

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4265245号
(P4265245)

(45) 発行日 平成21年5月20日(2009.5.20)

(24) 登録日 平成21年2月27日(2009.2.27)

(51) Int. Cl. F 1
G 0 6 F 1 2 / 0 0 (2 0 0 6 . 0 1) G 0 6 F 1 2 / 0 0 5 4 5 B

請求項の数 3 (全 19 頁)

(21) 出願番号	特願2003-71189 (P2003-71189)	(73) 特許権者	000005108 株式会社日立製作所 東京都千代田区丸の内一丁目6番6号
(22) 出願日	平成15年3月17日(2003.3.17)	(74) 代理人	100100310 弁理士 井上 学
(65) 公開番号	特開2004-280485 (P2004-280485A)	(72) 発明者	原 純一 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所 内
(43) 公開日	平成16年10月7日(2004.10.7)	(72) 発明者	岩崎 正明 神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所 内
審査請求日	平成17年12月19日(2005.12.19)		

最終頁に続く

(54) 【発明の名称】 計算機システム

(57) 【特許請求の範囲】

【請求項1】

ストレージ装置を有するサーバ計算機が複数と、前記サーバ計算機の各々の第一の評価項目の値と第二の評価項目の値を含むサーバ情報管理テーブル有する管理用計算機と、前記第一の評価項目の条件値と前記第一の評価項目よりも優先度の低い前記第二の評価項目の条件値とを含む配置規則情報を有する要求元計算機と、から構成される情報処理システムのファイル格納方法であって、

前記要求元計算機は、ユーザからファイルの作成指示を受信し、前記ファイルを分割することで複数の部分ファイルを生成し、前記第一の評価項目の条件値と前記第二の評価項目の条件値とを前記管理用計算機へ送信し、

前記管理用計算機は、前記サーバ情報管理テーブルに基づいて、前記要求元計算機から受信した前記第一の評価項目の条件値と前記第二の評価項目の条件値とを満たす前記サーバ計算機の第一の集合を特定し、前記第一の集合の前記サーバ計算機の情報を含む第一のリストを前記要求元計算機へ送信し、

前記要求元計算機は、受信した前記第一のリストに基づいて、前記第一の集合に含まれる前記サーバ計算機の数と前記複数の部分ファイルの数とを比較し、

前記第一の集合に含まれる前記サーバ計算機の数が前記複数の部分ファイルの数より少ない場合、前記管理用計算機は、前記第一の評価項目の条件値を前記管理用計算機へ送信し、

前記管理用計算機は、前記サーバ情報管理テーブルに基づいて、前記要求元計算機から

10

20

受信した前記第一の評価項目の条件値を満たす前記サーバ計算機の第二の集合を特定し、前記第二の集合の前記サーバ計算機の情報を含む第二のリストを前記要求元計算機へ送信し、

前記要求元計算機は、前記第二の集合の前記サーバ計算機の数の前記複数の部分ファイルの数以上の場合は、前記第二のリストによって指定された一つ以上の前記サーバ計算機へ前記複数の部分ファイルを送信し、前記第二の評価項目が除外されたことをユーザーへ示す、

ことを特徴とする情報処理システムのファイル格納方法。

【請求項 2】

請求項 1 記載のファイル格納方法であって、

前記第一の評価項目は信頼性であり、前記第二の評価項目はセキュリティであり、

前記配置規則情報は負荷に関する条件値を含み、

前記要求元計算機は、前記サーバ計算機の各々から負荷値を受信し、前記負荷に関する条件値を満たすサーバ計算機を特定する、

ことを特徴とする情報処理システムのファイル格納方法。

【請求項 3】

請求項 2 記載のファイル格納方法であって、

前記管理用計算機は、前記サーバ情報管理テーブルを作成するために、前記サーバ計算機の各々から前記第一の評価項目の値と前記第二の評価項目の値を受信する、

ことを特徴とする情報処理システムのファイル格納方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数の計算機がネットワークを介して接続された計算機システムにおいて、ファイルを分散して管理する分散ファイルシステムに関する。

【0002】

【従来の技術】

計算機システムにおける従来の分散ファイルシステム、具体的には、計算機等が使用するプログラム等が使用する一かたまりのデータ（以下「ファイル」と称する）を複数に分割し、その分割された部分（以下「部分ファイル」と称する）の各々を、複数の計算機に分散して配置する技術の例が、特許文献 1 に開示されている。

【0003】

また、ファイル分散の方法として、あらかじめ計算機間で各計算機の負荷、例えば、計算機が有するデータの最大転送レートの使用率や計算機が有するストレージ装置の使用率についての情報を共有し、この負荷情報に基づいて部分ファイルを各計算機に配置する方法が特許文献 2 に記載されている。また、特許文献 2 には、ある計算機に高い負荷が掛かった場合、負荷の集中している計算機に格納されている部分ファイルを他の計算機へ移動することも記載されている。

【0004】

一方、非特許文献 1 には、数千数万を超える計算機が接続されることを前提とした広域ネットワークにおける分散ファイルシステムが提案されている。本文献では、広域ネットワークに接続される各計算機のセキュリティレベルや信頼性レベルなどには差異があることから、分散ファイルシステムにおいて、ある一つの部分ファイルから他の部分ファイルの位置が特定できないようにすることが開示されている。これにより、分散ファイルシステムのセキュリティを確保することができる。

【0005】

【特許文献 1】

特開平 8 - 77054 号公報

【特許文献 2】

特開 2000 - 207370 号公報

10

20

30

40

50

【非特許文献1】

“ Maintenance - Free Global Data Storage ” (I
E E E I n t e r n e t C o m p u t i n g , V o l 5 , N o 5 , S e p t e m
b e r / O c t o b e r 2 0 0 1 , p p 4 0 - 4 9)

【0006】**【発明が解決しようとする課題】**

特許文献2に開示された分散ファイルシステムでは、負荷情報のみに基づいて部分ファイルの移動が行われる。したがって、他の計算機と比較して高性能な計算機がシステム内に存在する場合、一つのファイルを構成する全ての部分ファイルが同一の計算機へ移動される可能性がある。これでは、一つのファイルを構成する部分ファイルを分散した意味が無くなる。この弊害を解消するためには、部分ファイルの移動先に同一のファイルを構成する他の部分ファイルがあるかどうかを確認するようにすればよい。具体的には、ある部分ファイルの情報から、同一のファイルを構成する他の部分ファイルの位置を特定できるようにすればよい。しかし、この方法では、非特許文献1に開示されているように、分散ファイルシステムにおけるセキュリティが確保されない。

10

【0007】

また、従来の分散ファイルシステムにおいては、部分ファイルの配置を決定する規則又は部分ファイルを移動する規則（以下、「配置規則」という）の作成には、各計算機の負荷情報しか考慮されていない。したがって、従来技術の分散ファイルシステムでは、各計算機において性能やセキュリティレベル、信頼性レベル、使用料金等に差異がある場合に、

20

【0008】

更に、特許文献1、2及び非特許文献1に開示された分散ファイルシステムにおいては、複数の情報に基づく複数の配置規則を規定し、それを使い分けることは考慮されていない。

【0009】

本発明の目的は、ファイルの部分ファイルを移動する場合において、セキュリティを確保しつつ、同一ファイルの全ての部分ファイルが一つの計算機に集まってしまう問題を解決する分散ファイル管理方法、及びこの方法を用いた分散ファイルシステムを提供することにある。

30

【0010】

また、本発明の他の目的は、多様な配置規則に基づいて部分ファイルの配置を行う計算機システムにおけるファイル管理方法、及びこの方法を用いた分散ファイルシステムを提供することにある。

