Long Distance Fibre Channel

Live Webcast
October 10, 2017
10:00 am PT

Today's Presenters

J Metz
FCIA Board of Directors, Cisco

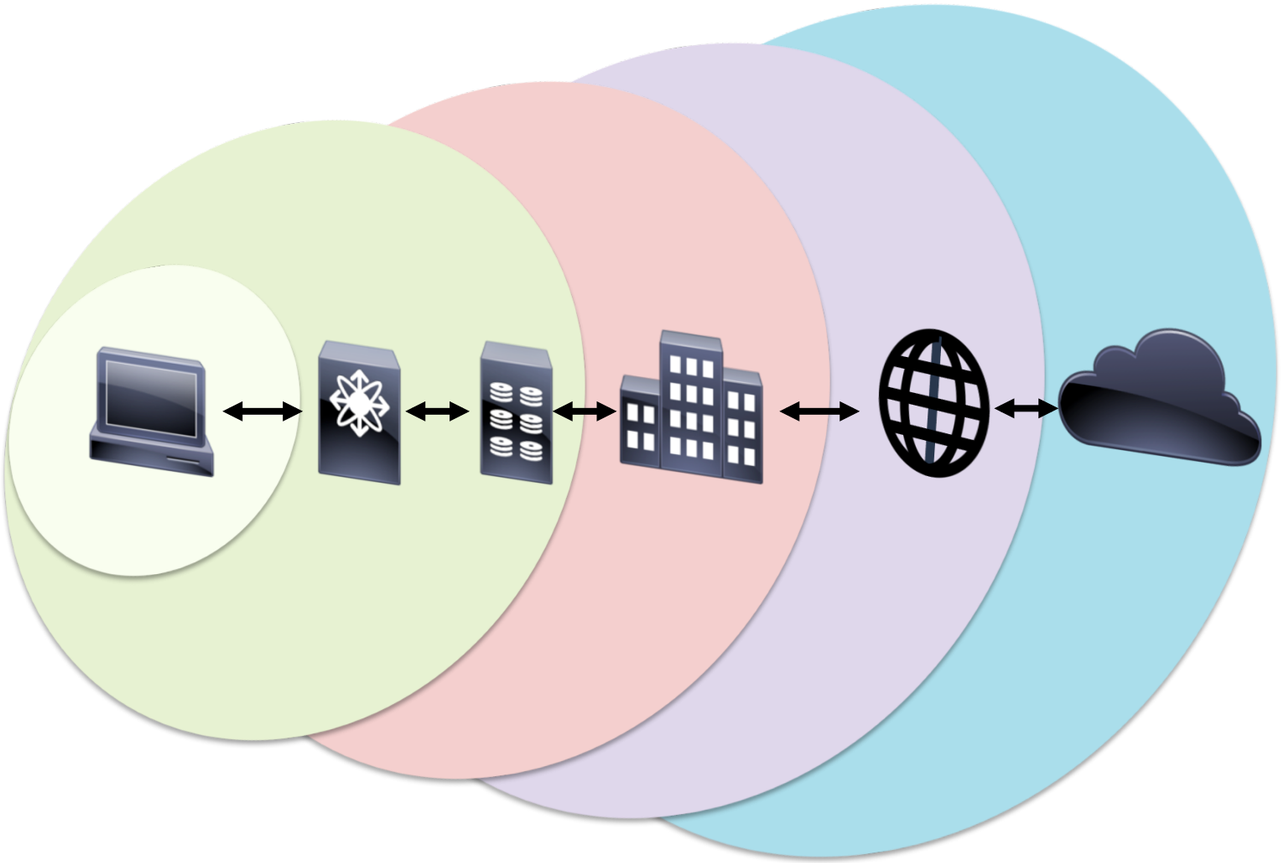
Mark Allen
Sr. Manager - Storage Networking
Technical Marketing, Cisco

Agenda

- Introduction
- Understanding the Physical Layer Requirements
- Understanding the Fibre Channel Requirements
- Understanding the Upper Layer Protocol Requirements
- Other Long Distance Options

Storage Big Picture

- Tendency is to look at storage from connectivity
- Another paradigm is to look at storage function



Intra-Host (PCIe/NVMe)

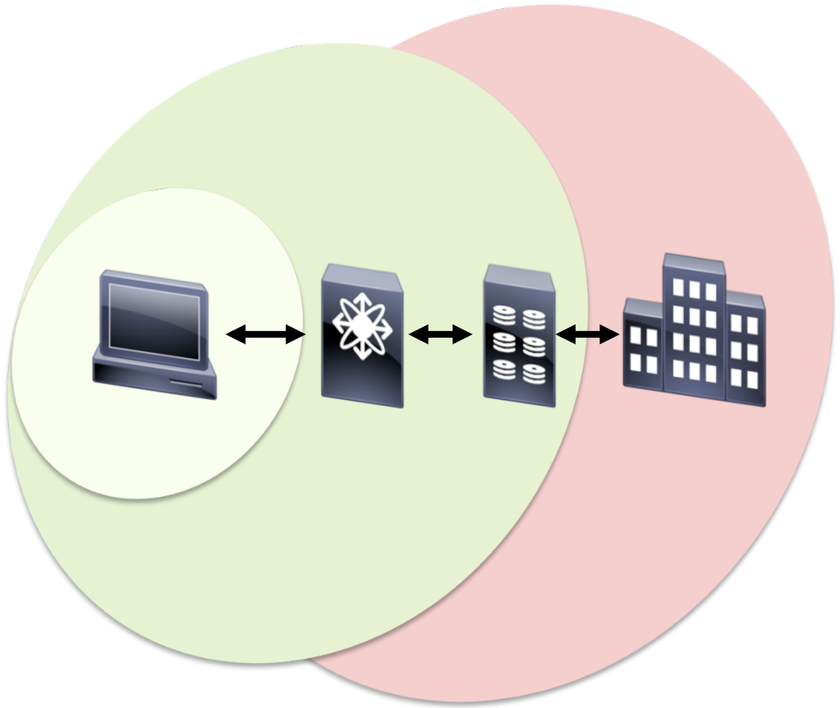
PCIe

- Co-Located
- Device limitations
- Remote storage access
  - Extending the bus architecture poses risks
  - Developing bus technology into a fabric poses risks
- Distance is a non-starter

Intra/Inter-Data Center

Fibre Channel, FCoE, InfiniBand & iSCSI

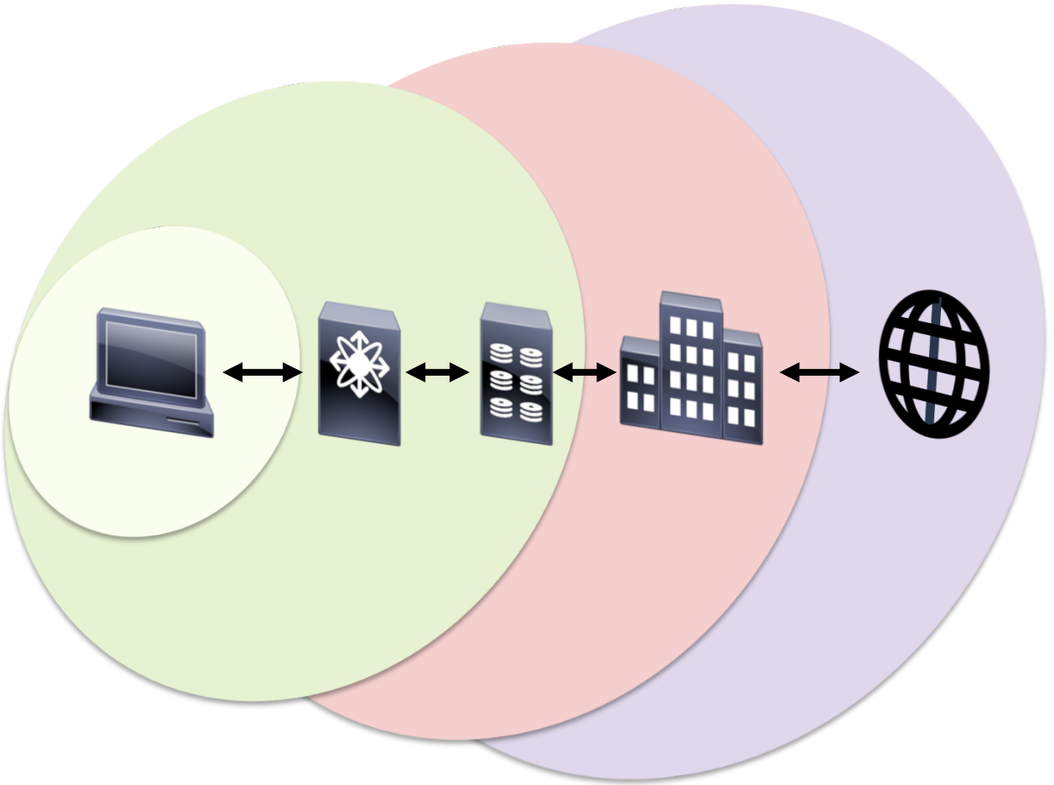
- Not Co-Located
- Primary usage: Intra/Inter-Data Center
- Distance considerations
  - Extending the architecture poses risks
  - Host-to-storage is not a good idea
- Global distance is not impossible for FC/iSCSI, but requires extra considerations



File Storage Scope

File Storage

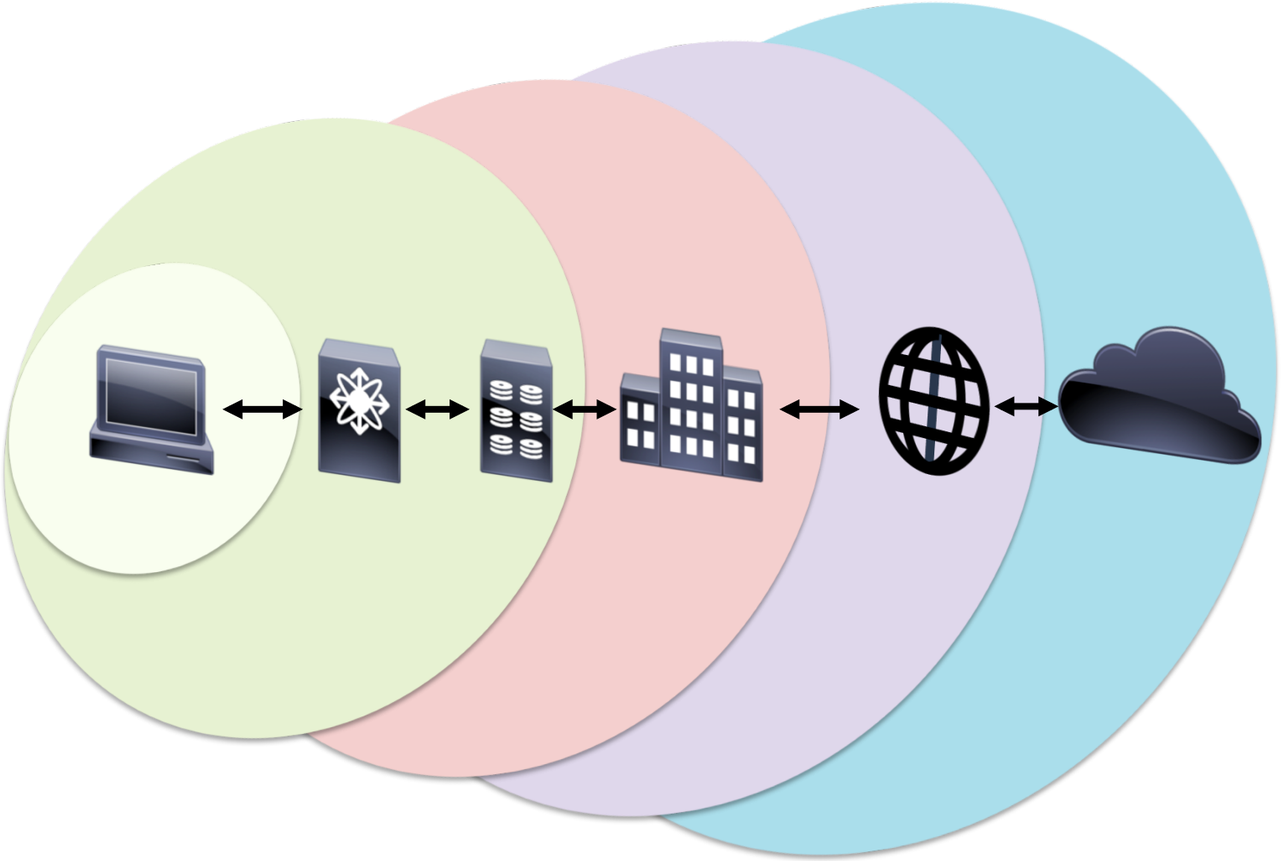
- Less rigid architecture, less performant than Block
- Inside and Outside Data Center
- Designed for sharing data among clients at scale
- Distance can be for normal operations, Disaster Recovery, and Backup



Object Storage Scope

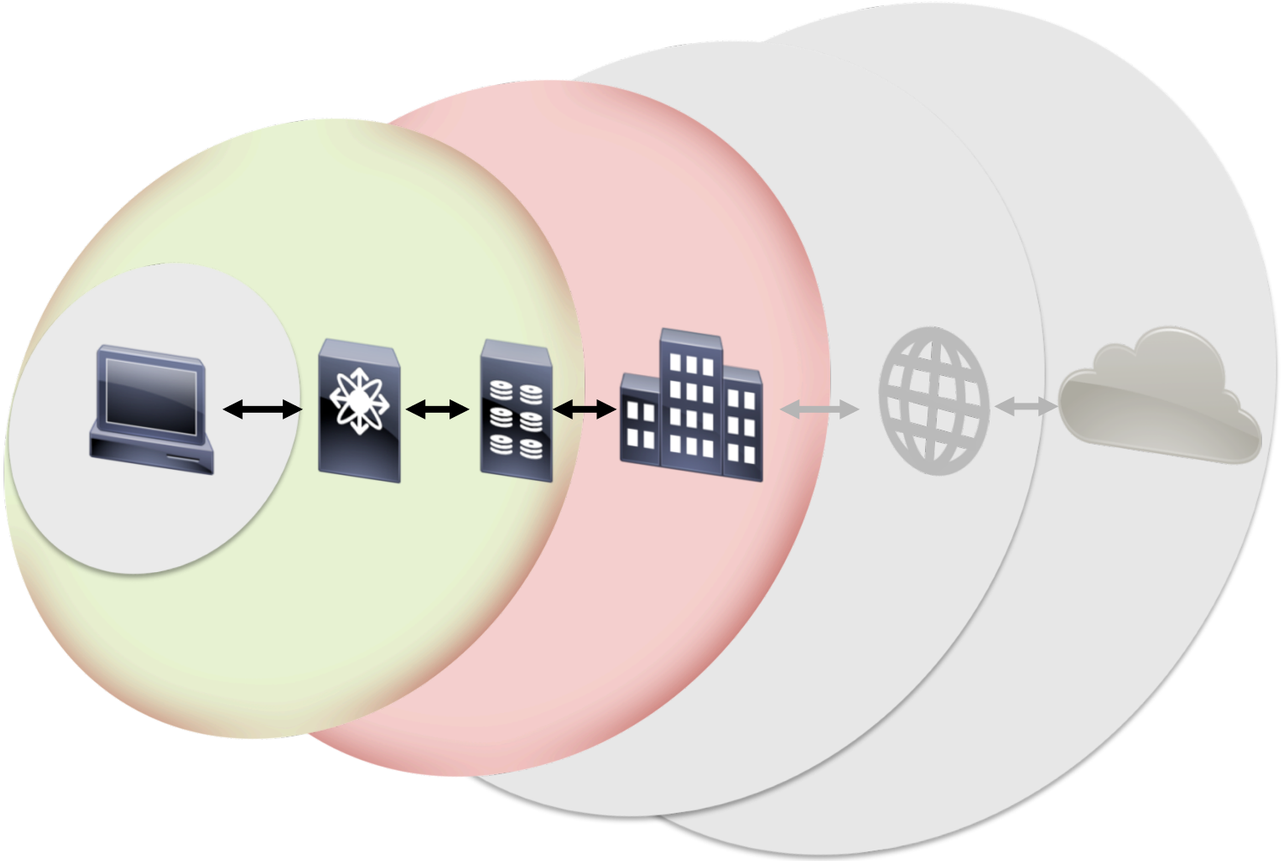
Object Storage

- Least performant
  - For data that doesn't change much, if at all
- Designed for scale and distance - access from anywhere



