特集論文

某金融機関向けIT-BCP対策におけるデータ同期方式

松本健太郎* 細川智洋*
森垣 努* 山路晃徳*
高篠智晴*

IT-BCP Data Synchronous Method for Certain Finance Institution

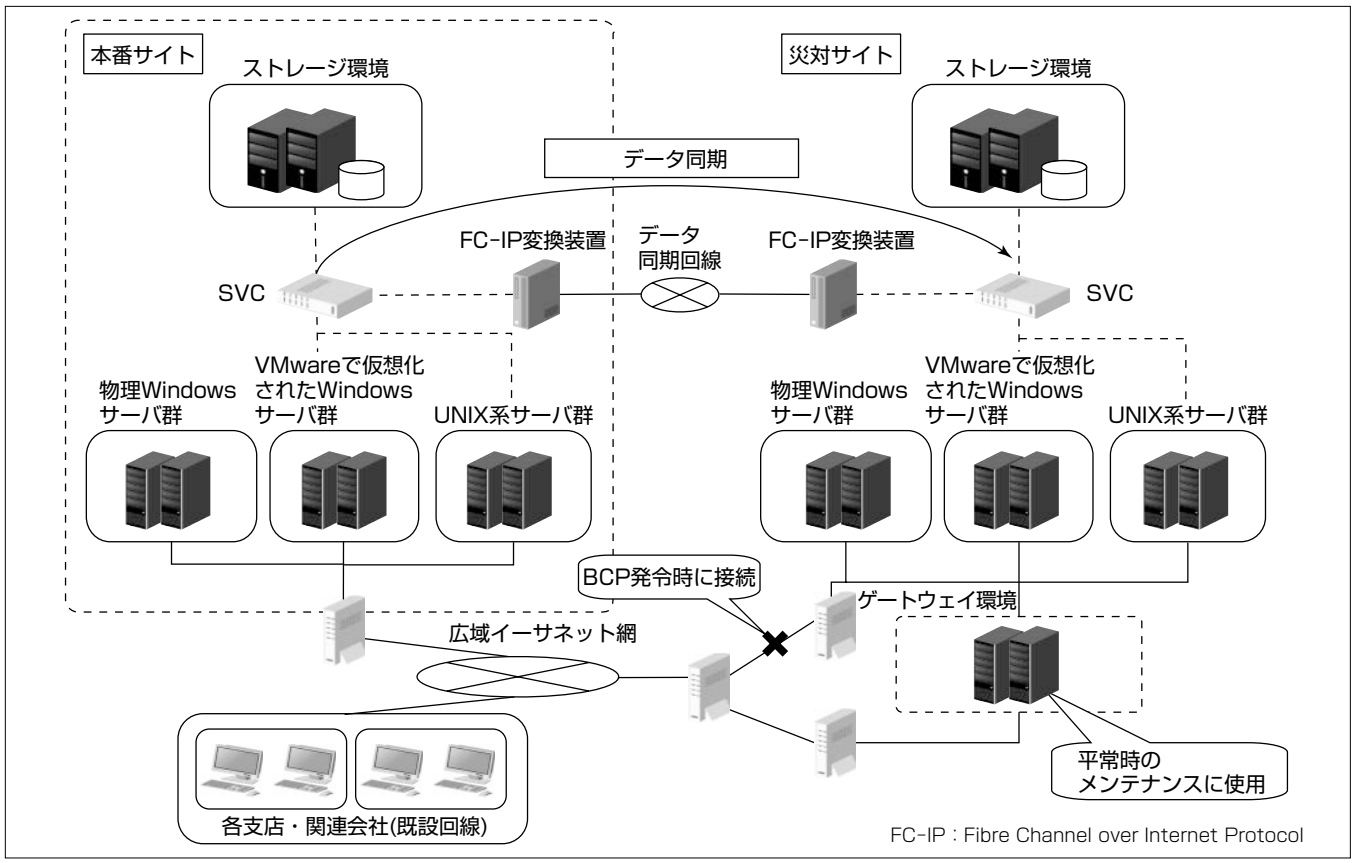
Kentarou Matsumoto, Tsutomu Morigaki, Tomoharu Takashino, Tomohiro Hosokawa, Akinori Yamaji

