(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2010-97533
(P2010-97533A)
(43)公開日　平成22年4月30日(2010.4.30)

| (51)Int.Cl. | | F I | | | テーマコード（参考） |
|---|---|---|---|---|---|
| G06F 9/46 | (2006.01) | G06F | 9/46 | 350 | 5B082 |
| G06F 12/00 | (2006.01) | G06F | 12/00 | 501B | |
| | | G06F | 12/00 | 545A | |

審査請求 未請求　請求項の数 11　OL　（全 51 頁）

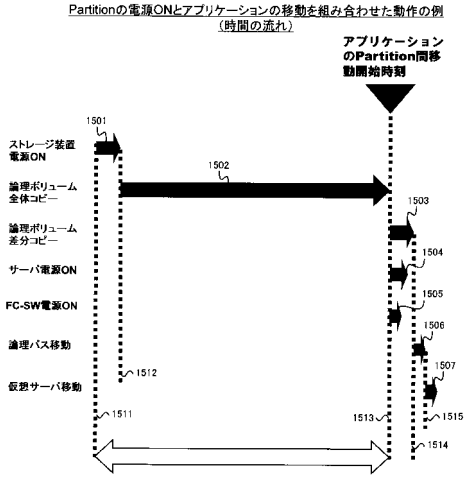
(21)出願番号　特願2008-269539 (P2008-269539)
(22)出願日　平成20年10月20日 (2008.10.20)

(71)出願人　000005108
株式会社日立製作所
東京都千代田区丸の内一丁目６番６号
(74)代理人　100075513
弁理士　後藤　政喜
(74)代理人　100114236
弁理士　藤井　正弘
(74)代理人　100120260
弁理士　飯田　雅昭
(72)発明者　鈴木　友彦
神奈川県小田原市中里３２２番２号　株式会社日立製作所ＳＡＮソリューション事業部内

Ｆターム(参考)　5B082 CA11　CA19　HA08

(54)【発明の名称】パーティションで区切られた計算機システムにおけるアプリケーション移動及び消費電力の最適化

(57)【要約】
【課題】論理分割された計算機システムにおいて、アプリケーション移動時の消費電力を削減する。
【解決手段】アプリケーションの移動元の論理ボリュームを含むストレージ装置は、前記移動元の論理ボリュームに格納されたデータを、アプリケーションの移動先の論理ボリュームにコピーし、前記コピーが開始された後に前記移動元の論理ボリュームに書き込まれたデータを前記移動元の論理ボリュームに格納せずに差分データとして保持し、管理計算機は、前記移動元の論理ボリュームに格納されたデータのコピーが終了すると、前記差分データのコピーを開始し、前記移動元の論理ボリュームに格納されたデータのコピーが終了してから前記差分データのコピーが終了するまでの間に、前記アプリケーションの移動先の計算機の電源を投入する。
【選択図】図１５



【特許請求の範囲】
【請求項１】
一つ以上の計算機と、ネットワークを介して前記一つ以上の計算機に接続される一つ以上のストレージ装置と、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される管理計算機と、を備える計算機システムであって、
前記各計算機は、ハードウェア資源として、前記ネットワークに接続される第１インタフェースと、前記第１インタフェースに接続される第１プロセッサと、前記第１プロセッサに接続される第１メモリと、前記管理計算機からの要求に従って前記ハードウェア資源への電源の投入及び遮断を制御する電源制御部と、を備え、
前記各計算機は、さらに、前記一つ以上の計算機のハードウェア資源に基づいて複数の仮想領域を提供する仮想化部を備え、
前記複数の仮想領域は、第１仮想領域及び第２仮想領域を含み、
前記第１仮想領域は、アプリケーションプログラムを実行する仮想計算機として動作し、
前記各ストレージ装置は、前記計算機によって書き込まれたデータを格納する記憶領域を提供する記憶媒体と、前記記憶媒体へのデータの入出力を制御するコントローラと、前記管理計算機からの要求に従って前記各ストレージ装置の電源の投入及び遮断を制御する第２電源制御部と、を備え、
前記一つ以上のストレージ装置の前記コントローラは、前記一つ以上のストレージ装置の前記記憶領域を複数の論理ボリュームとして前記計算機に提供し、
前記複数の論理ボリュームは、第１論理ボリューム及び第２論理ボリュームを含み、
前記第１論理ボリュームには、前記仮想計算機によって書き込まれたデータが格納され、
前記管理計算機は、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される第２インタフェースと、前記第２インタフェースに接続される第２プロセッサと、前記第２プロセッサに接続される第２メモリと、を備え、
前記管理計算機は、
前記第１論理ボリュームを含む前記ストレージ装置に、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を送信し、
前記第１論理ボリュームを含む前記ストレージ装置は、
前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を受信した後に前記仮想計算機から前記第１論理ボリュームへのデータの書き込み要求を受信すると、書き込みを要求されたデータを前記第１論理ボリュームに書き込まずに、差分データとして保持し、
前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求に従って、前記第１論理ボリュームに格納されたデータを読み出して、前記第２論理ボリュームを含む前記ストレージ装置に送信し、
前記管理計算機は、
前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出すると、前記保持された差分データを前記第２論理ボリュームにコピーする要求を、前記第１論理ボリュームを含む前記ストレージ装置に送信し、
前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出してから、前記差分データの前記第２論理ボリュームへのコピーが終了するまでの間に、前記一つ以上の計算機のハードウェア資源のうち、前記第２仮想領域に割り当てられたハードウェア資源の電源を投入する要求を、前記第２仮想領域に割り当てられたハードウェア資源を含む前記計算機に送信し、
前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信し、
前記仮想計算機は、前記第２仮想領域に移動した後、前記第２論理ボリュームへのデータ入出力を実行することを特徴とする計算機システム。

【請求項２】
前記第１論理ボリュームを含む前記ストレージ装置は、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求に従う前記データの読み出し及び送信が終了すると、終了メッセージを前記管理計算機に送信し、
前記管理計算機は、前記終了メッセージを受信したときに、前記コピーの終了を検出することを特徴とする請求項１に記載の計算機システム。
【請求項３】
前記管理計算機は、前記第１論理ボリュームの容量、及び、前記第１論理ボリュームから前記第２論理ボリュームへのデータの転送速度に基づいて、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーに要する時間を算出し、
前記算出された時間が経過したときに、前記コピーの終了を検出することを特徴とする請求項１に記載の計算機システム。
【請求項４】
前記ネットワークは、前記一つ以上の計算機と前記一つ以上のストレージ装置との間のデータ転送を中継する複数のスイッチ装置を備え、
前記各スイッチ装置は、複数のポートと、前記管理計算機に接続される第３インタフェースと、前記管理計算機からの要求に従って前記各スイッチ装置の電源の投入及び遮断を制御する第３電源制御部と、を備え、
前記各ポートは、前記計算機又は前記ストレージ装置に接続され、
前記複数のスイッチ装置は、第１スイッチ装置及び第２スイッチ装置を含み、
前記第１仮想領域で稼動する前記仮想計算機は、前記第１スイッチ装置を経由して前記第１論理ボリュームへのデータの入出力を実行し、
前記管理計算機は、
前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出してから、前記差分データの前記第２論理ボリュームへのコピーが終了するまでの間に、前記第２スイッチ装置に電源を投入する要求を送信し、
前記仮想計算機が前記第２スイッチ装置を経由して前記第２論理ボリュームへのデータの入出力を実行するように経路を切り替える要求を、前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信する前に送信し、
前記仮想計算機は、前記第２仮想領域に移動した後、前記第２スイッチ装置を経由して前記第２論理ボリュームへのデータ入出力を実行することを特徴とする請求項１に記載の計算機システム。
【請求項５】
一つ以上の計算機と、ネットワークを介して前記一つ以上の計算機に接続される一つ以上のストレージ装置と、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される管理計算機と、を備える計算機システムであって、
前記各計算機は、ハードウェア資源として、前記ネットワークに接続される第１インタフェースと、前記第１インタフェースに接続される第１プロセッサと、前記第１プロセッサに接続される第１メモリと、前記管理計算機からの要求に従って前記ハードウェア資源への電源の投入及び遮断を制御する電源制御部と、を備え、
前記仮想計算機は、さらに、前記一つ以上の計算機のハードウェア資源に基づいて複数の仮想領域を提供する仮想化部を備え、
前記複数の仮想領域は、第１仮想領域及び第２仮想領域を含み、
前記第１仮想領域は、アプリケーションプログラムを実行する仮想計算機として動作し、
前記各ストレージ装置は、前記計算機によって書き込まれたデータを格納する記憶領域を提供する記憶媒体と、前記記憶媒体へのデータの入出力を制御するコントローラと、前記管理計算機からの要求に従って前記各ストレージ装置の電源の投入及び遮断を制御する第２電源制御部と、を備え、

前記一つ以上のストレージ装置の前記コントローラは、前記一つ以上のストレージ装置の前記記憶領域を複数の論理ボリュームとして前記計算機に提供し、
前記複数の論理ボリュームは、第１論理ボリューム及び第２論理ボリュームを含み、
前記第１論理ボリュームには、前記仮想計算機によって書き込まれたデータが格納され、
前記管理計算機は、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される第２インタフェースと、前記第２インタフェースに接続される第２プロセッサと、前記第２プロセッサに接続される第２メモリと、を備え、
前記管理計算機は、
前記第１論理ボリュームを含む前記ストレージ装置に、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を送信し、
前記第１論理ボリュームを含む前記ストレージ装置は、
前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を受信した後に前記仮想計算機から前記第１論理ボリュームへのデータの書き込み要求を受信すると、書き込みを要求されたデータを前記第１論理ボリュームに書き込まずに、差分データとして保持し、
前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求に従って、前記第１論理ボリュームに格納されたデータを読み出して、前記第２論理ボリュームを含む前記ストレージ装置に送信し、
前記管理計算機は、
前記第１論理ボリュームに格納された全データのうち、所定の割合のデータの前記第２論理ボリュームへのコピーが終了すると、前記所定の割合のデータのコピーが終了した時点の前記差分データの量に基づいて、前記第１論理ボリュームに格納された全データのコピーが終了する時点の前記差分データの前記第２論理ボリュームへのコピーが終了する時刻を予測し、
前記予測されたコピーの終了時刻に前記計算機の起動処理が終了するように前記計算機の電源投入時刻を算出し、
前記算出された前記計算機の電源投入時刻が到来すると、前記一つ以上の計算機のハードウェア資源のうち、前記第２仮想領域に割り当てられたハードウェア資源の電源を投入する要求を、前記第２仮想領域に割り当てられたハードウェア資源を含む前記計算機に送信し、
前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出すると、前記保持された差分データを前記第２論理ボリュームにコピーする要求を、前記第１論理ボリュームを含む前記ストレージ装置に送信し、
前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信し、
前記仮想計算機は、前記第２仮想領域に移動した後、前記第２論理ボリュームへのデータ入出力を実行することを特徴とする計算機システム。
【請求項６】
前記ネットワークは、前記一つ以上の計算機と前記一つ以上のストレージ装置との間のデータ転送を中継する複数のスイッチ装置を備え、
前記各スイッチ装置は、複数のポートと、前記管理計算機に接続される第３インタフェースと、前記管理計算機からの要求に従って前記各スイッチ装置の電源の投入及び遮断を制御する第３電源制御部と、を備え、
前記各ポートは、前記計算機又は前記ストレージ装置に接続され、
前記複数のスイッチ装置は、第１スイッチ装置及び第２スイッチ装置を含み、
前記第１仮想領域で稼動する前記仮想計算機は、前記第１スイッチ装置を経由して前記第１論理ボリュームへのデータの入出力を実行し、
前記管理計算機は、
前記予測されたコピーの終了時刻に前記第２スイッチ装置の起動処理が終了するように

前記第２スイッチ装置の電源投入時刻を算出し、
前記算出された前記第２スイッチ装置の電源投入時刻が到来すると、前記第２スイッチ装置に電源を投入する要求を送信し、
前記仮想計算機が前記第２スイッチ装置を経由して前記第２論理ボリュームへのデータの入出力を実行するように経路を切り替える要求を、前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信する前に送信し、
前記仮想計算機は、前記第２仮想領域に移動した後、前記第２スイッチ装置を経由して前記第２論理ボリュームへのデータ入出力を実行することを特徴とする請求項５に記載の計算機システム。

【請求項７】
前記管理計算機は、
前記計算機の起動処理に要する時間を示す情報、及び、前記第２スイッチ装置の起動処理に要する時間をあらかじめ保持し、
前記所定の割合のデータのコピーが終了した時点の前記差分データの量を、前記所定の割合の逆数を乗ずることによって、前記前記第１論理ボリュームに格納された全データのコピーが終了する時点の前記差分データの量を予測し、
前記予測された差分データの量と、前記第１論理ボリュームを含む前記ストレージ装置から前記第２論理ボリュームを含む前記ストレージ装置へのデータ転送速度と、に基づいて、前記第１論理ボリュームに格納された全データのコピーが終了する時点の前記差分データの前記第２論理ボリュームへのコピーが終了する時刻を予測し、
前記予測されたコピーの終了時刻から前記計算機の起動処理に要する時間を減算することによって、前記計算機の電源投入時刻を算出し、
前記予測されたコピーの終了時刻から前記第２スイッチ装置の起動処理に要する時間を減算することによって、前記第２スイッチ装置の電源投入時刻を算出することを特徴とする請求項６に記載の計算機システム。

【請求項８】
一つ以上の計算機と、ネットワークを介して前記一つ以上の計算機に接続される一つ以上のストレージ装置と、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される管理計算機と、を備える計算機システムを制御する方法であって、
前記各計算機は、ハードウェア資源として、前記ネットワークに接続される第１インタフェースと、前記第１インタフェースに接続される第１プロセッサと、前記第１プロセッサに接続される第１メモリと、前記管理計算機からの要求に従って前記ハードウェア資源への電源の投入及び遮断を制御する電源制御部と、を備え、
前記各計算機は、さらに、前記一つ以上の計算機のハードウェア資源に基づいて複数の仮想領域を提供する仮想化部を備え、
前記複数の仮想領域は、第１仮想領域及び第２仮想領域を含み、
前記第１仮想領域は、アプリケーションプログラムを実行する仮想計算機として動作し、
前記各ストレージ装置は、前記計算機によって書き込まれたデータを格納する記憶領域を提供する記憶媒体と、前記記憶媒体へのデータの入出力を制御するコントローラと、前記管理計算機からの要求に従って前記各ストレージ装置の電源の投入及び遮断を制御する第２電源制御部と、を備え、
前記一つ以上のストレージ装置の前記記憶領域は、前記一つ以上のストレージ装置の前記コントローラによって、複数の論理ボリュームとして前記計算機に提供され、
前記複数の論理ボリュームは、第１論理ボリューム及び第２論理ボリュームを含み、
前記第１論理ボリュームには、前記仮想計算機によって書き込まれたデータが格納され、
前記管理計算機は、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される第２インタフェースと、前記第２インタフェースに接続される第２プロセッサと、前

記第２プロセッサに接続される第２メモリと、を備え、
前記方法は、
前記管理計算機が、前記第１論理ボリュームを含む前記ストレージ装置に、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を送信する手順と、
前記第１論理ボリュームを含む前記ストレージ装置が、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を受信した後に前記仮想計算機から前記第１論理ボリュームへのデータの書き込み要求を受信すると、書き込みを要求されたデータを前記第１論理ボリュームに書き込まずに、差分データとして保持する手順と、
前記第１論理ボリュームを含む前記ストレージ装置が、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求に従って、前記第１論理ボリュームに格納されたデータを読み出して、前記第２論理ボリュームを含む前記ストレージ装置に送信する手順と、
前記管理計算機が、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出すると、前記保持された差分データを前記第２論理ボリュームにコピーする要求を、前記第１論理ボリュームを含む前記ストレージ装置に送信する手順と、
前記管理計算機が、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出してから、前記差分データの前記第２論理ボリュームへのコピーが終了するまでの間に、前記一つ以上の計算機のハードウェア資源のうち、前記第２仮想領域に割り当てられたハードウェア資源の電源を投入する要求を、前記第２仮想領域に割り当てられたハードウェア資源を含む前記計算機に送信する手順と、
前記管理計算機が、前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信する手順と、
前記仮想計算機が、前記第２仮想領域に移動した後、前記第２論理ボリュームへのデータ入出力を実行する手順と、を含むことを特徴とする方法。
【請求項９】
前記方法は、さらに、前記第１論理ボリュームを含む前記ストレージ装置が、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求に従う前記データの読み出し及び送信が終了すると、終了メッセージを前記管理計算機に送信する手順を含み、
前記保持された差分データを前記第２論理ボリュームにコピーする要求を送信する手順は、前記管理計算機が、前記終了メッセージを受信したときに、前記コピーの終了を検出する手順を含むことを特徴とする請求項８に記載の方法。
【請求項１０】
前記保持された差分データを前記第２論理ボリュームにコピーする要求を送信する手順は、前記管理計算機が、前記第１論理ボリュームの容量、及び、前記第１論理ボリュームから前記第２論理ボリュームへのデータの転送速度に基づいて、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーに要する時間を算出する手順と、前記算出された時間が経過したときに、前記コピーの終了を検出する手順と、を含むことを特徴とする請求項８に記載の方法。
【請求項１１】
前記ネットワークは、前記一つ以上の計算機と前記一つ以上のストレージ装置との間のデータ転送を中継する複数のスイッチ装置を備え、
前記各スイッチ装置は、複数のポートと、前記管理計算機に接続される第３インタフェースと、前記管理計算機からの要求に従って前記各スイッチ装置の電源の投入及び遮断を制御する第３電源制御部と、を備え、
前記各ポートは、前記計算機又は前記ストレージ装置に接続され、
前記複数のスイッチ装置は、第１スイッチ装置及び第２スイッチ装置を含み、

前記第１仮想領域で稼動する前記仮想計算機は、前記第１スイッチ装置を経由して前記第１論理ボリュームへのデータの入出力を実行し、
前記方法は、さらに、
前記管理計算機が、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出してから、前記差分データの前記第２論理ボリュームへのコピーが終了するまでの間に、前記第２スイッチ装置に電源を投入する要求を送信する手順と、
前記管理計算機が、前記仮想計算機が前記第２スイッチ装置を経由して前記第２論理ボリュームへのデータの入出力を実行するように経路を切り替える要求を、前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信する前に送信する手順と、
前記仮想計算機が、前記第２仮想領域に移動した後、前記第２論理ボリュームへのデータ入出力を実行する手順は、前記第２スイッチ装置を経由して実行されることを特徴とする請求項８に記載の方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本願明細書に開示される技術は、計算機システムにおけるリソースの管理に関し、特に、パーティションで区切られた計算機システムにおけるアプリケーションの移動及び電源の制御に関する。
【背景技術】
【０００２】
情報処理システムに対しても地球環境への配慮が求められる中、仮想化によって計算機及びストレージ装置を集約し、それによってリソースの有効利用及び消費電力の削減を実現する技術が提案されている。
【０００３】
特許文献１には、仮想化されたサーバ、スイッチ及びストレージ装置を物理サーバ、物理スイッチ及び物理ストレージ装置に最適に配置することによって負荷を分散させる技術が開示されている。
【０００４】
特許文献２には、複数の物理サーバを備える計算機システムにおいて、仮想サーバを少数の物理サーバに集約し、残りの物理サーバの電源を遮断する技術が開示されている。これによって計算機システムの消費電力が削減される。
【０００５】
特許文献３には、仮想計算機と、その仮想計算機が使用する仮想ストレージ装置と、の対応を管理し、使用されていない仮想計算機及びそれに対応する仮想ストレージ装置の電源を遮断する技術が開示されている。これによって計算機システムの消費電力が削減される。
【特許文献１】特開２００７－４７９８６号公報
【特許文献２】特開２００７－３１０７９１号公報
【特許文献３】特開２００８－１０２６６７号公報
【発明の開示】
【発明が解決しようとする課題】
【０００６】
所定の利用方針に基づいてパーティションが区切られた計算機システムにおいては、アプリケーションを所望の時刻に所望のパーティションで実行することが求められる。そのために、仮想サーバ及び論理ボリュームを一つのパーティションから別のパーティションに移動させる場合がある。この移動の際に消費される電力を抑制するためには、これらの移動のタイミングを適切に制御する必要があるが、従来はそのための技術がなかった。
【課題を解決するための手段】

