

DESEMPENHO EM ARQUITETURAS MULTIPROCESSADORAS

(*) **Mário Maurício Fiallos Aguilar**

Este trabalho apresenta uma forma determinística de avaliar o desempenho de um sistema multiprocessador com organização de memória multiportas — Esta avaliação inclui cálculos de vazão e eficiência baseados no ciclo de memória e no tempo mínimo entre dois pedidos de acesso consecutivos à memória.

1. INTRODUÇÃO

Os progressos alcançados na área de circuitos integrados tem facilitado a evolução de novas opções no processamen-

to de informação. Dentro destas opções as arquiteturas "não Von Newmann" ou alternativas ocupam uma posição de destaque devido ao alto desempenho e custos razoáveis.

(*) Ingeniero electricista (Universidad Nacional A. de Honduras 1984) MSC. Engenharia Elétrica - Sistemas de Computação (PUC/RJ 1985); Professor Assistente do

DEE/Universidade Federal do Ceará; Áreas de Interesse: Arquiteturas de Processadores e Redes de Computadores.

As redes de computadores são sem dúvida alguma, uma opção atraente devido a sua excelente aplicação em várias áreas entre as quais estão: a automação de escritórios, [SOAR 86], o compartilhamento de recursos computacionais de custo elevado (impressoras de alta qualidade, dispositivos de alta capacidade de armazenamento, etc), [SCHA 86], o correio eletrônico, e de uma maneira geral, na comunicação entre computadores e periféricos de capacidades e desempenhos diversos.

Uma rede formada por computadores hospedeiros de alto desempenho é uma alternativa que utiliza duas das maiores linhas de pesquisa das arquiteturas "não Von Neumann": Os sistemas multiprocessadores (ou sistemas fortemente acoplados) e naturalmente, as redes de computadores (ou sistemas fracamente acoplados).

Os sistemas multiprocessadores [ENSL 77] são geralmente utilizados em aplicações que requerem o processamento de grande quantidade de informação em um espaço de tempo relativamente curto. Exemplos das áreas de aplicação são: o processamento paralelo de jobs, a computação gráfica, a realização de previsões meteorológicas e o comportamento dinâmico de fluidos. Maiores informações sobre a aplicação de multiprocessamento nestas áreas podem ser encontradas na referência [FIAL 85].

Por todo o anterior é compreensível que o desenvolvimento da indústria de computação brasileira provoque a médio e curto prazo, uma utilização maior da tecnologia de multiprocessamento. É compreensível também que existam estudos para calcular com o maior grau de aproximação o desempenho de sistemas multiprocessadores [FIAL 85], [BRAS 85], [BRAS 86], [BOWE 80].

No que se segue é descrita uma forma de avaliação do desempenho de um sistema multiprocessador com organização de memória multiportas simulada através de memórias rápidas [FIAL 85], baseada no ciclo de memória (t_c) e no tempo mínimo entre dois pedidos de acesso consecutivos à memória (t_2).

2. DESCRIÇÃO GERAL DA ARQUITETURA

Se o controle, a comutação e a lógica para resolver conflitos de acesso à memória (conflitos estes provocados por pedidos de acesso simultâneos de dois ou mais processadores) for concentrada na interface da memória, o sistema multiprocessador resultante possui uma organização de memória multiportas. A figura 1 mostra o sistema em questão.

O método usado para resolver conflitos no acesso à memória baseia-se na implementação de uma política de prioridades. Geralmente todas as portas da memória multiportas são elétrica e operacionalmente iguais. De fato, as portas são idênticas e geralmente não é feita diferença sobre o tipo de unidade funcional (processador ou canal) ligado a cada porta.

Entretanto, vários estudos mostram [ENSL 74], [FIAL 85] que considerando a complexidade e o desempenho de três tipos diferentes de arquiteturas (barra comum compartilhada no tempo, memória multiportas e barra cruzada), a organização que promete o menor desempenho, porém a menor complexidade é a organização de barra comum compartilhada no tempo sem memória local. Por sua vez as arquiteturas de barras cruzadas conseguem o máximo desempenho entre as três arquiteturas mencionadas, porém o seu grau de complexidade é maior.

Os sistemas multiprocessadores com organização de memória multiportas e os de barra compartilhada no tempo com memória local conseguem um desempenho e um grau de complexidade intermediários. É pelas razões enunciadas neste parágrafo que a arquitetura de memória multiportas representa uma opção atraente.

