

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2017-509998

(P2017-509998A)

(43) 公表日 平成29年4月6日(2017.4.6)

(51) Int.Cl.

G06F 12/08 (2016.01)

F 1

G06F 12/08

G06F 12/08

テーマコード(参考)

5 B 005

543B

審査請求 未請求 予備審査請求 未請求 (全 28 頁)

(21) 出願番号 特願2016-559352 (P2016-559352)
 (86) (22) 出願日 平成27年4月2日 (2015.4.2)
 (85) 翻訳文提出日 平成28年9月26日 (2016.9.26)
 (86) 國際出願番号 PCT/US2015/024030
 (87) 國際公開番号 WO2015/153855
 (87) 國際公開日 平成27年10月8日 (2015.10.8)
 (31) 優先権主張番号 14/245,356
 (32) 優先日 平成26年4月4日 (2014.4.4)
 (33) 優先権主張国 米国(US)

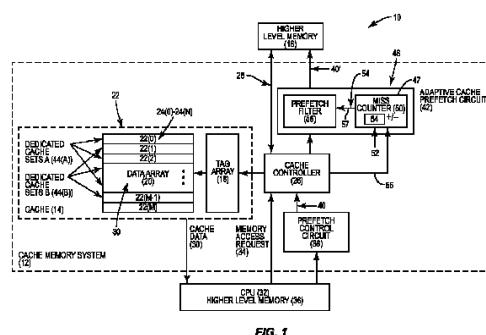
(71) 出願人 507364838
 クアルコム、インコーポレイテッド
 アメリカ合衆国 カリフォルニア 921
 21 サンディエゴ モアハウス ドラ
 イブ 5775
 (74) 代理人 100108453
 弁理士 村山 靖彦
 (74) 代理人 100163522
 弁理士 黒田 晋平
 (72) 発明者 ハロルド・ウェイド・ケイン・ザ・サード
 アメリカ合衆国・カリフォルニア・921
 21・サン・ディエゴ・モアハウス・ドラ
 イブ・5775

最終頁に続く

(54) 【発明の名称】キャッシュ汚染を低減するために専用キャッシュセットにおける専用プリフェッチポリシーを競合させることに基づいた適応キャッシュプリフェッチング

(57) 【要約】

キャッシュ汚染を低減するために専用キャッシュセットにおける専用プリフェッチポリシーを競合させることに基づいた適応キャッシュプリフェッチングが開示される。一態様では、適応キャッシュプリフェッチ回路は、データをキャッシュにプリフェッチするために提供される。適応キャッシュプリフェッチ回路は、キャッシュ内の専用キャッシュセットに適用される専用プリフェッチポリシーを競合させたことに基づいて、どのプリフェッチポリシーを置換ポリシーとして使用するかを決定するように構成される。各専用キャッシュセットは、所与の専用キャッシュセットのための、置換ポリシーとして使用される関連する専用プリフェッチポリシーを有する。専用キャッシュセットの各々へのアクセスについてのキャッシュミスは、適応キャッシュプリフェッチ回路によって追跡される。適応キャッシュプリフェッチ回路は、キャッシュ汚染を低減するために、そのそれぞれの専用キャッシュセットに対してより少ないキャッシュミスを生じた専用プリフェッチポリシーを使用して、プリフェッチポリシーをキャッシュ内の他のフォロワ(すな



【特許請求の範囲】**【請求項 1】**

キャッシュデータをキャッシュにプリフェッチするための適応キャッシュプリフェッチ回路であって、

少なくとも1つの第1の専用プリフェッチポリシーが適用されるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、前記少なくとも1つの第1の専用プリフェッチポリシーとは異なる少なくとも1つの第2の専用プリフェッチポリシーが適用される前記キャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、少なくとも1つのミス状態を更新するように構成されたミス追跡回路と、

前記ミス追跡回路の前記少なくとも1つのミス状態に基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中からプリフェッチポリシーを選択するように構成されたプリフェッチフィルタとを備える適応キャッシュプリフェッチ回路。

【請求項 2】

前記プリフェッチフィルタが、プリフェッチ制御回路によって発行された、前記キャッシュを充填させるためのプリフェッチ要求に適用されるべき前記プリフェッチポリシーを選択するようにさらに構成される、請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項 3】

前記少なくとも1つの第1の専用プリフェッチポリシーが、第1の専用プリフェッチポリシーから成り、

前記少なくとも1つの第2の専用プリフェッチポリシーが、第2の専用プリフェッチポリシーから成り、

前記プリフェッチフィルタが、前記ミス追跡回路の前記少なくとも1つのミス状態に基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から前記プリフェッチポリシーを選択するように構成される、

請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項 4】

前記第1の専用プリフェッチポリシーが、非プリフェッチポリシーから成り、

前記第2の専用プリフェッチポリシーが、常時プリフェッチポリシーから成る、請求項3に記載の適応キャッシュプリフェッチ回路。

【請求項 5】

前記ミス追跡回路が、少なくとも1つのミスカウンタから成り、前記少なくとも1つのミス状態が、少なくとも1つのミスカウントから成り、

前記少なくとも1つのミスカウンタが、前記少なくとも1つの第1の専用キャッシュセットおよび前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記少なくとも1つのミスカウントを更新するように構成され、

前記プリフェッチフィルタが、前記少なくとも1つのミスカウンタの前記少なくとも1つのミスカウントに基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から前記プリフェッチポリシーを選択するように構成される、

請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項 6】

前記ミス追跡回路が、ミス飽和インジケータから成り、前記少なくとも1つのミス状態が、ミス状態から成り、

前記ミス飽和インジケータが、前記少なくとも1つの第1の専用キャッシュセットおよび前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス状態を更新するように構成

され、

前記プリフェッチフィルタが、前記ミス飽和インジケータの前記ミス状態に基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から前記プリフェッチポリシーを選択するように構成される、

請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項7】

前記ミス飽和インジケータが、ミス飽和カウンタから成り、前記ミス状態が、ミス飽和カウントから成り、

前記ミス飽和カウンタが、前記少なくとも1つの第1の専用キャッシュセットおよび前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス飽和カウントを更新するように構成され、

前記プリフェッチフィルタが、前記ミス飽和カウンタの前記ミス飽和カウントに基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から前記プリフェッチポリシーを選択するように構成される、

請求項6に記載の適応キャッシュプリフェッチ回路。

【請求項8】

前記ミス飽和カウンタが、

前記少なくとも1つの第1の専用プリフェッチポリシーが適用される前記キャッシュ内の前記少なくとも1つの第1の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス飽和カウントを増分または減分することによって、前記ミス飽和カウントを更新し、

前記少なくとも1つの第1の専用プリフェッチポリシーとは異なる前記少なくとも1つの第2の専用プリフェッチポリシーが適用される前記キャッシュ内の前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス飽和カウントをそれぞれ減分または増分することによって、前記ミス飽和カウントを更新する

ように構成されることによって、前記ミス飽和カウントを更新するように構成される、請求項7に記載の適応キャッシュプリフェッチ回路。

【請求項9】

前記ミス追跡回路が、各々がミス状態を含む複数のミスインジケータから成り、前記複数のミスインジケータの各々が、前記少なくとも1つの第1の専用キャッシュセットおよび前記少なくとも1つの第2の専用キャッシュセットの中の専用キャッシュセットに関連付けられ、

前記複数のミスインジケータがそれぞれ、前記キャッシュ内の前記少なくとも1つの第1の専用キャッシュセットおよび前記少なくとも1つの第2の専用キャッシュセットの中の前記専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記関連するミス状態を更新するようにさらに構成され、

前記プリフェッチフィルタが、前記複数のミスインジケータにおける前記少なくとも1つのミス状態の比較に基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から前記プリフェッチポリシーを選択するように構成される、

請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項10】

前記プリフェッチフィルタが、前記ミス追跡回路の前記少なくとも1つのミス状態に基づいて、選択的に前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から前記プリフェッチポリシーを選択しないようにさらに構成される、請求項1に記載の適応キャッシュプリフェッチ回路。

10

20

30

40

50

【請求項 1 1】

前記プリフェッチフィルタが、前記ミス飽和カウンタの前記ミス飽和カウント内の少なくとも1つの有効ビットに基づいて、選択的に前記少なくとも1つの第1の専用プリフェッチポリシーおよび前記少なくとも1つの第2の専用プリフェッチポリシーの中から、プリフェッチ制御回路によって発行されたプリフェッチ要求に適用されるべき前記プリフェッチポリシーを選択しないようにさらに構成される、請求項7に記載の適応キャッシュプリフェッチ回路。

【請求項 1 2】

前記プリフェッチフィルタが、前記少なくとも1つの第1の専用プリフェッチポリシーまたは前記少なくとも1つの第2の専用プリフェッチポリシーを常に選択しないようにさらに構成される、請求項1に記載の適応キャッシュプリフェッチ回路。

10

【請求項 1 3】

前記プリフェッチフィルタが、

前記ミス追跡回路の前記少なくとも1つのミス状態に基づいて、前記少なくとも1つの第1の専用プリフェッチポリシーまたは前記少なくとも1つの第2の専用プリフェッチポリシーが、プリフェッチ制御回路によって発行されたプリフェッチ要求に適用されるべきかどうかを確率的に判断し、

前記確率的判断に基づいて、前記プリフェッチ制御回路によって発行された前記プリフェッチ要求に適用されるべき前記少なくとも1つの第1の専用プリフェッチポリシーまたは前記少なくとも1つの第2の専用プリフェッチポリシーを選択する

20

ようにさらに構成される、請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項 1 4】

各々が1つまたは複数のキャッシュエントリを記憶するように構成された複数のキャッシュセットを含む前記キャッシュであって、前記複数のキャッシュセットが、

前記少なくとも1つの第1の専用プリフェッチポリシーに基づいて、プリフェッチされたキャッシュデータを受信するように構成された前記少なくとも1つの第1の専用キャッシュセットと、

前記少なくとも1つの第2の専用プリフェッチポリシーに基づいて、前記プリフェッチされたキャッシュデータを受信するように構成された前記少なくとも1つの第2の専用キャッシュセットと、

30

前記少なくとも1つの第1の専用プリフェッチポリシーまたは前記少なくとも1つの第2の専用プリフェッチポリシーのいずれかに基づいて、前記プリフェッチされたキャッシュデータを受信するように構成された少なくとも1つのフォロワキャッシュセットと

を含む、前記キャッシュと、

メモリアドレスを含むメモリアクセス要求を受信し、前記メモリアドレスに対応するキャッシュエントリが前記キャッシュに含まれるかどうかを判断するように構成されたキャッシュコントローラと、

前記プリフェッチポリシーに従って前記プリフェッチされたキャッシュデータを前記キャッシュ内の前記複数のキャッシュセットにプリフェッチするためのプリフェッチ要求を発行するように構成されたプリフェッチ制御回路と

40

を備える、請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項 1 5】

前記プリフェッチフィルタが、前記キャッシュコントローラの外部に配設される、請求項14に記載の適応キャッシュプリフェッチ回路。

【請求項 1 6】

前記キャッシュコントローラが、前記プリフェッチフィルタを備える、請求項14に記載の適応キャッシュプリフェッチ回路。

【請求項 1 7】

集積回路(IC)の中に配設される、請求項1に記載の適応キャッシュプリフェッチ回路。

【請求項 1 8】

50

セットトップボックス、エンターテインメントユニット、ナビゲーションデバイス、通信デバイス、固定ロケーションデータユニット、モバイルロケーションデータユニット、モバイルフォン、セルラーフォン、コンピュータ、ポータブルコンピュータ、デスクトップコンピュータ、携帯情報端末(PDA)、モニタ、コンピュータモニタ、テレビジョン、チューナー、ラジオ、衛星ラジオ、音楽プレーヤ、デジタル音楽プレーヤ、ポータブル音楽プレーヤ、デジタルビデオプレーヤ、ビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、およびポータブルデジタルビデオプレーヤから成る群から選択されたデバイスに統合される、請求項1に記載の適応キャッシュプリフェッヂ回路。

【請求項19】

キャッシュデータをキャッシュにプリフェッヂするための適応キャッシュプリフェッヂ回路であって、

少なくとも1つの第1の専用プリフェッヂポリシーが適用されるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、前記少なくとも1つの第1の専用プリフェッヂポリシーとは異なる少なくとも1つの第2の専用プリフェッヂポリシーが適用される前記キャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、少なくとも1つのミス状態手段を更新するためのミス追跡手段と、

