ネットワークについて

木綿 芳行

1995年4月27日

1 相互結合網の分類

1.1 分類項目

以下に示す（1）～（4）を組み合わせると多様な相互結合網が存在することになる

1. 動作モード
   (a) 同期方式
   入出力ターミナル間でのデータ転送要求が同一時刻に一斉に生起し、網内の通信制御が一斉になされる.
   (b) 非同期方式
   入出力ターミナル間の個々のデータ転送要求がランダムに生起する.
2. 制御方式
   (a) 集中制御方式
   中央に設置された一つの制御方式によってすべてのスイッチが制御される.
   (b) 分散制御方式
   スイッチ自体に制御装置が内蔵されており, 各スイッチは自律的に動作する.
3. 交換方式
   (a) パケット交換方式
   メッセージに宛名をつけて結合網に送り出す. 入出力ターミナル間に物理的なリンクの確立をしないで, メッセージを送出できる。比較的短いメッセージの通信や通信相手が頻繁にかわる場合に有効である。また、網内にバッファを設けることによって, 接続要求の衝突の回避, 最適ルーティング, 故障対策などが容易に行なえる.

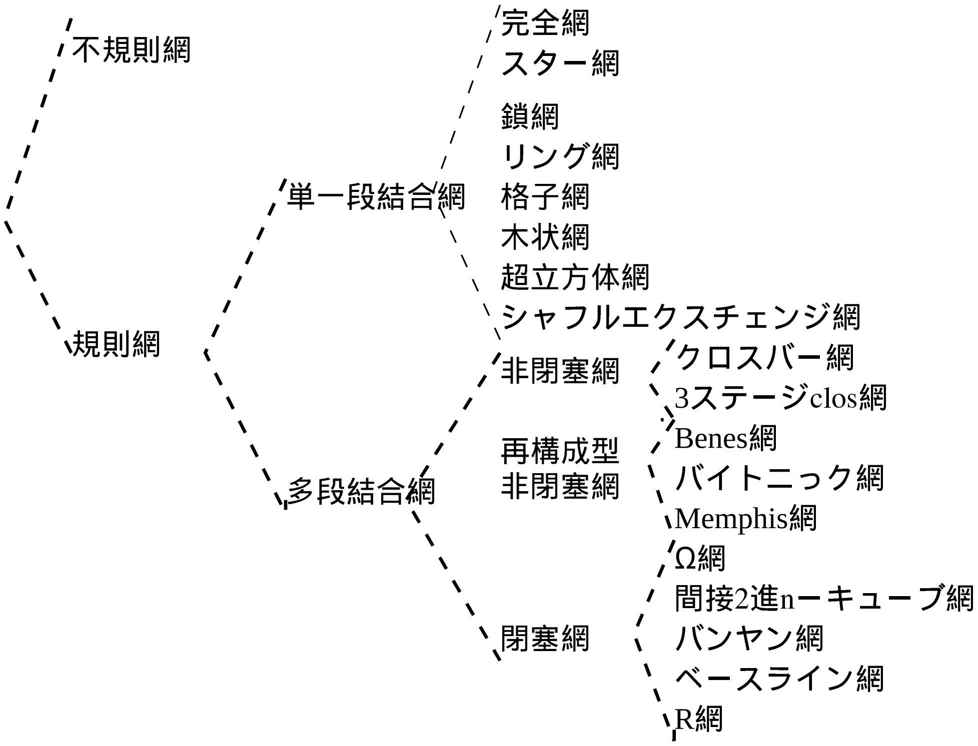

図 1: 相互結合網の分類

(b) 回線交換方式

電話による通信と類似しており, 入出力ターミナル間に物理的リンクを確立するまで時間がかかるが, 一旦リンク確立すると, 解放されるまで物理的なリンクが継続する。したがって, 大量データの高速転送などに適している.

4. トポロジー

相互結合網の幾何学的形状 （トポロジー) は, 図 1に示すように不規則網と規則網に大分類され規則網では, さらに単一段結合網と多段結合網に分類できる.

2 単一段結合網

単一段結合網は、直接結合網, 静的網とも呼ばれ, 木状網のようにあるターミナルから直接通信できるターミナルが限定されている.

