

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6095583号
(P6095583)

(45) 発行日 平成29年3月15日 (2017. 3. 15)

(24) 登録日 平成29年2月24日 (2017. 2. 24)

(51) Int. Cl.

F I

H O 4 L 12/42 (2006. 01)

H O 4 L 12/42 M

H O 4 L 12/437 (2006. 01)

H O 4 L 12/437 Z

請求項の数 5 (全 36 頁)

(21) 出願番号 特願2014-2571 (P2014-2571)
 (22) 出願日 平成26年1月9日 (2014. 1. 9)
 (65) 公開番号 特開2015-133538 (P2015-133538A)
 (43) 公開日 平成27年7月23日 (2015. 7. 23)
 審査請求日 平成28年3月17日 (2016. 3. 17)

(73) 特許権者 504411166
 アラクサラネットワークス株式会社
 神奈川県川崎市幸区鹿島田一丁目1番2号
 (74) 代理人 110001678
 特許業務法人藤央特許事務所
 (72) 発明者 持田 茂雄
 神奈川県川崎市幸区鹿島田一丁目1番2号
 アラクサラネットワークス株式会社内
 (72) 発明者 田中 啓久
 神奈川県川崎市幸区鹿島田一丁目1番2号
 アラクサラネットワークス株式会社内

審査官 宮島 郁美

最終頁に続く

(54) 【発明の名称】 データ転送装置システム及びネットワークシステム

(57) 【特許請求の範囲】

【請求項 1】

リング状のネットワークシステムに属するデータ転送装置システムであって、
 前記データ転送装置システムは、複数のデータ転送装置を含み、
 前記複数のデータ転送装置は第1のデータ転送装置及び前記第1のデータ転送装置に接
 続される第2のデータ転送装置を含み、
 前記第1のデータ転送装置は、前記ネットワークシステムに属し、前記データ転送装置
 システムに含まれない第3のデータ転送装置に接続される第1のポートを有し、
 前記第2のデータ転送装置は、前記ネットワークシステムに属し、前記データ転送装置
 システムに含まれない第4のデータ転送装置に接続される第2のポートを有し、
 前記第1のデータ転送装置は、前記ネットワークシステムの状態、前記データ転送装置
 システムにおける前記第1のデータ転送装置の属性、及び、前記ネットワークシステムに
 おける前記データ転送装置システムの属性に基づいて、前記第1のポート及び前記第2の
 ポートのデータ転送可否を制御し、

前記ネットワークシステムの状態は、前記ネットワークシステム内の前記第2のデー
 タ転送装置若しくはその他の部分に障害が発生している障害状態、又は、障害が発生してい
 ない正常状態のいずれかであり、

前記ネットワークシステムにおける前記データ転送装置システムの属性は、前記ネット
 ワークシステムの障害を監視するマスタノード、又は、前記ネットワークシステムの障害
 を監視しないトランジットノードであり、

前記データ転送装置システムにおける前記第1のデータ転送装置の属性は、前記第1のデータ転送装置及び前記第2のデータ転送装置によるデータ転送を制御し、前記データ転送装置システムが前記マスタノードである場合には前記マスタノードの処理を実行するマスタ装置、又は、前記マスタ装置である前記第2のデータ転送装置によって制御され、前記第2のデータ転送装置に障害が発生した場合に新たに前記マスタ装置となるバックアップ装置であり、

前記データ転送装置システムにおける前記第1のデータ転送装置の属性が前記バックアップ装置であり、前記ネットワークシステムにおける前記データ転送装置システムの属性が前記マスタノードである場合において、前記第2のデータ転送装置の障害が検出されたとき、前記第1のデータ転送装置は、前記データ転送装置システムにおける前記第1のデータ転送装置の属性を前記マスタ装置に変更し、前記第1のポートを、ユーザデータを転送可能な状態に制御し、

10

前記データ転送装置システムにおける前記第1のデータ転送装置の属性が前記マスタ装置であり、前記ネットワークシステムにおける前記データ転送装置システムの属性が前記マスタノードである場合において、前記ネットワークシステムの状態が前記障害状態であることが検出されたとき、前記第1のデータ転送装置は、前記第1のポート及び前記第2のポートを、ユーザデータを転送可能な状態に制御し、

前記第1のデータ転送装置は、前記ネットワークシステムの状態が前記障害状態から前記正常状態に復旧したことを検出した場合、前記第1のポート及び前記第2のポートのうち一方を、ユーザデータを転送可能な状態に制御し、他方を、ユーザデータを転送可能でない状態に制御することを特徴とするデータ転送装置システム。

20

【請求項2】

請求項1に記載のデータ転送装置システムであって、

前記データ転送装置システムにおける前記第1のデータ転送装置の属性が前記バックアップ装置であり、前記ネットワークシステムにおける前記データ転送装置システムの属性が前記トランジットノードである場合において、前記第2のデータ転送装置の障害が検出されたとき、前記第1のデータ転送装置は、前記データ転送装置システムにおける前記第1のデータ転送装置の属性を前記マスタ装置に変更することを特徴とするデータ転送装置システム。

【請求項3】

30

請求項1に記載のデータ転送装置システムであって、

前記第1のデータ転送装置は、第3のポートをさらに有し、

前記第2のデータ転送装置は、前記第3のポートに接続される第4のポートをさらに有し、

前記第1のデータ転送装置及び前記第2のデータ転送装置は、それぞれ、前記第3のポート及び前記第4のポートから第1の制御データを相互に送信し、

前記第1のデータ転送装置は、前記第2のデータ転送装置から前記第1の制御データを受信したか否かに基づいて、前記第2のデータ転送装置の障害を検出し、

前記データ転送装置システムにおける前記第1のデータ転送装置の属性が前記マスタ装置であり、前記ネットワークシステムにおける前記データ転送装置システムの属性が前記マスタノードである場合、前記第1のデータ転送装置は、前記第1のポート及び前記第3のポートから第2の制御データを送信し、前記第2の制御データを受信したか否かに基づいて、前記ネットワークシステムの状態を判定することを特徴とするデータ転送装置システム。

40

【請求項4】

請求項1に記載のデータ転送装置システムであって、

前記第1のデータ転送装置及び前記第2のデータ転送装置は、それぞれ、第3のポート及び前記第3のポートに接続される第4のポートを有し、

前記第1のデータ転送装置は、前記第1のデータ転送装置の前記第1のポート及び前記第2のデータ転送装置の前記第2のポートがそれぞれ前記ネットワークシステムにおいて

50

前記データ転送装置システムに隣接するデータ転送装置に接続されることを示す情報、及び、前記第1のデータ転送装置の前記第3のポート及び前記第2のデータ転送装置の前記第4のポートが相互に接続されることを示す情報を含む設定情報を保持し、

前記第2のデータ転送装置が前記データ転送装置システムから削除された場合、前記第1のデータ転送装置は、前記第1のデータ転送装置の前記第3のポート及び前記第2のデータ転送装置の前記第4のポートが相互に接続されることを示す情報を削除し、前記第1のデータ転送装置のポートのうち前記第1のポート以外のポートが前記データ転送装置システムに隣接するデータ転送装置に接続されるように前記設定情報を更新し、

前記第2のデータ転送装置が前記データ転送装置システムに追加された場合、前記第1のデータ転送装置は、前記第1のデータ転送装置の前記第3のポート及び前記第2のデータ転送装置の前記第4のポートが相互に接続されることを示す情報を追加し、前記第2のデータ転送装置の前記第2のポートが前記データ転送装置システムに隣接するデータ転送装置に接続されるように前記設定情報を更新することを特徴とするデータ転送装置システム。

10

【請求項5】

リング状に接続され、各々が一つ以上のデータ転送装置からなる複数のノードを含むネットワークシステムであって、

前記複数のノードのうち第1のノードは、第1のデータ転送装置及び前記第1のデータ転送装置に接続される第2のデータ転送装置を含み、

前記第1のデータ転送装置は、前記複数のノードのうち第2のノードに接続される第1のポートを有し、

20

前記第2のデータ転送装置は、前記複数のノードのうち第3のノードに接続される第2のポートを有し、

前記第1のデータ転送装置は、前記ネットワークシステムの状態、前記第1のノードにおける前記第1のデータ転送装置の属性、及び、前記ネットワークシステムにおける前記第1のノードの属性に基づいて、前記第1のポート及び前記第2のポートのデータ転送可否を制御し、

前記ネットワークシステムの状態は、前記ネットワークシステム内の前記第2のデータ転送装置若しくはその他の部分に障害が発生している障害状態、又は、障害が発生していない正常状態のいずれかであり、

30

前記ネットワークシステムにおける前記第1のノードの属性は、前記ネットワークシステムの障害を監視するマスタノード、又は、前記ネットワークシステムの障害を監視しないトランジットノードであり、

前記第1のノードにおける前記第1のデータ転送装置の属性は、前記第1のデータ転送装置及び前記第2のデータ転送装置によるデータ転送を制御し、前記第1のノードが前記マスタノードである場合には前記マスタノードの処理を実行するマスタ装置、又は、前記マスタ装置である前記第2のデータ転送装置によって制御され、前記第2のデータ転送装置に障害が発生した場合に新たに前記マスタ装置となるバックアップ装置であり、

前記第1のノードにおける前記第1のデータ転送装置の属性が前記バックアップ装置であり、前記ネットワークシステムにおける前記第1のノードの属性が前記マスタノードである場合において、前記第2のデータ転送装置の障害が検出されたとき、前記第1のデータ転送装置は、前記第1のノードにおける前記第1のデータ転送装置の属性を前記マスタ装置に変更し、前記第1のポートを、ユーザデータを転送可能な状態に制御し、

40

前記第1のノードにおける前記第1のデータ転送装置の属性が前記マスタ装置であり、前記ネットワークシステムにおける前記第1のノードの属性が前記マスタノードである場合において、前記ネットワークシステムの状態が前記障害状態であることが検出されたとき、前記第1のデータ転送装置は、前記第1のポート及び前記第2のポートを、ユーザデータを転送可能な状態に制御し、

前記第1のデータ転送装置は、前記ネットワークシステムの状態が前記障害状態から前記正常状態に復旧したことを検出した場合、前記第1のポート及び前記第2のポートのう

50

ち一方を、ユーザデータを転送可能な状態に制御し、他方を、ユーザデータを転送可能でない状態に制御することを特徴とするネットワークシステム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、データ転送装置システム及びそれを含むネットワークシステムに関する。

【背景技術】

【0002】

本技術分野の背景技術として、特開2008-136013号公報（特許文献1）、特開2013-102546号公報（特許文献2）及び特開2013-51481号公報（特許文献3）がある。

10

【0003】

特許文献1には、「リングネットワークを構成する装置として、監視装置とその他の中継装置という構成を取る。監視装置は自装置が属するリングネットワークのリング状態のみを監視する。監視装置は、リングポートからそれぞれヘルスチェックフレームを周期的に送信し、逆側のリングポートでこのヘルスチェックフレームを受信しているか否かを監視する。また、共有リンクありのマルチリングネットワークの構成では、補助ヘルスチェックフレームを2つある共有装置それぞれから共有リンク非監視リングの監視装置に対して周期的に送信する。監視装置は、ヘルスチェックフレームと共に、補助ヘルスチェックフレームも監視することで複数のリングネットワークを跨ったループを防止する。」と記載されている。

20

【0004】

特許文献2には、「複数の装置を連携させて仮想的に1台に見立てることができるパケット中継装置において、リンク最適化調整やパケット解析部を具備することによって、通信の流れ方や通信量などを自動的に判断し、主に稼働させる回線を最適化することによって、制御ポートに流れるトラフィックを管理して、制御ポートの帯域不足を補う。」と記載されている。

【0005】

特許文献3には、「旧監視ノードを備えるリングネットワークにおいて、旧監視ノードを新監視ノードに入れ替える。この入れ替えの際に、新監視ノードは、両リングポートがリンクアップの状態となったことが検出されると、中継ノードに対して、旧監視ノードのリングポートから送られたヘルスチェックフレームが新監視ノード側に転送されるように指示する。この結果、新監視ノードは、ヘルスチェックフレームを受信することにより、リングネットワークの監視を行う。」と記載されている。

30

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2008-136013号公報

【特許文献2】特開2013-102546号公報

【特許文献3】特開2013-51481号公報

40

【発明の概要】

【発明が解決しようとする課題】

【0007】

特許文献1又は3に記載されたリングネットワークにおいて、一つのリングは複数の中継装置をリング状に接続することによって形成される。一つのリングを構成する複数の中継装置のうち一つは、監視機能を有する中継装置（特許文献1では「監視装置」、特許文献3では「監視ノード」と記載）であり、残りは、監視機能を有しない中継装置（特許文献1では単に「中継装置」、特許文献3では「中継ノード」と記載）である。監視装置は、リングに接続される二つのリングポートのそれぞれからヘルスチェックフレームを送信することでリングの障害を監視し、障害が検出された場合には、リングポートの状態を変

50

更する。これによって、リングを通過するデータの通信経路が障害箇所を回避する経路に切り替えられる。しかし、監視装置に障害が発生した場合には、ヘルスチェックフレームを用いたリングの障害監視機能が停止し、経路の切り替えができなくなる。

【0008】

一方、特許文献2には、物理的に複数の中継装置を仮想的に一つの中継装置として扱うことによって中継装置の耐障害性を改善する、装置間スタックの技術が開示されている。この技術の特許文献1又は3に記載されたリングネットワークに適用することによって、リングネットワークの耐障害性が向上することを発明者らは見出した。

【0009】

しかし、特許文献1又は3に記載されたリングネットワークに、特許文献2に記載の装置間スタックの技術を適用しようとする、例えば、物理的に二つの中継装置をリングネットワークの仮想的な一つの監視装置として使用する、仮想的な一つの監視装置の二つのリングポートは、物理的には二つの中継装置に分散して配置されることとなる。そのため、一つの中継装置による監視を前提としている従来のリングネットワークの制御方法では、二つの中継装置に分散して配置されるリングポートを制御することができない。

【0010】

また、特許文献2に記載された装置間スタックの技術の制御方法は、物理的に複数の中継装置のうちの一つの中継装置が他の中継装置の動作を制御することで、仮想的に一つの中継装置として動作し冗長化を図っているため、仮想的な一つの中継装置を構成する他の中継装置は監視できるが、仮想的な一つの中継装置を構成する他の中継装置以外を監視することはできない。

【0011】

従って、本発明は、例えば、リングネットワークを構成する一ノードの冗長性を向上しつつ、リングネットワーク内でのループを防止するとともに、リングの障害を監視し通信の継続性を担保することを目的とする。

【課題を解決するための手段】

【0012】

上記の課題の少なくとも一を解決するために、本発明の一態様では、リング状のネットワークシステムに属するデータ転送装置システムであって、前記データ転送装置システムは、複数のデータ転送装置を含み、前記複数のデータ転送装置は第1のデータ転送装置及び前記第1のデータ転送装置に接続される第2のデータ転送装置を含み、前記第1のデータ転送装置は、前記ネットワークシステムに属し、前記データ転送装置システムに含まれない第3のデータ転送装置に接続される第1のポートを有し、前記第2のデータ転送装置は、前記ネットワークシステムに属し、前記データ転送装置システムに含まれない第4のデータ転送装置に接続される第2のポートを有し、前記第1のデータ転送装置は、前記ネットワークシステムの状態、前記データ転送装置システムにおける前記第1のデータ転送装置の属性、及び、前記ネットワークシステムにおける前記データ転送装置システムの属性に基づいて、前記第1のポート及び前記第2のポートのデータ転送可否を制御し、前記ネットワークシステムの状態は、前記ネットワークシステム内の前記第2のデータ転送装置若しくはその他の部分に障害が発生している障害状態、又は、障害が発生していない正常状態のいずれかであり、前記ネットワークシステムにおける前記データ転送装置システムの属性は、前記ネットワークシステムの障害を監視するマスタノード、又は、前記ネットワークシステムの障害を監視しないトランジットノードであり、前記データ転送装置システムにおける前記第1のデータ転送装置の属性は、前記第1のデータ転送装置及び前記第2のデータ転送装置によるデータ転送を制御し、前記データ転送装置システムが前記マスタノードである場合には前記マスタノードの処理を実行するマスタ装置、又は、前記マスタ装置である前記第2のデータ転送装置によって制御され、前記第2のデータ転送装置に障害が発生した場合に新たに前記マスタ装置となるバックアップ装置であり、前記データ転送装置システムにおける前記第1のデータ転送装置の属性が前記バックアップ装置であり、前記ネットワークシステムにおける前記データ転送装置システムの属性が前記マ