Com tudo isto, existem diversas limitações que tornam difícil conceber uma memória multiportas (MM) ideal; isto é: uma MM que permita "verdadeiras" transferências simultâneas entre os EPs e a mesma. Para contornar estes problemas decidiu-se adotar uma organização que simula MM. A filosofia adotada [FIAL 85] consiste na multiplexação no tempo do atendimento aos pedidos de acesso à memória utilizando memórias rápidas. Assim, a memória multiportas do sistema cujo desempenho será estudado, suporta transferências concorrentes e não simultâneas com os EPs. Isto significa que na realidade só um EP de cada vez, pode completar sua transferência com a memória multiportas. Repare que esta multiplexação pode ser realizada de duas formas diferentes: multiplexando os pedidos de acesso, de modo que cada EP permaneça acessando a MM durante todo seu ciclo de barra, ou permitindo que cada EP permaneça utilizando a MM apenas durante o ciclo de memória. Dependendo do ponto de vista ambas as formas, em particular a primeira, podem ser consideradas arquiteturas de barra compartilhada. A segunda forma de multiplexação permite maior rapidez, apesar de exigir lógica de controle não encontrada na primeira, e em consequência seu custo ser maior. Repare também que se as transferências entre os EPs e a MM são realizadas "rapidamente" pode obter-se a impressão de simultaneidade ou paralelismo nas mesmas.

Para a realização da arquitetura foram realizados vários estudos, entre os quais o da escolha da família de processadores que seriam utilizados na implementação dos elementos processadores (EPs). Este estudo compreendeu entre outros aspectos, a análise de microprocessadores de 16 bits e de características relacionadas com gerenciamento de memória e facilidades de multiprocessamento e memória virtual.

A família motorola M68000 através de sua unidade de gerenciamento de memória (MMU) MC68451 e da CPU MC68010 fornece um tratamento adequado aos aspectos anteriormente mencionados e por este motivo foi escolhida. A figura 2 mostra cada um dos EPs que formam a arquitetura.

Continuando com a descrição da arquitetura chegamos à vez da MM. A parte mais complexa da MM são as portas. Estas são formadas por 3 blocos lógicos básicos: um detector de instruções TAS [MOTO 83], Buffers/latches para

isolar ou comunicar as barras de endereço, dados e controle do EP com a memória física, e o controle interno. Estes blocos solicitam ao árbitro (sinais RQi na figura 1) o acesso à memória física, enquanto que os sinais ACKi outorgam o acesso à memória física. A memória física é formada por memórias estáticas rápidas (tempo de acesso entre 35 e 45 μ s).

Uma vez descrita a arquitetura já é possível apresentar os conceitos usados para a avaliação do desempenho do sistema multiprocessador.

3. AVALIAÇÃO DO DESEMPENHO

Existem dois parâmetros necessários para o cálculo ou avaliação do desempenho de um sistema multiprocessador: a vazão "throughput" e a eficiência.

A vazão de um sistema de computação, é a razão entre a quantidade de processamento realizada ("carga") e o tempo requerido para seu processamento. A eficiência por sua vez pode ser avaliada em função de sua vazão e do número de EPs que forma o sistema.

Idealmente, a vazão deveria aumentar proporcionalmente com o número de EPs adicionados ao sistema. Na prática, devido a diversos fatores, a relação entre a vazão e o número de EPs adicionados não é proporcional. Repare também que a eficiência ideal de um sistema deveria ser 100%, o que não ocorre sempre.

A forma de calcular a vazão e a eficiência da arquitetura baseia-se em dois modelos de acesso à memória que, utilizando parâmetros como tempo mínimo entre dois pedidos de acesso consecutivos à memória (t_2) e ciclo de memória (tc), consegue obter uma aproximação razoável com um grau de complexidade mínimo.

3.1 VAZÃO E EFICIÊNCIA EM SISTEMAS MULTIPROCESSADORES

Repetindo novamente: a vazão absoluta de um sistema de computação é a razão entre a quantidade de processamento realizada e o tempo requerido para sua realização.

Assim, se VA é a vazão absoluta:
 $VA = Q/TP = \text{QUANTIDADE DE PROCESSAMENTO REALIZADA/TEMPO DE PROCESSAMENTO.}$

A vazão relativa de um sistema (VR), comparada com um outro sistema referência é a razão entre sua vazão e a vazão do sistema referência. Assim, VR é razão entre a vazão absoluta do sistema sendo comparado (VAc) e a vazão absoluta do sistema referência ou base (VAb).

Quando o tempo de processamento é o mesmo para ambos os sistemas:

$$VR = VAc/VAb = (Qc/TP)/(Qb/TP) = Qc/Qb$$

Onde, Qc = quantidade de processamento realizada no sistema sendo comparado, e Qb = quantidade de processamento realizada no sistema base ou referência.

Quando a quantidade de processamento realizada é a

mesma para ambos sistemas:

$$VA = VAc/VAb = (Q/TPc)/(Q/TPb) = TPb/TPc$$

Onde, TPb = tempo de processamento no sistema base e TPc = tempo de processamento no sistema sendo comparado.

O cálculo da eficiência foi avaliado em função da sua vazão relativa a um sistema monoprocessador e do número de EPs que o formam. Neste caso a eficiência do sistema multiprocessador pode ser medida por:

$$EFI (\%) = (VR/N) \times 100$$

Onde, N = número de EPs do sistema.

Em outros termos, o sistema multiprocessador com N EPs terá uma eficiência de 100% em relação ao número de EPs, se sua vazão relativa for igual à vazão de N EPs.

