

(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2008-276281

(P2008-276281A)

(43) 公開日 平成20年11月13日(2008.11.13)

| (51) Int.Cl.                | F I             | テーマコード (参考) |
|-----------------------------|-----------------|-------------|
| <b>G06F 12/00 (2006.01)</b> | G06F 12/00 531M | 5B034       |
| <b>G06F 11/20 (2006.01)</b> | G06F 12/00 533J | 5B065       |
| <b>G06F 3/06 (2006.01)</b>  | G06F 11/20 310C | 5B082       |
|                             | G06F 3/06 304E  |             |
|                             | G06F 3/06 304F  |             |

審査請求 未請求 請求項の数 18 O L (全 11 頁)

(21) 出願番号 特願2007-115517 (P2007-115517)  
 (22) 出願日 平成19年4月25日 (2007. 4. 25)

(71) 出願人 00004237  
 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 100096231  
 弁理士 稲垣 清  
 (72) 発明者 井上 雅貴  
 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) 5B034 BB02 BB11 BB15 CC01 CC02  
 DD06  
 5B065 BA01 EA31 EA33  
 5B082 DE06 HA03

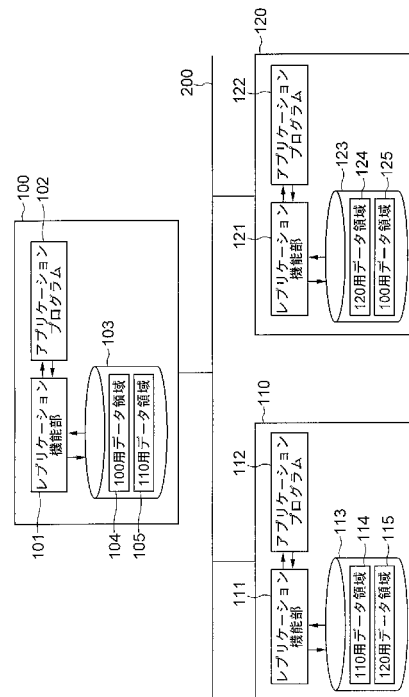
(54) 【発明の名称】 データ同期システム、方法、及び、プログラム

(57) 【要約】

【課題】 特別な外部記憶装置を必要とせずに、異なる機能を持つサーバ間でデータを同期させるデータ同期システムを提供する。

【解決手段】 ローカルディスク103は、自サーバ100のデータを記憶するオリジナルデータ領域104と、他のサーバ110のデータを記憶する同期データ領域105とを有する。レプリケーション機能部101は、アプリケーションプログラム102がデータの書き込み要求を発生すると、オリジナルデータ領域104にデータを書き込むと共に、サーバ120のレプリケーション機能部121を介して、ローカルディスク123の同期データ領域125に、データを書き込む。サーバ100に障害が発生したときには、サーバ120のローカルディスク123の同期データ領域125に記憶されたサーバ100のデータを、代替となるサーバのローカルディスクのオリジナルデータ領域にコピーすることで、復旧を行う。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

ネットワークを介して相互に接続された複数のサーバ間でデータを同期させるデータ同期システムにおいて、各サーバが、

自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを有するローカルディスクと、

自サーバのデータを、自サーバのローカルディスクの前記オリジナルデータ領域と、他のサーバが有するローカルディスクの前記同期データ領域とに書き込むレプリケーション機能部とを有することを特徴とするデータ同期システム。

**【請求項 2】**

前記レプリケーション機能部は、ネットワークにサーバが追加されると、各サーバから、前記オリジナルデータ領域及び前記同期データ領域にそれぞれどのサーバのデータを記憶しているかに関する情報を収集し、サーバ追加の原因を検出する、請求項 1 に記載のデータ同期システム。

**【請求項 3】**

前記レプリケーション機能部は、前記同期データ領域には記憶され、かつ、前記オリジナルデータ領域には記憶されないサーバのデータが存在すると判断すると、当該サーバに障害が発生し、該障害が発生したサーバの代替となるサーバが追加されたと判断する、請求項 2 に記載のデータ同期システム。

**【請求項 4】**

前記レプリケーション機能部は、代替サーバが追加されたと判断すると、前記障害が発生したサーバのデータを、当該データを記憶する前記同期データ領域から、前記代替サーバのオリジナルデータ領域にコピーし、前記障害が発生したサーバの前記同期データ領域に記憶されていたサーバのデータを、当該サーバのオリジナルデータ領域から、前記代替サーバの同期データ領域にコピーする、請求項 3 に記載のデータ同期システム。

**【請求項 5】**

前記レプリケーション機能部は、各サーバのデータが、前記オリジナルデータ領域及び前記同期データ領域の双方に記憶されていると判断すると、新規にサーバが追加されたと判断する、請求項 2 に記載のデータ同期システム。

