(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2014-103674
(P2014-103674A)
(43)公開日　平成26年6月5日(2014.6.5)

(51)Int.Cl.			F I			テーマコード（参考）
H04L	***12/70***	***(2013.01)***	H04L	12/70	Z	5B061
G06F	*13/362*	*(2006.01)*	G06F	13/362	510D	5K030

審査請求 有　請求項の数 12　ＯＬ　（全 63 頁）

(21)出願番号　特願2013-255681 (P2013-255681)
(22)出願日　平成25年12月11日 (2013.12.11)
(62)分割の表示　特願2013-503041 (P2013-503041)の分割
原出願日　平成23年4月8日 (2011.4.8)
(31)優先権主張番号　10003791.0
(32)優先日　平成22年4月8日 (2010.4.8)
(33)優先権主張国　欧州特許庁 (EP)

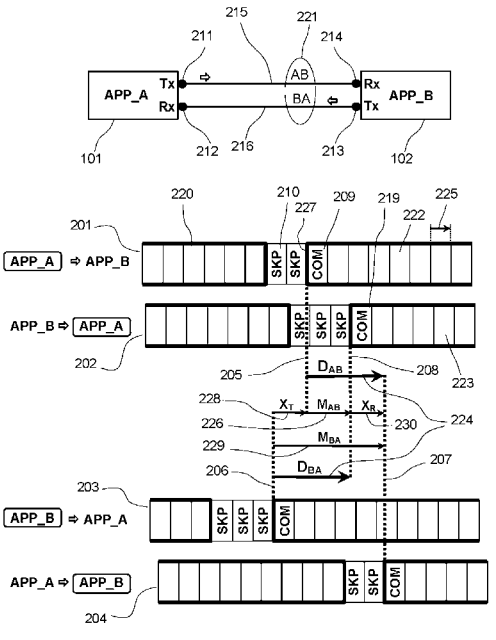
（特許庁注：以下のものは登録商標）
１．ＥＴＨＥＲＮＥＴ
２．イーサネット

(71)出願人　512245698
ヴァダース　イシュトヴァーン
ＶＡＤＡＳＺ，　Ｉｓｔｖａｎ
ドイツ連邦共和国　８１７３９　ミュンヘン　クルトーユルゲン通り　４
Ｃｕｒｄ－Ｊｕｒｇｅｎ－Ｓｔｒ．　４，
８１７３９　Ｍｕｎｉｃｈ　（ＤＥ）
(74)代理人　100127188
弁理士　川守田　光紀
(72)発明者　ヴァダース　イシュトヴァーン
ドイツ連邦共和国　８１７３９　ミュンヘン　クルトーユルゲン通り　４

Ｆターム(参考)　5B061 BB14　GG12
5K030 HB15　KA22　LA15

(54)【発明の名称】集中制御を用いないネットワークにおける、同期したセルロック送信を提供する装置及び方法

(57)【要約】　（修正有）
【課題】集中制御を必要とせずに、ネットワーク構造全体においてセル送信の開始を同時に行う。
【解決手段】少なくとも２つの通信手段と、少なくとも２つのリソースと、アービタとを備える装置１０１，１０２は、２つ以上の他の装置、およびネットワークを形成する実質的に同一のアービタに対して、能動的および受動的な通信を行うことと、活動期として時間区分を識別することと、他の装置の何れかが１活動期内に発行するアクティブ通信を１活動期内に受信することと、発行の活動期内に、受信したアクティブ通信要素を他の装置に転送することと、アクティブ通信要素を用いて、要求メッセージおよびリソースに関する状態メッセージを発信し、他の装置は、前記装置が発信した要求および状態メッセージを、発行の活動期内に受信し発信することと、次の活動期で、アービタの結果から排他的にリソースの状態割り当てを抽出することと、を実行する。
【選択図】図２



【特許請求の範囲】
【請求項１】
少なくとも２つの通信手段（２２１）と、少なくとも２つのリソースと、アービタ（１４０１）とを備える装置（１０１）であって、前記装置は：
同一の通信動作をする２つ以上の他の装置（１０２，１０３）、および前記２つ以上の他の装置（１０２，１０３）と共に前記装置（１０１）のネットワーク（４１０）を形成する実質的に同一のアービタ（１４０１）に対して、前記通信手段（２２１）を介して能動的および受動的な通信を行うことと；
活動期（６０３）として時間区分を識別することと；
前記ネットワーク（４１０）内の２つ以上の他の装置（１０２，１０３）の何れかの装置が１活動期（６０３）内に発行するアクティブ通信の２つの要素（２２２）の何れかを、１活動期（６０３）内に受信することと；
前記発行の活動期（６０３）内に、受信したアクティブ通信要素（２２２）を前記ネットワーク（４１０）内の他の装置（１０２，１０３）に転送することと；
前記アクティブ通信要素（２２２）を用いて、前記通信手段（２２１）経由の要求メッセージおよび前記リソースの現在および／または将来の状態に関して通知する状態メッセージを発信することであって、前記ネットワーク（４１０）内の他の装置（１０２，１０３）は、前記装置（１０１）が発信した要求および状態メッセージの各々を、前記発行の活動期（６０３）内に受信する、前記発信することと；
次の活動期（６０８）で、前記アービタ（１４０１）の結果から排他的に前記リソースの状態割り当てを抽出することと；
を実行するように構成され、
前記リソースは少なくとも２つの状態を有し、前記状態または状態シーケンスは、次の活動期（６０８）に前記アービタ（１４０１）によって決定され；
前記アービタ（１４０１）は：　前記発信された要求および状態メッセージの内容を入力として使用し、次の１または複数の活動期（６０８）で、前記ネットワーク（４１０）内の２つ以上の他の装置（１０２，１０３）と共に、前記装置（１０１）の全リソースの状態または状態シーケンスを計算することと；
前記通信要素（２２２）を介して受信された要求メッセージに従ってペイロード（７０４）を送信するために、１活動期（６０３）内に、前記通信手段（２２１）の直接接続パス（２１５）を割り当てる（９０２，９０４）ことであって、前記要求メッセージは、前記通信要素（２２２）を介して送信元である接続された装置（１０１）から送信先である別の接続された装置（１０２）に、該送信元装置（１０１）を該送信先装置（１０２）に直接接続する前記通信手段（２２１）を通じて送信される、前記割り当てる（９０２，９０４）ことと；
受信された要求メッセージに従って、前記送信元装置（１０１）を前記送信先装置（１０２）に直接接続する通信手段（２２１）を経由する送信に加えてペイロード（７０４）を送信する転送エージェントとしての装置（１０３）を割り当てることであって、該送信元装置（１０１）から該転送エージェント装置（１０３）への通信手段の直接通信パス（９０８）が、１活動期（６０３）内に割り当てられ（９０３，９０６）、かつ、該転送エージェント装置（１０３）から該送信先装置（１０２）への通信手段の直接通信パス（９１１）が、１活動期（６０３）内に割り当てられ（９０５，９０７）、前記割り当ては、前記通信手段の直接通信パス（９０８，９１１）が割り当て可能である場合に制約されている、前記割り当てる（９０５，９０７）ことと；
を実行するように構成され；
前記アービトレーションは各活動期（６０３）で実行される；
装置。
【請求項２】
前記通信手段（２２１）は、
個々のセル・ストリーム（６０１）用の双方向リンク（２２１）を提供するコンジット

として構成され、前記セルは、所定数のシンボル（２２２）を含み、
前記活動期（６０３）は、セル期間（６０３）毎に同期ロックされたシンボル（２２２）を評価する手段を提供し、
前記リソースは、パス（２１５）およびシンボル（２２２）を伝送する他の要素を備え、
前記要求メッセージは、前記要求メッセージは、ペイロード（７０４）送信のためのリクエストであって、前記送信先装置（１０１）毎に前記ペイロード（７０４）送信に必要なセル（６０１）数を特定するリクエストを運び、
前記状態メッセージは各リンク（２２１）の受信性能情報を含み、 前記アービトレーション処理は、要求されたペイロード（７０４）送信を、要求元装置（１０１）と送信先装置（１０２）とを直接接続するパス（２１５）に割り当て、さらに、転送エージェントである装置（１０３）に、要求元（１０１）から転送エージェント（１０３）への、および転送エージェント（１０３）から送信先（１０２）へのそれぞれのパス（９０８，９１１）を共に割り当て、
前記パス（２１５）は、前記リンク（２２１）の指示コンポーネントとして特定される、
請求項１に記載の装置。
【請求項３】
前記装置（１０１）は、
セル（６０１）のシンボル・サブセット（２２２）をペイロード・データとして他の装置（１０２）にセルロックリンク（２２１）を介して送信し、
セル（６０１）のシンボル・サブセット（２２２）をペイロード・データとして他の装置（１０２）からセルロックリンク（２２１）を介して受信するように構成され、
前記装置（１０１）は、セルロックリンク（２２１）における他の未使用パス（１０６，９０８）の組に対し、送信元である第２の装置（１０２）から受信したセル（６０１）のシンボル・サブセット（２２２）を、送信先である第３の装置（１０３）に再送する、転送エージェントの役割を担うように構成され、
前記シンボル・サブセット（２２２）の再送は、前記装置の受信時のセル期間（６０３）またはそれ以降のセル期間（６０８）の何れかの期間内に実行される、
請求項２に記載の装置。
【請求項４】
前記装置は、前記アービタ（１４０１）が要求元装置（１０１）から送信先装置（１０１）への複数の送信ルートを割り当てる際、ペイロード（７０４）・データ・ストリームの配信および再構築に関する規則を適用するように構成され；
前記規則は、セルロックネットワーク（４１０）内で相互接続された装置（１０１）の各々が従い、前記データ・パス（２１５）の所定の優先順位に基づき、または前記データ・パスの複数のレーンに基づいてもよく；
前記装置（１０１）はデータ・セグメントの送信に関連する転送エージェントとして割り当てられ；
前記優先順位が前記装置（１０１）に割り当てられる固有識別番号（１０６）に基づいてもよい、
請求項３に記載の装置。
【請求項５】
前記装置は、
セル（６０１）の中で等距離にある位置において、シンボル位置（８０５）またはシンボル位置のグループからなる専用サブセットを割り当て、
前記同一のセル期間（６０３）内で、前記シンボル位置の専用サブセットにおける特定の要素を介して、また前記シンボル位置の専用サブセットにおける別の要素を介して、シンボル（２２２）またはシンボル（２２２）のグループを再送する、
ように構成される、請求項１から４のいずれか１項に記載の装置。

【請求項６】
ベーシック・シンボル・レート（５０２，５０３）の倍数で動作する手段、前記リンク（２２１）を含む複数の並列レーンを使用する手段、または前記両手段の組み合わせによって、セル期間（６０３）毎にそのリンク（２２１）のサブセットを介して複数のセル（６０１）を送受信するように構成され；
前記ベーシック・シンボル・レートのシンボル期間（２２５）毎に対応するより多くのシンボル（２２２）を送信する場合、前記ベーシック・シンボル期間（２２５）内の各シンボル位置が、複数のインターリーブセル（６０１）の１つに対応するシンボル位置（２２２）に割り当てられる、
請求項１から５のいずれか１項に記載の装置。
【請求項７】
前記装置（１０１）は、前記リンク（２２１）のサブセットに広帯域幅を有するネットワーク・トポロジーをサポートするために強化されたアービトレーション処理を実装するよう構成され、前記アービトレーション処理は、次の手順でリソースに基礎伝送を完全に割り当てるよう構成され、前記手順は：
・　要求元から送信先装置への直接パスを用いる伝送を割り当てること；
・　高帯域幅リンクのみを利用する転送エージェント装置を用いる伝送を割り当てること；
・　特定の高帯域幅リンクを利用する転送エージェント装置を用いる伝送を割り当てること；
・　残りの全伝送を割り当てること；
を含む、請求項１から６のいずれか１項に記載の装置。
【請求項８】
前記装置（１０１）は、要素の循環的順序セットと見做されるネットワーク装置（１０１）用のアービトレーション処理を適用するように構成され、
前記アービトレーション方法が適用される場合に、特定の循環的順序が利用され、　前記要素は、前記特定の循環的順位における他の要素の相対位置を把握し、
前記要素の各々は、前記同一のアービトレーション処理を実行し、かつ、要件，性能および他の機能に関して前記要素と同一のデータセット、および要素間相互関係に基づいて、前記アービトレーション処理を実行し、
前記アービトレーション方法は実行サイクルのシーケンスで行われ、ただしリソースおよびその他の物が、同時実行ステップにおいて割り当てられ、前記要素の各々が前記同時実行ステップの１つに対する開始ポイントとして利用され、
前記同時実行ステップ後、更新データセットが次の実行ステップで利用され、各実行ステップが前記アービトレーション方法の結果のサブセットを生成し、
ここで、前記実行ステップは、前記アービトレーション方法のための結果の要素を生成し、
前記アービトレーション方法は、前記各開始ポイントから循環的に後のポイントで開始し、前記各開始ポイントから循環的に前のポイントで終了する要素の前記特定の循環的順位のシーケンス内における前記要素の有限集合を、前記順序セットから選択された要素と比べて考慮し、ここで考慮される要素の各開始ポイントと循環的に前のポイントは除外される、
請求項１に記載の装置。
【請求項９】
前記装置は、
パケット・プロトコル、ストレージ・インターフェース・プロトコル、またはその他の高水準プロトコルに関して、セルロックネットワーク（４１０）を介してペイロード（７０４）として送信または再送されるシンボル（２２２）を利用し、
さらに所定のシンボル位置に割り当てられたシンボルまたはシンボルのシーケンスを介して、前記プロトコルを識別するように構成される、

請求項１から８のいずれか１項に記載の装置。
【請求項１０】
３つ以上のコンピュータ装置（１０１）を備えるネットワーク構造でコンジットを介してデータ送信する方法であって、前記コンジットは、両方向に流れる個々のシンボル・ストリームのために双方向リンク（２２１）を提供し、前記方法は：
同一の通信動作および実質的に同一のアービタ（１４０１）を備える前記コンピュータ装置（１０１）を提供することと；
活動期（６０３）として時間区分を識別することと；
前記ネットワーク（４１０）内のコンピュータ装置（１０１）が１活動期（６０３）内に発行するアクティブ通信の２つの要素（２２２）の何れかを、１活動期（６０３）内に受信することと；
前記発行の活動期（６０３）内に、受信した通信要素（２２２）を前記ネットワーク（４１０）内の他の装置（１０１）に転送することと；
次の活動期（６０８）で、状態または状態シーケンスを決定することと；
要求メッセージおよび前記リソースの現在および／または将来の状態に関して通知する状態メッセージを発信することであって、前記ネットワーク（４１０）内のコンピュータ装装置（１０１）は、前記装置（１０１）が発信した要求および状態メッセージの各々を、前記発行の活動期（６０３）内に受信する、前記発信することと；
前記発信された要求および状態メッセージの内容を入力として使用することと；
次の１または複数の活動期（６０８）で、前記ネットワーク（４１０）内の装置（１０１）の全リソースの状態または状態シーケンスを計算することと；
送信元装置（１０１）から送信先装置（１０２）の直接接続リンク（２２１）を通じてペイロード（７０４）を送信するために、１活動期（６０３）内に、前記接続リンク（２２１）の直接パス（２１５）を割り当てることと；
前記送信元装置（１０１）から前記送信先装置（１０２）への直接通信リンク（２２１）に加えて、ペイロード（７０４）を送信する前記ネットワーク（４１０）内の転送エージェント装置（１０３）を割り当てることであって、該送信元装置（１０１）から該転送エージェント装置（１０３）への通信リンク（２２１）の直接パス（９０８）が、１活動期（６０３）内に割り当てられ、かつ、該転送エージェント装置（１０３）から該送信先装置（１０２）への通信リンクの直接パス（９１１）が、１活動期（６０３）内に割り当てられ、前記割り当ては、前記直接通信パス（９０８，９１１）がそれぞれ対応する通信期間（６０３）内に利用可能である場合に制約されている、前記割り当てることと；
次の活動期（６０８）で、前記アービタ（603）の結果から排他的に前記リソースの状態割り当てを抽出することであって、前記アービトレーションは各活動期（６０３）で実行される、前記抽出することと；
を含む、方法。
【請求項１１】
コンピュータ装置で実行されることにより、請求項１０に記載のステップを実行するように構成されるコンピュータプログラムコード手段を備える、コンピュータ・プログラム。
【請求項１２】
請求項１に記載の装置のハードウェア記述コードであって、ソースコードでもあり、合成結果が対象製造技術に使用される、および／または前記装置を生成するプログラマブルロジックデバイスをそれぞれ構成するコードストリームとして使用される、ハードウェア記述コード。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、複数のコンピュータ装置間で、セルロックデータ送信を実現および維持するための、コンピュータ装置および方法に関する。これら複数のコンピュータ装置は、全二

重データ伝送リンクと、動的なマルチパス・ルーティングを提供するセルベースのネットワーク・レイヤとにより、フルメッシュで相互接続可能である。本発明の装置および方法によって集中制御が不要となる。

【発明の背景】

【０００２】

フォーマットされたデータの伝送はセル送信およびパケット送信に分類することができる。両方の送信転送において、連続する複数のシンボルは、データ伝送媒体、いわゆるチャネルを通って送信される。シンボルは送信プロトコルの最小単位であり、複数のシンボルの有限集合が利用可能である。チャネルは、送信機から受信機へ逐次的に複数のシンボルを転送しうるコンジットとして実装されうる。チャネルは半二重であっても全二重であってもよい。半二重チャネルは単一方向の送信機能を、全二重チャネルは双方向の送信機能をそれぞれ可能にする。

【０００３】

セル送信とは、ある所定の数のシンボルを「セル」と呼ばれる要素として送信することを特徴とする。上位レベルのプロトコルは、あるセル内に送信されたシンボルを評価するよう要求される。パケット送信では、上位レベルのプロトコルに関連付けられる、可変長シンボルを含む要素を使用するが、この要素が「パケット」と呼ばれる。アプリケーションが送信しようとするデータは「ペイロード・データ」と呼ばれる。ペイロード・データの有用なセクションは、複数のセルやパケットを要求できる。何れの場合でも、ネットワーク全体に亘って送信されるペイロード・データのルーティングを制御するために、プロトコルが必要になる。パケット送信用キャリアとしてセル送信アーキテクチャを適用させるために、中間プロトコル・レイヤも使用できる。伝送リンクの電子的要素はしばしば「信号」と呼ばれる。

【０００４】

セル送信はいわゆる帯域外周波信号方式（out-band signaling）に依存することが多く、その場合、ペイロード・データは物理リンクの信号の１つのサブセット内に送信され、セル構成情報およびルーティング情報は物理リンクの他の信号を介して送信される。典型的なケースにおいて、シンボル・クロックおよび場合によってはセルの開始情報は、データ送信自体に利用されるもの以外のコンジットを介して中央リソースから提供される。多くのアプリケーションで高速差動信号方式が使用されるため、帯域外周波信号方式は旧式であると見做される。帯域外周波信号方式は関連する電子部品または電子モジュールに接続するために必要なコネクタの数が極めて多くなることがある。

【０００５】

セル送信には、セルのストレージが必要とするバッファが常に同じサイズである、という基本的な利点である。一方、ペイロード・データでは、多くの場合セルのサイズに適合せず、帯域幅の一部が無駄になる可能性がある。

【０００６】

セルの所定サイズ、つまりセルの「長さ」は、それぞれのセル送信アーキテクチャによって数バイトから数キロバイトまで様々である。セル送信はしばしば、ネットワーク参加者間の同期相互接続に基づいている。一方、多くのパケット送信プロトコルは接続しているモジュールの同期を必要としない。

【０００７】

セルベースおよびパケットベースのネットワークは、複数のクライアントが直接接続される代わりに、スイッチといわれる中央サービスに接続されるように実装されることが多い。スイッチはセルやパケットのフォームに含まれるデータを一旦受信し、それをターゲット・クライアントに向けて再送する。場合によっては、送信の同期やアービトレーションが中央リソースから提供される。

【０００８】

提供されるのがセル送信ネットワークであろうがパケット送信ネットワークであろうが、スイッチによって、接続可能なクライアント数のほかに、セル長またはパケット長がそ

れぞれ本質的に制限される。スイッチの帯域幅を増加する技術やサポートされるデータ伝送リンク数には常に限界がある。コンジットのデータ伝送性能をさらに高い水準に引き上げる努力もまた、現在行われている。現在の技術水準では、並列スイッチがネットワークパフォーマンスを向上させる最新の方法であり、最も高コストな方法でもある。代替方法として、ネットワーク参加者の各々がスイッチを備え、専用スイッチ装置を不要とすることである。スイッチを備えたネットワーク参加者は、ダイレクト・インターフェース・リンクを多くの参加者に提供し、究極的には各ネットワーク参加者が他のネットワーク参加者と直接のコネクション・リンクを持ついわゆるフルメッシュ型トポロジーを利用する。中央スイッチベースの、およびフルメッシュベースの両方のソリューションは実用性に限界がある。フルメッシュ型ネットワークは通常、ネットワーク参加者数が最大で16の場合に使用される。

【０００９】

フルメッシュ型ネットワークでは、ネットワーク参加者の全てのペアに直接接続が用いられる。各ネットワーク参加者は他の全てのネットワーク参加者およびローカル接続構造へのリンクを提供するスイッチ機能を備えている。ある理想的な実装では、データ・スイッチング・サービスを提供する。それによりネットワーク参加者二者間で多重経路を経由し、データ送信を同時に行える。ネットワーク参加者のペアの間で要求される帯域幅は非常に多種多様でありうるので、高帯域幅のデータ送信は、搭載されてリンクを介する送信容量を利用することで実現してもよい。スイッチをネットワーク接続された各ユニットに追加する確立された可能性は、ハードウェアの観点からは簡単な課題のように見える。しかし、ソフトウェアの観点では、動的なパス変更または複数パスの並存を介してデータ・ストリームを分配することは非常に複雑なことである。そのため、最新のネットワークソリューションにおいては中央スイッチが好まれ、またスイッチ用にできるだけ高い帯域幅が要求される。

【００１０】

ダイナミック・データ伝送パス割り当てを伴うネットワークにとって、一定長のセルおよび可変長でないパケットが使用されることは基本的利点である。ネットワーク内で一般的なサイズのセルを使用する大規模ネットワークにおけるセル送信構造にとって、セル送信周期の相対タイミングは重要な側面である。最も取り扱い易い構造は、完全に同期化されたものである。しかし、それは集中クロック制御が利用される場合にのみ利用可能であるが、一方で、上述の通り、集中制御および帯域外信号がわずかだが不利益を被ることになる。

【００１１】

非常に大規模なネットワークを集中クロック制御することはできないので、ジッタ／ワンダの影響を受けることになる。わずかなクロック・スピードの分散が原因で、セルのオフセットが許容制限を超過し、完全なセルがドロップされることがある。しかしながら、それでもこの構造は依然として有用である。

【００１２】

システムレベルのネットワークについては、セル送信期間はネットワーク全体において調整されロックされる必要がある。このことはセルの内容を転送するための必要条件と思われる。

【００１３】

小規模ネットワークは多くの場合ラックで実装され、相互接続はバックプレーンで提供される。同期セル送信の技術は存在しているが、最新技術の実装においては、EthernetやInfiniBand、Serial RapidIO、Serial Attached Small Computer System Interface(SAS)などのパケット送信が好まれる。

【００１４】

セル送信をベースとしたネットワークにとって、同期を実装することの利点は計り知れない。しかしそのような実装には課題が存在する。同期のソースはある特定のクロック・モジュールである。クロック・モジュールはクロック信号を全てのネットワーク参加者に

向けて送信する。信頼性の高いシステムはクロック供給において二重冗長性を必要とするが、冗長化可能なクロックの使用・供給により、従来の実装の一部を複雑なものにする。

【００１５】

既存のパケット送信技術では、パケット転送チェーンのオーバーフローまたはアンダーフロー状態を確実に回避できるようにするために、多数の制御されたパケット間SKIPシンボルが挿入される。

【００１６】

PICMG 3.0(r) AdvancedTCA(r)仕様においてはバックプレーン用のフルメッシュ相互接続が定義されている。しかし、既存のプロトコルを用いる場合においては、過度に帯域幅の高いこの相互接続アーキテクチャの性能を利用することは複雑であり、また非常にコストがかかる。フルメッシュ相互接続の利点の１つは、セントラル・スイッチ・リソースによって占有されている２つのスロットがメッシュ対応ボードと呼ばれるいかなるタイプにも使用可能である、ということである。

【発明の概要】

【００１７】

本発明の目的は、集中制御を必要とせずに、ネットワーク構造全体においてセル送信の開始を同時に行うことを提供することである。

【００１８】

本発明の別の目的は、同期ネットワークにおいて要求および性能メッセージを発信する装置におけるリソースのために反復的なアービトレーション処理を提供することである。

【００１９】

本発明の構成は、コンピュータ装置間でセルベースの同期ネットワーク通信を実現する。この同期ネットワーク通信は、フルメッシュ・ネットワークにおいて複数のパスを動的に利用し、また、集中制御を必要としない。第１の側面は、ネットワーク全体に亘る同期インフラストラクチャを確立することである。第２の側面は、要求および性能情報を全体的に発信することや、データ伝送用リソースの割り当てを行う反復的アービトレーション処理を通じて、ネットワークの可能性を広げることである。この構成を使用する際、これら２つの側面は互いに関連しあう。

【００２０】

実施形態は、両方向に流れる個々のシンボル・ストリームのために双方向リンクを提供するコンジットを通じて、ネットワーク構造の中で相互に接続可能なコンピュータ装置を対象とする。ここで、ある特定のシンボルは、セルの開始を識別する。セルは所定の数のシンボルの連続したシーケンスからなる。前記コンピュータ装置の各々は、多数のローカル制御アイドル・シンボルを次のセルの開始前に送信するように構成される。アイドル・シンボルの数は、調整のために、所定の最小値と最大値の間の許容範囲にある。

【００２１】

アイドル・シンボルは、セルの末尾シンボルの後に送信されてもよく、またはセルを構成する複数のシンボル間の任意の位置に挿入されてもよい。しかしこの場合は、アイドル・シンボル機能のために予め定義された特定のシンボルが利用される必要がある。

【００２２】

アイドル・シンボルはシンボル期間の増加によって置換されてもよい。１置換アイドル・シンボル当り１シンボル期間だけセル期間が延長される。

【００２３】

さらに、前記コンピュータ装置は、接続済みの全てのコンピュータ装置もしくは少なくとも一部のコンピュータ装置に対し、一斉にセル送信を開始することができる。各コンピュータ装置は自然数によって識別されるが、この自然数は、１からコンピュータ装置の所定の最大数までの値を取りうる。さらに、各コンピュータ装置は少なくとも１回、その機体識別番号を各接続されたコンピュータ装置に送信する。

【００２４】

前記コンピュータ装置は、自身で送信したセル開始シンボルに対する、受信した各セル

開始シンボルのタイミング・オフセットを測定するように構成される。ただし前記タイミング・オフセットは、前記装置自身のシンボル期間の単位で測定され、また正あるいは負の方向を含んで測定される。そして前記コンピュータ装置は、送信するセル開始シンボルと、受信するセル開始シンボルとの間に、前記測定に関連する接続中のコンピュータ装置へ向けて、測定したタイミング・オフセットを周期的に送信する。さらに具体的には、セル内の１つ又は複数のシンボル位置は、タイミング・オフセット測定データもしくは情報を送信するよう割り当てられる。

