



# Sun ONE Grid EngineとKNOPPIXを用いた 教育用PC群による 大規模グリッド構築に関する基礎的研究

岐阜高専 柴田良一



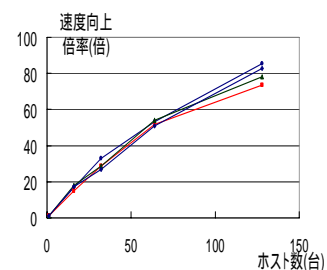
2004年2月20日  
グリッド協議会 第6回ワークショップ



## 発表のまとめ

**結論** Sun ONE Grid EngineとKNOPPIXを用いた  
グリッドシステムを開発し、教育用PC群  
による大規模グリッドの構築に成功した。  
(メガコンピューティングなのかな?)

**目次** 研究の背景や目的  
S1GEonKNOPPIXの開発  
グリッドの構成  
動作確認実験  
計算能力検証  
今後の展開とまとめ  
岐阜グリッドの紹介





## 共同研究チーム

|    |   |
|----|---|
| 構成 | 岐阜工業高等専門学校建築学科柴田研究室<br>柴田良一 + 渡辺輝樹 + 山口択 + 藤田憲<br>システム構築・実証実験作業<br>サン・マイクロシステムズ<br>林憲一 + 山方和昭 + 澁谷寿夫<br>グリッドエンジンの利用技術 |
| 協力 | 産業技術総合研究所情報処理研究部門<br>須崎有康<br>KNOPPIXに関する助言<br>豊橋技術科学大学建設工学系加藤研究室<br>加藤史郎 + 中澤祥二 + 島岡俊輔<br>建築構造解析への応用                  |

---



## 研究の背景

高精度かつ大規模な数値解析において  
計算能力に対する要求は留まらない  
HPCの1つとしてグリッドに期待

一般の研究者や技術者においても  
数値計算の重要性は変わらないが  
専用計算機の導入は不可能

|    |   |
|----|---|
| 着目 | 教育用演習PCやCAD業務用PCを転用し、<br>既存の限られた計算機を用いて<br>分散処理対応のグリッドが有効 |
|----|---|

---



## 教育用 P C 群の利用状況

### 現状

同じ仕様の P C が多数まとまって導入  
利用時間は授業時間に限定されている  
夜間や休日は全く利用されていない  
Windowsベースのシステムが大多数  
管理されソフトの追加は難しい  
導入後のデュアルブートは困難  
これらを解決すれば . . .



## 開発の前提

教育用 P C 群の遊休時間の有効活用  
既存システムの転用技術の開発  
1CD-LinuxとしてKNOPPIXを利用

手軽に利用できるグリッドシステム  
導入コストが低く効果が高い  
Sun ONE Grid Engine (S1GE)を利用

これらを組み合わせたシステムとして  
S1GEonKNOPPIXを開発  
[プロトタイプは林らによって開発]



## 開発の目的

数値解析ユーザーの立場で  
自分が必要としているものを作る  
これより  
「誰でも使えるグリッド」をめざす

計算能力が低くても：既存設備を使おう  
構築時間は短期間で：導入してすぐ活用  
並列処理は難しいよ：多数ジョブを分散  
仲間でしか使わない：ユーザ認証は簡易

その具体的な条件は . . .

---



## 具体的な条件：1

対象とする計算は？

有限要素法による数万自由度程度の数値解析  
を、パラメトリックに解析し、最大数万ケー  
スまでを処理する。  
多数のジョブを処理するスループットに期待

対象とする利用者は？

H P C 専用機を持たない技術者や研究者が、  
手元にある P C 群を活用する場面を想定。  
「簡単に構築、容易に運用、手軽に利用」を  
望んでいる数値解析ユーザー

