

# Open-**E** DATA STORAGE SERVER

を使用したレプリケーションと  
iSCSI フェール・オーバー



## 1 はじめに

### 1.1 Open-E DSS とレプリケーション

Open-E の Data Storage Server(DSS)はエンタープライズクラスのレプリケーション機能が搭載されており、3種類のレプリケーションをサポートしています。

それぞれボリューム・レプリケーション、データ・レプリケーションと iSCSI の自動フェールオーバーです。DSS は iSCSI、NAS、ファイバチャネルの各ボリュームを扱えるためシングルモジュールで様々な冗長構成を取ることが可能です。低コストで効果的なストレージ管理をご希望のすべてのお客様にお勧めします。

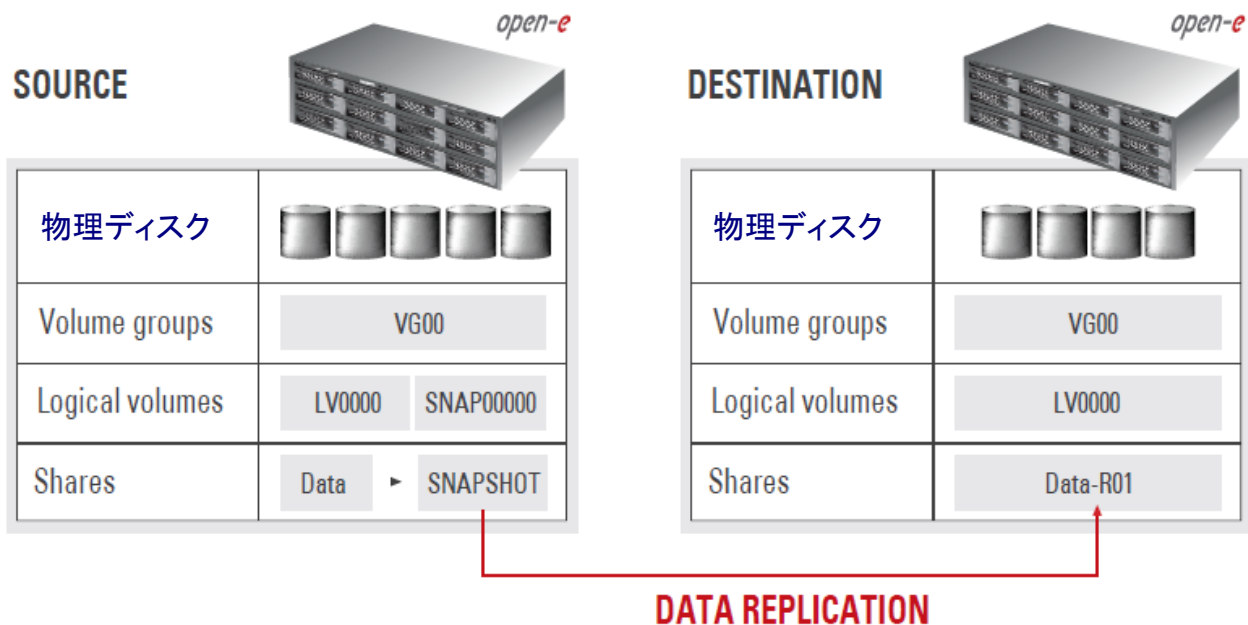
Open-E でサポートするレプリケーションの一覧は以下の通りです；

	モード		レプリケーション元/先			I/O		ボリューム		
	同期	非同期	システム内	LAN	WAN	ファイル・ベース	ブロック・ベース	NAS	iSCSI	FC
データ・レプリケーション		✓	✓	✓	✓	✓		✓		
ボリューム・レプリケーション	✓			✓			✓	✓	✓	✓
フェールオーバー機能付レプリケーション	✓			✓			✓		✓	

**データ・レプリケーションの使用例：** 効率的で素早いデータのバックアップ及びレストア

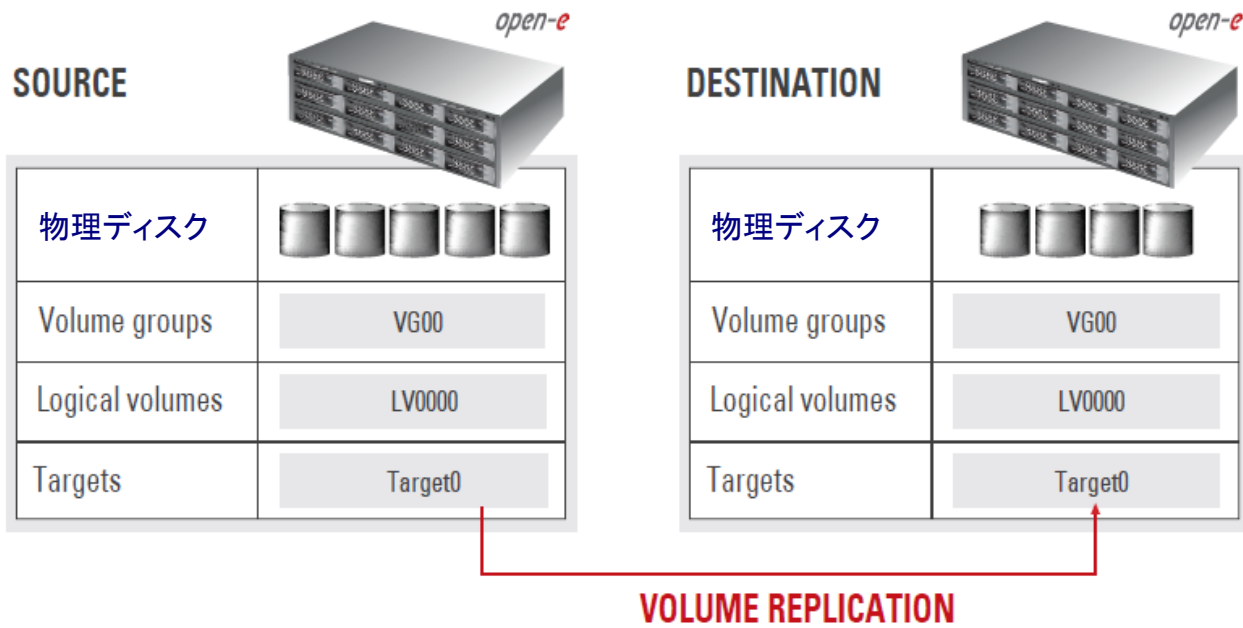
スナップショットを使用した非同期のデータ複製のため効率的に運用可能です。クロス・データ・レプリケーションが可能のためバックアップサーバーのコストを低く抑えながらデータ保護を行えます。

