

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2011-170667
(P2011-170667A)

(43) 公開日 平成23年9月1日(2011.9.1)

(51) Int.Cl.		F I		テーマコード (参考)
G06F 13/00	(2006.01)	G06F 13/00	520R	5B082
G06F 12/00	(2006.01)	G06F 13/00	520B	5B084
		G06F 12/00	545M	

審査請求 未請求 請求項の数 5 O L (全 14 頁)

(21) 出願番号 特願2010-34574 (P2010-34574)
(22) 出願日 平成22年2月19日 (2010.2.19)

(71) 出願人 000004237
日本電気株式会社
東京都港区芝五丁目7番1号
(74) 代理人 100079108
弁理士 稲葉 良幸
(74) 代理人 100109346
弁理士 大貫 敏史
(74) 代理人 100117189
弁理士 江口 昭彦
(74) 代理人 100134120
弁理士 内藤 和彦
(74) 代理人 100109586
弁理士 土屋 徹雄

最終頁に続く

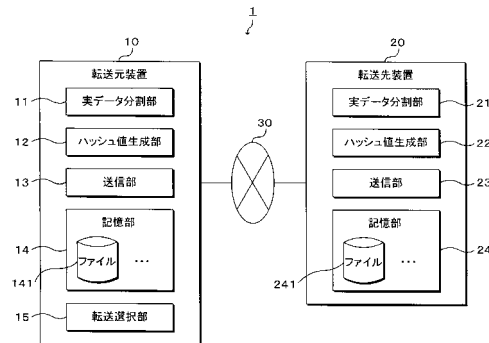
(54) 【発明の名称】 ファイル同期システム、ファイル同期方法およびファイル同期プログラム

(57) 【要約】

【課題】ファイル同期の効率を向上させる。

【解決手段】転送元装置10は、転送元ファイル141のデータを複数のブロックに分割する実データ分割部11と、各ブロックのデータからハッシュ値を生成するハッシュ値生成部12と、ハッシュ値を転送先装置20に送信する送信部13とを有し、転送先装置20は、転送先ファイル241のデータを、実データ分割部21によって分割された各ブロックと同サイズのブロックに分割する実データ分割部21と、各ブロックのデータからハッシュ値を生成するハッシュ値生成部22と、ハッシュ値生成部12および22によって生成されたハッシュ値をブロック番号ごとに比較し、不一致となったハッシュ値のブロック番号を転送元装置10に送信する送信部23とを有し、送信部13は、転送先装置20から受信したブロック番号に対応する実データを転送先装置20に送信する。

【選択図】図1



【特許請求の範囲】**【請求項 1】**

ファイル同期の対象となるファイルをそれぞれ有する転送元装置と転送先装置とを備えるファイル同期システムであって、

前記転送元装置は、

ファイル同期の対象となる転送元の前記ファイルのデータを複数のブロックに分割する第 1 の分割部と、

前記第 1 の分割部によって分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 1 の生成部と、

前記第 1 の生成部によって生成されたハッシュ値を前記転送先装置に送信する第 1 の送信部と、を有し、

前記転送先装置は、

ファイル同期の対象となる転送先の前記ファイルのデータを、前記第 1 の分割部によって分割された各ブロックと同サイズのブロックに分割する第 2 の分割部と、

前記第 2 の分割部によって分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 2 の生成部と、

前記第 1 の生成部によって生成されたハッシュ値と前記第 2 の生成部によって生成されたハッシュ値とを、各ハッシュ値の生成元データを含むブロックを識別するブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応する前記ブロック番号を前記転送元装置に送信する第 2 の送信部と、を有し、

前記転送元装置の前記第 1 の送信部は、前記転送先装置から受信した前記ブロック番号に対応する前記ブロックに含まれるデータを前記転送先装置に送信する、

ことを特徴とするファイル同期システム。

【請求項 2】

前記ファイルは、固定長のレコードを格納するファイルであることを特徴とする請求項 1 記載のファイル同期システム。

【請求項 3】

前記ファイルは、追記型のファイルであることを特徴とする請求項 1 または 2 記載のファイル同期システム。

【請求項 4】

ファイル同期の対象となるファイルをそれぞれ有する転送元装置と転送先装置とを備えるファイル同期システムで実行するファイル同期方法であって、

前記転送元装置が、

ファイル同期の対象となる転送元の前記ファイルのデータを複数のブロックに分割する第 1 の分割ステップと、

前記第 1 の分割ステップにおいて分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 1 の生成ステップと、

前記第 1 の生成ステップにおいて生成されたハッシュ値を前記転送先装置に送信する第 1 の送信ステップと、を実行し、

前記転送先装置が、

ファイル同期の対象となる転送先の前記ファイルのデータを、前記第 1 の分割ステップにおいて分割された各ブロックと同サイズのブロックに分割する第 2 の分割ステップと、

前記第 2 の分割ステップにおいて分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 2 の生成ステップと、

前記第 1 の生成ステップにおいて生成されたハッシュ値と前記第 2 の生成ステップにおいて生成されたハッシュ値とを、各ハッシュ値の生成元データを含むブロックを識別するブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応する前記ブロック番号を前記転送元装置に送信する第 2 の送信ステップと、を実行し、

