

Universidade Federal do Rio de Janeiro
Informática DCC/IM

Arquitetura de Computadores II

Redes de Interconexão

Gabriel P. Silva

Redes de Interconexão Estática

Redes de Interconexão Estáticas

◆ Topologias

- Unidimensionais e Bidimensionais
- Tridimensionais e Hipercúbicas

◆ Técnicas de Chaveamento

- Chaveamento por pacote (store-and-forward)
- Chaveamento por circuito
- Virtual Cut-Through
- Wormhole

◆ Algoritmos de Roteamento

- Determinístico
- Adaptativo

Redes de Interconexão Estáticas

◆ Normalmente utilizadas em arquiteturas paralelas por troca de mensagens (multicomputadores).

◆ Redes de interconexão estáticas são redes com topologia baseada em grafos, onde cada nó é um elemento processador e cada aresta do grafo representa um “link” entre dois elementos processadores.

Topologias de Redes de Interconexão Estática

◆ Linha

- Cada processador está conectado aos seus vizinhos da esquerda e da direita.
- A mensagem é repetidamente passada para o próximo nó até chegar ao seu destino.

◆ Bi-dimensional

- Anel
 - ◆ Quando o primeiro e últimos nós da topologia em linha estão interconectados
- Estrela
 - ◆ Um nó atua como nó de controle ao qual todos demais nós estão conectados
 - ◆ Por manipular toda a comunicação entre os nós, o nó central é o gargalo do sistema.

Topologias de Redes de Interconexão Estática

◆ Bi-dimensional

- Árvore
 - ◆ Uma árvore binária de profundidade d tem $2^d - 1$ nós.
 - ◆ As redes em árvore sofrem de um gargalo de comunicação nos níveis mais altos da árvore binária.
 - ◆ Este problema pode ser resolvido aumentando a capacidade de comunicação dos “links” que estão mais perto da raiz. Este rede é chamada de “Árvore Gorda”.

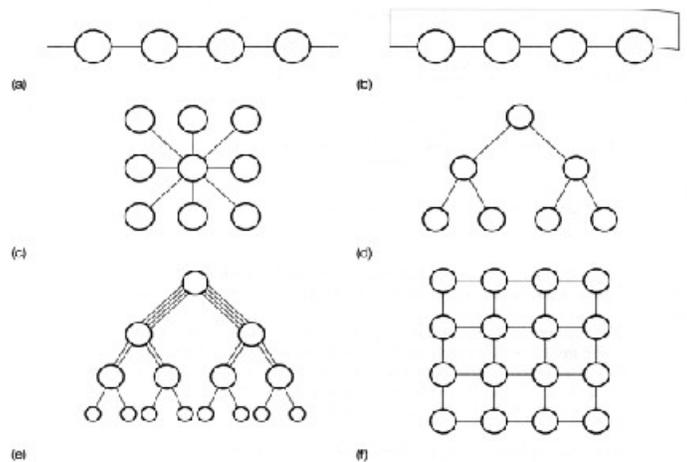
Topologias de Redes de Interconexão Estática

◆ Bi-dimensional

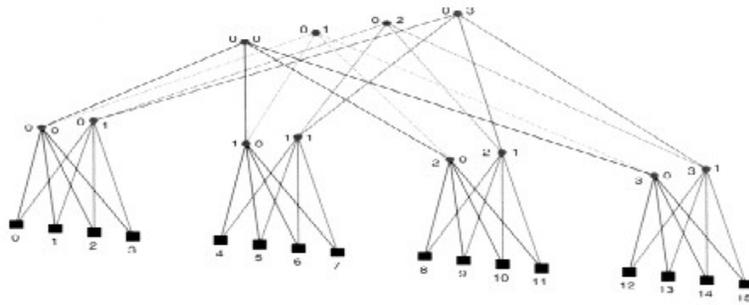
■ Malha Bi-dimensional

- ◆ Cada processador tem quatro vizinhos aos quais está conectado por um “link”.
- ◆ A malha bidimensional é uma extensão do vetor linear.
- ◆ Se as duas dimensões da malha não forem iguais, temos uma malha retangular.

Topologias de Redes de Interconexão Estática



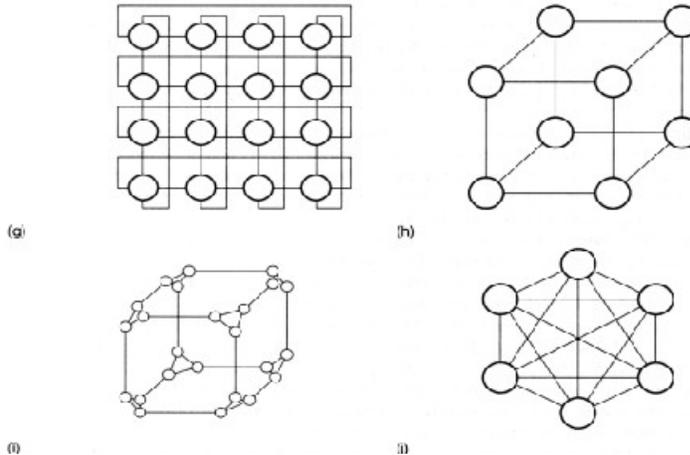
Topologias de Redes de Interconexão Estática



Interconexão Estática Tri-dimensional

- ◆ **Toro (Malha Conectada nas 2 dimensões)**
- ◆ **Cubo 3-D**
- ◆ **3-cube-connected cycle**
- ◆ **Totalmente conectada**
 - Todos os nós estão conectados entre si por um “link” direto.
 - O número de arestas do grafo totalmente conectado é dado por:
$$d = n(n-1)/2$$
 - Este tipo de rede é muito pouco utilizado devido aos altos custos de comunicação.

