(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第4426589号
(P4426589)

(45)発行日　平成22年3月3日(2010.3.3)　(24)登録日　平成21年12月18日(2009.12.18)

(51)Int.Cl.			F I		
H04N	7/173	(2006.01)	H04N	7/173	610A
G06F	3/06	(2006.01)	G06F	3/06	540
H04L	12/40	(2006.01)	H04L	12/40	Z

請求項の数 24　(全 16 頁)

(21)出願番号	特願2006-542714 (P2006-542714)
(86)(22)出願日	平成16年12月2日 (2004.12.2)
(65)公表番号	特表2007-513582 (P2007-513582A)
(43)公表日	平成19年5月24日 (2007.5.24)
(86)国際出願番号	PCT/US2004/040235
(87)国際公開番号	WO2005/057828
(87)国際公開日	平成17年6月23日 (2005.6.23)
審査請求日	平成18年6月1日 (2006.6.1)
(31)優先権主張番号	60/526,437
(32)優先日	平成15年12月2日 (2003.12.2)
(33)優先権主張国	米国(US)
(31)優先権主張番号	10/999,661
(32)優先日	平成16年11月30日 (2004.11.30)
(33)優先権主張国	米国(US)

(73)特許権者　504199079
インタラクティヴ　コンテント　エンジンズ，エルエルシー
アメリカ合衆国　ハワイ州　９６８１３　ホノルル　ビショップ・ストリート　１０８８　４１００号
(74)代理人　100070150
弁理士　伊東　忠彦
(74)代理人　100091214
弁理士　大貫　進介
(74)代理人　100107766
弁理士　伊東　忠重
(72)発明者　ローズ，スティーヴン，ダブリュ
アメリカ合衆国　ハワイ州　９６７６８　ハリイメイル　マイカイ　８６６

最終頁に続く

(54)【発明の名称】同期データ転送システム

(57)【特許請求の範囲】

【請求項１】

同期データ転送システムであって：

複数のプロセッサ・ノードと、

前記複数のプロセッサ・ノードに結合された、該複数のプロセッサ・ノード間の通信を可能にするためのバックボーンネットワークスイッチと、

前記複数のプロセッサ・ノードに結合され前記複数のプロセッサ・ノードにまたがって分散されており、複数のタイトルを保存する複数の記憶装置であって、各タイトルが複数のサブチャンクに分割され、該サブチャンクが前記複数の記憶装置にまたがって分散されるような複数の記憶装置と、

前記複数のプロセッサ・ノードのうちの対応するものの上でそれぞれ実行され、ソース・プロセッサ・ノードから宛先のプロセッサ・ノードに転送されるべき各サブチャンクについてのソース・ノード識別子および宛先ノード識別子を含むメッセージを送るよう動作する複数の転送プロセスと、

複数の逐次的な送信周期のそれぞれを開始させる送信コマンドを周期的にブロードキャストし、複数のメッセージを受け取り、各プロセッサ・ノードが各送信周期の間に高々一つのサブチャンクを送信して高々一つのサブチャンクを受信することを保証するよう各送信周期に先立って前記複数のメッセージの間から選択を行い、選択されたメッセージに対応する複数の送信要求を送る、前記複数のプロセッサ・ノードのうちの少なくとも一つの上で実行される同期スイッチマネージャ・プロセス、

とを有しており、少なくとも一つのメッセージを送り、前記同期スイッチマネージャ・プロセスから対応するサブチャンクを同定する送信要求を受信した各転送プロセスが、前記対応するサブチャンクをブロードキャストされた送信コマンドによって開始される次の送信周期の間に送ることを特徴とするシステム。

【請求項２】

請求項１記載の同期データ転送システムであって、前記複数のメッセージのそれぞれがタイムスタンプを有し、前記同期スイッチマネージャ・プロセスが前記複数のメッセージに対してタイムスタンプ順に基づいて優先度付けをし、前記複数の送信要求をタイムスタンプ順に送ることを特徴とするシステム。

【請求項３】

請求項２記載の同期データ転送システムであって、さらに：

前記複数のプロセッサ・ノードのうちの対応するものの上でそれぞれ実行され、複数のタイムスタンプ付き読み取り要求を送るよう動作する複数のユーザープロセスを有しており、

各転送プロセスが対応するタイムスタンプ付き読み取り要求からタイムスタンプを対応するメッセージに組み込むことを特徴とするシステム。

【請求項４】

請求項３記載の同期データ転送システムであって、前記同期スイッチマネージャ・プロセスが、前記複数のメッセージを準備完了メッセージリストにタイムスタンプ順に整理し、前記複数の逐次的送信周期の各周期の直前に前記準備完了メッセージリストをタイムスタンプ順にスキャンし、タイムスタンプの優先度に基づいてメッセージを選択することを特徴とするシステム。

【請求項５】

請求項４記載の同期データ転送システムであって、前記同期スイッチマネージャ・プロセスがあるメッセージを、同定されたソース・プロセッサ・ノードがすでに次の送信周期の間にサブチャンクを送信するために選択されているのでない場合であり、かつ同定された宛先プロセッサ・ノードがすでに前記次の送信周期の間にサブチャンクを受信するために選択されているのでない場合に選択することを特徴とするシステム。

【請求項６】

請求項１記載の同期データ転送システムであって、さらに：

前記複数の転送プロセスのそれぞれが、受け取ったサブチャンク読み取り要求を読み取り要求待ち行列に保存し、各サブチャンク読み取り要求がローカルに保存されたサブチャンクを示していることと、

前記複数の記憶装置のそれぞれが、ローカルな読み取り要求待ち行列で同定されているサブチャンクを物理的な順序で読み取ることと、

前記複数のプロセッサ・ノードのそれぞれが、対応する記憶装置による読み取りが成功したサブチャンクを成功読み取り待ち行列にリストすることと、

前記複数の転送プロセスのそれぞれが、対応する成功読み取り待ち行列中の各項目についてのメッセージを前記同期スイッチマネージャ・プロセスに送ること、

とを有することを特徴とするシステム。

【請求項７】

請求項６記載の同期データ転送システムであって、前記サブチャンク読み取り要求のそれぞれがタイムスタンプ付き読み取り要求を含んでおり、前記成功読み取り待ち行列のそれぞれにおける項目がタイムスタンプ順にリストされており、各転送プロセスが対応する成功読み取り待ち行列における各項目についてのメッセージをタイムスタンプ順に送ることを特徴とするシステム。

