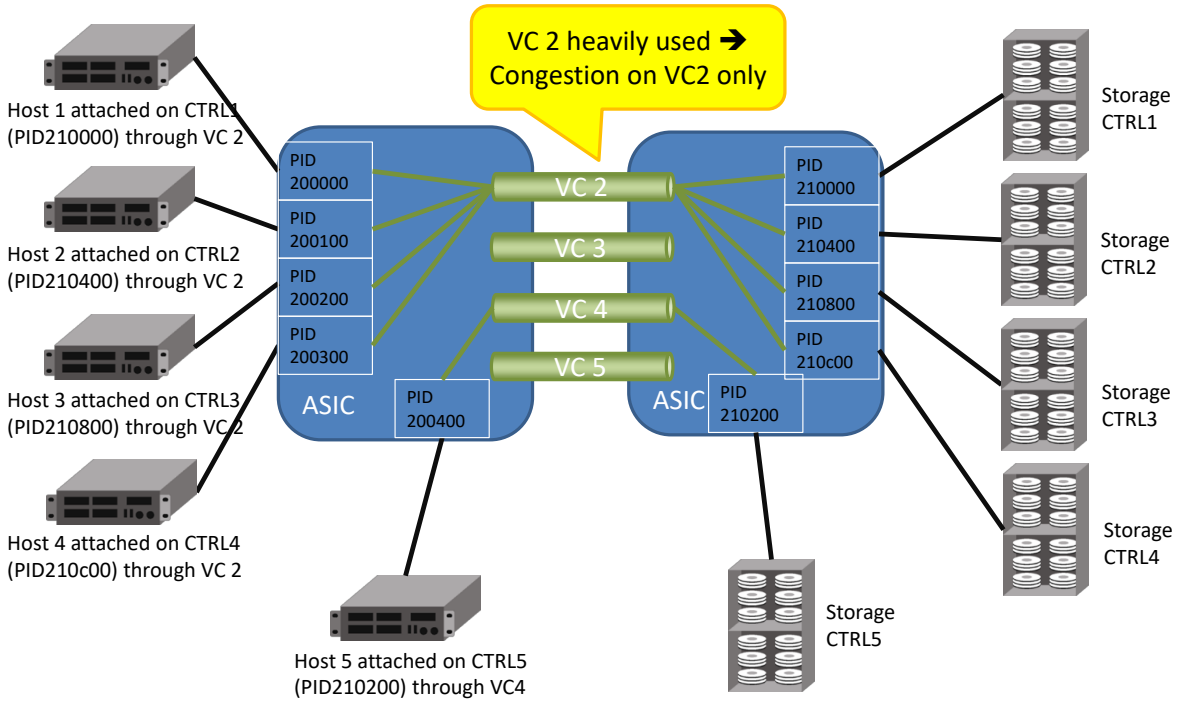
BROADCOM

Virtual Channels

VC Congestion risk...

Congestion can occur on a single “medium” VC with many high utilization devices are placed on ports sharing the same VC.

Example:



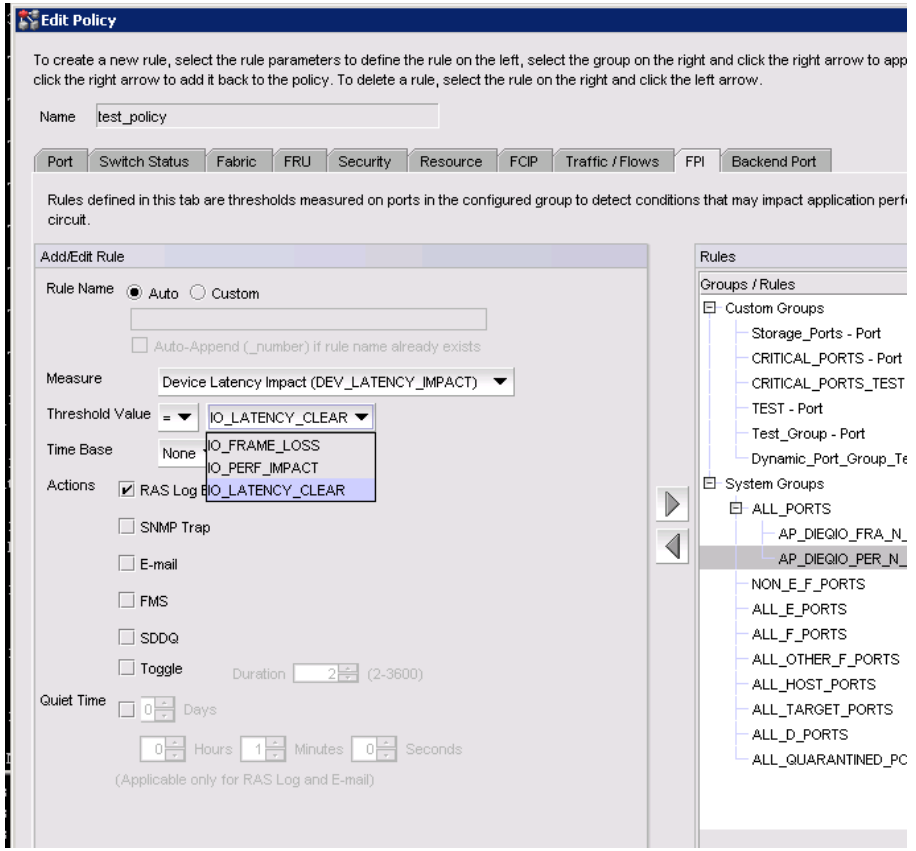
VC	PID	PID	PID	PID
VC 2	xxx0xx	xxx4xx	xxx8xx	xxxCxx
VC 3	xxx1xx	xxx5xx	xxx9xx	xxxDxx
VC 4	xxx2xx	xxx6xx	xxxAxx	xxxExx
VC 5	xxx3xx	xxx7xx	xxxBxx	xxxFxx

BROADCOM

Monitoring Fabric Performance Impact (FPI)

Performance Impact & Slow Drain Device quarantine FOS 7.4

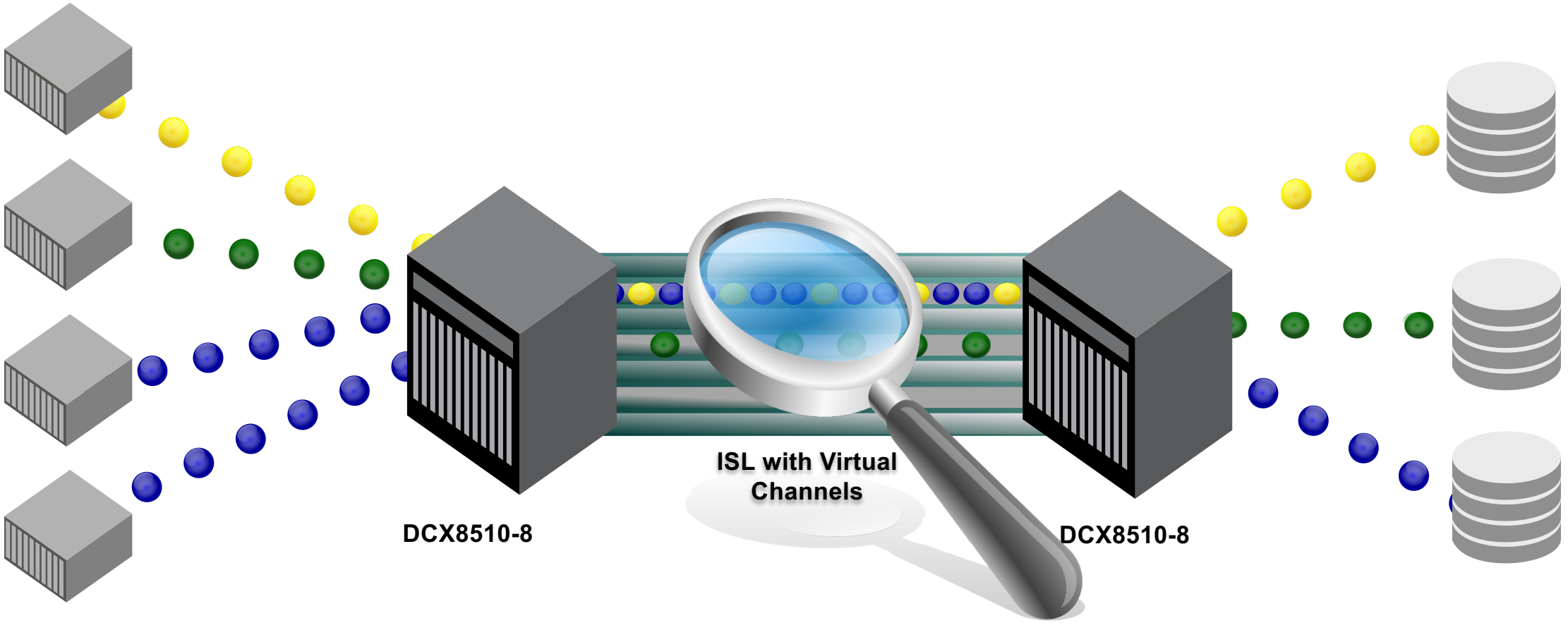
- FPI monitoring on F-Ports & **E-Ports**
- FOS 7.4 measures „Transient queue latency counter” F-Port & E-Ports
  - Measurement at 250ns cycles of TX queue on VC level
  - Frame is waiting in TX Queue > 10ms & < 80ms = IO_PERF_IMPACT
  - Frame is waiting in TX Queue >= 80ms IO_FRAME_LOSS
- FOS 7.3 & 7.4 measures on F-Ports „3 Windows method“ at 2.5us for Buffer Credit Zero Condition
- Actions **SDDQ** for F-port devices
- Action Toggle for F- & E-Ports
- Performance Impact Clear message



BROADCOM

Fabric Performance Impact (FPI & SDDQ)

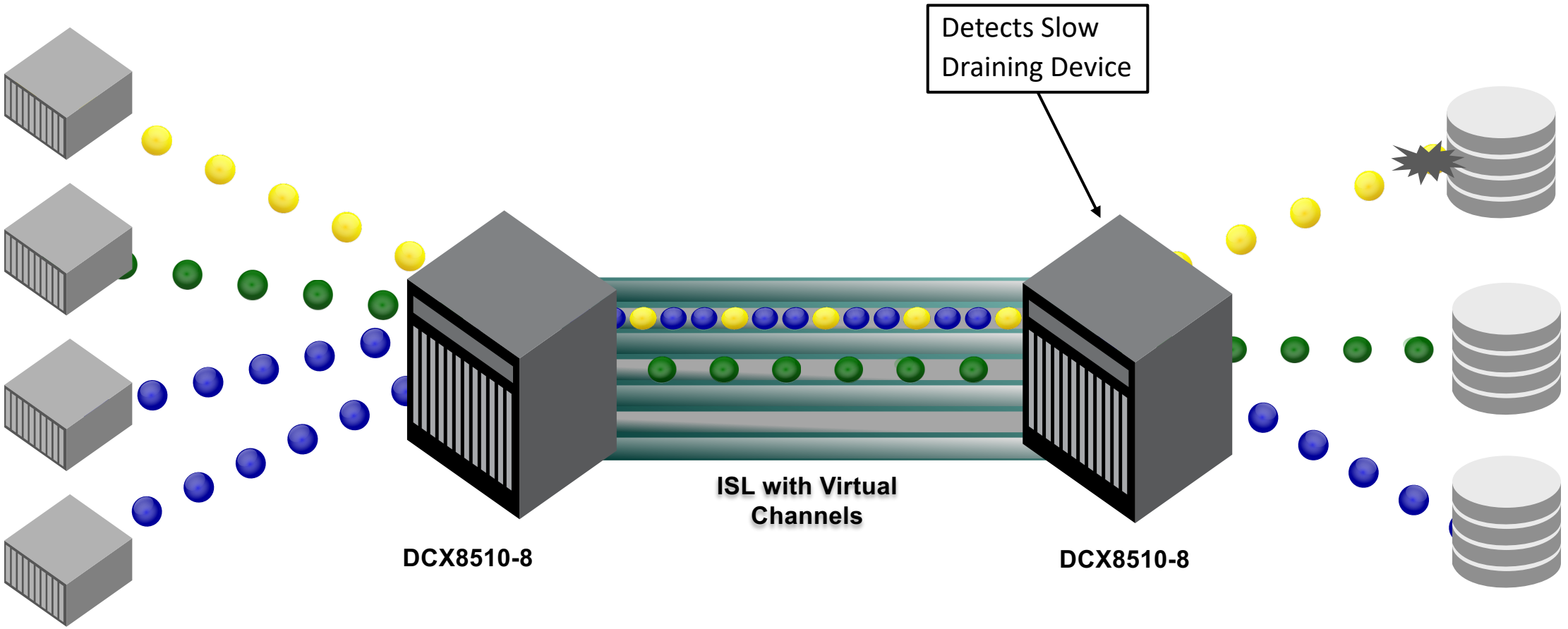
Regular operations



BROADCOM

Fabric Performance Impact (FPI & SDDQ)

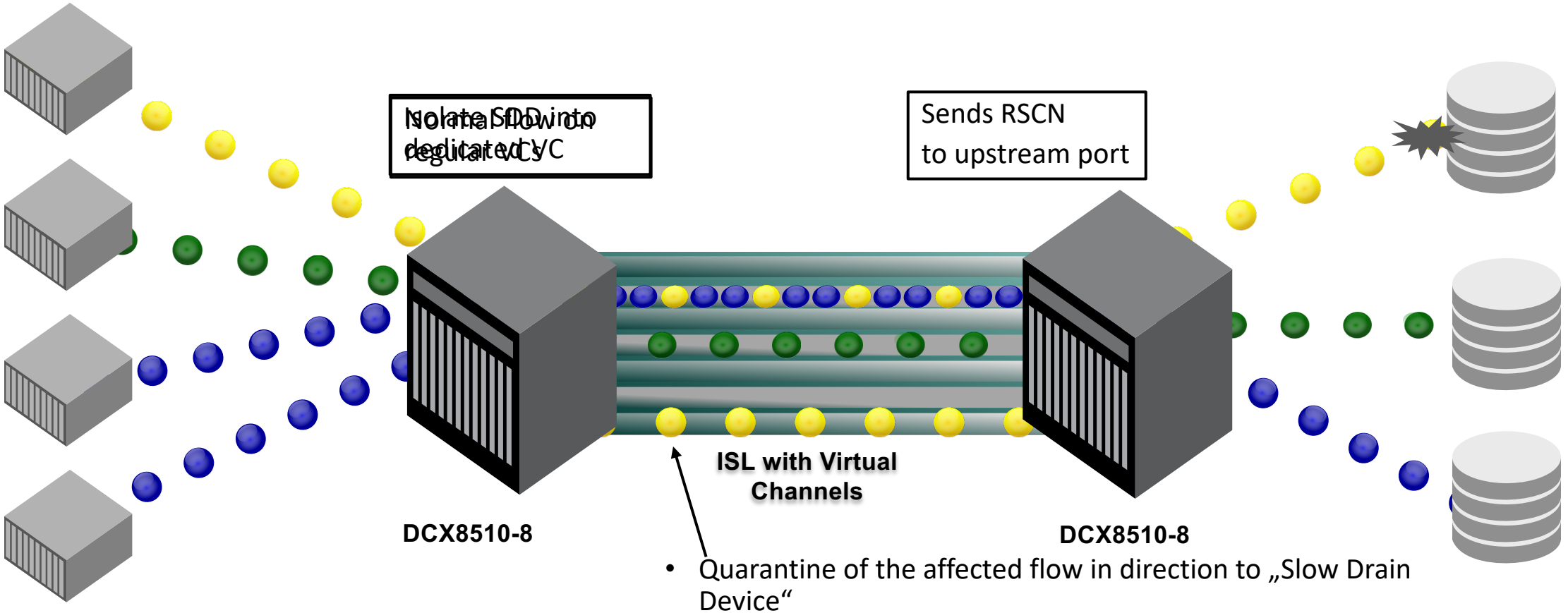
Slow drain device detection



BROADCOM

Fabric Performance Impact (FPI & SDDQ)

Slow drain device mitigation

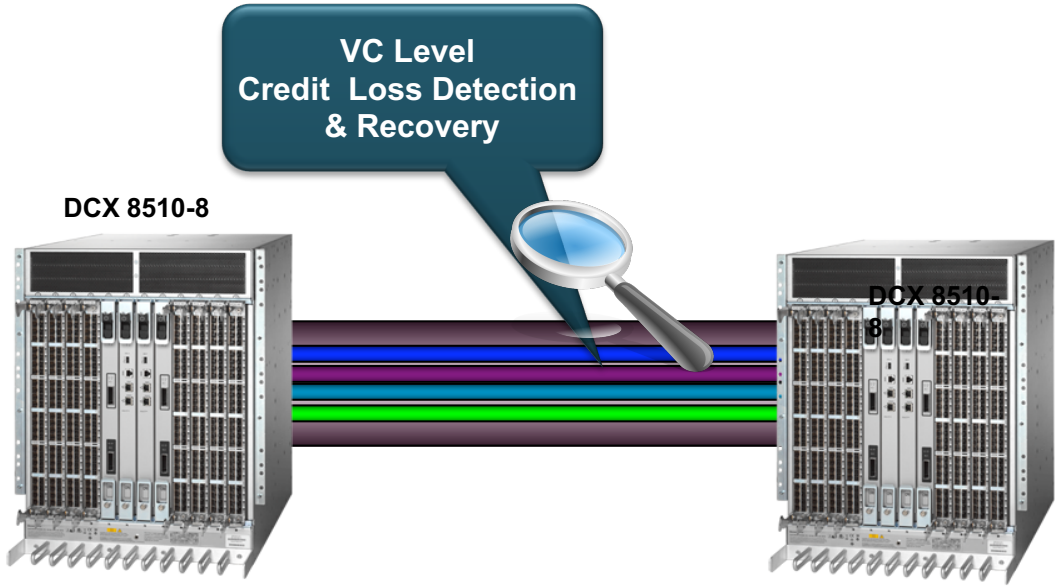


- Quarantine of the affected flow in direction to „Slow Drain Device“
- Archived by assigning the flow to a Low Priority VC
- QoS on E-Port needs to be enabled!