【0011】**【課題を解決するための手段】**

上記課題を解決するため、本発明の計算機システムは、以下の構成を有する。即ち、データを記憶するストレージ装置を備える複数の計算機（以下「計算機1」）、ユーザが使用する計算機（以下「計算機2」）及びこれら計算機を管理する計算機（以下「管理用計算機」）とがネットワークで接続された計算機システムである。本システムにおいて、計算機1は、自己の性能、セキュリティレベル、信頼性レベル、使用料金、及び負荷情報等を管理し、管理している情報を管理用計算機に送信する。管理用計算機は、送信されたこれらの情報をまとめて管理する。一方、計算機2は、ファイルを構成する部分ファイルの分散を規定する、複数の項目からなる情報（配置規則）を有する。

40

【0012】

ここで、ユーザの指示等により分散されるファイルの作成が計算機2に指示された場合、計算機2は、配置規則を満足する計算機1についての情報を得るために、管理用計算機に配置規則についての情報を送信する。配置規則を受信した管理用計算機は、自己が管理する各計算機1の情報を参照して、受信した配置規則に適合する計算機1を特定し、特定された計算機1の情報を計算機2に送信する。

50

【 0 0 1 3 】

配置規則に適合する計算機 1 の情報を受信した計算機 2 は、その情報に基づいて、ファイルの部分ファイルを配置規則を満たす各計算機 1 に送信する。

【 0 0 1 4 】

また、本発明の計算機システムは、ある特定の契機において部分ファイルの再配置を行う態様を有しても良い。

【 0 0 1 5 】

この場合、ある特定の契機において、計算機 2 は、再配置の対象となるファイルを構成する部分ファイルを有する計算機 1 から、その計算機 1 に関する情報を受信する。計算機 1 に関する情報を受信した計算機 2 は、自己が有する配置規則と受信した計算機 1 に関する情報とを比較して、その時点で配置規則を満たさない計算機 1 の有無を確認する。

10

【 0 0 1 6 】

配置規則を満たさない計算機 1 が存在する場合、計算機 2 は、配置規則を満たす計算機 1 を確認するために、管理用計算機に配置規則に関する情報を送信する。配置規則を受信した管理用計算機は、自己が管理する計算機 1 に関する情報を参照して、受信した配置規則を満たす計算機 1 を特定し、その計算機 1 についての情報を計算機 2 に送信する。

【 0 0 1 7 】

配置規則を満たす計算機 1 の情報を受信した計算機 2 は、受信した計算機 1 の情報から、特定の計算機 1 を選択する。その後、計算機 2 は、配置規則を満たさない計算機 1 に対して、配置規則を満たす特定の計算機 1 へ部分ファイルを転送するよう指示を送信する。

20

【 0 0 1 8 】

尚、計算機 1 や計算機 2 が管理用計算機を兼ねていてもよく、計算機 1 が計算機 2 を兼ねていても良い。

【 0 0 1 9 】

また、計算機 2 が有する配置規則には、配置規則を構成する項目の優先度が設定される態様も考えられる。この場合、計算機 2 は、配置規則を満足する計算機 1 が存在しない場合、優先度の低い項目を配置規則から除外することによって、配置規則を緩和し、ファイルの分散先（再配置先）となる計算機 1 を確保する。尚、優先度が 0 の場合には、その項目は配置規則の内容としては考慮されない態様としても良い。

【 0 0 2 0 】

更に、計算機 2 は、複数の配置規則を有し、分散されるファイルの種類ごとに異なる配置規則を適用してファイルを分散する態様としても良い。また、ファイルの分散時にユーザ等に配置規則の各項目の選択又は優先度を設定させ、その選択又は設定に応じたファイル分散を行う態様でも良い。

30

【 0 0 2 1 】

また、配置規則を満たす計算機 1 が不足する場合、計算機 2 は、複数の部分ファイルを 1 つの計算機 1 に格納するようにしても良い。

【 0 0 2 2 】

また、部分ファイルの移動を指示された計算機 1 で、一定の条件を満たすまで部分ファイルの移動を保留する態様でも良い。この場合、計算機 1 は、一つの部分ファイルに対して複数の移動指示を受けることになるが、計算機 1 は、これら複数の移動指示に含まれる移動先の計算機 1 に重複する計算機 1 があるかどうかを確認し、重複している計算機 1 がある場合にのみ、部分ファイルの転送を実行する。ここで一定の条件とは、一定時間、または一定数指示を受け取るまでとするとよい。

40

【 0 0 2 3 】

【 発明の実施の形態 】

以下、本発明の実施の形態について、図面を用いて詳細に説明する。

【 0 0 2 4 】

図 1 は、本発明を適用した計算機システムの構成の一実施形態を示すブロック図である。図 1 の計算機システムは、計算機 1 A、1 B、1 C、2 A、2 B、及び管理用計算機 4 を

50

有し、これらはすべてネットワーク 3 で相互に接続されている。

【 0 0 2 5 】

図 1 1 は、計算機 1 の構成例を示す図である。計算機 1 は、ファイルや部分ファイルといったデータを記憶するハードディスクやディスクアレイ等のストレージ装置 1 4 及び制御部 1 5 を有する。制御部 1 5 は、プロセッサ 1 6、メモリ 1 7、インタフェース 1 8、及びネットワークインタフェース 1 9 を有する。プロセッサ 1 6 は、ストレージ装置 1 4 に格納された部分ファイルやファイルを管理するためのオペレーティングシステム等のプログラムを実行する。これらのプログラムは、メモリ 1 7 に格納される。また、これらのプログラムは、フロッピー（登録商標）ディスク等の可搬記憶媒体又はネットワーク 3 を介して計算機 1 にインストールされる。

10

【 0 0 2 6 】

尚、ストレージ装置 1 4 は、制御部 1 5 と別筐体でも、同一筐体に組み込まれていても良い。制御部 1 5 とストレージ装置 1 4 は、S C S I やファイバチャネル等のインタフェース 1 8 を介して相互に接続される。また、制御部 1 5 とストレージ装置 1 4 が別筐体である場合には、制御部 1 5 とストレージ装置 1 4 が、ファイバチャネル等で構築されたストレージエリアネットワークで接続されていてもよい。更に、制御部 1 5 とストレージ装置 1 4 が別筐体の場合、ストレージ装置 1 4 は更にストレージ装置 1 4 全体の制御を行う制御部を有していても良い。

【 0 0 2 7 】

メモリ 1 7 には、自サーバ状態管理プログラム 1 1、部分ファイル管理プログラム 1 2、及び部分ファイル再配置プログラム 1 3 が格納される。

20

【 0 0 2 8 】

自サーバ状態管理プログラム 1 1 は、計算機 1 の性能、セキュリティレベル、信頼性レベル、使用料金、又は負荷の情報等（以下、「サーバ情報」）を制御部 1 5 が管理する際に実行される。サーバ情報に変更があった場合、制御部 1 5 は、同プログラム 1 1 を実行して、ネットワーク 3 を介して、管理用計算機 4 にその変更内容を通知する。尚、本実施形態においては、制御部 1 5 は、サーバ情報のうち、頻繁に更新されることの少ない情報（ここでは、性能、セキュリティレベル、信頼性レベル、使用料金）の変更のみ、管理用計算機 4 に通知するようにする。

【 0 0 2 9 】

また、制御部 1 5 は、部分ファイル管理プログラム 1 2 の実行結果に基づいて、計算機 2 へ計算機 1 自身のサーバ情報を送信する。

30

【 0 0 3 0 】

部分ファイル管理プログラム 1 2 は、計算機 1 のストレージ装置 1 4 に格納されている部分ファイルを制御部 1 5 が管理する際に実行される。又、同プログラム 1 2 は、計算機 2 からの要求に基づいて、制御部 1 5 が、部分ファイルの書き込み、読み出し、及び要求元となる計算機 2 への部分ファイルの送信を行う際にも実行される。