【請求項８】

請求項６記載の同期データ転送システムであって、さらに：

前記複数の転送プロセスのそれぞれが、ある成功要求待ち行列から、対応する送信要求によって同定されたあるサブチャンクと関連付けられている項目を除去することと、

前記複数のプロセッサ・ノードのうちの対応するものの上でそれぞれ実行されており、それぞれ同定されたサブチャンクを送信コマンドに反応して宛先プロセッサ・ノードに転送するために使われるネットワークパケットを構築するよう動作する複数のネットワーク転送プロセス、
とを有することを特徴とするシステム。
【請求項９】
請求項１記載の同期データ転送システムであって、前記ネットワークスイッチが複数のポートをもつギガビットイーサネット（登録商標）スイッチであり、前記複数のプロセッサ・ノードのそれぞれが前記ネットワークスイッチの対応するポートに結合されていることを特徴とするシステム。
【請求項１０】
請求項１記載の同期データ転送システムであって、前記複数のプロセッサ・ノードのうちに、前記同期スイッチマネージャ・プロセスを実行する管理ノードが含まれることを特徴とするシステム。
【請求項１１】
請求項１記載の同期データ転送システムであって、前記複数のプロセッサ・ノードのうちに、前記同期スイッチマネージャ・プロセスを実行する第一の管理ノードと、ミラーされた同期スイッチマネージャ・プロセスを実行する第二の管理ノードとが含まれることを特徴とするシステム。
【請求項１２】
ネットワークスイッチ結合された複数のプロセッサ・ノードの間でデータの分散されたサブチャンクを同期的に転送する方法であって：
前記プロセッサ・ノードのうちの少なくとも一つの上で実行される管理プロセスによって、複数の逐次的な送信周期のそれぞれを開始させる送信コマンドを周期的にブロードキャストし、
前記管理プロセスに対して、送るべき少なくとも一つのサブチャンクをもつ各プロセッサによって、送られるべき各サブチャンクについてのソース・プロセッサ・ノードおよび宛先プロセッサ・ノードを同定するメッセージを送り、
前記管理プロセスによって、各プロセッサ・ノードが次の送信周期の間に高々一つのサブチャンクを送信し、各プロセッサ・ノードが前記次の送信周期の間に高々一つのサブチャンクを受信することを保証するよう、前記プロセッサ・ノードから受け取られたメッセージを選択し、
前記管理プロセスによって、選択された対応するメッセージを送ったプロセッサ・ノードにそれぞれ送られる複数の送信要求を送り、
送信要求を受け取る各プロセッサ・ノードによって、前記受け取られた送信要求によって同定されたサブチャンクを次の送信コマンドに反応して宛先プロセッサ・ノードに送信する、
ことを含むことを特徴とする方法。
【請求項１３】
請求項１２記載の方法であって、さらに：
前記した送られるべき各サブチャンクについてのメッセージを送ることに先立って、各メッセージにタイムスタンプを付けることを含んでおり、
前記した選択がタイムスタンプ順に基づいて優先度を付けることであり、
前記した複数の送信要求を送ることがタイムスタンプ順に送信要求を送ることであることを特徴とする方法。
【請求項１４】
請求項１３記載の方法であって、さらに：
少なくとも一つのプロセッサ・ノードによって、複数のタイムスタンプ付き読み取り要求を送ることを含んでおり、
前記した各メッセージにタイムスタンプを付けることが、受け取られたタイムスタンプ付

き読み取り要求からのタイムスタンプを対応するメッセージに組み込むことを含んでいることを特徴とする方法。

【請求項１５】

請求項１４記載の方法であって、さらに：

前記管理プロセスによって、受け取られたメッセージを準備完了メッセージリスト中でタイムスタンプ順に整理し、

前記管理プロセスによって、各送信周期の直前に、前記準備完了メッセージリストをタイムスタンプ順にスキャンする、

ことを含むことを特徴とする方法。

【請求項１６】

請求項１５記載の方法であって、前記したスキャンすることが、あるメッセージを、同定されたソース・プロセッサ・ノードがすでに次の送信周期の間にサブチャンクを送信するために選択されているのでない場合であり、かつ同定された宛先プロセッサ・ノードがすでに前記次の送信周期の間にサブチャンクを受信するために選択されているのでない場合に選択することを含むことを特徴とする方法。

【請求項１７】

請求項１６記載の方法であって、前記したスキャンすることが、準備完了メッセージリスト全体がスキャンされたとき、あるいは全部のプロセッサ・ノードがサブチャンクを送信するために選択されてしまっている場合、あるいは全部のプロセッサ・ノードがサブチャンクを受信するために選択されてしまっている場合に、完了することを特徴とする方法。

【請求項１８】

請求項１２記載の方法であって、さらに：

受け取られたサブチャンク読み取り要求を読み取り要求待ち行列に保存し、各サブチャンク読み取り要求がローカルに保存されたサブチャンクを示しており、

ローカルなディスクドライブによって、読み取り要求待ち行列で同定されているサブチャンクを物理的な順序で読み取り、

読み取りが成功したサブチャンクの項目を成功読み取り待ち行列にリストする、

ことを有しており、前記した送られるべき各サブチャンクについてのメッセージを送ることが、前記成功読み取り待ち行列中に各項目についてのメッセージを送ることであることを特徴とする方法。

【請求項１９】

請求項１８記載の方法であって、各サブチャンク読み取り要求がタイムスタンプ付き読み取り要求であり、前記した読み取りに成功したサブチャンクの項目を成功読み取り待ち行列にリストすることが項目をタイムスタンプ順にリストすることを含んでおり、前記した前記構成読み取り待ち行列中の各項目についてのメッセージを送ることがタイムスタンプ順にメッセージを送ることを含んでいることを特徴とする方法。

【請求項２０】

請求項１８記載の方法であって、さらに：

前記成功要求待ち行列から、対応する送信要求によって同定されたあるサブチャンクと関連付けられている項目を除去し、

同定されたサブチャンクを送信コマンドに反応して宛先プロセッサ・ノードに転送するために使われるネットワークパケットを構築する、

ことを含むことを特徴とする方法。

【請求項２１】

請求項１２記載の方法であって、さらに、第一の管理ノード上で前記管理プロセスを実行し、前記第一の管理ノードをミラーするミラーされた管理ノード上でミラーされた管理プロセスを実行することを含むことを特徴とする方法。

【請求項２２】

同期データ転送システムであって：

第一および第二のユーザーノードおよび管理ノードを含む複数のストレージ・プロセッサ・ノードと、
前記複数のストレージ・プロセッサ・ノードに結合されたバックボーン通信スイッチと、
前記複数のストレージ・プロセッサ・ノードにまたがって分散される複数のサブチャンクにそれぞれ細分された複数のタイトルと、
それぞれが対応するサブチャンクを要求するための複数のタイムスタンプ付き読み取り要求を送る、前記第一のユーザーノード上で実行されるユーザープロセスと、
ローカルに保存されているサブチャンクを要求する受け取ったタイムスタンプ付き読み取り要求それぞれについてのソース・ノード識別子および宛先ノード識別子を含むメッセージを送る、前記第二のユーザーノード上で実行される転送プロセスと、
前記スイッチを通じて複数の逐次的な送信周期のそれぞれを開始させる送信コマンドを周期的にブロードキャストし、複数のメッセージを受け取り、各プロセッサ・ノードが各送信周期の間に高々一つのサブチャンクを送信して高々一つのサブチャンクを受信することを保証するよう各送信周期に先立って前記複数のメッセージの間から選択を行い、選択されたメッセージに対応する複数の送信要求を送る、前記管理ノードの上で実行される管理プロセス、
とを有しており、前記転送プロセスが、前記管理プロセスから送信要求を受け取るのに反応して、対応するサブチャンクを次のブロードキャストされる送信コマンドによって開始される次の送信周期の間に送ることを特徴とするシステム。

