Comunicações Multiponto Confiáveis Escaláveis através de Transmissão de Múltiplas Cópias

Daniel Antunes Maciel Villela daniel@gta.ufrj.br

Otto Carlos M. B. Duarte* otto@gta.ufrj.br

Grupo de Teleinformática e Automação – GTA
COPPE/EE - Programa de Eng. Elétrica/Departamento de Eletrônica
Universidade Federal do Rio de Janeiro
C. P. 68504 - CEP 21945-970 - Rio de Janeiro - RJ - Brasil
http://www.gta.ufrj.br

Resumo

Este trabalho analisa o desempenho de protocolos multiponto confiáveis que se servem de transmissões por múltiplas cópias com o objetivo de reduzir os efeitos do problema de implosão de reconhecimentos. A partir da adoção desta estratégia, são definidos protocolos que possuem diferentes abordagens e a análise de desempenho dos sistemas que operam sob estes protocolos é realizada. Os resultados obtidos mostram que protocolos com múltiplas cópias possuem um atraso pouco sensível ao aumento do número de receptores. O desempenho dos protocolos definidos revela-se superior, quando comparados com os convencionais, mesmo quando há aumento da taxa de perda de pacotes e da taxa de geração de dados. Por isso, a utilização desta estratégia é interessante e permite maior escalabilidade às comunicações multiponto confiáveis e diminuição do atraso médio gerado em sistemas multiponto.

Abstract

This paper proposes a transmission of multiple copies to be used in reliable multicast protocols in order to reduce the effects brought by the problem of acknowlegement implosion. This strategy is applied for definition of protocols that have different approaches. Performance evaluation of the systems that operate under these protocols is derived. The results show that protocols with multiple copies provide stability for a larger number of receivers. The performance of these systems with raise of rate of packet loss and data generation rate is higher than the systems with only one transmission. Thereby, this strategy becomes interesting, since it allows more scalability for the reliable multicast communications.

*Este trabalho foi realizado com recursos da UFRJ, FUJB, PROTEM-CC, CNPq, CAPES.

1 Introdução

Atualmente, a escalabilidade em sistemas de comunicação multiponto é fundamental, pois crescem o número de aplicações e número de participantes destes sistemas. O crescimento e interconexão das redes a nível mundial permite uma troca de dados intensa entre grupos integrados e espalhados. O MBone [1] encontra-se sobre a Internet e opera em multiponto permitindo aplicações como conferências em *whiteboard* e com grande número de usuários. As redes de alta velocidade e o crescimento da utilização de multimídia em estações de trabalho fazem aumentar cada vez mais a demanda por aplicações distribuídas. É importante também o desempenho final sob ponto de vista da qualidade de serviço proporcionada como também da utilização de recursos limitados existentes nos canais de comunicação. Por isso, há uma necessidade de protocolos de transporte que sejam confiáveis, eficientes e sejam perfeitamente escaláveis em relação ao número de participantes.

O requisito de confiabilidade exige informações de confirmação nas transferências de dados como os reconhecimentos (*acknowledgements*) que são retornados ao transmissor, em duas formas básicas: reconhecimento positivo, denominado ACK e reconhecimento negativo, NAK. À medida que o crescimento dos grupos em comunicação gera número alto de participantes nas aplicações envolvidas, a quantidade de reconhecimentos pode-se tornar muito alta acarretando uma sobrecarga para os transmissores. Trata-se do problema de implosão de reconhecimentos que gera efeitos nocivos ao desempenho do sistema. Podem ocorrer *overhead* em processamento, ocasionando atraso no envio dos dados, e escassez de recursos, como largura de faixa e espaço em *buffer*.

