

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4571576号
(P4571576)

(45) 発行日 平成22年10月27日 (2010.10.27)

(24) 登録日 平成22年8月20日 (2010.8.20)

(51) Int. Cl.	F I				
H04L 12/56	(2006.01)	H04L 12/56	100Z		
G06F 12/00	(2006.01)	G06F 12/00	545M		
G06F 13/00	(2006.01)	G06F 13/00	351N		
G06F 3/06	(2006.01)	G06F 3/06	304F		

請求項の数 11 (全 21 頁)

(21) 出願番号	特願2005-338333 (P2005-338333)	(73) 特許権者	000005108
(22) 出願日	平成17年11月24日 (2005.11.24)		株式会社日立製作所
(65) 公開番号	特開2007-150409 (P2007-150409A)		東京都千代田区丸の内一丁目6番6号
(43) 公開日	平成19年6月14日 (2007.6.14)	(74) 代理人	100064414
審査請求日	平成20年7月8日 (2008.7.8)		弁理士 磯野 道造
		(72) 発明者	天木 健介
			神奈川県横浜市戸塚区戸塚町5030番地
			株式会社日立製作所
			ソフトウェア事業部内
		(72) 発明者	渡辺 和彦
			神奈川県横浜市戸塚区戸塚町5030番地
			株式会社日立製作所
			ソフトウェア事業部内
		審査官	倉山 徹男
			最終頁に続く

(54) 【発明の名称】 リモートコピー記憶装置システムおよびリモートコピー方法

(57) 【特許請求の範囲】

【請求項1】

複数の業務処理を実行するセンタコンピュータと、
 前記センタコンピュータに接続され、前記センタコンピュータが前記業務処理を実行した結果得られる情報を記憶するセンタ記憶装置と、
 前記センタ記憶装置にそれぞれ通信回線を介して接続され、前記センタ記憶装置が送信するコピーデータをそれぞれ受信して記憶する複数のリモート記憶装置と
 を含んで構成されたりモートコピー記憶装置システムであって、
 前記センタコンピュータおよび前記センタ記憶装置の少なくとも一方は、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を計測する負荷計測手段を備え、
 前記センタコンピュータは、
 前記負荷計測手段を介して、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を取得し、
 前記取得した負荷情報が所定の閾値を超えた通信回線を経路切替対象通信回線として抽出し、
 前記抽出した経路切替対象通信回線を介して送信されるコピーデータに關与する業務処理の中から一部の業務処理を抽出し、
 前記抽出した業務処理の実行を停止することによってその業務処理に係るリモートコピーを停止し、

前記センタ記憶装置および前記リモート記憶装置に対して、前記停止した業務処理に係るコピーデータの送信経路を、前記経路切替対象通信回線と異なる他の通信回線に切り替える旨を指示する経路切替指示を通知し、

前記センタ記憶装置および前記リモート記憶装置は、それぞれ、

前記経路切替指示の通知を受けると、前記コピーデータの送信経路を切り替える処理を行い、

前記センタコンピュータは、その後、

前記停止した業務処理の実行を再開することによってリモートコピーを再開することを特徴とするリモートコピー記憶装置システム。

【請求項2】

前記センタ記憶装置および前記リモート記憶装置は、それぞれ自らに対する情報の書き込み動作についてその動作の履歴情報を記憶するジャーナルを、さらに、備え、

前記センタコンピュータは、

前記センタ記憶装置および前記リモート記憶装置に対し、前記経路切替指示を通知するに際しては、前記リモート記憶装置のジャーナルを前記センタ記憶装置のジャーナルに同期化させた後、前記経路切替指示を通知すること

を特徴とする請求項1に記載のリモートコピー記憶装置システム。

【請求項3】

前記センタコンピュータは、

前記抽出した経路切替対象通信回線を介して送信されるコピーデータに関与する業務処理の中から一部の業務処理を抽出するときには、その業務処理にあらかじめ設定された優先度が低い業務処理から順に抽出すること

を特徴とする請求項1または請求項2に記載のリモートコピー記憶装置システム。

【請求項4】

前記経路切替対象通信回線と異なる他の通信回線は、前記経路切替対象通信回線に接続された前記リモート記憶装置と、そのリモート記憶装置と異なる他のリモート記憶装置とをつなぐ通信回線であること

を特徴とする請求項1ないし請求項3のいずれか1項に記載のリモートコピー記憶装置システム。

【請求項5】

複数の業務処理を実行するセンタコンピュータと、

前記センタコンピュータに接続され、前記センタコンピュータが前記業務処理を実行した結果得られる情報を記憶するセンタ記憶装置と、

前記センタ記憶装置にそれぞれ通信回線を介して接続され、前記センタ記憶装置が送信するコピーデータをそれぞれ受信して記憶する複数のリモート記憶装置と

を含んで構成されたりモートコピー記憶装置システムに用いられるセンタコンピュータであって、

前記センタコンピュータおよび前記センタ記憶装置の少なくとも一方に備えられた通信回線の負荷情報を計測する負荷計測手段を介して、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を取得し、

前記取得した負荷情報が所定の閾値を超えた通信回線を経路切替対象通信回線として抽出し、

前記抽出した経路切替対象通信回線を介して送信されるコピーデータに関与する業務処理の中から一部の業務処理を抽出し、

前記抽出した業務処理の実行を停止することによってその業務処理に係るリモートコピーを停止し、

前記センタ記憶装置および前記リモート記憶装置に対して、前記停止した業務処理に係るコピーデータの送信経路を、前記経路切替対象通信回線と異なる他の通信回線に切り替える旨を指示する経路切替指示を通知し、

前記経路切替指示の通知を受けて、前記センタ記憶装置および前記リモート記憶装置が

10

20

30

40

50

前記コピーデータの送信経路を切り替えた後、前記停止した業務処理の実行を再開することによってリモートコピーを再開すること
を特徴とするセンタコンピュータ。

【請求項6】

前記センタ記憶装置および前記リモート記憶装置は、それぞれ自らに対する情報の書き込み動作についてその動作の履歴情報を記憶するジャーナルを、さらに、備え、

前記センタコンピュータは、

前記センタ記憶装置および前記リモート記憶装置に対し、前記経路切替指示を通知するに際しては、前記リモート記憶装置のジャーナルを前記センタ記憶装置のジャーナルに同期させた後、前記経路切替指示を通知すること

を特徴とする請求項5に記載のセンタコンピュータ。

10

【請求項7】

前記センタコンピュータは、

前記抽出した経路切替対象通信回線を介して送信されるコピーデータに關与する業務処理の中から一部の業務処理を抽出するときには、その業務処理にあらかじめ設定された優先度が低い業務処理から順に抽出すること

を特徴とする請求項5または請求項6に記載のセンタコンピュータ。

【請求項8】

複数の業務処理を実行するセンタコンピュータと、

前記センタコンピュータに接続され、前記センタコンピュータが前記業務処理を実行した結果得られる情報を記憶するセンタ記憶装置と、

前記センタ記憶装置にそれぞれ通信回線を介して接続され、前記センタ記憶装置が送信するコピーデータをそれぞれ受信して記憶する複数のリモート記憶装置と

を含んで構成されたりモートコピー記憶装置システムにおけるリモートコピー方法であって、

前記センタコンピュータおよび前記センタ記憶装置の少なくとも一方は、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を計測する負荷計測手段を備え、

前記センタコンピュータは、

前記負荷計測手段を介して、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を取得し、

前記取得した負荷情報が所定の閾値を超えた通信回線を経路切替対象通信回線として抽出し、

前記抽出した経路切替対象通信回線を介して送信されるコピーデータに關与する業務処理の中から一部の業務処理を抽出し、

前記抽出した業務処理の実行を停止することによってその業務処理に係るリモートコピーを停止し、

前記センタ記憶装置および前記リモート記憶装置に対して、前記停止した業務処理に係るコピーデータの送信経路を、前記経路切替対象通信回線と異なる他の通信回線に切り替える旨を指示する経路切替指示を通知し、

前記センタ記憶装置および前記リモート記憶装置は、それぞれ、

前記経路切替指示の通知を受けると、前記コピーデータの送信経路を切り替え、

前記センタコンピュータは、その後、

前記停止した業務処理の実行を再開することによってリモートコピーを再開すること

を特徴とするリモートコピー方法。

【請求項9】

前記センタ記憶装置および前記リモート記憶装置は、それぞれ自らに対する情報の書き込み動作についてその動作の履歴情報を記憶するジャーナルを、さらに、備え、

前記センタコンピュータは、

50

前記センタ記憶装置および前記リモート記憶装置に対し、前記経路切替指示を通知するに際しては、前記リモート記憶装置のジャーナルを前記センタ記憶装置のジャーナルに同期させた後、前記経路切替指示を通知すること

を特徴とする請求項 8 に記載のリモートコピー方法。

【請求項 10】

前記センタコンピュータは、

前記抽出した経路切替対象通信回線を介して送信されるコピーデータに関与する業務処理の中から一部の業務処理を抽出するときには、その業務処理にあらかじめ設定された優先度が低い業務処理から順に抽出すること

