

2º SIMPÓSIO BRASILEIRO SOBRE REDES DE COMPUTADORES (2º SBRC)

EQÜIDADE EM REDES DE COMPUTADORES

Jacques P. Sauvé

Departamento de Sistemas e Computação

e

Grupo de Redes de Computadores

Universidade Federal da Paraíba

58.100 - Campina Grande - Pb

RESUMO

O projeto de redes de computadores visando a eqüidade no compartilhamento dos recursos é examinado. O estudo inclui considerações sobre a dificuldade de se definir uma medida de eqüidade, algumas idéias sobre os fatores que poderiam ser considerados na medida de eqüidade, e considerações sobre os mecanismos disponíveis para se melhorar a eqüidade de uma rede de computadores.

ABSTRACT

The design of packet-switching networks is examined from the point of view of attaining fairness in the sharing of resources. The study includes considerations on the design of a network fairness measure and examines the mechanisms by which the fairness of a network may be improved.

1. Redes de Comutação de Pacotes

Redes de dados para a comunicação entre computadores apareceram durante a década de 70 e sua introdução no Brasil é planejada para o final de 1984. Essas redes empregam a técnica de "comutação de pacotes", desenvolvida na década de 60 devido a motivos que se aplicam ainda hoje. Primeiro, a comunicação entre computadores distantes utiliza canais de comunicação caros, fazendo com que uma fração significativa dos custos para interligar vários computadores advinha dos custos de transporte de dados. Segundo, a natureza dos dados emitidos por um computador (alta intensidade mas, de curta duração) permite que várias conexões compartilhem os mesmos recursos da rede (canais de comunicação e processadores de comutação) a fim de diminuir os custos de comunicação.

Este artigo examina brevemente um dos aspectos a serem considerados no projeto de uma rede de comutação de pacotes: o da equidade no compartilhamento dos recursos.

2. O Projeto de Redes de Comutação de Pacotes

O fato de se compartilhar os recursos da rede entre várias fontes gerando tráfego, imediatamente nos leva a considerar o seguinte problema no projeto da rede: como otimizar o compartilhamento de recursos. Esse problema tem sido abordado por vários pesquisadores, cada um com objetivos diferentes [1-13]. O projeto de uma rede de comutação de pacotes baseia-se num conjunto de medidas de desempenho que o projeto da rede tenta otimizar. As medidas de desempenho mais comuns são:

1. atraso médio fim-a-fim de cada usuário: o atraso em segundos sofrido por um pacote de um determinado usuário no caminho entre o hospedeiro fonte e o hospedeiro destino (um hospedeiro é um computador ou outro equipamento ligado a uma rede de computadores que gera e/ou recebe dados).
2. vazão média de cada usuário: o número de pacotes por segundo transportados pela rede para um determinado usuário.

3. atraso médio fim-a-fim da rede: o atraso médio fim-a-fim, através da rede, de um pacote escolhido randomicamente.
4. vazão média total da rede: o número total de pacotes transportados pela rede por segundo.
5. probabilidade de bloqueio: a probabilidade de um pacote não poder entrar na rede ou em qualquer processador de comutação devido ao excesso de tráfego.

Com o projeto adequado da topologia da rede (localização e dimensionamento dos processadores e dos canais de comunicação), é possível otimizar a medida de desempenho desejada.

3. Eqüidade

Um aspecto do projeto de redes de comutação de pacotes que tem recebido pouca atenção até recentemente, é o da eqüidade da rede. O compartilhamento dos recursos da rede por vários usuários causa uma interferência entre o tráfego desses usuários. O estudo de eqüidade tenta determinar até que ponto alguns usuários estão recebendo tratamento preferencial em detrimento aos outros usuários. Esse estudo é dificultado pela necessidade de se definir uma medida quantitativa do grau de eqüidade da rede para com seus usuários. Essa etapa é crítica e difícil porque cada usuário tem sua própria idéia do que seja uma rede equânime.

Demonstraremos a dificuldade de se definir uma medida de eqüidade através da apresentação daquelas medidas propostas por pesquisadores durante os últimos anos. Cada medida tem vantagens e desvantagens, demonstrando a dificuldade de uma abordagem não polêmica sobre eqüidade.

