FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Fibre Channel Data Center Interconnects:

64G FC and More

Live Webcast

July 18, 2024

10:00 AM PT/1:00 PM ET

Today's Speakers

David Rodgers
Technical Business Development Manager
EXFO

Gene Cannella
Principal Systems Architect
Adtran

Andy Adams
Product Line Management Director
Adtran

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

About the FCIA

"The Fibre Channel Industry Association (FCIA) is a mutual benefit, non-profit, international organization of manufacturers, system integrators, developers, vendors, industry professionals, and end users."

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

About the Fibre Channel Industry Association (FCIA)

25+ Years
Promoting Fibre Channel Technology

Industry Leading
Member Companies

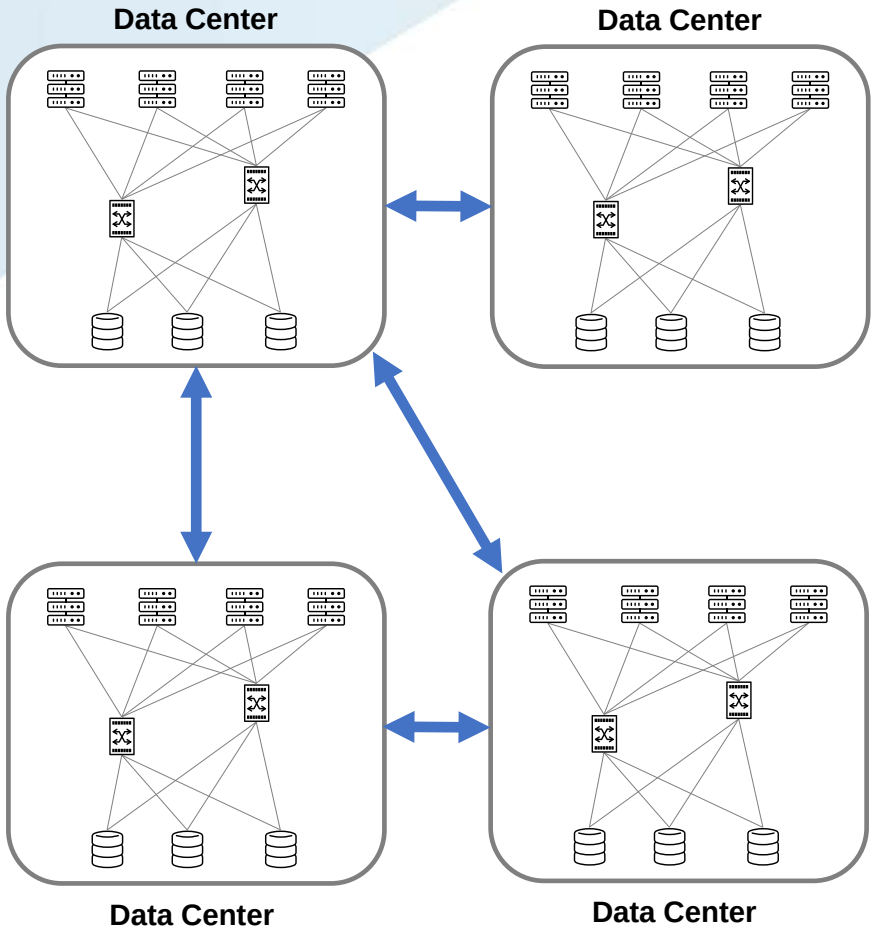
160M FC Ports
Shipped Since 2001

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Agenda

- Data Center Interconnect and Fibre Channel Extension
- Motivation *(Why do it?)*
- Concept *(What is it and how is it done?)*
- Types of DCI
- Fabric Extension
- Enabling Technologies
- DWDM *(Dense Wavelength Division Multiplexing – what is that?*
- OTN *(Optical Transport Network – what is that?)*
- Cryptography *(Layer 1 traffic security – How and Why)*
- Buffer Credits
- Protection
- Supervision

Data Center Interconnect and Fibre Channel Extension



Fibre Channel has been spectacularly successful in the modern data center.

Within the data center, Fibre Channel enables the creation of Storage Area Networks (SANs) that in turn facilitate the separate and independent scaling of compute and storage.

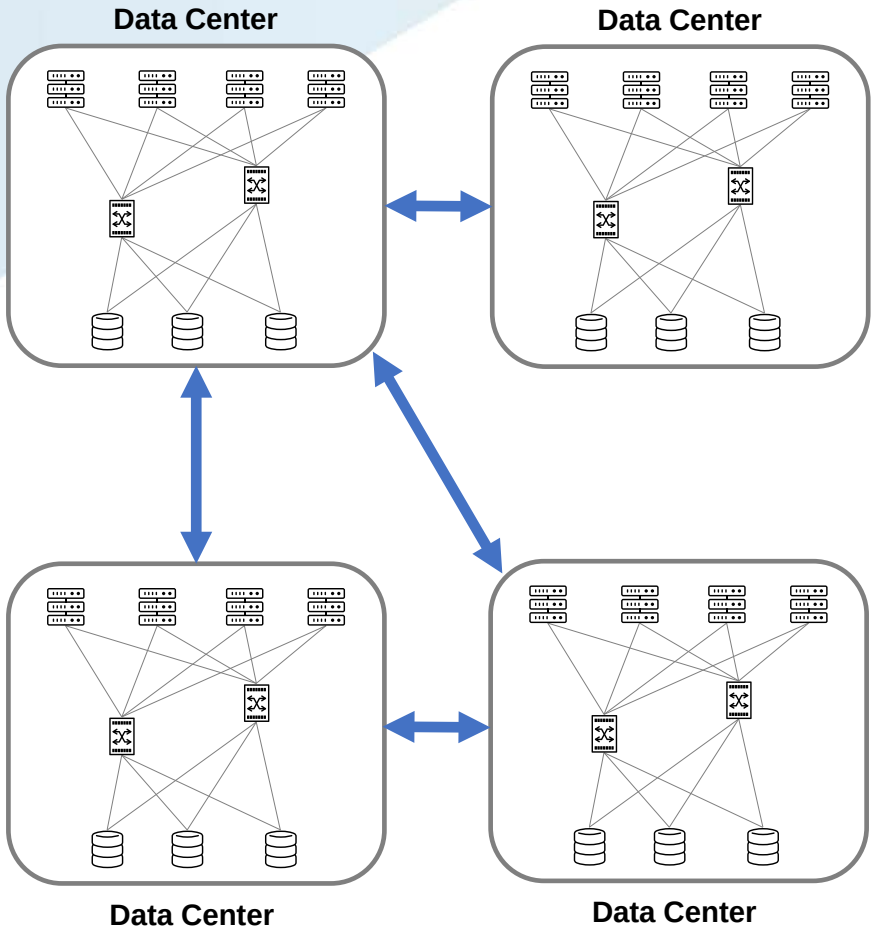
The data center environment integrates these SANs with Ethernet and IP-based Layer 2 and Layer 3 switched and routed networks to enable an unimaginable variety of applications.

