

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6073246号
(P6073246)

(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.

G06F 12/00 (2006.01)

F 1

G06F 12/00 545 A
G06F 12/00 514 E
G06F 12/00 501 B

請求項の数 33 (全 77 頁)

(21) 出願番号 特願2013-547958 (P2013-547958)
 (86) (22) 出願日 平成24年1月10日 (2012.1.10)
 (65) 公表番号 特表2014-506367 (P2014-506367A)
 (43) 公表日 平成26年3月13日 (2014.3.13)
 (86) 國際出願番号 PCT/IL2012/050009
 (87) 國際公開番号 WO2012/104847
 (87) 國際公開日 平成24年8月9日 (2012.8.9)
 審査請求日 平成26年12月25日 (2014.12.25)
 (31) 優先権主張番号 61/431,079
 (32) 優先日 平成23年1月10日 (2011.1.10)
 (33) 優先権主張国 米国(US)

(73) 特許権者 513174357
 ストローン リミテッド
 イスラエル国, 4366350 ラアナナ
 , 4 ハイエツツィラ ストリート, ピー
 . オー. ボックス 2660
 (74) 代理人 100114775
 弁理士 高岡 亮一
 (74) 代理人 100121511
 弁理士 小田 直
 (72) 発明者 ゴードン, ラズ
 イスラエル国, 38444 ハデラ, ピー
 . オー. ボックス 4437, 1 ハクネ
 ッセット ストリート

最終頁に続く

(54) 【発明の名称】大規模記憶システム

(57) 【特許請求の範囲】

【請求項 1】

分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードであって、前記インフラストラクチャ層は、相互接続コンピュータノードを含み、前記相互接続コンピュータノードのうちの少なくとも1つは、1つ以上の記憶関連リソースを備え、前記コンピュータノードは、

統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、前記UDSPエージェントは、少なくとも1つの割り当てを含むタスクを受け取り、

前記少なくとも1つの割り当てについて、前記コンピュータノードに対するグレードと、前記相互接続コンピュータノードのうちの少なくとも1つに対するグレードとを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、前記少なくとも1つの割り当てを実行するそれぞれのコンピュータノードの適合性を示すように計算し、1つ以上の接続される記憶関連リソースを有するコンピュータノードに対するグレードは、前記それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算され、

前記相互接続コンピュータノードのうちの少なくとも1つに対する計算されたグレードが、前記コンピュータノードに対する計算されたグレードより高い場合、前記コンピュータノードの前記少なくとも1つの割り当てを実行する能力に關わらず、前記受け取ったタスクを、前記相互接続コンピュータノードのうちの少なくとも1つにルートす

10

20

る

ように構成される、コンピュータノード。

【請求項 2】

前記 U D S P エージェントはさらに、

更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算し、前記更新されたインフラストラクチャ層は、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応答して作成されることと、

前記計算されたグレードに基づいて、以後のタスクの割り当てのうちの 1 つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、を行うように構成される、請求項 1 に記載のコンピュータノード。10

【請求項 3】

前記 U D S P エージェントはさらに、

更新されたインフラストラクチャ層の少なくとも 1 つの修正された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算するように構成され、前記更新されたインフラストラクチャ層は、少なくとも 1 つの相互接続コンピュータノードの修正に応答して作成され、前記修正は、(i) 少なくとも 1 つの新しい記憶関連リソースが前記それぞれのノードに接続されること、(i i) 少なくとも 1 つの既存の記憶関連リソースが前記それぞれのノードから切断されること、(i i i) 少なくとも 1 つの既存の記憶関連リソースが修正されることのうちの少なくとも 1 つを含み、かつ20

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの 1 つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートするように構成される、請求項 1 に記載のコンピュータノード。

【請求項 4】

前記以後のタスクの割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードに対して行われる、請求項 2 に記載のコンピュータノード。30

【請求項 5】

前記以後のタスクの割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも 1 つの修正された相互接続コンピュータノードに対して行われる、請求項 3 に記載のコンピュータノード。

【請求項 6】

前記以後のタスクの割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも 1 つの前記相互接続コンピュータノードに対して行われる、請求項 2 または 3 のいずれか 1 項に記載のコンピュータノード。40

【請求項 7】

前記更新されたインフラストラクチャ層は、動的に作成される、請求項 2 または 3 のいずれか 1 項に記載のコンピュータノード。

【請求項 8】

分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法であって、前記インフラストラクチャ層は、相互接続コンピュータノードを含み、前記相互接続コンピュータノードのうちの少なくとも 1 つは、1 つ以上の記憶関連リソースを備え、

少なくとも 1 つの割り当てを含むタスクを受け取ることと、

前記少なくとも 1 つの割り当てについて、前記コンピュータノードに対する50

グレードと、前記相互接続コンピュータノードのうちの少なくとも1つに対するグレードとを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、前記少なくとも1つの割り当てを実行するそれぞれのコンピュータノードの適合性を示すように計算し、1つ以上の接続される記憶関連リソースを有するコンピュータノードに対するグレードは、前記それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算されることと、

前記相互接続コンピュータノードのうちの少なくとも1つに対する計算されたグレードが、前記コンピュータノードに対する計算されたグレードより高い場合、前記コンピュータノードの前記少なくとも1つの割り当てを実行する能力に関わらず、前記受け取ったタスクを、前記相互接続コンピュータノードのうちの少なくとも1つにルートすることと、を含む、方法。
10

【請求項9】

更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算し、更新されたインフラストラクチャ層が、少なくとも1つの追加の相互接続コンピュータノードをそれに追加することに応答して作成されることと、

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの1つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、
20
をさらに含む、請求項8に記載の方法。

【請求項10】

更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードについて以後のタスクの割り当てについてのグレードを計算することであって、前記更新されたインフラストラクチャ層が、少なくとも1つの相互接続コンピュータノードの修正に応答して作成され、前記修正が、(i)少なくとも1つの新しい記憶関連リソースが前記それぞれのノードに接続されること、(ii)少なくとも1つの既存の記憶関連リソースが前記それぞれのノードから切断されること、(iii)少なくとも1つの既存の記憶関連リソースが修正されることのうちの少なくとも1つを含む、ことと、

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの1つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、
30
をさらに含む、請求項8に記載の方法。

【請求項11】

前記以後のタスクの割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードに対して行われる、請求項9に記載の方法。

【請求項12】

前記以後のタスクの割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードに対して行われる、請求項10に記載の方法。
40

【請求項13】

前記以後のタスクの割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算することは、前記作成された更新されたインフラストラクチャ層の少なくとも1つの相互接続コンピュータノードに対して行われる、請求項9または10のいずれか1項に記載の方法。

【請求項14】

前記更新されたインフラストラクチャ層は、動的に作成される、請求項9または10
50

のいずれか 1 項に記載の方法。

【請求項 15】

分散記憶システムであって、

相互接続コンピュータノードを含む、インフラストラクチャ層を備え、

前記相互接続コンピュータノードのうちの各 1 つは、統一分散記憶プラットフォーム (UDSP) エージェントを実行するように構成される、少なくとも 1 つの処理リソースを備え、

前記相互接続コンピュータノードのうちの少なくとも 1 つは、1 つ以上の記憶関連リソースを備え、

各所定のコンピュータノードのUDSP エージェントは、

10

少なくとも 1 つの割り当てを含むタスクを受け取ることと、

前記少なくとも 1 つの割り当てについて、前記コンピュータノードに対するグレードと、前記相互接続コンピュータノードのうちの少なくとも 1 つに対するグレードを、各グレードが、少なくとも 1 つのサービスレベル仕様 (SLS) 要件を満たしながら、前記少なくとも 1 つの割り当てを実行するそれぞれのコンピュータノードの適合性を示すように計算し、1 つ以上の接続される記憶関連リソースを有するコンピュータノードに対するグレードは、前記それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算されることと、

前記相互接続コンピュータノードのうちの少なくとも 1 つに対する計算されたグレードが、前記コンピュータノードに対する計算されたグレードより高い場合、前記コンピュータノードの前記少なくとも 1 つの割り当てを実行する能力に関わらず、前記受け取ったタスクを、前記相互接続コンピュータノードのうちの少なくとも 1 つにルートすることと、を行うように構成される、分散記憶システム。

20

【請求項 16】

前記UDSP エージェントはさらに、

更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算し、前記更新されたインフラストラクチャ層は、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応答して作成されることと、

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの 1 つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、を行うように構成される、請求項 15 に記載の分散記憶システム。

30

【請求項 17】

前記UDSP エージェントはさらに、

更新されたインフラストラクチャ層の少なくとも 1 つの修正された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算するように構成され、前記更新されたインフラストラクチャ層は、少なくとも 1 つの相互接続コンピュータノードの修正に応答して作成され、前記修正は、(i) 少なくとも 1 つの新しい記憶関連リソースが前記それぞれのノードに接続されること、(ii) 少なくとも 1 つの既存の記憶関連リソースが前記それぞれのノードから切断されること、(iii) 少なくとも 1 つの既存の記憶関連リソースが修正されることのうちの少なくとも 1 つを含み、かつ

40

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの 1 つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートするように構成される、請求項 15 に記載の分散記憶システム。

【請求項 18】

前記以後のタスクの前記割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラ

50

ストラクチャ層の少なくとも 1 つの前記追加された相互接続コンピュータノードに対して行われる、請求項 16 に記載の分散記憶システム。

【請求項 19】

前記以後のタスクの前記割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも 1 つの修正された相互接続コンピュータノードに対して行われる、請求項 17 に記載の分散記憶システム。

【請求項 20】

前記以後のタスクの前記割り当てについてのグレードを計算しながら、前記更新されたインフラストラクチャ層が作成され、前記計算は、前記作成された更新されたインフラストラクチャ層の少なくとも 1 つの前記相互接続コンピュータノードに対して行われる、請求項 16 または 17 のいずれか 1 項に記載の分散記憶システム。

10

【請求項 21】

前記更新されたインフラストラクチャ層は、動的に作成される、請求項 16 または 17 のいずれか 1 項に記載の分散記憶システム。

【請求項 22】

分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードであって、前記インフラストラクチャ層は、相互接続コンピュータノードを含み、前記コンピュータノードは、

統一分散記憶プラットフォーム (UDSP) エージェントを実行するように構成される、少なくとも 1 つの処理リソースを備え、前記UDSP エージェントは、

20

少なくとも 1 つの論理的記憶エンティティを参照するユーザ定義された記憶関連の要件と、前記相互接続コンピュータノードに接続された 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも 1 つのサービスレベル仕様 (SLS) を受け取ることと、

少なくとも、前記少なくとも 1 つの SLS および前記記憶関連リソースパラメータデータに基づいて、前記分散記憶システムの構成を計算することと、

前記計算された構成に従って、前記記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることと、

少なくとも 1 つの割り当てを含むタスクを受け取ることと、

30

前記少なくとも 1 つの割り当てについて、前記コンピュータノードに対するグレードと、前記相互接続コンピュータノードのうちの少なくとも 1 つに対するグレードとを、各グレードが、少なくとも 1 つのサービスレベル仕様 (SLS) 要件を満たしながら、前記少なくとも 1 つの割り当てを実行するそれぞれのコンピュータノードの適合性を示すように計算し、1 つ以上の接続される記憶関連リソースを有するコンピュータノードに対するグレードは、前記それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算されることと、

前記相互接続コンピュータノードのうちの少なくとも 1 つに対する計算されたグレードが、前記コンピュータノードに対する計算されたグレードより高い場合、前記コンピュータノードの前記少なくとも 1 つの割り当てを実行する能力に関わらず、前記受け取ったタスクを、前記相互接続コンピュータノードのうちの少なくとも 1 つにルートすることとを行うように構成される、コンピュータノード。

40

【請求項 23】

前記UDSP エージェントはさらに、

前記分散記憶システムの動的挙動に関する動的挙動パラメータデータを受け取ることと、

少なくとも 1 つの SLS が違反されたときに、少なくとも、前記少なくとも 1 つの SLS、前記記憶関連リソースパラメータデータ、および前記動的挙動パラメータデータに基づいて、前記分散記憶システムの再構成を計算することと、

前記計算された再構成に従って、前記記憶関連リソースのうちの 1 つの少な

50

くとも一部を自動的に割り付けることと、を行うように構成される、請求項 22 に記載のコンピュータノード。

【請求項 24】

前記 U D S P エージェントはさらに、

更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算し、前記更新されたインフラストラクチャ層は、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応答して作成されることと、

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの 1 つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、を行うように構成される、請求項 22 に記載のコンピュータノード。10

【請求項 25】

前記 U D S P エージェントはさらに、前記構成を判定する失敗に応答して、ユーザに、失敗通知、または構成の計算の成功を可能にするであろう、少なくとも 1 つの追加の記憶関連リソースの追加を示す推奨を提供するように構成される、請求項 22 に記載のコンピュータノード。

【請求項 26】

分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法であって、前記インフラストラクチャ層は、相互接続コンピュータノードを含み、前記相互接続コンピュータノードのうちの少なくとも 1 つは、1 つ以上の記憶関連リソースを備え。20

少なくとも 1 つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、前記相互接続コンピュータノードに接続された 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータと、を備える、少なくとも 1 つのサービスレベル仕様 (S L S) を受け取ることと、

少なくとも、前記少なくとも 1 つの S L S および前記記憶関連リソースパラメータデータに基づいて、前記分散記憶システムの構成を計算することと、

前記計算された構成に従って、前記記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることと。30

少なくとも 1 つの割り当てを含むタスクを受け取ることと、

前記少なくとも 1 つの割り当てについて、前記コンピュータノードに対するグレードと、前記相互接続コンピュータノードのうちの少なくとも 1 つに対するグレードとを、各グレードが、少なくとも 1 つのサービスレベル仕様 (S L S) 要件を満たしながら、前記少なくとも 1 つの割り当てを実行するそれぞれのコンピュータノードの適合性を示すように計算し、1 つ以上の接続される記憶関連リソースを有するコンピュータノードに対するグレードは、前記それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算されることと、

前記相互接続コンピュータノードのうちの少なくとも 1 つに対する計算されたグレードが、前記コンピュータノードに対する計算されたグレードより高い場合、前記コンピュータノードの前記少なくとも 1 つの割り当てを実行する能力に関わらず、前記受け取ったタスクを、前記相互接続コンピュータノードのうちの少なくとも 1 つにルートすることと

を含む、方法。

【請求項 27】

前記分散記憶システムの動的拳動に関する動的拳動パラメータデータを受け取ることと、

少なくとも 1 つの S L S が違反されたときに、少なくとも、前記少なくとも 1 つの S L S 、前記記憶関連リソースパラメータデータ、および前記動的拳動パラメータデータに基づいて、前記分散記憶システムの再構成を計算することと、50

前記計算された再構成に従って、前記記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることと、
をさらに含む、請求項 2 6 に記載の方法。

【請求項 2 8】

更新されたインフラストラクチャ層が、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応答して作成される、前記更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算することと、

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの 1 つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、
をさらに含む、請求項 2 6 に記載の方法。 10

【請求項 2 9】

前記構成を判定する失敗に応答して、ユーザに、失敗通知、または構成の計算の成功を可能にするであろう、少なくとも 1 つの追加のリソースの追加を示す推奨を提供することをさらに含む、請求項 2 6 に記載の方法。

【請求項 3 0】

分散記憶システムであって、

相互接続コンピュータノードを含む、インフラストラクチャ層を備え、

前記相互接続コンピュータノードのうちの各 1 つは、統一分散記憶プラットフォーム (U D S P) エージェントを実行するように構成される、少なくとも 1 つの処理リソースを備え、 20

各所定のコンピュータノードの U D S P エージェントは、
少なくとも 1 つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、前記相互接続コンピュータノードに接続された 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも 1 つのサービスレベル仕様 (S L S) を受け取ることと、

少なくとも、前記少なくとも 1 つの S L S 、および前記記憶関連リソースパラメータデータに基づいて、前記分散記憶システムの構成を計算することと、

前記計算された構成に従って、前記記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることと、 30

少なくとも 1 つの割り当てを含むタスクを受け取ることと、

前記少なくとも 1 つの割り当てについて、前記コンピュータノードに対するグレードと、前記相互接続コンピュータノードのうちの少なくとも 1 つに対するグレードとを、各グレードが、少なくとも 1 つのサービスレベル仕様 (S L S) 要件を満たしながら、前記少なくとも 1 つの割り当てを実行するそれぞれのコンピュータノードの適合性を示すように計算し、1 つ以上の接続される記憶関連リソースを有するコンピュータノードに対するグレードは、前記それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算されることと、

前記相互接続コンピュータノードのうちの少なくとも 1 つに対する計算されたグレードが、前記コンピュータノードに対する計算されたグレードより高い場合、前記コンピュータノードの前記少なくとも 1 つの割り当てを実行する能力に関わらず、前記受け取ったタスクを、前記相互接続コンピュータノードのうちの少なくとも 1 つにルートすることと

を行うように構成される、分散記憶システム。

【請求項 3 1】

前記 U D S P エージェントはさらに、

前記分散記憶システムの動的挙動に関する動的挙動パラメータデータを受け取ることと、

少なくとも 1 つの S L S が違反されたときに、少なくとも、前記少なくとも 50

1つのSLS、前記記憶関連リソースパラメータデータ、および前記動的挙動パラメータデータに基づいて、前記分散記憶システムの再構成を計算することと、

前記計算された再構成に従って、前記記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けることと、を行うように構成される、
請求項30記載の分散記憶システム。

【請求項32】

前記UDSPエージェントはさらに、

更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードに対する以後のタスクの割り当てについてのグレードを計算し、前記更新されたインフラストラクチャ層は、少なくとも1つの追加の相互接続コンピュータノードをそれに追加することに応答して作成されることと、10

前記計算されたグレードに基づいて、以後のタスクの前記割り当てのうちの1つ以上を実行するか、または前記以後のタスクを前記更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、を行うように構成される、
請求項30に記載の分散記憶システム。

【請求項33】

前記UDSPエージェントはさらに、前記構成を判定する失敗に応答して、ユーザに、失敗通知、または構成の計算の成功を可能にするであろう、少なくとも1つの追加の記憶関連リソースの追加を示す推奨を提供するように構成される、請求項30に記載の分散記憶システム。20

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、大規模記憶システムに関し、具体的には、そのようなシステムを実装するための装置および方法に関する。

【背景技術】

【0002】

分散記憶システムは、ネットワークが容量および速度を増大させるにつれて、過去10年にわたって急速に発展してきた。ネットワークがローカルエリアネットワーク(LAN)から大域的な広域ネットワーク(WAN)に拡張すると、事業がより大域的に分散しつつあり、分散記憶システムが遠隔の地理的場所にわたってデータ記憶およびアクセスを提供する要求をもたらしている。したがって、当技術分野では、汎用ネットワークにわたってデータ記憶を分散させるための新しい方法およびシステムの必要性がある。30

【0003】

背景として本開示主題に関連すると見なされる従来技術の参考文献が、以下に記載される。本明細書の参考文献の確認は、これらが本開示主題の特許性に何らかの点で関連すると意味するものとして推論されるものではない。

【0004】

米国特許公開第2009/0070337号「Apparatus and Method for a Distributed Storage Global Data base」は、「所与のデータ要素に対する要求が地理的慣性を負う、データ要素の分散を管理するための地理的分散記憶システム。地理的分散記憶システムは、地理的分散サイトを備え、それぞれ、データ要素と、データ要素のうちの1つに関する要求を受けるためのローカルアクセスポイントとを含む、大域的コヒーレント分散データベースの一部分をローカルに記憶するためのサイト記憶ユニットを備える。地理的分散記憶システムは、そこから要求が受け取られる、地理的分散サイトのうちの第1のサイトにおけるローカルアクセスポイントに少なくとも1つの要求されたデータ要素を転送し、第1のサイトにおいて少なくとも1つの要求されたデータ要素を記憶し、それにより、分散データベースの大4050

域的一貫性を維持しながら、第1のサイトからの将来の要求のためにデータ要素へのローカルアクセス可能性を提供するためのデータ管理モジュールを備える。」ことに関する。

【0005】

米国特許第5987505号「Remote Access and Geographically Distributed Computers in a Globally Addressable Storage Environment」は、「コンピュータシステムは、データがコンピュータハードディスク等の永続記憶デバイスおよび他の伝統的にアドレス指定不可能なデータ記憶デバイス上に記憶されるときさえも対処することによって、複数のネットワークコンピュータがデータにアクセスすることを可能にする、大域的アドレス可能記憶環境を採用する。コンピュータは、単一のコンピュータネットワーク上、または広域ネットワーク(WAN)によって連結される2つのローカルエリアネットワーク(LAN)等の複数の相互接続コンピュータネットワーク上に位置することができる。大域的アドレス可能記憶環境は、複数のネットワーク上の種々のコンピュータによって、かつそれらの間で、データがアクセスおよび共有されることを可能にする。」ことに関する。

【0006】

International Journal of Computer Applications 2010(0975-8887), Volume 1-No.22, “Unified Virtual Storage: Virtualization of Distributed Storage in a Network”, Ms. S. V. Patil et al.は、「ネットワーク上で接続されるデスクトップマシン上のディスクの空き容量を効率的に利用する方法。現在、多くのネットワークでは、クライアントノードのローカルディスクは、散発的にしか使用されていない。これは、ネットワーク内のデータ記憶を効率的に管理し、かつLAN内のデスクトップマシン上のディスク空間の共有のためのソフトウェアサポートを提供しようとする試行である。現在の状況では、従来のサーバ上の記憶拡張は、最大拡張制限、費用のかかる問題のような制約を有し、ハードウェア交換、機能向上の場合には、データの手動再配置が乱雑になる。UVS(統一仮想記憶)は、ネットワーク上で接続されるデスクトップマシン上の自由に利用可能なディスク空間を効率的に利用しようとする試行である。その目的は、クライアントノード上の空間を効率的に利用し、それにより、空間の浪費を回避するように、ネットワークサーバ上のデータトラフィックの負荷を低減することである。それはまた、記憶拡張のハードウェア制限を排除し、データ記憶の場所透過性を提供する。UVSの主要な利点は、それを既存のインフラストラクチャ(ローカルエリアネットワークシステム)にシームレスに組み込むことができる。仮想記憶は、拡張性のあるアーキテクチャをサポートし、事実上無限である。クライアントノードは、異なるサーバにわたる分散記憶のための一点アクセスとして統一仮想ドライブを使用し、それにより、サーバの個別アドレス指定を排除することができる。ネットワークおよび性能によって接続されるUVSサーバ上で実装されるプロトタイプの性能は、中央集権型システムよりも良好であり、フレームワークのオーバーヘッドは、高負荷中でさえも適度である。」と説明する。

【0007】

米国特許公開第2011/0153770号、「Dynamic Structural Management of a Distributed Caching Infrastructure」は、「n-階層分散キャッシングインフラストラクチャの動的構造管理のための方法、システム、およびコンピュータプログラム製品。本発明の実施形態では、n-階層分散キャッシングインフラストラクチャの動的構造管理の方法は、n-階層キャッシング内のある階層ノードの中に配設される複数のキャッシングサーバへの通信接続を確立することと、n-階層キャッシングのそれぞれの階層ノードの中のキャッシングサーバのうちのそれに対する性能測定基準を収集することと、閾値を横断するn-階層の階層ノードのうちの対応する1つの中の特定のキャッシングリソースの特性を識別することと、識別された特性を考慮するように、特定のキャッシングリソースを含む一式の

キャッシュリソースを動的に構造化することとを含む。」に関する。

【発明の概要】

【0008】

本開示主題の態様によれば、相互接続コンピュータノードを含むインフラストラクチャ層に接続されるように構成されるコンピュータノードが提供され、相互接続コンピュータノードのうちの少なくとも1つは、1つ以上の記憶関連リソースを備え、コンピュータノードは、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される少なくとも1つの処理リソースを備え、UDSPエージェントは、少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも1つについて割り当てのうちの少なくとも1つのグレードを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードは、該当する場合、それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関するパラメータデータにも基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの1つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、

更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャ層は、少なくとも1つの追加の相互接続コンピュータノードをそれに追加することに応じて作成されることと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることと、を行うように構成される。

【0009】

本開示主題のある特定の実施例によれば、さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードについて行われる、コンピュータノードが提供される。

【0010】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの相互接続コンピュータノードについて行われる、コンピュータノードが提供される。

【0011】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャが動的に作成される、コンピュータノードが提供される。

【0012】

本開示主題のある特定の実施例によれば、その上さらに、グレードが保留中の割り当てのみに対して計算される、コンピュータノードが提供される。

【0013】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも1つの記憶関連リソースを含む、コンピュータノードが提供される。

【0014】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリのうちのそれぞれの少なくとも1つの記憶関連リソースを含む、コンピュータノードが提供される。

【0015】

10

20

30

40

50

本開示主題のある特定の実施例によれば、その上さらに、割り当てグレードが最適化エンジンによって計算される、コンピュータノードが提供される。

【0016】

本開示主題のある特定の実施例によれば、その上さらに、最適化エンジンが、線形計画法、焼き鈍し法、遺伝的アルゴリズム等の最適化技法のうちの1つ以上を使用する、コンピュータが提供される。

【0017】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の最適化技法が、ヒューリスティックスまたは近似値を使用する、コンピュータノードが提供される。

【0018】

本開示主題のある特定の実施例によれば、その上さらに、より好適なコンピュータノードが、計算されたグレードに基づく、最も好適なコンピュータノードである、コンピュータノードが提供される。

【0019】

本開示主題のある特定の実施例によれば、その上さらに、UDS Pエージェントはさらに、計算された割り当てグレードに基づいて、統合グレードを計算するように構成され、割り当てのうちの1つ以上の実行、またはタスクをより好適なコンピュータノードにルートすることは、計算された統合グレードに基づく、コンピュータノードが提供される。

【0020】

本開示主題のある特定の実施例によれば、その上さらに、タスクがクライアントサーバから受け取られる、コンピュータノードが提供される。

【0021】

本開示主題のある特定の実施例によれば、その上さらに、タスクが、ゲートウェイリソースを通してクライアントサーバから受け取られる、コンピュータノードが提供される。

【0022】

本開示主題のある特定の実施例によれば、その上さらに、UDS Pエージェントはさらに、コンピュータノードの現在の状態、または該当する場合、コンピュータノードに接続された1つ以上の記憶関連リソースの現在の状態を示す、少なくとも1つのパラメータを監視し、少なくとも1つの監視されたパラメータへの変更を示す通知を、相互接続コンピュータノードのうちの少なくとも1つに伝搬するように構成される、コンピュータノードが提供される。

【0023】

本開示主題のある特定の実施例によれば、その上さらに、UDS Pエージェントはさらに、更新されたインフラストラクチャの少なくとも1つの修正された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算するように構成され、更新されたインフラストラクチャ層は、少なくとも1つの修正された相互接続コンピュータノードを含み、修正は、(i)少なくとも1つの新しい記憶関連リソースがそれぞれのノードに接続されること、(ii)少なくとも1つの既存の記憶関連リソースがそれぞれのノードから切断されること、(iii)少なくとも1つの既存の記憶関連リソースが修正されることのうちの少なくとも1つを含み、それに接続された新しい記憶関連リソースを有する、またはそこから切断された記憶関連リソースを有する、少なくとも1つの相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算するように構成され、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、

または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートするように構成される、コンピュータノードが提供される。

【0024】

本開示主題のある特定の実施例によれば、その上さらに、UDS Pエージェントはさらに、少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された1つ以上の記憶関連リソースに関する記憶関連

10

20

30

40

50

リソースパラメータデータと、を備える、少なくとも1つのサービスレベル仕様(SLS)を受け、少なくとも、少なくとも1つのSLSおよび記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算し、計算された構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される、コンピュータノードが提供される。

【0025】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、分散記憶システムの動的挙動に関する動的挙動パラメータデータを受け、少なくとも1つのSLSが違反されたときに、少なくとも、少なくとも1つのSLS、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、記憶システムの再構成を計算し、計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される、コンピュータノードが提供される。

10

【0026】

本開示主題のある特定の実施例によれば、その上さらに、相互接続コンピュータノードを含むインフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法が提供され、相互接続コンピュータノードのうちの少なくとも1つは、1つ以上の記憶関連リソースを備え、本方法は、少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも1つについて割り当てのうちの少なくとも1つのグレードを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードが、該当する場合、それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関するパラメータデータにも基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの1つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、更新されたインフラストラクチャ層が、少なくとも1つの追加の相互接続コンピュータノードをそれに追加することに応じて作成されると、更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算することと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることとを含む。

20

【0027】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードについて行われる、方法が提供される。

30

【0028】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算することは、作成された更新されたインフラストラクチャ層の少なくとも1つの相互接続コンピュータノードについて行われる、方法が提供される。