【０００７】
本願で開示する代表的な発明は、一つ以上の計算機と、ネットワークを介して前記一つ以上の計算機に接続される一つ以上のストレージ装置と、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される管理計算機と、を備える計算機システムであって、前記各計算機は、ハードウェア資源として、前記ネットワークに接続される第１インタフェースと、前記第１インタフェースに接続される第１プロセッサと、前記第１プロセッサに接続される第１メモリと、前記管理計算機からの要求に従って前記ハードウェア資源への電源の投入及び遮断を制御する電源制御部と、を備え、前記各計算機は、さらに、前記一つ以上の計算機のハードウェア資源に基づいて複数の仮想領域を提供する仮想化部を備え、前記複数の仮想領域は、第１仮想領域及び第２仮想領域を含み、前記第１仮想領域は、アプリケーションプログラムを実行する仮想計算機として動作し、前記各ストレージ装置は、前記計算機によって書き込まれたデータを格納する記憶領域を提供する記憶媒体と、前記記憶媒体へのデータの入出力を制御するコントローラと、前記管理計算機からの要求に従って前記各ストレージ装置の電源の投入及び遮断を制御する第２電源制御部と、を備え、前記一つ以上のストレージ装置の前記コントローラは、前記一つ以上のストレージ装置の前記記憶領域を複数の論理ボリュームとして前記計算機に提供し、前記複数の論理ボリュームは、第１論理ボリューム及び第２論理ボリュームを含み、前記第１論理ボリュームには、前記仮想計算機によって書き込まれたデータが格納され、前記管理計算機は、前記一つ以上の計算機及び前記一つ以上のストレージ装置に接続される第２インタフェースと、前記第２インタフェースに接続される第２プロセッサと、前記第２プロセッサに接続される第２メモリと、を備え、前記管理計算機は、前記第１論理ボリュームを含む前記ストレージ装置に、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を送信し、前記第１論理ボリュームを含む前記ストレージ装置は、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求を受信した後に前記仮想計算機から前記第１論理ボリュームへのデータの書き込み要求を受信すると、書き込みを要求されたデータを前記第１論理ボリュームに書き込まずに、差分データとして保持し、前記第１論理ボリュームに格納されたデータを前記第２論理ボリュームにコピーする要求に従って、前記第１論理ボリュームに格納されたデータを読み出して、前記第２論理ボリュームを含む前記ストレージ装置に送信し、前記管理計算機は、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出すると、前記保持された差分データを前記第２論理ボリュームにコピーする要求を、前記第１論理ボリュームを含む前記ストレージ装置に送信し、前記第１論理ボリュームに格納されたデータの前記第２論理ボリュームへのコピーの終了を検出してから、前記差分データの前記第２論理ボリュームへのコピーが終了するまでの間に、前記一つ以上の計算機のハードウェア資源のうち、前記第２仮想領域に割り当てられたハードウェア資源の電源を投入する要求を、前記第２仮想領域に割り当てられたハードウェア資源を含む前記計算機に送信し、前記差分データの前記第２論理ボリュームへのコピーが終了した後、前記仮想計算機を前記第２仮想領域に移動する要求を送信し、前記仮想計算機は、前記第２仮想領域に移動した後、前記第２論理ボリュームへのデータ入出力を実行することを特徴とする。

【発明の効果】
【０００８】
本発明の一実施形態によれば、アプリケーションの移動の際のリソースの電源投入タイミングを制御することによって、アプリケーションの移動の際に消費される電力を抑制することができる。

【発明を実施するための最良の形態】
【０００９】
本発明の実施形態を、図面を用いて詳細に説明する。

【００１０】
図１は、本発明の実施形態におけるパーティションの説明図である。

【００１１】
パーティションとは、ユーザの業務形態及び情報システムの利用方針に基づいて、情報システムを区切ることによって定義された領域である。各パーティションは、システムの設計者からは物理パーティションとして、ユーザからは論理パーティションとして認識される。
【００１２】
物理パーティションとは、物理的なリソースを区切ることによって定義された物理的な領域である。設計者は、各パーティションに割り当てる物理的なリソースの量を決めることができる。物理的なリソースの量とは、例えば、ＣＰＵのコア数、メモリの記憶領域の容量、ネットワークを構成するスイッチ装置のゾーンの大きさ、及び、ストレージ装置の記憶領域の容量等である。
【００１３】
論理パーティションとは、実行される業務（アプリケーション）が配置される領域である。ユーザは、アプリケーションが実行されるパーティション、及び、一つのパーティションで実行される複数のアプリケーションの組み合わせを決めることができる。
【００１４】
図２は、本発明の実施形態の計算機システムの構成を示すブロック図である。
【００１５】
本実施形態の計算機システムは、一つ以上のストレージ装置１００、一つ以上のサーバ１２０、一つ以上のファイバーチャネルスイッチ（ＦＣ－ＳＷ）１４０、管理サーバ１３０及び管理ネットワーク１５０を備える。
【００１６】
ストレージ装置１００は、サーバ１２０によって書き込まれたデータを格納する。具体的には、ストレージ装置１００は、ディスクドライブ１１０Ａ～１１０Ｃ及びコントローラ１０１を備える。
【００１７】
ディスクドライブ１１０Ａ～１１０Ｃは、サーバ１２０によって書き込まれたデータを格納する記憶媒体を含む記憶デバイスである。以下、ディスクドライブ１１０Ａ～１１０Ｃに共通する説明をする場合、これらを総称してディスクドライブ１１０とも記載する。
【００１８】
図２には三つのディスクドライブ１１０を示すが、ストレージ装置１００はいくつのディスクドライブ１１０を備えてもよい。複数のディスクドライブ１１０がＲＡＩＤ（Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙｓ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）を構成してもよい。
【００１９】
本実施形態のディスクドライブ１１０は、記憶媒体として磁気ディスクを備えるハードディスクドライブであるが、これはいかなる種類の装置によって置き換えられてもよい。例えば、ディスクドライブ１１０は、フラッシュメモリのような不揮発性の半導体記憶装置によって置き換えられてもよい。
【００２０】
コントローラ１０１は、ディスクドライブ１１０へのデータの書き込み及びディスクドライブ１１０からのデータの読み出しを制御する。コントローラ１０１は、相互に接続されたチャネルアダプタ（ＣＨＡ）１０２、ディスクアダプタ（ＤＡ）１０３、インタフェース（Ｉ／Ｆ）１０４及び電源制御部１０５を備える。
【００２１】
ＣＨＡ１０２は、ＦＣ－ＳＷ１４０のいずれかのポート（例えばポート１４１Ｂ）に接続され、サーバ１２０からのデータ入出力（Ｉ／Ｏ）要求、すなわち、データ書き込み要求及びデータ読み出し要求を処理する。
【００２２】
ＤＡ１０３は、ディスクドライブ１１０に接続され、ディスクドライブ１１０に対する

データの書き込み及びディスクドライブ１１０からのデータの読み出しを制御する。
【００２３】
ＣＨＡ１０２及びＤＡ１０３は、それぞれ、要求された処理を実行するためのＣＰＵ（図示省略）及びローカルメモリ（図示省略）を備えてもよい。
【００２４】
Ｉ／Ｆ１０４は、管理ネットワーク１５０に接続され、管理ネットワーク１５０を介して管理サーバ１３０と通信する。
【００２５】
電源制御部１０５は、ストレージ装置１００の電源の投入及び遮断（すなわちストレージ装置１００への電力の供給の開始及び停止）を制御する。より詳細には、電源制御部１０５は、Ｉ／Ｆ１０４を介して管理サーバ１３０から受信した制御情報に従って、ストレージ装置１００内の電源制御部１０５以外の部分への電力の供給を制御する。電源制御部１０５は、ストレージ装置１００のリソースの部分ごと（例えば、ディスクドライブ１１０ごと）の電力の供給を制御してもよい。電源制御部１０５は、例えば、いわゆるＢａｓｅｂｏａｒｄ Ｍａｎａｇｅｍｅｎｔ Ｃｏｎｔｒｏｌｌｅｒ（ＢＭＣ）によって実現されてもよい。
【００２６】
コントローラ１０１は、さらに、共有メモリ（図示省略）及びキャッシュメモリ（図示省略）を備えてもよい。共有メモリには、種々の制御情報が格納される。キャッシュメモリには、ディスクドライブ１１０に書き込まれるデータ及びディスクドライブ１１０から読み出されたデータが一時的に格納される。
【００２７】
コントローラ１０１は、ディスクドライブ１１０の記憶媒体によって実現される物理的な記憶領域を、複数の論理ボリューム１１１として管理する。図２に示す論理ボリューム１１１Ａ及び１１１Ｂは、それぞれ、複数の論理ボリューム１１１の一つである。コントローラ１０１は、任意の数の論理ボリューム１１１を管理することができる。
【００２８】
本実施形態の計算機システムは、複数のストレージ装置１００を備えてもよい。
【００２９】
サーバ１２０は、相互に接続されたＣＰＵ１２１、メモリ１２２、電源制御部１２３、ホストバスアダプタ（ＨＢＡ）１２４、Ｉ／Ｆ１２５及び仮想化部１２９を備える計算機である。
【００３０】
ＣＰＵ１２１は、メモリ１２２に格納されたプログラムを実行するプロセッサである。図２には一つのＣＰＵ１２１のみを示すが、サーバ１２０は複数のＣＰＵ１２１を備えてもよい。
【００３１】
メモリ１２２には、ＣＰＵ１２１によって実行されるプログラム及びＣＰＵ１２１によって参照されるデータが格納される。本実施形態のメモリ１２２には、少なくとも、オペレーティングシステム（ＯＳ）１２８、パス管理プログラム１２７及びアプリケーションプログラム１２６が格納される。
【００３２】
ＯＳ１２８は、サーバ１２０を管理する基本ソフトウェアであり、例えばＷｉｎｄｏｗｓ（登録商標）又はＵｎｉｘ（登録商標）のようなものであってもよい。後述のように、サーバ１２０上で複数のＯＳ１２８が実行される場合がある。その場合、同一の種類の複数のＯＳ１２８が実行されてもよいし、異なる種類の複数のＯＳ１２８が実行されてもよい。
【００３３】
パス管理プログラム１２７は、サーバ１２０から論理ボリューム１１１へのアクセス経路（パス）を制御する。

【００３４】
アプリケーションプログラム１２６は、種々の業務（アプリケーション）を実現するプログラムである。サーバ１２０のユーザは、所望のアプリケーションを実現するアプリケーションプログラム１２６をサーバ１２０にインストールし、実行することができる。アプリケーションプログラム１２６は、必要に応じて、論理ボリューム１１１へのデータＩ／Ｏ要求を発行する。
【００３５】
上記の各ソフトウェア（プログラム）は、ＣＰＵ１２１によって実行される。したがって、以下の説明において上記の各ソフトウェア（プログラム）が実行する処理は、実際にはＣＰＵ１２１によって実行される。
【００３６】
仮想化部１２９は、計算機システム内の一つ以上のサーバ１２０のリソース（すなわち、ＣＰＵ１２１及びメモリ１２２等）を用いて、複数の仮想的な領域（すなわち仮想的な計算機）を提供する。仮想化部１２９は、一つのサーバ１２０に複数のＯＳ１２８を実行させることができる。あるいは、仮想化部１２９は、複数のサーバ１２０を一つの仮想的な計算機として使用する、いわゆるクラスタリングを実現することもできる。
【００３７】
仮想化部１２９は、サーバ１２０内に実装されたハードウェアであってもよいし、メモリ１２２に格納されたプログラム（例えば、いわゆる仮想マシンモニタ又はハイパーバイザ等の仮想化ソフトウェア）であってもよい。仮想化部１２９がメモリ１２２に格納されたプログラムである場合、ＣＰＵ１２１が仮想化ソフトウェアを実行することによって仮想化部１２９の機能が実現される。
【００３８】
ＨＢＡ１２４は、ＦＣ－ＳＷ１４０のいずれかのポート（例えばポート１４１Ａ）に接続され、ストレージ装置１００との通信を実行するインタフェースである。
【００３９】
Ｉ／Ｆ１２５は、管理ネットワーク１５０に接続され、管理ネットワーク１５０を介して管理サーバ１３０と通信する。
【００４０】
電源制御部１２３は、サーバ１２０の電源を制御する。具体的には、電源制御部１２３は、Ｉ／Ｆ１２５を介して管理サーバ１３０から受信した制御情報に従って、サーバ１２０の電源の投入及び遮断を制御する。電源制御部１２３は、電源制御部１０５と同様のものであってもよい。電源制御部１２３は、サーバ１２０に含まれるリソースの部分ごと（例えばサーバ１２０が複数のＣＰＵ１２１を備える場合、ＣＰＵ１２１ごと）に電源の投入及び遮断を制御してもよい。
【００４１】
ＣＰＵ１２１は、アプリケーションプログラム１２６を実行し、必要に応じて、論理ボリューム１１１に対するデータ書き込み要求及びデータ読み出し要求を、ＨＢＡ１２４を介してストレージ装置１００に送信する。これらの要求の送信先（言い換えると、これらの要求によるデータＩ／Ｏに使用される論理パス）は、パス管理プログラム１２７によって制御される。
【００４２】
本実施形態の計算機システムは、複数のサーバ１２０を備えてもよい。
【００４３】
ＦＣ－ＳＷ１４０は、サーバ１２０とストレージ装置１００との間のデータＩ／Ｏを中継するネットワークを構成する。ＦＣ－ＳＷ１４０は、サーバ１２０とストレージ装置１００との間のデータＩ／Ｏ経路を切り替えることができる。本実施形態において、サーバ１２０とストレージ装置１００との間のデータＩ／Ｏは、ファイバーチャネル（ＦＣ）プロトコルに基づいてやり取りされる。
【００４４】

ＦＣ－ＳＷ１４０は、複数のポート１４１（図２の例では、ポート１４１Ａ～ポート１４１Ｄ）、Ｉ／Ｆ１４２及び電源制御部１４３を備える。

【００４５】

各ポートは、サーバ１２０のＨＢＡ１２４又はストレージ装置１００のＣＨＡ１０２に接続される。

【００４６】

ＦＣ－ＳＷ１４０は、各ポート１４１間の接続を設定することによって、サーバ１２０とストレージ装置１００との間のデータ通信経路を設定することができる。さらに、ＦＣ－ＳＷ１４０は、互いに独立したいわゆるゾーンを設定することができる。

【００４７】

Ｉ／Ｆ１４２は、管理ネットワーク１５０に接続され、管理ネットワーク１５０を介して管理サーバ１３０と通信する。

【００４８】

電源制御部１４３は、ＦＣ－ＳＷ１４０の電源を制御する。具体的には、電源制御部１４３は、Ｉ／Ｆ１４２を介して管理サーバ１３０から受信した制御情報に従って、ＦＣ－ＳＷ１４０の電源の投入及び遮断を制御する。電源制御部１４３は、電源制御部１０５と同様のものであってもよい。電源制御部１４３は、ＦＣ－ＳＷ１４０に含まれるリソースの部分ごと（例えばポート１４１ごと）に電源の投入及び遮断を制御してもよい。

【００４９】

本実施形態の計算機システムは、複数のＦＣ－ＳＷ１４０を備えてもよい。

【００５０】

管理サーバ１３０は、相互に接続されたＣＰＵ１３１、メモリ１３２、データベース１３３、及びＩ／Ｆ１３４を備える計算機である。

【００５１】

ＣＰＵ１３１は、メモリ１３２に格納されたプログラムを実行するプロセッサである。

【００５２】

メモリ１３２には、ＣＰＵ１３１によって実行されるプログラム及びＣＰＵ１３１によって参照されるデータが格納される。本実施形態のメモリ１３２には、少なくとも、管理プログラム１３５が格納される。

【００５３】

データベース１３３には、計算機システムを管理するための情報が格納される。データベース１３３は、例えば、管理サーバ１３０に接続された（又は内蔵された）ディスクドライブに格納されてもよい。

【００５４】

本実施形態のデータベース１３３には、管理テーブル１３６が格納される。管理テーブル１３６の全部又は一部が、必要に応じてメモリ１３２にコピーされ、ＣＰＵ１３１によって参照されてもよい。管理テーブル１３６の内容については後述する（図１６参照）。

【００５５】

Ｉ／Ｆ１３４は、管理ネットワーク１５０に接続され、管理ネットワーク１５０を介してストレージ装置１００、サーバ１２０及びＦＣ－ＳＷ１４０と通信する。例えば、ストレージ装置１００、サーバ１２０及びＦＣ－ＳＷ１４０の各々の電源を制御するための信号は、Ｉ／Ｆ１３４から管理ネットワーク１５０を介して送信される。

【００５６】

管理ネットワーク１５０は、いかなる種類のものであってもよい。典型的には、管理ネットワーク１５０は、いわゆるＬＡＮ（Ｌｏｃａｌ　Ａｒｅａ　Ｎｅｔｗｏｒｋ）のようなＩＰ（Ｉｎｔｅｒｎｅｔ　Ｐｒｏｔｏｃｏｌ）ネットワークである。その場合、Ｉ／Ｆ１０４、Ｉ／Ｆ１２５及びＩ／Ｆ１３４は、いわゆるネットワークインタフェースカードであってもよい。

【００５７】

図３は、本発明の実施形態における物理パーティションの第１の例の説明図である。

【００５８】
図３は、三つの物理パーティション、すなわち、パーティション１＿３００Ａ、パーティション２＿３００Ｂ及びパーティション３＿３００Ｃが定義される例を示す。この例において、各パーティションは、一つ以上のサーバ１２０、一つ以上のＦＣ－ＳＷ１４０及び一つ以上のストレージ装置１００を含む。
【００５９】
具体的には、パーティション１＿３００Ａは、三つのサーバ１２０、三つのＦＣ－ＳＷ１４０及び三つのストレージ装置１００を含む。これらの三つのサーバ１２０は、仮想化部１２９によって実現された一つのクラスタである。
【００６０】
同様に、パーティション２＿３００Ｂは、二つのサーバ１２０、二つのＦＣ－ＳＷ１４０及び二つのストレージ装置１００を含む。これらの二つのサーバ１２０は、仮想化部１２９によって実現された一つのクラスタである。
【００６１】
パーティション３＿３００Ｃは、一つのサーバ１２０、一つのＦＣ－ＳＷ１４０及び一つのストレージ装置１００を含む。
【００６２】
なお、図３において、サーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００の間の接続の図示は省略されている。
【００６３】
各パーティションに含まれる一つ以上のサーバ１２０では、他のパーティションとは独立にＯＳ１２８が実行され、そのＯＳ１２８の上で、他のパーティションとは独立にアプリケーションプログラム１２６が実行される。各パーティションで実行されるアプリケーションプログラム１２６は、そのパーティションに含まれる一つ以上のＦＣ－ＳＷ１４０を介して、そのパーティションに含まれる一つ以上のストレージ装置１００内の論理ボリュームへのデータＩ／Ｏを実行する。
【００６４】
図４は、本発明の実施形態における物理パーティションの第２の例の説明図である。
【００６５】
図４は、一つのサーバ１２０、一つのＦＣ－ＳＷ１４０及び一つのストレージ装置１００からなる計算機システムに、三つの物理パーティション、すなわち、パーティション１＿３００Ｄ、パーティション２＿３００Ｅ及びパーティション３＿３００Ｆが定義される例を示す。
【００６６】
具体的には、図４に示すサーバ１２０は、６個のＣＰＵ１２１を備え、それらのうち３個がパーティション１＿３００Ｄに、２個がパーティション２＿３００Ｅに、１個がパーティション３＿３００Ｆに割り当てられる。
【００６７】
図４に示すＦＣ－ＳＷ１４０は、１２個のポート１４１を備え、それらのうち４個がパーティション１＿３００Ｄに割り当てられ、残りの８個のうち４個がパーティション２＿３００Ｅに割り当てられ、残りの４個がパーティション３＿３００Ｆに割り当てられる。各パーティションに割り当てられた４個のポート１４１は、ゾーン４０１を形成する。
【００６８】
図４に示すストレージ装置１００が提供する記憶領域のうち、一部（例えば、一部のボリュームプール４０２）がパーティション１＿３００Ｄに割り当てられ、別のボリュームプール４０２がパーティション２＿３００Ｅに割り当てられ、残りがパーティション３＿３００Ｆに割り当てられる。なお、ボリュームプール４０２とは、一つ以上の論理ボリューム１１１からなる記憶領域の管理単位である。
【００６９】
この例において、各パーティションに割り当てられた１個以上のＣＰＵ１２１において

、ＯＳ１２８が他のパーティションとは独立に実行され、それらのＯＳ１２８の上でアプリケーションプログラム１２６が他のパーティションとは独立に実行される。各パーティションで実行されるアプリケーションプログラムは、そのパーティションに割り当てられた記憶領域へのデータＩ／Ｏを実行する。

【００７０】

上記のように、パーティションは、複数の装置のリソースを結合することによって定義されてもよいし（図３参照）、一つの装置のリソースを分割することによって定義されてもよい（図４参照）。いずれの場合も、システムの設計者は、各パーティションを、物理的なリソースを区切る物理パーティションとして認識する。

