

## SISTEMA HARDWARE PARA PROCESSAMENTO PARALELO

Cavalli, E.

Zabeu, M. C.

O trabalho apresenta uma estrutura hardware voltada para processamento paralelo baseada em microprocessador de 16 bits. O projeto foi desenvolvido no CPqD - TELEBRÁS. São delineadas as características do "hardware" e os estudos efetuados para a determinação da estrutura para processamento paralelo, junto às conclusões obtidas nestes estudos.

### 1.0 INTRODUÇÃO

O presente trabalho visa apresentar a arquitetura de um sistema para multiprocessamento composto por placas desenvolvidas no Centro de Pesquisa e Desenvolvimento da Telebrás - CPqD dentro do projeto "Processador Preferencial".

As placas que compõem o presente sistema são divididas funcionalmente em:

- placas de processamento (UCP's), baseadas no microprocessador iAPX 286;
- placas periféricas inteligentes, controladas por microprocessadores 8085. Funcionam em um esquema de comunicação com as placas UCP baseado em interrupções e uso de memória partilhada. Coordenam o acesso a uma gama ampla de periféricos como discos flexíveis e discos rígidos (DIS), fitas (FMG) e terminais de vídeo (SER, UCP, FMG, DIS), entre outras;
- placas de memória (MEM) de até 2M bytes de capacidade cada uma;
- barramento de interconexão com suporte para multiprocessamento, com capacidade de escoamento de até 10M bytes/s. Controla acessos concorrentes num esquema de "daisy-chain".

A estrutura do sistema é modular, permitindo implementações várias como a em questão, voltada para processamento paralelo com baixo acoplamento.

## 2.0 ESTRUTURA DO SISTEMA

O sistema está estruturado conforme mostra o esquema na figura 1.

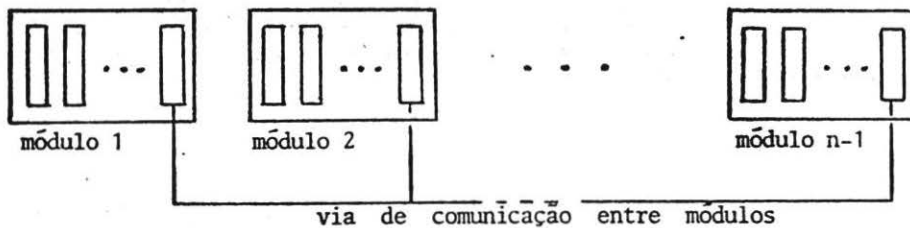


figura 1 - estrutura do sistema de multiprocessamento

Um número  $n$  de módulos são interligados por meio de uma via de comunicação. O número de módulos depende da aplicação e da capacidade da via de comunicação conforme será analisado a seguir.

O módulo é composto por um conjunto de placas, onde existe necessariamente uma placa de processamento (UCP) e uma placa para intercomunicação entre os módulos (UCC). O módulo inclui uma mecânica as fontes de alimentação e os eventuais equipamentos controlados pelas placas

### 2.1 Estrutura do módulo

Um exemplo de módulo é representado pelo esquema da figura 2.

