

Knights Landing (開発コード名) での Stream Triad のメモリー帯域幅の最適化

この記事は、インテル® デベロッパー・ゾーンに公開されている「[Optimizing Memory Bandwidth in Knights Landing on Stream Triad](#)」の日本語参考訳です。

概要

この記事は、インテル® Xeon Phi™ プロセッサ (開発コード名 Knights Landing) でピークメモリー帯域幅を評価する最適な手法を示します。ここでは、メモリー帯域幅を評価する業界標準である STREAM ベンチマークを使用します。

はじめに

STREAM ベンチマークは、4 つの単純なベクトルカーネル (Copy、Scale、Add、Triad) のメモリー帯域幅 (MB/秒) の制限と計算速度を評価するための合成ベンチマークです。ソースコードは <http://www.cs.virginia.edu/stream/> (英語) から入手できます。STREAM は、HPC チャレンジ・ベンチマーク (HPCC) にも含まれています。

STREAM の使用条件

STREAM では、各配列は、実行に使用されるすべての最終レベルキャッシュ (LLC) の合計サイズの 4 倍以上か、100 万要素以上のいずれか大きいほうでなければなりません。

標準バージョンとチューニング済みバージョン

「Standard (標準)」バージョンの結果は、C または Fortran コードを倍精度データ型を使用してプロダクション・ハードウェア上で実行したものです。「Tuned (チューニング済み)」バージョンの結果は、倍精度データ型を使用して実行したのですが、(アセンブリ言語コードを含む) コードが変更されています。この記事では、標準バージョンの結果を使用します。

新しい MCDRAM メモリーの帯域幅を評価する場合、必要に応じてコードを変更することができます。特定のメモリーモードでは、MCDRAM は個別の NUMA ドメインとして認識されるため、(numactl のように) 実行時に標準の NUMA API で評価することができます。

Knights Landing (開発コード名) のメモリー・アーキテクチャーの概要

Knights Landing (開発コード名) には、DDR4 と MCDRAM の 2 種類のメモリーが搭載されています。

- DDR4 は、低帯域幅で大容量のメモリーです。
- MCDRAM は、高帯域幅で低容量 (最大 16GB) のメモリーで、Knights Landing (開発コード名) シリコンに統合されています。

MCDRAM は、3 次キャッシュ (メモリー・サイド・キャッシュ)、個別の NUMA ノード、またはこれらの組み合わせとして構成可能で、キャッシュ、フラット、ハイブリッドのいずれかのモードで利用できます。この記事では、フラットモードの DDR と MCDRAM でピークメモリー帯域幅を評価する手法を示します (ここで、DDR と MCDRAM は、別々の NUMA ノードとして認識される、個別のアドレス指定可能なメモリーです)。

ソース変更

メモリー割り当てを 2MB 境界でアライメントします。

静的割り当ての例:

```
static double a[N+OFFSET] __attribute__((aligned(2097152)));
```

動的割り当ての例:

```
a = (STREAM_TYPE  
*)_mm_malloc(sizeof(STREAM_TYPE)*(STREAM_ARRAY_SIZE+OFFSET), 2097152);
```

上記を繰り返して、STREAM ベンチマークの 3 つの配列 a、b、c を割り当てます。

使用するコンパイラー・オプション (インテル® コンパイラー)

```
-mcmmodel medium -shared-intel -O3 -xMIC-AVX512 -DSTREAM_ARRAY_SIZE=134217728  
-DOFFSET=0 -DNTIMES=10 -qopenmp -qopt-streaming-stores always
```

注: STREAM の使用条件を満たす 1GB の配列を使用します。

メモリーモード: DDR4 + MCDRAM フラット (アドレス指定可能)

- フラットモードでは、DDR4 と MCDRAM を個別にアドレス指定可能なメモリーとして利用できます。
- 同じアプリケーションで DDR4 と MCDRAM を使用する場合、ソフトウェアの変更が必要になります。
- アプリケーションのメモリー割り当てを DDR4 または MCDRAM のいずれかに設定するには、NUMA API を使用します。
 - この場合、ソフトウェアの変更は不要です。メモリー・フットプリントが MCDRAM メモリーに収まる場合、ユーザーは、NUMA API を使用してアプリケーション全体を MCDRAM で実行することができます。
- DDR4 と MCDRAM がフラットモードなので、DDR と MCDRAM の帯域幅を個別に評価する必要があります。

DDR4 の帯域幅を評価する

- インテル® Xeon Phi™ プロセッサー 7250 上で次のコマンドを実行します。

```
export OMP_NUM_THREADS=68; export KMP_AFFINITY=scatter; numactl -m  
0 ./stream
```

システムで利用可能な NUMA ノードを特定するには、numactl -H コマンドを使用します。

MCDRAM の帯域幅を評価する (1 つ目の手法)

NUMA API を使用して実行時にすべてのメモリーを MCDRAM に割り当てて、フラットモードで MCDRAM の帯域幅を評価します。ソースの変更は不要です。

- インテル® Xeon Phi™ プロセッサ 7250 上で次のコマンドを実行します。

```
export OMP_NUM_THREADS=68; export KMP_AFFINITY=scatter; numactl -m
1 ./stream
```

MCDRAM の帯域幅を評価する (2 つ目の手法)

ソースを変更して、MEMKING/HBW API を使用します。

以下は、STREAM の標準のメモリー割り当てです。

```
static STREAM_TYPE a[STREAM_ARRAY_SIZE+OFFSET],
b[STREAM_ARRAY_SIZE+OFFSET],
c[STREAM_ARRAY_SIZE+OFFSET];
```

例えば、次のように MCDRAM (HBM) にメモリーを割り当てます。

```
hbw_posix_memalign ((void **)&a, 2097152, sizeof(STREAM_TYPE) *
(STREAM_ARRAY_SIZE + OFFSET)
hbw_posix_memalign ((void **)&b, 2097152, sizeof(STREAM_TYPE) *
(STREAM_ARRAY_SIZE + OFFSET)
hbw_posix_memalign ((void **)&c, 2097152, sizeof(STREAM_TYPE) *
(STREAM_ARRAY_SIZE + OFFSET)
```

上記のように、hbw_posix_memalign(void **memptr, size_t alignment, size_t size) 構文を使用します。

- 次のコマンドを実行します。

```
export OMP_NUM_THREADS=68; export KMP_AFFINITY=scatter; ./stream_hbw_malloc
```

まとめ

MCDRAM は Knights Landing (開発コード名) で必要な高帯域幅を提供します。メモリー構成ごとに帯域幅が異なるため、それぞれのモードで評価する必要があります。

インテル® Xeon Phi™ プロセッサ 7250 で予測される STREAM Triad のメモリー帯域幅 (GB/秒):

クラスターモード: クアドラント (4 分割)、メモリーモード: DDR + MCDRAM フラット

MCDRAM フラット	DDR4 フラット
約 475 ~ 490GB/秒	約 90GB/秒

コンパイラーの最適化に関する詳細は、[最適化に関する注意事項](#)を参照してください。