The services provided and the data processed have become fantastically valuable and business-critical

Scalability, availability, security and survivability of the data center resources, applications and data are paramount. All these are enhanced by the networked interconnection of data centers.

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Motivating Reasons For Connecting Data Centers



- Business Continuity/Disaster Recovery (BCDR)
- Workload balancing and workload mobility
- Resource sharing and scaling
- Data exchange and data replication
- Data center migration

Business Continuity: *Risk management strategy for maintaining normal or near-normal operations during and after an unexpected incident or disaster.*

Disaster Recovery*: Risk management strategy for the restoration and resumption of limited to full operations after an incident.*

Redundancy and/or the ability to restore systems, processes, infrastructure, data and applications must be in place before any incident and fully integrated with normal operations.

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Types of Connectivity

Layer 3 Extension:

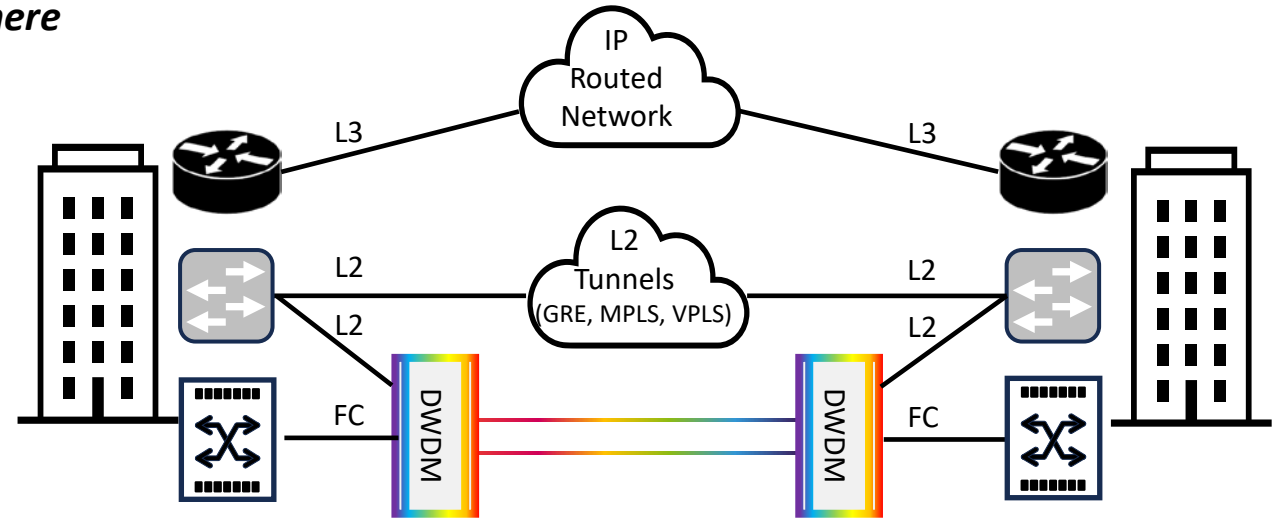
- IP routed connectivity between data centers.
- Used for segmentation/virtualization and file server backup applications.
- Asynchronous transactions

…FCIP exists here

Fibre Channel SAN Extension:

- Use native Fibre Channel methods
- Extend and scale the FC Fabric
- Synchronous transactions

…ISLs over OTN/DWDM



Layer 2 LAN Extension:

- Extend a single Layer 2 domain across data centers.
- Enables distribution of physical resources, facilitates mobility of the active machine
- Interconnect high availability clusters
- Server migration, application mobility.
- Can take DWDM path or L2 Tunnel

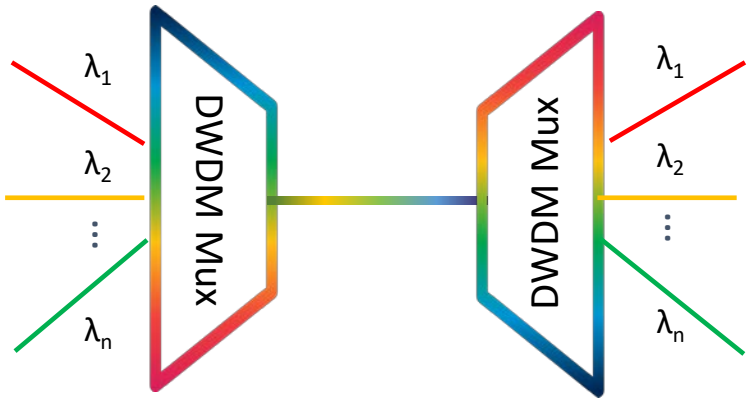
…FCoE exists here

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Fibre Channel Extension Via DWDM & OTN

Dense Wave Division Multiplexing

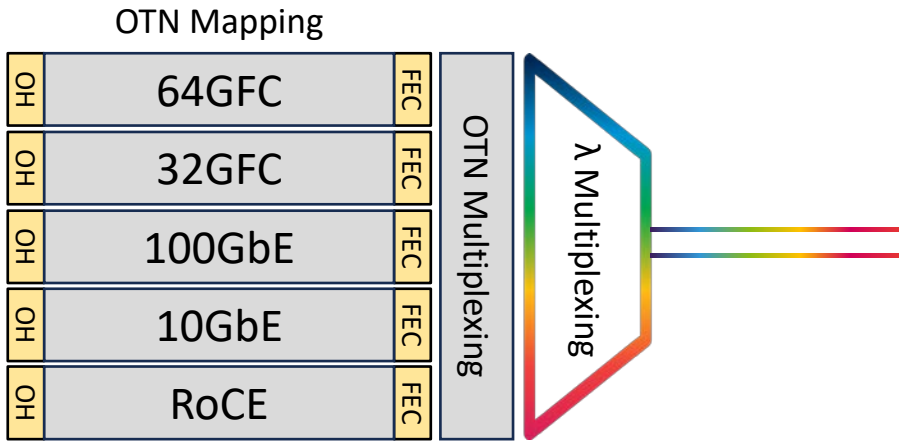
- DWDM gives us a very high bandwidth, very cost efficient solution for long distance interconnects



…Scales to over one hundred wavelengths

Optical Transport Network

- OTN provides a low-overhead, low-latency system for supervising all the wavelengths on DWDM as well as all the client services in a multi-protocol environment