---



## 具体的な条件：2

### 利用する機器は？

Windows系OSで利用されている教育用PC端末を利用する。ネットワークは、通常のLANを流用する。よって、導入コストは殆ど必要ない。

### 何が実現されるのか？

グリッドの導入により、教育用PC端末の遊休時間の有効利用が進み、潜在的な数値解析の需要に応えることが可能になる。

---



## PC群起動の条件

教育用PC群：多くがWindowsベースで動作  
一方

グリッドシステム：殆どがUNIXベースで動作

Windowsのままのグリッドは困難

Windows用グリッド?として

岐阜グリッドを開発し検証中

デュアルブートやエミュレータの導入

既存PCへは運用面作業面で無理

PC群のHDDを利用せず独立して起動する

1CD-LinuxによるCDブート

---



## 1CD - LinuxとしてのKNOPPIX

### KNOPPIXとは

CDのみでブート可能なLinuxディストリビューションで、ドイツKnopper氏が開発。  
日本語化は産総研須崎氏などが推進。  
<http://unit.aist.go.jp/it/knoppix/>

### 選定理由は

ハードウェアの認識能力が高い  
利用者が多く安定して動作している  
再構築の情報も入手可能

---



## グリッドソフトウェアの条件

開発の前提より  
システムを構築した本人や仲間が利用する  
パーソナルな利用が可能  
グリッドの目的は  
大規模数値計算の並列処理のピーク性能追及  
ではなく  
多数の中規模ジョブを処理するスループット  
に期待する

手軽に利用するためには無償で利用でき  
上位のグリッドとも連携可能なシステム

---



## グリッドとしてのS1GE

### S1GEとは

サン・マイクロシステムズが提供している  
グリッドソフトウェアで、標準版は無償。  
容易にインストール可能で高度な機能を持つ

<http://jp.sun.com/products/software/serverpref/gridware/>

### 選定理由は

標準版は無償で利用できる  
初心者にもわかりやすいGUIで操作可能  
計算ノードのシステム構築が容易  
Globusなどとの連携実績がある



## S1GEonKNOPPIXの概要：1

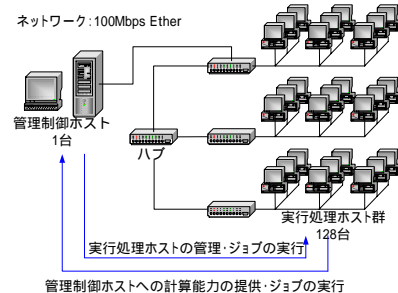
### グリッドとしてのS1GE

3種類のホストにより構成される

管理制御ホスト：グリッド全体を管理

実行依頼ホスト：ジョブの実行を依頼

実行処理ホスト：ジョブを引受け処理





## S1GEonKNOPPIXの概要：2

### 管理制御ホスト

NFSやグリッドのサービスを提供するため  
専用のPCを用いて構築した

### 実行処理ホスト

管理制御ホストに含める

### 実行処理ホスト

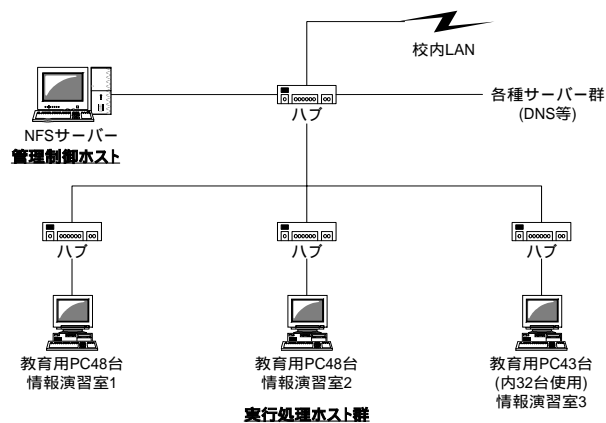
S1GEonKNOPPIX-CDにより起動する

---



## S1GEonKNOPPIXの概要：3

### ネットワーク構成





## 教育用PC群によるグリッド

### 利用したPCの仕様

#### 動作確認実験の構成

|         | 管理制御ホスト       | 実行処理ホスト       |
|---------|---------------|---------------|
| OS      | RedHatLinux8  | KNOPPIX3.3    |
| CPU     | Celeron433MHz | Celeron667MHz |
| RAM     | 256MB         | 128MB         |
| NETWORK | 100MEther     | 100MEther     |
| CD読込    | -             | 12倍速          |

