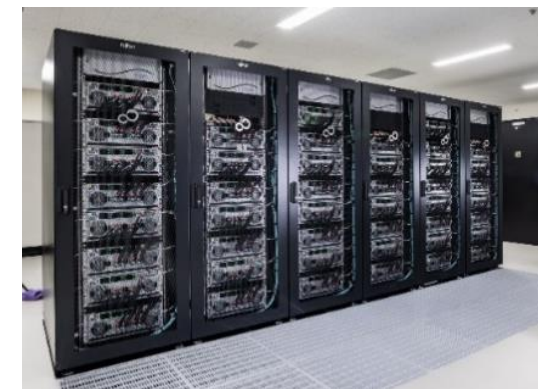


スーパーコンピュータはなぜ速い！？ その仕組みを徹底解説(13:30-)

住元 真司 (すみもと しんじ)

スーパーコンピューティング研究部門
特任教授



スーパーコンピュータって何で速いのでしょうか？

- 東京大学・情報基盤センターではスーパーコンピュータを運用しています。
 - Wisteria/BDEC-01(写真右)
- 本講演では、スーパーコンピュータの仕組みについて徹底解説します。
 - コンピュータの構成と仕組み
 - スマートホンとスーパーコンピュータの違い
 - スーパーコンピュータの速さの仕組み



Odyssey



Wisteria/BDEC-01



Aquarius

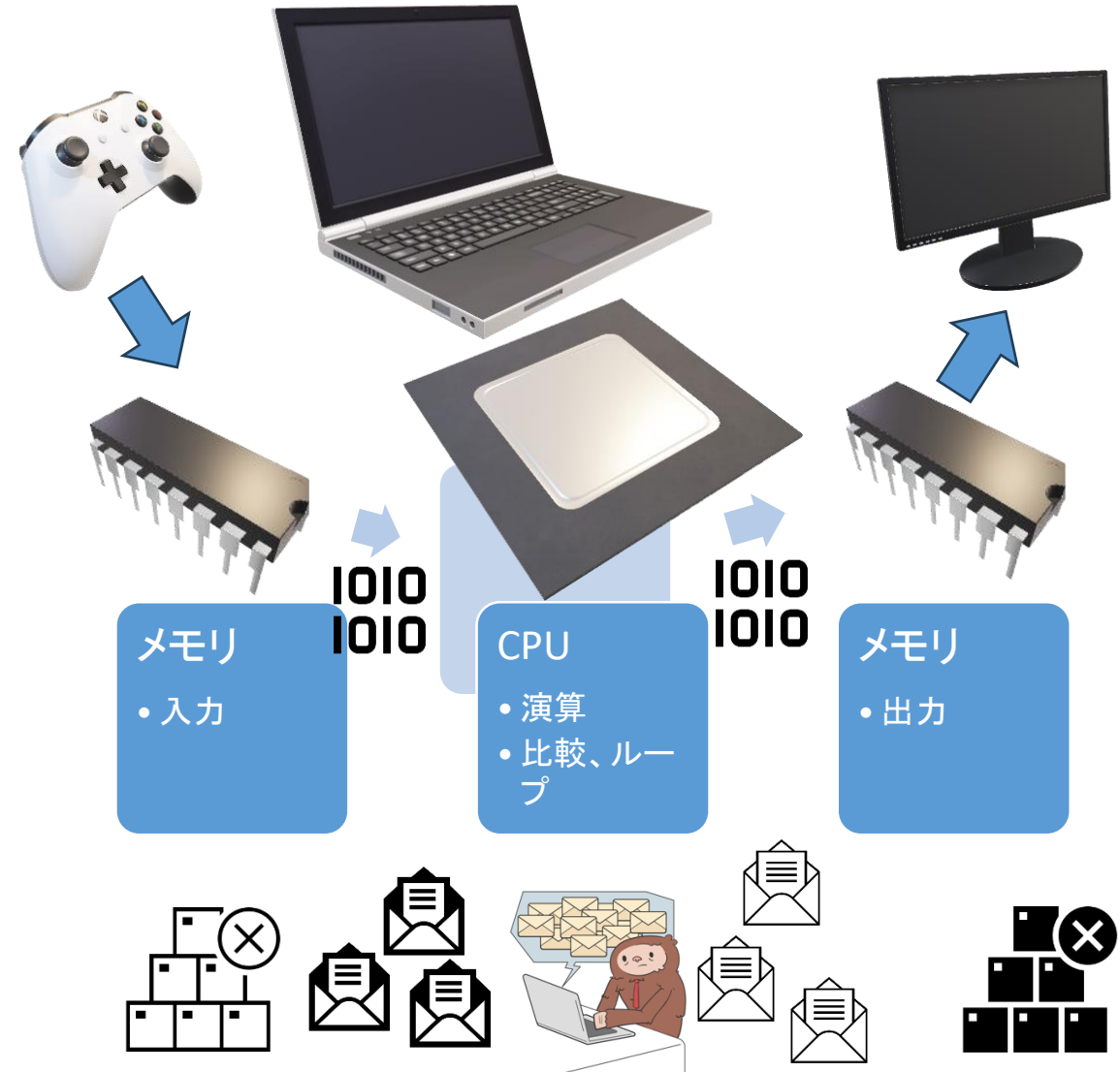
コンピューターの構成

- 記憶装置（メモリ、ディスク）：
データを置く倉庫
- 中央演算装置（CPU）：デー
タを加工する装置
- I/O装置：プリンタ、ディスプレイ、
マウス、キーボード：データを
入出力する
- ネットワーク：他のコンピュータと
データ交換する装置



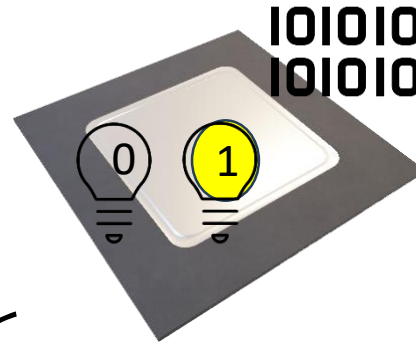
コンピュータの動作の仕組み

- 記憶装置(メモリ)からデータを入力してCPUで加工してメモリに出力します
 - オフィスでの書類処理
- メモリその他、I/O処理装置 (キーボード、ディスプレイなど) からも入出力できます。
 - 入出力データはメモリを介して行われます



コンピューターの動作原理！

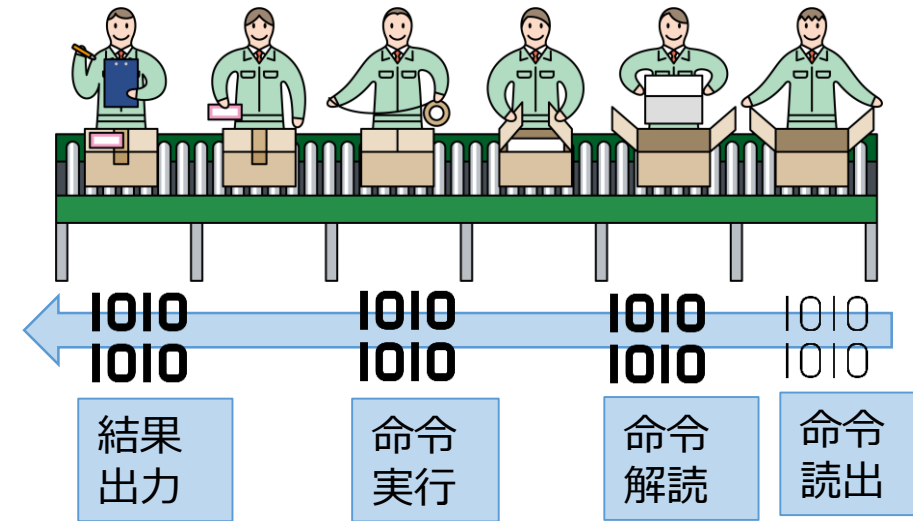
- コンピューターは2進数で動作
0 or 1 の集まりで動作
 - 0と1を電圧の有る無しで表現
- 0と1をたくさん組み合わせると
組み合わせは無限大 ∞
 - 組み合わせを命令とデータ表現に利用
- 命令列を作ることがコンピュータープログラミング
 - 命令列： 足し算、引き算、掛け算、割り算、比較と繰り返し、など単純
- プログラムの処理を行うのが中央演算装置(CPU)



