

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4504606号
(P4504606)

(45) 発行日 平成22年7月14日(2010.7.14)

(24) 登録日 平成22年4月30日(2010.4.30)

(51) Int.Cl.

F I

H04L 12/56 (2006.01)

H04L 12/56 200C

請求項の数 9 (全 23 頁)

(21) 出願番号	特願2001-500501 (P2001-500501)	(73) 特許権者	500477458
(86) (22) 出願日	平成12年5月26日 (2000.5.26)		ネットワーク・イクイPMENT・テクノロ
(65) 公表番号	特表2003-501885 (P2003-501885A)		ジーズ・インコーポレイテッド
(43) 公表日	平成15年1月14日 (2003.1.14)		アメリカ合衆国 カリフォルニア州 94
(86) 国際出願番号	PCT/US2000/014549		555 フリーモント パセオ パードレ
(87) 国際公開番号	W02000/074321		パークウェイ 6550
(87) 国際公開日	平成12年12月7日 (2000.12.7)	(74) 代理人	100059959
審査請求日	平成19年5月28日 (2007.5.28)		弁理士 中村 稔
(31) 優先権主張番号	60/136,953	(74) 代理人	100067013
(32) 優先日	平成11年5月28日 (1999.5.28)		弁理士 大塚 文昭
(33) 優先権主張国	米国 (US)	(74) 代理人	100082005
			弁理士 熊倉 禎男
		(74) 代理人	100065189
			弁理士 穴戸 嘉一

最終頁に続く

(54) 【発明の名称】 ネットワークスイッチにおいてトラフィックを成形する装置及び方法

(57) 【特許請求の範囲】

【請求項1】

成型器ユニットであって、

複数のセル記述子を受信し処理する処理ブロックと、接続IDから形状IDへのマッピングテーブルを保持している形状RAMと、“後刻”リストの先頭及び末尾を格納しているCOIN RAMと、当該複数のセル記述子を格納するデータバッファアレイとを具備するセル記述子(CD)処理ブロックと、

複数の形状IDを処理する処理ブロックと、形状毎のGCRA構成及び状態データを格納する総称セルレートアルゴリズム(GCRA)RAMと、カレンダー待ち行列リンクドリストアレイを格納するリンクRAMとを具備する形状ID処理ブロックと、

を備え、

上記セル記述子処理ブロックは、受信した上記複数のセル記述子に対応する上記複数の形状IDを上記形状ID処理ブロックへ出力し、

上記形状ID処理ブロックは、上記複数の形状IDそれぞれの適合性を検査し、適合した形状IDを上記セル記述子処理ブロックへ返送し、そして、適合していない形状IDをカレンダー待ち行列内へ挿入し、さらに、成熟した形状IDを当該カレンダー待ち行列から当該セル記述子処理ブロックへ返送するように構成されていることを特徴とする成型器ユニット。

【請求項2】

上記形状ID処理ブロックは、上記カレンダー待ち行列を格納するMINT RAMを更に

10

20

備えていることを特徴とする請求項 1 に記載の成型器ユニット。

【請求項 3】

ネットワークスイッチ内でセルトラフィックを成形する方法であって、
セル記述子処理ブロックにおいてセル記述子（CD）を受信するステップと、
上記セル記述子から形状 ID をデコードし、上記セル記述子を“後刻”リスト内に格納するステップと、

上記形状 ID を形状 ID 処理ブロックへ出力するステップと、

上記形状 ID の適合性を検査するステップと、

もし上記形状 ID が適合していれば、上記形状 ID を上記セル記述子処理ブロックへ返送するステップと、

もし上記形状 ID が適合していなければ、上記形状 ID をカレンダー待ち行列内へ挿入し、上記形状 ID が成熟した時に、上記形状 ID を当該カレンダー待ち行列から成熟リストへ転送し、次いで上記形状 ID を上記セル記述子処理ブロックへ返送するステップと、

上記セル記述子処理ブロックが上記形状 ID を受信した時に、上記セル記述子を“今”リストへ移動させるステップと、

上記セル記述子を上記セル記述子処理ブロックから出力するステップと、
を含むことを特徴とする方法。

【請求項 4】

各形状 ID は異なる接続に対応し、上記“今”及び“後刻”リストは、仮想接続（VC）の優先順位を特定することを特徴とする請求項 3 に記載の方法。

【請求項 5】

上記形状 ID が成熟した時には、たとえより高い優先順位の VC が当該形状 ID を生成しなくても、上記セル記述子処理ブロックが VC を決定して送出し、これにより、高い優先順位の VC がより低い優先順位の VC の前に送られることを特徴とする請求項 4 に記載の方法。

【請求項 6】

各接続は、異なるレートで成形されることを特徴とする請求項 3 に記載の方法。

【請求項 7】

複数の接続は全て同一の形状 ID にセットされることを特徴とする請求項 3 に記載の方法。

【請求項 8】

通信システムであって、

情報を供給する複数の源と、

上記複数の源からの上記情報を受信する複数の行先と、

上記複数の源と上記複数の行先とを接続するネットワークを形成する 1 つまたはそれ以上のノードと、
を備え、

上記ネットワークは、上記情報を輸送するための複数のチャンネルを有し、上記ノードのそれぞれはキューイング制御ユニットを含み、上記キューイング制御ユニットは、

待ち行列管理部と、

破棄ブロックと、

成型器と、

を含み、上記成型器は、

セル記述子（CD）処理ブロックと、

形状 ID 処理ブロックと、

を含み、

上記セル記述子処理ブロックは、

複数のセル記述子（CD）を受信し、

上記複数のセル記述子から対応する複数の形状 ID をデコードし、そして、
当該複数のセル記述子を“後刻”リスト内に格納し、

10

20

30

40

50

上記複数の形状 I D を上記形状 I D 処理ブロックに出力し、
上記形状 I D 処理ブロックは、
上記複数の形状 I D それぞれの適合性を検査し、
適合した形状 I D を上記セル記述子処理ブロックへ直接返送し、
適合していない形状 I D をカレンダー待ち行列へ挿入し、成熟した形状 I D を当該カレンダー待ち行列から上記セル記述子処理ブロックへ返送することを特徴とし

上記キューイング制御ユニットは、更に、
上記成型器によって出力された上記複数のセル記述子を受信及び処理するポート毎の待ち行列ユニットと、
上記ポート毎の待ち行列ユニットからの出力を受信及び処理するデキューユニットと、

上記デキューユニットからの出力を受信及び処理するマルチキャストサーバーと、
上記マルチキャストサーバーからの出力を受信及び記憶する自由バッファリストユニットと、
を含むことを特徴とする通信システム。

【請求項 9】

上記セル記述子処理ブロックが上記形状 I D を上記形状 I D 処理ブロックから受信した時に、上記セル記述子処理ブロックは、上記形状 I D に対応するセル記述子を“今”リストへ移動させ、そして上記形状 I D に対応するセル記述子を出力することを特徴とする請求項 8 に記載の通信システム。

【発明の詳細な説明】

【0001】

(技術分野)

本発明は、一般的にはネットワーク通信の分野に関し、特定的にはネットワークスイッチにおけるトラフィックを成形する(シェーピング)装置及び方法に関する。

【0002】

(従来技術)

一般的に、ネットワーク通信システムは、ネットワーク内の多くのユーザを相互接続する。各ユーザは、ポートにおいてネットワークに接続される。ネットワークは多くのノードの相互接続によって形成され、源(ソース)の1ユーザからの入力ポートへの情報入力はネットワークを通してノードからノードへ、出力ポートへ、そして行先(デスティネーション)の別のユーザへ渡される。源から行先へ転送される情報はパケット化され、各ノードは入力ポートにおける入力パケットを出力ポートにおける出力パケットにスイッチする。ATM(非同期転送モード)ネットワークの場合には、パケットは更にセルに分割される。

【0003】

現在の技術では、高速パケットスイッチは、各スイッチポートにおいて毎秒数十万パケットを転送する。各スイッチポートは、広帯域統合サービスデジタル回路網(BISDN)の場合、典型的には50Mビット/秒から2.4Gビット/秒のレートで情報を転送するように設計されている。スイッチサイズは、数ポートから数千ポートまでの範囲である。

【0004】

“高速パケットスイッチ”なる用語は、可変長パケット及び固定長パケットの両方を取扱うことができるスイッチを含む。固定長パケットの使用は、スイッチ設計を簡易化することができる。短い固定長パケット(セル)を使用する高速パケットスイッチを、ATMスイッチと呼ぶ。高速パケットスイッチは、単一の統合回路網内で異なる型の通信サービスを取扱う。これらのサービスは、音声、ビデオ、及びデータ通信を含むことができる。ネットワークを通して伝送される音声及びビデオサービスが許容できる遅延及び遅延変動は制限されるから、ATMスイッチはこれらのサービスに適している。広帯域ISDNのためのATM標準は、5バイトのヘッダー及び48バイトのデータの53バイトの長さを有する

セルを定義している。A T Mフォーラムトラフィック管理仕様は、以下のような多くのサービスクラス定義を指定している。

C B R：連続ビットレート。音声及びビデオのような厳密に制約された遅延及び遅延変動を要求する実時間応用のために。C B Rサービスクラスは、一貫して使用可能な固定された量の帯域幅を要求する。

R T - V B R：実時間可変ビットレート。源が、時間に伴って変化するレートで伝送し（当分野においては“バースティ”という）、しかも厳密に制約された遅延及び遅延変動で受信しなければならないような応用用。

N R T - V B R：非実時間可変ビットレート。遅延またはその変動に関してのサービスは要求しないが、損失に対して感応性を有するバースティ応用用。

U B R：無指定ビットレート。ファイル転送及びeメールのような非実時間応用用。これは、関連するサービスを保証せず、従って割振られた帯域幅資源を用いず、セル損失比またはセル転送遅延に関して保証せず、そして現在のネットワーク輻輳レベルに関する明示フィードバックを用いずに、セルの非連続バーストを伝送する