**【請求項 6】**

前記レプリケーション機能部は、新規にサーバが追加されたと判断すると、該新規追加サーバを含めて、各サーバのデータが、前記オリジナルデータ領域と前記同期データ領域との双方に記憶されるように、前記同期データ領域にデータを記憶するサーバの再割当てを行う、請求項 5 に記載のデータ同期システム。

**【請求項 7】**

ネットワークを介して相互に接続された複数のサーバ間でデータを同期させるデータ同期方法において、

各サーバが、自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを有するローカルディスクの前記オリジナルデータ領域と、他のサーバが有するローカルディスクの前記同期データ領域とに、自サーバのデータを書き込むことを特徴とするデータ同期方法。

**【請求項 8】**

前記サーバが、ネットワークにサーバが追加されると、各サーバから、前記オリジナルデータ領域及び前記同期データ領域にそれぞれどのサーバのデータを記憶しているかに関する情報を収集し、サーバ追加の原因を検出する、請求項 7 に記載のデータ同期方法。

**【請求項 9】**

前記サーバは、前記同期データ領域には記憶され、かつ、前記オリジナルデータ領域には記憶されないサーバのデータが存在すると判断すると、当該サーバに障害が発生し、該障害が発生したサーバの代替となるサーバが追加されたと判断する、請求項 8 に記載のデータ同期方法。

10

20

30

40

50

**【請求項 10】**

前記サーバは、代替サーバが追加されたと判断すると、前記障害が発生したサーバのデータを、当該データを記憶する前記同期データ領域から、前記代替サーバのオリジナルデータ領域にコピーし、前記障害が発生したサーバの前記同期データ領域に記憶されていたサーバのデータを、当該サーバのオリジナルデータ領域から、前記代替サーバの同期データ領域にコピーする、請求項 9 に記載のデータ同期方法。

**【請求項 11】**

前記サーバは、各サーバのデータが、前記オリジナルデータ領域及び前記同期データ領域の双方に記憶されていると判断すると、新規にサーバが追加されたと判断する、請求項 8 に記載のデータ同期方法。

10

**【請求項 12】**

前記サーバは、新規にサーバが追加されたと判断すると、該新規追加サーバを含めて、各サーバのデータが、前記オリジナルデータ領域と前記同期データ領域との双方に記憶されるように、前記同期データ領域にデータを記憶するサーバの再割当てを行う、請求項 11 に記載のデータ同期方法。

**【請求項 13】**

ネットワークを介して相互に接続された複数のサーバ間でデータを同期させる処理を、前記サーバに実行させるプログラムであって、前記サーバに、

自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを有するローカルディスクの前記オリジナルデータ領域と、他のサーバが有するローカルディスクの前記同期データ領域とに、前記自サーバのデータを書き込む処理を実行させることを特徴とするプログラム。

20

**【請求項 14】**

前記サーバに、ネットワークにサーバが追加されると、各サーバから、前記オリジナルデータ領域及び前記同期データ領域にそれぞれどのサーバのデータを記憶しているかに関する情報を収集し、サーバ追加の原因を検出する処理を実行させる、請求項 13 に記載のプログラム。

**【請求項 15】**

前記サーバ追加の原因を検出する処理では、前記同期データ領域には記憶され、かつ、前記オリジナルデータ領域には記憶されないサーバのデータが存在すると判断すると、当該サーバに障害が発生し、該障害が発生したサーバの代替となるサーバが追加されたと判断する、請求項 14 に記載のプログラム。

30

**【請求項 16】**

前記サーバ追加の原因を検出する処理で代替サーバが追加されたと判断されると、前記サーバに、前記障害が発生したサーバのデータを、当該データを記憶する前記同期データ領域から、前記代替サーバのオリジナルデータ領域にコピーし、前記障害が発生したサーバの前記同期データ領域に記憶されていたサーバのデータを、当該サーバのオリジナルデータ領域から、前記代替サーバの同期データ領域にコピーする処理を実行させる、請求項 15 に記載のプログラム。

**【請求項 17】**

前記サーバ追加の原因を判断する処理では、各サーバのデータが、前記オリジナルデータ領域及び前記同期データ領域の双方に記憶されていると判断すると、新規にサーバが追加されたと判断する、請求項 14 に記載のプログラム。

40

**【請求項 18】**

前記サーバ追加の原因を検出する処理で新規にサーバが追加されたと判断されると、前記サーバに、該新規追加サーバを含めて、各サーバのデータが、前記オリジナルデータ領域と前記同期データ領域との双方に記憶されるように、前記同期データ領域にデータを記憶するサーバの再割当てを行う処理を実行させる、請求項 17 に記載のプログラム。

**【発明の詳細な説明】****【技術分野】**

50

## 【0001】

本発明は、データ同期システム、方法、及び、プログラムに関し、更に詳しくは、ネットワークを介して接続された複数のサーバ間で各サーバのデータを同期させるデータ同期システム、方法、及び、プログラムに関する。