#### 計算能力評価の構成

|         | 管理制御ホスト           | 実行処理ホスト       |
|---------|-------------------|---------------|
| OS      | RedHatLinux8      | KNOPPIX3.3    |
| CPU     | Athron1666MHz × 2 | Celeron667MHz |
| RAM     | 2048MB            | 128MB         |
| NETWORK | 100MEther         | 100MEther     |
| CD読込    | -                 | 12倍速          |



## S1GEonKNOPPIXの開発：1

### 管理制御ホストの構築

通常のLinuxBOXで構築するグリッドエンジンと同様に行う。詳しくは...

管理制御ホストにとっては実行処理ホストが、HDD上のLinuxでもCD上のKNOPPIXであっても影響は無い。ただし、実行処理ホストの登録については、工夫が必要。

S1GEの初心者向けガイドとして、技術評論社SoftwareDesignの平成16年1月号より連載中、ご参考までに



## S1GEonKNOPPIXの開発：2

### 実行処理ホストの構築

1CD-Linuxなので固定した情報から起動して  
個別の実行処理ホストにする工夫が必要

ポイントは、  
DHCPによる動的なネットワーク設定  
グリッドシステムのAutoMountの設定

この部分の開発にあたっては共同研究者  
のサン・マイクロシステムズ、  
林憲一、山方和昭、澁谷寿夫、らの助言  
に基づいて進めた。

---



## S1GEonKNOPPIXの開発：3

### 実行処理ホストの起動

現在のS1GEonKNOPPIXでは、  
グリッド起動時にCDを入れて起動する。  
多数の実行処理ホストを起動するのは手間  
が掛かりそうだが、単に電源投入時にCD  
を入れるだけなので、それほどでもない。  
(スタッフ学生はそうは思っていないかも)

既存PCに極力手を入れない方針では、  
この方法はそれなりに有効だと思っている。

よい方法があれば、ご助言をお願いします。

---



## 動作確認のためのパラメータ

### 動作確認用のジョブ

S1GEで動作確認用に用いるsimple.shを用いる  
動作は、date; sleep 10; date の実行のみ  
(これではジョブにならないかな?)