【００７１】

図５は、本発明の実施形態における論理パーティションの例の説明図である。

【００７２】

計算機システムのユーザは、各パーティションを、実行する業務が配置される領域として認識する。以下の説明において、「業務」を「アプリケーション」と記載する。例えば、図３に示すパーティション１＿３００Ａ、パーティション２＿３００Ｂ及びパーティション３＿３００Ｃが、それぞれ、論理パーティション、すなわち、パーティション１＿３００Ｇ、パーティション２＿３００Ｈ及びパーティション３＿３００Ｉとしてユーザに認識されてもよい。あるいは、図４に示すパーティション１＿３００Ｄ、パーティション２＿３００Ｅ及びパーティション３＿３００Ｆが、それぞれ、論理パーティション、すなわち、パーティション１＿３００Ｇ、パーティション２＿３００Ｈ及びパーティション３＿３００Ｉとしてユーザに認識されてもよい。

【００７３】

図５の例では、パーティション１＿３００Ｇで三つの業務（アプリケーション）５００が実行され、パーティション２＿３００Ｈで二つのアプリケーション５００が実行され、パーティション３＿３００Ｉで一つのアプリケーション５００が実行される。

【００７４】

図６は、本発明の実施形態におけるアプリケーション５００の構成の説明図である。

【００７５】

図５に示す各アプリケーション５００は、図６に示すように、仮想サーバ（ＶＭ）６０１、論理パス６０２及び論理ボリューム（ＬＵ）６０３によって構成される。

【００７６】

仮想サーバ６０１は、仮想化部１２９によって実現される仮想的な計算機である。

【００７７】

論理パス６０２は、仮想サーバ６０１が論理ボリューム６０３にアクセスするために使用される経路である。論理パス６０２は、ＨＢＡ１２４からＦＣ－ＳＷ１４０を経由してストレージ装置１００に至る物理的な経路によって実現される。

【００７８】

論理ボリューム６０３は、ストレージ装置１００が仮想サーバ６０１に提供する論理的な記憶領域である。仮想サーバ６０１は、一つの論理ボリューム６０３を一つの記憶デバイスとして認識する。例えば、一つの論理ボリューム１１１が一つの論理ボリューム６０３として提供されてもよいし、複数の論理ボリューム１１１が一つの論理ボリューム６０３として提供されてもよい。あるいは、論理ボリューム６０３の記憶領域へのデータ書き込み要求を受信したときに、その記憶領域に、論理ボリューム１１１の記憶領域が割り当てられてもよい。

【００７９】

仮想サーバ６０１上でＯＳ１２８が実行され、そのＯＳ１２８上でアプリケーションプログラム１２６が実行される。そのアプリケーションプログラム１２６は、論理パス６０２を介して、論理ボリューム６０３へのデータＩ／Ｏを実行する。これによって、アプリケーション５００が実現される。

【００８０】

各アプリケーション５００は、一つのパーティションから別のパーティションへ移動することができる。このようなアプリケーション５００の移動については後述する。

【００８１】

図７は、本発明の実施形態の全体システム構成の説明図である。

【００８２】

図７に示すパーティション１＿３００Ａ、パーティション２＿３００Ｂ及びパーティション３＿３００Ｃは、図３に示したものと同様の物理パーティションとしてシステム管理者に認識される。

【００８３】

一方、ユーザは、パーティション１＿３００Ａ、パーティション２＿３００Ｂ及びパーティション３＿３００Ｃを論理パーティションとして認識する。パーティション１＿３００Ａには、二つのアプリケーション（ＡＰＰ）５００（すなわち、アプリケーション５００Ａ及び５００Ｂ）が配置される。パーティション２＿３００Ｂには、二つのアプリケーション５００（すなわち、アプリケーション５００Ｃ及び５００Ｄ）が配置される。パーティション３＿３００Ｃには一つのアプリケーション５００（すなわち、アプリケーション５００Ｅ）が配置される。

【００８４】

アプリケーション５００Ａを構成する仮想サーバ６０１、論理パス６０２及び論理ボリューム６０３を、それぞれ、仮想サーバ６０１Ａ、論理パス６０２Ａ及び論理ボリューム６０３Ａと記載する。同様に、アプリケーション５００Ｂは、仮想サーバ６０１Ｂ、論理パス６０２Ｂ及び論理ボリューム６０３Ｂによって構成される。アプリケーション５００Ｃは、仮想サーバ６０１Ｃ、論理パス６０２Ｃ及び論理ボリューム６０３Ｃによって構成される。アプリケーション５００Ｄは、仮想サーバ６０１Ｄ、論理パス６０２Ｄ及び論理ボリューム６０３Ｄによって構成される。アプリケーション５００Ｅは、仮想サーバ６０１Ｅ、論理パス６０２Ｅ及び論理ボリューム６０３Ｅによって構成される。

【００８５】

以下の説明において、アプリケーション５００Ａ～５００Ｅに共通する説明をする場合には、これらを総称してアプリケーション５００と記載する。仮想サーバ６０１Ａ～６０１Ｅに共通する説明をする場合、これらを総称して仮想サーバ６０１と記載する。論理パス６０２Ａ～６０２Ｅに共通する説明をする場合、これらを総称して論理パス６０２と記載する。論理ボリューム６０３Ａ～６０３Ｅに共通する説明をする場合、これらを総称して論理ボリューム６０３と記載する。

【００８６】

管理サーバ１３０の管理プログラム１３５は、各パーティションを管理する。具体的には、管理サーバ１３０は、管理ネットワーク１５０を介して各パーティションに含まれるサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００に接続され、それらの装置を管理する。

【００８７】

さらに、管理プログラム１３５は、各パーティションへのハードウェアリソースの割り当て、及び、各パーティションへの各アプリケーション５００の配置を管理する。例えば、管理プログラム１３５は、アプリケーション５００のパーティション間の移動（具体的には、論理ボリューム６０３の移動、論理パス６０２の切り替え、及び、仮想サーバ６０１の移動）を制御することができる。

【００８８】

ここで、本実施形態において実行されるアプリケーション５００のパーティション間の移動について説明する。

【００８９】

管理プログラム１３５は、アプリケーション５００をあるパーティションから別のパーティションに移動することができる。このような移動は、種々の目的、例えば、消費電力の削減又は負荷分散のために実行される。

【００９０】
具体的には、例えば、図７に示すアプリケーション５００Ａ～５００Ｅが夜間にはほとんど使用されない場合、アプリケーション５００Ａ～５００Ｅが夜間には一つのパーティションに配置され、昼間には複数のパーティションに分散されるように、アプリケーション５００の移動が制御されてもよい。例えば、昼間（例えば８：００から翌日の０：００まで）は図７に示すように各アプリケーション５００が配置され、夜間（例えば０：００から８：００まで）はアプリケーション５００Ａ～５００Ｅがすべてパーティション３＿３００Ｃに配置されてもよい。
【００９１】
この場合、０：００にアプリケーション５００Ａ及び５００Ｂをパーティション１＿３００Ａからパーティション３＿３００Ｃに、同じく０：００にアプリケーション５００Ｃ及び５００Ｄをパーティション２＿３００Ｂからパーティション３＿３００Ｃに移動する処理を管理プログラム１３５が実行する。その後、パーティション１＿３００Ａ及びパーティション２＿３００Ｂの電源を遮断する（すなわち、それらのパーティションに割り当てられた物理的なリソースへの電力の供給を遮断する）ことによって、計算機システムの消費電力を削減することができる。
【００９２】
その後、８：００にアプリケーション５００Ａ及び５００Ｂをパーティション３＿３００Ｃからパーティション１＿３００Ａに、同じく８：００にアプリケーション５００Ｃ及び５００Ｄをパーティション３＿３００Ｃからパーティション２＿３００Ｂに移動する処理を管理プログラム１３５が実行する。０：００の移動の後に移動先のパーティションの電源が遮断されている場合、８：００の移動の前にそれらのパーティションの電源を投入する必要がある。
【００９３】
図８は、本発明の実施形態において実行されるアプリケーション５００のパーティション間の移動処理を示すフローチャートである。
【００９４】
管理プログラム１３５は、時刻等をトリガーにして、アプリケーション５００の移動処理を開始する（ステップ８０１）。例えば、上記の図７の例の場合、０：００及び８：００に移動処理が開始されてもよい。あるいは、移動処理は、０：００及び８：００から終了処理に要する時間を減算した時刻開始されてもよい。その場合、移動処理は、０：００及び８：００に終了する。
【００９５】
次に、管理プログラム１３５は、これから移動するアプリケーション５００が現在どのパーティションに配置されているかを確認する（ステップ８０２）。この確認のために、後に説明する管理テーブル１３６が参照される。
【００９６】
次に、管理プログラム１３５は、移動先のパーティションを選択する（ステップ８０３）。上記の図７の例のようにあらかじめ移動先のパーティションが決定されている場合、そのパーティションが選択されてもよい。移動先として選択可能なパーティションが複数存在する場合には、それらのうち一つが選択されてもよい。
【００９７】
次に、管理プログラム１３５は、移動先として選択されたパーティションの状態を確認する（ステップ８０４）。具体的には、管理プログラム１３５は、移動先として選択されたパーティションに既に配置されているアプリケーション５００の数、それらのアプリケーション５００によって既に使用されている論理ボリューム６０３の容量、及び、そのパーティションを使用可能な時間帯、等の情報を参照する。ステップ８０４で実行される詳細な処理については、後に図２７を参照して説明する。
【００９８】
次に、管理プログラム１３５は、ステップ８０４で参照した情報に基づいて、選択され

たパーティションへのアプリケーション５００の移動が可能であるか否かを判定する（ステップ８０５）。ステップ８０４の参照及びステップ８０５の判定のために、管理テーブル１３６が参照される。

【００９９】

ステップ８０５において、選択されたパーティションへのアプリケーション５００の移動が可能でないと判定された場合、管理プログラム１３５は、ステップ８０３に戻って、別のパーティションを移動先として選択する。なお、選択可能な全てのパーティションについて、移動が可能でないと判定された場合、管理プログラム１３５は、アプリケーション５００の移動を実行せずに図８の処理を終了してもよい。

【０１００】

ステップ８０５において、選択されたパーティションへのアプリケーション５００の移動が可能であると判定された場合、管理プログラム１３５は、アプリケーション５００の移動を実行する（ステップ８０６）。この移動については、後に図９及び図１０を参照して詳細に説明する。

【０１０１】

次に、管理プログラム１３５は、ステップ８０６の移動が反映されるように、アプリケーション５００の配置に関する情報（具体的には、管理テーブル１３６に含まれる情報）を更新する（ステップ８０７）。この更新については、後に管理テーブル１３６を参照して詳細に説明する。

【０１０２】

以上で、アプリケーション５００の移動処理が終了する（ステップ８０８）。

【０１０３】

図９は、本発明の実施形態において実行されるアプリケーション５００のパーティション間の移動処理の詳細な手順を示す説明図である。

【０１０４】

図９は、例として、図７の計算機システムにおいてアプリケーション５００Ａをパーティション１＿３００Ａからパーティション２＿３００Ｂに移動する手順を示す。ただし、説明を簡単にするため、説明に必要のない部分の図示は省略されている。

【０１０５】

図９（Ａ）は、初期状態、すなわち、アプリケーション５００Ａが移動する前の状態を示す。この状態において、仮想サーバ６０１Ａは、パーティション１＿３００Ａ内のサーバ１２０で稼動している。論理ボリューム６０３Ａは、パーティション１＿３００Ａ内のストレージ装置１００によって管理されている。仮想サーバ６０１Ａが論理ボリューム６０３Ａにアクセスするために使用される論理パス６０２Ａは、パーティション１＿３００Ａ内のＦＣ－ＳＷ１４０を経由する。

【０１０６】

この状態において、アプリケーション５００Ａの移動が開始されると、最初に、論理ボリューム６０３Ａの移動が実行される。具体的には、論理ボリューム６０３Ａの複製がパーティション２＿３００Ｂ内のストレージ装置１００に作成される。

【０１０７】

図９（Ｂ）は、論理ボリューム６０３Ａの移動が終了した状態を示す。図９（Ｂ）に示す論理ボリューム６０３Ｆは、論理ボリューム６０３Ａの複製である。この複製を作成するために、論理ボリューム６０３Ａに格納された全てのデータを読み出して、それを論理ボリューム６０３Ｆにコピーする処理が実行される。図９（Ｂ）の状態において、コピーは終了しているが、仮想サーバ６０１Ａはまだ論理ボリューム６０３Ａにアクセスし、論理ボリューム６０３Ｆはまだ使用されない。

【０１０８】

次に、論理パス６０２Ａの移動が実行される。具体的には、論理パス６０２Ａがパーティション２＿３００Ｂ内のＦＣ－ＳＷ１４０を経由して論理ボリューム６０３Ｆへのアクセスを可能とするように、論理パス６０２Ａの設定が切り替えられる。この切り替えは、

パーティション１＿３００Ａ内のサーバ１２０のパス管理プログラム１２７が、論理ボリューム６０３Ｆを新たに論理ボリューム６０３Ａとして認識し、さらに、論理ボリューム６０３Ａへのデータ Ｉ／Ｏ要求の送信先を、パーティション１＿３００Ａ内のＦＣ－ＳＷ１４０のポート１４１から、パーティション２＿３００Ｂ内のＦＣ－ＳＷ１４０のポート１４１に変更することによって実行されてもよい。

【０１０９】

図９（Ｃ）は、論理パス６０２Ａの移動が終了した状態を示す。この状態において、論理パス６０２Ａは、パーティション１＿３００Ａ内の仮想サーバ６０１Ａがパーティション２＿３００Ｂ内のＦＣ－ＳＷ１４０を経由して論理ボリューム６０３Ｆにアクセスできるように設定されている。

【０１１０】

次に、仮想サーバ６０１Ａの移動が実行される。この移動は、例えば、パーティション１＿３００Ａ内のサーバ１２０のメモリ１２２のイメージを、パーティション２＿３００Ｂ内のサーバ１２０のメモリ１２２にコピーすることによって実行される。このような移動は、仮想化部１２９の機能によって実現されてもよい。

【０１１１】

図９（Ｄ）は、仮想サーバ６０１Ａの移動が終了した状態を示す。この状態において、パーティション２＿３００Ｂ内のサーバ１２０に移動した仮想サーバ６０１Ａは、パーティション２＿３００Ｂ内のＦＣ－ＳＷ１４０を経由する論理パス６０２Ａを用いて、論理ボリューム６０３Ｆ（すなわち新たな論理ボリューム６０３Ａ）へのデータ Ｉ／Ｏを実行することができる。これによって、アプリケーション５００Ａがパーティション２＿３００Ｂに移動する。

【０１１２】

後述するように、論理ボリューム６０３Ａの移動は、アプリケーション５００Ａの稼動を停止することなく実行することができる。一方、論理パス６０２Ａ及び仮想サーバ６０１Ａの移動を実行するためには、アプリケーション５００Ａの稼動を停止する必要がある。しかし、その停止時間は十分に短いため、ユーザの利便性を損ねることなくアプリケーション５００Ａを移動することができる。

【０１１３】

図１０は、本発明の実施形態において実行されるアプリケーション５００のパーティション間の移動処理の詳細な手順を示すフローチャートである。

【０１１４】

図１０に示す移動処理は、図８のステップ８０６において実行を開始される（ステップ１００１）。すなわち、この処理の実行が開始される時点で、移動先のパーティションが既に選択されている。選択されたパーティションを識別する情報が引数として与えられる。

【０１１５】

次に、管理プログラム１３５は、選択されたパーティションへのアプリケーション５００の移動を実行する（ステップ１００２）。具体的には、管理プログラム１３５は、図９に示したように、最初に論理ボリューム６０３を移動し、次に論理パス６０２を移動し、最後に仮想サーバ６０１を移動する。ステップ１００２において実行される詳細な処理については、後に図１７等を参照して説明する。

【０１１６】

次に、管理プログラム１３５は、移動の結果を反映するように管理テーブル１３６を更新する（ステップ１００３）。

【０１１７】

以上で、図１０に示す移動の処理が終了する（ステップ１００４）。その後、処理は図８のフローチャートに戻り、ステップ８０７以降の処理が実行される。

【０１１８】

次に、論理ボリューム６０３の移動の詳細な手順を説明する。

【０１１９】
図１１は、本発明の実施形態において実行される論理ボリューム６０３の移動の説明図である。
【０１２０】
上記のように、論理ボリューム６０３の移動は、アプリケーション５００を停止することなく実行される。このため、論理ボリューム６０３の移動のためのコピーが開始された後、それが終了するまでの間に、移動中の論理ボリューム６０３へのデータＩ／Ｏが実行される場合がある。本実施形態では、コピーが開始された後のデータＩ／Ｏは、論理ボリューム６０３に反映されず、差分ストックとして保持される。そして、論理ボリューム６０３の全データのコピーが終了した後、差分ストックとして保持されているデータＩ／Ｏがコピー先の論理ボリューム６０３に反映される。
【０１２１】
例えば、図９に示すように、論理ボリューム６０３Ａのデータが論理ボリューム６０３Ｆにコピーされる場合について説明する。ストレージ装置１００のコントローラ１０１は、コピーが開始された後に論理ボリューム６０３Ａへのデータ書き込み要求を受信すると、そのデータを論理ボリューム６０３Ａ及び論理ボリューム６０３Ｆのいずれにも書き込まずに差分ストックとして保持する。この差分ストックは、例えばストレージ装置１００内のいずれかの論理ボリューム１１１に格納されてもよい。
【０１２２】
論理ボリューム６０３Ａに格納されている全てのデータの論理ボリューム６０３Ｆへのコピー（以下、全体コピーと記載する）が終了すると、次に、差分ストックとして格納されたデータを、論理ボリューム６０３Ｆに書き込む。差分ストックとして格納されたデータのコピー先への書き込みを、以下、差分コピーと記載する。
【０１２３】
このような全体コピー及び差分コピーは、コントローラ１０１によって制御される。例えば、コントローラ１０１内の制御プロセッサ（図示省略）が、制御メモリ（図示省略）に格納されたＩ／Ｏ制御プログラム（図示省略）を実行することによって全体コピー及び差分コピーが実現されてもよい。
【０１２４】
本実施形態では、全体コピー処理の終了を契機として、サーバ１２０及びＦＣ－ＳＷ１４０の電源が投入される。これは、計算機システムによる消費電力を削減するためである。電源投入のタイミングの制御については、後に詳細に説明する（図１４、図１５等参照）。
【０１２５】
なお、従来のストレージ装置では、全体コピーと差分コピーとの組が一つのコピー処理として扱われるため、ストレージ装置の外部の装置（例えば管理サーバ）は、全体コピーのみが終了したことを知ることができなかった。このため、全体コピー処理の終了（又は差分コピー処理の開始）とその他の処理とを関連付けて制御することができなかった。これに対して、本実施形態では、上記のような制御を実現するため、全体コピー処理が終了したことを示すメッセージをストレージ装置１００が管理サーバ１３０に送信してもよいし（後述する図２０参照）、管理サーバ１３０が全体コピー処理の終了時刻を算出してもよい。
【０１２６】
図１２は、本発明の実施形態において実行される論理ボリューム６０３の移動の別の例の説明図である。
【０１２７】
ストレージ装置１００がボリュームコピー機能（いわゆるリモートコピー機能又はローカルコピー機能）を持つ場合、そのボリュームコピー機能を用いて論理ボリューム６０３を移動することができる。
【０１２８】

例えば、論理ボリューム６０３Ａを正ボリューム、論理ボリューム６０３Ｆを副ボリュームとするボリュームペアが構成される場合、論理ボリューム６０３Ａに格納されたデータの更新が、論理ボリューム６０３Ｆにも反映される。

【０１２９】

論理ボリューム６０３Ａの更新が直ちに論理ボリューム６０３Ｆに反映されるように設定されている場合、論理ボリューム６０３Ｆには、ほぼ常に論理ボリューム６０３Ａと同一のデータが格納されている。このため、図９のようにアプリケーション５００Ａが移動する場合の論理ボリューム６０３Ａの移動に要する時間はほぼ０となる。

【０１３０】

一方、論理ボリューム６０３Ａの更新が直ちに論理ボリューム６０３Ｆに反映されないように設定されている場合もある。例えば、定期的に更新が反映されるように設定されていてもよいし、論理ボリューム６０３Ａと論理ボリューム６０３Ｆとの間のデータ転送経路のトラフィック量が所定の閾値より小さいときに更新が反映されるように設定されていてもよい。このような場合、図９のようなアプリケーション５００Ａの移動における論理ボリューム６０３Ａの移動は、前回の反映が実行された後で更新された論理ボリューム６０３Ａのデータ（すなわち差分データ）のみがコピーされる。

