FIBRE CHANNEL OVER DISTANCE SOLUTIONS

Storage replication has historically occurred within the data center, having the primary—or production data—situated beside a back-up data copy. World events, weather phenomenon, and the drive to global 24/7 business operations have required enhanced data resiliency by having additional copies geographically separated from the primary data repository; in some cases, regulations have been defined that mandate data separation to ensure continued business operations in lieu of damage or destruction of the production data environment. All of these facts have caused a major focus on Business Continuity and Disaster Recovery (BCDR) solutions and all leverage the WAN as a critical component of an implementation's infrastructure.

Considerations for Storage Extension

There are four major considerations when extending storage over distance:

Response Time—the sensitivity to Round Trip Time (RTT) latency. This is a concern at both the protocol and application levels, as the stringent nature of replication normally has application timeouts and extensive protocol handshaking built in to ensure zero frame loss. In many cases, the storage application will specify maximum latency values to ensure proper operation.

Data Loss Sensitivity—also called the Recovery Point Objective (RPO), is the tolerance of the administrator/company to the loss of data in the event of a disruption. RPOs vary by application and vertical, for instance stock transaction databases for a financial institution will mandate an RPO of "0" whereas a laptop back-up database may have a longer RPO (of 24 hours for example) as the data is not critical for business operations and a previous back-up file could be leveraged.

Recovery Time Sensitivity—also called the Recovery Time Objective (RTO), is the tolerance of the administrator/company to have the database offline during recovery operations after an event. Using the example above, the stock transaction database will have an RTO of near "0" whereas the back-up database will have a longer RTO. RTOs are dictated by how impactful lack of access to the information held within the database is to business operations.

Infrastructure Requirements—the requirements with regards to amount of bandwidth required to fulfill normal replication from the primary to secondary data center, excess bandwidth available to address deferred replication data (due to a maintenance window or other situation), and the resiliency of the WAN to limit the impact of fiber cuts, etc. during normal replication operations.

RPO and RTO are primarily addressed at the application level. These objectives dictate the type of data replication undertaken (synchronous, asynchronous, and point-in-time) and won't be a focus of this application note. However, RPOs do typically impact the distance permitted between source and target data repositories based on latency allowances.

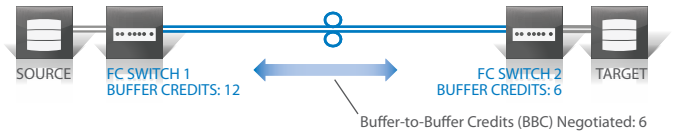
Response time and infrastructure requirements are primarily addressed at the protocol level. These requirements necessitate an optimized Fibre Channel communications link between sites that ensures proper transactional processes can operate and that there is sufficient bandwidth available to fulfill transfer requirements.

The Fibre Channel Protocol and Distance

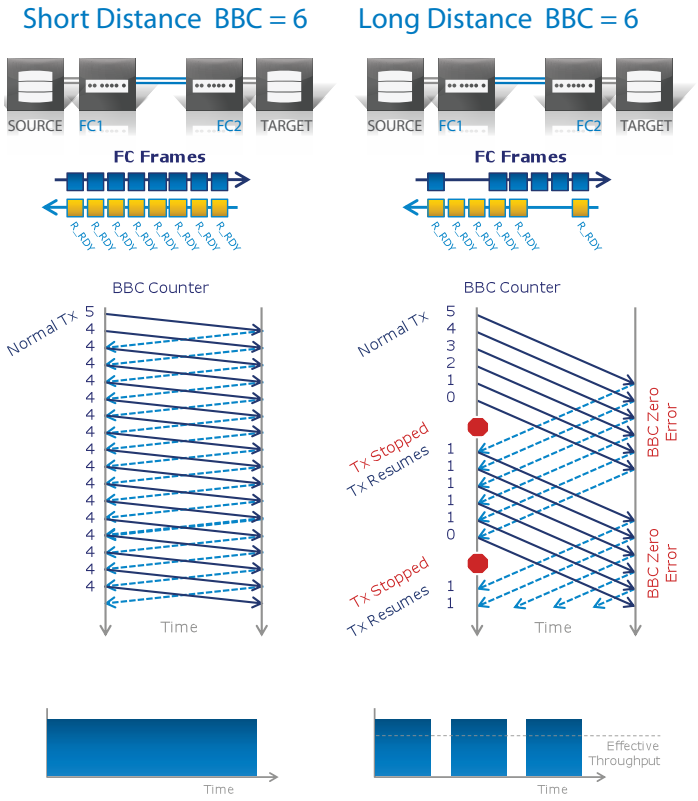
Fibre Channel has strict transactional processes driven by the critical nature of the data being transferred and ensuring deterministic, in-order delivery and guaranteed committment to the target data repository. The requirement for a limited Recovery Point Objective (RPO), or difference in data sets between the production and back-up repositories, is also required, especially in some enterprise verticals such as financial. These characteristics mean that Fibre

