

グループコミュニケーションの中断を防止する サーバ機能の端末間移行方式

金友大 中島一彰

NEC サービスプラットフォーム研究所
E-mail: d-kanetomo@ce.jp.nec.com

情報共有ツールによるグループコミュニケーションシステムを容易に実現する方法として、コミュニケーション内容の一貫性の保証やコミュニケーションの参加者管理のための会議サーバを設置せず、参加者の端末にサーバ機能を持たせる方法がある。この方法でグループコミュニケーションシステムを実現した場合、サーバ機能を持つ端末の故障や離脱により、他の参加者の継続意図に反してコミュニケーションが終了してしまう問題がある。この問題を解決するためには、コミュニケーションを中断せずにサーバ機能を端末間で移動できることが必要である。本論文ではコミュニケーションと同時実行可能なサーバレプリケーション方式と短時間での切替方式を提案する。

Transition of server function between terminals for avoiding unintended termination of group communication

Dai Kanetomo and Kazuaki Nakajima

Service Platforms Research Laboratories, NEC Corporation
E-mail: d-kanetomo@ce.jp.nec.com

Allocating server function on a terminal attending the communication is an easy way to introduce group communication system with data sharing tools. However, in such cases, when the terminal with server function drops off from the communication, other participants must terminate the communication if they intend to continue it. For avoiding such termination, this paper proposes server replication and server switching method for realizing transition of server function between terminals during the communication.

1. はじめに

近年、企業内においては時間や金銭コストの削減を目的として、複数拠点間を通信で結び、会議に代表される複数人でのリアルタイムコミュニケーションを行なうことが一般的になっている。そのようなグループコミュニケーションにおいては、意思疎通を効率的に行なうため対面での会議と同様に、音声だけでなく PC を端末としてホワイトボードや資料を共有して議論することも行なわれている。

ホワイトボードや資料を共有するためのコミュニケーションシステムでは、コミュニケーション内容の一貫性の保証や参加者の管理のためにサーバ装置を必要とする場合がある。コミュニケーション内容の一貫性の保証は、複数の端末で同時にホワイトボードへの書き込みや資料のページ移動をした場合でも全ての端末でホワイトボードや資料の状態が同一になるために必要であり、全ての操作を一旦サーバ装置に集約し、処理した順に結果を全端末に反映させる

ことで実現できる。このようなコミュニケーションシステムをより容易に導入するために、参加者がクライアントとして利用する端末のいずれかにサーバ機能を持たせる方法もある。この方法では端末以外にサーバ用のマシンを別途必要としないという利点がある一方、サーバ機能を持つ端末の故障やコミュニケーションからの離脱が発生すると、他の参加者がコミュニケーションを継続したくても、コミュニケーションを一旦中断せざるを得ないという問題がある。この問題を解決するためには、コミュニケーションを中断せずに、サーバ機能を端末間で移行できるようにすることが必要である。本稿では、サーバ機能の端末間移行方式として、サーバ機能の他の端末へのレプリケーション方式と、サーバ機能間の切替方式を提案する。

2. サーバ機能移行方式の要件と課題

2.1 サーバ機能移行方式の要件

リアルタイムコミュニケーションのサーバ機能移行方式の要件は 2 つある。1 つはサーバ機

能のレプリケーションがコミュニケーションへ悪影響を与えないことである。コミュニケーションを実行しながらのレプリケーションにより、データの欠落や、応答が困難になるような大きな遅延が起こってはならない。もう1つはサーバ機能の切替において切替時間をなるべく短くすることである。切替時間が長くなれば一旦中断して再開するのと変わらないからである。

2.2 従来技術の課題

クライアント/サーバシステムにおいてホットスタンバイでのサーバ切替を可能にする方式の一つとして「プロセスレプリケーション」[1]がある。サーバからプロセスの処理情報を逐次、別のノード(マシン)に転送することで、サーバのプロセスイメージの複製をレプリカとして生成する。本方式はサーバからレプリカにクライアントからの処理情報を転送し、サーバで実行するだけでなく、レプリカで実行したことを確認した後にクライアントに応答を返すため、サーバ・レプリカ間の通信条件によっては、遅延が大きくなることが考えられる。また切替についても、一旦クライアントを停止させてサーバからレプリカへの接続の切替を行なうため、切替に要する時間が長くなる可能性がある。

