(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2013 年 10 月 3 日(03.10.2013)

WIPO | PCT

(10) 国際公開番号
WO 2013/146808 A1

(51) 国際特許分類:
H04L 12/701 (2013.01) *G06F 13/00* (2006.01)

(21) 国際出願番号: PCT/JP2013/058846

(22) 国際出願日: 2013 年 3 月 26 日(26.03.2013)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2012-074936 2012 年 3 月 28 日(28.03.2012) JP

(71) 出願人 (米国を除く全ての指定国について): 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 Tokyo (JP).

(72) 発明者; および
(71) 出願人 (米国についてのみ): 濱田 欣孝(HAMADA Yoshitaka) [JP/JP]; 〒1088001 東京都港区芝五丁目７番１号 日本電気株式会社内 Tokyo (JP).

(74) 代理人: 工藤 実(KUDOH Minoru); 〒1400013 東京都品川区南大井六丁目２４番１０号カドヤビル６階 Tokyo (JP).

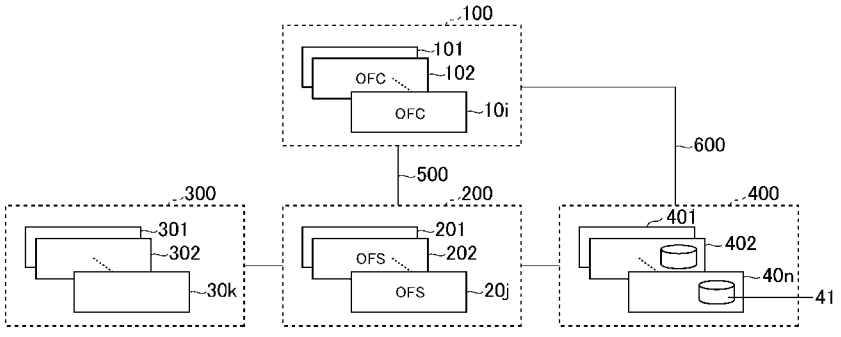
(81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

添付公開書類:
— 国際調査報告（条約第 21 条(3)）

(54) **Title**: COMPUTER SYSTEM AND COMMUNICATION PATH MODIFICATION MEANS

(54) 発明の名称 : コンピュータシステム、及び通信経路変更方法



(57) **Abstract**: This computer system is equipped with a controller, a plurality of switches for performing relay processing of reception packets in accordance with flow entries set by the controller, and a plurality of storage systems. On the basis of resource information obtained from the plurality of storage systems, the controller modifies the flow entries of the plurality of switches and thereby modifies the storage system that will become the access destination of a client terminal. This configuration enables the overall network system to be managed and used without regard to the state each of the plurality of storage systems or the overall network.

(57) 要約:

[続葉有]

WO 2013/146808 A1

本発明によるコンピュータシステムは、コントローラと、コントローラによって設定されたフローエントリに従って、受信パケットの中継処理を行う複数のスイッチと、複数のストレージシステムとを具備する。コントローラは、複数のストレージシステムから取得したリソース情報に基づいて、複数のスイッチのフローエントリを変更することで、クライアント端末のアクセス先となるストレージシステムを変更する。これにより、複数のストレージシステムのそれぞれやネットワーク全体の状況を意識せずに、ネットワークシステム全体を管理又は利用することができる。

明　細　書

発明の名称 ： コンピュータシステム、及び通信経路変更方法

技術分野

[0001] 本発明は、コンピュータシステム、及び通信経路変更に関し、特に、オープンフロー（オープンフロー）技術を利用したコンピュータシステム、及びその通信経路変更方法に関する。

背景技術

[0002] 従来のストレージシステムでは、各ストレージのコントローラ部が、主に配下のストレージシステムのクラスタ制御を行っている。このため、グローバルネームスペースを利用して、クラウド環境のストレージシステムにおけるクラスタ制御を実現する場合、管理者はストレージシステム間の連携機能を各ストレージのコントローラに設定する必要がある。又、ストレージシステムにはその制御に対応した機能が必要である。

[0003] 又、従来のストレージシステムでは、各ストレージシステムと、ＡＤＳ（Ａｃｔｉｖｅ　Ｄｉｒｅｃｔｏｒｙ　Ｓｅｒｖｅｒ）やＬＤＡＰ（Ｌｉｇｈｔｗｅｉｇｈｔ　Ｄｉｒｅｃｔｏｒｙ　Ａｃｃｅｓｓ　Ｐｒｏｔｏｃｏｌ）サーバなどの認証サーバとの連携により、ユーザの当該ストレージシステムへの優先アクセス制御が実現される。

[0004] 一方、クラウド環境のシステムでは、ポリシに基づいて新しいネットワークアプライアンス製品やストレージシステムに継続移行させることで、特定のネットワークアプライアンス製品やストレージシステムの製造中止や、ストレージシステムやネットワークの状況を意識させずに、ユーザに対するサービス提供を維持させる必要がある。このようなシステムは、例えば、各スイッチの転送動作等を外部のコントローラによって一元的に制御するオープンフロー技術によって実現され得る（非特許文献１参照）。

[0005] オープンフローを利用したネットワークスイッチ（以下、オープンフロースイッチ（ＯＦＳ）と称す）は、プロトコル種別やポート番号等の詳細な情

報をフローテーブルに保持し、フローの制御と統計情報の採取を行うことができる。ＯＦＳが保持するフローテーブルは、ＯＦＳとは分離して設けられるコントローラ（以下、オープンフローコントローラ（ＯＦＣ）と称す）によって設定される。ＯＦＣは、ノード間の通信経路の設定や、経路上におけるＯＦＳに対する転送動作（中継動作）等の設定を行う。この際、ＯＦＣは、ＯＦＳが保持するフローテーブルに、フロー（パケットデータ）を特定するルールと、当該フローに対する処理を規定するアクションとを対応付けたフローエントリを設定する。フローテーブルに設定されるエントリの内容については、例えば非特許文献１で規定されている。

[0006] 通信経路上のＯＦＳは、ＯＦＣによって設定されたフローエントリに従って受信パケットデータの転送先を決定し、転送処理を行う。これにより、ネットワーク上のノードは、ＯＦＣによって設定された通信経路を利用して他のノードとの間でパケットデータの送受信が可能となる。すなわち、オープンフローを利用したコンピュータシステムでは、転送処理を行うＯＦＳと分離して設けられたＯＦＣによって、システム全体の通信を一元的に制御及び管理することが可能となる。

[0007] ＯＦＣは、ＯＦＳからの要求に応じて通信経路の算出及び通信経路上のＯＦＳにおけるフローテーブルの更新を行う。詳細には、ＯＦＳが自身のフローテーブルに規定されていないパケットデータを受信した場合、当該パケットデータをＯＦＣに通知する。ＯＦＣは、通知されたパケットデータのヘッダ情報に基づいて転送元及び転送先を特定し通信経路やＯＦＳに設定するフローエントリ（ルール＋アクション）を生成し、各ＯＦＳのフローテーブルを更新する。

[0008] ＯＦＣは、Ｌ１～Ｌ４のヘッダ情報によって規定されるフロー単位で、クライアント端末間の転送を制御できるため、ネットワークを任意に仮想化することができる。これにより、物理構成の制約は緩和され、仮想テナント環境の構築が容易となるとともに、スケールアウトによる初期投資費用を低減することができる。

[0009] 又、ストレージシステムに関する技術が、例えば特開２００２－２０７６２９（特許文献１参照）、特開２００４－２１８１８（特許文献２参照）、特開２００５－３２３２４５（特許文献３参照）、特表２００６－５１６０５４（特許文献４参照）に記載されている。

[0010] 特開２００２－２０７６２９に記載のシステムでは、ストレージに対する使用可能容量の上限が２種類設定され、ユーザが取り扱うデータ容量が２つの上限のいずれかを超えるか否かにより、使用するストレージが決まる。特開２００４－２１８１８に記載のシステムは、ストレージアクセスフローを識別する識別子と、その送信元ストレージ識別子と送信先ストレージ識別子を対応付けたテーブルを利用してアクセスフローを識別分離し、アクセスフロー毎の優先制御等を行う。特開２００５－３２３２４５に記載の仮想スイッチは、通信品質を保証する処理とアクセスフローとを対応付けたフィルタリングテーブルを利用して、通信品質に応じたフロー制御を行う。特表２００６－５１６０５４には、通過パケット分類、ポリシ処理、及びセキュリティ処理を実行し、最大回線速度を維持するＩＰプロセッサが記載されている。

