



メインフレーム

FICON エクステンションの新しい標準規格の提唱 - Brocade FICON エミュレーションとテープ・パイプラインング -

Brocade エクステンション・プロダクトが持つエミュレーションとテープ・パイプラインング技術は、FICON テープと FICON IBM z/OS Global Mirror(XRC:eXtended Remote Copy)アプリケーションにおいて、実質的に距離無制限のエクステンションを提供します。本文では FICON エクステンション・アプリケーションにおける、これらの技術の内容と利点を詳しくご説明いたします。

BROCADE

目次

はじめに.....	3
FICON プロトコル効果の背景.....	3
FICON プロトコル上での長距離ネットワークの影響.....	3
IBM EXTENDED REMOTE COPY (XRC) エミュレーション.....	4
XRC 環境における BROCADE エミュレーション機能.....	5
FICON テープ・アクセスのための BROCADE エミュレーション・プロセス.....	5
デバイス・エミュレーションによる性能向上.....	7
BROCADE がサポートする FICON テープの長距離接続.....	8
まとめ.....	8

はじめに

Brocade® 7500 エクステンション・スイッチと、Brocade 48000 ダイレクタ/DCX バックボーン用の FR4-18i エクステンション・ブレードは、これまでの距離制限を克服した Brocade 初の FICON エクステンション製品であり、Fibre Channel over Internet Protocol (FCIP) 上で FICON Inter-Switch Link (ISL) エクステンションを提供します。FCIP による FICON ISL は、ネットワークバンド幅や遅延の影響などの考慮点があるものの、FICON プロトコルとアプリケーションの要件にあった非常に効果的なソリューションです。Brocade の次の目標は、現在 ISL エクステンションの課題となっている、ネットワークバンド幅や遅延の影響を克服し FICON エクステンションを可能とするための FICON デバイス・エミュレーションをプロダクト・ラインに追加することでした。

FICON デバイス・エミュレーションと Read/Write テープ・パイプラインの技術は、もともと Brocade USD-X と Brocade Edge M3000 エクステンション・プロダクト(旧型名: UltraNet Storage Director-eXtended/UltraNet Edge Storage Router) で提供されていたものです。これらの技術は FICON テープや IBM の eXtended Remote Copy (XRC) と呼ばれる一般的な汎用機におけるディスク・ミラーリング・ソリューション (XRC は HDS や EMC からライセンス販売されています) のエクステンションにおいて、実質的な距離の無制限化を実現するものです。Brocade USD-X と M3000 は FICON エミュレーションとパイプライン機能を提供することで、FICON 長距離エクステンションでの業界標準を作り上げ、世界中の汎用機環境において数千の実績を誇っています。

Brocade はこれらの技術を Brocade 7500/FR4-18i に実装することで、FICON エクステンション性能の向上を示すとともに、FICON エクステンションの可能性を拡大します。

FICON プロトコル効果の背景

標準 FICON は、これまでの ESCON やパラレル・チャネル・ブロック・マルチプレクサと比べ、はるかに有効なプロトコルです。FICON プロトコルは、これまでのデバイス・アクセス方法や I/O ドライバをサポートするために必要な複数のエンド・トゥ・エンド・プロトコルのやりとりを、縮小または削減するようにデザインされています。それによって、FICON は累積的遅延が増大するこれまでのプロトコルよりも、さらに効率的なプロトコルを提供できるのです。多くのアプリケーションにおいて、透過的な FICON フレームのやりとりは、顕著で優れた総合的性能を提供できると予想できます。(但しクリティカルな長距離エクステンション・アプリケーション(数百から数千 km) においてはその限りではありません。) さらに、FICON は接続性におけるはるかに大きな柔軟性(チャネルあたり最大 64,000 デバイス)と複数同時データ転送(最大 32 同時オペレーション)、そして完全相互のチャネルオペレーション(複数同時の Read/Write)を提供します。

