

UMA PROPOSTA PARA UMA REDE DE ÁREA METROPOLITANA COM PRIORIDADE PARA O TRÁFEGO SÍNCRONO*

MAGDA PATRÍCIA C. ARANTES E SHUSABURO MOTOYAMA
UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE TELEMÁTICA
13.081-970 - CAMPINAS - SP - BRASIL
FONE: 0192 398361
FAX: 0192 391395
E-MAIL: magda@dt.fee.unicamp.br,
motoyama@ccvax.unicamp.br

SUMÁRIO

Uma rede de área metropolitana (MAN) alternativa é proposta neste trabalho. Diferentemente das MANs sob padronização como a FDDI ou DQDB a rede proposta dá prioridade para o tráfego síncrono (voz) e usa técnica de comutação digital distribuída. O objetivo deste tipo de rede é reduzir o custo de investimento da rede telefônica local que possui assinantes espalhados em uma grande área. A rede é flexível o suficiente para se compatibilizar com a emergente B-ISDN de acordo com a tendência internacional. A configuração da rede proposta bem como o protocolo de acesso ao meio serão discutidos neste trabalho. O modelo matemático e a análise de desempenho da rede são também apresentados.

1 - INTRODUÇÃO

Muito esforço está sendo feito atualmente para padronizar as redes de área metropolitanas (MANs) distribuídas e de alta velocidade. Neste sentido o "American National Standards Institute" (ANSI) está padronizando a "Fiber Distributed Data Interface" (FDDI) que se trata de um duplo anel de 100 Mbps e cujo protocolo de acesso é o "token-passing"; já o grupo de trabalho que está padronizando o IEEE 802.6 propõe o "Distributed Queueing Dual Bus" (DQDB) que consiste de um par de barramentos unidirecionais operando em direções opostas e que utiliza o protocolo "slotted channel" para controle de acesso ao meio.

* Este trabalho foi financiado em parte pelo CPqD da Telebrás e pela Fapesp.

Embora as redes FDDI e DQDB tenham sido projetadas para aplicações em comunicações de dados, elas evoluíram no sentido de suportar o tráfego síncrono (tráfego de voz). A rede FDDI evoluiu para a FDDI-II passando a ter capacidade de comutação para o tráfego síncrono. Assim, estas duas redes possuem capacidade de manipular tanto o tráfego síncrono quanto o tráfego assíncrono e além disso são compatíveis com a emergente B-ISDN que deverá integrar todos os serviços em um futuro próximo.

As redes FDDI e DQDB, ambas sob padronização, são apropriadas para países que possuem altas taxas de tráfego de dados e que possuem a demanda de serviços telefônicos plenamente atendida. Entretanto, para países em desenvolvimento como o Brasil, que possuem um serviço telefônico incipiente e baixo tráfego de dados, é importante objetivar projetos de redes telefônicas de baixo custo para o atendimento da demanda telefônica reprimida. Entretanto, essas redes devem ser suficientemente flexíveis para serem compatíveis com a emergente B-ISDN de acordo com a tendência internacional.

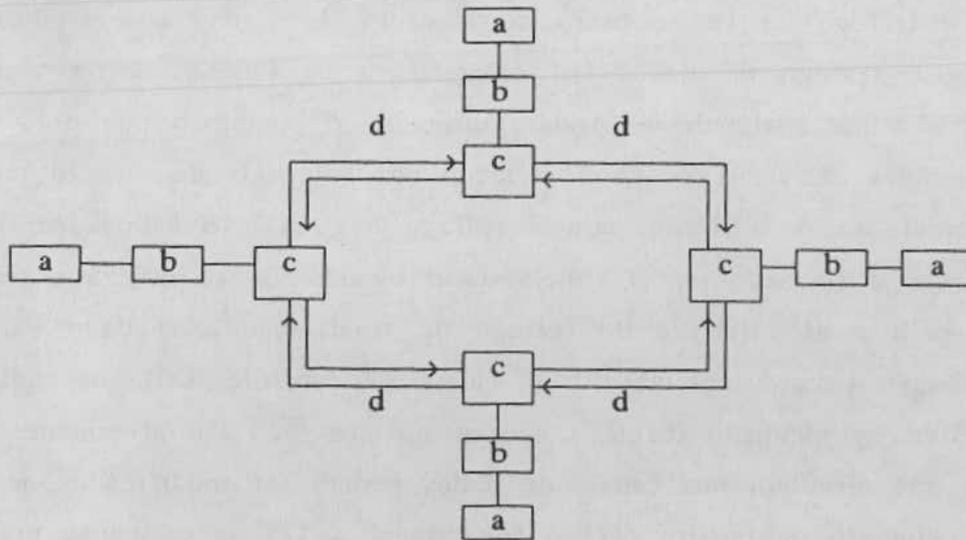
Redes telefônicas de baixo custo podem ser obtidas pelo uso de técnicas de comutação digital distribuída, associadas a um meio de transmissão de faixa-larga como a fibra óptica. A opção pela técnica de comutação digital distribuída é muito importante para países que possuem assinantes telefônicos espalhados em áreas extensas, como é o caso do Brasil. O custo total do investimento de implantação da rede pode ser reduzido trazendo os circuitos da comutação, confinados em pequenas caixas metálicas, para perto dos assinantes. Este procedimento proporciona uma sensível redução no comprimento dos pares de fios que ligam os assinantes aos circuitos de comutação e ainda torna desnecessária a construção de dispendiosos edifícios que abrigariam as centrais telefônicas como é o caso dos atuais sistemas de comutação centralizados.