Interconexão Estática Tri-dimensional



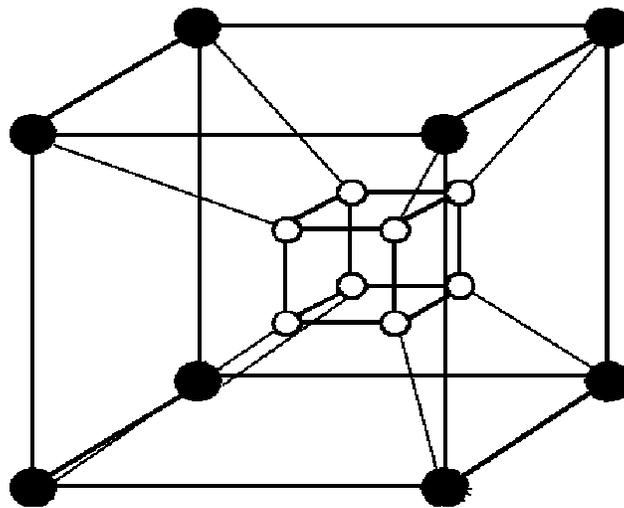
Redes de Interconexão Hipercúbica

- ◆ Um hipercubo é uma malha multidimensional de nós processadores com exatamente dois nós em cada dimensão.
- ◆ Um hipercubo com dimensão d possui um total de $n = 2^d$ processadores:
 - $d = 0$ é um hipercubo de dimensão zero, com apenas um nó;
 - $d = 1 \rightarrow n = 2$, um hipercubo com 2 nós conectados por um “link”;
 - Um hipercubo de dimensão $d+1$ consiste de dois hipercubos de dimensão d .

Propriedades da Rede Hipercúbica

1. Dois nós estão conectados por um “link” se sua numeração binária difere apenas de **um** “bit”.
2. Cada nó está diretamente conectado a outros **d** processadores .
3. O número total de posições de bits diferentes entre dois nós é chamada de Distância de Hamming, **HD**, entre eles. Esta distância é caminho mais curto para uma mensagem trafegar entre esses dois nós. Por exemplo, a HD entre o nó 3(011) e o nó 5(101) é 2.

Rede Hipercúbica



Propriedades das Redes de Interconexão Estáticas

- 1) Grau do nó:** Número de canais que incidem em um nó da rede.
- 2) Diâmetro:** é a distância máxima entre quaisquer dois nós da rede.
- 3) Conectividade:** é a medida da multiplicidade de caminhos entre dois nós quaisquer.
- 4) Largura da Biseção:** é definida como o número mínimo de “links” de comunicação que necessitam ser removidos para particionar a rede em duas metades iguais.

Propriedades das Redes de Interconexão Estáticas

- 5) Largura de Banda da Biseção:** é definido como o volume mínimo de comunicação entre duas metades da rede com igual número de nós.
- 6) Largura do Canal:** número de bits de cada “link” físico de comunicação.
- 7) Taxa do Canal:** taxa de pico de transmissão dos bits através de cada “link” físico.
- 8) Custo:** O custo de uma rede pode ser avaliado pela contagem do número de “links” requeridos por toda a rede de interconexão.

Parâmetros de Desempenho

- **Banda Passante:**
 - Taxa máxima com que a rede é capaz de transmitir a informação.
- **Latência:**
 - Intervalo de tempo gasto por uma mensagem para atravessar a rede;
 - Latência = overhead + (tam. mensagem / banda passante).
- **Escalabilidade**

Propriedades das Redes de Interconexão Estáticas

Topologia	Grau Nó	Diâmetro	Largura Bisseção	Custo (Links)
Linha	1 ou 2	$N-1$	1	$N-1$
Anel	2	$N/2$	2	N
Estrela	1 ou $N-1$	2	1	$N-1$
Árvore Binária	1, 2 ou 3	$2 \log_2^* ((N+1)/2)$	1	$N-1$
Malha 2-D	2, 3, ou 4	$2(N^{1/2}-1)$	$N^{1/2}$	$2(N-N^{1/2})$

Propriedades das Redes de Interconexão Estáticas

Topologia	Grau Nó	Diâmetro	Largura Bisseção	Custo (Links)
Toro	4	$(N^{1/2} - 1)$	$2 * (N^{1/2})$	$2 * N$
Cubo 3-D	3, 4, 5 ou 6	$3(N^{1/3} - 1)$	$N^{2/3}$	$2(N - N^{2/3})$
Hipercubo	$\log_2 N$	$\log_2 N$	$N/2$	$(N \log_2 N) / 2$
Comp. Conectada	$N - 1$	1	$(N^2) / 4$	$N(N-1) / 2$

