

# 4: SBRC

RECIFE - 24 A 26 DE MARÇO 86

---

IMPLEMENTAÇÃO DE UMA ESTAÇÃO DE VOZ NUM  
COMPUTADOR PESSOAL DE 16 BITS

Wagner Luiz Zucchi, EPUSP, FDTE

End.: Cidade Universitária "Armando Salles de Oliveira"

Av. Prof. Luciano Gualberto, Travessa 3 nº 158

Caixa Postal 11455 - CEP.: 05508 - São Paulo - SP

Wilson Vicente Ruggiero, PhD., SCOPUS Tecnologia S.A.

End.: Rua Bela Cintra, 881, 8º andar, Cerqueira César

CEP.: 01415 - São Paulo - SP

## RESUMO

Este trabalho apresenta alguns aspectos da implementação de uma estação de voz e de dados sobre um computador pessoal de 16 bits (NEXUS) que está sendo realizada na EPUSP.

A estação de voz é implementada numa placa de expansão do computador pessoal e inclui elementos de 'hardware' e de 'software' que são aqui discutidos.

## 1. INTRODUÇÃO

A integração de serviços de voz e de dados numa rede local de computadores vem chamando a atenção dos pesquisadores da área nos últimos anos, sendo que, no exterior, diversos equipamentos com essas características já são produzidos e comercializados (1).

O mercado para esse tipo de equipamento está, no exterior, em rápida expansão. A figura 1 resume os dados apresentados na referência citada.

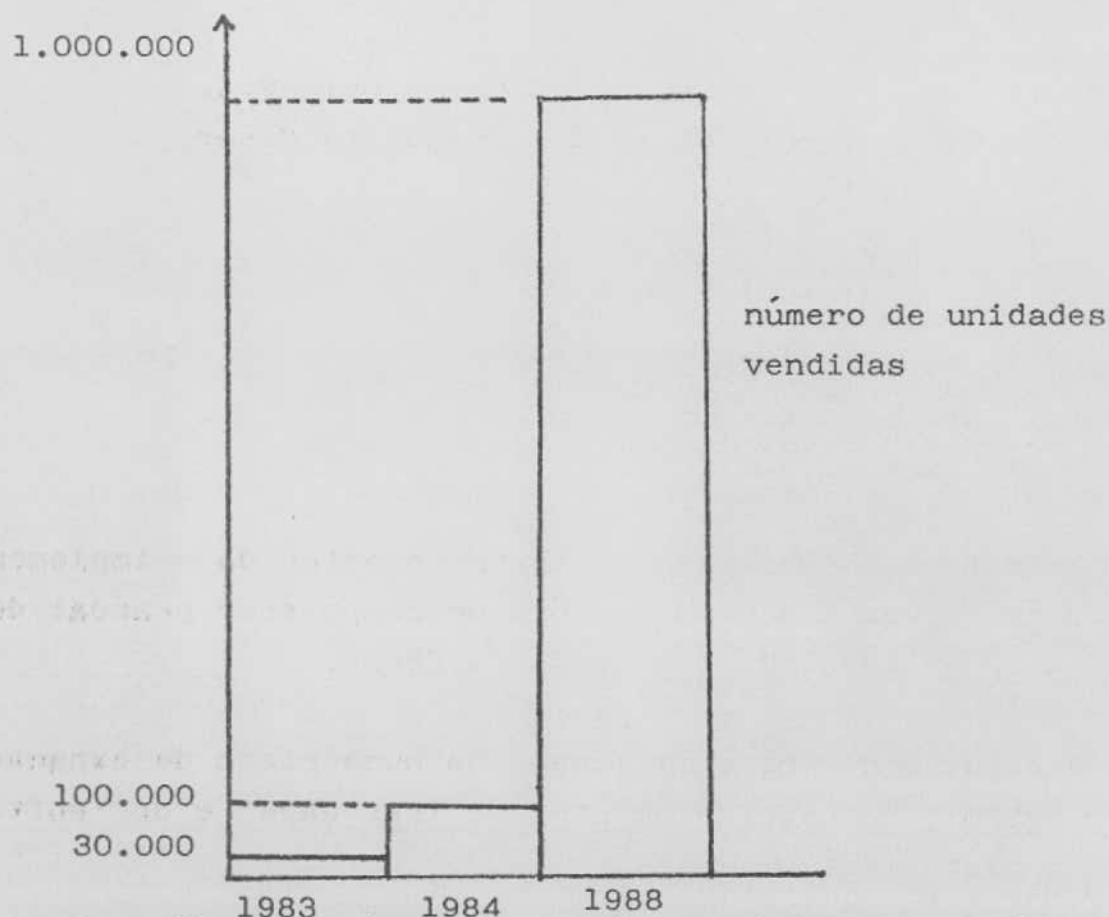


Figura 1. - Mercado atual e mercado previsto para equipamentos de voz e dados integrados (fonte: ref.1) (os dados referem-se aos EUA)

No Brasil, mantidas as proporções, é razoável esperar-se uma tendência de crescimento análoga nos próximos anos.

O equipamento de integração básico é um terminal ou microcomputador reunindo as funções de display/teclado alfanumérico e base de dados, com as funções de discagem, sinalização e comunicação vocal. A figura 2 mostra uma possível apresentação desse tipo de terminal:

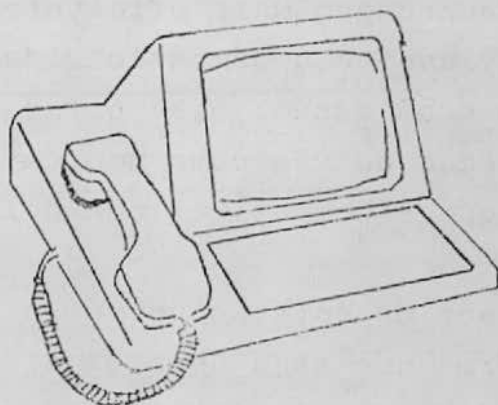


Figura 2.-Terminal de voz e dados integrados

Muitas vezes o terminal de voz e dados (TVD) reúne também funções de processamento local, tais como, processamento de palavras, edição de documentos, desenvolvimento de programas, etc.