A rede metropolitana alternativa proposta neste trabalho, ao contrário das MANs sob padronização (FDDI e DQDB), proporciona prioridade ao tráfego síncrono em um estágio inicial e usa técnica de comutação distribuída com o objetivo de se obter redes telefônicas de baixo custo. Ainda assim, esta MAN alternativa é suficientemente flexível para se compatibilizar com a emergente B-ISDN.

2 - CONFIGURAÇÃO DA REDE

A configuração da rede proposta é mostrada na Fig. 2.1. O anel possui um

conjunto de Interfaces Ópticas (IO) interconectadas através de fibras ópticas. Existe, em cada IO, um mecanismo para operar sob condições de falha. Adota-se um duplo anel para se obter maior confiabilidade da rede.



- a - Terminais
- b - Unidade de Acesso Local (UAL)
- c - Interface Óptica (IO)
- d - Fibra Óptica

Fig. 2.1 - Configuração básica da rede.

A cada Interface Óptica é conectado um bloco chamado Unidade de Acesso Local (UAL) e a esse bloco são conectados os terminais (aparelhos telefônicos e/ou terminais de dados). A UAL e a IO constituem um nó dentro da rede. Um dos nós da rede é chamado nó monitor. Este nó realiza as funções de supervisão, manutenção e inicialização telefônica, e ainda realiza algumas funções na sinalização telefônica. Em caso de falha no nó monitor qualquer outro nó em condições perfeitas poderá ser usado para os propósitos de gerenciamento da rede.

As principais funções da IO são a comunicação com todos os outros nós da rede, a multiplexação da capacidade de transmissão da fibra óptica e o controle de acesso ao meio. A unidade de acesso local é responsável por todo o processamento da comunicação de seus terminais, tais como o estabelecimento de chamadas, a supervisão das chamadas aceitas e o término de chamadas, bem como o empacotamento (colocar em forma de pacotes) dos sinais de dados. Adicionalmente, a UAL desempenha funções de comutação local, tornando possível uma eventual comunicação direta entre seus terminais.

3 - MÉTODO DE ACESSO AO MEIO

O método de acesso ao meio proposto para a MAN alternativa é composto por dois mecanismos distintos, um de alocação de canais com comutação de circuito para o tráfego de voz e outro de alocação de canais com comutação de pacotes para o tráfego de dados (ou sinalização). A Fig.3.1 mostra a estrutura de quadro a ser utilizado na rede. O quadro é composto por dois subquadros: o subquadro de voz (V) que é formado por K canais de oito bits onde se utiliza a comutação de circuito para o tráfego de voz e o subquadro de sinalização (D) que é formado por J canais de n bytes onde se utiliza a comutação de pacotes para o atendimento do tráfego de sinalização e/ou dados. Cada canal comutado por pacote tem um bit no campo de controle (CC) que indica se o canal está livre ou ocupado (L/O). Como os nós da rede são ativamente acoplados os bits que circulam nos canais de dados podem ser modificados ou repetidos. O comprimento de quadro (Tf) é feito igual a 125 μ s e apenas um quadro pode ser acomodado no anel. Isto é possível com algum compromisso entre os parâmetros da rede e com adoção de memórias elásticas. Outros parâmetros relevantes da rede a serem considerados para a análise de desempenho são o comprimento, L, do meio de transmissão, a capacidade de transmissão, R, e o número, N, de nós da rede.

