# 計算機アーキテクチャ特論
# (Advanced Computer Architectures)

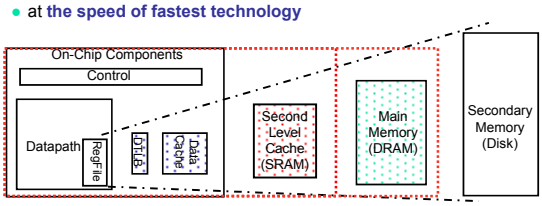
4. キャッシュの実装と評価

吉瀬 謙二 計算工学専攻
kise _at_ cs.titech.ac.jp www.arch.cs.titech.ac.jp
W831 講義室 木曜日 15:05 - 16:35



## Memory **Hierarchy**

- By taking advantage of **the principle of locality**（局所性）
  - Present **much memory** in **the cheapest technology**
  - at **the speed of fastest technology**

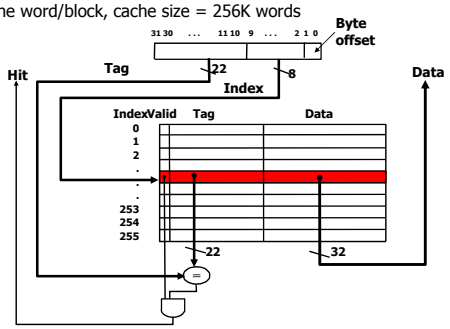



| **Speed (%cycles):** | ½'s | **1** | **9** | **90** |
|---|---|---|---|---|
| **Size (bytes):** | 100's | 1KB | 4KB | 512KB |



## 1KB L1 Data Cache

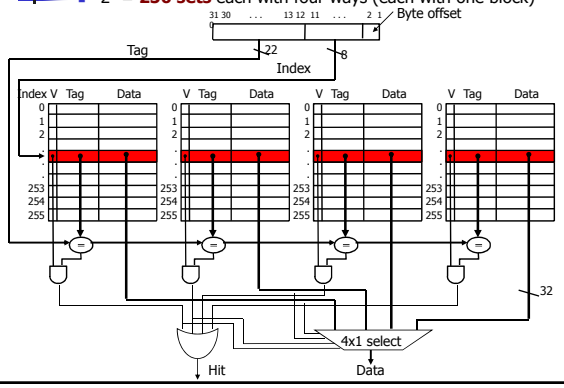
- 1KB cache
- One word/block, cache size = 256K words






## 4KB L2 Data Cache: 4-Way Set Associative Cache

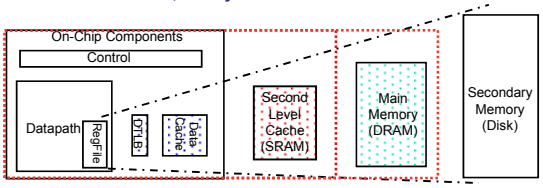
- $2^8$ = **256 sets** each with four ways (each with one block)






## Memory **Hierarchy**

- **L1 hit, 1 cycle for both LW/SW instruction execution**
- **L1 miss and L2 hit, 10 cycle for both LW/SW instruction ex.**
- **L1 miss and L2 miss, 100 cycle for both LW/SW instruction ex.**



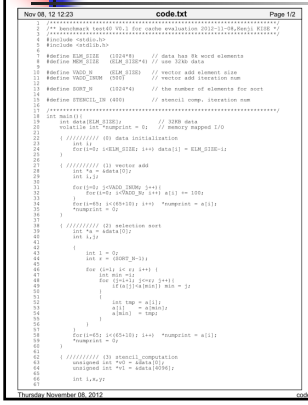

| **Speed (%cycles):** | ½'s | **1** | **9** | **90** |
|---|---|---|---|---|
| **Size (bytes):** | 100's | **1KB** | **4KB** | **512KB** |



## Sample code sequence, total of 229 cycles

```
lw     $2,80($fp)   # L1 hit, 1 cycle
nop                 # 1 cycle
sll    $2,$2,2      # 1 cycle
move   $3,$2        # 1 cycle
lw     $2,84($fp)   # L1 miss, L2 miss, 100 cycle
nop                 # 1 cycle
addu   $2,$3,$2     # 1 cycle
lw     $2,0($2)     # L1 miss, L2 hit, 10 cycle
nop                 # 1 cycle
addiu  $2,$2,100    # 1 cycle
sw     $2,0($4)     # L1 miss, L2 miss, 100 cycle
lw     $2,80($fp)   # L1 hit, 1 cycle
sw     $2,0($4)     # L1 miss, L2 hit, 10 cycle
```

## Benchmark program test40






## キャッシュに関する論文

- M.K. Qureshi, D. Thompson, Y.N. Patt.
  **The V-Way Cache: Demand Based Associativity via Global Replacement**.
  In Proc. of Int. Symp. Computer Architecture, ISCA 2005.
- Kaushik Rajan and R.Govindarajan.
  **Emulating optimal replacement with a shepherd cache**,
  In MICRO-40, pp. 445-454, 2007.

どちらか1つを選択して，プロセッサシミュレータ上に実装，性能評価する．



## レポート キャッシュの実装

1. SimCore/MIPS functional simulatorにキャッシュを追加して，そのヒット率とプロセッサ性能を評価する．
   1. ベンチマークプログラム 40bench を用いる．
   2. 命令は考慮しない，LW, SW命令で参照するデータキャッシュを対象．
   3. キャッシュの構成は講義スライドを参照

## レポート キャッシュの実装

1. 評価項目
   1. ベンチマークプログラムを実行した際の，ロード命令の実行回数，ストア命令の実行回数を測定してまとめること．メモリマップされたアドレス0へのロード，ストアはカウントしてはならない．
   2. 1KBのL1データキャッシュを実装し，そのミス率を測定せよ．また，L1キャッシュを追加することによるプロセッサ性能の向上率を示すこと．
      1. ライトアロケート方式（書き込みでミスした場合も，データをキャッシュに格納）とする．
   3. 4KBの標準的な4-way set associative方式のL2データキャッシュを実装し（L1データキャッシュも利用），そのミス率を測定せよ．また，L2キャッシュを追加することによるプロセッサ性能の向上率を示すこと．
   4. 講義で紹介した2つのキャッシュ方式から1つを選択し，その方式をL2キャッシュとして実装すること．ミス率を測定せよ．L2キャッシュを変更することによるプロセッサ性能の向上率を示すこと．標準的なL2より高い性能を実現すること．想定した性能が得られたか議論せよ．

## レポート 提出方法

- **12月13日の講義開始時に印刷したものを提出**．
- A4サイズで10ページ以内にまとめること．
  - 実装したキャッシュのコードを掲載すること．
  - 実装した各種の方式のミス率をグラフにして示すこと．
  - 実装したキャッシュの挙動は本当に正しいのか，どうして正しいと主張できるか説明すること．

## 講義計画

http://www.arch.cs.titech.ac.jp/sub5.html

- 導入：マイクロプロセッサ
- スーパースカラプロセッサの基礎と命令レベル並列性
- **キャッシュ**
- **分岐予測**
- 動的命令スケジューリングと投機処理
- メモリデータフローとデータキャッシュ
- 組込技術，低消費電力技術
- チップマルチプロセッサ
- オンチップネットワーク，メニーコアアーキテクチャ

- 【成績評価】レポートおよび，期末レポートにより評価する．