10

20

30

40

50

タノードである場合において、前記第 2 のデータ転送装置の障害が検出されたとき、前記第 1 のデータ転送装置は、前記データ転送装置システムにおける前記第 1 のデータ転送装置の属性を前記マスタ装置に変更し、前記第 1 のポートを、ユーザデータを転送可能な状態に制御し、前記データ転送装置システムにおける前記第 1 のデータ転送装置の属性が前記マスタ装置であり、前記ネットワークシステムにおける前記データ転送装置システムの属性が前記マスタノードである場合において、前記ネットワークシステムの状態が前記障害状態であることが検出されたとき、前記第 1 のデータ転送装置は、前記第 1 のポート及び前記第 2 のポートを、ユーザデータを転送可能な状態に制御し、前記第 1 のデータ転送装置は、前記ネットワークシステムの状態が前記障害状態から前記正常状態に復旧したことを検出した場合、前記第 1 のポート及び前記第 2 のポートのうち一方を、ユーザデータを転送可能な状態に制御し、他方を、ユーザデータを転送可能でない状態に制御することを特徴とする。

10

【発明の効果】

【0013】

本発明の一態様によれば、リングネットワークの耐障害性を高めることができる。

【0014】

上記以外の課題、構成及び効果は、以下の実施形態の説明により明らかにされる。

【図面の簡単な説明】

【0015】

【図 1】本発明の実施例 1 のネットワークシステムの構成を示すブロック図である。

20

【図 2 A】本発明の実施例 1 のネットワークシステムにおけるスイッチ間の接続を示すブロック図である。

【図 2 B】本発明の実施例 1 のネットワークシステムに含まれるスイッチの構成を示すブロック図である。

【図 3】本発明の実施例 1 のスイッチが保持するコンフィグレーションテーブルの説明図である。

【図 4】本発明の実施例 1 のスイッチが保持するスタック状態管理テーブルの説明図である。

【図 5】本発明の実施例 1 のスイッチが保持するポートリンク状態管理テーブルの説明図である。

30

【図 6】本発明の実施例 1 のスイッチが保持するリングポート状態管理テーブルの説明図である。

【図 7】本発明の実施例 1 のネットワークシステムにおいてマスタノードを構成するマスタスイッチに障害が発生した場合の動作を示す説明図である。

【図 8】本発明の実施例 1 のネットワークシステムにおいてマスタノードを構成するバックアップスイッチに障害が発生した場合の動作を示す説明図である。

【図 9】本発明の実施例 1 のネットワークシステムにおいてマスタノードを構成するスイッチが障害から復旧した場合の動作を示す説明図である。

【図 10】本発明の実施例 1 のネットワークシステムにおいてトランジットノードを構成するマスタスイッチに障害が発生した場合の動作を示す説明図である。

40

【図 11】本発明の実施例 1 のネットワークシステムにおいてトランジットノードを構成するスイッチが障害から復旧した場合の動作を示す説明図である。

【図 12】本発明の実施例 1 のインタフェース制御部が実行する処理のフローチャートである。

【図 13】本発明の実施例 1 のプロトコル共通制御部が実行する処理のフローチャートである。

【図 14】本発明の実施例 1 のスタック管理 / 制御部が実行する処理のフローチャートである。

【図 15 A】本発明の実施例 1 のリング制御部が実行する処理のフローチャートの第 1 の部分である。

50

【図 1 5 B】本発明の実施例 1 のリング制御部が実行する処理のフローチャートの第 2 の部分である。

【図 1 6】本発明の実施例 2 のネットワークシステムにおけるトランジットノードへのスイッチの増設の説明図である。

【図 1 7】本発明の実施例 2 のネットワークシステムにおけるトランジットノードからのスイッチの減設の説明図である。

【発明を実施するための形態】

【0016】

以下、図面を参照して本発明の実施例を説明する。

【実施例 1】

10

【0017】

図 1 は、本発明の実施例 1 のネットワークシステムの構成を示すブロック図である。

【0018】

図 1 に示すネットワークシステム 100 は、本実施例が適用されるネットワークシステムの一例であり、データ転送装置 101A ~ 101J によって構成される。データ転送装置 101A ~ 101J の各々は、任意の種類のものであってよいが、一例を挙げるとすれば、OSI 参照モデルの第 2 層（データリンク層）のデータに基づいて、受信したデータを転送するデータ転送装置である。以下の説明において、データ転送装置 101A ~ 101J をスイッチ 101A ~ 101J と記載する。また、スイッチ 101A ~ 101J のいずれにも共通する説明をする場合等、スイッチ 101A ~ 101J のいずれかを特定する必要がない場合には、スイッチ 101 と記載する。後述する他のスイッチについても同様である。他の構成要素についても、同種の構成要素に共通する説明をする場合に、A、B、C 等を省略した参照符号を記載する場合がある。

20

【0019】

ネットワークシステム 100 にはリングプロトコルが適用される。リングプロトコルとは、複数のノードをリング状に接続したネットワークにおける障害の検出及びそれに伴う経路切り替えを高速に行うネットワークの冗長化プロトコルである（特許文献 1）。各ノードは、一つのスイッチ 101 であってもよいし、後述するスタックを構成する複数のスイッチ 101 であってもよい。図 1 の例では、スイッチ 101G、101H、101I 及び 101J の各々が一つのノードを構成する。以下、スイッチ 101G、101H、101I 及び 101J の各々によって構成されるノードを、それぞれ、ノード 101G、101H、101I 及び 101J と記載する。一方、スイッチ 101A 及び 101B は一つのノード 102A を、スイッチ 101C 及び 101D は一つのノード 102B を、スイッチ 101E 及び 101F は一つのノード 102C を構成する。

30

【0020】

スタックとは、物理的に相互に接続された複数台のスイッチ 101 を仮想的に一台のスイッチとして動作させる機能であり、この機能によって、複数台の一部に障害が発生した場合にも通信を継続することができる（特許文献 2）。以下、スタック機能による仮想的な一台のスイッチによって構成されるノード（例えば図 1 のノード 102A ~ 102C）を、「スタック構成のノード」のように記載する場合がある。また、仮想的な一台のスイッチを構成する複数のスイッチ 101 の集合（例えば図 1 のスイッチ 101A 及び 101B の集合）を「スタック」、スタックに含まれる各々のスイッチ 101（例えば図 1 のスイッチ 101A 及び 101B の各々）を「スタックを構成するスイッチ」又は「スタックのメンバスイッチ」のように記載する場合がある。

40

【0021】

図 1 の例では、リンク 104A を介してノード 102A をノード 102C に接続し、リンク 104B を介してノード 102C をノード 101G に接続し、リンク 104C を介してノード 101G をノード 101H に接続し、リンク 104D を介してノード 101H をノード 102A に接続することによって、リング 103A が構成される。さらに、リンク 104G を介してノード 102A をノード 101J に接続し、リンク 104H を介してノ

50

ード１０１Ｊをノード１０２Ｂに接続し、リンク１０４Ｅを介してノード１０２Ｂをノード１０１Ｉに接続し、リンク１０４Ｆを介してノード１０１Ｉをノード１０２Ａに接続することによって、もう一つのリング１０３Ｂが構成される。各リングを構成する複数のノードのうち一つはマスタノードであり、残りはトランジットノードである。これらの詳細な定義については後述する。図１の例では、ノード１０１Ｇがリング１０３Ａのマスタノードであり、ノード１０２Ｂがリング１０３Ｂのマスタノードであり、その他のノードはそれぞれのリングのトランジットノードである。

【００２２】

また、スタック機能によって仮想的な一台のスイッチとして動作する二つのスイッチ１０１のうち、一方はマスタスイッチ、もう一方はバックアップスイッチである。これらの詳細な定義についても後述する。図１の例では、スイッチ１０１Ａ、１０１Ｃ及び１０１Ｅがマスタスイッチであり、スイッチ１０１Ｂ、１０１Ｄ及び１０１Ｆがバックアップスイッチである。なお、本実施例では代表的な例として二台のスイッチ１０１がスタックを構成する例を示すが、本発明は、三台以上のスイッチ１０１がスタックを構成する場合にも適用できる。その場合、三台以上のメンバスイッチのうち一つがマスタスイッチ、別の一つがバックアップスイッチとなる。

【００２３】

なお、上記のようにスタック構成のノード１０２Ｂがノード１０１Ｉ及び１０１Ｊに接続される場合、物理的には、ノード１０２Ｂを構成するスイッチ１０１Ｃ及び１０１Ｄのうち一方がリンク１０４Ｅを介してスイッチ１０１Ｉに接続され、もう一方がリンク１０４Ｈを介してスイッチ１０１Ｊに接続され、スイッチ１０１Ｃ及び１０１Ｄがリンク１０４Ｋによって接続されることによって、リング１０３Ｂが形成される。他のスタック構成のノード１０２Ａ及び１０２Ｃについても同様である。スイッチ１０１Ａ及び１０１Ｂを接続するリンク１０４Ｉ、スイッチ１０１Ｅ及び１０１Ｆを接続するリンク１０４Ｊ及び上記のリンク１０４Ｋのように、スタックを構成する二つのスイッチ間を接続するリンクは、スタックリンクとも記載される。

【００２４】

また、図１では省略されているが、実際には、図１に示したそれぞれのノードにさらに一つ以上のネットワーク装置が接続されてもよい。例えば、図１に示したいずれかのノードは、他のリングに属するノード（図示省略）が接続されることによって、当該他のリングにも属することができる。あるいは、図１に示したいずれかのノードに、いずれのリングにも属しないネットワーク装置（例えばデータ転送装置又はサーバ等）が接続されてもよい。それらのネットワーク装置は、図１に示したリングを通過するデータの転送元又は転送先となり得る。

【００２５】

図２Ａは、本発明の実施例１のネットワークシステムにおけるスイッチ１０１間の接続を示すブロック図である。

【００２６】

図１に示すように、本実施形態は種々の形態のリングに適用され得るが、図２Ａには、説明を容易にするために、スイッチ１０１Ｋ～１０１Ｎによって構成される一つのリング１０３Ｃを示す。この例において、スイッチ１０１Ｋ及び１０１Ｌは、リンク１０４Ｌによって接続され、スタック構成のマスタノード１０２Ｄとなる。一方、スイッチ１０１Ｍ及び１０１Ｎは、それぞれトランジットノードとなる。以下、スイッチ１０１Ｍ及び１０１Ｎをそれぞれノード１０１Ｍ及び１０１Ｎのように記載する場合がある。

【００２７】

各スイッチ１０１は、複数のポートを有する。図２Ａには、例としてスイッチ１０１Ｋが有するポート２０１Ａ～２０１Ｃ及びスイッチ１０１Ｌが有するポート２０１Ｄ～２０１Ｆのみを示し、他のポートの図示は省略する。リングを構成するノード間を接続するポートは、リングポートとも記載される。図２Ａでは、ポート２０１Ａ及び２０１Ｄがリングポートである。スイッチ１０１Ｍ及び１０１Ｎもそれぞれ二つのリングポート（図示省

10

20

30

40

50

略)を有する。ポート201Aはリンク104Oを介してスイッチ101Nの一方のリングポートに接続され、ポート201Dはリンク104Mを介してスイッチ101Mの一方のリングポートに接続される。スイッチ101Mのもう一方のリングポートは、リンク104Nを介してスイッチ101Nのもう一方のリングポートに接続される。

【0028】

上記のように、一つのリングに属するノードは二つのリングポートを有するが、マスタノードのリングポートは、一方がプライマリポート、もう一方がセカンダリポートとして扱われる。リングに障害が発生していない場合、マスタノードは、セカンダリポートを、データフレームを送受信しないように論理的にブロックする。図2Aの例では、ポート201Aがプライマリポート、ポート201Dがセカンダリポートであるため、障害が発生していない場合には、ポート201Dがブロックされている。これらの詳細については後述する。

【0029】

スイッチ101Kのポート201Bは、リンク104Lを介してスイッチ101Lのポート201Eに接続される。このように、スタックを構成する二つのスイッチ間を接続するポート(すなわちスタックリンクに接続されるポート)は、スタックポートとも記載される。

【0030】

なお、各リンク104は、物理的には、例えばポート201間を接続するネットワークケーブル等によって実現される。

【0031】

リング103Cを構成する各スイッチ101は、上記のリングポート以外のポート(上記のリングポートとは別のリングポート、又はリングポートではないポート)を介して、さらに他の一つ以上のネットワーク装置と接続され得る。図2Aの例では、ポート201C及び201Fが、それぞれリンク104P及び104Qを介して、リング103Cに属しないスイッチ101Oに接続される。この例において、リンク104P及び104Qは、いわゆるリンクアグリゲーションによって仮想的に一つのリンクとして扱われる。これによって、ノード102Dは、スイッチ101K又は101Lの一方に障害が発生した場合にも、正常なもう一方を使用してスイッチ101Oとの通信を継続することができる。ただし、リンクアグリゲーションの適用は一例であり、リンク104P及び104Qがそれぞれに別のスイッチ(図示省略)に接続されてもよい。同様に、スイッチ101M及び101Nには、それぞれリンク104R及び104Sを介して、図示しない任意のネットワーク装置が接続されてもよい。

【0032】

図2Bは、本発明の実施例1のネットワークシステムに含まれるスイッチ101の構成を示すブロック図である。

【0033】

図2Bには、例として、リング103Cに属し、スタックを構成するスイッチ101Kの構成を示す。他のスイッチ101の構成も基本的には図2Bと同様であるが、それぞれのスイッチ101の属性(例えば当該スイッチ101がリングに属するか否か及びスタックを構成するか否か等)によって保持される情報が相違する場合がある。また、ユーザがコマンドを入力するのに使用する管理装置240をリンク104Tを介してポート201G経由でスイッチ101Kに接続する。管理装置240の接続先ポートはポート201の何れかであればよい。また、管理装置は、図2Bのようにスイッチ101Kの外に存在する形態でもよいし、管理装置の機能をスイッチ101K内部に持たせてもよい。

【0034】

スイッチ101Kは、相互に接続された制御部210、メモリ220、転送エンジン230、及び転送エンジン230に接続される複数のポート201を有する。制御部210は、プロトコル共通制御部211、スタック管理/制御部212、リング制御部213及びインタフェース制御部214を含む。これらは、専用のハードウェアによって実現され

てもよい。あるいは、制御部 2 1 0 がプロセッサであり、これがメモリ 2 2 0 に格納されたプログラムを実行することによって、プロトコル共通制御部 2 1 1 等の各制御部が実現されてもよい。その場合、以下の説明においてプロトコル共通制御部 2 1 1 等の各制御部が実行する処理は、実際には制御部 2 1 0 に相当するプロセッサによって実行される。

