B-051

並列計算機の通信ネットワークトポロジの 3 次元表示手法
Three Dimensional Display Method for Communication Network Topology of Parallel Computer

鈴木 遼平† 石畑 宏明†
Ryohei Suzuki Hiroaki Ishihata

1. はじめに

並列計算機システムの実行性能を下げる大きな原因として，通信の競合から発生する通信待ち時間が挙げられる．通信の競合を回避し，効率の良い通信を行うためには通信アルゴリズムを工夫する必要がある．

通信アルゴリズムの開発では，通信状況を把握して通信の競合が発生する原因を特定することが重要になる．そのため通信経路やリンクの混雑状況といった情報を分かりやすく提供するツールが求められる．

我々の研究室では通信アルゴリズム設定環境の一環としてネットワークトポロジを 3 次元上に表示して，通信バンド幅や輻輳状況をリンクに反映させる可視化ツールを作成している [1]．このようなツールではネットワークトポロジをどのように表示するかで，通信状況の把握し易さが異なる．そのためリンクに反映された情報が把握し易いネットワークトポロジの表示方法が重要となる．

従来行われていた，ネットワークトポロジを 2 次元上で表示するツールでは，ネットワークが大規模になると，多数の通信路（リンク）が交差する場所ができる．そのためリンクの始点と終点が判断しにくく，通信状況の把握が難しくなる．FatTree トポロジはリンクの交差が多く，1000 ノード程度で全体像から，リンクがどのノードを接続したものか分からなくなる．

本研究では，可視化ツールに適用する FatTree トポロジの 3 次元表示手法を提案する．

2. トポロジ可視化の機能要件

ネットワークトポロジを可視化するにあたり，以下のことを機能要件とする．

- 表示するリンクどうしの交差をなくす
  リンクどうしが交差せずに描画できるようスイッチを 3 次元空間中に配置する．リンクには，描画の際にそこを通過するメッセージ数やデータ量などが判断できるように，形や色を変える．リンクどうしが交差するとこれらの情報の判読が困難になる．特に FatTree トポロジでは，Level が上のスイッチほど遠くのスイッチとリンクで接続するため，交差が発生しやすい．
- 規則的にスイッチを配置する
  トポロジの全体から，スイッチの位置関係を直感的に判断できるようにスイッチを配置する．通信の競合の発生箇所を直感的に理解できるようにするためには，スイッチを規則的に配置する必要がある．
- ネットワーク全体をコンパクトに表示する
  大規模なネットワークトポロジの表示では，全体が著しく縦長や横長になりがちである．そのためスイッチの配置を工夫する

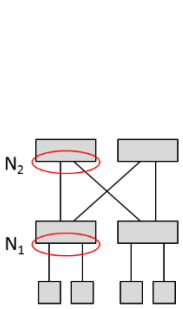
† 東京工科大学 Tokyo University of Technology

3. FatTree の表示手法

3.1 FatTree トポロジの表示手法

FatTree は，複数のクロスバスイッチを多段に組みあわせて構成する．$2N_1$ ポートのクロスバスイッチに N_1 個のノードを接続し，残りのポートを N_1 個の N_2 ポートスイッチで個別に接続すると 2 段の FatTree 構成となる（図 1a）．

ここでスイッチ同士を接続している N_2 ポートのスイッチを $2N_2$ ポートのスイッチに置き換え，空いている N_2 ポートを個別に $N_1 \times N_2$ 個のスイッチで接続すると 3 段の FatTree 構成ができる(図 1b).



