TS Open Day– Data Center Fibre Channel over IP

Presented by:
Rong Cheng- TAC LAN & DCN China
Jan 30th, 2015

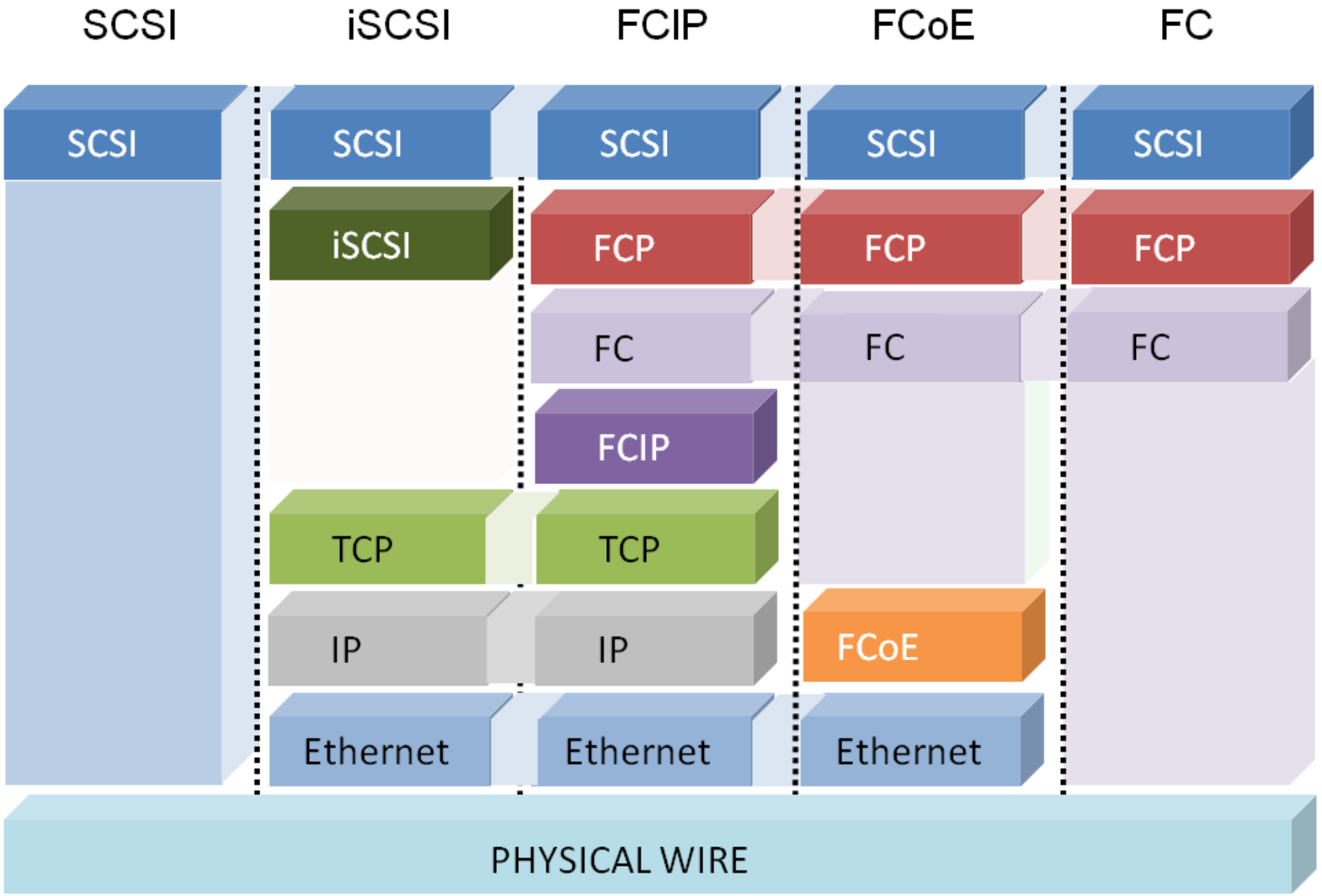
Agenda

- FCIP Introduction
- FCIP Configuration on MDS9000
- Troubleshooting FCIP

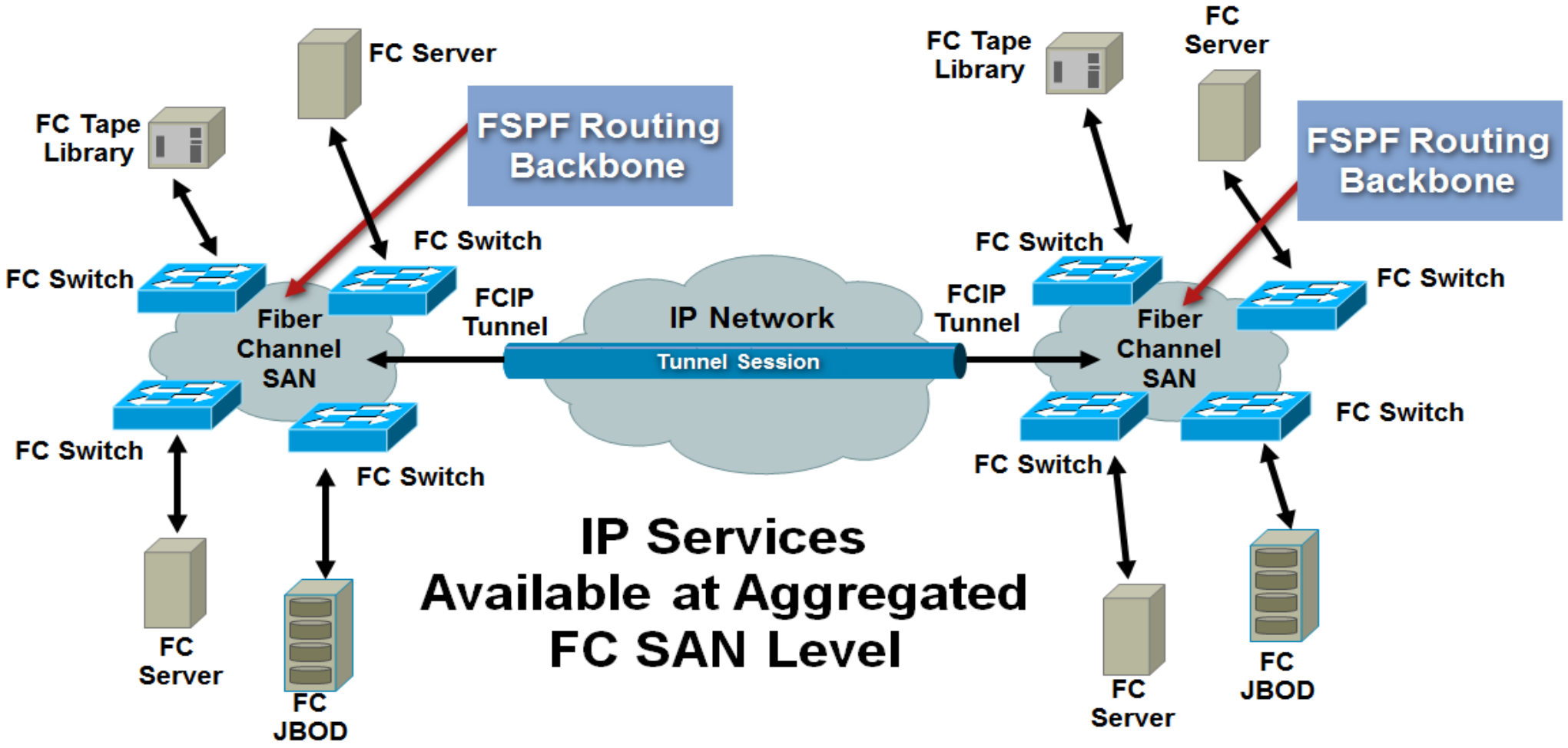
FCIP - Fibre Channel Over IP

- FCIP provides a standard way of encapsulating FC frames within TCP/IP, allowing islands of FC SANs to be interconnected over an IP-based network
