

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 6月16日現在

機関番号：32613

研究種目：若手研究(B)

研究期間：2010～2012

課題番号：22700039

研究課題名（和文）クラウド仮想計算機に適した省エネルギーI/Oソフトウェアの開発

研究課題名（英文）Energy-Saving I/O software for Cloud Computing

研究代表者

山口 実靖（YAMAGUCHI SANEYASU）

工学院大学・工学部・准教授

研究者番号：50439262

研究成果の概要（和文）：

本研究では、既存の仮想化システムのI/Oスケジューラの性能の評価を行い既存のI/Oスケジューラが仮想化環境にて十分な性能を発揮できないことを確認した。つづいて、ホストOSとゲストOSを改変し、スケジューリング結果の保持、動作の解析が可能となる解析システムを構築した。そして、当該解析システムを用いて動作解析を行い、仮想化環境にて十分な性能を発揮できない理由が巨大なイメージファイル間の長距離シークにあることを明らかにし、この問題を解決する手法として飢餓状態判定時間の拡大など提案した。提案手法を、オープンソースのOS、仮想化システムに実装し、複数の仮想化システムで評価したところ、I/O性能の向上、統合サーバ数の向上、消費電力の低減などが確認できた。

研究成果の概要（英文）：

In this work, we have evaluated performance of the existing I/O scheduler in virtualized environment. Our evaluation has demonstrated that the existing I/O scheduler have not been able to provide enough performance in virtualized environment. In order to analyze the behavior of the I/O schedulers, we have developed I/O scheduler analyzing system. Analyses with the system have clearly shown the reason of their insufficient performance. After the analyses, we have proposed a method for improving I/O performance in virtualized environment. Our performance evaluation has demonstrated that the proposed method has been able to improve performance.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合計 |
|------|-----------|---------|-----------|
| 22年度 | 1,500,000 | 450,000 | 1,950,000 |
| 23年度 | 900,000 | 270,000 | 1,170,000 |
| 24年度 | 600,000 | 180,000 | 780,000 |
| 年度 | | | |
| 年度 | | | |
| 総計 | 3,000,000 | 900,000 | 3,900,000 |

研究分野：総合領域

科研費の分科・細目：情報学，ソフトウェア

キーワード：I/Oスケジューラ，仮想化，オペレーティングシステム，クラウド

1. 研究開始当初の背景

企業などには多数の常時稼働サーバ計算機が存在し、その膨大な消費電力や設置ス

ペースが問題になっていた。この解決策として、サーバを仮想化し複数のサーバを1台の物理計算機に統合する「サーバ仮想化」や、ネット

ワーク上の仮想計算機クラスタから必要な量だけ計算資源を借りて使用する「クラウドコンピューティング」があり、注目を集めていた。サーバ仮想化では、物理的に個別の計算機で動作していたプログラムとOSを仮想計算機ソフトの上で動作させ、実稼働の物理計算機の数を減少させる。

サーバ仮想化による省電力をさらに発展させるには、物理計算機あたりの統合仮想サーバ数を増加させる必要があるが、仮想計算機にて得られる性能は非常に低く、一台の物理計算機に多くの仮想サーバを統合することは現実的でなかった。特に、クラウドコンピューティングの主たるアプリケーションの一つであるI/Oインテンシブアプリケーションを実行するには、得られるI/O性能が十分ではなかった。

計算機のI/O処理の最適化、I/O性能の向上はI/Oスケジューラによって行われ、I/O性能はI/Oスケジューラの振る舞いに大きな影響をうける。しかし、当時のOSには仮想化環境を考慮したI/Oスケジューラが存在せず、仮想化環境では全く意味のない性能向上スケジューリングや最適化が行われていた。これが理由で仮想計算機にて得られるI/O性能が低くなっていた。

