(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第6069742号
(P6069742)

(45)発行日　平成29年2月1日(2017.2.1)　(24)登録日　平成29年1月13日(2017.1.13)

(51)Int.Cl.		F I		
G06F 3/06	(2006.01)	G06F	3/06	301J
G06F 11/10	(2006.01)	G06F	3/06	305C
		G06F	3/06	540
		G06F	11/10	676

請求項の数 22　(全 36 頁)

(21)出願番号　特願2015-530281 (P2015-530281)
(86)(22)出願日　平成25年8月9日 (2013.8.9)
(65)公表番号　特表2015-528973 (P2015-528973A)
(43)公表日　平成27年10月1日 (2015.10.1)
(86)国際出願番号　PCT/CN2013/081182
(87)国際公開番号　WO2015/018065
(87)国際公開日　平成27年2月12日 (2015.2.12)
審査請求日　平成27年1月6日 (2015.1.6)

(73)特許権者　504277388
▲ホア▼▲ウェイ▼技術有限公司
HUAWEI TECHNOLOGIES CO., LTD.
中華人民共和国518129広東省深▲セン▼市龍岡区坂田華為本社ビル
Huawei Administration Building, Bantian, Longgang District Shenzhen, Guangdong 518129 (CN)
(74)代理人　100146835
弁理士　佐伯　義文
(74)代理人　100140534
弁理士　木内　敬二

最終頁に続く

(54)【発明の名称】ファイル処理方法およびファイル処理装置、ならびに記憶デバイス

(57)【特許請求の範囲】

【請求項１】

独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)内に記憶されることになるF個のファイルを受信するステップであって、前記RAIDがT個の記憶装置によって形成され、Fが2以上の自然数であり、Tが3以上の自然数である、受信するステップと、

前記RAIDのストリップサイズに従って、前記F個のファイルの各々を1以上のデータブロックに分割するステップと、

前記F個のファイルのデータブロックに従って、T個の行を有する第1の行列を取得するステップであって、前記各ファイルの全ての前記データブロックが異なる行に現われないように、F個のファイルの各々の全てのデータブロックが前記第1の行列内の1個且つ1個だけの行内に配置され、前記第1の行列の列の各々はデータブロックと、対応する列内の前記データブロックを計算することによって取得されたチェックブロックとを含み、各列の2個のデータブロックが同じファイルに属することはない、取得するステップと、

前記第1の行列内の各列内のデータブロックと、前記列内の前記データブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込むステップと

を含むファイル処理方法。

【請求項２】

前記F個のファイルのデータブロックに従って、T個の行を有する第1の行列を取得する

前記ステップが、
分割することによって取得された、前記F個のファイルの前記データブロックをD個の行を有する第2の行列に配列するステップであって、前記各ファイルの全ての前記データブロックが前記第2の行列の異なる行に現われないように、F個のファイルの各々の全てのデータブロックが前記第1の行列内の1個且つ1個だけの行内に配置され、Dが前記RAID内のデータ記憶装置の数量である、配列するステップと、
チェックブロックを前記第2の行列の各列にそれぞれ挿入することによって、T個の行を有する前記第1の行列を取得するステップであって、前記挿入されたチェックブロックが、前記第1の行列内の前記チェックブロックが配置された列内のデータブロックに従って計算することによって取得される、取得するステップと
を含む、請求項1に記載のファイル処理方法。
【請求項３】
前記RAIDが独立チェック記憶装置を含むとき、チェックブロックを前記第2の行列の各列にそれぞれ挿入することによって、T個の行を有する前記第1の行列を取得する前記ステップが、
前記RAID内の前記独立チェック記憶装置の位置に従って、チェックブロックを前記第2の行列に挿入するための位置を判断するステップと、
前記RAIDのチェックアルゴリズムに従って、前記第2の行列の各列内の前記データブロックに関するチェック計算を実行して、各列内の前記データブロックのチェックブロックを取得するステップと、
第2の行列の各列内の前記データブロックに従って計算することによって取得された前記チェックブロックの前記判断された位置に従って、前記チェックブロックを前記各列に挿入することによって、T個の行を有する前記第1の行列を取得するステップと
を含む、請求項2に記載のファイル処理方法。
【請求項４】
前記RAIDが独立チェック記憶装置を含まないとき、チェックブロックを前記第2の行列の各列にそれぞれ挿入することによって、T個の行を有する前記第1の行列を取得する前記ステップが、
チェックブロックを前記第2の行列の各列に挿入するための位置A[x,y]を判断するステップであって、前記第2の行列がN個の列を有し、xとyが両方とも整数であり、xの値が0からD-1に徐々に増大し、yの値が0からN-1に徐々に増大する、判断するステップと、
前記第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを前記第x番目の行内の第(y+1)番目の列から第N番目の列までの位置に順次に移動させるステップと、
前記RAIDのチェックアルゴリズムに従って、第ｙ番目の列内の前記データブロックに関するチェック計算を実行して、第ｙ番目の列内の前記データブロックのチェックブロックを取得するステップと、
第ｙ番目の列内の前記データブロックの前記チェックブロックを前記第2の行列の第ｙ番目の列内の前記位置A[x,y]に挿入することによって、T個の行を有する前記第1の行列を取得するステップと
を含む、請求項2に記載のファイル処理方法。
【請求項５】
前記第1の行列内の各列内のデータブロックと、前記列内の前記データブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込む前記ステップが、
前記第1の行列の第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記データブロックに従って計算することによって取得された前記チェックブロックとからなるストライプが完全に占有されているとき、第ｙ番目の列内の前記データブロックと、前記チェックブロックとを、前記RAIDを形成する前記T個の記憶装置内に書き込むステップであって、第ｙ番目の列が前記第1の行列内の前記列のうちの1つである、書き込むステップを

含む、請求項1から4のいずれか一項に記載のファイル処理方法。
【請求項６】
前記第1の行列がM個の列を有し、前記第1の行列内の各列内のデータブロックと、前記列内の前記データブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込む前記ステップが、
前記第1の行列の第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記データブロックに従って計算することによって取得された前記チェックブロックとからなるストライプが完全に占有されていないとき、第ｙ番目の列内の欠けているデータブロックの数量を判断するステップであって、第ｙ番目の列が前記第1の行列内の前記列のうちの1つである、判断するステップと、
第ｙ番目の列内のデータブロックを有さない位置を前記第1の行列内の第(M-1)番目の列から第(y+1)番目の列までから選択された数量のデータブロックで充填するステップと、
充填した後の第ｙ番目の列内の前記データブロックに従って、第ｙ番目の列内の前記チェックブロックを更新するステップと、
第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記更新されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込むステップと
を含む、請求項1から4のいずれか一項に記載のファイル処理方法。
【請求項７】
前記第1の行列内の各列内のデータブロックと、前記列内の前記データブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込む前記ステップが、
前記第1の行列の第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記データブロックに従って計算することによって取得された前記チェックブロックとからなるストライプが完全に占有されていないとき、第ｙ番目の列内のデータブロックを有さない位置を0で充填するステップと、
0で充填した後の第ｙ番目の列内の前記データブロックと、前記チェックブロックとからなるストライプを、前記RAIDを形成する記憶装置内に書き込むステップであって、第ｙ番目の列が前記第1の行列内の前記列のうちの1つである、書き込むステップと
を含む、請求項1から4のいずれか一項に記載のファイル処理方法。
【請求項８】
ホストのアクセス要求を受信するステップであって、前記アクセス要求が前記RAID内に記憶されたファイルを読み取るために使用され、前記アクセス要求が前記ファイルに関する論理アドレスを搬送する、受信するステップと、
前記論理アドレスに従って、前記ファイルのデータブロックが記憶された物理アドレスに問い合わせるステップと、
前記物理アドレスに従って、前記ファイルが記憶された1個の記憶装置を判断するステップと、
前記記憶装置内に記憶された前記ファイルの前記データブロックを前記ホストに返すステップと
をさらに含む、請求項1から7のいずれか一項に記載のファイル処理方法。
【請求項９】
独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)内に記憶されることになるF個のファイルを受信するステップであって、Fが2以上の自然数である、受信するステップと、
前記RAIDのストリップサイズに従って、前記F個のファイルの各々を1以上のデータブロックに分割するステップと、
前記F個のファイルの前記データブロックを1次元アレイに配列するステップであって、前記アレイ内に、1個のファイルに属する2個の隣接するデータブロック同士の間にD-1個

の位置の間隔が存在し、Dの値が前記RAID内のデータ記憶装置の数量であり、FはD以上である、配列するステップと、
前記アレイのD個のデータブロックの順序と、前記D個のデータブロックの順序に従って計算することによって取得されたP個のチェックブロックとからなるストライプを、前記RAIDを形成する記憶装置内に書き込むステップであって、Pの値が前記RAID内の独立チェック記憶装置の数量である、書き込むステップと
を含むファイル処理方法。
【請求項１０】
ホストのアクセス要求を受信するステップであって、前記アクセス要求が前記RAID内に記憶されたファイルを読み取るために使用され、前記アクセス要求が前記ファイルに関する論理アドレスを搬送する、受信するステップと、
前記論理アドレスに従って、前記ファイルのデータブロックが記憶された物理アドレスに問い合わせるステップと、
前記物理アドレスに従って、前記ファイルが記憶された1個の記憶装置を判断するステップと、
前記記憶装置内に記憶された前記ファイルの前記データブロックを前記ホストに返すステップと
をさらに含む、請求項9に記載のファイル処理方法。
【請求項１１】
独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)内に記憶されることになるF個のファイルを受信するように構成された受信モジュールであって、前記RAIDがT個の記憶装置によって形成され、Fが2以上の自然数であり、Tが3以上の自然数である、受信モジュールと、
前記RAIDのストリップサイズに従って、前記F個のファイルの各々を1以上のデータブロックに分割するように構成された分割モジュールと、
前記F個のファイルのデータブロックに従って、T個の行を有する第1の行列を取得するように構成された処理モジュールであって、前記各ファイルの全ての前記データブロックが異なる行に現われないように、F個のファイルの各々の全てのデータブロックが前記第1の行列内の1個且つ1個だけの行内に配置され、前記第1の行列の列の各々はデータブロックと、対応する列内の前記データブロックを計算することによって取得されたチェックブロックとを含み、各列の2個のデータブロックが同じファイルに属することはない、処理モジュールと、
前記第1の行列内の各列内のデータブロックと、前記列内の前記データブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込むように構成された書込みモジュールと
を含むファイル処理装置。
【請求項１２】
前記処理モジュールが、具体的には、
分割することによって取得された、前記F個のファイルの前記データブロックをD個の行を有する第2の行列に配列することであって、前記各ファイルの全ての前記データブロックが前記第2の行列の異なる行に現われないように、F個のファイルの各々の全てのデータブロックが前記第1の行列内の1個且つ1個だけの行内に配置され、Dが前記RAID内のデータ記憶装置の数量である、配列することと、
チェックブロックを前記第2の行列の各列にそれぞれ挿入することによって、T個の行を有する前記第1の行列を取得することであって、前記挿入されたチェックブロックが、前記第1の行列内の前記チェックブロックが配置された列内のデータブロックに従って計算することによって取得される、取得することと
を行うように構成される、請求項11に記載のファイル処理装置。
【請求項１３】
前記RAIDが独立チェック記憶装置を含むとき、前記処理モジュールが、具体的には、

前記RAID内の前記独立チェック記憶装置の位置に従って、チェックブロックを前記第2の行列に挿入するための位置を判断することと、
前記RAIDのチェックアルゴリズムに従って、前記第2の行列の各列内の前記データブロックに関するチェック計算を実行して、各列内の前記データブロックのチェックブロックを取得することと、
前記第2の行列の各列内の前記データブロックに従って計算することによって取得された前記チェックブロックの前記判断された位置に従って、前記チェックブロックを前記列に挿入することによって、T個の行を有する前記第1の行列を取得することと
を行うように構成される、請求項12に記載のファイル処理装置。
【請求項１４】
前記RAIDが独立チェック記憶装置を含まないとき、前記処理モジュールが、具体的には、
チェックブロックを前記第2の行列の各列に挿入するための位置A[x,y]を判断することであって、前記第2の行列がN個の列を有し、xとyが両方とも整数であり、xの値が0からD-1に徐々に増大し、yの値が0からN-1に徐々に増大する、判断することと、
前記第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを前記第x番目の行内の第(y+1)番目の列から第N番目の列までの位置に順次に移動させることと、
前記RAIDのチェックアルゴリズムに従って、第ｙ番目の列内の前記データブロックに関するチェック計算を実行して、第ｙ番目の列内の前記データブロックのチェックブロックを取得することと、
第ｙ番目の列内の前記データブロックの前記チェックブロックを前記第2の行列の第ｙ番目の列内の前記位置A[x,y]に挿入することによって、T個の行を有する前記第1の行列を取得することと
を行うように構成される、請求項12に記載のファイル処理装置。
【請求項１５】
前記書込みモジュールが、具体的には、
前記第1の行列の第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記データブロックに従って計算することによって取得された前記チェックブロックとからなるストライプが完全に占有されているとき、第ｙ番目の列内の前記データブロックと、前記チェックブロックとを、前記RAIDを形成する前記T個の記憶装置内に書き込むことであって、第ｙ番目の列が前記第1の行列内の前記列のうちの1つである、書き込むこと
を行うように構成される、請求項11から14のいずれか一項に記載のファイル処理装置。
【請求項１６】
前記第1の行列がM個の列を有し、前記書込みモジュールが、具体的には、
前記第1の行列の第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記データブロックに従って計算することによって取得された前記チェックブロックとからなるストライプが完全に占有されていないとき、第ｙ番目の列内の欠けているデータブロックの数量を判断することであって、第ｙ番目の列が前記第1の行列内の前記列のうちの1つである、判断することと、
第ｙ番目の列内のデータブロックを有さない位置を前記第1の行列内の第(M-1)番目の列から第(y+1)番目の列までから選択された数量のデータブロックで充填するステップと、
充填した後の第ｙ番目の列内の前記データブロックに従って、第ｙ番目の列内の前記チェックブロックを更新することと、
第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記更新されたチェックブロックとからなるストライプを、前記RAIDを形成する前記T個の記憶装置内に書き込むことと
を行うように構成される、請求項11から14のいずれか一項に記載のファイル処理装置。
【請求項１７】
前記書込みモジュールが、具体的には、