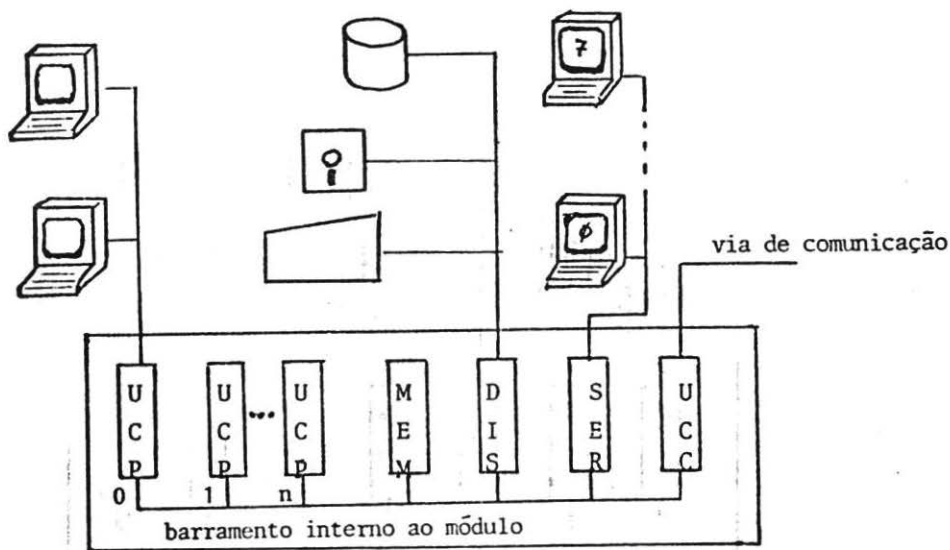


figura 2 - Exemplo de estrutura do módulo

O módulo é constituído de um sub-bastidor com um conjunto de placas processadoras (UCP) e pelo menos uma placa de comunicação (UCC). A placa de memória (MEM) pode ser necessária se o módulo contiver mais de uma UCP. As placas controladoras de periféricos (DIS, SER, etc) assim como a utilização das duas interfaces seriais das UCPs podem estar ou não presentes dependendo da configuração do sistema,

É evidente que estas funções deverão estar presentes em pelo menos um módulo.

O módulo compreende, além do sub-bastidor com as placas, as fontes de alimentação e os equipamentos locais controlados diretamente pelo módulo (eventuais terminais, impressoras, unidades de disco, etc).

Os fatores que levam a escolher o número n de UCP a serem colocadas no módulo são descritos a seguir.

#### 2.1.1 Determinação do número de placas UCP -

As placas UCP têm como características básicas:

- relógio de funcionamento de 8 ou 10MHz;
- memória RAM interna de até 512K bytes, com lógica de detecção e correção de erros;
- memória EPROM residente de até 64K bytes;
- 2 interfaces seriais padrão RS-232 ou RS-422;

- controlador de interrupção;
- temporizadores

A quantidade de 8 placas UCP em um módulo foi determinada com base em estudos de utilização efetiva destas e do barramento interno ao módulo conforme será descrito a seguir.

O barramento interno tem uma capacidade nominal de escoamento de dados de 10M bytes/s. A avaliação da largura de banda efetiva deste meio é feita a partir da estimativa do tempo de acesso a dados na UCP. Para o funcionamento com relógio de 8MHz (10MHz), a taxa nominal de acesso a dados na UCP é de 8M bytes/s (10M bytes/s). Entretanto, esta taxa é reduzida em função do tempo de acesso à memória.

O tempo de acesso à memória é elevado do valor potencial do iAPX286 de 250 ns para cerca de 375 ns. Com isso a velocidade do iAPX 286 é reduzida em função desse tempo, de forma que a largura efetiva de banda de acesso à memória é de cerca de 5,33M bytes/s (tomando acesso a palavras). Sendo o acesso à memória mais restritivo em questões de tempo dentro do funcionamento global do sistema a taxa de acesso ao barramento é estimada em função deste dado. Assim, a largura de faixa efetiva do barramento é também de cerca de 5,33M bytes/s (o barramento é ocupado durante todo o ciclo de transferência).

Para estimar o uso do barramento por parte da UCP considera-se que um software normal tem mais de 80% dos seus acessos à memória voltados à busca de instruções e acesso à pilha.

Desta maneira, não mais que 20% dos acessos são dedicados à busca de dados e destes pode-se assumir que somente 1/5 residem fora do contexto do programa (dados de comunicação). Se o código referente à operação residir integralmente na UCP (ou seja, os programas são limitados em até 512K bytes RAM e 64K bytes EPROM), tem-se que 4% dos acessos são executados no barramento. Assim uma aplicação normal, residente em uma só UCP, toma o barramento numa taxa de 4% de seus acessos à memória.

Considerando a largura de banda efetiva do barramento de cerca de 5,33M bytes, tem-se que a utilização do mesmo por uma UCP leva a um fluxo de dados externos de cerca de 200K bytes, o que é razoável para várias aplicações.

A restrição de se manter o código a ser executado interno à UCP vem do fato de que, se 80% dos acessos para busca de instruções e pilha forem feitos sobre o barramento, este ficará saturado com apenas uma UCP, ou seja, a função básica de intercomunicação fica prejudicada frente à utilização excessiva para acesso ao código. Assim, a manutenção do código interno à UCP garante a utilização efetiva do barramento para transferência de dados entre processadores e para operações de E/S.