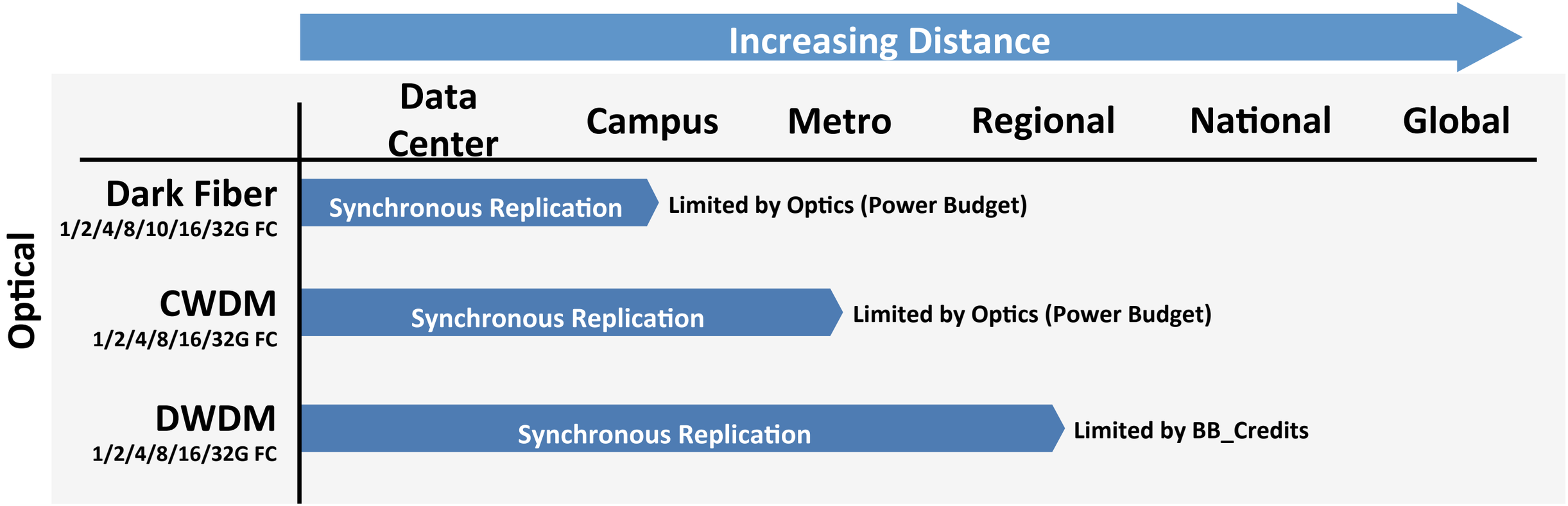
Long Distance Fibre Channel

- **Intra-DC and Inter-DC solutions**
- **Business and regulatory requirements drive the need to have data in multiple data centers**
  - Active–Active Data Centers
  - Active–Standby Data Centers
  - Long Term Data Backup and Warehousing



Physical Layer Requirements

Long Distance FC Technology Options



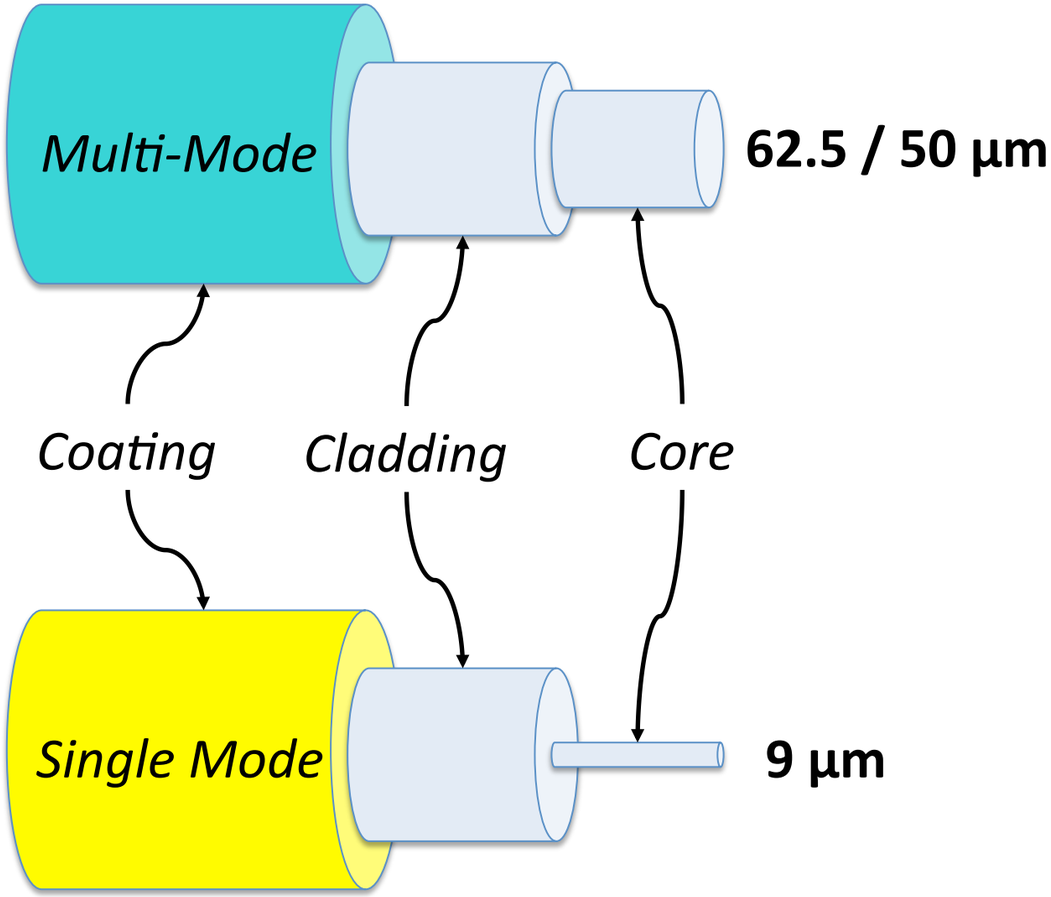
Dark Fiber

- **Simplest form of Intra or Inter-Data Center connectivity**
  - Point-to-point connections
  - Use of either Multi-Mode (Intra) or Single Mode (Intra/Inter) fiber
- **No amplification of signal**
  - Distances limited by optics power
- **No multiplexing of signal**
  - One connection per cable

Multi-Mode vs. Single Mode Fiber

Primary difference – core size

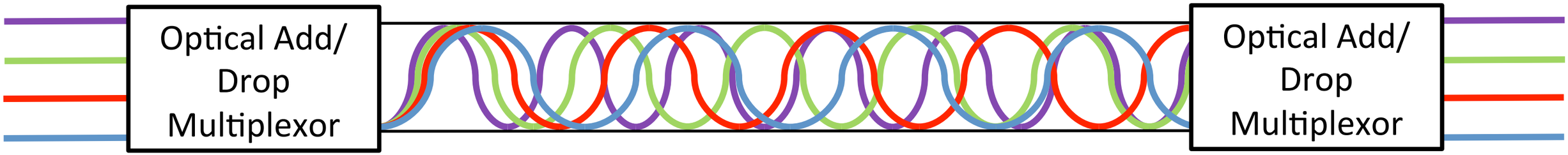
- Larger core of Multi-Mode less expensive with larger optical dispersion
  - Works with LED and laser transmitters
  - OM-3 and OM-4 required for higher speeds
- Smaller core of Single Mode reduces optical dispersion – allows greater distances



xWDM

Wave Division-Multiplexing technologies are used to transmit multiple wavelengths over the same dark fiber

- Multiple connections per fiber pair
- Spacing of wavelengths differs by solution
  - CWDM – 20nm wavelength spacing
  - DWDM – 0.4 / 0.8nm wavelength spacing
- Uses Optical Add/Drop Multiplexors (OADM) to combine wavelengths

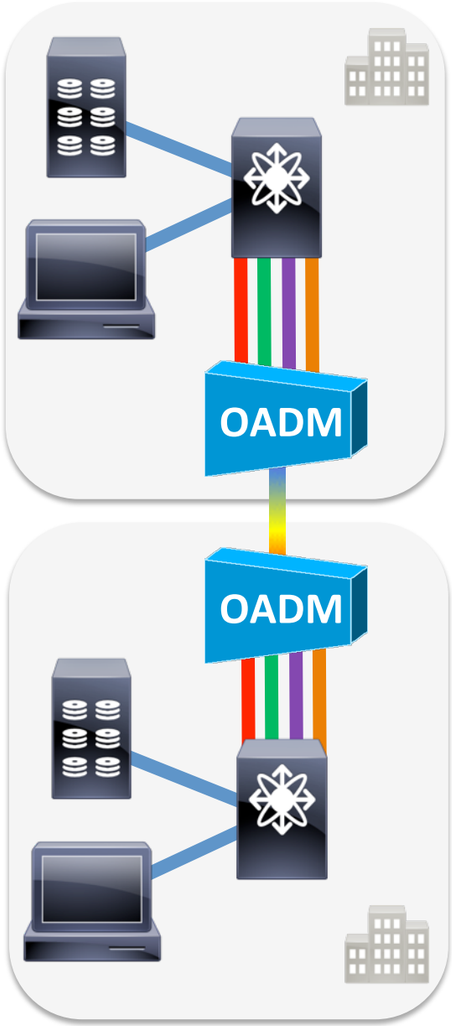


Coarse Wave Division Multiplexing (CWDM)

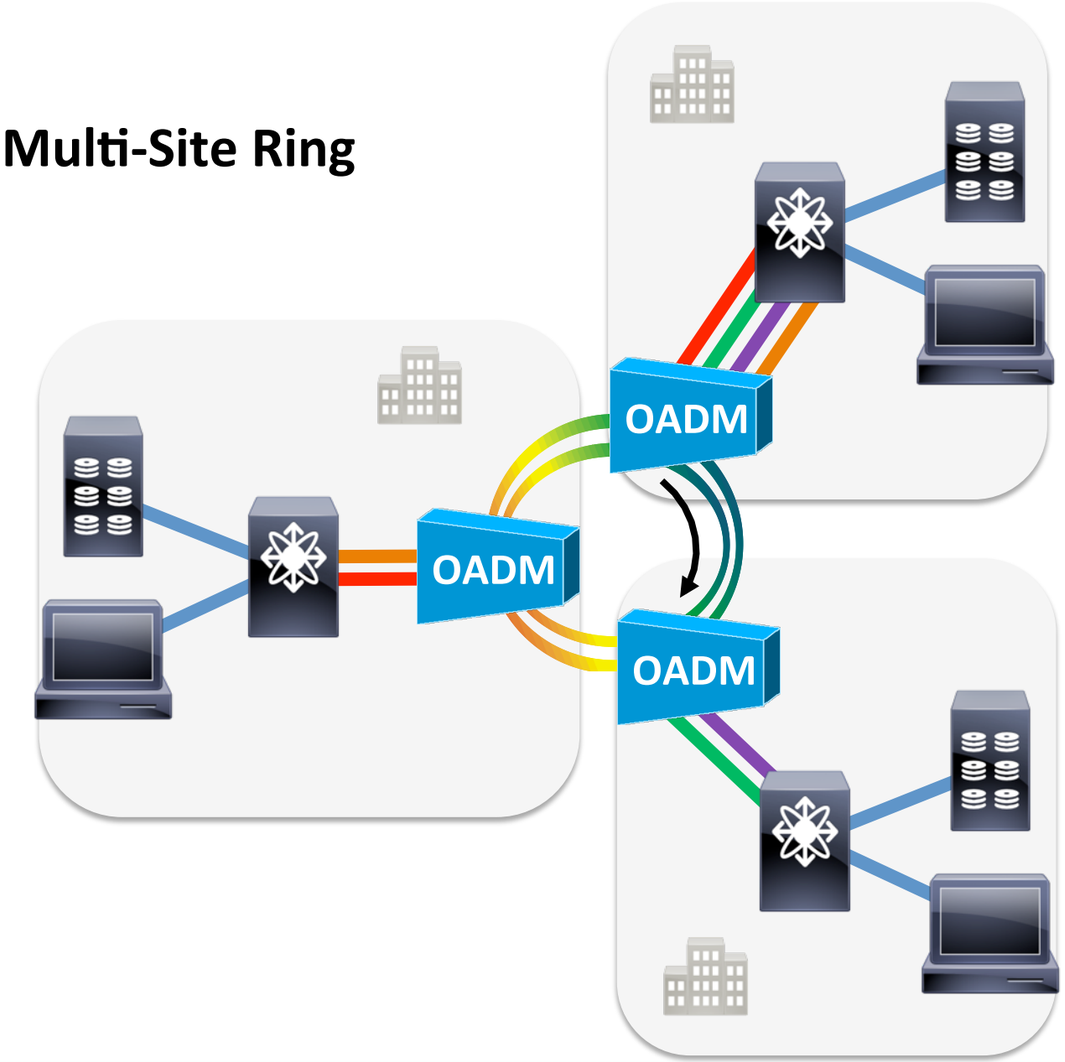
- Lower cost xWDM technology
- 18 available wavelengths
  - Spectrum of 1270nm to 1610nm with 20nm wavelength spacing
  - 1470 - 1610nm typical range
- Optical multiplexing done with passive CWDM OADM
- Optical power budget of optics primary driver of distance
- Distance also varies by topology and speed
  - Ring topology < Point-to-Point topology
  - Higher speed < Lower speed