別のサーバ機能移行方式として、Web サービスを対象とした「Active/Active Replication」[2]がある。本方式は同じサービスを提供する複数のサーバを特別なグループコミュニケーション方式で接続することで、サーバの状態をほぼリアルタイムに同期させておき、障害発生時にはサーバ群がクライアントと協調することでサーバ・クライアント間の接続をフェイルオーバーさせる。本方式を適用した場合、サービスを遅延させることなく切り替えることができるが、利用しているグループコミュニケーション方式[3]がマルチキャストを利用するため、通信中のデータの欠落が起こりうる。また全てのサーバが入力データを処理し、結果をクライアントに返すため、その分ネットワーク帯域を余計に消費するため、コミュニケーションに悪影響を与えることも懸念される。

3. 提案するサーバ機能移行方式

2章で挙げた要件を満たすサーバ機能移行方式について、本章で説明する。

3.1 前提となるサーバ機能の状態モデル

3.1.1 対象とするコミュニケーションシステ

ムの特徴

ホワイトボードや資料の共有システムといった提案方式が対象とするコミュニケーションシステムは、ストリームデータが送受信される映像通信や音声通話とは異なり、ホワイトボードの書き込みやスライド資料のページ遷移などの参加者の操作毎に、対応するメッセージがサーバへの入力として送信され、サーバでの処理後、コミュニケーションに参加している各端末にサーバの出力内容がメッセージとして送信される。サーバでは共有対象毎に状態が管理されており、各端末での操作毎に更新される。管理された共有対象の状態は、コミュニケーションに途中参加者があった場合に、クライアントとの同期を行なうことで、途中参加者がコミュニケーションをキャッチアップすることを可能にする。またコミュニケーション内容と同様に、コミュニケーション参加者もサーバで管理されており、コミュニケーション開始時には、参加の可否を判断して、意図しない参加者の存在を防止する。

3.1.2 サーバ機能の状態遷移のモデル化

3.1.1で述べたコミュニケーションシステムの状態とその遷移は次のようにモデル化できる。

- あるコミュニケーションシステムでの共有対象 t に対し、そのサーバ機能の状態 S^t およびクライアント機能の状態 C^t は初期状態 S_0^t および C_0^t からそれぞれ $S_0^t, S_1^t, S_2^t, \dots, C_0^t, C_1^t, C_2^t, \dots$ と遷移する。
- S_x^t から S_{x+1}^t への状態遷移は決定的であり、クライアント機能での操作により生成されるメッセージ I_x^t のサーバ機能への入力により発生する。
- S_x^t から S_{x+1}^t への状態遷移により、サーバ機能の出力メッセージ O_x^t が生成され、 O_x^t がクライアント機能で受信されると、 C_x^t から C_{x+1}^t への状態遷移が発生する。

3.2 設計方針

サーバ機能移行方式の設計は、3.1.2で示したシステムモデルと、以下の2つの方針に従って行なった。

1つは遠隔地の端末間でもサーバ機能の移行に支障がない方式とすることである。対象はコミュニケーションシステムであることから、通常参加者およびその端末は遠隔拠点に存在するためである。

もう1つは機能移行のために新たなノード(マシン)を必要としないことである。PCを端末と

し、それらの端末のみでホワイトボードや資料を共有するリアルタイムコミュニケーションシステムを構成する利点は、導入が容易であることである。サーバ機能移行方式においてその利点をなくさないためには、端末以外に新たなマシンが必要になることは避けなければならない。

3.3 移行方式の概要

サーバ機能移行方式は、サーバ機能のレプリケーションと、レプリケーション先のサーバ機能に対する切替方式から構成される。

3.3.1 サーバ機能のレプリケーション

コミュニケーション開始時に、本来のサーバ機能（以降、現用系と呼ぶ）が稼働する端末とは別の端末で、予備のサーバ機能（以降、予備系と呼ぶ）を起動させる。起動時の状態は両サーバ機能とも同一（ S_0^i ）の初期状態である。

サーバ機能を切り替えてコミュニケーションを継続するためには、両サーバ機能が切替時点で同一の状態であればならない。そのために現用系サーバ機能に対する入力メッセージ（ I_x^i ）を全て同一順序で予備系サーバ機能に対しても入力することで、状態を同一にする