As propostas para reduzir ou mesmo eliminar o problema de implosão de reconhecimentos possuem diversas vertentes. Alguns autores propõem hierarquias como topologias de árvore, anel [2, 3, 4, 5] e divisão em grupos locais [6, 7]. Biersack *et al.* propõem a utilização da recuperação por retransmissão (*Automatic Repeat Request* - ARQ) associada à correção automática (*Forward Error Control* - FEC) em um esquema conhecido como ARQ híbrido [8, 9]. Em De Lima e Duarte [10] uma estratégia ARQ com *selective repeat* é analisada para comunicações ponto a multiponto. Floyd *et al.* [11] introduzem o conceito de enquadramento a nível de aplicação (*Application Level Framing* - ALF). Nunes e Duarte [12] apresentam a utilização de ALF e suporte a mobilidade em um protocolo multidestinatário. Pingali, Kurose e Towsley [13] fazem uma análise de modelos para protocolos multidestinatários confiáveis em uma comparação de protocolos orientados pelo transmissor, orientados pelo receptor e com um mecanismo de supressão de NAKs, proposto inicialmente em [14, 11]. Seu estudo baseia-se em análise de vazão dos protocolos. Yamamoto, Kurose e Towsley [15] realizam o estudo destes modelos para análise de atraso. Villela e Duarte [16, 17] apresentam a possibilidade de utilização de uma política de prioridades como forma de distinção entre dados e reconhecimentos para aumentar a robustez de um sistema de comunicação multiponto confiável.

Este trabalho propõe a transmissão de múltiplas cópias de pacotes de dados,

como nova possibilidade de reduzir o número de reconhecimentos retornados à rede e também diminuir o atraso gerado. Com o auxílio de várias abordagens distintas, define-se três protocolos multiponto confiáveis que adotam esta estratégia e que estão baseados nos modelos propostos em [13, 15], para efeito de comparação. É realizada a avaliação de desempenho em sistemas que utilizem estes protocolos, observando-se o atraso gerado. Com a análise efetuada, torna-se possível o conhecimento das conseqüências em desempenho e escalabilidade por meio do atraso médio gerado nas diferentes abordagens e o ganho em escalabilidade proporcionado pela utilização da estratégia proposta, variando-se a probabilidade de perda de pacotes, a taxa de geração de dados e o número de cópias enviadas a cada transmissão.

Neste trabalho, a seção 2 descreve o funcionamento da transmissão por múltiplas cópias, seus propósitos e define os três protocolos que a adotam. Em seguida, na seção 3, é realizada a análise de desempenho destes protocolos em questão. Na seção 4, são apresentados resultados numéricos obtidos a partir desta análise, sob variação da probabilidade de perda, da taxa de geração de dados no emissor do grupo de comunicação multiponto e do número de cópias enviadas. Por fim, a seção 5 apresenta as principais conclusões.

2 Transmissão por Múltiplas Cópias aplicada a Protocolos Multiponto Confiáveis

Para reduzir os efeitos da implosão de reconhecimentos, neste trabalho é proposta a transmissão por múltiplas cópias a ser utilizada em protocolos multiponto confiáveis. Esta seção apresenta o funcionamento detalhado da transmissão com múltiplas cópias e define protocolos que a utilizam juntamente com abordagens como orientação pelo transmissor e pelo receptor [13] e também mecanismo de supressão de NAKs.

2.1 Transmissão por Múltiplas Cópias

Normalmente, os transmissores de um sistema de comunicação multiponto enviam pacotes de dados e esperam o momento em que se pode considerar que o pacote foi perdido ou, em caso contrário, teve sucesso na transmissão em questão. Reconhecimentos positivos (ACKs) poderiam determinar a transmissão realizada com sucesso. Se houver erro, após um intervalo de tempo pré-definido (*timeout*) sem o devido reconhecimento, o pacote é considerado perdido e o transmissor será responsável por sua retransmissão. Por outro lado, o protocolo utilizado pode basear-se em reconhecimentos negativos (NAKs) que assinalam as ausências de pacotes e, consequentemente, causam a retransmissão de pacotes perdidos. Se os pacotes são perdidos a taxas pequenas, em geral, apenas uma transmissão será suficiente para uma correta recepção para a maior parte dos pacotes gerados. Em um caso ideal, onde não há perda, todos os pacotes são recebidos perfeitamente em uma transmissão.