Channel over distance solutions must be precisely engineered. This is definitely the case with regards to flow control mechanisms that guarantee proper Fibre Channel frame delivery.

Fibre Channel has a built in flow control mechanism called Buffer-to-Buffer Credit (BBC) Flow Control. In BBC Flow Control, the source and destination negotiate the number of Buffer Credits available when the link between them is provisioned. The negotiated number is the maximum number of Buffer Credits that either the switches or directors can support.



Buffer Credits can be considered equivalent to the maximum allowable unacknowledged frames, or frames-in-flight, over the WAN. A BBC Counter at the source is decremented by 1 whenever an FC frame is sent. A frame must be committed to the target data repository prior to an acknowledgement frame, called an R_Ready, being sent back to the source to increment the BBC counter. If the BBC Counter reaches "0" the source switch stops sending data and transmission will resume when the BBC Counter is incremented by receiving an R_Ready. This state indicates that there are not enough Buffer Credits available between the source and target switches to support full-rate transmission. The illustration below compares two situations where the BBC count is sufficient or insufficient to address full-rate throughput between end points. When the BBC Counter reaches "0" a BBC Zero error is registered. A BBC Zero error does not indicate transmission errors, only the status of the counter and the link's inability to support full-rate throughput.

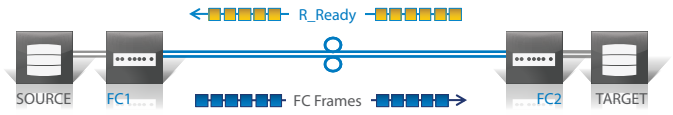


Buffer Credit link extension operations are typically provided by the Fibre Channel switch or director at the SAN-WAN demarcation. Note that vendors provide a minimum baseline of Buffer Credits with switches and customers must purchase extended distance licenses to enable more Buffer Credits within the system. Customers are sometimes reluctant to add extended licenses, or not familiar enough with Fibre Channel extension to understand the necessity of Buffer Credits in a storage extension solution. Buffer credit link extension is the protocol-level assurance that Fibre Channel frames have been delivered and that the target database is not corrupted. BBC Flow control ensures that the target database is not corrupt and is capable to support business operations in the event of a failure of the production environment.

Fibre Channel over Distance Implementation Options

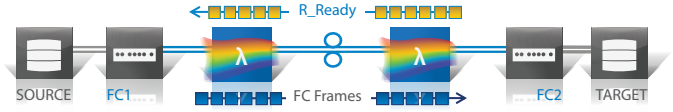
There are multiple ways to implement a Fibre Channel over distance solution:

Direct on Fiber



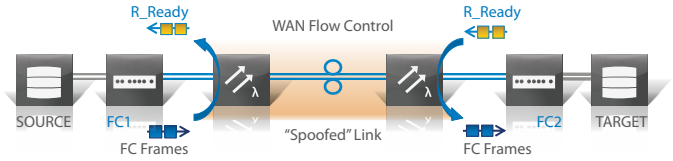
The Direct on Fiber solution (illustrated above) is fiber-intensive as each link necessitates a fiber pair. In addition, Fibre Channel switches and directors typically can't address long span distances so optical reach limits are possible. No protection options are available. BBC requirements are based on fiber distance/link latency only.