要 旨

東日本大震災以降，各企業で事業継続・災害対策が急務となっている。三菱電機インフォメーションシステムズ㈱(MDIS)の顧客である金融機関でも，一部業務に対するBCP(Business Continuity Planning)対策を実施しており，首都圏直下型地震の災害による電源供給断などのインフラ途絶，データセンター倒壊で，基幹システムの復旧に1か月以上を要する事態となった場合を想定し，基幹システムについても二重化を行うこととなった。全体の推進・構築は，基幹系・情報系のインフラ全般にかかわっているMDISが中心となって実施している。

対象システムは基幹システムと基幹システムの稼働に必要となる周辺システム計10システム(UNIX(注1)系サーバ：約15台，Windows(注2)系サーバ：約150台)で，グループ会社も含めたシステムを対象としている。エンドユーザーへの影響やアプリケーション改修による工数増をさけるため，BCPサイト(以下“災対サイト”という。)のサーバは，関東データセンター(以下“本番サイト”という。)のホスト名，IPアドレスなどの固有値をそのまま引き継ぐ構成とした。また，既存環境への変更を加えないように構築することとなりこの案件では，ストレージ・リソースの管理を一元化するストレージ仮想化システムIBM(注3) SAN Volume Controller(以下“SVC”という。)とVMware(注4) SRM(VMware vCenter(注4) Site Recovery Manager(注4))を利用した同期方式を採用して，多種多様なシステムが対象となる，各システム，サーバに合わせたRPO(Recovery Point Objective)の設定やデータ同期設計を実現した。

(注1) UNIXは，The Open Groupの登録商標である。
(注2) Windowsは，Microsoft Corp. の登録商標である。
(注3) IBM System Storage，IBMロゴ，ibm.comは，世界の多くの国で登録されたIBM Corp. の登録商標である。
(注4) VMware，vCenter，Site Recovery Managerは，VMware, Inc. の登録商標である。



基幹システムのデータ同期方式

IBM社のSVCを利用した非同期でデータ連携を行う。災害発生時にはネットワークの切替えでセンターの切替えとシステム復旧を行う。災対サイトの平常時のメンテナンスはゲートウェイ環境を経由して行う。

*三菱電機インフォメーションシステムズ㈱

1. まえがき

東日本大震災以降，MDISの顧客である金融機関でも，首都圏直下型地震の災害による電源供給断などのインフラ途絶，データセンター倒壊等を想定し，基幹システムの二重化を行うこととなった。MDISは，既存の基幹システム構築時，SIer(System Integrator)としてインフラ全般の設計・構築を担当し，また，この基幹システムの構築後から現在に至るまで，顧客のインフラ全般の運用・保守を委託されており，基幹システムについての運用ノウハウが多く蓄積されている。さらに，基幹系・情報系のシステムの運用・保守に加え，端末管理，ヘルプデスクまで広く顧客の業務に携わっていることから，この案件では，同じくSIerとして，他のベンダーをも取りまとめる立場で，全体の推進を行うこととなった。

このシステムの利用環境に関しては，既に顧客で一部システムの災対サイトとして利用している西日本のデータセンターを使用することとなっている。このような背景の中，このプロジェクト推進に当たり，最も苦労したポイントは次の4点である。

⑴ 多種多様なシステムに合わせたRPOの設定
⑵ 既存環境への変更を抑えた構築方針の策定
⑶ データ更新量と必要回線速度の見極め
⑷ 基盤に合わせた統一的なデータ同期方式の設計

本稿では，これらの課題を踏まえたITインフラBCP対策システムの構築について述べる。

2. 要件定義

2.1 RPOとRTO

IT-BCPの構築では，目標とする復旧地点であるRPOの設定と復旧にかかる時間の目標値であるRTO(Recovery Time Objective)の設定が最も重要な部分であり，実現性を評価しながら慎重に設計を進めていく必要がある。RTOについては顧客の業務要件を考慮したうえで，対象システムすべてに対して共通で5時間と設定した。一方，RPOについては，この案件の対象が基幹系システムとその他周辺システムで，BCPの対象が複数システムであることを考慮して設定を行う必要があったため，業務時間帯によって異なる値を設定することとした。例えば，オンライン中は1時間前の状態に復旧させるが，夜間のオンライン時間外は，対象サーバのデータ更新頻度などによって要件が異なるRPO値を設定した。なぜなら，夜間処理におけるデータ更新量が非常に多く，オンラインと同じRPOを設定した場合，必要なデータを時間内に反映できなくなってしまうおそれがあるためである。