前記転送元装置の前記第 1 の送信ステップが、前記転送先装置から受信した前記ブロック番号に対応する前記ブロックに含まれるデータを前記転送先装置に送信する、

10

20

30

40

50

ことを特徴とするファイル同期方法。

【請求項 5】

請求項 4 に記載の各ステップをコンピュータに実行させるためのファイル同期プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ファイル同期システム、ファイル同期方法およびファイル同期プログラムに関する。

【背景技術】

【0002】

同じファイルを複数のシステムで取り扱う場合に、ファイル同期を行ってそれぞれのファイルの内容が同じ内容になるように管理することがある。ファイル同期を行う方法としては、例えば、対象ファイル全体を転送して同期をとる方法や、対象ファイルの更新前後における差分情報のみを転送して同期をとる方法等がある。差分情報のみを転送する方法を採用することで、同期処理のターンアラウンドタイムを短縮することが可能となる。

【0003】

下記特許文献 1 には、差分情報のみを転送する方法でファイル同期を行うシステムが開示されている。このシステムでは、ファイルの転送元装置が更新前後のファイルをそれぞれ管理し、更新前後のファイルから抽出した差分データを転送先装置に送信している。そして、転送先装置が差分データと更新前のファイルとを用いて更新後のファイルを生成することで、ファイル同期を行っている。

【先行技術文献】

【特許文献】

【0004】

【特許文献 1】特開 2003 - 337723 号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

上記特許文献 1 に記載のシステムでは、更新後のファイルに加え、更新前のファイルも管理している。したがって、記憶領域の増大やファイルメンテナンス作業の増加等を招来し、コストが増大する要因になっている。

【0006】

本発明は、上述した課題を解決するためになされたものであり、ファイル同期の効率を向上させることができるファイル同期システム、ファイル同期方法およびファイル同期プログラムを提供することを目的とする。

【課題を解決するための手段】

【0007】

本発明のファイル同期システムは、ファイル同期の対象となるファイルをそれぞれ有する転送元装置と転送先装置とを備えるファイル同期システムであって、前記転送元装置は、ファイル同期の対象となる転送元の前記ファイルのデータを複数のブロックに分割する第 1 の分割部と、前記第 1 の分割部によって分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 1 の生成部と、前記第 1 の生成部によって生成されたハッシュ値を前記転送先装置に送信する第 1 の送信部と、を有し、前記転送先装置は、ファイル同期の対象となる転送先の前記ファイルのデータを、前記第 1 の分割部によって分割された各ブロックと同サイズのブロックに分割する第 2 の分割部と、前記第 2 の分割部によって分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 2 の生成部と、前記第 1 の生成部によって生成されたハッシュ値と前記第 2 の生成部によって生成されたハッシュ値とを、各ハッシュ値の生成元データを含むブロックを識別するブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応する前記ブロック番号

10

20

30

40

50

を前記転送元装置に送信する第2の送信部と、を有し、前記転送元装置の前記第1の送信部は、前記転送先装置から受信した前記ブロック番号に対応する前記ブロックに含まれるデータを前記転送先装置に送信する、を備える。

【0008】

本発明のファイル同期方法は、ファイル同期の対象となるファイルをそれぞれ有する転送元装置と転送先装置とを備えるファイル同期システムで実行するファイル同期方法であって、前記転送元装置が、ファイル同期の対象となる転送元の前記ファイルのデータを複数のブロックに分割する第1の分割ステップと、前記第1の分割ステップにおいて分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第1の生成ステップと、前記第1の生成ステップにおいて生成されたハッシュ値を前記転送先装置に送信する第1の送信ステップと、を実行し、前記転送先装置が、ファイル同期の対象となる転送先の前記ファイルのデータを、前記第1の分割ステップにおいて分割された各ブロックと同サイズのブロックに分割する第2の分割ステップと、前記第2の分割ステップにおいて分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第2の生成ステップと、前記第1の生成ステップにおいて生成されたハッシュ値と前記第2の生成ステップにおいて生成されたハッシュ値とを、各ハッシュ値の生成元データを含むブロックを識別するブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応する前記ブロック番号を前記転送元装置に送信する第2の送信ステップと、を実行し、前記転送元装置の前記第1の送信ステップが、前記転送先装置から受信した前記ブロック番号に対応する前記ブロックに含まれるデータを前記転送先装置に送信する。

10

20

【0009】

本発明のファイル同期プログラムは、上記ファイル同期方法に含まれる各ステップをコンピュータに実行させる。

【発明の効果】

【0010】

本発明によれば、ファイル同期の効率を向上させることができる。

【図面の簡単な説明】

【0011】

【図1】実施形態におけるファイル同期システムの構成を例示するブロック図である。

【図2】転送元装置におけるハッシュ値生成レートテーブルのデータ構成を例示する図である。

30

【図3】転送元装置における転送レートテーブルのデータ構成を例示する図である。

【図4】転送元装置におけるヒット率テーブルのデータ構成を例示する図である。

【図5】転送先装置におけるハッシュ値生成レートテーブルのデータ構成を例示する図である。

【図6】ファイル同期システムにおけるファイル転送の処理手順を説明するためのフローチャートである。

【図7】図6の処理手順に含まれる差分転送の処理手順を説明するためのフローチャートである。

【発明を実施するための形態】

40

【0012】

以下、添付図面を参照して、本発明に係るファイル同期システム、ファイル同期方法およびファイル同期プログラムの好適な実施形態について説明する。

【0013】

まず、実施形態におけるファイル同期システムの構成について説明する。図1は、実施形態におけるファイル同期システムの構成を例示するブロック図である。図1に示すように、ファイル同期システム1は、転送元装置10と、転送先装置20とを有する。転送元装置10および転送先装置20は、ネットワーク30を介して接続される。