G F R：保証（された）フレームレート。これも非実時間応用用。このサービスカテゴリは、契約した最小レートで、またはそれ以下でトラフィックを伝送する源のために損失保証を与える。源が契約した最小レートを超えれば、そのレートより上のトラフィックはどのような損失保証も受けられない。

A B R：使用可能ビットレート。ネットワーク内で使用可能な帯域幅の量に依存して、情報転送レートを変化させることができる非実時間応用用。

【 0 0 0 5 】

典型的なA T Mスイッチにおいては、セル処理機能はネットワークのノード内で遂行される。各ノードはA T Mスイッチであり、入力コントローラ（I C）、スイッチファブリック（S F）、出力コントローラ（O C）、及びノード制御（C）を含む。ノード制御は、接続の確立及び解放、帯域幅予約、バッファリング制御、輻輳制御、保守、及びネットワーク管理を含む機能のために使用される。

【 0 0 0 6 】

各スイッチにおいては、入力コントローラは典型的に同期しているので、入力コントローラからの全てのセルは同時にスイッチファブリックに到着し、セルはそれらの優先順位に従って受入れられたり、または拒絶されたりすることができる。スイッチファブリックを通るトラフィックはスロットされ、スイッチファブリック遅延は、タイムスロット持続時間、パイプライン遅延、及びキューイング遅延の合計に等しい。

【 0 0 0 7 】

ノード制御は、スイッチファブリックをバイパスする直接通信経路によって、またはスイッチファブリックを通して伝送される制御セルを介しての何れかによって、入力コントローラ及び出力コントローラと通信する。

【 0 0 0 8 】

スイッチへの外部接続は、一般的に双方向である。双方向接続は、入力コントローラ（C）及び出力コントローラ（O C）を互いにグループにしてポートコントローラ（P C）を形成させることによって形成する。

【 0 0 0 9 】

仮想チャンネル内のセルの入力シーケンスはスイッチファブリックにまたがって保存されるので、各仮想チャンネル内のセルの出力シーケンスは入力シーケンスと同一である。セルは、セルヘッダー内にそのセルが属している接続を識別する仮想チャンネル識別子（V C I）を含んでいる。各セルのヘッダー内の各入力V C Iは入力コントローラにおいて変換され、出力識別子を指定する。この変換は入力コントローラにおいて、典型的には、入力V C Iを使用してテーブルを表引きし、接続テーブルをアドレスすることによって遂行される。この接続テーブルは、その接続がルートされているスイッチファブリックの出力ポートを指定するルート指定フィールドをも含んでいる。接続の優先順位、サービスのクラス、及びトラフィックの型のような他の情報も、接続毎に接続テーブル内に含ませるこ

10

20

30

40

50

とができる。

【 0 0 1 0 】

A T Mスイッチにおいては、セルの到着はスケジュールされていない。典型的な動作では、同一の出力ポートを各々が要求している複数のセルが異なる入力ポートに同時に到着するかも知れない。これらの要求が出力ポートの出力容量を超えるような動作を、出力競合（回線争奪）と称する。出力ポートは一時に固定された数（例えば、1つ）のセルしか伝送できないから、伝送のためには固定された数のセルだけしか受入れることができず、従ってそのポートへルートされた他のセルは破棄するか、または待ち行列内にバッファしなければならない。スイッチモジュールを通るセルをルート指定するために、例えば自己ルート指定及びラベルルート指定のような、異なる方法が使用される。

10

【 0 0 1 1 】

自己ルート指定ネットワークは、各セルにルート指定タグをプレフィックスする入力コントローラと共に動作する。典型的には、入力コントローラはルート指定テーブルからルート指定タグを入手する表引きを使用する。ルート指定タグはセルが送給される出力ポートを指定する。各スイッチング要素は、ルート指定タグを検査することによって迅速にルート指定決定を行うことができる。自己ルート指定ネットワークは、セルが入るスイッチポートには無関係に、要求された行先へ各セルが到着することを保証する。

【 0 0 1 2 】

ラベルルート指定ネットワークは、各スイッチング要素内の変換テーブルを参照する各セル内のラベルと共に動作する。ラベルは各スイッチング要素内で変換され、従ってスイッチング要素の任意のネットワークを使用することができる。

20

【 0 0 1 3 】

スイッチは2つの主要設計、即ち時分割、及び空間分割を有している。時分割スイッチングファブリックでは、全てのセルは、全ての入力及び出力ポートによって共通に共用される単一の通信チャンネルを通して流れる。空間分割スイッチングでは、入力ポートと出力ポートとの間に複数の経路が設けられている。これらの経路は同時に動作するので、スイッチファブリックを横切って同時に多くのセルを伝送することができる。従って、スイッチファブリックの合計容量は、各経路の帯域幅と、セルを同時に伝送できる経路の平均数との積である。

【 0 0 1 4 】

トラフィックロードがネットワーク内の使用可能なシステム資源を越えると、輻輳が発生して性能が劣化する。セルの数がネットワークの輸送容量内にある時には、送給されるセルの数が、輻輳を生じさせずに送られるセルの数に等しいように、全てのセルを送給することができる。しかしながら、もしノードがトラフィックを処理できないレベルまでトラフィックが増加すれば、輻輳が発生する。

30

【 0 0 1 5 】

輻輳は、幾つかの要因によって惹起される。もしネットワーク内のノードがそれらに要求される種々のタスク（キューイングバッファ、更新テーブル等）を遂行するのに遅過ぎれば、たとえ余分なライン容量が存在しているとしても、待ち行列が累積する。一方、たとえノードが無限に速いとしても、入力トラフィックレートが何等かの出力の特定グループのための出力トラフィックレートの容量を超えれば、待ち行列が累積する。

40

【 0 0 1 6 】

もしノードがキューイングセルのための自由バッファを有していなければ、ノードは新たに到着するセルを破棄しなければならない。パケットデータトラフィックの場合、あるセルが破棄されると、破棄されたセルを含んでいたパケットが多分多数回伝送され、輻輳エポックを更に拡張させる。

【 0 0 1 7 】

A T Mスイッチにおいては、あるサービスレートを保証するためにセルの到着をスケジュールする必要はなく、それによって設計者は十分なバッファ空間を設けることができる。惹起される1つの問題は、セルが均一な分布で到着しないことである。事実、殆どのトラ

50

フィックは、それらの間にランダムな持続時間の遅延を設けて伝送される“バースト”(ランダムなサイズを有するセルのグループ)で到着する。より予測可能なデータストリームを形成させるために、セルバーストは、当分野においては“成形器”(シェーパー)として知られているデバイスによって成形される。成形器はセルバーストを受入れて、所定の“形状”に従ってセルを均一に分散させる。異なる形状には異なる仮想チャンネル(VC)が必要であり、従って各VCを独立的に成形する成形器を有することが望ましい。

【0018】

(発明の概要)

一般的に言えば、本発明は、接続毎に成形を行うような、ネットワークスイッチにおいてトラフィックを成形するための装置、及び方法に関する。本発明による成型器は、2つの基本ブロック、即ちセル記述子(CD)処理ブロック、及び形状ID処理ブロックを備えている。CD処理ブロック及び形状ID処理ブロックは、CDの管理をCD出力時点のスケジューリングから切り離すように動作する。CD処理ブロックは、トークン(形状ID)を、形状IDブロックへ出力する。もしこのトークンが確認されれば、それは直ちにCD処理ブロックへ返送され、確認されなければそれは処理されない。トークンが“成熟”していれば、そのトークンはCD処理ブロックへ返送される。次いで、CD処理ブロックはCDを出力する。

【0019】

接続毎の形状IDと共に“今”及び“後刻”リストを使用し、それぞれ仮想接続(VC)及び仮想経路(VP)内の優先権を与える。これは、VP内で成形中の接続の相対的優先順位を保存する。換言すれば、たとえより高い優先順位のVCがトークンを生成しなくとも、それを最初に送ることができ、それによってセルの優先順位が保存される。また、カレンダー待ち行列を使用することによって“仮想終了時”(VFT)計算の複雑さが減少する。

【0020】

(実施の形態)

本発明は、以下の添付図面に基づく詳細な説明から容易に理解されよう。なお、類似構造的要素に対しては類似参照番号を付してある。

【0021】

以下の説明は、当分野に精通している者が本発明を製造し、使用できるようになされるものであり、本発明を遂行するために発明者が企図した最良モードを記述している。しかしながら、以下に本発明の基本的原理を詳細に記載してネットワークスイッチにおいてトラフィックを成形する装置及び方法を説明するので、当分野に精通していれば種々の変更は容易に明白であろう。これら全ての変更、等価、及び代替は、本発明の思想及び範囲内にあることを意図している。

【0022】

先ず図1を参照する。複数のネットワークユーザが、源/行先(S/D)4で表されている。各ユーザは、典型的に、源(S)として情報を送り、行先(D)として情報を受信する。S/Dユニット4の源(S)は、他のS/Dユニット4の行先(D)へ情報を送る。情報を源から行先へ転送するために、各S/Dユニット4はマルチノード(N)ネットワーク1を通して接続されている。ネットワーク1は、多くのノード(N)5を含んでいる。ノードは、ノードからノードへ接続されているから、一般的に、ネットワーク1内にノード5のチェーンを形成することによって、S/Dユニット4のどの特定の1つをも、他のS/Dユニット4のどの1つへも接続することができる。一般的に、S/Dユニット4とノード5との間の接続、及びノード5間の接続は、情報を両方向に転送することを可能にする双方向リンク8によっている。

【0023】

明瞭化のために、図1に図示したノード(N)5の数は比較的少数であるが、ネットワークは数百またはそれ以上のノードを含むことができる。またS/Dユニット4は、Sユーザ、即ち4-0、4-1、4-2、4-3、4-4、...、4-(S-2)、4-(S-1)を含