次に，具体的に実施した“多種多様なシステムに合わせたRPOの設定”について述べる。

2.1.1 システム単位別のRPO

RPOの定義について，システムを構成するサーバによって，データの更新頻度が異なる点に着目して設計を行った。対象サーバは，更新頻度・周期別に次の3種類に分類した。

⑴ オンライン，夜間処理中にデータ更新が行われるもの
⑵ 夜間処理中にだけデータ更新が行われるもの
⑶ 不定期にデータ更新が行われるもの

⑴でオンライン中に被災した場合，1時間前の状態に復旧することとするが，⑴の夜間処理中又は，⑵では，夜間処理で更新されるデータ量が非常に多いため，夜間処理中のデータ同期は行わず，夜間処理開始前，夜間処理完了後にRPOを設定した。⑶は主に管理系のシステムで不定期に手動での更新が発生するものであるため，更新時に手動でのデータ反映を行う運用を検討したが，手順漏れによってデータ反映が行えない場合のリスクを考慮し，1日1回データ同期を行うこととし，RPOを1日前と設定した。

2.1.2 被災時間帯別のRPO

オンライン中に被災した場合は1時間前の状態へ復旧することになっていたが，夜間処理時間帯に切替えが行われた場合は，災対サイトでの復旧後，夜間処理を始めからやり直す運用としたため，夜間処理前の状態に復旧することにした。また，オンライン開始時点で被災した場合は，夜間処理で更新された状態まで復旧することを要件とした(**図1**)。