前記第1の行列の第ｙ番目の列内の前記データブロックと、第ｙ番目の列内の前記データブロックに従って計算することによって取得された前記チェックブロックとからなるストライプが完全に占有されていないとき、第ｙ番目の列内のデータブロックを有さない位置を0で充填することと、

0で充填した後の第ｙ番目の列内の前記データブロックと、前記チェックブロックとからなるストライプを、前記RAIDを形成する記憶装置内に書き込むことであって、第ｙ番目の列が前記第1の行列内の列のうちの1つである、書き込むことと

を行うように構成される、請求項11から14のいずれか一項に記載のファイル処理装置。

【請求項１８】

前記受信モジュールが、ホストのアクセス要求を受信することであって、前記アクセス要求が前記RAID内に記憶されたファイルを読み取るために使用され、前記アクセス要求が前記ファイルに関する論理アドレスを搬送する、受信することを行うようにさらに構成され、

前記ファイル処理装置が、

前記論理アドレスに従って、前記ファイルのデータブロックが記憶された物理アドレスに問い合わせて、前記物理アドレスに従って、前記ファイルが記憶された1個の記憶装置を判断して、前記RAIDの前記1個の記憶装置内に記憶された前記ファイルの前記データブロックを前記ホストに返すように構成された読取りモジュール

をさらに含む

請求項11から17のいずれか一項に記載のファイル処理装置。

【請求項１９】

コントローラと、独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)とを含む記憶デバイスであって、

前記RAIDがファイルを記憶するように構成され、

前記コントローラが、

プロセッサと、メモリと、通信バスと、通信インターフェースとを含み、前記プロセッサ、前記メモリ、および前記通信インターフェースが、接続されて、前記通信バスを使用することによって互いと通信し、

前記通信インターフェースが、ホストおよび前記RAIDと通信するように構成され、

前記メモリが、コンピュータ実行命令を記憶するように構成され、

前記プロセッサが、前記コンピュータ実行命令を実行して、請求項1から8のいずれか一項に記載の方法を実行するように構成される記憶デバイス。

【請求項２０】

プログラムコードを含むコンピュータプログラムであって、前記プログラムコード内に含まれた命令が、請求項1から8のいずれか一項に記載の方法を実行するために使用される、コンピュータプログラム。

【請求項２１】

コントローラと、独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)とを含む記憶デバイスであって、

前記RAIDがファイルを記憶するように構成され、

前記コントローラが、

プロセッサと、メモリと、通信バスと、通信インターフェースとを含み、前記プロセッサ、前記メモリ、および前記通信インターフェースが、接続されて、前記通信バスを使用することによって互いと通信し、

前記通信インターフェースが、ホストおよび前記RAIDと通信するように構成され、

前記メモリが、コンピュータ実行命令を記憶するように構成され、

前記プロセッサが、前記コンピュータ実行命令を実行して、請求項9から10のいずれか一項に記載の方法を実行するように構成される記憶デバイス。

【請求項２２】

プログラムコードを含むコンピュータプログラムであって、前記プログラムコード内に

含まれた命令が、請求項9から10のいずれか一項に記載の方法を実行するために使用される、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、通信技術の分野に関し、詳細には、ファイル処理方法およびファイル処理装置、ならびに記憶装置に関する。

【背景技術】

【０００２】

安価なディスクの冗長アレイ(Redundant Array of Inexpensive Disks、RAID)としても知られている独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)は、略して、ディスクアレイと呼ばれる。RAIDの原理は、性能が、巨大容量を有する高価なハードディスクの性能に達するか、またはそれをさらに超え、加えて、分散型データ配列の設計と組み合わせてデータセキュリティが改善されるように、複数の比較的安価なディスクを組み合わせることによってディスクアレイグループを形成することである。選択された種々のバージョンに従って、RAIDは、単一のディスクと比較して、記憶容量を増大させるだけでなく、データ統合レベルおよびデータ障害耐性能力(data fault tolerance capability)をも高めることが可能である。加えて、コンピュータにとって、ディスクアレイは、独立ディスクまたは論理記憶ユニットのようにある。

【０００３】

アーカイブシナリオでは、多くのファイルをアーカイブする必要がある。したがって、先行技術では、RAIDは、一般に、アーカイブされたファイルを記憶するために使用され、データセキュリティを改善する目的で、アーカイブされたファイルを記憶するために、たとえば、RAID3、RAID4、RAID5、またはRAID6の形式の、チェック機能を備えたRAIDが一般に使用される。先行技術では、データアクセス速度を改善するために、ファイルは、一般に、いくつかのデータブロックに分割されて、1個のファイルに属する複数のデータブロックと、チェックブロックとがRAIDのストライプ(stripe)を形成して、そのストライプはRAIDを形成する複数のディスク内に書き込まれる。アーカイブされたファイルはそれほど頻繁にアクセスされないため、エネルギー節約目的を達成するために、ファイルがアーカイブされた後、記憶システム内のディスクは、一般に、休止状態であるか、または電源切断状態である。アーカイブされたファイルにアクセスする必要があるときだけ、ファイルを読み取るために、ファイルのデータブロックが記憶された複数のディスクが起動されるか、またはそれらのディスクに電源投入される。

【発明の概要】

【課題を解決するための手段】

【０００４】

本発明の実施形態は、ファイル処理方法およびファイル処理装置、ならびに、ファイル記憶装置のセキュリティを確実にしながら、1個のファイルをRAIDの1個の記憶装置内に記憶することができ、かつエネルギー節約効果を達成することができる記憶デバイスを提供する。

【０００５】

第1の態様によれば、本発明の一実施形態は、

独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)内に記憶されることになるF個のファイルを受信するステップであって、RAIDがT個の記憶装置によって形成され、Fが2以上の自然数であり、Tが3以上の自然数である、受信するステップと、

RAIDのストリップサイズに従って、F個のファイルを少なくとも2個のデータブロックに分割するステップと、

少なくとも2個のデータブロックに従って、T個の行を有する第1の行列を取得するステップであって、1個のファイルに属するデータブロックが第1の行列内の1個の行内に配置される、取得するステップと、

第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むステップと
を含むファイル処理方法を提供する。

【０００６】

第1の態様の第1の可能な実装様式では、少なくとも2個のデータブロックに従って、T個の行を有する第1の行列を取得するステップは、

分割することによって取得された、少なくとも2個のデータブロックをD個の行を有する第2の行列に配列するステップであって、1個のファイルに属するデータブロックが第2の行列の1個の行内に配置され、DがRAID内のデータ記憶装置の数量である、配列するステップと、

チェックブロックを第2の行列の各列にそれぞれ挿入することによって、T個の行を有する第1の行列を取得するステップであって、挿入されたチェックブロックが、第1の行列内のチェックブロックが配置された列内のデータブロックに従って計算することによって取得される、取得するステップと
を含む。

【０００７】

第1の態様の第1の可能な実装様式を参照しながら、第2の可能な実装様式では、RAIDが独立チェック記憶装置を含むとき、チェックブロックを第2の行列の各列にそれぞれ挿入することによって、T個の行を有する第1の行列を取得するステップは、

RAID内の独立チェック記憶装置の位置に従って、チェックブロックを第2の行列に挿入するための位置を判断するステップと、

RAIDのチェックアルゴリズムに従って、第2の行列の各列内のデータブロックに関するチェック計算を実行して、各列内のデータブロックのチェックブロックを取得するステップと、

第2の行列の各列内のデータブロックに従って計算することによって取得されたチェックブロックの判断された位置に従って、チェックブロックをその列に挿入することによって、T個の行を有する第1の行列を取得するステップと
を含む。

【０００８】

第1の態様の第1の可能な実装様式を参照しながら、第3の可能な実装様式では、RAIDが独立チェック記憶装置を含まないとき、チェックブロックを第2の行列の各列にそれぞれ挿入することによって、T個の行を有する第1の行列を取得するステップは、

チェックブロックを第2の行列の各列に挿入するための位置A[x,y]を判断するステップであって、第2の行列がN個の列を有し、xとyが両方とも整数であり、xの値が0からD-1に徐々に増大し、yの値が0からN-1に徐々に増大する、判断するステップと、

第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを第x番目の行内の第(y+1)番目の列から第N番目の列までの位置に順次に移動させるステップと、

RAIDのチェックアルゴリズムに従って、第y番目の列内のデータブロックに関するチェック計算を実行して、第y番目の列内のデータブロックのチェックブロックを取得するステップと、

第y番目の列内のデータブロックのチェックブロックを第2の行列の第y番目の列内の位置A[x,y]に挿入することによって、T個の行を有する第1の行列を取得するステップと
を含む。

【０００９】

第1の態様、または第1の態様の第1から第3の可能な実装様式のうちのいずれか1つを参照しながら、第4の可能な実装様式では、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むステップは、

第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されているとき、第y番目の列内のデータブロックと、チェックブロックとを、RAIDを形成するT個の記憶装置内に書き込むステップであって、第y番目の列が第1の行列内の列のうちの1つである、書き込むステップ
を含む。

【００１０】

第1の態様、または第1の態様の第1から第3の可能な実装様式のうちのいずれか1つを参照しながら、第5の可能な実装様式では、第1の行列はM個の列を有し、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むステップは、

第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されていないとき、第y番目の列内の欠けているデータブロックの数量を判断するステップであって、第y番目の列が第1の行列内の列のうちの1つである、判断するステップと、

第y番目の列内のデータブロックを有さない位置を第1の行列内の第(M+1)番目の列から第(y+1)番目の列までから選択された数量のデータブロックで充填するステップと、

充填した後の第y番目の列内のデータブロックに従って、第y番目の列内のチェックブロックを更新するステップと、

第y番目の列内のデータブロックと、第y番目の列内の更新されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むステップと
を含む。

【００１１】

第1の態様、または第1の態様の第1から第3の可能な実装様式のうちのいずれか1つを参照しながら、第6の可能な実装様式では、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むステップは、

第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されていないとき、第y番目の列内のデータブロックを有さない位置を0で充填するステップと、

0で充填した後の第y番目の列内のデータブロックと、チェックブロックとからなるストライプを、RAIDを形成する記憶装置内に書き込むステップであって、第y番目の列が第1の行列内の列のうちの1つである、書き込むステップ
を含む。

【００１２】

第1の態様、または第1の態様の第1から第6の可能な実装様式のうちのいずれか1つを参照しながら、第7番目の可能な実装様式では、この方法は、

ホストのアクセス要求を受信するステップであって、アクセス要求がRAID内に記憶されたファイルを読み取るために使用され、アクセス要求がファイルに関する論理アドレスを搬送する、受信するステップと、

論理アドレスに従って、ファイルのデータブロックが記憶された物理アドレスに問い合わせるステップと、

物理アドレスに従って、ファイルが記憶された1個の記憶装置を判断するステップと

記憶装置内に記憶されたファイルのデータブロックをホストに返すステップと
をさらに含む。

【００１３】

第2の態様によれば、本発明の一実施形態は、

独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)内に記憶されることになるF個のファイルを受信するステップと、
RAIDのストリップサイズに従って、F個のファイルを少なくとも2個のデータブロックに分割するステップと、
分割することによって取得された、少なくとも2個のデータブロックをアレイに配列するステップであって、アレイ内に、1個のファイルに属する2個の隣接するデータブロック同士の間にD-1個の位置の間隔が存在し、Dの値がRAID内のデータ記憶装置の数量である、配列するステップと、
アレイのD個のデータブロックと、D個のデータブロックに従って計算することによって取得されたP個のチェックブロックとからなるストライプを、RAIDを形成する記憶装置内に書き込むステップであって、Pの値がRAID内の独立チェック記憶装置の数量である、書き込むステップと
を含むファイル処理方法を提供する。
【００１４】
第2の態様の第1の可能な実装様式では、この方法は、
ホストのアクセス要求を受信するステップであって、アクセス要求がRAID内に記憶されたファイルを読み取るために使用され、アクセス要求がファイルに関する論理アドレスを搬送する、受信するステップと、
論理アドレスに従って、ファイルのデータブロックが記憶された物理アドレスに問い合わせるステップと、
物理アドレスに従って、ファイルが記憶された1個の記憶装置を判断するステップと
記憶装置内に記憶されたファイルのデータブロックをホストに返すステップと
をさらに含む。
【００１５】
第3の態様によれば、本発明の一実施形態は、
独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)内に記憶されることになるF個のファイルを受信するように構成された受信モジュールであって、RAIDがT個の記憶装置によって形成され、Fが2以上の自然数であり、Tが3以上の自然数である、受信モジュールと、
RAIDのストリップサイズに従って、F個のファイルを少なくとも2個のデータブロックに分割するように構成された分割モジュールと、
少なくとも2個のデータブロックに従って、T個の行を有する第1の行列を取得するように構成された処理モジュールであって、1個のファイルに属するデータブロックが第1の行列内の1個の行内に配置される、処理モジュールと、
第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むように構成された書込みモジュールと
を含むファイル処理装置を提供する。
【００１６】
第3の態様の第1の可能な実装様式では、処理モジュールは、具体的には、
分割することによって取得された、少なくとも2個のデータブロックをD個の行を有する第2の行列に配列することであって、1個のファイルに属するデータブロックが第2の行列の1個の行内に配置され、DがRAID内のデータ記憶装置の数量である、配列することと、
チェックブロックを第2の行列の各列にそれぞれ挿入することによって、T個の行を有する第1の行列を取得することであって、挿入されたチェックブロックが、第1の行列内のチェックブロックが配置された列内のデータブロックに従って計算することによって取得される、取得することと
を行うように構成される。
【００１７】
第3の態様の第1の可能な実装様式を参照しながら、第2の可能な実装様式では、RAIDが