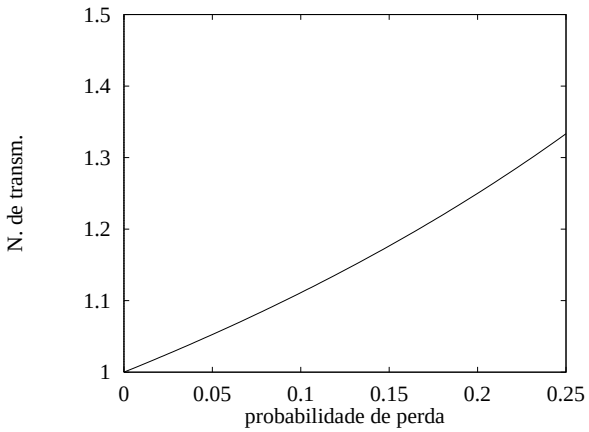

Figura 1: Número de transmissões necessárias para recepção correta por um receptor.

Entretanto, o cenário é bem diferente se a taxa de perda de pacotes é relativamente alta. Nesta situação, algumas retransmissões podem ser previstas até a completa recepção do pacote para um receptor escolhido dentre os participantes do sistema de comunicação. Ainda, o número de transmissões necessárias para todos os receptores receberem corretamente o pacote será maior. Mesmo taxas de perda consideradas aceitáveis para um único receptor podem ser críticas quando há grande quantidade de participantes no sistema. Por isso, torna-se necessário soluções para alcance de escalabilidade nestes sistemas.

Por isso, transmitir cópias dos pacotes é uma possibilidade a ser considerada em comunicações multiponto. Esta é a base da estratégia de transmissões múltiplas proposta neste trabalho. Seu funcionamento é bastante simples: a cada transmissão o transmissor envia m cópias em vez de uma única. Isto irá reduzir a taxa de perda de pacotes. A simplicidade da estratégia também implica em fácil implementação e pouca carga de processamento adicionada.

A figura 1 exemplifica o número de transmissões necessárias até a correta recepção de um pacote de dados por parte de um único receptor, para o caso em que se assume uma perda de pacote inteiramente independente de qualquer possível perda anterior. Para probabilidades de perda mais altas o número de transmissões necessárias sofre acréscimo de mais de 20% do caso ideal em que uma única transmissão é suficiente. De posse de um conjunto de receptores, este acréscimo sofre acréscimo ainda maior. A figura 2 mostra o número de transmissões necessárias para completa recepção de um pacote por parte de todos os receptores participantes da comunicação multiponto. Neste caso, trabalhando-se com grande número de receptores, o número de transmissões necessárias chega a exceder o valor 7. Logo, com o envio de m cópias e considerando-se o uso de reconhecimentos negativos, os pacotes estarão sendo en-

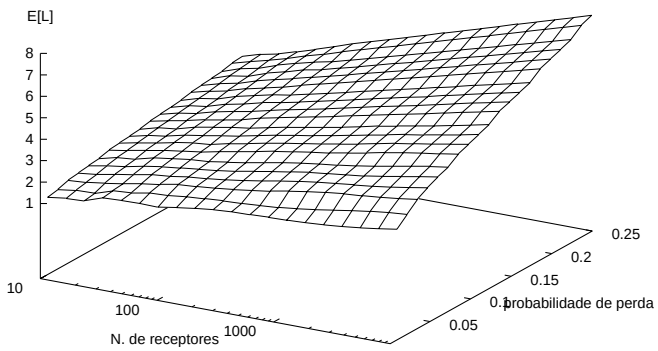

Figura 2: Número de transmissões necessárias até a completa recepção por parte de todos os receptores integrantes do grupo de comunicação.

viados em avanço, reduzindo-se o número de NAKs retornados pelos receptores aos transmissores. Além disso, o atraso médio gerado diminuirá uma vez que elimina-se uma parcela de tempo referente ao retorno de NAKs e também se reduz o número de respostas dos receptores, diminuindo outra parcela referente a processamento de NAKs.