Uma vez avaliada a taxa de utilização do barramento por parte de uma UCP estimou-se qual a taxa efetiva de trabalho realizada por um conjunto de placas UCP. Para isso, foi levado em conta que existe uma quantidade máxima de trabalho realizada por um conjunto de processadores, independente do número destes, que é função do montante de acessos "externos" implícitos em uma dada

aplicação. Ou seja, dado que uma aplicação necessita de uma taxa  $x$  de trabalho "externo" (acessos portanto ao barramento), o uso efetivo  $T$  de um processador em um conjunto de  $n$  processadores é dado por:

$$T = \frac{1}{1 + x(n - 1)} \quad (1)$$

Com isso, para vários conjuntos de processadores utilizando o barramento a uma taxa de 4% conforme deduzido anteriormente tem-se:

no. de processadores	uso efetivo de 1 processador	equivalência de processadores
4	89%	3,57
8	78%	6,25
16	62,5%	10

Tab. 1 - Comparação entre quantidade de processadores e uso efetivo dos mesmos

Por esta visão tem-se que a quantidade de 8 placas UCP em um módulo é um valor razoável dentro de um ponderação entre o uso efetivo de um processador e a equivalência de utilização do conjunto de processadores. Este valor, utilizado também nas estimativas de consumo, custo, mecânica ligadas ao módulo mostrou-se adequado para a arquitetura em questão.

#### 2.1.2 Estrutura de intercomunicação em um módulo -

A comunicação entre as placas UCP em um módulo leva em conta a não possibilidade de acesso à memória interna de uma UCP por outra UCP. Desta maneira, a comunicação entre placas deve ser feita via envio e recepção de mensagens ou via a existência de uma área comum de dados. Esta área comum de dados é implementada na forma de uma memória externa e é obtida com uso de placas do sistema "processador preferencial" com capacidade de armazenamento de até 2M bytes.

Com a alocação preferencial do barramento para fins de transferência de dados entre placas tornou-se mais razoável a implementação da intercomunicação via memória externa.

Para isso, foi necessária a determinação de um mecanismo de indicação da existência de mensagens entre os processadores. Duas opções foram avaliadas para este fim, conforme descrito a seguir:

- a) detecção periódica de mensagem: implementada através de rotina acionada por uma interrupção de relógio, tendo como função determinar a existência de mensagens em uma região da memória externa. Esta opção tem como desvantagem o tempo de latência introduzido pela característica periódica

da avaliação e o tempo gasto pela UCP em verificação infrutíferas. O tempo gasto para verificação é pequeno (de ordem de 5 microseg.), mas o tempo de latência introduzido é função da frequência de relógio e este não pode ser muito alto para não sobrecarregar o sistema. Com um relógio de 10mseg. a latência média é de 5 mseg.

- b) Ativação do processo de tratamento de mensagens via interrupção: Esta opção é possível devido à existência de um controlador de interrupções na placa UCP capaz de gerar até 8 interrupções separadamente. O tempo de latência introduzido neste caso é da ordem da demora ao atendimento desta interrupção. Esta está contida no bloco das interrupções de alta prioridade, sendo precedida apenas pela interrupção relacionada ao relógio do sistema. Cada UCP gera a interrupção para os demais sete UCP. Dos 8 bits 7 são saídas e uma é entrada habilitada a gerar interrupção. Para possibilitar o controle das linhas por mais que uma UCP os mesmos devem ser do tipo coletor aberto e a interrupção deverá ser gerada por borda.

A seguinte topologia foi adotada:

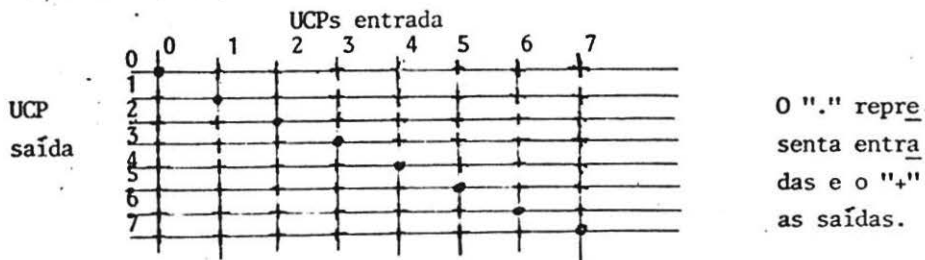


figura 3 - topologia de interligação das UCPs

A gerência da memória deve garantir a integridade da informação nela contida. Para isso contém marcações contra multiprocessamento sobre a mesma área (a marcação é feita mantendo em "lock" o barramento) e sinalização de uso da área por parte do produtor ou do consumidor por meio de "flags". Este controle é feito pelo software de controle do sistema.