4. Medidas de Eqüidade

4.1. Eqüidade no Atraso

Em [11], uma medida de eqüidade é apresentada baseada no atraso dos pacotes de cada usuário, e se fundamenta nas se

guintes considerações. O que realmente importa para o usuário é o desempenho recebido e o custo deste desempenho. Considerando o desempenho recebido como sendo caracterizado pelo atraso médio dos pacotes do usuário, podemos comparar o desempenho e o custo deste para se chegar a uma medida de equidade. Por exemplo, certas redes (e.g., Datapac [14]) implantaram um esquema de tarifação tal que a taxa de comunicação é dependente da distância entre os hospedeiros fonte e destinatário. Para redes com uma estrutura não uniforme de tarifação, poderíamos dizer que, para a rede ser equânime, um usuário que esteja pagando mais devido à sua localização, não deveria ser duplamente penalizado com um maior atraso. Este usuário está pagando mais e deveria receber pelo menos o mesmo desempenho que outros usuários. Tal consideração leva à seguinte medida de equidade:

$$F = \frac{1}{T^2} \sum_r \frac{\lambda_r}{\lambda} (T_r - T)^2 \quad (1)$$

onde

- F = medida de equidade
- T = atraso médio fim-a-fim na rede
- λ_r = vazão média do usuário r
- λ = vazão média total da rede
- T_r = atraso médio fim-a-fim do usuário r.

A idéia é de medir o afastamento dos atrasos sofridos pelos diferentes usuários e isto é convenientemente feito definindo F como sendo o quadrado do coeficiente de variação entre as medidas T_r . Numa rede idealmente equânime, teríamos $T_r = T$ para todos os usuários e $F = 0$.

Uma das desvantagens dessa medida é que ela não se aplica a qualquer rede pois a estrutura de tarifação muda de rede para rede. Em [11], os autores oferecem uma medida generalizada de equidade que permite que um atraso ideal, diferente de T, seja escolhido para cada usuário. Definindo T_r^i como sendo o atraso médio ideal para o usuário i, temos:

$$F = \frac{1}{T^2} \sum_r \frac{\lambda_r}{\lambda} (T_r - T_r^i)^2 \quad (2)$$

A busca de maior eqüidade (F pequeno) procura aproximar T_r do seu valor ideal T_r^i . A equação (2) reflete a eqüidade da rede à medida em que possamos escolher os T_r^i de maneira equânime. To memos alguns exemplos.

- i) Para $T_r^i = T, \forall r$, (2) é equivalente a (1) e se aplica a redes com tarifação não uniforme.
- ii) Para uma rede com tarifação independente da localização dos hospedeiros, o usuário espera um desempenho que piore em função da distância percorrida por seus pacotes. Denotando o número de canais na rota dos pacotes do usuário r por $|\Pi_r|$, podemos fazer

$$T_r^i \propto |\Pi_r|$$

para atingir esse tipo de eqüidade.

- iii) O atraso ideal do usuário r , T_r^i , pode ser escolhido em função de objetivos menos óbvios. Por exemplo, se considerarmos que os recursos mais escassos da rede sejam os buffers para armazenar os pacotes, podemos definir uma rede equânime como sendo uma na qual cada usuário utiliza um número igual de buffers. Usando o famoso teorema de Little da Teoria das Filas, isso equivale a fazer T_r^i proporcional a $1/\lambda_r$. Isto é, numa rede perfeitamente equânime, daríamos desempenho melhor (atraso menor) aos usuários grandes (vazão alta) tal que a combinação de vazão e atraso resulte na utilização igual de buffers por parte de cada usuário.
- iv) Como quarta opção, podemos escolher o atraso ideal em função da tarifação cobrada do usuário. Essa função seria escolhida pelo gerente da rede. Por exemplo, se V_r for a taxa cobrada do usuário r , pode-se escolher T_r^i tal que $T_r^i \cdot V_r$ seja igual para cada usuário. Isto é, quanto mais o usuário paga, melhor o seu desempenho. O controle da eqüidade da rede se resumiria à escolha das taxas de tarifação cobradas.

4.2. Eqüidade na Vazão

Uma das desvantagens das medidas de eqüidade apresenta das na seção anterior é a caracterização do desempenho em função apenas do atraso médio fim-a-fim dos pacotes dos usuários. Outra medida importante é a vazão média obtida pelo usuário e a referência [15] examina a definição de medidas de eqüidade baseadas nesta vazão. De forma semelhante à medida dada pela equação (2), podemos definir

$$F = \frac{R}{\lambda^2} \sum_r (\lambda_r - \lambda_r^i)^2 \quad (3)$$

onde

R = número total de usuários

λ_r^i = vazão média ideal do usuário r

observe que uma rede idealmente equânime tem $F = 0$. A vazão média ideal de cada usuário pode ser escolhida de várias maneiras dependendo do conceito de eqüidade que quisermos implementar. Três exemplos seguem:

- i) Para $\lambda_r^i = \frac{\lambda}{R}$, $\forall r$, teríamos o objetivo de manter as vazões dos usuários iguais, o que seria aplicável a uma rede com tarifação não uniforme.
- ii) De forma alternativa, podemos argumentar da seguinte forma: o usuário r oferece tráfego à rede com taxa γ_r mas só recebe uma vazão de λ_r ($\lambda_r \leq \gamma_r$) devido às estratégias operacionais de rede, tais como controle de fluxo. Uma estratégia razoável seria a de manter a vazão do usuário r , λ_r , proporcional à sua demanda, γ_r . Isto é, colocamos

$$\lambda_r^i = \frac{\gamma_r}{\sum_r \gamma_r} \cdot \lambda$$

- iii) Se quisermos incluir a tarifação cobrada na medida de eqüidade, poderíamos por exemplo, escolher λ_r^i tal que λ_r^i / V_r seja igual para cada usuário. Neste caso, uma rede idealmente equânime forneceria tan

to mais desempenho a um usuário quanto este seria cobrado pelo serviço.

