


東工大の TSUBAMEのご紹介

Tokyo-tech
Supercomputer &
UBiquitously
Accessible
Mass-storage
Environment)

- ・ 東工大  のスーパーコンピュータ
(最高で)
- ・ 世界ランキングで**第？位(来年ご期待！)**
- ・ 世界グリーンランキングで
ここ数年間、運用スパコン世界**第1位**

グリーンIT

人類の未来のための
情報処理技術

おまけ: 東工大の TSUBAME

なぜスパコン??

大きな数の呼び方

K	キロ	1,000	千
M	メガ	1,000,000	百万
G	ギガ	1,000,000,000	十億
T	テラ	1,000,000,000,000	兆
P	ペタ	1,000,000,000,000,000	兆
E	エクサ	1,000,000,000,000,000,000	兆

注) 1京(けい) = 10¹⁶

開発費 1000 億円~

スーパーコンピュータ

その時代の通常のコンピュータに比べはるかに高度な計算能力(とくに非常に高速な計算能力)を持つコンピュータ

FLOPS

1秒間に実行できる演算回数

注) 演算 = 小数の掛け算

注) 理論値と実測値で差がある

速度だけでないよ!

演算速度を比較してみると

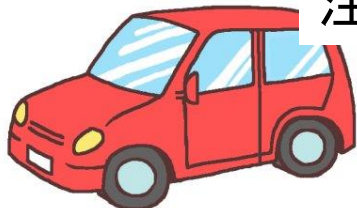
携帯 0.6G FLOPS ?

パソコン 50G FLOPS

スパコン 100T ~ 3P FLOPS

2000倍~

注) 1976年のスパコンは 0.16 G FLOPS



スパコンの利用法

⇒ シミュレーション

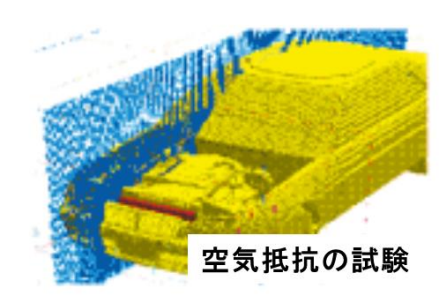
||
コンピュータ上での擬似実現

シミュレーションって何のため
近場の関係はわかっている.
but 全体での振る舞いがわからない
⇒ 全体の振る舞いを見る
パラメータ間の関係を見る

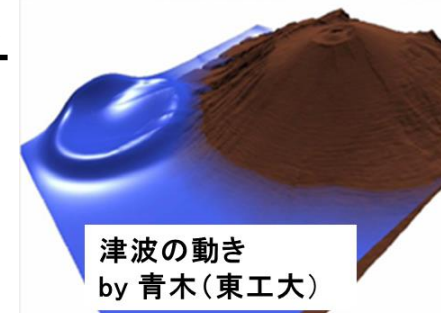
なぜスパコン??

