

1.2.2 研究開発項目①-1「光電ハイブリッドスイッチ制御技術」

【成果の概要】

光電ハイブリッドスイッチシステムプロジェクトにおいて、本研究開発項目では光電ハイブリッドスイッチシステムを実現するシステム制御技術の開発を進めてきた。本章ではシステムの特徴である高速伝送経路切り替えを実現するために課題となる、伝送経路切り替えによって発生する光信号断時においても Ethernet Switch のリンクアップを維持する手法について述べる。また、システムレベルでの経路切り替えを行うには本研究で開発を行った光コアスイッチ等の装置に加えて、システムに接続されるサーバ等のネットワーク機器を含めた多くの機器を連携させて制御する必要がある。本章では、その解決策として OpenFlow 機能と EtherCAT を用いた場合のシステム制御シーケンスについて述べる。さらに、物理的な経路切り替えが完了してから IP パケットレベルでの疎通が可能となるまでに要する時間の要因とその対処法について述べる。また、この対処法を用いて経路切り替えを行った場合に IP パケットロスが生じないことの実証実験については、1.2.9 研究開発項目⑤「光電ハイブリッドスイッチシステム実証」にて述べる。

【最終目標】

研究開発項目⑤の連携実験の結果を踏まえてコントローラ制御のシーケンスなどの改良を行い、光電ハイブリッドスイッチ制御技術を確立させる。

【成果】

＜システム実現における課題と対応＞

1. Ethernet Switch のリンクダウンについて

本システム（光電ハイブリッドスイッチシステム）において、サーバ間でパケットの送受信を行う場合のサーバ間で IP パケットを送受信する場合の伝送経路は図 1.2.2-1 のように 2 つ存在する。すなわち、2 台以上の Ethernet Switch が縦続接続された“電気経路”と、光コアスイッチを介して 2 台の Ethernet Switch が接続された“光経路”である。

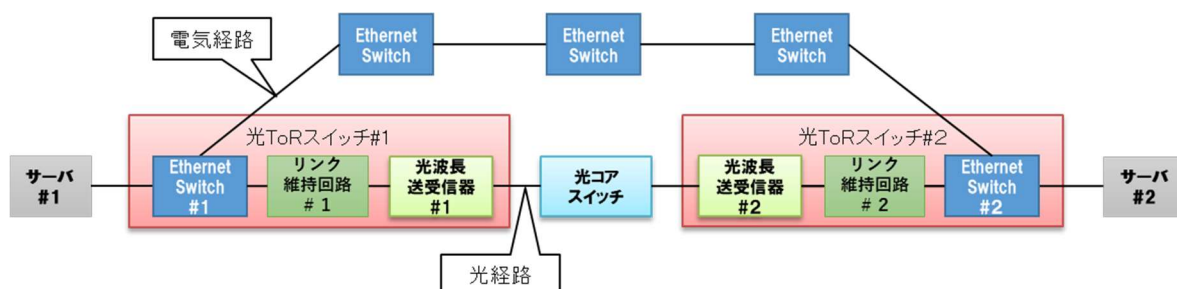


図 1.2.2-1 光電ハイブリッドスイッチシステム伝送経路例

本システムでは電気経路での通信から光経路での通信へ切り替え（オフロード）を行うが、その際に光コアスイッチの切り替えを行って物理的に光経路を新たに形成することとなる。しかしながら、現行の Ethernet Switch を用いると物理的な光経路が形成されても直ちに IP パケットは疎通せず、IP パケットの疎通が可能となるまで数秒を要することを確認している。この原因は、あらたに光経路を形成する時点で光経路に対応した Ethernet Switch の通信ポートがリンクダウン状態となっているため

である。このリンクダウンは、光コアスイッチを切り替えて光経路どうしを切り替えることによって発生する。この原因は、OSI (Open Systems Interconnection)参照モデルでいうところの物理層において信号断、あるいは信号同期に問題が生じると Ethernet 通信に用いられるデータリンク層においてデータ通信が停止するためである。

したがって、高速で経路を切り替えるには、光経路が形成されていない場合、また光コアスイッチを切り替えて光伝送経路が一時的に切れる際にも Ethernet Switch がリンクアップを維持できるよう物理層での信号同期を常に維持しておくことが必須である。