【0014】

転送元装置10および転送先装置20は、ファイル同期の対象となるファイル141お

50

よび 241 をそれぞれ備えている。本実施形態では、説明の便宜のために、転送元装置 10 のファイル 141 が更新された場合に、この更新された情報を、転送先装置 20 のファイル 241 に反映させる場合について説明する。なお、転送先装置 20 のファイル 241 が更新され、この更新された情報を転送元装置 10 のファイル 141 に反映させる場合には、以下に説明する転送元装置 10 の各機能と転送先装置 20 の各機能とを入れ替えることで実現することができる。転送元装置 10 および転送先装置 20 は、それぞれの各機能を併せて備えることができる。

【0015】

ここで、転送元装置 10 および転送先装置 20 は、物理的には、例えば、CPU (Central Processing Unit) と、記憶装置と、入出力インターフェースとを含んで構成される。記憶装置には、例えば、CPU で処理されるプログラムやデータを記憶する ROM (Read Only Memory) や HDD (Hard Disk Drive)、主として制御処理のための各種作業領域として使用される RAM (Random Access Memory) 等が含まれる。転送元装置 10 および転送先装置 20 の HDD には、ファイル同期の対象となるファイル 141 および 241 がそれぞれ格納される。CPU が、ROM に記憶されたプログラムを実行し、RAM に展開されるデータを処理することで、後述する転送元装置 10 および転送先装置 20 の各機能が実現する。なお、ファイル同期の対象となるファイルは、一組であることには限定されず、複数組設定することができる。

【0016】

転送元装置 10 は、実データ分割部 11 と、ハッシュ値生成部 12 と、送信部 13 と、記憶部 14 と、転送選択部 15 と、を有する。

【0017】

実データ分割部 11 は、ファイル同期の対象となる転送元のファイル 141 の実データを複数のブロックに分割する。分割するブロック数は、分割ブロック数 N として予め設定し、記憶部 14 に記憶させる。N 個に分割した各ブロックには、ブロックを識別するブロック番号を付与する。実データ分割部 11 は、各ブロックのサイズを、分割ブロックサイズ D として記憶部 14 に一時的に記憶させる。

【0018】

ハッシュ値生成部 12 は、実データ分割部 11 によって N 個に分割された各ブロックに含まれる実データからそれぞれハッシュ値を生成する。ハッシュ値は、所定のハッシュ関数に各実データを入力することで生成する。ハッシュ値生成部 12 は、ハッシュ値の生成に要した時間を積算してハッシュ値生成時間を算出する。ハッシュ値生成部 12 は、全てのブロックの実データからハッシュ値を生成した後に、ファイルサイズ A およびハッシュ値生成時間を用いて、図 2 に示す転送元側のハッシュ値生成レートテーブルに新規レコードを追加する。

【0019】

図 2 に示す転送元側のハッシュ値生成レートテーブルは、データ項目として、例えば、転送処理名項目、ファイルサイズ項目、ハッシュ値生成時間項目および重み付け項目を有する。転送処理名項目は、転送処理を識別するための情報を格納する。ファイルサイズ項目は、転送元のファイルのファイルサイズを格納する。積算時間項目は、ハッシュ値生成部 12 でハッシュ値を生成する際に要した時間の積算値を格納する。重み付け項目は、各レコードを用いてハッシュ値生成レートを算出する際に、各レコードに付与する重み付けを格納する。重み付けは、新しく生成されたレコード程、大きくなるように設定する。これにより、直近の転送状況を反映した生成レートを算出することが可能となる。

【0020】

送信部 13 は、ハッシュ値取得要求を転送先装置 20 に送信する。ハッシュ値取得要求には、例えば、ハッシュ値生成部 12 によって生成されたハッシュ値、転送元ファイル 141 のファイル名、ファイルサイズ A、分割ブロック数 N、分割ブロックサイズ D およびハッシュ値サイズ S が含まれる。ファイル名、ファイルサイズ A およびハッシュ値サイズ S は、予め設定し、記憶部 14 に記憶させる。

10

20

30

40

50

【 0 0 2 1 】

送信部 1 3 は、転送先装置 2 0 から受信したブロック番号に対応するブロックに含まれる実データを転送先装置 2 0 に送信する。送信部 1 3 は、送信した実データのサイズを積算して転送データサイズを算出するとともに、実データの送信に要した時間を積算して転送時間を算出する。送信部 1 3 は、転送対象となる全ての実データを送信した後に、算出した転送データサイズおよび転送時間を用いて、図 3 に示す転送レートテーブルに新規レコードを追加する。