## 【背景技術】

## 【0002】

耐障害性を高めるために、サーバ間でデータを同期させるシステムがある。データ同期に関し、複数サーバのファイル二重化方式として、特許文献1に記載の技術がある。特許文献1では、共有ディスクを用意し、共有ディスクに、各サーバのデータを記憶する。共有ディスクを用いないレプリケーション方式としては、特許文献2に記載の技術がある。特許文献2では、運用系（主系）でローカルディスクにデータを書き込むと、待機系（従系）のローカルディスクにも同じデータを書き込む。

10

## 【0003】

ファイル冗長化に関し、同一のファイルを、複数のファイルサーバに分散して保存するシステムが、特許文献3に記載されている。特許文献3では、ファイル保存システム内の各ドメインは、ファイルサーバと監視サーバとクライアントとがLANによって接続されることにより構成されている。ファイルサーバは、システム内に存在する他のファイルサーバに障害が発生したことを検出すると、障害が発生したファイルサーバに保存されているファイルの冗長度が低下したものとみなし、その冗長度を回復させる主体になるか否かを判定する。主体となると判定した場合には、ファイルのコピー元となるファイルサーバと、コピー先となるファイルサーバとを所定の条件に従って選択し、この選択に基づいて、コピー元ファイルサーバからコピー先ファイルサーバに、冗長度の低下したファイルと同一のファイルをコピーさせる。

20

## 【0004】

【特許文献1】特開平6-175788号公報

【特許文献2】特開2001-109642号公報

【特許文献3】特開2005-141528号公報

## 【発明の開示】

## 【発明が解決しようとする課題】

## 【0005】

特許文献1では、各サーバのローカルディスクと共有ディスクとを同期させることで、ファイル二重化を実現している。しかし、共有ディスクを用いる方式では、共有ディスクに、全てのサーバのデータを保持できるだけの容量が要求され、コストが高くなるという問題がある。また、特許文献2では、共用ディスクを用いることによるコスト高は回避できるものの、適用可能なシステムは、運用系と待機系とからなるクラスタシステムに限定され、特許文献2に記載の技術を、異なる機能を持つサーバ群に対して適用することはできないという問題がある。

30

## 【0006】

特許文献3では、例えば、サーバBに障害が発生したときには、サーバBが管理していたファイルを、他のサーバにコピーすることで、ファイルの冗長度を回復する。しかし、特許文献3に記載の技術は、ファイルサーバに関するものであり、サーバ間でのデータ同期を目的としたものではなく、ファイルが、複数のファイルサーバに記憶されていれば、その保存場所は問わない構成である。このため、サーバBの代替サーバを用意したとしても、その代替サーバに、サーバBが管理していたファイルがコピーされることはない。

40

## 【0007】

サーバが管理するファイルには、カーネルや各種設定ファイルが含まれており、代替サーバを用いて業務を再開するためには、障害が発生したサーバの動作環境を代替サーバに構築することが必要である。しかし、特許文献3では、ファイルの冗長性を回復することが目的であり、ファイルの保存場所は問わない構成であるので、特許文献3に記載の技術を、代替サーバに、障害が発生したサーバの環境を構築するという用途に適用することは

50

できない。

【0008】

本発明は、上記従来技術の問題点を解消し、共有ディスクなどの特殊な外部記憶装置を使用せずに、異なる機能を持ったサーバ間でのデータ同期を実現できるデータ同期システム、方法、及び、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】

【0009】

上記目的を達成するために、本発明のデータ同期方法は、ネットワークを介して相互に接続された複数のサーバ間でデータを同期させるデータ同期システムにおいて、各サーバが、自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを有するローカルディスクと、自サーバのデータを、自サーバのローカルディスクの前記オリジナルデータ領域と、他のサーバが有するローカルディスクの前記同期データ領域とに書き込むレプリケーション機能部とを有することを特徴とする。

10

【0010】

本発明のデータ同期方法は、ネットワークを介して相互に接続された複数のサーバ間でデータを同期させるデータ同期方法において、各サーバが、自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを有するローカルディスクの前記オリジナルデータ領域と、他のサーバが有するローカルディスクの前記同期データ領域とに、自サーバのデータを書き込むことを特徴とする。

20

【0011】

本発明のプログラムは、ネットワークを介して相互に接続された複数のサーバ間でデータを同期させる処理を、前記サーバに実行させるプログラムであって、前記サーバに、自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを有するローカルディスクの前記オリジナルデータ領域と、他のサーバが有するローカルディスクの前記同期データ領域とに、前記自サーバのデータを書き込む処理を実行させることを特徴とする。