本研究では、図 1.2.2-1 に示すように光伝送経路の切り替え時に Ethernet Switch のリンクダウンを生じさせないため、Ethernet Switch と光波長送受信器の間に「リンク維持回路」を配置することとした。本回路は、光波長送受信器からの入力信号に異常を感知した場合に Ethernet Switch に対してアイドル信号を送出し、物理層において信号の同期をビット単位で常に維持する機能を有するもので FPGA (Field Programmable Gate Array) によって実現している。また、本回路は入力した信号状態をモニタする機能を有している。モニタ可能項目については、後に表 1.2.9-5 において述べる。

2. 連携制御の実現

本システムでは、2 台以上の Ethernet Switch が縦続接続された伝送経路である“電気経路”から、光コアスイッチを介して 2 台の Ethernet Switch が接続された伝送経路である“光経路”へ切り替えを行う際には、複数の装置を連携して制御することが必要である。本研究では、連携制御を実現するため、制御コマンドの通信プロトコルとして EtherCAT を用いることとした。さらに、制御対象装置の同期制御の精度を高めるため、各装置は Distributed Clocks (DC)モードにて動作させることとした。

制御対象となる装置とその制御シーケンスについては次の<システム制御シーケンス>において述べる。

<システム制御シーケンス>

本研究項目では、光 Top of Rack (ToR)スイッチ内の Ethernet Switch における電気経路と光経路の切り替え手段として OpenFlow 機能を用いている。具体的には、Ethernet Switch のあるポートから入力した IP パケットに対して、その IP パケット内の特定のビットに定められている Type of Service (TOS) の値によって出力ポートが定まるよう入出力ポートの関係を OpenFlow Table によってあらかじめ定めしておく。こうすることで、サーバ等から入力される IP パケットはその TOS 値により Ethernet Switch によって異なる出力ポート、すなわち電気経路と光経路への振り分けが自律的に行われる。ただし、OpenFlow 機能を用いるには通信に用いる Ethernet Switch のポートがリンクアップしていることが前提であり、そのためにもリンク維持回路を用いることが必須である。

OpenFlow 機能を用いた経路設定の方法として、Ethernet Switch 上の OpenFlow Table で入出力ポート間の経路を定め、外部のコントローラからその経路を指示して入出力経路を切り替える方法がある。しかしながら、この方法では、コントローラ側のソフトウェア処理に起因する経路指定コマンドの送出時間のばらつきや、制御コマンドを伝達するネットワーク構成によって、Ethernet Switch 内の経路切り替え時間と光経路切り替えタイミングを高精度で合わせることが困難であるため、高速経路切り替えを行うには不向きである。したがって後に述べるシステム検証では、経路切り替えには上に述べた IP パケット送信側での TOS 値変更による経路切り替えを用いている。

ここで、光電ハイブリッドスイッチシステムの構成例を図 1.2.2-2 に示す。図 1.2.2-2 は、3 台の光

Top of Rack (ToR)スイッチと光コアスイッチから構成される場合を示している。光 ToR スwitchは 100GbE Ethernet Switch、100GbE リンク維持回路が実装されたリンク維持送受信ボードおよび光波長送受信器から構成され、光波長送受信器には、光波長送受信ボードとバースト多値プロセッサとが内蔵されている。また、Ethernet Switch にはサーバおよび映像送受信器が接続されている。

さらに、光コアスイッチと光波長送受信ボードは EtherCAT スレーブ機能を有しており、EtherCAT を用いてコントローラより連携制御される。

加えて、光波長送受信ボードは波長可変光源を有し、コントローラから波長切り替え指示を受けて波長を変化させるとともに、バースト多値プロセッサに対して再同期トリガ信号を発出する。