を特徴とする請求項 8 または請求項 9 に記載のリモートコピー方法。

10

【請求項 11】

前記経路切替対象通信回線と異なる他の通信回線は、前記経路切替対象通信回線に接続された前記リモート記憶装置と、そのリモート記憶装置と異なる他のリモート記憶装置とをつなぐ通信回線であること

を特徴とする請求項 8 ないし請求項 10 のいずれか 1 項に記載のリモートコピー方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、センタ記憶装置に記憶する情報と同じ情報を複数のリモート記憶装置に記憶するリモートコピー記憶装置システム、センタコンピュータおよびリモートコピー方法に関する。

20

【背景技術】

【0002】

コンピュータとその通信ネットワークは、今や人間の経済活動を支えるための不可欠のインフラとなっている。そのため、インフラとしてのコンピュータに障害が発生すると、企業の経済活動のみならず一般の人々の社会生活にも大きな影響が及ぶことになる。このようなコンピュータの障害の中でも最も深刻な影響があるとされているものは、記憶装置に記憶されているデータの喪失である。コンピュータを構成するハードウェアやソフトウェアは、障害が生じて、正常なもので代替することが可能であるが、コンピュータの記憶装置に記憶されているデータは、いったん失われると元に戻すことができない。

30

【0003】

そこで、地震など大規模な災害などによって記憶装置に記憶されているデータが失われるのを防ぐことを目的に、互いに離れた 2 箇所の拠点にそれぞれ記憶装置を設置し、その 2 箇所の拠点に設置された記憶装置間でデータを互いに複製して記憶するリモートコピー技術が実用化されている。そして、記憶装置の市場ニーズは、記憶装置を設置する拠点を 3 箇所以上にすることができるリモートコピー技術を求めている。

【0004】

例えば、特開 2003 - 122509 号公報には、3 箇所以上の拠点に設置された記憶装置間で、データ更新の順序性を確保した上で、常時、データを複製することが可能なリモートコピー技術が開示されている。そのリモート技術によれば、災害などによって、ある 1 つの記憶装置に障害が発生し、その記憶装置からデータが喪失されても、他の 2 つ以上の記憶装置に記憶されているデータに基づき、順序性を保証したデータを再構築することができる。

40

【特許文献 1】特開 2003 - 122509 号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

しかしながら、特許文献 1 に開示された技術では、3 つ以上の記憶装置間においてリモートコピーを行う場合、コピー元からコピー先へデータを送信するときに、その通信回線における通信負荷が過負荷になることへの配慮がなされていない。そのために、その通信

50

回線における通信負荷が過負荷になったときには、コピー処理が遅延し、その遅延のためにコンピュータにおける業務処理の性能低下を招く問題があることが判明してきた。

【0006】

すなわち、コピー先が複数になると、そのコピー先の数に比例して、コピー元から送信されるデータ量が増加することになる。その場合、コピー元の送信能力やコピー元とコピー先とをつなぐ通信回線の転送帯域が、そのリモートコピーのデータ送信によって消費され、元々備えているコピー元の送信能力や通信回線の転送帯域の余裕が小さくなってしまふ。一方、コンピュータの業務処理によって記憶装置へ書き込まれるデータ量は、日時や季節により大きく変動することが多い。従って、リモートコピーのコピー先を複数にした場合には、コピー元から送信するデータ量が、コピー元の送信能力を超えたり、コピー元とコピー先とをつなぐ通信回線の転送帯域を超えたりする危険性、つまり、その通信回線が過負荷になる危険性が增大することになる。

10

【0007】

現実に、コピー元から送信する送信データ量がコピー元の送信能力を超えたり、コピー元とコピー先とをつなぐ通信回線の転送帯域を超えたりした場合、つまり、コピー元とコピー先とをつなぐ通信回線の通信負荷が過負荷になった場合には、データの転送が大幅に遅延し、コンピュータ本来の業務処理の性能が低下したり、障害や災害が発生したときには、コピーすべきデータを喪失したりする。従って、通信回線の通信負荷が過負荷にならないようにする方策が必要であるが、特許文献1に開示された技術では、その方策は何ら示されていない。

20

【0008】

以上の従来技術の問題に鑑み、本発明の目的は、センタ記憶装置に記憶する情報と同じ情報を複数のリモート記憶装置に記憶するリモートコピー記憶装置システムにおいて、コンピュータが実行する業務処理によるセンタ記憶装置への書き込みデータが増加しても、センタ記憶装置（コピー元）とリモート記憶装置（コピー先）とをつなぐ通信回線の通信負荷が過負荷になるのを防止し、さらには、前記コンピュータにおける業務処理のスループットの低下を防止することが可能なリモートコピー記憶装置システム、センタコンピュータおよびリモートコピー方法を提供することにある。

【課題を解決するための手段】

【0009】

以上の目的を達成するために、本発明においては、リモートコピー記憶装置システムを、複数の業務処理を実行するセンタコンピュータと、前記センタコンピュータに接続され、前記センタコンピュータが前記業務処理を実行した結果得られる情報を記憶するセンタ記憶装置と、前記センタ記憶装置にそれぞれ通信回線を介して接続され、前記センタ記憶装置が送信するコピーデータをそれぞれ受信して記憶する複数のリモート記憶装置とを含むように構成し、前記センタコンピュータおよび前記センタ記憶装置の少なくとも一方は、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を計測する負荷計測手段を備えるものとした。そして、そのリモートコピー記憶装置システムにおいて、(1)前記センタコンピュータは、(1a)前記負荷計測手段を介して、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷情報を取得し、(1b)前記取得した負荷情報が所定の閾値を超えた通信回線を経路切替対象通信回線として抽出し、(1c)前記抽出した経路切替対象通信回線を介して送信されるコピーデータに關与する業務処理の中から一部の業務処理を抽出し、(1d)前記抽出した業務処理の実行を停止することによってその業務処理に係るリモートコピーを停止し、(1e)前記センタ記憶装置および前記リモート記憶装置に対して、前記停止した業務処理に係るコピーデータの送信経路を、前記経路切替対象通信回線と異なる他の通信回線に切り替える旨を指示する経路切替指示を通知し、(2)前記センタ記憶装置および前記リモート記憶装置は、それぞれ、前記経路切替指示の通知を受けると、前記コピーデータの送信経路を切り替える処理を行い、(3)前記センタコンピュータは、その後、前記停止した業務処理の実行を再開することによってリモートコピーを再開することを

30

40

50

特徴とする。

【0010】

以上のように、本発明によれば、前記センタコンピュータは、前記センタ記憶装置と前記複数のリモート記憶装置とをそれぞれ接続する通信回線の負荷を監視し、過負荷になる恐れのある通信回線を検出する。そして、前記センタ記憶装置と前記複数のリモート記憶装置とに対し、その通信回線を使用して送信されている情報のうち一部の業務処理の情報についてその送信経路を切り替えることを指示する。その指示を受けたセンタ記憶装置およびリモート記憶装置は、その送信経路を前記の過負荷になる恐れのある通信回線を使用しない送信経路に切り替える。従って、その過負荷になる恐れのある通信回線は、通信負荷が軽減されるので、その通信回線が過負荷になることを防止することができる。また、その結果、通信回線が過負荷になることによるセンタコンピュータの処理性能の低下を防止することができる。

10

【発明の効果】

【0011】

本発明により、リモートコピー記憶装置システムにおいて、センタ記憶装置とリモート記憶装置とをつなぐ通信回線の通信負荷が過負荷になるのを防止することができるようになり、その結果、コンピュータにおける業務処理の性能が低下するのを防止することができるようになる。

【発明を実施するための最良の形態】

【0012】

以下、本発明の実施形態について、適宜、図面を参照しながら詳しく説明する。

20

【0013】

まず、図1および図2を参照して、本発明の実施形態に係るリモートコピー記憶装置システムの概略構成および通信回線の過負荷を防止する方法について説明する。図1は、本発明の実施形態に係るリモートコピー記憶装置システムの概略構成の例、および、通常状態におけるコピーデータの送信経路の例を示した図である。

【0014】

図1に示すように、リモートコピー記憶装置システム1は、センタコンピュータ10およびセンタ記憶装置20が設置された業務センタ2と、リモート記憶装置30a, 30bがそれぞれ設置された付近地データセンタ3および遠隔地データセンタ4とを含んで構成される。そして、センタ記憶装置20と付近地データセンタ3のリモート記憶装置30aとは通信回線5によって接続され、センタ記憶装置20と遠隔地データセンタ4のリモート記憶装置30bとは通信回線6によって接続される。また、付近地データセンタ3のリモート記憶装置30aと遠隔地データセンタ4のリモート記憶装置30bとは、通信回線7によって接続される。このとき、通信回線5, 6, 7は、好ましくは、それぞれが独立した光ファイバ回線などによって構成される。

