
RELATÓRIO TÉCNICO

UMA AVALIAÇÃO DO IMPACTO DAS OPERAÇÕES DE E/S NO DESEMPENHO DO MULTIPROCESSADOR MULTIPLUS

Sidney de Castro Oliveira
Júlio Salek Aude
NCE/UFRJ

NCE — 12/92
dezembro



Núcleo de Computação Eletrônica
Universidade Federal do Rio de Janeiro

Tel.: 598-3212 - Fax.: (021) 270-8554
Caixa Postal 2324 - CEP 20001-970
Rio de Janeiro - RJ

UMA AVALIAÇÃO DO IMPACTO DAS OPERAÇÕES DE E/S NO DESEMPENHO DO MULTIPROCESSADOR MULTIPLUS

RESUMO

Este trabalho apresenta um estudo sobre o impacto das operações de E/S no desempenho dos elementos processadores do Multiplus, um multiprocessador com memória global compartilhada em desenvolvimento no NCE/UFRJ. As questões são avaliadas a partir de dados oriundos de simulações que procuraram retratar a operacionalidade do subsistema de E/S dentro do contexto do Multiplus

THE IMPACT OF I/O OPERATIONS ON THE PERFORMANCE OF THE MULTIPLUS MULTIPROCESSOR

ABSTRACT

This paper discusses the impact of I/O operations on the performance of the processing nodes within Multiplus, a shared-memory multiprocessor under development at NCE/UFRJ. The different aspects of such impact are evaluated through simulation experiments which try to model the I/O system operation within the Multiplus architecture.

UMA AVALIAÇÃO DO IMPACTO DAS OPERAÇÕES DE E/S NO DESEMPENHO DO MULTIPROCESSADOR MULTIPLUS

Sidney de Castro Oliveira
NCE/UFRJ

Júlio Salek Aude
NCE/UFRJ

RESUMO

Este trabalho apresenta um estudo sobre o impacto das operações de E/S no desempenho dos elementos processadores do Multiplus, um multiprocessador com memória global compartilhada em desenvolvimento no NCE/UFRJ. As questões são avaliadas a partir de dados oriundos de simulações que procuraram retratar a operacionalidade do subsistema de E/S dentro do contexto do Multiplus

ABSTRACT

This paper discusses the impact of I/O operations on the performance of the processing nodes within Multiplus, a shared-memory multiprocessor under development at NCE/UFRJ. The different aspects of such impact are evaluated through simulation experiments which try to model the I/O system operation within the Multiplus architecture.

Afiliação: Núcleo de Computação Eletrônica/UFRJ
Caixa Postal 2324
20001 - Rio de Janeiro, RJ
Tel: 021 - 598 3212 R. 3171/3160
e-mail: sidney@vidigal.nce.ufrj.br

S. C. Oliveira - Analista de Sistemas do NCE/UFRJ. M.Sc. em Engenharia de Sistemas e Computação pela COPPE/UFRJ (1992). Áreas de interesse: Arquitetura de Computadores, Processamento Paralelo e Estações Gráficas de Trabalho.

J. S. Aude - Analista de Sistemas do NCE/UFRJ. Professor Adjunto do Instituto de Matemática da UFRJ, Ph.D. em Ciência da Computação pela University of Manchester (1986). Áreas de interesse: CAD para VLSI, Arquitetura de Computadores e Processamento Paralelo.

1 - INTRODUÇÃO

A tecnologia de microeletrônica tem permitido o surgimento de microprocessadores cada vez mais velozes e poderosos. A utilização destes dispositivos em arquiteturas avançadas resulta em computadores de alto desempenho. Diversas áreas de aplicação desta classe de computadores demandam grandes quantidades de dados, como por exemplo, o processamento científico em geral: meteorologia, análise estrutural, interpretação de imagens, etc. Estes dados são mantidos em dispositivos de armazenamento e, periodicamente, intercambiados com os elementos processadores, conforme sua necessidade. O equilíbrio deste intercâmbio de dados é fundamental para o alto desempenho destes computadores.

O objetivo deste trabalho é avaliar o impacto das operações de E/S no desempenho dos elementos processadores do Multiplus, constituindo ferramenta auxiliar na definição de um subsistema de E/S que atenda satisfatoriamente a demanda de dados do Multiplus. A proposta é obter, através de simulações, informações que retratem a operacionalidade do Multiplus, permitindo a avaliação de questões inerentes ao seu *cluster* de elementos processadores. O procedimento de análise foi comparativo entre o comportamento do Multiplus com e sem as operações de E/S. Dentre as questões que se procurou avaliar está a duração média dos períodos dos acessos externos executados pelas unidades inteiras dos elementos processadores do Multiplus e a banda passante efetiva do barramento do *cluster* para as operações de E/S.

2 - O MULTIPLUS E SEU SUBSISTEMA DE E/S

O Multiplus é um multiprocessador científico de alto desempenho com arquitetura modular e memória global compartilhada, em desenvolvimento no NCE/UFRJ. A arquitetura é capaz de suportar até 2048 elementos processadores ou nós de processamento. Eles são baseados em microprocessadores RISC de 32 *bits* com arquitetura SPARC e capacidade de processamento de 25 MIPS VAX a 40 MHz. Além do microprocessador, cada nó de processamento possui um co-processador de ponto flutuante, 32 MBytes de memória pertencentes ao espaço de endereçamento global de 32 GBytes, *cache* de instrução e *cache* de dado separados, com 64 KBytes cada um, e *hardware* de suporte à gerência de memória.

Até 16 elementos processadores podem ser interconectados através de dois barramentos de 64 *bits* de largura cada um, formando um *cluster* de processadores. Cada *cluster* possui uma placa multifunção responsável pela arbitragem dos barramentos e geração dos sinais de *reset*. Os diferentes *clusters* de processadores se comunicam através de uma rede de interconexão multiestágio do tipo n-cubo invertido. Interfaces inteligentes, com capacidade de armazenamento de mensagens de até 128 *bytes* e de realização de operações de DMA nos barramentos, acoplam a rede de interconexão aos *clusters*.

A figura 1 mostra o esquema geral da arquitetura do Multiplus. Com esta arquitetura pode-se criar uma família de computadores cobrindo um espectro que vai desde estações de trabalho de alto desempenho, com configurações de 1 a 8 nós de processamento, passando por mini-supercomputadores, de 16 a 128 nós, e chegando aos supercomputadores com 256 ou mais nós de processamento. Maiores informações sobre o Multiplus podem ser obtidas em [AUDE91].

A definição de um subsistema de E/S para o Multiplus teve o compromisso de atender satisfatoriamente às diversas configurações que o sistema pode assumir. Primeiramente cabe distinguir dois tipos de processamento associado a um subsistema de E/S. Um processamento orientado a caracteres, controlando operações de E/S com terminais, impressoras e redes como ETHERNET, e outro orientado a blocos, controlando operações de E/S com discos e fitas magnéticas. Para cada um dos dois tipos de processamento existe um processador de E/S (PES) específico encarregado de controlar as operações de E/S correspondentes. Este par de processadores é independente em suas funções e, associado a cada *cluster* de elementos processadores, forma o subsistema de E/S do Multiplus. Neste trabalho não serão abordados os problemas relacionados à E/S de caracteres, até mesmo porque eles não representam risco de se configurarem como um gargalo para o sistema computacional.