CWDM Topologies

Point-to-Point



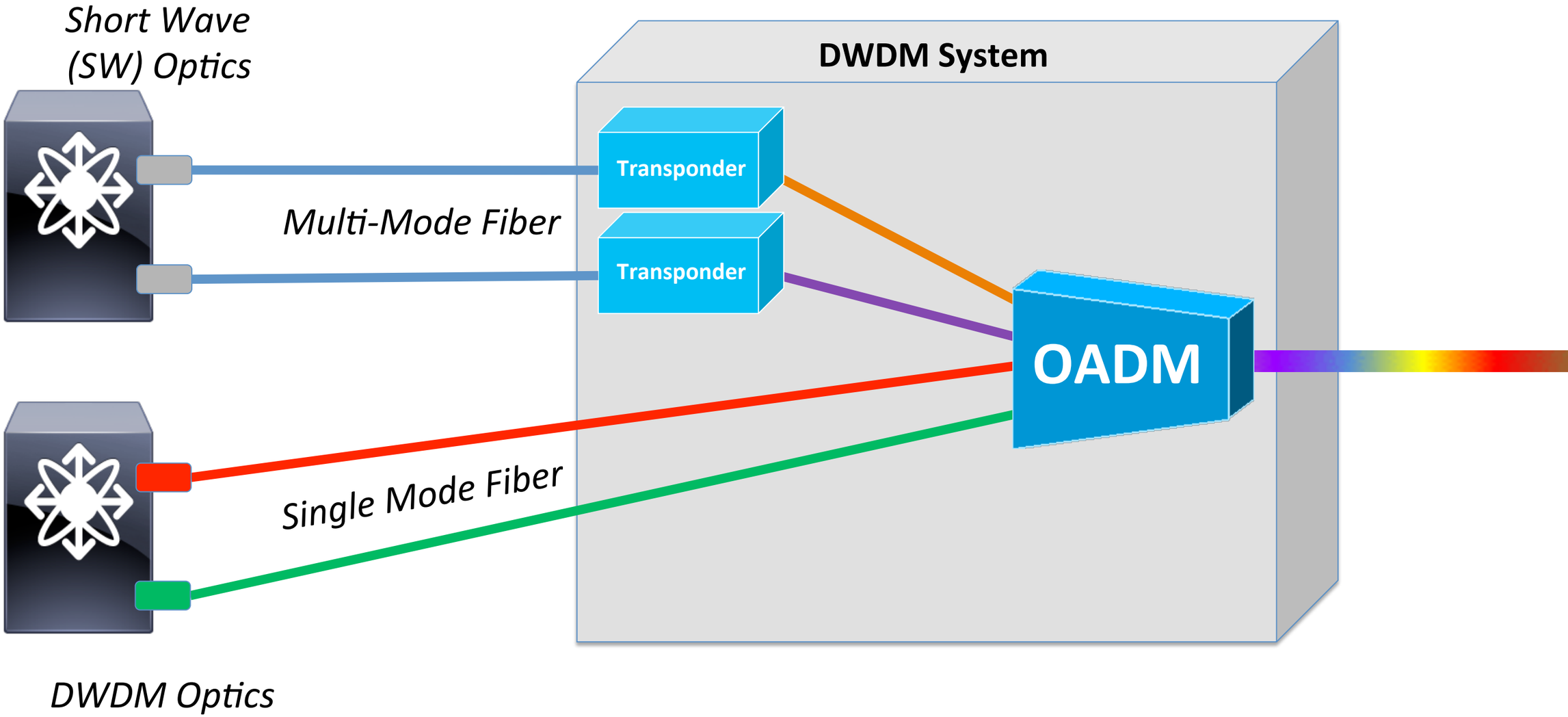
Multi-Site Ring



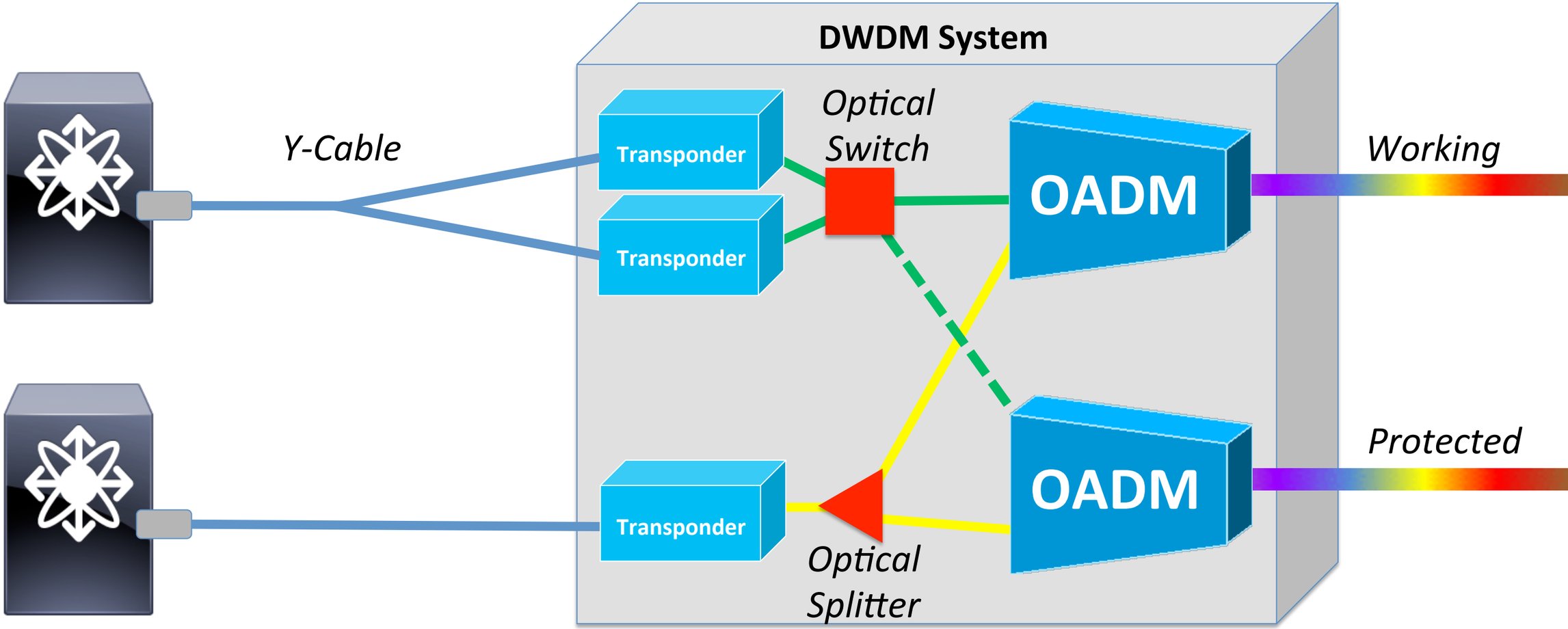
Dense Wave Division Multiplexing (DWDM)

- Higher density xWDM technology compared with CWDM
- DWDM primarily uses ITU C band and L band
  - C band – 1529–1561 nm
  - L band – 1570–1603 nm
  - Channel spacing ~.8nm @ 100 Ghz frequency
  - DWDM wavelengths mapped to ITU Channels
    - IE. ITU Channel 50 = 195000 Ghz = 1537.40 nm
  - Potential to carry 192 wavelengths per fiber pair
- Erbium-Doped Fiber Amplifiers (EDFA) allow for longer distances than CWDM
  - Use every 80-100km to amplify signal
- Multiple Protection options: Optical Splitter or Y-cable/line card

DWDM Connection Options



DWDM Protection Options



Physical Layer Summary

- **Three Primary Optical Solutions**
  - Dark Fiber
  - CWDM
  - DWDM
- **Increased Distance = Increased Cost/Complexity**
- **Physical Layer protection available with DWDM**

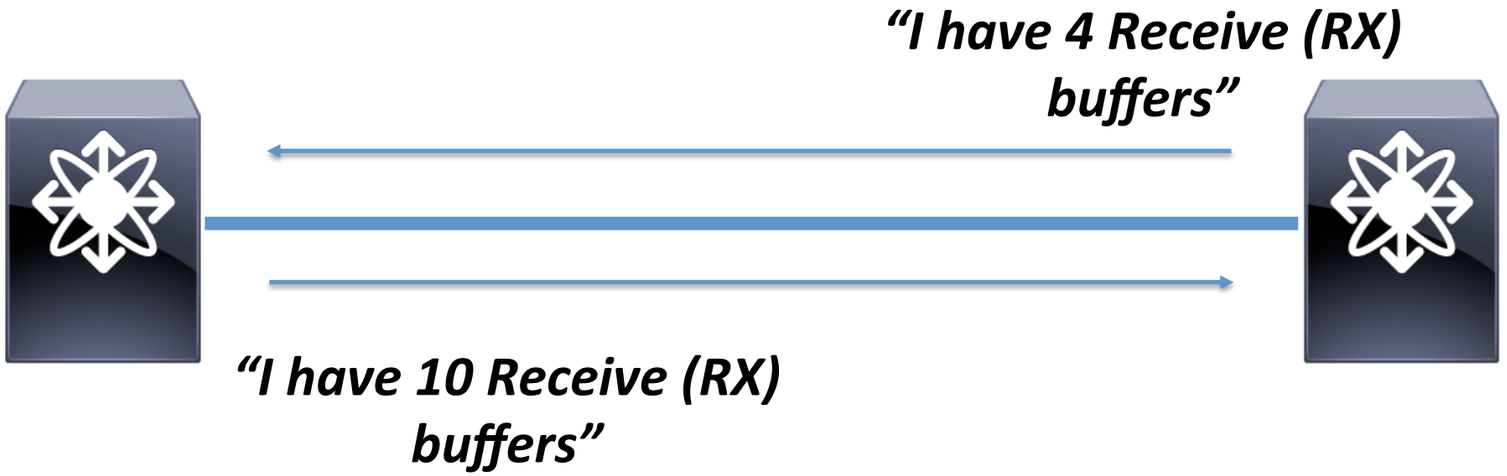
Fibre Channel Protocol Requirements

Fibre Channel Flow Control

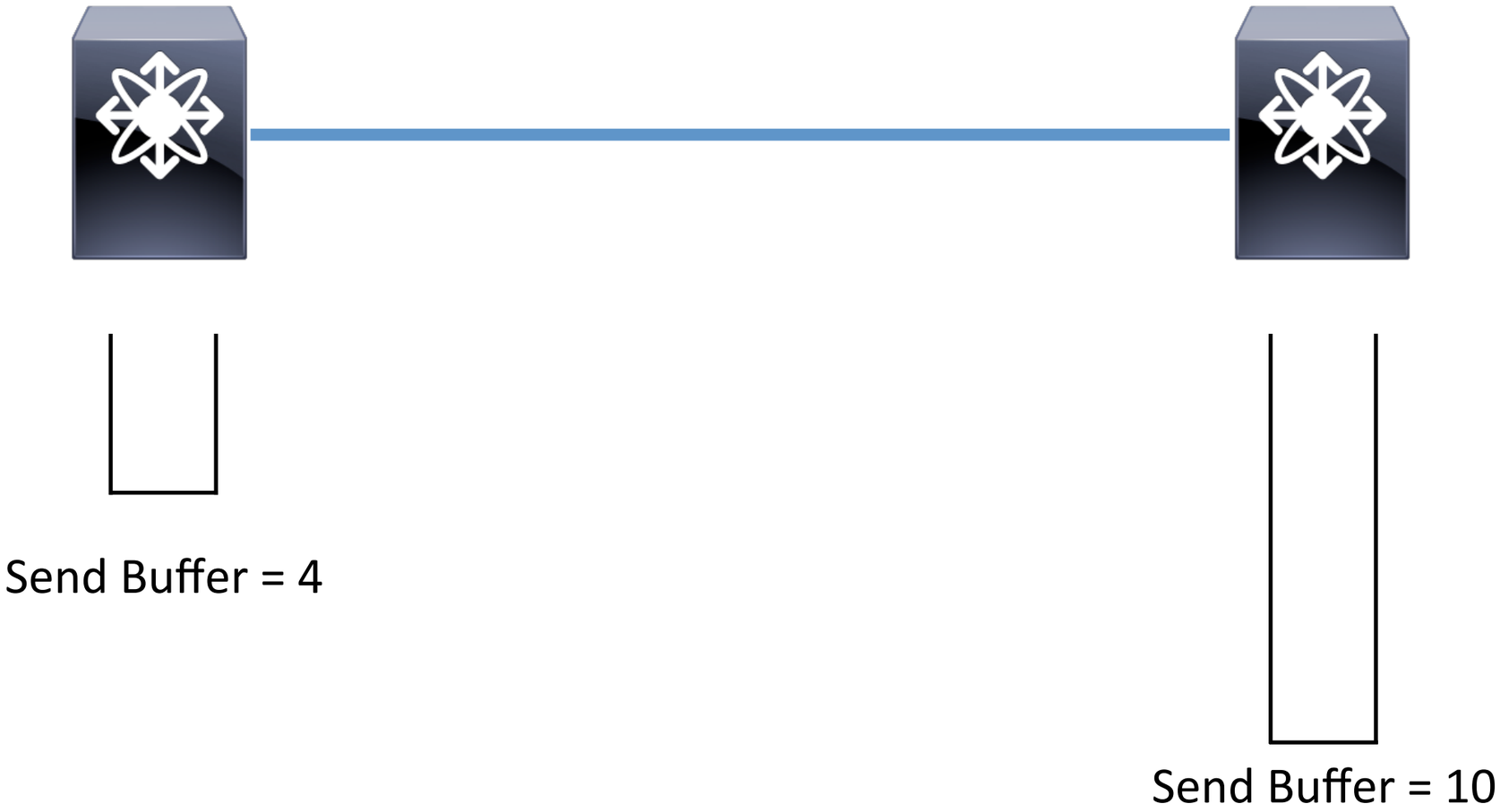
Fibre Channel uses concept of Buffer-to-Buffer Credits (BB_Credits) for flow control

- Every Frame to be transmitted requires one buffer
- Each buffer holds one frame regardless of the size of the frame
  - 64 byte frame uses the same buffer as 2112 byte frame
- BB_Credits negotiated between every pair of devices
  - Each side tells the other how many frames it can receive without being acknowledged
- Frame flow controlled by return of Receiver Ready (R_Rdy) frames

