

分散マルチメディアシステム DMSIC における同期機構の実現と評価

大野 隆一, 相田 仁, 齊藤 忠夫

東京大学工学部

本論文では, Ethernet でつながれた複数の UNIX ワークステーション上に現在構築中の分散マルチメディアシステム (DMSIC) において蓄積型情報源から取り出されたメディア情報 (RMS) と参加者の声などの生のメディア情報 (LMS) との間の同期 (R&L 同期) を維持する機構を示す. 本機構の評価を行った結果, 本機構を用いることで R&L 同期のセマンティクスを満たしながら各ノードで情報提示が行なわれることを確認した.

Implementation and Evaluation of a Media Synchronization Mechanism in a Distributed Multimedia System *DMSIC*

Ryuichi Ohno, Hitoshi Aida, and Tadao Saito

Faculty of Engineering, the University of Tokyo

In this paper, we propose a technique to realize R&L synchronization (synchronization between *Retrieved Media Streams* (RMS) and *Live Media Streams* (LMS)) at the time of presentation in a *Distributed Multimedia System with Interactive Control* (DMSIC) which has been implemented on UNIX workstations connected by Ethernet. An evaluation of this technique shows that media units are presented at each node, satisfying the semantics of R&L synchronization.

1 まえがき

近年, ネットワーク技術及びコンピュータ技術の発展に伴い, ネットワークを介してマルチメディア情報を送ることで会議, 教育などの様々な活動を支援するシステムの研究・開発が盛んに行なわれてきている.

このようなシステムにおいては様々な情報源から送られてくる複数のメディア情報を各ノードで提示する際にメディア間の同期を取ることが必要となる.

従来, ビデオ・オン・デマンドのように一つの蓄積型装置から取り出されたマルチメディア情報を一つのノードで提示するシステムにおいてはメディア間同期について様々な検討がなされてきたが, 多地点間の会議や教育を支援するシステムのように共通の情報が複数のノードにおいてほぼ同時に提示されるようなシステムにおいてはメディア間同期の問題はあまり取り上げられてこなかった. 我々はディスク, テープなどの蓄積型情報源から引き出されたメディア情報 (RMS) とその説明を行なう参加者の声などの生のメディア情報 (LMS) を分けて考え, その間の同期 (R&L 同期) を維持することで多地点間の会議や教育を支援するシステムにおけるメ

ディア間同期の問題に対応する.

本論文では, 現在試作中のシステムにおいて R&L 同期を取る機構を示す. 論文中には, まず, システムの概念, 及び, 現在試作中のシステムの詳細について述べる. 次に, 試作中のシステムにおいて R&L 同期を維持する手法について述べる. 最後に, 本同期方式の性能評価について述べる.

2 システムの概要

2.1 システムの概念

システムを概念を図 1 に示す.

各ノードにはディスプレイ装置などのマルチメディア表示用の装置, 及び, メディア間のずれの吸収などに用いられるメモリが備わっている. また, システム中の幾つかのノードにはメディア情報の入ったディスク, テープなどの情報発生源となる装置が備わっている.

本システムの基本的な動作を以下に示す.

- システム中の一つのノードの利用者が操作者となる. 操作者ノードは再生, 巻き戻しなどの指示を出

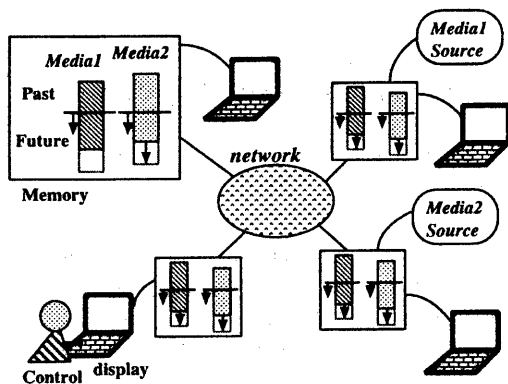


図 1: システムの概念

す。この操作命令はシステム中の各ノードにマルチキャストされる。

- 情報発生源ノードでは操作者ノードからの指示に従いメディア送出位置を変化させる。
- 各情報提示ノードでは情報発生源ノードから送られてきたメディア情報を一旦バッファ中に蓄える。各ノードでは操作者ノードからの指示に従い情報提示位置を変えつつも、メディア間同期を保ちながらメディア情報の提示を行なう。