10101010101010101010101010101010
10101010101010101010101010101010

(例)
16個の0と1(16ビット)命令
0100 0011 0100 0000
読む どこへ どこから
1001 0001 0010 0011
たし算 なにと なにをどこへ

流れ作業＝パイプライン

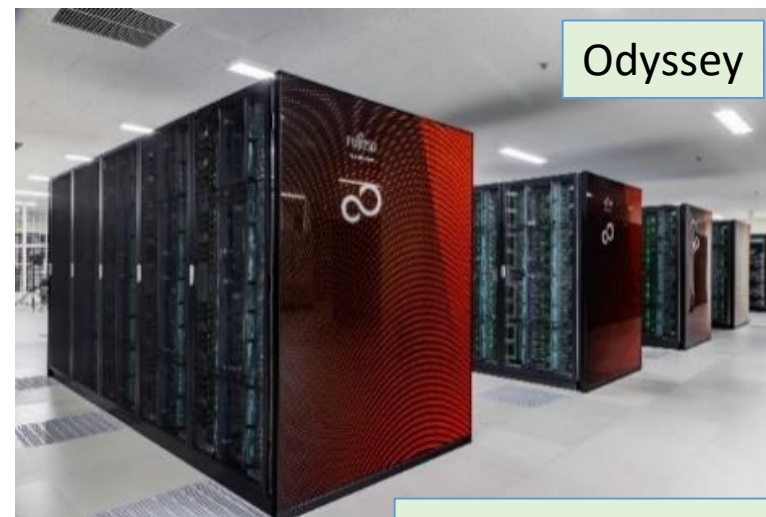


スーパーコンピュータはなぜ速いの？

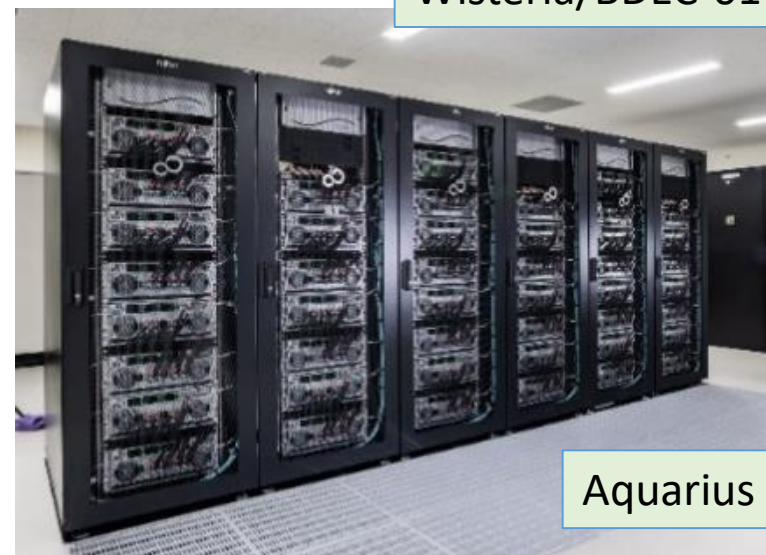
- スマートフォンと比べてみましょう！



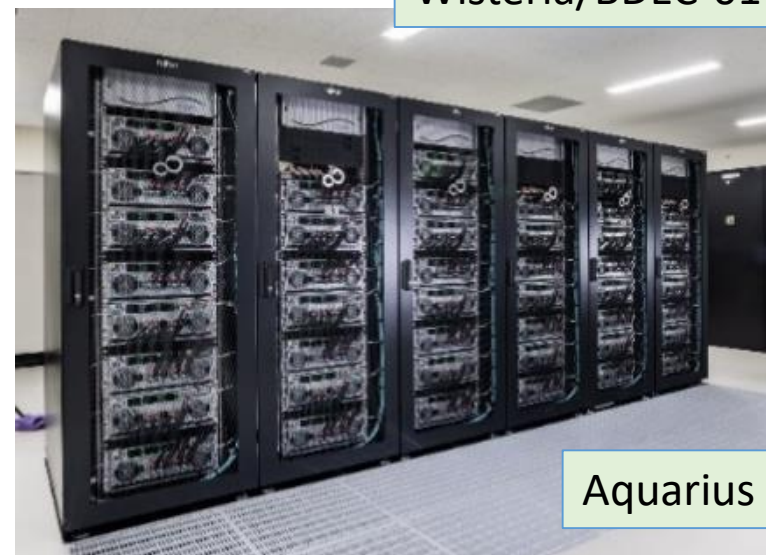
VS.



Odyssey



Wisteria/BDEC-01

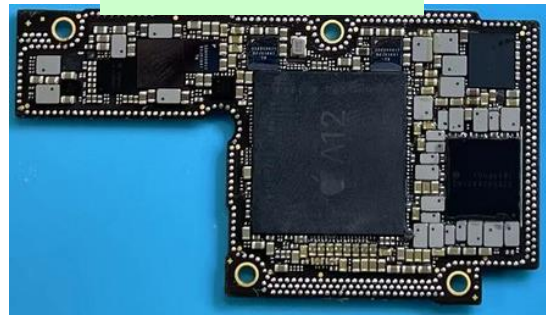


Aquarius

スマートフォンとスーパーコンピュータの違いは？

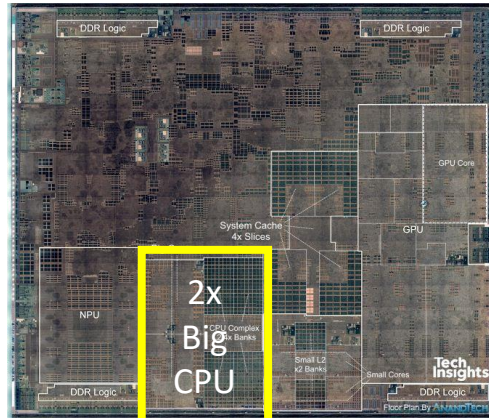


スマートフォン

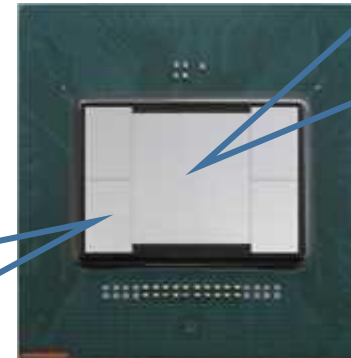


VS.

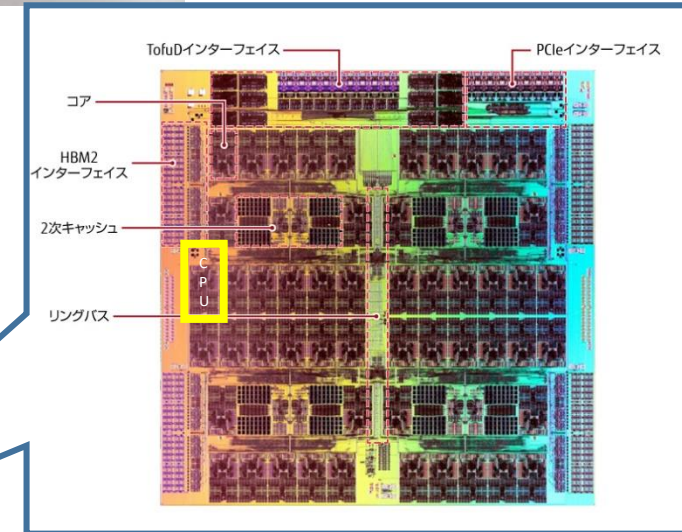
<https://www.anandtech.com/show/13392/the-iphone-xs-xs-max-review-unveiling-the-silicon-secrets/2>



スーパーコンピュータ
Odyssey



高性能
メモリ



<https://www.fujitsu.com/jp/about/resources/publications/technicalreview/2020-03/article03.html>
https://www.hpci-office.jp/invite2/documents2/ws_material_191018_shimizu.pdf