前記ミス追跡手段の前記少なくとも1つのミス状態手段に基づいて、前記少なくとも1つの第1の専用プリフェッヂポリシーおよび前記少なくとも1つの第2の専用プリフェッヂポリシーの中からプリフェッヂポリシーを選択するためのプリフェッヂフィルタ手段とを備える適応キャッシュプリフェッヂ回路。

【請求項20】

専用キャッシュセットにおける専用プリフェッヂポリシーを競合させることに基づいた適応キャッシュプリフェッヂングの方法であって、

キャッシュ内でアドレス指定されるべきメモリアドレスを含むメモリアクセス要求を受信するステップと、

前記メモリアドレスに対応する、前記キャッシュ内の複数のキャッシュエントリの中のアクセスされたキャッシュエントリが、前記キャッシュに含まれるかどうかを判断することによって、前記メモリアクセス要求がキャッシュミスであるかどうかを判断するステップと、

少なくとも1つの第1の専用プリフェッヂポリシーが適用される前記キャッシュ内の少なくとも1つの第1の専用キャッシュセットと、前記少なくとも1つの第1の専用プリフェッヂポリシーとは異なる少なくとも1つの第2の専用プリフェッヂポリシーが適用される前記キャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、ミス追跡回路の少なくとも1つのミス状態を更新するステップと、

キャッシュデータを前記キャッシュ内の複数のキャッシュセットの中のフォロワキャッシュセットにおけるキャッシュエントリにプリフェッヂするためのプリフェッヂ要求を発行するステップと、

前記ミス追跡回路の前記少なくとも1つのミス状態に基づいて、前記少なくとも1つの第1の専用プリフェッヂポリシーおよび前記少なくとも1つの第2の専用プリフェッヂポリシーの中から、前記プリフェッヂ要求に適用されるべきプリフェッヂポリシーを選択するステップと、

前記選択されたプリフェッヂポリシーに基づいて、前記プリフェッヂされたキャッシュデータを前記フォロワキャッシュセットにおける前記キャッシュエントリに充填するステップと

を含む方法。

【請求項21】

前記ミス追跡回路の前記少なくとも1つのミス状態を更新するステップが、非プリフェッヂポリシーが適用される前記キャッシュ内の前記少なくとも1つの第1の専

10

20

30

40

50

用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス追跡回路の前記少なくとも1つのミス状態を更新するステップと、

常時プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス追跡回路の前記少なくとも1つのミス状態を更新するステップと

を含む、請求項20に記載の方法。

【請求項 2 2】

前記ミス追跡回路の前記少なくとも1つのミス状態を更新するステップが、前記少なくとも1つの第1の専用プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第1の専用キャッシュセットと、前記少なくとも1つの第1の専用プリフェッчポリシーとは異なる前記少なくとも1つの第2の専用プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第2の専用キャッシュセットとにおける、前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、少なくとも1つのミスカウンタの少なくとも1つのミスカウントを更新するステップを含み、

前記プリフェッчポリシーを選択するステップが、前記少なくとも1つのミスカウンタの前記少なくとも1つのミスカウントに基づいて、前記少なくとも1つの第1の専用プリフェッчポリシーおよび前記少なくとも1つの第2の専用プリフェッчポリシーの中から、前記プリフェッч要求に適用されるべき前記プリフェッчポリシーを選択するステップを含む、

請求項20に記載の方法。

【請求項 2 3】

前記少なくとも1つのミスカウンタの前記少なくとも1つのミスカウントを更新するステップが、前記少なくとも1つの第1の専用プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第1の専用キャッシュセットと、前記少なくとも1つの第1の専用プリフェッчポリシーとは異なる前記少なくとも1つの第2の専用プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第2の専用キャッシュセットとにおける、前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、少なくとも1つのミス飽和カウンタの少なくとも1つのミス飽和カウントを更新するステップを含み、

前記プリフェッчポリシーを選択するステップが、前記少なくとも1つのミス飽和カウンタの前記少なくとも1つのミス飽和カウントに基づいて、前記少なくとも1つの第1の専用プリフェッчポリシーおよび前記少なくとも1つの第2の専用プリフェッчポリシーの中から、前記プリフェッч要求に適用されるべき前記プリフェッчポリシーを選択するステップを含む、

請求項22に記載の方法。

【請求項 2 4】

前記少なくとも1つのミス飽和カウンタの前記少なくとも1つのミス飽和カウントを更新するステップが、

前記少なくとも1つの第1の専用プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第1の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記少なくとも1つのミス飽和カウンタの前記少なくとも1つのミス飽和カウントを増分または減分するステップと、

前記少なくとも1つの第1の専用プリフェッчポリシーとは異なる前記少なくとも1つの第2の専用プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記少なくとも1つのミス飽和カウンタの前記少なくとも1つのミス飽和カウントをそれぞれ減分または増分するステップと

を含む、請求項23に記載の方法。

10

20

30

40

50

【請求項 25】

前記選択されたプリフェッчポリシーとしての前記少なくとも1つの第1の専用プリフェッчポリシーまたは前記選択されたプリフェッчポリシーとしての前記少なくとも1つの第2の専用プリフェッчポリシーを無視するステップをさらに含む、請求項20に記載の方法。

【請求項 26】

前記少なくとも1つの第1の専用プリフェッчポリシーまたは前記少なくとも1つの第2の専用プリフェッчポリシーが前記選択されたプリフェッчポリシーとして選択されるべきかどうかを確率的に判断するステップをさらに含み、

前記プリフェッчされたキャッシュデータを充填するステップが、前記確率的に判断されたプリフェッчポリシーに基づいて、前記プリフェッчされたキャッシュデータを前記フォロワキャッシュセットにおける前記キャッシュエントリに充填するステップを含む、請求項20に記載の方法。

10

【請求項 27】

プロセッサベースの適応キャッシュプリフェッч回路に、

少なくとも1つの第1の専用プリフェッчポリシーが適用されるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、前記少なくとも1つの第1の専用プリフェッчポリシーとは異なる少なくとも1つの第2の専用プリフェッчポリシーが適用される前記キャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、ミス追跡回路の少なくとも1つのミス状態を更新し、

20

前記ミス追跡回路の前記少なくとも1つのミス状態に基づいて、前記少なくとも1つの第1の専用プリフェッчポリシーおよび前記少なくとも1つの第2の専用プリフェッчポリシーの中から、プリフェッч制御回路によって発行された、前記キャッシュを充填させるためのプリフェッч要求に適用されるべきプリフェッчポリシーを選択することによって、キャッシュデータを前記キャッシュにプリフェッчさせるコンピュータ実行可能命令を記憶した非一時的コンピュータ可読記憶媒体。

20

【請求項 28】

非プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第1の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス追跡回路の前記少なくとも1つのミス状態を更新し、

30

常時プリフェッчポリシーが適用される前記キャッシュ内の前記少なくとも1つの第2の専用キャッシュセットにおける前記アクセスされたキャッシュエントリから生じる前記キャッシュミスに基づいて、前記ミス追跡回路の前記少なくとも1つのミス状態を更新することによって前記ミス追跡回路の前記少なくとも1つのミス状態を更新することによって、前記プロセッサベースの適応キャッシュプリフェッч回路に、キャッシュデータを前記キャッシュにプリフェッчさせる前記コンピュータ実行可能命令を記憶した、請求項27に記載の非一時的コンピュータ可読記憶媒体。

【請求項 29】

前記選択されたプリフェッчポリシーとしての前記少なくとも1つの第1の専用プリフェッчポリシーまたは前記選択されたプリフェッчポリシーとしての前記少なくとも1つの第2の専用プリフェッчポリシーを無視することによって、前記プロセッサベースの適応キャッシュプリフェッч回路に、キャッシュデータを前記キャッシュにプリフェッчさせる前記コンピュータ実行可能命令を記憶した、請求項27に記載の非一時的コンピュータ可読記憶媒体。

40

【発明の詳細な説明】**【技術分野】****【0001】****優先権主張**

本出願は、その全体が参照により本明細書に組み込まれる、2014年4月4日に出願された

50

「ADAPTIVE CACHE PREFETCHING BASED ON COMPETING DEDICATED PREFETCH POLICIES IN DEDICATED CACHE SETS TO REDUCE CACHE POLLUTION」と題する米国特許出願第14/245,356号の優先権を主張する。

【0002】

本開示の技術は一般に、コンピュータシステムにおいて提供されるキャッシュメモリに関し、より詳細には、キャッシュミスを低減するためにキャッシュラインをキャッシュメモリにプリフェッチすることに関する。

【背景技術】

【0003】

メモリセルは、「メモリ」としても知られるコンピュータデータストレージの基本ビルディングブロックである。コンピュータシステムは、メモリからデータを読み出すかまたはデータをメモリに書き込むことができる。メモリは、一例として、中央処理装置(CPU)システムにおいてキャッシュメモリを提供するために使用され得る。単に「キャッシュ」とも呼ばれ得るキャッシュメモリは、メモリアクセスレイテンシを低減するために、頻繁にアクセスされるメモリアドレスにおいて記憶されたデータのコピーをメインメモリまたは上位レベルキャッシュメモリに記憶する、より小型でより高速のメモリである。したがって、キャッシュは、メモリアクセス時間を低減するためにCPUによって使用され得る。たとえば、キャッシュは、より高速な命令実行のために、CPUによってフェッチされた命令を記憶するために使用されてもよい。別の例として、キャッシュは、より高速なデータアクセスのために、CPUによってフェッチされるべきデータを記憶するために使用されてもよい。

10

20

30

30

【0004】

キャッシュは、タグアレイおよびデータアレイから成る。タグアレイは、「タグ」としても知られるアドレスを含む。タグは、データアレイ内のデータ記憶ロケーションに対するインデックスを提供する。タグアレイ内のタグおよびデータアレイ内のタグのインデックスにおいて記憶されたデータは、「キャッシュライン」または「キャッシュエントリ」としても知られる。メモリアクセス要求の一部として、キャッシュに対するインデックスとして提供されるメモリアドレスまたはその部分がタグアレイ内のタグと一致する場合、これは「キャッシュヒット」として知られる。キャッシュヒットは、一致するタグのインデックスにおいて含まれるデータアレイ内のデータが、メインメモリおよび/または上位レベルキャッシュ内の要求されたメモリアドレスに対応するデータを含むことを意味する。一致するタグのインデックスにおけるデータアレイに含まれるデータは、より大きいメモリアクセスレイテンシを有するメインメモリまたは上位レベルキャッシュメモリにアクセスしなければならないことに対立するものとして、メモリアクセス要求に使用され得る。しかしながら、メモリアクセス要求のためのインデックスがタグアレイ内のタグと一致しない場合、またはキャッシュラインが他の点で無効である場合、これは「キャッシュミス」として知られる。キャッシュミスでは、データアレイは、メモリアクセス要求を満たすことができるデータを含まないと見なされる。

【0005】

キャッシュにおけるキャッシュミスは、様々なコンピュータシステムで動作する多くのアプリケーションの性能低下の実質的な原因である。キャッシュミスの数を低減するために、コンピュータシステムは、プリフェッチャとしても知られるプリフェッチエンジンを利用することができる。プリフェッチャは、将来のメモリアクセスを予測するために、コンピュータシステムにおけるメモリアクセスパターンを検出するように構成され得る。これらの予測を使用して、プリフェッチャは、キャッシュラインをキャッシュに投機的にプリロードするための要求を上位レベルメモリに対して行う。したがって、これらのキャッシュラインが必要とされるとき、これらのキャッシュラインはキャッシュにすでに存在しており、結果としてキャッシュミスペナルティが生じない。

40

【0006】

多くのアプリケーションがプリフェッチングから恩恵を受けるが、いくつかのアプリケ

50

ーションは、予測するのが困難であるメモリアクセスパターンを有する。これらのアプリケーションについてプリフェッティングを有効にすることは、結果として性能を著しく低減する場合がある。これらの場合、プリフェッチャは、キャッシュラインがアプリケーションによって決して使用され得ないキャッシュに充填されるよう要求してもよい。さらに、キャッシュにおいてプリフェッチされたキャッシュラインのための場所をあけるために、次いで有用なキャッシュラインが取って代わられる場合がある。以前に取って代わられたキャッシュラインがアクセスされる前にプリフェッチされたキャッシュラインが後でアクセスされない場合、以前に取って代わられたキャッシュラインへのアクセスについてキャッシュミスが生成される。このシナリオにおけるキャッシュミスは事実上、プリフェッチ動作によって引き起こされたものであった。後でアクセスされるキャッシュラインが参照されないプリフェッチされたキャッシュラインに取って代わられるプロセスは、「キャッシュ汚染」と呼ばれる。キャッシュ汚染はキャッシュミス率を増加させる場合があり、このことは性能を低下させる。

