

(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2005-293598  
(P2005-293598A)

(43) 公開日 平成17年10月20日(2005.10.20)

(51) Int.C1.<sup>7</sup>

F 1

テーマコード(参考)

G 06 F 3/06

G 06 F 3/06

3 0 2 Z

5 B 0 6 5

G 11 B 20/10

G 11 B 20/10

3 1 1

5 D 0 4 4

G 11 B 20/12

G 11 B 20/12

審査請求 有 請求項の数 11 O.L. (全 9 頁)

(21) 出願番号 特願2005-108214 (P2005-108214)  
 (22) 出願日 平成17年4月5日 (2005.4.5)  
 (31) 優先権主張番号 10/816897  
 (32) 優先日 平成16年4月5日 (2004.4.5)  
 (33) 優先権主張国 米国(US)

(71) 出願人 503003854  
 ヒューレット-パッカード デベロップメント カンパニー エル.ピー.  
 アメリカ合衆国 テキサス州 77070  
 ヒューストン 20555 ステイト  
 ハイウェイ 249  
 (74) 代理人 10009623  
 弁理士 奥山 尚一  
 (74) 代理人 100096769  
 弁理士 有原 幸一  
 (74) 代理人 100107319  
 弁理士 松島 鉄男  
 (74) 代理人 100114591  
 弁理士 河村 英文

最終頁に続く

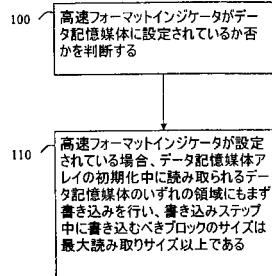
(54) 【発明の名称】データ記憶媒体を所望のブロックフォーマットにフォーマットする方法およびシステム

## (57) 【要約】

【課題】データ記憶媒体を所望のブロックフォーマットにフォーマットする方法およびシステムを提供する。

【解決手段】データ記憶媒体アレイ内にある、または該データ記憶媒体アレイに追加すべきデータ記憶媒体を所望のブロックフォーマットにブロックフォーマットする方法であって、該方法は、一実施態様においては、データ記憶媒体アレイの初期化中に読み取られるデータ記憶媒体のいずれの領域にもまず書き込みを行うステップと、初期化後に実行される書き込みステップ用の書き込みサイズを、読み取りサイズ以上のサイズに設定するステップと、を含む。

【選択図】図1



**【特許請求の範囲】****【請求項 1】**

データ記憶媒体アレイ内にある、または該データ記憶媒体アレイに追加すべきデータ記憶媒体を所望のブロックフォーマットにブロックフォーマットする方法であって、

高速フォーマットインジケータが前記データ記憶媒体に設定されているか否かを判断するステップと、

前記高速フォーマットインジケータが設定されている場合、前記データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行うステップと、

を含む、方法。

10

**【請求項 2】**

前記高速フォーマットインジケータが設定されており、前記データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行うステップが行われる場合、その後、すべての書き込みが最大読み取りサイズ以上のサイズで行われるライト・ビフォア・リード・モデルに従う、請求項 1 に記載の方法。

**【請求項 3】**

前記書き込みを行うステップが行われた後に、前記高速フォーマットインジケータを変更するステップをさらに含む、請求項 1 に記載の方法。

**【請求項 4】**

前記データ記憶媒体アレイが R A I D データ記憶媒体アレイを含む、請求項 1 に記載の方法。

20

**【請求項 5】**

自動データ記憶媒体ブロックフォーマット機能を有するデータ記憶媒体アレイであって、

複数のデータ記憶媒体と、

高速フォーマットインジケータが前記データ記憶媒体に設定されているか否かを判断し、前記高速フォーマットインジケータが設定されている場合、前記複数のデータ記憶媒体の初期化を行うトポジマネージャシステムであって、該データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行う、トポジマネージャシステムと、

を備えている、データ記憶媒体アレイ。

30

**【請求項 6】**

前記高速フォーマットインジケータが設定されており、前記トポジマネージャが、前記データ記憶媒体アレイの初期化が行われる間に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行う場合、前記トポジマネージャはその後、すべての書き込みが最大読み取りサイズ以上のサイズで行われるライト・ビフォア・リード・モデルに従う、請求項 5 に記載のデータ記憶媒体アレイ。