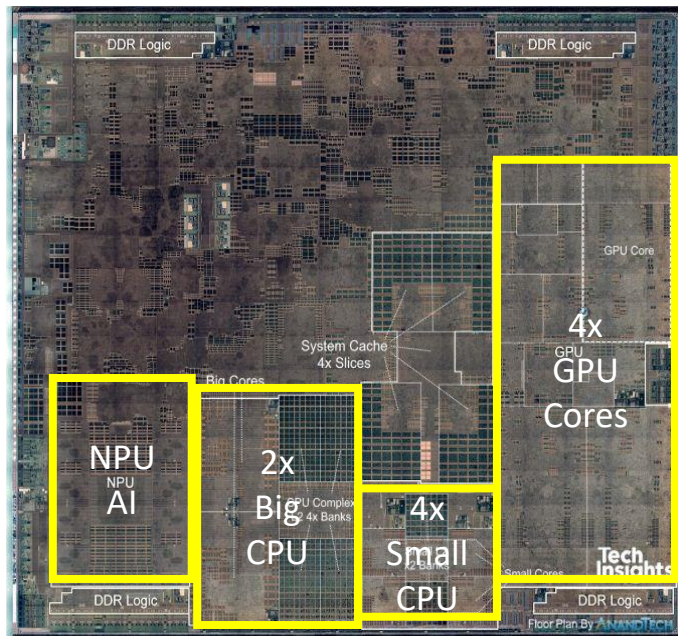


- 同じ半導体プロセス (TSMC 7nm)

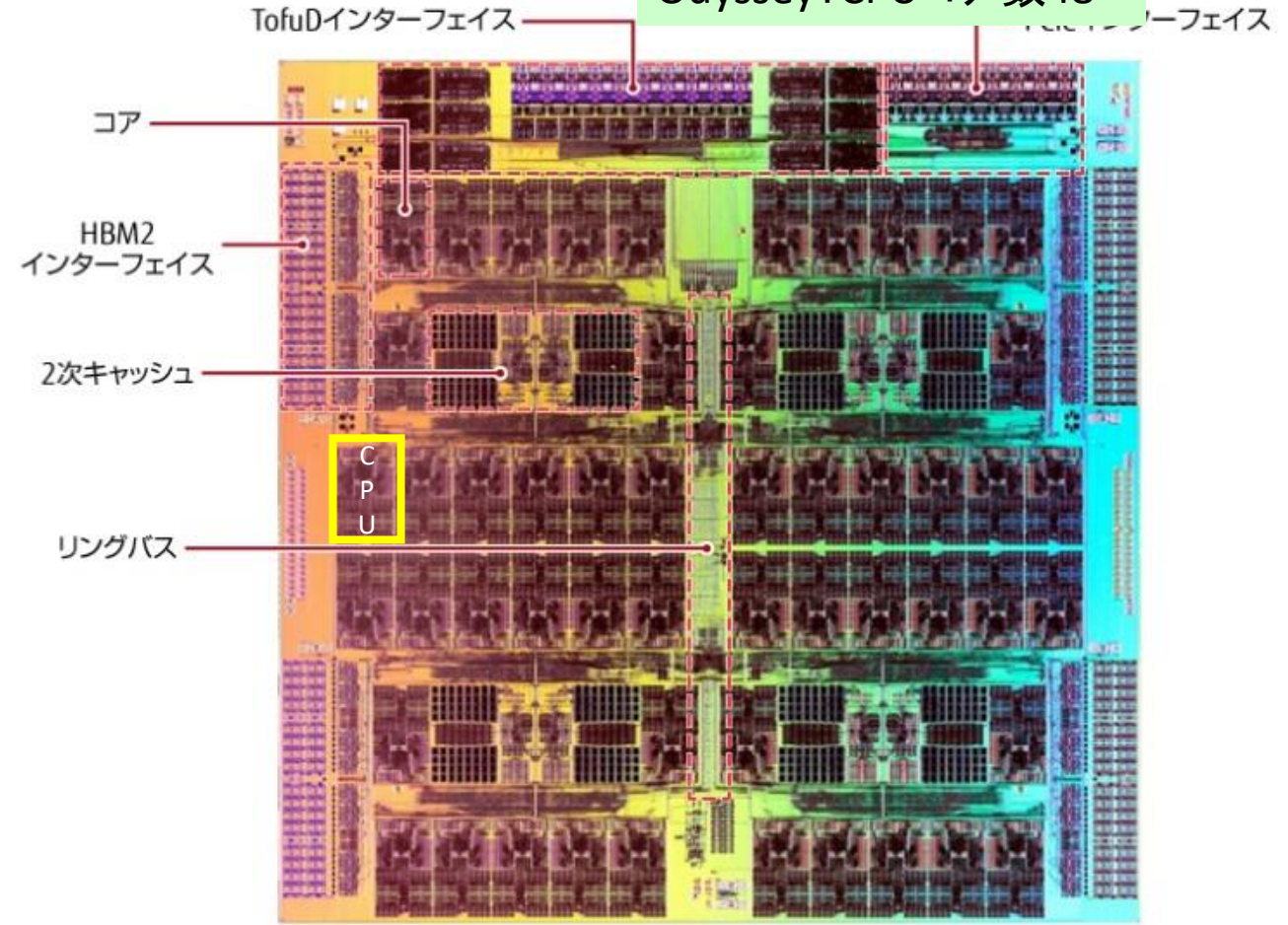
CPUを構成するLSIをのぞいてみよう

- スマートフォンもスーパーコンピューターもCPUコア自体は同じArmと呼ばれるCPU
- スマートフォン用はGPUや動作に必要な回路がたくさん面積を占める
- スーパーコンピューター用は必要最小限の回路で構成、たくさんのCPUを持つ

スマートフォン
iPhone
CPUコア数2+4



スーパーコンピューター
Odyssey: CPUコア数48



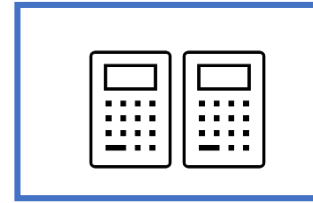
スマートフォンとスーパーコンピュータのCPU比較

- スマートフォンのCPUとOdysseyのCPUは同じArmプロセッサ
- CPUの速さを決める周波数(1秒間にカウントできる数)もほぼ同じ
A12=2.5GHz, A64FX=2.2GHz
- 違いはCPUコアの数と同時計算できる数の違い(SIMD)：複数の電卓を一度に操作できる

	スマートフォン	スーパーコンピューター
CPUコアあたり同時計算数	2	8
CPUコア数	2	48
総演算数	4	384

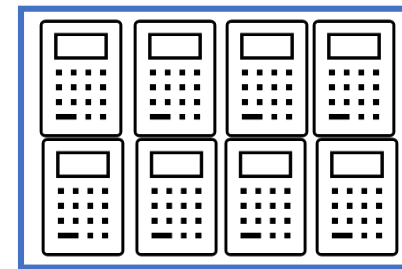
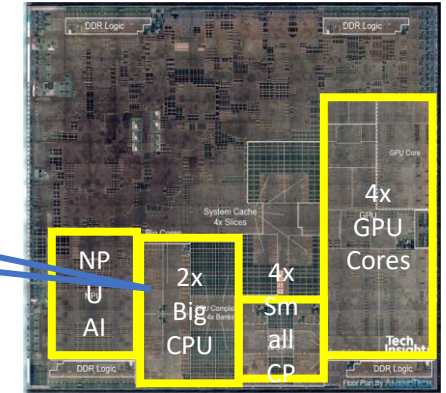
- スーパーコンピューターのCPUの方が96倍
たくさんの演算器を搭載

