

1. はじめに

近年、インターネット技術の進展と共に、企業内ネットワークもインターネット技術導入によるイントラネット化が進み、80年代から急速に進められたダウンサイジングとそれに伴うシステムの分散化は更に加速されたかに見える。しかし、企業内においては逆に Web サーバを窓口として業務サービスを行なうサーバシステムの集中化が起こりつつある。また、計算機性能の急速な向上により従来に比べより少数のサーバによりサービスが提供可能となっている。結果として、1台のサーバの異常がシステム全体に与える影響はより深刻となっている。このようなシステムの可用性を向上させるべく、複数サーバによる相互バックアップが可能なクラスタシステムが導入されてきた。今後ネットワークを通じたサービスの重要性が増すに従い、広域災害、テロといったサイトレベルの物理的障害に対する対策が不可欠になると考えられる。本研究では、サイトの機能喪失時にネットワーク上の他のサイトにおいて、サービスを継続可能とする広域クラスタシステムの構築方式に関する検討を行なっている。本稿では、広域クラスタシステム構築の課題の一つである系切替え方式に関し検討を行なった結果を報告する。

2. 広域クラスタシステムとは

本研究の対象としている広域システムは、図1に示すように、組織内において階層構造化されたモデルである。これは、最上位に組織中枢の計算機システムを持ち、次段に地理的に分離された拠点の計算機システムを持ち、最下層に各々の拠点の端末機器が存在し、最上位、及び、拠点の計算機システムは、そのシステム内において、可用性向上のために冗長化されている。最上位の計算機システムは、組織内の主要データベース、拠点に対する各種業務サービス、又は制御を集約的に提供しており、各拠点の計算機システムは、拠点内の端末機器に対し、業務サービスの中継や、ローカル業務サービスを提供しているものとする。広域クラスタシステムとは、このようなモデルにおいて、最上位の計算機システムがクラッシュし、機能喪失した際に、その上で動作している業務サービスを広域網上の異なる地点において処理を引き継ぐバックアップシステムを持つものである。最も単純な形態は、1:1のバックアップを可能とする2サイトの広域クラスタシステムである。

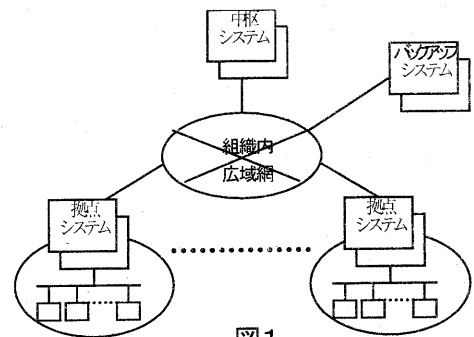


図1

従来、最上位に位置する計算機システムが保持するデータベースをバックアップする目的で、バックアップサイトにデータベースの複製を持たせるアプローチは行なわれていたが、業務サービスを含めたバックアップを行なうに至っていないのが現状であると思われる。

3. 系切替えの課題

広域クラスタシステム構築の課題として、大きくは、(a) 系の切替え方式の問題、(b) 系切替え時のデータ引継ぎ問題、の2点であるが、本稿では、(a)の系切替え方式の問題に焦点を絞って検討を行なった。系切替えにおける課題としては、(1) 異常検知の方法、(2) 頭脳分離問題の解決、(3) 切替えの自動化、の3点である。

(1) 異常検知の方法

中枢システムの異常を検知する方法として、冗長化ハートビートが考えられるが、ハートビート異常を検知した段階では、中枢システムの異常か、回線の異常かが判別できない。

(2) 頭脳分離問題の解決

中枢システムとバックアップシステム間の通信が途絶し異常と判断した際に、中枢システム自体に問題がない状況で系切替えを行なうと、データが非同期に更新されたり、制御対象に対する二重制御が発生する頭脳分離問題が発生する。従って、系切替えを発生させる場合には、中枢システムを確実に停止させる必要がある。

(3) 切替えの自動化

業務サービスの重要性が増すほど、可用性の向上、すなわち、サービス停止時間の短縮が必要である。異常の兆候を検知しても、人手を介して切替えを行なうのでは、時間もコストもかかるので、自動化が求められる。自動化を行なうためには、(1)、(2)を解決するルール設定が必要となる。

4. 自動系切替え方式

前節の課題解決を目的として、サービス受容サイトによる選挙による切替え要否判定に従って自動的に系を切替える方式を考案した。この方式は、(a) サービス、又は、制御を受ける拠点システムの過半数からの支持を得たサイト(中枢システム又はバックアップシステム)がサービスを継続する、(b) 選挙の実施は、バックアップシステムが中枢システム間との通信途絶と判断した場合、及び、拠点システムが中枢システムからのサービス、又は制御を受けられないと判断した場合に行なわれる、(c) 拠点システムの支持は、中枢システムを優先とし、中枢システムとの通信が途絶したと判断した場合のみに、バックアップシステムを支持する、の3点を原則としている。

本方式の特徴としては、(i) より多数の拠点に対しサービス継続が可能となる、(ii) 回線異常かサイトクラッシュかに関わらず、切替えの要否判断を可能とする、などであるが、残念ながら、頭脳分離問題を完全に解くことは出来ていない。これは、図2の(a)のように2分割された場合には、サービスを継続するシステムを判定できるが、(b)のように網が寸断され、中枢システム、又は、バックアップシステムが過半数の支持を集められなかった場合には、本方式では、双方共にサービスを停止してしまう点である。

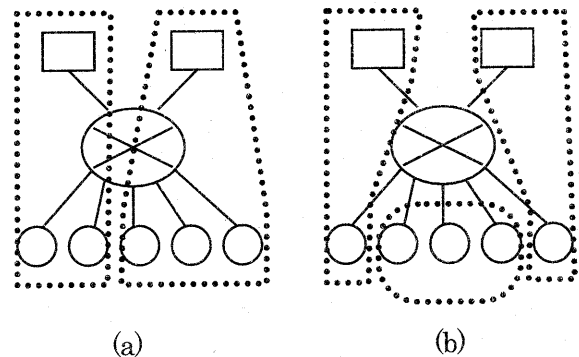


図2

従って、実装においては、極力頭脳分離状態を回避する方

策として、組織内の広域網が分断された場合に、衛星回線などの地上回線以外の通信路を用いて、中枢システムとバックアップシステム間で支持数の比較を行いサービス継続可能なシステムを決定する方式を考慮している。

5. まとめ

以上のように、広域クラスタシステムにおける自動系切替え方式の検討を行なった。頭脳分離問題の完全な解法は得られなかったが、サービス受容サイトによる選挙方式による切替え要否判定を原則として、実装段階において、衛星回線を用いて、支持数の比較を行なうことにより、頭脳分離問題を極力回避する方式を提案した。今後更に、組織内の計算機システムの集中管理と重要性が増すに従い、その可用性向上策が要求されると確信しており、このようなモデルの重要性が増すと考えている。

6. 今後の課題

今後、本方式の試作・評価を行なった後、実システムへの適用を検討する。また、1:1バックアップモデルから、N:1バックアップモデルへの拡張を検討する予定である。