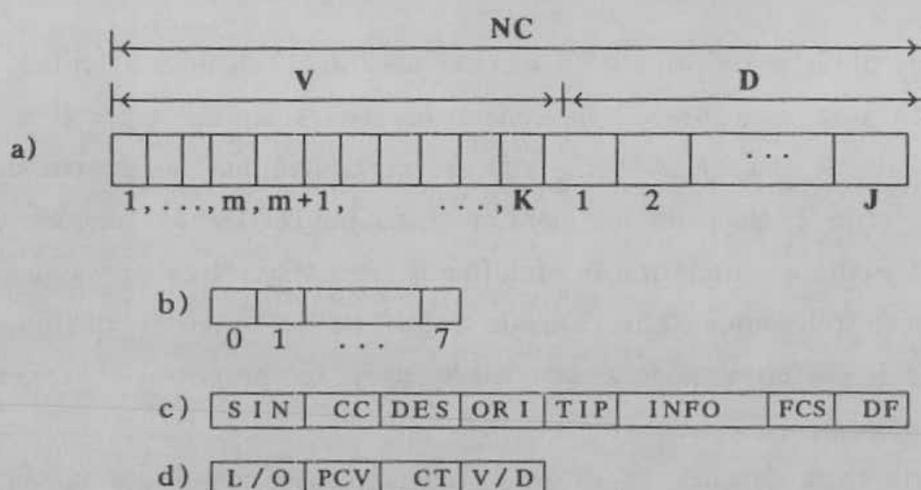


Fig.3.1 - a) estrutura do quadro
 b) canais de voz
 c) canais de sinalização
 d) campo de controle (CC) do canal de sinalização.

Neste método um canal do subquadro de voz é alocado aleatoriamente e permanece alocado durante toda a duração da chamada (comutação de circuito),

atendendo assim a restrição de continuidade de atendimento imposta pelo tráfego de voz. Ao final da chamada o canal é liberado e não pode ser reutilizado pela mesmo nó tendo em vista a equidade de acesso dos nós da rede. O canal alocado a uma conversação é utilizado tanto para a transmissão como para a recepção pelos dois terminais envolvidos.

A estratégia adotada para a sinalização do estado de ocupação dos canais de voz, utiliza alguns dos canais do sub quadro de voz. Isto é possível uma vez que o tráfego de voz varia muito lentamente se comparado com a duração do quadro. Assim, a cada quadro são utilizados m canais para informar a situação de um subconjunto de d canais, sendo necessário portanto q quadros para que o status dos $(K - m)$ canais sejam informados. Este conjunto de q quadros constitui um super-quadro. O primeiro canal de um super-quadro contém uma palavra de sincronismo que indica o seu início. Quando o início de um quadro chega ao nó monitor ele atualiza a tabela de alocação dos canais de voz e no quadro seguinte ele envia a informação sobre o estado de ocupação dos próximos m canais. O nó monitor não trabalha como um nó centralizador para a sinalização, ele apenas atualiza as informações de sinalização pois, separadamente cada nó não possui o controle do super-quadro inteiro.

A relação entre K , m , d e q é dada por,

$$q = \left\{ \frac{K - m}{d} + \frac{1}{m} \right\}^+,$$

onde:

$$\{x\}^+ = \begin{cases} x & \text{se } x \text{ é inteiro,} \\ \{x\}+1 & \text{se } x \text{ fraacionário e } \{x\} \text{ é a parte inteira de } x \end{cases}$$

Os serviços a serem comutados por pacotes são a sinalização telefônica, mensagens de controle e dados em geral.

Existem J canais reservados para a comutação de pacotes circulando continuamente no anel, com o campo (L/O) determinando o seu estado de ocupação. Se um nó particular tem informação a ser transmitida, ele espera por um canal vazio, modifica sua condição para ocupado e envia sua informação utilizando esse canal. Os canais são acessados aleatoriamente pelos nós. Quando um nó utiliza um canal e este retorna ao nó origem, ele é marcado livre e passado ao nó seguinte, promovendo assim equidade no uso do meio de transmissão. Entretanto se um canal tem mais de um pacote a transmitir, ele pode utilizar

todos os canais livres que encontrar obedecendo a restrição acima. Na recepção, os nós analisam continuamente os campos de endereço de origem e de destino dos pacotes. Se um nó reconhece seu endereço no campo de destino, então ele aceita o pacote, envia-o a sua unidade de acesso local sem o campo de sincronismo, e ao mesmo tempo retransmite o pacote de volta ao anel. Por outro lado, se o nó reconhece seu endereço no campo de origem, ele remove o pacote do anel e muda o conteúdo do campo relativo à ocupação do canal para a condição de canal livre.

Nos parágrafos anteriores discutiu-se a alocação de canais para sinais de voz e dados, para a alocação fixa, ou alocação com fronteira fixa, onde os canais reservados para o tráfego de voz não podem ser utilizados para o tráfego de dados e vice-versa. Como o tráfego de voz tem característica aleatória é possível ocorrer situações em que existem canais de comutação de circuito vazios. Estes canais vazios poderiam ser utilizados para o tráfego de dados aumentando assim a eficiência da rede.