独立チェック記憶装置を含むとき、処理モジュールは、具体的には、
RAID内の独立チェック記憶装置の位置に従って、チェックブロックを第2の行列に挿入するための位置を判断することと、
RAIDのチェックアルゴリズムに従って、第2の行列の各列内のデータブロックに関するチェック計算を実行して、各列内のデータブロックのチェックブロックを取得することと、
第2の行列の各列のデータブロックに従って計算することによって取得されたチェックブロックの判断された位置に従って、チェックブロックをその列に挿入することによって、T個の行を有する第1の行列を取得することと
を行うように構成される。

【００１８】

第3の態様の第1の可能な実装様式を参照しながら、第3の可能な実装様式では、RAIDが独立チェック記憶装置を含まないとき、処理モジュールは、具体的には、
チェックブロックを第2の行列の各列に挿入するための位置A[x,y]を判断することであって、第2の行列がN個の列を有し、xとyが両方とも整数であり、xの値が0からD-1に徐々に増大し、yの値が0からN-1に徐々に増大する、判断することと、
第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを第x番目の行内の第(y+1)番目の列から第N番目の列までの位置に順次に移動させることと、
RAIDのチェックアルゴリズムに従って、第y番目の列内のデータブロックに関するチェック計算を実行して、第y番目の列内のデータブロックのチェックブロックを取得することと、
第y番目の列内のデータブロックのチェックブロックを第2の行列の第y番目の列内の位置A[x,y]内に挿入することによって、T個の行を有する第1の行列を取得することと
を行うように構成される。

【００１９】

第3の態様、または第3の態様の第1から第3の可能な実装様式のうちのいずれか1つを参照しながら、第4の可能な実装様式では、書込みモジュールは、具体的には、
第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されているとき、第y番目の列内のデータブロックと、チェックブロックとを、RAIDを形成するT個の記憶装置内に書き込むことであって、第y番目の列が第1の行列内の列のうちの1つである、書き込むこと
を行うように構成される。

【００２０】

第3の態様、または第3の態様の第1から第3の可能な実装様式のうちのいずれか1つを参照しながら、第5の可能な実装様式では、第1の行列はM個の列を有し、書込みモジュールは、具体的には、
第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されていないとき、第y番目の列内の欠けているデータブロックの数量を判断することであって、第y番目の列が第1の行列内の列のうちの1つである、判断することと、
第y番目の列内のデータブロックを有さない位置を第1の行列内の第(M+1)番目の列から第(y+1)番目の列までから選択された数量のデータブロックで充填することと、
充填した後の第y番目の列内のデータブロックに従って、第y番目の列内のチェックブロックを更新することと、
第y番目の列内のデータブロックと、第y番目の列内の更新されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むことと
を行うように構成される。

【００２１】

第3の態様、または第3の態様の第1から第3の可能な実装様式のうちのいずれか1つを参照しながら、第6の可能な実装様式では、書込みモジュールは、具体的には、
第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されていないとき、第y番目の列内のデータブロックを有さない位置を0で充填することと、
0で充填した後の第y番目の列内のデータブロックと、チェックブロックとからなるストライプを、RAIDを形成する記憶装置内に書き込むことであって、第y番目の列が第1の行列内の列のうちの1つである、書き込むことと
を行うように構成される。

【００２２】

第3の態様、または第3の態様の第1から第6の可能な実装様式のうちのいずれか1つを参照しながら、第7の可能な実装様式では、受信モジュールは、ホストのアクセス要求を受信することであって、アクセス要求が、RAID内に記憶されたファイルを読み取るために使用され、アクセス要求がアクセスされることになるファイルに関する論理アドレスを搬送する、受信することを行うようにさらに構成され、
ファイル処理装置は、
論理アドレスに従って、ファイルのデータブロックが記憶された物理アドレスに問い合わせて、物理アドレスに従って、ファイルが記憶された1個の記憶装置を判断して、RAIDの1個の記憶装置内に記憶されたファイルのデータブロックをホストに返すように構成された読取りモジュール
をさらに含む。

【００２３】

第4の態様によれば、本発明の一実施形態は、コントローラと、独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)とを含む記憶デバイスであって、
RAIDがファイルを記憶するように構成され、
コントローラが、
プロセッサと、メモリと、通信バスと、通信インターフェースとを含み、プロセッサ、メモリ、および通信インターフェースが、接続されて、通信バスを使用することによって互いと通信し、
通信インターフェースが、ホストおよび独立ディスクの冗長アレイ(Redundant Array of Independent Disks、RAID)と通信するように構成され、
メモリが、コンピュータ実行命令を記憶するように構成され、
プロセッサが、コンピュータ実行命令を実行して、第1の態様または第2の態様によるファイル処理方法を実行するように構成される
記憶デバイスを提供する。

【００２４】

第5の態様によれば、本発明の一実施形態は、プログラムコードを記憶したコンピュータ可読記憶媒体を含むコンピュータプログラム製品であって、プログラムコード内に含まれた命令が、第1の態様または第2の態様によるファイル処理方法を実行するために使用されるコンピュータプログラム製品を提供する。

【００２５】

本発明の実施形態で提供されるファイル処理方法では、記憶デバイスは、受信されたF個のファイルを複数のデータブロックに分割して、それらの複数のデータブロックに従って、T個の列を有する第1の行列を取得する。1個のファイルに属するデータブロックは、第1の行列の1個の列内に配置される。記憶デバイスは、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとを使用することによって、ストライプを形成して、同じファイルに属するデータブロックがRAIDの1枚のディスク内に記憶され得るように、そのストライプをRAID内に記憶する。加えて、ファイルが損傷したとき、記憶デバイスは、他のファイルと、チェッ

クブロックとに従って、損傷したファイルを復元し、それによって、ファイル記憶のセキュリティを確実にすることができる。さらに、アーカイブシナリオでは、RAID内のファイルにアクセスする必要があるとき、記憶デバイスは、ファイルが記憶された1個の記憶装置を起動させて、操作するだけでよく、これは、明らかなエネルギー節約効果をもたらす。

【００２６】

本発明の実施形態または先行技術の技術的解決策をより明瞭に説明するために、以下では、これらの実施形態を説明するために必要とされる添付の図面を手短に紹介する。明らかに、以下の説明において添付の図面は、単に本発明のいくつかの実施形態を示す。

【図面の簡単な説明】

【００２７】

【図１－Ａ】本発明の一実施形態によるファイル処理方法のアプリケーションシナリオの図である。

【図１－Ｂ】本発明の一実施形態による記憶デバイス110の概略構造図である。

【図２－Ａ】本発明の一実施形態によるファイル処理方法の流れ図である。

【図２－Ｂ】本発明の一実施形態による別のファイル処理方法の流れ図である。

【図３】本発明の一実施形態によるファイル処理方法における、チェックブロックを挿入するための方法の流れ図である。

【図４－Ａ】本発明の一実施形態による、記憶されることになるファイルのデータブロックの配列の概略図である。

【図４－Ｂ】本発明の一実施形態による、記憶されることになるファイルのデータブロックの配列の概略図である。

【図４－Ｃ】本発明の一実施形態による、記憶されることになるファイルのデータブロックの配列の概略図である。

【図４－Ｄ】本発明の一実施形態によるファイル記憶構造の概略図である。

【図５－Ａ】本発明の一実施形態によるファイル処理方法における、チェックブロックを挿入するための別の方法の流れ図である。

【図６－Ａ】本発明の一実施形態による、記憶されることになるファイルのデータブロックの別の配列の概略図である。

【図６－Ｂ】本発明の一実施形態による、記憶されることになるファイルのデータブロックの別の配列の概略図である。

【図６－Ｃ】本発明の一実施形態による別のファイル記憶構造の概略図である。

【図６－Ｄ】本発明の一実施形態による、記憶されることになるファイルのデータブロックの別の配列の概略図である。

【図７】本発明の一実施形態によるファイル処理方法における、RAIDを形成するディスク内にデータを書き込むための方法の流れ図である。

【図８】本発明の一実施形態によるさらに別のファイル処理方法の流れ図である。

【図９】本発明の一実施形態による、記憶されることになるファイルのデータブロックのさらに別の構成の概略図である。

【図１０】本発明の一実施形態によるファイル読取り方法の概略流れ図である。

【図１１】本発明の一実施形態によるファイル処理装置の概略構造図である。

【発明を実施するための形態】

【００２８】

当業者に本発明の技術的解決策をより良好に理解させるために、以下では、本発明の実施形態の添付の図面を参照して、本発明の実施形態の技術的解決策を明瞭かつ完全に説明する。明らかに、説明される実施形態は、本発明の実施形態のすべてではなく、その一部にすぎない。

【００２９】

図1-Aに示すように、図1-Aは、本発明の一実施形態によるアプリケーションシナリオの図である。図1-Aに示すアプリケーションシナリオでは、記憶システムは、ホスト100と、

接続デバイス105と、記憶デバイス110とを含む。

【００３０】

ホスト100は、先行技術で知られている任意のコンピューティングデバイス、たとえば、アプリケーションサーバまたはデスクトップコンピュータを含み得る。オペレーティングシステムおよび他のアプリケーションプログラムがホスト100内に設置され、複数のホスト100が存在し得る。

【００３１】

接続デバイス105は、ファイバースイッチまたは別の既存のスイッチなど、先行技術で知られている、記憶デバイスとホストとの間の任意のインターフェースを含み得る。

【００３２】

記憶デバイス110は、先行技術で知られている記憶デバイス、たとえば、記憶アレイ、単なるディスクの束(Just a Bunch Of Disks、JBOD)、または直接アクセス記憶装置(Direct Access Storage Device、DASD)の1つもしくは複数の相互接続されたディスクドライブを含むことが可能であり、この場合、直接アクセス記憶装置は、1つもしくは複数の記憶ユニットのテープライブラリ、またはテープ記憶デバイスを含み得る。

【００３３】

図1-Bは、本発明の一実施形態による記憶デバイス110の概略構造図であり、図-Bに示す記憶デバイスは記憶アレイである。図1-Bに示すように、記憶デバイス110は、コントローラ115と、ディスクアレイ125とを含むことが可能であり、この場合、ディスクアレイは、本明細書では、独立ディスクの冗長アレイ(Redundant Arrays of Independent Disks、RAID)を指す。複数のディスクアレイ125が存在する場合があり、ディスクアレイ125は、複数のディスク130によって形成される。

【００３４】

コントローラ115は、記憶デバイス110の「中枢部」であり、主に、プロセッサ(processor)118と、キャッシュ(cache)120と、メモリ(memory)122と、通信バス(略して、バス)126と、通信インターフェース(Communication Interface)128とを含む。プロセッサ118、キャッシュ120、メモリ122、および通信インターフェース128は、通信バス126を使用することによって互いと通信する。

【００３５】

通信インターフェース128は、ホスト100およびディスクアレイ125と通信するように構成される。

【００３６】

メモリ122は、プログラム124を記憶するように構成され、メモリ122は、高速RAMメモリを含むことが可能であるか、または不揮発性メモリ(non-volatile memory)、たとえば、少なくとも1つのディスクメモリを含み得る。メモリ122は、ランダムアクセスメモリ(Random-Access Memory、RAM)、磁気ディスク、ハードディスク、USBフラッシュドライブ、リムーバブルハードディスク、光ディスク、固体ディスク(Solid State Disk、SSD)、または不揮発性メモリなど、プログラムコードを記憶することが可能な任意の非一時的(non-transitory)機械可読媒体であってよいことを理解されよう。

【００３７】

プログラム124は、プログラムコードを含むことが可能であり、この場合、プログラムコードはコンピュータ動作命令を含む。

【００３８】

キャッシュ120(Cache)は、アレイの性能および信頼性を改善するために、ホスト100から受信されたデータをバッファリングして、ディスクアレイ125から読み取られたデータをバッファリングするように構成される。キャッシュ120は、RAM、ROM、フラッシュメモリ(Flash memory)、または固体ディスク(Solid State Disk、SSD)など、データを記憶することが可能な任意の非一時的(non-transitory)機械可読媒体であってよく、これは、本明細書で限定されない。

【００３９】

プロセッサ118は、中央処理装置CPU、もしくは特定用途向け集積回路ASIC(Application Specific Integrated Circuit)であってよく、または本発明の本実施形態を実装する、1つもしくは複数の集積回路として構成されてもよい。オペレーティングシステム、および他のソフトウェアプログラムがプロセッサ118内に設置され、異なるソフトウェアプログラムを、異なる機能、たとえば、ディスク130に関する入出力(Input/output、I/O)要求の処理、ディスク内のデータに関する他の処理の実行、または記憶デバイス内に記憶されたメタデータの修正を備えた処理モジュールと見なすことが可能である。したがって、コントローラ115は、I/O動作およびRAID管理機能を実装することが可能であり、スナップショット、ミラーリング、および複写など、様々なデータ管理機能を提供することも可能である。本発明の本実施形態では、プロセッサ118は、プログラム124を実行するように構成され、具体的には、以下の方法実施形態の関連ステップを実行することが可能である。

【００４０】

図1-Aを参照すると、任意の記憶デバイス110は、接続デバイス105を使用することによって、1つまたは複数のホスト100によって送信された複数のファイルを受信して、受信した複数のファイルを複数のデータブロックに分割して、それらのデータブロックを、ディスクアレイ125を形成する複数のディスク130内に記憶することができる。任意の記憶デバイス110は、任意のホスト100によって送信されたファイル読取り要求を受信して、そのファイル読取り要求に従って、ディスク130内に記憶されたファイルのデータブロックをホストに返すことも可能である。

【００４１】

ディスク130は、ディスクアレイ125を形成する記憶装置の単なる例であり、実際のアプリケーションでは、複数のディスクを含むキャビネット同士の間でディスクアレイが形成される実装様式が存在し得ることに留意されたい。したがって、本発明の本実施形態の記憶装置は、磁気ディスク、固体ディスク(Solid State Disk、SSD)、または複数の磁気ディスクによって形成されたキャビネットもしくはサーバなど、任意の装置を含むことが可能であり、これは本明細書で限定されない。