【 0 0 3 5 】

メモリ 2 2 0 には、コンフィグレーションテーブル 2 2 1、スタック状態管理テーブル 2 2 2、ポートリンク状態管理テーブル 2 2 3 及びリングポート状態管理テーブル 2 2 4 が格納される。コンフィグレーションテーブル 2 1 1 は、スイッチが属するノードのリングネットワークを構成するネットワークシステムにおける属性（マスタノードかトランジットノード）を示す情報を含むリングネットワークを構成するスイッチの情報を格納する。スタック状態管理テーブル 2 2 2 は、スタックを構成するデータ転送装置システムにおける属性（マスタかバックアップか）を示す情報を含む。ポートリンク状態管理テーブル 2 2 3 は、スタックを構成するスイッチが有するポートの状態に関する情報を保持する。リングポート状態管理テーブル 2 2 4 は、自スイッチの二つのリングポートに関する情報を含む。

10

【 0 0 3 6 】

制御部 2 1 0 は、外部からのコマンドの入力に応じて、コンフィグレーションテーブル 2 2 1 を更新することによりリングプロトコル及びスタックに関するスイッチの設定を行う。

【 0 0 3 7 】

制御部 2 1 0（インターフェース制御部 2 1 4）は、スイッチのポートの状態変化を検出し、ポートリンク状態管理テーブル 2 2 3 を更新する。

20

【 0 0 3 8 】

制御部 2 1 0（プロトコル共通制御部 2 1 1）は、リングにおけるヘルスチェックフレームの送信制御を、リングネットワークシステムにおける自ノードの属性（マスタノードかトランジットノード）に基づいて行う。また、スタックにおけるヘルスチェックフレームの送信制御を、スタックを構成するデータ転送装置システムにおける属性（マスタかバックアップか）に基づいて行う。

【 0 0 3 9 】

制御部 2 1 0（スタック管理 / 制御部 2 1 2）は、スタックを構成するスイッチの状態を監視し、スイッチの障害を検知した場合に、スタック状態管理テーブル 2 2 2 に基づいてスタックを構成するデータ転送装置システムにおける属性（マスタかバックアップか）の切り替えを行う。

30

【 0 0 4 0 】

制御部 2 1 0（リング制御部 2 1 3）は、コンフィグレーションテーブル 2 2 1 とスタック状態管理テーブル 2 2 2 に基づいてリングポート状態管理テーブル 2 2 4 を更新する。また、リングネットワークシステムの状態（障害が発生していない正常状態か、障害が発生している障害状態か）を監視し、監視結果に基づいてリングポート状態管理テーブル 2 2 4 を更新することにより、スタックを構成するスイッチのリングポートの設定を制御する。

40

【 0 0 4 1 】

転送エンジン 2 3 0 は、いずれかのポート 2 0 1 がリンク 1 0 4 を介して受信したデータを他のポート 2 0 1 に転送する。当該他のポート 2 0 1 は、転送されたデータを、そのポート 2 0 1 に接続されたリンク 1 0 4 を介して送信する。これによってスイッチ 1 0 1 K のデータ転送機能が実現される。例えば、スイッチ 1 0 1 K がレイヤ 2 スイッチである場合、スイッチ 1 0 1 K は図示しない M A C（Media Access Control）アドレステーブルを保持し、転送エンジン 2 3 0 は、受信したフレームのヘッダに含まれる宛先 M A C アドレスと M A C アドレステーブルとを照合して決定したポート 2 0 1 に当該フレームを転送する。また、転送エンジン 2 3 0 は、制御部 2 1 0 からの指示に従ってポート 2 0 1 の状態を変更する。

50

【 0 0 4 2 】

次に、スイッチ 1 0 1 が保持するテーブルについて、図 3 ~ 図 6 を参照して説明する。ここでは、例として、図 2 A 及び図 2 B に示すスイッチ 1 0 1 K が保持するテーブルを説明する。

【 0 0 4 3 】

図 3 は、本発明の実施例 1 のスイッチ 1 0 1 K が保持するコンフィグレーションテーブル 2 2 1 の説明図である。

【 0 0 4 4 】

コンフィグレーションテーブル 2 2 1 は、各スイッチ 1 0 1 の設定情報を保持する。このテーブルに保持される設定情報は、後述するように、ユーザがコマンドを入力して設定するものである。保持された設定情報は、ユーザがコマンドを入力して変更するまで維持される。ユーザは、各スイッチ 1 0 1 の何れかのポートに接続される管理装置 2 4 0 に対してコマンドを入力し、管理装置 2 4 0 はコマンドによって入力された情報をコンフィグレーションテーブル 2 2 1 に設定する。また、管理装置 2 4 0 は、コマンドを入力する際に接続し、入力が完了すると、各スイッチ 1 0 1 から外してもよい。なお、後述するスタックのヘルスチェックフレームによって、スイッチ 1 0 1 K とスイッチ 1 0 1 L とが同じコンフィグレーションテーブル 2 2 1 を保持するように同期処理を行ってもよい。また、スイッチ 1 0 1 K とスイッチ 1 0 1 L とが保持するコンフィグレーションテーブル 2 2 1 の設定情報が同じになるように、スイッチ 1 0 1 K とスイッチ 1 0 1 L のそれぞれに対してユーザがコマンドを入力して設定してもよい。

【 0 0 4 5 】

図 3 に示すコンフィグレーションテーブル 2 2 1 は、スタックを構成するスイッチ 1 0 1 K 及び 1 0 1 L に対応するレコードを含み、各レコードは、スイッチ 3 0 1、スタックプライオリティ 3 0 2、スタックポート番号 3 0 3、リング ID 3 0 4、リングモード 3 0 5、リングポート番号 3 0 6 及びポートシャットダウン設定 3 0 7 を含む。

【 0 0 4 6 】

スイッチ 3 0 1 は、各スイッチ 1 0 1 の識別情報である。図 3 の例において、「スイッチ 1」はスイッチ 1 0 1 K を識別し、「スイッチ 2」はスイッチ 1 0 1 L を識別する。

【 0 0 4 7 】

スタックプライオリティ 3 0 2 は、各スイッチ 1 0 1 のスタックにおける優先度を示す。例えば、スタックプライオリティ 3 0 2 の値が大きいほど優先度が高いことを示す。この値は、スタックの運用を開始するときに、二つのスイッチのいずれをマスタスイッチとするかを決定するために参照される。例えば、スタックプライオリティ 3 0 2 が示す優先度が高いスイッチ 1 0 1 K が最初にマスタスイッチとなる。

【 0 0 4 8 】

スタックポート番号 3 0 3 は、スタックリンクが接続されるスタックポートの識別情報である。図 3 の例では、「1 / 0 / 1」がポート 2 0 1 B を識別し、「2 / 0 / 5」がポート 2 0 1 E を識別する。この例において、例えば「2 / 0 / 5」の最初の「2」は当該ポート 2 0 1 E を有するスイッチ 1 0 1 L の番号、「0」はスロット番号、「5」はスイッチ 1 0 1 L におけるポート 2 0 1 E の番号であるが、このような識別情報の形式は一例であり、実際にはどのような形式が使用されてもよい。また、図 3 に示すスイッチ 1 0 1 K 及び 1 0 1 L は、それぞれ一つのスタックポートを有しているが、実際には複数のスタックポートを有し、複数のスタックリンクによって相互に接続されてもよい。

【 0 0 4 9 】

リング ID 3 0 4 は、スイッチ 1 0 1 K 及び 1 0 1 L が属するリング 1 0 3 C の識別情報である。図 3 の例では、リング 1 0 3 C のリング ID 3 0 4 は「2」である。図 1 に示すようにネットワークシステム 1 0 0 が複数のリング 1 0 3 を有する場合、それぞれのリング 1 0 3 に固有の識別情報が与えられる。

【 0 0 5 0 】

リングモード 3 0 5 は、スイッチが属するノードのリングネットワークにおける属性を

示す。具体的には、リングモード305は、スイッチ101K及び101Lを含むノードがマスタノード又はトランジットノードのいずれであるかを示す。図3の例では、ノード102Dがマスタノードであるため、リングモード305は「マスタ」となる。

【0051】

詳細は後述するが、リングモード305が「マスタ」の場合、スタックを構成するスイッチ101K及び101Lを含むノードは、リングネットワークのマスタノードとして、二つのリングポートの各々からリングのヘルスチェックフレームをリングネットワーク上の隣接するノードに送信し、各々のヘルスチェックフレームをリングネットワーク上の隣接するノードから受信し終端して、障害を監視する。一方、リングモード305が「トランジット」の場合、スタックを構成するスイッチ101K及び101Lを含むノードは、

10

【0052】

リングポート番号306は、隣接するノードに接続されるリングポートの識別情報である。図3の例では、「1/0/40」がポート201Aを識別し、「2/0/35」がポート201Dを識別する。ポートシャットダウン設定307は、各リングポートがシャットダウンされているか否かを示す。ポートシャットダウン設定307の値「shutdown」はリングポートがシャットダウンされていることを、「No shutdown」はシャットダウンされていないことを示す。図3の例では、いずれのリングポートもシャットダウンされていない。

20

【0053】

以下、コンフィグレーションテーブル221に値を設定するためのコマンドの例を、図2A及び図3に沿って示す。スイッチ101K又は101Lのリング103Cにおけるリングモード305をマスタモードに設定する場合、次のようなコマンドが使用される。

```
(config)# ring 2
```

```
(config-ring)# mode master
```

【0054】

スイッチ101Kのポート201A及びスイッチ101Lのポート201Dをリングポートとして設定する場合、次のようなコマンドが使用される。

30

```
(config)# interface 1/0/40
```

```
(config-if)# ring-port 2
```

```
(config-if)# exit
```

```
(config)# interface 2/0/35
```

```
(config-if)# ring-port 2
```

【0055】

一方、スイッチ101M又は101Nのリング103Cにおけるリングモード305をトランジットモードに設定する場合、次のようなコマンドが使用される。

```
(config)# ring 2
```

```
(config-ring)# mode transit
```

40

【0056】

スイッチ101Kのスタックプライオリティを「20」に設定する場合、次のようなコマンドが使用される。

```
(config)# switch 1 priority 20
```

【0057】

スイッチ101Kのポート201Bをスタックポートとして設定する場合、次のようなコマンドが使用される。

```
(config)# interface 1/0/1
```

```
(config-if)# switchport mode stack
```

```
(config-if)# exit
```

50

【 0 0 5 8 】

スイッチ 1 0 1 K のポート 2 0 1 A をシャットダウンする場合、次のようなコマンドが使用される。

```
(config)# interface 1/0/40
```

```
(config-if)# shutdown
```

【 0 0 5 9 】

スイッチ 1 0 1 K のポート 2 0 1 A のシャットダウンを解除する場合、次のようなコマンドが使用される。

```
(config)# interface 1/0/40
```

```
(config-if)# no shutdown
```

10

【 0 0 6 0 】

図 4 は、本発明の実施例 1 のスイッチ 1 0 1 K が保持するスタック状態管理テーブル 2 2 2 の説明図である。

【 0 0 6 1 】

スタック状態管理テーブル 2 2 2 は、スタックを構成するスイッチ 1 0 1 がスタック状態を管理するために保持するテーブルである。

【 0 0 6 2 】

図 4 に示すスタック状態管理テーブル 2 2 2 は、スタックを構成するスイッチ 1 0 1 K 及び 1 0 1 L に対応するレコードを含み、各レコードは、スイッチ 4 0 1 及びスタック状態 4 0 2 を含む。スイッチ 4 0 1 は、図 3 のスイッチ 3 0 1 と同様に、スタックを構成する各スイッチ 1 0 1 の識別情報である。スタック状態 4 0 2 は、スタックにおける各スイッチの属性を示す。具体的には、スタック状態 4 0 2 は、スタックにおける各スイッチ 1 0 1 の状態がマスタ、バックアップ又は初期状態のいずれであることを示す。図 3 の例は、スイッチ 1 0 1 K がマスタスイッチであり、スイッチ 1 0 1 L がバックアップスイッチであることを示す。例えば各スイッチ 1 0 1 の電源が投入された直後等、各スイッチ 1 0 1 がマスタスイッチ又はバックアップスイッチのいずれであるか定まっていない時点では、スタック状態 4 0 2 の値は「初期状態」となる。

20

【 0 0 6 3 】

図 5 は、本発明の実施例 1 のスイッチ 1 0 1 K が保持するポートリンク状態管理テーブル 2 2 3 の説明図である。

30

【 0 0 6 4 】

ポートリンク状態管理テーブル 2 2 3 は、各スイッチ 1 0 1 の各ポート 2 0 1 の状態を管理するためのテーブルである。図 5 に示すポートリンク状態管理テーブル 2 2 3 は、スタックを構成するスイッチ 1 0 1 K 及び 1 0 1 L の各ポート 2 0 1 に対応するレコードを含み、各レコードは、スイッチ 5 0 1、ポート番号 5 0 2 及びポート状態 5 0 3 を含む。

【 0 0 6 5 】

スイッチ 5 0 1 は、図 3 のスイッチ 3 0 1 と同様に、各スイッチ 1 0 1 の識別情報である。ポート番号 5 0 2 は、各スイッチ 1 0 1 の各ポート 2 0 1 の識別情報である。ポート状態 5 0 3 は、各ポート 2 0 1 の状態を示す情報であり、その値は「UP」又は「DOWN」のいずれかとなる。あるポート 2 0 1 が他のポート 2 0 1 に接続され、それらのポート 2 0 1 間の通信が可能である場合、それらのポート 2 0 1 のポート状態 5 0 3 は「UP」となる。あるポート 2 0 1 又はそのポート 2 0 1 に接続された他のスイッチ 1 0 1 のポート 2 0 1 のいずれかに障害が発生している場合、又は、そのポート 2 0 1 に何も接続されていない場合、そのポート 2 0 1 のポート状態 5 0 3 は「DOWN」となる。

40

【 0 0 6 6 】

さらに、障害等が発生しておらず、リンクも正常に接続されているポート 2 0 1 であっても、そのポートに関するポートシャットダウン設定 3 0 7 が「shutdown」となっている場合には、そのポート 2 0 1 のポート状態 5 0 3 は「DOWN」となる。ポート状態 5 0 3 が「DOWN」となっているポート 2 0 1 はデータフレーム、リングのヘルスチェックフレーム、スタックのヘルスチェックフレーム及びフラッシュ制御フレーム等を送受信す

50

ることができない。

【 0 0 6 7 】

図 6 は、本発明の実施例 1 のスイッチ 1 0 1 K が保持するリングポート状態管理テーブル 2 2 4 の説明図である。

【 0 0 6 8 】

リングポート状態管理テーブル 2 2 4 は、各ノードの各リングポートの状態を管理するためのテーブルである。具体的には、図 6 に示すリングポート状態管理テーブル 2 2 4 は、スイッチ 1 0 1 K を含むノード 1 0 2 D の各リングポートに対応するレコードを含み、各レコードは、スイッチ 6 0 1、リング ID 6 0 2、リングポート番号 6 0 3、リングポート 6 0 4、リングポート状態 6 0 5 及びリング障害状態 6 0 6 を含む。

10

【 0 0 6 9 】

スイッチ 6 0 1、リング ID 6 0 2 及びリングポート番号 6 0 3 は、それぞれ、図 3 のスイッチ 3 0 1、リング ID 3 0 4 及びリングポート番号 3 0 6 に対応する。

【 0 0 7 0 】

リングポート 6 0 4 は、各リングポートがプライマリポート又はセカンダリポートのいずれであるかを示す。図 6 の例では、ポート 2 0 1 A がプライマリポート、ポート 2 0 1 D がセカンダリポートである。マスタノードの二つのリングポートの一方が必ずプライマリポートとなり、もう一方が必ずセカンダリポートとなるように決定される限り、リングポート 6 0 4 はどのような方法で決定されてもよい。例えば、より若いスイッチ番号が与えられたスイッチ 1 0 1 のリングポートがプライマリポートとなってもよい。二つのリン

