

(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
2020年2月13日(13.02.2020)



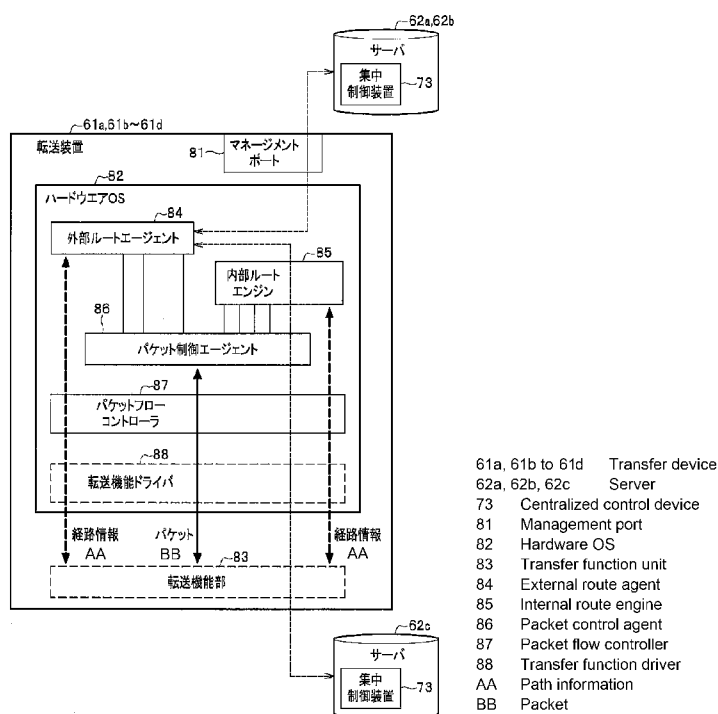
(10) 国際公開番号

WO 2020/031687 A1

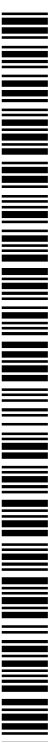
- (51) 国際特許分類:
H04L 12/717 (2013.01) H04L 12/707 (2013.01)
- (21) 国際出願番号: PCT/JP2019/028883
- (22) 国際出願日: 2019年7月23日(23.07.2019)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2018-148897 2018年8月7日(07.08.2018) JP
- (71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).
- (72) 発明者: 平澤 崇佳 (HIRASAWA, Takayoshi); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP).
入野 仁志 (IRINO, Hitoshi); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 Tokyo (JP).
- (74) 代理人: 特許業務法人磯野国際特許商標事務所 (ISONO INTERNATIONAL PATENT OFFICE, P.C.); 〒1050001 東京都港区虎ノ門一丁目1番18号 ヒューリック虎ノ門ビル 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,

(54) Title: TRANSFER DEVICE, TRANSFER SYSTEM, TRANSFER METHOD, AND PROGRAM

(54) 発明の名称: 転送装置、転送システム、転送方法及びプログラム



(57) Abstract: [Problem] To make it possible to avoid a single point of failure in a connection between a centralized control device and a group of transfer devices, to distribute traffic to a plurality of paths, and to select a bypass path in the event of failure in a switch cluster. [Solution] Transfer devices 61a to 61d, with respect to a centralized control device 73 which subjects a switch cluster comprising a group of transfer devices to centralized control externally, performs path control communication via the same path as a path for a D-plane (main signal). Each of the transfer devices 61a to 61d



WO 2020/031687 A1

CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO,
 DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
 HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH,
 KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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, ST, 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, KM, ML, MR, NE, SN, TD, TG).

添付公開書類 :

- 一 国際調査報告 (条約第21条(3))

is provided with: a packet flow controller 87 serving as a separating unit for separating an internal packet for a cluster 61 and an external packet for the cluster that are transmitted on the same path; and an internal route engine 85 which performs path control for determining a path that tracks a plurality of paths in the cluster as desired. The packet flow controller 87 separates an internally directed path control packet for the inside of the cluster. The engine 85, in the event of communication failure of the internally directed path control packet that has been separated, performs path control to generate a path bypassing the path of communication failure.

- (57) 要約 : 【課題】 集中制御装置と転送装置群との接続における単一障害点を回避できると共にトラフィックを複数経路に分散でき、スイッチクラスタ内の障害時に迂回経路を選択する。 【解決手段】 転送装置 61 a ~ 61 d は、転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置 73 に対して、D - p l a n e (主信号) と同様の経路と同じ経路を介して経路制御の通信を行う。前記同様の経路を伝送する、クラスタ 61 の内部用パケットと、当該クラスタの外部用パケットとを分離する分離部としてのパケットフローコントローラ 87 と、クラスタの内部の複数経路を自在に辿る経路を求める経路制御を行う内部ルートエンジン 85 とを備える。パケットフローコントローラ 87 は、クラスタの内部向けの経路制御パケットを分離し、前記エンジン 85 は、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行う。

明 細 書

発明の名称： 転送装置、転送システム、転送方法及びプログラム 技術分野

[0001] 本発明は、通信ネットワークを構成する転送装置群（スイッチクラスタ）の通信経路の拡張制御、並びに、通信経路における障害発生時の自動経路切替制御を行う転送装置、転送システム、転送方法及びプログラムに関する。

背景技術

[0002] 近年、OTT（Over The Top）の発展に伴い、汎用的な転送機能に特化したデータセンタ向けの汎用スイッチ市場が盛んになっている。通信キャリアとしても従来から用いられてきた転送装置として、キャリア向け専用開発装置を分離して汎用装置群とする検討が進められている（特許文献1）。

[0003] 例えば、図10に従来の転送装置の構成を示す。この転送装置10は、転送装置10の制御を行うマネジメント機能10aと、他の転送装置と接続する経路を制御するルーティング機能10bと、パケットを転送するパケット転送機能10cの3つの機能を1パッケージに収容して構成されている。

[0004] 転送装置10の内部には、図11に示すように、NOS（Network Operating System）12で制御されるファブリックカード13a、13bと、ラインカード14a、14b、14cとが組み込まれている。ファブリックカード13a、13bと、ラインカード14a、14b、14cという2種類の装置コンポーネント間はClos型トポロジ15で接続されている。

[0005] また、図12に示す従来の転送システムは、外部のサーバ11に切り離して収容したマネジメント機能10aと、このマネジメント機能10aにより制御される転送装置11a、11b、11c、11d、11e群とで構成される。この際、各転送装置11a～11eはパケット転送機能10cを備える。更に、各転送装置11a～11e及びサーバ11の何れか一方でルーティング機能10bを備える。

[0006] 図12に示す伝送システムの各転送装置11a～11eには、図13に示

す構成要素が次のように收容されている。転送装置11aはスパインSW (Switch) 17aであり、転送装置11bはスパインSW17b、転送装置11cはリーフSW18a、転送装置11dはリーフSW18b、転送装置11eはリーフSW18eである。

[0007] 図13に示す転送システムは、サーバ16に搭載されたCTL (Controller) 16aで制御されるスパインSW17a, 17bと、リーフSW18a, 18b, 18cという2種類の装置コンポーネント間が、Clos型トポロジ19で接続されて構成されている。Clos型トポロジ19で接続されたスパインSW17a, 17b及びリーフSW18a~18cは、スイッチクラスタ20を構成している。この構成により大容量転送を実現している（非特許文献1）。

[0008] 図13に示すような、従来の転送装置と同様の転送容量を担保するために将来期待される通信ネットワーク構成においては、複数の転送装置を組み合わせさせてトラフィックを処理する構成が検討されている。

[0009] また、複数の汎用の転送装置群を制御する既存のアーキテクチャとして、自律分散型（非特許文献2）と、集中制御型（非特許文献3, 4）がある。図14に自律分散型転送システム21を示し、図15に集中制御型転送システム31を示す。

[0010] 図14に示す自律分散型転送システム21は、コントロール機能としてのNOSCTL22aが搭載された外部のサーバ22と、NOSCTL22aによりマネジメントSW23を介して制御される転送装置24a~24e群とで構成されている。各転送装置24a~24eは、ルーティング機能とパケット転送機能を実現するNOSを備える。このように、各転送装置24a~24eに自律的な経路解決機能としてのNOSを配備することで、NOSCTL22aとの接続が切断された場合においても、各転送装置24a~24e上のNOS自身で自律的に経路構築できるため、従来同様の通信ネットワークの信頼性が維持できる利点がある。

[0011] しかし、物理的な転送装置24a~24eの台数の増加に伴い経路情報も

増加し、通信ネットワーク全体に影響が生じる。既存のネットワークで用いられている専用の転送装置 24 a ~ 24 e は、図 14 に示すシステム 21 で置き換えを想定すると、装置台数の増加に伴う経路情報の増加により、ネットワーク全体に影響が生じるため、周囲に影響を与えずに置き換えることは困難である。そこで、下記集中制御型の転送システムが実現されている。

[0012] 図 15 に示す集中制御型転送システム 31 は、R/F (Routing/Forwarding) 制御を行うコントローラである R/FCTL 32 a が搭載された外部のサーバ 32 と、転送装置 34 a ~ 34 e 群とを備えて構成されている。R/FCTL 32 a は、ルーティング機能とフォワーディング制御機能とを有する。転送装置 34 a ~ 34 e 群は、パケット転送機能を備え、R/FCTL 32 a によりマネジメント SW 33 を介して制御される。

[0013] このような集中制御型の転送システム 31 は、複数の転送装置 34 a ~ 34 e に対して、外部に切り離れた R/FCTL 32 a で一元的に経路構築を行うため、経路情報交換時に物理的な転送装置の台数を意識せずに、転送装置を論理的なノードとして制御できる利点がある。

[0014] 非特許文献 5 における転送システムの実装例を図 16 に示す。ここで、上述の図 15 に示した集中制御型転送システム 31 では、R/FCTL 32 a から各転送装置 34 a ~ 34 e に破線で示す 1 本の制御線しか接続できず冗長性に欠けていた。しかし、図 16 の集中制御型の転送システム 41 では、データ転送用の線 43 も用いて、上記図 15 の 1 本の制御線のみ依存しない構成が実現されている。

[0015] 図 16 では、CTL 45 から、D-plane (主信号) と同様の経路も用いて集中制御するために、追加機能 47 をマネジメントポート 48 を介して各転送装置 42 a, 42 b に接続している。また、各外部サーバ 43 a, 43 b と各転送装置 42 a, 42 b とは、マネジメントポート 48 で接続されている。更に、各転送装置 42 a, 42 b に D-plane 用ポート 46 経由で外部装置 44 a, 44 b が接続されている。

先行技術文献

特許文献

[0016] 特許文献1：特開2017-38209号公報

非特許文献

[0017] 非特許文献1：C.Clos, “A Study of Non-Blocking Switching Networks.” The Bell System technical Journal, 32(2):406-424, March 1953.

非特許文献2：“Multi-Service Fabric (MSF)”， [online]， 2018 GitHub, Inc, [平成30年7月18日検索]， インターネット <URL: <https://github.com/multi-service-fabric/msf>>

非特許文献3：ONF, “Enabling next-generation solutions in service provider networks”， [online]， 2018, [平成30年7月18日検索]， インターネット <URL: <https://onosproject.org/>>

非特許文献4：ONF, “CORD Installing a physical PoD”， [online]， 2018, [平成30年7月18日検索]， インターネット <URL: https://guide.openCORD.org/cord-5.0/install_physical.html>

非特許文献5：平澤崇佳 他2名, “複数ノードの論理ノード化に関する高信頼化の一検討”，信学技報, NS2017-172, pp. 27-32.

発明の概要

発明が解決しようとする課題

[0018] 上述した非特許文献3～5では、通信ネットワークを構成する複数の転送装置を、1台の集中制御装置（CTL）で集中制御して管理する集中制御技術（OpenFlow）が用いられている。しかし、集中制御技術に対応した転送装置は、集中制御装置と接続するためのマネジメントポートを1つしか持たないものが多い。そこで、集中制御装置と複数の転送装置との接続には、マネジメントスイッチ（図15のマネジメントSW23, 33参照）で集約して接続するため、この集約接続部分が単一障害点となってしまう問題があった。更に、集中制御装置と転送装置とのトラフィックが特定経路に集中してしまう問題があった。

[0019] この他、図16に一例を示したように、外部サーバ等に搭載した集中制御装置（CTL45）から各転送装置を制御するための通信をD-plane（主信号）と同様の経路で行う技術も提案されている。しかし、スイッチクラスタ内のデータ転送は転送先ポートを静的（固定的）に指定して行っている。このため、スイッチクラスタ内の障害（例えば、図16に×印で示す経路障害）が発生すると、迂回経路の選択機能が無いため、集中制御装置との通信が途切れてしまう。

[0020] つまり、図17に示すように、図示せぬ外部サーバ内に配備されて集中制御を行うCTL51が、矢印Y2～Y5で示すように、各転送装置52a～52dとD-plane（主信号）と同様の経路により通信を行っている場合に次のように通信が行えなくなる。即ち、図17に×印で示すように、転送装置52aと転送装置52d間の経路に断線等の障害が発生したとすると、迂回経路の選択機能が無いので、CTL51は、転送装置52b及び52dと矢印Y2、Y3で示す通信が行えなくなる問題が生じる。

[0021] 本発明は、このような事情に鑑みてなされたものであり、集中制御装置と転送装置群との接続における単一障害点を回避できると共にトラフィックを複数経路に分散でき、スイッチクラスタ内の障害時に迂回経路を選択できる転送装置、転送システム、転送方法及びプログラムを提供することを課題とする。

課題を解決するための手段

[0022] 上記課題を解決するための手段として、請求項1に係る発明は、通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置であって、前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部の経路制御を行う外部用パケットとを分離する分離部と、前記スイッチクラスタの内部の経路制御を行う内部経路制御部とを備え、前記分離部は前記スイッチクラスタの内部向けの経路制御パケットを分離し、前記内部経路制御部は、当該分離さ

れた内部向けの経路制御パケットで前記スイッチクラスタ内の経路制御を行うことを特徴とする転送装置である。

[0023] 請求項 7 に係る発明は、通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置による転送方法であって、前記転送装置は、前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部用パケットとを分離するステップと、前記スイッチクラスタ内部の複数経路を自在に辿る経路を求める経路制御を行うステップと、前記分離において前記スイッチクラスタの内部向けの経路制御パケットを分離し、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行うステップとを実行することを特徴とする転送方法である。

[0024] 請求項 8 に係る発明は、通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置としてのコンピュータを、前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部用パケットとを分離する手段、前記スイッチクラスタ内部の複数経路を自在に辿る経路を求める経路制御を行う手段、前記分離において前記スイッチクラスタの内部向けの経路制御パケットを分離し、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行う手段、前記スイッチクラスタ内部の経路情報及び前記スイッチクラスタ外部の経路情報を、パケット転送機能が用いる経路情報として通知する手段として機能させるためのプログラムである。