30

【0015】

ここで、付近地データセンタ3は、業務センタ2とそれほど遠く離れていない所（例えば、10km～50km程度離れた所）に設けられ、遠隔地データセンタ4は、業務センタ2からかなり遠くに離れた所（例えば、100km以上離れた所）に設けられる。具体的には、業務センタ2を、例えば、東京の都心に設けたとすれば、付近地データセンタ3は、例えば、さいたま市に設けられ、遠隔地データセンタ4は、例えば、大阪市に設けられる。

40

【0016】

次に、図1を参照してセンタコンピュータ10、センタ記憶装置20およびリモート記憶装置30a, 30bの構成および機能について説明するが、ここでは、その概略構成および概略機能の説明にとどめ、詳しい構成および機能は、別途、図3および図4を用いて説明する。

【0017】

センタコンピュータ10は、所定の業務を処理するために、業務A処理部100、業務

50

B 処理部 200、業務 C 処理部 300 などのプログラムを備え、それらのプログラムを実行し、その結果として得られる情報をセンタ記憶装置 20 のハードディスク装置 50 に、それぞれ業務 A データ 51、業務 B データ 52、業務 C データ 53 などのデータとして書き込む。なお、ここでいう業務処理とは、例えば、銀行のオンライン業務、銀行社員の給与計算業務などのように、全く異なるデータベースに基づきコンピュータによって処理される業務処理をいう。

【0018】

また、センタコンピュータ 10 は、経路切替処理部 800 を備え、経路切替処理部 800 によってセンタ記憶装置 20 とリモート記憶装置 30a、30b とをそれぞれ接続する通信回線 5、6 の通信負荷情報を取得する。そして、センタコンピュータ 10 は、その通信負荷情報が所定の閾値を超える通信回線 5 (または 6) を検出し、その通信回線 5 (または 6) を用いてセンタ記憶装置 20 から送信される情報の一部を別経路により送信するようにするための送信経路切替指示を、センタ記憶装置 20 とリモート記憶装置 30a、30b とに対し出力する。ここで、別経路とは、リモート記憶装置 30a、30b 間を接続する通信回線 7 を指す。なお、この経路切替処理部 800 の詳しい処理フローについては、別途、図を用いて説明する。

10

【0019】

センタ記憶装置 20 は、ハードディスク装置 50 を備え、センタコンピュータ 10 が業務 A 処理部 100、業務 B 処理部 200、業務 C 処理部 300 などのプログラムを実行した結果得られる情報を、それぞれ業務 A データ 51、業務 B データ 52、業務 C データ 53 などのように記憶する。また、センタ記憶装置 20 は、ハードディスク装置 50 に記憶した情報と同じ情報 (以下、コピーデータという) を通信回線 5、6 へ出力するためのポート 22、23 を備える。図 1 の例では、コピーデータは、ポート 22 から通信回線 5 へ出力されて、通信回線 5 を介してリモート記憶装置 30a へ送信される。また、それと同じコピーデータが、ポート 23 から通信回線 6 へ出力されて、通信回線 6 を介してリモート記憶装置 30b へ送信される。

20

【0020】

リモート記憶装置 30a は、ハードディスク装置 50a を備え、さらに、ポート 22a、23a を備える。リモート記憶装置 30a は、センタ記憶装置 20 から通信回線 5 を介して送信されるコピーデータを、ポート 22a を介して受信する。そして、その受信したコピーデータを、業務 A データ 51a、業務 B データ 52a、業務 C データ 53a などのようにセンタコンピュータ 10 の業務処理に対応させてハードディスク装置 50a に記憶する。

30

【0021】

また、同様に、リモート記憶装置 30b は、ハードディスク装置 50b を備え、さらに、ポート 22b、23b を備える。リモート記憶装置 30b は、センタ記憶装置 20 から通信回線 6 を介して送信されるコピーデータを、ポート 22b を介して受信する。そして、その受信したコピーデータを、業務 A データ 51b、業務 B データ 52b、業務 C データ 53b などのようにセンタコンピュータ 10 の業務処理に対応するようにハードディスク装置 50b に記憶する。

40

【0022】

以上のように構成されたリモートコピー記憶装置システム 1 において、センタコンピュータ 10 が業務 A 処理部 100、業務 B 処理部 200、業務 C 処理部 300 などのプログラムを実行し、その実行結果をセンタ記憶装置 20 に書き込むと、そのコピーデータがリモート記憶装置 30a、30b それぞれのハードディスク装置 50a、50b にも書き込まれる。このとき、センタコンピュータ 10 がセンタ記憶装置 20 および通信回線 6、7 を介してリモート記憶装置 30a、30b のハードディスク装置 50a、50b へコピーデータを書き込む方式、つまり、リモートコピーのデータ転送方式には、同期方式と非同期方式の 2 つがある。

【0023】

50

同期方式においては、センタコンピュータ10は、記憶装置に対し書き込み命令を発行したとき、その記憶装置から書き込み完了の報告を待って次の命令を実行する。従って、センタコンピュータ10は、同期方式でリモートコピーを実行しているときには、センタ記憶装置20におけるハードディスク装置50へのデータの書き込みの完了を待つだけでなく、そのリモートコピー先のリモート記憶装置30a, 30bにおけるハードディスク装置50a, 50bへのデータ書き込みの完了をも待つ。

【0024】

一方、非同期方式においては、センタコンピュータ10は、記憶装置に対し書き込み命令を発行したとき、その記憶装置から書き込み完了の報告を待たずに次の命令を実行する。この場合には、センタコンピュータ10の書き込み命令に対し、記憶装置における書き込み動作は、多少遅れてなされることになる。従って、センタコンピュータ10がリモート記憶装置30a, 30bに対し、非同期方式でリモートコピーを実行しているときには、センタ記憶装置20のハードディスク装置50に記憶されているデータと、リモート記憶装置30a, 30bのハードディスク装置50a, 50bに記憶されているデータとの間には相違を生じる。

10

【0025】

一般に、リモートコピーにおいては、付近地データセンタ3のリモート記憶装置30aに対しては、同期方式のデータ転送が行われ、遠隔地データセンタ4のリモート記憶装置30bに対しては、非同期方式のデータ転送が行われる。そこで、本実施形態においては、特に断らない限り、同様の方法によってデータ転送方式が行われるものとする。

20

【0026】

ここまで説明したリモートコピー記憶装置システム1の構成は、特許文献1に開示された従来技術に準じたものとなっている。しかしながら、このような従来技術には、次の2つの問題があることを示しておく。

【0027】

(1) センタ記憶装置20からリモート記憶装置30aへ同期方式で送信されるコピーデータ量が通信回線5の通信帯域を超えた場合には、データの送信に待ちが生じるために、リモート記憶装置30aにおける書き込み完了までの時間が長くなる。そのため、リモート記憶装置30aからセンタコンピュータ10へ通知する書き込み完了の報告が遅延する。その結果、センタコンピュータ10の待ち時間が長くなるので、センタコンピュータ10が業務A処理部100、業務B処理部200、業務C処理部300などを処理する能力が低下することになる。

30

【0028】

(2) センタ記憶装置20からリモート記憶装置30bへ非同期方式で送信されるコピーデータ量が通信回線6の通信帯域を超えた場合にも、同様にリモート記憶装置30bにおける書き込み動作が遅延する。この場合には、センタコンピュータ10は、リモート記憶装置30bにおける書き込み動作の完了を待たずに次の命令を実行するので、センタコンピュータ10が記憶装置への書き込み命令を発行した時点、リモート記憶装置30bがコピーデータを受信した時点、ハードディスク装置50bへの書き込み動作が完了した時点などの間に大きな時間差を生じることになる。このようなとき、地震などの災害によってセンタコンピュータ10に障害が生じた場合には、センタコンピュータ10は、リモート記憶装置30bにおいてどの書き込み命令までのデータが実際にリモート記憶装置30bに書き込まれたかというような情報を保持していないので、場合によっては、障害の復旧が困難になることがある。

40

【0029】

そこで、本実施形態においては、センタ記憶装置20のポート22, 23にその通信負荷量を計測する負荷計測部(図1には図示せず、図3にて図示)を設け、その負荷計測部によりポート22, 23で実際に行われているデータの入出力時間や入出力スループットを計測することにより、通信回線5, 6の通信負荷量を計測する。そして、その通信負荷量が所定の閾値を超えたときには、その通信回線5, 6を介して送信されるコピーデータ