【００４２】

図2-Aは、本発明の一実施形態によるファイル処理方法の流れ図である。この方法は、図1-Bに示した記憶デバイス110のコントローラ115によって実行されることが可能であり、この方法は、ファイルアーカイブシナリオに適用され得る。図2-Aに示すように、この方法は、以下を含む。

【００４３】

ステップ200で、記憶デバイス110は、RAID内に記憶されることになるF個のファイルを受信し、Fは2以上の自然数である。本発明の本実施形態では、記憶デバイス110のコントローラ115は、1つまたは複数のホスト100によって送信されたファイル記憶要求を受信することが可能であり、ファイル記憶要求は、ファイルを記憶デバイス110の第1のRAID内に記憶することを要求するために使用され、第1の記憶要求は、F個の記憶されることになるファイルを含み得る。第1のRAIDはT個の記憶装置を含み、Tの値は3以上の自然数である。

【００４４】

図1-Bを参照すると、記憶デバイス110は複数のRAIDを含み得る。本実施形態で説明される第1のRAIDまたは第2のRAIDは、記憶デバイス110内に含まれる複数のRAIDのうちの任意の1つである。本発明の本実施形態の第1のRAIDおよび第2のRAIDは、単に、異なるRAIDを区別することを目的とする。同じ記憶デバイス110内に含まれた複数のRAIDの構成形式は同じであってよく、たとえば、第1のRAIDと第2のRAIDは両方ともRAID5の構成形式である。当然、同じ記憶デバイス110内に含まれた複数のRAIDの構成形式は異なってよく、たとえば、第1のRAIDはRAID3であり、第2のRAIDはRAID5であり、これは本明細書で限定されない。実際の動作では、受信されたF個のファイルは、まず、キャッシュ120内でバッファリングされることが可能であり、処理された後、F個のファイルはディスクアレイ125内に書き込まれることを理解されよう。

【００４５】

ステップ205で、記憶デバイス110は、第1のRAIDのストリップサイズ(strip size)に従って、F個のファイルを少なくとも2個のデータブロックに分割する。ストリップ(strip)は、ある程度、連続的なアドレスブロックである。ディスクアレイ内で、コントローラは、一般に、ストリップを使用することによって、仮想ディスクのブロックアドレス(block address)をメンバーディスクのブロックアドレスにマッピングする。ストリップはまた、ストライプ要素(stripe element)と呼ばれる。ブロックサイズ、チャンクサイズ、または粒度と呼ばれる場合もあるストリップサイズ(strip size)は、各ディスクに書き込まれたストリップデータブロックのサイズを指す。一般に、RAIDのストリップサイズは、2KBから512KB(または、それ以上)の間であり、ストリップサイズの値は2のn乗、たとえば、2KB、4KB、8KB、16KBなどである。

【００４６】

受信されたファイルが第1のRAIDのストリップサイズに従って分割されるとき、ファイルのサイズが第1のRAIDのストリップサイズ未満である場合、ファイルは1個のデータブロックとして使用され得る。ファイルが分割された後の残りのデータブロックがストリップサイズの値未満である場合、ファイルの残りのデータは1個のデータブロックとして使用される。たとえば、図4-Aに示すように、コントローラ115は、5個の記憶されることになるファイルF1～F5を受信する。第1のRAIDのストリップサイズに従って分割された後、ファイルF1は5個のデータブロックF1-1、F1-2、F1-3、F1-4、およびF1-5に分割される。ファイルF2は、3個のデータブロックF2-1、F2-2、およびF2-3に分割される。ファイルF3は、1個のデータブロックF3-1に分割される。ファイルF4は、5個のデータブロックF4-1、F4-2、F4-3、F4-4、およびF4-5に分割される。ファイルF5は、4個のデータブロックF5-1、F5-2、F5-3、およびF5-4に分割される。

【００４７】

ステップ210で、記憶デバイス110は、少なくとも2個のデータブロックに従って、T個の行を有する第1の行列を取得し、1個のファイルに属するデータブロックは第1の行列の1個の行内に配置される。加えて、第1の行列内の各列は、その列内のデータブロックに従って計算することによって取得されたチェックブロックを含み、Tの値は、第1のRAIDを形成するディスクの数量に等しい。

【００４８】

たとえば、第1のRAIDが計4枚のディスクを有する場合、ファイルF1～F5を分割することによって取得された前述のデータブロックに従って、4個の行を有する第1の行列を取得することが可能であり、1個のファイルに属するデータブロックは、第1の行列内の1個の行内に配置される。図4-Cに示すように、ファイルF1のデータブロックF1-1、F1-2、F1-3、F1-4、およびF1-5はすべて、第1の行列の第0番目の行内に配置され、ファイルF2のデータブロックF2-1、F2-2、およびF2-3はすべて、第2の行列の第1番目の行内に配置される。

【００４９】

具体的には、第1の行列を取得するプロセスで、第1のファイルの第1のデータブロックを第1の行列内の位置A[0,0]に位置するデータブロックと判断することが可能であり、第2のデータブロックが第1のファイルに属するかどうかが判断される。第2のデータブロックが第1のファイルに属する場合、第2のデータブロックは、第1のデータブロックと同じ行に配列され、第2のデータブロックが第1のファイルに属さない場合、第2のブロックは、初めに見出された空の行に配列されるか、または第2のデータブロックは、第2の行列の最短の行に配列され、他のデータブロックは、分割することによって取得されたすべてのデータブロックが配列されるまで、類比によって処理される。当然、1個のファイルに属するデータブロックが第1の行列内の1個の行内に配置されることが保証される限り、分割することによって取得されたデータブロックを第1の行列に配列するために別の方法を使用することが可能であり、これは本明細書で限定されないことを理解されよう。配列後のT個の行を有する第1の行列を図4-C、または図6-Bに示すことができる。

【００５０】

配列後のT個の行を有する第1の行列内の各列内で、その列内のデータブロックに従って

計算することによって取得されたチェックブロックが、たとえば、図4-Cに示す第1の行列内に含まれ、第0番目の列は、第0番目の列内のデータブロックF1-1、F2-1、およびF3-1に従って取得されたチェックブロックP1を含み、第1番目の列は、第1番目の列内のデータブロックF1-2、F2-2、およびF3-2に従って取得されたチェックブロックP2を含む。

【００５１】

本発明の本実施形態では、ファイルが第1の行列内に配置される特定の位置は限定されず、1個のファイルのデータブロックが第1の行列の1個の行内に配置されることが保証される限り、1個のファイルに属するデータブロックの配列順序は限定されない。実際のアプリケーションでは、1個のファイルに属するデータブロックは、第1の行列の1個の行内に順次に配列される。

【００５２】

ステップ215で、記憶デバイス110は、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプ(stripe)を、第1のRAIDを形成するT個の記憶装置内に書き込む。

【００５３】

第1の行列が取得された後、RAIDのストライプは、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなることが可能であり、ストライプを、第1のRAIDを形成するディスク内に書き込むことが可能である。たとえば、ある状況では、第1の行列内の各列内のデータブロックと、図4-Cで示した列内のデータブロックに従って計算することによって取得されたチェックブロックとからそれぞれなるストライプがディスク内に書き込まれた後、図4-Dに示す記憶構造が形成される。F1-1、F2-1、F3-1、およびP1は、第1のRAIDのあるストライプを形成し、F1-2、F2-2、F3-2、およびP2は、第1のRAIDの別のストライプを形成する、等々である。別の状況では、第1の行列内の各列内のデータブロックと、図6-Bに示した列内のデータブロックに従って計算することによって取得されたチェックブロックとからそれぞれなるストライプがディスク内に書き込まれた後、図6-Cに示した記憶構造が形成され得る。本発明の本実施形態で説明されるストライプ(stripe)は、RAIDを形成する記憶装置の各々の中に同時に書き込まれるデータブロックの収集物を指し、ストライプ内の各データブロックのサイズは同じであり、1つのストライプ内のデータブロックは各記憶装置の1つの変位位置に配置されることに留意されたい。

【００５４】

実際のアプリケーションでは、図4-Cまたは図6-Bに示した第1の行列内の列内のデータブロックは、その列内のチェックブロックを取得するために計算され得ることを理解されよう。チェックブロックがその列に挿入された後、その列内のデータブロックと、チェックブロックとからなるストライプが第1のRAIDを形成するディスク内に記憶される。オプションで、計算することによって、第1の行列の各列内のチェックブロックがすべて取得された後、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとを使用することによってストライプをそれぞれ形成することが可能であり、ストライプは第1のRAIDを形成するディスク内に記憶される。たとえば、図4-Cに示した第1の行列内の第0番目の列内のデータブロックF1-1、F2-1、およびF3-1に従って計算することによってチェックブロックP1がまず取得された後、第0番目の列内のデータブロックF1-1、F2-1、およびF3-1と、図4-Cに示したP1とからなるストライプが、第1のRAIDを形成するディスクD1～D4内に記憶される。オプションで、計算することによってチェックブロックP1～P7がすべて取得された後、第1の行列内の各列内のデータブロックと、チェックブロックとからそれぞれなるストライプがディスクD1～D4内に記憶され、これは、本明細書で限定されない。

【００５５】

本発明の本実施形態で説明されるファイル処理方法では、記憶されることになるファイルが分割および配列された後、異なるファイルに属するデータブロックからなるストライプが第1のRAIDを形成するディスク内に記憶され、これはファイル書込み効率を確実にす

るだけでなく、1個のファイルに属するデータブロックが1枚のディスク内に記憶されることをも可能にすることを前述の説明から理解されよう。たとえば、ファイルF1に属するすべてのデータブロックはディスクD1内に記憶され、ファイルF2に属するすべてのデータブロックはディスクD2内に記憶される。本発明の本実施形態の方法を使用することによって、複数のファイルがRAID内に記憶された後、記憶アレイ内のファイルにアクセスする必要があるとき、記憶デバイス110は、RAID内のすべてのディスクを起動する必要がなく、そのファイルが記憶されたディスクを起動させて、そのディスク内のファイルをホストに返すことだけを必要とし得、それによって、より良好なエネルギー節約効果を達成する。加えて、本発明の本実施形態の技術的解決策では、データブロックが損傷した場合、同じストライプのチェックブロックまたは他のファイルのデータブロックを使用することによって、損傷したデータブロックを復元することが可能であり、それによって、ファイル記憶装置のセキュリティを確実にする。

【００５６】

本発明の本実施形態で説明されるデータブロックは、複数のデータによって形成されるデータユニットを指すことに留意されたい。本発明の本実施形態で説明されるチェックブロックは、チェックデータによって形成されるデータユニットを指す。本発明の本実施形態で説明される行列は、データブロックを使用することによって形成された複数のアレイを含むことが可能であるか、またはデータブロックとチェックブロックとを使用することによって形成された複数のアレイを含むことが可能である。本発明の本実施形態で、行は、1個のファイルに属するすべてのデータブロックを含む1個のアレイを指す。本発明の本実施形態で、列は、行に対して直角のアレイを指す。すなわち、本発明の本実施形態で規定される行は、通常の行列内で規定される水平アレイに限定されない。通常の行列内の水平アレイが1個のファイルに属するデータブロックを含むとき、水平アレイ(たとえば、図4-Cに示した水平アレイ)は、本発明の本実施形態で行と呼ばれる場合がある。通常の行列内の垂直アレイが1個のファイルに属するデータブロックを含むとき、垂直アレイは、本発明の本実施形態で行と呼ばれる場合もあり、これは本明細書では限定されない。

【００５７】

図2-Bは、本発明の一実施形態による別のファイル処理方法の流れ図である。この方法は、図1-Bに示した記憶デバイス110のコントローラ115によって実行されることが可能であり、この方法は、ファイルアーカイブシナリオに適用され得る。図2-Bに示すように、この方法は、図2-Aに示した方法と類似し、ステップ207およびステップ209は、図2-Aで示した方法のステップ210の詳細な記述である。図2-Bに示すように、この方法は以下を含む。

【００５８】

ステップ200で、記憶デバイス110は、第1のRAID内に記憶されることになるF個のファイルを受信し、Fは2以上の自然数であり、第1のRAIDはT個の記憶装置を含み、Tの値は3以上の自然数である。

【００５９】

ステップ205で、記憶デバイス110は、第1のRAIDのストリップサイズに従って、F個のファイルを少なくとも2個のデータブロックに分割する。

【００６０】

ステップ207で、記憶デバイス110は、少なくとも2個のデータブロックをD個の行を有する第2の行列に配列し、1個のファイルに属するデータブロックは、第2の行列の1個の行内に配置され、Dは第1のRAID内のデータディスクの数量である。

【００６１】

ファイルを分割することによって複数のデータブロックが取得された後、取得された複数のデータブロックは、D個の行*N個の列の第2の行列に配列されることが可能であり、Dは、第2の行列の行の数量を表すために使用され、Dの値は、第1のRAIDを形成するデータディスクの数量に従って判断され、Nは整数である。第1のRAID内のデータディスクの数量は、第1のRAIDの構成形式に従って判断される必要があることを理解されよう。たとえば

、RAID3は、複数のデータディスクと1枚の独立チェックディスクとを含み、RAID4は、複数のデータディスクと2枚の独立チェックディスクとを含む。一方、RAID5は、複数のデータディスクだけを含み、独立チェックディスクは含まない。Dの値は、第1のRAID内のデータディスクの数量だけに従って判断される必要がある。たとえば、第1のRAIDの構成形式がRAID3であり、第1のRAIDが計4枚のディスクを含む場合、データディスクの数量は3であり、チェックディスクの数量は1である。したがって、第2の行列内の行の数量は3であり、これらは、図4-Bに示した第2の行列に配列され得る。第1のRAIDの構成形式がRAID5であり、第1のRAIDが計4枚のディスクを含む場合、データディスクの数量は4であり、チェックディスクの数量は0である。したがって、第2の行列内の行の数量は4であり、これは、図6-Aに示す第2の行列に配列され得る。