40

【0029】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャが動的に作成される、方法が提供される。

【0030】

本開示主題のある特定の実施例によれば、その上さらに、グレードが保留中の割り当てのみに対して計算される、方法が提供される。

【0031】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連

50

リソースカテゴリの少なくとも 1 つの記憶関連リソースを含む、方法が提供される。

【 0 0 3 2 】

本開示主題のある特定の実施例によれば、その上さらに、1 つ以上の記憶関連リソースは、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリのうちのそれぞれの少なくとも 1 つの記憶関連リソースを含む、方法が提供される。

【 0 0 3 3 】

本開示主題のある特定の実施例によれば、その上さらに、グレードを計算することが、最適化エンジンによって行われる、方法が提供される。

【 0 0 3 4 】

本開示主題のある特定の実施例によれば、その上さらに、最適化エンジンが、線形計画法、焼き鈍し法、遺伝的アルゴリズム等の最適化技法のうちの 1 つ以上を使用する、方法が提供される。

【 0 0 3 5 】

本開示主題のある特定の実施例によれば、その上さらに、1 つ以上の最適化技法が、ヒューリスティックスまたは近似値を使用する、方法が提供される。

【 0 0 3 6 】

本開示主題のある特定の実施例によれば、その上さらに、より好適なコンピュータノードが、計算されたグレードに基づく、最も好適なコンピュータノードである、方法が提供される。

10

【 0 0 3 7 】

本開示主題のある特定の実施例によれば、その上さらに、計算された割り当てグレードに基づいて、統合グレードを計算することをさらに含む、方法が提供され、割り当てのうちの 1 つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることは、計算された統合グレードに基づく。

20

【 0 0 3 8 】

本開示主題のある特定の実施例によれば、その上さらに、タスクがクライアントサーバから受け取られる、方法が提供される。

【 0 0 3 9 】

本開示主題のある特定の実施例によれば、その上さらに、タスクが、ゲートウェイリソースを通じてクライアントサーバから受け取られる、方法が提供される。

30

【 0 0 4 0 】

本開示主題のある特定の実施例によれば、その上さらに、コンピュータノードの現在の状態、または該当する場合、コンピュータノードに接続された 1 つ以上の記憶関連リソースの現在の状態を示す少なくとも 1 つのパラメータを監視することと、少なくとも 1 つの監視されたパラメータへの変更を示す通知を、相互接続コンピュータノードのうちの少なくとも 1 つに伝搬することとをさらに含む、方法が提供される。

【 0 0 4 1 】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャ層の少なくとも 1 つの修正された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャが、少なくとも 1 つの修正された相互接続コンピュータノードを含み、修正が、(i) 少なくとも 1 つの新しい記憶関連リソースがそれぞれのノードに接続されること、(i i) 少なくとも 1 つの既存の記憶関連リソースがそれぞれのノードから切断されること、(i i i) 少なくとも 1 つの既存の記憶関連リソースが修正されることのうちの少なくとも 1 つを含み、それに接続された新しい記憶関連リソースを有する、またはそこから切断された記憶関連リソースを有する、少なくとも 1 つの相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算することと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの 1 つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることとをさらに含む、方法が提供される

40

50

。

【 0 0 4 2 】

本開示主題のある特定の実施例によれば、その上さらに、少なくとも 1 つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも 1 つのサービスレベル仕様 (S L S) を受けることと、少なくとも、少なくとも 1 つの S L S および記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算することと、計算された構成に従って、記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることをさらに含む、方法が提供される。

【 0 0 4 3 】

10

本開示主題のある特定の実施例によれば、その上さらに、分散記憶システムの動的拳動に関する動的拳動パラメータデータを受けることと、少なくとも 1 つの S L S が違反されたときに、少なくとも、少なくとも 1 つの S L S 、記憶関連リソースパラメータデータ、および動的拳動パラメータデータに基づいて、分散記憶システムの再構成を計算することと、計算された再構成に従って、記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることをさらに含む、方法が提供される。

【 0 0 4 4 】

本開示主題のある特定の実施例によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層を備える、分散記憶システムが提供され、相互接続コンピュータノードの各々は、統一分散記憶プラットフォーム (U D S P) エージェントを実行するように構成される少なくとも 1 つの処理リソースを備え、相互接続コンピュータノードのうちの少なくとも 1 つは、1 つ以上の記憶関連リソースを備え、 U D S P エージェントは、少なくとも 1 つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも 1 つについて割り当てのうちの少なくとも 1 つのグレードを、各グレードが、少なくとも 1 つのサービスレベル仕様 (S L S) 要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードは、該当する場合、それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにもに基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの 1 つ以上を実行するか、またはタスクをインフラストラクチャ層のより好適なコンピュータノードにルートするように構成され、更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャ層は、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応じて作成されることと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの 1 つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることを行うように構成される。

【 0 0 4 5 】

30

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードについて行われる、分散記憶システムが提供される。

【 0 0 4 6 】

40

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも 1 つの相互接続コンピュータノードについて行われる、分散記憶システムが提供される。

【 0 0 4 7 】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャが動的に作成される、分散記憶システムが提供される。

50

【 0 0 4 8 】

本開示主題のある特定の実施例によれば、その上さらに、グレードが保留中の割り当てのみに対して計算される、分散記憶システムが提供される。

【 0 0 4 9 】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも1つの記憶関連リソースを含む、分散記憶システムが提供される。

【 0 0 5 0 】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリのうちのそれぞれの少なくとも1つの記憶関連リソースを含む、分散記憶システムが提供される。 10

【 0 0 5 1 】

本開示主題のある特定の実施例によれば、その上さらに、割り当てグレードが、最適化エンジンによって計算される、分散記憶システムが提供される。

【 0 0 5 2 】

本開示主題のある特定の実施例によれば、その上さらに、最適化エンジンが、線形計画法、焼き鈍し法、遺伝的アルゴリズム等の最適化技法のうちの1つ以上を使用する、分散記憶システムが提供される。 20

【 0 0 5 3 】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の最適化技法が、ヒューリスティックスまたは近似値を使用する、分散記憶システムが提供される。

【 0 0 5 4 】

本開示主題のある特定の実施例によれば、その上さらに、より好適なコンピュータノードが、計算されたグレードに基づく、最も好適なコンピュータノードである、分散記憶システムが提供される。

【 0 0 5 5 】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、計算された割り当てグレードに基づいて、統合グレードを計算するように構成され、割り当てのうちの1つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることは、計算された統合グレードに基づく、分散記憶システムが提供される。 30

【 0 0 5 6 】

本開示主題のある特定の実施例によれば、その上さらに、タスクがクライアントサーバから受け取られる、分散記憶システムが提供される。

【 0 0 5 7 】

本開示主題のある特定の実施例によれば、その上さらに、タスクがゲートウェイリソースを通してクライアントサーバから受け取られる、分散記憶システムが提供される。

【 0 0 5 8 】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、コンピュータノードの現在の状態、または該当する場合、コンピュータノードに接続された1つ以上の記憶関連リソースの現在の状態を示す、少なくとも1つのパラメータを監視し、少なくとも1つの監視されたパラメータへの変更を示す通知を、相互接続コンピュータノードのうちの少なくとも1つに伝搬するように構成される、分散記憶システムが提供される。 40

【 0 0 5 9 】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、更新されたインフラストラクチャの少なくとも1つの修正された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算するように構成され、更新さ 50

れたインフラストラクチャ層は、少なくとも1つの修正された相互接続コンピュータノードを含み、修正は、(i)少なくとも1つの新しい記憶関連リソースがそれぞれのノードに接続されること、(ii)少なくとも1つの既存の記憶関連リソースがそれぞれのノードから切断されること、(iii)少なくとも1つの既存の記憶関連リソースが修正されることのうちの少なくとも1つを含み、それに接続された新しい記憶関連リソースを有する、またはそこから切断された記憶関連リソースを有する、少なくとも1つの相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算するように構成され、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートするように構成される、分散記憶システムが提供される。

10

【0060】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件を備える、少なくとも1つのサービスレベル仕様(SLS)を受け、少なくとも、少なくとも1つのSLSおよび記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算し、計算された構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される、分散記憶システムが提供される。

【0061】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、分散記憶システムの動的挙動に関する動的挙動パラメータデータを受け、少なくとも1つのSLSが違反されたときに、少なくとも、少なくとも1つのSLS、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、分散記憶システムの再構成を計算し、かつ計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される、分散記憶システムが提供される。

20

【0062】

本開示主題のある特定の実施例によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層を備える、分散記憶システムが提供され、相互接続コンピュータノードのうちの各1つは、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、相互接続コンピュータノードのうちの少なくとも1つは、1つ以上の記憶関連リソースを備え、UDSPエージェントは、少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも1つについて割り当てのうちの少なくとも1つのグレードを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードは、該当する場合、それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにもに基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの1つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャは、少なくとも1つの相互接続コンピュータノードの修正に応じて作成され、修正は、(i)少なくとも1つの新しい記憶関連リソースがそれぞれのノードに接続されること、(ii)少なくとも1つの既存の記憶関連リソースがそれぞれのノードから切断されること、(iii)少なくとも1つの既存の記憶関連リソースが修正されることのうちの少なくとも1つを含み、それに接続された新しい記憶関連リソースを有し、またはそこから切断された記憶関連リソースを有し、更新されたインフラストラクチャ層を生じる、少なくとも1つの相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算することと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることを行うよ

30

40

50

うに構成される。

【0063】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、分散記憶システムの動的挙動に関する動的挙動パラメータデータを受け、少なくとも1つのSLS要件が違反されたときに、少なくとも、少なくとも1つのSLS要件、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、分散記憶システムの再構成を計算し、かつ計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される、分散記憶システムが提供される。

【0064】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードについて行われる、分散記憶システムが提供される。

【0065】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの相互接続コンピュータノードについて行われる、分散記憶システムが提供される。

【0066】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャが、動的に作成される、分散記憶システムが提供される。

【0067】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも1つの記憶関連リソースを含む、分散記憶システムが提供される。

【0068】

本開示主題のある特定の実施例によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層に接続されるように構成される、コンピュータノードが提供され、相互接続コンピュータノードのうちの少なくとも1つは、1つ以上の記憶関連リソースを備え、コンピュータノードは、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、UDSPエージェントは、少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも1つについて割り当てのうちの少なくとも1つのグレードを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すようにし計算し、グレードは、該当する場合、それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関するパラメータデータにも基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの1つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、更新されたインフラストラクチャの少なくとも1つの修正された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャ層は、少なくとも1つの相互接続コンピュータノードの修正に応じて作成され、修正は、(i)少なくとも1つの新しい記憶関連リソースがそれぞれのノードに接続されること、(ii)少なくとも1つの既存の記憶関連リソースがそれぞれのノードから切断されること、(iii)少なくとも1つの既存の記憶関連リソースが修正されることのうちの少なくとも1つを含み、それに接続された新しい記憶関連リソースを有し、またはそこから切断された記憶関連リソースを有し、更新されたインフラストラクチャ層を生じる、少なくとも1つの相互接続コンピュータノードについて以後のタスク

10

20

30

40

50

の割り当てのグレードを計算することと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることを行いうように構成される。

【0069】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、記憶システムの動的挙動に関する動的挙動パラメータデータを受け、少なくとも1つのSLS要件が違反されたときに、少なくとも、少なくとも1つのSLS要件、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、分散記憶システムの再構成を計算し、かつ計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される、コンピュータノードシステムが提供される。

【0070】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードについて行われる、コンピュータノードシステムが提供される。

【0071】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの前記割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算は、作成された更新されたインフラストラクチャ層の少なくとも1つの相互接続コンピュータノードについて行われる、コンピュータノードシステムが提供される。

【0072】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャが、動的に作成される、コンピュータノードシステムが提供される。

【0073】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも1つの記憶関連リソースを含む、コンピュータノードシステムが提供される。

【0074】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法が提供され、相互接続コンピュータノードのうちの少なくとも1つは、1つ以上の記憶関連リソースを備え、本方法は、

少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも1つについて割り当てのうちの少なくとも1つのグレードを、各グレードが、少なくとも1つのサービスレベル仕様(SLS)要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードが、該当する場合、それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータにも基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの1つ以上を実行するか、またはタスクをインフラストラクチャ層のより好適なコンピュータノードにルートし、更新されたインフラストラクチャ層が、少なくとも1つの相互接続コンピュータノードの修正に応じて作成され、修正が、(i)少なくとも1つの新しい記憶関連リソースがそれぞれのノードに接続されること、(ii)少なくとも1つの既存の記憶関連リソースがそれぞれのノードから切断されること、(iii)少なくとも1つの既存の記憶関連リソースが修正されることのうちの少なくとも1つを含む、更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算することと、それに接続された

10

20

30

40

50

新しい記憶関連リソースを有し、またはそこから切断された記憶関連リソースを有し、更新されたインフラストラクチャ層を生じる、少なくとも1つの相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算することと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることとを含む。

【0075】

本開示主題のある特定の実施例によれば、その上さらに、記憶システムの動的挙動に関する動的挙動パラメータデータを受けることと、少なくとも1つのSLS要件が違反されたときに、少なくとも、少なくとも1つのSLS要件、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、分散記憶システムの再構成を計算することと、計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けることとをさらに含む、方法が提供される。10

【0076】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算することは、作成された更新されたインフラストラクチャ層の少なくとも1つの修正された相互接続コンピュータノードについて行われる、方法が提供される。

【0077】

本開示主題のある特定の実施例によれば、その上さらに、以後のタスクの割り当てのグレードを計算しながら、更新されたインフラストラクチャ層が作成され、計算することは、作成された更新されたインフラストラクチャ層の少なくとも1つの相互接続コンピュータノードについて行われる、方法が提供される。20

【0078】

本開示主題のある特定の実施例によれば、その上さらに、更新されたインフラストラクチャが、動的に作成される、方法が提供される。

【0079】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースは、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも1つの記憶関連リソースを含む、方法が提供される。30

【0080】

本開示主題のある特定の実施例によれば、その上さらに、分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードが提供され、インフラストラクチャ層は、相互接続コンピュータノードを含み、コンピュータノードは、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、UDSPエージェントは、
少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも1つのサービスレベル仕様(SLS)を受け、少なくとも、少なくとも1つのSLSおよび記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算し、40

計算された構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付け、分散記憶システムの動的挙動に関する動的挙動パラメータデータを受け、少なくとも1つのSLSが違反されたときに、少なくとも、少なくとも1つのSLS、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、分散記憶システムの再構成を計算し、かつ計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される。

【0081】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノー50

ドのうちの少なくとも 1 つについて割り当てのうちの少なくとも 1 つのグレードを、各グレードが、SLS の少なくとも 1 つのユーザ定義された記憶関連要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードは、該当する場合、それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関するパラメータデータにも基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの 1 つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャ層は、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応じて作成され、計算されたグレードに基づいて、以後のタスクの割り当てのうちの 1 つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることを行うように構成される、コンピュータノードが提供される。10

【0082】

本開示主題のある特定の実施例によれば、その上さらに、構成が最適化エンジンによって計算される、コンピュータノードが提供される。

【0083】

本開示主題のある特定の実施例によれば、その上さらに、最適化エンジンが、線形計画法、焼き鈍し法、遺伝的アルゴリズム等の最適化技法のうちの 1 つ以上を使用する、コンピュータノードが提供される。20

【0084】

本開示主題のある特定の実施例によれば、その上さらに、1 つ以上の最適化技法が、ヒューリスティックスまたは近似値を使用する、コンピュータノードが提供される。

【0085】

本開示主題のある特定の実施例によれば、その上さらに、記憶要件は、場所、ローカル保護レベル、バックアップ保持ポリシー、遠隔保護レベル、性能レベル、暗号化レベル、重複排除、圧縮、記憶方法のうちの少なくとも 1 つを含む、コンピュータノードが提供される。

【0086】

本開示主題のある特定の実施例によれば、その上さらに、1 つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも 1 つの記憶関連リソースを含む、コンピュータノードが提供される。30

【0087】

本開示主題のある特定の実施例によれば、その上さらに、記憶関連リソースが、少なくとも 2 つの場所の間で分散している、コンピュータノードが提供される。

【0088】

本開示主題のある特定の実施例によれば、その上さらに、構成が、記憶システムリソースをオーバーコミットすることを含む、コンピュータノードが提供される。40

【0089】

本開示主題のある特定の実施例によれば、その上さらに、UDSP エージェントはさらに、構成の判定の失敗に応答して、ユーザに、失敗通知、または構成の計算の成功を可能にするであろう、少なくとも 1 つの追加の記憶関連リソースの追加を示す推奨を提供するように構成される、コンピュータノードが提供される。

【0090】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法が提供され、相互接続コンピュータノードのうちの少なくとも 1 つは、1 つ以上の記憶関連リソースを備え、本方法は、50

少なくとも 1 つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された 1 つ以上の記憶関連リソースに関する記憶関連リソースパラメータと、を備える、少なくとも 1 つのサービスレベル仕様 (SLS) を受けることと、少なくとも、少なくとも 1 つの SLS および記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算することと、

計算された構成に従って、記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることと、分散記憶システムの動的挙動に関する動的挙動パラメータデータを受けることと、少なくとも 1 つの SLS が違反されたときに、少なくとも、少なくとも 1 つの SLS、記憶関連リソースパラメータデータ、および動的挙動パラメータデータに基づいて、分散記憶システムの再構成を計算することと、計算された再構成に従って、記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることとを含む。10

【0091】

本開示主題のある特定の実施例によれば、その上さらに、少なくとも 1 つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも 1 つについて割り当てのうちの少なくとも 1 つのグレードを、各グレードが、少なくとも 1 つのサービスレベル仕様 (SLS) 要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードが、該当する場合、それぞれのコンピュータノードに接続される 1 つ以上の記憶関連リソースに関するパラメータデータにも基づいて計算されことと、計算されたグレードに基づいて、割り当てのうちの 1 つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、更新されたインフラストラクチャ層が、少なくとも 1 つの追加の相互接続コンピュータノードをそれに追加することに応じて作成される、更新されたインフラストラクチャ層の少なくとも 1 つの追加された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算することと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの 1 つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることとをさらに含む、方法が提供される。20

【0092】

本開示主題のある特定の実施例によれば、その上さらに、構成を計算することが、最適化エンジンによって行われる、方法が提供される。30

【0093】

本開示主題のある特定の実施例によれば、その上さらに、最適化エンジンが、線形計画法、焼き鈍し法、遺伝的アルゴリズム等の最適化技法のうちの 1 つ以上を使用する、方法が提供される。

【0094】

本開示主題のある特定の実施例によれば、その上さらに、1 つ以上の最適化技法が、ヒューリスティックスまたは近似値を使用する、方法が提供される。

【0095】

本開示主題のある特定の実施例によれば、その上さらに、記憶要件が、場所、ローカル保護レベル、バックアップ保持ポリシー、遠隔保護レベル、性能レベル、暗号化レベル、重複排除、圧縮、記憶方法のうちの少なくとも 1 つを含む、方法が提供される。40

【0096】

本開示主題のある特定の実施例によれば、その上さらに、1 つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも 1 つの記憶関連リソースを含む、方法が提供される。

【0097】

本開示主題のある特定の実施例によれば、その上さらに、記憶関連リソースが、少なくとも 2 つの場所の間で分散している、方法が提供される。

【0098】

本開示主題のある特定の実施例によれば、その上さらに、構成が、記憶システムリソース

50

スをオーバーコミットすることを含む、方法が提供される。

【0099】

本開示主題のある特定の実施例によれば、その上さらに、構成の判定の失敗に応答して、ユーザに、失敗通知、または構成の計算の成功を可能にするであろう、少なくとも1つの追加のリソースの追加を示す推奨を提供することをさらに含む、方法が提供される。

【0100】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層を備える、分散記憶システムが提供され、相互接続コンピュータノードのうちの各1つは、統一分散記憶プラットフォーム（UDSP）エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、UDSPエージェントは、少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも1つのサービスレベル仕様（SLS）を受け、少なくとも、少なくとも1つのSLS、および記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算し、計算された構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付け、分散記憶システムの動的拳動に関する動的拳動パラメータデータを受け、少なくとも1つのSLSが違反されたときに、少なくとも、少なくとも1つのSLS、記憶関連リソースパラメータデータ、および動的拳動パラメータデータに基づいて、分散記憶システムの再構成を計算し、かつ計算された再構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される。

10

【0101】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、少なくとも1つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも1つについて割り当てのうちの少なくとも1つのグレードを、各グレードが、SLSの少なくとも1つのユーザ定義された記憶要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、グレードは、該当する場合、それぞれのコンピュータノードに接続される1つ以上の記憶関連リソースに関するパラメータデータにも基づいて計算されることと、計算されたグレードに基づいて、割り当てのうちの1つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることと、更新されたインフラストラクチャ層の少なくとも1つの追加された相互接続コンピュータノードについて以後のタスクの割り当てのグレードを計算し、更新されたインフラストラクチャ層は、少なくとも1つの追加の相互接続コンピュータノードをそれに追加することに応じて作成されることと、計算されたグレードに基づいて、以後のタスクの割り当てのうちの1つ以上を実行するか、または以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノードにルートすることを行うように構成される、分散記憶システムが提供される。

20

【0102】

本開示主題のある特定の実施例によれば、その上さらに、構成が最適化エンジンによって計算される、分散記憶システムが提供される。

30

【0103】

本開示主題のある特定の実施例によれば、その上さらに、最適化が、線形計画法、焼き鈍し法、遺伝的アルゴリズム等の最適化技法のうちの1つ以上を使用する、分散記憶システムが提供される。

【0104】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の最適化技法が、ヒューリスティックスまたは近似値を使用する、分散記憶システムが提供される。

40

【0105】

本開示主題のある特定の実施例によれば、その上さらに、記憶要件が、場所、ローカル

50

保護レベル、バックアップ保持ポリシー、遠隔保護レベル、性能レベル、暗号化レベル、重複排除、圧縮、記憶方法のうちの少なくとも1つを含む、分散記憶システムが提供される。

【0106】

本開示主題のある特定の実施例によれば、その上さらに、1つ以上の記憶関連リソースが、キャッシュリソース、データ記憶リソース、およびネットワークリソースの記憶関連リソースカテゴリの少なくとも1つの記憶関連リソースを含む、分散記憶システムが提供される。

【0107】

本開示主題のある特定の実施例によれば、その上さらに、記憶関連リソースが、少なくとも2つの場所の間で分散している、分散記憶システムが提供される。 10

【0108】

本開示主題のある特定の実施例によれば、その上さらに、構成が、記憶システムリソースをオーバーコミットすることを含む、分散記憶システムが提供される。

【0109】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、構成の判定の失敗に応答して、ユーザに、失敗通知、または構成の計算の成功を可能にするであろう、少なくとも1つの追加の記憶関連リソースの追加を示す推奨を提供するように構成される、分散記憶システムが提供される。

【0110】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層を備える、分散記憶システムが提供され、相互接続コンピュータノードのうちの各1つは、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、UDSPエージェントは、少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも1つのサービスレベル仕様(SLS)を受け、少なくとも、少なくとも1つのSLS、および記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算し、計算された構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される。 30

【0111】

本開示主題の態様によれば、その上さらに、分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードが提供され、インフラストラクチャ層は、相互接続コンピュータノードを含み、コンピュータノードは、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースを備え、UDSPエージェントは、

少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも1つのサービスレベル仕様(SLS)を受け、少なくとも、少なくとも1つのSLS、および記憶関連リソースパラメータデータに基づいて、分散記憶システムの構成を計算し、 40

かつ計算された構成に従って、記憶関連リソースのうちの1つの少なくとも一部を自動的に割り付けるように構成される。

【0112】

本開示主題の態様によれば、その上さらに、分散記憶システムのインフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法が提供され、インフラストラクチャ層は、相互接続コンピュータノードを含み、本方法は、少なくとも1つの論理的記憶エンティティを参照するユーザ定義された記憶要件と、相互接続コンピュータノードに接続された1つ以上の記憶関連リソースに関する記憶関連リソースパラメータデータと、を備える、少なくとも1つのサービスレベル仕様(SLS)を受けることと、 50

少なくとも、少なくとも 1 つの SLS、および記憶関連リソースパラメータデータに基づいて、記憶システムの構成を計算することと、計算された構成に従って、記憶関連リソースのうちの 1 つの少なくとも一部を自動的に割り付けることを含む。

【0113】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層を備える、分散記憶システムが提供され、コンピュータノードのうちの各 1 つは、統一分散記憶プラットフォーム (UDSP) エージェントを実行するように構成される、少なくとも 1 つの処理リソースを備え、相互接続コンピュータノードのうちの少なくとも 1 つは、1 つ以上の他の記憶関連リソースを備え、UDSP エージェントは、少なくとも 1 つの割り当てを含むタスクを受けることと、相互接続コンピュータノードのうちの少なくとも 1 つについて割り当てのうちの少なくとも 1 つのグレードを、各グレードが、少なくとも 1 つのサービスレベル仕様 (SLS) 要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算し、計算されたグレードに基づいて、割り当てのうちの 1 つ以上を実行するか、またはタスクをインフラストラクチャ層のより好適なコンピュータノードにルートすることを行なうように構成される。

【0114】

本開示主題の態様によれば、その上さらに、統一分散記憶プラットフォーム (UDSP) エージェントを実行するように構成される、少なくとも 1 つの処理リソースを備える、相互接続コンピュータノードを含む、インフラストラクチャ層に接続されるように構成されるコンピュータノードが提供され、UDSP エージェントは、少なくとも 1 つの割り当てを含むタスクを受けることと、インフラストラクチャ層に接続され、かつ 1 つ以上の記憶関連リソースを備える、少なくとも 1 つの他のコンピュータノードについて割り当てのうちの少なくとも 1 つのグレードを、各グレードが、少なくとも 1 つのサービスレベル仕様 (SLS) 要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算することと、計算されたグレードに基づいて、割り当てのうちの 1 つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることを行なうように構成される。

【0115】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層に接続されるように構成されるコンピュータノードを動作させる方法が提供され、本方法は、

少なくとも 1 つの割り当てを含むタスクを受けることと、インフラストラクチャ層に接続され、かつ 1 つ以上の記憶関連リソースを備える、少なくとも 1 つの他のコンピュータノードについて割り当てのうちの少なくとも 1 つのグレードを、各グレードが、少なくとも 1 つのサービスレベル仕様 (SLS) 要件を満たしながら、割り当てのうちのそれぞれの割り当てを実行する相互接続コンピュータノードのうちのそれぞれのコンピュータノードの適合性を示すように計算することと、計算されたグレードに基づいて、割り当てのうちの 1 つ以上を実行するか、またはタスクをより好適なコンピュータノードにルートすることとを含む。

【0116】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを備える、分散記憶システムが提供され、コンピュータノードのうちの各 1 つは、統一分散記憶プラットフォーム (UDSP) エージェントを実行するように構成される、少なくとも 1 つの処理リソースを備え、コンピュータノードのうちの少なくとも 1 つは、オブジェクトをキャッシュに格納するように構成され、かつ対応するキャッシュ関連パラメータを有する、少なくとも 1 つのキャッシュリソースを含む、1 つ以上のリソースを備え、少なくとも 1 つのキャッシュリソースを有するそれぞれのコンピュータノードの少なくとも 1 つのUDSP エージェントは、キャッシュ関連パラメータが少なくとも 1 つの第 1 の SLS 基準を満

10

20

30

40

50

たすかどうかを判定するために、それぞれのコンピュータノードに接続された少なくとも1つのキャッシュリソースのキャッシュ関連パラメータを監視し、少なくとも1つの第1のSLS基準が満たされない場合、少なくとも1つの他のコンピュータノードへの少なくとも1つのキャッシュリソースの1つ以上のキャッシュオブジェクト空間の少なくとも一部のハンドオフを開始するように構成され、1つ以上のキャッシュオブジェクト空間の少なくとも一部を受けた後、そのキャッシュ関連パラメータは、少なくとも1つの第2のSLS基準を満たす。

【0117】

本開示主題のある特定の実施例によれば、その上さらに、少なくとも1つのキャッシュリソースを有する、それぞれのコンピュータノードの各UDSPエージェントはさらに、キャッシュ関連パラメータが、対応するキャッシュリソースの利用不足を示す少なくとも1つの第3のSLS基準を満たすことを監視し、少なくとも1つの第3のSLS基準を満たす場合、少なくとも1つの他のコンピュータノードへの少なくとも1つのキャッシュリソースの1つ以上のキャッシュオブジェクト空間の少なくとも一部のハンドオフを開始するように構成され、1つ以上のキャッシュオブジェクト空間の少なくとも一部を受けた後、そのキャッシュ関連パラメータは、少なくとも1つの第2のSLS基準を満たす、分散記憶システムが提供される。

