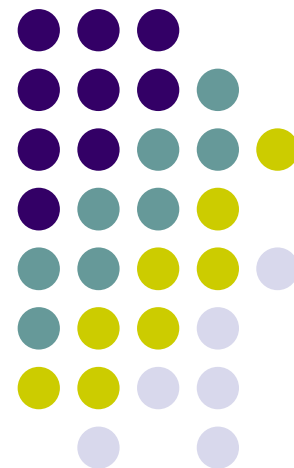


2014年度 実践的並列コンピューティング 第1回

授業の概要と
TSUBAMEスパコンガイダンス

遠藤 敏夫

endo@is.titech.ac.jp



本授業のねらい



- 並列計算・高性能計算技術を実践的に学ぶ
 - 講義＋**実習**
 - この場で実際に東工大TSUBAME2スーパーコンピュータを使います
- 予定内容
 - HPCのためのライブラリや言語
 - OpenMP
 - MPI
 - CUDA (GPUのためのプログラミング)
 - その他追加の可能性あり

お知らせがあるときはOCW webページにて

- <http://ocw.titech.ac.jp> → 「実践的並列コンピューティング」で検索



必要となる基礎知識

- 本講義ではC言語をベースに
 - ポインタの概念、malloc/free
 - Pthread, Javaスレッドは知っているが良いが必須ではない
- 基礎的なLinuxコマンド
 - コンパイル、makeコマンド