2.2 既存環境への変更を抑えた災対サイト構築方針

業務アプリケーションについては，BCP対象となり得る

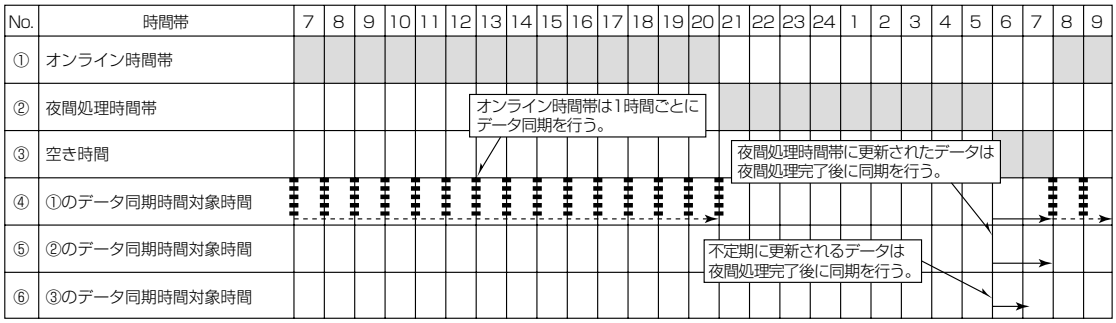


図1．被災時間帯別のRPO

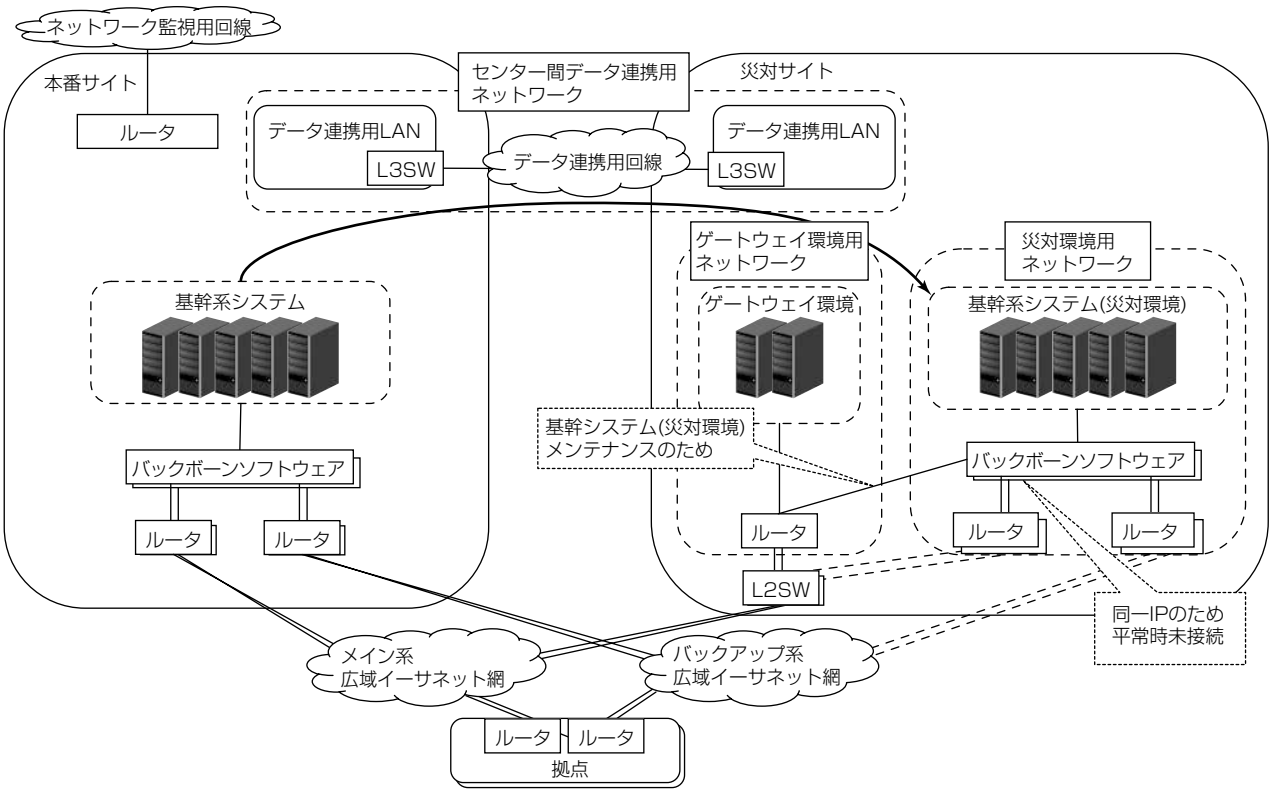


図 2．災対サイト構築方式

(1)　各サーバのオンライン時間帯の最大更新量の調査

会社	対象システム	対象時間帯	データ更新量(GB)
A社	システム#1	オンライン	57
	システム#2	オンライン	27
B社	システム#1	オンライン	24
	システム#2	オンライン	4.75
C社	システム#1	オンライン	87.7
合計			200.45

(2)　1時間当たりに転送可能なデータ量の算出

項目	値	計算式
1時間間隔で同期を行うデータ量	約200GB	−
1セッション当たりの転送量	17Mbps	65,535byte(ウィンドウサイズ)×8/0.03(遅延)
1セッション当たりの転送量(1時間当たり)	6GB/h	17Mbps(1セッション当たりの転送量)/8×3,600
必要セッション数	34本	200GB(1時間間隔で同期を行うデータ量)/6GB(1セッション当たりの転送量(1時間あたり))
1時間当たりの転送可能容量(データ非圧縮)	204GB/h	6GB(1セッション当たりの転送量(1時間あたり))×34(必要セッション数)
圧縮率	0.5	−
1時間当たりの転送可能容量(データ圧縮)	408GB/h	204GB(1時間後当たりの転送可能容量(データ非圧縮))/0.5(圧縮率)
必要回線帯域	578Mbps	17Mbps(1セッション当たりの転送量)×34(必要セッション数)