【請求項２３】
請求項２２記載の同期データ転送システムであって、前記管理プロセスが前記複数のメッセージからタイムスタンプ優先度に基づいて選択を行うことを特徴とするシステム。

【請求項２４】
請求項２３記載の同期データ転送システムであって、前記管理プロセスがまず最高優先度のタイムスタンプをもつメッセージを選択し、次いでその後の各メッセージを、同定されたソース・ノードがすでに次の送信周期の間にサブチャンクを送信するために選択されているのでない場合であり、かつ同定された宛先ノードがすでに前記次の送信周期の間にサブチャンクを受信するために選択されているのでない場合に選択することを特徴とするシステム。

【発明の詳細な説明】
【技術分野】
【０００１】
関連出願への相互参照
本出願は、2003年12月2日に出願された米国仮出願第60/526,437号の恩恵を主張するものであり、2002年11月26日に出願された“Interactive Broadband Server System”と題する係属中の米国特許出願、通し番号第10/304,378号の部分継続出願であり、この後者自身も2001年11月28日に出願された米国仮出願第60/333,856号の恩恵を主張している。これらすべては共通の発明者を有し、本出願の出願人に譲渡されており、ここにあらゆる面において参照によって組み込まれる。
本発明の技術分野
本発明は、対話式ブロードバンドサーバーシステムに、より詳細には、多重同時アイソクロナス・データストリーム（multiple simultaneous isochronous data stream）の高速配送を容易にするための同期（synchronized）データ転送システムを利用する対話式コンテンツエンジンに関するものである。
【背景技術】
【０００２】
ストリーミングメディアコンテンツの記憶および配送のためのソリューションを提供することが望まれている。スケーラビリティーについての初期の目標は、ストリームあたり4メガビット毎秒（Mbps）での100ないし100万の同時の個別的なアイソクロナス・コンテ

ンツストリームである。ただし、異なるデータレートも考えられている。利用可能な全通信帯域幅は利用可能な最大のバックプレーンスイッチによって制約される。現在最大のスイッチはテラビット毎秒の範囲、すなわち20万同時出力ストリーム程度である。出力ストリームの数は一般にストリームあたりのビットレートに反比例する。

【発明の開示】

【発明が解決しようとする課題】

【０００３】

コンテンツ記憶装置の最も単純なモデルは、単一のネットワークコネクタを有する単一のプロセッサに接続された単一のディスクドライブである。データはディスクから読み出され、メモリに入れられ、パケットとしてネットワークを通じて各ユーザーに配布される。ウェブページなどのような伝統的なデータは非同期的に配送できる。つまり、ランダムな時間的遅延をもったランダムな量のデータがある。低ボリューム、低解像度のビデオはウェブサーバーから配送できる。ビデオおよびオーディオのようなリアルタイムのメディアコンテンツはアイソクロナス伝送、すなわち配送時間を保証された伝送を必要とする。このシナリオでは、通信帯域幅の制約はディスクドライブに存在する。ディスクはアームの動きや回転による遅延と競合しなければならない。システムが任意の所与の時点においてドライブからプロセッサへの連続的なコンテンツの同時ストリームを6つしかサポートできないとすると、7番目のユーザーの要求は先の6人のユーザーの一人がコンテンツストリームを放棄するのを待たねばならない。この設計の長所は単純さである。短所は、当該設計中で唯一の機械的装置であるディスクがその速さでしかデータのアクセス・転送ができないということである。

【０００４】

一つまたは複数の別のドライブを追加したり、ドライブアクセスをインターリーブしたりすることによって改良することもできる。また、重複コンテンツを各ドライブに保存して冗長性およびパフォーマンスを稼ぐこともできる。これは前進ではあるが、それでもいくつかの問題がある。ローカルな単数または複数のドライブ上には一定量のコンテンツしか置くことができない。ディスクドライブ、CPUおよびメモリはそれぞれ破綻につながりかねない単一障害点である。このシステムは、ディスクコントローラが扱うことのできるドライブ数までしか拡大できない。ユニット数がたくさんあったとしても、タイトルの配布に問題がある。実世界では、誰もが最新の映画を見たいと思う。大まかにいって、コンテンツ要求の80%はほんの20%のタイトルについてのものである。機械の通信帯域幅のすべてを一つのタイトルにつぎ込むわけにはいかない。それではその機械にしか保存されていない人気の劣るタイトルへのアクセスが遮断されてしまう。結果として、「高需要」タイトルはほとんど、あるいはすべての機械に登録されなければならなくなる。端的にいえば、ユーザーが古い映画を見たいとしたら、そのユーザーは不運なことになりうる――たとえシステムに登録されていたとしても。大きなライブラリでは、先の比はこの例で用いた80/20の割合よりずっと大きくなりうる。

【０００５】

システムがデータ処理で使われる標準的な構内ネットワーク（LAN: Local Area Network）に基づいているとしたら、他の不十分な点もあるだろう。現代のイーサネット（登録商標）ベースのTCP/IPシステムは配送が保証される驚くべきものであるが、パケットの衝突や部分的に紛失したパケットの再送信によって引き起こされる時間的な代価ならびにいっさいを機能させるために必要とされる管理を伴っている。コンテンツストリームのセットが適時に利用可能である保証はない。また、各ユーザーが一つのスイッチポートを消費し、各コンテンツサーバーが一つのスイッチポートを消費するので、スイッチポート数はサーバー数の2倍でなければならず、全体的なオンライン通信帯域幅を制限することになる。

