

平成 21 年 5 月 25 日現在

研究種目：基盤研究 (C)  
 研究期間：2006～2008  
 課題番号：18500024  
 研究課題名 (和文) 継続点付き非中断型スレッドによる汎用高性能プロセッサ向け OS 構成法  
 研究課題名 (英文) Operating system on commodity processor using continuation-based non-interruptible threads  
 研究代表者  
 日下部 茂 (KUSAKABE SHIGERU)  
 九州大学・大学院システム情報科学研究所・准教授  
 研究者番号：70234416

研究成果の概要：

交付額

(金額単位：円)

	直接経費	間接経費	合計
2006 年度	1,300,000	0	1,300,000
2007 年度	1,100,000	330,000	1,430,000
2008 年度	1,000,000	300,000	1,300,000
年度			
年度			
総計	3,400,000	630,000	4,030,000

研究分野：計算機ソフトウェア

科研費の分科・細目：情報学・ソフトウェア

キーワード：オペレーティングシステム, マルチスレッド

## 1. 研究開始当初の背景

データフロー方式は本質的に並列・並行処理、非同期実行に適した実行方式とされ、逐次プロセッサシステムでも並行制御や非同期イベント処理を行うようなオペレーティングシステムにもデータフロー概念を導入することが有効と考えた。プロセッサアーキテクチャにおいてはデータフロー計算方式から派生した様々な技術が応用され対応したコンパイラ技術もある。しかしながら、オペレーティングシステムは論理的には多数の非同期イベント・並行処理を扱う一方、データフロー概念の積極的な導入はなされていない。近年は汎用的な環境でもスレッドがサポートされているものの、それらは逐次処理を前提に作成された複数のプログラムを多重に実行するためのプロセスモデルから派生

したものである。計算機利用形態の多様化にともない、サーバシステムなど多数の非同期イベントを効率的に扱う重要性は増大すると考える。イベントが非同期かつ大量に発生する場合には、逐次実行方式に忠実であるより、本質的に非同期実行であるデータフローモデルをベースにしたスレッド実行方式を積極的に導入する方がよいと考える。本研究の成果は、特に多層のメモリ階層を持ちモード切替が高コストな汎用高性能プロセッサのシステムで非同期イベントを大量に処理するような場合には、大きな効果が期待できると考える。

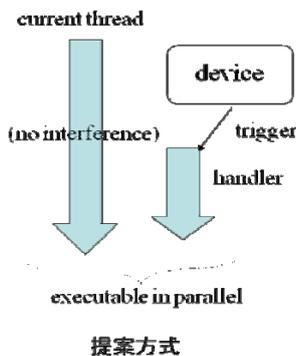
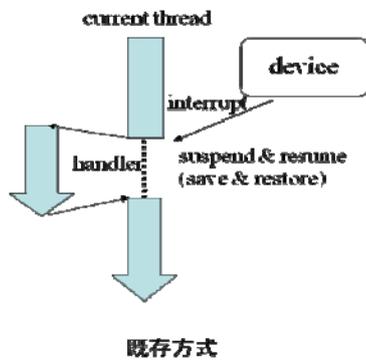
## 2. 研究の目的

本研究は高性能かつ柔軟な多重並行処理環

境を、現在広く利用されている慣習的なマルチスレッド実行方式とはベースとするモデルが異なるマルチスレッド実行方式にもとづいて構築しようとする構想を背景に持つ。前提としているマルチスレッド実行方式は簡単には次のような特徴を持つ；プログラムは半順序関係を持つスレッドから構成され、スレッドは実行開始後には終了まで実行が中断されない命令列として構成される。この場合プログラムは、逐次実行列であるスレッドをノードとし、スレッド間の内継続関係をエッジとするようなデータフローグラフとなり、本質的に多重並行処理に適合しやすいものとなる。また、スレッド内の命令は中断せずに実行されスレッド実行中は動的な環境管理は不要である。そのため逐次処理部と並行処理制御を系統立てて扱うことでシステム構成の単純化や性能の向上が期待できる。

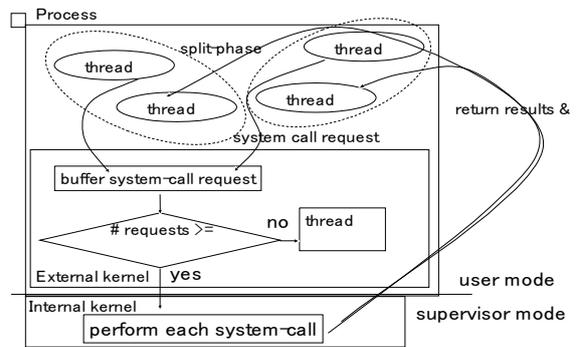
### 3. 研究の方法

(1) 非同期に大量のイベントが発生するようなシステムに対する効率の良い実行時環境を提供する枠組みについて、外部イベントを割り込みで処理する既存方式に代わるような継続を用いた方式の検討を行う。以下の図は従来方式を概念的に示しており、外部イベントに対する処理には実行中の処理を中断するための処理が必要となる。

