

## 審査の結果の要旨

氏名 藤田 肇

インターネット上のサーバ性能向上や可用性の向上を目的に単一IPアドレスクラスタ実現手法が提案され実用化されてきた。しかし、既存手法では台数効果、高可用性、信頼性を同時に達成することが出来なかった。本論文では、単一IPアドレスクラスタ実現手法の一つであるブロードキャスト型方式を基に、新しい負荷分散機構および投機的SYNパケット受信機構を提案し、既存手法よりも高い台数効果、既存手法では達成できなかった高可用性、の2つを同時に満たすことができる単一IPアドレスクラスタを実現した。さらにクライアントサイドチェックポイント機構を提案し、システム全体の信頼性向上に貢献した。本論文は、以下の通り8章から構成されている。

第1章では、インターネットサーバへの要求事項の整理を行い、それに対する従来の解決手法という観点からブロードキャスト型単一IPアドレスクラスタサーバの問題点と、本論文で提案する手法の概要を紹介している。

第2章では、インターネット上においてクラスタサーバを構築する際に用いられる技術について紹介し、それらと単一IPアドレスクラスタサーバ方式との比較について議論した後に、既存の単一IPアドレスクラスタ実現法の詳細と、ブロードキャスト型クラスタの負荷分散性能、可用性、信頼性の問題について指摘している。

第3章では、提案手法を実際に動作するシステムとして実装するにあたり、既に広く普及しているインターネット環境に導入するという観点から実装手法に求められる要件を議論し、透過性のために新機構をサーバOSカーネル内のパケットフィルタとして実装すること、もしくはミドルウェアとして実装し、アプリケーションへの変更を最低限に抑えるという方針を示している。

第4章では、ブロードキャスト型単一IPアドレスクラスタサーバにおいて柔軟な負荷分散を実現するためのサーバOSカーネル向け機構を提案している。本章では、クライアントからのリクエストの情報を利用する度合いに応じて、リクエスト内容を考慮しないレイヤ4負荷分散と、考慮するレイヤ7負荷分散とに分け、それぞれに向けた設計方式を提案している。いずれの手法もTCP接続の確立時にのみスケジューラノードが関与することで、それ以後の通信は他のサーバ故障に影響されない。また提案手法を用いてwebサーバを構築し、SPECweb2005 Supportベンチマークによる性能比較実験を行っている。その結果、レイヤ4負荷分散において既存手法よりも総リクエスト処理数を13%向上、レイヤ7負荷分散ではダウンロード速度基準に満たないリクエスト数を48%削減している。

第5章では、サーバノード故障時のダウンタイムを削減するために必要となる投機的SYNパケット受信機構をサーバOSカーネル機構として提案している。投機的SYNパケット受信機構により、クライアントの新しいTCP接続要求に応えようとしているサーバノードが故障した時でも、代替ノードが即座にクライアントからのSYNパケットを受信し応答することにより、ダウンタイムを削減できるようになった。クラスタサーバの1ノードに故障を注入する評価実験を行い、従来手法では3~9秒のダウンタイムが観測されていたのに対し、提案手法を用いるとクライアントからのリクエストのほぼ全てが0.4秒以内に処理されることを示している。

第6章では、クライアントサイドチェックポイントを行うミドルウェア層として通信ライブラリ機構を提案している。この機構はサーバ状態をクライアント側のメモリ上に記録し、サーバ故障時にはTCP接続の復元とサーバ状態の復元を行うことによって実現される。ストリーム配信アプリケーションffserverを用いた評価で、本機構の適用による使用メモリ量の増加は元のプログラムのメモリ使用量を超えないことを示している。

第7章では、本論文で提案している単一IPアドレスクラスタ上でのリクエスト分散機構、分散システムにおける負荷分散アルゴリズム、TCP通信の性能や信頼性を向上させるための機構、サーバ向け耐故障機能に関して、既存の研究と比較し提案の新規性を明らかにしている。

第8章では、論文全体のまとめと、今後の研究課題を述べている。

以上、本論文では、ブロードキャスト型単一IPアドレスクラスタにおける柔軟な負荷分散手法を組み込める負荷分散機構、サーバ故障時に即座に代替サーバが対応可能とするための投機的SYNパケット受信機構を提案した。これにより、従来ブロードキャスト型単一IPアドレスクラスタにおいて実現できていなかった高い台数効果と高い可用性を実現することができた。さらに、ブロードキャスト型クラスタでもなお対処できないサーバノード故障時の問題に対処するためにクライアント側ミドルウェア層にチェックポイント機構を提案し信頼性向上を達成した。本論文で提案しているシステム構成法により、利用者やアプリケーションに意識させずにインターネット上のサーバシステムの性能が向上し、またソフトウェア資産のOS間移行コストを低減することができる。これら本研究の成果は、高性能計算機クラスタのシステムソフトウェアの発展に顕著な貢献をしたといえる。

よって本論文は博士（情報理工学）の学位請求論文として合格と認められる。