10

【0118】

本開示主題のある特定の実施例によれば、その上さらに、第1のコンピュータノードのUDSPデータリポジトリが、第2のコンピュータノードの1つ以上のキャッシュリソースに対応するキャッシュ関連パラメータについてのデータを含むように、第1のコンピュータノードが、それぞれのUDSPデータリポジトリを有し、第1のコンピュータノードのUDSPエージェントは、それぞれの第2のコンピュータノードのキャッシュ関連パラメータが少なくとも1つの第1のSLS基準を満たすかどうかを判定するために、少なくとも1つの第2のコンピュータノードの1つ以上のキャッシュリソースについて、そのUDSPデータリポジトリから抽出されたデータに従って、キャッシュ関連パラメータを監視し、少なくとも1つの第1のSLS基準が満たされない場合、第1のコンピュータノードのUDSPデータリポジトリから抽出されるような、それぞれの第2のコンピュータノードのキャッシュ関連パラメータのうちの少なくとも1つを含む、それぞれの第2のコンピュータノードへのメッセージを開始し、それにより、第1のコンピュータノードから受け取られたキャッシュ関連パラメータが、第2のコンピュータノードのキャッシュ関連パラメータよりも好ましいかどうかを、それぞれの第2のコンピュータノードが判定することを可能にするように構成される、分散記憶システムが提供される。

20

【0119】

本開示主題のある特定の実施例によれば、その上さらに、第1のコンピュータノードが、それに接続された少なくとも1つのキャッシュリソースを有する、分散記憶システムが提供される。

30

【0120】

本開示主題のある特定の実施例によれば、その上さらに、第1のコンピュータノードのUDSPデータリポジトリが、第1のコンピュータノードのキャッシュリソースに対応するセルフキャッシュ関連パラメータについてのデータも含む、分散記憶システムが提供される。

40

【0121】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、ハンドオフの開始の失敗に応答して、ユーザに、失敗通知、またはハンドオフの開始の成功を可能にする、少なくとも1つの追加のキャッシュリソースの追加を示す推奨を提供するように構成される、分散記憶システムが提供される。

【0122】

本開示主題のある特定の実施例によれば、その上さらに、ゲートウェイリソースまたはクライアントサーバのUDSPデータリポジトリが、コンピュータノードの1つ以上のキ

50

キャッシングリソースに対するキャッシング関連パラメータについてのデータを含むように、ゲートウェイリソースまたはクライアントサーバは、それぞれのUDSPデータリポジトリを有し、ゲートウェイリソースまたはクライアントサーバのUDSPエージェントはさらに、それぞれのコンピュータノードのキャッシング関連パラメータが少なくとも1つの第1のSLS基準を満たすかどうかを判定するために、少なくとも1つのコンピュータノードの1つ以上のキャッシングリソースについて、そのUDSPデータリポジトリから抽出されたデータに従って、キャッシング関連パラメータを監視し、少なくとも1つの第1のSLS基準が満たされない場合、ゲートウェイリソースまたはクライアントサーバのUDSPデータリポジトリから抽出されるような、それぞれのコンピュータノードのキャッシング関連パラメータのうちの少なくとも1つを含む、それぞれのコンピュータノードへのメッセージを開始し、それにより、ゲートウェイリソースまたはクライアントサーバから受け取られたキャッシング関連パラメータが、コンピュータノードのキャッシング関連パラメータよりも好みいかどうかを、それぞれのコンピュータノードが判定することを可能にするように構成される、分散記憶システムが提供される。

【0123】

本開示主題のある特定の実施例によれば、その上さらに、それぞれのコンピュータノードのUDSPエージェントはさらに、1つ以上のオブジェクト空間の少なくとも一部の新しい場所を示す、ハンドオフ後キャッシングマッピングを作成し、ハンドオフ後キャッシングマッピングを、他のコンピュータノードに、および該当する場合、1つ以上のキャッシングオブジェクト空間の少なくとも一部と関連付けられる1つ以上のクライアントサーバに送り、かつ1つ以上のキャッシングオブジェクト空間の少なくとも一部の所有権を他のコンピュータノードに譲渡するように構成される、分散記憶システムが提供される。

【0124】

本開示主題のある特定の実施例によれば、その上さらに、それぞれのコンピュータノードのUDSPエージェントはさらに、ハンドオフ中に、ハンドオフ後キャッシングマッピング内のオブジェクトに関する要求を受け、オブジェクトが、ハンドオフ後キャッシングマッピングに従って、それぞれのコンピュータノードによって所有されず、それぞれのコンピュータノードが、オブジェクトの所有権を持たない場合、要求を他のコンピュータノードに伝えるように構成される、分散記憶システムが提供される。

【0125】

本開示主題の態様によれば、その上さらに、相互接続コンピュータノードを含む、インフラストラクチャ層に接続されるように構成されるコンピュータノードが提供され、コンピュータノードは、オブジェクトをキャッシングに格納するように構成され、かつ対応するキャッシング関連パラメータを有する、少なくとも1つのキャッシングリソースと、統一分散記憶プラットフォーム(UDSP)エージェントを実行するように構成される、少なくとも1つの処理リソースとを備え、UDSPエージェントは、キャッシング関連パラメータが少なくとも1つの第1のSLS基準を満たすかどうかを判定するために、コンピュータノードに接続された少なくとも1つのキャッシングリソースのキャッシング関連パラメータを監視し、少なくとも1つの第1のSLS基準が満たされない場合、相互接続コンピュータノードのうちの少なくとも1つの他のコンピュータノードへの少なくとも1つのキャッシングリソースの1つ以上のキャッシングオブジェクト空間の少なくとも一部のハンドオフを開始するように構成され、1つ以上のキャッシングオブジェクト空間の少なくとも一部を受けた後、そのキャッシング関連パラメータは、少なくとも1つの第2のSLS基準を満たす。

【0126】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、キャッシング関連パラメータが、対応するキャッシングリソースの利用不足を示す、少なくとも1つの第3のSLS基準を満たすことを監視し、少なくとも1つの第3のSLS基準を満たす場合、相互接続コンピュータノードのうちの少なくとも1つの他のコンピュータノードへの少なくとも1つのキャッシングリソースの1つ以上のキャッシングオブジェクト

10

20

30

40

50

空間の少なくとも一部のハンドオフを開始するように構成され、1つ以上のキャッシュオブジェクト空間の少なくとも一部を受けた後、そのキャッシュされたパラメータは、少なくとも1つの第2のSLS基準を満たす、コンピュータノードが提供される。

【0127】

本開示主題のある特定の実施例によれば、その上さらに、コンピュータノードが提供され、コンピュータノードはさらに、相互接続コンピュータノードのうちの少なくとも1つの他のコンピュータノードの1つ以上のキャッシュリソースに対応するキャッシュ関連パラメータについてのデータを含む、UDSPデータリポジトリを備え、UDSPエージェントはさらに、それぞれの他のコンピュータノードのキャッシュ関連パラメータが少なくとも1つの第1のSLS基準を満たすかどうかを判定するために、少なくとも1つの他のコンピュータノードの1つ以上のキャッシュリソースについて、そのUDSPデータリポジトリから抽出されたデータに従って、キャッシュ関連パラメータを監視し、少なくとも1つの第1のSLS基準が満たされない場合、UDSPデータリポジトリから抽出されるような、それぞれの他のコンピュータノードのキャッシュ関連パラメータのうちの少なくとも1つを含む、それぞれの他のコンピュータノードへのメッセージを開始し、それにより、コンピュータノードから受け取られたキャッシュ関連パラメータが、他のコンピュータノードのキャッシュ関連パラメータよりも好ましいかどうかを、それぞれの他のコンピュータノードが判定することを可能にするように構成される。10

【0128】

本開示主題のある特定の実施例によれば、その上さらに、UDSPデータリポジトリは、コンピュータノードのキャッシュリソースに対応するセルフキャッシュ関連パラメータについてのデータも含む、コンピュータノードが提供される。20

【0129】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、ハンドオフの開始の失敗に応答して、ユーザに、失敗通知、またはハンドオフの開始の成功を可能にする、少なくとも1つの追加のキャッシュリソースの追加を示す推奨を提供するように構成される、コンピュータノードが提供される。

【0130】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、1つ以上のキャッシュオブジェクト空間の少なくとも一部の新しい場所を示す、ハンドオフ後キャッシュマッピングを作成し、ハンドオフ後キャッシュマッピングを、他のコンピュータノードに、および該当する場合、1つ以上のキャッシュオブジェクト空間の少なくとも一部と関連付けられる1つ以上のクライアントサーバに送り、かつ1つ以上のキャッシュオブジェクト空間の少なくとも一部の所有権を他のコンピュータノードに譲渡するように構成される、コンピュータノードが提供される。30

【0131】

本開示主題のある特定の実施例によれば、その上さらに、UDSPエージェントはさらに、ハンドオフ中に、ハンドオフ後キャッシュマッピング内のオブジェクトに関する要求を受け、オブジェクトが、ハンドオフ後キャッシュマッピングに従って、それぞれのコンピュータノードによって所有されず、それぞれのコンピュータノードが、オブジェクトの所有権を持たない場合、要求を他のコンピュータノードに伝えるように構成される、コンピュータノードが提供される。40

【0132】

本開示主題の態様によれば、その上さらに、それに接続された少なくとも1つのキャッシュリソースを有し、かつ相互接続コンピュータノードを含むインフラストラクチャ層に接続されるように構成される、コンピュータノードを動作させる方法が提供され、本方法は、キャッシュ関連パラメータが少なくとも1つの第1のSLS基準を満たすかどうかを判定するために、対応するキャッシュ関連パラメータを有し、かつオブジェクトをキャッシュに格納するように構成される、少なくとも1つのキャッシュリソースのキャッシュ関連パラメータを監視することと、少なくとも1つの第1のSLS基準が満たされない場合50

、相互接続コンピュータノードのうちの少なくとも1つの他のコンピュータノードへの少なくとも1つのキャッシュリソースの1つ以上のキャッシュオブジェクト空間の少なくとも一部のハンドオフを開始することとを含み、1つ以上のキャッシュオブジェクト空間の少なくとも一部を受けた後、そのキャッシュ関連パラメータは、少なくとも1つの第2のSLS基準を満たす。

【0133】

本開示主題のある特定の実施例によれば、その上さらに、キャッシュ関連パラメータが、対応するキャッシュリソースの利用不足を示す、少なくとも1つの第3のSLS基準を満たすことを監視することと、少なくとも1つの第3のSLS基準を満たす場合、相互接続コンピュータノードのうちの少なくとも1つの他のコンピュータノードへの少なくとも1つのキャッシュリソースの1つ以上のキャッシュオブジェクト空間の少なくとも一部のハンドオフを開始することとを含む、方法が提供され、1つ以上のキャッシュオブジェクト空間の少なくとも一部を受けた後、そのキャッシュ関連パラメータは、少なくとも1つの第2のSLS基準を満たす。

10

【0134】

本開示主題のある特定の実施例によれば、その上さらに、コンピュータノードがさらに、相互接続コンピュータノードのうちの少なくとも1つの他のコンピュータノードの1つ以上のキャッシュリソースに対応する、キャッシュ関連パラメータについてのデータを含む、UDSPデータリポジトリを備える、方法が提供され、本方法はさらに、それぞの他のコンピュータノードのキャッシュ関連パラメータが少なくとも1つの第1のSLS基準を満たすかどうかを判定するために、少なくとも1つの他のコンピュータノードの1つ以上のキャッシュリソースについて、そのUDSPデータリポジトリから抽出されたデータに従って、キャッシュ関連パラメータを監視することと、少なくとも1つの第1のSLS基準が満たされない場合、UDSPデータリポジトリから抽出されるような、それぞの他のコンピュータノードのキャッシュ関連パラメータのうちの少なくとも1つを含む、それぞの他のコンピュータノードへのメッセージを開始し、それにより、コンピュータノードから受け取られたキャッシュ関連パラメータが、他のコンピュータノードのキャッシュ関連パラメータよりも好ましいかどうかを、それぞの他のコンピュータノードが判定することを可能にすることとを含む。

20

【0135】

30

本開示主題のある特定の実施例によれば、その上さらに、UDSPデータリポジトリが、コンピュータノードのキャッシュリソースに対応するセルフキャッシュ関連パラメータについてのデータも含む、方法が提供される。

【0136】

本開示主題のある特定の実施例によれば、その上さらに、ハンドオフの開始の失敗に応答して、ユーザに、失敗通知、またはハンドオフの開始の成功を可能にする、少なくとも1つの追加のキャッシュリソースの追加を示す推奨を提供することをさらに含む、方法が提供される。

【0137】

40

本開示主題のある特定の実施例によれば、その上さらに、1つ以上のキャッシュオブジェクト空間の少なくとも一部の新しい場所を示す、ハンドオフ後キャッシュマッピングを作成することと、ハンドオフ後キャッシュマッピングを、他のコンピュータノードに、および該当する場合、1つ以上のキャッシュオブジェクト空間の少なくとも一部と関連付けられる1つ以上のクライアントサーバに送ることと、1つ以上のキャッシュオブジェクト空間の少なくとも一部の所有権を他のコンピュータノードに譲渡することとをさらに含む、方法が提供される。

【0138】

本開示主題のある特定の実施例によれば、その上さらに、ハンドオフ中に、ハンドオフ後キャッシュマッピング内のオブジェクトに関する要求を受けることと、オブジェクトが、ハンドオフ後キャッシュマッピングに従って、それぞのコンピュータノードによって

50

所有されず、それぞれのコンピュータノードが、オブジェクトの所有権を持たない場合、要求を他のコンピュータノードに伝えることとをさらに含む、方法が提供される。

【図面の簡単な説明】

【0139】

本開示主題を理解し、それが実践で実行され得る方法を確認するために、以下、添付図面を参照して、非限定的実施例のみとして、本主題が説明される。

【図1】本発明の例示的実施形態による、インフラストラクチャ層を含む、分散記憶システムの最上位アーキテクチャを概略的に図示する。

【図2】本開示主題による、分散記憶システムを構成するための簡略化した例となるシステムを概略的に図示する。10

【図3】本開示主題による、目的ベースの管理システムによって行われる最適化プロセスの簡略化した例となるフロー図を概略的に図示する。

【図4】本開示主題による、目的ベースの管理システムによって行われる構成プロセスの例となる動作アルゴリズムの簡略化したフロー図を概略的に図示する。

【図5】本開示主題のある特定の実施例による、分散記憶システムに接続された例となるコンピュータノードを概略的に図示するブロック図である。

【図6】本開示主題のある特定の実施例による、タスクを作成するために実行される一連の動作を図示する、フローチャートである。

【図7】本開示主題のある特定の実施例による、例となる記憶ブロック書きタスクを作成するために実行される一連の動作を図示する、フローチャートである。20

【図8】本開示主題のある特定の実施例による、UDSPエージェントによって受け取られたタスクを管理するために実行される一連の動作を図示する、フローチャートである。

【図9】本開示主題のある特定の実施例による、保留中のタスク割り当てを実行するノードの適合性にグレードを付けるために実行される一連の動作を図示する、フローチャートである。

【図10】本開示主題のある特定の実施例による、コンピュータノード上で保留中の割り当てを実行するために実行される一連の動作を図示する、フローチャートである。

【図11】本開示主題のある特定の実施例による、分散記憶システム(DSS)の再構成を管理するために実行される一連の動作を図示する、フローチャートである。30

【図12】本開示主題のある特定の実施例による、コンピュータノードおよびそれに接続されたリソースのローカルパラメータを監視するために実行される一連の動作を図示する、フローチャートである。

【図13】本開示主題のある特定の実施例による、コンピュータノードに接続されたリソースを検出および管理するために実行される一連の動作を図示する、フローチャートである。

【図14】本開示主題のある特定の実施例による、新しいコンピュータノードを分散記憶システム(DSS)に接続するために実行される一連の動作を図示する、フローチャートである。

【図15】本開示主題のある特定の実施例による、遠隔コンピュータノードから通知を受け、それに従って統一分散記憶プラットフォーム(UDSP)データリポジトリを更新するために実行される一連の動作を図示する、フローチャートである。40

【図16】本開示主題のある特定の実施例による、キャッシュ管理モジュールを概略的に図示するブロック図である。

【図17】本開示主題のある特定の実施例による、コンピュータノードのローカルキャッシュリソースを管理するために実行される一連の動作を図示する、フローチャートである。

【図18】本開示主題のある特定の実施例による、遠隔コンピュータノードの遠隔キャッシュリソースを管理するために実行される一連の動作を図示する、フローチャートである。50

【図19】本発明の例示的実施形態による、キャッシュリソースを分散させることの種々のシナリオを概略的に図示する。

【図20】本開示主題のある特定の実施例による、ハンドオフイニシエータによってキャッシュハンドオフを行うために実行される一連の動作を図示する、フローチャートである。

【図21】本開示主題のある特定の実施例による、ハンドオフ標的によってキャッシュハンドオフを行うために実行される一連の動作を図示する、フローチャートである。

【図22】本開示主題のある特定の実施例による、ハンドオフ中にハンドオフイニシエータによって受け取られるオブジェクト関連要求を処理するために実行される一連の動作を図示する、フローチャートである。

【図23】本開示主題のある特定の実施例による、ハンドオフ中にハンドオフ標的によって受け取られるオブジェクト関連要求を処理するために実行される一連の動作を図示する、フローチャートである。

【発明を実施するための形態】

【0140】

記載される図面および説明では、同一の参照数字は、異なる実施形態または構成に共通する構成要素を示す。

【0141】

特に記述がない限り、以下の考察から明白であるように、本明細書の全体を通して、「受ける」、「計算する」、「実行する」、「ルートする」、「監視する」、「伝搬する」、「割り付ける」、「提供する」、等の用語を用いる考察は、データを他のデータに操作および/または変換するコンピュータのアクションおよび/またはプロセスを含み、該データは、電子量等の物理量として表され、および/または該データは、物理的な物体を表すことが理解される。「コンピュータ」という用語は、非限定的例として、パーソナルコンピュータ、サーバ、コンピュータシステム、通信デバイス、プロセッサ（例えば、デジタル信号プロセッサ（DSP）、マイクロコントローラ、フィールドプログラマブルゲートアレイ（FPGA）、特定用途向け集積回路（ASIC）等）、任意の他の電子コンピュータデバイス、および/またはそれらの任意の組み合わせを含む、データ処理能力を伴う任意の種類の電子デバイスを対象とすると広範囲に解釈されるべきである。

【0142】

本明細書の教示による動作は、所望の目的のために特別に構築されたコンピュータによって、またはコンピュータ可読記憶媒体に記憶されたコンピュータプログラムによる所望の目的のために特別に構成された汎用コンピュータによって行われてもよい。

【0143】

本明細書で使用されるように、「例えば（for example）」、「等の」、「例えば（for instance）」、およびそれらの变形といった表現は、本開示主題の非限定的実施形態を表す。本明細書での「1つの場合」、「いくつかの場合」、「他の場合」、またはそれらの变形という言及は、実施形態と関連して説明される特定の特徴、構造、または特性が、本開示主題の少なくとも1つの実施形態に含まれることを意味する。したがって、「1つの場合」、「いくつかの場合」、「他の場合」、またはそれらの变形といった表現の出現は、必ずしも同じ実施形態を指すわけではない。

【0144】

明確にするために、別個の実施形態との関連で説明される、本開示主題のある特徴はまた、単一の実施形態と組み合わせて提供されてもよいことが理解される。逆に、簡潔にするために、単一の実施形態との関連で説明される、本開示主題の種々の特徴はまた、別々に、または任意の好適な副次的組み合わせで提供されてもよい。

【0145】

本開示主題の実施形態では、図3、4、6～18および20～23に示されるものよりも少ない、多い、および/または異なる段階が実行されてもよい。本開示主題の実施形態では、図3、4、6～18および20～23で図示される1つ以上の段階が、異なる順序

で実行されてもよく、および／または段階の1つ以上のグループが、同時に実行されてもよい。図1、2、5、および19は、本開示の実施形態によるシステムアーキテクチャの概略図を図示する。図1、2、5、および19の各モジュールは、本明細書で定義および説明されるような機能を果たす、ソフトウェア、ハードウェア、および／またはファームウェアの任意の組み合わせで構成することができる。図1、2、5、および19のモジュールは、1つの場所に集中させられ、または1つよりも多くの場所にわたって分散させられてもよい。本開示主題の他の実施形態では、本システムは、図1、2、5、および19に示されるものよりも少ない、多い、および／または異なるモジュールを備えてもよい。

【0146】

これを念頭に置いて、本発明の例示的実施形態による、インフラストラクチャ層を含む、分散記憶システムの最上位アーキテクチャを概略的に図示する、図1に注目する。本開示主題の実施例によれば、分散記憶システム(DSS)200は、インフラストラクチャ層201、統一分散記憶プラットフォーム(UDSP)層202、およびAPI/フレームワーク層203等の層のうちの1つ以上を備えることができる。

【0147】

本開示主題のいくつかの実施例によれば、インフラストラクチャ層201は、1つ以上の相互接続コンピュータノード205(例えば、とりわけ、1つ以上の処理ユニット等の1つ以上の処理リソース、メモリ等の1つ以上のメモリリソース、および1つ以上のネットワークインターフェースを含む、任意のタイプのコンピュータ)、および場合によっては、2つ以上の相互接続コンピュータノード205を含むことができ、それについてのより詳細な説明は、とりわけ図5を参照して本明細書で提供される。インフラストラクチャ層201はさらに、以下の記憶関連リソースのうちの1つ以上を含むことができる：(a)データ記憶リソース(例えば、データ記憶デバイス204、RAID(独立ディスク冗長アレイ)206、DAS(ダイレクトアタッチストレージ)208、JBOD(単なるディスクの束)210、ネットワーク記憶アプライアンス207(例えば、SAN、NAS等)、SSD213等)、(b)揮発性および／または不揮発性である、メモリリソース(例えば、RAM、DRAM等)等のキャッシュリソース212、および／または場合によっては、加えて、または代替として、キャッシュリソースとして使用することができるデータ記憶リソース(例えば、SSD213)等、(c)ネットワークリソース214、および(d)さらなる機能性をDSS200に提供し、および／またはその性能を強化する追加のリソース(圧縮アクセラレータ、暗号化アクセラレータ209、SANリソースとの通信を可能にするホストバスアダプタ(HBA)等)。

【0148】

場合によっては、リソースは、1つよりも多くの同じタイプのデバイス、および／または1つよりも多くの異なるタイプのデバイスを含むことができる。リソースのうちのいくつかのさらに詳細な説明が、本明細書で以下に続く。

【0149】

本開示主題のいくつかの実施例によれば、コンピュータノード205は、ネットワーク(例えば、汎用ネットワーク)によって相互接続することができる。

【0150】

場合によっては、インフラストラクチャ層201のリソースのうちの1つ以上を、1つ以上のコンピュータノードに205直接接続することができる。場合によっては、インフラストラクチャ層201のリソースのうちの1つ以上は、コンピュータノード205内で構成され、その一部を形成することができる。場合によっては、インフラストラクチャ層201のリソースのうちの1つ以上を、ネットワーク(例えば、汎用ネットワーク)によって、コンピュータノード205のうちの1つ以上に(例えば、iSCSI222等の論理接続によって)接続することができる。

【0151】

随意に、ネットワークは、汎用ネットワークであり得る。随意に、ネットワークは、WANを含むことができる。随意に、WANは、例えば、インターネット等のグローバルW

10

20

30

40

50

ANであり得る。随意に、ネットワークリソースは、IPネットワークインフラストラクチャを使用して相互接続することができる。随意に、ネットワークは、記憶領域ネットワーク(SAN)であり得る。随意に、ネットワークは、記憶の仮想化を含むことができる。随意に、ネットワークは、LANを含むことができる。随意に、ネットワークインフラストラクチャは、イーサネット(登録商標)(Ethernet(登録商標))、インフィニバンド(Infiniband)、FC(Fibre Channel;ファイバチャネル)217、FCoE(ファイバチャネルオーバーイーサネット(登録商標);Fibre Channel over Ethernet(登録商標))等、または2つ以上のネットワークインフラストラクチャの任意の組み合わせを含むことができる。随意に、ネットワークは、汎用ネットワークおよび/または記憶ネットワークを含む、当技術分野で公知の任意の種類のネットワークであり得る。随意に、ネットワークは、本明細書でさらに詳述されるように、ネットワーク内でリソースを割り付けて管理するための目的ベースの管理システムを適用するために好適な任意のネットワークであり得る。随意に、ネットワークは、いずれか2つ以上のネットワークタイプ(とりわけ、本明細書で開示されるネットワークタイプを含む)の組み合わせであり得る。

【0152】

本開示主題のいくつかの実施例によれば、インフラストラクチャ層201の少なくとも1つのリソース(とりわけ、コンピュータノード205、データ記憶リソース、キャッシュリソース、ネットワークリソース、コンピュータノード205に接続される追加のリソース、または任意の他のリソースを含む)は、ネットワークおよび/または1つ以上のコンピュータノード205に接続された、既製商品の意図的に構築されていないリソースであり得る。そのようなリソースは、例えば、製造業者、サイズ、計算能力、容量等のリソース特性に関わらず、本明細書で詳述されるように相互接続できることに留意されたい。したがって、その製造業者に関わらず、コンピュータノード205と通信することができる、任意のリソース(とりわけ、コンピュータノード205を含む)を、本明細書でさらに詳述されるように、インフラストラクチャ層201に接続し、DSS200によって利用することができる。場合によっては、任意の数のリソース(とりわけ、コンピュータノード205を含む)を、ネットワークおよび/または1つ以上のコンピュータノード205に接続し、DSS200によって利用することができ、したがって、DSS200の拡張性を可能にする。場合によっては、任意の数のコンピュータノード205をネットワークに接続することができ、任意の数のリソースを1つ以上のコンピュータノード205に接続し、DSS200によって利用することができ、したがって、DSS200の拡張性を可能にする。新しいリソース(とりわけ、コンピュータノード205を含む)をDSS200に接続するプロセスに関するより詳細な説明が、とりわけ図5を参照して、本明細書でさらに詳述されることに留意されたい。

【0153】

UDSP層202を参照すると、本開示主題のいくつかの実施例によれば、それは、コンピュータノード205のうちの1つ以上にインストールする(または別様に関連付ける、あるいは内側で構成する)ことができる、1つ以上のUDSPエージェント220を含むことができる。場合によっては、UDSPエージェント220は、コンピュータノード205のうちのそれぞれにインストールする(または別様に関連付ける)ことができる。場合によっては、UDSPエージェント220は、加えて、ゲートウェイリソース216(とりわけ、本明細書でさらに詳述されるようなプロトコル変換器の役割を果たすことができる)のうちの1つ以上に、および場合によっては、ゲートウェイリソース216のうちのそれぞれにインストールする(または別様に関連付ける)ことができる。場合によっては、UDSPエージェント220は、加えて、クライアントサーバ218(例えば、クライアントとしてDSS200に接続されるサーバおよび/または他のデバイス)のうちの1つ以上に、および場合によっては、クライアントサーバ218のうちのそれぞれにインストールする(または別様に関連付ける)ことができる。場合によっては、クライアントサーバ218は、随意的であるいずれのゲートウェイリソース216も必要とすること

10

20

30

40

50

なく、DSS200と直接相互作用できることに留意されたい。さらに、場合によっては、そのインストール場所またはその関連性に従って、UDSPエージェント220の差異（例えば、その機能性および／またはその能力の差異等）があり得る（例えば、コンピュータノード205にインストールされた、または関連付けられたUDSPエージェント220、ゲートウェイリソース216にインストールされた、または関連付けられたUDSPエージェント220、クライアントサーバ218にインストールされた、または関連付けられたUDSPエージェント220等の間の差異があり得る）ことに留意されたい。

【0154】

とりわけ図5を参照して、UDSPエージェント220の詳細な説明が本明細書で提供されることに留意されたい。それでもやはり、本開示主題のいくつかの実施例によれば、UDSPエージェント220は、DSS200の種々の動作（とりわけ、インフラストラクチャ層201のリソースを自動的に割り付けて管理すること、データ経路動作を処理すること等）を制御および管理するように構成できることに留意されたい。場合によっては、UDSPエージェント220は、DSS200のインフラストラクチャ層201への新しいコンピュータノード205の接続を管理するように構成することができる。場合によっては、UDSPエージェント220は、それらがインストールされるコンピュータノード205に接続されたリソースを検出するように、かつそのようなリソースを管理するように構成することができる。上記で示されるように、とりわけ図5を参照して、UDSPエージェント220のより詳細な説明が本明細書で提供される。