【００２５】

ある特定の実装においては、全てのコンピュータ装置は、全ての接続されたコンピュータ装置に、関連するタイミング・オフセット測定情報を相互に周期的に送信する。

【００２６】

別の側面によると、コンピュータ装置は、自分の装置のセル開始タイミング・オフセットと、接続している各コンピュータ装置の受信したタイミング・オフセットとの差を計算する。タイミング・オフセットの差は、コンピュータ装置が次に送信するセルの開始以前に送信されうるアイドル・シンボルの数を決定するために用いられる。それによって、コンピュータ装置は、自身のセル送信開始タイミングを、自身に接続されている他の複数のコンピュータのセル送信開始タイミングのマジョリティに合わせることができる。あるいは、接続されている他の複数のコンピュータ装置のサブセットのみが、使用アイドル・シンボルの数を計算もしくは決定するために考慮されることとしてもよい。ここで、各コンピュータ装置に割り当てられた固有の識別コードが、コンピュータ装置のサブセットを形成するために利用してもよい。当該サブセットは、セル送信開始タイミングの一致度を向上させるための決定の基準として利用される。

【００２７】

コンピュータ装置は該装置の各リンクのためにセル同期状態情報を生成してもよい。この場合、コンピュータ装置は自分のリンクのセル同期状態情報のセットを、各接続されているコンピュータ装置に対して周期的に送信する。当該セル同期状態情報は所定のシンボル位置で符号化される。接続されているコンピュータ装置の固有識別番号の数値順に割り当てられるように送信される。コンピュータ装置はさらに、各接続されているコンピュータ装置から相互に受信したセル同期状態情報を保存もしくは評価することができる。このことは、例えば、接続されているコンピュータ装置から受信したセル同期状態情報を利用して、次のセル開始シンボルより先に送信される必要なアイドル・シンボルの数を計算する、ということである。特定されたアイドル・シンボルの数の最小値と最大値の中点は、アイドル・シンボル数のデフォルト値として定義されてもよい。例えば、適用されたアイドル・シンボルの数は、定義済みデフォルト値に引き寄せられるよう制御されてもよい。追加オプションとして、適用されるアイドル・シンボル数の最大値と最小値の間の中点が、アイドル・シンボル数のデフォルト値から特定の値を超えて外れることがないように、適用されるアイドル・シンボル数の修正変更が制御されてもよい。さらなるオプションとして、コンピュータ装置は他のコンピュータ装置の協調的な動作に依存してもよく、当該他のコンピュータ装置に、適用されるアイドル・シンボル数に関する必要な修正変更を適用してもよい。ある特定の実装によれば、セル間で送信されるアイドル・シンボル数が計算可能であり、それによって、細かくタイミング調整された又はロックされたコンピュータ装置のサブセットにおいて、最初のセル開始タイミングと最後のセル開始タイミングとの中点が特定され、前記適用される数のアイドル・シンボルは、前記中点に向けて自身のセル開始タイミングをシフトする。

【００２８】

別の特定の側面によれば、コンピュータ装置は、自分のセル開始シンボルと接続されている各コンピュータ装置のセル開始シンボルの各々との間のタイミング・オフセットが十分に小さいと判断される場合に、自分のセル同期状態情報を用いて、特定のリンクがデータ送信に利用可能である、という宣言を行うことができる。

【００２９】

さらなる特定の側面によれば、コンピュータ装置は、セルロック状態を維持し、適用されたアイドル・シンボル数に対する複数の要件を順守するために、リクエストをネットワーク内にある全てのコンピュータ装置に広く配信（発信）し、一括でアイドル・シンボル数を増加あるいは減少することができる。

【００３０】

セル毎のシンボル数は、初期化期間中に接続されているコンピュータ装置との折衝の結果に基づいて定義されてもよい。

【００３１】

さらに、各コンピュータ装置は専用コンジットによって他の各コンピュータ装置と相互に接続されてもよい。

【００３２】

複数の実施形態が、通信インターフェース、リソースおよびアービタを有し、前記通信インターフェースを介して他の装置と能動的および受動的な通信を行う能力を有する装置に関連する。前記他の通信装置は、通信の振る舞いが実質的に同一であり、実質的に同一のアービタを備え、ネットワーク内で相互接続されている。ただし、時間区分が前記装置によって活動期として識別され、任意の装置により１つの活動期内に発行される２つのメッセージの何れかが他の装置の１つの活動期内に他の装置によって受信される。前記装置は、受信した通信要素を、これらが発行された活動期内に他の装置へ転送できてもよい。前記リソースは２つ又はそれ以上の状態を有する。前記装置は要求および状態メッセージを発信することができ、後者の状態メッセージは前記リソースの現在の及び／又は未来の状態に関する情報を伝達する。前記メッセージが発信されるにあたり、その発行活動サイクルの間に、前記ネットワークの前記装置が各々を他の各装置から前記メッセージを受信し、さらに同一活動サイクル内で、前記送受信済みメッセージが該装置の各々にあるアービタによって同様に評価されるように、前記メッセージは発信される。次の１または複数の活動サイクルで、状態割り当て情報が計算される。次の活動サイクルで、前記装置は、自身のアービタによるアービトレーション結果から排他的にリソースの状態割り当て情報を抽出する。

【００３３】

他の側面によれば、前記装置はネットワーク構造内で相互接続可能なコンピュータ装置であってもよい。通信手段は双方向リンクを提供するコンジットを介して実現されてもよい。能動的および受動的な通信は、個々のセル・ストリームによって伝達されてもよい。セルは所定の数の複数のシンボルから構成され、ネットワークにおける全てのセル送信は活動サイクルに類似してタイミング調整される。セルは、セル期間内に送信されるシンボルの同期度の評価が可能になるのに十分短い時間内に開始される。セルの特定の位置において、各コンピュータ装置は送信リクエストのリストからなる要求メッセージと、受信機性能および利用できないリソースのリストからなる状態メッセージを、他の全てのコンピュータ装置に発信する。相互接続されているコンピュータ装置の各々は、送信リクエスト、受信機性能および、利用できないリソースの同一データのセットに基づいて同じアービトレーション処理を実行する。結果としてリソース状態が設定されうる。直接送信のためにパスが割り当てられてもよい。要求元から転送エージェントへの、および転送エージェントから要求先へのパスのリソース状態割り当て情報と共に、コンピュータ装置が転送エージェントとして割り当てられてもよい。パスは、方向を有するリンク・コンポーネントである。

【００３４】

本発明の第２の側面は集中制御を除外しておらず、活動サイクルが集中制御から提供されてもよいことに留意されたい。一方で、本発明の第１の側面に従ってセルロック状態が確立されている場合、当該セルロック状態は、本発明の第２の側面の前記活動サイクルに従うタイミングに関する要件が満たされるように実装されてもよい。「セルロック状態」という語句は、本明細書においては「活動サイクルのための要件に従ってタイミング調整されている」ということと同じ概念で使用される。

【００３５】
別のオプションによれば、コンピュータ装置は受信したシンボルのサブセットを、同じセル期間内に、所定のより大きな番号が付けられたシンボル・ポジションによって、１または複数の他のコンピュータ装置に再送することができる。その結果、セルのペイロード・コンテンツの転送が達成される。さらにコンピュータ装置は、１つのセル期間内に状態及びペイロード・データを受信し、格納してもよく、また、これらのデータもしくはこれらのデータの一部を、次のセル期間の間に他の接続されているコンピュータ装置に再送してもよい。再送の対象となった末尾シンボルに続くシンボルのサブセットは、パケット・プロトコルとして利用されうる。代替手段もしくは追加手段として、再送シンボルを含むセル内で、最初に再送されたシンボルより前のシンボルによるサブセットが、パケット・プロトコルとして利用されうる。さらなる代替手段もしくは追加手段として、セルロックネットワークを介して１つのコンピュータ装置から他のコンピュータ装置に送信又は再送されたシンボルが、パケット・プロトコルやストレージ・インターフェース・プロトコルまたは、その他の上位レベルのプロトコルとして利用されうる。制御シンボルまたは制御シンボルのグループは、セル内において等間隔に位置するよう割り当てられてもよく、それによって、データシンボルの再送は、各データシンボルの定数オフセットで達成されうる。

【００３６】
送信元コンピュータ装置の識別はペイロード・データと一緒に送信先コンピュータ装置に向けて送信されないかもしれない。しかし送信先コンピュータ装置は、この情報をアービトレーション結果から抽出することができてもよい。

【００３７】
アービトレーション処理は、ある送信リクエストを送信元コンピュータ装置と送信先コンピュータ装置とを直接的に接続しているパスに割り当てることができ、また、残りの送信リクエストを、送信元コンピュータ装置からのデータ入力パス及び送信先コンピュータ装置へのデータ出力パスに割り当てることができる。後者の割り当ては、使用されていない転送エージェント役のコンピュータ装置を割り当てるアルゴリズムを介して行われる。この場合、アービトレーション処理は、異なる転送エージェント役のコンピュータ装置を各送信リクエスタ・コンピュータ装置に割り当てる各々のステップを伴う、循環的シーケンスにおけるペイロード転送パス割り当てを計算することができる。パスが各セル期間に複数のセルを送信するよう設定されている場合、アービトレーション処理は、１つのステップ内で、同じ送信元および送信先コンピュータ装置間で同じ送信パスに向けて複数の送信リクエストを割り当ててもよく、それによってアービトレーション手順を迅速に進める。あるいは、アービトレーション処理は、１つのステップ内で、同じ送信元および送信先コンピュータ装置間で同じ送信パスに向けて可能な最大送信回数を割り当ててもよい。送信パスは、あるセル期間内に複数のセルを伝達することができる送信パスであり、優先順位が指定されてもよい。

【００３８】
一部もしくは全てのコンピュータ装置は、複数のレーンで構成されるリンクと相互接続されてもよく、また、追加の複数のレーンは接続されているコンピュータ装置間の追加のセルの同時送信に利用してもよい。複数のデータ・パスが要求元から送信先コンピュータ装置に割り当てられる場合、該データ・パスの順次割り当てを統制するための規則が確立される必要がある。

【００３９】
セル転送レイヤ上に割り当てられる複数のパケット・プロトコルが、ネットワークに共存可能である。さらに、複数のパケット・プロトコル・インターフェースがネットワークに提供されてもよく、一方でパケット・プロトコル・インターフェースはコンポーネントまたはネットワーク構成要素内にのみ現れることが出来る。

【００４０】
ある特定の実装において、コンピュータ装置は、自分のシンボル期間を単位として、接

続されているコンピュータ装置から所定の数のシンボルを受信する時間を測定してもよい。次いで、コンピュータ装置は、測定した結果をネットワーク内の他のコンピュータ装置に送信してもよい。それにより、例えば、その測定結果を利用して、障害の検出を行うことができる。

【００４１】

さらなる側面によれば、同期されているコンピュータ装置は、セル配列のナンバリング方法を適用してもよい。それにより、例えば、少なくとも一部のセルは連続した番号を含むことが可能となる。

【００４２】

ある特定の実装において、あるセルにおける末尾シンボルとそれに続くセルの先頭シンボルとの間に適用されたシンボルは、制御情報または追加のペイロード・データを転送するために利用することができる。

【００４３】

さらなる側面によれば、外部クロックが、外部クロック・ソースに接続されているコンピュータ装置によってネットワーク内に配信される。そこでコンピュータ装置は、所定のシンボル位置を介して、外部クロックの重要な境界に適合するセル開始シンボルに対する、セルおよびシンボル位置を識別することができる。セル周期は、外部クロック周期がセル周期の整数倍であるように選択されうる。なおセル周期は、セルのシンボル数と、セル間に適用される定数個のアイドル・シンボルのシンボル数の合計であり、外部クロック周期がセル周期の整数の倍数であるように選択されうる。クロック・タイミング情報は送信されたセル内のみに送信されてもよく、もしくは複数のセルに分散して送信されてもよい。コンピュータ装置は他のコンピュータ装置のクロック情報を自分の送信セルを介して転送することができる。さらに、接続されているコンピュータ装置は、外部参照クロック・ソースのクロックの品質レベルについての情報を広く知らせることができる。

【００４４】

さらなる側面によれば、セルにおける１または複数のシンボル位置がパケット・プロトコルに割り当てられ、それによって、送信されるパケットのシンボルが、後に続く複数のセルにおいて割り当てられたシンボル位置に配置される。さらに具体的には、低レベル情報がパケット内に周期的に配され、その結果、セル内の１つの所定のシンボル位置は、一定のものを沢山、もしくは低帯域幅のデータを発信するのには十分である。

【００４５】

さらなる側面によれば、リンクのサブセットは複数のベーシック・シンボル・レートで動作する。ある特定の実装において、リンクは複数のN個のベーシック・シンボル・レートで動作可能であり、そのためN個のシンボルがベーシック・シンボル・レートのあるシンボル期間に送信される。この場合、ベーシック・シンボル・レートでセル期間のシンボル期間番号Kで送信された次のN個のシンボルは、N個のセルのシンボル番号Kに関連付けられる。ここで、N個のセルのシンボルは、シンボル位置毎にインターリーブ送信されてもよい。コンピュータ装置は、外部リソースから自分の識別番号に依存した自分のリンクのための特定のシンボル・レート乗率についての情報を取得することができる。第１のシンボル・ストリームまたはその他はっきりと識別されたシンボル・ストリームは、セル同期の初期化や保守に割り当てられたシンボルを提供してもよい。

【００４６】

別の側面によれば、複数の実施形態はメモリ装置を含んでもよく、該メモリ装置は：

・　循環的アドレス指定シーケンスにおいてストレージ・ロケーションに書き込み可能であり、また、少なくとも１つの同様に指示される別のサイクリック・アドレス指定シーケンスにおいストレージ・ロケーションから読み出し可能な、ストレージ・アレイ；

・　前記ストレージ・アレイのアドレスのためのアドレス・レジスタであって、前記ストレージ・アレイは、書き込みシーケンスと同相のトリガイベントにおいて、書き込みアドレス生成器から出力された現在のアドレスによってロードされ、前記アドレス・レジスタのコンテンツは、前記ストレージ・アレイの読み出しアドレス生成器のためのプリロード

のソースとして利用され、前記プリロードは読み出しシーケンスと同相に行われるイベントに適用される、アドレス・レジスタ；
とを備える。
【００４７】
さらなる側面によれば、イベント・シンクロナイザーが上述のメモリデバイスに提供され、イベント・トリガをストレージ・アレイの書き込み側から読み出し側に向けて同期する。該メモリ装置は例えば複数のインスタンスにおいて並列に利用され、各インスタンスのための読み出しアドレス生成器のプリロードは同時にトリガされる場合がある。
【００４８】
さらなる側面によれば、前記書き込みアドレス生成器および前記読み出しアドレス生成器は、読み出しカウンタがプリロードされるステップは含められないがアドレス指定サイクルが一巡する時は含められるアドレス指定ステップ毎に、アドレスコードのちょうど１ビットを変更することで、同じスキームに従ってそれらの循環的アドレス指定シーケンスを生成してもよい。
【００４９】
さらなる側面によれば、ネットワーク構造は、上記に定義された複数のネットワーク接続可能なコンピュータ装置を含んでよく、１つのセル期間内の所定のシンボル位置において送受信されるシンボルについての評価が行われてもよい。それは、前記シンボルにより運ばれるデータコードのビットのシリアル送信をエミュレートするためである。エミュレートされるシリアル送信はオープン・コレクタ・バス相互接続を介して達成される。前記エミュレートは、各コンピュータ装置が個別に、他のコンピュータ装置から特定のシンボル位置で受信したシンボルを評価するように実装される。この評価は、最上位ビットから始まるシンボルのビット・シーケンシャル評価のためのビットが、チェックされたシーケンスにおいて識別された所定のビット値またはビットパターンを伴うデータに優先権を与えるように行われる。前記チェックは、ただ１つの送信機が優先権を有すると識別されるまで1ビットずつ続けられ、そして、この送信機からのデータがこのシンボル送信の結果でる。ある特定のシンボル位置における所定の非データシンボルの送信は遅延として扱われてもよく、前記現在のセルのこの特定シンボル位置における全てのシンボルが無視される。
【００５０】
ある実施形態は、１つまたは複数のコンピュータ装置を有するネットワーク構造の中でデータを送信する方法を対象としてもよい。この方法において、前記１つまたは複数のコンピュータ装置は、それぞれ、前記ネットワークへのインターフェースとして実質的に同一のインターフェースを有する。前記インターフェースは、両方向に流れる個々のシンボル・ストリームのために双方向リンクを提供するコンジットを介するインターフェースである。前記方法は、所定の数のシンボルの連続したシーケンスから構成されるセルの先頭を特定するセル開始シンボルである１つの特定のシンボルを提供することと；前記複数のコンピュータ装置のいずれかから、ローカルで決定された数のアイドル・シンボルを次のセルの開始前に送信することと；含み、前記アイドル・シンボルの数は、所定の下限値と上限値との間で調整するための許容範囲を有し、アイドル・シンボルの挿入は、動的変更によりシンボル期間に代替可能であり、それによって、セル送信期間が、代替されたアイドル・シンボルの期間によって延長される。さらに前記方法は、前記セル開始シンボルを、接続している全てのコンピュータ装置に一斉に送信することと；接続している全てのコンピュータ装置の各々に、同じ数の前記アイドル・シンボルを適用することと；を含む。
【００５１】
ある実施形態は、１つまたは複数のコンピュータ装置を有するネットワーク構造の中でデータを送信する方法を対象としてもよい。この方法において、前記１つまたは複数のコンピュータ装置は、それぞれ、前記ネットワークへのインターフェースとして実質的に同一のインターフェースを有し；前記インターフェースは、両方向に流れる個々のシンボル・ストリームのために双方向リンクを提供するコンジットを介するインターフェースで

ある。前記方法は、実質的に同じ通信動作および実質的に同じアービタを備える前記コンピュータ装置を提供することと；時間区分を活動期として識別することと；１つの活動期内に前記ネットワーク上でコンピュータ装置によって発行されるアクティブ通信の任意の２つの要素を１つの活動期内で受信することと；受信した通信要素を前記ネットワーク上の他の装置にそれらが発行されたた活動期内に転送することと；次の活動サイクルで状態または状態シーケンスを決定することと；リソースの状態及び／又は未来の状態を通知する要求メッセージ及び状態メッセージを発信すること、ただし前記メッセージを、前記ネットワーク上の前記コンピュータ装置が、他の前記コンピュータ装置から前記メッセージの各々を、それらが発行された活動サイクル内に受信できるように、発信することと；前記発信されたリクエストおよび状態メッセージのコンテンツを入力データとして使用することと；次の１または複数の活動サイクルで、前記ネットワーク上の全ての装置の全てのリソースに対する前記状態または状態シーケンスを計算することと；次の活動サイクルで、各々のアービトレーション結果から排他的に各々のリソースの状態割り当て情報を抽出することであって、前記アービトレーションは各活動サイクル内で実行される、該抽出することと；を含む。

【００５２】

別の側面によれば、コンピュータ装置は、該コンピュータ装置の動作を記述したコンピュータ・プログラムのデータセットとして、具現化されてもよい。及び／又は、ある特定の製造技術をターゲットとしたコンピュータ装置の物理的なインスタンス化を表すデータセットに変換可能なソース・データ・セットとして具現化されてもよい。ターゲットとなる技術はプログラマブルロジックデバイス（programmable logic device：PLD）であってもよく、その場合、データセットは該PLDの構成に使用されるビットストリームとして現れてもよい。

【００５３】

別の側面によれば、コンピュータ・プログラムはコード手段により上に特定された方法の各ステップを実行してもよい。

【００５４】

さらなる好適な改良は、従属請求項において定義される。

【図面の簡単な説明】

【００５５】

本発明を、以下の図を参照つつ、複数の実施形態に基づいて説明する。

【図１】本発明を実装可能なネットワーク・アーキテクチャを示した略ブロック図。

【図２】コンピュータ装置間の双方向リンクの例。

【図３】コンピュータ装置の初期化シーケンスを示したフローチャート。

【図４】相互接続されたコンピュータ装置が第２のセルロックネットワークをどのように利用するか方法について、いくつかの方法を示した略ブロック図。

【図５】複数の高速回線によるインターリーブ送信の例。

【図６】シンボルの送信ストリームの構造。

【図７】ある実施形態における、２つのセルのフォーマットの例。

【図８】追加セルの構成の例。

【図９】ペイロード転送手順の簡単な例。

【図１０】セルロックネットワークのオペレーションについて、特にペイロード転送メカニズムに関する例。

【図１１】６台のコンピュータ装置からなるクラスター・ネットワーク構造の例。

【図１２】ある独特な非同期FIFOメモリの概念を示した回路図。

【図１３】複数のプロトコルをサポートするセルロックネットワークの例。

【図１４】ある実施形態のネットワーク・インターフェースを示した略ブロック図。

【図１５】コンピュータおよびネットワークシステムの関連項目を階層に示した表。

【図１６】ある実施形態における、セル内のシンボル位置割り当て一覧。

【図１７】ある実施形態における、セルロック状態および細かいタイミング調整状態のた

めの制御シンボル位置割り当て一覧。
【図１８】ある実施形態における、粗いタイミング調整状態のための制御シンボル位置割り当て一覧。
【図１９】ある実施形態におけるパケット・コンテンツの一覧。
【図２０】ある実施形態における接続の符号化状態の一覧。
【図２１】ある実施形態における、送信コンピュータ装置のための、リンクの完全動作のためのビット・エンコード一覧。
【図２２】ある実施形態における送信リクエスト・コード一覧。
【実施形態の説明】
【００５６】
以下は、全二重データ送信リンクによってフルメッシュ相互接続可能な複数のコンピュータ装置101間におけるロックされたセルデータの送信に基づいて、本発明の実施形態を説明する。
【００５７】
ある実施形態は、共に高性能ネットワーキングおよび多用途性の可能性を開く２つのコンセプトに基づく。第１のコンセプトは、セルロック中のネットワーク（CLN： Cell Locked Network；セルロックネットワーク）410を形成する。CLNはフルメッシュ・ネットワーク105全体におけるセルベースの同期マルチパス・データ転送アーキテクチャであり帯域外シグナリングや集中制御を伴わない。第２のコンセプトは、セル期間603毎に、データ・パスへのデータ転送リクエストの自動割り当てを生じさせ、それによって、セルロックネットワーク410において、ハードウェア制御によるダイナミック・マルチパス・ルーティング制御が可能になる。
【００５８】
複数のコンピュータ装置101が、全二重ポイント・ツー・ポイント接続リンク221により相互接続されている。これらのコンピュータ装置101は相互に接続されているため、任意の２つのコンピュータ装置101間で直接接続リンク221を利用することができる。このようなトポロジーは、フルメッシュ・相互接続トポロジーと呼ばれる。図１は４つのコンピュータ装置APP_A 101、APP_B 102、APP_C 103、APP_D 104間のフルメッシュ・相互接続トポロジーの例を示す。
【００５９】
コンピュータ装置101という言葉は、ここではコンピュータやストレージ・要素、入／出力ノードなどを意味し、セルロックネットワーク410である実施形態の主人公であるネットワークに接続されたどのコンポーネントを意味してもよい。
【００６０】
リンク221の実装においては、エンベデッド・クロックを伴うシリアル・ビット・ストリームを利用してもよい。一般的には、リンク221は、両方向においてシンボル222のためのデータ送信媒体を提供する任意のコンジットになりうる。シンボル222は、データ、フレーミング、その他必要に応じた制御を示す。
【００６１】
コンピュータ装置101は共振器を備えることができ、また、所定の周波数許容誤差内でクロックを生成する。シンボル送信レートの基準であるデータ送信クロック周波数は、共振器もしくはそれに類するものの周波数から計算される。共振器およびクロック・ジェネレータは各コンピュータ装置101にローカルに備えられている。各コンピュータ装置101は、複数のシンボル222を、一応同じとされているが実際は個々にわずかに違っているシンボル・レートで送信する。環境条件により、シンボル送信レートはまた、共振器の周波数域内で動的に変化してもよい。
【００６２】
セルロックネットワーク410内の各コンピュータ装置101は、それぞれジオグラフィック・アドレス106と呼ばれる固有のアドレスを持つ。任意の実施形態で最大N個のコンピュータ装置について取り扱う場合、ジオグラフィック・アドレス106は1からNまでの自然数と

なる。相互接続している各コンピュータ装置101はセルロックネットワーク410内において各々のジオグラフィック・アドレス106によって識別される。ジオグラフィック・アドレス106は、例えばラックマウント型のシステム内の特にスロット毎に、コーディングピン接続を介して各コンピュータ装置に割り当てられる。また、自立型コンピュータ装置を相互接続している場合には、セットアップ・ジャンパによって各コンピュータ装置に割り当てられる。リンク221の初期化の間、コンピュータ装置101は自分のジオグラフィック・アドレス106を少なくとも複数の送信セル601の一部に含め、さらにコンピュータ装置101はリンク221のもう一方のエンドで同じ動作を行う。これにより、各リンク221にジオグラフィック・アドレス106を持つ接続済みコンピュータ装置が相互に認識可能となる。

【００６３】

図６はシンボルの送信ストリームの構造である。各コンピュータ装置は、シンボル222を、接続中の他の複数のコンピュータ装置に向けて自分のシンボル・レートで連続して送信する。シンボル222の送信ストリームは、各セル601の先頭に設けられる固有のセル開始シンボル209によって、セル601として構築される。

【００６４】

セル601は、開始シンボル209から始まるシンボル222が所定の数606連続したものとして定義される。次のセル602が開始する前に、セルの末尾シンボル220の後ろに少数のアイドル・シンボル210が続いてもよい。

【００６５】

各セルのシンボル数606や、アイドル・シンボル210のデフォルト数605、アイドル・シンボル210の最小数や最大数は、実施形態によって適宜設定される。

【００６６】

コンピュータ装置101は、他の接続されている全てのコンピュータ装置に向けて、同じタイミングで送信を実行する。コンピュータ装置101は他の接続されている全てのコンピュータ装置に向けてセル開始シンボル209を同時に送信し、次いでセル601の残りのシンボル222を送信し、セル601の末尾シンボル202の後に同じ数のアイドル・シンボル210を送信する。

【００６７】

各コンピュータ装置101はその装置のタイミングに基づいて動作するため、セル開始シンボル209の送信タイミングはコンピュータ装置101毎に異なる。

【００６８】

通常、伝送遅延224はリンク221毎に異なり、また、送信路の構造によってどのような遅延が生じるかは不明である。伝送遅延224は、同一リンクにおいては双方向とも見かけ上は全く同じになると考えられる。伝送遅延224は一定のわずかな許容範囲内に留まる。セルロックネットワーク410のオペレーションについて、シンボルの伝送遅延224はセルの長さ603よりもかなり短いが、いくつかのシンボル225にかぶることはある。

【００６９】

セルロックネットワーク410でセル601の送信の同期を維持するために、各コンピュータ装置上での動的な調整機能が必要とされる。動的な調整機能を提供するために、送信されるアイドル・シンボル210の数は、各コンピュータ装置101におけるセル周期604毎に個別に決定される。セル601の送信の同期が実現されるために、リンク221上でセルロック状態310が確立される。そのような状態を、セルロックネットワーク410においてセルがロックされたという。

【００７０】

ロック状態を達成および維持するために適用されるアイドル・シンボル210の数をコンピュータ装置が決定する方法は、複数ある。これら複数の実施形態は特定のアルゴリズムを限定するためのものではない。

【００７１】

アイドル・シンボルのデフォルト数605を定義するために、セル606毎のシンボル数およびシンボル・レートの許容誤差値をもとに最適な数が計算される。アイドル・シンボルの