A arquitetura da memória foi estudada com duas formas:

- memória com reserva de áreas específicas ("canais" para cada processador);
- memória gerenciada por tabelas, onde as áreas de uso de cada processador são alocadas dinamicamente numa filosofia de lista ligada.

Da forma (a) tem-se, para o uso de uma placa de memória de 2M bytes no módulo, canais de 256K bytes para cada UCP. Da forma (b) este tamanho pode variar conforme as necessidades da aplicação para intercomunicação. Entretanto, a implementação de estruturas de controle de acesso à memória é facilitada na forma (b) uma vez que pode ser contida na própria tabela de gerência.

Cumpra-se notar que 256K bytes é uma quantidade razoável de dados para intercomunicação para a maioria das aplicações.

### 2.1.3 Comunicação com placas controladoras de E/S -

As placas periféricas existentes não têm previsão de tratamento de multiprocessamento, ou seja, existe um mecanismo de comunicação entre as placas e UCP's por interrupção mas não há maneiras de diferenciar qual o solicitante do serviço. Duas opções se apresentaram:

- a) Realocação dinâmica do tratamento de entrada/saída a cada processador conforme necessidade do processo. A identificação do processador requisitante deve ficar na área de memória partilhada entre a placa UCP e a placa periférica em questão, exigindo assim a manutenção de estrutura de dados mais complexas que as atualmente existentes e assim mudando a filosofia genérica de utilização das placas. A existência de um mecanismo de verificação periódica da disponibilidade da placa a ser utilizada conjunto à interrupção associada ao fim de operação da placa periférica introduz complexidade adicional sem nenhuma vantagem.
- b) Alocação fixa de recursos para determinadas UCP's de forma a que a aplicação deve estar preparada para acessar os periféricos através de UCP's pré-estabelecidas na configuração do sistema. Esta opção, apesar de mais restritiva, apresenta-se mais simples e imediata para a estruturação do módulo, e portanto foi a escolhida para a implementação.

## 2.2 Comunicação entre módulos

Com base no assumido no item 2.1.1, quanto à utilização do barramento por uma aplicação, tornando-se os dados referentes à comunicação entre módulos como 25% dos dados externos a UCP, tem-se para cada UCP uma banda de 50K bytes/seg de comunicação entre os módulos. Assim, se para uma dada aplicação tem-se o uso de  $n$  módulos, a via de comunicação entre eles tem uma largura de banda efetiva de  $(n * 50K)/2$  bytes/s. Para evitar gargalos a via de comunicação deverá ter uma capacidade superior àquela indicada na fórmula anterior.

As placas existentes atualmente para intercomunicação baseiam-se no protocolo Ethernet e em um sistema de comunicação desenvolvido para uma central telefônica CPA-T.

### 2.2.1 Placa Ethernet -

Pode ser configurada com taxas de transmissão a 1 ou 10M bits/s. O protocolo de acesso à placa é o CSMA/CD (aleatório com tratamento de colisão). O controle da placa é feito na filosofia de memória partilhada. A interface é implementada usando o integrado 82586. O meio físico de conexão é padrão Ethernet, implementado pela interface serial CI 82501.

Com esta interface podem ser ligados não mais que 4 - 5 módulos completos com 8 UCP devido a sobrecarga na via de comunicação.

### 2.2.2 Sistema de comunicação da Central Telefônica CPA-T "TRÓPICO" -

O sistema de comunicação do Trópico visa interligar um grande número de módulos (até 1024) que, na grande maioria, contém placas específicas para o controle do processo de comutação telefônica. Muitos módulos contêm um único processador e grande parte dos processadores é baseada no microprocessador de 8 bits 8085H da Intel. O tráfego não é muito intenso e a grande maioria dos sinais tem tamanho pequeno (10 - 20 bytes de dados úteis).