[0025] 請求項 1 の構成と請求項 7 の方法と請求項 8 のプログラムによれば、スイッチクラスタ（クラスタともいう）の内部用パケットと、クラスタの外部用パケットとを分離できるので、クラスタ内部では、クラスタ外部の経路と独立に動的な経路制御が可能となる。このため、クラスタと、クラスタ外部の

集中制御装置との単一障害点が回避可能となると共に、クラスタ内の転送装置の故障や経路の故障においても、この故障経路を迂回する経路（迂回経路）を生成できるので、故障前の通信を継続できる。

[0026] 請求項 2 に係る発明は、前記内部経路制御部が、前記スイッチクラスタの内部に転送装置群の接続経路が複数ある場合、前記主信号経路に係る通信を複数経路に分散することも可能な経路制御を行うことを特徴とする請求項 1 に記載の転送装置である。

[0027] この構成によれば、クラスタ内部に経路が複数ある場合は、集中制御装置との通信を複数経路に分散できる。これによって、クラスタと、クラスタ外部の集中制御装置との単一障害点が回避可能となる。

[0028] 請求項 3 に係る発明は、前記内部経路制御部が、前記スイッチクラスタの内部経路を、当該スイッチクラスタの外部経路と独立して経路制御することを特徴とする請求項 1 又は 2 に記載の転送装置である。

[0029] この構成によれば、クラスタの内部経路は、外部経路と独立して制御されるので、クラスタ内部の故障を検知した際に迂回経路を生成し、故障前の通信を継続できる。

[0030] 請求項 4 に係る発明は、前記分離部で分離された前記スイッチクラスタの外部向けの経路制御パッケージを、前記集中制御装置へ転送する外部ルートエージェントを更に備えることを特徴とする請求項 1～3 の何れか 1 項に記載の転送装置である。

[0031] この構成によれば、クラスタの外部向けの経路制御パッケージを、内部向けの経路制御パッケージと独立して集中制御装置へ転送できるので、集中制御装置がクラスタ外部の経路制御を容易に行うことができる。

[0032] 請求項 5 に係る発明は、前記内部経路制御部が、前記集中制御装置と前記スイッチクラスタの内部の転送装置との通信を、当該スイッチクラスタの内部経路を介して行うように制御することを特徴とする請求項 1～4 の何れか 1 項に記載の転送装置である。

[0033] この構成によれば、集中制御装置がクラスタ内部の転送装置と通信する際

に、内部経路制御部が生成した迂回経路等を介して行うことができる。

[0034] 請求項6に係る発明は、通信の経路で接続された転送装置群によるスイッチクラスタと、当該スイッチクラスタの外部から集中制御を行う集中制御装置とを有し、当該集中制御装置で主信号経路を介して当該スイッチクラスタの経路制御の通信を行う転送システムであって、前記集中制御装置は、前記スイッチクラスタの外部向けの経路制御パケットに応じて、当該スイッチクラスタの外部からこのスイッチクラスタが単一ノードと見做せるように、当該外部との経路制御を行うことを特徴とする転送システムである。

[0035] この構成によれば、クラスタの外部から、クラスタ内の転送装置群を単一ノードと見做して通信を行うことができる。このため、転送装置群への通信を単純化できる。

発明の効果

[0036] 本発明によれば、集中制御装置と転送装置群との接続における単一障害点を回避できると共にトラフィックを複数経路に分散でき、スイッチクラスタ内の障害時に迂回経路を選択できる転送装置、転送システム、転送方法及びプログラムを提供することができる。

図面の簡単な説明

[0037] [図1]本発明の実施形態に係る転送装置を用いた転送システムの構成を示すブロック図である。

[図2]本実施形態の転送システムのスイッチクラスタ内部経路の構成を示すブロック図である。

[図3]本実施形態の転送システムのスイッチクラスタ外部経路及び集中制御装置とスイッチクラスタとの集中制御接続の構成を示すブロック図である。

[図4]本実施形態の転送装置の構成を示すブロック図である。

[図5]本実施形態の転送システムのスイッチクラスタ内部の経路故障時の動作を説明するための基本構成を示すブロック図である。

[図6]本実施形態の転送システムのスイッチクラスタ内部の経路故障を示すブロック図である。

[図7]本実施形態の転送システムのスイッチクラスタ内部の経路故障後の迂回経路を示すブロック図である。

[図8]本実施形態の転送システムのスイッチクラスタ内部の転送装置の故障を示すブロック図である。

[図9]本実施形態の転送システムのスイッチクラスタ内部の転送装置故障後の迂回経路を示すブロック図である。

[図10]従来の転送装置の構成を示す図である。

[図11]従来の転送システムにおいて、図10の転送装置に收容される構成要素であるファブリックカードとラインカード間がC I o s型トポロジで接続された構成を示すブロック図である。

[図12]転送装置群によるスイッチクラスタと、当該スイッチクラスタ外部のサーバとの分離構成を示す図である。

[図13]従来の転送システムにおいて、図12の転送装置に收容される構成要素であるスパインSWとリーフSW間がC I o s型トポロジで接続された構成を示すブロック図である。

[図14]自律分散型転送システムの構成を示すブロック図である。

[図15]集中制御型転送システムの構成を示すブロック図である。

[図16]従来の集中制御型転送システムにおいて1本の制御線のみ依存しない構成を示すブロック図である。

[図17]従来の転送システムの転送装置間に経路故障が生じた様態を示すブロック図である。

発明を実施するための形態

[0038] 以下、本発明の実施形態を、図面を参照して説明する。但し、本明細書の全図において機能の対応する構成部分には同一符号を付し、その説明を適宜省略する。

<実施形態の構成>

図1は、本発明の実施形態に係る転送装置を用いた転送システムの構成を示すブロック図である。

[0039] 図1に示す転送システム60は、複数の転送装置61a, 61b, 61c, 61dと、この転送装置61a~61d群によるスイッチクラスタ（クラスタ）61に接続された複数のサーバ62a, 62bとを備えて構成されている。クラスタ61の外部には、外部ルータ63が接続されている。これら構成要素は、光ファイバや導電線等の経路65a~65gで接続されている。

[0040] 本実施形態の特徴は、各転送装置61a~61dに、クラスタ61の内部（クラスタ内部という）からのD-plane（主信号）と同様の経路による経路制御パケットと、クラスタ61の外部（クラスタ外部という）からのD-plane（主信号）と同様の経路による経路制御パケットとを分離する機能（図4のパケットフローコントローラ87）を備える。この分離機能で、クラスタ内部の経路制御パケットと、クラスタ外部の経路制御パケットとを分離し、クラスタ内部の経路制御パケットに応じてクラスタ内部向けの経路制御エンジン（図4の内部ルートエンジン85）が経路制御を行う。また、外部ルートエージェント84（図4）がクラスタ外部の経路制御パケットをサーバ62a, 62bに搭載される集中制御装置（図4の集中制御装置73）に転送するようにした。

[0041] これによって、クラスタ内では、クラスタ外部の経路と独立に動的な経路制御が可能となるため、クラスタ61とこの外部の集中制御装置73（図4）との単一障害点が回避可能となると共に、クラスタ内の障害においても迂回経路を生成して障害を回避可能とした。更に、クラスタ内部に経路が複数ある場合は、集中制御装置73との通信を複数経路に分散可能とした。

[0042] 図2に示すように、各転送装置61a~61dに、上記分離されたクラスタ内の経路制御パケットから自律的な経路切替制御処理を行う自律分散制御機能85（図4の内部ルートエンジン85に対応）を備える。つまり、クラスタ外部の経路とは独立してクラスタ内部経路の制御が行われる。

[0043] 自律分散制御機能85は、D-plane（主信号）と同様の経路上で、転送装置61a~61dの故障（装置故障という）や、クラスタ内の経路6

5 c ~ 6 5 f の故障（経路故障という）を契機に、自動で集中制御装置 7 3 との通信（集中制御用通信）の経路 6 5 a ~ 6 5 f を、自律的に故障を回避できる経路に切り替えて通信を継続する機能を有する。また、自律分散制御機能 8 5 は、転送装置 6 1 a ~ 6 1 d、クラスタ内の経路の増減設においても、自律的に経路を切り替えて通信を継続する機能を有する。

[0044] 但し、集中制御用通信を D - p l a n e（主信号）と同様の経路上で流通させるために、集中制御用通信に用いる I P（Internet Protocol address）アドレスの経路交換を、各転送装置 6 1 a ~ 6 1 d 上のクラスタ内部向けの自律分散制御機能 8 5 が D - p l a n e（主信号）と同様の経路上で行っている。

[0045] また、自律分散制御機能 8 5 は、クラスタ内の経路 6 5 c ~ 6 5 f の状態を監視し、経路障害時に経路を選択するルーティング機能を有する。ルーティング機能は、クラスタ内の経路 6 5 c ~ 6 5 f の故障解決に用いるルーティングプロトコル（プロトコル）を有する。また、ルーティング機能は、前述した自律分散型転送システム 2 1（図 1 4）による自律分散型の転送と同様に、各転送装置 6 1 a ~ 6 1 d と集中制御装置 7 3 と間の転送による疎通性を自律的に解決する処理を行う。つまり、クラスタ内部に経路が複数ある場合は、各転送装置 6 1 a ~ 6 1 d と集中制御装置 7 3 との通信を複数経路に分散する処理も可能である。

[0046] クラスタ内部において、上記プロトコルは、O S P F（Open Shortest Path First）や I S - I S（Intermediate System to Intermediate System）等の動的に経路状態を取得し、自律的に経路計算を行うプロトコルであればどれでもよい。一方、クラスタの外部からくる経路情報においても同プロトコルを用いる場合があるが、その経路情報は、集中制御装置 7 3 で処理される。

[0047] 集中制御装置 7 3 から各転送装置 6 1 a ~ 6 1 d へのパケットの転送は、図 3 に示すように、転送装置 6 1 a ~ 6 1 d に外部ルートエージェント 8 4 を備え、この外部ルートエージェント 8 4 が集中制御装置 7 3 からの転送を

仲介する。この仲介は、集中制御装置 7 3 において、特定の転送装置 6 1 a からのパケットの出力を命令する場合は、一度、集中制御装置 7 3 から外部ルートエージェント 8 4 にそのパケット出力指示を出し、この外部ルートエージェント 8 4 から出力されたパケットが転送装置 6 1 a から他の転送装置 6 1 b へ出力される様にするをいう。

[0048] また、転送装置 6 1 a ~ 6 1 d 上の外部ルートエージェント 8 4 は、集中制御装置 7 3 より集中制御により経路指示を受けるが、その経路としてはマネージメントポート（図 4 のマネージメントポート 8 1）経由か、D - p l a n e 通信用ポート経由の何れか一方又は両方でもよい。クラスタ 6 1 の外部からの通信は、外部からの経路制御パケットを受信した集中制御装置 7 3 が集中制御をクラスタ 6 1 に行う。この集中制御の指示を、外部ルートエージェント 8 4 が受ける。

[0049] また、上述したクラスタ 6 1 の外部からの通信は、図 3 に示すように、転送システム 6 0 の外部に存在する外部ルータ 6 3 の外部ルーティング機能 7 4 が外部の経路制御パケットを、直接接続するクラスタ 6 1 内の転送装置 6 1 a ~ 6 1 d とこの装置上の外部ルートエージェント 8 4 を経由して集中制御装置 7 3 に転送する。

[0050] この構成により、クラスタ内の装置故障や経路故障時に、各転送装置 6 1 a ~ 6 1 d が自律的にクラスタ内の経路接続を解決する処理を行い、クラスタ 6 1 とマネージメントポート（図 4 のマネージメントポート 8 1）及び D - p l a n e 通信用ポートの何れか一方又は両方を介した集中制御装置 7 3 間の通信を維持可能としている。また、ルーティング機能がルーティングプロトコルに応じて、クラスタ内の経路 6 5 c ~ 6 5 f を複数経路同時に利用可能とし、これにより効率的な帯域利用を可能としている。

[0051] また、クラスタ内部の経路故障解決のための通信（内部通信）と、クラスタ外部からの通信（外部通信）とは、各転送装置 6 1 a ~ 6 1 d に配備される上述した分離機能により区別される。しかし、区別するための条件は、M A C（Media Access Control）アドレス、I P アドレス、V L A N（Virtual

Local Area Network) 番号等の固有情報を用いて行われている。なお、各固有情報を組み合わせて用いてもよい。

- [0052] 上記の内部通信と外部通信との分離機構により、内部で自律分散的に解決した経路の影響をクラスタ外部の転送装置へ与えることが無いため、クラスタ外部からはクラスタ61の転送装置61a~61d群を論理的な単一ノードと見做すことが可能となる。また、クラスタ内部では複数の転送装置61a~61dの自律分散的な連携を実現可能としている。
- [0053] 次に、図4は同一構成の転送装置61a~61dの内、転送装置61aの構成を代表して示すブロック図である。
- [0054] 図4に示す転送装置61aは、マネージメントポート81と、ASICによる高速・大容量のハードウェア転送を実現する転送機能部83と、CPUによるソフトウェア処理が行われる計算処理部で構成されている。計算処理部にはハードウェアOS82があり、このOS82上で、外部ルートエージェント84と、内部ルートエンジン85と、パケット制御エージェント86と、パケットフローコントローラ87と、転送機能ドライバ（ドライバともいう）88とが動作している。
- [0055] 転送機能部83は、ASIC (Application Specific Integrated Circuit) 等の半導体集積回路で形成されており、これにより高速で大容量の転送が可能となっている。また、転送機能部83は、転送先のポートを備え、当該ポートを使用して書き込まれたフォワーディングルールに従ってパケットを転送可能となっている。
- [0056] なお、内部ルートエンジン85は、上述した自律分散制御機能であり、請求項記載の内部経路制御部を構成する。パケットフローコントローラ87は、請求項記載の分離部を構成する。
- [0057] マネージメントポート81は、集中制御指示のパケットや、集中制御装置73で処理されるべきクラスタ外部からの経路情報パケットを、集中制御装置73があるサーバへ接続することができる。なお、集中制御指示のパケットや、集中制御装置73で処理されるべきクラスタ外部からの経路情報パケ