半年の流れと単位認定

- OpenMP: 共有メモリ並列プログラミング
- MPI: 分散メモリ並列プログラミング
- GPUプログラミング
- その他



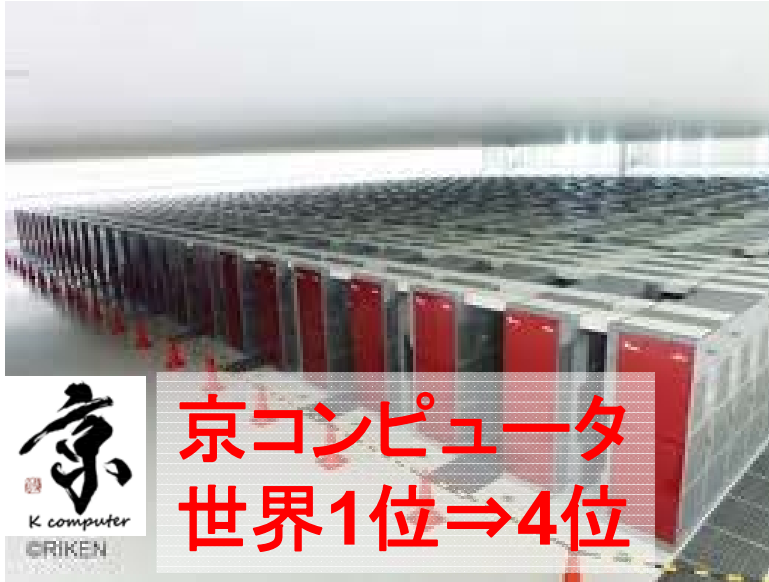
下記により採点・単位認定

- 各パートで課題を出す。2つ以上のパートのレポート提出を必須とする
 - ×切は、パート終了の2~3週間後
- 出席点



スーパーコンピュータとは

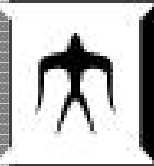
色々なスーパーコンピュータ



京コンピュータ
世界1位⇒4位



東工大TSUBAME2
世界4位⇒11位



BlueGene/P



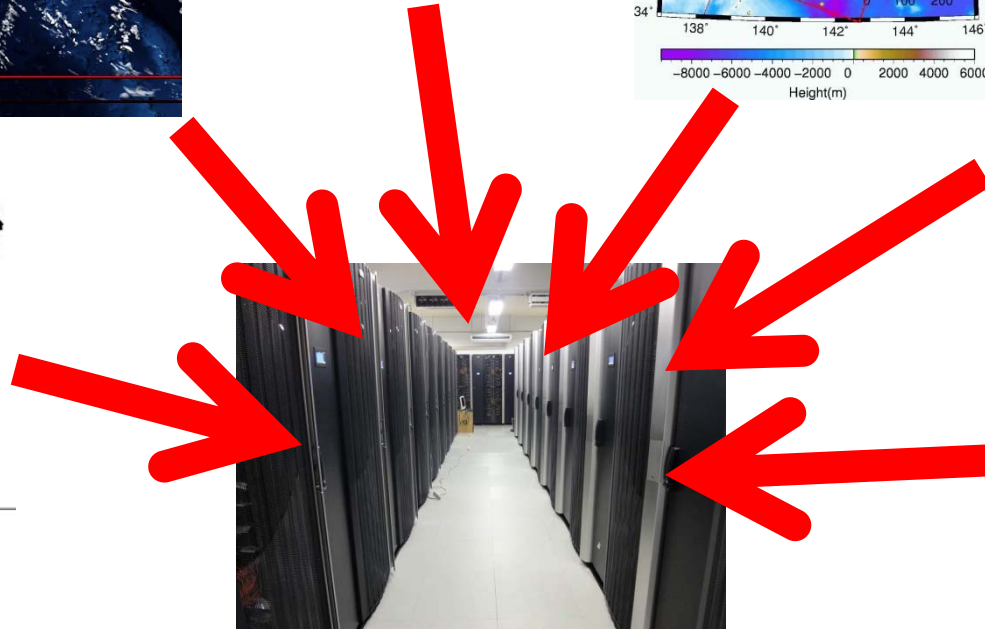
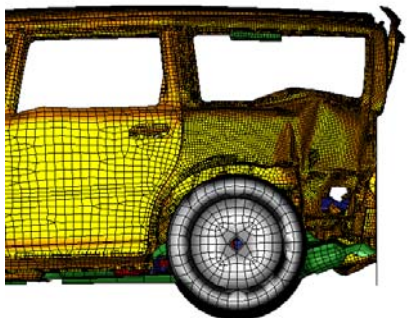
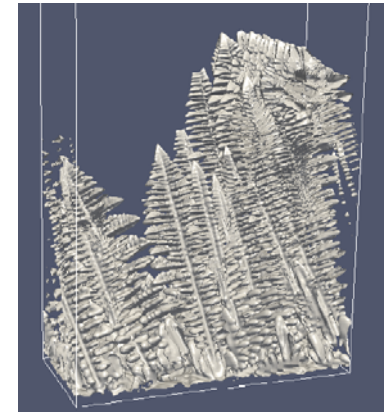
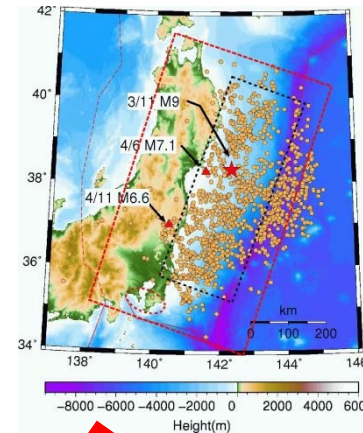
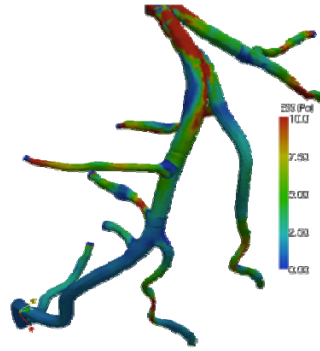
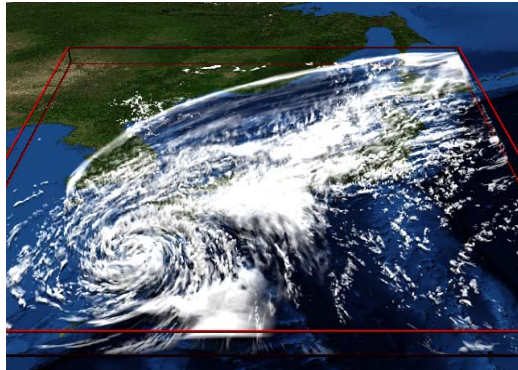
Jaguar



Cray-1 in 1976
160MFlops

スパコンは何に使われる？

スパコンはあらゆる科学分野の**仮想実験場**



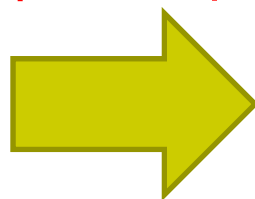


普通のコンピュータとの違い

- スパコンは、普通のコンピュータよりも**ずっと計算速度が速い**コンピュータ
 - 速度の単位はFlops: 一秒間に加・減・乗を何回できるか
 - 倍精度浮動小数(double)の演算速度がポピュラー



約10万倍!