なぜ, さらに速くしたい

(例1) 何度もテストしたい 100 倍速くなると
1 時間 \times 1万回 = 400 日 \longrightarrow 4 日

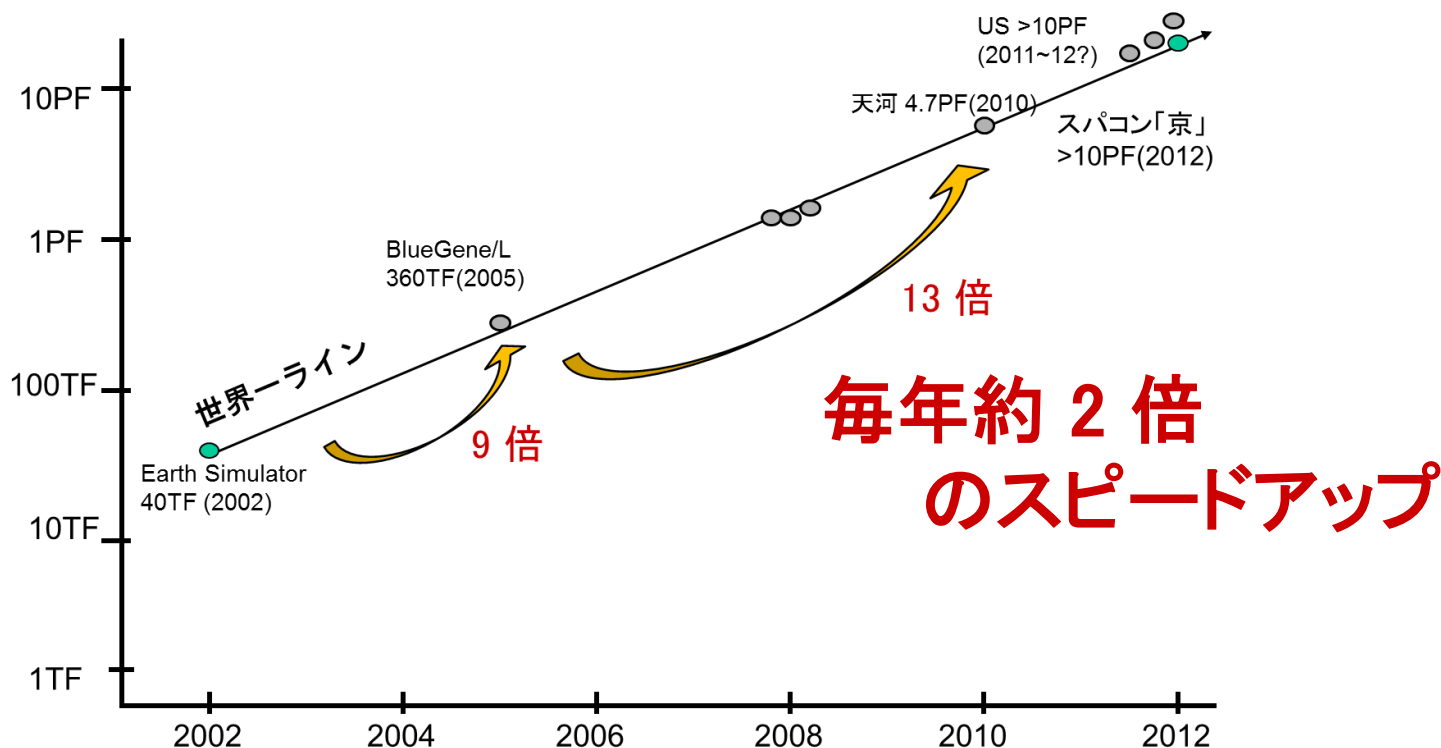


(例2) 精度を上げたい 100 倍速くなると
5 km \times 5km の格子 \longrightarrow 500m \times 500m の格子
同じ時間で!



おまけ: 東工大の TSUBAME

スパコンの計算速度の驚くべき進化

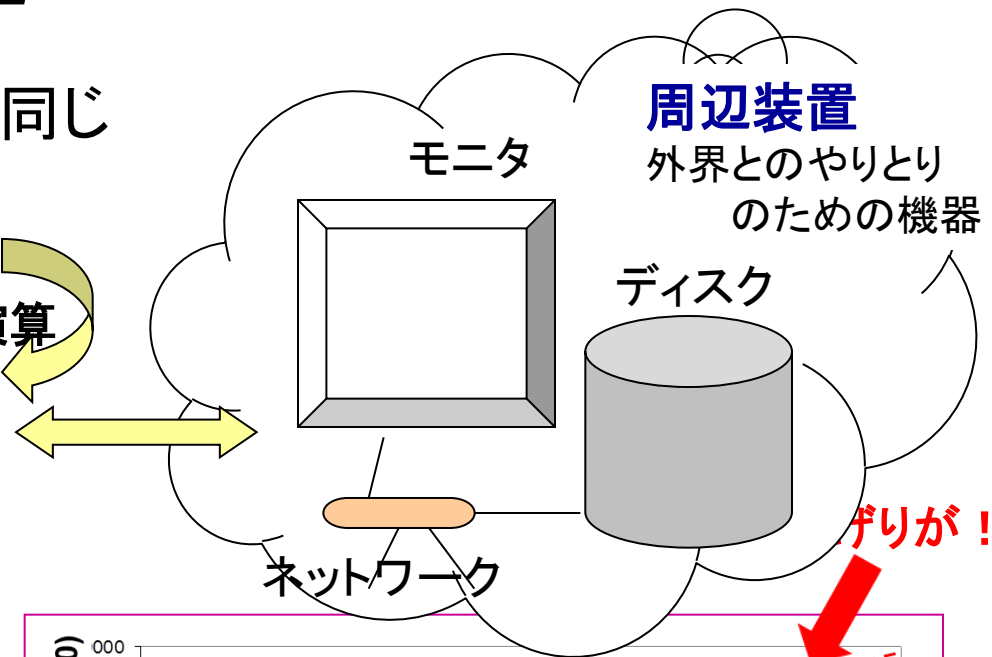
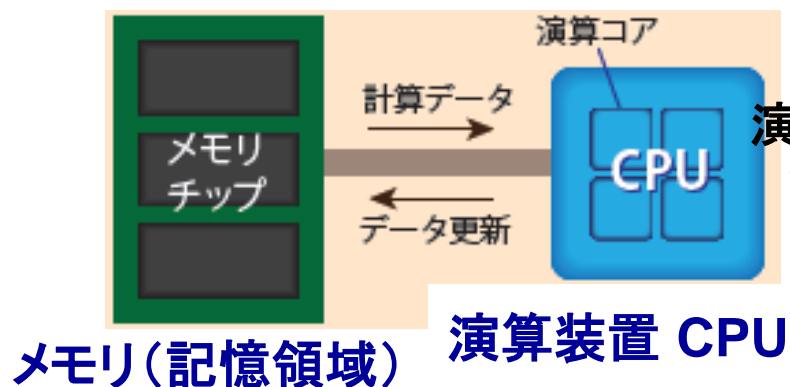


様々な工夫, 技術革新の成果!

