

先進計算機構成論 導入

東京大学大学院 情報理工学系研究科 創造情報学専攻

塩谷 亮太

shioya@ci.i.u-tokyo.ac.jp

自己紹介

■ 塩谷 亮太（しおや りょうた）

◇ 創造情報学専攻 准教授

■ 経歴：

◇ 2011 年：情報理工 電情 博士取得

□ 当時は 坂井・五島研（現 入江・門本研）でした

◇ 2011 年：名古屋大学 助教

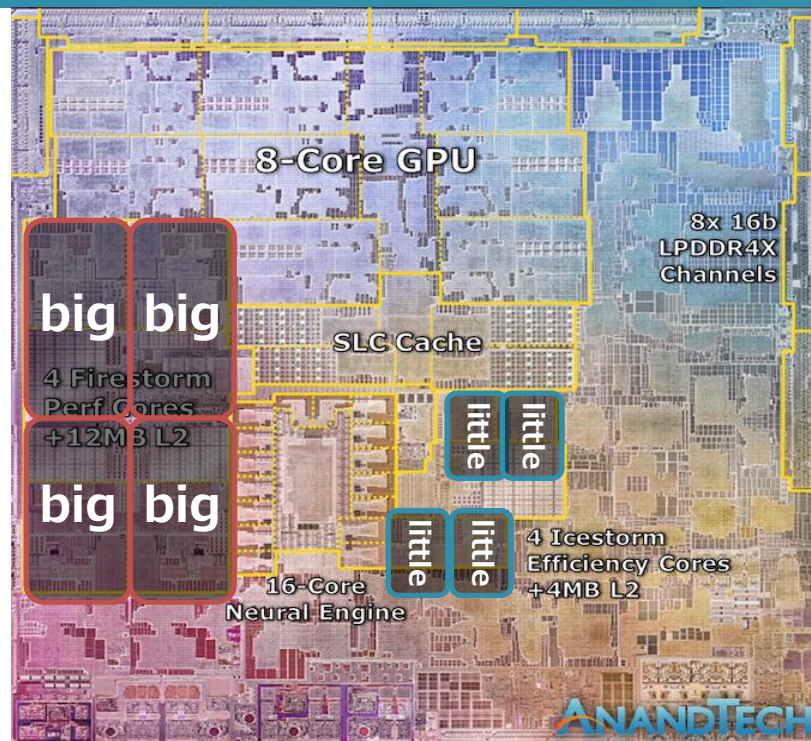
◇ 2018 年：情報理工 創造情報学専攻 准教授

ちょっと考えてみよう

1. スーパーコンピュータとパソコンは何が違うのか？
 - ◇ スーパーコンピュータはパソコンの上位互換なのか？
2. コンピュータのプログラムを高速に実行できるとはどういうことなのか？
3. なぜ GPU は CPU よりグラフィックや機械学習で速いのか？
 - ◇ そもそも「専用」のマシンだとなぜ速いのか？
4. 大きくて速いキャッシュは作れないのか？
5. 物理的な量がなさそうな「情報」を処理してるのに、なぜコンピュータは電力を消費するのか？

Apple M1

写真は <https://www.anandtech.com/show/16226/apple-silicon-m1-a14-deep-dive> より



- Apple M1（2020年11月発売の MacBook Pro に搭載）
 - ◇ big core (firestorm) と little core (icestorm) の2種類を持つ
 - ◇ チップ面積の大部分は GPU や様々な専用回路に割かれている
- big や little, あるいは GPU が大きいとはどういう事なのか？
 - ◇ 大きいと何が違うのか？ なぜサイズが異なるのか？

- この講義では,
 - ◇ この手のコンピュータに関する疑問に
ざっくり答えられるようになることを目指します



- 学部とかでコンピュータのハードのことをあまり勉強していない人でも大丈夫・・・なようにしたい
 - ◇ 2回目ぐらいまでは学部の講義レベルぐらいの復習
 - ◇ できればC言語と論理回路の基本がわかってほしい
 - ◇ (わかってなくてもなんとかしたいが、一部きついかも)

コンピュータ・アーキテクチャ

- 「アーキテクチャ」

- ◇ 建築そのものや, 建築における設計や様式

- 「コンピュータ・アーキテクチャ」

- ◇ コンピュータにおける設計や様式

コンピュータ・アーキテクチャ



プログラムに 書かれている命令に 従って計算&出力

◇ 命令セット・アーキテクチャ

□ 命令をどう規定するか

◇ マイクロアーキテクチャ

□ 命令を CPU 内でどう処理するよう作るか

◇ システム・アーキテクチャ

□ CPU やメモリなどを含めた全体をどう作るか



他の分野との関係：ソフトとハードの境界にある

アプリケーション・ソフトウェア
画像処理 / 音声認識 / 言語処理 / 機械制御
WEB サービス / 暗号 ...