上記の動作により、本システムでは各ノードにおいて 2.2 節で述べる R&L 同期のタイミングを満たしながら同一のメディア情報が提示される。

主なメディア操作としては、再生、一時停止、早送り、巻き戻し、コマ送りなどが考えられる。但し、操作を行なえる人は一時に一人に限るとし、操作権がノード間を移動するといった方法で操作権の委譲が行なわれるとする。

本システムでは、すべてのメディア情報源でのメディアフレーム発生間隔を 33 ミリ秒といった共通の値とする。各フレームには予めフレーム番号が付けられ、同じフレーム番号を持つメディアフレームを同時に提示することでメディア間の同期を取る。

このフレーム番号は RMS 中の各フレーム、操作命令パケット、バッファリングされている情報などに付加される。また、LMS 中の各フレームにも、そのフレームが送信された時に提示されていた RMS 中のメディアフレームのフレーム番号を付加する。

2.2 R&L 同期のセマンティクス

R&L 同期のセマンティクスを以下に示す。ここで、RMS 間の同期はすべて厳密に取るものとする。

sync_with_one RMS と操作者からの LMS との同期。すべてのノードにおいて、操作者ノードにおいて LMS が取り込まれたのと同じタイミングで、LMS が RMS と同期して提示されること。これは、操作者以外のノードにおける再生点を操作者ノードでの再生点よりも、LMS の圧縮、伝送、及び、伸長にかかる時間の分だけ遅らせることで実現できる。

このセマンティクスは、例えば、遠隔教育を支援するシステムにおいて教師が教育用教材を操作しながら授業を進めていくような場合に適用される。

sync_with_all RMS と全ての参加者からの LMS との同期。すべてのノードでの再生点が各時刻において等しいこと。

このセマンティクスは、例えば、遠隔会議を支援するシステムにおいて全ての参加者が会議資料を見ながら議論に参加するような場合に適用される。

sync_with_multi RMS と操作者を含む複数の参加者(発信者ノード)からの LMS との同期。

すべての発信者ノードでの各時刻における再生点が等しく、それ以外のノードでは、発信者ノードにおいて LMS が取り込まれたのと同じタイミングで、LMS が RMS と同期して提示されること。

このセマンティクスは、例えば、遠隔討論会を支援するシステムにおいて討論参加者が討論会資料を見ながら議論を進めていくような場合に適用される。

sync_with_none RMS と LMS との同期は考えず、RMS 間の同期のみ保証する。

3 試作システムにおける同期方式

3.1 試作システムの概要

本節では Ethernet (10Mbps) で結ばれた複数台の UNIX ワークステーション上に実装した試作システムの概要について述べる。

試作システムではシステム中に一つ存在し、初期化などの際に用いられる管理プロセスと各参加ノードごとに起動される動作プロセスの 2 種類のプロセスを用いる。また、これらのプロセス間に各メディア情報及び制御情報ごとに connection を張ることで情報をやりとりする。情報送出、表示などは各動作プロセス中で定期的にスケジュールされる procedure を用いて行なう。

情報源としてはハードディスク中に格納された動画像 (1 フレーム 21,600 bytes)、テキスト情報 (1 つにつき 20 bytes)、及び音声 (8000 bytes/sec) の 3 つを用い、操作命令としては再生、一時停止、逆再生、早送り、巻き戻しなどの命令をサポートしている。

LMS としては音声情報のみを用いる。操作者ノードを含む複数のノードが発信者ノードとなり、発信者ノードから各ノードに音声情報が送られる。

発信者ノードでは、話したい時のみ (talk ボタンを押すことで) 音声デバイスを open する (つまり、read 用の音声 queue に情報を取り込む。)

3.2 同期方式の概要

本システムのように QOS を指定することで資源を確保することができないネットワークと端末により構成されるシステムの場合、他のプロセスによる影響を排除できない。従って、このような環境でメディア間の同期を取るには各ノードでの負荷変動、遅延時間変動などを吸収するための何らかの機構が必要となる。

