

HiTactix-BSD 連動システムを応用した大規模双方向メディアサーバの設計と実装

竹内 理, レモアル・ダミエン
(株)日立製作所 システム開発研究所

1. はじめに

CATV 網や FTTH(Fiber To The Home)計画の進行により、一般家庭から大容量のメディアデータの発信が可能になりつつある[1]。それに伴い、クライアント（一般家庭）からの大容量メディアデータのアップロードと、クライアントへの大容量メディアデータの配送を並行実行可能な、大規模双方向メディアサーバの実現に対する要求が高まりつつある。

この大規模双方向メディアサーバは、以下の要件を充足する必要がある。

- A) 高い I/O 実行性能を保持すること
- B) メディアデータアップロード処理時にも、メディアデータの配送レート、すなわち、
 - B 1) ディスク読み出しレート、及び、
 - B 2) ネットワーク送信レート
 を保証可能であること
- C) 他アプリケーションとの連動や、ネットワーク経路での管理が可能であること

我々は、HiTactix-BSD 連動システム[2]上に、新規の I/O 方式、及びディスク・スケジューリング方式（サイクリック・パイプライン I/O 方式、及びサイクリック・ディスク I/O スケジューリング方式）を実装し、上記要件を同時に充足可能な大規模双方向メディアサーバを実現した。本稿では、本メディアサーバのシステム構成の概要、新規に実装した I/O 方式の概要、及びその性能評価結果につき述べる。

2. システム構成概要

本メディアサーバのシステム構成概要を図 1 に示す。

- 1) 本メディアサーバは SMP ハードウェア（2CPU 搭載）上で動作する。SMP ハードウェアの 1CPU を占有して汎用 OS（BSD/OS）が、他 CPU を占有してメディア処理向きカーネル（HiTactix）が動作する。
- 2) 「メディア・サーバアプリケーション」は、BSD/OS 上で動作する。BSD/OS の機能を用いることにより、本アプリケーションを他アプリケーションと連動させたり、ネットワーク経由で管理することができる。
- 3) 「メディアサーバ・アプリケーション」は、メディアデータの配送開始/停止要求を受理した場合、HiTactix 上の「メディアデータ配送制御スレッド」に対して当該要求を転送する。開始要求発行から停止要求発行までの間、「メディアデータ配送制御スレッド」は周期的に、定レートでのディスク読み出し、及びネットワーク送信を繰り返す（サイクリック・パイプライン I/O 方式）。
- 4) 「メディアサーバ・アプリケーション」は、メディアデータのアップロード要求を受理した場合、アップロードデータのディスクへの書き込み要求を「ディスク I/O 制御スレッド」に対して発行する。「ディスク I/O 制御スレッド」は、サイクリック・ディスク I/O スケジューリング方式（詳細は本稿では省略する）を用いて、当該メディアデータの書き込みが発生しても、「メディアデータ配送制御スレッド」のディスク読み出しレートが変動しないことを保証する。

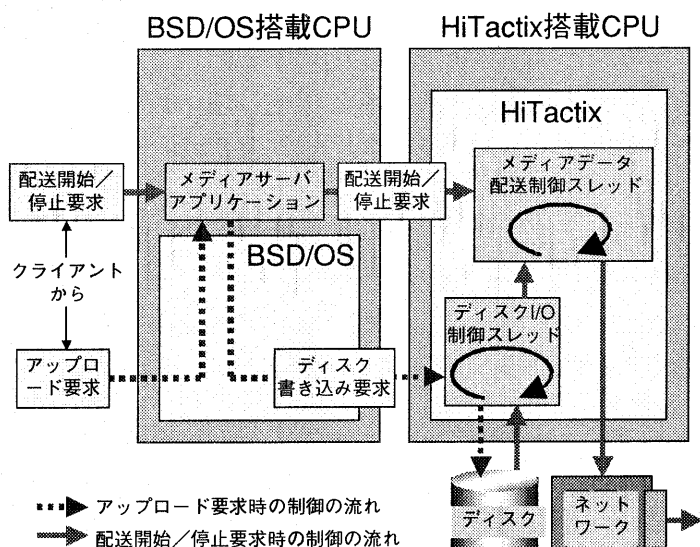


図1 メディアサーバのシステム構成

Design and Implementation of a Highly Scalable Bi-directional Media Server Using Cooperating HiTactix-BSD System
Tadashi Takeuchi, Le Moal Damien
Systems Development Laboratory, Hitachi, Ltd.

上記 2) から 1 節で示した要件 C) を、また上記 4) から 1 節で示した要件 B 1) を本メディアサーバは充足していることがわかる。次節でサイクリック・パイプライン I/O 方式の概要を示し、本メディアサーバが 1 節で示した要件 A) 及び要件 B 2) を充足していることを示す。

3. サイクリック・パイプライン I/O 方式の概要

既存 OS における従来 I/O 方式を図 2 に示す。大規模双方向メディアサーバに従来 I/O 方式を用いると、以下の問題が発生する。

- 1) 以下の理由により、高い I/O 実行性能を確保できない。
 - イ) ユーザバッファとカーネルバッファ間のメモリコピーが多発する。
 - ロ) アプリケーションからの API 呼び出し（システムコール発行）が多発する。
 - 2) アプリケーションの周期的なスケジューリングを保証できないため、ネットワーク送信レートの一定性も保証できない。
- サイクリック・パイプライン方式は以下の特徴を持った