【0155】

場合によっては、UDSP層202は、DSS200用の管理システムを含む、UDSP225を含むことができる。随意に、管理システム処理を、インフラストラクチャ層201内のコンピュータノード205にインストールされた1つ以上のUDSPエージェント220を通して、またはゲートウェイリソース216に、あるいは（例えば、直接、および／またはゲートウェイリソース216を通して）DSS200へのアクセスを伴うクライアントサーバ218にインストールされた1つ以上のUDSPエージェント220を通して、またはそれらの任意の組み合わせで、実装することができる。

【0156】

管理システムは、ユーザが、サービスレベル仕様（SLS）と関連付けることができる新しい論理記憶エンティティ（論理ユニット、オブジェクト記憶部、ファイルシステムインスタンス等）を作成すること（場合によっては、各論理的記憶エンティティが単一のSLSと関連付けられる）、論理記憶エンティティを更新すること、論理的記憶エンティティのアクセス許可をゲートウェイリソース216に、および／またはクライアントサーバ218に付与すること、スナップショットを作成すること、バックアップを作成すること、遠隔サイトへのフェイルオーバー、一次サイトへのフェイルバック、DSS200の動的挙動を監視すること、SLSコンプライアンスを監視すること、種々の（例えば、事前定義および／またはユーザ定義された等）レポート（例えば、性能レポート、リソース可用性レポート、在庫レポート、コンピュータノード205と他のリソースとの間の関係を示す関係レポート、傾向レポート、主要業績評価指標を含む種々のパラメータの予測レポート等）の生成、DSS200の異なる範囲を参照すること（例えば、DSS200全体の分解能では、あるサイト、あるSLS用等のある使用のタイプ、あるリソース等）、DSS200によって提供される種々の警告（故障したハードウェアの警告等）を管理すること等の、DSS200に関する種々の管理タスク（とりわけ、監視および報告タスク）を行うことを可能にすることができる。上記の管理タスクは、非限定的実施例のみとして提供されることに留意されたい。場合によっては、本明細書でさらに詳述されるように、論理的記憶エンティティは、SLSに従ってDSS200によって自動的に作成できることに留意されたい。論理的記憶エンティティのうちのそれぞれは、1つ以上のデータ記憶リソースと関連付けることができることに留意されたい。

【0157】

本明細書の全体を通して、ユーザが言及されるとき、これは、システム管理者等の人間

10

20

30

40

50

のオペレータ、または任意のタイプの補助エンティティを指すことができることに留意されたい。補助エンティティは、例えば、いずれの人間の介入も必要としない補助エンティティ等を含む、外部管理システム等の外部アプリケーションを指すことができる。

【0158】

場合によっては、管理システムは、ユーザがDSS200に満たすように要求する種々の要件を特定する、サービスレベル仕様(SLS)を定義する、ユーザ定義された記憶要件を、ユーザがDSS200に提供することを可能にすることができる。場合によっては、SLSは、論理的記憶エンティティと関連付けることができる。隨意に、SLSは、例えば、データが記憶および/または処理される1つ以上の地理的な場所の仕様、可用性、保持、復旧パラメータ(例えば、RPO-目標復旧点、RTO-目標復旧時間)を定義するローカル保護レベル、どれだけ長く情報が保持されるべきかを定義するバックアップ保持ポリシー、種々の障害シナリオ下で特定された可用性、保持、および復旧目標を達成するために、1つ以上の遠隔の地理的な場所を定義する、障害復旧(DR)の遠隔保護レベル、ローカルおよび/または遠隔の複製ポリシー、IOPS(毎秒の入力/出力動作)、応答時間、およびスループット等の測定基準を使用して定義される(隨意にコミットされた)性能レベル、暗号化要件、重複排除要件、圧縮要件、記憶方法(物理的容量、シン容量/プロビジョニング)等の情報を含むことができる。

【0159】

場合によっては、管理システムは、種々のサービスレベルグループ(SLG)の管理(作成、更新、および削除を含む)を可能にすることができる。SLGは、複数の論理的記憶エンティティの間で共有することができるテンプレートSLSである。SLGは、部分SLS(増強を必要とする)であり得る、および/またはオーバーライドすることができる設定を含有することができる。したがって、例えば、SLGは、それぞれがSLSパラメータを追加する、および/またはオーバーライドすることができる、種々のSLSによって受け継ぐことができる、種々の復旧パラメータのみを定義することができる。

【0160】

本開示主題のいくつかの実施例によれば、UDSP225は、リソースを割り付け、DSS200の中でリソースを管理するための自動管理システムを含むことができる。隨意に、自動管理システムは、とりわけ図2および図5に関して本明細書でさらに詳述されるように、とりわけ、1つ以上のサービスレベル仕様(SLS)によって定義される、ユーザ定義された要件、コンピュータノード205および/またはそれに接続されたリソースに関する種々のパラメータのデータ、DSS200またはその部品を指す種々のパラメータのデータ(例えば、最大許容サイトレベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等)、ならびにDSS200および環境(例えば、クライアントサーバ218、ゲートウェイリソース216等)の動的挙動を指す種々のパラメータのデータのうちのいずれか1つ、またはそれらの任意の組み合わせに基づいて、ネットワーク内でリソースを割り付けて管理するように構成することができる、目的ベースの管理システム(OBMS)100である。隨意に、OBMS100処理は、インフラストラクチャ層201内のコンピュータノード205のうちの1つ以上にインストールされた1つ以上のUDSPエージェント220を通して、またはゲートウェイリソース216に、あるいは(例えば、直接、および/またはゲートウェイリソース216を通して)DSS200へのアクセスを伴うクライアントサーバ218にインストールされた1つ以上のUDSPエージェント220を通して、実装することができる。

【0161】

本開示主題のいくつかの実施例によれば、API/フレームワーク層203は、DSS200へのソフトウェア拡張(プラグイン)の追加を促進する、プラグイン層を含む。そのようなプラグインは、例えば、プロセスをデータに適用する、新しい機能性および特徴をDSS200に導入する、DSS200を特定のアプリケーションとインターフェース接続し、アプリケーション特有のタスク(例えば、記憶関連タスク等)を実装する、種々のリソース特有のドライバを実装する、(例えば、プラグイン機能性および/または目標

10

20

30

40

50

に関して)新しいSLSパラメータおよび/または1つ/複数のパラメータグループを導入する、管理機能を実装するため等に利用することができる。場合によっては、プラグイン層はまた、種々のハードウェア構成要素(例えば、暗号化カード等)と関連付けられるドライバを含むこともできる。

【0162】

場合によっては、プラグインは、1つ以上のUDSPエージェント220上で展開することができる。場合によっては、プラグインは、例えば、プラグイン仕様に従って(例えば、ソフトウェア暗号化プラグインを任意のUDSPエージェント220にインストールすることができる)、コンピュータノード205に、および/またはゲートウェイリソース216に、および/またはUDSPエージェント220がインストールされるクライアントサーバ218に接続された種々のリソースに従って(例えば、ハードウェアアクセラレータプラグインを、そのようなハードウェアアクセラレータと関連付けられるコンピュータノード205と関連付けられる、各UDSPエージェント220上で自動的に展開することができる)、自動管理システム(例えば、OBMS100)の決定に従って、またはシステム管理者の選択に従って等、1つ以上のUDSPエージェント220上で展開することができる。場合によっては、プラグインは、自動管理システム(例えば、OBMS100)によって、および/またはコンピュータノード205によって、自動的に展開することができる。随意に、例えば、ソフトウェア拡張は、例えば、DSS200上に記憶されたデータの重複排除を可能にするデータ重複排除プラグイン、例えば、DSS200上に記憶されたデータの暗号化/復号を可能にするデータ暗号化プラグイン、例えば、DSS200上に記憶されたデータの圧縮/復元を可能にするデータ圧縮プラグイン等の、データ処理プラグイン226を含むことができる。随意に、例えば、ソフトウェア拡張は、例えば、DSS200上に記憶されたデータのインデックス作成を可能にするコンテンツインデキシングプラグイン、DSS200上に記憶されたデータのスナップショットの管理を可能にするスナップショット管理プラグイン、例えば、DSS200上に記憶されたデータの階層化を可能にする階層化管理プラグイン、例えば、障害復旧に関係するプロセス、ポリシー、およびプロシージャの管理を可能にする障害復旧プラグイン、例えば、DSS200上に記憶されたデータの連続またはリアルタイムバックアップを可能にする連続データ保護プラグイン等の記憶機能プラグイン228を含むことができる。随意に、例えば、ソフトウェア拡張は、例えば、クエリ処理を加速するデータベースプラグイン、例えば、種々のDSS200管理タスク、ならびにユーザ、クライアントサーバ218、およびDSS200に接続された他のエンティティとの他の相互作用の実施を可能にする管理プラグイン233、および他の好適なアプリケーションプラグイン等のアプリケーションプラグイン230を含むことができる。

【0163】

本明細書で示されるように、場合によっては、プラグインは、(例えば、プラグイン機能性および/または目標に関して)新しいSLSパラメータおよび/またはパラメータグループを導入することができる。そのような場合、プラグインの機能に従って、それぞれのSLSパラメータおよび/またはパラメータグループをDSS200に導入することができる。そのような導入されたSLSパラメータは、例えば、ユーザによって、および/または自動管理システム(例えば、OBMS100)によって自動的に等、プラグイン関連要件を設定するために使用することができる。

【0164】

場合によっては、ソフトウェア拡張は、コンピュータノード205のうちの1つの上に記憶するか、または1つよりも多くのコンピュータノード205の上に分散させることができる。場合によっては、ソフトウェア拡張は、1つ以上のコンピュータノード205に接続された1つ以上のデータ記憶リソースの上に記憶することができる。場合によっては、ソフトウェア拡張は、UDSPエージェント220によって共有することができる、仮想ソフトウェア拡張ライブラリの中に記憶することができる。

【0165】

10

20

30

40

50

場合によっては、ソフトウェア拡張は、自動的に、および／または手動で（例えば、システム管理者によって）管理することができる。そのような管理は、時として、管理プラグイン233を利用することによって実施することができる。そのような場合、管理プラグイン233は、ソフトウェア拡張のDSS200への追加／からの除去、種々のソフトウェア拡張の1つ以上のUDSPエージェント220への追加／からの除去等を可能にすることができる。

【0166】

DSS200の最上位アーキテクチャの説明の後に、目的ベースの管理システム（OBMS）100によって行うことができるDSS200構成プロセスの詳細な説明が本明細書で提供される。この目的で、本開示主題による、分散記憶システム200を構成するための簡略化した例となるシステムを概略的に図示する、図2にここで注目する。この目的で、OBMS100は、とりわけ、インフラストラクチャ層201の中でリソースを自動的に割り付けて管理するように構成することができる。OBMS100は、入力モジュール102と、1つ以上のプロセッサ104と、出力モジュール106とを含むことができる。

10

【0167】

場合によっては、入力モジュール102は、入力データを受けるように構成することができる。そのような入力データは、とりわけ、1つ以上のサービスレベル仕様（SLS）によって定義される、ユーザ定義された記憶要件、1つ以上の論理的記憶エンティティの定義、コンピュータノード205および／またはそれに接続されたリソース（記憶関連リソースデータとも呼ばれる、記憶関連リソースを含む）に関する種々のパラメータのデータ、DSS200またはその部品を指す種々のパラメータのデータ（例えば、最大許容サイトレベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等）、ならびにDSS200および環境（例えば、クライアントサーバ218、ゲートウェイリソース216等）の動的挙動（動的挙動パラメータデータ）を指す種々のパラメータのデータ等のうちのいずれか1つ、またはそれらの任意の組み合わせを含むことができる。

20

【0168】

場合によっては、ユーザ定義された要件は、1人以上のユーザがDSS200および／または1つ以上の論理的記憶エンティティに満たすように要求する種々の要件を特定する、1つ以上のサービスレベル仕様（SLS）を定義することができる。

30

【0169】

場合によっては、DSS200および環境の動的挙動に関する種々のパラメータのデータ（動的挙動パラメータデータ）は、DSS200構成要素（コンピュータノード205およびそれに接続されたリソースを含む）のうちの1つ以上の現在の状態を示す、種々のパラメータデータを含むことができる。そのようなデータは、1つ以上のコンピュータノード205、1つ以上のゲートウェイリソース216、1つ以上のクライアントサーバ218等に関するデータを含む、リソースのうちの1つ以上に関する、存在および／または負荷および／または可用性および／または故障および／または能力および／または応答時間および／または接続性および／または費用（例えば、ネットワーククリンク、異なるタイプのデータ記憶リソースの費用）のデータ、および／または任意の他のデータを含むことができる。場合によっては、そのようなデータは、とりわけ、種々の統計データを含むことができる。

40

【0170】

場合によっては、コンピュータノード205および／またはそれに接続されたリソース（記憶関連リソースデータとも呼ばれる、記憶関連リソースを含む）に関する種々のパラメータのデータは、例えば、以下等の、記憶関連リソースを含むハードウェアリソースを含む、DSS200のリソースを示す種々のパラメータのデータを含むことができる。
a.（例えば、そのハードドライブのうちのそれぞれに対する）データ記憶リソースに関するパラメータ：

50

1. ハードドライブカテゴリパラメータ（例えば、ハードドライブサイズ、インターフェース（例えば、S A S、S A T A、F C、U l t r a - S C S I 等）、キャッシングサイズ、特殊な特徴（例えば、オンドライブ暗号化等）等）

2. ハードドライブ性能パラメータ（例えば、応答時間、平均待ち時間、ランダムシーク時間、データ転送速度等）

3. ハードドライブ電力消費

4. ハードドライブ信頼性パラメータ（例えば、平均故障間隔（M T B F）、年間故障率（A F R）等）

b. コンピュータノード205パラメータ：

1. C P U および C P U あたりのコアの数 10

2. 周波数、L2 および L3 キャッシュサイズ等の各 C P U および / またはコアの性能パラメータ

3. アーキテクチャ（例えば、C P U および / またはコアが 64 ビット計算をサポートするか、それがリトルエンディアンまたはビッグエンディアンであるか）

4. ある命令セットのサポート（例えば、A E S - N I、A E S 暗号化を加速するための新しい命令セット）

5. 利用可能なハードドライブスロットの数

6. 利用可能な記憶インターフェース（S A T A、S A S 等）

7. メモリの最大量

8. サポートされたメモリ構成 20

c. キャッシュリソースパラメータ：

1. キャッシュリソースタイプ（例えば、D R A M、S S D）、サイズ、および性能

2. キャッシュされた記憶空間がローカルまたは遠隔にあるか

3. N U M A パラメータ

d. ゲートウェイリソースパラメータ

1. C P U および C P U あたりのコアの数

2. 周波数、L2 および L3 キャッシュサイズ等の各 C P U および / またはコアの性能パラメータ

3. アーキテクチャ（例えば、C P U および / またはコアが 64 ビット計算をサポートするか、それがリトルエンディアンまたはビッグエンディアンであるか） 30

4. ある命令セットのサポート（例えば、A E S - N I、A E S 暗号化を加速するための新しい命令セット）

5. エンクロージャの中で利用可能なハードドライブスロットの数

6. 利用可能な記憶インターフェース（S A T A、S A S 等）

7. メモリの最大量

8. サポートされたメモリ構成

9. ゲートウェイに関するネットワーキングパラメータ（ポートの数、各ポートの速度およびタイプ等）

e. ネットワークリソースパラメータ：

1. 切替およびルーティング能力 40

2. ネットワークタイプ

3. セキュリティパラメータ

【0171】

これらは実施例にすぎず、追加のおよび / または代替の種々のパラメータを使用できることに留意されたい。

【0172】

場合によっては、D S S 2 0 0 および環境の動的挙動に関するデータ（動的挙動パラメータデータ）は、例えば、以下等の、ハードウェアリソースを含む、D S S 2 0 0 のリソースを示す種々のパラメータのデータを含むことができる。

a.（例えば、そのハードドライブのうちのそれぞれに対する）データ記憶リソースに関する 50

するパラメータ：

1. ハードドライブ空き容量
2. ハードドライブの S . M . A . R . T . パラメータ
3. ハードドライブの電力状態（オフにされている、回転段階にある、準備ができている等）

4. ハードドライブ上の最近および現在の負荷

5. 既存の割り付けおよび保留

b. コンピュータノード 205 パラメータ：

1. 各コアに対する最近および現在の負荷統計

2. 既存の割り付けおよび保留

10

3. メモリの現在の量

c. キャッシュリソースパラメータ：

1. 利用可能なサイズ

2. キャッシュの占有レベル

3. 最近および現在のスワッピング / ページフォールト統計

4. 既存の割り付けおよび保留

d. ゲートウェイリソースパラメータ：

1. 最近および現在のネットワーク接続統計

2. 最近および現在のノード負荷統計

3. 最近および現在の待ち時間統計

20

4. (ゲートウェイによって DSS の中ヘルートされるコマンドの) 最近および現在のルーティング費用統計

5. 既存の割り付けおよび保留

e. ネットワークリソースパラメータ：

1. ネットワークセグメントの最近および現在の負荷

2. ネットワークセグメントの最近および現在の信頼性および品質パラメータ

3. 既存の割り付けおよび保留

【 0173 】

これらは実施例にすぎず、追加のおよび / または代替的な種々のパラメータを使用できるように留意されたい。

30

【 0174 】

場合によっては、入力モジュール 102 は、入力データを 1 つ以上のプロセッサ 104 に転送するように構成することができる。示されるように、OBMS100 処理は、（例えば、とりわけ図 5 を参照して、本明細書でさらに詳述されるように、目的ベースの構成モジュール 380 を利用しながら）1 つ以上の UDSPEージェント 220 を通して、例えば、インフラストラクチャ層 201 内のコンピュータノード 205 のうちの 1 つ以上にインストールされた UDSPEージェント 220 を通して、またはゲートウェイリソース 216 にインストールされた UDSPEージェント 220 を通して、または（例えば、直接、またはゲートウェイリソース 216 を通して）DSS200 へのアクセスを伴うクライアントサーバ 218 にインストールされた UDSPEージェント 220 を通して、またはそれらの任意の組み合わせで、実装することができる。そのような場合、1 つ以上のプロセッサ 104 は、そのような UDSPEージェント 220 と関連付けられた 1 つ以上の処理リソース（例えば、処理ユニット）であり得る（例えば、処理がコンピュータノード 205 にインストールされた UDSPEージェント 220 を通して実装される場合には、プロセッサがそのコンピュータノード 205 の処理ユニットであり得る等）。例えば、並列および / または分散処理の場合に、1 つよりも多くの処理リソース（例えば、処理ユニット）を使用できることに留意されたい。

40

【 0175 】

1 つ以上のプロセッサ 104 は、入力モジュール 102 から入力データを受けるように、かつとりわけそれらが影響を及ぼすエンティティ（そのような SLS と関連付けられる

50

論理的記憶エンティティ等)に関して、DSS200の1人以上のユーザによって提供されるユーザ定義された記憶要件(例えば、SLS)の全てを満たす構成要件を決定するために、入力データに基づいて最適化プロセスを行うように構成することができる。最適化プロセスおよび判定された構成要件のより詳細な説明が、とりわけ図3に関して本明細書で提供される。

【0176】

構成要件は、場合によっては、現在のDSS200リソースが決定された構成要件を満たすために十分であるかどうかを判定することができる、出力モジュール106に転送することができる。したがって、出力モジュール106は、構成要件をシステムによって満たすことができる場合、リソースの割り付け、保留、コミット、またはオーバーコミット(例えば、インフラストラクチャ層201内で利用可能な実際のリソースよりも多くのリソースを事実上割り付けること)、あるいは、リソースを追加することおよび/またはプラグインを追加することを含み得る、ユーザによって作用される改良推奨、および/またはシステムが構成要件を満たすことを可能にするための任意の他の推奨を発行することを含む、ソリューション駆動型アクションを行うように構成することができる。そのような改良推奨は、例えば、1つ以上のリソースを追加する、1つ以上のプラグインを追加またはアップグレードする、追加のおよび/または異なる場所(ローカルおよび/または遠隔)にわたってインフラストラクチャを広げる推奨等を発行することを含むことができる。

【0177】

場合によっては、初めてDSS200および/または1つ以上の論理的記憶エンティティを展開するときに、および/またはDSS200および/または1つ以上の論理的記憶エンティティに適用された1つ以上の変更(例えば、事前定義された変更)(例えば、DSS200へのコンピュータノード205、キャッシュリソース、データ記憶リソース、ネットワークリソース、プラグイン、または任意の他のリソース等のリソースの追加/除去、1つ以上のユーザ定義された記憶要件の変更等)後に、および/または(とりわけ図5および図11に関して、以下でさらに詳述されるように)DSS200の動的挙動に従って等、構成プロセスまたはその部分を開始できることに留意されたい。加えて、または代替として、構成プロセスまたはその部分は、半連続的に(例えば、所定の時間間隔等で)開始することができる。加えて、または代替として、構成プロセスまたはその部分は、連続的に行うことができる。

【0178】

さらに、図2を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および/または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

【0179】

ここで、本開示主題による、目的ベースの管理システムによって行われる最適化プロセスの簡略化した例となるフロー図を概略的に図示する、図3に注目されたい。場合によっては、1つ以上のプロセッサ104は、(例えば、入力モジュール102から)入力データを受け、場合によっては、受け取られた入力データを、最適化エンジンによる処理に好適な形式に(例えば、最適化問題表現に)変換するように構成することができる(ブロック112)。

【0180】

1つ以上のプロセッサ104と関連付けられる最適化エンジンは、(とりわけ図2を参照して、本明細書でさらに詳述されるように)入力データによって定義されるような要件を満たす必要構成に到達するよう、元のおよび/または変換された入力データに基づいて最適化プロセスを行うように構成することができる(ブロック114)。いくつかの場合では、最適化プロセスに、それが見出した第1の有効ソリューションを返信するように命

10

20

30

40

50

令できる一方で、他の場合では、最適化プロセスに、一式の計算された有効ソリューションの中から最適ソリューションを検索するように命令できることに留意されたい。随意に、最適化プロセスで使用される最適化技法は、線形計画法、焼き鈍し法、遺伝的アルゴリズム、または当技術分野で公知である任意の他の好適な最適化技法のうちのいずれか1つ、またはそれらの任意の組み合わせを含むことができる。随意に、最適化技法は、ヒューリスティックスまたは近似値を利用することができる。随意に、部分的および／または非最新情報に基づいて、最適化決定を行うことができる。

【0181】

場合によっては、最適化エンジンの出力は、1つ以上のプロセッサ104によって、最適化ソリューション表現から構成要件表現に変換することができる（ブロック116）。

10

【0182】

場合によっては、構成要件は、例えば、下記のうちのいずれか1つ、またはそれらの任意の組み合わせとして、1つ以上のプロセッサ104によって出力される：場所要件（例えば、少なくとも1つの追加のサイトの可用性、1つ／複数の追加のサイト内のある記憶空間の可用性、サイト間の最大待ち時間、例えば、障害復旧目的でのサイト間の最短地理的距離等）、キャッシュリソース要件（例えば、必要キャッシュサイズ、必要キャッシュタイプ、必要キャッシュ場所、必要キャッシュ性能パラメータ等）、ゲートウェイリソース要件（例えば、必要ファイバチャネル帯域幅、必要処理性能パラメータ等）、ネットワークリソース要件（例えば、必要ネットワーク帯域幅、必要ネットワークタイプ等）、コンピューティングリソース要件（例えば、コンピュータノード処理性能パラメータ、CPUコアのコンピュータノードの数等）、データ記憶リソース要件（例えば、必要記憶空間、必要記憶装置タイプ等）、追加のリソース要件（例えば、必要圧縮性能、必要暗号化性能等）、プラグイン要件（例えば、必要データベースプラグイン等）、環境要件（例えば、必要物理的セキュリティレベル等）等（ブロック117）。

20

【0183】

図3を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

30

【0184】

図4を参照すると、本開示主題による、目的ベースの管理システムによって行われる構成プロセスの例となる動作アルゴリズムの簡略化したフロー図の概略図が示されている。場合によっては、上で示されるように、入力モジュール102は、入力データを受け、データを1つ以上のプロセッサ104に転送することができる（ブロック110）。上でさらに示されるように、1つ以上のプロセッサ104は、場合によっては、最適化エンジンによって、入力データを処理に好適な形式に（例えば、最適化問題表現に）変換することができる（ブロック112）。

40

【0185】

1つ以上のプロセッサ104と関連付けられる最適化エンジンは、（とりわけ図2を参照して、本明細書でさらに詳述されるように）入力データによって定義されるような要件を満たす必要構成に到達するよう、元のおよび／または変換された入力データに基づいて最適化プロセスを行うように構成することができる（ブロック114）。場合によっては、最適化エンジンの出力は、1つ以上のプロセッサ104によって、最適化ソリューション表現から構成要件表現に変換することができる（ブロック116）。

【0186】

場合によっては、出力モジュールは、DSS200が必要構成を満たすことができるかどうかを判定するために、必要構成を、DSS200リソース（例えば、コンピュータノード205、記憶関連リソース等）および／または環境の実際のデータと比較することが

50

できる（ブロック118）。場合によっては、実際のDSS200リソースは、現在利用可能であるDSS200リソースの部分を指すことができることに留意されたい。実際のDSS200リソースおよび／または環境が必要構成を満たすことができる場合、必要構成に従って、リソースを保留する、および／または割り付けるように、OBMS100を構成することができる（ブロック126）。場合によっては、OBMS100は、DSS200構成を設定する、および／または任意の誘発された展開アクションを行うように構成することができる（ブロック128）。場合によっては、設定および／または展開アクションは、とりわけ、SLSと関連付けられる、新しい論理的記憶エンティティ（論理ユニット、オブジェクト記憶部、ファイルシステムインスタンス等）を自動的に作成することを含むことができる。場合によっては、各論理的記憶エンティティは、単一のSLSと関連付けられる。10

【0187】

記憶装置構成を設定すること、および／または任意の誘発された展開アクションを行うことの一部として、関連設定および／または展開アクション要求を、UDSPエージェント205に送ることができ、場合によっては、そのような要求は、要求された設定および／または展開アクションに関連する記憶関連リソースと関連付けられる、UDSPエージェント205に送られる。場合によっては、そのような要求を受けるUDSPエージェント205は、とりわけ図5に関して、以下でさらに詳述されるように、DSS200によって使用されるように要求された設定および／または展開に関して、それと関連付けられるデータリポジトリを更新するように構成することができる。場合によっては、展開後に、DSS200を展開するプロセスが無事に終了する（ブロック130）。20

【0188】

実際のDSS200リソースおよび／または環境が必要構成を満たすことができない場合、失敗通知および／または必要インフラストラクチャ構成の実装を可能にするためにユーザによって講じられる是正アクションに関する推奨をユーザに提供する、メッセージをユーザ（例えば、システム管理者）に送るように、OBMS100を構成することができる（ブロック120）。随意に、アクションは、構成の計算の成功を可能にするであろう、インフラストラクチャリソースを追加することを含むことができる。随意に、アクションは、関連プラグインを追加することを含むことができる。随意に、アクションは、追加のおよび／または代替的な場所にわたってインフラストラクチャリソースを広げることを含むことができる。本明細書で開示される推奨は実施例にすぎず、加えて、または代替として、他の推奨をユーザに発行できることに留意されたい。場合によっては、OBMS100は、随意に、ある間隔／遅延後に、必要インフラストラクチャ構成が再評価されるべきか否かに関する決定を行うように構成することができる（ブロック122）。もし再評価されるべきであれば、OBMS100は、ブロック112に戻るように構成することができる。随意に、出力モジュール106は、随意に、連続モードに設定された場合、ある間隔／遅延後に、自動的に112へ進む。随意に、再試行するか否かの決定は、再試行命令のユーザ入力に基づく。もしそうでなければ、DSS200を展開するプロセスは失敗している。場合によっては、OBMS100は、失敗を報告するように構成することができる。30

【0189】

図4を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に關しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。40

【0190】

ここで、本開示主題のある特定の実施例による、分散記憶システムに接続された例となるコンピュータノードを概略的に図示する、ブロック図が示される、図5に注目する。50

【0191】

本開示主題のいくつかの実施例によれば、コンピュータノード205は、1つ以上の処理リソース310を備えることができる。1つ以上の処理リソース310は、処理ユニット、マイクロプロセッサ、マイクロコントローラ、あるいは関連コンピュータノード205リソースおよび／またはコンピュータノード205に接続された記憶関連リソースを制御するため、かつコンピュータノード205リソースおよび／またはコンピュータノード205に接続された記憶関連リソースに関する動作を可能にするために、独立して、または協働してデータを処理するように適合される、複数および／または並列および／または分散処理ユニットを含む、任意の他のコンピュータデバイスまたはモジュールであり得る。