- FC frames are treated the same as datagrams
- TCP/IP is used as the underlying transport to provide congestion control and in-order delivery of error-free data
- It is not FCoE, FCP or iSCSI.

Network Stack Comparison



FCIP Design



FCIP Architecture Model

Key:

FC-0 - Fibre Channel Physical Media Layer

FC-1 - Fibre Channel Encode and Decode Layer

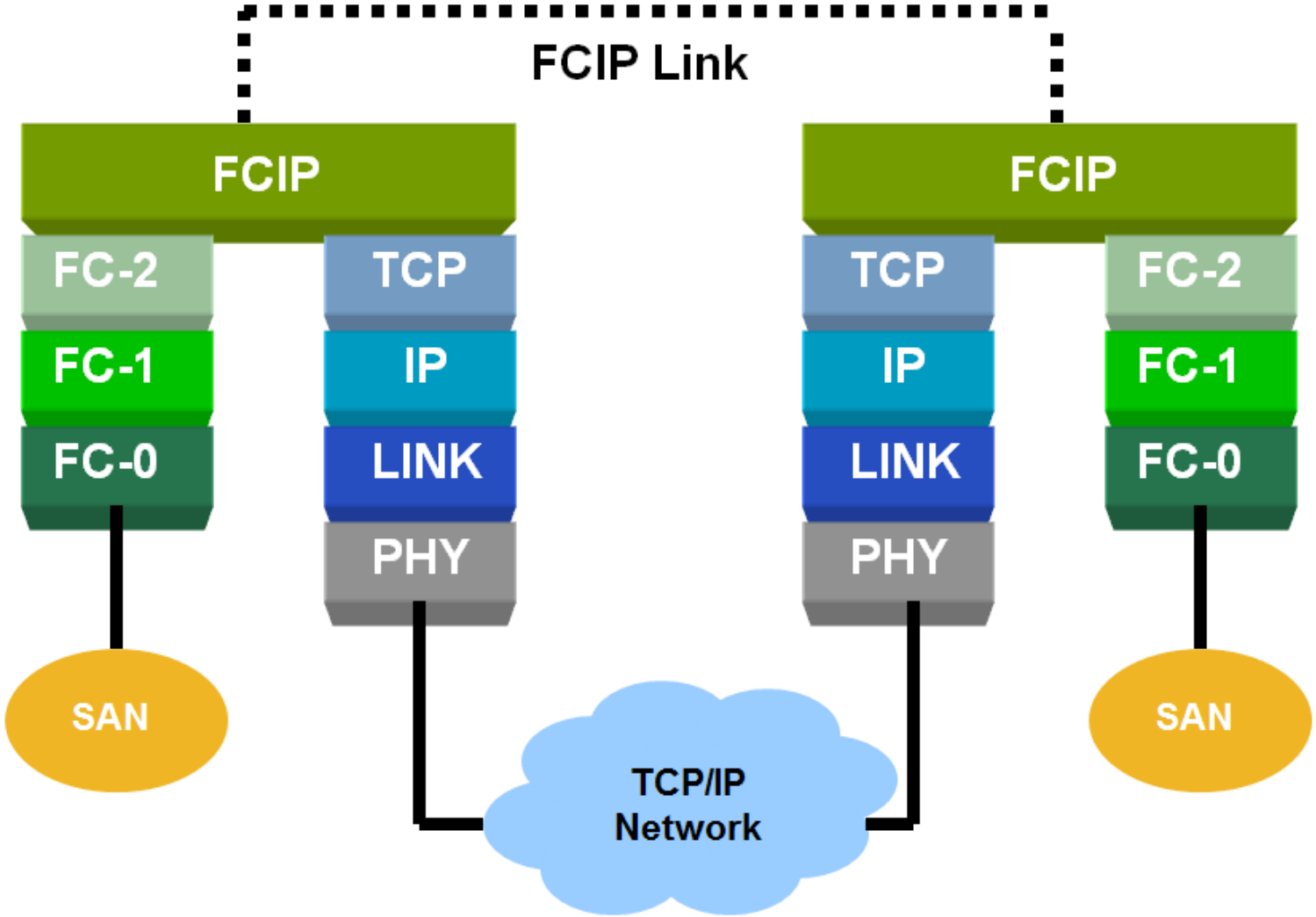
FC-2 - Fibre Channel Framing and Flow Control Layer

