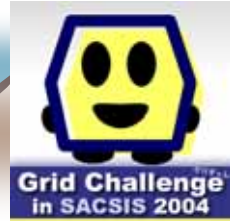


SAC SIS2005併設企画 Grid Challenge: 狙いと実現

Grid Challenge実行委員会
田浦健次郎(東大)



◆ SAC SIS 2005

- 毎年行われている, 基盤システム(Grid, 並列, 分散, ミドルウェア, アーキテクチャ, DB, WWW, etc.)に関する国内最大の会議
- 2005は5月@筑波

◆ **Grid Challenge** in SAC SIS 2005

- SAC SIS 2005の併設企画
- 活動自体は2月-3月

2005年1月7日

グリッド協議会シンポジウム



企画概要

- ◆ 1,000 CPU程度の分散環境を参加者に提供する
 - **規定課題部門**
 - 与えられた問題を{早く・速く}解く
 - **自由課題部門**
 - 製品, 研究成果物の実験場を提供
 - 条件: SACSISにてポスター発表をする

2005年1月7日

グリッド協議会シンポジウム



実際の人・手間

- ◆ 多数(8 研究{所・室})のクラスタを結合
 - 各拠点で管理担当者をたててもらおう
 - アカウント作成
 - コンテスト期間中の利用の便宜, ひどい故障には対処
 - 共有利用期間 1ヶ月
 - 排他利用期間 ≤ 日中5日間 × 2
- ◆ 問題の検討, 開発用パッケージ作成

2005年1月7日

グリッド協議会シンポジウム



拠点と人

- ◆ 産業技術総合研究所 グリッド研究センター
 - 田中良夫さん, 中田さん, 建部さん, 横川さん
- ◆ 筑波大学 HPCS研究室 (朴・佐藤研究室)
- ◆ 電通大 弓場研究室・本多研究室
 - 吉瀬さん
- ◆ 東工大 松岡研究室
 - 田中康司さん
- ◆ 東工大 合田研究室
- ◆ 徳島大 小野研究室
- ◆ 東大 平木研究室
- ◆ 東大(柏) 近山研究室
 - 横山さん
- ◆ 東大(本郷) 田浦研究室(+COE情報科学技術戦略コア)
 - 遠藤さん

2005年1月7日

グリッド協議会シンポジウム



とりあえず今回の企画としての狙い

- ◆ 能力のある人材(学生)の発掘, 売り出し場の提供, この分野への引き付け
 - JSPP PSCと同様
- ◆ 研究者の実験場
 - スケーラビリティテスト
 - 分散環境でのテスト
- ◆ 使えるソフトウェア・研究成果の売り出し場

2005年1月7日

グリッド協議会シンポジウム



中期的な狙い この企画が継続的に行われれば...

- ◆ 過去の並列コンテスト(PSC)との違い
 - 基盤ソフト, ツール, 利用形態, etc.が, 現在も, これからも, 研究対象として「発展」中
 - 全員が, 用意されたMPIで性能を競うだけではない
 - 本企画も, その「発展」を助ける結果になると期待

Grid, 並列, 分散を研究するコミュニティにも,
得るものがあるはず

2005年1月



期待(1)

- ◆ この分野の研究者がいつも念頭におく(あてにすることができる), ソフトウェア研究・開発のターゲットを提供する
 - 1,000CPU規模が, 万人の身近な目標に
 - この数字自体, 年とともに増えていくべき
 - その増え方もコミュニティの「感覚」として共有(台数はMooreの法則のように増えない)
 - 実現できていない, でたらめな台数を書くことを戒めることにも
 - 研究者にpaperwareから脱却することをencourageする
 - 実用的ソフト(論文の寿命を超えて使われるソフト)を作る努力に正しく報いる場を提供
 - 時と場所(設定)を選ばずに動作するソフトウェアの開発をencourageする

2005年1月7日

グリッド協議会シンポジウム



期待(2)