1 CPUコアあたり



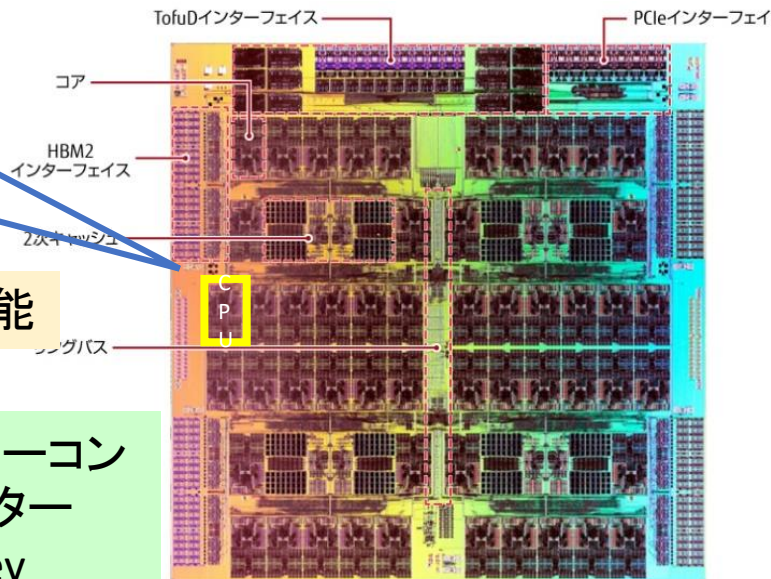
スマートフォン
iPhone

同時に2つ演算実行可能



同時に8つ演算実行可能

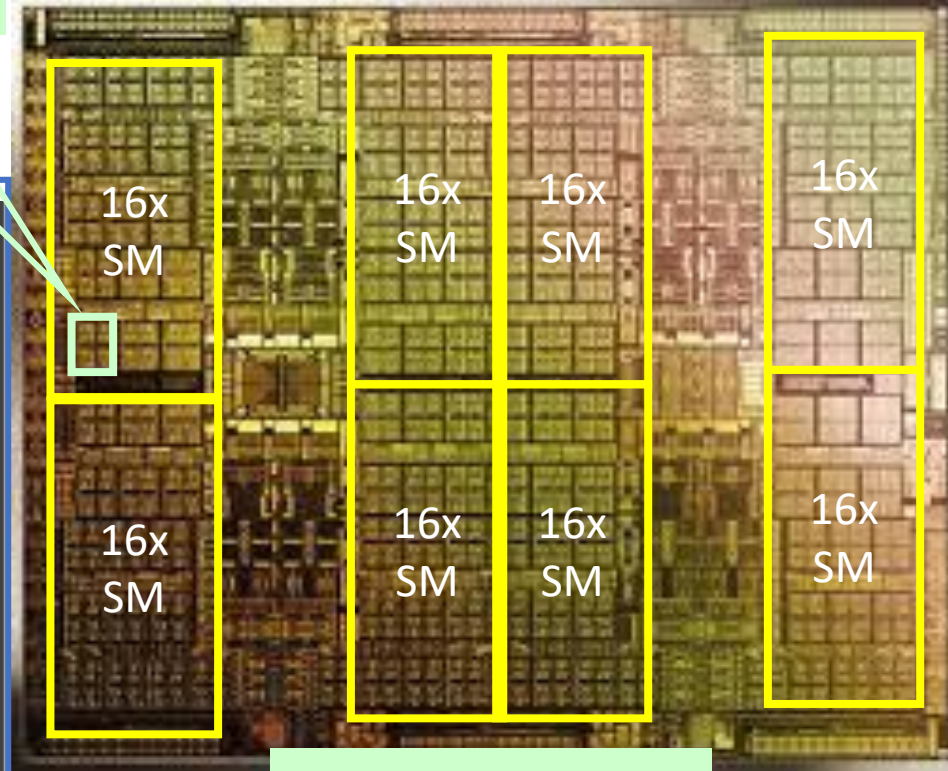
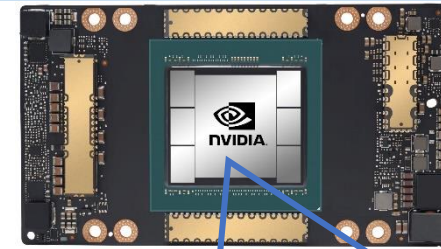
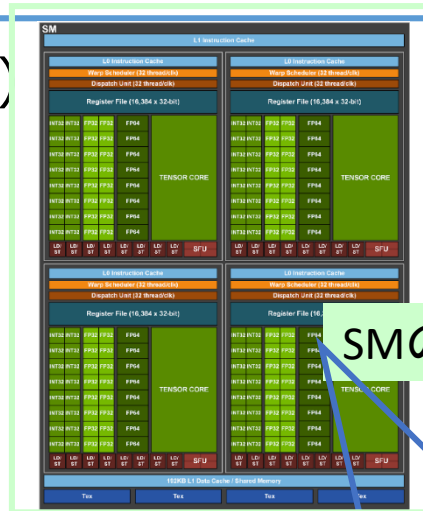
スーパーコン
ピューター
Odyssey



Aquarius搭載: GP-GPU NVIDIA A100

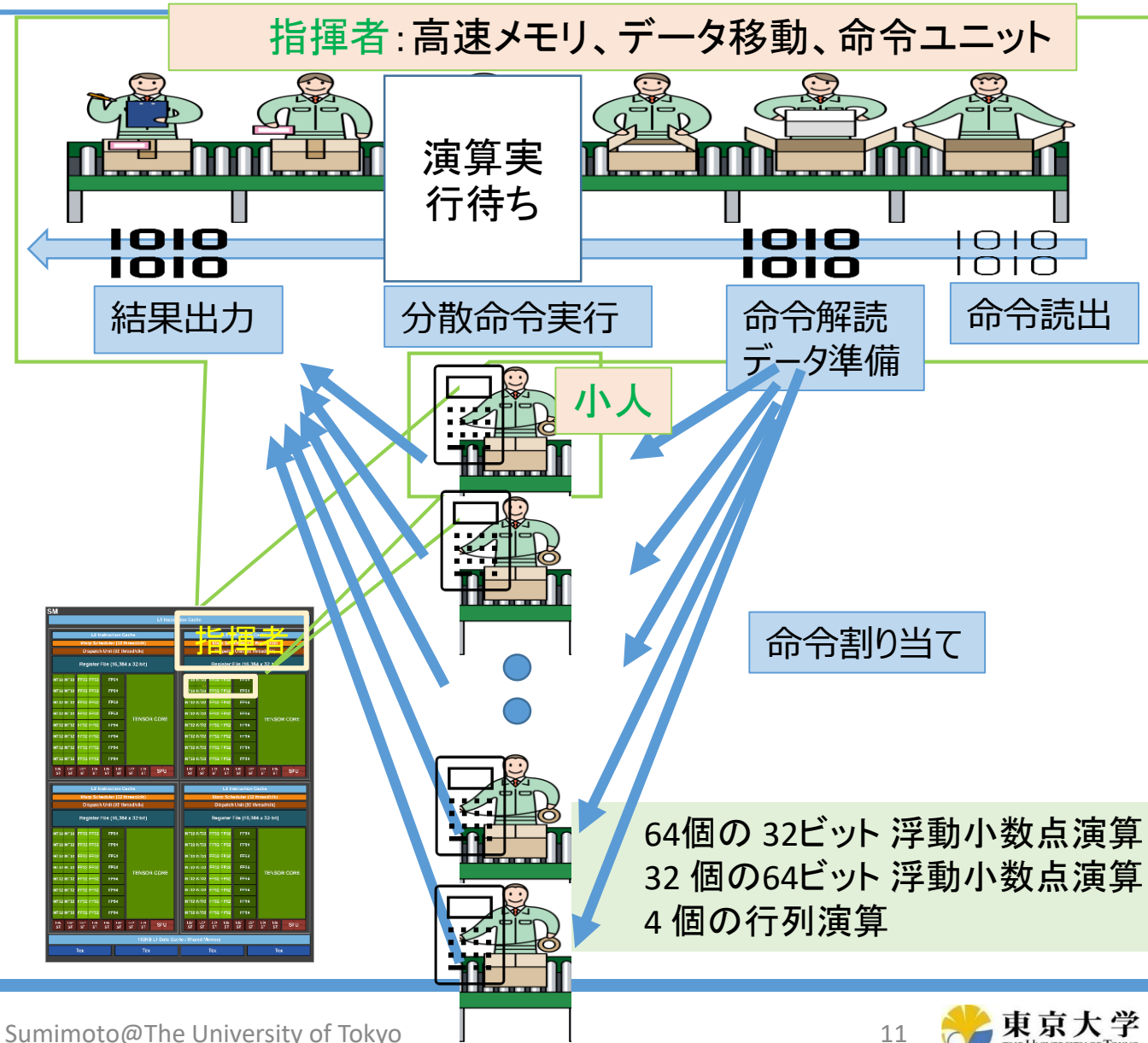


- 1つのSM(Streaming Multiprocessor)
 - 64 32ビット 浮動小数点演算(cuda)
 - 64 32ビット 整数演算(Cuda)
 - 32 64ビット 浮動小数点演算(cuda)
 - 4 行列演算(Tensor)
 - 高速メモリ、データ移動、命令ユニット
- 16つだと
 - 1024個の32ビット 浮動小数点演算
 - 1024個の32ビット 整数演算
 - 512個の64ビット 浮動小数点演算
 - 64個の行列演算
- 全体で
 - 8192個の32ビット 浮動小数点演算
 - 8192個の32ビット 整数演算
 - 4096個の64ビット 浮動小数点演算
 - 512個の行列演算
- Odysseyの演算数に比べて10倍多い