2.1 単一段結合網の評価項目

1. 距離 (D、d)

二つのターミナル間で通信を行なうとき, 中継ターミナルが必要である. 二つのターミナル間の最小中継ターミナル数 + 1 をターミナル間距離と定義する. 任意の二つの

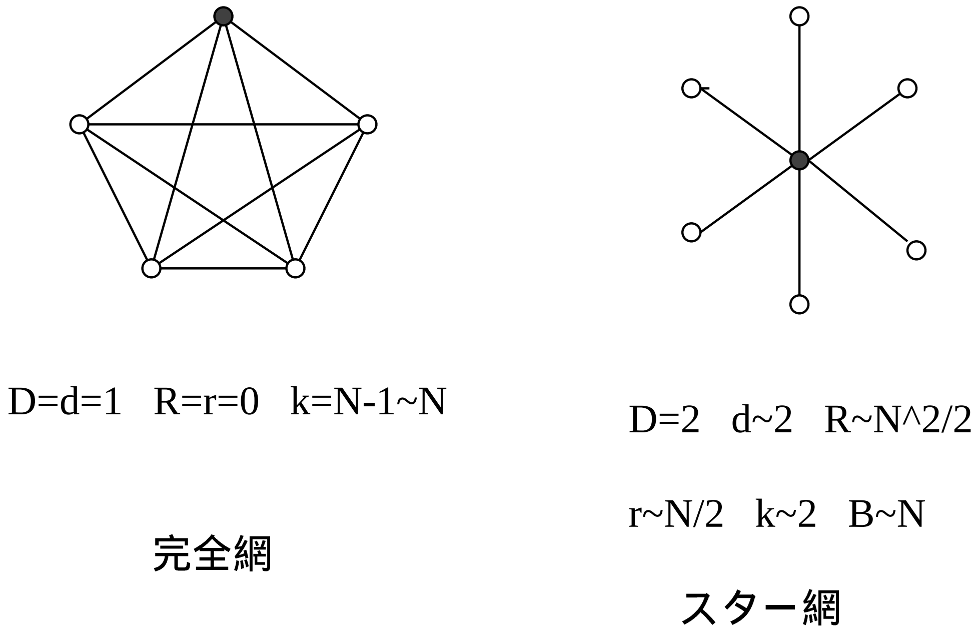

図 2: 種々の単一段結合網

ターミナル間距離のうち, 最大のものを最大距離（D）, 平均値を平均距離（d）とする.

2. 中継量（R、r）

   あるターミナルは他のターミナルの中継ターミナルとして使用される. 中継量は、中継規則に基づいてあるターミナルが他のターミナル間通信の中継として使用される数を示す. 網内での最大中継量, 平均中継量をそれぞれ R、r と表す. R の大きいターミナルが網内に偏在する場合は、そこが通信のボトルネックの可能性がある. 中継規則が, 中継回数が最小となるようにきめられているとすると,N 個のターミナルがある場合は, 入出力ターミナル. ペアは N(N-1)/2 本, 一つのペア当たり (d-1) 個なので

   $r = N(N-1)(d-1)/2N = (N-1)(d-1)/2$

3. 接続数（k）

   ターミナル一つ当たりに直接結合されるターミナル数を接続数（k）と定義する.

4. 放送時間（B）

   網内の管理的な役割を果たすターミナルから同報通信を行なう場合, 全ターミナルにメッセージが伝わるまでの最短時間を放送時間 B と定義する.

2.2 種々の単一段結合網

1. 完全網

図 2に完全網を示す．ハードウェア量の制約により, 大規模システムの実現は困難である.

2. スター網

   中央に設置されたターミナルに対する接続数が (N-1) となっており, ここが通信のボトルネックとなっている.

   スター網の特殊な形態に単一バス方式とループ方式がある.

   単一バス方式では, 単一バスがスター網の中央に設置されているターミナルに相当する. 制御方式として, 集中制御方式と分散方式がある.

   (a) 集中制御方式

      単一バスに一台の制御装置が設置され, 割り込み方式では, 各ターミナルが非同期に要求を制御装置に知らせ, 制御装置がアービタ (要求調停回路) によって唯一のターミナルにバスの使用権を与える.

      ポーリング方式では, 逆に制御装置がサイクリックにターミナルの要求を調べて, 要求のあるターミナルにバスの使用権を与える.

      集中制御方式では, ターミナル番号の識別のため, 多数の制御網を必要とする.

   (b) 分散制御方式

      CSMA/DA 方式では, 各ターミナルはバスの使用状況を調べて, バスが空いていれば, データ転送を開始し, もしバス上のデータと出力中のデータが異なり, 他のターミナルからもデータ転送が開始されていることをしれば, データ転送を中止し, 一定時間後に再送を試みる負荷が高くなると衝突の頻度が増して, 再送の回数が増える.

