

SUPORTE PARA DESENVOLVIMENTO DE SISTEMAS DISTRIBUIDOS BASEADO NUM BARRAMENTO PARALELO CENTRALIZADO

C. KIRNER*

E. MARQUES**

SUMÁRIO

Este artigo apresenta o projeto de um sistema distribuído baseado num barramento paralelo centralizado. Descreve-se a arquitetura do sistema e analisa-se as questões de projeto do barramento paralelo e do processador de comunicação.

ABSTRACT

This paper presents the design of a distributed systems based in a centralized parallel bus. The architecture of the systems is described and the design of the bus and communication processor is analysed.

(*) Engenheiro Eletricista(EESC-USP, 1973), Mestre em Engenharia Eletrônica(ITA, 1978) Doutorado em Engenharia de Sistemas e Computação (COPPE-UFRJ, em andamento), Sistemas Operacionais, Arquitetura, Redes. Prof. Assistente. DCEs-UFSCar - C.P. 384. 13560 - São Carlos - S.P. - Fone (0162) 71-1100 - Ramal 143.

(**) Bacharelado em Ciência da Computação (UFSCar - 1985), Bolsista de Iniciação Científica do CNPq em 1984 e 1985; Arquitetura, Redes, Sistemas Operacionais. DCEs - UFSCar; C.P. 384 - 13560 - São Carlos - S.P. - Fone (0162) 71-1100 - Ramal 143.

1. INTRODUÇÃO

Um grupo de pesquisa do Departamento de Computação e Estatística da UFSCar in vestiu, durante 2 anos, no desenvolvimento de um subsistema de comunicação com conexão ponto a ponto [6] para servir como parte de um suporte para o desenvolvimento de Siste mas Distribuídos. O fator determinante na concepção do projeto foi única e exclusivamen te o interesse acadêmico.

Um dos objetivos do projeto era o de gerar conhecimento e motivação para que o grupo, ou mesmo membros do grupo, passassem a atuar na área de redes e sistemas distri buídos, de forma mais decisiva, com um enfoque que permitisse, ao mesmo tempo, uma espe culação científica de vanguarda e uma proximidade com as necessidades atuais e futuras do mercado.

Assim, há pouco mais de 1 ano, motivados pela análise de dois tipos de rede que enfatizava a proposta de uma rede estrela coincidente [7], passamos a trabalhar na proposta de um barramento paralelo centralizado. Para isso contamos com o auxílio do CNPq que, através de uma bolsa de iniciação científica, propiciou as condições para o estudo de alguns barramentos e a elaboração, em linhas gerais, da proposta do barramen to paralelo centralizado e de suas interfaces de acesso [8].

O subsistema de comunicação, composto de vários processadores de comunicação, interligados por um barramento paralelo centralizado ("backplane"), baseia-se na exis tência de "buffers" de emissão e de recepção ligados às linhas paralelas de alta veloci dade de transmissão. Essa característica faz com que o tempo de comunicação entre dois processadores de comunicação, já preparados, seja muito pequeno, ou praticamente nulo, restando assim o tempo de preenchimento do buffer emissor e o tempo de esvaziamento do buffer receptor, perfeitamente previsíveis, de forma que o tempo total de comunicação se ja bem determinado.

Além disso procurou-se usar linhas redundantes, no barramento paralelo, com ca pacidade para substituição de algumas partes, proporcionando tolerância à falha com ca pacidade de reconfiguração das linhas, em alguns casos.

Estas duas características, ou seja, tempo determinístico na comunicação e to lerância à falha recomendam a utilização desse tipo de subsistema na solução de proble mas de aplicação em tempo real, particularmente os de controle de processo, mesmo a um custo mais alto, o que normalmente acontece nesses casos.

O projeto visa gerar a infraestrutura necessária ao estudo de primitivas de comunicação e sincronização para sistemas de aplicação em tempo real, e ao desenvolvi mento de sistemas operacionais distribuídos.

Outro aspecto, a ser destacado, consiste na possibilidade de comparação de uso, em sistemas distribuídos, de conexão ponto a ponto e de conexão por barramento pa ralelo centralizado, dos pontos de vista de implementação, desempenho, resistência à

Cabe citar que o projeto está sendo desenvolvido com o uso de microprocessadores, onde estiver sendo possível, aproveitando-se as vantagens de custo e flexibilidade desses componentes, mas tomando-se todo o cuidado para que a estruturação do projeto permita a utilização de estruturas mais rápidas e mais adequadas a determinadas exigências de aplicações especiais.

O uso de barramento centralizado já vem sendo explorado, especialmente com o esquema de contenção de barramento, conforme publicação recente [2].

2. ARQUITETURA DO SISTEMA

O sistema constitui-se de um subsistema de comunicação composto por vários processadores de comunicação (PC), interligados através de um barramento paralelo disposto num "backplane", e de um conjunto de processadores de trabalho, denominados processadores operadores (PO), conforme a figura 1.

O subsistema de comunicação será montado em um gabinete, contendo uma fonte confiável ("no break"), o barramento paralelo com 72 linhas dispostas no painel traseiro e 64 placas de circuito ligadas ao barramento, cada uma correspondendo a um processador de comunicação. Este conjunto poderá ser colocado dentro de uma sala bem protegida, de onde sairão todas as ligações para os processadores operadores remotos e, eventualmente, para dispositivos periféricos, sensores e atuadores que estiverem ligados a processadores de comunicação expandidos. Cabe salientar que um processador de comunicação expandido conte

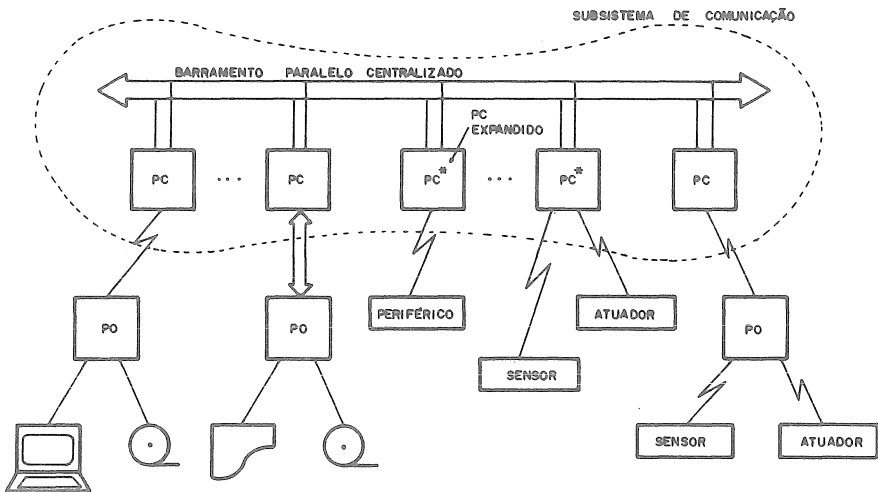


Figura 1 - Configuração Genérica do Sistema Distribuído

rã, além das funções de um processador de comunicação normal, funções simples de controle de alguns periféricos, tais como terminal de vídeo e impressora, e funções de leitura de dispositivos sensores e acionamento de dispositivos atuadores. A limitação, a 64 processadores de comunicação num barramento, deve-se às restrições de componentes eletrônicos "tri-state" e à necessidade de manter-se o comprimento do barramento a dimensões reduzidas. A centralização do subsistema de comunicação, além de facilitar a manutenção, contribui bastante para o aumento da segurança do sistema, uma vez que os maiores riscos de ruptura, incidindo com maior frequência nos cabos externos, comprometerão somente a comunicação individual de cada processador operador com o subsistema, sem afetar sensivelmente o funcionamento do conjunto.

A comunicação entre processadores operadores ocorrerá através dos processadores de comunicação correspondentes e do barramento paralelo centralizado de alta velocidade, que é o meio compartilhado de comunicação. A alta taxa de transferência de informação entre processadores de comunicação será obtida muito mais pela existência de várias linhas paralelas, do que pela velocidade de cada linha. Para permitir que todos os processadores de comunicação utilizem o barramento, é necessário que cada um utilize-o no menor tempo possível, sem monopolizá-lo, e para isso é necessário a existência de "buffers" de emissão e "buffers" de recepção em cada processador de comunicação. Esses "buffers" (figura 2) servirão para a montagem da informação, a ser transferida, ou para a sua desmontagem para posterior utilização, ou retransmissão, em taxas menores, para um processador operador específico.

Um processador operador emissor fará a transferência de informação para um processador operador receptor da seguinte forma. Primeiro, o processador operador emissor transferirá a informação para o processador de comunicação origem, a ele ligado, por bit ou por byte, dependendo da ligação ser serial ou paralela. Esse processador de comunicação irá coletando as partes da informação, conforme a sua chegada, e montando-as convenientemente no "buffer" de emissão. Quando a montagem terminar, o processador de comunicação origem acionará o seu circuito de disputa do barramento e, tão logo ganhar acesso, verificará se o processador de comunicação destino está pronto para receber a informação ("buffer" de recepção vazio). No caso do processador de comunicação destino não encontrar-se pronto, o processador de comunicação origem desistirá do barramento e aguardará um tempo para nova tentativa. No caso do processador de comunicação destino estar pronto, então haverá a transferência da informação, do "buffer" de emissão do processador de comunicação origem, para o "buffer" de recepção do processador de comunicação destino (figura 2). Este último processador encarregar-se-á de desmontar a informação recebida e de transferi-la ao processador operador, a ele ligado, num padrão de transferência compatível com a ligação.

O tempo máximo de comunicação entre dois processadores operadores, no pior caso, ou seja, quando n processadores estiverem competindo pelo barramento, e esses dois

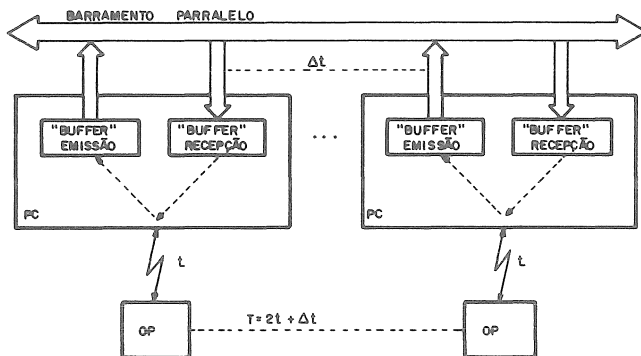


Figura 2 - Comunicação entre 2 Processadores Operadores

forem atendidos por último, deverá ser a soma: tempo gasto na transmissão PO-PC e no preenchimento do "buffer" de emissão mais o tempo de esvaziamento do "buffer" receptor e respectiva transmissão PC-PO mais n vezes o tempo de transferência de um "buffer" pelo barramento, correspondendo ao tempo de espera mais o tempo de transferência. Como o tempo de transferência de um "buffer" pelo barramento deverá ser da ordem de centenas ou milhares de vezes menor do que o tempo gasto na transmissão PO-PC e no preenchimento do "buffer", ou vice-versa, ele terá pouca influência no tempo total de comunicação entre processadores operadores, com a ressalva de que os processadores de comunicação sempre deverão estar prontos. Sabendo-se que esta ressalva, muitas vezes, não corresponde a realidade, a garantia de tempo bem definido de comunicação desaparecerá, a menos que se garanta, a nível de modelamento do sistema, que nunca ocorrerá tentativa de transferência de "buffer" para um processador de comunicação ocupado, ou se consiga formas de contornar o problema. Uma dessas formas consiste na introdução, em cada processador de comunicação, de "buffers" adicionais em número suficiente para armazenar todas as transferências, que possam ocorrer dentro de curtos períodos de tempo, para serem atendidas de acordo com uma ordem de prioridade. Nessa situação, o tempo máximo de comunicação entre 2 processadores operadores poderá variar, dependendo das prioridades envolvidas, mas dentro de valores perfeitamente previsíveis.

É interessante notar, em toda essa análise, que é quase impossível que um processador de comunicação monopolize o uso do barramento ou impeça a transferência de informação de processadores de comunicação de baixa prioridade, ao tentar transferir sequências de informação, pois o preenchimento de "buffer", que é uma atividade bem mais demorada do que a transferência pelo barramento, propiciará o intervalo de tempo necessário para as outras transferências pendentes.