また、ディスクシステムをテープシステムの代替のバックアップシステムとして使用するとレストア速度が劇的に速くなります。



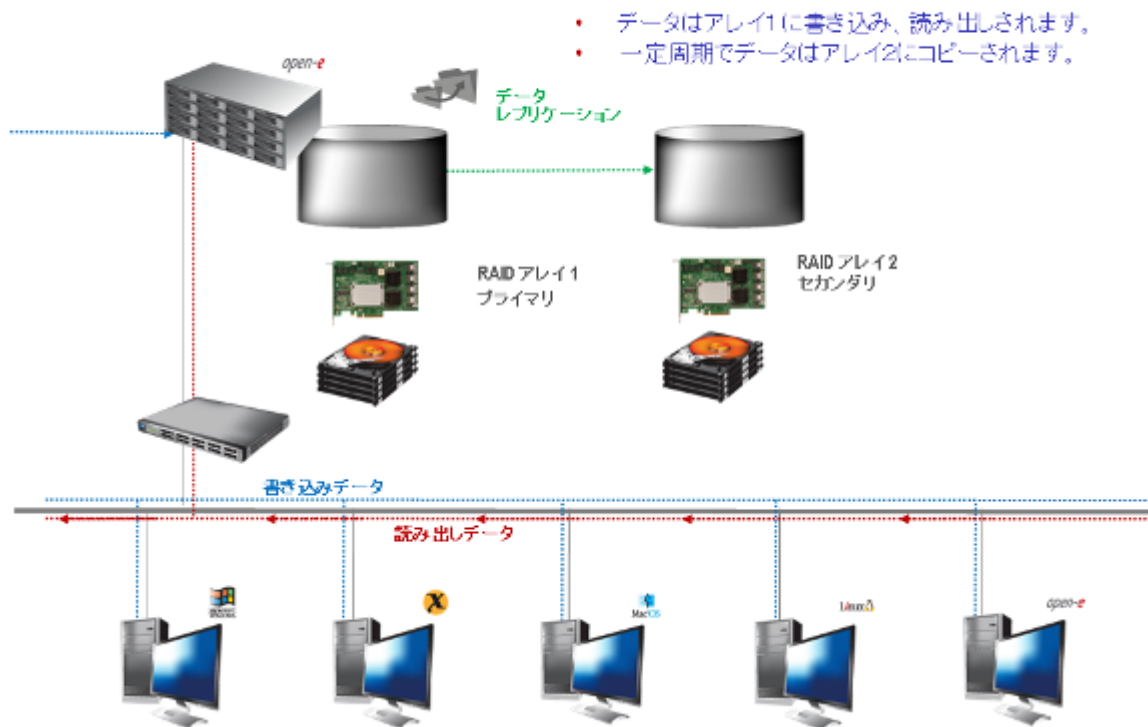
ボリューム・レプリケーションの使用例： デザスタ・リカバリの最適ソリューション

DSS は iSCSI、NAS、FC の各ボリュームを同期でレプリケーションしますので貴社の大事なデータを保護するのに最適です。プライマリ・サーバーに障害が発生した場合は、数マウスクリックで引き続き運用可能です。また、iSCSI ボリュームでは自動フェールオーバーをサポートしていますのでシームレスな運用が可能です。



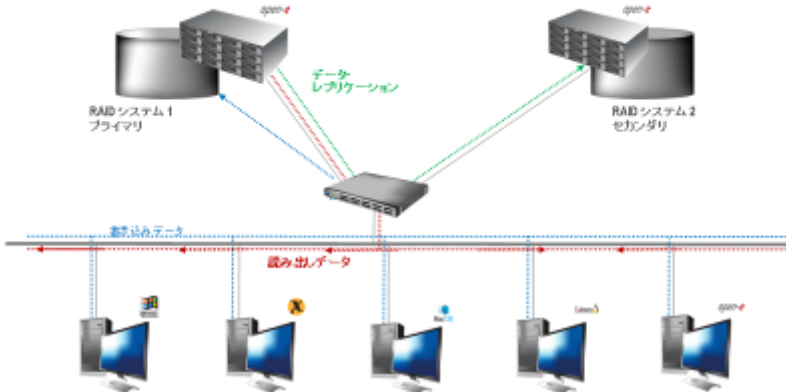
## 1.2 データ・レプリケーションの様々な形態

### 同じシステム内のデータ・レプリケーション



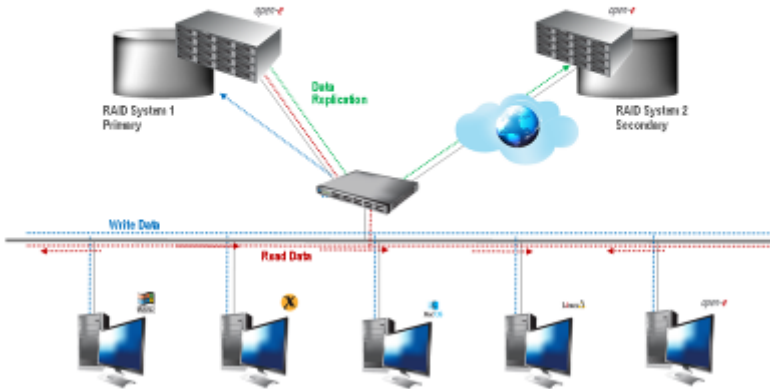
### LAN内の別システムへのデータ・レプリケーション

- データはシステム1に書き込まれ、読み出されます。
- LANを通じてシステム1の内容がシステム2にコピーされます。

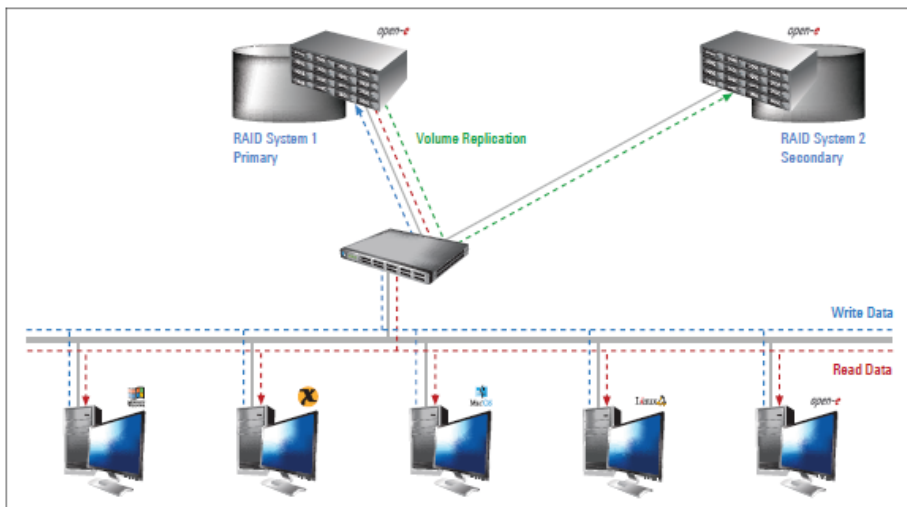


### WAN経由の非同期データレプリケーション

- データはシステム1で読み、書きされる。
- インターネットを通じて非同期でデータはシステム2にレプリケーションされる。



### 1.3 ボリューム・レプリケーション 図



## 2 レプリケーション

### 2.1 レプリケーション

ボリューム・レプリケーションはソフトウェア RAID の RAID1と同じミラーリング機能をイーサネットを通じて実行します。そのため、完全同期でデータの複製が可能です。ただし、ネットワーク負荷がかかり、システムのパフォーマンスが落ちるという欠点があります。OPEN-E ではレプリケーション専用で別ネットワークを組むことをお勧めしています。iSCSI ボリュームの場合、フェールオーバー機能もサポートしています。iSCSI フェールオーバー機能の詳細については次章をご参照ください。

データ・レプリケーションは rsync+snapshot 技術を使用し、スナップショットで固定したデータを非同期でレプリケーション先に複製します。LAN 外でも使用可能ですが、NAS ボリュームのみのサポートとなります。

**注意！！ GUI の作成時期の関係でレプリケーションと複製、レプリケーション先と実行先、コピー先など同じことを意味するのに異なる用語が使用されている場合があります。ご注意ください。**

**注意！！ レプリケーションでは両方のサーバーの時刻設定が同じである必要があります。必ず、時刻設定を同じに設定してください。**

### 2.2 データ・レプリケーションの設定方法

データ・レプリケーションの設定方法はいろいろなタブに移動する必要があるため、ここで設定方法をご説明します。データ・レプリケーションはバックアップの一方法で、NAS ボリュームでご使用いただきます。

**注意！！ レプリケーション元と先はかならず同じ容量に設定してください。**

全体の流れは以下の通りです：

1. レプリケーションする元のデータがある DSS サーバーの論理ボリュームにスナップショットを設定します。
2. また、ローカルバックアップを使用設定します。
3. 次にレプリケーションする先の DSS サーバーでレプリケーション・エージェントを使用設定します。
4. その後、レプリケーション先として使用する共有を作成、選択し、その共有部の設定に入り、レプリケーション・エージェントを設定にします。
5. レプリケーションする元のデータがある DSS サーバーでレプリケーション・タスクを作成します。
6. レプリケーション・タスクのスケジュールを設定し、レプリケーションを起動します。

### 2.3 設定の実際

#### A レプリケーションする元のデータがある DSS サーバーでの設定項目

1. 構成->ボリュームマネージャーから論理ボリュームを選択してユニットマネージャーを立ち上げます。

システム ボリューム	サイズ (GB)
スワップ用に予約されています	4.00
スナップショット用に予約されています	0.00
システム用に予約されています	1.00
複製用に予約されています	0.00
空き	293.03

アクション: 新しい NAS ボリューム

ボリューム複製の使用  
 WORM

追加:  GB

適用する

アクションからスナップショットの作成を選択し、必要な容量を指定して“適用する”をクリックします。一般的にはスナップショットを行うターゲット論理ボリュームの 10%がスナップショットの容量値として適しています。

2. ユニットマネージャーの下の方に進み、スナップショットの定義を表示します。



左記の例のように論理ボリュームに対するスナップショットを設定します。  
表示されるリストからターゲットの論理ボリュームを指定すれば完了です。その後、“適用する”をクリックします。