      トークンパッシング法では一個のトークンがリング上を流れ, トークンを捕らえたターミナルがアクセス権を得る．パケットの衝突が生じないが, フリートークンを捕まえるまでパケットの伝送ができない.

      デージーチェイン方式では上流でバスを利用していることを下流に知らせる方式である.

   ループ方式はループ状のバスに一だいの制御装置があり, 時分割方式を採用している.

   光ファイバーを利用したループ網が主として LAN を中心に利用されている.

3. 鎖網とリング網

   リング網は鎖網の両端を結合して平均距離 d を小さくできる．高速計算機の例では,Cyberplus(CDC 社) などがある.

4. 格子網

   格子網には, 正方格子, 三角格子, 六角格子などがある．最左端と最右端, 最上端と最下端同士のターミナルを結ぶとトーラス網 (図 3) となる．格子網は, 単純であり, 画

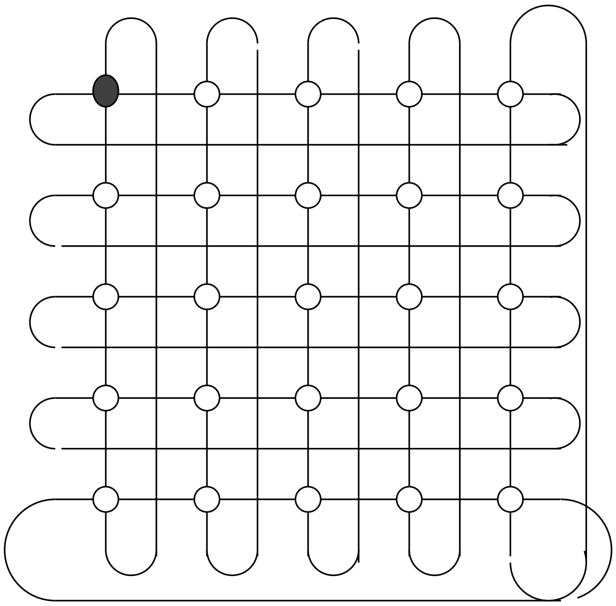

図 3: トーラス網

像処理や偏微分方程式の解法などに向いているので,MMP(グッドイヤ社),PAX(筑波大) など多数のシステムが開発されている.

5. 木状（トリー）網

   葉ノードにのみターミナルを結合し, その他のノードはすべて中継スイッチとする方式や, 木の各レベル内に部分的に鎖状結合をする方式（X トリー）などがある. VLSI チップ内に効率良く埋め込むことができ, 網の拡張も容易である. しかし, 中継ノードが一つでも故障するとシステムダウンにつながる恐れがある.

6. 超立方体網（2 進 n-キューブ, ハイパーキューブ）

   図 4 2 進 3 ーキューブの例を示す. ターミナル a を 2 進数で表し, それとハミング距離が 1 だけ違うターミナルに結合している. 接続数が他の網より大きく, またリング網, 木状網などを含んでおり汎用並列処理向きの相互結合網として注目されている.

7. 環立方体網

   2 進 p ーキューブの各ターミナルを p 個のターミナルからなるリングで置き換えたものである.

   図 5に $p = 3$ の場合を示す. 二次元平面上にうまくマッピングできる特徴を有してる.

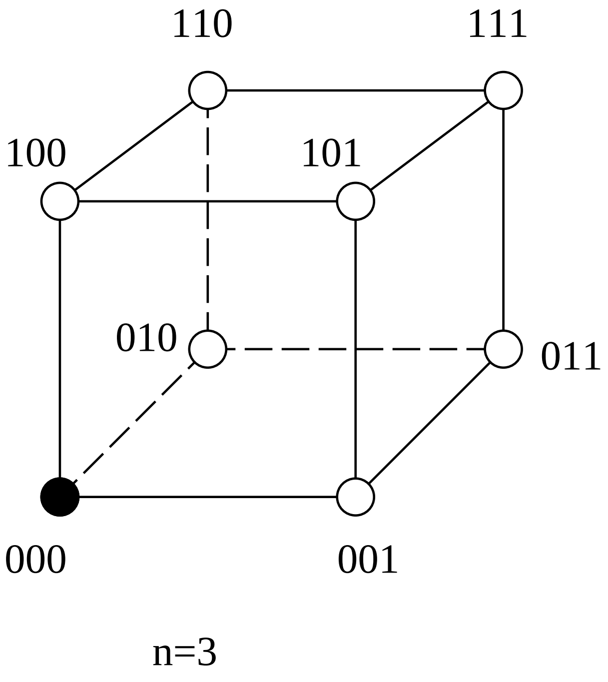

図 4: 2 進 n-キューブ網

3 多段結合網

多段結合網では N 入力ー N 出力ターミナル間にはスイッチ群が設定されており, 各スイッチを制御することによって, 所定のターミナル間の通信が可能となる.