【００６２】

Nは、第2の行列内の列の数量を表すために使用され、Nは整数であり、Nの値は、限定されなくてよく、具体的には、データブロックの数量に従って判断され得る。受信された複数の記憶されることになるファイルがキャッシュ120内でバッファリングされる場合、Nの値は、キャッシュ120のサイズに従って判断されることが可能であり、Nとストリップサイズの積はキャッシュ120の容量以下であることを理解されよう。具体的には、第2の行列を配列するプロセスは、前述の図2-Aのステップ210で説明した、第1の行列を配列するための方法に類似し、本明細書で詳細は繰り返して説明されない。

【００６３】

本発明の本実施形態のデータディスクは、データブロックを記憶するデータ記憶装置の単なる例であり、独立チェックディスクは、チェックデータを記憶するために特に使用される独立チェック記憶装置の単なる例であることに留意されたい。本発明の本実施形態のデータ記憶装置は、データブロックを記憶するために使用される記憶装置を指し、独立チェック記憶装置は、チェックブロックを記憶するために特に使用される記憶装置を指し、記憶装置は、磁気ディスク、または磁気ディスクを含むキャビネットもしくはサーバなどの装置を含むが、これらに限定されない。

【００６４】

本発明の本実施形態では、データブロックが具体的に配列されるとき、1個のファイルに属するデータブロックが第2の行列の1個の行内に配置されることが確実にされなければならない。たとえば、図4-Bに示すように、第1のRAIDの構成形式がRAID3であり、第1のRAIDが3枚のデータディスクを含む場合、ファイルF1からF5を分割した後で取得されるデータブロックを3個の行*7個の列の第2の行列に配列することが可能であり、この場合、ファイルF1のデータブロックF1-1、F1-2、F1-3、F1-4、およびF1-5はすべて、第2の行列の第0番目の行内に配置され、ファイルF2のデータブロックF2-1、F2-2、およびF2-3はすべて、第2の行列の第1番目の行内に配置され、ファイルF5のF5-1、F5-2、F5-3、およびF5-4も第2の行列の第1番目の行内に配置される。

【００６５】

ステップ209で、記憶デバイス110は、チェックブロックを第2の行列の各列にそれぞれ挿入することによって、T個の行を有する第1の行列を取得する。

【００６６】

第2の行列の各列に挿入されたチェックブロックは、第1のRAIDの構成形式によって判断されたチェックアルゴリズムに従って、列内のデータブロックを計算することによって取得され、Tの値とDの値との間の差は、第1のRAID内の独立チェックディスクの数量である。たとえば、Tの行を有する第1の行列は、(D+P)個の行*M個の列の第1の行列であり得、この場合、Pは第1のRAID内のチェックディスクの数量であり、MはN以上の整数であり、Mとストリップサイズの積は、RAID内の単一のディスクの容量以下である。

【００６７】

記憶デバイス110のコントローラ115は、第1のRAIDの構成形式に従って、チェックアルゴリズム(すなわち、チェックブロックの計算方法)を判断して、判断されたチェックアルゴリズムに従って、第2の行列の各列内のデータブロックのチェックブロックを計算して

、各列のデータブロックのチェックブロックを第2の行列に挿入することによって、(D+P)個の行*M個の列の第1の行列を取得することができ、この場合、Pは第1のRAID内の独立チェックディスクの数量であり、Mの値はNの値以上であるべきであり、Mとストリップサイズの積は、RAID内の単一のディスクの容量以下である。受信された複数の記憶されることになるファイルがキャッシュ120内でバッファリングされる場合、Mとストリップサイズの積は、やはりキャッシュ120のサイズの容量以下であることを理解されよう。

【００６８】

実際の動作では、受信された複数のファイルが一時記憶領域(すなわち、キャッシュ120)内でまずバッファリングされる場合、チェックブロックは一時記憶領域の容量を依然として占有する必要があることを考慮して、チェックブロックを第2の行列に挿入する条件を設定すること、たとえば、一時記憶領域内のデータ量が設定されたしきい値を超えるとき、チェックブロックが第2の行列の各列に挿入されるという条件を設定することが可能であることを理解されよう。当然、設定された記憶制限時間に達したとき、チェックブロックが第2の行列の各列に挿入されることを設定することも可能である。記憶制限時間は、ファイルを、第1のRAIDを形成するディスク内に書き込むための事前設定された制限時間である。たとえば、記憶は一時間に一度実行されることを指定することが可能であり、その場合、記憶制限時間は一時間である。記憶制限時間は、ディスク内に書き込む必要があるデータの量など、実際の状況に従って判断され得る。記憶は一日に一度実行されてよく、または記憶は10分ごとに実行されてよく、これは本明細書で限定されない。

【００６９】

ステップ209で、チェックブロックが第2の行列に挿入されるとき、第1のRAIDの構成形式に従って、異なる処理をそれぞれ実行することが可能である。詳細については、図3および図５－Ａの関連記述を参照することが可能である。

【００７０】

一事例では、第1のRAIDの構成形式が独立チェックディスクを有するRAIDであるとき、たとえば、第1のRAIDがRAID3またはRAID4であるとき、コントローラ115は、図3に示した方法のプロセスに従って、チェックブロックを挿入することができる。図3に示すように、この方法は、以下を含む。

【００７１】

ステップ305で、記憶デバイス110は、第1のRAID内の独立チェックディスクの位置に従って、チェックブロックを第2の行列に挿入するための位置を判断する。

【００７２】

たとえば、第1のRAIDがRAID3である場合、第1のRAIDは独立チェックディスクを有する。図4-Dに示すように、第1のRAIDが4枚のディスクを有する場合、ディスクD1、D2、D3、およびD4のうちのいずれか1つを独立チェックディスクとして使用することができる。たとえば、図4-Dに示した第1のRAID内の独立チェックディスクとしてD4が使用される。チェックブロックを第2の行列に挿入するための位置は、独立チェックディスクの判断された位置に従って判断され得る。たとえば、図4-Dに示した独立チェックディスクD4の位置に従って、図4-Bに示した第2の行列の最後の行の後に、チェックブロックの行が追加されることを判断することができる。このように、図4-Bに示した第2の行列は3個の行を有し、その場合、チェックブロックを挿入するために、第4番目の行が第2の行列に追加される。

【００７３】

当然、独立チェックディスクとしてD2が使用される場合、3個の行*7個の列の第2の行列が4個の行*7個の列の第1の行列になるように、図4-Bに示した第2の行列内のデータの第1の行とデータの第2の行との間に行が挿入されて、チェックブロックの位置として使用されることを理解されよう。独立チェックディスクの位置の前述の例は、独立チェックディスクの位置に対する何らかの制限となることは意図されない。

【００７４】

ステップ310で、記憶デバイス110は、第1のRAIDのチェックアルゴリズムに従って、第2の行列の各列内のデータブロックに関するチェック計算を実行して、各列内のデータブロ

ックのチェックブロックを取得する。

【００７５】

たとえば、第1のRAIDのチェックアルゴリズムがパリティチェックアルゴリズムである場合、パリティチェックアルゴリズムに従って、図4-Bに示した第2の行列内の各列内のデータに関するチェック計算をそれぞれ実行して、各列内のデータブロックのチェックブロックを取得することができる。たとえば、チェック計算は、チェックブロックP1を取得するために、図4-Bに示した第0番目の列内のデータブロックF1-1、F2-1、およびF3-1に従って実行される。チェック計算は、チェックブロックP2を取得するために、第1番目の列内のデータブロックF1-2、F2-2、およびF4-1に従って実行される。本実施形態では、パリティチェックアルゴリズムは、単なる例であり、チェックアルゴリズムを限定しないことに留意されたい。ステップ305およびステップ310の順序は限定されない。

【００７６】

ステップ315で、記憶デバイス110は、第2の行列の各列内のデータブロックに従って計算することによって取得されたチェックブロックの判断された位置に従って、チェックブロックをその列に挿入することによって、T個の行を有する第1の行列を取得する。

【００７７】

たとえば、チェックブロックが図4-Bに示した第2の行列に挿入された後、図4-Cに示す、4個の行*7個の列の第1の行列を取得することが可能であり、この場合、P1は、第2の行列内の第0番目の列内のデータブロックF1-1、F2-1、およびF3-1に従って計算することによって取得されたチェックブロックであり、P2は、第2の行列内の第1番目の列内のデータブロックF1-2、F2-2、およびF4-1に従って計算することによって取得されたチェックブロックである、等々である。

【００７８】

独立チェックディスクの場合、第2の行列内の各列に挿入されることになるチェックブロックの数量は、独立チェックディスクの数量に従って判断され得る。したがって、チェックブロックが挿入された後、第2の行列内の行の数量は変化するが、第2の行列内の列の数量は変化せずに残る。言い換えれば、第1のRAIDの構成形式が、独立チェックディスクを有するRAIDである場合、第1の行列内のMの値は、第2の行列内のNの値と等しい。

【００７９】

別の事例では、第1のRAIDの構成形式が独立チェックディスクを有するRAIDではなく、分散型チェックブロックを有するRAIDであるとき、たとえば、第1のRAIDがRAID5またはRAID6であるとき、コントローラ115は、図５－Ａに示した方法のプロセスに従って、チェックブロックを挿入することができる。図５－Ａに示すように、この方法は、以下を含む。

【００８０】

ステップ505で、記憶デバイス110は、チェックブロックを第2の行列の各列に挿入するための位置A[x,y]を判断する。

【００８１】

実際のアプリケーションでは、第1のRAIDの構成形式と、第1のRAID内のチェックブロックの分散様式とに従って、チェックブロックを第2の行列の各列に挿入するための位置A[x,y]を判断することが可能である。RAID5内のディスク上にチェックブロックを分散する様式は、左同期(後方パリティ、すなわち、Left Synchronous)、左非同期(後方動的、すなわち、Left Asynchronous)、右同期(前方パリティ、すなわち、Right Synchronous)、または右非同期(前方動的、すなわち、Right Asynchronous)であってよいことを当業者は理解されよう。「左」または「右」は、チェック情報がどのように分散されるかを示し、「同期」または「非同期」は、データがどのように分散されるかを示す。「左」のアルゴリズムでは、最後のディスクから始めて、チェックブロックを(必要な場合、環状に繰り返し分散されるように)第1のディスクに向かう方向に各ストライプ内で1枚のディスク位置だけ移動させる。「右」のアルゴリズムでは、第1のディスクから始めて、チェックブロックを(必要な場合、環状に繰り返し分散されるように)最後のディスクに向かう方向に各ストライプ内で1枚のディスク位置だけ移動させる。RAID5に基づいて、チェックブロック

の別のグループを有するRAID6が追加される。

【００８２】

独立チェックディスクを有さないRAIDの構成形態では、チェックブロックがディスク内でどのように具体的に分散されるかは、第1のRAIDの構成形式と、チェックブロックの分散様式とに従って判断され得る。たとえば、第1のRAIDの構成形式がRAID5であり、チェックブロックの分散様式が左同期である場合、チェックブロックは、最後のディスクから始めて、第1のディスクの位置に向かう方向に各ストライプ内の1枚のディスク位置だけ移動されるようにディスク内で分散されることを理解されよう。

【００８３】

本発明の本実施形態では、チェックブロックを第2の行列に挿入するための位置A[x,y]は、第1のRAID内のチェックブロックの分散様式に従って判断されることが可能であり、この場合、xは、0以上、(D-1)以下の整数であり、yは、0以上、(N-1)以下の整数であり、すなわち、0≦x≦(D-1),0≦y≦(N-1)である。加えて、xおよびyの値は、第2の行列内のチェックブロックの様々な位置とともに変化し、xの値は0から(D-1)に徐々に増大し、yの値は0から(N-1)に徐々に増大する。たとえば、図6-Aに示すように、第1のRAIDの構成形式がRAID5であり、チェックブロックの分散様式が左同期である場合、チェックブロックを第2の行列の第0番目の列に挿入するための位置はA[3,0]であり、チェックブロックを第1番目の列に挿入するための位置はA[2,1]であり、チェックブロックを第2番目の列に挿入するための位置はA[1,2]であり、チェックブロックを第3番目の列に挿入するために位置はA[0,3]である。次の循環は第4番目の列から始まり、すなわち、チェックブロックを第4番目の列に挿入するための位置はA[3,4]であり、チェックブロックを第5番目の列に挿入するための位置はA[2,5]である、等々である。具体的には、位置は、図6-Bに示すP1～P7であり得る。

【００８４】

ステップ510で、記憶デバイス110は、第2行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータを第x番目の行内の第(y+1)番目の列から第N番目の列までの位置に順次に移動させる。

【００８５】

チェックブロックを第2の行列の各列に挿入するための位置A[x,y]が判断された後、第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを第x番目の行の第(y+1)番目の列から第N番目の列までの位置に順次に移動する必要があり、すなわち、元の位置A[x,y]からA[x,N-1]までの中のすべてのデータブロックを1つの位置だけ右に向かって移動させて、位置A[x,y+1]からA[x,N]まで順次に移動させる必要がある。たとえば、図6-Aに示した、チェックブロックが第2の行列の第0番目の列に挿入するための位置がA[3,0]になると判断されたとき、第2の行列の第3番目の行内の位置A[3,0]からA[3,4]までの中のすべてのデータブロックを1つの位置だけ後方に移動させて、A[3,1]からA[3,5]の位置に順次に移動させる必要がある。このようにして、元の位置A[3,0]内のデータブロックF4-1をA[3,1]に移動させることができ、元の位置A[3,1]内のデータブロックF4-2をA[3,2]に移動させることができる、等々である。xおよびyの値は、第2の行列内のチェックブロックの様々な位置とともに変化し、チェックブロックの位置A[x,y]が判断されるたびに、第x番目の行内の元の位置A[x,y]からA[x,N-1]までの中のすべてのデータブロックを1つの位置だけ後方に移動させる必要がある。本発明の実施形態では、各列内のデータブロックの数量は限定されない。

【００８６】

ステップ515で、記憶デバイス110は、第1のRAIDのチェックアルゴリズムに従って、第y番目の列内のデータブロックに関するチェック計算を実行して、第y番目の列内のデータブロックのチェックブロックを取得する。