A estratégia para usar os canais vazios do subquadro de voz pelo tráfego de dados é chamada alocação dinâmica ou alocação com fronteira móvel e é descrita a seguir: Quando o nó monitor encontra um número suficiente de canais contíguos livres no subquadro de voz ele transmite um pacote endereçado à todos os nós da rede, usando o subquadro de dados, sinalizando que estes canais constituem temporariamente um canal de dados. Qualquer nó com dados a transmitir poderá utilizar este canal, de acordo com a estratégia de alocação para dados, conforme discussão anterior. Este canal adicional de dados pode ser utilizado até que qualquer nó que tenha sinais de voz para transmitir não encontra canais de voz livres. Nesta situação o nó marca o campo PCV (Pedido de Conexão de Voz) no campo CC (Campo de Controle) do pacote de dados. Quando este pacote chega ao nó monitor o campo CT (Controle de Tráfego) é marcado, a fim de desautorizar a utilização dos canais de voz pelos canais de dados, e é feita a liberação dos canais de voz atualizando a tabela de alocação dos canais de voz.

Com a estratégia de alocação dinâmica dos canais todos os canais de voz podem se tornar canais de dados, o que torna a rede flexível o suficiente para acomodar novas tecnologias tais como o ATM que seriam usadas na B-ISDN. Na MAN alternativa proposta, as células de ATM poderiam ser transmitidas dentro dos campos de informação dos pacotes.

4 - ANÁLISE DE DESEMPENHO

Uma vez que na MAN alternativa proposta, o mecanismo de acesso não permite o uso contínuo de um mesmo canal por um nó, todos os nós têm a mesma oportunidade de acesso aos $(K-m)$ canais de voz. Então se todos os processos de chegada de chamadas têm distribuição de Poisson e se a distribuição do tempo de duração da conversação têm duração exponencial, a fórmula Erlang-B pode ser usada para estimar a probabilidade de bloqueio do tráfego de voz.

Para fins de modelamento, o método de acesso proposto pode ser visto como um sistema de múltiplas filas e múltiplos servidores cíclicos. Entretanto não existe na literatura solução exata para este modelo. Neste trabalho, um modelo aproximado desenvolvido por Zafirovic [1] para modelar o mecanismo de acesso a um anel do tipo "slotted ring" será usado. O modelo desenvolvido por Zafirovic é baseado no modelo de servidor cíclico de Takagi [2]. O modelo de Takagi tem solução exata para um sistema de multi-filas e servidor cíclico e único, ou seja, para anéis com apenas um "slot". Caso existam mais de um "slot" no anel, o mesmo modelo será usado, porém, para um anel com J "slots", o tempo de circulação do "slot" será considerado J vezes mais rápido que para o sistema com um único "slot" [1].

A avaliação de desempenho será feita em termos do atraso médio de transferência da mensagem (TM_i), do tempo médio de espera de uma mensagem (WM_i), do tempo médio de espera de um pacote (WP_i), que são definidos à seguir, e ainda da sensibilidade destes elementos às variações de alguns parâmetros da rede (comprimento de mensagens, taxa de chegada de mensagens, etc.).

- *Atraso médio de transferência de uma mensagem* (TM_i) é o tempo médio de duração de serviço a nível da camada MAC, ou seja, é a média dos atrasos desde a chegada da mensagem no nó origem até a sua entrega no nó destinatário;
- *Tempo médio de espera de uma mensagem* (WM_i) é o tempo médio que uma mensagem permanece na fila até começar a ser atendida, ou seja, é o tempo decorrido desde o momento de chegada da mensagem no nó E_i até o início do atendimento do seu primeiro pacote;
- *Tempo médio de espera de um pacote* (WP_i) é o tempo médio que o pacote permanece na fila até que comece a ser atendido, ou seja, é o tempo decorrido desde a chegada do pacote ao nó E_i até o início de sua transmissão.

A seguir serão feitas algumas considerações nas características do sistema afim de se obter a relação entre TM_i , WM_i e WP_i :

O processo de chegada de mensagens no nó E_i é do tipo Poissoniano com taxa λ_i . O comprimento das mensagens tem distribuição exponencial com média $1/\mu_i$ bits e uma distribuição constante para o comprimento dos pacotes é assumido. O número de pacotes por mensagem, portanto tem distribuição geométrica com valor esperado de G_i pacotes, que caracteriza a chegada de pacotes em lotes.

O modelo $M^B / G / 1$ considerando a distribuição geométrica para o número de pacotes por mensagem foi estudado por Halfin [3] e por Whitt [4] e adaptado por Zafirovic-Vukotic [1] que obteve os seguintes resultados:

$$WM_i = WP_i - (G_i - 1) t_i, \quad i = 1, 2, \dots, N$$

onde: t_i é o tempo decorrido entre o início de serviço de dois pacotes consecutivos de uma mesma mensagem, na estação E_i .