10

【0007】

キャッシュラインをキャッシュにプリフェッチした結果としてのキャッシュ汚染を制限しようするために、様々なキャッシュデータ置換ポリシー（「プリフェッチポリシー」と呼ばれる）が存在する。たとえば、1つのキャッシュプリフェッチポリシーは、プリフェッチャによってキャッシュにプリフェッチされるキャッシュラインの数を動的に調整するために、プリフェッチ精度、遅延、および汚染レベルなどの様々なメトリックを追跡する。しかしながら、そのようなメトリックを追跡することは、コンピュータシステムにおける追加のハードウェアオーバーヘッドを必要とする。たとえば、参照ビットがキャッシュ内のキャッシュウェイごとに追加され得るおよび/またはブルームフィルタがキャッシュにおいて利用され得る。別のキャッシュプリフェッチポリシーは、キャッシュ汚染を制限するために、所望の時間枠の中でアクセスされなかったキャッシュ内のデッドキャッシュラインのみを、プリフェッチされたキャッシュデータに置き換える。デッドラインではない、したがって有用なデータを含むキャッシュラインは、キャッシュミスを低減するためにキャッシュから追い出されない。しかしながら、このデッドラインのみを置換するキャッシュプリフェッチポリシーは、キャッシュ内のキャッシュラインへのアクセスのタイミングを追跡するためにハードウェアオーバーヘッドを追加する。

20

【発明の概要】

30

【発明が解決しようとする課題】

【0008】

したがって、プリフェッティングの性能利点を低減すること、および電力消費を増大させる場合がある大幅な追加のハードウェアオーバーヘッドを生じることなしに、キャッシュにおいてキャッシュ汚染を制限するキャッシュデータのプリフェッティングを提供することが望まれる。

【課題を解決するための手段】

【0009】

発明を実施するための形態で開示する態様は、キャッシュ汚染を低減するために専用キャッシュセットにおける専用プリフェッチポリシーを競合させることに基づいた適応キャッシュプリフェッティングを含む。一態様では、適応キャッシュプリフェッチ回路は、データをキャッシュにプリフェッチするために提供される。キャッシュのための最適な置換ポリシーを決定しようとする代わりに、適応キャッシュプリフェッチ回路は、キャッシュ内の専用キャッシュセットに適用される専用プリフェッチポリシーを競合させた結果に基づいて、どのプリフェッチポリシーを使用するかを決定するように構成される。この点について、キャッシュ内のキャッシュセットのサブセットは、「専用」キャッシュセットであるものとして割り振られる。その他の非専用キャッシュセットは、「フォロワ」キャッシュセットである。各専用キャッシュセットは、所与の専用キャッシュセットのための関連する専用プリフェッチポリシーを有する。専用キャッシュセットの各々へのアクセスについてのキャッシュミスは、適応キャッシュプリフェッチ回路によって追跡される。適応キ

40

50

ヤッショウプリフェッチ回路は、そのそれぞれの専用キャッシュセットに対してより少ないキャッシュミスを生じた専用プリフェッチポリシーを使用して、プリフェッチポリシーをキャッシュ内の他のフォロワーキャッシュセットに適用するように構成され得る。たとえば、ある専用プリフェッチポリシーは決してプリフェッチしないものとしてよく、別の専用プリフェッチポリシーはデュエリング専用プリフェッチポリシーをキャッシュに提供するために常にプリフェッチするものとしてよい。このようにして、キャッシュ内の専用キャッシュセットに対する実際のキャッシュミス結果は、フォロワーキャッシュセットのためのプリフェッチポリシーとして使用される場合にどの専用プリフェッチポリシーがキャッシュ内により少ないキャッシュ汚染をもたらすことになるかについてのより良い指標であり得るので、キャッシュ汚染が低減され得る。キャッシュ汚染の低減は、性能の向上、メモリ競合の低減、およびキャッシュによる電力消費の減少をもたらすことができる。

10

20

30

40

50

【0010】

この点について、一態様では、キャッシュデータをキャッシュにプリフェッチするための適応キャッシュプリフェッチ回路が提供される。適応キャッシュプリフェッチ回路は、少なくとも1つの第1の専用プリフェッチポリシーが適用されるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、少なくとも1つの第1の専用プリフェッチポリシーとは異なる少なくとも1つの第2の専用プリフェッチポリシーが適用されるキャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、少なくとも1つのミス状態を更新するように構成されたミス追跡回路を備える。一例では、ミス追跡回路は、少なくとも1つの第1の専用キャッシュセットと少なくとも1つの第2の専用キャッシュセットの両方についてのキャッシュミスを追跡するために、単一のミス状態として少なくとも1つのミス状態を提供する場合がある。別の例として、ミス追跡回路は、少なくとも1つの第1の専用キャッシュセットおよび少なくとも1つの第2の専用キャッシュセットの各々についてのキャッシュミスを別個に追跡するために、少なくとも1つの第1の専用キャッシュセットおよび少なくとも1つの第2の専用キャッシュセットの各々について別個のミス状態を含む場合がある。適応キャッシュプリフェッチ回路は、プリフェッチフィルタをさらに備える。プリフェッチフィルタは、ミス追跡回路の少なくとも1つのミス状態に基づいて、少なくとも1つの第1の専用プリフェッチポリシーおよび少なくとも1つの第2の専用プリフェッチポリシーの中からプリフェッチポリシーを選択するように構成される。

【0011】

別の態様では、キャッシュデータをキャッシュにプリフェッチするための適応キャッシュプリフェッチ回路が提供される。適応キャッシュプリフェッチ回路は、少なくとも1つの第1の専用プリフェッチポリシーが適用されるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、少なくとも1つの第1の専用プリフェッチポリシーとは異なる少なくとも1つの第2の専用プリフェッチポリシーが適用されるキャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、少なくとも1つのミス状態手段を更新するためのミス追跡手段を備える。適応キャッシュプリフェッチ回路はまた、ミス追跡手段の少なくとも1つのミス状態手段に基づいて、少なくとも1つの第1の専用プリフェッチポリシーおよび少なくとも1つの第2の専用プリフェッチポリシーの中からプリフェッチポリシーを選択するためのプリフェッチフィルタ手段を備える。

【0012】

別の態様では、専用キャッシュセットにおける専用プリフェッチポリシーを競合させることに基づいた適応キャッシュプリフェッチングの方法が提供される。方法は、キャッシュ内でアドレス指定されるべきメモリアドレスを含むメモリアクセス要求を受信するステップを含む。方法はまた、メモリアドレスに対応する、キャッシュ内の複数のキャッシュエントリの中のアクセスされたキャッシュエントリが、キャッシュに含まれるかどうかを判断することによって、メモリアクセス要求がキャッシュミスであるかどうかを判断するステップを含む。方法はまた、少なくとも1つの第1の専用プリフェッチポリシーが適用さ

れるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、少なくとも1つの第1の専用プリフェッчポリシーとは異なる少なくとも1つの第2の専用プリフェッчポリシーが適用されるキャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、ミス追跡回路の少なくとも1つのミス状態を更新するステップを含む。方法はまた、キャッシュデータをキャッシュ内の複数のキャッシュセットの中のフォロワキャッシュセットにおけるキャッシュエントリにプリフェッчするためのプリフェッч要求を発行するステップを含む。方法はまた、ミス追跡回路の少なくとも1つのミス状態に基づいて、少なくとも1つの第1の専用プリフェッчポリシーおよび少なくとも1つの第2の専用プリフェッчポリシーの中から、プリフェッч要求に適用されるべきプリフェッчポリシーを選択するステップを含む。方法はまた、選択されたプリフェッчポリシーに基づいて、プリフェッчされたキャッシュデータをフォロワキャッシュセットにおけるキャッシュエントリに充填するステップを含む。

【0013】

別の態様では、プロセッサベースの適応キャッシュプリフェッч回路にキャッシュデータをキャッシュにプリフェッчさせるコンピュータ実行可能命令を記憶した非一時的コンピュータ可読媒体が提供される。コンピュータ実行可能命令は、プロセッサベースの適応キャッシュプリフェッч回路に、少なくとも1つの第1の専用プリフェッчポリシーが適用されるキャッシュ内の少なくとも1つの第1の専用キャッシュセットと、少なくとも1つの第1の専用プリフェッчポリシーとは異なる少なくとも1つの第2の専用プリフェッчポリシーが適用されるキャッシュ内の少なくとも1つの第2の専用キャッシュセットとにおける、アクセスされたキャッシュエントリから生じるキャッシュミスに基づいて、ミス追跡回路の少なくとも1つのミス状態を更新することによって、キャッシュデータをキャッシュにプリフェッчさせる。コンピュータ実行可能命令はまた、プロセッサベースの適応キャッシュプリフェッч回路に、ミス追跡回路の少なくとも1つのミス状態に基づいて、少なくとも1つの第1の専用プリフェッчポリシーおよび少なくとも1つの第2の専用プリフェッчポリシーの中から、プリフェッч制御回路によって発行された、キャッシュを充填させるためのプリフェッч要求において適用されるべきプリフェッчポリシーを選択することによって、キャッシュデータをキャッシュにプリフェッчさせる。

【図面の簡単な説明】

【0014】

【図1】キャッシュと、キャッシュ汚染を低減するために専用キャッシュセットにおける専用プリフェッчポリシーを競合させることに基づいてキャッシュエントリをプリフェッчするように構成された例示的な適応キャッシュプリフェッч回路とを含む、例示的なキャッシュメモリシステムの概略図である。

【図2】図1のキャッシュメモリシステムのキャッシュにおいて提供されるデータアレイの概略図であり、キャッシュは、複数のフォロワキャッシュセットと、各々がキャッシュデータをそれぞれの専用キャッシュセットにプリフェッчするために使用される専用プリフェッчポリシーに関連付けられる、複数の専用キャッシュセットとから成る。

【図3A】所与の専用プリフェッчポリシーが適用されたキャッシュ内の専用キャッシュセットがアクセスされたときに、キャッシュミスが発生したかどうかに基づいて、ミス追跡回路においてミス状態を更新するための例示的なプロセスを示すフローチャートである。

【図3B】専用キャッシュセット間の競合を追跡するミスインジケータのミス状態に基づいてデータをフォロワキャッシュセットにプリフェッчするために、専用キャッシュセットにプリフェッчするために使用される専用プリフェッчポリシーの中の選択されたプリフェッчポリシーを使用する、適応キャッシュプリフェッチングのための例示的なプロセスを示すフローチャートである。

【図4】専用キャッシュセットにおける専用プリフェッчポリシーを競合させることに基づいた適応キャッシュプリフェッチングが提供されるときの、図1のキャッシュメモリシ

10

20

30

40

50

システム内のキャッシングに対する例示的なプリフェッチング性能を示すグラフである。

【図5】キャッシングと、キャッシングへのアクセスを制御するように構成されたキャッシングコントローラと、キャッシングコントローラ内部で提供され、キャッシング汚染を低減するために、データを専用キャッシングセットにプリフェッチするために使用される専用プリフェッチポリシーを競合させることに基づいて、プリフェッチポリシーをプリフェッチされたキャッシングエントリに適用するように構成された例示的なプリフェッチフィルタとを含む、例示的な代替キャッシングメモリシステムの概略図である。

【図6 A】図5のキャッシングメモリシステムにおいて提供され得る例示的なキャッシングの概略図であり、キャッシングは、複数のフォロワキャッシングセットと、各々が所与の専用キャッシングセットのための関連する専用プリフェッチポリシーを有する、複数の専用キャッシングセットとから成る。

【図6 B】図5のキャッシング内の各専用キャッシングセットに対するキャッシングミスに基づいて複数のミスカウントを更新するように構成された例示的な代替ミスカウンタの概略図である。

【図7】図1のキャッシングメモリシステムを含むことができる例示的なプロセッサベースのシステムのブロック図である。

【発明を実施するための形態】

【0015】

次に図面を参照しながら、本開示のいくつかの例示的な態様について説明する。「例示的」という言葉は、本明細書では「例、事例、または例示として役立つ」ことを意味するように使用される。「例示的」として本明細書で説明するいずれの態様も、必ずしも他の態様よりも好ましいか、または有利であると解釈されるべきではない。

