システムプログラム概論
Introduction to Systems Programming
メモリ管理 (1)

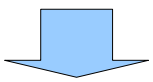
2006/4/17

門林　雄基
Youki Kadobayashi
NAIST

Copyright(C)2005 Youki Kadobayashi. All rights reserved.

今日の講義のポイント

- **問題は何か？**
  - memory hierarchy (**メモリ階層**)
  - **この複雑な技術を、単純なプログラミングで使いこなせるように出来ないか**



- **OSにおけるメモリ管理**
  (memory management)

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

今日の講義のポイント

- メモリ管理～その概念の発展
  - モノプログラミング
  - → 固定区画でのマルチプログラミング
  - → 仮想記憶

- 仮想記憶 (virtual memory)
  - Paging, page table, segmentation
  - TLB

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

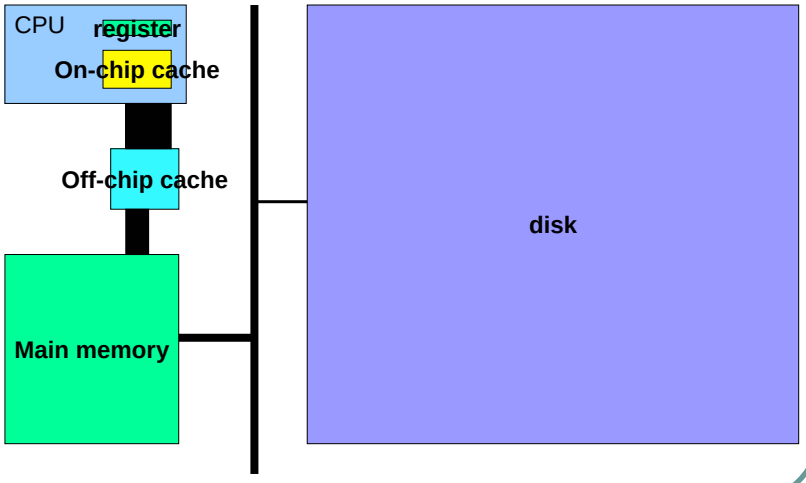
メモリ階層 (memory hierarchy)

- メモリ技術: アクセス速度、容量、価格のトレードオフ

	Latency	Size	cost
Register	0.13ns	512bytes	On chip
On-chip cache	4.7ns	2MB	On chip
Main memory	20ns	1GB	$0.1/MB
Disk	13ms	500GB	$0.0003/MB

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

メモリ階層



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

メモリ管理の目標

- 複数のプロセスがメモリという資源を競合して使用するのでこれをうまく調停する
- メモリ技術を組み合わせ、非常に高速かつ大容量のアドレス空間をユーザプログラムに提供する (virtual memory)
- アドレス空間を分離し、他のプロセスのバグ等の悪影響を受けないようにする (protection, software fault isolation)

- cache, swap, paging, segmentation 等のメモリ管理技術
  - 今日のシステムはこれらすべてを組み合わせている。
  - 以下では、これらを順にみていく。

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

資源割り当て問題としてのメモリ管理

- メモリ管理とはマルチプログラミングにおける資源割り当て問題の一種
- **競合** (contention)
  - 複数のプロセスがメモリという限られた資源を奪い合う
  - → これだけではうまくいかない。
- **調停** (arbitration)
  - 限られたメモリ上に複数のプログラムを共存させるために、メモリの割り当てを管理する必要がある。
  - → **メモリ管理** (memory management)
- 資源に対する競合と調停は OS の基本的な機能

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Part 1: memory management without hardware support

- ハードウェアの機構を用いないメモリ管理

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Monoprogramming (or uni-)

- 一度にひとつのプログラムだけを実行
- アドレス固定、境界チェックなし
- 最も原始的なメモリ管理
- メモリ上に他のプロセスが存在しない
  => protection なし

- MOS p. 191

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

アドレス可変 (relocatable)、境界チェックあり (bounds checking)

- マルチプログラミングが可能に
- メモリサイズの制限を受ける

- MOS p.192

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

スワップ

- **一つのプロセス全体をディスクに退避**
- swap out, swap in (MOS p. 197)
- **メモリサイズの制限を受けないが、**
  I/O **オーバーヘッドが問題** (p.4)