【 0 0 2 2 】

図 3 に示す転送レートテーブルは、データ項目として、例えば、転送処理名項目、転送データサイズ項目、転送時間項目および重み付け項目を有する。転送処理名項目は、転送処理を識別するための情報を格納する。転送データサイズ項目は、送信部 1 3 で送信した実データのサイズの積算値を格納する。転送時間項目は、送信部 1 3 で実データを送信する際に要した時間の積算値を格納する。重み付け項目は、各レコードを用いて転送レートを算出する際に、各レコードに付与する重み付けを格納する。重み付けは、新しく生成されたレコード程、大きくなるように設定する。これにより、直近の転送状況を反映した生成レートを算出することが可能となる。

【 0 0 2 3 】

送信部 1 3 は、送信対象の全ての実データを送信した後に、ファイルサイズ A および転送データサイズを用いて、図 4 に示すヒット率レートテーブルに新規レコードを追加する。

【 0 0 2 4 】

図 4 に示すヒット率テーブルは、データ項目として、例えば、転送ファイル名項目、ファイルサイズ項目およびヒットサイズ項目を有する。転送ファイル名項目は、転送元のファイルの名称を格納する。ファイルサイズ項目は、転送元のファイルのファイルサイズを格納する。ヒットサイズ項目は、転送元ファイルのファイルサイズから、送信部 1 3 で送信した転送データサイズを減算した値、つまり、転送元ファイルの実データのうち未送信の実データサイズを格納する。

【 0 0 2 5 】

転送選択部 1 5 については、後述する。

【 0 0 2 6 】

転送先装置 2 0 は、実データ分割部 2 1 と、ハッシュ値生成部 2 2 と、送信部 2 3 と、記憶部 2 4 と、を有する。

【 0 0 2 7 】

実データ分割部 2 1 は、ファイル同期の対象となる転送先のファイル 2 4 1 の実データを、実データ分割部 1 1 によって分割された各ブロックと同サイズのブロックに分割する。実データ分割部 1 1 によって分割された各ブロックのサイズは、転送元装置 1 0 から受信した分割ブロックサイズ D を参照することで得られる。

【 0 0 2 8 】

ハッシュ値生成部 2 2 は、実データ分割部 2 1 によって分割された各ブロックに含まれる実データからそれぞれハッシュ値を生成する。ハッシュ値は、所定のハッシュ関数に各実データを入力することで生成する。ハッシュ値生成部 1 2 は、ハッシュ値の生成に要した時間を積算してハッシュ値生成時間を算出する。ハッシュ値生成部 2 2 は、全てのブロックの実データからハッシュ値を生成した後に、ファイルサイズおよびハッシュ値生成時間を用いて、図 5 に示す転送先側のハッシュ値生成レートテーブルに新規レコードを追加する。

【 0 0 2 9 】

図 5 に示す転送先側のハッシュ値生成レートテーブルは、データ項目として、例えば、転送処理名項目、ファイルサイズ項目、ハッシュ値生成時間項目および重み付け項目を有する。転送処理名項目は、転送処理を識別するための情報を格納する。ファイルサイズ項目は、転送元のファイルのファイルサイズおよび転送先のファイルのファイルサイズのう

10

20

30

40

50

ち、小さい方のファイルサイズを格納する。ハッシュ値生成時間項目は、ハッシュ値生成部 22 でハッシュ値を生成する際に要した時間の積算値を格納する。重み付け項目は、各レコードを用いてハッシュ値生成レートを算出する際に、各レコードに付与する重み付けを格納する。重み付けは、新しく生成されたレコード程、大きくなるように設定する。これにより、直近の転送状況を反映した生成レートを算出することが可能となる。

【0030】

送信部 23 は、ハッシュ値生成部 12 によって生成されたハッシュ値とハッシュ値生成部 22 によって生成されたハッシュ値とを、同じブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応するブロック番号を含む応答メッセージを転送元装置 10 に送信する。

10

【0031】

転送元装置 10 の転送選択部 15 は、ファイル同期時の転送方法として、転送元のファイル 141 を、通常転送で送信するのか、差分転送で送信するのかを選択する。通常転送は、転送元ファイルの全ての実データを転送先装置に送信する転送処理である。差分転送は、データ更新されたブロックに含まれる実データを転送先装置に送信する転送処理である。

【0032】

転送選択部 15 は、転送元のファイル 141 のファイルサイズ A が、分割ブロック数 N とハッシュ値サイズ S とを乗算した値よりも大きい場合に、差分転送を選択する。転送選択部 15 は、転送元のファイル 141 のファイルサイズ A が、分割ブロック数 N とハッシュ値サイズ S とを乗算した値以下である場合に、通常転送を選択する。総実データサイズが総ハッシュ値サイズよりも大きい場合に、差分転送を選択することで、転送効率を向上させることができる。一方、総実データサイズが総ハッシュ値サイズ以下である場合に、通常転送を選択することで、差分転送により転送効率が低下する事態を防止することができる。

20

【0033】

転送選択部 15 は、転送元のファイル 141 を通常転送したときの推定総処理時間（以下、「通常転送処理時間」という。）が、転送元のファイル 141 を差分転送したときの推定総処理時間（以下、「差分転送処理時間」という。）よりも長い場合に、差分転送を選択し、通常転送処理時間が差分転送処理時間以下である場合に、通常転送を選択する。以下に、通常転送処理時間および差分転送処理時間を算出する方法について、転送元のファイル 141 を転送元装置 10 から転送先装置 20 に転送する場合を例にして説明する。

