

仮想化環境における I/O スケジューラの動作解析

新居 健一[†] 山口 実靖[†]

[†]工学院大学大学院電気・電子工学専攻

1. はじめに

多くの企業で計算機の消費電力や設置スペースの増大が問題視され、仮想化によるサーバ統合などが重要視されている。しかし、既存の OS の I/O スケジューラは仮想化環境に対して最適化されていないため、仮想化環境における I/O 性能が高くないという問題がある。

仮想化環境における I/O スケジューラの性能に関する研究としては Boutcher らの性能評価[1]がある。彼らは、ホスト OS の I/O スケジューラとしては単純に I/O 要求を到着順に処理する NOOP が優れているとしており、ゲスト OS の I/O スケジューラとして適するものはワークロードに依存するとしている。しかし、必ずしもこれが汎用的な結論であるとは言い切れず、さらに多くの評価、考察を行うことが好ましいと考えることができる。

そこで我々は、1 台の物理計算機上で 6 台の VM(仮想計算機)を稼働させ、単一の仮想化システムを用いて I/O スケジューラごとの I/O 性能を評価し、I/O 性能を向上させる手法を提案した[2]。

本稿では、代表的な仮想化システムである VMware Server と Xen を用いて、Linux に標準で搭載されている 4 つの I/O スケジューラが仮想化環境において I/O 性能に対してどのような影響を与えるか仮想化システムごとに考察する。そしてその動作解析を行い、動作結果をもとに仮想化環境に適した I/O スケジューラの改善手法を提案する。

2. I/O スケジューラ

I/O スケジューラは、アプリケーション群から発行された I/O 要求群を適切な順に並び替え、高性能、高公平性などを実現するソフトウェアである。

Linux には、NOOP, Deadline, AS, CFQ の 4 つの I/O スケジューラが存在し、順に、単純に I/O を到着順に処理する、待ち時間の長い I/O を生じさせないように処理する、近隣の I/O が到着することを期待し I/O の発行を遅延させる、公平性を重視し処理する、などの特徴を持つ。

3. 仮想化環境における I/O スケジューラの性能

一台の物理計算機上に、5 台の VM (仮想計算機)を起動し、各 VM 上でベンチマークソフト FFSB(Flexible File System Benchmark)を実行し、I/O 性能の測定を行った。その際ホスト OS と、ゲスト OS の I/O スケジューラを変更して、それぞれの性能を比較した。仮想化システムとして Xen を用いた測定結果を図 1、VMware Server を用いたものを図 2 に示す。横軸がホスト OS の I/O スケジューラを表し、グラフの線がゲスト OS の I/O スケジューラを表す。図内の CFQ8 については後述する。

図よりホスト OS の I/O スケジューラは CFQ, AS, Deadline, NOOP の順に優れており、ゲスト OS の I/O スケジューラは Xen では AS が優れ、VMware Server では Deadline が優れていることがわかる。

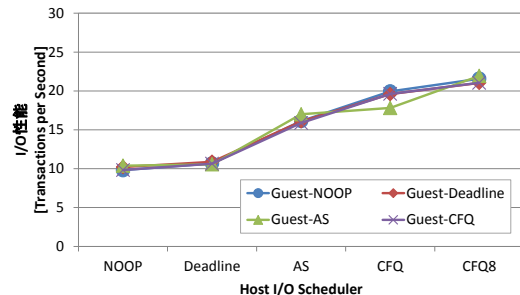


図 1. 仮想化環境における I/O スケジューラの性能 (Xen)

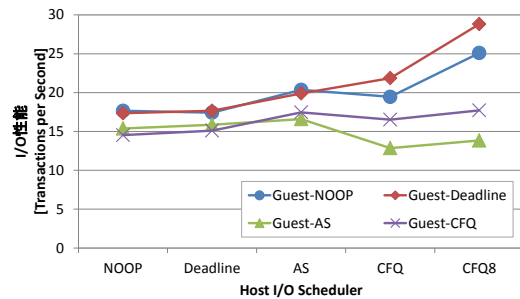


図 2. 仮想化環境における I/O スケジューラの性能 (VMware Server)