2. 研究の目的

仮想化環境に適したI/Oスケジューラの開発を行い、仮想化環境におけるI/O性能の向上、集約サーバ数の向上と消費電力の低減を最終的な目的とした。

そして、このこれらの大きな目的を達成するために、以下にあげる小さな目標の達成をめざし研究を進めた。小さな目的とは、仮想計算機におけるI/Oスケジューラの性能の調査、動作の解析、性能劣化原因の特定、仮想計算機を考慮したI/Oスケジューラの開発、それによる仮想サーバのI/O性能の向上、被統合サーバ数の向上、消費電力の低減である。

3. 研究の方法

オープンソースのOSであるLinuxを主たるターゲットとして、これに対応したI/Oスケジューラを開発することにより研究を行った。同様に仮想化システムとしては、オープンソースの仮想化システムであるXenやKVMを用いて開発を進めた。

詳細は以下の通りである。まず、LinuxとPCを用いてPCクラスタを構築し、その上に仮想計算機クラスタを構築し、これをクラウドコンピューティングプラットフォームとした。そしてクラウド基盤上で動作する仮想計算機の上でI/Oインテンシブアプリケーションを動作させ、I/O性能の調査、OS内部のI/Oスケジューラの動作解

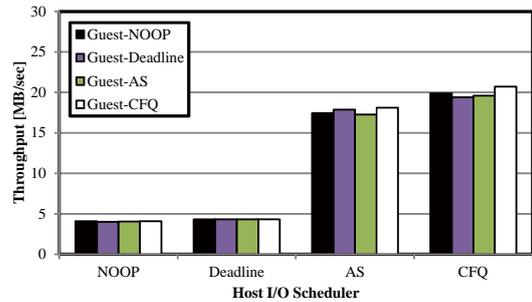


図1 仮想化環境におけるI/OスケジューラのI/O性能 (ベアメタル型仮想化システム, シーケンシャルリード)

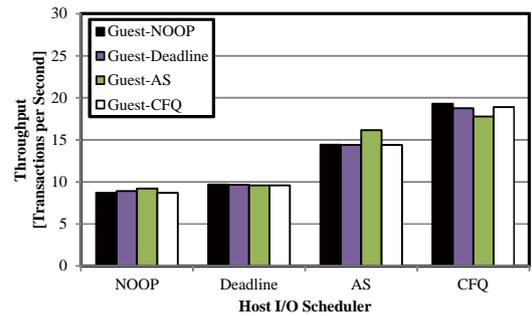


図2 仮想化環境におけるI/OスケジューラのI/O性能 (ベアメタル型仮想化システム, FFSSB)

析を行った (オープンソースOSであるため、ソースコードの変更を行った)。そして、動作解析を元に性能劣化原因を特定し、改善手法の考察、提案、実装と評価を行った。

4. 研究成果

主な研究成果を以下に示す。

(1) 仮想化環境における性能の評価

まず、既存のI/Oスケジューラが仮想化環境でどのような性能を示すかを調べるために、性能評価実験を行った。性能評価は、マイクロベンチマークとアプリケーションベンチマークで行い、マイクロベンチマークとしてはシーケンシャルリードを、アプリケーションベンチマークとしてはファイルシステムベンチマークFFSSBを用いた。

ベアメタル型仮想化システムを用いたときのマイクロベンチマークのI/O性能の測定結果を図1に、アプリケーションベンチマークのI/O性能の測定結果を図2に示す。ホスト型仮想化システムを用いたときのマイクロベンチマークのI/O性能の測定結果を図3に、アプリケーションベンチマークのI/O性能の測定結果を図4に示す。縦軸がI/O性能を表し、横軸がホストOSのI/Oスケジューラ、各線がゲストOSのI