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

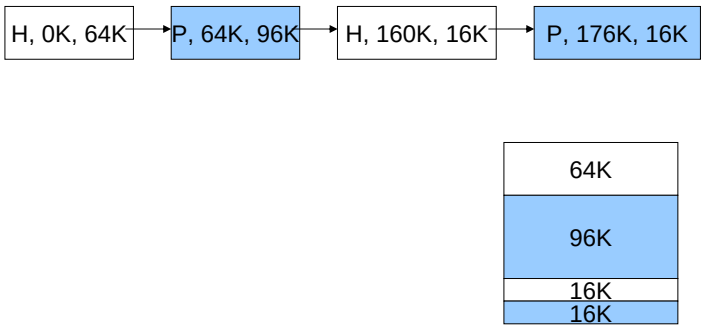
可変区画 (variable partition) 割付け

- **メモリセグメントのリストを維持**
  (H/P, start, length)
- **メモリの連続した空き領域** (hole) **を探す**
  - first fit: **充分な大きさの空き領域を検出するまでリストを走査**
  - next fit: **前回終了したところから検索を開始**
  - best fit: **該当する空き領域の中で最も小さいものを採用**

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

可変区画割付け

メモリ空間の計算機内におけるリスト表現



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

可変区画 (variable partition) 割付け

- best fit は first fit, next fit よりメモリの無駄が多い。
- first fit は平均的に大きなholeを作り出す。
- プロセスが要するメモリ容量が既知でない場合はどうするのか？

- brute-force approach: memory compaction
- better approach: virtual memory

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Part 2: memory management with hardware support

- ハードウェアの機構を用いたメモリ管理

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

仮想記憶が誕生した背景

- **参照の局所性** (locality of reference)
  - プロセスは実行中に全メモリ空間の一部しか必要としない
  - ⇩
  - プロセスの実行に必要な一部だけをメモリに置き、残りはディスクなどに置けば良い。
  - ⇩
  - 実メモリと、論理的なアドレス空間の分離

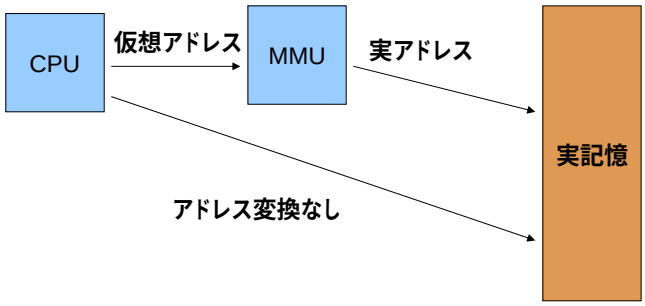
Copyright(C)2007 Youki Kadobayashi. All rights reserved.

仮想記憶 (virtual memory) と実記憶 (physical memory)

- **仮想アドレス、仮想アドレス空間** (virtual address space)
- **実アドレス、実アドレス空間** (physical address space)
- Address translation: **仮想アドレスから実アドレスへのマッピング**
- Memory Management Unit (MMU)
  - **通常** CPU **内に実装される**

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

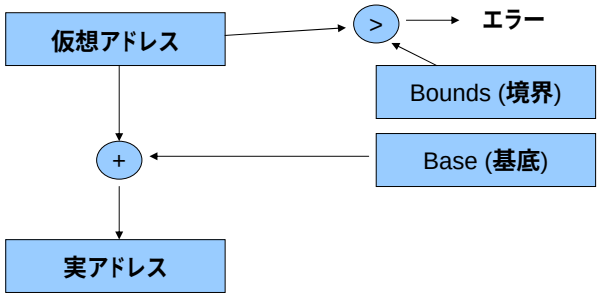
Address translation



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Base and bounds

- **仮想アドレス** + base = **実アドレス**
  - bounds **を超えるメモリアクセスは違反**
- **実メモリにおいて、可変区画割付けが必要**