10

【0192】

コンピュータノード205はさらに、コンピュータノード205が、とりわけ、他のコンピュータノードおよび／またはDSS200に接続された他のリソースと通信することを可能にするために、1つ以上のネットワークインターフェース320（例えば、ネットワークインターフェースカード、または任意の他の好適なデバイス）を備えることができる。

【0193】

本開示主題のいくつかの実施例によれば、コンピュータノード205は、とりわけ、SLSを定義する種々のユーザ定義された記憶要件のデータ、および／または各SLSと関連付けられる論理的記憶エンティティのデータ、および／またはコンピュータノード205および／またはそれに接続されたリソースに関する種々のパラメータのデータ、および／またはDSS200またはその部品を指す種々のパラメータに関するデータ、および／またはDSS200および環境（例えば、クライアントサーバ218、ゲートウェイリソース216等）の動的挙動に関するデータ、および／またはDSS200設定および／または展開に関するデータ、および／または任意の他のデータを含む、データを記憶するように構成される、UDSPデータリポジトリ330と関連付けることができる。場合によっては、UDSPデータリポジトリ330はさらに、記憶されたデータの読み出し、更新、および削除を可能にするように構成することができる。場合によっては、UDSPデータリポジトリ330は、コンピュータノード205上、コンピュータノード205に接続された記憶関連リソース（例えば、データ記憶リソース、キャッシュリソース、または任意の他の好適なリソース）上、クライアントサーバ218上、ゲートウェイリソース216上、または任意の他の好適な場所でローカルに位置できることに留意されたい。場合によっては、UDSPデータリポジトリ330は、2つ以上の場所の間で分散させることができる。場合によっては、UDSPデータリポジトリ330は、加えて、または代替として、DSS200内の1つ以上の論理的記憶エンティティの上に記憶することができる。場合によっては、加えて、または代替として、UDSPデータリポジトリ330は、複数のコンピュータノードの間で共有することができる。

20

【0194】

本開示主題のいくつかの実施例によれば、コンピュータノード205はさらに、例えば、1つ以上の処理リソース310によって実行することができる、UDSPエージェント220を備えることができる。上記で示されるように、UDSPエージェント220は、とりわけ、コンピュータノード205および／またはDSS200の種々の動作を制御して管理するように構成することができる。UDSPエージェント220は、タスク管理モジュール335、マルチキャストモジュール340、タスク作成モジュール345、実行モジュール350、ローカルパラメータ監視モジュール360、遠隔ノードパラメータ監視モジュール370、クラウドプラグアンドプレイモジュール380、リソース検出および管理モジュール385、目的ベースの構成モジュール390、キャッシュ管理モジュール397、および目的ベースのルーティングモジュール395等の、モジュールのうちの1つ以上を備えることができる。

30

【0195】

40

50

本開示主題のいくつかの実施例によれば、タスク管理モジュール335は、とりわけ図8に関してさらに詳述されるように、データ経路動作（例えば、読み取り／書き込み動作）等の受け取られたタスクを管理するように構成することができる。

【0196】

マルチキャストモジュール340は、（例えば、ユニキャスト／マルチキャスト／リキヤスト伝送によって）種々の通知を種々のUDSPエージェント220（例えば、他のコンピュータノードにインストールされたUDSPエージェント、ゲートウェイリソース216、クライアントサーバ218等）に伝搬するように構成することができる。そのような通知は、例えば、リソース状態変化の通知、新しいリソースの追加の通知、リソースの切断の通知、ローカルパラメータの変化の通知等を含むことができる。加えて、マルチキャストモジュール340は、種々のUDSPエージェント220とDSS200の他のエンティティならびに外部エンティティ（外部管理システム等）との間で、任意のプロトコルを処理するように構成することができる。10

【0197】

タスク作成モジュール345は、とりわけ図8および9に関してさらに詳述されるように、DSS200の中で実行するための新しいタスクを作成するように構成することができる。

【0198】

実行モジュール350は、とりわけ図10に関してさらに詳述されるように、受け取られたタスクと関連付けられる1つ以上の割り当てをローカルで実行するように構成することができる。20

【0199】

ローカルパラメータ監視モジュール360は、とりわけ図12に関してさらに詳述されるように、コンピュータノード205および／またはそれに接続された任意のリソースの動的挙動を示すパラメータ等の、種々のローカルパラメータを監視し、（例えば、マルチキャストモジュール340を利用しながら）1つ以上のローカルパラメータへの変更を示す通知を伝搬するように構成することができる。場合によっては、ローカルパラメータは、監視が行われる特定のコンピュータノード205（あるいは必要な変更を加えて、ゲートウェイリソース216またはクライアントサーバ218）および／またはそれに接続されたリソースに関するパラメータであることに留意されたい。30

【0200】

遠隔ノードパラメータ監視モジュール370は、とりわけ図15に関してさらに詳述されるように、1つ以上の遠隔コンピュータノード205および／またはそれに接続されたリソースの1つ以上のパラメータの変化を示す通知を受け、それに従ってUDSPデータリポジトリ330を更新するように構成することができる。場合によっては、遠隔ノードパラメータ監視モジュール370は、別のコンピュータノード205に（例えば、他のコンピュータノード205と関連付けられるUDSPエージェント220に）登録して、そこから選択的通知を受けるように構成することができる。場合によっては、遠隔ノードパラメータ監視モジュール370は、独立して、および／または活発に、任意の必要情報について、遠隔コンピュータノード205にクエリを行うように構成できることに留意されたい。40

【0201】

クラウドプラグアンドプレイモジュール380は、とりわけ図14に関してさらに詳述されるように、DSS200へのコンピュータノード205の自律および／または自動接続を可能にするように構成することができる。

【0202】

リソース検出および管理モジュール385は、とりわけ図13に関してさらに詳述されるように、コンピュータノード205に接続されたリソースを検出して管理するように構成することができる。

【0203】

10

20

30

40

50

目的ベースの構成モジュール 390 は、とりわけ図 2～4 および 11 に関してさらに詳述されるように、DSS200 を構成および／または再構成するように構成することができる。

【0204】

目的ベースのルーティングモジュール 395 は、とりわけ図 6 および 8 に関してさらに詳述されるように、受け取られたタスクをコンピュータノード 205 にルートするように構成することができる。

【0205】

キャッシュ管理モジュール 397 は、とりわけ図 16～22 に関してさらに詳述されるように、とりわけ、キャッシュリソースに関するパラメータを監視するように、およびコンピュータノードに接続されたキャッシュリソースを管理するように（とりわけ、キャッシュハンドオフを行うことを含む）構成することができる。10

【0206】

1 つ以上の処理リソース 310 は、UDSP エージェント 220 およびその内で構成されるモジュールのうちのいずれをも実行するように構成できることに留意されたい。

【0207】

本開示主題のいくつかの実施例によれば、UDSP エージェント 220 モジュールのうちのいくつかまたは全てを、組み合わせて単一のモジュールとして提供でき、または一例として、それらのうちの少なくとも 1 つを、2 つ以上のモジュールの形態で実現できることに留意されたい。さらに、場合によっては、UDSP エージェント 220 は、加えて、または代替として、1 つ以上のゲートウェイリソース 216 および／またはクライアントサーバ 218 等にインストールできることに留意されたい。そのような場合、UDSP エージェント 220 の部分的または修正バージョンを、1 つ以上のゲートウェイリソース 216 および／またはクライアントサーバ 218 等にインストールするか、またはそれによって使用することができる。20

【0208】

ここで図 6 を参照すると、本開示主題のある特定の実施例による、タスクを作成するために実行される一連の動作を図示する、フローチャートが示されている。DSS200 によって受け取られる、要求された動作（例えば、読み取り／書き込み動作、管理動作等）を実行するために、タスクを生成することができる。場合によっては、タスクは、要求された動作の一部として実行される、1 つ以上の割り当てのリストを備えることができる。30

【0209】

場合によっては、タスク作成モジュール 345 は、タスク作成プロセス 500 を行うことができる。この目的で、場合によっては、タスク作成モジュール 345 は、例えば、クライアントサーバ 218、ゲートウェイリソース 216、コンピュータノード 205、または任意の他のソースから発信する、要求された動作を受けることができる（ブロック 510）。受け取られた要求された動作は、動作のタイプ（例えば、読み取り、書き込み、管理等）を示すデータ、および／または要求された動作に関する任意の他のデータ（例えば、書き込み要求では、動作が行われる関連論理的記憶エンティティ、書き込まれるブロック等を示すデータ）を含むことができる。40

【0210】

タスク作成モジュール 345 は、タスクコンテナを作成するように構成することができる（ブロック 520）。タスクコンテナは、とりわけ、要求された動作発信元（例えば、そのネットワーク識別子）を示すデータ、動作が行われる関連論理的記憶エンティティを示すデータ、動作特有のデータ（例えば、ブロック書き込み動作の場合、書き込むブロック）および空の割当リストのうちの 1 つ以上を備えることができる。

【0211】

場合によっては、例えば、要求が論理的記憶エンティティと関連付けられるときに、タスク作成モジュール 345 は、論理的記憶エンティティと関連付けられる SLS を取り出し、SLS に従って行われる 1 つ以上の割り当てを作成する（例えば、SLS が、データ

が暗号化されることを要求する場合、暗号化割り当てを自動的に作成することができる等)ように構成することができる(ブロック530)。

【0212】

タスク作成プロセス500は、コンピュータノード205と関連付けられるUDSPエージェント220のタスク作成モジュール345によって行うことができるに留意されたい。しかしながら、加えて、および/または代替として、タスク作成プロセス500は、クライアントサーバ218および/またはゲートウェイリソース216、あるいはタスク作成モジュール345を有する任意の他のソースと関連付けられる、UDSPエージェント220のタスク作成モジュール345によって行うことができる。したがって、場合によっては、コンピュータノード205は、例えば、クライアントサーバ218および/またはゲートウェイリソース216等によってすでに作成されている1つ以上のタスクを受けることができる。

【0213】

図6を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および/または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

【0214】

タスク作成のプロセスをより良く理解するために、本開示主題のある特定の実施例による、例となる記憶ブロック書込タスクを作成するために実行される一連の動作を図示する、フローチャートを示す、図7に注目する。本明細書で提供される実施例では、タスク作成モジュール345は、DSS200に書き込まれるブロックデータ、およびブロックが書き込まれる関連論理的記憶エンティティを示すデータを受けることができる(ブロック605)。

【0215】

場合によっては、タスク作成モジュール345は、新しいタスクコンテナを作成するように構成することができる。タスクコンテナは、とりわけ、動作が起源とした発信元(例えば、そのネットワーク識別子)を示すデータ、ブロックが書き込まれる関連論理的記憶エンティティを示すデータ、論理的記憶エンティティに書き込まれる記憶ブロックデータ、および空の割当リストを備えることができる(ブロック610)。

【0216】

場合によっては、各タスクには、生成番号を割り当てることができる。そのような生成番号は、対立を解決し、故障シナリオに対処するために、種々のプラグインおよびリソースによって使用することができる、一意的な順次(または任意の他の順序の値)識別子であり得る。例えば、第1のタスク(FT)が第2の対立するタスク(ST)の前に発行され、STが最初に処理するために受け取られると仮定することができる。そのような場合、実行モジュール350は、FTの生成番号がSTの生成番号よりも早いかどうかをチェックするように構成することができ、そのような場合、実行モジュール350は、STに従って以前に更新されたデータを上書きしないように構成することができる。

【0217】

タスク作成モジュール345はまた、動作が行われる論理的記憶エンティティと関連付けられるSLSを取り出し(ブロック615)、それに従って、タスクと関連付けられる割り当てリストに関連割り当てを導入するように構成することもできる。したがって、タスク作成モジュール345は、SLSに従って圧縮が必要とされるかどうかをチェックするように構成することができ(ブロック620)、もしそうであれば、タスク作成モジュール345は、関連割り当てを割り当てリストに追加する(例えば、データを圧縮する)ように構成することができる(ブロック625)。タスク作成モジュール345はさらに、SLSに従って暗号化が必要とされるかどうかをチェックするように構成することができ

10

20

30

40

50

き（ブロック 630）、もしそうであれば、タスク作成モジュール 345 は、関連割り当てを割り当てリストに追加する（例えば、データを暗号化する）ように構成することができる（ブロック 635）。

【0218】

これらが、SLS に従って行われる 2 つだけの割り当てであると仮定すると、タスク作成モジュール 345 は、新しいタスクの作成に成功し、新しいタスクは、実行の準備ができる（ブロック 640）。

【0219】

図 7 を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および / または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。10

【0220】

タスクおよびそれらの作成に関する簡潔な説明の後に、本開示主題のある特定の実施例による、UDSP エージェントによって受け取られたタスクを管理するために実行される一連の動作を図示する、フローチャートを示す、図 8 に注目する。

【0221】

場合によっては、UDSP エージェント 220 のタスク管理モジュール 335 は、タスクを受けるように構成することができる（ブロック 405）。タスクを、（例えば、直接、またはとりわけプロトコル変換器の役割を果たすことができるゲートウェイリソース 216 を通して）クライアントサーバ 218 から、ゲートウェイリソース 216 から、別のコンピュータノード 205 から、外部エンティティ（例えば、アプリケーション等）から、または任意の他のソースから受けることができることに留意されたい。20

【0222】

タスクの受領後、タスク管理モジュール 335 は、DSS200 リソース（例えば、コンピュータノードおよび / または記憶関連リソース等）の全てまたは一部の動的挙動を示す、データの全てまたは一部を取り出すように構成することができる（ブロック 410）。

【0223】

場合によっては、タスク管理モジュール 335 は、タスクが SLS と関連付けられる（例えば、タスクが特定の論理的記憶エンティティに関する等）かどうかをチェックし（ブロック 412）、もしそうであれば、（例えば、UDSP データリポジトリ 330 から、または UDSP データリポジトリ 330 の中で入手可能でなければ、別のコンピュータノードの UDSP データリポジトリ等から）タスクと関連付けられる論理的記憶エンティティと関連付けられる、SLS を取り出すように構成することができる（ブロック 413）。

【0224】

タスク管理モジュール 335 は、DSS200 コンピュータノード 205 のうちの 1 つ以上の適合性にグレードを付けて、1 つ以上の保留中のタスク割り当てを実行するために、目的ベースのルーティングモジュール 395 を利用するように構成することができる（ブロック 415）。

【0225】

保留中のタスク割り当ては、その実行前に満たされていない必須条件を持たない割り当てである。例えば、圧縮割り当ては、重複排除割り当ての以前の実行に依存し得る、暗号化割り当ては、圧縮割り当ての以前の実行に依存し得る、等である。

【0226】

保留中のタスク割り当てを実行するコンピュータノード 205 の適合性、したがって、それらのグレードは、例えば、それらの記憶関連リソースを含む、それらのリソース（例

10

20

30

40

50

えば、それらの処理能力)に、および/またはタスクが論理的記憶エンティティに関する場合、1つ以上のSLS要件(例えば、そのような論理的記憶エンティティの範囲でタスク割り当てのうちの1つ以上を実行するために使用されることが可能なりソースを有する)が存在するならば、そのような要件を満たすそれらの能力に、および/またはそれらの動的挙動および現在の状態等に依存し得る。グレード付けプロセスのより詳細な説明が、図9に関して提供される。

【0227】

計算されたグレードに基づいて、タスク管理モジュール335は、グレード付け結果によって、タスクを、例えば、より好適なコンピュータノード205に、および時には最も好適なコンピュータノードにルートするために、目的ベースのルーティングモジュール395を利用するように構成することができる(例えば、タスクを、最高グレードを有するコンピュータノード205にルートすることができる)(ブロック420)。

10

【0228】

タスク管理モジュール335は、タスクが別のコンピュータノードにルートされたかどうかをチェックするために構成することができる(ブロック425)。タスクが別のコンピュータノードにルートされた場合には、ローカルコンピュータノード205(例えば、プロセスを実行するコンピュータノード205)に関するプロセスが終了する(ブロック440)。しかしながら、ローカルコンピュータノード205が最も好適なものである場合には、例えば、UDSPEージェントの220実行モジュール350を利用することによって、保留中のタスク割り当てのうちの1つ以上をローカルコンピュータノード205上で実行することができる(ブロック430)。

20

【0229】

場合によっては、ローカルコンピュータノード205が実行することができる、保留中のタスク割り当ての全てが、それによって実行されるわけではなく、それが最も好適なものとして選択された、保留中のタスク割り当てのみが実行されることに留意されたい。したがって、例えば、タスクが3つの保留中のタスク割り当てを備え、そのうちの2つをローカルコンピュータノード205によって実行することができ、そのうちの1つに対してローカルコンピュータノードが最高グレードを有し、そのうちのもう1つに対して最高グレードを持たない場合、ローカルコンピュータノード205と関連付けられるUDSPEージェント220は、ローカルコンピュータノード205が最高グレードを有する割り当てのみを実行するように構成することができる。さらに、ローカルコンピュータノード205のUDSPEージェント220は、場合によっては、1つ以上の割り当ての並列および/または同時処理のためにローカルコンピュータノード205の1つよりも多くの処理リソースを(そのようなもののが存在すれば)利用できることに留意されたい。場合によっては、1つよりも多くの割り当てのそのような並列および/または同時処理のために、ローカルコンピュータノード205は、遠隔処理リソース(例えば、1つ以上の遠隔コンピュータノード205と関連付けられる処理リソース)を利用して実行することができる。1つ/複数の割り当ての実行のより詳細な説明が、とりわけ図10に関して提供される。

30

【0230】

タスク管理モジュール335はさらに、ローカルコンピュータノード205上の割り当ての実行後に追加の割り当てが存在するかどうか、および/またはローカルコンピュータノード205上の割り当ての実行が1つ以上の新しいタスク(例えば、複製割り当てが、それぞれ異なる場所に向かった、複数の書き込みタスクの生成をもたらすことができる)および/または割り当ての作成をトリガしたかどうかをチェックするように構成することができる(ブロック435)。もしそうでなければ、プロセスが終了する(ブロック440)。もしそうであれば、プロセスはブロック405に戻り、そこで残りの割り当ておよび/または1つ以上の新しいタスクが、ローカルコンピュータノード205と関連付けられるUDSPEージェント220によって受け取られ、タスクのうちのそれぞれを管理するプロセスが始まる。

40

【0231】

50

場合によっては、例えば、1つ以上の相互接続コンピュータノード205をインフラストラクチャ層に追加することによって、インフラストラクチャ層から1つ以上のコンピュータノード205を除去することによって、インフラストラクチャ層の1つ以上の既存のコンピュータノード205を修正すること（例えば、処理リソース310および／または他の記憶関連リソースをそれに追加すること、処理リソース310および／または他の記憶関連リソースをそこから除去すること等）等によって、インフラストラクチャ層を更新することができる。場合によっては、インフラストラクチャ層へのそのような変更は、DSS200の動作中を含む、動的に（例えば、ユーザが所望するときはいつでも）行うことができる

【0232】

10

タスク管理モジュール335は、場合によっては、追加または修正されている更新されたインフラストラクチャ層コンピュータノード205のうちの1つ以上の適合性にグレードを付けて、以後のタスクの1つ以上の保留中のタスク割り当てを実行するために、目的ベースのルーティングモジュール395を利用するように構成することができる。場合によっては、更新されたインフラストラクチャ層は、そのようなグレード付け計算中に作成することができ、計算は、更新されたインフラストラクチャ層の1つ以上のコンピュータノード205に関して行うことができる。場合によっては、計算は、更新されたインフラストラクチャ層の1つ以上の追加の、または修正されたコンピュータノード205に関して行うことができる。

【0233】

20

タスク管理モジュール335は、場合によっては、以後のタスクの該保留中の割り当てのうちの1つ以上を実行するか、または計算されたグレードに基づいて、該以後のタスクを更新されたインフラストラクチャ層のより好適なコンピュータノード205（および場合によっては、最も好適なコンピュータノード205）に送るように構成することができる。

【0234】

図8を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

30

【0235】

本明細書で詳述されるように、タスク管理モジュール335は、保留中のタスク割り当てを実行するDSS200コンピュータノード205のうちの1つ以上の適合性にグレードを付けるために、目的ベースのルーティングモジュール395を利用するように構成することができる。本開示主題のある特定の実施例による、保留中のタスク割り当てを実行するノードの適合性にグレードを付けるために実行される一連の動作を図示する、フローチャートを図示する、図9に注目する。

【0236】

40

グレード付けプロセス700は、目的ベースのルーティングモジュール395が、行われるタスク、DSS200リソース（コンピュータノードおよび／または記憶関連リソース等を含む）の全てまたは一部の動的挙動を示すデータ、またはグレード付けプロセスによって使用することができる任意の他のデータのうちの少なくとも1つを受けることによって始まることができる（ブロック710）。場合によっては、タスクが特定の論理的記憶エンティティと関連付けられたときに、目的ベースのルーティングモジュール395はまた、タスクと関連付けられる論理的記憶エンティティと関連付けられる、SLSも受けることができる。

【0237】

目的ベースのルーティングモジュール395は、保留中のタスク割り当てのうちのそれ

50

それを実行する 1 つ以上のコンピュータノード 205 の適合性にグレードを付けるように構成することができる（ブロック 720）。グレート付けは、とりわけ、受け取られたデータに基づいて行うことができる。

【0238】

グレードを、DSS200 に接続された各コンピュータノード 205 に対して、またはコンピュータノード 205 のうちのいくつかのみ（例えば、ローカルコンピュータノード 205 からの地理的な距離である、ネットワークトポロジーに従って、1 つ以上の保留中のタスク割り当てを実行するのに好適である十分な数のコンピュータノード 205 が見出されるまで、ランダムおよび／または確定的にコンピュータノード 205 を選択する等）に対して計算できることに留意されたい。さらに、保留中のタスク割り当てを実行するコンピュータノード 205 の適合性にグレードを付けるために、種々のグレード付けアルゴリズムを使用できることに留意されたい。なおもさらに、グレード付けプロセスは、ヒューリスティックスおよび／または近似値を含有および／または使用できることに留意されたい。加えて、または代替として、グレード付けは、部分的および／または非最新情報に基づくことができる。

10

【0239】

場合によっては、グレードが計算される各コンピュータノード 205 について、目的ベースのルーティングモジュール 395 は、各保留中のタスク割り当てについて、コンピュータノード 205 が保留中のタスク割り当てを実行できるかどうかをチェックするよう構成することができる。タスクが論理的記憶エンティティと関連付けられる場合、目的ベースのルーティングモジュール 395 はまた、コンピュータノード 205 が、それぞれの SLS によって定義される要件を満たしながら、保留中のタスク割り当てを実行できるかどうかをチェックするように構成することもできる。コンピュータノード 205 が保留中のタスク割り当てを実行することができない（または関連するときに SLS によって定義される要件を満たすことができない）場合、そのノードのグレードは、（関連するときに SLS によって定義される要件を満たしながら）保留中のタスク割り当てを実行することが可能であるコンピュータノード 205 のグレードよりも低くなるであろう。場合によっては、グレードは、それぞれのコンピュータノード 205 に接続された 1 つ以上の記憶関連リソースに関するパラメータデータ（例えば、記憶関連リソースと関連付けられる存在および／または負荷および／または可用性および／または故障および／または能力および／または応答時間および／または接続性および／または費用に関するパラメータのデータ）、および（関連するときに SLS によって定義される要件を満たしながら）保留中のタスク割り当てを実行するそのような記憶関連リソースの能力にも基づいて計算される。

20

【0240】

例となる方式で、理解を容易にするために、（関連するときに SLS によって定義される要件を満たしながら）保留中のタスク割り当てを実行することができないコンピュータノード 205 のグレードがゼロである一方で、（関連するときに SLS によって定義される要件を満たしながら）保留中のタスク割り当てを実行することができるコンピュータノード 205 のグレードは、ゼロよりも大きい。

30

【0241】

場合によっては、計算されたグレードは、非スカラー値によって、例えば、多次元値によって表すことができることに留意されたい。さらに、計算されたグレードが順序付けられたセットに属さなくてもよいことに留意されたい。なおもさらに、好適なノードおよび／または最も好適なノードの決定（例えば、どのグレードが「より高い」かという決定）は、恣意的であり得る（例えば、グレードが順序付けられたセットに属しないとき等）ことに留意されたい

40

【0242】

場合によっては、割り当てを実行するローカルコンピュータノード 205 の適合性が、1 つ以上の遠隔コンピュータノード 205 の適合性と同一であれば、それら全てがタスクをそれに伝達する同一の通信費を有した場合、タスクを任意の遠隔コンピュータノード 2

50

05に伝達することと関連付けられる費用により、ローカルコンピュータノード205のグレードがより高くなるであろう。

【0243】

場合によっては、グレードが計算される各コンピュータノード205については、目的ベースのルーティングモジュール395は、各保留中のタスク割り当てに対して計算されたグレードに基づいて、統合グレードを計算するように構成することができる（ブロック730）。そのような統合グレードは、例えば、コンピュータノード205の割り当てグレードの要約、コンピュータノード205の割り当てグレードの平均、または計算されたコンピュータノード205の割り当てグレードに基づく任意の他の計算の概要であり得る。

10

【0244】

図9を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

【0245】

図10を参照すると、本開示主題のある特定の実施例による、コンピュータノード上で保留中の割り当てを実行するために実行される一連の動作を図示する、フローチャートが示されている。

20

【0246】

本明細書で詳述されるように、タスク管理モジュール335は、保留中のタスク割り当てのうちの1つ以上を実行するための割り当て実行プロセス800を行うために実行モジュール350を利用するように構成することができる。そのような場合、実行モジュール350は、1つ以上の保留中のタスク割り当てを実行するように構成することができる（ブロック810）。

【0247】

本明細書で示されるように、場合によっては、ローカルコンピュータノード205が実行できる、保留中のタスク割り当ての全てが、それによって実行されるわけではなく、それが最も好適なものとして選択された、保留中のタスク割り当てのみが実行されることに留意されたい。加えて、さらに、ローカルコンピュータノード205と関連付けられるUDSページェント220は、場合によっては、1つ以上の割り当ての並列および／または同時処理のために1つよりも多くの処理リソースを（そのようなものが存在すれば）利用できることに留意されたい。場合によっては、1つよりも多くの割り当てのそのような並列および／または同時処理のために、ローカルコンピュータノード205は、遠隔処理リソース（例えば、1つ以上の遠隔コンピュータノード205と関連付けられる処理リソース）を利用することができる。

30

【0248】

1つ以上の保留中のタスク割り当ての実行後に、実行モジュール335は、実行された割り当ての状態を更新して、割り当てが実行されたことを示すように構成することができる（ブロック820）。

40

【0249】

場合によっては、割り当てを部分的に実行することができ、またはそれらの実行は失敗し得る。そのような場合、実行モジュール335は、関連指示を用いて割り当て状態を更新するように構成することができる。場合によっては、状態はまた、実行結果のデータを含有することもできる。

【0250】

場合によっては、実行モジュール335は、現在のDSS200構成（とりわけ、リソース可用性および割り付けを含む）をチェックする必要性があるかどうかをチェックする

50

ように構成することができる（ブロック830）。そのような必要性は、例えば、論理的記憶エンティティと関連付けられる、実行された割り当てのうちの1つ以上の実行が、それぞれのSLS要件を満たさなかった（または、例えば、事前定義された閾値等に従って、満たさなくなりそうになった等）場合、および／または1つ以上の割り当ての実行が失敗した、および／または割り当ての実行が、事前定義または計算された閾値を超えるコンピュータノード205および／またはそれに接続されたリソースに関するパラメータのデータの変更（記憶空間または任意の他のリソースの不足等）をもたらす場合、および／または任意の他の理由により、存在し得る。

【0251】

DSS200の現在の構成をチェックする必要性がある場合、実行モジュール335は、再構成が必要とされるかどうかをチェックするよう、1つ以上のコンピュータノード205と関連付けられるUDSPエージェント220に推奨するように構成することができる（ブロック840）。場合によっては、推奨は、1つ以上の割り当てが実行されるコンピュータノード205と関連付けられるUDSPエージェント220の目的ベースの構成モジュール390によって処理できることに留意されたい。他の場合において、推奨は、再構成プロセスを行う責任を持つことができる1つ以上のコンピュータノード205（例えば、専用コンピュータノード）と関連付けられるUDSPエージェント220に送ることができる。再構成チェックに関するさらなる説明が、とりわけ図11に関して本明細書で提供される。