…Scales from 2.5 Gbps to n x 100Gbps per wavelength

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

DWDM With OTN Used To Extend the Fabric

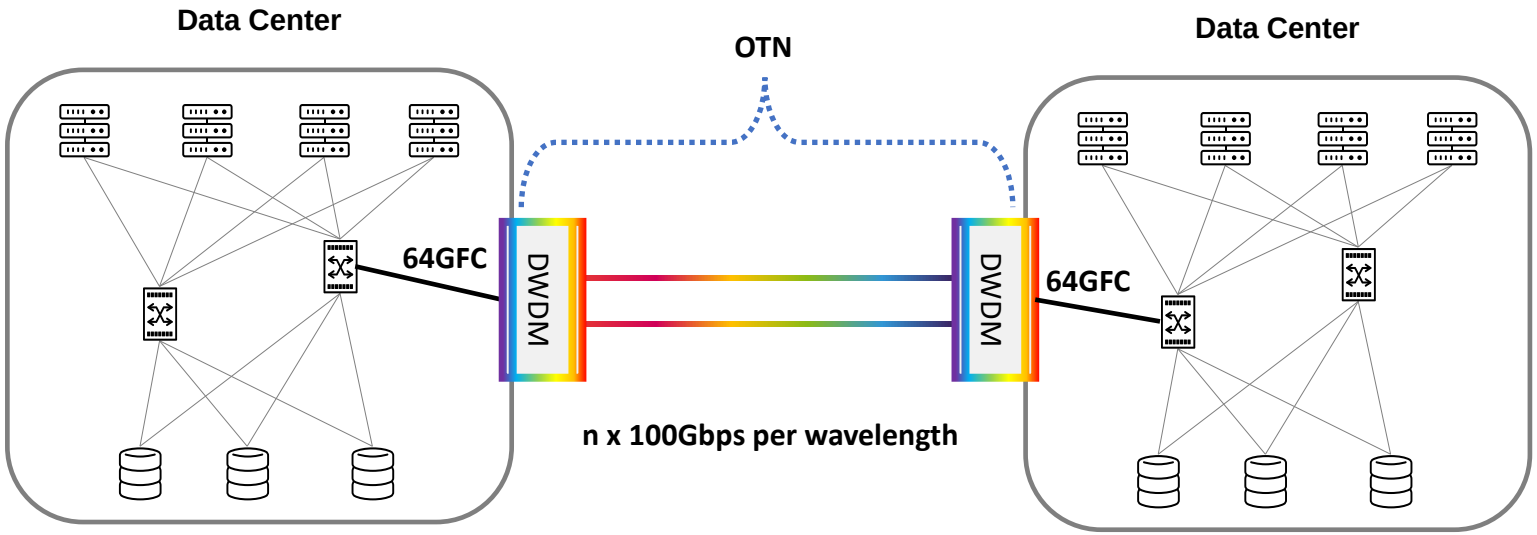
Typically, in fact nearly universally, used to transport ISLs (Inter-Switch Links)

- Effectively this “remotes” the associated E_Ports
- Generally not best practice to remote F_Ports or N-Ports. Inflexible and difficult to scale

This creates an logical architecture that extends the Fibre Channel SAN fabric across multiple data centers

Best practice to efficiently use the DWDM system is to transport ISLs at the highest rate supported by the fabric.

Today, that rate is becoming 64GFC. Of course, lower rates such as 32GFC and 16GFC are still supported on the DWDM system.



The DWDM system can support >100 wavelengths per fiber.

Each wavelength might provide 100, 200 or 400Gbps of payload bandwidth, with 800Gbps per wavelength now emerging

The DWDM system thus can provide ~ 10-80 Tbps per fiber

100Gbps OTN can carry 1x 64GFC ISL *(plus 1x 32GFC + 1x 16GFC ISLs)*
200Gbps OTN can carry 3x 64GFC ISLs *(plus 1x 32GFC + 1x 4GFC ISLs)*
400Gbps OTN can carry 7x 64GFC ISLs *(plus 1x 16GFC ISL)*
800Gbps OTN can carry 14x 64GFC ISLs *(plus 1x 32GFC ISL)*

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Latency Suitable for Synchronous Transactions

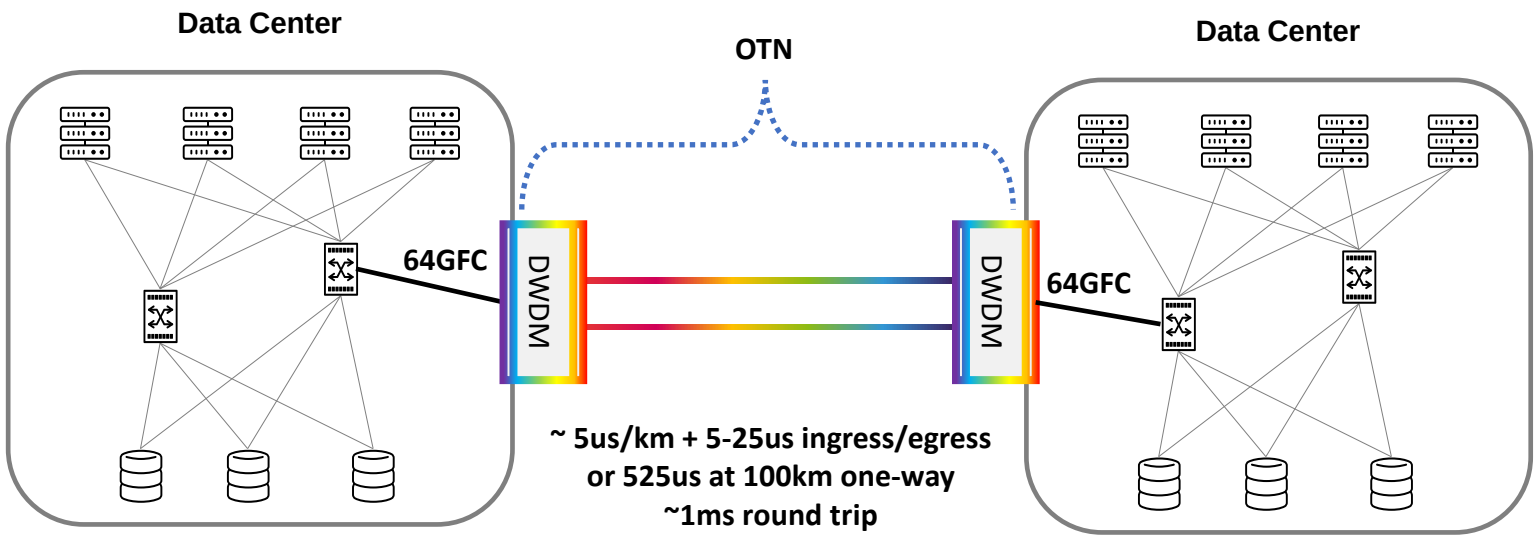
Synchronous transactions can be supported between multiple data centers at up 100-150km distances

…as long as the speed of light is the only significant contributor to the latency of the extended ISL

OTN and DWDM exhibit deterministic, minimal latency for native Fibre Channel traffic

- End-to-end OTN/DWDM latency typically 5-25 microseconds, plus the speed of light contribution
- This means at 100km the OTN/DWDM latency represents approximately 1-5 percent of the latency contributed by speed of light over that distance

…But we're going to need a lot of buffer credits!