【0016】

発明を実施するための形態で開示する態様は、キャッシング汚染を低減するために専用キャッシングセットにおける専用プリフェッチポリシーを競合させることに基づいた適応キャッシングプリフェッチングを含む。一態様では、適応キャッシングプリフェッチ回路は、データをキャッシングにプリフェッチするために提供される。キャッシングのための最適な置換ポリシーを決定しようとする代わりに、適応キャッシングプリフェッチ回路は、キャッシング内の専用キャッシングセットに適用される専用プリフェッチポリシーを競合させた結果に基づいて、プリフェッチポリシーを決定するように構成される。この点について、キャッシング内のキャッシングセットのサブセットは、「専用」キャッシングセットであるものとして割り振られる。その他の非専用キャッシングセットは、「フォロワ」キャッシングセットである。各専用キャッシングセットは、所与の専用キャッシングセットのための関連する専用プリフェッチポリシーを有する。専用キャッシングセットの各々へのアクセスについてのキャッシングミスは、適応キャッシングプリフェッチ回路によって追跡される。適応キャッシングプリフェッチ回路は、そのそれぞれの専用キャッシングセットに対してより少ないキャッシングミスを生じた専用プリフェッチポリシーを使用して、プリフェッチポリシーをキャッシング内のその他のフォロワキャッシングセットに適用するように構成され得る。たとえば、ある専用プリフェッチポリシーは決してプリフェッチしないものとしてよく、別の専用プリフェッチポリシーはデュエリング専用プリフェッチポリシーをキャッシングに提供するために常にプリフェッチするものとしてよい。このようにして、キャッシング内の専用キャッシングセットに対する実際のキャッシングミス結果は、フォロワキャッシングセットのためのプリフェッチポリシーとして使用される場合にどのプリフェッチポリシーがキャッシング内のより少ないキャッシング汚染をもたらすことになるかについてのより良い指標であり得るので、キャッシング汚染が低減され得る。キャッシング汚染の低減は、性能の向上、メモリ競合の低減、およびキャッシングによる電力消費の減少をもたらすことができる。

【0017】

この点について、図1は、例示的なキャッシングメモリシステム12を含む例示的なコンピュータシステム10である。専用キャッシングセットにおける専用プリフェッチポリシーを競合させることに基づいた、キャッシングメモリシステム12で利用される適応キャッシングプリ

10

20

30

40

50

フェッチフィルタリングについて説明する前に、例示的なキャッシュメモリシステム12について最初に説明する。

【0018】

この点について、図1のキャッシュメモリシステム12はキャッシュ14を含む。キャッシュ14は、上位レベルメモリ16からキャッシュ14にロードされた、キャッシュされたデータを記憶するように構成されたメモリである。例として、上位レベルメモリ16は上位レベルキャッシュまたはメインメモリであってもよい。この例では、キャッシュ14はセットアソシエティブキャッシュである。キャッシュ14は、タグアレイ18およびデータアレイ20を含む。データアレイ20は複数のキャッシュセット22(0)～22(M)を含み、ここで、「M+1」はキャッシュセット22の数に等しい。一例として、1024個のキャッシュセット22(0)～22(1023)がデータアレイ20において提供され得る。複数のキャッシュセット22(0)～22(M)の各々は、キャッシュデータを1つまたは複数のキャッシュエントリ24(0)～24(N)に記憶するように構成され、ここで、「N+1」はキャッシュセット22当たりのキャッシュエントリ24の数に等しい。キャッシュコントローラ26もキャッシュメモリシステム12において提供される。キャッシュコントローラ26は、上位レベルメモリ16からデータアレイ20にキャッシュデータを充填するように構成される。たとえば、キャッシュコントローラ26は、上位レベルメモリ16から、所与のメモリアドレスにおいて記憶されたデータに対応する、データアレイ20に記憶されるべきデータ28を受信するように構成される。受信されたデータ28は、メモリアドレスに従って、データアレイ20内のキャッシュエントリ24(0)～24(N)にキャッシュデータ30として記憶される。このようにして、中央処理装置(CPU)32は、上位レベルメモリ16からキャッシュデータ30を取得しなければならないことに対立するものとして、キャッシュ14に記憶されたキャッシュデータ30にアクセスすることができる。

【0019】

引き続き図1を参照すると、キャッシュコントローラ26はまた、CPU32または下位レベルメモリ36からメモリアクセス要求34を受信するように構成される。キャッシュコントローラ26は、メモリアクセス要求34内のメモリアドレスを使用して、キャッシュ14内のタグアレイ18をインデックス付けする。メモリアドレスによってインデックス付けされたタグアレイ18内のインデックスにおいて記憶されたタグがメモリアクセス要求34内のメモリアドレスと一致し、タグが有効である場合、キャッシュヒットが発生する。これは、メモリアクセス要求34のメモリアドレスに対応するキャッシュデータ30がデータアレイ20内のキャッシュエントリ24(0)～24(N)に含まれることを意味する。それに応答して、キャッシュコントローラ26は、メモリアクセス要求34のメモリアドレスに対応するインデックス付けされたキャッシュデータ30をCPU32または下位レベルメモリ36に戻して提供する。キャッシュミスが発生した場合、キャッシュコントローラ26はキャッシュデータ30をCPU32または下位レベルメモリ36に提供しない。

【0020】

キャッシュ14において発生するキャッシュミスは、キャッシュメモリシステム12の性能低下の原因である。キャッシュメモリシステム12におけるキャッシュミスの数を低減するために、プリフェッチ制御回路38がキャッシュメモリシステム12において提供される。プリフェッチ制御回路38は、将来のメモリアクセスを予測するために、CPU32または下位レベルメモリ36によってメモリアクセスパターンを検出するように構成され得る。これらの予測を使用して、プリフェッチ制御回路38は、キャッシュコントローラ26に対するプリフェッチ(すなわち、置換)ポリシーに基づいて、キャッシュエントリ24(0)～24(N)に記憶された既存のキャッシュデータを置換するためにキャッシュデータをキャッシュ14内のキャッシュエントリ24(0)～24(N)に投機的にプリロードするためのプリフェッチ要求40を行うことができる。したがって、近い将来に必要とされると投機的に予測されたキャッシュデータが要求されたとき、キャッシュデータはキャッシュ14内のキャッシュエントリ24(0)～24(N)にすでに存在している。したがって、結果としてキャッシュミスペナルティが生じない。しかしながら、キャッシュ14内の置換されたキャッシュデータがプリフェッチされたキャッシュデータの前に必要とされる場合、キャッシュデータをキャッシュ14にプリ

10

20

30

40

50

フェッチすることはキャッシュ汚染を引き起こす場合もある。

【0021】

図1のキャッシュ14のための最適なプリフェッチポリシーを決定しようとする代わりに、適応キャッシュプリフェッチ回路42がキャッシュメモリシステム12において提供される。以下でより詳細に説明するように、適応キャッシュプリフェッチ回路42は、キャッシュ14内の専用キャッシュセットに適用される専用プリフェッチポリシーを競合させた結果に基づいて、どのプリフェッチポリシーを使用するかを決定するように構成される。

【0022】

この点について、図2は、図1のキャッシュメモリシステム12のキャッシュ14において提供されるデータアレイ20を示す。そこに示されているように、データアレイ20は複数のキャッシュセット22(0)～22(M)を含む。しかしながら、データアレイ20内のキャッシュセット22(0)～22(M)のある一定のサブセットが専用キャッシュセット44として指定される。この例では、キャッシュセット22(0)～22(M)の中のいくつかのキャッシュセットが専用キャッシュセット44(A)として指定される。記号(A)は、キャッシュデータ30としてのデータ28を専用キャッシュセット44(A)にプリフェッチするために、第1の専用プリフェッチポリシーAがキャッシュコントローラ26によって使用されることを示す。キャッシュセット22(0)～22(M)の中の他のキャッシュセットが専用キャッシュセット44(B)として指定される。記号(B)は、キャッシュデータ30としてのデータ28を専用キャッシュセット44(B)にプリフェッチするために、第1の専用プリフェッチポリシーAとは異なる第2の専用プリフェッチポリシーBがキャッシュコントローラ26によって使用されることを示す。キャッシュセット22(0)～22(M)の中のその他の非専用キャッシュセットがフォロワキャッシュセット46として指定される。専用キャッシュセット44(A)、44(B)の各々へのアクセスについてのキャッシュミスは、適応キャッシュプリフェッチ回路42によって追跡される。適応キャッシュプリフェッチ回路42は、アクセスされたときに専用キャッシュセット44(A)、44(B)により少ないキャッシュミスを生じさせた専用プリフェッチポリシーAまたはBを使用して、プリフェッチポリシーをキャッシュセット22(0)～22(M)の中の他のフォロワキャッシュセット46に適用するように構成される。言い換えれば、図2のデータアレイ20内の専用キャッシュセット44(A)、44(B)は、互いに競合して設定される。このようにして、それぞれの専用プリフェッチポリシーAまたはBを用いてプリフェッチされた専用キャッシュセット44(A)、44(B)の各々に関連付けられた実際のキャッシュミス結果は、キャッシュセット22(0)～22(M)の中のフォロワキャッシュセット46のためのプリフェッチポリシーとして使用された場合にどのプリフェッチポリシーがキャッシュ14内のより少ないキャッシュ汚染をもたらすことになるかについてのより良い指標であり得るので、キャッシュ汚染が低減され得る。キャッシュ汚染の低減は、キャッシュメモリシステム12における性能の向上、メモリ競合の低減、およびキャッシュ14による電力消費の減少をもたらすことができる。

【0023】

図1および図2に関して以下でより詳細に説明するように、専用キャッシュセット44(A)、44(B)内のキャッシュエントリ24(0)～24(N)へのアクセスから生じるキャッシュミスは、図1のキャッシュメモリシステム12内のミス追跡回路47において追跡される。この例では、ミス追跡回路47は、プリフェッチポリシーを決定するために、専用キャッシュセット44(A)、44(B)へのアクセスから発生するキャッシュミスを追跡するように構成される。この例におけるミス追跡回路47は、ミスカウント50の形で提供されるミスインジケータ48を含む。ミスカウント50は、ミス状態52に基づいて、専用キャッシュセット44(A)、44(B)へのアクセスから発生するキャッシュミスを追跡するように構成される。ミス状態52は、この例ではミスカウント54の形で提供される。この例では、ミスカウント50は単一のミス飽和カウントである。しかしながら、以下で説明する他の態様では、別個のミスカウント50が、専用キャッシュセット44(A)、44(B)の各々に対するキャッシュミスを別個に追跡するために、専用キャッシュセット44(A)、44(B)の各々に提供される場合がある。図1のミスカウント50は、第1の専用プリフェッチポリシーAが適用される第1の専用キャッシュセット44(A)内のアクセスされたキャッシュエントリ24(0)～24(N)から生じる、キャッシュヒ

10

20

30

40

50

ット/ミスライン55を介してキャッシュコントローラ26によって報告されるキャッシュミスに基づいて、ミスカウント54を更新するように構成される。ミスカウンタ50はまた、第2の専用プリフェッчポリシーBが適用される第2の専用キャッシュセット44(B)内のアクセスされたキャッシュエントリ24(0)～24(N)から生じるキャッシュミスに基づいて、ミスカウント54を更新するように構成される。

【0024】

引き続き図1を参照すると、適応キャッシュプリフェッч回路42において提供されるプリフェッчフィルタ56は、ミスカウンタ50のミスカウント54に基づいて、第1の専用プリフェッчポリシーAおよび第2の専用プリフェッчポリシーBの中からプリフェッчポリシーを選択するように構成される。この例では、ミスカウンタ50は、専用キャッシュセット44(A)、44(B)のうちの一方へのアクセスについてキャッシュミスが発生するときに増分し、専用キャッシュセット44(B)、44(A)のうちの他方へのアクセスについてキャッシュミスが発生するときに減分するか、またはその逆も同様であるように構成されるミス飽和カウンタである。専用キャッシュセット44(A)、44(B)の各々に別個のミスカウンタを提供することは可能であり、本明細書ではオプションとして企図されるが、ミスカウンタ50としてミス飽和カウンタを提供することは、専用キャッシュセット44(A)、44(B)の各々に別個のミスカウンタを提供することに代わる、より低コストの代替策であり得る。ミスカウンタ50は、経時的にアクセスされたときに専用キャッシュセット44(A)、44(B)のどちらがより少ないキャッシュミスを生じるかを追跡する。プリフェッчフィルタ56は、フォロワーキャッシュセット46のためのプリフェッчポリシーとして使用されるべき、より少ないキャッシュミスを生じた専用キャッシュセット44(A)、44(B)に対応する専用プリフェッчポリシーAまたはBを選択するために、ミスカウントライン57を介してミスカウンタ50を受信する。この例では、プリフェッчフィルタ56は、キャッシュコントローラ26からプリフェッч要求40を受信する。プリフェッчフィルタ56は、ミスカウンタ50に基づいて、選択された専用プリフェッчポリシーAまたはBを、プリフェッч要求40'としてキャッシュコントローラ26から受信されたプリフェッч要求40に適用する。