【0252】

DSS200の現在の構成をチェックする必要性がない場合、または再構成が必要とされるかどうかをチェックする推奨後に、実行モジュール335は、1つ以上の保留中のタスク割り当ての実行後にタスクが終了した（例えば、タスクと関連付けられる割り当ての全てが実行された）かどうかをチェックするように構成することができる（ブロック850）。

【0253】

タスクが終了していない場合、プロセスが終了する（ブロック860）。タスクが終了した場合、実行モジュール335は、タスクが終了したことを示す任意の通知（例えば、タスク発信元への通知等）が必要とされるかどうかをチェックするように構成することができる（ブロック870）。通知が必要とされない場合、プロセスが終了する（ブロック860）。通知が必要とされる場合、実行モジュール335は、必要に応じてタスク実行の通知を発行するように構成することができ（ブロック880）、プロセスが終了する（ブロック860）。

【0254】

本開示主題のいくつかの実施例によれば、各必要通知について、例えば、本明細書で説明されるタスク作成プロセス中に、必要通知を送ることの専用割り当てを作成することができる。そのような場合、随意に、ブロック850～880を無視することができる。

【0255】

図10を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素についても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施することにも留意されたい。

【0256】

ここで、本開示主題のある特定の実施例による、DSSの再構成を管理するために実行される一連の動作を図示する、図11に注目する。

【0257】

本開示主題のいくつかの実施例によれば、場合によっては、DSS200の再構成が必要とされるかどうかをチェックする、再構成プロセス900を行うことができる。場合に

10

20

30

40

50

よっては、そのようなチェックは、周期的に（例えば、事前定義された時間間隔に従って、例えば、1分ごと、5分ごと、1時間ごと、または任意の他の事前定義された時間間隔）、連続的に（例えば、繰り返しループ等で）、トリガイベント（例えば、監視されたパラメータが事前定義または計算された閾値を超える、とりわけ図10に関して詳述されるように、コンピュータノード205と関連付けられるUDSPエージェント220からの推奨の受領等）後等に、行うことができる。

【0258】

本明細書で示されるように、場合によっては、コンピュータノード205と関連付けられる各UDSPエージェント220は、例えば、目的ベースの構成モジュール390を利用しながら、再構成プロセス900を行うように構成することができる。場合によっては、1つ以上のコンピュータノード205（例えば、専用コンピュータノードと関連付けられる）UDSPエージェント220は、例えば、目的ベースの構成モジュール390を利用しながら、再構成プロセス900を行う責任を持つことができる。

10

【0259】

場合によっては、目的ベースの構成モジュール390は、DSS200の中の1つ以上の論理的記憶エンティティと関連付けられるSLS、DSS200ならびにそのリソースおよび環境の動的挙動を示すデータ、DSS200の現在の構成のデータ、DSS200に関する統計データおよび履歴データ等のうちのいずれか1つ、またはそれらの任意の組み合わせを受けるように構成することができる（ブロック910）。場合によっては、加えて、または代替として、データの全てまたは一部を、再構成プロセス900が実行されるコンピュータノード205と関連付けられるUDSPデータリポジトリ330から取り出すことができることに留意されたい。

20

【0260】

場合によっては、目的ベースの構成モジュール390は、SLSのうちのいずれかが違反されている（または、例えば、事前定義された閾値等に従って、違反される寸前である）かどうか、および／またはDSS200の再構成を行うための、なんらかの他の理由（例えば、SLSに関係なく、1つ以上の割り当てを実行する失敗等）があるかどうかをチェックするために、受け取られたデータを利用するように構成することができる（ブロック920）。

【0261】

30

いくつかの場合では、SLSが違反されるたびに（SLSの違反は、時には、例えば、事前定義された閾値等に従って、そのような違反に近いことを含むことができることに留意されたい）、DSS200の再構成を開始できるが、他の場合では、そのようなDSS200の再構成は、ある事前定義された基準を満たすことに応じて開始できることに留意されたい。そのような基準は、例えば、事前定義された時間枠内で、または時間に関係なくのいずれか等で、要求される、事前定義された数の検出されたSLS違反が満たされることであり得る。したがって、例えば、例となる基準は、3つのSLS違反の検出、または1日以内に3つのSLS違反の検出等であり得る。場合によっては、加えて、または代替として、違反の重要性を基準と見なすことができる。この目的で、目的ベースの構成モジュール390は、DSS200に関する統計データおよび履歴データを利用するように構成することができる。

40

【0262】

DSS200を再構成する必要性がある場合、目的ベースの構成モジュール390は、とりわけ図2～4に関して上記で詳述されるように、DSS200構成プロセスを行うための目的ベースの管理システム（OBMS）100を起動するように構成することができる（ブロック930）。本明細書で示されるように、DSS200の再構成の場合、OBMS100は、構成プロセスのための入力の一部としてDSS200の現在の構成を受け、DSS200を再構成するときにそれを考慮できることに留意されたい。場合によっては、そのような再構成中に、OBMS100は、リソースの全てまたは一部を保留する、および／または割り付ける、および／または再び割り付ける、および／または解放するよ

50

うに構成することができる。

【0263】

SLSが違反されていない（または違反される寸前ではない）、かつ再構成を行うための他の理由がない場合、またはDSS200の再構成の開始後に、再構成プロセス900が終了する（ブロック940）。

【0264】

図11を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

10

【0265】

ここで、本開示主題のある特定の実施例による、コンピュータノードおよびそれに接続されたリソースのローカルパラメータを監視するために実行される一連の動作を図示する、図12に注目する。

【0266】

場合によっては、ローカルパラメータ監視モジュール360は、コンピュータノード205および／またはそれに接続された記憶関連リソースの種々のパラメータを監視するよう構成することができる（ブロック1010）。本明細書で示されるように、監視されたパラメータは、存在および／または負荷および／または可用性および／または故障および／または能力および／または応答時間および／または接続性および／または費用（例えば、ネットワークリンク、異なるタイプのデータ記憶リソースの費用）を示す任意のパラメータ、および／またはコンピュータノード205および／またはそれに接続された任意の記憶関連リソースの動的挙動を示す任意の他のパラメータ、および／またはコンピュータノード205および／またはそれに接続された記憶関連リソースのうちの1つ以上に関する任意の他のデータであり得る。場合によっては、ローカルパラメータ監視モジュール360は、必要な変更を加えて、クライアントサーバ218および／またはゲートウェイリソース216の種々のパラメータを監視するように構成することができる。

20

【0267】

そのような監視は、周期的に（例えば、事前定義された時間間隔に従って、例えば、1分ごと、5分ごと、1時間ごと、または任意の他の事前定義された時間間隔）、連続的に（例えば、繰り返しループ等で）、トリガイベント（例えば、コンピュータノード205への新しいリソースの接続等）後等に、行うことができることに留意されたい。

30

【0268】

場合によっては、ローカルパラメータ監視モジュール360は、新しいパラメータ、または監視されたパラメータのうちのいずれかの値の変化が検出されたかどうかをチェックするように構成することができる（ブロック1020）。もしそうでなければ、ローカルパラメータ監視モジュール360は、パラメータを監視し続けるように構成することができる。しかしながら、新しいパラメータ、または監視されたパラメータのうちのいずれかの値の変化が検出された場合、ローカルパラメータ監視モジュール360は、（例えば、マルチキャストモジュール340を利用しながら）1つ以上のローカルパラメータへの変更を示す通知を伝搬するように構成することができる。場合によっては、そのような通知は、（例えば、ユニキャスト／マルチキャスト／リキャスト伝送によって）1つ以上のコンピュータノード205および／またはクライアントサーバ218および／またはゲートウェイリソース216に送ることができる（ブロック1030）。

40

【0269】

場合によっては、ローカルパラメータ監視モジュール360は、種々の所定の期間で、または種々のトリガイベントに応答して、種々の指示（例えば、1つ以上のローカルパラメータの種々のグループの指示等）を備えることができる、種々の種類の通知を送るよう

50

に構成することができることに留意されたい。さらに、いくつかの通知は、例えば、そのような通知を受けるように登録した1つ以上のコンピュータノード205に選択的に送ることができることに留意されたい。

【0270】

場合によっては、ローカルパラメータ監視モジュール360は、パラメータ値、場合によっては、加えて、または代替として、UDSPデータリポジトリ330の中のそれらの派生物（例えば、パラメータに関係する種々の統計データ）を更新するように構成することができる（ブロック1040）。

【0271】

場合によっては、ローカルパラメータ監視モジュール360は、現在のDSS200構成をチェックする必要性があるかどうかをチェックするように構成することができる。そのような必要性は、例えば、監視されたパラメータのうちの1つが、それと関連付けられる事前定義または計算された閾値を超えた場合、および／または任意の他の理由で存在し得る。

10

【0272】

DSS200の現在の構成をチェックする必要性がある場合、ローカルパラメータ監視モジュール360は、再構成が必要とされるかどうかをチェックするよう、1つ以上のコンピュータノード205と関連付けられるUDSPエージェント220に推奨するように構成することができる。場合によっては、ローカルパラメータ監視モジュール360が作動しているローカルコンピュータノード205と関連付けられるUDSPエージェント220の目的ベースの構成モジュール390によって、推奨を処理できることに留意されたい。他の場合では、推奨は、再構成プロセスを行う責任を持つことができる1つ以上のコンピュータノード205（例えば、専用コンピュータノード）と関連付けられるUDSPエージェント220に送ることができる。再構成チェックに関するさらなる説明が、とりわけ図11に関して本明細書で提供される。

20

【0273】

図12を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に關しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施することにも留意されたい。

30

【0274】

ここで、本開示主題のある特定の実施例による、コンピュータノードに接続されたリソースを検出および管理するために実行される一連の動作を図示する、フローチャートを図示する、図13に注目する。

【0275】

場合によっては、リソース検出および管理モジュール385は、検出および管理プロセス1200を行うように構成することができる。場合によっては、リソース検出および管理モジュール385は、1つ以上のコンピュータノード205に接続された記憶関連リソースについてスキャンするように構成することができる（ブロック1210）。場合によっては、リソース検出および管理モジュール385は、連続的および／または周期的等に（例えば、所定の期間ごとに、例えば、1分ごと、5分ごと、1時間ごと等）スキャンを行うように構成することができる。場合によっては、スキャンは、ユーザ（例えば、システム管理者等）によって開始することができる。

40

【0276】

リソース検出および管理モジュール385は、いずれか新しい記憶関連リソースが見出されたかどうかをチェックするように構成することができる（ブロック1220）。新しい記憶関連リソースが見出されない場合、リソース検出および管理モジュール385は、記憶関連リソースについてスキャンし続けるように構成することができる。1つ以上の新

50

しい記憶関連リソースが見出された場合、記憶関連リソース検出および管理モジュール385は、そのような記憶関連リソースを使用する必要性が1つ以上のプラグインにあるかどうかをチェックし、もしそうであれば、プラグインがローカルに（例えば、新しいリソースが取り付けられる／接続されるコンピュータノード205上に）存在するかどうかをチェックするように構成することができる（ブロック1230）。

【0277】

1つ以上のプラグインの必要性があり、それらが全てローカルに存在する場合、リソース検出および管理モジュール385は、プラグインを新しい記憶関連リソースと関連付けるように構成することができ、記憶関連リソースをローカルリソースプールに追加することができる（ブロック1240）。

10

【0278】

ローカルに存在しない1つ以上のプラグインの必要性がある場合、リソース検出および管理モジュール385は、（例えば、マルチキャストモジュール340を利用しながら）1つ以上の欠落したプラグインが、例えば、1つ以上のコンピュータノード205および／またはクライアントサーバ218および／またはゲートウェイリソース216の上に、および／または本明細書で詳述されるように共有仮想ソフトウェア拡張ライブラリの中に（ブロック1250）、および／またはDSS200上の任意の他の場所に、および／または任意の補助エンティティ上に存在するかどうかをチェックするように構成することができる。

【0279】

リソース検出および管理モジュール385が必要なプラグインを見出した場合、リソース検出および管理モジュール385は、プラグインを新しい記憶関連リソースと関連付けるように構成することができ、記憶関連リソースをローカルリソースプールに追加することができる（ブロック1240）。

20

【0280】

場合によっては、リソース検出および管理モジュール385が必要なプラグインを見出さなかった場合、リソース検出および管理モジュール385は、1つ以上のプラグイン要求を発行するように構成することができる。そのようなプラグイン要求は、場合によっては、ユーザに送ることができ（ブロック1270）、したがって、そのようなユーザが関連プラグインをDSS200に追加することを可能にする（例えば、それを購入した、それをインターネットからダウンロードした後等に）。そのような要求を送信した後に、リソース検出および管理モジュール385は、記憶関連リソースについてスキャンし続けるように構成することができる（ブロック1210）。

30

【0281】

場合によっては、必要なプラグインが見出され、取り出され（要求される場合）、インストールされるまで、新しい記憶関連リソースを、記憶関連リソースのスキャンが行われるたびに識別される新しい記憶関連リソースとしてマークすることができ、したがって、必要なプラグインが見出されるまで本明細書で詳述されるプロセスを繰り返すことに留意されたい。

【0282】

場合によっては、加えて、または代替として、リソース検出および管理モジュール385は、記憶関連リソースのスキャン後に記憶関連リソース除去が検出されたかどうかをチェックするように構成することができる（ブロック1280）。そのような場合において、記憶関連リソース除去が検出された場合、リソース検出および管理モジュール385は、ローカルリソースプールから記憶関連リソースを除去し、随意に、もはや必要とされないあらゆるプラグインを消去する（例えば、そのようなプラグインを利用したリソースが除去されるという事実を踏まえて）ように構成することができる（ブロック1290）。

40

【0283】

場合によっては、加えて、または代替として、リソース検出および管理モジュール385は、必要な変更を加えて、1つ以上のクライアントサーバ218および／またはゲート

50

ウェイリソース 216 に接続された / から切断された記憶関連リソースに対する検出および管理プロセス 1200 を行うように構成することができることに留意されたい。さらに、リソース検出および管理モジュール 385 の利用は、DSS200 の動作中を含んで、場合によっては、ユーザによるいずれの管理アクション（とりわけ、いずれの予備管理アクションを含む）も行わずに、記憶関連リソースのコンピュータノード 205 へ、および / またはクライアントサーバ 218 および / またはゲートウェイリソース 216 への途切れない追加および / または除去および / またはアタッチおよび / またはデタッチ（例えば、「プラグアンドプレイ」）を可能にできることに留意されたい。

【0284】

さらに、場合によっては、記憶関連リソースのローカルリソースプールへの追加 / からの除去が、コンピュータノード 205 の監視されたローカルパラメータへの変更（例えば、種々のローカルパラメータの追加および / または除去および / または更新および / または任意の他の変更）をもたらし得ることに留意されたい。本明細書で示されるように、新しいパラメータが検出されたときに、場合によっては、とりわけ図 12 に関して本明細書で詳述されるように、適切な通知をローカルパラメータ監視モジュール 360 によって送ることができる。場合によっては、そのような通知が再構成をトリガできることに留意されたい。

10

【0285】

図 13 を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および / または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

20

【0286】

ここで、本開示主題のある特定の実施例による、新しいコンピュータノードを分散記憶システム（DSS）に接続するために実行される一連の動作を図示する、フローチャート図 14 に注目する。

【0287】

場合によっては、UDSP エージェント 220 を備える、新しいコンピュータノード 205 がネットワークに接続するときに、新しいコンピュータノード 205 のクラウドプラグアンドプレイモジュール 380 は、新しいネットワーク接続および / または既存のネットワーク接続への変更（例えば、コンピュータノード 205 上で、クラウドプラグアンドプレイモジュール 380 が、新しい、または異なるネットワークに接続されること）を検出するように構成することができる（ブロック 1305）。新しいネットワーク接続の検出後に、クラウドプラグアンドプレイモジュール 380 は、例えば、マルチキャストモジュール 340 を利用することによって、発見メッセージを（例えば、ユニキャスト / マルチキャスト / リキャスト伝送によって）送るように構成することができる（ブロック 1310）。そのような発見メッセージは、例えば、少なくとも DSS200 識別子（各 DSS200 が、その識別を可能にする一意的な識別子を有することができる）を含む応答を送ることによって応答するように、任意の受け取りコンピュータノード 205 をトリガすることができる。

30

【0288】

クラウドプラグアンドプレイモジュール 380 は、所定の時間間隔（例えば、受信コンピュータノード 205 が発見メッセージに応答することを可能にすることができる時間間隔）内に受け取られる任意の応答をリッスンし、任意の応答が受け取られたかどうかをチェックするように構成することができる（ブロック 1315）。応答が受け取られず、コンピュータノード 205 が DSS200 に接続しなかった場合、クラウドプラグアンドプレイモジュール 380 は、ブロック 1310 を繰り返し、発見メッセージを再送るように構成することができる。

40

50

【0289】

応答が受け取られた場合、クラウドプラグアンドプレイモジュール380は、（例えば、受け取られたDSS200識別子に従って）応答が単一のDSS200を指すかどうかをチェックするように構成することができる（ブロック1320）。もしそうであれば、クラウドプラグアンドプレイモジュール380は、コンピュータノード205を検出されたDSS200に接続せるように構成することができる（ブロック1325）。DSS200に参加する結果として、コンピュータノード205は、本明細書で詳述されるように、種々の通知を自動的に送信および受け始めることができることに留意されたい。

【0290】

1つよりも多くのDSS200が検出された（例えば、1つよりも多くのDSS200識別子が発見メッセージへの応答として受け取られた）場合、クラウドプラグアンドプレイモジュール380は、デフォルトDSS200が存在するかどうかをチェックするように構成することができる（ブロック1330）。この目的で、場合によっては、ローカルレジストリ（例えば、ローカルネットワーク上でアクセス可能なデータリポジトリ）から、（例えば、事前定義されたDNS記録等の下で）ドメイン名システムから等、デフォルトDSS200の指示を取り出すことができる。場合によっては、応答がデフォルトDSS200の指示を含むことができる、応答コンピュータノード205のうちの1つによって、デフォルトDNS200の指示を送ることができる。デフォルトDSS200を識別するための他の方法および技法も使用できることに留意されたい。

【0291】

そのようなデフォルトDSS200が存在する場合、クラウドプラグアンドプレイモジュール380は、コンピュータノード205をデフォルトDSS200に接続せるように構成することができる（ブロック1325）。デフォルトDSS200が検出されない場合、新しいコンピュータノード205の指示を、新しいコンピュータノード205が接続するDSS200の選択のためにユーザに提供することができ、クラウドプラグアンドプレイモジュール380は、そのような選択を待つように構成することができる（ブロック1335）。いったん選択が行われると、クラウドプラグアンドプレイモジュール380は、コンピュータノード205を選択されたDSS200に接続せるように構成することができる（ブロック1325）。

【0292】

場合によっては、新しいネットワーク接続の検出時に（ブロック1305）、クラウドプラグアンドプレイモジュール380は、加えて、または代替として、例えば、事前定義されたネットワークアドレスおよび/またはディレクトリサービス（例えば、DNS、アクティブディレクトリ等）で、ローカルレジストリ（例えば、ローカルネットワーク上でアクセス可能なデータリポジトリ）および/またはグローバルレジストリ（例えば、インターネット上でアクセス可能なデータリポジトリ）レジストリサービスを調べるように構成することができる（ブロック1340）。そのようなレジストリサービスは、とりわけ、利用可能なDSS200および/またはデフォルトDSS200の特定を可能にすることができる。

【0293】

クラウドプラグアンドプレイモジュール380は、ローカルレジストリが見つかったかどうかをチェックするように構成することができ（ブロック1345）、もしそうであれば、それをローカルレジストリ上に登録するように構成することができる（まだ登録されていない場合）（ブロック1355）。そのような登録は、レジストリの中にローカルコンピュータノード205に関係する種々の構成パラメータを記憶することを含むことができる。クラウドプラグアンドプレイモジュール380はさらに、ローカルレジストリによって定義されるポリシーがグローバル登録を可能にするかどうかをチェックするように構成することができる（ブロック1355）。もしそうであれば、またはローカルレジストリが見出されない場合、クラウドプラグアンドプレイモジュール380は、グローバルレジストリが見出されたかどうかをチェックするように構成することができる（ブロック150）。

360)。もしそうであれば、クラウドプラグアンドプレイモジュール380は、グローバルレジストリ上に登録するように構成することができる（まだ登録されていない場合）（ブロック1365）。そのような登録は、レジストリの中にローカルコンピュータノード205に関係する種々の構成パラメータを記憶することを含むことができる。

【0294】

グローバルレジストリ上の登録後に、またはローカルレジストリによって定義されるポリシーがグローバル登録を可能にしない場合、クラウドプラグアンドプレイモジュール380は、ブロック1320に飛び、そこから継続するように構成することができる。

【0295】

自動的および手動の両方で、新しいコンピュータノード205をDSS200に参加させるために、他の方法を使用することができ、本明細書で提供される方法は実施例にすぎないことに留意されたい。 10

【0296】

クラウドプラグアンドプレイモジュール380の利用は、UDSPEージェント220がコンピュータノード205にインストールされているならば（UDSPEージェント220の詳細な説明が本明細書で提供される）、DSS200の動作中を含んで、場合によっては、ユーザによるいずれの管理アクション（とりわけ、いずれの予備管理アクションも含む）も行わずに、任意の時に、コンピュータノード205が、シームレスに、ネットワークに追加および/または除去および/またはアタッチおよび/またはデタッチされることを可能にできることに留意されたい。さらに、随意に、ネットワークからの1つ以上のコンピュータノード205の追加および/または除去および/またはアタッチおよび/またはデタッチ後に、DSS200の継続動作を可能にするために、いずれのユーザも必要とされないことに留意されたい。 20

【0297】

図14を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および/または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素についても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。 30

【0298】

ここで、本開示主題のある特定の実施例による、遠隔コンピュータノードから通知を受け、それに従って統一分散記憶プラットフォーム（UDSPE）データリポジトリを更新するために実行される一連の動作を図示する、図15に注目する。

【0299】

場合によっては、コンピュータノード205のUDSPEージェント220の遠隔ノードパラメータ監視モジュール370は、他のコンピュータノード205および/またはクライアントサーバ218および/またはゲートウェイリソース216および/またはユーザ等から発信する、種々の通知（一般的な通知、および/またはメッセージを受けるためにコンピュータノード205が登録したソースから発信する通知）を受けるように構成することができる（ブロック1410）。 40

【0300】

場合によっては、遠隔ノードパラメータ監視モジュール370は、それに従ってUDSPEデータリポジトリ330を更新するように構成することができる（ブロック1420）。

【0301】

本明細書で詳述されるように、コンピュータノード205によって実行されるプロセスに関連する、DSS200の状態（例えば、その動的挙動等）またはその部分の知識をローカルで維持するために、UDSPEデータリポジトリ330に記憶されたそのようなデータを使用できることに留意されたい。 50

【0302】

図15を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施することにも留意されたい。

【0303】

DSS200を説明した後に、DSS200の中のキャッシュリソースを管理するためのシステムおよび方法の説明が続く。本明細書で示されるように、インフラストラクチャ層201は、1つ以上の、場合によっては、2つ以上のコンピュータノード205を備えることができる。インフラストラクチャ層201はさらに、1つ以上のキャッシュリソース212および／またはキャッシュリソースとして使用することができるリソース（例えば、RAM、DRAM、SSD213等）を備えることができる。各キャッシュリソース212および／またはキャッシュリソースとして使用することができるリソースは、（例えば、直接、ネットワークによって等）1つ以上のコンピュータノード205に接続することができる。上記でさらに示されるように、各コンピュータノード205は、その上にインストールされた（または別様にそれと関連付けられた）UDSPエージェント220を有することができる。

【0304】

本明細書で示されるように、UDSPエージェント220は、キャッシュ管理モジュール397を備えることができる。キャッシュ管理モジュール397は、とりわけ標準および／または専用キャッシングアルゴリズム、方法、および技法が動作している、1つ以上のキャッシュリソース上で、種々のキャッシュ関連動作を処理するように構成することができる。キャッシュ管理モジュール397は、DSS200に接続された1つ以上のコンピュータノード205に接続された1つ以上のキャッシュリソース上に記憶されたオブジェクト空間のキャッシュマッピングを管理するように構成することができる。

【0305】

図16は、本開示主題のある特定の実施例による、キャッシュ管理モジュールを概略的に図示するブロック図である。

【0306】

本開示主題のいくつかの実施例によれば、キャッシュ管理モジュール397は、ローカルキャッシュリソース管理モジュール2510、遠隔キャッシュリソース監視モジュール2520、キャッシュハンドオフモジュール2530、およびオブジェクト要求管理モジュール2540のモジュールのうちの1つ以上を備えることができる。

【0307】

ローカルキャッシュリソース管理モジュール2510は、とりわけ図17に関して、本明細書でさらに詳述されるように、コンピュータノード205のローカルキャッシュリソースを管理するように構成することができる。

【0308】

遠隔キャッシュリソース監視モジュール2520は、とりわけ図18に関して、本明細書でさらに詳述されるように、遠隔コンピュータノード205の遠隔キャッシュリソースを監視し、それに従ってハンドオフ推奨および／または命令を発行するように構成することができる。

【0309】

キャッシュハンドオフモジュール2530は、とりわけ図20および21に関して、本明細書でさらに詳述されるように、種々のキャッシュハンドオフ関連プロセスを行うように構成することができる。

【0310】

オブジェクト要求管理モジュール2540は、とりわけ図22および23に関して、本

10

20

30

40

50

明細書でさらに詳述されるように、ハンドオフ中に受け取られたオブジェクト関連要求を管理するように構成することができる。

【0311】

本開示主題のいくつかの実施例によれば、キャッシュ管理モジュール397モジュールのうちのいくつかまたは全てを組み合わせて単一のモジュールとして提供でき、または一例として、それらのうちの少なくとも1つを、2つ以上のモジュールの形態で実現できることに留意されたい。

【0312】

図17に注目する。図17は、本開示主題のある特定の実施例による、コンピュータノードのローカルキャッシュリソースを管理するために実行される一連の動作を図示する、

10

【0313】

本開示主題のいくつかの実施例によれば、ローカルキャッシュリソース管理モジュール2510は、コンピュータノード205に接続されたキャッシュリソースのパラメータを含む、種々のキャッシュ関連パラメータを監視するように構成することができる（ブロック1510）。

【0314】

キャッシュ関連パラメータ205は、コンピュータノード205および／またはそれに接続されたリソース（キャッシュリソースを含む）に対応する、ノードレベルキャッシュ関連パラメータ（例えば、負荷パラメータ、性能パラメータ、存在パラメータ、可用性パラメータ、故障パラメータ、能力パラメータ、応答時間パラメータ、接続性パラメータ、費用パラメータ、場所パラメータ等）を含むことができる。キャッシュ関連パラメータは、加えて、または代替として、キャッシュされたオブジェクト（例えば、キャッシュリソース上にキャッシュされたオブジェクト）に関するオブジェクトレベルキャッシュ関連パラメータ（例えば、キャッシュされたオブジェクトの場所、オブジェクトがキャッシュされる媒体のタイプ等）を含むことができる。キャッシュ関連パラメータはなおも、加えて、または代替として、キャッシュを使用する種々のエンティティ（例えば、クライアントサーバ218等）のパラメータ等の外部キャッシュ関連パラメータを含むことができる。

20

【0315】

ローカルキャッシュリソース管理モジュール2510はなおもさらに、（とりわけ、図18に関して以下でさらに説明されるように）キャッシュハンドオフを行う第三者推奨（例えば、遠隔コンピュータノード205からの推奨）を受けるように構成することができる。そのような第三者推奨は、1つ以上のキャッシュ関連パラメータのデータを含むことができ、そのようなデータは、場合によっては、それに従って第三者がキャッシュハンドオフを行うことを推奨する理由を示す、キャッシュ関連パラメータを含むことができることに留意されたい。第三者推奨を受ける場合、ローカルキャッシュリソース管理モジュール2510は、そのような受け取られたパラメータがローカルで既知のパラメータ（例えば、コンピュータノード205と関連付けられるUDSPデータリポジトリ330上に記憶されたパラメータ）よりも好ましい（例えば、新しいおよび／またはより最新のパラメータが受け取られる）かどうかを判定し、もしそうであれば、以後のブロックで、ローカルで既知のパラメータに加えて、および／またはその代わりに、それらを考慮するように構成することができる。

30

【0316】

そのような監視は、周期的に（例えば、事前定義された時間間隔に従って、例えば、1分ごと、5分ごと、1時間ごと、または任意の他の事前定義された時間間隔）、連続的に（例えば、繰り返しループ等で）、トリガイベント（例えば、キャッシュリソースを含む記憶関連リソースのコンピュータノード205への接続／からの切断等）後等に、行うことができることに留意されたい。