先行技術文献

特許文献

[0011] 特許文献1：特開２００２－２０７６２９号公報

特許文献2：特開２００４－２１８１８号公報

特許文献3：特開２００５－３２３２４５号公報

特許文献4：特表２００６－５１６０５４号公報

非特許文献

[0012] 非特許文献1：ＯｐｅｎＦｌｏｗ Ｓｗｉｔｃｈ Ｓｐｅｃｉｆｉｃａｔｉｏｎ Ｖｅｒｓｉｏｎ １．０．０ (Ｗｉｒｅ Ｐｒｏｔｏｃｏｌ ０ｘ０１)　Ｄｅｃｅｍｂｅｒ ３１，　２００９

発明の概要

[0013] 従来、クラスタを活用することによって高信頼なストレージシステムが実現されてきた。クラスタでは、ネットワークに接続するコントローラ部を冗長化して、配下の単一ディスクコントローラ部の動作を保証している。昨今、クラウド環境のストレージシステムや、グローバルネームスペースにより仮想化されたストレージシステムでは、より多くのコントローラ部によって１つのシステムを制御する構成が可能となっている。そのため、各ストレージシステムやストレージシステム間の複雑な連携機能と設定が必要となってきている。

[0014] 又、このようなシステムにおいてユーザの優先アクセス制御を実施するには、各ストレージシステムとＡＤＳやＬＤＡＰサーバなどの認証サーバとの連携が必要となる。このため、クラウド環境における複数のストレージシステムにおいてユーザの優先アクセス制御を実施するには、ストレージシステム毎にユーザの情報を設定する必要がある。

[0015] 従来のネットワークシステムは、機能毎の複数のネットワークアプライアンス機器で構成され、複雑なネットワーク構成となっている。ネットワークアプライアンス機器の保守のためには、それぞれのメンテナンス方法に精通する必要がある。

[0016] 以上のように、複数のコントローラによって制御される複数のシステムを有するネットワークシステムでは、システム間の連携や優先アクセス制御のための設定をシステム毎に行う必要がある。この場合、ユーザは、システムを管理又は利用する場合、コントローラによって制御されるシステム毎の状況や、ネットワークシステム全体の状況を把握する必要がある。又、アプライアンス機器に応じたメンテナンスを行わなければならないため、保守管理が煩雑となる。

[0017] 従って、本発明の目的は、複数のシステムを有するネットワークシステムにおいて、複数のシステムのそれぞれやネットワーク全体の状況を意識せずに、ネットワークシステム全体を管理又は利用することにある。

[0018] 本発明の他の目的は、複数のシステムを有するネットワークシステムにお

けるユーザ単位の優先アクセス制御を簡易に設定することにある。

[0019] 一の観点において、本発明によるコンピュータシステムは、コントローラと、コントローラによって設定されたフローエントリに従って、受信パケットの中継処理を行う複数のスイッチと、複数のストレージシステムとを具備する。コントローラは、複数のストレージシステムから取得したリソース情報に基づいて、複数のスイッチのフローエントリを変更することで、クライアント端末のアクセス先となるストレージシステムを変更する。

[0020] 他の観点において、本発明による通信経路変更方法は、コントローラが、通信経路上のスイッチに対して、スイッチにおける受信パケットの中継先を規定するフローエントリを設定するステップと、コントローラが、複数のストレージシステムから取得したリソース情報に基づいて、スイッチのフローエントリを変更することで、クライアント端末のアクセス先となるストレージシステムを変更するステップとを具備する。

[0021] 本発明による通信経路変更方法は、記録媒体に格納され、コンピュータによって実行されるプログラムによって実現されることが好ましい。

[0022] 本発明によれば、複数のストレージシステムを有するネットワークシステムにおいて、複数のストレージシステムのそれぞれやネットワーク全体の状況を意識せずに、ネットワークシステム全体を管理又は利用することができる。

[0023] 又、本発明によれば、複数のシステムを有するネットワークシステムにおけるユーザ単位の優先アクセス制御を簡易に設定することが可能となる。

図面の簡単な説明

[0024] 上記発明の目的、効果、特徴は、添付される図面と連携して実施の形態の記述から、より明らかになる。

[図1]図１は、本発明によるコンピュータシステムの実施の形態における構成を示す図である。

[図2]図２は、本発明によるオープンフローコントローラの実施の形態における構成を示す図である。

[図3]図３は、本発明に係るリソース情報の構造の一例を示す図である。

[図4A]図４Ａは、本発明に係るロケーションポリシ情報の構造の一例を示す図である。

[図4B]図４Ｂは、本発明に係るロケーションポリシ情報の構造の他の一例を示す図である。

[図5]図５は、本発明に係るアクセスポリシ情報の構造の一例を示す図である。

[図6]図６は、本発明に係るオープンフロースイッチの実施の形態における構成を示す図である。

[図7]図７は、本発明によるコンピュータシステムの実施の形態におけるフロー切り替え動作の一例を示すフロー図である。

発明を実施するための形態

[0025] （概要）

本発明に係るＯＦＣは、複数のストレージシステムのそれぞれのリソース情報（例えば、ストレージシステムの性能、負荷、ディレクトリ／ファイルのロケーションに関する情報）を保持し、リソース情報に基づいたフロー制御により、ユーザのアクセス先となるストレージシステムを決定する。更に、本発明に係るＯＦＣは、ユーザの優先度に基づいたフロー制御を行うことで、優先度に応じたアクセス制御を実現する。このように、本発明によるコンピュータシステムは、ＯＦＣによってストレージシステムのリソース管理を行うことによって、ストレージシステムに特殊な機能を追加することなく、容易にストレージシステムのロケーション管理と優先アクセス制御を実現する。本発明によるコンピュータシステムは、企業内ネットワークやインターネット／クラウド環境などにおいて利用されることが好適である。

[0026] 以下、添付図面を参照しながら本発明の実施の形態を説明する。図面において同一、又は類似の参照符号は、同一、類似、又は等価な構成要素を示す。

[0027] （コンピュータシステムの構成）

本発明によるコンピュータシステムでは、オープンフロープロトコルに従って通信経路の構築及びパケットデータの転送制御が行われる。図１は、本発明によるコンピュータシステムの実施の形態における構成を示す図である。図１を参照して、本発明によるコンピュータシステムは、複数のオープンフローコントローラ１０１～１０ｉ（以下、ＯＦＣ１０１～１０ｉと称す）を有するコントローラ群１００、複数のオープンフロースイッチ２０１～２０ｊ（以下、ＯＦＳ２０１～２０ｊと称す）を有するスイッチ群２００、複数のクライアント端末３０１～３０ｋを有するクライアント端末群３００、複数のストレージシステム４０１～４０ｎを有するストレージシステム群４００を具備する。ただし、ｉ、ｊ、ｋ、ｎは自然数である。

[0028] ＯＦＣ１０１～１０ｉのそれぞれは、セキュアな制御ネットワーク５００を介してＯＦＳ２０１～２０ｊのそれぞれに接続されるとともに、ネットワーク６００を介してストレージシステム４０１～４０ｎのそれぞれに接続される。ＯＦＣ１０１～１０ｉのそれぞれは、制御ネットワーク５００を介してＯＦＳ２０１～２０ｊのそれぞれのフローテーブル２３にフローエントリ（ルール＋アクション）を設定する。ＯＦＳ２０１～２０ｊのそれぞれは、制御ネットワーク５００を介してＯＦＣ１０１～１０ｉのそれぞれに対しファーストパケットを通知する。

[0029] ＯＦＣ１０１～１０ｉのそれぞれが制御するＯＦＳは予め決められていることが好ましく、複数のＯＦＣが同じＯＦＳを制御しても構わない。この場合、当該ＯＦＳは、複数のＯＦＣに対しファーストパケットを通知し、複数のＯＦＣのそれぞれは、ファーストパケットのヘッダ情報に基づいて当該ＯＦＳにフローエントリを設定するか否かを決める。例えば、ＯＦＣ毎に設定可能なフローが決められており、ファーストパケットが、自身が設定可能なフロー（ルール）に適合するＯＦＣが、当該ファーストパケットの通知元のＯＦＳのフローテーブルを設定する。

[0030] クライアント端末３０１～３０ｋは、図示しないＣＰＵ、ネットワークインタフェース（Ｉ／Ｆ）、及びメモリを備えるコンピュータ装置であり、メ

モリ内のプログラムを実行することでストレージシステム群４００との間で通信を行う。クライアント端末群３００は、スイッチ群２００を介してストレージシステム群４００に接続される。クライアント端末３０１～３０ｋのそれぞれは、複数のＯＦＳ２０１～２０ｊのいずれかを介して複数のストレージシステム４０１～４０ｎのいずれかにアクセスする。この際、ＯＦＳ２０１～２０ｊのそれぞれは、制御ネットワーク５００を介してＯＦＣ１０１～１０ｉのいずれかによって設定されたフローテーブル１８に従い、クライアント端末３０１～３０ｋのそれぞれから転送されたパケットの中継処理を行う。