Por fim, deve-se acrescentar que transmissões em múltiplas cópias foram anteriormente propostas por Sastry [18] em protocolos com retransmissão contínua (*Go-Back-N*). Uma variação foi apresentada por Birrel [19] em uma estratégia semelhante a que é apresentada neste trabalho. Ademais, os própositos que foram visados nestes trabalhos são diferentes do que se objetiva neste trabalho em que se busca escalabilidade e alto desempenho.

2.2 Um Protocolo orientado pelo Transmissor

Será definido inicialmente um protocolo orientador pelo **T**ransmissor e com transmissões por **M**últiplas cópias (**TM**). Como a orientação se faz pelo transmissor, a cada pacote recebido corretamente o receptor deve retornar um reconhecimento positivo ao transmissor que mantém informação sobre ACKs recebidos dos receptores em uma lista. Caso não receba os devidos reconhecimentos de um determinado pacote por todos os receptores dentro de um certo intervalo de tempo, deverá iniciar

retransmissões para os destinatários que não enviaram esta resposta. São utilizados temporizadores para determinar a retransmissão de pacotes. As transmissões ocorrem por múltiplas cópias, por isso, a cada novo pacote a ser enviado, o transmissor se encarregará de enviar m cópias. Ao estouro de temporizador, são realizadas retransmissões também com m réplicas.

A estratégia de retransmissão empregado é a retransmissão seletiva (*selective repeat*). Este protocolo pode ser relacionado com o protocolo **A** descrito em [13] que também é controlado por transmissor.

2.3 Protocolos orientados pelo receptor

Distintamente do protocolo definido anteriormente, uma outra opção é a orientação pelo receptor. Com isso, define-se um segundo protocolo denominado **RM**, orientado pelo **R**eceptor com transmissão por **M**últiplas cópias. Esta abordagem coloca nos receptores a responsabilidade pela confiabilidade da transferência dos dados. O transmissor irá transmitir pacotes em conjuntos de m réplicas até que receba um reconhecimento negativo (NAK) por parte de um receptor, indicando que determinado pacote foi recebido com erro ou simplesmente não chegou. Neste caso, ao emissor caberá retransmitir o pacote mais uma vez em grupo de m cópias. A operação efetuada nos receptores para deteção de pacotes perdidos baseia-se em duas possibilidades. Primeiramente, um receptor pode constatar perda de pacote por lacuna na numeração de sequenciamento dos pacotes. Ou, ainda, quando uma retransmissão é requisitada, associa-se um temporizador para o pacote desejado. Se o período de temporização expirar, o receptor enviará reconhecimento negativo ao transmissor.

Pode-se, ainda, definir um terceiro protocolo com o acréscimo de um mecanismo de supressão de NAKs. Este protocolo será chamado **RMS**, pois é orientado pelo **R**eceptor, com transmissões por **M**últiplas Cópias e com **S**upressão de NAKs. Um esquema de supressão de NAKs foi proposto por Ramakrishnan e Jain [14] para redes locais (LANs). Um mecanismo semelhante é usado também por Floyd e Jacobson [11] com o mesmo intuito em WANs e pode ser descrito de forma breve. Os receptores, ao transmitirem um NAK, o fazem em multiponto de tal forma que todos os participantes do sistema terão condições de perceber o pedido de retransmissão. Entretanto, o envio de NAK não é feito de forma imediata, pois determina-se um intervalo de tempo aleatório para que isto aconteça a partir da deteção da perda de um pacote. Se os demais receptores receberem um NAK de outro receptor, enquanto estão prontos para enviarem seus NAKs e aguardam durante seus respectivos períodos aleatórios de tempo, irão suspender a transmissão de seu próprio NAK. Com isto, obtém-se o efeito de redução de carga de NAKs colocados na rede em respostas.

Os protocolos **RM** e **RMS** também podem ser relacionados com os protocolos **N1** e **N2**, respectivamente, apresentados em [13], que apresentam os conceitos de controle por receptores e supressão de NAKs.

