

「新しい情報社会を支える高性能情報処理技術」
平成 15 年度採択研究代表者

横田 治夫

(東京工業大学 教授)

「ディペンダブルで高性能な先進ストレージシステム」

1. 研究実施の概要

新しい情報化社会に求められるディペンダブルで高性能な先進ストレージシステムを構築するための基本的な技術の確立を目指している。そのため、並列分散ストレージシステムを対象に、その管理コストの削減を目的とする。さらに、情報社会におけるストレージシステムの新たな位置づけを目指して、ストレージシステムをより賢くする方法も考える。

これまでに、本研究のために準備した 160 台規模のシミュレータと 2 種類のプロトタイプシステムを用いて、管理コスト削減のためのストレージのアクセス負荷と容量バランスの両立、通常処理とデータ管理処理のプライオリティ管理、それらのための管理データ構造の同時実行制御等の機能を実現・評価し、当該分野のトップレベルの国際会議や論文誌等で発表を行った。また、チーム内の研究の進捗状況の確認とそれに対する海外のトップレベルの専門家からのコメントの収集および関連国際会議での情報収集等を目的に、昨年度に引き続いて平成 17 年 12 月に米国でワークショップを開催し、米国の最新のストレージ研究の内容を確認するとともに、当該分野における中心的な研究者との交流によって今後の関係を強固にした。

今後も、これまでの研究を引き続き行うと同時に、コンテンツの特徴を活かしたストレージ管理の手法等を検討し、提案した基本技術が商用システムに取り入れられることを目標に研究を進める。

2. 研究実施内容

[研究目的]

高度な情報化社会の多様な要求に対応できるディペンダブルで高性能な先進ストレージシステムを構築するための基本的な技術の確立を目指す。高いディペンダビリティと性能を実現するために並列分散ストレージシステムを対象とし、柔軟で効率的な並列分散ストレージシステムを実現する時に問題となる管理コストの削減を目的とする。さらに、ストレージシステムをさらに賢くすることによって、情報社会での新たな位置づけを持つ先進ストレージシステムの実現を目指す。

[研究方法・結論]

ディペンダブルで高性能な並列分散ストレージシステムを実現するため、各ストレージ

ジドライブをインテリジェント化してストレージクラスタを構成する。そのストレージクラスタ内で自律的にデータ移動を行うことで、障害回復を行うと同時にアクセス負荷やデータ容量のバランスを取る自律ディスクの研究を行っている。

平成 17 年度は、本研究のために準備した 160 台規模のシミュレータと 2 種類のプロトタイプシステムを用いて、管理コスト削減のための機能を実現し、評価を行ってその効果を示した。なお、シミュレータは主に管理機能とその拡張性の確認を目的とし、プロトタイプシステムは提案する自律ディスクがサイズ、電力的にフィージブルに実現可能であることを示すことを目的としている。それらの実験・評価の結果を当該分野のトップレベルの国際会議や国内会議等で口頭発表を行うと同時に、多くの論文投稿や特許申請を行った。

また、昨年度に引き続いて今年度も、研究進捗状況の確認とトップレベルの専門家からのコメントの収集を目的として、平成 17 年 12 月に米国カリフォルニア州サンフランシスコ市内において Advanced Storage System (ADSS) ワークショップを開催し、成功裏に終了した。以下がワークショップの様子である。

