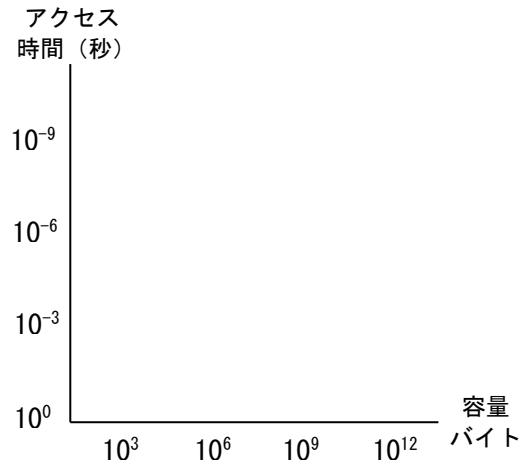


1) メモリ階層(記憶の階層)の考え方を、整理してみよう

ア) いろいろなメモリ素子技術を、アクセス速度と容量制約(大きさ・発熱・コスト)の点で比較してみよう。右図に素子技術を書き込んでみよう。



イ) メモリの速度と容量の間にトレードオフが存在することを、説明してみよう

.....

.....

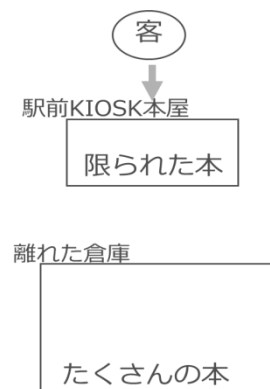
.....

2) キャッシュメモリの原理を、「駅前 KIOSK 本屋」のアナロジーで説明してみよう

ア) 本の置き方

.....

.....



イ) 客の求める本が KIOSK 内にあったときの動作

.....

.....

ウ) 客の求める本が KIOSK 内になかった時の動作

.....

.....

3) キャッシュメモリのアクセス時間のモデル式を立ててみよう

ア) ヒット率とは、.....

.....

イ) (手前の)キャッシュメモリにデータがあって、アクセスがキャッシュメモリまでで終わるときのアクセス時間 = T_c
 キャッシュメモリにはデータがなくて、(向こう側の)メインメモリまで取りに行くときのアクセス時間 = T_m とするとき、
 平均のメモリアクセス時間 T_{av} は、

.....

.....

4) せっかくキャッシュメモリを置くのだから、平均のメモリアクセス時間 T_{av} が、できればキャッシュメモリのみ T_c とあまり変わらないのがうれしい。(もし、 T_{av} がほとんど T_m と同じだったら、キャッシュメモリを置いた意味がない。) 仮に、 $T_m = 1$ ナノ秒 (= 10^{-9} 秒)、 $T_c = 0.1$ ナノ秒とするとき、 T_{av} が T_c の 1.5 倍 (5割増し) に収まるためには、ヒット率 h はいくらでなければならぬだろうか?

.....

.....

.....

.....

5) キャッシュメモリは容量が小さいので、メインメモリのデータのうち、ごく一部しか置かない。そうすると、キャッシュメモリのブロックは、メインメモリのいろいろなところのブロックが交代にやってきて置かれる。なんとなく、いろいろな家から出てきた車が、駐車場のスペースに駐車するのに似ている。駐車場の何番のブロックを使うかの選択方式に対応するのが、アドレスマッピングの方式である。いくつかの方式があるので、比較してみよう

ア) ダイレクトマッピング方式とは、どのようなやり方か？

イ) フルアソシアティブ方式とは、どのようなやり方か？

ウ) セットアソシアティブ方式とは、どのようなやり方か？

エ) ダイレクトマッピング方式と、フルアソシアティブ方式を、比較してみよう

	衝突の起こり具合、スペースの利用効率	アクセスに要する手間・時間
ダイレクトマッピング		
フルアソシアティブ		

6) キャッシュメモリは小さいので、一杯になると、すでに使われている場所を追い出して使わなければならない。その追い出しのやり方(アルゴリズム)には、いくつかのものがあるので、比較してみよう。

ア) FIFOとは

イ) OPTとは

ウ) LRUとは

エ) 追い出しのアルゴリズムの良し悪しは、ヒット率で評価できる。つまり、追い出したものがすぐまた必要とされて読み出されるようだと効率が悪いことになる。

アクセスされるブロックの番号の列を仮定し(参照列と呼ぶ)、その参照列に対して追い出しアルゴリズムを動作させて、どれだけヒット率が変わるか(どれだけキャッシュにない「ミス」になるか)を数えることができる。次の例で、実際にミスの数を数えてみよう。

キャッシュメモリは、フルアソシアティブ方式とする。(ダイレクトマッピングだとどこに置くか決まってい面白くない)

参照列: 0 ⇒ 1 ⇒ 2 ⇒ 3 ⇒ 4 ⇒ 0 ⇒ 1 ⇒ 2 ⇒ 5 ⇒ 0 ⇒ 1 ⇒ 2 ⇒ 3 ⇒ 4 ⇒ 5

キャッシュメモリ内のブロック数は、4つとする

FIFOの場合、右図を埋めてみよう

0	1	2	3	4	0	1	2	5	0	1	2	3	4	5
0	0	0	0											
	1	1	1											
		2	2											
			3											

LRUの場合、右図を埋めてみよう

0	1	2	3	4	0	1	2	5	0	1	2	3	4	5
0	0	0	0											
	1	1	1											
		2	2											
			3											