FICON は Fibre Channel FC-2 プロトコル・レイヤー(1/2/4/8 Gbit/sec)上にマッピングされます。Fibre Channel スタandardにおいて、FICON は"SB-2/3"(接続される I/O デバイスのための IBM シングル・バイト・コマンド・アーキテクチャの総称)という名前で Level-4 レイヤーとして定義されています。"FICON"と"SB-2/3"は同義語であり、共にコネクションレスのポイント・トゥ・ポイントもしくはスイッチング、コネクションレスのポイント・トゥ・ポイント Fibre Channel トポロジーとして提供されます。

FICON プロトコル上での長距離ネットワークの影響

インフォメーション・ユニット(IU)ペーシングは SB-2 level-4 機能の一つであり、Channel Command Word (CCW) の数を制限しているもので、この数によって、コントロール・ユニットが生成する"コマンド・レスポンス"と呼ばれる Ack を待たずして行うことができる Write/Read の数が決まります。フロー制御自体は FC-PH レベルのバッファ・トゥ・バッファ・クレジット機能によるため、IU ペーシングは厳密にはフロー制御メカニズムではありません。むしろ IU ペーシングは、他の同時に発生した I/O が非常に大きなデータ転送によって Fibre Channel 環境を独占してしまうことから防止するためのメカニズムと言えるでしょう。

実質的には、IU ペーシングは、複数の競合するチャンネル・プログラムのための負荷分散もしくは”公正なアクセス”を提供するものです。この機能は負荷の高いチャンネル上の I/O 応答時間をさらに伸ばすものであって、長距離接続環境においては、逆に効率を悪くする結果となります。遅延が大きくなればなるほど、チャンネルや WAN 上の休止時間が増えます。休止時間は、予期されるコマンド応答の遅れによって、本来性能最適化を保證するための CCW の実行を禁止するペーシング・ウィンドウのポイントが増加することによって起こります。通常 1,2,4,8G bit/sec FICON 実装のペーシング・ウィンドウは 16 以上の IU を受け付けません。ペーシング・クレジットは特別なプロトコル・シーケンスのためにコントロール・ユニットの要求によって動的に調整することができますが、チャンネルは大きなペーシング・ウィンドウのための権限を持ったコントロール・ユニットには関連付けられません。

IBM eXTENDED Remote Copy (XRC) エミュレーション

汎用機向けの IBM Global Mirror は、z/OS 環境で利用できる、最も包括的なディスク複製アプリケーションの一つです。Global Mirror(旧名:XRC)は、ストレージ・コントローラとホスト・ベースのシステム・データ・ムーバー (SDM)からなるミラーリング・ソリューションです。システム・データ・ムーバー(SDM)はミラーリングをコントロールし、プライマリ・ストレージ上の全ての変更のあったデータをリモート・コピーするために読み込み、セカンダリ・ストレージにそれを書き込みます。ビジネス継続とディザスタ・リカバリ(BC/DR)のために、プライマリとセカンダリのストレージは通常地域的に分かれて置かれます。Brocade の独自の FICON エミュレーション機能は、プライマリ DASD を実質距離の制限無くエクステンドすることができ、地域的な配置において最大の柔軟性を提供します。