TCP - Transmission Control Protocol

IP - Internet Protocol

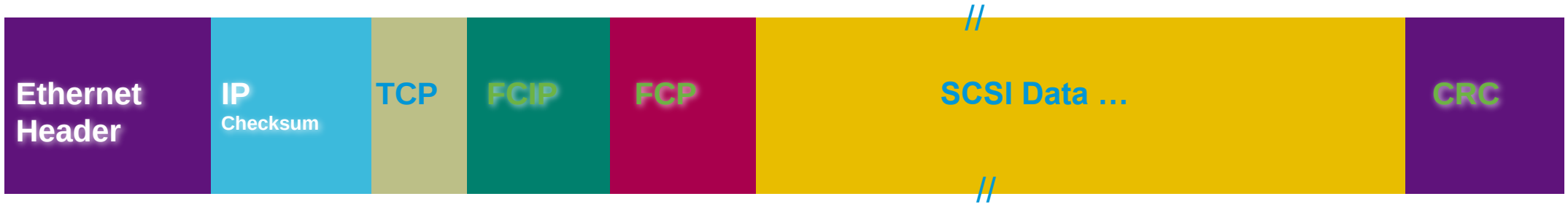
LINK - IP Link Layer

PHY - IP Physical Layer



FCIP

- End-station addressing, address resolution, message routing, and other fundamental elements of the network architecture remain unchanged from the Fibre Channel model, with IP introduced exclusively as a transport protocol for an inter-network bridging function
- IP is unaware of the Fibre Channel Payload and the Fiber channel fabric is unaware of IP



FCIP Connection Establishment

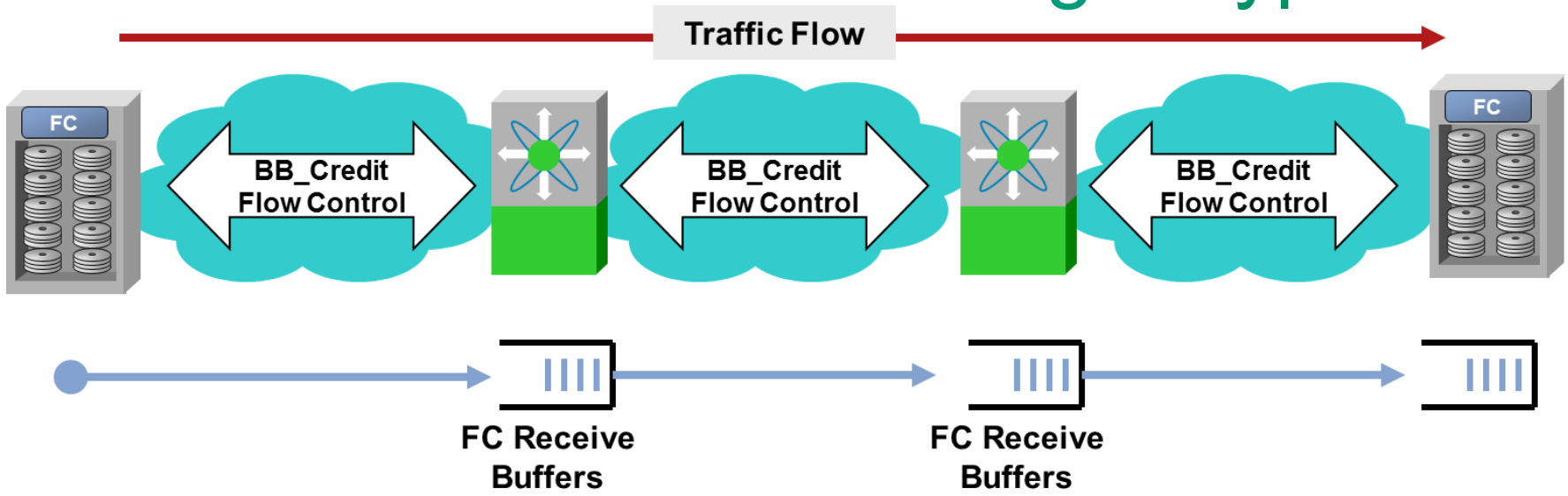
- Non-Dynamic **(This is the connection type the MDS9000 uses)**

  TCP connection to a specific IP address

- Dynamic **(Feature not supported or used with MDS9000 FCIP implementation)**

  Discovery of FCIP entities using SLPv2

Fibre Channel Frame Buffering – Typical FC



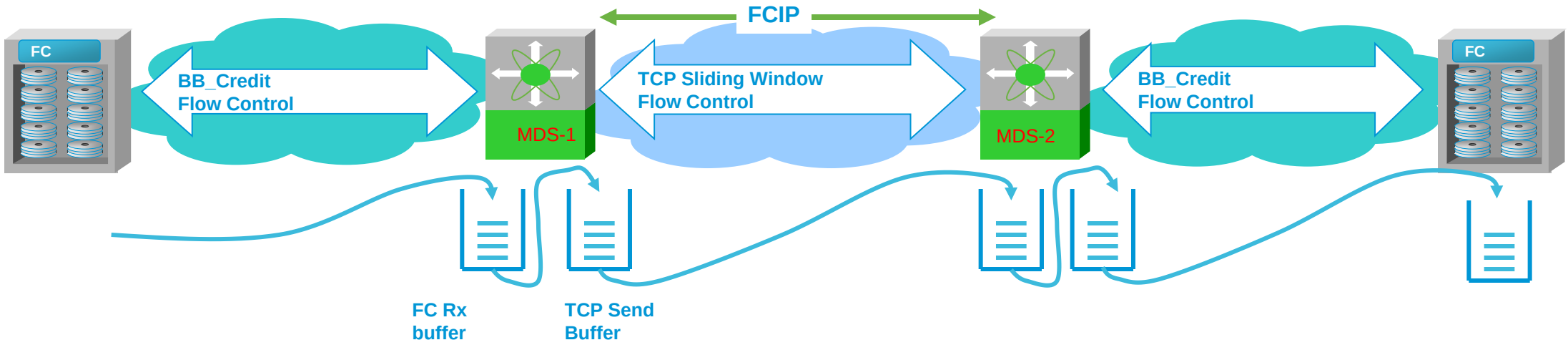
- FC frames buffered and queued in intermediate switches
- Buffer depth controlled by Rx BB_Credit parms