デフォルト数605は、挿入するアイドル・シンボル210の数を増減させることで修正・変更できなければならない。

【００７２】

各コンピュータ装置101は、自身が送信したセル開始シンボル209と、実装されたリンク221によって受信したセル開始シンボル219との間のオフセット226を測定する構成を備える。オフセット226は、コンピュータ装置101の送信したセル開始シンボル209と受信したセル開始シンボル219の間のシンボル期間225をカウントすることで決定される。なおシンボル期間225はコンピュータ装置101におけるものである。オフセット226の測定値が0ということは、受信したセル開始シンボル219とコンピュータ装置101のセル開始シンボル209の送信が同時であったということである。オフセット226の測定値が正の値ということは、セル開始シンボル219の受信がセル開始シンボル209の送信よりも遅かったということである。オフセット226の測定値が負の値ということは、セル開始シンボル219の受信がセル開始シンボル209の送信よりも早かったということである。測定方法および正の値・負の値の割り当ては、ある実装の要素であって、本発明の技術範囲を制限するものではない。

【００７３】

各コンピュータ装置101は、リンク221を介して接続されている他のコンピュータ装置101に対して、当該リンク221について測定したセル開始オフセット値226を伝える機能を追加的に備える。測定データはシンボル222に符号化され、セル601の所定のシンボル位置で送信される。

【００７４】

図２はコンピュータ装置APP_A 101およびAPP_B 102間の双方向リンク221の例である。この例においては、リンク221は、すでにセルロック状態（310）となっている。

【００７５】

図２の上部に概略を示す。APP_A 101はリンク221を介してAPP_B 102に接続されており、リンク221は向きが固定された２つのデータ転送パスAB215およびBA216で構成される。シンボル・シーケンス・タイミングの例が以下の測定ポイントで示される：

・測定ポイント211におけるシンボル・シーケンス201
・測定ポイント212におけるシンボル・シーケンス202
・測定ポイント213におけるシンボル・シーケンス203
・測定ポイント214におけるシンボル・シーケンス204

【００７６】

シンボル・シーケンス201はAPP_A 101によってパスAB 215を介してAPP_B 102に送信される。この例においては、APP_A 101はセル601の最終シンボル222の後にアイドル・シンボル210である２つのSKP(スキップ)シンボルを送信する。次のセル602は、セル開始シンボル209であるCOM（コンマ）シンボルで始まる。空のボックスはここでは説明されていないが、追加のシンボル222を表す。

【００７７】

APP_A 101が受信するシンボル・シーケンス202は、APP_B 102が送信したシンボル・シーケンス203と同一のものである。しかし、接続パスBA 216における伝送遅延によりDBA 224だけ遅延する。

【００７８】

シンボル・シーケンス203はパスBA 216を介してAPP_B 102からAPP_A 101 に送信される。APP_B 102は、次のセル601開始前に３つのSKPシンボル210を送信する。

【００７９】

APP_B 102が受信するシンボル・シーケンス204はAPP_A 101によって送信されたものであり、シンボル・シーケンス201がDAB 224だけ遅延したバージョンとして出現する。

【００８０】

シンボル・シーケンスのタイミング201および202を見ることができるのはAPP_A 101のみであり、一方でシンボル・シーケンスのタイミング203および204を見ることができるのはAPP_B 102のみである。

【００８１】
図２はAPP_A 101によるセル開始シンボル209の送信時間205、APP_B 102によるセル開始シンボル209の送信時間206、セル開始シンボル209がAPP_A 101からAPP_B 102に到達するまでの時間207、セル開始シンボル209がAPP_B 102からAPP_A 101に到達するまでの時間208についての時間関係を示している。
【００８２】
$D_{AB} = D_{BA}$ 224はABパス215およびBAパス216のそれぞれの遅延であり、等しくなるように定義されている。
【００８３】
M_{AB} 226はAPP_A 101において、APP_B 102から受信したセル開始シンボル219のオフセット値として測定される。
【００８４】
M_{BA} 229はAPP_B 102において、APP_A 101から受信したセル開始シンボル219のオフセット値として測定される。
【００８５】
X_T 228は、APP_A 101のAPP_B 102に対する、セル開始シンボル209の送信時間のタイミング・オフセットである。
【００８６】
X_R 230は、APP_A 101のAPP_B 102に対する、それぞれ他のコンピュータ装置101からセル開始シンボル209を受信する時間のタイミング・オフセットである。
【００８７】
APP_A 101はM_{AB} 226の測定値を認識している。
【００８８】
APP_B 102は符号化した自分の測定値M_{BA} 229をAPP_A 101の所定のシンボル位置で送信する。
【００８９】
M_{AB} 226およびM_{BA} 229を利用して、APP_A 101は適用すべきアイドル・シンボル210の数を決定することができる。
【００９０】
$D_{AB} = D_{BA}$ と $X_T + D_{AB} = D_{BA} + X_R$ から、$X_R = X_T$ が導かれ、$X_T + M_{AB} + X_R = M_{BA}$ with $X_R = X_T$ から、$2 * X_T = M_{BA} - M_{AB}$ が導かれる。
【００９１】
X_T 0となるためには、接続されているコンピュータ装置101の受信・測定したオフセット値が限りなく0となる必要がある。そのためにはAPP_A 101およびAPP_B 102を接続するリンク221のセルロック状態（310）が達成され維持される必要がある。
【００９２】
フルメッシュ・トポロジー105において複数のコンピュータ装置101が接続される場合には、状況はさらに複雑になる。この状況においてセルロック状態を維持するための２つの方法を示す。
【００９３】
第１の方法は次の通りである。ある特定のコンピュータ装置101を、タイミングに関する、他のコンピュータ装置全ての参照先とし、当該他のコンピュータ装置101の全ては上述の原則に従って調整を行う。また、参照先であると考えられるコンピュータ装置101は、常にアイドル・シンボルのデフォルト数605を適用する。上述の、ジオグラフィック・アドレス106を利用してセルロックネットワーク410上の相互接続している複数のコンピュータ装置101を識別するという可能性は、例えば、セルロックネットワーク410の同期を図るために、一番小さいジオグラフィック・アドレスを有するコンピュータ装置101をタイミングの参照先として定義することを許す。タイミングの参照先となることを宣言するコンピュータ装置101は、この情報を各セル601に広める。しかし、例え一時的であっても、複数のコンピュータ装置101がタイミングの参照先になることは決して無い。

【００９４】
第２の方法は次の通りである。コンピュータ装置101は、測定したオフセット226のデータおよび受信したオフセット229の情報を利用して、接続中の他の全てのコンピュータ装置101のそれぞれについて差$M_{AB} - M_{BA}$を計算し、それによって、自分のセル開始シンボル209のタイミングと、他のコンピュータ装置101のセル開始シンボル209のタイミングとを比較する。自身のセル開始タイミング227を0とした場合の、他のコンピュータ装置101のセル開始タイミング227の時系列（chronology）が得られる。このリストの要素にはゼロが含まれると共に、セル開始タイミング227が最も早かったもの対応する値が最小値となり、もっと遅かったものに対応する値が最大値となる。セル開始タイミング227の最小値と最大値の中間の値が、タイミング調整の対象値となる。自身のセル開始タイミング227の値が「0」なので、中間の値はアイドル・シンボルのデフォルト数605からの必要な調整値についての正しい方向を示している。しかし、その値は、アイドル・シンボル210について、許容される最大値と最小値の間に残るように選択されなければならない。この計算は、以降の実施形態で詳述する。

【００９５】
前述の内容を、セルロックネットワーク410の簡単な定義を用いて要約する。

【００９６】
相互接続されているコンピュータ装置101は、ある所定の許容誤差の範囲内で、特定のシンボル・レートでシンボル222を送信する。複数のコンピュータ装置101は、双方向にシンボル222を送信するためにリンク221に相互接続されており、双方向のパスAB215およびパスBA216における遅延は一致していることになっている。コンピュータ装置101に接続しているリンク221の遅延はかなりばらつきがあってもよいが、所定の最大値を超えてはならない。セル601はユニーク・セル開始シンボル209で始まるシンボル222の、所定の長さのシーケンス606として特定される。セルロック状態は、以下の条件が持続する場合に成立する。各々のコンピュータ装置101が、当該コンピュータ装置101にリンク221を介して接続している他のコンピュータ装置101に、セル601を同時に送信する。他のコンピュータ装置101によって行われるセル601の送信は、セルの長さ603に比べて極めて短い時間内に開始される。各セル601の送信後には、いくつかのアイドル・シンボル210が送信される。同じコンピュータ装置101は、全てのリンク221において同じ数のアイドル・シンボル210を送信するが、異なるコンピュータ装置101は異なる数のアイドル・シンボル210を用いてもよい。接続されているコンピュータ装置の多くにとって、次のセル送信を開始するタイミング227の同期性が向上するのであれば、アイドル・シンボル210の数として、定義されたデフォルト数605やその他の数を用いてもよい。この方法は、接続している複数のコンピュータ装置101のシンボル・レートのわずかな差によるセル開始タイミング227のずれを動的に正す。

【００９７】
アイドル・シンボル605のデフォルト数の正しい値を見つけ出す方法の一例を次に示す。

【００９８】
セル601は3.000のシンボル222で構成されており、クロック許容誤差は±300ppmである。この許容誤差値は、1.000.000のシンボル222の送信が、見かけ上の時間±300のシンボル期間225を要とするということを意味する。換言すれば、シンボル・レートの許容誤差の一番大きな端にある第１のコンピュータ装置101が1.000.300のシンボル222を送信する間に、シンボル・レートの許容誤差の一番低い側にある第２のコンピュータ装置101は999.700のシンボルを送信するということである。これは、3.000個のシンボル222によるセルの長さ606のために再計算することができる。第１のコンピュータ装置101が3.001個のシンボル222を送信するのと同じ時間で、第２のコンピュータ装置101は2.999個のシンボルを送信する。すなわち、第２のコンピュータ装置101は第１のコンピュータ装置から、セル601を、の自身の2998個のシンボル期間225内に受信し、第１のコンピュータ装置は第２のコンピュータ装置101から、セル601を、自身の3002のシンボル期間225内に受信する。

【００９９】
次のような量や関数を定義する。

T: 許容誤差量の比
例）T: = 0.000300

n: 各セル601のシンボル222の数

margin: ある程度の緩みを確保するためのシンボル期間225の割合

trunc(): 切り捨て関数

P_{min} , P_{nom} , P_{max}: セル期間603の最小値・公称値・最大値
$P_{min} := P_{nom} - n * T$
$P_{max} := P_{nom} + n * T$
従って、$P_{max} - P_{min} := 2 * n * T$

【０１００】
このとき、ある１つのコンピュータ装置101がタイミング参照先として宣言されている場合（第１の方法）に計算されるアイドル・シンボル210の最小数・デフォルト数605・最大数（$IS1_{min}$, $IS1_{default}$, $IS1_{max}$）は、次の通りとなる。

$P_{max} + IS1_{default} \leqq P_{min} + IS1_{max}$
$P_{min} + IS1_{default} \geqq P_{max} + IS1_{min}$

ここで $IS1_{min} := 0$ とすると、
$IS1_{max} - IS1_{default} \geqq P_{max} - P_{min}$
$IS1_{default} \geqq P_{max} - P_{min}$

従って、
$IS1_{default} := trunc(P_{max} - P_{min} + 1 + margin)$

そして、$P_{max} - P_{min} = 2 * n * T$ であるので、

$IS1_{default} := trunc(2 * n * T + 1 + margin)$
$IS1_{max} := 2 * IS1_{default}$

【０１０１】
全てのコンピュータ装置101が協働する場合（第２の方法）に計算されるアイドル・シンボル210の最小数・デフォルト数605・最大数（$IS2_{min}$, $IS2_{default}$, $IS2_{max}$）は、次の通りとなる。

$P_{min} + D2_{max} \leqq P_{max} + D2_{min}$
$IS2_{max} - IS2_{min} \leqq P_{max} - P_{min}$

ここで$IS2_{min} := 0$ とすると、
$IS2_{max} \leqq P_{max} - P_{min}$

従って、
$IS2_{max} := trunc(P_{max} - P_{min} + 1 + margin)$

$P_{max} - P_{min} = 2 * n * T$であるから、

$IS2_{max}$:= trunc(2 * n * T + 1 + margin)
$IS2_{default}$:= $IS2_{max}$ / 2
【０１０２】
この場合の結果は：
$IS1_{default}$:= trunc(2 * 3000 * 0.000300 + 1 + 0.3)
$IS1_{default}$:= trunc(1.8 + 1.3) = 3
$IS1_{max}$:= 2 * 3 = 6

また、
$IS2_{max}$:= trunc(2 * 3000 * 0.000300 + 1 + 0.3)
$IS2_{max}$:= trunc(1.8 + 1.3) = 3
$IS2_{default}$:= 3 / 2 = 1.5

となる。
【０１０３】
上述の計算によれば、全てのコンピュータ装置101が、セルロック状態を確立し維持する上で協働し合う場合（第２の方法）の例において、アイドル・シンボル210の数は0から3の範囲であり、アイドル・シンボルのデフォルト数$IS2_{default}$は1.5である。
【０１０４】
セルロック状態310を確立し維持するために、宣言されたタイミング参照を伴う第１の方法を採用する場合は、状況が異なってくる。シンボル・レート許容誤差の幅の中でタイミング参照先のコンピュータ装置101がどこにあるのか不明なので、最悪の両ケースに備えるための体制が必要となる。計算されたとおり、アイドル・シンボルのデフォルト数605である$IS1_{default}$:= 3と割り当てられなければならず、また、許容範囲は0から6となる。この計算の例のように、参照シンボル・レートの許容誤差がゆるい場合においては、この方法では不十分だということがわかる。高精度のクロックを有するコンピュータ装置101がシンボル・レートの参照先として利用される場合においてのみ、この影響による問題が生じない。
【０１０５】
よりきついクロック許容誤差とより短いセル606は、アイドル・シンボルのデフォルト数605が必要とする数を減少させる。例えば、クロック許容誤差が±50 ppm（ppm:百万分の一）、および実質的なセルサイズ606がシンボル222の1.000個分である場合、アイドル・シンボル210のデフォルト数605の値は1、最小値は0、最大値は2となりうる。
【０１０６】
コンピュータ装置101はシンボル・レート許容誤差の幅の中での自分の位置情報を所有していない。コンピュータ装置101は、接続されている各コンピュータ装置101のシンボル送信レートを、自身のシンボル期間225を単位として測定しなければならない。あるコンピュータ装置が別のコンピュータ装置101のずれを観測する場合、両者のうちのどちらが許容誤差の範囲の外にあるのかは不明である。複数のコンピュータ装置101が同様にシンボル送信レートを測定し、ほぼ全ての装置が健全な状態にあると仮定する場合、クロック・レートが許容誤差外にあるコンピュータ装置101を特定することは容易である。挿入できるアイドル・シンボル210の数は、所定の範囲内に定められているので、ある閾値を超えると、アイドル・シンボル210の挿入による調整によってはシンボル・レートの逸脱を補正することができなくなり、ネットワーク410のセルロック状態310を達成および維持することは不可能となる。この場合、閾値を超えるコンピュータ装置101は動作を取り止めなければならない。エラー修復等についての詳細を以下に説明する。
【０１０７】
セルロック状態が成立している場合、全てのコンピュータ装置101は各々のセル開始シンボル209をほぼ一斉に送信する。リンク遅延224が異なる場合には、異なる送信元からの

セル開始シンボル209はそれぞれの送信先コンピュータ装置101に同時には受信されない、という結果になる。セルロック状態一般および複数の実施形態の範囲は、ここに説明されたセルロック状態を成立し維持する方法を強制するものではない。

【０１０８】

セルロックネットワーク410のある実施形態は、セル開始オフセット測定226のデータを符号化したものや、セル601におけるそのシンボル位置を割り当てなければならない。

【０１０９】

図３は任意のコンピュータ装置101の初期化手順を簡単に表したフローチャートである。

【０１１０】

ステップS0 301: コンピュータ装置101の起動時は、該装置のセルロックネットワーク410へのリンク221が論理的・物理的に非アクティブである。

【０１１１】

ステップS1 302: いくつかの基本的な初期化が行われた後、コンピュータ装置101は周期的なビーコンなどにより、サポートされる全てのリンク221にプレゼンス信号を送信し始める。

【０１１２】

ステップS2 303: すでにセル601を他のコンピュータ装置101に送信している別のコンピュータ装置101が、このコンピュータ装置101にセル送信を開始することを許可するために、短い待機期間が必要である。

【０１１３】

ステップS3 304: コンピュータ装置101がリンク221上で検出された送信機の数をチェックする。送信機がセル601を送信しているか、ただプレゼンス信号を発信しているかに関わらず、その数をカウントする。検出された送信機数が0もしくは1であるか、それ以上であるかによって、カウントが継続される。送信機が検出されなければこの状態が持続され、チェック処理は無期限に繰り返される。

【０１１４】

ステップS4 305: ステップS3 304において、複数の送信機が検出された場合、セル601を送信しているコンピュータ装置101の数が決定される。

【０１１５】

ステップS5 306: ステップS4 305において、接続されている複数のコンピュータ装置101がセル601を送信していると同定された場合、それらのコンピュータ装置101へのリンク221がすでにロック状態S9 310になっているかどうかがチェックされる。もしロック状態になっていなければ、ステップS5 306は他のコンピュータ装置101へのリンク221がロック状態S9 310になるまで、チェックを続ける。

【０１１６】

ステップS6 307: ステップS3 304において、接続されたコンピュータ装置101が１つだけ検出された場合、またはステップS4 305 においてセル601を送信しているコンピュータ装置101が0ないし1つしか検出されなかった場合、またはステップS5 306においてセル601を送信しているコンピュータ装置101へのリンク221が全てロック状態（S9 310）であるか、のいずれかが検出された場合、このコンピュータ装置101は、セル601の送信を開始しなければならない。セル開始シンボル209はすでに作動している送信機に合わせて、コンピュータ装置101が接続している全てのリンク221上で同時に送信されなければならない。最初に送信されるセル601は粗いタイミング調整状態（S7 308）を要求するフォーマットでなければならない。

【０１１７】

ステップS7 308: 接続されている全てのコンピュータ装置101のセル開始タイミング227が所定の比較的大きな時間にタイミング調整されるまで、粗いタイミング調整処理が行われる。

【０１１８】

ステップS8 309: 細かいタイミング調整処理によりロック状態S9 310が実現される。

【０１１９】

ステップS9 310: 最終的にはリンク221がロック状態になる。

【０１２０】

ネットワーク410へのリンク221を起動するコンピュータ装置101は、セルロックネットワーク410の全ての実装されているインターフェース・リンク221を介して無条件にプレゼンス信号の送信を開始する。待機期間によって、すでにセル601の送信を行っている接続されているコンピュータ装置101が、他のリンク221を介して、プレゼンス信号の伝達を開始したばかりのコンピュータ装置101に対してセル601の送信を開始することが可能になる。待機期間の後、コンピュータ装置101は送信機の数を検出するべくリンク221をチェックする。何も検出されなければ、コンピュータ装置101は接続されるコンピュータ装置101の起動を待つ状態S3 304にとどまる。複数のコンピュータ装置101が検出された場合、セル601を送信しているコンピュータ装置の数をチェックする。２つ以上のコンピュータ装置101がセル601を送信している場合、コンピュータ装置101はセル601が送信されている全てのリンクがセルロック状態S9 310になるまで待機する。接続されているコンピュータ装置101のうち0ないし１つのコンピュータ装置のみがセル601を送信している場合、または接続されているコンピュータ装置101のうち１つだけしか存在しないと検出された場合、コンピュータ装置101はセル601の送信を開始しなければならない。セル601の送信は、セル601を送信している他のコンピュータ装置101にできるだけ合うように開始される。従って、粗いタイミング調整処理および細かいタイミング調整処理は、コンピュータ装置101のリンク221をロック状態S9 310へと導く。

【０１２１】

ある実施形態におけるタイミング調整メカニズムのロバスト性によって、さらなる自由な状態のシーケンスが可能となる。

【０１２２】

全てのコンピュータ装置101は、あらかじめ特定の１つの所定の実施形態に合うように設定されることになっている。この実施形態はセルロック状態のネットワーク（410）アーキテクチャを有する。初期化シーケンスの間に、ある特定のコンフィギュレーションを解決するあるレベルのコンフィギュアビリティを追加することができる。

【０１２３】

粗いタイミング調整状態S7 308のリンク221上で利用されているセル601のコンテンツと、ロック状態S9 310に関連付けられる、完全動作状態のペイロード転送モードで利用されているコンテンツとは、一般的に異なるものである。しかし、異なる動作モードにあってもセル601コンテンツの不必要な相違は避けるべきである。

【０１２４】

リンク221が粗いタイミング調整状態S7 308にある間の、セル601のプロトコルの使用もしくはフォーマット割り当ては特に指定されない。フォーマットは、少なくとも以下の情報を送信することができるものであることができる。

・ コンピュータ装置101のジオグラフィック・アドレス106。ただし、ジオグラフィック・アドレス106毎のリンク221のマッピングが任意の実施形態において事前に定義されている場合を除く。

・ 接続されているコンピュータ装置101についてのジオグラフィック・アドレス106毎のリンク221の状態テーブル。

・ 接続されているコンピュータ装置101についてのジオグラフィック・アドレス106毎のリンク221の完全動作テーブル。

【０１２５】

上記の加えて、リンク221の細かいタイミング調整状態309では、セル開始オフセットの測定情報226が各セル601に送信される。

【０１２６】

多数の接続されたコンピュータ装置101が同時に起動した場合、これらのコンピュータ

装置101は、それぞれセル開始シンボル209を保有しようとするが、それらは、セル周期604に関連して予測不能なパターンに分散する。この点において、完全なフルメッシュ・トポロジー105内に配されないネットワークは考慮しなくてもよい。図３のステップに忠実に従ったとしても、セル開始タイミング227の予測不可能な分散パターンは起こり得る。

【０１２７】

ジオグラフィック・アドレス106の優先順位における、接続されているコンピュータ装置101を考慮した任意の優先順位づけスキームを適用することで、タイミング調整がなされる。例えば、接続されているコンピュータ装置101のうち最も小さいジオグラフィック・アドレス106を有するものをタイミング調整の参照先として検討する。各コンピュータ装置101は、各々のセル開始タイミング227を、最も小さいジオグラフィック・アドレス106を有するコンピュータ装置101のセル開始タイミング227に合わせる。この第１の調節処理は粗いタイミング調整（coarse alignment）と呼ばれる。

【０１２８】

セルロック状態310は、もともとリンク221の状態として特定された。セルロックリンク221を有するコンピュータ装置101は、セルロック状態221における全てのリンク221が同時に操作され互いに絡み合っているので、セルがロックされている、と言える。セルロックネットワーク410のネーミングもまた、セルロック状態の、普遍的な性質を反映している。

【０１２９】

すでにセルロック状態にあるネットワーク410に、コンピュータ装置101が追加される場合、すでにセルロック状態のコンピュータ装置101は、追加されるコンピュータ装置101の初期タイミング調整をサポートするにおいていかなる変更も行ってはならない。前述の優先順位付けスキームは、デッドロックの可能性が排除されるまで、粗いタイミング調整にのみ適用可能である。セルロック状態のネットワーク410に追加された、すでに動作しているコンピュータ装置101は、そのセル開始タイミングを、セルロック状態のコンピュータ装置101のすでに確立されているタイミングに合わせなければならない。追加されたコンピュータ装置101がセルロック状態に達した場合、ネットワーク410のセルロック状態を維持する処理を行っている他のコンピュータ装置101と同じように取り扱われることが望ましい。

【０１３０】

粗いタイミング調整のために、リンク遅延224は無視され、ローカルで生成されたセル開始シンボル209に対する、受信したセル開始シンボル219の、ローカルで測定されたオフセット226のみが考慮される。あるレベルのタイミング調整が達成される場合、より精度の高い方法に変更する必要がある。このより精度の高い方法をサポートするために、コンピュータ装置101はセル601における所定のシンボル位置で測定したオフセット・データ226を送信する必要がある。

【０１３１】

より上位レベルでの利用のための基盤を確立するために、セルロックネットワーク410上の全てリンク221においてセルロック状態を獲得することが必要であるが、まずは個々のリンク221がセルロック状態に達する。セルロックリンク221の独立したサブセットという現象を回避することが重要である。フルメッシュ・ネットワーク・トポロジー105において、これは、図３に示される状態シーケンスがフォローされる場合に回避される。

【０１３２】

複数の相互接続されているセルロックネットワーク410に基づくアーキテクチャにおいて、互いに同期していない２つのセルロックネットワーク410間に、追加のコンピュータ装置101を接続することは可能である。任意の順位付け構造が利用可能であるように、２つのセルロックネットワーク410が統合され、１つのセルロック環境を成立させるまで、任意の順位付け法を利用して、一方のセルロックネットワーク410におけるセルの送信が、アイドル・シンボル210の適用可能な適切なバリエーションを利用するようにしてもよい。セルロック環境が確立するまでは、２つのネットワーク410のリンクは成立できない

。
【０１３３】
接続されている全てのコンピュータ装置101へ送信が同時に行われることは、ネットワーク410における意図されたセルロックにとって、制限ではなく、イネーブラであるということを、理解しなければならない。
【０１３４】
ある実施形態において、接続されている任意の２つのコンピュータ装置101の最大オフセット229は、次の上位レベルのプロトコルが成立する際に決定され、また考慮される。
【０１３５】
コンピュータ装置101がリンク221のセルロック状態を宣言する場合、この状態情報を接続されている全てのコンピュータ装置101に送信する。リンク221によって接続されている双方のコンピュータ装置101がロック状態を宣言した場合、リンク221はペイロードの送信にも利用可能である。リンク221の全機能は、ロック状態でなければ利用できず、セル期間603毎に確認されなければならない。
【０１３６】
リンク221にエラーが発生した場合、それを検出した次のセル開始タイミング227を伴うコンピュータ装置101はリンク221の完全動作および、リンク221を開始する初期化シーケンスを無効にする。他のコンピュータ装置101に接続する他のリンク221がセルロック状態にあり、セル開始シンボル209に他のセル開始シンボル209が同時に適用される場合、リンク221はいくつかのセル周期604の間にセルロック状態を確立することができる。もしリンク221の１つのパス215のみにおいて失敗が起きたとしても、リンク221の反対方向のパスは次のアービトレーションが使用できない、ということを考慮に入れなければならない。
【０１３７】
健全なシステムにおいては、全てのリンク221はセルロック状態を短期間で達成しなければならない。セルロックネットワーク410を始動する時、全ての確立されたリンク221がセルロック状態を達成し、完全動作（full functionality）を表す信号を発信するまでは、ペイロード・トラフィックの開始を待機することが有用である。
【０１３８】
以下に、セルロック状態を維持する方法について述べる。
【０１３９】
リンク221のセルロック状態が確立すると、リンク221は、ペイロード・データ701を双方向に伝達することができる。
【０１４０】
セルロック状態である間、セル601はセル開始シンボルのオフセット測定情報226を、セル601における所定のシンボル位置で提供する必要がある。例えば、8ビットの符号付き整数はほとんどの実装で十分なはずである。
【０１４１】
セルロック状態にある全てのコンピュータ装置101は、タイミング維持処理に参加する。このことは、自分の測定に基づくセル開始オフセット・データ226および接続していてかつセルがロックされている（310）他のコンピュータ装置101から受信したセル開始オフセット測定データ226を利用することを意味し、セルロック状態にある各コンピュータ装置101において、全てのセル期間601において補正を行いうるための準備がなされていることを意味する。計算された数のアイドル・シンボル210は、現在のセル601の末尾シンボル220が送信されると直ちに適用されうる。適用されたアイドル・シンボル210の数は、定義された範囲内のアイドル・シンボルの数でなければならない。タイミングについての見直しが行われないのであれば、アイドル・シンボルのデフォルト数605が適用されなければならない。接続されている全てのコンピュータ装置101において適用されているアイドル・シンボル210の数の最小値と最大値との間の中央値は、（default-1）と（default+1）の範囲内に収まらなければならない。
【０１４２】