[0031] ストレージシステム４０１～４０ｎのそれぞれは、図示しないコントローラとストレージ部４１を備える。ストレージ部４１は、複数の物理ディスクを有するディスクアレイ（ＲＡＩＤ：Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙｓ ｏｆ Ｉｎｄｅｐｅｎｄｅｎｔ Ｄｉｓｋｓ）、論理ディスク、ＳＡＮ（Ｓｔｏｒａｇｅ Ａｒｅａ Ｎｅｔｗｏｒｋ）、あるいはＮＡＳ（Ｎｅｔｗｏｒｋ Ａｔｔａｃｈｅｄ Ｓｔｏｒａｇｅ）に例示される。ストレージシステム４０ｎがディスクアレイである場合、ストレージ部４１は、図示しないディスクコントローラによって制御される複数の物理ディスクによって実現される。あるいは、ストレージシステム４０ｎがＳＡＮやＮＡＳである場合、ストレージ部４１は、図示しないサーバによって制御されるディスクアレイ等の記憶装置によって実現される。

[0032] ストレージ部４１には、自身が属するストレージシステム４０ｎのシステム情報、性能情報、負荷情報、ディレクトリ／ファイルロケーション情報が格納される。ここで、システム情報とはストレージシステム４０ｎを特定する情報（識別子やアドレス情報）や通信規格等を含む。性能情報は、ストレージシステム４０ｎにおけるデータ転送速度、時間当たりのＩ／Ｏ数、応答時間、記憶容量等に例示されるシステム性能を示す情報を含む。負荷情報は、ストレージシステム４０ｎへのアクセス負荷（応答時間、トランザクション量、データ転送量、レイテンシ等）やストレージシステム４０ｎのコント

ローラ（図示なし）における処理負荷（ＣＰＵロードアベレージ等）を示す情報を含む。又、負荷情報は、ストレージシステム４０ｎにおける障害有無や、発生した障害内容を示す情報（障害情報と称す）を含むことが好ましい。更に、負荷情報は、ストレージシステム４０ｎが使用不可となる情報（例えば、メンテナンス中を示す情報）を含むことが好ましい。ストレージシステム４０ｎでは、定期的又はユーザによって指定された時刻において負荷情報を自身の負荷情報を収集（計測）しストレージ部４１に記録する。ディレクトリ／ファイルロケーション情報は、ストレージ部４１に設定されたディレクトリやファイルのロケーションを特定する情報である。ストレージシステム４０ｎのシステム情報、性能情報、負荷情報、ディレクトリ／ファイルロケーション情報は、ストレージシステム毎に管理され、それぞれのストレージ部４１に記録されても良いし、ストレージ部４１と異なる図示しない記憶装置（例えばコントローラとして機能するサーバ内の記憶装置）に記録されても良い。

[0033] ストレージシステム群４００において、ストレージシステム４０１～４０ｎは、筐体間レプリケーションによる一般的な同期機能を有する。

[0034] 図２から図５を参照して、本発明に係るＯＦＣ１０ｉの構成の詳細を説明する。以下では、ＯＦＣ１０ｉの構成について説明するが、他のＯＦＣ１０１～１０ｉ－１も同様な構成であるため、その説明は省略する。

[0035] 図２は、本発明によるＯＦＣ１０ｉの実施の形態における構成を示す図である。ＯＦＣ１０ｉは、オープンフロー技術により、システム内におけるパケット転送に係る通信経路パケット転送処理を制御するフロー制御部１３を備える。オープンフロー技術とは、コントローラ（ここではＯＦＣ１０ｉ）が、ルーティングポリシー（フローエントリ：フロー＋アクション）に従い、マルチレイヤ及びフロー単位の経路情報をＯＦＳ２０ｊに設定し、経路制御やノード制御を行う技術を示す（詳細は、非特許文献１を参照）。これにより、経路制御機能がルータやスイッチから分離され、コントローラによる集中制御によって最適なルーティング、トラフィック管理が可能となる。オ

ープンフロー技術が適用されるＯＦＳ２０ｊは、従来のルータやスイッチのようにパケットやフレームの単位ではなく、ＥＮＤ２ＥＮＤのフローとして通信を取り扱う。

[0036] ＯＦＣ１０ｉは、図示しないＣＰＵ、ネットワークＩ／Ｆ、及びメモリを備えるコンピュータによって実現される。ＯＦＣ１０ｉでは、ＣＰＵがメモリ等の記録媒体に格納されたプログラムを実行することで、図２に示すリソース管理部１１、ロケーション決定部１２、フロー制御部１３の各機能を実現する。又、ＯＦＣ１０ｉのメモリ（図示なし）には、リソース情報１５、ロケーションポリシ情報１６、アクセスポリシ情報１７、フローテーブル１８が記録される。ＯＦＣ１０ｉは、設定コンソール１０に接続され、設定コンソール１０によって、ロケーションポリシ情報１６が設定される。又、スイッチ群２００によるネットワークに接続した後は、当該ネットワーク経由でロケーションポリシ情報１６が設定され得る。尚、リソース情報１５、ロケーションポリシ情報１６は、他のＯＦＣに共有されても良い。例えば、リソース情報１５、ロケーションポリシ情報１６は、コントローラ群１００における他のＯＦＣからアクセス可能な状態で図示しないメモリに記録される。あるいは、リソース情報１５、ロケーションポリシ情報１６は、ＯＦＣ１０１～１０ｉで共通の記憶装置（図示なし）に記録される。

[0037] リソース管理部１１は、ストレージシステム４０１～４０ｎのそれぞれのシステム情報、性能情報、負荷情報、ディレクトリ／ファイルロケーション情報を収集し、リソース情報１５として図示しない記憶装置に記録する。例えば、リソース管理部１１は、ＳＮＭＰ（Ｓｉｍｐｌｅ Ｎｅｔｗｏｒｋ Ｍａｎａｇｅｍｅｎｔ Ｐｒｏｔｏｃｏｌ）を利用した方法や、ストレージシステム４０ｎ内におかれた図示しない負荷収集エージェントから情報を収集する方法等によりリソース情報１５を収集する。

[0038] 図３は、本発明に係るリソース情報１５の構造の一例を示す図である。ストレージシステム４０１～４０ｎのそれぞれのシステム情報、性能情報、負荷情報、ディレクトリ／ファイルロケーション情報は、システム情報１５１

、性能情報１５２、負荷情報１５３、ロケーション情報１５４として対応付けられて記録される。これにより、ＯＦＣ１０ｉは、ストレージシステム毎の性能、負荷、障害状況、ディレクトリ／ファイルロケーションを管理することが可能となる。

[0039] ロケーションポリシ情報１６には、クライアント端末毎（ユーザ毎）に使用可能なストレージシステムが対応付けられて記録される。図４Ａ及び図４Ｂは、本発明に係るロケーションポリシ情報１６の構造の一例を示す図である。図４Ａを参照して、ロケーションポリシ情報１６には、クライアント端末を特定するユーザＩＤ１６１と、当該クライアント端末が使用可能なストレージシステムを特定する使用可能システム情報１６２とが対応付けて設定される。例えば、クライアント端末３０１の使用可能なストレージシステムが４０１～４０３として設定される場合、クライアント端末３０１を示すユーザＩＤ１６１とストレージシステム４０１～４０３を示すシステム情報１５１とが対応付けられて、ロケーションポリシ情報１６として設定される。あるいは、図４Ｂに示すように、ユーザＩＤ１６１に対し優先度１６３が更に対応付けられても良い。例えば、クライアント端末３０１がクライアント端末３０２よりも高い優先度に設定される場合、クライアント端末３０１を示すユーザＩＤ１６１に優先度１６３“１”が対応付けられ、クライアント端末３０２を示すユーザＩＤ１６１に優先度１６３“１”より低い優先度１６３“２”が対応付けられて、ロケーションポリシ情報１６として設定される。

[0040] ロケーション決定部１２は、ストレージシステム群４００が有するディレクトリ／ファイルのロケーション情報全体（リソース情報１５）から、クライアント端末群３００におけるディレクトリ／ファイルのロケーションをストレージシステム４０１～４０ｎに分配し、その情報をアクセスポリシ情報１７として記録する。この際、ロケーション決定部１２は、負荷情報１５３に応じてロケーション毎の負荷を把握し、負荷や優先度に応じてアクセス先を割り当てる。これにより、クライアント端末群３００によるファイルアク