システム・ソフトウェア
OS / コンパイラ / インタプリタ

ここ → コンピュータ・アーキテクチャ

論理回路

集積回路 / デバイス

この講義の内容

■ 話題：

- ◇ コンピュータ全体や、主に最近の CPU
- ◇ 他に GPU や 機械学習用ハード、FPGA など

■ 内容：

- ◇ どのような工夫によって高速化されているのか
 - 主に CPU が主題
 - GPU や機械学習専用ハードも程度扱う
- ◇ それぞれのハードは一体何が違うのか？
 - 動作速度や消費電力や熱の違いはどのように決まるのか
 - CPU と比較しながら説明
- ◇ セキュリティなどの話題も取り扱う

キーワード

■ CPU

- ◇ CISC / RISC
- ◇ 命令パイプライン
- ◇ Out-of-order スーパースカラ, In-order スーパースカラ, VLIW
- ◇ 投機実行, キャッシュ, プリフェッチ

■ その他のプロセッサや計算機

- ◇ GPU, FPGA, 機械学習とかの専用プロセッサ
- ◇ プロセッサ以外のメモリ

■ 回路

- ◇ 遅延, 電力, 熱

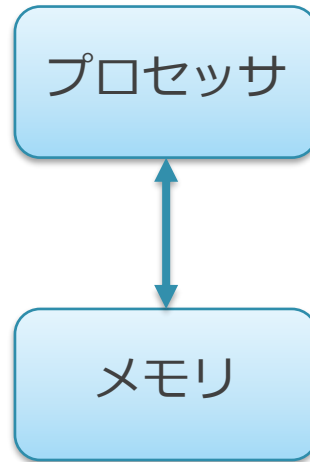
■ セキュリティ

- ◇ Spectre/Meltdown などのアタック

補足 : CPU と GPU

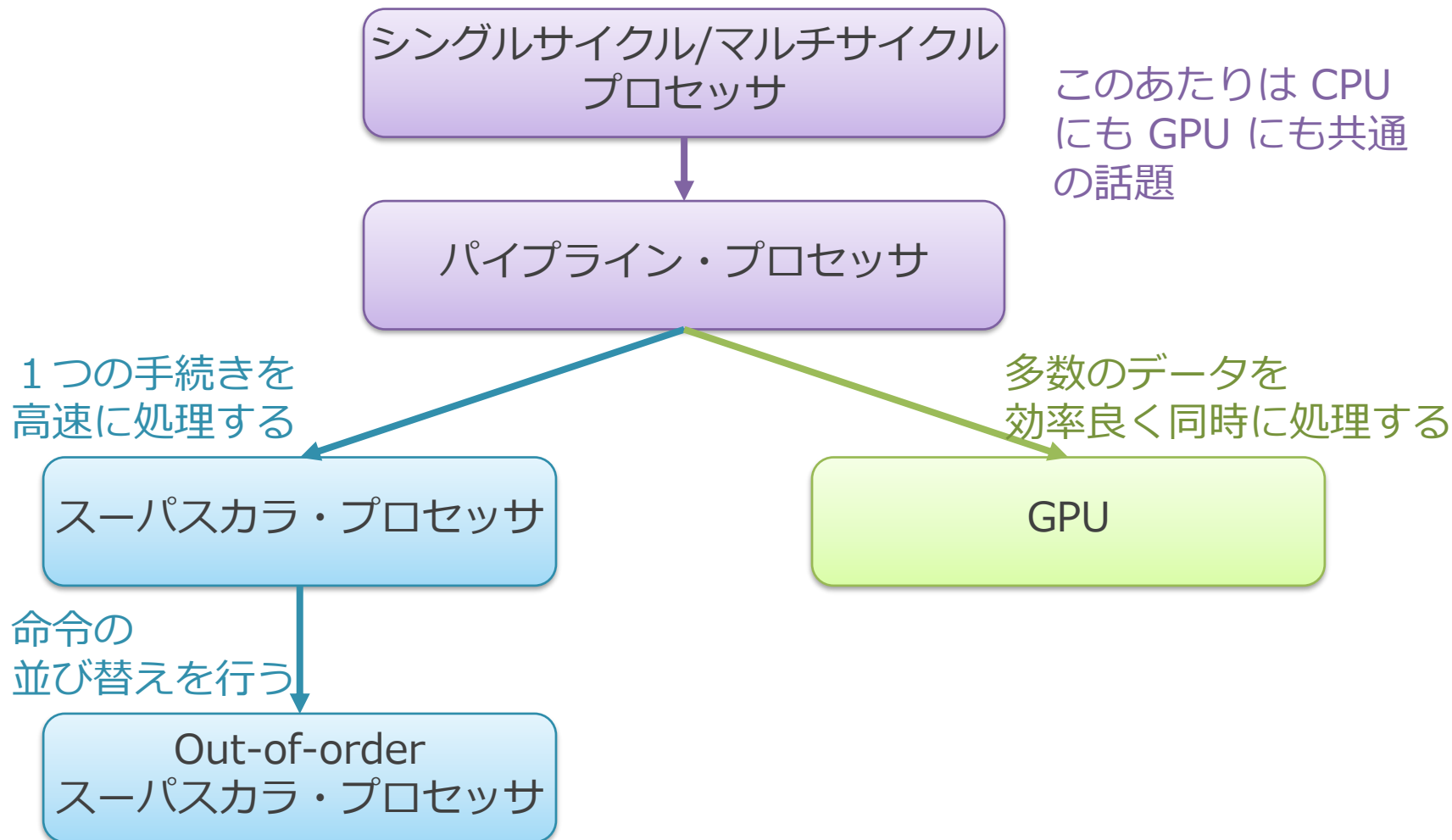
- 最近は全体に GPU に興味を持つ人が多い
 - ◇ 機械学習でよく使うし, 速いから?
 - ◇ CPU にはあんま興味がないんだよな, とかで飛ばされると悲しい
- 背景 : CPU と GPU の関係
 - ◇ コンピュータの基本構造
 - ◇ 各回の講義内容が CPU や GPU にどのように関わるかの関係を説明