約20GFlops (一秒で200億回)

5.7PFlops (一秒で2400兆回)



1980年のパソコン



Cray-1 (1976)
160MFlops



現代の携帯電話

スパコン開発競争の激化—世界—スパコンの変遷



参考: www.top500.org

Linpack演算速度 (Gflops)

Jun-93	CM-5/1024	60	LANL	US
Nov-93	Numerical Wind		AL	Japan
Jun-94	XP/S140		NL	US
Nov-94	Numerical Wind Tunnel		IAL	Japan
Jun-95	↓	↑		Japan
Nov-95	↓			
Jun-96	SR2201			
Nov-96	CP-PACS			
Jun-97	ASCI			
Nov-97	↓			
Jun-98	↓			
Nov-98	↓			
Jun-99	↓	2121	↓	US
Nov-99	↓	2379	↓	US
Jun-00	↓	2379	↓	US
Nov-00	ASCI White	4938	LLNL	US
Jun-01	↓	7226	↓	US
Nov-01	↓	7226	↓	US
Jun-02	Earth-Simulator	35860	ES Center	Japan
Nov-02	↓	35860	↓	Japan
Jun-03	↓	35860	↓	Japan

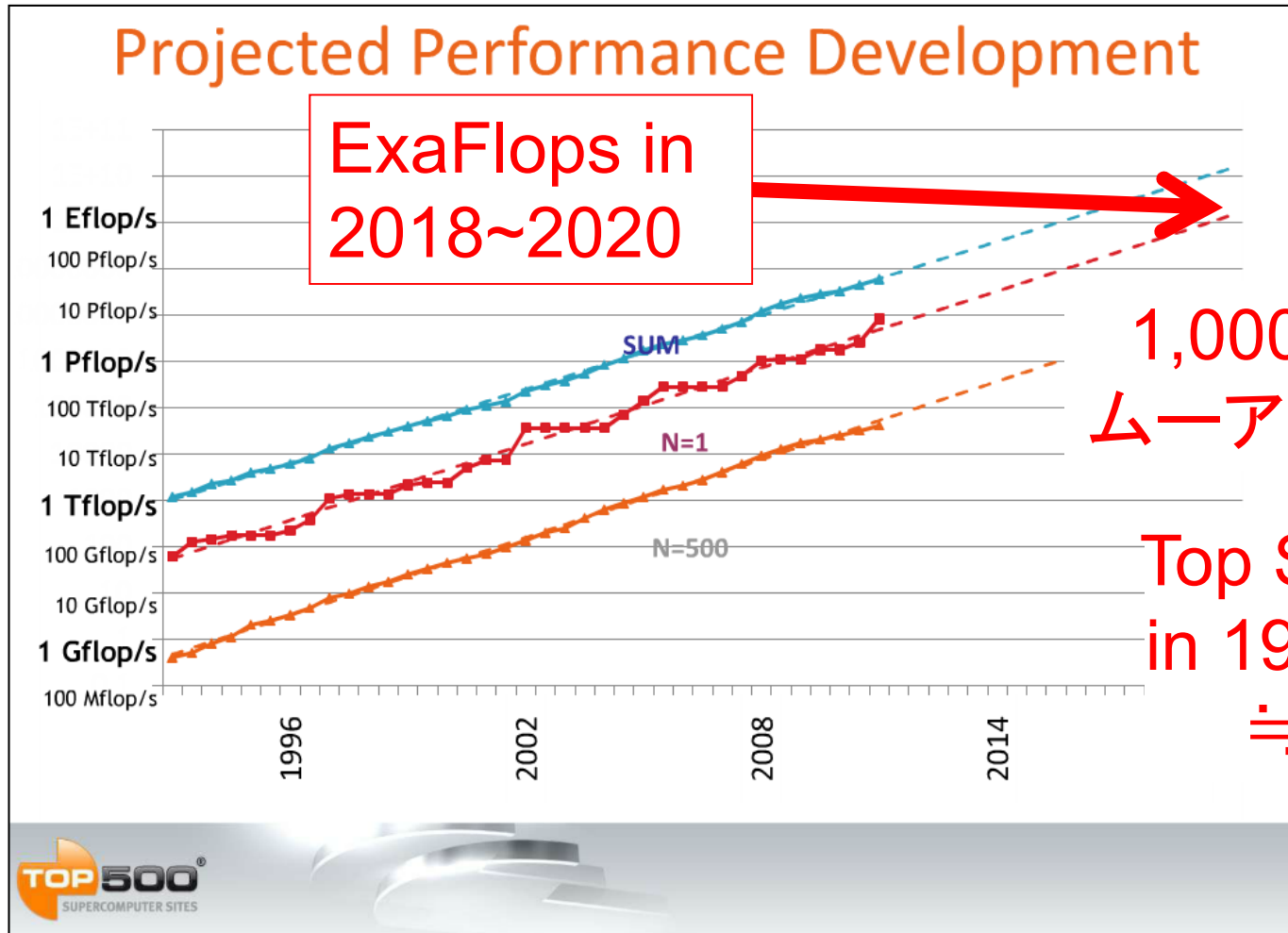
60GFlops

20年間で性能500,000倍
アメリカと日本の一騎打ち
⇒中国の台頭

Nov-03	↓	35860	↓	Japan
Jun-04	↓	35860	↓	Japan