Transparent Optical Transport



The Transparent Optical Transport solution (illustrated above) requires the implementation of a WAN solution—such as WDM for transport of Fibre Channel. WAN solutions provide network multiplexing to ensure the efficient use of physical fiber and provide enhanced reach relative to direct on fiber solutions. In addition, performance monitoring enhances solution visibility and integrated protection options provide higher availability. BBC requirements are based on fiber distance and WAN equipment latency.

Optical Transport with Fibre Channel Spoofing



The Optical Transport with Fibre Channel Spoofing solution (illustrated above) requires the implementation of a WAN solution, but adds a WAN flow-control domain which locally acknowledges flow control transactions. As there is no longer an end-to-end flow control, the system is considered to be "spoofing" the Fibre Channel switches at either end. The system provides similar value as the Transparent Optical Transport solution—with the exception that there is a potential for corrupt data being deposited in the target repository due to the WAN equipment getting involved in the flow control of a link (i.e., it's not transparent to the solution). BBC requirements are reduced based on spoofing—this overcomes latency with the consequence of potentially having a damaged database to recover from.

Addressing Storage Extension over Distance

The impact of latency on storage extension must be addressed. The primary way to address latency requirements, without implementing a spoofing solution, is to ensure there are sufficient Buffer Credits to optimize throughput. Effective throughput could either be full-rate, or a sufficient level to address link utilization requirements (i.e., 1.4 Gbps on a 2 Gbps FC link). Ensuring optimal streaming of data over the distance between the two data centers is critical and is a function of fiber distance, transmission speed, Fibre Channel frame size, and the latency of transport equipment.

The physical distance between sites is a primary factor for buffer credit optimization. It is typically the largest contributor to link latency. It must be kept in mind that even though the physical distance is fixed, the number of frames that can exist in transit along a fiber-optic path that connects two adjacent nodes/sites is dictated by the transmission speed and the frame size:

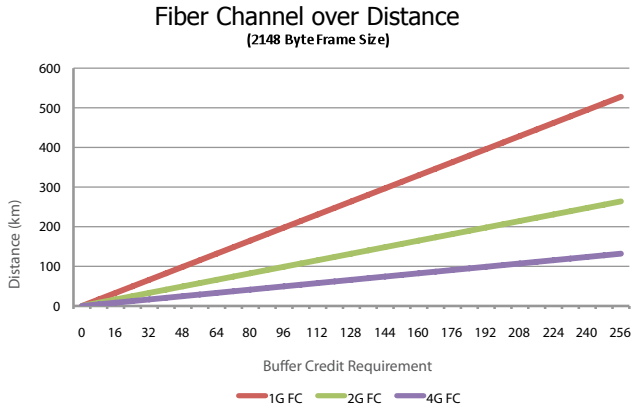
Increases linearly as the transmission speed increases

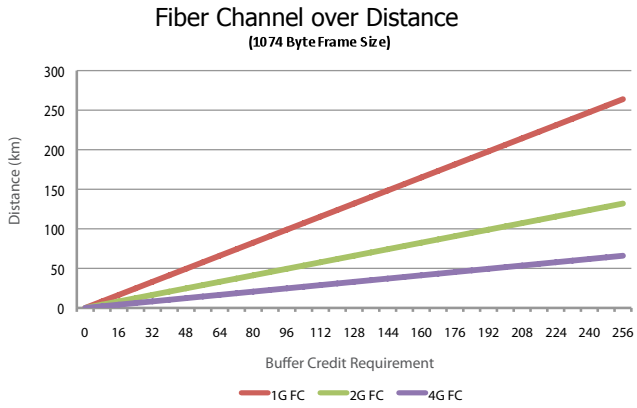
1G → 2G → 4G

Increases linearly as the size of the frame decreases

1000Bytes → 500Bytes

These two factors—transmission and frame size—impact the quantity of frames-in-flight that a link can support over the physical fiber, and thus dictate how many Buffer Credits are required to ensure the optimal streaming of data. The graphs below show how Fibre Channel transmission rates and frame size alter requirements for Buffer Credits over distance. SAN and WAN administrators must understand that increasing the link speed will not enable increased Fibre Channel throughput alone—careful engineering of Buffer Credits is also a necessity.