セス（ファイルおよびディレクトリの書き込み／読み出し／ファイル属性操作を含む）は分散される。

[0041] 詳細には、ロケーション決定部１２は、リソース情報１５とロケーションポリシ情報１６とに基づいて、クライアント端末３０１～３０ｋのそれぞれがアクセス（書き込み／読み出し／ファイル属性操作を含む）するディレクトリ／ファイルのロケーション（ストレージシステム）を決定する。ここで、ロケーション決定部１２は、リソース情報１５を参照して、ストレージシステムの状態（例えば、性能、負荷）を把握するとともに、ストレージシステムに設定されたディレクトリ／ファイルロケーションを確認する。又、ロケーション決定部１２は、ロケーションポリシ情報１６を参照してクライアント端末３０ｋが利用可能なストレージシステムを特定する。ロケーション決定部１２は、ストレージシステムの性能や負荷を考慮して、クライアント端末３０ｋのアクセス先となるロケーション（ストレージシステム）を、当該端末が使用可能なストレージシステムから選択する。例えば、ストレージシステム４０ｎの負荷が予め設定された閾値を超える場合、ロケーション決定部１２は、当該ストレージシステム４０ｎをアクセス先としたクライアント端末のアクセス先を他のストレージシステムに変更する。この際、ロケーション決定部１２は、ストレージシステムの性能や負荷を考慮し、例えば性能の高いストレージシステムや負荷の小さいストレージシステムを優先的にアクセス先として設定することが好ましい。ロケーション決定部１２は、クライアント端末３０ｋのアクセス先が決まると、アクセスポリシ情報１７として記憶装置に記録する。

[0042] 図４Ｂに示すアクセスポリシ情報１７のように、クライアント端末毎に優先度１６３が設定されている場合、ロケーション決定部１２は、優先度１６３を考慮してクライアント端末のアクセス先となるディレクトリ／ファイルのロケーション（ストレージシステム）を決定することが好ましい。例えば、優先度に応じてアクセス可能なロケーション（ストレージシステム）を予め設定しておき、設定対象のクライアント端末の優先度に対応するストレー

ジの中からアクセス先が設定される。尚、アクセス先として設定するアルゴリズムは、本実施の形態に例示した場合に限らず、ストレージシステムの負荷や性能、ユーザの優先度等を考慮して決定すれば、任意に設定され得る。

[0043] 図５は、本発明に係るアクセスポリシ情報１７の構造の一例を示す図である。アクセスポリシ情報１７は、ユーザ毎のアクセス対象となるディレクトリ／ファイルのロケーション（例えば格納対象となるストレージシステム）を示す情報である。図５を参照して、アクセスポリシ情報１７には、クライアント端末を特定するユーザＩＤ１６１と、当該クライアント端末のアクセス対象となるストレージシステム及びディレクトリ／ファイルロケーションを特定するアクセス対象システム情報１７２とが対応付けて設定される。クライアント端末のアクセス対象となるストレージシステムが変更された場合、アクセス対象システム情報１７２には、変更後のストレージシステム及びディレクトリ／ファイルロケーションを特定する情報とが記録される。この際、アクセス対象システム情報１７２には、アクセス対象変更前のストレージシステムを特定する情報が変更後の情報に対応付けられて記録されることが好ましい。アクセス対象システム情報１７２として記録されたストレージシステムを特定する情報は対応するシステム情報１５１に紐付けられ、ディレクトリ／ファイルロケーションを特定する情報は対応するロケーション情報１５４に紐付けられる。このようなアクセスポリシ情報１７を参照することで、ロケーション決定部１２及びフロー制御部１３は、現時点においてクライアント端末３０ｋのアクセス対象として設定されているストレージシステム及びディレクトリ／ファイルロケーションや、アクセス対象が変更される前にアクセス対象として設定されていたストレージシステムを特定することが可能となる。すなわち、本発明に係るＯＦＣ１０ｉは、クライアント端末毎のファイルを、どのストレージシステムのどのディレクトリ／ファイルロケーションに格納したかを管理することができる。

[0044] フロー制御部１３は、フローテーブル１８に従ってＯＦＳ２０ｊ毎にフローエントリ（ルール＋アクション）の設定又は削除を行う。ＯＦＳ２０ｊは

、設定されたフローエントリを参照し、受信パケットのヘッダ情報に応じたルールに対応するアクション（例えばパケットデータの中継や破棄）を実行する。ルール、及びアクションの詳細は後述する。

[0045] フロー制御部１３は、ＯＦＳ２０ｊからのファーストパケットの受信通知や、ロケーション決定部１２によるアクセスポリシ情報１７の変更に応じてＯＦＳ２０ｊに対するフローエントリ（フロー＋アクション）の設定、削除又は更新を行う。ここでファーストパケットとは、ＯＦＳ２０ｊに設定されたフローエントリ（ルール）に適合しないパケットデータを示す。

[0046] フローテーブル１８には、フローエントリを特定するためのフロー識別子、当該フローエントリの設定対象（ＯＦＳ２０ｊ）を識別する識別子、経路情報、フローエントリ（ルール、アクション情報）が対応付けられて設定される。フローテーブル１８には、ＯＦＣ１０ｉの制御対象となる全てのＯＦＳ２０ｊに対して生成されたフローエントリが設定される。又、フローテーブル１８には、フロー毎のＱｏＳや暗号化に関する情報など、通信の扱い方が定義されても構わない。

[0047] フローエントリに設定されるルールには、例えば、ＴＣＰ／ＩＰのパケットデータにおけるヘッダ情報に含まれる、ＯＳＩ（Ｏｐｅｎ Ｓｙｓｔｅｍｓ Ｉｎｔｅｒｃｏｎｎｅｃｔｉｏｎ）参照モデルのレイヤ１からレイヤ４のアドレスや識別子の組み合わせが規定される。例えば、レイヤ１の物理ポート、レイヤ２のＭＡＣアドレス、ＶＬＡＮタグ（ＶＬＡＮ ｉｄ）、レイヤ３のＩＰアドレス、レイヤ４のポート番号、のそれぞれの組み合わせがルールとして設定される。尚、ＶＬＡＮタグには、優先順位（ＶＬＡＮ Ｐｒｉｏｒｉｔｙ）が付与されていても良い。

[0048] ここで、ルールに設定されるポート番号等の識別子やアドレス等は、所定の範囲で設定されても構わない。又、宛先や送信元のアドレス等を区別してルールとして設定されることが好ましい。例えば、ＭＡＣ宛先アドレスの範囲や、接続先のアプリケーションを特定する宛先ポート番号の範囲、接続元のアプリケーションを特定する送信元ポート番号の範囲がルールとして設定

される。更に、データ転送プロトコルを特定する識別子をルールとして設定してもよい。

[0049] アクション情報には、例えばＴＣＰ／ＩＰのパケットデータを処理する方法が規定される。例えば、受信パケットデータを中継するか否かを示す情報や、中継する場合はその送信先が設定される。又、アクション情報には、パケットデータの複製や、破棄することを指示する情報が設定されてもよい。経路情報は、フローエントリ（ルール＋アクション情報）を適用する経路を特定する情報である。

[0050] フロー制御部１３は、アクセスポリシ情報１７に基づいて通信経路を算出し、当該通信経路上のＯＦＳ２０ｊに設定するフローエントリを生成する。例えば、ファーストパケットの通知に応じてフローエントリを生成する場合、フロー制御部１３は、ファーストパケットのヘッダ情報に基づいて転送元、又は転送先のクライアント端末を特定するとともに、アクセスポリシ情報１７を参照して当該クライアント端末（ユーザＩＤ１６１）のアクセス対象として設定されたストレージシステムやディレクトリ／ファイルロケーション（アクセス対象システム情報１７２）を特定する。フロー制御部１３は、特定したクライアント端末とストレージシステム（ファイルロケーション）の間の通信経路を算出するとともに当該通信経路上のＯＦＳ２０ｊに設定するフローエントリを生成し、フローテーブル１８及び当該ＯＦＳ２０ｊのフローテーブル２３に設定する。

[0051] あるいは、アクセスポリシ情報１７の変更（更新）に応じてフローエントリを生成する場合、フロー制御部１３は、変更されたアクセスポリシ情報１７のユーザＩＤ１６１からクライアント端末を特定するとともに当該ユーザＩＤ１６１に対応付けられたアクセス対象システム情報１７２から、アクセス対象のストレージシステムを特定する。フロー制御部１３は、特定したクライアント端末とストレージシステムの間の通信経路を算出するとともに当該通信経路上のＯＦＳ２０ｊに設定するフローエントリを生成し、フローテーブル１８及び当該ＯＦＳ２０ｊのフローテーブル２３に設定する。

[0052] 以上のような構成により、本発明によるＯＦＣ１０ｉは、ＯＦＳ２０ｊからのファーストパケットの受信通知のみならず、ストレージシステムの状況変化に応じたアクセスポリシ情報１７の変更をトリガとして、当該パケットを転送させるためのフローエントリを生成するとともに、当該フローエントリを、算出した通信経路上のＯＦＳ２０ｊに設定する。負荷分散やロケーションポリシに応じた通信経路の設定及び変更が可能となる。本発明では、オープンフロープロトコルを利用して通信経路制御を行っているため、負荷分散時や障害回避のためのメンテナンス時においても、ユーザに意識させることなくアクセス先となるディレクトリ／ファイルロケーションが決まり、システムとして分散配置させることが可能となる。