WM_i representa também o tempo que o primeiro pacote da mensagem espera na fila.

O tempo de espera da mensagem (TM_i) é constituído pelas seguintes componentes: 1) Tempo médio de espera da mensagem em fila (WM_i); 2) Tempo médio de serviço de todos os pacotes de uma mensagem com exceção do último: $(G_i - 1)t_i$; 3) Tempo de transmissão do último pacote (Tu); 4) Tempo médio de propagação do último pacote da mensagem do nó E_i até o seu destino ($T\rho$); 5) Tempo médio de espera de passagem dos canais de voz ($T\theta$), o qual corresponde à metade do tempo de duração do sub quadro de voz em função da chegada tipicamente aleatória dos pacotes ($T\theta = 8K/2R$).

Portanto:

$$TM_i = WM_i + (G_i - 1) t_i + Tu + T\rho + T\theta$$

e pelas expressões anteriores têm-se:

$$TM_i = WP_i + Tu + T\rho + T\theta$$

$$i = 1, 2, \dots, N.$$

Considerando que a carga é idêntica em todos os nós e que todos os nós estão distribuídos no anel equidistantes uns dos outros os resultados para o modelo de servidor cíclico e limitado, obtidos por Takagi, podem ser aplicados. Assim, o tempo médio de espera de um pacote no nó E_i , WP_i , é dado por [2]:

$$W_P = \left\{ \frac{\lambda G T_u (N + 1) + 2 G (N + 1) - 1}{2 [1 - \lambda G T_u (N + 1)]} \right\} T_u$$

onde o índice i foi suprimido devido às condições de simetria da rede.

Uma vez que o tráfego de voz possui prioridade sobre o tráfego de dados, os resultados do cálculo da probabilidade de bloqueio para os canais de voz no caso de fronteira fixa são válidos para fronteira móvel. Isto se verifica porque quando ocorre uma nova chamada, os canais de voz utilizados temporariamente para transmissão de dados são liberados no quadro seguinte, não havendo portanto, tempo para que a chamada seja bloqueada. O bloqueio ocorre quando todos os canais de voz são efetivamente utilizados para voz.

Considerando o método de acesso com alocação dinâmica para os tráfegos de voz e dados, proposto na seção 3, duas são as situações possíveis para o sistema: 1) Todos os canais de voz estão sendo utilizados pela comutação de circuitos para transmissão de sinais de voz. Neste caso apenas os canais dedicados ao tráfego de dados estão sendo utilizados para a comutação de pacotes; 2) O tráfego de voz é baixo podendo não haver canais de voz sendo utilizados.

Então o número de canais de dados da rede varia de um mínimo de J canais de dados até um máximo de $J + Z_{MAX}$ canais de dados, onde Z_{MAX} representa o número máximo de canais comutados por pacotes que pode ser obtido no subquadro de voz.

O valor de Z_{MAX} é dado por:

$$Z_{MAX} = \left\{ \frac{(K - m) * 8}{Cpd} \right\}$$

onde : (x) representa a parte inteira de x , K é o número de canais de voz, m é o número de canais de voz dedicados à sinalização do estado dos mesmos, Cpd é o Comprimento do canal de dados em bits.

Para o estudo analítico será considerado que o número de canais do subquadro de voz gera um número inteiro de canais de dados. Em casos reais, em função do uso aleatório dos canais de voz, esse fator tem efeito mínimo sobre as medidas de desempenho da rede.

O modelo de servidor único e cíclico de Takagi [2] pode também ser aplicado no caso de alocação dinâmica com as adaptações descritas anteriormente e considerando não apenas os canais adicionais como também a probabilidade deles ocorrerem.

Como um canal comutado por pacote equivale a S canais de voz, a probabilidade de se gerar Z canais de dados adicionais equivale ao complemento da probabilidade de ocupação dos (K-m-Z.S) canais de voz. Esta probabilidade é dada por:

$$P(Z) = 1 - \frac{A^{(K-m-Z.S)}}{(K-m-Z.S) \sum_{i=0}^{K-m-Z.S} \frac{A^i}{i!}}$$

onde: $Z = 1, 2, \dots, \frac{K-m}{S}$

Como a ocupação dos canais de voz é de natureza aleatória, o número médio dos canais comutados por pacotes adicionais é dado por:

$$E(Z) = \sum_{Z=1}^{Z_{MAX}} Z \cdot P(Z)$$

Para cálculo do atraso médio de mensagem no caso de alocação dinâmica serão utilizados os resultados obtidos por Takagi [2] onde o número de canais (ou servidores) disponíveis para dados no anel é $J + E(Z)$. Como foi dito anteriormente, o modelo é conseguido considerando um único servidor com velocidade $J + E(Z)$ vezes maior.

