

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 20 日現在

機関番号：32613

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26730040

研究課題名(和文) 高速な規模伸縮が可能な低消費電力クラウド分散データベース管理システム

研究課題名(英文) Power Effective Distributed Database Management System Capable of High-speed Scaling

研究代表者

山口 実靖 (Yamaguchi, Saneyasu)

工学院大学・情報学部(情報工学部)・准教授

研究者番号：50439262

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：クラウド用分散DBMS(KVS)の性能伸縮性の向上を目標とし、基礎性能調査と、システムの動作の考察、考察に基づく動的性能伸縮性の改善と省電力化に取り組み、目標を達成した。具体的には、これらの目標に対して二つの手法を提案し、性能評価によりこれらの有効性を検証した。そして、性能評価によりこれらの提案した手法が上記目標に対して有効であることを確認した。

研究成果の概要(英文)：In this work, I have aimed to improve scaling performance of KVS (Key-Value Store) DBMS (Database Management System). For this purpose, I have proposed two methods. One is a method for improving cache behavior. The other is a method for improving behavior of I/O scheduler. My evaluation has demonstrated that these method have been able to improve scaling performance of KVS system.

研究分野：オペレーティングシステム

キーワード：ビッグデータ 分散KVS 省電力 データストレージ ネットワーク

1. 研究開始当初の背景

大規模 SNS(ソーシャル・ネットワーキング・サービス)の普及、スマートフォンなどの携帯デバイスの発達、クラウドコンピューティング基盤の整備などにより、多数の計算機を用いた大規模分散 DB(データベース)処理が行われるようになっていた。たとえば Twitter や Facebook では毎日膨大なデータが処理されていた。(この状況は今も変わってはいない。)このような超大規模 DB 処理は、従来の高機能な RDBMS(リレーショナルデータベース管理システム)だけでなく、スケーラブルな分散 DBMS(データベース管理システム)である KVS(Key-Value Store)を用いても行われるようになっていた。これらの大規模システムの消費電力は膨大であり、省電力化が非常に重要な課題となっていた。

ソーシャル・ネットワーキング・サービスなどの単位時間あたりの投稿数は時刻により大きく異なり、ネットワークサービスの負荷は時刻や曜日などとともに大きく変わっていた。つまり、負荷が低い時間帯は並列データベースシステムを構成する計算機の台数を削減して省電力化に努め、負荷が高い時間帯は計算機の台数を増やし高性能化に努めることが重要である状況であった。ただし、主要な分散 DBMS の一つである KVS は、計算機台数の増加や減少に非常に長い時間がかかるという課題を抱えていた。

2. 研究の目的

研究の目的は、高速に規模伸縮が可能である低消費電力クラウド分散データベース管理システムの構築である。

具体的には、

- ・大規模 KVS システムの基礎性能(データベース性能)の評価
- ・規模伸縮性(システムに計算機を追加する処理に要する時間と、システムから計算機を削除するのに要する時間)の評価
- ・規模伸縮時のシステムの振る舞いの調査(計算機の追加や削除中に発行されるデータ転送要求やストレージへの I/O 要求の調査)
- ・KVS システムの動的規模伸縮性の向上(計算機の追加や削除時間の削減)を目標とする。

3. 研究の方法

PC クラスタと、オープンソース OS の Linux を用いてクラウドコンピューティングシステムを構築した。また、KVS システムは著名なオープンソース KVS の一つである Cassandra を用いて構築した。

最初に、小規模のクラウドコンピューティングシステムを構築し、この簡易なクラウドコンピューティングシステム上に KVS システムを構築した。そして、その小規模 KVS システムの基礎性能(データベース性能)の評価、規模伸縮性の評価(システムに計算機を追加する処理に要する時間と、システムから計算機を削除するのに要する時間)の評価、規模伸縮時のシステムの振る舞いの調査(計算機

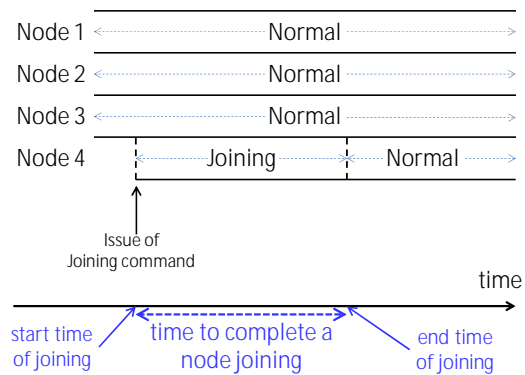


図1

の追加や削除中に発行されるデータ転送要求やストレージへの I/O 要求の調査)を行った。

考察が容易な小規模システムで振る舞いの調査を行い、性能劣化原因の考察(例えば CPU バウンドとなっているのか、I/O バウンドとなっているのかなどの調査)を行い、大規模システムへの考察につなげた。