Buffer to Buffer Negotiation



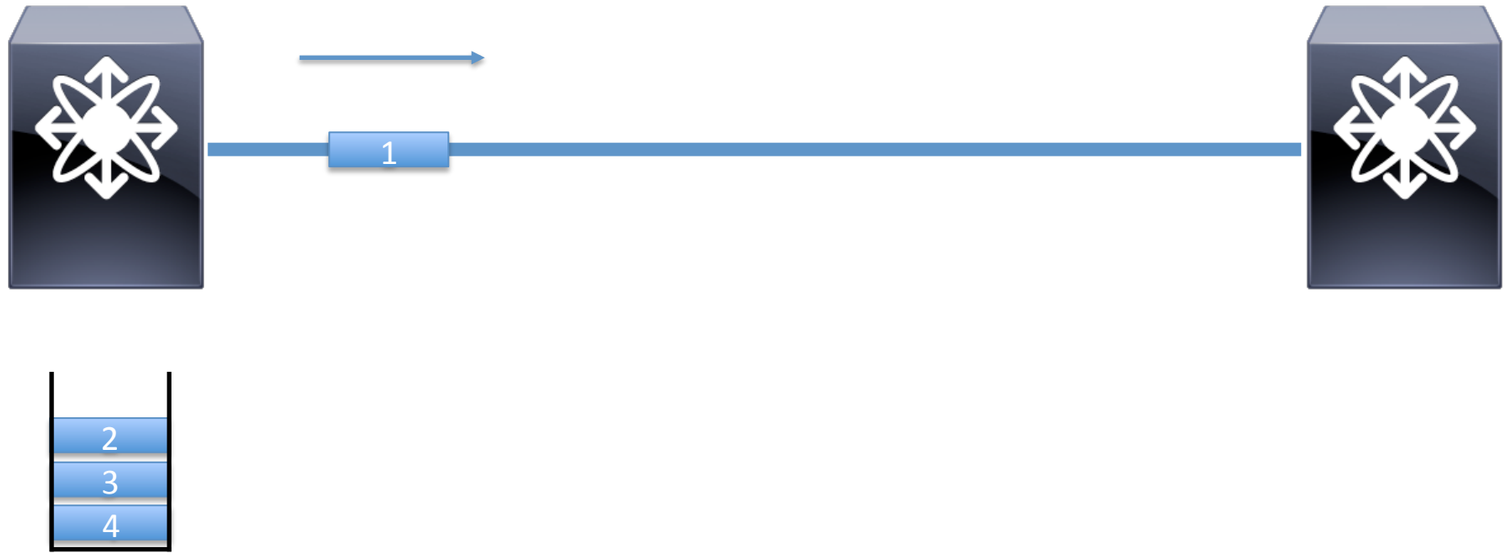
Buffer to Buffer Negotiation



Buffer to Buffer Function

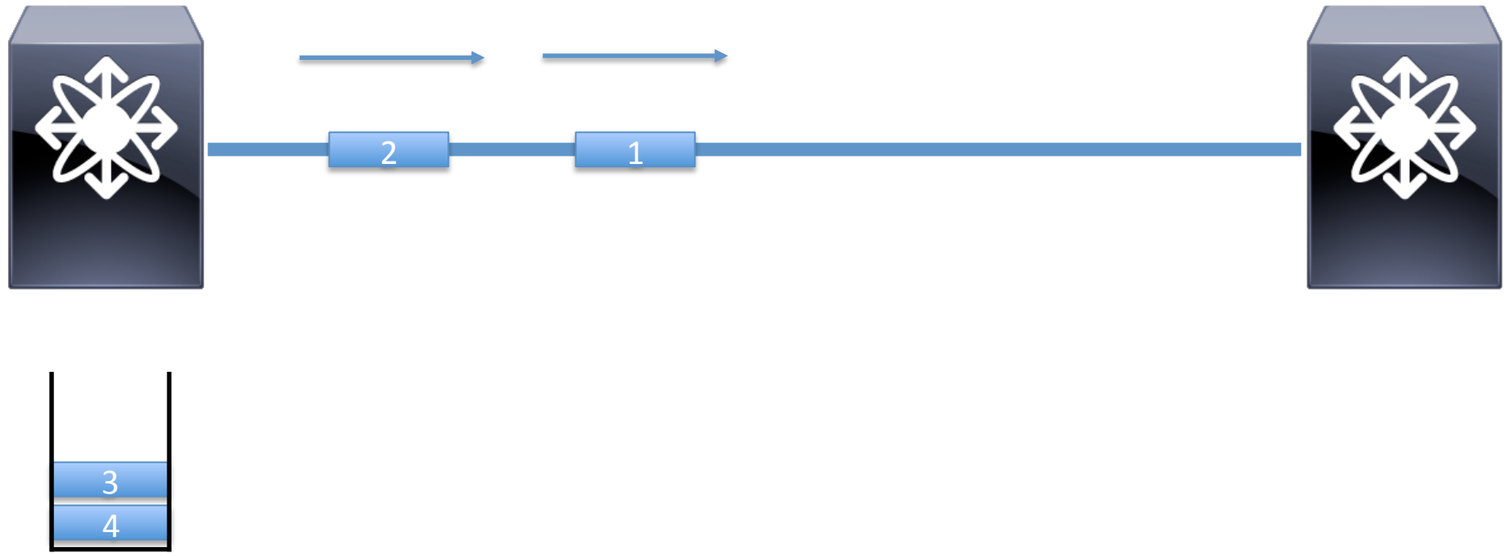
Send Buffer = 4

Buffer to Buffer Function



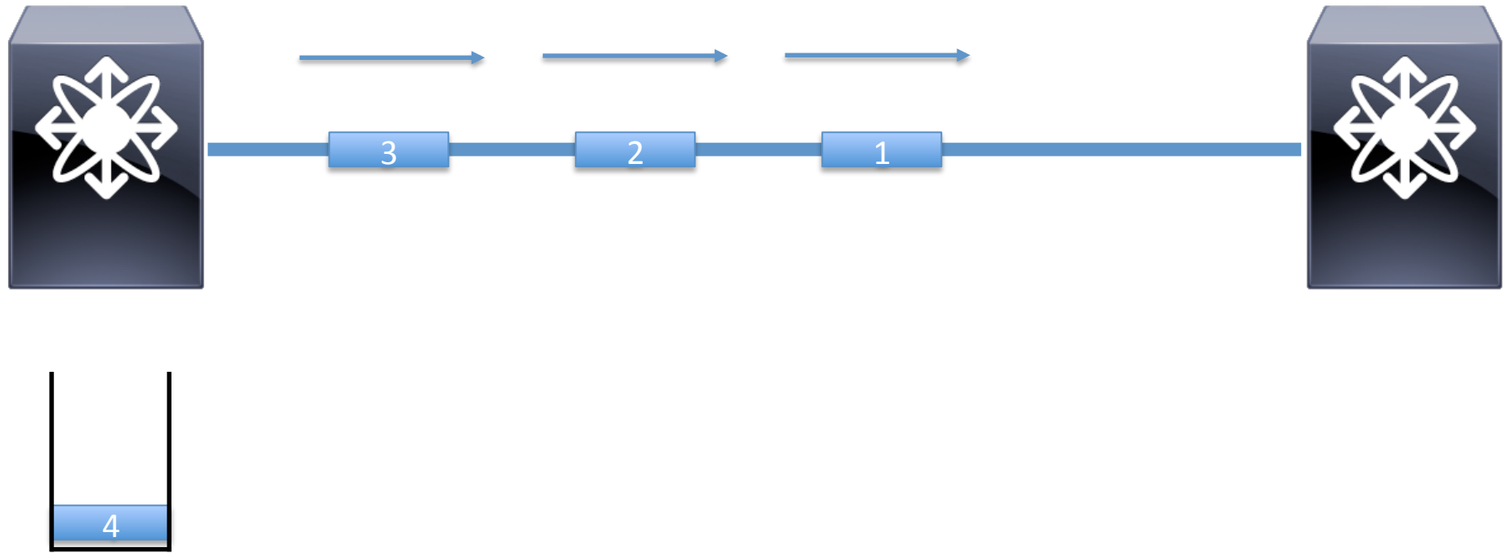
Send Buffer = 3

Buffer to Buffer Function



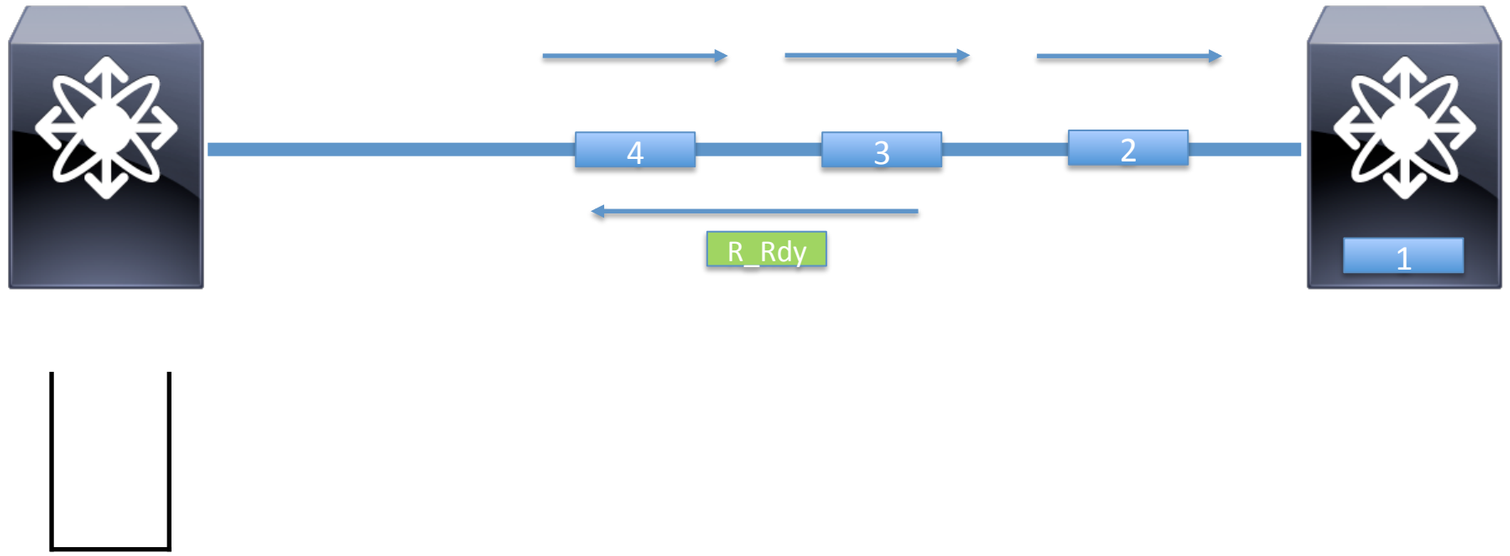
Send Buffer = 2

Buffer to Buffer Function



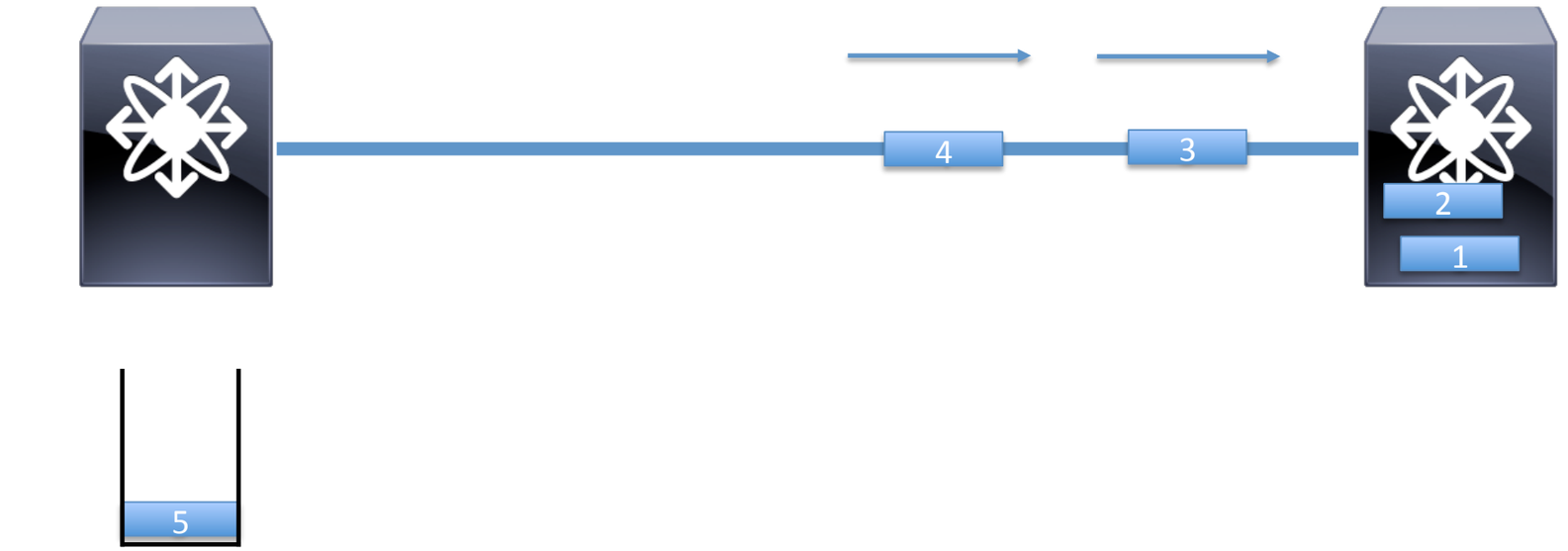
Send Buffer = 1

Buffer to Buffer Function



Send Buffer = 0

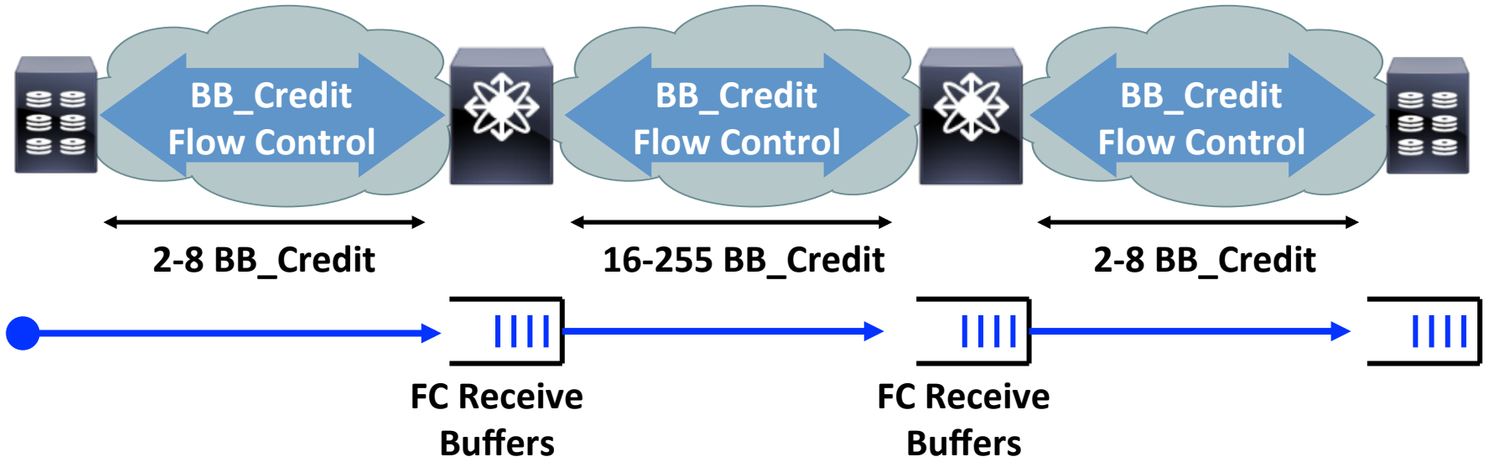
Buffer to Buffer Function



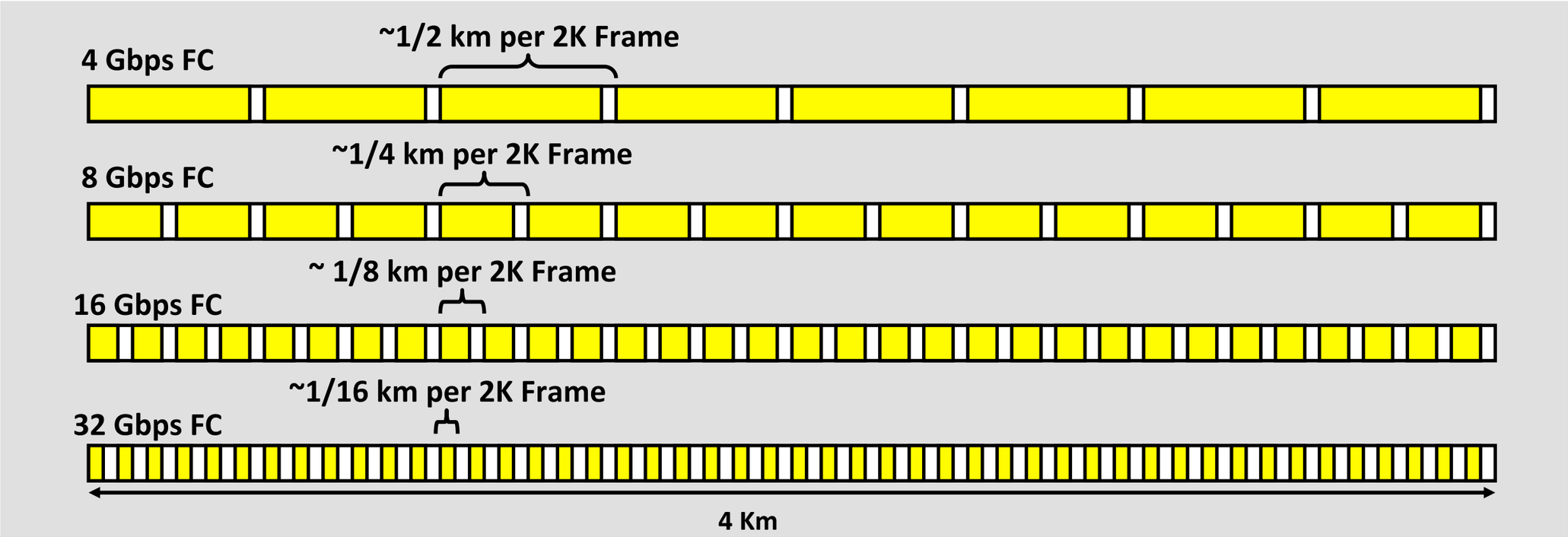
Send Buffer = 1

Flow Control in Multi-hop FC

- FC frames buffered and queued in intermediate switches
- Each device can only transmit up to the number of BB_Credits before traffic is throttled



