

Intel CPU の行方

戸松 祐太, 千田 智治
Yuta TOMATSU, Tomoharu SENDA

1 はじめに

近年, 市場が CPU に求めるニーズは, サーバ, 一般 PC 用途などの高性能 CPU や, モバイル機器用途の省エネルギー CPU である.

前者のニーズに応えるために, Intel ではクロック増加の限界により, シングルコア CPU から 1 つの CPU に複数の CPU コアを実装したマルチコア CPU に力を入れてきた. 現在, CPC コア数を増やす事と, MCM (Multi-Chip Module) と呼ばれる 1 個の CPU に他のチップセットを実装するマルチコア化も進めている. また, 一般 PC 向けに CPU コア自身の性能を上げる命令セットの新しい拡張にも力を入れている.

省エネルギー CPU に関しては, 今年の春に CPU 「Atom」が発表され, 近々発売される予定である.

本稿では, はじめにモバイル系 CPU 「Atom」の消費電力に関して述べ, その後にマルチコア化と命令セットの拡張による Intel の新しい技術について述べ, 今後の推察を行う.

2 Atom

Fig. 1 に Intel のロードマップを示す. Fig. 1 で括弧で囲まれている名前はファミリーネームを示している.

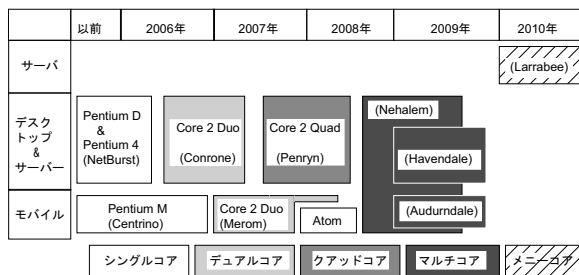


Fig.1 Intel のロードマップ (出典: 自作)

Atom¹⁾ は今年の春に発売予定である携帯電話やモバイル系の低コストノート PC などの携帯機器向けの CPU である. 新しく設計した 45nm の CPU で 2GHz 近くのクロック数で動作する. また, OpenGL や DX9L に対応する高いグラフィックス性能があり, flash や JavaScript にも対応する. 命令セットは Core2Duo の命令セットと互換性をもっているため, Windows 系や Linux などの OS を動かすことが出来る. また, Atom の最大の特徴はその消費電力の低さである. 携帯機器向けであるため, 消費電力は今までの Intel CPU の中で一番低くなっている.

2.1 消費電力

Atom はモバイル機器に適切な 1W 未満~2.5W の熱電力設計 (TDP) 仕様になっている. また, アイドル状態の省電力は約 0.3W に抑え, モバイル機器の電源を常時オンにしても消費電力を抑制することが可能である.

2.1.1 アイドル状態の段階的な Sleep

Atom CPU がアイドル状態になると, 段階を得て Sleep 状態になる.

第一段階は CPU がアイドル状態に入ると CPU コアは停止する「Stop Clock」という段階である. この段階では 1 μ s 以下のレイテンシで復帰できるが, 電圧はあまり下がらない.

次の段階では, 1L キャッシュのデータを 2L キャッシュにコピーし, 1L キャッシュ内のデータは削除される「Deep Sleep」という段階である. この段階では, CPU コアと 1L キャッシュとの情報伝達であるスヌープの必要がなくなる. この間の電圧は 0.3~0.6W の間になる.

まだアイドル状態が続くと, CPU 供給電圧がキャッシュのデータ内容を維持できる限界レベルまで下がる. この段階を「Deeper Down」と言い, 2L キャッシュ内のデータはメインメモリに書き戻され, 回路設計技術で動的にキャッシュの容量を変更する「Dynamic Smart Cache Sizing」の技法によりキャッシュメモリの SRAM をオフにしていく. 最終的には 2L は完全に空になり, 完全にパワーオフにされる. この段階からの復帰は 30 μ s かかる.