Nov-04	BlueGene/L beta	70720	IBM/DOE	US
Jun-05	BlueGene/L	136800	DOE/NNSA/LLNL	US
Nov-05	↓	280600	↓	US
		280600	↓	US
		280600	↓	US
		280600	↓	US
		478200	↓	US
		1026000	↓	US
		1105000	↓	US
		1105000	↓	US
Nov-09	Jaguar	1759000	ORNL	US
Jun-10	↓	1759000	↓	US
	Tianhe-1A	2566000	NSC	China
	K computer	8162000	RIKEN AICS	Japan
	↓	10510000	↓	Japan
	Keukaia	16324000	DOE/NNSA/LLNL	US
Nov-12	↓		DOE/SC/ORNL	US
Jun-13	Tianhe-2	33863000	NSC	China
Nov-13	↓	33863000	↓	China

33.9PFlops

Development of Supercomputers (from www.top500.org)



1,000x in 10 years
ムーアの法則より速い

Top Supercomputer
in 1993 (60GFlops)
≡ PC in 2011



なぜ計算速度が重要？

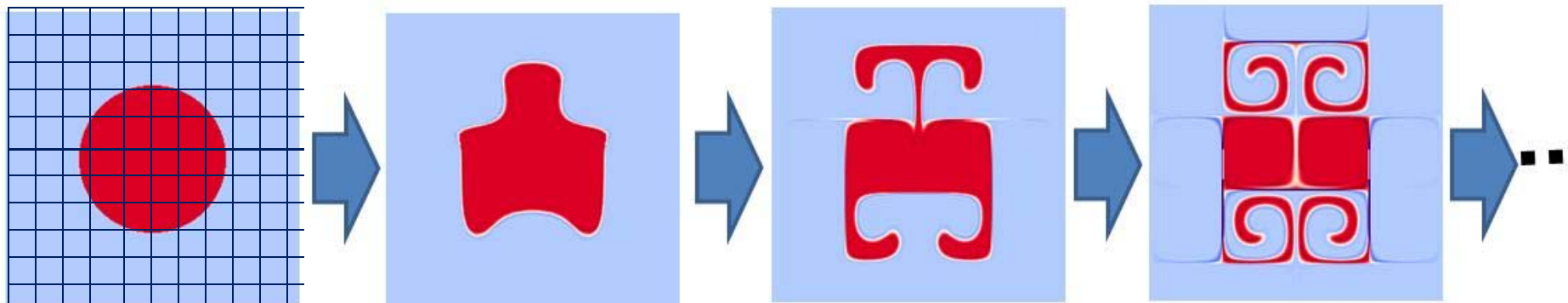


- 仮想実験(シミュレーション)のためには、**ばく大な量**の計算を、**タイムリー**にこなさなければならないから
⇒明日の天気計算に、一年かかっては意味がない！