【0012】

本発明のデータ同期システム、方法、及び、プログラムでは、各サーバは、自サーバのデータを、自サーバのローカルディスクのオリジナルデータ領域と、他のサーバが有するローカルディスクの同期データ領域とに書き込む。各サーバのローカルディスクに、自サーバのデータを記憶するためのオリジナルデータ領域と、他のサーバのデータを記憶する同期データ領域とを用意し、自サーバのデータを、自サーバのローカルディスクのオリジナルデータ領域と、他サーバのローカルディスクの同期データ領域との双方に書き込むことで、双方の領域のデータをリアルタイムに同期させることができ、データの二重化を実現できる。何れかのサーバに障害が発生したときには、障害が発生したサーバのデータを記憶するサーバの同期データ領域から、データをコピーすることで、障害復旧が可能となる。本発明では、各サーバのローカルディスクは、サーバ2台分の容量があればよく、共有ディスクなどの高価な設備は必要がない。

30

【0013】

本発明のデータ同期システム、方法、及び、プログラムは、ネットワークにサーバが追加されると、各サーバから、前記オリジナルデータ領域及び前記同期データ領域にそれぞれのサーバのデータを記憶しているかに関する情報を収集し、サーバ追加の原因を検出する構成を採用できる。何れかのサーバに障害が発生した場合には、何れかのサーバの同期データ領域には記憶されているものの、何れのサーバのオリジナルデータ領域にも記憶されていないサーバのデータが存在することになるので、サーバ追加の原因は、代替サーバの追加であると判断できる。また、既存サーバのデータが、オリジナルデータ領域と同期データ領域との双方に記憶されているときには、特に障害が発生しているサーバはないと判断できるので、サーバ追加の原因は、新規サーバの追加であると判断できる。

40

【0014】

50

本発明のデータ同期システム、方法、及び、プログラムは、前記同期データ領域には記憶され、かつ、前記オリジナルデータ領域には記憶されないサーバのデータが存在すると判断すると、当該サーバに障害が発生し、該障害が発生したサーバの代替となるサーバが追加されたと判断する構成を採用できる。また、本発明のデータ同期システム、方法、及び、プログラムでは、代替サーバが追加されたと判断すると、前記障害が発生したサーバのデータを、当該データを記憶する前記同期データ領域から、前記代替サーバのオリジナルデータ領域にコピーし、前記障害が発生したサーバの前記同期データ領域に記憶されていたサーバのデータを、当該サーバのオリジナルデータ領域から、前記代替サーバの同期データ領域にコピーする構成を採用できる。この場合、障害サーバのローカルディスクに記憶されていた自サーバのデータ及び他のサーバのデータを、代替サーバのローカルディスクにコピーすることで、障害発生前の状態に復旧することができ、代替サーバを用いた運用が可能となる。

10

**【0015】**

本発明のデータ同期システム、方法、及び、プログラムは、各サーバのデータが、前記オリジナルデータ領域及び前記同期データ領域の双方に記憶されていると判断すると、新規にサーバが追加されたと判断する構成を採用できる。また、本発明のデータ同期システム、方法、及び、プログラムでは、新規にサーバが追加されたと判断すると、該新規追加サーバを含めて、各サーバのデータが、前記オリジナルデータ領域と前記同期データ領域との双方に記憶されるように、前記同期データ領域にデータを記憶するサーバの再割当てを行う構成を採用できる。この場合、追加サーバを含めて、各サーバのデータが、オリジナルデータ領域と同期データ領域とに記憶されるように、同期データ領域に記憶するサーバのデータの割当てを行うことで、追加サーバを含めて、データを二重化できる。

20

**【発明の効果】****【0016】**

本発明のデータ同期システム、方法、及び、プログラムでは、各サーバは、自サーバのデータを、自サーバのローカルディスクのオリジナルデータ領域と、他のサーバが有するローカルディスクの同期データ領域とに書き込む。このようにすることで、自サーバのローカルディスクのオリジナルデータ領域と、他サーバのローカルディスクの同期データ領域とを同期させることができ、共有ディスクなどの特殊な外部記憶装置を使用せずに、異なる機能を持つサーバ間で、データ同期を実現できる。また、何れかのサーバに障害が発生したときには、障害が発生したサーバのデータを記憶するサーバの同期データ領域からデータをコピーすることで、障害復旧が可能となる。

30

**【発明を実施するための最良の形態】****【0017】**

以下、図面を参照し、本発明の実施の形態を詳細に説明する。図1は、本発明の一実施形態のデータ同期システムの構成を示している。図1では、3つのサーバ(100、110、120)により、データ同期システムが構成されている。これら3つのサーバは、それぞれ異なる機能を持つサーバである。各サーバ(100、110、120)は、レプリケーション機能部(101、111、121)、アプリケーションプログラム(102、112、122)、及び、ローカルディスク(103、113、123)を有する。各サーバは、ネットワーク200を介して接続されており、ネットワーク200を介して、相互に通信可能である。