コンピュータの基本的な構造



- 超大ざっぱなコンピュータの構造：
 - ◇ 処理対象となるデータを格納するメモリと,
 - ◇ それを処理するプロセッサからなる
- プロセッサあるいはプロセッシング・ユニット：
 - ◇ CPU (Central Processing Unit)
 - ◇ GPU (Graphic Processing Unit)
 - ◇ … いろいろある

いろいろなプロセッサ



- 上の2つと左下は CPU

対象プログラムの違い

■ 逐次型（例：スーパスカラ CPU）

```
1: *a += 1;  
2: if (b != NULL) {  
3:     c += *b;  
4: }
```

1行目と3行目を並列実行するには？

- a と b が同じ場所を指すかも？
 - ポインタのアドレス一致予測
- b が NULL ではない必要？
 - 分岐予測

■ データ並列型（例：GPU）

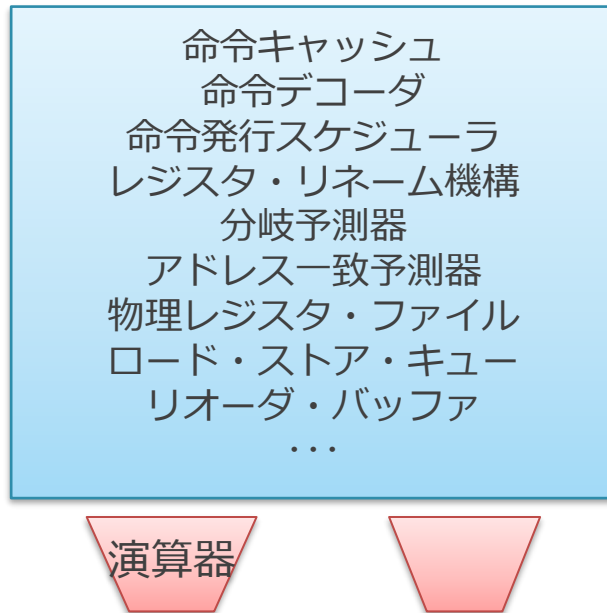
```
1: for (i=0; i<1000; i++) {  
2:     c[i] += a[i] * b[i];  
3: }
```