40

【0317】

場合によっては、キャッシュ関連パラメータの監視は、とりわけ図12に関して、本明

50

細書で詳述されるローカルパラメータ監視モジュール 360 によって行うことができるこ
とに留意されたい。

【0318】

監視されたパラメータのうちの 1 つ以上の値の変化が検出される、および / または任意の新しいパラメータが検出される、および / または任意のパラメータがもはや検出されない（例えば、あるキャッシュリソースが除去された等）および / または周期的に（例えば、事前定義または計算された時間間隔に従って、例えば、1 分ごと、5 分ごと、1 時間ごと、または任意の他の事前定義された時間間隔）検出される場合、ローカルキャッシュリソース管理モジュール 2510 は、監視されたパラメータが、1 つ以上の SLS によって定義される 1 つ以上のキャッシュ関連要件の違反を示すかどうかをチェックするように構成することができる。そのようなチェックは、例えば、キャッシュリソース上に現在キャッシュされているオブジェクトと関連付けられる SLS を踏まえて、監視されたパラメータを、および / または例えば、（本明細書で詳述されるマッピングのうちのいずれかを使用して）キャッシュリソースにマップされるオブジェクトが関連付けられる、論理的記憶エンティティと関連付けられる SLS を踏まえて、そのようなオブジェクトを評価することによって行うことができる。

【0319】

ローカルキャッシュリソース管理モジュール 2510 はさらに、監視されたパラメータが、1 つ以上のそのような SLS の違反に近いこと（または違反）を示す、そのようなキャッシュ関連 SLS と関連付けられる、1 つ以上の高ウォーターマーク（例えば、事前定義された最大閾値、計算された最大閾値等）等の 1 つ以上の第 1 の SLS 基準を満たすかどうかを判定するように構成することができる。

【0320】

ローカルキャッシュリソース管理モジュール 2510 はなおもさらに、監視されたパラメータが、DSS200 あるいはその部分を指す 1 つ以上の閾値および / または SLS（例えば、最大許容サイトレベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等）の違反またはそのような違反に近いことを（例えば、1 つ以上の高ウォーターマーク等の 1 つ以上の第 1 の SLS 基準に従って）示すかどうかを判定するように構成することができる（ブロック 1520）。

【0321】

これらの目的で、ローカルキャッシュリソース管理モジュール 2510 は、UDSP データリポジトリ 330 から関連 SLS を取り出すように構成することができることに留意されたい。

【0322】

SLS 違反がある、または SLS が違反される寸前である、または DSS200 あるいはその部分を指す 1 つ以上のパラメータの違反がある場合、全てのキャッシュ関連 SLS 、および / または DSS200 あるいはその部分を指す閾値および / または SLS（例えば、最大許容レベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等）が譲渡後に満たされるように、ローカルキャッシュリソース管理モジュール 2510 は、第 1 のローカルコンピュータノード 205（ハンドオフイニシエータ）が 1 つ以上のキャッシュオブジェクト空間の所有権（例えば、処理する責任）を譲渡することができる 1 つ以上のハンドオフ標的（例えば、1 つ以上のキャッシュリソース 212、および / またはそれに接続されたキャッシュリソースとして使用することができる、1 つ以上のリソースを有する、別の遠隔コンピュータノード 205）を検索するように構成することができる（ブロック 1530）。キャッシュ関連 SLS は、1 つ以上のキャッシュされたオブジェクト、および / またはそのようなオブジェクトおよび / またはそれと関連付けられるキャッシュリソースを含有する 1 つ以上のコンピュータノード 205 、および / またはそのようなコンピュータノード 205 と関連付けられる任意のエンティティを参照する、および / またはそれらに影響を及ぼす、要件を含有する任意の SLS である。

10

20

30

40

50

【 0 3 2 3 】

ハンドオフィニシエータは、1つ以上の論理的記憶エンティティに関する1つ以上のオブジェクト空間を処理する責任を持つことができることに留意されたい。1つ以上のキャッシュオブジェクト空間またはその部分のハンドオフ（所有権の譲渡）は、ハンドオフ標的に、譲渡された1つ以上のキャッシュオブジェクト空間またはその部分の所有権を受容させる。

【 0 3 2 4 】

場合によっては、キャッシュ関連SLSのうちの1つ以上と関連付けられるハンドオフ標の中間ウォーターマーク（例えば、事前定義された中間閾値、計算された中間閾値等）等の1つ以上の第2のSLS基準が満たされるように、場合によっては、ローカルキャッシュリソース管理モジュール2510はさらに、ローカルコンピュータノード205（ハンドオフィニシエータ）が1つ以上のキャッシュオブジェクト空間の所有権を譲渡することが可能である、ハンドオフ標的を検索するように構成することができる。

10

【 0 3 2 5 】

ローカルキャッシュリソース管理モジュール2510はさらに、1つ以上のハンドオフ標的が見出されたかどうかをチェックするように構成することができる（ロック1540）。もしそうでなければ、場合によっては、ローカルキャッシュリソース管理モジュール2510は、不十分なキャッシュリソース（不十分なリソースエラー）があることをユーザに報告し、随意に、キャッシュリソースが1つ以上のキャッシュ関連SLSの要件（例えば、キャッシュリソースを追加すること等）を満たすことを可能にするために行われるアクションをユーザに推奨するように構成することができる。

20

【 0 3 2 6 】

1つ以上のハンドオフ標的が見出される場合、とりわけ図20を参照してさらに詳述されるように、ローカルキャッシュリソース管理モジュール2510は、検出されたハンドオフ標的のうちの1つ以上を用いてハンドオフプロセスを開始するように構成することができる（ロック1560）。場合によっては、ハンドオフプロセスの開始後に、ローカルキャッシュリソース管理モジュール2510は、ロック1510に戻り、キャッシュパラメータを監視し続けるように構成することができる。

【 0 3 2 7 】

場合によっては、非限定的実施例として、1つ以上のハンドオフ標的の選択は、ロック1530で見出された、考えられるハンドオフ標的のうちの1つ以上をランダムに選択すること（または任意の他の規則等に従って選択すること）によって行うことができることに留意されたい。別の非限定的な実施例として、1つ以上のハンドオフ標的の選択は、ロック1530で見出された、可能性のあるハンドオフ標的のうちの1つ以上の適合性をランク付けするように任意のタイプのランキングアルゴリズムを行い、最も好適なものを選択することによって行うことができる。

30

【 0 3 2 8 】

ローカルコンピュータノード205に接続されたキャッシュリソースに、および／またはキャッシュハンドオフを行う第三者推奨の受領に関する、キャッシュ関連パラメータを監視した後、SLS違反がない、または1つ以上の第1のSLS基準（例えば、高ウォーターマーク）の横断がない場合、ローカルキャッシュリソース管理モジュール2510は、監視されたキャッシュ関連パラメータのうちのいずれかが、1つ以上のキャッシュ関連SLSおよび／またはDSS200あるいはその部品を指す閾値および／またはSLS（例えば、最大許容レベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等）と関連付けられる、低ウォーターマーク（例えば、事前定義された最小閾値、計算された最小閾値等）等の1つ以上の第3のSLS基準を満たさず、よって、それが十分活用されていないと見なされるかどうかをチェックするように構成することができる（ロック1570）。ローカルコンピュータノード205のキャッシュリソースが十分活用されない場合、ローカルキャッシュリソース管理モジュール2510は、ロック1510に戻り、キャッシュパラメータを監視し続けるように構成することができる。

40

50

きる。

【0329】

ローカルコンピュータノード205のキャッシュリソースのうちの1つ以上、またはそれらの部分が十分活用されない場合、全てのキャッシュ関連SLS、および／またはDSS200あるいはその部品を指す閾値および／またはSLS（例えば、最大許容レベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等）が譲渡後に満たされるように、ローカルキャッシュリソース管理モジュール2510は、ローカルコンピュータノード205が1つ以上のキャッシュオブジェクト空間の所有権を譲渡することが可能である、1つ以上のハンドオフ標的を検索するように構成することができる（ブロック1580）。場合によっては、キャッシュ関連SLSのうちの1つ以上と関連付けられる中間ウォーターマーク（例えば、事前定義された中間閾値、計算された中間閾値等）等のハンドオフ標的の第2のSLS基準が満たされるように、ローカルキャッシュリソース管理モジュール2510はさらに、ローカルコンピュータノード205が1つ以上のキャッシュオブジェクト空間の所有権を譲渡することができる、ハンドオフ標的を検索するように構成することができる。

【0330】

場合によっては、非限定的実施例として、1つ以上のハンドオフ標的の選択は、ブロック1530で見出された、考えられるハンドオフ標的のうちの1つ以上をランダムに選択すること（または任意の他の規則等に従って選択すること）によって行うことができることに留意されたい。別の非限定的な実施例として、1つ以上のハンドオフ標的の選択は、ブロック1530で見出された、可能性のあるハンドオフ標的のうちの1つ以上の適合性をランク付けするように任意のタイプのランキングアルゴリズム（キャッシュされたオブジェクトマッピングの連結を推進するアルゴリズム等）を行い、最も好適なものを選択することによって行うことができる。

【0331】

キャッシュされたオブジェクトのそのような譲渡は、例えば、可能であれば、他の目的で使用することができる、DSS200または任意の他のエンティティのうちのいずれか一方のキャッシュリソースを解放させることができることに留意されたい。キャッシュされたオブジェクトのそのような譲渡はまた、完全に解放された（それ以上オブジェクトがその上にキャッシュされておらず、いずれのエンティティもそれを使用していない）場合にキャッシュリソースをオフにすることを可能にさせ、したがって、電力消費の削減をもたらすこともできる。

【0332】

ローカルキャッシュリソース管理モジュール2510はさらに、1つ以上のハンドオフ標的が見出されるかどうかをチェックするように構成することができる（ブロック1590）。1つ以上のハンドオフ標的が見出される場合、ローカルキャッシュリソース管理モジュール2510は、とりわけ図20を参照してさらに詳述されるように、検出されたハンドオフ標的のうちの1つ以上を用いてハンドオフプロセスを開始するように構成することができる（ブロック1560）。

【0333】

場合によっては、ハンドオフプロセスの開始後に、ローカルキャッシュリソース管理モジュール2510は、ブロック1510に戻り、キャッシュパラメータを監視し続けるように構成することができる。

【0334】

図17を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に關しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

【0335】

ここで、本開示主題のある特定の実施例による、遠隔コンピュータノードの遠隔キャッシュリソースを監視するために実行される一連の動作を図示する、図18に注目する。

【0336】

本開示主題のいくつかの実施例によれば、遠隔キャッシュリソース監視モジュール2520は、1つ以上の遠隔コンピュータノード205に接続されたキャッシュリソースのパラメータを含む、種々のキャッシュ関連パラメータを監視するように構成することができる（ブロック1510）。キャッシュ関連パラメータ205は、遠隔コンピュータノード205および／またはそれに接続されたリソース（キャッシュリソースを含む）のノードレベルキャッシュ関連パラメータ（例えば、負荷パラメータ、性能パラメータ、存在パラメータ、可用性パラメータ、故障パラメータ、能力パラメータ、応答時間パラメータ、接続性パラメータ、費用パラメータ、場所パラメータ等）を含むことができる。キャッシュ関連パラメータは、加えて、または代替として、キャッシュされたオブジェクトに関するオブジェクトレベルキャッシュ関連パラメータ（例えば、キャッシュされたオブジェクトの場所、オブジェクトがキャッシュされる媒体のタイプ等）を含むことができる。キャッシュ関連パラメータはなおも、加えて、または代替として、キャッシュを使用する種々のエンティティ（例えば、クライアントサーバ218等）のパラメータ等の外部キャッシュ関連パラメータを含むことができる（ブロック1610）。

【0337】

この目的で、遠隔キャッシュリソース監視モジュール2520は、UDSPデータリポジトリ330から関係するキャッシュ関連パラメータを取り出すように構成することができる。

【0338】

遠隔キャッシュリソース監視モジュール2520は、1つ以上のハンドオフ推奨が発効されるべきであるかどうかを判定するために、そのようなパラメータを利用するように構成することができる（ブロック1620）。

【0339】

例えば、監視されたキャッシュ関連パラメータが、1つ以上のキャッシュ関連SLS（1つ以上のキャッシュされたオブジェクト、および／またはそのようなオブジェクトおよび／またはそれと関連付けられるキャッシュリソースを含有する1つ以上のコンピュータノード205、および／またはそのようなコンピュータノード205と関連付けられる任意のエンティティを参照する、および／またはそれらに影響を及ぼす、要件を含有する任意のSLS）、および／またはDSS200あるいはその部品を指す閾値および／またはSLS（例えば、最大許容レベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等）の違反を示す、またはそのような違反に近い（例えば、1つ以上の高ウォーターマーク等の1つ以上の第1のSLS基準に従って）場合に、ハンドオフ推奨が発行されるべきであることに留意されたい。

【0340】

別の実施例として、場合によっては、監視されたキャッシュ関連パラメータのうちのいずれかが、低ウォーターマーク（例えば、事前定義された最小閾値、計算された最小閾値等）等の1つ以上の第3のSLS基準を横断し、よって、それが十分活用されていないと見なされる場合に、ハンドオフ推奨が発行されるべきである。

【0341】

この目的で、遠隔キャッシュリソース管理モジュール2520は、UDSPデータリポジトリ330から関連SLSを取り出すように構成できることに留意されたい。

【0342】

ハンドオフ推奨が発行されるべきである場合に、遠隔キャッシュリソース監視モジュール2520は、本明細書で詳述されるように、監視されたパラメータが、SLS違反、そのような違反に近いこと、または利用不足を示す、1つ以上のコンピュータノード205に、そのような通知を発行するように構成することができる（ブロック1630）。

10

20

30

40

50

【0343】

場合によっては、遠隔キャッシュリソース監視モジュール2520は、加えて、または代替として、1つ以上のハンドオフ命令を提供するように構成することができる。ハンドオフ推奨を、ハンドオフ標的ノイニシエータによって拒否することができる一方で、ハンドオフ命令は、キャッシュハンドオフを行うコマンドであり得る。場合によっては、そのようなハンドオフ命令を、1つ以上のハンドオフイニシエータに発行することができ、それにハンドオフ標的を検索させ、それを用いてハンドオフを開始させる。場合によっては、そのようなハンドオフ命令を、1つ以上のハンドオフイニシエータ、およびハンドオフを行うことができる1つ以上のそれぞれのハンドオフ標的に発行することができ、1つ以上のハンドオフイニシエータに、それぞれの1つ以上のハンドオフ標的を用いてハンドオフを開始させる。10

【0344】

図18を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

【0345】

ここで、本発明の例示的実施形態による、キャッシュリソースを分散させることの種々のシナリオを図示する、図19に注目する。20

【0346】

図19を参照すると、キャッシュを伴うコンピュータノード2010（キャッシュを伴うコンピュータノードが言及されるとき、それは、それに接続されたキャッシュリソースを有するコンピュータノード205を指す）が、場合によっては、1つ以上のキャッシュオブジェクト空間またはその部分の所有権を、キャッシュを伴うコンピュータノード2020に譲渡するハンドオフイニシエータの役割、および例えば、キャッシュを伴うコンピュータノード2030から、1つ以上のキャッシュオブジェクト空間またはその部分の所有権を受容するハンドオフ標的の役割を果たすことができると理解することができる。場合によっては、キャッシュを伴うコンピュータノード2010は、ハンドオフイニシエータの役割およびハンドオフ標的の役割を同時に果たすことに留意されたい。30

【0347】

場合によっては、あるキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2030は、1つよりも多くの他のハンドオフ標的を用いて、例えば、場合によっては同時に、キャッシュを伴うコンピュータノード2010およびキャッシュを伴うコンピュータノード2040を用いて、ハンドオフを行うハンドオフイニシエータの役割を果たすことができる。

【0348】

場合によっては、あるキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2040は、1つよりも多くの他のハンドオフ標的を用いて、例えば、場合によっては同時に、キャッシュを伴うコンピュータノード2030およびキャッシュを伴うコンピュータノード2050を用いて、ハンドオフを行うハンドオフ標的の役割を果たすことができる。40

【0349】

場合によっては、あるキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2090は、ハンドオフイニシエータを用いて、例えば、キャッシュを伴うコンピュータノード2095を用いて、ハンドオフを行うハンドオフ標的の役割を果たし、かつ同じハンドオフイニシエータを用いて、例えば、キャッシュを伴うコンピュータノード2095を用いて、ハンドオフを行うハンドオフイニシエータの役割を果たすことができ、場合によっては、同時にハンドオフ標的の役割を果たしている。したがつ50

て、例えば、場合によっては同時に、キャッシュを伴うコンピュータノード2095が、キャッシュを伴うコンピュータノード2090を用いてハンドオフを開始することができる一方で、キャッシュを伴うコンピュータノード2090は、キャッシュを伴うコンピュータノード2095を用いてハンドオフを開始することができる。

【0350】

場合によっては、ある特定のコンピュータノード（それに接続されたキャッシュリソースを伴う、または伴わない）、例えば、コンピュータノード205は、1つ以上のハンドオフ標的を用いて1つ以上のハンドオフを開始するよう、1つ以上のキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2080およびキャッシュを伴うコンピュータノード2070に推奨するように構成することができる。

10

【0351】

場合によっては、あるクライアントサーバ、例えば、クライアントサーバ218は、1つ以上のハンドオフ標的を用いて1つ以上のハンドオフを開始するよう、1つ以上のキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2070に推奨するように構成することができる。

【0352】

場合によっては、あるゲートウェイリソース、例えば、ゲートウェイリソース216は、1つ以上のハンドオフ標的を用いて1つ以上のハンドオフを開始するよう、1つ以上のキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2080に推奨するように構成することができる（キャッシュを伴うコンピュータノード2080が、そのような推奨に従ってハンドオフを開始している、そのようなハンドオフは画面中に示されていない）。

20

【0353】

場合によっては、あるキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2070およびキャッシュを伴うコンピュータノード2080は、1つ以上のコンピュータノード（それに接続されたキャッシュリソースを伴う、または伴わない）および／またはクライアントサーバおよび／またはゲートウェイリソース、例えば、コンピュータノード205、ゲートウェイリソース216、クライアントサーバ218から1つ以上の推奨を受けて、1つ以上のハンドオフ標的を用いて1つ以上のハンドオフを開始するように構成することができる。

30

【0354】

場合によっては、あるキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2070は、1つ以上のコンピュータノード（それに接続されたキャッシュリソースを伴う、または伴わない）および／またはクライアントサーバおよび／またはゲートウェイリソース、例えば、コンピュータノード205、クライアントサーバ218から1つ以上の推奨を受けて、1つ以上のハンドオフ標的を用いて1つ以上のハンドオフを開始し、場合によっては同時に、別のキャッシュを伴うコンピュータノード、例えば、キャッシュを伴うコンピュータノード2080に1つ以上のキャッシュオブジェクト空間またはその部分の所有権を譲渡する、ハンドオフインシエータの役割を果たすように構成することができる。

40

【0355】

当業者であれば理解することができるよう、上記で例示されるシナリオは実施例にすぎず、図19で提供される説明図で提示されていない多数の他のシナリオが存在し得ることに留意されたい。

【0356】

ここで、本開示主題のある特定の実施例による、ハンドオフインシエータによってキャッシュハンドオフを行うために実行される一連の動作を図示する、図20に注目する。

【0357】

本開示主題のいくつかの実施例によれば、ハンドオフインシエータ（それに接続されたキャッシュリソースを有する別のコンピュータノード205への1つ以上のキャッシュオ

50

プロジェクト空間またはその部分の所有権の譲渡を開始している、それに接続されたキャッシュリソースを有するコンピュータノード205)と関連付けられるキャッシュハンドオフモジュール2530は、1つ以上のキャッシュオブジェクト空間またはその部分の所有権のハンドオフを開始するように構成することができる(ブロック1710)。

【0358】

この目的で、キャッシュハンドオフモジュール2530は、ハンドオフプロセスの影響を受ける、キャッシュオブジェクト空間またはその部分のうちのそれぞれのハンドオフ後の場所を示す、新しいハンドオフ後キャッシュマッピングを作成するように構成することができる(ブロック1720)。

【0359】

各クライアントサーバ218(またはDSS200の任意の他のユーザ)は、クライアントサーバ218(またはDSS200の任意の他のユーザ)が参照する1つ以上の論理的記憶エンティティに関する、キャッシュオブジェクト空間を処理する1つ以上のコンピュータノード205を示す、1つ以上のローカルキャッシュマッピングを有することができることに留意されたい。キャッシュハンドオフは、とりわけブロック1770を参照してさらに詳述されるように、適宜にそのような1つ以上のキャッシュマッピングの更新を必要とすると理解することができる。

【0360】

いくつかの非限定的実施例では、そのようなキャッシュマッピングは、ハンドオフプロセスに起因する、キャッシュオブジェクト空間の任意の分割、マージ、および場所の変更をコンパクトに反映する、分割関数の階層構造を使用することによって表すことができる。そのような分割関数のいくつかの実施例は、奇数および偶数のアドレス指定されたオブジェクト等を分割する、ハッシュ関数である。そのようなキャッシュマッピング記述およびそのような分割関数は実施例にすぎず、加えて、または代替として、任意の他の既知の方法および/または技法を利用できることに留意されたい。

【0361】

キャッシュハンドオフモジュール2530はさらに、例えば、「ハンドオフ進行中」状態を示すローカルフラグを設定することによって、そのような状態になり(ブロック1730)、それを用いてハンドオフプロセスを監視する要求を示す、ハンドオフ後キャッシュマッピングを含む「ハンドオフ開始」通知を、ハンドオフが要求される、ハンドオフ標的(ハンドオフインシエータから、1つ以上のキャッシュオブジェクト空間またはその部分の所有権を受領するように選択される、それに接続されたキャッシュリソースを有するコンピュータノード205)に送るように構成することができる(ブロック1740)。キャッシュハンドオフモジュール2530はさらに、ハンドオフ標的からのハンドオフ要求への応答を待つ(例えば、事前定義または計算された時間枠にわたって等)ように構成することができる。ハンドオフ標的は、そのようなハンドオフ要求を容認する、そのようなハンドオフ要求を拒絶する、またはそのようなハンドオフ要求を部分的に容認する(例えば、ハンドオフインシエータがそれに譲渡しようとしている、1つ以上のキャッシュオブジェクト空間またはその部分の一部のみの所有権を受容することを容認する等)ことができることに留意されたい。

【0362】

キャッシュハンドオフモジュール2530はまた、ハンドオフ要求がハンドオフ標的によって受容されたかどうかをチェックするように構成することもできる(ブロック1750)。ハンドオフ要求が拒絶された場合(または場合によっては、応答が所定の時間枠内で受けられなかった場合等)、キャッシュハンドオフモジュール2530は、「ハンドオフ進行中」状態を終了するように構成することができる(ブロック1760)。

【0363】

しかしながら、要求が容認された場合、キャッシュハンドオフモジュール2530は、ハンドオフ後キャッシュマッピングを、1つ以上のクライアントサーバ218(またはDSS200の任意の他の関連ユーザ)に、例えば、譲渡される1つ以上のキャッシュオブ

10

20

30

40

50

ジェクト空間（例えば、それと関連付けられる論理的記憶エンティティへのアクセス権を有する）またはその部分と関連付けられるクライアントサーバ218に、送るように構成することができる（ブロック1770）。場合によっては、新しいキャッシュマッピングを、DSS200に接続された全てのクライアントサーバ218（またはDSS200の任意の他の関連ユーザ）に送ることができることに留意されたい。

【0364】

加えて、キャッシュハンドオフモジュール2530は、例えば、ハンドオフ後キャッシュマッピングにおいてハンドオフ標的にマップされた非ダーティなキャッシュされたオブジェクト（関連キャッシュオブジェクト空間と関連付けられる永続記憶装置に最後に保存されて以来、修正されなかったキャッシュオブジェクト）を、（例えば、それらのデータを伝送することによって）ハンドオフ標的に送ることによって、および／またはそれらを「忘れる」ことによって（例えば、それらを削除されたとしてマークすること、またはハンドオフインシエータのキャッシュリソースからそれらを物理的に削除することによって）、そのようなキャッシュされたオブジェクトの所有権をハンドオフ標的に譲渡するよう構成することができる（ブロック1780）。

10

【0365】

なおも加えて、キャッシュハンドオフモジュール2530は、ハンドオフ後キャッシュマッピングにおいてハンドオフ標的にマップされるダーティキャッシュオブジェクト（関連キャッシュオブジェクト空間と関連付けられる永続記憶装置に最後に保存されて以来、修正されたキャッシュオブジェクト）を、一気に消去すること（それらを関連キャッシュオブジェクト空間と関連付けられる永久記憶装置に保存すること）によって、および／または（例えば、それらにデータを伝送することによって）そのようなキャッシュされたオブジェクトをハンドオフ標的に送り、ついで、それらを「忘れる」ことによって、それらの所有権を譲渡するよう構成することができる（ブロック1790）。

20

【0366】

場合によっては、キャッシュハンドオフモジュール2530はさらに、ハンドオフが行われたことを示す「ハンドオフ終了」通知をハンドオフ標的に送るように（ブロック1795）、かつ「ハンドオフ進行中」状態を終了するように（ブロック1760）構成することができる。場合によっては、ブロック1795は、ブロック1780および1790の実行が完了した後のみ実施することができる。

30

【0367】

図20を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および／または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施できることにも留意されたい。

【0368】

ここで、本開示主題のある特定の実施例による、ハンドオフ標的によってキャッシュハンドオフを行うために実行される一連の動作を図示する、図21に注目する。

40

【0369】

本開示主題のいくつかの実施形態によれば、ハンドオフ標的と関連付けられるキャッシュハンドオフモジュール2530は、ハンドオフインシエータから「ハンドオフ開始」通知（ハンドオフ後キャッシュマッピングを含む）を受け、ローカルのハンドオフ前キャッシュマッピングのバックアップコピーを保存し、ハンドオフインシエータから受け取られるハンドオフ後キャッシュマッピングに従ってキャッシュマッピングを更新し、「ハンドオフ進行中」状態になるように構成することができる（ブロック1810）。

【0370】

キャッシュハンドオフモジュール2530はさらに、場合によっては、例えば、キャッシュ関連SSL、および／またはDSS200あるいはその部品を指す閾値および／また

50

は S L S (例えば、最大許容サイトレベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等) 、および / またはそれと関連付けられる U D S P データリポジトリ 3 3 0 の中に記憶されたキャッシュ関連パラメータのハンドオフ標的知識に従って、ハンドオフイニシエータから受け取られるハンドオフ要求が容認可能であるかどうかをチェックするように構成することができる (ブロック 1 8 2 0) 。

【 0 3 7 1 】

場合によっては、ハンドオフ標的は、ハンドオフイニシエータがアクセスできるそのような情報 (および場合によっては、新しいおよび / またはより最新の情報) とは異なる、キャッシュ関連 S L S 、および / または D S S 2 0 0 あるいはその部分を指す閾値および / または S L S (例えば、最大許容サイトレベルオーバーコミット、最大許容全体オーバーコミット、種々のセキュリティパラメータ等) 、および / またはそれと、および / またはそれに接続されたキャッシュリソースと関連付けられるキャッシュ関連パラメータに関する情報へのアクセスを有することができると理解することができる。例えば、場合によっては、ハンドオフ要求がハンドオフ標的によって受け取られる時間までに、それに接続されたキャッシュリソースに関する 1 つ以上のパラメータが、すでに変化している。

【 0 3 7 2 】

(例えば、ハンドオフ標的知識に基づいて) ハンドオフ要求が容認できない場合、キャッシュハンドオフモジュール 2 5 3 0 は、拒絶通知 (ハンドオフ標的がハンドオフイニシエータによって送られたハンドオフ要求を容認しないことを示す) をハンドオフイニシエータに送り、ローカルのハンドオフ前キャッシュマッピング (ブロック 1 8 1 0 でバックアップのために保存された) を修復し、「ハンドオフ進行中」状態を終了するように構成することができる (ブロック 1 8 3 0) 。

【 0 3 7 3 】

ハンドオフ要求が容認できる場合、キャッシュハンドオフモジュール 2 5 3 0 は、容認通知 (ハンドオフ標的がハンドオフイニシエータによって送られたハンドオフ要求を容認することを示す) をハンドオフイニシエータに送るように構成することができる (ブロック 1 8 4 0) 。そのような場合、キャッシュハンドオフモジュール 2 5 3 0 は、1 つ以上のキャッシュオブジェクト空間またはその部分がハンドオフ標的の責任に譲渡されたことを示す、「ハンドオフ終了」通知を待ち、いったんそのような通知が受け取られると、「ハンドオフ進行中」状態を終了するように構成することができる (ブロック 1 8 5 0) 。

【 0 3 7 4 】

図 2 1 を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および / または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に関しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施することにも留意されたい。