40

**【0018】**

アプリケーションプログラム102は、サーバ100の機能を実現するプログラムである。ローカルディスク103は、サーバ100自身のデータを保持するためのオリジナルデータ領域104と、他サーバの同期データを保持するための同期データ領域105とを有する。サーバ110、120についても、サーバ100と同様な構成となっており、サーバ110のローカルディスク113は、サーバ110自身のデータを保持するためのオリジナルデータ領域114と、他サーバの同期データ(バックアップデータ)を保持するための同期データ領域115とを有する。また、サーバ120のローカルディスク123は

50

、サーバ120自身のデータを保持するためのオリジナルデータ領域124と、他サーバの同期データを保持するための同期データ領域125とを有する。各サーバのデータには、各サーバの動作に必要なカーネルや各種設定ファイルが含まれる。

【0019】

レプリケーション機能部101は、サーバ100自身のデータを、オリジナルデータ領域104に書き込むと共に、他サーバのレプリケーション機能部を介して、他サーバ上の同期データ領域に、オリジナルデータ領域104に書き込んだデータと同じデータを書き込む機能を有する。また、ネットワークにサーバが追加された際に、残りサーバから追加サーバへコピーすべきデータを割り出す機能を有する。レプリケーション機能部101の機能は、サーバ100上で所定のプログラムを実行することで実現される。サーバ110、120のレプリケーション機能部111、121についても、サーバ100のレプリケーション機能部101と同様な機能を有する。

10

【0020】

例えば、レプリケーション機能部101は、アプリケーションプログラム102からデータの書き込み要求を受け取ると、データをローカルディスク103のオリジナルデータ領域104に書き込む。レプリケーション機能部101は、オリジナルデータ領域104へのデータ書き込みが正常に完了すると、サーバ120のレプリケーション機能部121を介して、ローカルディスク123の同期データ領域125に、オリジナルデータ領域104に書き込んだデータと同じデータを書き込む。

【0021】

同様に、サーバ110のレプリケーション機能部111は、ローカルディスク113のオリジナルデータ領域114にデータを書き込むと共に、サーバ100のレプリケーション機能部101を介して、ローカルディスク103の同期データ領域105にもデータを書き込む。また、サーバ120のレプリケーション機能部121は、ローカルディスク123のオリジナルデータ領域124にデータを書き込むと共に、サーバ110のレプリケーション機能部111を介して、ローカルディスク113の同期データ領域115にもデータを書き込む。このようにすることで、各サーバのデータは、自サーバのローカルディスクと、他サーバのうちの何れかのローカルディスクとの双方に記憶されることになる。

20

【0022】

また、レプリケーション機能部は、ネットワーク200に新たにサーバが追加されると、各サーバから、オリジナルデータ領域にどのサーバのデータを記憶し、同期データ領域にどのサーバのデータを記憶するかを示す情報を収集し、サーバ追加の原因を検出する。レプリケーション機能部は、各サーバのデータが、オリジナルデータ領域と同期データ領域との双方に記憶されているときには、サーバが、新規に追加されたと判断する。この場合には、各サーバで保持する同期データの組合せを振り直し、追加サーバを含めて、各サーバのデータを、自サーバと他サーバとの双方で保持するようにする。

30

【0023】

一方、レプリケーション機能部は、何れかのサーバの同期データ領域には記憶されているものの、オリジナルデータ領域には記憶されていないサーバが存在するときには、そのサーバに障害が生じ、障害サーバの代替となるサーバが追加されたと判断する。代替サーバが追加されたときには、レプリケーション機能部は、障害サーバのデータを記憶する同期データ領域から、代替サーバのオリジナルデータ領域にデータをコピーする。また、障害サーバの同期データ領域に記憶されるサーバのデータを、そのサーバのオリジナルデータ領域からコピーする。

40

【0024】

図2は、運用時のデータ同期手順を示している。以下では、サーバ100のレプリケーション機能部101の動作を中心に説明する。アプリケーションプログラム102は、レプリケーション機能部101に、データ書き込みを要求する(ステップA1)。この要求を受けたレプリケーション機能部101は、ローカルディスク103上にあるオリジナルデータ領域104に、データを書き込む(ステップA2)。レプリケーション機能部101

50

は、書込みに成功したか否かを判断し（ステップ A 3）、書込みに失敗したときには、ステップ A 2 に戻り、リトライする。

【 0 0 2 5 】

レプリケーション機能部 1 0 1 は、データ書込みに成功すると、サーバ 1 0 0 の同期データを記憶するサーバを検索する（ステップ A 4）。検索の結果、同期データを記憶するサーバは、サーバ 1 2 0 であることが判明すると、レプリケーション機能部 1 0 1 は、サーバ 1 2 0 のレプリケーション機能部 1 2 1 に対して、書込み要求を行う（ステップ A 5）。レプリケーション機能部 1 2 1 は、この書込み要求を受けて、ステップ A 2 で書き込まれたデータと同じデータを、ローカルディスク 1 2 3 の同期データ領域 1 2 5 に書き込む（ステップ A 6）。レプリケーション機能部 1 2 1 は、書込みに成功したか否かを判断し（ステップ A 7）、書込みに失敗したときには、ステップ A 6 へ戻り、リトライする。成功したときには、処理を終了する。