A parte orientada a bloco do subsistema de E/S processa operações cujas características estão fortemente relacionadas à aplicação submetida ao sistema como um todo. O Multiplus é voltado para aplicações científicas. Assim sendo, espera-se que suas operações de E/S sejam esparsas e extensas. Isto caracteriza a necessidade de um canal de comunicação rápido e eficiente entre o subsistema de E/S e os elementos processadores. Visualizando a arquitetura do Multiplus, pode-se notar que ela sugere uma distribuição do processamento de E/S. Como são

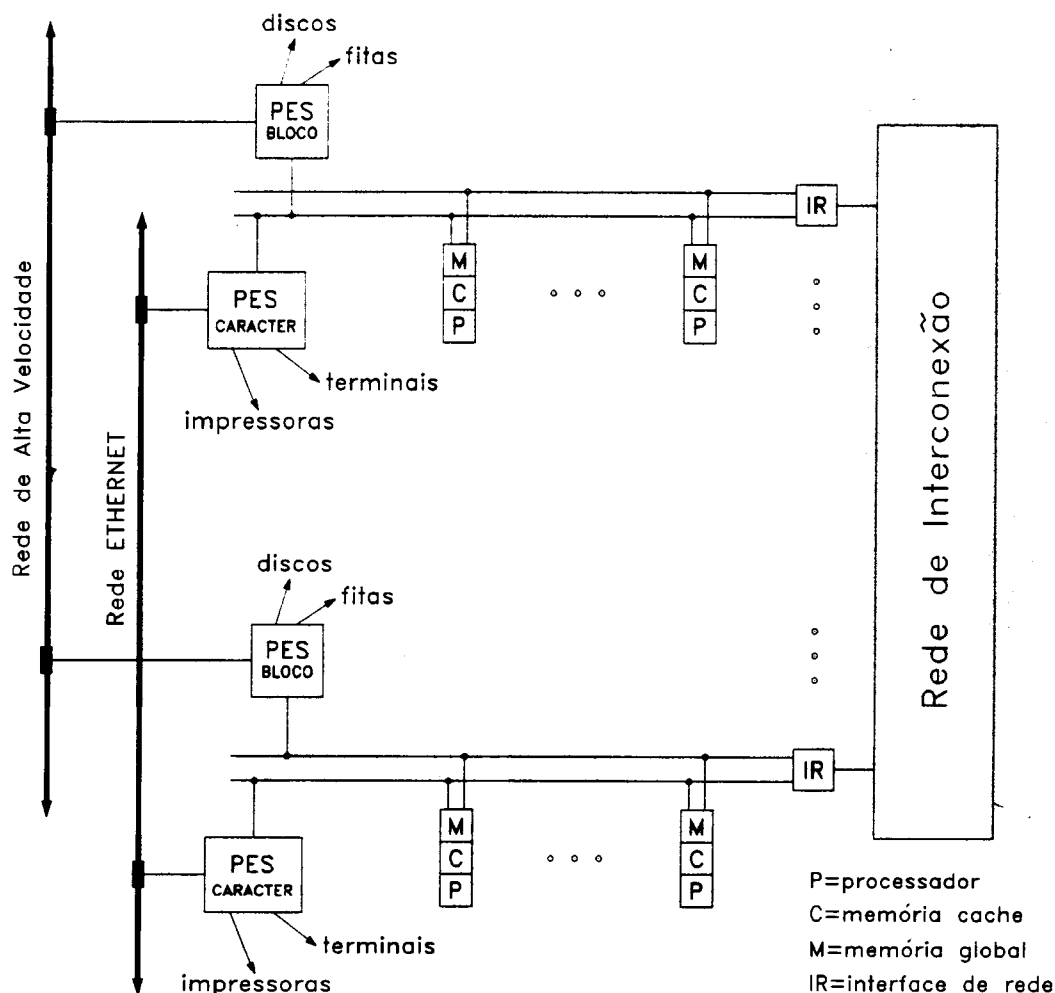


FIGURA 1: Esquema geral da arquitetura do Multiplus.

vários barramentos interligados por uma chave, é sugestivo associar um processador de E/S a cada barramento. Desta forma, o *cluster* de elementos processadores formado por cada barramento possui, através do próprio barramento, um canal de comunicação de alta velocidade entre os elementos processadores e o subsistema de E/S. Além do mais, a própria finalidade a que o Multiplus se destina reforça esta configuração. Uma máquina paralela sugere a realização de várias operações em paralelo. E dentro destas operações também se encontram as de E/S de dados e instruções para se ter um maior equilíbrio e eficiência no sistema. Este paralelismo é tanto mais explorado quanto mais distribuído for o processamento de E/S.

Para evitar uma eventual saturação da comunicação pela chave, em detrimento de seu desempenho, buscou-se um caminho alternativo para as transferências de E/S entre os *clusters*, incorporando-se uma rede de comunicação de alta velocidade ao subsistema de E/S. Para manter a homogeneidade da arquitetura, esta comunicação é restrita às transferências de entrada e saída. Uma característica interessante é que, com esta rede de comunicação, cada elemento processador solicita sempre serviços de E/S ao seu PES associado. Caso a informação solicitada esteja nos dispositivos de armazenamento controlados por este próprio PES, a transferência é realizada normalmente pelo barramento. Caso contrário, o PES solicita a informação ao PES apropriado, recebendo-a através da rede de comunicação e transferindo-a, posteriormente, ao elemento processador via barramento do *cluster*. Entretanto, não é essencial que todo *cluster* possua um PES associado, podendo existir *clusters* só de elementos processadores. Neste caso, qualquer transferência de E/S é realizada obrigatoriamente pela rede de chaves.

Desta forma, o Multiplus se configura com um subsistema de E/S com duas forças de acoplamento aos elementos processadores. Cada *cluster* pode ter um PES fortemente acoplado aos seus elementos processadores, permitindo transferências com um custo de tempo bastante reduzido. Caso as informações de E/S não se encontrem no *cluster* do elemento processador solicitante, o acoplamento se enfraquece, aumentando o custo de tempo das transferências. Consegue-se, assim, explorar, a nível de *cluster*, as vantagens de uma configuração concentrada, embora o subsistema de E/S como um todo seja distribuído.

3 - SIMULAÇÃO

Com objetivo de avaliar o impacto das operações de E/S no desempenho dos elementos processadores do Multiplus, procurou-se gerar dados, através de simulações, que pudessem retratar a operacionalidade do subsistema de E/S dentro do contexto do Multiplus, monitorando o comportamento dos elementos processadores e dos barramentos do *cluster* em função das operações de E/S. As principais questões avaliadas foram:

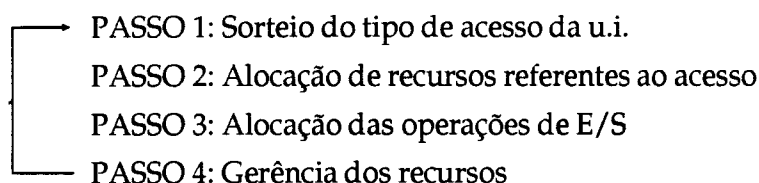
- A escolha do barramento onde serão efetuadas as operações de E/S: barramento de código ou de dado.
- A degradação do desempenho dos elementos processadores medido pela duração média dos acessos executados.
- A taxa de transferência média efetivamente obtida nas operações de E/S.

3.1 - O Simulador

O simulador utilizado foi uma evolução de um simulador desenvolvido no NCE/UFRJ por [MESLIN91] com objetivo de estudar o desempenho dos elementos processadores do Multiplus frente às diversas opções de política de *cache* disponíveis. Para permitir a interpretação dos resultados oriundos da simulação, será descrito, a seguir, o princípio básico de funcionamento do simulador.

Para avaliar as políticas de *cache*, o simulador retrata a operacionalidade dos elementos processadores do Multiplus. Cada elemento processador possui uma unidade inteira (u.i.), que demanda acessos continuamente, memórias *cache* de dado e código separadas e uma memória local. Como a memória do Multiplus é global e compartilhada, é permitido a um elemento processador acessar a memória local de outro. Este acesso pode ser interno ao *cluster* ou via rede de interconexão, caso o elemento processador referenciado pertença a um *cluster* distinto. A partir de uma distribuição de probabilidade de ocorrência dos diversos tipos de acessos possíveis, o simulador avalia a influência da interdependência destes acessos no desempenho global do Multiplus.