Exemplos de Arquiteturas com Redes Estáticas

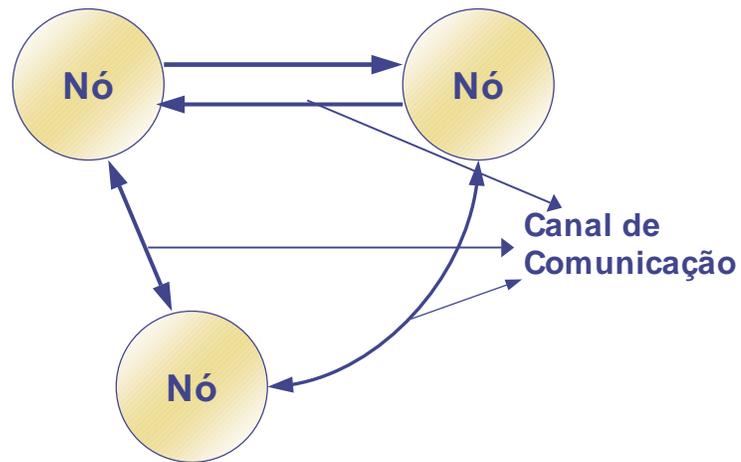
- ◆ Cray T3-E: Torus 3-D, 600 Mbytes/ s por link.
- ◆ Cray T3-D: Torus 3-D, 300 Mbytes/ s por link.
- ◆ Intel iPSC-2: Hipercubo, 2.8 Mbytes/ s por link.
- ◆ Chaos Router: Torus 2-D, 360 Mbytes/ s por link.
- ◆ MIT M-Machine: Malha 3-D, 800 Mbytes/ s por link.

Técnicas de Roteamento

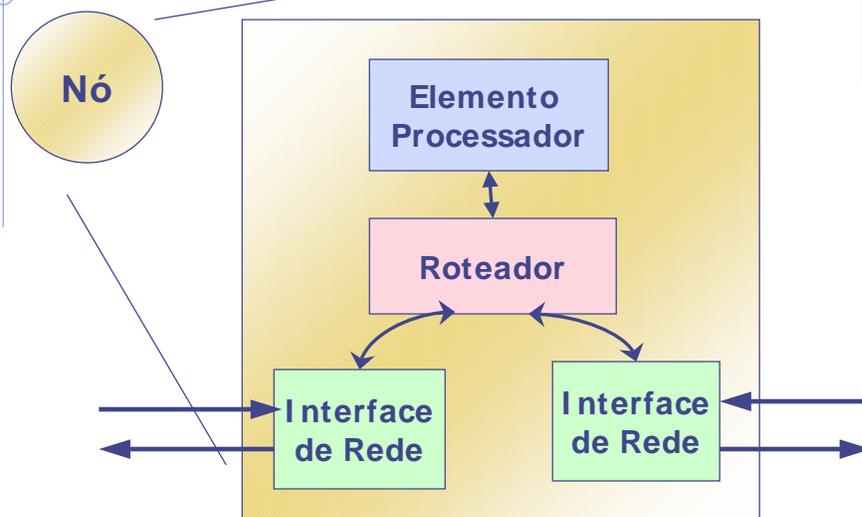
Elementos de uma Rede de Interconexão Estática

- ◆ **Nós:** são os elementos ativos da rede, que realizam a computação e o roteamento das mensagens. São compostos por:
 - Interface de Rede
 - Roteador
 - Elemento Processador
- ◆ **Canais:** conexões ponto-a-ponto por onde trafegam as mensagens.

Elementos de uma Rede de Interconexão Estática



Elementos de uma Rede de Interconexão Estática



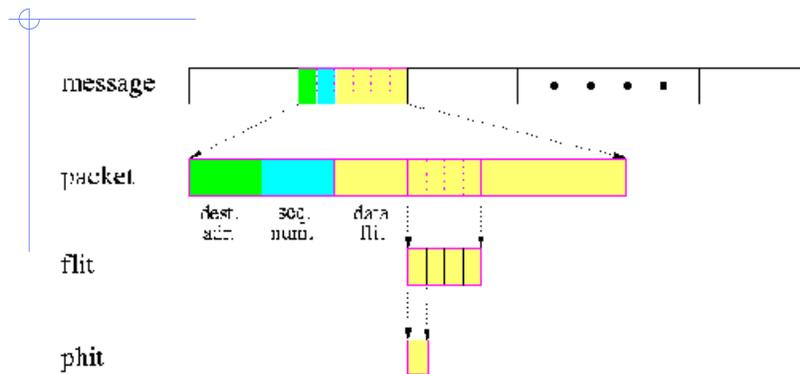
Transmissão da Mensagem

- ◆ Como as mensagens são transmitidas através da rede de interconexão?
- ◆ Existem dois métodos básicos:
 - Chaveamento de Pacotes
 - ◆ A mensagem é dividida em pacotes que são enviados individualmente pela rede.
 - Chaveamento por Circuito
 - ◆ Um circuito físico é estabelecido entre o destino e a origem para a transmissão da mensagem.

Chaveamento de pacotes

- ◆ Conhecido também como “store-and-forward”
- ◆ A mensagem é dividida em pacotes que são enviados independentemente através da rede de conexão
- ◆ O pacote será transmitido para o nó vizinho apenas se houver espaço disponível para o armazenamento
- ◆ $L = (P / B) * D$
 - P → comprimento do pacote
 - B → largura de banda do canal
 - D → distância entre os nós
- ◆ A latência é proporcional à distância entre os nós

Formato da Mensagem



Formato da Mensagem

- ◆ **mensagem:** A unidade de comunicação do ponto de vista do programador. Seu tamanho é limitado apenas pelo espaço na memória de usuário.
- ◆ **pacote:** Menor unidade de comunicação de tamanho fixo contendo informação de roteamento (p. ex., endereço de destino) e de sequenciamento no seu cabeçalho. Seu tamanho é da ordem de centenas a dezenas de bytes ou palavras.