Outra componente do atraso médio de mensagem que se altera com a alocação dinâmica é o tempo médio de espera de passagem dos canais de voz (T_{ϕ}). Esta componente varia de:

$$T_{\phi_{max}} = 8 * K / 2R \quad , \quad \text{até um} \quad T_{\phi_{min}} = 8 * m / 2R.$$

Para cálculo do atraso médio de mensagem será tomado o valor de $E(T_{\phi})$ obtido pela expressão abaixo.

$$E(T_{\phi}) = \frac{8 (K - E(Z) \cdot S)}{2 R}$$

Com as expressões acima obtém-se resultados numéricos para a avaliação de desempenho no caso de se adotar a estratégia de alocação dinâmica de canais.

Ainda para a avaliação de desempenho, tanto para alocação fixa como para a alocação dinâmica, adotou-se uma Taxa de Transmissão R igual a 65,536 Mbps, que corresponde à multiplexação de dois sistemas PCM de terceira ordem proporcionando 1024 canais. Adotou-se L igual a 5 Km para o comprimento do anel e N igual a 30 estações. Os valores adotados para L e N são valores arbitrados. Deve-se lembrar que os parâmetros da rede devem ser adotados com o compromisso de manter fixo o comprimento de quadro T_f igual a 125 μ s.

Para o caso de Fronteira Fixa, as Fig.4.1, Fig.4.2 e Fig.4.3 mostram o atraso médio de transferência de mensagem em termos das taxas de chegada de mensagens para três diferentes comprimentos de mensagens ($L_m = 1/\mu$): 48 bits, 1024 bits ou 5120 bits e para diferentes parâmetros de sub-quadro de dados. Observa-se que um melhor desempenho é obtido para o caso de maior número de canais de dados.

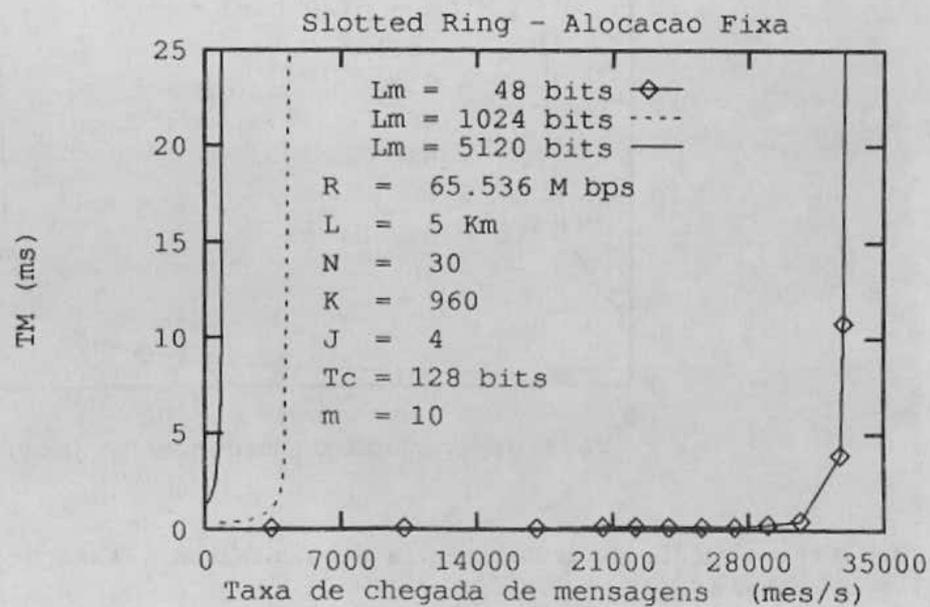


Fig. 4.1 : Atraso médio de transferência de mensagens x Taxa de chegada de mensagens.