【０００６】

本発明の恩恵、特徴および利点は以下の記述および付属の図面に関してよりよく理解されるであろう。

【発明を実施するための最良の形態】
【０００７】
以下の記述は、特定の用途と要件のコンテキストにおいて与えられる本発明を通常の当業者が作成し、利用できるようにするために提示される。ただし、好ましい実施形態に対するさまざまな修正が当業者には明白であろうし、ここに定義される一般的な原理は他の実施形態にも適用されうる。したがって、本発明はここに示され、説明される特定の実施形態に限定されることを意図したのではなく、ここに開示される原理および新規の特徴と一致する最も広範な範囲を与えられることを意図したものである。
【０００８】
ここに記載されるアーキテクチャはさまざまな性能の個々の構成要素を受け入れられるので、導入が最初のシステム購入時点に限定されるのを避けることができる。利便性（commodity）構成要素の使用によって、近年の十分に実証済みの技術、単一ソースの回避およびストリームあたりのコスト最小化を保証する。個別の構成要素の障害は耐えられる。多くの場合には、ユーザーの観点から目につく動作の変化はない。その他の場合には、短い「自己修復」サイクルがある。多くの場合には多重障害にも耐えうる。また、すべてではないまでもほとんどの場合には、システムはすぐに対処してやらなくても復旧できるので、「無人」（lights-out）運転にとっては理想的となる。
【０００９】
コンテンツ記憶割り当ておよび内部通信帯域幅は、最長未使用時間（LRU: Least Recently Used）アルゴリズムによって自動的にやりくりされる。このアルゴリズムは、RAMキャッシュおよびハードディスクアレイキャッシュにある内容が現在の需要に対して適切であり、バックプレーンスイッチ通信帯域幅が最も効率的な仕方で使用されることを保証する。システム内の通信帯域幅が申し込み過多になることは、よしあるとしてもまれなので、パケットの伝送を破棄または遅延することは必要でない。本アーキテクチャは、各構成要素の複合的な通信帯域幅を最大限に利用する機能を提供し、よって保証が満たされることができ、当該ネットワークはプライベートでありかつ完全なコントロール下にあり、予期しないピーク需要の状況であってもどのデータ経路も過負荷にならない。どんなビットレートのストリームでも受け入れることができるが、典型的なストリームは1ないし20Mbpsの範囲に留まると期待される。非同期コンテンツは利用可能な通信帯域幅に基づいて受け入れられる。アプリケーションが要求すればその目的のために通信帯域幅をリザーブしてもよい。ファイルはいかなるサイズでもよく、記憶の不効率が最小限となる。
【００１０】
図１は、本発明の例示的な実施形態に基づいて実装された対話的コンテンツエンジン（ICE: Interactive Content Engine）１００の一部分の単純化されたブロック図である。本発明の十分にして完全な理解のために適用できない部分は明快のため示されていない。ICE１００は、適切な多重ポート（またはマルチポート）ギガビットイーサネット（登録商標）（GbE）スイッチ１０１を、いくつかのストレージ・プロセッサ・ノード（SPN: Storage Processor Node）１０３に結合された複数のイーサネット（登録商標）ポートを有するバックプレーンファブリックとして含んでいる。各SPN１０３は簡略化したサーバーであり、２つのギガビットイーサネット（登録商標）ポート、一つまたは複数のプロセッサ１０７、メモリ１０９（たとえばランダムアクセスメモリ（RAM））および適切な数（たとえば４ないし８）のディスクドライブ１１１を含んでいる。各SPN１０３上の第一のGbポート１０５は全二重動作（各SPN／ポート接続における同時の送信および受信）のためにスイッチ１０１の対応するポートに接続されており、ICE１００内でデータを移動させるのに使う。もう一方のGbポート（図示せず）はコンテンツ出力を下流のユーザー（図示せず）に配送する。
【００１１】
各SPN１０３はそのローカルディスクドライブへの、および５つのSPNからなる各グループ内の他の４つのSPNの他のディスクドライブへの高速アクセスを有する。スイッチ１０１は単なるSPN１０３間の通信デバイスではなく、ICE１００のためのバックプレーンであ

る。図解のためにSPN１０３は５つしか示されていないが、ICE１００は典型的にはより多くのサーバーを含んでいることは理解されている。各SPN１０３はコンテンツの記憶装置、処理器および送信機の役をする。図示した構成では、各SPN１０３は、市販の構成要素を使って構成され、通常の意味におけるコンピュータではない。標準的なオペレーティングシステムが考えられているが、そのような割り込み駆動型のオペレーティングシステムは無用なボトルネックを生じることもある。

【００１２】

各タイトル（たとえばビデオ、映画または他のメディアコンテンツ）はどの単一のディスクドライブ１１１にも全体としては保存されていない。そうではなく、インターリーブ・アクセスによるスピード上の恩恵を実現するため、各タイトルのデータは分割されてICE１００内のいくつかのディスクドライブにまたがって保存されている。単一のタイトルの内容は複数のSPN１０３の複数のディスクドライブにまたがっている。タイトルコンテンツの短い「時間フレーム」がラウンドロビン式に各SPN１０３内の各ドライブから収集される。この仕方では、物理的な負荷は拡散されてSCSIおよびIDEのドライブ数限界を免れ、ある形のフェイルセーフ動作が得られ、タイトルの大きな集合が組織・管理される。

【００１３】

図示した特定の構成では、各コンテンツタイトルは固定サイズのばらばらなチャンク（かたまり）（典型的にはチャンク１つあたり２メガバイト（MB）程度）に分割される。各チャンクはラウンドロビン式に異なるSPN１０３のセットに保存される。各チャンクは４つのサブチャンクに分割され、パリティを表す５番目のサブチャンクが生成される。各サブチャンクは異なるSPN１０３のディスクドライブ上に保存される。図示して説明される構成では、約512キロバイト（KB）（ここで「K」は1024である）のサブチャンクのサイズはディスクドライブ１１１のそれぞれのデータの名目的な単位に一致する。SPN１０３は５つずつグループ化され、各グループすなわちSPNセットがタイトルのデータの一つのチャンクを保存する。図示したように、前記５つのSPN１０３は１～４、そして「パリティ」とラベル付けされ、これらが集団的にチャンク１１３を、それぞれＳＰＮ１、２、３、４および「パリティ」に保存される５つの別個のサブチャンク１１３ａ、１１３ｂ、１１３ｃ，１１３ｄ、１１３ｅとして保存する。サブチャンク１１３ａ～１１３ｅは各異なるSPNのための異なるドライブ上に（たとえばSPN1/ドライブ1、SPN2/ドライブ2、SPN3/ドライブ3など）分散式に保存されて示されているが、他のいかなる可能な組み合わせで保存されてもよい（たとえばSPN1/ドライブ1、SPN2/ドライブ1、SPN3/ドライブ3など）。サブチャンク１～４はデータを含み、サブチャンク「パリティ」はデータサブチャンクのためのパリティ情報を含む。各SPNセットのサイズは典型的には５であるが任意であり、たとえば２つのSPNから10のSPNまでといったその他いかなる好適な数であっても全く同じようにできる。２つのSPNはその記憶の50%を冗長性のために使用することになり、10個ならそれは10%である。５は記憶の効率と障害の確率との間の妥協点である。