3.2 DESEMPENHO DA ARQUITETURA

A avaliação ou cálculo da vazão relativa e da eficiência da arquitetura é realizada tomando como hipótese que os EPs solicitem o acesso à memória com a taxa máxima $1/t_2$. Entretanto, serão considerados dois diferentes casos: o primeiro que inclui além da hipótese anterior um número de EPs menor ou igual do que a razão entre t_2 e tc e o segundo incluindo também a mesma hipótese, porém com um número de EPs maior do que a razão entre t_2 e tc.

A figura 3-a serve para ilustrar o primeiro caso, isto é: um sistema com N EPs ($EP_0 \dots EP_{N-1}$) que solicitam o acesso à memória com a taxa constante $1/t_2$ entre dois pedidos de acesso consecutivos, sendo que o tempo entre dois pedidos de acesso consecutivos (t_2) é maior ou igual do que o tempo necessário para realizar N ciclos de memória: $t_2 \geq Ntc$. Já a figura 3-b mostra o segundo caso: o tempo entre dois pedidos de acesso consecutivos é menor do que Ntc : $t_2 < Ntc$.

Nas figuras 3-a e 3-b, a↑ representa o início de um acesso do EP à MM, e P↑ representa o pedido de acesso do EP à MM. É importante não confundir a↑ que representa o instante no qual o EP começa acessar a memória com P↑ que representa unicamente o instante no qual o EP solicita (e não consegue) realizar o acesso.

Por simplicidade na figura 3-a, o eixo tempo é dividido em intervalos contíguos de valor Ntc (na figura 3-a são mostrados três destes intervalos). Cada um destes intervalos é por sua vez dividido em outros N subintervalos de tempo (cada subintervalo com um valor de 1tc) que representam N acessos à memória, um acesso realizado por cada um dos EPs. Assim por exemplo, no primeiro intervalo Ntc à esquerda, o primeiro subintervalo representa o acesso do EP₀, o segundo subintervalo representa o acesso do EP₁, etc. . .

A partir do intervalo de sincronização, nenhum dos EPs sofre retardos no acesso à memória e permanecem todos ocupados ora acessando a memória ora realizando processamento interno. Pode-se concluir por tanto que a quantidade de processamento produzida pelos N EPs é N vezes superior à quantidade de processamento produzida por um único EP. Deste modo a vazão relativa VR = N e a eficiência EFI (%) = $\frac{VR}{N} \times 100\% = 100\%$.

N

Na figura 3-b a partir do intervalo de sincronização, o EP₀ acessa a memória seguido dos acessos à memória de ca-

da um dos restantes EPs. Como $N_{tc} > t_2$, isto é, o tempo entre dois pedidos de acesso consecutivos à memória de cada um dos EPs é menor do que o tempo necessário para realizar N ciclos de memória (N_{tc}), cada um dos EPs pode sofrer um retardo $\Delta = N_{tc} - t_2$. Repare que todos os EPs realizam um acesso à memória e pedem de novo o acesso em um intervalo de tempo N_{tc} . Ver por exemplo, no segundo intervalo N_{tc} da figura 3-b, que o EPo além de acessar a memória (no primeiro subintervalo) realiza um novo pedido de acesso à memória no mesmo intervalo de tempo N_{tc} . Isto é: em um intervalo de tempo N_{tc} , cada processador processa um tempo igual a t_2 e fica esperando um tempo igual a Δ . A vazão de cada processador em um intervalo de tempo N_{tc} é por tanto t_2/N_{tc} . Como existem N EPs a vazão total é N vezes o valor acima e portanto $VR = t_2/tc = a$. A eficiência do sistema é:

$$EFI (\%) = \frac{VR}{N} \times 100\% = \frac{a}{N} \times 100\%$$

Os gráficos da figura 4 servem para acelerar alguns pontos importantes do desempenho obtido a partir dos dois casos apresentados.

A figura 4-a mostra a VR do sistema multiprocessador em função da razão entre t_2 e tc . Pode-se apreciar que quando $N \geq a$, a vazão relativa do sistema cresce proporcionalmente com o valor de "a" e é independente de N . Quando $N < a$, a vazão relativa é constante (N) e independe do valor de "a".

Na figura 4-b é mostrada a vazão relativa em função do número de EPs (N) mantendo para cada um dos casos a razão "a" constante. Note por exemplo que quando $N \leq a$ a VR resultante é N . Quando o número de EPs é maior do que a razão "a", a vazão relativa máxima é igual a "a" independentemente do número de EPs adicionados ao sistema.

A eficiência em função do número de EPs é apresentada na figura 4-c. Pode-se confirmar facilmente que a eficiência de 100% ocorre quando o número de EPs é menor ou igual do que razão "a".