30

【0034】

通常転送処理時間は、例えば、下記式 1 を用いて算出することができる。

【0035】

通常転送処理時間 T_1 = ファイルサイズ A / 転送レート R_3 ... 式 1

【0036】

上記式 1 のファイルサイズ A は、転送元のファイル 141 のファイルサイズである。上記式 1 の転送レート R_3 は、例えば、図 3 に示す転送レートテーブルのデータを下記式 2 に代入して算出することができる。

40

【0037】

転送レート R_3 = (送信データサイズ / 転送時間 × 重み付け) / (重み付け) ... 式 2

【0038】

差分転送処理時間は、例えば、下記式 3 を用いて算出することができる。

【0039】

差分転送処理時間 T_2 = {通常転送処理時間 T_1 × (1 - ヒット率 H)} + ハッシュ値生成時間 T_3 + ハッシュ値転送時間 T_4 + ハッシュ値生成時間 T_5 + メッセージ送信時間 T_6 ... 式 3

【0040】

50

上記式 3 のヒット率 H は、転送元のファイル 1 4 1 と転送先のファイル 2 4 1 とに同じデータが存在する確率であり、例えば、図 4 に示すヒット率テーブルのデータを下記式 4 に代入して算出することができる。

【 0 0 4 1 】

ヒット率 H = ファイル 1 4 1 のヒットサイズ / ファイルサイズ A ... 式 4

【 0 0 4 2 】

上記式 3 のハッシュ値生成時間 T 3 は、転送元装置 1 0 でハッシュ値を生成するのに要する時間であり、例えば、下記式 5 を用いて算出することができる。

【 0 0 4 3 】

ハッシュ値生成時間 T 3 = ファイルサイズ A / ハッシュ値生成レート R 1 ... 式 5 10

【 0 0 4 4 】

上記式 5 のハッシュ値生成レート R 1 は、例えば、図 2 に示すハッシュ値生成レートテーブルのデータを下記式 6 に代入して算出することができる。

【 0 0 4 5 】

ハッシュ値生成レート R 1 = (ファイルサイズ / ハッシュ値生成時間 × 重み付け) / (重み付け) ... 式 6

【 0 0 4 6 】

上記式 3 のハッシュ値転送時間 T 4 は、転送元装置 1 0 から転送先装置 2 0 にハッシュ値を転送するのに要する時間であり、例えば、下記式 7 を用いて算出することができる。

【 0 0 4 7 】

ハッシュ値転送時間 T 4 = 分割ブロック数 N × ハッシュ値サイズ S / 転送レート R 3 ... 式 7 20

【 0 0 4 8 】

上記式 3 のハッシュ値生成時間 T 5 は、転送先装置 2 0 でハッシュ値を生成するのに要する時間であり、例えば、下記式 8 を用いて算出することができる。

【 0 0 4 9 】

ハッシュ値生成時間 T 5 = ファイルサイズ B / ハッシュ値生成レート R 2 ... 式 8

【 0 0 5 0 】

上記式 8 のファイルサイズ B は、転送先のファイル 2 4 1 のファイルサイズである。上記式 8 のハッシュ値生成レート R 2 は、図 5 に示すハッシュ値生成レートテーブルのデータを下記式 9 に代入して算出することができる。

【 0 0 5 1 】

ハッシュ値生成レート R 2 = (ファイルサイズ / ハッシュ値生成時間 × 重み付け) / (重み付け) ... 式 9 30

【 0 0 5 2 】

上記式 3 のメッセージ送信時間 T 6 は、転送処理の際に転送元装置 1 0 と転送先装置 2 0 との間でやり取りされる各種メッセージの送信に要する時間であり、例えば、下記式 1 0 を用いて算出することができる。

【 0 0 5 3 】

メッセージ送信時間 T 6 = メッセージのデータサイズ C / 転送レート R 3 ... 式 1 0 40

【 0 0 5 4 】

各種メッセージとしては、例えば、ハッシュ値取得要求や、不一致となったハッシュ値に対応するブロック番号を含む応答メッセージ等が該当する。

【 0 0 5 5 】

次に、実施形態におけるファイル同期システムの動作について説明する。図 6 は、ファイル同期システムにおけるファイル転送の処理手順を説明するためのフローチャートである。

【 0 0 5 6 】

最初に、転送元装置 1 0 の転送選択部 1 5 は、ファイル同期の対象となる転送元ファイル 1 4 1 のファイルサイズ A を取得し (ステップ S 1 0 1)、ファイルサイズ A が、分割ブロック数 N とハッシュ値サイズ S とを乗算した値よりも大きいか否かを判定する (ステップ S 1 0 2)。この判定が NO である場合 (ステップ S 1 0 2 ; NO) に、転送元装置 1 0 の転送選択部 1 5 は、通常転送を選択して本処理手順を終了する。

【 0 0 5 7 】

一方、上記ステップ S 1 0 2 の判定でファイルサイズ A が分割ブロック数 N とハッシュ値サイズ S とを乗算した値よりも大きいと判定した場合 (ステップ S 1 0 2 ; YES) に、転送元装置 1 0 の転送選択部 1 5 は、転送元ファイル 1 4 1 のファイル名を転送先装置 2 0 に送信する (ステップ S 1 0 3)。