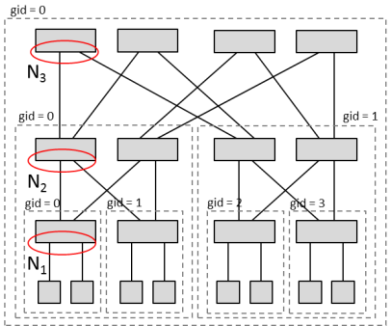


図 1a 2×2 の FatTree　　図 1b 2×2×2 の FatTree

同様に L 段の構成を構築することができる．このとき，この FatTree のトポロジを $N_1 \times N_2 \times \cdots \times N_L$ の構成の FatTree と呼び，N の添字 1 ～L を Level と呼ぶ．

ノードは Level 0 のスイッチとして扱うものとする．またその Level までで通信が可能な範囲をグループと呼ぶ．FatTree はより小さなグループを再帰的に上位のスイッチで接続した構成である．図 1b の点線の枠はグループを囲んだものである．

Level i のスイッチ数 swn(i)は式(1)および(2)で求められる．なお dport(i)は Level i のスイッチの下向きポート数である．

$$swn(0) = \sum_{i=0}^{n} dport(i) \quad \cdots \quad (1)$$

$$swn(\mathrm{i}) = swn(0)/dport(i) \cdots \quad (2)$$

Level i の 1 グループあたりのスイッチ数 nsw(i)は式(3)で求められる．Level 0 と 1 はグループを組まないため nsw(0) = nsw(1) = 1 である．

$$nsw(\mathrm{i}) = swn(\mathrm{i}\text{-}1) \times dports(\mathrm{i}\text{-}1) \quad \cdots \quad (3)$$

ここから各レベルの j 番目のスイッチが，属するグループ番号 gid(j)を式 4 のように求めることができる．これにより Level i の j 番目のスイッチが接続する Level i+1 スイッチのポート番号 p(j)が式(5)で求められる．

$$gid(\mathrm{j}) = \mathrm{j}/nsw(\mathrm{i}) \quad \cdots \quad (4)$$

$$p(\mathrm{j}) = gid(\mathrm{j})/dports(\mathrm{i}+1) \quad \cdots \quad (5)$$

3.2 FatTree の 3 次元表示における課題と解決策

各機能要件を満たすための課題と方針を以下に示す．

- リンクどうしの交差をなくす

  直接リンクで接続したスイッチどうしは，親子関係があり，Level の高い方が親となる．FatTree では子が複数の親を持っており，そこでリンクが交差しやすい．そのため親と子が離れすぎていたり，親どうしが近づきすぎたりすると交差が起こる．

  本手法では，親の下に子を配置するようした．また Level 3 以降は，同じ子を持つ親スイッチを z 方向にずらして配置した．このときずらす距離は，Level i が奇数のときスイッチが x 方向に長いため，スイッチサイズと同じにし，偶数のときスイッチが z 方向に長いためスイッチサイズの 1/4 にした．

- 規則に従った配置

  配置の規則は，x 方向にグループ内のスイッチを配置し，z 方向に次のグループを配置する．また y 軸は Level の高さを表しており，Level が高いものほど高い位置に配置する．図 2 に 2×2×2 構成の FatTree の Level ごとに，上 z 軸方向から見たときのスイッチの配置を示す．点線で囲んだ領域はその Level のグループを表している．

  Level i は x 方向に 1 グループあたりのスイッチ数 nsw(i)ずつ配置する．次のグループはそのグループの z 方向に配置していく．前に述べたように Level 3 以降は，同じ子を持つ親スイッチを z 方向にずらして配置した．

  なお Level 0 は nsw(0)が 1 だが，1 列にすると z 方向に長くなってしまう．そのため，x 方向に Level 1 の下向きポート数(ndport(1))ずつ z 方向に配置していく．

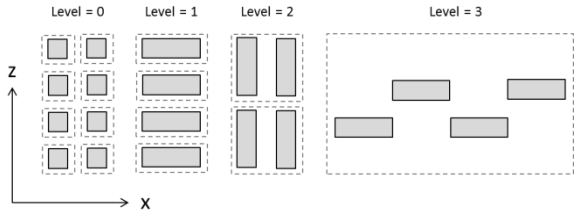


図 2　2×2×2 構成の FatTree の各 Level の配置

- 全体をコンパクトに表示

  各 Level でスイッチの向きを変えることにより，Level を区別し易くした．またグループ内のスイッチを x 方向に，グループを z 方向に配置する規則により，x と z 両方に大きくなる構造にした．そのため横長や縦長にならず，画面上により多くのスイッチを表示できる．これにより大規模なネットワークを把握することができる．