図 3．サーバごとのデータ更新量の必要回線帯域

範疇（はんちゅう）が広く，また複数他社のベンダーで構築・管理されている。

これによって，本番環境への変更を加えた場合，エンドユーザーへの影響を極小化する意味でも，“既存システムへの変更は行わない”という方針は顧客からの重要な要求であった。そのため，災対サイトに構築するサーバのホスト名，IPアドレス等の固有値を変更した場合の，アプリケーションへの影響範囲を見極めることが重要であったが，そもそもIPアドレスを変更してしまうことで，エンドユーザーの利用にも影響が出てしまうおそれがあったため，災対サイトに構築するシステムは，ホスト名，IPアドレスともに変更せず，本番サイトのサーバと全く同じ構成のサーバを構築することとした。その際，同一ネットワーク上でIPアドレスが重複することによって問題が発生しないように，平常運用時は災対サイトに設置するルータのネットワークケーブルを抜線しておき，BCP発動時に接続する運用とした。災対サイトでの訓練，平常時のメンテナンスを考慮し，一部対処済みBCPで構築された環境のネットワークを流用し，ゲートウェイ環境を構築することとした（**図 2**）。

2.3　データ更新量とデータ同期用回線

この案件で構築するシステムはデータ更新量が非常に多く，データ同期を行うに当たり，回線帯域不足が懸念された。そこでRPOを満足する仕組みを構築するため，データ同期のシミュレーションを行った。対象は，更新量の多いUNIX系システムとし，オンライン中に更新されたデータを規定時間内に送信完了できるかについて，次によってシミュレーションを実施した（**図 3**）。

(1)　各サーバのオンライン時間帯の最大更新量の調査

⑵　1時間当たりに転送可能なデータ量の算出

この結果から，1Gbpsのデータ同期用回線の6割を割り当てれば効率的に同期を行えることが確認できた。

3. IT-BCPの同期設計

3.1　ストレージ構成とデータ同期方式

先に述べた通り，災対サイトとのデータ同期方式の検討に当たり，次の2つの要件を満足し，かつ統一的な同期方式を採用する必要があった。

⑴　システム，サーバごとに異なるRPOを満たす

⑵　原則，既存環境への変更を加えない

そこで，各サーバが利用しているストレージの機能によるデータ同期方式とSVCを利用したデータ同期方式についての検討を行ったが，使用しているストレージに左右されないというメリットを考慮し，SVCを利用したデータ同期方式を採用した。なお，SVCの管理下にないサーバについては，運用，切替え時の手順によってデータの最新化が行えるよう対応した。

SVCを利用したデータ同期の実装に当たり，対象サーバを次の3つに分け，サーバごとに災対サイトへ反映が必要なデータを洗い出し，それらの反映方法について検討を行った。

⑴　UNIX系サーバ

⑵　VMwareで仮想化されているWindows系サーバ

⑶　物理Windows系サーバ

以降，このグループごとに行った同期対象データの洗い出しと同期方式について述べる。

3.2　UNIX系サーバでのデータ同期方式

災対サイトのUNIX系システムにおける同期対象データの洗い出し，データ同期方式の検討内容を述べる。

この案件の対象となっているUNIX系サーバはローカルディスクとSVC管理下のストレージで構成されている。SVC管理下のストレージに保存されているデータは，SVCの機能によって災対サイト側への反映が行えるため，それ以外のローカルディスクに保存されているデータについてだけ，同期要否と同期方法の検討を行った。

ローカルディスクは主にシステム領域とアプリケーションのインストールディレクトリとして利用されており，各データの更新頻度は低いものがほとんどであった。またUNIX系サーバは切替え作業時間短縮のため，通常時から起動しておく方針としていた。そこで，UNIX系サーバのローカルディスクに保存されているデータについては，ゲートウェイ環境経由で災対サイトの対象サーバにアクセスし，手動で更新を反映する運用とした。

UNIX系システムでのデータ同期方式は，データ書き込みの確実性からリモートコピー方式の中でもMetro Mirrorが最有力候補であった。しかし，その仕組みから，遠隔地への同期を行った場合に性能劣化が懸念され，本番環境の業務へ影響が出る恐れがあること，また，複数のサーバの同期を同一タイミングで行うことができないといったデメリットがあった。そこで，本番サイトの対象ボリュームからFlashCopy(注5)にて取得したボリュームについてMetro Mirrorを利用して同期を行うことで本番環境業務の性能劣化を回避し，複数サーバの同期タイミングをそろえたデータ同期が行える仕組みを採用した。また，同期対象データが非常に大きいため，Metro Mirror間の同期にFC-IP(Fibre Channel Over IP)変換装置を採用し，同期データの圧縮と合わせて伝送速度の向上を図った(**図4**)。