50

の一部を、通信回線 7 を介して迂回送信するように、その送信経路を切り替える。

【 0 0 3 0 】

図 2 は、本発明の実施形態に係るリモートコピー記憶装置システムにおいて、コピーデータの一部を迂回送信するときの送信経路の例を示した図である。

【 0 0 3 1 】

図 2 に示すように、ポート 2 2 で通信負荷量が所定の閾値を超えたときには、業務 A データ 5 1 および業務 B データ 5 2 は、それ以前と同様に、ポート 2 2、通信回線 5、ポート 2 2 a を介してリモート記憶装置 3 0 a へ送信されるが、業務 C データ 5 3 については、この経路による送信が停止される。このとき、一方では、業務 C データ 5 3 は、ポート 2 3、通信回線 6、ポート 2 3 b を介してリモート記憶装置 3 0 b へ送信されている。そこで、リモート記憶装置 3 0 b は、その受信した業務 C データ 5 3 をハードディスク装置 5 0 b へ書き込むとともに、ポート 2 2 b、通信回線 7、ポート 2 3 a を介してリモート記憶装置 3 0 a へ送信する。すなわち、業務 C データ 5 3 は、センタ記憶装置 2 0 からリモート記憶装置 3 0 b へ送信され、そのリモート記憶装置 3 0 b からリモート記憶装置 3 0 a へ送信される。

10

【 0 0 3 2 】

このようにして、リモート記憶装置 3 0 a は、そのハードディスク装置 5 0 a に業務 A データ 5 1 a、業務 B データ 5 2 a、業務 C データ 5 3 a を記憶することができる。この場合には、ポート 2 2 および通信回線 5 における通信負荷が軽減されるので、センタコンピュータ 1 0 が業務処理 1 0 0、2 0 0、3 0 0 を処理する能力の低下を防止することができる。

20

【 0 0 3 3 】

同様に、ポート 2 3 で通信負荷量が所定の閾値を超えたときには、業務 A データ 5 1 および業務 C データ 5 3 は、それまでと同様に、ポート 2 3、通信回線 6、ポート 2 3 b を介してリモート記憶装置 3 0 b へ送信されるが、業務 B データ 5 2 については、この経路による送信が停止される。このとき、一方では、業務 B データ 5 2 は、ポート 2 2、通信回線 5、ポート 2 2 a を介してリモート記憶装置 3 0 a へ送信されている。そこで、リモート記憶装置 3 0 a は、その受信した業務 B データ 5 2 をハードディスク装置 5 0 a へ書き込むとともに、ポート 2 3 a、通信回線 7、ポート 2 2 b を介してリモート記憶装置 3 0 b へ送信する。すなわち、業務 B データ 5 2 は、センタ記憶装置 2 0 からリモート記憶装置 3 0 a へ送信され、そのリモート記憶装置 3 0 a からリモート記憶装置 3 0 b へ送信される。

30

【 0 0 3 4 】

このようにして、リモート記憶装置 3 0 b は、そのハードディスク装置 5 0 b に業務 A データ 5 1 b、業務 B データ 5 2 b、業務 C データ 5 3 b を記憶することができる。この場合には、ポート 2 3 および通信回線 6 における通信負荷が軽減されるので、業務 A データ 5 1 および業務 C データ 5 3 については、センタコンピュータ 1 0 による記憶装置への書き込み命令発行時点とハードディスク装置 5 0 b への書き込み動作完了時点との時間差を短縮することができる。しかしながら、業務 B データ 5 2 については、リモート記憶装置 3 0 a を余分に経由して送信されるので、その時間差はむしろ大きくなってしまふ。この問題は、前記した従来技術の問題の (2) と同じ問題である。

40

【 0 0 3 5 】

そこで、本実施形態においては、各記憶装置 2 0、3 0 a、3 0 b が、それぞれのハードディスク装置 5 0、5 0 a、5 0 b に書き込んだデータについての履歴情報を記録するジャーナル (図 1、図 2 には図示せず、図 3 にて図示) を備えるようにした。そして、そのジャーナルを用いることによって、センタコンピュータ 1 0 における業務処理を実行中であっても、そのコピーデータの順序性を損なうことなく、図 2 に示したようなコピーデータの転送経路の切り替えを行うことができるようにした。これは、同時に、コンピュータがいつ停止しても、各記憶装置 2 0、3 0 a、3 0 b 間でジャーナルの同期化を行うことができることを意味するので、障害からの復旧に際しても、データを喪失することなく

50

業務処理を再開することができるようになる。なお、これらの詳細については後記する。

【 0 0 3 6 】

図 3 は、本発明の実施形態に係るリモートコピー記憶装置システムの詳細な構成を示した図である。以下、リモートコピー記憶装置システム 1 の詳細な説明を行う。ただし、図 1 および図 2 の説明と重複する説明については、その一部を省略する。

【 0 0 3 7 】

図 3 に示すように、リモートコピー記憶装置システム 1 は、センタコンピュータ 1 0 およびセンタ記憶装置 2 0 が設置された業務センタ 2 と、リモート記憶装置 3 0 a , 3 0 b がそれぞれ設置された付近地データセンタ 3 および遠隔地データセンタ 4 とを含んで構成される。

10

【 0 0 3 8 】

センタコンピュータ 1 0 は、CPU (Central Processing Unit) 1 3、メモリ 1 2、負荷計測部 1 4、入出力 I F (Interface) 部 1 5 を含んで構成される。ここで、メモリ 1 2 は、コンピュータのいわゆる主記憶装置であり、半導体メモリなどによって構成される。そのメモリ 1 2 には、センタコンピュータ 1 0 が所定の業務を処理するための業務 A 処理部 1 0 0、業務 B 処理部 2 0 0、業務 C 処理部 3 0 0 などのプログラムが格納されている。さらに、メモリ 1 2 には、図 2 で概略を説明した経路切替処理部 8 0 0 のプログラムが格納され、加えて、その経路切替処理部 8 0 0 のプログラムを実行するときを使用される負荷管理テーブル 4 0 0、装置管理テーブル 5 0 0、業務プライオリティ管理テーブル 6 0 0 などが設けられる。

20

【 0 0 3 9 】

メモリ 1 2 に格納された業務 A 処理部 1 0 0、業務 B 処理部 2 0 0、業務 C 処理部 3 0 0、経路切替処理部 8 0 0 などのプログラムは、CPU 1 3 によって実行され、所定の機能が実現される。経路切替処理部 8 0 0 によって実現される機能およびその処理の内容については後記する。

【 0 0 4 0 】

入出力 I F 部 1 5 は、通信回線 8 , 9 を介してセンタ記憶装置 2 0 や付近地データセンタ 3 のリモート記憶装置 3 0 a などに接続され、それらの記憶装置 2 0 , 3 0 a に対して読み書きするデータを入出力するとともに、制御データなどを出力する。また、負荷計測部 1 4 は、入出力 I F 部 1 5 に付属するように設けられ、入出力 I F 部 1 5 における入出力動作の所要時間や単位時間当たりに行われる入出力動作の回数 (入出力スループット) など入出力 I F 部 1 5 の入出力に係る負荷情報を計測する。

30

【 0 0 4 1 】

センタ記憶装置 2 0 は、制御 CPU 6 0、メモリ 4 0、ハードディスク装置 5 0、ポート 2 1 , 2 2 , 2 3、負荷計測部 1 4 a , 1 4 b , 1 4 c などを含んで構成される。ここで、メモリ 4 0 は、半導体メモリなどによって構成され、センタ記憶装置 2 0 自体を制御するプログラムのほか、コピー制御処理部 4 1、入出力処理部 4 2 などのプログラムが格納される。これらのプログラムは、制御 CPU 6 0 によって実行され、センタ記憶装置 2 0 に含まれる所定の機能が実現される。

【 0 0 4 2 】

40

ハードディスク装置 5 0 は、しばしば複数の (多数の) ハードディスクドライブで構成された大容量の記憶装置である。そのハードディスク装置 5 0 は、センタコンピュータ 1 0 が業務 A 処理部 1 0 0、業務 B 処理部 2 0 0、業務 C 処理部 3 0 0 などのプログラムを実行した結果のデータを、業務 A データ 5 1、業務 B データ 5 2、業務 C データ 5 3 などのようにそれぞれのプログラムに対応させて記憶する。さらに、ハードディスク装置 5 0 は、業務 A データ 5 1、業務 B データ 5 2、業務 C データ 5 3 などのデータが書き込まれたとき、その書き込みに係る履歴データをそれぞれジャーナル A 5 4、ジャーナル B 5 5、ジャーナル C 5 6 などのように記憶する。

【 0 0 4 3 】

ポート 2 1 , 2 2 , 2 3 は、センタコンピュータ 1 0 の入出力 I F 部 1 5 との通信回線

50

8 またはリモート記憶装置 30 a , 30 b とをつなぐ通信回線 5 , 6 に接続され、それらの通信回線 8 , 5 , 6 を介して、センタコンピュータ 10、リモート記憶装置 30 a , 30 b との間でデータの送受信を行う。また、負荷計測部 14 a , 14 b , 14 c は、ポート 21 , 22 , 23 それぞれに付属するように設けられ、それぞれのポート 21 , 22 , 23 における入出力動作のスループットなど入出力に係る負荷情報を計測する。

【0044】

付近地データセンタ 3 に設置されたリモート記憶装置 30 a および遠隔地データセンタ 4 に設置されたリモート記憶装置 30 b については、図 3 にその概略構成を示し、図 4 にその詳細構成を示す。図 3 から分かるように、リモート記憶装置 30 a とリモート記憶装置 30 b とは、その基本構成は同じである。また、センタ記憶装置 20 とほとんど同じ構成であるが、リモート記憶装置 30 a , 30 b の場合には、ポート 21 a , 22 a , 23 a , 21 b , 22 b , 23 b に負荷計測部が付属しない。

10

【0045】

図 4 は、本実施形態に係るリモート記憶装置の構成を示した図である。リモート記憶装置 30 a は、制御 CPU 60 a、メモリ 40 a、ハードディスク装置 50 a、ポート 21 a , 22 a , 23 a などを含んで構成される。ここで、メモリ 40 a は、半導体メモリなどによって構成され、リモート記憶装置 30 a 自体を制御するプログラムのほか、コピー制御処理部 41 a、入出力処理部 42 a などのプログラムが格納される。これらのプログラムは、制御 CPU 60 a によって実行され、リモート記憶装置 30 a に含まれる所定の機能が実現される。

20

【0046】

ハードディスク装置 50 a は、センタコンピュータ 10 が業務 A 処理部 100、業務 B 処理部 200、業務 C 処理部 300 などのプログラムを実行した結果得られ、センタ記憶装置 20 が送信するデータを、業務 A データ 51 a、業務 B データ 52 a、業務 C データ 53 a などのようにそのプログラムに対応させて記憶する。さらに、ハードディスク装置 50 a は、業務 A データ 51 a、業務 B データ 52 a、業務 C データ 53 a などのデータが書き込まれたとき、その書き込みに係る履歴データをそれぞれジャーナル A 54 a、ジャーナル B 55 a、ジャーナル C 56 a などのように記憶する。

【0047】

ポート 21 a , 22 a , 23 a は、センタコンピュータ 10、センタ記憶装置 20 または他のリモート記憶装置 30 b とをつなぐ通信回線 9 , 5 , 7 などに接続され、それらの通信回線を介して、センタコンピュータ 10、センタ記憶装置 20、他のリモート記憶装置 30 b との間でデータの送受信を行う。

30

【0048】

なお、リモート記憶装置 30 b の構成は、基本的に、以上に説明したリモート記憶装置 30 a の構成と同じなので、その説明を省略する。

【0049】

続いて、図 5 ~ 図 11 および図 3 を参照して、センタコンピュータ 10 の経路切替処理部 800 のプログラムによって実現される機能およびその処理フローの詳細について説明する。ここで、図 5、図 6 および図 7 は、それぞれ、経路切替処理部 800 によって使用される負荷管理テーブル 400、装置管理テーブル 500 および業務プライオリティ管理テーブル 600 の構成の例を示した図である。

40

【0050】

図 5 に示すように、負荷管理テーブル 400 は、ハードウェア識別子、入出力時間、入出力スループット、転送帯域、業務識別子、コピータイプなどのフィールドによって構成される。

【0051】

ここで、ハードウェア識別子は、負荷計測対象の装置にユニークに付された識別情報である。本実施形態では、負荷計測はセンタ記憶装置 20 の負荷計測部 14 a , 14 b , 14 c で行うほか、センタコンピュータ 10 の負荷計測部 14 でも行う。負荷計測を負荷計

50

測部 1 4 a , 1 4 b , 1 4 c で行う場合には、その負荷計測対象の装置は、ポート 2 1 , 2 2 , 2 3 である。また、負荷計測を負荷計測部 1 4 で行う場合には、その負荷計測は、業務 A 処理部 1 0 0、業務 B 処理部 2 0 0、業務 C 処理部 3 0 0 など業務処理ごとに分けて行う。そこで、この場合には、各業務処理部 1 0 0 , 2 0 0 , 3 0 0 によるデータの格納先をそのハードウェア識別子に対応させる。

【 0 0 5 2 】

また、入出力時間は、入出力 I F 部 1 5 またはポート 2 1 , 2 2 , 2 3 における入出力動作の 1 回当たりの入出力所要時間である。入出力スループットは、入出力 I F 部 1 5 またはポート 2 1 , 2 2 , 2 3 における単位時間当たりの入出力実行回数である。転送帯域は、入出力 I F 部 1 5 またはポート 2 1 , 2 2 , 2 3 における単位時間当たりのデータ転送量である。これらの情報は、いずれもポート 2 1 , 2 2 , 2 3 に接続される通信回線 8 , 5 , 6 の通信負荷に直接に影響を受けるデータである。すなわち、通信回線 8 , 5 , 6 の通信負荷が大きくなると（通信回線が混雑すると）、入出力の所要時間は長くなり、入出力のスループットは低下し、転送帯域は小さくなる。そこで、本実施形態では、入出力時間、入出力スループット、転送帯域の少なくとも一つを含むデータによって、通信負荷を表わすものとする。ただし、入出力時間は大きいければ大きいほど、通信負荷も大きくなるが、入出力スループットおよび転送帯域は小さければ小さいほど通信負荷が大きくなる。なお、センタコンピュータ 1 0 は、所定時間ごとに（例えば、1 秒とか 1 分ごとに）負荷計測部 1 4 , 1 4 a , 1 4 b , 1 4 c からこれらの情報を取得し、負荷管理テーブル 4 0 0 に格納する。

【 0 0 5 3 】

また、業務識別子は、ハードウェア識別子のフィールドで指定される装置が関与する業務処理を指示する情報である。また、コピータイプは、同じくその装置がどのようなコピーデータの転送（同期方式 / 非同期方式、付近地接続 / 遠隔地接続 / 別経路接続など）で使用されるかを示す情報である。なお、これら情報はオペレータなどによってあらかじめ設定される。

【 0 0 5 4 】

次に、図 6 に示すように、装置管理テーブル 5 0 0 は、ハードウェア識別子、ジャーナル識別子、付近地接続情報、遠隔地接続情報、別経路転送情報、状態情報などのフィールドによって構成される。

【 0 0 5 5 】

ここで、ハードウェア識別子は、図 5 で説明したハードウェア識別子と同じものである。ジャーナル識別子は、そのハードウェア識別子のフィールドで指定される装置が関与するジャーナルを指示する情報である。すなわち、ジャーナル識別子によって、どのデータのコピーが行われているか判断することができる。

【 0 0 5 6 】

また、付近地接続情報は、センタ記憶装置 2 0 が前記ハードウェア識別子のフィールドで指定される装置を介して、付近地データセンタ 3 のリモート記憶装置 3 0 a と接続されている状態を示す情報である。また、遠隔地接続情報は、センタ記憶装置 2 0 が前記ハードウェア識別子のフィールドで指定される装置を介して、遠隔地データセンタ 4 のリモート記憶装置 3 0 b と接続されている状態を示す情報である。また、別経路接続情報は、付近地データセンタ 3 のリモート記憶装置 3 0 a と遠隔地データセンタ 4 のリモート記憶装置 3 0 b とを接続する状態を示す情報である。なお、その接続状態としては、「転送中」「一時停止中」「未転送」「未接続」などがある。また、状態情報は、当該装置の稼働状態を示す情報で、装置は存在するが転送業務を停止している場合には、「無効」を表示する。

【 0 0 5 7 】

次に、図 7 に示すように、業務プライオリティ管理テーブル 6 0 0 は、業務プライオリティ、業務識別子、ハードウェア識別子、負荷上限値、負荷下限値、コピータイプなどのフィールドによって構成される。なお、このテーブルに格納する情報は、ユーザの指示に

10

20

30

40

50

基づき、あらかじめオペレータなどによって設定されるものとする。

【 0 0 5 8 】

ここで、業務識別子は、業務 A 処理部 1 0 0、業務 B 処理部 2 0 0、業務 C 処理部 3 0 0 などの各業務処理を識別する情報である。業務プライオリティは、各業務処理の処理の優先度を示す情報である。図 7 の例では、「 A 」は優先度が高く、「 C 」は優先度が低いとしている。

【 0 0 5 9 】

負荷上限値は、通信負荷の上限の閾値であり、通信負荷がこの閾値を超えた場合には、通信負荷を減ずるために、コピータイプの状態を変更する。本実施形態では、ハードウェア識別子で指定される装置で転送される業務識別子で指定される業務のデータについて別経路で転送するように変更する。また、負荷下限値は、通信負荷の下限の閾値であり、通信負荷がこの閾値を超えて小さくなった場合には、いったん変更したコピータイプをもとに戻す。また、コピータイプのフィールドでは、コピータイプを変更する対象の部位を指定する。なお、通信負荷を入出力スループットまたは転送帯域で表現する場合には、小さい値が負荷上限値となり、大きい値が負荷下限値となる。

【 0 0 6 0 】

次に、図 8 および図 9 を参照して、ジャーナルの同期化について説明する。ここで、図 8 は、本実施形態に係る記憶装置におけるデータ書き込みの履歴情報であるジャーナルの構成の例を示した図である。