20

【 0 0 7 1 】

リングポート状態 6 0 5 は、各リングポートの状態がフォワーディング又はブロッキングのいずれの状態であるかを示す。図 6 の例では、ポート 2 0 1 A の状態がフォワーディング、ポート 2 0 1 D の状態がブロッキングである。なお、状態がフォワーディングであるポート 2 0 1 は、後述するヘルスチェックフレーム及びフラッシュ制御フレームのような、リングを制御又は管理するためのフレームだけでなく、ユーザデータを含むデータフレームも送受信する。これに対して、状態がブロッキングであるポート 2 0 1 は、リングを制御又は管理するためのフレームは送受信するが、データフレームは送受信しない。

30

【 0 0 7 2 】

初期状態においては、マスタノードの二つのリングポートのうち一方がフォワーディング、もう一方がブロッキングである必要がある。例えば、プライマリポートがフォワーディング、セカンダリポートがブロッキングであってもよい。一方、トランジットノードの二つのリングポートの初期状態はいずれもブロッキングであり、フラッシュ制御フレームを受信したときにいずれもフォワーディングに変更される。

【 0 0 7 3 】

リング障害状態 6 0 6 は、ノードが属するリングに障害が発生しているか否かを示す。図 6 の例では、リング 1 0 3 C に障害が発生していない（すなわち正常である）。

【 0 0 7 4 】

40

図 3 ~ 図 6 には、図 2 A 等 に示すスイッチ 1 0 1 K が保持する情報を例示した。この例が示すように、スタックを構成するスイッチ 1 0 1 は、自スイッチ 1 0 1 に関する情報だけでなく、スタックを構成する相手方のスイッチに関する情報も保持する。そのために、スタックを構成するマスタスイッチ及びバックアップスイッチは、自スイッチに関する情報を相互に通信することによってテーブルの内容を同期する。したがって、図 2 A 等 に示すスイッチ 1 0 1 L も、スタックに障害が発生していない限り、図 3 ~ 図 6 に示したものと同一の情報を保持する。

【 0 0 7 5 】

ただし、リングポート状態管理テーブル 2 2 4 のリング障害状態 6 0 6 は、スイッチ 1 0 1 K 及び 1 0 1 L の間で引き継がれない。後述するようにマスタスイッチに障害が発生

50

してバックアップスイッチが新たなマスタスイッチに遷移した場合、当該新たなマスタスイッチが改めてリング障害を検知して、当該リングが障害状態にあることを確認する。

【0076】

一方、リングに属しているがスタックを構成しないスイッチ101（例えば図2Aのスイッチ101M及び101N）は、スタックに関する情報を保持しない。例えばスイッチ101M及び101Nのコンフィグレーションテーブル221は、それぞれ、自スイッチ101M及び101Nが有する二つのリングポートに関する設定情報を含み、スタックプライオリティ302及びスタックポート番号303を含まない。また、スイッチ101M及び101Nは、スタック状態管理テーブル222を保持しない。スイッチ101M及び101Nのポートリンク状態管理テーブル223は、それぞれ、自スイッチ101M及び101Nのポートに関する情報のみを含む。スイッチ101M及び101Nのリングポート状態管理テーブル224は、自スイッチ101の二つのリングポートに関する情報を含む。

10

【0077】

また、本実施例では、トランジットノードであるスイッチ101M及び101Nのリングポート状態管理テーブル224は、リング障害状態606を含まない。

【0078】

ここで、図2A～図6を参照して、正常時（すなわちリング103Cのどの部分にも障害が発生していない場合）の各スイッチ101の動作について説明する。

【0079】

20

図2Aの例において、ノード102Dはリング103Cのマスタノードであり、ノード101M及び101Nはトランジットノードである。また、ノード102Dは、スイッチ101K及び101Lからなるスタック構成のノードであり、スイッチ101Kがマスタスイッチ、スイッチ101Lがバックアップスイッチである。

【0080】

ここで、まず、マスタノード及びトランジットノードについて説明する。マスタノードは、一つのリングに一つのみ存在するノードであり、ヘルスチェックフレームを用いてリングの障害状態を監視する。具体的には、マスタノードは、二つのリングポートからヘルスチェックフレームを送信し、所定の時間内に両方向のヘルスチェックフレームを受信するか否かを監視する。マスタノードは、監視によってリングの障害が検出されない場合、二つのリングポートのうち一方の状態をフォワーディング、もう一方の状態をブロッキングに設定し、障害が検出された場合、ブロッキングとなっていたリングポートの状態をフォワーディングに変更する。

30

【0081】

一方、トランジットノードは、一方のリングポートがヘルスチェックフレームを受信すると、そのヘルスチェックフレームを、もう一方のリングポートから、リング上の隣接するノードに転送する。また、トランジットノードの二つのリングポートのいずれも、データフレームを送受信する。マスタノードは、トランジットノードをこのように経由したヘルスチェックフレームの受信状況により障害監視を行う。なお、トランジットノードは、リングのヘルスチェックフレームを用いた障害監視を行わなくてもよい。

40

【0082】

次に、マスタスイッチ及びバックアップスイッチについて説明する。マスタスイッチは、スタックを構成する全てのメンバスイッチ及びその機能を制御するスイッチであり、バックアップスイッチは、マスタスイッチからの制御に従って動作するスイッチである。特に、リングに属するスタック構成のノードのマスタスイッチは、マスタスイッチが有するリングポートの状態（フォワーディング又はブロッキング）を設定し、さらに、バックアップスイッチが有するリングポートの状態も設定する。当該ノードがマスタノードである場合、マスタスイッチはマスタノードとしての機能を実現するための処理を実行する。一方、バックアップスイッチは、マスタスイッチに障害が発生した場合にマスタスイッチとなる。具体的には、マスタスイッチ及びバックアップスイッチは、スタックポートを介し

50

て相互にヘルスチェックフレームを送信し、所定の期間ヘルスチェックフレームを受信できなかった場合は相手方のスイッチに障害が発生したと判定する。

【 0 0 8 3 】

なお、以下の説明において、リングの障害監視のためにマスタノードが送受信するヘルスチェックフレームをリングのヘルスチェックフレーム、スタックの障害監視のためにマスタスイッチ及びバックアップスイッチが相互に送信するヘルスチェックフレームをスタックのヘルスチェックフレームとも記載する。

【 0 0 8 4 】

図 2 A の例において、マスタノード 1 0 2 D のマスタスイッチ 1 0 1 K は、ポート 2 0 1 A 及び 2 0 1 B から、それぞれリングのヘルスチェックフレームを送信する。ポート 2 0 1 A から送信されたリングのヘルスチェックフレームは、トランジットノードであるスイッチ 1 0 1 N 及び 1 0 1 M によって順次転送され、バックアップスイッチ 1 0 1 L のポート 2 0 1 D に到着する。バックアップスイッチ 1 0 1 L は、受信したリングのヘルスチェックフレームをポート 2 0 1 E からスタックリンク 1 0 4 L を経由してマスタスイッチ 1 0 1 K のポート 2 0 1 B に送信する。

10

【 0 0 8 5 】

その反対に、ポート 2 0 1 B から送信されたリングのヘルスチェックフレームは、バックアップスイッチ 1 0 1 L、トランジットノードであるスイッチ 1 0 1 M 及び 1 0 1 N によって順次転送され、マスタスイッチ 1 0 1 K のポート 2 0 1 A に到着する。マスタスイッチ 1 0 1 K は、ポート 2 0 1 A から送信したリングのヘルスチェックフレームをポート 2 0 1 B から受信し、ポート 2 0 1 B から送信したリングのヘルスチェックフレームをポート 2 0 1 A から受信すると、リング 1 0 3 C に障害が発生していないと判定する。

20

【 0 0 8 6 】

このようにリング 1 0 3 C に障害が発生していない場合、マスタノード 1 0 2 D の二つのリングポートのうちプライマリポートであるポート 2 0 1 A の状態はフォーワーディング、セカンダリポートであるポート 2 0 1 D の状態はブロッキングである。この状態において、例えば、ノード 1 0 2 D がスイッチ 1 0 1 O からデータフレームを受信し、その宛先がリンク 1 0 4 S の接続先のネットワーク装置（図示省略）である場合の当該データフレームの流れを説明する。

【 0 0 8 7 】

スイッチ 1 0 1 O から送信されたデータフレームは、リンク 1 0 4 P を経由してポート 2 0 1 C が、又はリンク 1 0 4 Q を経由してポート 2 0 1 F が受信する。ポート 2 0 1 D の状態がブロッキングであるため、ポート 2 0 1 F がデータフレームを受信した場合、スイッチ 1 0 1 L は、当該データフレームをポート 2 0 1 E からスイッチ 1 0 1 K に送信する。スイッチ 1 0 1 K は、ポート 2 0 1 B 又は 2 0 1 C から受信したデータフレームをポート 2 0 1 A から送信する。スイッチ 1 0 1 N は、リンク 1 0 4 O を経由して受信したデータフレームを、リンク 1 0 4 S の接続先に転送する。

30

【 0 0 8 8 】

仮に、データフレームの送信先がリンク 1 0 4 R の接続先のネットワーク装置であれば、スイッチ 1 0 1 N は、受信したデータフレームをスイッチ 1 0 1 M に転送する。スイッチ 1 0 1 M は、受信したデータフレームをリンク 1 0 4 R の接続先のネットワーク装置に転送する。このように、リング 1 0 3 C に障害が発生していなければ、リング 1 0 3 C に属する任意の二つのノード間にデータフレームの通信経路を設定することができる。その一方で、ポート 2 0 1 D の状態がブロッキングであるため、データフレームがリング 1 0 3 C 内を循環することは防止される。

40

【 0 0 8 9 】

次に、リングに障害が発生した場合のスイッチ 1 0 1 の動作について説明する。

【 0 0 9 0 】

図 7 は、本発明の実施例 1 のネットワークシステムにおいてマスタノードを構成するマスタスイッチに障害が発生した場合の動作を示す説明図である。

50

【 0 0 9 1 】

図 7 に示すリング 1 0 3 D は、図 2 A のリング 1 0 3 C と同様のスイッチ 1 0 1 K ~ 1 0 1 N に加えて、スイッチ 1 0 1 P を含む。スイッチ 1 0 1 K がスイッチ 1 0 1 P に接続され、スイッチ 1 0 1 P がスイッチ 1 0 1 N に接続されることを除いて、リング 1 0 3 D のスイッチ 1 0 1 間の接続関係は、リング 1 0 3 C に示したものと同様である。また、図 7 ではポート 2 0 1 等が省略されているが、スイッチ 1 0 1 K ~ 1 0 1 N の構成は図 2 A ~ 図 6 を参照して説明した通りであり、スイッチ 1 0 1 P の構成はスイッチ 1 0 1 N 等と同様である。

【 0 0 9 2 】

以下、リング 1 0 3 D の外のネットワーク装置（例えばスイッチ 1 0 1 O）からマスタノード 1 0 2 D が受信したデータフレームが、リング 1 0 3 D を通過して、スイッチ 1 0 1 N からリング 1 0 3 D の外のネットワーク装置に送信される例を説明する。マスタスイッチであるスイッチ 1 0 1 K は、リングのヘルスチェックフレームによる障害監視を実行する。リングに障害が発生していない時点において、データフレームは、フォワーディング状態のポート 2 0 1 A から送信され、スイッチ 1 0 1 P 及び 1 0 1 N によって順次転送される。さらに、スイッチ 1 0 1 K 及び 1 0 1 L は、相互にスタックのヘルスチェックフレームによる障害監視を実行する。上記のような障害監視が実行されている状態を以下の説明において障害監視状態とも記載する。

【 0 0 9 3 】

その後、スイッチ 1 0 1 K に障害が発生すると、スイッチ 1 0 1 K からリングのヘルスチェックフレーム及びスタックのヘルスチェックフレームが送信されなくなる。バックアップスイッチであるスイッチ 1 0 1 L は、スイッチ 1 0 1 K からのスタックのヘルスチェックフレームを受信しなくなったときにスイッチ 1 0 1 K の障害を検知し、自身が保持するスタック状態管理テーブル 2 2 2 のスイッチ 1 0 1 L に対応するスタック状態 4 0 2 を「バックアップ」から「マスタ」に変更する。これによって、スイッチ 1 0 1 L が新たなマスタスイッチとなる。

【 0 0 9 4 】

次に、スイッチ 1 0 1 L は、自身が保持するリングポート状態管理テーブル 2 2 4 のリングポート 2 0 1 D に対応するリングポート状態 6 0 5 を「ブロッキング」から「フォワーディング」に変更する。さらに、スイッチ 1 0 1 L は、リング 1 0 3 D にフラッシュ制御フレームを送信する。フラッシュ制御フレームは、スイッチ 1 0 1 L から、スイッチ 1 0 1 M、1 0 1 N 及び 1 0 1 P に順次転送される。スイッチ 1 0 1 L 及びフラッシュ制御フレームを受信したスイッチ 1 0 1 M、1 0 1 N 及び 1 0 1 P は、各々が保持する MAC アドレステーブル（図示省略）をクリアする。

【 0 0 9 5 】

これによって、データフレームの転送経路が、スイッチ 1 0 1 L からスイッチ 1 0 1 M を経由してスイッチ 1 0 1 N に至る経路に切り替えられる。スイッチ 1 0 1 K に障害が発生することによってスイッチ 1 0 1 K からスイッチ 1 0 1 P を経由してスイッチ 1 0 1 N に至る経路を使用できなくなるが、上記のように経路を切り替えることによってデータフレームの転送を継続することができる。スイッチ 1 0 1 K に障害が発生しているため、リングポート 2 0 1 D の状態をフォワーディングに変更してもデータフレームはリング 1 0 3 D 内を循環しない。

【 0 0 9 6 】

さらに、スイッチ 1 0 1 L は、マスタスイッチとなったため、リングのヘルスチェックフレームの双方向（すなわちスイッチ 1 0 1 K の方向及びスイッチ 1 0 1 M の方向）への送信を開始する。また、スイッチ 1 0 1 L は、引き続きスタックのヘルスチェックフレームによる障害監視も実行する。これによって、スイッチ 1 0 1 L は、スイッチ 1 0 1 K が引き続き障害状態にある場合には、リング 1 0 3 C が障害状態にあることを認識し、スイッチ 1 0 1 K が復旧した場合にはその復旧を検知することができる。このように、障害からの復旧を監視する状態を、以下の説明において復旧監視状態とも記載する。

【0097】

図8は、本発明の実施例1のネットワークシステムにおいてマスタノードを構成するバックアップスイッチに障害が発生した場合の動作を示す説明図である。

【0098】

図8に示すリング103Dは、ノード102Dの二つのリングポートのうち、スイッチ101Kのポート201Aがセカンダリポートであり、スイッチ101Lのポート201Dがプライマリポートであることを除いて、図7に示したリング103Dと同様である。ただし、この変更は説明のためのものであり、例えばスタックプライオリティ302に応じてリングポート604が決定される場合には、実際には、コンフィグレーションテーブル221を変更しない限り、プライマリポート及びセカンダリポートは変更されない。

10

【0099】

この例において、リング103Dに障害が発生していない場合、ポート201Dの状態がフォーディング、ポート201Aの状態がブロッキングであるため、ノード102Dからスイッチ101Nに至るデータフレームは、スイッチ101L、101M及び101Nによって順次転送される。

【0100】

マスタスイッチであるスイッチ101Kは、リングのヘルスチェックフレームによる障害監視を実行する。さらに、スイッチ101K及び101Lは、スタックのヘルスチェックフレームを用いてスタックの障害監視を実行する。スイッチ101Kは、スタックの障害監視によってスイッチ101Lの障害を検知した場合、リング103Dにフラッシュ制御フレームを送信する。スイッチ101K及びフラッシュ制御フレームを受信したスイッチ101P、101N及び101Mは、各々が保持するMACアドレステーブルをクリアする。