Por essas razões, o TVD constitui um bloco básico de um sistema de integração de escritório. Para tal, os diversos terminais instalados numa empresa devem ser interligados por uma rede com capacidade de comutação de voz e dados, tal como um PABX integrado ou uma rede local. A segunda possibilidade é interessante por permitir que os dois tipos de tráfego sejam manipulados por um único tipo de comutação (comutação dinâmica ou de pacotes).

A possibilidade técnica de transmitir voz por pacotes já foi demonstrada na literatura especializada (ref. 2, 3, 4).

Também do ponto de vista econômico, a integração é interessante em redes geograficamente distribuídas (ref. 5). A mesma relação parece valer em redes locais, onde a integração de serviços reduz a quantidade de equipamentos e simplifica as instalações necessárias.

A nosso ver, a abordagem mais eficiente para que um TVD seja economicamente viável é a compatibilidade com os sistemas já existentes. Por essa razão, esse estudo concentra-se na implementação da estação de voz como uma extensão de um computador pessoal NEXUS, que é compatível com o IBM PC.

Também a interface da estação de voz com a rede local deve ser projetada de forma que essa seja o mais independente possível do tipo de método de acesso utilizado.

O esquema básico do terminal integrado é mostrado na figura 3:

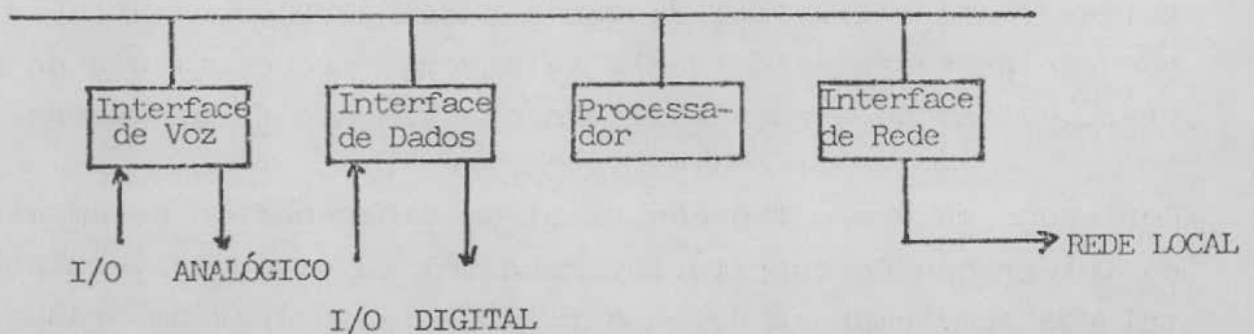


Figura 3. -Componentes da estação de voz

A independência entre o nível de tratamento de voz e os níveis de transmissão é conseguida definindo os serviços oferecidos e solicitados em cada nível. Em trabalhos anteriores (8) demonstrou-se que os métodos de acesso convencionais em redes locais oferecem um serviço de transmissão de pacotes altamente confiável, porém com atrasos variáveis e, em alguns casos, não limitados.

A estação de voz deve realizar diversas funções para adaptar as características do serviço de transmissão às exigências do serviço de voz. Essas funções e suas características também já foram discutidas anteriormente (9). A figura 4 apresenta um diagrama de blocos representando as diversas funções da estação de voz.

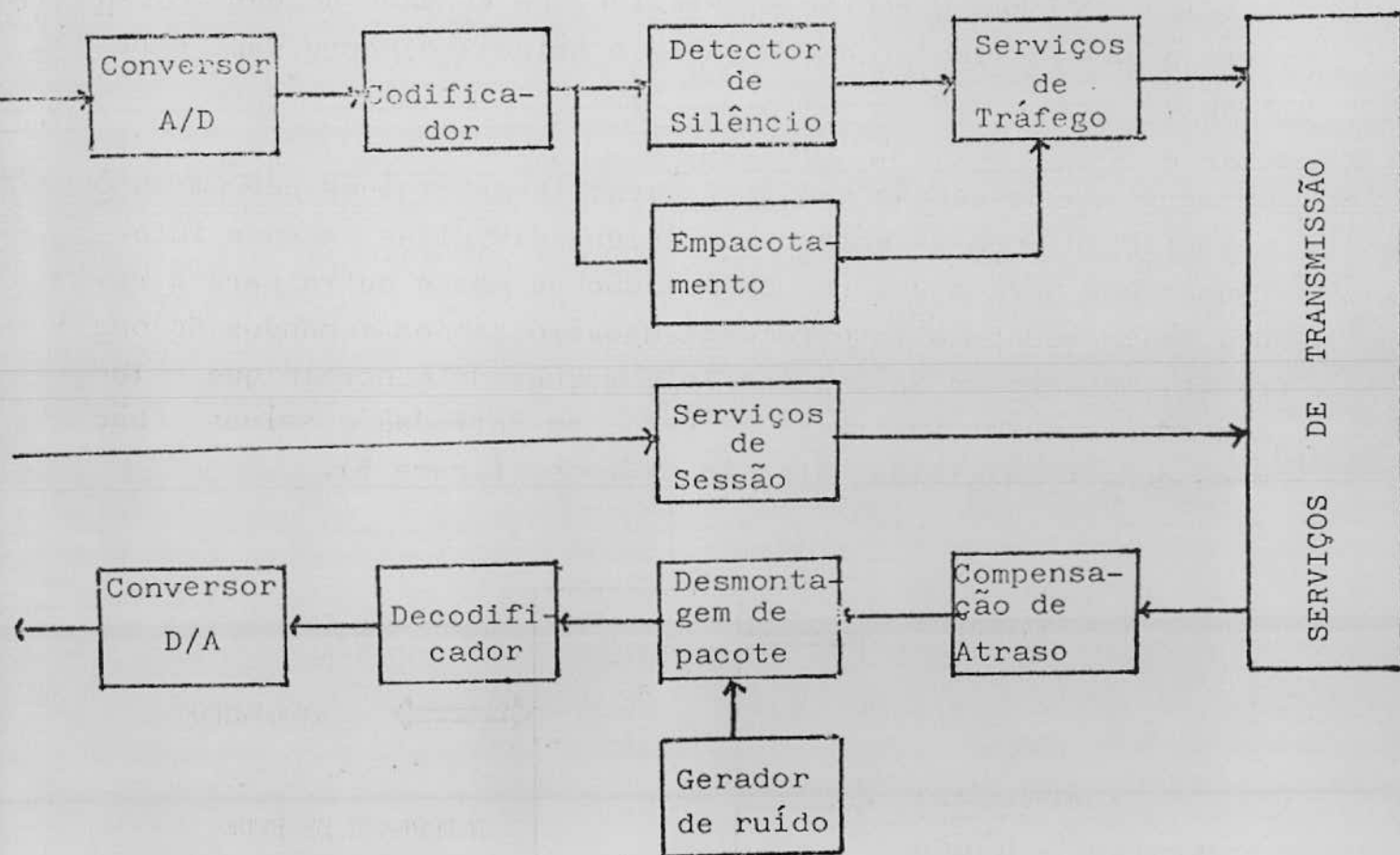


Figura 4. - Diagrama de funções da estação de voz

O objetivo desse trabalho é descrever as linhas gerais de uma implementação dessa estação de voz que está sendo realizada na EPUSP. Essa implementação visa, principalmente, a análise das características do serviço de voz e o dimensionamento de um sistema integrado.