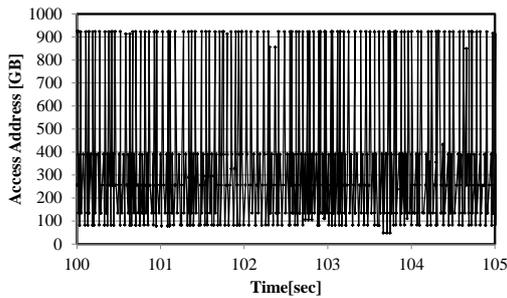


図 3 発行 I/O 要求 (HostOS--NOOP, GuestOS--NOOP) (ベアメタル型仮想化システム, FFSB)

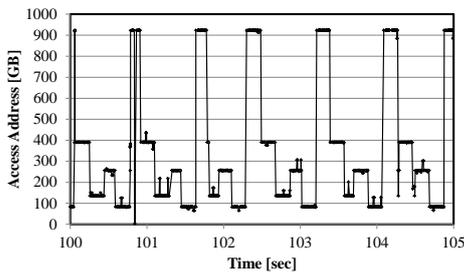


図 4 発行 I/O 要求 (HostOS--CFQ, GuestOS--NOOP) (ベアメタル型仮想化システム, FFSB)

ノスケジューラを表している。

(2) 動作解析

シーケンシャルリードおよび FFSB をゲスト OS 上で実行した際に発行された I/O 要求をホスト OS 上でモニタリングした。

まず、ベアメタル型仮想化システムを用いたときの発行 I/O について考察する。FFSB 実行時に発行された I/O の解析結果を図 3, 図 4 に示す。図 3 はホスト OS とゲスト OS の I/O スケジューラが NOOP, NOOP, 図 4 は CFQ, NOOP における発行 I/O を表している。前者が前述の評価にて性能が低かった組み合わせであり、後者が性能が高かった組み合わせである。図より、I/O 性能の低い図 3 では、アクセス対象の VM イメージの変更が頻繁に発生し、ストレージは長距離のシークを頻繁に行っていることがわかる。これに対して I/O 性能の高い図 4 では、アクセス対象 VM の変更が少なく、I/O 性能が高くなっていることが確認できる。このことから、VM イメージファイル間の長距離シークが性能劣化の主たる原因であることが分かった。

同様に、ベアメタル型仮想化システム上でシーケンシャルリードを行った際の発行 I

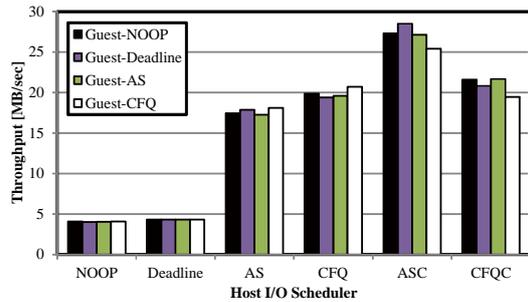


図 5 I/O 提案手法 I/O 性能 (ベアメタル型仮想化システム, シーケンシャルリード)

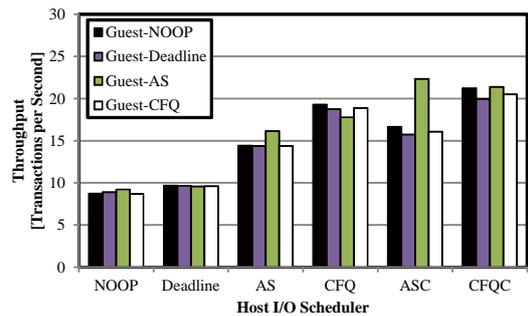


図 6 I/O 提案手法 I/O 性能 (ベアメタル型仮想化システム, FFSB)

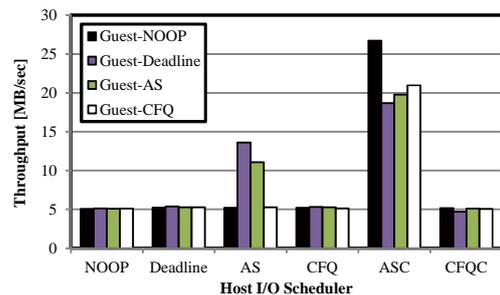


図 7 I/O 提案手法 I/O 性能 (ホスト型仮想化システム, シーケンシャルリード)