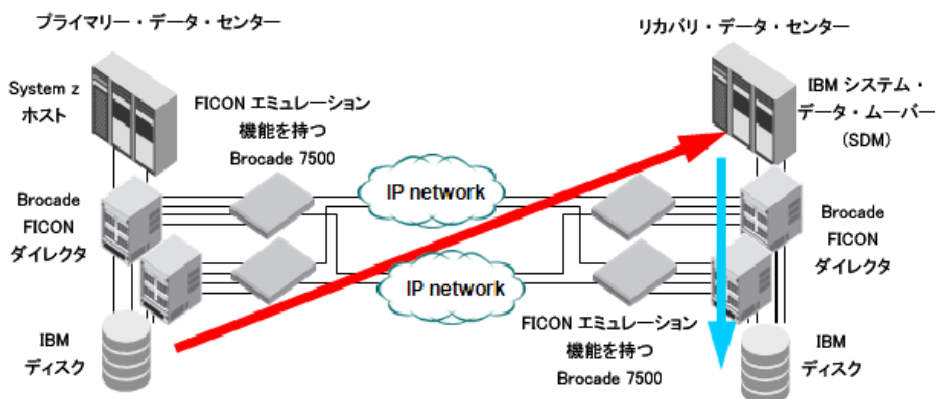


図 1. IBM eXtended Remote Copy (XRC) エミュレーション

System z 向けの IBM Global Mirror は、以下の二つの基本的タイプで動作します。

- Read Track Sets (RTS): 同期フェーズにおいて全てのトラックを読むために使われます。
- Read Record Sets (RRS): プライマリ DASD において変更のあったレコードを読むために使われます。

XRC アプリケーションでは、RRS チャンネル・プログラムの長い連結を効率よく実行することで高性能を得ることができます。これらのチャンネル・プログラムは最大で数百もの RRS コマンドを連結することができるので、長距離/高遅延の WAN においては、IU ペーシングが性能面でボトルネックになってしまいます。デバイスへ送られるコマンドの流れやコマンド・データ IU が、デバイスからのタイムリーに返ってこなくてはならないコマンド応答 IU の遅れによって、遮られてしまうからです。

新しい System z10 の機能の一つとして、IBM は長距離における XRC において、重大な性能影響を与えることのない IU ペーシングの強化を発表しました。この機能は新しい System z10 と最新の DS8000 ファームウェアの組み合わせによって可能となりますが、その他のシステムにおいては、依然 IU ペーシング制約による性能影響を受けることとなります。またこの新しい機能は FICON ISL エクステンションのみに有効です。IP WAN にて接続する場合には、これまでどおり FCIP チャンネル・エクステンション装置が必要です。

注: IU ペーシング強化機能の恩恵を受けるには、コントロール・ユニットと System z10 プロセッサをアップデートしなくてはなりません。Brocade の System z 向け Global Mirror(XRC)エミュレーションは、これまでのシステム環境(z10 プロセッサ以前の環境)においても、性能向上を提供することができます。

XRC 環境における BROCADE エミュレーション機能

Brocade の System z 向け Global Mirror(XRC)エミュレーション処理は、RRS チャネル・プログラムに対して有効であり、長距離接続での IU ペーシングや遅延の増加によって発生する休止状態を緩和するように努めます。遅延の原因には、間でのバッファリングによる遅延や、信号伝搬遅延、そして WAN バンド幅(1Gbps または 2Gbps)が Fibre Channel よりも低いバンド幅であることからくるバンド幅制約などがあります。更に IU ペーシング機能は、RRS オペレーションにおいて読み込まれる実際のデータサイズがどのくらいであるかを推測するには役に立たず、結果として IU 応答が少なくなることで、リンクの休止状態を増長する可能性があります。

Brocade の XRC エミュレーションは、前述の要因の全ての組み合わせにおいて、エミュレーションの無い状態と比べ性能向上を提供します。XRC エミュレーションの効果を適切に評価するためには、関連する全ての要因の正しい理解が必要です。情報収集の手段としては、システム・データ・ムーバー(SDM)のプライマリ・ボリュームの実チャネル・プログラムのトレースによるデータ分析も含まれます。更に、平均的な伝搬遅延やトータル平均バンド幅の情報については、WAN の回線プロバイダから得なくてはなりません。

WAN を介した場合、IU ペーシングがもとで起こる過度の遅延を緩和するために、XRC RRS エミュレーション・アルゴリズムが実装されています。RRS チャネル・プログラムは、独自の Define Subsystem Operation(DSO)コマンドのコマンド・データ IU によって他のチャネル・プログラムと区別されます。このコマンドに関連するデータ・バッファに含まれる情報には、後に続く RRS コマンド IU の全てのバッファ・サイズと番号が含まれています。この独自の DSO コマンドと関連するチャネル・プログラムのみがエミュレーション・プロセスをとおります。その他の全てのチャネル・プログラムとそれに関連する IU は、特別な処理なしで WAN リンク上を行き来します。このようにして、RRS オペレーションに含まれる Channel Command Words(CCW)に対して、性能向上のためにエミュレーションが行われます。その他のサポートされる機能については、ある程度 WAN 遅延の影響を受けます。