O passo seguinte na avaliação do desempenho da arquitetura é o cálculo da vazão e a eficiência utilizando parâmetros específicos da família M68000. Deverá se obter o número de EPs que segundo a maneira de avaliar o desempenho apresentada produz a máxima vazão. A eficiência também é avaliada. Primeiramente os cálculos são efetuados utilizando unicamente os aspectos dos EPs já modelados; em segundo lugar, são considerados além dos aspectos já modelados, aspectos mais realistas da CPU e da MMU que fornecem valores menos otimistas e mais próximos da realidade.

Para realizar a avaliação do desempenho segundo as hipóteses assumidas, é necessário encontrar o tempo mínimo entre dois pedidos de acesso consecutivos à memória em função de parâmetros dos EPs formados a partir da Motorola. É utilizado o sinal MAS ("mapped address strobe") [MOTO 83 1] como pedido de acesso à MM. Com isto último, $t_2 = 4/f$ onde, f = frequência de clock do 68010 e então: a vazão relativa máxima e a eficiência são dadas pelas expressões:

$$VR = t_2/tc = 4/ftc$$

$$EFI (\%) = \frac{VR}{N} \times 100\% = (4/ftcN) \times 100\%$$

A tabela 1 mostra valores de t_2/tc , o número mínimo de EPs para obter a máxima vazão relativa e a correspondente

eficiência para diferentes valores do "CLOCK" e para dois diferentes ciclos de memória (45 μ s e 35 μ s).

É claro (ver tabela 1) que para aumentar a vazão do sistema, geralmente os EPs devem funcionar com a frequência máxima de operação. Entretanto, quanto maior for esta frequência, maior deverá ser a frequência de operação do controle da memória multiportas a fim de evitar retardos nos acessos.

Os aspectos relacionados com o "timing" dos EPs que não foram considerados para a avaliação da vazão e da eficiência da arquitetura e que entretanto podem influenciar bastante as figuras obtidas acima são:

a) Características internas à MMU que obrigam o uso de "latches" externos à mesma para armazenar o endereço físico (PA). Seria lógico utilizar o sinal MAS para ativar os "latches". Entretanto a referência [MOTO 83 1] mostra que o sinal é gerado no máximo 10 μ s antes que o endereço físico exista. Esta é uma característica bastante surpreendente da MMU e que provoca outros problemas que serão vistos adiante.

Como o sinal MAS pode ser ativado até 10 μ s antes de que o endereço físico exista, este sinal precisa ser retardado no mínimo 10 μ s, a fim de fornecer o tempo de "Set-up" entre o endereço físico e MAS nos "latches" externos à MMU. Este tempo de "Set-up" é chamado tpo. Se os latches usados fossem os SN745373, tpo é 18 μ s e tdelay = 10 μ s

b) Devido à presença antecipada do sinal MAS, existe a necessidade de utilizar a barra interna da MM um tempo maior que o tempo de ciclo da memória física (MF). Este tempo adicional deve-se, no caso da leitura, a que o acesso somente poderá ser feito após o aparecimento do endereço físico (PA), [MOTO 83 1] e no caso da escrita, após que os sinais de validação de dados (PUDS e/ou PLDS) indicarem o byte ou palavra na qual a mesma será realizada.

c) A existência do tempo de "Set-up" entre o sinal "DTACK" [MOTO 83 1] e o ciclo de relógio S_4 . Se este tempo não for satisfeito, os processadores inserem ciclos de "Wait" que não estavam previamente modelados.

Entretanto, os efeitos apresentados acima podem ser atenuados (de modo que o desempenho do sistema não diminua e também de forma que o comportamento do sistema seja aproximadamente o da figura 3), se algumas modificações forem introduzidas:

a) Para permitir que a utilização da barra interna da MM ocorra unicamente durante ciclo de leitura, o sinal PA (endereço físico) será interpretado como um pedido de acesso, enquanto que nos ciclos de escrita, o pedido será realizado unicamente quando o sinal PUDS/PLDS (validação de dados) [MOTO 83 1] se encontrar ativo.

b) No controle da porta, o sinal DTACK não precisa ser ativado após o acesso à MF, e sim tão logo os sinais PA (em leitura) e PUDS/PLDS (em escrita) se encontrarem ativos.

Para encontrar a frequência máxima de operação dos EPs que evite a existência de ciclos de "Wait" e que possibilite o cálculo da vazão relativa máxima, foram analisados os ciclos de leitura e escrita e encontrou-se que a frequência máxima sem ciclos de "Wait" era determinada [FIAL 85] pelo "timing" do ciclo de escrita (na realidade o ciclo de escrita no referente ao comportamento do sistema representava o pior caso). Assim da figura 5:

$$20\mu s + 20\mu s < T/2 - 20\mu s$$

$$20\mu s < T$$

T = Período

Conseqüentemente, a freqüência máxima de operação é 5 Mhz.