20

【0101】

さらに、スイッチ101Kは、自身が保持するリングポート状態管理テーブル224のリングポート201Aに対応するリングポート状態605を「ブロッキング」から「フォーディング」に変更する。

【0102】

これによって、データフレームの転送経路が、スイッチ101Kからスイッチ101Pを経由してスイッチ101Nに至る経路に切り替えられる。スイッチ101Lに障害が発生することによってスイッチ101Lからスイッチ101Mを経由してスイッチ101Nに至る経路を使用できなくなるが、上記のように経路を切り替えることによってデータフレームの転送を継続することができる。

30

【0103】

スイッチ101Kは、復旧監視状態へ遷移する。具体的には、スイッチ101Kは、引き続き、リングのヘルスチェックフレームを双方向に送信する。また、スイッチ101Kは、引き続きスタックのヘルスチェックフレームによる障害監視も実行する。これによって、スイッチ101Kは、スイッチ101Lが復旧した場合にはその復旧を検知することができる。

【0104】

図9は、本発明の実施例1のネットワークシステムにおいてマスタノードを構成するスイッチが障害から復旧した場合の動作を示す説明図である。

40

【0105】

具体的には、図9には、図8に示した障害が復旧した場合の動作を示す。スイッチ101Lは、障害から復旧すると、バックアップスイッチとしてスタックに加入する。次に、スイッチ101Kは、セカンダリポートであるポート201Aの状態をブロッキングに変更する。さらに、スイッチ101Kは、フラッシュ制御フレームを送信する。その結果、リング103Dの各スイッチのMACアドレステーブルがクリアされる。これによって、データフレームの転送経路が、スイッチ101Lからスイッチ101Mを経由してスイッチ101Nに至る経路に切り替えられる。その後、スイッチ101K及び101Lは、障

50

害監視状態へ遷移する。

【0106】

図10は、本発明の実施例1のネットワークシステムにおいてトランジットノードを構成するマスタスイッチに障害が発生した場合の動作を示す説明図である。

【0107】

図10に示すリング103Dにおけるスイッチ101間の接続関係は図7～図9と同様であるが、スイッチ101Mがマスタノードであり、スイッチ101K及び101Lがスタック構成のトランジットノードである点は相違する。また、障害が発生する前の時点で、スイッチ101Kがマスタスイッチである。

【0108】

この例において、スイッチ101Kに障害が発生すると、バックアップスイッチであったスイッチ101Lがマスタスイッチに遷移する。一方、スイッチ101Mは、リングのヘルスチェックフレームを用いた障害監視によってリング障害を検出すると、障害動作を実行する。具体的には、スイッチ101Mは、セカンダリポートの状態をフォワーディングに変更し、フラッシュ制御フレームを送信する。これによって、スイッチ101M及びフラッシュ制御フレームを受信したスイッチ101（スイッチ101Lを含む）のMACアドレステーブルがクリアされる。

【0109】

これによって、データフレームの転送経路が、スイッチ101Lからスイッチ101Mを経由してスイッチ101Nに至る経路に切り替えられる。そして、スイッチ101Mは、復旧監視状態へ遷移する。

【0110】

図11は、本発明の実施例1のネットワークシステムにおいてトランジットノードを構成するスイッチが障害から復旧した場合の動作を示す説明図である。

【0111】

具体的には、図11には、図10に示した障害が復旧した場合の動作を示す。スイッチ101Kは、障害から復旧すると、バックアップスイッチとしてスタックに加入する。次に、スイッチ101Mは、リングのヘルスチェックフレームを用いた障害監視によってリング障害の復旧を検出して、復旧動作を実行する。具体的には、スイッチ101Mは、セカンダリポートの状態をブロッキングに変更し、フラッシュ制御フレームを送信する。これによって、スイッチ101M及びフラッシュ制御フレームを受信したスイッチ101（スイッチ101Lを含む）のMACアドレステーブルがクリアされる。

【0112】

これによって、データフレームの転送経路が、スイッチ101Lからスイッチ101Pを経由してスイッチ101Nに至る経路に切り替えられる。そして、スイッチ101Mは、障害監視状態へ遷移する。

【0113】

なお、図10及び図11に示すように、本実施例では、障害発生前にマスタスイッチであったスイッチ101Kは、障害から復旧した後、バックアップスイッチとしてスタックに加入する。しかし、このような動作は一例であり、実際にはスイッチ101Kが再びマスタスイッチとしてスタックに加入してもよい。例えば、スイッチ101Kの復旧が検出されると、スタックプライオリティ302に基づいて、優先度の高いスイッチ101Kがマスタスイッチとなってもよい。その場合、スイッチ101Lはマスタスイッチからバックアップスイッチへ遷移する。ノード102Dがマスタノードである場合も同様である。

【0114】

次に、ここまで説明した動作を実現するためにスイッチ101の各部が実行する具体的な処理について、フローチャートを参照して説明する。

【0115】

図12は、本発明の実施例1のインタフェース制御部214が実行する処理のフローチャートである。

【 0 1 1 6 】

インタフェース制御部 2 1 4 は、ポートの状態変化（リンクアップ又はリンクダウン）を検出すると（ステップ 1 2 0 1 ）、その変化がリンクアップ又はリンクダウンのいずれであるかを判定する（ステップ 1 2 0 2 ）。リンクアップが検出された場合、インタフェース制御部 2 1 4 は、ポートリンク状態管理テーブル 2 2 3 における当該ポート（すなわち状態変化が検出されたポート）のポート状態 5 0 3 を「UP」に変更する（ステップ 1 2 0 3 ）。

【 0 1 1 7 】

次に、インタフェース制御部 2 1 4 は、スタックを構成するスイッチ間でポートリンク状態管理テーブル 2 2 3 を同期する（ステップ 1 2 0 4 ）。例えばスイッチ 1 0 1 K のインタフェース制御部 2 1 4 がいずれかのポート 2 0 1 の状態変化を検出した場合、ステップ 1 2 0 3 で「スイッチ 1 」に対応するいずれかのポート状態 5 0 3 が変更され、ステップ 1 2 0 4 でその変更の内容がスタックリンク 1 0 4 L を経由してスイッチ 1 0 1 L に通知される。スイッチ 1 0 1 L のインタフェース制御部 2 1 4 はその通知に従ってポートリンク状態管理テーブル 2 2 3 の当該ポートのポート状態 5 0 3 を変更する。これによって、スタックを構成する二つのスイッチのポートリンク状態管理テーブル 2 2 3 の内容が同一になる。

【 0 1 1 8 】

次に、インタフェース制御部 2 1 4 は、コンフィグレーションテーブル 2 2 1 又はリングポート状態管理テーブル 2 2 4 を参照して、当該ポートがリングポートであるか否かを判定する（ステップ 1 2 0 5 ）。当該ポートがリングポートである場合、インタフェース制御部 2 1 4 は、スタック状態管理テーブル 2 2 2 を参照して、当該スイッチ（すなわちこの処理を実行しているインタフェース制御部 2 1 4 が属するスイッチ 1 0 1 ）のスタック状態 4 0 2 がマスタスイッチであるか否かを判定する（ステップ 1 2 0 6 ）。当該スイッチのスタック状態 4 0 2 がマスタスイッチである場合、インタフェース制御部 2 1 4 は、ポートリンク状態変更をリング制御部 2 1 3 に通知する（ステップ 1 2 0 7 ）。

【 0 1 1 9 】

次に、インタフェース制御部 2 1 4 は、コンフィグレーションテーブル 2 2 1 を参照して、当該ポートがスタックポートであるか否かを判定する（ステップ 1 2 0 8 ）。当該ポートがスタックポートである場合、インタフェース制御部 2 1 4 は、ポートリンク状態変更をスタック管理 / 制御部 2 1 2 に通知する（ステップ 1 2 0 9 ）。

【 0 1 2 0 】

ステップ 1 2 0 5 において当該ポートがリングポートでないと判定されたか、又は、ステップ 1 2 0 6 において当該スイッチのスタック状態 4 0 2 がマスタスイッチでない（すなわちバックアップスイッチ又は初期状態である）と判定された場合、インタフェース制御部 2 1 4 はステップ 1 2 0 7 を実行しない。ステップ 1 2 0 8 において当該ポートがスタックポートでないと判定された場合、インタフェース制御部 2 1 4 はステップ 1 2 0 9 を実行しない。

【 0 1 2 1 】

ステップ 1 2 0 2 においてリンクダウンが検出されたと判定された場合、インタフェース制御部 2 1 4 は、ポートリンク状態管理テーブル 2 2 3 における当該ポートのポート状態 5 0 3 を「DOWN」に変更する（ステップ 1 2 1 0 ）。インタフェース制御部 2 1 4 がステップ 1 2 1 0 に続いて実行するステップ 1 2 1 1 ~ 1 2 1 6 は、それぞれステップ 1 2 0 4 ~ 1 2 0 9 と同様であるため、説明を省略する。

【 0 1 2 2 】

以上でインタフェース制御部 2 1 4 が実行する処理が終了する。インタフェース制御部 2 1 4 は、ポートの状態変化を検出する度に上記のステップ 1 2 0 1 ~ 1 2 1 6 を実行する。

【 0 1 2 3 】

図 1 3 は、本発明の実施例 1 のプロトコル共通制御部 2 1 1 が実行する処理のフローチ

10

20

30

40

50

ャートである。

【 0 1 2 4 】

プロトコル共通制御部 2 1 1 は、当該スイッチ（すなわちこの処理を実行しているプロトコル共通制御部 2 1 1 が属するスイッチ 1 0 1）がスタックポートを有しているか否かを判定する（ステップ 1 3 0 1）。当該スイッチがスタックポートを有している（すなわち当該スイッチがスタックを構成している）場合、プロトコル共通制御部 2 1 1 は、当該スタックポートからスタックのヘルスチェックフレームを送信する（ステップ 1 3 0 2）。送信されるスタックのヘルスチェックフレームは、そのフレームがスタックのヘルスチェックフレームであることを示す情報、当該スイッチのスタック状態 4 0 2、当該スイッチのスタックプライオリティ 3 0 2、当該スイッチの識別情報（すなわちスイッチ 3 0 1）、及び当該スイッチのスタックポート番号 3 0 3 等のデータを含む。各スイッチ 1 0 1 は、必要に応じて、例えば図 3 ～ 図 6 に示したテーブルの情報のような、それぞれが保持している情報をスタックリンク経由で（例えばスタックのヘルスチェックフレームに含めて）通信することによって、保持している情報を相手スイッチとの間で同期することができる。

10

【 0 1 2 5 】

次に、プロトコル共通制御部 2 1 1 は、当該スイッチがリングのマスタノードであるか否かを判定する（ステップ 1 3 0 3）。当該スイッチがリングのマスタノードである場合、プロトコル共通制御部 2 1 1 は、当該スイッチがスタックのマスタスイッチであるか否かを判定する（ステップ 1 3 0 4）。当該スイッチがスタックのマスタスイッチである場合、プロトコル共通制御部 2 1 1 は、当該スイッチのリングポートからリングのヘルスチェックフレームを送信する（ステップ 1 3 0 5）。リングのヘルスチェックフレームは、少なくとも、そのフレームがリングのヘルスチェックフレームであることを示す情報を含む。

20

【 0 1 2 6 】

ステップ 1 3 0 1 において当該スイッチがスタックポートを有しないと判定された場合、プロトコル共通制御部 2 1 1 はステップ 1 3 0 2 を実行しない。ステップ 1 3 0 3 において当該スイッチがリングのマスタノードでないと判定されたか、又は、ステップ 1 3 0 4 において当該スイッチがスタックのマスタスイッチでないと判定された場合、プロトコル共通制御部 2 1 1 はステップ 1 3 0 5 を実行しない。

30

【 0 1 2 7 】

以上でプロトコル共通制御部 2 1 1 が実行する処理が終了する。プロトコル共通制御部 2 1 1 は、繰り返し（例えば定期的に）上記のステップ 1 3 0 1 ～ 1 3 0 5 を実行する。なお、実際には、プロトコル共通制御部 2 1 1 は、ステップ 1 3 0 1 ～ 1 3 0 2 とステップ 1 3 0 3 ～ 1 3 0 5 とを独立に実行してもよい。例えば、プロトコル共通制御部 2 1 1 がスタックのヘルスチェックフレームを送信する間隔とリングのヘルスチェックフレームを送信する間隔とが異なってもよい。

【 0 1 2 8 】

図 1 4 は、本発明の実施例 1 のスタック管理 / 制御部 2 1 2 が実行する処理のフローチャートである。

40

【 0 1 2 9 】

スタック管理 / 制御部 2 1 2 は、最初に、当該スイッチ（すなわちこの処理を実行しているスタック管理 / 制御部 2 1 2 が属するスイッチ 1 0 1）のスタック状態 4 0 2 が初期状態であるか否かを判定する（ステップ 1 4 0 1）。例えば、当該スイッチがスタックの新たな（又は障害から復旧した）メンバスイッチとして追加された場合、当該スイッチのスタック状態 4 0 2 は初期状態である。

【 0 1 3 0 】

当該スイッチのスタック状態 4 0 2 が初期状態である場合、スタック管理 / 制御部 2 1 2 は、スタックのヘルスチェックフレームを受信したか否かを判定する（ステップ 1 4 0 2）。スタック管理 / 制御部 2 1 2 が所定の時間内にスタックのヘルスチェックフレーム

50

を受信しなかった（すなわちタイムアウトした）場合、当該スイッチが属するスタックの相手スイッチ（例えば当該スイッチがスイッチ 1 0 1 Kである場合、スイッチ 1 0 1 L）に障害が発生している可能性があるため、当該スイッチがマスタスイッチとなる必要がある。このため、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタック状態 4 0 2 をマスタに変更する（ステップ 1 4 0 3）。

【 0 1 3 1 】

ステップ 1 4 0 2 においてスタック管理 / 制御部 2 1 2 が所定の時間内にスタックのヘルスチェックフレームを受信した場合、スタック管理 / 制御部 2 1 2 は、受信したスタックのヘルスチェックフレームに含まれるスタック状態を参照して、相手スイッチの状態がマスタであるか否かを判定する（ステップ 1 4 0 4）。相手スイッチの状態がマスタである場合、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタック状態 4 0 2 をバックアップに変更する（ステップ 1 4 0 5）。例えば、図 9 又は図 1 1 に示すように障害から復旧したスイッチ 1 0 1 L のスタック管理 / 制御部 2 1 2 は、相手スイッチ 1 0 1 K がマスタスイッチであるため、ステップ 1 4 0 5 においてバックアップスイッチとしてスタックに加入される。

10

【 0 1 3 2 】

ステップ 1 4 0 4 において、相手スイッチの状態がマスタでないと判定された場合、スタック管理 / 制御部 2 1 2 は、当該スイッチと相手スイッチのスタックプライオリティ 3 0 2 を比較する（ステップ 1 4 0 6）。その結果、当該スイッチの優先度が高い場合、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタック状態 4 0 2 をマスタに変更し（ステップ 1 4 0 7）、当該スイッチの優先度が低い場合、当該スイッチのスタック状態 4 0 2 をバックアップに変更する（ステップ 1 4 0 8）。