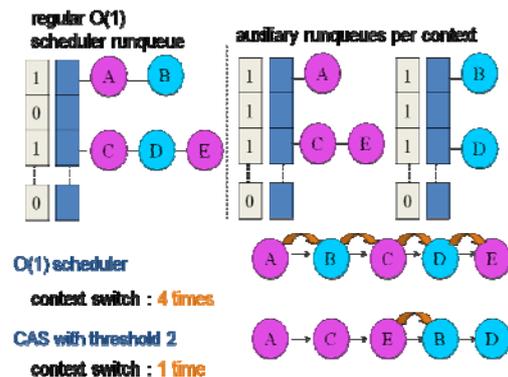


一方、提案方式は、外部イベントに対する処理も、内部的な計算も統一的なスレッド実行機構で処理するため、実行機構の単純化、低オーバーヘッド化、並列 I/O でのスケーラビリティが期待できる。

(2) 高性能汎用プロセッサの内部動作を考慮し、ユーザプログラムとオペレーティングシステムを連携させるようなスレッド分割のために、プロセッサでの命令列実行において干渉となるような内部イベントの地点でもスプリットフェーズ方式を使う。システムコールにおいて、ユーザプログラムからの要求発行と、ユーザプログラムの結果受け取り、カーネル内のシステムコール本体を別スレッドとしてオーバーヘッド削減を目指す一括システムコールについて、汎用プロセッサでの効果を調べる。下図に示すように、システムコールの要求は、事前に定められた方針で一旦バッファリングされる。これにより、実行モード変更などのオーバーヘッドの削減を目指す。



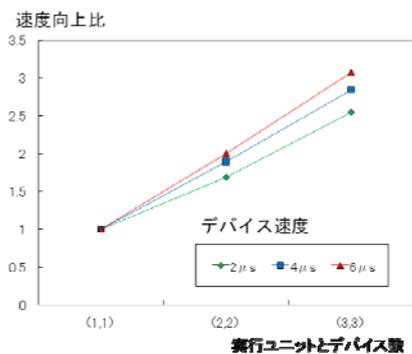
(3) 汎用プロセッサは性能向上のためにメモリに速度と容量が異なる階層を設けており、高い実行性能の達成にはそのようなメモリ階層の活用が重要である。スレッドのスケジューリングとメモリ階層の活用に着目し、メモリ空間の活用状況に反映できるようなスケジューラの研究を行う。以下は Linux 0(1) スケジューラへ提案手法を用いた場合の概念図を示す。



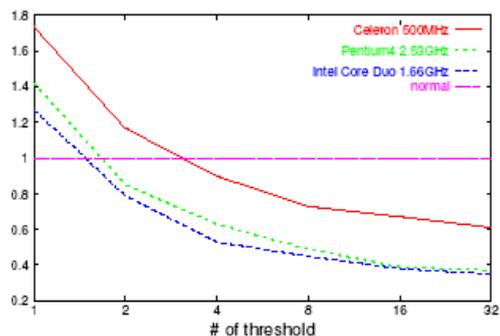
通常の実行可能なスレッドのキューに加え、実行可能なスレッドが属しているコンテキストの状況を反映できるような構造を付加し、スケジューリング時にコンテキストの共有状況に応じてスレッドを選択する。これによりコンテキスト切り替えのオーバーヘッド削減、メモリ階層の有効活用を目指す。

#### 4. 研究成果

(1) 非同期に大量のイベントが発生するようなシステムに対する効率の良い実行時環境を提供する枠組みについて、外部イベントを割り込みで処理する既存方式にかわるような継続を用いた方式の検討を行った。この点について汎用プロセッサ上での実現方式のための予備評価として、専用プロセッサ上での効果を確認し、国際会議等で発表を行った。下図は提案方式が良好な台数効果を示したものである。



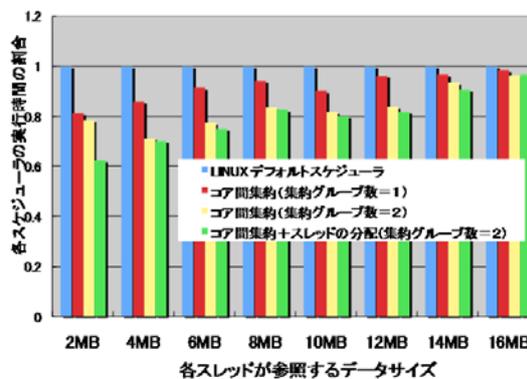
(2) 高性能汎用プロセッサの内部動作を考慮し、ユーザプログラムとオペレーティングシステムを連携させるようなスレッド分割のために、プロセッサでの命令列実行において干渉となるような内部イベントの地点でもスプリットフェーズ方式を使う。システムコールにおいて、ユーザプログラムからの要求発行と、ユーザプログラムの結果受け取り、カーネル内のシステムコール本体を別スレッドとしてオーバーヘッド削減を目指す一括システムコールについて、汎用プロセッサでの効果を調べ、国際会議等で発表を行った。