[0053] 図６は、本発明によるオープンフロースイッチ２０ｊの実施の形態における構成を示す図である。ＯＦＳ２０ｊは、ＯＦＣ１０ｉによって設定（更新）されたフローテーブル２３に従って受信パケットの処理方法（アクション）を決定する。ＯＦＳ２０ｊは、転送処理部２１とフロー管理部２２を備える。転送処理部２１とフロー管理部２２は、ハードウェアで構成されても、ＣＰＵによって実行されるソフトウェアで実現してもどちらでも良い。

[0054] ＯＦＳ２０ｊの記憶装置（図示なし）には、ＯＦＣ１０ｉによって設定されるフローテーブル２３が格納されている。フロー管理部２２は、ＯＦＣ１０ｉから取得したフローエントリ（ルール＋アクション情報）をフローテーブル２３に設定する。フロー管理部２２は、受信パケットのヘッダ情報が、フローテーブル２３に記録されたルールに適合（又は一致）しない場合、当該パケットデータをファーストパケットとして判定し、ファーストパケットを受信したことをＯＦＣ１０ｉに通知するとともにフローエントリの設定要求を発行する。この際、ＯＦＳ２０ｊは、コントローラ群１００の中から予め決められたＯＦＣに対してファーストパケットの通知を行う。

[0055] フロー管理部２２は、ＯＦＣ１０ｉから送信されるフローエントリ（ルール＋アクション情報）を、フローテーブル２３に設定する。本発明では、ＯＦＳ２０ｊにおけるファーストパケットの受信やストレージシステムの負荷

の変化等をトリガとして、ＯＦＳ２０ｊのフローテーブル２３が更新される。

[0056] 受信パケットのヘッダ情報が、フローテーブル２３に記録されたルールに適合（一致）する場合、当該パケットデータは、転送処理部２１によって、他のＯＦＳ又はストレージシステム４０ｎに転送される。詳細には、転送処理部２１は、パケットデータのヘッダ情報に適合（又は一致）するルールに対応するアクション情報を特定する。転送処理部２１は、当該アクション情報で指定された転送先のノード（ＯＦＳ２０ｊ又はストレージシステム４０ｎ）に当該パケットデータを転送する。ＯＦＳ２０ｊのフローテーブル２３の設定及びパケット転送動作は、オープンフロープロトコルに準拠する（参考：非特許文献１）。

[0057] （動作）

図１及び図７を参照して、本発明によるフロー切り替え動作の一例を説明する。図７は、本発明によるコンピュータシステムの実施の形態におけるフロー切り替え動作の一例を示すフロー図である。以下では、ＯＦＣ１０ｉによるフロー切り替え動作について説明する。

[0058] 先ず、ＯＦＣ１０ｉは、ストレージシステム４０１～４０ｎからリソース情報１５（システム情報、性能情報、負荷情報、ディレクトリ／ファイルロケーション情報）を収集する（ステップＳ１０１）。リソース情報１５の収集は、ＯＦＣ１０ｉからの要求に応じて行われても良いし、ストレージシステム４０１～４０ｎのそれぞれが自主的に通知することで行われても良い。又、リソース情報１５の収集は、定期的に行われても良いし、予めストレージシステム毎に設定されたタイミングや、負荷の増大により閾値を超えた場合や障害の発生をトリガとして行われても良い。

[0059] ＯＦＣ１０ｉのロケーション決定部１２は、リソース情報１５が更新されると、負荷情報１５３を参照し、障害が発生したストレージシステムや負荷が閾値を超えたストレージシステムを特定する（ステップＳ１０２、Ｓ１０３）。ここで、ＯＦＣ１０ｉのロケーション決定部１２は、自身が管理する

ストレージシステムの中で、いずれのストレージシステムにおいても障害がなく負荷が閾値を超えていない場合、リソース情報１５が更新されるまで待機する（ステップＳ１０２Ｎｏ、Ｓ１０３Ｎｏ）。

[0060] 更新された負荷情報１５３を参照し、障害が発生したストレージシステムや負荷が閾値を超えたストレージシステムがある場合、ロケーション決定部１２は、アクセス先のロケーションを変更する（ステップＳ１０４）。詳細には、ロケーション決定部１２は、障害が発生したストレージシステムや負荷が閾値を超えたストレージシステムをアクセス対象としたクライアント端末を、アクセスポリシ情報１７を参照して特定するとともに、当該クライアント端末がアクセス可能なストレージシステムを、ロケーションポリシ情報１６を参照して特定する。ロケーション決定部１２は、特定したクライアント端末及びアクセス可能なストレージシステムの組合せの中から、性能情報１５２や負荷情報１５３に基づいて最適な組合せを決定し、当該クライアント端末のアクセス先となるストレージシステム及びディレクトリ／ファイルロケーションを決定し、アクセスポリシ情報１７を変更する。この際、図４Ｂに示すように、ロケーションポリシ情報１６に、クライアント端末毎の優先度１６３が設定されている場合、ロケーション決定部１２は、優先度１６３を考慮してクライアント端末のアクセス先ストレージシステムを決定することが好ましい。これにより、クライアント端末に応じた優先制御が可能となる。

[0061] アクセスポリシ情報１７が更新されると、フロー制御部１３は、変更されたアクセスポリシ情報１７に基づいてフローの切り替えを行う（ステップＳ１０５）。詳細には、先ず、フロー制御部１３は、変更されたアクセスポリシ情報１７を参照して、切り替え対象となるフローの端点（クライアント端末及び変更前のアクセス先ストレージシステム）と切り替え後（変更後）のフローの端点（クライアント端末及び変更後のアクセス先ストレージシステム）を特定する。続いてフロー制御部１３は、切り替え後のフローに対応するフローエントリを生成し、フローテーブル１８及び当該フローに対応する

通信経路上のＯＦＳ２０ｊに設定する。フローに対応する通信経路の算出やフローエントリの設定は、オープンフロープロトコルに準拠して行われる（非特許文献１参照）。この際、新たに設定されるフローエントリと同じルールを持つフローエントリは、フローテーブル１８、２３から削除又は使用できない状態に設定されることが好ましい。

[0062] 以上のように本発明に係るＯＦＣ１０ｉは、ストレージシステム群４００における負荷状態や障害発生有無に応じて、ＯＦＳ２０ｊを外部から制御することで、フロー単位の経路制御、障害回復、負荷分散を実現する。例えば、クライアント端末３０１、３０２、３０３がアクセスしているストレージシステム４０１において障害が発生した場合や高負荷状態となった場合、ＯＦＣ１０ｉは、クライアント端末３０１、３０２、３０３を端点とするストレージシステム４０１へのフローＡを、クライアント端末３０１及びストレージシステム４０２を端点とするフローＢ、クライアント端末３０２及びストレージシステム４０３を端点とするフローＣ、クライアント端末３０３及びストレージシステム４０３を端点とするフローＤに切り替える。このように、ＯＦＣによるフローの切り替え制御により、ストレージシステム４０１がサービスを停止しても、クライアント端末３０１～３０３への影響なく、ストレージシステム４０１のハードウェア交換、ソフトウェア改修等のメンテナンスを実施することが可能となる。すなわち、本発明によれば、オープンフロープロトコルを利用してＮクラスタ制御機能が実現され、アクセス先のストレージシステムが停止したり高負荷状態となっても、クライアント端末３０１、３０２、３０３からストレージシステムへのアクセスの継続が可能となり、ストレージシステムの冗長性が確保される。

[0063] 図７に示す一例では、ストレージシステムにおける障害発生や負荷に応じてフローを切り替える態様について説明したが、これに限らず、ストレージシステムを停止する情報に基づいて行われてもよい。例えば、ストレージシステム４０ｎでメンテナンスが必要な場合、ストレージシステムを停止する必要がある。このような場合、例えば、ロケーションポリシ情報１６の使用

可能システム情報１６２を変更し、当該ストレージシステム４０ｎを使用できないようにすることができる。あるいは、ストレージシステム４０ｎのリソース情報に当該システムを停止することを示す情報を設定し、これをＯＦＣ１０ｉが取得することで、ＯＦＣ１０ｉは、停止するシステムをアクセス先とするフローを他のフローに切り替えることができる。

[0064] 又、図４Ｂに示すロケーションポリシ情報１６のように、クライアント端末毎に設定された優先度１６３を利用することで、ＯＦＣによるクライアント端末毎の優先フロー制御が可能となる。以下では一例として、クライアント端末３０１の優先度１６３“１”がクライアント端末３０２の優先度１６３“２”よりも高く設定されている場合のフローの切り替え（設定）について説明する。