## 2. SERVIÇOS DE SESSÃO

Os serviços de sessão dizem respeito ao estabelecimento, manutenção e finalização das conexões de voz. Esses serviços devem oferecer um mecanismo seguro e confiável para o usuário controlar suas conexões de voz. Seguro no sentido de que o usuário é sempre informado do estado da conexão solicitada e de possíveis falhas detectadas. Flexível no sentido de dar ao usuário a possibilidade de controlar o estabelecimento das conexões.

Os serviços de sessão são fisicamente constituídos por um autômata finito, com as entradas e saídas divididas em duas interfaces: uma para o usuário da estação de voz e outra para a rede. As entradas na interface de usuário são os comandos do operador, enquanto as saídas são informações de controle que são enviadas. Na interface de rede, as entradas e saídas são mensagens transmitidas através da rede (figura 5).



Figura 5. - Interfaces dos serviços de sessão



É importante observar que num terminal de voz e dados existe mais flexibilidade para implementação da interface de usuário do que num sistema telefônico tradicional, pois além dos sinais sonoros podem ser utilizadas informações visuais no display.

Os seguintes comandos foram definidos na interface entre o usuário e os serviços de sessão:

- a) Comando CONECTE: é originado pelo usuário da estação da voz para estabelecer uma conexão vocal. Esse comando é análogo a "tirar o fone do gancho" no sistema telefônico tradicional e pode ser implementado por um interruptor no TVD ou por uma tecla especial.
- b) Comando SELEÇÃO: é utilizado para o usuário informar à estação de voz o endereço da estação de voz remota com a qual a conexão deve ser estabelecida. Num TDV esse comando pode ser implementado pelo teclado, o que dispensa o disco do sistema tradicional. Além disso, o teclado permite correções do endereço sendo discado.
- c) Comando DESCONECTE: permite ao usuário encerrar uma conexão de voz, de forma unilateral, em qualquer estado da conexão. Esse comando corresponde a "colocar o fone no gancho" e pode ser implementado de forma simétrica ao comando CONECTE.

As informações que a estação de voz envia aos usuários são:

- a) Informação REDE-PRONTA: informa ao usuário a disponibilidade da rede para estabelecer uma conexão de voz e solicita o início do processo de seleção. Essa informação pode ser implementada como um sinal sonoro e visual.
- b) Informação SINAL-DE-PROGRESSO: é utilizada para informar ao usuário o estado de uma conexão de voz. Os estados possíveis são: 'ocupado', 'bloqueado', 'chamando' e 'não conectado'. Essa informação será implementada por uma mensagem vi-

sual no display para simplificar a interface analógica.

- c) Informação DESCONEXÃO: informa ao usuário que a sua conexão de voz foi desfeita, quer por falha na rede, quer por um comando do usuário remoto. A informação DESCONEXÃO não possui um equivalente exato no sistema telefônico tradicional.

Além desses comandos e informações, os serviços de sessão de diferentes estações de voz comunicam-se através de mensagens que estabelecem ou encerram as conexões através da rede e definem os parâmetros específicos de cada conexão. Três mensagens foram definidos com essa finalidade.

- a) PEDIDO-DE-CONEXÃO: essa mensagem é trocada pelas estações de voz para solicitar o estabelecimento de uma conexão. Ela possui os seguintes parâmetros:

- . identificador da estação chamante;
- . identificador da estação chamada;
- . taxa de codificação;
- . mínimo intervalo de silêncio;
- . tamanho dos pacotes.

- b) ESTADO-CHAMADO: essa mensagem permite que uma estação seja informada a respeito do estado de outra estação, com a qual a primeira pretende estabelecer uma conexão de voz. Os estados previstos são: DISPONÍVEL, OCUPADO e DESATIVO.

- c) INDICAÇÃO DE DESCONEXÃO: essa mensagem informa a estação de voz que o usuário remoto encerrou a conexão vocal. Dada a natureza do serviço de voz, uma conexão é forçosamente encerrada quando um dos usuários desconecta e as mensagens pendentes são perdidas.

Além das mensagens acima descritas a estação de voz deve implementar também algumas temporizações para evitar 'deadlocks'. Na atual implementação duas temporizações foram previstas:



- a) T1: protege a estação de voz quando o usuário local não atende a uma chamada remota. Quando esse temporizador se esgota, a estação de voz aborta a tentativa de estabelecimento de conexão.
- b) T2: protege a estação de voz quando o usuário remoto não responde a uma INDICAÇÃO-DE-DESCONEXÃO. Quando esse temporizador se esgota, a conexão é encerrada unilateralmente.

A figura 6 apresenta o diagrama de estados dos serviços de sesção. Dada a alta confiabilidade oferecida pelos serviços de rede não foram previstas retransmissões de mensagens.

