

仮想計算機環境を考慮した 仮想 HDD イメージファイルの再配置手法の性能評価

山田 将也[†] 山口 実靖[†]

工学院大学大学院工学研究科電気・電子工学専攻[†]

1 はじめに

I/O 性能の向上を実現する手法の一つに HDD 上のデータの配置を変更する手法があり、これまで研究されてきている[1-9]。しかし、これまでに提案されてきた再配置手法は仮想計算機(VM)環境を想定しておらず、これらの手法を工夫なく VM 環境に適用しても得られる性能向上は限定的であると予想される。また VM は Web サーバや DB サーバなどに用いられることが多く、このような使用方法の場合は頻繁にアクセスされる領域(ホットスポット)が既知であることが多い。よって、その領域の移動による I/O 性能の向上が効果的に行えると期待できる。そこで我々はこれまでに VM 環境に適した再配置手法を提案し、単一種類のアプリケーションが動作する環境にて性能評価を行い、既存手法よりよい性能を示すことを確認してきた[10]。本稿では、1 台の物理ストレージ上に複数の仮想 HDD イメージファイルが存在し、各 VM 内で複数種類のアプリケーションが動作する VM 環境でデータ再配置手法を適用した際の性能評価を行う。

2 既存研究

初期のディスクレイアウトの理論に関する研究として文献[1]がある。文献[1]では、最高頻度アクセスデータをストレージの中央に配置する organ pipe 手法がランダムアクセスに適していることが示されている。organ pipe手法を現実のワークロードに適用した研究として cylinder shufflingがある[2]。本手法ではシリンダー単位で並び替えを行う。並び替えの単位をブロックとすることによりさらなる高速化を実現した研究として文献[3]がある。FFSB[4]にて、関連するデータブロックと inode をディスク上の近隣に配置することにより I/O 速度を向上させる方法が提案されている。ログ構造化ファイルシステム[6]では、大幅な書き込みの性能の向上が実現されている。また文献[5]にて、参照の局所性ではなく、微小ファイルの距離を近づけることに着目して性能を上げる方法が提案されている。HFS[7]や Smart File System[8]では、ファイルシステム(FS)が実行時アクセスパターンを観察し、高頻度アクセスデータを予約領域に移動を行っている。

FS2[9]では、ファイルの複製を用いて連続アクセス Performance Evaluation of Virtualized Environment Aware Virtual HDD Image File Layout Optimization

[†]Masaya Yamada [†]Saneyasu Yamaguchi

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

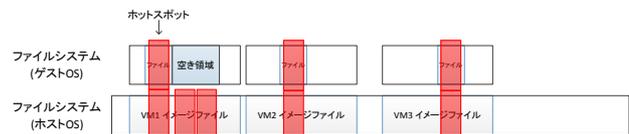


図1 二重FS構造を考慮したデータ再配置手法

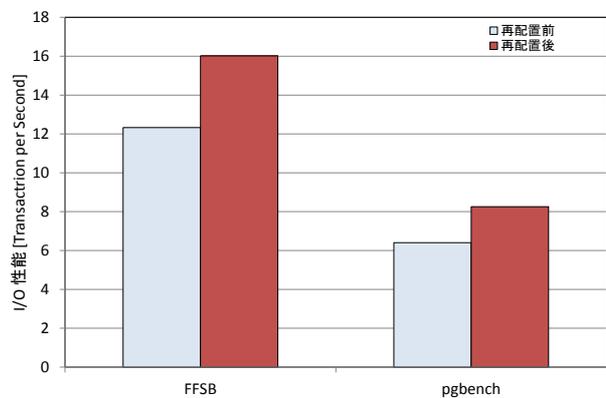


図2 I/O性能

されるファイルを近隣に配置することにより I/O の高速化を実現している。

3 二重FS 構造を考慮したデータ再配置手法

VM環境では、ホストFS上に巨大なイメージファイルが作成され、そのイメージファイルの中にゲストFSが構築される。ホストFSのブロック使用状況管理においては、イメージファイルにより確保されている全領域が使用中としてみなされる。しかし、このイメージファイルの中にゲストFSが構築され、ゲストFSがこのイメージファイルの各ブロックを使用中あるいは未使用として分類する。すなわち、ホストFSにおいては使用中と扱われるがゲストFSにおいては未使用と扱われる領域があり、この領域は実質的に未使用領域であり、ゲストFSにおける新規領域確保時を除いて使用されることがない。本手法では図1の様に、ゲストFSにおける未使用領域にホットスポットを移動する。これにより、巨大なシークを削減しVMのI/O性能を向上させることが可能である[10]。

4 性能評価

1 台の物理計算機上に 3 台の VM を稼働させ、それらの VM 上でベンチマークソフト FFSB と pgbench を同時に実行し、提案手法を適用し性能を測定した。実 HDD の容量は 1[TB]で各 VM の仮想 HDD イメージファイルの大きさは 100[GB]である。