Formato da Mensagem

- ◆ **flit**: Menor unidade de informação no nível do “link”, com o tamanho de uma ou várias palavras. Os Flits podem ser de diversos tipos e o protocolo para o envio de um flit consome diversos ciclos.
- ◆ **phit**: A menor unidade de informação no nível físico que é transferida através de um link físico em um ciclo.

Chaveamento por circuito

- ◆ Todo um caminho é estabelecido pelo envio de uma pequena mensagem de “sonda” antes do envio da mensagem principal.
- ◆ Os canais que constituem o circuito são reservados exclusivamente.
- ◆ $L = (P/B) * D + M/B$
 - $P \rightarrow$ comprimento da mensagem de “sonda”
 - $M \rightarrow$ comprimento da mensagem
 - Se $P \ll M$, latência independe da distância

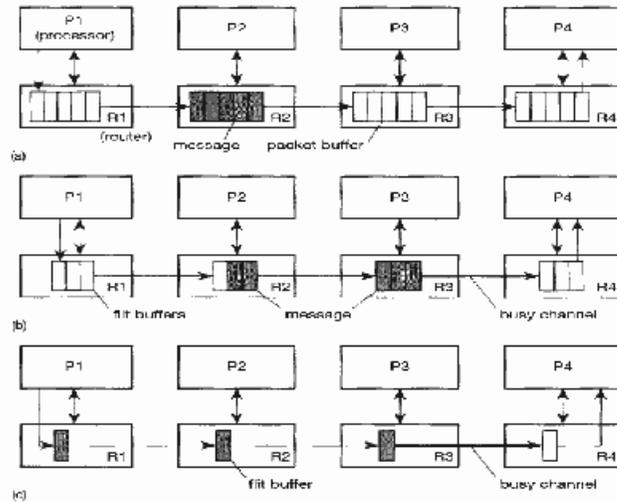
Virtual cut-through

- ◆ Solução de compromisso.
- ◆ A mensagem é dividida em pequenas unidades chamadas “flow control digits” ou “flits”.
- ◆ Os flits são enviados, em modo pipeline, enquanto os canais estiverem disponíveis. Se algum canal requisitado estiver ocupado, os flits são armazenados nos nós intermediários .
- ◆ $L = (HF/ B) D + M/ B$
 - HF → comprimento do flit de cabeçalho
 - Se $HF \ll M$, então independe da distância

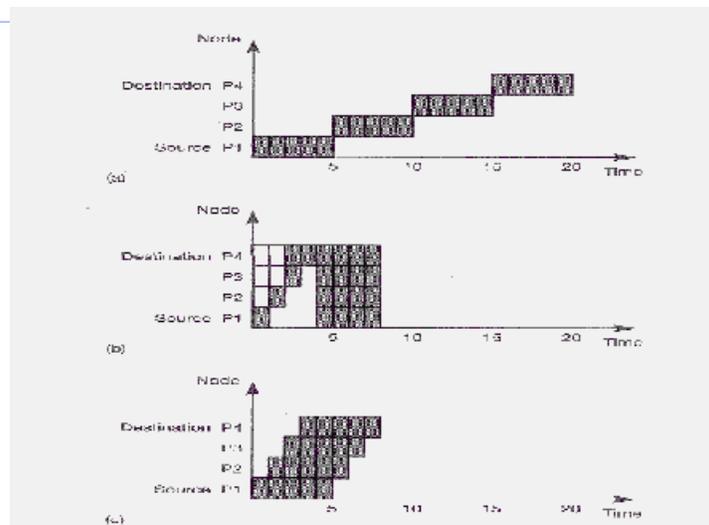
Wormhole Routing

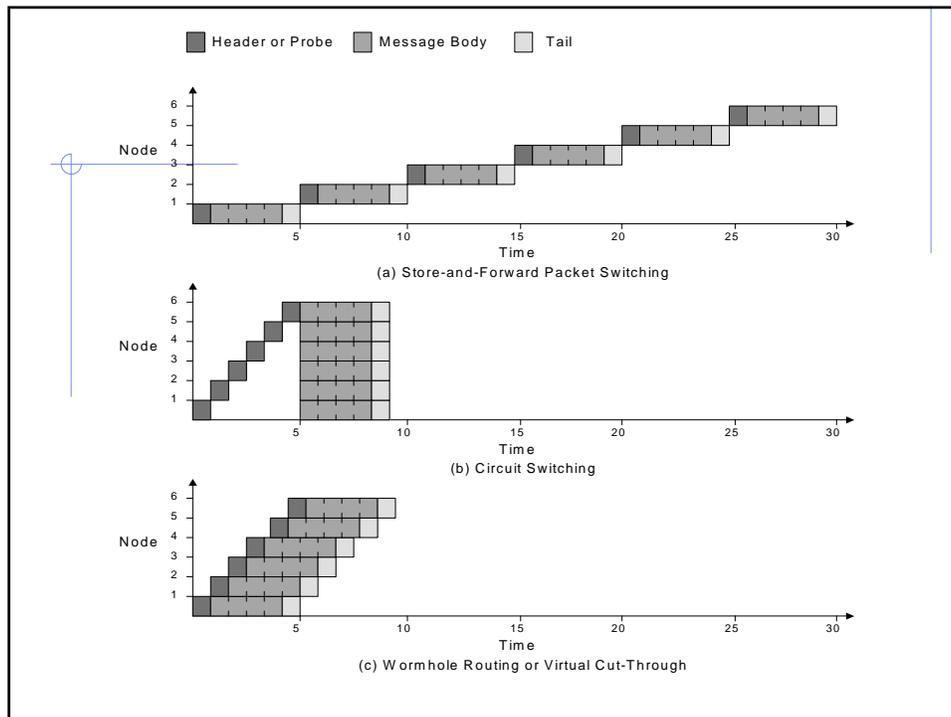
- ◆ É um caso especial do “virtual cut-through” onde a capacidade de armazenamento nos nós intermediários é igual a 1 flit.
- ◆ Pode realizar replicação de pacotes, enviando cópias de flits para diversos canais de saída para implementar “multicast” e “broadcast”.
- ◆ Com o uso de múltiplos “buffers” para cada canal, é possível implementar canais virtuais, para que diversas mensagens possam compartilhar o mesmo canal físico.
- ◆ $L = (HF/ B) * D + M/ B$
 - HF → comprimento do flit de cabeçalho
 - Se $HF \ll M$, então independe da distância