10

20

30

40

50

【0025】

この例では、図1および図2のデータアレイ20において利用される2つ(2)の専用プリフェッчポリシーAおよびBしかないので、図2のデータアレイ20内の専用キャッシュセット44(A)、44(B)はデュエリング専用キャッシュセットであると言える。しかしながら、プリフェッчフィルタ56が2つ(2)以上の専用プリフェッчポリシーから選択することを可能にするために、各々が専用プリフェッчポリシーで指定される2つ(2)以上のタイプの専用キャッシュセット44が提供され得ることに留意されたい。図2では、データアレイ20内に示されている、プリフェッчポリシーAに関連付けられた「Q」個の専用キャッシュセット44(A)(1)～44(A)(Q)、およびプリフェッчポリシーBに関連付けられた「Q」個の専用キャッシュセット44(B)(1)～44(B)(Q)がある。たとえば、図2のデータアレイ20が1024個のキャッシュセット22(すなわち、22(0)～22(M)、ここで、「M」は1023に等しい)を含んでいた場合、キャッシュセット22(0)～22(1023)のうちの30個(32)は専用キャッシュセット44(A)として指定され得、キャッシュセット22(0)～22(1023)のうちの30個(32)は専用キャッシュセット44(B)として指定され得る。この例では、「Q」は32(32)に等しくなる。これにより、キャッシュセット22(0)～22(M)のうちの960個(960)がフォロワーキャッシュセット46として残ることになる。同数の専用キャッシュセット44が各専用プリフェッчポリシーAおよびBに専用であることは必要とされないことに留意されたい。

30

40

50

【0026】

データアレイ20内のより大きい数のキャッシュセット22(0)～22(M)を専用キャッシュセット44として指定することは、それぞれの専用キャッシュセット44(A)、44(B)へのアクセスがより頻繁に発生し得るので、より頻繁に更新されるべき競合する専用プリフェッчポリシーAおよびBを提供し得る。しかしながら、データアレイ20内のより大きい数のキャッシュセット22(0)～22(M)を専用キャッシュセット44として指定することはまた、競合するプリフェッчポリシーAまたはBが適用され得るキャッシュセット22(0)～22(M)の中のフォ

口ワキャッシュセット46の数を制限する。専用キャッシュセット44(A)、44(B)として選択されるキャッシュセット22(0)～22(M)の数、ならびにデータアレイ20内部の専用キャッシュセット44(A)および44(B)のロケーションは、データアレイ20内のキャッシュセット22(0)～22(M)へのアクセスの分布を確率的に決定するためのサンプリングなどの設計上の考慮事項に基づいて選択され得る。

【0027】

さらに、プリフェッчポリシーAおよびBが異なるプリフェッчポリシーである限り、専用プリフェッчポリシーAおよびBは任意の所望のプリフェッчポリシーとして提供され得る。そうでない場合、同じプリフェッчポリシーがフォロワキャッシュセット46に適用されることになり、このことにより、適応キャッシュプリフェッч回路42を利用することなしにすべてのキャッシュセット22(0)～22(M)について単一のプリフェッчポリシーを使用することに比べて、キャッシュ汚染を低減する機会がなくなる。たとえば、データ28を専用キャッシュセット44(A)(1)～44(A)(Q)にプリフェッчするために使用されるプリフェッчポリシーAは決してプリフェッчしないものとしてよく、一方、プリフェッчポリシーBはデータ28を専用キャッシュセット44(B)(1)～44(B)(Q)に常にプリフェッчするものとしてよい。

10

【0028】

専用キャッシュセット44(A)、44(B)における専用プリフェッчポリシーを競合させることに基づいて図1のキャッシュメモリシステム12上で実行される適応プリフェッчングについてさらに説明するために、図3Aおよび図3Bが提供される。図3Aは、専用キャッシュセット44(A)、44(B)の競合を追跡するために、キャッシュ14内の専用キャッシュセット44(A)、44(B)がアクセスされたときにキャッシュミスが発生するかどうかに基づいてミスカウント50のミスカウント54を更新するための例示的なプロセス60のフローチャートである。図3Bは、専用キャッシュセット44(A)、44(B)間の競合を追跡するミスカウント50のミスカウント54に基づいてデータ28をキャッシュ14内のフォロワキャッシュセット46にプリフェッчするために、専用プリフェッчポリシーA、Bの中の選択されたプリフェッчポリシーを使用する、適応キャッシュプリフェッчングのための例示的なプロセス80のフローチャートである。図1のキャッシュメモリシステム12を参照して、両方のプロセス60、80について説明する。

20

【0029】

30

図3Aを参照すると、キャッシュ14のキャッシュコントローラ26は、キャッシュ14内でアドレス指定されるべきメモリアドレスを含むメモリアクセス要求34を受信する(ブロック62)。キャッシュコントローラ26は、メモリアクセス要求34のメモリアドレスに対応する、キャッシュ14内のキャッシュエントリ24(0)～24(N)の中のアクセスされたキャッシュエントリ24が、キャッシュ14のデータアレイ20に含まれるかどうかを判断するために、タグアレイ18に照会する(ブロック64)。メモリアクセス要求34のメモリアドレスがキャッシュ14のデータアレイ20に含まれる、すなわち、キャッシュヒットが発生した場合(決定66)、ミスカウント50のミスカウント54は更新されず(ブロック66)、プロセスは終了する(ブロック68)。しかしながら、メモリアクセス要求34がキャッシュ14のデータアレイ20に含まれない(決定66)、すなわち、キャッシュミスが発生した場合、キャッシュコントローラ26はキャッシュミスを適応キャッシュプリフェッч回路42に通信する。キャッシュミスが専用キャッシュセット44(A)または44(B)に対するものである場合(決定70)、ミスカウント50のミスカウント54は、専用キャッシュセット44(A)、44(B)に対するアクセスされたキャッシュエントリ24から生じるキャッシュミスに基づいて更新され(ブロック72、74)、プロセスは終了する(ブロック68)。たとえば、ミスカウント50のミスカウント54は、アクセスされたキャッシュエントリ24から生じるキャッシュミスが専用キャッシュセット44(A)で発生した場合は増分され、アクセスされたキャッシュエントリ24から生じるキャッシュミスが専用キャッシュセット44(B)で発生した場合は減分され得る。したがって、図3Aのこの例示的なプロセス60は、専用キャッシュセット44(B)に対するキャッシュミスの完了を追跡するようにミスカウント50のミスカウント54を維持する。キャッシュミスが専用キャッシュ

40

50

ュセット44(A)または44(B)に対するものではない場合(決定70)、ミスカウント54は更新されず、プロセスは終了する(ブロック68)。

【0030】

上記で説明したように、図3Bのプロセス80は、ミスカウント50のミスカウント54に基づいて、専用キャッシュセット44(A)、44(B)に関連付けられた専用プリフェッчポリシーA、Bの中の選択されたプリフェッчポリシーを使用して、データ28をキャッシュ14にプリフェッчするために使用される。この点について、プリフェッч要求40は、データ28をキャッシュ14内のキャッシュセット22(0)～22(M)の中のアクセスされたキャッシュセット22内のキャッシュエントリ24にプリフェッчするために、CPU32または下位レベルメモリ36によって発行される(ブロック82)。適応キャッシュプリフェッч回路42のプリフェッчフィルタ56は、キャッシュコントローラ26から受信された情報に基づいて、アクセスされたキャッシュセット22が専用キャッシュセット44(A)、44(B)であるかどうかを判断する(決定84)。アクセスされたキャッシュセット22が専用キャッシュセット44(A)、44(B)である場合(決定84)、プリフェッчフィルタ56によって適用されるプリフェッчポリシーは、アクセスされた特定の専用キャッシュセット44(A)、44(B)に関連付けられたそれぞれの専用プリフェッчポリシーAまたはBである(ブロック88)。しかしながら、アクセスされたキャッシュセット22が専用キャッシュセット44(A)、44(B)ではなく(決定84)、代わりにフォロワキャッシュセット46である場合、プリフェッчフィルタ56は、ミスカウント50のミスカウント54に基づいて、専用プリフェッчポリシーAまたはBの中からプリフェッч要求40に適用されるべきプリフェッчポリシーを選択する(ブロック86)。たとえば、アクセスされたときに専用キャッシュセット44(A)が専用キャッシュセット44(B)よりも少ないキャッシュミスを生じたことをミスカウント54が示す場合、プリフェッчフィルタ56は、フォロワキャッシュセット46へのプリフェッч要求40に使用されるべきプリフェッчポリシーAを選択し得る。また、ブロック86において、追加または代替の特徴として、適応キャッシュプリフェッч回路42のプリフェッчフィルタ56は、ミスカウントに基づいて、第1の専用プリフェッчポリシーAまたは第2の専用プリフェッчポリシーBがプリフェッч要求40に適用されるべきかどうかを確率的に判断するように制御されることも可能である。いずれの場合も、アクセスされたキャッシュセット22が専用キャッシュセット44(A)、44(B)であるかフォロワキャッシュセット46であるかにかかわらず、プリフェッчフィルタ56によって適用される選択されたプリフェッчポリシーは、プリフェッчされたキャッシュデータ30をアクセスされたキャッシュセット22のキャッシュエントリ24に充填するために使用され(ブロック90)、プロセスは終了する(ブロック92)。

【0031】

上記で説明したように、専用プリフェッчポリシーAまたは専用プリフェッчポリシーBをバイモーダル的に(bimodally)選ぶためにミスカウント54を固定しきい値に適用するのではなく、ミスカウント54は、ミスカウント54の値(magnitude)に基づいて専用プリフェッчポリシーAを使用するかまたは専用プリフェッчポリシーBを使用するかを選択する確率を制御するために使用され得る。たとえば、ミスカウント54の大きい値は、専用プリフェッчポリシーAを選ぶ高い確率(および逆に、専用プリフェッчポリシーBを選ぶ低い確率)を示すために使用されてもよい。ミスカウント54の小さい値は、専用プリフェッчポリシーAを選ぶ低い確率(および逆に、専用プリフェッчポリシーBを選ぶ高い確率)を示すために使用されてもよい。一例として、そのような確率的関数は、ミスカウント54と比較されるべきランダム整数を生成することによって実装され得る。たとえば、ミスカウント54が6(6)ビットカウントを使用して実装される場合、ランダム6ビット整数が生成され、ミスカウント54と比較される。ミスカウント54がランダムに生成された整数以下である場合、専用プリフェッчポリシーAが使用され、そうでない場合、専用プリフェッчポリシーBが使用される。

【0032】

図4は、適応キャッシュプリフェッчングが適応キャッシュプリフェッч回路42によって実行されるときの、図1のキャッシュメモリシステム12のキャッシュ14に対する例示的

10

20

30

40

50

なプリフェッティング性能を示すグラフ94である。この点について、キャッシュ汚染96はY軸上に示されている。キャッシュ汚染96のより高いレベルは、グラフ94のY軸上のより大きい幅によって示されている。キャッシュ汚染96は、上記で説明した適応キャッシュプリフェッチ回路42によって提供される非プリフェッチポリシー100のみ、常時プリフェッチポリシー102のみ、およびプリフェッチデュエリングポリシー104を使用した、X軸上に示す例示的なアプリケーション98(1)～98(X)についてベンチマークされる。図示のように、適応キャッシュプリフェッチ回路42によって提供されるプリフェッチデュエリングポリシー104を利用するキャッシュ汚染96は、非プリフェッチポリシー100のみまたは常時プリフェッチポリシー102のみを使用することに対して、大半のアプリケーション98(1)～98(X)についてより少ないキャッシュ汚染96(すなわち、より小さい幅のキャッシュ汚染96)をもたらす。