おまけ: 東工大の TSUBAME

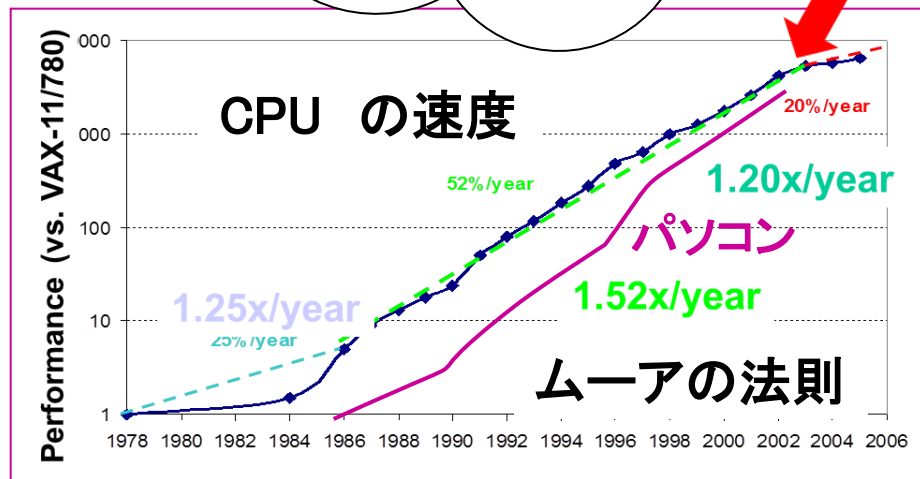
そもそもスパコンってなぜ速い？

PC もスパコンも原理は同じ



スパコンでは超高級品を使用

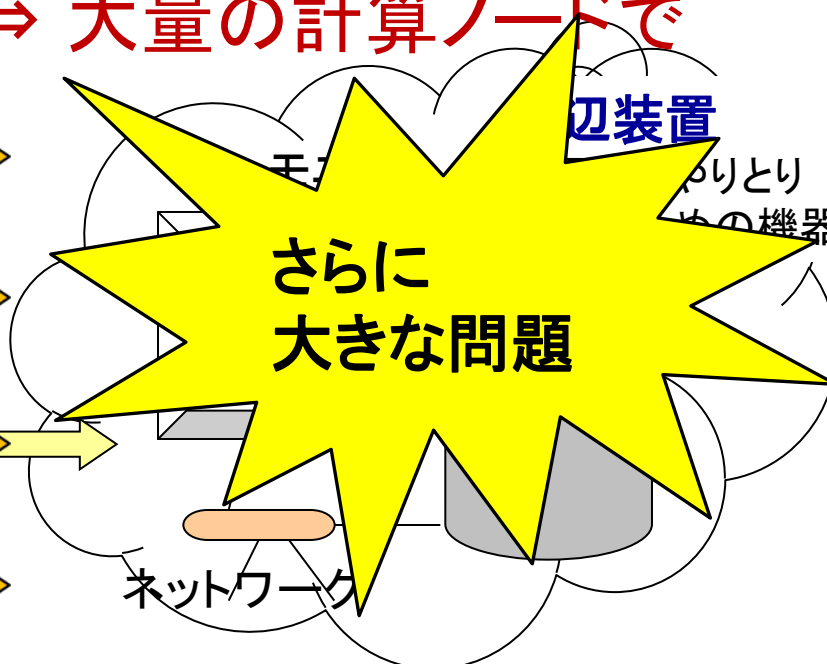
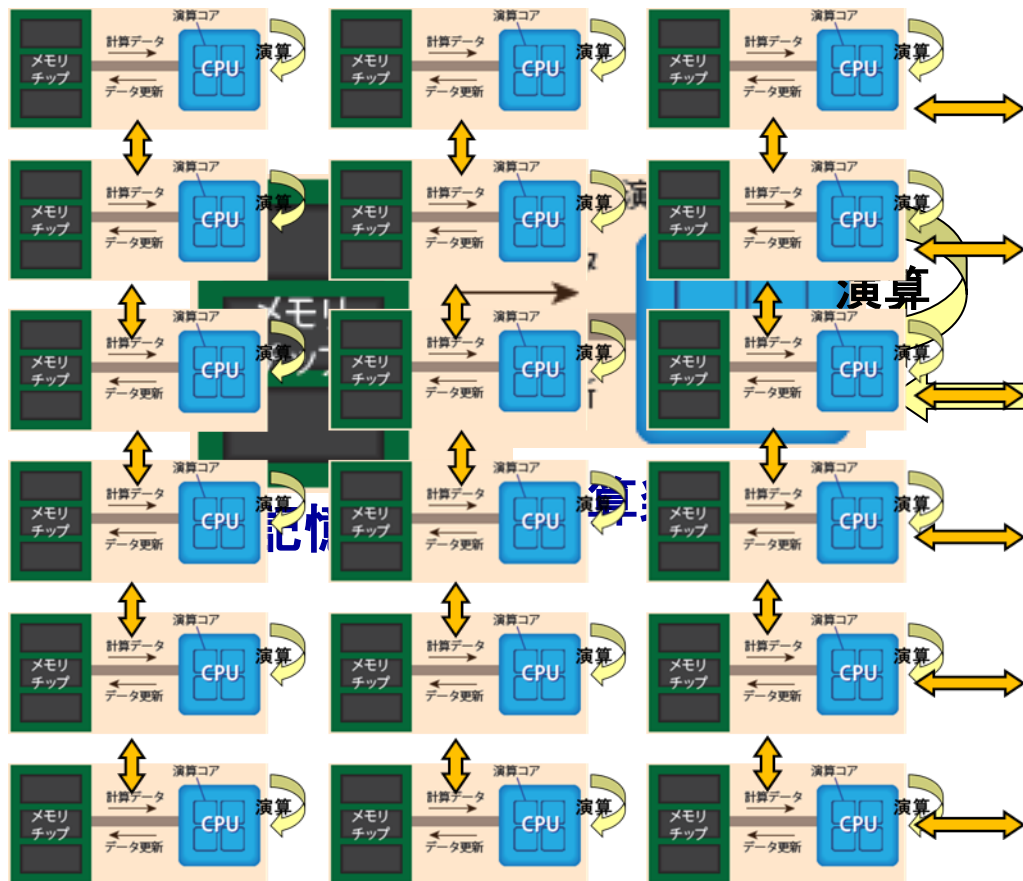
昔はよかったのお **一昔前の話 ~ 2000年**