天気予報はどうやって計算？

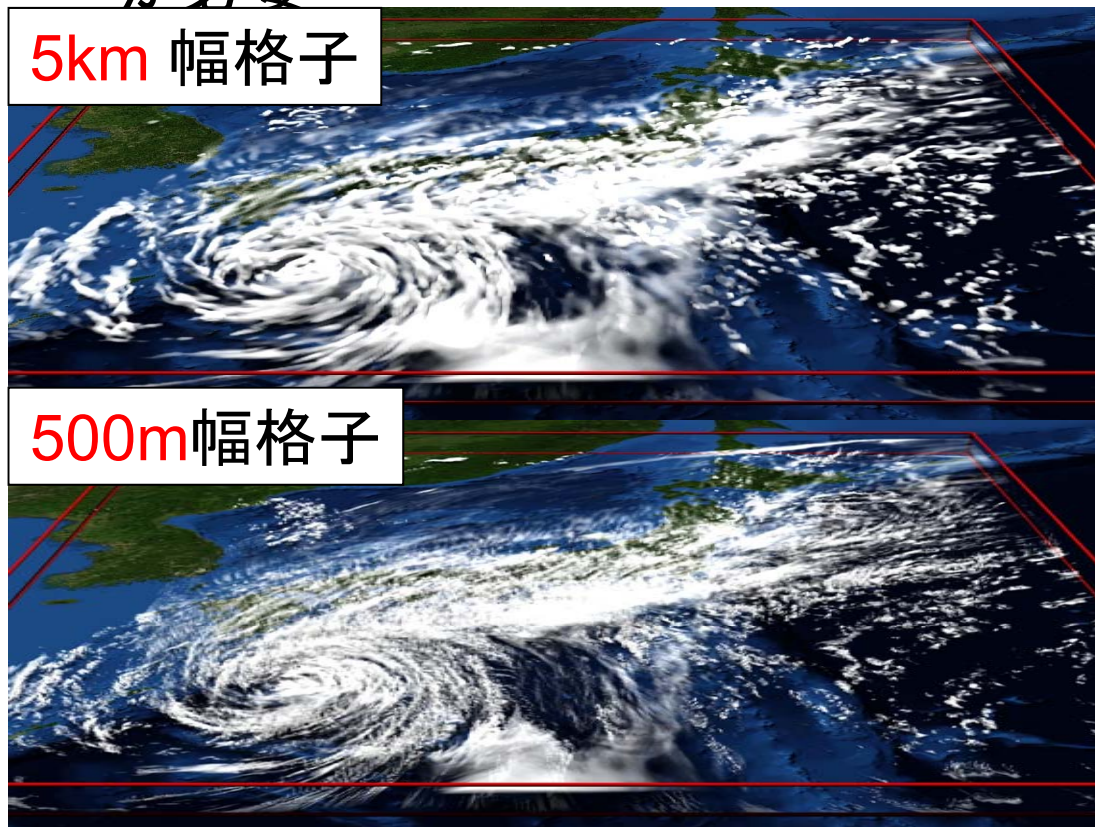
空間を、細かいマス目に分割 ⇒ 1点ずつ計算が必要
ある瞬間の計算が終了したら、次の瞬間の計算へ

- ・ 因果律：過去⇒現在⇒未来⇒もっと未来
- ・ パラパラ漫画のように計算を続ける



なぜ計算速度が向上し続ける必要？

- より厳密なシミュレーションを行うには、細かい解像度が必要

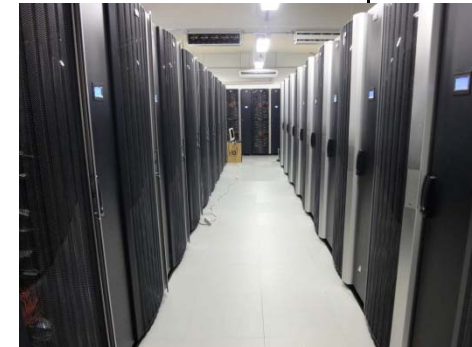


より細かい雲の
挙動が把握可能に！

10倍の高解像度のためには
X方向10倍 × Y方向10倍 × Z方向10倍 × 時間方向10倍 = **10000倍**
の計算をこなす必要！

一般的なスパコンの構造

- システム = 多数の**計算ノード** + **外部ストレージ**
 - パーツ間は**ネットワーク**で接続
- 計算ノード = 1以上の**プロセッサ** + **メモリ** + **ローカルストレージ**
 - パーツ間は**PCI-e, QPIなどの通信路**で接続
- プロセッサ = 1以上の**コア** (+ **L3 キャッシュ**等)