ットは、転送機能部 83 経由で集中制御装置 73 があるサーバへ接続することもできる。

[0058] 外部ルートエージェント 84 は、内部ルートエンジン 85 が後述のように生成した内部経路の情報を用いてマネージメントポート 81 や転送機能部 83 を介して集中制御装置 73 に接続し、集中制御装置 73 に経路制御を問い合わせる処理を行う。また、外部ルートエージェント 84 は、集中制御装置 73 が処理したクラスタ外部からの経路情報を外部ルートエージェント 84 内の記憶部（図示せず）に記憶する。パケットフローコントローラ 87 は、その外部ルートエージェント 84 に記憶された経路情報をドライバ 88 を介して転送機能部 83 にコピーする。更に、外部ルートエージェント 84 は、内部ルートエンジン 85 及び転送機能部 83 を介して、他のサーバ 62c の集中制御装置 73 に接続することも可能である。即ち、外部ルートエージェント 84 は、クラスタ外部からの経路情報を受け取り、集中制御装置 73 へ転送することで集中制御されるためのエージェントであり、集中制御装置 73 間との通信により、クラスタ内部のトポロジ解決用通信パケットと、外部からの経路情報パケットを集中制御装置 73 まで転送する。

[0059] 内部ルートエンジン 85 は、外部ルートエージェント 84 と集中制御装置 73 の間の通信経路を解決するための、内部経路情報を処理する。この内部ルートエンジン 85 は、ルーティングプロトコルに応じて内部経路の解決を次のように行う。即ち、パケット制御エージェント 86 に、転送機能部 83 の各ポートを模擬した仮想 I/F (Interface) を用意する。次に、パケットフローコントローラ 87 がパケット制御エージェント 86 に用意された各ポートの中から転送先に対応するポートへブリッジして、転送機能部 83 の内部経路に係るパケットを内部ルートエンジン 85 へ転送する。この転送されたパケットを内部ルートエンジン 85 が処理する。

[0060] また、内部ルートエンジン 85 は、クラスタ内の各転送装置 61a ~ 61d に配備された内部ルートエンジン 85 同士で経路交換を行い、転送装置 61a ~ 61d の故障や経路故障時に自動で迂回経路を生成する等の経路再計

算を行う。この経路再計算により内部経路の自動切り替えが可能となる。

- [0061] パケット制御エージェント 86 は、転送機能部 83 からパケットフローコントローラ 87 により引き上げられたパケットを内部ルートエンジン 85 や外部ルートエージェント 84 に受け渡す。
- [0062] 転送機能ドライバ 88 は、転送機能部 83 の制御を可能とするものであり、経路情報の転送機能部 83 への書き込みと、パケット制御エージェント 86 に入ってきたパケットを、パケットフローコントローラ 87 経由で転送機能部 83 のポートから転送させる機能と、転送機能部 83 のポートに入ってきたパケットの内、自装置宛となるパケットをパケットフローコントローラ 87 経由でパケット制御エージェント 86 まで転送することを可能とする機能を担っている。
- [0063] パケットフローコントローラ 87 は、後述のパケット分離と転送機能部の制御とを行う。即ち、パケット分離は、転送機能部 83 からパケットを取得し、内部ルートエンジン 85 と外部ルートエージェント 84 とに分離して出力する制御を行う処理である。転送機能部 83 の制御は、内部ルートエンジン 85 と外部ルートエージェント 84 に保存された経路情報をドライバ 88 を介して転送機能部 83 にコピーするものである。
- [0064] 転送機能部 83 に経路情報をコピーしない場合、パケットが入力された際に経路情報が無いためパケットの転送が不可能となる。転送機能部 83 は、経路情報があれば自転送機能部 83 自体でパケットの転送先を判断できる。
- [0065] また、経路故障や装置故障の障害発生時の迅速な切り替えを実現するため、パケットフローコントローラ 87 は、転送機能部 83 の各ポートの経路ダウンを検知した際に、内部ルートエンジン 85 に接続される仮想 I/F の内の対応するポートをダウンさせる。この機能により、ルーティングプロトコルのタイマ待ちによる経路再計算を行う場合と比較して高速な切り替えが可能となっている。
- [0066] <実施形態の動作>

次に、本実施形態に係る転送システム 60 の経路故障時の動作を、図 5 ~

図7を参照して説明する。

図5は2台のサーバ62a, 62bと、4台の転送装置61a~61dで構成されるスイッチクラスタとの構成を示すブロック図である。図5の矢印Y11~Y14は、集中制御装置73a, 73bと各転送装置61a~61d間で集中制御を行うための通信を示す。矢印Y15は、2つある集中制御装置73a, 73b同士で情報を同期するための通信を示す。

[0067] この通信時に、図6に×印で示すように経路65dに断線等の通信を不通とする故障が発生したとする。この場合、転送装置61d, 61b間では、経路65dを経由していたクラスタ内で処理されるパケットやクラスタ外からのパケットが不通となる。つまり、矢印Y14で示す転送装置61dと集中制御装置73b間の通信と、矢印Y15で示す転送装置61a, 61d, 61bを介した集中制御装置73a, 73b間の通信とが不通となる。この際、故障経路65dの両側の転送装置61d, 61bの内部ルートエンジン85（図4）が、経路65dの故障を検知する。

[0068] この経路故障検知後、クラスタ内の各転送装置61a~61dの内部ルートエンジン85は、内部ルートエンジン85同士で経路交換を行い、ルーティングプロトコルに応じて自律的に故障経路65dを経由しない迂回経路を生成する経路再計算を行う。

[0069] この経路再計算により図7に矢印Y17で示す迂回経路（新経路Y17）と、矢印Y18で示す迂回経路（新経路Y18）とが生成される。新経路Y17は、転送装置61dが転送装置61aを介して集中制御装置73aと繋がる経路である。新経路Y18は、集中制御装置73a, 73b同士が転送装置61a, 61c, 61bを介して繋がる経路である。

[0070] また、矢印Y17で示す経路上の転送装置61d, 61aの外部ルートエージェント84が新経路Y17を介して集中制御装置73aに接続する。この結果、集中制御装置73はクラスタ内部の経路構成変更を認識した上で、継続してクラスタを集中制御することが可能となる。

[0071] 更に、矢印Y18で示す経路上の転送装置61a, 61c, 61bの外部

ルートエージェント 84 が新経路 Y 18 を介して集中制御装置 73 a, 73 b に接続し新経路 Y 18 の経路情報を転送する。これにより集中制御装置 73 a, 73 b に新経路 Y 18 の経路情報が記憶され、集中制御装置 73 a, 73 b 間で新経路 Y 18 を経由した通信が可能となる。この後、集中制御装置 73 a, 73 b a による外部経路の処理が行われ、外部ルートエージェント 84 (図 4) 経由で新たな経路情報を転送機能部 83 (図 4) が認識する。

[0072] 新経路 Y 17 や Y 18 経由で外部ルートエージェント 84 や内部ルートエンジン 85 に蓄積された経路情報を、各転送装置 61 a ~ 61 d 上のパケットフローコントローラ 87 が各転送装置 61 a ~ 61 d の転送機能部 83 に更新する。

[0073] 次に、図 5 に示した通信時に、図 8 に×印で示すように転送装置 61 d に部品故障等の通信不能となる故障が発生したとする。この場合、故障した転送装置 61 d に繋がる経路 65 e、65 d が故障状態となるため、矢印 Y 14 で示す転送装置 61 d と転送装置 61 b を介した集中制御装置 73 b 間の通信が不通となり、矢印 Y 15 で示す転送装置 61 d を介した集中制御装置 73 a, 73 b 間の通信が不通となる。この際、故障した転送装置 61 d に繋がる転送装置 61 a, 61 b の内部ルートエンジン 85 が、経路 65 e, 65 d の故障を検知する。

[0074] この経路故障検知後、正常な経路 65 c, 65 f で繋がる転送装置 61 a ~ 61 c は、内部ルートエンジン 85 同士で経路交換を行い、次の経路再計算を行う。即ち、自律的なプロトコルに応じて故障経路 65 e, 65 d を経由しない迂回経路を生成する経路再計算を行う。この経路再計算により図 9 に矢印 Y 19 で示す迂回経路 (新経路 Y 19) が生成されたとする。新経路 Y 19 は、集中制御装置 73 a, 73 b 同士が、正常な経路 65 c, 65 f で繋がる転送装置 61 a, 61 c, 61 b を介して繋がる経路である。

[0075] その正常な転送装置 61 a, 61 c, 61 b の外部ルートエージェント 84 経由で新たな経路情報を転送機能部 83 が認識する。

- [0076] また、集中制御装置 73 a、73 b 間は、各転送装置上の内部ルートエンジン 85 が構築した新経路 Y 19 の経路を用いて接続することが可能となる。その後、集中制御装置 73 a、73 b は故障前と同様に、クラスタ外部からの新たな経路情報の処理が継続可能となる。
- [0077] 但し、経路故障を各転送装置 61 a、61 c、61 b 上のソフトウェアで認識できていない場合は、内部ルートエンジン 85 での経路故障検知による内部経路の再計算はできない。しかし、ルーティングプロトコルの応答も無くなるため、無応答のプロトコルに通常設定されているタイマのタイムアップを契機に、内部ルートエンジン 85 が別経路での再計算を行う。これにより経路の疎通が回復する。
- [0078] このようなパケット転送不能だが、各転送装置 61 a、61 c、61 b 上のソフトウェアで経路が未故障状態と認識されている場合に、高速な経路切替を実現するために、プロトコルのタイマを最小化することや、BFD (Bidirectional Forwarding Detection) 等の障害検知技術を併用してもよい。この機能により、経路故障又は、プロトコルの応答が無くなる場合の経路故障及び装置故障が発生した場合においても、集中制御装置 73 a、73 b による論理ノードとしての転送装置 61 a～61 d の制御を維持するための、従来のプロトコルタイマよりも高速な経路自動切替が実現可能となる。
- [0079] <実施形態の効果>
- 本実施形態に係る転送装置 61 a～61 d の効果について説明する。転送装置 61 a～61 d は、転送装置 61 a～61 d 群によるクラスタ 61 に外部から集中制御を行う集中制御装置 73 と、主信号経路である D - p l a n e (主信号) と同様の経路を介して経路制御の通信を行う。
- [0080] (1) 転送装置 61 a～61 d は、D - p l a n e (主信号) と同様の経路を伝送する、クラスタ 61 の内部用パケットと、当該クラスタ 61 の外部用パケットとを分離する分離部としてのパケットフローコントローラ 87 と、クラスタ 61 の内部の複数経路を自在に辿る経路を求める経路制御を行う内部ルートエンジン 85 とを備える。パケットフローコントローラ 87 は、

クラスタ 6 1 の内部向けの経路制御パケットを分離し、内部ルートエンジン 8 5 は、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行う構成とした。

[0081] 但し、内部用パケットは、クラスタ 6 1 のみで転送され、クラスタ 6 1 の集中制御を維持するために通信されるパケットである。具体的には、D - p l a n e（主信号）と同様の経路を介した集中制御通信用パケットと、内部ルートエンジン 8 5 がクラスタ 6 1 内部の経路情報を交換するために D - p l a n e（主信号）と同様の経路を介する通信パケットの 2 つである。また、外部用パケットは、クラスタ 6 1 外より送られてくるクラスタ 6 1 外部の経路情報等のパケット {ルーティングプロトコルのパケットや、A R P（Address Resolution Protocolパケット）等} である。

[0082] この構成によれば、クラスタ 6 1 の内部用パケットと、クラスタ 6 1 の外部用パケットとを分離できるので、クラスタ内部では、クラスタ外部の経路と独立に動的な経路制御が可能となる。このため、クラスタ 6 1 と、クラスタ外部の集中制御装置 7 3 との単一障害点が回避可能となると共に、クラスタ内部の転送装置の故障や経路の故障においても、この故障経路を迂回する経路（迂回経路）を生成できるので、故障前の通信を継続できる。

[0083] （2）内部ルートエンジン 8 5 は、クラスタ 6 1 の内部に各転送装置 6 1 a ~ 6 1 d と集中制御装置 7 3 との間に接続経路が複数ある場合、それらの集中制御用通信を複数経路に分散することを可能とする経路制御を行う構成とした。

[0084] この構成によれば、クラスタ内部に経路が複数ある場合は、集中制御装置 7 3 との通信を複数経路に分散できる。これによって、クラスタ 6 1 と、クラスタ外部の集中制御装置 7 3 との単一障害点が回避可能となる。

[0085] （3）内部ルートエンジン 8 5 は、クラスタ 6 1 の内部経路 6 5 c ~ 6 5 f を、当該クラスタ 6 1 の外部経路と独立して経路制御する構成とした。

[0086] この構成によれば、クラスタの内部経路 6 5 c ~ 6 5 f は、外部経路と独立して制御されるので、クラスタ内部の故障を検知した際に迂回経路を生成

し、故障前の通信を継続できる。

[0087] (4) 転送装置 61a～61d は、パケットフローコントローラ 87 で分離されたクラスタ 61 の外部向けの経路制御パケットを、集中制御装置 73 へ転送する外部ルートエージェント 84 を更に備える構成とした。

[0088] この構成によれば、クラスタの外部向けの経路制御パケットを、内部向けの経路制御パケットと独立して集中制御装置 73 へ転送できるので、集中制御装置 73 がクラスタ外部の経路制御を容易に行うことができる。

[0089] (5) 内部ルートエンジン 85 は、集中制御装置 73 とクラスタ 61 の内部の転送装置との通信を、当該クラスタ 61 の内部経路 65a～65f を介して行うように制御する構成とした。

[0090] この構成によれば、集中制御装置 73 がクラスタ内部の転送装置と通信する際に、内部ルートエンジン 85 が生成した迂回経路等を介して行うことができる。

[0091] また、本実施形態の転送システムは、通信の経路で接続された転送装置群によるクラスタ 61 と、当該クラスタ 61 の外部から集中制御を行う集中制御装置 73 とを有し、集中制御装置 73 で主信号経路を介してクラスタ 61 の経路制御の通信を行う。集中制御装置 73 は、クラスタ 61 の外部向けの経路制御パケットに応じて、クラスタ 61 の外部から当該クラスタ 61 が単一ノードと見做せるように、当該外部との経路制御を行う構成とした。

[0092] この構成によれば、クラスタの外部から、クラスタ内の転送装置 61a～61d 群を単一ノードと見做して通信を行うことができる。このため、転送装置 61a～61d 群への通信を単純化できる。つまり、纏めて1つのノードとして見せることで、クラスタ外部に対して経路情報が増加するのを抑圧できる。

[0093] また、本実施形態のコンピュータで実行されるプログラムについて説明する。コンピュータは、経路接続された転送装置 61a～61d 群によるクラスタ 61 に外部から集中制御を行う集中制御装置 73 に対して、D - p | a n e (主信号) と同様の経路を介して経路制御の通信を行う当該クラスタ 6