BROADCOM

Hardware Buffer Credit Loss Detection and Recovery Feature

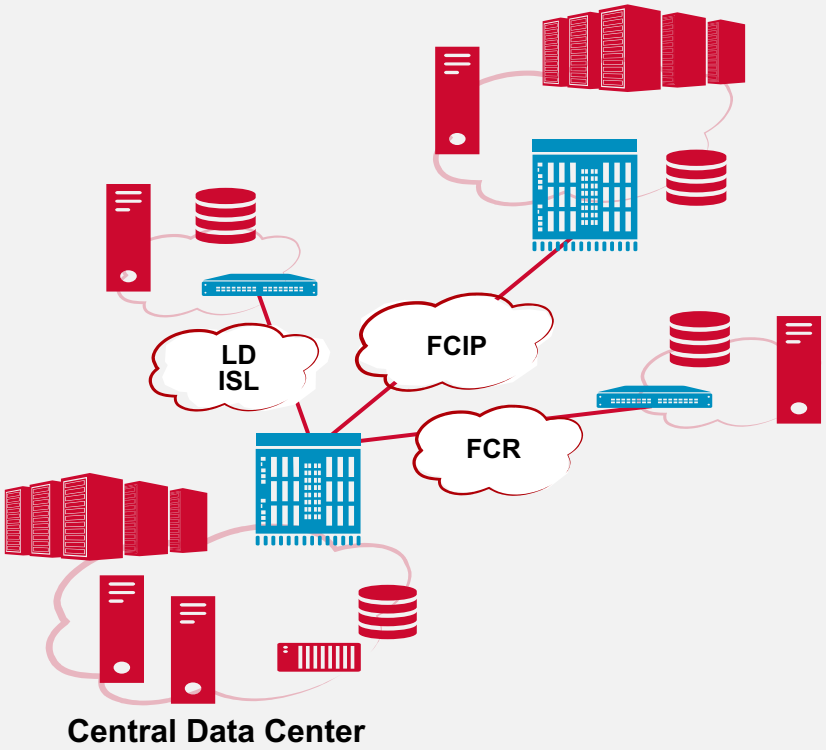
- Condor3 supports Hardware Buffer Credit Loss and Detection
  - Enabled by default
- Buffer credit loss is detected and recovered automatically at the **Virtual Channel** (VC) level on Condor3 ISLs
  - Both sides of the link must be Condor3
  - Supported on normal and long distance ISLs
  - Port-level credit recovery is supported on Condor2 ISLs



BROADCOM

Advanced Switch Features and Tools

- **Access Gateway**
- **Fabric Extension**
  - Long Distance (LD) ISLs
  - Fibre Channel over IP (FCIP)
- **Fibre Channel Routing (FCR)**
- **Trunking**
- **Virtual Fabrics**

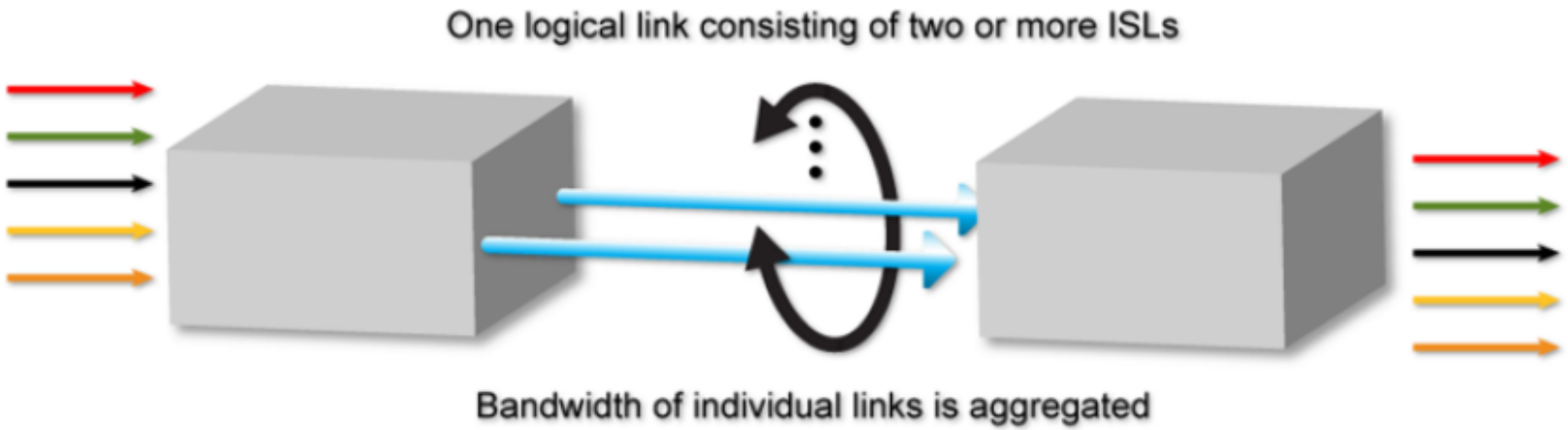


BROCADE
A Broadcom Company

Brocade ISL Trunking

What is Brocade ISL Trunking?

- Trunking combines two or more physical ISLs into a single logical link
- Trunking Goals
  - Reduces individual ISL congestion
  - Forms a fault-tolerant high bandwidth logical ISL (called a trunk or trunk group) that withstands the failure of individual ISLs
- Trunk group characteristics:
  - Frames are multiplexed across ISLs in the trunk group
  - One port in the trunk group represents the link in the routing database
  - ASICs preserve in-order delivery



BROADCOM

2, 4, 8,10,16 and 32 Gbps Trunking Overview

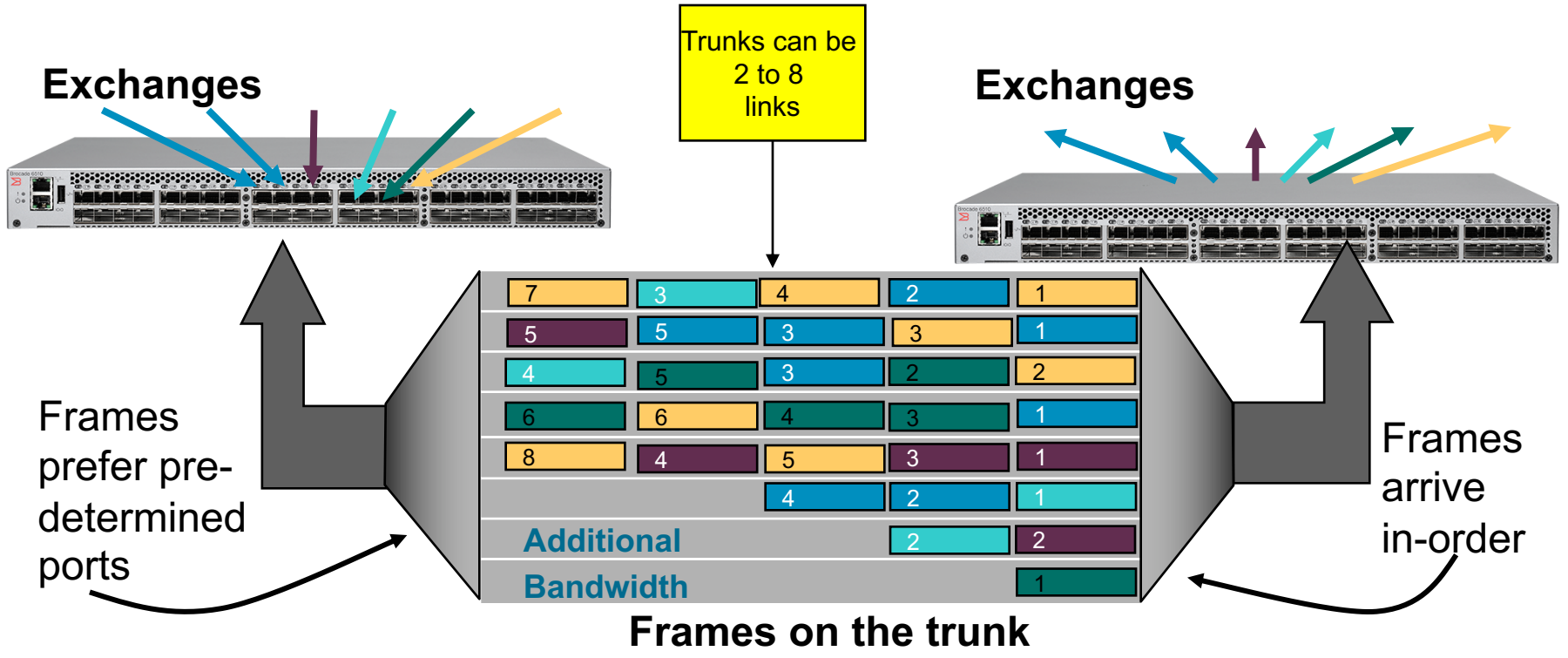
- Automatically aggregates up to 8 ISLs when the switches are connected
  - Condor2 ASICs provide up to 64 Gbps of aggregate bandwidth
  - Condor3 ASICs provide up to 128 Gbps of aggregate bandwidth
  - Condor4 ASICs provide up to 256 Gbps of aggregate bandwidth
- All ports in a trunk group must operate at the same speed
  - Condor2 ASICS support multiple 2/4/8 Gbps trunks between switches
  - Condor3 ASICs support multiple 2/4/8/10/16 Gbps trunks between same switches
  - Condor4 ASICs support multiple 4/8/10/16/32 Gbps trunks between same switches

Trunking Requirements

- Trunking license required for all switches participating in trunking
  - Available when the license is installed and ports are reinitialized
- Trunking is enabled by default
  - if previously disabled, it must be re-enabled (`portcfgtrunkport`) on the trunk ports
- Trunk ports must operate at a common speed and long distance setting
- Trunk ports must originate and end in a valid port group from the same ASIC
  - Trunking port groups include: ports 0-7, 8-15 and so on
- Cable length difference between shortest ISL and longest ISL in a trunk
  - Max cable difference between ISLs is 400 m
  - Differences > than 30 m could introduce performance degradation
- When trunking criteria is met the trunk forms automatically

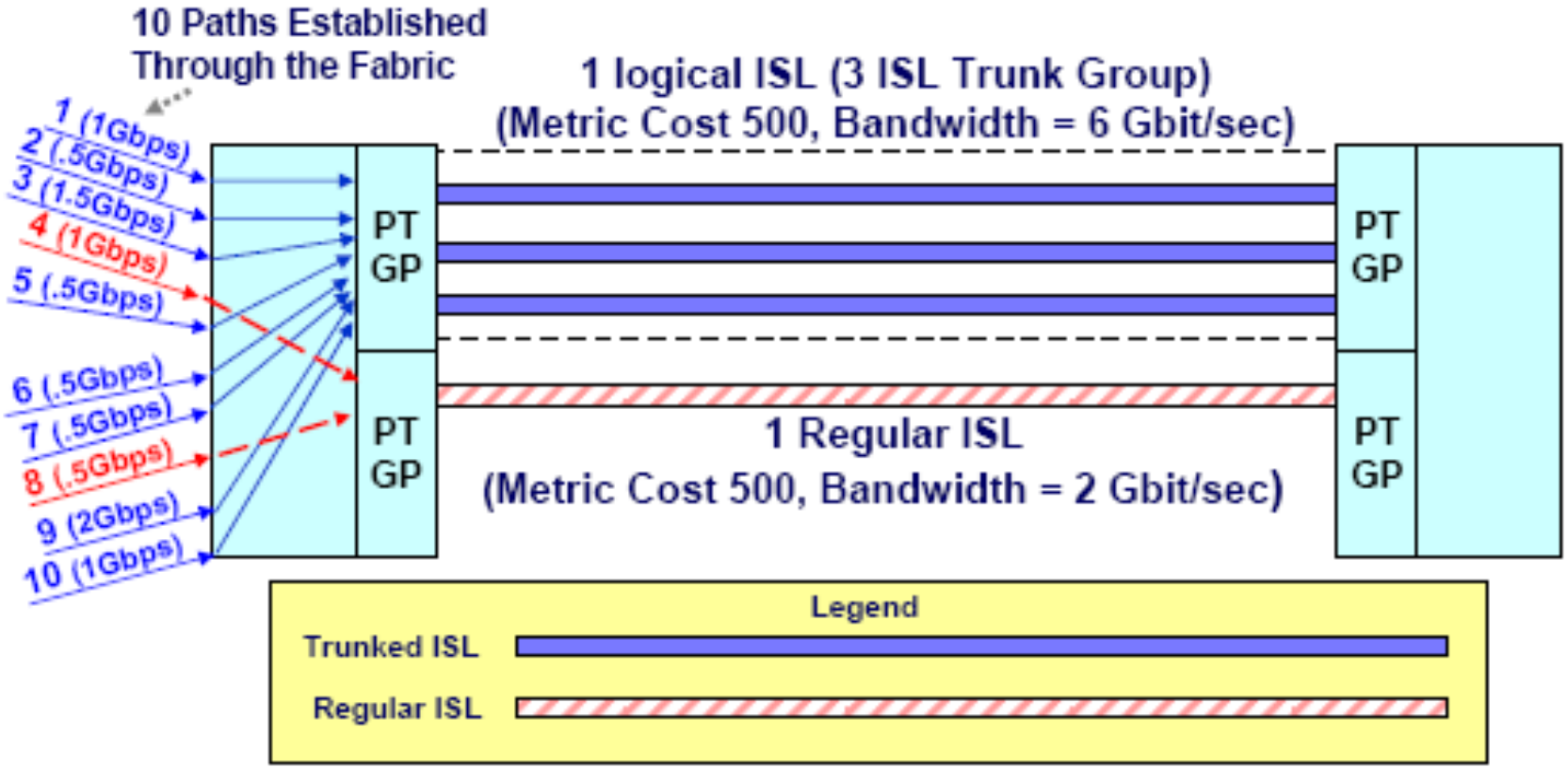
Trunking Frame Allocation

- ASICs evenly distribute frames when bandwidth of trunk is fully utilized
- Frames at low bandwidth will not appear to be evenly distributed
- No interruption of traffic if the trunk master goes offline



Port-based Routing over Trunks

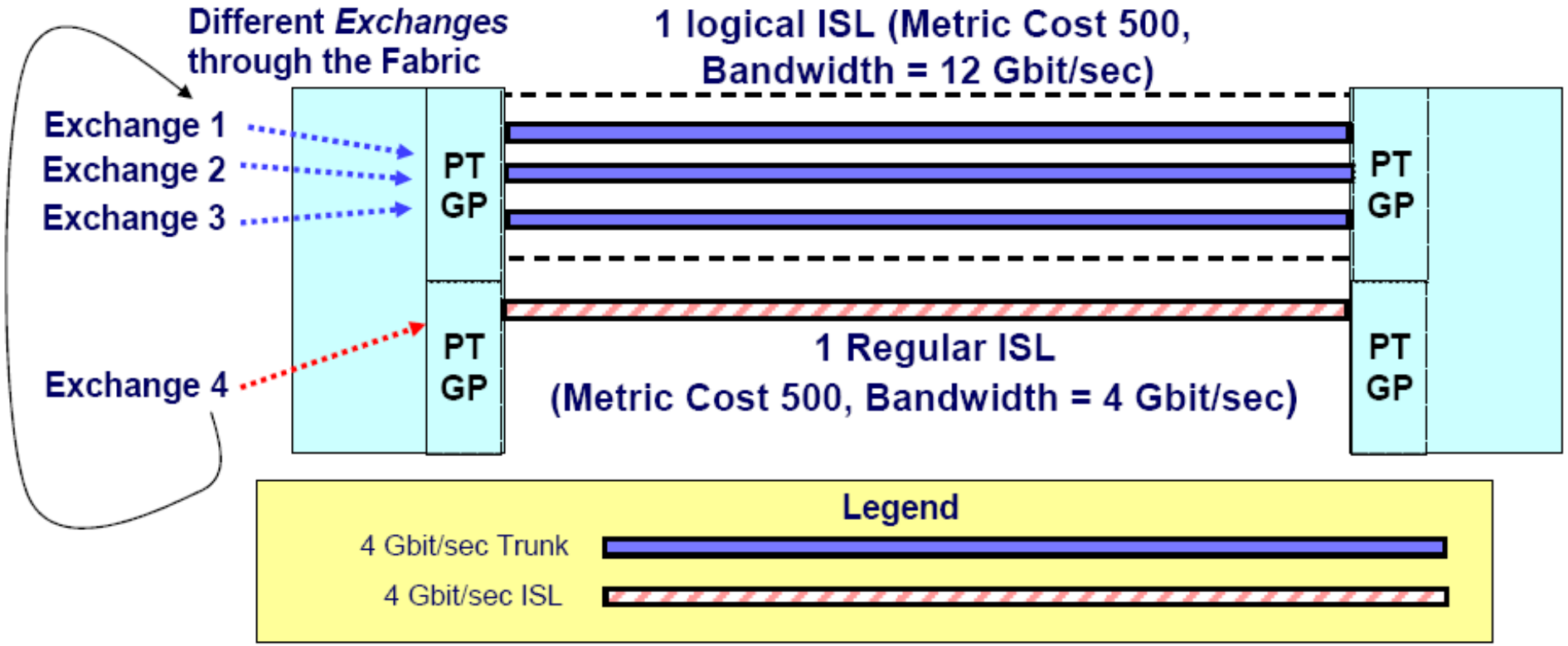
- Port based routing is supported on 2, 4, 8, 16 and 32 Gbit/sec switches
- A trunk group is considered one logical ISL identified by the trunk master
- The load assigned to a trunk group is primarily based on the ISL link cost and secondarily on total bandwidth metrics of all the ISLs in the group



BROADCOM

Exchange-based Routing over Trunks

- Exchange based routing is only supported on 4, 8, 16 and 32 Gbit/sec switch trunks
- The exchanges assigned to a trunk group/ISL are based primarily on the link cost & secondarily on total bandwidth metrics of all the ISLs



The Deskew Counter

- Deskew values are related to distance and link quality
  - Deskew units represent the time difference for traffic to travel over each ISL as compared to the shortest ISL in the group
  - The system automatically sets the minimum deskew value of the ISL with the least latency (shortest round-trip time) to 15 deskew units
  - The deskew for the remaining ISLs is calculated in relation to the ISL with the least latency
- The deskew value is a representation of an ISLs transmission capabilities
  - Differences in deskew can be caused by signal degradation which affects the transmission time of frames through the link
  - Can also be caused by excessive differences in cable length
- Deskew values are displayed in the `trunkshow` command output

Trunking and Cable Lengths

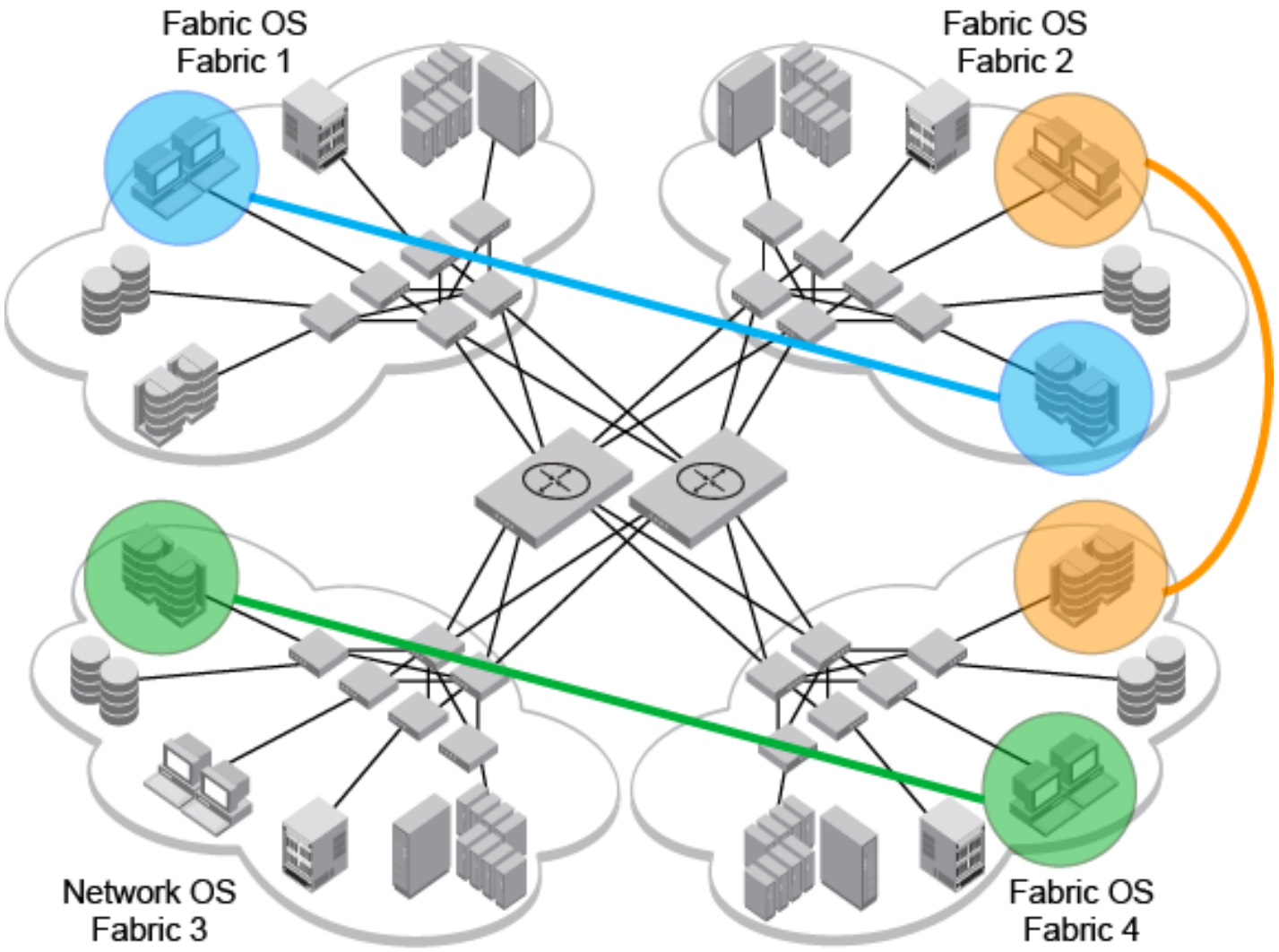
- The maximum supported difference in cable length between the shortest and the longest ISL in a trunk group is 400 meters
  - This is to help ensure in-order delivery of frames
  - Consider this when creating long distance trunks over WDMs