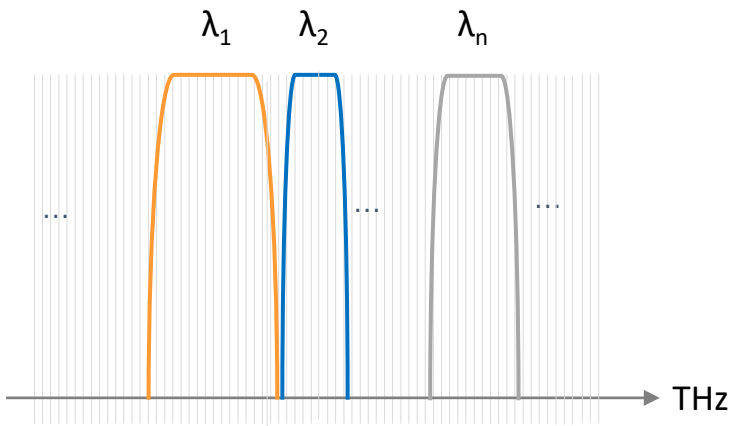
FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

DWDM (Dense Wavelength Division Multiplexing)

Dense Wavelength Division Multiplexing fits many optical signals onto a single optical fiber (pair).

Each optical signal is a modulated carrier occupying a channel defined by its center frequency and occupied bandwidth.

Channel spacings are standardized according to an ITU-defined grid. Center frequency spacings commonly range from 37.5GHz to 150GHz in increments of 6.25GHz.



Transmit spectrum

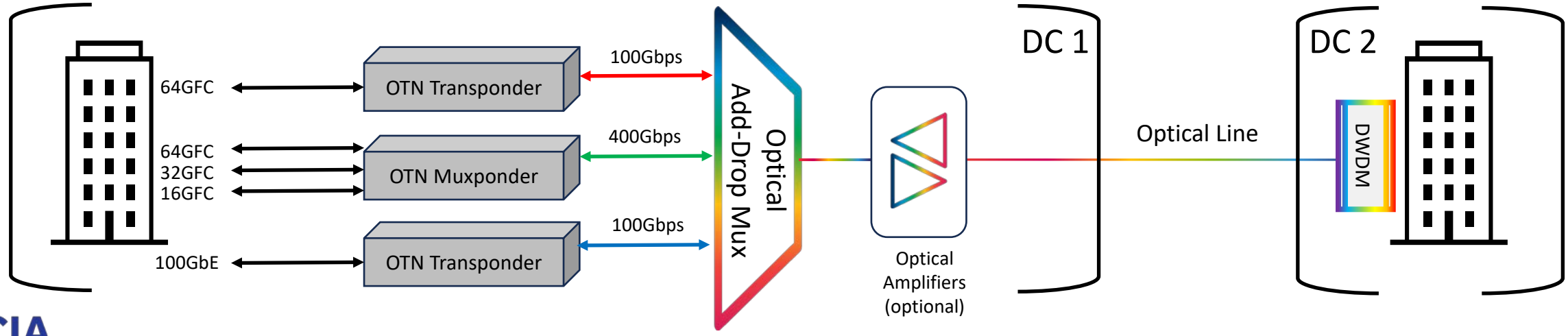
DWDM System Components

DWDM systems employ transponders and muxponders to:

1. Adapt the native client protocol (Fiber Channel, Ethernet, etc.) to the network transport protocol (OTN)
2. Retransmit the client optical signal using a tunable laser on a DWDM grid frequency, with higher power and longer reach.

Optical Add-Drop Multiplexers combine the various single frequency laser outputs onto one optical fiber, and separate the signals back out at the far end.

Optical amplifiers optionally may be used to increase transmission distances. This is generally not necessary at the distances common for data center interconnection.



OTN (The Optical Transport Network)

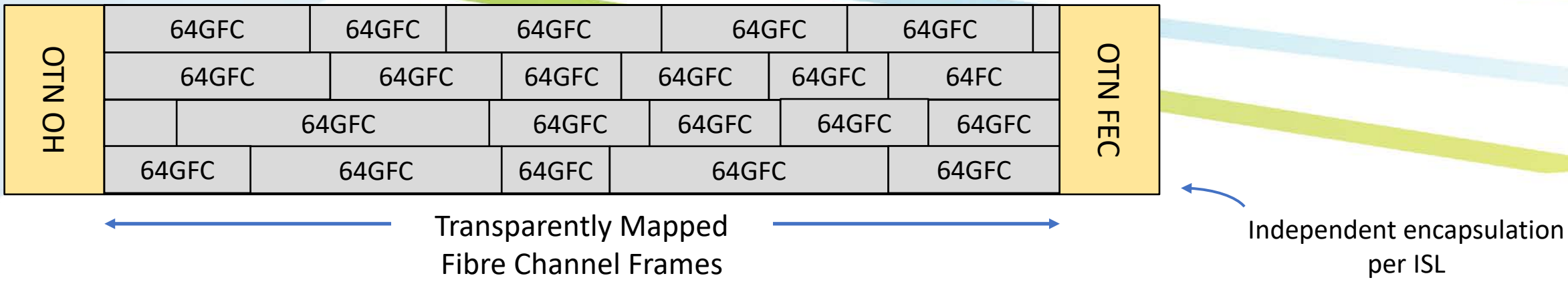
Deployment of DWDM networks for Data Center Interconnect is largely driven by the transport of Ethernet and Fibre Channel, but operators must also interconnect using other protocols such as InfiniBand or RoCE or many others.

…OTN provides a system for the generalized adaptation of diverse traffic types into the optical DWDM network.

With DWDM, users need to manage tens to hundreds of traffic-carrying wavelengths, and as many or more client services.

…OTN provides a protocol and infrastructure for the management of both the adapted individual client services and the wavelengths carrying them.

Optical Transport Network (OTN)



Transparent at the bit level

- All PCS characters are transmitted per FC-FS-x Fibre Channel - Framing and Signaling standard

Low Latency

- Mapping is deterministic. Does not experience routing delay or delivery time variation

Forward Error Correction

- FEC scheme ensures Fibre Channel data is preserved end to end. Very low bit error rates

Operations, Administration, Maintenance, and Provisioning

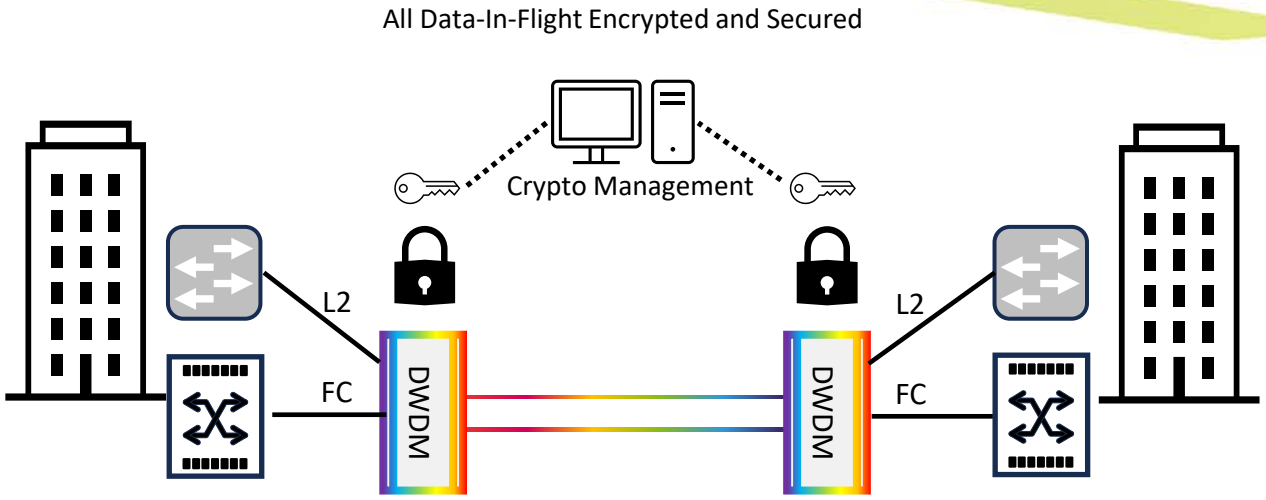
- OTN Overhead supports OAM&P with byte channels for fault detection, performance monitoring, protection, service tracing, etc.

FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

Cryptography - Layer 1 Security

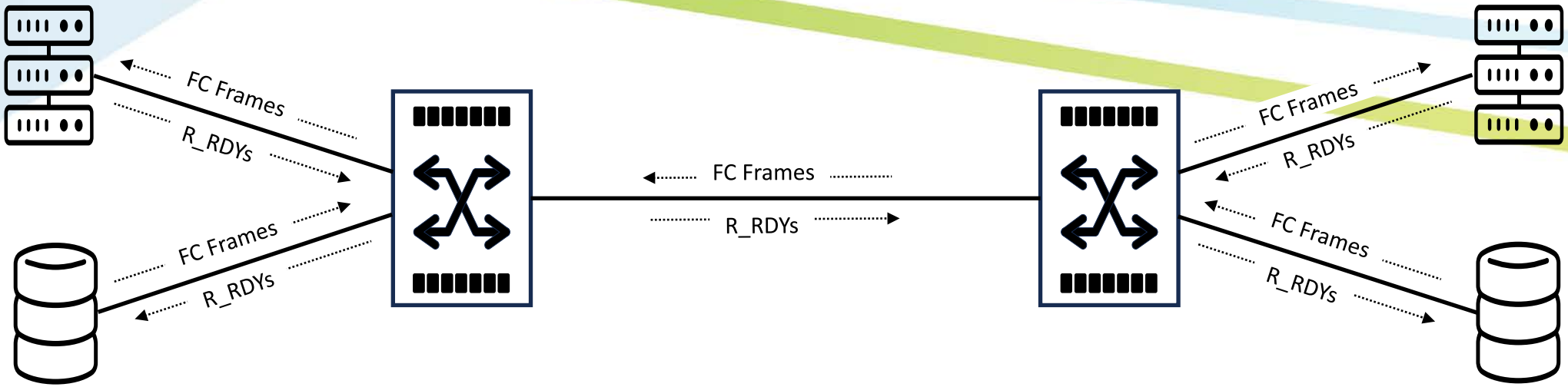
- Often simply a requirement of the regulatory environment or customers of data center services
  - Industry and governmental certification and approvals
- Part of a layered security strategy – Defense in Depth
  - Encrypts and secures Data-In-Motion
  - Complementary to Data-At-Rest encryption strategies
- Layer 1 encryption is agnostic to higher layer protocols
  - Protects data at all layers, even those with their own encryption and security methods
  - Protects metadata that MACsec and IPsec leave exposed – (MAC & IP addresses, protocol types, etc.)

- Strictly Separated Encryption Domain Management
- Quantum-safe options



- Leverages AES methods and integrates with Public Key Infrastructure

DCI Requirements for Buffer-to-Buffer Credits



Fibre Channel achieves link based flow control using the notion of Buffer-to-Buffer credits (BB_Credits)

In this scheme, a device may only transmit a frame if it knows that frame will not overrun the resources of the device at the other end of the link

BB_Credits let a transmitting device know how many frames it may transmit before that happens.

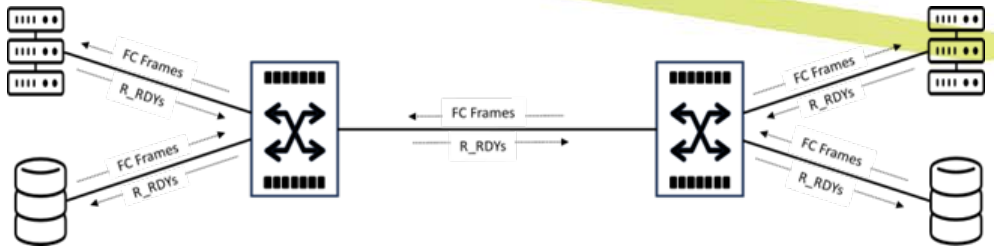
…How does that work?

DCI Requirements for Buffer-to-Buffer Credits

1

Devices are configured with a static number of buffer credits. Each credit is an indication that the device has the memory resources to accept and process a single arriving frame.

- For example, if a device is provisioned with 10 BB_Credits, this means that it can accept up to 10 arriving frames before it finishes processing the first one.



2

Devices on a link inform each other of the numbers of credits with which each has been configured

- E_Ports do this in the Exchange Link Parameters (ELP)
- F_Ports do this in the Fabric Login (FLOGI)

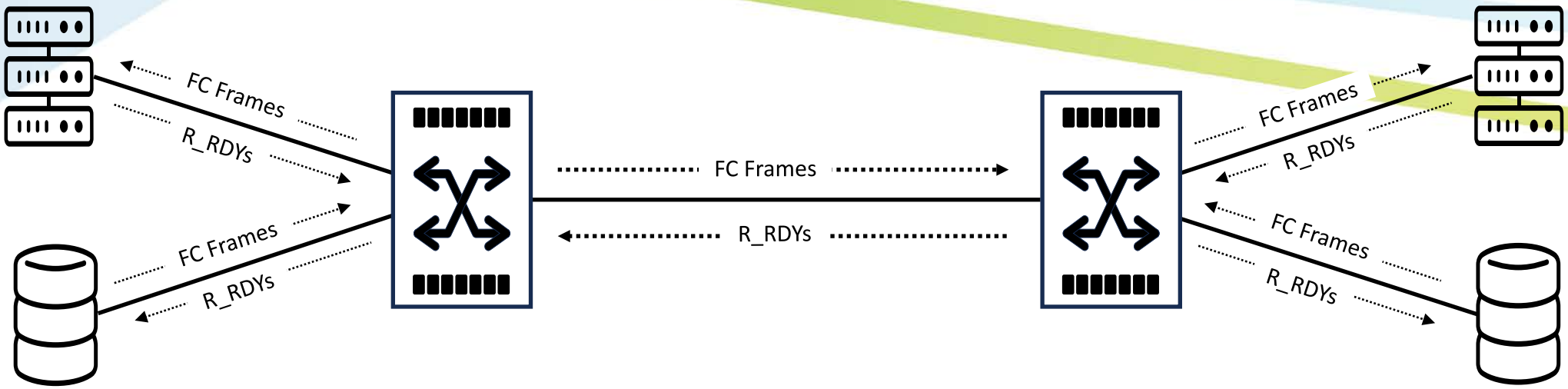
3

Each device on a link keeps a counter that ticks up when it transmits a frame, and ticks back down when that frame is acknowledged.