おまけ: 東工大の TSUBAME

じゃあ、どうやって高速化！？

⇒ 大量の計算ノードで



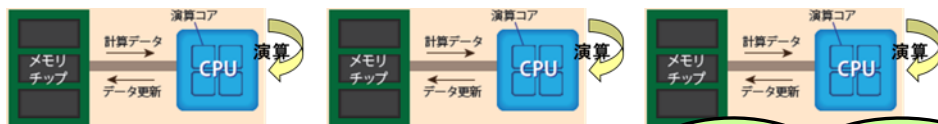
- ▶ 千個ならば千倍？
No! 要: 高度な技術
- ▶ PCを沢山並べたら？
No! 要: 超高速通信

並列処理 = みんなで手分け！

おまけ: 東工大の TSUBAME

じゃあ、どうやって高速化！？

⇒ 大量の計算ノードで



東工大では
この問題に
正面から取り組みました

さらに
大きな問題

体育館並み？
専用発電所も！

巨大化！ 大消費電力！



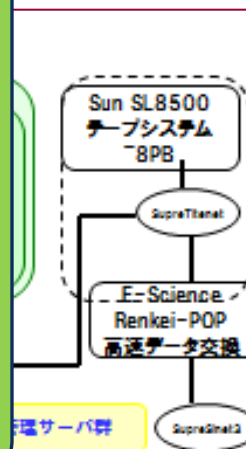
TSUBAME2.0 のご紹介

いまは TSUBAME2.5 だよ!



3つの特徴

1. 日本初のペタコン
2. 世界一グリーンなマシン
3. みんなのスパコン



計算ノード: 2.4PFlops (CPU+GPU), 224.69TFlops (CPU), ~100TBメモリ, 200TB SSD



ThinPEL/ノード
HP製CPU搭載サーバ 100nodes
CPU Intel Westmere-EP 2.93GHz (Turbo boost 3.199GHz) 19Core/node
Mem:55.9GB (+40GB) 102GB (+40GB)
GPU NV410 10200 5150Flops/3GPU/node
SSD 60GB x 2 120GB @55.9GB/40GB/node
120GB x 2 240GB @103GB/40GB/node
OS: Suse Linux Enterprise Server
Windows HPC Server
CPU Total: 215.99TFLOPS (Turbo boost 3.199GHz)
CPU+GPU: 224.69TFlops
Memory Total: 80.55TB (CPU) + 12.7TB (GPU)
SSD Total: 172.88TB

MediumPEL/ノード
HP製Socketサーバ 50nodes
CPU Intel Nehalem-EX 2.0GHz 32Core/node
Mem:127GB (+10GB)
SSD 120GB x 4 480GB
OS: Suse Linux Enterprise Server
CPU Total: 8.07TFLOPS