10

20

30

40

50

む。S の値はどのような整数であることもできるが、典型的には S は数百またはそれ以上である。

【 0 0 2 4 】

典型的な実施の形態では、図 1 の通信システムは A T M ネットワークであり、情報転送の単位はセルである。複数のセルが、情報のパケットを形成する。ネットワーク 1 は、画像、音声、及びデータを含む異なる型の情報を支援するようにセル及びパケットと通信する。

【 0 0 2 5 】

図 2 において、S / D ユニット 4 - x は、複数 C のノード (N) 5 - 0、5 - 1、...、5 - (C - 1) を通して S / D ユニット 4 - y に接続されている。

10

【 0 0 2 6 】

図 2 に示す S / D ユニット 4 - x は、図 1 の S / D ユニット 4 の何れかの典型である。例えば、S / D ユニット 4 - x は、図 1 の S / D ユニット 4 - 2 を表すことができる。同様に、図 2 の S / D ユニット 4 - y は、図 1 の S / D ユニット 4 の何れかを表すことができる。例えば、図 2 の S / D ユニット 4 - y は、図 1 の S / D ユニット 4 - 4 を表すことができる。この例では、ノード 5 - 0、5 - 1、...、5 - (C - 1) は、図 1 の S / D ユニット 4 - 2 を S / D ユニット 4 - 4 に接続するために使用されているネットワーク 1 内の C ノードを表している。

【 0 0 2 7 】

図 2 において、双方向リンク 8 - 0、8 - 1、...、8 - (C - 1)、8 - (C) は、複数 C のノード (N) 5 - 0、5 - 1、...、5 - (C - 1) を通して S / D ユニット 4 - x から S / D ユニット 4 - y に接続されている。図 2 において、情報は、S / D ユニット 4 - x 内の源 (S) から S / D ユニット 4 - y 内の行先 (D) へ転送することができる。同様に、情報を S / D ユニット 4 - y 内の源 (S) から S / D ユニット 4 - x 内の行先 (D) へ転送することができる。情報は、図 2 の何れかの方向へも転送することはできるが、説明の目的から、源 (S) と行先 (D) との間の転送が、S / D ユニット 4 - x から S / D ユニット 4 - y へであるのか、または S / D ユニット 4 - y から S / D ユニット 4 - x へであるのかを考えると便利である。方向には関係なく、各転送は源 (S) から行先 (D) へである。

20

【 0 0 2 8 】

図 3 に、図 2 の仮想チャンネル内での源 (S) から行先 (D) への転送に使用される回路を示す。図 3 において、図 2 の S / D ユニット 4 - x 内の源ユニット 4 - (S) は、図 2 の S / D ユニット 4 - y 内の源ユニット 4 - (D) に接続されている。

30

【 0 0 2 9 】

図 3 において、各リンク 8 - 0、8 - 1、...、8 - (C - 1)、8 - (C) は、情報を順方向に転送するための順方向 (F) チャンネルと、情報を逆方向に転送するための逆方向 (R) チャンネルとを含んでいる。図 3 の逆方向チャンネルは、源ユニット 4 - (S) から行先ユニット 4 - (D) への情報の転送に関している。図 3 の逆方向チャンネルは、図 1 のネットワークに関連して使用される制御情報を送る目的のためのものである。逆方向チャンネル (R) は、図 2 に関連して説明した S / D ユニット 4 - y から S / D ユニット 4 - x への情報の順方向の転送に使用される順方向チャンネル (F) とは区別される。順方向 (F) 及び逆方向 (R) の両チャンネルは、源ユニット 4 - (S) から行先ユニット 4 - (D) への転送に関している。図 3 の各ノードは、順方向 (F) 回路 6 及び逆方向 (R) 回路 7 を含んでいる。図 3 において、順方向チャンネル 8 - 0 P、8 - 1 F、...、8 - (C - 1) F は、入力としてそれぞれ順方向回路 6 - 0、6 - 1、...、6 - (C - 1) に接続されている。順方向チャンネル 8 - (C) F は、ノード 6 - (C - 1) からユニット 4 - (D) まで接続されている。同様に、逆方向チャンネル 8 - 0 R、8 - 1 R、...、8 - (C - 1) R は、逆方向回路 7 - 0、7 - 1、...、7 - (C - 1) から接続されている。逆方向チャンネル 8 - (C) R は、D ユニット - (D) から逆方向回路 7 - (C - 1) まで接続されている。

40

【 0 0 3 0 】

50

図 3 において、各ノード 5 は、順方向 (F) 回路 6 から逆方向 (R) 回路 7 まで接続されているフィードバック接続 9 を有している。詳述すれば、フィードバックチャンネル 9-0、9-1、...、9-(C-1) は、それぞれノード 5-0、5-1、...、5-(C-1) 内の順方向 (F) 回路 6 から逆方向 (R) 回路 7 まで接続されている。図 3 の回路においては、仮想チャンネル接続は、順方向チャンネルセッティングアップに沿って、S ユニット 4-(S) と D ユニット 4-(D) との間に順方向の通信経路を作っている。図 1 のネットワーク 1 内には他の仮想チャンネルも確立されているので、各ノード、及び図 3 のノードを含む行先においてバッファリングが必要である。

【0031】

図 4 に、図 3 の信号経路を有するノードの 1 つの典型的な実施の形態を示す。図 4 において、ノード 5 は、N リンク 18-0、18-1、...、18-n、...、18-(N-1) を含む。図 4 の各リンク 18 は、図 2 の双方向リンク 8 に類似している。図 4 では、リンク 18-0、18-1、...、18-n、...、18-(N-1) は、ポートコントローラ 11-0、11-1、...、11-n、...、11-(N-1) に接続されている。

10

【0032】

図 4 のノードは、例えば、リンク 18 の 1 つ (例えば、図 4 の入力リンク 18-0) を、スイッチファブリック 10 を通して別のリンク 18 の 1 つ (例えば、リンク 18-n) に接続することによって、図 3 の情報転送に関連して使用される。この例では、スイッチファブリック 10 は、リンク 18-0 をリンク 18-n に接続するように機能する。

【0033】

20

図 4 のノードが図 2 のノード 5-1 を表すような例では、図 2 のリンク 8-1 は図 4 のリンク 18-0 であり、図 2 のリンク 8-2 は図 4 のリンク 18-n である。

このような接続では、図 4 のノードは、情報を例えばリンク 18-0 からリンク 18-n へ接続し、また情報を逆方向にリンク 18-n からリンク 18-0 へ接続する。リンク 18-0 及び 18-n は、説明のために任意に選択したものである。N リンク 18 の何れかを、他の何れかのリンク 18 へ接続するために、図 2 の回路内で選択することができる。

【0034】

図 4 のノードを図 3 の仮想チャンネル接続 (源 (S) が左側、行先 (D) が右側) に使用する場合、説明の目的からリンク 18-0 がノード 5 への順方向入力であり、リンク 18-n がノードからの順方向出力であるものとする。

30

【0035】

図 4 において、ポートコントローラ (PC) 11-0、11-1、...、11-n、...、11-(N-1) は、それぞれ入力コントローラ 14-0、14-1、...、14-n、...、14-(N-1) を有し、またそれぞれ出力コントローラ 15-0、15-1、...、15-n、...、15-(N-1) を有している。図 4 において、順方向情報セルは、図 3 の源 4-S からバス 18-01、入力コントローラ 14-0、バス 20-n0、スイッチファブリック 10、バス 20-n1、コントローラ 15-n、及びバス 18-n0 を通して図 3 の行先 4-(D) へ送られる。ポートコントローラは、共用キューイングユニット 51 内に配置されている共通バッファ記憶装置を共用し、バス 41-0、41-1、41-n、...、41-(N-1) を通してユニット 51 に双方向に接続されている。

40

【0036】

図 5 に、図 4 のキューイングユニット 51 の詳細を示す。キューイングユニット 51 は、データ待ち行列ユニット 52 及び待ち行列制御ユニット 53 を含む。データ待ち行列ユニット 52 及び待ち行列制御ユニット 53 は、各々双方向バス 41-0、41-1、...、41-n、...、41-(N-1) に接続されている。バス 41 上の制御情報は待ち行列ユニット 53 へ接続され、バス 41 上のデータはデータ待ち行列ユニット 52 へ接続される。

【0037】

図 5 に示す待ち行列制御ユニット 53 は、データ待ち行列ユニット 52 及びキューイングユニット 51 の総合動作を制御する待ち行列管理者 54 を含んでいる。待ち行列管理者は、典型的に、ソフトウェアを実行することができる処理ユニットを含む。バス 41 上の入

50

力情報がデータ待ち行列ユニット52内への格納を要求すると、待ち行列管理者54は自由バッファリストユニット59から使用可能なバッファ位置を検出し、その使用可能なデータ位置をデータ待ち行列ユニット52に割当てて、待ち行列管理者の一般的機能及び動作は公知である。キューイングに加えて、及び本発明の方法で動作させるために、総合通信ネットワークの効率的な動作を促進するように若干のセルを時折破棄する必要がある。待ち行列管理者54の制御の下にある破棄ユニット55は、先に割振られた待ち行列割当てを破棄する時点を決する。キューイング動作の結果は、ポート毎の待ち行列ユニット56内に格納され、ユニット56自体はデキューユニット57を活動させ、ユニット57自体はマルチキャストサーバーを通して動作して先に割振られたバッファ位置を除去する。除去されると、デキューされたバッファ位置はユニット59内の自由バッファリストへ戻されて追加され、再割当てのために使用可能になる。

10

【0038】

破棄ユニット55は、3つのユニット、即ちFIFOユニット61(サブユニット61-1及び61-2を含む)、破棄ユニット62、及びポインタ完全性ユニット63からなる。破棄ユニット55は、