4. スケジューリング結果の解析

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。仮想化システムとして Xen を用い、ホスト OS の I/O スケジューラが NOOP、ゲスト OS の I/O スケジューラが AS のときの発行 I/O を図 3 に示す。縦軸が発行された I/O の物理ディスクにおけるアクセスアドレス、横軸が I/O の処理時刻を表している。同様にホスト OS の I/O スケジューラが CFQ、ゲスト OS の I/O スケジューラが AS のときの発行 I/O を図 4 に示す。仮想化システムとして VMware Server を用い、ホスト OS の I/O スケジューラが NOOP で、ゲスト OS の I/O スケジューラが Deadline のときの発行 I/O を図 5 に示し、ホスト OS の I/O スケジューラが CFQ、ゲスト OS の I/O スケジューラが Deadline のときの発行 I/O を図 6 に示す。

図 3 より、仮想化システムとして Xen を用い、ホスト OS の I/O スケジューラに NOOP を用いた場合は、アクセス対象の VM イメージの変更が多く発生していることがわかる。よって、HDD ヘッドの移動距離が長く多くの時間を要する I/O が多数発生していることになる。これに対して図 4 の例では、アクセス対象 VM の変更が図 3 より少なく、I/O 性能が高くなっていることがわかる。

また、図 5 より仮想化システムとして VMware Server を用い、ホスト OS の I/O スケジューラに NOOP を用いた場合は、図 3 よりアクセス対象 VM の変更が少ないことがわかる。これに対して図 6 の例では、アクセス対象 VM の変更が図 3 よりさらに少ないことがわかる。

I/O Schedulers Behavior Analysis in Virtualized Environment
Kenichi Nii[†], Saneyasu YAMAGUCHI[†]

[†]Electrical Engineering and Electronics, Kogakuin University Graduate School

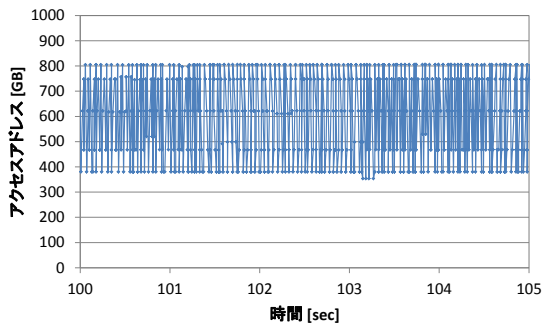


図 3. Xen における発行 I/O
(ホスト OS-NOPP, ゲスト OS-AS)

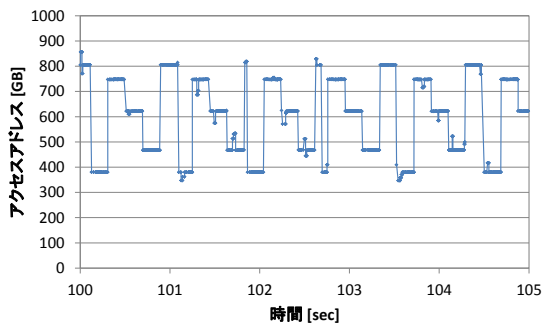


図 4. Xen における発行 I/O
(ホスト OS-CFQ, ゲスト OS-AS)

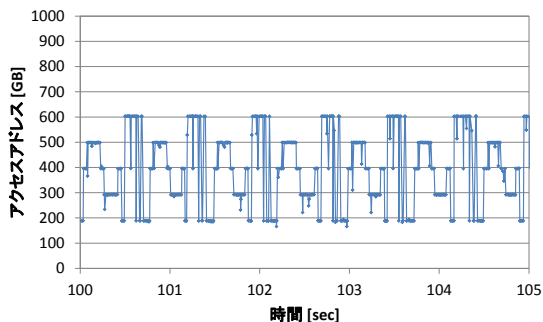


図 5. VMware Server における発行 I/O
(ホスト OS-NOPP, ゲスト OS-Deadline)