【０１３１】

アプリケーション５００Ａがパーティション１＿３００Ａに配置されているときには論理ボリューム６０３Ａが、アプリケーション５００Ａがパーティション２＿３００Ｂに配置されているときには論理ボリューム６０３Ｆが使用される。

【０１３２】

上記のように、ボリュームコピー機能を使用することによって、図１１に示すような全体コピー及び差分コピーを実行する場合と比較して短時間で論理ボリューム６０３の移動を実行することができる。

【０１３３】

図１３は、本発明の実施形態における物理パーティションの第３の例の説明図である。

【０１３４】

図１３の例では、図３と同様に、一つ以上のサーバ１２０が各パーティションに割り当てられるが、図４と同様に、一つのＦＣ－ＳＷ１４０及び一つのストレージ装置１００が複数のパーティションに共有される。

【０１３５】

このような場合であっても、ＦＣ－ＳＷ１４０及びストレージ装置１００が部分ごとに電源を制御することができれば、上記と同様の電源制御を実行することができる。例えば、ストレージ装置１００が複数のディスクドライブ１１０を備え、かつ、ディスクドライブ１１０ごと（又は、複数のディスクドライブからなるＲＡＩＤグループ（図示省略）ごと）の電源の投入、切断を制御できる場合、移動元の論理ボリューム６０３を格納するディスクドライブ１１０及び移動先の論理ボリューム６０３が属するディスクドライブ１１０の電源を適切なタイミングで制御することによって、後述するように消費電力を削減することができる。

【０１３６】

ただし、計算機システム全体の消費電力に対して、上記のようにストレージ装置１００等の部分ごとの制御によって削減される電力の量が十分に小さい場合、そのような部分ごとの制御による消費電力削減効果は小さい。このような場合、複数のパーティションによって共有されるストレージ装置１００等の電源は常時投入し、サーバ１２０のみの電源を制御してもよい。

【０１３７】

図１４は、本発明の実施形態において実行されるアプリケーション５００の移動及び電源制御の説明図である。

【０１３８】

図１４は、図９を参照して説明したアプリケーション５００の移動の手順に加えて、そ

の移動の際に実行される電源制御の手順を示す。ただし、図９はアプリケーション５００がパーティション１＿３００Ａからパーティション２＿３００Ｂに移動する例を示したが、図１４は、アプリケーション５００がパーティション２＿３００Ｂからパーティション１＿３００Ａに移動する例を示す。例えば、図９の処理によってパーティション２＿３００Ｂに移動したアプリケーション５００Ａが、パーティション１＿３００Ａに戻るときに図１４の処理が実行される。なお、図９の処理が実行される場合にも図１４と同様の電源制御が実行されてよい。

【０１３９】

図１４に示すサーバ１２０等のハードウェア及び仮想サーバ６０１Ａ等のアプリケーションの構成は図９と同様であるため説明を省略する。ただし、図１４のパーティション１＿３００Ａのサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００に表示された「サーバＰ１」、「ＦＣ－ＳＷＰ１」及び「ストレージＰ１」、並びに、パーティション２＿３００Ｂのサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００に表示された「サーバＰ２」、「ＦＣ－ＳＷＰ２」及び「ストレージＰ２」は、それぞれ、各ハードウェアの識別子である。これらの識別子は、後述するテーブルに登録される（図１６等参照）。

【０１４０】

図１４の処理が実行される直前の時点において、パーティション１＿３００Ａに割り当てられたサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００の電源は切断されている。

【０１４１】

図１４の例では、最初に、移動先であるパーティション１＿３００Ａのストレージ装置１００の電源が投入される。

【０１４２】

次に、パーティション２＿３００Ｂからパーティション１＿３００Ａへの論理ボリューム６０３の移動のための全体コピーが開始される。図１４の例では、論理ボリューム６０３Ｆに格納されている全データが論理ボリューム６０３Ａにコピーされる。このコピーも、図１１に示したように実行される。すなわち、この全体コピーが開始された後の論理ボリューム６０３Ｆの更新は禁止され、その更新の内容は差分ストックとして保持される。

【０１４３】

次に、全体コピーが終了する。

【０１４４】

全体コピーの終了を契機として、パーティション１＿３００Ａのサーバ１２０及びＦＣ－ＳＷ１４０の電源が投入される。本実施形態では仮想サーバ６０１Ａを移動する前に論理パス６０２Ａを移動する必要があるため、サーバ１２０の起動処理がＦＣ－ＳＷ１４０の起動処理より先に終了する必要はない。このため、サーバ１２０の電源を投入する前にＦＣ－ＳＷ１４０の電源を投入してもよい。しかし、一般に、サーバ１２０の起動に要する時間は、ＦＣ－ＳＷ１４０の起動に要する時間より長い。このため、図１４に示すように、先にサーバ１２０の電源を投入してもよい。

【０１４５】

次に、論理ボリューム６０３Ｆから論理ボリューム６０３Ａへの差分コピーが実行される。具体的には、差分ストックとして保持されているデータが論理ボリューム６０３Ａに書き込まれる。

【０１４６】

上記のような論理ボリューム６０３の移動は、例えば、従来のオンラインマイグレーションの機能によって実現される。

【０１４７】

差分コピーが終了すると、論理パス６０２Ａがパーティション１＿３００Ａに移動し、続いて、仮想サーバ６０１Ａがパーティション１＿３００Ａに移動する。

【０１４８】

図１５は、本発明の実施形態において実行されるアプリケーション５００の移動及び電源制御のタイミングの説明図である。

【０１４９】

具体的には、図１５は、図１４に示した各処理の実行のタイミング及びその実行に要する時間を示す。図１５に示された片方向の矢印の基点（図１５の例では矢印の左端）、先端（図１５の例では矢印の右端）及び長さは、それぞれ、各処理の開始時刻、終了時刻及び処理時間に対応する。

【０１５０】

例えば、矢印１５０１は、ストレージ装置１００の起動処理を示す。矢印１５０１の基点は、ストレージ装置１００の電源が投入された時刻（時刻１５１１）を示し、矢印１５０１の先端は、ストレージ装置１００の起動処理が終了した時刻（時刻１５１２）を示す。同様に、矢印１５０２は論理ボリューム６０３Ｆの全体コピー処理を、矢印１５０３は論理ボリューム６０３Ｆの差分コピー処理を、矢印１５０４はサーバ１２０の起動処理を、矢印１５０５は、ＦＣ－ＳＷ１４０の起動処理を、矢印１５０６は論理パス６０２Ａの移動処理を、矢印１５０７は仮想サーバ６０１Ａの移動処理を示す。

【０１５１】

図１４にも示したように、最初にストレージ装置１００の電源が投入され（時刻１５１１）、ストレージ装置１００が起動すると（時刻１５１２）、論理ボリューム６０３Ｆの全体コピーが実行される。全体コピーに要する時間は、論理ボリューム６０３Ｆに格納されているデータ量、及び、コピーに使用されるデータ転送経路の転送性能に依存するが、一般には、各ハードウェアの起動に要する時間及び差分コピーに要する時間のいずれと比較しても十分に長い場合が多い。

【０１５２】

全体コピーの終了（時刻１５１３）を契機として、サーバ１２０及びＦＣ－ＳＷ１４０の電源が投入され、さらに、論理ボリューム６０３Ｆの差分コピーが開始される。

【０１５３】

サーバ１２０及びＦＣ－ＳＷ１４０の起動が終了し、さらに差分コピーも終了すると（時刻１５１４）、論理パス６０２Ａの移動が実行される。論理パス６０２Ａの移動が終了すると（時刻１５１５）、仮想サーバ６０１Ａの移動が実行される。

【０１５４】

従来は、全体コピー処理及び差分コピー処理が一つのコピー処理として実行されたため、管理サーバ１３０が全体コピー処理の終了時刻（すなわち時刻１５１３）を知ることができなかった。このため、全体コピーの終了を契機とする制御を実行することができなかった。その場合、制御の基準として、例えば時刻１５１１又は時刻１５１４が使用される。

【０１５５】

時刻１５１１にサーバ１２０及びＦＣ－ＳＷ１４０の電源が投入された場合、図１５の矢印１５０４及び矢印１５０５は、それらの基点が時刻１５１１となるように移動する。しかし、論理パス６０２Ａ及び仮想サーバ６０１Ａの移動が可能になるのは時刻１５１４以降であるので、サーバ１２０及びＦＣ－ＳＷ１４０は、それぞれの起動が終了した後、時刻１５１４までの時間、何も実行せずにただ電力を消費しながら待機する。すなわち、その時間、サーバ１２０及びＦＣ－ＳＷ１４０は電力を浪費する。

【０１５６】

一方、時刻１５１４（すなわち差分コピーが終了した時刻）にサーバ１２０及びＦＣ－ＳＷ１４０の電源が投入された場合、図１５の矢印１５０６及び１５０７が後の時刻（すなわち右方向）に移動する。この場合、サーバ１２０及びＦＣ－ＳＷ１４０の起動処理が終了するまで論理パス６０２Ａ及び仮想サーバ６０１Ａの移動を実行することができない。その結果、ストレージ装置１００の電源が投入されてからアプリケーション５００Ａの移動が終了するまでの時間が図１５に示す場合より長くなる。すなわち、差分コピー処理が終了してからサーバ１２０及びＦＣ－ＳＷ１４０の起動処理が終了するまでの時間、ス

トレージ装置１００が電力を浪費する。

【０１５７】

計算機システム全体の消費電力量を最小にするには、差分コピー処理、サーバ１２０の起動処理及びＦＣ－ＳＷ１４０の起動処理が同時刻に終了するように（すなわち、矢印１５０３、１５０４及び１５０５の先端が同一の時刻に対応するように）、サーバ１２０及びＦＣ－ＳＷ１４０の電源投入を制御することが望ましい。しかし、既に述べたように、差分コピー処理に要する時間を正確に予測することは困難である。このため、本実施形態では、全体コピー処理の終了を契機としてサーバ１２０及びＦＣ－ＳＷ１４０の電源が投入される。

【０１５８】

差分ストックとして格納されているデータ量が多ければ、図１５に示すように、サーバ１２０及びＦＣ－ＳＷ１４０の起動が終了してもまだ差分コピーが終了しない場合がある。その場合、サーバ１２０及びＦＣ－ＳＷ１４０の起動が終了した後、差分コピーが終了するまでの時間、サーバ１２０及びＦＣ－ＳＷ１４０が電力を浪費する。一方、差分ストックとして格納されているデータ量が少なければ、サーバ１２０及びＦＣ－ＳＷ１４０の起動が終了する前に差分コピーが終了する場合がある。その場合、差分コピーが終了した後、サーバ１２０及びＦＣ－ＳＷ１４０の起動が終了するまでの間、ストレージ装置１００が電力を浪費する。

【０１５９】

このように、全体コピー処理の終了時刻は、サーバ１２０及びＦＣ－ＳＷ１４０の電源投入のタイミングとして厳密に最適であるとは限らない。しかし、一般には、全体コピーに要する時間と差分コピーに要する時間とをあわせた論理ボリューム６０３の移動時間は、サーバ１２０及びＦＣ－ＳＷ１４０の起動処理に要する時間より大幅に長い。このため、サーバ１２０及びＦＣ－ＳＷ１４０の電源の最適な投入時刻は、論理ボリューム６０３の移動処理の終了間際であることが多い。これに対して、一般に、全体コピーに要する時間は、差分コピー処理に要する時間より大幅に長いため、全体コピー処理の終了時刻は、論理ボリューム６０３の移動処理の終了間際であることが多い。このため、全体コピー処理の終了時刻（時刻１５１３）を、サーバ１２０及びＦＣ－ＳＷ１４０の電源を投入する近似的に最適な時刻として用いることができる。

【０１６０】

ストレージ装置１００と同時にサーバ１２０及びＦＣ－ＳＷ１４０の電源を投入した場合、又は、差分コピー処理が終了した後にサーバ１２０及びＦＣ－ＳＷ１４０の電源を投入した場合のいずれの場合と比較しても、全体コピー処理の終了を契機として（具体的には、全体コピー処理が終了した後、差分コピー処理が終了する前に）サーバ１２０及びＦＣ－ＳＷ１４０の電源を投入することによって、計算機システム全体の消費電力を削減することができる。

【０１６１】

以下、上記のようなアプリケーション５００の移動の処理について詳細に説明する。

【０１６２】

図１６は、本発明の実施形態の管理テーブル１３６の説明図である。

【０１６３】

本実施形態の管理テーブル１３６は、パーティション管理テーブル１３６Ａ及びアプリケーション管理テーブル１３６Ｂを含む。説明を簡単にするため、図１６には、図１４の例においてアプリケーション５００Ａの移動が実行される前の時点のパーティション管理テーブル１３６Ａ及びアプリケーション管理テーブル１３６Ｂを示す。

【０１６４】

図１６（Ａ）は、パーティション管理テーブル１３６Ａの説明図である。パーティション管理テーブル１３６Ａは、管理サーバ１３０が管理する計算機システム上に定義された各パーティションを管理するための情報を含む。

【０１６５】

具体的には、パーティション管理テーブル１３６Ａは、パーティション番号１６０１、ハードウェア種別１６０２、ハードウェア名１６０３、リソース量１６０４、残リソース量１６０５、電源１６０６、配置アプリケーション（ＡＰＰ）１６０７及びアプリケーションリソース量１６０８を含む。

【０１６６】

パーティション番号１６０１は、計算機システム上に定義された各パーティションを識別する情報である。図１６（Ａ）の例では、パーティション番号１６０１として、「１」及び「２」が保持される。この例において、「１」は、パーティション１＿３００Ａの識別子であり、「２」は、パーティション２＿３００Ｂの識別子である。

【０１６７】

ハードウェア種別１６０２は、各パーティションに割り当てられたハードウェアの種類を識別する情報である。具体的には、ハードウェア種別１６０２は、各パーティションに割り当てられた各ハードウェアがサーバ１２０、ＦＣ－ＳＷ１４０又はストレージ装置１００のいずれであるかを識別する情報である。

【０１６８】

図１６（Ａ）の例において、「サーバ１」、「ＦＣ－ＳＷ１」及び「ストレージ１」は、それぞれ、パーティション１＿３００Ａのサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００を示す。一方、「サーバ２」、「ＦＣ－ＳＷ２」及び「ストレージ２」は、それぞれ、パーティション２＿３００Ｂのサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００を示す。

【０１６９】

ハードウェア名１６０３は、各パーティションに割り当てられたハードウェアを識別する情報である。図１６（Ａ）の例では、パーティション１＿３００Ａに対応するハードウェア名１６０３として、「サーバＰ１」（エントリ１６１１）、「ＦＣ－ＳＷＰ１」（エントリ１６１２）及び「ストレージＰ１」（エントリ１６１３）が保持される。この例において、「サーバＰ１」、「ＦＣ－ＳＷＰ１」及び「ストレージＰ１」は、それぞれ、パーティション１＿３００Ａに割り当てられたサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００の識別子である。

【０１７０】

さらに、図１６（Ａ）の例では、パーティション２＿３００Ｂに対応するハードウェア名１６０３として、「サーバＰ２」（エントリ１６１４）、「ＦＣ－ＳＷＰ２」（エントリ１６１５）及び「ストレージＰ２」（エントリ１６１６）が保持される。この例において、「サーバＰ２」、「ＦＣ－ＳＷＰ２」及び「ストレージＰ２」は、それぞれ、パーティション２＿３００Ｂに割り当てられたサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００の識別子である。

【０１７１】

図１６（Ａ）には、図１４に示したものと整合するように、一つのパーティションに一つのサーバ１２０、一つのＦＣ－ＳＷ１４０及び一つのストレージ装置１００が割り当てられる例を示す。しかし、実際には、一つのパーティションに複数のサーバ１２０、複数のＦＣ－ＳＷ１４０及び複数のストレージ装置１００が割り当てられてもよい。その場合、それらの複数のハードウェアの識別子がハードウェア名１６０３として保持される。

【０１７２】

例えば、パーティション１＿３００Ａに二つのサーバ１２０が割り当てられ、それぞれの識別子が「サーバＰ１」及び「サーバＰ１０」（図示省略）である場合、ハードウェア種別１６０２の値「サーバ１」に二つのエントリが対応付けられ、それぞれのエントリのハードウェア名１６０３として「サーバＰ１」及び「サーバＰ１０」が保持される。

【０１７３】

リソース量１６０４は、各ハードウェアに含まれる全リソースの量である。リソースは、どのように計量されてもよい。例えば、サーバ１２０のリソース量は、そのサーバ１２０が備えるＣＰＵ１２１の数であってもよいし、ＣＰＵ１２１の使用率であってもよい。

ＦＣ－ＳＷ１４０のリソース量は、そのＦＣ－ＳＷ１４０が備えるポート１４１の数であってもよいし、接続することができる論理パス６０２の数であってもよい。ストレージ装置１００のリソース量は、論理ボリューム１１１として提供することができる記憶容量であってもよい。

【０１７４】

図１６（Ａ）の例では、各ハードウェアのリソース量１６０４として「１０」が保持される。図１６（Ａ）の例では単位が省略されているが、実際には、各リソース量の単位（例えば、ストレージ装置１００のリソース量の場合、「テラバイト」等）が明示されてもよい。

【０１７５】

残リソース量１６０５は、各ハードウェアに含まれる全リソースのうち、まだいずれのアプリケーション５００にも割り当てられていないものの量である。言い換えると、残リソース量１６０５は、各ハードウェアに含まれる全リソースの量から、既にいずれかのアプリケーション５００に割り当てられたものの量を減算した残りの量である。

【０１７６】

電源１６０６は、各ハードウェアの電源の状態を示す情報である。図１６（Ａ）において、電源１６０６の値「ＯＮ」は電源が投入されている状態、「ＯＦＦ」は電源が遮断されている状態を示す。

【０１７７】

配置アプリケーション１６０７は、各パーティションに配置されたアプリケーション５００を識別する情報である。

【０１７８】

アプリケーションリソース量１６０８は、各ハードウェアに含まれる全リソースのうち、アプリケーション５００に割り当てられたリソースの量である。

【０１７９】

図１６（Ａ）の例は、図１４においてアプリケーション５００Ａがパーティション２＿３００Ｂからパーティション１＿３００Ａに移動する前の時点に対応する。すなわち、この時点においてパーティション１＿３００Ａ内の各ハードウェアのリソースはまだいずれのアプリケーション５００にも割り当てられていない。

【０１８０】

このため、パーティション１＿３００Ａ内の各ハードウェアに対応する残リソース量１６０５の値は、リソース量１６０４の値と同じである。パーティション１＿３００Ａ内の各ハードウェアの電源はまだ投入されていないため、対応する電源１６０６として「ＯＦＦ」が保持される。さらに、パーティション１＿３００Ａ内の各ハードウェアに対応する配置アプリケーション１６０７及びアプリケーションリソース量１６０８は空白である。

【０１８１】

一方、図１４のアプリケーション５００Ａが移動する前の時点において、パーティション２＿３００Ｂにはアプリケーション５００Ａが配置されている。このため、パーティション２＿３００Ｂに対応する配置アプリケーション１６０７として、アプリケーション５００Ａの識別子（図１６（Ａ）の例では「ＡＰＰ１」）が保持される。この状態において、パーティション２＿３００Ｂ内の各ハードウェアの電源が投入されているため、各ハードウェアに対応する電源１６０６として「ＯＮ」が保持される。

【０１８２】

さらに、パーティション２＿３００Ｂにおいて、サーバ１２０の全リソース量「１０」のうち「５」、ＦＣ－ＳＷ１４０の全リソース量「１０」のうち「１」、ストレージ装置１００の全リソース量「１０」のうち「６」がアプリケーション５００に割り当てられている場合、パーティション２＿３００Ｂのサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００に対応するアプリケーションリソース量１６０８として、それぞれ「５」、「１」及び「６」が保持される。そしてそれらに対応する残リソース量１６０５として、それぞれ「５」、「９」及び「４」が保持される。