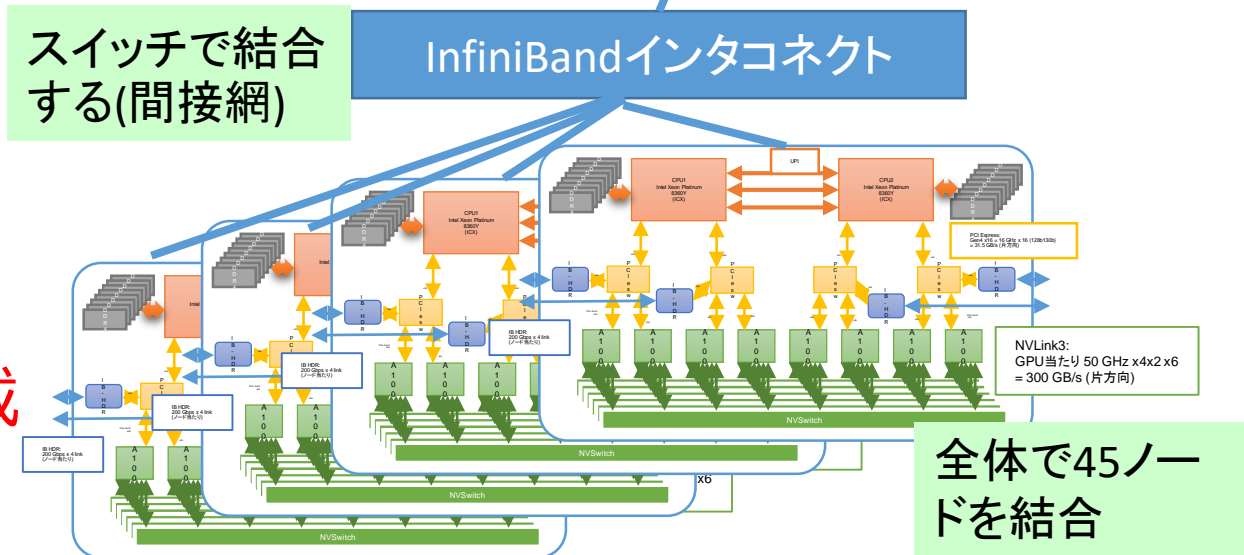
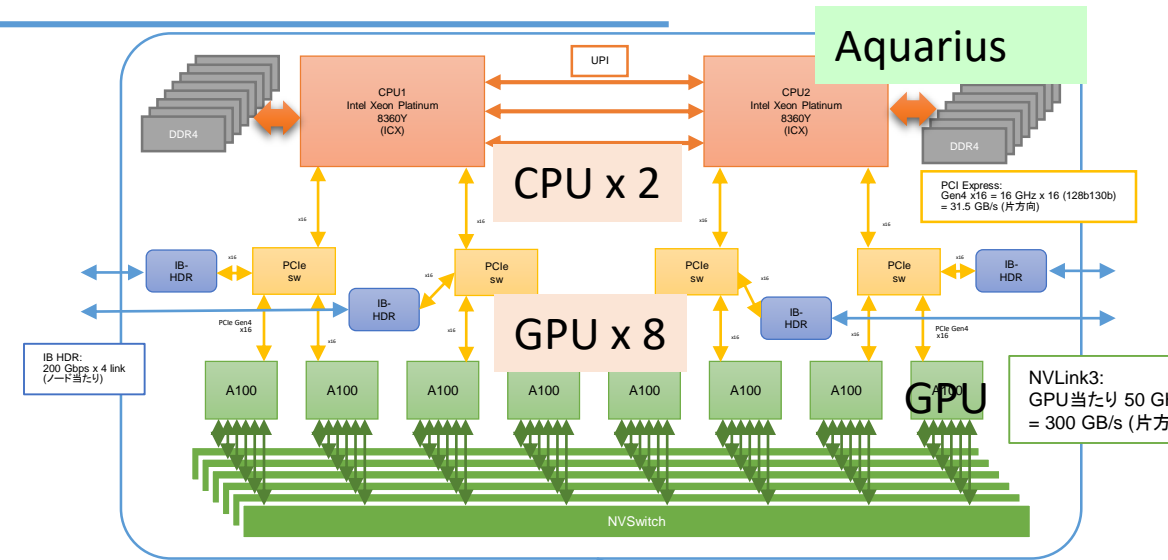
GPUはなぜ速いの？

- CPUの命令実行部分を多数のGPUコアに分散して並列処理
 - 指揮者が演算しかできない多数の小人に手分けして並列処理
 - 多数の小人に仕事を与える指揮者がデータ用意して演算を指示
 - たくさんの小人が同時に演算を実行するので速く計算可能
 - 但し掛け算や足し算など単純な計算しかできない