上述のコンセプトは、全てのコンピュータ装置101がそれぞれに直接リンク221を有する場合に適用可能である。もし直接リンク221が提供されないのであれば、セル開始タイミング227は参照チェーン越しに同期を計ることが可能である。しかし、参照チェーンが長ければ長いほど、セル開始タイミング227の総範囲も広くなってしまう。

【０１４３】

上述のある特定のケースが一般的なコンポーネントである場合、コンピュータ装置101は一方でリンク221を大きなセルロックネットワーク410に提供し、もう一方でリンク221を第２のセルロックネットワーク(S_CLN) 405に提供する。図４を参照すると、セルロックネットワーク410および405は、いかなる上位レベルのコミュニケーション・プロトコルにも依存せずに動作している。上位レベルのコミュニケーション・プロトコルのインターフェースは、第２のセルロックネットワーク405に接続されているコンピュータ装置101の別のプロトコル特有のアダプタに実装される。あるいは、第２のセルロックネットワーク405に接続されているコンピュータ装置101は、高パフォーマンス・ローカル・バスに接続可能であり、またI/Oデバイスを備えているので、I/Oインターフェースが物理的には見えることはない。

【０１４４】

図４はセルロックネットワーク410上で相互接続されているコンピュータ装置APP_A 101、APP_B 102、 APP_C 103およびAPP_D 104が第２のセルロックネットワーク405を使用可能ないくつかの異なる方法を示す。第２のセルロックネットワーク405はイーサネット、SAS等のI/Oインターフェースに接続するコンバータ・コンポーネント（CONV）402に接続してもよい。別の可能性として、統合コンバータおよびI/O機能部404からなるI/O機能を接続してもよい。I/O機能部404はローカル・バス、およびI/Oプロトコルを第２のセルロックネットワーク405のインターフェースに変換するコンバータ・コンポーネント402を接続する。プロトコル・コンバータ・コンポーネント402は標準プロトコルのうちの超高帯域幅をサポート可能であり、対応するデータキューは直接第２のセルロックネットワーク405に転送される。

【０１４５】

ゼロは、アイドル・シンボル210の数に許される最小の数でなくてはならない。適用されるアイドル・シンボル210の数は、シンボル・レート許容誤差の範囲内でリンク221間のタイミング調整を認めなければならないので、ある程度の余裕を含む必要がある。このことはまた、少なくとも１つのアイドル・シンボル210が現れる場合が多いことを保証する。そのため、アイドル・シンボルを介していくつかの制御情報を発信することが可能である。セルの最終シンボル220の後に送信されるアイドル・シンボルは特定のSKPシンボルである必要はないが、セル開始シンボル209を除く任意のシンボル222でなければならない。

【０１４６】

高精度のクロッキングを使用し、セル601毎のシンボル222の数がそれほど多くないのであれば、0をアイドル・シンボル210の最小数として、1をアイドル・シンボル210の最大数として指定することが可能である。この場合、アイドル・シンボルのデフォルト数605は0.5である。これが実現できれば、0または1のアイドル・シンボル210が交互に適用される。この方法はアイドル・シンボル210を余計な負荷なしに最小限に抑えることができる。

【０１４７】

どのような理由であれ、ある１つのコンピュータ装置101がセルロックネットワーク410のタイミング参照先であると宣言される必要がある。優先順位付け構造は、複数のコンピュータ装置101がタイミング参照先であると主張することを解決するのに利用される。タイミング参照先コンピュータ装置101は、常にデフォルト数605のアイドル・シンボルを送信する。あるコンピュータ装置101がタイミング参照先であると主張しているが、ネットワーク410のセルロック中のサブセットの一部ではないような場合に、セルロック中のコンピュータ装置101は、セルロック状態を解除することなくタイミング参照先にアプローチしなければならない。

【０１４８】

ここからは、外部クロックがセルロックネットワーク410を介してどのように届けられるのかについて述べる。

【０１４９】

以下は、外部クロック・ソースが利用可能となるケースについて簡単に述べたものである。個々のクロック配信経路を利用する代わりに、セルロックネットワーク410は接続されているコンピュータ装置101に外部クロック・ソースを提供する。

【０１５０】

特定の外部クロック・ソースに接続されているコンピュータ装置101は、位相ロックループ（PLL）を利用してローカル・クロックと外部クロック・ソースの同期を図る。特定の外部クロック・ソースに接続されているコンピュータ装置101は、それがローカル・クロックの元で作動している他のコンピュータ装置101よりも優先度が高いことを宣言する。もし、複数の外部クロック・ソースに接続されているコンピュータ装置が利用可能であれば、実装においてさらなる優先順位付けが必要となる。自分の優先順位が最も高いと同定しているコンピュータ装置101は、常に一定の数605のアイドル・シンボルを利用し、他のコンピュータ装置101はセルロック状態を達成し維持するために、最も優先順位が高いコンピュータ装置101を参照して、アイドル・シンボル210の数を調整する。

【０１５１】

例えば8kHzのデータ転送クロックをセルロックネットワーク410全体に配信し、相互接続による専用クロック配信を不要にするというアイディアが実現可能である。この目的のために構造をアレンジし、その結果、セル送信期間603および一定数のアイドル・シンボル607の時間の総和であるセル周期604は、外部クロックの周期の除数となる。クロックを配信するコンピュータ装置101はタイミング参照先であることを宣言しなければならず、また、セル601およびシンボル222の位置が外部クロックの次のエッジに一致することを特定しなければならない。あるいは、シンボル期間225と外部クロック周期の直接的な関係が利用可能であり、セル長606の制約を回避することができる。

【０１５２】

ここからは、情報ブロックを複数のセルの間でどのように発信するのかについて述べる。

【０１５３】

ここで述べられたセルロックネットワーク410の最下位ハードウェアプロトコルにで発信される必要がある、いくつかの変化しない情報もしくはあまり変化をしない情報がある。セル601のある所定の位置にある任意のシンボル222は、情報ブロックを逐次的に送信するために割り当てられる。発信される情報には、ジオグラフィック・アドレス106、グローバルユニーク識別子、製品・ベンダー情報、コンピュータ装置101に接続しているリンク221の状態コード等が含まれてもよい。

【０１５４】

この目的のために指定されるシンボル位置は、次のセル601における同一のシンボル位置において送信されるシンボルと共に、情報のパケットを送信するのに利用可能である。非データ・シンボルはパケットの開始を特定するのに必要とされる。

【０１５５】

ある実装例では、パケット長が可変であるか固定であるかを特定することができる。両方のケースにおいて、割り当てられるプロトコル、規約等は特定される必要がある。

【０１５６】

低レベル情報が複数の長さの等しいパケット、つまり複数の情報ブロックで送信される場合、これらの長さの等しいパケットは互いにタイミングを合わせてセルロックネットワーク410全域に送信されるほうがよりよい。これらの等長パケットの同期を図る簡単な方法は、開始前非データシンボル222を利用することである。コンピュータ装置101の１つがセル601内の開始前シンボルを送信するとき、次のセル602において、全てのコンピュータ装置101はパケット開始シンボルを送信し、パケットの送信を続ける。コンピュータ装置101が、パケット周期が許容するよりも長いパケット開始シンボルを持たないということを

認める場合、コンピュータ装置101は開始前シンボルおよびパケット開始シンボルの送信を介して、パケット通信を再開するための権限を与えられる。

【０１５７】

上述の同期パケット送信は、効率の良い集積回路間通信（Inter-Integrated Circuit（I2C)インター・インテグレイティド・サーキット（アイ・ツー・シー））エミュレーションをセルロックネットワーク410内に埋め込むことができる。各コンピュータ装置101は、同期パケット内の特定のシンボル位置によって、そのI2C信号ストリームをセル601あたり１バイトで送信する。全てのコンピュータ装置101は他のコンピュータ装置101からのI2C送信シンボルを受信する。複数のコードの内の１つが有効であると識別されると、同じデータがI2Cネットワーク越しに送信されていれば、その１つは優先権を得る。優先順位付けスキームをサポートするために、ビットは最上位ビットがはじめに送信されるよう考慮される。このタイプのアービトレーションを失っているコンピュータ装置101は、次の複数のセル602によって、「1111 1111」データコードを送信する。これは、コンピュータ装置101がI2Cプロトコル・エミュレーションにデータを送信する権利を得るまで続く。これはセルロックネットワーク410および当該ネットワークの同期パケットの同期の性質であり、このような簡易な方法でそれを実装することができる。I2Cのクロック伸長機能は、特定の非データ・シンボルの送信を介して追加可能である。この非データ・シンボルが１つのコンピュータ装置101から受信された場合には、他の全てコンピュータ装置101から同じセル601内に送信されたI2Cデータは無視されなければならない。

【０１５８】

以下に、ベーシックなシンボル・レートの倍数でリンク221を動作させるオプションについて述べる。

【０１５９】

実施形態の基本的なバージョンにおいては、コンピュータ装置101に接続している全てのリンク221は名目上は同じシンボル・レートで動作する。ネットワークケーブルもしくはバックプレーン・ルーティングは、大型システムにおいては、適用可能なシンボル・レートの制限要因となりやすい。接近しているコンピュータ装置101間のリンク221は、より高いシンボル・レートで動作できる。そこで、セルロックネットワーク410のベーシックなシンボル・レートの整数倍で、そのサブセットを動作させるオプションを説明する。

【０１６０】

シンボル・レートに適用される係数に対応して、ベーシック・シンボル・レートにおける１つのセル期間603に複数のセル601が送信されうる。高いレベルの互換性を維持するために、複数のシンボル・レートが利用され、ベーシック・シンボル・レートにおける各シンボル期間225の間に、対応する数のシンボル222が送信される。１つのシンボル位置が送信される各セル601のためにサポートされる。

【０１６１】

図５は複数の高速リンク221を介したインターリーブ送信の例を示す。リンクLNK_A 501はベーシック・シンボル・レートで動作するパス215で、シンボル・シーケンス「ABCDE」を送信する。リンクLNK_B 502は2倍のシンボル・レートで動作するパス215で、シンボル・シーケンス「ABCDE」および「abcde」をインターリーブ送信する。リンクLNK_C 503は3倍のシンボル・レートで操作されるパス215で、シンボル・シーケンス「ABCDE」「abcde」および「12345」をインターリーブ送信する。LNK_Cにおいて描かれるシンボル222のシーケンスにおいて、セル開始209は、ここではCOMシンボルである1/3長の504で識別され、ここではSKPを利用した1/3長の２つのフィルタ・シンボル507および508がその後に続く。このシーケンス503においては、ベーシック・シンボル・レートの１つのアイドル・シンボル位置が、３つの1/3長のフィルタ・シンボル505によって埋められる。一般的には、複数位置の最初のものを介して送信されるセル601は、全ての制御シンボルを提供し、追加のセルはデータのみを送信し、制御シンボル位置は割り当てられない。

【０１６２】

セルロックネットワーク410の主たる特徴は、全てのセル601がほぼ同時に送信されると

いうことである。言い換えれば、セルロックネットワーク410全域で送信されるセル228のオフセットには保証された限界がある。よく構成されたある実施形態は、シンボル期間225のセル・オフセット228をほんのわずかしか許さない。実施形態はどのようなセルロックネットワーク410にも適用され、上述の実施形態に従ってセルロックが確立する場合に限定されるものではない。

【０１６３】

以下は、全てのコンピュータ装置101がフルメッシュ・ネットワーク105によって相互接続され、各々がジオグラフィック・アドレス106によって識別される実施形態である。

【０１６４】

以下の実施形態は、ネットワーク参加者が自分のニーズを超えてデータの送受信を行い、また、他の参加者のデータの転送エージェント役を行い、その結果リンクを必要としない場合にそのリンクの活用のためのサービスを生み出す、というコンセプトに基づく。データ送信の集中制御は必要とされない。さらに、ベーシック・プロトコル・レベルでのフロー制御も必要とされない。伝送ルーティングはより長期間の情報に基づいてはならない。エラーチェックおよび再送も必要とされない。個々のネットワーク参加者のスタベーション（starvation）は回避されなければならない。優先レベルはオプションでサポートされる。１つのリンクが破損した場合でも、高可用性サポートによりサービスを続けることができる。稼働中のネットワークへの参加者の追加および削除がサポートされる。ルーティング制御は上位レベルのプロトコルに依存しない。ネットワークの運用やルーティング制御は完全に自動化され、ソフトウェアの全てのレイヤからは見ることができない。制御のオーバーヘッドが許容リミットを超えてはいけない。

【０１６５】

複数の実施形態において、送信予定のデータは出力キュー1404から利用可能であり、到着したデータは入力キュー1405に格納されうることが想定されている。両方のキューは実装された各リンク221に独立して存在する。

【０１６６】

フルメッシュ・ネットワーク105によって提供される広帯域幅の能力は、全てのコンピュータ装置101が絶えずデータを直接交換するような場合に活用されうる。実際の多くのネットワークは別の方法で利用されている。通常、あるコンピュータ装置101が他のコンピュータ装置101との間で高帯域幅を必要とする時間は短い場合も長い場合もあり、一方で、別のコンピュータ装置101は低帯域幅もしくは不定期のデータ交換しか必要としない。

【０１６７】

ある実施形態は、アイドル信号の伝達経路215の動的利用により、コンピュータ装置101間での広帯域データ転送能力を必要な時に提供する。これは、転送エージェント機能をコンピュータ装置101に追加することで達成可能である。

【０１６８】

他の複数のコンピュータ装置101間の送信において転送エージェント役として機能するコンピュータ装置101は、送信を補助するために、一時的にデータを格納することができる。ネットワーク410およびセル長606の最大サイズは実施形態に依存し、この要件は制限をするものではない。

【０１６９】

上述のコンセプト項目を解決するために、各アービトレーションラウンドのために、各コンピュータ装置101はその送信リクエストおよびその受信機性能を全ての接続されているコンピュータ装置101に等しく発信する、というソリューションが提示される。これによって、各コンピュータ装置101におけるアービトレーションを等しく実行することが可能になり、また、アービトレーション結果を配信する必要もなくなる。

【０１７０】

送信リクエストおよび受信機性能は、各コンピュータ装置101のジオグラフィック・アドレス106毎に関連付けられるシンボル位置を利用して発信される。

【０１７１】
以下は、ペイロード転送が提供可能となる、別のソリューションである。
【０１７２】
第1の転送ソリューションによれば、セル601の完全なデータ・セグメントはペイロード・データの最小単位として利用される。このソリューションは2ないし３つのセル周期を動的に変更することで展開される。第１のパイプラインステージにおいて、送信リクエストおよび受信機性能が発信され、第２のパイプラインステージにおいて、直接接続パス215を介する送信が、転送エージェント役のコンピュータ装置101への送信と一緒に実行される。第３のパイプラインステージにおいて、転送エージェント役のコンピュータ装置101は、格納したデータをターゲットのコンピュータ装置101に送信する。このソリューションは、データ送信のために、セルのコンテンツの最大量を利用する。送信データを転送するために転送エージェント役のコンピュータ装置101を介して２番目のセル期間603を利用することで、ペイロード・データに対する追加遅延が発生する。パイプライン処理の結果、リンク221の予約情報は、セルロックネットワーク410に新規の参加する可能性のある者に通知される必要がある。通知されなければ、そのような可能性のある者は、セルロックネットワーク410に参加することができないか、または、障害発生の状況から回復できない。直接接続パス215は、それが既に転送データの送信に割り当てられている場合は、接続されているコンピュータ装置101の直接通信の送信のための新たな要求にすぐに対応できるわけではない。このソリューションは、各サポートされたリンク221のために、転送エージェント役のコンピュータ装置601において、ペイロード・データの完全なセル601の機能のための一時的な格納を必要とする。
【０１７３】
第２の転送ソリューションによれば、セル601は複数の等長ペイロード・データ・セクタに分割される。分割されたものは、個々にルーティングされ、セルロックネットワーク410を通って転送される。データ・セクタは同じセル期間603に、転送エージェントを介して次のデータ・セクタとして転送される。末尾のデータ・セクタは、転送エージェントにデータを送信するのには使用できない。第２の転送ソリューションは、セルのシーケンスを介してアービトレーションする必要がなくなる。前述の第１のソリューションと比較すると、リソース割り当ての粒度が細かければ細かいほど、転送エージェント役のコンピュータ装置101に対するストレージ要求は少なくなりうる。一方で、要求および性能の送信のためのオーバーヘッドは大きくなり、アービトレーション処理はより複雑になるかもしれない。さらに、アービトレーション粒度および応答時間はセル周期604より短くはない。
【０１７４】
第３の転送ソリューションは、同じセル期間603内でペイロード転送が提供される。２つのセルフォーマットが利用される。１つはペイロード・データ範囲を伴うCF1 702であり、先に開始する。もう１つはペイロード・データ範囲を伴うCF2 703であり、遅れて開始する。第３の転送ソリューションは、実装しやすく、非常に効率的である。ペイロード・データ・セグメント704は先の２つの転送によるソリューションで送信されるのに比べて、シンボル222の数が少ない。データ送信に利用することはできないセル601の位置も、実は無駄にならない。そのような位置は、IPネットワークなどの二次通信構造を伝達するために割り当てることができる。転送エージェント役のコンピュータ装置101では、いくつかのシンボル222のためだけにしか一時的データバッファを必要としないという利点がある。データバッファのサイズは従って、セル長606に依存しない。第３の転送ソリューションは、短いパイプラインを利用する。送信リクエストおよび受信機性能は１つのセル期間603内で発信され、アービトレーションはこのセル期間603が終了する前に行われ、さらに、全ての結果データの送信は、次のセル期間608内で行われる。
【０１７５】
転送による３つのソリューション全てに共通するのは、要求情報および性能情報がセル601の送信中に所定のシンボル位置で送信される、ということである。たとえ複雑なシナ

リオであったとしても、アービトレーションアルゴリズムを実行するのに十分な時間が残るように、シンボル位置を割り当てることが可能である。制御情報にどのシンボル位置が割り当てられても、データ送信は制御通信とインターリーブされる。接続されている全てのコンピュータ装置から受信したリクエストおよび機能情報は、次のセル期間608中のデータ送信に利用されるリソースの利用について決定する、アービトレーションアルゴリズムのパラメータを備える。

【０１７６】

以下は、第３の転送ソリューションに基づく更に詳細な実施形態である。

【０１７７】

図７は２つのセルフォーマットCF1 702およびCF2 703の略図である。セル・ペイロード701はデータ・セグメント(D) 704と、このプロトコル・レイヤに関して使用できない別のセグメントである廃棄セグメント(W) 705に分割される。廃棄セグメントW 705は、データ・セグメントD 704より著しく小さい。セルフォーマットCF1 702はDWシーケンスを含み、一方で、セルフォーマットCF2 703はWDシーケンスを含む。データ・セグメントD 704はセルロックネットワーク410上で最小構成単位のデータを１つ伝達する。送信機がセルフォーマットCF1 702を使用する場合、廃棄セグメントW 705が適宜選択され、転送エージェント役のコンピュータ装置101は、セルフォーマットCF2 703を使用している同じセル期間603内に、セルフォーマットCF1 702で受信したセル601のデータ・セグメントD 704を再送することができる。

【０１７８】

必要なフォワーディング・シフト706は、セル開始228の可能な最大オフセットから計算可能である。この最大オフセットは、複数のコンピュータ装置101のロックされたサブセット、リンク遅延の最大値224、セル長606およびシンボル・レート許容誤差に関係する。オーバーヘッドの固定的な処理を、転送エージェント役のコンピュータ装置101に追加可能である。転送エージェント役のコンピュータ装置101は、データのフェッチ、ローカル・シンボル・クロックの利用および他のシンボル・ストリームを多重化してデータを送信することを必要とする。この計算のための上記考慮すべき点は、フォワーディング・シフトがセル601の先頭だけではなく、セル期間603中に動作しなければならない点である。8ビットから10ビットの符号化スキームが使用される場合には、シンボル222の毎秒500メガシンボルのシンボル・レートに対応する5 Gb/sの信号レートおよび最大50ppmのクロック許容誤差を用いると、要求されたフォワーディング・シフト706はおよそ25ないし30シンボル期間となる。実際の実装はシミュレーションを通じて、必要なフォワーディング・シフト706を定義し実証しなければならない。廃棄セグメント705内のシンボル222が別の用途に割り当てられないのであれば、1000シンボル／セル606のセル・ベース・ネットワークは、ペイロード転送方法によって、3%近い帯域幅の損失被ることになる。

【０１７９】

セル601内に廃棄セグメントW705がいつも存在するため、実装においては、廃棄セグメントの有用な目的への割り当てが望まれる。例えば、リンク221によって接続されているコンピュータ装置101間の直接的な通信のための、安定して存在するデータチャネルとして、廃棄セグメントW705が割り当てられうる。このことは、高帯域幅のセルロックネットワーク410に負荷をかけることなく、低帯域幅通信を可能にする。

【０１８０】

ペイロード転送による第３のソリューションは、次のセル601の送信において、２ステージパイプラインとして稼働する。第１のパイプラインステージにおいて、リクエストおよび性能情報を含む制御シンボル805が発信され、第２のパイプラインステージにおいて、データ・セグメント704が送信される。

【０１８１】

セルフォーマットの組み合わせは、アービトレーション結果によってリカバーされるため、送信されるセル601内で、２つのセルフォーマットCF1 702 およびCF2 703を区別する必要はない。各セグメントの長さは、任意の実施形態において事前に定義されるため、デ

ータ・セグメントD704および廃棄セグメントW705との間の境界を特定するシンボル位置を無駄にする必要はない。

【０１８２】

制御シンボル805には、セル同期およびセルロックに必要とされるシンボル222や、アービトレーションに関する情報を伝達するシンボル222が含まれるが、そのような制御シンボル805のためのセル601内のシンボル位置割り当ては、使用されるセルフォーマットの種類にかかわらず、固定したままにしておかなければならない。固定しておかないと、不必要に複雑化してしまう。

【０１８３】

図８は、セル601における実装の複雑さを制限するための追加の構成を示す。この例において、C1～C13として特定される制御シンボル805は、セル601の均一なグリッド位置に配され、残りの位置は、サブセグメントD1～D11およびW1～W2として分類されるペイロード・データ701が使用可能である。制御シンボル805が任意の位置ではなく規則的なパターンに配置されれば、実装は極めてシンプルなものとなる。Wサブセグメント705およびDサブセグメント704の長さは、前述の規則的なパターンに一致するように指定されている。この構造により、転送エージェント役のコンピュータ装置101がデータD1～D11のサブセグメント803をセルフォーマットCF1 702で受信し、セルフォーマットCF2 703で再送する場合に、各データシンボル位置803は一定量だけシフトする。制御シンボル805も転送される必要がある場合、それらは同じシフトに従うように割り当てられる。セル601が非常に長い場合には、相当多くの制御シンボル位置が生じる可能性がある。この場合、多くの不必要な制御シンボル位置はデータシンボルに再割り当てする可能性がある。しかし、この再割り当ては、CF1 702 および CF2 703に対してそれぞれ異なる。この割り当てについての例が、特定の実施形態を示す図１６の表に含まれる。

【０１８４】

図９はペイロード転送の第３のソリューションの簡単な例である。例示システム901の概略図には、コンピュータ装置APP_A 101、 APP_B 102およびAPP_C 103がフルメッシュ相互接続リンク221とともに示される。コンピュータ装置APP_A 101およびAPP_B 102の間のリンク221で構成されるデータ送信パスは、それぞれAB、BAとして識別され、パスAB215はAPP_A 101からAPP_B 102に向かうデータ伝送を、パスBA216はAPP_B 102からAPP_A 101に向かうデータ伝送を示す。APP_A 101およびAPP_C 103間のパスはそれぞれAC 908とCA 909、APP_B 102およびAPP_C 103間のパスはそれぞれBC 910とCB 911である。このシンプルな例では、コンピュータ装置APP_A 101は２つのセル601の合計のペイロード・データ704をコンピュータ装置APP_B 102に送信し、これら２つのペイロード・データ704のセグメントを、AB1およびAB2と識別する。図９はペイロード・データ704の２つのセグメントが、１セル期間603内でどのように送信されるかを示す。AB1はセルフォーマットCF1 702を使用し、ダイレクト・パスAB215を介して送信される。APP_A 101は送信機902、APP_B 102は受信機904として示される。AB2は、パスがすでにAB1の送信に割り当てられているため、ダイレクト・パスAB215を介して送信することができない。パスAC 908およびCB 910は割り当てられていないので、これらがペイロード転送に割り当て可能である。コンピュータ装置APP_A 101はAB2コンテンツを、ACパス908を介してコンピュータ装置APP_C 103にセルフォーマットCF1 702で送信する。コンピュータ装置APP_A 101は送信機903、コンピュータ装置APP_C 103は転送エージェント役受信機906として示される。転送エージェント役のコンピュータ装置APP_C 103はAB2コンテンツを、CBパス911を介してセルフォーマットCF2 703で再送する。APP_C 103は送信転送エージェント役907として、APP_B 102は受信機905として示される。コンピュータ装置APP_B 102は、パスAB215およびCB911の両方から受信したペイロード・データ704をその入力キューにコンピュータ装置APP_A 101から受信したデータのために保存する。

【０１８５】

図１０は、セルロックネットワーク・オペレーションの複雑な例、とりわけ、ペイロード転送メカニズムについて示す。３つのコンピュータ装置APP_A 101、 APP_B 102およびA

PP_C 103がフルメッシュ相互接続901を構成するリンク221とともに示される。５セル期間603 CP_1、 CP_2、 CP_3、 CP_4およびCP_5のシーケンス1002の例を示す。各コンピュータ装置の各セル期間1002の出力キューコンテンツ1010が、APP_A 101のものをEQ_A 1003として、APP_B 102のものをEQ_B 1008として、また、APP_C 103のものをEQ_C 1009として示す。各コンピュータ装置101はそれぞれ２つの出力キューを有しており、それらはリンク221ごとの出力キューである。出力キューコンテンツは、図９の規則に従って識別される。各セル期間603に送信される出力キューコンテンツは、太字で区別されている。図１０の送信割り当てテーブルは、出力キューコンテンツ1010の各セル期間CP_1～CP_5 1002のデータ送信パスへの割り当てを示す。左揃えのテーブルコンテンツ1006はセルフォーマットCF1 702での送信を示し、右揃えのテーブルコンテンツ1007はセルフォーマットCF2 703での送信を示す。一部のデータ転送パスは使用されておらず1007、この例において、これらがダイレクト送信に利用可能な割り当てであり、小さなフォントで示されている。

【０１８６】

図１０に示す送信シーケンスは、各セル期間603について述べたものである。

【０１８７】

第１のセル期間CP_1において、コンテンツアイテム1010 AB1、AB2、AB3、BA1、BA2、BC1およびCA1は、アービトレーションの実行のために状況が凍結されている場合に、各出力キューにおいて利用可能である。

【０１８８】

第２のセル期間CP_2において、AB1、BA1、BC1、およびCA1は各々ダイレクト接続パスAB 215、BA 216、BC 910およびCA 909を介して送信される。AB2は転送エージェント役のAPP_C 103を介して送信される。これはつまり、CF1 702の送信にパスAC 908が使用され、CF2 703の送信にはパスCB 911が使用される、ということである。また、保留されている送信リクエストAB3およびBA2は、第２のセル期間CP_2においては達成されることはない。これらおよび細字で示される出力キューコンテンツ1010のサブセットである追加のアイテムAC1、AC2およびBA3は、アービタによって、次のセル期間603の割り当てを決定するために利用される。