10

【 0 0 5 8 】

続いて、転送先装置 2 0 は、転送元ファイル 1 4 1 のファイル名に基づいて、転送元ファイル 1 4 1 と同期をとるファイルに設定されている転送先ファイル 2 4 1 を特定し (ステップ S 1 0 4)、転送先ファイル 2 4 1 のファイルサイズ B を転送元装置 1 0 に送信する (ステップ S 1 0 5)。

【 0 0 5 9 】

続いて、転送元装置 1 0 の転送選択部 1 5 は、通常転送処理時間が差分転送処理時間よりも長いかなかを判定する (ステップ S 1 0 6)。この判定が NO である場合 (ステップ S 1 0 6 ; NO) に、転送元装置 1 0 の転送選択部 1 5 は、通常転送を選択して本処理手順を終了する。

20

【 0 0 6 0 】

一方、上記ステップ S 1 0 6 の判定で通常転送処理時間が差分転送処理時間よりも長いと判定した場合 (ステップ S 1 0 6 ; YES) に、転送元装置 1 0 の転送選択部 1 5 は、差分転送を選択し、転送元装置 1 0 および転送先装置 2 0 は、差分転送を実行する (ステップ S 1 0 7)。差分転送の処理手順について、以下に説明する。

【 0 0 6 1 】

図 7 は、図 6 のステップ S 1 0 7 で実行される差分転送の処理手順を詳細に説明するためのフローチャートである。

【 0 0 6 2 】

最初に、転送元装置 1 0 の実データ分割部 1 1 は、ファイル同期の対象となる転送元のファイル 1 4 1 の実データを分割ブロック数 N 分のブロックに分割する (ステップ S 2 0 1)。

30

【 0 0 6 3 】

続いて、転送元装置 1 0 のハッシュ値生成部 1 2 は、上記ステップ S 2 0 1 で N 個に分割された各ブロックに含まれる実データをそれぞれハッシュ関数に入力してハッシュ値を生成し、ハッシュ値の生成に要した時間を積算する (ステップ S 2 0 2)。

【 0 0 6 4 】

続いて、転送元装置 1 0 の送信部 1 3 は、上記ステップ S 2 0 2 で生成された各ハッシュ値、転送元ファイル 1 4 1 のファイル名、ファイルサイズ A、分割ブロック数 N、分割ブロックサイズ D およびハッシュ値サイズ S を含むハッシュ値取得要求を転送先装置 2 0 に送信する (ステップ S 2 0 3)。

40

【 0 0 6 5 】

続いて、転送先装置 2 0 の実データ分割部 2 1 は、ハッシュ値取得要求に基づいて、転送先のファイル 2 4 1 の実データを、分割ブロックサイズ D と同サイズのブロックに分割する (ステップ S 2 0 4)。

【 0 0 6 6 】

続いて、転送先装置 2 0 のハッシュ値生成部 2 2 は、上記ステップ S 2 0 4 で分割された各ブロックに含まれる実データをそれぞれハッシュ関数に入力してハッシュ値を生成し、ハッシュ値の生成に要した時間を積算する (ステップ S 2 0 5)。

50

【 0 0 6 7 】

続いて、転送先装置 20 の送信部 23 は、上記ステップ S 202 で生成されたハッシュ値と上記ステップ S 205 で生成されたハッシュ値とを、同じブロック番号ごとにそれぞれ比較し（ステップ S 206）、比較対象のハッシュ値が不一致となるブロック番号を特定する（ステップ S 207）。送信部 23 は、不一致であると特定したブロック番号を含む応答メッセージを転送元装置 10 に送信する（ステップ S 208）。

【 0 0 6 8 】

続いて、転送元装置 10 の送信部 13 は、転送先装置 20 から受信したブロック番号に対応するブロックに含まれる実データを転送先装置 20 に送信し（ステップ S 209）、実データの転送時間を積算する（ステップ S 210）。

10

【 0 0 6 9 】

上述したように、実施形態におけるファイル同期システム 1 によれば、転送元装置 10 は、更新後ファイルを管理することで差分転送を行うことができる。これにより、更新前ファイルや差分情報を管理するための記憶領域を削減することができ、更新前ファイルや差分情報のメンテナンス作業を省くことができる。それゆえに、ファイル同期の効率を向上させることができる。

【 0 0 7 0 】

このような効果は、ファイル同期の対象となるファイルが、固定長のレコードを格納するファイルやデータ追記型のファイルである場合に、特に有効となる。例えば、固定長のレコードを格納するファイルで特定のレコードが更新された場合には、更新されたレコードを含むブロックの実データのみを差分転送すればよい。データ追記型のファイルで実データが追記された場合には、追記された実データを含むブロックの実データのみを差分転送すればよい。これにより、同期処理のターンアラウンドタイムを大幅に短縮することができる。このような効果は、ファイルの容量が大きくなるほど増大する。

20

【 0 0 7 1 】

なお、上述した各実施形態は、単なる例示に過ぎず、実施形態に明示していない種々の変形や技術の適用を排除するものではない。すなわち、本発明は、その趣旨を逸脱しない範囲で様々な形態に変形して実施することができる。