### パラメータ設定

起動する実行処理ホストの数

1,16,32,64,128 : 5種類

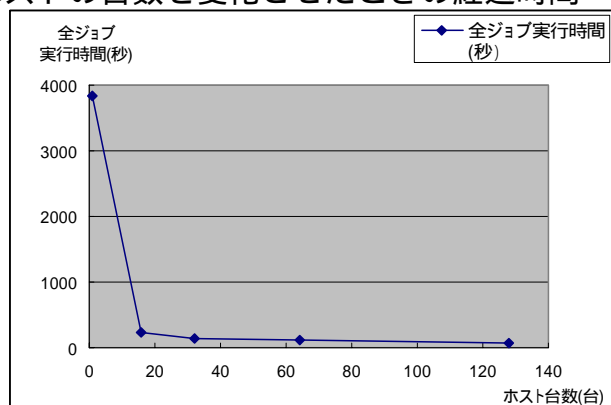
投入するジョブの数

128,256,512,1024 : 4種類



## 動作確認の結果 : 1

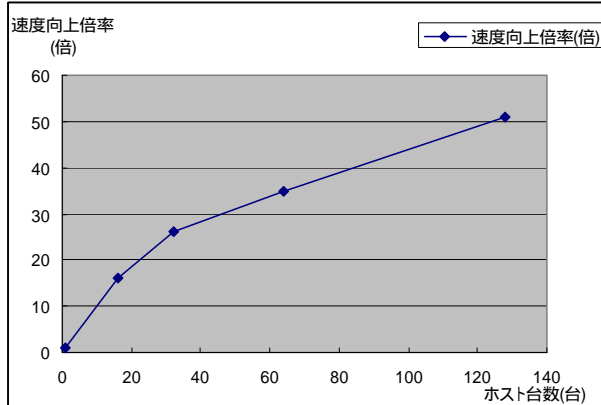
128個のジョブを処理するとき  
実行処理ホストの台数を変化させたときの経過時間





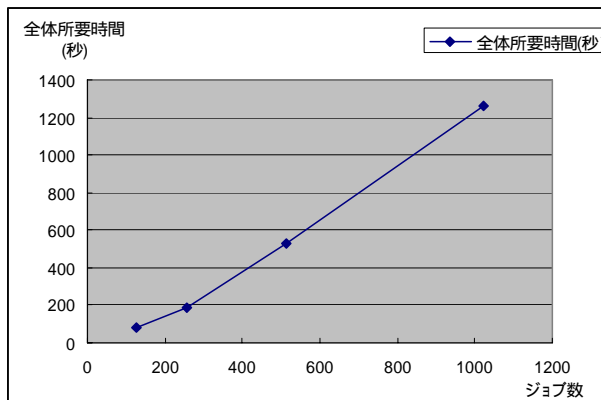
### 動作確認の結果：2

128個のジョブを処理するとき  
実行処理ホストの台数を変化させたときの速度向上



### 動作確認の結果：3

128台の実行処理ホストで  
処理するジョブ数を変化させたときの所要時間





## 動作確認の結果：4

ジョブの処理時間に当たるパラメータとして Simple.shのsleepの時間を変えて確認を行った

その結果、  
S1GEは15秒間隔でジョブを投入するので  
それより短い時間でジョブが終了しても  
次のジョブが投入されていない。

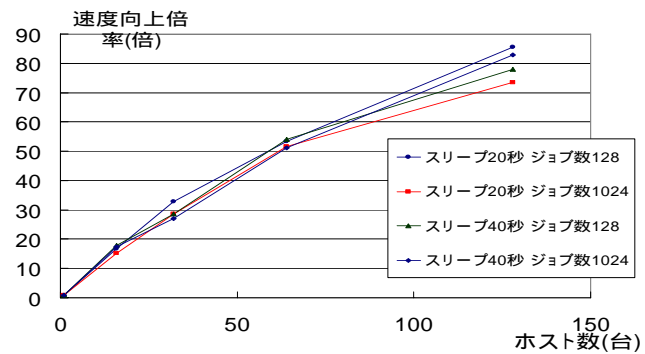
そこで、sleepの時間をパラメータとして  
20秒、40秒  
を検討した。



## 動作確認の結果：5

パラメータ設定： 実行するジョブ数：128, 1024 個  
sleepの時間：20, 40 秒

このときの実行処理ホスト数との関係





## 動作確認のまとめ

最後の結果から  
今回対象としたパラメータの範囲であれば  
緩やかな速度向上率低減があるが  
実行処理ホストの台数に比例して  
安定して分散処理速度が向上している

概数として、  
128台の実行処理ホストで分散処理すると  
全体で80倍程度の速度で処理が行われる

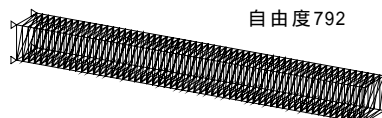
---



## 計算能力の検証：1

具体的な数値計算において  
S1GEonKNOPPIXの計算能力を検証する

解析例 3次元トラス構造物の線形弾性応力解析  
荷重パターンが多数あるものとして  
入力データが異なるジョブを多数繰り返す  
全数解析による最適化評価問題が目標