【０１８９】

第３のセル期間CP_3において、アイテムAB3、AC1およびBA2は、各々ダイレクト接続パスAB 215、AC 908およびBA 216を介して送信される。BA3は転送エージェント役のAPP_C 103を介して、パスBC 910およびCA 909を利用し、それぞれセルフォーマットCF1 702 およびCF2 703で送信される。AC2の送信リクエストは第３のセル期間CP_3においては実行されることはない。AC2および追加のアイテムBA4およびBA5は、次のアービトレーションのためのデータ送信リクエスト入力アイテムである。

【０１９０】

第４のセル期間CP_4において、アイテムAC2およびBA4は、各々ダイレクト接続パスAC 908およびBA 216を介して送信される。BA5は転送エージェント役のAPP_C 103を介して、パスBC 910およびCA 909を利用し、それぞれセルフォーマットCF1 702 およびCF2 703で送信される。AC3の送信リクエストは第４のセル期間CP_4においては実行されることはない。AC4、CA2およびCA3は追加アイテムである。

【０１９１】

第５のセル期間CP_5において、AC3 およびCA2は、各々ダイレクト接続パスAC 908およびCA 909を介して送信される。AC4はパスAB 215を介して転送エージェント役であるAPP_B 102に向けて送信され、さらにAC4はパスBC 910を介してAPP_C 103に転送する。CA3はセルフォーマットCF1 702でパスCB 911を介して転送エージェント役であるAPP_B 102にむけて送信され、さらにCA3はセルフォーマットCF2 703でパスBA 216を介してAPP_A 101にむけて再送される。

【０１９２】

以下に、アービトレーションアルゴリズムについてより詳細に説明する。

【０１９３】

アービトレーションの対象は、セル・ペイロードのデータ・セグメント704の一部において、個別にサポートされているリンク221のためのプロトコル非依存型出力キュー（protocol agnostic egress queues：PAEQ）で利用可能なデータである。データがPAEQ 1404において利用可能ならば、必要とされているセル601の数に従って、アービトレーションを目的としてリクエストが生成される。アービタ1401のタスクは、データ・パス215を割り当ててダイレクト接続パス215を介してこれらのデータ・セグメント704を配信することである。または、転送エージェント役のコンピュータ装置101およびデータ・パス215を介して、追加データを転送エージェントにむけて/から配信する。フルメッシュの、セルロックネットワーク410において、多くのデータ・パス215は常に個々の直接接続されているコンピュータ装置間でのデータ送信のために必要とされているわけではないので、多くのデータ・パス215はペイロード転送に利用可能である、ということが仮定される。

【０１９４】

本発明の複数の実施形態は、送信リクエストおよび受信機の可用性と性能に応じた、データ送信パスの割り当てのためのある特定のアービトレーションアルゴリズムに限定されるものではない。発信される情報セットのコンテンツまたはフォーマットは、特定の何かに制限されるものではない。

【０１９５】

提案されたアービトレーションメカニズムもしくは処理は、完全なデータ送信要件および受信機の性能情報が各コンピュータ装置101によって他のコンピュータ装置101に発信される状況を基にしており、複数のコンピュータ装置101がアービトレーション処理をローカルで実行する場合、それらはデータ送信リソースの割り当てのために同時かつ独立して同じ結果となる。このことは集中制御もしくは集中タイミング調整なしに起こる。アービトレーション後、データは、ローカルで利用可能なアービトレーション結果によって制御された、セルロックネットワーク410全体にルーティングされる。

【０１９６】

アービトレーションの結果による送信パスの割り当ては、送信パス割り当てが送信リクエストを満たす結果として得られたものなのか、あるいはアービトレーションアルゴリズムの別の処理で得られたものなのかにかかわらず、１セル期間603中は有効である。その結果、データが出力キューPAEQ 1404に出現すると送信パス215が直ちに利用可能である、ということが起こりうる。

【０１９７】

受信機の性能情報はいくつかの説明を必要とする。ある特定の数のセル・データ・セグメント・コンテンツのために利用可能なバッファという意味での受信機の機能情報は、このコンテクストにおいては有用でない。受信機は莫大な数のセル・データ・セグメント・コンテンツ704のためのストレージを有することになっている。また、受信機は受信したセル・データ・セグメント・コンテンツ704を異なる上位レベルのプロトコルの個別の出力キュー1306に供給可能となっている。受信機のビルディング・ブロックは、セルロックネットワーク410が配信した任意の量のセル・データ・セグメント704を受け入れることができるはずである。データフローが停止または絞られる必要がある場合、データフローは上位プロトコル・レイヤのフロー制御メカニズムを介して通信することができる。これはセルロックネットワーク410をコンジットとして動作させることを考慮するというコンセプトに完全に沿ったものである。ここで、コンジットはフロー制御を提供しない。万一、他の手段が受信機の役に立たない場合、下部構造が、受信データを取得することのできない入力キューに向けられたデータをドロップすることで、受信機のバッファを解放するための教示された方法を提供することができる。それでもなお、バッファの可用性情報を、アービトレーション関連情報に追加することおよびそれに応じて利用することは確実に可能である。利用可能な受信機を有しているコンピュータ装置101のジオグラフィック・アドレス106のリストが提供される。このリストは、完全動作中のリンク221のリストと同一であってもよい。

【０１９８】

データ送信パス215が、シンボル期間225あたり複数個のシンボル分の帯域幅を提供するケースにおいては、アービトレーションや、そのサポート中に発信されたデータセットは、それに応じて強化されなければならない。

【０１９９】

より高い帯域幅を提供する２つのアプローチがある。

・　複数のシンボル・レートを実装する

・　複数の並列なレーンを備えるリンクを実装する

【０２００】

これら２つのアプローチは組み合わされてもよい。

【０２０１】

全ての制御情報はリンク毎に１つ定義されているセルを介して伝達可能である。例えば、複数の並列なレーンの第１のレーン上における複数のインターリーブ送信の第１の送信を介して、制御情報を送信する。他の送信シーケンスにおける制御シンボル位置は、この発明によっては割り当てられていない。

【０２０２】

コンピュータ装置101は、フルメッシュ・ネットワーク上のリンク221におけるクラスター内部サブセット1103のための高帯域幅相互接続を伴うクラスター1102にグループ化されることができる。

【０２０３】

図１１は６台のコンピュータ装置APP_A 101、APP_B 102、 APP_C 103、 APP_D 104、 APP_E 1105、およびAPP_F 1106からなるクラスター・ネットワーク構造の例を示す。コンピュータ装置APP_A 101およびAPP_B 102はクラスターCLUS_AB 1102を、PP_C 103およびAPP_D 104はクラスターCLUS_CD 1107を、APP_E 1105およびAPP_F 1106はクラスターCLUS_EF 1108をそれぞれ形成する。クラスター1102のコンポーネントは３つの並列レーン1103として実装されるリンク221により接続される。複数のリンク221は、別のクラスター1102のコンピュータ装置101に接続しており、単レーンのリンク221の接続1104とともに示される。

【０２０４】

このクラスター化アーキテクチャは、クラスター1102を形成しているコンピュータ装置101同士が近接しているため、コストがかからず、または設置場所に関するの不利益を被ることなく、実装可能である。コンピュータ装置101同士が近接していることは、2倍もしくは3倍のシンボル・レート設定を可能にし、また、追加レーンの実装も可能にする。クラスター1102の内部接続1103のための帯域幅は、できるだけ広くするべきである。

【０２０５】

クラスター・アーキテクチャを最大限に活用するために、アービトレーションは以下のパス割り当てステップを順次適用することで強化される。各ステップは網羅的にリソースを割り当てる。

・　ダイレクト接続パスが、ダイレクト送信に割り当てられる

・　クラスター内部送信が、クラスター内部の転送エージェント役のコンピュータ装置に割り当てられる

・　クラスター・トゥ・クラスター送信のための転送エージェント役のコンピュータ装置が、送信元コンピュータ装置のクラスター内で割り当てられる

・　クラスター・トゥ・クラスター送信のための転送エージェント役のコンピュータ装置が、送信先コンピュータ装置のクラスター内で割り当てられる

・　クラスター・トゥ・クラスター送信のための転送エージェント役のコンピュータ装置が、他のクラスター内で割り当てられる

・　クラスター内部送信のための転送エージェント役のコンピュータ装置が、ネットワークの全体に亘って割り当てられる。

【０２０６】

アービトレーションは、保留データ送信リクエストを送信パス215に割り当てなければ

ならない。それによって、フルメッシュ・トポロジーは最適に利用されるが、それと同時にその接続がブロックされていないことが確実でなければならない。これら２つのコンセプトは、PAEQ 1404の第１の送信を各送信先コンピュータ装置101に各ダイレクト接続パス215を介して割り当てることで、提供される。追加データの送信は転送エージェント役のコンピュータ装置101を介して接続パス215を利用して行われる。これはダイレクト送信パス215の割り当てがなされた後にも利用されていない状態を保つ。平等な機会のために、ラウンドロビン分布の変形が用いられる。アービトレーションの時間は、最後のアービトレーション関連情報の受信と次のセル期間603におけるペイロード・データ701の先頭シンボル222との間で利用可能なシンボル期間225の数に制約されることに留意する必要がある。

【０２０７】

可用性情報は全てのコンピュータ装置101に同じように発信され、アービトレーションが各コンピュータ装置101において同じ結果が生成されることを可能にする。リンク221は対称的に処理されなければならない。一方のパス215が利用不可能であることが宣言される場合、リンク221の反対向きのパスであるもう一方のパス216は同じように、利用不可能である、と処理される。

【０２０８】

アービトレーションによって生成されたルーティング割り当てを利用し、セルロックネットワーク401を介してデータを送信する順番を定める規則を確立することが可能である。これはアービトレーションアルゴリズムとは独立に行うことができる。これが必要とされるのは、複数のルートが同時に、同じ目的のために利用可能になるためである。上記規則は、受信機が正しい順番でデータ・ストリームを再構成することを可能にするための、送信機によるデータ・セグメントの指定方法のシーケンスを制御する。規則の例を後に示す。

【０２０９】

以下に、コンピュータ装置(101)のネットワーク・インターフェース（network interface: NWIF）ブロック構築におけるシンボル入力ストリーム処理を示す。

【０２１０】

すでに述べてきたように、各コンピュータ装置(101)は独立したローカル・クロックを有する。ローカル・クロックはシンボル・レートをローカルで生成するためのソースであり、シンボル222の送信に使用され、また、コンピュータ装置(101)の他の構造によっても使用可能である。

【０２１１】

各入力シンボル・ストリームは、送信するコンピュータ装置101のシンボル・レートで他のコンピュータ装置(101)に到着する。各入力シンボル・レートは、受信するコンピュータ装置(101)自身のシンボル・レートよりわずかに低かったりわずかに高かったりしてもよい。この状態は、シンボル受信機アーキテクチャ内で解決可能である。

【０２１２】

シンボル入力部の正に入口のハードウェア構造は、各リンク２２１毎に独立に入力シンボル・レートで動作する。PLLは入力信号からのクロックをリカバーするために使われる。シンボル入力クロックは、例えば適正な分周機でもってリカバーされてもよい。

【０２１３】

わずかに異なる内部シンボル・レートにより、入力シンボル223の二重の読み出しまたは入力シンボル223の削除が起こりうるため、入力シンボル223を、ローカル・シンボル・レートを使用するローカル構造に直接渡すことは不可能である。

【０２１４】

図１２は、シンボル入力ストリームをサポートするためのある独特な非同期先入れ先出し(FIFO: first-in-first-out)メモリ1299の回路図を示す。この回路図は回路全体を示すことを意図したものではない。ここに述べる複数の実施形態を理解するために関連する部分のみを示している。

【０２１５】
細かいタイミング調整プロセス309の間、およびセルロック状態310の間、機能性の維持は次の通りである：入力シンボル・レートで入力クロック(IN_CLK)を利用して、入力シンボル223を非同期FIFO 1299に書き込む。シンボルはローカル・シンボル・レートでローカル・クロック(L_CLK)で読み出される。書き込みアドレス(ADDR_W)がカウンタ(CNT)1202によって生成される。CNT1202は次の書き込みアドレスを各IN_CLKクロックで展開する。読み出しアドレス(ADDR_R)は事前設定可能なカウンタ(P_CNT)1203によって生成される。P_CNT1203は次の読み出しアドレスを各L_CLKクロックで展開する。CNT 1202およびP_CNT 1203はこのオペレーションモードの間は継続して有効である。次のアドレスが生成される場合、CNT 1202およびP_CNT 1203は、グレイ・コードと呼ばれるバイナリ符号化を反映した、同一の循環的アドレス・シーケンスを生成する。IN_CLKクロック毎に、シンボル入力(SYM_IN)を介して利用可能な入力シンボル223は、アドレスADDR_Wのデュアルポート・ストレージ配列(DPR)1201に格納される。SYM_INで利用可能なシンボル223は、符号化ビルディング・ブロック（DECD）で、受信されたセル開始シンボル219であるかどうかをチェックされる。もしセル開始シンボルであれば、カンマ検出(COM_DET)がアクティブになる。DECD 1204からのアクティブなCOM_DET信号は、レジスタ(REG)1205が、セル開始シンボル219がDPR 1201内の格納されているアドレスである現在のADDR_Wを格納できるようにする。格納されたADDR_W値の後のいくつかのL_CLK期間は、REG 1223へ転送される。DECD 1204からのアクティブなCOM_DET信号はまた、モノフロップ・カウンタ(MF_CNT) 1206のトリガとしても利用される。MF_CNT 1206のアウトプットはフリップフロップ（FF）1207および1208によってL_CLKに同期され、受信したセル開始シンボル219が検出されたという検出アウトプット信号（DET）を生成する。信号DETは、MF_CNT 1206がリセットもしくはモノフロップ・カウンタが終了したときに、解除される。ADDR_W値に同期しているFF 1210および1211と、ADDR_W値を遅延させる別のFF1212および1213と、L_CLKに同期するバイナリ数LVLを生成する組み合わせロジック（DIFF）1209とで構成されるコンパレータ構造が形成される。このコンパレータ構造を利用して、FIFO 1299に現時点で格納されているシンボル223の数を決定することが可能である。LVLはFIFO 1299において現時点で利用されている格納位置の数を反映する。L_COMはFIFO 1299への入力信号である。L_COMはセル開始シンボルCOMが送信される時に、L_CLK期間でパルスを伝達する。シフトレジスタ（SHFT）1214はこのパルスをL_CLK期間に遅延させる。L_CLK期間内に、転送されたセル・コンテンツ704がフォーマットCF2 703で開始される。遅延されたパルスは、P_CNT 1203のプリセットを、REG 1223から受信したセル開始シンボル219の格納されたアドレスに読み込ませることのできる適切な信号である。ビルディング・ブロックの次のグループは、送信したセル開始シンボル209および受信したセル開始シンボル219の間の、測定されたオフセット(OFFS) 226の値を生成する。CNT 1221はオフセット測定のためのシンボル期間225をカウントする。セル601の完全な長さ606はカウント可能でなければならない。CNT 1221はリセットおよび有効化するために同期入力を行う。FF 1219およびゲート1220はアクティブなDET信号上にパルスを生成する。このパルスは設定可能なFF 1217を追加で設定し、これはゲート1218と一緒にCNT1221のイネーブル入力（EN）をアクティブな状態に保持する。SHFT 1214の出力信号は、FF 1215およびFF 1216を介してさらに遅延され、遅延したパルスはFF 1216のアウトプットに到達し、該パルスはCNT 1221のイネーブル入力を無効にする。その後、CNT 1221のアウトプットは次の受信したセル開始シンボル219が検出されるまで安定している。CNT 1221の安定した状態のアウトプットはセル開始オフセット226と固定の関係である。OFFS 226の結果は補正器(CORR) 1222で計算され、入力セル開始シンボル219および出力セル開始シンボル209の間で測定されたオフセット226に対応するOFFSバイナリコードへ提供される。CORR 1222における補正は、OFFS226が0であるという結果が確実なものであるケースについて、つまり、コンピュータ装置101のセル開始シンボル出力をセル開始シンボル入力として受信する可能性のあるケースについて、実装に応じて定数を減算する。CNT1221が有効でない場合にOFFSアウトプット値のみが有効である、ということを考慮する必要がある。

【０２１６】
DPR 1201における必要なバッファの程度は、シンボル・レート許容誤差、セル長606および最大リンク遅延224に依存する。FIFO 1299は、入力シンボル・レートがローカル・シンボル・レートよりも速い場合、１セル期間603中に余剰シンボル223を格納でなければならない。FIFO 1299は、１セル期間603以内にL_CLKでの継続的なシンボル・ストリームを保証するために、入力シンボル・ストリーム中の十分な数のシンボル223を格納でなければならない。非同期FIFO 1299のバッファ許容量は、クロック許容誤差により引き起こされる変動に加え、セルロックネットワーク410のサポートされた相互接続の最短遅延と最長遅延の範囲をカバーする必要がある。各セルロック状態のネットワーク410リンク221入力は、個々に独立してこの特別な非同期FIFO 1299を必要とする。
【０２１７】
入力シンボル223は、受信した通りのシーケンスで、FIFO 1299に書き込まれる。セル601の読み出しは、各入力シンボル・ストリームとも、受信したセル開始シンボル219から同時に開始される。ローカル・アーキテクチャは複数の非同期入力FIFO 1299を、それぞれそのように読み込む。そのため、各非同期FIFO 1299において、ローカル側はADDR_Rを受信したセル開始シンボル209を格納したアドレスに設定しなければならない。セル601の読み出し開始の正しい時間は重要である。セル601の一部である受信シンボル223を読み出す間、非同期FIFO 1299はローカル・クロックL_CLKによるクロックで、SYM_RECへ逐次的に読み出される。
【０２１８】
セルロックネットワーク410の複数の非同期FIFO 1299は全て、複数の非同期FIFO 1299からのセル開始シンボル209の読み出しと同時に開始される。ペイロード転送メカニズムをサポートするために、セル開始シンボル209の同時読み出しは同時に起こるだけではなく、現在のコンピュータ装置101によりセル開始シンボル209のセルロックネットワーク410への送信後、所定の回数のL_CLK期間においても、起こりうる。
【０２１９】
送信の遅延により、セル601の最後に受信したシンボル223は出力構造がすでに後続セル602の送信を開始した後に入力構造によりフェッチされる。それゆえ、受信したシンボル223のルーティング制御は、セル601の最後に受信したシンボル223が正確に送信されるまで維持されなければならない。
【０２２０】
ある特定の実装において、非同期FIFO 1299およびそのストレージDPR 1201はデータのための8ビットバイトだけではなく非データシンボル223も格納できるように構築されてもよい。ｚ
【０２２１】
上述のメカニズムはリンク221がロック状態である場合にも適用される。
【０２２２】
図１２の示すコンセプトは、ロック状態に加え、図３の細かいタイミング調整状態および粗いタイミング調整状態をサポートするという追加事項を含む。DECD 1204ビルディング・ブロックはセル601における制御シンボル805専用の、および規則的なパターンのシンボル位置を特定する構造を含む。入力信号LOCKEDは、このインターフェースに対して、セルロック状態，粗いタイミング調整状態，細かいタイミング調整状態，のいずれが有効であるかを制御する。セルロック状態がアクティブなLOCKED制御入力を用いて伝達される場合、イネーブル信号IN_CNTおよびL_CNTは常に、またはほぼ常にアクティブである。LOCKEDが解除された場合、IN_CNT出力は制御シンボル位置におけるIN_CLK周期のためだけにCNT 1202を有効にする。DECD 1204はローカル側のL_CNTのために対応する制御を生成する。これにより、制御シンボル位置のみにおいてP_CNT 1203を前に進めることができる。DECD 1204はLOCKED 入力に対応するIN_CNTおよびL_CNT制御のアクティブ化および非アクティブ化を確かに処理可能であるので、モードの切り替わりはセルの境界で発生する。DPR 1201のサイズおよびそのアドレス指定構造は、セル601の制御シンボル805の数よりも多いス

トレージを提供することで、このケースをサポートする。粗いタイミング調整は、クロックが合わないことに起因して、読み出しの重複や脱落が生じる可能性を含んでいる。しかし、このような結果は、セル601全体についてのみ起こるべきであって、受信したセル601内のシンボル223について起こるべきではない。DPR 1201はしばしば２つのセル開始シンボル209を含む。一方で、REG 1205およびREG 1223は両方のアドレスを格納し、REG 1223はP_CNT 1203をあらかじめロードするためにその１つを提供する。

【０２２３】

セルロックネットワーク410は複数の接続されたコンピュータ装置101の協調した動作に基づく。実装されるコンポーネントが正確に作動するということもまた仮定されている。実際の導入においては障害が発生するということも起こりうる。長期間の連続使用を提供することを目的とした大規模な導入においては、単独のシステムコンポーネント障害にかかわらずにオペレーションを維持することが重要である。追加的なオーバーヘッドにより、セルロックネットワーク410を強化して、どこかで障害が起こったときのシナリオに耐えられるようにすることは可能である。

【０２２４】

ここで考慮される障害は、コンピュータ装置101が１つの接続パス215を介してシンボル223を受信することをやめる、ということである。障害は、一過性の、もしくは持続的な性質のものでありうる。一過性の障害は高速シリアル相互接続による同期の失敗により引き起こされる可能性がある。持続的な障害の理由となりうるのは：

・　接続部の破損
・　送信機のハードウェア障害
・　受信機のハードウェア障害

【０２２５】

データフローへのダメージを抑えるために、障害の起きているパス215は利用可能なリソースから出来るだけ速やかに切り離されなければならない。従って、利用できないという情報は、現在のまたは次のセル期間603内で全てのコンピュータ装置101に向けて発信されなければならない。

【０２２６】

アービトレーションのための構造は、各コンピュータ装置101に関連する情報を、他のコンピュータ装置101へ発信することすなわち広く知らせることことに基づく。使用不可能なパス215を通ってルーティングされるデータ・セグメントは確実に見失われ、セルロックネットワーク410はこのケースについての構造を修復できなくなってしまう。上位レベルのプロトコルは、見失ったデータを検出し、要求があればその再送を開始することができる。しかし、アービトレーション関係情報が全てのコンピュータ装置101から全ての他のコンピュータ装置101へ正確に通信されていない場合には、各アービトレーション結果が相違したものとなってしまい、データ・ストリームが正確にルーティングされなくなってしまう。従って、全てのアービトレーション関係情報は、代替パス215上でも利用可能であるべきである。この冗長化構造は、先行するコンピュータ装置101から他の全てのコンピュータ装置101に送信されるアービトレーション関係情報の再送を介して提供される。再送されたアービトレーション関係情報が利用される場合、ダメージの範囲を、障害パス215を通ってルーティングされるデータ・セグメントの損失に限定することができる。

【０２２７】

高い可用性の実装にあたっては、転送エージェント役のコンピュータ装置101において転送データに起こりうるダメージや、アービタまたはネットワーク・インターフェース構造における別のパスの障害についても、考慮しなければならない。独立した第２のパラレルなセルロックネットワーク410を実装することで、冗長化のためのより強固なサポートが可能となる。

【０２２８】

実施形態は、単一接続障害シナリオをカバーするために、アービトレーションパラメー

タの冗長化分散を包含してもよい。

【０２２９】

コンジットのような動作の副次的影響は、リクエストされている/されていないにかかわらず、コンジットが送信リクエストのアービトレーションからの結果を無条件に実行してしまう、ということである。導入されたアービタは、データ送信ウィンドウが要求されていない場合とデータ送信要求が何も成されていない場合とで、同じ結果を生成するというルールに従わなければならない。あるルートに関して、要求のない伝送パス215やデータが利用可能である場合、そのルートを利用することが許される。このような利用については、受信機についても同じであるということが認識されるべきである。最下位レベルの送信リクエストは、優先レベルがサポートされるのであれば、リクエストされうる送信ルートの確保に利用してもよい。

【０２３０】

本発明の実施形態は、特定の送信リクエスト優先レベルの実装/非実装方法に限られない。さらに、セルロックネットワーク410上の上位レベルのプロトコルを符号化する特定の方法に限られない。シンボル222の割り当てルールは、セル開始シンボル209がセル開始信号の伝達227以外の目的のために利用されてはいけない、というものであってもよい。

【０２３１】

大きなネットワークに基づくシステムに必須の特徴は、システムオペレーションの停止もしくは中断を行わずに、ネットワーク・コンポーネントの導入または削除ができる、ということである。セル601内の情報ブロックにおけるデータ送信機能を無効にすること、および次のセル期間603でセルロックネットワーク410のインターフェース・リンク221を遮断することは、セルロックネットワーク410の観点から、安全である。

【０２３２】

ネットワーク参加者の最大数は、有用な実装のために重要である。一目瞭然であるので、プロトコルのオーバーヘッドはネットワークサイズと連動する。それはつまり、より小さなネットワークは異なるプロトコルが定義されない限り、オーバーヘッドを減らすことができない、ということである。また、より小さなネットワークは帯域幅の優位性についてより大きなネットワークほどの可能性を有していないということがある。一方で、大きなネットワークは、リンクのルーティングについて課題を抱える傾向にある。構造の物理的な大型化は信号伝達の遅延を増加させる。ネットワークがより多くのコンポーネントを有するほど、アービトレーションの機構も複雑化する。これらを考慮すると、コンピュータ装置の数が12ないし16というのが、セルロックネットワークにとっては最適であろう。しかし、この実施形態はネットワークのある特定のサイズを制限するものではない。

【０２３３】

ラックマウント型システムにおいて、スロットの一部はある特定のアプリケーションのために使用せずに空けておいてもよい。このようなケースでは、空気の循環および電磁両立性（electromagnetic compatibility：EMC）を考慮し、ラックを閉じるのにフィルタパネルが使われることが多い。複数の実施形態に従ってセルロックネットワーク・アーキテクチャを利用するシステムにおいて、空のスロットをフィルタパネルで塞ぐことより、セルロックネットワーク上でペイロード転送性能を提供するコンピュータ装置で塞ぐことの方が有益であり、従って推奨される。これは、システムの帯域幅をできるだけ高くすることを確実にする。

【０２３４】

ローカルで生成されるシンボル・レートの誤差が小さく、セル601のセル長がそれほど長くない場合、実施形態によっては、セル開始シンボル209の後ろにアイドル・シンボル210およびセル開始シンボル209を除く所定の数のセル・コンテンツが続くことが可能である。この発展型は、シンボル期間を節約し、セル601の個別処理のための柔軟性を維持するために、極めて有益である。

【０２３５】

複数のレーンを利用することで、特定リンク221の帯域幅を増加することが可能である

ことはすでに述べた。複数のレーンがリンク221のために利用される場合、各々のレーンにおいてセル開始シンボル209が利用される必要があるかもしれない、ということに注意する必要がある。リンク221上の全てのレーンの長さは原則的に同じでなければならない。

【０２３６】

セルロックネットワーク410が、全てのリンク221上の双方向において複数のレーンを使用するような実装も可能である。そのような実装において、特定の数のレーンをまたいで制御シンボル805を配信することは確実に可能である。これはプロトコルにおける制御オーバーヘッドの割合を減らすか、または制御オーバーヘッドの割合を増大させずによりフレキシブルな制御プロトコルを可能にする。複数のレーンを有するリンク221を使用したペイロード転送による上記第３のソリューションでは、セル長606は長くなってしまう。このことは、現実に送信されたデータの合計がある程度小さい場合に悪影響をもたらす。この点で、ペイロード転送についての上記第１のソリューションが、より適切な場合があり、それはセル長606を適切な長さにすると同時にネットワーク410の応答時間を短縮することを可能にする。

【０２３７】

上述の基盤構造は、様々なプロトコル・フォーマットを呈するシンボル・ストリームの伝送能力を高める可能性がある。

【０２３８】

図１３は、複数のプロトコルをサポートするセルロックネットワークの例についての略ブロック図である。

【０２３９】

複数のコンピュータ装置101のネットワーク・インターフェース(NWIF) 1302の構成は、複数のFIFO 1299、複数のマルチプレクサ、アービトレーション回路1401、複数のプロトコル非依存型入力キュー(PAIQ) 1405および複数のプロトコル非依存型出力キュー(PAEQ) 1404等を含む。