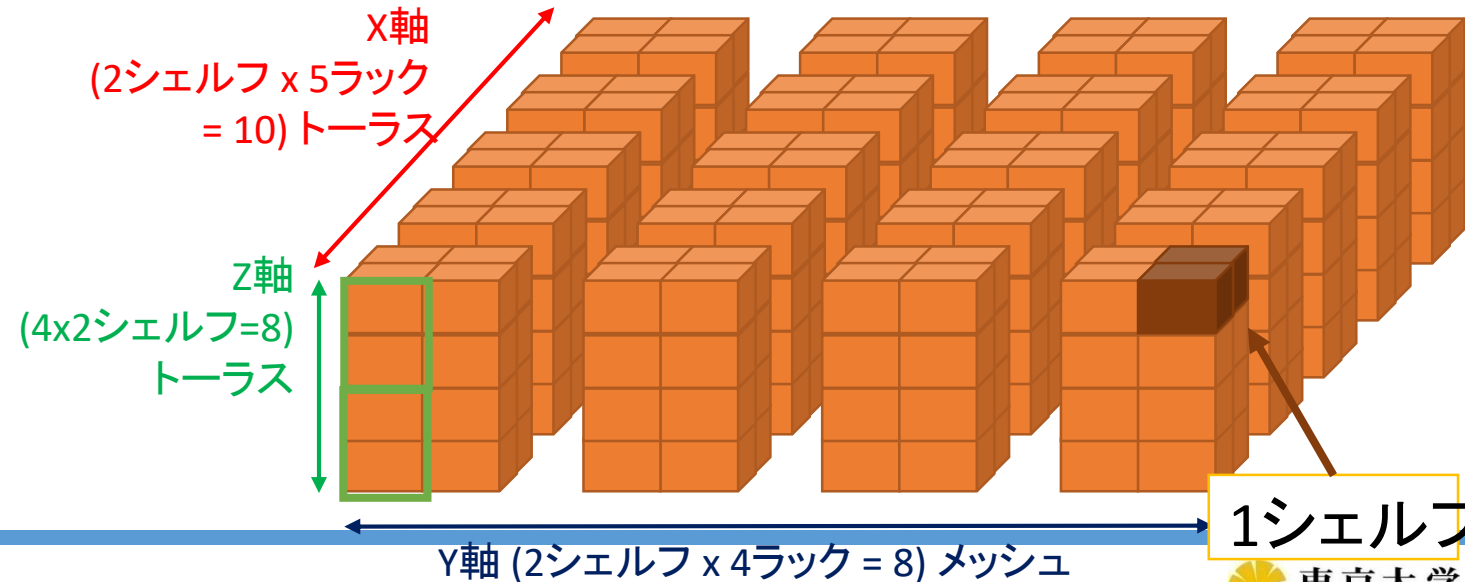
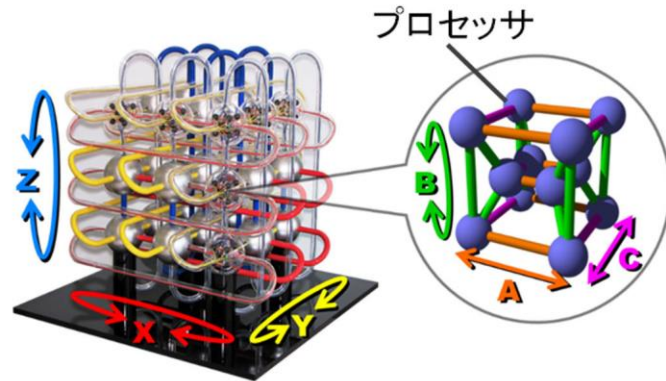
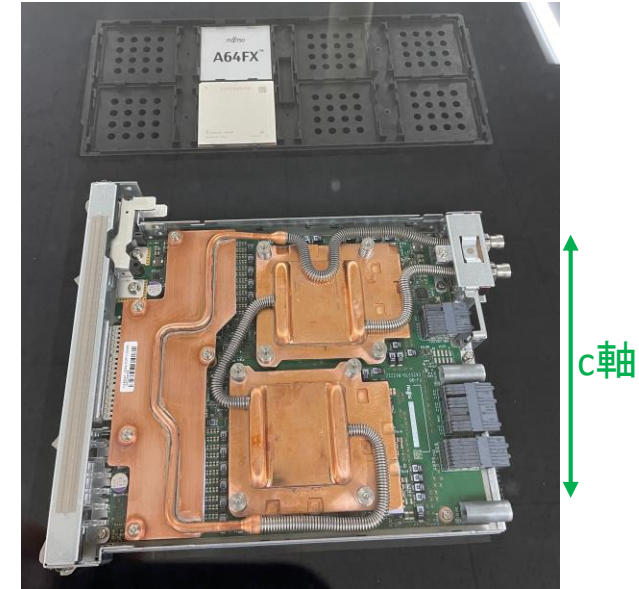
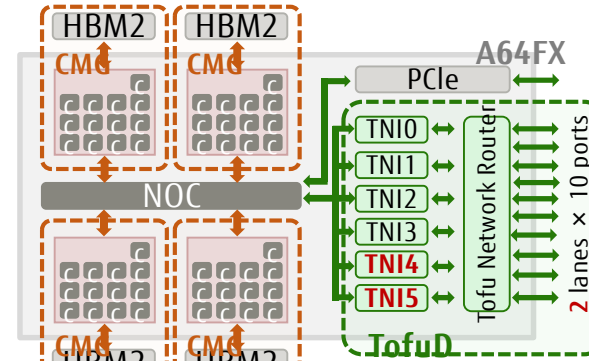
スーパーコンピュータの速さはこれだけではありません！

- 高い性能のCPU & GPUを搭載した計算機をたくさんつないで並列に計算
 - マルチコアCPU： 教室内（オフィス内） -100並列(人) = CPU単体
 - Odyssey： 48CPUコア
 - マルチソケットCPU： 建物内（オフィス） -400並列(人) = ノード(サーバー、計算機)
 - Aquarius： 2 CPU + 8 GPUs
- 計算機間はインタコネクト(ネットワーク)
 - スーパーコンピューター： 複数の建物（支店）
 - Odyssey： 7680ノード
 - スーパーコンピュータ 富岳:約16万ノード
- Aquarius: **スマートフォンに比べて1024 x 8x45 = 368,640倍たくさんの演算器を搭載**



OdysseyのTofu-Dインタコネクト

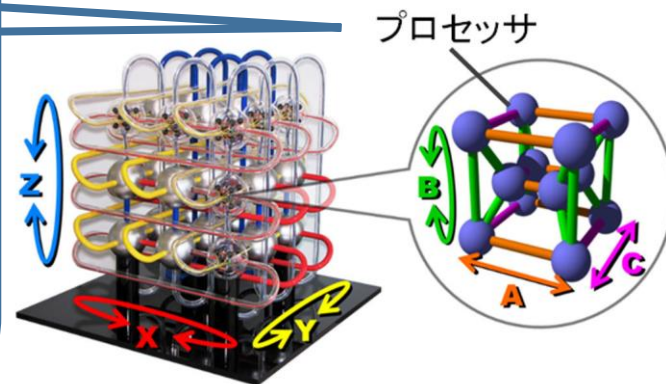
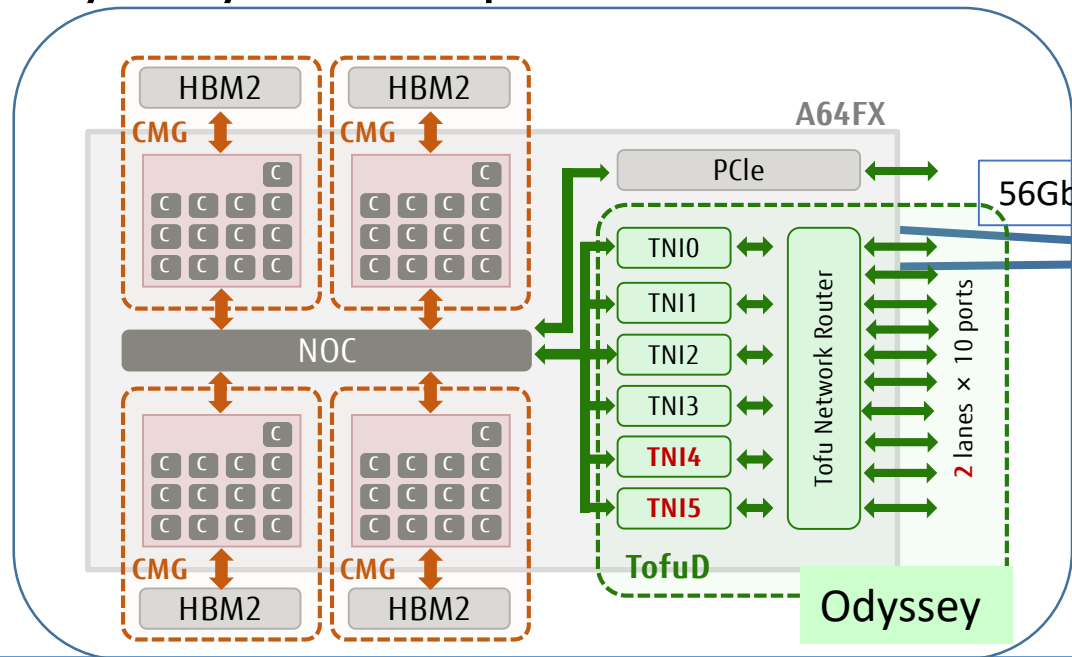
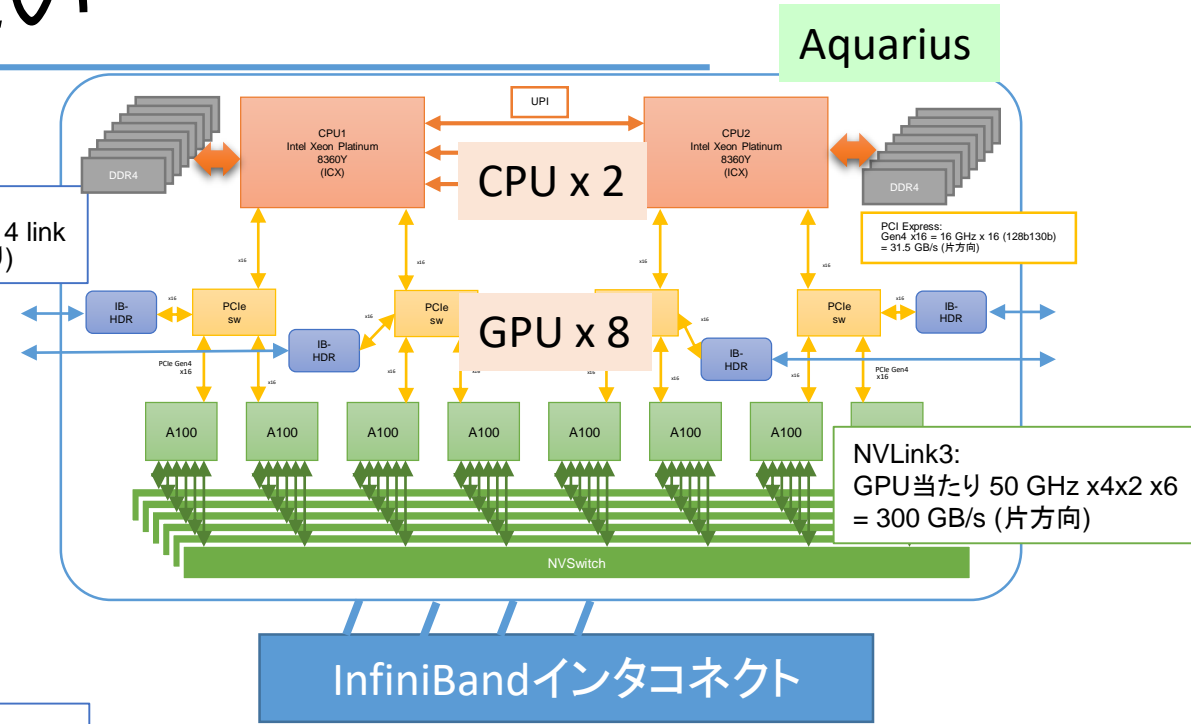
- OdysseyのA64FX CPUはインタコネクトを内蔵
 - イカと同じ10本足で、複数のノード(計算機)と直接手をつなぐ(直接網)
 - 6次元トーラスメッシュという構成
- Tofu-D 形状: 全20ラック
 - $(X, Y, Z, a, b, c) = (10, 8, 8, 2, 3, 2)$
- 3次元、2次元、1次元に展開して指定可能
- 典型的な例 (全系) 3次元: $20 \times 24 \times 16$ ($X*a, Y*b, Z*c$)