計算機の追加時間は図 1 の様に定義した。すなわち、KVS システムに計算機の追加命令を発行すると、新規計算機は Joining 状態でシステムに参加する。そして、データ転送などを終えシステムの構成要素として稼働できるようになると Normal 状態に移行し、システムの一部として KVS サービスを提供するようになる。本研究では、図内の「追加コマンド発行時刻」から「状態が Joining から Normal に変わる時刻」までを計算機追加時間と定めた。

次に、大規模 PC クラスタシステム上に大規模クラウドコンピューティングシステムを構築し、その上に大規模 KVS システムを構築した。そして、同 KVS システムを用いて評価や考察を行った。

4. 研究成果

前章の調査により、KVS のインデックスファイルの様な頻繁にアクセスされる様なファイルでも大規模な I/O アクセスによりオペレーティングシステムのページキャッシュから破棄されてしまい、これが計算機の追加削除時の I/O 性能を大幅に低下させていることが分かった。

図 2 に、インデックスファイルとデータファイルのサイズと、観測されたインデックスファイルとデータファイルのアクセス数を示す。図より、インデックスファイルはデータファイルと比較してサイズが圧倒的に小さいにもかかわらず、少なくないアクセスを受けていることが分かる。すなわち、単位バイトあたりのアクセス数で比較すると、インデックスファイルは圧倒的に多くのアクセスを受けていることが分かり、このデータに対するアクセスをページキャッシュヒットとすることが重要であることが分かった。

これを受けて、図 3 の様にインデックスファイルへのアクセスを行い、インデックスファイルを積極的にページキャッシュ内に格

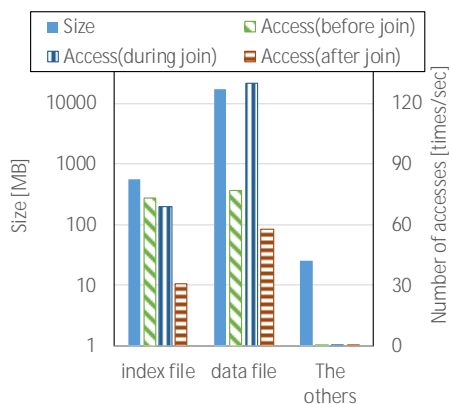


図2

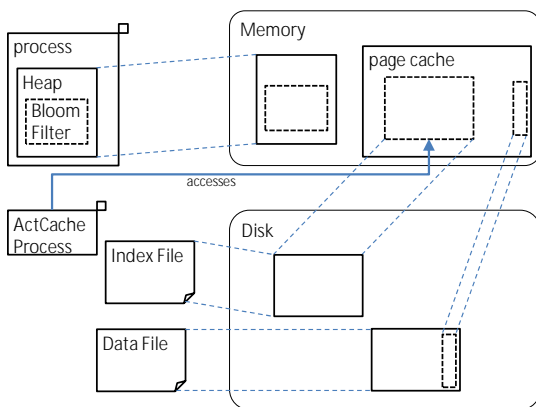


図3

納させる手法を提案した。本研究では、この手法を Act Cache あるいは AC と呼んだ。

次に、KVS システムに計算機を追加中には大規模なシーケンシャルアクセス要求が実行されるため、I/O スケジューラをこれに最適化する手法を提案した。本研究では、この手法を NLL と呼んだ。

最後に、提案した手法の有効性を検証するために性能評価実験を行った。具体的には、通常の Cassandra KVS システムと、Act Cache を用いた KVS システムと、NLL 手法を用いた KVS システムと、両手法を用いた KVS システムにて、稼働中の KVS システムに計算機を追加することに要する時間を評価した。評価結果の例を図4に示す。図より、本研究において開発された Act Cache 手法や、NLL 手法を用いることにより、あるいはこれらの両手法を併用することにより、稼働中の KVS システムの規模伸縮性の改善(計算機の追加などに要する時間を短縮)することが可能であることが分かった。

5. 主な発表論文等

[雑誌論文](計 3 件)

Saneyasu Yamaguchi, Yuki Morimitsu, "Improving Dynamic Scaling Performance of Cassandra", IEICE TRANSACTIONS on Information and Systems (2017)