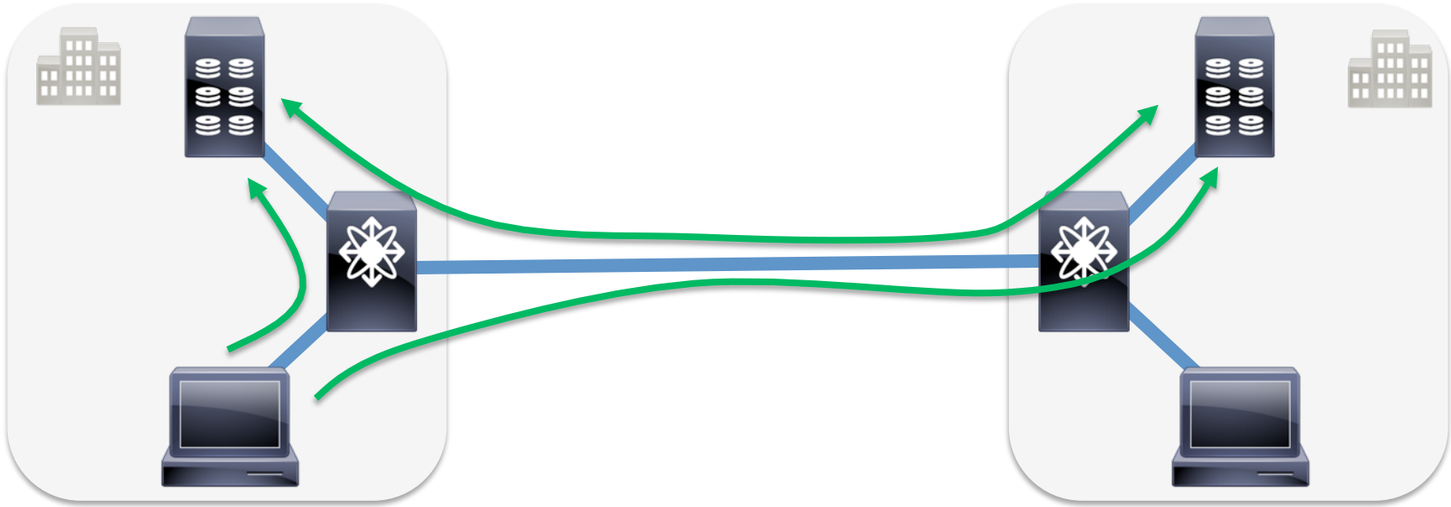
Extending Optical FC SAN Extension



- BB_Credits are used to ensure enough FC frames in flight
- As distance increases, the number of available BB_Credits need to increase as well
- Insufficient BB_Credits will throttle performance - no data will be transmitted until R_RDY is returned

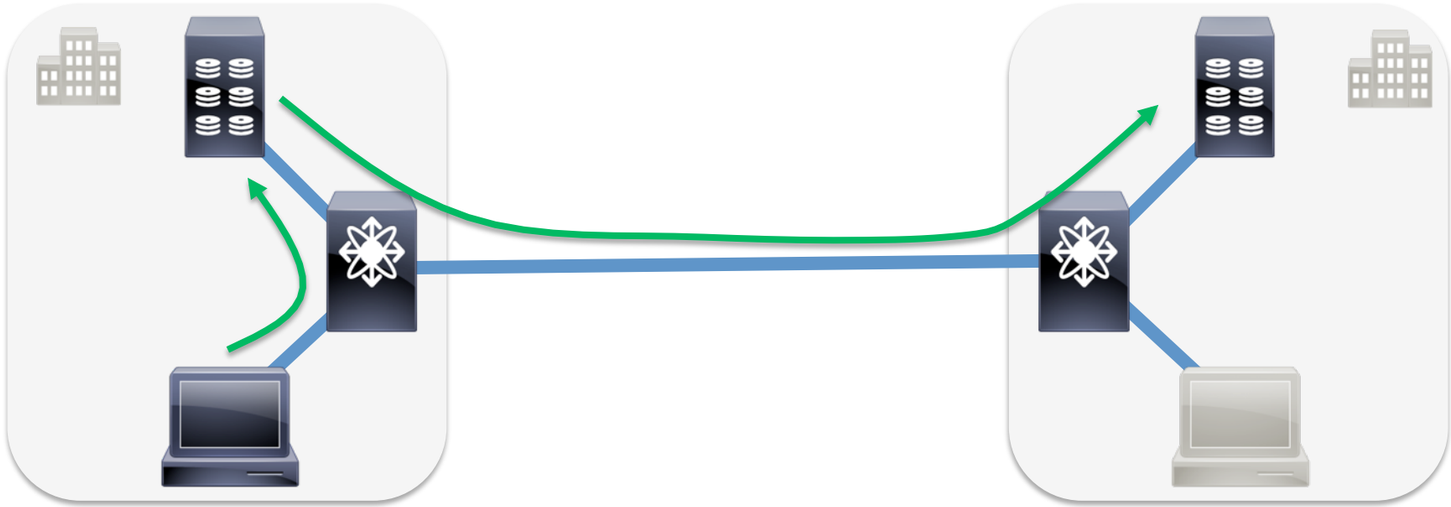
Long Distance Fibre Channel Uses Cases

- Active-Active Data Center
  - Servers access Data in both Data Centers
  - Disk Quorum / Synchronization required
  - Mix of Frame Sizes and Write/Read Traffic



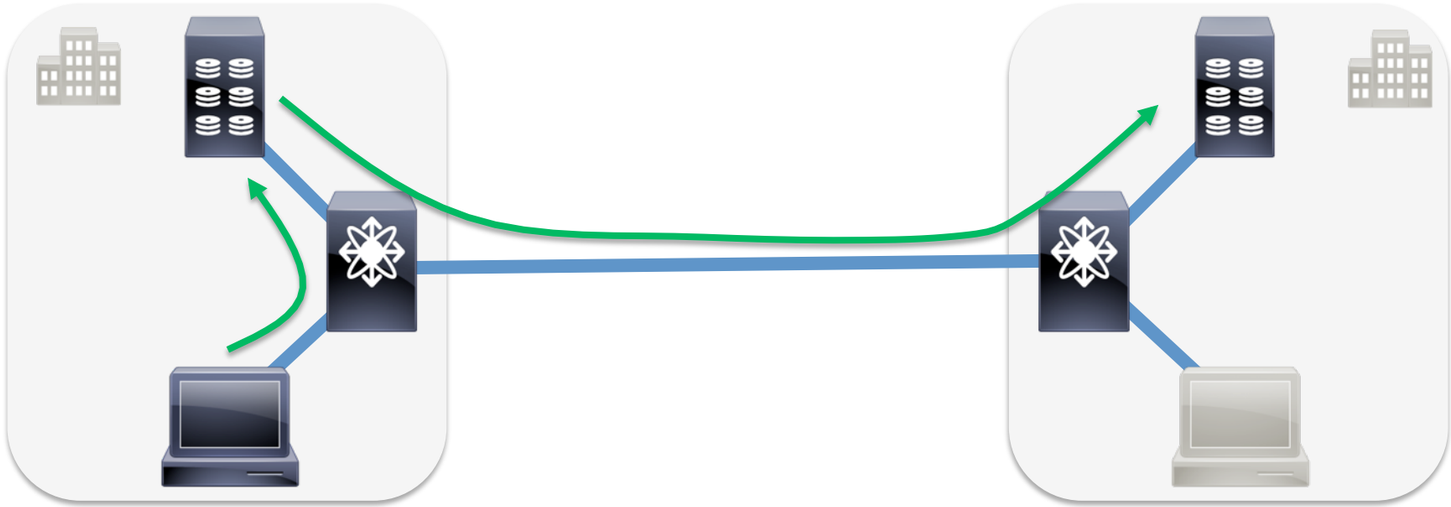
Long Distance Fibre Channel Uses Cases

- Synchronous Replication
  - Used when all data transactions are required to be in two data centers prior to acknowledging application (Zero Recovery Point Objective (RPO))
  - Write Traffic with a Mix of Frame Sizes



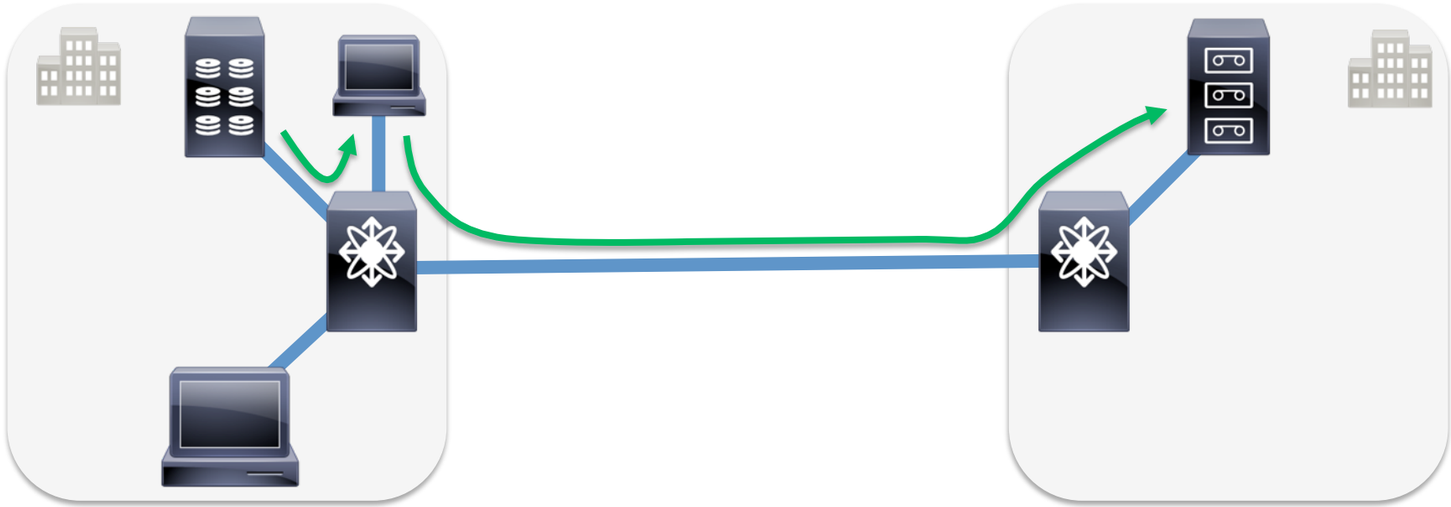
Long Distance Fibre Channel Uses Cases

- Asynchronous Replication
  - Used when data between data centers are not required to be identical (Low RPO)
  - Write Traffic with Large Frames Sizes



Long Distance Fibre Channel Uses Cases

- Data Backup
  - Archival backup from backup server to real or virtual tape devices
  - Primarily Write Traffic with Large Frame Sizes
  - Highest RPO of all solutions



FC Protocol Requirements Summary

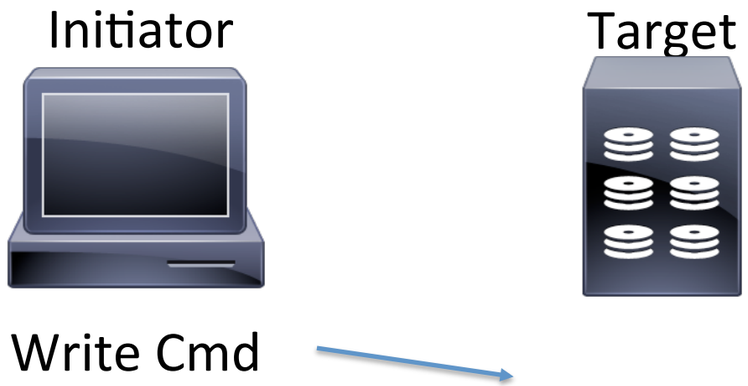
- BB_Credits negotiated between every all FC devices
- Need sufficient BB_Credits to fill link between Data Centers
  - Insufficient BB_Credits throttle performance
- Smaller frames require more BB_Credits
- Higher Speeds require more BB_Credits

Upper Layer Protocol Requirements

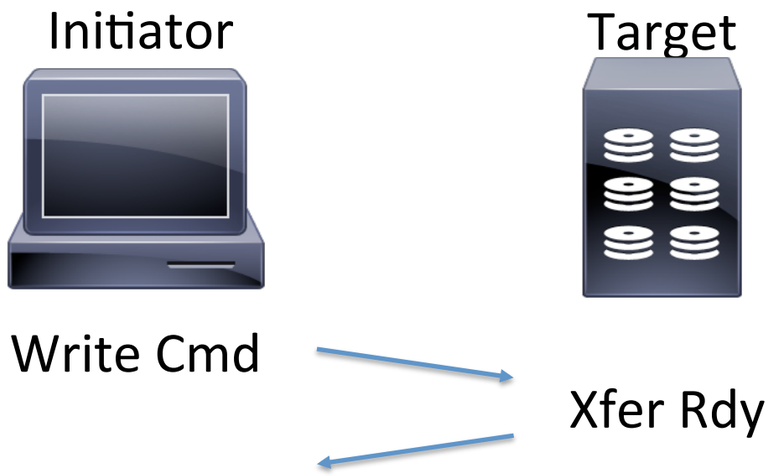
SCSI's Role

- SCSI is the most common protocol to run on Fibre Channel
  - FICON also used in Mainframe environments
- SCSI built on handshake protocol
  - SCSI Write commands acknowledged prior to Data transfer