【 0 0 3 1 】

部分ファイル再配置プログラム 1 3 は、計算機 2 からの部分ファイルの移動の指示に基づいて、制御部 1 5 が部分ファイルの移動先を決定したり、制御部 1 5 が部分ファイルの他の計算機 1 への転送を制御する際に実行される。また、同プログラム 1 3 は、他の計算機 1 の制御部 1 5 から部分ファイルの移動の受け入れ要求を受信した制御部 1 5 が、移動を要求された部分ファイルを受信してストレージ装置 1 4 へ保存する際にも実行される。

40

【 0 0 3 2 】

管理用計算機 4 は、計算機 1 と同様の構成であり、プロセッサ、メモリ、入出力部及びストレージ装置を有する。メモリには、分散ファイル情報管理テーブル 4 3、サーバ情報管理テーブル 4 4、分散ファイル管理プログラム 4 1 及びサーバ情報管理プログラム 4 2 が格納されている。

【 0 0 3 3 】

分散ファイル管理プログラム 4 1 は、分散されて管理されるファイル（以下「分散ファイ

50

ル」)を構成する各部分ファイルの位置を管理用計算機4が管理する際に実行される。管理用計算機4は、計算機2からの分散ファイル作成の通知に基づいて、通知された分散ファイルを構成する各部分ファイルにそれぞれ一意のID(以下「部分ファイルID」)を決定し、分散ファイル名、部分ファイルID、及び各部分ファイルを格納する計算機名を関連付けて管理する。この管理に使用されるテーブルが分散ファイル情報管理テーブル43である。本テーブル43の内容については後述する。

【0034】

また、同プログラム41は、管理用計算機4が、計算機2からの、分散ファイルを構成する部分ファイルのID及び位置に関する問い合わせに応じて、部分ファイルID及び部分ファイルの位置のリストを、問い合わせ元の計算機2に送信する際にも実行される。

10

【0035】

サーバ情報管理プログラム42は、管理用計算機4が、各計算機1の制御部15から送信されるサーバ情報を管理する際に実行される。サーバ情報の管理に使用されるテーブルがサーバ情報管理テーブル44である。本テーブル44の内容例については後述する。また、同プログラム42は、管理用計算機4が、計算機2からの問い合わせに対し、計算機2からの問い合わせに含まれる配置規則を満足する計算機1をサーバ情報管理テーブル44に基づいて検索し、その結果を計算機2へ返信する際にも実行される。

【0036】

尚、計算機1が分散ファイル管理プログラム41等を有することで、管理用計算機4を兼ねても良い。

20

【0037】

計算機2は、計算機1と同様の構成であり、プロセッサ、メモリ、入出力部、及びストレージ装置を有する。計算機2は、入出力部を介してユーザの指示を受け付ける。尚、計算機1又は管理用計算機4がユーザの指示を受け付ける入出力部を有することで計算機2を兼ねても良い。メモリには、配置規則23、分散ファイル操作プログラム22及び部分ファイル配置評価プログラム21とが格納されている。

【0038】

分散ファイル操作プログラム22は、計算機2が、ユーザからの分散ファイルの作成要求に対し、分散ファイルの部分ファイルへの分割、配置先となる計算機1の確認、及び各計算機1への部分ファイルの作成要求を発信する際に実行される。

30

【0039】

部分ファイル配置評価プログラム21は、計算機2が部分ファイルの配置先となる計算機1を確認する際に、予め計算機2に設定されている性能、セキュリティレベル、信頼性レベル、使用料金、負荷情報等の項目に基づく配置規則を満足する計算機1の有無を計算機2が管理用計算機4へ問い合わせの際に実行される。又、問い合わせの結果得られる配置規則を満たす計算機1の数が、配置すべき部分ファイル数未満である場合、計算機2は、同プログラム21を実行して配置規則から優先度の最も低い項目を除外し、再度管理用計算機4へ項目を除いた配置規則に合致する計算機1の有無を問い合わせる。

【0040】

また、計算機2は、各計算機1から受け取ったサーバ情報に基づき、各計算機1が予め設定されている配置規則を満足しているかどうかをチェックする。チェックした結果、配置規則を満たさない計算機1が存在する場合、計算機2は、配置規則を満たさない計算機1が保存している部分ファイルの移動先となる計算機1を決定する。

40

【0041】

具体的には、計算機2は、配置規則を満足する計算機1の有無を管理用計算機4へ問い合わせる。管理用計算機4は、計算機2からの要求に従って、要求に合致する計算機1を検索し、その結果を計算機2に送信する。その結果を受信した計算機2は、要求に合致する計算機1のうち、移動すべき部分ファイルが属する分散ファイルに属する他の部分ファイルを保存している計算機1を除外して、移動先となる計算機1を決定する。

【0042】

50

図1は、分散ファイルA、B、Cが作成された後の状態を例示している。計算機1Aのストレージ装置14には分散ファイルAの部分ファイルA1(141)、計算機1Bのストレージ装置14には分散ファイルAの部分ファイルA2(51)及び分散ファイルBの部分ファイルB1(52)、計算機1Cのストレージ装置14には分散ファイルBの部分ファイルB2(61)及び分散ファイルCの部分ファイルC1(62)が保存されている。

【0043】

図2は、図1の状態における、管理用計算機4が有する分散ファイル情報管理テーブル43の例を示したものである。管理用計算機4は、各分散ファイルの各部分ファイルに付された一意の部分ファイルIDと部分ファイルが格納される位置との対応関係を本テーブル43で管理する。ここで、各分散ファイルに属性情報がある場合には、この属性情報を各部分ファイルに対応付けて管理する。図2では、分散ファイルの属性情報の例として、分散ファイルの種類を管理している。分散ファイルの属性情報としては、他にファイルの所有者、ファイル作成時刻、ファイル更新時刻などが考えられる。これらの情報は、各計算機2が分散ファイル作成要求をユーザから受取った際に、管理用計算機4に送信される。

【0044】

図3は、管理用計算機4において管理されているサーバ情報管理テーブル44の例を示したものである。ここで、本実施形態の各計算機1では、性能としてストレージ装置14のデータ転送レート(Mバイト/秒)、セキュリティレベルとして各計算機1の管理者が設定した使用者等のレベル、信頼性レベルとして各計算機1の管理者が設定したレベル、使用料金として、各計算機1のストレージ装置14の使用容量に応じた料金(円/Mバイト)、負荷情報として、各計算機1のストレージ装置14の記憶領域の使用率(%)がサーバ情報として管理されているが、各計算機1から管理用計算機4へは、値の更新の頻度が少ない性能、セキュリティレベル、信頼性レベル、及び使用料金の情報のみが送信されると仮定する。これは、管理用計算機4と各計算機1との間の通信頻度を抑えるためである。したがって、管理用計算機4は、図3のように、各計算機1から送信された情報のみをサーバ情報管理テーブル44で管理する。

【0045】

尚、管理用計算機4へのサーバ情報の登録は、分散ファイルシステムの初期設定時に、ネットワーク3で接続された各計算機1から管理用計算機4へ自動的にサーバ情報が送信される構成としても良いし、ユーザ又は管理者が手動で設定する構成としても良い。また、サーバ情報に変更があった場合には、その都度、計算機1から管理用計算機4へ情報が転送されるか、ユーザ等が手動で登録を行う。