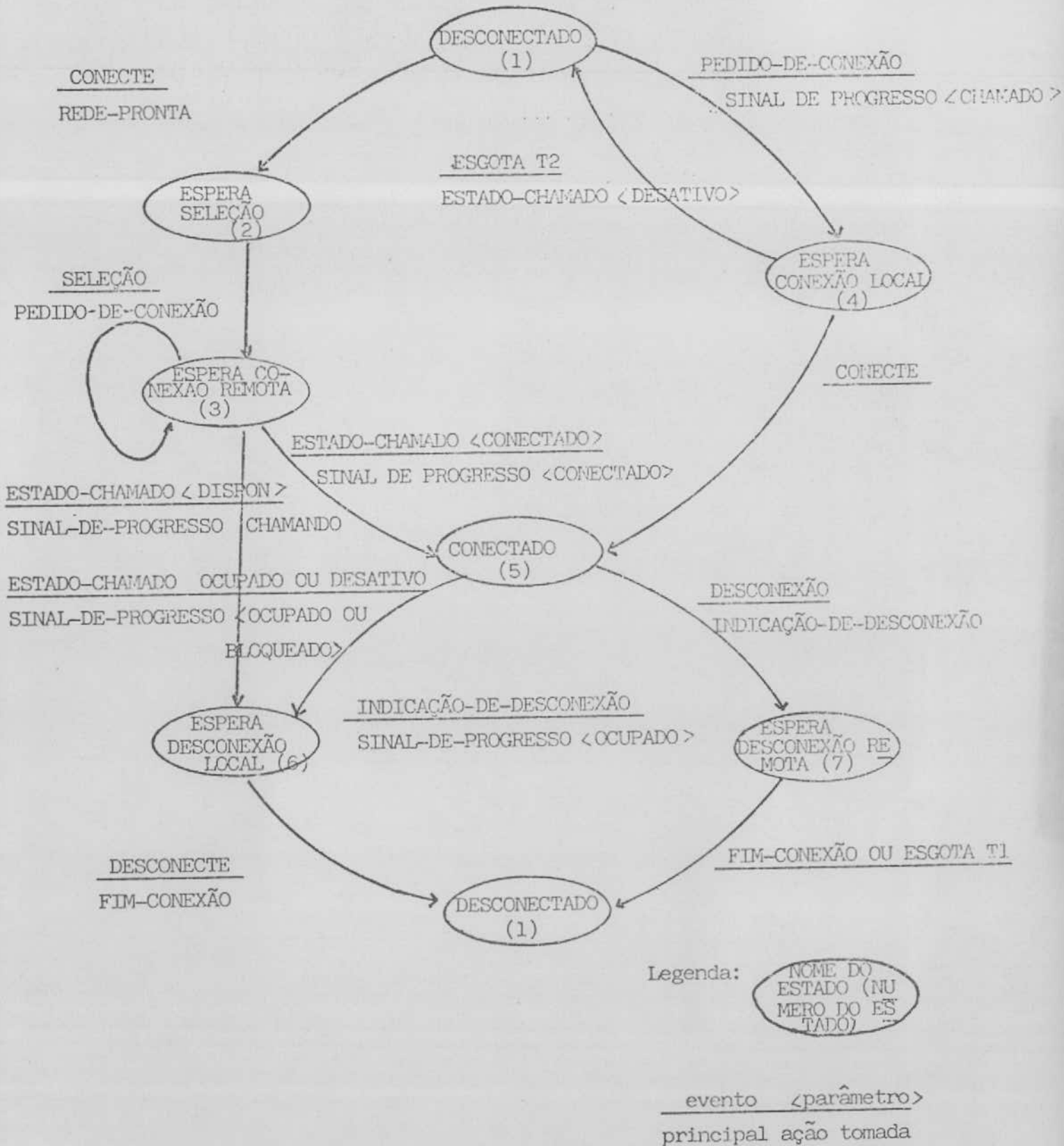


Figura 6. - Diagrama de estados dos serviços de sessão

### 3. FUNÇÕES DA ESTAÇÃO DE VOZ IMPLEMENTADAS POR 'HARDWARE'

Algumas das funções da estação de voz foram implementadas por 'hardware' devido às suas características de tempo real. Entre estas incluem-se:

- . A conversão analógico/digital e a codificação;
- . a decodificação e a conversão digital/analógica;
- . a detecção de silêncio;
- . o empacotamento/desmontagem das mensagens.

A seguir são descritas a implementação dessas funções.

#### a) Conversão analógico/digital e codificação

Essa função é implementada por um CODEC PCM integrado com saída paralela (National TP 3051). Este componente já possui integrados os filtros necessários e os circuitos de compensação (lei -  $\mu$ ), além disso, o seu modo de operação é controlado por software.

#### b) Decodificação e conversão digital/analógica

O mesmo circuito integrado, acima descrito, realiza as funções de decodificação PCM e conversão digital/analógica, oferecendo uma potência de 7,2 dbm para uma carga de 600  $\Omega$  (faixa dinâmica na saída:  $\pm 2,8V$ ).

#### c) Detector de silêncio

A detecção de silêncio em conversações telefônicas não é uma tarefa trivial, devido a presença de ruídos ambientais com nível de energia médio igual ou superior a alguns fonemas da voz humana. Por causa disso, uma simples medida de energia nos enlaces telefônicos é insuficiente para garantir uma detecção de silêncio eficiente. Se o limiar de energia adotado for bai-

xe, o ruído será frequentemente confundido com a voz e o efeito desejado de compressão de voz será pequeno. Por outro lado, se for adotado um limiar de energia elevado haverá frequentes cortes na voz, o que comprometerá a qualidade da comunicação.

A figura 7 (ref.6) apresenta um exemplo dessas dificuldades. A partir do ponto indicado por "B" nota-se uma variação significativa do nível de energia, que poderia ser confundida com o início da palavra. Uma análise mais detalhada mostra, porém, que o início real da palavra ocorre no ponto indicado pela letra "A". No exemplo mostrado a palavra pronunciada é "four", cujo início sibilante e não vocálico se confunde com o ruído ambiente, dificultando a detecção.

Estudos do comportamento espectral da voz humana (ref.7), mostram que a energia dos sons não vocálicos é concentrada nas regiões de alta frequência do espectro da voz. Por essa razão, os sons não vocálicos podem ser detectados por uma medida da frequência do sinal, ou da taxa de cruzamento por zero.

O esquema básico do mecanismo de detecção de silêncio é apresentado na figura 8.