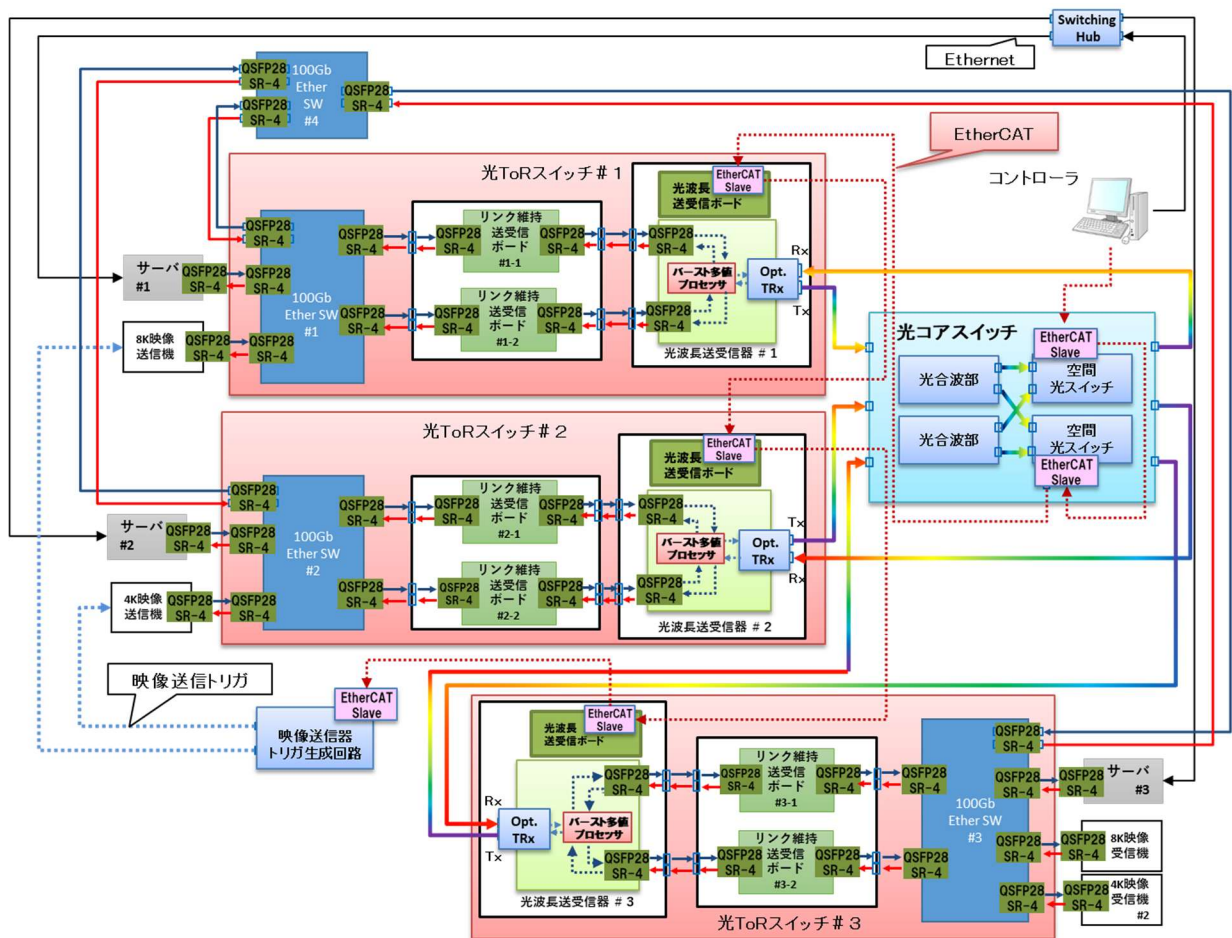


図 1.2.2-2 光電ハイブリッドスイッチシステムの一例

次に、図 1.2.2-2 の制御系のみを抽出した制御系構成とコントローラ機能構成を図 1.2.2-3 に示す。本システムの制御を行うコントローラはシステムコントローラ機能部と EtherCAT マスタ機能部とから構成される。図 1.2.2-3 に示すようにシステム制御は Ethernet と EtherCAT の両方を用いて行われる。システムコントローラ機能部は以下の機能を有する。

- ・ 光コアスイッチの経路設定、光波長送受信ボードの波長設定内容を EtherCAT マスタ機能部へ送信する。
- ・ サーバ等のネットワークノードに対し Ethernet を用いて TOS 値設定指示およびパケット送信・停止指示を行う。

また、EtherCAT マスタ機能部は EtherCAT スレーブである光コアスイッチ、光波長送受信ボードに対して EtherCAT を用いて設定コマンドを送出する。

