

スーパーコンピューティングの 将来

牧野淳一郎
東京大学理学系研究科
天文学専攻

詳しくは

[http://jun.artcompsci.org/
articles/future_sc/face.html](http://jun.artcompsci.org/articles/future_sc/face.html)

どういいう話

をすするか。

ベクトルを

買ってはい

けなさい。

何故話をす

るか。

和田君に

命令された。

スパコンと

は何か。

歷史

1960年代

単に、

高くくて速い

計算機

代表例:

CDC 6(7)600

(Cray 設計)

1970年代

ベクトル

プロセッサ

Cray-1, Cyber

1980年代

いろいろしいるあ

った

日本のベクトル

レイの並列

ベクトル

並列計算機

1990年代

ベクトル

衰退

並列計算機

会社倒産

PC

クラスタ

生き残る

何故そうな

ったか？

ベクトルは

高くなった

並列計算機

やっぱり

高くなった

比較対象

PC

クラスタ

高い | 計算機と

安い | 計算機を

比べると

Cray-1

100Mflops

10M\$

PDP-11/70

10kflops?

50K\$?

Cray-1

50倍お得

1985

Cray XMP

1 Gflops

10M\$

PC-AT

30kflops?

5K\$

Cray-XM1P

20倍お得

1995

VPP-500

100 Gflops

30M\$?

Dec Alpha

300Mflops

30K\$

VP P

3倍損

2005

SX-8

10TF

50M\$?

Intel PD

12 Gflops

1K\$

SX-8

60 倍損

スカラ並列

はどうか？

日立

SR-11000

(IBM製)

SX-8 と変

わらない

東大、北大

x86 サーバ

はどうか？

1 ノード

100 万円

SX より

10 倍まし

普通のPC

より5倍損

ベクトルは

何故高くな

ったか

メモリシステム

容量が高い

ランダムア

クセスでき

る

超多バンク数

PC の

1000 倍

バンド幅も大

きい

PC の 10 倍

共有メモリに

してあるのでも

っと高い

売れないから

高い

というのも

スカラ並列は
何故高いか

チップ当りで

ベクトルより

遅い

共有メモリな

ので高い

売れないから

高い

x86 サーバは

何故高いか

高くても買う

人が少しはいい

るから

儲けは大きい

保守とかでお

金がかかる、

ともいう

信頼性が高い

というのは嘘

計算機は高い

ほど良く壊れ

る

作る台数が少
ないと問題点
をつぶせない

日本のロケット
トが良く失敗
するのと同じ

PC クラスタ
を買えばいい
か？

それでもない

ネットワーク

どうするか

プログラムど

うやっで書く

か

MPI?

HPF?