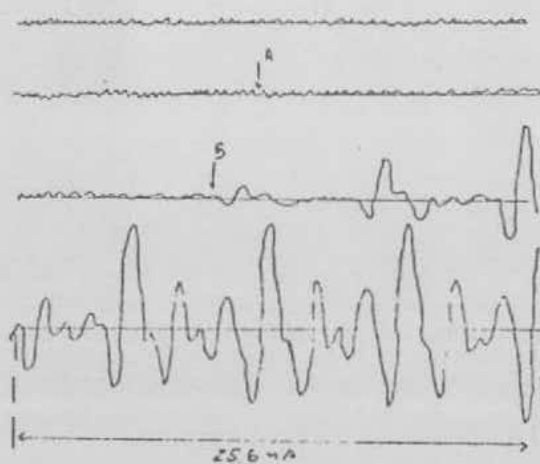


Figura 7. - Espectro temporal do início da palavra "four"

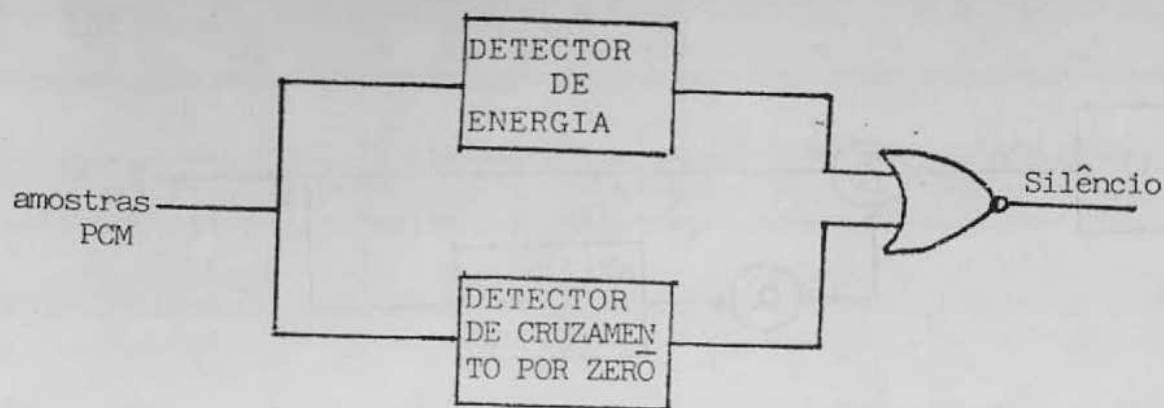


Figura 8. - Detector de Silêncio

O detector de energia possui seu limiar fixado num nível elevado de tal forma que apenas sons vocálicos são detectados. O detector de cruzamento por zero é ajustado de forma a identificar sinais cuja composição espectral seja diferente da do ruído. Os sinais gerados por ambos os detectores são considerados para produzir o sinal de silêncio.

O detector de energia consiste num filtro digital passa-baixas cuja transformada-z da resposta ao impulso é uma função racional. Na implementação aqui discutida, o filtro utilizado possui resposta ao impulso.

$$h(n) = \begin{cases} a^n & n \geq 0 \\ 0 & n < 0 \end{cases} \quad 0 < a < 1$$

A transformada-z dessa seqüência é:

$$H(z) = \frac{1}{1 - az^{-1}}$$

Esse filtro pode ser implementado de acordo com o diagrama mostrado na figura 9.

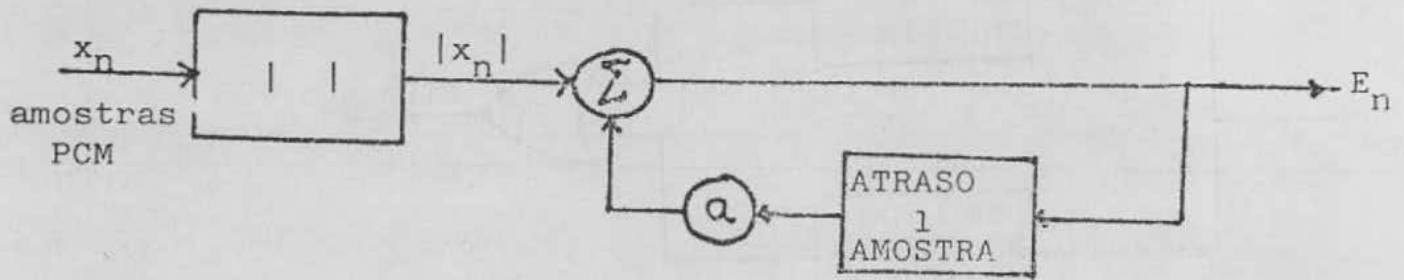


Figura 9.-Detector de energia

O detector de taxa de cruzamento por zero consiste num circuito que conta as trocas de polaridade do sinal PCM num dado intervalo de tempo. O diagrama de blocos do detector utilizado está mostrado na figura 10.

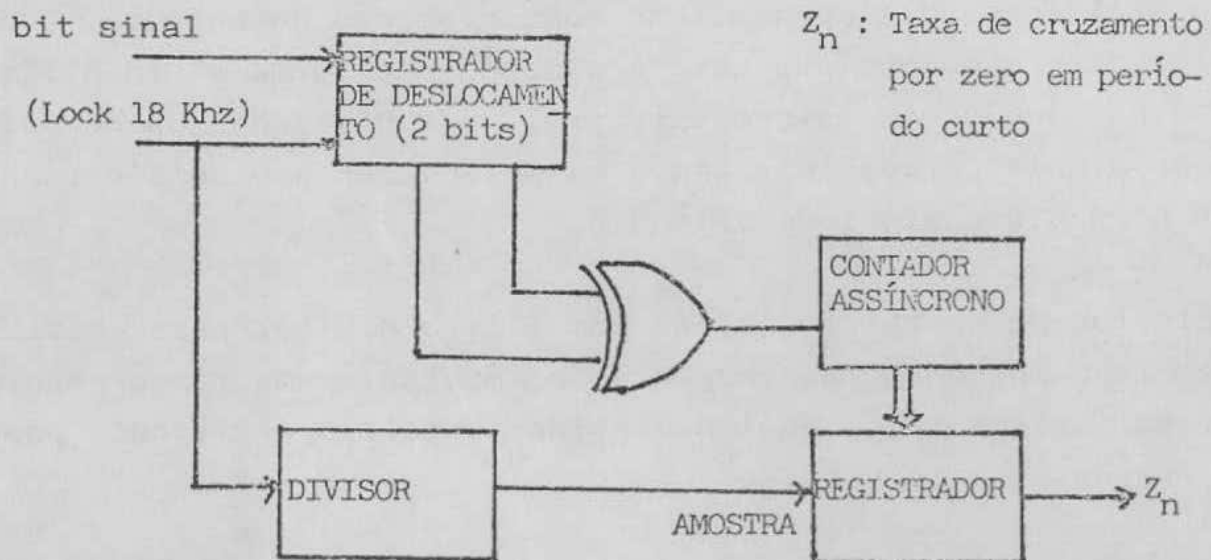


Figura 10.-Detector de taxa de cruzamento por zero.