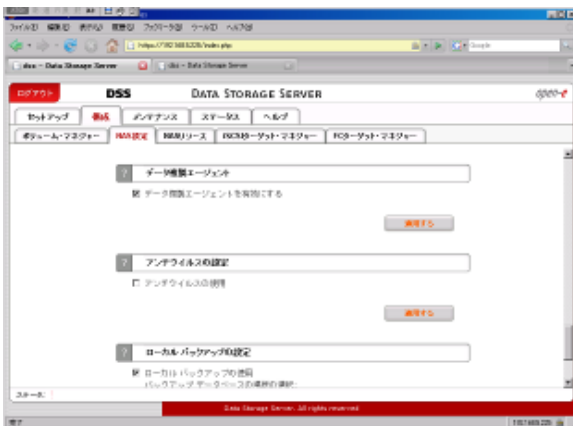
3. データ・レプリケーションはバックアップの一方法のため、ローカルバックアップを使用設定にします。



構成->NAS 設定->ローカルバックアップの設定でローカルバックアップの使用にチェックし、その他の必要な設定項目を入力します。

## B データをレプリケーションする先(Destination)の DSS サーバー設定

1. 構成->NAS 設定からデータ複製エージェントを開き、エージェント設定を有効にします。



2. 既にレプリケーション先として使用する共有部が作成されているものとします。

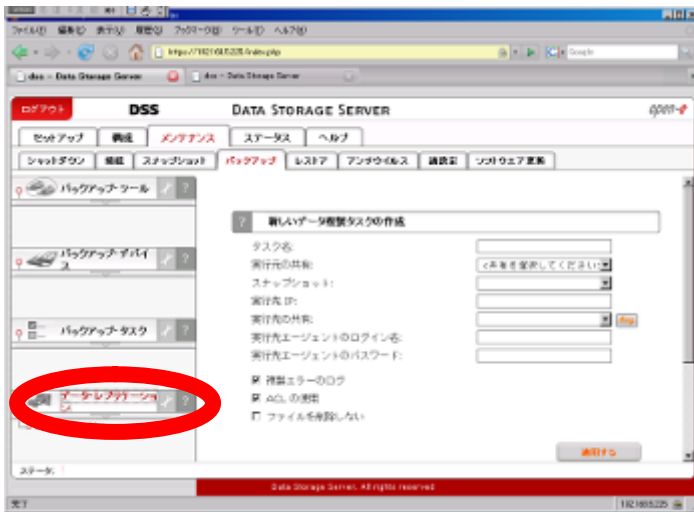


構成->NAS リソースからレプリケーション先の共有を開きます。左図では office という共有を開いています。ここで、データ複製の使用にチェックするとこの共有へのレプリケーションとしてアクセスする ID とパスワードを設定可能です。  
データの安全面からパスワードを使用することをお勧めします。これでこの DSS サーバーがデータ・レプリケーションのコピー先として登録されます。

レプリケーション先の DSS サーバーで設定する項目は以上です。

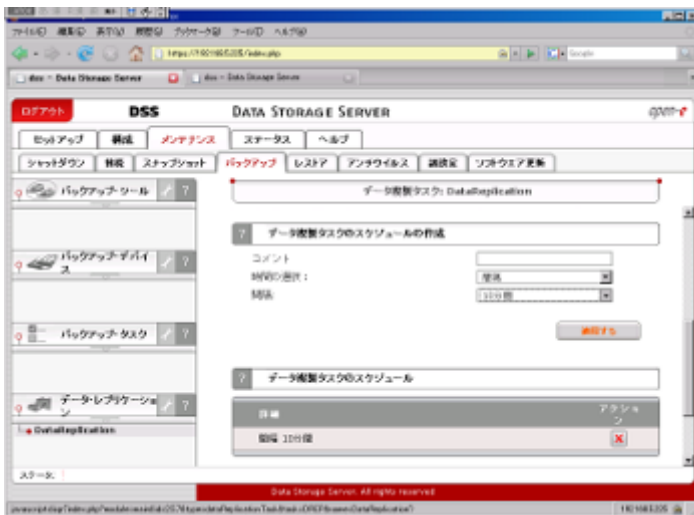
### C レプリケーションする元のデータがある DSS サーバーでの設定項目

3. メンテナンス→バックアップに表示されている左欄のタスクからデータ・レプリケーションをクリックします。すると新しいデータ複製タスクの作成が表示されます。ここに必要な情報を入力してください。



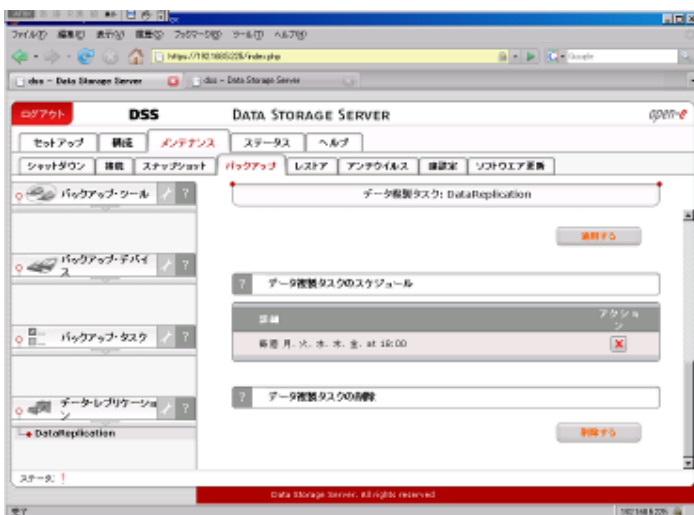
特にレプリケーション先の DSS サーバーの IP アドレスを正しく入力してください。その後、実行先の共有 右の参照ボタンをクリックするとエージェント設定してある共有が表示されます。その中からレプリケーション先に使用する共有を選択して下さい。実行先エージェントで設定した ID、パスワードを入力します。

4. レプリケーションタスクのスケジュールを設定する。



メンテナンス→バックアップよりデータレプリケーション内に表示されているタスクを選択します。左図では DatReplication を選択しています。データ複製タスクのスケジュールの作成を開きます。ここでレプリケーションを実施するスケジュールを入力します。

何時間間隔でも結構ですし、何曜日の何時という指定でも結構です。



設定されたタスクのスケジュールはその下のデータ複製タスクのスケジュールで確認可能です。例えば、月～金の 18:00 にレプリケーションする場合の設定は左図の通りです。

## 5. レプリケーション結果の確認

The screenshot shows the DSS web interface. The main content area displays a table of tasks. The table has columns for '時間' (Time), '名前' (Name), 'タイプ' (Type), '状態' (Status), and 'アクション' (Action). The tasks listed are 'DataReplication' tasks with various completion times and statuses (OK, 終了, 開始).

時間	名前	タイプ	状態	アクション
2008-02-07 13:50:15	DataReplication	データ複製	OK	終了
2008-02-07 13:50:12	DataReplication	データ複製	OK	開始
2008-02-07 13:40:15	DataReplication	データ複製	OK	終了
2008-02-07 13:40:13	DataReplication	データ複製	OK	開始
2008-02-07 13:30:16	DataReplication	データ複製	OK	終了

ステータス->タスクを開き、左欄のデータ複製を選択します。実行中のタスク及びタスクログが表示されます。

## 2.4 ボリューム・レプリケーションの設定方法

ボリューム・レプリケーションはボリュームのミラーリングを LAN を通じて行います。ボリューム・レプリケーションはブロック・レベルのレプリケーションを行います。ただし、仮想化ソフトウェアのパフォーマンスを向上させるためにファイル I/O も選択可能です。ボリュームは iSCSI、NAS、FC のいずれでもレプリケーション可能です。また、iSCSI フェールオーバーを使用する場合は下層ではボリューム・レプリケーションが必要になります。

### ノート！ ブロック I/O とファイル I/O について

Open-E DSS では標準ではデータ・レプリケーションにファイル I/O を、ボリューム・レプリケーションにブロック I/O を使用します。iSCSI ボリュームの場合はボリューム・レプリケーションでもファイル I/O を選択可能です。

ファイル I/O はファイルを複製するのに導入が簡単な方法です。不利な点は速度が遅い、遅延時間があるということと、ほとんどのアプリケーションがネットワーク共有にインストールできない、と言う事実です。

ブロック I/O はディスクに直接アクセスする方法で速度が速く、遅延時間が最小限で、可用性も高くなります。不利な点はファイルを共有する能力をそのままでは持たないということ、内部動作が複雑なこと、そして一般的にコストが高いことです。