- A two meter difference is approximately equal to one deskew unit
  - Differences greater than 30 meters could introduce performance degradation
  - Since the shortest ISL is set to a deskew of 15, an ISL with a difference of 30 meters has a deskew of approximately 30
    - A 400 meter difference would yield a deskew value of 215

FC-to-FC routing

FC-FC Routing

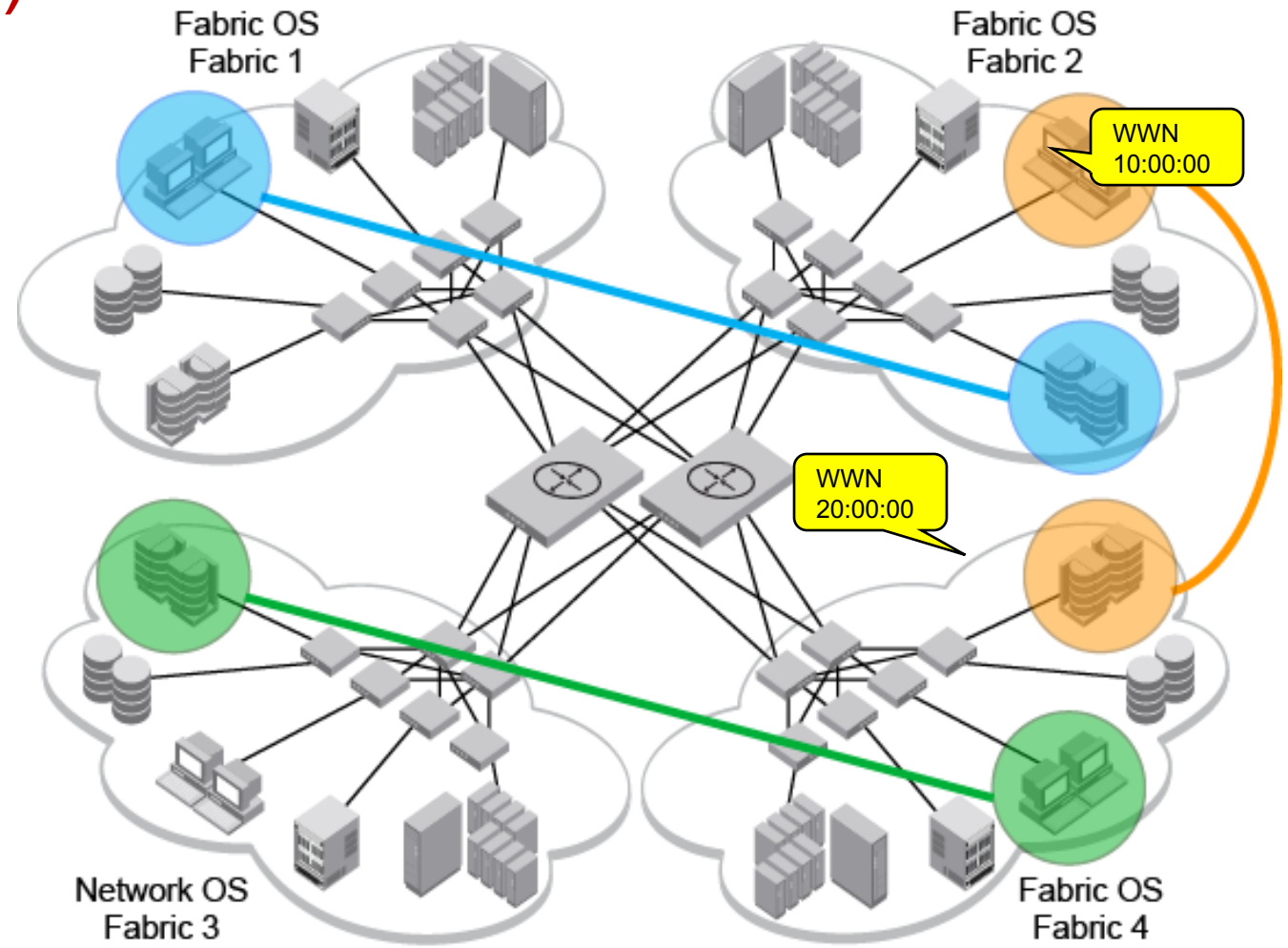
- Fabric OS provides Layer 3 Fibre Channel-to-Fibre Channel routing (FC-FC Routing) between fabrics
- Allows device access between two or more fabrics without merging the fabrics
- FC-FC Routing is supported between the following fabric types:
  - Fabric OS-to-Fabric OS



BROADCOM

FC-FC Routing (cont.)

- Physical connectivity is accomplished through the use of a Fibre Channel router (FC router)
- Logical connectivity is accomplished through the use of Logical Storage Area Networks (LSANs), by creating uniquely named zones called “LSAN zones”



LSAN Zone Fabric 2
• Storage WWN 20:00:00:
• Host WWN 10:00:00:

LSAN Zone Fabric 4
• Storage WWN 20:00:00:
• Host WWN 10:00:00:

BROADCOM

FC-FC Routing Benefits

- **Scalability**: The scaling of a routed fabric has no effect on other fabrics
  - Overall network size can vastly exceed that of a conventional, non-routed fabric
- **Fabric isolation**
  - Fabric reconfigurations do not propagate between separate fabrics
  - Faults in fabric services are contained
  - Zoning errors do not propagate
- **Interoperability**: Interoperability between different SAN operating systems can be achieved
- **Seamless migrations**: Using dual routers/fabrics, devices can be migrated to other switches seamlessly

Autonomous Fabrics

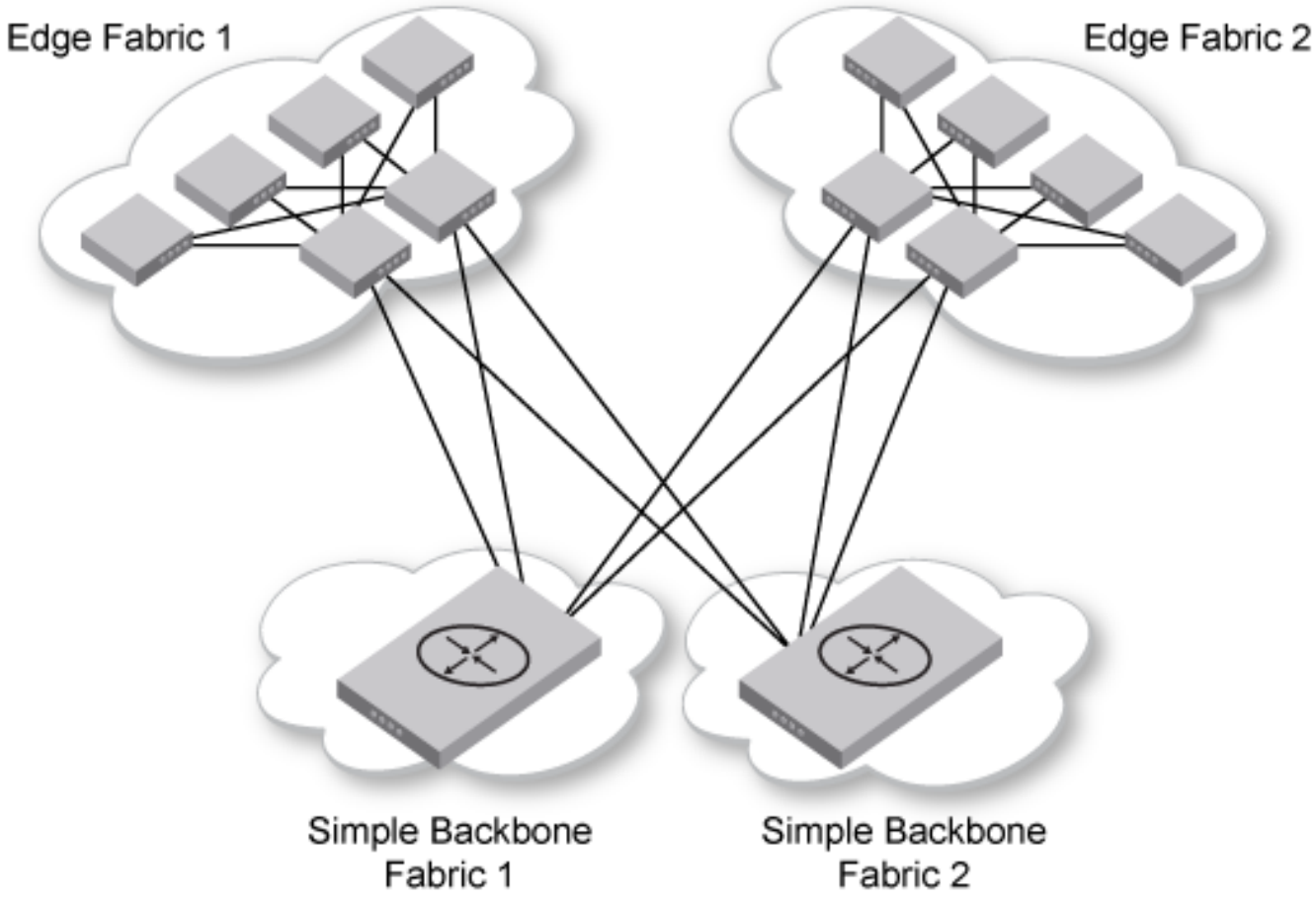
What does Autonomous mean?

- **FSPF database is contained within the fabric**
  - Cannot use the physical PID of the remote device; must use logical PID assigned to the routed device
- **Name Server**
  - Separate name servers for each fabric
  - Cannot merge fabric information with remote domain
- **Domain Space**
  - Duplicate domains may exist in separate fabrics
- **Zoning Database**
  - Each fabric has its own zoning database
- **Routing Tables**
  - Different for each fabric
  - Specific to each fabric
- **SCR/RSCNs**
  - RSCNs are specific to the fabric
- **Time-out values**
  - Parameters specified in fabric.ops can be different in each fabric

BROADCOM

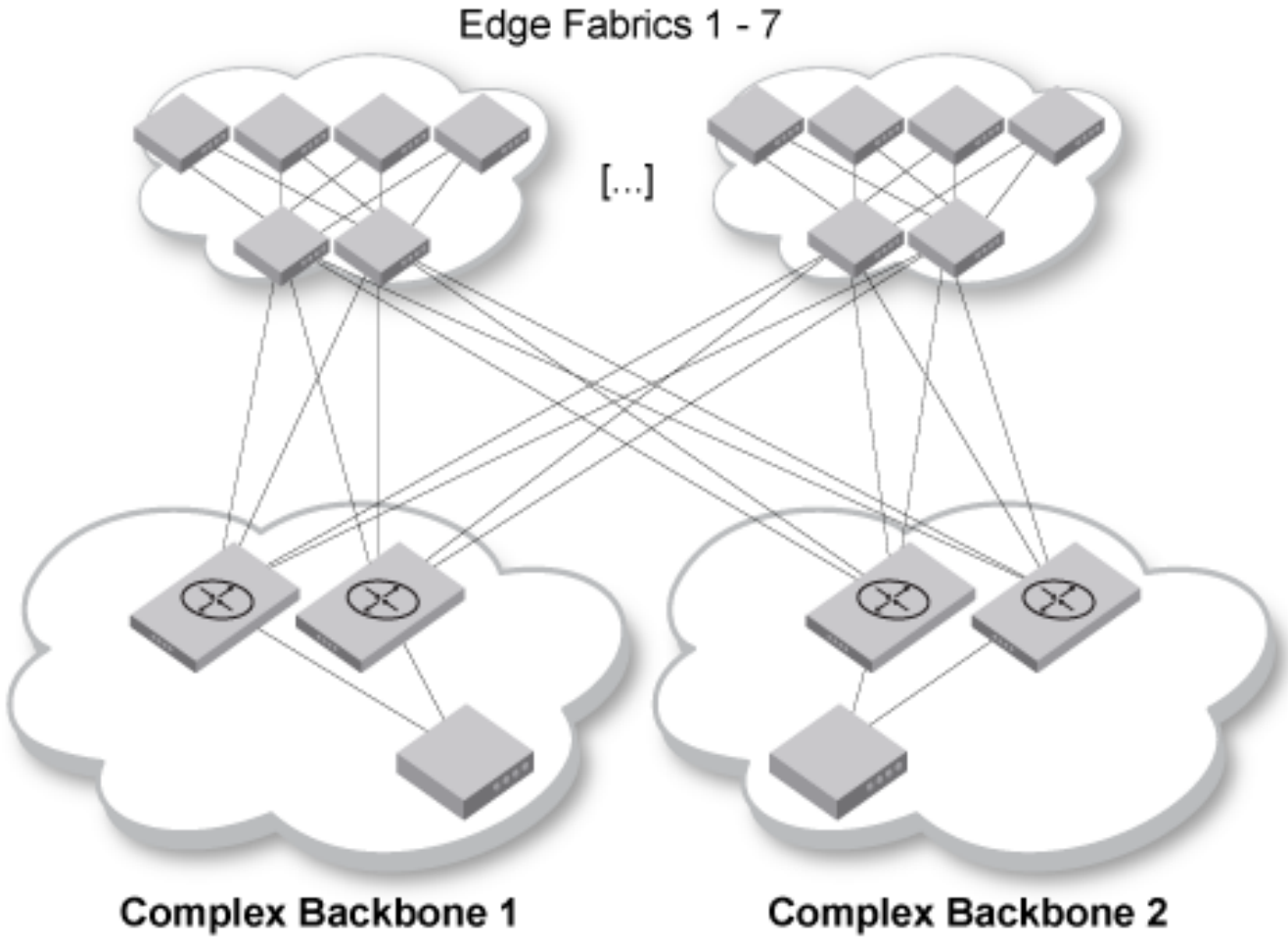
FC-FC Routing Terminology
Autonomous fabrics

- **Edge Fabric**: A fabric that is attached to one or more FC router EX_Ports
  - Maintains its own fabric settings as an autonomous system
  - Can communicate with devices in other edge fabrics and devices in the backbone fabric with the use of LSAN zones
- **Backbone Fabric (BB)**: The interconnection point for edge fabrics, containing at least one FC router
  - A simple backbone fabric has one FC router fabric with no E_Port and VE_Port connectivity to other FC routers



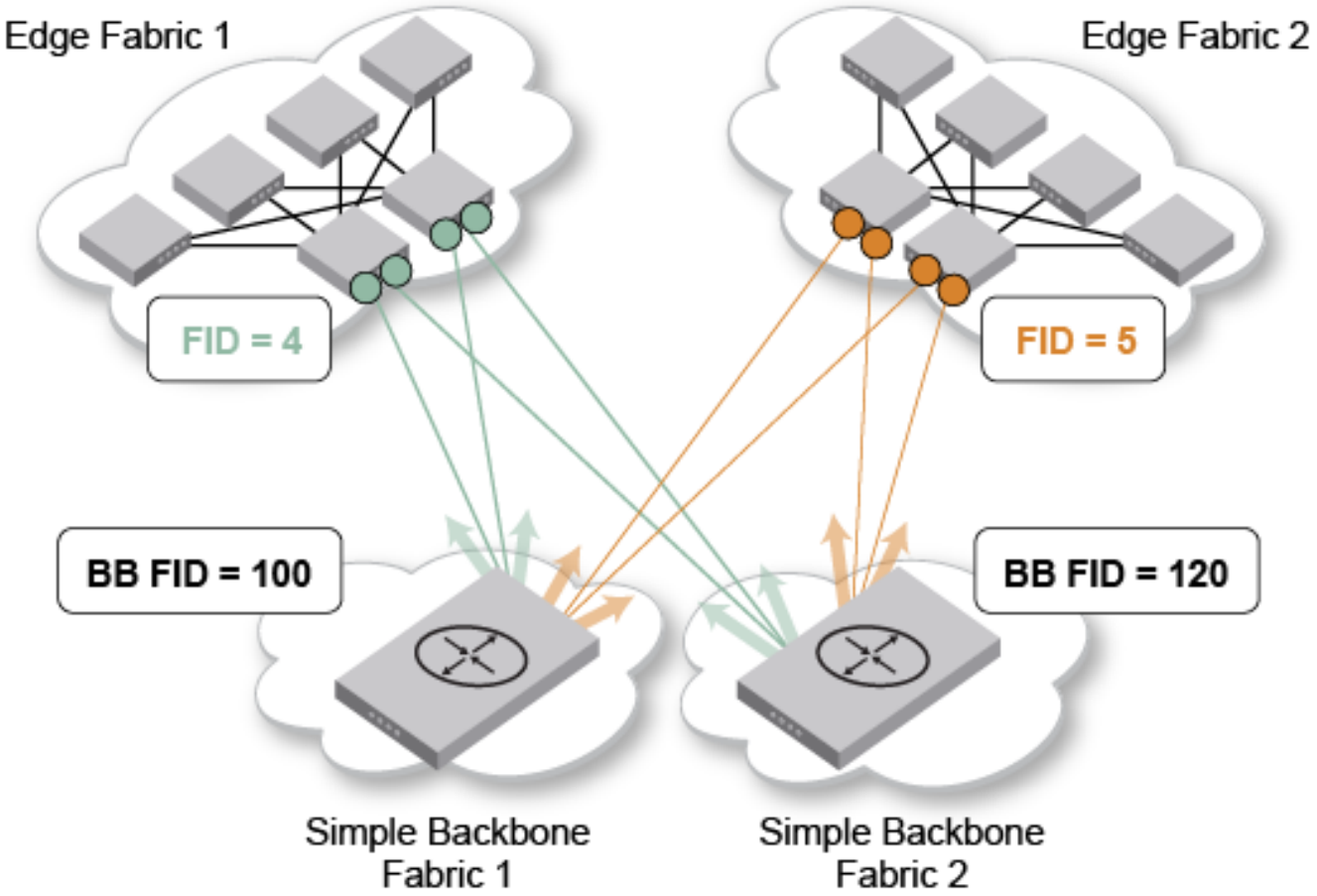
FC-FC Routing Terminology (cont.)
Autonomous fabrics

- **Complex Backbone**: Allows scalable routed fabrics by allowing multiple FC routers to be networked together using E_Ports in the backbone fabric
  - All ISLs that participate in the backbone fabric are configured as E_Ports or VE_Ports
  - Any combination of Fabric OS switches and routers can make up the backbone



FC-FC Routing Terminology (cont.)

- **Fabric ID (FID)**: Uniquely identifies each fabric participating in routed fabrics
  - Every edge and backbone fabric requires a unique FID
  - FIDs can be assigned any number between1-128 (default is 1)
  - ***Both backbone and edge FIDs are configured on the FC router***

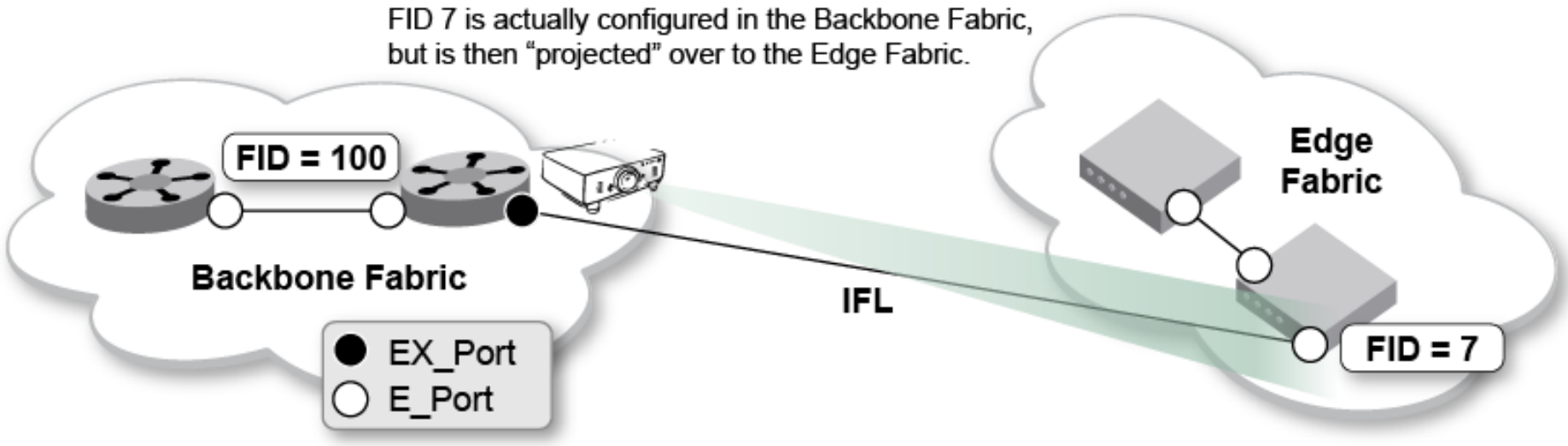


BROADCOM

FC-FC Routing Terminology (cont.)