【0046】

図4は、計算機2Aにおいて予め規定されている配置規則のテーブル23の例である。本図においては、配置規則の項目として、性能、セキュリティ、信頼性、使用料金及び負荷が採用されているが、他の組み合わせでも構わない。

【0047】

また、本配置規則のテーブル23では、各項目に優先度が設定される。計算機2は、配置規則を満足する計算機1の数が不足する場合、各項目に設定された優先度に応じて、配置規則のテーブル23から項目を除外し、配置規則の条件を緩和する。これにより、次善の配置規則で部分ファイルを分散することが可能となる。

【0048】

本図においては、性能の項目を優先度3とし、性能値が40(Mバイト/秒)以上である計算機1に部分ファイルを配置することを規定している。また、セキュリティの項目を優先度1とし、セキュリティレベルの値が4以上である計算機1に部分ファイルを配置すること、また、信頼性の項目を優先度2とし、信頼性レベルの値が2以上である計算機1に部分ファイルを配置することを規定している。さらに、使用料金及び負荷の規則は優先度0とし、配置規則の項目としては適用しないことを示している。以下の説明では、優先度1が、優先度3よりも優先されるとして説明する。

【0049】

10

20

30

40

50

尚、本実施形態では、配置規則の項目の組み合わせが一つの場合について説明する。しかし、他の実施形態として、分散ファイルの配置時に、ユーザに項目の優先度を決定させることで、個々の分散ファイルの配置ごとに、ユーザが望む項目から構成される配置規則を使用する態様が考えられる。

【0050】

この場合、計算機2は、分散ファイル作成時に、入出力部に配置規則の項目を表示し、ユーザに項目の優先度を入力させる。ここで、ユーザは、優先したい項目について高い優先度を設定する。逆にユーザは優先度0を入力することで、その項目を配置規則から除外することも出来る。更に、新たな項目がユーザによって追加されても良い。

【0051】

このようにして設定された項目及びその優先度を、計算機2はその分散ファイルの配置規則のテーブル23としてメモリに記録すると共に、設定された配置規則の情報を管理用計算機4に送信する。管理用計算機4は、図2の分散ファイル情報管理テーブル43を図10のように拡張し、各分散ファイルごとの配置規則についても分散ファイルの属性情報として管理する。

【0052】

また、本実施形態では、分散ファイルの種類を問わず同じ配置規則を適用しているが、他の実施態様として、分散ファイルの種類によって適用する規則の選択/優先度の設定を自動で行う態様も考えられる。この場合、計算機2は、あらかじめ、図9に示すように、図4に示す配置規則のテーブル23をファイルの種類ごとに有する。そして、計算機2は、ユーザから分散ファイル作成が要求された場合に、要求されたファイルの種類に応じて図9の配置規則のテーブルから適用する配置規則を自動で選択し、必要があれば優先度を設定する。

【0053】

尚、計算機2に設定されるファイルの種類ごとの配置規則の情報は、計算機1の使用者等が手動で入力したり、新たに計算機2がネットワーク3に接続された際に、管理用計算機4から配置規則の情報が自動的に転送されても良い。さらに、新たに追加される配置規則の項目の情報は、各計算機1及び管理用計算機4に転送され、各計算機1及び管理用計算機4は、追加された配置規則の項目に対応するサーバ情報の収集等を開始する。

【0054】

次に、図1の計算機システムにおいて、ユーザから計算機2Aに対し分散ファイルAの作成要求があり、部分ファイルA1(141)及び部分ファイルA2(51)が作成される場合の処理手順を説明する。

【0055】

ユーザから入出力部を介して計算機2Aに分散ファイルAの作成要求があると、計算機2Aは、ユーザから指定されたファイルを分散ファイルAとし、分散ファイルAを予め規定されたサイズに分割して部分ファイル(ここでは、2つの部分ファイルA1(141)及び部分ファイルA2(51))を作成する。その後、計算機2Aは、管理用計算機4に、作成する部分ファイル数(本例では2)を通知する。

【0056】

管理用計算機4は、管理用計算機4が管理する計算機システム内で一意の部分ファイルIDを、計算機2Aから通知された部分ファイル数と同じ数だけ作成し、作成した部分ファイルIDを計算機2Aに通知する。管理用計算機4より部分ファイルIDを受け取った計算機2Aは、受取った部分ファイルIDを各部分ファイルに付与する。さらに、計算機2Aは、部分ファイルを配置すべき計算機1を決定する。以下、決定処理の手順を説明する。

【0057】

図5は、計算機2Aが、部分ファイルを配置すべき計算機1のリストを確定するまでの手順を示すフローチャートである。

【0058】

10

20

30

40

50

まず、計算機 2 A は、自身が有する配置規則のテーブル 2 3 の項目の中から、1 以上の優先度が設定されている項目全てを選択する。この際、ファイルごとに特定の配置規則のテーブル 2 3 が決まっている場合には、計算機 2 は、項目を選択する前に、その特定の配置規則のテーブル 2 3 を選択する（ステップ 5 0 0）。その後、計算機 2 A は、管理用計算機 4 に対し、選択された項目で規定される配置規則を満足する計算機 1 を問い合わせるためのコマンドを送信する（ステップ 5 0 1）。ただし、管理用計算機 4 に登録されていない項目の情報は管理用計算機 4 の返答では考慮されない。

【 0 0 5 9 】

問い合わせコマンドの結果を受信した計算機 2 A（ステップ 5 0 9）は、次に、配置規則の項目として負荷が選択されているかどうかを判定する。これは、本実施形態においては、負荷の情報は管理用計算機 4 では管理されていないため、管理用計算機 4 から送信された計算機 1 のリストについて、改めて負荷の条件を満たすかどうかを判断する必要があるからである。尚、本例では負荷の項目を例示するが、管理用計算機 4 で管理されない項目が他の項目である場合もある。このような項目がある場合、計算機 2 は、その項目についても、ここで説明した処理と同様の処理を行う（ステップ 5 0 2）。

10

【 0 0 6 0 】

負荷の項目が選択されている場合には、計算機 2 A は、ステップ 5 0 1 で得られた計算機 1 のリストに含まれる全ての計算機 1 に対して、現在の負荷情報を問い合わせるコマンドを発行する。そして、その応答に基づいて、計算機 2 A は、配置規則の負荷の項目に設定された条件を満足していない計算機 1 をリストから除外する（ステップ 5 0 3）。

20

【 0 0 6 1 】

次に、計算機 2 A は、ステップ 5 0 3 で修正された計算機 1 のリストに、作成した部分ファイル数（ここでは 2）以上の計算機 1 が含まれているかどうかを判定する（ステップ 5 0 4）。部分ファイル数以上の計算機 1 がリストに含まれていれば、計算機 2 は、適用した配置規則及び計算機 1 のリストを確定する（ステップ 5 0 7）。

【 0 0 6 2 】

また、ステップ 5 0 4 において、計算機 1 のリストに部分ファイル数以上の数の計算機 1 が含まれない場合、計算機 2 A は、ステップ 5 0 1 で配置規則として選択された項目のうち、優先度が最も低い項目を除外する（ステップ 5 0 5）。その後、計算機 2 A は、除外の結果、優先度が 1、即ち最も優先度の高い項目が配置規則のテーブル 2 3 に残っているかどうかを判定する（ステップ 5 0 6）。優先度が 1 の項目が残っていない（すべての項目を除外してしまった）場合、計算機 2 A は、ユーザに対して警告を表示するとともに、直前に得られた計算機 1 のリスト（優先度 1 の項目を除外する前の計算機 1 のリスト）を表示する（ステップ 5 0 8）。また、ステップ 5 0 6 において優先度 1 の項目が配置規則のテーブル 2 3 に残っている場合、計算機 2 A は、ステップ 5 0 1 以降の処理を繰り返し、残っている項目で構成される配置規則を満たす計算機 1 のリストを確定する。