4.3. Eqüidade no Gargalo

O enfoque dado à eqüidade em [13] também se baseia na vazão de cada usuário mas, a vazão ideal de cada usuário depende do gargalo no percurso dos pacotes. O gargalo de um percurso é aquele canal de comunicação que fica estrangulado, limitando as vazões de todos os usuários cujos pacotes o percorrem. O critério de eqüidade apresentado é que todos os usuários cujas vazões são limitadas pelo mesmo gargalo tenham vazões idênticas. Infelizmente, o estudo em [13] fornece uma definição de eqüidade sem fornecer uma medida pela qual o grau de eqüidade de uma rede possa ser avaliada. Poderíamos, porém, utilizar a medida de eqüidade dada pela equação (3) e escolher λ_r^i de acordo com o critério dado em [13].

4.4. Eqüidade na Potência

A grande desvantagem das medidas de eqüidade apresentadas nas seções anteriores é que elas caracterizam o desempenho recebido pelo usuário apenas pelo atraso médio fim-a-fim ou pela vazão média. Na realidade o tipo de desempenho de maior importância para um usuário depende do tipo de tráfego gerado por ele. Por exemplo, para um usuário fazendo uma transferência de arquivos, a vazão recebida é a medida de importância enquanto o atraso médio de um pacote não importa. O contrário é verdadeiro para um usuário gerando tráfego iterativo como seria o caso de uma consulta a um banco de dados. Para tráfego em tempo real, o usuário requer uma alta vazão e um atraso pequeno. A fim de caracterizar o tipo de desempenho desejado pelo usuário, introduzimos a "potência" recebida pelo usuário r:

$$\text{Potência recebida pelo usuário } r = P_r = \frac{\lambda_r^{\beta_r}}{T_r}$$

O parâmetro β_r permite que possamos escolher a vazão como sendo prioritária (β_r grande) ou o atraso como sendo de maior importância (β_r pequeno). De posse dos parâmetros β_r (escolhidos pelos usuários), podemos definir uma medida de eqüidade como

segue:

$$F = \sum_r (P_r - P_r^i)^2 \quad (4)$$

Como no caso das equações (1), (2) e (3), essa equação dá uma medida do afastamento entre o conjunto de potências recebidas pelos usuários (P_r) e o conjunto de potências ideais (P_r^i). As potências ideais podem ser escolhidas de maneira a refletir o conceito de equidade desejado. Poderíamos, por exemplo, fazer todos os P_r^i iguais ou então incluir a tarifação, fazendo P_r^i/V_r igual para todos os usuários.

4.5. Equidade de Penalidade Versus Congestionamento

Um conceito muito interessante de equidade foi apresentado em [12]. A medida de equidade proposta tem a vantagem da seção anterior de incluir ambas as medidas importantes: atraso médio fim-a-fim e vazão média.

A premissa básica é a de que a falta de equidade de uma rede não é medida pelo desempenho final dado aos usuários mas pela interferência entre os usuários resultante do compartilhamento de recursos. Por exemplo, se tivéssemos apenas dois usuários na rede e eles não compartilhassem nenhum recurso, deveríamos considerar a rede como sendo perfeitamente equânime, por não haver possibilidade de interferência.

A fim de se chegar a uma medida de equidade que se case com esse conceito, duas medidas intermediárias são propostas em [12] para refletir o grau de interferência entre usuários:

1) Medida de Penalidade:

$$PE_r = \frac{\lambda_r^s - \lambda_r}{\lambda_r}$$

onde λ_r^s é a vazão que o usuário r receberia caso não houvesse interferência dos demais usuários que compartilham com ele os recursos da rede. PE_r mede o quanto o usuário r deixa de transmitir (é penalizado) por causa dos demais.

2. Medida de Congestionamento (*):

$$C_r = \sum_{j \neq r} \frac{\lambda_j}{\lambda} \cdot \frac{T_j - T_j^S}{T_j^S}$$

onde T_j^S é o atraso médio fim-a-fim do usuário j se fosse retirado o usuário r . C_r mede quanto o determinado usuário afeta (interfere) os demais, aumentando o atraso destes.

As duas medidas se completam, enfocando o problema de ambos os lados: de um usuário em relação aos demais e vice-versa. De posse dessas medidas, uma rede é considerada equânime se, para cada usuário, o ganho que este tem ao congestionar os demais deve ser pago penalizando-se na mesma proporção. Isto é, o desempenho de cada usuário da rede deve obedecer a relação:

$$\frac{PE_r}{C_r} = \text{Constante}, \forall r$$

Uma medida de equidade pode ser obtida calculando o afastamento dos valores PE_r/C_r com relação à constante desejada (pode-se usar, por exemplo, o desvio padrão do conjunto $\{PE_r/C_r\}$).

5. Otimização da Equidade

As medidas de equidade apresentadas nas seções anteriores foram propostas recentemente e introduziram um assunto bastante polêmico já que dificilmente os usuários e o administrador da rede concordarão sobre o melhor conceito de equidade. A princípio, nos parece que uma medida de equidade poderia incluir a satisfação dos usuários, a tarifação da rede, a localização dos usuários e uma medida dos graus da interferência entre usuários. A medida deveria ser parametrizada de modo a permitir um ajuste dos desempenhos ideais a serem alcançados.

(*) - A medida C_r é definida de uma maneira diferente em [12], mas essa diferença não influe na discussão apresentada aqui.

Qualquer que seja a medida de equidade usada, surge imediatamente o problema de se melhorar a equidade de uma rede. Isto pode ser feito, durante o projeto da rede, bem como durante sua operação, com a escolha judiciosa de certas políticas operacionais. Os parâmetros da rede que podem ser utilizados para melhorar a equidade da rede são: a escolha do encaminhamento dos pacotes dos usuários [11], a escolha dos parâmetros de controle de fluxo [11, 12, 13] e o escalonamento prioritário dos pacotes [11]. O ajuste desses parâmetros durante a operação da rede deve ser baseado em algoritmos de otimização da equidade que podem ser implementados de maneira centralizada ou distribuída.

Finalmente, devemos salientar que a otimização da equidade de uma rede não é um problema isolado desde que sempre haverá um compromisso entre uma distribuição equânime dos recursos e um desempenho global (atraso médio e vazão total) aceitável da rede.

REFERÊNCIAS

1. L. Kleinrock, Queueing Systems: Volume II. Computer Applications. New York: Wiley Interscience, 1976.
2. M. Schwartz e T.E. Stern, Routing Techniques used in Computer Communications Networks, IEEE Transactions on Communications, Vol COM-28, pp. 539-552, abril 1980.
3. J.M. McQuillan, Routing Algorithms for Computer Networks - A Survey, Anais da 1977 National Telecommunications Conference, dezembro 1977.
4. M. Reiser, A Queueing Network Analysis of Computer Communication Networks with Window Flow Control, IEEE Transaction on Communications, Vol. COM-27, pp. 1199-1209, agosto 1979.
5. M.C. Pennotti e M. Schwartz, Congestion Control in Store-and-Forward Tandem Links, IEEE Transactions on Communications, Vol. COM-23, pp. 1434-1443, dezembro 1975.
6. J.W. Wong, Flow Control in Message-Switched Communications Networks, Computer Communication, Vol. 1, pp. 67-74, abril 1978.
7. M. Gerla e L. Kleinrock, Flow Control: A Comparative Survey, IEEE Transactions on Communications, Vol. COM-28, pp. 553-574, abril 1980.
8. M. Irland, Buffer Management in a Packet Switch, IEEE Transaction on Communications, Vol. COM-26, pp. 328-337, março 1978.
9. F. Kamoun e L. Kleinrock, Analysis of Shared Finite Storage in a Computer Network Environment under General Traffic Conditions. IEEE Transactions on Communications, Vol. COM-28, pp. 992-1003, 1980.

10. M. Gerla, The Design of Store-and-Forward Networks for Computer Communications, School of Engineering and Applied Science, University of California at Los Angeles, Rep. UCLA-ENG-7319, janeiro 1973.
11. J.W. Wong, J.P. Sauv e e J.A. Field, A Study of Fairness in Packet-Switching Networks, IEEE Transactions on Communications, Vol. COM-30, pp. 346-353.
12. F. Wong, Eq uidade entre Circuitos Virtuais de uma Rede de Computadores, CETUC-D-ST-03-83, PUC/RJ, junho 1983.
13. J.M. Jaffe, Bottleneck Flow Control, IEEE Transactions on Communications, Vol. COM-29, pp. 954-962.
14. Bell Telephone Co. Canada, General Tariff CTC(TP), N  6716, Item 4200 Datapac Service, Efetivo junho 1977.
15. J.P. Sauv e, J.W. Wong e J.A. Field, On Throughput and Fairness in Packet-Switching Networks with Window Flow Control, Relat rio E-100, CCNG, University of Waterloo, dezembro 1981.