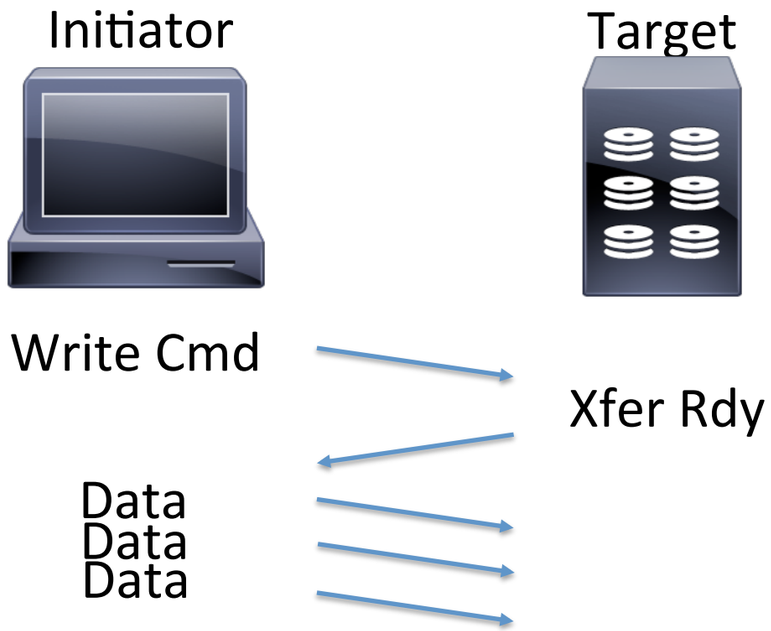
SCSI Write Exchange Sequence



SCSI Write Exchange Sequence



SCSI Write Exchange Sequence



SCSI Write Exchange Sequence

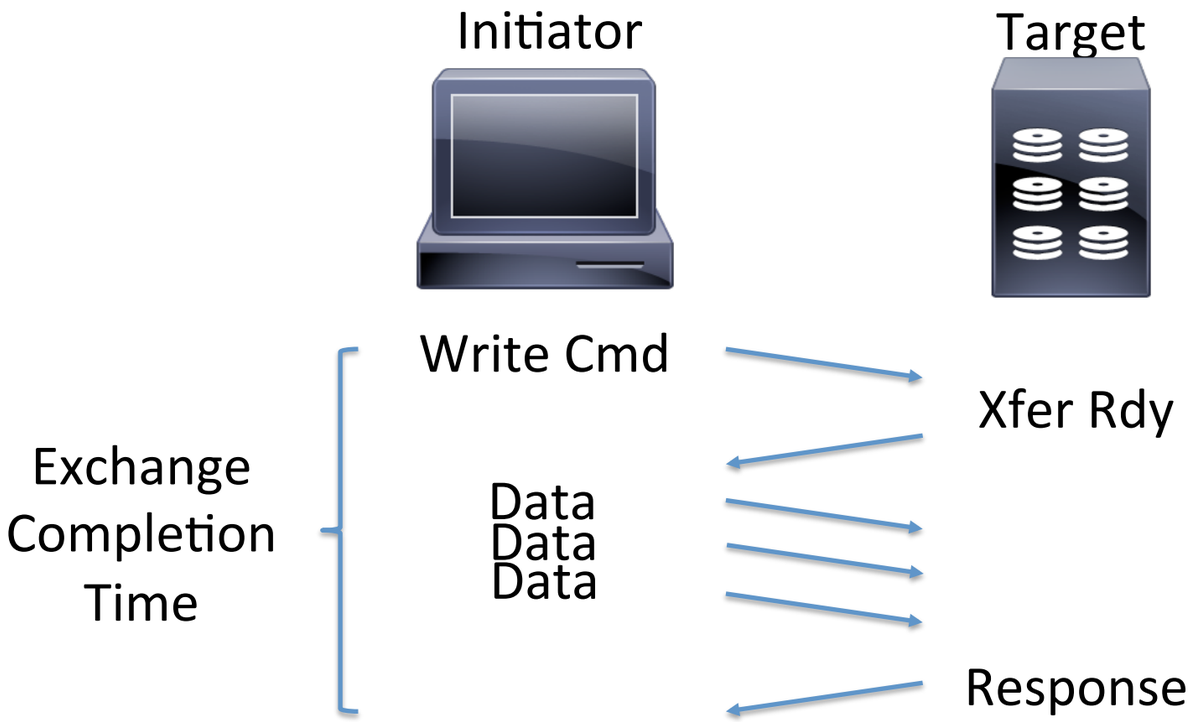
Write Cmd

Xfer Rdy

Data
Data
Data

Response

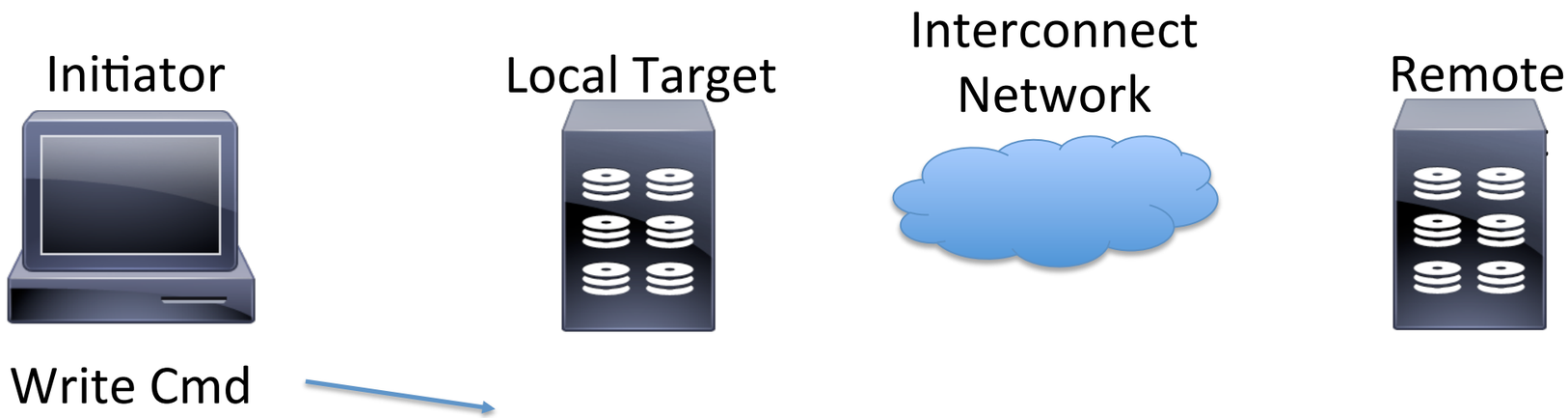
SCSI Write Exchange Sequence



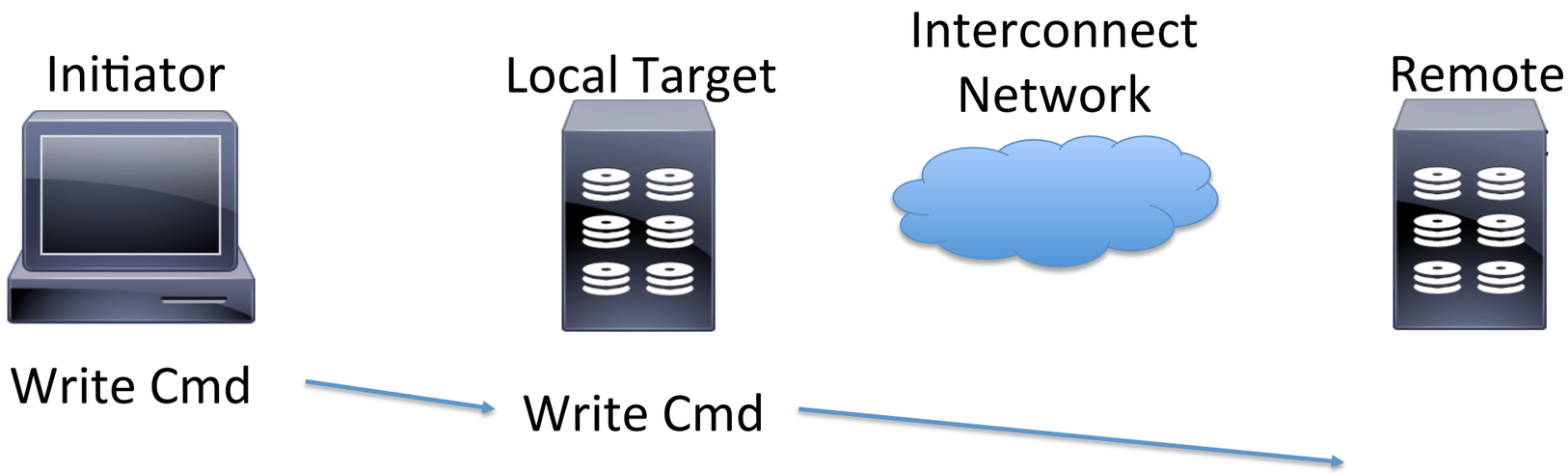
Affect of Distance on Applications

- When extending FC over distance, a delay of 5us per km is introduced
- SCSI write sequences require 2 round trips, so 20us per km per write command is the minimum added latency