FatPEL/ノード
HP製Socketサーバ 10nodes
CPU Intel Nehalem-EX 2.0GHz 32Core/node
Mem:270GB (+25GB) @6nodes
59GB (+10GB) @4nodes
SSD 120GB x 4 480GB
OS: Suse Linux Enterprise Server
CPU Total: 2.56TFLOPS

PCI-E gen2 x16 x2/10node
QDR InfiniBand Tera 5-10700GPU



**~50 compute racks + 6 switch racks
Two Rooms, Total 160m²**

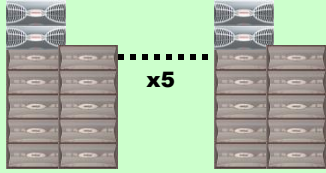
1.4MW (Max, Linpack), 0.48MW (Idle)



TSUBAME2.0 システム概念図

ペタバイト級HDD ストレージ: Total **7.13PB** (Lustre+ home)

並列ファイルシステム領域
5.93PB



MDS,OSS
HP DL360 G6 30nodes
Storage
DDN SFA10000 x5
(10 enclosure x5)
Lustre(5File System)
OSS: 20 OST: 5.9PB
MDS: 10 MDT: 30TB

OSS x20 MDS x10

ホーム領域
1.2PB



Storage Server
HP DL380 G6 4nodes
BlueArc Mercury 100 x2
Storage
DDN SFA10000 x1
(10 enclosure x1)

NFS,CIFS用 x4 NFS,CIFS,iSCSI用 x2

Sun SL8500
テープシステム
~8PB

SupreTitenet

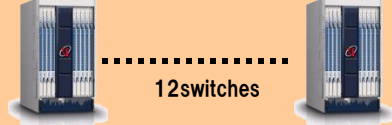
E-Science
Renkei-POP
高速データ交換

SupreSinet3

管理サーバ群

ノード間相互結合網: フルバイセクション ノンプロッキング 光 QDR Infiniband ネットワーク

Core Switch



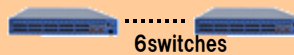
12switches
Voltaire Grid Director 4700 12switches
IB QDR: 324port

Edge Switch



179switches
Voltaire
Grid Director 4036 179switches
IB QDR : 36 port

Edge Switch (10GbE port付き)



6switches
Voltaire
Grid Director 4036E 6 switches
IB QDR:34port
10GbE: 2port

計算ノード: **2.4PFlops (CPU+GPU)**, **224.69TFlops CPU**, **~100TBメモリ**, **~200TB SSD**

Thin計算ノード



1408nodes (32node x44 Rack)

HP製GPU搭載サーバ 1408nodes
CPU Intel Westmere-EP 2.93GHz
(Turbo boost 3.196GHz) 12Core/node
Mem:55.8GB (=52GiB)
103GB (=96GiB)
GPU NVIDIA M2050 515GFlops,3GPU/node
SSD 60GB x 2 120GB ※55.8GBメモリ搭載node
120GB x 2 240GB ※103GBメモリ搭載node
OS: Suse Linux Enterprise Server
Windows HPC Server

CPU Total: 215.99TFLOPS (Turbo boost 3.196GHz)
CPU+GPU: 2391.35TFlops
Memory Total:80.55TB (CPU) + 12.7TB (GPU)
SSD Total: 173.88TB

Medium計算ノード



HP製4Socketサーバ 24nodes
CPU Intel Nehalem-EX 2.0GHz
32Core/node
Mem:137GB (=128GiB)
SSD 120GB x 4 480GB
OS: Suse Linux Enterprise Server
CPU Total: 6.14TFLOPS

Fat計算ノード



HP製4Socketサーバ 10nodes
CPU Intel Nehalem-EX 2.0GHz
32Core/node
Mem:274GB (=256GiB) ※8nodes
549GB (=512GiB) ※2nodes
SSD 120GB x 4 480GB
OS: Suse Linux Enterprise Server
CPU Total: 2.56TFLOPS