本システムでは各ノードのバッファ中に理想的な再生点を定め、この理想的な再生点に各メディアの再生点を近付けることで各ノードにおけるメディア間の同期を取る (R&L 同期^[2])。また、この理想的な再生点を、R&L 同期のセマンティクスを満たすようにノード間で調整することで R&L 同期を実現する (図 2)。

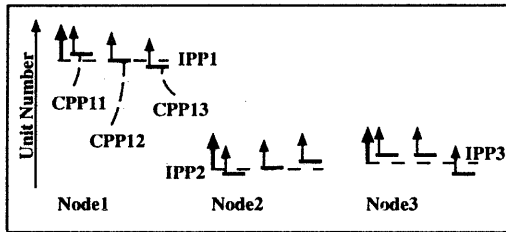


図 2: A media synchronization mechanism

4 R&L 同期方式

4.1 ノード間での再生点調整

R&L 同期のいずれかのセマンティクスを満たすためには、システム中の各ノードの再生点は以下に示す 2 つのタイプのどちらかに属さなければならない。

e_type グローバルな時間軸上で再生点が操作者ノードの再生点と完全に一致するノード。

d_type グローバルな時間軸上で再生点が操作者ノードの再生点よりも幾分遅れるノード。

ここで、各ノードの時計のずれ程度の同期ずれは許容するものとし、グローバルな時間軸を各ノードの時間とする。

また、操作命令への付加情報として、命令開始位置以外にその命令が出された時刻も付加するものとする。

4.1.1 操作命令に対する提示開始タイミング

一時停止状態で操作者ノードから再生開始命令が発せられた場合、(操作者ノードを含む) 各ノードでは以下のように対応する。

e_type のノード [再生命令が操作者ノードで発せられた時間] + [一定の待ち時間 (T_w)] 後に再生を開始する。再生開始命令受信時に既にこの時間を過ぎていた場合、即座に再生を開始する。

d_type のノード [再生命令を受け取った時間] + [一定の待ち時間 (T_w)] + [audio queue が吐き出されるのにかかる時間 (T_q)] 後に再生を開始する。

ここで、 T_w は `sync_with_one` の場合 0、それ以外の場合命令の伝送遅延時間程度の値とする。

また、負荷変動の影響を避けるため audio queue は一定に保たれる必要がある^[2]。従って、**d_type** のノードでは T_q だけ余分に再生点を遅らせる必要がある。

一時停止命令に関してはこのような待ち時間を考える必要はないが、それ以外の命令に関しては同様にして各ノードでの情報提示開始タイミングを調整することで R&L 同期が満たされるようにする。

`sync_with_multi` の場合の情報提示開始タイミングの一例を図 3 に示す。

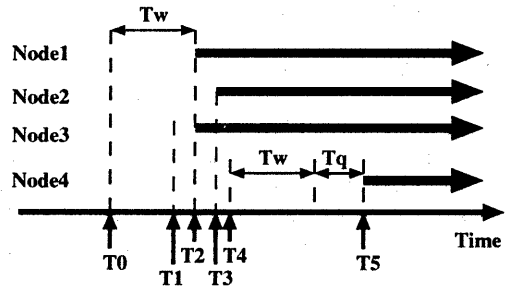


図 3: `sync_with_multi` における提示開始タイミング

図 3 において Node1 (操作者ノード)、Node2、Node3 が **e_type** のノード、Node4 が **d_type** のノードとする。

時刻 T_0 において Node1 において命令が出されると、 T_w 後の時刻 T_2 が **e_type** のノードに対する理想的な再生開始時間となる。Node3 への命令到着時間を T_1 とすると、まだ T_2 に達していないので Node3 は Node1 と同時に T_2 から情報提示を開始できる。Node2 への命令到着時間を T_3 とすると、既に T_2 を過ぎているため即座に情報提示を開始する。このように、 T_w が小さいと **e_type** のノードが同時に提示開始を行なうことが難しくなるが、 T_w が大きくなると待ち時間が大きくなるため望ましくない。Node4 への命令到着時間