**【請求項 7】**

データ記憶媒体アレイ内にある、または該データ記憶媒体アレイに追加すべきデータ記憶媒体を所望のブロックフォーマットにブロックフォーマットするプログラム製品であって、実行されると、

高速フォーマットインジケータが前記データ記憶媒体に設定されているか否かを判断するステップと、

前記高速フォーマットインジケータが設定されている場合、前記データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行うステップと、

を含む方法を機械に実行させる機械可読プログラムコードを含む、プログラム製品。

**【請求項 8】**

前記高速フォーマットインジケータが設定されており、前記データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行うス

40

50

ステップが行われる場合、その後、すべての書き込みが最大読み取りサイズ以上のサイズで行われるライト・ビフォア・リード・モデルに従う、請求項7に記載のプログラム製品。

【請求項9】

前記書き込みを行うステップが行われた後に、前記高速フォーマットインジケータを変更するプログラムコードをさらに含む、請求項7に記載のプログラム製品。

【請求項10】

データ記憶媒体アレイ内にある、または該データ記憶媒体アレイに追加すべきデータ記憶媒体を所望のブロックフォーマットにブロックフォーマットするシステムであって、

高速フォーマットインジケータが前記データ記憶媒体に設定されているか否かを判断する手段を備え、

前記高速フォーマットインジケータが設定されている場合、前記データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行う、システム。

【請求項11】

自動データ記憶媒体ブロックフォーマット機能を有するデータ記憶媒体アレイであって、

複数のデータ記憶媒体と、

高速フォーマットインジケータが複数の前記データ記憶媒体のそれぞれに設定されているか否かを判断する手段と、

前記高速フォーマットインジケータが設定されている場合、前記データ記憶媒体アレイの初期化中に読み取られる前記データ記憶媒体のいずれのブロックにもまず書き込みを行う手段と、

を備えているデータ記憶媒体アレイ。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、包括的にはデータ記憶媒体アレイに関する。より詳細には、本発明は、データ記憶媒体アレイ内のデータ記憶媒体をフォーマットする方法に関する。より詳細には、本発明は、高速フォーマットされた媒体をデータ記憶媒体に包含する方法に関する。

【背景技術】

【0002】

通常、従来のコンピュータネットワークは、Hewlett Packard社によるH p SureStore Virtual Array 7100等、中央データ記憶用に少なくとも1つの共通データ記憶装置を有する。SureStore VA 7100は、耐故障性およびパフォーマンス考慮事項のために組み合わせられた複数の独立したディスク（たとえば、ファイバチャネルハードドライブ）を採用する1つの例示的な独立ディスク冗長アレイ（RAID）である。従来のRAIDアレイおよびデータ記憶媒体アレイを採用する他の装置は、アレイ内のデータ記憶媒体間の一貫しないブロックフォーマットに悩まされている。本明細書において利用するブロックフォーマットは、データ記憶媒体のブロックサイズおよび／またはブロックデータ内容のフォーマットを指し、区分化および／またはデータフォーマットを包含するように広く解釈することができることを理解されたい。

【0003】

具体的には、データを記憶するデータ記憶媒体アレイはそれぞれ、所望のシステムブロックフォーマットを有する。データ記憶媒体アレイに記憶されているデータファイルが開始および終了する場所をオペレーティングシステムが適宜規定することができるような特定のブロックフォーマットが要求される。しかし、従来のデータ記憶媒体は通常、所望のシステムブロックフォーマットと異なり得る所定のブロックフォーマットを有して製造される。たとえば、Seagate製の標準的な18GBハードドライブ（部品番号ST318451F）は、標準の512バイトブロックフォーマットを備えている。しかし、H

ewlett Packard SureStore VA 7100 には、様々な故障検出およびエラー訂正手順に使用される 8 バイトチェックサムを記憶データ 512 バイト毎に提供する 520 バイトブロックフォーマットが要求される。したがって、520 バイトデータ記憶媒体は、非標準ブロックフォーマットを有するシステムに向けて特別に製造する必要があり、これはシステムを構築するコストを増大させる。この問題に対する従来の解決策（オートフォーマット）は、「Method for Automatically Converting Block Size and Formatting Backend Fibre Channel Disks in an Auto-Inclusive Storage Array Environment」という名称の特許第 6,681,308 号に記載された。