O princípio básico de funcionamento do simulador pode ser representado por um *loop* de quatro passos:



O primeiro passo determina, através de um sorteio aleatório, o tipo de acesso das unidades inteiras (u.i.) em estado livre de cada um dos elementos processadores. Após o sorteio do acesso, as unidades inteiras passam para o estado espera. Os tipos de acessos possíveis às unidades inteiras, juntamente com suas probabilidades de ocorrência, são mostrados na figura 2.

O segundo passo faz a alocação dos recursos necessários à execução de cada um dos acessos determinados no passo anterior. Caso a alocação seja bem sucedida, ou seja, todos os recursos necessários à execução do acesso estejam disponíveis, eles são ocupados, e a unidade inteira correspondente é colocada no estado execução, caso contrário, a u.i. permanece no estado espera. Cada um dos acessos ocupa os recursos necessários por uma quantidade determinada de ciclos de relógio. Os diversos recursos disponíveis são: *cache* de dado, *cache* de código, memória local, barramento de dado, barramento de código e rede de interconexão.

O terceiro passo é o responsável pela alocação das operações de E/S e dos recursos envolvidos nestas operações. Pode-se notar, pela posição deste passo em relação aos demais, que a prioridade na alocação dos recursos necessários às operações de E/S é inferior à das unidades inteiras.

O quarto passo gerencia os recursos e faz a computação dos ciclos, isto é, incrementa o contador total de ciclos e controla os acessos em execução. Quando um acesso se finda, os recursos envolvidos são liberados e a unidade inteira correspondente é colocada no estado livre.

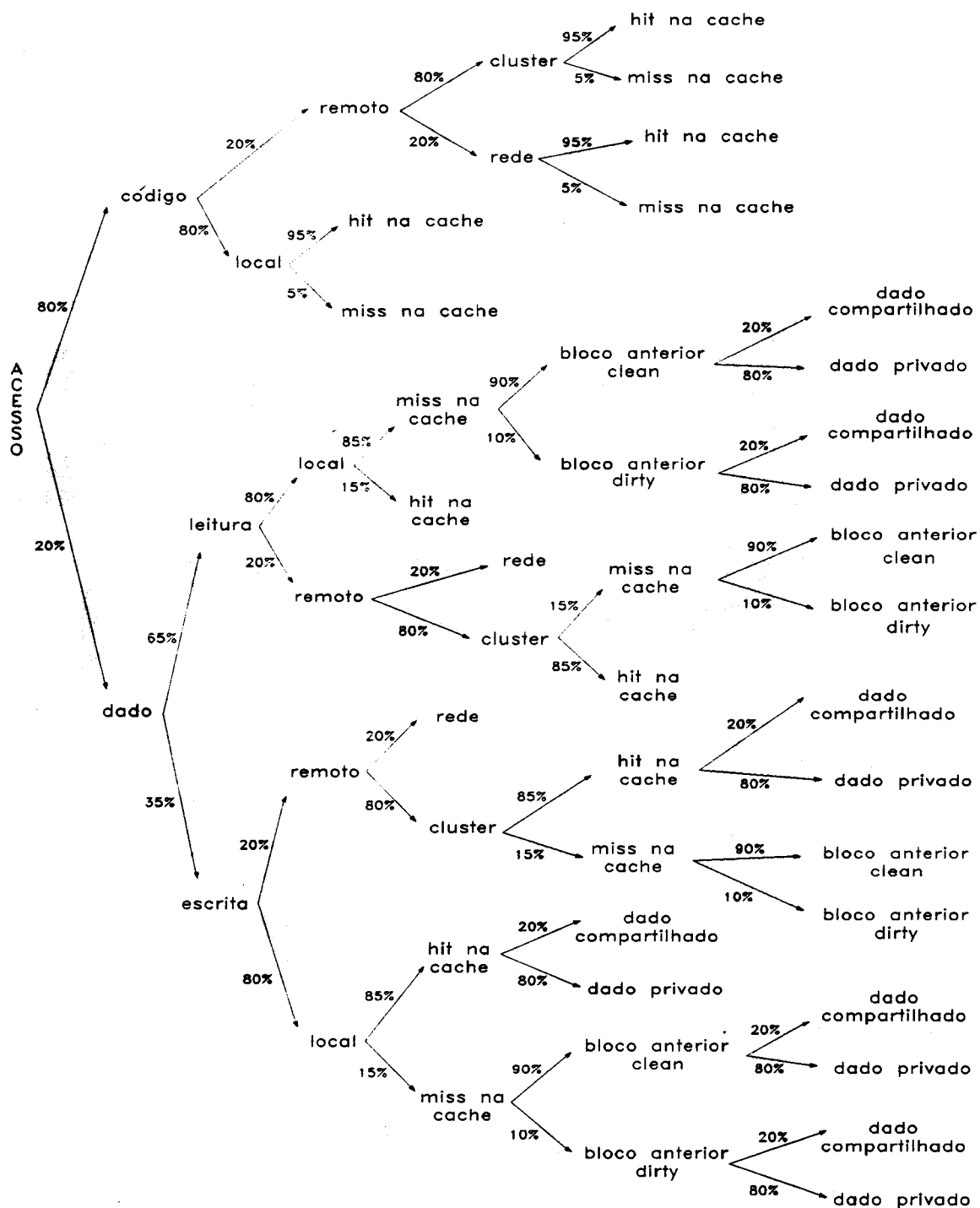


FIGURA 2: Tipos de acessos possíveis às unidades inteiras e suas probabilidades de ocorrência.

3.2 - Os Resultados

Antes de iniciar a análise dos resultados das simulações é necessário esclarecer alguns dos critérios adotados. Como se sabe, o perfil esperado para as operações de E/S do Multiplus é de transferências esparsas e extensas. Entretanto, estas transferências ocorrem, no barramento,

em conjunto com os acessos externos das unidades inteiras dos elementos processadores. Devido à disparidade na periodicidade destes dois eventos, a quantidade de ciclos de relógio simulados necessários para conciliá-los é muito grande, resultando num empecilho prático. A alternativa encontrada para contornar este problema foi simular o comportamento do sistema durante uma fração de tempo na qual as transferências de E/S estão ocorrendo e estender o resultado, analiticamente, para diversas cargas de E/S passíveis de serem impostas pelos elementos processadores do Multiplus.

As transferências de E/S são efetuadas em várias rajadas ininterruptas de 128 *bytes* cada. Toda rajada é precedida de uma nova arbitração. Desta forma, durante as transferências de E/S, os PES estão sistematicamente tentando alocar os recursos necessários à realização das rajadas. Como a prioridade dos PES na alocação dos recursos é menor que a das unidades inteiras, a quantidade de rajadas realizadas por ciclo de relógio é função da taxa de ocupação dos recursos compartilhados. Quanto mais livres estiverem os recursos, maior a banda passante efetivamente obtida nas transferências de E/S.

Por último, cabe esclarecer que a avaliação do impacto das operações de E/S no comportamento dos elementos processadores do Multiplus foi restrita às possibilidades reais de implementação. Desta forma, alguns parâmetros do simulador não foram variados, mantendo-se fixos nos valores correspondentes à definição adotada para implementação do Multiplus:

- Política de *cache*: WRITE THROUGH
- Barramento de código e barramento de dado separados
- Largura de cada barramento: 64 *bits*
- Velocidade de processamento das unidades inteiras: 25 MHz

Para analisar os resultados das simulações cabe, inicialmente, mostrar algumas características do comportamento dos elementos processadores do Multiplus sem o