## 計算能力の検証：2

### 数値解析の規模の検証

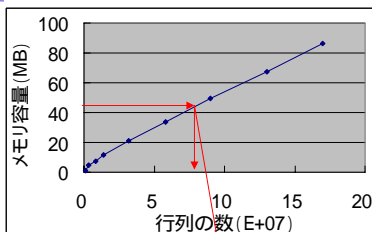
実行処理ホストは128MBのメモリを実装し  
KNOPPIXの動作に80MB程度を必要とする  
よって、数値解析に利用できる実メモリは  
48MB程度になる

そこで、

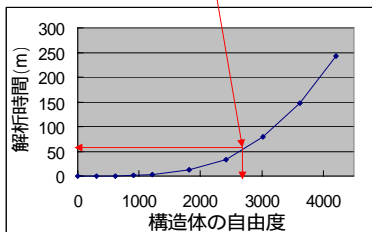
以下のパラメータ間の関係を調査した  
数値解析が使用する配列個数  
メモリ使用量  
構造体の自由度  
解析時間



## 計算能力の検証：3



実メモリ限界の計算  
規模は  
自由度2700  
解析時間60分程度  
となる





## 計算能力の検証：4

グリッドで多数のジョブを検証するために  
解析規模を小さくした

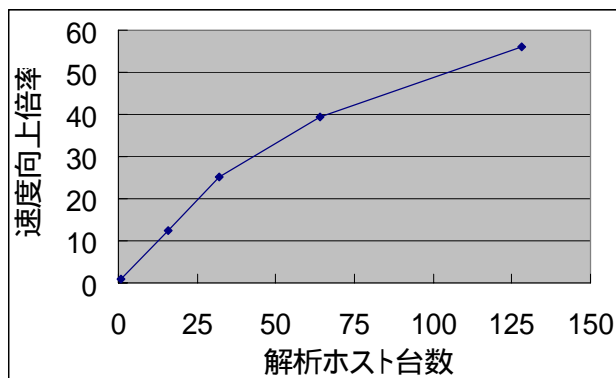
自由度数：792  
実行処理時間：ほぼ1分  
入力データファイル：30KB程度  
出力データファイル：40KB程度  
実行形式ファイル：66KB

以上のジョブを512個処理するための時間を  
実行処理ホストの数を変えて検証した

---



## 計算能力の検証：5



ホスト数が32台までは線形で向上するが  
その後、緩やかになってゆく

---



## 計算能力検証のまとめ

最後の結果から  
今回対象としたパラメータの範囲であれば  
緩やかな速度向上率低減があるが  
実行処理ホストの台数に比例して  
安定して分散処理速度が向上している

概数として、  
128台の実行処理ホストで分散処理すると  
全体で5.5倍程度の速度で処理が行われる

---



## 今後の展開とまとめ

**大規模化** 現在は、演習室3つの128台構成であるが  
LAN内であれば増設が可能であり  
全ての演習室を合わせ400台規模を目指す

**起動方法** 大規模化に備え  
CD以外で起動する方法を検討する  
Windows上で容易にLinuxを起動する方法  
coLinux, UML for Win32などを調査中

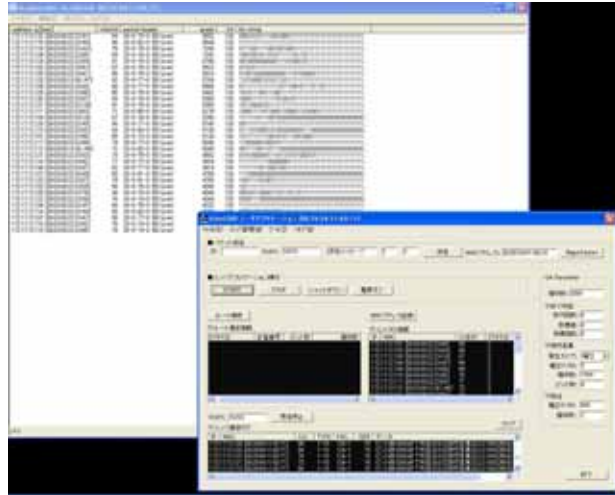
同様な環境の研究者にとっては  
グリッドを活用する1つの方法  
となりうることを確認できた