10

#### 【0004】

ドライブのオートフォーマット (SCSI フォーマット) に伴う冗長的なフォーマット時間を回避する 1 つの選択肢は、ブロック書き込みのフォーマットテンプレートのみが（本実施形態では、512 から 520 に）変更され、ブロックに対して行われるフォーマットが最小であるか、またはない「高速フォーマット」として知られる方法を用いることである。具体的には、「書き込み」動作を所与のブロックに対して行うべきときまでブロックはリフォーマットされない。「書き込み」動作が呼び出されると、520 フォーマットテンプレートが使用されて、特定のブロックがリフォーマットされる。高速フォーマット方法を実行可能にする仮定は、各ブロックが、読み取り前に書き込まれる (write before read: ライト・ビフォア・リード) というものである。1 つの「高速フォーマット」方法が Seagate のドライブによって提供され、この方法では、ブロック書き込み境界のテンプレートのみが 512 から 520 に変更され、次いで、完全なフォーマットが完了したかのように戻されることにより、行われるドライブの準備は最小である。このような高速フォーマット動作は、各ブロックが記憶媒体の使用に先立って個々にフォーマットされている標準フォーマット技法を使用する新しい 512 バイトブロックドライブの導入と比較して、VA 7100 / 7400 / 7410 / 7110 等の記憶アレイが略瞬時にこのディスクを包含し使用することができることを意味する。不都合なことに、この高速フォーマットは、ディスクをドライブが 520 バイトブロックフォーマットとして報告する状態にするが、すべてのブロックは実際には 512 バイトブロックサイズにフォーマットされたままである。しかし、記憶アレイはこの新しい想定に気づかず、ブロックを書き込み前に読み取る場合がある。書き込み前に読み取ると、誤った故障予測に繋がり得る冗長的な再試行および累積エラーが発生し得る。

20

30

#### 【0005】

したがって、高速フォーマット済み媒体をデータ記憶媒体アレイに統合する方法が必要である。

#### 【発明を実施するための最良の形態】

#### 【0006】

本発明の一態様による例示的なデータ記憶媒体アレイ（本明細書中以下、「アレイ」）200 を図 2 のブロック図に示す。アレイ 200 は、一実施形態では、アレイ 200 の一部であっても、またはリンク 230（たとえば、ネットワークまたは他の都合のよいリンク）を介して接続されてもよいトポロジマネージャ 220 によって制御される。トポロジマネージャ 220 は、アレイ内の様々なデータ記憶媒体を初期化することができ、アレイ 200 にデータを記憶することができ、かつ／またはブロックフォーマット制御信号を生成して、アレイ 200 内のデータ記憶媒体（本明細書中以下、「ドライブ」）210 のブロックフォーマットを開始することができる。図 2 は、例として 8 つのドライブ 210 を示す。当業者にとって容易に理解されるように、他のアレイ構成を実施することも可能である。したがって、図 2 のアレイは単なる例示にすぎず、本発明の範囲を限定するものではない。

40

#### 【0007】

本発明によるデータ記憶媒体という語は、テープおよびディスク等のデータ記憶媒体、

50

および当業者にとって容易に理解される他のデータ記憶媒体を含むものとして広く解釈されることを意図し、したがって本発明の範囲の限定を意図するものではない。説明のみを目的とし、本発明を限定せず、本発明の実施形態について、RAIDアレイにおいて一般に実施されるハードディスクドライブを参照して以下説明する。

#### 【0008】

図1を参照すると、ブロック100において、本発明の方法の一実施形態の第1のステップは、高速フォーマットインジケータがデータ記憶媒体に設定されているか否かを判断することである。本発明の一実施形態は、ドライブの追加および「高速フォーマット」の直後に実施することができる。この実施形態では、フォーマットテンプレートが設定された後、高速フォーマットビット等の高速フォーマットインジケータを設定することができる。高速フォーマットインジケータは、書き込み用のフォーマットテンプレートのみが設定され、初期化中に読み取られるエリアにはまだ書き込まれていないことを示すために使用される。