【０１８３】
図１６（Ｂ）は、アプリケーション管理テーブル１３６Ｂの説明図である。アプリケーション管理テーブル１３６Ｂは、管理サーバ１３０が管理する計算機システム上で稼動する各アプリケーション５００を管理するための情報を含む。
【０１８４】
具体的には、アプリケーション管理テーブル１３６Ｂは、アプリケーション名１６２１、配置パーティション番号１６２２、配置サーバ名１６２３、配置ＦＣ－ＳＷ名１６２４、配置ストレージ名１６２５、サーバリソース量１６２６、ＦＣ－ＳＷリソース量１６２７及びストレージリソース量１６２８を含む。
【０１８５】
アプリケーション名１６２１は、計算機システム内で稼動するアプリケーション５００を識別する情報である。
【０１８６】
配置パーティション番号１６２２は、各アプリケーション５００が配置されているパーティションを識別する情報である。
【０１８７】
配置サーバ名１６２３は、各アプリケーション５００の仮想サーバ６０１が配置されているサーバ１２０を識別する情報である。
【０１８８】
配置ＦＣ－ＳＷ名１６２４は、各アプリケーション５００の論理パス６０２が配置されているＦＣ－ＳＷ１４０を識別する情報である。
【０１８９】
配置ストレージ名１６２５は、各アプリケーション５００の論理ボリューム６０３が配置されているストレージ装置１００を識別する情報である。
【０１９０】
サーバリソース量１６２６は、各アプリケーション５００に割り当てられているサーバ１２０のリソース量を示す。
【０１９１】
ＦＣ－ＳＷリソース量１６２７は、各アプリケーション５００に割り当てられているＦＣ－ＳＷ１４０のリソース量を示す。
【０１９２】
ストレージリソース量１６２８は、各アプリケーション５００に割り当てられているストレージ装置１００のリソース量を示す。
【０１９３】
図１６（Ｂ）の例は、図１４においてアプリケーション５００Ａがパーティション２＿３００Ｂからパーティション１＿３００Ａに移動する前の時点に対応する。すなわち、この時点においてアプリケーション５００Ａは、パーティション２＿３００Ｂに配置されている。このため、図１６（Ｂ）の例において、アプリケーション名１６２１の値「ＡＰＰ１」に対応する配置パーティション番号１６２２、配置サーバ名１６２３、配置ＦＣ－ＳＷ名１６２４及び配置ストレージ名１６２５として、それぞれ、「２」、「サーバＰ２」、「ＦＣ－ＳＷＰ２」及び「ストレージＰ２」が保持される。これらの値は、図１６（Ａ）の例と整合する。さらに、図１６（Ａ）に示すように、サーバリソース量１６２６、ＦＣ－ＳＷリソース量１６２７及びストレージリソース量１６２８として、それぞれ、「５」、「１」及び「６」が保持される。
【０１９４】
次に、アプリケーション５００の移動及びそれに伴う電源制御の処理について、フローチャートを参照して説明する。以下の説明において、処理の具体例として、図１４及び図１５に示したアプリケーション５００Ａの移動処理及びそのための電源制御処理を参照する。
【０１９５】

図１７は、本発明の実施形態において実行されるアプリケーション５００の移動及び電源制御の処理の全体を示すフローチャートである。

【０１９６】

この処理は、図１０のステップ１００２において実行される。

【０１９７】

管理サーバ１３０の管理プログラム１３５は、アプリケーション５００の移動処理を開始すると（ステップ１７０１）、最初に、ストレージ電源ＯＮ処理を実行する（ステップ１７０２）。これは、移動先のストレージ装置１００（図１４の例では、パーティション１＿３００Ａのストレージ装置１００）の電源を投入する処理である。

【０１９８】

ステップ１７０２において実行されるストレージ電源ＯＮ処理については、後に図１８を参照して詳細に説明する。

【０１９９】

ストレージ電源ＯＮ処理が終了する（すなわち、移動先のストレージ装置１００の起動処理が終了する）と、管理プログラム１３５は、論理ボリューム６０３の全体コピー処理を実行する（ステップ１７０３）。ステップ１７０３において実行される全体コピー処理については、後に図１９及び図２０を参照して詳細に説明する。

【０２００】

ステップ１７０３の全体コピー処理が終了すると、管理プログラム１３５は、サーバ電源ＯＮ処理を実行する（ステップ１７０４）。これは、移動先のサーバ１２０の電源を投入する処理である。ステップ１７０４において実行されるサーバ電源ＯＮ処理については、後に図２１を参照して詳細に説明する。

【０２０１】

ステップ１７０３の全体コピー処理が終了すると、管理プログラム１３５は、さらに、ＦＣ－ＳＷ電源ＯＮ処理を実行する（ステップ１７０５）。これは、移動先のＦＣ－ＳＷの電源を投入する処理である。ステップ１７０５において実行されるＦＣ－ＳＷ電源ＯＮ処理については、後に図２２を参照して詳細に説明する。

【０２０２】

ステップ１７０３の全体コピー処理が終了すると、管理プログラム１３５は、さらに、論理ボリューム６０３の差分コピー処理を実行する（ステップ１７０６）。ステップ１７０６において実行される差分コピー処理については、後に図２３及び図２４を参照して詳細に説明する。

【０２０３】

管理プログラム１３５は、全体コピー処理の終了を、ストレージ装置１００から受信した終了メッセージに基づいて判定してもよいし、管理プログラム１３５が算出した全体コピー処理時間に基づいて判定してもよい。

【０２０４】

前者の場合、管理サーバ１３０がストレージ装置１００から終了メッセージを受信した時刻が全体コピー処理の終了時刻として扱われ、その時刻を契機にステップ１７０４～ステップ１７０６が実行される。終了メッセージについては後述する（図２０参照）。

【０２０５】

後者の場合、全体コピー処理が開始されてから、全体コピー処理時間が経過した時刻が全体コピー処理の終了時刻として扱われ、その時刻を契機にステップ１７０４～ステップ１７０６が実行される。全体コピー処理時間は、コピー元の論理ボリューム６０３Ｆの容量を、論理ボリューム６０３Ｆからコピー先の論理ボリューム６０３Ａへのデータ転送速度によって除算することによって算出される。データ転送速度は、ハードウェアのスペックに基づいて算出されてもよいし、実測された値であってもよい。

【０２０６】

なお、図１７は、便宜上、ステップ１７０４からステップ１７０６が順次実行されるように記載されているが、この順序は一例にすぎない。これらの三つのステップは、ステッ

プ１７０３の全体コピー処理の終了を契機として実行される必要があるが、どのような順序で実行されてもよい。可能であれば、これらの三つのステップが同時に実行されてもよい。ただし、図１５を参照して説明したように、本実施形態による消費電力削減の効果を得るためには、ステップ１７０４及びステップ１７０５は、遅くとも差分コピー処理が終了する前に開始する必要がある。

【０２０７】

サーバ電源ＯＮ処理、ＦＣ－ＳＷ電源ＯＮ処理及び差分コピー処理（ステップ１７０４～ステップ１７０６）が全て終了すると、管理プログラム１３５は、論理パス６０２の移動処理を実行する（ステップ１７０７）。この処理については、後に図２５を参照して詳細に説明する。

【０２０８】

ステップ１７０７の論理パス６０２の移動処理が終了すると、管理プログラム１３５は、仮想サーバ６０１の移動処理を実行する（ステップ１７０８）。この処理については、後に図２６を参照して詳細に説明する。

【０２０９】

ステップ１７０８の仮想サーバ６０１の移動処理が終了すると、管理プログラム１３５は、アプリケーション５００の移動処理を終了する（ステップ１７０９）。なお、仮想サーバ６０１の移動処理が終了した後（例えばステップ１７０９において）、移動元のパーティションのハードウェア（図１４の例では、パーティション２＿３００Ｂのサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００）の電源が遮断されてもよい。

【０２１０】

図１８は、本発明の実施形態において実行されるストレージ電源ＯＮ処理を示すフローチャートである。

【０２１１】

この処理は、図１７のステップ１７０２において実行される。

【０２１２】

管理プログラム１３５は、ストレージ電源ＯＮ処理が開始されると（ステップ１８０１）、対象ハードウェアを確認する（ステップ１８０２）。具体的には、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、移動先のパーティションに含まれるストレージ装置１００（すなわち移動先のストレージ装置）の状態を確認する。

【０２１３】

例えば、図１４のようにパーティション１＿３００Ａが移動先として指定されている場合、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、パーティション番号１６０１の値「１」及びハードウェア種別１６０２の値「ストレージ１」に対応する全てのエントリを特定する。そして、管理プログラム１３５は、特定された各エントリの電源１６０６の値を確認する。

【０２１４】

次に、管理プログラム１３５は、移動先のストレージ装置１００の電源が遮断された状態（すなわち「ＯＦＦ」状態）であるか否かを判定する（ステップ１８０３）。具体的には、管理プログラム１３５は、ステップ１８０２で取得した値が「ＯＮ」又は「ＯＦＦ」のいずれであるかを判定する。

【０２１５】

移動先のストレージ装置１００の電源が既に投入されている場合、さらに電源を投入する処理を実行する必要はないため、管理プログラム１３５は、ストレージ電源ＯＮ処理を終了する（ステップ１８０６）。

【０２１６】

移動先のストレージ装置１００の電源が遮断されている場合、管理プログラム１３５は、管理ネットワーク１５０を介して、移動先のストレージ装置１００に電源ＯＮ命令を送信する（ステップ１８０４）。ストレージ装置１００の電源制御部１０５は、Ｉ／Ｆ１０４を介して電源ＯＮ命令を受信すると、ストレージ装置１００の電源を投入する。なお、

このような処理を実行するために、ストレージ装置１００のうち、少なくともＩ／Ｆ１０４及び電源制御部１０５の電源は、ストレージ電源ＯＮ処理が実行される時点で既に投入されている必要がある。

【０２１７】

ストレージ装置１００は、電源ＯＮ命令に従って電源が投入され、起動処理が終了すると、起動処理の終了の報告を管理サーバ１３０に送信してもよい。

【０２１８】

次に、管理プログラム１３５は、ストレージ電源ＯＮ処理の結果が反映されるように、管理テーブル１３６の電源に関する情報を更新する（ステップ１８０５）。

【０２１９】

ステップ１８０２で複数のエントリが特定された場合（すなわち、移動先のパーティションが複数のストレージ装置１００を含む場合）、各エントリについてステップ１８０３～ステップ１８０５が実行される。

【０２２０】

以上でストレージ電源ＯＮ処理が終了する（ステップ１８０６）。

【０２２１】

図１６（Ａ）の例では、ステップ１８０２において、ハードウェア名「ストレージＰ１」を含むエントリ１６１３のみが特定される。エントリ１６１３の電源１６０６の値が「ＯＦＦ」であるため、ステップ１８０４において、管理プログラム１３５は、移動先のストレージ装置１００（すなわち「ストレージＰ１」によって識別されるストレージ装置１００）に電源ＯＮ命令を送信する。これによって移動先のパーティション１＿３００Ａに含まれるストレージ装置１００の電源が投入され、ステップ１８０５において、エントリ１６１３の電源１６０６の値が「ＯＮ」に更新される。

【０２２２】

図１９は、本発明の実施形態の管理サーバ１３０が実行する論理ボリューム６０３の全体コピー処理のフローチャートである。

【０２２３】

論理ボリューム６０３の全体コピー処理は、図１５を参照して説明したように、論理ボリューム６０３の移動処理の一部である。この処理は、図１７のステップ１７０３において実行される。

【０２２４】

管理プログラム１３５は、論理ボリューム６０３の移動処理を開始すると（ステップ１９０１）、論理ボリューム６０３の全体コピー処理を開始し（ステップ１９０２）、全体コピーを実行して（ステップ１９０３）、全体コピー処理を終了する（ステップ１９０４）。

【０２２５】

具体的には、管理プログラム１３５は、ステップ１９０３において、全体コピー命令を移動元のストレージ装置１００に送信する。この命令を受信したストレージ装置１００が実行する処理については、後に図２０を参照して詳細に説明する。

【０２２６】

図２０は、本発明の実施形態のストレージ装置１００が実行する論理ボリューム６０３の全体コピー処理のフローチャートである。

【０２２７】

この処理は、図１９のステップ１９０３において送信された命令を受信したストレージ装置１００（図１４の例では、パーティション２＿３００Ｂのストレージ装置１００）のコントローラ１０１によって実行される。

【０２２８】

コントローラ１０１は、論理ボリューム６０３の全体コピー処理を開始すると（ステップ２００１）、差分Ｉ／Ｏを格納する記憶領域を確保する（ステップ２００２）。この記憶領域は、例えば、ストレージ装置１００が管理するいずれかの論理ボリューム１１１の

空き記憶領域に確保されてもよい。

【０２２９】

次に、コントローラ１０１は、これから移動する論理ボリューム６０３へのデータＩ／Ｏの実行先を、ステップ２００２において確保した記憶領域に切り替える（ステップ２００３）。その後、コントローラ１０１は、これから移動する論理ボリューム６０３（図１４の例では、論理ボリューム６０３Ｆ）へのデータＩ／Ｏを受信すると、それによって更新されるデータを論理ボリューム６０３に反映せず、ステップ２００２において確保した記憶領域に差分ストックとして格納する（図１１参照）。

【０２３０】

次に、コントローラ１０１は、論理ボリューム６０３の全体コピーを実行する（ステップ２００４）。図１４の例では、コントローラ１０１は、論理ボリューム６０３Ｆの全データを読み出し、読み出したデータを論理ボリューム６０３Ａに書き込む要求をパーティション１＿３００Ａのストレージ装置１００に送信する。なお、コピー元及びコピー先の論理ボリューム６０３は、図１９のステップ１９０３において送信された命令に含まれる引数によって指定される。

【０２３１】

読み出した全データの送信が終了すると、コントローラ１０１は、論理ボリューム６０３の全体コピー処理を終了する（ステップ２００５）。このとき、コントローラ１０１は、全体コピー処理が終了したことを示す終了メッセージを管理サーバ１３０に送信する。

【０２３２】

管理プログラム１３５は、ステップ２００５で送信された終了メッセージを受信したことを契機として、ステップ１７０４～ステップ１７０６を実行する。

【０２３３】

なお、図１７を参照して説明したように、管理プログラム１３５は、ハードウェアスペック等に基づいて全体コピー処理時間を算出し、その算出された全体コピー処理時間に基づいてステップ１７０４～ステップ１７０６の実行を制御してもよい。その場合、コントローラ１０１は、ステップ２００５において終了メッセージを送信しなくてもよい。

【０２３４】

図２１は、本発明の実施形態において実行されるサーバ電源ＯＮ処理を示すフローチャートである。

【０２３５】

この処理は、図１７のステップ１７０４において実行される。すなわち、この処理は、図２０のステップ２００５において送信された終了メッセージを管理サーバ１３０が受信したことを契機として実行される。

【０２３６】

管理プログラム１３５は、ストレージ電源ＯＮ処理が開始されると（ステップ２１０１）、対象ハードウェアを確認する（ステップ２１０２）。具体的には、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、移動先のパーティションに含まれるサーバ１２０（すなわち移動先のサーバ）の状態を確認する。

【０２３７】

例えば、図１４のようにパーティション１＿３００Ａが移動先として指定されている場合、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、パーティション番号１６０１の値「１」及びハードウェア種別１６０２の値「サーバ１」に対応する全てのエントリを特定する。そして、管理プログラム１３５は、特定された各エントリの電源１６０６の値を確認する。

【０２３８】

次に、管理プログラム１３５は、移動先のサーバ１２０の電源が遮断された状態（すなわち「ＯＦＦ」状態）であるか否かを判定する（ステップ２１０３）。具体的には、管理プログラム１３５は、ステップ２１０２で取得した値が「ＯＮ」又は「ＯＦＦ」のいずれであるかを判定する。

【０２３９】
移動先のサーバ１２０の電源が既に投入されている場合、さらに電源を投入する処理を実行する必要はないため、管理プログラム１３５は、サーバ電源ＯＮ処理を終了する（ステップ２１０６）。
【０２４０】
移動先のサーバ１２０の電源が遮断されている場合、管理プログラム１３５は、管理ネットワーク１５０を介して、移動先のサーバ１２０に電源ＯＮ命令を送信する（ステップ２１０４）。サーバ１２０の電源制御部１２３は、Ｉ／Ｆ１２５を介して電源ＯＮ命令を受信すると、サーバ１２０の電源を投入する。なお、このような処理を実行するために、サーバ１２０のうち、少なくともＩ／Ｆ１２５及び電源制御部１２３の電源は、サーバ電源ＯＮ処理が実行される時点で既に投入されている必要がある。
【０２４１】
サーバ１２０は、電源ＯＮ命令に従って電源が投入され、起動処理が終了すると、起動処理の終了の報告を管理サーバ１３０に送信してもよい。
【０２４２】
次に、管理プログラム１３５は、サーバ電源ＯＮ処理の結果が反映されるように、管理テーブル１３６の電源に関する情報を更新する（ステップ２１０５）。
【０２４３】
ステップ２１０２で複数のエントリが特定された場合（すなわち、移動先のパーティションが複数のサーバ１２０を含む場合）、各エントリについてステップ２１０３～ステップ２１０５が実行される。
【０２４４】
以上でサーバ電源ＯＮ処理が終了する（ステップ２１０６）。
【０２４５】
図１６（Ａ）の例では、ステップ２１０２において、ハードウェア名「サーバＰ１」を含むエントリ１６１１のみが特定される。エントリ１６１１の電源１６０６の値が「ＯＦＦ」であるため、ステップ２１０４において、管理プログラム１３５は、移動先のサーバ１２０（すなわち「サーバＰ１」によって識別されるサーバ１２０）に電源ＯＮ命令を送信する。これによって移動先のパーティション１＿３００Ａに含まれるサーバ１２０の電源が投入され、ステップ２１０５において、エントリ１６１１の電源１６０６の値が「ＯＮ」に更新される。
【０２４６】
なお、図１４の例のように、各パーティションに一つ以上のサーバ１２０が割り当てられている場合、ステップ２１０４において、移動先のパーティションのサーバ１２０の電源を投入する命令（すなわち、そのサーバ１２０に含まれるリソース全体の電源を投入する命令）が送信される。しかし、例えば図４に示すように、一つのサーバ１２０のリソースの各部分（例えば各ＣＰＵ１２１）が各パーティションに割り当てられ、かつ、その部分ごとに電源を投入することができる場合、ステップ２１０４では、移動先のパーティションに含まれるサーバ１２０のリソースの電源を投入する命令が、そのリソースを含むサーバ１２０に送信される。この命令を受信したサーバ１２０の電源制御部１２３は、命令によって指定されたリソースの部分（例えば指定されたＣＰＵ）の電源を投入する。
【０２４７】
図２２は、本発明の実施形態において実行されるＦＣ－ＳＷ電源ＯＮ処理を示すフローチャートである。
【０２４８】
この処理は、図１７のステップ１７０５において実行される。すなわち、この処理は、図２０のステップ２００５において送信された終了メッセージを管理サーバ１３０が受信したことを契機として実行される。
【０２４９】
管理プログラム１３５は、ＦＣ－ＳＷ電源ＯＮ処理が開始されると（ステップ２２０１

）、対象ハードウェアを確認する（ステップ２２０２）。具体的には、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、移動先のパーティションに含まれるＦＣ－ＳＷ１４０（すなわち移動先のＦＣ－ＳＷ）の状態を確認する。

【０２５０】

例えば、図１４のようにパーティション１＿３００Ａが移動先として指定されている場合、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、パーティション番号１６０１の値「１」及びハードウェア種別１６０２の値「ＦＣ－ＳＷ１」に対応する全てのエントリを特定する。そして、管理プログラム１３５は、特定された各エントリの電源１６０６の値を確認する。

【０２５１】

次に、管理プログラム１３５は、移動先のＦＣ－ＳＷ１４０の電源が遮断された状態（すなわち「ＯＦＦ」状態）であるか否かを判定する（ステップ２２０３）。具体的には、管理プログラム１３５は、ステップ２２０２で取得した値が「ＯＮ」又は「ＯＦＦ」のいずれであるかを判定する。

【０２５２】

移動先のＦＣ－ＳＷ１４０の電源が既に投入されている場合、さらに電源を投入する処理を実行する必要はないため、管理プログラム１３５は、ＦＣ－ＳＷ電源ＯＮ処理を終了する（ステップ２２０６）。

【０２５３】

移動先のＦＣ－ＳＷ１４０の電源が遮断されている場合、管理プログラム１３５は、管理ネットワーク１５０を介して、移動先のＦＣ－ＳＷ１４０に電源ＯＮ命令を送信する（ステップ２２０４）。ＦＣ－ＳＷ１４０の電源制御部１４３は、Ｉ／Ｆ１４２を介して電源ＯＮ命令を受信すると、ＦＣ－ＳＷ１４０の電源を投入する。なお、このような処理を実行するために、ＦＣ－ＳＷ１４０のうち、少なくともＩ／Ｆ１４２及び電源制御部１４３の電源は、ＦＣ－ＳＷ電源ＯＮ処理が実行される時点で既に投入されている必要がある。

【０２５４】

ＦＣ－ＳＷ１４０は、電源ＯＮ命令に従って電源が投入され、起動処理が終了すると、起動処理の終了の報告を管理サーバ１３０に送信してもよい。

【０２５５】

次に、管理プログラム１３５は、ＦＣ－ＳＷ電源ＯＮ処理の結果が反映されるように、管理テーブル１３６の電源に関する情報を更新する（ステップ２２０５）。

【０２５６】

ステップ２２０２で複数のエントリが特定された場合（すなわち、移動先のパーティションが複数のＦＣ－ＳＷ１４０を含む場合）、各エントリについてステップ２２０３～ステップ２２０５が実行される。