Técnicas de Chaveamento



Técnicas de Chaveamento





Broadcasting

- ◆ Algumas redes de interconexão possuem recursos de hardware no roteamento de mensagens para suportar diferentes tipos de operações de comunicação.
- ◆ Todas as redes suportam a comunicação ponto-a-ponto ou “unicast”.
- ◆ Operações coletivas:
 - “broadcasting” (um nó origem envia uma mesma mensagem para todos os outros)
 - “multicasting” (um nó origem envia a mesma mensagem para um grupo especificado de nós destino).

Algoritmos

◆ Determinísticos

- O caminho é completamente determinado pelo endereço dos nós fonte e destino. Os nós intermediários, mesmo no caso de um congestionamento, não podem redirecionar as mensagens.

◆ Adaptativos

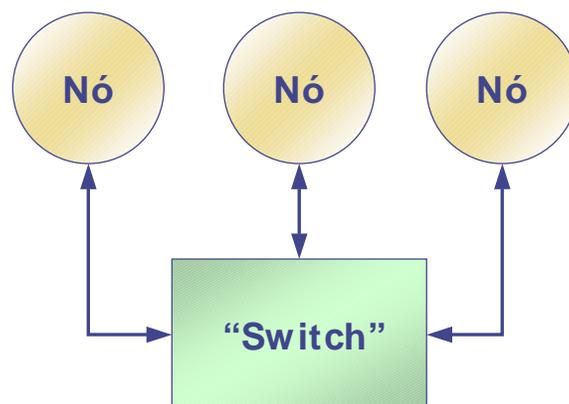
- No roteamento adaptativo os nós intermediários levam em conta o estado atual da rede para determinar a direção para qual a mensagem deve ser enviada.

Redes de Interconexão Dinâmicas

Redes de Interconexão Dinâmicas

- ◆ Também conhecidas como redes indiretas.
- ◆ A comunicação entre os nós se dá através de elementos de chaveamento (“switchs”).
- ◆ Diversas topologias possíveis podem ser utilizados nos elementos de chaveamento.

Redes de Interconexão Dinâmicas



Redes de Interconexão Dinâmicas

- ◆ **As duas grandes topologias possíveis que um elemento de chaveamento ou roteador pode ter são:**
 - **“Crossbar”**
 - **Multiestágio**

Redes Crossbar

- ◆ **Permitem o acesso simultâneo de todas as entradas a todas as suas saídas.**
- ◆ **Conflito só existe se dois ou mais processadores quiserem se comunicar com o mesmo módulo de memória/ processador.**
- ◆ **Custo muito elevado de interconexão**
- ◆ **Requer $p \times m$ elementos de chaveamento.**

Redes Multiestágio

- ◆ Representam um solução de compromisso entre o barramento único e a “crossbar switch”.
- ◆ As entradas são conectadas às saídas através de vários estágios, onde cada estágio é composta por diversas chaves do tipo “crossbar”.
- ◆ Elas podem ser classificadas de acordo com o número de estágios, a topologia de interconexão entre os estágios e o tipo de chave empregado em cada estágio.

Redes Multiestágio

- ◆ As conexões entre os estágios são também chamadas de permutações.
- ◆ Entre as diversas permutações possíveis, destacamos:
 - Embaralhamento Perfeito
 - Reversão de Dígitos (Bits)
 - Butterfly
 - Cúbica
 - Baseline

Redes Multiestágio

- ◆ **Embaralhamento Perfeito**
 - Faz uma rotação para a esquerda de um bit no endereço fonte para obter o destino.
- ◆ **Reversão de Dígitos**
 - É feito um espelhamento dos bits do endereço.
- ◆ **Butterfly**
 - Os bits de ordem 0 e i bits do endereço são trocados entre si.
- ◆ **Cúbica**
 - Complementa o bit i do endereço.
- ◆ **Baseline**
 - Faz uma rotação para a direita dos $i+1$ bits menos significativos do endereço fonte .

Permutações Perfeita e Reversão de Bit

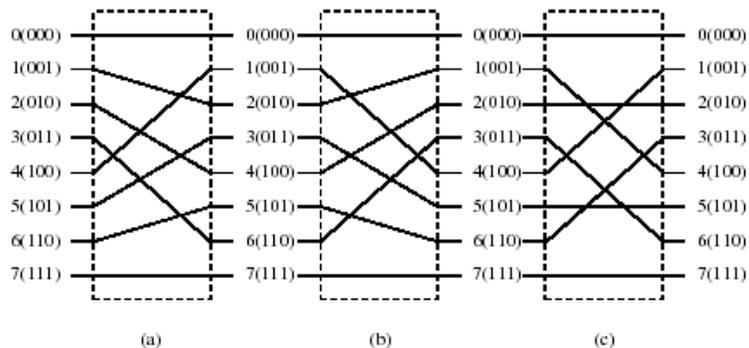


Figure 1.13 (a) The perfect shuffle, (b) inverse perfect shuffle, and (c) bit reversal permutations for $N = 8$.