O barramento paralelo foi concebido inicialmente para ter 72 linhas, sendo 40 para transferência de informação (pacotes de mensagem), 4 para atuar como reserva das linhas de informação, 8 para endereço origem, 8 para endereço destino, e as 12 restantes para controle (figura 3).

As linhas de controle, além de indicar o estado do barramento, servirão também: para implementar o protocolo necessário à transferência de informações; para alterar o padrão de transferência; e para realizar testes e ajustar o subsistema de comunicação. O estado do barramento poderá ser verificado através das linhas *desocupado* e *er*

Nº DE LINHAS	NOME	
32	dados	- transferir a informação útil de um pacote.
8	cabeçalho	- transferir o cabeçalho de um pacote.
4	reserva	- servir de linhas sobressalentes para as linhas de transferência de um pacote.
8	endereço origem	- permitir a disputa pelo acesso ao barramento e indicar a origem do pacote a ser transferido.
8	endereço destino	- indicar qual processador de comunicação irá receber o pacote.
1	desocupado	- indicar se o barramento está ocioso.
1	clock-in	- disparar o carregamento do pacote no processador de comunicação destino.
1	reset	- colocar todos os processadores de comunicação na condição inicial.
1	desconectar	- isolar processador de comunicação com falha
1	erro-tensão	- indicar tensão no barramento fora de especificação.
1	limpo-para envio	- indicar que o processador de comunicação destino está pronto para receber o pacote.
1	teste	- solicitar a execução das rotinas de teste.
1	modo-transm	- indicar transferência de 8 ou de 32 bits.
2	reformul.	- indicar qual dos 4 bytes das linhas de dados estará sendo usado.
1	não-autor.	- indicar que o processador de comunicação destino não está pronto para recepção.
1	paridade	- indicar a paridade do pacote.
72	barramento	- total das linhas

Figura 3 - Tabela do Detalhamento do Barramento Paralelo

no-tensão. O protocolo de transferência utilizará as linhas *limpo-para-envio*, *não-autorizo* e *clock-in*. O padrão de transferência da parte útil do pacote poderá ser de 32 bits, com uso pleno do barramento, ou de 8 bits, quando houver necessidade de transmissão mais rápida de pacotes menores, como acontece com pacotes de controle. O uso do barramento no padrão de 8 bits também será necessário quando aparecerem algumas falhas não contornáveis em transmissão com 32 bits. A alteração de padrão de transferência utilizará as linhas *modo-transmissão* e *reformulação*. A realização de testes e ajustamento do sistema de comunicação será feita através das linhas *reset*, *teste* e *desconectar*.

Estima-se que o barramento poderá funcionar a uma taxa da ordem de 1 a 10 Mega pacotes/segundo, levando em conta as atividades de requisição do barramento, transferência propriamente dita, e liberação do barramento, em cada transferência. Como cada pacote de informação possui 40 bits (8 para o cabeçalho e 32 para a informação útil), o barramento paralelo proposto corresponderia a um barramento serial funcionando na faixa de 40 a 400 Mega bits/segundo. Desta maneira, pela ampliação do número de linhas de dados, poder-se-ia chegar, com facilidade, num barramento paralelo que trabalhasse a uma taxa correspondente a Giga bits/segundo em um barramento serial.

4. PROJETO DO PROCESSADOR DE COMUNICAÇÃO

4.1. DESCRIÇÃO DO PROCESSADOR DE COMUNICAÇÃO

O processador de comunicação é o elemento responsável pela conexão do usuário (processador operador) à rede. Cabe a ele controlar todo o fluxo de informação entre o processador operador e vice-versa, e, dentre as suas diversas tarefas, destacam-se as seguintes:

- comunicar-se com o processador operador, recebendo e enviando informação;
- manipular os "buffers", colocando e retirando informação;
- controlar o acesso ao barramento;
- estabelecer a comunicação com outro processador de comunicação, cuidando do protocolo de transferência de informação.

O processador de comunicação é formado por 3 módulos: módulo de interface com o barramento; módulo de interface com o processador operador; e módulo de controle das interfaces (figura 4).

O módulo de interface com o barramento consiste dos seguintes elementos:

- circuito de acesso ao barramento;
- sincronizadores;
- "buffers" de emissão e de recepção.

O circuito de acesso ao barramento é quem se encarregará da disputa pelo barramento. Ele decidirá, através de prioridade, se o processador de comunicação, no qual está incluído, terá conseguido ou não o direito acessar o barramento naquele instante. A prioridade de cada processador de comunicação será baseada no seu endereço origem dâfe

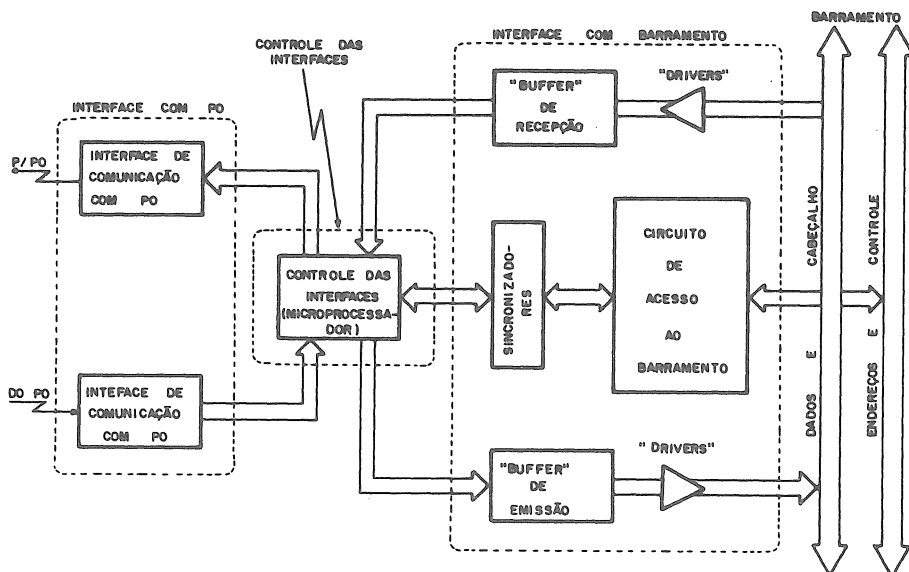


Figura 4 - Configuração do Processador de Comunicação

rente de todos os outros, o que garante que a disputa terá sempre um único vencedor.

Os sincronizadores são elementos que retardam ou armazenam sinais necessários à sequencialização das ações de disputa, uso, e liberação do barramento.

Os "buffers" são registradores de dados que colocam ou retiram informação do barramento, em determinados instantes definidos pelos sincronizadores. Os "drives", que estão entre os "buffers" e o barramento, e que também são usados em outras linhas, são circuitos "tri-state", que funcionam como isoladores nesse caso, ou que implementam lógica de fiação em outros casos.

O módulo de *interface com processador operador* consiste de interfaces padrão de entrada e saída de microprocessadores (8251, SIO, etc.) ou de circuitos especiais. A conexão com o processador operador pode assumir muitas variações: as ligações podem ser feitas com o uso de fios ou fibras óticas, de maneira serial ou paralela, com velocidade baixa ou alta, e, quando não se utilizar processador operador, devido a expansão do processador de comunicação, essas linhas poderão estar ligadas a periféricos, sensores, atuadores e outros elementos.

O módulo de *controle das interfaces* consiste de circuitos de microprocessador ou de um hardware especial equivalente. Sua função se resume na coordenação dos trabalhos das interfaces, e na transferência de informação entre elas.

Quando a conexão com processador operador, ou elemento correspondente, exigir

velocidades muito altas, não suportadas por microprocessadores e suas interfaces, os $\overline{m\bar{o}}$ dulos: controle das interfaces, e interface com P0, deverão ser construídos com técnicas e componentes adequados às exigências.

4.2. CIRCUITO DE ACESSO AO BARRAMENTO

O circuito de acesso ao barramento (figura 5), existente em cada processador de comunicação, é o elemento que permitirá que o processador de maior prioridade, dentre aqueles que estiverem concorrendo ao acesso pelo barramento, saia vencedor e ganhe o direito de usar o barramento por um ciclo. Este circuito é um aperfeiçoamento daqueles apresentados em [4] e [9].

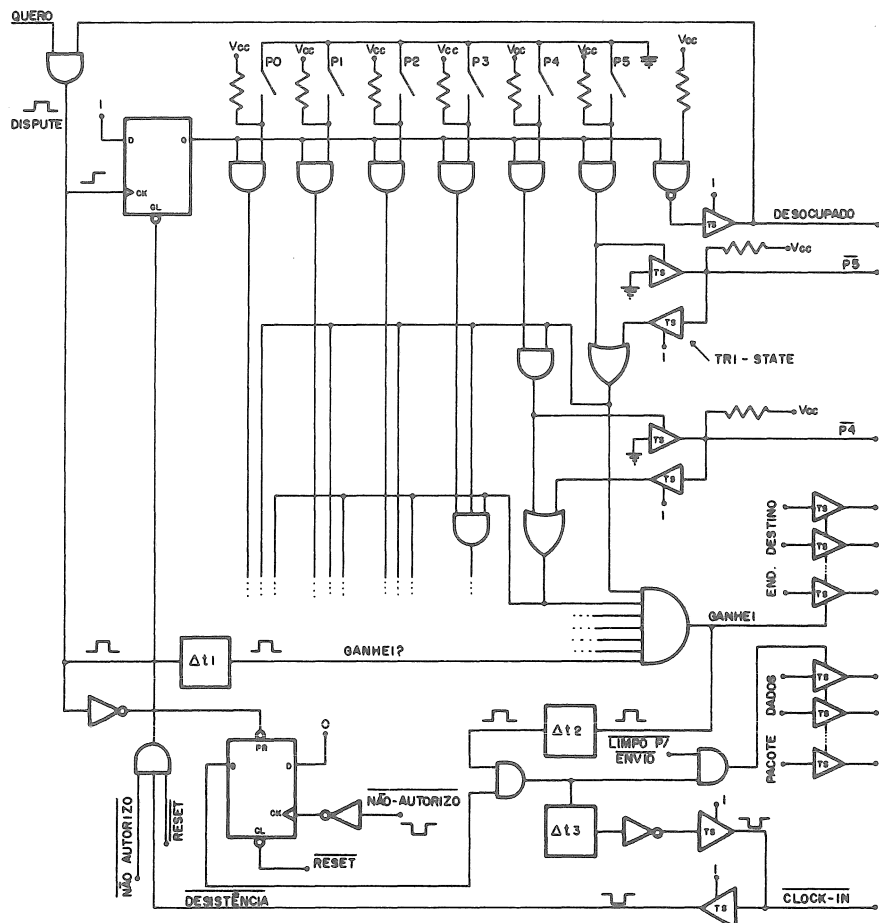


Figura 5 - Circuito de Acesso ao Barramento

O circuito está projetado para 64 processadores de comunicação e funciona da seguinte maneira. Quando um processador de comunicação desejar o domínio do barramento, ele deverá manifestar-se através do sinal *quero*. Se o barramento estiver ocupado (*desocupado* = 0), o pedido ficará pendente até o mesmo desocupar. No caso do barramento estar desocupado, ou ser desocupado (*desocupado* = 1) com pedido pendente, então todos os processadores de comunicação, que estiverem nestas condições, iniciarão simultaneamente uma disputa e inibirão a entrada de outros, indicando que o barramento estará ocupado (*desocupado* = 0). Aqueles, que conseguirem passar, colocarão seus endereços (prioridades) no barramento de endereço origem (arbitragem) e, através da lógica por fiação com componentes "tri-state", será feita a comparação simultânea dos bits de endereço (o maior endereço será prioritário). Os circuitos, que perderem a disputa, deverão esperar a liberação do barramento para participarem da nova disputa (o sinal *quero* de cada um permanecerá). O circuito, que ganhar a disputa, retirará seu sinal *quero* e colocará o endereço destino e o pacote no barramento. O processador de comunicação destino, identificando seu endereço, emitirá o sinal *limpo-para-envio* = 0 ou *não-autorizo* = 0 se o buffer de recepção estiver livre ou ocupado respectivamente. O sinal *não-autorizo* = 0 provocará a liberação do barramento, fazendo *desistência* = 0. O sinal *limpo-para-envio* = 0 permitirá a emissão do sinal *clock-in* = 0, que acionará o "buffer" de recepção destino para receber o pacote, e provocará a liberação do barramento para nova disputa.