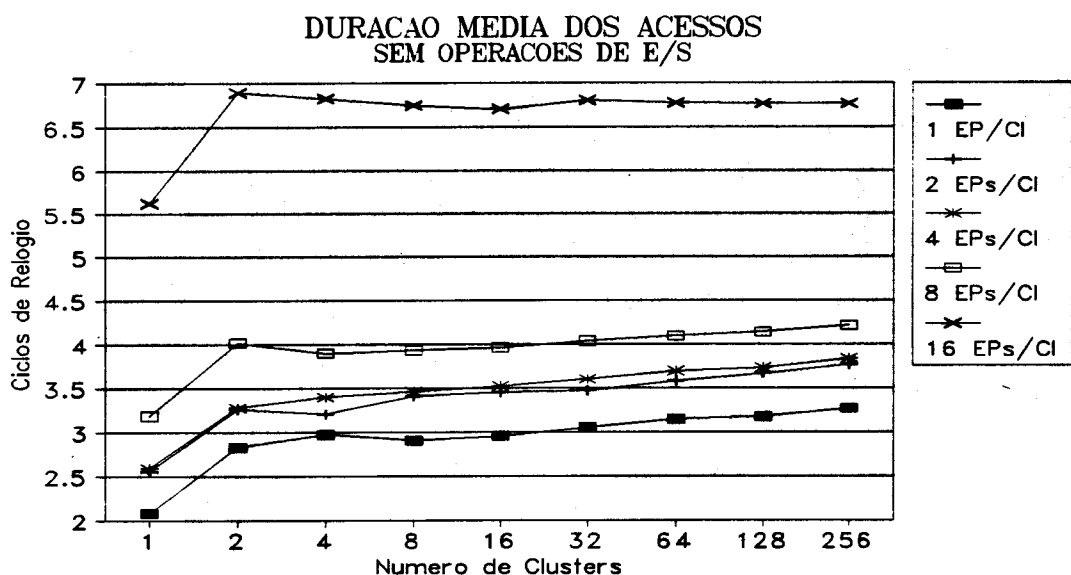


FIGURA 3: Duração média dos acessos externos executados pelas unidades inteiras sem as operações de E/S.

modelamento do subsistema de E/S. Isto permitirá avaliar o impacto das operações de E/S no desempenho destes elementos. O gráfico da figura 3 ilustra a duração média dos acessos externos executados pelas unidades inteiras dos elementos processadores do Multiplus para as diversas possibilidades de configuração. Correspondentemente, as taxas de ocupação resultantes para os barramentos de código e dado são mostradas nos gráficos das figuras 4 e 5.

A primeira questão avaliada foi a escolha do barramento do *cluster* no qual serão realizadas as operações de E/S. Como as informações contidas nas operações de E/S são tratadas como dados pelos elementos processadores, é sugestivo pensar que, para manter uma homogeneidade, os PES devem utilizar o barramento de dado para realizá-las. Entretanto, mais importante que a homogeneidade é o desempenho do sistema. As figuras 4 e 5 mostram que o

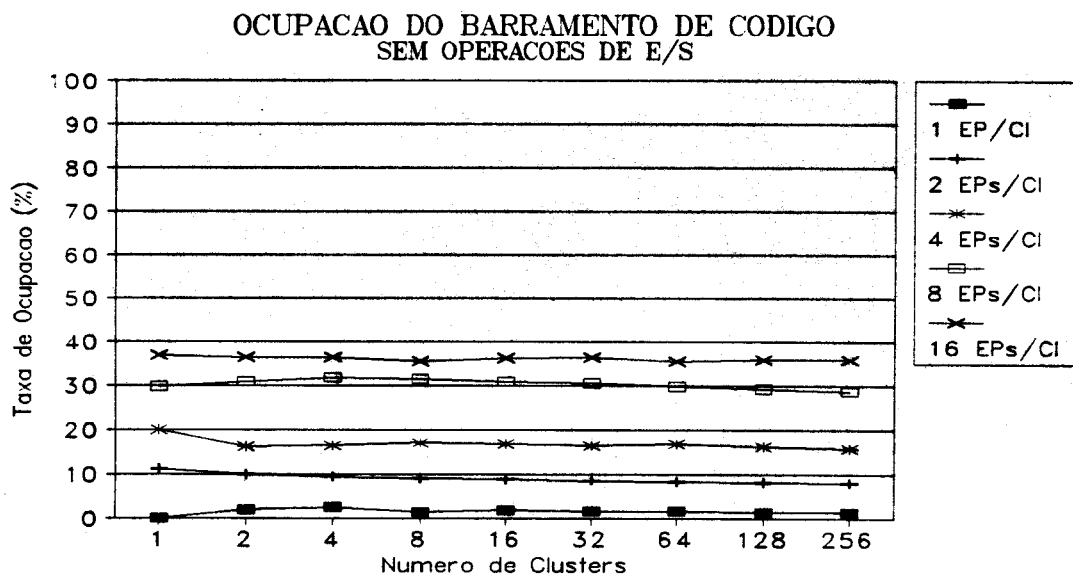


FIGURA 4: Taxa de ocupação do barramento de código do *cluster* de elementos processadores sem operações de E/S.

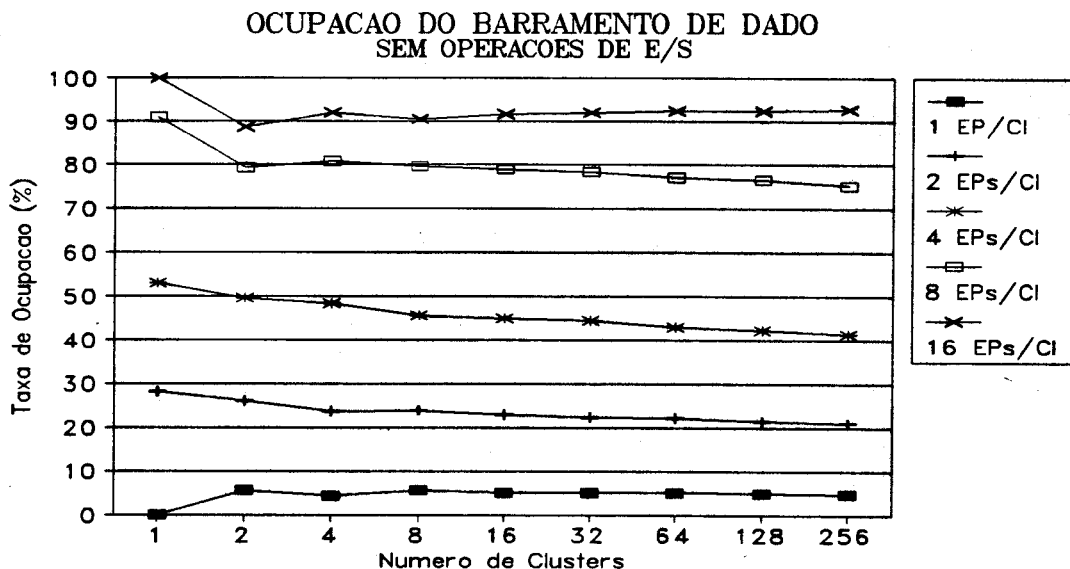


FIGURA 5: Taxa de ocupação do barramento de dado do *cluster* de elementos processadores sem operações de E/S.

barramento de dado, pelas próprias características do Multiplus, é bem mais congestionado que o barramento de código, principalmente nas configurações com muitos elementos processadores por *cluster*. Desta forma, optou-se por realizar as operações de E/S através do barramento de código, aproveitando sua maior ociosidade. Todas as questões subsequentes foram avaliadas considerando a utilização deste barramento para realização das operações de E/S. Porém, com o objetivo de ratificar esta opção, avaliações semelhantes utilizando o barramento de dado para as operações de E/S serão mostradas oportunamente ao longo do texto, permitindo o confronto dos resultados.

A segunda questão avaliada foi a influência das operações de E/S no desempenho dos elementos processadores. Esta influência pode ser determinada pela degradação da duração média dos acessos externos executados pelas unidades inteiras. O procedimento adotado foi simular o comportamento das unidades inteiras durante uma sequência de rajadas provenientes de uma transferência de E/S e estender o resultado, analiticamente, para diversas cargas de E/S.