1. 全ての接続の契約されたサービス品質(QoS)を保証(損なう接続を破棄することによって)
2. バッファ輻輳の監視及び制御
3. バッファが輻輳し始めるとATMヘッダー内に明示順方向輻輳指示(EFCI)のタグ付けを遂行
4. 輻輳が過大になると接続毎にセル及びフレームの破棄を遂行
5. 非保証接続(ABR、GFR、及びUBR)間の公平さの保証
6. 種々のEFCI及び破棄しきい値を支援することによって、ABR、GFR、及びUBRの異なる品質の提供
7. ポインタ完全性検証(ポインタの重複が発生していないことを確認する)に責を負う。

20

【0039】

前述したように、成型器ブロック60はセルバーストを離間させ、セルを均一に分散させる。図6(A)は、サンプル伝送ストリームを示している。セルは1ms離間し、これらがバーストとして知られるグループに束ねられ、バースト間には不規則な遅延が存在している。成型器はセルバーストを受入れて、図6(B)に示すように、セルが均一な3ms間隔で伝送されるようにセルを均一に分配する。

30

【0040】

一般的に言えば、図7に示すように、本発明により構成された成型器60は、2つの基本ブロック、即ちセル記述子(CD)処理ブロック70、および形状ID処理ブロック72を備えている。これらの機能ブロックは、分離したASICとして実現することも、または同一チップ上に実現することもできる。以下に、CD処理ブロック70をDALEK70と称し、また形状ID処理ブロック72をTARDIS72と称する。公知のように、セル記述子(CD)は、各セルを表す記述子である。より効率的な処理を行うために、各セルの代わりに各セル毎のCDが制御経路を通してルートされる。破棄サブシステム55及び成型器60がCDを処理すると、対応するセルがメモリから出力される。CDフォーマットの例を図8に示す。

40

【0041】

DALEK70はCDを格納し、トークン(形状ID)を生成する。形状IDは基本的には、セルの伝送可能なレートを指定する所定の“形状”である。動作中本発明の成型器は、成形されたセルレートをユーザが指定することを可能にし、またはユーザはソフトウェア制御の決定を据え置くことができる。トークンはDALEK70からTARDIS72へ出力される。TARDIS72は形状IDを処理し、トークンをDALEK70へ戻す。詳細を後述するように、DALEK70自体は適切なCDを出力する。

【0042】

50

各セル毎のＣＤ内の接続識別子（接続ＩＤ）から、DALEK 7 0 は適切な形状ＩＤを決定する。TARDIS 7 2 は、各独自の形状ＩＤ毎にセル間の最小時間間隔を指定するテーブルを含んでいる。トークンが“成熟”する（即ち、セルが特定の接続のために出て行くことができるようになる）と、トークンがDALEK 7 0 へ返送される。次いで、DALEK はどのＶＣが優先順位を有しているのかを正確に決定し、セルを送出する。即ち、たとえより高い優先順位のＶＣ上のセルが始めにトークンを生成しなくとも、それが送られる。本発明によれば、特定の接続を、他の接続には無関係に成形することができる。また、多数の異なる接続を、同一の形状ＩＤに従って成形することができる。従って、高及び低優先順位トラフィックを同一の物理的接続内へ送ることができる。

【 0 0 4 3 】

図 9 は、本発明の 1 実施の形態のより詳細なブロック図である。DALEK 7 0 は 3 つの分離したメモリアレイ、即ち形状 RAM 7 0 1、COIN RAM 7 0 2、及びデータ RAM 7 0 3 を使用する。同様に、TARDIS は 3 つのアレイ、即ちGCRA（総称セルレートアルゴリズム）RAM 7 2 1、リンク RAM 7 2 2、及びMINT RAM 7 2 3 と対話する。DALEK 7 0 及びTARDIS 7 2 は、それらの関連 RAM アレイと共に、成型器 6 0 の完全論理機能を実現している。

【 0 0 4 4 】

TARDIS 7 2 とDALEK 7 0 との間関係は、それぞれマスター及びスレーブの一方である。TARDIS 7 2 は 2 つのブロックを接続するインタフェースを制御し、主タイミングシーケンス信号をDALEK 7 0 へ供給する。対話は、形状ＩＤ及び管理データを含む。形状ＩＤはTARDIS 7 2 とDALEK 7 0 との間で交換され、ＣＤの管理をＣＤ出力時点のスケジューリングから切り離す。前者はDALEK 7 0 の責任であり、後者はTARDIS 7 2 の責任である。6 つ（各方向に 3 つ）の形状ＩＤを、各主タイミングシーケンスでDALEK 7 0 とTARDIS 7 2 との間で渡すことができる。

【 0 0 4 5 】

DALEK 7 0 は、TARDIS 7 2 を介して外部 CPU によって管理される。TARDIS 7 2 は、各主タイミングシーケンス毎に 1 回、全てのDALEK 7 0 読み出しレジスタを読み出してローカルコピーを維持し、これらのコピーは CPU によって読み出される。同様に、DALEK 7 0 のために意図されている CPU 書き込みデータは、CPU から到着する 1 主タイミングシーケンス中にTARDIS 7 2 からDALEK 7 0 へ転送される。DALEK 7 0 状態レジスタの若干のビットは、TARDIS 7 2 の割込み出力を表明することができる。これらの各割込み源は、個々に動作可能にされる。TARDIS 7 2 からDALEK 7 0 へ転送される全てのイベントフラグは CPU によって読み出されるまで捕捉され、保持される。DALEK 7 0 とTARDIS 7 2 との間の通信は、共用データバス + 制御信号を使用して達成される。形状ＩＤ及び管理データの両者は、同一のバスを共用する。主タイミングシーケンスに基づく時分割多重化が、全ての要求されたデータの転送のために必要なタイミング及び帯域幅を保証する。

【 0 0 4 6 】

TARDIS ブロック

図 1 0 は、TARDIS 7 2（及び関連 RAM）のブロック図であって、このブロックを通る形状ＩＤトークンのデータの流れを示している。まず、形状ＩＤトークンがDALEK 7 0 から受信され、その適合性が調べられる。適合する形状ＩＤトークンは直ちにDALEK 7 0 へ返送され、一方適合しない形状ＩＤトークンはカレンダー待ち行列内へ挿入される。形状ＩＤトークンはカレンダー待ち行列から“成熟”リストへ転送され、次いで形状ＩＤトークンはDALEK 7 0 へ伝送される。TARDIS 7 2 は、主タイミングシーケンスに同期しているシーケンス（後述）を使用して動作し、シーケンス同期をDALEK 7 0 へ供給する。TARDIS 7 2 によって管理されるデータ構造は、DALEK 7 0 へ直ちに出力するための、1 組のGCRA構成及び状態データ、スケジュールされた形状ＩＤのカレンダー待ち行列リンクド（リンクされた）リストアレイ、及び形状ＩＤの“成熟”リンクドリストを含む。

【 0 0 4 7 】

形状毎のGCRA構成及び状態データは、TARDIS 7 2 によってGCRA RAM 7 2 1 内に維持さ

10

20

30

40

50

れる。構成データは、形状のレートを定義する最小セル間隔を含む。状態データは、スケジュール時及びカウントフィールドを含む。スケジュール時は、次の形状IDトークンの出力時点である。カウントは、現在TARDIS 72内に存在している形状IDトークンの数である。最小セル間隔は、主CPUからアクセス可能である。GCRAデータは、各主タイミングシーケンス中に6回まで、形状IDトークンの出力時点をスケジュールするために使用される。若干のスケジュールされた形状IDトークンは、後述するように、カレンダー待ち行列へ挿入され、他の形状IDトークンは形状のカウントフィールド内に保持される。

【0048】

カレンダー待ち行列リンクドリストアレイは、TARDIS 72によってMINT RAM 723及びリンクRAM 722内に維持される。この構造は、各カレンダー毎に1つの64Kリンクドリストのアレイである。カレンダー待ち行列をリンクされたリストのアレイとして実現すると、複数の形状の形状IDトークンを同時にスケジュールすることができる。MINT RAM 723はリンクされたリストの先頭と末尾を保持する。各スケジュールされた形状IDトークンは、通常は、計算されたスケジュール時のためのカレンダー待ち行列に添付される。若干の環境の下では、形状IDは現時点+1のためのリストに添付される。

【0049】

各主タイミングシーケンス中に、カレンダー時が前進する。新しい現時点のためのカレンダー待ち行列リストは、“成熟”リンクドリストの末尾に転送される。このようにすると、“古い”カレンダー時のためのカレンダー待ち行列リストは自動的に空にされる。“成熟”リンクドリストは、内部論理及びリンクRAM 722を使用するTARDIS 72によって維持される。この構造は、直ちにDALEK 70へ出力するために待ち行列に入れられた形状IDの単一のリンクされたリストである。

【0050】

各主タイミングシーケンス中に、3つまでの形状IDトークンをDALEK 70へ転送することができる。シーケンス中に受信された適合形状IDトークンに優先順位が与えられ、次の順位は“成熟”リンクドリストからの形状IDトークンである。これにより、適合セルのストリームに対して輻輳が与える衝撃を最小にすることができる。カレンダー待ち行列及び“成熟”リンクドリストのためのリンクは、共にリンクRAM 722を使用する。各形状からは、単一の形状IDトークンだけしかスケジュールすることができないから、即ちこれらのリスト構造の何れにも1つしか存在しないから、16Kリンクだけでよい。リンクRAM 722のアドレスは形状IDであり、戻されるデータは同一リスト内の次の形状IDトークンである。図11はカレンダー待ち行列を示し、図12は“成熟”リンクドリスト構造を示している。

【0051】

TARDIS 72内では、時間は、1主タイミングシーケンスの分解能及び64K主タイミングシーケンスの範囲を与える16ビット2進フィールドで表される。現時点は各主タイミングシーケンスの開始時に1回インクリメントされる。最小セル間隔は、主タイミングシーケンスの1/256の分解能、及び64K主タイミングシーケンスの範囲を与える24ビット2進フィールドで表される。間隔の16の最上位ビットは、“整数部”として知られている。間隔の8つの最下位ビットは、“小数部”として知られている。各形状のピークセルレート（PCR）は、レートの逆数である最小セル間隔として定義されている。最小及び最大許容レートを、図13のテーブルに示す。