20

【 0 1 3 3 】

ステップ 1 4 0 5、1 4 0 7 又は 1 4 0 8 でスタック状態 4 0 2 が変更された場合、次に、スタック管理 / 制御部 2 1 2 は、スタックを構成するスイッチ間でポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 を同期する（ステップ 1 4 0 9）。この同期は、当該スイッチが有するポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 に含まれる各情報のうち当該スイッチに関する各情報を相手スイッチへ送信し、相手スイッチが送信した相手スイッチに関するポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 に含まれる各情報によって、ポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 に含まれる各情報のうち相手スイッチに関する各情報を更新することによって実行される。また、ポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 に含まれる各情報のうち当該スイッチに関する各情報の全てを更新対象として送信して更新してもよいし、変更された各情報のみを更新対象として送信して更新してもよい。例えば、当該スイッチが変更されたスタック状態 4 0 2 をスタックのヘルスチェックフレームに含めて送信し、相手スイッチが受信したスタック状態 4 0 2 に基づいてスタック状態管理テーブル 2 2 2 を変更することによって実行される。また、本実施例では明記していないが、スタックのヘルスチェックフレームは、各スイッチが固有で持つ情報（MAC 学習情報など）を定期的に同期することに用いる。

30

40

【 0 1 3 4 】

ステップ 1 4 0 3 でスタック状態 4 0 2 が変更されたか、又はステップ 1 4 0 9 でテーブルが同期した場合、次に、スタック管理 / 制御部 2 1 2 は、当該スイッチが追加されたことをリング制御部 2 1 3 に通知する（ステップ 1 4 1 0）。

【 0 1 3 5 】

ステップ 1 4 0 1 において、当該スイッチのスタック状態 4 0 2 が初期状態でない（すなわちマスタ又はバックアップである）と判定された場合、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタックポートのポート状態 5 0 3 が変更されたか否かを判定する（ステップ 1 4 1 1）。この状態変更は、図 1 2 のステップ 1 2 0 9 又は 1 2 1 9 において通知されるものである。ポート状態 5 0 3 が変更されていない場合、スタック管理 / 制御

50

部 2 1 2 は、スタックのヘルスチェックフレームを受信したか否かを判定する（ステップ 1 4 1 2）。スタックのヘルスチェックフレームを受信した場合、スタック管理 / 制御部 2 1 2 は、スタックを構成するスイッチ間でポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 を同期する（ステップ 1 4 1 3）。同期方法については、ステップ 1 4 0 9 と同様である。

【 0 1 3 6 】

ステップ 1 4 1 1 において、当該スイッチのスタックポートのポート状態 5 0 3 が変更されたと判定された場合、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタックポートの変更されたポート状態 5 0 3 を判定する（ステップ 1 4 1 4）。その結果、全てのスタックポートのポート状態 5 0 3 が「DOWN」であると判定された場合、当該スイッチは、相手スイッチの障害又はスタックリンクの障害等によって相手スイッチと通信できない状態にある。この場合、当該スイッチがバックアップスイッチであればマスタスイッチとなる必要があるため、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタック状態 4 0 2 を判定する（ステップ 1 4 1 5）。なお、ステップ 1 4 1 2 においてタイムアウトが検出された場合にもスタックが上記と同様の状態であるため、スタック管理 / 制御部 2 1 2 はステップ 1 4 1 5 を実行する。

【 0 1 3 7 】

ステップ 1 4 1 5 においてスタック状態 4 0 2 がバックアップであると判定された場合、スタック管理 / 制御部 2 1 2 は、当該スイッチのスタック状態 4 0 2 をマスタに変更する（ステップ 1 4 1 6）。次に、スタック管理 / 制御部 2 1 2 は、当該スイッチのリングモード 3 0 5 がマスタであるか否かを判定する（ステップ 1 4 1 7）。リングモード 3 0 5 がマスタである場合、スタック管理 / 制御部 2 1 2 は、当該スイッチが新たにマスタスイッチとなったことをリング制御部 2 1 3 に通知する（ステップ 1 4 1 8）。

【 0 1 3 8 】

ステップ 1 4 1 7 においてリングモード 3 0 5 がトランジットであると判定された場合、スタック管理 / 制御部 2 1 2 はステップ 1 4 1 8 を実行しない。ステップ 1 4 1 5 においてスタック状態 4 0 2 がマスタであると判定された場合、スタック管理 / 制御部 2 1 2 はステップ 1 4 1 6 ~ 1 4 1 8 を実行しない。

【 0 1 3 9 】

ステップ 1 4 1 4 において、全てのスタックポートのポート状態 5 0 3 が「DOWN」でない（すなわち少なくとも一つのスタックポートのポート状態 5 0 3 が「UP」である）と判定された場合、スタック管理 / 制御部 2 1 2 は、スタックのヘルスチェックフレームを受信したか否かを判定する（ステップ 1 4 1 9）。スタックのヘルスチェックフレームを受信した場合、スタック管理 / 制御部 2 1 2 は、スタックを構成するスイッチ間でポートリンク状態管理テーブル 2 2 3 及びスタック状態管理テーブル 2 2 2 を同期する（ステップ 1 4 2 0）。この同期方法については、ステップ 1 4 0 9 と同様である。例えば、当該スイッチが、ステップ 1 4 1 1 で変更されたと判定されたポート状態 5 0 3 及びそれに対応するスイッチ 5 0 1 及びポート番号 5 0 2 の値をスタックのヘルスチェックフレームに含めて送信し、相手スイッチが受信した値に基づいてポートリンク状態管理テーブル 2 2 3 を変更することによって実行される。

【 0 1 4 0 】

ステップ 1 4 1 9 においてタイムアウトが検出された場合、スタック管理 / 制御部 2 1 2 はステップ 1 4 2 0 を実行しない。

【 0 1 4 1 】

以上でスタック管理 / 制御部 2 1 2 の処理が終了する。図 1 4 のフローは、定期的に行ってもよく、リングネットワークに新たにスイッチを追加した場合や、当該スイッチを再起動した場合、または外部からの処理の開始指示の入力を受けた場合に行ってもよい。

【 0 1 4 2 】

図 1 5 A 及び図 1 5 B は、本発明の実施例 1 のリング制御部 2 1 3 が実行する処理のフローチャートである。

【0143】

最初に、リング制御部213は、スイッチ追加通知を受けたか否かを判定する（ステップ1501）。このスイッチ追加通知は、スタック管理/制御部212がステップ1410で送信するものである。スイッチ追加通知を受けた場合、リング制御部213は、当該スイッチ（すなわちこの処理を実行しているリング制御部213が属するスイッチ101）のスタック状態402を判定する（ステップ1502）。スタック状態402がバックアップである場合、リング制御部213は、当該スイッチと当該スイッチが属するスタックの相手スイッチとの間でリングポート状態管理テーブル224を同期する（ステップ1503）。この同期方法については、図14のステップ1409と同様である。

【0144】

ステップ1502において、スタック状態402がマスタであると判定された場合、リング制御部213は、当該スイッチのリングモード305を判定する（ステップ1504）。リングモード305がトランジットである場合、リング制御部213は、リングポート状態を初期化する（プライマリポート、セカンダリポート共にブロッキング）する（ステップ1505）。一方、リングモード305がマスタである場合、リング制御部213は、リングポート状態を初期化（プライマリポートはフォワーディング、セカンダリポートはブロッキング）し（ステップ1506）、さらに、リングのフラッシュ制御フレームを送信する（ステップ1507）。

【0145】

ステップ1501においてスイッチ追加通知を受けなかったと判定された場合、リング制御部213は、当該スイッチのリングモード305を判定する（ステップ1508）。当該スイッチのリングモード305がマスタである場合、リング制御部213は、当該スイッチのスタック状態402を判定する（ステップ1509）。当該スイッチのスタック状態402がマスタである場合、リング制御部213は、リングのヘルスチェックフレームを受信したか否かを判定する（ステップ1510）。ステップ1510においてタイムアウトが検出された場合、当該スイッチが属するリングのいずれかの部位（例えば当該スイッチが属するスタックの相手スイッチ、又は当該リングのトランジットノード）に障害が発生しているため、リング制御部213は、当該スイッチのリング障害状態606を障害に変更する（ステップ1511）。

【0146】

次に、リング制御部213は、当該スイッチが属するノードのセカンダリポートのリングポート状態605がブロッキングであるか否かを判定する（ステップ1512）。セカンダリポートのリングポート状態605がブロッキングである場合、リング制御部213は、セカンダリポートのリングポート状態605をフォワーディングに変更する（ステップ1513）。これによって、図8に示すセカンダリポートの状態変更が実現される。次に、リング制御部は、リングのフラッシュ制御フレームを送信する（ステップ1514）。

【0147】

ここで、図7及び図8とは異なるいくつかの場合におけるステップ1513の処理を説明する。図8にはバックアップスイッチであるスイッチ101Lに障害が発生した例を示しているが、スイッチ101Lは正常であり、いずれかのトランジットノード（例えばスイッチ101M等）に障害が発生した場合にも、上記と同様の処理が実行され、セカンダリポート201Aのリングポート状態605がフォワーディングに変更される。その結果、プライマリポート201D及びセカンダリポート201Aのいずれのリングポート状態605もフォワーディングとなる。これによって、通信経路の切り替えが可能となる一方で、トランジットノードの障害のため、データフレームが循環するループは形成されない。

【0148】

一方、図7に示すように、マスタスイッチ101Kのポート201Aがプライマリポートであり、バックアップスイッチ101Lのポート201Dがセカンダリポートである場

10

20

30

40

50

合において、マスタスイッチ 101K 及びバックアップスイッチ 101L のいずれでもなく、いずれかのトランジットノードに障害が発生したとき、マスタスイッチ 101K のリング制御部 213 は、ステップ 1513 において、ポート 201D の状態をブロッキングからフォワーディングに変更することを決定し、その変更の指示を、スタックポート 201B を介してバックアップスイッチ 101L に送信する。バックアップスイッチ 101L のリング制御部 213 は、受信した指示に従ってポート 201D の状態をフォワーディングに変更する。これによって、上記の例と同様に、ループを形成することなく、通信経路の切り替えが可能となる。

【0149】

図 7 に示すように、マスタスイッチ 101K のポート 201A がプライマリポートであり、バックアップスイッチ 101L のポート 201D がセカンダリポートである場合において、マスタスイッチ 101K ではなくバックアップスイッチ 101L に障害が発生しているときも、マスタスイッチ 101K のリング制御部 213 は、ステップ 1513 において、ポート 201D の状態をブロッキングからフォワーディングに変更することを決定し、その変更の指示を送信する。しかし、バックアップスイッチ 101L のリング制御部 213 は、障害のため、送信された指示に従う制御を実行できない場合がある。その場合にも、バックアップスイッチ 101L 以外のスイッチを利用した通信経路の切り替えが可能となる一方で、バックアップスイッチ 101L の障害のため、ループは形成されない。

【0150】

ステップ 1509 において、当該スイッチのスタック状態 402 がバックアップであると判定された場合、リング制御部 213 は、マスタ切り替え通知を受けたか否かを判定する（ステップ 1515）。この通知は、スタック管理 / 制御部 212 がステップ 1418（図 14 参照）で送信するものである。マスタ切り替え通知を受けた場合、当該スイッチがバックアップスイッチからマスタスイッチに変更された（すなわち当該スイッチの相手スイッチに障害が発生しており、その結果、当該スイッチが属するリングにも障害が発生している）ため、リング制御部 213 は、ステップ 1512 以降の処理を実行する。これによって、図 7 に示すセカンダリポートの状態変更が実現される。一方、マスタ切り替え通知を受けていない場合、当該スイッチは依然としてバックアップスイッチであるため、リング制御部 213 はステップ 1510 ~ 1514 を実行しない。

【0151】

ステップ 1512 において、セカンダリポートのリングポート状態 605 がフォワーディングであると判定された場合、リング制御部 213 は、リングポート状態 605 を変更する必要がないため、ステップ 1513 及び 1514 を実行しない。これによって、スタックの相手スイッチ又はトランジットノードに障害が発生する前からセカンダリポートがフォワーディングであった場合には、障害が発生した後もフォワーディングの状態が維持されるように、セカンダリポートが制御される。

【0152】

ステップ 1510 においてリングのヘルスチェックフレームを受信した場合、リング制御部 213 は、当該スイッチのリング障害状態 606 が障害であるか否かを判定する（ステップ 1516）。リング障害状態 606 が障害である場合、当該スイッチが属するリングに障害が発生していたが、ステップ 1510 においてリングのヘルスチェックフレームを受信したことによってその障害が解消されたことが確認されたため、リング制御部 213 は、当該スイッチのリング障害状態 606 を正常に変更する（ステップ 1517）。

【0153】

次に、リング制御部 213 は、セカンダリポートのリングポート状態 605 をブロッキングに変更し（ステップ 1518）、プライマリポートのリングポート状態 605 がブロッキングであった場合にはそれをフォワーディングに変更し（ステップ 1519）、リングのフラッシュ制御フレームを送信する（ステップ 1520）。これによって、図 9 に示すセカンダリポートの状態変更が実現される。なお、ステップ 1513 の場合と同様に、ステップ 1518 又は 1519 の制御対象のポートがバックアップスイッチ側にある場合

10

20

30

40

50

には、マスタスイッチのリング制御部 2 1 3 がリングポート状態 6 0 5 を変更する指示をバックアップスイッチに送信し、バックアップスイッチのリング制御部 2 1 3 がその指示に従ってリングポート状態 6 0 5 を変更する。

【 0 1 5 4 】

ステップ 1 5 1 6 において、当該スイッチのリング障害状態 6 0 6 が正常であると判定された場合、既に、プライマリポートのリングポート状態 6 0 5 はフォワーディングに、セカンダリポートのリングポート状態 6 0 5 はブロッキングに設定されている。この状態を維持するため、リング制御部 2 1 3 はステップ 1 5 1 7 ~ 1 5 2 0 を実行する必要がない。

【 0 1 5 5 】

以上のように、本実施例のスタック構成のマスタノードにおいて、マスタスイッチは、リングの状態が正常であるときに、二つのリングポートの一方の状態をフォワーディング、もう一方の状態をブロッキングに制御し、リングの障害を検出したときに、両方の状態をフォワーディングに制御する（ステップ 1 5 1 3）。具体的には、マスタスイッチは、マスタスイッチのリングポートの状態がブロッキングである場合、リングの障害を検出したときにそのポートの状態をフォワーディングに変更し、バックアップスイッチのリングポートの状態がブロッキングである場合、リングの障害を検出したときに、マスタスイッチのリングポートの状態をフォワーディングに維持しながら、バックアップスイッチにリングポートの状態をフォワーディングに変更する指示を送信する。そして、マスタノードのマスタスイッチは、リングが障害から回復したことを検出したときに、再び一方の状態をフォワーディング、もう一方の状態をブロッキングに制御する（ステップ 1 5 1 8、1 5 1 9）。

【 0 1 5 6 】

本実施例では、マスタノードの二つのリングポートの一方にプライマリポート、もう一方にセカンダリポートという属性（リングポート 6 0 4）が与えられ、プライマリポートの状態は、リングの正常時及び障害時のいずれもフォワーディングとなるように制御され、セカンダリポートの状態は、リングの正常時にブロッキング、障害時にフォワーディングとなるように制御される。しかし、このような制御方法は一例であり、正常時に一方がフォワーディング、もう一方がブロッキングとなるように制御され、障害時に両方がフォワーディングとなるように制御される限り、どのような制御方法を用いてもよい。