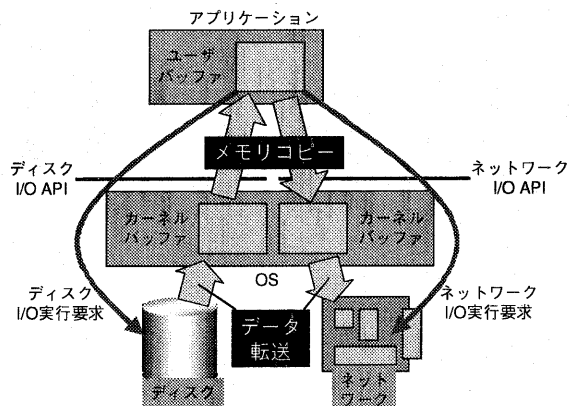


図2 従来 I/O 方式

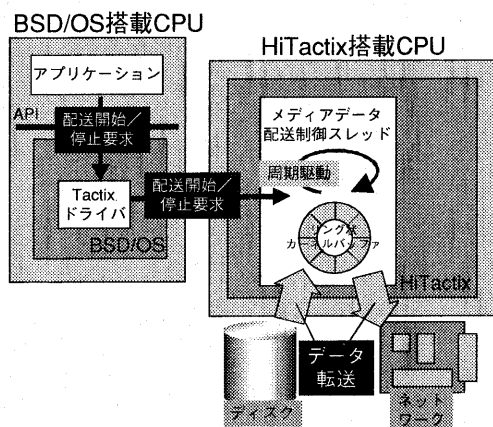


図3 サイクリック・パイプライン I/O 方式

め、上記問題を解決できる（図3参照）。

- 以下の理由により高い I/O 実行性能を保持できる。
 - リング状カーネルバッファを介し、メモリコピーなく I/O を実行する。
 - アプリケーションからの API 呼び出しが、配送開始時と停止時以外に必要な。
- アイソクロナス・スケジューラ[3]を保持する HiTactix が、周期的に「メディアデータ配送制御スレッド」をスケジューリングすることを保証するため、ネットワーク送信レートの一定性も保証できる。

4. 性能評価

本メディアサーバの

- I/O 実行性能
- 配送レート保証性能

の性能評価を行なった。本稿では、上記 B) の評価実験の概要と、その評価結果につき述べる。

配送レート保証性能評価実験システム概要を以下に示す。

- PC/AT 互換機 (PentiumIII 600MHz 搭載) をメディアサーバとして用いて、メディアデータ (1 ストリームあたりの要求配送レートは 38.4Mbps) のディスクからの読み出し、及びネットワークへの送信を実行した。同時に読み出すメディアデータの数 (ストリーム数) を 1 と 2 の場合を実行し、それぞれの場合における配送レートの変動を測定した。

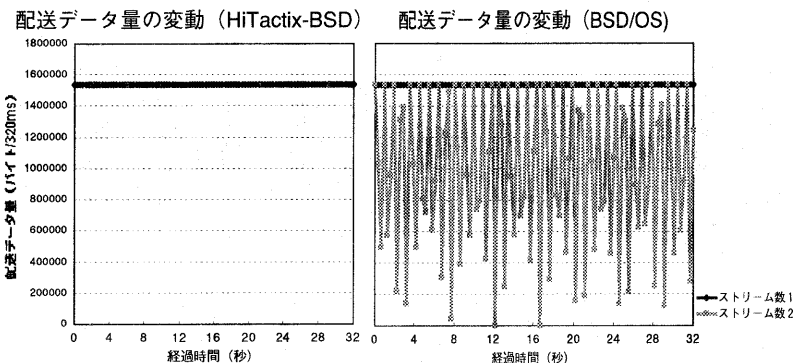


図4 配送レート保証性能評価実験結果

- 実験は、HiTactix-BSD 連動システム上に構築したメディアサーバを用いた場合と、BSD/OS 上に構築したメディアサーバを用いた場合との、2 つの場合につき行ない両者の結果を比較した。

実験結果を図4に示す。図中のグラフの横軸は経過時間を、縦軸が単位時間 (320ms) あたりに配送したデータ量を示す。この結果から明らかなように、ストリーム数が1の場合は、両メディアサーバ間の配送レート保証性能に差異はない。しかし、ストリーム数が2になると、BSD/OS 上のメディアサーバでは 36.0% の配送レートの変動が発生しているのに対し、HiTactix-BSD 連動システム上のメディアサーバではこの変動が発生しない。なお、ここで言う「配送レートの変動」とは、以下の式で求まる値を意味する。

$$\begin{aligned}
 (\text{配送レートの変動}) = & \\
 & | (\text{要求配送レート}) - (\text{実際の配送レート}) | \\
 & / (\text{要求配送レート}) \times 100 (\%)
 \end{aligned}$$

5. まとめ

本稿では、HiTactix-BSD 連動システム上に構築した大規模双方向メディアサーバの概要につき述べた。さらに、その配送レート保証性能の評価結果につき述べた。評価の結果、BSD/OS 上のメディアサーバでは 36.0% の配送レートの変動が発生するが、本メディアサーバでは上記変動発生を抑制可能であることが明らかになった。

参考文献

- 渡辺隆市他, 「多彩なサービスの経済的な提供を実現する光アクセス技術」, NTT 技術ジャーナル, pp12-40, Mar. 2000.
- 竹内 理他, 「OS 接続モジュール Symbiose を用いた BSD-HiTactix 連動システムの設計と実装」, 情報処理学会研究報告, 2000-OS-83, pp31-36, Feb. 2000.
- 竹内 理他, 「連続メディア処理向き OS の周期駆動保証機構の設計と実装」, 情報処理学会論文誌, Vol. 40, No. 3, pp1204-1215, Mar. 2000

BSD/OS は BSDI 社の登録商標です。

PentiumIII は米国 Intel 社の登録商標です。