【００１４】

このようにしてコンテンツを分散させることによって、少なくとも二つの目標が達成される。まず、単一のタイトルを視聴できるユーザーの数が単一のSPNセットによってサービスを受けられる数に限定されず、その限界はSPNセットすべてを合わせた通信帯域幅によって決まるようになる。したがって、各コンテンツタイトルのコピーは一つしか必要とされない。その代償は、毎秒立ち上げることのできる所与のタイトルのための新規視聴者の数の制限であるが、これは冗長記憶による無駄なスペースおよび管理上のオーバーヘッドに比べればはるかに取るに足りない制約である。第二の目標とは、ICE１００の全体的な信頼性の向上である。単一ドライブの障害はパリティドライブを使ってのその内容のリアルタイムの再生成によって隠蔽される。独立したディスクの冗長なアレイ（RAID: redundant array of independent disks）と同様である。SPN１０３の障害は、それが、それぞれは動作し続けるいくつかのRAIDセットのうちのそれぞれから一つのドライブを含んでいるという事実によって隠蔽される。障害のあったSPNに接続していたユーザーは非常にすばやく他のSPNで走っている影のプロセスによって引き継がれる。ディスクドライ

ブの、あるいはあるSPN全体の障害の場合には、障害のあった装備を修理または交換するよう運用者に通知される。紛失したサブチャンクがユーザープロセスによって再構築されると、それはそれを提供したはずだったSPNに送り返され、そこでRAM内にキャッシュされる（あたかもそのローカルディスクから読まれた場合のように）。これにより、人気のあるタイトルについて同じ再構築をすることにおいて他のユーザープロセスの時間を無駄にすることが回避される。その人気がサブチャンクがキャッシュされ続けるのに十分なほどでありさえすれば、その後の要求はRAMから充填されるからである。

【００１５】

各「ユーザー」SPN１０３で走っているユーザープロセス（UP: user process）の目標は、自分のディスクからのサブチャンクと他のユーザーSPNからの対応する４つのサブチャンクとを収集して、配送のためにビデオコンテンツのチャンクを組み立てることである。ユーザーSPNは一つまたは複数の管理用の（management）MGMT SPNからは区別される。後者は同じように構成されるが、後述するように異なる機能を実行する。信頼性およびパフォーマンスを向上させるために、冗長なMGMT SPNのペアが考えられる。各UPによって実行される収集および組み立て機能は、各ユーザーSPN１０３上の多くのユーザーのために何度もなされる。結果として、ユーザーSPN１０３どうしの間にはかなりの量のデータトラフィックが行き交うことになる。そうでなければパケット衝突検出および再試行をもつ典型的なイーサネット（登録商標）プロトコルは圧倒されるだろう。典型的なプロトコルはランダムな送信のために設計されており、それらのイベントの間の不活発な時間をあてにしている。よって、このアプローチは使用されない。ICE１００では、衝突は全二重の、フルスイッチ式（fully switched）アーキテクチャを使うことによって、そして通信帯域幅を注意深く管理することによって回避される。ほとんどの通信は同期的になされる。のちにさらに説明するように、スイッチ１０１そのものは同期的な仕方で管理され、伝送の調整がされる。どのSPN１０３がいつ送信できるかが決定されているので、ポートが所与の期間の間にさばける以上のデータで圧倒されることはない。実際、データはまずユーザーSPN１０３のメモリ１０９において収集され、次いでその転送が同期的に管理される。調和の一環として、ユーザーSPN１０３どうしの間に状態信号がある。エンドユーザーに向かう実際のコンテンツとは異なり、ユーザーSPNユニットどうしの間の信号のためのデータサイズはきわめて小さい。

【００１６】

各サブチャンクの長さ（約512Kバイト；ここで「K」は1024）は、もしそうではなくサブチャンクの送信がランダムまたは非同期的に行われることが許されるとしたら、GbEスイッチ１０１において利用可能ないかなるバッファリングをも圧倒してしまうだろう。これだけの情報を送信するための期間は約4ミリ秒（ms）であり、複数のポートが単一のポートに同時に送信しようとしないことを保証することが望まれる。したがって、のちにさらに説明するように、スイッチ１０１は、すべてのポートが最大負荷条件のもとでフルに利用されて同期的に動作するように管理される。

【００１７】

ファイルシステム（あるいは仮想ファイルシステム（virtual file system）すなわちVFS）を管理する冗長ディレクトリプロセスは、ユーザーによる要求があったときに所与のコンテンツタイトルがどこに保存されているかを報告することを受け持っている。それはまた、新たなタイトルをロードすべき時に必要になる記憶スペースを割り当てることも受け持っている。すべての割り当ては一体の諸チャンクにおいてなされ、その各チャンクは５つのサブチャンクからなる。各ディスクドライブ上のスペースは当該ドライブ内では論理ブロックアドレス（LBA: Logical Block Address）によって管理される。一つのサブチャンクはあるディスクドライブ上で連続的なセクタまたはLBAアドレスに保存される。ICE１００における各ディスクドライブの容量は、その最大LBAアドレスをサブチャンクあたりのセクタ数で割ったもので表される。

【００１８】

各タイトルマップまたは「ディレクトリ項目」は、そのタイトルの諸チャンクがどこに

保存されているか、そしてより特定的には各チャンクの各サブチャンクがどこに位置しているかを示すリストを含んでいる。図示した実施例では、リスト中で一つのサブチャンクを表す各項目には、特定のユーザーSPN１０３を同定するSPNID、同定されたユーザーSPN１０３の特定のディスクドライブ１１１を同定するディスクドライブ番号（DD#）およびサブチャンクポインタ（または論理ブロックアドレスすなわちLBA）が64ビット値としてパックされて含まれている。各ディレクトリ項目は、公称4Mbpsでの約30分のコンテンツのためのサブチャンクリストを含む。これは450チャンク、すなわち2250サブチャンクに等しい。各ディレクトリ項目は補助データを含めて約20KBである。SPN上で実行されているUPがディレクトリ項目を要求すると、その項目全体が送られ、対応するユーザーのためにローカルに保存される。たとえSPNが1000ユーザーをサポートしていたとしても、ローカルなリストまたはディレクトリ項目のために消費されるメモリは20MBでしかない。

【００１９】

ICE１００はあるユーザーに利用可能な全タイトルのデータベースを維持している。このリストはローカルな光ディスクライブラリ、リアルタイムネットワークプログラミングおよび使用許諾および転送の手配がなされているところではリモート位置にあるタイトルを含む。このデータベースは各タイトルについての全メタデータを含む。それには管理情報（使用許諾期間、ビットレート、解像度など）とともにユーザーにとって関心のある情報（プロデューサー、監督、キャスト、スタッフ、原作者など）も含まれる。ユーザーが選択をすると、仮想ファイルシステム（VFS）２０９（図２）のディレクトリに問い合わせがされて、そのタイトルがすでにディスクアレイにロードされているかどうかが判別される。まだであれば、ロードプロセス（図示せず）がそのコンテンツ作品について開始され、必要ならいつ視聴のために利用可能になるかについてUPに通知される。たいていの場合、遅延は光ディスク取得ロボット（図示せず）の機械的遅延、すなわち約30秒を超えない。