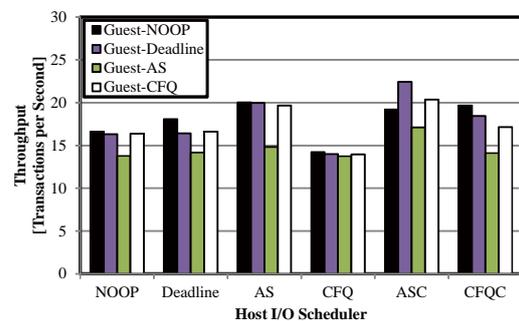


図 8 I/O 提案手法 I/O 性能 (ホスト型仮想化システム, FFSB)

ノの解析, ホスト型仮想化システム上でシーケンシャルリードと FFSB を行った際の発行 I/O の解析を行った。これらの解析

結果からも同様に、VMイメージファイル間の長距離シークが主たる性能劣化原因であることが分かった。

(3) 改善手法

アクセス対象VMの変更頻度を低減させるためにホストOSのI/OスケジューラのASとCFQに対して変更を行い、I/O性能を向上させる手法を提案した。具体的には、ASの近隣I/O発行の予測時間の限界値、飢餓状態判定の閾値、連続処理の許容時間などを拡大し、CFQのタイムスライス、飢餓状態判定の閾値などを拡大した。改善したASをASCと、改善したCFQをCFQCと呼ぶ。

これらの改善手法の性能を評価して、図5から図8を得た。図より、ホスト型仮想化システムにてシーケンシャルリードを行った場合のCFQCを除き、提案手法の適用によりI/O性能が大きく向上することが確認された。ホスト型仮想化システムにてシーケンシャルリードを行ったときのCFQCのみ、性能向上が見られなかったが、性能の低下はなく、提案手法はホスト型仮想化システムのI/O性能向上に関して有効であることが確認された。

5. 主な発表論文等

[学会発表] (計 23 件)

- ① 新居健一、山口実靖，“仮想化環境におけるI/Oスケジューラの動作と性能に関する考察”，インターネットコンファレンス2010，pp. 33-40 (2010)
- ② 越智俊介，山口実靖，“高集約サーバ統合環境におけるWebアプリケーションの性能に関する考察”，研究報告 - マルチメディア通信と分散処理 (DPS) Vol. 2011-DPS-146, No. 50 (2011)
- ③ Yosuke Kuno, Kenichi Nii and Saneyasu Yamaguchi, "A Study on Performance of Processes in Migrating Virtual Machines," The Fourth International Workshop on Ad Hoc, Sensor and P2P Networks (AHSP2011) (2011)
- ④ Yusuke Nagasako and Saneyasu Yamaguchi, "A Server Cache Size Aware Cache Replacement Algorithm for Block Level Network Storage," The Fourth International Workshop on Ad Hoc, Sensor and P2P Networks (AHSP2011) (2011)
- ⑤ Yosuke Kuno, Kenichi Nii and Saneyasu Yamaguchi, "A Study on Performance of Processes in Migrating Virtual Machines," The Fourth International Workshop on Ad Hoc, Sensor and P2P Networks (AHSP2011) (2011)
- ⑥ Yusuke Nagasako and Saneyasu Yamaguchi, "A Server Cache Size Aware Cache Replacement Algorithm for Block Level Network Storage," The Fourth International Workshop on Ad Hoc, Sensor and P2P Networks (AHSP2011) (2011)
- ⑦ 山田 将也，山口 実靖，“仮想計算機環境におけるVMイメージファイルの配置に関する一考察”，マルチメディア，分散，協調とモバイル(DICOMO2011)シンポジウム，6H-4，pp. 1234 - 124 (2011)
- ⑧ 久野 陽介，新居 健一，山口実靖，“ライブマイグレーション中の仮想計算機上のプロセスの性能に関する考察”，マルチメディア，分散，協調とモバイル(DICOMO2011)シンポジウム，8H-3，pp. 1611 - 1620 (2011)
- ⑨ 山田将也，山口実靖，“仮想計算機環境における二重ファイルシステム構造を考慮した仮想HDDイメージファイルの配置”，第4回Webとデータベースに関するフォーラム(WebDB Forum 2011) (2011)
- ⑩ 新居健一，山口実靖，“I/Oスケジューラの改善による仮想計算機環境におけるI/O性能の向上”，第4回Webとデータベースに関するフォーラム(WebDB Forum 2011) (2011)
- ⑪ 渡邊有貴，山口実靖，“仮想化環境におけるVMメモリ割当量とキャッシュ利用に関する一考察”，情報処理学会第153回データベースシステム研究発表会(2011)
- ⑫ 竹内 洗祐，長廻 雄介，山口 実靖，“多段キャッシュ環境の下位キャッシュにおける参照の時間的局所性の解析”，第150回マルチメディア通信と分散処理研究会(31) (2012)
- ⑬ 渡邊 有貴，山口 実靖，“仮想計算機キャッシュと仮想記憶を考慮した仮想計算機メモリ割当量に関する一考察”，第150回マルチメディア通信と分散処理研究会(39) (2012)
- ⑭ Ryo Oura, Saneyasu Yamaguchi, "Fairness Comparisons Among Modern TCP Implementations," The 6th International Workshop on Telecommunication Networking, Applications and Systems in conjunction with AINA2012 (2012)
- ⑮ Kyosuke Nagata, Saneyasu Yamaguchi, "An Android Application Launch Analyzing System," 8th ICCM: 2012 International Conference on Computing Technology and Information Management (2012)