【 0 0 6 1 】

図 8 に示すように、ジャーナル 7 0 0 は、ジャーナル識別子 7 0 1 とジャーナルデータ 7 0 2 とによって構成される。ここで、ジャーナルデータ 7 0 2 のフィールドには、ハードディスク装置 5 0、5 0 a、5 0 b へのデータ書き込みの履歴情報（ログ情報）が記録される。ジャーナル識別子 7 0 1 は、そのジャーナルデータ 7 0 2 をユニークに識別するための識別情報である。また、ジャーナル識別子 7 0 1 は、同一の記憶装置 2 0、3 0 a、3 0 b 内において、ハードディスク装置 5 0、5 0 a、5 0 b へ書き込まれたデータの順序性を確保する情報でもある。従って、ジャーナル識別子 7 0 1 としては、通しの連続番号や年月日時分秒の情報など順序性を有する識別情報が用いられる。

【 0 0 6 2 】

なお、ジャーナル識別子 7 0 1 は、通常、センタコンピュータ 1 0 またはセンタ記憶装置 2 0 で順序性のある識別情報が付与され、その識別情報が付与されたジャーナル識別子 7 0 1 は、コピーデータに付してコピー元のセンタ記憶装置 2 0 からコピー先のリモート記憶装置 3 0 a、3 0 b へ送信される。そして、異なる複数の記憶装置 2 0、3 0 a、3 0 b 間において、ジャーナルデータ 7 0 2 を同定するのに使用される。

【 0 0 6 3 】

ジャーナル 7 0 0 は、前記したように、各記憶装置 2 0、3 0 a、3 0 b において、業務処理 1 0 0、2 0 0、3 0 0 ごとに作成され、各々のハードディスク装置 5 0 a、5 0 b、5 0 c にジャーナル A 5 4、5 4 a、...、ジャーナル B 5 4、5 4 a、...、ジャーナル B 5 4、5 4 a、...、として記憶される（図 3 参照）。このとき、それぞれ記憶されるジャーナルデータの数は、多くても 1 0 ~ 2 0 程度である。

【 0 0 6 4 】

ここで、コピーデータが付近地データセンタ 3 のリモート記憶装置 3 0 a へ同期方式で送信され、遠隔地データセンタ 4 のリモート記憶装置 3 0 b へ非同期方式で送信される場合を想定する。一般に、リモート記憶装置 3 0 b におけるコピーデータのハードディスク装置 5 0 b への書き込みが遅延する。従って、この場合には、ある時点をもって両者のハードディスク装置 5 0 a、5 0 b を比較すると、その内容は一致しない。そのため、災害などによるセンタコンピュータ 1 0 の突然の停止や、次に示すようなコピー経路変更のための業務処理の停止があった後に、その処理を再開するためには、ハードディスク装置 5 0 a、5 0 b の内容を同期化しておく必要がある。ジャーナル 7 0 0 は、その同期化のために使用される。

10

20

30

40

50

【 0 0 6 5 】

図 9 は、本実施形態における記憶装置において、ある業務処理を一時停止したとき、各記憶装置に記憶されているジャーナルの状況の例を示した図である。図 9 に示すように、業務センタ 2 のセンタ記憶装置 2 0 とコピーデータが同期方式で送信される付近地データセンタ 3 のリモート記憶装置 3 0 a との間では、その各々に記憶されているジャーナル 9 0 1 とジャーナル 9 0 2 とは一致している。すなわち、業務センタ 2 のジャーナル 9 0 1 と付近地データセンタ 3 のジャーナル 9 0 2 とは、いずれも、ジャーナル識別子が「 n 」から「 m 」までのジャーナルを有している。ここで、「 n 」は「 m 」よりも新しいことを示す識別情報であるとする。また、図 9 で、有効なジャーナルとは、各記憶装置 2 0 , 3 0 a , 3 0 b が有効な情報として記憶している状態にあるジャーナルである。

10

【 0 0 6 6 】

一方、業務センタ 2 のセンタ記憶装置 2 0 とコピーデータが非同期方式で送信される遠隔地データセンタ 4 のリモート記憶装置 3 0 b との間では、その各々に記憶されているジャーナル 9 0 1 とジャーナル 9 0 3 とは一致していない。すなわち、業務センタ 2 のジャーナル 9 0 1 は、ジャーナル識別子が「 n 」から「 m 」までのジャーナルを有しているが、遠隔地データセンタ 4 のジャーナル 9 0 3 は、ジャーナル識別子が「 q 」から「 p 」までのジャーナルを有している。ここで、「 q 」は「 p 」よりも新しいことを示す識別情報であるとする。

【 0 0 6 7 】

このとき、「 q 」が「 m 」よりも新しいことを示す識別情報である場合（「 q 」と「 m 」とが同じである場合を含む）には、業務センタ 2 のジャーナル 9 0 1 と遠隔地データセンタ 4 のジャーナル 9 0 3 とは共通のジャーナルデータ 7 0 2 を有する。この場合には、業務センタ 2 のセンタ記憶装置 2 0 または付近地データセンタ 3 のリモート記憶装置 3 0 a から遠隔地データセンタ 4 のリモート記憶装置 3 0 b へ、ジャーナル識別子が「 n 」のジャーナルデータからジャーナル識別子が「 q 」より 1 つ新しいジャーナル識別子までのジャーナルデータを送信することにより、遠隔地データセンタ 4 のジャーナル 9 0 3 を業務センタ 2 のジャーナル 9 0 1 に同期化することができる。

20

【 0 0 6 8 】

一方、「 q 」が「 m 」よりも古いことを示す識別情報であった場合には、業務センタ 2 のジャーナル 9 0 1 と遠隔地データセンタ 4 のジャーナル 9 0 3 とは共通のジャーナルデータ 7 0 2 を有しない。この場合には、ジャーナル識別子が「 m 」のジャーナルデータとジャーナル識別子が「 q 」のジャーナルデータとの間のデータの連続性を保証することができないので、遠隔地データセンタ 4 のジャーナル 9 0 3 を業務センタ 2 のジャーナル 9 0 1 に同期化することはできない。

30

【 0 0 6 9 】

次に、図 3、図 1 0 および図 1 1 を参照して、センタコンピュータ 1 0 がコピーデータの送信経路を切り替えることによって、大きな通信負荷がかかった通信経路の通信負荷を軽減する手順について説明する。ここで、図 1 0 は、本実施形態に係るセンタコンピュータ 1 0 における経路切替処理部 8 0 0 の処理フローの例を示した図である。

【 0 0 7 0 】

図 1 0 において、センタコンピュータ 1 0 は、所定の時間ごとに（例えば、1 秒とか 1 分ごとに）、負荷計測部 1 4 , 1 4 a , 1 4 b , 1 4 c から負荷情報を取得し、その取得した負荷情報によって負荷管理テーブル 4 0 0（図 5 参照）の情報を更新する（ステップ S 1 1）。このとき取得する負荷情報は、入出力 I F 部 1 5、ポート 2 1 , 2 2 , 2 3 のそれぞれにおける入出力時間、入出力スループットおよび転送帯域の少なくとも 1 つを含む情報である。

40

【 0 0 7 1 】

次に、センタコンピュータ 1 0 は、負荷管理テーブル 4 0 0 と業務プライオリティ管理テーブル 6 0 0 とを参照して、負荷が上限値を超えたハードウェアを抽出する（ステップ S 1 2）。ここで、負荷は、負荷管理テーブル 4 0 0 における入出力時間、入出力スル-

50

ットまたは転送帯域の計測データである。また、上限値は、業務プライオリティ管理テーブル600で指定される値である。また、ハードウェアとは、負荷管理テーブル400および業務プライオリティ管理テーブル600のハードウェア識別子によって指示される構成要素である。

【0072】

次に、センタコンピュータ10は、業務プライオリティ管理テーブル600を参照して、ステップS12で抽出されたハードウェアが関与する業務処理のうち、最もプライオリティが低い業務処理を抽出する（ステップS13）。そして、抽出した業務処理の実行、つまり、そのリモートコピーを一時停止する（ステップS14）。

【0073】

次に、センタコンピュータ10は、装置管理テーブル500を参照して、当該ハードウェア（ステップS13で抽出されたハードウェア）の当該業務処理（ステップS14で抽出された業務処理）について転送可能な別経路を抽出する（ステップS15）。本実施形態の場合には、別経路とは、センタ記憶装置20のポート21, 22, 23を経由しない経路であって、リモート記憶装置30a, 30b同士をつなぐ経路を指す。装置管理テーブル500において、当該ハードウェアの別経路接続情報が「未転送」の場合には、別経路で転送可能であると判断する。なお、転送可能な別経路がない場合には、ステップS13に戻って次にプライオリティが低い業務処理を抽出して、ステップS14、ステップS15を実行する。

【0074】

次に、センタコンピュータ10は、ステップS15で抽出された別経路のコピー元およびコピー先となるリモート記憶装置30a, 30bから、当該業務処理（ステップS14で抽出された業務処理）に係るジャーナルを参照して、そのジャーナルを比較し、ジャーナルの同期化が可能な場合、そのジャーナルを同期化する（ステップS16）。この処理の詳細な処理フローについては、図11を参照してこの後に説明するが、その基本的な考え方については、図9を参照して説明した通りである。