FICON テープ・アクセスのための BROCADE エミュレーション・プロセス

FICON プロトコルの利点として、かつての ESCON やパラレル・チャネルと比較して、テープ・オペレーションをサポートするのに必要なエンド・トゥ・エンドのやりとりの数を減らすことが挙げられます。しかしながら、多くの一般的なアクセス方法は、小さな read もしくは write CCW から成り立つ小さなチャネル・プログラムを生成し、通常オペレーティング・システムが供給する mode-set コマンドや、時には no-op コマンドの終了よりも優先します。このように、テープ・オペレーションをサポートする小さなチャネル・プログラムは、テープの read や write といったオペレーションに代表されるコマンド・データ・ステータスのやりとりをもとに、未だ順次的に処理されます。

これらのエンド・トゥ・エンドのやりとりは、FICON テープデバイスを接続するにあたり、納得のいく I/O アクセス時間や FICON が使う WAN の使用効率に対して重大な障害となることを危惧しなければなりません。テープ・オペレーションに必要なコマンド・データ・ステータスのやりとりに加え、IU ペーシングの影響によって WAN 経由で接続される FICON テープ・デバイスに対し、好ましくない遅延をさらに付け加えることも懸念され、特にテープ write オペレーションにおいては、IU ペーシング・アルゴリズムの中で書き込みデータ・フレームが重要な役割を握っていると言えるでしょう。Brocade のエミュレーションとパイプライン機能の組み合わせは、これらのやりとりからくる好ましくない遅延の影響を少なくし、WAN によって長距離接続される FICON テープや仮想テープ・デバイスに対し総合的な性能向上を提供します。

テープ・パイプラインは、あくまでも WAN 越しにホストとデバイスが接続される環境において I/O オペレーションの連続性を維持するものであり、通常の FICON における CCW や一つのコマンド・チェーン内のデータの

ストリーミングと混同してはいけません。通常、テープ・アクセス方式はリードの場合 end-of-file デリミタ(テープ・マーク)に達するまで、ライトの場合はデータセットがクローズされるか end-of-tape 状態が起こる(マルチ・ボリューム・ファイル)まで、データを順次読み込む(書き込む)ことを予想できます。エミュレーションは、連続する read/write オペレーションの性能を最適化するように設計されています。write オペレーションはテープ・デバイスの I/O オペレーションの大きな比率を占めることが予想できるため(アーカイブ目的)最初に取り組みます。

