筑波大学計算科学研究センター CCS HPCサマーセミナー 「最適化1」

高橋大介
daisuke@cs.tsukuba.ac.jp
筑波大学大学院システム情報工学研究科
計算科学研究センター

講義内容

- 並列計算機システムの計算ノード単体におけるプログラムの最適化手法
  - レジスタブロッキング
  - キャッシュブロッキング
  - ストリーミングSIMD命令の活用
- 性能評価
  - ベンチマークプログラムの例

最適化

- 最適化の対象はいろいろある.
  - コード量の削減
  - データ量の削減
  - 実行時間の削減
- 今回は，実行時間を削減するためにプログラムを書き換えることを「最適化」と呼ぶことにする.

最適化の利点

- 最適化を行って実行時間を削減することにより，
  - 計算機の有効活用
  - 電気代（または課金）の削減
  - 同じ時間でより多くの計算ができる
- プログラムを書く時間＋実行時間の観点から考えると，長時間実行されるプログラムであるほど，最適化のメリットを享受できる.
  - 最適化によって性能が仮に3割向上したとすれば，それは，3割性能の高いマシンを使用しているのと同じことになる.
- 1回しか実行されず，かつ実行時間の短いプログラムは，最適化してもあまり意味がない.

最適化を行う前に

- そもそも，最適化を行う必要があるか？
- 現在用いているアルゴリズムは最適か？
- 効率の悪いアルゴリズムを最適化しても，意味がない.
  - バブルソートのプログラムを最適化しても，クイックソートよりは速くならない.
- 最適なアルゴリズムは
  - 解くべき問題の性質
  - 使おうとする計算機のアーキテクチャ，メモリ量など

  に大きく依存する

最適化の方針

- ベンダー提供の高速なライブラリが使える場合には，できるだけ使うようにする.
  - BLAS，LAPACKなど
- 最近のコンパイラの最適化能力は非常に高くなっている.
- コンパイラでもできる最適化は，ユーザー側では行わない.
  - 手間が掛かるだけ.
  - プログラムが複雑になりバグが入り込む余地が出てくる.
  - コンパイラの最適化能力を過信しない.
- 人間はアルゴリズムの改良に専念する.
- アセンブラはやむを得ない場合を除き，使わない.

最適化の第一歩

- まず，自分のプログラムでどの位の演算性能が出ているかを調べる.
- 演算性能の指標として，FLOPS（Floating Operations Per Second）がある.
  - 1秒間に実行可能な浮動小数点演算の回数を表す単位
  - MFLOPS（10^6），GFLOPS（10^9), TFLOPS（10^12）
- プログラム全体（または一部）の実行時間と，演算回数から，FLOPS値を算出し，プロセッサの理論ピーク性能と比較する.
  - Pentium4であれば，クロックの2倍のFLOPS値
  - Intel Core2であれば，クロックの4倍のFLOPS値

時間計測

- 時間計測を行う対象として
  - 経過時間（elapsed time）
  - CPU時間（CPU time）

  がある.
- 対象とするプログラムの実行時間が短い場合，タイマーの精度が足りない場合がある.
  - 何回か外側にループを回して測定する.
- この場合，コンパイラの最適化により，ループが回っていないことになる場合があるので注意する.
  - ダミールーチンを入れるか，測定対象をサブルーチンにして，分割コンパイルする.

ホットスポット

- 計算時間の大半を占有する部分を「ホットスポット」という.
- まず，どこがホットスポットかを調べる.
- 便利なツールとして，プロファイラがある.
  - Linuxではgprofコマンドが使える.
    - 「gcc –pg foo.c」のように，コンパイラオプションに「-pg」を付けることにより，gprofによって使用されるプロファイル情報を書き込む特別なコードが生成される.
  - a.outを実行し，その後にgprof a.outとすることで，ホットスポットを特定することができる.

gprofの出力例

Flat profile:

Each sample counts as 0.01 seconds.