10

【 0 0 2 6 】

図 3 は、障害発生などによりサーバ交換をした際のデータコピーの手順を示している。ここでは、図 4 に示すように、サーバ 1 0 0 に障害が発生し、その代替として、追加サーバ 1 3 0 を追加する場合を考える。障害が発生したサーバ 1 0 0 の代替として、新たにサーバ 1 3 0 がネットワーク 2 0 0 に接続されると（ステップ B 1）、各サーバのレプリケーション機能部は、各サーバのローカルディスクのオリジナルデータ領域及び同期データ領域に、どのサーバのデータを記憶しているかを示す情報を収集し、サーバ追加の原因を検出する（ステップ B 2）。

20

【 0 0 2 7 】

ステップ B 2 では、レプリケーション機能部は、各サーバから収集した情報を基に、保持データ管理テーブルを作成する。保持データ管理テーブルの一例を、図 5 に示す。この保持データ管理テーブルを参照することで、サーバ 1 1 0 のレプリケーション機能部 1 1 1 が管理するデータは、サーバ 1 1 0 自身のデータ、及び、サーバ 1 2 0 のバックアップのデータであり、サーバ 1 2 0 のレプリケーション機能部 1 2 1 が管理するデータは、サーバ 1 2 0 自身のデータ、及び、サーバ 1 0 0 のバックアップのデータであることがわかる。

【 0 0 2 8 】

図 5 では、サーバ 1 0 0 のバックアップデータはサーバ 1 2 0 に存在するものの、オリジナルデータが存在しないことがわかるので、ステップ B 2 では、サーバ 1 0 0 が障害により切断させたことがサーバ追加の原因であると検出して、追加サーバ 1 3 0 を、障害サーバ 1 0 0 の代替機と判断する（ステップ B 3）。代替機が追加された場合には、障害サーバ 1 0 0 のデータを記憶するサーバ 1 2 0 のレプリケーション機能部 1 2 1 は、同期データ領域 1 2 5 から、サーバ 1 0 0 のデータを追加サーバ 1 3 0 に送信する。追加サーバ 1 3 0 のレプリケーション機能部 1 3 1 は、受信したサーバ 1 0 0 のデータを、ローカルディスク 1 3 3 のオリジナルデータ領域 1 3 4 に書き込む（ステップ B 4）。

30

【 0 0 2 9 】

また、図 5 に示す保持データ管理テーブルを参照すると、サーバ 1 1 0 のデータを同期データ領域に保持するサーバが存在していないことがわかる。サーバ 1 1 0 のレプリケーション機能部 1 1 1 は、オリジナルデータ領域 1 1 4 の保持している自サーバのデータを、追加サーバ 1 3 0 に送信する。追加サーバ 1 3 0 のレプリケーション機能部 1 3 1 は、受信したサーバ 1 1 0 のデータを、ローカルディスク 1 3 3 の同期データ領域 1 3 5 に書き込む（ステップ B 5）。コピー完了後、レプリケーション機能部 1 3 1 は、通常のレプリケーション処理（図 2）を開始する（ステップ B 6）。その後、追加サーバ 1 3 0 は、アプリケーションプログラム 1 3 2 が動作することで、サーバ 1 0 0 のデータを引き継いで動作する。

40

【 0 0 3 0 】

図 6 は、ネットワークに新規サーバが追加された際のデータコピー手順を示している。新規サーバの追加は、図 4 にて、サーバ 1 0 0 に障害が発生せず、サーバ 1 0 0 が動作を

50



続けた状態で、サーバ130が追加された状態に相当する。各サーバのレプリケーション機能部は、追加されたサーバ130を検出する(ステップC1)。追加サーバ130を検出すると、各サーバのレプリケーション機能部は、各サーバのローカルディスクのオリジナルデータ領域及び同期データ領域に、どのサーバのデータを記憶しているかを示す情報を収集し、サーバ追加の原因を検出する(ステップC2)。

#### 【0031】

ステップC2では、レプリケーション機能部は、各サーバから収集した情報を基に、保持データ管理テーブルを作成する。保持データ管理テーブルの一例を、図7に示す。この保持データ管理テーブルを参照すると、各サーバのデータは、オリジナルデータ領域と、同期データ領域との双方に存在するため、特に障害となっているサーバがないことがわかる。この場合、レプリケーション機能部は、追加サーバ130を、新規追加サーバと判断する(ステップC3)。新規にサーバが追加された場合は、レプリケーション機能部は、新規追加サーバ130を含めた全サーバで、同期データを保持するように、各サーバがどのサーバの同期データを保持するかを割当てし直す(ステップC4)。