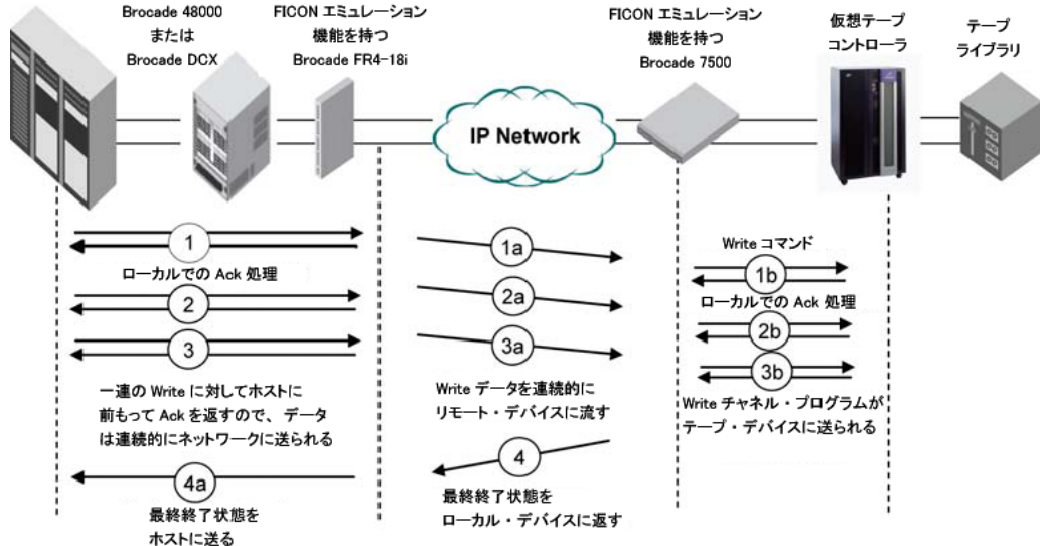


図 2. Brocade FICON Write テープ・パイプラインング

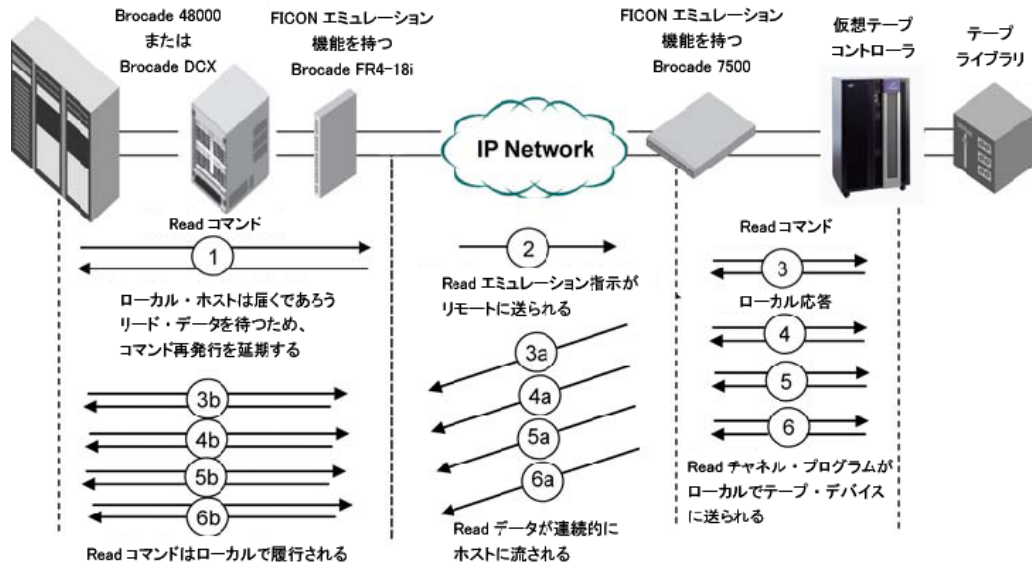


図 3. Brocade FICON Read テープ・パイプラインング

テープ・エミュレーション機能は、これよりも前に開発された XRC エミュレーション機能の拡張であり、既存の XRC エミュレーション基盤の多くを活用しています。分離されたコントロール・ユニットを動的に切り替えることができるなら、XRC エミュレーション機能とテープ・エミュレーション機能の両方を同時にサポートすることも可能です。XRC エミュレーションと同様に、テープ・エミュレーション技術は Fibre Channel フレーム・レベルに適用され、バッファやコマンド・レベルへの適用ではありません。更に XRC エミュレーション同様、テープ・アクセスにおいてエミュレーションを行うのは、テープ・アクセスの一部のテープ・オペレーションの性能に大きくかかわってくる mode-set や no-op を伴う標準の write と read のみです。他の全てのコントロール機能はエミュレーションされず、エミュレーション処理において、ホストに対してコントロール・ユニット/デバイスとして振舞ったり、デバイスに対してホストとして振舞ったりしません。

デバイス・エミュレーションによる性能向上

比較するにあたり”Shuttle mode”という言葉をしばしば使用します。”Shuttle mode”とは FICON をエクステンションするにあたり、全てのオペレーションが端から端まで流れ、それぞれのオペレーションの伝搬遅延によって性能に直接影響を与える動きのことを言います。