【 0 0 7 2 】

例えば、上述した実施形態では、ハッシュ値生成レートや転送レートを算出する際に、過去の複数の転送処理に対して重み付けを付与しているが、例えば、最新の転送処理にのみ重み付けを付与して、他の転送処理には重み付けを付与しないこととしてもよい。これにより、CPU の使用状況やネットワークの使用状況を機敏に反映させたレートを算出することが可能となる。また、例えば、過去の複数の転送処理に対して均等に重み付けを付与することとしてもよい。これにより、CPU の使用状況やネットワークの使用状況が急激に変化するような不安定な環境下では、最新の転送処理の状況にのみ左右されることなく、過去の複数の転送処理の状況を同等に考慮したレートを算出することが可能となる。

30

【 0 0 7 3 】

上記の実施形態の一部または全部は、以下の付記のようにも記載され得るが、本発明を以下に限定するものではない。

40

【 0 0 7 4 】

（付記 1）ファイル同期の対象となるファイルをそれぞれ有する転送元装置と転送先装置とを備えるファイル同期システムであって、前記転送元装置は、ファイル同期の対象となる転送元の前記ファイルのデータを複数のブロックに分割する第 1 の分割部と、前記第 1 の分割部によって分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 1 の生成部と、前記第 1 の生成部によって生成されたハッシュ値を前記転送先装置に送信する第 1 の送信部と、を有し、前記転送先装置は、ファイル同期の対象となる転送先の前記ファイルのデータを、前記第 1 の分割部によって分割された各ブロックと同サイズのブロックに分割する第 2 の分割部と、前記第 2 の分割部によって分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 2 の生成部と、前記第 1 の生

50

成部によって生成されたハッシュ値と前記第 2 の生成部によって生成されたハッシュ値とを、各ハッシュ値の生成元データを含むブロックを識別するブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応する前記ブロック番号を前記転送元装置に送信する第 2 の送信部と、を有し、前記転送元装置の前記第 1 の送信部は、前記転送先装置から受信した前記ブロック番号に対応する前記ブロックに含まれるデータを前記転送先装置に送信する、ことを特徴とするファイル同期システム。

【 0 0 7 5 】

(付記 2) 前記ファイルは、固定長のレコードを格納するファイルであることを特徴とする付記 1 記載のファイル同期システム。

【 0 0 7 6 】

(付記 3) 前記ファイルは、追記型のファイルであることを特徴とする付記 1 または 2 記載のファイル同期システム。

【 0 0 7 7 】

(付記 4) ファイル同期の対象となるファイルをそれぞれ有する転送元装置と転送先装置とを備えるファイル同期システムで実行するファイル同期方法であって、前記転送元装置が、ファイル同期の対象となる転送元の前記ファイルのデータを複数のブロックに分割する第 1 の分割ステップと、前記第 1 の分割ステップにおいて分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 1 の生成ステップと、前記第 1 の生成ステップにおいて生成されたハッシュ値を前記転送先装置に送信する第 1 の送信ステップと、を実行し、前記転送先装置が、ファイル同期の対象となる転送先の前記ファイルのデータを、前記第 1 の分割ステップにおいて分割された各ブロックと同サイズのブロックに分割する第 2 の分割ステップと、前記第 2 の分割ステップにおいて分割された各ブロックに含まれるデータからそれぞれハッシュ値を生成する第 2 の生成ステップと、前記第 1 の生成ステップにおいて生成されたハッシュ値と前記第 2 の生成ステップにおいて生成されたハッシュ値とを、各ハッシュ値の生成元データを含むブロックを識別するブロック番号ごとにそれぞれ比較し、不一致となったハッシュ値に対応する前記ブロック番号を前記転送元装置に送信する第 2 の送信ステップと、を実行し、前記転送元装置の前記第 1 の送信ステップが、前記転送先装置から受信した前記ブロック番号に対応する前記ブロックに含まれるデータを前記転送先装置に送信する、ことを特徴とするファイル同期方法。

【 0 0 7 8 】

(付記 5) 付記 4 に記載の各ステップをコンピュータに実行させるためのファイル同期プログラム。

【 符号の説明 】

【 0 0 7 9 】

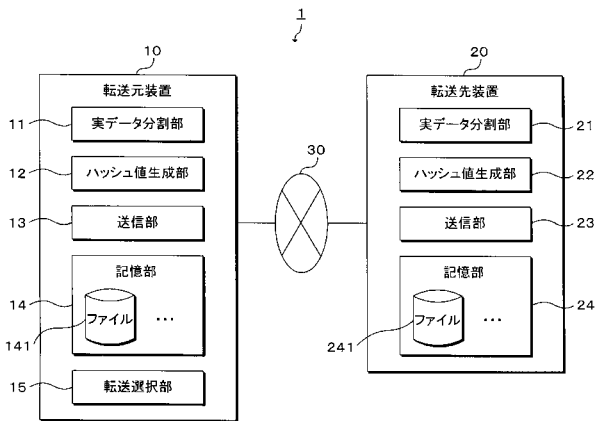
1 ... ファイル同期システム、 1 0 ... 転送元装置、 1 1 ... 実データ分割部、 1 2 ... ハッシュ値生成部、 1 3 ... 送信部、 1 4 ... 記憶部、 1 5 ... 転送選択部、 2 0 ... 転送先装置、 2 1 ... 実データ分割部、 2 2 ... ハッシュ値生成部、 2 3 ... 送信部、 2 4 ... 記憶部、 3 0 ... ネットワーク、 1 4 1、 2 4 1 ... ファイル。