【0075】

センタコンピュータ10は、ステップS16でジャーナルの同期化を行った場合には、当該業務処理（ステップS13で抽出された業務処理）のリモートコピーの経路を別経路（ステップS15で抽出された別経路）に切り替え、ステップS14で一時停止していた業務処理についてのリモートコピーを再開する（ステップS17）。なお、ジャーナルの同期化を行うことができなかつた場合には、リモートコピーの経路を切り替えることなく、その業務処理についてリモートコピーを再開する。

【0076】

なお、以上の説明においては、センタコンピュータ10がリモートコピーの経路を別経路に切り替える（ステップS17）としているが、実際には、センタコンピュータ10は、センタ記憶装置20またはリモート記憶装置30a（または30b）に対し、切り替え後のコピー元およびコピー先となるリモート記憶装置30a（または30b）の情報と切り替え対象の業務処理の情報とを付して、リモートコピーの経路の経路切り替えを指示する通知を送信するだけでよい。すなわち、その経路切り替えを指示する通知を受信したセンタ記憶装置20またはリモート記憶装置30a（または30b）が、どの業務処理のコピーデータを、どのポートを使用し、どの通信回線を経由し、どのリモート記憶装置へ送信するかを決定し、その決定した経路に基づき、コピーデータを送信する。

【0077】

また、同様に、ジャーナルの同期化（ステップS17）についても、センタコンピュータ10は、同期化の可否判定を行い、同期化が可能な場合に、同期化の対象となるセンタ記憶装置20またはリモート記憶装置30a（または30b）に対し、ジャーナルの同期化を指示する通知を送信するだけでよい。すなわち、ジャーナルの同期化を指示する通知を受信したセンタ記憶装置20またはリモート記憶装置30a（または30b）が、その相互間において同期化に必要なジャーナルデータの送受信を行い、両者のジャーナルを同

10

20

30

40

50

期化する。

【 0 0 7 8 】

以上のようにして、センタコンピュータ10は、ポート21, 22, 23などで計測される通信負荷が所定の上限値(閾値)を超えた場合には、プライオリティが低い業務処理のデータについて、そのリモートコピーの経路を別経路に切り替えることができる。すなわち、そのプライオリティが低い業務処理のコピーデータの送信を別経路に切り替えることができるので、通信負荷が所定の上限値(閾値)を超えていた元の通信経路の通信負荷を軽減することができる。なお、通信経路が具体的にどのように変更されるかについては、図2を用いて説明したとおりである。

【 0 0 7 9 】

また、元の通信経路の通信負荷が業務プライオリティ管理テーブル600の負荷下限値を超えて軽減された場合には、前記した別経路に切り替えられていた業務処理のコピーデータの送信経路を元に戻すことができる。その処理フローは、図10に示した処理フローと似たものであるので、その説明を省略する。

【 0 0 8 0 】

図11は、本実施形態に係るセンタコンピュータ10における経路切替処理部800に含まれるジャーナルの同期化処理(ステップS16)の詳細な処理フローの例を示した図である。この処理フローの基本的な考え方は、図9の説明に基づく。本実施形態では、別経路とは、付近地データセンタ3のリモート記憶装置30aと遠隔地データセンタ4のリモート記憶装置30bとをつなぐ通信回線7を指す。そこで、図11の処理フローは、当初のリモートコピーが、付近地データセンタ3のリモート記憶装置30aへは同期方式で行われ、遠隔地データセンタ4のリモート記憶装置30bへは非同期方式で行われたことを前提として記載している。

【 0 0 8 1 】

図11において、センタコンピュータ10は、まず、付近地データセンタ3のリモート記憶装置30aから当該業務処理(図10ステップS13で抽出した業務処理)の業務データに係るジャーナルの最終のジャーナル識別子(ここでは、これを「m」と表記する)を取得する(ステップS21)。そして、同様に、遠隔地データセンタ4のリモート記憶装置30bから当該業務処理の業務データに係るジャーナルの最新のジャーナル識別子(ここでは、これを「q」と表記する)を取得する(ステップS22)。

【 0 0 8 2 】

次に、センタコンピュータ10は、ジャーナル識別子「m」とジャーナル識別子「q」との時間的な順序関係を比較し、ジャーナル識別子「q」がジャーナル識別子「m」と同じ、または、新しいものであるか否かを判定する(ステップS23)。その結果、ジャーナル識別子「q」がジャーナル識別子「m」と同じ、または、新しいものであった場合には(ステップS23でYes)、ジャーナルの同期化が可能であるので、センタコンピュータ10は、付近地データセンタ3のリモート記憶装置30aおよび遠隔地データセンタ4のリモート記憶装置30bに対し、ジャーナルの同期化を指示する通知を送信する。その通知を受信した付近地データセンタ3のリモート記憶装置30aは、遠隔地データセンタ4のリモート記憶装置30bが未受信のジャーナルデータを遠隔地データセンタ4のリモート記憶装置30bへ送信する(ステップS24)。遠隔地データセンタ4は、そのジャーナルデータを受信することによって、当該業務処理の業務データに係るジャーナルを付近地データセンタ3のものと同期化することができる。

【 0 0 8 3 】

一方、ジャーナル識別子「q」がジャーナル識別子「m」よりも古いものであった場合には(ステップS23でNo)、ジャーナルの同期化ができないことを示すエラー応答を返し(ステップS25)、処理を終了する。

【 0 0 8 4 】

以上、本実施形態によれば、センタ記憶装置20とリモート記憶装置30a, 30bとをつなぐ通信回線の通信負荷が過負荷になる前に、リモートコピーが行われている途中で

10

20

30

40

50

あっても、そのコピーデータの順序性を損なうことなく、その間で行われるリモートコピーデータの送信の一部を別経路、つまり、リモート記憶装置30a, 30b間の通信回線を介した経路による送信に切り替えることができる。従って、通信回線の過負荷によって生じるセンタコンピュータ10における業務処理の低下を防止することができる。

【図面の簡単な説明】

【0085】

【図1】本発明の実施形態に係るリモートコピー記憶装置システムの概略構成の例、および、通常状態におけるコピーデータの送信経路の例を示した図である。

【図2】本発明の実施形態に係るリモートコピー記憶装置システムにおいて、コピーデータの一部を迂回送信するときの送信経路の例を示した図である。

10

【図3】本発明の実施形態に係るリモートコピー記憶装置システムの詳細な構成を示した図である。

【図4】本実施形態に係るリモート記憶装置の構成を示した図である。

【図5】本発明の実施形態に係るセンタコンピュータにおける負荷管理テーブルの構成の例を示した図である。

【図6】本発明の実施形態に係るセンタコンピュータにおける装置管理テーブル構成の例を示した図である。

【図7】本発明の実施形態に係るセンタコンピュータにおける業務プライオリティ管理テーブルの構成の例を示した図である。

【図8】本実施形態に係る記憶装置におけるデータ書き込みの履歴情報であるジャーナルの構成の例を示した図である。

20

【図9】本実施形態における記憶装置において、ある業務処理のリモートコピーを一時停止したとき、各記憶装置に記憶されているジャーナルの状況の例を示した図である。

【図10】本実施形態に係るセンタコンピュータにおける経路切替処理部の処理フローの例を示した図である。

【図11】本実施形態に係るセンタコンピュータにおける経路切替処理部に含まれるジャーナルの同期化処理の詳細な処理フローの例を示した図である。

【符号の説明】

【0086】

- 1 リモートコピー記憶装置システム
- 2 業務センタ
- 3 付近地データセンタ
- 4 遠隔地データセンタ
- 5, 6, 7, 8, 9 通信回線
- 10 センタコンピュータ
- 12 メモリ
- 13 CPU
- 14, 14a, 14b, 14c 負荷計測部
- 15 入出力IF部
- 20 センタ記憶装置
- 21, 22, 23, 21a, 22a, 23a, 21b, 22b, 23b ポート
- 30a, 30b リモート記憶装置
- 40, 40a メモリ
- 41, 41a コピー制御処理部
- 42, 42a 入出力処理部
- 42a 入出力処理部
- 50, 50a, 50b ハードディスク装置
- 100 業務A処理部
- 200 業務B処理部
- 300 業務C処理部