30

【 0 0 6 3 】

その後、計算機 2 A は、確定した計算機 1 のリストから、ランダムに部分ファイルと同じ数の計算機 1 を選択する。そして、計算機 2 A は、選択された各計算機 1（この場合、計算機 1 A 及び 1 B）へそれぞれ部分ファイル A 1 及び A 2 の作成を指示する。この指示には、部分ファイルのデータ及び部分ファイル ID 等が含まれる。さらに、計算機 2 A は、管理用計算機 4 に対し、作成された分散ファイルに関する情報、具体的には、分散ファイル名、各部分ファイル ID、及び各部分ファイルの配置場所（計算機 1 を特定する情報）を通知する。

40

【 0 0 6 4 】

尚、上述したように、計算機 2 A は、図 5 のステップ 5 0 8 の処理に基づいて、ユーザに配置規則を満足する計算機 1 が足りない旨の警告を出す。このとき、作成された計算機 1 のリストに計算機 1 が 1 つでも存在する場合、計算機 2 A が、部分ファイルのうち複数または全てを同一の計算機 1 に配置してもよいかどうかをユーザに選択させても良い。そして、ユーザがよいと選択した場合は、計算機 2 は、複数の部分ファイルを同一の計算機 1

50

に配置する。

【0065】

また、計算機2Aは、図5のステップ508の処理に基づいて優先度の低い項目を除外した配置規則を満たす計算機1のリストが得られた場合、除外された項目を入出力部を介してユーザへ示し、警告を発するようにしても良い。

【0066】

次に、部分ファイル作成の指示を受けた計算機1の動作手順を説明する。

【0067】

部分ファイル作成の指示を受けた各計算機1（ここでは1A及び1B）は、指示と共に転送されたデータに基づいて自計算機1内のストレージ装置14内に部分ファイルを作成する。また、計算機1は、計算機2Aから受信した部分ファイルIDと部分ファイルのストレージ装置14内の記憶位置を対応付けるテーブルを作成する。

10

【0068】

図5のステップ501で計算機2Aから問い合わせコマンドを受けた管理用計算機4は、以下のように動作する。管理用計算機4は、計算機2Aから送信された配置規則の項目（図4）に基づいて、図3に示したサーバ情報管理テーブル44の中から図4の配置規則を満足する計算機1（本例の場合計算機1A及び1B）を抽出する。そして、抽出した計算機1の情報を、配置規則を満足する計算機1のリストとして計算機2Aに送信する。

【0069】

次に、図1の計算機システムにおいて、ユーザから計算機2Bに対する分散ファイルBの参照要求があった場合の処理手順及びこの参照要求を契機として、部分ファイルの移動（再配置）を行う場合の処理動作について説明する。尚、部分ファイルの移動（再配置）を行う契機としては、参照要求の他に、部分ファイルの内容を書き換える要求を計算機2が受け付けた場合や、ある一定周期、若しくはユーザが指定した時刻等が考えられる。また、本例では、計算機2Bには、図6のような配置規則の項目及び優先度が設定されていると仮定する。

20

【0070】

まず、入出力部を介してユーザから計算機2Bに分散ファイルBの参照要求があると、計算機2Bは、分散ファイルBを構成する各部分ファイルがどの計算機1に配置されているかを問い合わせるコマンドを管理用計算機4に送信する。

30

【0071】

問い合わせのコマンドを受信した管理用計算機4は、分散ファイル情報管理テーブル43を参照し、分散ファイルBの各部分ファイルの部分ファイルID及び部分ファイルを保存している計算機1の計算機名を計算機2Bに送信する。本例では、管理用計算機4は、部分ファイルB1（52）及び部分ファイルB2（61）が各々計算機1B及び計算機1Cに保存されていることを示す情報を送信する。

【0072】

管理用計算機4から分散ファイルBの部分ファイルについての情報を受信した計算機2Bは、計算機1Bへ部分ファイルB1（52）の読み出し要求を、計算機1Cへ部分ファイルB2（61）の読み出し要求を送信する。各計算機1は、計算機2Bの要求に応じて該当する部分ファイルをストレージ装置14から読み出して、計算機2Bへ送信する。この際、各計算機1は、各計算機1のサーバ情報も併せて計算機2Bに送信する。本例においては、図8に示されるようなサーバ情報が計算機2Bに送信されるものとする。尚、ファイルの再配置が他の契機で行われる場合には、計算機2は、その契機において、分散ファイルを構成する部分ファイルを有する計算機1に対して、各計算機1が有するサーバ情報を要求するコマンドを発行しても良い。

40

【0073】

計算機2Bは、受信した部分ファイルB1（52）及び部分ファイルB2（61）から分散ファイルBを復元する。また、各計算機1から受信したサーバ情報に基づいて、分散ファイルを構成する全ての部分ファイルの配置の評価を行う。

50

【 0 0 7 4 】

部分ファイルの配置の評価は、具体的には次のように行われる。即ち、計算機 2 B は、部分ファイルを保存している各計算機 1 から送信されたサーバ情報と計算機 2 B に登録されている配置規則とを比較して、分散ファイル B の部分ファイルを格納する計算機 1 が、分散ファイル B に対応する配置規則を満足しているかをチェックする。本例においては、計算機 2 B は、図 6 の配置規則及び図 8 に示されるサーバ情報を比較する。尚、分散ファイルごとに配置規則が決まっている場合には、計算機 2 B は、自計算機 2 B が有する配置規則のテーブル 2 3 の中から、分散ファイル B に対応する配置規則を決定した上で比較を行う。

【 0 0 7 5 】

評価の結果、計算機 2 B が有する配置規則を満足していない計算機 1 が存在する場合、計算機 2 B は、部分ファイルの移動処理を行う。本例においては、計算機 1 C が負荷の項目で配置規則を満たしていないので、計算機 1 C に格納された部分ファイル B 2 (6 1) が移動の対象となる。以下、部分ファイルの移動処理について説明する。

【 0 0 7 6 】

部分ファイルの移動を行う場合、計算機 2 B は、まず配置規則を満足していない計算機 1 の数及び配置規則を満足していない計算機 1 に格納されている部分ファイルの数を確認する。通常、各計算機 1 には 1 つの部分ファイルしか格納されていないので、計算機 1 の数と部分ファイルの数は一致する。しかし、1 つの分散ファイルの部分ファイルを重複して格納している計算機 1 が存在する場合には、配置規則を満足しない計算機 1 の数に比べて、それら計算機 1 に含まれる部分ファイルの数は多くなる。ここで確認された部分ファイルが、移動対象の部分ファイルとなる。

【 0 0 7 7 】

その後、計算機 2 B は、移動先の計算機 1 の候補として、配置規則を満足する計算機 1 のリストを管理用計算機 4 から取得する。そして、計算機 2 B は、計算機 1 のリストから移動先となる計算機 1 を決定し、移動対象となる部分ファイルを有する計算機 1 に対して、移動先となる計算機 1 へのデータの移動を指示する。

【 0 0 7 8 】

図 7 は、計算機 2 B が、移動先となる計算機 1 の候補となる計算機 1 のリストを確定する手順を示すフローチャートである。

【 0 0 7 9 】

まず、計算機 2 B は、配置規則のテーブル 2 3 で、1 以上の優先度が設定されている全ての項目を選択する (ステップ 7 0 0) 。その後、計算機 2 B は、管理用計算機 4 に対し、選択された項目で構成される配置規則を満足する計算機 1 の有無を問い合わせるコマンドを発行する (ステップ 7 0 1) 。