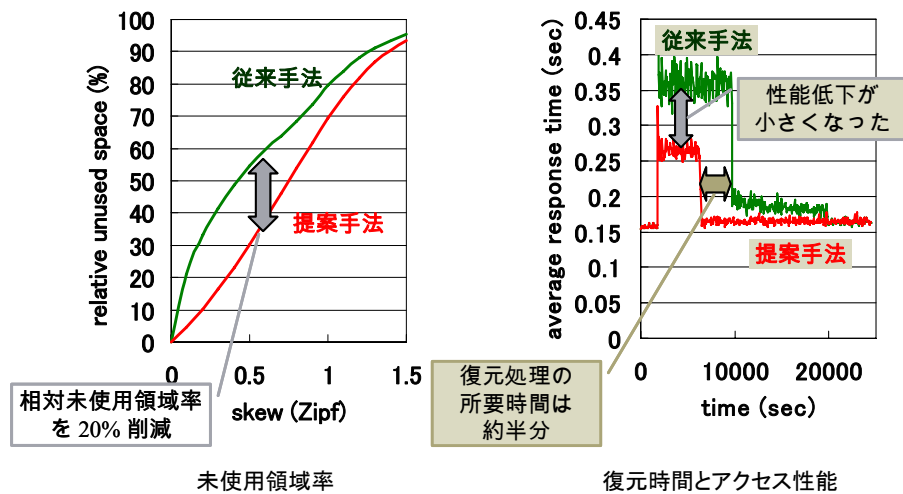


ADSS2005 ワークショップの様子

[研究実施の具体的内容]

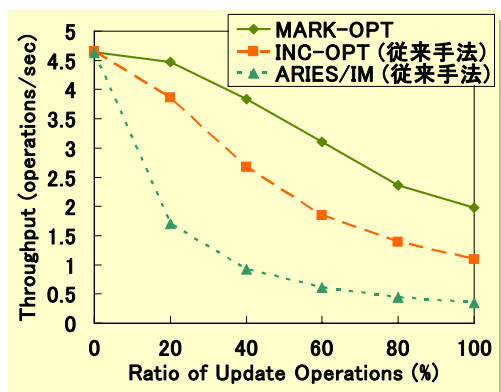
並列分散ストレージシステムの管理コスト削減を目指して、シミュレータとプロトタイプを用いて各ストレージデバイスのアクセス負荷と容量バランスの両立、通常処理とデータ管理処理のプライオリティ管理、それらのための管理データ構造の同時実行制御等の機能を実現し、評価を行った。また、モバイル環境での先進ストレージシステムの利用に関しても検討を行った。

アクセス負荷と容量バランスの両立に関しては、信頼性向上とサービス品質を保つために採用しているプライマリ・バックアップ方式を拡張して、プライマリデータでアクセス負荷バランスを取り、バックアップデータで容量バランスを取る Adaptive Overlapped Declustering 法を提案し、シミュレーションで未使用領域率と復元時間における効果を確認した。



また、バージョン管理を前提に格納されるデータのバージョン間のアクセス頻度に着目して、アクセス負荷と容量バランスの両立させる手法である COBALT の提案も行い、シミュレーションによってアクセス負荷のバランスを保ちながら容量のバランスが確保される効果を確認した。一方、通常サービス処理とデータ管理処理のプライオリティ管理に関しては、データ管理のためのデータ移動時にバックアップを利用して負荷を軽減させる手法の提案を行い、シミュレーションによってその効果を確認した。

さらに、それらのデータ管理のためには何らかの管理データ構造が必要となる。そのひとつである並列 Btree のアクセス効率化のための同時実行制御方法に関しても、衝突時のリスタートを削減する MARK-OPT 法を提案し、従来の手法と比較して性能が大幅に改善されることを示した（下図参照）。さらに、CREST シンポジウムにおいて、プロトタイプシステムを用いて、振動等でディスクアクセスにおけるエラーの頻度が上がった場合に、自律的に故障の発生を予測して、他のディスクにデータを移動させるデモを行った。



MARK-OPT 法と従来の手法の比較



プロトタイプを用いたデモの様子

この他、複数のディスクのアクセス頻度を考慮して電源管理を行い、システム全体と

しての電源消費を抑える方法、モバイル環境でのストレージ構成として、移動先の近傍に先行してストレージ間でデータを移動させておくことでデータアクセスの遅延を少なくする手法、ストレージに格納されるコンテンツに付加するメタデータを記述する XML の効率的な格納方法およびその XML を分散して格納して効率よくアクセスする方法、ストレージ中に格納された扱いの異なるコンテンツを統合させて提供する方法、分散格納されたコンテンツの間のワークフローを高信頼で管理する方法、等に関しても検討、実験、評価等を行い、それぞれにその効果を示した。