予備系サーバ機能へメッセージを入力するために、クライアント機能と両サーバ機能間で入力メッセージをルーティングする。

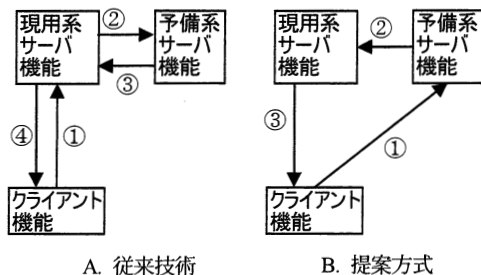


図1 メッセージルーティングの比較

従来技術[2]（図1のA）はクライアントが現用系サーバ機能にメッセージを送信し、現用系サーバ機能が予備系サーバ機能に対しメッセージを迂回させるのに対し、提案する方式では図1のBに示すように、クライアント機能は先に予備系サーバ機能にメッセージを送信し、予備系機能が現用系サーバ機能にメッセージを転送する方式を採用する。2つの方式を比較した場合、従来技術はクライアント機能部から現用系サーバ機能までにメッセージが通過するパス数

が短い、予備系サーバ機能部への状態反映を待ってクライアント機能に出力するために、クライアント機能で応答を受信するまでのメッセージのパス数では提案方式のほうが短い。

3.3.2 サーバ機能の切替方式

切替やその原因となる事象の発生時でもサーバ機能からの出力メッセージ（ O_x^i ）が全クライアント機能で受信されるようにし、クライアント機能間での状態ズレを防止する。

3.4 システム構成

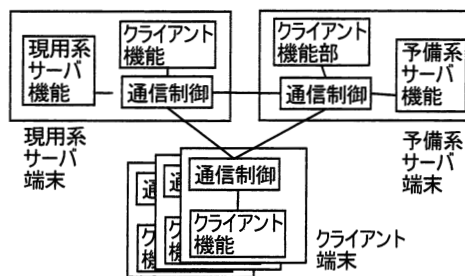


図2 システム構成

本節より詳細に方式を説明する。図2に端末種別とシステム構成を示す。端末は現用系サーバ機能がある端末（現用系サーバ端末）、予備系サーバ機能がある端末（予備系サーバ端末）とクライアント機能だけがある端末（クライアント端末）の3種類がある。図2ではクライアント端末にサーバ機能部は存在しないが、実際の運用としては参加者の全端末にサーバ機能部・クライアント機能部・通信制御部のセットがインストールされており、コミュニケーション開始時に端末の役割を決定し、それに従って必要な端末でのみサーバ機能部を実行することを想定している。

端末内の機能ブロックとして他に通信制御部が各端末に存在する。サーバ機能部・クライアント機能部それぞれの通信プロキシとして、メッセージ送信の制御やバッファリングなどを行なう。

3.5 処理の流れ

本節ではサーバ機能部のレプリケーションとサーバ切替のそれぞれについて処理の流れを説明する。

3.5.1 サーバ機能部のレプリケーション

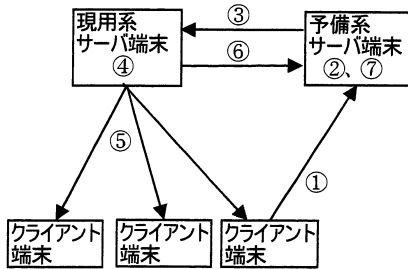


図3 コミュニケーション中の処理の流れ

コミュニケーション中のサーバ機能部のレプリケーション処理は、クライアント端末上で操作が行なわれた場合に以下のステップで実行される(図3を参照)。

1. 端末上での操作により生成されたサーバに対する入力メッセージが、予備系サーバ端末に送信される
2. 予備系サーバ端末は入力メッセージを一旦バッファする
3. バッファ後、現用系サーバ端末に対し入力メッセージを転送する
4. 現用系サーバ端末では受信した入力メッセージをサーバ機能が処理する
5. メッセージ処理により出力メッセージが生成されると、全クライアント端末に送信される
6. 送信が完了すると、現用系サーバ端末から予備系サーバ端末に対し、処理完了通知が送信される
7. 処理完了通知を受信した予備系サーバ端末では、バッファされていた元の入力メッセージを読み出して、サーバ機能部でメッセージ処理を実行する。このメッセージ処理に伴って発生する出力メッセージは破棄され、外部には影響しない。また処理を行なったメッセージはバッファより削除する