[0065] クライアント端末３０１、３０２がストレージシステム４０１にアクセスしている状況で、ストレージシステム４０１に障害発生状態もしくは高負荷状態により、ストレージシステム４０１へのアクセスがスローダウンしている場合、ＯＦＣ１０ｉは、クライアント端末３０１、３０２を端点とするストレージシステム４０１へのフローＡを、クライアント端末３０１及びストレージシステム４０２を端点とするフローＢ、クライアント端末３０２及びストレージシステム４０３を端点とするフローＣに切り替える。この際、ＯＦＣ１０ｉは、クライアント端末毎に設定された優先度に応じてアクセス先となるストレージシステムを決定する。例えば、クライアント端末３０１、３０２のアクセス先として使用可能なストレージシステムがストレージシステム４０１～４０３であり、ストレージシステム４０２の負荷がストレージシステム４０３より低い場合、高優先度のクライアント端末３０１のアクセス対象は低負荷のストレージシステム４０２に変更され、低優先度のクライアント端末３０２のアクセス対象は、ストレージシステム４０２よりも高負荷のストレージシステム４０２に変更される。これにより、本発明によれば、高優先のクライアント端末がより快適なサービスを継続して受けることが可能となる。

[0066] 更に、フローの切り替えは、ストレージシステムの状態（例えば停止状態を示す情報、障害情報、負荷の大きさ）のみならず、ストレージシステムの性能を考慮して行われても構わない。例えば、ストレージシステム４０２に比較して、ストレージシステム４０３が高性能である場合、ＯＦＣ１０ｉは、高性能なストレージシステム４０３を高優先度のクライアント端末３０１のアクセス先として設定し、低性能なストレージシステム４０３を低優先度のクライアント端末３０２のアクセス先として設定して、フローの切り替えを行う。あるいは、高優先度のクライアント端末（例えばクライアント端末３０１）よりも多くのクライアント端末（例えばクライアント端末３０２～３０ｋ）が低優先度に設定されている場合、ＯＦＣ１０ｉは、低性能なストレージシステム４０２を、少ない数の高優先度のクライアント端末３０１のアクセス先として設定し、高性能なストレージシステム４０３を数の多い低優先度のクライアント端末３０２～３０ｋのアクセス先として設定して、フローの切り替えを行う。これにより、低性能のストレージシステム４０２への負荷をストレージシステム４０３に比較して、より低く抑え、高優先度のクライアント端末３０１に対するサービス価値を高く継続することが可能となる。

[0067] 本発明では、クライアント端末毎に優先度を設定しつつ、ストレージシステムの状態や性能に応じてそのアクセス先を変更できる。このため、例えば、高価なストレージシステム４０１は主に高優先アクセス権のユーザ用、比較的安価なストレージシステム４０２はそれ以外の低優先アクセス権のユーザ用、安価なストレージシステム４０３は比較的アクセスが少ないファイルを格納し、バックアップ用とすることができる。尚、ロケーションポリシ情報１６に設定される優先度１６３は、２段階のみならず複数段階に順位付けされても構わない。

[0068] 更に、フローの切り替えは、ストレージシステムが使用できなくなる場合に限らず、ＯＦＳ２０ｊが使用できなく場合に行われてもよい。この場合、ＯＦＳ２０ｊは、定期的、所定の時刻、あるいは障害の発生や負荷が閾値を

越えたタイミングにおいて、負荷情報（例えば、トラフィック量や輻輳量等）や障害発生有無等の情報を、制御ネットワーク５００を介してＯＦＣ１０ｉに通知する。ＯＦＣ１０ｉは、ストレージシステム４０ｎにおける負荷情報等と同様にＯＦＳ２０ｊの負荷情報等を管理することで、上述と同様にＯＦＳ２０ｉの状態に応じたフローの切り替えが可能となる。

[0069] ＯＦＣ１０ｉは、一定期間内の負荷情報について閾値を越える場合、通信経路やアクセス先ストレージシステムを変更することが好ましいが、スケジュールにより、所定の割合の通信経路を強制的に他の経路又は他のアクセス先に変更してもよい。これにより、リアルタイムでのトラヒックの高い状態を検知し、分散し、以上奈トラフィックの収束を短時間に行うことが可能となる。又、定期的なフローの切り替えを行った場合、異常なトラヒックの発生や負荷の増大を事前に回避することが可能となる。

[0070] 本発明に係るＯＦＣ１０ｉは、ネットワークの経路制御のみならず、ストレージシステム群４００のリソース管理やロケーション管理を行っている。このため、ネットワーク全体を１台の仮想的なスイッチのように扱いながら、クラウド環境のストレージシステムにおけるクラスタ制御機能や、ユーザ（クライアント端末）毎の優先アクセス制御を実現することが可能となる。

[0071] 又、ＯＦＣ１０ｉがフロー単位で経路制御、障害回復、負荷分散を行っているため、障害回復後や負荷分散後においてもユーザ（クライアント端末）毎のフローの可視化が可能となる。ＯＦＣ１０ｉは、ストレージシステム群４００のリソース情報１５を管理しているため、フロー毎の性能情報や障害の箇所を容易に特定することができ、ストレージシステムやネットワークを含めた冗長化が容易となる。従って、本発明によれば、システム全体の安定した業務サービスの提供が可能となる。又、複雑化するネットワーク環境において、サービスを停止することなく、メンテナンスを容易に実施することが可能となる。

[0072] 本発明によるコンピュータシステムでは、オープンフロープロトコルを利用しているため、Ｌ１／Ｌ２／Ｌ３／Ｌ４機能毎のネットワーク機器やＡＤ

ＳやＬＤＡＰ等の認証サーバ等、さまざまなネットワークアプライアンス機器の接続とその運用管理の煩雑さを解消し、シンプルなネットワーク構成により、負荷分散や優先アクセス制御を実現することが可能となる。又、ストレージシステムの機能（性能）や、異種混在のストレージシステムに依存しないシステムを構築することができ、クラウド環境においても、ユーザ（クライアント端末）は、ストレージシステムやネットワークの状況やファイルロケーションを意識することなく、ストレージシステムを利用することが可能となる。

[0073] ユーザは、ストレージシステムやネットワークの状況を意識する必要がないため、ストレージシステムをスケールアウトに拡張／縮小させることが可能となる。このように、ストレージシステムの機能によらずにクラウド環境のストレージシステムをスケールアウトに拡張することが可能であることから、ユーザは、ネットワーク環境に時宜に応じたストレージシステムを容易に導入することが可能となる。

[0074] 以上、本発明の実施の形態を詳述してきたが、具体的な構成は上記実施の形態に限られるものではなく、本発明の要旨を逸脱しない範囲の変更があっても本発明に含まれる。本実施の形態では、単一のＯＦＣによる優先アクセス制御、Ｎクラスタ制御について説明したが、これに限らず、ロケーションポリシ情報１６を共有する複数のＯＦＣによっても実現され得る。又、単一のＯＦＣによって優先アクセス制御、Ｎクラスタ制御が行われる場合、ロケーションポリシ情報１６を他のＯＦＣと共有することで、当該他のＯＦＣをバックアップ系として利用することが可能となる。現用系のＯＦＣが故障等によりネットワークから離脱する場合、バックアップ系のＯＦＣは、現用系ＯＦＣからリソース情報１５を引き継ぐ必要はなく、新たにストレージシステム４０１～４０ｎからリソース情報１５を収集すればよい。

[0075] 又、新たなストレージシステムがコンピュータシステムに追加された場合、既設のＯＦＣに当該ストレージシステムの情報が登録されることにより、新たなストレージシステムに対しても即座に上述のＮクラスタ制御や優先ア

クセス制御を実行することができる。尚、ＯＦＣに新たなストレージのクレデンシャルを予め登録しておくことで、ＯＦＣは、当該ストレージシステムの追加を自動検出により登録することができる。

[0076] 更に、本実施の形態ではストレージシステムへの通信経路の変更を一例に説明したが、これに限らず、クライアント端末にサービスを提供するサーバやコンピュータシステムでも良い。

[0077] 尚、本出願は、日本出願番号特願２０１２－０７４９３６に基づき、日本出願番号特願２０１２－０７４９３６における開示内容は引用により本出願に組み込まれる。

請求の範囲

［請求項1］ コントローラと、

前記コントローラによって設定されたフローエントリに従って、受信パケットの中継処理を行う複数のスイッチと、

複数のストレージシステムと

を具備し、

前記コントローラは、前記複数のストレージシステムから取得したリソース情報に基づいて、前記複数のスイッチのフローエントリを変更することで、クライアント端末のアクセス先となるストレージシステムを変更する