Permutação Butterfly

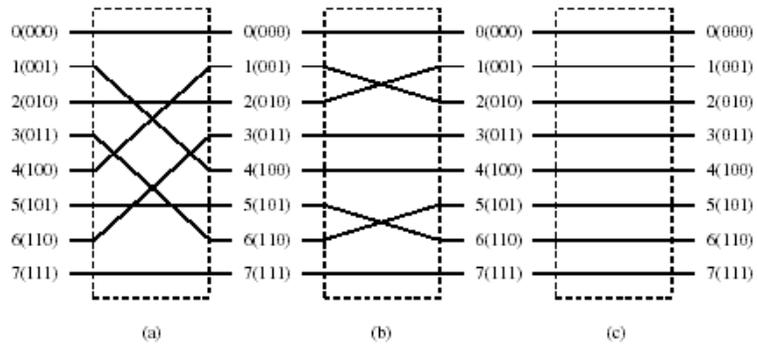


Figure 1.14 The butterfly permutation for $N = 8$: (a) second butterfly, (b) first butterfly, and (c) zeroth butterfly.

Permutação Cúbica

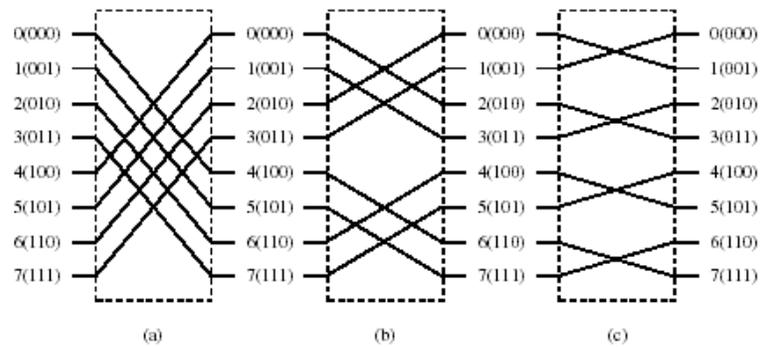


Figure 1.15 The cube permutation for $N = 8$: (a) second cube, (b) first cube, and (c) zeroth cube.

Permutação Baseline

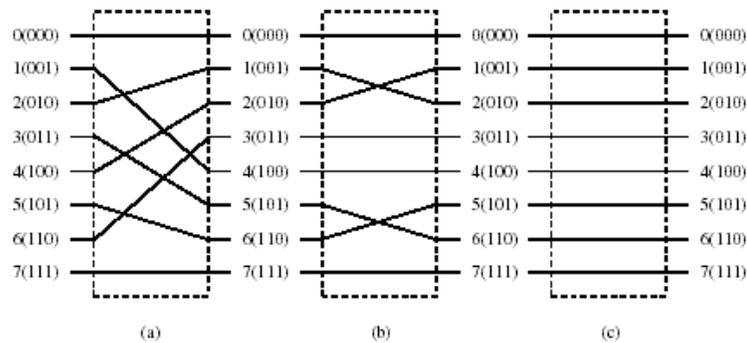


Figure 1.16 The baseline permutation for $N = 8$: (a) second baseline, (b) first baseline, and (c) zeroth baseline.

Redes Multiestágio

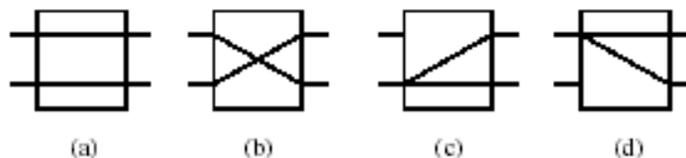
- ◆ A partir das permutações mostradas entre os estágios, vários tipos de redes podem ser montadas.
- ◆ Normalmente o elemento de chaveamento das redes que iremos mostrar é 2×2 , mas a mesma topologia pode ser encontrada com chaves 4×4 ou mesmo 8×8 .
- ◆ Uma rede com chaves de tamanho $k \times k$ vai precisar de $\log_k N$ estágios para interligar N elementos processadores.

Redes Multiestágio

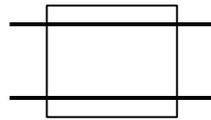
- ◆ As redes multiestágio podem ser classificadas como:
 - Bloqueantes – Nem sempre é possível a conexão entre um par livre de processadores;
 - Não-bloqueantes – Qualquer entrada livre pode ser ligada a qualquer saída livre sem alterar as demais conexões;
 - Reconfiguráveis – Qualquer entrada pode ser ligada a qualquer saída livre, mas pode ser necessário que algumas ligações sejam refeitas.

Redes Multiestágio

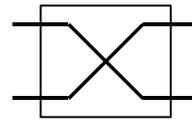
- ◆ Quatro estados possíveis de uma chave 2x2:
(a) direto, (b) permutação, (c) broadcast inferior, (d) broadcast superior.