A estrutura de comunicação é representada de forma simplificada, pela figura 4

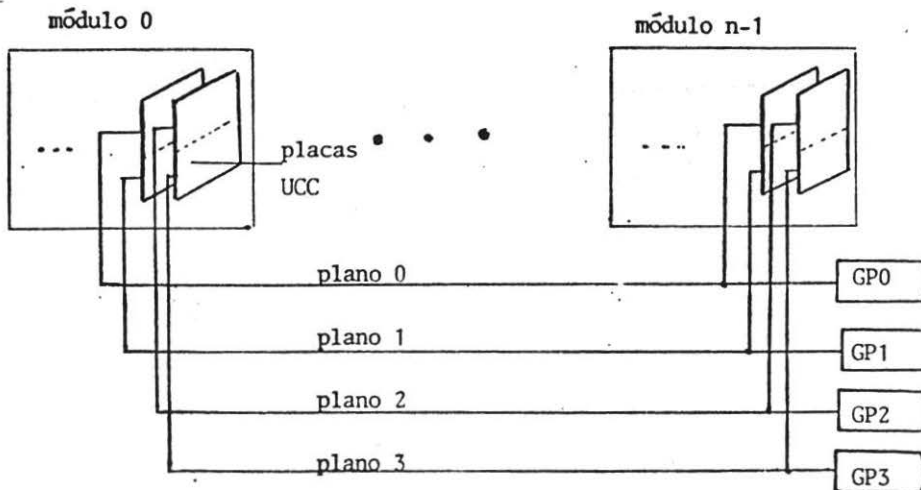


figura 4 - Sistema de comunicação

Nesta estrutura existem 2 ou 4 planos de comunicação. Cada plano é controlado por um gerenciador de plano (GP). Cada placa UCC dá acesso a uma via de comunicação dentro de dois planos. Em cada plano existem 6 vias de comunicação de 2M bits/seg. A velocidade efetiva de transferência de dados úteis depende (muito) do tamanho do sinal e consegue ser maior que 1M bits/seg, em caso de sinais que contêm mais que 50 bytes de dados úteis. Contribuem para a perda de velocidade os seguintes elementos:



- cabeçalho do sinal (22 bytes com endereço de origem, endereço de destino, identificador do sinal, etc)
- protocolos de camada e de liberação do sinal
- transmissão do código de CRC (2 bytes)

Ignorando as perdas na velocidade efetiva de transmissão teremos em caso de utilização de 4 planos as seguintes velocidades de comunicação:

- velocidade teórica de comunicação global de um módulo:

$$4 * 2/8 = 1M \text{ bytes/seg}$$

- velocidade teórica de comunicação global da rede

$$4 * 6 * 2/8 = 6 M \text{ bytes/seg}$$

A velocidade efetiva de comunicação é reduzida não somente pelos problemas acima, mas também pela ocorrência de bloqueio nos planos.

Este sistema de comunicação pode ser utilizado eficientemente para conectar até 12 - 16 módulos completos (cada um com 8 UCP) nas hipóteses de tráfego feitas anteriormente (400K Bytes/seg. por módulo) e considerando que pelo menos um plano deve ser considerado como reserva.

Uma vantagem deste sistema de comunicação é o baixo custo das placas UCC. Por outro lado os gerenciadores de plano (GP) são relativamente complexos e introduzem no sistema 3 ou 4 tipos de placas diferentes.

### 3.0 PERSPECTIVAS DE EXPANSÃO

A utilização do processador iAPX 386 na placa UCP deve abrir novas perspectivas de expansão da atual arquitetura, notadamente devido à menor utilização do barramento pela possibilidade de acesso à memória interna entre UCP's, maior capacidade de processamento, maior memória interna a placa (de 2 a 8M bytes), utilização de memória "cache" para atender às necessidades de velocidade ao acesso à memória (a 16MHz, o iAPX 386 requer um ciclo de memória de 125 ns), além de processamento numérico de maior performance com o uso dos processadores numéricos 387 ou 1167.

Observe-se que o acesso direto à memória interna de um processador por outro processador reduz pela metade os ciclos do barramento utilizados para comunicação entre processadores e isso compensa parcialmente a maior produção de ciclos devido à maior capacidade de elaboração do 386 em relação ao 286.

Pelas análises anteriores revelou-se também a necessidade de se desenvolver uma via de comunicação mais veloz de forma a conectar-se pelo menos 32 módulos.

Uma possível estrutura de comunicação a ser estudada é aquela do hipercubo, mas outras soluções mais eficientes e menos custosas deverão também ser procuradas e analisadas.