コンピュータシステム。

［請求項2］ 請求項１に記載のコンピュータシステムにおいて、

前記リソース情報は、前記複数のストレージシステムのそれぞれにおける負荷情報を含み、

前記コントローラは、前記負荷情報を参照し、負荷が閾値を超えるストレージシステムをアクセス先とするクライアント端末のアクセス先を変更する

コンピュータシステム。

［請求項3］ 請求項１又は２に記載のコンピュータシステムにおいて、

複数のクライアント端末を更に具備し、

前記コントローラは、前記複数のクライアント端末のそれぞれに設定された優先度に応じて、前記複数のクライアント端末のアクセス先となるストレージシステムを決定する

コンピュータシステム。

［請求項4］ 請求項３に記載のコンピュータシステムにおいて、

前記コントローラは、前記複数のクライアント端末のそれぞれのディレクトリ又はファイルのロケーション情報を保持し、クライアント端末のアクセス先となるストレージシステムの変更に応じて、前記ロ

ケーション情報を変更する

コンピュータシステム。

[請求項5] 請求項１から４のいずれか１項に記載のコンピュータシステムで利用されるコントローラ。

[請求項6] コントローラが、通信経路上のスイッチに対して、前記スイッチにおける受信パケットの中継先を規定するフローエントリを設定するステップと、

前記コントローラが、複数のストレージシステムから取得したリソース情報に基づいて、前記スイッチのフローエントリを変更することで、クライアント端末のアクセス先となるストレージシステムを変更するステップと

を具備する

通信経路変更方法。

[請求項7] 請求項６に記載の通信経路変更方法において、

前記リソース情報は、前記複数のストレージシステムのそれぞれにおける負荷情報を含み、

前記アクセス先を変更するステップにおいて、前記コントローラは、前記負荷情報を参照し、負荷が閾値を超えるストレージシステムをアクセス先とするクライアント端末のアクセス先を変更する

通信経路変更方法。

[請求項8] 請求項６又は７に記載の通信経路変更方法において、

前記アクセス先を変更するステップにおいて、前記コントローラは、複数のクライアント端末のそれぞれに設定された優先度に応じて、前記複数のクライアント端末のアクセス先となるストレージシステムを決定する

通信経路変更方法。

[請求項9] 請求項８に記載の通信経路変更方法において、

前記コントローラは、前記複数のクライアント端末のそれぞれのデ

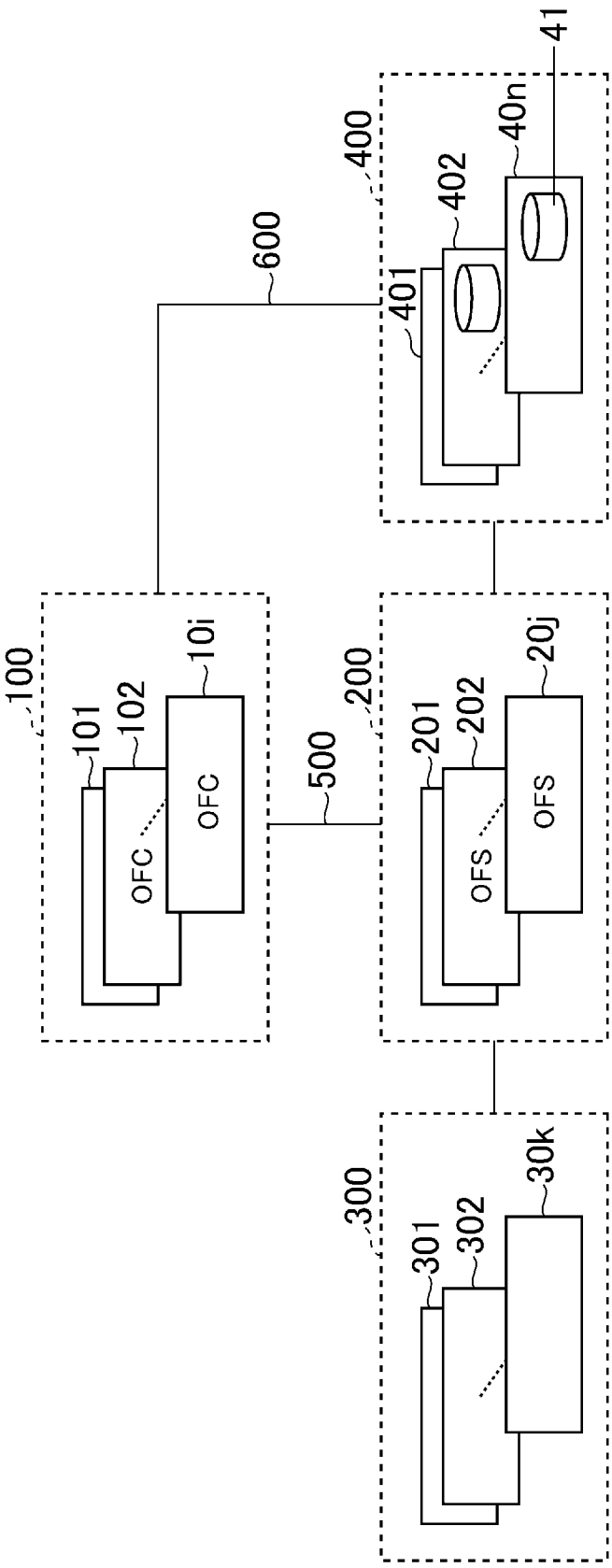
ィレクトリ又はファイルのロケーション情報を保持し、

前記アクセス先を変更するステップは、前記コントローラが、クライアント端末のアクセス先となるストレージシステムの変更に応じて、前記ロケーション情報を変更するステップを更に備える

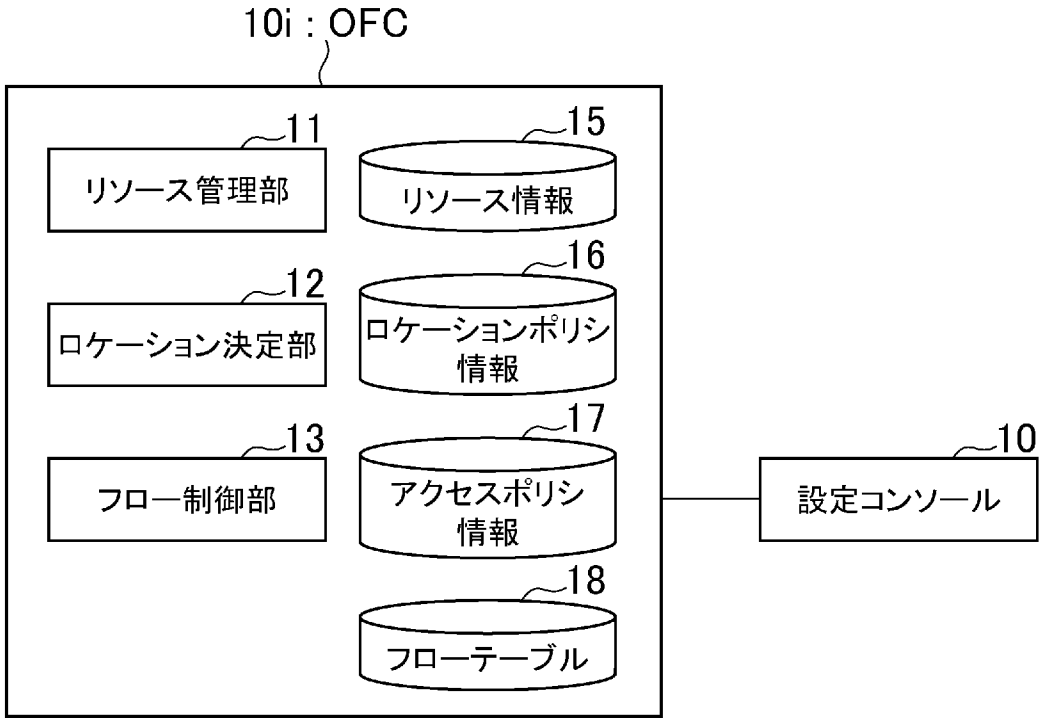
通信経路変更方法。

［請求項10］ 請求項６から９のいずれか１項に記載の通信経路変更方法をコンピュータに実行させる通信経路変更プログラムが記録された記録媒体。

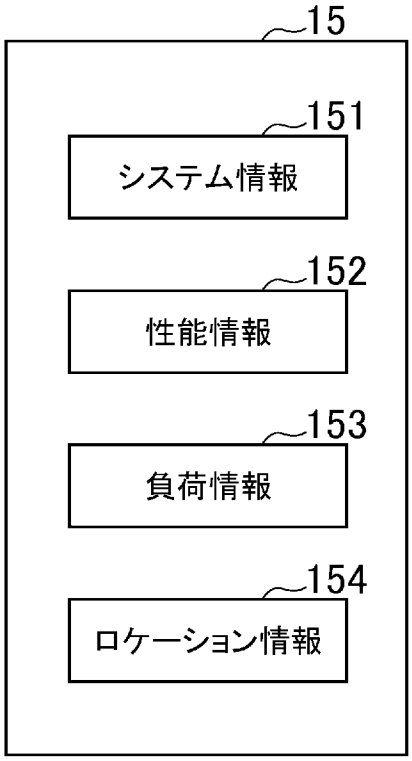
[図1]