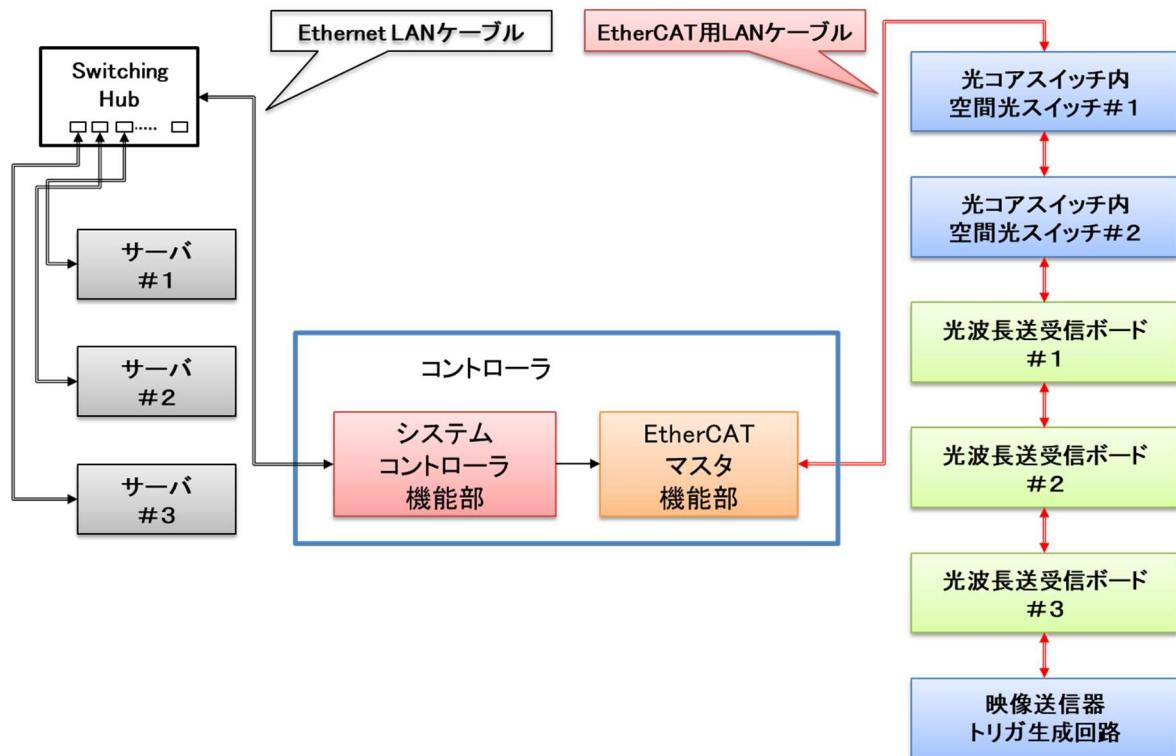


図 1.2.2-3 システム制御系構成例とコントローラ機能構成

次に、装置制御シーケンスの詳細について示す。ただし、経路切り替えリクエストがサーバ等からコントローラに上がり、コントローラが EtherCAT マスタに設定コマンドを送信するまでの詳細アルゴリズムについては、1.2.3 項において述べられているため、本項目では主にコントローラが、サーバ、光 ToR スイッチ内の光波長送受信ボードとバースト多値プロセッサおよび、光コアスイッチを制御するシーケンスについて述べる。

光電ハイブリッドスイッチシステムにおいて、対向するサーバ等の装置間で IP パケットを送受信する伝送経路を切り替える場合、以下の 3 通りの場合が考えられる。

1. 電気経路⇒光経路
2. 光経路⇒電気経路
3. 光経路⇒光経路

以下に各場合の制御シーケンスについて述べる。

1. 電気経路⇒光経路

2 つのサーバ # 1 と # 2 との間で IP パケットの送受信に電気経路を用いて行っている際に、光経路への切り替えを行う場合の制御手順を図 1.2.2-4 に示す。