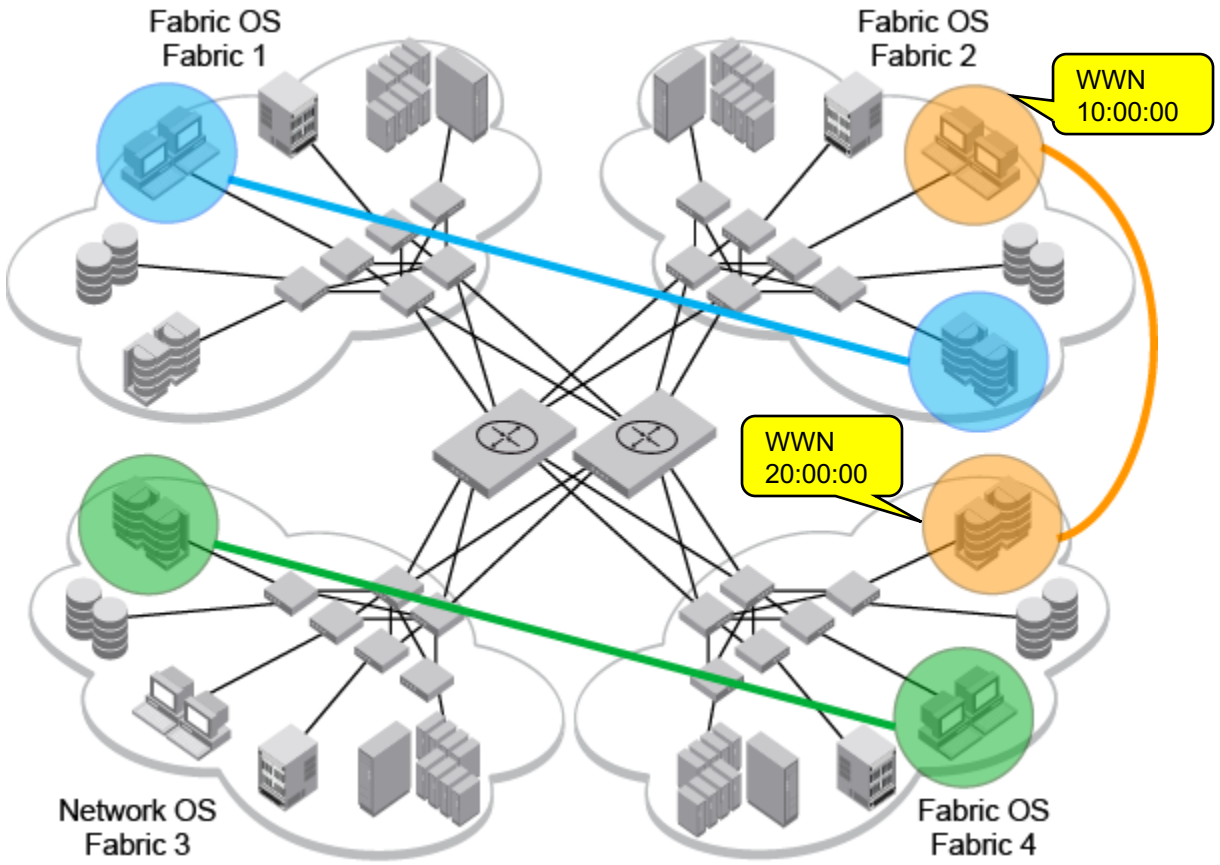
- **EX_Port**: A type of E_Port used to connect an FC router port to an edge fabric without merging the two
  - EX_Ports on a router connect to E_Ports in an edge fabric
  - Use the `portcfgexport` command to configure router EX_Ports

- **Inter-Fabric Link (IFL)**: The connection between a backbone fabric EX_Port and an edge fabric E_Port



FC-FC Routing Terminology (cont.)

- **LSAN**: A logical storage area network that spans multiple physical fabrics
  - Allowing devices in different fabrics to communicate with each other
- **LSAN zone**: Zones that define which devices are to be shared between fabrics
  - Defined in each fabric that will share devices (edge or backbone)
  - Traditional zone created using normal zoning tools but uses a special naming prefix LSAN_

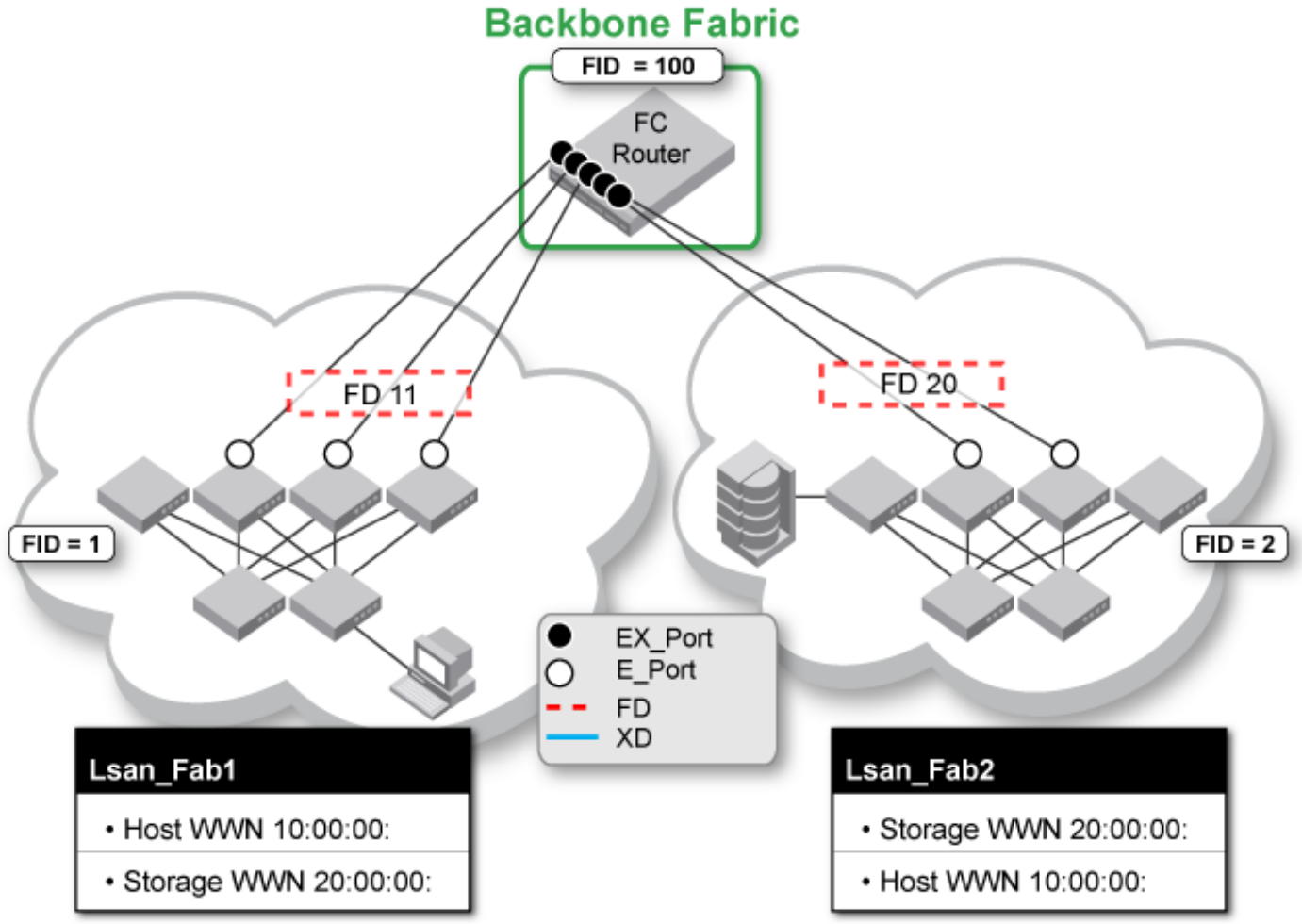


LSAN Zone Fabric 2
• Storage WWN 20:00:00:
• Host WWN 10:00:00:

LSAN Zone Fabric 4
• Storage WWN 20:00:00:
• Host WWN 10:00:00:

FC-FC Routing Terminology (cont.)

- **Front Domain (fd)**: A logical domain created in the ***edge*** fabric when edge fabrics are connected to backbone fabrics
- **Translate Domain (xd)**: A logical domain created in the ***edge*** fabric when routed fabrics share devices
  - Sharing is accomplished through the creation and enabling of LSAN zones
  - This logical domain is where the imported devices logically exist

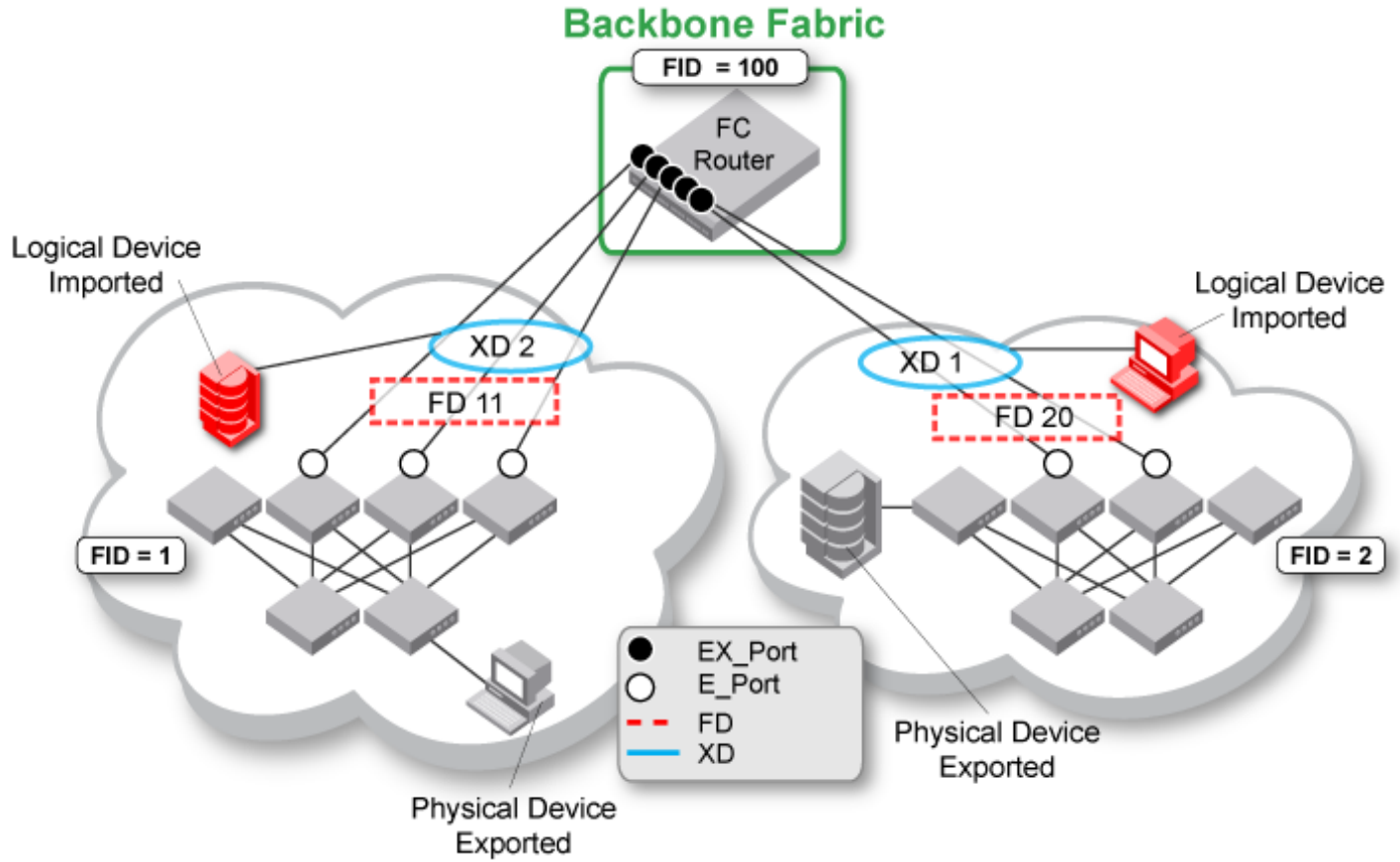


BROADCOM

FC-FC Routing Terminology (cont.)

- **Exported device**: A *physical* device defined in an LSAN zone that the router shares-out from a fabric (edge or backbone)

- **Imported device**: A *logical* device defined in an LSAN zone that acts as a proxy device for the physical device in a different routed fabric (edge or backbone)



BROADCOM

FC-to-FC routing (cont.)

Hardware

- 48000 with FR4-18i
- DCX with IR license or FR4-18i
- DCX8510 with IR license. FR4-18i not supported.
- X6 Directors with IR license.
- 5100/5300 with IR license
- 6510/6520 with IR license
- G620/G630 with IR license.
- 7500/7800/7840/7810
- Not supported on 4G switches and low-end (300, 6505, G610)
- Not supported on embedded switches

Brocade Virtual Fabrics

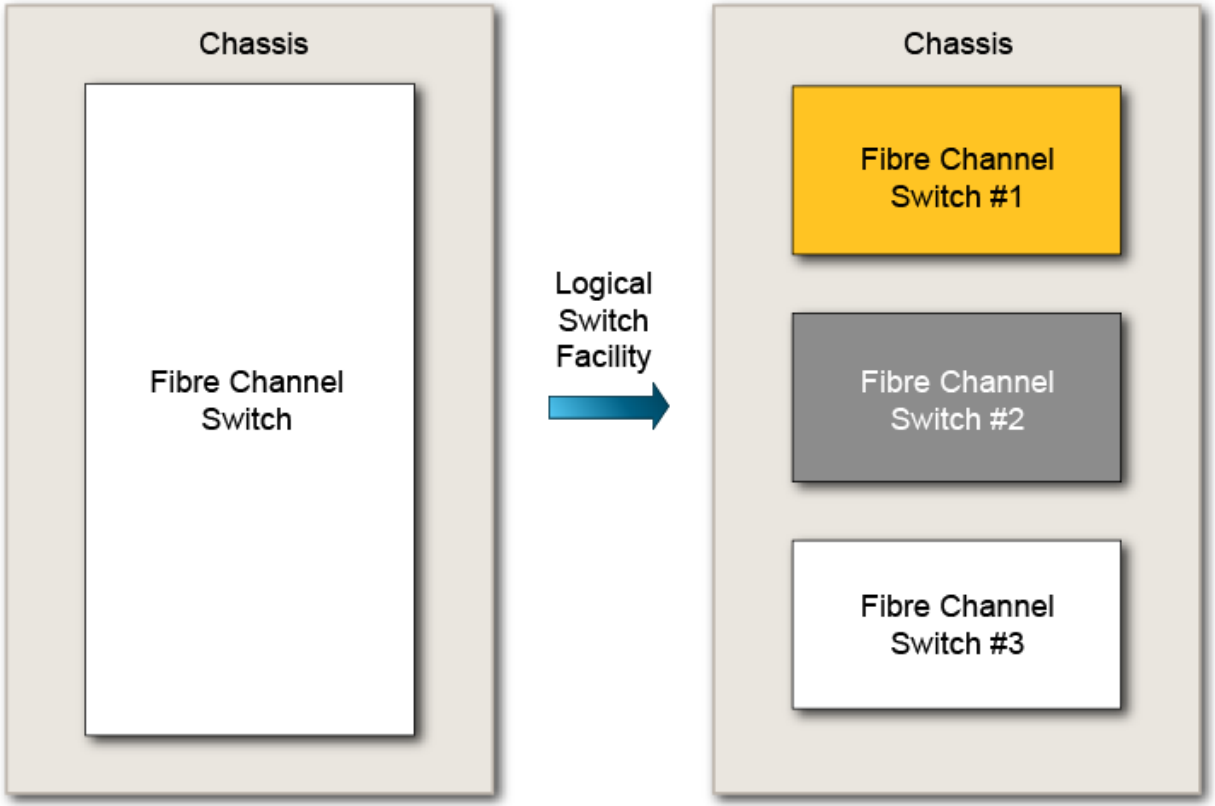
BROADCOM®

Introduction to Virtual Fabrics

- **Virtual Fabrics are an architecture to virtualize hardware boundaries**
  - Allows multiple switches and fabrics within the same hardware
- **Increases flexibility**
  - Allows movement of ports from fabric to fabric without rewiring
  - Allows easy assignment of unused ports as needed
- **Increases network security**
  - Can be used to create separate secure logical fabrics for sensitive data
- **Further reduces fault impact**
  - Maintains separate fabric services per logical fabric
  - Separate FICON and Open Systems traffic
- **Multi-tennancy** — managing fabrics for multiple companies or groups on the same hardware

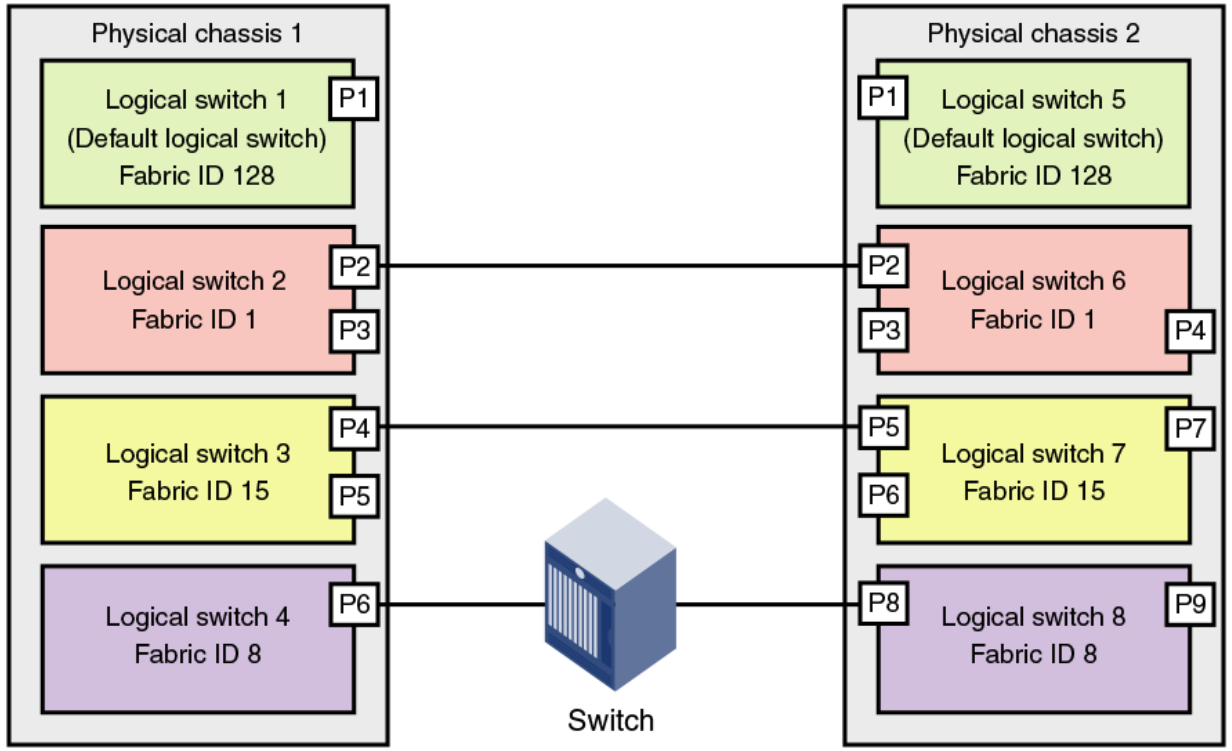
Virtual Fabrics
Logical switch defined

- Allows a physical switch to be divided into multiple logical switches
- Building block for Virtual Fabrics and device sharing