1の転送装置であるとする。このコンピュータは請求項に記載されており、転送装置内のCPU、又はCPU及びASIC等の半導体チップの両方を備えるものである。

[0094] このプログラムは、上記コンピュータを、D-plane（主信号）と同様の経路又は主信号と同様の経路を伝送する、クラスタ61の内部用パケットと、当該クラスタ61の外部用パケットとを分離する手段、クラスタ61内部の複数経路を自在に辿る経路を求める経路制御を行う手段、分離においてクラスタ61の内部向けの経路制御パケットを分離し、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行う手段、クラスタ61内部の経路情報及びクラスタ61外部の経路情報を、パケット転送機能（転送機能部83）が用いる経路情報として通知する手段として機能させる。

[0095] このプログラムによれば、上述した転送装置61a～61dと同様な効果を得ることができる。

[0096] その他、具体的な構成について、本発明の主旨を逸脱しない範囲で適宜変更が可能である。

符号の説明

- [0097] 60 転送システム
61a～61d 転送装置
62a, 62b サーバ
63 外部ルータ
65a～65g 経路
74 外部ルーティング機能
81 マネージメントポート
82 ハードウェアOS
83 転送機能部
84 外部ルートエージェント
85 内部ルートエンジン又は自律分散制御機能（内部経路制御部）

- 86 パケット制御エージェント（分離部）
- 87 パケットフローコントローラ（分離部）
- 88 転送機能ドライバ

請求の範囲

- [請求項1] 通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置であって、
- 前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部の経路制御を行う外部用パケットとを分離する分離部と、
- 前記スイッチクラスタの内部の経路制御を行う内部経路制御部とを備え、
- 前記分離部は前記スイッチクラスタの内部向けの経路制御パケットを分離し、前記内部経路制御部は、当該分離された内部向けの経路制御パケットで前記スイッチクラスタ内の経路制御を行う
- ことを特徴とする転送装置。
- [請求項2] 前記内部経路制御部は、
- 前記スイッチクラスタの内部に転送装置群の接続経路が複数ある場合、前記主信号経路に係る通信を複数経路に分散することも可能な経路制御を行う
- ことを特徴とする請求項1に記載の転送装置。
- [請求項3] 前記内部経路制御部は、
- 前記スイッチクラスタの内部経路を、当該スイッチクラスタの外部経路と独立して経路制御する
- ことを特徴とする請求項1又は2に記載の転送装置。
- [請求項4] 前記分離部で分離された前記スイッチクラスタの外部向けの経路制御パケットを、前記集中制御装置へ転送する外部ルートエージェントを更に備える
- ことを特徴とする請求項1～3の何れか1項に記載の転送装置。
- [請求項5] 前記内部経路制御部は、
- 前記集中制御装置と前記スイッチクラスタの内部の転送装置との通

信を、当該スイッチクラスタの内部経路を介して行うように制御することを特徴とする請求項1～4の何れか1項に記載の転送装置。

[請求項6]

通信の経路で接続された転送装置群によるスイッチクラスタと、当該スイッチクラスタの外部から集中制御を行う集中制御装置とを有し、当該集中制御装置で主信号経路を介して当該スイッチクラスタの経路制御の通信を行う転送システムであって、

前記集中制御装置は、

前記スイッチクラスタの外部向けの経路制御パケットに応じて、当該スイッチクラスタの外部からこのスイッチクラスタが単一ノードと見做せるように、当該外部との経路制御を行う

ことを特徴とする転送システム。

[請求項7]

通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置による転送方法であって、

前記転送装置は、

前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部用パケットとを分離するステップと、

前記スイッチクラスタ内部の複数経路を自在に辿る経路を求める経路制御を行うステップと、

前記分離において前記スイッチクラスタの内部向けの経路制御パケットを分離し、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行うステップと

を実行することを特徴とする転送方法。

[請求項8]

通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の

通信を行う当該スイッチクラスタの転送装置としてのコンピュータを
、

前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部用パケットとを分離する手段、

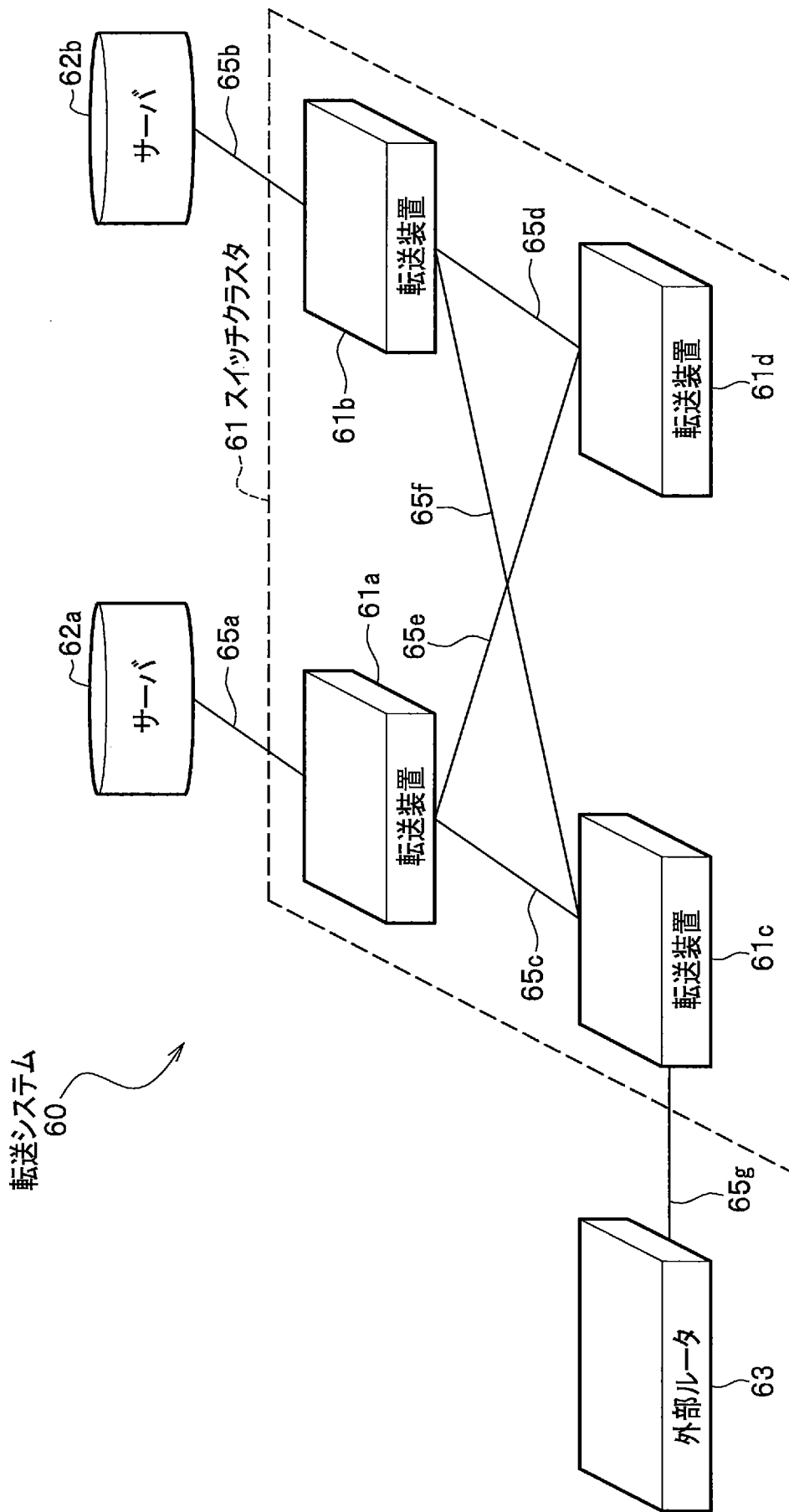
前記スイッチクラスタ内部の複数経路を自在に辿る経路を求める経路制御を行う手段、

前記分離において前記スイッチクラスタの内部向けの経路制御パケットを分離し、当該分離された内部向けの経路制御パケットの不通時に当該不通の経路を迂回する経路を生成する経路制御を行う手段、

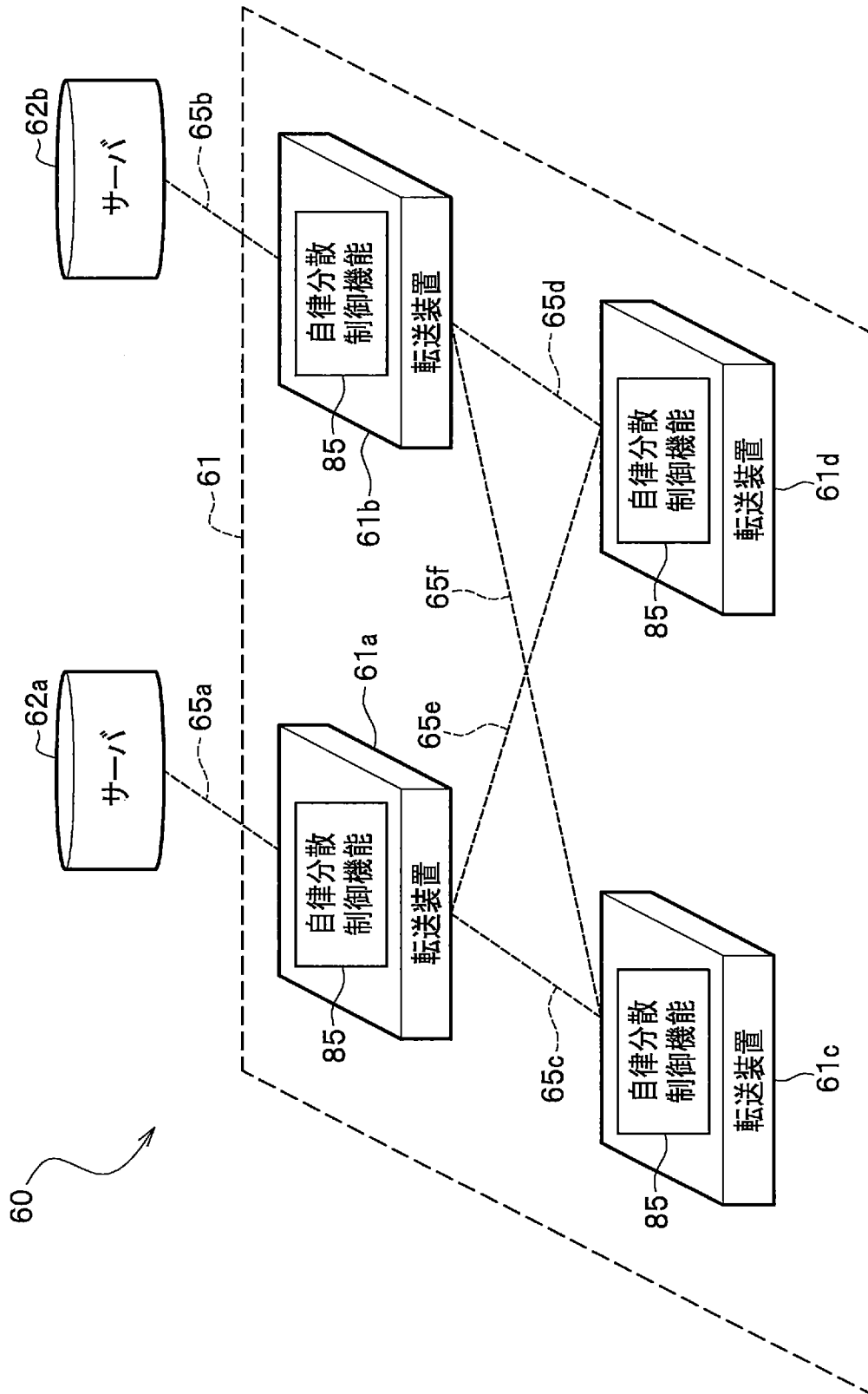
前記スイッチクラスタ内部の経路情報及び前記スイッチクラスタ外部の経路情報を、パケット転送機能が用いる経路情報として通知する手段

として機能させるためのプログラム。

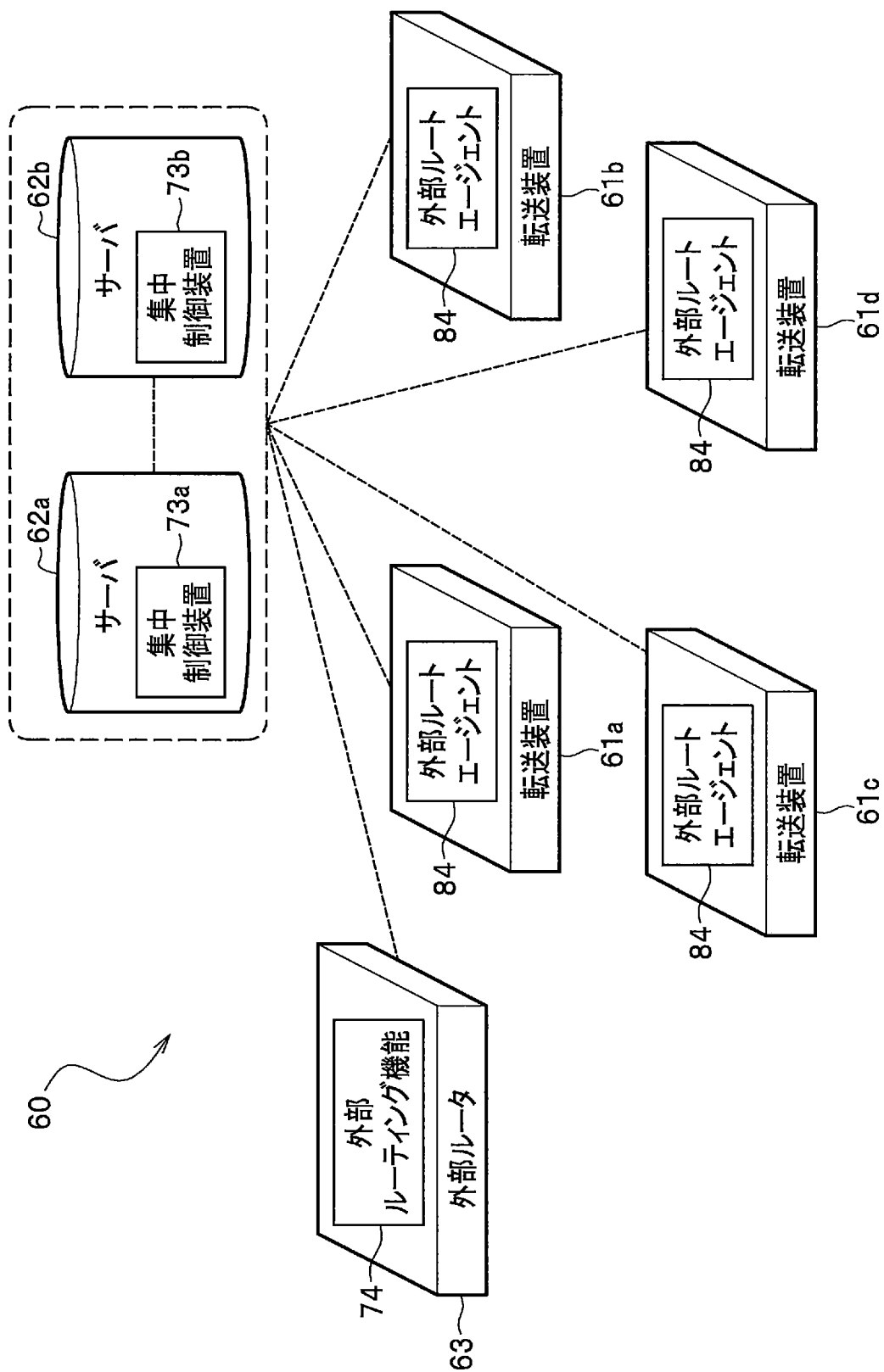
[図1]