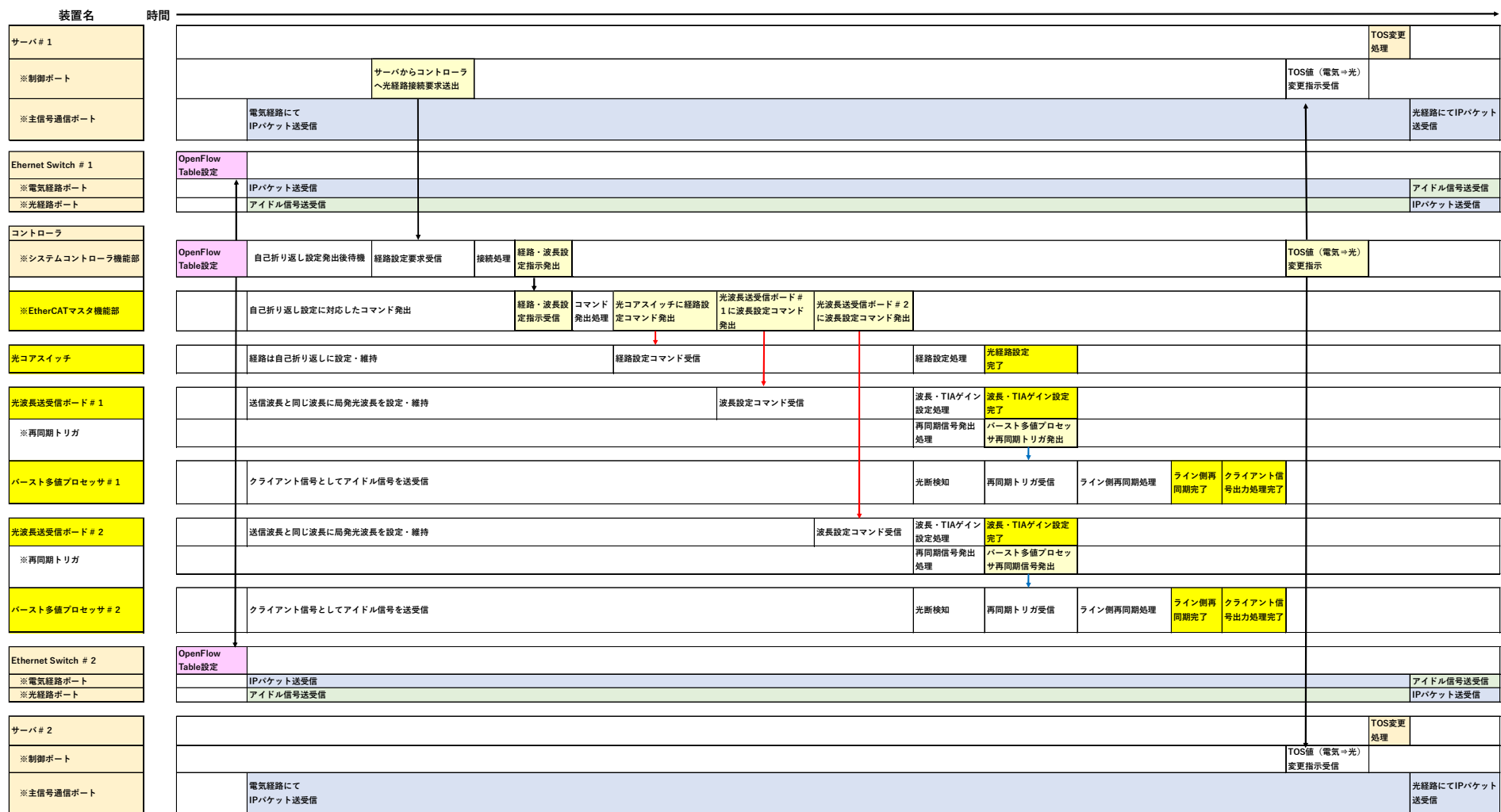


図 1.2.2-4 電気経路から光経路へ切り替えを行う際の制御シーケンス

図 1.2.2-4 では、光コアスイッチによって光伝送経路が形成する前の段階においても Ethernet Switch の通信ポートのリンクアップが維持されていることを前提としている。

また、当初サーバが電気経路を用いて IP パケットの送受信を行っているとき、光コアスイッチと光波長送受信は自己折り返しを行うよう設定しておく。これは、光経路設定前後での光波長送受信器への入力光信号パワーの変動幅をできるだけ抑え、光波長送受信器内集積型受信器の TIA ゲインの変動幅を低減し、制御の安定度をより高めるためである。TIA のゲイン制御の詳細については、1.2.5 にて述べる。