【００２０】

光ディスク（図示せず）に保存されている情報は全メタデータ（ディスクが最初にライブラリにロードされたときにデータベースに読み込まれる）とともに、当該タイトルを表す圧縮されたデジタルビデオおよびオーディオならびにそれらのデータストリームについて事前に知得できる全情報を含む。たとえば、クロック値およびタイムスタンプのようなデータストリーム中のあらゆる有意な情報へのポインタを含んでいる。それはすでにサブチャンクに分割されており、パリティサブチャンクも事前計算されてディスク上に保存されている。一般に、ロード時間と処理オーバーヘッドを節約するために事前になしうることは何でも光ディスク上に含められる。

【００２１】

リソース管理システムに含まれるものに、ディスパッチャー（図示せず）がある。UPはこのディスパッチャーを参照してそのストリームについての開始時刻を受け取る（通例要求から数ミリ秒以内に）。ディスパッチャーはシステムへの負荷が均一のままであり、遅延が最小化され、いかなる時点でもICE１００内で要求される通信帯域幅が利用可能な値を超えることのないことを保証する。ユーザーが停止、一時停止、早送り、巻き戻しあるいはその他ストリームの流れを中断する動作を要求したときには常に、その通信帯域幅は割り当てを解除され、新たに要求される何らかのサービス（たとえば早送りストリーム）のために新たな割り当てが行われる。

【００２２】

図２は、本発明のある実施形態に基づいて実装された同期データ転送システム２００を図解する、ICE１００の一部分の論理ブロック図である。スイッチ１０１は、いくつかの例示的なSPN１０３に結合して示されている。SPN１０３には第一のユーザーSPN２０１、第二のユーザーSPN２０３および管理（MGMT）SPN２０５が含まれる。先に注記したように、多くのSPN１０３がスイッチ１０１に結合されており、本発明を解説するために二つのユーザーSPN２０１、２０３のみが示されており、先に述べたようにどのSPN１０３として物理的に実装されてもよい。MGMT　SPN２０５は物理的実装は他のどのSPN１０３とも同様

であるが、一般に特定のユーザー機能よりむしろ管理機能を実行する。SPN２０１は各ユーザーSPN１０３のある種の機能を図解し、SPN２０３は他の機能を図解する。しかし、各ユーザーSPN１０３が同様の機能を実行するよう構成され、SPN２０１について述べた機能（およびプロセス）はSPN２０３で与えることもでき、その逆も言えることは理解されるものである。

【００２３】

先に述べたように、スイッチ１０１はポートあたり1Gbpsで動作し、よって各サブチャンク（約512KB）があるSPNから別のSPNに渡されるのに約4msかかる。各ユーザーSPN１０３は一つまたは複数のユーザープロセス（UP）を実行するが、そのそれぞれが下流のユーザーをサポートするためのものである。あるタイトルの新しいチャンクがユーザー出力バッファ（図示せず）を再充填するために必要とされるとき、リストからの次の５つのサブチャンクが、それらのサブチャンクを保存している他のユーザーSPNに対して要求される。多くのUPが複数サブチャンクを実質同時に要求する可能性があるため、サブチャンク伝送期間はほうっておいたら単一ポートのためのGbEスイッチでさえほとんどどんなものでもバッファリング容量を圧倒してしまうだろう。ましてや全体スイッチの場合はそうである。これは図示したスイッチ１０１について成り立つ。もしサブチャンク伝送が管理されていなければ、それは可能性として各UPのための５つのサブチャンクすべてが同時に返されて出力ポートの通信帯域幅を圧倒する結果につながりうる。ICE１００のSPNの伝送のタイミングを緊密にし、最も枢要なデータが最初に無傷で伝送されるようにすることが望まれる。

【００２４】

SPN２０１は、対応する下流のユーザーにサービスを提供するためにUP２０７を実行しているものとして示されている。ユーザーがタイトル（たとえば映画）を要求し、その要求はUP２０７に転送される。UP２０７はタイトル要求（TR: title request）をMGMT SPN２０５に位置しているVFS２０９（のちにさらに説明する）に送信する。VFS２０９はディレクトリ項目（DE: directory entry）をUP２０７に返し、該UP２０７は２１１で示されるDEをローカルに保存する。DE２１１は当該タイトルの各サブチャンク（SC1、SC2など）の位置を示すリストを含んでおり、各項目には、特定のユーザーSPN１０３を同定するSPNID、同定されたSPN１０３の特定のディスクドライブ１１１を同定するディスクドライブ番号（DD#）および同定されたディスクドライブ上のサブチャンクの特定の位置を与えるアドレスすなわちLBAが含まれている。SPN２０１は、DE２１１内の各サブチャンクについて同時に一つずつタイムスタンプ付き読み出し要求（TSRR: time stamped read request）を開始する。ICE１００では、要求はすぐに直接なされる。換言すれば、SPN２０１はサブチャンクの要求を行うことを、そのデータを保存している特定のユーザーSPN１０３に対してすぐに直接に開始する。図示した構成では、たとえローカルに保存されていても同じようにして要求がなされる。つまり、要求されたサブチャンクがSPN２０１のローカルディスクドライブ上にあったとしても、リモートにあるかのようにスイッチ２０１を通じて要求を送出するのである。ネットワークは、あるSPNからその同じSPNに要求が送られようとしていることを認識するよう構成されていてもよい位置である。すべての場合を同じように扱うほうが簡単である。特に、要求が実際にローカルであることが比較的ありそうもないような大きな施設ではそうである。

【００２５】

要求はすぐに直接に送出されるが、サブチャンクはそれぞれ完全に管理された仕方で返される。各TSRRは当該SPNIDを使っている特定のユーザーSPNに対するものであり、対象となるユーザーSPNがそのデータを取得して返すためのDD#およびLBAを含んでいる。TSRRはさらに、要求されたサブチャンクが適正に適切な要求者に返されることを保証し、要求者がそのサブチャンクを識別することを可能にするのに十分な他のいかなる識別情報をも含みうる（たとえば宛先SPN上で実行されている複数UPの間の区別をするUP識別子、各データチャンクについてのサブチャンクの間の区別をするサブチャンク識別子など）。各TSRRはもともとの要求がなされた特定の時点を同定するタイムスタンプ（TS: timestamp）を

も含む。TSは同期伝送のために要求の優先度を同定する。ここで、優先度は時間に基づいており、早い要求ほど高い優先度を得る。要求されたタイトルの返されたサブチャンクは、受け取られると、さらなる処理およびそのタイトルを要求したユーザーへの配送のためにローカルなタイトルメモリ２１３に保存される。