3.5.2 サーバ機能切替

次にサーバ機能切替時の処理について説明する。

現用系サーバ端末の障害発生を予備系サーバ端末が検知すると、予備系サーバ端末のサーバ機能がバッファしてある未処理の入力メッセージがある限り、入力順にメッセージ処理を行なう。ここで処理される入力メッセージは現用系サーバ機能では未処理であるため、その際に出

力メッセージは破棄せず全端末に送信する。現用系サーバ端末の障害発生タイミングが出力メッセージの送信中であった場合、出力メッセージを受信できない端末が発生するが、この場合でも予備系サーバ端末のバッファには該当のメッセージが残っているため、改めてメッセージ処理を行なうことにより、全端末で出力メッセージを受信して、クライアント機能の状態を整合することができる。

サーバ機能が切り替わっても、端末からのメッセージ送信は、元々予備系サーバ端末に送信されていたため、送信先の変更はない。また先に述べたバッファされていた入力メッセージの処理中でも、端末からの新たな入力メッセージは受信可能であり、バッファの処理順の最後に追加される。バッファされた入力メッセージの処理が完了した以降は、受信した入力メッセージを即時にサーバ機能部で処理し、出力メッセージを全端末に送信する。

4. 評価

本章では提案した方式が2章で述べた要件を満たすことを試算により検証する。また従来方式[1]との比較により、提案方式の有効性を明らかにする。

4.1 前提とするネットワーク構成

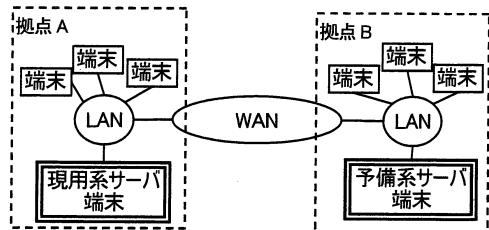


図4 2拠点を結ぶネットワーク構成

試算の前提として、図4に示すような遠隔の2拠点間でコミュニケーションを行なう場合のネットワーク構成を仮定する。拠点内の複数端末は各拠点のLANに接続され、LAN間をWANで結んでいる。またリスク分散の観点から両拠点の端末がそれぞれ現用系と予備系のサーバ機能を提供している。このときWAN、LANそれぞれにおける通信時間を以下のように仮定する。

表 1： 試算の前提となる通信遅延時間

| ネットワーク種別 | 通信遅延時間 |
|-----------|--------|
| WANにおける通信 | 150 |
| LANにおける通信 | 0.5 |

4.2 コミュニケーションへの影響

提案方式のコミュニケーションに対する影響を、クライアント端末がメッセージを送信してから応答である出力メッセージを受信するまでのエンドトゥエンドの遅延時間を試算することで評価する。表 1 の前提に加えて端末での処理時間を表 2 のように仮定して、クライアント端末のある拠点毎に試算した。

表 2： 試算の前提条件

| 処理内容 | 時間 (ms) |
|------------|---------|
| 端末でのパッパリング | 10 |
| メッセージ処理 | 150 |

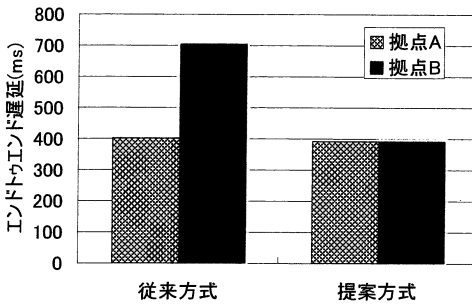


図 5 遠隔 2 拠点での遅延時間

試算結果を図 5 に示す。提案方式が拠点の位置によって遅延時間に差がないのに対し、従来方式では遅延の少ない拠点では提案方式と同程度だが、遅延の大きな拠点では約 1.8 倍の遅延時間となった。エンドトゥエンドの遅延時間のうち通信遅延時間がどのように構成されているかを表 3 に示す。従来方式では、拠点の違いにより、特に遅延の大きな WAN 回線の経由数に違いがあるため差が発生していることが分かる。この結果より、遅延の大きな WAN 回線を介して現用系サーバ端末と予備系サーバ端末が存在する場合には、提案方式が優位であることが分かる。