% time	cumulative seconds	self seconds	calls	self s/call	total s/call	name
48.90	2.90	2.90	2	1.45	2.83	zfft1d0_
32.38	4.82	1.92	49152	0.00	0.00	fft8b_
14.17	5.66	0.84	16384	0.00	0.00	fft8a_
4.55	5.93	0.27	1	0.27	5.93	MAIN__
0.00	5.93	0.00	16384	0.00	0.00	fft235_
0.00	5.93	0.00	4	0.00	0.00	factor_
0.00	5.93	0.00	3	0.00	1.89	zfft1d_
0.00	5.93	0.00	2	0.00	0.00	settbl_
0.00	5.93	0.00	1	0.00	0.00	settbls_

gprofの結果から分かること

- ホットスポットは
  - zfft1d0_
  - fft8b_
  - fft8a_

  の3つであり，この3つで全実行時間の95%以上を消費している.
- これらのホットスポットのみに着目して最適化すればよい.
- プログラムを記述する際にはホットスポットが集中するように配慮する.
- ホットスポットが多くあると，コードの改良に手間が掛かる.
  - 最初からコードを書き直した方がましな場合もある.

コンパイルオプション

- コンパイルオプションの指定の仕方によって，性能が大きく変化する.
- コンパイラのマニュアルを参考に，いろんなコンパイルオプションを試してみる.
  - 「-fast」，「-O3」，「-O2」，など
  - Intel Compilerでは「-xP」（最新のXeon向け）
- 必ずしも最適化レベルを高くしたからといって，速いコードを出力するとは限らない.
  - コンパイラが余計な最適化を行う可能性があるため.
  - 計算結果が合わない場合もあるので注意する.

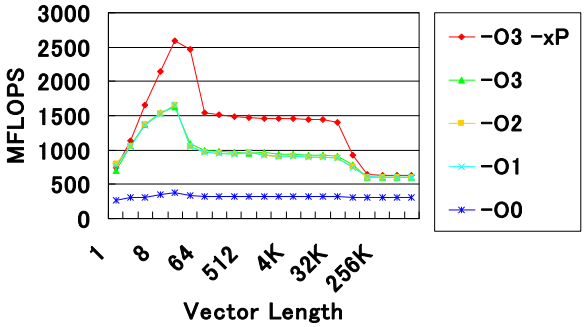
コンパイラディレクティブ

- コンパイラディレクティブ（指示行）は，コンパイラにプログラマの意図を伝え，最適化を支援する．
  - コンパイルオプションと違い，ループ単位で最適化をコントロールできる．
- ディレクティブの例
  - ベクトル化を行う際に，ループの依存性がないことをコンパイラに指示する．
  - ベクトル化の抑止
- C言語では「#pragma」，Fortranでは「!dir$」や「cpgi$l」などで記述することが多い．
  （コンパイラによって違うことがあるので注意）

Fortranで記述したZAXPY

```
      subroutine zaxpy(n,a,x,y)
      complex*16 a,x(*),y(*)
!dir$ vector aligned
      do i=1,n
         y(i)=y(i)+a*x(i)
      end do
      return
      end
```

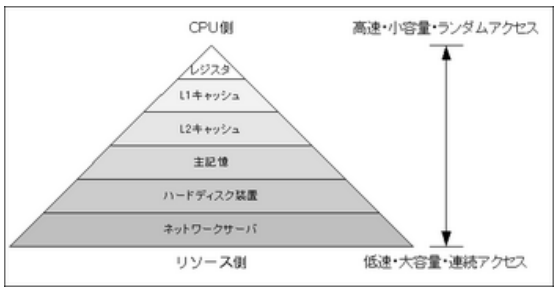

ZAXPYの性能
（Xeon 2.8GHz, 1CPU, Intel Fortran ）



プログラムを記述する際の注意点

- CやFortranの文法をきちんと守る．
  - コンパイラによっては，warningが出るだけのこともあるが，多くの場合バグの原因になる．
- コンパイラに依存する拡張機能は，やむを得ない場合（例えばディレクティブなど）を除き，できるだけ使わないようにする．
  - g77における自動割付配列
    - real*8 a(n)で，a(n)が仮引数でなく，かつnが変数のような場合
  - プログラムの移植性が悪くなる．
  - 思わぬエラーの原因となる．
- あまり使われていない（と思われる）関数や機能はなるべく使わないようにする．
  - コンパイラのバグが取りきれていない可能性がある．

記憶階層（1/4）



出典：Wikipedia

記憶階層（2/4）

- データを保持する記憶装置のコストバランス
  - 「小容量×高速≒大容量×低速」が成り立つ
- 「小容量×高速」記憶装置
  - レジスタ
- 「大容量×低速」記憶装置
  - ハードディスクや磁気テープ
- 「大容量×高速」はコストパフォーマンスが悪く実現困難