【0052】

高帯域幅限界は、TARDIS 72によって強化されてはいない。従って、より高い帯域幅（即ち、より小さい最小セル間隔）を有する形状IDが正しく成形される保証はない。成型器60の出力帯域幅が制限されているので、他の成形された接続が存在する場合には、このような形状IDは重大なセル遅延変動を受ける恐れがある。低帯域幅限界は、TARDIS 72によって強化されている。この限界よりも大きい最小セル間隔を有するように構成された形状IDは成形されない（即ち、それはあたかも最小セル間隔が0.001:00であるかのように処理される）。図14は、本発明の1実施の形態により、TARDIS 72内で構成するこ

10

20

30

40

50

とができる最小セル間隔の例を示している。

【 0 0 5 3 】

TARDIS 7 2 内のスケジューリングは、

1 . DALEK 7 0 から形状 I D トークンを受信した時（主タイミングシーケンス内で 3 つまで）、

2 . “ 成熟 ” リストの先頭の形状 I D トークンが DALEK 7 0 へ伝送される時（主タイミングシーケンス内で 3 つまで）、

に遂行される。図 1 5 は、スケジュール動作のための真理値表である。この表の以下の説明では、上述した形状 I D トークンが単一の形状に属していることに注目されたい。TARDIS 7 2 によって支援される 16 K 形状は独立的に処理される。

10

【 0 0 5 4 】

形状 I D トークンが DALEK 7 0 から受信され、TARDIS 7 2 内に形状 I D トークンが存在しない（カウント 0 によって指示される）場合に、スケジューラの “ 最初のイン ” 結果が発生する。“ 最初のイン ” は、形状 I D トークンを DALEK 7 0 へ戻させ（それが適合している場合）、またカレンダー待ち行列内へ挿入させる。更に、カウントがインクリメントされる。これは、TARDIS 7 2 内に “ 実 ” 形状 I D が存在しないにも拘わらず、“ ゴースト ” 形状 I D トークンが TARDIS 7 2 内に存在し続けるという、このアルゴリズムの重要な特性を示している。カウントは、実際に、“ 実 ” 形状 I D トークンの数 + 1 “ ゴースト ” である。

【 0 0 5 5 】

“ 次のイン ” のスケジューラ結果は、DALEK 7 0 から 1 つまたは複数の形状 I D トークンが受信され、TARDIS 7 2 内に既に形状 I D トークンが存在している（非 0 カウントによって指示される）場合に発生する。形状 I D トークン内の “ 次のイン ” 結果は、カウントに対するインクリメントの形状で TARDIS 7 2 内に保持される。この形状は、現在は適合していないので、形状 I D トークンは DALEK 7 0 へは戻されない。また、形状 I D トークンが既に存在しているので、それはカレンダー待ち行列内へ挿入もされない。

20

【 0 0 5 6 】

“ 次のアウト ” のスケジューラ結果は、“ 成熟 ” リストの先頭の形状 I D トークンが DALEK 7 0 へ送られ、TARDIS 7 2 内に複数の形状 I D トークンが存在している（1 より大きいカウントによって指示される）場合に発生する。“ 次のアウト ” により、カレンダー待ち行列内へ形状 I D トークンが挿入され、カウントがデクレメントされる。“ ゴーストアウト ” のスケジューラ結果は、“ 成熟 ” リストの先頭の形状 I D トークンが DALEK 7 0 へ送られ、TARDIS 7 2 内に “ ゴースト ” 形状 I D トークンだけが存在している（1 のカウントによって指示される）場合に発生する。“ ゴーストアウト ” により、カウントは 0 にセットされる。この特別な “ ゴースト ” 形状 I D は、システムへ出力すべき C D が見出されないで DALEK 7 0 によって無視される。

30

【 0 0 5 7 】

形状 I D トークン内の “ 最初のイン ” 及び “ 次のアウト ” スケジューリング結果は、カレンダー待ち行列リスト（スケジュール時のためのリスト）に添付しなければならない。各形状 I D を配置するのは正確に何処であるかの決定は、以下の 2 つの要因によって複雑にされる。

40

1 . カレンダー待ち行列が 64 K のエントリを有しているので、ポインタは規則的に循環（ラップアラウンド）する。

2 . “ 成熟 ” リストにおける輻輳が、スケジュール時を “ 過去 ” に配置することができる。

図 1 6 のテーブルは、カレンダー待ち行列挿入時計算のための真理値表を定義している。もし “ 現時点 ” が選択されれば、形状 I D トークンは（現時点 + 1）カレンダー待ち行列に配置される。次いで、それは、次の主タイミングシーケンスにおいて “ 成熟 ” リストに添付される。

【 0 0 5 8 】

50

TARDIS 7 2 によって遂行される動作シーケンスは、主タイミングシーケンスに密に結合されている。これらのシーケンスを、スケジュール、成熟、及び管理と名付ける。

【 0 0 5 9 】

スケジュールシーケンス

このシーケンスは、形状 I D のスケジューリングを遂行する。このシーケンスは、DALEK 7 0 から形状 I D トークンを受信することによって、または形状 I D トークンを “ 成熟 ” リストから DALEK 7 0 へ伝送することによって開始される。このシーケンスは、形状 I D エントリをカレンダー待ち行列内へ挿入し、据置きカウントを更新する。図 1 7 のテーブルは、このシーケンスを示している。

- 1 . GCRA R A M : 形状 I D のための現在の GCRA 構成及び状態を読出せ。
- 2 . 内部論理内のスケジューリングアルゴリズムの実行せよ。
- 3 . GCRA R A M : 更新された GCRA 構成及び状態の書込み。
- 4 . MINT R A M : スケジュール時カレンダー待ち行列の現在の先頭 / 末尾を読出せ。
- 5 . MINT R A M : スケジュール時カレンダー待ち行列の更新された先頭 / 末尾を書込み。
- 6 . リンク R A M : 古いカレンダー待ち行列末尾からリンクを新しい末尾に書込み。

MINT R A M 及びリンク R A M 動作は、スケジューリングアルゴリズムが “ 最初のイン ” または “ 次のアウト ” の結果を戻した時に限って遂行される。

【 0 0 6 0 】

成熟シーケンス

このシーケンスは、形状 I D トークンのリストを、現時点カレンダー待ち行列から “ 成熟 ” リンクドリストの末尾へ転送し、最初の 3 つの形状 I D を TARDIS 7 2 内へロードする。それは、各主タイミングシーケンス中に 1 回開始される。図 1 8 のテーブルは、

MINT R A M : カレンダー待ち行列から現時点リストを読出せ、

MINT R A M : カレンダー待ち行列内の現時点リストをクリアせよ、

リンク R A M : 現時点リストを “ 成熟 ” リストの末尾へリンクせよ、

リンク R A M : “ 成熟 ” リスト内の次の (第 2 の) 形状 I D トークンを読出せ、

リンク R A M : “ 成熟 ” リスト内の次の (第 3 の) 形状 I D トークンを読出せ、

のシーケンスを示している。

【 0 0 6 1 】

管理シーケンス

このシーケンスは、最小セル間隔を CCRA R A M へ / から書込む、または読出す。これらの動作により、C P U は最小セル間隔を構成し、監視することができる。図 1 9 のテーブルはこのシーケンスを示している。テーブルは、次のシーケンスを示している。

1 . 書込みレジスタ W R _ S I D によって指し示されたアドレス (形状 I D) が読出され、データ (M C I) が読出しレジスタ R R _ M C I _ I N T 及び R R _ M C I _ F R A 内に配置される。読出しレジスタは、読出し要求に関してだけロードされる。

2 . 書込みレジスタ W R _ S I D によって指し示されたアドレス (形状 I D) が、データ (M C I) を使用して書込みレジスタ W R _ M C I _ I N T 及び W R _ M C I _ F R A 内に書込まれる。このステップは、書込み要求に関してだけ発生する。

【 0 0 6 2 】

総合シーケンス例

TARDIS 7 2 が遂行する総合シーケンスの例を図 2 0 に示す。このようなシーケンスは、各主タイミングシーケンス中に走る。各総合シーケンスは、上述したスケジュール、成熟、及び監視シーケンスを組合わせる。図 2 0 の例は、

- 1 . DALEK 7 0 からの 3 つの形状 I D トークンが、全てスケジュール “ 最初のイン ” を有している、
- 2 . DALEK 7 0 からの 3 つの “ 成熟 ” 形状 I D トークンが、全て “ 次のアウト ” スケジュール結果を有している、
- 3 . C P U が GCRA R A M 構成書込みを要求している、

という最悪シナリオを示している。

【 0 0 6 3 】

DALEKブロック

DALEKは、成型器内に現在存在しているセル記述子（C D）の格納を制御する（各接続 I D 毎のリンクされたリストの管理を含む）。図 2 1 は、C D 処理機能ブロック、即ちDALEK 7 0 内へ、及びそれからのC D及び関連形状 I D トークンの流れを示している。システムからC Dを受信すると、形状 I D の表引きが先ず遂行される。C D は“後刻”リスト内に格納され、形状 I D トークンはTARDIS 7 2 へ出力される。形状が確認されると、形状 I D トークンがTARDIS 7 2 からDALEK 7 0 へ入力される。C D は“今”リストへ移動し、C D はシステムへ返送される。

10

【 0 0 6 4 】

DALEK 7 0 は、システム主タイミングシーケンスに同期しているシーケンスを使用して動作する。シーケンス同期は、TARDIS 7 2 によって与えられる。主タイミングシーケンスは、37クロック期間の長さである。これは約685 n s、またはS T S - 1 2 Cをベースとするシステムにおいては1セル時間である。接続 I D 毎の構成可能なC L P オプションフィールドにより、各C D は“C L P クリア”か、“C L P 不変”の何れかとして処理することができる。“C L P クリア”接続 I D 上のC D は、DALEK 7 0 へエントリされるとリセットされるC L P ビットを有している。“C L P 不変”接続 I D 上のC D は、不変のまま渡されるC L P ビットを有している。C L P 及びその関連パリティビットは、DALEK 7 0 によって変更される唯一のC D のフィールドである。