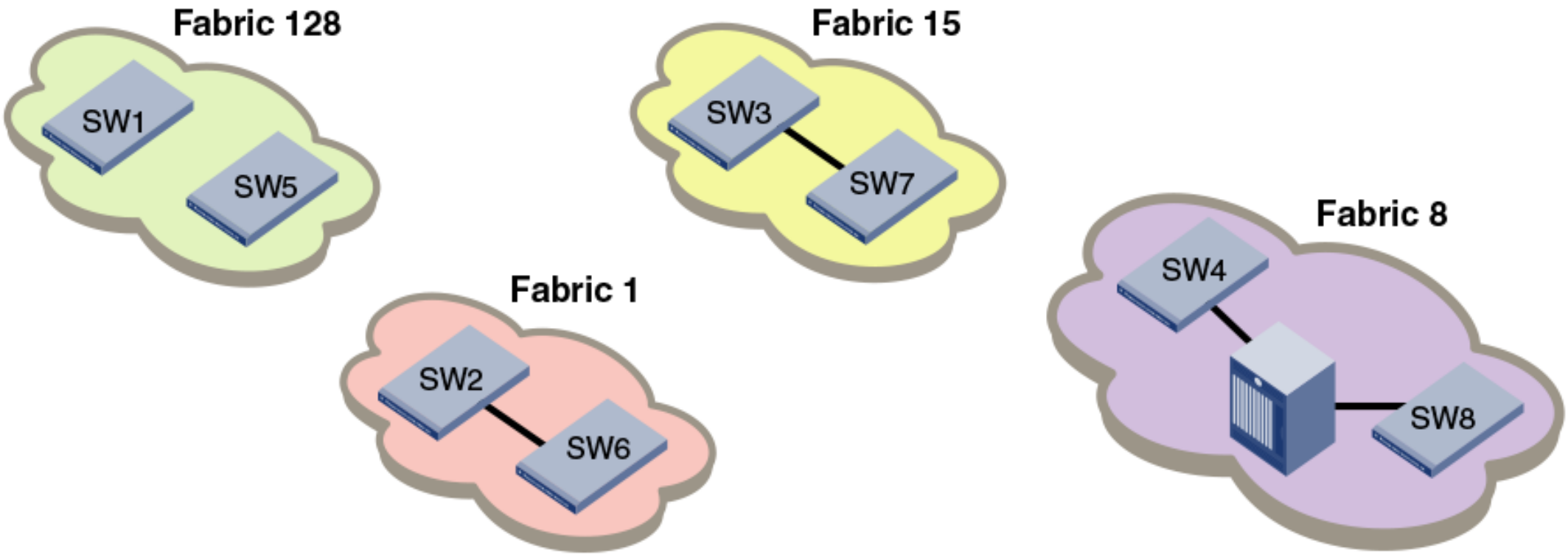
Logical fabric and ISLs

- Below shows two physical chassis divided into logical switches. In ISLs are used to connect the logical switches with FID 1 and the logical switches with FID 15. The logical switches with FID 8 are each connected to a non-Virtual Fabrics switch. The two logical switches and the non-Virtual Fabrics switch are all in the same fabric, with FID 8.



BROADCOM

Logical switches connected to form logical fabrics



BROADCOM

Virtual Fabrics Terminology
Logical switch

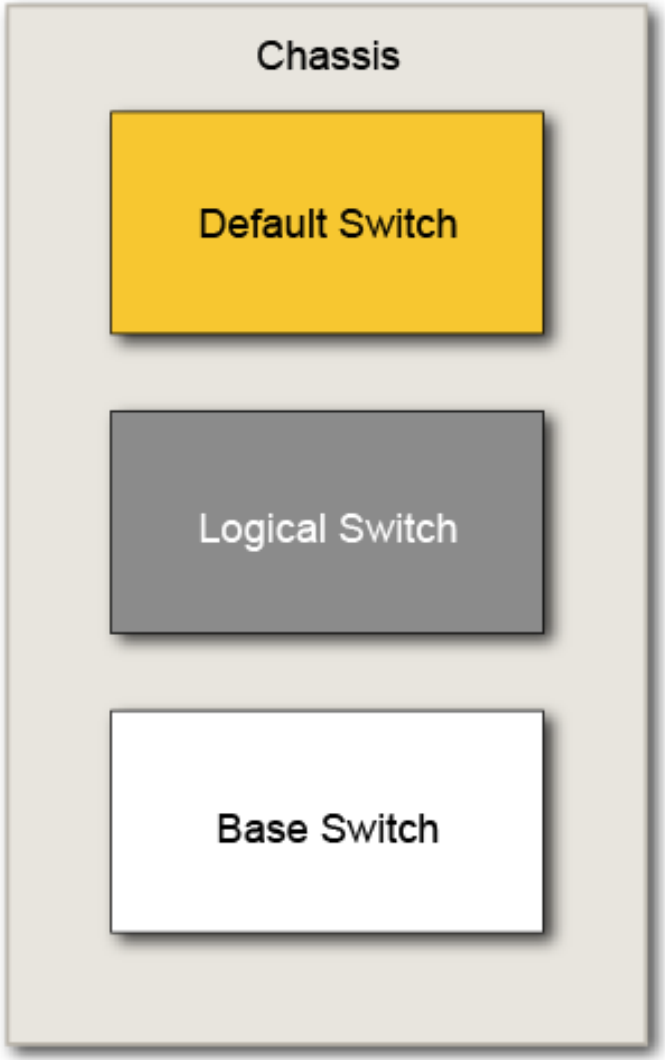
- A logical switch is a complete and self-contained FC switch
  - Fabric services (e.g. Name Server, zoning)
  - Configuration (e.g. port, switch, fabric)
  - Fabric characteristics (e.g. addressing)

- A collection of zero or more user ports that acts as a single Fibre Channel switch
  - Ports can be F / FL /E / VE / EX / VEX / D_Port / Mirror

- Logical switches can be created and deleted without impact to other logical switches

Virtual Fabrics
Logical switch types

- Three types of Brocade logical switches:
  - **Default logical switch**
    - Initially this contains all the ports in the chassis
    - Default fabric ID is 128
    - Supports F/FL and E/VE Ports
  - **Logical switch**
    - Standard generic Logical Switch
    - Support for F/FL and E/VE Ports
  - **Base switch**
    - Used for sharing ports/devices between Logical Switches
    - Supports E/EX and VE/VEX ports only



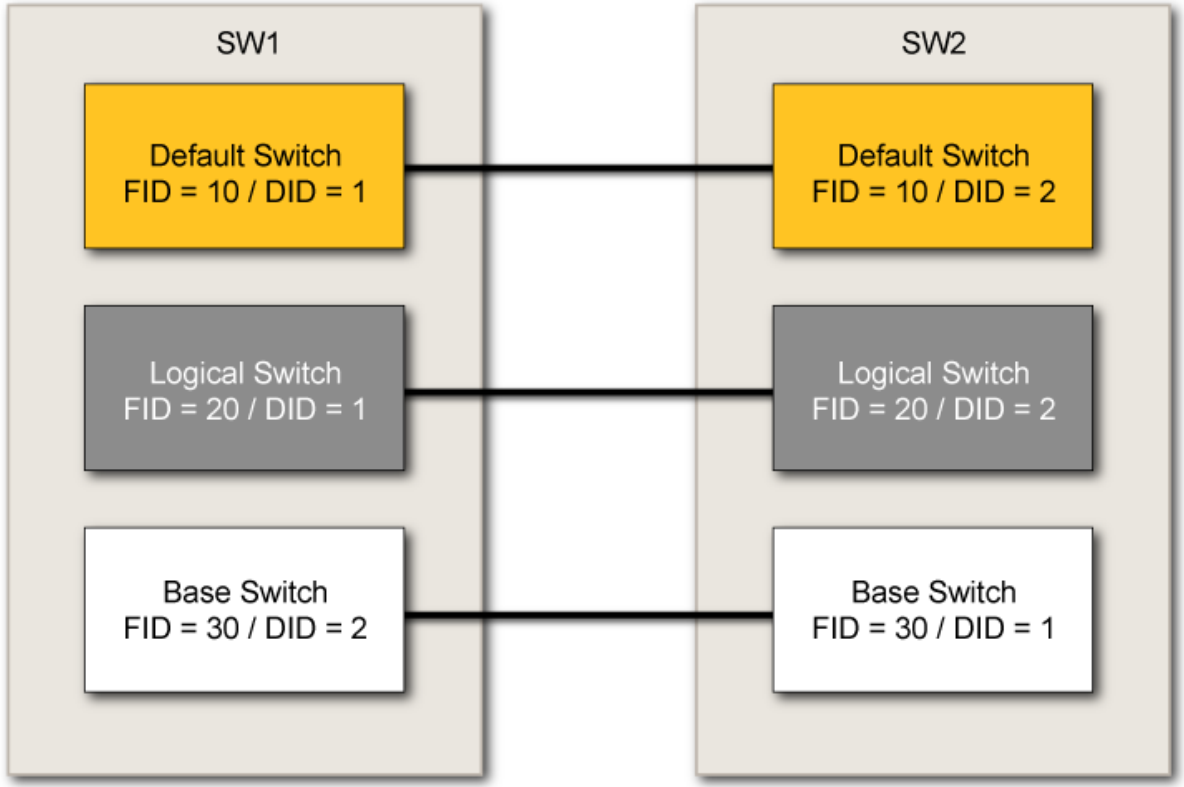
Virtual Fabrics
Logical fabric and fabric identifier

- **Logical Fabric (LF)**
  - Consists of at least one logical switch
  - Provides connectivity with traffic isolation by fabric on a common shared infrastructure

- **Fabric Identifier (FID)**
  - Identifies a Layer 2 fabric
  - Every logical switch is created with a unique FID in the chassis
  - All logical switches in a logical fabric must have the same FID
  - Non-VF switches can exist in the fabric
  - A number between 1-128
  - Supports Fabric Naming

Virtual Fabrics
Fabric ID vs domain ID

- A **FID** identifies the fabric to which a logical switch belongs
- A **Domain ID (DID)** identifies the switch within that fabric
- Within a fabric, FIDs must be the same; DIDs must be unique

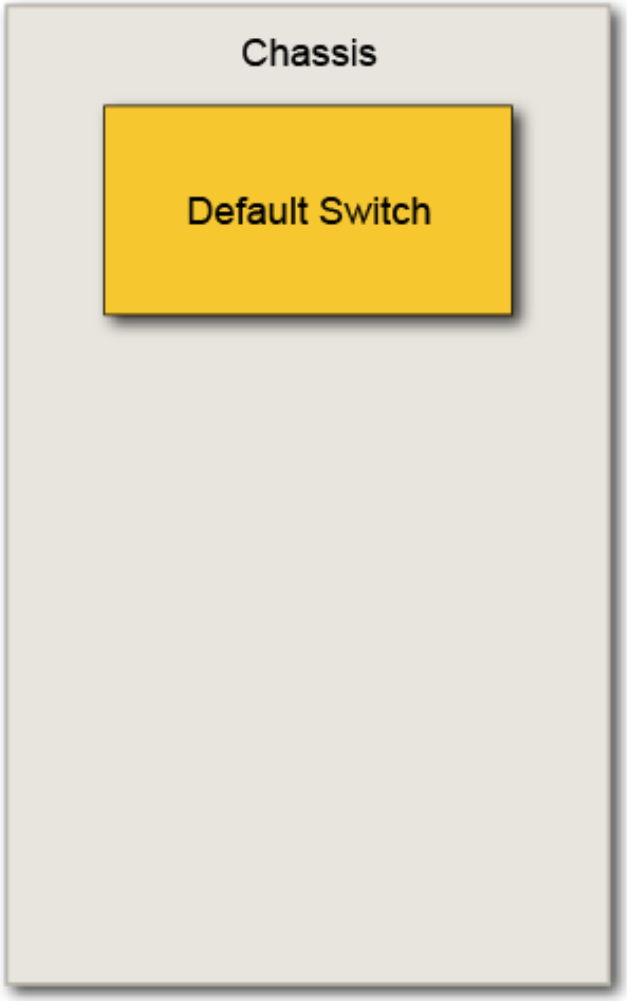


BROADCOM

Logical Switches
Default logical switch

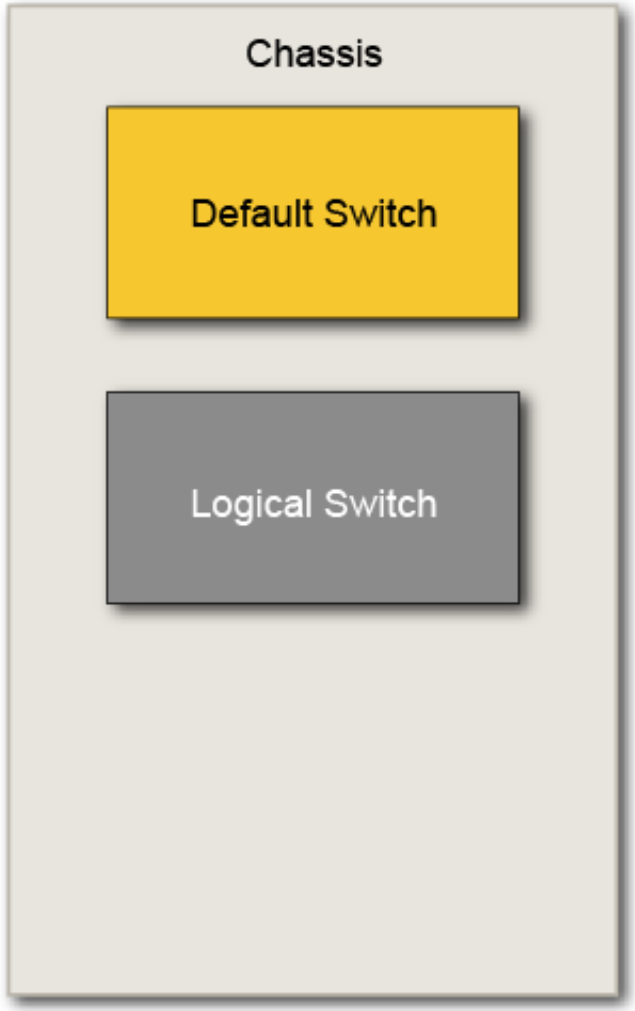
- The default FID is 128 – this can be changed

- All ports are initially assigned to the default logical switch



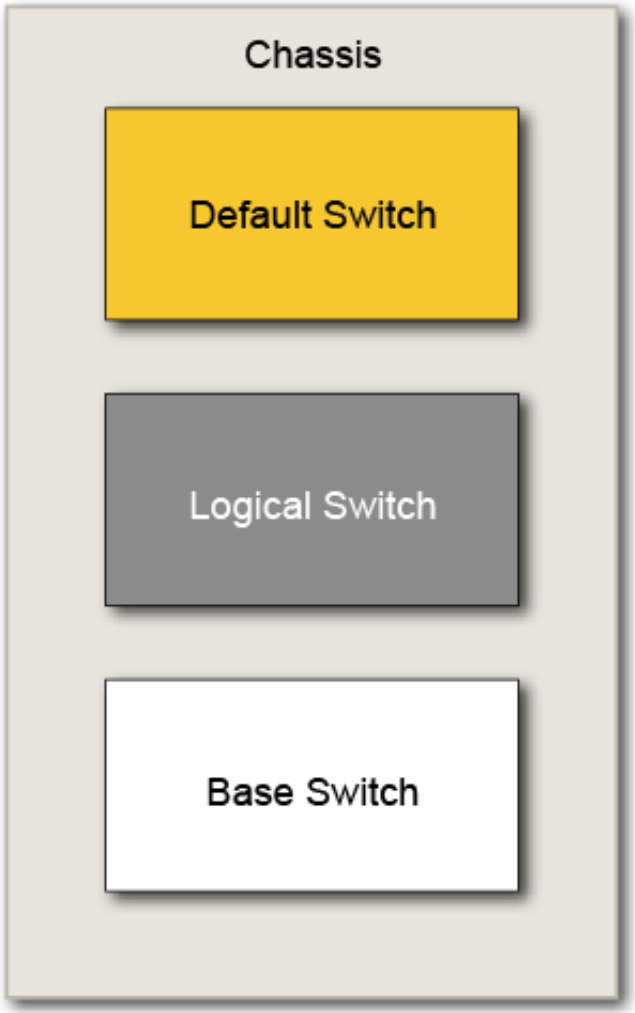
Logical Switches
Standard logical switch

- The standard logical switch is most commonly used for connecting devices
- Supports F, FL, E, VE, D and M_Ports connectivity
- Supports:
  - Brocade Native Mode
  - FICON



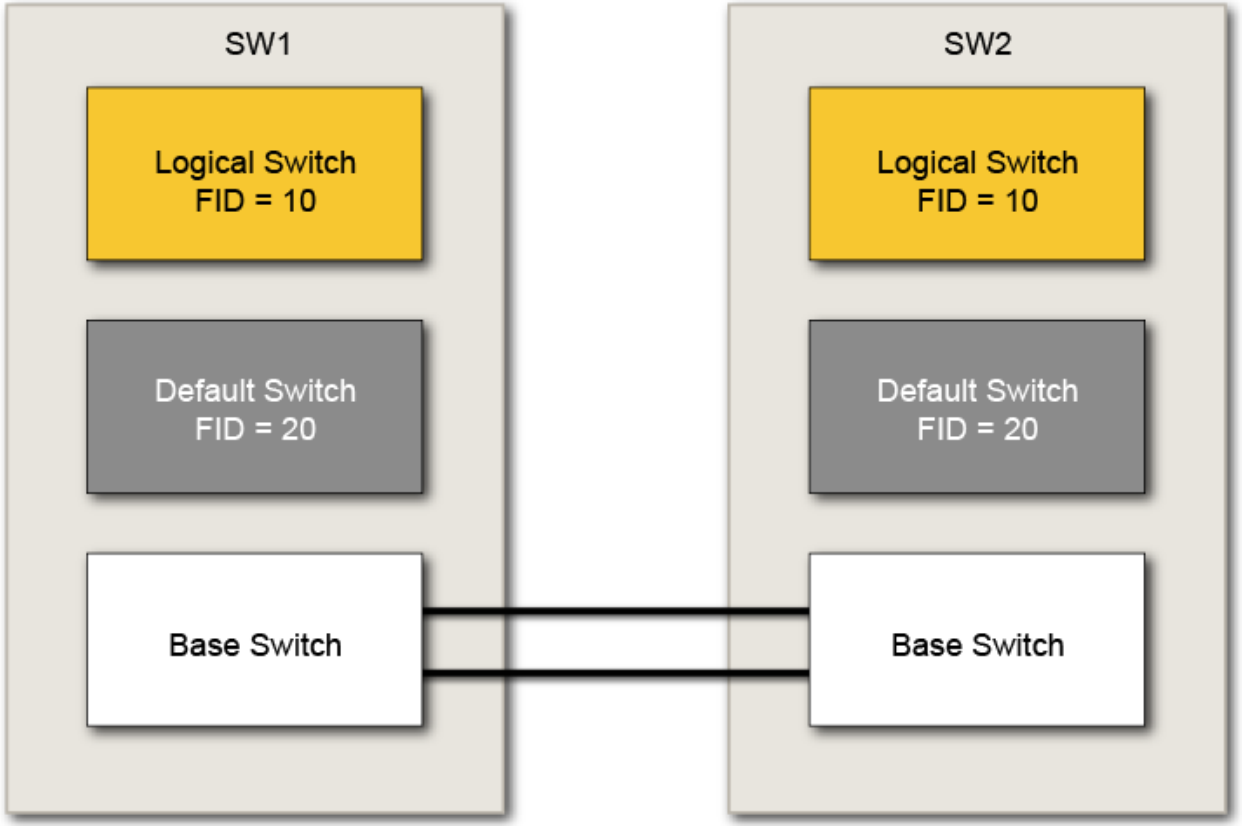
Logical Switches
Base switch

- The base switch is a special logical switch used for FC-FC routing and connectivity
  - It can be used for ISL connectivity from chassis-to-chassis
  - It can be used for FC-FC routing (IFL) between logical fabrics
  - Support only for E, EX, VE, VEX, D, and M_Ports
    - No support for F or FL ports
- Contains shared XISLS (extended ISLs) for use by other logical switches