- スマートフォンに比べて $96 \times 7680 = 737,280$ 倍たくさんの演算器を搭載

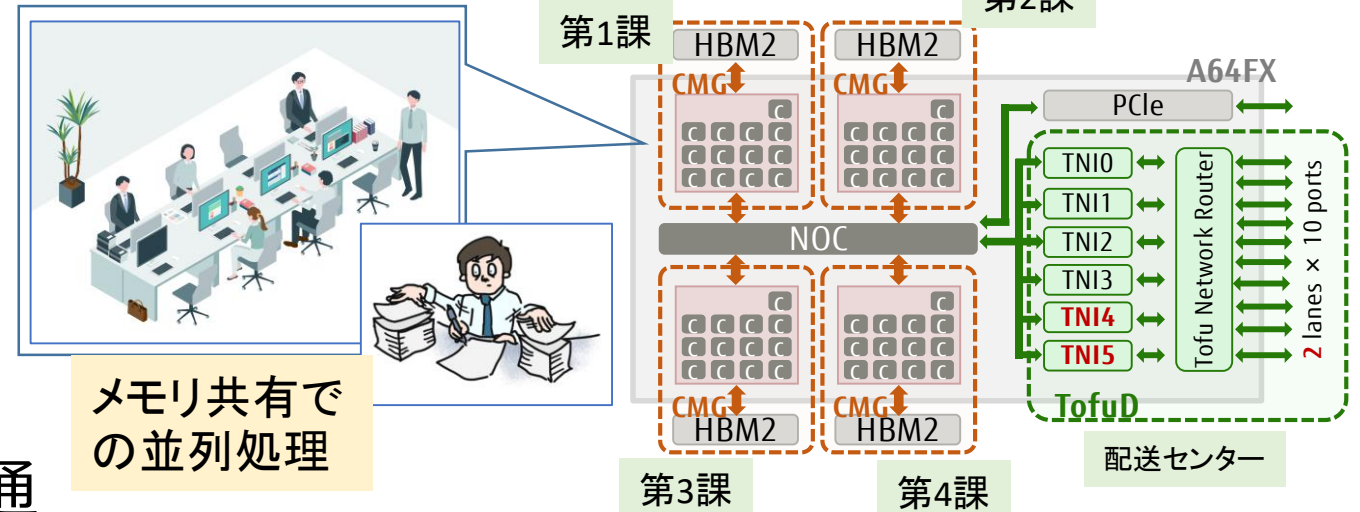
インタコネク(ネットワーク)の違い

- スーパーコンピュータのインタコネクは接続できるノードの数だけではありません
- 家庭のネットワーク：1Gbpsに対して
 - Aquarius：200Gbps x 4 = **800倍高速**
 - Odyssey：56Gbps x 10 = **560倍高速**



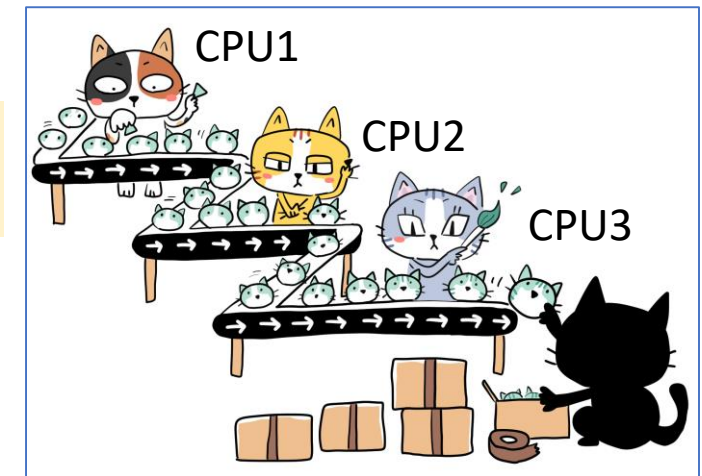
それでは、たくさんのノードにどうやって分散してプログラムを実行するのでしょうか？

- 人間の世界と同じように、仕事を複数の人に分担して処理を進めます。
 - CPU:メモリ上でデータを共有して協調して並列処理：オフィスに相当
 - GPU:マスタ(課長)がワーカ(課員)指示して処理分散：(指揮者と小人)
- ネットワークを介する場合は外へのデータ通信が必要=トラックでの書類送付に相当
 - ある程度まとめて書類を送る
 - 流れ作業にすることで各ノードの処理のバラツキを吸収するなど
- 仕事の分散処理方式は工夫次第で無限大の組み合わせ：社会活動の縮図
 - 日夜、研究者・技術者のみなさまが改善・高速化に取り組んでいます。