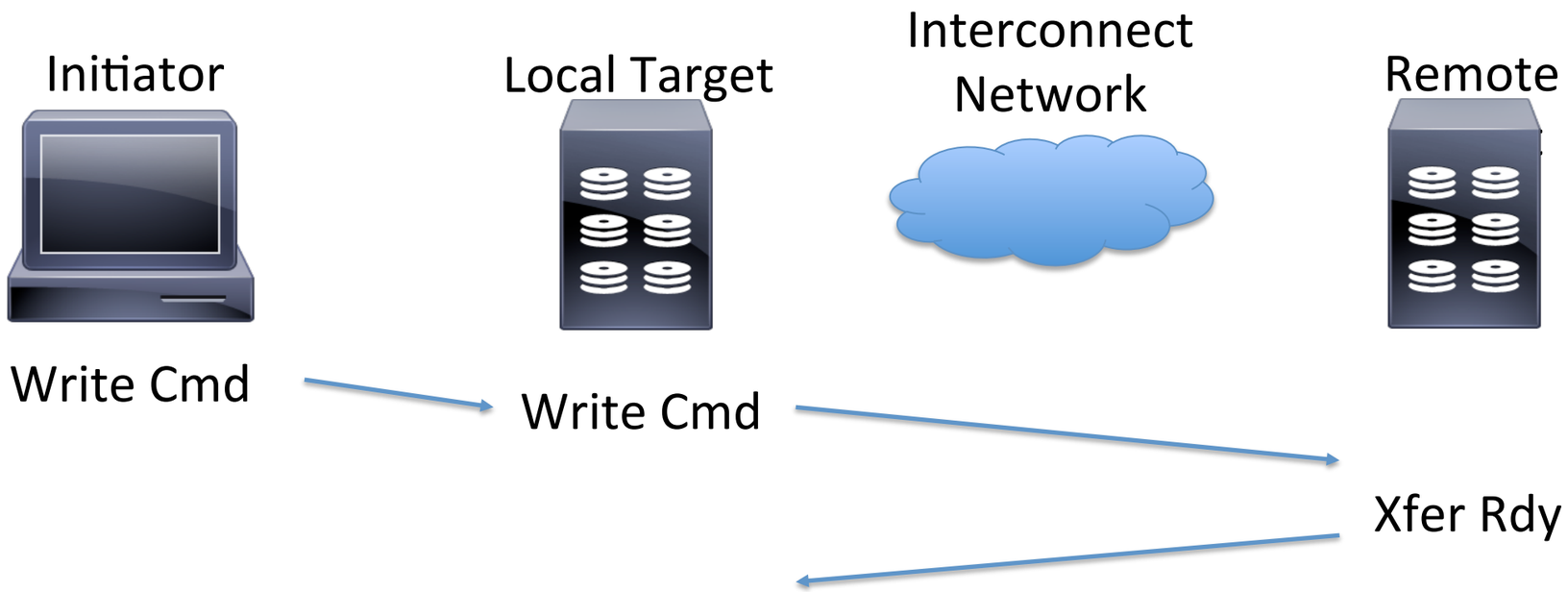
SCSI Write Exchange Sequence



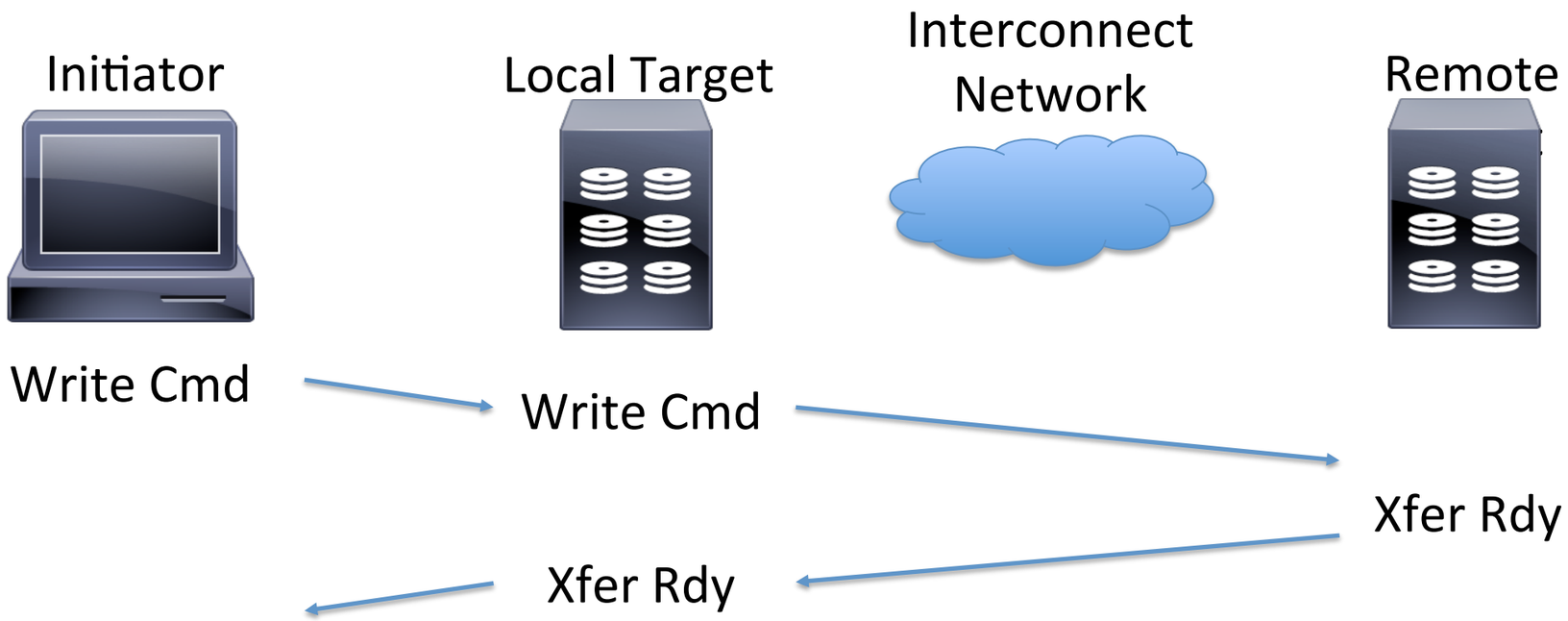
SCSI Write Exchange Sequence



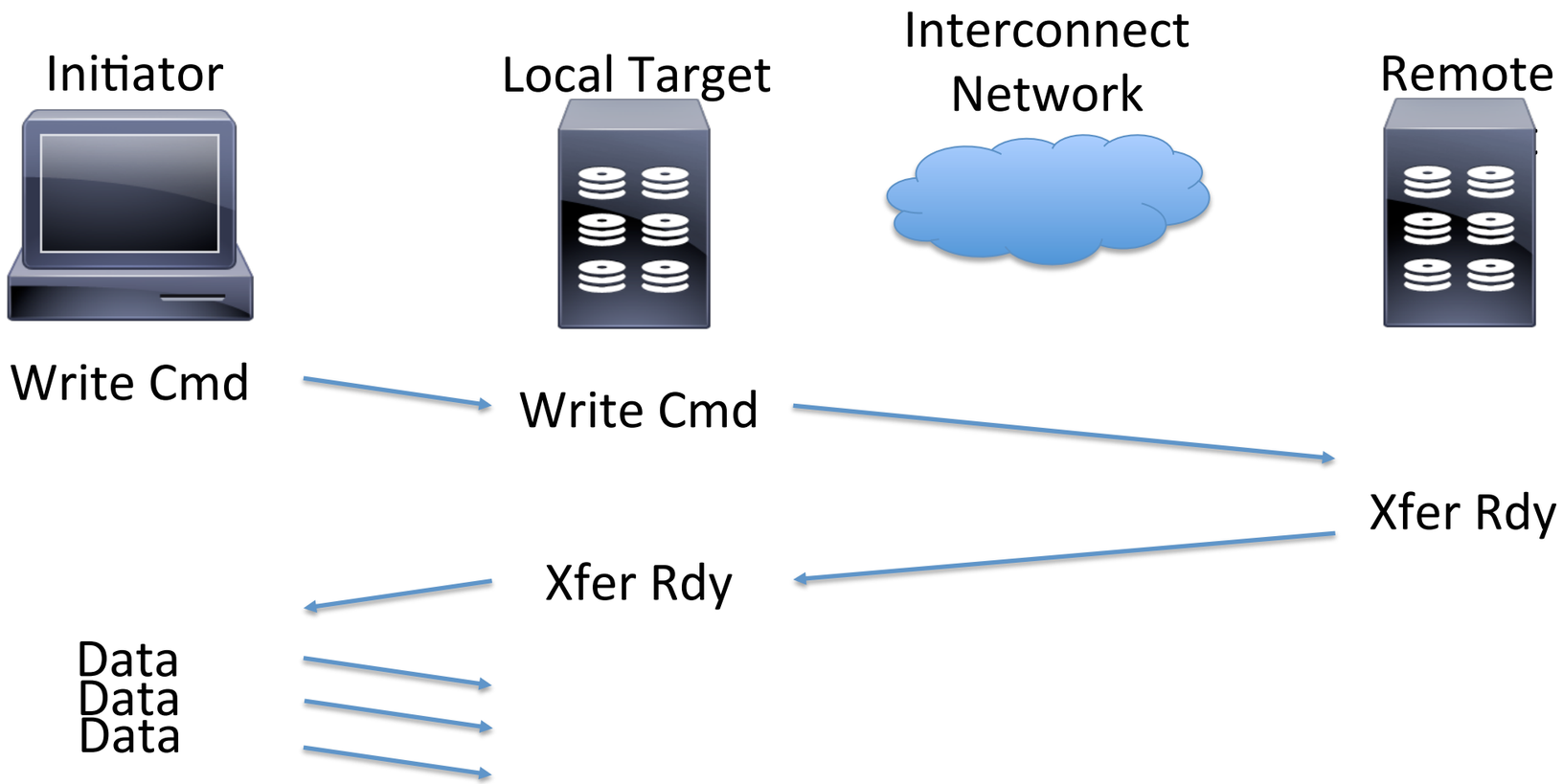
SCSI Write Exchange Sequence



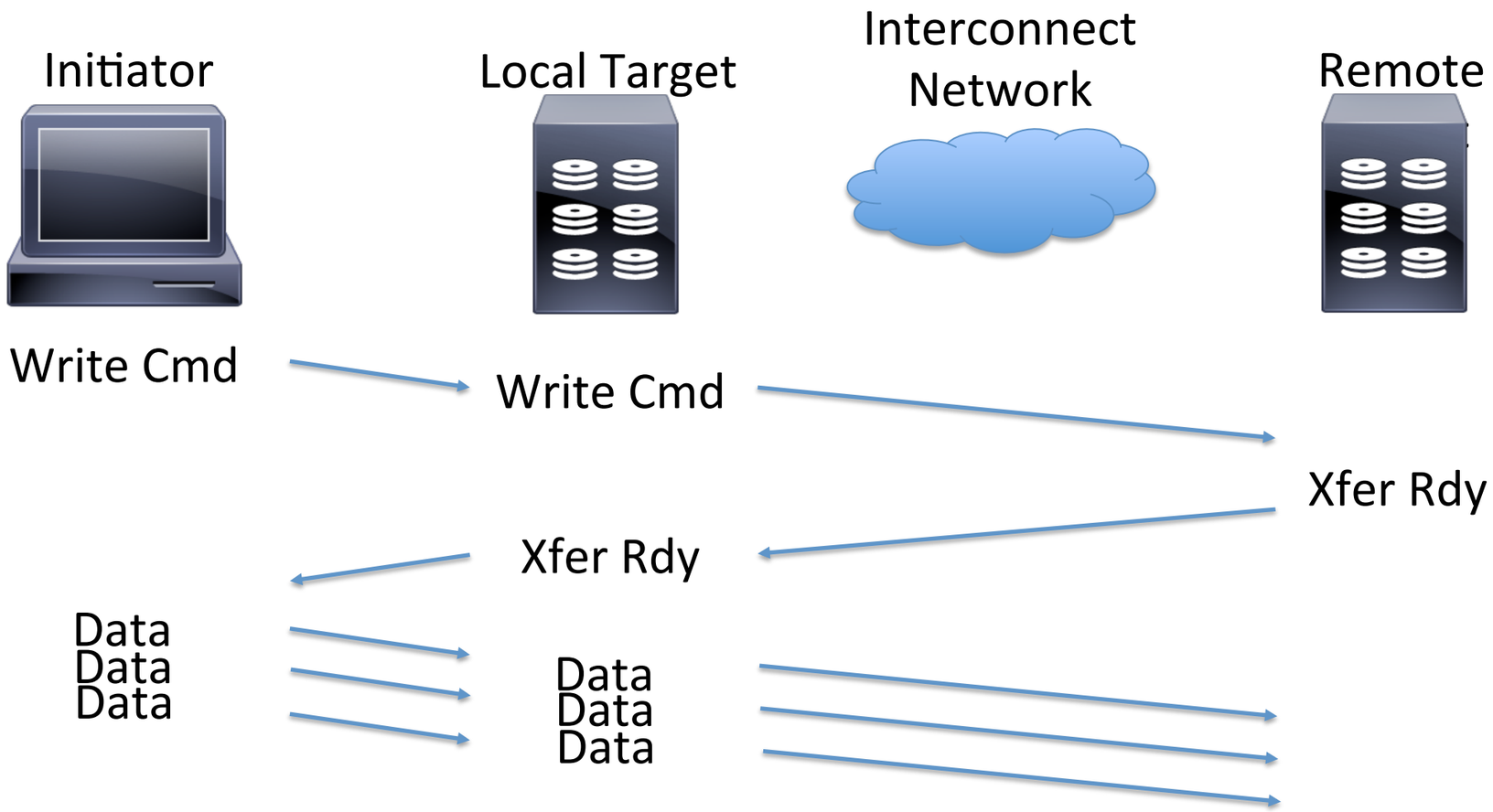
SCSI Write Exchange Sequence



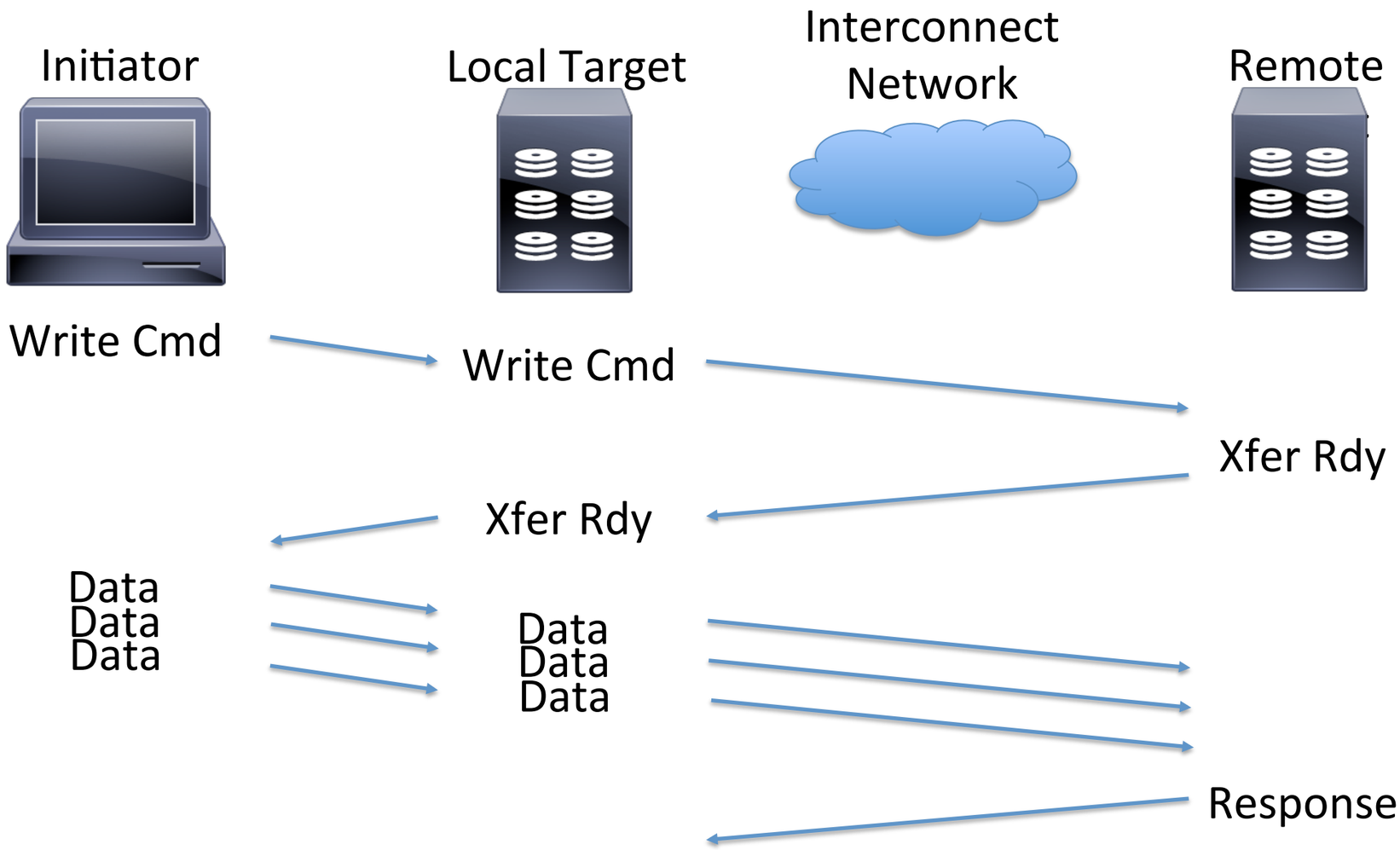
SCSI Write Exchange Sequence



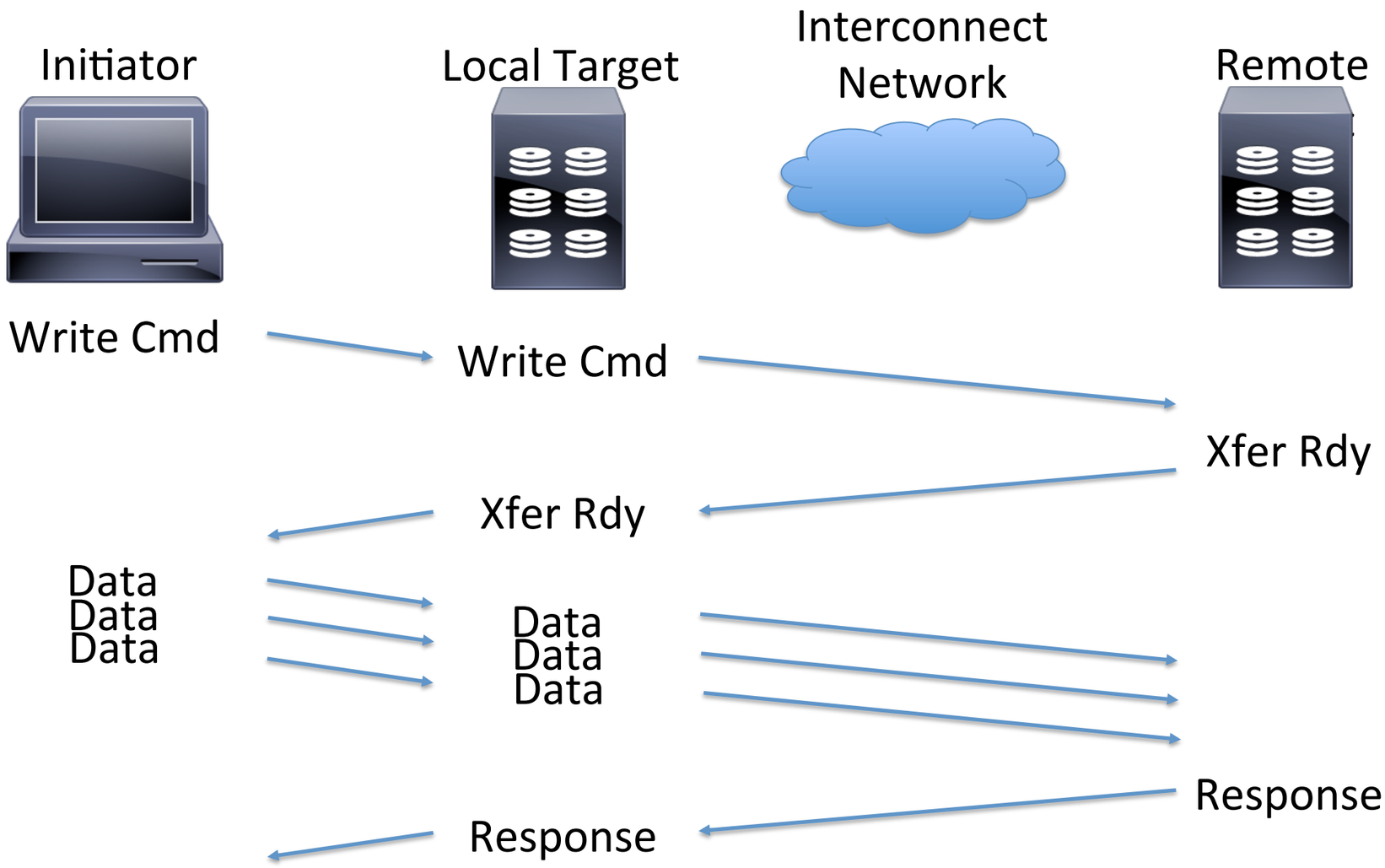
SCSI Write Exchange Sequence



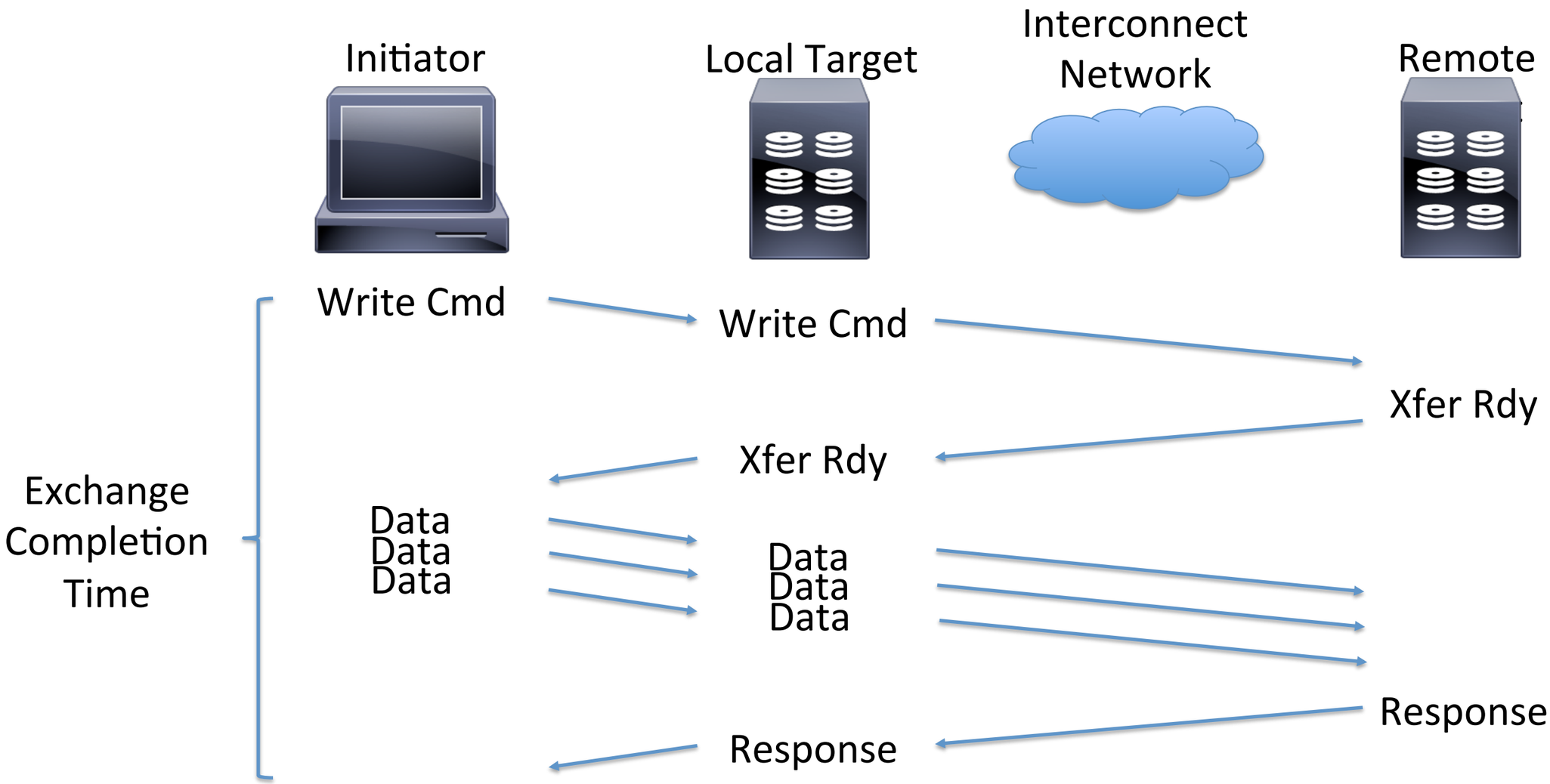
SCSI Write Exchange Sequence



SCSI Write Exchange Sequence

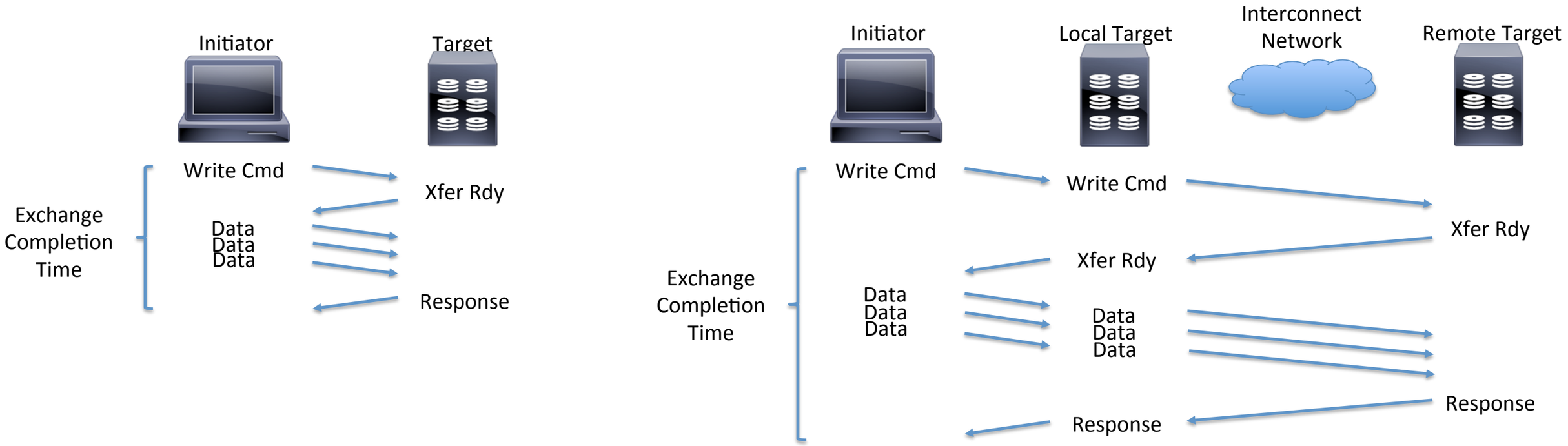


SCSI Write Exchange Sequence

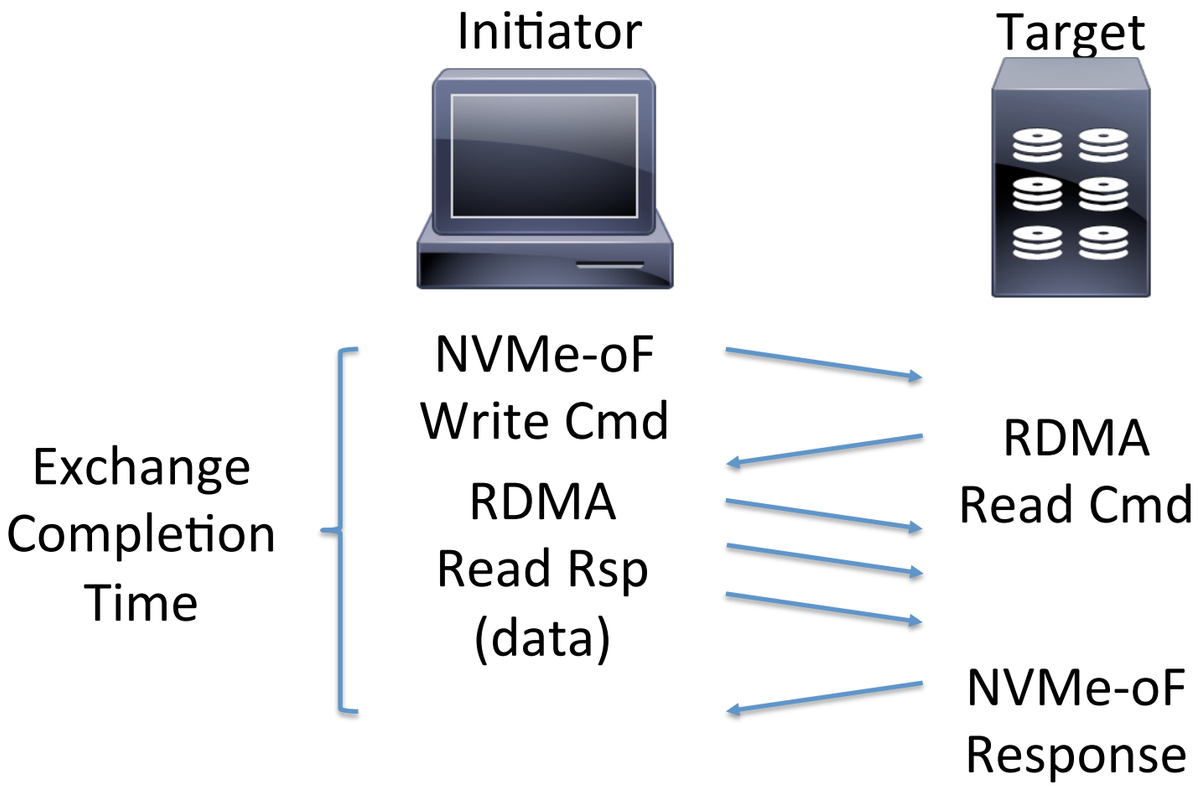


SCSI Write Exchange Sequence

- For Synchronous Replication, delay introduced by distance can be significant; 100km would add at least 2ms of application latency



NVMe-oF Write Exchange Sequence

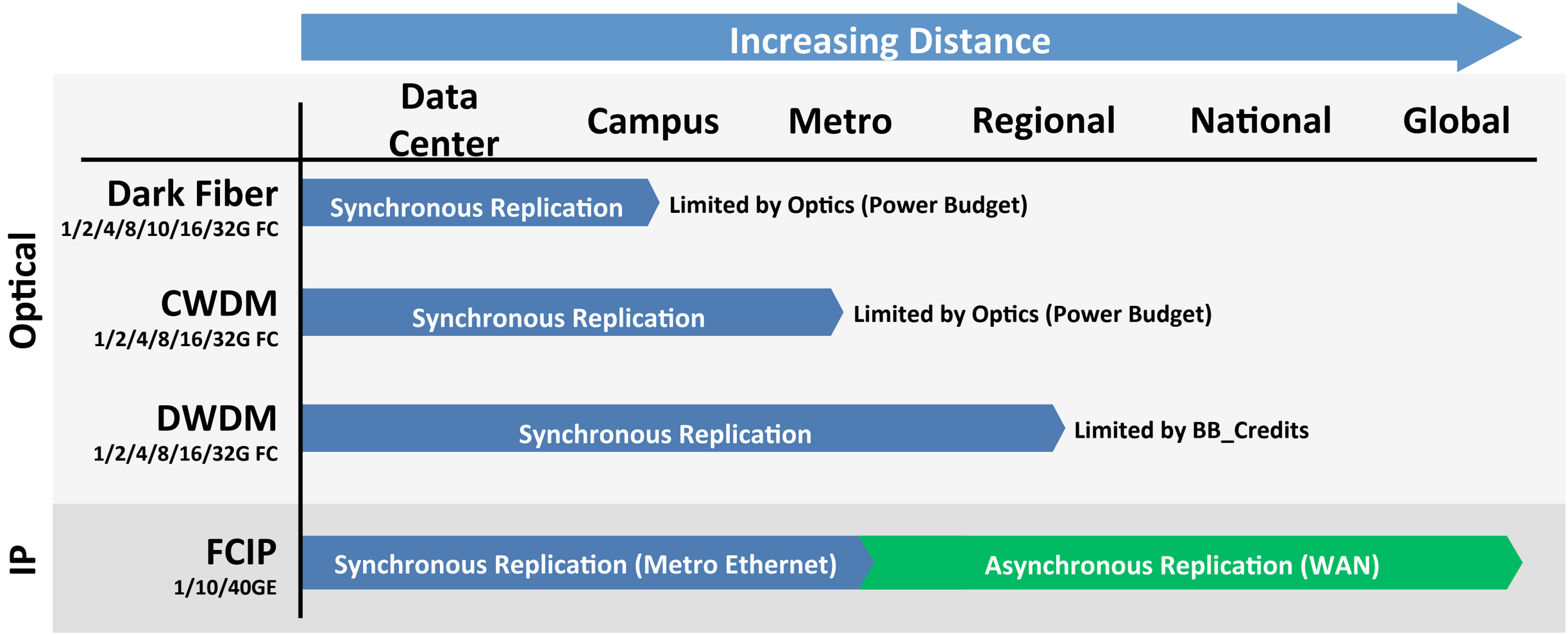


ULP Requirements Summary

- SCSI write commands require 2 Round Trips
- Latency introduced by distance may impact application performance in Active-Active or Synchronous Replication environments
- No NVMe-of distance solutions today, but distance latency will have greater impact

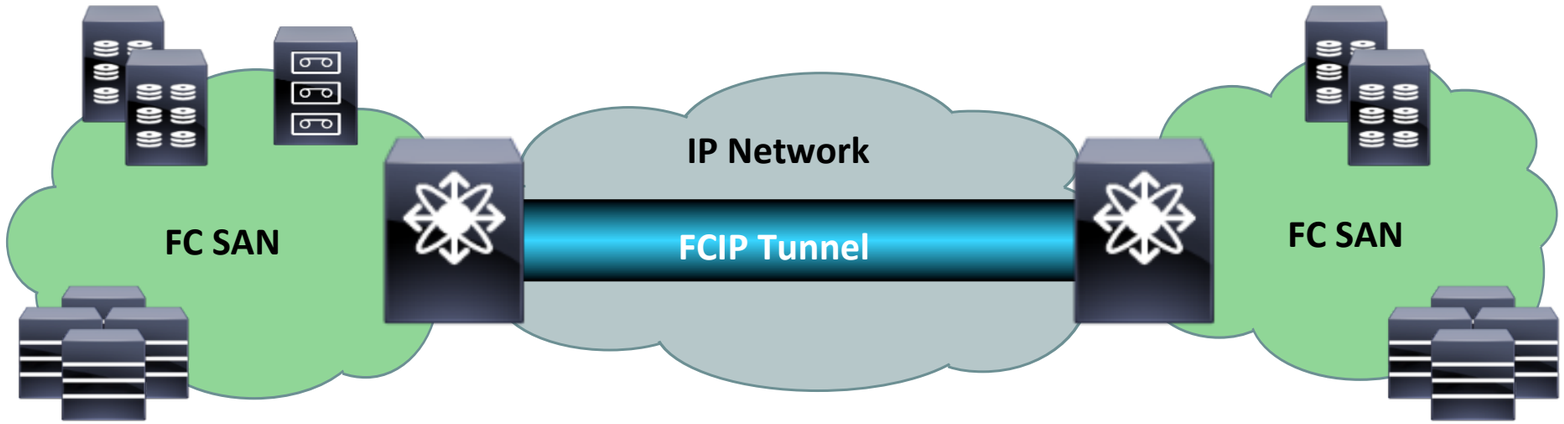