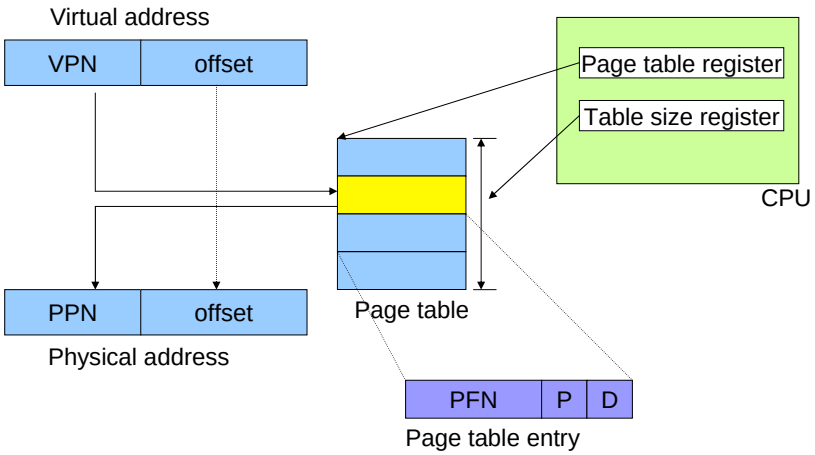
Copyright(C)2007 Youki Kadobayashi. All rights reserved.

ページング(1)

- **ページサイズを単位とした** address translation
- **(仮想ページ番号、オフセット) → (物理ページ番号、オフセット)**
  - (MOS p. 202〜)
- **ページテーブルを実メモリ上に維持**
- **ページテーブル･レジスタ、テーブルサイズ･レジスタ**
- present/absent bit, page fault

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

ページング(2)



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

ページング(3)

- メモリ管理が非常に単純になる -- ビットマップを用いたメモリ管理
  - 連続領域を探す必要がない
- 仮想アドレスを sparse に使う場合、ページテーブル長が爆発

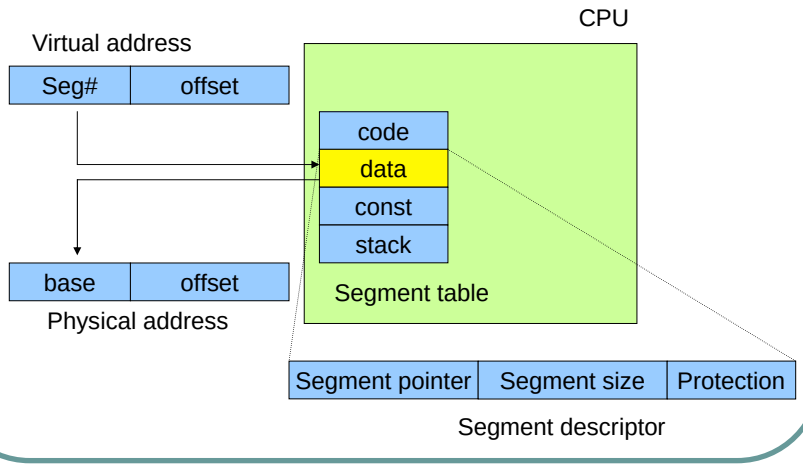
Copyright(C)2007 Youki Kadobayashi. All rights reserved.

セグメンテーション(1)

- base and bounds **の拡張。**
  コード、データ、スタック毎に base, bounds
  - (MOS p. 249〜)
- segment descriptor (**セグメント記述子**) **に以下の情報が含まれる**:
  - **セグメント・ポインタ、セグメント・サイズ、保護ビット**
- **セグメントテーブルを** CPU **内部に維持**

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

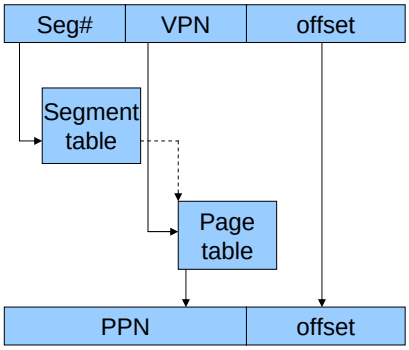
セグメンテーション(2)



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Multi-level translation: ページング + セグメンテーション

- **今日のプロセッサアーキテクチャではこれが主流。**
- **仮想アドレス** = (segment#, page#, offset)



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Multi-level page tables

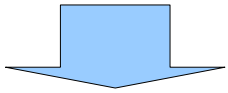
- **この方式をとるプロセッサもある**
- MOS p. 207〜 (**独習**)