【 0 0 8 0 】

コマンドを受信した管理用計算機 4 は、受信した項目で構成される配置規則を満たす計算機 1 をサーバ情報管理テーブル 4 4 から検索し、その結果得られる計算機 1 のリストを計算機 2 B に送信する。本例の場合、管理用計算機 4 は、図 6 の配置規則を満足する計算機 1 として計算機 1 A を選択し、計算機 2 B に通知する (ステップ 7 1 0) 。

【 0 0 8 1 】

配置規則を満たす計算機 1 のリストを受信した計算機 2 B は、計算機 2 B が有する配置規則の項目で負荷の項目が選択されているかどうかを判定する (ステップ 7 0 2) 。負荷の項目が選択されている場合には、計算機 2 B は、受信した計算機 1 のリストに含まれる全ての計算機 1 に、現在の負荷情報を問い合わせるコマンドを発行する。コマンドの送信先である計算機 1 全てから応答が帰ってきたら、計算機 2 B は、その応答に基づいて、負荷の項目を満足していない計算機 1 を、取得された計算機 1 のリストから除外する (ステップ 7 0 3) 。

【 0 0 8 2 】

次に、計算機 2 B は、移動対象となっている部分ファイルが属する分散ファイル (この場

10

20

30

40

50

合分散ファイルB)に含まれる他の部分ファイルを格納している計算機1がリストに含まれているかを判断する。具体的には、分散ファイル参照の際に管理用計算機4から受信した、他の部分ファイルを格納している計算機1の計算機名と、リストに含まれる計算機1の計算機名とを逐一比較する。そして、計算機2Bは、該当する計算機1がリストに含まれていれば、その計算機1をリストから除外する(ステップ704)。

【0083】

次に、計算機2Bは、ステップ704までに得られた計算機1のリストに、移動すべき部分ファイルの数以上の計算機1が含まれているかどうかを判定する(ステップ705)。部分ファイル数以上の計算機1がリストに含まれていれば、計算機2Bは、適用した配置規則及び計算機1のリストを確定する(ステップ708)。本例においては、管理計算機4から通知された計算機1Aについてステップ702~705までの条件の確認が行われ、最終的に、計算機1Aが移動先の計算機1のリストとして確定される。

10

【0084】

一方、ステップ705において、計算機1のリストに移動すべき部分ファイルの数以上の計算機1が含まれていない場合、計算機2Bは、配置規則のテーブル23から、優先度が最も低い項目を除外する(ステップ706)。その後、計算機2Bは、除外の結果、優先度が1の項目が配置規則のテーブル23に残っているかどうかを判定する(ステップ707)。優先度が1の項目が残っていない(すべての項目を除外してしまった)場合、計算機2Bは、ユーザに警告を発生し、最後に得られた計算機1のリスト(最後に優先度1の項目を除外する前の計算機1のリスト)を、移動先となる計算機1のリストとする(ステップ709)。

20

【0085】

また、ステップ707において優先度1の項目が配置規則のテーブル23に残っている場合、計算機2Bは、ステップ701の処理に戻り、管理用計算機4に対し、残っている項目で構成された配置規則を満たす計算機1を問い合わせるコマンドを送信し、以下、同様のステップを実行して所望の計算機1のリストを確定する。

【0086】

その後、計算機2Bは、確定した移動先となる計算機1のリストから、移動すべき部分ファイル数分の計算機1を移動先の計算機1としてランダムに選択する。そして、計算機2Bは、配置規則を満たしていないと判断された計算機1に対し、移動すべき部分ファイルの部分ファイルID及び移動先となる計算機1の情報を送信する。

30

【0087】

計算機2Bから移動すべき部分ファイルの部分ファイルID及び移動先となる計算機1の情報を受信した計算機1は、受信した部分ファイルIDを用いて部分ファイルの格納位置を特定し、指定された部分ファイルを指定された計算機1に送信する。

【0088】

本実施形態においては、計算機2Bは、計算機1Aを移動先の計算機1として選択し、部分ファイルB2(61)の部分ファイルID及び計算機1Aを示す情報を計算機1Cに送信する。指示を受けた計算機1Cは、計算機2Bの指示に従い、計算機1Aへ部分ファイルB2(61)を転送する。

40

【0089】

ここで、移動先となる計算機1のリストに移動すべき部分ファイル数より多い計算機1が含まれる場合、計算機2Bは、移動すべき部分ファイルそれぞれに、重複しないように全ての計算機1を割り振り振っても良い。この場合、ある部分ファイルには、複数の計算機1が移動先として割り振られることになる。その後計算機2Bは、部分ファイルの移動指示を部分ファイルを格納している計算機1に送信する際に、複数の計算機1を移動先の候補として指定する情報を送信する。移動先として複数の計算機1の情報を受取った計算機1は、受信した複数の計算機1に関する情報を用いて、転送先となる計算機1を自由に選択して部分ファイルを転送する。

【0090】

50

またこの場合、移動すべき部分ファイルを保持している計算機 1 で、計算機 2 からの移動指示の実行をある一定の条件を満たすまで保留しておく構成とすることが出来る。さらに、部分ファイルの移動を保留している計算機 1 が、移動を保留している部分ファイルへの移動指示を重ねて受け取った場合、本計算機 1 は、各々の移動指示で指定される移動先となる計算機 1 のうち、全ての移動指示で移動先として指定されている計算機 1 を移動先として選択する。

【 0 0 9 1 】

尚、計算機 2 B は、図 7 のステップ 7 0 9 の処理に基づき、ユーザに対し、部分ファイルが配置規則を満たさない計算機 1 に配置されているが、配置規則を満たす移動先の計算機 1 が見つからない旨の警告を入出力部に表示しても良い。また、移動先となる計算機 1 のリストに部分ファイル数未満の計算機 1 しか含まれていない場合、計算機 2 B は、ユーザに対し、部分ファイルが配置規則を満たさない計算機 1 に配置されているが、配置規則を満たす移動先の計算機 1 の数が足りない旨の警告を入出力部に表示しても良い。

10

【 0 0 9 2 】

更にこの場合、計算機 2 B は、配置規則を満たさない部分ファイルのうちの一部または全てを同一の計算機 1 に移動してもよいかどうかをユーザに選択させるようにしても良い。ここで、ユーザが部分ファイルの重複を許した場合、計算機 2 B は、移動先の計算機 1 のリストから計算機 1 を無作為に抽出し、抽出された計算機 1 に部分ファイルを移動するよう、移動対象である部分ファイルを有する複数の計算機 1 に指示する。

【 0 0 9 3 】

又、上述したように、計算機 1 は、一定時間、または一定数の移動指示を計算機 2 B から受信するまで、部分ファイルの移動処理の実行を保留しても良い。これにより、短時間の間に複数の計算機 2 から同一の部分ファイルの参照があった場合、計算機 2 毎の配置規則の違いによって、参照ごとに毎回部分ファイルが移動してしまう現象を避け、頻繁に部分ファイルが移動することによる計算機システムへ与える負荷を低減することができる。

20

【 0 0 9 4 】

尚、部分ファイルの移動処理を保留している計算機 1 は、その保留された移動処理を以下のパターンで実行する。

【 0 0 9 5 】

まず、保留している移動処理が 1 つしかなく、かつその移動処理の対象となる移動先の計算機 1 が一つしか指定されていない場合、移動処理を保留している計算機 1 は、移動先として指定された計算機 1 に部分ファイルを転送する。

30

【 0 0 9 6 】

次に、保留している移動処理が一つしかないが、移動処理の対象となる移動先の計算機 1 が複数指定されている場合、移動処理を保留している計算機 1 は、複数指定された計算機 1 からランダムに計算機 1 を選択し、選択した計算機 1 に部分ファイルを転送する。

【 0 0 9 7 】

また、保留している移動処理が複数あり、かつその各々の移動処理において移動先の計算機 1 が一つ又は複数指定されている場合、移動処理を保留している計算機 1 は、保留している各々の処理で指定される移動先の計算機 1 の中で、共通して存在する計算機 1 に部分ファイルを転送する。共通に存在する計算機 1 がない場合、移動処理を保留している計算機 1 は移動処理を行わない。

40

【 0 0 9 8 】

計算機 2 B は、部分ファイル移動の完了を移動先となった計算機 1 から通知された後、管理用計算機 4 に対し、部分ファイル ID 及び部分ファイルの移動先である計算機 1 についての情報（ネットワークアドレス等）を通知する。

【 0 0 9 9 】

本例においては、計算機 2 B は、部分ファイル B 2 (6 1) の部分ファイル ID と移動先となる計算機 1 A の情報を、管理計算機 4 に通知する。

【 0 1 0 0 】

50

尚、上述した実施形態における部分ファイル転送の方法とは別の方法として、以下が考えられる。

【0101】

すなわち、移動処理における移動先となる計算機1が複数指定されている場合、移動処理を実行する計算機1が、その全ての移動先となる計算機1に部分ファイルを転送する形態である。尚、この場合、移動処理を実行する計算機1のストレージ装置に格納されている部分ファイルは削除される。

【0102】

また、移動先となる計算機1全てに部分ファイルを転送するのではなく、移動処理を行う計算機1が、移動先の候補となる複数の計算機1から複数の計算機1を選択し、その選択された計算機1に対して部分ファイルを転送することも考えられる。

10

【0103】

更に、保留している移動処理が複数あり、かつその各々の移動処理において移動先の計算機1が一つ又は複数指定されている場合で共通する計算機1がない場合、移動処理を保留している計算機1は、各々の移動処理の移動先として指定されている計算機1全て（又は一部）に部分ファイルを転送する。

【0104】

尚、本方法の場合、移動処理を行う計算機1は、複数の計算機1に同じ部分ファイルを転送したことを管理用計算機4に通知しなければならない。

【0105】

図12は、本方法における管理用計算機4で保持されている分散ファイル情報管理テーブル43の例を示す図である。本図においては、図2で説明した分散ファイル情報管理テーブル43において、部分ファイルの保存先となる計算機1の情報を登録するエントリが複数用意される。

20

【0106】

また、本方法において、部分ファイルを複数の計算機1に配置（以下「部分ファイルのコピー」）してよいかどうかを、以下のように設定することが考えられる。

【0107】

一つは、ユーザが指定する場合である。この場合、分散ファイル情報管理テーブル43には、図13に示すようなエントリが、ファイルの属性情報として設定される。そして、ユーザは、分散ファイル作成時に、計算機2を介して図13に示すエントリに値を設定することで、ある分散ファイルの部分ファイルのコピーの可否を設定する。計算機1は、この設定された情報を参照して部分ファイルのコピーを行う。

30

【0108】

もう一つは、あらかじめファイルの種類によって部分ファイルのコピーの可否を設定しておき、分散ファイルの再配置時に、部分ファイルの転送を行う計算機1がその情報に基づいて自動的に部分ファイルのコピーを行う場合である。この場合、管理用計算機4又は計算機1は、図14に示すようなテーブルをあらかじめ有する。

【0109】

このようにして設定された情報に基づいて、計算機1は、部分ファイルのコピーを行う。尚、設定された情報が管理用計算機4にしか存在しない場合、移動処理を行う計算機1は、管理用計算機4に部分ファイルのコピー処理の可否を問い合わせ、その結果に応じて部分ファイルのコピーを行う。

40

【0110】

更に、本方法においては、同じ部分ファイルが複数の計算機1に存在するので、計算機2が分散ファイルを参照する場合、部分ファイルをどの計算機1から参照するかを決定する必要がある。決定方法としては以下の例が考えられる。

【0111】

第一の例は、管理用計算機4で参照先を決定する場合である。この場合、計算機2から部分ファイルの配置場所の問い合わせを受信したら、管理用計算機4は、部分ファイル情報

50

管理テーブル 4 3 に登録された複数の配置先からランダムに一つの計算機 1 を選択し、その情報を計算機 2 に通知する。

【 0 1 1 2 】

第二の例は、分散ファイルを参照する計算機 2 が決定する場合である。この場合、部分ファイルの配置場所の問い合わせを受けた管理用計算機 4 は、同じ部分ファイルを格納する複数の計算機 1 に関する情報を全て計算機 2 へ送信する。複数の計算機 1 についての情報を受取った計算機 2 は、その複数の計算機 1 から任意に計算機 1 を選択して部分ファイルの送信要求を発行する。計算機 2 における計算機 1 の選択の仕方には、以下のような例が考えられる。

【 0 1 1 3 】

一つ目は、計算機 2 が、管理用計算機 4 から受信した情報に含まれる計算機 1 全てに部分ファイルの送信要求を出し、最初に応答が返ってきたものを採用する。この場合、計算機 2 は、その他の計算機 1 に対しては部分ファイル送信の中止命令を送信する。

【 0 1 1 4 】

二つ目は、計算機 2 が、管理用計算機 4 から受信した情報に含まれる全計算機 1 から並列して部分ファイルを取得する。具体的には、計算機 2 は、各々の計算機 1 に対して、それぞれの部分ファイルのオフセットを指定した取得要求を出す（「部分ファイルの最初の何バイト目から何バイトを送信」という要求を各計算機 1 に分割して出す）。

【 0 1 1 5 】

本方法によれば、部分ファイルをより柔軟に計算機システム内の計算機に再配置することができる。

【 0 1 1 6 】

上述した実施形態によれば、計算機システムは、部分ファイル自体が他の部分ファイルに関する情報を有していなくても、配置規則を満足する計算機 1 の情報を管理用計算機 4 から取得するだけで、部分ファイルを重複せずに配置することができる。これにより、計算機システムのセキュリティを確保しつつ、ファイルの分散を行うことが出来る。

【 0 1 1 7 】

【 発明の効果 】

本発明によれば、分散ファイルの全ての部分ファイルが特定のサーバへ集まってしまうことを回避し、かつ計算機システムのセキュリティを確保しながら分散ファイルを配置することができる。