【０２４０】

ラッパー(WRP) 1310の構成は、ネットワーク・インターフェースNWIF 1302構造のためのプロトコル固有の出力キュー1305およびプロトコル固有の入力キュー1306用インターフェースを含む。

【０２４１】

コンピュータ装置のローカル構造は、プロトコル固有の性質により出力キュー1305および入力キュー1306をサポートする。出力キュー1305および入力キュー1306はそれぞれ２つの番号で特定される。１つ目の番号はキューによって供給されるプロトコルを特定し、２つ目の番号は各ターゲットまたは送信元コンピュータ装置101のジオグラフィック・アドレス106を特定する。

【０２４２】

コンピュータ装置APP_A 101およびAPP_B 102は、CPUおよびI/O（input/output：I/O）1303の両方を備える。ローカルインターネットプロトコル（Local Internet Protocol：IP）およびSAS通信チャネルはスイッチングおよびブリッジコンポーネント1304に組み合わせられ、出力キュー1305および入力キュー1306に接続する。コンピュータ装置APP_C 103は、接続した２つのサブユニット1307およびディスク1308に接続した任意のSASを有する構成を呈する。コンピュータ装置APP_D 104はSASディスクまたはサブシステム1309として示される。コンピュータ装置APP_A 101、APP_B 102およびAPP_C 103は、セルロックネットワーク410を介して送信および受信を行うためのIPおよびSASパケットの両方を有する。

【０２４３】

NWIF 1302は、サポートされた各リンク221のために、信号PAEQ 1404および信号PAIQ 1405をメンテナンスする。WRP 1310は、プロトコル固有の出力キュー1305のシンボル・ストリームをを信号PAEQ 1404に統合するための機能および、プロトコル固有の入力キュー1306のシンボル・ストリームを信号PAIQ 1403から分離する機能を有する。

【０２４４】
ある実装は、たまたま使用されていないセル601内のシンボル位置を埋めるために、非データシンボルを特定してもよい。
【０２４５】
セルロックネットワーク401内の異なるプロトコルの数はそれほど大きな数ではないと思われるので、１バイトで表現されうる２５６個の値で十分なはずである。プロトコルの種類は各セル601の最初のペイロード・シンボルの位置で宣言されてもよい。図１３の例において、IPプロトコルはコード「10」で識別され、SASプロトコルはコード「30」で識別される。
【０２４６】
セル内でプロトコルの変更が認められるのであれば、非データシンボルSWIが指定され利用されることが必要とされ、SWIシンボルに続くシンボル222が新しい種類のプロトコルを宣言する。識別されたSWIシンボルは、いかなるプロトコル・データ・ストリーム内においても使用されてはならない、という規則が守られねばならない。パケット・プロトコルによっては、パケット内においてデータ・ストリームの割り込みを許可しないかもしれないことを考慮することは特に重要である。そこで、そのようなインターフェースを利用する構成は、パケットの送信を開始するのであれば、当該パケットが当該パケットのためのデータセットの全てを含んでいるか、データの続きが時間内に到着し、障害が発生した場合には再送が手配される、という場合のみに、送信を開始するべきである。インターフェースの構成は、特定のプロトコルの、タイムクリティカル連続性要件は優先されるべきであるという側面を考慮して構築されなければならない。全てのセル送信期間は、少なくとも２つの接続されているコンピュータ装置101の間の直接接続の帯域幅を付与することができることにここで留意すべきである。このことは、セルロックネットワーク410全体で、タイムクリティカル・プロトコルのためのデータのストリーミングの十分な基盤になりうる。
【０２４７】
一般に、全ての出力および入力キューは、コンジットのような動作を行うべきである。データが損なわれてしまう、という最悪なケースにおいてさえも、キューおよびブロックの有用なリソース内に未使用データを蓄積しないことが大事である。不完全なデータは出力キューから除去または送信されなければならない。でなければ不完全なデータがリソースをブロックしてしまう。同様に、不完全な受信データもいずれかの方法で入力キューから除去されるべきである。
【０２４８】
図１５は、コンピューティングシステムおよびネットワーキングシステムにおける関連項目の階層一覧表である。セルロックネットワーク410が既存のネットワーク・アーキテクチャとどのように関連しているのか理解するのに役立つだろう。
【０２４９】
以下に、典型的なパラメータ値を伴う実施形態をより詳しく説明する。
【０２５０】
フルメッシュ「ファブリック・インターフェース」を伴うPICMG 3.0(r) AdvancedTCA(r)シェルフの実装例は、プロトタイプな実施形態に使用可能な基盤を示す。「ポート0」は「ファブリック・インターフェース」のフルメッシュ・ネットワークにおける各「リンク」に利用される。プロトタイプな実施形態のコンピュータ装置101は、「メッシュが有効な」AdvancedTCA(r)「ボード」を実装しうる。
【０２５１】
実施形態は、図１２に示されるシンボル入力アーキテクチャを実装するので、他のコンピュータ装置101の制御情報は粗いタイミング調整の間にすでに利用可能である。
【０２５２】
以下の割り当ては、この例のための設定である。この実施形態は最大16の相互接続されたコンピュータ装置101からなる。バックプレーンによる相互接続はフルメッシュ・ネッ

トワークとなる。電気的インターフェースはLVDS（low voltage differential signaling：低電圧作動信号送信）である。この例においては、シンボル・フロー方向毎に１つの異なるペアが使用される。シグナリングは8ビット/10ビットの符号化を使用する。全二重リンクは両方向とも公称上同じ遅延を有する。送信ビットレートは3.125 Gbit/sである。送信クロックの許容誤差範囲は±50 ppmである。ペイロード転送を行う第３のソリューションに従う。制御シンボル805のグリッドは25シンボル期間である。ペイロード転送のためのオフセット706は、25シンボル期間である。

【０２５３】

以下の非データ・シンボル定義が使用される：

・ COM: セル開始シンボルとして使用される'カンマ'シンボルである。KコードK.28.5に割り当てられる。

・ SKP: アイドル・シンボルとして使用されるSKIPシンボルである。割り当てはK.28.0である。'カンマ'以外のシンボル222は、アイドル・シンボル210位置で使用されてもよい。

・ RST: パケット送信用に指定されるC2制御シンボル位置で送信される場合、パケット・カウンタをリセットするためにPADシンボルK.23.7が割り当てられる。

・ SWI: ペイロード・データ位置におけるSKIPシンボルは、プロトコルのスイッチングに利用される。割り当てはK.28.0である。

・ PST: 信号を事前開始するために、C2制御シンボル位置にSKIPシンボルが利用される。つまり、次のセルはRSTシンボルを含み、次のパケットの開始を合図する。

・ PAD: ペイロード・データ位置において繰り返されるPADシンボルが、未使用のシンボル位置を埋めるために利用される。割り当てはK.23.7である。

・ Data symbols: 8ビットのバイトデータを10ビットに符号化。

【０２５４】

他の非データ・シンボルが、セルのペイロード・データ位置で、上位レベルのプロトコルによって利用してもよい。

【０２５５】

次に、アイドル・シンボルの数を導き出す計算のための、正確な式が説明される。

【０２５６】

式に用いられる表記は以下の通りである。

min(): 値の集合に対して適用される最小関数
max(): 値の集合に対して適用される最大関数
abs(): 絶対値関数 (例) abs(x) := |x|
sgn(): -1, 0または+1を返す符号関数
trunc(): 切り捨て関数
mod: モジュロ演算子
Φ: 空集合
∈: 集合要素
∉: 集合要素ではない
∪: 集合の和
∃: 少なくとも１つ存在
∧: 論理積演算子
∨: 論理和演算子

【０２５７】

本実施形態においては、以下の通り定数を指定する。説明は後になされる。

N := 16
n := 1025
defidle := 1/2
rangein := 4

entrlock := 1024
rangeout := 6
【０２５８】
N: 本実装によってサポートされるコンピュータ装置101の数
【０２５９】
n: セル606毎のシンボルの数
【０２６０】
minidle: アイドル・シンボルの最小値
minidle := 0
【０２６１】
maxidle: アイドル・シンボルの最大値
maxidle := 2 * defidle
【０２６２】
defidle: 本実施形態におけるアイドル・シンボルのデフォルト数605。次の式で求められる：

defidle := (minidle + maxidle) / 2

または、発信されたシンボル・レート許容誤差値に従って、およびコンピュータ装置101が同期の参照先として宣言されているかどうかを考慮することで、動的に求められる。
【０２６３】
ga: ジオグラフィック・アドレス106 $1 \leqq ga \leqq N$
gax: ジオグラフィック・アドレス106 $1 \leqq gax \leqq N$
【０２６４】
APP_{ga}: ジオグラフィック・アドレス106 gaを有するコンピュータ装置101
【０２６５】
$L_{ga,gax}$: APP_{ga}およびAPP_{gax}を接続するリンク221
【０２６６】
GA: 現在のコンピュータ装置101のジオグラフィック・アドレス106のインデックス。この計算においてはAPP_{GA}のように表記される
【０２６７】
SUB9: ロック状態S9（310）にあるリンク221を有するコンピュータ装置101を表す、ジオグラフィック・アドレス106のサブセット。
SUB9 := {ga: ga∈{1..N}, ∃ gax∈{1..N}} （状態S9の$L_{ga,gax}$で）
【０２６８】
SUB8: ロック状態にあるリンク221は有さないが、細かいタイミング調整がなされた状態S8（309）にあるリンク221を有するコンピュータ装置101を表す、ジオグラフィック・アドレス106のサブセット。

$SUB8 := \{ga: ga \in \{1..N\}, ga \notin SUB9, \exists\ gax \in \{1..N\}\}$（状態S8の$L_{ga,gax}$で）

【０２６９】
SUB7: 粗いタイミング調整のみがなされた状態S7（308）にあるリンク221を有するコンピュータ装置101を表す、ジオグラフィック・アドレス106のサブセット。

$SUB7 := \{ga: ga \in \{1..N\}, ga \notin SUB9, ga \notin SUB8, \exists\ gax \in \{1..N\}\}$（状態S7の$L_{ga,gax}$で）

【０２７０】
gacpr: 周期的処理に関して先行するコンピュータ装置101のジオグラフィック・アドレス106
gacpr := min({max({ga: ga ∈ SUB9, ga＜GA}), max(SUB9)})
【０２７１】
APP_{gacpr} : サイクリックに先行タスクをおこなうコンピュータ装置101

【０２７２】
$moffset_{ga}$ ： ローカルで測定されたAPP_{ga}のセル開始シンボルのオフセット226
【０２７３】
$moffset_{GA}$ ： APP_{ga}について値0
【０２７４】
$roffset_{ga}$ ： APP_{ga}からのセル開始シンボル209について受信したオフセットの測定データ。利用できる測定データがない場合の値は0となる。値が0となるのは粗いタイミング調整の間のみ。
【０２７５】
$roffset_{GA}$ ： APP_{GA}について値0
【０２７６】
$diffoffset_{ga}$ ： この符号値は、APP_{GA}のセル開始タイミング227からAPP_{ga}のセル開始タイミング227のオフセット228の2倍の値に等しい。
$diffoffset_{ga} := moffset_{ga} - roffset_{ga}$
【０２７７】
moffsetrange: 測定されたオフセット値226の範囲
$moffsetrange := max(\{moffset_{ga}: ga \in SUB7\}) - min(\{moffset_{ga}: ga \in SUB7\})$
【０２７８】
rangein: 本実施形態で使用される、絶対値関数$abs(diffoffset_{ga})$の最大値。セルロック状態に入ることをリンク221 $L_{GA,ga}$を許可するための値。
【０２７９】
entrlock: リンク221がロック状態になる以前に、リンク221 $L_{GA,ga}$が$abs(diffoffset_{ga}) \leqq rangein$を満たすように調整されるまで待機するための、所定の数のセルサイクル604。
【０２８０】
rangeout: セルロック状態を解除するための閾値。
$abs(diffoffset_{ga}) \geqq rangeout$であるとき、本実装はリンク221 $L_{GA,ga}$がセルロック状態ではないことを示す。
【０２８１】
garef: ネットワーク410のタイミング参照先であることを通知しているコンピュータ装置101のジオグラフィック・アドレス106。参照先であるとするコンピュータ装置101が複数存在する場合は、それらは無視されなければならない。
【０２８２】
APP_{garef} ： ネットワーク410のタイミング参照先であると宣言されているコンピュータ装置101。
【０２８３】
garefvalid: タイミング参照先が特定された場合に、garefvalid = TRUE でなければ、 garefvalid = FALSEである。garefの値はgarefvalid = FALSEの場合に無視される。タイミング参照先が導入され、同時にネットワーク410においてロック状態が続いている場合に、セルロック中のコンピュータ装置101$\{APP_{ga}: ga \in SUB9\}$のサブセットがセルロック状態を解除することなく$APP_{garef}$のタイミングに向かうようにするための特別な予防策が必要となる。APP_{garef}は$abs(diffoffset_{garef}) \leqq rangein$の場合に、直ちにロックされたサブセットを追加されなければならない。この手続きの間は、ロックされたサブセットに別の追加がなされることはない。
【０２８４】
trefvalid: タイミング参照先が有効であるかどうかを反映した値であり、指定または自動選択される。粗いタイミング調整プロセスの間、より細かく調整されたサブセットが存在せず、さらにセル開始タイミング227がセル期間603の少なくとも1/4で配信されている場合、最小ジオグラフィック・アドレス106を有するコンピュータ装置101がタイミング参照に使用される。

trefvalid := garefvalid ∨ ((SUB9=Φ) ∧ (SUB8=Φ) ∧ (moffsetrange ≧ n/4)

【０２８５】

tref: タイミング参照のインデックスであり、最初の粗いタイミング調整のためにgarefから引き継いだものと自動的に割り当てられたものの両方。

```
if (garefvalid=TRUE)
then tref := garef
else tref := min(SUB7)
```

【０２８６】

ALIG7: よりよくタイミング調整されたサブセットが存在しない場合に、タイミング調整のために考慮されるSUB7のサブセット。

```
if (SUB9 = Φ) ∧ (SUB8 = Φ) ∧ (trefvalid = TRUE)
      then ALIG7 := {tref}
else if (SUB9 = Φ) ∧ (SUB8 = Φ) ∧ (trefvalid = FALSE)
      then ALIG7 := SUB7
else ALIG7 := Φ
```

【０２８７】

ALIG8: タイミング調整のために考慮されるSUB8のサブセット。

```
if SUB9 = Φ
      then ALIG8 := SUB8
else ALIG8 := Φ
```

【０２８８】

ALIG : タイミング調整のために考慮すべきサブセット。

ALIG := SUB9 ∪ ALIG8 ∪ ALIG7

【０２８９】

first: 現在のコンピュータ装置101のセル開始シンボル209における最初のセル開始シンボル209の、符号付きオフセット228の2倍の値。

first := min({diffoffset $_{ga}$: ga ∈ ALIG})

【０２９０】

last: 現在のコンピュータ装置101のセル開始シンボル209における最後のセル開始シンボル209の、符号付きオフセット228の2倍の値。

last := max({diffoffset $_{ga}$: ga ∈ ALIG})

【０２９１】

ISF_{ga} : APP $_{ga}$ が現在のセルの直前に適用したアイドル・シンボルの数。

【０２９２】

maxis : 現在のセルの直前に適用されたアイドル・シンボルの最大数。

maxis := max({ISF_{ga} : ga∈ALIG})

【０２９３】

minis : 現在のセルの直前に適用されたアイドル・シンボルの最小数。

minis := min({ISF_{ga} : ga∈ALIG})

【０２９４】

midis : 現在のセルの直前に適用されたアイドル・シンボルの中間数。

midis := (maxis + minis) / 2

【０２９５】

midis2 : 常に整数値を扱うための、midisの2倍の値。

midis2 := maxis + minis

【０２９６】

gravis : 適用されるアイドル・シンボル210の数がアイドル・シンボルのデフォルト数605に近づくことを保証するために計算に加えられる必要のある値。この値は、コンピュータ装置101が同じ数のアイドル・シンボル210を接続している各コンピュータ装置101に向けて適用するために、各コンピュータ装置101で個々に計算される。

gravis := defidle - midis

【０２９７】

gravis2 : 常に整数値を扱うための、gravisの2倍の値。

gravis2 := maxidle - midis2

【０２９８】

trg: 次のセル開始のための第１のターゲット。

trg := (first + last) / 2

【０２９９】

trg2: 常に整数値を扱うための、trgの2倍の値。

trg2 := first + last

【０３００】

chg: 接近される必要があるタイミング参照先がない場合にdefidleに相対して適用される変化の絶対値である。タイミング参照先がSUB8またはSUB9内である場合、次の式が適用される。

chg := min({abs(trg + gravis / 2), defidle})

【０３０１】

chg2: 常に整数値を扱うための、chgの2倍の値。

chg2 := min({abs(trg2 + gravis), maxidle})

【０３０２】

chgappr: タイミングが調整されたサブセットの範囲外にあり、タイミング参照先にアプローチされる必要のある場合に、defidleに相対して適用される変化の絶対値である。APP_{garef}のセル開始タイミング227が調整されたサブセットの範囲外にある場合、ロック状態S9 310もしくは細かいタイミング調整状態S8 309にあるコンピュータ装置101によってアプローチされなければならない。ただし、タイミング調整レベルは維持されなければならない。相対的な位置とは関係なく、アプローチは常に同じ方向でなされる。これにより、それらが調整されたサブセットに結びつけられている間にコンピュータ装置101が反対方向に調整移動する必要がなくなる。なお、修正条件下では、同期を行うために、アイドル・シンボルのデフォルト数に必要な値を増やさなくてはならない可能性にも留意すべきである。

chgappr := min(abs(trg),max(defidle,1))

【０３０３】

nxtidle: 現在のセル601の末尾シンボル220の後ろに、APP_{GA}によって全てのリンク221に適用される、計算されうるアイドル・シンボル210の数。

【０３０４】

計算は以下の通りである。これらの式のシーケンスの後にdefidle = 1/2であれば、nxtidleの結果はnxtidleの最終値である。計算が整数のみとなるように、以下に示す式の一部にdefidleの2倍であるmaxidleを用いる。

```
nxtidle := trunc(defidle)
if (trefvalid=TRUE) ∧ (GA=tref)
      then nxtidle := defidle - (nxtidle - defidle)
else if (trefvalid=FALSE) ∨ (tref∈ALIG)
      then nxtidle := trunc(sgn(trg2*2 + gravis2) * chg2 + 1 + maxidle*2)
/ 4
else nxtidle := max(0, sgn(trg2) * trunc(chgappr) + minis)
```

【０３０５】

細かいタイミング調整状態からセルロック状態への状態の移行は、以下の通り制御される。

【０３０６】

セル周期604のentrlock数が連続したシーケンスである間、またはAPP_{ga}がリンク221 $L_{GA,ga}$をセルロック状態であると認めている間に、

(last - first) ≦ rangein
であると確認されているとき、
(trg - rangein/2) ≦ diffoffset $_{gax}$ ≦ (trg + rangein/2)
もしくは、計算を簡単にするために、
(trg2 - rangein) ≦ diffoffset $_{gax}$ * 2 ≦ (trg2 + rangein)
である全てのリンク221 $L_{GA,gax}$の状態は、細かいタイミング調整状態からセルロック状態へと変化される。

【０３０７】

セルロック状態のリンク221がabs(diffoffset $_{ga}$) ≧ rangeouである場合、直ちにリンク221をアービトレーションのために利用することができないことと、リンク221の状態は適宜細かいタイミング調整状態もしくは粗いタイミング調整状態へと格下げされることとが宣言されなければならない。

【０３０８】

粗いタイミング調整状態から細かいタイミング調整状態への状態の移行は、以下の通りに制御される。

【０３０９】

リンク221の開始時、両端にあるコンピュータ装置101は、図１８の表に指定されている通りに粗いタイミング調整状態への割り当て（assignment）に対応するセル601を送信する。セル開始タイミング227が一定の限度を超えてずれている間、リンク221の粗いタイミング調整状態が持続する。この例示的実施形態において、細かいタイミング調整の間に許容される最大限のずれは、およそ±100のシンボル期間225であり、これはセル601の末尾制御シンボルC24のシンボル位置による（図１６の表を参照）。細かいタイミング調整状態を識別する前に、タイミング調整はすでにタイミング調整済みのサブセット{APP$_{ga}$: ga∈ALIG}の全ての対象は±100の範囲でなければならない。このことが達成される場合、コンピュータ装置101はリンク221において送信が細かいタイミング調整状態であるように変化させる。リンク221の別の一端にあるコンピュータ装置101もまた、リンク221にとって細かいタイミング調整状態であるように送信する。次のセル期間603内で、両方のコンピュータ装置101が図１７の表で指定されるようにセル601のフォーマットを変更する。

【０３１０】

ネットワーク・インターフェースの構成についての例の詳細な説明は以下の通りである。

【０３１１】

図１４は、例示的実施形態についてのネットワーク・インターフェース(NWIF) 1302のブロック図である。コンピュータ装置101のこの機能については、図１３を参照のこと。

【０３１２】

NWIF 1302の構成は、プロトコル非依存型出力キュー(PAEQx15) 1404およびプロトコル非依存型入力キュー(PAIQx15) 1405を介してWRP 1310の構成に向かって複数のインターフェースを提示する。15の個々の出力/入力キューは相互に、接続されているコンピュータ装置101のジオグラフィック・アドレス106に関連付けられる。これらのキューはアービタ(ARB) 1401の制御下にある。ARB 1401は、セル期間603後いくつのペイロード704のセル・ロードがそれぞれ受信・送信されたかという情報を個々のキューに知らせる。一方で、ARB 1401は出力キュー1404において利用可能な送信可能データの合計についての情報を受け取る。

【０３１３】

セルロックネットワーク・インターフェースの入力信号は受信したデータ・ストリームをバッファしタイミング調整するFIFOx15 1299の回路に接続される。そのため、受信したデータ・ストリームはローカル・シンボル・レートと同期している内部パス1419上で利用可能となりタイミングが調整され、その結果、コンピュータ装置101によって生成されるセル601送信の所定のオフセットを有することになる。FIFOx15 1299の回路は、入力シンボル・ストリームをバッファし再び時間調整をするために必要とされる全ての機能を包含

する。セル・ロッキング制御（CLC）1408ブロックは、正しい数のアイドル・シンボル210の挿入について、特に関わる。

【０３１４】

パス1419上で利用可能な15の接続の入力データ・ストリームは、３つの回路構成にルーティングされる。マルチプレクサ(MUX15:15) 回路1411は、入力データ・ストリームをソートして入力キューPAIQx15 1405をマッチさせる。PAIQx15 1405は、コンピュータ装置101に接続しているジオグラフィック・アドレス106の順に割り当てられている。セル・ペイロード704はこのマルチプレクサを通って対応するPAIQx15 1405内の入力キューに運ばれる。制御シンボル805は制御シンボル・エクストラクター(CSE) 回路1407により、入力データ・ストリームから抽出される。APP_{gacpr}のデータ・ストリームはマルチプレクサ(MUX15:1) 1420により選択される。入力データ・ストリームは、マルチプレクサ(MUX15:15) 1414内で、転送エージェント役のコンピュータ装置101への各出力パスに再ソートされる。別のマルチプレクサ(MUX2:1x15) 1423は、制御シンボル805の転送をサポートする。任意の共通制御がこれらのマルチプレクサに利用され、APP_{gacpr}からの制御シンボル805は、MUX2:1x15 1423内で、各マルチプレクサへの入力の一つとして利用可能である。

【０３１５】

CSE1407の回路はオフセット226の情報、要件および性能、および受信したセルの制御シンボル805位置からの別の情報を収集する。それらはARB 1401およびCLC1408の回路にそれぞれ提供される。

【０３１６】

出力キューPAEQx15 1404のデータ・ストリームはセル601のための制御情報を持っていない。出力キューPAEQx15 1404からの制御情報は、制御シンボル挿入（CSI）1406の回路およびARB 1401にルーティングされる。CLC 1408の回路はまたCSI 1406に情報を提供する。

【０３１７】

マルチプレクサ(MUX2:1x15) 1409はCSI 1406からの制御シンボル805を出力キューPAEQx15 1404からのデータ・ストリームに挿入する。

【０３１８】

MUX2:1x15 1409の複数のデータ・パスはジオグラフィック・アドレスごとに並べられるので、外部リンク221の割り当てに従って再度並べられなければならない。これはMUX15:15 1410内で行われる。

【０３１９】

出力シンボル・ストリームは、MUX15:15 1410を経由するコンピュータ装置101自身に、またはMUX2:1x15 1423を介して利用可能な転送データ・ストリームのどちらかによって供給される。マルチプレクサの配列MUX2:1x15 1412は、各2:1マルチプレクサに対して個別に制御を有する。

【０３２０】

制御シンボル挿入回路CSI1406により提供される制御シンボル805は、マルチプレクサ配列1423の出力からはなくなっている。これらの制御シンボル805は、MUX2:1x15 1409およびMUX15:15 1410を通過する。パス1418を経由するシンボル送信が、正しい制御シンボル805の位置で選択されるように、MUX2:1x15 1412が制御される。

【０３２１】

各々の識別子に「x15」を持つ図１４の回路は、並列に搭載される15の同一のデバイスから構成されることに留意する。

【０３２２】

MUX15:15 1411およびMUX15:15 1410のマルチプレクサ配列は物理的に存在する必要はない。いくつかの点で、また実装に応じて、非常に柔軟に割り当てられうるデータキューは、とりわけ、割り当ての特定の順番に制限されたストレージ利用を必要とせずに、アドレス指定可能な要素として現れる。

【０３２３】

以下は、例示的実施形態における、セル601のシンボル位置への割り当て方法の更なる詳細である。

【０３２４】

図１６の表に示されるように、シンボル位置1～1025は、セルフォーマットCF1 702またはCF2 703において、C1～C24、D1～D975およびW1～W26に割り当てられる。

【０３２５】

制御シンボル805はセル601内に図８のやり方に従って位置づけられる。制御シンボル位置805間のペイロード・データ・シンボル803の数は24である。セルフォーマットCF1 702からCF2 703へのペイロード転送オフセット706はシンボル期間225の25個分である。

【０３２６】

図１６の表において、制御シンボル位置805，C1～C24は結合されたボックスに描かれており、セルフォーマットCF1 702およびCF2 703の両方で同じように割り当てられることが強調される。W1～W26として割り当てられるシンボルは、W 705エリアのシンボルや制御シンボル位置805を含む。これらは、ペイロード転送オフセット706にバインドされているためにDシンボルに変化させることができない。

【０３２７】

W1～W26の位置は分散しており、セルフォーマットに依存して配置される。が、これらはIP送信において全て割り当てられてうる。

【０３２８】

C1～C24の制御シンボル位置805は、25シンボルのグリッド内の２つのグループに割り当てられる。第１のグループC1～C4はセル601の前半層に割り当てられ、C5～C24はセル601の後半層に割り当てられる。重要なのは、C24の後にアービトレーションを実行するのに十分な時間がある、ということである。

【０３２９】

図１７の表は、実施形態においてリンク221がセルロック状態および細かいタイミング調整状態にある間に、セル601内における制御シンボル805　C1～C24の割り当てを示す。表の最終列は、それぞれの指定がどの状態に適用されているかを示す。シンボルC1～C24は以下の通り割り当てられる。