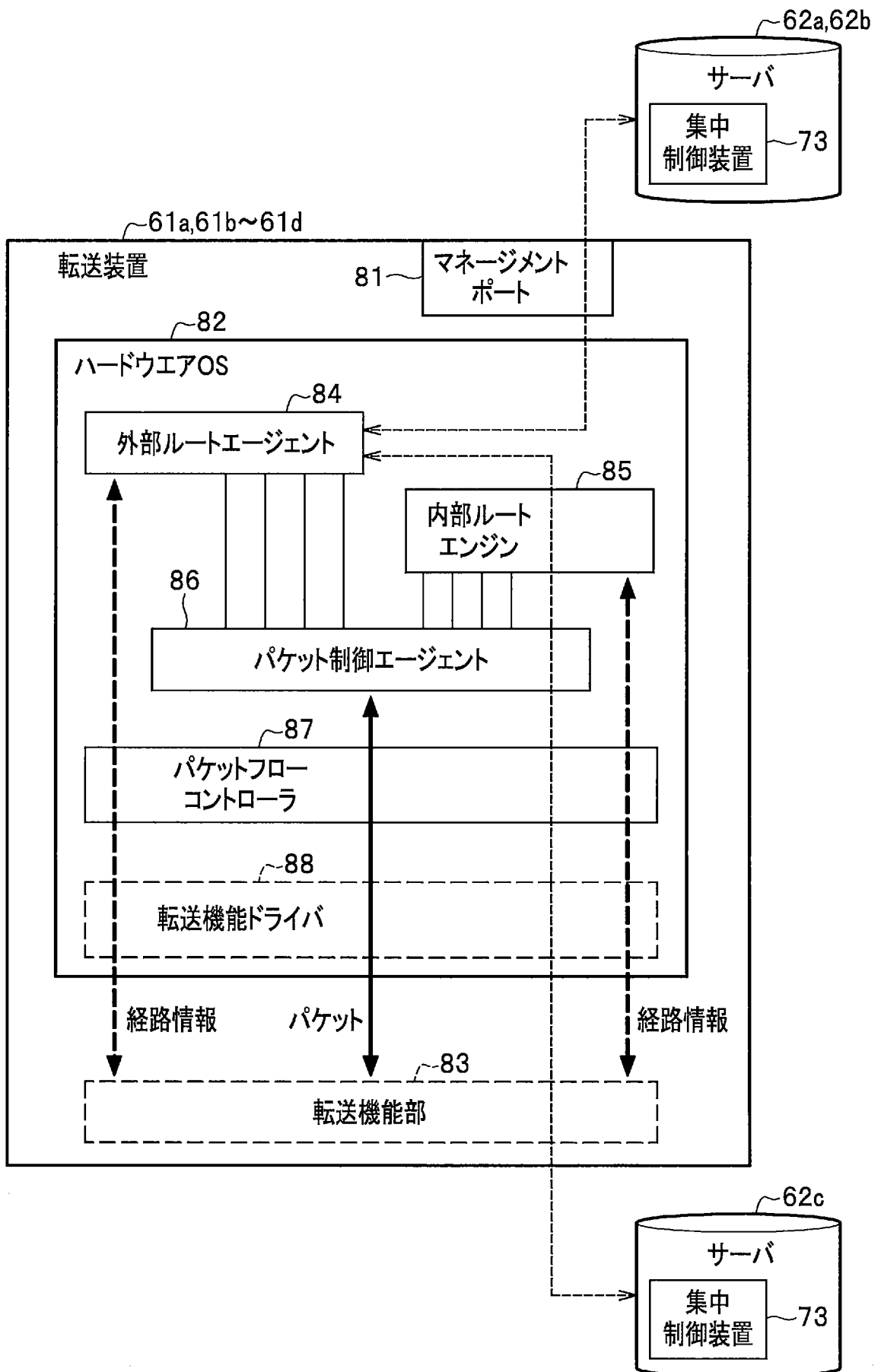
[図2]



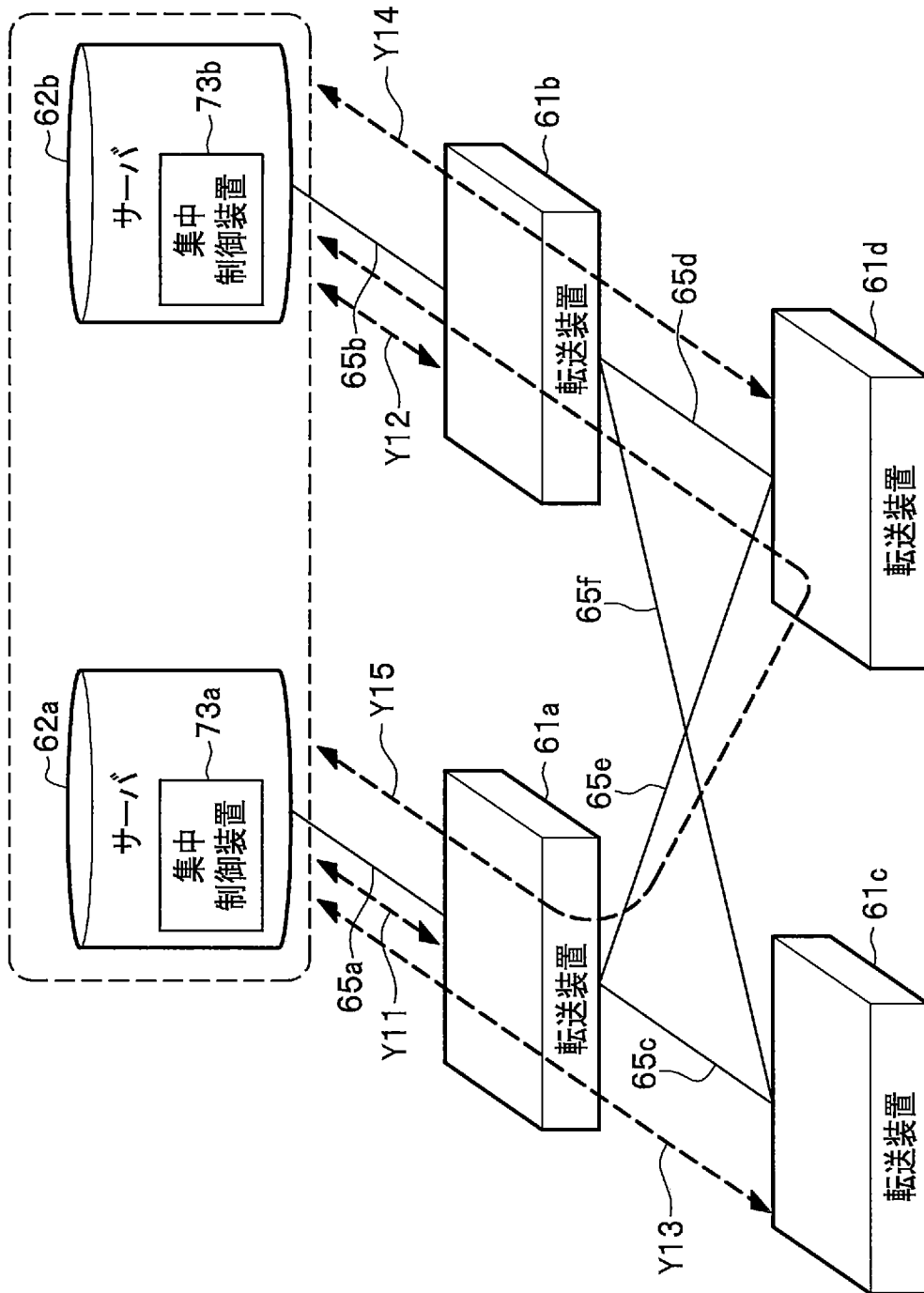
[図3]



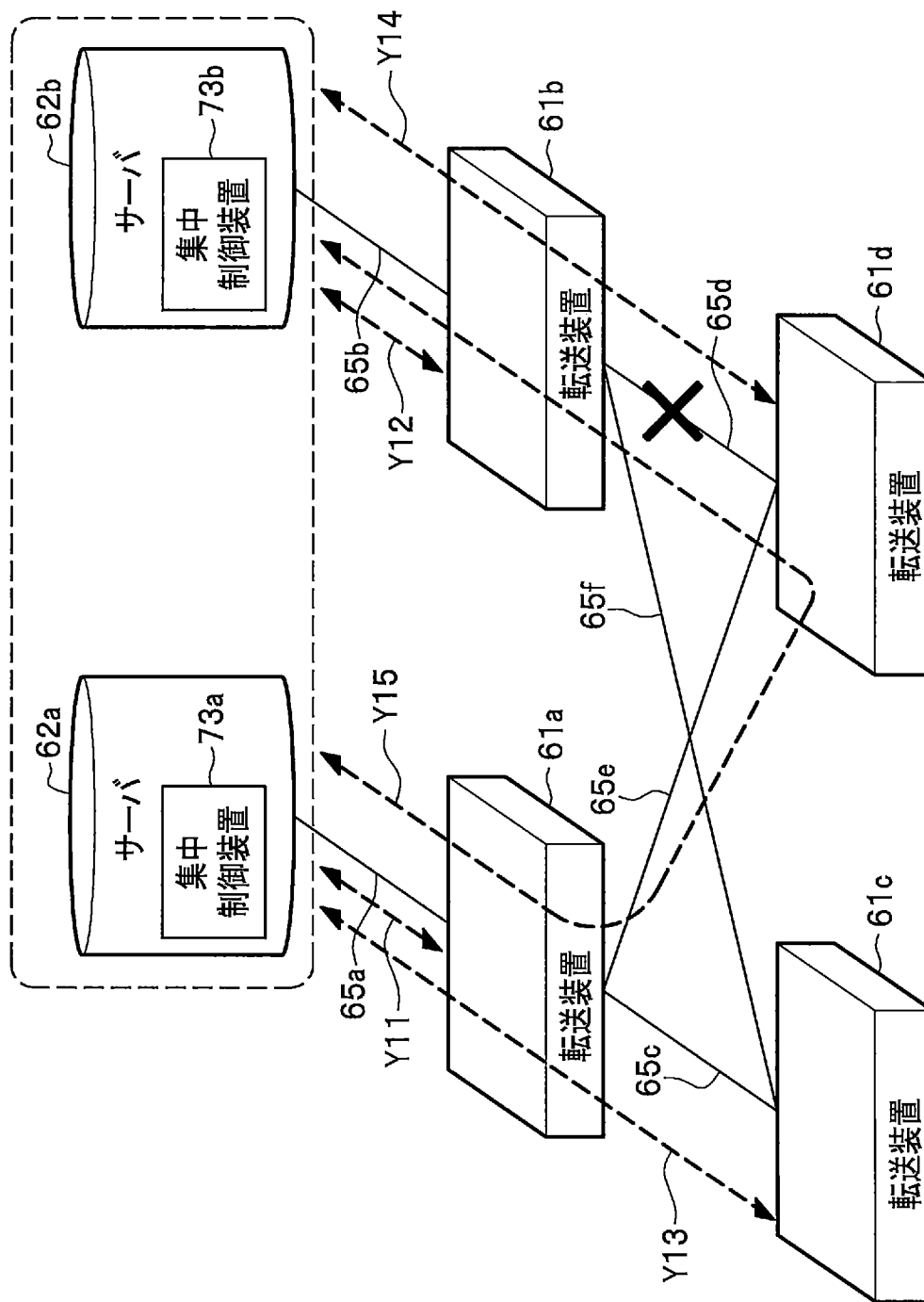
[図4]



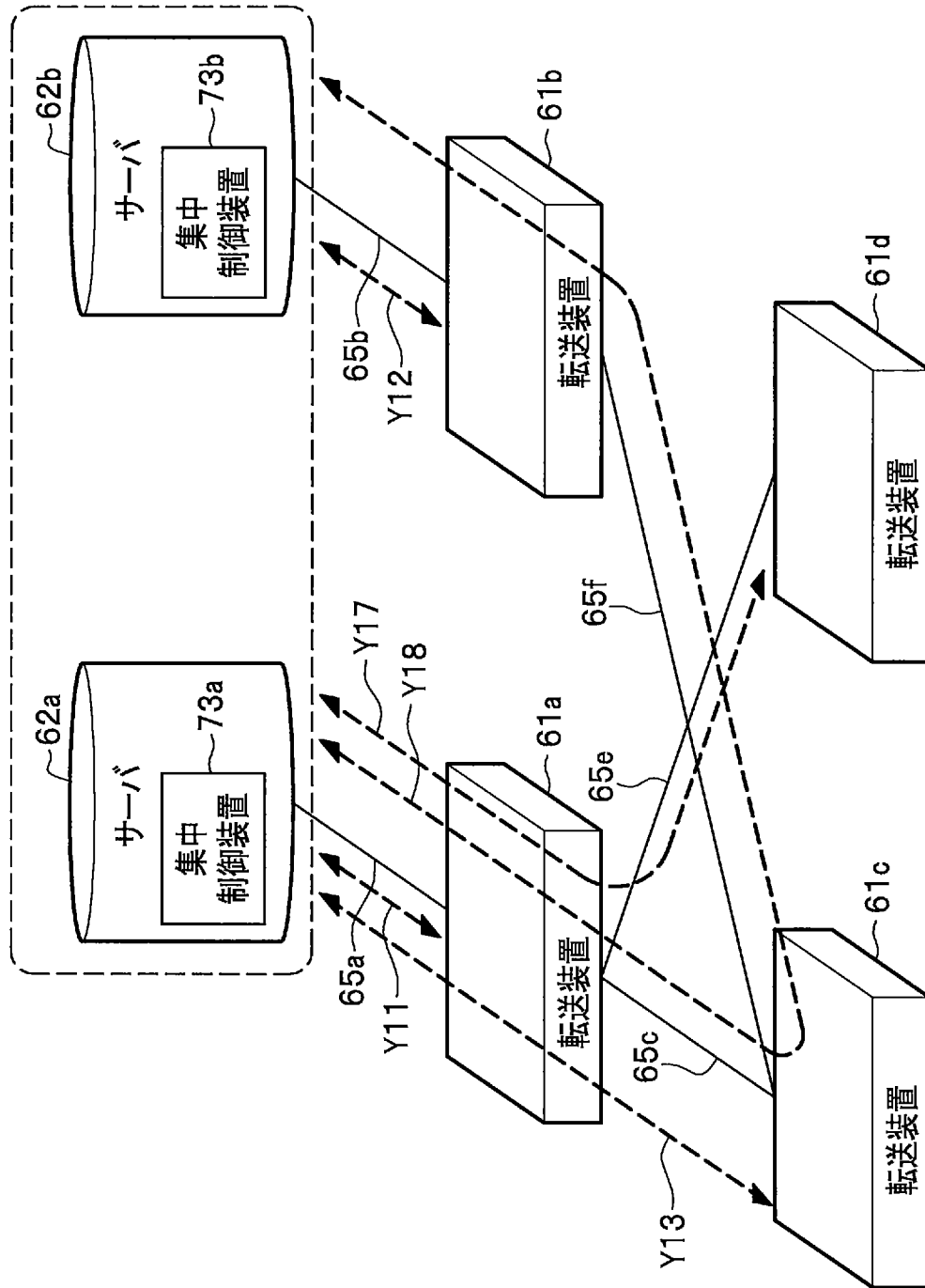
[図5]