【００８７】

チェックブロックの位置がA[x,y]であると判断されて、第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを第x番目の行の第(y+1)番目の列か

ら第N番目の列までの位置に順次に移動させた後、第1のRAIDのチェックアルゴリズムに従って、第y番目の列内のデータブロックに関するチェック計算を実行して、第y番目の列内のデータブロックのチェックブロックを取得することが可能である。チェックブロックは、A[x,y]の位置に挿入される必要があるチェックブロックである。たとえば、図6-Bに示すように、第0番目の列内のチェックブロックの位置がA[3,0]であると判断されて、第2の行列内の元の位置A[3,0]内のデータブロックF4-1を位置A[3,1]に移動させるとき、第0番目の列内のチェックブロックP1は、第0番目の列内の新しいデータブロックF1-1、F2-1、およびF3-1に従って計算することによって取得され得る。

【００８８】

ステップ520で、記憶デバイス110は、第y番目の列内のデータブロックのチェックブロックを第2の行列の第y番目の列内の位置A[x,y]に挿入することによって、T個の行を有する第1の行列を取得する。

【００８９】

計算することによってチェックブロックが取得された後、T個の行を有する第1の行列を取得することができるように、チェックブロックは、判断されたチェックブロックの位置A[x,y]に挿入され得る。たとえば、(D+P)個の行*M個の列の第1の行列を取得することが可能である。独立チェックディスクを有さないRAIDの構成形式では、チェックブロックは、第2の行列の各列に挿入される必要があり、そのチェックブロックが挿入された位置内の元のデータブロックを順次に後方に移動させる必要があり、したがって、取得された第1の行列のMの値は、第2の行列のNの値よりも大きい。たとえば、第1のRAIDの構成形式がRAID5であり、チェックブロックが左同期様式で分散される場合、チェックブロックが、図6-Aに示した4個の行*5個の列の第2の行列の各列に挿入された後、図6-Bに示した4個の行*7個の列の第1の行列を取得することが可能である。

【００９０】

ステップ215で、記憶デバイス110は、第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、第1のRAIDを形成するディスク内に書き込む。実際のアプリケーションでは、ステップ215で、第1の行列の各列内のデータブロックが、第1のRAIDを形成するディスク内にストライプの形で書き込まれるとき、以下の状況が発生する場合があり、それらの状況がそれぞれ処理され得る。

【００９１】

一事例では、第1の行列の第y番目の列内のデータブロックと、チェックブロックとからなるストライプが完全に占有されているとき、第y番目の列内のデータブロックおよびチェックブロックを、第1のRAIDを形成するディスク内に直接書き込むことが可能であり、この場合、第y番目の列は第1の行列内のM個の列のうちの1つである。たとえば、図6-Bに示した第0番目の列が完全に占有されているとき、すなわち、第0番目の列内のデータブロックと、チェックブロックとからなるストライプが完全に占有されているとき、第0番目の列内のデータブロックと、チェックブロックとからなるストライプがディスク内に書き込まれる。

【００９２】

別の事例では、第1の行列の第y番目の列内のデータブロックと、チェックブロックとからなるストライプが完全に占有されていない場合、第y番目の列内のデータブロックを有さない位置内に0が充填され得る。0で充填した後の第y番目の列内のデータブロックと、チェックブロックとからなるストライプは第1のRAIDを形成するディスク内に書き込まれ、この場合、第y番目の列は第1の行列内の列である。たとえば、図6-Bに示した第1の行列内の第4番目の列が十分に占有されていないとき、すなわち、位置A[1,4]内に何のデータも書き込まれていないとき、位置A[1,4]内に0が充填され得る。次いで、第4番目の列内のデータブロックと、チェックブロックとからなるストライプがディスク内に書き込まれ、すなわち、データブロックF1-4およびF5-3と、チェックブロックP5とからなるストライプがディスク内に書き込まれる。

【００９３】
さらに別の事例では、第1の行列の第y番目の列内のデータブロックと、チェックブロックとからなるストライプが十分に占有されていないとき、かつ記憶制限時間に達し、他のファイルが受信されない場合、処理のために図7に示す方法を使用することが可能である。図7に示すように、この方法は、以下を含む。
【００９４】
ステップ700で、記憶デバイス110は、第y番目の列内の欠けているデータブロックの数量を判断する。
【００９５】
たとえば、図4-Cに示した第1の行列内の第5番目の列内のデータブロックと、チェックブロックとからなるストライプは十分に占有されていない。たとえば、すなわち、第5番目の列内に、何のデータも書き込まれていない位置A[0,5]が存在し、別の例では、図6-Bに示した第1の行列の第y(y=4)番目の列内に、何のデータも書き込まれていない位置A[1,4]がやはり存在する。その時点で記憶制限時間に達する場合、図4-Cに示した第1の行列の第5番目の列内の欠けているデータブロックの数量は1であり、図6-Bに示した第1の行列の第4番目の列内の欠けているデータブロックの数量も1であると判断することができる。
【００９６】
ステップ705で、記憶デバイス110は、第1の行列内の第(M-1)番目の列から第(y+1)番目の列までの数量のデータブロックを選択して、第y番目の列内のデータブロックを有さない位置をそれらのデータブロックで充填する。
【００９７】
第y番目の列内に、データブロックを有さない位置が存在し、記憶制限時間に達し、記憶デバイス110が記憶されることになる何の他のファイルもホストから受信しない場合、ディスクの記憶空間を節約するために、記憶デバイス110は、第1の行列内の第(M-1)番目の列から第(y+1)番目の列までの対応する数量のデータブロックを順次に選択して、第y番目の列内のデータブロックを有さない位置をそれらのデータブロックで充填する。言い換えれば、第1の行列内の第y番目の列内に、データブロックを有さない位置が存在すると判断されたとき、記憶デバイス110は、第1の行列の最後の列から始めて、最後の列から第0番目の列への方向に従って、対応する数量のデータブロックを選択して、第y番目の列内のデータブロックを有さない位置をそれらのデータブロックで充填することができる。
【００９８】
たとえば、記憶デバイス110は、図4-Cに示した第1の行列の第6番目の列からデータブロックF5-4を選択して、第5番目の列内のデータブロックが欠けている位置をそれらのデータブロックで充填することができ、すなわち、位置A[0,5]は第1の行列内の位置A[1,6]内のデータブロックF5-4で充填される。記憶デバイス110は、図6-Bに示した第1の行列の第6番目の列内の任意のデータブロック(すなわち、データブロックF5-4およびF5-5)を選択して、第4番目の列の位置A[1,4]をそのデータブロックで充填する。
【００９９】
ステップ710で、記憶デバイス110は、充填した後の第y番目の列内のデータブロックに従って、第y番目の列内のチェックブロックを更新する。
【０１００】
第y番目の列内のデータブロックを有さない位置は新しいデータで充填されているため、記憶デバイス110は、判断されたチェックアルゴリズムと、充填後の第1の行列の第y番目の列内のすべてのデータブロックとに従って、第y番目の列内のチェックブロックを計算および更新する必要がある。yの値は、第1の行列内のデータブロックが欠けている様々な位置とともに変化する。たとえば、図6-Bに示すように、第y番目の列が図6-Bに示した第(M-1)番目の列内のデータブロックF4-5で充填される場合、充填後の第y番目の列内のデータは、図6-Dの第y番目の列内のデータブロック内に示すように、チェックブロックP5を更新するために、充填後の第y番目の列内のデータブロックF1-4、F4-5、およびF5-3に従って再計算される必要がある。図6-Bに示した第1の行列内の第(y+1)番目の列内に何のデ

ータブロックも有さない位置がやはり存在するため、記憶デバイス110は、第(M-1)番目の列から1個のデータを選択して、第(y+1)番目の列内のデータブロックを有さない位置を1個のデータブロックで充填し、次いで、第(y+1)番目の列内の更新されたデータブロックF1-5、F5-4、およびF4-4に従って、チェックブロックP6を再計算および更新することも可能であることを理解されよう。

【０１０１】

ステップ715で、記憶デバイス110は、第y番目の列内のデータブロックと、チェックブロックとからなるストライプを、第1のRAIDを形成するT枚のディスク内に書き込む。

【０１０２】

第1の行列の第y番目の列内のデータブロックが欠けている位置が新しいデータブロックで充填されて、第y番目の列内のチェックブロックが更新されるとき、記憶デバイス110は、第y番目の列内の更新されたデータブロックと、チェックブロックとからなるストライプをT枚のディスク内に書き込むことができる。

【０１０３】

図7に示した方法が使用された後、第1の行列内のデータブロックを有さない位置が依然として存在するとき、記憶制限時間に達した場合、ストライプがディスク内に書き込まれる前に、データを有さないデータブロックを0で充填することが可能であることを理解されよう。詳細については、前述の説明を参照することが可能であり、本明細書で詳細は繰り返して説明されない。データを有さないデータブロックを0で充填することは、そのデータブロックが使用されていないことを示すために使用されることを当業者は理解されよう。

【０１０４】

図7に示した方法を使用することによって、第1の行列内のデータブロックがディスク内に書き込まれるとき、1個のファイルが可能な限り少数のディスク内に記憶されて、ディスク空間が節約され得ることを確実にし得ることを前述の説明から理解されよう。

【０１０５】

さらに別の状況では、図2-Aまたは図2-Bで説明した方法を使用することによってファイルが記憶されるとき、記憶制限時間に達し、第1のRAIDがすでに満杯である場合、第1の行列内にあり、かつ第1のRAID内に書き込まれていないデータを第2のRAID内に書き込むことが可能である。第1のRAID内に書き込まれていないデータが第2のRAID内に書き込まれるとき、第2のRAIDの構成形式が第1のRAIDの構成形式と同じであり、第2の行列内のメンバーディスクの数量が第1の行列内のメンバーディスクの数量と同じである場合、たとえば、第1のRAIDと第2のRAIDが両方ともRAID5である場合、第2のRAID内のメンバーディスクの数量は、第1のRAIDのメンバーディスクの数量と同じであることを理解されよう。第1の行列内にあり、かつ第1のRAID内に書き込まれていないデータを、ステップ215の方法に従って、第2のRAIDを形成するディスク内に書き込むことが可能である。第2のRAIDの構成形式が第1のRAIDの構成形式と異なる場合、または第2の行列内のメンバーディスクの数量が第1の行列内のメンバーディスクの数量と異なる場合、たとえば、第1のRAIDはRAID3であり、第2のRAIDはRAID5である場合、残りのデータブロックを、前述のファイル処理方法に従って、第2のRAID内に再度書き込む必要がある。

【０１０６】

図8は、本発明の一実施形態による別のファイル処理方法の流れ図である。この方法は、独立チェックディスクを有するRAIDの構成形式だけに適用され得る。この方法は、図1-Aに示した記憶デバイス110によって実行されることも可能である。図8に示すように、この方法は以下を含む。

【０１０７】

ステップ800で、記憶デバイス110は、第1のRAID内に記憶されることになるF個のファイルを受信する。

【０１０８】

ステップ805で、記憶デバイス110は、第1のRAIDのストリップサイズに従って、F個のフ

ァイルを少なくとも2個のデータブロックに分割する。

【０１０９】

ステップ800およびステップ805の関連説明については、図2-Aのステップ200およびステップ205の関連説明を参照することが可能である。

【０１１０】

ステップ810で、記憶デバイス110は、分割することによって取得された、少なくとも2個のデータブロックを1個のアレイに配列する。アレイ内の(D-1)個の位置の間隔は1個のファイルに属する2個の隣接するブロック同士の間であり、Dの値は第1のRAID内のデータディスクの数量である。

【０１１１】

具体的には、データブロックがアレイに配列されるとき、第1のRAIDの構成形式と、第1のRAID内のデータディスクの数量とに従って、少なくとも2個のデータブロックをどのように配列するかを判断する必要がある。第1のRAIDの構成形式が独立チェックディスクを有するRAIDであるとき、たとえば、第1のRAIDがRAID3またはRAID4であるとき、配列されたアレイ内で、1個のファイルに属する2個の隣接するデータブロックを(D-1)個の位置に対して間隔を開ける必要があり、この場合、Dの値は第1のRAID内のデータディスクの数量である。たとえば、図4-Dを参照すると、RAIDの構成形式はRAID3である。第1のRAIDは4枚のディスクを含み、この場合、D1、D2、およびD3はデータディスクであり、D4は独立チェックディスクである。分割することによって取得されたデータブロックは、図9に示したアレイに配列され得る。ファイルF1のデータブロックF1-1およびF1-2は、2つの位置に対して間隔が開けられて、ファイルF2のデータブロックF2-1およびF2-2も、2つの位置に対して間隔が開けられる、等々である。

【０１１２】

ステップ815で、記憶デバイス110は、アレイ内のD個のデータブロックと、D個のデータブロックに従って計算することによって取得されたP個のチェックブロックとからなるストライプを、第1のRAIDを形成するディスク内に書き込み、この場合、Pの値は、第1のRAID内の独立チェックディスクの数量である。

【０１１３】

具体的には、データブロックを、第1のRAIDを形成するディスク内に記憶するプロセスでは、P個のチェックブロックを取得するために、データグループから順次に選択されたD個のデータに関して、第1のRAIDのチェックアルゴリズムに従って、チェック計算を実行する必要がある。ストライプは、D個のデータと、計算することによって取得されたP個のチェックブロックとから順になり、第1のRAIDを形成するディスク内に書き込まれる。チェックブロックがディスク内に書き込まれるとき、チェックブロックを第1のRAID内の独立チェックディスク内に書き込む必要があることを当業者は理解されよう。たとえば、図9に示した第1のアレイ内のデータがディスク内に書き込まれた後、図4-Dに示した記憶構造を取得することが可能である。

【０１１４】

図8に示したファイル処理方法を使用することによって、ファイルをRAID内に同時に書き込むことが可能であり、それによって、ファイル書込み効率を確実にして、1個のファイルが1枚のディスク内に記憶されることを確実にする。加えて、ストライプは異なるファイルのデータブロックからなり、ファイルが損傷したとき、他のファイルに従って、損傷したファイルを復元することが可能であり、これは、ファイル記憶装置のセキュリティを確実にする。