(注5)　FlashCopy，IBMロゴ，ibm.comは，世界の多くの国で登録されたIBM Corp. の登録商標である。

3.3　Windows系サーバでのデータ同期方式

この案件で構築するWindows系サーバには，VMwareで仮想化されているWindows系サーバ，物理Windows系サーバの2種類があった。その中で災対サイトへデータの反映が必要なものは，VMwareにて仮想化されているWindows系サーバであり，かつSVC管理下のストレージにデータが保存されているものだけであった。統一的な復旧方式を採用し復旧手順を簡易化するため，VMwareSRMによる復旧方式を採用した。SVCの機能によるデータ

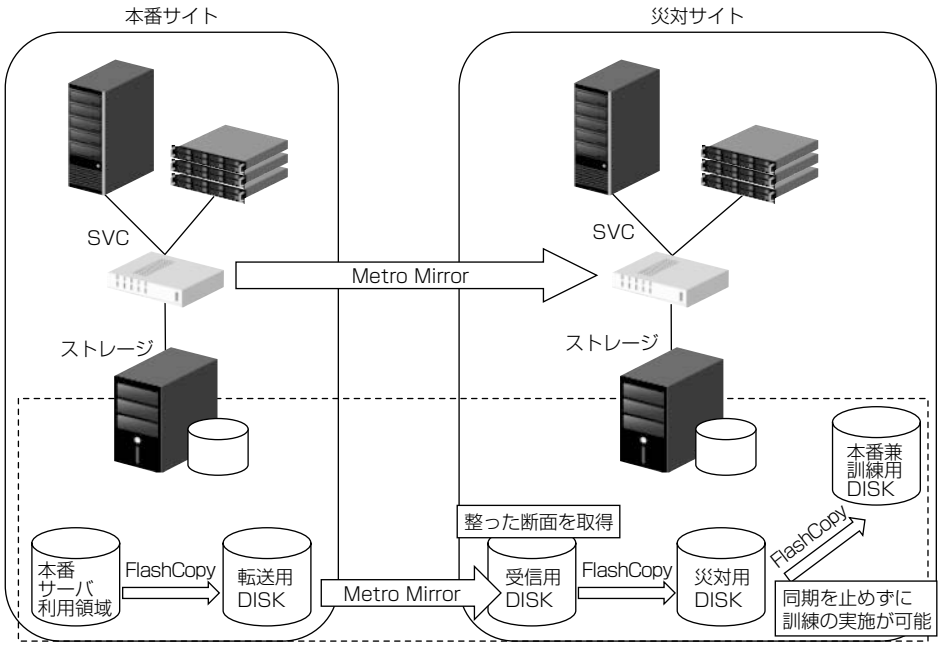
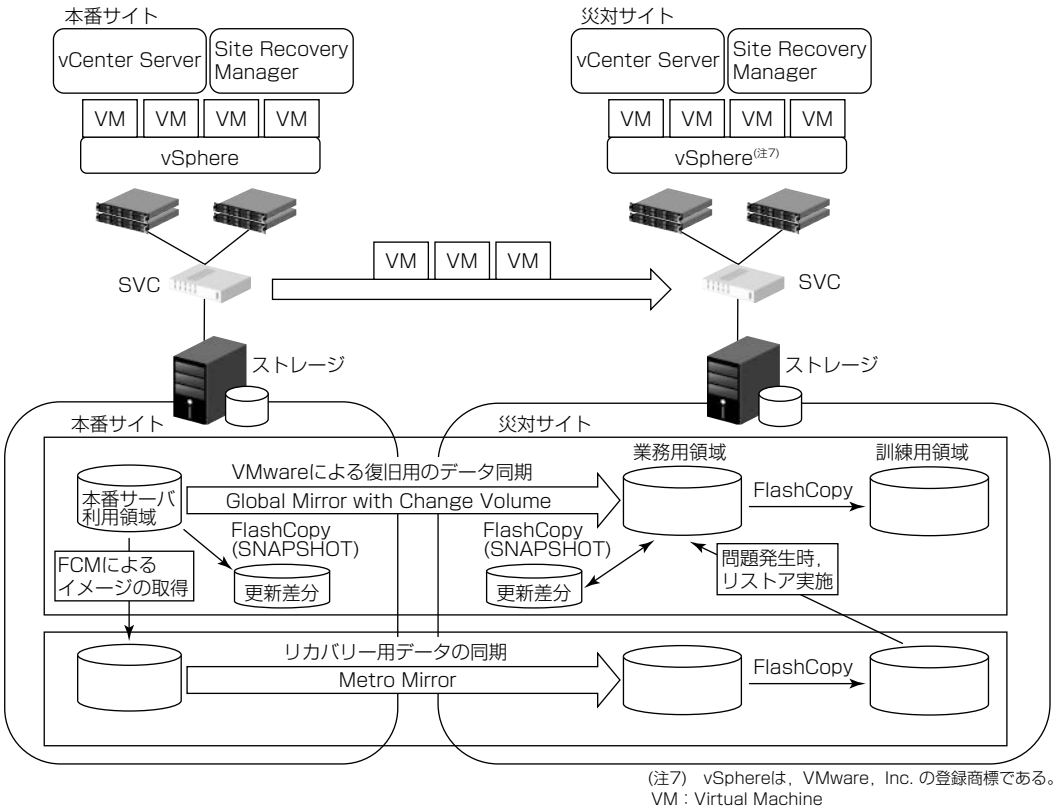