10

#### 【0032】

ステップC4での割当て変更後の保持データ管理テーブルの具体例を、図8に示す。レプリケーション機能部は、どのサーバが、どのサーバの同期データを保持するかの割当てを変更することで、図8に示すように、各サーバのデータが、オリジナルデータ領域と同期データ領域との双方に記憶されるようにする。割当てが決定すると、決定した割当てに従って、各サーバの同期データをコピーする(ステップC5)。ステップC5では、図8に示す割当てに従って、例えば、サーバ120のローカルディスク123のオリジナルデータ領域に記憶されたデータを、新規追加サーバ130のローカルディスク133の同期データ記憶領域135にコピーする。コピーの完了後、レプリケーション機能部131がレプリケーション処理(図2)を開始する(ステップC6)。

20

#### 【0033】

本実施形態では、各サーバのローカルディスクに、自サーバのデータを記憶するためのオリジナルデータ領域と、他サーバのデータを記憶するための同期データ領域とを設ける。レプリケーション機能部は、自サーバのローカルディスクのオリジナルデータ領域に、データを書き込むと共に、他サーバのうちの何れかのローカルディスクの同期データ領域に、オリジナルデータ領域に書き込んだデータと同じデータを書き込む。このようにすることで、共有ディスクなどの高価な設備を必要とせずに、各サーバのデータ同期を行うことができる。また、サーバに障害が発生する直前までデータを同期しているため、定期的にバックアップを行う場合に比して、多くのデータを救済できる。更に、各サーバのローカルディスクに必要な容量は、サーバ2台分に収まり、コスト増が抑えられる。データ同期システムの対象は、現用系と待機系とから構成されるクラスタシステムに限られず、別個の機能を持ったサーバ群にも適用できる。

30

#### 【0034】

以上、本発明をその好適な実施形態に基づいて説明したが、本発明のデータ同期システム、方法、及び、プログラムは、上記実施形態にのみ限定されるものではなく、上記実施形態の構成から種々の修正及び変更を施したものも、本発明の範囲に含まれる。

40

#### 【産業上の利用可能性】

#### 【0035】

本発明は、処理速度よりも耐障害性が重視される計算機、例えばファイルサーバなどにおいて安価にミラーリングを実現するといった用途に適用できる。

#### 【図面の簡単な説明】

#### 【0036】

【図1】本発明の一実施形態のデータ同期システムの構成を示すブロック図。

【図2】通常運用時の動作手順を示すフローチャート。

【図3】代替サーバ追加時の動作手順を示すフローチャート。

【図4】代替サーバ追加時のデータ同期システムを示すブロック図。

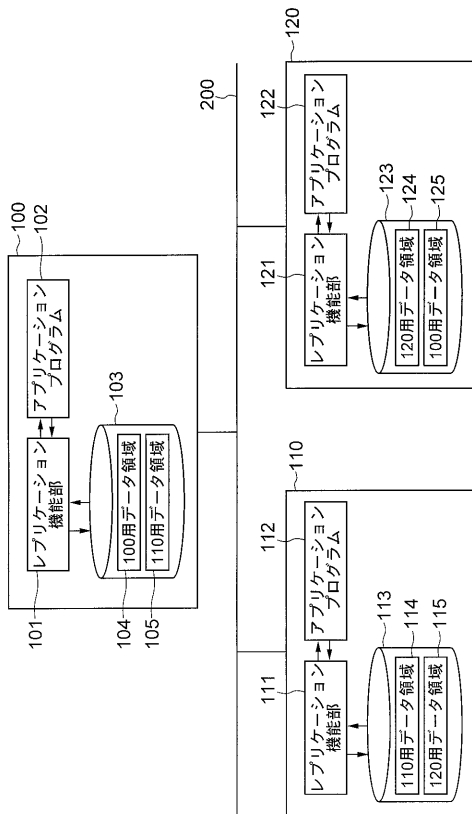
50

- 【図5】代替サーバ追加時の保持データ管理テーブルを示す図。
- 【図6】新規サーバ追加時の動作手順を示すフローチャート。
- 【図7】新規サーバ追加時の保持データ管理テーブルを示す図。
- 【図8】新規サーバ追加時の再割当て後の保持データ管理テーブルを示す図。
- 【符号の説明】