多段結合網には, 非閉塞網, 再構築型非閉塞網, および閉塞網がある.

3.1 非閉塞網

非閉塞網では, 任意の入出力ターミナル間に接続要求があった際に接続中の他のターミナル間通信のためのスイッチ設定を変更しないで要求が満たされる. 接続中の他のものの経路を変更することによって接続要求が満たされる網を再構築型非閉塞網と呼んで区別する.

1. クロスバー網

   N 入力 N 出力のクロスバー網は, 閉塞なしで任意の時刻に任意の入出力ターミナルを高速に結合できる.

   結合パターン数は N!個, 入出力ターミナル間の通信に必要な時間はスイッチ段数で 1 である. スイッチの個数は N^2 となり,N が大きいほど実現が困難である.

   (a) 集中制御方式

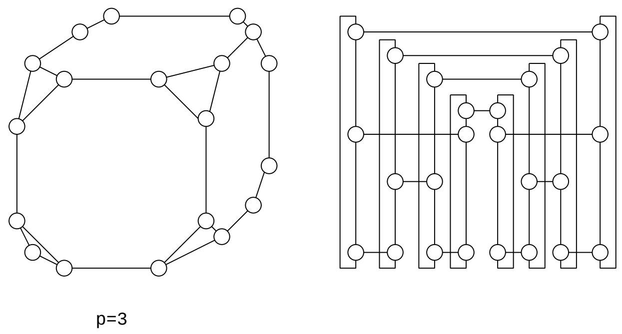

図 5: 超立方体環

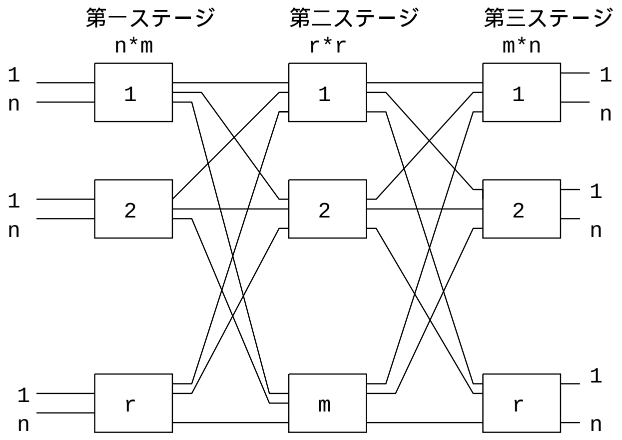

図 6: 3ステージclos網

SIMD型計算機のように使用する結合パターン数が限定される場合には, 各スイッチの制御パターンをメモリに記憶させる. メモリ容量分の結合パターンが使用でき, かつスイッチは単純で高速化できる。しかし, ピンネック問題が生じる. マルチプレクサを用いる方式では, N個必要で, それぞれの制御にlogNビット必要である.

(b) 分散制御方式

行上のスイッチ群に対して, 簡単な制御回路を割り当てる方式がある。入力側ターミナルから出力ターミナル番号などの制御情報が入力されると, 当該列を管理する制御回路が行と列のクロスポイントにあるスイッチの開閉制御を行なう. しかし, 二つ以上の異なる入力ターミナルから同一の出力ターミナルにアクセスする場合には, 列方向のスイッチ間で調停をする必要がある.