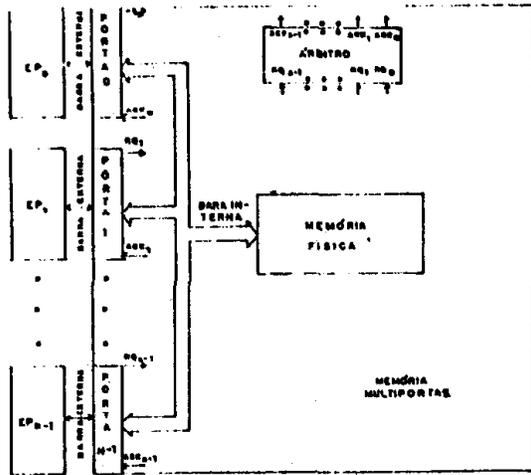


FIGURA 1

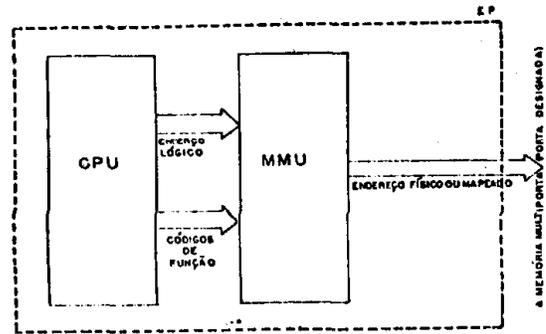


FIG. 2

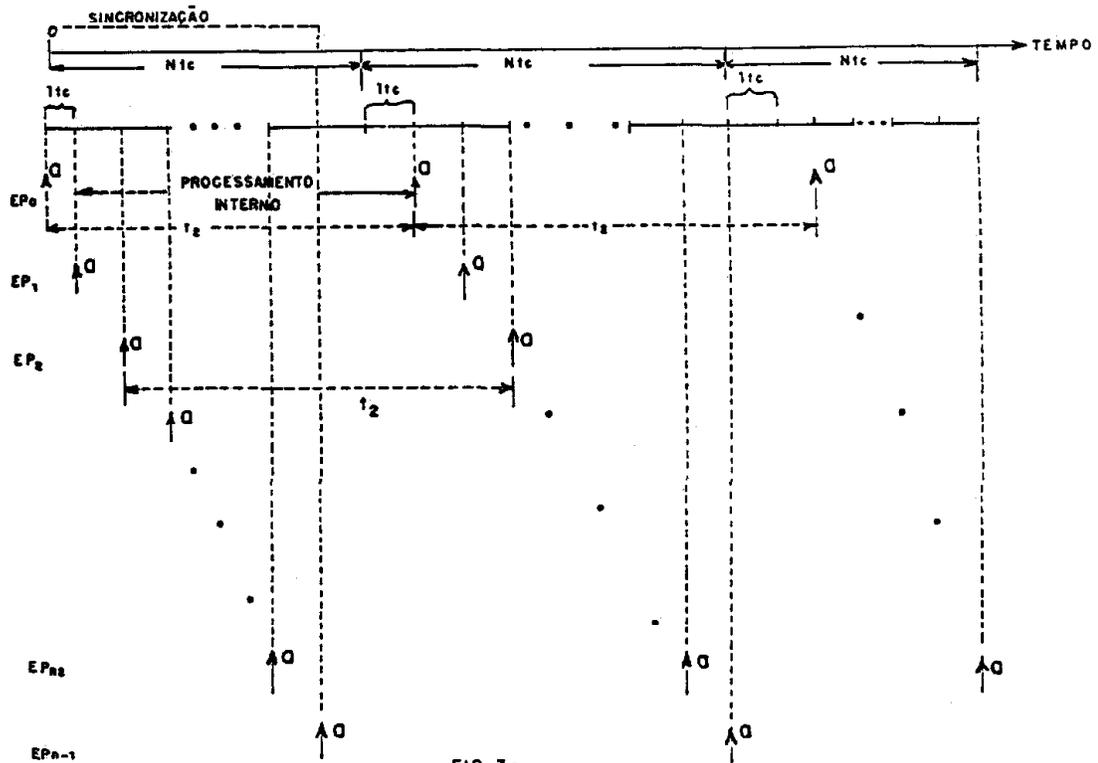


FIG. 3a