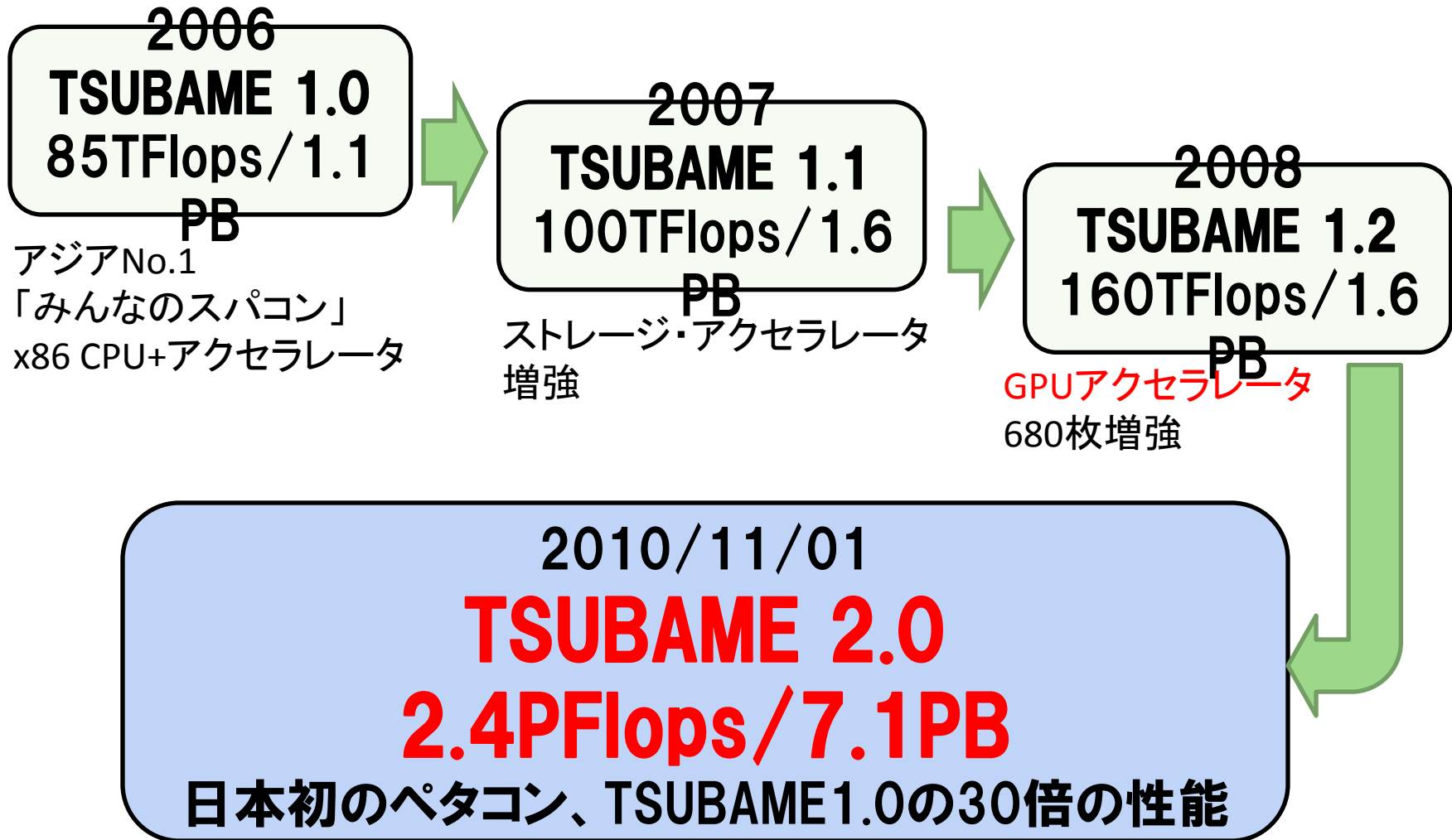
PCI-E gen2 x16 x2slot/node

GSIC:NVIDIA Tesla S1070GPU





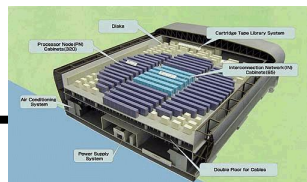
TSUBAMEの歴史



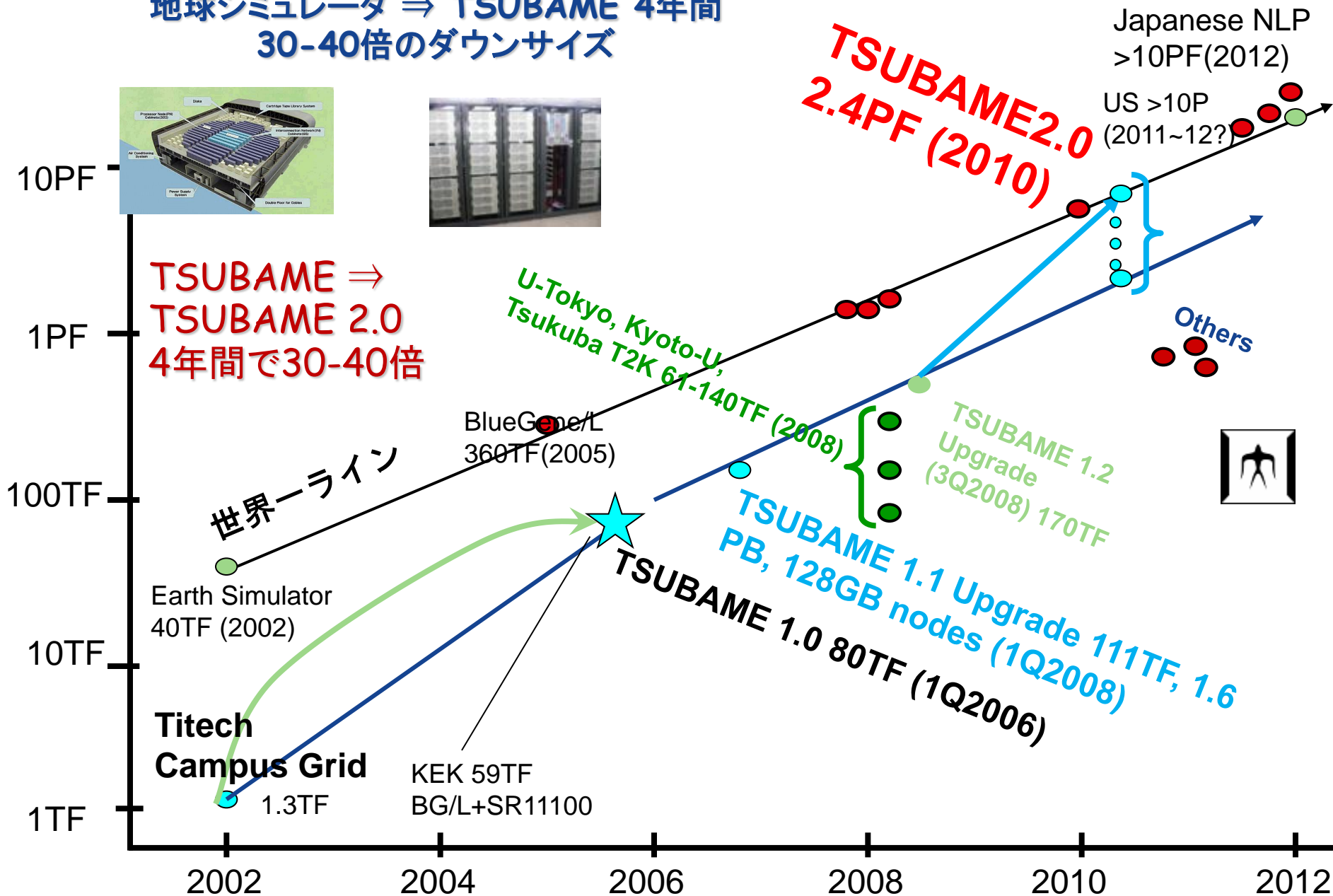