Other Long Distance Options

Long Distance FC Technology Options



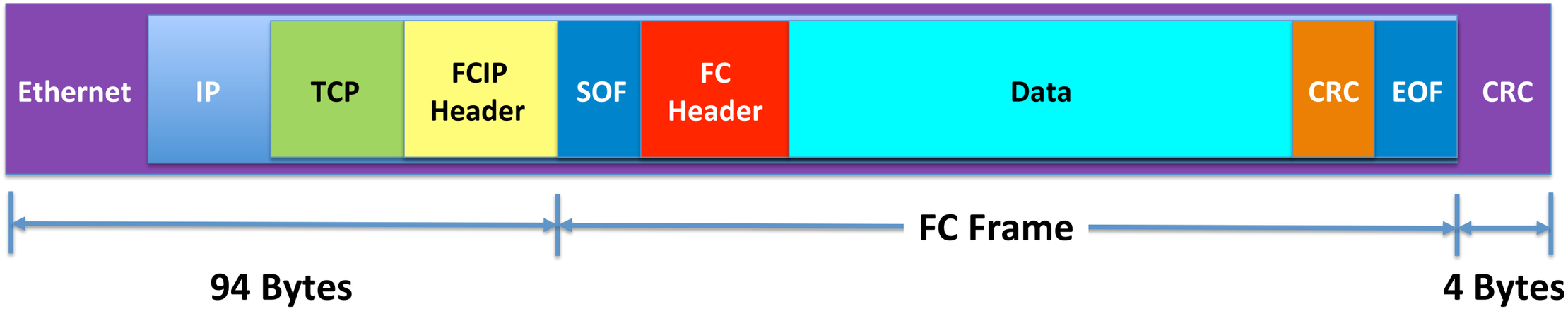
FCIP: Fibre Channel over IP

- FCIP is a standard from the IETF IP Storage WG for Linking Fibre Channel SANs over IP (RFCs 3821 and 3643)
  - Point-to-point tunnel between FCIP link end-points
  - Creates one logical FC fabric with single FSPF routing domain



FCIP Frame Detail

- Max Fibre Channel frame is 2148 bytes
- FCIP will segment and reassemble FC frames if MTU too small (no FCIP header on subsequent packets)
  - Jumbo frames may increase performance



Summary

Fibre Channel over Distance

- **Many Physical Layer options**
  - Choose according to availability and requirements
- **Fibre Channel requirements need to be considered**
  - Make sure sufficient BB_Credits based on speed, distance and application
- **Latency introduced by distance a factor to consider**
  - What impact to application performance is acceptable?

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 all the great questions we received today
- Next webcast: “FCIP (Extension): Data Protection and Business Continuity”
- Follow us @FCIAnews for announcements and details

Thank you!