  `switchport fcrxbbcredit <x>`

  Performance buffers added by default unless configured otherwise

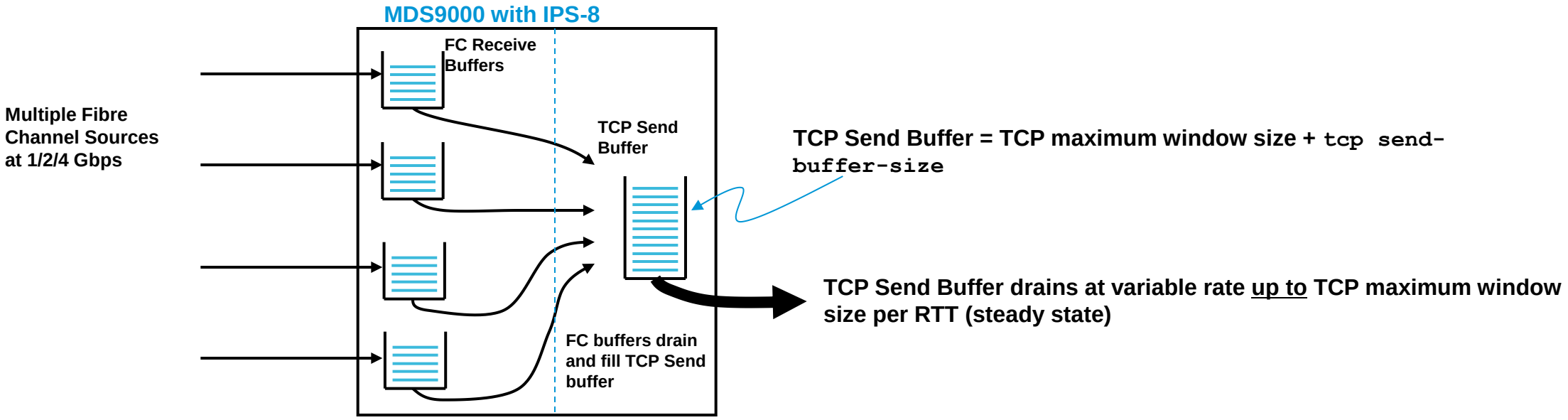
- Hop-by-hop traffic flow paced by return of R_RDY **(Receiver ready)** frames
- Frames cannot sit in any switch for >500ms (or they will expire)

FC & FCIP Combined



- TCP uses a sliding window for flow control
  - end to end between FCIP entities
  - BB_Credits not relevant over FCIP
- FC → FCIP is typical point where problems may occur
  - Multiple 1 or 2 Gbps FC sources to slow, high latency WAN path

FC and TCP Buffering – an example



An example:
Four FC sources connected to ports each with 32 buffers (configured as switchport fcrxbbcredits 32) = total 128 buffers

To avoid frame expiry, each buffer must be able to empty in less than 500ms.
So, minimum aggregate drain rate = 2 x 128 or 256 frames/second

Assuming TCP send Buffer is already full, TCP must be able to send at > 256 Frames/second. At 2148 Bytes/frame, this is only (2148 x 8 x 256) = 4.4Mbps.

Agenda

- FCIP Introduction
- FCIP Configuration on MDS9000
- Troubleshooting FCIP

Cisco IPS Module – FCIP Basic Configuration Steps

- Perform these basic configuration steps on both MDS 9000 switches to configure IPS modules and FCIP links.
  - Step 1 - Configure the Gigabit Ethernet interface(s).
  - Step 2 - Create a FCIP profile and assign a Gigabit Ethernet interface IP address to that profile.
  - Step 3 - Create a FCIP interface and assign a profile to that interface.
  - Step 4 - Configure peer information for the FCIP interface(s).
  - Step 5 - Enable the interface(s).

Gigabit Ethernet Interface Configuration

Basic Gigabit Ethernet configuration

IP address and subnet mask, IP MTU frame size, and port enable

Static IP Routing

IP Address and Subnet Mask

	Command	Purpose
Step 1	switch# config terminal switch(config)#	Enters configuration mode.
Step 2	switch(config)# interface gigabitethernet 3/1 switch(config-if)#	Enters the interface configuration mode on the Gigabit Ethernet interface (slot3, port 1).
Step 3	switch(config-if)# ip address 10.100.1.25 255.255.255.0	Enters the IP address (10.100.1.25) and subnet mask (255.255.255.0) for the Gigabit Ethernet interface.
Step 4	switch(config-if)# no shutdown	Enables the interface.

IP MTU Frame Size

- Maximum Transmit Unit (MTU) Size
  - Default IP MTU = 1500 bytes
  - Jumbo frames supported (9000 bytes)

	Command	Purpose
Step 1	switch# config terminal switch(config)#	Enters configuration mode.
Step 2	switch(config-if)# switchport mtu 2300 switch(config-if)#	Changes the IP maximum transmission unit (MTU) to 2300. The default is 1500.

Managing IP Routing

- To configure static IP routing through the Gigabit Ethernet interface, follow these steps:

	Command	Purpose
Step 1	switch1# config terminal switch1(config)#	Enters configuration mode.
Step 2	switch1(config)# ip route 10.1.1.1 255.255.255.255 10.100.1.x interface gigabit 3/1 switch1(config-if)#	Enter the routing entry to SW2(10.1.1.1) with configuring the next hop(10.100.1.x), which is the IP address of the router connected to the Gigabit Ethernet 3/1 interface.