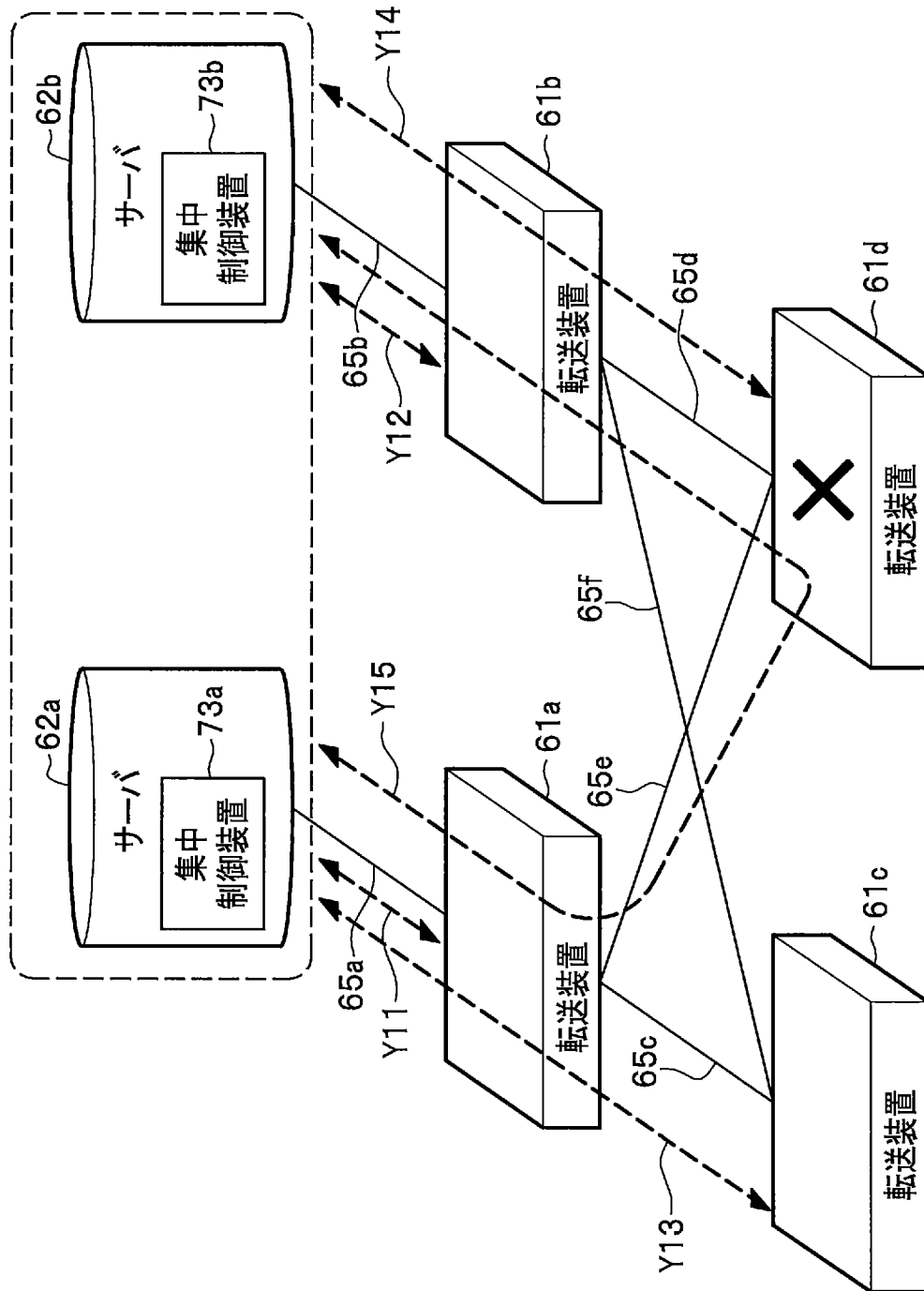
[図6]



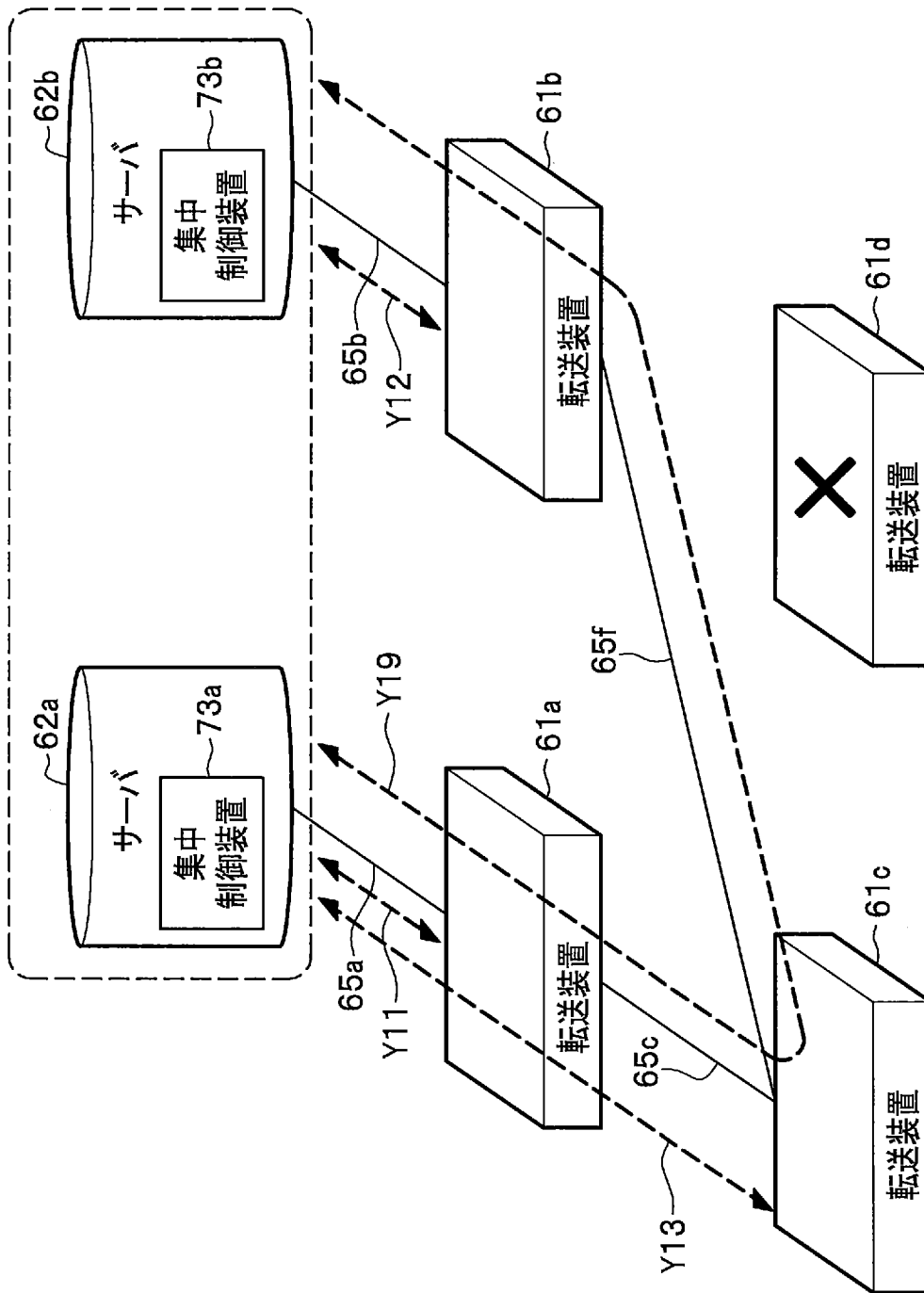
[図7]



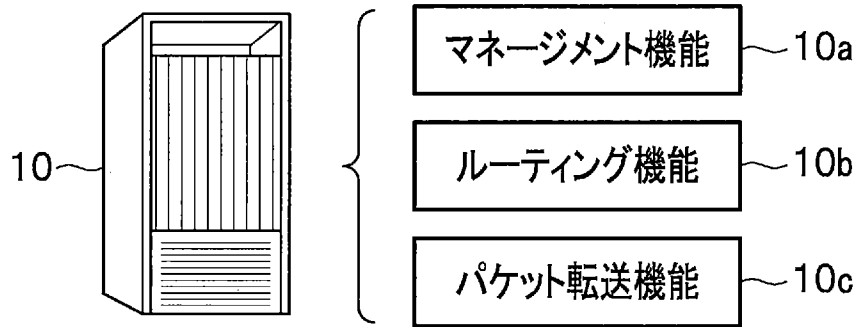
[図8]



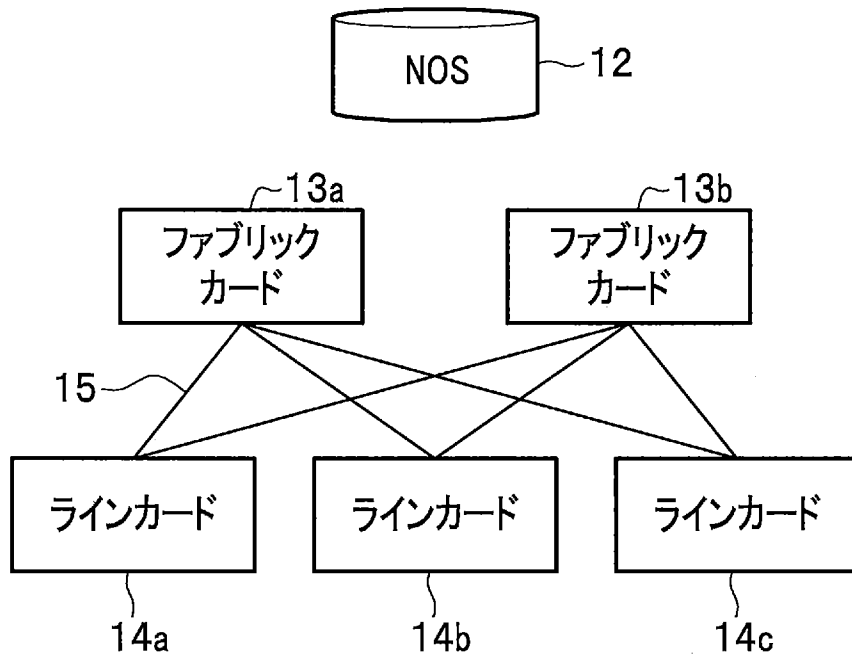
[図9]



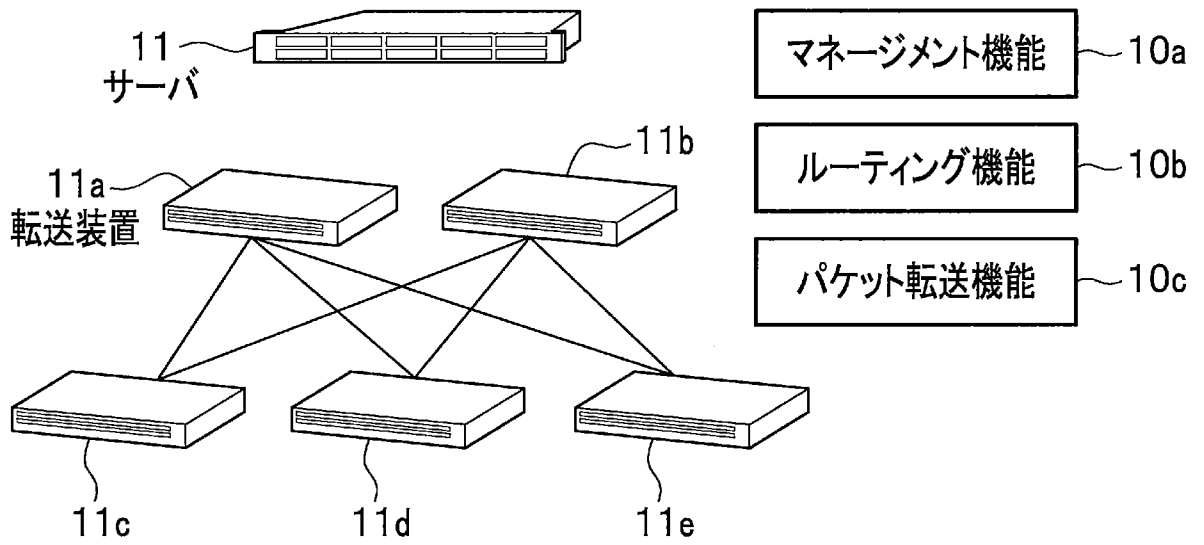
[図10]