【０３３０】

C1：セル開始シンボル209 COM

【０３３１】

C2：この位置は情報に関するパケットの送信に使用され、一定もしくは低いレートで変化している。コンピュータ装置101は、パケットをセル601毎に1つのシンボル222と共にこの制御シンボル位置805で送信する。RSTシンボルはパケット開始信号に使用される。RSTの後に、後続セル601が逐次的にパケットの次のシンボルを含む。パケットの最後にはPSTシンボルが送信され、それによって、他の全てのコンピュータ装置101に対して次のセル601のパケット開始が伝えられる。図１８の表はパケット・コンテンツの割り当てを示す。受信機は、この制御シンボル位置でPSTまたはRSTシンボルが検出されるまで受信したシンボル223を無視しなければならない。この例示的実施形態はパケット長を固定し、図１９の表で示される割り当てと一緒に情報ブロックとしての利用を提供する。

【０３３２】

C3：APP_{gacpr}受信したC2制御シンボルを転送するために利用される。

【０３３３】

C4：この制御シンボル805は、セル・オフセット測定データ226を指定した8ビットの整数値として伝達する。この整数値は各リンク221毎に個別に評価され送信される。値が±127の範囲を超えてしまった場合、粗いタイミング調整状態を特定するバイナリ・パターン「1000 0000」が送信され、そのようなケースにおいては、図１８の表は図１７の表の代わりに利用される。

【０３３４】

C5：この制御シンボル805は、ジオグラフィック・アドレスが範囲1から8にあるコンピ

ュータ装置101にリンク221の完全動作情報を提供する。ビット割り当ては図２１の表に示される。故障しているリンク221またはいかなる理由であれロック状態を失っているリンク221の完全動作ビットは直ちに無効にされなければならない。コンピュータ装置のジオグラフィック・アドレス106に割り当てられたビット位置は、コンピュータ装置101がセルロックネットワーク410のタイミング参照先であると宣言する場合は「1」に設定され、そうでなければビットは「0」にクリアにされる。

【０３３５】

C6：APP_{gacpr} から受信されるC5制御シンボル805の転送に使用される。

【０３３６】

C7：この制御シンボル805は、ジオグラフィック・アドレスが範囲9から16にあるコンピュータ装置101にリンク221完全動作情報を提供する。ビット割り当ては図２１の表に示される。故障しているリンク221またはいかなる理由であれロック状態を失っているリンク221の完全動作ビットは直ちに無効にされなければならない。コンピュータ装置101のジオグラフィック・アドレス106に割り当てられたビット位置は、コンピュータ装置101がセルロックネットワーク410のタイミング参照先であると宣言する場合は「1」に設定され、そうでなければビットは「0」にクリアにされる。

【０３３７】

C8：APP_{gacpr} から受信されるC7制御シンボル805の転送に使用される。

C9、C11、C13、C15、C17、C21、C23：これらの制御シンボル位置805は送信リクエスト・コードを運ぶ。各シンボル805は２つの送信ターゲットに対する送信リクエスト・コードを含み、シンボル位置は各ジオグラフィック・アドレスに関連づけられる。送信するコンピュータ装置101自身のジオグラフィック・アドレス106のコード位置において、4ビットはgacprの送信ビット3..0に利用される。なお、gacprは周期的処理に関して先行するコンピュータ装置101 APP_{gacpr}のジオグラフィック・アドレス106である。細かいタイミング調整状態S8 309において、これらの制御シンボル805は無視される。

【０３３８】

C10、C12、C14、C16、C18、C20、C22、C24：これらの制御シンボル位置805は送信リクエスト・コードの冗長サポートを伝達する。APP_{gacpr}から先行する制御シンボル位置805で受信した制御コードは、これらの位置で、全てのアクティブなリンク221を介して再送される。細かいタイミング調整状態S8 309において、これらの制御シンボル805は利用されずに無視される。

【０３３９】

図１８の表は、実施形態における粗いタイミング調整状態のリンク221のみについて、セル601内の制御シンボル805 C1～C24の割り当てを示している。しかしながら、このフォーマットが非対称な状況で使用される場合、すなわち、リンク221が、粗いタイミング調整状態のコンピュータ装置101と、細かいタイミング調整状態及び／又はロック状態にあるコンピュータ装置101とを接続する状況で使用される場合も考えなければならない。この表は、リンク221が粗いタイミング調整状態にある場合に、リンク221のいずれの端にあるコンピュータ装置101によっても共に使用されなければならない。シンボルC1～C24は次のように割り当てられる。

【０３４０】

C1：セル開始シンボル209 COM

【０３４１】

C2：図１９の表のP2に記載の通り、ジオグラフィック・アドレスおよびクロック・クオリティ情報を送信する

【０３４２】

C4：粗いタイミング調整状態308において、バイナリコード「1000 0000」が送信される。C4における他の値とともに、図１７の表は図１８の表の代わりに有効になる。

【０３４３】

C5、C7：図１７の表に同じ。ロックされたリンク221のセットを識別するためのリソー

スである。
【０３４４】
C9、C11、C13、C15、C17、C19、C21、C23：これらの制御シンボル位置は、リンク221が粗いタイミング調整状態の間に、セル601の任意レベルの調整に依存することなく、全てのセル601において状態情報を交換するのに利用される。状態情報の符号化は細かいタイミング調整状態もしくはロック状態の間にパケットが使用するのと同じである。
【０３４５】
C3、C6、C8、C10、C12、C14、C16、C18、C20、C22、C24：　無視してもよい。
【０３４６】
図１９はパケット・コンテンツの表である。パケット・コンテンツ送信および同期オペレーションの確立は完全に自動で行われなければならない。以下は、この例示的実施形態におけるコンテンツの定義である：
【０３４７】
P1：パケット先頭のRSTシンボル信号
【０３４８】
P2 ビット3..0：このコンピュータ装置101のジオグラフィック・アドレス106のビット3..0。セルロックネットワーク410は、ユニークジオグラフィック・アドレス106毎に、接続されているコンピュータ装置101の識別情報を頼りにする。接続されているコンピュータ装置101は、このシンボル位置を介して送信された検出されたジオグラフィック・アドレス106によって、コードリンク221の接続を識別する。
【０３４９】
P2 ビット7..4：クロック・クオリティ・コード。最高質のクロック、例えばこの位置において最も小さい値のもの、をもつコンピュータ装置101は、セル601のタイミング調整の基準として使われてもよい。クロック・クオリティ情報はこの実施形態に置いては評価されない。
【０３５０】
P3：IPMIメッセージ・バイトを利用するマネジメント・バス。このシンボル位置を利用するI^{2}Cバス・エミュレーションによって、マネジメントをベースにしたIPMI標準がサポートされてもよい。I^{2}Cベースのマネジメント・バスの個別ハードウェアを実装する必要はない。受信したI^{2}Cバス信号アーキテクチャの値は、セルロックネットワーク410上の他の全てのコンピュータ装置101からこの位置で受信されたデータに基づいて、再構築されうる。
【０３５１】
P4～P11：ジオグラフィック・アドレス106毎に順番付けされたコンピュータ装置101の接続状態情報。シンボル位置毎に２つの4ビット状態コードが提示される。状態コードの割り当ては図２０の表に示す。コンピュータ装置101のジオグラフィック・アドレス106の位置における値は未使用であり、「0000」に設定される。
【０３５２】
P12：パケットの末尾シンボルは、開始準備シンボルPSTである。ここでは、この機能のためにSKPシンボルが指定される。別のコンピュータ装置101がPSTシンボルを検出する場合は、該コンピュータ装置101は次のセル601においてRSTシンボル送信と一緒にそのパケット送信を開始しなければならない。
【０３５３】
図１９は接続状態の表である。状態は4ビットコードで特定される。リンク221の状態は、セルロック状態への信号がないことから特定される。さらに、シンボル・レート許容誤差の測定情報も含む。
【０３５４】
図２１は、送信コンピュータ装置101の観点からリンク221が完全に動作していることを表したビット符号化の表である。接続されているコンピュータ装置101のジオグラフィック・アドレス106に準じて、完全に動作中であることはビット位置において「1」で示され

る。アービトレーション処理によって機能しないパス215が割り当てられないようにするために、これらのビットは最新の状態を反映する。コンピュータ装置101がセルロックネットワーク410から切り離されようとしている場合、送信データを破損しないようにするために、コンピュータ装置101の切断前に、これらのビットを無効化する。

【０３５５】

コンピュータ装置101のジオグラフィック・アドレス106のためのビット位置は、コンピュータ装置101がネットワーク410上でクロック参照を提供しているかどうかを表す信号に割り当てられる。

【０３５６】

図２２は送信リクエスト・コードの表である。制御のためのオーバーヘッドを制限内に保つために、送信リクエストはターゲットのコンピュータ装置101毎に4ビットで符号化される。ここで示される実施形態におけるアービトレーション処理は、はセル期間603毎に１つのセル601が使用可能な場合に制限されているが、この符号化は高帯域幅の接続を許す形態でも提示される。より多くの数の送信が送信パスにリクエストされる場合、送信準備の整っている追加のデータがリクエスト・コードの計算後と次のセル期間603の開始の間で現れる可能性が高い。そのため、より高い送信リクエストの値が生じる可能性もある。待機送信の優先レベルによっては、帯域幅の無駄をなくすために２番目に低いリクエスト境界が使用される可能性がある。残りの送信リクエストはその結果、次のセル期間603まで保留にされる。

【０３５７】

以下に、複数の実施形態に関連したアービトレーション処理の例を詳細に説明する。使用される記号は以下の通りである。

【０３５８】

X：任意のジオグラフィック・アドレス106を表す：　$1 \leqq x \leqq N$

APP_x：ジオグラフィック・アドレス106 xを有するコンピュータ装置101を表す。

N：本実施形態においてサポートされるコンピュータ装置101の最大数

dec()：デクリメント関数

mod：モジュロ演算子

nxt()：サイクリック後行タスクを決定する関数：

nxt(x) := (x mod N) +1

pred()：サイクリック先行タスクを決定する関数：

pred(x) := ((x+N-2) mod N) + 1

APP_g：記述が適用されるコンピュータ装置101を表す

g：コンピュータ装置101 APP_gのジオグラフィック・アドレス106：$1 \leqq g \leqq N$

【０３５９】

それぞれ接続されているコンピュータ装置101のジオグラフィック・アドレス106に準じて番号付けられるため、パス215の識別子が間接的に現れることを考慮する。

【０３６０】

アービトレーション入力パラメータはテーブルに書き込まれる：

RCODE[1..N,1..N]：図２２の表に指定される4ビットコードを表す

A[1..N,1..N]：送信パス215の可用性を表す１ビット値を表す

【０３６１】

アービトレーション結果は、各コンピュータ装置101個別に以下のテーブルに現れる：

OUT[1..N]：符号付き数値を表す

IN[1..N]：符号付き数値を表す

【０３６２】

コンピュータ装置101は、出力キュー1404および入力キュー1405を、接続している各コンピュータ装置101のために、個々のジオグラフィック・アドレス106に従って準備する。

【０３６３】

テーブルの第１のインデックスRCODE[1..N,1..N]およびA[1..N,1..N]は送信元コンピュ

ータ装置101のジオグラフィック・アドレス106であり、第２のインデックスは送信先コンピュータ装置101のジオグラフィック・アドレス106である。

【０３６４】

OUT[x]：パス215 APP_g → APP_xのソースを識別する。

【０３６５】

OUT[x]=xの時、APP_gの出力キュー1404はパス215 APP_g → APP_xのソース・データとして選択される。これは直接接続リンク221を経由した送信を実現する。

【０３６６】

OUT[x]≠xかつOUT[x]>0の時、APP_gはデータ・ソースであり、APP_xは転送エージェントとして利用され、$APP_{OUT[x]}$はデータが送信されようとする送信先コンピュータ装置101である。そのため、APP_gにおける$APP_{OUT[x]}$への出力キュー1404はパス215 APP_g → APP_xのソース・データとして選択される。

【０３６７】

OUT[x] ≠ -xかつOUT[x]＜0の時、APP_gは転送エージェントであり、APP_gは$APP_{-OUT[x]}$から受信したデータをパス215 APP_g → APP_xへ転送しなければならない。

【０３６８】

IN[x]は送信元コンピュータ装置101から入力パス215 APP_g → APP_xを介して期待されるデータを識別する。

【０３６９】

IN[x]=xの時、期待されるデータはAPP_xにより直接提供され、セルフォーマットCF1 702を有する。

【０３７０】

IN[x]≠xかつIN[x]>0の時、APP_gは転送エージェントであり、APP_xからのデータをセルフォーマットCF1 702で受信する。受信するデータはセルフォーマットCF2 703で$APP_{IN[x]}$に転送される。

【０３７１】

IN[x] ≠ -xかつIN[x]＜0の時、APP_gは転送エージェントを経由したデータ送信のターゲットである。入力パス215 APP_g → APP_xを介して到着するデータは$APP_{-IN[x]}$より提供され、セルフォーマットCF2 703を有する。

【０３７２】

両方のテーブルにおいて、正の数がセルフォーマットCF1 702に関連し、負の数がセルフォーマットCF2 703に関連することに留意する。

【０３７３】

アービトレーション処理自体は、特定のジオグラフィック・アドレス106を持つコンピュータ装置101の存在する/しないに依存しないことに留意しなければならない。パス215の可用性マトリクスA[1..N,1..N]は、対応パス215を無効であるとしてマークすることで、すでに情報を取り扱っている。

【０３７４】

例示的なアービトレーション処理は以下のステップからなる。

【０３７５】

アービトレーションステップ１：送信リクエスト・コードがテーブルRCODE[1..N,1..N]において利用可能である。シングルビット数A[1..N,1..N]の送信パス可用性マトリクスは利用可能なパス215の数で埋められる。利用可能なパス215は1、利用不可能なパス215は0となる。

【０３７６】

アービトレーションステップ２：ソースAPP_sおよびターゲットAPP_tの両方において、RCODE[s,t]の符号化値が、セル601の必要な送信数にコンバートされ、R[s,t]に格納される。

【０３７７】

アービトレーションステップ３：全てのパス215に、直接接続がデフォルトとして指定

される。

```
tが1からNについて、同時に：
    OUT[t] := t
sが1からNについて、同時に：
    IN[s] := s
```

【０３７８】

アービトレーションステップ４：テーブルRの各R[s,t]位置のために同時に実行される。

R[s,t] > 0の時、対応するソース-ターゲットのペアのための１つの送信リクエストがAPP$_s$ からAPP$_t$へのダイレクト送信パス215に割り当てられる。ダイレクト接続は全てのパス215のデフォルト設定によって成立するので、IN[1..N]もしくはOUT[1..N]を変更する必要はない。割り当てられたパス215はビジー状態とマークされ、R[s,t]値はデクリメントされる：

全てのs値およびt値について同時に：

```
if R[s,t]>0 then A[s,t] := 0
if R[s,t]>0 then dec(R[s,t])
```

【０３７９】

アービトレーションステップ５：ジオグラフィック・アドレス106 mが割り当てられた各コンピュータ装置101に、転送エージェント機能に関与するように割り当てる初期化ステップ。送信元コンピュータ装置101sはジオグラフィック・アドレス106を割り当てられ、転送エージェント機能に関して周期的処理の後行タスクに関与させられる。最初の送信先コンピュータ装置101tは、送信元コンピュータ装置101の周期的処理の後行タスクに割り当てられる。ジオグラフィック・アドレス106 fを有するコンピュータ装置101をターゲットの候補として割り当てることは禁止される。なぜなら、転送エージェントが別のパラレルな実行処理のソースに割り当てられる場合に、関連するパスが割り当てられることが可能であるからである。

```
m := g
s := nxt(m)
t := nxt(s)
f := pred(m)
```

【０３８０】

アービトレーションステップ６：各コンピュータ装置101が転送エージェントとして段階的に並行して計算される場合。各アービトレーションステップは、完全に並行して実行された先行ステップの更新された値を使用する。各ステップにおいて、全ての転送エージェントが異なるソースに対して検証されることに留意する。

A[s,m]>0のとき、tは現在の値で開始し、R[s,t])>0 and A[m,t]>0 and t ≠ fとなるまで、t:=nxt(t)が適用される。これは現在の１つのステップ内で行われる。そのような値が存在しない場合はt:=0となる。tの値が見つかった場合、転送エージェントが同定され、そしてリンク221が以下の通りに指定される：

```
if t>0 then assign:
        if s=g then OUT[m] := t
        if m=g then IN[s] := t
        if m=g then OUT[t] := -s
        if t=g then IN[m] := -s
        dec(R[s,t])
        A[s,m] := 0
        A[m,t] := 0
```

【０３８１】

アービトレーションステップ７：次段の送信元コンピュータ装置101および対応する禁止ターゲットが割り当てられる。値は転送エージェント毎に個々に異なる。

s := nxt(s)
t := nxt(s)
f := pred(f)

【０３８２】

アービトレーションステップ８：
is ≠ m ならば「アービトレーションステップ６」に進む
is ≠ mでないならば終了。

【０３８３】

アービトレーションステップ６について、さらに説明が必要である：

【０３８４】

各コンピュータ装置101は、各転送エージェント役のコンピュータ装置101と一緒にこのアービトレーションステップを同時に実行する。しかしINおよびOUTテーブルは結果がAPP_gに関する場合にのみ維持される。

【０３８５】

上述のアービトレーション処理は、さまざまな送信リクエストや送信先コンピュータ装置101の割り当てに依存する送信要求の達成に関して、著しい不均衡の可能性を負うものである。これらの影響は、アービトレーションが２回または３回の反復処理に細分化される場合に、著しく軽減する。アービトレーションの最初の２回の反復ステップは、ソースからターゲットへの関係１回につき、それほど多くない数の送信を割り当てるだけである。

【０３８６】

以下の規則は送信シーケンス割り当てを統制する。

【０３８７】

N個のコンピュータ装置101のフルメッシ・ネットワーク105において、コンピュータ装置101はN1個の出力パス215およびN1個の入力パス216、およびそれらのパス毎のプロトコル非依存型出力/入力キューを提供する。各パスはそれぞれ、一貫性のあるシーケンスの仕様が、任意の２つのコンピュータ装置101間の複数ルート送信を忠実にサポートすることを必要とする。

【０３８８】

送信機が送信先コンピュータ装置101へのダイレクト送信パスに加えて、転送エージェントを経由するパスに割り当てられた場合、送信機はターゲット向けられているデータのセグメントを分配する。その結果最初のセグメントはダイレクト接続パスに割り振られ、次いで次の複数のセグメントはそれぞれ複数の転送エージェント役コンピュータ装置101のジオグラフィック・アドレス106の昇べきの順に割り振られる。送信先コンピュータ装置101は、受信したデータを、各送信元コンピュータ装置101から受信したデータのための入力キュー1405に格納する。ダイレクト接続パス215を介してくるセグメントのデータがはじめに受信され、続いて、複数の転送エージェント役のコンピュータ装置101に関連するデータ・セグメントが、ジオグラフィック・アドレス106の昇順で受信される。

【０３８９】

まとめると、コンピュータ装置および方法は、全二重データ送信リンクによるフルメッシュ相互接続されうる複数のコンピュータ装置間において、セルロック中のデータの送信を取得し維持するために説明された。また、ネットワーク上のアービトレーション関係情報を広く配信する装置や、複製されたアービトレーション処理の結果を利用してマルチパス・データ送信を実行するように制御される装置のリソースも説明された。集中制御を行わずに、ネットワーク全体においてセル送信を名目上同時に開始することを可能にするため、セルとして送信される所定の数のシンボルに次いで、可変数のアイドル・シンボルが送信される。セルの特定の位置において、各コンピュータ装置は、送信リクエストや受信機性能、ブロックされたリソースを含むリストを他の全てのコンピュータ装置にブロードキャストする。相互接続している各コンピュータ装置は、同じ送信リクエスト、受信機性能およびブロックされたリソースのデータセットに基づいて、同じアービトレーション処

理を実行する。結果として、送信パスはダイレクト送信やペイロード転送に割り当てられる。各リンクのいずれの方向についても、個々のセル期間毎に送信パスを割り当て可能である。セル送信レイヤに割り当てられた複数のパケット・プロトコルが、ネットワーク上で共存可能である。

【０３９０】

本発明はいかなるメッシュ型ネットワークにおいても実装もしくは利用可能でありうる。説明されたメモリデバイスは装置の各リンクのために受信機に実装されてもよい。

【０３９１】

上記コンピュータ装置(101)の実装のためのデザインプロセスは、ハードウェア記述言語(HDL)で書かれる。または図式デザインツールを介して生成されるソースコードの生成から始まる。ソースコードは２つのレベルで存在してもよく、第１のレベルは振る舞いレベルのコード（behavioural level code）であり、第２のレベルはレジスタ送信レベル（register transfer level：RTL）コードである。上記コンピュータ装置(101)のソフトウェア表現は、両方のレベルにおいて生成されうる。シミュレーション・ソフトウェアを備えるコンピュータがデザインの機能性検証のために使用される。シミュレーションのためのデータセットは１つまたは複数該コンピュータ装置(101)のソフトウェア表現のインスタンスを含む。つまり、コンピュータ装置(101)の複数のインスタンス化を用いたシステム・レベル・シミュレーションは、ネットワーク(410)上における該コンピュータ装置(101)の上述のオペレーションに対応する。RTLレベル・ソース・コードは、合成(synthesis)とよばれるプロセスにおいて、要素コンポーネントのリストから構成されるデータセットおよびターゲット生産技術のための相互接続のリストに変換されうる。ターゲット技術はプログラマブルロジックデバイス（programmable logic device: PLD）であってもい。PLDにおいては、データセットが、製造プロセスにおいてPLDを設定するために利用されるビットストリームとしてあらわれたり、ストレージ・デバイスによって提供されたりする。データセットは、PLDを備える要素の電源投入時にPLDにアップロードされ、PLDを構成するために用いられる。設定されたPLD、もしくは該ターゲット生産技術における該データセットを用いて組み立てられたデバイスは、該コンピュータ装置(101)のインスタンス化である。

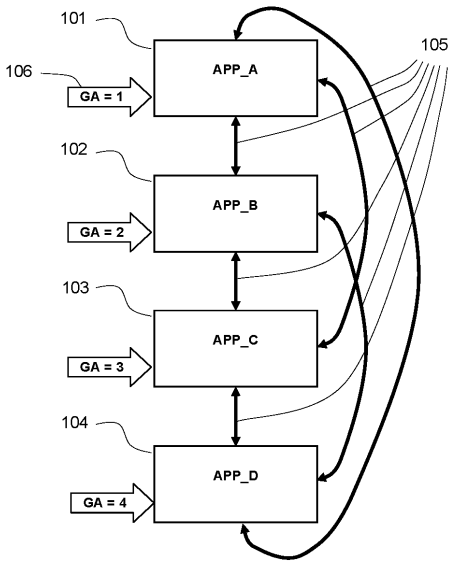
【０３９２】

複数の実施形態が、ハードウェア、ソフトウェアおよびそれらの組み合わせにおいて実現されうる。実施形態は、１つの処理システムにおいて集中的に、もしくは、様々な要素が相互接続された複数の処理システムに拡張されている分散方式において実現されうる。本発明に記載の方法を実行するのに適したいかなる種類の処理システムまたはその他の装置も、要求を満たす。ハードウェアおよびソフトウェアの典型的な組み合わせは、アプリケーションを伴う処理システムであり、アプリケーションがロードされ実行される場合に、アプリケーションが処理システムを制御することで本発明に記載の方法が実行される。複数の実施形態はまた、アプリケーション製品に組み込まれることができ、アプリケーション製品は本発明に記載の方法の実装を可能にする全ての機能を含み、また、アプリケーション製品は処理システムにロードされる場合にこれらの方法を実行可能である。

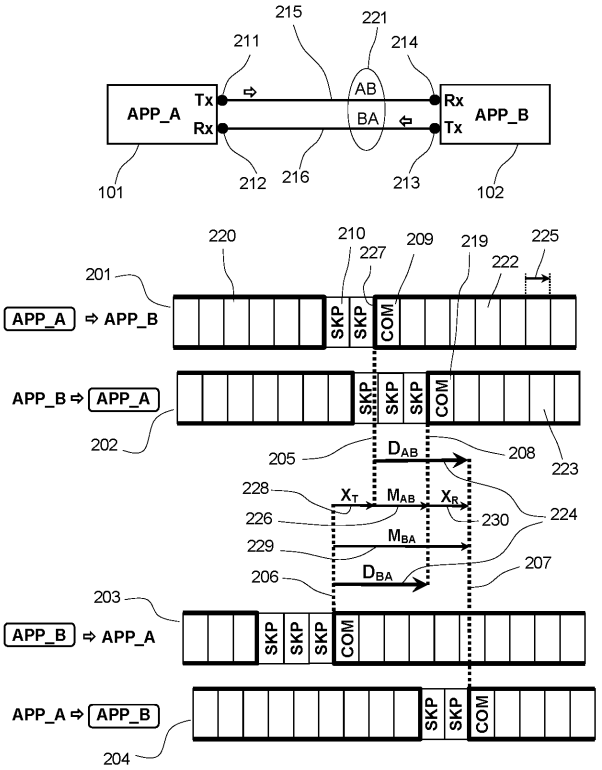
【０３９３】

「ある」という語は1または1以上を意味する。本発明において「複数」という語は2または2以上を意味する。本発明において「別の」という語は少なくとも第２以上を意味する。本発明において「含む」及び／又は「有する」という語は備えるということを意味する（つまりオープンランゲージである）。従って、上述の所定の複数の実施形態は、以下の特許請求の範囲内で変化することができる。

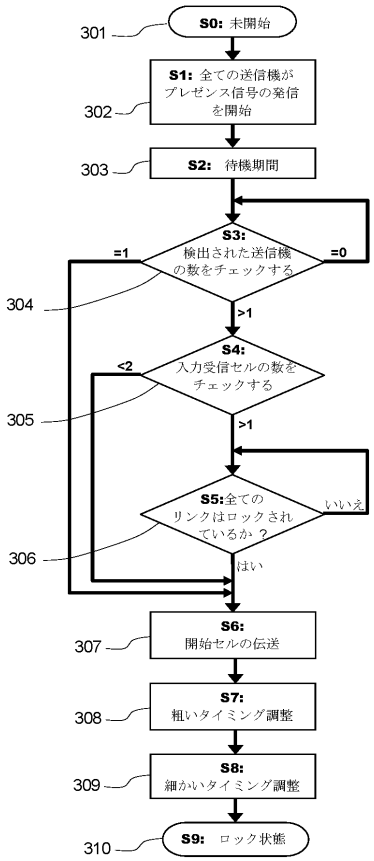
【図１】


101
APP_A
105
106
GA = 1
102
APP_B
GA = 2
103
APP_C
GA = 3
104
APP_D
GA = 4

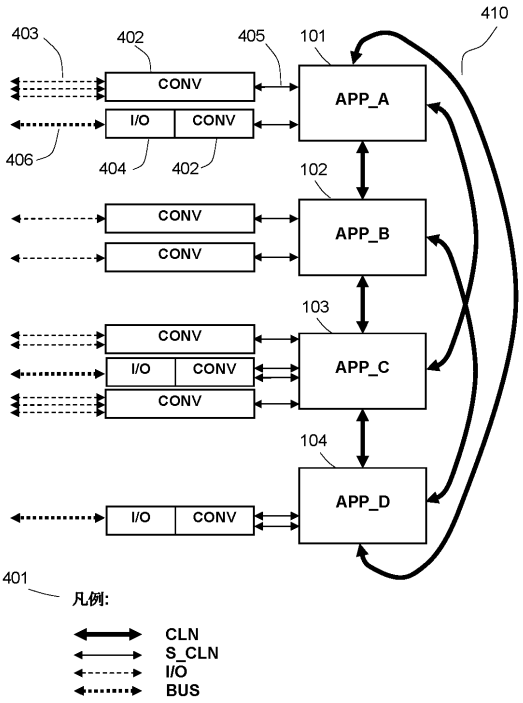
【図２】


215
221
211
214
APP_A
Tx
Rx
AB
BA
Rx
Tx
APP_B
101
212
216
213
102
220
210
227
209
219
225
222
201
APP_A ⇨ APP_B
SKP
SKP
COM
APP_B ⇨ APP_A
SKP
SKP
SKP
COM
202
205
DAB
208
223
228
XT
MAB
XR
226
230
224
MBA
229
207
DBA
206
203
APP_B ⇨ APP_A
SKP
SKP
SKP
COM
APP_A ⇨ APP_B
SKP
SKP
COM
204