をT4とすると、Node4ではさらにTw+Tq時間だけ待った後で時刻T5から情報提示を開始する。

4.1.2 理想的な再生点の決定

理想的な再生点(F_{IPP})は4.1.1節で述べた理想的な命令開始時間(T_{Istart})からの経過時間(T_{elapse})と、命令開始フレーム番号(F_{Istart})から式(1)により算出する。

$$F_{IPP} = F_{Istart} + T_{elapse} \times P_{rate} + F_{offset} \quad (1)$$

ここで、 P_{rate} は情報提示速度(33frame/sec等、逆再生の際には負の値)を示す。また、 F_{offset} は4.1.3節で述べる情報提示中の再生点調整による理想的な再生点の変化分を示す。

4.1.3 情報提示中の調整

ネットワークの遅延変動などにより、以下のような問題が生じる。

1. ネットワークの負荷が大きくなると、d.typeのノードでLMSがRMSよりも遅れて到着するようになる。
2. 様々な変動により、各ノードでの再生点がバッファ上で適切な位置を取れなくなる。

1に関しては、各d.typeノードにおいてLMSを受信した際に、 F_{offset} を調整することで対処することとする。

2に関しては、グローバルな調整が必要となる。そこで、システム中の一つ存在する管理プロセスが情報提示中に各ノードの状態情報を収集し、理想的な再生点を動的に変化させることで情報提示中の再生点の調整を行なうこととする。

例えば、図4では、全体的に各ノードの再生点がバッファのTailに近い位置にある。この場合、ノード間の再生点の位置関係を保ったまま各ノードの再生点をバッファの先頭に近付けることでpastバッファをある程度確保することが望ましい。

このようなシステムワイドな再生点調整を以下のように実装した。

1. 各ノードは定期的に各ノードの状態情報を管理プロセス(初期化プロセスと共用)に送る。ここで、各ノードの状態情報としては、各バッファの先頭フレーム番号、各バッファの大きさ、理想的な再生点、現在の動作モード(再生中等)及び、状態情報送出時刻とする。また、一時停止中には状態情報は送らない。

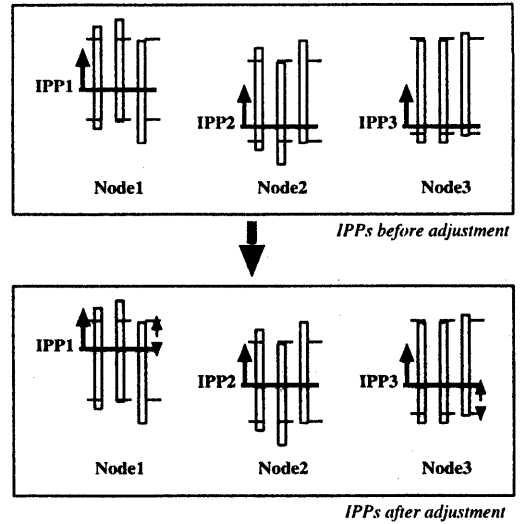


図4: Adjustment of IPPs during Presentation

2. 管理プロセスでは、全てのノードからの同一時刻での状態情報が揃った時点で、理想的な再生点を全体として前後にどれだけ動かすか決める。ここで、IPPの調整は各ノードで利用可能なfutureバッファの大きさのminimumとpastバッファのminimumとが等しくなる位置にIPPが来るように行なった。
3. 管理プロセスはIPPの調整値を各ノードに送る。
4. 各ノードでは送られてきた調整値を F_{offset} に加える。

4.2 Liveメディア情報

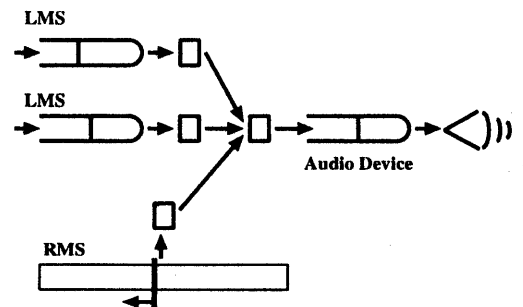


図5: 音声情報の提示

LMSとしては、教育システムなどで先生が教材について説明する声、会議システムなどで資料について説明する人の声など、人の声を用いることが多いであろう。

そこで、本実装でも LMS として音声情報を用いることとする。

Sun の audio device は read, write 用につづつしか open できないため、各ノードでは図 5 のように音声情報源、発信者ノードなどから来る複数の音声情報をつつにまとめて audio device に write する必要がある。