#### 【0009】

ブロック110を参照すると、高速フォーマットインジケータが設定されている場合、データ記憶媒体アレイの初期化中に読み取られるデータ記憶媒体のいずれのブロックにもまず書き込みが行われる。記憶媒体アレイの初期化中に読み取られる記憶媒体のブロック領域としては、たとえば、ドライブのメンバーシップメタデータスタンプエリアを挙げることができる。「初期化」は、電源投入中またはディスクをアレイに包含する間のすべての動作として広く定義される。記憶媒体の初期化中に読み取るべきブロックは、経験的またはその他の方法で決定することができる。記憶媒体によっては、媒体内の固定エリアのごくわずかな部分が読み取られる。しかし、記憶媒体によっては、初期化中に読み取るエリアがはるかに大きいものもある。一実施形態では、トポロジマネージャシステムを使用して、ドライブサイズまたは他の或る基準に基づいて、メタデータスタンプをデータ記憶媒体に書き込むべき場所を回転立ち上げ中に求めることができる。初期化中に書き込むべきブロックを求める方法は本発明において限定されない。

#### 【0010】

上記書き込みステップが行われた後、本方法が行われたことを示すように高速フォーマットインジケータを変更することができる。

#### 【0011】

初期化後、データチャンクが書き込まれる際に、ブロックセット（チャンク）の書き込み動作の粒度は、最大読み取りサイズ以上であるべきである。たとえば、データ復元動作中、別のディスクに移すべきデータが8メガチャンクで読み取られる場合、書き込むべきブロック数は、その8メガチャンクのごくわずかな部分が実際に読み取られる場合であっても8メガチャンクの包含に十分なブロック数になる。したがって、高速フォーマットインジケータが設定され、データ記憶媒体アレイの初期化中に読み取られるデータ記憶媒体のいずれのブロックにもまず書き込みを行うステップが行われる場合、システムはその後、すべての書き込みが最大読み取りサイズ以上のサイズで行われるライト・ビフォア・リード・モデルに従う。本発明の一実施形態では、図2のトポロジマネージャ220は、ブロック100および110のステップを実行するように設定することができる。上記例でのサイズ数は本発明を限定しないことに留意する。

#### 【0012】

一実施形態では、データ書き込みおよび読み取りを行う（初期化書き込みおよび読み出しとは対照的に）任意のファームウェア構成要素がライト・ビフォア・リード・モデルに従わないかどうかを経験的にまたは他の方法で判断する動作を提供することができる。このような動作は、まだ書き込まれていないため、正しいフォーマットにまだリフォーマットされていないブロックの読み取りを妨げる。このようなファームウェアが特定される場合、適正なブロックフォーマットに設定するために、このようなファームウェアによって読み取られるいずれの領域にも書き込みを行わなければならない。

#### 【0013】

10

20

30

40

50

一実施形態では、別個のステップを行って記憶システムの読み取りサイズを求めることができる。ステップは経験的に求めて、または他の或る都合のよい様式で求めてよい。

#### 【0014】

図3は、本発明のさらなる実施形態のフローチャートである。ブロック300は、記憶システムの最大読み取りサイズを求めるために行われるステップである。ブロック310において、高速フォーマットインジケータが設定されているか否かを判断する動作が行われる。ブロック320において、高速フォーマットインジケータが設定されている場合、データ記憶媒体アレイの初期化中に読み取られるデータ記憶媒体のいずれのブロック（ドライブのメンバーシップメタデータスタンプエリア等）にもまず書き込みを行うステップが行われる。ブロック330において、高速フォーマットインジケータを変更する動作が行われる。ブロック340において、その後、すべての書き込みが最大読み取りサイズ以上のサイズで行われるライト・ビフォア・リード・モデルに従う。

#### 【0015】

一実施形態では、データ記憶媒体アレイはRAIDデータ記憶媒体アレイを含むことができ、データ記憶媒体はハードドライブを含むことができる。所望のブロックフォーマットは520バイトであることができ、現在のブロックフォーマットは512バイトであることができる。