O gráfico da figura 6 ilustra os resultados obtidos pela simulação. Comparando-o com o gráfico da figura 3, pode-se obter o fator de degradação da duração média dos acessos externos das unidades inteiras durante uma sequência de rajadas de E/S. Este fator de degradação é mostrado na figura 7. Nota-se que, como a prioridade do PES na alocação das operações de E/S é menor que a das unidades inteiras, o fator de degradação decresce com o aumento da quantidade de elementos processadores por *cluster*. Isto deve à maior ocupação do barramento do *cluster* pelos elementos processadores, resultando numa menor interferência por parte das operações de E/S e, conseqüentemente, numa menor banda passante do barramento do *cluster* para estas operações.

É interessante notar, também pelo gráfico da figura 6, que esta degradação decrescente resultou num ponto de mínimo na duração média dos acessos, percebido quando se fixa o número de *clusters* e varia-se o número de elementos processadores por *cluster*. Isto deve à

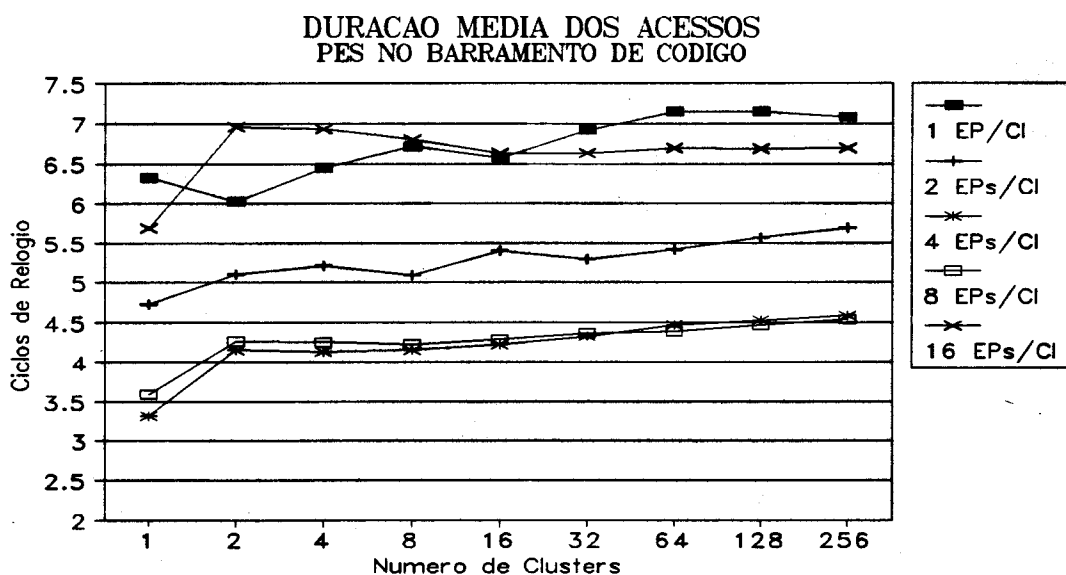


FIGURA 6: Duração média dos acessos externos das unidades inteiras durante uma sequência de rajadas de E/S.

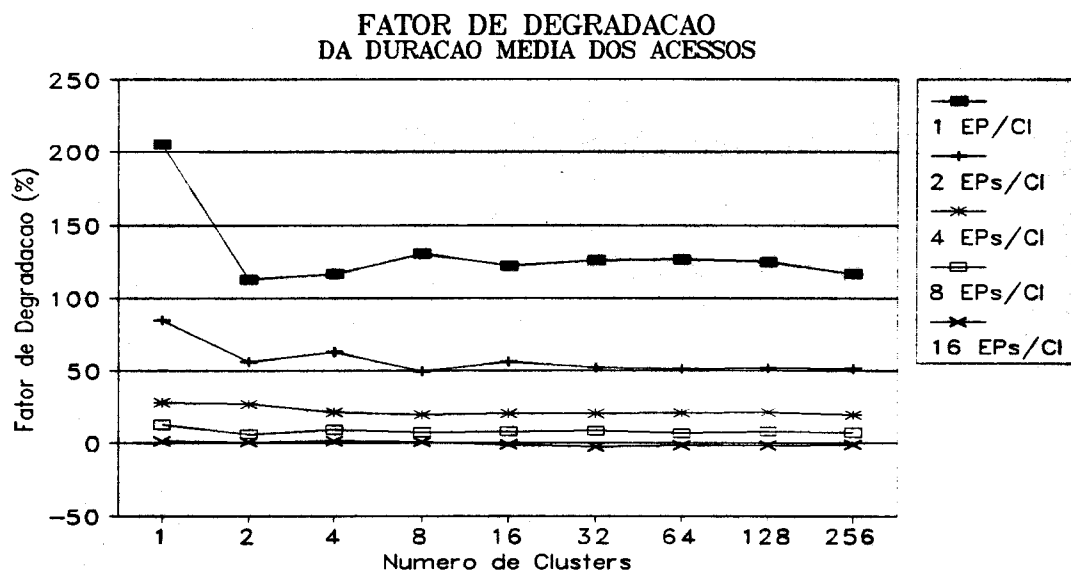


FIGURA 7: Degradação da duração média dos acessos externos das unidades inteiras durante as transferências de E/S.

acentuada elevação da duração média dos acessos nos extremos da faixa. Para 1 EP por *cluster*, a duração média dos acessos é alta em decorrência das operações de E/S. Já para 16 EP por *cluster*, o congestionamento do barramento do *cluster* é, pelas próprias operações das unidades inteiras, o responsável pela elevação da duração média dos acessos.

Para estender o resultado analiticamente, supôs-se quatro cargas de E/S passíveis de serem impostas pelos elementos processadores: 0,5 MBytes/s, 2 MBytes/s, 4 MBytes/s e 16 MBytes/s. A escolha destes valores teve como base a suposição de que cada MIPS do processador demanda 1 MBit/s de E/S [SIEWIOR83]. Como cada processador SPARC tem a capacidade de processar 16 Mips a 25 MHz, a expectativa é de que cada um deles imponha uma carga de 2 MBytes/s ao subsistema de E/S. A partir deste valor, decorreram-se as demais cargas de E/S. Procurou-se cobrir uma faixa de valores que pudesse, tanto conter as cargas de E/S esperadas para o Multiplus, quanto evidenciar seu impacto no desempenho dos elementos processadores.

Cada uma das cargas de E/S supostas reflete, nas unidades inteiras, um comportamento distinto. Durante as transferências de E/S provenientes destas cargas, a duração média dos acessos executados cresce (ver figura 6). Quanto maior a carga de E/S, maior o tempo em que os acessos das unidades inteiras permanecem nesse patamar mais elevado de duração. Calculando-se, para cada uma destas cargas, a média, ponderada em relação ao tempo, da duração média dos acessos durante as transferências de E/S (figura 6) e fora delas (figura 3), pode-se obter os gráficos das figuras 8, 9, 10 e 11. Cabe lembrar que, para elaboração destes gráficos, considerou-se que a frequência dos ciclos de relógio, responsável pela temporização dos acessos das unidades inteiras, é de 25 MHz, frequência que será utilizada na implementação do Multiplus.