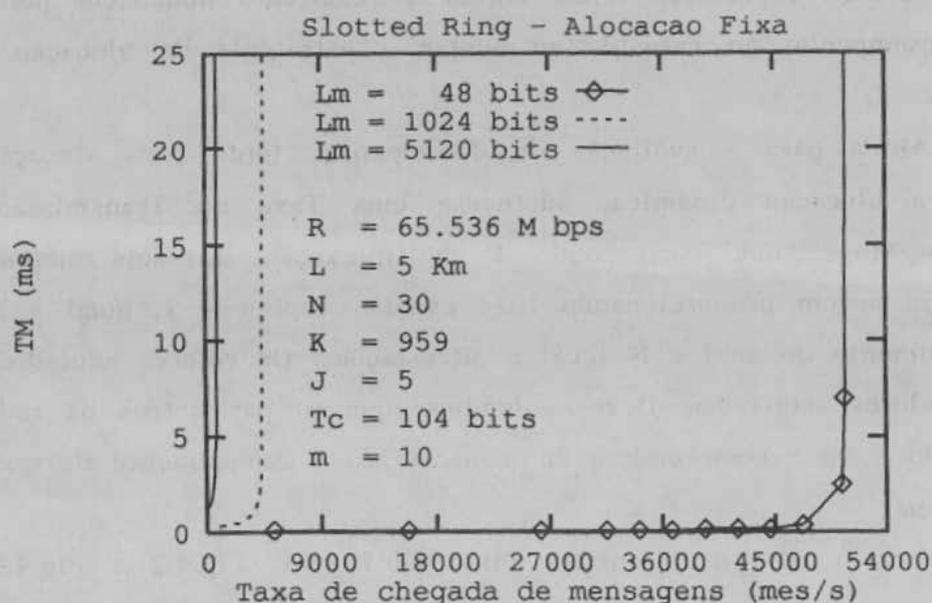


Fig. 4.2 : Atraso médio de transferência de mensagens x Taxa de chegada de mensagens.

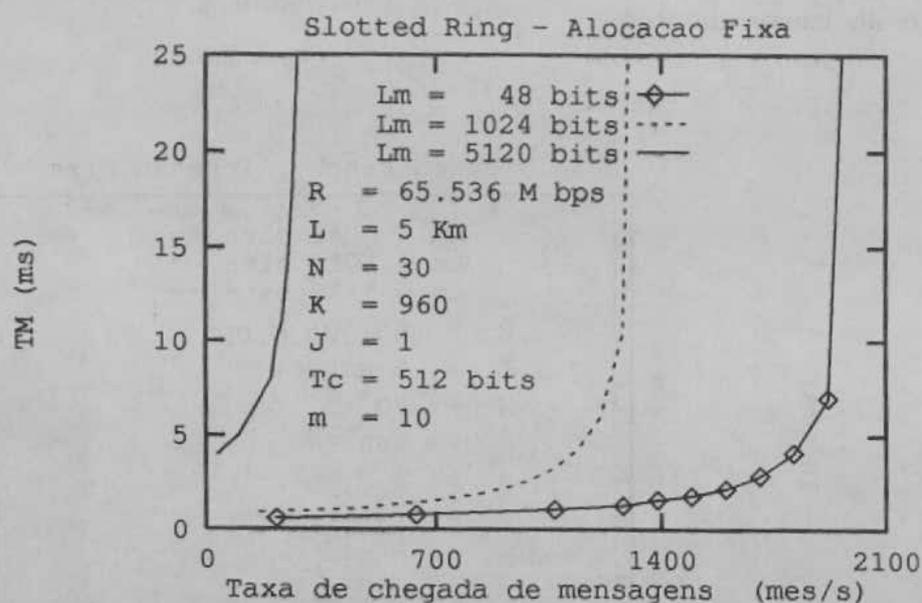


Fig. 4.3 : Atraso médio de transferência de mensagens x Taxa de chegada de mensagens.

Para o caso de fronteira dinâmica a Fig. 4.4 mostra o atraso médio de transferência de mensagens em termos da taxa de chegada de mensagens com probabilidade de bloqueio (B) para o tráfego de voz de 1 %, e para tráfegos de voz que proporcionam uma Taxa de Ocupação dos Canais de Voz (TOCV) de 75% e 50%. Observa-se que para ocasiões de tráfego de voz leve a rede poderá suportar maiores taxas de tráfego de dados.

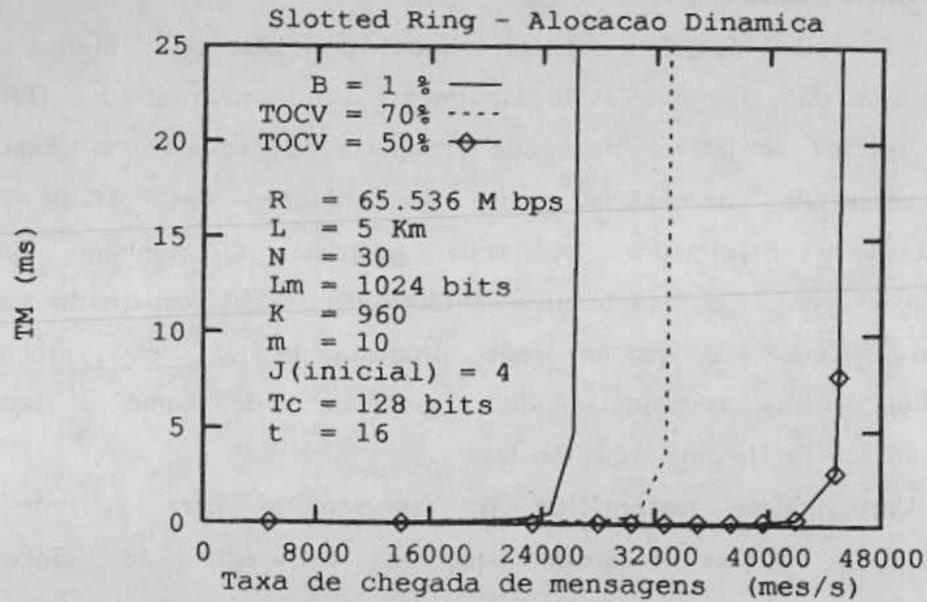


Fig. 4.4 : Atraso médio de transferência de mensagens x Taxa de chegada de mensagens.