【０１１５】

本発明の本実施形態では、前述の図2-A、図2-B、または図8に示したファイル処理方法を使用することによって、ファイルがRAIDを形成するディスク内に記憶された後、アーカイブシナリオでは、記憶されたファイルは比較的低い頻度でアクセスされる。したがって、エネルギー節約目的を達成するために、ディスクは、一般に、休止状態または電源切断状態にされる。ファイルを読み取る必要があるとき、図10で説明された方法に従ってファ

イルを読み取ることが可能である。以下では、図1-Aと図1-Bとを参照して、図10を説明する。この方法は以下を含む。

【０１１６】

ステップ225で、記憶デバイス110は、ホスト100のアクセス要求を受信し、この場合、アクセス要求はRAID内に記憶されたファイルを読み取るために使用され、アクセス要求は、読み取られることになるファイルに関する論理アドレスを搬送する。アクセス要求は、アクセスされることになるファイルの名前を搬送することも可能であることを理解されよう。

【０１１７】

ステップ230で、記憶デバイス110は、論理アドレスに従って、ファイルのデータブロックが記憶された物理アドレスに問い合わせる。一般に、記憶デバイス110がデータを記憶した後、データを記憶するための物理アドレスと論理アドレスとの間のマッピング関係のマッピング表が形成される。ファイルを読み取るためのアクセス要求を受信した後、記憶デバイス110は、ディスク内のデータの物理アドレスに問い合わせるために、アクセス要求内で搬送される論理アドレスに従って、マッピング表をチェックすることができる。RAID内で、マッピング表は、キャッシュ120内のデータとディスク130内のデータの両方に関して形成され得ることを理解されよう。物理アドレスに問い合わせるとき、一般に、キャッシュ120のマッピング表にまず問い合わせることができ、次いで、ディスク130のマッピング表に問い合わせる。データがキャッシュ内にある場合、キャッシュ内のデータはホストに直接返される。

【０１１８】

ステップ235で、記憶デバイス110は、物理アドレスに従って、ファイルを記憶するためのディスクを判断する。本発明の本実施形態では、ファイルが記憶された後、前述の実施形態のファイル処理方法を使用することによって、RAIDを形成するディスク内で、1個のファイルを1枚のディスク内に記憶することが可能である。したがって、このステップでは、記憶デバイス110は、物理アドレスに従って、ファイルを記憶するための1枚のディスクを判断することができる。

【０１１９】

ステップ240で、記憶デバイス110は、ディスク内に記憶されたファイルのデータブロックをホスト100に返す。具体的には、記憶デバイス110は、物理アドレスに従って、ファイルが配置されたディスク130を起動させて、取得された物理アドレスに従って、ディスク130内のデータを読み取り、そのデータをホスト100に返すことができる。

【０１２０】

本発明の本実施形態では、ファイルは図2-A、図2-B、または図8に示した方法に従って記憶されるため、1個のファイルに属するデータは可能な限り少数のディスク内に記憶される。したがって、ファイルが読み取られるとき、そのファイルが記憶された1枚のディスクだけを起動させる必要があり、そのファイルのデータは、起動された1枚のディスクから読み取られ、ホストに返され、RAID全体を形成するすべてのディスクを起動させる必要はなく、それによって、明らかなエネルギー節約効果をもたらす。

【０１２１】

本発明の本実施形態では、メタデータを記憶するためのディスクおよびキャッシュ120は、ホストのアクセス要求に適時に応答するために、常に、電源投入状態にとどまることを当業者は理解されよう。メタデータはRAID内に記憶されたデータを記述するデータであり、データ、たとえば、メタデータの環境は、論理アドレスと物理アドレスとの間のマッピング関係を含み得ることを当業者は理解されよう。

【０１２２】

ディスクの頻繁な電源投入または電源切断は、記憶システムのエネルギー節約効果に影響を及ぼす可能性があり、ディスクの寿命に影響を及ぼす可能性もある。したがって、ディスクが頻繁に電源投入または電源切断されるのを避けるために、実際のアプリケーションでは、ディスクを等級づけることが可能である。少数の高性能ディスクは常に電源投入

状態にとどまると同時に、多数の大容量ディスクはエネルギー節約状態に入る。高性能ディスクは、本明細書では、比較的低いアクセス遅延を有するディスク、または1秒あたりの入出力回数(Input/Output Operations Per Second、IOPS)が比較的多いディスク、たとえば、固体ディスク(Solid State Disk、SSD)を指す。大容量ディスクは、比較的大きな容量のディスクを指す。記憶した後、ディスクが電源投入または起動される回数を削減して、応答速度を改善するために、ファイルアクセス条件に従って、高いアクセス頻度を有するファイルを、常に電力投入状態にとどまる少量の確保された高性能ディスクに移動させることができる。

【０１２３】

さらに、ディスクが頻繁に電源投入または電源切断されるのを避けるために、本発明の本実施形態の記憶システムは、警告機構と保護機構とを提供することも可能である。RAID内の各ディスクが電源投入または電源切断された累積回数に関する統計が収集される。事前設定された期間内のディスクの電源投入および電源切断の回数が事前設定されたしきい値を超えるとき、システムは、催促または警告を与えて、何らかの保護対策をとることができる。たとえば、設定されたしきい値は、1日あたり10回または1月あたり100回であり得る。保護対策は、設定された時間内にディスクに電源投入または電源切断を行わない、等々と設定されてよく、これは本明細書で限定されない。

【０１２４】

図11は、本発明の一実施形態によるファイル処理装置の概略構造図である。図11に示すように、ファイル処理装置1100は、以下を含む。

RAID内に記憶されることになるF個のファイルを受信するように構成された受信モジュール1102であって、RAIDがT個の記憶装置によって形成され、Fが2以上の自然数であり、Tが3以上の自然数である、受信モジュール1102と、

RAIDのストリップサイズに従って、F個のファイルを少なくとも2個のデータブロックに分割するように構成された分割モジュール1104と、

少なくとも2個のデータブロックに従って、T個の行を有する第1の行列を取得するように構成された処理モジュール1106であって、1個のファイルに属するデータブロックが第1の行列内の1個の行内に配置される、処理モジュール1106と、

第1の行列内の各列内のデータブロックと、その列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むように構成された書込みモジュール1108と

を含む。

【０１２５】

具体的には、処理モジュール1106は、

分割することによって取得された、少なくとも2個のデータブロックをD個の行を有する第2の行列に配列することであって、1個のファイルに属するデータブロックが第2の行列の1個の行内に配置され、DがRAID内のデータ記憶装置の数量である、配列することと、

チェックブロックを第2の行列の各列にそれぞれ挿入することによって、T個の行を有する第1の行列を取得することであって、挿入されたチェックブロックが、第1の行列内のチェックブロックが配置された列内のデータブロックに従って計算することによって取得される、取得することと

を行うように構成される。

【０１２６】

一事例では、RAIDが独立チェック記憶装置を含むとき、処理モジュールは、具体的には、

RAID内の独立チェック記憶装置の位置に従って、チェックブロックを第2の行列に挿入するための位置を判断することと、

RAIDのチェックアルゴリズムに従って、第2の行列の各列内のデータブロックに関する計算を実行して、各列内のデータブロックのチェックブロックを取得することと、

第2の行列の各列内のデータブロックに従って計算することによって取得されたチェッ

クブロックの判断された位置に従って、チェックブロックをその列に挿入することによって、T個の行を有する第1の行列を取得することと
を行うように構成される。
【０１２７】
別の事例では、RAIDが独立チェック記憶装置を含まないとき、処理モジュールは、具体的には、
チェックブロックを第2の行列の各列に挿入するための位置A[x,y]を判断することであって、第2の行列がN個の列を有し、xとyが両方とも整数であり、xの値が0からD-1に徐々に増大し、yの値が0からN-1に徐々に増大する、判断することと、
第2の行列の第x番目の行内の第y番目の列から第(N-1)番目の列までのデータブロックを第x番目の行内の第(y+1)番目の列から第N番目の列までの位置に順次に移動させることと、
RAIDのチェックアルゴリズムに従って、第y番目の列内のデータブロックに関するチェック計算を実行して、第y番目の列内のデータブロックのチェックブロックを取得することと、
第y番目の列内のデータブロックのチェックブロックを第2の行列の第y番目の列内の位置A[x,y]に挿入することによって、T個の行を有する第1の行列を取得することと
を行うように構成される。
【０１２８】
一事例では、書込みモジュール1108は、具体的には、
第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されているとき、第y番目の列内のデータブロックと、チェックブロックとを、RAIDを形成するT個の記憶装置内に書き込むことであって、第y番目の列が第1の行列内の列のうちの1つである、書き込むこと
を行うように構成される。
【０１２９】
別の事例では、書込みモジュール1108は、具体的には、
第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されていないとき、第y番目の列内の欠けているデータブロックの数量を判断することであって、第y番目の列が第1の行列内の列のうちの1つである、判断することと、
第1の行列内の第(M-1)番目の列から第(y+1)番目の列までの数量のデータブロックを選択して、第y番目の列内のデータブロックを有さない位置をそれらのデータブロックで充填することと、
充填した後の第y番目の列内のデータブロックに従って、第y番目の列内のチェックブロックを更新することと、
第y番目の列内のデータブロックと、第y番目の列内の更新されたチェックブロックとからなるストライプを、RAIDを形成するT個の記憶装置内に書き込むことと
を行うように構成される。
【０１３０】
さらに別の事例では、書込みモジュール1108は、具体的には、
第1の行列の第y番目の列内のデータブロックと、第y番目の列内のデータブロックに従って計算することによって取得されたチェックブロックとからなるストライプが完全に占有されていないとき、第y番目の列内のデータブロックを有さない位置を0で充填することと、
0で充填した後の第y番目の列内のデータブロックと、チェックブロックとからなるストライプを、第1のRAIDを形成する記憶装置内に書き込むことであって、第y番目の列が第1の行列内の列のうちの1つである、書き込むことと
を行うように構成される。

【０１３１】
さらに、さらに別の事例では、受信モジュール1102は、ホストのアクセス要求を受信することであって、アクセス要求はRAID内に記憶されたファイルを読み取るために使用され、アクセス要求は、アクセスされることになるファイルに関する論理アドレスを搬送する、受信することを行うようにさらに構成され得、
ファイル処理装置は、
論理アドレスに従って、ファイルのデータブロックが記憶された物理アドレスに問い合わせて、物理アドレスに従って、ファイルが記憶された1個の記憶装置を判断して、記憶装置内に記憶されたファイルのデータブロックをホストに返すように構成された読取りモジュール1110
をさらに含む。
【０１３２】
本発明の本実施形態で提供されるファイル処理装置は、前述の実施形態で説明したコントローラ内に配置されることが可能であり、かつ前述の実施形態で説明したファイル処理方法を実行するように構成される。各ユニットの機能の詳細な説明について、これらの方法実施形態の説明を参照することができ、本明細書で詳細は繰り返して説明されない。
【０１３３】
本発明の本実施形態で説明したファイル処理装置は、1個のファイルに属するデータを1枚のディスク内に記憶することができる。加えて、本発明の本実施形態で説明したファイル処理装置は、異なるファイルのデータブロックを使用することによって、ストライプを形成して、そのストライプをディスク内に書き込むことができる。データブロックが損傷したとき、ファイル処理装置は、同じストライプのチェックブロックまたは他のファイルのデータブロックを使用することによって、損傷したデータブロックを復元することができ、それによって、ファイル記憶装置のセキュリティを改善する。さらに、ファイルを読み取るとき、本発明の本実施形態で説明されたファイル処理装置は、ファイルが記憶された1枚のディスクだけを起動させるか、またはそのディスクだけに電源投入して、そのディスクからファイルのデータを読み取って、そのデータをホストに返せばよく、RAID内のすべてのディスクを起動させるか、またはそれらのディスクに電源投入する必要はなく、それによって、より良好なエネルギー節約効果を達成する。
【０１３４】
本発明の本実施形態は、プログラムコードを記憶したコンピュータ可読記憶媒体を含む、データ処理のためのコンピュータプログラム製品であって、プログラムコード内に含まれた命令が、前述の方法実施形態のうちのいずれか1つで説明した方法プロセスを実行するために使用される、コンピュータプログラム製品をさらに提供する。前述の記憶媒体は、USBフラッシュドライブ、リムーバブルハードディスク、磁気ディスク、光ディスク、ランダムアクセスメモリ(Random-Access Memory、RAM)、固体ディスク(Solid State Disk、SSD)、または不揮発性メモリ(non-volatile memory)など、プログラムコードを記憶することが可能な任意の非一時的(non-transitory)機械可読媒体を含むことが可能であることを当業者は理解されよう。
【０１３５】
本出願で提供されたいくつかの実施形態では、開示された装置および方法を他の様式で実装することが可能であることを理解されたい。たとえば、上で説明した装置実施形態は単なる例である。たとえば、モジュール分割は、単なる論理的な機能分割であり、実際の実装形態では、他の分割であり得る。たとえば、複数のモジュールもしくは構成要素を組み合わせるかまたは統合して別のデバイスにすることが可能であり、あるいは、いくつかの特徴を無視すること、または実行しないことが可能である。加えて、表示もしくは議論された相互結合または直接結合あるいは通信接続は、いくつかの通信インターフェースを介して実装され得る。装置間またはモジュール間の間接的結合または通信接続は、電子形態、機械形態、または他の形態で実装され得る。
【０１３６】

別個の部分として説明されたモジュールは、物理的に分離されてよく、もしくは物理的に分離されなくてもよく、モジュールとして表示された部分は、物理ユニットであってよく、もしくはそうでなくてもよく、1つの位置に配置されてよく、または複数のネットワークユニット上に分散されてもよい。モジュールの一部またはすべては、これらの実施形態の解決策の目的を達成するための実際のニーズに従って、選択され得る。