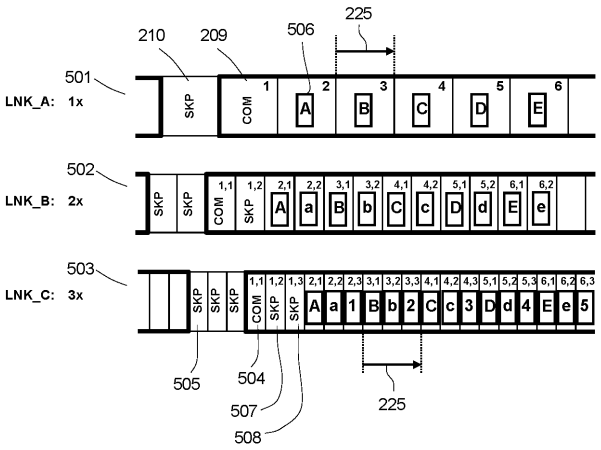
【図３】


301
S0: 未開始
S1: 全ての送信機がプレゼンス信号の発信を開始
302
S2: 待機期間
303
S3: 検出された送信機の数をチェックする
=1
=0
304
>1
S4: 入力受信セルの数をチェックする
<2
305
>1
S5:全てのリンクはロックされているか？
いいえ
306
はい
S6: 開始セルの伝送
307
S7: 粗いタイミング調整
308
S8: 細かいタイミング調整
309
S9: ロック状態
310

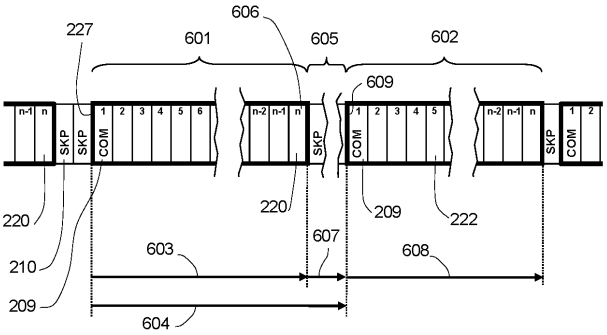
【図４】


403
402
405
101
410
CONV
I/O
CONV
APP_A
406
404
402
102
CONV
CONV
APP_B
103
CONV
I/O
CONV
CONV
APP_C
104
I/O
CONV
APP_D
401
凡例:
CLN
S_CLN
I/O
BUS

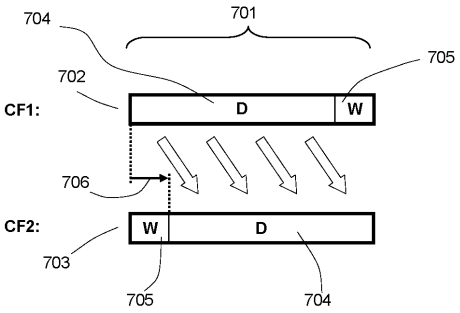
【図５】


LNK_A: 1x
LNK_B: 2x
LNK_C: 3x
SKP
COM
501
502
503
210
209
506
225
505
504
507
508

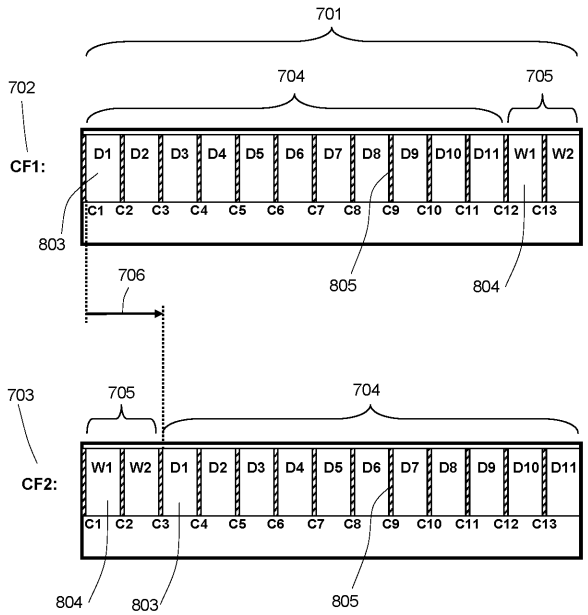
【図６】


227
601
606
605
602
609
220
210
209
607
603
608
604
222

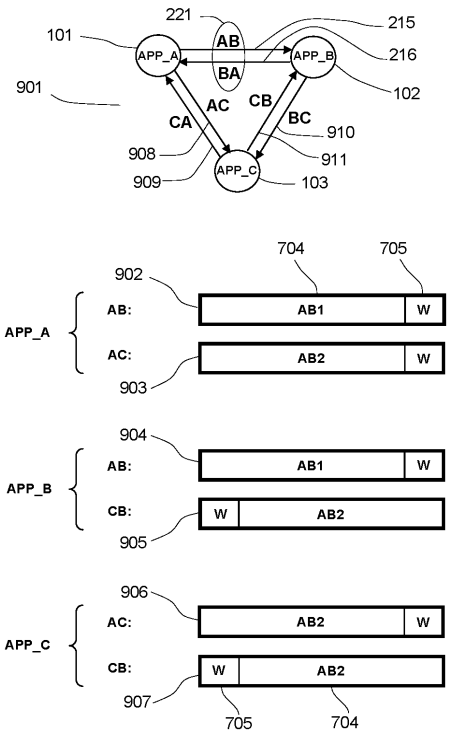
【図７】


CF1:
CF2:
D
W
701
702
704
705
706
703

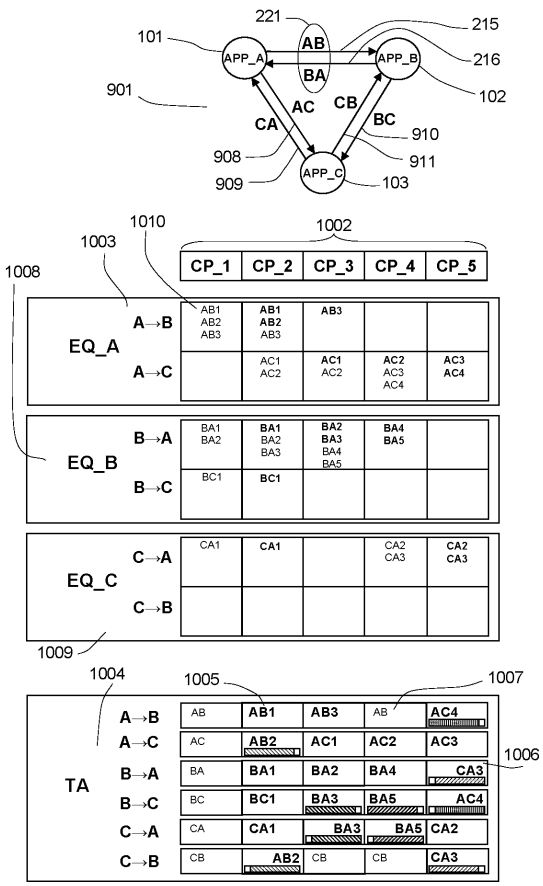
【図８】


CF1:
CF2:
D1 D2 D3 D4 D5 D6 D7 D8 D9 D10 D11 W1 W2
C1 C2 C3 C4 C5 C6 C7 C8 C9 C10 C11 C12 C13
701
702
703
704
705
706
803
804
805

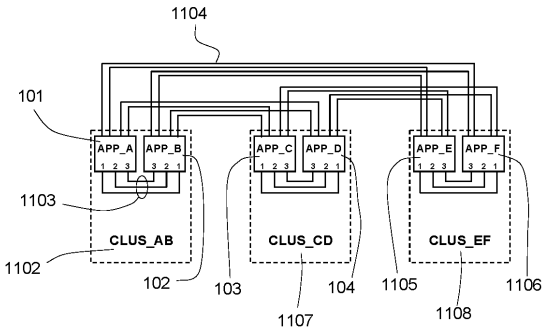
【図９】


APP_A
APP_B
APP_C
AB
BA
AC
CA
CB
BC
101
102
103
215
216
221
901
908
909
910
911
APP_A
AB:
AC:
APP_B
AB:
CB:
APP_C
AC:
CB:
AB1
AB2
W
902
903
904
905
906
907
704
705

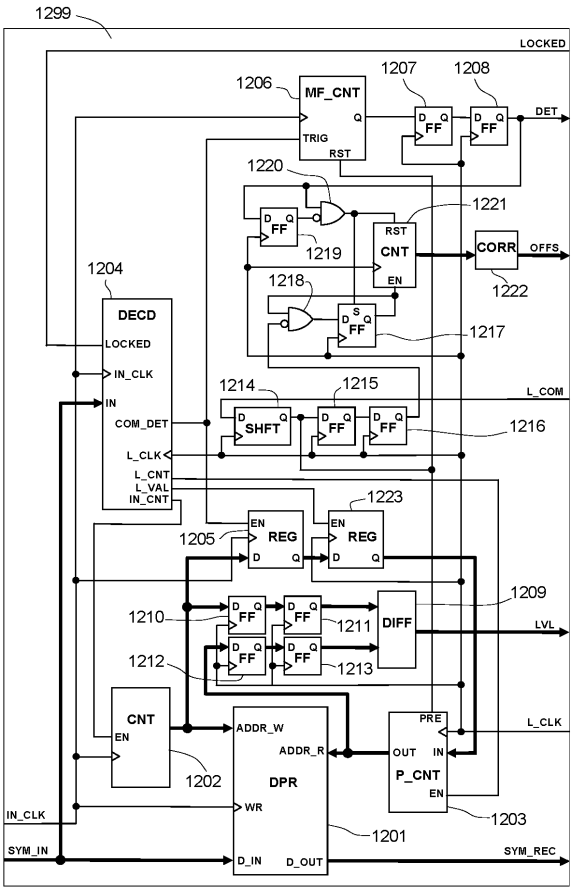
【図１０】



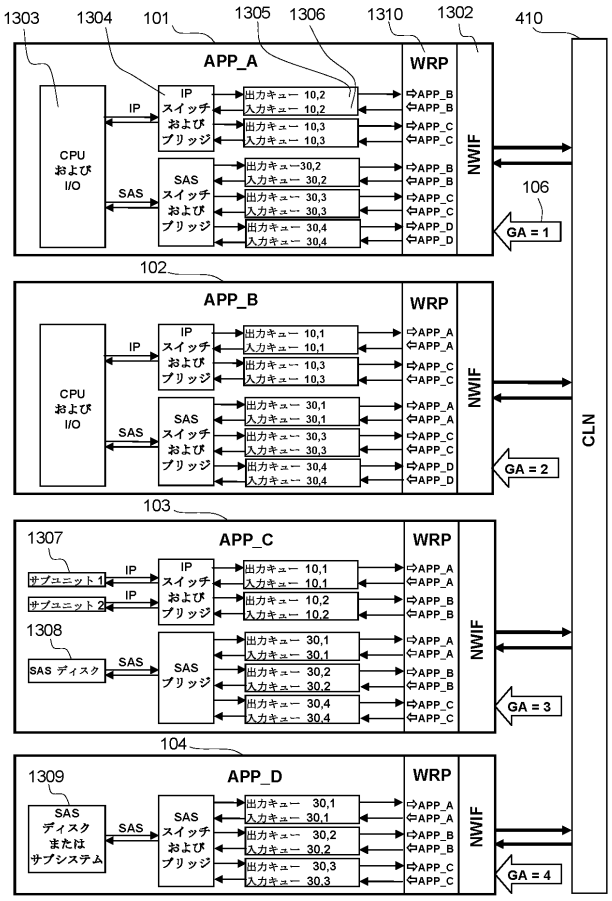
【図１１】



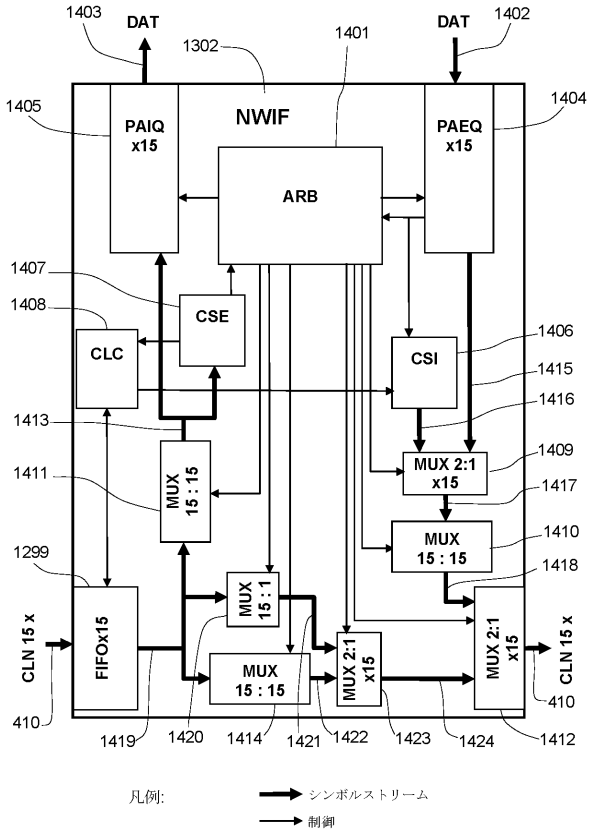
【図１２】



【図１３】



【図１４】



【図１５】

階層の概要	
階層	例
フルメッシュ相互接続	信号技術の例: バックプランにおける異なる信号のペア ツイストペアケーブル・インターフェース （例：CAT5 オプティカル・インターフェース）
シンボル送信	差動信号符号化例： 8 ビット/10 ビットの平衡容量結合 64 ビット/66 ビットの平衡容量結合 128 ビット/130 ビットの平衡容量結合 磁気結合、 1000BASE-T 様な符号化
セルロックの構造	クロッキングおよびその他オプション: タイミング基準としての外部の 8kHz ネットワーク全体における 8kHz ディストリビューション コンピュータ装置 101 間の 2 倍のシンボル・レート セル 601 あたり 1 個のシンボル 222 による埋め込みパケット 埋め込みパケットを経由したシステム・マネジメント・チャネル
ペイロード転送のための調停の分配	転送オプション: セル期間 **603** 内 次のセル期間 **603** 内 帯域幅制御優先順
上位レベルのプロトコル・キュー	プロトコル・キューの例: Anonymous, s-CLN405 へのインターフェース Ethernet InfiniBand SAS SATA
標準物理インターフェース	物理インターフェースの例: 1000BASE-T、10GBASE-T 等 InfiniBand 1X、4X、12X SAS、SATA HyperTransport RapidIO

【図１６】

ある実施形態におけるセル 601 内におけるシンボル位置の割り当て一覧

シンボル位置	セルフォーマット CF1 **702**	セルフォーマット CF2 **703**
1	**C1**	
2-25	D1-D24	**W1-W24**
26	**C2**	
27-50	D25-D48	D1-D24
51	**C3**	
52-75	D49-D72	D25-D48
76	**C4**	
77-100	D73-D96	D49-D72
101	D97	**W25**
102-125	D98-D121	D73-D96
126	D122	D97
127-150	D123-D146	D98-D121
151	D147	D122
152-175	D148-D171	D123-D146
176	D172	D147
177-200	D173-D196	D148-D171
201	D197	D172
202-225	D198-D221	D173-D196
227	D222	D197
228-250	D223-D246	D198-D221
251	D247	D222
202-275	D248-D271	D223-D246
221	D272	D247
222-300	D273-D296	D248-D271
241	D297	D272
242-325	D298-D321	D273-D296
261	D322	D297
262-350	D323-D346	D298-D321
351	D347	D322
352-375	D348-D371	D323-D346
376	**W1**	D347
377-400	D372-D395	D348-D371
401	**C5**	
402-425	D396-D419	D498-D516
426	**C6**	
427-450	D420-D443	D517-D535
451	**C7**	
452-475	D444-D467	D536-D554
476	**C8**	
477-500	D468-D491	D555-D573

シンボル位置	セルフォーマット CF1 **702**	セルフォーマット CF2 **703**
501	**C9**	
502-525	D492-D515	D468-D491
526	**C10**	
527-550	D516-D539	D492-D515
551	**C11**	
552-575	D540-D563	D516-D539
576	**C12**	
577-600	D564-D597	D540-D563
601	**C13**	
602-625	D598-D611	D564-D597
626	**C14**	
627-650	D612-D635	D598-D611
651	**C15**	
652-675	D636-D659	D612-D635
676	**C16**	
677-700	D660-D683	D636-D659
701	**C17**	
702-725	D684-D607	D660-D683
726	**C18**	
727-750	D708-D731	D684-D607
751	**C19**	
752-775	D732-D755	D708-D731
776	**C20**	
777-800	D756-D779	D732-D755
801	**C21**	
802-825	D780-D803	D756-D779
826	**C22**	
827-850	D804-D827	D780-D803
851	**C23**	
852-875	D828-D851	D804-D827
876	**C24**	
877-900	D852-D875	D828-D851
901	D876	**W26**
902-925	D877-D900	D852-D875
926	D901	D876
927-950	D902-D925	D877-D900
951	D926	D901
952-975	D927-D950	D902-D925
976	D951	D926
977-1000	D952-D975	D927-D950
1001	**W2**	D951
1002-1025	**W3-W26**	D952-D975

【図１７】

細かいタイミング調整状態 S8 309 およびセルロック状態 S9 310 での実施例のリンク 221 における、制御シンボル位置 805 の割り当て

制御シンボル位置	シンボル	割り当て CF1 **702 および** CF2 **703**	状態におけるパスへの共通出力値：
C1	COM	セル開始シンボル **209**	S7, S8, S9
C2	PST, RST, Data	このシンボル位置を経由したパケット送信。図 19 を参照。	S8, S9
C3	PST, RST, Data	APP_{gacpr} から C2 で受信したシンボルの送信	S9
C4	Data	8 ビット符号付き整数：受信したセル開始シンボル 219 および自分のセル開始シンボル出力 209 との間の測定されたオフセット 226。 送信コンピュータ装置 101 の地点を反映した値の割り当て。 0000 0000：最適なタイミング調整 0000 0001 – 0111 1111: 自分の COM の後に検出された受信した COM 0000 0001: 1 シンボル期間 225 のオフセットを表す 1000 0001 – 1111 1111: 自分の COM の前に検出された受信した COM 1111 1111: 1 シンボル期間 225 のオフセットを表す	S8, S9 個別の値
		1000 0000: このリンク 221 および図 18 のために認識された粗いタイミング調整状態 S7 **308** がこのテーブルのかわりに適用する。	S7
C5	Data	APP_1 … APP_5 における受信機の可用性	S7, S8, S9
C6	Data	APP_{gacpr} から C5 で受信したシンボルの転送	S9
C7	Data	APP_5 … APP_{16} における受信機の可用性	S7, S8, S9
C8	Data	APP_{gacpr} から C7 で受信したシンボルの転送	S9
C9	Data	APP_1 (bits 7..4)および APP_2 (bits 3..0)への送信要求	S9
C10	Data	APP_{gacpr} から C9 で受信したシンボルの転送	S9
C11	Data	APP_3 (bits 7..4)および APP_4 (bits 3..0)への送信要求	S9
C12	Data	APP_{gacpr} から C11 で受信したシンボルの転送	S9
C13	Data	APP_5 (bits 7..4)および APP_6 (bits 3..0)への送信要求	S9
C14	Data	APP_{gacpr} から C13 で受信したシンボルの転送	S9
C15	Data	APP_7 (bits 7..4)および APP_8 (bits 3..0)への送信要求	S9
C16	Data	APP_{gacpr} から C15 で受信したシンボルの転送	S9
C17	Data	APP_9 (bits 7..4)および APP_{10} (bits 3..0)への送信要求	S9
C18	Data	APP_{gacpr} から C17 で受信したシンボルの転送	S9
C19	Data	APP_{11} (bits 7..4)および APP_{12} (bits 3..0)への送信要求	S9
C20	Data	APP_{gacpr} から C19 で受信したシンボルの転送	S9
C21	Data	APP_{13} (bits 7..4)および APP_{14} (bits 3..0)への送信要求	S9
C22	Data	APP_{gacpr} から C21 で受信したシンボルの転送	S9
C23	Data	APP_{15} (bits 7..4)および APP_{16} (bits 3..0)への送信要求	S9
C24	Data	APP_{gacpr} から C23 で受信したシンボルの転送	S9

【図１８】

粗いタイミング調整状態 S7 308 での、実施例のリンク 221 における制御シンボル位置 805 割り当て			
制御シンボル位置	シンボル	割り当て	状態におけるパスへの共通出力値：
C1	COM	セル開始シンボル **209**	S7, S8, S9
C2	Data	図 19 からの値 P2	S7
C3	Data	割り当てなし	
C4	Data	1000 0000: このリンク 221 における粗いタイミング調整状態 S7 308 を識別する。	S7
		その他のすべての値: このテーブルのかわりに図 17 を適用する	S8, S9: 個別の値
C5	Data	APP_1 ... APP_8 のための受信機の可用性	S7, S8, S9
C6	Data	割り当てなし	
C7	Data	APP_9 ... APP_{16} のための受信機の可用性	S7, S8, S9
C8	Data	割り当てなし	
C9	Data	図 19 からの値 P4	S7
C10	Data	割り当てなし	
C11	Data	図 19 からの値 P5	S7
C12	Data	割り当てなし	
C13	Data	図 19 からの値 P6	S7
C14	Data	割り当てなし	
C15	Data	図 19 からの値 P7	S7
C16	Data	割り当てなし	
C17	Data	図 19 からの値 P8	S7
C18	Data	割り当てなし	
C19	Data	図 19 からの値 P9	S7
C20	Data	割り当てなし	
C21	Data	図 19 からの値 P10	S7
C22	Data	割り当てなし	
C23	Data	図 19 からの値 P11	S7
C24	Data	割り当てなし	

【図１９】

実施例における C2 シンボル位置のためのパケット・コンテンツ・テーブル 図 17 および図 18 を参照		
パケット内での位置	割り当ておよび利用方法	
P1	RST: パケット開始シンボル	
P2	Bits 3..0: このコンピュータ装置 101 の GA, GA=16.の場合に、コード 0000 が使われる。	
	Bits 7..4: クロックの質:小さいほどよい	0000-0100: 通信の質, 外部ソース
		0101-1001: 通信の質, 内部ソース
		1010: ± 15 ppm
		1011: ± 30 ppm
		1100: ± 50 ppm
		1101: ± 100 ppm
		1110: ± 200 ppm
		1111: ± 300 ppm
P3	I^2C プロトコル・エミュレーションを介するマネジメント・パス	
P4	bits 3..0: APP_1 への接続状態 bits 7..4: APP_2 への接続状態	
P5	bits 3..0: APP_3 への接続状態 bits 7..4: APP_4 への接続状態	
P6	bits 3..0: APP_5 への接続状態 bits 7..4: APP_6 への接続状態	
P7	bits 3..0: APP_7 への接続状態 bits 7..4: APP_8 への接続状態	
P8	bits 3..0: APP_9 への接続状態 bits 7..4: APP_{10} への接続状態	
P9	bits 3..0: APP_{11} への接続状態 bits 7..4: APP_{12} への接続状態	
P10	bits 3..0: APP_{13} への接続状態 bits 7..4: APP_{14} への接続状態	
P11	bits 3..0: APP_{15} への接続状態 bits 7..4: APP_{16} への接続状態	
P12	PST: 開始前シンボル	

【図２０】

実施例における接続状態一覧、図 19 参照		
Bits 3..0 および 7..4	信号、タイミング調整	シンボル・レートのチェック
0000	接続なしまたは自分のジオグラフィック・アドレス	チェック結果なし
0001	接続は検出されたが COM は未検出、状態 S6 **307**	チェック結果なし
0010	-- 未使用	
0011	-- 未使用	
0100	粗いタイミング調整状態 S7 **308**	チェック結果なし
0101	粗いタイミング調整状態 S7 **308**	許容範囲外
0110	粗いタイミング調整状態 S7 **308**	許容範囲内
0111	粗いタイミング調整状態 S7 **308**	許容範囲の 50%内
1000	細かいタイミング調整状態 S8 **309**	チェック結果なし
1001	細かいタイミング調整状態 S8 **309**	許容範囲外
1010	細かいタイミング調整状態 S8 **309**	許容範囲内
1011	細かいタイミング調整状態 S8 **309**	許容範囲の 50%内
1100	ロック状態 S9 **310**	チェック結果なし
1101	ロック状態 S9 **310**	許容範囲外
1110	ロック状態 S9 **310**	許容範囲内
1111	ロック状態 S9 **310**	許容範囲の 50%内

【図２１】

実施例における全機能性テーブル GA は、図 12 の P2 に準拠したジオグラフィック・アドレス **106**						
制御シンボル位置	ビット	ジオグラフィックアドレス ga	ga ≠ GA の場合、リンク $L_{GA,ga}$ が完全動作するか		ga = GA の場合 APP_{GA} が自分をタイミング参照先と宣言	
			0	**1**	**0**	**1**
C5	0	1	いいえ	はい	いいえ	はい
	1	2	いいえ	はい	いいえ	はい
	2	3	いいえ	はい	いいえ	はい
	3	4	いいえ	はい	いいえ	はい
	4	5	いいえ	はい	いいえ	はい
	5	6	いいえ	はい	いいえ	はい
	6	7	いいえ	はい	いいえ	はい
	7	8	いいえ	はい	いいえ	はい
C7	0	9	いいえ	はい	いいえ	はい
	1	10	いいえ	はい	いいえ	はい
	2	11	いいえ	はい	いいえ	はい
	3	12	いいえ	はい	いいえ	はい
	4	13	いいえ	はい	いいえ	はい
	5	14	いいえ	はい	いいえ	はい
	6	15	いいえ	はい	いいえ	はい
	7	16	いいえ	はい	いいえ	はい

【図２２】

実施例における図17の制御シンボル805 C9, C11, C13, C15, C17, C19, C21, C23に適用される送信リクエスト・コード ga は図17の APP_{ga} インデックス GA は図19の P2 に準拠したジオグラフィック・アドレス **106**	
バイナリコード bits 3..0 or 7..4	送信リクエストの APP_{ga} への割り当て ga ≠ GA
0000	送信リクエストなし
0001	1つのセルの送信をリクエスト
0010	2つのセルの送信をリクエスト
0011	3つのセルの送信をリクエスト
0100	4つのセルの送信をリクエスト
0101	5つのセルの送信をリクエスト
0110	6つのセルの送信をリクエスト
0111	8つのセルの送信をリクエスト
1000	10つのセルの送信をリクエスト
1001	15つのセルの送信をリクエスト
1010	25つのセルの送信をリクエスト
1011	40つのセルの送信をリクエスト
1100	60つのセルの送信をリクエスト
1101	90つのセルの送信をリクエスト
1110	130つのセルの送信をリクエスト
1111	180つのセルの送信をリクエスト
Bits 3..0 or 7..4	ga = GA の場合割り当てる
Bits 3..0 of gacpr	周期的処理における先行装置と識別されるコンピュータ装置101のジオグラフィック・アドレス **106**： APP_{gacpr}