10

【0033】

さらに、図3Aおよび図3Bの例示的なプロセスにおける図1の適応キャッシュプリフェッチ回路42の動作は、選択的に無効化されるように構成され得ることに留意されたい。たとえば、図1の適応キャッシュプリフェッチ回路42は、図3Bのブロック86において第1の専用プリフェッチポリシーAおよび第2の専用プリフェッチポリシーBの中からプリフェッチポリシーを選択しないように構成される場合がある。代わりに、デフォルトのプリフェッチポリシーまたはプリフェッチ要求40に提供されるもしくはプリフェッチ要求40に関連付けられたプリフェッチポリシーが、データ28をフォロワーキャッシュセット46にプリフェッチするために使用されることになる。たとえば、有効化/無効化機能は、有効化/無効化ビットとして指定されるミスカウント54内のビットに基づいて制御される場合がある。たとえば、ミスカウント54内の最上位ビットは、適応キャッシュプリフェッチ有効化/無効化ビットとして指定される場合がある。ミスカウンタ50は、キャッシュコントローラ26からの命令に基づいて、ミスカウント54において有効化/無効化ビットを設定するように構成される場合がある。適応キャッシュプリフェッチ回路42は、ミスカウント54に基づいて、プリフェッチフィルタ56が専用プリフェッチポリシーをプリフェッチ要求40に適用するべきかどうかを判断するために、ミスカウンタ50からミスカウント54を受信することの一部としてその有効化/無効化ビットを検討するように構成される場合がある。同様に、所望される場合、プリフェッチフィルタ54が専用プリフェッチポリシーA、Bのうちの1つを使用すべきではないことを示すために、インジケータが適応キャッシュプリフェッチ回路42において提供される場合がある。

20

【0034】

図1では、適応キャッシュプリフェッチ回路42は、キャッシュメモリシステム12内のキャッシュコントローラ26の外部で提供される。上記で説明したように、適応キャッシュプリフェッチ回路42は、キャッシュセット22(0)～22(M)の中のフォロワーキャッシュセット46へのプリフェッチのための、専用プリフェッチポリシーAまたはBの中の選択されたプリフェッチポリシーを適用するためのプリフェッチ要求40を受信する。しかしながら、図1の適応キャッシュプリフェッチ回路42の機能は、キャッシュコントローラ26内部で提供されるか、またはキャッシュコントローラ26に組み込まれる場合もある。さらに、ミス追跡回路47は、キャッシュコントローラ26内部で提供される場合もある。この点について、図5は、代替キャッシュメモリシステム12(1)を含む代替コンピュータシステム10(1)を示す。図1のキャッシュメモリシステム12と図5のキャッシュメモリシステム12(1)との間で共通の構成要素は共通の要素番号を用いて示されており、したがって、ここでは改めて説明しない。この態様では、図1の適応キャッシュプリフェッチ回路42の機能を含む代替キャッシュコントローラ26(1)が提供される。キャッシュコントローラ26(1)の外部に示されているミスカウンタ50が提供されるが、ミスカウンタ50はキャッシュコントローラ26(1)内部に含まれる場合もある。

40

【0035】

さらに、上記で説明した図1および図2のデータアレイ20内の複数のキャッシュセット22(0)～22(M)の中のキャッシュセット22は専用キャッシュセット44(A)、44(B)として指定さ

50

れ、そこではミスカウンタ50はミス飽和カウンタであったが、そのように限定されないことに留意されたい。たとえば、データアレイ20内の複数のキャッシュセット22(0)～22(M)の中の2つ(2)以上のタイプのキャッシュセット22が専用キャッシュセット44として指定され得る。これは、適応キャッシュプリフェッチ回路42によって適用され得る2つ(2)以上の専用プリフェッチポリシーを提供するために所望され得る。この場合、図1および図5のキャッシュメモリシステム12、12(1)においてそれぞれ提供される単一のミスカウンタ50を使用する代わりに、2つ(2)以上の専用キャッシュセット44の各々に対するキャッシュミスを別個に追跡するために複数のミスカウンタが提供され得る。

【0036】

この点について、図6Aは、2つ(2)以上のタイプの専用キャッシュセット44を有する、キャッシュメモリシステム12、12(1)内のデータアレイ20の図である。図6Aのデータアレイ20には、3つ(3)のタイプの専用キャッシュセット44(A)、44(B)、および44(C)があり、専用プリフェッチポリシーA、B、およびCは、それぞれ専用キャッシュセット44(A)、44(B)、44(C)の各々に関連付けられる。さらに、専用キャッシュセット44内部で指定されるキャッシュセット22の数は変動する場合がある。たとえば、専用キャッシュセット44(A)、44(B)はそれぞれ、「Q」個のキャッシュセット22(すなわち、44(A)(1)～44(A)(Q)および44(B)(1)～44(B)(Q))を含む。しかしながら、専用キャッシュセット44(C)は、「R」個のキャッシュセット22(すなわち、44(C)(1)～44(C)(R))を含む。このようにして、適応キャッシュプリフェッチ回路42は、専用キャッシュセット44(A)、44(B)、および44(C)に対する追跡されたキャッシュミスの競合に基づいて、キャッシュセット22(0)～22(M)の中のフォロワーキャッシュセット46にプリフェッチするために専用プリフェッチポリシーA、B、またはCのいずれかを適用することができる。

【0037】

図6Bは、代替ミスカウンタ50(1)の形の代替ミスインジケータ48(1)を有する代替ミス追跡回路47(1)を示す。ミスカウンタ50(1)は、図6Aの専用キャッシュセット44(A)、44(B)、および44(C)に対するキャッシュミスを追跡するように構成される。この態様では、2つ(2)のタイプの専用キャッシュセット44(A)、44(B)があるだけではないので、競合する専用キャッシュセット44(A)、44(B)、44(C)ごとのミスカウント54(1)を追跡するための追加のミスカウンタが必要とされる。この点について、ミスカウンタ50(1)は複数のミスカウント54(1)～54(D)から成り、ここで、「D」は、図6Aのデータアレイ20内の専用キャッシュセット44(A)、44(B)、44(C)として提供されるキャッシュセット22(0)～22(M)の中のキャッシュセット22の総数である。このようにして、プリフェッチフィルタ56は、データ28をデータアレイ20のフォロワーキャッシュセット46にプリフェッチするために専用プリフェッチポリシーA、B、およびCの中のどの専用プリフェッチポリシーを使用すべきかを決定するために、ミスカウンタ50(1)内のミスカウント54(1)～54(D)の各々を比較することができる。

【0038】

本明細書で開示する態様による適合キャッシュプリフェッチ回路および/またはキャッシュメモリシステムは、任意のプロセッサベースのデバイスにおいて提供されるか、そのデバイスに統合され得る。例としては、限定はしないが、セットトップボックス、エンターテインメントユニット、ナビゲーションデバイス、通信デバイス、固定ロケーションデータユニット、モバイルロケーションデータユニット、モバイルフォン、セルラーフォン、コンピュータ、ポータブルコンピュータ、デスクトップコンピュータ、携帯情報端末(PDA)、モニタ、コンピュータモニタ、テレビジョン、チューナー、ラジオ、衛星ラジオ、音楽プレーヤ、デジタル音楽プレーヤ、ポータブル音楽プレーヤ、デジタルビデオプレーヤ、ビデオプレーヤ、デジタルビデオディスク(DVD)プレーヤ、およびポータブルデジタルビデオプレーヤがある。

【0039】

この点について、図7は、図1および図5のキャッシュメモリシステム12、12(1)および/または適応キャッシュプリフェッチ回路42、42(1)を利用することができるプロセッサベ

10

20

30

30

40

50

ースのシステム110の一例を示す。この例では、プロセッサベースのシステム110は、各々が1つまたは複数のプロセッサ114を含む1つまたは複数のCPU112を含む。CPU112はマスタデバイスであり得る。CPU112は、一時的に記憶されたデータに迅速にアクセスするための、プロセッサ114に結合されたキャッシュメモリシステム12または12(1)を含むことができる。CPU112はシステムバス116に結合され、プロセッサベースのシステム110に含まれるマスタデバイスとスレーブデバイスを相互結合することができる。よく知られているように、CPU112は、システムバス116を介してアドレス情報、制御情報、およびデータ情報を交換することによって、これらの他のデバイスと通信する。たとえば、CPU112は、バストラシザクション要求をスレーブデバイスの一例としてのメモリコントローラ118に通信することができる。図7には示されていないが、複数のシステムバス116が提供される場合があり、各システムバス116は異なるファブリックを構成する。

10

【0040】

他のマスタデバイスおよびスレーブデバイスは、システムバス116に接続され得る。図7に示すように、これらのデバイスは、例として、メモリシステム120、1つまたは複数の入力デバイス122、1つまたは複数の出力デバイス124、1つまたは複数のネットワークインターフェースデバイス126、および1つまたは複数のディスプレイコントローラ128を含むことができる。入力デバイス122は、限定はしないが、入力キー、スイッチ、音声プロセッサなどを含む、任意のタイプの入力デバイスを含むことができる。出力デバイス124は、限定はしないが、オーディオインジケータ、ビデオインジケータ、他の視覚インジケータなどを含む、任意のタイプの出力デバイスを含むことができる。ネットワークインターフェースデバイス126は、ネットワーク130との間のデータの交換を可能にするように構成された任意のデバイスとすることができます。ネットワーク130は、限定はしないが、ワイヤードまたはワイヤレスネットワーク、プライベートまたは公共ネットワーク、ローカルエリアネットワーク(LAN)、ワイドローカルエリアネットワーク(WLAN)、およびインターネットを含む、任意のタイプのネットワークとすることができます。ネットワークインターフェースデバイス126は、任意のタイプの所望の通信プロトコルをサポートするように構成され得る。

20

【0041】

CPU112はまた、1つまたは複数のディスプレイ132に送られる情報を制御するために、システムバス116を介してディスプレイコントローラ128にアクセスするように構成され得る。ディスプレイコントローラ128は、1つまたは複数のビデオプロセッサ134を介して表示されるべき情報をディスプレイ132に送り、1つまたは複数のビデオプロセッサ134は、表示されるべき情報をディスプレイ132に適したフォーマットに処理する。ディスプレイ132は、限定はしないが、陰極線管(CRT)、液晶ディスプレイ(LCD)、プラズマディスプレイなどを含む、任意のタイプのディスプレイを含むことができる。

30

【0042】

当業者は、本明細書で開示する態様に関して説明する様々な例示的な論理ブロック、モジュール、回路、およびアルゴリズムが、電子ハードウェア、メモリもしくは別のコンピュータ可読媒体に記憶されプロセッサもしくは他の処理デバイスによって実行される命令、または両方の組合せとして実装され得ることをさらに諒解されよう。本明細書で開示するメモリは、任意のタイプおよびサイズのメモリであり得、任意のタイプの所望の情報を記憶するように構成され得る。この互換性を明確に説明するために、様々な例示的な構成要素、ブロック、モジュール、回路、およびステップについて、上記では概してこれらの機能に関して説明した。そのような機能がどのように実装されるかは、特定の適用例、設計上の選定、および/または全体的なシステムに課された設計制約に依存する。当業者は、説明した機能を特定の適用例ごとに様々な方法で実装し得るが、そのような実装の決定は、本開示の範囲からの逸脱を引き起こすものとして解釈されるべきではない。

40

【0043】

本明細書で開示する態様に関して説明する様々な例示的な論理ブロック、モジュール、および回路は、プロセッサ、デジタル信号プロセッサ(DSP)、特定用途向け集積回路(ASIC

50

)、フィールドプログラマブルゲートアレイ(FPGA)もしくは他のプログラマブル論理デバイス、個別ゲートもしくはトランジスタ論理、個別ハードウェア構成要素、または本明細書で説明する機能を実行するように設計されたそれらの任意の組合せを用いて実装または実行され得る。プロセッサはマイクロプロセッサであり得るが、代替として、プロセッサは任意の従来のプロセッサ、コントローラ、マイクロコントローラ、または状態機械であり得る。プロセッサはまた、コンピューティングデバイスの組合せ、たとえば、DSPとマイクロプロセッサの組合せ、複数のマイクロプロセッサ、DSPコアと連携した1つもしくは複数のマイクロプロセッサ、または任意の他のそのような構成として実装され得る。

【0044】

本明細書で開示する態様は、ハードウェアおよびハードウェアに記憶された命令において具現化され得、たとえば、ランダムアクセスメモリ(RAM)、フラッシュメモリ、読み取り専用メモリ(ROM)、電気的プログラマブルROM(EPROM)、電気的消去可能プログラマブルROM(EEPROM)、レジスタ、ハードディスク、リムーバブルディスク、CD-ROM、または当技術分野で知られている任意の他の形態のコンピュータ可読媒体に存在し得る。例示的な記憶媒体は、プロセッサが記憶媒体から情報を読み取り、記憶媒体に情報を書き込むことができるようプロセッサに結合される。代替として、記憶媒体はプロセッサと一体であり得る。プロセッサおよび記憶媒体はASICに存在し得る。ASICはリモート局に存在し得る。代替として、プロセッサおよび記憶媒体は個別構成要素としてリモート局、基地局、またはサーバに存在し得る。