20

【 0 0 6 5 】

DALEK 7 0 によって管理されているデータ構造、及びDALEK 7 0 を通るデータの流れを以下に説明する。任意の時点に、DALEK 7 0 内の各C D は、2つのリンクされたリスト構造の一方内に格納される。各形状 I D 毎に1つの、1組の“後刻”リンクドリストは、C D を受信した時からそれらが伝送のための準備が整うまで、C D を保持する。“今”リンクドリストは、伝送のための準備が整っている全てのC D を保持する。

【 0 0 6 6 】

各主タイミングシーケンス中に、3つまでのC D をシステムから受信することができる。各C D は、ToShapeビット及び接続 I D フィールドを含む。ToShapeビットがセットされている（形状 I D マッピングに対する有効接続 I D が存在している）各C D は、DALEK 7 0 によって外部 R A M アレイ（即ち、データバッファ 7 0 3）内に格納される。一旦格納されたC D はリスト間で転送されても移動せず、代わりにリンクが操作される。リンクは、C D の一部としてデータバッファ 7 0 3 内に格納される。

30

【 0 0 6 7 】

形状 R A M 7 0 1 と呼ばれる外部 R A M アレイは、接続 I D から形状 I D へのマッピングテーブルを保持している。成形は、形状 I D に対して遂行される。複数の接続 I D を単一の形状 I D へマップすることができる。接続 I D のためのC L P オプションフィールドは、その形状 I D と共に形状 R A M 7 0 1 内に格納される。ToShapeビットがセットされているC D は、16K “後刻”リンクドリストの1つに添付される。“後刻”リストは、C D 内のフィールドから4レベルの優先順位を適用する優先順序をベースとしている。このフィールドは、成形された接続内の優先順位（通常は、V C 優先順位）を定義する。“後刻”リストの先頭及び末尾は、COIN R A M 7 0 2 と呼ばれる分離した外部 R A M 内に格納される。

40

【 0 0 6 8 】

受信したC D の格納と同時に、DALEK 7 0 は、GCRA評価のために形状 I D をTARDIS 7 2 へ送る。そのC D はそれがリストの先頭に到達し、形状 I D がTARDIS 7 2 から入力されるまで“後刻”リスト内に留まる。TARDIS 7 2 から入力される形状 I D トークンは、その形状 I D を有するC D をシステムへ出力できることを指示する。C D 選択は、最高優先順位においてその形状 I D のためのリストを占有することである。それは、“後刻”リストの先頭から“今”リストの末尾へ転送される。

50

【 0 0 6 9 】

“ 今 ” リストは、直ちに出力する準備が整っている C D を受入れるための出力待ち行列を提供する。各主タイミングシーケンス中に 1 つの C D しか出力することができないにも拘わらず、3 つまでの形状 I D を TARDIS 7 2 から入力することができるので、このリストが必要である。“ 今 ” リストは、C D 内のフィールドから 4 レベルの優先順位を適用する優先順序をベースとしている。このフィールドは、成形された接続間の優先順位（通常は、V P 優先順位）を定義する。“ 今 ” リストの先頭及び末尾は、1 つの “ 今 ” リストだけが存在しているので、DALEK 7 0 内に格納される。

【 0 0 7 0 】

3 つの全ての外部 R A M アレイ内に保持されているデータは、パリティビットによって保護される。全メモリ読出し動作に続いてパリティが調べられ、エラーがあればフラグが立てられる。同様に、システムから受信した C D のパリティが調べられ、エラーにフラグが立てられる。図 2 2 は、これらのデータ構造及び DALEK 7 0 を通るデータの流れを示している。

【 0 0 7 1 】

DALEK 7 0 によって遂行される動作シーケンスは、主タイミングシーケンスと密に結合されている。これらのシーケンスには、受信、転送、伝送、及び管理と命名されている。

【 0 0 7 2 】

受信シーケンス

このシーケンスは、システムから C D を受入れ、形状 I D をデコードし、そしてその C D を形状 I D “ 後刻 ” リンクドリストに添付する。形状 I D トークンは、このシーケンス中に TARDIS 7 2 へ渡される。図のテーブルはこのシーケンスを示している。

- 1 . 形状 R A M : C D 接続 I D フィールドからデコードされた形状 I D を読出せ。
- 2 . COIN R A M : 形状 I D / 優先順位リストの先頭 / 末尾を読出し、次いで更新されたデータを書込め。
- 3 . データバッファ : C D 及びヌルリンクを書込み、次いでリストの古い末尾へのリンクを書込め。

【 0 0 7 3 】

転送シーケンス

このシーケンスは、C D を “ 後刻 ” リンクドリストから “ 今 ” リンクドリストへ転送する。転送は、TARDIS 7 2 から形状 I D トークンを受信することによって開始される。図 2 4 のテーブルはこのシーケンスを示している。

- 1 . COIN R A M : 全ての 4 優先順位 “ 後刻 ” リストの先頭 / 末尾を読出せ。
- 2 . データバッファ : “ 今 ” 優先順位、及び選択された “ 後刻 ” リストの先頭のリンクを読出せ。
- 3 . COIN R A M : “ 後刻 ” リストの新しい先頭 / 末尾を書込め（データバッファリンクから）。
- 4 . データバッファ : “ 今 ” リストの新しい末尾へのリンクを書込め。

【 0 0 7 4 】

伝送シーケンス

このシーケンスは、“ 今 ” リンクドリストから C D を読出し、システムへその C D を出力する。図 2 5 はこのシーケンスを示している。

- 1 . データバッファ : C D をワード毎に読出せ。
- 2 . C D __ データバス駆動。
- 3 . C D __ S H P __ R D Y 表明。

【 0 0 7 5 】

管理シーケンス

このシーケンスは、形状 I D を形状 R A M へ書込み（もし要求されれば）、形状 R A M から形状 I D を読出す。これらの動作により、DALEK 7 0 内の形状 I D マッピングへの接続 I D の構成及び監視が可能になる。図 2 6 のテーブルはこのシーケンスを示している。

1. 書込みレジスタCPU__WR__CIDによって指し示されているアドレス（接続ID）が書込みレジスタ内のデータ（形状ID）を使用して書込まれる。

2. CPU__WR__CIDによって指し示されているアドレス（接続ID）が読出され、そのデータ（形状ID）が読出しレジスタCPU__RD__CID内に配置される。

【0076】

総合シーケンス例

図27に、DALEK70が遂行する総合シーケンスの例を示す。このようなシーケンスは、各主タイミングシーケンス中に走る。各総合シーケンスは、前項に記載した受信、転送、伝送、及び管理シーケンスを組み合わせる。ここで選択した総合シーケンスの例は、次の最悪シナリオである。

1. システムから3つのCDを受信して3つの受信シーケンスが開始される。

2. TARDIS72から3つの形状IDトークンを受信して3つの転送シーケンスが開始される。

3. “今”リストが占有されて伝送シーケンスが開始される。

4. CPU__SR__WRREQビットが表明されて管理シーケンスが開始される。

【0077】

先に定義したように、接続毎の形状IDと共に“今”及び“後刻”を本発明に使用する場合には、それぞれ仮想接続（VC）及び仮想経路（VP）内に優先順位を与える。これにより、VP内で成形されている接続のための相対優先順位が効果的に保存される。また、カレンダー待ち行列を使用すると、“仮想終了時間”（VFT）計算の複雑さが減少し、得られたVFTはそのアルゴリズムの複雑さ $[O(1)$ 対 $O(N \log N)]$ と境を接する一定の時間を有している。最後に、“活動リスト”を使用することにより、接続毎のスケジューリングの複雑さが減少する。

【0078】

当分野に精通していれば、本発明の範囲及び思想から逸脱することなく本発明に種々の変更をなしえることが理解されよう。特許請求の範囲内で上述した以外に実施できることは明白である。