【 0 1 5 7 】

ステップ 1 5 0 8 において、当該スイッチのリングモード 3 0 5 がトランジットであると判定された場合、リング制御部 2 1 3 は、当該スイッチのスタック状態 4 0 2 を判定する（ステップ 1 5 2 1）。当該スイッチのスタック状態 4 0 2 がマスタである場合、リング制御部 2 1 3 は、当該スイッチが属するトランジットノードのいずれかのリングポートのポート状態 5 0 3 が変更されたか否かを判定する（ステップ 1 5 2 2）。この状態変更は、図 1 2 のステップ 1 2 0 7 又は 1 2 1 4 において通知されるものである。いずれかのリングポートのポート状態 5 0 3 が変更された場合、リング制御部 2 1 3 は、当該変更が UP から DOWN への変更であるか否かを判定する（ステップ 1 5 2 3）。当該変更が UP から DOWN への変更である場合、リング制御部 2 1 3 は、ポート状態 5 0 3 が UP から DOWN へ変更されたリングポートのリングポート状態 6 0 5 をブロッキングに変更する（ステップ 1 5 2 4）。このように、ステップ 1 5 2 2 からステップ 1 5 2 4 を実行することで、ポート状態 5 0 3 が変更されたリングポートのリングポート状態 6 0 5 を初期化（ブロッキング）し、リングネットワークにループが発生することを防ぐ。

【 0 1 5 8 】

当該トランジットノードのリングポートのポート状態 5 0 3 が変更されなかったか（ステップ 1 5 2 2）、又は、変更されたとしても当該変更が DOWN から UP への変更であった場合（ステップ 1 5 2 3）には、リング制御部 2 1 3 はステップ 1 5 2 4 を実行しない。

【 0 1 5 9 】

次に、リング制御部 213 は、フラッシュ制御フレームを受信したか否かを判定する（ステップ 1525）。フラッシュ制御フレームを受信した場合、リング制御部 213 は、当該スイッチが属するトランジットノードにブロッキング状態のリングポートがあるか否かを判定する（ステップ 1526）。ブロッキング状態のリングポートがある場合、当該リングポートのリングポート状態 605 をフォワーディングに変更する（ステップ 1527）。

【0160】

当該スイッチがフラッシュ制御フレームを受信しなかったか（ステップ 1525）、又は、受信したとしても当該スイッチが属するトランジットノードにブロッキング状態のリングポートがない場合（ステップ 1526）には、リング制御部 213 はステップ 1527 を実行しない。

10

【0161】

ステップ 1521 において、当該スイッチのスタック状態 402 がバックアップであると判定された場合、リング制御部 213 は、マスタ切り替え通知（図 14 のステップ 1418）を受けたか否かを判定する（ステップ 1528）。マスタ切り替え通知を受けた場合、当該スイッチがバックアップスイッチからマスタスイッチに変更されたため、リング制御部 213 は、ポートの設定内容に基づいてリングポート状態 605 を更新する（ステップ 1529）。

【0162】

具体的には、当該スイッチがバックアップスイッチからマスタスイッチに切り替えられる前に、マスタスイッチによって当該スイッチのリングポートの状態がフォワーディングに設定されていた場合、リングポート状態 605 をフォワーディングに更新し、ブロッキングに設定されていた場合、リングポート状態 605 をブロッキングに更新する。

20

【0163】

一方、マスタ切り替え通知を受けていない場合、当該スイッチは依然としてバックアップスイッチであるため、リング制御部 213 はステップ 1529 を実行しない。

【0164】

図 15 のフローは、定期的に行ってもよく、リングネットワークに新たにスイッチを追加した場合や、当該スイッチを再起動した場合、または外部からの処理の開始指示の入力を受けた場合に行ってもよい。

30

【0165】

以上の本発明の実施例 1 によれば、リングネットワークのノードにスタックを適用し、スタックを構成する各スイッチが、相手スイッチの状態（すなわち障害が発生しているか否か）、スタックにおける自スイッチの属性（すなわちマスタスイッチであるか否か）及び自スイッチが属するノードの属性（すなわちマスタノードであるか否か）に応じてポートのデータフレーム転送を適切に制御することによって、リングネットワークの耐障害性を高めることができる。

【実施例 2】

【0166】

次に、本発明の実施例 2 について説明する。以下に説明する相違点を除き、実施例 2 のネットワークシステムの各部は、実施例 1 の同一の符号を付された各部と同一の機能又は内容を有するため、それらの説明は省略する。

40

【0167】

図 16 は、本発明の実施例 2 のネットワークシステムにおけるトランジットノードへのスイッチの増設の説明図である。

【0168】

図 16 を参照して、リングを構成するトランジットノードのうち、単独のスイッチによって構成される（すなわちスタック構成でない）トランジットノードを、新たなスイッチの増設によってスタック構成のトランジットノードを構築する手順を説明する。

【0169】

50

本実施例のネットワークシステムは、図16に示すリング103Eを含む。図16(1)には、スイッチを増設する前の正常状態のリング103Eを示す。このリング103Eは、スイッチ101Q~101Vがリング状に接続されたものであり、スイッチ101Qがマスタノードであり、スイッチ101R~101Vの各々が一つのトランジットノードである。各スイッチ101Q~101Vは実施例1のスイッチ101と同様であるため、詳細な説明は省略する。

【0170】

実施例1で説明したマスタノードと同様、正常状態において、マスタノードであるスイッチ101Qのプライマリポート201I及びセカンダリポート201Jの状態はそれぞれフォワーディング及びブロッキングである。この状態において、スイッチ101Qがス

10

【0171】

上記のリング103Eにおいてスイッチ101Tからなるトランジットノードに新たにスイッチを増設することによって、当該トランジットノードをスタック構成のノードに変更する手順を説明する。まず、図16(2)に示すように、増設対象であるスイッチ101Tに隣接するスイッチ101S及び101Uの、スイッチ101Tに接続されているリングポートのシャットダウンが実行される。これは、例えばユーザがスイッチ101Sにポート201Kを対象とするshutdownコマンド(図3の説明参照)を入力し、スイッチ101Uにポート201Nを対象とするshutdownコマンドを入力することによって実行され

20

【0172】

これによって、スイッチ101Sのコンフィグレーションテーブル221のポート201Kに関するポートシャットダウン設定307がshutdownに変更され、さらに、ポートリンク状態管理テーブル223のポート201Kに関するポート状態503がDOWNに変更される。スイッチ101Uのポート201Nについても同様である。

【0173】

これによって、リングのヘルスチェックフレームがリング103Eを1周することができなくなるため、スイッチ101Qはリング103Eの障害を検出する。その結果、スイッチ101Qは、図16(3)に示すように、セカンダリポート201Jの状態をフォ

30

【0174】

さらに、スイッチ101Tのリングポートのうち一方が削除される。図16(3)の例では、スイッチ101Tの二つのリングポート201L及び201Mのうちスイッチ101S側のポート201Lが削除される。具体的には、この削除は、例えば、ユーザがポート201Kと201Lとを接続していたネットワークケーブルを物理的に取り外し、さら

40

【0175】

次に、新たなスイッチ101Wを追加することで、スイッチ101T及び101Wからなるスタック構成のノード102Eが構築される(図16(4))。具体的には、例えばユーザがスイッチ101Wをスイッチ101Tに接続し、さらに、それぞれにコマンドを入力することによってスイッチ101T及び101Wのコンフィグレーションテーブル221を作成する(図3の説明参照)。そして、スイッチ101T及び101Wが各々図14の動作フローを実行してスタック構成のノード102Eが構築される。

【0176】

50

次に、ユーザがスイッチ 101W のリングポート 201O とスイッチ 101S のリングポート 201K とを接続する (図 16 (5))。

【0177】

次に、構築されたノード 102E に隣接するスイッチ 101S 及び 101U の、ノード 102E に接続されているリングポートのシャットダウン解除が実行される (図 16 (6))。これは、例えばユーザがスイッチ 101S にポート 201K を対象とする No shutdown コマンド (図 3 の説明参照) を入力し、スイッチ 101U にポート 201N を対象とする No shutdown コマンドを入力することによって実行される。その結果、リングのヘルスチェックフレームがリング 103E を 1 周できるようになるため、スイッチ 101Q はリング 103E が障害から復旧したことを検出する (図 15 のステップ 1510)。この時点で、構築されたノード 102E のリングポート 201O 及び 201M、並びにそれらに接続されるリングポート 201K 及び 201N の状態はブロッキングである (図 15 のステップ 1521 からステップ 1524)。

10

【0178】

リング 103E の復旧を検出したスイッチ 101Q は、セカンダリポート 201J の状態をブロッキングに変更する (図 16 (7))。さらに、スイッチ 101Q がフラッシュ制御フレームを送信することによって、ポート 201K、201M、201N 及び 201O の状態がブロッキングからフォワーディングに変更される。これによって、リング 103E の状態が正常になる (図 15 のステップ 1516 からステップ 1520、図 15 のステップ 1525 からステップ 1527)。

20

【0179】

図 16 には、リング 103E の一部として稼働していたスイッチ 101T に新たにスイッチ 101W を追加することによってスタック構成のノード 102E を構築する例を示したが、初期状態においてスイッチ 101T を含まないリング 103E に、スイッチ 101T 及び 101W からなるスタック構成のノード 102E を追加することもできる。具体的には、例えば初期状態においてスイッチ 101S がスイッチ 101U に接続されていた場合に、それらのスイッチのリングポート 201K 及び 201N をシャットダウンし、それらの間にスタック構成のノード 102E を接続し、シャットダウンを解除することで、ノード 102E が追加される。このようなシャットダウン、シャットダウンの解除及びそれらの前後の処理は、図 16 に示したものと同様の手順で実現できるため、詳細な説明を省略する。

30

【0180】

図 17 は、本発明の実施例 2 のネットワークシステムにおけるトランジットノードからのスイッチの減設の説明図である。

【0181】

図 17 を参照して、リングを構成するトランジットノードのうち、スタック構成のトランジットノードから一方のスイッチを減設することによって、当該スタック構成のトランジットノードを単一のスイッチによるトランジットノードに変更する手順を説明する。

【0182】

減設前のリング 103E は、図 17 (1) に示すように、スイッチ 101Q ~ 101W がリング状に接続されたものであり、スイッチ 101Q がマスタノードであり、スイッチ 101R、101S、101U 及び 101V の各々が一つのトランジットノードである。一方、スイッチ 101S と 101U の間に接続されたトランジットノード 102E は、スイッチ 101T 及び 101W からなるスタック構成のノードである。各スイッチ 101Q ~ 101W は図 16 に示したものと同様であってよい。また、リング 103E は、初めから図 17 (1) に示される形態に構築されたものであってもよいし、図 16 に示す手順によって構築されたものであってもよい。

40

【0183】

上記のリング 103E においてトランジットノード 102E からスイッチ 101W を減設することによって、当該トランジットノードをスイッチ 101T のみからなるノードに

50

変更する手順を説明する。まず、図 17 (2) に示すように、減設対象であるスイッチ 101 T に隣接するスイッチ 101 S 及び 101 U の、ノード 102 E に接続されているリングポート 201 K 及び 201 N のシャットダウンが実行される。これは、図 16 (2) と同様に実行される。その結果、スイッチ 101 Q は、リング 103 E の障害を検出する (図 15 のステップ 1510 からステップ 1511) 。

【 0184 】

そして、スイッチ 101 Q は、セカンダリポート 201 J の状態をフォワーディングに変更する (図 17 (3)) 。さらに、スイッチ 101 Q は、フラッシュ制御フレームを送信する (図 15 のステップ 1512 から 1514) 。これによって、ノード 102 E がリング 103 E から切り離された後も、リング 103 E 上の通信経路のうちノード 102 E を含まない通信経路を経由したデータフレームの通信を継続することが可能になる。さらに、スイッチ 101 W のリングポート 201 O が削除される。これらの手順は、図 16 (3) と同様に実行される。

【 0185 】

次に、スイッチ 101 W が取り外される。 (図 17 (3) 、 (4)) 。例えばユーザがスイッチ 101 W をスイッチ 101 T から切り離し、さらに、スイッチ 101 T にコマンドを入力することによってスイッチ 101 T のコンフィグレーションテーブル 221 を作成する (図 3 の説明参照) 。なお、スイッチ 101 W 取り外し後、スイッチ 101 T が、図 14 の動作フローを実行してスタック構成のノード 102 E を解除することも可能である。具体的には、ステップ 1401 にて「その他」と判断し、ステップ 1414 にて「全スタックポート DOWN」と判断することで、スタック状態としては「マスタ」であり、リングモードとしては「トランジット」であるスイッチ 101 T となる。そして、ユーザがスイッチ 101 T のリングポート 201 L とスイッチ 101 S のリングポート 201 K とを接続する (図 17 (5)) 。

【 0186 】

次に、スイッチ 101 T に隣接するスイッチ 101 S 及び 101 U の、スイッチ 101 T に接続されているリングポートのシャットダウン解除が実行される (図 17 (6)) 。この処理は、 (図 16 (6)) と同様に実行される。スイッチ 101 Q はリング 103 E が障害から復旧したことを検出する (図 15 のステップ 1510) 。この時点で、ポート 201 K ~ 201 N の状態はブロッキングである (図 15 のステップ 1521 からステップ 1524) 。

【 0187 】

リング 103 E の復旧を検出したスイッチ 101 Q は、セカンダリポート 201 J の状態をブロッキングに変更する (図 17 (7)) 。さらに、スイッチ 101 Q がフラッシュ制御フレームを送信することによって、ポート 201 K ~ 201 N の状態がブロッキングからフォワーディングに変更される。これによって、リング 103 E の状態が正常になる (図 15 のステップ 1516 からステップ 1520 、図 15 のステップ 1521 からステップ 1527) 。

【 0188 】

図 17 には、スタック構成されたノード 102 E からスイッチ 101 W を減設し、スイッチ 101 T をトランジットノードとして残す例を示したが、ノード 102 E 全体を減設し、スイッチ 101 S をスイッチ 101 U に接続することによって 5 個のスイッチからなるリング 103 E を構築することもできる。これは、例えば、スイッチ 101 T のポート 201 K 及びスイッチ 101 U のポート 201 N をシャットダウンし、ポート 201 K とポート 201 N とを新たに接続し、シャットダウンを解除することで実現される。このようなシャットダウン、シャットダウンの解除及びそれらの前後の処理は、図 17 に示したものと同様の手順で実現できるため、詳細な説明を省略する。

【 0189 】

以上のように、実施例 2 によれば、リングを経由するデータフレームの通信を継続したまま、リングのトランジットノードをスタック構成のトランジットノードに変更する増設

10

20

30

40

50

又はスタック構成のトランジットノードを単独のスイッチからなるトランジットノードに変更する減設を実行することができる。

【0190】

なお、本発明は上述した実施例に限定されるものではなく、様々な変形例が含まれる。例えば、上記した実施例は本発明を分かりやすく説明するために詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。また、ある実施例の構成の一部を他の実施例の構成に置き換えることが可能であり、また、ある実施例の構成に他の実施例の構成を加えることも可能である。また、各実施例の構成の一部について、他の構成の追加・削除・置換をすることが可能である。

【0191】

上記の各構成、機能、処理部、処理手段等は、それらの一部または全部を、例えば集積回路で設計する等によってハードウェアで実現してもよい。また、上記の各構成、機能等は、プロセッサがそれぞれの機能を実現するプログラムを解釈し、実行することによってソフトウェアで実現してもよい。各機能を実現するプログラム、テーブル、ファイル等の情報は、メモリ、ハードディスクドライブ、SSD (Solid State Drive) 等の記憶装置、または、ICカード、SDカード、DVD等の計算機読み取り可能な非一時的データ記憶媒体に格納することができる。

【0192】

また、リングプロトコルは、上述のプロトコルには限定されず、たとえば、EAPS (Ethernet Automatic Protection System (Ethernetは登録商標)) のような、図1のような複数のノードをリング状に接続したネットワークにおける障害の検出及びそれに伴う経路切り替えを行う冗長化プロトコルであってもよい。

【0193】

また、図面には、実施例を説明するために必要と考えられる制御線及び情報線を示しており、必ずしも、本発明が適用された実際の製品に含まれる全ての制御線及び情報線を示しているとは限らない。実際にはほとんど全ての構成が相互に接続されていると考えてもよい。