Open-E DSS の iSCSI ターゲットの場合にはブロック I/O とファイル I/O を選択可能です。ブロック I/O の場合は論理ボリュームに直接アクセスする I/O となります。ファイル I/O の場合はファイル・キャッシュなどの追加レイヤを使用します。安全なのはブロック I/O ですが、VMware やデータベースを使用した場合に十分な安全性とアクセス速度の向上を図れるのがファイル I/O です。ファイル I/O は技術的には IET (iSCSI エンタープライズ・ターゲット) にマウントする XFS パーティションに LUN を構成し、iSCSI ターゲットを割り当てます。

**注意！！ GUI の作成時期の関係でレプリケーションと複製、レプリケーション先と実行先、コピー先など同じことを意味するのに異なる用語が使用されている場合があります。ご注意ください。**

**注意！！ レプリケーション元と先はかならず同じ容量に設定してください。**

**重要！！ レプリケーションでは両方のサーバーの時刻設定が同じである必要があります。必ず、時刻設定を同じに設定してください。**

### 1. レプリケーションする元のボリュームがある DSS サーバーの設定



ボリューム作成の際にボリューム複製の使用にチェックを入れておきます。

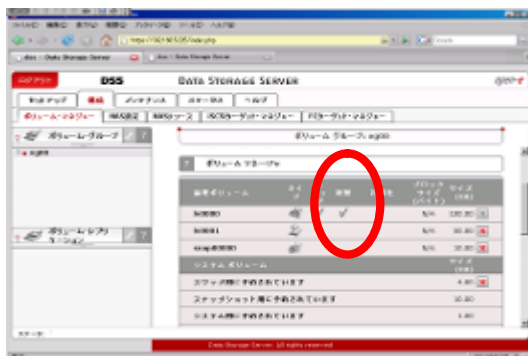
一度作成した論理ボリュームを複製する場合はアクションから該当論理ボリュームを選択して (展開 lv001 など)、チェックして適用してください。さらに IO タイプを選択してください。

File I/O にチェックを入れない場合は自動的にブロック I/O となります。

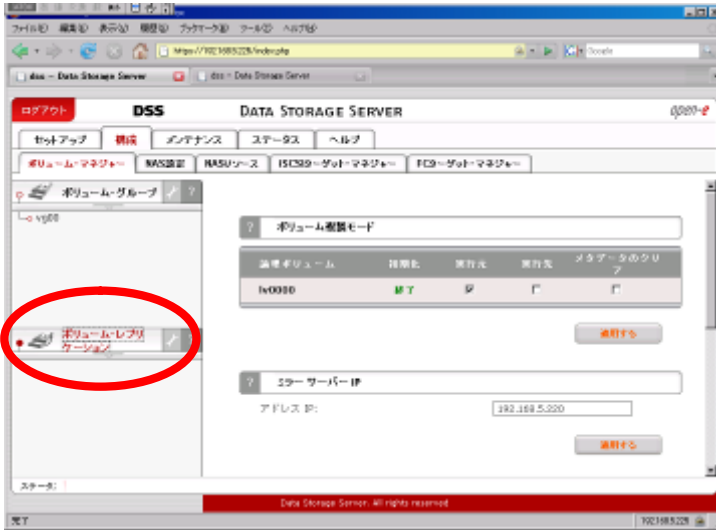
基本的にはブロック I/O をおすすめしますが、VMware、Virtualiron などの仮想化ソフトウェアと組み合わせる場合はパフォーマンスの観点からファイル I/O をおすすめします。

ファイル I/O でデータ保護を行う場合はディスクに 0 を書き込む初期化が必要になります。

ファイル I/O の初期化はディスク容量によっては大量の RAM、CPU 能力が必要になります。



作成されると左図のように複製欄にチェックが入り、IO タイプが表示されます。



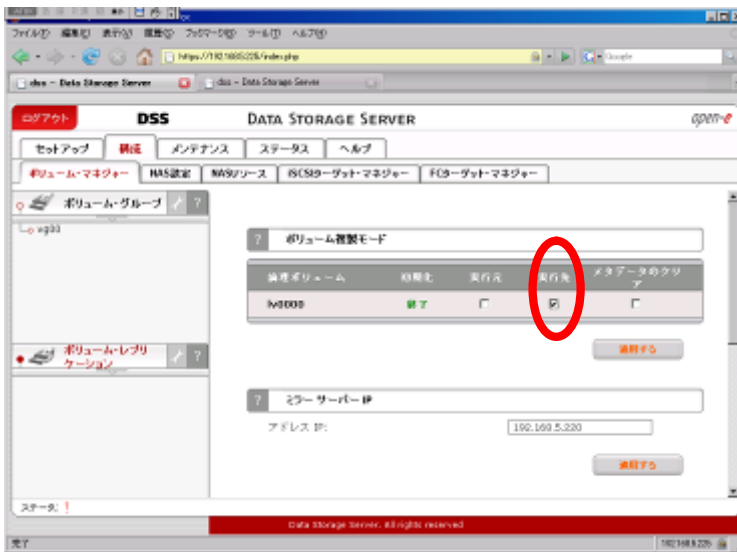
左側のボリューム・レプリケーションの文字をクリックして、ボリューム・レプリケーションを開きます。ボリューム複製を使用するにチェックを入れた論理ボリュームがボリューム複製モード欄に表示されます。

右図では lv0000。

ここでこのボリュームがレプリケーション元であるのか、レプリケーション先であるのか、レプリケーションするときメタデータを消去するのかを設定可能です。

ここではレプリケーション元の DSS サーバーなので実行元にチェックが入っています。設定後は適用するを必ずクリックして変更を確定してください。

## 2. レプリケーションする先のボリュームがある DSS サーバーの設定

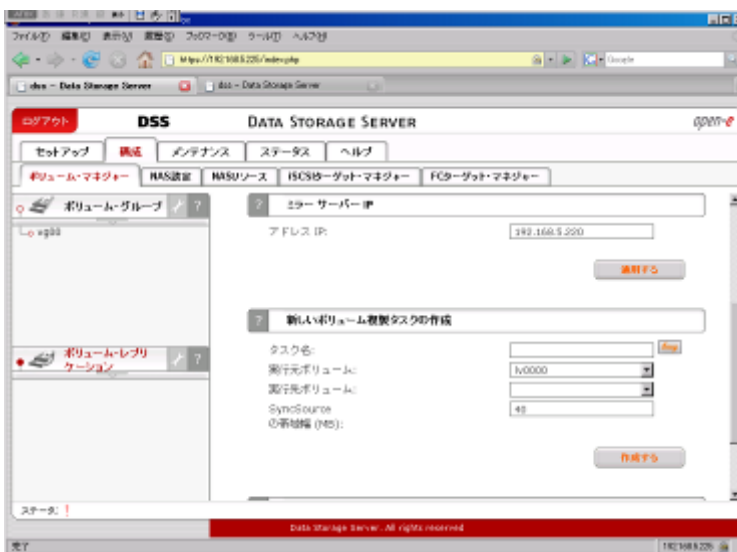


レプリケーション先の DSS サーバーでも、レプリケーション元と同様に論理ボリュームの複製を使用するにチェックを入れた論理ボリュームを作成します。

そして、ボリューム・レプリケーションの文字をクリックして表示されるボリューム複製モードで実行先にチェックを入れます。設定後は適用するを必ずクリックして変更を確定してください。

次にそのすぐ下のミラーサーバー IP を開きます。ここでレプリケーション元の DSS サーバーの IP アドレスを入力します。その後適用するをクリックします。

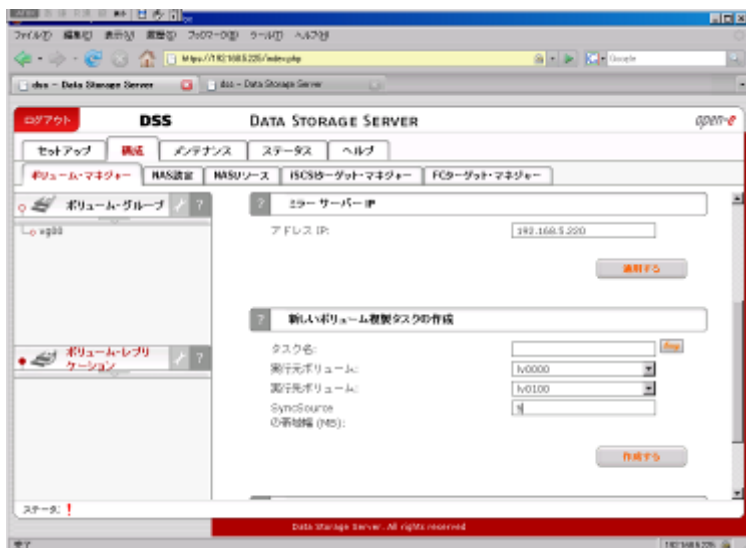
## 3. レプリケーション元のデータがある DSS サーバーに戻ります。




ボリューム・レプリケーションを押して表示されるボリューム複製モードが実行元になっていることを確認して、下のミラーサーバー IP を開きます。ここでレプリケーション先の DSS サーバーの IP アドレスを入力します。