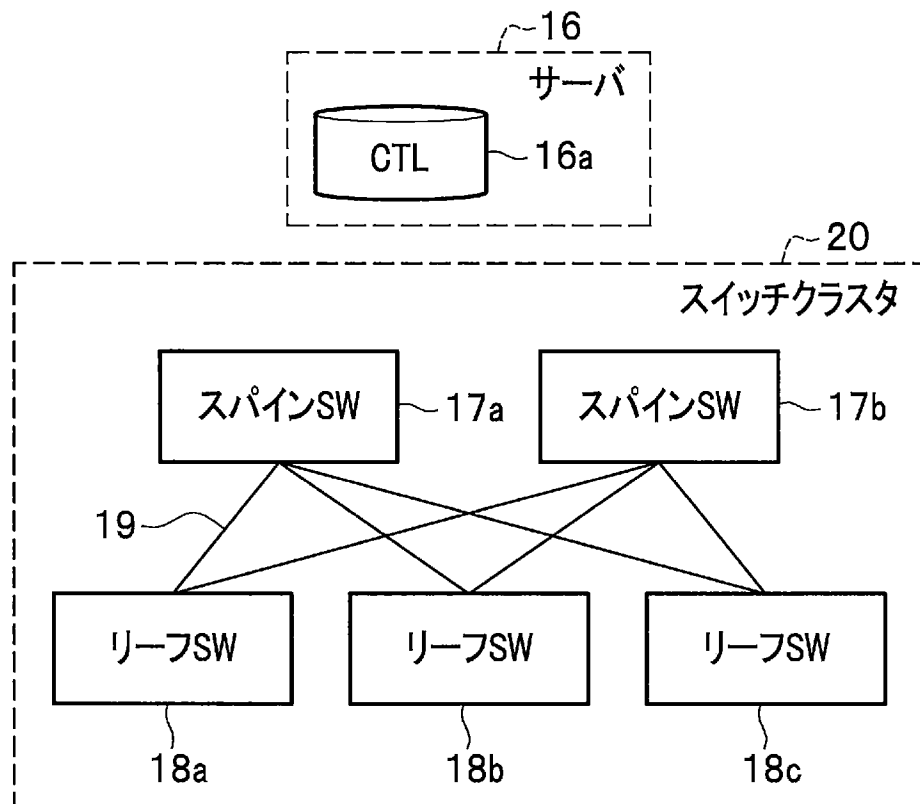
[図11]



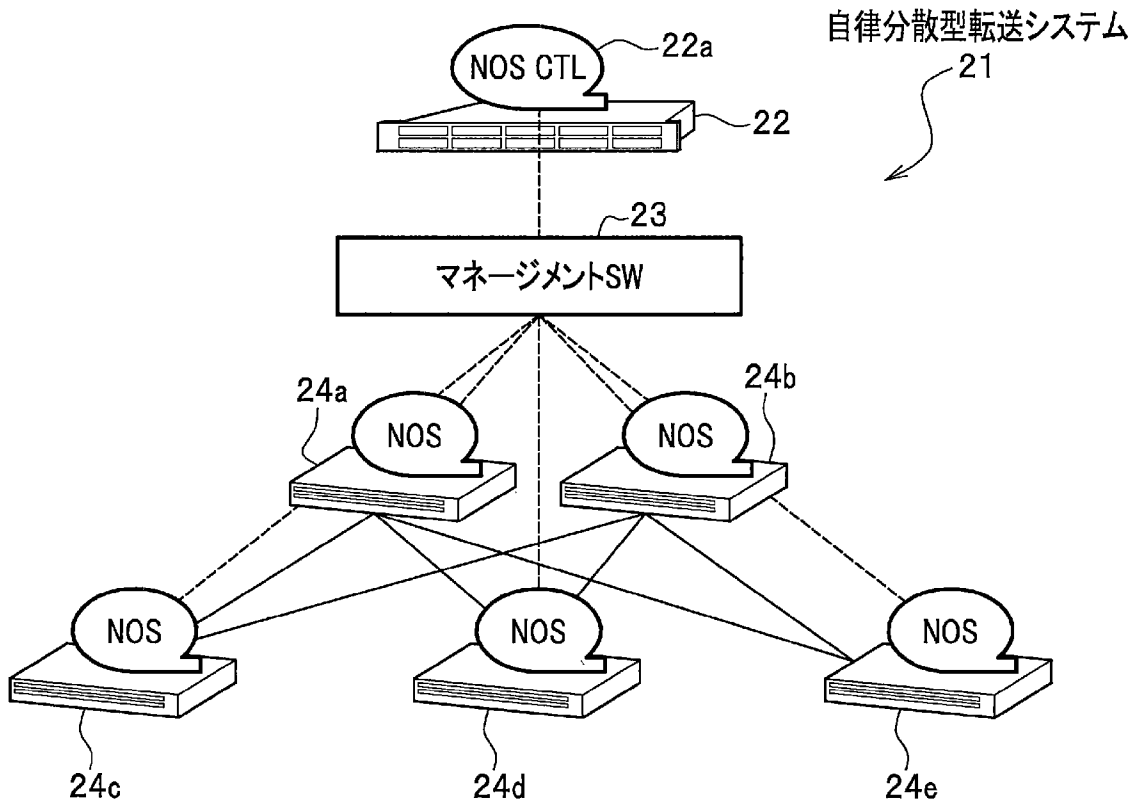
[図12]



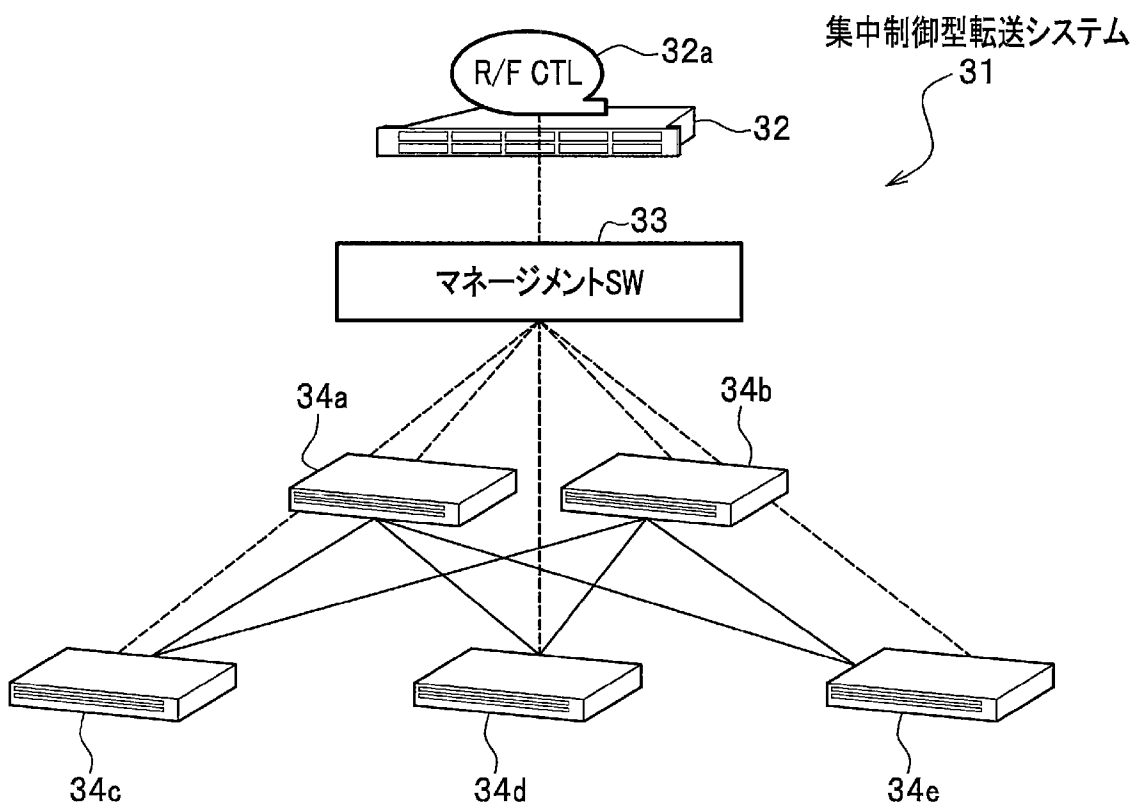
[図13]



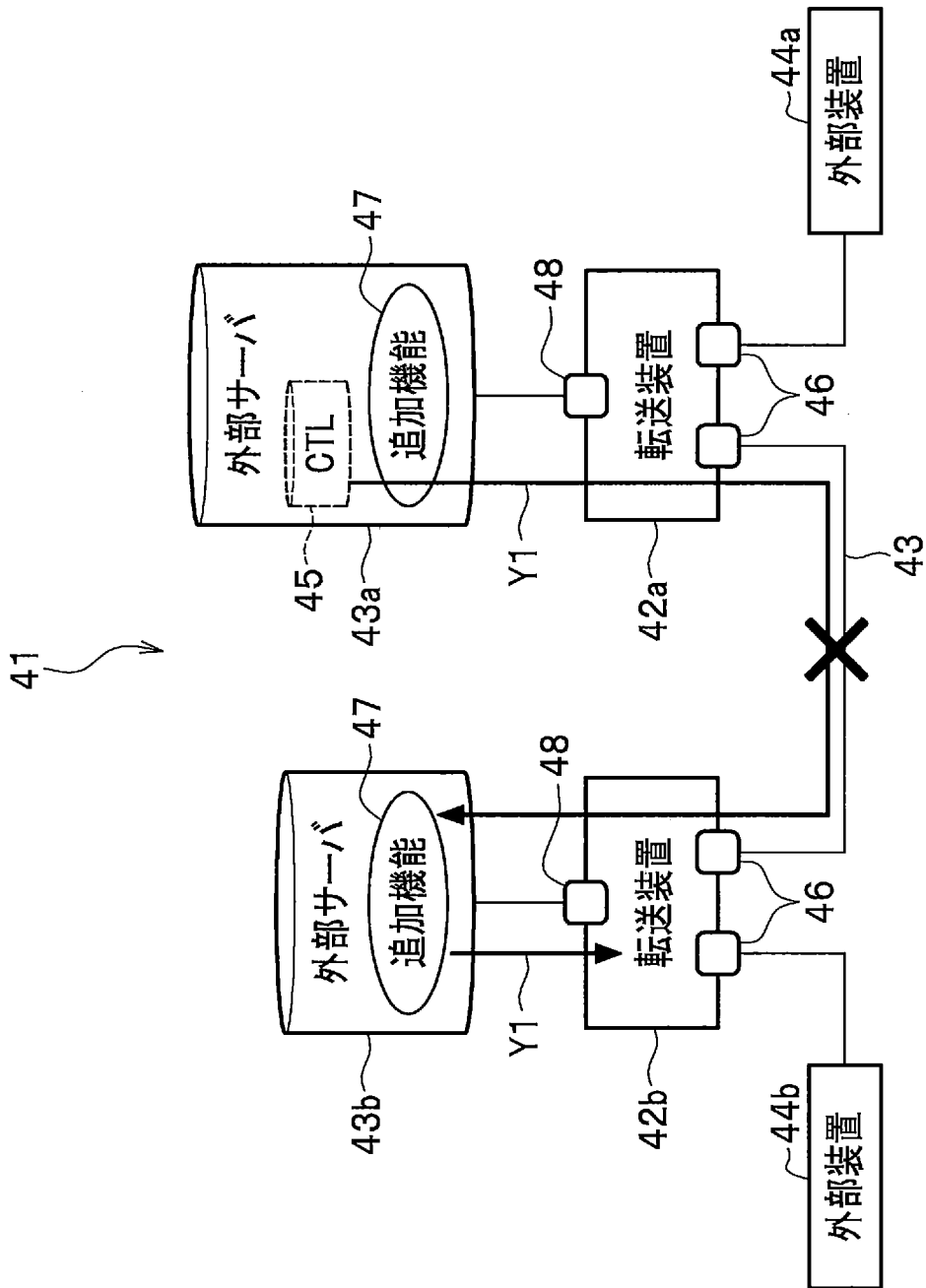
[図14]



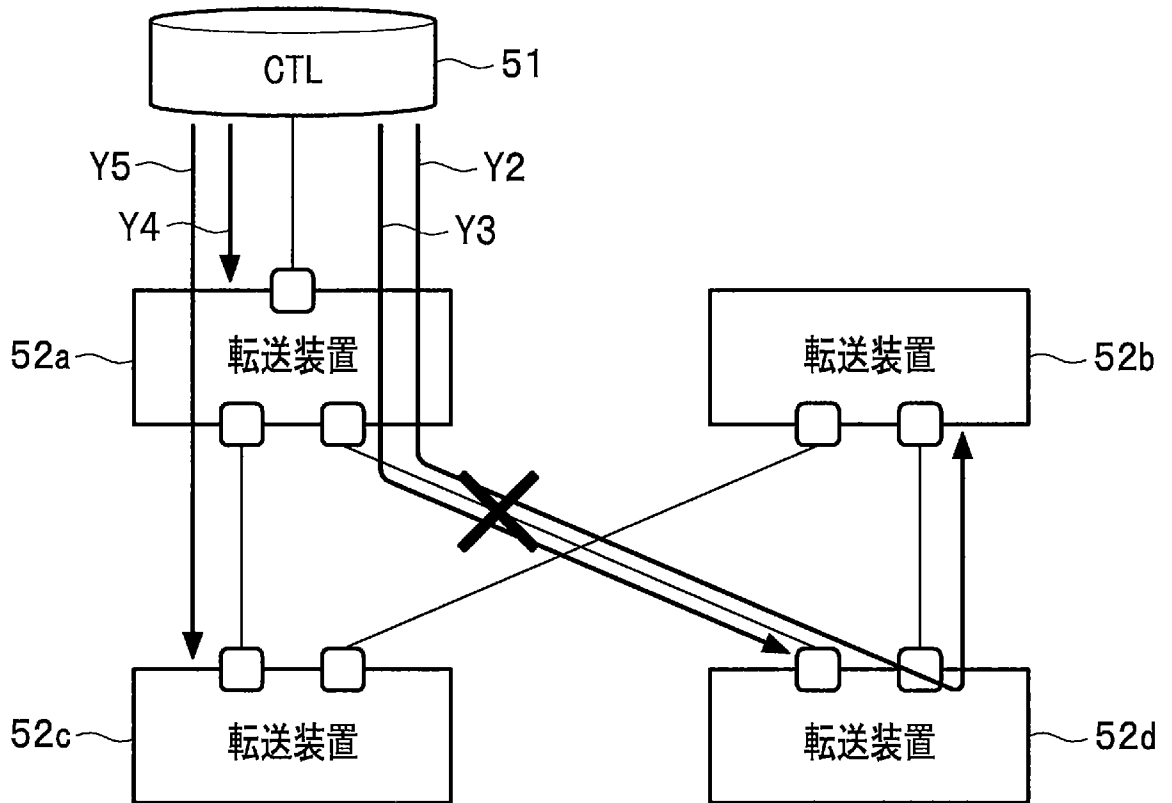
[図15]



[図16]



[図17]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/028883

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. H04L12/717 (2013.01) i, H04L12/707 (2013.01) 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)