【 0 3 7 5 】

ここで、本開示主題のある特定の実施例による、ハンドオフ中にハンドオフイニシエータによって受け取られるオブジェクト関連要求を処理するために実行される一連の動作を図示する、図 2 2 に注目する。

【 0 3 7 6 】

本開示主題のいくつかの実施例によれば、ハンドオフイニシエータと関連付けられるオブジェクト要求管理モジュール 2 5 4 0 は、例えば、ハンドオフ中に、クライアント (例えば、クライアントサーバ 2 1 8 、ゲートウェイリソース 2 1 6 、または任意の他のソース) から、ハンドオフ前および / またはハンドオフ後キャッシュマッピング内のオブジェクトに関するオブジェクト関連要求 (例えば、読み取り / 書込要求) を受けるように構成することができる (ブロック 1 9 1 0) 。

【 0 3 7 7 】

場合によっては、要求クライアントが、ハンドオフ後キャッシュマッピングに従ってそ

10

20

30

40

50

のローカルキャッシュマッピングを更新する前に、要求を送信できる（場合によっては、ハンドオフ後キャッシュマッピングが、例えば、ネットワーク上の重いトラフィックにより、または任意の他の理由で、即時にクライアントによって受け取られないため）ことに留意されたい。したがって、そのようなオブジェクト関連要求を、ハンドオフインシエータに送ることができる一方で、要求されたオブジェクトは、それによってもはや所有（例えば、処理）されていない。したがって、そのようなオブジェクト関連要求の受領時に、オブジェクト要求管理モジュール 2540 は、ハンドオフ後キャッシュマッピングに従って、要求されたオブジェクトがハンドオフインシエータの所有権の下にあるかどうかをチェックするように構成することができ（ブロック 1920）、もしそうであれば、オブジェクト要求管理モジュール 2540 は、要求を処理するように構成することができる（ブロック 1930）。

【0378】

しかしながら、ハンドオフインシエータが、ハンドオフ後キャッシュマッピングに従って、要求されたオブジェクトの所有者ではない場合には、オブジェクト要求管理モジュール 2540 は、（例えば、ハンドオフインシエータがまだ要求されたオブジェクトに所有権を譲渡していなかったため）要求されたオブジェクトが依然としてハンドオフインシエータの所有権の下にあるかどうかをチェックするように構成することができ（ブロック 1940）。

【0379】

そのような要求されたオブジェクトが依然としてハンドオフインシエータによって所有されている場合、オブジェクト要求管理モジュール 2540 は、要求を処理するように構成することができる（ブロック 1930）。しかしながら、そのような要求されたオブジェクトがもはやハンドオフインシエータによって所有されていない場合、オブジェクト要求管理モジュール 2540 は、オブジェクト関連要求をハンドオフ標的に伝えるように構成することができる（ブロック 1950）。

【0380】

図 22 を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および / または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に關しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施されることにも留意されたい。

【0381】

ここで、本開示主題のある特定の実施例による、ハンドオフ中にハンドオフ標的によって受け取られる、ハンドオフ後キャッシュマッピングに含まれるオブジェクトに関する、オブジェクト関連要求を処理するために実行される一連の動作を図示する、図 23 に注目する。

【0382】

本開示主題のいくつかの実施例によれば、ハンドオフ標的と関連付けられるオブジェクト要求管理モジュール 2540 は、例えば、クライアント（例えば、クライアントサーバ 218、ゲートウェイリソース 216、または任意の他のソース）から、またはハンドオフインシエータから、ハンドオフ後キャッシュマッピング内のオブジェクトに関するオブジェクト関連要求（例えば、読み取り / 書込要求）を受けるように構成することができる（ブロック 2110）。そのような要求の受領時に、オブジェクト要求管理モジュール 2540 は、要求がハンドオフインシエータから発信したかどうかをチェックするように構成することができる（ブロック 2120）。

【0383】

本明細書で示されるように、ハンドオフインシエータは、ハンドオフ後キャッシュマッピングに従って、ハンドオフインシエータが要求されたオブジェクトの所有者ではない、およびハンドオフインシエータが要求されたオブジェクトの現在の現在の所有者ではない

10

20

30

40

50

という判定後に、オブジェクト関連要求をハンドオフ標的に伝えるであろう（ハンドオフイニシエータが依然として所有者であるという指示は、例えば、オブジェクトがハンドオフイニシエータのキャッシュの中で依然としてダーティであること等であり得る）。したがって、要求がハンドオフイニシエータから発信するときはいつでも、これは、ハンドオフイニシエータが要求されたオブジェクトの所有権をハンドオフ標的にすでに譲渡していることを示すので、オブジェクト要求管理モジュール 2540 は、要求を処理するように構成することができる（ブロック 2130）。

【0384】

要求がハンドオフイニシエータから発信しなかった場合、オブジェクト要求管理モジュール 2540 は、（例えば、ハンドオフ標的がハンドオフイニシエータから要求されたオブジェクトの所有権をすでに受けている等のため）要求されたオブジェクトがハンドオフ標的によって所有されるかどうかをチェックするように構成することができる（ブロック 2140）。それがハンドオフ標的によって所有される場合、オブジェクト要求管理モジュール 2540 は、要求を処理するように構成することができる（ブロック 2130）。

10

【0385】

（例えば、要求されたオブジェクトの所有権がまだ転送されず、ハンドオフプロセスが終了しなかったため）要求されたオブジェクトがハンドオフ標的によって所有されない場合、オブジェクト要求管理モジュール 2540 は、要求をハンドオフイニシエータに伝えるように構成することができる（ブロック 2150）。そのようなシナリオは、例えば、あるユーザが、更新されたハンドオフ後キャッシュマッピングを有することができるという事実に起因し得、要求されたオブジェクトがハンドオフ標的にマップされる一方で、それが要求されたオブジェクトを要求している間に、ハンドオフプロセスが依然として進行中であり、したがって、要求されたオブジェクトの所有権がまだハンドオフ標的に譲渡されていなかったことを示すことに留意されたい。

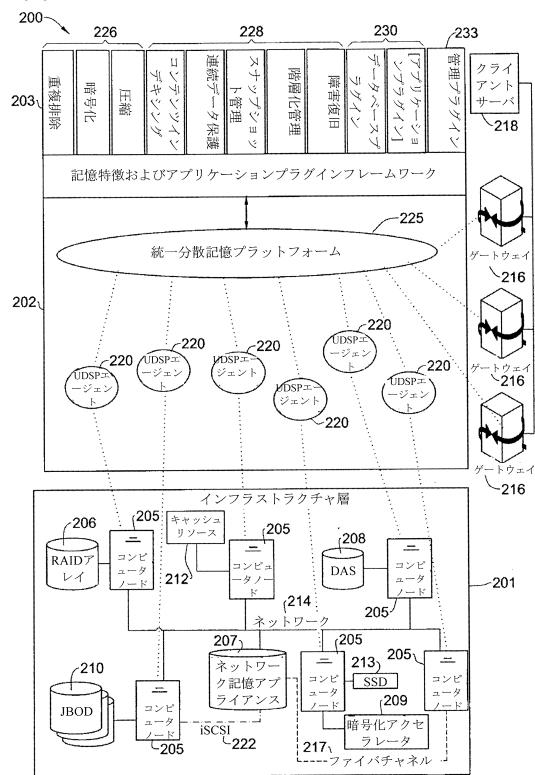
20

【0386】

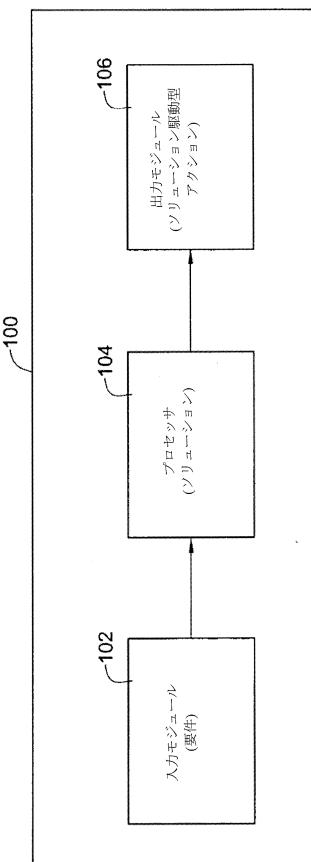
図 23 を参照すると、ブロックのうちのいくつかを連結ブロックに統合することができ、またはいくつかのブロックに分解することができ、および / または他のブロックが追加されてもよいことに留意されたい。さらに、場合によっては、ブロックは、本明細書で説明されるのとは異なる順序で実施することができる。また、それらを実現するシステム要素に關しても、フロー図が説明されるが、これは決して義務的ではなく、ブロックは、本明細書で説明される以外の要素によって実施することにも留意されたい。

30

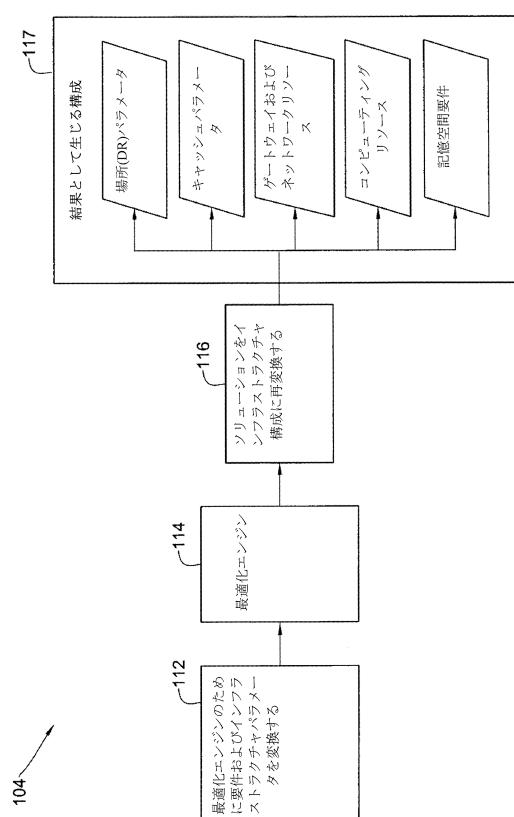
〔 四 1 〕



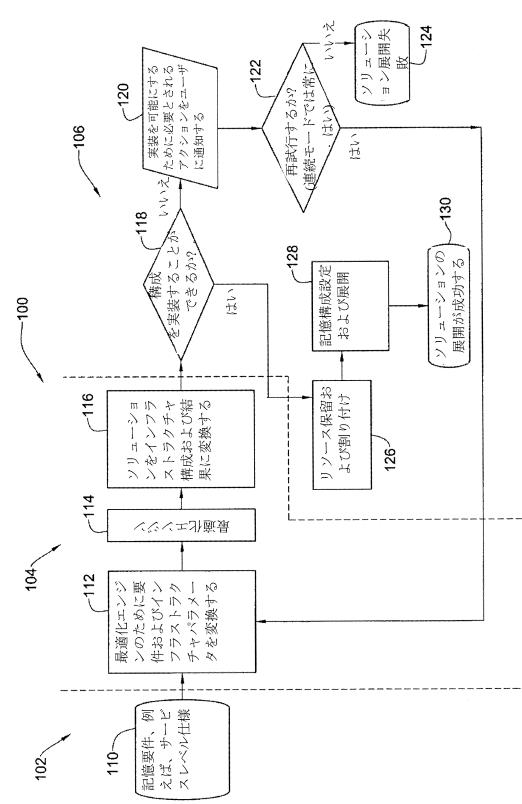
【図2】



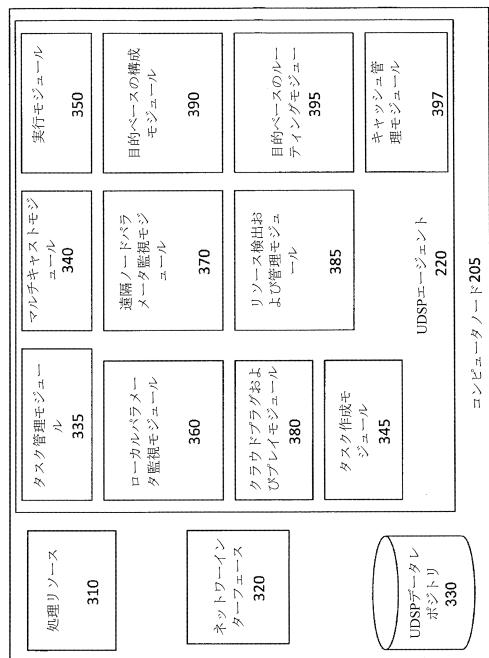
(3)



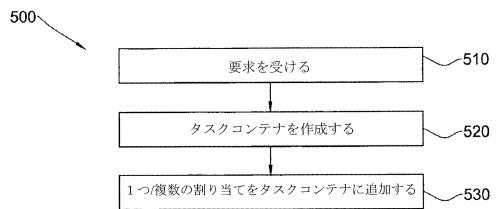
【 4 】



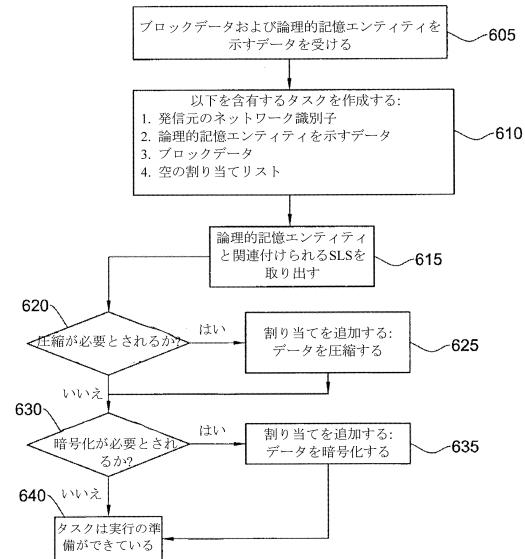
【図5】



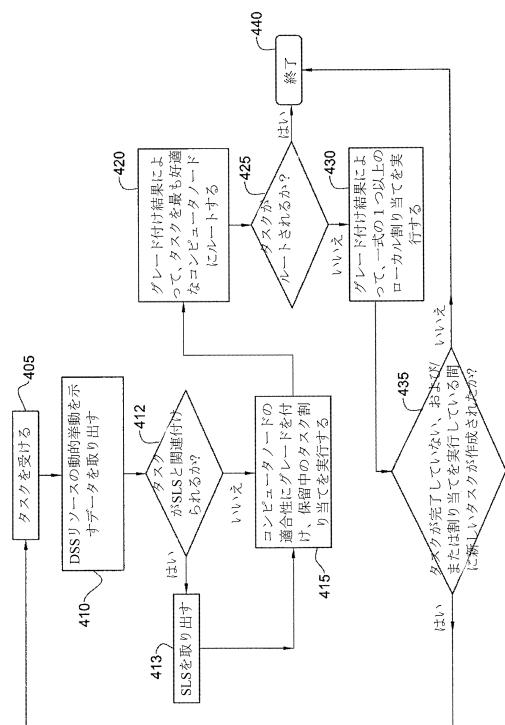
【図6】



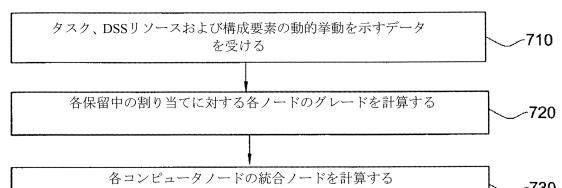
【図7】



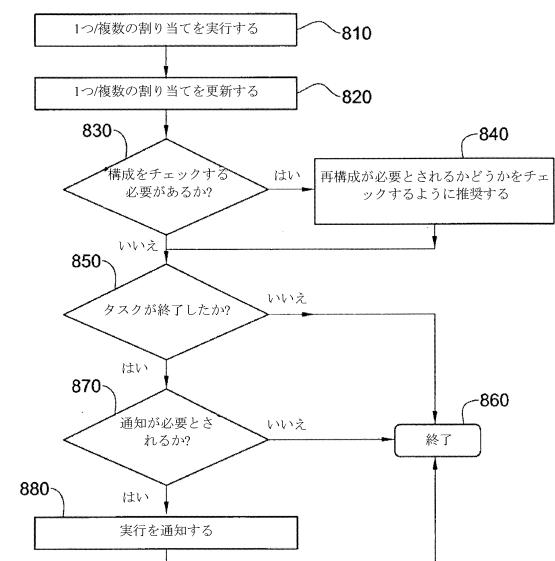
【図8】



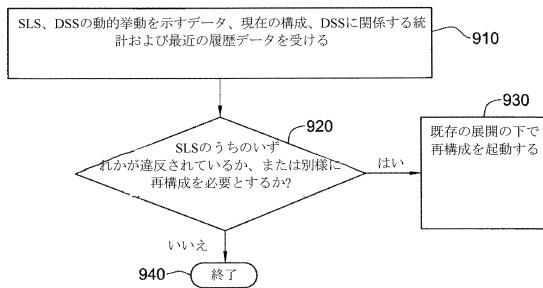
【図9】



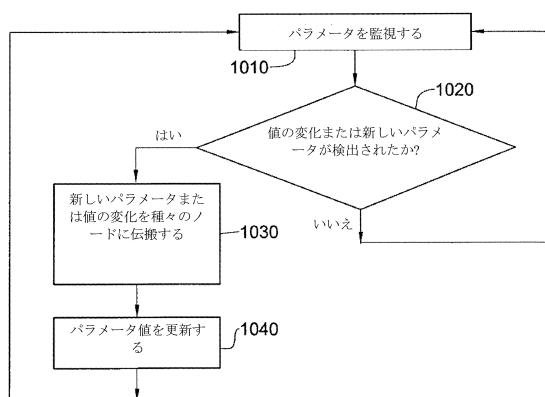
【図10】



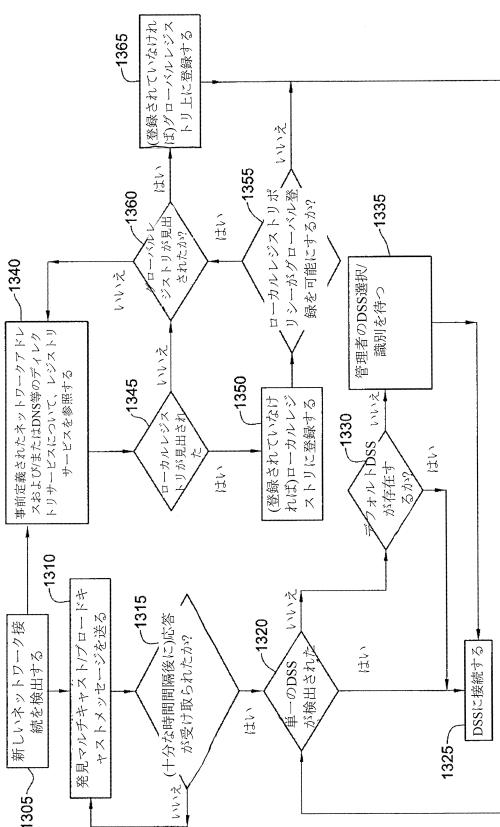
【図11】



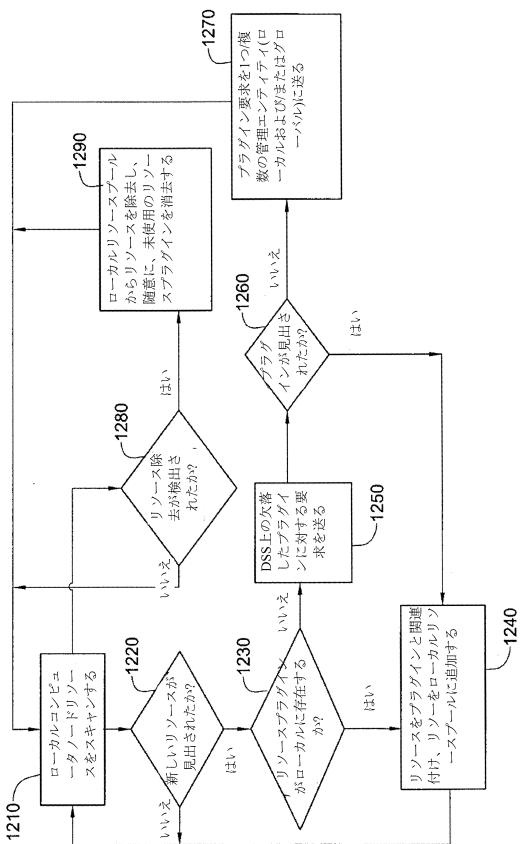
【 図 1 2 】



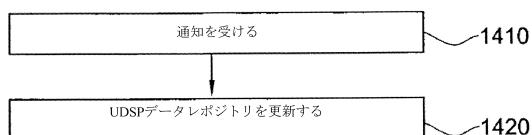
【図14】



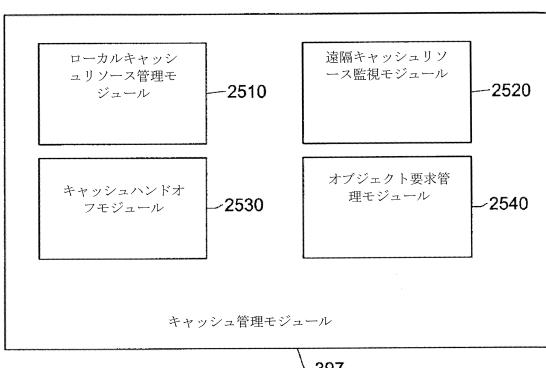
【図13】



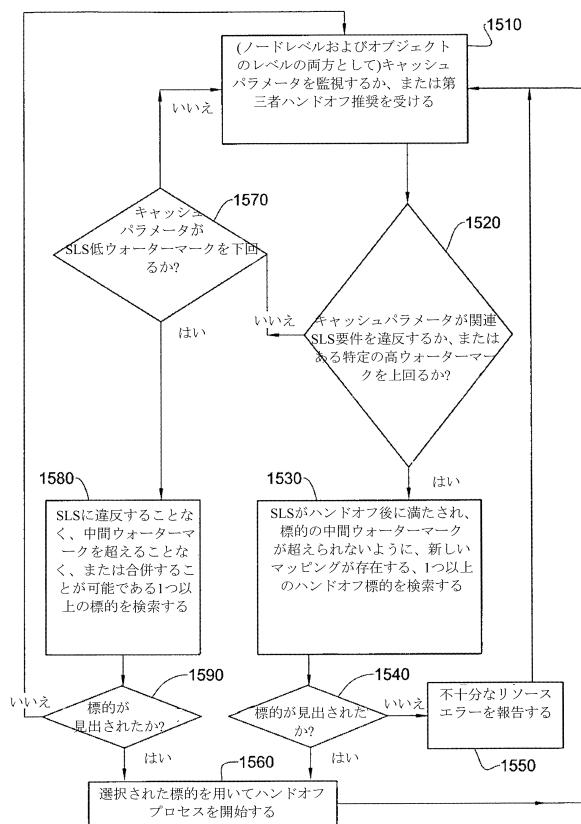
【 15 】



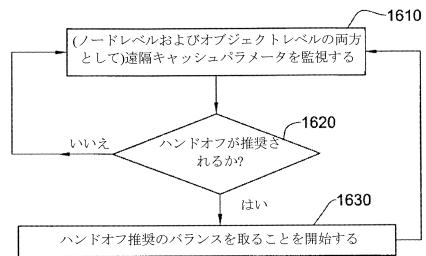
【 义 1 6 】



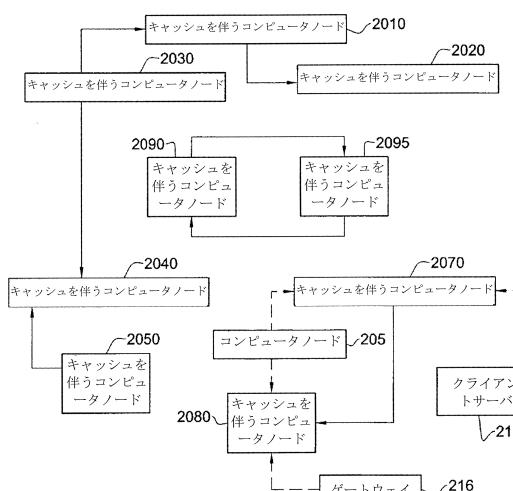
【図17】



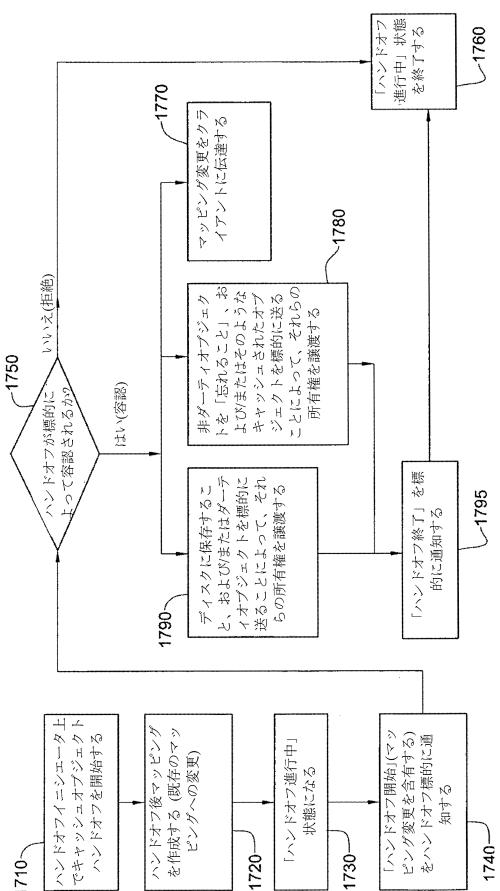
【図18】



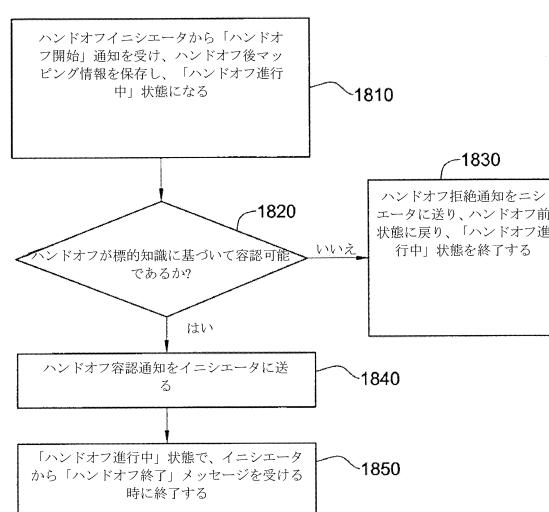
【図19】



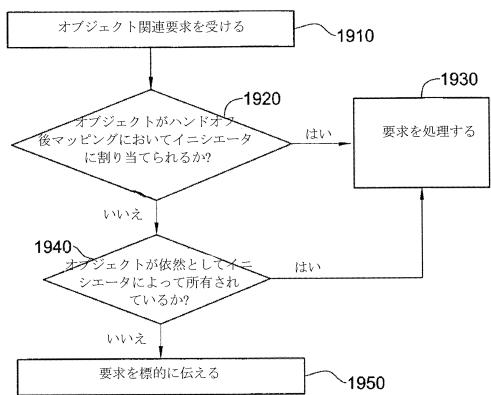
【 20 】



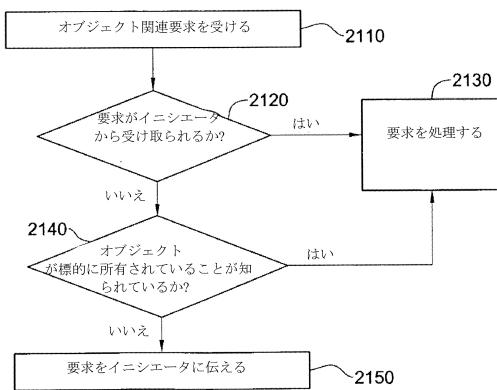
【図21】



【図2-2】



【図2-3】



フロントページの続き

(72)発明者 キプニス, タル

イスラエル国, 53235 ギバアタイム, 4/3 モーデイ ハゲタオット

(72)発明者 ローウェンバーグ, ガイ

イスラエル国, 53361 ギバタイム, 20 ジャボティンスキー ストリート

審査官 漆原 孝治

(56)参考文献 特表2007-531130 (JP, A)

特表2005-535949 (JP, A)

林 通秋, 大規模資源の管理・制御に関する技術の実証実験, 電子情報通信学会技術研究報告,

日本, 社団法人電子情報通信学会, 2010年11月11日, 第110巻, 第286号, pp.39-44

(58)調査した分野(Int.Cl., DB名)

G 06 F 12/00