Displaying the IP Route Table

```
switch#
```

show ips ip route interface ethernet

- **Takes the Ethernet interface as a parameter and returns the route table for that interface**

```
SW1# show ips ip route interface gigabitethernet 3/1
Codes: C - connected, S - static

No default gateway

S 10.1.1.1/32 via 10.100.1.x, GigabitEthernet3/1
C 10.100.1.0/24 is directly connected, GigabitEthernet3/1
```

Connected (C) identifies the subnet in which the interface is configured (directly connected to the interface). Static (S) identifies the static routes that go through the router.

Verify Gigabit Ethernet Connectivity

```
switch1#
```

ping

- **Sends echo request packets out to a remote device at a specified IP address**

```
switch1# ping 10.1.1.1
PING 10.1.1.1 (10.1.1.1): 56 data bytes
64 bytes from 10.1.1.1: icmp_seq=0 ttl=255 time=0.1 ms
64 bytes from 10.1.1.1: icmp_seq=1 ttl=255 time=0.1 ms
64 bytes from 10.1.1.1: icmp_seq=2 ttl=255 time=0.1 ms
--- 10.1.1.1 ping statistics ---
3 packets transmitted, 3 packets received, 0% packet loss
round-trip min/avg/max = 0.1/0.1/0.1 ms
```

Note: Extended Ping Parameters are also supported on MDS, use them to troubleshoot

Creating FCIP Profiles – Switch 1

- Create a profile for each Gigabit Ethernet Interface

To create a FCIP profile for Switch 1 above, follow these steps:

	Command	Purpose
Step 1	switch1# config terminal switch1(config)#	Enters configuration mode.
Step 2	switch1(config)# fcip profile 10 switch1(config-profile)#	Creates a profile for the FCIP connection. The valid range is from 1 to 255.
Step 3	switch1(config-profile)# ip address 10.100.1.25	Associates the profile (10) with the local IP address of the Gigabit Ethernet interface (3/1).

Creating FCIP Profiles – Switch 2

- Create a profile for each Gigabit Ethernet Interface

To create a FCIP profile for Switch 2 above, follow these steps:

	Command	Purpose
Step 1	switch2# config terminal switch2(config)#	Enters configuration mode.
Step 2	switch2(config)# fcip profile 20 switch2(config-profile)#	Creates a profile for the FCIP connection.
Step 3	switch2(config-profile)# ip address 10.1.1.1	Associates the profile (20) with the local IP address of the Gigabit Ethernet interface.

Local TCP Port Association

To enter the switch(config-profile)# prompt, follow these steps:

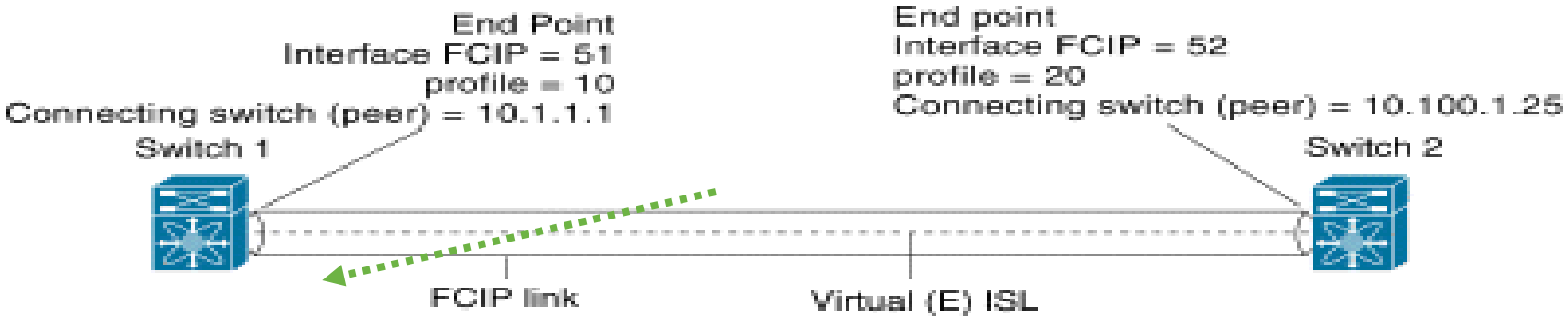
	Command	Purpose
Step 1	switch# config terminal switch(config)#	Enters configuration mode.
Step 2	switch(config)# fcip profile 20 switch(config-profile)#	Creates the profile (if it does not already exist). The valid range is from 1 to 255.

- To change the default FCIP port number (3225), follow this step:

	Command	Purpose
Step 1	switch(config-profile)# port 5000	Associates the profile with the local port number (5000).
	switch(config-profile)# no port	Reverts to the default 3225 port.

Peer will have to be configured to open FCIP connection to non-default TCP port

Creating FCIP Interfaces – Switch 1

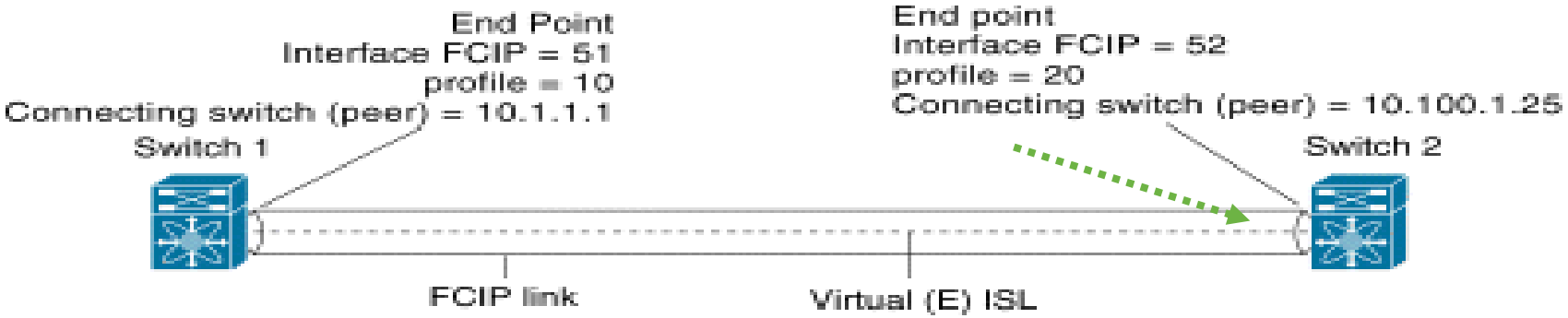


- Create FCIP interface, assign a profile, set peer address, and enable

To create and enable a FCIP interface for Switch 1 above, follow these steps:

	Command	Purpose
Step 1	switch1# config terminal switch(config)#	Enters configuration mode.
Step 2	switch1(config)# interface fcip 51 switch1(config-if)#	Creates a FCIP interface (51).
Step 3	switch1(config-if)# use-profile 10	Assigns the profile (10) to the FCIP interface.
Step 4	switch1(config-if)# peer-info ipaddr 10.1.1.1 *(port)*	Assigns the peer IP address information (10.100.1.25 for switch 2) to the FCIP interface. If not specified the port will be 3225.
Step 5	switch1(config-if)# no shutdown	Enables the interface.

Creating FCIP Interfaces – Switch 2



- Create FCIP interface, assign a profile, set peer address, and enable

To create and enable a FCIP interface for Switch 2 above, follow these steps:

	Command	Purpose
Step 1	switch2# config terminal switch(config)#	Enters configuration mode.
Step 2	switch2(config)# interface fcip 52 switch2(config-if)#	Creates a FCIP interface (52).
Step 3	switch2(config-if)# use-profile 20	Binds the profile (20) to the FCIP interface.
Step 4	switch2(config-if)# peer-info ip address 10.100.1.25 *(port)*	Assigns the peer IP address information (10.1.1.1 for switch 1) to the FCIP interface
Step 5	switch1(config-if)# no shutdown	Enables the interface.

TCP Minimum Retransmit Timeout

- **Controls the minimum amount of time TCP waits before retransmitting. By default, this value is 300 milliseconds.**

	Command	Purpose
Step 1	switch(config-profile)# tcp min-retransmit-time 500	Specifies the minimum TCP retransmit time for the TCP connection in milliseconds (500). The default is 300 milliseconds and the range is from 250 to 5000 milliseconds.
	switch(config-profile)# no tcp min-retransmit-time 500	Reverts the minimum TCP retransmit time to the factory default of 300 milliseconds.

TCP Keepalive Timeout

- **Configure the interval between which the TCP connection verifies if the FCIP link is functioning.**

	Command	Purpose
Step 1	switch(config-profile)# tcp keepalive-timeout 120	Specifies the keepalive timeout interval for the TCP connection in seconds (120). The default is 60 seconds. The range is from 1 to 7200 seconds.
	switch(config-profile)# no tcp keepalive-timeout 120	Reverts the keepalive-timeout to 60 seconds.

TCP Maximum Retransmissions

- **Specifies the maximum number of times a packet is retransmitted before TCP decides to close the connection.**

	Command	Purpose
Step 1	switch(config-profile)# tcp max-retransmissions 6	Specifies the maximum number of retransmissions (6). The default is 4 and the range is from 1 to 8 retransmissions.
	switch(config-profile)# no tcp max-retransmissions 6	Reverts to the default of 4 retransmissions.

TCP Window Management

- **Optimal TCP window size is computed using three options.**
  - **Maximum Bandwidth**
  - **Minimum Available Bandwidth**
  - **Round Trip Time**

	Command	Purpose
Step 1	switch(config-profile)# tcp max-bandwidth-mbps 900 min-available-bandwidth-mbps 300 round-trip-time-ms 10	Configures the maximum available bandwidth at 900 Mbps, the minimum slow start threshold as 300 Mbps, and the round trip time as 10 milliseconds.
	switch(config-profile)# no tcp max-bandwidth-mbps 900 min-available-bandwidth-mbps 300 round-trip-time-ms 10	Reverts to the factory defaults. The defaults are max-bandwidth = 1G, min-available-bandwidth = 2 Mbps and round-trip-time is 10ms.
	switch(config-profile)# tcp max-bandwidth-kbps 2000 min-available-bandwidth-kbps 2000 round-trip-time-us 200	Configures the maximum available bandwidth at 2000 Kbps, the minimum slow start threshold as 2000 Kbps, and the round trip time as 200 microseconds.

TCP Buffer Size

- **Defines the required additional buffering—beyond the normal send window size —that TCP allows before flow controlling the switch's egress path for the FCIP interface. The default buffer size is 0 KB.**