3 Análise de Desempenho

Neste trabalho, é enfocada particularmente a análise de atraso gerado pelos protocolos definidos. Este parâmetro compreende o intervalo de tempo decorrido desde a chegada do pacote ao *buffer* do transmissor até sua recepção correta e completa pelo receptor sob os protocolos **TM**, **RM** e **RMS**. O atraso sofrerá influência de fatores como atraso de propagação fim-a-fim, probabilidade de perda de pacote, tempos de processamento tanto no transmissor como no receptor e ainda alguns parâmetros dos protocolos como o tempo aleatório gerado no protocolo **RMS** pelo mecanismo de supressão de NAKs.

A análise apresentada assume algumas premissas a exemplo do que é apresentado por Villela e Duarte [16]. O grupo de participantes que compõem o sistema de comunicação multiponto é formado por um transmissor e R receptores. Os pacotes novos a serem enviados aos receptores são gerados no transmissor de acordo com um processo de Poisson com taxa λ. A probabilidade de perda de um pacote de dados enviado por um transmissor a um receptor é p_r. Esta probabilidade é a mesma para todos os receptores. As perdas de pacotes são assumidas independentes no espaço e no tempo, ou seja, independem dos receptores e de alguma eventual perda de pacote ocorrida anteriormente. Assume-se, ainda, que os reconhecimentos nunca são perdidos. Considera-se que o tempo de processamento é constituído em sua maior parte por procedimentos comuns às m cópias, como construção de cabeçalhos por exemplo, e despreza-se o tempo de transmissão adicional. O atraso de propagação entre o transmissor e cada um dos receptores possui valor τ, que é o mesmo para todos os participantes do sistema de comunicação multiponto.

O tempo de processamento necessário para o devido tratamento a pacotes de dados é uma variável aleatória X. Da mesma forma, o tempo necessário para processar um reconhecimento, seja ACK ou NAK, seja para enviar ou receber, é outra variável aleatória chamada Y. Estas variáveis aleatórias X e Y possuem médias $E[X]$ e $E[Y]$ e segundos momentos $E[X^2]$ e $E[Y^2]$, respectivamente.

Como a probabilidade de perda de um pacote é dada por p_r, a probabilidade de um grupo inteiro de m cópias ser perdido será $p = p_r^m$. Pode ser definida a váriavel aleatória $N_i^{(r)}$ correspondente ao número de vezes que um grupo de pacotes de dados deve ser transmitido até que um pacote deste grupo seja corretamente recebido pelo receptor i. Outra variável aleatória importante será N, equivalente ao número de transmissões necessárias para que um pacote de dados seja recebido por todos os receptores corretamente. O conhecimento da distribuição destas variáveis aleatórias permite obter relações importantes a serem utilizadas na análise:

$$E[N^{(r)}] = \frac{1}{1-p} \text{ e } E[N] = \sum_{n=1}^{\infty}(1-(1-p^{n-1})^R). \quad (1)$$

De posse destes dados, pode-se estudar os atrasos observados nos três modelos de protocolos propostos, **TM**, **RM** e **RMS**.

3.1 Atraso para Protocolo Orientado pelo Transmissor com Múltiplas Cópias

Serão calculados, inicialmente, os tempos de espera de pacotes em *buffer* para o protocolo **TM**. A taxa de pacotes retransmitidos devido a estouro de temporizador será equivalente a $\lambda(E[N]-1)$. Portanto, o fluxo de dados será denotado por $\lambda_d^T = \lambda E[N]$. Para o cálculo de taxa de chegada de ACKs, é necessário observar a taxa de chegada dos pacotes nos receptores. De um conjunto m de réplicas, o número de pacotes que chega a um receptor com sucesso é uma variável aleatória com distribuição binomial. Logo, como a probabilidade de perda para um simples pacote (réplica) é p_r, a média desta variável aleatória é $m(1-p_r)$. Sabendo-se que são necessárias em média $E[N]$ transmissões a fim de que todos os R receptores recebam corretamente um pacote e que a taxa de chegada e sua distribuição são as mesmas para todos receptores, a taxa de chegada de ACKs será denotada λ_a^T, com valor $m\lambda RE[N](1-p_r)$. Lembrando que as variáveis aleatórias Y e X correspondem ao tempo de processamento de um ACK e um pacote de dados, respectivamente, pode-se escrever a carga introduzida ao transmissor como

$$\rho_{TM}^T = m\lambda RE[N](1-p_r)E[Y] + \lambda E[N]E[X]. \tag{2}$$

É importante observar que no transmissor não haverá m réplicas do pacote de dados em *buffer*. As cópias dos pacotes são geradas pelo protocolo e enviadas. Trata-se de apenas processamento comum do pacote e repetição do procedimento de envio determinado na implementação em m vezes. O fato de transmitir réplicas não implica em cópia de memória, o que causaria degradação ao desempenho do sistema. Ao contrário, esta estratégia é muito simples de se implementar e o tempo adicional ao processamento pode ser considerado desprezível.

Como o processo de chegada de dados em *buffer* é Poisson e assume-se que os tempos de serviço são variáveis aleatórias independentes, modela-se este sistema como uma fila M/G/1. Pode-se, portanto, utilizar as relações para tempo de espera em tal fila [20]. Com isso, o tempo de espera para pacotes de dados em *buffer* é

$$E[W_{TM}^T] = \frac{\lambda_d^T E[X^2] + \lambda_a^T E[Y^2]}{2(1-\rho_{TM}^T)} \ . \tag{3}$$

De forma mais simples, nos receptores há apenas um fluxo ocasionado por chegada de dados. Sua taxa é denotada por $\lambda^R = m\lambda E[N](1-p_r)$. Sabendo-se que o recebimento correto de pacote de dados será seguido por transmissão de ACK, o tempo de processamento necessário será $X+Y$. Por conseguinte, a carga de trabalho no receptor é $\rho_{TM}^R = \lambda E[N](1-p_r)(E[X]+E[Y])$. Mais uma vez, assumindo-se que as chegadas ocorrem segundo processo de Poisson, encontra-se o tempo médio de espera para um pacote no *buffer* de recepção pela equação de Pollaczec-Kinchin [20]:

$$E[W_{TM}^R] = \frac{\lambda^R(E[X^2]+E[Y^2]+2E[X]E[Y])}{2(1-\rho_{TM}^R)}. \tag{4}$$

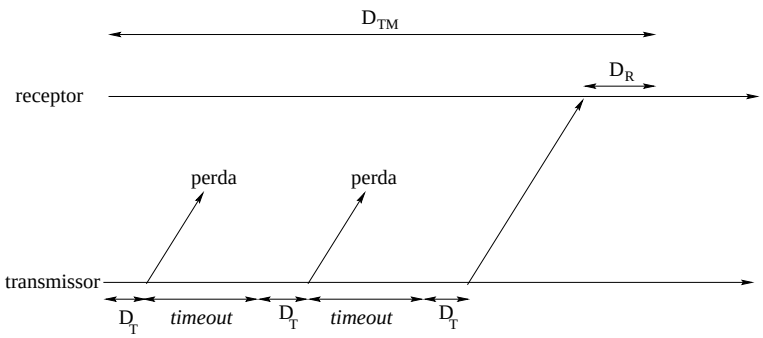

Figura 3: Operação possível para o protocolo TM.

A figura 3 mostra uma possível operação realizada por este protocolo, que é similar a desempenhada pelo protocolo orientado pelo transmissor apresentado em [16]. Nesta figura, as váriaveis D_T e D_R correspondem ao tempo de espera em *buffer* somado ao tempo de processamento no transmissor e no receptor, respectivamente. O intervalo de tempo reservado pelo temporizador será T_T. O tempo médio decorrido entre a chegada de um pacote ao *buffer* do transmissor e a recepção completa e com sucesso pelo receptor é denotada por D_{TM}. Observando-se que a probabilidade de um receptor escolhido precisar de exatamente j transmissões é $p^{j-1}(1-p)$ e que o atraso de propagação é τ, pode-se obter D_{TM}:

$$\begin{aligned} E[D_{TM}] &= \left(\sum_{j=1}^{\infty} p^{j-1}(1-p)(j(E[W_{TM}^{T}] + E[X]) + (j-1)T_T) \right) \\ &\quad +\tau + E[W_{TM}^{R}] + E[X] \\ &= \frac{(E[W_{TM}^{T}] + E[X] + pT_T)}{(1-p)} + \tau + E[W_{TM}^{R}] + E[X]. \end{aligned}$$