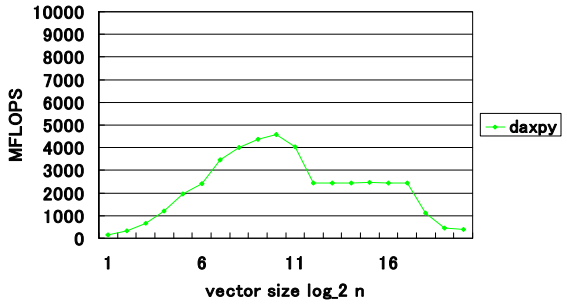
記憶階層(3/4)

- 記憶階層は記憶域に対するアクセスパターンの局所性(locality)を前提に設計されている.
- 局所性には
  - 時間的局所性
    - ある一定のアドレスに対するアクセスは，比較的近い時間内に再発するという性質
  - 空間的局所性
    - ある一定時間内にアクセスされるデータは，比較的近いアドレスに分布するという性質

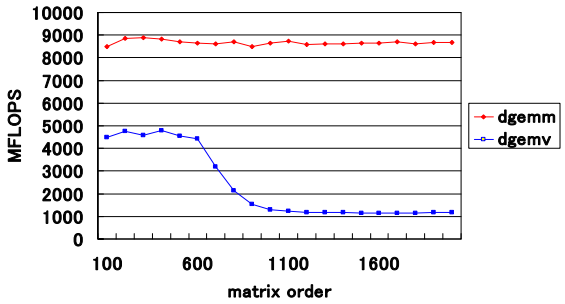
記憶階層(4/4)

- これらの傾向は，事務計算などの非数値計算には当てはまることが多いが，数値計算プログラムでは一般的ではない.
- 特に大規模な科学技術計算においては，データ参照に時間的局所性がないことが多い.
- これが，科学技術計算でベクトル型スーパーコンピュータが有利であった大きな理由.

BLASの性能(Woodcrest 2.4GHz 4MB L2 cache，Intel MKL 9.1)



BLASの性能(Woodcrest 2.4GHz 4MB L2 cache，Intel MKL 9.1)



BLASの演算回数

BLAS	ロード回数＋ストア回数	浮動小数点演算回数	比 $n=m=k$
Level 1 DAXPY $y=y+\alpha x$	$3n$	$2n$	$3:2$
Level 2 DGEMV $y=\beta y+\alpha Ax$	$mn+n+2m$	$2mn$	$1:2$
Level 3 DGEMM $C=\beta C+\alpha AB$	$2mn+mk+kn$	$2mnk$	$2:n$

Byte/Flopの概念(1/2)

- 1回の浮動小数点演算を行う際に必要なメモリアクセス量をByte/Flopで定義することができる.

```
subroutine daxpy(n, a, x, y)
real*8 a, x(*), y(*)
do i = 1, n
   y(i) = y(i) + a * x(i)
end do
```

- daxpyでは，1回のiterationにつき，2回の倍精度浮動小数点演算に対して3回の倍精度実数データ(合計24Byte)のload/storeが必要.
  - この場合，24Byte/2Flop = 12Byte/Flopとなる.
- Byte/Flop値は，小さいほど良い.

Byte/Flopの概念（2/2）

- デュアルコアのIntel Core2プロセッサ（3GHz）では，
  - 理論ピーク性能は12Gflops x 2コア=24Gflops
  - メモリバンド幅は最大21GB/s（4チャネルの場合）
  - Byte/Flop値は21/24=0.875
- daxpyでは，ワーキングセットがキャッシュの容量を超えた場合，メモリバンド幅（21GB/s）が律速となるので，21/12=1.75Gflops以上は出せない．
  - 理論ピーク性能のたった7%！
- メモリバンド幅が実効性能を左右することになる．

ループアンローリング（1/2）

- ループアンローリングとは，ループを展開することにより，
  - ループのオーバーヘッドを減らす
  - レジスタブロッキングを行う
- あまり展開し過ぎると，レジスタ不足や命令キャッシュミスを引き起こすので注意が必要．

```
double A[N], B[N], C;
for (i = 0; i < N; i++) {
  A[i] += B[i] * C;
}
```

⇒

```
double A[N], B[N], C;
for (i = 0; i < N; i += 4) {
  A[i] += B[i] * C;
  A[i+1] += B[i+1] * C;
  A[i+2] += B[i+2] * C;
  A[i+3] += B[i+3] * C;
}
```