4.3. EXPANSÃO DO BARRAMENTO E INTERCONEXÃO DE REDES

O problema de expansão do barramento pode ser resolvido com a interconexão de redes através de um elemento denominado conector ("gateway"). A função do conector é receber pacotes de um barramento e colocá-los em outro, com o devido endereçamento. Para isso pode-se adotar diversas estruturas, sendo uma delas a estrutura em árvore (figura 6).

O conector deverá entrar no lugar de um processador de comunicação e consumirá o endereço destino, ao receber um pacote. Para poder continuar sua *viagem*, o pacote deverá ter embutido, em si mesmo, os *passaportes* necessários para que cada conector possa encaminhá-lo adequadamente. Os passaportes serão os endereços dos outros conectores intermediários e do processador de comunicação destino, e deverão, desta forma, reduzir a parte útil do pacote. O fato do pacote, que circular entre barramentos, precisar passar por conectores e ter uma informação útil menor, faz com que a taxa de transmissão da informação sofra uma queda considerável.

5. ASPECTOS DE SOFTWARE

O software de um sistema distribuído envolve várias camadas: o software do processador de comunicação, o núcleo do sistema, abrangendo o subsistema de comunicação e parte dos processadores operadores, o sistema operacional distribuído, e o sistema aplicativo.

O funcionamento do processador de comunicação baseia-se na existência de *proce*

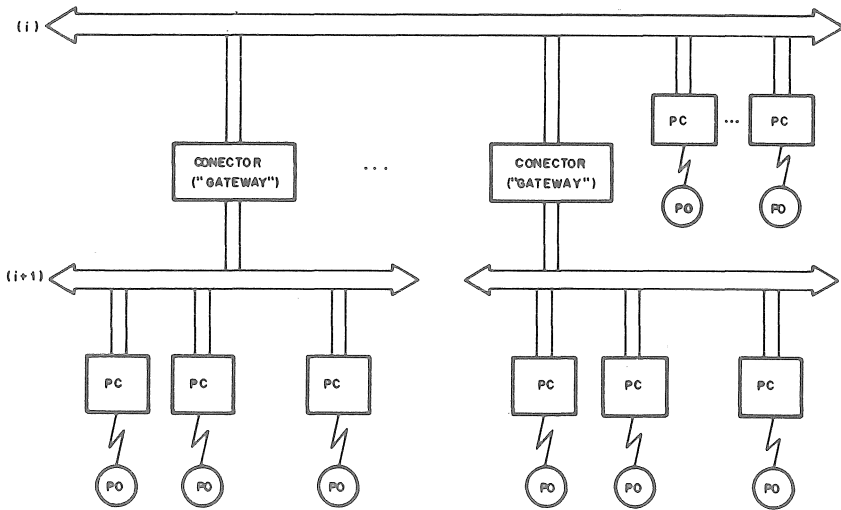


Figura 6 - Uma Forma de Expansão do Barramento (Interconexão de Redes)

dimentos que serão executados no microprocessador do módulo de controle das interfaces. Estes procedimentos ẽ que colocarão o processador de comunicação nas condições iniciais, e depois, farão a monitoração dos sinais de controle e do estado das interfaces, cuidando da transferência interna das informações, e do acionamento de seus circuitos.

A ampliação dos procedimentos do processador de comunicação com primitivas de sincronização e comunicação entre processos e outras primitivas básicas de interesse de sistemas distribuídos, implementará a parcela do núcleo, referente ao processador de comunicação, ficando a parcela complementar para ser implementada no processador operador. O núcleo do sistema distribuído será obtido pela reunião de todos os núcleos de cada par PC-PO[6]. E nessa fase que serão exploradas as questões de implementação de primitivas apropriadas ao controle de aplicações em tempo real, particularmente aquelas relativas à reconfiguração dinâmica de sistema e ao tratamento de exceções.