Finalmente a Fig. 4.5 mostra uma comparação do desempenho da rede para os três casos apresentados: Alocação fixa com $J = 1$ e $J = 4$ canais de dados e para alocação dinâmica com um número inicial de canais de dados, $J(\text{inicial})$, igual à 4. Nos três casos o comprimento médio de mensagem ($L_m = 1/\mu$) é de 1024 bits. Nota-se que houve uma melhora significativa no desempenho da rede para o caso de alocação dinâmica de canais.

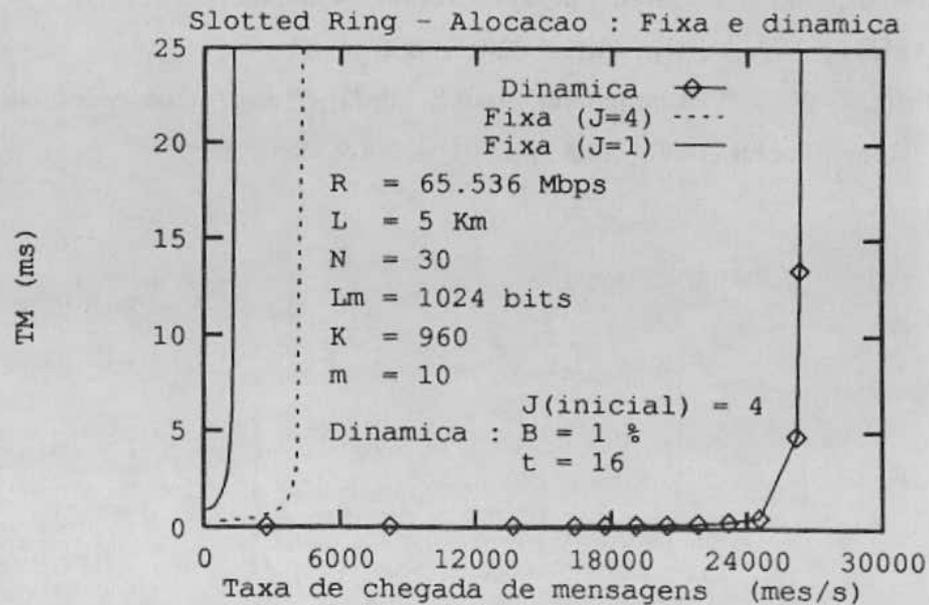


Fig. 4.5 : Atraso médio de transferência de mensagem x Taxa de chegada de mensagens.

5 - CONCLUSÕES

A rede de área metropolitana proposta tem muitas características diferentes daquelas que estão atualmente sendo padronizadas (FDDI e DQDB). O principal objetivo da rede proposta é reduzir o custo inicial de infra-estrutura necessária para implantação dos atuais sistemas de comutação centralizados. A rede proposta é também apropriada para aplicações onde os assinantes telefônicos estão espalhados em uma área extensa. Uma vez que a rede proposta utiliza um protocolo de acesso flexível, novas tecnologias de comutação, tais como a tecnologia ATM, poderão ser facilmente acomodadas.

Um modelo matemático foi apresentado para a rede proposta. Os resultados obtidos mostram que a estratégia de alocação dinâmica proporciona uma significativa melhora na utilização dos canais. Os principais parâmetros para o projeto da rede foram salientados e poderão ser de grande valia para uma eventual implementação da rede.

6 - BIBLIOGRAFIA

- [1] Zafirovic - Vukotic, M. & Niemegeers, J. G., " Performance modelling of HSLAN SlottedRing protocol ", ACM Sigmetrics 88, Santa Fé, New Mexico, p. 163 - 170, Mai 1988.
- [2] Takagi, H., " Analysis of polling systems ", The MIT Press, Cambridge, Massachusetts, 1986 , 197 p.
- [3] Halfin, S., " Batch delays versus customers delays ", Bell Systems Techn., Vol 62, p. 2011 - 2015, Set 1983.
- [4] Whitt, W., " Comparing batch delays and customers delays ", Bell Systems Techn., Vol 62, p.2001 - 2009, Set 1983.