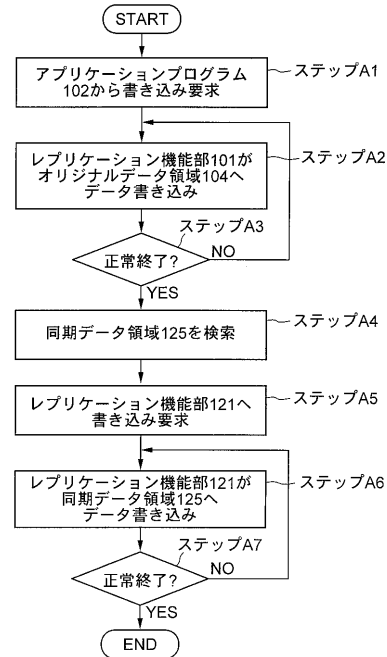
【0037】

- 100、110、120、130：サーバ
- 101、111、121、131：レプリケーション機能部
- 102、112、122、132：アプリケーションプログラム
- 103、113、123、133：ローカルディスク
- 104、114、124、134：オリジナルデータ領域
- 105、115、125、135：同期データ領域

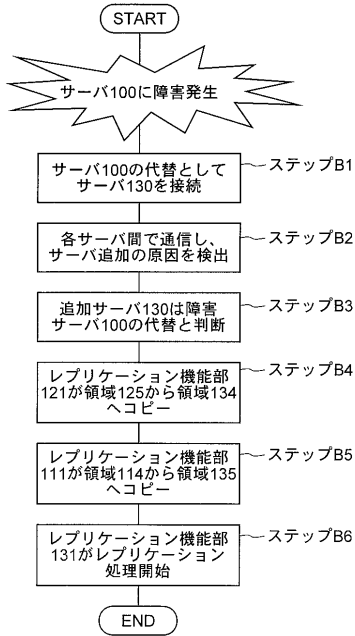
【図1】



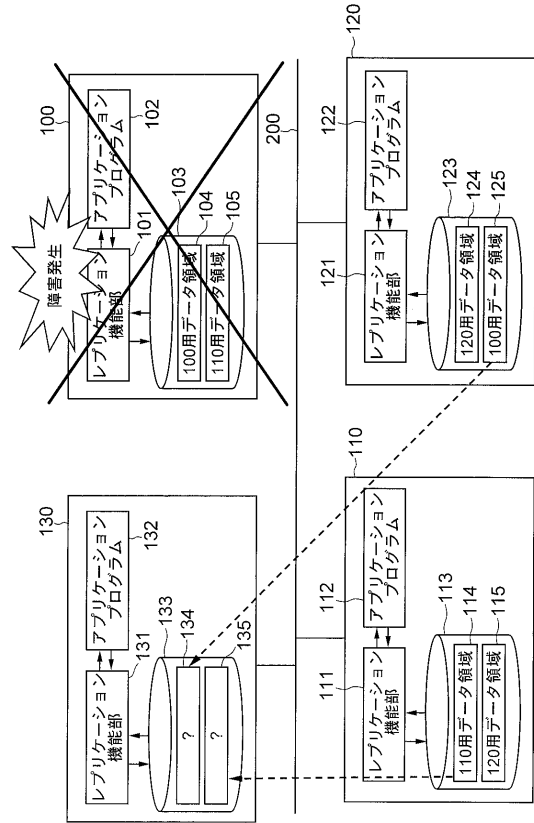
【図2】



【 図 3 】



【 図 4 】



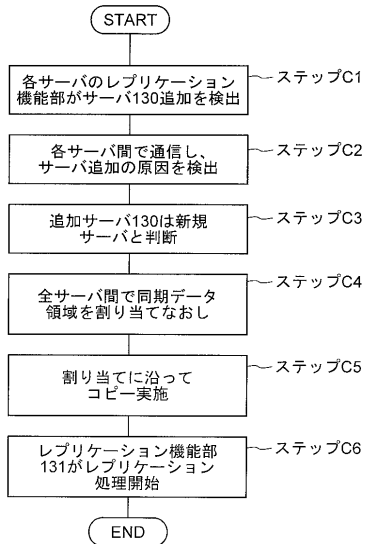
【 図 5 】

|            | サーバ100     | サーバ110     | サーバ120     | サーバ130 |
|------------|------------|------------|------------|--------|
| オリジナルデータ領域 | 障害         | サーバ110のデータ | サーバ120のデータ | 初期状態   |
| 同期データ領域    | サーバ100のデータ | サーバ120のデータ | サーバ110のデータ | 初期状態   |

【 図 7 】

|            | サーバ100     | サーバ110     | サーバ120     | サーバ130 |
|------------|------------|------------|------------|--------|
| オリジナルデータ領域 | サーバ100のデータ | サーバ110のデータ | サーバ120のデータ | 初期状態   |
| 同期データ領域    | サーバ120のデータ | サーバ120のデータ | サーバ100のデータ | 初期状態   |

【 図 6 】



【 図 8 】

|            | サーバ100     | サーバ110     | サーバ120     | サーバ130     |
|------------|------------|------------|------------|------------|
| オリジナルデータ領域 | サーバ100のデータ | サーバ110のデータ | サーバ120のデータ | サーバ130のデータ |
| 同期データ領域    | サーバ130のデータ | サーバ100のデータ | サーバ110のデータ | サーバ120のデータ |