【００２６】

ユーザーSPN２０３は、TSRRを受け取るため、そして要求されたサブチャンクを返すために各ユーザーSPN（たとえば２０１、２０３）上で実行されている、転送プロセス（TP: transfer process）２１５および支援機能の動作を図解している。TP２１５はストレージ・プロセス（図示せず）を含むか、あるいは他の仕方でストレージ・プロセスとのインターフェースをもつ。ストレージ・プロセスは、保存されているサブチャンクの要求およびアクセスのためのSPN２０３上のローカルなディスクドライブ１１１のインターフェースとなるものである。ストレージ・プロセスは状態機械などのような、いかなる所望の仕方で実装されてもよく、TP２１５とローカルディスクドライブ１１１との間でインターフェースされる別個のプロセスであってもよい。これは当業者には既知のとおりである。図示したように、TP２１５は他のユーザーSPN１０３上で実行されている一つまたは複数のTSRRを一つまたは複数のUPから受け取り、各要求をローカルメモリ１０９内の読み取り要求待ち行列（RRQ: read request queue）２１７に保存する。RRQ２１７はサブチャンクSCA、SCBなどのための要求のリストを保存する。要求されたサブチャンクを保存しているディスクドライブは、対応する要求をRRQ２１７から除去し、物理的な順番にソートし、それからそれぞれの読み出しをソートされた順番で実行する。各ディスク上のサブチャンクへのアクセスはグループで管理される。各グループは「エレベーター・シーク」動作に従って物理的順序にソートされている（低位から高位へ一回の掃引、次に高位から低位への掃引などといった具合にディスクヘッドがシーケンス上次のサブチャンクを読むために一時停止しながらディスク表面を行き来する）。うまくいった読み込みの要求は成功読み込み待ち行列（SRQ: successful read queue）２１８にTS順にソートされて保存される。失敗した読み込みの要求（もしあれば）は失敗読み込み待ち行列（FRQ: failed read queue）２２０に保存され、失敗した情報がネットワーク管理システム（図示せず）に転送され、該ネットワーク管理システムがエラーおよび適切な是正動作を決定する。図示した構成では待ち行列２１７、２１８および２２０が実際のサブチャンクではなく要求情報を保存していることを注意しておく。

【００２７】

読み込みに成功した各サブチャンクは最近要求されたサブチャンクのLRUキャッシュのためにリザーブされているメモリに入れられる。取得された各サブチャンクについて、TP２１５は対応するメッセージ（MSG: message）を生成するが、これはそのサブチャンクのTS、そのサブチャンクのソース（SRC: source）（たとえばそのサブチャンクが送信されてくるもとのSPNIDおよびその物理メモリ位置ならびにその他任意の識別情報）ならびにそのサブチャンクを送信すべき先の宛先（DST: destination）SPN（たとえばSPN２０１）を含む。図示したように、SRQ２１８はサブチャンクSCA、SCBなどに対してそれぞれメッセージMSGA、MSGBなどを含んでいる。要求されたサブチャンクが読み出されてキャッシュされたのち、TP２１５は対応するMSGを、MGMT SPN２０５で実行されている同期スイッチマネージャ（SSM: synchronized switch manager）２１９に送る。

【００２８】

SSM２１９は諸ユーザーSPNのTPから複数MSGを受け取って優先順位付けし、最終的には、SRQ２１８内のMSGのうちの一つをメッセージ識別子（MSGID: message identifier）などを利用して同定する送信要求（TXR: transmit request）をTP２１５に送る。SSM２１９がSRQ２１８内のサブチャンクを同定するMSGIDをもつTXRをTP２１５に送ると、その要求リスト項目はSRQ２１８からネットワーク転送プロセス（NTP: network transfer process）２２１に移され、そこで当該サブチャンクを宛先のユーザーSPNに転送するのに使われるパケットが構築される（ここで、「移す」とはその要求をSRQ２１８からは除去することを表している）。サブチャンク要求リスト項目がSRQ２１８から除去される順序

は、リストがタイムスタンプ順になっているにもかかわらず、必ずしもその順ではない。適正な順序を決定するのはSSM２１９だけだからである。SSM２１９は、送るべきサブチャンクを少なくとも一つもつ他のすべてのSPN１０３に対して一つのTXRを送る。ただし、サブチャンクを送るべき先がすでに同等以上の優先度のサブチャンクを受け取るようスケジュールされているSPN１０３上のUPである場合は別であり、これについてはのちにさらに説明する。SSM２１９は次いで全ユーザーSPN１０３に対して単一の送信コマンド（TX CMD: transmit command）をブロードキャストする。SSM２１９によってブロードキャストされたTX CMDコマンドに反応して、TP２１５はNTP２２１に、そのサブチャンクをユーザーSPN１０３の要求UPに送信するよう命令する。このようにして、SSM２１９からTXRを受け取った各SPN１０３は同時に別の要求元ユーザーSPN１０３に送信を行う。

【００２９】

MGMT SPN２０５上のVFS２０９はタイトルのリストおよびICE１００内でのその位置を管理する。典型的なコンピュータシステムでは、ディレクトリ（データの情報）は通例当該データが存在するのと同じディスク上にある。ところがICE１００では、VFS２０９は中央に位置していて分散データを管理する。各タイトルのデータがディスクアレイ中の複数のディスクにまたがって分散しており、前記複数のディスク自身も複数のユーザーSPN１０３にまたがって分散しているからである。先に述べたように、ユーザーSPN１０３上のディスクドライブ１１１は主としてタイトルのサブチャンクを保存する。VFS２０９は、先に述べたようにSPNID、DD#およびLBAを通じて各サブチャンクの位置を決めるための識別子を含んでいる。VFS２０９はまた、光記憶装置のようなICE１００の外部にある他の部分のための識別子をも含んでいる。ユーザーがあるタイトルを要求すると、そのユーザー要求を受け付けたユーザーSPN１０３上で実行されているUPにはディレクトリ情報（ID／アドレス）の完全なセットが利用可能となる。そこからは、すべきことはサブチャンクをディスクドライブからメモリ（バッファ）へと転送し、該サブチャンクをスイッチ１０１を通じて要求元ユーザーSPN１０３に移すことである。要求元ユーザーSPN１０３はバッファ中で完全なチャンクを組み立て、それをユーザーに配送し、終了するまで繰り返す。