ループアンローリング（2/2）

```
double A[N][N], B[N][N],
        C[N][N], s;
for (j = 0; j < N; k++) {
  for (i = 0; i < N; j++) {
    s = 0.0;
    for (k = 0; k < N; k++) {
      s += A[i][k] * B[j][k];
    }
    C[j][i] = s;
  }
}
```

行列積の例

⇒

```
double A[N][N], B[N][N],
        C[N][N], s0, s1;
for (j = 0; j < N; k += 2)
  for (i = 0; i < N; i++) {
    s0 = 0.0;  s1 = 0.0;
    for (k = 0; k < N; k++) {
      s0 += A[j][k] * B[j][k];
      s1 += A[j+1][k] * B[j][k];
    }
    C[j][i] = s0;
    C[j+1][i] = s1;
  }
```

行列積を最適化した例

ループの入れ換え

- ループの入れ換えは，主にストライドの大きなメモリ参照による悪影響を軽減する手法．
- コンパイラが判断して入れ換えてくれることもある．

```
double A[N][N], B[N][N], C;
for (j = 0; j < N; j++) {
  for (k = 0; k < N; k++) {
    A[k][j] += B[k][j] * C;
  }
}
```

ループ入れ替え前

⇒

```
double A[N][N], B[N][N], C;
for (k = 0; k < N; k++) {
  for (j = 0; j < N; j++) {
    A[k][j] += B[k][j] * C;
  }
}
```

ループ入れ替え後

パディング

- 複数の配列がキャッシュの同じ位置にマッピングされてしまい，スラッシングが生じる場合に有効．
  - 特にサイズが2のべきとなる配列の場合
- 二次元配列の定義サイズを少し変えてみる．
- コンパイルオプションを指定すると行ってくれるものもある．

```
double A[N][N], B[N][N];
for (k = 0; k< N; k++) {
  for (j = 0; j < N; j++) {
    A[j][k] = B[k][j];
  }
}
```

パディングを行う前

⇒

```
double A[N][N+1], B[N][N+1];
for (k = 0; k < N; k++) {
  for (j = 0; j < N; j++) {
    A[j][k] = B[k][j];
  }
}
```

パディングを行った後

ブロック化（1/2）

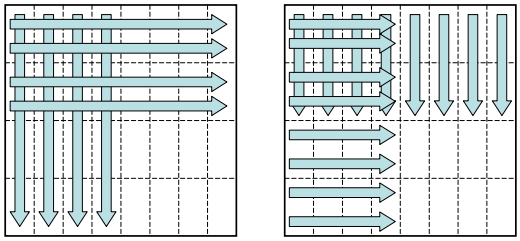
- メモリ参照を最適化するための有効な方法．
- キャッシュミスをできるだけ減らす．

```
double A[N][N], B[N][N], C;
for (i = 0; i < N; i++) {
  for (j = 0; j < N; j++) {
    A[i][j] += B[j][i] * C;
  }
}
```

⇒

```
double A[N][N], B[N][N], C;
for (i = 0; i < N; i += 4) {
  for (j = 0; j < N; j += 4) {
    for (ii = i; ii < i + 4; ii++) {
      for (jj = j; jj < j + 4; jj++) {
        A[ii][jj] += B[jj][i] * C;
      }
    }
  }
}
```

ブロック化（2/2）



ブロック化しない場合の　　ブロック化した場合の

メモリアクセスパターン　　メモリアクセスパターン

ストリーミングSIMD命令の活用

- 浮動小数点演算をより高速に処理するために，最近のプロセッサではストリーミングSIMD命令と呼ばれるものを搭載しているものが多い．
  - Intel Pentium4/XeonのSSE/SSE2/SSE3
  - AMD Athlonの3DNow!
  - Motorola PowerPCのAltiVec
- 最新のIntel Core2では，SSE3命令を活用することで，浮動小数点演算性能を4倍にすることができる．

Intel SSE3命令

- SSE2命令とは，Intel Pentium4/Xeonから導入された，x87命令に代わる新しい演算命令であるが，SSE3命令はSSE2命令に加えて，新たに13個の命令を付け加えたもの．
  - 128bit長のデータに対して，SIMD処理を行うことができる．
  - Intel Pentium4およびXeonプロセッサには128bitのXMMレジスタがXMM0～XMM7の8個搭載されている．
  - AMD Opteronプロセッサや，Xeon EM64Tプロセッサには XMM0～XMM15の16個搭載されている．
- SSE3のベクトル命令を用いることによって，64bitの倍精度浮動小数点演算ではベクトル長が2のベクトル演算（加減乗除，平方根，論理演算）を行うことができる．