上記グラフは一括化するシステムコール数が多いほど実行コストが削減されている結果を示している。

(3) 汎用プロセッサは性能向上のためにメモリに速度と容量が異なる階層を設けており、高い実行性能の達成にはそのようなメモリ階層の活用が重要である。スレッドのスケジューリングとメモリ階層の活用に着目し、メモリ空間の活用状況に反映できるようなスケジューラの研究を行った。一定の効果を確認し、国際会議等で発表を行った。下のグラフはコンテキストを共有するスレッドを集約して実行することでスループットの向上が達成できたことを示すグラフである。

SysBench memoryプログラム実行時の各集約スケジューラの効果



#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計7件)

1. Shigeru Kusakabe, Hideo Taniguchi, Makoto Amamiya, "A Study of a Continuation-based Fine-grain Multithreaded Operating System CEFOS", Proc. of Workshop on Unique Chips and Systems co-located with IEEE ISPASS'09, 2009年4月, 査読有
2. Satoshi Yamada, Shigeru Kusakabe "Impact of priority bonuses of Inter-Core Aggregation Scheduler on a commodity CMP platform", Workshop on Managed Many-Core Systems (MMCS) co-located with ASPLOS'09, <http://www.cercs.gatech.edu/mmcs09/program.htm>, 2009年3月, 査読有
3. Satoshi Yamada, Shigeru Kusakabe, "Effect of Context Aware Scheduler on TLB," Proc. of Workshop on Multi-Threaded Architectures and Applications, (Proc. of IEEE International Parallel &

- Distributed Processing Symposium), Published in CD, 2008年4月, 査読有
4. Satoshi Yamada, Shigeru Kusakabe, "Development of a Thread Scheduler for Global Aggregation of Sibling Threads," Research Reports on Information Science and Electrical Engineering of Kyushu University, Vol.13, No.2, pp.69-74, 査読有
  5. Shigeru Kusakabe 他 6 名(第一著者), "Scalability of Continuation-based Fine-grained Multithreading in Handling Multiple I/O Requests on Fuce," Proc. of ACM International Conference on Computing Frontiers, pp.225-235, 2007年5月, 査読有
  6. Shigeru Kusakabe 他 7 名(第一著者), "OS Mechanism for Continuation-based Fine-grained Threads on Dedicated and Commodity Processors" Proc. of Workshop on Multi-Threaded Architectures Applications (Proc. of IEEE International Parallel & Distributed Processing Symposium), Published in CD, 2007年3月, 査読有
  7. Satoshi Yamada, Shigeru Kusakabe, Hideo Taniguchi, Makoto Amamiya, "Impact of Wrapped System Call Mechanism on Commodity Processors," Proc. of ICSOFT 2006, Vol.1, pp.308-315, 2006年9月, 査読有

[学会発表] (計4件)

1. 山田 賢, 日下部 茂 "アドレス空間を共有するスレッドの空間的な集約を行うスケジューラの評価", 第104回情報処理学会研究報告, 2007-OS-104, Vol.2007, No.10, pp.41-48, 2007年1月, 滋賀
2. 山田 賢, 日下部 茂, 谷口 秀夫, "高多重処理におけるスループット向上のためのスケジューリング機構の検証" 火の国情報シンポジウム予稿集 in CD, A-2-3, 2007年3月, 鹿児島
3. 青野光洋, 泉雅昭, 松崎隆哲, 日下部茂, 乃村能成, 谷口秀夫, 雨宮真人 "Zero-Wait方式を用いた多重 I/O 要求に対する高並列処理の提案と評価" 火の国情報シンポジウム予稿集, in CD, C-2-4, 2007年3月, 鹿児島
4. 青野光洋, 泉雅昭, 松崎隆哲, 日下部茂, 乃村能成, 谷口秀夫, 雨宮真人 "Zero-Wait方式による多数の I/O 要求に対する処理の Fuce 上での実現と評価" 情報処理学会システムソフトウェアとオペレーティング・システム研究発表会報告 (2006-OS-103), Vol.2006 No.86, pp.25-32, 2006年8月, 高知

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

名称:  
 発明者:  
 権利者:  
 種類:  
 番号:  
 出願年月日:  
 国内外の別:

○取得状況 (計0件)

名称:  
 発明者:  
 権利者:  
 種類:  
 番号:  
 取得年月日:  
 国内外の別:

[その他]

[www.aie.csce.kyushu-u.ac.jp](http://www.aie.csce.kyushu-u.ac.jp)

6. 研究組織

(1) 研究代表者

日下部 茂 (KUSAKABE SHIGERU)

九州大学・大学院システム情報科学研究所・

准教授

研究者番号: 70234416

(2) 研究分担者

( )

研究者番号:

(3) 連携研究者

( )

研究者番号: