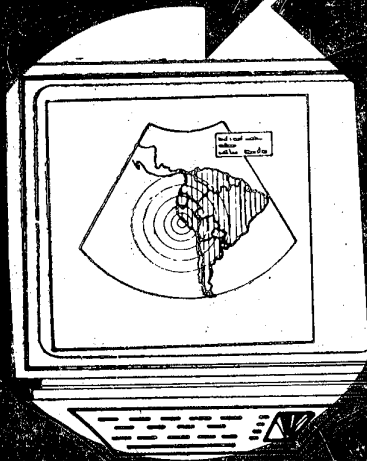


**PANEL
INFO '82**



IX CONFERENCIA LATINOAMERICANA DE INFORMATICA



**16 AL 20 DE AGOSTO DE 1982
LIMA - PERU**

**III
Convención Peruana
de
Computación é
Informática**

004.06
C748
V.2

UM PROTOCOLO PARA REDES LOCAIS DO TIPO DIFUSAO ("BROADCAST")

Luiz Fernando Gomes Soares

Daniel Alberto Menascé

Pontificia Universidad Católica
Rio de Janeiro-RJ-Brasil

1. Introdução

Nos últimos anos bastante atenção tem sido dada a sistemas de arquitetura distribuida. Como definido em [1], estes sistemas caracterizam-se pela capacidade de realizar operações concorrentes, onde a noção de estado global é substituida pela de estado local. Em tais sistemas o compartilhamento de memória deve ser evitado. O sistema é composto de um número ilimitado, mas finito, de elementos de processamento-armazenamento, capazes de cooperar concorrentemente.

Uma máquina distribuída vai ser então formada por um conjunto destes módulos processadores (MP), interligados por um sistema de comunicação(SC), conforme a figura 1. Exemplos de SCs em redes locais podem ser encontrados em [1,4 e 5].

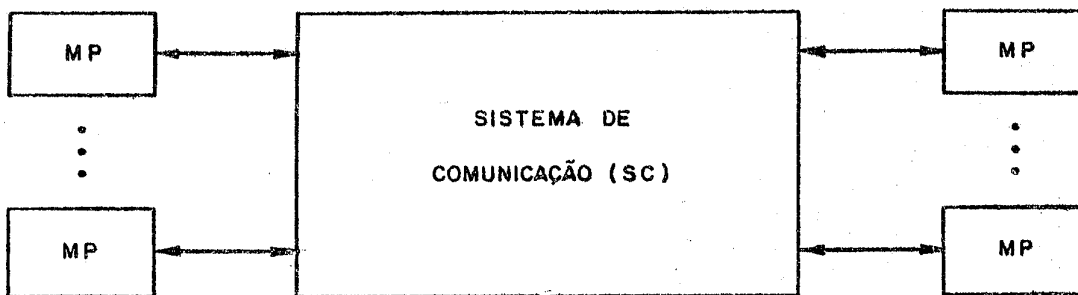


figura 1: Esboço de uma máquina de arquitetura distribuída.

O sistema de comunicação que provê a infra-estrutura para a comunicação dos MP's, é formado por um Sistema Físico de Interconexão (SFI) e um protocolo de comunicação interna.

O SFI por sua vez é composto por um arranjo topológico interligando através de enlaces físicos os Processadores de Comunicação (PC), que são os responsáveis pela realização do protocolo de comunicação interna ao SFI. A cada módulo processador está associado um ou mais PC's.

Este artigo descreve o protocolo de comunicação utilizado em um SFI do tipo difusão ("broadcast") composto por dois embarramentos seriais compartilhado pelos diversos PCs (vide figura 2). Este SFI faz parte de uma rede local em desenvolvimento na PUC/RJ [2 e 3].

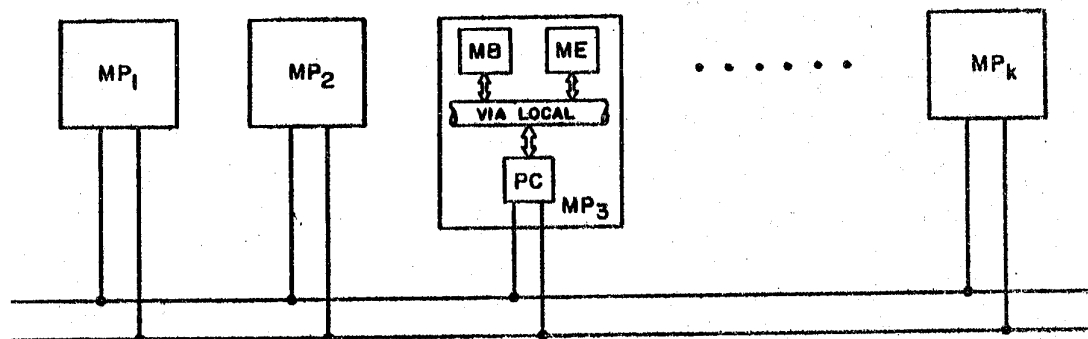


figura 2: A estrutura da REDPUC

Obviamente, pela própria definição de máquina de arquitetura distribuída, o controle do acesso ao embarramento é distribuído. A não existência de um elemento centralizador evidentemente aumenta a disponibilidade do sistema.

2. Protocolo de Acesso ao Embarramento (PAE)

O protocolo de acesso ao embarramento (PAE) compreende três níveis, independentes entre si. Os dois primeiros níveis podem ser substituídos por qualquer outro protocolo que realiza as mesmas funções. O terceiro nível é o que dá à interface do PAE sua identidade, e a sua substituição resultará em uma interface diferente. Aos três níveis, 1, 2 e 3, podemos dar o nome, respectivamente, de nível físico, nível de embarramento e nível de transporte.

O nível físico especifica as características físicas e elétricas do PAE, tais como níveis de tensão, conexões físicas e métodos de transmissão. As características importantes deste nível, vista pelo nível imediatamente acima são: transmissão serial bit a bit, síncrona, full-duplex em um sistema de comunicação do tipo difusão usando técnica de codificação Manchester. O uso da codificação Manchester vem da necessidade de se transmitir o relógio juntamente com os dados para a sincronização entre o transmissor e os receptores. Esta técnica vai possibilitar o controle distribuído da sincronização nos embarramentos.

O nível de embarramento, é o responsável pela ligação e desligamento de um módulo na rede, bem como do controle de acesso ao embarramento. Dentre as suas funções destacam-se: transferência, temporização, validação dos quadros, deferência para transmissões, transferência do controle do

do embarramento, recepção de mensagens, gerenciamento da lista de controle do ciclo, tratamento das omissões de transmissão (desligamento de um módulo), tratamento de um intervalo de contenção (ligação de um módulo).

Finalmente, o nível 3, nível de transporte converte a linha-física, propensa a erros, em uma ligação lógica, relativamente livre de erros, para a transferência de dados entre o transmissor e o receptor. Mais ainda, transforma a ligação lógica, em vários canais lógicos, realizando a função de um concentrador. Este nível realiza o controle de erro e de fluxo nos vários canais lógicos, bem como a seqüenciação das mensagens no mesmo.

2.1 Nível de Embarramento

O formato da mensagem a nível de embarramento é mostrada na figura 3.a

FLAG	END. PRÓXIMO	INFORMAÇÃO (NÍVEL 3)	FCS	FLAG
------	--------------	----------------------	-----	------

3.a - Quadro de informação: END Próximo \neq 00.

FLAG	00 _H	MOD. ORIGEM	FCS	FLAG
------	-----------------	-------------	-----	------

3.b - Quadro de supervisão do nível 2: pedido de inserção no ciclo.

Figura 3 - Quadros do nível de embarramento

Transparência

O protocolo do nível de embarramento é orientado a bit. Um quadro é delimitado por flags de oito bits. A transparência de dados é conseguida através da técnica de inserção e deleção de bits a exemplo do que ocorre no HDLC.

Transferência do controle do embarramento

A cada mensagem transmitida está associado o endereço do próximo módulo a assumir o controle do embarramento. Ao receber um quadro cujo conteúdo do campo "endereço do próximo" foi idêntico ao seu código de identificação, o módulo habilita o seu transmissor tão logo o embarramento se torne disponível.

Deferência para transmissão

Cada módulo só transmite quando recebe o controle do acesso ao embarramento, seja pela recepção de uma mensagem cujo conteúdo do campo "endereço do próximo" é igual ao seu código de identificação, seja pela execução do procedimento de inicialização/reinicialização. O módulo ao receber o controle de acesso ao embarramento espera que a transmissão em andamento termine para então começar a transmitir.

Temporização

Após a detecção do flag final de quadro, caracterizando o final de uma transmissão, cada módulo ativa um temporizador para controlar o tempo máximo que o próximo módulo dispõe para iniciar a transmissão. Caso a temporização se esgote, sem que ocorra transmissão, procedimentos são executados visando recuperar a situação. Devido a esta forma de controle, módulos, que ao receberem o controle do embarramento, não dispuserem de mensagens do nível de transporte para transmitir, devem transmitir mensagens com o campo de informação vazio, a título de "atestado de vida".

Validação

O campo FCS ("frame checking sequence") contém um código cíclico que permite ao receptor efetuar a detecção de erros conforme é feito no HDLC.

Recepção de mensagens

O PAE toma vantagem do fato do meio de transmissão ser um meio do tipo difusão, isto é, todos os módulos escutam o módulo transmissor. Cada módulo recebe todas as mensagens transmitidas, mesmo que não seja ele o módulo de destino, e passa o campo de informação para o nível de transporte, se o quadro for considerado isento de erros.

Gerenciamento da lista de controle de ciclo

O acesso ao embarramento por parte dos módulos é cíclico e está baseado no conceito de lista de controle de ciclo, que é o elemento responsável pela seqüenciação lógica das transmissões. A lista de controle do ciclo contém todos os endereços possíveis (256 ao todo) e associa a cada endereço de módulo o seu número de ordem no ciclo e o seu estado operacional (ativo ou inativo), com exceção dos endereços $(00)_H$ e $(FF)_H$.

Entende-se como gerenciamento da lista de controle do ciclo a sua criação e a manutenção de sua atualidade, seja por inclusões ou exclusões de módulos, de modo que ela sempre indique quantos e quais são os módulos ativos em um dado instante e qual o seu interrelacionamento lógico para fins de seqüenciação do acesso ao embarramento. As inclusões/exclusões são processadas alterando-se, em cada lista, o estado do módulo de inativo/ativo para ativo/inativo e reordenando-se a lista; isto é, calculando os novos números de ordem dos módulos com endereço superior ao do incluído/excluído.

Tratamento de Omissões de Transmissão

Quando um módulo deixa de transmitir por três vezes consecutivos, ele é retirado da lista de controle do ciclo de todos os módulos. A ausência de transmissão é caracterizada na codificação Manchester por ausência de transições no cabo e pelo esgotamento da temporização. Este procedimento é descrito a seguir:

```
<< esgotamento da temporização 3 vezes seguidas >>
BEGIN
  retira módulo não transmissor da lista;
  IF endereço do proximo = endereço do não transmissor
  THEN altera endereço do próximo para o novo proximo da lista;
END
```

Tratamento do intervalo de contenção

O intervalo de contenção é o mecanismo através do qual há a inclusão de módulos em um ciclo já estabelecido. O intervalo de contenção é o último intervalo de cada ciclo e é ativado pelo último módulo do ciclo por meio de um quadro com o conteúdo do campo endereço do próximo com o valor $(FF)_H$. Existem três modos como o intervalo de contenção pode terminar:

- a) a temporização se esgota sem que ocorra qualquer transmissão.
- b) uma única transmissão ocorre, sem que sofra qualquer interferência, fazendo com que o módulo que transmitiu seja incluído em todas as listas.
- c) dois ou mais módulos tentam a inclusão gerando interferência entre os quadros, neste caso, é executado o procedimento de reinicialização com todos os módulos sendo colocados como ativo.

O quadro de inclusão no ciclo está apresentado na figura 3.b.

Sempre que o intervalo de contenção se encerra com uma mensagem bem sucedida, o módulo a ganhar lugar na lista de controle do ciclo deve criar sua própria lista de controle, a partir das leituras do endereço do próximo durante um ciclo completo.

Interferência na transmissão

No PAE, a interferência não tem como ser distinguida de um erro de transmissão, uma vez que a mesma é detetada pelo FCS. Tanto para interferência, como então para o erro, é adotado o seguinte procedimento: sempre que ocorrer interferência ou erro o quadro é simplesmente descartado. Se um módulo notar que por três vezes consecutivas ele tenta transmitir e é mal sucedido, ele suspeita de uma falha e algum mecanismo de recuperação deve ser acionado.

Inicialização e Reinicialização

Cada módulo, ao ser ligado ou ao recuperar-se de uma falha interna, executa um procedimento de inicialização. Nas situações em que parem dúvidas sobre a real situação do ciclo de acesso ao embarramento cada módulo passa a executar um procedimento de reinicialização.

O procedimento de inicialização/reinicialização consiste basicamente da observação da existência ou não de mensagens no embarramento, tomando por base uma lista de controle de ciclo internamente criada com todos os módulos (fisicamente ligados) ativos. Neste ponto, cada módulo aguarda que algum dos módulos anteriores (de menor número de ordem que o seu) transmite. Caso ocorra alguma transmissão, o módulo em questão entende que já há um ciclo estabelecido e aguarda que o controle lhe seja passado diretamente, por um endereçamento específico, ou indiretamente, via um intervalo de contenção. A cada esgotamento da temporização, o módulo de número de ordem mais baixo é excluído da lista, de modo que, após a ocorrência de n esgotamentos de temporização, o módulo n , caso ativo, enviará um de seus quadros, passando o acesso ao embarramento ao próximo módulo da sua tabela de ciclo corrente. A partir deste instante, todos os módulos continuam a criação de suas respectivas listas de controle de ciclo a partir de leituras do endereço do próximo, durante um ciclo por completo.

A manutenção da integridade do ciclo é feita da seguinte forma: cada módulo testa continuamente a sua permanência na lista de controle do ciclo que está sendo utilizada no sistema (pode ser até uma diferente para cada módulo). Sempre que algum módulo perceber a ocorrência de dois ciclos consecutivos - através do intervalo de contenção - sem que, neste período, ele tenha recebido explicitamente o controle do embarramento, o módulo em questão assume que foi retirado indevidamente da lista e executa uma reinicialização.

Resta observar que o nível 2 do protocolo de acesso ao embarramento deve ser realizado um para cada embarramento global, e estes dois protocolos são totalmente independentes.

2.2 O Nível de Transporte

A figura 4 apresenta o formato de quadros deste nível do PAE, cuja sintaxe é descrita logo após as considerações sobre endereçamento.

MÓDULO ORIGEM	O	CEF	CANAL ORIGEM		I	CEF	CANAL ORIGEM	$\theta\theta_H$	FUNÇÃO	INFO
---------------	---	-----	--------------	--	---	-----	--------------	------------------	--------	------

a) FORMATO GERAL DE QUADROS NÃO NUMERADOS

FUNÇÃO = $\theta\theta_H$	ENDEREÇO PROCESSO DESTINO	ENDEREÇO PROCESSO ORIGEM	OO CANAL ORIGEM
---------------------------	---------------------------	--------------------------	-----------------

a.1) CAMPO DE FUNÇÃO E INFORMAÇÃO DO QUADRO CON (CONEXÃO)

FUNÇÃO = θ^1_H	MÓDULO DESTINO	OO CANAL DESTINO	θ_H	ENDEREÇO PROCESSO ORIGEM	OO CANAL ORIGEM
-----------------------	----------------	------------------	------------	--------------------------	-----------------

a.2) CAMPO DE FUNÇÃO E INFORMAÇÃO DO QUADRO REC (RECONHECIMENTO DE CONEXÃO)

FUNÇÃO = θ^2_H	MÓDULO DESTINO	OO CANAL DESTINO	OO CANAL ORIGEM
-----------------------	----------------	------------------	-----------------

a.3) CAMPO DE FUNÇÃO E INFORMAÇÃO DO QUADRO DISC (DESCONEXÃO)

FUNÇÃO = θ^3_H	MÓDULO DESTINO	OO CANAL ORIGEM
-----------------------	----------------	-----------------

a.4) CAMPO DE FUNÇÃO E INFORMAÇÃO DO QUADRO RED1 (RECONHECIMENTO DE DESCONEÇÃO)

FUNÇÃO = θ^4_H	ENDEREÇO PROCESSO	...	ENDEREÇO PROCESSO;
-----------------------	-------------------	-----	--------------------

a.5) CAMPO DE FUNÇÃO E INFORMAÇÃO DO QUADRO MPRO (MUDANÇA DE PROCESSO)

MÓDULO ORIGEM	O	CEF	CANAL ORIGEM		I	CEF	CANAL ORIGEM	MÓDULO DESTINO	CANAL DESTINO	M	N	INFO = DADOS
---------------	---	-----	--------------	--	---	-----	--------------	----------------	---------------	---	---	--------------

b) FORMATO GERAL DE QUADROS DE DADOS

Figura 4

Endereçamento

Existem cinco tipos de endereçamento no PAE: endereçamento de módulo, endereçamento de processo, endereçamento de grupo, endereçamento global e endereçamento de canal. Cada módulo é identificado por um endereço fixo de 8 bits, com exceção dos endereços 00_H e FF_H .

A cada módulo é atribuído um ou mais processos. Cada processo é identificado univocamente por um endereço, que não tem relação alguma com o endereço do módulo. Esta independência vai dar uma maior facilidade na distribuição de processos pelos módulos, bem como vai facilitar em muito o mecanismo de reconfiguração do sistema. Cada módulo deve conhecer apenas os endereços dos processos que ele suporta. Cada processo pode mover-se a qualquer tempo de um módulo para outro.

Cada módulo permite até 63 canais de comunicação entre processos em diferentes módulos. Cada canal é estabelecido por procedimentos que serão vistos mais adiante, e é alocado exclusivamente para a comunicação entre apenas dois processos. O endereçamento de canal é realizado através de 6 bits. O canal 0 é reservado para endereçamento global e de grupo.

O protocolo de acesso ao embarramento provê facilidades para endereçamento de grupo e endereçamento global ao se enviar um quadro [2].

Controle de Erro e de Fluxo

Quando um processo deseja se comunicar com outro processo residente em outro módulo, é estabelecida uma ligação entre eles. Esta ligação é identificada por dois endereços de canais - que podem ser diferentes - um em cada módulo. A ligação é identificada pelo endereço de um dos dois canais e pelo endereço do módulo correspondente. O controle de erro e de fluxo é realizado em cada canal. O procedimento utilizado para a realização destes controles explora o fato do meio ser do tipo "broadcast", e mais ainda, explora também o fato de que o acesso ao embarramento é cíclico e que cada módulo uma vez tendo realizado uma transmissão só vai poder realizar nova transmissão depois que todos os outros módulos ativos transmitirem (ou tiverem chance de transmitir).

O controle de erro e controle de fluxo são realizados através de uma variação do protocolo do bit alternado, ilustrado no diagrama da figura 5. As transições de estados $1 \rightarrow 4 \rightarrow 8 \rightarrow 1$ e $3 \rightarrow 5 \rightarrow 9 \rightarrow 3$ no transmissor e $0 \rightarrow 4 \rightarrow 0$ e $2 \rightarrow 5 \rightarrow 2$ no receptor dizem respeito ao controle de fluxo.

Note que o procedimento descrito na figura 5 mantém a sequência dos quadros por canal, uma vez que a janela é de valor 1.

Na figura 4, o bit N, no quadro de dados, é o bit de paridade do quadro. Os bits CEF's em todos os quadros, é o bit de reconhecimento. O reconhecimento é feito por canal de origem. Pode-se aproveitar qualquer tipo de quadro para o envio do reconhecimento, e pode-se enviar mais de um sinal de reconhecimento por quadro.

Os quadros com endereços globais ou de grupos são tratados em um mesmo buffer de transmissão e recepção e só pode haver uma transmissão pendente deste buffer sem confirmação. Isto vai permitir que o mesmo mecanismo

de controle de erro e de fluxo da figura 5 possa ser utilizado neste caso.

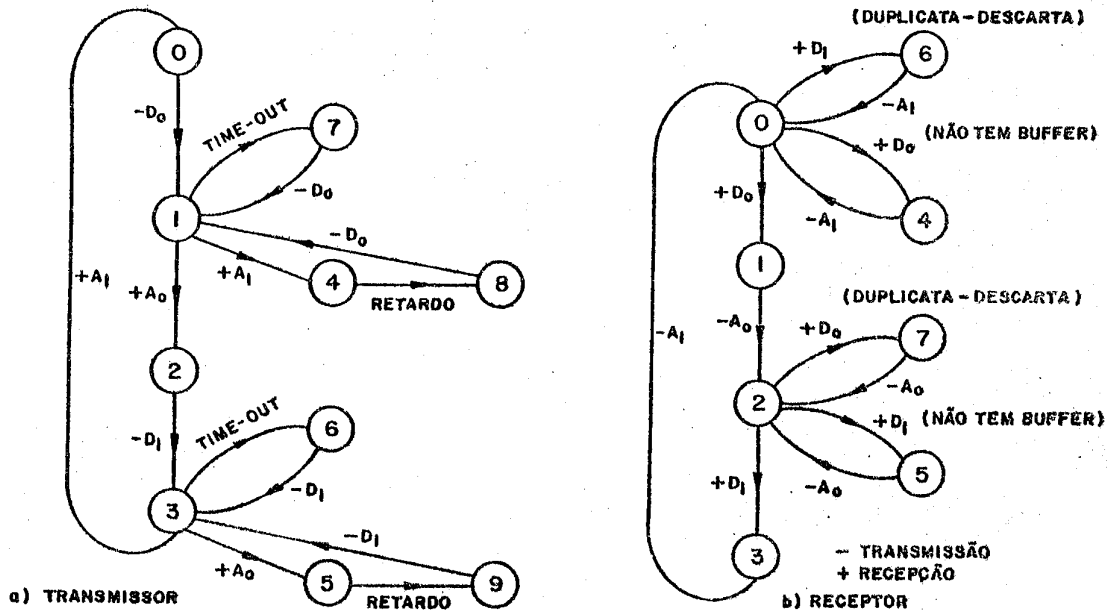


figura 5: Diagrama de estados do controle de erro e fluxo

Quadros não numerados

Antes de um processo X se comunicar com um processo Y, o processo X deve saber em que módulo está o processo Y e qual o canal de Y envolvido na ligação. As mesmas informações deve o processo Y conhecer do processo X. Os quadros CON e REC são os responsáveis pelo estabelecimento da ligação e pelas trocas de informações mencionadas. Assim que um processo X deseja se comunicar com o processo Y ele deve enviar um quadro dando suas coordenadas e pedindo a Y que envie as suas. Assim, ao enviar o quadro CON, X se anuncia: <<sou um processo "X" (endereço do processo de origem), estou no canal "a" (canal de origem) módulo "M" (módulo de origem) e quero me comunicar com o processo "Y" (endereço do processo de destino).>> Após o processo Y reconhecer que X quer se comunicar com ele, responde com REC: <<Processo X (módulo de destino e canal de destino), sou o processo Y (endereço do processo de origem) e estou no canal "b" (canal de origem), e módulo "V" (módulo de origem). A ligação está então estabelecida e os processos podem então começar a se comunicarem.

Cada módulo deve manter uma lista dos canais usados, em ordem de tempo, de forma a sempre terminar a ligação que a mais tempo não está sendo utilizada, para estabelecer nova ligação.

O procedimento de desconexão é semelhante. Um processo X querendo se desconectar de um processo Y envia o quadro DISC dizendo: <<Processo Y (módulo de destino e canal de destino) se desconecte>>. O processo Y responde com um quadro REDI: <<Processo X (módulo de destino e canal de destino), me desconectei>>.

Finalmente o quadro MPRO, serve para avisar à rede da mudança de processos entre módulos. No caso, o módulo que recebeu os processos avisa à re

de que agora os processos estão residentes nele. Se existir em algum módulo alguma comunicação com qualquer dos processos, ela tem de ser refeita. A corrupção deste quadro não traz maiores consequências, uma vez que serão feitas tentativas para comunicação com os processos nos módulos antigos e uma vez que não se tenha conseguido a ligação é desconectada e nova conexão é tentada. Este quadro simplesmente vai diminuir o fluxo de mensagens inúteis no embarramento.

Vale aqui observar que qualquer transmissão de qualquer quadro é tentada um número limitado de vezes depois do qual a conexão é desfeita.

O bit M

O bit M dos quadros que transportam dados, servem para enviar mensagens de tamanho maior do que o comprimento máximo permitido no quadro.

Resta apenas observar, ao encerrar este item, que o nível de transporte, nível 3, do protocolo de acesso ao embarramento não depende do fato de se estar usando um, dois ou mais embarramentos. Ele é totalmente independente desta especificação do embarramento global.