#### 【0016】

図4を参照して、一実施形態の動作の論理フローを説明する。ステップ400において、システムに追加すべきドライブのブロックサイズが何であるかが求められる。ステップ402において、ブロックサイズが512である場合、高速ブロックフォーマットが設定されているか否かが判断される。設定されている場合、このビットは設定されていたが、フォーマットコマンドは送られていなかった、または失敗したため、ステップ404においてエラー状態が設定される。エラー状態は、一実施形態では、0バイトブロックサイズを報告して不首尾フォーマットを示すことができる。高速ブロックフォーマットが設定されていない場合、ステップ406において、高速フォーマットビットがイネーブルされる。ステップ408において、ブロックサイズテンプレートが520に設定される。ステップ410において、フォーマットユニットコマンドがドライブに送られる。ステップ412において、高速フォーマットが完了したか否かが判断される。完了している場合、メタデータが、書き込み前にまず読み取られるドライブのエリアに書き込まれる。次いで、ステップ416において、高速フォーマットビットがディセーブルされ、ドライブは520バイトブロックと併用する準備が整う。

#### 【0017】

別法として、アレイサイズがすでに520であると判断される場合、ステップ420において、高速フォーマットビットが設定されているか否かが判断される。設定されている場合、エラー状態がブロック424において設定され、書き込み前に読み取るべきメタデータがドライブに書き込まれていなかった、または高速フォーマットビットをリセット前にクリアできなかったことを示す。高速フォーマットビットが設定されていない場合、ステップ422において、ドライブが適宜フォーマットされ、メタデータを読み取ることができ、ドライブが520バイトブロックと併用する準備が整っていることが列挙される。

#### 【0018】

本発明の実施形態のいくつかの利点は時間の節減である。製造する際に、本発明を使用して、システムを出荷に向けて準備するのにかかる時間に影響を及ぼすことができる。たとえば、（製造時にシステムにオートフォーマットを準備する時間） - （システムに高速オートフォーマットを準備する時間） = 時間の節減、（時間の節減） × （1分当たりの動作コスト） = 総節減。

#### 【0019】

別法として、ドライブをフォーマットする際に実現される現場での節減は、以下のよう

10

20

30

40

50

に説明することができる。（ディスクをオートフォーマットする時間） - （ディスクを高速オートフォーマットする時間） = 時間節減、（時間節減） × （1分当たりのカスタマーエンジニアコスト） = 総節減。

#### 【0020】

一例として、（73ギガフォーマット40分） - （73ギガ高速フォーマット3分） ~ = 37分、（37分） × （\$300 / 60分） = （1訪問 / 取り付け）当たりの総節減 \$ 185。いくつかの実施形態では、ディスクのフォーマットに必要な時間量は容量に伴って線形に増減することがわかっている。将来には、ドライブ容量が伸びるにつれ（たとえば、ディスク容量が2倍の146GBになると、低速フォーマットに80分、高速フォーマットに6分かかる）、本発明のいくつかの実施形態によって生み出される節減量は線形に増大することができる（1訪問当たり \$185 から \$370 に）。製品の寿命にわたり、ディスクを現場で交換する、または新しいディスクを追加する場合、1訪問当たりのこのコストは、総保証コストに対して大きな影響を有することができる。

#### 【0021】

本明細書に提供するフローチャートは特定の順序の方法ステップを示すが、これらステップの順序は図示の順序と異なってもよいことが理解されることに留意されたい。また、2つ以上のステップを同時に、または一部同時にを行うことも可能である。このような変形は、主に設計者の選択で選択されるソフトウェアおよびハードウェアシステムに依存する。このような変形はすべて本発明の範囲内であることが理解される。同様に、本発明のソフトウェアおよびウェブ実施は、ルールベースの論理および他の論理を使用して様々なデータベース探索ステップ、関連付けステップ、比較ステップ、および判断ステップを実現する標準のプログラミング技法を使用して実現することができる。本明細書および特許請求の範囲において使用される「構成要素」という言葉は、ソフトウェアコードの1つまたは複数のライン、および / またはハードウェア実施態様、および / または手動入力を受け取る機器を使用する実施態様を包含することにも留意されたい。

#### 【0022】

本発明の好ましい実施形態の上記説明は、例示および説明を目的として提示された。網羅的である、すなわち本発明を開示された厳密な形態に限定する意図はなく、上記教示に鑑みて変形形態および変形形態が可能であるか、または本発明の実施から得ることができる。実施形態は、当業者が意図する特定の用途に適するよう様々な実施形態で、また様々な変更を行って本発明を利用することができるよう、本発明の原理および本発明の実際の応用を説明するために選択され説明された。本発明の範囲は本明細書に添付の特許請求の範囲およびその等価物に規定されることを意図する。