図4．UNIX系サーバのデータ同期方式



(注7) vSphereは，VMware，Inc. の登録商標である。
VM：Virtual Machine

図5．VMwareで仮想化されているデータ同期方式

同期方式を採用し，VMware SRMを利用する場合は，“Global Mirror with ChangeVolume”というデータ同期方式で同期環境を構成する必要があった。

しかし，このデータ同期方式ではクラッシュ・コンシステンシーレベルの整合性しか担保されておらず，アプリケーションの稼働を保障することができなかった。そこで，VMware SRMで復旧できなかった場合の対策として，このデータ同期方式の同期周期と合わせて，オンライン中でも対象サーバの静止状態のデータを取得できるFCM（IBM Tivoli(注6) Storage FlashCopy Manager for VMware）を本番サイトへ新規導入した。FCMを用いたバックアップを災対サイトへ送信しておき，問題発生等には，そのバックアップからリストアする運用とした（**図5**）。

（注6） Tivoli，IBMロゴ，ibm.comは，世界の多くの国で登録されたIBM Corp. の登録商標である。

4. むすび

災対サイトを構築する場合，“顧客の業務継続性維持”“最低限，保持すべきリソースは何か”という視点から顧客の業務運用を考慮した上でシステム要件を決定していくプロセスが必要となる。また，システム的にはできるだけ既存環境に変更を加えず，さらに既存システムへの影響を最小限にとどめるように設計・構築を行うことが重要となる。この事例では次について述べた。

(1) 既存システムへの影響を最低限に抑えた災対サイト構築方針の検討
(2) 既存システムに適したデータ同期方式の検討
(3) データ転送シミュレーション

データ同期方式の検討では，原則既存環境を変更できないという制約があったが，基幹システムが利用しているSVCを利用したデータ同期方式を採用することで，顧客要件を満足させることができた。さらに災対サイトに構築するサーバを本番サイトと同じものとすることで運用負荷，運用コストの低減も実現することができた。

災対サイトは，今後更に一般化し，より低コスト，短期間で構築できるようにするためのノウハウの蓄積，システム構築プロセスの汎用化が求められるため，今回検討した内容がおおいに役に立つと考える。また，このプロジェクトのスコープは災対サイトの構築であったが，バックアップや準本番環境としての利用など，リソースの有効利用という活用方法も考えられ，今後は，これらも踏まえた提案，構築を行っていく。BCP策定時に最も重要視されているのは運用コストである。今回運用負荷，運用コストの低減も加味した構築を行ったが，運用保守を任されているベンダーとして，システム構築だけにとどまらず，運用改善，コストの低減を図っていく。

今回はデータセンターの被災というリスクに対する対策を行ったが，セキュリティリスクなど，世の中にはほかにも多くのリスクが存在する。今後も，運用・保守を通して更なる改善を行い，顧客システムをより良いものにしていくとともに，被災以外のリスクを踏まえたBCP対策ソリューションの提供を行っていく所存である。