- **セグメントを用いた方式と比較したときの得失は?**
  → MOS p. 252

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

メモリアクセスの高速化技法

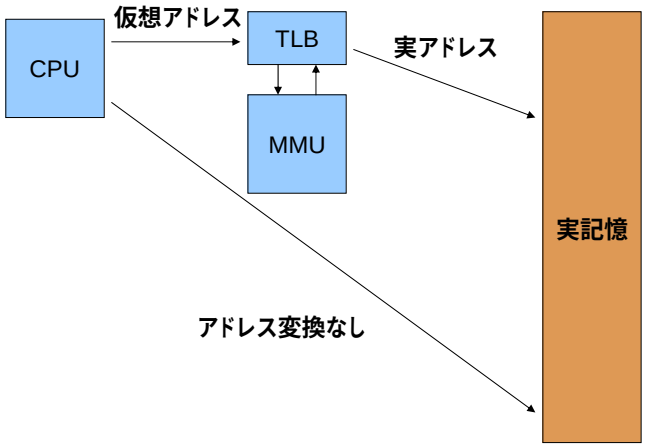
- address translationのオーバーヘッド
  - セグメントテーブル参照
    => ページテーブル参照
    => メモリアクセス



- Translation Look-aside Buffer (TLB)
  - 変換早見表 -- メインメモリへのアクセスなしにアドレス変換

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

TLB
Translation look-aside buffer



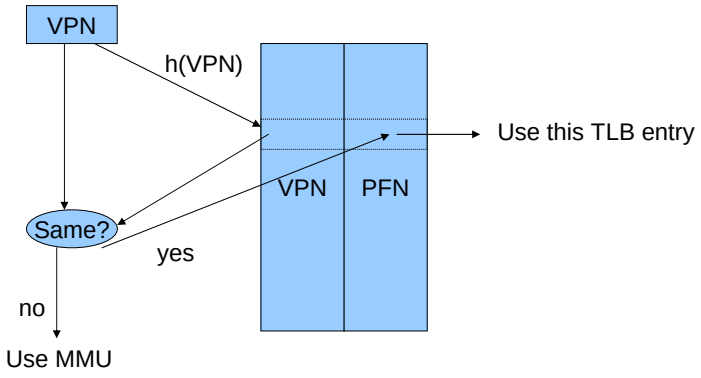
Copyright(C)2007 Youki Kadobayashi. All rights reserved.

TLB types

- Direct mapped TLB
- N-way set associative TLB
- Fully associative TLB (**略**)

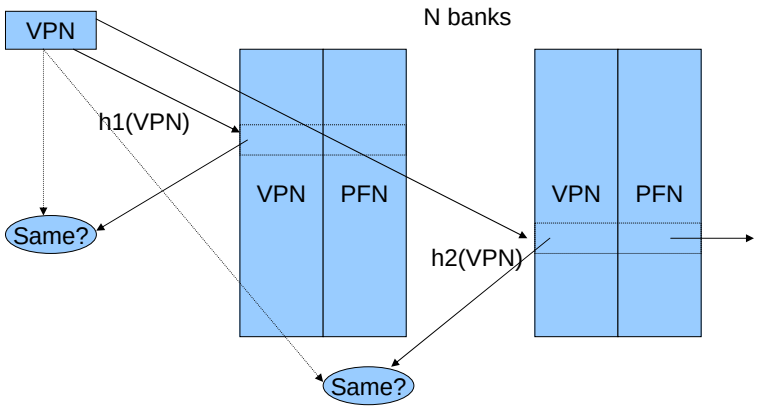
Copyright(C)2007 Youki Kadobayashi. All rights reserved.

Direct mapped TLB



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

N-way set associative TLB



Copyright(C)2007 Youki Kadobayashi. All rights reserved.

TLB secrets

- **テーブルサイズ** vs **ハッシュ衝突** (TLB **キャッシュミス**) **のトレードオフ**
- **小さなテーブルでも** 90% **のアドレス変換を高速化できる。**
- Q. **コンテクストスイッチすると** TLB **は作り直しか？**

Copyright(C)2007 Youki Kadobayashi. All rights reserved.

まとめ

- メモリ管理は OS における資源管理メカニズムのうち最も興味深いもの
- OS がプロセッサのアドレス変換機構を活用することで、さまざまな機能が実現される：
  - 広大なアドレス空間
  - メモリ階層を活かした、費用対効果の高いシステム
  - ソフトウェア障害の検出と隔離
  - プロセス間のセキュリティ
  - プロセス間のメモリ共有

次回

Copyright(C)2007 Youki Kadobayashi. All rights reserved.