【０２５７】

以上でＦＣ－ＳＷ電源ＯＮ処理が終了する（ステップ２２０６）。

【０２５８】

図１６（Ａ）の例では、ステップ２２０２において、ハードウェア名「ＦＣ－ＳＷＰ１」を含むエントリ１６１２のみが特定される。エントリ１６１２の電源１６０６の値が「ＯＦＦ」であるため、ステップ２２０４において、管理プログラム１３５は、移動先のＦＣ－ＳＷ１４０（すなわち「ＦＣ－ＳＷＰ１」によって識別されるＦＣ－ＳＷ１４０）に電源ＯＮ命令を送信する。これによって移動先のパーティション１＿３００Ａに含まれるＦＣ－ＳＷ１４０の電源が投入され、ステップ２２０５において、エントリ１６１２の電源１６０６の値が「ＯＮ」に更新される。

【０２５９】

図２３は、本発明の実施形態の管理サーバ１３０が実行する論理ボリューム６０３の差分コピー処理のフローチャートである。

【０２６０】

論理ボリューム６０３の差分コピー処理は、図１５を参照して説明したように、論理ボリューム６０３の移動処理の一部である。この処理は、図１７のステップ１７０６において実行される。

【０２６１】

管理プログラム１３５は、論理ボリューム６０３の差分コピー処理を開始すると（ステップ２３０１）、差分コピーを実行する（ステップ２３０２）。具体的には、管理プログラム１３５は、差分コピー命令を移動元のストレージ装置１００に送信する。この命令を受信したストレージ装置１００が実行する処理については、後に図２４を参照して詳細に説明する。

【０２６２】

次に、管理プログラム１３５は、実行されたコピーの結果が反映されるように、管理テーブル１３６を更新する（ステップ２３０３）。例えば、図１４に示すように論理ボリューム６０３Ｆから論理ボリューム６０３Ａへのコピーが実行された場合、パーティション管理テーブル１３６Ａのうち、ストレージ装置１００に関するエントリ１６１３及び１６１６が更新される。

【０２６３】

図１６（Ａ）の例では、論理ボリューム６０３Ａを含むストレージ装置１００がエントリ１６１３に対応し、論理ボリューム６０３Ｆを含むストレージ装置１００がエントリ１６１６に対応する。この場合、コピー前の論理ボリューム６０３Ｆが使用しているリソース量が「６」である。論理ボリューム６０３Ｆから論理ボリューム６０３Ａへのコピーによって、アプリケーション５００Ａの一部としての論理ボリューム６０３Ｆはリソースを使用しなくなり、一方、論理ボリューム６０３Ａは、アプリケーション５００Ａの一部として新たにリソースを使用するようになる。

【０２６４】

このため、ステップ２３０３において、エントリ１６１３の残リソース量１６０５、配置アプリケーション１６０７及びアプリケーションリソース量１６０８が、それぞれ、「４」、「ＡＰＰ１」及び「６」に更新される。一方、エントリ１６１６の残リソース量１６０５は「１０」に更新され、エントリ１６１６の配置アプリケーション１６０７及びアプリケーションリソース量１６０８は空白に更新される。

【０２６５】

さらに、ステップ２３０３において、アプリケーション管理テーブル１３６Ｂも更新される。具体的には、アプリケーション名「ＡＰＰ１」に対応する配置ストレージ名１６２５が「ストレージＰ１」に更新される。

【０２６６】

管理プログラム１３５は、ステップ２３０３を実行すると、論理ボリューム６０３の差分コピー処理を終了する（ステップ２３０４）。これによって、論理ボリューム６０３の移動処理が終了する（ステップ２３０５）。

【０２６７】

図２４は、本発明の実施形態のストレージ装置１００が実行する論理ボリューム６０３の差分コピー処理のフローチャートである。

【０２６８】

この処理は、図２３のステップ２３０２において送信された命令を受信したストレージ装置１００（図１４の例では、パーティション２＿３００Ｂのストレージ装置１００）のコントローラ１０１によって実行される。

【０２６９】

コントローラ１０１は、論理ボリューム６０３の差分コピー処理を開始すると（ステップ２４０１）、コピー元の論理ボリューム６０３（図１４の例では論理ボリューム６０３Ｆ）へのデータＩ／Ｏを停止する（ステップ２４０２）。ただし、既に図２０のステップ２００３においてデータＩ／Ｏの実行先が切り替えられているため、ステップ２４０２では、差分データＩ／Ｏの記憶領域へのデータＩ／Ｏが停止される。

【０２７０】
次に、コントローラ１０１は、差分コピーを実行する（ステップ２４０３）。具体的には、コントローラ１０１は、差分データＩ／Ｏの記憶領域からデータを読み出し、読み出したデータをコピー先の論理ボリューム６０３Ａに書き込む要求をパーティション１＿３００Ａのストレージ装置１００に送信する。
【０２７１】
全ての差分データの送信が終了すると、コントローラ１０１は、確保した差分データＩ／Ｏの記憶領域を削除する（ステップ２４０４）。
【０２７２】
以上で、コントローラ１０１は、論理ボリューム６０３の差分コピー処理を終了する（ステップ２４０５）。
【０２７３】
なお、ステップ２４０２において停止されたデータＩ／Ｏは、後述する論理パス６０２の移動処理によって再開される（図２５参照）。
【０２７４】
図２５は、本発明の実施形態において実行される論理パス６０２の移動処理のフローチャートである。
【０２７５】
この処理は、図１７のステップ１７０７において実行される。
【０２７６】
管理プログラム１３５は、論理パス６０２の移動処理を開始すると（ステップ２５０１）、論理パス６０２を移動する（ステップ２５０２）。例えば、図１４に示すように論理パス６０２Ａを移動する場合、管理プログラム１３５は、論理パス６０２Ａを移動させる命令を、パーティション２＿３００Ｂのサーバ１２０に送信してもよい。この命令を受信したサーバ１２０は、論理パス６０２Ａがパーティション１＿３００ＡのＦＣ－ＳＷ１４０を経由して論理ボリューム６０３Ａに接続されるようにパス管理プログラム１２７の設定を変更してもよい。
【０２７７】
ステップ２５０２において論理パス６０２Ａが移動すると、パーティション１＿３００Ａのストレージ装置１００は、論理ボリューム６０３Ａへのデータ Ｉ／Ｏを再開する。
【０２７８】
次に、管理プログラム１３５は、ステップ２５０２における論理パス６０２の移動が反映されるように管理テーブル１３６を更新する（ステップ２５０３）。例えば、図１４に示すように論理パス６０２Ａが移動した場合、パーティション管理テーブル１３６Ａのうち、ＦＣ－ＳＷ１４０に関するエントリ１６１２及び１６１５が更新される。
【０２７９】
図１６（Ａ）の例では、パーティション１＿３００ＡのＦＣ－ＳＷ１４０がエントリ１６１２に対応し、パーティション２＿３００ＢのＦＣ－ＳＷ１４０がエントリ１６１５に対応する。この場合、移動前の論理パス６０２Ａが使用しているパーティション２＿３００ＢのＦＣ－ＳＷ１４０のリソース量が「１」である。ステップ２５０２の移動の結果、論理パス６０２Ａは、パーティション２＿３００ＢのＦＣ－ＳＷ１４０のリソースを使用しなくなり、代わりにパーティション１＿３００ＡのＦＣ－ＳＷ１４０のリソースを使用するようになる。
【０２８０】
このため、ステップ２５０３において、エントリ１６１２の残リソース量１６０５、配置アプリケーション１６０７及びアプリケーションリソース量１６０８が、それぞれ、「９」、「ＡＰＰ１」及び「１」に更新される。一方、エントリ１６１５の残リソース量１６０５は「１０」に更新され、エントリ１６１５の配置アプリケーション１６０７及びアプリケーションリソース量１６０８は空白に更新される。
【０２８１】

さらに、ステップ２５０３において、アプリケーション管理テーブル１３６Ｂも更新される。具体的には、アプリケーション名「ＡＰＰ１」に対応する配置ＦＣ－ＳＷ名１６２４が「ＦＣ－ＳＷＰ１」に更新される。

【０２８２】

以上で論理パス６０２の移動処理が終了する（ステップ２５０４）。

【０２８３】

図２６は、本発明の実施形態において実行される仮想サーバ６０１の移動処理のフローチャートである。

【０２８４】

この処理は、図１７のステップ１７０８において実行される。

【０２８５】

管理プログラム１３５は、仮想サーバ６０１の移動処理を開始すると（ステップ２６０１）、仮想サーバ６０１を移動する（ステップ２６０２）。この移動は、例えば、図９において説明したように、メモリ１２２のイメージをコピーすることによって実行される。

【０２８６】

次に、管理プログラム１３５は、ステップ２６０２における論理パス６０２の移動が反映されるように管理テーブル１３６を更新する（ステップ２６０３）。例えば、図１４に示すように仮想サーバ６０１Ａが移動した場合、パーティション管理テーブル１３６Ａのうち、サーバ１２０に関するエントリ１６１１及び１６１４が更新される。

【０２８７】

図１６（Ａ）の例では、パーティション１＿３００Ａのサーバ１２０がエントリ１６１１に対応し、パーティション２＿３００Ｂのサーバ１２０がエントリ１６１４に対応する。この場合、移動前の仮想サーバ６０１Ａが使用しているパーティション２＿３００Ｂのサーバ１２０のリソース量が「１」である。ステップ２６０２の移動の結果、仮想サーバ６０１Ａは、パーティション２＿３００Ｂのサーバ１２０のリソースを使用しなくなり、代わりにパーティション１＿３００Ａのサーバ１２０のリソースを使用するようになる。

【０２８８】

このため、ステップ２６０３において、エントリ１６１１の残リソース量１６０５、配置アプリケーション１６０７及びアプリケーションリソース量１６０８が、それぞれ、「５」、「ＡＰＰ１」及び「５」に更新される。一方、エントリ１６１４の残リソース量１６０５は「１０」に更新され、エントリ１６１４の配置アプリケーション１６０７及びアプリケーションリソース量１６０８は空白に更新される。

【０２８９】

さらに、ステップ２６０３において、アプリケーション管理テーブル１３６Ｂも更新される。具体的には、アプリケーション名「ＡＰＰ１」に対応する配置サーバ名１６２３が「サーバＰ１」に更新される。

【０２９０】

以上で仮想サーバ６０１の移動処理が終了する（ステップ２６０４）。

【０２９１】

図２７は、本発明の実施形態において実行される移動先パーティションの状態確認処理のフローチャートである。

【０２９２】

この処理は、図８のステップ８０４において実行される。

【０２９３】

管理プログラム１３５は、移動先パーティションの状態確認処理を開始すると（ステップ２７０１）、これから移動しようとするアプリケーション５００が現在配置されているパーティションを確認する（ステップ２７０２）。この確認のために、アプリケーション管理テーブル１３６Ｂの配置パーティション番号１６２２から配置ストレージ名１６２５までが参照される。

【０２９４】

次に、管理プログラム１３５は、移動先のリソース、すなわち、移動先のパーティションを選択する（ステップ２７０３）。

【０２９５】

次に、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、移動先として選択されたパーティションに含まれるストレージ装置１００（以下、選択されたストレージ装置）の状態を確認する（ステップ２７０４）。

【０２９６】

次に、管理プログラム１３５は、ステップ２７０４で確認した結果に基づいて、選択されたストレージ装置１００へのアプリケーション５００の移動（すなわち、論理ボリューム６０３のデータのコピー）が可能であるか否かを判定する（ステップ２７０５）。

【０２９７】

選択されたストレージ装置１００へのアプリケーション５００の移動が可能であると判定された場合、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、移動先として選択されたパーティションに含まれるＦＣ－ＳＷ１４０（以下、選択されたＦＣ－ＳＷ）の状態を確認する（ステップ２７０６）。

【０２９８】

次に、管理プログラム１３５は、ステップ２７０６で確認した結果に基づいて、選択されたＦＣ－ＳＷ１４０へのアプリケーション５００の移動（すなわち論理パス６０２の移動）が可能であるか否かを判定する（ステップ２７０７）。

【０２９９】

選択されたＦＣ－ＳＷ１４０へのアプリケーション５００の移動が可能であると判定された場合、管理プログラム１３５は、パーティション管理テーブル１３６Ａを参照して、移動先として選択されたパーティションに含まれるサーバ１２０（以下、選択されたサーバ）の状態を確認する（ステップ２７０８）。

【０３００】

次に、管理プログラム１３５は、ステップ２７０８で確認した結果に基づいて、選択されたサーバ１２０へのアプリケーション５００の移動（すなわち仮想サーバ６０１の移動）が可能であるか否かを判定する（ステップ２７０９）。

【０３０１】

ステップ２７０９において、選択されたサーバ１２０へのアプリケーション５００の移動が可能であると判定された場合、結局、移動先として選択されたパーティションへのアプリケーション５００の移動が可能である。このため、管理プログラム１３５は、アプリケーションが移動可能であることを示す応答を返して処理を終了する（ステップ２７１０）。この場合、図８のステップ８０５において、選択したパーティションへのアプリケーション５００の移動が可能であると判定される。

【０３０２】

一方、ステップ２７０５、ステップ２７０７又はステップ２７０９の少なくとも一つにおいて、アプリケーション５００の移動が可能でないと判定された場合、結局、移動先として選択されたパーティションへアプリケーション５００を移動することはできない。このため、管理プログラム１３５は、アプリケーションを移動できないことを示す応答を返して処理を終了する（ステップ２７１１）。この場合、図８のステップ８０５において、選択したパーティションへのアプリケーション５００の移動が可能でないと判定される。

【０３０３】

ここで、図１４に示すアプリケーション５００Ａの移動を例として、図２７の処理の具体例を説明する。

【０３０４】

図１４に示すアプリケーション５００Ａが移動する前の時点の管理テーブル１３６は、図１６に示すとおりである。この場合、ステップ２７０２において、配置パーティション番号１６２２の値「２」、配置サーバ名１６２３の値「サーバＰ２」、配置ＦＣ－ＳＷ名１６２４の値「ＦＣ－ＳＷＰ２」及び配置ストレージ名１６２５の値「ストレージＰ２」

から、アプリケーション５００Ａが、パーティション２＿３００Ｂに配置されていることが確認される。

【０３０５】

そして、ステップ２７０３において、パーティション番号１６０１が参照され、現在アプリケーション５００Ａが配置されているパーティション２＿３００Ｂ以外のパーティション、すなわち、パーティション１＿３００Ａが移動先として選択される。

【０３０６】

ステップ２７０４において、パーティション１＿３００Ａに含まれるストレージ装置１００に対応するエントリ（図１６（Ａ）の例ではエントリ１６１３）の残リソース量１６０５が参照される。

【０３０７】

ステップ２７０５において、ステップ２７０４で参照された残リソース量１６０５の値とストレージリソース量１６２８の値とが比較される。残リソース量１６０５の値がストレージリソース量１６２８の値より小さい場合には、アプリケーション５００Ａを移動すると、移動先のパーティション１＿３００Ａのリソース（例えばストレージ装置１００の容量）が不足するため、アプリケーション５００Ａを移動することはできない。図１６の例では、残リソース量１６０５の値「１０」がストレージリソース量１６２８の値「６」より大きいため、アプリケーション５００Ａを移動可能であると判定される。

【０３０８】

ステップ２７０６において、パーティション１＿３００Ａに含まれるＦＣ－ＳＷ１４０に対応するエントリ（図１６（Ａ）の例ではエントリ１６１２）の残リソース量１６０５が参照される。

【０３０９】

ステップ２７０７において、ステップ２７０６で参照された残リソース量１６０５の値とＦＣ－ＳＷリソース量１６２７の値とが比較される。残リソース量１６０５の値がＦＣ－ＳＷリソース量１６２７の値より小さい場合には、アプリケーション５００Ａを移動すると、移動先のパーティション１＿３００Ａのリソース（例えばＦＣ－ＳＷ１４０のポート１４１）が不足するため、アプリケーション５００Ａを移動することはできない。図１６の例では、残リソース量１６０５の値「１０」がＦＣ－ＳＷリソース量１６２７の値「１」より大きいため、アプリケーション５００Ａを移動可能であると判定される。

【０３１０】

ステップ２７０８において、パーティション１＿３００Ａに含まれるサーバ１２０に対応するエントリ（図１６（Ａ）の例ではエントリ１６１１）の残リソース量１６０５が参照される。

【０３１１】

ステップ２７０９において、ステップ２７０８で参照された残リソース量１６０５の値とサーバリソース量１６２６の値とが比較される。残リソース量１６０５の値がサーバリソース量１６２６の値より小さい場合には、アプリケーション５００Ａを移動すると、移動先のパーティション１＿３００Ａのリソース（例えばサーバ１２０のＣＰＵ１２１）が不足するため、アプリケーション５００Ａを移動することはできない。図１６の例では、残リソース量１６０５の値「１０」がサーバリソース量１６２６の値「５」より大きいため、アプリケーション５００Ａを移動可能であると判定される。

【０３１２】

図２８は、本発明の実施形態においてアプリケーション５００の移動が実行された後の管理テーブル１３６の説明図である。

【０３１３】

具体的には、図２８は、図１６に示すパーティション管理テーブル１３６Ａ及びアプリケーション管理テーブル１３６Ｂが、図１７から図２６に示す処理によって更新された結果を示す。すなわち、図２８は、図１４に示すアプリケーション５００Ａの移動が実行された後のパーティション管理テーブル１３６Ａ及びアプリケーション管理テーブル１３６

Ｂを示す。

【０３１４】

「ＡＰＰ１」によって識別されるアプリケーション５００Ａは、パーティション１＿３００Ａに移動した。このため、アプリケーション管理テーブル１３６Ｂにおいて、「ＡＰＰ１」に対応する配置パーティション番号１６２２、配置サーバ名１６２３、配置ＦＣ－ＳＷ名１６２４及び配置ストレージ名１６２５は、それぞれ、「１」、「サーバＰ１」、「ＦＣ－ＳＷＰ１」及び「ストレージＰ１」に更新される。

【０３１５】

さらに、パーティション管理テーブル１３６Ａにおいて、パーティション１＿３００Ａに対応するエントリ１６１１～１６１３の電源１６０６は全て「ＯＮ」、それらの配置アプリケーション１６０７は全て「ＡＰＰ１」に更新される。さらに、「サーバＰ１」に対応する残リソース量１６０５及びアプリケーションリソース量１６０８がそれぞれ「５」及び「５」に、「ＦＣ－ＳＷＰ１」に対応する残リソース量１６０５及びアプリケーションリソース量１６０８がそれぞれ「９」及び「１」に、「ストレージＰ１」に対応する残リソース量１６０５及びアプリケーションリソース量１６０８がそれぞれ「４」及び「６」に更新される。

【０３１６】

一方、パーティション２＿３００Ｂに対応するエントリ１６１４～１６１６の残リソース量１６０５はリソース量１６０４と同じ「１０」に、電源１６０６は全て「ＯＦＦ」に、配置アプリケーション１６０７及びアプリケーションリソース量１６０８は全て空白に更新される。

【０３１７】

図２９は、本発明の実施形態において実行されるアプリケーション５００の移動及び電源制御の処理の変形例の説明図である。

【０３１８】

既に説明した図１７の処理では、サーバ電源ＯＮ処理（ステップ１７０４）及びＦＣ－ＳＷ電源ＯＮ処理（ステップ１７０５）が、いずれも論理ボリューム６０３の全体コピー処理の終了を契機として開始される。これに対して、図２９では、サーバ電源ＯＮ処理及びＦＣ－ＳＷ電源ＯＮ処理の開始時刻が算出される。

【０３１９】

図１５を参照して説明したように、差分コピー処理、サーバ１２０の起動処理及びＦＣ－ＳＷ１４０の起動処理が同時刻に終了するように（すなわち、矢印１５０３、１５０４及び１５０５の先端が同一の時刻に対応するように）、サーバ電源ＯＮ処理及びＦＣ－ＳＷ電源ＯＮ処理を開始することが望ましい。しかし、差分コピー処理に要する時間を正確に予測することは困難であるため、図１７の処理では、全体コピー処理の終了時刻を、サーバ電源ＯＮ処理及びＦＣ－ＳＷ電源ＯＮ処理を開始する近似的に最適な時刻として用いている。

【０３２０】

これに対して、図２９の処理では、差分コピー処理に要する時間を近似的に算出することによって、サーバ電源ＯＮ処理及びＦＣ－ＳＷ電源ＯＮ処理を開始する時刻が設定される。

【０３２１】

以下、図２９の処理を説明する。図２９の処理は、図１７の処理の代わりに、すなわち、図１０のステップ１００２において実行される。

【０３２２】

管理サーバ１３０の管理プログラム１３５は、アプリケーション５００の移動処理を開始すると（ステップ２９０１）、最初に、ストレージ電源ＯＮ処理を実行する（ステップ２９０２）。これらの処理は、それぞれ図１７のステップ１７０１及びステップ１７０２と同様であるため、説明を省略する。

【０３２３】

ストレージ電源ＯＮ処理が終了すると、管理プログラム１３５は、論理ボリューム６０３の全体コピー処理を実行する（ステップ２９０３）。この処理は、図１７のステップ１７０３と同様である。

【０３２４】

論理ボリューム６０３の全体コピー処理が終了すると、管理プログラム１３５は、論理ボリューム６０３の差分コピー処理を実行する（ステップ２９０４）。この処理は、図１７のステップ１７０６と同様である。

【０３２５】

さらに、管理プログラム１３５は、ステップ２９０３及び２９０４と並行して、ステップ２９０５から２９０９を実行する。

【０３２６】

具体的には、管理プログラム１３５は、論理ボリューム６０３の全体コピー処理が開始された後、論理ボリューム６０３の差分コピー処理の終了予定時間を算出する（ステップ２９０５）。この処理は、全体コピー処理によってコピーされるべきデータのうち所定の割合（例えば８０％、９０％又は１００％等）のデータのコピーが終了したことを契機に実行される。ここでは、所定の割合が９０％である場合を例としてステップ２９０５を説明する。なお、１回の全体コピー処理において複数回（例えば、８０％が終了した時点及び９０％が終了した時点）ステップ２９０５が実行されてもよい。

【０３２７】

管理プログラム１３５は、全体コピー処理によってコピーされるべきデータのうち９０％のコピーが終了した時点で差分ストックとして格納されているデータ（以下、差分データ）の量から、全体コピー処理が全て（すなわち１００％）終了した時点の差分データ量を予測する。例えば、差分Ｉ／Ｏが一定の間隔で発生すると仮定した場合、全体コピー処理の９０％が終了した時点の差分データ量を０．９で除算する（すなわち０．９の逆数を乗ずる）ことによって、全体コピー処理が全て終了した時点の差分データ量を算出することができる。一般には、差分Ｉ／Ｏの間隔は一定ではないが、上記のように算出された値を近似的な予測値として使用してもよい。

【０３２８】

そして、管理プログラム１３５は、差分データ量の予測値と、コピーに使用されるデータ転送経路の転送速度から、差分データのコピーに要する時間を算出する。この算出の際に、データ転送経路の転送速度として、既に実行された全体コピー処理の際に実測された値が用いられてもよい。

【０３２９】

このようにして算出された時間を全体コピー処理の終了時刻に加算することによって、差分コピー処理の終了予定時刻が算出される。

【０３３０】

なお、既に説明したように、論理ボリューム６０３の全体コピー処理の終了時刻は、論理ボリューム６０３のデータ量及びデータ転送経路の転送速度に基づいて算出することができる。

【０３３１】

次に、管理プログラム１３５は、算出された差分コピー処理の終了予定時刻に基づいて、サーバ電源ＯＮ処理を開始する時刻及びＦＣ－ＳＷ電源ＯＮ処理を開始する時刻を設定する（それぞれステップ２９０６及びステップ２９０７）。

【０３３２】

具体的には、管理プログラム１３５は、差分コピー処理の終了予定時刻からサーバ１２０の起動処理に要する時間を減算することによってサーバ電源ＯＮ処理を開始する時刻を算出する。そして、その時刻が到来したら、サーバ電源ＯＮ処理を開始する（ステップ２９０８）。この処理は、図１７のステップ１７０４と同様である。

【０３３３】

さらに、差分コピー処理の終了予定時刻からＦＣ－ＳＷ１４０の起動処理に要する時間

を減算することによってＦＣ－ＳＷ電源ＯＮ処理を開始する時刻を算出する。そして、その時刻が到来したら、ＦＣ－ＳＷ電源ＯＮ処理を開始する（ステップ２９０９）。この処理は、図１７のステップ１７０５と同様である。

【０３３４】

なお、管理プログラム１３５は、サーバ１２０の起動処理に要する時間及びＦＣ－ＳＷ１４０の起動処理に要する時間を示す情報をあらかじめ保持している。これらの時間は、各装置のハードウェアスペックから算出されてもよいし、過去に実測された値であってもよい。

【０３３５】

ステップ２９０４、ステップ２９０８及びステップ２９０９のいずれも終了すると、管理プログラム１３５は、論理パス６０２の移動処理（ステップ２９１０）及び仮想サーバ６０１の移動処理（ステップ２９１１）を実行し、アプリケーション５００の移動処理を終了する（ステップ２９１２）。これらは、それぞれ、図１７のステップ１７０７、ステップ１７０８及びステップ１７０９と同様である。

【０３３６】

差分Ｉ／Ｏの偏在、及び、転送速度の変動の程度が小さいほど、上記のような差分コピー処理の終了時刻の予測は正確になる。その結果、計算機システムの消費電力を効果的に削減することができる。

【０３３７】

なお、上記のように、図２９の処理は、図１７の処理の代わりに実行されるものである。図１７の処理及び図２９の処理は、いずれも、計算機システムの消費電力を削減するために、サーバ１２０及びＦＣ－ＳＷ１４０の電源を近似的に最適な時刻に投入するためのものである。いずれの処理においても、サーバ１２０及びＦＣ－ＳＷ１４０の電源を投入する近似的に最適な時刻が、全体コピー処理及び差分コピー処理に要する時間に基づいて算出される。このように、図１７の処理及び図２９の処理の技術的な意義は共通する。

【０３３８】

図３０は、本発明の実施形態の計算機システムの規模と効果との関係を示す説明図である。

【０３３９】

図３０（Ａ）は、比較的小規模な計算機システム、例えば図３又は図４に示したように、一つの物理パーティションに含まれるサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００がそれぞれ１台～数台程度である計算機システムにおけるアプリケーションの移動及び電源制御のタイミングを示す。図３０（Ａ）の内容は、図１５に示したものと同様であるため、説明を省略する。

【０３４０】

図３０（Ｂ）は、比較的大規模なシステム、例えば、一つの物理パーティションに含まれるサーバ１２０、ＦＣ－ＳＷ１４０及びストレージ装置１００がそれぞれ数十台～数百台程度である計算機システムにおけるアプリケーションの移動及び電源制御のタイミングを示す。このような物理パーティションは、例えばデータセンタとして使用される。図３０（Ｂ）の内容も、基本的には図１５に示したものと同様である。しかし、一般に、計算機システムの規模が大きくなるほど、各装置の起動及びアプリケーションの移動に要する時間は長くなる。

【０３４１】

例えば、論理ボリューム６０３の移動に要する時間は、前述のように転送されるデータ量等に依存するが、一般には小規模なシステムでは数時間程度である場合が多い。これに対して、データセンタのような大規模なシステムでは、論理ボリューム６０３の移動に数週間以上を要する場合もある。

【０３４２】

本実施形態では、図１５を参照して説明したように、サーバ１２０及びＦＣ－ＳＷ１４０の電源投入時刻を論理ボリューム６０３の移動の開始時刻より遅らせることによって、

論理ボリューム６０３の移動の終了時刻と、サーバ１２０及びＦＣ－ＳＷ１４０の起動処理の終了時刻とが概ね一致するように制御される。すなわち、論理ボリューム６０３の移動が実行されている時間の大部分（例えば、全体コピー処理が実行されている時間）は、サーバ１２０及びＦＣ－ＳＷ１４０の電源が投入されていない。このため、論理ボリューム６０３の移動に長時間を要する大規模な計算機システムに適用するほど、本実施形態の効果は大きくなる。

【０３４３】

図３１は、本発明の実施形態における物理パーティションの第４の例の説明図である。

【０３４４】

図３１に示す物理パーティションは、例えばデータセンタとして使用される大規模なものである。図３１の例では、パーティション１＿３００Ａ及びパーティション２＿３００Ｂそれぞれが２０台のサーバ１２０、９台のＦＣ－ＳＷ１４０及び３台のストレージ装置１００を備える。なお、これらの台数は一例であり、大規模なデータセンタではさらに台数が多い場合もある。

【０３４５】

このような計算機システムにおいて、データセンタを移動（マイグレーション）するために、パーティション内に配置されたアプリケーションを移動する必要が生じる場合がある。例えば、移動すべきアプリケーションに含まれる論理ボリューム群（すなわち論理ボリューム６０３の集合）３１０１Ａを移動先のパーティション内の論理ボリューム群３１０１Ｆに移動するために、図３０において説明したように数週間以上を要する場合もある。このような計算機システムに本実施形態を適用することによって、サーバ１２０及びＦＣ－ＳＷ１４０の消費電力を削減することができる。

【０３４６】

仮に、本実施形態を適用することによって、サーバ１２０及びＦＣ－ＳＷ１４０の電源投入時刻を従来より１日遅らせることができた場合、１日分のサーバ１２０及びＦＣ－ＳＷ１４０の消費電力が削減される。サーバ１２０及びＦＣ－ＳＷ１４０の台数が多いほど、消費電力削減の効果が大きくなる。

【０３４７】

図３２は、本発明の実施形態における物理パーティションの第５の例の説明図である。

【０３４８】

図３２は、パーティションによって階層化されたデータセンタの例を示す。この例において、パーティション１＿３００Ａは高い階層、パーティション２＿３００Ｂは低い階層に相当する。すなわち、パーティション１＿３００Ａは、パーティション２＿３００Ｂより高い処理性能を有する。例えば、パーティション１＿３００Ａには、パーティション２＿３００Ｂに割り当てられているものより高性能のハードウェアが割り当てられてもよい。

【０３４９】

例えば、あるプロジェクトで使用されるアプリケーション群を、ある期間のみ、高い階層の（すなわち高性能の）パーティションで実行する必要がある場合、その期間の開始時点にアプリケーション群の移動が終了するように、あらかじめそのアプリケーション群の論理ボリューム６０３のコピーを開始することができる。

【０３５０】

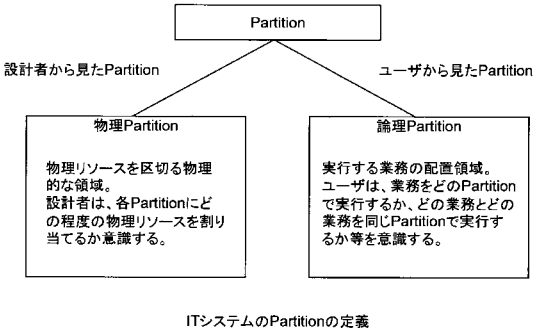
図３２の例では、低い階層（Ｔｉｅｒ２）に相当するパーティション２＿３００Ｂに、プロジェクトＡで使用されるアプリケーション群３２０１Ａ及びプロジェクト（又は業務）Ｂで使用されるアプリケーション群３２０１Ｂが配置されている。このような場合、本実施形態によって、アプリケーション群ごとに、必要に応じて高い階層（Ｔｉｅｒ１）のパーティション１＿３００Ａへの移動を制御することができる。これによって、計算機システムの消費電力を削減しつつ、所望の時刻におけるアプリケーション群の移動を実現することができる。

【図面の簡単な説明】

【０３５１】
【図１】本発明の実施形態におけるパーティションの説明図である。
【図２】本発明の実施形態の計算機システムの構成を示すブロック図である。
【図３】本発明の実施形態における物理パーティションの第１の例の説明図である。
【図４】本発明の実施形態における物理パーティションの第２の例の説明図である。
【図５】本発明の実施形態における論理パーティションの例の説明図である。
【図６】本発明の実施形態におけるアプリケーションの構成の説明図である。
【図７】本発明の実施形態の全体システム構成の説明図である。
【図８】本発明の実施形態において実行されるアプリケーションのパーティション間の移動処理を示すフローチャートである。
【図９】本発明の実施形態において実行されるアプリケーションのパーティション間の移動処理の詳細な手順を示す説明図である。
【図１０】本発明の実施形態において実行されるアプリケーションのパーティション間の移動処理の詳細な手順を示すフローチャートである。
【図１１】本発明の実施形態において実行される論理ボリュームの移動の説明図である。
【図１２】本発明の実施形態において実行される論理ボリュームの移動の別の例の説明図である。
【図１３】本発明の実施形態における物理パーティションの第３の例の説明図である。
【図１４】本発明の実施形態において実行されるアプリケーションの移動及び電源制御の説明図である。
【図１５】本発明の実施形態において実行されるアプリケーションの移動及び電源制御のタイミングの説明図である。
【図１６】本発明の実施形態の管理テーブルの説明図である。
【図１７】本発明の実施形態において実行されるアプリケーションの移動及び電源制御の処理の全体を示すフローチャートである。
【図１８】本発明の実施形態において実行されるストレージ電源ＯＮ処理を示すフローチャートである。
【図１９】本発明の実施形態の管理サーバが実行する論理ボリュームの全体コピー処理のフローチャートである。
【図２０】本発明の実施形態のストレージ装置が実行する論理ボリュームの全体コピー処理のフローチャートである。
【図２１】本発明の実施形態において実行されるサーバ電源ＯＮ処理を示すフローチャートである。
【図２２】本発明の実施形態において実行されるＦＣ－ＳＷ電源ＯＮ処理を示すフローチャートである。
【図２３】本発明の実施形態の管理サーバが実行する論理ボリュームの差分コピー処理のフローチャートである。
【図２４】本発明の実施形態のストレージ装置が実行する論理ボリュームの差分コピー処理のフローチャートである。
【図２５】本発明の実施形態において実行される論理パスの移動処理のフローチャートである。
【図２６】本発明の実施形態において実行される仮想サーバの移動処理のフローチャートである。
【図２７】本発明の実施形態において実行される移動先パーティションの状態確認処理のフローチャートである。
【図２８】本発明の実施形態においてアプリケーションの移動が実行された後の管理テーブルの説明図である。
【図２９】本発明の実施形態において実行されるアプリケーションの移動及び電源制御の処理の変形例の説明図である。
【図３０】本発明の実施形態の計算機システムの規模と効果との関係を示す説明図である

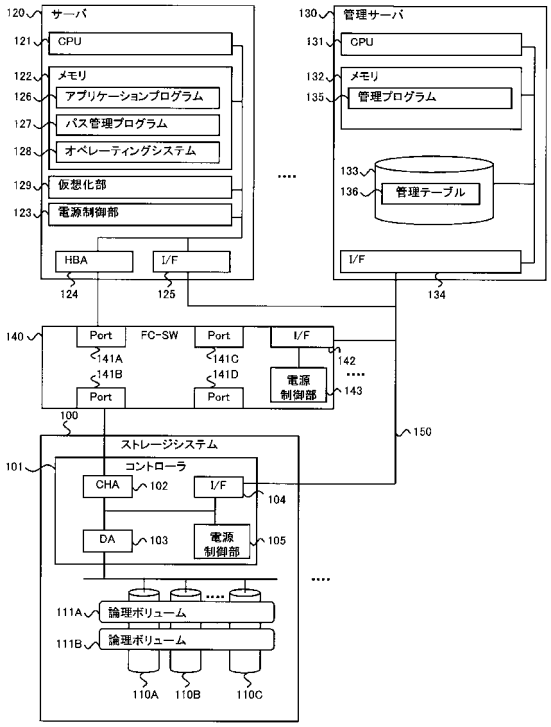
。
【図３１】本発明の実施形態における物理パーティションの第４の例の説明図である。
【図３２】本発明の実施形態における物理パーティションの第５の例の説明図である。
【符号の説明】
【０３５２】
１００　ストレージ装置
１０１　コントローラ
１１０Ａ～１１０Ｃ　ディスクドライブ
１１１Ａ、１１１Ｂ、６０３、６０３Ａ～６０３Ｆ　論理ボリューム
１２０　サーバ
１３０　管理サーバ
１３５　管理プログラム
１３６　管理テーブル
１３６Ａ　パーティション管理テーブル
１３６Ｂ　アプリケーション管理テーブル
３００Ａ～３００Ｉ　パーティション
５００、５００Ａ～５００Ｅ　アプリケーション
６０１、６０１Ａ～６０１Ｅ　仮想サーバ
６０２、６０２Ａ～６０２Ｅ　ＦＣ－ＳＷ

【図１】

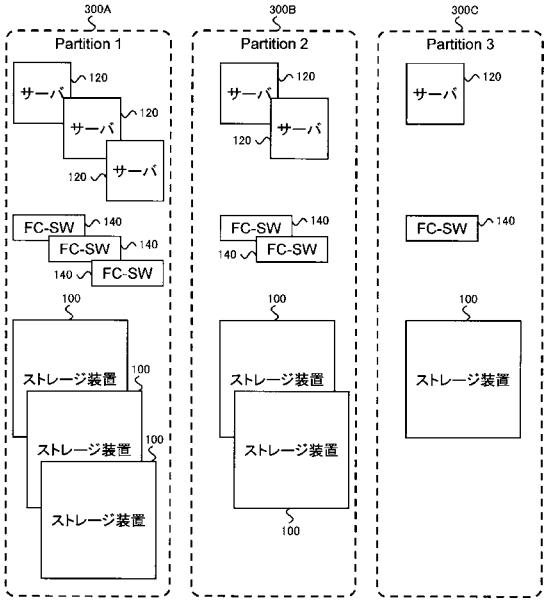



ITシステムのPartitionの定義

【図２】

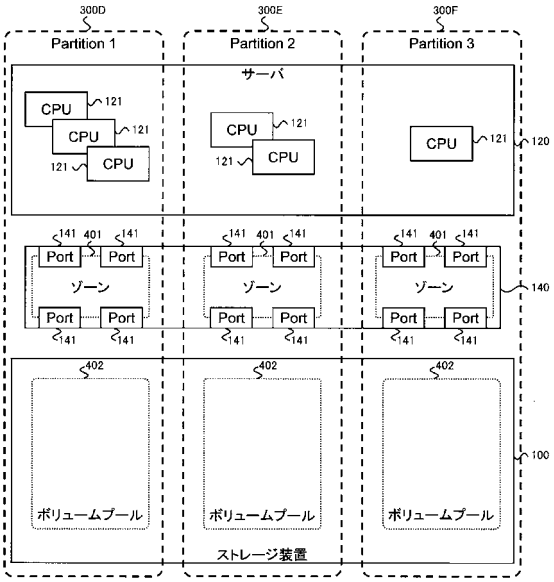


【図３】



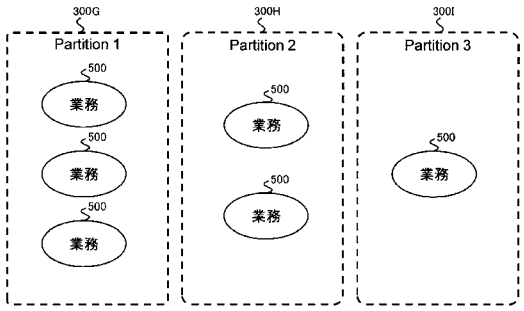

物理パーティションの第１の例

【図４】



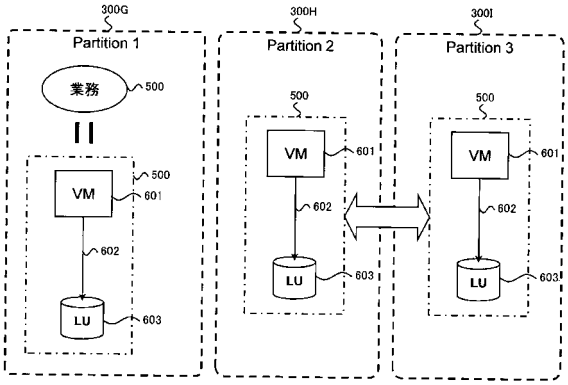

物理パーティションの第２の例

【図５】

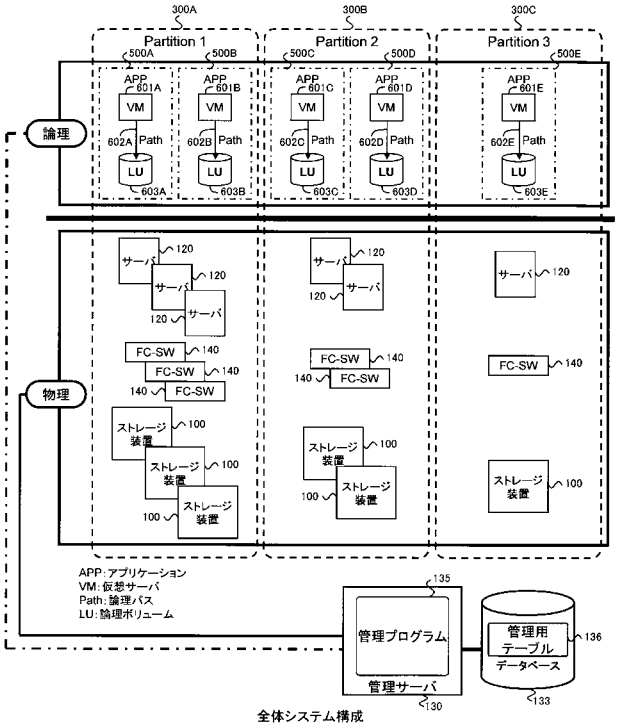



論理パーティション

【図６】

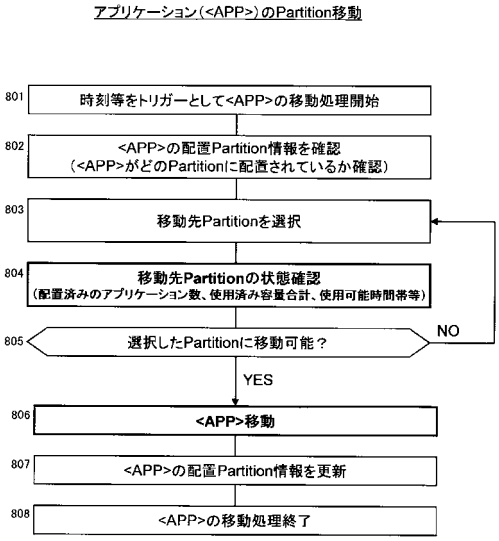


【図７】

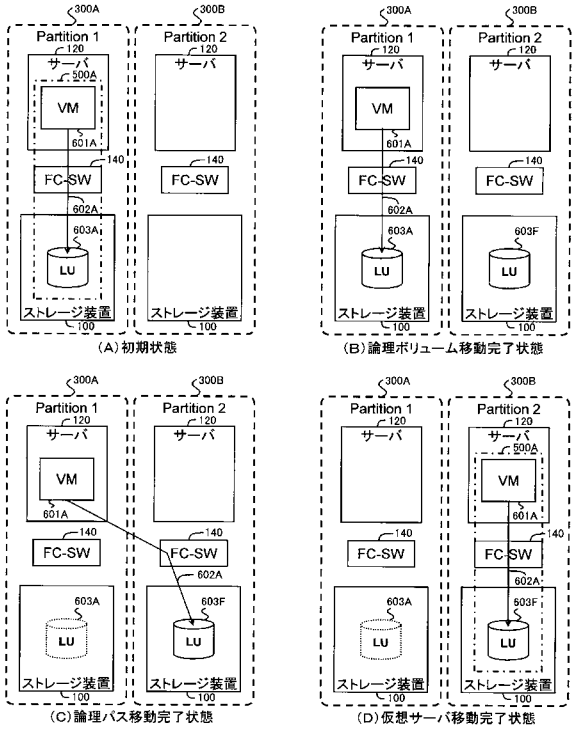



全体システム構成

【図８】

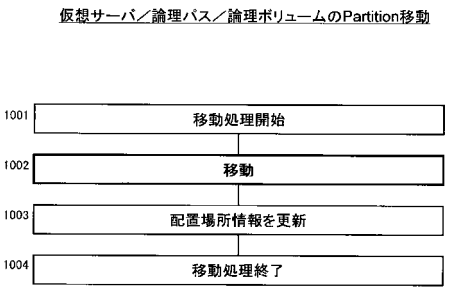



【図９】

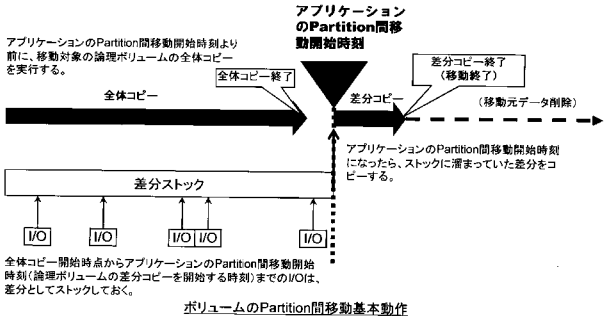



アプリケーションのパーティション間移動

【図１０】

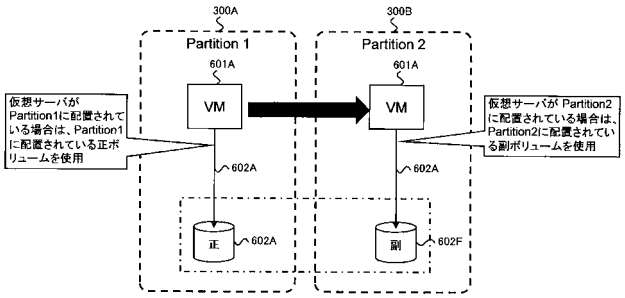


【図１１】



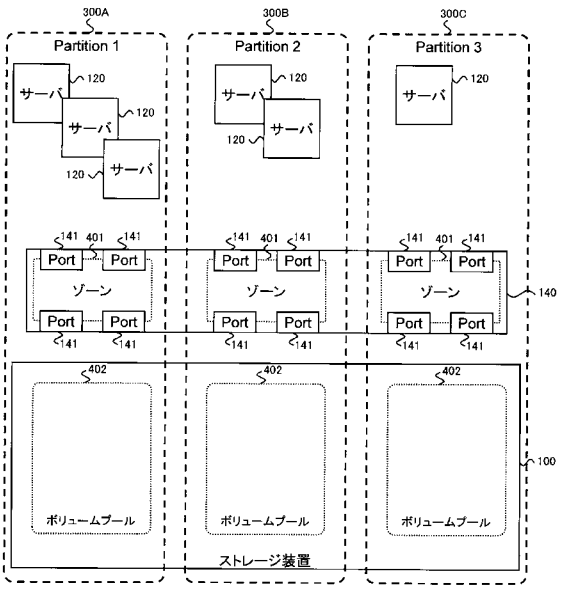

ボリュームのPartition間移動基本動作

【図１２】




ボリュームペアを利用する例

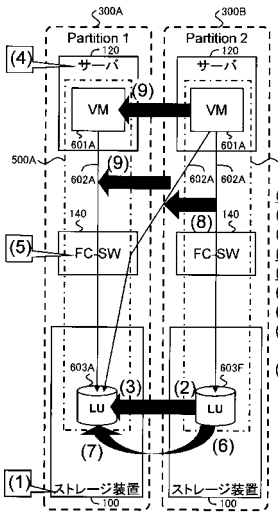
【図１３】




物理パーティションの第3の例（サーバのみ電源断）

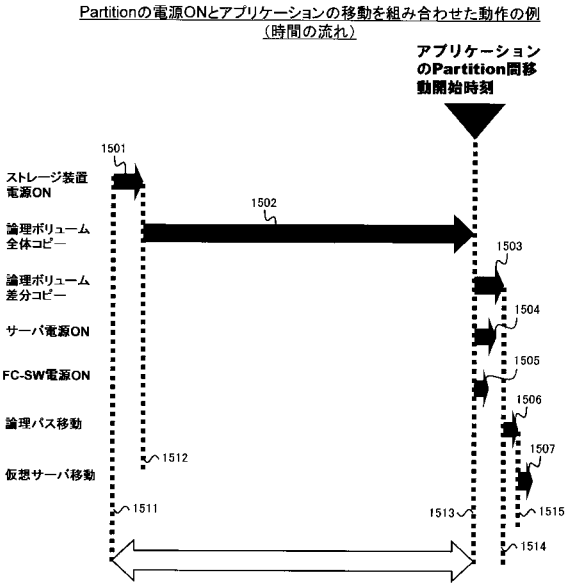
【図１４】

Partitionの電源ONとアプリケーションの移動を組み合わせた動作の例




(1) Partition1のストレージ装置の電源ON
(2) Partition1へ論理ボリューム全体のコピー開始
(3) Partition1への論理ボリューム全体コピー終了
(4) Partition1のサーバの電源ON
(5) Partition1のFC-SWの電源ON
(6) Partition1へ論理ボリュームの差分コピー開始
(7) Partition1への論理ボリューム差分コピー終了
(8) 論理パスをPartition1側の論理ボリュームへ移動
(9) 仮想サーバをPartition1へ移動

【図１５】




【図１６】

パーティション管理テーブル 136A

(A)

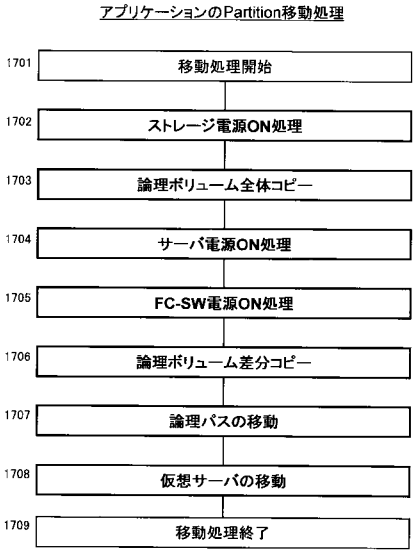
| 1601 パーティション番号 | 1602 ハードウェア種別 | 1603 ハードウェア名 | 1604 リソース量 | 1605 残リソース量 | 1606 電源 | 1607 配置APP | 1608 APPリソース量 | 136A |
|---|---|---|---|---|---|---|---|---|
| 1 | サーバ1 | サーバP1 | 10 | 10 | OFF | | | 1611 |
| | FC-SW1 | FC-SWP1 | 10 | 10 | OFF | | | 1612 |
| | ストレージ1 | ストレージP1 | 10 | 10 | OFF | | | 1613 |
| 2 | サーバ2 | サーバP2 | 10 | 5 | ON | APP1 | 5 | 1614 |
| | FC-SW2 | FC-SWP2 | 10 | 9 | ON | APP1 | 1 | 1615 |
| | ストレージ2 | ストレージP2 | 10 | 4 | ON | APP1 | 6 | 1616 |

アプリケーション管理テーブル 136B

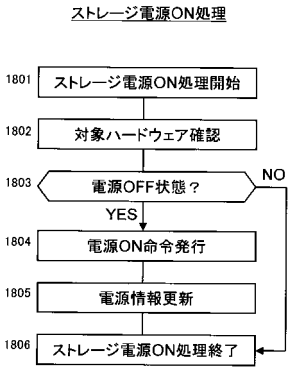
(B)

| 1621 アプリケーション名 | 1622 配置Partition# | 1623 配置サーバ名 | 1624 配置FC-SW名 | 1625 配置ストレージ名 | 1626 サーバリソース量 | 1627 FC-SWリソース量 | 1628 ストレージリソース量 | 136B |
|---|---|---|---|---|---|---|---|---|
| APP1 | 2 | サーバP2 | FC-SWP2 | ストレージP2 | 5 | 1 | 6 | |

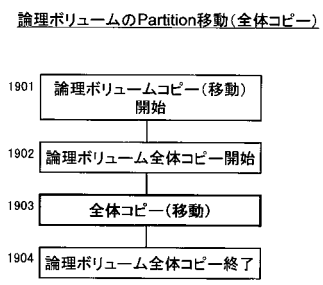
【図１７】



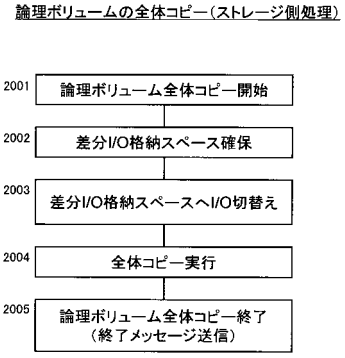

【図１８】



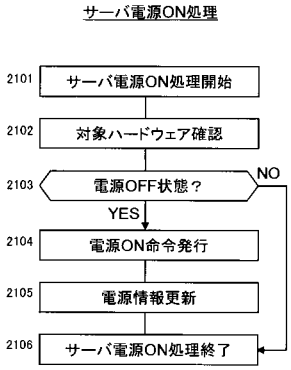
【図１９】



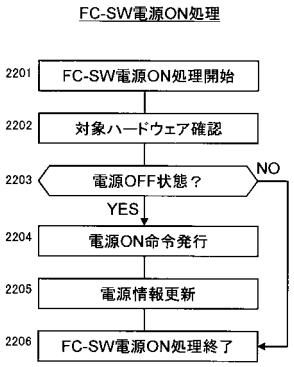

【図２０】



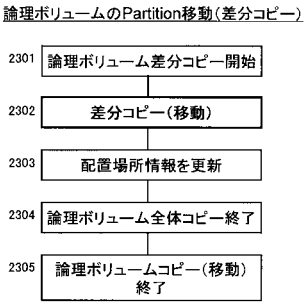

【図２１】



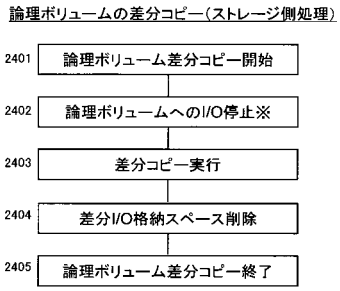

【図２２】



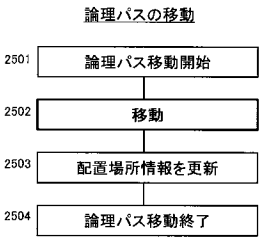
【図２３】



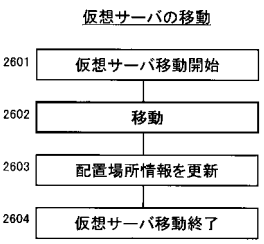

【図２４】



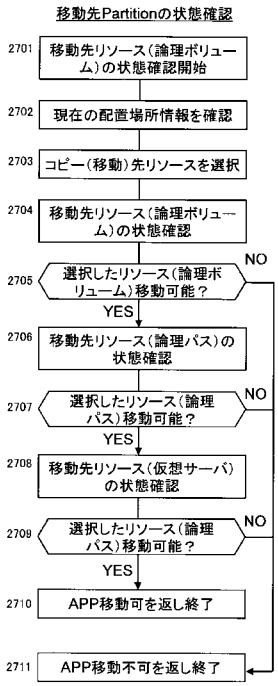

【図２５】




【図２６】



【図２７】




【図２８】

パーティション管理テーブル 136A

(A)

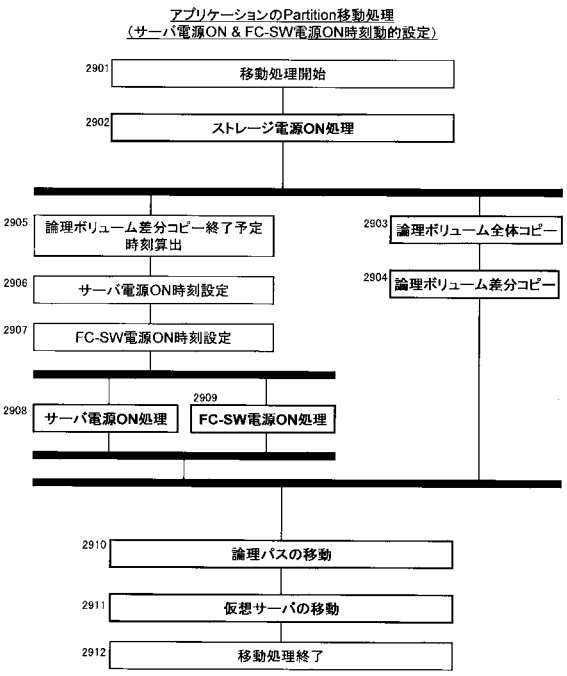
| 1601 パーティション番号 | 1602 ハードウェア種別 | 1603 ハードウェア名 | 1604 リソース量 | 1605 残リソース量 | 1606 電源 | 1607 配置APP | 1608 APPリソース量 | 136A |
|---|---|---|---|---|---|---|---|---|
| 1 | サーバ1 | サーバP1 | 10 | 5 | ON | APP1 | 5 | 1611 |
| | FC-SW1 | FC-SWP1 | 10 | 9 | ON | APP1 | 1 | 1612 |
| | ストレージ1 | ストレージP1 | 10 | 4 | ON | APP1 | 6 | 1613 |
| 2 | サーバ2 | サーバP2 | 10 | 10 | OFF | | | 1614 |
| | FC-SW2 | FC-SWP2 | 10 | 10 | OFF | | | 1615 |
| | ストレージ2 | ストレージP2 | 10 | 10 | OFF | | | 1616 |

アプリケーション管理テーブル 136B

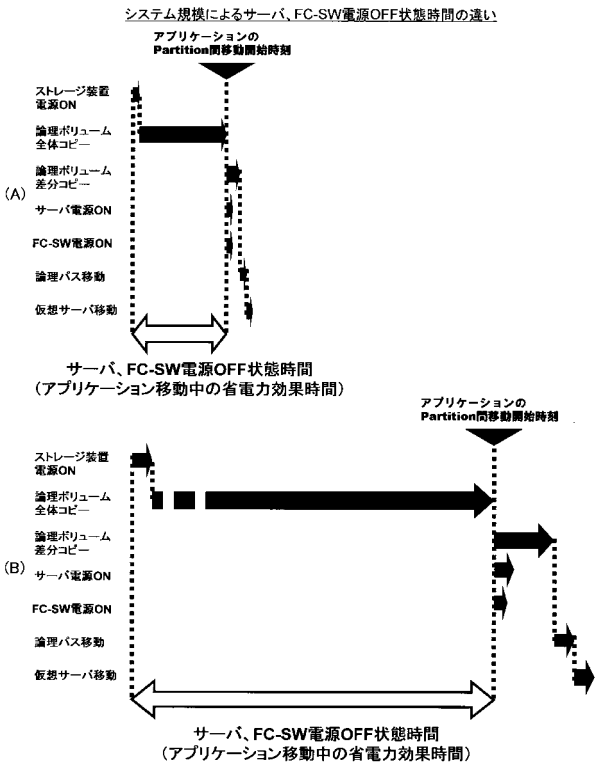
(B)

| 1621 アプリケーション名 | 1622 配置Partition# | 1623 配置サーバ名 | 1624 配置FC-SW名 | 1625 配置ストレージ名 | 1626 サーバリソース量 | 1627 FC-SWリソース量 | 1628 ストレージリソース量 | 136B |
|---|---|---|---|---|---|---|---|---|
| APP1 | 1 | サーバP1 | FC-SWP1 | ストレージP1 | 5 | 1 | 6 | |

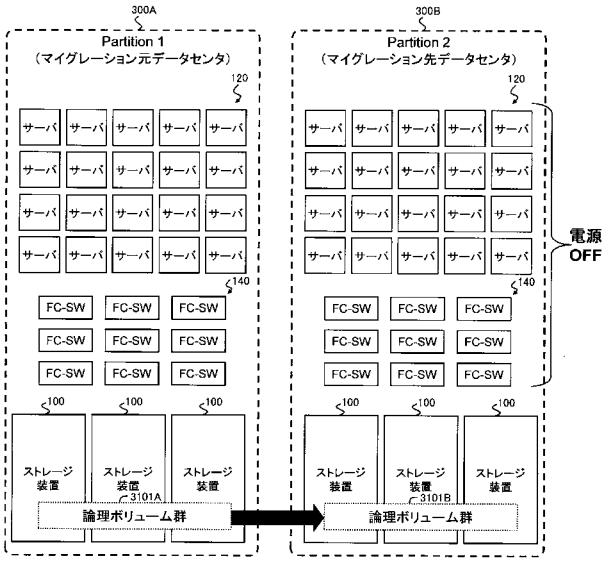
【図２９】




【図３０】

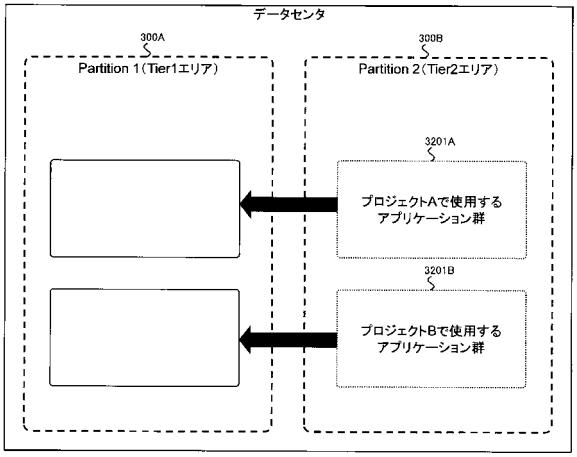


【図３１】




物理パーティションの第４の例
（データセンタを1つのPartitionと定義する例）

【図３２】




物理パーティションの第５の例
（データセンタをTierのエリアで区切る例）