Deeper Down の次の段階が「Deep Power Down」である. この段階では, CPU コア内の命令フェッチやメモリリード・ライトなどの情報である CPU ステイトを別の SRAM に記憶させ, CPU コアには電力が供給されなくなる. CPU ステイトを記憶させている SRAM は, 「State Storage」と呼ばれ CPU 内に実装されている. この State Storage には独自の電源供給口がある. CPU 電源電圧がどこまで下がっても, この独立した電源供給口により State Storage の記憶が消えることがない. CPU 内に実装されていることにより, 復帰の際 CPU ステイトを比較的早く CPU に転送することができる.

2.1.2 その他の技法による省電力化

Atom ではプロセス技術やマイクロアーキテクチャ的に消費電力化を図っている.

プロセス技術では 45nm CMOS と呼ばれる技法が用いられている. これは大量の電流を流せる高誘電率な素材を用いた絶縁膜である High-k ゲート酸化膜に使うことにより, 大幅にリーク電流を低減している. また, 全て

のトランジスタが低速だがリーク電流が小さい。

マイクロアーキテクチャ的にはCPUコアには4,700万個のトランジスタを搭載させている。これはCore2DuoのCPUコア数の半分のトランジスタ数である。行ユニット部分ではトランジスタ数が少なくてすむ整数演算については128-bit分のSIMD演算ユニットを搭載しているが、トランジスタ数が多くなる浮動小数点演算については64-bit分のSIMD演算ユニットとなっている。

3 マルチコア化とメニーコア化

Intelでは今まで、CPUコアを2つ実装したデュアルコアや、4つ実装したクアッドコアなど、1つのCPUに数個のコアの実装を目指してきた。1つのCPUに1~9個のCPUコアや他のコアを実装したものをマルチコアCPUと呼ぶ。また、IntelではマルチコアCPUの延長線上に、10個以上のCPUコアを1つのCPUに実装することにも力を入れている。そのCPUをメニーコアCPUと呼ぶ。

Intelでは、マルチコアCPUとしては数個のCPUコアを1つのCPUに実装させたNehalemや、GPUコアとCPUコアを1つのCPUに実装したHavendaleを出す予定である。また、メニーコアには数十個のCPUコアを実装するLarrabeeを開発中である。また、ここで示している「Nehalem」「Havendale」「Larrabee」はファミリーネームであり、実際に商品化するときは違う名前にあると思う。

3.1 Nehalem

Nehalemとは、Intelが2008年の後半に生産開始としているマルチコアCPUである。このNehalemでは1つのCPUに実装されるCPUコア数は1~8コアまでのバリエーションがあり、市場の要望の異なるコア数のCPUを購入する事が出来る。

Nehalemのアーキテクチャは、Core Microarchitectureを基本にして様々な技術を用いている。その技術は3階層キャッシュとメモリコントローラの実装、ハイパースレディングテクノロジーの採用とQPI(QuickPath Interconnect)の実装である。

3.1.1 3階層のキャッシュとメモリコントローラの実装

Nehalemでは、各CPUコアにアクセスレイテンシが短い1Lキャッシュと2Lキャッシュが結合されており、3Lキャッシュはレイテンシは長いが各コアがアクセス出来る大容量の共有キャッシュとなっている。Fig. 2に、デュアルコア版のNehalemのアーキテクチャを示す。一般的にキャッシュの容量の増加と、1つのキャッシュにアクセスするコア数が増える事により、レイテンシが伸びる。Nehalemの3階層構造はこの大容量の3Lキャッシュと1Lキャッシュの間のレイテンシを埋める形として、2Lキャッシュを挟んでいる。これにより、2階層キャッシュよりもレイテンシを減少することが出来る。また、この3階層キャッシュ構造はCPUコア数が4コア以上の場合実装する。

また、従来ではCPUとメモリのアクセスをバスを通

して行っていた事を、メモリコントローラの実装により、CPUとメモリが直接アクセスすることが可能となった。これにより、メモリへのアクセスが高速になる。CPUからのメモリへのアクセス回数の減少とメモリアクセスの高速化により、メモリアクセスレイテンシは劇的に減少する。