その後適用するをクリックします。

その後、下欄にある新しいボリューム複製タスクの作成に移動します。




まず、レプリケーションのタスク名を入力します。

ここでタスク名の右の参照ボタンをクリックするとボリューム・レプリケーションの複製先をミラーサーバーIPアドレスに探しに行きます。検索されると実行先ボリューム欄に選択肢として表示されますので、レプリケーション先を設定してください。必要な情報を設定後、**適用する**をクリックすればボリューム複製タスクが作成されます。必ずタスク名を確認してください。




左図のように複製タスクマネジャー内に作成されたタスクが表示されます。

ここでをクリックしていただくと、ボリューム・レプリケーションが開始されます。これはリアルタイム同期のため開始時期のみ表示されます。



左図が開始したレプリケーションのタスクマネジャー画面

です。この画面でをクリックしていただくと、ボリューム・レプリケーションが停止します。

このとき、レプリケーション先のボリュームでは**自動で** replication1\_reverse というタスクが作成され、開始されています。そのため、レプリケーション先のボリュームをその他の用途に割り当てることはできません。

**注意！！** ボリューム・レプリケーションはリアルタイムで行われるためシステムの反応が遅くなり、ネットワークをデータが複数回転送されるためネットワーク負荷が大きくなります。

OPEN-E ではレプリケーション用に専用ネットワークを構築することをお勧めしています。

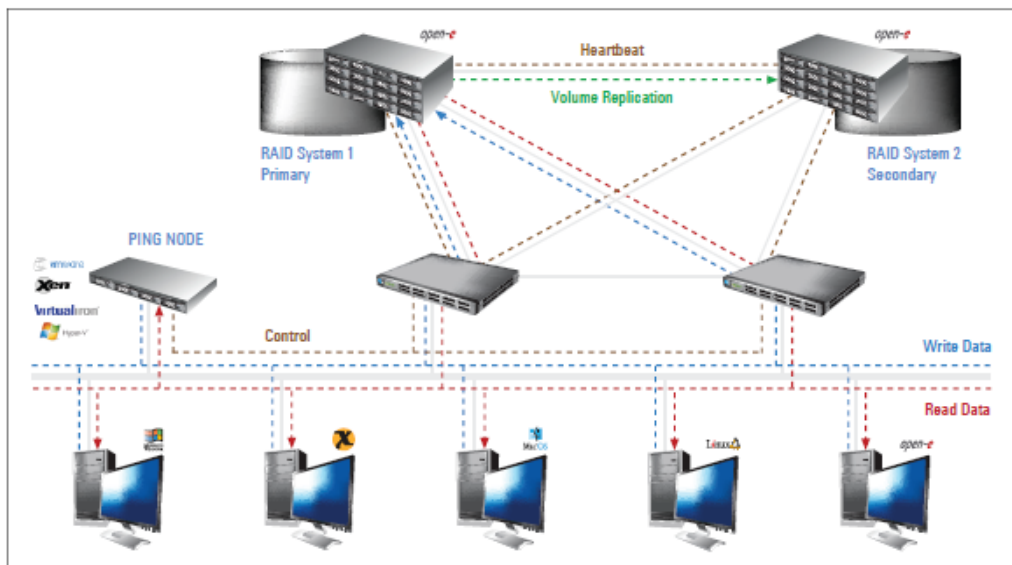
### 3 iSCSI フェールオーバー機能

iSCSI フェールオーバー機能とは、ボリューム・レプリケーションでミラーリングされている 2 個の iSCSI ボリュームに仮想 IP アドレスを設定し、プライマリ側の iSCSI ボリュームに障害が発生した場合に、自動的にセカンダリ側の iSCSI ボリュームに切り替わり、アクセスしているユーザーにはシームレスに iSCSI ボリュームを提供するソリューションです。

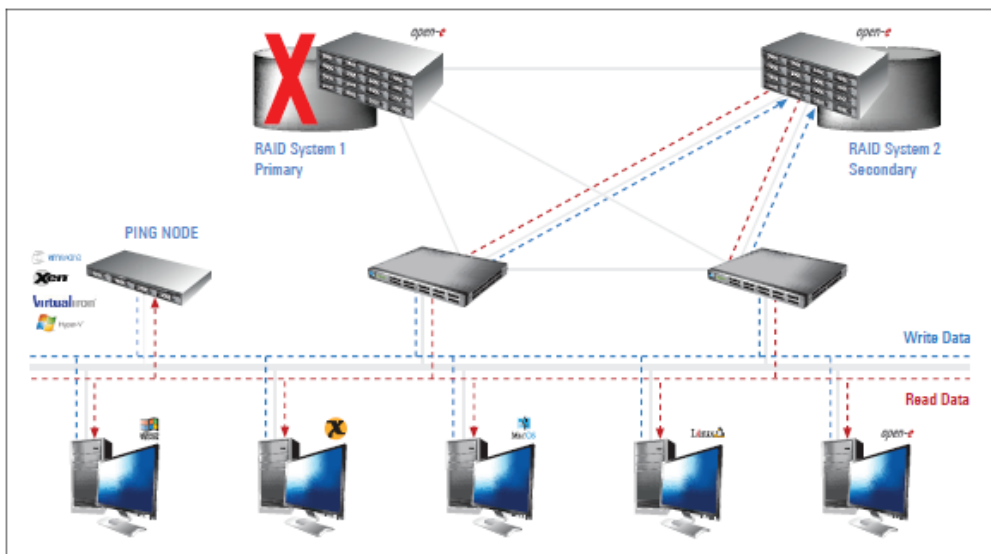
現在サポートしているのは Active/Passive、マニュアル・フェールバックモデルのみです。

すなわち、常時プライマリ側が動作しており、プライマリ側に障害が発生した場合に、自動的にセカンダリ側ボリュームに切り替わります。プライマリ側の障害が復旧した場合は、手動で両ボリュームの同期を取ってから、手動でフェールバックします。

また、現在フェールオーバー情報は SNMP に反映されません。



正常時  
すべての  
アクセスは  
プライマリ側



障害時  
すべての  
アクセスは  
セカンダリ側

#### 3.1 iSCSI フェールオーバー要件

1. 2台の iSCSI システム
2. システムごとに2枚以上の NIC  
一枚が iSCSI データ用、もう一枚が同期確認用に使用されます。
3. ネットワーク上の Ping node  
Ping できるネットワーク内のノードであればシステムは問いません。スイッチに設定することも可能です。ただし、必ず時刻設定がレプリケーション用サーバーと同じものがが必要です。

**重要！！** 仮想化ソフトの仕様によって設定が必要な場合があります。仮想化ソフトウェアのマニュアルなどを参考に正しい設定を行ってください。特にブロック I/O、ファイル I/O どちらか、LUN 番号などにご注意ください。

## 3.2 iSCSI フェールオーバー設定の順序

1. iSCSI フェールオーバー機能設定のためには iSCSI ボリュームのボリューム・レプリケーションが必須です。前章のボリューム・レプリケーション設定を参考に、ボリューム・レプリケーションを設定します。VMwareではFailoverIにファイルI/Oをサポートしています。ブロック I/O とファイル I/O の詳細については 9 ページのコラムをご参照ください。

**重要！！** レプリケーションでは両方のサーバーの時刻設定が同じである必要があります。必ず、時刻設定を同じに設定してください。

2. それぞれのボリュームを iSCSI ターゲットに設定します。その際、iSCSI ターゲット名、SCSI ID、LUN などは両方のボリュームで完全に同一にする必要があります。また、VMware、Virtualiron などの仮想化ソフトウェアを使用する場合、LUN は0以外を選択してください。また、RO、WB には絶対にチェックを入れないでください。CHAP の設定は仮想化ソフトウェアによってはサポートしていないので十分ご注意ください。さらに LUN は0以外の数字を選択してください。
3. プライマリ側、セカンダリ側それぞれのシステムで iSCSI データ用の NIC に仮想 IP アドレスを設定します。
4. プライマリ側、セカンダリ側それぞれのシステムで同期用の NIC に互いを参照する設定をします。
5. プライマリ側、セカンダリ側それぞれのシステムで iSCSI フェールオーバーを設定します。
6. プライマリ側で failover タスクを作成します。
7. プライマリ側の failover manager を使用して failover を開始します。セカンダリ側では自動的にそのタスクの reverse タスクが開始します。