【図面の簡単な説明】

【図1】 マルチノードネットワークを通して接続されている複数の源/行先（S/D）ユーザの概要ブロック図である。

【図2】 図1のネットワーク内のノードのシーケンスを通して1つのS/Dユーザを別のS/Dユーザに接続する回路の概要図である。

【図3】 情報を源（S）から順方向（F）に行先（D）へ送り、制御信号を逆方向（R）へ源（S）に伝送する仮想チャンネル接続を有する図2の回路の概要図である。

【図4】 図1のネットワーク内のノード（N）の典型的な1つの概要図である。

【図5】 図4のノード内のキューイングユニットの概要図である。

【図6A】 各セルが1ms離間し、トラフィックの“バースト”がランダムに離間しているセルトラフィックを示す図である。

【図6B】 セルが3msの均一間隔に“成形”された後の図6Aのセルトラフィックを示す図である。

【図7】 本発明により構成された成型器の機能ブロックのブロック図である。

【図8】 セル記述子（CD）フォーマットの例である。

【図9】 本発明により構成された成型器の一例のブロック図である。

【図10】 形状ID処理ブロックを通る形状IDのデータの流れを示す図である。

【図11】 本発明により構成されたカレンダー待ち行列を示す図である。

【図12】 形状IDの“成熟”リンクドリストを示す図である。

【図13】 本発明の一実施の形態による最小及び最大セル間隔のテーブルである。

【図14】 最小セル間隔の例のテーブルである。

【図15】 スケジューリング動作のための真理値表である。

【図16】 カレンダー待ち行列挿入時点計算の真理値表である。

【図 17】 形状 I D をスケジュールするためのスケジュールシーケンスを示す図である。

【図 18】 形状 I D 処理ブロックの“成熟”シーケンスの動作を示す図である。

【図 19】 形状 I D 処理ブロックの管理シーケンスの動作を示す図である。

【図 20】 形状 I D 処理ブロックが遂行する総合シーケンスの例を示す図である。

【図 21】 C D 処理ブロックを通る C D 及び形状 I D のデータの流れを示す図である。

【図 22】 C D 処理ブロック内のデータ構造及びデータの流れを示す図である。

【図 23】 C D 処理ブロックにおける受信シーケンスの動作を示す図である。

【図 24】 C D 処理ブロックにおける転送シーケンスの動作を示す図である。

【図 25】 C D 処理ブロックにおける伝送シーケンスの動作を示す図である。

【図 26】 C D 処理ブロックにおける管理シーケンスの動作を示す図である。

【図 27】 C D 処理ブロックによって遂行される総合シーケンスの例を示す図である。

10

【図 1】

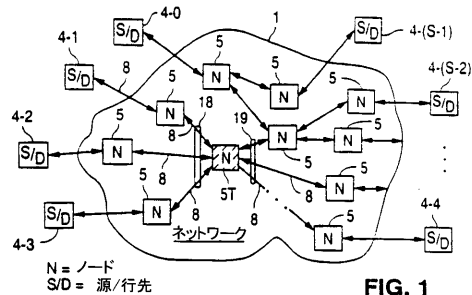
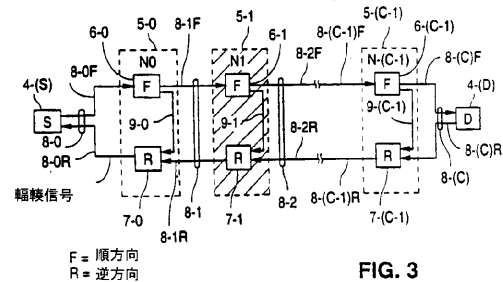


FIG. 1

【図 3】



【図 4】

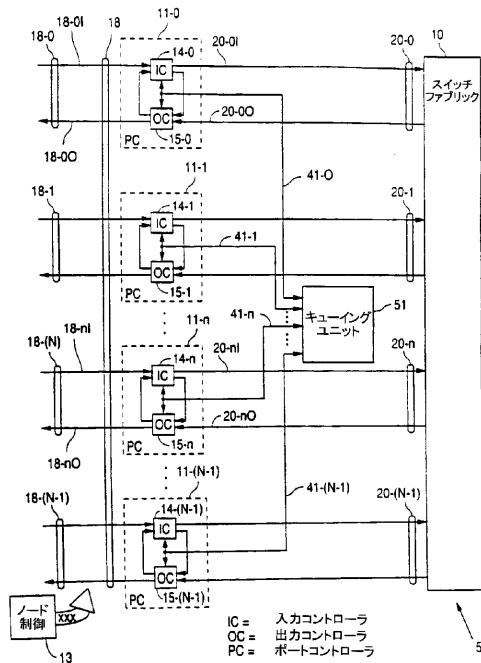


FIG. 4

【図 5】

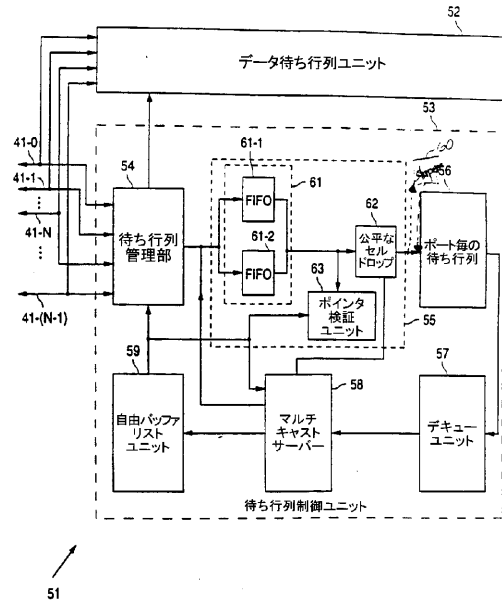


FIG. 5

【図 6 A】

FIG. 6(A)

【図 6 B】

FIG. 6(B)

【図 7】

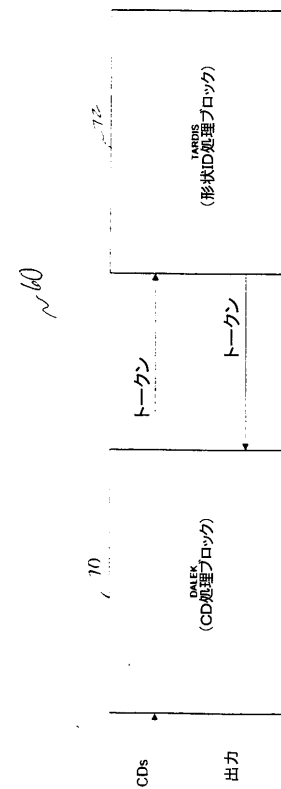


Figure 7

【図 8】

ワード	ラベル	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
0	CD_0	セルポインタ															
1	CD_1	3	ATMC	接続ID													
2	CD_2	M	C	破棄選択	待ち行列	ポート	COS										
3	CD_3	TS	TS	S	S	A	M	00000000	PTI	0							
		VP	VC	A	A	C	C	00000000									
		Priority	Priority	R	R	O	O	00000000									
				E	E	P	P	00000000									
								00000000									

(1) 成形器はCLPビットをリセットし、パリティのビット3を変化させることができる

Figure 8

【図 9】

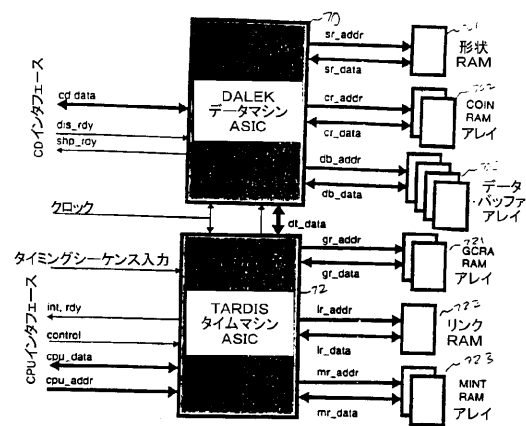


Figure 9

【図 10】

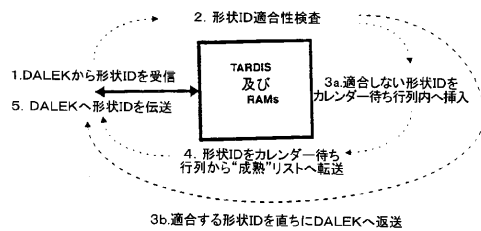


Figure 10

【図 12】

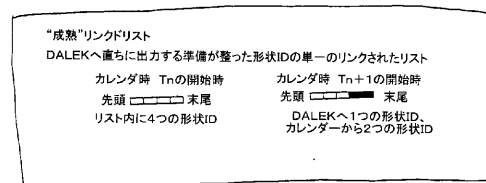


Figure 12

【図 11】

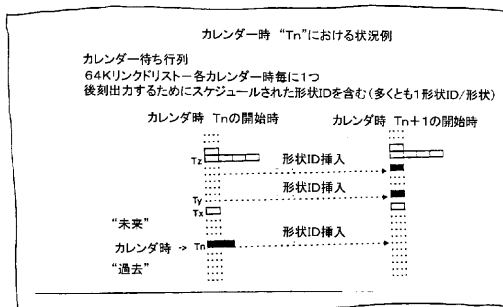


Figure 11

【図 13】

構成された最小セル間隔	整数・少数 (16進:16進)	記述
最小	0002:00	最高帯域幅 約309.4Mbps
最大	FFFF:FF	最低帯域幅 約9.1kbps

Figure 13

【図 14】

最小セル間隔	セルレート (cps)	ステップサイズ (cps)	ビットレート (Mbps)	ステップサイズ (%)
0002:00	730*10 ³	1422	309.41	0.195%
0002:01	728*10 ³		308.80	
0002:9D	558*10 ³	834	236.79	0.1492
0004:00	365*10 ³		154.70	
0004:01	364*10 ³	356	154.55	0.098%
0010:00	91.2*10 ³		38.68	
0010:01	91.2*10 ³	22	38.67	0.024%
0100:00	5.70*10 ³		2.4172	
0100:01	5.70*10 ³	0.08	2.4172	0.001%
1000:00	356	small	151.1kbps	small
FFFF:FF	23.87	very small	9.1kbps	very small

ステップサイズは558478cps(236.79Mbps)以下のレートでITU推奨に合致

Figure 14

【図 15】

ラベル	入力		出力			
	スケジューリング動作 のトリガ	カウン ト	SID から DALEK へ	SIDから カレンダー 待ち行列 へ	カウン ト	スケジュール時
最初のイン	DALEKから	=0	Yes	Yes	+1	CT+MCI
次のイン	SIDを受信した時	!=0	No	No	+1	NC
次のアウト	「成熟」リストの先頭 のSIDがDALEKへ	>1	Yes	Yes	-1	ST+MCI
ゴーストアウト	送られた時	<=1	Yes	No	0	NC

CT=現時点、NC=不変、ST=スケジュール時、MCI=最小セル間隔

Figure 15

【図 18】

DT インタ フェース	GCRA RAM	MINT RAM	リンク RAM	注
	先頭/末尾 の読み出し			現時点カレンダー待ち 行列リストの読み出し
	スル先頭 /末尾の 書き込み	古い末尾 の書き込み		現時点カレンダー待ち行列リストを消去 リストを「成熟」リストの末尾へリンク
		次の形状ID の読み出し		「成熟」リストの第2の形状IDを取得 (そのために必要なそれをDALEKへ受け 渡すことができる)
形状ID				「成熟」リストの先頭をDALEKへ
		次の形状ID の読み出し		「成熟」リストの形状IDを取得 (そのために必要なそれをDALEKへ