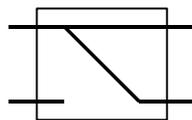
Redes Multiestágio



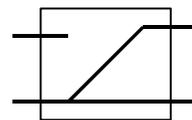
(a) Direto



(b) Permutação



(c) Broadcast Superior

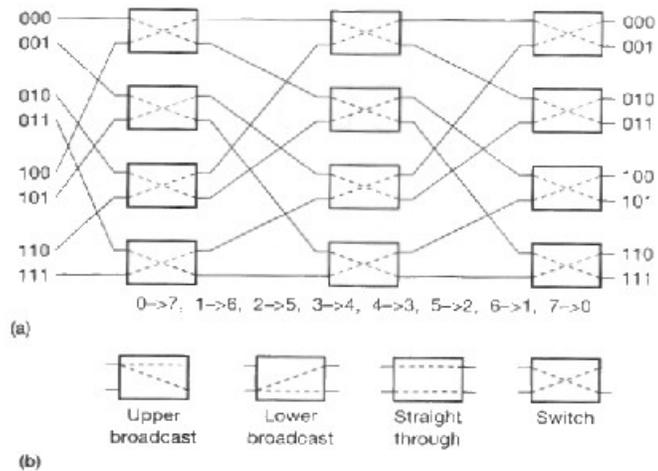


(d) Broadcast Inferior

Rede Omega

- ◆ É o tipo mais simples de rede multiestágio.
- ◆ Possui $\log_2 N$ estágios com $N/2$ chaves 2×2 em cada estágio.
- ◆ Cada chave tem 4 posições possíveis de interconexão.
- ◆ O padrão de interconexão C_i é um embaralhamento perfeito para $0 \leq i \leq n-1$. O padrão de interconexão C_n é a permutação identidade, isto é cada origem está ligada no destino de mesmo endereço.

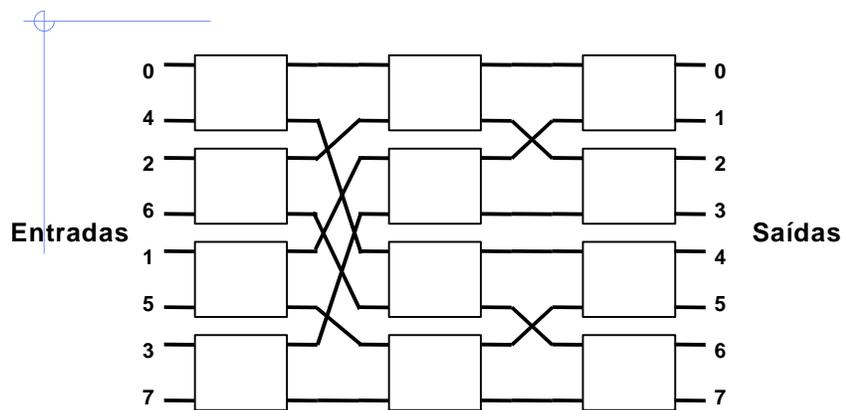
Rede Omega



Rede Baseline

- ◆ Possui $\log_2 N$ estágios com $N/2$ chaves 2×2 em cada estágio.
- ◆ O padrão de interconexão C_i é uma permutação “baseline” de ordem $(n-1)$ para $1 \leq i \leq n$. O padrão de interconexão C_0 é um embaralhamento perfeito.
- ◆ Tem como particularidade controlar a chave do estágio i com o $(n-i)$ -ésimo bit do endereço de destino.

Rede Baseline

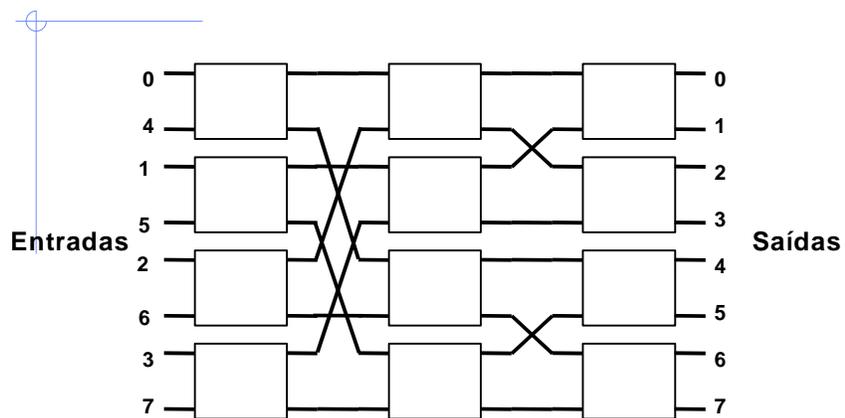


(a) Rede Baseline

Rede Cubo Generalizado

- ◆ Possui $\log_2 N$ estágios com $N/2$ chaves 2×2 em cada estágio como a rede omega.
- ◆ Difere apenas na topologia com que os estágios são interconectados após o primeiro estágio. Ao invés do embaralhamento perfeito são ligadas como descrito a seguir.
- ◆ O padrão de interconexão C_i é uma permutação “butterfly” de ordem $(n-i)$ para $1 \leq i \leq n$. O padrão de interconexão C_0 é um embaralhamento perfeito.

