

## 仮想計算機上 OS と VMM 間の制御移行に着目した性能評価

島谷 隼生†

山内 利宏‡

谷口 秀夫‡

† 岡山大学大学院自然科学研究科

‡ 岡山大学学術研究院自然科学学域

## 1 はじめに

仮想計算機（以降、VM）環境では、VM 上で動作するプロセスの処理内容に応じて、VM 上 OS から VM モニタ（以降、VMM）への制御移行が発生し、VMM による処理を必要とする。このため、VM 上で動作するプロセスは、仮想化による影響を受け、物理計算機（以降、BM）上で動作する場合よりも実行時間が長大化する。

本稿では、ベンチマークアプリケーションを用いて、BM 環境と VM 環境におけるプログラムの実行時間を評価した結果について述べる。また、VM 上 OS と VMM 間の制御移行と VM 上のプログラム実行時間の長大化との関係を明らかにする。

## 2 評価

## 2.1 観点

VM 上のプログラムは、VMM による処理を必要とする場合（例：ディスク入出力処理）、BM 環境で動作する場合と比べ、VMM による処理が加わるため、実行時間が長大化する。この実行時間の長大化は、VMM の処理内容に依存する。このため、VM 上 OS から VMM への制御移行数に着目して評価した。

## 2.2 評価方法と評価環境

評価には、PARSEC Benchmark スイート [1] 中のベンチマークプログラムを使用した。具体的には、13 個のベンチマークプログラムのうち、FreeBSD 11.3-RELEASE 上でビルド可能な次の 7 つを使用した。blackscholes：財務分析、bodytrack：画像処理、dedup：重複排除処理、ferret：画像検索、fluidanimate：物理シミュレート処理、freqmine：トランザクションデータ分析、および streamcluster：クラスタリング処理である。

プログラムを 10 回実行し、平均実行時間を測定した。また、VM と VMM 間の制御移行数と制御移行により発生した VMM の処理時間を測定した。

評価には、Intel Xeon E5-2630 v3（8 コア、2.4 GHz、20 MB L3 キャッシュ）、メモリ 32 GB、および HDD WD5000AZRZ（5,400 rpm、64 MB キャッシュ）を搭載した計算機を用いた。BM と VM の OS は、FreeBSD 11.3-RELEASE を使用し、VMM として bhyve を使用した。また、VM には、1 コアと 8GB のメモリを割り当て、BM 評価時も同等とし、コア数とメモリを同じ 1 コアと 8GB に制限した。

Performance Evaluation Focusing on Control Transitions between VMM and OS on VM.

Toshiki Shimatani†, Toshihiro Yamauchi‡, Hideo Taniguchi‡  
†Okayama University ‡Okayama University

## 2.3 結果と考察

## 2.3.1 BM/VM 環境におけるプログラムの実行時間

各環境の平均実行時間を表 1 に示し、BM 環境を 1 とした場合の VM 環境の平均実行時間を図 1 に示す。表 1 と図 1 より以下のことがわかる。

(1) dedup において、BM 環境に比べ、VM 環境の平均実行時間の増加が最も大きい（約 1143 ms）。これは、ディスクに対する書き込み処理が他のベンチマークよりも多いためである。dedup ベンチマークでは、与えられた入力データセットの重複を排除した後、処理結果をファイルに書き出す。VM 環境では、準仮想化ドライバを使用しないため、ディスクに対する入出力には VMM による処理を必要とする。

(2) ferret において、BM 環境に比べ、VM 環境の平均実行時間の増加が最も小さい（約 8 ms）。これは、VM 上で実行した ferret 全体の実行時間に占める VMM の処理時間の割合が小さいためである。ferret は、全体の実行時間が長く、ディスク入出力処理のような VMM による処理を必要とする処理が少ないため、VM 環境における実行時間の長大化の影響が小さい。

## 2.3.2 VM から VMM への制御移行数と VMM の処理時間

各ベンチマークについて、VM から VMM への制御移行数を表 2 に示し、VMM 移行後の VMM 内での処理時間を表 3 に示す。表 1、表 2、および表 3 より、以下のことがわかる。

(1) 表 1 と表 3 より、ferret を除いた全てのベンチマークにおいて、VM 環境で長大化した実行時間よりも、VMM 移行後の VMM 内での処理時間が短い。例えば、dedup において、VM 環境で長大化した実行時間は、約 1143 ms であり、VMM 移行後の VMM 内での処理

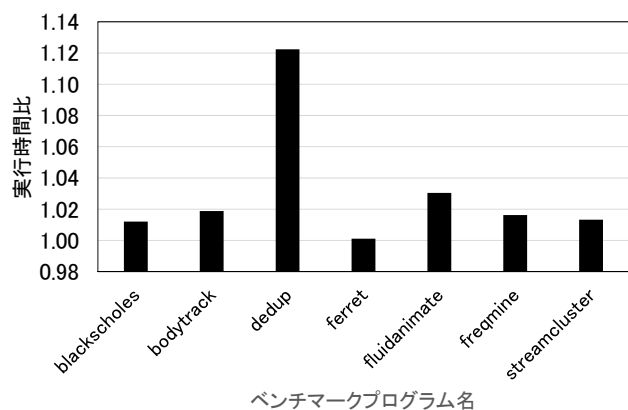


図 1 実行時間比

表1 BM/VM 環境におけるベンチマークの平均実行時間 [ms]

	blackscholes	bodytrack	dedup	ferret	fluidanimate	freqmine	streamcluster
BM	952	2,374	9,343	6,933	1,995	6,804	4,566
VM	963	2,419	10,486	6,940	2,055	6,915	4,626