41
400
40n
402
600
401
100
10i
101
102
OFC
OFC
500
200
20j
201
202
OFS
OFS
300
30k
301
302

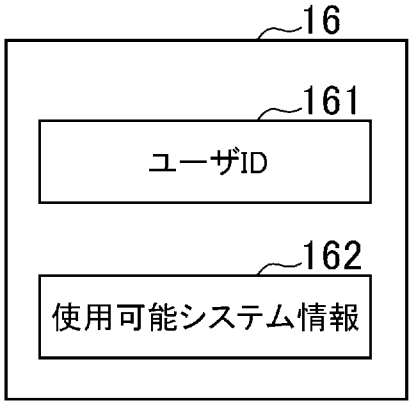
[図2]


10i : OFC
11
リソース管理部
12
ロケーション決定部
13
フロー制御部
15
リソース情報
16
ロケーションポリシ情報
17
アクセスポリシ情報
18
フローテーブル
10
設定コンソール

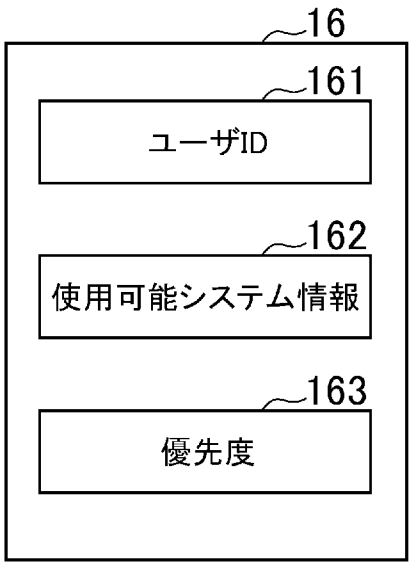
[図3]


15
151
システム情報
152
性能情報
153
負荷情報
154
ロケーション情報

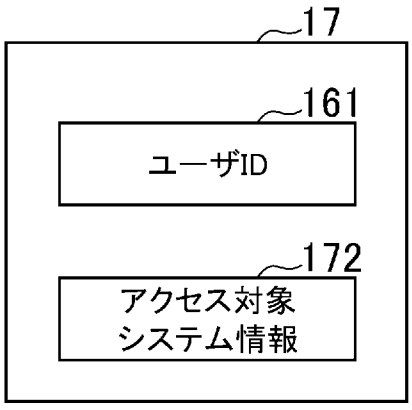
[図4A]


16
161
ユーザID
162
使用可能システム情報

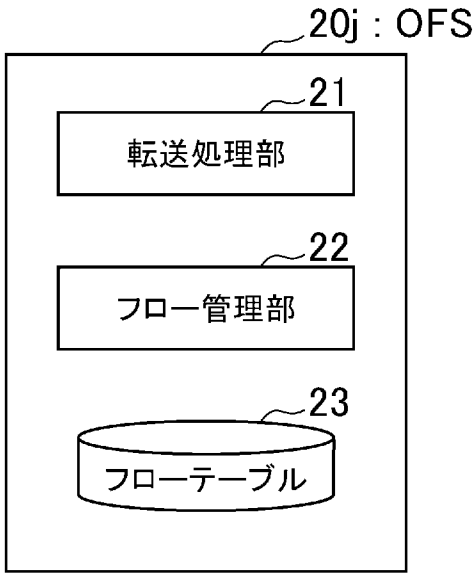
[図4B]


16
161
ユーザID
162
使用可能システム情報
163
優先度

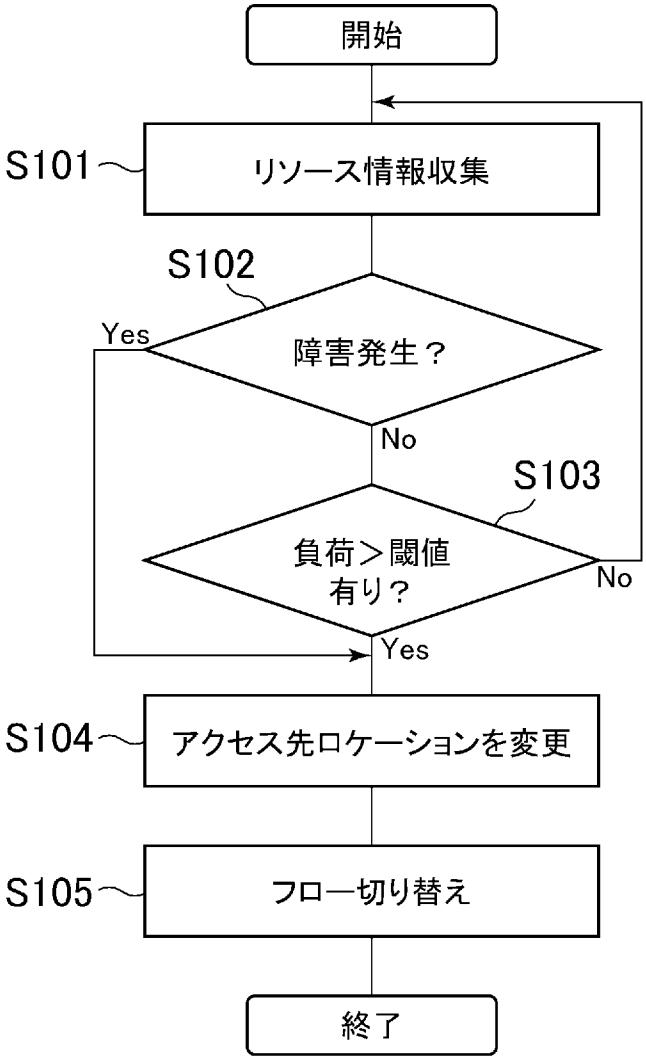
[図5]


17
161
ユーザID
172
アクセス対象
システム情報

[図6]


20j : OFS
21
転送処理部
22
フロー管理部
23
フローテーブル

[図7]


開始
S101
リソース情報収集
S102
障害発生？
Yes
No
S103
負荷＞閾値
有り？
No
Yes
S104
アクセス先ロケーションを変更
S105
フロー切り替え
終了

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2013/058846

A. CLASSIFICATION OF SUBJECT MATTER
H04L12/701(2013.01)i, *G06F13/00*(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04L12/701, G06F13/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2013
Kokai Jitsuyo Shinan Koho	1971-2013	Toroku Jitsuyo Shinan Koho	1994-2013

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2012-22555 A (Hitachi Systems & Services, Ltd.), 02 February 2012 (02.02.2012), paragraphs [0049], [0068], [0084] to [0090]; fig. 1 (Family: none)	1-10
Y	JP 2011-82799 A (NEC Corp.), 21 April 2011 (21.04.2011), entire text; all drawings & US 2012/0204051 A1 & EP 2487837 A1 & WO 2011/043317 A1 & CN 102549973 A	1-10
Y	JP 2000-293496 A (NEC Corp.), 20 October 2000 (20.10.2000), entire text; all drawings (Family: none)	2-5,7-10

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search 16 May, 2013 (16.05.13)	Date of mailing of the international search report 28 May, 2013 (28.05.13)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2013/058846

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2006-332825 A (Fujitsu Ltd.), 07 December 2006 (07.12.2006), paragraph [0259] & US 2006/0271700 A1 & US 2010/0057935 A1	3-5,8-10

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１３／０５８８４６

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　H04L12/701(2013.01)i, G06F13/00(2006.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　H04L12/701, G06F13/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１３年
日本国実用新案登録公報	１９９６－２０１３年
日本国登録実用新案公報	１９９４－２０１３年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2012-22555 A (株式会社日立システムアンドサービス) 2012.02.02, 【0049】, 【0068】, 【0084】-【0090】, 第１図 (ファミリーなし)	1-10
Y	JP 2011-82799 A (日本電気株式会社) 2011.04.21, 全文, 全図 & US 2012/0204051 A1 & EP 2487837 A1 & WO 2011/043317 A1 & CN 102549973 A	1-10

☑　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー

「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの

「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「Ｏ」口頭による開示、使用、展示等に言及する文献

「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献

「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
１６．０５．２０１３	２８．０５．２０１３

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｋ　９５６７
小宮　慎司

電話番号　０３－３５８１－１１０１　内線　３５５６

Ｃ（続き）．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2000-293496 A (日本電気株式会社) 2000.10.20, 全文,全図 (ファミリーなし)	2-5,7-10
Y	JP 2006-332825 A (富士通株式会社) 2006.12.07, 【0259】 & US 2006/0271700 A1 & US 2010/0057935 A1	3-5,8-10