ベンチマークのデータサイズは VM メモリの 8 倍となる 2[GB]とした。これはキャッシュの影響が大きく現れないようにするためである。

測定結果を図 2 に示す。図 2 より、データ再配置を実行することにより、VM の I/O 性能が約 30% 向上することが確認できた。再配置実行前のベンチマーク実行時の実 HDD へのアクセス分布を図 3 に、再配置実行後のアクセス分布を図 4 に、再配置実行前のシーク距離とシーク時間の関係を図 5 に、再配置後のシーク距離とシーク時間の関係を図 6 に示す。図 3, 4 から、再配置を実行する前は 100[GB]以上の長距離シークが多数発生しているが再配置を実行することによってそれらのシークを削減できたことが確認できる。これにより、VM の I/O 性能が向上したと考えられる。

5 おわりに

本稿では VM 内に複数種類のアプリケーションが動作する仮想化環境にて我々の提案手法の性能評価を行った。評価の結果、データ再配置を行うことによって VM の I/O 性能が約 30% 向上することが確認できた。今後は、VM ごとに異なるアプリケーションが動作している環境での性能評価を行う予定である。

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] C.K.Wong, "Algorithmic Studies in Mass Storage Systems," Computer Sciences Press, 1983.
- [2] C.Ruemmler and J. Wilkes, "Disk Shuffling," HP Technical Report, HPL-CSP-91-30, 1991.
- [3] P.Vongsathorn and S. D. Carson, "A System for Adaptive Disk Rearrangement," Software Practice Experience, 20(3): 225-242, 1990.
- [4] M.K.McKusick et al., "A Fast File System for UNIX," ACM Transactions on Computing Systems (TOCS), 2(3), 1984.
- [5] Greg Ganger and Frans Kaashoek, "Embedded Inodes and Explicit Groups: Exploiting Disk Bandwidth for Small Files," USENIX Annual Technical Conference, 1-17. 1997.
- [6] M.Rosenblum and J. Ousterhout, "The Design and Implementation of a Log-Structured File System," ACM Transactions on Computer Systems, 26-52, 1992.
- [7] HFS Plus Volume Format, <http://developer.apple.com/technotes/tn/tn1150.html>
- [8] C.Staelin and H.Garcia-Molina, "Smart Filesystems," USENIX Winter, 45-52, 1991. 2001.
- [9] Hai Huang, Wanda Hung, Kang G. Shin "FS2: Dynamic Data Replication in Free Disk Space for Improving Disk Performance and Energy Consumption," SOSPO'05, pp.263-276, October. 2005.

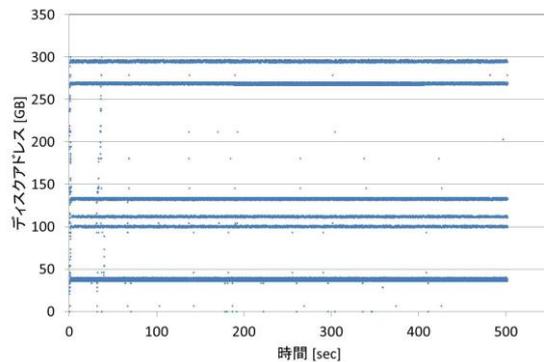


図 3 実 HDD へのアクセス分布(再配置前)

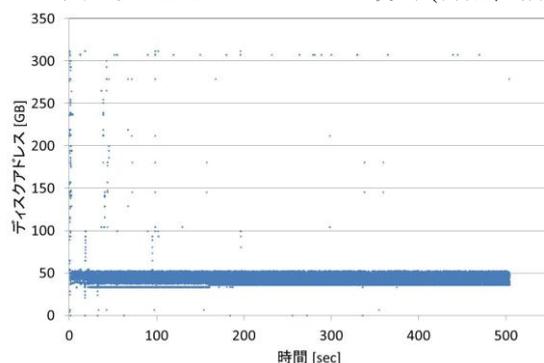


図 4 実 HDD へのアクセス分布(再配置後)

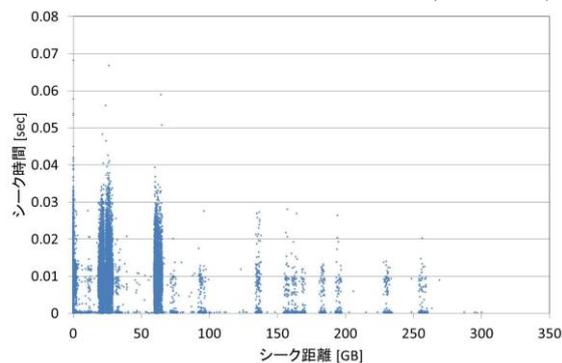


図 5 シーク距離とシーク時間の関係(再配置前)

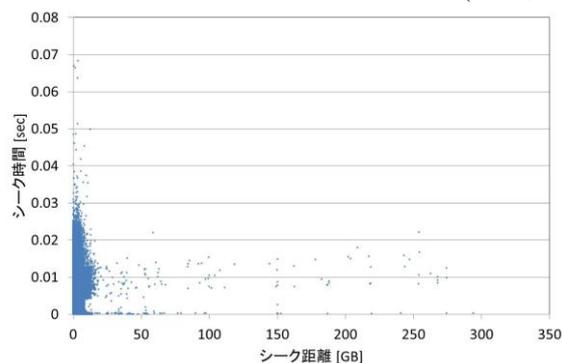


図 6 シーク距離とシーク時間の関係(再配置前)

- [10] 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", WebDBF 2011, 2011 年 11 月