3.2 Atraso para o Protocolo orientado pelo Receptor com Múltiplas Cópias

Para a análise do protocolo **RM** será seguido o mesmo raciocínio desenvolvido para o protocolo **TM**, pois serão calculados os tempos de espera, para, em seguida, obter o tempo total decorrido entre a chegada do pacote ao *buffer* do transmissor e a recepção com sucesso pelo transmissor.

O transmissor processa mais uma vez dados e reconhecimentos. O fluxo de pacotes de dados é composto por transmissões de dados novos, denotado por λ_t^T, com valor λ. O conjunto de NAKs será composto por NAKs que irão determinar retransmissões e outros que, ao contrário, não as geram. A taxa de chegada de NAKs que determinam retransmissão de pacotes e a taxa de chegada do conjunto restante

de NAKs são, respectivamente, $\lambda_r^T = \lambda(E[N]-1)$ e $\lambda_n^T = \lambda(E[N^{(r)}]-E[N]-R+1)$.

Os tempos de processamento de um pacote e de um NAK correspondem às variáveis aleatórias X e Y. Todavia, os NAKs que geram retransmissões determinam um processamento no transmissor de acordo com a soma $X+Y$, enquanto a outra parcela de NAKs determina processamento simples com tempo Y. Com isso, da mesma forma que foi feito para o protocolo **TM**, pode-se calcular a carga submetida ao transmissor. Será então dada por

$$\rho_{RM}^T = \lambda E[N]E[X] + \lambda R(E[N^{(r)}]-1)E[Y]. \tag{5}$$

De posse destas relações, obtém-se, analogamente ao protocolo **TM**, os tempos médios de espera de um pacote de dados e de um reconhecimento no transmissor:

$$E[W_{RM}^T] = \frac{\lambda_t^T E[X^2] + \lambda_r^T (E[X^2]+E[Y^2]+2E[X]E[Y]) + \lambda_n^T E[Y^2]}{2(1-\rho_{RM}^T)} \; . \tag{6}$$

Para o receptor, há igualmente chegada de dados e reconhecimentos. Percebe-se que a chegada de dados ocorrerá de forma semelhante a que foi discutido na análise do protocolo **TM** e que os NAKs são gerados exclusivamente pelo próprio receptor. Como o envio do NAK será determinado após a recepção de pacote com número de sequenciamento superior, o NAK será enviado caso nenhum dos pacotes do conjunto de m réplicas tenha sido recebido com sucesso e será único para todo o conjunto. Isto significa que não haverá um fator referente às m cópias na taxa de NAKs gerados. Por isso, as taxas de chegada de dados e NAKs são, respectivamente, $\lambda_d^R = m\lambda E[N](1-p_r)$ e $\lambda_n^R = \lambda(E[N^{(r)}]-1)$.

Os respectivos tempos de serviço são X e Y, simplesmente. Com isso, pode-se obter a carga imposta ao receptor como $\rho_{RM}^R = m\lambda E[N](1-p_r)E[X] + \lambda(E[N^{(r)}]-1)E[Y]$. O tempo médio de espera em *buffer* no receptor será

$$E[W_{RM}^R] = \frac{\lambda_d^R E[X^2] + \lambda_n^R E[Y^2]}{2(1-\rho_{RM}^R)} \; . \tag{7}$$

A figura 4 mostra um possível comportamento de um pacote de dados em transmissão com o protocolo **RM**. Para a análise do tempo total decorrido entre a chegada de um pacote ao *buffer* do transmissor até o completo recebimento por parte de um receptor tomado aleatoriamente, pode-se observar duas fases distintas na operação executada pelo protocolo. Esta operação é análoga a que é realizada pelo protocolo orientado pelo receptor com prioridades apresentado em [16].