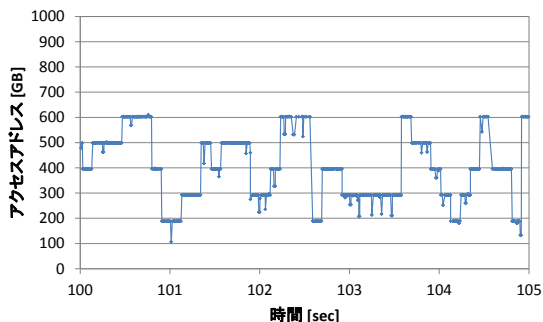


図 6. VMware Server における発行 I/O
(ホスト OS-CFQ, ゲスト OS-Deadline)

5. VM 間移動を削減するスケジューリング法

前章の考察を踏まえ、VM 間の移動をさらに削減するために、CFQ スケジューラのスライス時間及び遅延時間を大きくした。その性能を図 1, 図 2 の CFQ8 に示す。この変更により I/O 性能が向上しており、仮想化システムとして Xen を用いている図 1 では約 20%, VMware Server を用いている図 2 では約 30%の性能向上が確認された。

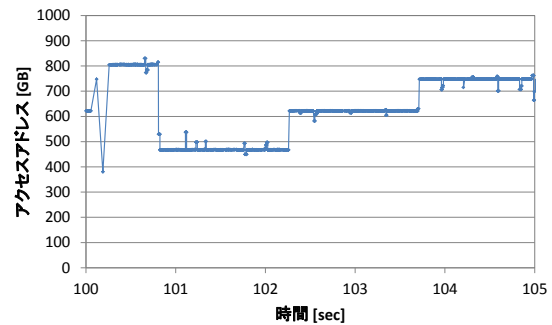


図 7. Xen における発行 I/O
(ホスト OS-CFQ8, ゲスト OS-AS)

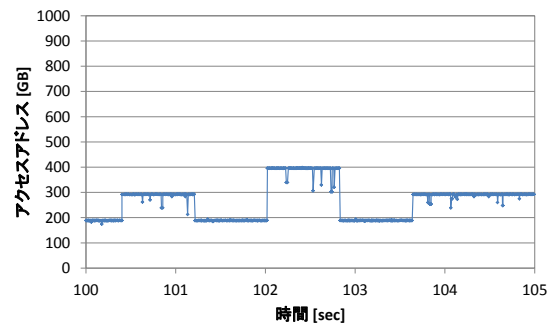


図 8. VMware Server における発行 I/O
(ホスト OS-CFQ8, ゲスト OS-Deadline)

また、前章と同様に発行された I/O をホスト OS 上でモニタリングした結果を図 7, 図 8 に示す。図より、Xen または VMware Server の二つの仮想化システムにおいて、ホスト OS の I/O スケジューラとして改善手法である CFQ8 を用いることで、VM 間移動が削減されることが確認された。

6. まとめ

本稿では、仮想化システムとして Xen, VMware Server を用いた場合における I/O スケジューラの性能評価と、スケジューリング結果の解析を行い、仮想化システムによらず性能低下の大きな原因として処理対象 VM の切り替えがあることを示した。

そして、VM 切り替え回数を削減する手法を提案し、その性能を評価した。性能評価の結果、改善手法により I/O 性能が Xen の場合約 20%, VMware Server の場合約 30%向上することが確認された。

今後は、ゲスト OS の I/O スケジューラの動作解析と性能向上に関する考察、公平性に関する考察を行う予定である。

謝辞

本研究は科研費(22700039)の助成を受けたものである。

参考文献

- [1] David Boutcher and Abhishek Chandra, "Does Virtualization Make Disk Scheduling Passé?," SOSP Workshop on Hot Topics in Storage and File System (Hot Storage '09)
- [2] 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 論文集 pp33-40