10

20

30

【 図 1 】



【 図 2 】

転送処理名	ファイルサイズ	ハッシュ値生成時間	重み付け
最新の処理	1000	10	0.4
1つ前の処理	1000000	1400	0.3
2つ前の処理	50000	100	0.2
3つ前の処理	100000	150	0.1
⋮	⋮	⋮	⋮

【 図 3 】

転送処理名	転送データサイズ	転送時間	重み付け
最新の処理	900	1000	0.4
1つ前の処理	1000000	100000	0.3
2つ前の処理	30000	18000	0.2
3つ前の処理	90000	12000	0.1
⋮	⋮	⋮	⋮

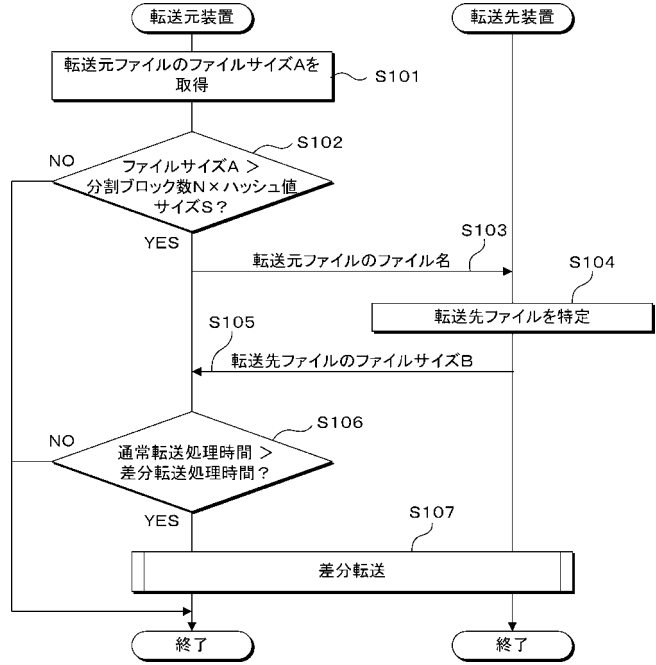
【 図 4 】

転送ファイル名	ファイルサイズ	ヒットサイズ
Sale. dat	900	900
Tax. html	1000000	950000
Explain. txt	30000	10
Price. csv	90000	80000
⋮	⋮	⋮

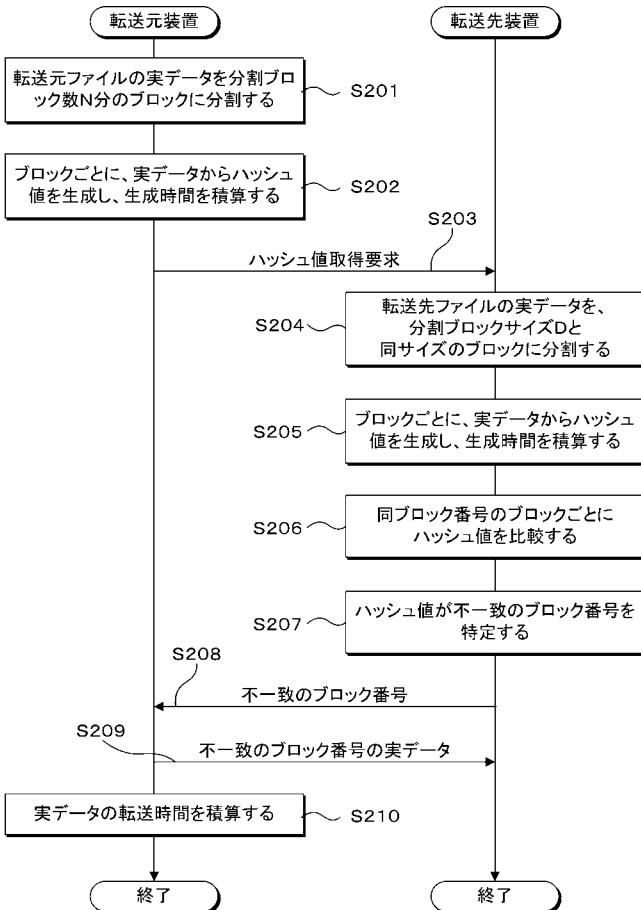
【 図 5 】

転送処理名	ファイルサイズ	ハッシュ値生成時間	重み付け
最新の処理	900	20	0.4
1つ前の処理	1000000	2000	0.3
2つ前の処理	30000	200	0.2
3つ前の処理	90000	140	0.1
⋮	⋮	⋮	⋮

【 図 6 】



【 図 7 】



フロントページの続き

(72)発明者 水野 佐年士

東京都港区芝五丁目7番1号 日本電気株式会社内

Fターム(参考) 5B082 HA05

5B084 BB18 CB03 DC22