次にシステム内の各装置の動作状態と制御シーケンスについて説明する。

- ① コントローラまたは設定用外部 PC 等を用いて、光 ToR スイッチ内の各 Ethernet Switch の OpenFlow Table をあらかじめ設定する。入力 IP パケットの TOS 値の違いにより、上位に位置する Ethernet Switch への接続ポートまたは光波長送受信器への接続ポートに IP パケットが振り分けられるよう設定を行う。このとき、光経路はまだ設定されておらず、サーバ # 1 と # 2 は電気経路を用いて通信を行っている。光波長送受信器に接続された Ethernet Switch ポートはアイドル信号を送受信している。
- ② サーバ # 1 からコントローラに対して、光経路への接続要求が発出される。
- ③ コントローラは、システム内のネットワーク機器の通信状態を考慮し、光経路での接続についての可否判断と可能である場合のルートを決め、そのルートを実現するための光コアスイッチの経路設定情報、光波長送受信器内の可変波長光源の設定波長を決定する。経路設定および波長設定に関わる詳細アルゴリズムについては 1.2.3 で述べる。
- ④ コントローラ内のシステムコントローラ機能部より、EtherCAT マスタ機能部に対して、光コアスイッチおよび光波長送受信器の制御情報が送信される。
- ⑤ EtherCAT マスタ機能部より、光コアスイッチ、光波長送受信器 # 1、光波長送受信器 # 2 に対して EtherCAT を用いて制御コマンドが発出される。
- ⑥ 光コアスイッチにより物理的な光経路が接続される。
- ⑦ ⑥と同時に光波長送受信ボード上の局発光の波長を切り替えと光波長送受信器内の TIA ゲインの設定が行われる。
- ⑧ 光波長送受信器内の光波長送受信ボードより、バースト多値プロセッサへ再同期トリガ信号を発出。光波長送受信ボードの詳細仕様については 1.2.5 で述べる。
- ⑨ 光波長送受信器においてクライアント信号処理完了後、TOS 値を変更した IP パケットをサーバより送信。

ここで注意すべきは、送信側光 ToR スイッチより光信号を受信し、受信側光 ToR スイッチ内のバースト多値プロセッサでライン側同期確立後、クライアント信号である Ethernet 信号が光波長送受信器から出力されるまで一定時間を要することである。この原因は、バースト多値プロセッサ内でのクライアント信号に関わる処理と光波長送受信器で用いられている 100GbE プラガブルトランシーバ（市販品）の信号断からの回復時間である。さらに IP パケットを送受信するサーバ上で動作するアプリケーションソフトウェアの処理時間によって、TOS が変更された IP パケットの送受信が可能になるまで一定時間を要する。システムとして IP パケットレベルでのスループットを高めるにはこれらの時間を考慮して TOS 値の変更のタイミングを定めることが必要である。