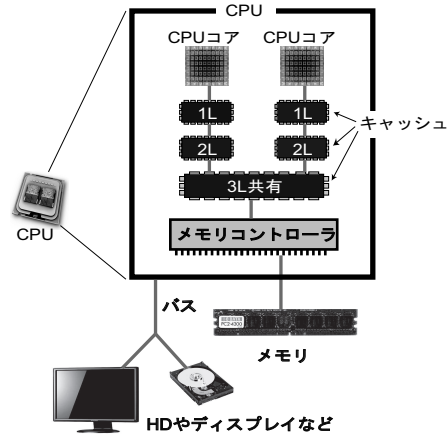


Fig.2 Nehalemのアーキテクチャ(文献²⁾より参照)

3.1.2 ハイパースレディングテクノロジーの採用

ハイパースレディングテクノロジーとは、1つのCPUコアのレジスタやパイプライン回路の空き時間を利用して、1つのCPUコアに2つのスレッドを実行させる技術である。Nehalemでは全てのコアにこの技術が施されているので、8コアのCPUでは16スレッドを実行させることが可能である。これにより、大量のデータを並列実行するスーパーコンピューティング分野や、多種多様なタスクを同時に実行するサーバなどの処理を高速に行うことが出来る。

3.1.3 QPI(QuickPath Interconnect)の実装

QPIは、CPUとチップセットの間とCPU同士のダイレクトコネクタに使われる。Core2Duoなどに使われているFSBより転送レートが約4倍になっており最高6.4Gtpsの高転送レートが可能となっている。また、QPIの帯域がFSBの約2倍の2.5GB/secとなっている。この帯域により、CPU同士を直接接続させる時、メモリからの読み出しが早い低レイテンシで柔軟なマルチプロセッサ構成が実現できる。

3.2 Havendale/Audurndale

HavendaleとAudurndaleは2009年に製造開始になるGPU統合型CPUであり、前者がデスクトップ用CPUで後者がモバイル用CPUである。この2つのCPUは、デュアルコア版のNehalemとグラフィックス統合チップセットGMCH(Graphics Memory Controller Hub)を1つのCPUに収めたMCM(Multi-Chip Module)である。

CPUコアやキャッシュメモリ等の接続部から構成されているCPUダイに用いられている技術はNehalemとほぼ同様である。しかし、キャッシュ階層は4MBの共有

キャッシュをもった3階層より浅い。CPU コアは2つで、3階層キャッシュにしても2階層キャッシュよりレイテンシを小さくしにくいからである。また、CPU ダイにはメモリコントローラとPCIデバイスと接続できるPCI Express Genが入っており、2個のCPU コアと共有キャッシュ、GMCH とのQPI という構成である。また、GMCH ダイにはCPU ダイに無かったメモリコントローラとパソコン内部の各パーツ間を結ぶバスであるPCI Express Genが搭載されている。Fig. 3に、Havendaleのアーキテクチャを示す。

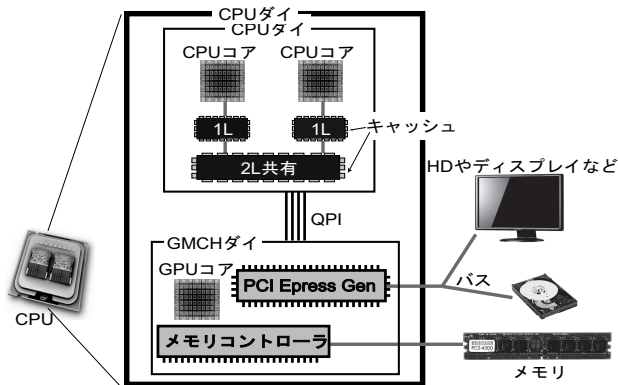


Fig.3 Havendaleのアーキテクチャ（文献²⁾より参照）

メモリコントローラをGMCHダイに入れている理由のため、発熱が多いGMCHのサイズを小さくし過ぎると電圧密度が高まり冷却が難しくなる。このことにより、メモリコントローラをCPUダイに含めず、サイズを小さくし、GMCHダイを従来通りの大きさに実装させている。

GPUコアとCPUコアの統合は、2つのコアをQPIで結合させていることによりグラフィックス処理をCPUコアとGPUコアで高速並列処理をすることが出来るという利点がある。また、GPUコアはグラフィックス処理に特化しているが、演算処理も行えるためグラフィックス処理以外の処理にも活用できる。つまり、ハイパフォーマンスコンピューティングの分野にも適応できるといった柔軟性をもっている。