杉本 洋輝, 山口 実靖, "二重キャッシュ環境における負の参照の時間的局所性を考慮したキャッシュ管理手法", 情報処理学

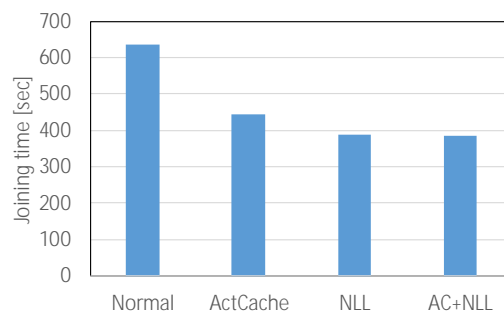


図4

会論文誌コンシューマ・デバイス&システム (CDS) (2015)

Tomoki Kozu, Yuria Akiyama, Saneyasu Yamaguchi, "Improving RTT Fairness on CUBIC TCP", International Journal of Networking and Computing, Vol 4, No. 2, pp.291-306 (2014)

[学会発表](計 15 件)

Eita Fujishima, Kenji Nakashima, Saneyasu Yamaguchi, "Performance improvement of I/O intensive OLAP with dynamic control of file storing location", ACM IMCOM 2017: International Conference on Ubiquitous Information Management and Communication (2017)

Kenji Nakashima, Eita Fujishima, Saneyasu Yamaguchi, "File Placing Control for Improving the I/O Performance of Hadoop in Virtualized Environment", 4th International Workshop on Computer Systems and Architectures (CSA'16) (2016)

Saneyasu Yamaguchi, Eita Fujishima, "Optimized VM memory allocation based on monitored cache hit ratio", In Proceedings of the 4th Workshop on Distributed Cloud Computing (DCC '16). (2016)

中島健司, 藤島永太, 山口実靖, "ホスト OS ファイルシステムにおける VM イメージファイルの非連続配置による仮想化環境における Hadoop I/O 性能の向上", 研究報告システムソフトウェアとオペレーティング・システム(OS) (2016)

Masaki Sakamoto, Saneyasu Yamaguchi, "Dynamic Memory Resource Management in Virtual Machines with Different Applications", 2016 International Symposium of Information and Internet Technology (SYMINTech) (2016)

Eita Fujishima, Saneyasu Yamaguchi, "Dynamic File Placing Control for Improving the I/O Performance in the Reduce Phase of Hadoop", the Tenth International Conference on Ubiquitous Information Management and Communication (IMCOM2016), 8-2 (2016)

Shohei Miyokawa, Taiki Tokuda, Saneyasu Yamaguchi, "Elasticity

Improvement of Cassandra", the Tenth International Conference on Ubiquitous Information Management and Communication (IMCOM2016), S2-3 (2016)

Hiroki Sugimoto, Kenichi Kourai, Saneyasu Yamaguchi, "Host OS Page Cache Hit Ratio Improvement based on Guest OS Page Drop", The 17th International Conference on Information Integration and Web-based Applications & Services (2015)

Masaki Sakamoto and Saneyasu Yamaguchi, "Dynamic Memory Allocation in Virtual Machines based on Cache Hit Ratio", 6th International Workshop on Advances in Networking and Computing (WANC'15), pp. 613-615, Poster (2015)

Shunsuke Yagai and Saneyasu Yamaguchi, "Reducing Storage Power Consumption with File Layout Optimization in Virtualized Environment", 6th International Workshop on Advances in Networking and Computing (WANC'15), pp. 590-594 (2015)

Eita Fujishima and Saneyasu Yamaguchi, "Improving the I/O Performance in the Reduce Phase of Hadoop", The Third International Symposium on Computing and Networking (CANDAR'15), pp. 82-88 (2015)

Shohei Miyokawa, Taiki Tokuda and Saneyasu Yamaguchi, "Active Caching for KVS Dynamic Scaling", The Third International Symposium on Computing and Networking (CANDAR'15), pp. 161-166 (2015)

Saneyasu Yamaguchi and Shunsuke Yagai, "Power Effective File Layout with Application Support in Virtualized Environment", 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE2015), pp. 425-430 (2015)

藤島永太・山口実靖, "ファイル格納位置制御による Hadoop MapReduce ジョブの性能の向上", FIT2015 第14回情報科学技術フォーラム, RC-003 (2015)

Shunsuke Yagai and Saneyasu Yamaguchi, "Energy Efficient Storage Management Cooperated with Data Intensive Applications in Virtual Machines", The 1st International Workshop on Future Technologies for Smart Information Systems (FTSIS 2014) in conjunction with the 33rd IEEE Symposium on Reliable Distributed Systems (SRDS 2014) (2014)

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

山口 実靖 (YAMAGUCHI, Saneyasu)
工学院大学・情報学部 (情報工学部)・准教授

研究者番号 : 5 0 4 3 9 2 6 2

(2) 研究分担者

(3) 連携研究者

(4) 研究協力者