## 3.3 iSCSI フェールオーバー設定の実際

### 3.3.1 iSCSI ボリュームの作成

構成→ボリュームマネージャーを使用して iSCSI ボリュームを作成します。その際、ボリューム複製の使用にチェックを入れておきます。


### 3.3.2 ボリューム複製の作成

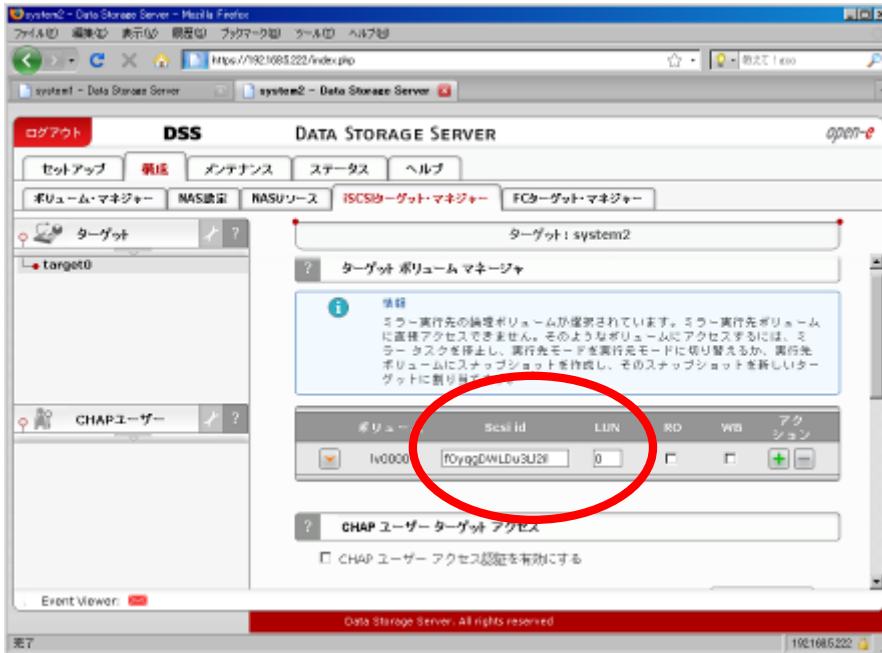
P36 のボリューム・レプリケーションを参考に、レプリケーションタスクを作成し、実行します。

### 3.3.3 iSCSI ターゲットの作成

プライマリ側システムの WebGUI で構成→iSCSI ターゲットマネージャーを開きます。ここで iSCSI ターゲットの名前を入力します。

デフォルトのターゲット名のチェックを外し、iSCSI ターゲット名を入力します。作成されると左欄に新規ターゲットが表示されます。

そして、SCSI ID、RO、WB の設定を確認後、Action 欄の  をクリックしてターゲットを開始します。



同様にセカンダリ側でも iSCSI ターゲットを作成し、ターゲット名、SCSI ID、RO、WB、LUN を設定します。この際、ターゲット名、SCSI ID、RO、WB、LUN はプライマリ側、セカンダリ側とも全く同じに設定してください。

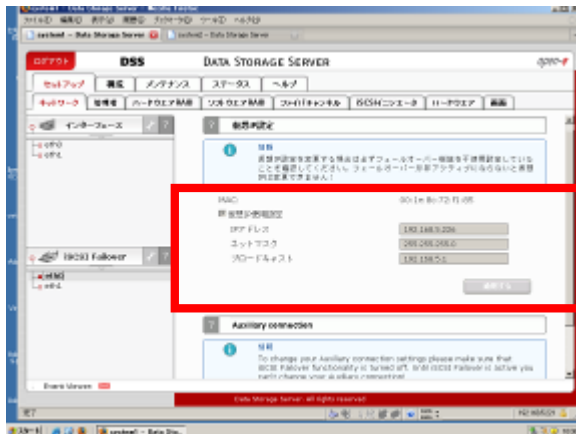
RO (Read only)、WB (Write back cache)にはチェックを入れないでください。iSCSI イニシエータと同期がとれない場合があります。WB にチェックが入らない場合は write through 設定となります。

また、VMware、Virtualiron などの仮想化ソフトウェアを使用する場合、ターゲット側の LUN は0以外を選択してください。

また、CHAP の設定は仮想化ソフトウェアによってはサポートしていないので十分ご注意ください。

### 3.3.4 仮想 IP アドレスを設定

プライマリ側システムの WebGUI でセットアップ→ネットワークを開きます。ここで左欄の iSCSI Failover 内の NIC を一枚選択します。ここでは eth0 を選択します。

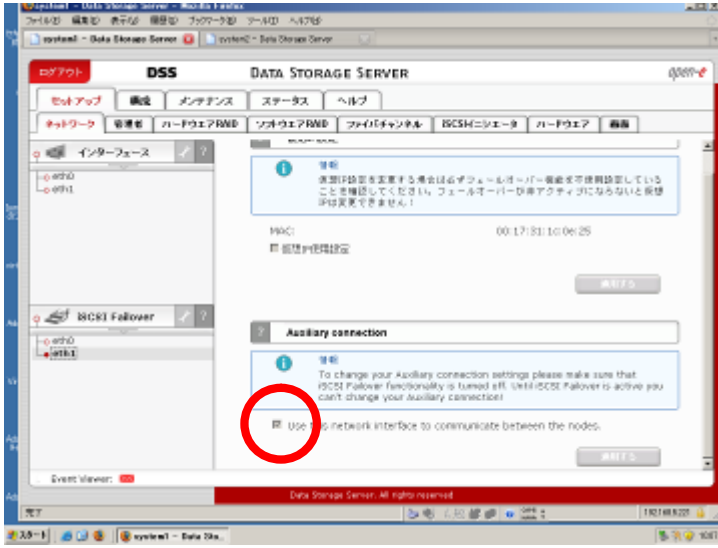


そして仮想 IP 使用設定にチェックを入れ、現れた設定ボックスに仮想 IP アドレス、ネットマスク、ブロードキャストアドレスなどの情報を入力します。その後、適用するをクリックして確定します。仮想 IP を使用設定すると自動的にその下欄の Auxiliary connection は使用設定となります。

同様にセカンダリ側でも仮想 IP アドレスを設定します。

### 3.3.5 同期用(heart beat)NIC 設定

プライマリ側システムの WebGUI でセットアップ→ネットワークを開きます。ここで左欄の iSCSI Failover 内の同期用 NIC を一枚選択します。ここでは eth1 を選択します。



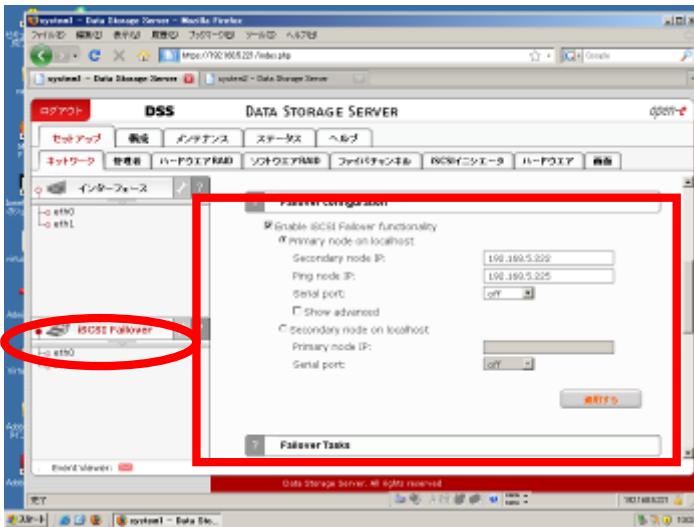
そして、Auxiliary connection のチェックボックスをチェックし、適用するをクリックして確定します。

同様にセカンダリ側でも Auxiliary connection のチェックボックスをチェック設定します。

これでプライマリ側、セカンダリ側それぞれのシステムで同期用の NIC で互いを参照する設定になりました。

### 3.3.6 iSCSI フェールオーバーを設定

プライマリ側システムの WebGUI でセットアップ→ネットワークを開きます。ここで左欄の iSCSI Failover の文字を選択し、フェールオーバーを設定します。