【０１３７】

加えて、本発明のこれらの実施形態の機能モジュールは、1個の処理モジュール内に統合されてよく、もしくはそれらのモジュールの各々は物理的に単独で存在してよく、または2個以上のモジュールを統合して、1個のモジュールにしてもよい。

【０１３８】

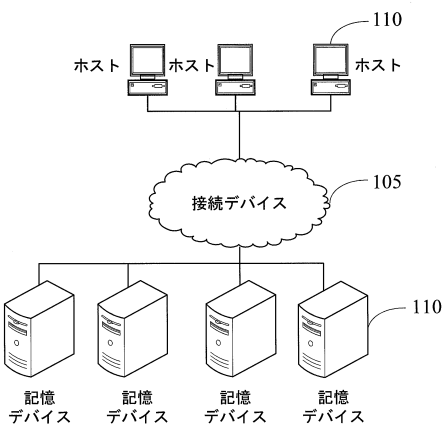
最後に、前述の実施形態は、本発明の限定するのではなく、本発明の技術的解決策を説明することを単に意図することに留意されたい。本発明は前述の実施形態を参照して詳細に説明されたが、そのような修正または置換が対応する技術的な解決策の本質を本発明の実施形態の技術的解決策の範囲から逸脱させない限り、当業者は前述の実施形態で説明された技術的解決策に依然として修正を行うこと、またはそれらのいくつかのもしくはすべての技術的特徴に均等の置換を行うことが可能であることを当業者は理解されたい。

【符号の説明】

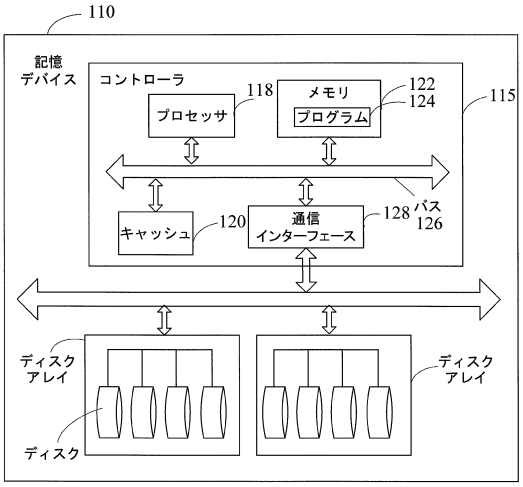
【０１３９】

100　ホスト
105　接続デバイス
110　記憶デバイス
115　コントローラ
118　プロセッサ
120　キャッシュ
122　メモリ
124　プログラム
125　ディスクアレイ
126　通信バス(バス)
128　通信インターフェース
130　ディスク
1100　ファイル処理装置
1102　受信モジュール
1104　分割モジュール
1106　処理モジュール
1108　書込みモジュール
1110　読取りモジュール

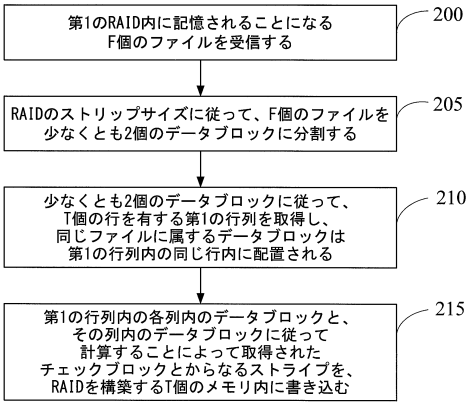
【図１－Ａ】



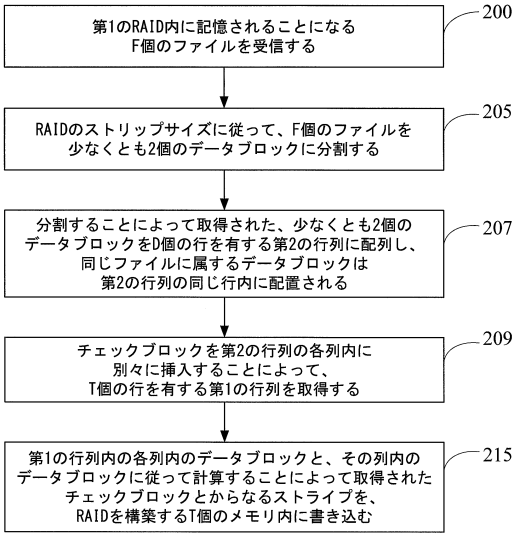
【図１－Ｂ】



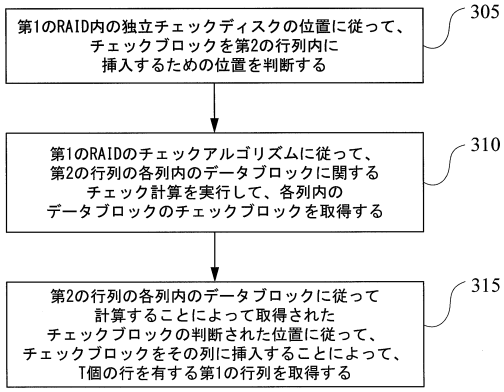
【図２－Ａ】



【図２－Ｂ】



【図３】



【図４－Ａ】

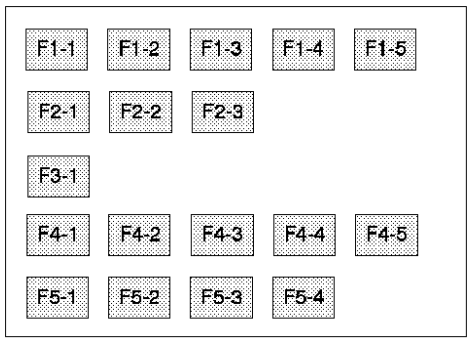


图 4-A

【図４－Ｂ】

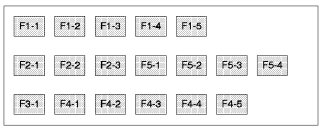


图 4-B

【図４－Ｃ】

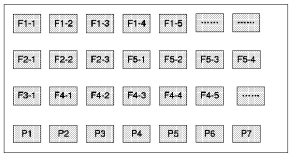


图 4-C

【図４－Ｄ】

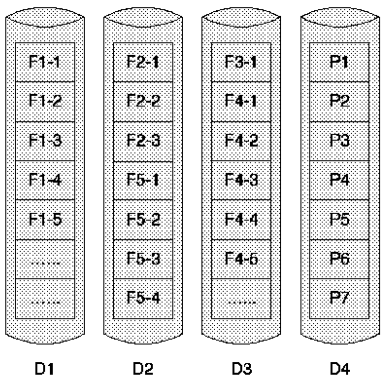
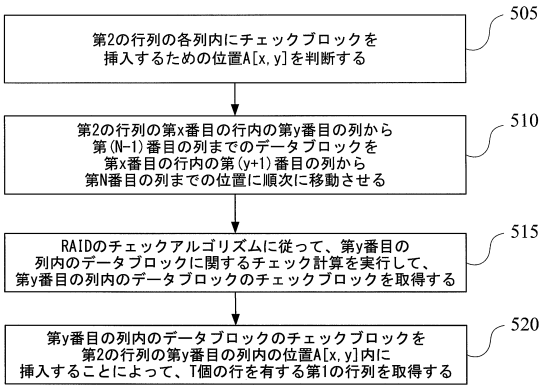


图 4-D

【図５－Ａ】



【図６－Ａ】

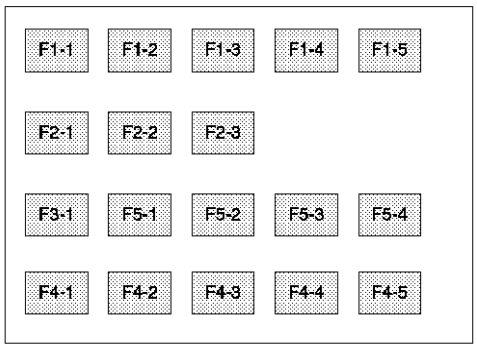
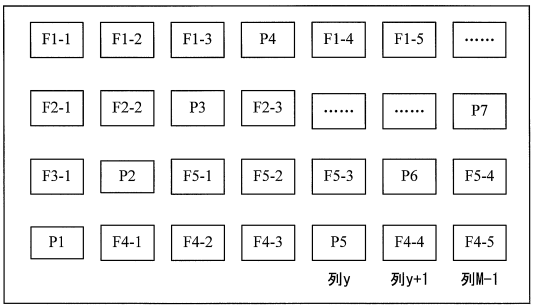


図 6-A

【図６－Ｂ】



【図６－Ｃ】

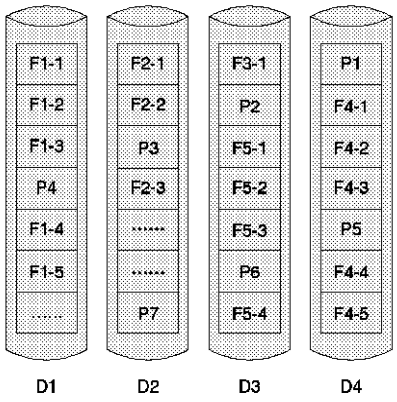
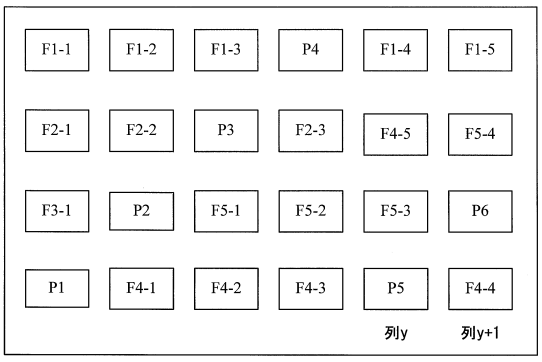
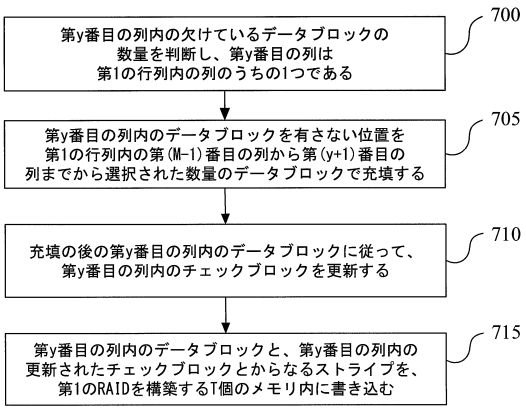


図 6-C

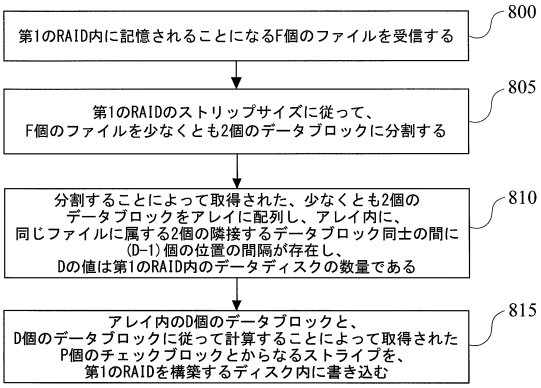
【図６－Ｄ】



【図７】



【図８】



【図９】

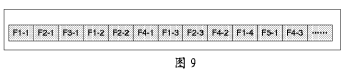
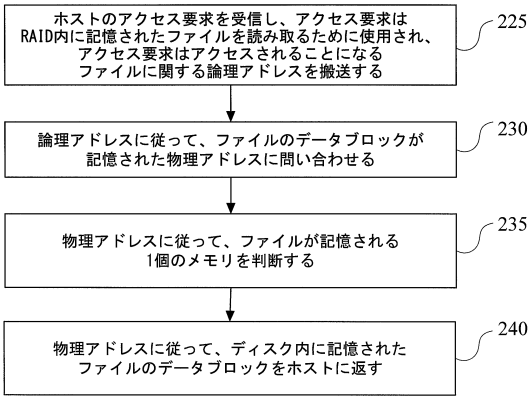
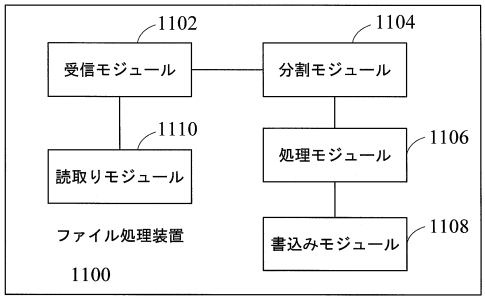


図 9

【図１０】


ホストのアクセス要求を受信し、アクセス要求は
RAID内に記憶されたファイルを読み取るために使用され、
アクセス要求はアクセスされることになる
ファイルに関する論理アドレスを搬送する
225
論理アドレスに従って、ファイルのデータブロックが
記憶された物理アドレスに問い合わせる
230
物理アドレスに従って、ファイルが記憶される
1個のメモリを判断する
235
物理アドレスに従って、ディスク内に記憶された
ファイルのデータブロックをホストに返す
240

【図１１】


1102
受信モジュール
1104
分割モジュール
1110
読取りモジュール
処理モジュール
1106
書込みモジュール
1108
ファイル処理装置
1100

フロントページの続き

(72)発明者　孔　▲ハン▼
中華人民共和国５１８１２９広東省深▲セン▼市龍岡区坂田華為本社ビル
(72)発明者　王　静
中華人民共和国５１８１２９広東省深▲セン▼市龍岡区坂田華為本社ビル

審査官　田中　啓介

(56)参考文献　特開２００６－２６０５８２（ＪＰ，Ａ）
米国特許第０８００６１１１（ＵＳ，Ｂ１）
米国特許第０５８６００９０（ＵＳ，Ａ）
特開平０７－１５２４９８（ＪＰ，Ａ）
特開平１１－２８８３５９（ＪＰ，Ａ）
特開２０１０－０７９９２８（ＪＰ，Ａ）
特開２００４－１４５４０９（ＪＰ，Ａ）
特開平１１－２８８３８７（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ０６Ｆ３／０６－３／０８
１１／０８－１１／１０
１２／００、１３／１０－１３／１４