2. 3ステージclos網

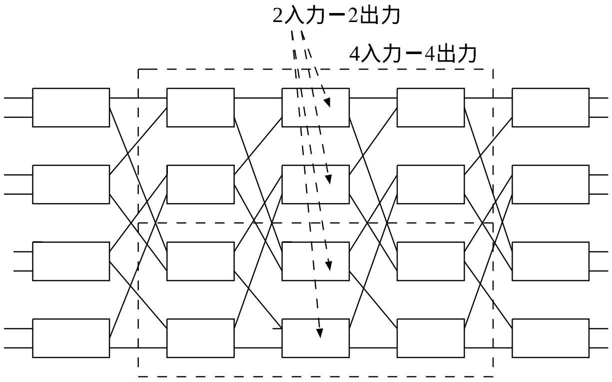

図 7: 8 入力 8 出力 Benes 網

図 6に３ステージ clos 網を示す. 四角のボックスはそれぞれ $n * m$,$r * r$,$m * n$ のクロスバー網を示す. N = 256,512 の時, クロスバースイッチでは,65536 個,262144 個で, ３ステージ clos 網では、22146 個,63487 個ですみ, 大幅なスイッチ数の削減となっている.

3.2 再構築型非閉塞網

1. Benes 網

   N 入力ー N 出力の Benes 網を NBenes 網と略記する.

   Benes 網では, スイッチ設定に時間がかかり任意の結合パターンに対するスイッチ設定を動的に行ない得ない欠点を有する. したがって, 頻繁に使用する結合パターンに対するスイッチ制御データをあらかじめ計算で求め, 記憶しておき, 結合要求にスイッチを集中制御する方式がもちいられる.

2. バイトニックソータを用いた結合網

   バイトニックソータで, 宛先情報に基づいてソーティングすると, 所定の入出力ターミナル間の通信が実現できる. スイッチの設定が各スイッチで分散的に高速なされる特徴がある.

3. Memphis 網

   ３ステージ Clos 網において $m = n = r$ の場合 IBM GF-11 などに利用されている. 制御メモリに記憶された 1024 個の可能な置換群の中から、一つが高速選択される.

3.3 閉塞網

閉塞網を理解するために, 基本置換を理解しておく必要がある.
基本置換には

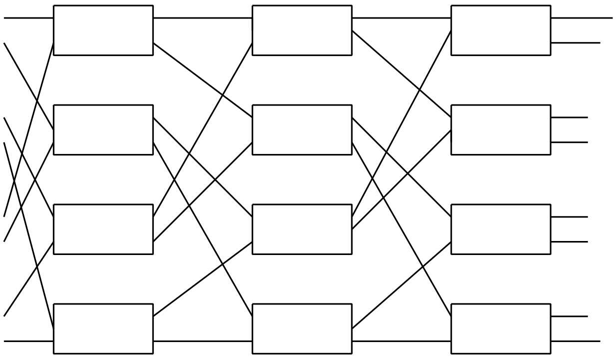

図 8: Ω網

- エクスチェンジ置換
- シャフル置換
- バタフライ置換
- ビット逆転置換
- シフト置換

閉塞網では, 可能な入出力結合パターン総数 N!のうちのごく一部を実現するものである.

1. オメガ網

   Ω網は, シャフル・エクスチェンジを logN 段連結して得られる.

   Ω網は繰り返し構造をしているため, LSI 化しやすい特徴がある.

2. 間接 2 進 n-キューブ

   先に述べた n-キューブを多段結合網で実現したもの.

3. バンヤン網

   間接 2 進 n-キューブの最終段の逆シャフル置換がないもの.

4. ベースライン網

   Benes 網の前半部分に相当する構成となっている.

3.4　多段結合網の比較

分類	網名	通信時間	スイッチ数	制御
非閉塞網	クロスバー網	1	N^2	集中 / 分散
	3 ステージ Clos 網	3	$4\sqrt{2}N^{\frac{3}{2}} - 4N$	集中 / 分散
再構成型閉塞網	Benes 網	$2\log N - 1$	$N\log N - \frac{N}{2}$	集中
	バイトニックソータ網	$\frac{\log N}{2}(\log N + 1)$	$\frac{N}{4}\log N(\log N + 1)$	集中 / 分散
	Memphis 網	3	$3N^{\frac{3}{2}}$	集中
閉塞網	Ω 網など	$\log N$	$\frac{N}{2}\log N$	集中 / 分散

現状では, 物理的な結合網を意識した並列アルゴリズムの開発がなされている. 高速多段結合網が実現できれば, プロセッサ間結合を論理レベルで考えることができる. このため, 多段結合網をガリウムヒ素などの新素子で構成したり, 光を利用したりすることが考えられている.

参考文献

[1] 富田 眞治. 並列計算機構成論. 昭晃堂, 1986.

[2] 宮原 秀夫 / 尾家 祐二 . コンピューターネットワーク. 森北出版,1992.