- TSUBAME初の完全リプレイス

TSUBAME2.0の性能向上

地球シミュレータ ⇒ TSUBAME 4年間
30-40倍のダウンサイズ

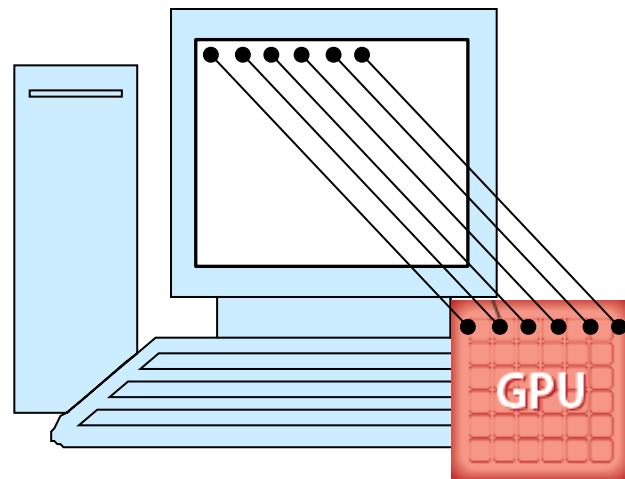
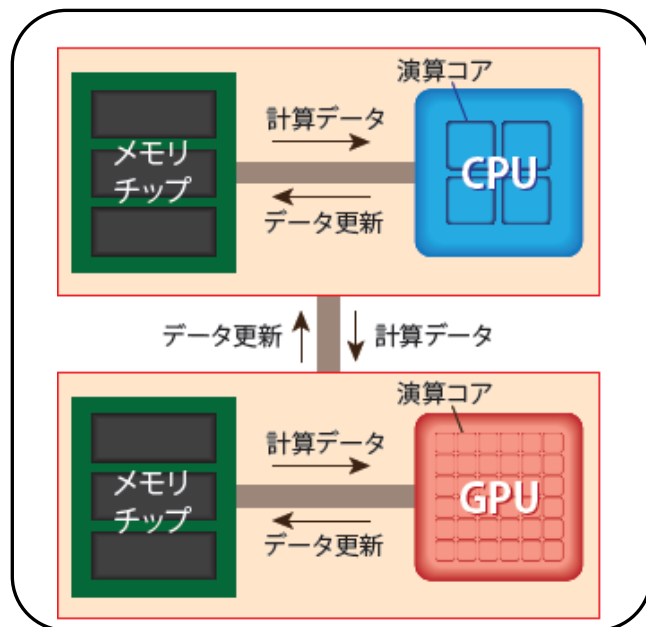
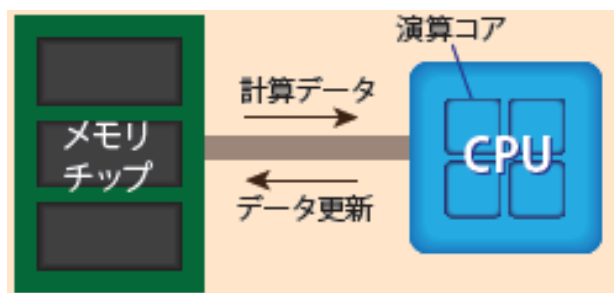