【００３０】

SSM２１９は、「準備完了」メッセージ（RDY MSG: ready message）リスト２２３中にタイムスタンプ順の準備完了メッセージのリストを作成する。ユーザーSPN１０３上でTPからメッセージが受け取られる順番は必ずしもタイムスタンプ順ではないが、RDY MSGリスト２２３中ではTS順に整理されている。次の一組の転送の直前に、SSM２１９は最も早いタイムスタンプから始まってRDY MSGリスト２２３をスキャンする。SSM２１９はまずRDY MSGリスト２２３内で最も早いTSを同定し、対応するTXRメッセージを生成して、対応するサブチャンクを保存しているユーザーSPN１０３のTP２１５に送り、それによりそのサブチャンクのペンディングの転送を開始させる。SSM２１９は後続の各サブチャンクについてTS順にリスト２２３のスキャンを続け、すでにペンディングのサブチャンク転送に関わっているのでないソースと宛先をもつ各サブチャンクについてTXRメッセージを生成する。全ユーザーSPN１０３にブロードキャストされる一つ一つのTX CMDに対しては、各ユーザーSPN１０３が送信するサブチャンクは同時に一つのみであり、受信するサブチャンクは同時に一つのみである。ただし、この両者は同時にできる。たとえば、SPN#2へのペンディングのサブチャンク転送をスケジュールするためにTXRメッセージがSPN#10のTPに送られた場合、SPN#10は同時に別のサブチャンクを送ることはできない。しかし、SPN#10は同時に別のSPNからサブチャンクを受信することはできる。さらに、SPN#2はSPN#10から前記サブチャンクを受信しながら同時に別のサブチャンクを受信することはできないが、SPN#2は同時に別のSPNに送信をすることはできる。これはスイッチ１０１の各ポートの全二重性のためである。

【００３１】

SSM２１９はユーザーSPN１０３が全部ふさがってしまうまで、あるいはRDY MSGリスト２２３の終わりに達するかするまでRDY MSGリスト２２３のスキャンを続ける。RDY MSGリスト２２３中のTXRメッセージに対応する各項目は最終的にはRDY MSGリスト２２３か

ら除去される（TXRメッセージが送られたとき、あるいは転送が完了したあとに）。前の周期の最後の転送が終わったとき、SSM２１９は全ユーザーSPN１０３に次のラウンドの送信を開始するよう合図するTX CMDパケットをブロードキャストする。各転送は、図示した特定の構成については約4ないし5msの周期内で同期的に起こる。各転送ラウンドの間、追加的なMSGがSSM２１９に送られ、次の送信ラウンドをスケジュールするために新しいTXRメッセージが諸ユーザーSPN１０３に送出され、そのプロセスは繰り返される。TX CMDから次のTX CMDまでの間の周期は近似的には、一つのサブチャンクのバイトすべてを送信するために必要な、パケットオーバーヘッドおよびパケット間遅延を含めた期間に、そのサブチャンクの送信中にスイッチ内に発生したかもしれないすべてのキャッシュをクリアするための期間（典型的には60マイクロ秒（μs））と、個々のSPNによるTX CMDの認識における遅延によって引き起こされるあらゆるジッターのための期間（典型的には100μs未満）とを加えたものに等しい。

【００３２】

ある実施形態では、重複した、すなわちミラーされたMGMT SPN（図示せず）が主MGMT SPN２０５をミラーしており、SSM２１９、VFS２０９およびディスパッチャーがそれぞれ一対の冗長な専用MGMT SPNの上で重複される。ある実施形態では、同期TX CMDブロードキャストがMGMT SPN２０５の健全性を示す鼓動としてはたらく。鼓動とは、副次的なMGMT SPNに対する万事良好という信号である。鼓動がないと、たとえば5msなど所定の時間期間内のいっさいの管理機能を副次的MGMT SPNが引き継ぐのである。

【００３３】

本発明はある種の好ましいバージョンを参照しつつかなりの詳細において記載されてきたが、他のバージョンおよび変形も可能であり、考えられている。当業者は開示された概念および具体的な実施例を、請求項によって定義される本発明の精神および範囲から外れることなく、本発明と同じ目的に資するための他の構造を考案したり修正したりする基礎として容易に使うことができる。

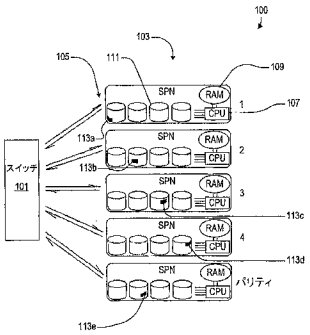
【図面の簡単な説明】

【００３４】

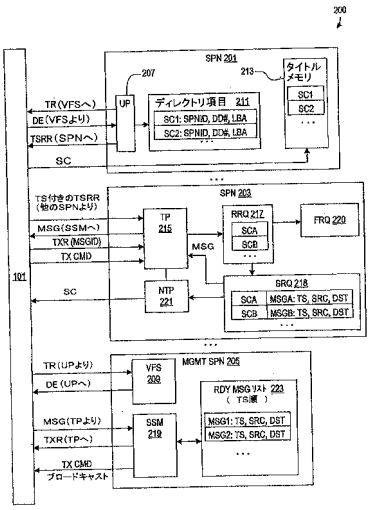
【図１】本発明の例示的な実施形態に基づいて実装された対話的コンテンツエンジン（ICE: Interactive Content Engine）の一部分の単純化したブロック図である。

【図２】本発明のある実施形態に基づいて実装された同期データ転送システムを図解する、図１のICEの一部分の論理ブロック図である。

【図１】


100
103
105
111
109
107
SPN
RAM
CPU
1
113a
2
113b
3
113c
4
113d
スイッチ
101
パリティ
113e

【図２】


200
SPN 201
207
213
タイトル
メモリ
SC1
SC2
UP
TR(VFSへ)
DE(VFSより)
TSRR(SPNへ)
SC
ディレクトリ項目 211
SC1: SPNID, DD#, LBA
SC2: SPNID, DD#, LBA
TS付きのTSRR
(他のSPNより)
SPN 203
MSG(SSMへ)
TXR(MSGID)
TX CMD
TP
215
MSG
RRQ 217
SCA
SCB
FRQ 220
101
SC
NTP
221
SRQ 218
MSGA: TS, SRC, DST
MSGB: TS, SRC, DST
TR(UPより)
DE(UPへ)
MGMT SPN 205
VFS
209
RDY MSGリスト 223
(TS順)
MSG(TPより)
TXR(TPへ)
SSM
219
MSG1: TS, SRC, DST
MSG2: TS, SRC, DST
TX CMD
ブロードキャスト

フロントページの続き

審査官　古川　哲也

(56)参考文献　特開平０７－１４１２３２（ＪＰ，Ａ）
特開平０７－２３６１３２（ＪＰ，Ａ）
特開平０８－１０７４２２（ＪＰ，Ａ）
特開平０８－１０７５４２（ＪＰ，Ａ）
特開平０９－０９７１３６（ＪＰ，Ａ）
特開平０９－１３８７３５（ＪＰ，Ａ）
米国特許第０６１３４５９６（ＵＳ，Ａ）
国際公開第０１／１０１２５（WO，Ａ１）

(58)調査した分野(Int.Cl.，ＤＢ名)
H04N　7/16　-　7/173
G06F　3/06
H04L　12/40