ソフトウェア

分散共有メモ

リ？

MPI で書か
ない と性能で
ない

MPI なんかに
で書きたくない
いのが人情

別にサボリた

いいわけじゃない

い

MPI でプロ
グラムするのは
生産性低い

ちゃんとそう

いう研究結果

もある

同じことをす
るのに余計に
時間がかかる

サイエンスす

る時間が無く

なる

何故 MPI で

書かないかとい

けないか

そもそもそもそもコン

パイラがない

コンパイラが
あっても性能
でない

性能でないか

ら誰も使わな

い

誰も使わない

から性能上が

らない

惡循環

1990年代

並列言語結構

使えていた

CM-5

Cray T3E

PC クラスタ

では初めから

MPI とか

商用コンパイ

ラが高かった

せいも

性能でない1の
が大きい1

性能でない1の

は通信レイテ

ンシのせい1

Cray T3D

Shmem $3\mu\text{s}$

その頃の PC

500 μ s くらい

100 倍遅い

並列プログラ

ムのスケーラ

ビリティ

通信レイテン

シが効く

台数増えると

計算量はもち

ろん減る

通信量は普通

減る

通信回数と同

じが増える

レイテンシが
最後は効く

100倍遅いと

並列化は

「できない」

MPI で頑張
って通信回数
減らすと

ある程度は性

能がでる

コンパイラが
通信入れるの
では

なかななか**性能**

でない

どうすればいい

いか？

良くわからない

い

本来、計算機

屋さんが考え

るべきこと

速くて安い！ネ

ットワークを

作る

まともなコン

パイラを作る

でも、あんまり役に立つ話はない

ネットワーク

が遅いのはど

うして？

普通のPCク

ラスタ

実はソフトウェア

エアが遅いだけ

け

OS 層を全部

飛ばして

ネットワーク

カードを直接

アクセス

レイテンシ

5 μ s くらいに

はできる

似鳥君の修士

論文

1月末迄切

文章書いてる

かどうか心配

関係ない話で

失礼

でも、スイッチ
チがはいると
 $5\mu\text{s}$ 増える。

スイッチ重ね

るとやっぱり

遅くなる

専用

ネットワーク

カード

高いけれど速い

Quadrics

QsNet-II

多段スイッチ

通して $2\mu\text{s}$

GbE の 10

倍速い

値段も 10 倍

PC 1 台の 2

倍以上する

ネットワーク

つけると 3 倍

損

何故高いか？

台数が少ない
から。

ハードウェア

はGbEより

ずっと単純。

でも高い

それでも

ベクトルより

10倍まし

スカラ並列よ

りも10倍まし

MPI でも

ちよつとは楽

PC より安い

のはないか？

FPGA

Field

Programmable

Gate

Array

論理回路を書

き換えられる

LSI

ハードウェア

設計しない

いけない

数値計算向け

のソフトウェア

アない

PGR

濱田・中里

パイプライン

の演算器を書

く

後はソフトウェア

エアが自動で

なんかする

割合使えるか

も？

計算精度低い

ものでない」と

厳しい

ゲーム機

発表当時、

PS2 は使え

そうだった

実は使えなかつた

た

5年間新製品

がでない

スパコン並

5年たつと10

倍遅くなる

単精度しかな

い

丸めが変

使えなかつた

PS3, Xbox

も倍精度遅い

多分使えない

IBM BG/L

コロナビニア大

の QCD

専用計算機

ちよつとしいし

ったのが

BG/L

QCD 屋さん

はとっても偉い

高いネット

ワーク + PC

並の性能

電気は喰わない

い

スペースも喰

わなない

IBM が保守

にくる

計算センター

向け

プログラム書

くのはとても

大変

「性能でた」と

いいうのが論文

になる

Cray XT3

XD1

T3D の子孫

Opteron +

速いネット

ワーク

結構よさげ

でも高い

高い $\times 86$ サー

バの倍くらい

ベクトルに比

べると10倍い

い

GP GPU

グラフィック

カードを計算

に使う

单精度

値段が高い

結構疑問

GRAPPE は?

今作ってるの

は実は「汎用」

LSI 設計・試

作の費用が[”]高

く[”]な[”]った

GRAPPE-4

の時は 2500

万

GRAPPE-6

の時は 1億ち

よつと

今なんかが作る

と 4 億くら

い

IBM だと

10 億

特別推進でも

そんなの無理

天文専用

GRAPE で

予算とる

多分無理

GRAPE-DR

(GDR)

SIMID

のおお化け

1 チップに

512

プロセッサ

マイクロプロ

セッサの

200倍

どうしてそんな
なに入るか

元々プロセツ

サってそれく

らい

2 μ m でプロ

セッサ作れた

トランジスタ

数今の

1/500

沢山入れない

のは何故？

いれても性能

でないから

メモリバンド

が足りない

GDDR ではメ

モリバンド幅

は頑張らない

あまりいいらな

い計算がメイ

ン

GRAPPE を

エミュレーシ

ョン

行列計算もで

きる

SPH だって

できる

3体問題を沢
山並列に計算
できる

メモリアクセ

ス当りの計算

数が多いなら

結構性能でる

FFT は駄目

スキームによ
つては流体計
算もいけそう

チップ

単体性能

512Gflops

倍精度だと

半分

でも

256 Gflops

デーブアウト

した

もうすぐでき

るはず

システム完成

は 2009/3

振興調整費が

5年だから

完成した時の

速度 (予定)

2 Pflops

天文台の

VPP の

4000 倍

地球シミュ

レータの

50 倍

あまり本当と

は思えない

チップ 4000

個

512 ノード

クラスタ

ネットワーク

は未定

消費電力

300KW

電気代払えな

いかも

出来てから考

える

京速計算機

???

来年度予算は

ついでにらしい

35 億

何に使うのか

不明

文科省計画は

1100 億

2011/3 に

10P flops

「特定処理部」

先に完成

全体完成は

2013/3

ベクタ 0.5P

スカラ ー 2P?

特定 10P?

実施機関

理研

魔球の

姫野さん

「特定」って

何？

公式には決ま
ってない

非公式にも多

分決まっとな

い

まだ何ができ

るかわからな

い

来年度前半に

は見えてくる

かも

GRAPPE-DR

の技術をもと

に

とか書いてある
資料もある

3種混合を本
当にするかど
うかは

まだ不透明

大学センター

はどうなる？

多分つぶされ

る？

ディスク課金

が買うより高

いし、、、

そんな

センター

いらなしい

大学センター
生き残り作戦
例

東工大

SX-8

Sun Opteron

ClearSpeed

全部で

100T

天文台はどうか

するべき？

ベクトル

使い、やすい、け

ど遅い、

PC クラスタ

ピークは高い
けど性能で
とは限らない

GRAPPE

-DR

ピークはもっ

と高い

使いやすいさっ

て何？

究極の三択

良く考えまし
ょう

おしましい