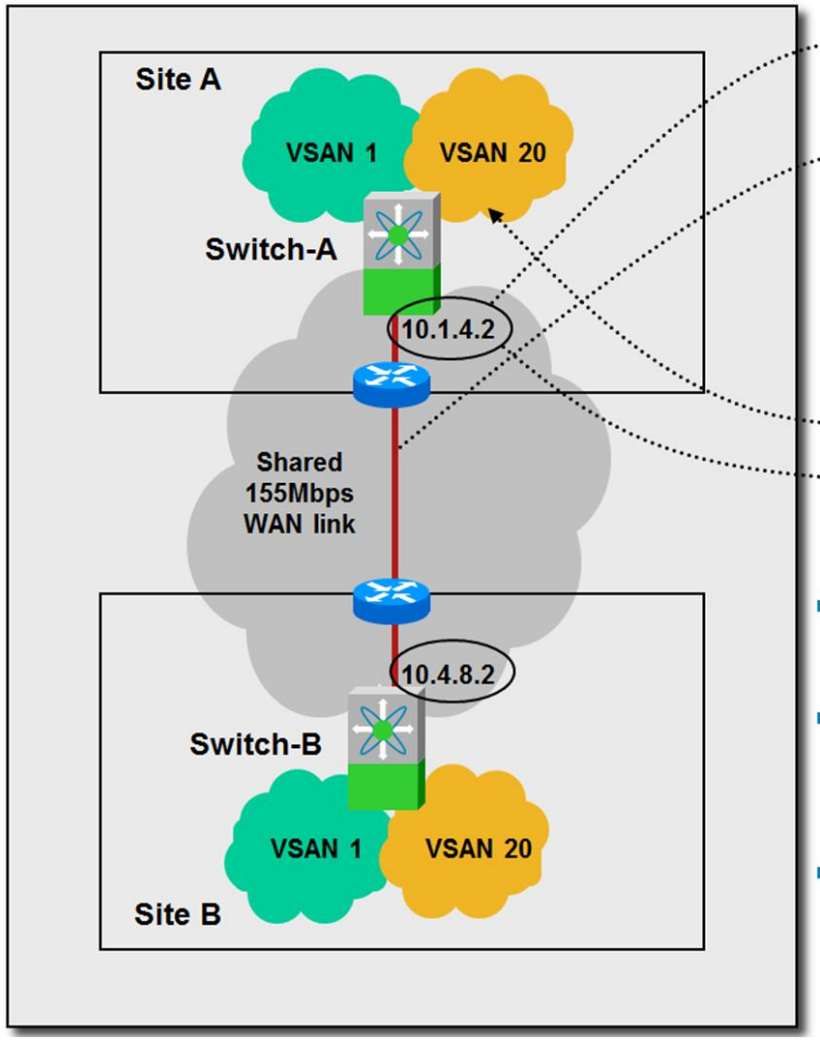
	Command	Purpose
Step 1	switch(config-profile)# tcp send-buffer-size 5000	Configure the advertised buffer size to 5000 KB. The valid range is from 0 to 8192 KB.
	switch(config-profile)# no tcp send-buffer-size 5000	Reverts the switch to its factory default (0 KB).

Enable or Disable Time Stamps

- If the time-stamp option is enabled, be sure to configure the Network Time Protocol (NTP) on both switches. Do not use with tape or write acceleration.
- To enable or disable the time-stamp option, follow these steps on the FCIP interface:

	Command	Purpose
Step 1	switch(config-if)# time-stamp Please enable NTP with a common time source on both MDS Switches that are on either side of the FCIP link	Enables time stamp checking for received packets with a default acceptable time difference of 2000 milliseconds.
	switch(config-if)# no time-stamp	Disables (default) time stamps.
Step 2	switch(config-if)# time-stamp acceptable-diff 4000	Configures the acceptable time within which a packet is accepted. The default difference is a 2000 millisecond interval from the network time. The valid range is from 1 to 60,000 milliseconds.
	switch(config-if)# no time-stamp acceptable-diff 4000	Deletes the configured time difference and reverts the difference to factory defaults.
Step 3	switch(config-if)# no shutdown	Enables the interface.

FCIP Configuration Example: MDS9000




```
fcip profile 10
  ip address 10.1.4.2
  tcp max-bandwidth-mbps 155 min-available-bandwidth-mbps 20 round-
trip-time-ms 1

interface fcip50
  switchport mode E
  no shutdown
  switchport trunk allowed vsan 1
  switchport trunk allowed vsan add 20
  use-profile 10
  peer-info ipaddr 10.4.8.2

interface GigabitEthernet2/5
  ip address 10.1.4.2 255.255.255.0
  switchport mtu 2300
  no shutdown
```

RTT will autconfigure and adapt to network changes during idle periods

Jumbo Frame MTU - 2300 Bytes will handle largest FC frame

- Three steps for FCIP config – Profile, GigE i/f and FCIP i/f
- Min-bandwidth set to minimum bandwidth available (through QoS or other means). Sender will start at this rate
- Peer FCIP interface configured similarly

Agenda

- FCIP Introduction
- FCIP Configuration on MDS9000
- Troubleshooting FCIP

Verifying FCIP Status

```
MDS# show fcip summary

--------------------------------------------------------------------------------
Tun prof     Eth-if      peer-ip          Status T W T Enc Comp   Bandwidth     rtt
                                                 E A A                max/min     (us)
--------------------------------------------------------------------------------
2    1                   192.168.10.3     DOWN   N N N N   N      1000M/773M  1000
3    3                   10.6.32.251      DOWN   N N N N   N      1000M/500M  1000
10   10     GE1/1        10.10.1.2        TRNK   Y N N N   N      1000M/950M  250
11   11     GE1/4        10.10.2.2        TRNK   Y N N N   N      1000M/950M  70
```

Verifying FCIP Profile status

```
MDS# show fcip profile 10      <----- this is the profile linked at 10 (fcip10)
FCIP Profile 10
    Internet Address is 10.10.1.1  (interface GigabitEthernet1/1)
    Tunnels Using this Profile: fcip10
    Listen Port is 3225
    TCP parameters
        SACK is enabled
        PMTU discovery is enabled, reset timeout is 3600 sec
        Keep alive is 60 sec
        Minimum retransmission timeout is 200 ms
        Maximum number of re-transmissions is 4
        Send buffer size is 0 KB
        Maximum allowed bandwidth is 1000000 kbps
        Minimum available bandwidth is 950000 kbps
        Configured round trip time is 250 usec
        Congestion window monitoring is enabled, burst size is 50 KB
        Auto jitter detection is enabled
```

```
MDS# show fcip profile 1      <--- this is the profile linked to tunnel 2 (fcip2)
FCIP Profile 1
    Internet Address is 192.168.10.2
    Tunnels Using this Profile: fcip2
    Listen Port is 3225
    TCP parameters
        SACK is enabled
        PMTU discovery is enabled, reset timeout is 3600 sec
        Keep alive is 60 sec
        Minimum retransmission timeout is 200 ms
        Maximum number of re-transmissions is 4
        Send buffer size is 0 KB
        Maximum allowed bandwidth is 1000000 kbps
        Minimum available bandwidth is 773560 kbps
        Configured round trip time is 1000 usec
        Congestion window monitoring is enabled, burst size is 50 KB
        Auto jitter detection is enabled
```