The far end device acknowledges that it has received ***and*** processed a frame by returning an R_RDY primitive

If more frames are transmitted than R_RDYs are received, the transmit counter will continue to tick upwards until the BB_Credits limit is reached, at which point the transmitter must stop.

DCI Requirements for Buffer-to-Buffer Credits

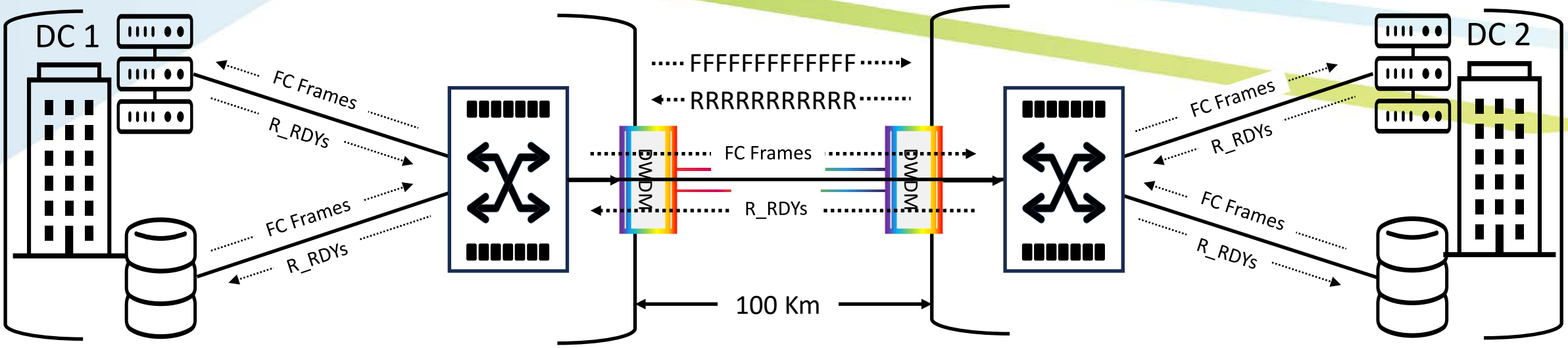


Every frame transmitted should be acknowledged with a Receiver_Ready (R_RDY) primitive.

Frames do not have to wait for an R_RDY if the transmit buffer counter has not reached the BB_Credits limit.

…For short distance links, this scheme is fundamentally about accommodating the instantaneous differences between the receiver's frame processing times and the transmitter's frame emission rate. This factor generally will dominate the buffer credit allocation as long as the link stays inside the building.

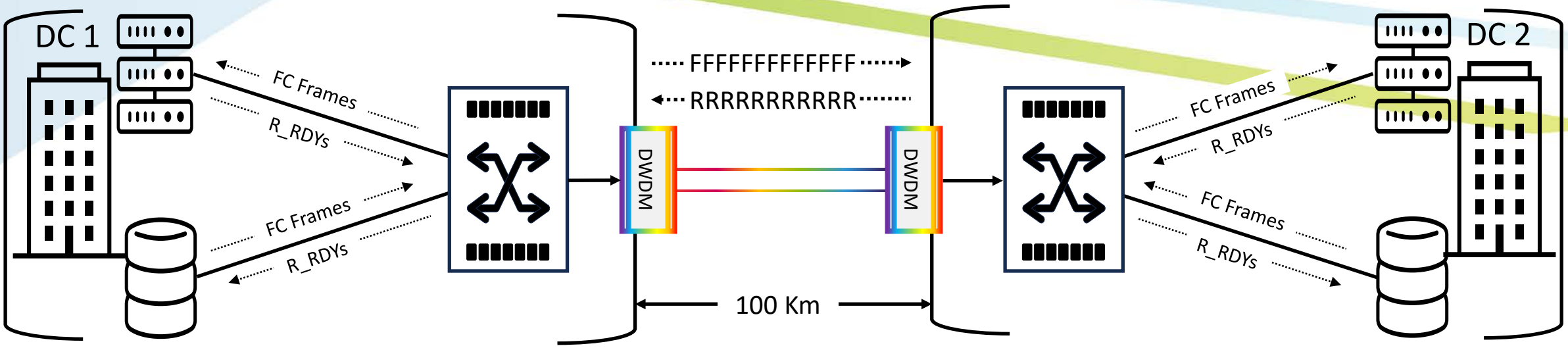
DCI Requirements for Buffer-to-Buffer Credits



When the SAN is extended to a remote data center, the extended ISL will now have many, many frames on the fiber at any time

There must be enough BB_Credits for as many frames as can be transmitted in the time it takes light to make the round trip on the fiber.

DCI Requirements for Buffer-to-Buffer Credits



When the SAN is extended to a remote data center, the extended ISL will now have many, many frames on the fiber at any time

There must be enough BB_Credits for as many frames as can be transmitted in the time it takes light to make the ***round trip*** on the fiber.

DCI Requirements for Buffer-to-Buffer Credits

A 64GFC Example

Givens (approx):

- Speed of Light (c) = $3x10^8$ m/s
- Optical fiber Index of Refraction (N) = 1.5
- Speed of light in the fiber = c x N = $2x10^8$ m/s
- 64GFC bit rate = 57.8 Gbps
- Assume 2000 Bytes/Frame
  - This is the most important unknown!
- Link Distance = 100 Km

Calculations:

- Transmission time (t) for a 2000 Byte Frame
  = (8 bits per Byte) x 2000 Bytes / 57.8 Gbps
  = 277 ns
- Length (L) of a 2000 Byte frame on the fiber
  = c x N x t
  = $2x10^8$ m/s x 277 ns = 55 meters
- Buffer Credits needed
  = 2 x Link Distance / L
  = 2 x 100 Km / 55 m = **3636 BB_Credits**

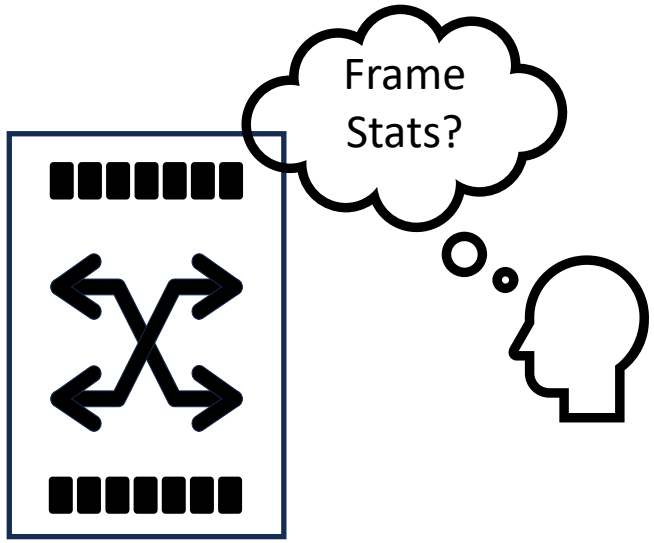
FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