4.2.1 発信者ノードからの情報送付

送信者ノードでは talk ボタンが押されると定期的に“音声情報を audio device から read し他ノードに送出する procedure”が起動される。ただし、この procedure は図 6 のように必ずしも定期的に起動されるとは限らず、従って一度に audio device から read するデータの大きさも変動する。

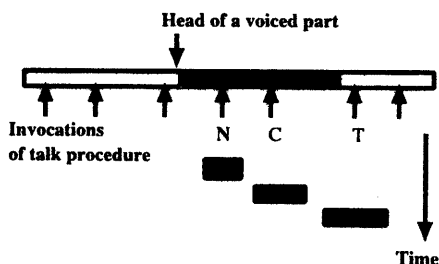


図 6: 発信者ノードからの送出フレーム

audio device から read したデータは有音、無音の判定を行ない、無音の場合、情報送出は行なわない。

有音の場合には図 6 のように read したデータが有音部の先頭 (図中の N) か、途中 (図中の C) か、末尾 (図中の T) か判別され、そのことを示す情報も送出フレームに付加する。また、有音部の先頭の場合、その時点で提示されているメディアフレーム番号から有音部の先頭が read された瞬間に提示されていたメディアフレーム番号を推定し、そのフレーム番号と、これまでに受信し発行された命令総数、及び、現在の状態 (一時停止中など) も送出フレームに付加する。

4.2.2 各情報受信ノードでの情報提示

各情報受信ノードでは発信者ノードから次々と送られてくる音声情報を発信者ごとにバッファに保持する。これらの音声情報は図 7 のように、各有音部単位 (図中の Voiced Part) でまとめられ、それらの有音部は受信した順序で保持される。

図 7 において、Frame Num は有音部の先頭に対応するフレーム番号を、Com Count, Com Mode はその有音部 (の先頭) が送信者ノードで read された時点で送信者ノードにおいてそれまでに発行された命令総数、ノードの状態をそれぞれ示している。これらの情報は送

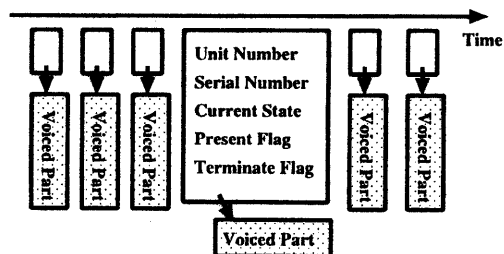


図 7: LMS 受信用 queue

信者ノードからの送出フレームの付加情報として取得できる。

Present Flag はその有音部が提示可能状態になったことを示すフラグである。このフラグは各受信ノードにおける再生点、送信者ノードにおいてこの有音部が read された時点における再生点に到達した際に ON になる。この判定は [Frame Num, Com Count, Com Mode] と各受信ノードでの再生点、命令発行総数、状態とを比較することで行なう。sync_with_all などの e_type のノードでは、有音部を受信した時点で受信ノードにおける再生点は既に送信者ノードでの read の時点での再生点を過ぎていたので、受信と同時に Present Flag も ON になる。

定期的に起動される音声提示 procedure は、Present Flag が ON になっている有音部を、先頭から順番に audio device に write していく。

audio device に write された有音部はバッファから解放される。

Terminate Flag はこの有音部に関する情報を受信している際に OFF、受信が完了した時点で ON になるフラグである。音声提示 procedure はこのフラグが OFF である有音部に関しては、Present Flag が ON の際にも、音声提示の連続性が保証できる程度の情報が受信されていないければこの有音部を audio device に write しない。

