

研究テーマ: 高性能計算のためのグリッド環境実現に向けての フィジビリティスタディ(1/2)

(プロジェクト番号 JGN2-A18025)

研究機関: 東京大学情報基盤センター、筑波大学計算科学研究センター

研究の概要:

科学技術計算の大規模化への要求に応えるため、複数の並列計算機を接続して大規模な並列計算機として使用する「計算グリッド」の必要性が高まってきている。並列計算機間の接続に使用するネットワークは、広域ネットワークであることが想定されるため、低遅延ネットワークでの接続を前提に設計されている既存ソフトウェアは遅延を考慮したソフトウェアに作り替える必要がある。本プロジェクトは、計算グリッド向けに開発が進められているシステムソフトウェアを実際の高遅延環境で動作させ性能を測定することで、計算グリッドの実現可能性を調査するものである。

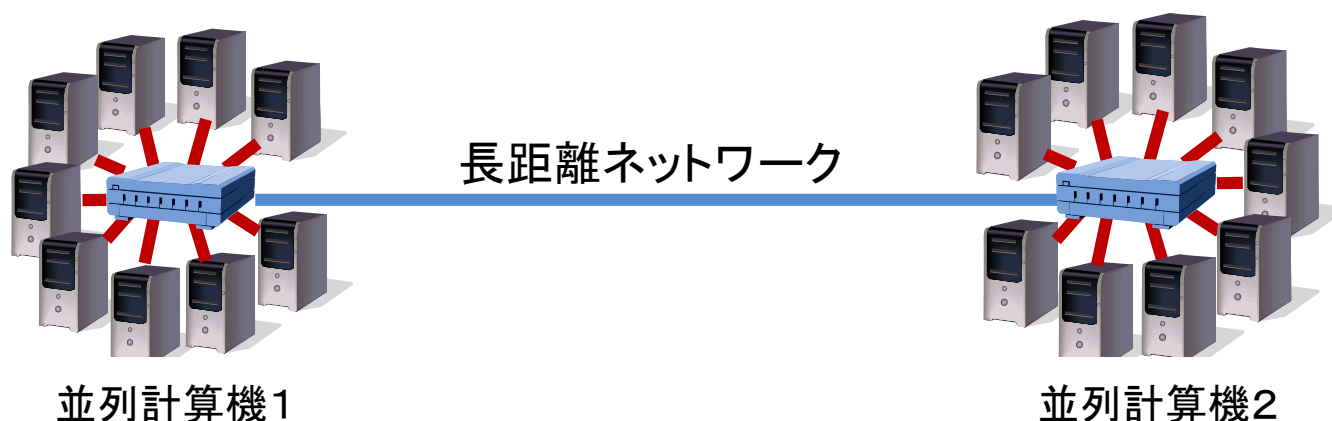
性能を測定したシステムソフトウェアは、並列計算用通信ライブラリであるMPIをグリッド環境に適応させた GridMPI、およびグリッド向け分散ファイルシステムである Gfarm である。実験では、東京大学と筑波大学にそれぞれ設置されているPCクラスタをJGN2で接続した環境にこれらのシステムソフトウェアをインストールし、その性能を測定した。性能測定の結果、GridMPIの性能測定では高遅延環境でもバンド幅が確保されていれば十分な性能が得られるアプリケーションが存在すること、Gfarmの性能測定ではディスク性能を上回るネットワークバンド幅も有効に活用できる可能性があることがそれぞれ示された。

研究の目的:

計算グリッドは下図のように複数の並列計算機をネットワークで接続したものである。

密に結合された古くからある形式の並列計算機と比較して、以下の特徴を持つ

- 並列計算機間の接続は物理的距離によってもたらされる高い遅延を持つ
- 並列計算機間の接続は並列計算機内で使用されるネットワークのスイッチ容量には及ばないため、バンド幅ボトルネックとなる