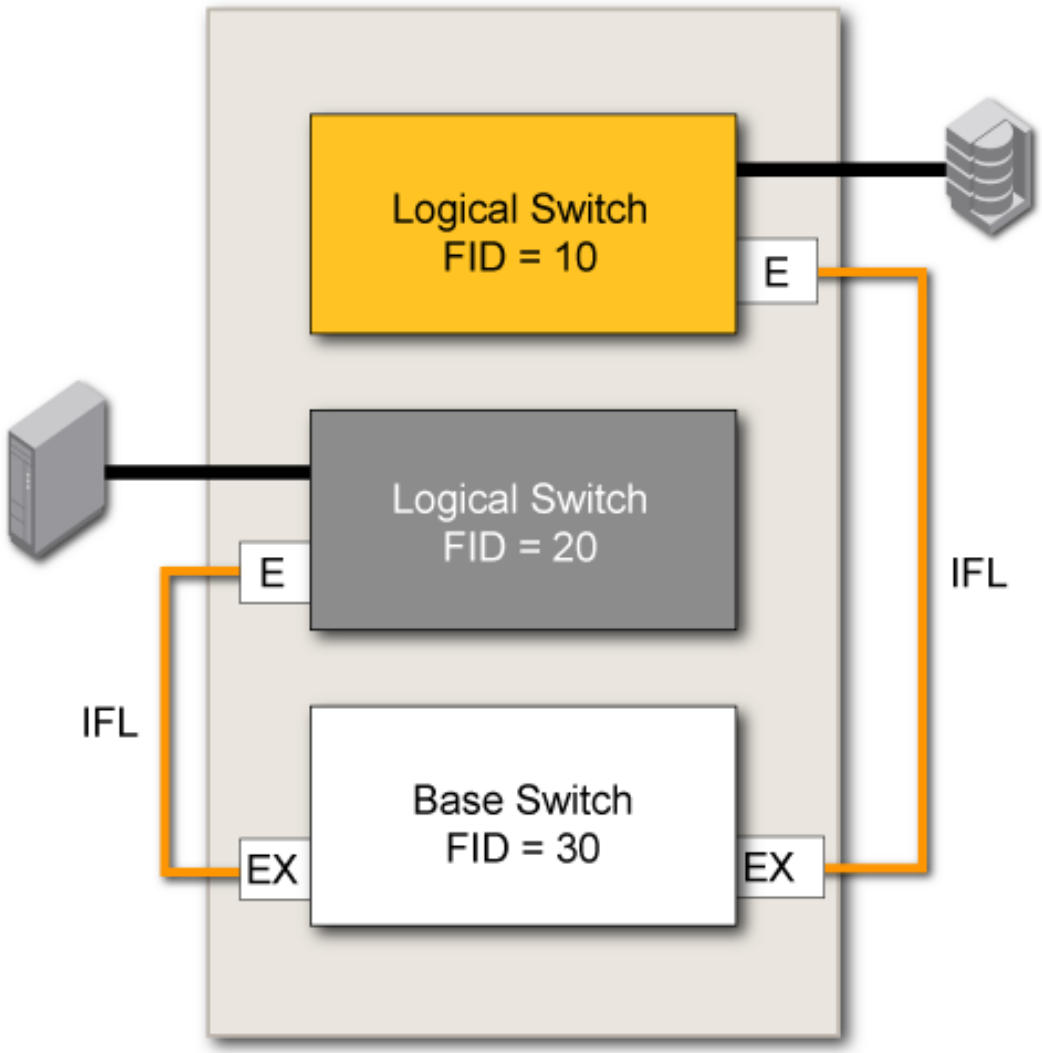
Logical Switches
Base switch (cont.)

- In the configuration below, all logical switches (FID 10 and FID 20) can use the ISL connections between the base switches
- This base switch-to-base switch connectivity forms a base fabric



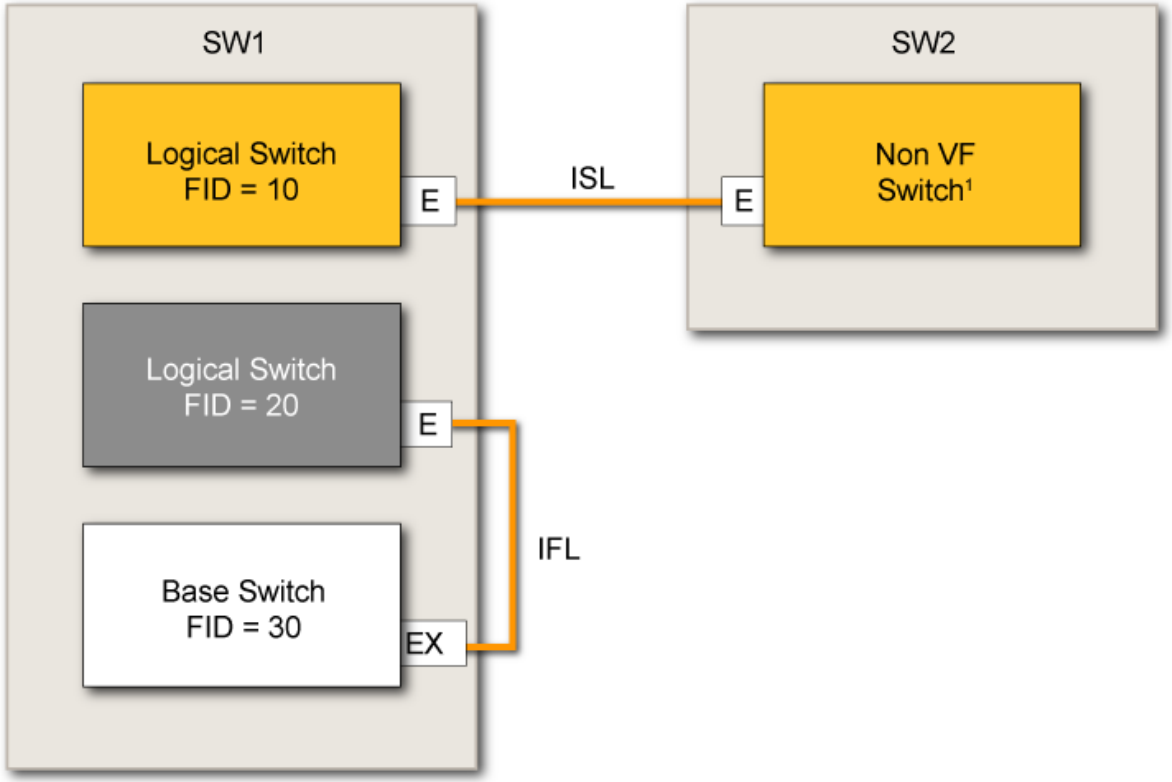
Logical Switches
Base switch (cont.)

- The base switch can be used as an FC router
  - The host connected to FID 20 can access the storage connected to FID 10
- EX_Ports are created on the Base Switch to establish a connection from each FID to the FC router



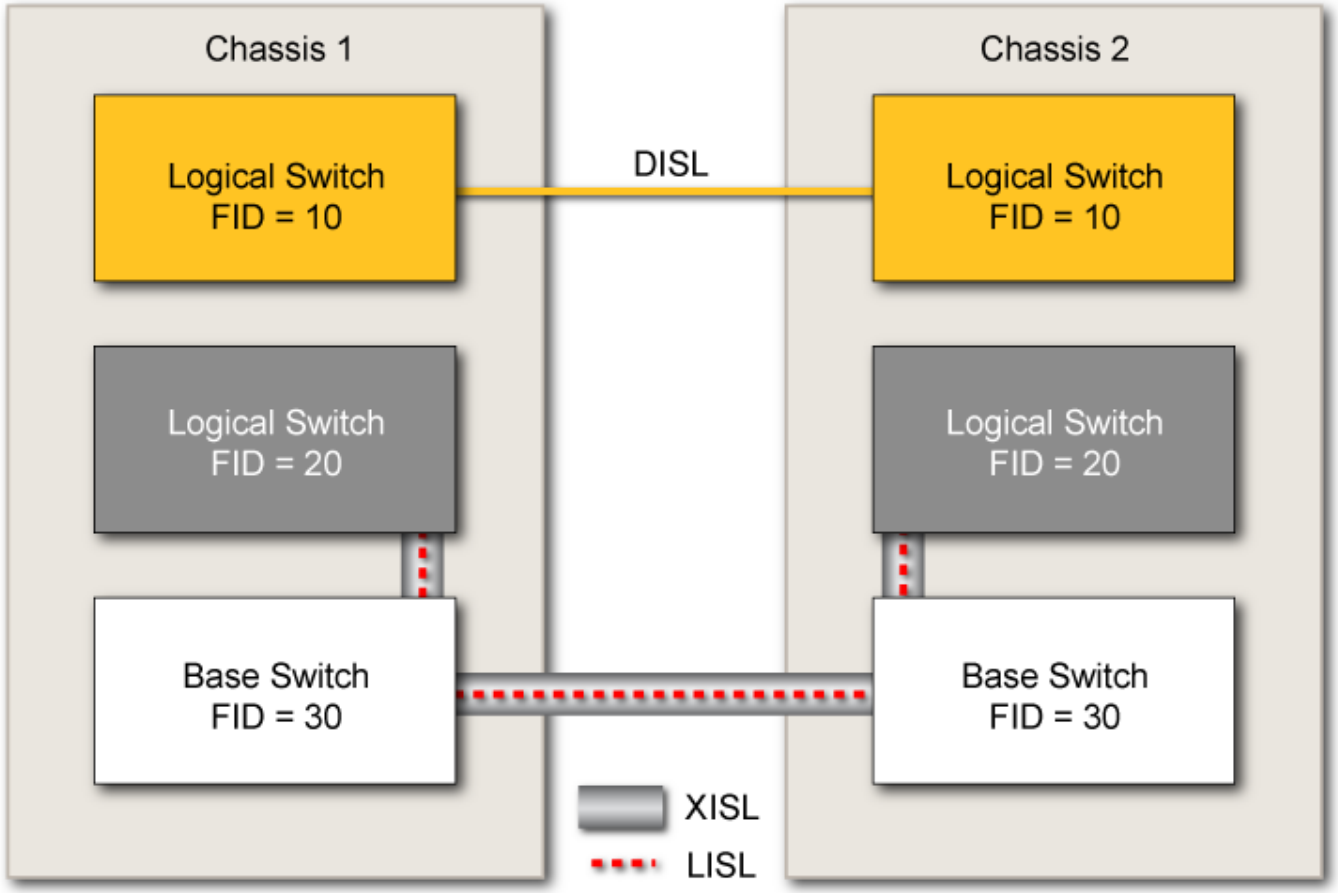
Logical Switch Link Types
ISL and IFL

- **Inter Switch Link (ISL)** – Connects two non-virtual switches or a virtual switch to a non-virtual switch in the same fabric
- **Inter Fabric Link (IFL)** – Connects between an edge fabric E_Port and an FC Router EX_Port



Logical Switch Link Types (cont.)
DISL, XISL and LISL

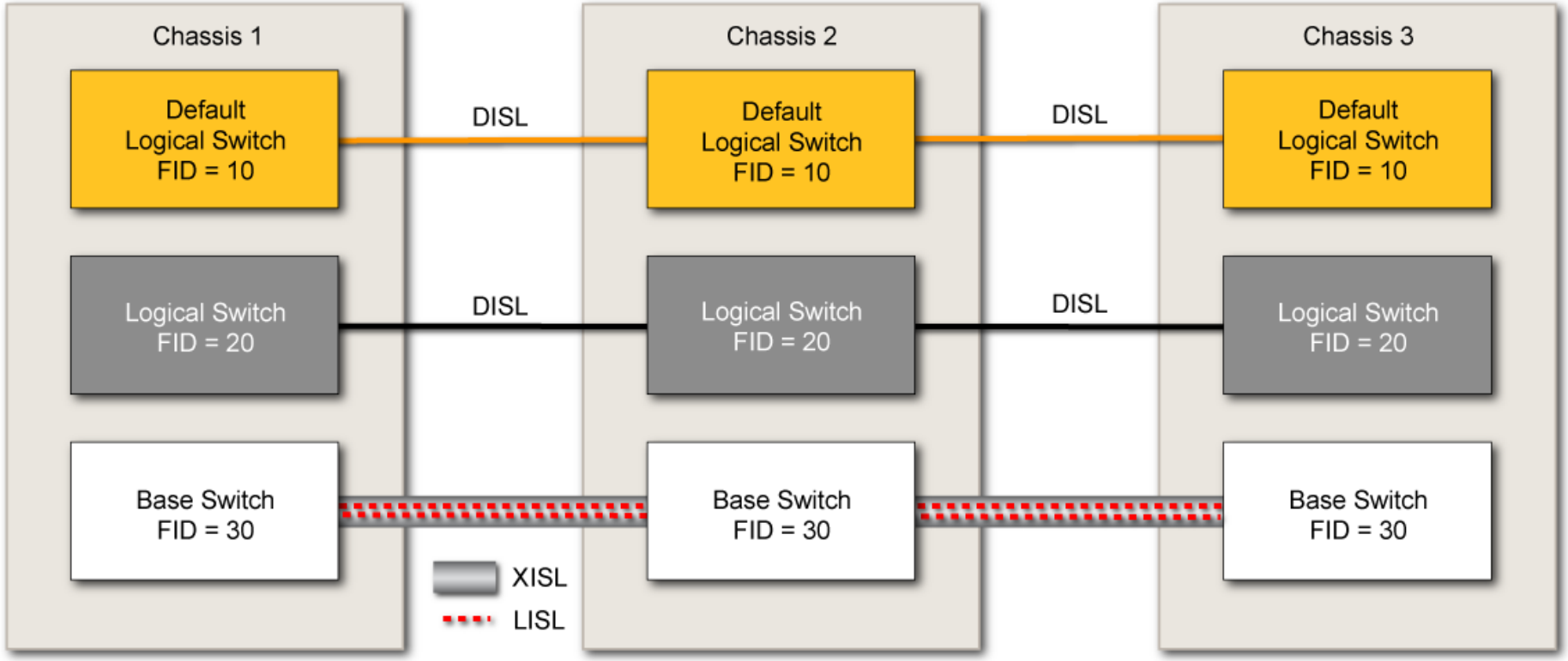
- **Dedicated Inter Switch Link (DISL)** - Used to directly connect two (non-base) logical switches in two different chassis
- **Extended Inter Switch Link (XISL)** - Used to directly connect two base switches in two different chassis
- **Logical Inter Switch Link (LISL)** - Used to connect two logical (non-base) switches in two different chassis through the XISL on the base switch
  - Not considered a hop



BROADCOM

Route Selection Example

- FIDs 10 and 20 below will use the lower cost DISL connections



- DISL favored over LISL
  - LISL cost = cost of path in base fabric + 10

BROADCOM

Brocade Virtual Fabrics

Hardware

- DCX & DCX-4S
- DCX8510
- X6 Directors
- 5100/5300
- 6510/6520
- G620/G630
- 7800/7810/7840
- Not supported on 4G switches and low-end (300, 6505, G610)
- Not supported on embedded switches
- Maximum number of logical switches: 3-16 per chassis (depending on hardware platform and FOS version)

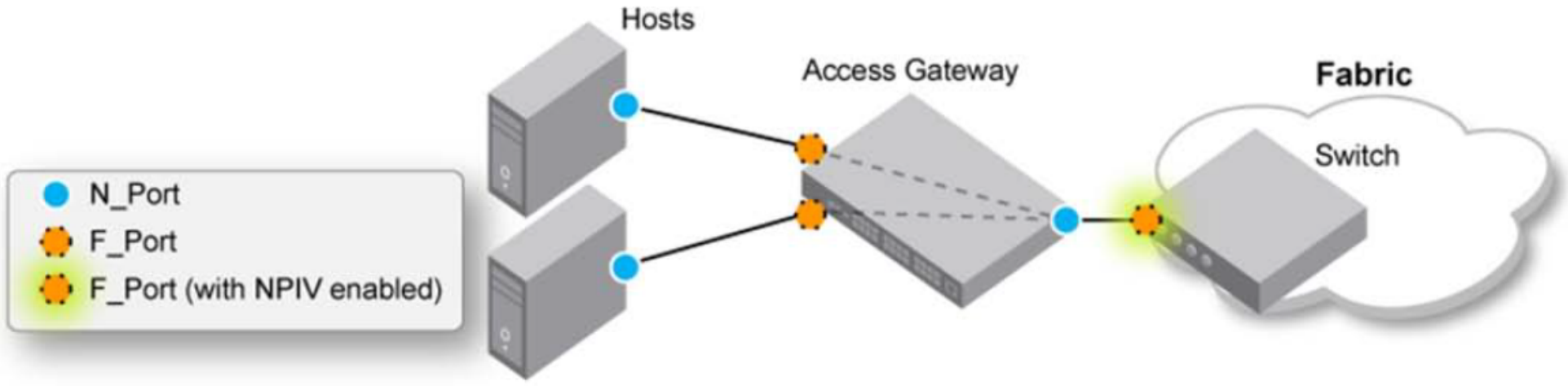
Access Gateway

Access Gateway – Overview

- Access Gateway (AG) is a Fabric OS feature that enables:
  - Seamless connectivity to a fabric
  - Enhanced scalability
  - Simplified manageability
- Designed to connect numerous servers with minimal impact to **any** existing fabric
- Focus is connectivity, bandwidth is shared
- Included in the base Fabric OS
- Attached F_Port devices must be Fibre Channel Protocol (FCP) initiators or targets
  - Not supported: loop devices, FICON, virtual iSCSI initiators, logical switches

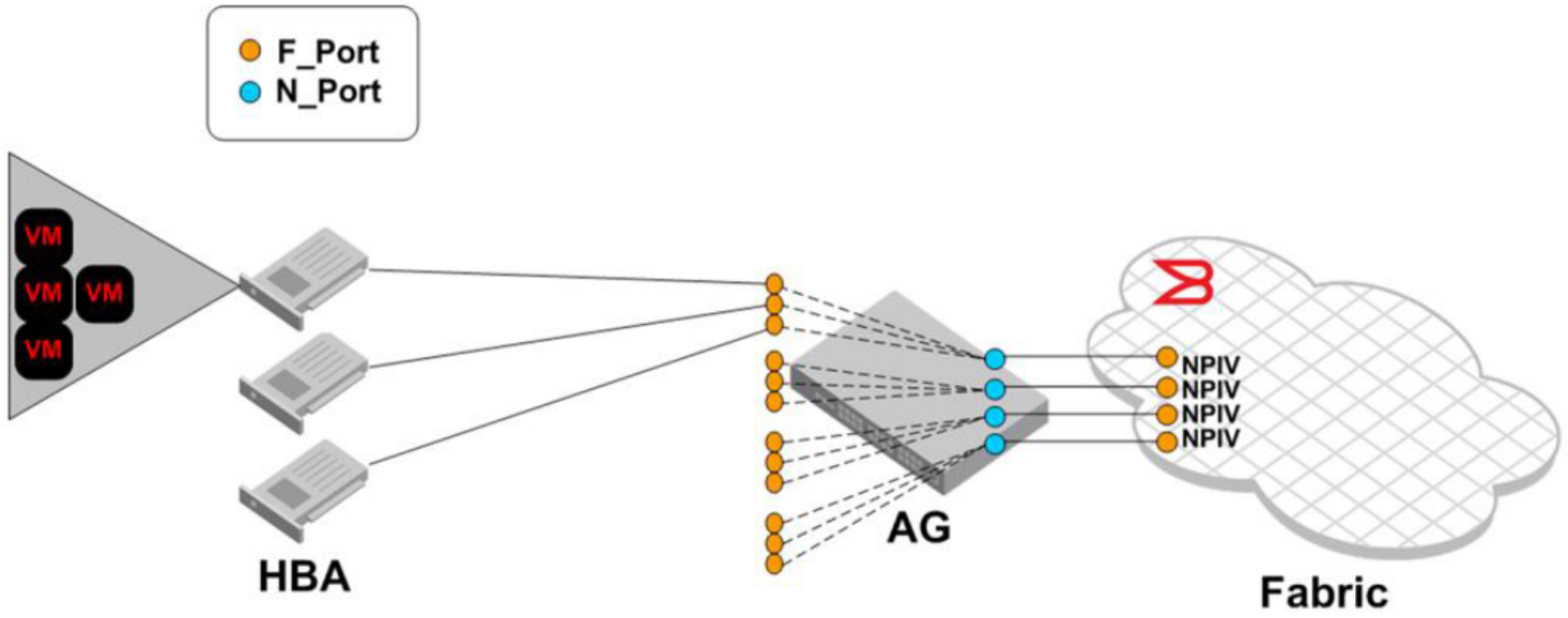
Access Gateway – Overview (cont.)