Nota-se que a duração média dos acessos externos executados pelas unidades inteiras é pouco influenciada pelas operações de E/S, desde que estas operações sejam provenientes de

cargas de E/S pouco expressivas. Para cargas de 0,5 MBytes/s, 2 MBytes/s e 4 MBytes/s a degradação foi da ordem de 1%, 4% e 8%, respectivamente. Este resultado pode ser explicado pela própria origem dos dados. Primeiramente, a degradação da duração média dos acessos durante as rajadas de E/S decresce significativamente com o aumento do número de elementos processadores por *cluster* (ver figura 7). Segundo, a carga de E/S total é proporcional ao número de elementos processadores por *cluster*. Desta forma, ao se calcular a média ponderada, verifica-se que, quando a degradação é mais significativa, a carga de E/S é pequena, influenciando pouco o resultado. À medida em que cresce o número de elementos processadores por *cluster*, o peso das operações de E/S no cálculo da média ponderada é maior. Entretanto, a

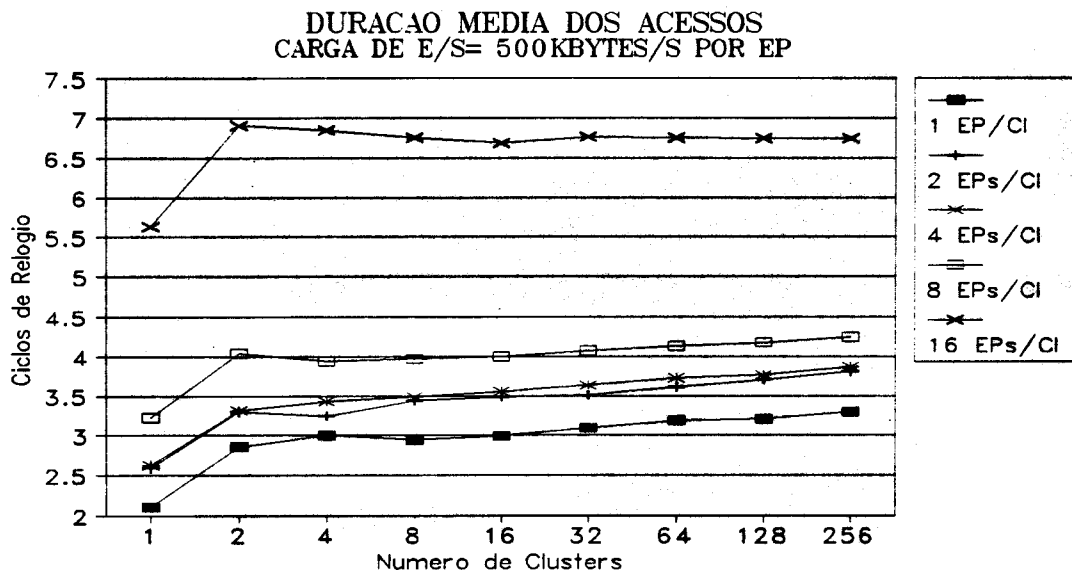


FIGURA 8: Duração média dos acessos externos das unidades inteiras.
Carga de E/S de 0,5 MBytes/s por elemento processador.

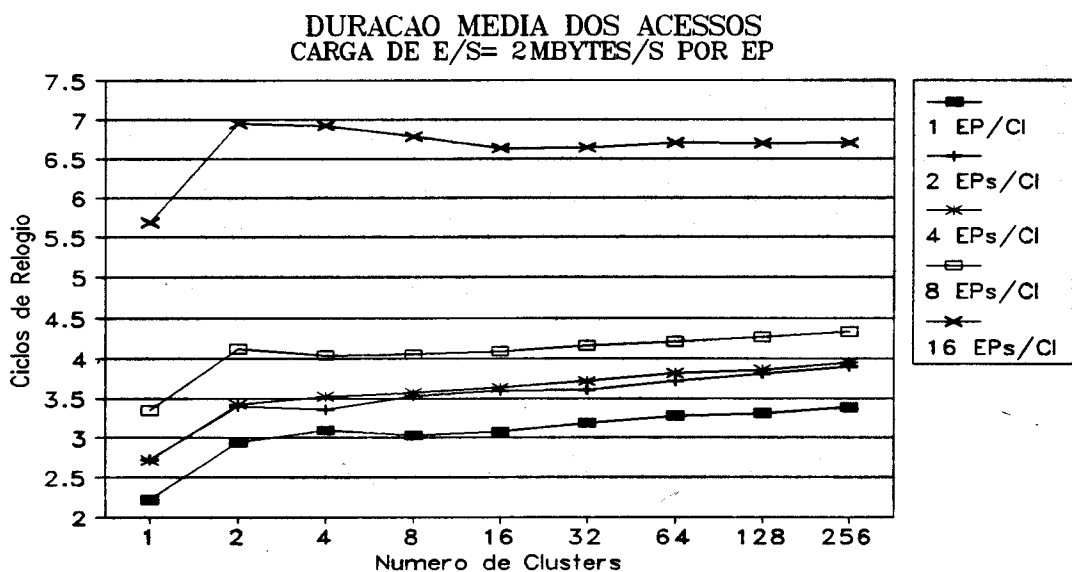


FIGURA 9: Duração média dos acessos externos das unidades inteiras.
Carga de E/S de 2 MBytes/s por elemento processador.

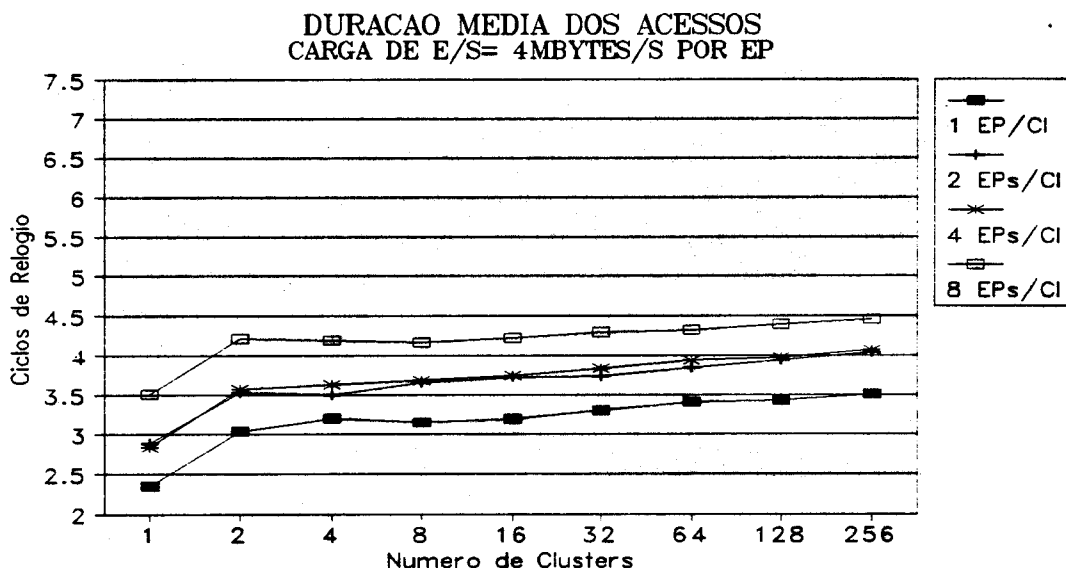


FIGURA 10: Duração média dos acessos externos das unidades inteiras.
Carga de E/S de 4 MBytes/s por elemento processador.

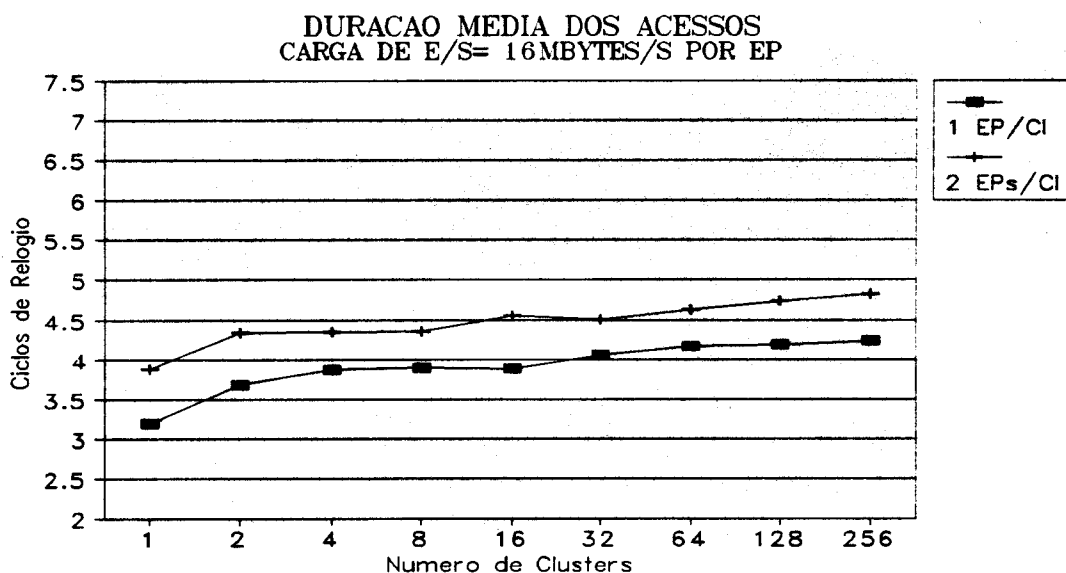


FIGURA 11: Duração média dos acessos externos das unidades inteiras.
Carga de E/S de 16 MBytes/s por elemento processador.

degradação da duração dos acessos é pouco expressiva, o que, novamente, ocasiona pouca influência no resultado.