【 0 1 1 8 】

また、複数かつ多様な配置規則を規定し、それらに従って部分ファイルを柔軟に配置することができる。

【 図面の簡単な説明 】

【 図 1 】 本発明の実施形態による計算機システムを示すブロック図である。

【 図 2 】 分散ファイル情報管理テーブルの一例を示す図である。

【 図 3 】 サーバ情報管理テーブルの一例を示す図である。

【 図 4 】 配置規則及び優先度設定情報の一例を示す図である。

【 図 5 】 部分ファイルの配置先計算機を選定する処理動作の一例を示すフローチャートである。

【 図 6 】 配置規則及び優先度設定情報の一例である。

【 図 7 】 部分ファイルの移動先計算機を選定する処理動作の一例を示すフローチャートである。

【 図 8 】 配置規則及び優先度設定情報の一例である。

【 図 9 】 分散ファイル種類適用規則対応表の例である。

【 図 1 0 】 分散ファイル情報管理テーブルの一例を示す図である。

【 図 1 1 】 計算機の構成例を示す図である。

【 図 1 2 】 部分ファイル情報管理テーブルの例を示す図である。

【 図 1 3 】 部分ファイル情報管理テーブルのファイル属性情報の例を示す図である。

10

20

30

40

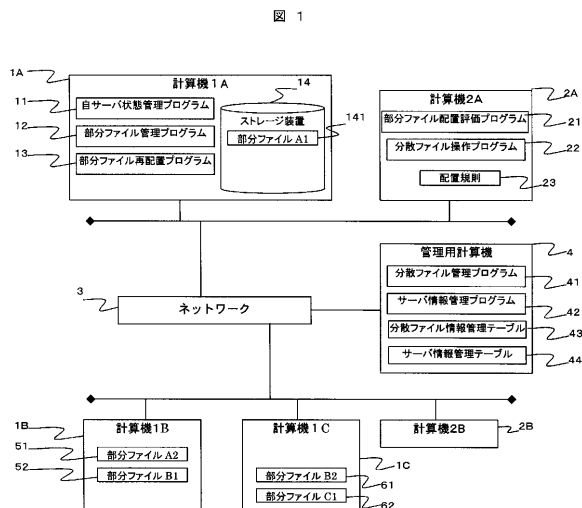
50

【図14】部分ファイルのコピーの可否を示す情報が登録されたテーブルの例を示す図である。

【符号の説明】

1、2... 計算機、3... ネットワーク、4... 管理用計算機、14... ストレージ装置。

【図1】



【図3】

図 3

計算機名	性能 [Mbyte/sec]	セキュリティ レベル	信頼性レベル	使用料金 [円/Mbyte]
計算機 1 A	40	5	3	6
計算機 1 B	50	4	2	4
計算機 1 C	100	3	1	8

【図4】

図 4

項目名	優先度	値
性能	3	40
セキュリティ	1	4
信頼性	2	2
使用料金	0	-
負荷	0	-

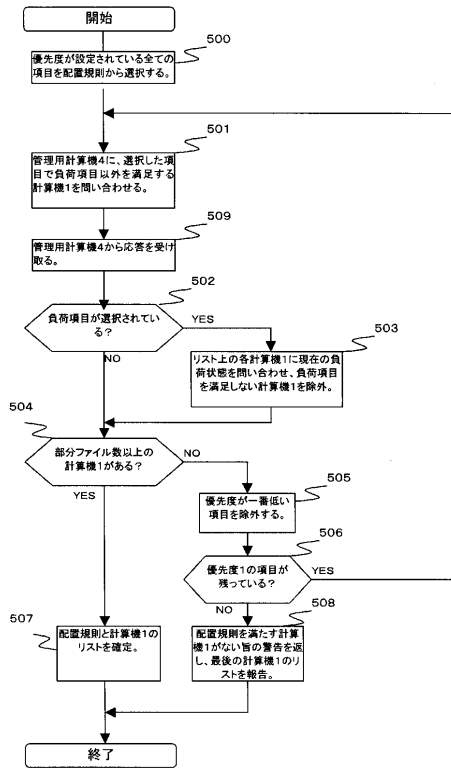
【図2】

図 2

分散ファイル名	分散ファイル属性情報 (ファイルの種類)	部分ファイルID	保存計算機名
分散ファイルA	データベースファイル	A 1	計算機 1 A
		A 2	計算機 1 B
分散ファイルB	テキストファイル	B 1	計算機 1 B
		B 2	計算機 1 C
分散ファイルC	一時ファイル	C 1	計算機 1 C

【図5】

図5



【図6】

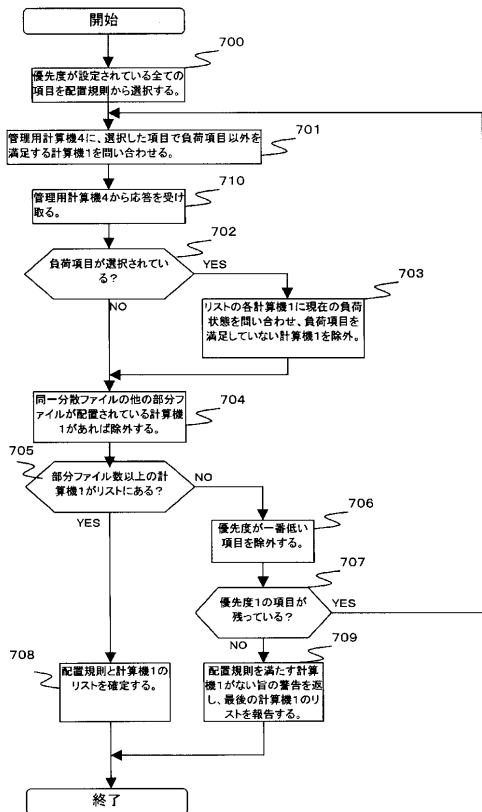
図6

23

項目名	優先度	値
性能	0	--
セキュリティ	0	--
信頼性	0	--
使用料金	1	8
負荷	2	70

【図7】

図7



【図8】

図8

計算機名	性能 [Mbyte/sec]	セキュリティ レベル	信頼性レベル	使用料金 [円/Mbyte]	負荷 [%]
計算機1B	50	4	2	6	50
計算機1C	100	3	1	8	80

【図9】

図9

分散ファイル種類	性能		セキュリティ		信頼性		使用料金		負荷	
	優先度	値	優先度	値	優先度	値	優先度	値	優先度	値
データベース	1	10	1	4	2	4	0	--	3	--
テキスト	0	--	2	--	0	--	1	3	0	--
テンポラリー	0	--	0	--	0	--	1	1	0	--

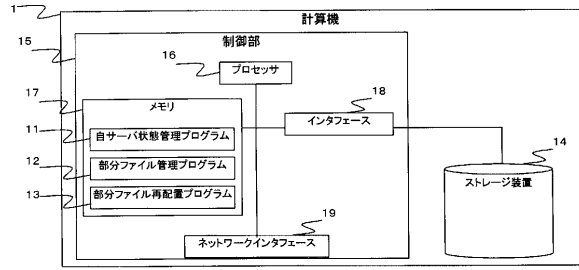
【図 10】

図 10

分散ファイル名	性能		セキュリティ		信頼性		使用料金		負荷		部分ファイル	
	優先度	値	優先度	値	優先度	値	優先度	値	優先度	値	部分ファイル ID	保存位置
分散ファイルA	0	-	1	4	2	4	0	-	3	-	A1	計算機1A
分散ファイルB	3	5	2	-	0	-	1	3	0	-	B1	計算機1B
分散ファイルC	0	-	0	-	0	-	1	1	0	-	C1	計算機1C

【図 11】

図 11



【図 12】

図 12

分散ファイル名	分散ファイル属性情報 (ファイルの種類)	部分ファイル ID	保存計算機名
分散ファイルA	データベースファイル	A1	計算機1A
		A2	計算機1B
		B1	計算機1C
分散ファイルB	テキストファイル	B2	計算機1A
			計算機1C
分散ファイルC	一時ファイル	C1	計算機1C
			計算機1B

【図 13】

図 13

分散ファイル種類	コピー許可?
データベース	NO
テキスト	YES
テンポラリ	YES

【図 14】

図 14

分散ファイル名	コピー許可?	部分ファイル ID	部分ファイル保存位置
分散ファイルA	NO	A1	計算機1A
		A2	計算機1B
分散ファイルB	YES	B1	計算機1B
		B2	計算機1C
分散ファイルC	NO	C1	計算機1C

フロントページの続き

(72)発明者 園田 浩二

神奈川県川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

審査官 田川 泰宏

(56)参考文献 米国特許出願公開第2003/0046270(US, A1)

特開平08-147113(JP, A)

特開平09-265359(JP, A)

特開2002-222061(JP, A)

(58)調査した分野(Int.Cl., DB名)

G06F 12/00