スパコンの計算性能は何で決まる？



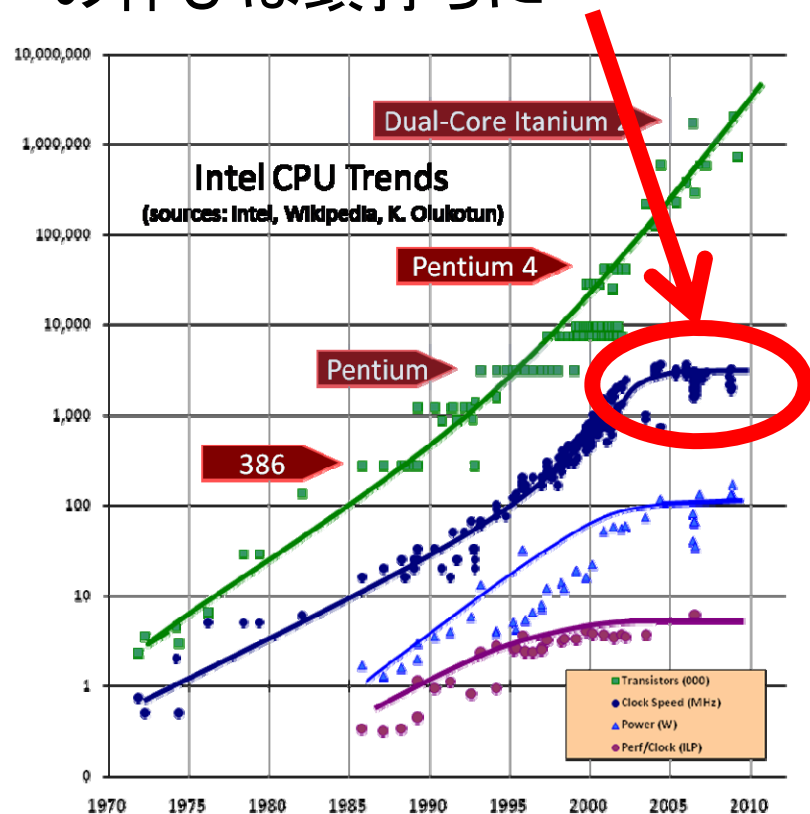
理論ピーク演算性能:

システムが仮に浮動小数演算のみを続けたときのFlops値。
実効演算性能とは区別が必要。以下の積となる:

- クロック周波数(Hz=1/sec)
 - 1~3GHz程度。2003ごろより頭打ち
- 1クロックあたりの同時計算数(flop)
 - TSUBAME2のCPUでは4, Sandy Bridge世代(AVX)で8
 - 単精度はその2倍
- プロセッサあたりのコア数
 - 4~16程度, これからも伸びる見込み
- 計算ノードあたりのプロセッサ数
 - 1~4程度
- 計算ノード数
 - TSUBAME2では1400, 京コンピュータでは88000

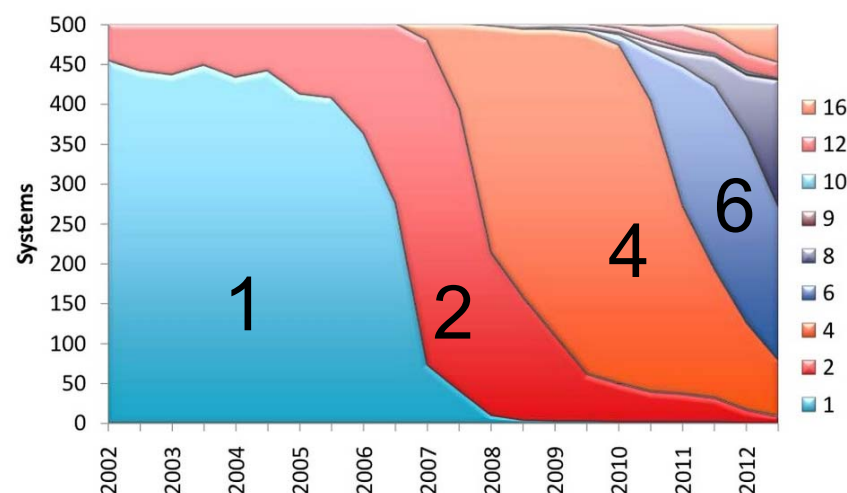
近年のスパコンの動向

2003年ごろからクロック周波数の伸びは頭打ちに



プロセッサあたりのコア数を増やして性能を稼ぐ方向に

Cores per Socket



90年代までの常識:「5年待てば私のプログラムは速くなる」
今の常識:「並列プログラミング覚えないと！」

GPUやアクセラレータ



- 「周波数はあまり高くせず、並列度で稼ぐ」という方針をさらに推し進めたのが、GPUやアクセラレータ
- GPU(graphic processing unit)はもともと画像出力用専用プロセッサだが、演算へ転用(GPGPU)