- ◆ Gridの大きな価値のひとつのはずである「資源の柔軟な融通・プーリング」の実経験をつみ、加速する
 - 障壁が何かを理解・共有する
 - 精神的なもの(持てる者の「使われ損」感覚)?
 - 課金(利用量のトラッキング)機構の欠落?
 - 優先度(自分の保有資源は取り戻せる)の欠落?
 - 安定性・信頼性の低下(激しく使われると故障率大幅アップ)?
 - 開発しているソフトにそこまでのスケーラビリティはない。ゆえに台数を増やしてもあまり利益がない。“得をするのはアプリ屋だけ”感覚

2005年1月7日

グリッド協議会シンポジウム



今回の企画内容の紹介

- ◆ 規定課題問題
- ◆ 各拠点の資源

2005年1月7日

グリッド協議会シンポジウム



規定問題設定の苦労話

- ◆ 条件 : 3つの ×
 - × 計算/通信比の小さすぎる計算
 - × 通信量0の計算
 - × 敷居が高い(解くだけで大変な)問題
 - もともと, クラスタ・ドメインをまたがったプログラミングは敷居が高い

敷居が低く(解くだけなら簡単で), 少ない通信量で
1,000CPU規模までスケールし, かつ並列化(通信)を工夫
のしがいがある問題

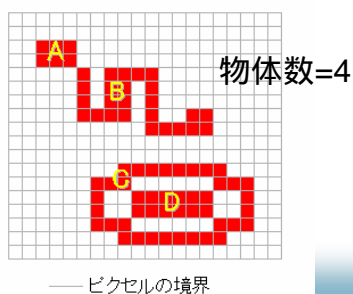
2005年1月7日

グリッド協議会シンポジウム



規定問題

- ◆ 多数のマシに分散する画像ファイルに含まれる「物体」の総数を求めよ。



2005年1月7日

グリッド協議会シンポジウム



予選と本選

- ◆ 予選(1/31-3/2): **早く(early)**結果を出すと勝ち
 - これ自体, 賞の対象
 - 世の中の基準(MPP時代にはなかった基準)
 - 実行速度よりも, 開発時間の短縮が大事
 - 一度結果を出してしまえばおしまい
- ◆ 本選(3/7-3/18): **速く(fast)**結果を出すと勝ち

2005年1月7日

グリッド協議会シンポジウム



お試しパック

- ◆ 予選前からプログラムが開発できるパッケージ
 - Commodity環境ゆえのやりかた
 - Linux数台あればプログラムが書ける
 - 場所を選ばず動くソフトをencourage
 - その分, 予選期間を短めにできる
 - 予選期間の負荷も軽めにできる(トラブルも?)

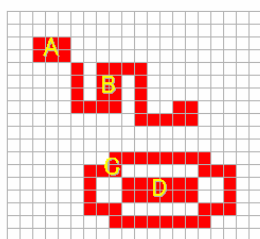
2005年1月7日

グリッド協議会シンポジウム

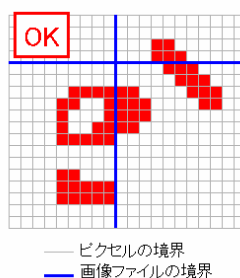
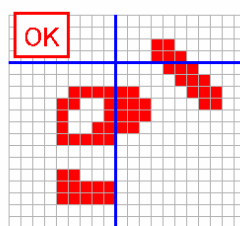


初級と上級

- ◆ 初級: 「またぎ」なし
- ◆ 上級: 「またぎ」あり



— ピクセルの境界



— ピクセルの境界
— 画像ファイルの境界

2005年1月7日

グリッド協議会シンポジウム



初級と上級

- ◆ 初級: 必要な通信は物体数の総計 (reduction)のみ
 - 通信量少ない
 - マスターワークスタイルで容易に記述可能
- ◆ 上級: 縁データの通信が必要
 - マスターにすべて集約してもOK
 - 近隣同士で通信すればbetter?