#### 【図面の簡単な説明】

#### 【0023】

【図1】本発明の一実施形態によるデータ記憶媒体アレイ内にある、またはデータ記憶媒体アレイに追加すべきデータ記憶媒体をロックフォーマットする方法のフローチャートである。

【図2】トポロジマネージャと併せて記憶媒体のロックを示す概略ブロック図である。

【図3】本発明のさらなる一実施形態によるデータ記憶媒体アレイ内にある、またはデータ記憶媒体アレイに追加すべきデータ記憶媒体をロックフォーマットする方法のフローチャートである。

【図4】本発明の全体的な動作の一実施形態のフローチャートである。

#### 【符号の説明】

#### 【0024】

200：データ記憶媒体アレイ

210：データ記憶媒体（ドライブ）

220：トポロジマネージャ

230：リンク

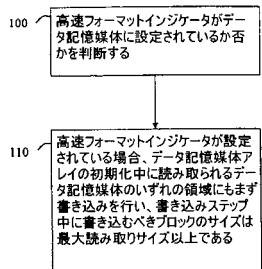
10

20

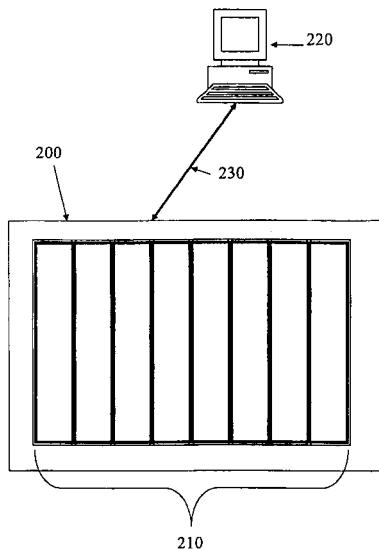
30

40

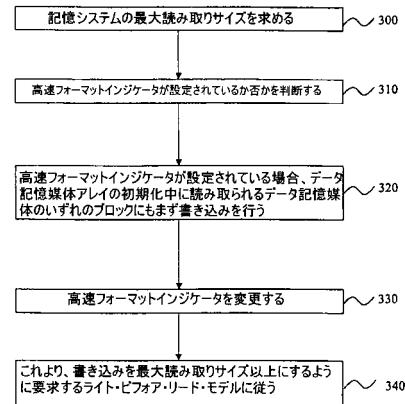
【図1】



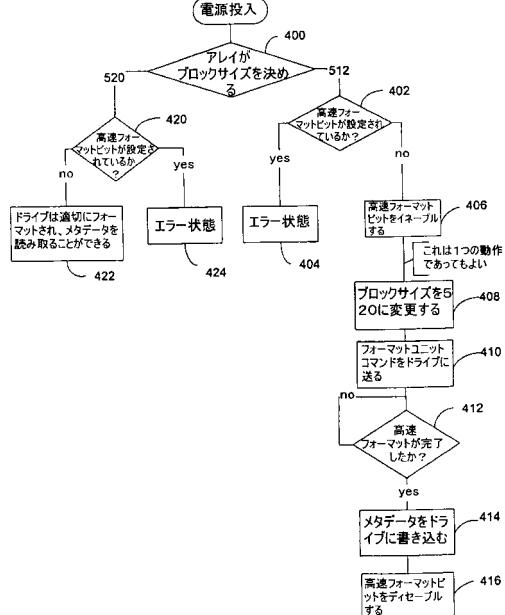
【図2】



【図3】



【図4】



## フロントページの続き

(72)発明者 デイヴィッド・クリストファー・ダルマン

アメリカ合衆国アイダホ州83642,メリディアン,ブラウンフィールド・ドライブ 621

(72)発明者 アンドリュー・ヒルズ・ダルマン

アメリカ合衆国アイダホ州83642,メリディアン,ノース・メンデルソン・アヴェニュー 5

309

F ターム(参考) 5B065 BA01 CA30 CH20

5D044 BC01 CC01 CC04 DE02