

仮想化環境における I/O スケジューラと I/O 性能の関係に関する一考察

新居 健一[†] 山口 実靖[†]

[†]工学院大学情報通信工学科

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などが問題視され、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]があり、ホスト OS の I/O スケジューラとして NOOP が優れるとしているが、これは必ずしも汎用的な結論ではなくさらに多くの評価が必要な状況にあるといえる。

本稿では、代表的な仮想化システムである Xen を用いて、Linux 標準の 4 つの I/O スケジューラが仮想化環境において I/O 性能にどのような影響を与えるかについて考察する。そして、その動作解析と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, CFQ, AS 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処理する、公平性を重視し処理する、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、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、2 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。測定結果を図 1 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 AS16 については 5 章で後述する。図 1 より、ホスト OS、ゲスト OS ともに、AS を用いることが良いことが分かる。

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 をホスト OS 上でモニターし、その読み込みアドレスを図 2, 3 に示す。ホスト OS で NOOP, ゲスト OS で CFQ の組み合わせのものが図 2、ホスト OS で AS, ゲスト OS で AS の組み合わせのものが図 3 である。図 2, 3 では、両 VM のイメージファイルは、ブロックアドレス 540GB 付

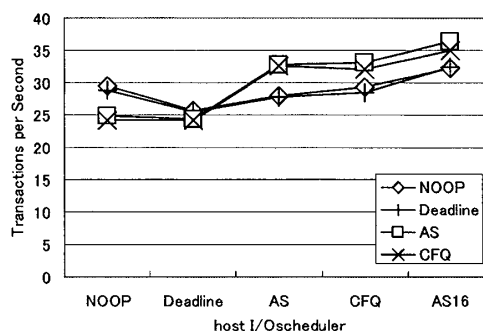


図 1. FFSB における TPS の比較

近、690GB 付近に存在している。

両図より、ホスト OS の I/O スケジューラに NOOP を用いると HDD ヘッドが両 VM 間を頻繁に移動し性能が低下するが、ホスト OS の I/O スケジューラに AS を用いると I/O が遅延され、近隣の I/O を待ち、VM 間の移動が少なくなっていることがわかる。

また、I/O 要求の処理に要した時間とシーク距離の関係についても調査を行った。FFSB を実行した際に発行された連続した 500 個の I/O 要求をグラフ化し、図 4, 5 に示す。横軸がシーク距離で単位が 4096[byte]のブロック、縦軸が I/O に要した時間を表している。

図 4, 5 では、同一の VM の I/O 処理を連続して処理した場合要求同士の間隔が大きくなることはないが、ホスト OS で NOOP, ゲスト OS で CFQ の場合のように VM 間を頻繁に移動する場合は横軸 40,000,000 の周辺にプロットが集中してしまう。図 4 では、10,000,000[block]以上の移動を行ったものが全体の約 75% を占めている。ホスト OS で AS, ゲスト OS で AS の場合では 10,000,000[block]以上移動したものが全体の約 10% であった。

以上のように、HDD ヘッドの VM 間の移動の有無が性能の優劣の原因となっていると考えられる。

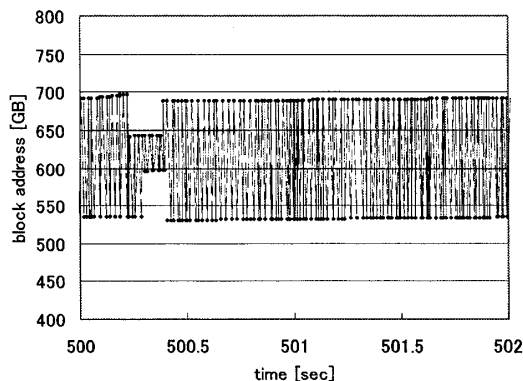


図 2. ホスト OS-NOOP ゲスト, OS-CFQ の Read アドレス

A Study of Performance of I/O Schedulers in Virtualized Environment

Kenichi NII[†], Saneyasu YAMAGUCHI[†]

