

MKng プロジェクトにおけるマイクロカーネル移植技法： 3 Z-7 MIPS 用マイクロカーネルおよび UNIX サーバの性能評価†

光澤 敦 林 和則 中野 貴之 田中 浩一 徳田 英幸

NTT 情報通信研究所

ソニー IT カンパニー

慶應義塾大学 環境情報学部

1 はじめに

近年オペレーティングシステムの研究プラットフォームとして、PC 互換機が多く用いられている。i386 互換 CPU ならびに PC 互換機が安価だからであろう。しかしながら、PC 互換機のアーキテクチャは、そのままでは将来性に疑問が残る。

例えば、最も重要なソフトウェア動向の一つとして、プログラムやデータによって使用されるメモリ量が年に 1.5 倍から 2 倍の割合で増大している事実がある。また最近の研究によれば、キャッシュや TLB(Translation Lookaside Buffer) で構成されるメモリシステムが計算機の性能上のボトルネックの一つになることが予想されている。しかしながら、i386 互換 CPU は未だに 32 ビット CPU であり、メモリシステムにも根本的な変更は見られない。そこで我々は、慶應義塾大学環境情報学部を中心に行なっている次世代マイクロカーネル (MKng) 研究プロジェクト [1] のサブタスクとして、MIPS アーキテクチャ上で次世代マイクロカーネルの研究を行なっている [2]。

本研究では、その基盤環境構築を目的として、Mach3 マイクロカーネルおよび 4.4BSD Lite サーバを、MIPS プロセッサを搭載したソニー製 NEWS ワークステーション (以下、NEWS と呼ぶ) に移植し、その性能を測定する。本稿では、まず MIPS アーキテクチャにおける研究課題について紹介し、3 章で NEWS プラットフォームへの移植について述べる。4 章で性能評価について述べ、5 章でまとめる。

2 MIPS アーキテクチャにおける研究課題

第一の研究課題は、64 ビットアドレッシングのサポートである。64 ビットアドレス空間を効率良く管理するには、(1)TLB のヒット率を上げること、(2)TLB ミス時のページテーブル検索時間を短縮すること、および (3) ページテーブルの空間効率の良い実装、が必要である。(1) に対するアーキテクチャでの解決技術として、スーパーページおよびサブブロックが存在する [3]。(2) および (3) に対するオペレーティングシステムにおける解決技術に、クラスタ化ページテーブル [3] とガード付きページテーブル [4] がある。しかしながら、これらの技術はその有効性がシミュレーションを通して議論されているに過ぎない。そこで、R4400 および R10000 をはじめとする最近の MIPS プロセッサに実装されたスーパーページおよびサブブロックの機能を用いて、これらの技術を用いた新しいアドレス変換機構を実装し、性能評価する。

第二の研究課題は、マルチプロセッサのサポートである。こ

Microkernel Porting Technology in the MKng Project:
Performance Evaluation of Mach3 Microkernel and 4.4BSD Lite Server on MIPS Architectures

Atsushi MITSUZAWA¹, Kazunori HAYASHI², Takashi NAKANO², Kouichi TANAKA², and Hideyuki TOKUDA³

¹NTT Information and Communication Systems Laboratories, 1-1 Hikarinooka, Yokosuka-shi, Kanagawa 239 Japan, ²Information Technology Company, Sony Corporation, 3-3-1 Tsujido Shinmachi, Fujisawa-shi, Kanagawa 251 Japan, ³Faculty of Environmental Information, Keio University, 5322 Endo, Fujisawa-shi, Kanagawa 252 Japan

†この研究は、情報処理振興事業協会 (IPA) が実施している創造的ソフトウェア育成事業「次世代マイクロカーネル研究プロジェクト」のもとに行われた。

れは、以下の 4 つの機能により実現される。

- リアルタイムプロセッサ同期機構。
- リアルタイムプロセッサ間通信機構。プロセッサ間のデータ転送を実現するための高速なプロセッサ間通信機構で、1 対 1 および 1 対多の通信の両方を実現。
- マルチプロセッサに対応した仮想記憶機能。上記のプロセッサ間通信機構による実装。
- マルチプロセッサ対応割り込み処理機構。デバイスドライバの割り込み処理を特定のプロセッサに割り当てて行なうか、複数のプロセッサに割り当てて行なうかを、デバイスドライバ毎かつ動的にコンフィギュレーション可能にするための機構。
- マルチプロセッサ対応のリアルタイムプロセス間通信 (Real-Time Mach IPC)。