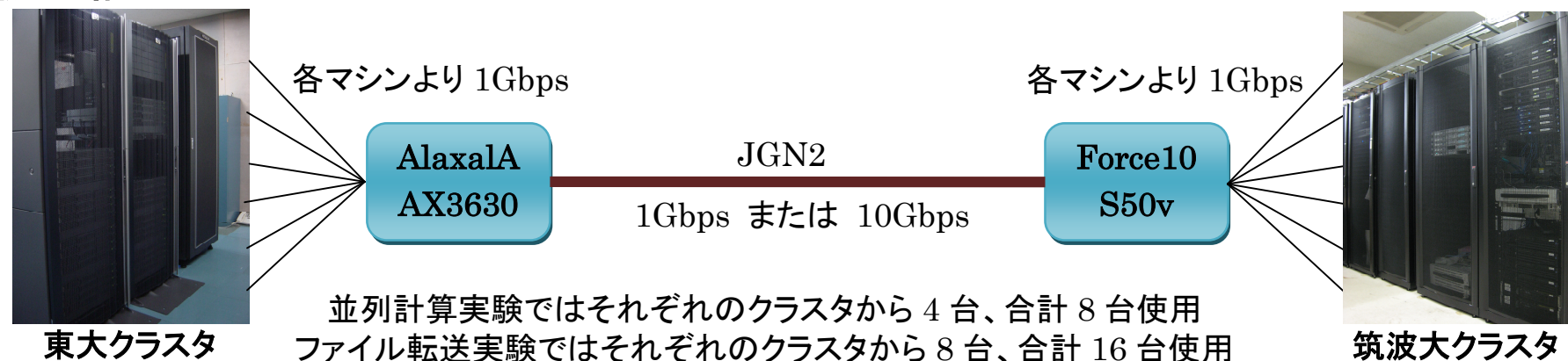
本プロジェクトではJGN2およびグリッド用システムソフトウェアを用いて計算グリッドを実際に構築する。その上でアプリケーション性能を測定することで、遅延とボトルネックリンクが性能に及ぼす影響を調べる。また、ファイル転送試験を行い、高遅延環境においても並列同時アクセスによるスループットの向上が得られることを確認する。

研究テーマ: 高性能計算のためのグリッド環境実現に向けての フィジビリティスタディ(2/2)

(プロジェクト番号 JGN2-A18025)

研究機関: 東京大学情報基盤センター、筑波大学計算科学研究センター

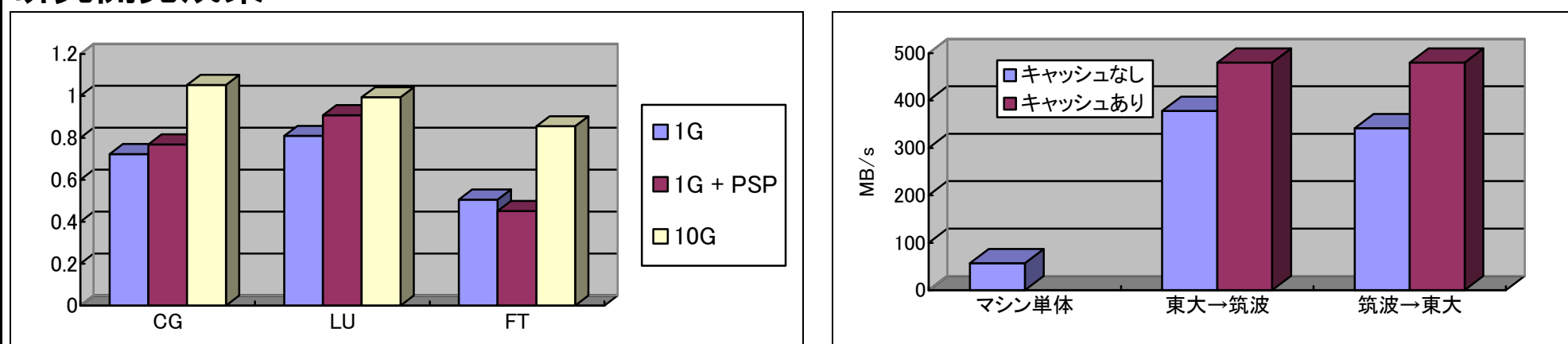
実験機器構成:



ソフトウェア

- GridMPI および PSPacer: 東京大学と産業技術総合研究所で開発されているグリッド向けMPIライブラリおよびソフトウェアによるページング機構
- Gfarm: 筑波大学および産業技術総合研究所で開発されているグリッド向け分散ファイルシステム

研究開発成果:



左のグラフはGridMPI上でNAS Parallel BenchmarksのCG(共役勾配法)、LU(LU分解)、FT(高速フーリエ変換)を実行した結果である。値はグリッド環境でないローカルクラスタでの結果を1とした相対性能である。3本のグラフはそれぞれ、JGN2で使用したバンド幅とソフトウェアページング(PSP)の有無を示す。3種類のアプリケーション共、10Gの場合はローカルに近い性能が出ており、バンド幅の重要性を示す結果となっている。また、アプリケーションによってはページングが有効であることが示されている。

右のグラフはGfarm上のファイル転送性能であり、クラスタを東大、筑波大でそれぞれ8台使用した並行ファイル読み込みとマシン単体の性能が示されている。複数の計算機にあるディスクを同時に使用することで、ディスクより高速なネットワーク帯域が有効に活用できることが示されている。

プロジェクトのアピールポイント:

アプリケーションを選べば計算グリッドで実用的な性能が得られる可能性のあるデータが得られた。また、ファイルシステムに関してはノード数を増やせば10Gのバンド幅を十分に生かせる可能性が示された。

プロジェクトの自己評価:

グリッドとしては非常に規模の小さい並列計算機での実験にとどまっているため、今後はより大規模な環境で計算グリッド実現可能性を検証していく必要がある。