- The Access Gateway feature was introduced in FOS v5.2.1
- Access Gateway ports can be configured as N_Ports, which connect to the edge fabric
- No change in domain count - improves the scalability of the fabric
- Hosts/HBAs are mapped (through NPIV) to the N_Ports, and connect to the edge fabric through the N_Ports
- No fabric management or zoning on the Access Gateway



AG Provides Scalability

- Multiple F_Ports on an AG are mapped to a single N_Port on the same AG
- Several N_Ports on an AG can be connected to a fabric
- Every connection from an AG to a fabric can support a maximum of 255 devices, providing scalability for device attachment



Supported Platforms

- Supported on the Brocade 300, 5100, VA-40FC, embedded blade server switches, 6505, 6510, 8000, G610, G620

Brocade 300

Brocade 5100

Brocade VA-40FC

Brocade 4xxx and 5xxx Series
Embedded Blade Server Switches

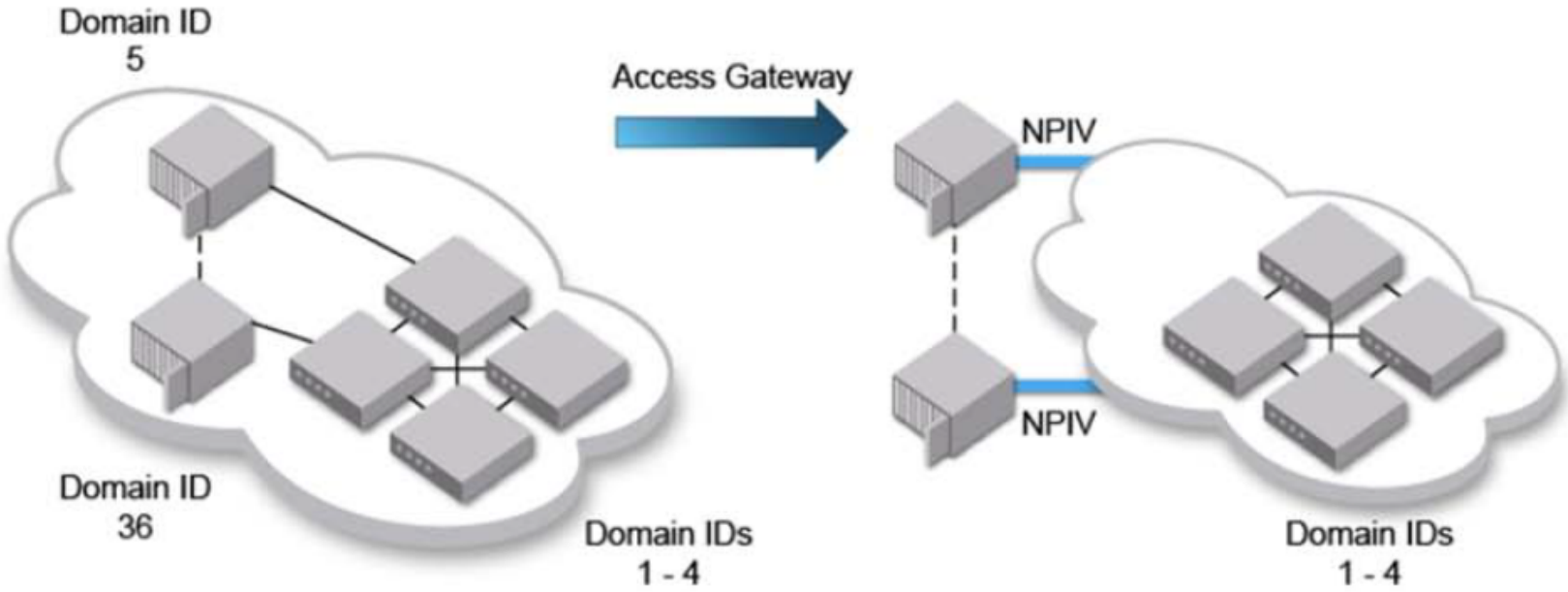
Brocade 6510

Brocade 8000

BROADCOM

Access Gateway Use Cases

- AG works well in large server count environments where management of multiple fabric domains is increasingly complex and limiting
- Mixed-vendor SAN configurations can utilize their full capabilities, without the restrictions of interoperability modes



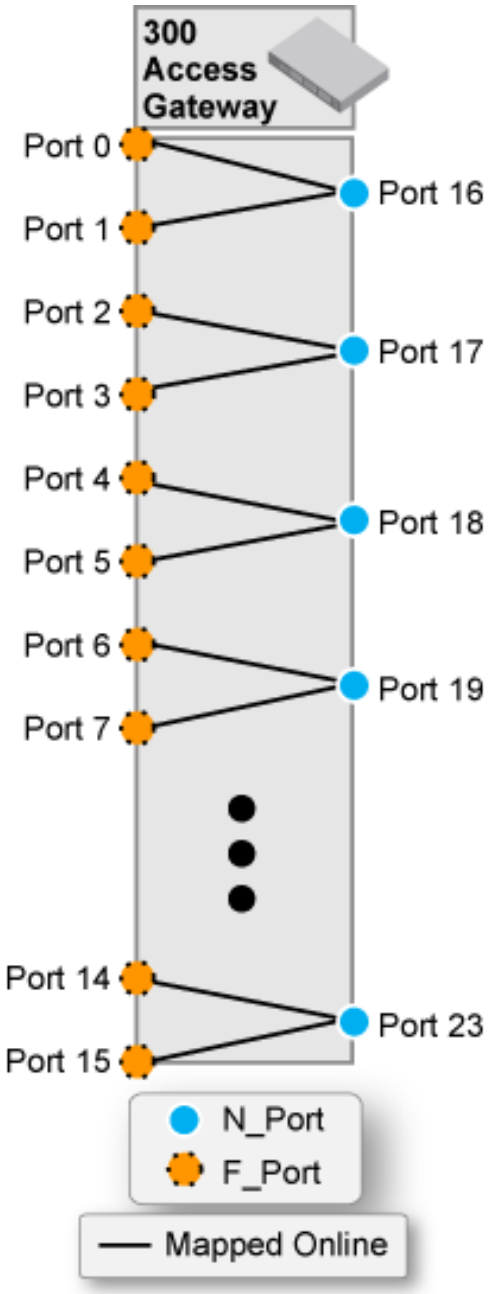
Traditional Brocade Blade Server SAN Switches Attached to SAN Fabric = 36 Domain IDs	Brocade Blade Server SAN Switches in Access Gateway Mode Attached to SAN Fabric = 4 Domain IDs

Access Gateway – Details

- When a switch is configured as an Access Gateway, the following are some of the features that are not supported:
  - Extended Fabrics
  - FICON (includes CUP)
  - IP over FC
  - Zoning
  - Virtual Fabrics
  - FC-FC Routing
  - Management Services
  - Name Services (SNS)
  - Port mirroring
  - SMI-S

Default Port Mapping

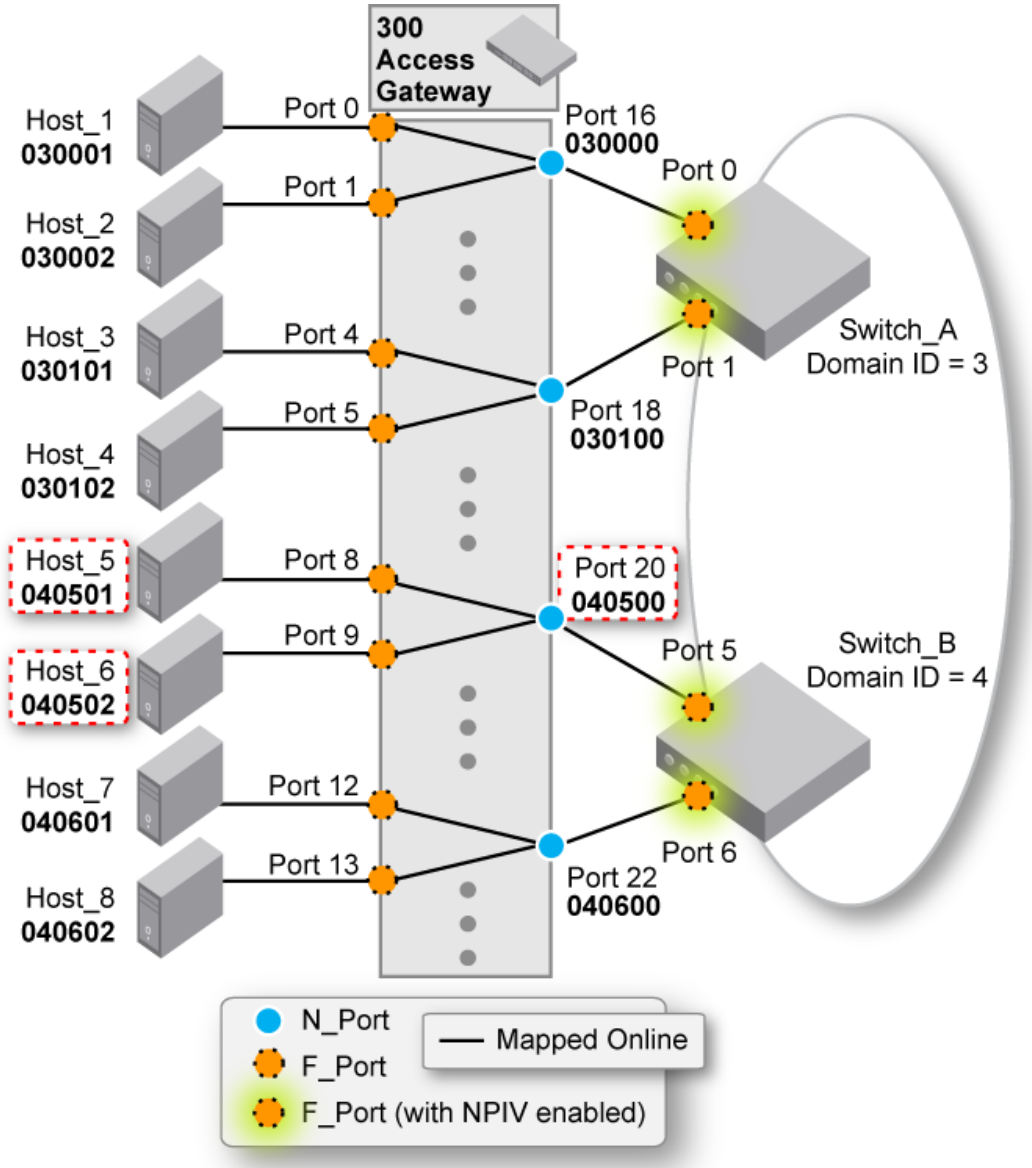
- Access Gateway uses a port map to direct traffic from host HBAs to the N_Ports that connect to the fabric
  - Each F_Port has a primary N_Port that is used to connect to the fabric
  - The port map and N_Port configuration can be changed
- Enabling Access Gateway enables the default port group. In this B300 example:
  - N_Ports: 16 - 23
  - Two F_Ports are mapped to each N_Port



BROADCOM

Access Gateway – Port Mapping (cont.)

- Access Gateway uses NPIV to assign the 24-bit FC address based on the port map
  - F_Ports (devices) share the same domain and area values as the N_Ports to which they are mapped
  - The last byte is assigned in the order in which the devices login to the fabric
- Example:
  - Port 20 address = 040500
  - Host_5 address = 040501
  - Host_6 address = 040502



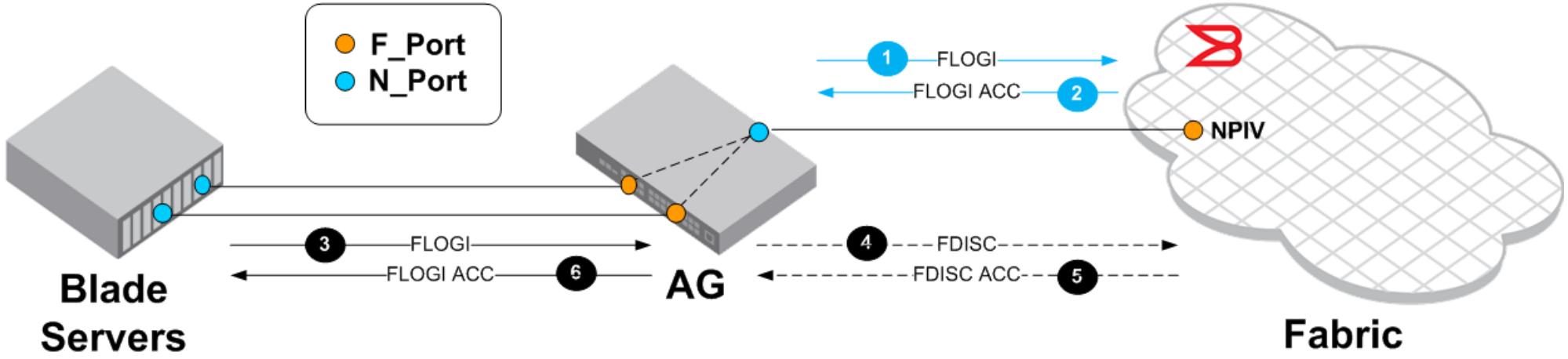
BROADCOM

Access Gateway – Initialization

- Access Gateway is logically transparent to the fabric, so the initialization of an Access Gateway occurs in stages:
  1. N_Ports come online first
  2. F_Ports are mapped to N_Ports according to the port map
  3. If the N_Port does not come online and failover is enabled, F_Ports are remapped
  4. F_Ports are enabled
  5. Hosts log in to the fabric

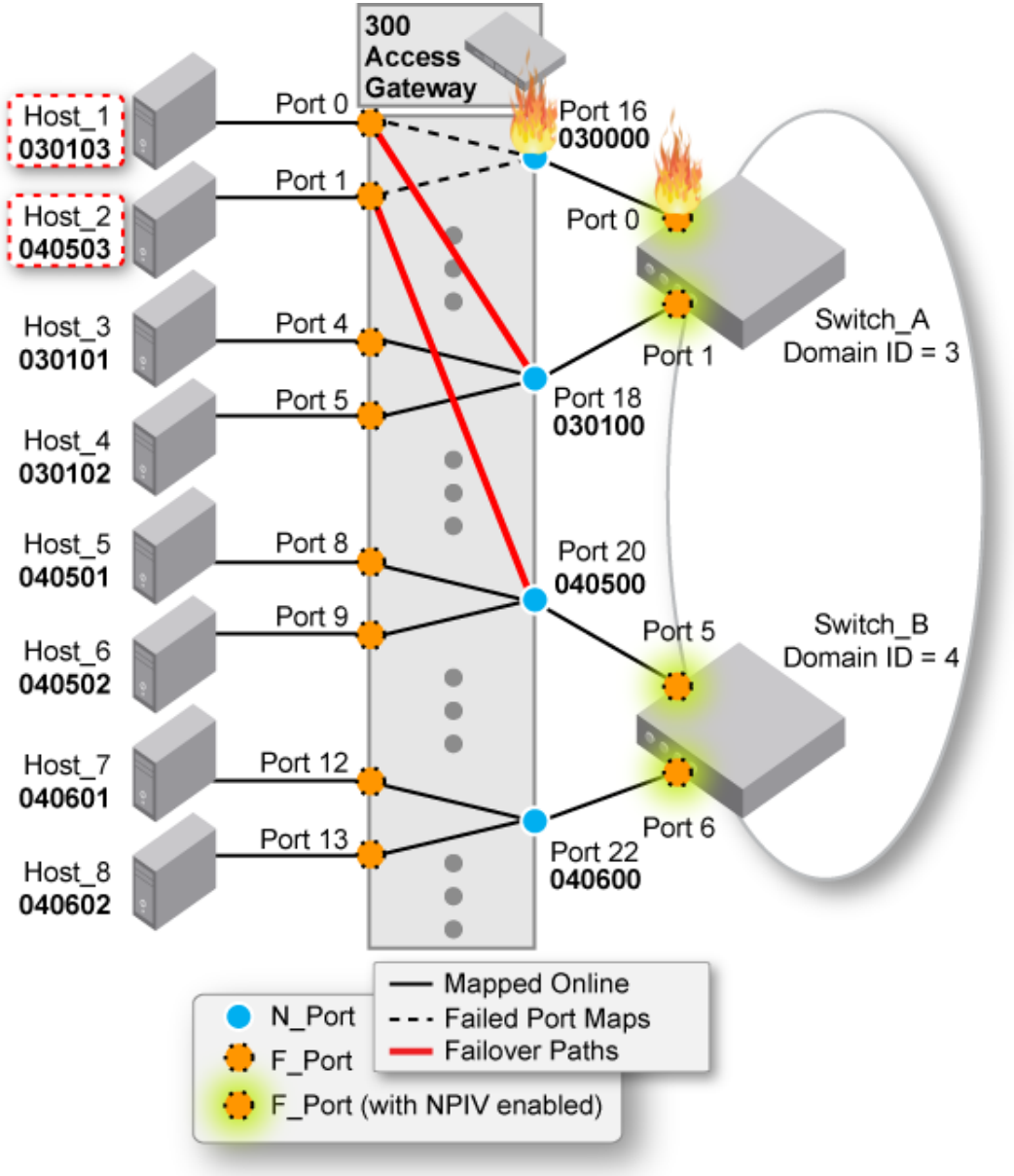
Access Gateway – Device Login

1. The N_Port of the AG logs into the fabric using a FLOGI
2. Fabric responds with a FLOGI ACC
3. Host issues a FLOGI request to the Access Gateway F_Port to which it is attached
4. Access Gateway passes the Fabric Login request to the N_Port to which the F_Port is mapped and the N_Port transforms the FLOGI request to a Fabric Discovery (FDISC)[1] request, which is transmitted to the fabric
5. Fabric Discovery Accept (FDISC ACC) response is received from the fabric
6. F_Port transforms the FDISC ACC response to a Fabric Login Accept (FLOGI ACC) response, which is transmitted to the host