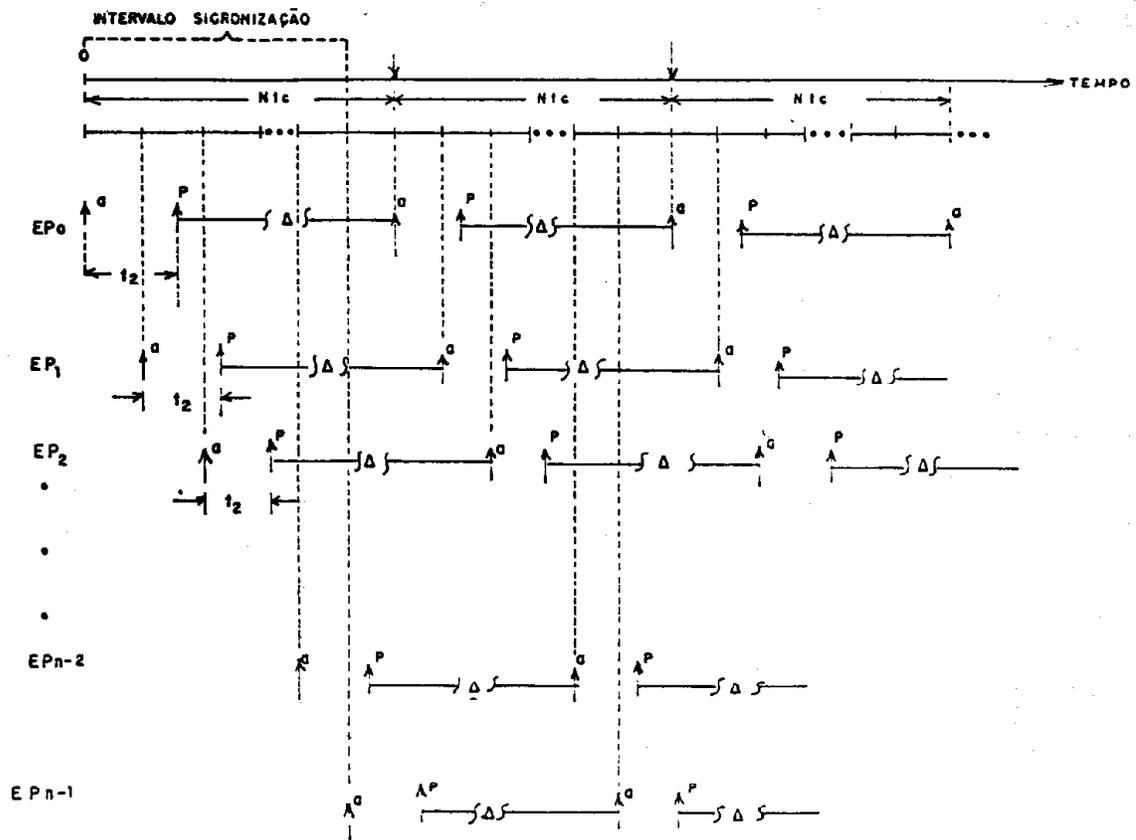
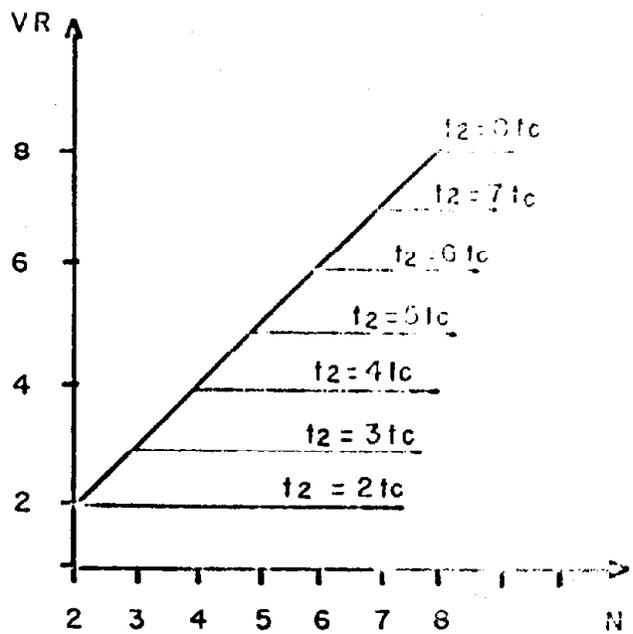
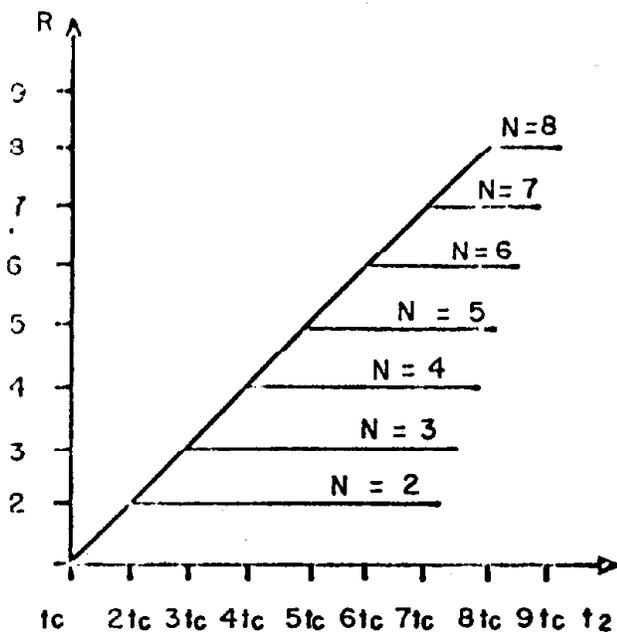


FIG. 3 b

(a)

(b)



(G)