ループの各周回は自明に並列実行できる

■ 並列性の違い

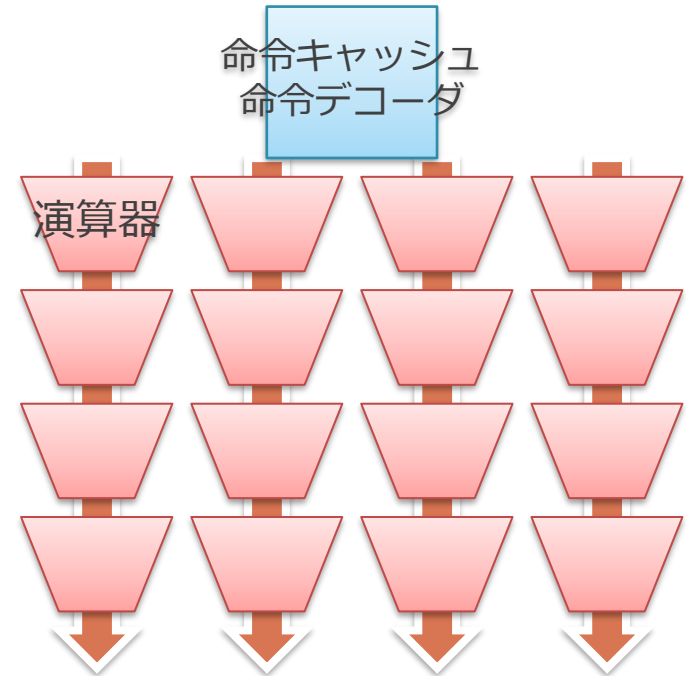
- ◇ 逐次型では各種の投機や並び替えを行って並列実行
- ◇ データ並列型では自明に並列実行できる

基本構造の違い



■ 逐次型：スーパスカラ

- ◇ 頑張って並列実行できる部分を探し、小数の演算器で実行
- ◇ 主に CPU
PC やサーバ, スマホ



■ データ並列型：

- ◇ 1つの命令を多数の演算器にブロードキャスト
- ◇ 主に GPU やアクセラレータ
スパコン

両者では、やるべき事も全く異なる

講義内容との関係

紫：共通，青：CPU，緑：GPU



- 下記で公開しています

- ◇ <https://github.com/shioyadan/advanced-computer-organization>
- ◇ よくわからない部分への質問や、タイポや間違いの指摘は issue を投げてもらうのもあります

- 内容のアップデート予定

- ◇ 前回講義時（2年前）と比べて実物に関わる経験が増えた
- ◇ 特に GPU や機械学習ハードに対する解像度が上がった
 - これらの知見を反映させたい

参考図書など

- デイビッド A パターソン, ジョン L ヘネシー :
「コンピュータの構成と設計」
 - ◇ 通称「パタヘネ」
- ジョン L ヘネシー, デイビッド A パターソン :
「コンピュータアーキテクチャ 定量的アプローチ」
 - ◇ 通称「ヘネパタ」（著者順が逆）
 - ◇ こっちの方が発展的内容
- 安藤秀樹 :
「命令レベル並列処理—プロセッサアーキテクチャとコンパイラ
(並列処理シリーズ)」
 - ◇ 主にスーパスカラプロセッサと VLIW を扱う

他の講義との関係

- 各地の学部：
 - ◇ 論理回路, デジタル回路
 - ◇ コンピュータ・アーキテクチャ, 計算機アーキテクチャ
- 大学院：東大だと, 下記の先生方が近い内容を扱っているはず
 - ◇ 電子情報学（入江先生, 門本先生）
 - ◇ システム情報学（中村先生, 高瀬先生）
 - ◇ コンピュータ科学（中田先生, 高前田先生）

- 出席感想とレポート・・・で主につけていました
- 今年は変えるかもしれない
 - ◇ LLM がレポート課題を難しくしてしまった
 - ◇ 別の選択課題も出そうと考えています
 - プログラムの特性の解析など

来週 4/15 は休講です

- 次回は 4/22 です

今日の出席

■ 出欠

- ◇ パスワード：, 本日24時まで
- ◇ 1回講義時間中に出席を LMS でつけて, 後から感想を書いてください

■ 下記について, それぞれ知っているかどうかを LMS の出席欄に書いて下さい

1. 組み合わせ回路と順序回路
2. 命令とは何か
3. 命令パイプライン
4. キャッシュ
5. スーパスカラ

■ 注意：

- ◇ 回答は「わかる」「聞いたことがある」「知らない」等をお願いします
- ◇ この講義は上記があんまわかってない人向けの説明から始まります
- ◇ その他コメントや, 「こう言う話題も取り上げて欲しい」や質問などあればお願いします