3. Conclusões

O protocolo descrito neste trabalho está sendo implementado na PUC/RJ. Modelos analíticos e de simulação foram desenvolvidos para avaliar o desempenho do embarramento. As curvas da figura 6 ilustram a vazão relativa (fração utilizada da capacidade do embarramento) em função do retardo médio de uma mensagem.

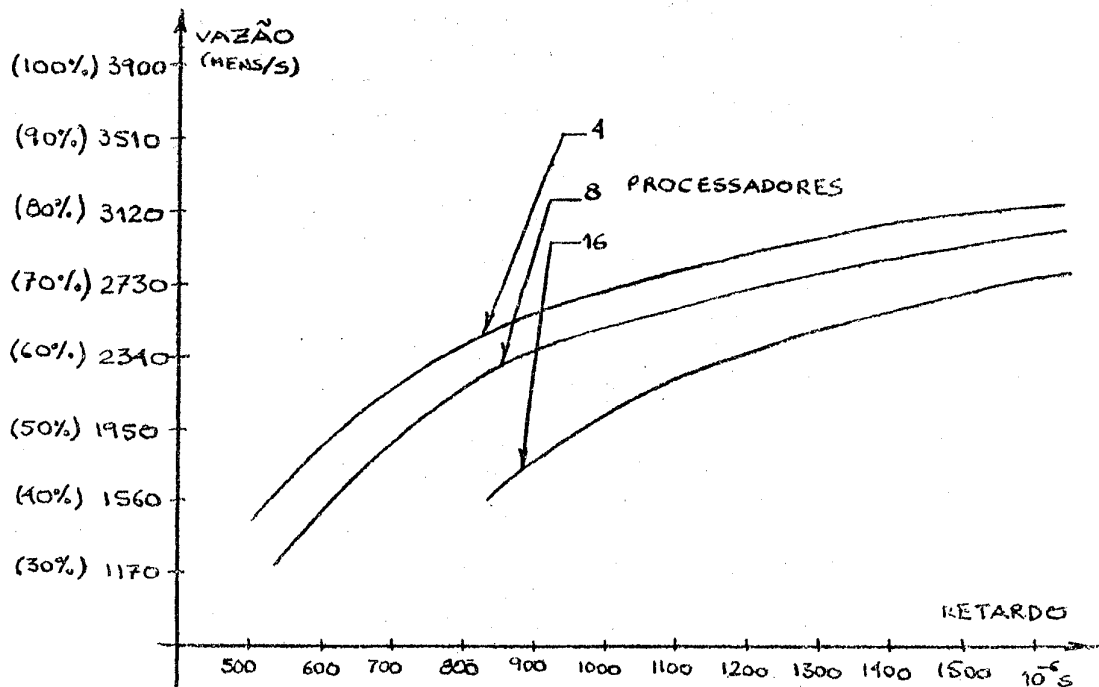


FIGURA 6

Agradecimento

Os autores gostariam de agradecer a Leonardo Lellis P. Leite que desenvolveu o programa de simulação do embarramento.

Referências

- [1] SHIMIZU, Egmont Y.; RUGGIERO, Wilson V.; MOSCATO, Lucas A. "Componentes Básicos de uma Máquina de Arquitetura Distribuída: Identificação e Implementação". Anais do VII Seminário Integrado de Software e Hardware, 1980.
- [2] Gomes Soares, Luiz F. e Daniel A. Menascê, "Arquitetura de Uma Máquina Distribuída", Relatório Técnico Pl.2 (PUC-TELEBRÁS), PUC/RJ, Outubro de 1981.
- [3] Gomes Soares, Luiz F. e Daniel A. Menascê, "Central de Comutação de Pacotes da PUC/RJ", Relatório Técnico Pl.3 (PUC-TELEBRÁS), PUC/RJ, Outubro de 1981.
- [4] Liu, M.T., "Distributed loop Computer Networks", em Advances in Computers, M.C. Yovits (ed.) New York, Academic Press, pp.163-221, 1978.
- [5] Intel, Xerox and Digital Equipments Corporation, "The Ethernet: A Local Area Network. Data Link Layer and Physical Specification", 30 de setembro, 1980.