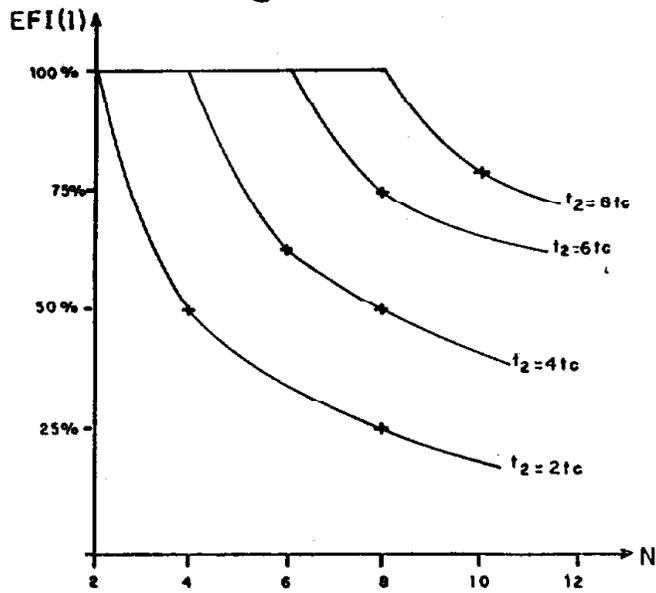


Fig. 4

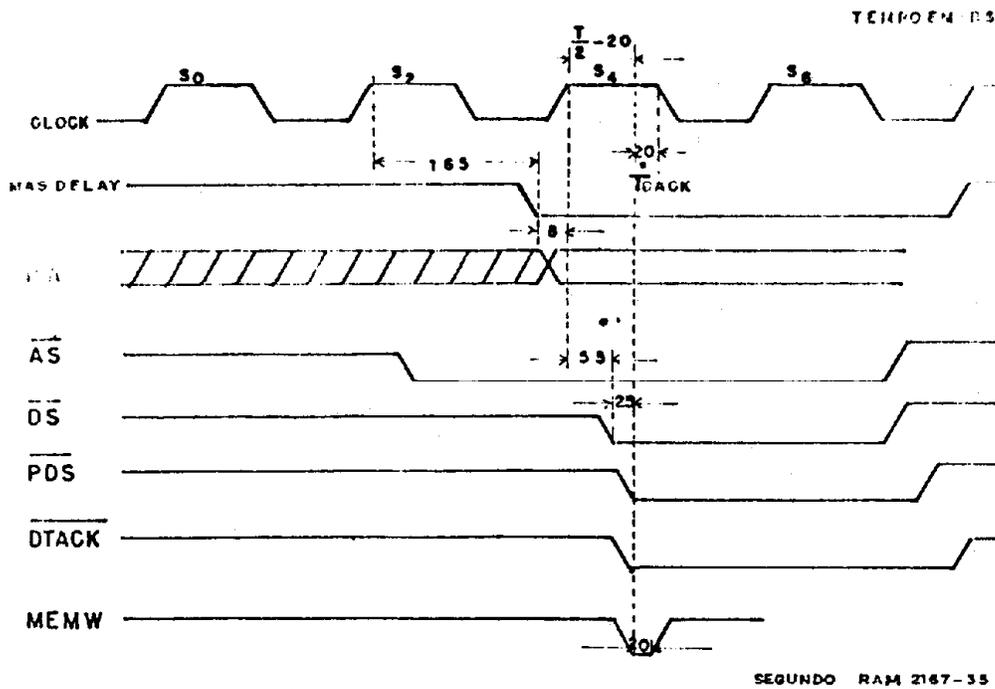


FIG. 5

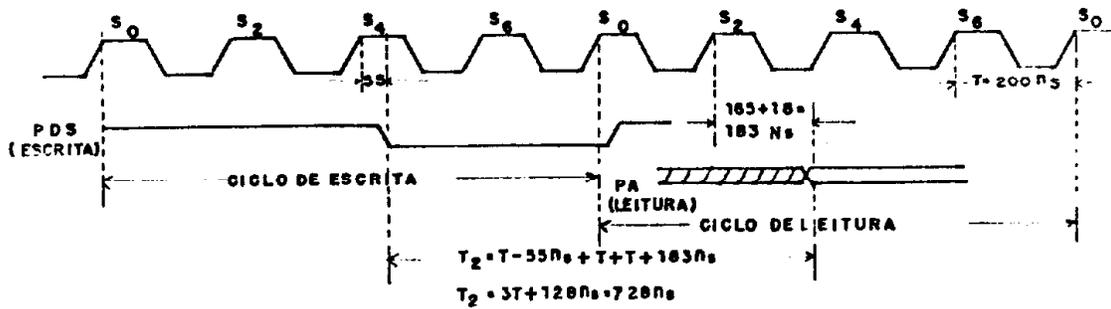


FIG. 6

Para encontrar a vazão relativa máxima é necessário obter o valor de t_2 em termos da frequência. Quando o processamento inclui unicamente ciclos de leitura ou unicamente ciclos de escrita, o valor de t_2 é $4T$: este é o intervalo de tempo mínimo entre a existência de dois endereços físicos (no caso de leitura) ou de dois sinais PUDS/PLDS ativos (escrita). Entretanto, o valor de t_2 é menor quando um ciclo de escrita é seguido de um ciclo de leitura, como é mostrado na figura 6, neste caso o valor de t_2 é $3T + 128 \mu s$. Como $T = 200 \mu s$, então: $t_2 = 728 \mu s$.