3.3 実装

提案した表示方法を，OpenGL と GLUT を使用して実装した．スイッチは glutSolidCube()を利用し，リンクは gluClinder()で描画した．リンクは送受信を区別するため，上りと下りの 2 つを表示している．リンクにはテクスチャマップを使い通信の方向が分かるように矢印を表示した．図 3 に実装した 3×3×3 構成の FatTree を示す．各スイッチのリンクに交差がないことが分かる．

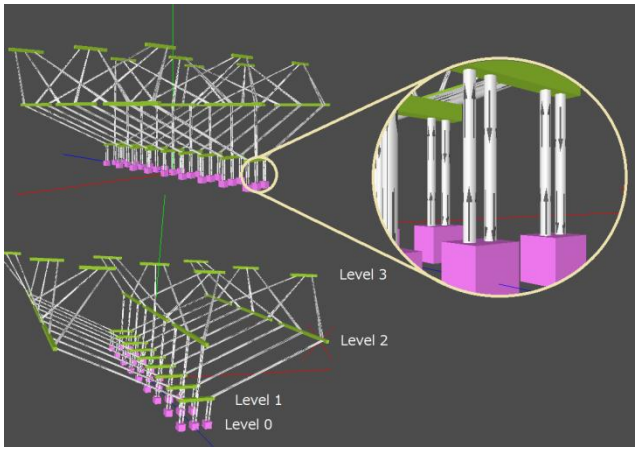


図 3　提案した表示手法の実装（3×3×3 の FatTree）

4. 応用

4.1 ネットワークの状態の表示

作成したネットワークトポロジ表示方法を使用して，リンクの通信状況を色を変えて表示させた例を図 4 に示す．リンクの交差がないため，リンクに反映された情報が見やすい．またグループを z 方向に，グループ内のスイッチを x 方向に配置するように分けるなど，規則的に配置されているため，全体の中で競合発生箇所を直感的に理解できる．

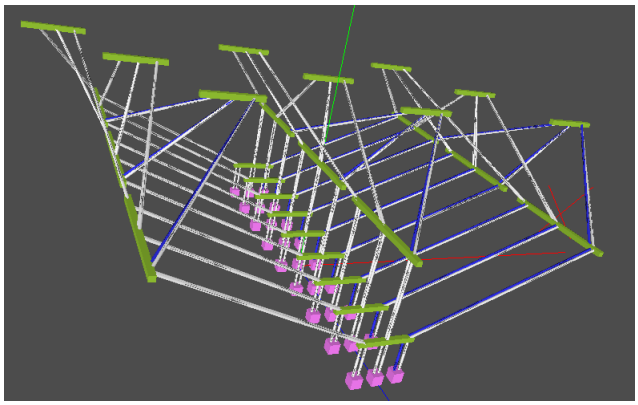


図 4　トポロジのリンクに通信状況を表示させた例

5. まとめ

本研究では大規模並列計算機で用いられる通信ネットワークトポロジの 1 つである FatTree を 3 次元表示する手法を提案した．リンクの交差をなくすことでリンクに反映された情報の判別を容易にし，規則に従い配置することで直感的にネットワーク全体の状態を把握できるようになった．

ネットワークシミュレータの通信ログを表示するインターフェースの作成が課題である．

謝辞

本研究の一部は科研費(22500052)の助成を受けたものである．

参考文献

[1] 石井省吾, “大規模並列コンピュータを対象とした通信アルゴリズムの可視化ツールの設計”, 情報処理学会第 7 3 回全国大会予稿, 4J-2 (2011).