3.3 Larrabee

Larrabeeは、2008年末～2009年に試作製品を登場する予定の最初のメニーコアCPUである。搭載するコア数は32コアと48コアである。このCPUコアは浮動小数点演算に特化したインオーダー実行プロセッサである。インオーダーで浮動小数点演算に特化すると、CPUコアサイズは今までのCPUコアの数分の1にする事が可能である。また、ハイパースレッディングテクノロジーを改良して、1つのCPUコアで4スレッド実行できる。このため、3Dグラフィックスアプリケーション向けに数百万個ものピクセルを演算してモニタにレンダリングする動作を毎秒数百回実行するような多くのデータを並列に処

理するストリームコンピューティングが得意とするCPUである。

Larrabeeのコアの命令セットは、高並列ワークロードのために特化した命令セットに拡張されている。また、コアにまたがる形でキャッシュが正しいデータにアクセスできるキャッシュコヒーレンシをとる共有キャッシュを搭載している。

Larrabeeは汎用CPUであるが、命令セットの拡張によって特定のワークロードに対応することができる。そのため、IA (Intel Architecter) 命令セットアーキテクチャの互換をとる(x86アーキテクチャの拡張)。この互換より、既存のソフトウェア、ライブラリなどを継続して使用することが可能である。しかし、既存のソフトがLarrabeeを効率よく使うことが出来るかは疑問が残る。この理由を次節で説明する。

3.4 マルチコア化の問題

現在一般向けPCのCPUのCPUコア数を増やしてもCPU性能の向上につながるとは限らない。なぜなら、アムダールの法則が関わってくるからである。アムダールの法則とは、逐次実行するしかない部分がタスク全体の割合が大きいと、CPUに多くのCPUコアを実装させてもCPU性能が上がらなくなるというものである。一般家庭や個人が使用しているソフトウェアの多くがマルチスレッド対応ではない。この現状でCPUに多くのコアを実装しても、ソフトウェアが追いついていないことにより、CPUの性能を持て余しマルチ・メニーコア化による一般PC向けのCPUの高速化は難しいとされている。

4 Intelの命令拡張

IntelはCPUに複数のコアを実装することによりCPUの高速化を図る他に、命令セットを見直し、新しく拡張することによりCPUコア自体の性能を上げようとしている。Intelがこのような多くのソフトウェアがマルチスレッド対応でないために、一般PC向けのCPUをマルチコア化しても性能が上がらないことに起因する。本章では従来の命令セットの拡張の欠点と、その欠点を補う新拡張のAVX(Advanced Vector Extensions)について記述する。

4.1 従来の命令セットの欠点

今主流になっている命令セットはx86アーキテクチャである。x86命令セットでは、命令中のオペコードの前に1バイト長のプリフィックスをつけることにより、新たに命令を加えたり、レジスタを拡張することができる。しかし、このプリフィックスにより命令フォーマットが複雑になり命令長が長くなる。これにより命令デコードが難しくなり処理に時間がかかることや、命令デコードが肥大化して電力を消費することが、CPUの性能を低下させてきた。

また命令レベルでの並列性であるILP (Instruction Level Parallelism) 技法の向上を図ろうとしてハードウェアを追加しても、電力消費が増え、またハードウェア

アが複雑になるだけで、CPU 性能が向上することが困難となっている。上記の理由のために Intel は新しい命令セットの拡張を行おうとしている。

4.2 AVX

AVX(Advanced Vector Extensions)³⁾ とは Intel が 2010 年以降の CPU に搭載させる予定の拡張した新しい命令セットアーキテクチャである。AVX は SIMD (Single Instruction Multiple Data) 型の浮動小数点演算の性能に集中した拡張命令セットである。そのため、従来の x86 命令セットと SIMD 演算の命令セットである SSE を変更させる。以下にその変更点をあげる。