[†]Department of Information and Communications Engineering, Kogakuin University

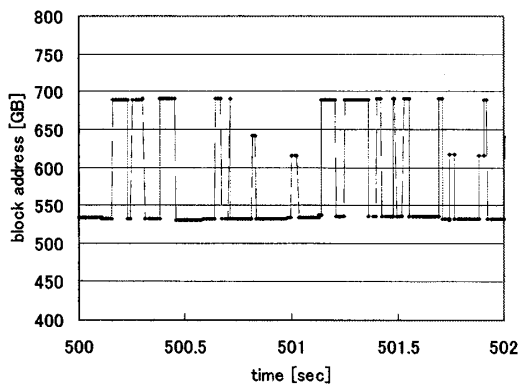


図 3. ホスト OS-AS, ゲスト OS-AS の Read アドレス

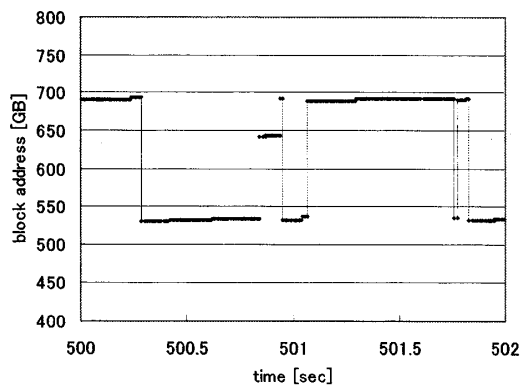


図 6. ホスト OS-AS16, ゲスト OS-AS の Read アドレス

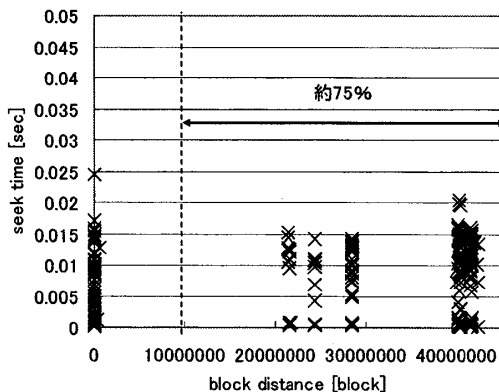


図 4. ホスト OS-NOOP, ゲスト OS-CFQ のシーク距離と時間

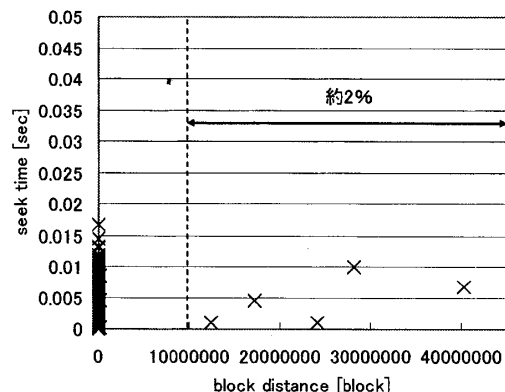


図 7. ホスト OS-AS16, ゲスト OS-AS のシーク距離と時間

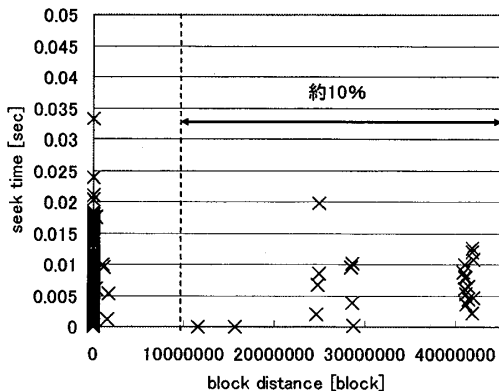


図 5. ホスト OS-AS, ゲスト OS-AS のシーク距離と時間

5. I/O スケジューラの改善

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、AS スケジューラの遅延時間を大きくした。その性能を図 1 の AS16 に示す。この変更により I/O 性能がさらに向上することが確認された。また、前章と同様にホスト OS からモニターした読み込み要求をグラフ化したものを図 6, 7 に示す。

ホスト OS の I/O スケジューラが AS であった場合の図

3 では、図中の 500 秒から 502 秒の間に 30 回程度 VM 間の移動が発生しているが、遅延時間を延長させた図 6 では 7 回程度の移動しか行われていない。このことにより、VM 間の移動がさらに削減されていることが確認できる。

また図 5, 図 7 を比較すると、10,000,000 [block] 以上の移動がさらに削減されており、図内では全体の 2% 弱となっている。

6. まとめ

本稿では、仮想化環境における I/O スケジューラの性能の評価と、スケジューリング結果の解析を行い、HDD ヘッドの VM 間移動を軽減させることが I/O 性能において重要であることを明らかにした。

そして、VM 間移動をさらに削減する改善手法として遅延時間を延長させることで I/O 性能の向上を図る手法を提案し、性能の向上を確認した。

今後は、仮想化環境に適した I/O スケジューラの作成を行い、その評価を行う予定である。

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 Systems (HotStorage '09)