表2 VM から VMM への制御移行数

通番	VMMへの制御移行の種類	blackscholes	bodytrack	dedup	ferret	fluidanimate	freqmine	streamcluster
1	PAUSE命令の発行	157	695	12,973	3,864	522	487	278
2	IN/OUT命令の発行	230	441	851	<b>4,201</b>	391	574	211
3	CPUID命令の発行	16	16	16	17	16	17	16
4	VMMによるエミュレートが必要な命令の発生	78	182	<b>3,882</b>	36	522	24	12
5	外部割込みの発生	1,191	3,018	12,103	8,389	2,661	8,786	5,680
6	HLT命令の発行	1	1	1	1	1	1	1
7	仮想APIC領域への書き込み	420	1,060	4,244	2,914	908	2,982	2,026
8	VMからVMMへの制御移行数の合計	2,093	5,413	<b>34,070</b>	<b>19,422</b>	5,021	12,871	8,224

表3 VMM 移行後の VMM 内での処理時間 [ms]

通番	VMMへの制御移行の種類	blackscholes	bodytrack	dedup	ferret	fluidanimate	freqmine	streamcluster
1	PAUSE命令の発行	0.708	3.090	57.953	17.759	2.270	2.241	1.266
2	IN/OUT命令の発行	1.179	2.300	4.334	<b>22.852</b>	2.093	2.932	1.109
3	CPUID命令の発行	0.007	0.007	0.008	0.008	0.008	0.008	0.007
4	VMMによるエミュレートが必要な命令の発生	1.096	2.778	<b>79.163</b>	0.427	9.318	0.301	0.152
5	外部割込みの発生	1.272	3.347	15.782	9.993	2.978	9.109	6.229
6	HLT命令の発行	0.145	0.120	0.128	0.129	0.121	0.129	0.115
7	仮想APIC領域への書き込み	0.086	0.229	1.215	0.746	0.208	0.650	0.395
8	VMM移行時のVMMの処理時間の合計	4.494	11.872	<b>158.582</b>	<b>51.914</b>	16.996	15.371	9.273

時間は約 159 ms である。これは、VM 環境における実行時間の長大化が VMM への制御移行以外の要因によっても発生するためである。VMM は、VM 環境におけるアドレス変換処理のために通常のページテーブルとは別のページテーブル（拡張ページテーブル）を用意している。これにより、VM 環境では、TLB ミス時に 2 段階のアドレス変換処理が発生し、BM 環境よりもアドレス変換コストが増加する。このため、VMM への制御移行以外にも VM 環境におけるプログラム実行時間が長大化する。

また、VM 環境で長大化した実行時間のうち、VMM 移行後の VMM 内での処理時間が占める割合を求めると、ferret を除き、最も大きい場合 (blackscholes) で約 39 % であり、最も小さい場合 (dedup) で約 14 % である。このことから、VM 環境におけるプログラム実行時間の長大化は、VMM への制御移行に伴う VMM の処理よりもメモリ仮想化による影響が大きいことが推測できる。

(2) 表 1 と表 3 より、ferret では、VM 環境で長大化した実行時間よりも、VMM 移行後の VMM 内での処理時間が長い。具体的には、VM 環境で長大化した実行時間が約 8 ms であり、VMM 移行後の VMM 内での処理時間は、約 52 ms である。これは、VMM 移行後の VMM 内での処理時間に含まれている tty デバイスへの書き込み処理時間が BM 環境でも同様に生じるためである。

(3) 表 1 と表 2 より、VM 環境におけるプログラム実行

時間の長大化が最も大きい dedup は、他のベンチマークよりも、VMM によるエミュレートが必要な命令の発生を理由とする制御移行が多い。例えば、ferret よりも 3,846 回多く発生している。これは、ファイルへの書き込みに伴い発生するメモリマップド I/O によるものである。

(4) 表 2 と表 3 より、VMM によるエミュレートが必要な命令の発生を理由とする制御移行の 1 回あたりの VMM の処理時間を求めると、約 11~20  $\mu$ s である。一方、IN/OUT 命令の発行を理由とする制御移行の 1 回あたりの VMM の処理時間を求めると、約 5  $\mu$ s である。このことから、VMM によるエミュレートが必要な命令の発生に伴う VMM の処理時間が、他の理由に伴う VMM の処理時間よりも長いため、VMM によるエミュレートが必要な命令の発生が多い dedup では、VM 環境におけるプログラム実行時間の長大化が大きい。

### 3 おわりに

VM 環境において長大化したプログラム実行時間のうち、VMM への制御移行に伴う VMM の処理時間が占める割合は、最大で約 39 % であることを示した。また、VMM によるエミュレートが必要な命令の発生が多い場合に VM 環境のプログラム実行時間がより長くなることを示した。

#### 参考文献

- [1] Bienia, C., Kumar, S., Singh, J. P. and Li, K.: The PARSEC Benchmark Suite: Characterization and Architectural Implications, PACT '08, ACM, p. 72–81 (2008).