2. 光⇒電気経路

次に伝送経路を光経路から電気経路へ切り替える際の制御シーケンスを図 1.2.2-5 に示す。

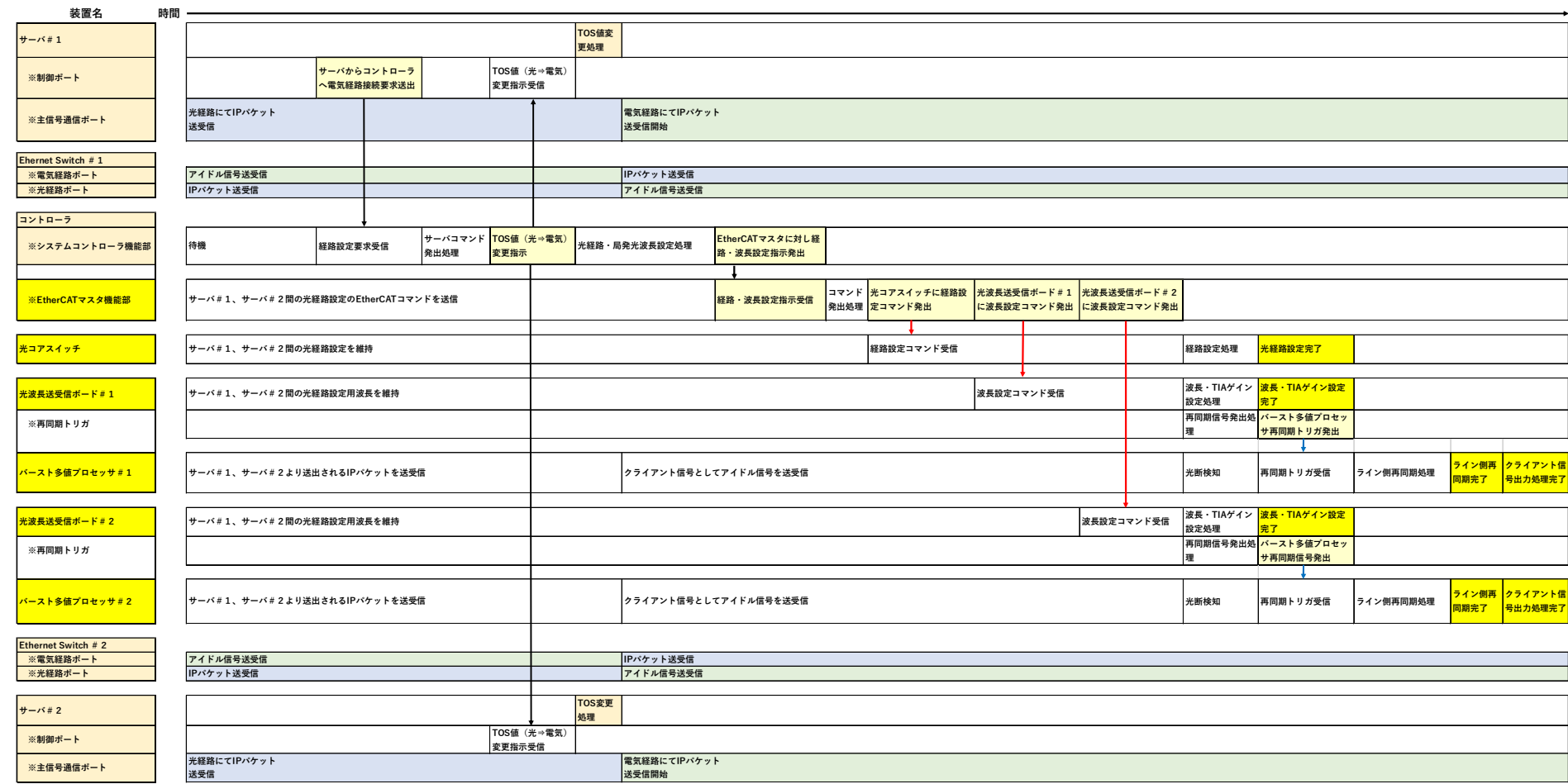


図 1.2.2-5 光経路から電気経路へ切り替えを行う際の制御シーケンス

光経路から電気経路への切り替えを行う場合には、IPパケットの伝送経路が光経路から電気経路への変更が完了し、電気経路でのIPパケットの疎通が可能となった後に光コアスイッチおよび光波長送受信器の制御を行い、光経路を解放することが重要である。光コアスイッチ、光波長送受信器はそれぞれ100μs以下で物理的な切り替えが高速で可能であるが、サーバ上のソフトウェアがTOS値変更によくの処理時間を要する場合、あるいは、コントローラからのコマンドがサーバ等へ到達するのに時間を要するネットワーク構成である場合、IPパケットの送受信を行っている途中で光伝送路が切断されることになるため、これを避ける必要がある。

以下にシステムコントローラ機能部からの指示を起点とした場合の制御シーケンスの詳細を示す。

- ① コントローラよりサーバ等のIPパケット送受信装置に対して電気経路に対応したTOS値に変更を行うよう指示
- ② IPパケットのTOS値を電気経路に対応した値に変更し、IPパケットを電気経路経由に変更
- ③ IPパケットの電気経路での疎通完了
- ④ コントローラ内のシステムコントローラ機能部より、EtherCATマスタ機能部に対して、光コアスイッチおよび光波長送受信器の制御情報が送信される。
- ⑤ EtherCATマスタ機能部より、光コアスイッチ、光波長送受信器#1、光波長送受信器#2に対してEtherCATを用いて制御コマンドが発出される。
- ⑥ 光コアスイッチにより物理的な光経路を折り返し経路に設定。
- ⑦ ⑥と同時に光波長送受信器内の局発光の波長を送信光と同じ波長に切り替え、TIAゲインの設定が行われる。
- ⑧ 光波長送受信器内の光波長送受信ボードより、バースト多値プロセッサへ再同期トリガ信号を発出し自己折り返しの光経路形成。

3. 光経路⇄電気経路

本システムにおいて光経路間で切り替えを行う場合の伝送経路の概略を図1.2.2-6に示す。

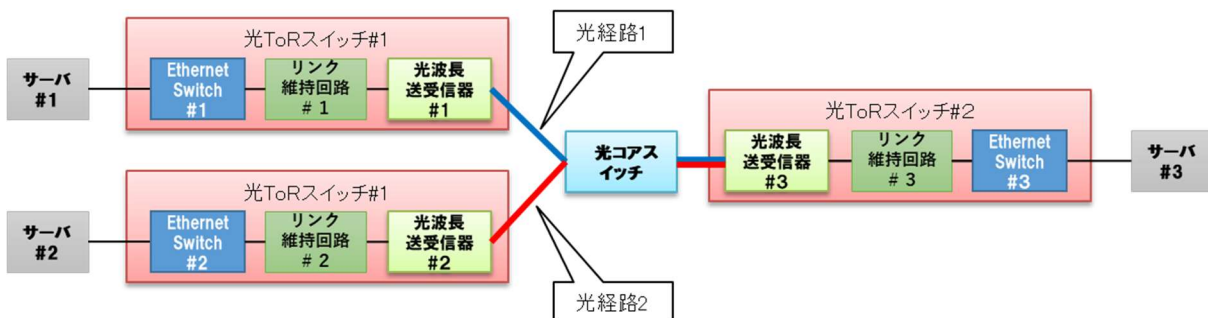
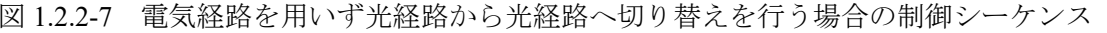