Para que as operações de E/S degradem significativamente a duração média dos acessos externos executados pelas unidades inteiras é necessário que se tenha uma alta carga de E/S e poucos elementos processadores por *cluster*. O gráfico da figura 11 ilustra esta situação, onde cada elemento processador impõe uma carga de E/S de 16 MBytes/s. A degradação resultante é da ordem de 30%.

Nas figuras 10 e 11, os gráficos não foram traçados para algumas configurações. Isto se deve ao fato de, nestas configurações, a carga de E/S imposta ser superior à banda passante máxima do barramento de código para as operações de E/S, impossibilitando a análise.

Cabe lembrar que estas conclusões são referenciadas no critério de que a prioridade do PES para realização de operações de E/S é menor que a dos elementos processadores. Se a prioridade fosse invertida, chegaria-se a resultados mais favoráveis às operações de E/S, entretanto, a degradação do desempenho dos elementos processadores seria maior.

Considerando, agora, que as operações de E/S se realizam através do barramento de dado, avaliações semelhantes podem ser feitas. O gráfico da figura 12 ilustra a duração média dos acessos externos executados pelas unidades inteiras dos elementos processadores durante uma transferência de E/S no barramento de dado para diferentes configurações do Multiplus. Pode-se notar que não ocorreram alterações significativas na duração média dos acessos se compararmos com os resultados ilustrados na figura 6, onde utilizou-se o barramento de código para as operações de E/S. Este resultado pode ser explicado observando-se as taxas de ocupação dos barramentos de código e dado sem as operações de E/S (figuras 4 e 5). Verifica-se que elas são bastante semelhantes nas configurações do Multiplus mais susceptíveis à degradação da duração média dos acessos, isto é, poucos elementos processadores por *cluster*. Quando a diferença entre as taxas de ocupação se acentua, a configuração correspondente é pouco sensível às operações de E/S, o que justifica o resultado.

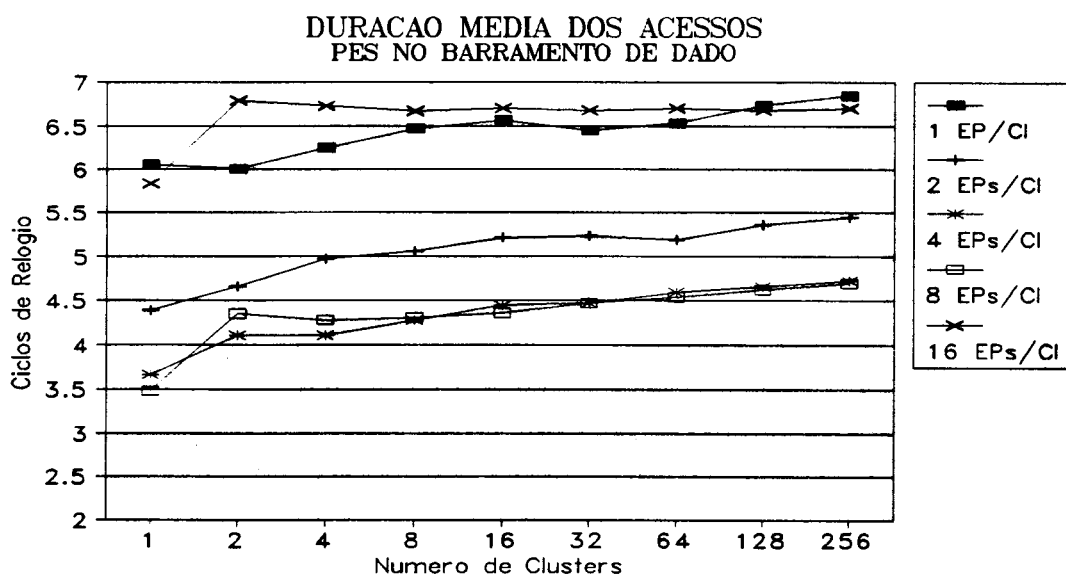


FIGURA 12: Duração média dos acessos externos das unidades inteiras durante uma sequência de rajadas de E/S no barramento de dado.

Desta forma, a degradação do desempenho dos elementos processadores mostra-se pouco sensível ao barramento onde as operações de E/S se realizam. Por outro lado, o fato de o barramento de dado ser mais congestionado que o de código, restringe bastante sua banda passante para as operações de E/S, como será mostrado a seguir.

A última questão avaliada foi a taxa de transferência média efetivamente obtida nas operações de E/S, isto é, a banda passante do barramento do *cluster* para este tipo de operação. O procedimento adotado foi simular o comportamento do barramento de código do Multiplus durante uma transferência de E/S monitorando a quantidade de rajadas efetuadas por unidade de ciclo de relógio.

Para permitir a interpretação dos resultados obtidos é necessário destacar alguns critérios adotados. Cada acesso ao barramento de código dentro das rajadas tem duração de 2 ciclos. Como as rajadas são de 128 *bytes* e a largura do barramento é de 64 *bits*, são necessários 16 acessos, totalizando 32 ciclos para transferí-los. Somando-se a estes 32 ciclos outros 8 ciclos referentes ao atraso até a aquisição do barramento, a duração de cada rajada totaliza 40 ciclos de relógio, resultando numa taxa de transferência máxima de 3,2 *bytes* por ciclo nas operações de E/S. Entretanto, devido à menor prioridade do PES para realização de rajadas, elas são espaçadas entre si. Isto resulta numa taxa efetiva de transferência nas operações de E/S menor, que é função da taxa de ocupação do barramento de código pelos elementos processadores. O gráfico da figura 13 ilustra o resultado desta simulação para diferentes configurações do Multiplus. Pode-se notar que quanto maior o número de elementos processadores por *cluster*, menor a taxa efetiva de transferência de E/S obtida no barramento de código.

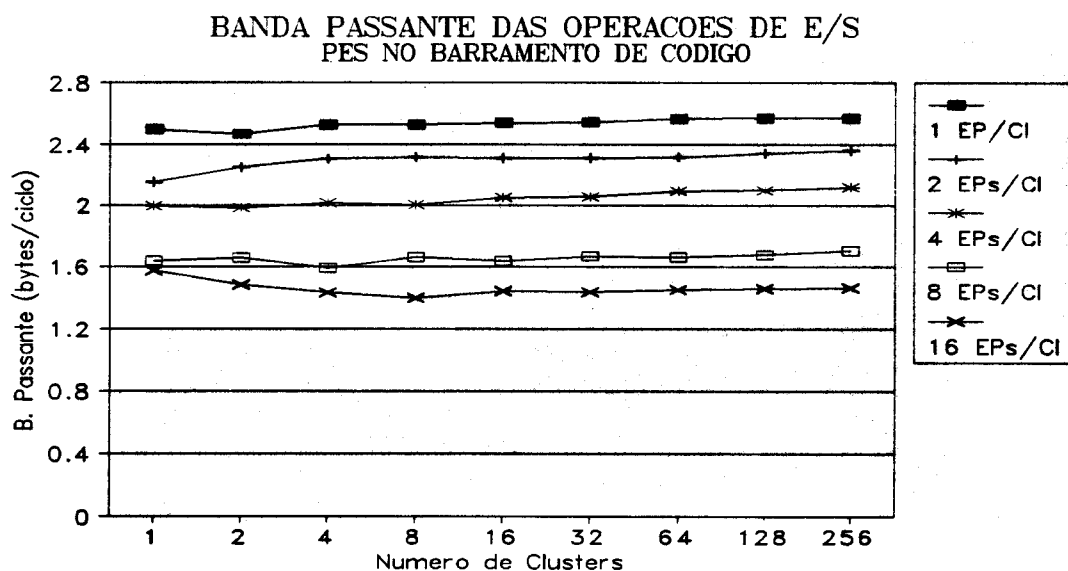


FIGURA 13: Banda passante efetiva do barramento de código para as operações de E/S.