TSUBAME ⇒
TSUBAME 2.0
4年間で30-40倍



グリーン世界一の秘密

その1: GPU の本格導入



GPU (graphic processing unit) 画面の制御のために開発された演算装置. 多数の演算コアが並列処理を行う設計.

CPU よりちょっと劣るけど...
1 GPU に ~~448~~ コアは使える!
2688

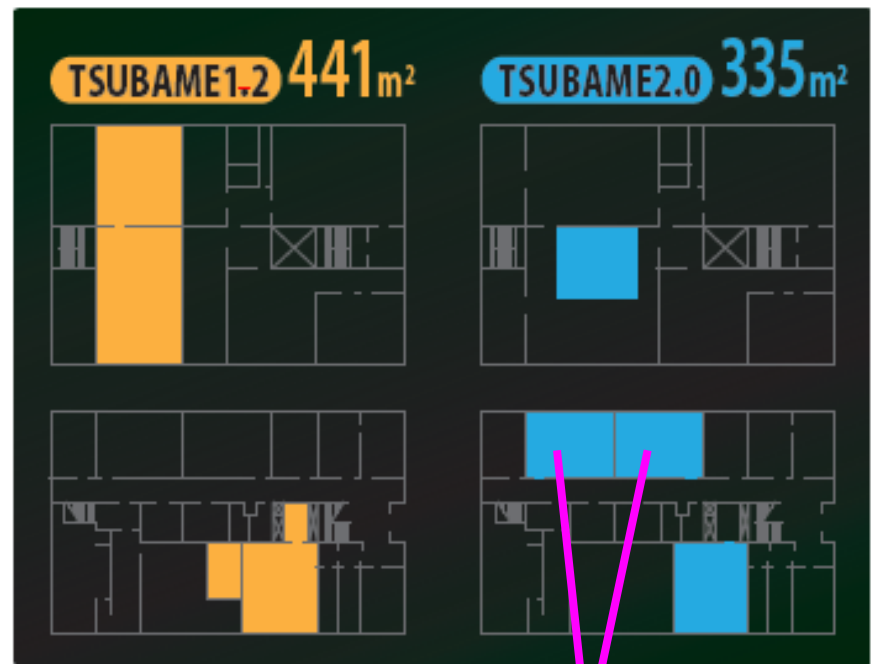
スパコンに導入したのは
TSUBAME 1.0 が世界初

グリーン世界一の秘密

その1: GPU の本格導入



これ1つで
地球シミュレータと同等



スパコン本体は教室2つ分でOK

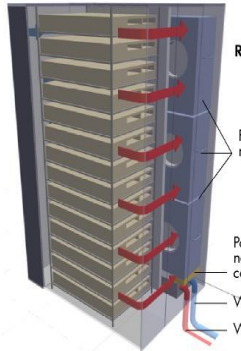
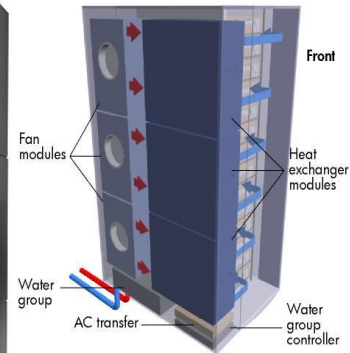
グリーン世界一の秘密

その2: 高密度実装・水冷キャビネット冷却

ラック内に熱交換システムを内蔵した密閉型水冷システム



TSUBAME3.0 に向けて
新たな冷却システム実験



みんなのスパコン

学部1年生から

学生証

→ TSUBAME

25Gバイトフォルダ



無線 LAN で
学内のどこからでも



スーパーコン

SuperComputing Contest

対象: 高校生・高専生

時期: 8月下旬 (**22年目!**)

場所: 東工大, 阪大



2011	開成高校	甲陽学院高校	早稲田高校
2012	灘高校	開成高等学校	栄光学園高校
2013	久留米工業高専	麻布高等学校	早稲田高校
2014	大阪府立大学工業高専	渋谷教育学園渋谷高校	香川県立高松高校
2015	筑波大学付属駒場高校	久留米工業高専	明石工業高専
2016	久留米工業高専	慶應義塾高校	奈良工業高専

おしまい