DCI Requirements for Buffer-to-Buffer Credits

Some guidelines:

- BB_Credits needed are proportional to rate
  - *64GFC requires twice as many as 32GFC*
- BB_Credits needed are proportional to distance
  - *100 Km requires twice as many as 50 Km*
- BB_Credits needed are inversely proportional to average frame length
  - *2000 Byte frames require half as many as 1000 Byte frames*

- You can't really know *a priori* what the average frame length will be over a relevant period of time
- Your switch user interface will give you statistics for recent time periods that will approximately but usefully predict future performance



DCI Requirements for Buffer-to-Buffer Credits

ISL Link Utilization

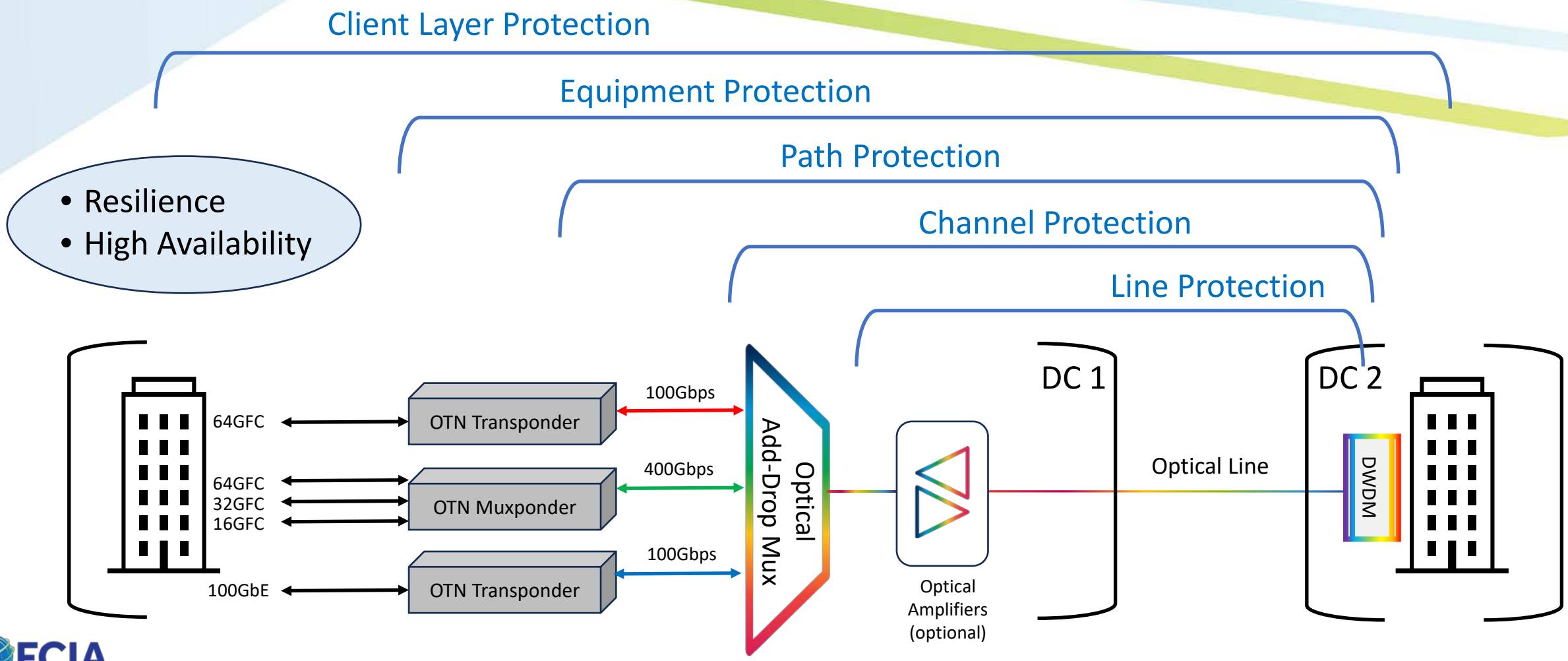
If you never run out of BB_Credits, your link should be fully utilized.

But the transient behavior of the link is not always predictable. It may not be possible to provision as many BB_Credits as it would take to be certain the link never stalls.

Perhaps 99.9% link utilization is OK. Or 99%?

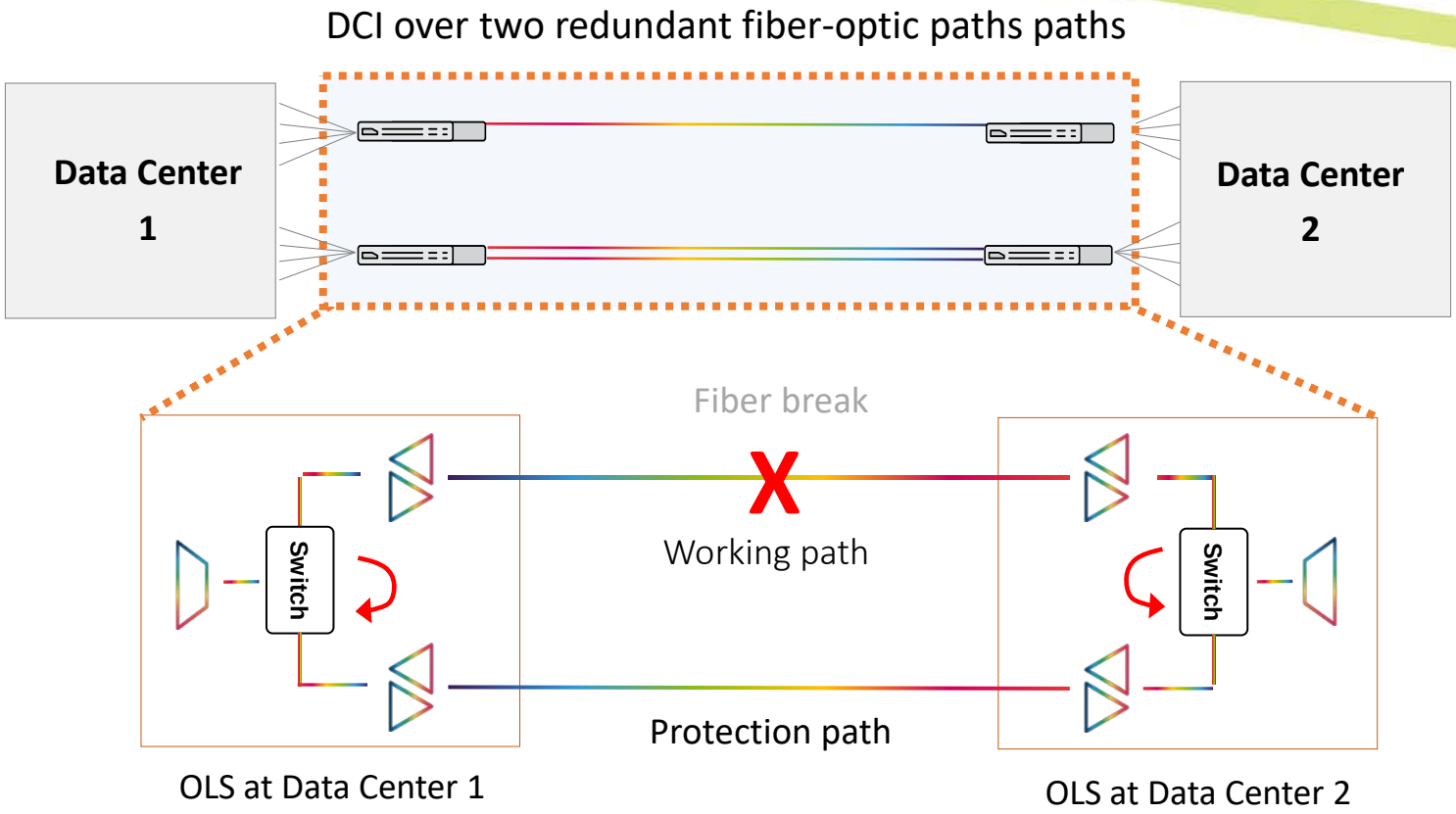
The occasional stall will not result in frame drops – that's the purpose of the buffer-to-buffer flow control scheme. No retransmissions necessary.