Access Gateway – N_Port Failover

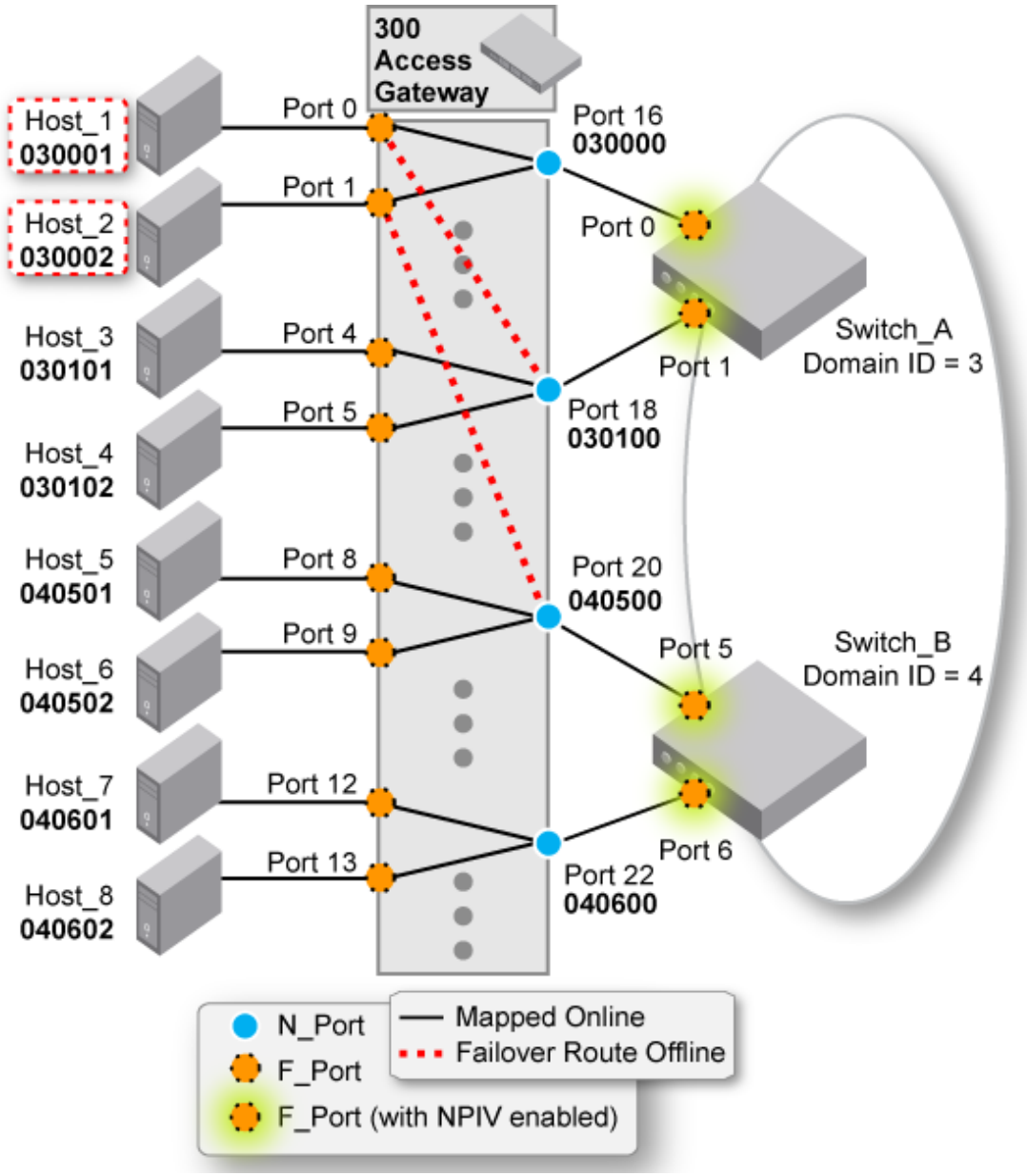
- If an N_Port goes offline, the Access Gateway **N_Port Failover policy** allows hosts to be automatically remapped to another online N_Port
  - F_Ports connected to failed N_Port are evenly distributed across N_Ports connected to the same fabric
  - F_Ports receive a new FC address based on the new N_Port
  - Enforced at N_Port initialization as well (*cold failover*)
  - The default configuration requires all N_Ports to be connected to the same fabric
  - Enabled by default; managed on a per-N_Port basis



BROADCOM

Access Gateway – N_Port Failback

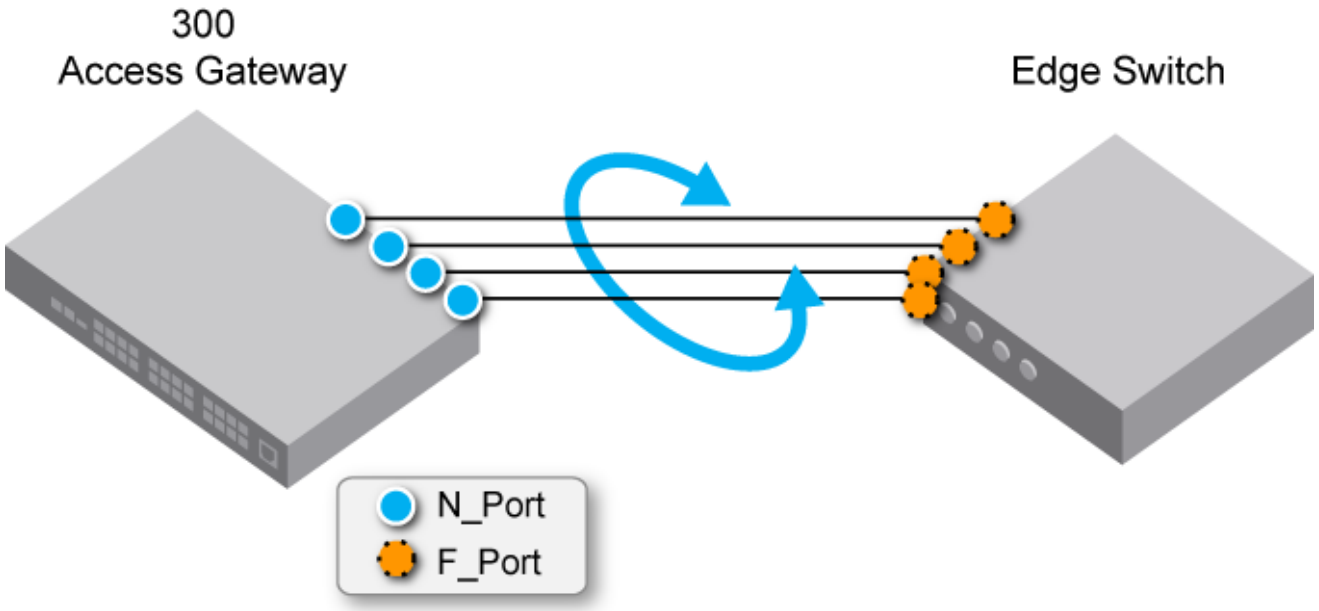
- If an N_Port comes back online, the Access Gateway **N_Port Failback policy** automatically remaps F_Ports *back* to the originally mapped Primary N_Port
  - Only the originally mapped F_Ports fail back
  - With multiple N_Port failures, only F_Ports that were mapped to the recovered N_Port experience failback
  - Failed-back F_Ports return to their original FC address
  - Enabled by default, can be changed; managed on a per-N_Port basis



BROADCOM

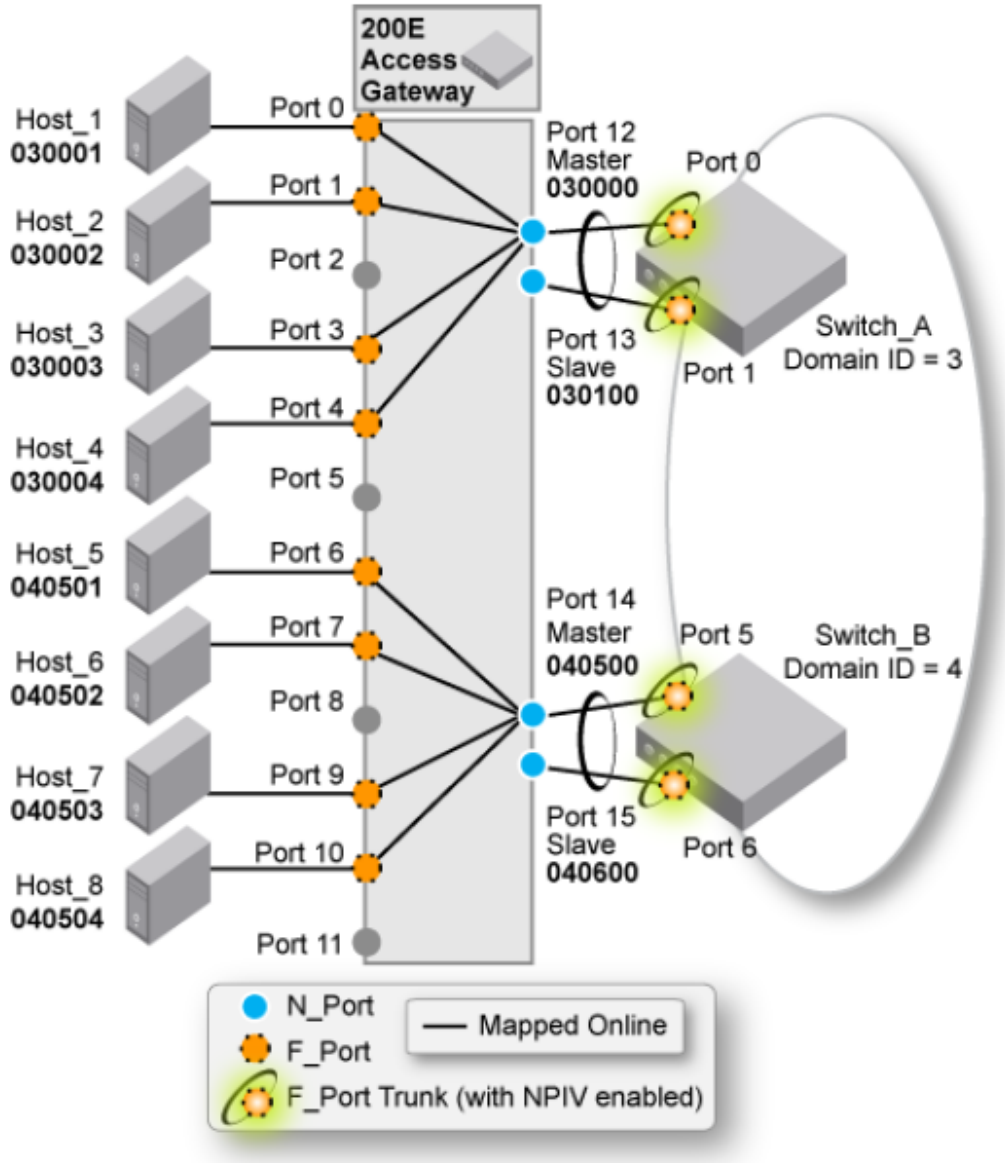
Access Gateway – F_Port Trunking

- Trunking aggregates the bandwidth of the ports within the trunk group
- Needs to be configured
  - F_Port Trunks do not automatically form
- Has the same requirements as ISL Trunking
  - Trunking license on both AG and edge fabric switch
  - Port group-to-port group
  - Same speed



Access Gateway – F_Port Trunking (cont.)

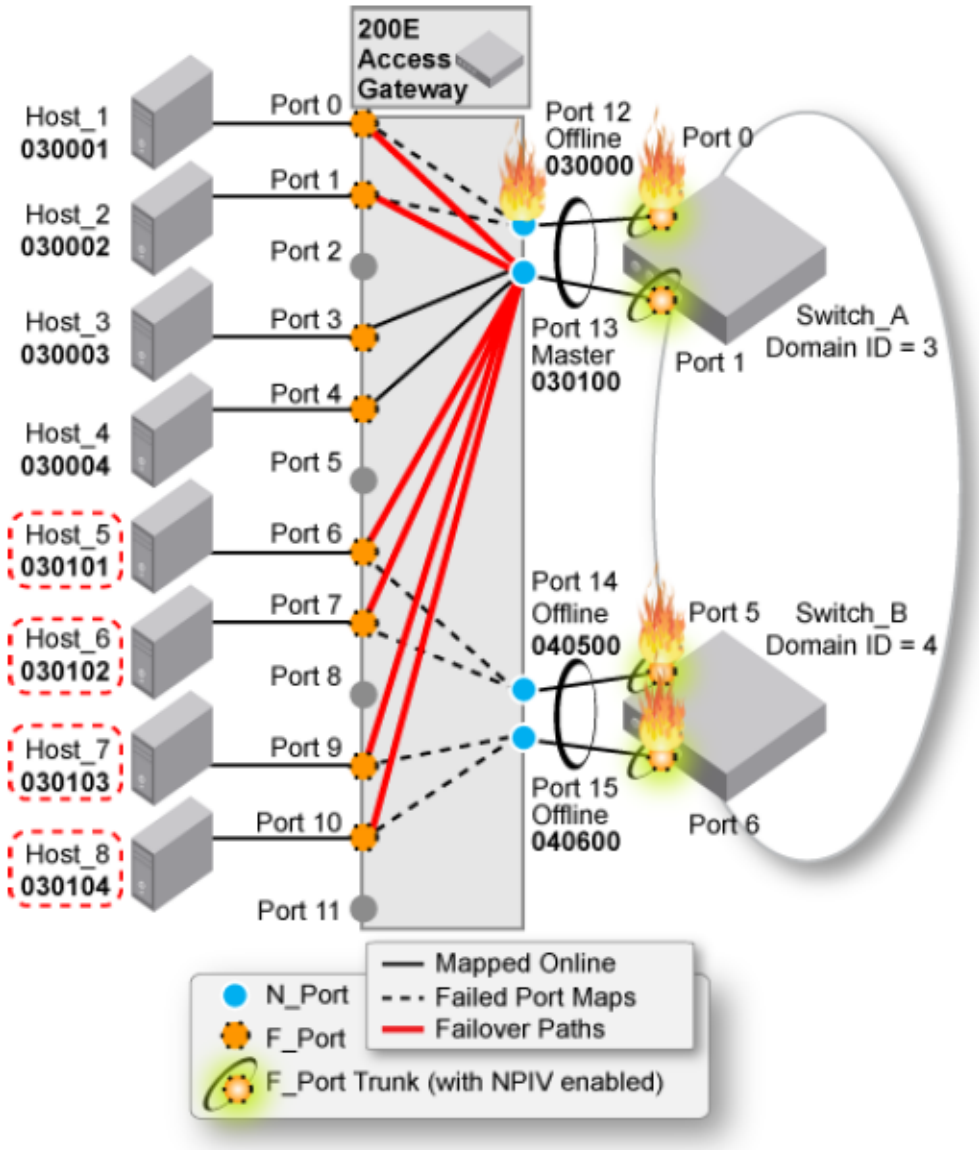
- On the Access Gateway no special configuration required other than trunking license
  - All AG N_Ports within the same F_Port Trunk TA/TI share the same Port ID
- F_Ports on the AG configured to any of the Trunk member N_Ports are mapped to the trunk master N_Port
- Frames received on any of these F_Ports are load balanced across all member N_Ports



BROADCOM

Access Gateway – F_Port Trunking Failover

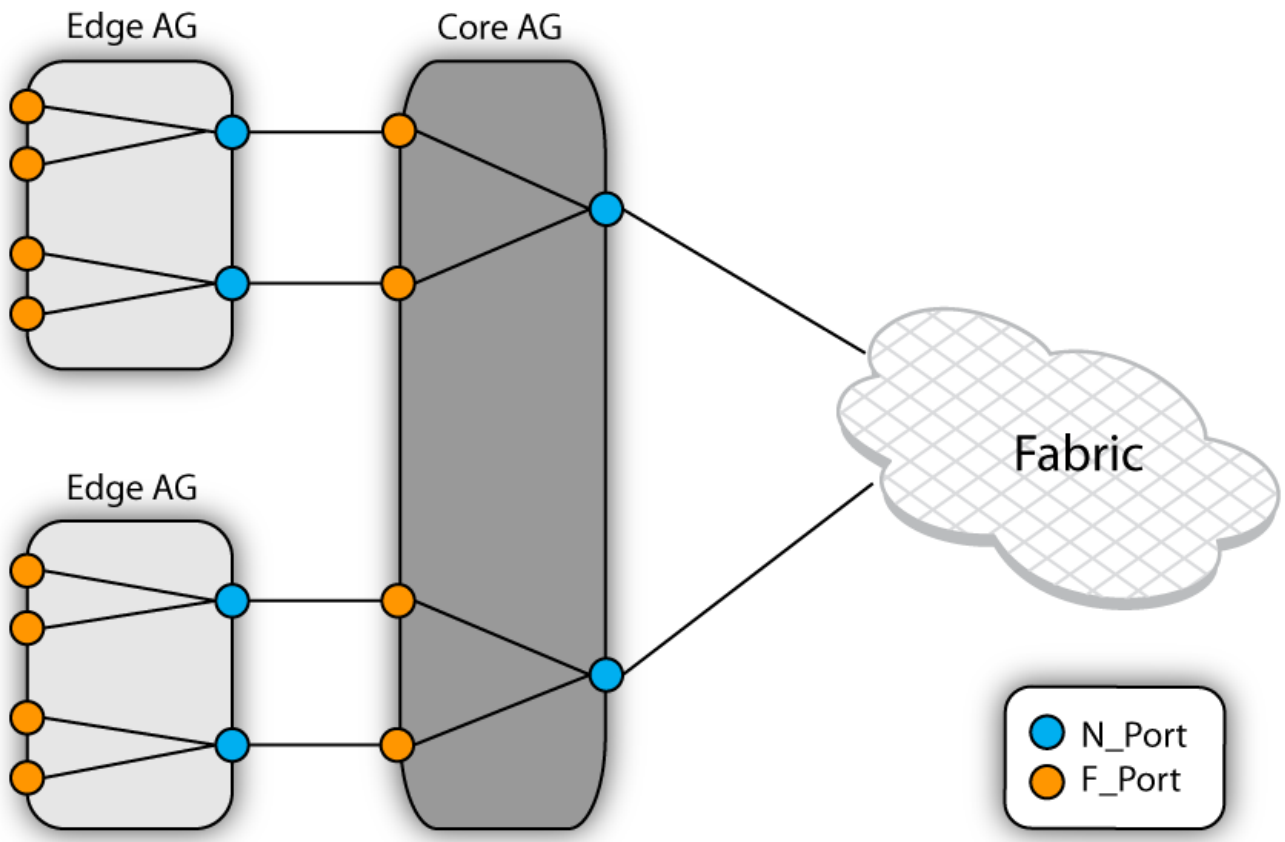
- If trunk master goes down, F_Ports are mapped to the new master N_Port without any PID change. i.e. No failover happens
  - Does not cause fabric disruptions (no Build Fabric or RSCN's)
- Failover is triggered when last member in the trunk goes offline
  - When trunk goes offline, F_Ports are failed over to other N_Ports as per failover policy and preferred failover settings



BROADCOM

Access Gateway – Cascading

- Connects two Access Gateways, linking one end as an N_Port and the other end as an F_Port
  - **Core AG**: Access Gateway connected to the fabric
  - **Edge AG**: Access Gateway connected to the devices
- Higher over-subscription while consolidating ports to main fabric
- No license requirement for cascading
- Device F_Ports may also be attached directly to core AG



Brocade Access Gateway

SOLVING INTEROPERABILITY, SCALABILITY, AND MANAGEMENT CHALLENGES

Simplify SAN Connectivity through NPIV Technology

- Deploys as a full-fabric switch or Access Gateway
- Connects transparently to Brocade and Cisco fabrics
- Eliminates additional switch domains and switch management tasks
- Accelerates server deployment and replacement with no disruptions for fabric reconfiguration

Maximizes Performance and Availability of the Fabric

- Supports frame-based trunking to optimize and balance performance, bandwidth, and availability
- Increases availability with non-disruptive fault recovery from path failure
- Isolates the SAN from disruptions due to server maintenance
- Leverages QoS to assure bandwidth for critical servers, virtual servers, or applications

BROADCOM

Continue the Conversation with Brocade Education

Connect with us today

- Brocade Education Website: www.Broadcom.com/brocade-education
  - Registering, Accessing and utilizing Brocade Education resources with videos
  - Course Catalog, Fundamentals Learning Path, FAQs, etc.
- LinkedIn: Brocade Education Group
  - New training announcements, Education & product updates, webinars, whitepapers
  - Automation, NVMe, PyFOS, Ansible, SANnav, Use-cases
- Facebook Brocade Education: www.facebook.com/BrocadeEducation
  - New training announcements, Education updates, etc.
- Broadcom Community: https://community.broadcom.com/
  - New training announcements, Blogs: what's new
  - Technical Q&As
- YouTube: Search Brocade Education
  - Overview videos
  - Course highlights, etc.

BROADCOM

Advance your skills with free on-demand Brocade SAN Training

Available 24x7 via the Learning Portal

Brocade GEN6 FIBRE CHANNEL

Brocade SANnav Management Portal

Rest API

Brocade Fabric Vision Technolog

Brocade SANnav Global View

Brocade Product Training

- **Accessing Brocade Education via Learning Portal**
  - Log On to Learning Portal here: www.broadcom.com/education
    - Register for Customer Support Portal: (New Users) www.broadcom.com/registration
    - Forgot Password for Account? Reset here: www.broadcom.com/forgotpassword
  - From Learning Portal use the search bar to search for training

- **Additional Learning Portal and Customer Support Portal Training:**
  - Guide to Brocade Education SAN Technical Training https://youtu.be/3HnAN2SMhfo
  - Brocade Fundamentals Curriculum Path Overview https://youtu.be/h-AvYYjZ-DA
  - Broadcom Learning Portal Overview Training https://youtu.be/1vAy-IApUJo
    - Accessing the Learning Portal, navigation, registering and completing training
  - Broadcom Customer Support Portal (CSP) Overview https://youtu.be/tdRoccA5Els
    - Registering, Logging in and navigation of the CSP

BROADCOM

EOF

Спасибо!

Узнайте больше:

www.brocade.com
http://www.linkedin.com/groups?gid=4246353
https://t.me/BrocadeRussiaSAN

BROADCOM®