3 NEWS への移植

3.1 NEWS プラットフォーム

前章で述べた課題を研究するには、その基盤環境を構築することが必要である。そこで本研究では、NEWS に Mach3 マイクロカーネル (MK83) および 4.4BSD Lite サーバ (Lites-1.1.u3) を移植する。移植に用いた NEWS の種類およびその簡単な仕様を示したのが表 1 である。

表 1: MIPS プラットフォーム

NWS-	3150	5000	7000
CPU(MHz)	R3081(40)	R4400(200)	R1M(200)
ISA(bit)	I(32bit)	III(32/64)	IV(32/64)
NCPU	1	1	1,2,4
MM	8M-32M	32M-512M	32M-2G
PCache(I+D)	16K+4K	16K+16K	32K+32K
SCache	-	1M	1M
I/O	-	APbus	APbus

ここで、プロセッサの性能としては既に見劣りする R3081 を搭載した NWS-3150 へも移植を行なうのは、同等のプロセッサである R3000 を搭載した DECstation で動作する Mach3 マイクロカーネルがあるので、プラットフォームの相違点だけに着目して移植することができるからである。相違点としては、(1) エンディアン、(2) バックプレーンバスおよび割り込み、および (3) デバイスである。NWS-5000 プラットフォームは、二つの目的に使用する。一つは、R4000 以降の MIPS プロセッサに採用された (1) ページテーブルエントリおよび (2) TLBmiss に関する例外ベクタ構成への対応である。このことは、Mach3 の物理メモリ管理および例外処理を大幅に変更しなければならぬことを意味する。それで NWS-5000 における最初の移植では、R4400 を 32 ビットモードで動作させ新しい物理メモリ管理および例外処理の実装を行なう。もう一つは、研究課題の一つである 64 ビットアドレッシングにおける新しいアドレス変換機構の実装である。NEWS-7000 は、マルチプロセッサの研究を行なうためのプラットフォームである。前章で挙げた研究

課題の他に、デバイスドライバをマルチプロセッサ対応にする必要がある。

3.2 戦略

前節で述べた各プラットフォームの役割から、表2に示す4つのステップに分けて移植を行なう。ステップを経て移植することにより、システムに対する理解が徐々に深まるため、結果的に移植に要する時間を短縮できると考えられる。表2に、各ステップでのプラットフォーム、ホスト環境、コンパイラ、および実行可能形式を記す。

表2: プログラミング環境

Step	1	2	3	4
機種	3150	5000	5000	7000
ホスト	NEWS		SGI	
OS	NEWS-OS4(BSD)		IRIX6	
cc	GCC 2.7		MIPSPPro 7.0	
as	MIPS as-2.11		MIPSPPro 7.0	
ld	MIPS ld		MIPSPPro 7.0	
a.out	ECOFF(static)		ELF(static)	

ホスト環境としては、ステップ2まではNEWS上で動作するNEWS-OS4を、ステップ3からはSGI社のIRIX6を使用した。なぜなら、(1)64ビットアドレッシングのコード生成、(2)MIPS-ISA III以降のコード生成、および(3)マルチプロセッサのためのコード生成、をすべて行なう必要からである。

3.3 移植

移植作業はステップ2まで終了している。すなわち、R3081を搭載したNWS-3150およびR4400を搭載したNWS-5000上で、MK83およびLites-1.1.u3が動作中である。現在サポートしているデバイスに、シリアル(Zilog Z85230 ESCC)、SCSI(HP SPIFF3)、イーサネット(SONIC)、フレームバッファ(NWB-1102)、マウス、およびキーボードがある。