SSE3命令の利用方法

- SSE3命令の利用方法としては，以下の方法が挙げられる．
  (1) コンパイラによりベクトル化する方法
  (2) SSE3組み込み関数を使用する方法
  (3) インラインアセンブラを使用する方法
  (4) アセンブラで「.s」ファイルを直接記述する方法
- (1)～(4)の順にコーディングが複雑になるが，性能という観点からは有利になる．

倍精度複素数の積和演算（a + b * c）をSSE3組み込み関数で記述した例

```
#include <pmmintrin.h>      /* SSE3命令を使う場合のヘッダファイル */

static __inline  __m128d ZMULADD(__m128d a, __m128d b, __m128d c)
{
  __m128d br, bi;                          /* 128bitのデータ型で変数を定義 */

  br = _mm_movedup_pd(b);                  /* br = [b.r b.r]  実部のみを取り出す*/
  br = _mm_mul_pd(br, c);                  /* br = [b.r*c.r b.r*c.i] */
  a = _mm_add_pd(a, br);                   /* a = [a.r+b.r*c.r a.i+b.r*c.i] */
  bi = _mm_unpackhi_pd(b, b);              /* bi = [b.i b.i]  虚部のみを取り出す */
  c = _mm_shuffle_pd(c, c, 1);             /* c = [c.i c.r]  実部と虚部を入れ替える */
  bi = _mm_mul_pd(bi, c);                  /* bi = [-b.i*c.i b.i*c.r] */

  return _mm_addsub_pd(a, bi);              /* [a.r+b.r*c.r-b.i*c.i a.i+b.r*c.i+b.i*c.r] */
}
```


Cで記述したZAXPY

```
typedef struct { double r, I; } doublecomplex;

void zaxpy(int n, doublecomplex a, doublecomplex *x, doublecomplex *y)
{
  int i;

  if (a.r == 0.0 && a.i == 0.0) return;

#pragma unroll(8)
#pragma vector aligned
  for (i = 0; i < n; i++) {
    y[i].r += a.r * x[i].r – a.i * x[i].i,
    y[i].i += a.r * x[i].i + a.i * x[i].r;
}
```

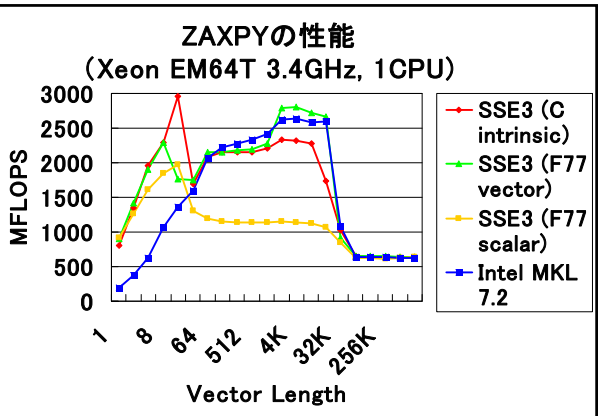
SSE3組み込み関数によるZAXPY

```
#include <pmmintrin.h>

typedef struct { double r, i; } doublecomplex;
__m128d ZMULADD(__m128d a, __m128d b, __m128d c);

void zaxpy(int n, doublecomplex a, doublecomplex *x, doublecomplex *y)
{
  int i;
  __m128d a0;

  if (a.r == 0.0 && a.i == 0.0) return;
  a0 = _mm_loadu_pd(&a);
#pragma unroll(8)
  for (i = 0; i < n; i++)
    _mm_store_pd(&y[i], ZMULADD(_mm_load_pd(&y[i]), a0, _mm_load_pd(&x[i])));
}
```

性能評価の目的（1/3）

- 計算機システムを実際に使ってみて，
  - 「性能が高いと思って使ってみたけれども，実際は思ったほど性能が出なかった」という経験はないか？
- これには，大きく分けて2つの理由がある．
  - 「性能が高い」といわれていたのは，その計算機システムの得意なある一面を指したものであり，ユーザが実行しようとした計算には向いていなかった．
  - 本当はその計算機システムは高い性能を秘めていたはずであるが，ユーザの使い方に問題があり，その高い性能を引き出せなかった．