表 3 端末の位置による通信遅延の違い

| クライアント端末の位置 | 現用系サーバ端末の拠点 | 予備系サーバ端末の拠点 |
|-------------|-------------|-------------|
| 従来方式 | 2W + 6L | 4W + 8L |
| 提案方式 | 2W + 5L | 2W + 5L |

※ W=WANの通信遅延、L=LANの通信遅延

また IP 電話による音声通話においてはコミュ

ニケーションが成立するためにはエンドトゥエンドの遅延時間が 400ms 以下でなければならない [4]が、情報共有システムによるコミュニケーションも同様であると考え、提案方式における遅延のコミュニケーションに対する影響は許容範囲内であり、2 章で挙げたレプリケーションに関する要件を満たすと言える。

4.3 切替時間

次に図 4 の構成で現用系サーバ端末に障害が発生した場合にサーバ機能が切り替わり通信可能になるまでの時間を試算する。

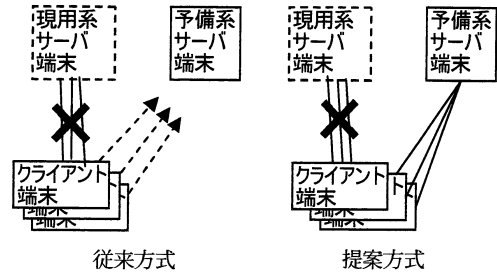


図 6 障害時の接続先変更の有無

図 6 に示すように、両方式ともに障害が発生した現用系サーバ端末と他の端末の接続は切断されるが、提案方式が元より予備系サーバ端末と接続しているのに対し、従来方式では、各端末は新たに予備系サーバ端末に接続してグループコミュニケーションを再開しなければならない。また新たに接続してグループコミュニケーションを行なう際には、意図しないクライアントの参加を防ぐためにサーバ機能がクライアントの接続認可処理を行なうことが必要である。このことを踏まえて表 4 のように条件を仮定して試算した結果を図 7 に示す。両方式を比較すると、提案方式では接続先変更が不要であるため、ほぼ 1/4 の時間で済んでいる。このことより、2 章で挙げたサーバ切替に関する要件を満たすと言える。

表 4 切替時間試算の前提条件

| 処理内容 | 時間 (ms) |
|----------|---------|
| 全端末への通知 | 30 |
| 端末での切替処理 | 50 |
| 接続認可処理 | 100 |

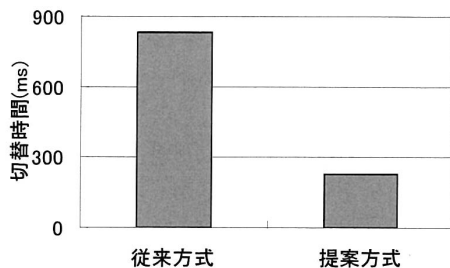


図 7 切替時間の比較

5. おわりに

本稿では、グループコミュニケーションシステムにおけるサーバ機能の端末間移行方式を提案した。本提案方式では、サーバ機能に対する入力メッセージのルーティングによるレプリケーションと、既存の接続を活かしたサーバ切替方式により、コミュニケーションに影響を与えることなくサーバ機能を端末間で移行させることが可能となった。今後、提案した方式の試作および試作システムを利用したの評価を行なう予定である。

参考文献

- 1) 小椋則樹、最所圭三、福田晃一、「クライアント/サーバシステムにおける信頼性向上の一手法」、ユニシス技報 Vol.16 No.4, 1997
- 2) I Engelmann, I Scott, I Leangsuksun, I He, "Transparent Symmetric Active/Active Replication for Service-Level High Availability," *ccgrid*, pp. 755-760, 2007.
- 3) Transis Project at Hebrew University of Jerusalem, Israel, <http://www.cs.huji.ac.il/labs/transis>
- 4) 総務省 IP ネットワーク技術に関する研究会報告書概要、
http://www.soumu.go.jp/s-news/2002/pdf/020222_3_01.pdf