図 1.2.2-6 光電ハイブリッドスイッチシステム光伝送経路概略

図1.2.2-6において、サーバ#1とサーバ#3を接続する光伝送経路を光経路1、サーバ#2とサーバ#3を接続する光伝送経路を光経路2とする。光経路1を用いてサーバ#1とサーバ#3とが通信を行っている状態から、伝送経路2を用いてサーバ#2とサーバ#3が通信を行う状態へ切り替えを行う場合の制御シーケンスを図1.2.2-7に示す。



光経路 1 から光経路 2 への切り替えを行う場合の制御シーケンスを以下に示す。

- ① サーバ # 2 よりコントローラに対して、光経路を用いてサーバ # 3 と接続を行いたいとのリクエストが上がる。
- ② 光経路の切り替え動作に入る前に、コントローラは各サーバに対して通信を完了するよう指示を行う。
- ③ 通信の停止後、コントローラはシステムコントローラ機能部より、EtherCAT マスタ機能部に対して、光コアスイッチおよび光波長送受信器の制御情報を送信する。
- ④ EtherCAT マスタ機能部より、光コアスイッチ、光波長送受信器 # 1、光波長送受信器 # 2、光波長送受信器 # 3 に対して EtherCAT を用いて制御コマンドが発出される。
- ⑤ 光コアスイッチにより物理的な光経路が切り替わり、光波長送受信器 # 1 は送受折り返し経路に設定される。
- ⑥ ⑤と同時に、光波長送受信器 # 1 の局発光波長を送信光と同じ波長に切り替え、光波長送受信器 # 2、# 3 の局発光波長をそれぞれ光波長送受信器 # 3、# 2 の送信波長に切り替える。さらに設定波長に対応した TIA ゲインへの設定を行う。
- ⑦ 光波長送受信内の光波長送受信ボードより、バースト多値プロセッサへ再同期トリガ信号を発出する。
- ⑧ バースト多値プロセッサのライン側受信部の同期が確立され、さらにバースト多値プロセッサのクライアント信号処理が完了
- ⑨ 光波長送受信器での IP パケットの送受信が可能となる。

光電ハイブリッドスイッチシステムでは、サーバ間の電気経路は常にリンクアップを維持した状態にある。しかしながら、サーバ # 1 とサーバ # 3 が光経路を用いて通信を行っている状態からサーバ # 2 とサーバ # 3 が光経路を用いて通信を行うよう、電気経路を介さず光経路から光経路へ直接切り替える場合は、サーバ # 3 は連続的に IP パケットを受信できずアイドル信号のみを受信する時間が存在する。その要因は光コアスイッチの経路切り替え時間、光波長送受信器内の波長可変光源の波長切り替え時間およびバースト多値プロセッサのライン側再同期時間に加えて、バースト多値プロセッサがライン同期後、クライアント信号に関わる処理、さらに光波長送受信器内の 100GbE プラガブルトランシーバ（市販品）およびサーバでの処理に時間を要するためであり、物理的な光経路が形成された後、IP パケットの送信を開始するにはこの時間を考慮する必要がある。

<まとめ>

光電ハイブリッドスイッチシステムは、通信経路として Ethernet Switch のみから構成される電気経路と光スイッチを含む光経路とが存在する。本研究項目では、2つの経路を切り替えるための装置制御シーケンスについて述べた。また、電気経路から光経路へ切り替えを行う際には TOS 値変更まで待ち時間を設けること、光経路から電気経路への切り替えを行う際には光経路解放まえに IP パケットをあらかじめ電気経路に移動させておくことが必要であることを述べた。これらシステム制御技術を用いることにより、光電ハイブリッドスイッチシステムに必要な電気経路、光経路の切替をスムーズに実現できる。