6. CONCLUSÕES

O projeto do subsistema de comunicação encontra-se detalhado a nível de circuito, e o software mínimo, necessário para o funcionamento do processador de comunicação, já está especificado.

Uma proposta do subsistema de comunicação, com 5 nós de comunicação, está sendo objeto de pedido de financiamento que, se for concedido, permitirá passarmos a fase de implementação e testes.

É interessante notar que, apesar da existência do barramento de alta velocidade, o sistema formado por "BUFFER" DO PC - BARRAMENTO-"BUFFER" DO PC constitui-se numa ligação frouxa, que corresponde a uma das características essenciais de sistemas distribuídos.

O barramento centralizado, como foi proposto, permitirá a exploração das vantagens de qualquer barramento, sem os inconvenientes da dispersão física que o tornaria vulnerável, e com a vantagem adicional de ser tolerante à falha em algumas situações.

A possibilidade de utilização de fibra ótica nas ligações PO-PC, e a centralização do subsistema de comunicação, que poderá ficar completamente isolado, tornam este tipo de sistema adequado para aplicações que exijam alta imunidade a ruídos.

O projeto permite muitas variações em todos os seus níveis: ao nível do hardware, pode-se sofisticar o barramento centralizado, o processador de comunicação e as interfaces de conexão com outras redes; ao nível do software básico pode-se conceber vários conjuntos de primitivas, apropriadas a aplicações diferentes, formando vários núcleos; da mesma maneira ao nível do sistema operacional e ao nível da aplicação, o sistema poderá ir de situações bastante simples até as mais complexas.

Um projeto de sistema distribuído, por natureza, é um projeto de fôlego, e esse, particularmente, por abranger a área de aplicações em tempo real, apresenta-nos um desafio a mais. É nossa intenção atuar até ao nível do sistema operacional, mas o projeto estará à disposição dos pesquisadores do Departamento e de fora, interessados em testar variações de hardware e software, e de continuar trabalhando em outros níveis.

7. BIBLIOGRAFIA

- [1] ABRAMS, M. et al. - *Computer Networks: A Tutorial*. IEEE Computer Society, N.Y., 1980.
- [2] ACAMPORA, A.S. & HLUCHYJ, M.G. - "A New Local Area Networks Architecture Using a Centralized Bus". *IEEE Communications Magazine*, august, 1984, p.12-21.
- [3] BRAYER, K. & LAFLEUR, V. - "A Testbed Approach to the Design of a Computer Communications Network". *Computer*, october, 1982, p. 14-23.
- [4] CORSO, D.D. & VERRUA, L. - "Contention Delay in Distributed Priority Networks". *The Euromicro Journal - Microprocessing and Microprogramming*, january, 1984, p.21-29.
- [5] DAVIES, D.W. et al. - *Computer Networks and Their Protocols*. John Wiley & Sons, N.Y., 1981.
- [6] KIRNER, C. - "Suporte para Desenvolvimento de Sistemas Distribuídos: Uma Implementação". *Anais do IV Congresso da Sociedade Brasileira de Computação - XI Seminário Integrado de Software e Hardware*, Viçosa, MG, 1984, p. 109-121.
- [7] LINDSAY, D.C. - "Local Area Networks: Bus and Ring Vs. Coincident Star". *Computer Communication Review*, july/october, 1982, p. 83-91.
- [8] MARQUES, E. - "Projeto de um Barramento Compacto para Sistemas Distribuídos". *Publicação interna*, Departamento de Computação e Estatística - UFSCar, 1984.

- [9] TEIXEIRA, C. & TREVELIN, L.C. - "Protocolo de Acesso e Superação de Falhas em Redes de Topologia Estrela Coincidente". *Anais do 2º Simpósio Brasileiro sobre Redes de Computadores*, Campina Grande, PB, 1984, p. 161 - 165.
- [10] THURBER, K.J. & FREEMAN, H.A., ed. - *Local Computer Networks: A Tutorial*. IEEE Computer Society, N.Y., 1981.