Intel Xeon X5670

$$2.93 \text{ GHz} \times 4 \text{ Flop} \times 6 \text{ core} = 70.4 \text{ GFlops}$$



NVIDIA Tesla K20X

$$0.73 \text{ GHz} \times 128 \text{ Flop} \times 14 \text{ SM} = 1310 \text{ GFlops}$$

- TSUBAME2.5, Titan(2012 No.1), Tianhe-1A(2010 No.1)などが採用
- NVIDIA GPU, AMD/ATI GPU, Intel Xeon Phiなど



- ここで、「TSUBAME2.0利用ガイダンス」の説明



本授業では

- 本授業の受講者は t2g-ppcomp ユーザグループに入ってもらう予定
 - t2subコマンドで “-W grouplist=t2g-ppcomp”
 - ちょっと大きなプログラムも動かせるように
- 授業に関する質問は担当教員あてに
 - 「レポートの×切はいつですか？」などをsoudan MLに投げないように

次回までに準備してください



- ノートPCを持参のこと / Please bring your laptop PC
 - 無線ネットワーク(Wifi)が使えること
 - titech-pubnetの使い方を調べてください
 - SSHの使える端末ソフトをインストールすること
 - WindowsノートPCなら、Putty, Teraterm+SSHなど
- TSUBAME2のアカウントを作ってください / Please make your account on TSUBAME2
 - 東工大ポータル ⇒ TSUBAME2.0利用ポータルより
 - 作成後、下記をメール連絡

Subject: TSUBAME2 ppcomp account

To: ppcomp@el.gsic.titech.ac.jp

- 専攻・研究室
- 学年
- 氏名
- アカウント名

TSUBAME情報源



- <http://tsubame.gsic.titech.ac.jp>

- マニュアル
- 統計情報(混雑具合など)
- メンテナンス予定

- スパコン相談ML:

- soudan@cc.titech.ac.jp

授業に関することは遠藤など
教員に！

TSUBAME計算サービス
High-Performance Computing at Tokyo Tech

メニュー

- Top
- お知らせ
- FAQ
- 利用について
 - TSUBAME2計算サービス種別
 - アカウント取得方法
 - 有料サービス利用方法
 - ログイン方法
 - TSUBAMEポータル
 - 各種利用の手引き
 - TSUBAME1利用者の方へ
- TSUBAME2システム構成
 - TSUBAME2ハードウェア構成
 - TSUBAME2ソフトウェア構成

重要なお知らせ

- 2011.10.02 現在のTSUBAMEの運用状況はこちら(グランドチャレンジ実施中)
- 2011.08.15 平成23年度秋期グランドチャレンジ制度実施に伴うサービス休止のお知らせ(9/29更新)
- 2011.07.22 グループディスク課金本運用開始について
- 2011.06.27 7月以降の夏季のTSUBAME運用・停止について(9/29更新)

最新のお知らせ

- 2011.9.27 : 【障害報告】2011.9.16発生: /work0ストレージ障害
- 2011.9.22 : Hキュー予約システムメンテナンスのお知らせ(9/26実施)
- 2011.9.22 : Yキューの終了について
- 2011.9.13 : 【障害報告】2011.9.12発生: ジョブスケジューラ障害
- 2011.9.13 : 【障害報告】2011.9.9発生: ジョブスケジューラ障害
- 2011.9.06 : Materials Studio 5.5.3, Discovery Studio 3.1公開、学内配布開始について
- 2011.9.02 : 【障害報告】2011.9.2発生: ジョブスケジューラ障害



連絡先/Contact

- 授業全体に関すること
 - ppcomp@el.gsic.titech.ac.jp (遠藤あて)
- ノートPCの準備が困難な場合も、早急宛先に上記へ相談のこと
- 東工大OCW
 - <http://ocw.titech.ac.jp> → 情報理工学研究科 → 実践的並列コンピューティング (平成26年度前期)