図 4 に示す FICON テープ性能グラフは、一つの IBM 仮想テープコントローラ(VTC)を Brocade 7500 エクステンション装置でエクステンションしたテスト結果です。ここでは 8 個の 32KB ブロック・サイズのテープ・ジョブでテストしています。Read/Write テープ・パイプラインと Shuttle Mode の性能を様々な遅延において比較しています。Read 性能は Write 性能に比べ低くなっていますが、これは VTC の制限からくるものです。しかしながら Shuttle Mode と比較すると、FICON エミュレーションによる性能向上は Read/Write とともに図られていることが分かります。

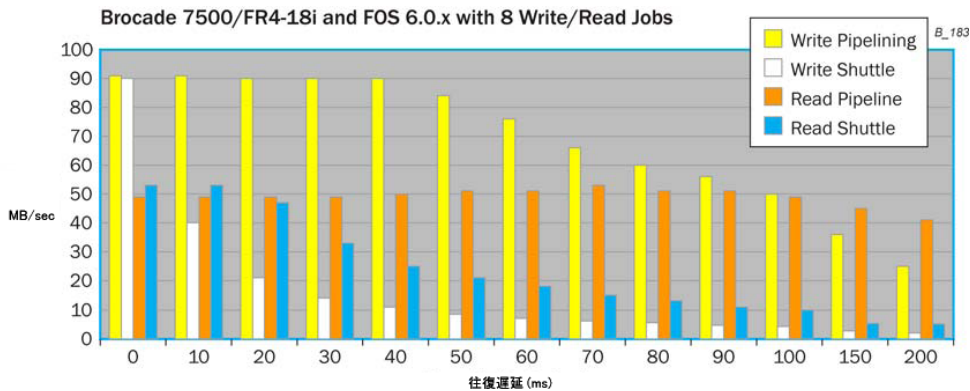


図 4. 一つの VTC 上で 8 個の Read/Write ジョブを実行した時の FICON テープ・パイプライン性能

図 5 に示すテスト結果は、四つの VTC を使い 16 個の同時テープ・ジョブを走らせたときのテープ・パイプライン性能です。リード・パイプラインにおいては 25% から 15 倍までの性能向上が長距離接続において確認できます。エミュレーションが無くても、VTC の数と同時テープ・ジョブを増やすことによって、WAN 回線を効率よく使用できるのが分かります。FICON Read エミュレーションでは、少ないテープ・ジョブ環境においても、WAN 回線の効率的な使用と高いスループットが得られることが確認できます。

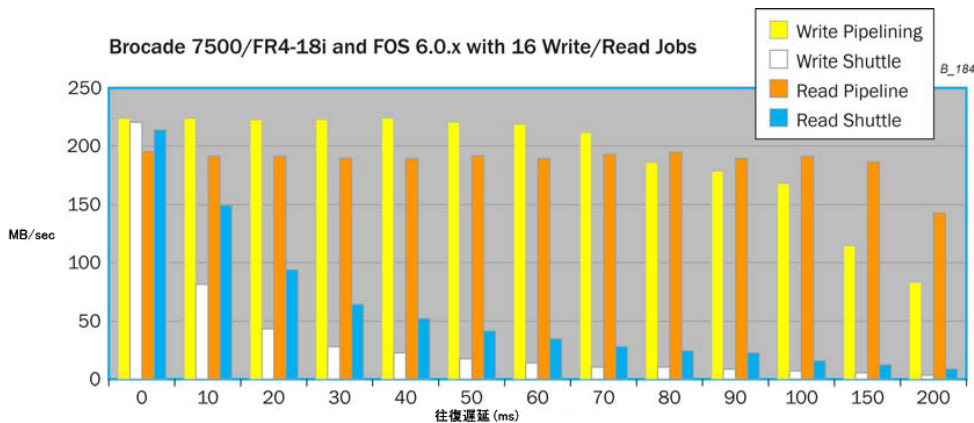


図 5. 四つの VTC 上で 16 個の Read/Write ジョブを実行した時の FICON テープ・パイプライン性能

BROCADE がサポートする FICON テープの長距離接続

Brocade の FICON 長距離接続の機能の中には、IBM の仮想テープ・サーバ(VTS)と Sun の仮想ストレージ・マネージャー(VSM)のエミュレーションサポートと、FICON テープ長距離接続のための Read/Write パイプラインが含まれます。幅広いコンフィグレーションが可能のため、ディザスタ・リカバリ・システムの構築において、FICON カスケードを含む様々なスイッチ環境の配置など、柔軟な対応が可能となっています。FICON テープ長距離接続構成には以下のものが含まれます。