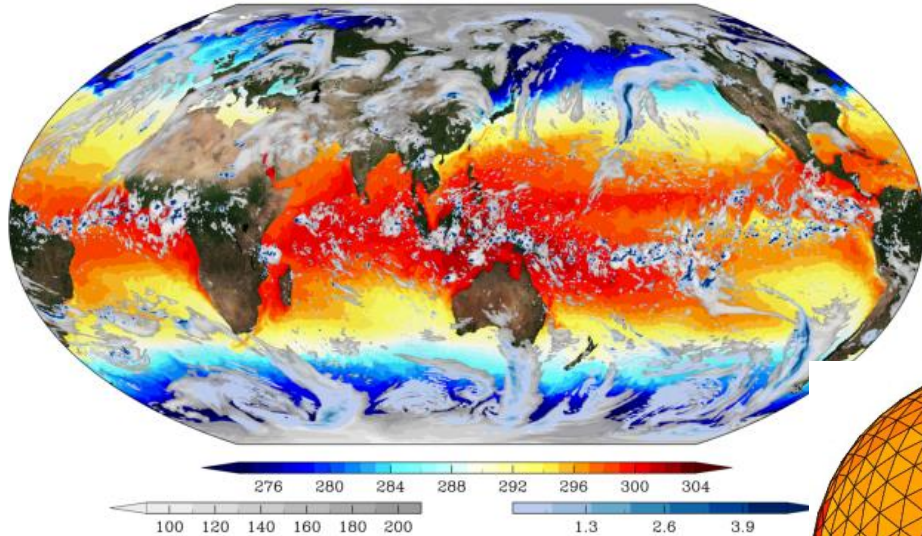
メモリ共有での並列処理

流れ作業(パイプライン)での並列処理

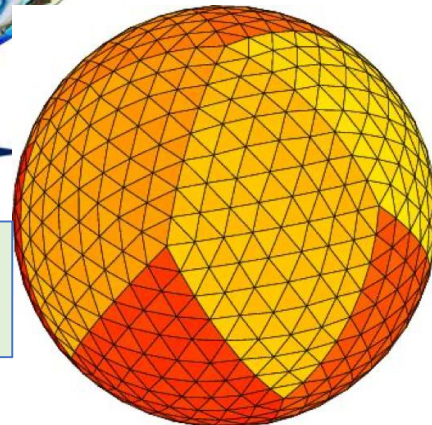


大規模計算をどうやって実行するの？

- 計算対象毎に区間分割して各分割毎に各CPUに割り当て
- 計算量が偏る場合は部分的に細かく分割
- ノードをまたがる場合はインタコネクトによる通信が必要

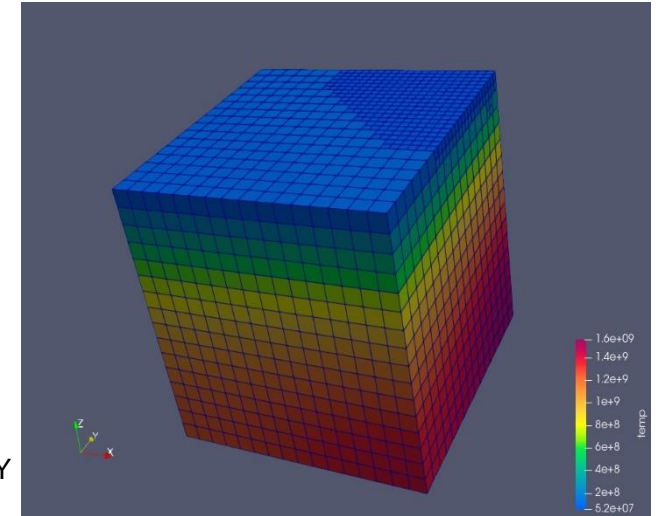
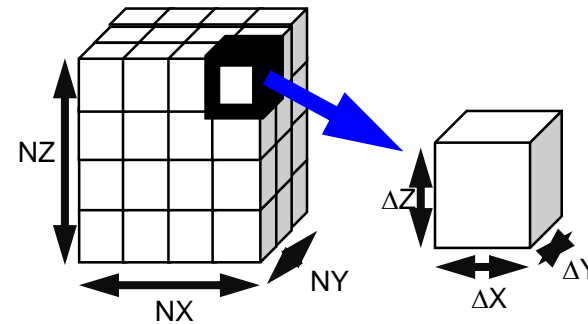


NICAM: 大気・海洋アプリケーション
球体の場合は正20面体格子割当て



$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} + f(x_c, y_c, z_c) = 0$$

$$\phi = 0 @ Z = Z_{\max}$$

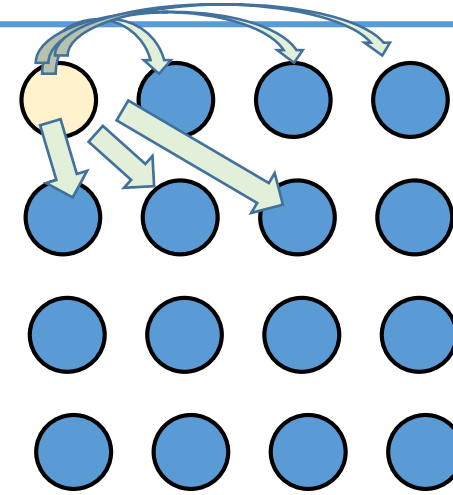


pHEAT-3D: 熱伝導アプリケーション
直方体の場合はX-Y-Z 分割割り当て

インタコネク上で通信最適化

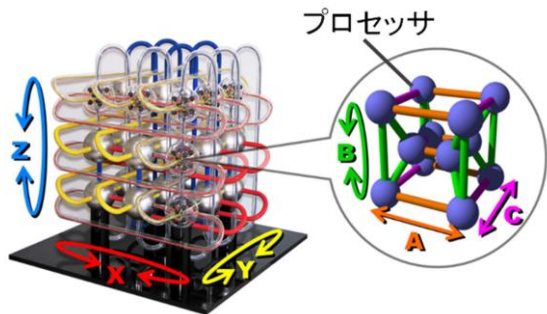
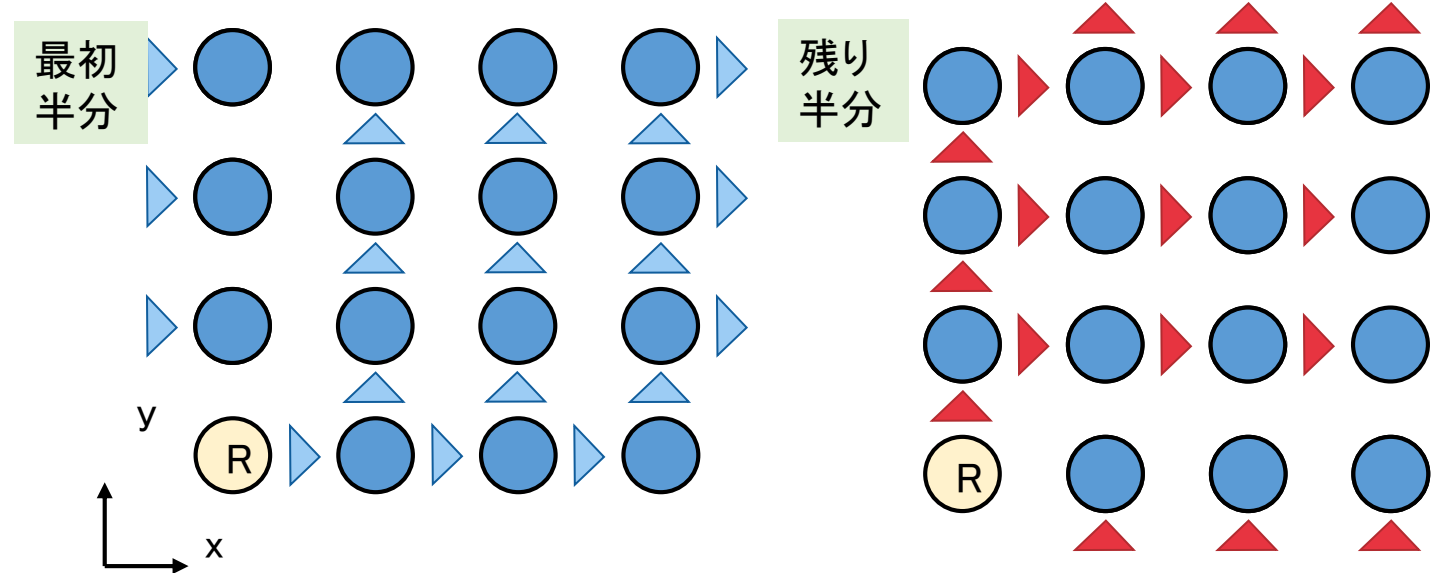
- 同じデータを他のすべてのノードに転送する：
Broadcast

- 一人の人がすべての人に手渡しする：
人数が少ないと手間ではないかも
- 数百人に手渡しするのは大変



- 通信の最適化：バケツリレー方式

- 複数の手を使ってバケツリレーを繰り返す
- データの大きさが速さが違う



まとめ

- 本講演では、スーパーコンピュータの仕組みについて徹底解説しました。

- コンピュータの構成と仕組み
- スマートホンとスーパーコンピュータの違い
- スーパーコンピュータの速さの仕組み



- スーパーコンピューターの中には人間社会の中でも活用可能な技術(工夫)がいっぱいです。
 - 人との違いは単純な計算しかできませんが、ばらつきがなく均一で高速に処理できることです。
- 面白いと感じた方：**スーパーコンピュータ・プログラミングを始めてみませんか？**

Questions?