---

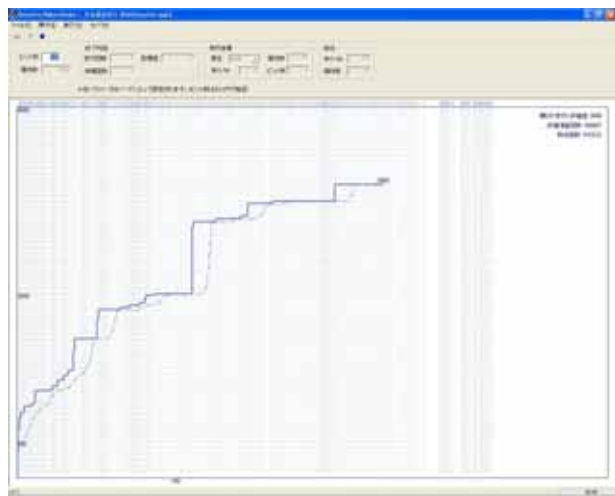




## 岐阜グリッドの概要：2



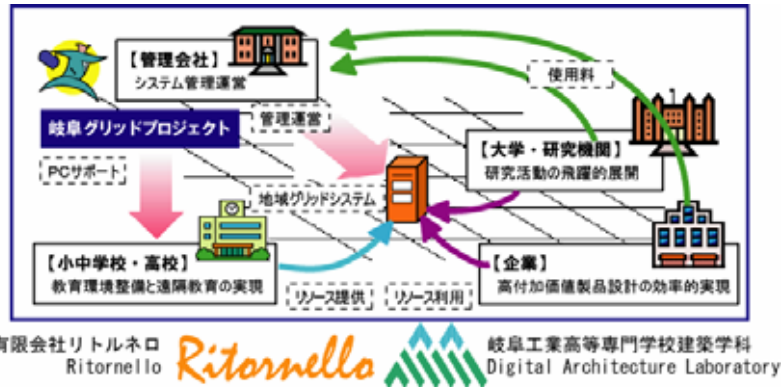
## 岐阜グリッドの概要：3



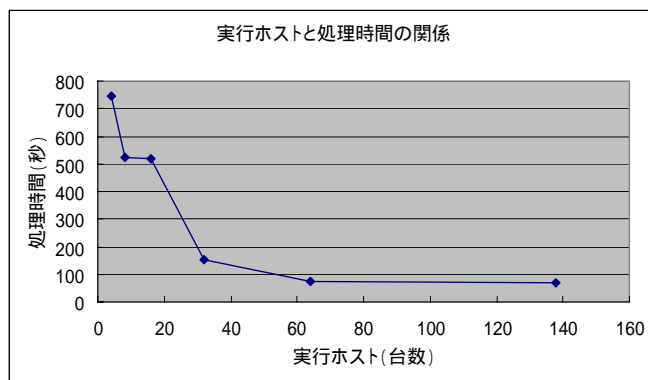


## 岐阜グリッドの概要：4

将来的には、地域グリッド(仮称)のような大きな組織的なグリッドにしてゆきたい！



## 岐阜グリッドの動作確認



並列遺伝的アルゴリズムを実行し  
目的の解に到達するまでの処理時間



## 最後にお願

共同研究のお誘い  
研究を推進するためには  
情報と資本が必要です

(独法化を控えた営業活動で  
す)

このような研究に興味をお持ちの方で  
研究資金を提供して下さる方  
応用課題を共同ですすめてくれる方

柴田 [ryos@gifu-nct.ac.jp](mailto:ryos@gifu-nct.ac.jp) までご連絡ください  
お待ちしております

---