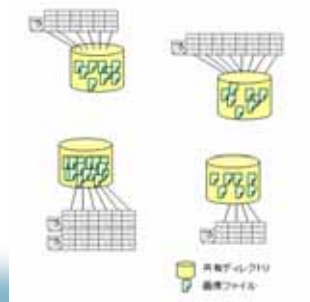
2005年1月7日

グリッド協議会シンポジウム



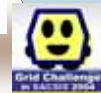
データ配置

- ◆ クラスタ内共有, クラスタ間非共有
- ◆ 配置を偏らせれば, 「負荷移動」が問題となる



2005年1月7日

グリッド協議会シンポジウム

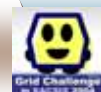


マスターとレプリカ

- ◆ 各ファイルにはいくつかのコピーが存在
- ◆ レプリカをうまく用いることで, (ファイルの中身を動かさずに) 負荷の移動ができる
- ◆ 実際にファイルの中身を動かす負荷の移動に意味があるかどうかは, 微妙(やってみてのお楽しみ)

2005年1月7日

グリッド協議会シンポジウム



ホスト

昨日の時点で
ログイン設定
完了&確認

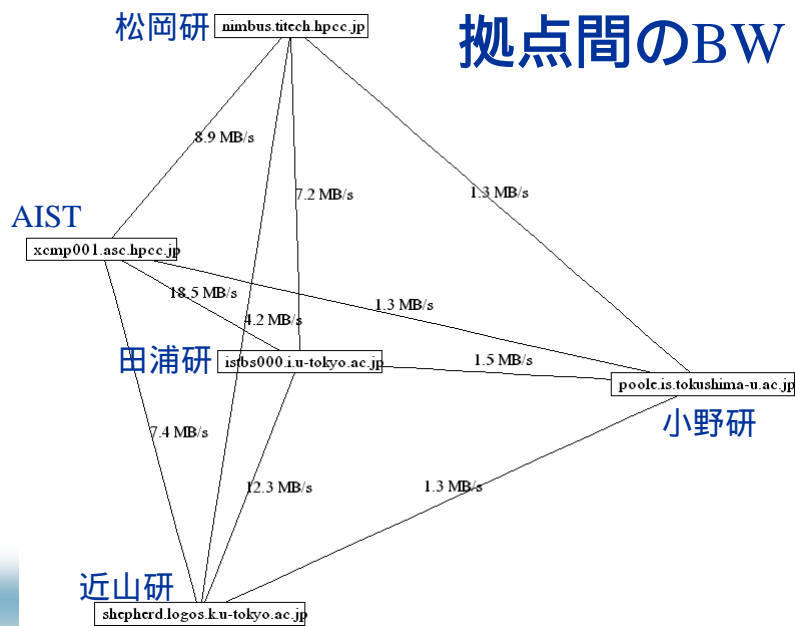
拠点	ホスト	ホスト数	故障中	CPU数
AIST	xcmp001-xcmp042	40	2	80
松岡研	pad000-pad114	100	15	200
小野研	mp20_001-mp20_052	50	2	100
近山研	sheep01-sheep62	60	2	120
COE& 田浦研	istbs000-istbs108	107	2	214
合計		357	23	714
HPCS研		20?		40?
平木研		40?		80?
合田研		30?		60?
				894?

2005年1月7日

グリッド協議会シンポジウム



拠点間のBW



2005年1月7日

グリッド協議会シンポジウム



プログラミング環境

- ◆ 委員会でインストールするもの
 - Globus + Ninf
 - GXP
 - 役に立ってインストールが楽なものは歓迎
 - 各サイトにユーザレベルでインストールできるものは委員であれば誰でも提供できる
- ◆ 各参加者も、何を使っても基本的には自由(制約しない)

2005年1月7日

グリッド協議会シンポジウム



2005年1月7日

グリッド協議会シンポジウム



参考

- ◆ 16K x 16K pixels x 1600 files
 - 圧縮後50KB – 1MB程度 . 合計750MB
 - 試作プログラム処理時間: 650秒程度 .
 - ⇒ 1 CPUで1MB/sec程度

2005年1月7日

グリッド協議会シンポジウム