表3に、NWS-5000へ移植するにあたっての主な変更点を挙げる。

表3: R4400を搭載したNWS-5000への移植

MIPS 非依存部分 (行)			
Tools(COFF)	約 150	MK(endian)	約 50
MK(SCSI)	約 500	Lites(fixes)	約 50
MIPS 依存部分 (行)			
MK(mips.cpu.h)	525	MK(mips.init.c)	326
MK(trap.c)	72	MK(pmap.c)	56
MK(as.macro.h)	234	MK(start.s)	383
MK(locore.s)	622	MK(tlbctl.s)	1209
MK(cache.s)	1140	MK(endian)	約 150
Lites(signal)	1315		
NEWS 依存部分 (行)			
MK(ROM 関係)	約 300	MK(diskless)	1209
MK(APbus)	3881	MK(ESCC)	約 170
MK(SCSI)	2302	Lites(NEWS-OS4)	約 500

UNIX環境としては、Lites-1.1.u3が4.4BSD LiteおよびNEWS-OS4の両方をエミュレートする。これは、Lites-1.1.u3に4.4BSD LiteとNEWS-OS4の両方のシステムコールを実装し、システムコール番号によって対応するシステムコールを呼び出すことで実現される。現在は、ifconfig、mount等システムの動作に密接に関わるものは、4.4BSD Liteのコマンドをソースコードから作成し、Xウィンドウ等のコマンドはNEWS-OS4のバイナリをそのまま使用している。

4 性能評価

表4が、NWS-5000に移植したシステムのByteBenchmarkによる性能評価である。ただし、評価に用いたNWS-5000の主記憶は32Mバイトである。なお、測定にあたっては、精度を高くするために適宜ループ回数を増やした。そのため、表4の数値は相対値としてのみ意味を持つ。

評価プログラムは、表2におけるステップ2のホスト環境上で、以下のオプションでコンパイルした。

`-O -DMORE_LOOP -s`

また、上記の共通オプションに、fcall、fcalla、dreadのコンパイルでは、それぞれ-EMPTYZ、-DASSIGN、-DTHIRTYTWOのオプションを加えてコンパイルした。

表4: NEWS-OS4との性能比較 (ms)

項目	NEWS-OS4			MK+Lites		
	real	user	sys	real	user	sys
pipes	0.6	0.0	0.2	8.6	0.5	1.3
scall	6.1	0.7	5.3	6.0	3.3	2.7
fcall	3.5	3.5	0.0	4.0	3.9	0.0
fcalla	0.2	0.2	0.0	0.3	0.2	0.0
sieve	0.1	0.1	0.0	0.2	0.2	0.0
dwrite	2.0	0.0	0.1	1.5	0.0	0.3
dread	0.1	0.0	0.1	2.0	0.2	0.5
sh	0.5	0.0	0.0	0.6	0.0	0.2

移植したシステムの性能は、NEWS-OS4と比較して全般的に悪い。原因の一つは、現在のNWS-5000の実装では、(1)PTEをR3000と同一サイズにしていること、(2)TLBmissをR3000と同じように処理していることであり、最終的な実装では表4より良い結果が得られると予想している。

5 まとめと課題

本稿では、Mach3マイクロカーネル(MK83)および4.4BSD Liteサーバ(Lites-1.1.u3)を、MIPSプロセッサを搭載したソニー製NEWSワークステーションへ移植した。このうち、NEWS-5000へ移植したシステムについて、ByteBenchmarkを用いて性能を測定し、NEWS-OS4と比較した。

今後の課題として、残されたステップ3以降を行なうと同時に、MPTS[5]によるマイクロカーネルの性能評価を行なう予定である。

参考文献

- [1] 徳田 英幸. MKng: 次世代マイクロカーネル研究プロジェクトの概要. In 第55回情報処理学会論文集(1Z-2), 1997.
- [2] 光澤 敦, 林 和則, and 田中 浩一. MachマイクロカーネルシステムのMIPSアーキテクチャへの移植. In 情報処理学会研究報告97-OS-75, pages 19-24, June 1997.
- [3] Madhusudhan Talluri, Mark D Hill, and Yousef A Khalidi. A New Page Table for 64-bit Address Spaces. In *Proceedings of the 15th ACM Symposium on Operating System Principles*, 1995.
- [4] Jochen Liedtke. Address Space Sparsity and Fine Granularity. *Operating Systems Review*, 29(1), January 1995.
- [5] François Barbou des Places and Philippe Bernadat and G.N.Madhusudan and Yves Paindaveine and Nick Stephen. *OSF Microkernel Performance Test Suite MPTS Release 1.1*. OSF Research Institute, April 1996.