Figure 18

【図 19】

DTI インタフェース	GCRA RAM	MINT RAM	リンク RAM	注
	CPUのため にGCRAを 読み出し			MCI及び状態情報を読み出し
				スケジュールシーケンスに関する読み出しを ここですることができ、この場合Management はスルにされる
	CPUのため にGCRA を書込み			新しいMCI及び状態情報を書込み (Requestのみを書き込み)

Figure 19

【図 16】

ラベル	入力	出力
時間循環なし、 窓内にST	(CT < CT+MCI) 及び (ST < CT+MCI)	スケジュール時
時間循環、 時間循環後窓内にST	(CT > CT+MCI) 及び (ST < CT+MCI)	スケジュール時
時間循環、 時間循環前窓内にST	(CT > CT+MCI) 及び (ST > CT+MCI)	スケジュール時
窓外にST	その他	現時点

CT=現時点、MCI=最小セル間隔、ST=スケジュール時
“窓”は(現時点)から(現時点+最小セル間隔)までの時間
時間循環とは(カレンダー時=D)が窓内にあることをいう

Figure 16

【図 17】

DT インタ フェース	GCRA RAM	MINT RAM	リンク RAM	注
形状ID				DALEKとの形状IDの送受
				GCRA構成及び状態の取得
				結果を計算する数サイクル (内部ロジックで)
				更新されたGCRA状態の書き込み、 スケジュール待ち行列待ち行列リストの取得
				スケジュール時カレンダー待ち行列リストの更新 古い末尾を新しい末尾へリンク

Figure 17

【図 20】

T 状態	DALEK インタフェース	GCRA RAM	MINT RAM	リンク RAM	注
6	cid_error				
7		CPUのため にGCRAを 読み出し			GR.WREQ処理の開始
8					
9	later_count		先頭/末尾 の読み出し		DALEKからのSID Aのスケジューリング の開始
10	now_count				成熟シーケンスの開始
11		CPUのため にGCRA を書込み			GR.WREQ処理の終了
12					
13					
14	dalek_sig1		先頭/末尾 を書込み	古い末尾 の書き込み	DALEKからのSID Bの スケジューリングの開始
15	dalek_sig0				
16	dlr_test				リストの第2 の形状IDを 読み出し
17	dlr_test				
18	high-Z				
19					
20	SID X				SID Aのスケジューリングの完了 DALEKからのSID Cの スケジューリングの開始
21					
22					リストの第3 の形状IDを 読み出し

T 状態	DALEK インタフェース	GCR RAM	MINT RAM	リンク RAM	注
23					
24	SID Y	Write data	Write data	Write data	SID Bのスケジューリングの完了 「成熟」からのSID 1の スケジューリングの開始
25		Write data	Write data	Write data	
26		Write data	Write data	Write data	
27		Write data	Write data	Write data	
28					
29	SID Z	Write data	Write data	Write data	SID Cのスケジューリングの完了 「成熟」からのSID 1の スケジューリングの開始
30		Write data	Write data	Write data	
31		Write data	Write data	Write data	
32		Write data	Write data	Write data	
33					
34	config	Write data	Write data	Write data	SID 10のスケジューリングの完了 「成熟」からのSID Kの スケジューリングの開始
35	cpu_wrcid	Write data	Write data	Write data	
36	cpu_wrcid	Write data	Write data	Write data	
37	dis_test	Write data	Write data	Write data	
1					
2	high-Z		Write data	Write data	SID Jのスケジューリングの完了
3	dis_test		Write data	Write data	
4	status	Write data	Write data	Write data	
5	cpu_wrcid	Write data	Write data	Write data	
6	cid_error				
7	high-Z		Write data	Write data	SID Kのスケジューリングの完了
8					

【図 22】

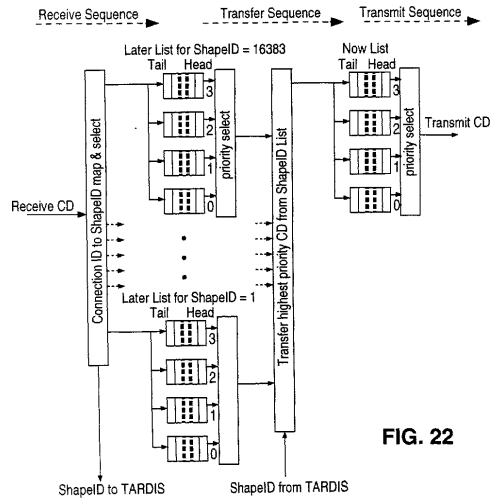


FIG. 22

【図 21】

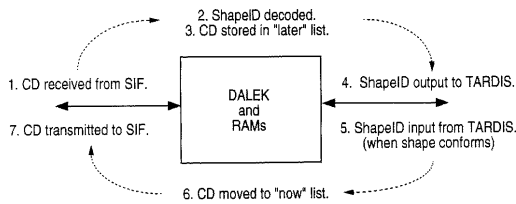


FIG. 21

【図 23】

CD インタフェース	形状 RAM	データ バッファ	COIN RAM	TARDIS インタフェース	注
CDが有効					
					DIS_RDYがアサート
					形状ID CDをデータバッファに書き込み
					先頭/末尾を更新
					リンクを更新

Figure 23

【図 25】

CD インタフェース	形状 RAM	データ バッファ	COIN RAM	TARDIS インタフェース	注
					バッファ(BUFFER)からCDを取得
					「今」リスト の先頭に あるCDの 読み出し (4読み 出し)
					CDを出力可能
					…次のメインタイミングシーケンスのCD出力時まで待機…
CDの出力					SHP_RDYがアサート

Figure 25

【図 24】

CD インタフェース	形状 RAM	データ バッファ	COIN RAM	TARDIS インタフェース	注
					形状ID
					優先順位3の リストの先頭の 読み出し
					優先順位2
					優先順位1
					優先順位0
					古い「後刻」 先頭のリンクの 読み出し
					新しい「後刻」先頭の取得
					古い「今」末 尾のリンクの 書き込み
					優先順位の 新しい先頭の 書き込み
					新しいリンク及び先頭の保存

Figure 24

【図 26】

CD インタフェース	形状 RAM	データ バッファ	COIN RAM	TARDIS インタフェース	注
					cpu_wrrreq=1なら書き込み
					CPUから SIDを 書き込み
					CPUのた めにSIDを 読み出し
					読み出しはいつでもよい

Figure 26

【図 27】

T 状態	CD インタフェース	形状 RAM	データ バッファ	COIN RAM	TARDIS インタフェース	注
37					du_test	cpu_sr_wrrreqがアサート
1					high-Z	DT_SYNCがアサート
2			CPUのために SIDを読み出し		du_test	
3					status	
4					cpu_rd_sid	DIS_RDYがアサート
5					cid_error	
6					cid_error	
7					cid_error	
8					cid_error	
9					cid_error	
10					cid_error	DIS_RDYがアサート
11					cid_error	
12					cid_error	
13					cid_error	
14					cid_error	
15					cid_error	DIS_RDYがアサート

T 状態	CD インタフェース	形状 RAM	データ バッファ	COIN RAM	TARDIS インタフェース	注
16			古い今の 末尾のリンク の書き込み		du_test	
17			古い今の 末尾のリンク の書き込み		du_test	
18			古い今の 末尾のリンク の書き込み		high-Z	
19			古い今の 末尾のリンク の書き込み		high-Z	
20			古い今の 末尾のリンク の書き込み		high-Z	
21			古い今の 末尾のリンク の書き込み		high-Z	
22			古い今の 末尾のリンク の書き込み		high-Z	
23			古い今の 末尾のリンク の書き込み		high-Z	
24			古い今の 末尾のリンク の書き込み		high-Z	
25			古い今の 末尾のリンク の書き込み		high-Z	
26			古い今の 末尾のリンク の書き込み		high-Z	
27			古い今の 末尾のリンク の書き込み		high-Z	
28			古い今の 末尾のリンク の書き込み		high-Z	
29			古い今の 末尾のリンク の書き込み		high-Z	
30			古い今の 末尾のリンク の書き込み		high-Z	
31			古い今の 末尾のリンク の書き込み		high-Z	
32			古い今の 末尾のリンク の書き込み		high-Z	
33			古い今の 末尾のリンク の書き込み		high-Z	
34			古い今の 末尾のリンク の書き込み		high-Z	
35			古い今の 末尾のリンク の書き込み		high-Z	
36			古い今の 末尾のリンク の書き込み		high-Z	
37			古い今の 末尾のリンク の書き込み		high-Z	
1			古い今の 末尾のリンク の書き込み		high-Z	
2			古い今の 末尾のリンク の書き込み		high-Z	
3			古い今の 末尾のリンク の書き込み		high-Z	
4			古い今の 末尾のリンク の書き込み		high-Z	

フロントページの続き

- (74)代理人 100096194
弁理士 竹内 英人
- (74)代理人 100074228
弁理士 今城 俊夫
- (74)代理人 100084009
弁理士 小川 信夫
- (74)代理人 100082821
弁理士 村社 厚夫
- (74)代理人 100086771
弁理士 西島 孝喜
- (74)代理人 100084663
弁理士 箱田 篤
- (72)発明者 レイノルズ スタンレイ エム
オーストラリア ウェスト オーストラリア 6 0 0 8 スピアコ パーカー ロード 1 6 4
- (72)発明者 ロウペンスキ トッド エル
アメリカ合衆国 カリフォルニア州 9 4 0 1 0 ヒルズボロウ サミット ドライヴ 2 4 3 0
- (72)発明者 レメイア リチャード
カナダ ケベック ジェイ4ゼット 1シー1 ブロッザード リュー アルフォンズ 5 7 3 0

審査官 矢頭 尚之

- (56)参考文献 特開平11-068771(JP,A)
特開平11-055276(JP,A)
特開平10-135957(JP,A)

- (58)調査した分野(Int.Cl., DB名)
H04L 12/56