Int.Cl. H04L12/717, H04L12/707

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

| | |
|--|-----------|
| Published examined utility model applications of Japan | 1922-1996 |
| Published unexamined utility model applications of Japan | 1971-2019 |
| Registered utility model specifications of Japan | 1996-2019 |
| Published registered utility model applications of Japan | 1994-2019 |

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 A | JP 2017-228935 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 28 December 2017, paragraphs [0004], [0036], [0037], [0041], [0045], [0046], [0049], [0050], [0059], [0060], [0065]-[0069], fig. 4, 5, 11 (Family: none) | 1-6 7, 8 |
| Y A | 平澤 崇佳 他, 分散連携型 Openflow コントローラにおける連携通信帯域の考察, 電子情報通信学会 2017 年通信ソサイエティ大会講演論文集 2, 29 August 2017, p. 19 (B-6-1), 「2. ネットワーク装置の提案装置制御手法」, fig. 1, (HIRASAWA, Takayoshi et al., Proceedings of the 2017 IEICE Society Conference 2), non-official translation (Consideration of cooperative communication band in distributed cooperative Openflow controller, "2. Proposed device control method for network device") | 1-6 7, 8 |

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

| | |
|---|--|
| * Special categories of cited documents: | "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 |
| "A" document defining the general state of the art which is not considered to be of particular relevance | "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 |
| "E" earlier application or patent but published on or after the international filing date | "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 |
| "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) | "&" document member of the same patent family |
| "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 | |

| | |
|---|--|
| Date of the actual completion of the international search 07.08.2019 | Date of mailing of the international search report 27.08.2019 |
|---|--|

| | |
|--|---|
| Name and mailing address of the ISA/ Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan | Authorized officer Telephone No. |
|--|---|

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/028883

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|-----------|---|-----------------------|
| Y A | JP 2014-160922 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 04 September 2014, paragraphs [0034], [0051]-[0054], fig. 7, 8 (Family: none) | 6 1-5, 7, 8 |
| A | JP 2015-80105 A (NEC CORPORATION) 23 April 2015 (Family: none) | 1-8 |
| A | 小野 敢一郎 他, マルチサービスファブリック (MSF) 新たなネットワーク構成・制御技術で柔軟な運用・サービス提供を実現, ビジネスコミュニケーション, 01 July 2015, vol. 52, no. 7, pp. 10-13, (ONO, Kanichiro et al., Business Communication), non-official translation (Multi service fabric (MSF) Realize flexible operation and service provision with new network configuration and control technology) | 1-8 |
| A | 芦田 優太 他, OpenFlow ネットワークにおけるドメイン縮約を用いた階層型ネットワーク構築手法の提案, 電子情報通信学会技術研究報告 (信学技報), 12 April 2012, vol. 112, no. 10, pp. 35-40 (CQ2012-7), (ASHIDA, Yuta et al., A hierarchical control method with domain aggregation on OpenFlow network, IEICE Technical Report) | 1-8 |

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/028883

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17 (2)(a) for the following reasons:

- 1. Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

- 2. Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

- 3. Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
[see extra sheet]

- 1. As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
- 2. As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
- 3. As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

- 4. No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/028883

(Continuation of Box No. III)

(Invention 1) Claims 1-5, 7, and 8

The invention as in claim 1 relates to "a transfer device for a switch cluster, which is configured to perform communication of path control through a main signal path with a centralized control device for performing centralized control on the switch cluster from the outside by a transfer device group connected by a communication path, comprising:

a separation unit for separating an internal packet for the switch cluster and an external packet for performing path control of the outside of the switch cluster, which are transmitted through the main signal path; and

an internal path control unit for performing path control of the inside of the switch cluster,

wherein the separation unit separates a path control packet for the inside of the switch cluster, and the internal path control unit performs path control in the switch cluster by the separated path control packet for the inside", and has a special technical feature "a transfer device for a switch cluster, which is configured to perform communication of path control through a main signal path with a centralized control device for performing centralized control on the switch cluster from the outside by a transfer device group connected by a communication path, wherein a path control packet for the inside of the switch cluster and an external packet for performing path control of the outside of the switch cluster are separated, and path control in the switch cluster is performed on the basis of the path control packet for the inside.

The invention as in claims 2-5 is dependent on the invention as in claim 1, and has the same special technical feature as claim 1.

The invention as in claims 6 and 7 relates to an invention of a transfer method by a transfer device for a switch cluster, which is configured to perform communication of path control through a main signal path with a centralized control device for performing centralized control on the switch cluster from the outside by a transfer device group connected by a communication path, or a program for causing a computer to function as such a transfer device, wherein a path control packet for the inside of the switch cluster and an external packet for the switch cluster are separated, and a bypass path is generated when the path control packet for the inside cannot be transmitted. The invention as in claims 6 and 7 has the same special technical feature as claim 1 in that path control is performed on the basis of a communication state of the path control packet for the inside.

Thus, the invention as in claims 1-5, 7, and 8 is classified as invention 1.

(Continuation of Box No. III)

(Invention 2) Claim 6

The invention as in claim 6 relates to "a transfer system, comprising: a switch cluster by a transfer device group connected by a communication path; and a centralized device for performing centralized control from the outside of the switch cluster, for performing communication of path control of the switch cluster through a main signal path by the centralized device, wherein

the centralized device performs path control to the outside such that the switch cluster is regarded as a single node from the outside of the switch cluster in accordance with a path control packet for the outside of the switch cluster".

The invention as in claim 1 and the invention as in claim 6 have a common technical feature in that "a switch cluster by a transfer device group connected by a communication path" and "a centralized device for performing centralized control from the outside of the switch cluster" are provided to "perform communication of path control through a main signal path".

However, the following document indicates that, in an open flow, in an open flow system comprising: a plurality of clusters (corresponding to "switch clusters" in claims 1 and 6) formed of a plurality of network devices (corresponding to a "transfer device" in claims 1 and 6); and a controller (corresponding to a "centralized device" in claims 1 and 6) for controlling a plurality of network devices handled by itself (clusters handled by itself), OF channel communication for exchanging path information between the controller and the network devices is implemented on D-plane (corresponding to a "main signal path" in claims 1 and 6). A matter that the system includes "a switch cluster by a transfer device group connected by a communication path" and a "centralized device for performing centralized control on the switch cluster from the outside", wherein "communication of path control is performed through a main signal path" cannot be considered to be a special technical feature.

There are no common technical features between the invention as in claim 1 and the invention as in claim 6.

Thus, the invention as in claim 6 is classified as invention 2.

Document: 平澤 崇佳 他, 分散連携型 Openflow コントローラにおける連携通信帯域の考察, 電子情報通信学会 2017 年通信ソサイエティ大会講演論文集 2, 29 August 2017, p. 19(B-6-1), 「2. ネットワーク装置の提案装置制御手法」, fig. 1, (HIRASAWA, Takayoshi et al., Proceedings of the 2017 IEICE Society Conference 2), non-official translation (Consideration of cooperative communication band in distributed cooperative Openflow controller, "2. Proposed device control method for network device")

| | | | |
|---|---|--|---------------|
| A. 発明の属する分野の分類（国際特許分類（IPC）） Int.Cl. H04L12/717(2013.01)i, H04L12/707(2013.01)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. H04L12/717, H04L12/707 | | | |
| 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 1971-2019年 日本国実用新案登録公報 1996-2019年 日本国登録実用新案公報 1994-2019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 A | JP 2017-228935 A（日本電信電話株式会社）2017.12.28, [0004], [0036]-[0037], [0041], [0045]-[0046], [0049]-[0050], [0059]-[0060], [0065]-[0069], Fig. 4, 5, 11（ファミリーなし） | 1-6 7, 8 | |
| Y A | 平澤 崇佳 他, 分散連携型Openflowコントローラにおける連携通信帯域の考察, 電子情報通信学会2017年通信ソサイエティ大会講演論文集2, 2017.08.29, p.19(B-6-1), 「2. ネットワーク装置の提案装置制御手法」, Fig.1 | 1-6 7, 8 | |
| <input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 | | <input type="checkbox"/> 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 07.08.2019 | | 国際調査報告の発送日 27.08.2019 | |
| 国際調査機関の名称及びあて先 日本国特許庁（ISA/J P） 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） 玉木 宏治 | 5 X 3 0 4 7 |
| | | 電話番号 03-3581-1101 内線 | 3 5 9 6 |

| 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 A | JP 2014-160922 A (日本電信電話株式会社) 2014.09.04, [0034], [0051]-[0054], Fig. 7, 8 (ファミリーなし) | 6 1-5, 7, 8 |
| A | JP 2015-80105 A (日本電気株式会社) 2015.04.23, (ファミリーなし) | 1-8 |
| A | 小野 敢一郎 他, マルチサービスファブリック(MSF) 新たなネットワーク構成・制御技術で柔軟な運用・サービス提供を実現, ビジネスコミュニケーション, 2015.07.01, 第52巻, 第7号, pp. 10-13 | 1-8 |
| A | 芦田 優太 他, OpenFlow ネットワークにおけるドメイン縮約を用いた階層型ネットワーク構築手法の提案, 電子情報通信学会技術研究報告(信学技報), 2012.04.12, 第112巻, 第10号, pp. 35-40(CQ2012-7) | 1-8 |

第II欄 請求の範囲の一部の調査ができないときの意見 (第1ページの2の続き)

法第8条第3項 (PCT17条(2)(a))の規定により、この国際調査報告は次の理由により請求の範囲の一部について作成しなかった。

1. 請求項 _____ は、この国際調査機関が調査をすることを要しない対象に係るものである。つまり、

2. 請求項 _____ は、有意義な国際調査をすることができる程度まで所定の要件を満たしていない国際出願の部分に係るものである。つまり、

3. 請求項 _____ は、従属請求の範囲であってPCT規則6.4(a)の第2文及び第3文の規定に従って記載されていない。

第III欄 発明の単一性が欠如しているときの意見 (第1ページの3の続き)

次に述べるようにこの国際出願に二以上の発明があるときの国際調査機関は認めた。
(特別ページ参照)

1. 出願人が必要な追加調査手数料をすべて期間内に納付したので、この国際調査報告は、すべての調査可能な請求項について作成した。
2. 追加調査手数料を要求するまでもなく、すべての調査可能な請求項について調査することができたので、追加調査手数料の納付を求めなかった。
3. 出願人が必要な追加調査手数料を一部のみしか期間内に納付しなかったため、この国際調査報告は、手数料の納付のあった次の請求項のみについて作成した。
4. 出願人が必要な追加調査手数料を期間内に納付しなかったため、この国際調査報告は、請求の範囲の最初に記載されている発明に係る次の請求項について作成した。

追加調査手数料の異議の申立てに関する注意

- 追加調査手数料及び、該当する場合には、異議申立手数料の納付と共に、出願人から異議申立てがあった。
- 追加調査手数料の納付と共に出願人から異議申立てがあったが、異議申立手数料が納付命令書に示した期間内に支払われなかった。
- 追加調査手数料の納付はあったが、異議申立てはなかった。

(発明1) 請求項1-5、7、8

請求項1に係る発明は、

「通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置であって、

前記主信号経路を伝送する、前記スイッチクラスタの内部用パケットと、当該スイッチクラスタの外部の経路制御を行う外部用パケットとを分離する分離部と、

前記スイッチクラスタの内部の経路制御を行う内部経路制御部とを備え、

前記分離部は前記スイッチクラスタの内部向けの経路制御パケットを分離し、前記内部経路制御部は、当該分離された内部向けの経路制御パケットで前記スイッチクラスタ内の経路制御を行う

ことを特徴とする転送装置。」

というものであり、通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置において、スイッチクラスタの内部向けの経路制御パケットと、当該スイッチクラスタの外部の経路制御を行う外部用パケットとを分離して、内部向けの経路制御パケットに基づいてスイッチクラスタ内の経路制御を行う、という特別な技術的特徴を有する。

請求項2-5に係る発明は、請求項1に係る発明の従属項であり、請求項1と同一の特別な技術的特徴を有する。

請求項6、7に係る発明は、通信の経路で接続された転送装置群によるスイッチクラスタに外部から集中制御を行う集中制御装置と、主信号経路を介して経路制御の通信を行う当該スイッチクラスタの転送装置による転送方法、又は、そのような転送装置としてコンピュータを機能させるためのプログラムの発明であって、スイッチクラスタの内部向けの経路制御パケットと、当該スイッチクラスタの外部用パケットとを分離して、内部向けの経路制御パケットの不通時に迂回経路を生成するというものであり、内部向けの経路制御パケットの通信状態に基づいて経路制御を行っているという点で、請求項1と同一の特別な技術的特徴を有する。

したがって、請求項1-5、7、8に係る発明を発明1に区分する。

(次の特別ページに続く。)

(発明 2) 請求項 6

請求項 6 に係る発明は、

「通信の経路で接続された転送装置群によるスイッチクラスタと、当該スイッチクラスタの外部から集中制御を行う集中制御装置とを有し、当該集中制御装置で主信号経路を介して当該スイッチクラスタの経路制御の通信を行う転送システムであって、

前記集中制御装置は、

前記スイッチクラスタの外部向けの経路制御パケットに応じて、当該スイッチクラスタの外部からこのスイッチクラスタが単一ノードと見做せるように、当該外部との経路制御を行うことを特徴とする転送システム。」

というものである。

請求項 1 に係る発明と、請求項 6 に係る発明とは、「通信の経路で接続された転送装置群によるスイッチクラスタ」と、「スイッチクラスタに外部から集中制御を行う集中制御装置」とを有し、「主信号経路を介して経路制御の通信を行う」ものである点で、共通の技術的特徴を有する。

しかしながら、下記の文献には、オープンフローにおいて、複数のネットワーク装置（請求項 1、6 における「転送装置」に相当）からなる複数のクラスタ（同「スイッチクラスタ」）と、自身の担当する複数のネットワーク装置（自身の担当するクラスタ）を制御するコントローラ（同「集中制御装置」）とからなるオープンフローシステムにおいて、コントローラネットワーク装置間の経路情報をやり取りする OFchannel 通信を D-plane（同「主信号経路」）上で実現することが記載されており、「通信の経路で接続された転送装置群によるスイッチクラスタ」と、「スイッチクラスタに外部から集中制御を行う集中制御装置」とを有し、「主信号経路を介して経路制御の通信を行う」ものである点は、特別な技術的特徴とはいえない。

請求項 1 に係る発明と、請求項 6 に係る発明との間に、他に共通の技術的特徴は存在しない。

したがって、請求項 6 に係る発明を発明 2 に区分する。

文献：平澤 崇佳 他，分散連携型 Openflow コントローラにおける連携通信帯域の考察，電子情報通信学会 2017 年通信ソサイエティ大会講演論文集 2，2017 年 8 月 29 日，p. 19 (B-6-1)，特に、「2. ネットワーク装置の提案装置制御手法」，Fig. 1