5 R&L 同期方式の評価

本節では、試作システムにおける R&L 同期方式の評価について述べる。

5.1 評価環境

Ethernet で接続された 3 台の UNIX workstation を用いて R&L 同期機構の評価を行なった。3 台のマシンの役割を以下に示す。

SS10 (Sparc station 10) RMS 情報源 (video, text), LMS 発信 (voice), 情報提示の 3 つの役割を兼ねる。

IPX (Sparc station IPX) 情報源 (audio), LMS 発信 (voice), 操作者, 情報提示の4つの役割を兼ねる。

SS1 (Sparc station 1) 情報提示のみを行なう。

バッファの大きさは video, text に対しては 20 フレーム分, audio に対しては 20+20 フレーム分とし, $T_{int} = 100ms$ とした。初期状態としては各メディアバッファが 0 フレーム目からのメディア情報で半分だけ満たされている状態とした。

また, 本測定の際には他のプロセスは走っていないかった。

5.2 評価結果

IPX において再生命令が出され, 各ノードで 0 フレーム目からの再生が行なわれた際に各ノードにおいて実際に各フレームが提示された時間を測定した。図 8 に sync_with_multi の場合の RMS の提示タイミングを示す。

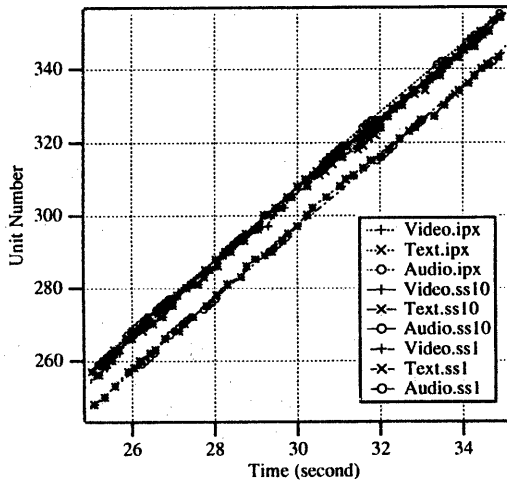


図 8: sync_with_multi の場合の情報提示タイミング

表 1 に LMS の提示タイミングを示す。ただし, 表 1 中の値は有音部の先頭が提示される際の RMS に対する LMS の遅れの平均, 及び, 最大値 (括弧中の値) を示している。

5.3 考察

図 8 から, 以下に示すことがわかる。

- 各ノードにおいて 3 つのメディアはほぼ同期して提示されている。

presentation nodes \ LMSs	SS10	IPX	SS1
a LMS generated at SS10	0 units (0 units)	10.95 units (15 units)	1.95 units (5 units)
a LMS generated at IPX	10.14 units (13 units)	0 units (0 units)	0.43 units (2 units)

表 1: sync_with_multi の場合の各ノードでの LMS の情報提示タイミングの遅れ

- d.type ノードでは e.type ノードに比べて CPP が遅れている。
- e.type ノードでは CPP はほぼ等しくなっている。

以上の結果から, 4.1.1 節, 4.1.3 節等で示した R&L 同期実現のための再生点調整はきちんと動作していることがわかる。

表 1 から以下に示すことがわかる。

- d.type ノードでは LMS の提示遅れは小さい (5 unit 以内)。
- e.type ノードでは LMS の提示遅れは大きい (平均 10 unit 程度)。

以上の結果から, LMS は R&L 同期のセマンティクスをほぼ満たしながら提示されていることがわかる。

6 むすび

本稿では, DMSIC において, R&L 同期を保つ機構を示し, 試作システムを用いた評価によりその有効性を示した。

試作システムでは LMS として音声のみを用いたが, 参加者をビデオカメラで撮ったものを LMS として用いるような場合にも各フレームを取り込んだ瞬間に提示されていた RMS の unit 番号をその LMS 中のフレームに付加することで同様にして R&L 同期を実現できる。

また, DMSIC は付加変動に柔軟に対応する機構^[2]も備えており, 様々な性能のワークステーション上で R&L 同期を保ちつつメディア情報を提示することができる。

参考文献

- [1] Anderson D.P., and Homsy G.: "A Continuous Media I/O Server and Its Synchronization Mechanism", IEEE COMPUTER, 24, 10, pp.51-57(1991).
- [2] 大野, 相田, 齊藤: "会話型操作を伴う分散マルチメディアシステムにおける同期機構", 情報処理学会マルチメディアと分散処理研究会, 66-21, 1994.