Em conseqüência, para memórias com tempo de ciclo igual a $35 \mu s$ e $45 \mu s$ a vazão relativa máxima e o número mínimo de EPs são:

$$\begin{array}{llll} tc = 45 \mu s & VR = 728/45 & e & N = 17 \\ tc = 35 \mu s & VR = 728/35 & e & N = 21 \end{array}$$

4. CONCLUSÕES

As duas formas de avaliação do desempenho apresentadas, oferecem diferentes alternativas para o cálculo da vazão e a eficiência do sistema. A segunda forma por levar em consideração aspectos mais realistas do que a primeira, oferece uma melhor aproximação. Repare por exemplo, que segundo a primeira avaliação, quando $tc = 35 \mu s$ e $f = 5 \text{ Mhz}$, a $VR = 22,86$, enquanto que na segunda, este valor é 20,8. A diferença de valores deve-se a que na segunda modelagem, t_2 foi avaliado em função do tempo mínimo que existe entre dois pedidos de acesso quando o ciclo de escrita precede ao de leitura.

Devido às restrições encontradas no árbitro da MM [FIAL 85] que tornam proibitivas implementações com mais do que 8 EPs, os valores acima encontrados revelam a existência de uma folga bastante confortável que permite uma grande flexibilidade na construção do árbitro, na escolha da memória física ou ainda compensar a existência de parâmetros específicos do controle do sistema desprezados na avaliação realizada.

E finalmente, acredita-se que a avaliação realizada permite obter um primeiro cálculo do desempenho do sistema com um grau de aproximação razoável e uma complexidade muito pequena.

$$tc = 45 \mu s$$

FREQUÊN- CIA(Mhz)	t_2/tc	NÚMERO DE EPs (N)	EFI(%)
2	44,4	45	97,86
4	22,2	23	95,56
5	17,7	18	98,72
6	14,81	15	98,73
8	11,11	12	92,58
10	8,89	9	98,78
12,5*	7,11	8	88,88
16,67**	5,53	6	88,83

$$tc = 35 \mu s$$

FREQUÊN- CIA(Mhz)	t_2/tc	NÚMERO DE EPs	EFI(%)
2	57,14	58	98,52
4	28,57	29	98,52
5	22,86	23	99,39
6	19,05	20	95,25
8	14,29	15	95,27
10	11,43	12	95,25
12,5*	9,14	10	91,40
16,67**	6,86	7	98,00

TABELA 1

*Não se tem notícia de MMU operando a esta frequência.

**Não se tem notícia de MMU e CPU (68010) operando a esta frequência.

BIBLIOGRAFIA

- [BOWE 80] The Logical Desing of Multi-microprocessor Systems. B.A. Bowen and R. J.A. Buhr. — Prentice Hall Inc. — 1980.
- [BRAS 85] Técnicas de Multiplexação Por Divisão no Tempo para Interconexão Processador — Memória em Sistemas de Multiprocessadores: Um estudo comparativo. — M.A.G. Brasileiro e J. A. Moura. — Anais V Congresso da SBC, pp. 60 — 71. — 1985.
- [BRAS 86] Interconexão Processador — Memória Via Técnica de Multiplexação Por Divisão de Tempo. — M.A.G. Brasileiro, J.A. Moura. — Anais VI Congresso da SBC, pp 151 — 163, 1986.
- [ENSL 74] Multiprocessors and Parallel Processing. — P. H. Enslow, Jr. — John Wiley e Sons, Inc, 1974.
- [ENSL 77] Multiprocessor Organization A. Survey. — P. H. Enslow. Computer Surveys, Vol. 9 no. 1, pp 103 — 129. — 1977.
- [FIAL 85] Um Estudo e Uma Proposta de um Sistema Multiprocessador Com Organização de Memória Multiportas Simulada através de Memórias Rápidas. — Mario Mauricio Fiallos Aguilar, tese de mestrado DEE PUC/RJ. — 1985.
- [MOTO 83] MC 68010 16 bit Microprocessor Manual. — Motorola Inc, Agosto 1983.
- [MOTO 83] MC 68451 Memory Management Unit. — Motorola Inc, 1983.
- [SCHA 86] O Servidor de Impressão da REDPUC. — D. Schwabe e outros. — 4 SBRC, pp 263 — 280, 1986.
- [SOAR 86] Rede Experimental da PUC/RJ Para Interligação de Voz e Dados. — L.F.G. Soares e outros, 4 SBRC, 1986.