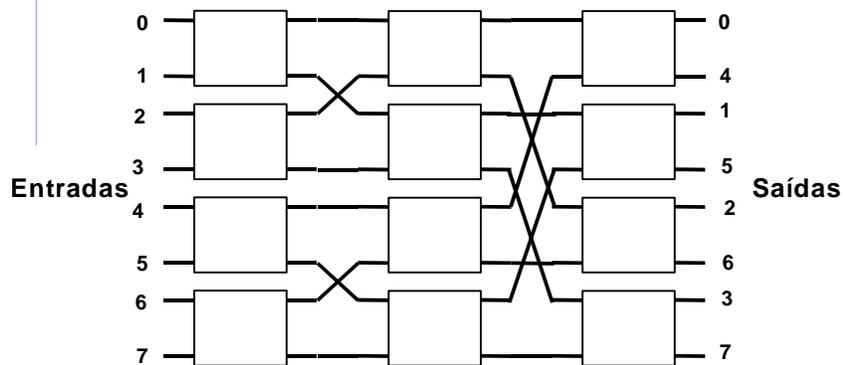
Rede Cubo Generalizado



(a) Rede Cubo Generalizado

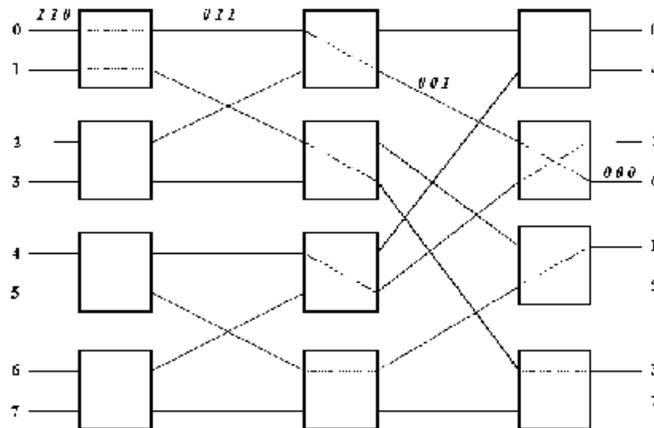
Rede N-Cubo Binário Indireto

◆ É o espelho da rede cubo generalizado.



(a) Rede Cubo Binário Indireto

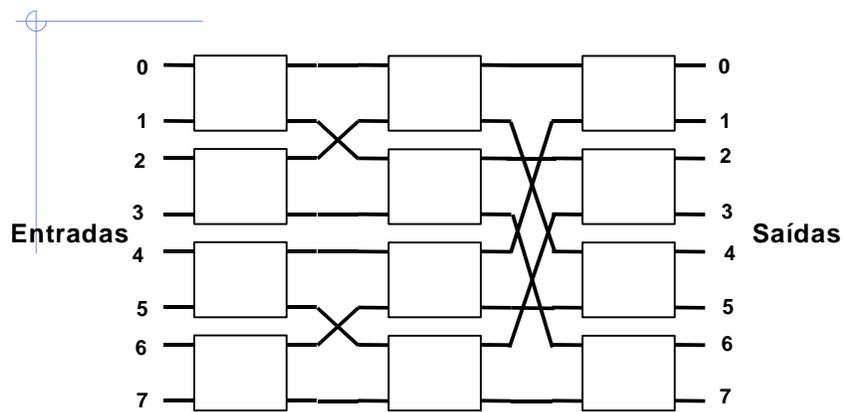
Rede do Multiplus



Redes Butterfly

- ◆ Nesta rede o padrão de interconexão C_i é a permutação “butterfly” de ordem i , para $0 \leq i \leq n-1$.
- ◆ O padrão de conexão C_n é a identidade.
- ◆ Uma versão muito utilizada substitui as chaves 2×2 por chaves 8×8 .
- ◆ Neste caso, o número de estágios diminui para $\log_8 N$ e o número de chaves por estágio para $N/8$.
- ◆ Possui as mesmas propriedades da rede omega.

Redes Butterfly



(a) Rede Butterfly

Rede Benes

- ◆ As redes anteriormente apresentadas são do tipo bloqueante, ou seja, certas combinações de endereçamento não são possíveis.
- ◆ A rede do tipo Benes possui estágios adicionais para prover caminhos redundantes no esquema de interconexão.
- ◆ O preço pago é um aumento no custo, latência e tamanho da rede de interconexão.
- ◆ Exemplos de arquitetura com rede multiestágio são NYU Ultracomputer, CEDAR, HEP e Multiplus.

Rede Benes

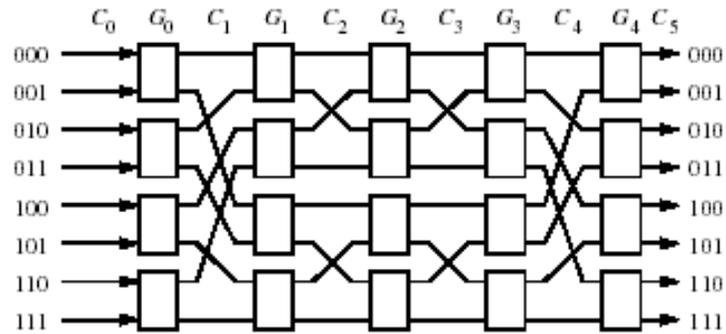


Figure 1.17 An 8×8 Benes network.

Resumo das Propriedades

Tipo da Rede	Nº de Estágios	Nº de Chaves / Estágio	Topologia entre os estágios	Tam. da Chave	Modo de Oper.
Omega	$\log_2 N$	$N/2$	Perfect Shuffle	2×2	Bloq.
Butterfly	$\log_8 N$	$N/8$	Butterfly	2×2 ou 8×8	Bloq.
Cubo Generalizado	$\log_2 N$	$N/2$	Perfect Shuffle (C_0) Butterfly ($C_1 \rightarrow C_n$)	2×2	Bloq.
Benes	$(2 * \log_2 N) - 1$	$N/2$	Perfect Shuffle (C_1, C_{n-1}) Butterfly ($C_2 \rightarrow C_{n-2}$)	2×2	Não Bloq. Rearranj.

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.