【0045】

本明細書の例示的な態様のいずれかで説明した動作ステップは、例および議論を提供するために説明したものであることにも留意されたい。説明した動作は、図示した順序以外の多数の異なる順序で実行され得る。さらに、単一の動作ステップで説明した動作は、実際にはいくつかの異なるステップで実行され得る。加えて、例示的な態様で説明した1つまたは複数の動作ステップが組み合わされ得る。当業者には容易に明らかになるように、フローチャート図に示す動作ステップは多数の異なる修正を受ける場合があることを理解されたい。当業者はまた、情報および信号が様々な異なる技術および技法のいずれかを使用して表され得ることを理解されよう。たとえば、上記の説明全体にわたって言及され得るデータ、命令、コマンド、情報、信号、ビット、シンボル、およびチップは、電圧、電流、電磁波、磁場もしくは磁性粒子、光場もしくは光学粒子、またはそれらの任意の組合せによって表され得る。

【0046】

本開示の前述の説明は、いかなる当業者も本開示を作製または使用できるようにするために提供される。本開示に対する様々な修正は、当業者に容易に明らかになり、本明細書で定義する一般原理は、本開示の趣旨または範囲から逸脱することなく、他の変形形態に適用され得る。したがって、本開示は、本明細書で説明する例および設計に限定されるものではなく、本明細書で開示する原理および新規の特徴と一致する最も広い範囲を与えられるべきである。

【符号の説明】

【0047】

- 10 コンピュータシステム
- 10(1) 代替コンピュータシステム
- 12 キャッシュメモリシステム
- 12(1) 代替キャッシュメモリシステム
- 14 キャッシュ
- 16 上位レベルメモリ
- 18 タグアレイ
- 20 データアレイ
- 22、22(0)～22(M) キャッシュセット
- 24(0)～24(N) キャッシュエントリ

10

20

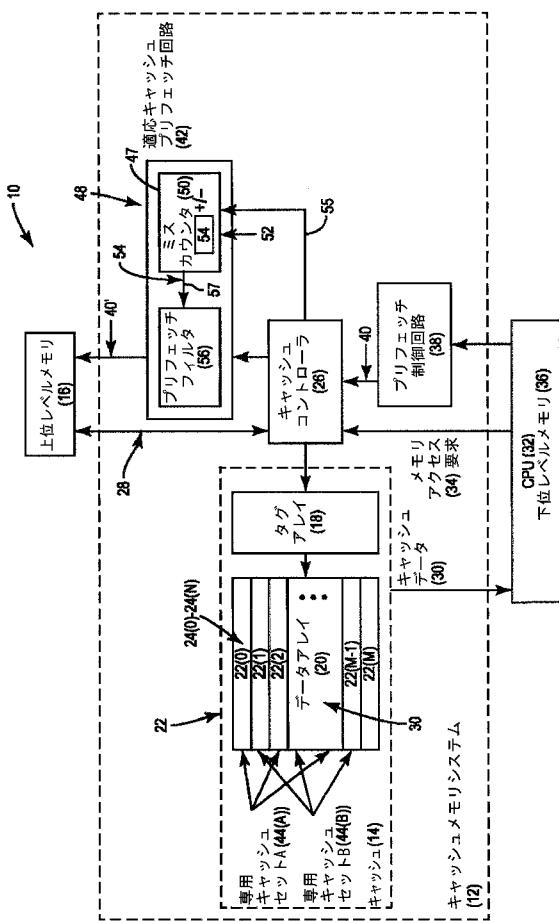
30

40

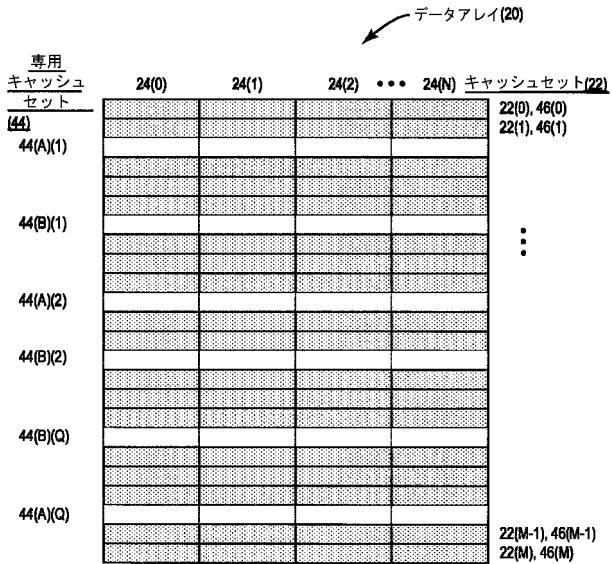
50

| | | |
|--|--------------------|----|
| 26 | キャッシュコントローラ | |
| 26(1) | 代替キャッシュコントローラ | |
| 28 | データ | |
| 30 | キャッシュデータ | |
| 32 | CPU | |
| 34 | メモリアクセス要求 | |
| 36 | 下位レベルメモリ | |
| 38 | プリフェッヂ制御回路 | |
| 40、40' | プリフェッヂ要求 | |
| 42、42(1) | 適応キャッシュプリフェッヂ回路 | 10 |
| 44(A)、44(B)、44(C)、44(A)(1) ~ 44(A)(Q)、44(B)(1) ~ 44(B)(Q)、44(C)(1) ~ 44(C) | | |
| (R) | 専用キャッシュセット | |
| 46 | フォロワーキャッシュセット | |
| 47 | ミス追跡回路 | |
| 47(1) | 代替ミス追跡回路 | |
| 48 | ミスインジケータ | |
| 48(1) | 代替ミスインジケータ | |
| 50 | ミスカウンタ | |
| 50(1) | 代替ミスカウンタ | |
| 52 | ミス状態 | 20 |
| 54、54(1) ~ 54(D) | ミスカウント | |
| 55 | キャッシュヒット/ミスライン | |
| 56 | プリフェッヂフィルタ | |
| 57 | ミスカウントライン | |
| 60 | プロセス | |
| 80 | プロセス | |
| 94 | グラフ | |
| 96 | キャッシュ汚染 | |
| 100 | 非プリフェッヂポリシー | |
| 102 | 常時プリフェッヂポリシー | 30 |
| 104 | プリフェッヂデュエリングポリシー | |
| 98(1) ~ 98(X) | アプリケーション | |
| 110 | プロセッサベースのシステム | |
| 112 | CPU | |
| 114 | プロセッサ | |
| 116 | システムバス | |
| 118 | メモリコントローラ | |
| 120 | メモリシステム | |
| 122 | 入力デバイス | |
| 124 | 出力デバイス | 40 |
| 126 | ネットワークインターフェースデバイス | |
| 128 | ディスプレイコントローラ | |
| 130 | ネットワーク | |
| 132 | ディスプレイ | |
| 134 | ビデオプロセッサ | |