【符号の説明】

【0194】

101A ~ 101W スイッチ
 102A ~ 102E ノード
 103A ~ 103E リング
 104A ~ 104S リンク
 201A ~ 201O ポート
 210 制御部
 211 プロトコル共通制御部
 212 スタック管理 / 制御部
 213 リング制御部
 214 インタフェース制御部
 220 メモリ
 221 コンフィグレーションテーブル
 222 スタック状態管理テーブル
 223 ポートリンク状態管理テーブル
 224 リングポート状態管理テーブル
 230 転送エンジン

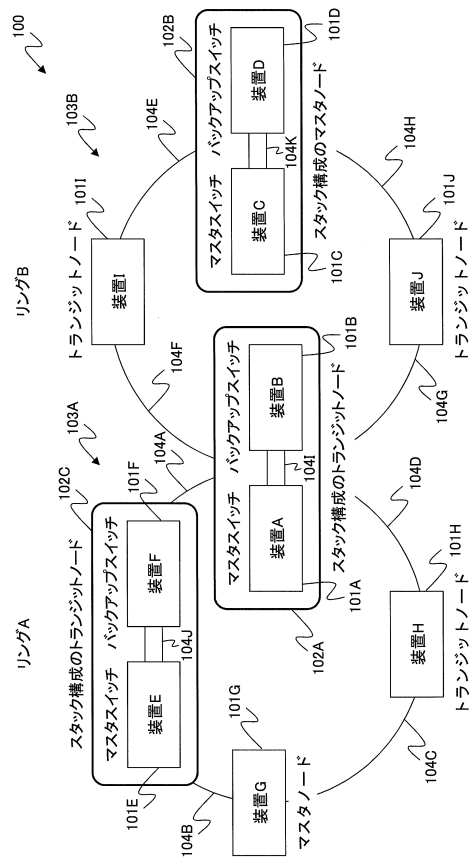
10

20

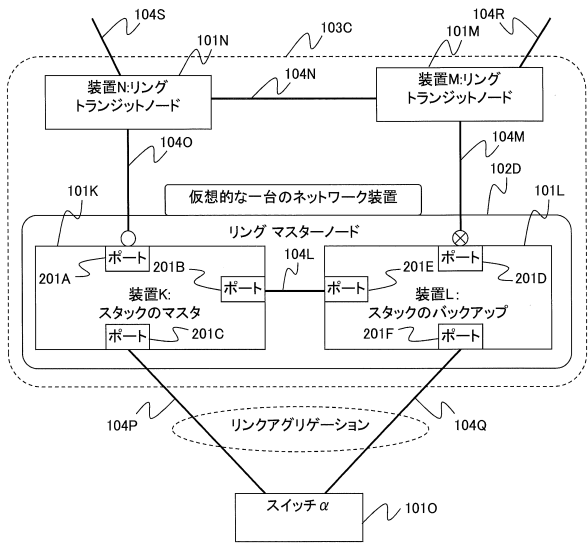
30

40

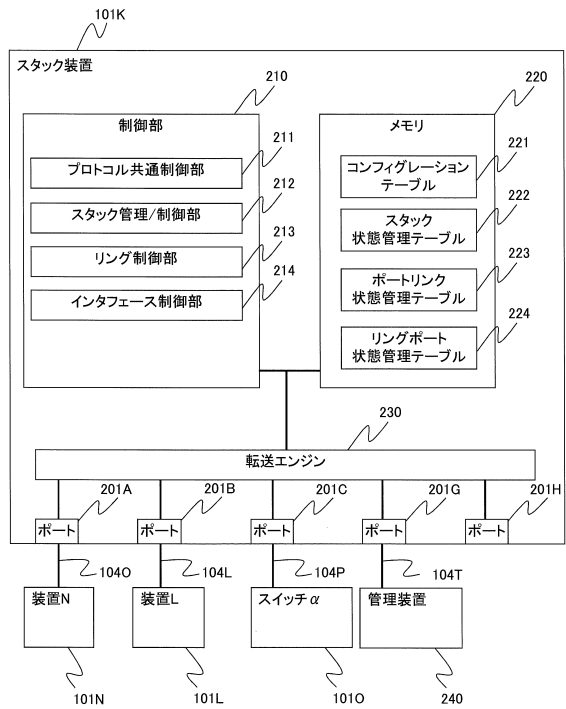
【図 1】



【図 2 A】



【図 2 B】



【図 3】

コンフィグレーションテーブル

301	302	303	304	305	306	307	221
スイッチ	スタック プライオリティ	スタックポート 番号	リング ID	リング モード	リングポート 番号	ポート シャットダウン 設定	
スイッチ1	20	1/0/1	2	マスタ	1/0/40	No shutdown	
スイッチ2	10	2/0/5	2	マスタ	2/0/35	No shutdown	

【図 4】

スタック状態管理テーブル

401	402	222
スイッチ	スタック状態	
スイッチ1	マスタ	
スイッチ2	バックアップ	

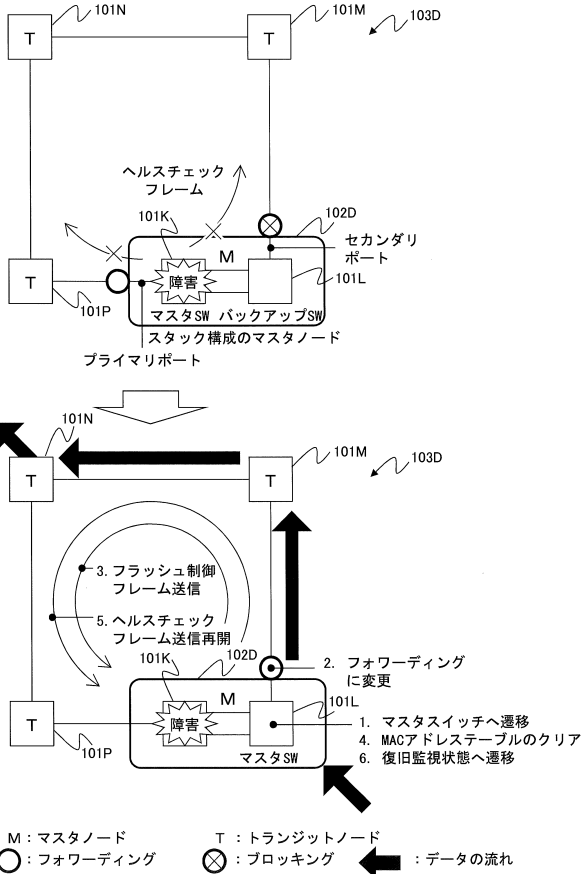
【図 5】

ポートリンク状態管理テーブル		
501 スイッチ	502 ポート番号	503 ポート状態
スイッチ1	1/0/1	UP
⋮	⋮	⋮
スイッチ1	1/0/40	UP
⋮	⋮	⋮
スイッチ2	2/0/5	UP
⋮	⋮	⋮
スイッチ2	2/0/35	UP
⋮	⋮	⋮

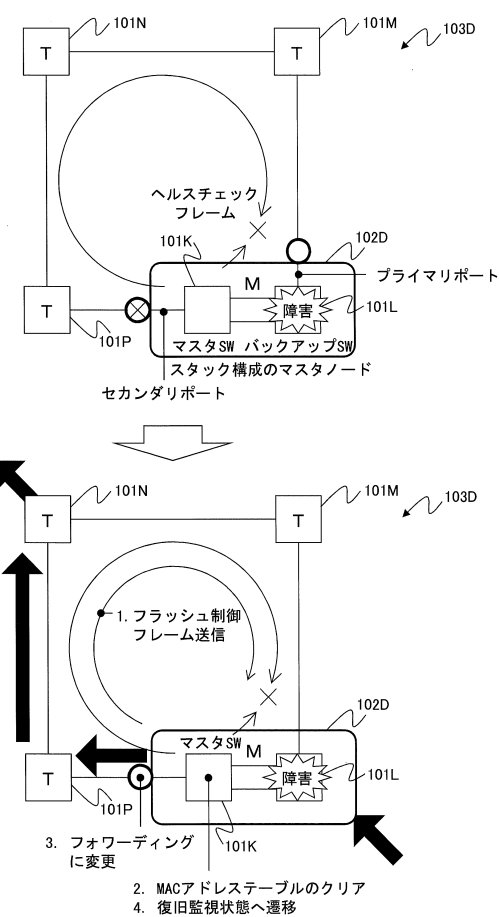
【図 6】

リングポート状態管理テーブル					
601 スイッチ	602 リング ID	603 リングポート番号	604 リングポート	605 リングポート状態	606 リング障害状態
スイッチ1	2	1/0/40	プライマリ	フォワーディング	正常
スイッチ2	2	2/0/35	セカンダリ	ブロッキング	正常

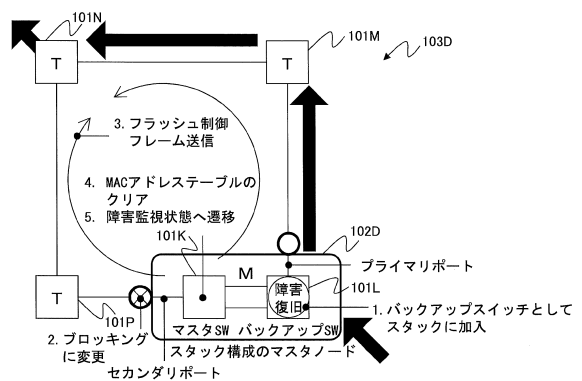
【図 7】



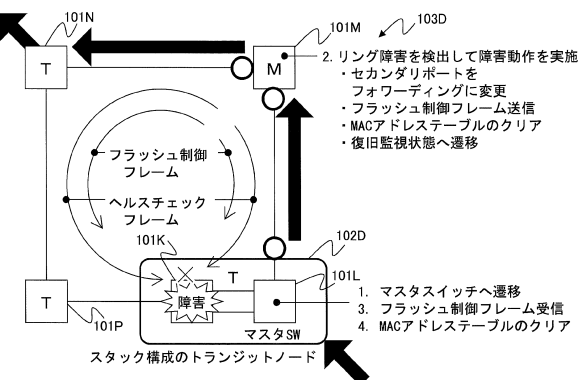
【図 8】



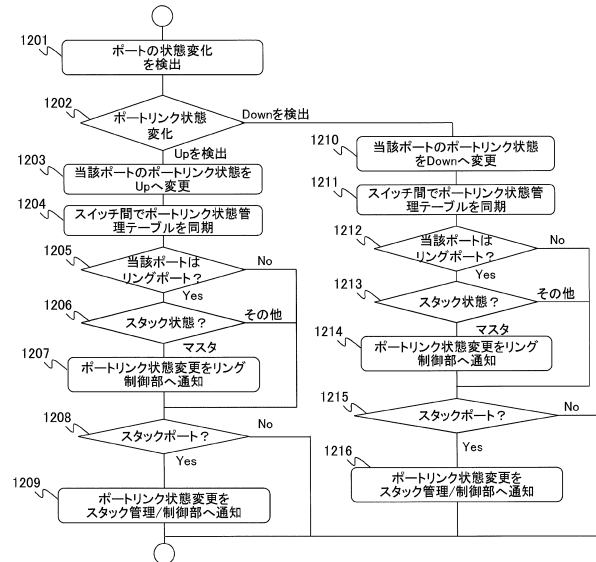
【図 9】



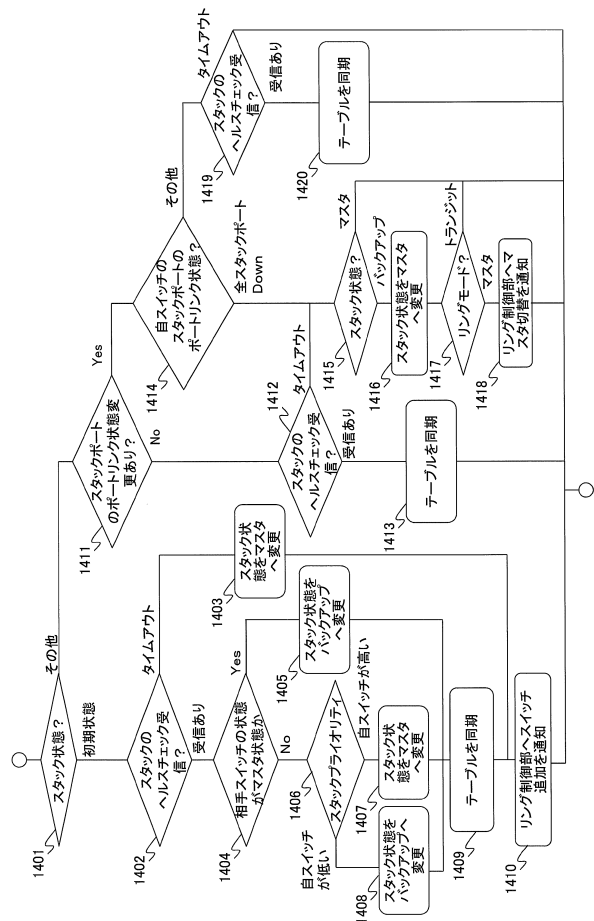
【図 10】



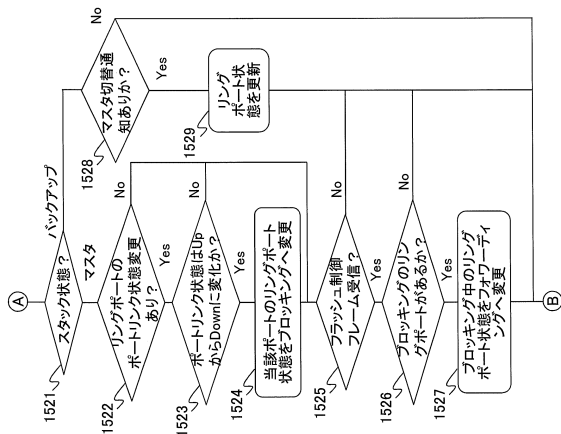
【 図 1 2 】



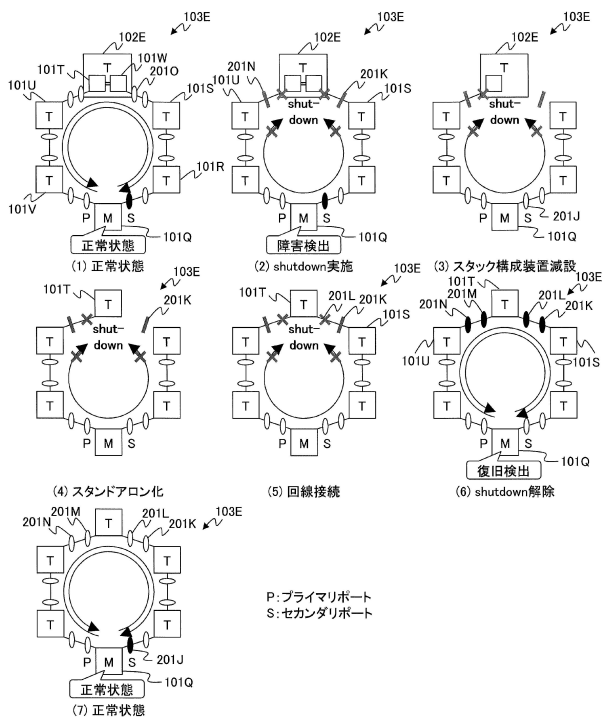
【 図 1 4 】



【 ㊦ 1 5 B 】



【 図 1 7 】



フロントページの続き

(56)参考文献 特開2013-098839(JP,A)
特開2005-130049(JP,A)
特開2008-136013(JP,A)
特開2012-138902(JP,A)
国際公開第2010/120556(WO,A1)
米国特許出願公開第2003/0072259(US,A1)

(58)調査した分野(Int.Cl., DB名)
H04L12/42-12/437