Enable iSCSI failover functionality にチェックを入れた後、

プライマリ側のシステムで作業しているのでセカンダリ側ノードの IP アドレス、ping node を設定します。

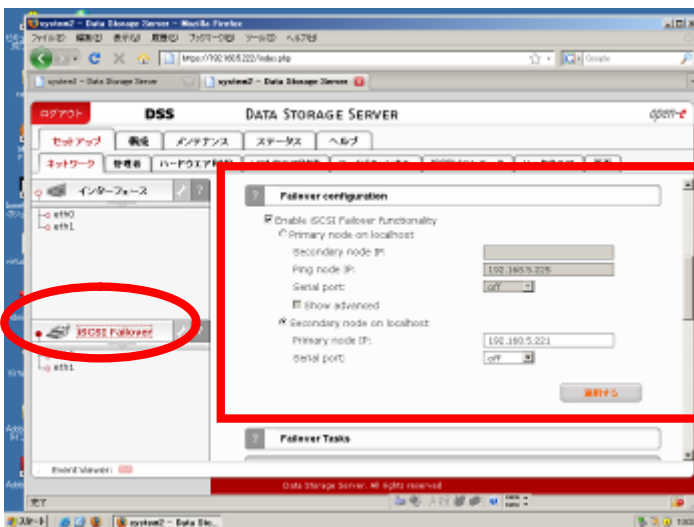
ping node は同一ネットワーク内にあり、ping で応答するシステムであればシステムは問いません。

ただし、時刻設定がフェールオーバーする両システムと同一であることを確認してください。

**注意！！！！**

ping node はメンバーの ICMP リクエストに回答する必要があるため、フェールオーバー設定している 2 台のシステムは使用できません。

同様にセカンダリ側で設定を行います。

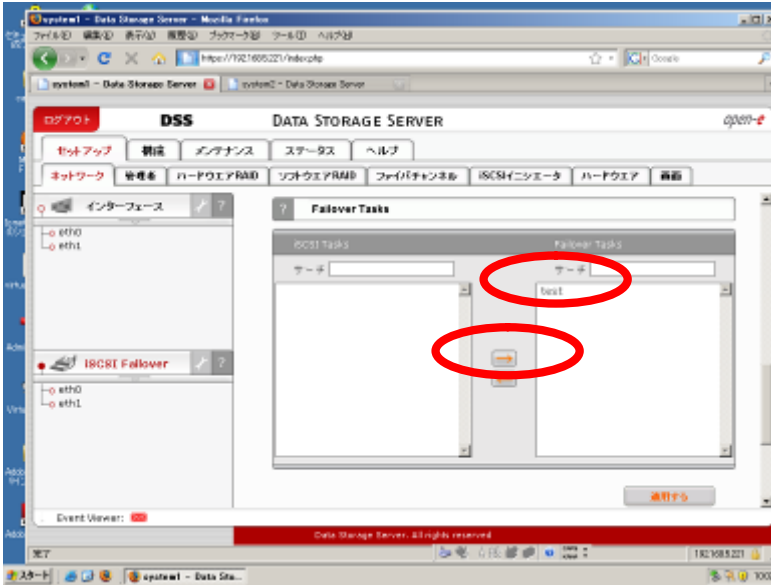


Enable iSCSI failover functionality にチェックを入れた後、

セカンダリ側のシステムで作業しているのでプライマリ側ノードの IP アドレスを設定します。

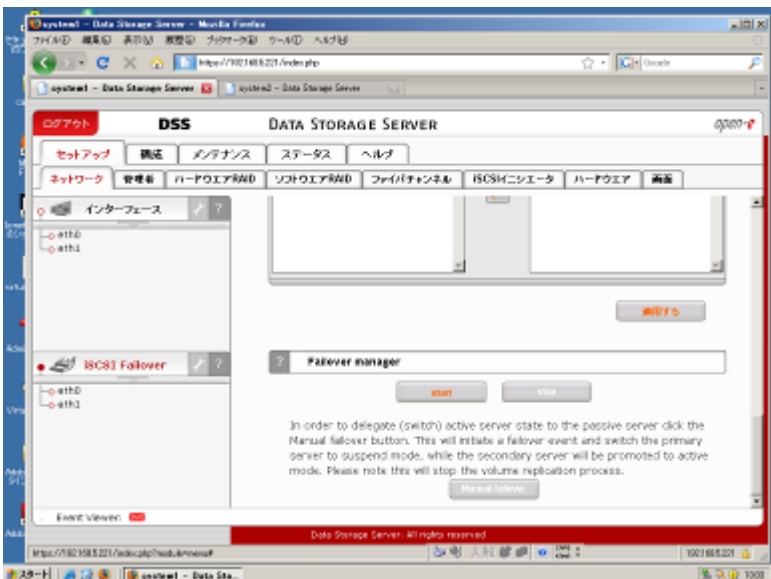
### 3.3.7 failover タスクを作成

プライマリ側システムの WebGUI でセットアップ→ネットワークを開きます。ここで左欄の iSCSI Failover の文字を選択します。

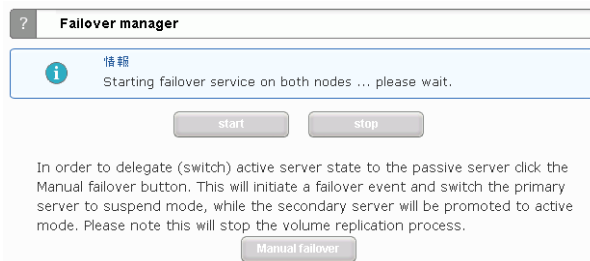


Failover task 欄に移動し、表示されているタスクのうち、実行するものを右側の欄に移動します。移動後、適用するをクリックして確定します。

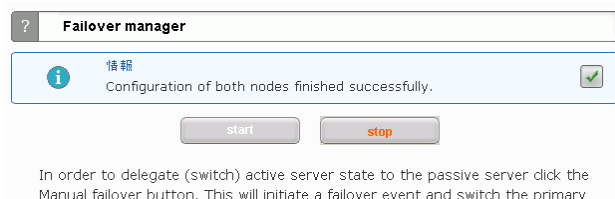
### 3.3.8 failover manager を使用して failover を開始

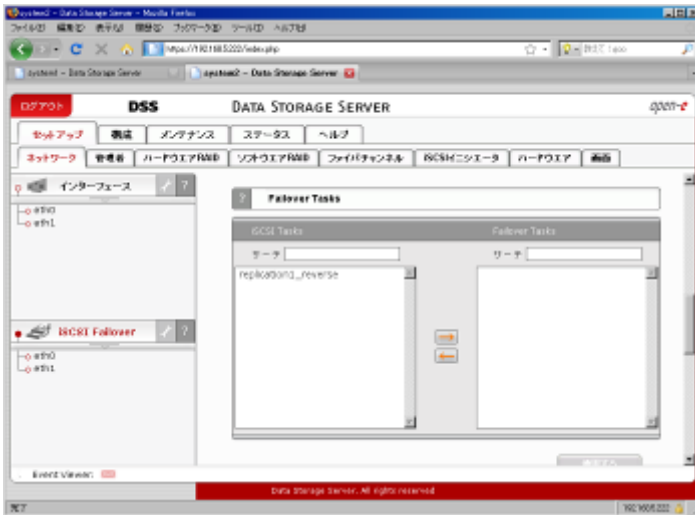


プライマリ側の failover task のすぐ下の failover manager を開きます。いままでの設定がすべて問題なければここでフェールオーバーが開始可能として、Start ボタンがハイライトされているはずです。Start がハイライトされていない場合はすべての設定をもう一度確認してください。