Protection Strategies



FCIA
FIBRE CHANNEL INDUSTRY ASSOCIATION

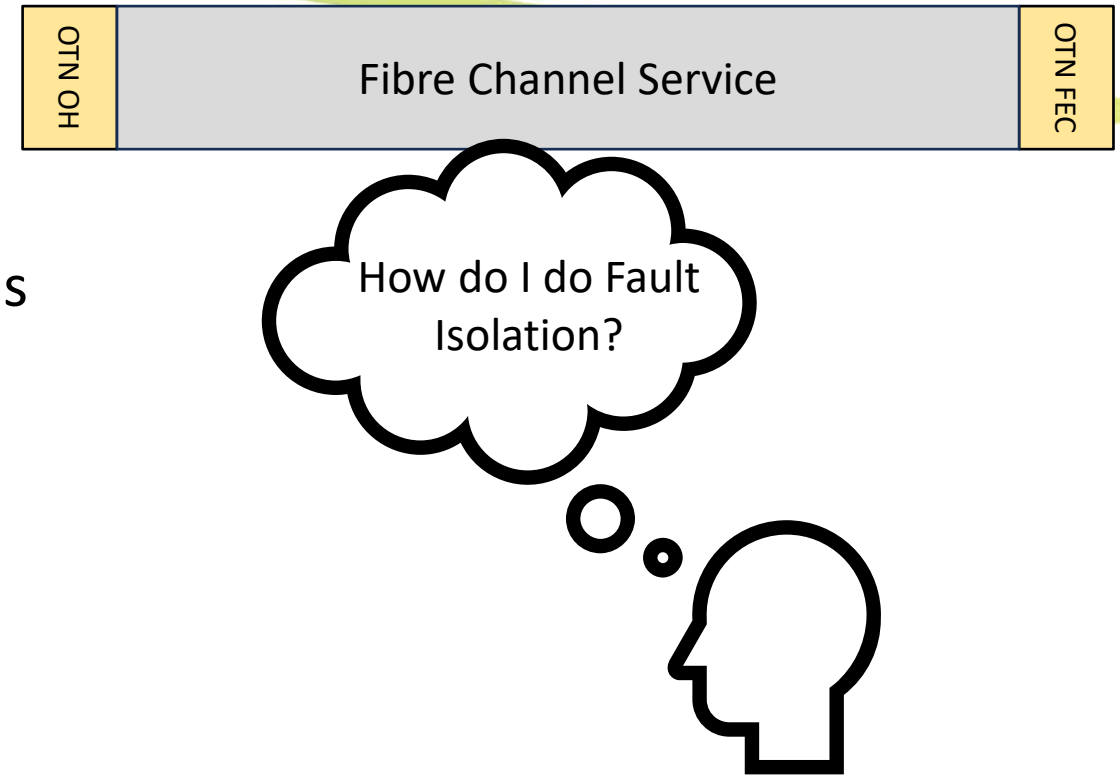
Line Protection Example



Supervision

- Performance Monitoring
- Fault Monitoring
- Consequent Actions
- Error Forwarding and Maintenance Signals (e.g., NOS, LOS)

- Non-Intrusive FC layer monitoring

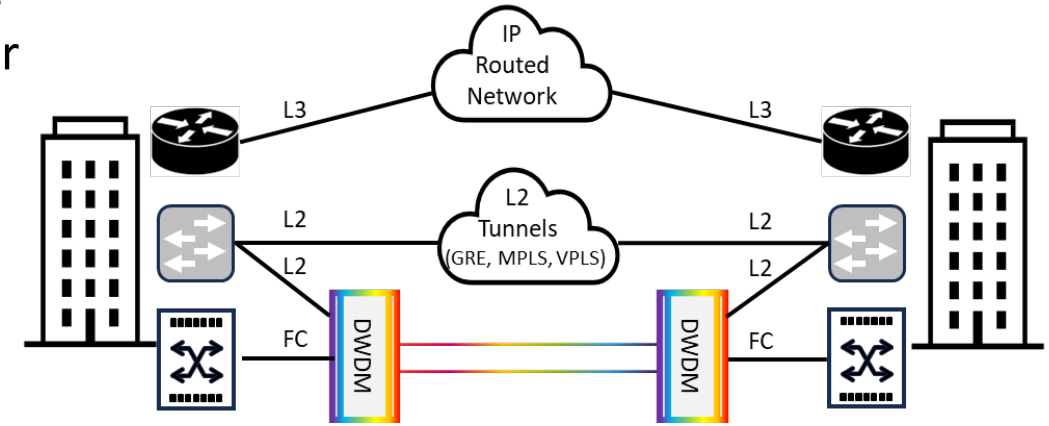


Summary

Data Center Interconnect is now table stakes, and mission-critical.

L2, L3 and SAN networks and applications all benefit from extension

- DWDM can extend the SAN fabric as far as 100-150 Km while supporting synchronous replication
  - Uses hundreds or thousands of buffer credits: approximately 40 BB_Credits per kilometer for 64GFC ISLs
  - 64GFC is the latest and highest rate to be supported on DWDM
- FCIP can be used to tunnel asynchronous transactions through the routed IP network



Questions

After this Webcast

- Please rate this event – we value your feedback
- We will post a Q&A blog at http://fibrechannel.org/ with answers to today's questions
- Our library of FCIA on-demand webcasts http://fibrechannel.org/webcasts/ includes:
  - Fibre Channel Fundamentals
  - FC-NVMe
  - FC-NVMe for ML
  - Long Distance Fibre Channel
  - Fibre Channel Speedmap
  - FCIP (Extension): Data Protection and Business Continuity
  - Fibre Channel Performance
  - FICON
  - Fibre Channel Cabling
  - 64GFC and 128GFC
  - FC Zoning Basics
- Follow us @FCIAnews for updates on future FCIA webcasts

Thank You