性能評価の目的（2/3）

- 世の中に計算機が1種類しかなく，今後も進歩がないのであれば，「性能評価」はあまり必要がない．
- しかし，現実には世界に非常に多くのプロセッサや計算機システムが普及している．
- 自分の解決したい問題をどの計算機システムが効率よく計算してくれるか，ということをユーザ自身が判断する必要がある．
- また，計算機システムの開発者側からは，計算機の性能をハードウェアやソフトウェアの観点から改良していく際には，どうしても「汝自身」を知るために，「性能評価」を行い，改良に役立てる必要がある．

性能評価の目的（3/3）

- 性能評価を行うことで，
  - 計算機システムの性能がどの程度のもので，またどういった問題に向いているかということを知ることができる．
  - また，問題が大き過ぎて実行するのに非常に時間がかかる計算に要する時間を，実行する前に予測することができる．
- さらに，コストパフォーマンスの高い計算を行うには，計算機システムを使用する際のコストと，性能の両面からユーザは判断することになる．

性能評価の指標

- MIPS（Million Instructions Per Second）
  - CPUが1秒間に何百万回の命令を実行できるかということを表したもの．
  - MIPSはあくまでも命令実行回数であるので，アーキテクチャが異なるコンピュータ間の性能比較には適していない．
- FLOPS（Floating Operations Per Second）
  - 1秒間に実行可能な浮動小数点演算の回数を表す単位
  - MFLOPS，GFLOPS，TFLOPS
- SPEC（The Standard Performance Evaluation Corporation）
  - SPECベンチマークの値であり，SPECintが整数演算性能，SPECfpが浮動小数点演算性能である．

ベンチマークプログラムの例

- Dhrystone
- Whetstone
- Livermore Fortran Kernels
- LINPACK
- SPEC

各種ベンチマークの概要（1/4）

- Dhrystone
  - かなり昔から使われている整数演算性能を評価するベンチマークテストの1つ.
  - 複数の小さなループ処理を主体とし，通常はL1キャッシュに常駐させることが可能.
  - L1キャッシュを無限大と仮定した場合のプロセッサ性能を表わすことになり，L2キャッシュの容量やメモリアクセス性能によって特性はほとんど左右されない.
  - このような理由により，Dhrystoneベンチマークはシステム全体の性能を正確に表わすことができないので，最近はSPECベンチマークが使われることが多い.

各種ベンチマークの概要（2/4）

- Whetstone
  - かなり昔から使われている浮動小数点演算性能を評価するベンチマークテストの1つ.
  - Dhrystoneベンチマークと同様に，通常はループをL1キャッシュに常駐させることが可能である.
  - sin，cosなどの関数，整数および浮動小数点数の混合計算，分岐，スカラ変数などが総合的に測定される.
  - コンパイラやハードウェアだけではなく，数学関数ライブラリに重点を置いてテストする.

各種ベンチマークの概要（3/4）

- Livermore Fortran Kernels
  - 主にベクトル型のスーパーコンピュータで広く用いられてきたベンチマークテスト
  - 14個のカーネルからなる，リバモア14ループと，24個のカーネルからなる，リバモア24ループがある.
- LINPACK
  - テネシー大学のJack Dongarraによって開発された，浮動小数点演算能力を評価するためのベンチマークテスト.
  - ガウス消去法を用いて連立一次方程式の解を求めるのに要する時間を測定.
  - 「TOP500 Supercomputer」のベンチマークにも用いられている.

各種ベンチマークの概要（4/4）

- SPEC（Standard Performance Evaluation Corporation）
  - 主要なベンダーが資金提供している非営利団体
  - http://www.spec.orgで測定結果を公表している
- SPEC CPU2006：CPU，メモリ，コンパイラの総合的な性能評価
  - CINT2006（SPECint）：整数演算を評価する
  - CFP2006（SPECfp）：浮動小数点演算を評価する
- 他にも，SPECweb2005やSPECjvm98などがある.

まとめ

- 実行時間を短縮するために，最適化を行うことは重要.
  - ただ，本当に最適化が必要なケースであるかどうかは判断する必要がある.
- メモリバンド幅が律速にならないように最適化を行うことが今後のプロセッサでは重要になる.
- 性能評価は，これから使おうとしているコンピュータの性能を事前に知るうえで有効.