3. 研究実施体制

「先進ストレージ研究統轄・推進」グループ

①研究分担グループ長：横田 治夫（東工大、教授）

②研究項目：

- ・ ディペンダブルで高性能な先進ストレージシステムの研究の統括・推進
- ・ 先進ストレージシステムにおけるデータ管理機能の検討、実装、評価

「高度メディア蓄積・管理手法研究」グループ

①研究分担グループ長：植村 俊亮（奈良先端大、教授）

②研究項目：

- ・ ユビキタスストレージシステム実現に向けてのシミュレーションによる評価
- ・ 異なるアプローチによる大規模 XML データの分散処理
 - + ネイティブ XML データベースの分散化によるアプローチ
 - + XML データベースの関係データベースへの写像に基づくアプローチ

「コンテンツ対応メタデータ研究」グループ

①研究分担グループ長：上原 年博（NHK 放送技術研究所、主任研究員）

和泉 吉則（NHK 放送技術研究所、主任研究員）

②研究項目：

- ・ ストレージに格納されるメタデータの検討
- ・ コンテンツ対応データ管理の検討

「システムアーキテクチャ研究」グループ

①研究分担グループ長：太田 光彦（富士通㈱、部長付）

②研究項目：

- ・ 自律ディスクプロトタイプ試作機の研究開発、製作
- ・ 自律ディスクの実機実装及び検証実験

「ストレージ技術動向調査」グループ

①研究分担グループ長：藺田 浩二（日立製作所システム開発研究所、主任研究員）

②研究項目：最新のストレージシステムの技術動向に関する調査

4. 主な研究成果の発表（論文発表および特許出願）

（1）論文（原著論文）発表

- 趙剛, 小林亜樹, 酒井善則, 「ラフ集合論の縮約による多グループフィードバックを用いた対話型画像検索」, 映像情報メディア学会誌, Vol.59, No.6, pp.884-893 2005.6
- 吉原朋宏, 小林大, 田口亮, 上原年博, 横田治夫, 「並列 Btree 構造における SMO 発生時の処理性能を改善する並行性制御」, DBSJ Letters, Vol.4, No. 1, pp.21-24, 2005.6.
- 吉田誠, 小林隆志, 横田治夫, 「リサーチマイニング手法を用いた研究の発展経緯確認ツールの実装」, DBSJ Letters, Vol.4, No. 1, pp.81-84, 2005.6.
- 中野真那, 小林大, 渡邊明嗣, 上原年博, 田口亮, 横田治夫, 「バージョン管理用差分情報のアクセス頻度に着目したデータ配置」, DBSJ Letters, Vol.4, No. 1, pp.121-124, 2005.6.
- 小林大, 渡邊明嗣, 田口亮, 上原年博, 横田治夫, 「データ移動コストとキャッシュを考慮した複製へのアクセス分散制御」, DBSJ Letters, Vol.4, No. 1, pp.125-128, 2005.6.
- 中尾伸章, 天笠俊之, 的野晃整, 植村俊亮, “アクセス頻度を考慮した XML 文書分割方式の提案”, 日本データベース学会 Letters, Vol.4, No.1, pp.113-116, 2005.6.
- 加藤英之, 小林隆志, 横田治夫, 「Web サービスを用いたワークフローにおける負荷分散手法のシミュレーションによる評価」, DBSJ Letters, Vol.4, No. 2, pp.25-28, 2005.10.
- 渡邊明嗣, 横田治夫, 「値域分割に基づく分散ストレージにおける効率向上のための複製管理」, DBSJ Letters, Vol.4, No. 2, pp.69-72, 2005.10.
- Kenta YASUKAWA, Ken-ichi BABA, Katsunori YAMAOKA, “Tentative Accommodating and Congestion Confirming Strategy -A Novel Admission Control Strategy for Packet Switching Networks-“, IEICE Transactions on Communications, Vol.E89-B, No.2, pp.373-382 2006.2.

（2）特許出願

H17 年度出願件数：5 件（CREST 研究期間累積件数：13 件）