- ホストとリモート・テープ・コントローラ間の長距離接続(フロントエンド長距離接続)
- ローカル・テープ・コントローラとリモート・テープ・ライブラリ間の長距離接続(バックエンド長距離接続)
- 高可用性のためのクラスタ間の長距離接続

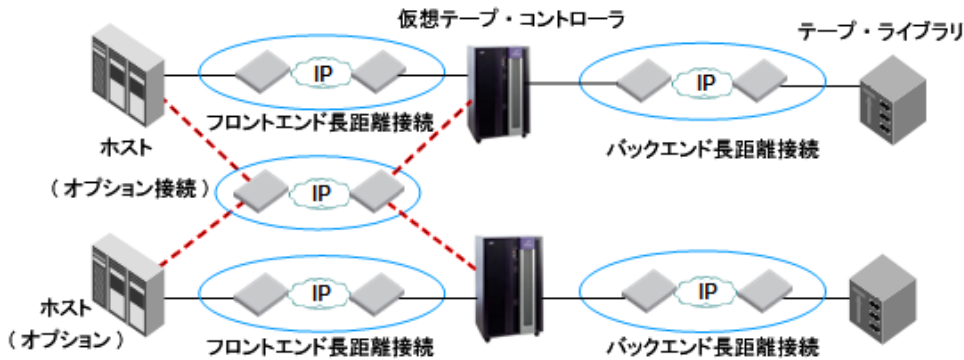


図 6. 構成可能なディザスタ・リカバリ・システムにおけるテープ環境

まとめ

FICON チャンネルは、IBM System z 向けに、接続距離、性能、拡張性ともに大幅に拡大する設計となっています。Brocade USD-X と Edge M3000 のエクステンション製品が提供する FICON エミュレーションとパイプライン機能は FICON の可能性をさらに拡大し、テープと IBM XRC を事実上距離無制限に接続することを可能とし、拡張性ばかりか WAN のバンド幅の実行効率を最大限に引き上げるなど、劇的にバックアップ/リカバリの性能を向上させました。

この優れた Brocade の技術は、Brocade の多くの汎用機カスタマーにおいて広くご使用いただいております。今、この FICON エミュレーションとパイプライン機能が、業界をリードする Brocade 7500 エクステンション・スイッチと Brocade 48000 ダイレクタ/DCX バックボーン向けの FR4-18i エクステンション・ブレードに搭載されました。Brocade の全てのエクステンション製品において、リモート・テープと XRC を事実上距離の制限を受けずに長距離接続することが可能となり、FICON の長距離接続の市場において、Brocade のリーダーシップが拡大するだけでなく、Brocade の汎用機カスタマーに対し、前例を見ない柔軟性と性能とデータ保護の向上をお届けできるようになりました。

© 2008 Brocade Communications Systems, Inc. All Rights Reserved.

Brocade, Fabric OS, File Lifecycle Manager, MyView, and StorageX are registered trademarks and the Brocade B-wing symbol, DCX, and SAN Health are trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. All other brands, products, or service names are or may be trademarks or service marks of, and are used to identify, products or services of their respective owners.

Notice: This document is for informational purposes only and does not set forth any warranty, expressed or implied, concerning any equipment, equipment feature, or service offered or to be offered by Brocade. Brocade reserves the right to make changes to this document at any time, without notice, and assumes no responsibility for its use. This informational document describes features that may not be currently available. Contact a Brocade sales office for information on feature and product availability. Export of technical data contained in this document may require an export license from the United States government.