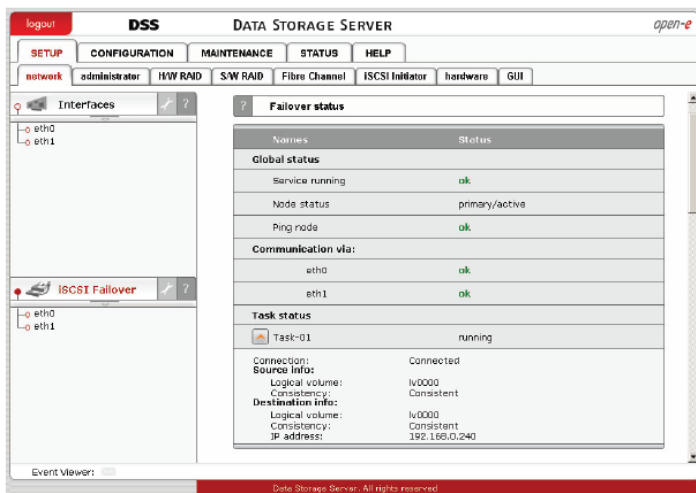


Startボタンをクリックすると、左図のように接続が進行し、接続が確立すると下図のようにフェールオーバーマネージャーに Stopボタンが表示されます。





セカンダリ側で設定する項目はありません。Failoverタスクがプライマリ側で開始すると、それに対応した reverse タスクが自動的にセカンダリ側で開始されます。

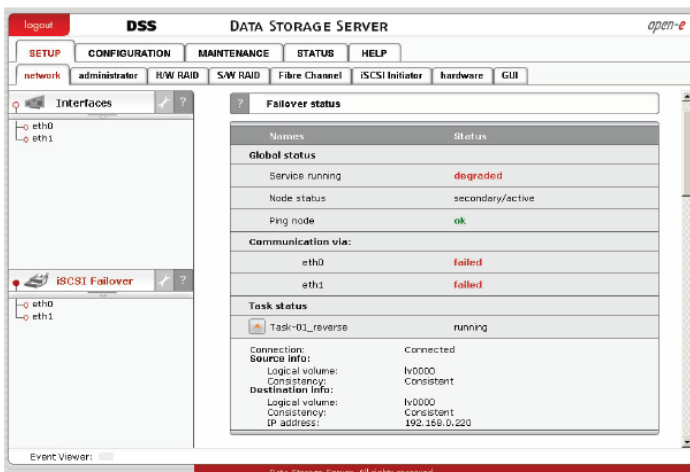


Failoverの状態は上欄の Failover status で確認可能です。プライマリ側は左図のようになります。

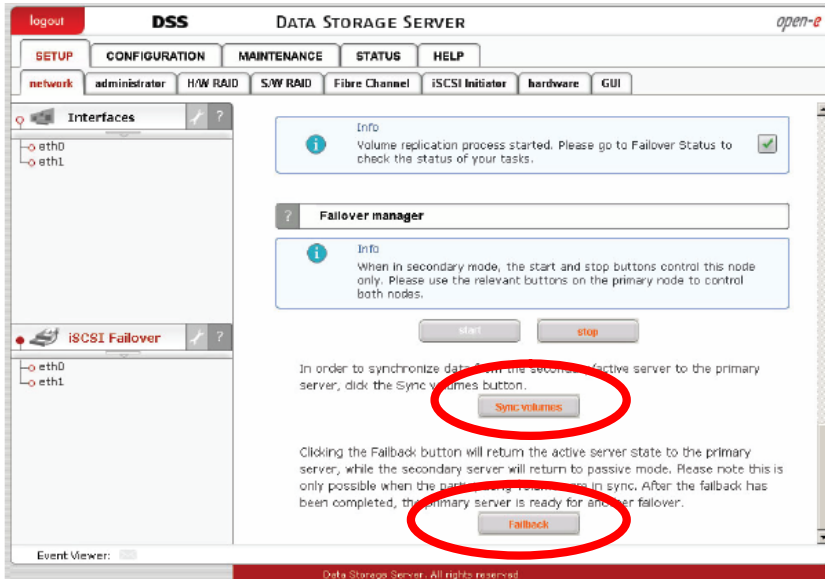
タスクの内容は以下のように確認可能です。

eth0	unknown
<b>Task status</b>	
replication1	running
接続:	Connected
複製元の情報:	
論理ボリューム:	lv0000
整合性:	Consistent
複製先の情報:	
論理ボリューム:	lv0000
整合性:	Consistent
IP アドレス:	192.168.5.222

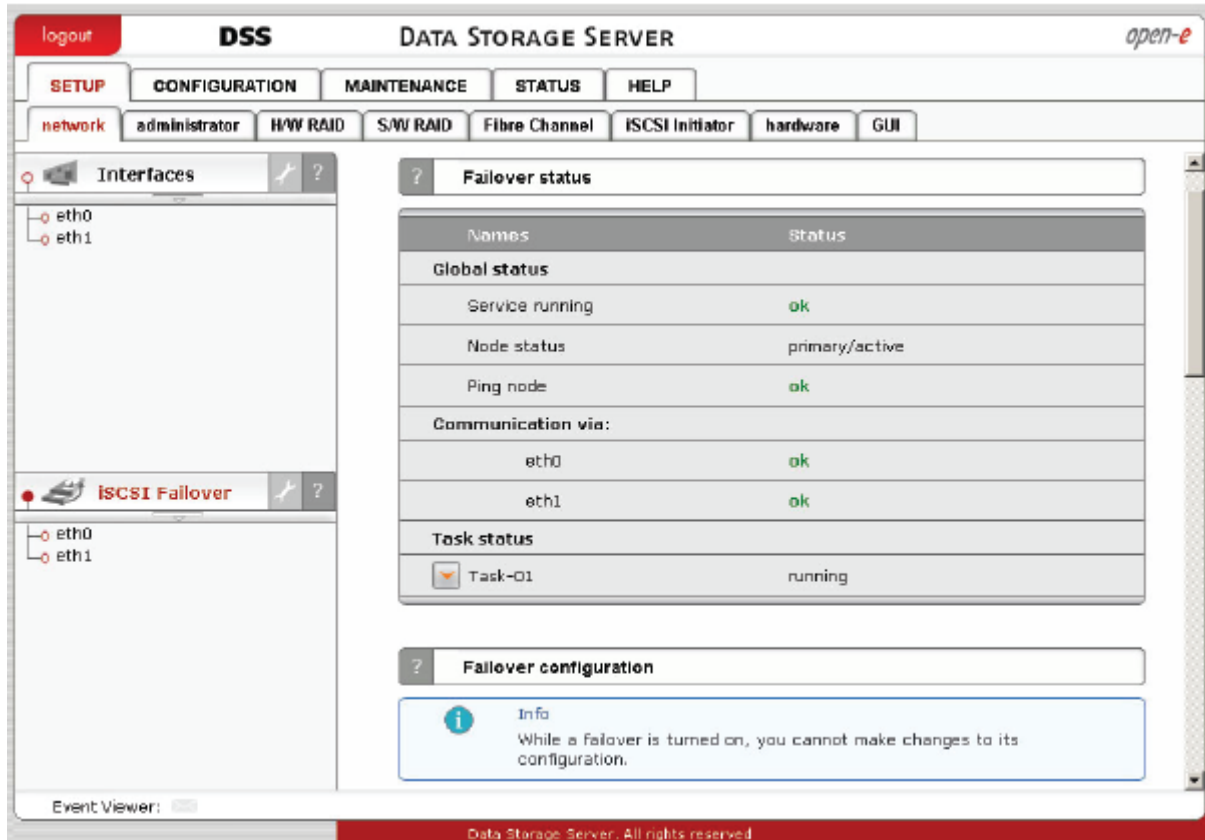
Failover設定後は必ず Manual Failover ボタンをクリックして、手動でFailoverさせてください。Manual Failover するとセカンダリ側は以下のように表示が変わります。



セカンダリ側がActiveになった時点でボリューム・レプリケーションは停止します。  
従いまして、再度プライマリ側をActiveに戻すためには、ボリュームの整合性を取った上で(Sync.  
volume)、手動で元に戻します。(Fail back)



完了後のプライマリ側の iSCSI Failover status は以下ようになります。



Open-E は IP ベースのストレージソリューションを提供しているストレージ専門ソフトウェア会社です。  
Open-E の各製品 DSS、NAS-R3、iSCSI-R3 は高い柔軟性と低コストを兼ね添え、簡単なストレージ管理を  
していただけます。1998 年創業、本社所在地：ドイツ、プフハイム市  
<http://www.open-e.com>

Open-E 日本事務所  
102-71 東京都千代田区富士見 2-3-3 タロービル3F  
Tel: 03-3238-7804 Fax: 03-3238-5332  
[hiro.hashimoto@open-e.com](mailto:hiro.hashimoto@open-e.com)

代理店：  
株式会社 アスク  
102-0076 東京都千代田区五番町 2-4 カサドタク1F  
Tel: 03-5215-5650 Fax: 03-5215-5651  
<http://www.ask-sss.com>  
[info@ask-sss.com](mailto:info@ask-sss.com)