O funcionamento do detector de taxa de cruzamento por zero é afetado pela presença de nível DC e de zumbido de 60 Hz no sinal analógico. Por isso, esse sinal deve ser cuidadosamente filtrado antes da conversão análogo/digital.



Os sinais gerados pelos detectores de energia e de taxa de cruzamento por zero são utilidades para obtenção da informação de silêncio. A CPU não transmite pacotes na conexão de voz enquanto, o sinal de silêncio está ativo. A figura 11 apresenta o detector de silêncio em mais detalhes:

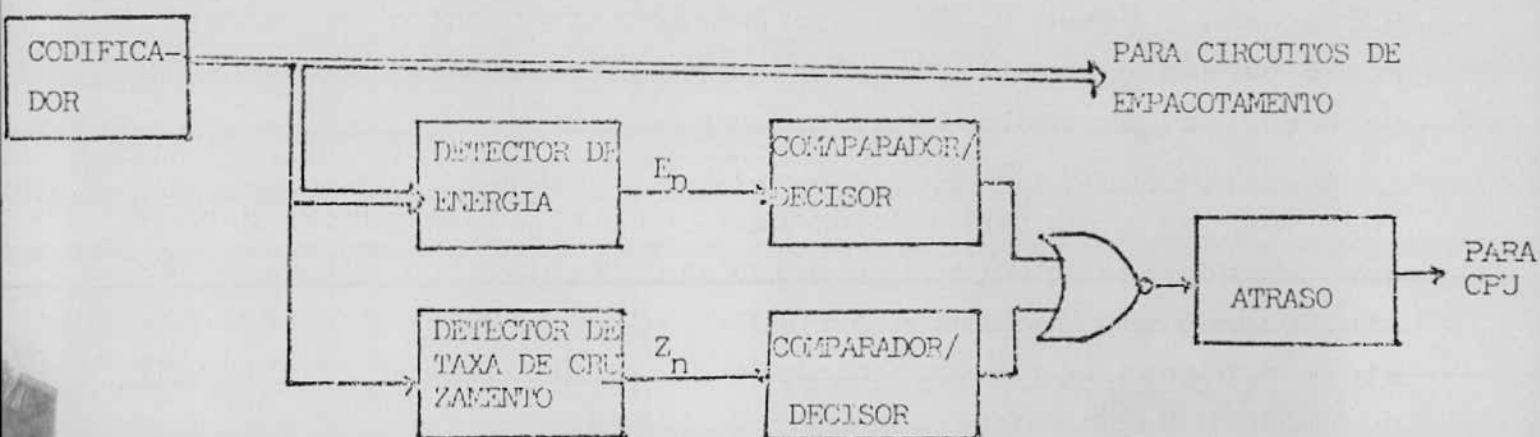


Figura 11: Diagrama do detector de silêncio

O bloco 'ATRASO' que aparece na figura 11 corresponde a um circuito mono-estável utilizado para evitar a detecção de silêncio em intervalos muito curtos do sinal de voz.

#### d) Empacotamento

A função de formação dos pacotes de voz deve ser totalmente implementada por 'hardware', pois o intervalo de tempo entre amostras sucessivas (125  $\mu$ s em cada sentido) é muito pequeno para permitir o tratamento por 'software' concorrendo com a transmissão de dados.

Fazendo a montagem dos pacotes na estação de voz, a CPU será interrompida somente quando um pacote estiver pronto, o que a descarrega significativamente.

A montagem dos pacotes é feita na memória da estação de voz. Essa memória é mapeada na memória de sistema, podendo seus en-

dereços serem configurados por 'dip-switches'.

A memória da estação de voz foi dividida em dois blocos de 2 K bytes cada (expansível até 4 K bytes). Esses dois blocos são acessíveis pela CPU e pelo CODEC através dos mesmos endereços, pois, num dado instante, a CPU acessa o bloco 1 enquanto o CODEC acessa o bloco 2. Quando o pacote está pronto, os acessos aos blocos são invertidos. A figura 12 esquematiza o mecanismo de acesso à memória.

Toda vez que a montagem de um novo pacote é iniciada, a geração de endereço aponta o primeiro endereço disponível na memória. Quando uma amostra é produzida ela é armazenada nessa posição e o endereço é incrementado. Esse processo é repetido até que todo o pacote esteja montado.

O final da montagem de um pacote também é verificado por 'hardware'. Quando essa condição é detectada, um sinal de interrupção é gerado para a CPU, que reconhece a interrupção e inverte o sinal de seleção.

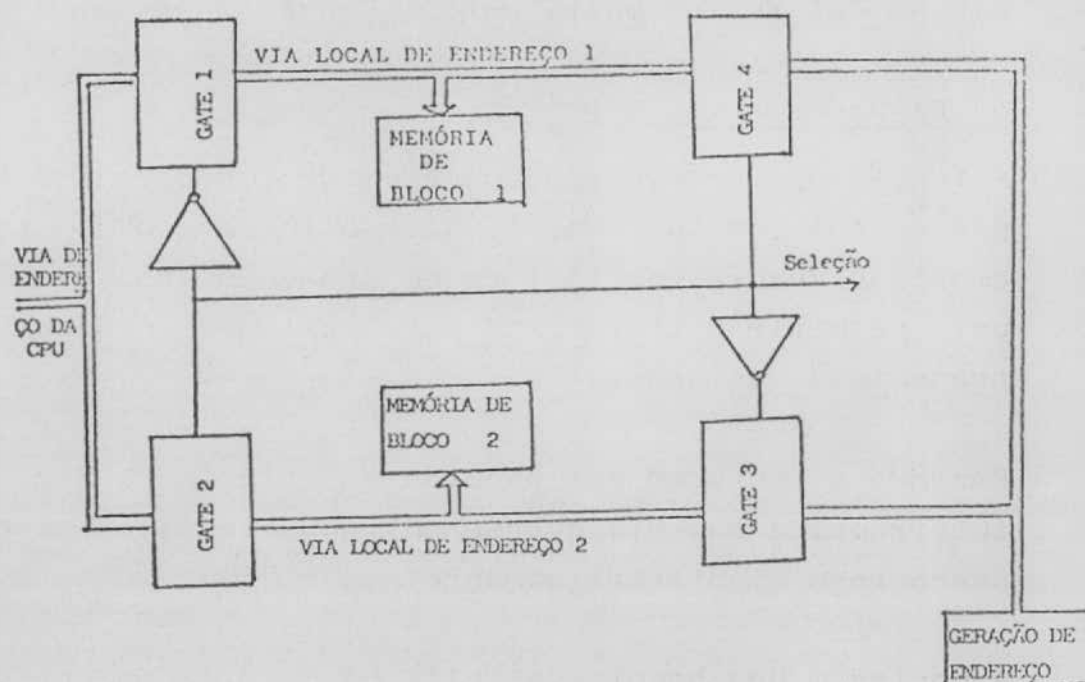


Figura 12.-Vias de endereço na estação de voz

O acesso à via de dados da memória é análogo ao acesso à via de endereço, porém são utilizadas portas bidirecionais, pois, os dados podem ser lidos ou escritos em ambos os blocos de memória.

A recepção e desmontagem de pacotes é feita por 'hardware' simultaneamente com a transmissão. Cada bloco de memória contém dois pacotes: um sendo montado para transmissão e outro sendo desmontado para recepção. O tamanho dos pacotes é configurável por 'dip-switches', sendo que os pacotes ocupam os endereços mais altos em cada bloco de memória. A figura 13 ilustra a utilização de um dos blocos da memória.

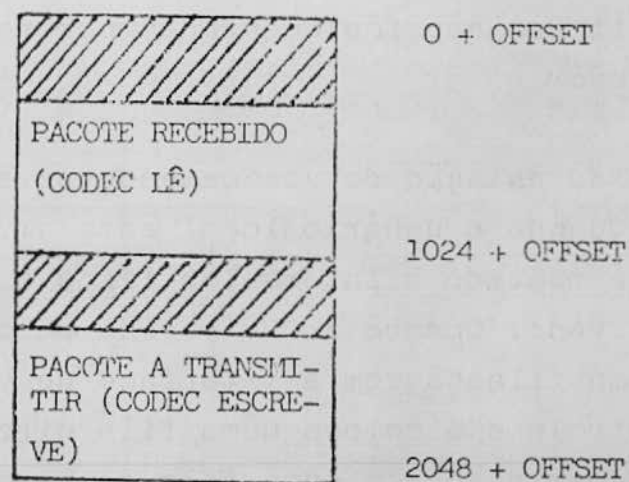


Figura 13. - Mapeamento da memória

Os sinais de controle para escrita e leitura na memória são gerados na própria estação de voz. Isso implica numa dificuldade adicional pois esses sinais devem ser sincronizados com o CODEC.

A operação do CODEC é controlada por um sinal de "chip select". Quando esse sinal desce (transição "1" para "0") e o 'clock' está baixo, o CODEC lê os dados de memória durante o próximo meio ciclo positivo do 'clock' (1024 Mhz). Se o 'clock' está alto quando o sinal 'chip select' desce, o CODEC coloca os seus

dados na via na próxima transição de descida do 'clock'.

A estação de voz contém todos os circuitos para sincronização das operações de leitura/escrita na memória com a operação do CODEC.

Além disso, a interface de voz também contém uma lógica de seleção de portas para que o CODEC possa ser programado pela CPU.

#### 4. FUNÇÕES IMPLEMENTADAS POR 'SOFTWARE'

As funções da estação de voz que não possuem requisitos de tempo real críticos são implementadas por software na CPU do micro computador.

A operação da estação de voz depende do estado dos usuários em conexão. Quando o usuário local está em conversação, um pacote de voz é montado a intervalos fixos e o sinal de silêncio fica desativado. Quando a CPU recebe um pacote ela o retira da memória compartilhada com a interface de voz, completa os campos de controle e o coloca numa fila circular de pacotes à espera de transmissão. Se essa fila se esgota, o último pacote recebido é colocado sobre o pacote mais antigo, destruindo-o.

O segundo estado da estação de voz se caracteriza pelo usuário local em silêncio e pela ativação do respectivo sinal. Toda vez que a CPU recebe um pacote da interface de voz ela examina o sinal de silêncio e descarta o pacote, se ele estiver ativado. Quando o sinal de silêncio é desativado a estação de voz volta a transmitir pacotes, sendo que o primeiro pacote transmitido possui um bit especial, no seu cabeçalho, ativado para indicar a interrupção.

A parte receptora da estação de voz também opera com dois estados. Quando o usuário remoto está falando, a estação recebe pacotes com atrasos variáveis e os coloca numa fila para a inter

face de voz. Toda vez que a interface apanha um pacote na área de memória compartilhada, o pacote mais antigo é retirado dessa fila e colocado na memória compartilhada.

Se o usuário remoto fica em silêncio, a estação de voz deixa de receber pacotes e a fila de recepção se esvazia. Quando isso acontece, a estação gera pacotes com ruído para a interface de voz. A figura 14 ilustra o funcionamento da estação de voz.

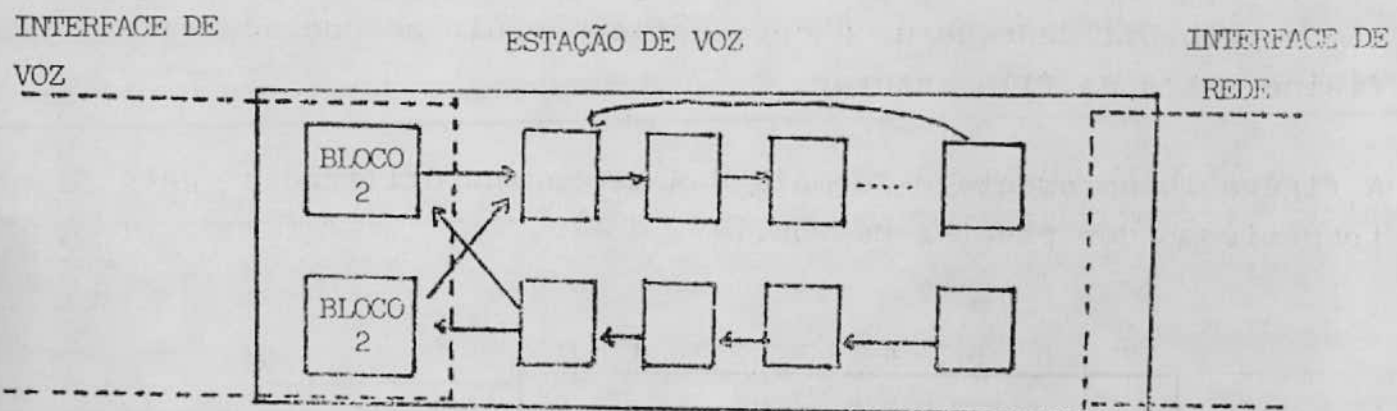


Figura 14.- Funções de 'software' da estação de voz

Toda vez que a estação de voz recebe um pacote de um usuário remoto após um período de silêncio, ela examina se o bit de silêncio está ativado. Caso isso não aconteça, a estação entende que o atraso introduzido na recepção não está sendo suficiente para a compensação do atraso variável introduzido pela rede e transmite uma mensagem para a outra estação solicitando o aumento da compensação de atraso.