Verifying TCP Status

```
MDS# show ips stats tcp interface gigabitethernet 1/1

TCP statistics for port GigabitEthernet1/1
    Connection Stats
      0 active openings, 4 accepts
      0 failed attempts, 0 reset received, 4 established
    Segment stats
      5160052 received, 5151305 sent, 0 retransmitted
      0 bad segments received, 0 reset sent

    TCP Active Connections
      Local Address          Remote Address          State       Send-Q     Recv-Q
      10.10.1.1:3225         0.0.0.0:0               LISTEN      0          0
      10.10.1.1:3225         10.10.1.2:65525         ESTABLISH   0          0
      10.10.1.1:3225         10.10.1.2:65523         ESTABLISH   0          0
```

Checking FCIP Counters

```
MDS# show fcip counters
fcip10
    TCP Connection Information
      2 Active TCP connections
        Control connection: Local 10.10.1.1:3225, Remote 10.10.1.2:65523
        Data connection: Local 10.10.1.1:3225, Remote 10.10.1.2:65525
      0 Attempts for active connections, 2 close of connections
    TCP Parameters
      Path MTU 2300 bytes
      Current retransmission timeout is 200 ms
      Round trip time: Smoothed 2 ms, Variance: 4 Jitter: 150 us
      Advertized window: Current: 1023 KB, Maximum: 24580 KB, Scale: 4
      Peer receive window: Current: 31 KB, Maximum: 31 KB, Scale: 4
      Congestion window: Current: 29 KB, Slow start threshold: 30 KB
      Current Send Buffer Size: 25 KB, Requested Send Buffer Size: 0 KB
      CWM Burst Size: 50 KB
      Measured RTT : 500000 us Min RTT: 41 us Max RTT: 0 us
    5 minutes input rate 1016 bits/sec, 127 bytes/sec, 1 frames/sec
    5 minutes output rate 1240 bits/sec, 155 bytes/sec, 1 frames/sec
      2478970 frames input, 254854240 bytes
        2478964 Class F frames input, 254853112 bytes
        6 Class 2/3 frames input, 1128 bytes
        0 Reass frames
        0 Error frames timestamp error 0
      2648632 frames output, 344279856 bytes
        2648622 Class F frames output, 344279056 bytes
        10 Class 2/3 frames output, 800 bytes
        0 Error frames
```

Verifying MTU
Is it what you think it is?

Musky-9506# clear counters interface gigabitethernet 3/3

show interface GiG 3/3

MTU configured under switchport settings

```
Musky-9506# sh interface gigabitethernet 3/3
GigabitEthernet3/3 is up
    Port description is FCIP
    Hardware is GigabitEthernet, address is 0005.3000.a408
    Internet address is 50.1.2.1/24
    MTU 2300  bytes
    Port mode is IPS
    Speed is 1 Gbps
    Beacon is turned off
    Auto-Negotiation is turned on
    5 minutes input rate 16 bits/sec, 2 bytes/sec, 0 frames/sec
    5 minutes output rate 16 bits/sec, 2 bytes/sec, 0 frames/sec
    3782808 packets input, 427481218 bytes
      0 multicast frames, 0 compressed
      0 input errors, 0 frame, 0 overrun 0 fifo
    5240056 packets output, 982552407 bytes, 0 underruns
      0 output errors, 0 collisions, 0 fifo
      0 carrier errors
```

From: show interface FCIP x

```
TCP Parameters
      Path MTU 2300 bytes
      Current retransmission timeout is 200 ms
      Round trip time: Smoothed 2 ms, Variance: 1
      Advertized window: Current: 101 KB, Maximum: 39 KB, Scale: 4
      Peer receive window: Current: 66 KB, Maximum: 116 KB, Scale: 4
      Congestion window: Current: 10 KB, Slow start threshold: 966 KB
    5 minutes input rate 1256 bits/sec, 157 bytes/sec, 0 frames/sec
    5 minutes output rate 304 bits/sec, 38 bytes/sec, 0 frames/sec
      7751 frames input, 14420052 bytes
        0 Class F frames input, 0 bytes
        7751 Class 2/3 frames input, 14420052 bytes
        0 Reass frames
        0 Error frames timestamp error 0
      8050 frames output, 15809456 bytes
        0 Class F frames output, 0 bytes
        8050 Class 2/3 frames output, 15809456 bytes
        0 Error frames
```

This is controlled via PMTU enabled, If Path MTU discovery times out, this will remain 1500

Check for reassembly of frames, should be zero if MTU set proper to fit FCP frame

Monitor FC Frame Timeouts

MDS9216-TOP# **show ips stats dma int gigabitethernet 2/4**

Dma-bridge ASIC Statistics for port GigabitEthernet2/4

Hardware Egress Counters

1030008889 Good, 0 bad protocol, 0 bad header cksum, 0 bad FC CRC

Hardware Ingress Counters

1843269599 Good, 0 protocol error, 0 header checksum error

0 FC CRC error, 0 iSCSI CRC error, 0 parity error

Software Egress Counters

1030008581 good frames, 0 bad header cksum, 0 bad FIFO SOP

0 parity error, 0 FC CRC error, 0 timestamp expired error

0 unregistered port index, 0 unknown internal type

0 RDL ok, 0 RDL drop (too big), 0 RDL ttl_1

1278210315 idle poll count, 116279 loopback

0 FCC PQ, 0 FCC EQ, 0 FCC generated

Flow Control: 0 [0], 0 [1], 0 [2], 0 [3]

Software Ingress Counters

1843269472 Good frames, 0 header cksum error, 0 FC CRC error

0 iSCSI CRC error, 0 descriptor SOP error, 0 parity error

141281 frames soft queued, 0 current Q, 732 max Q, 0 low memory

0 out of memory drop, 0 queue full drop

0 RDL ok, 0 RDL drop (too big)

Flow Control: 0 [0], 141281 [1], 0 [2], 0 [3]

Monitor for 500ms Timeout issues

Thank you.