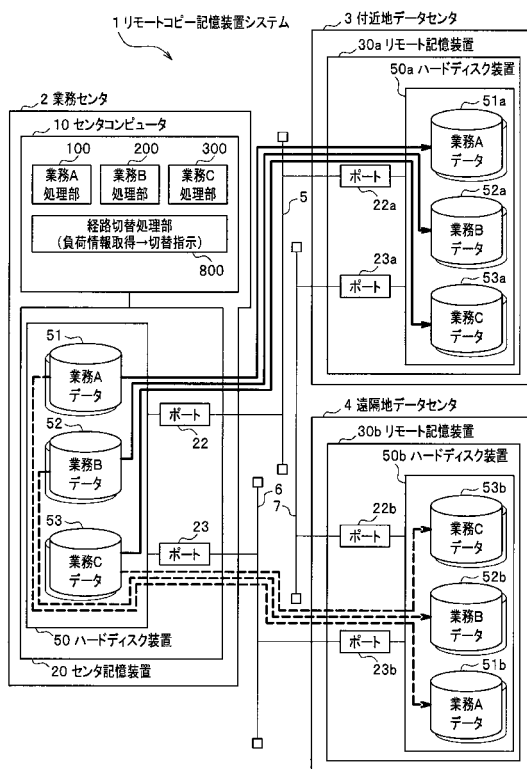
30

40

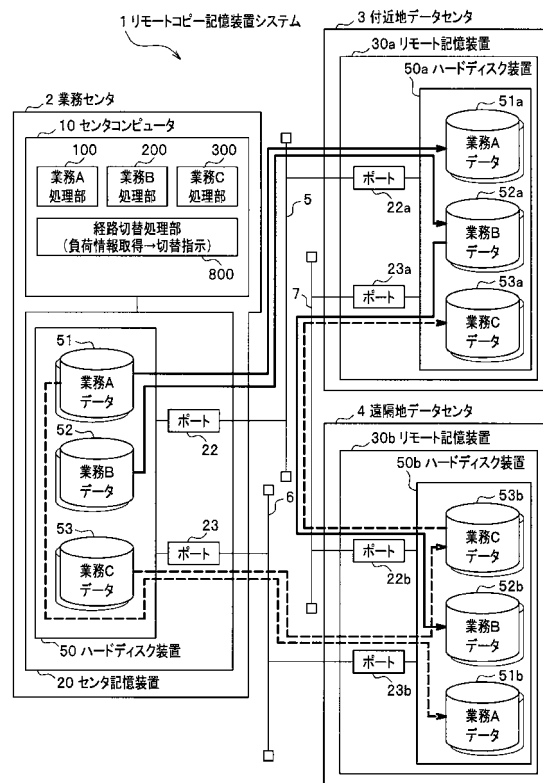
50

- 400 負荷管理テーブル
- 500 装置管理テーブル
- 600 業務プライオリティ管理テーブル
- 700 ジャーナル

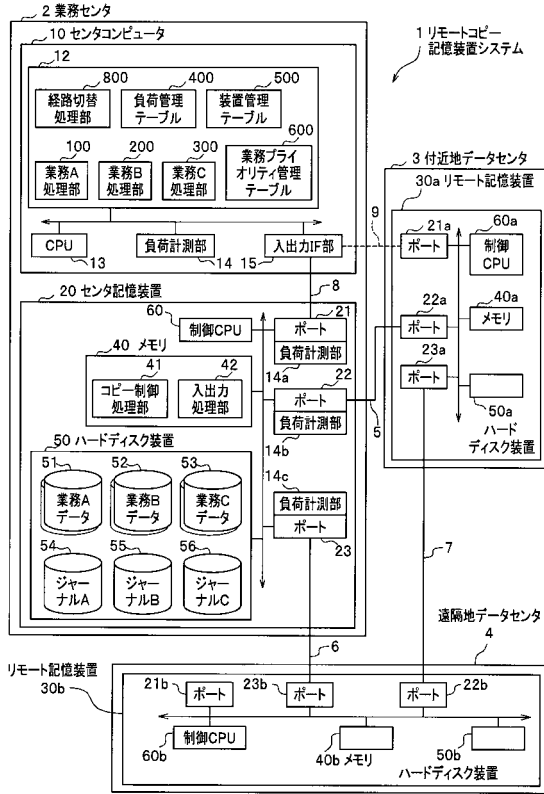
【図1】



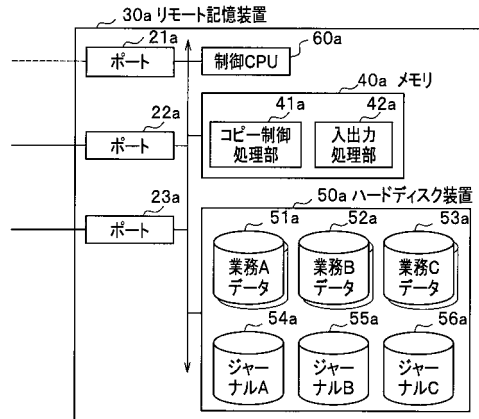
【図2】



【図3】



【図4】

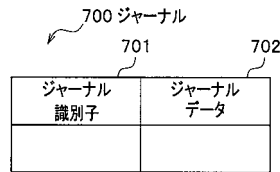


【図5】

400 負荷管理テーブル

ハードウェア識別子	入出力時間	入出力スループット	転送帯域	業務識別子	コピータイプ
業務Aデータ	10ms	100 IOPS	40MB/s	業務A	同期+非同期
業務Bデータ	5ms	200 IOPS	80MB/s	業務B	同期+非同期
業務Cデータ	1ms	1000 IOPS	100MB/s	業務C	同期+非同期
ポート23	10ms	100 IOPS	40MB/s	業務A,B,C	非同期
ポート22	5ms	200 IOPS	80MB/s	業務A,B,C	同期

【図8】

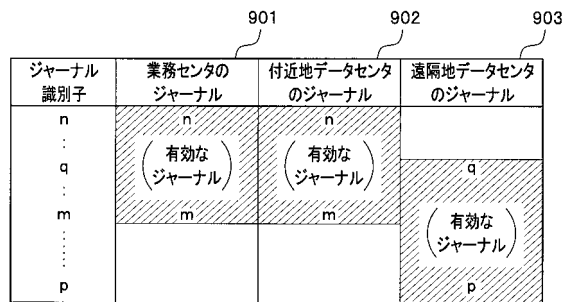


【図6】

500 装置管理テーブル

ハードウェア識別子	ジャーナル識別子	付近地接続情報	遠隔地接続情報	別経路接続情報	状態情報
業務Aデータ	ジャーナルA	転送中	転送中	未転送	有効
ポート23	ジャーナルC	未接続	転送中	未転送	有効
業務Bデータ	ジャーナルB	転送中	転送中	未転送	有効
業務Cデータ	ジャーナルC	転送中	転送中	未転送	有効
ポート22	ジャーナルB	転送中	未接続	未転送	有効

【図9】

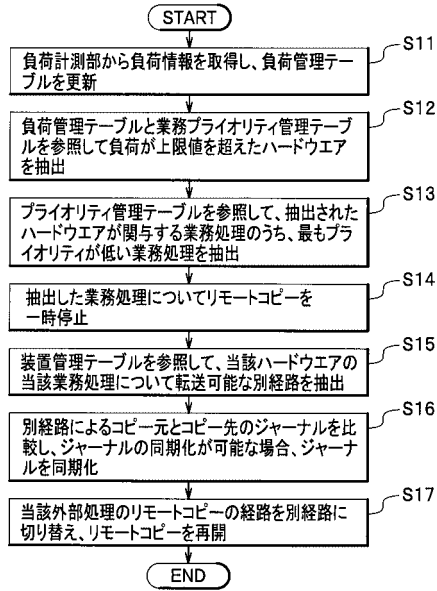


【図7】

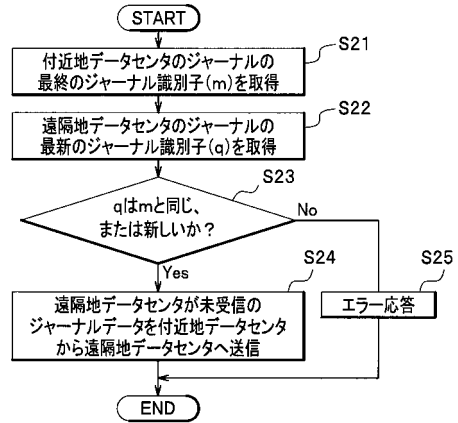
600 業務プライオリティ管理テーブル

業務プライオリティ	業務識別子	ハードウェア識別子	負荷上限値	負荷下限値	コピータイプ
C	業務C	業務Cデータ	50 IOPS	500 IOPS	付近地(同期)
B	業務B	ポート22	8ms	2ms	付近地(同期)
B	業務B	業務Bデータ	70 IOPS	500 IOPS	付近地(同期)
A	業務A	業務Aデータ	150 IOPS	500 IOPS	遠隔地(非同期)
C	業務C	ポート23	8ms	2ms	遠隔地(非同期)

【図10】



【図11】



フロントページの続き

- (56)参考文献 特開2005-122235(JP,A)
特開2003-131917(JP,A)
特開2005-267167(JP,A)
特開2003-122509(JP,A)
特開2003-258903(JP,A)
宮村 崇, MPLSのExplicit Routingを用いた負荷分方式に関する一検討 (SSE2000-54), 電子情報通信学会技術研究報告, 社団法人電子情報通信学会, 2000年 6月23日, 第100巻 第154号, pp.25~30

(58)調査した分野(Int.Cl., DB名)

H04L 12/56
G06F 3/06
G06F 12/00
G06F 13/00