Ao receber essa mensagem a estação diminui a fila de transmissão, se for possível, para manter o atraso total constante e autoriza o aumento da fila de transmissão.



Essa troca de informação é realizada através de bits especiais no cabeçalho da mensagem de voz.

A estação receptora também observa a seqüência das mensagens recebidas e encerra automaticamente a conexão de voz, caso a confiabilidade de transmissão esteja baixa.

Essa possível baixa qualidade de transmissão não é devida à rede local, mas sim à reduções na fila de transmissão que aumentam a probabilidade de um pacote não conseguir acesso ao meio físico antes da fila saturar.

A figura 15 apresenta o formato das mensagens utilizadas para transmissão dos pacotes de voz.

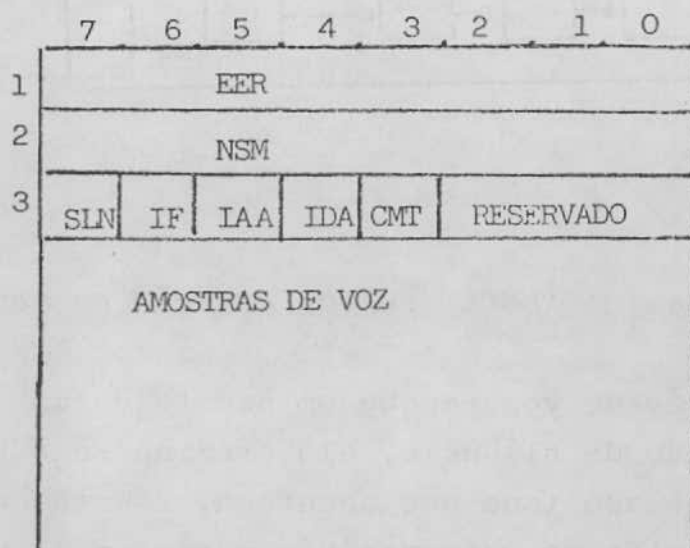


Figura 15. Formato da mensagem de voz

O significado dos campos mostrados na figura 15 é descrito a seguir.

EER (Endereço da Estação Remota): indica a estação remota a qual a mensagem se destina. Esse campo é necessário porque um nó da rede pode implementar mais do que uma estação de voz.



NSM (Número Sequêncial de Mensagem): é utilizado para numerar as mensagens sendo transmitidas e para detectar perdas. A numeração é seqüencial em módulo 256. Nenhum algoritmo é utilizado para recuperação de perdas.

SLN (Silêncio): esse bit é utilizado para indicar que a mensagem é a primeira após um período de silêncio.

IF (Indicador de Formato): indica se a mensagem contém (IF=1), ou não (IF= $\emptyset$ ), um pacote de voz.

IAA (Indicador de Aumento de Atraso): esse bit é utilizado para a estação receptora indicar que deseja aumentar o atraso de compensação de variância por um intervalo de tempo igual à duração de 1 pacote.

IDA (Indicador de Diminuição do Atraso): esse bit é utilizado para a estação receptora indicar que deseja diminuir o atraso de compensação de variância por um intervalo de tempo igual à duração de 1 pacote.

CMT (Confirmação de Modificação do Atraso): esse bit autoriza a alteração do valor do atraso de compensação de variância, conforme solicitado pelos bits anteriormente descritos.

## 5. CONCLUSÕES

A estação de voz aqui descrita está sendo implementada na EPUSP com o patrocínio da FDTE e da SCOPUS Tecnologia S.A.. O 'software', sendo desenvolvido, implementa um serviço de conversações interativas em tempo real. Pretende-se implementar duas estações de voz num computador pessoal NEXUS, de tal forma que medidas da eficiência e do desempenho da estação possam ser realizadas.

Outros serviços podem ser implementados no futuro, tais como , teleconferência, conversações homem-máquina, armazenamento de mensagens de voz, etc.

#### AGRADECIMENTOS

Agradecemos à EPUSP pelo ambiente propício à pesquisa, e à FDTT e SCOPUS Tecnologia S.A. pelo apoio técnico-financeiro recebido.

1. Jerry Borrell, "Voice-Data Integration Enters War for desk tops ", Mini-Micro Systems, February 1985.
2. Howard Frank e Israel Gitman, "Study shows packet switching best for voice traffic, too", Data Communications, March 1979
3. James W. Forgie e Alan G. Nemeth, "An efficient packetized voice/data network using statistical flow control", ICC 1977.
4. P.C. Revasio, et al., "Voice Transmission over an Ethernet backbone", Proc. of the IFIP, Symposium on Local Area Networks, Florence, Italy, April. 1982.
5. Israel Gitman e Howard Frank, "Economic Analysis of Integrated voice and data networks: a case study" Proc. of the IEEE, Vol. 66, nº 11, November 1978.
6. L.R.Rabiner e M.R. Sambur "An Algoritm for determining the endpoints of isolated utterances" BSTJ, vol. 54, nº2, february 1975.
7. L.R. Rabiner e R.W. Schafer, "Digital Processing of Speech Signals", Prentice-Hall, 1978.
8. W.L.Zucchi e W.V.Ruggiero - "Redes Locais com Integração de Serviços de Voz e Dados" - II SBRC, Campina Grande, Abril 1984.
9. W.L. Zucchi e W.V. Ruggiero, "Caracterização de uma Estação de Voz para Redes Locais de Computadores, III SBRC, Rio de Janeiro, Abril 1985.