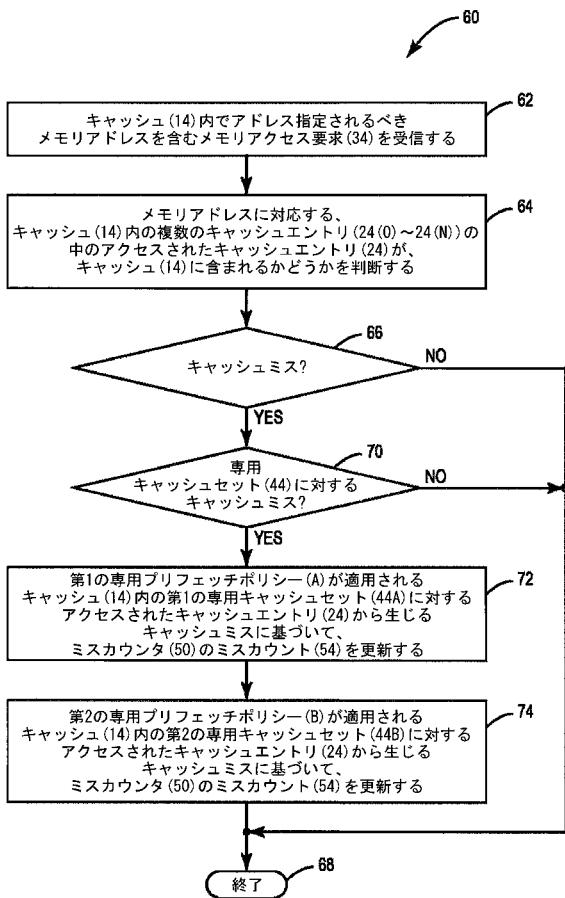
【 図 1 】



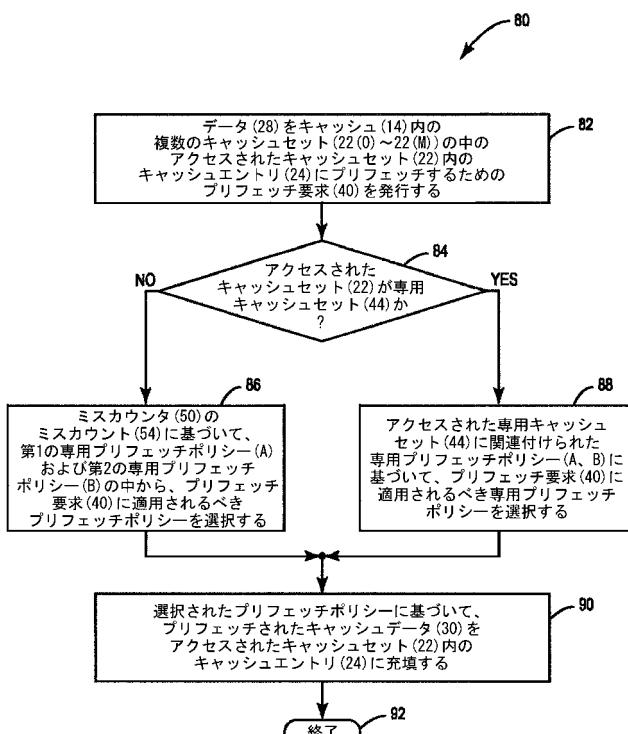
【 図 2 】



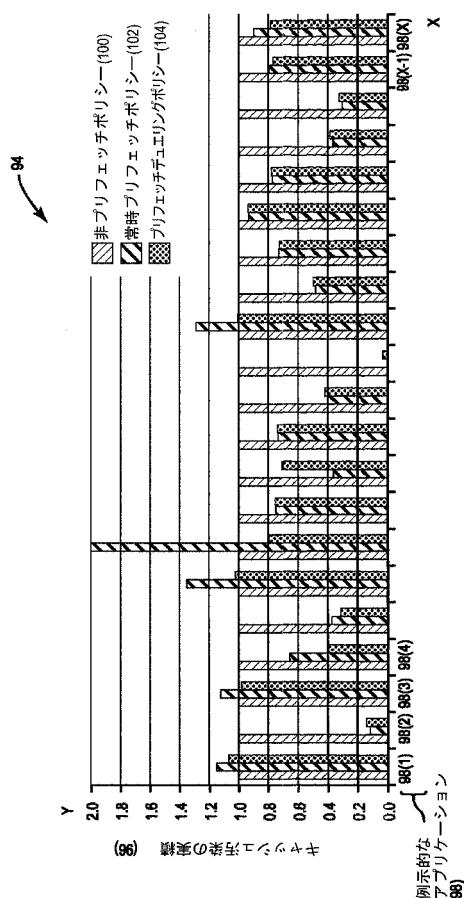
【 図 3 A 】



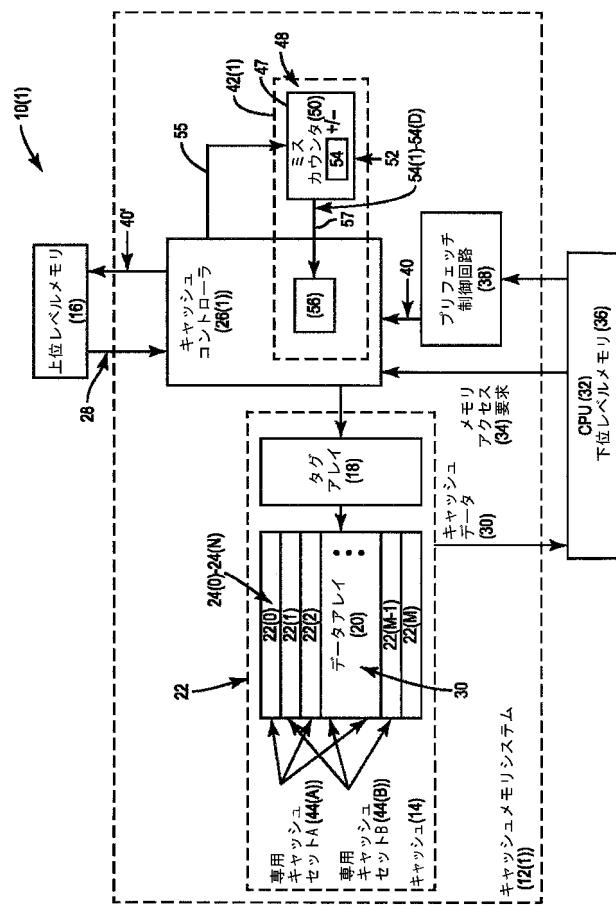
【 図 3 B 】



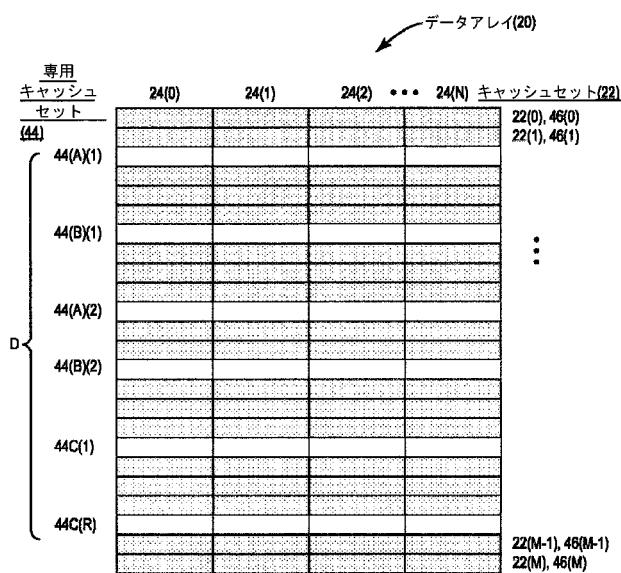
【 図 4 】



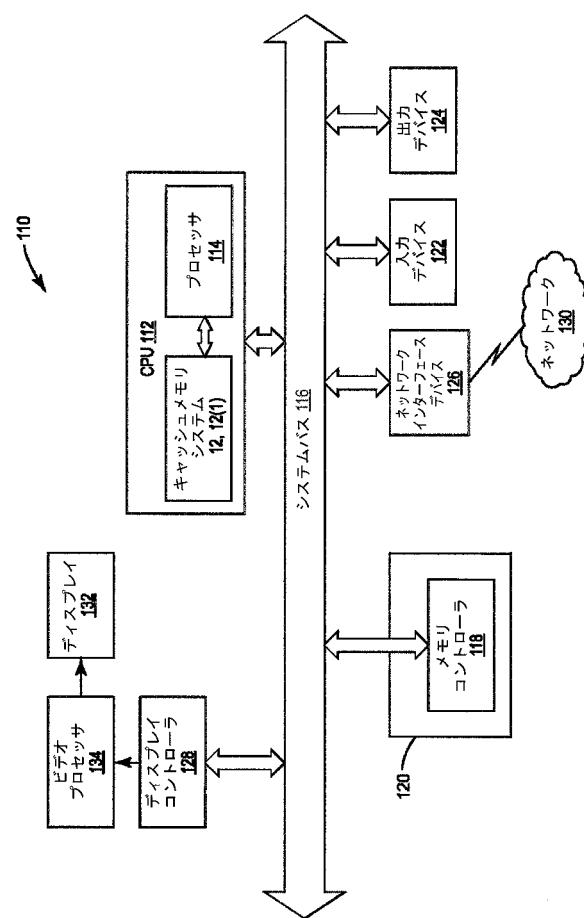
【 図 5 】



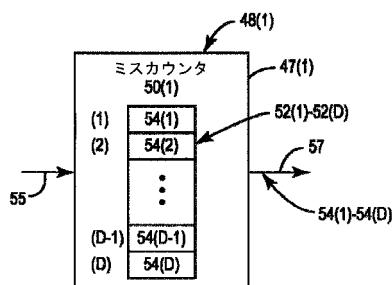
【 図 6 A 】



【 図 7 】



【 図 6 B 】



【国際調査報告】

| INTERNATIONAL SEARCH REPORT | | International application No PCT/US2015/024030 | | | | | | |
|---|---|--|-----------|--|-----------------------|---|---|------|
| A. CLASSIFICATION OF SUBJECT MATTER INV. G06F12/08 ADD. | | | | | | | | |
| According to International Patent Classification (IPC) or to both national classification and IPC | | | | | | | | |
| B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06F | | | | | | | | |
| Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched | | | | | | | | |
| Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal, WPI Data | | | | | | | | |
| C. DOCUMENTS CONSIDERED TO BE RELEVANT <table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="text-align: left; padding: 2px;">Category*</th> <th style="text-align: left; padding: 2px;">Citation of document, with indication, where appropriate, of the relevant passages</th> <th style="text-align: left; padding: 2px;">Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td style="text-align: center; padding: 2px;">Y</td> <td style="padding: 2px;"> ALAA R ALAMELDEEN ET AL: "Interactions Between Compression and Prefetching in Chip Multiprocessors", HIGH PERFORMANCE COMPUTER ARCHITECTURE, 2007. HPCA 2007. IEEE 13TH INTERNATIONAL SYMPOSIUM ON, IEEE, PI, 1 February 2007 (2007-02-01), pages 228-239, XP031072910, ISBN: 978-1-4244-0804-7 abstract page 228, right-hand column, paragraph 3 page 229, right-hand column, paragraph 5 page 230, left-hand column, last paragraph page 230, right-hand column ----- -/-/ </td> <td style="text-align: center; padding: 2px;">1-29</td> </tr> </tbody> </table> | | | Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. | Y | ALAA R ALAMELDEEN ET AL: "Interactions Between Compression and Prefetching in Chip Multiprocessors", HIGH PERFORMANCE COMPUTER ARCHITECTURE, 2007. HPCA 2007. IEEE 13TH INTERNATIONAL SYMPOSIUM ON, IEEE, PI, 1 February 2007 (2007-02-01), pages 228-239, XP031072910, ISBN: 978-1-4244-0804-7 abstract page 228, right-hand column, paragraph 3 page 229, right-hand column, paragraph 5 page 230, left-hand column, last paragraph page 230, right-hand column ----- -/-/ | 1-29 |
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. | | | | | | |
| Y | ALAA R ALAMELDEEN ET AL: "Interactions Between Compression and Prefetching in Chip Multiprocessors", HIGH PERFORMANCE COMPUTER ARCHITECTURE, 2007. HPCA 2007. IEEE 13TH INTERNATIONAL SYMPOSIUM ON, IEEE, PI, 1 February 2007 (2007-02-01), pages 228-239, XP031072910, ISBN: 978-1-4244-0804-7 abstract page 228, right-hand column, paragraph 3 page 229, right-hand column, paragraph 5 page 230, left-hand column, last paragraph page 230, right-hand column ----- -/-/ | 1-29 | | | | | | |
| <input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. | | <input checked="" type="checkbox"/> See patent family annex. | | | | | | |
| * Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed | | | | | | | | |
| Date of the actual completion of the international search | | Date of mailing of the international search report | | | | | | |
| 12 June 2015 | | 25/06/2015 | | | | | | |
| Name and mailing address of the ISA/ European Patent Office, P.O. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel: (+31-70) 340-2040, Fax: (+31-70) 340-3016 | | Authorized officer Jardon, Stéphan | | | | | | |

INTERNATIONAL SEARCH REPORT

| |
|------------------------------|
| International application No |
| PCT/US2015/024030 |

| C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT | | |
|--|--|-----------------------|
| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| Y | US 6 560 676 B1 (NISHIMOTO AKIRA [JP] ET AL) 6 May 2003 (2003-05-06) abstract; figure 3 column 2, line 59 - column 3, line 12 column 4, line 40 - column 6, line 24 claims 1-14 ----- | 1-29 |
| A | US 5 732 242 A (MOWRY TODD C [CA]) 24 March 1998 (1998-03-24) column 3, line 41 - column 3, line 59; claims 1, 7; figures 3-8 ----- | 1-29 |
| A | US 2012/096227 A1 (DUBROVIN LEONID [IL] ET AL) 19 April 2012 (2012-04-19) paragraphs [0003] - [0005], [0014], [0018] - [0037]; claim 1; figures 1,3-4 ----- | 1-29 |
| A | US 2011/145508 A1 (PELLEG DAN [IL] ET AL) 16 June 2011 (2011-06-16) the whole document ----- | 1-29 |
| A | WO 02/39283 A2 (EMC CORP [US]) 16 May 2002 (2002-05-16) abstract; figure 6 page 2, line 29 - page 3, line 27 ----- | 1-29 |
| A | US 2004/268050 A1 (CAI ZHONG-NING [US] ET AL) 30 December 2004 (2004-12-30) abstract; figures 1-4, 7-11 paragraphs [0004], [0017], [0025] - [0031]; claim 1 ----- | 1-29 |
| A | US 2004/205298 A1 (BEARDEN BRIAN S [US] ET AL) 14 October 2004 (2004-10-14) abstract paragraphs [0006] - [0010], [0021], [0097] - [0115]; claim 1; figures 7, 8, 9 ----- | 1-29 |
| A | US 2007/239940 A1 (DOSHI KSHITIJ A [US] ET AL) 11 October 2007 (2007-10-11) paragraphs [0002] - [0004], [0013] - [0030]; figures 3a, 3b ----- | 1-29 |
| A | US 2006/174228 A1 (RADHAKRISHNAN RAMESH [US] ET AL) 3 August 2006 (2006-08-03) abstract; claim 1; figure 1 paragraphs [0006], [0008] - [0009], [0024] - [0025], [0042] - [0056]; claims 1-21; figures 1-3 ----- | 1-29 |
| A | US 7 899 996 B1 (LEVIN-MICHAEL ORIT [US]) 1 March 2011 (2011-03-01) column 8, line 34 - column 11, line 22; claims 1-25; figures 6-9 ----- | 1-29 |

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2015/024030

| Patent document cited in search report | | Publication date | Patent family member(s) | | Publication date |
|--|----|------------------|-------------------------|----|------------------|
| US 6560676 | B1 | 06-05-2003 | JP 3812258 | B2 | 23-08-2006 |
| | | | JP 2001195304 | A | 19-07-2001 |
| | | | US 6560676 | B1 | 06-05-2003 |
| US 5732242 | A | 24-03-1998 | US 5732242 | A | 24-03-1998 |
| | | | US 6240488 | B1 | 29-05-2001 |
| | | | US 2002010838 | A1 | 24-01-2002 |
| | | | US 2006149904 | A1 | 06-07-2006 |
| US 2012096227 | A1 | 19-04-2012 | NONE | | |
| US 2011145508 | A1 | 16-06-2011 | US 2011145508 | A1 | 16-06-2011 |
| | | | US 2012203974 | A1 | 09-08-2012 |
| | | | US 2012203975 | A1 | 09-08-2012 |
| WO 0239283 | A2 | 16-05-2002 | DE 10196886 | T1 | 04-12-2003 |
| | | | GB 2383658 | A | 02-07-2003 |
| | | | JP 3810738 | B2 | 16-08-2006 |
| | | | JP 2004513456 | A | 30-04-2004 |
| | | | US 6529998 | B1 | 04-03-2003 |
| | | | WO 0239283 | A2 | 16-05-2002 |
| US 2004268050 | A1 | 30-12-2004 | NONE | | |
| US 2004205298 | A1 | 14-10-2004 | NONE | | |
| US 2007239940 | A1 | 11-10-2007 | CN 101082861 | A | 05-12-2007 |
| | | | US 2007239940 | A1 | 11-10-2007 |
| US 2006174228 | A1 | 03-08-2006 | NONE | | |
| US 7899996 | B1 | 01-03-2011 | NONE | | |

フロントページの続き

(81)指定国 AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,AT,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC,MK,MT,NL,NO,PL,PT,R0,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,NE,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,CZ,DE,DK,DM,D0,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IR,IS,JP,KE,KG,KN,KP,KR,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,MZ,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,SM,ST,SV,SY,TH,TJ,TM,TN,TR,TT,TZ,UA,UG,US

(72)発明者 デイヴィッド・ジョン・バルフレイマン

アメリカ合衆国・ウィスコンシン・53705・マディソン・プラフ・ストリート・3102・アパートメント・4

F ターム(参考) 5B005 JJ13 MM01 NN22 SS12 VV04

【要約の続き】

わち、非専用)キャッシュセットに適用するように構成され得る。