Analogamente, se as operações de E/S forem realizadas através do barramento de dado, os resultados se mostram bem diferentes. A quantidade de rajadas efetuadas por unidade de ciclo de relógio reduz-se acentuadamente, refletindo de forma significativa na banda passante do barramento de dado para as operações de E/S. O gráfico da figura 14 ilustra os resultados da simulação. Comparando-o com o gráfico da figura 13 pode-se perceber que a banda passante do barramento de dado para as operações de E/S é bastante inferior ao do barramento de código. Desta forma, sua utilização para este fim restringe bastante a carga de E/S suportada pelo subsistema de E/S, em detrimento da versatilidade e desempenho do Multiplus.

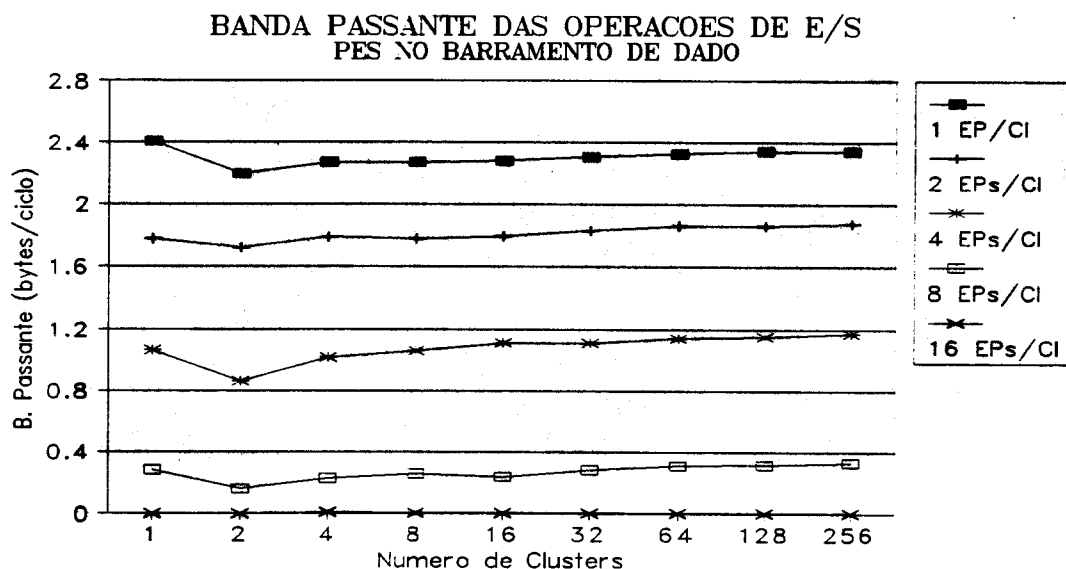


FIGURA 14: Banda passante efetiva do barramento de dado para as operações de E/S.

4 - CONCLUSÃO

Os resultados das simulações permitem concluir que o barramento do *cluster*, canal de comunicação entre o PES e os elementos processadores do Multiplus, é capaz de suportar satisfatoriamente as operações de E/S. Mesmo supondo que os elementos processadores demandem cargas de E/S excessivamente elevadas, a banda passante efetiva do barramento do *cluster* para as operações de E/S é bastante confortável, permitindo o compartilhamento do barramento sem interferência significativa na operacionalidade dos elementos processadores do Multiplus.

Estas conclusões se baseiam na suposição de que o PES é capaz de transferir dados na taxa máxima do barramento do *cluster*. Como o disco magnético, principal dispositivo de armazenamento secundário, tem uma banda passante bem inferior à do barramento do *cluster*, é necessário que o PES "bufferize" os dados antes de fornecê-los. No Multiplus, este *buffer* é representado por uma memória *cache* de disco. O bom dimensionamento desta *cache* é fundamental para o desempenho do subsistema de E/S.

5 - AGRADECIMENTOS

Os autores agradecem ao CNPq e à FINEP o apoio dado ao desenvolvimento deste projeto.

6 - REFERÊNCIAS

- [AKELLA91] AKELLA, J *et alii*. "Modeling and Measurement of the Impact of Input/Output on System Performance", Proceedings of the 18th Annual International Symposium on Computer Architecture, Toronto, Canada, pp 390-399, May 1991..

- [AUDE91] AUDE, J. Salek *et alii*. "MULTIPLUS: A Modular High- Performance Multiprocessor", Proceeding of the EUROMICRO, Microprocessing and Microprogramming, North-Holland, vol 32, pp 45-52, 1991.
- [BERRA89] BERRA, P. Bruce *et alii*. "Optical and Supercomputing". Proceedings of the IEEE, Special Issue on Supercomputer Technology, pp 1797-1815, Dec 1989.
- [BORRILL89] BORRILL, Paul L. "High-speed 32-bit Buses for Forward- looking Computers". IEEE Spectrum, pp 34-37, Jul 1989.
- [BRONST91] BRONSTEIN, Gerson. "Projeto de uma Rede de Interconexão para uma Máquina Paralela de Alto Desempenho". Tese M.Sc. COPPE/UFRJ, Novembro de 1991, 99 pp.
- [HWANG87] HWANG, Kai. "Advanced Parallel Processing with Supercomputer Architectures". Proceeding of IEEE, pp 1348-1379, vol 75, no 10, Oct 1987.
- [KATZ89] KATZ, Randy H. *et alii*. "Disk System Architectures for High Performance Computing", Proceedings of the IEEE, pp 1842- 1858, vol 77, no 12, Dec 1989.
- [MESLIN91] MESLIN, Alexandre M. "Estudos de Arquiteturas de Memórias Cache para o Multiprocessador Multiplus". Tese M.Sc. COPPE/UFRJ, Agosto de 1991, 114 pp.
- [OLIVEIRA90] OLIVEIRA, Sidney C. *et alii*. "O Subsistema de Memória de Massa do Multiprocessador Multiplus". Anais do III-Simpósio Brasileiro de Arquitetura de Computadores - Processamento Paralelo, pp. 298 - 313, Rio de Janeiro, Novembro, 1990.
- [OLIVEIRA92] OLIVEIRA, Sidney C. "Uma Proposta de Arquitetura de E/S para o Multiprocessador Multiplus". Tese M.Sc. COPPE/UFRJ, Fevereiro de 1992, 84 pp.
- [PATTER90] PATTERSON, David A. *et alii*. "Computer Architecture: A Quantitative Approach". Morgan Kaufmann Publishers, 1990.
- [PETER85] PETERSON, J. L. *et alii*. "Operating System Concepts". Addison Wesley Publishing Company, 1985.
- [PIEPER89] PIEPER, John S. "Parallel I/O Systems for Multicomputers". School of Computer Science, Carnegie Mellon University, Pittsburgh, CMU-CS-89-143, 1989.
- [SIEWIOR83] SIEWIOREK, D. P. *et alii*. "Computer Structures: Principles and Examples", McGraw-Hill Book Company, New York, NY, 1983.
- [SMITH85] SMITH, A. J. "Disk Cache: Miss Ratio Analysis and Design Considerations". ACM Transactions on Computer Systems, New York, pp 161-203, vol 3, no 3, Aug 1985.