- SIMD 長を伸ばす (128bits から 256bits へ)
- 命令エンコーディングフォーマットを変える
- 積和算を実装する

SIMD 長を伸ばすと言うことは、SIMD 演算ユニットの演算幅が増大し、理論値では CPU コアあたりの浮動小数点演算が 2 倍になる。また、新しい命令フォーマットでは、1 バイトの「VEX」と呼ぶプリフィックスとそれに続く 1~2 バイトの「ペイロード」を導入する。これにより、従来のプリフィックスで表現していた情報をペイロードのビットフィールドの中で表現することが可能になり、命令長を短くすることが出来る。積和算を実装することにより、積算と加算を 1 命令で出来るので計算の速度が上がる。以上より、CPU の性能は大幅に上がると考えられる。

しかし、この AVX の搭載の CPU を効率よく扱うためには、SIMD の浮動小数点演算を多様するソフトウェアになる。既存のソフトウェアには整数演算中心が多いため、AVX の登場によりソフトウェアの変更が必要になる。

5 今後の展望

Intel CPU の今後の展望として、CPU の高性能化に伴うソフトウェアの変化に対する Intel の対応と、CPU に搭載されるチップの増大の 2 点について述べる

- ・ CPU の高性能化に伴うソフトウェアの変化

Intel は、CPU コアの性能アップとメニーコア化を並列に研究していく方針である。6 年後にはこの 2 つの技術により CPU は飛躍的に性能が向上する。しかし、その CPU を効率よく扱うにはソフトウェアが「マルチスレッド対応のソフトウェア」かつ「浮動小数点中心のソフトウェア」である必要がある。前者に対して Intel はメニーコアの CPU に活用できる新しいプログラミング言語「Baker」を開発しており、後者に対してはプログラマーの設計に係っている。今後 Intel は、CPU 自体の性能向上と平行して、その CPU を使いこなすソフトウェアの開発のサポートや言語などサービス面でも力を入れなければならないと考えられる。

- ・ CPU に搭載されるチップの増大

Intel は、GPU やメモリコントローラなど本来 CPU

の外側にあったアーキテクチャを CPU と統合させる。この先、Intel は CPU に他のアーキテクチャを実装させていくと思われる。Intel 自身が開発している他のアーキテクチャとして、RAID コントローラや通信関連のチップセットがある。Intel はサーバやスーパーコンピュータ向けの CPU も開発しているので、大容量の記憶が可能な RAID の制御や、クラスタに用いる通信を制御するチップが 1 つの CPU に搭載すると思われる。そのことにより、RAID や通信コネクタと CPU を直接接続することが可能になり、CPU の性能が向上すると考えられるからである。また、RAID 制御のチップの搭載のように、CPU の処理よりも遅いアーキテクチャとの直接結合が生じる。そのためレイテンシが伸びることが予想される為、CPU 内のキャッシュ構造が多階層になると考えられる。今後の CPU は演算処理だけでなく、多種多彩な処理ができるようになると思われる。

6 まとめ

市場は省電力モバイル CPU と、高性能 CPU の 2 つのニーズがある。CPU の高性能のニーズに応えるために、高クロック数の向上の限界によりシングルコアから、Nehalem と呼ばれる 1 つの CPU に数個の CPU コアを実装させるマルチコア CPU に変えてきた。また、モバイル機器が普及したことによりマルチコア化と並列に消費電力が低い CPU である Atom の開発も進めてきた。マルチコア CPU には CPU コア以外に GPU コアを実装させた Havendale の開発が進んでいる。Intel では、ハイパフォーマンスコンピューティング向けのマルチコアの先の 10 コア以上のメニーコアまでを目指している。また、一般 PC 向けには命令セットの新拡張により CPU コア自身の性能向上化も目指している。

現在、Intel ではマルチ・メニーコア化と命令セットの新拡張を並列に開発することにより、CPU の性能向上を図ろうとしている。

参考文献

- 1) Intel - インテル® Atom ブランド発表
<http://www.intel.co.jp/jp/intel/pr/press2008/080303.htm>
- 2) Intel のマイクロアーキテクチャ改革「Nehalem」
<http://pc.watch.impress.co.jp/docs/2007/0330/kaigai348.htm>
- 3) x86 からの脱却を図る Intel の新ロードマップ
<http://pc.watch.impress.co.jp/docs/2008/0407/kaigai434.htm>