- ⑩ Masaya Yamada, Yuki Watanabe, Saneyasu Yamaguchi, "An Integrated I/O Analyzing System for Virtualized Environment," 8th ICCM: 2012 International Conference on Computing Technology and Information Management (2012)
- ⑪ 竹内 洗祐, 長廻 雄介, 山口 実靖, "複数サーバ接続単一ネットワークストレージ環境での下位キャッシュへの参照の局所性の解析", マルチメディア、分散、協調とモバイル DICOM2012 シンポジウム, pp. 418 - 424 (2012)
- ⑫ 竹内洗祐・山口実靖, "仮想化環境下におけるホスト OS キャッシュの参照の時間的局所性の解析", FIT2012 第 11 回情報科学技術フォーラム, B-010 (2012)
- ⑬ Masaya Yamada, Saneyasu Yamaguchi, "Filesystem Layout Reorganization in Virtualized Environment," The 9th IEEE International Conference on Autonomic and Trusted Computing (IEEE ATC 2012), ATC4-2 (2012)
- ⑭ Nagata Kyosuke, Saneyasu Yamaguchi, Hisato Ogawa, "A Power Saving Method with Consideration of Performance in Android Terminals," The 9th IEEE International Conference on Autonomic and Trusted Computing (IEEE ATC 2012), ATC7-3 (2012)
- 21 竹内洗祐 山口実靖, "複数サーバ接続ネットワークストレージ環境での参照の局所性の解析", 第 24 回 コンピュータシステム・シンポジウム (ComSys 2012) (2012)
- 22 竹内洗祐, 日名川幸矢, 山口実靖, "仮想化環境下における下位キャッシュの参照の時間的局所性の解析", 第 156 回データベースシステム研究発表会 (2012)
- 23 日名川幸矢, 竹内洗祐, 山口実靖, "仮想化環境におけるキャッシュヒット率を考慮した VM メモリ割り当て", 第 154 回 DPS・第 60 回 CSEC 合同研究発表会 (2013)

6. 研究組織

(1) 研究代表者

山口 実靖 (YAMAGUCHI SANEYASU)

工学院大学・工学部・准教授

研究者番号：50439262