Buffer Credit Availability

Depending on the type of Fibre Channel switch or director to be used, Buffer Credits available for distance extension can be minimal. Most vendors offer extended fabric licenses to make available a larger pool of Buffer Credits which can be assigned to Inter-Switch Links (ISLs). Typically a maximum of 256 Buffer Credits per port can be assigned from the extended fabric license Buffer Credit pool; this defines a theoretical maximum distance for a Fibre Channel link. In addition to fiber latency, the latency of WAN equipment must be considered as it will also impact the Buffer Credit requirements for a solution.

Calculating Buffer Credit Requirements

$$\text{BBC Requirement} = \frac{\text{Transmission Rate}_{FC}}{\text{Average Frame Size}} \times \text{RTT}_{\text{Link \& Equipment}}$$

Transmission Rate$_{FC}$

Interface rate at which the link is provisioned [bits/second]

Protocol	Transmission Rate [Gbps]	Transmission Rate [bps]
1G Fibre Channel	1.0625	1062500000
2G Fibre Channel	2.125	2125000000
4G Fibre Channel	4.250	4250000000

Average Frame Size

Typical frame size provisioned [bits]

Maximum frame size is 2148 bytes; typical frame size is 1074 bytes

Note: frame sizes are a multiple of 4 bytes
Note: 8 bits in a byte—the equation must use a value in bits
Note: Fibre Channel leverages 8B/10B encoding. The average frame size used for the buffer to buffer credit calculation must be converted.

$$\text{8B/10B encoding} = \text{Avg. frame size [bits]} \times \frac{10}{8}$$

RTT$_{\text{Link \& Equipment}}$

Round Trip Physical link and equipment latency [seconds]

C_{Fiber} = *Speed of Light in Fiber = 203940.488 km/s*

WAN equipment latency can be ascertained from product technical specification. Ensure that equipment latency for the calculation is for a round trip.

$$\text{RTT}_{\text{Link}} = \frac{(\text{Distance}_{\text{Source to Target}} + \text{Distance}_{\text{Target to Source}})}{C_{\text{Fiber}}}$$

Note: There is typically a round trip latency command within most Fibre Channel switch or director command line interface (CLI) dictionaries that will provide cumulative latency for both the physical fiber and WAN equipment. The command will provide an approximation for the cumulative time delay for a frame to get from one port to another, the time for the other port to forward the frame, and the time for the frame to propagate back to the originator of the frame

Conclusion

Storage extension of Fibre Channel (and FICON) is critical, based on the importance of the data being transferred for Business Continuity and Disaster Recovery (BCDR) operations. Protocol transmission rate advancements, typical Fibre Channel frame sizes, and WAN equipment latency all alter the number of "frames in flight". All factors dictate an increased focus and proper architecting with respect to the Buffer-to-Buffer Credit (BBC) Flow Control mechanism as it is a critical factor that drives link performance.

Fibre Channel-capable BTI Client Service Modules Overview

BTI's transponder and muxponder client service modules provide support for Fibre Channel.

	1G DTPR	2.5G DTPR	4G DTPR	8-port MXP	10-port MXP
Line Rate	1 Gbps	2.5 Gbps	4 Gbps	2.5 Gbps/5 Gbps	10 Gbps
Client Ports	2	2	2	8	10
Line Ports	2	2	2	2	2
Client: Line	1:1	1:1	1:1	8:1/8:2	10:1
Maximum Client Signals Supported					
1G FC	2	2	2	7	9
2G FC		2	2	3	4
4G FC			2		2

2009-BTI-WP0055.1