Há um tempo correspondente à primeira fase que ocorre desde a chegada do pacote em *buffer* de transmissão até o momento em que o receptor percebe a perda do pacote, por lacuna no sequenciamento de pacotes, e envia o reconhecimento negativo.

A operação prossegue com a fase de recuperação do pacote. Nesta etapa, o receptor já enviou o NAK e aguarda pela retransmissão. Pedirá novas retransmissões sucessivamente sempre que ocorrer estouro de temporizador associado ao pacote até que receba com sucesso o pacote requisitado.

O atraso médio gerado na fase de deteção de perda ($E[D_{RM}^{F1}]$) é encontrado aplicando-se raciocício similar ao desenvolvido para o protocolo com prioridades de

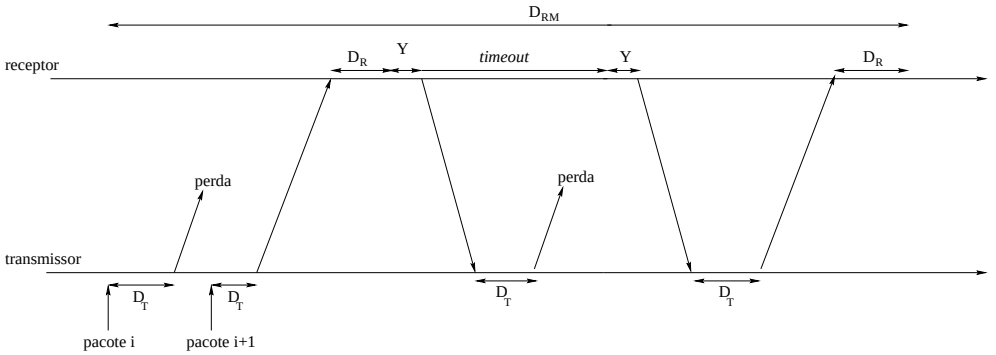

Figura 4: Possível operação para o protocolo RM.

[16]. O atraso médio gerado na segunda fase ($E[D_{RM}^{F2}]$) também segue de forma similar e é análogo a parte da operação de **TM**, como pode ser visto na figura 4. Portanto, estas duas componentes do atraso total são dadas, respectivamente, por:

$$\begin{aligned} E[D_{RM}^{F1}] &= \sum_{k=0}^{R-1} \left(\begin{array}{c} R-1 \\ k \end{array} \right) p^k (1-p)^{R-k-1} \frac{p^{k+1}/\lambda}{1-p^{k+1}} \\ &\quad + E[W_{RM}^T] + E[W_{RM}^R] + \tau + 2E[X] + E[Y] \end{aligned}$$

$$\text{e } E[D_{RM}^{F2}] = \frac{p(T_R + E[W_{RM}^R] + E[Y])}{(1-p)} + E[W_{RM}^T] + E[W_{RM}^R] + 2(E[X] + \tau).$$

Finalmente, pode-se obter o atraso total médio ocorrido neste protocolo. Com probabilidade $1-p$, o pacote de dados não é perdido e chega corretamente ao receptor escolhido aleatoriamente. Neste caso, o atraso é resultado da soma dos tempos de processamento e tempos de espera em *buffer* em transmissor e receptor com o atraso de propagação. No caso de perda de pacote, que ocorre com probabilidade p, o atraso será dado pela soma das duas parcelas identificadas pela fase de deteção da perda e recuperação do pacote. Portanto, o tempo médio decorrido desde a chegada do pacote ao *buffer* de transmissão até a completa recepção do mesmo será

$$E[D_{RM}] = (1-p)(E[W_{RM}^T] + E[W_{RM}^R] + 2E[X] + \tau) + p(E[D_{RM}^{F1}] + E[D_{RM}^{F2}]). \quad (8)$$

3.3 Atraso para o Protocolo orientado pelo Receptor com Múltiplas Cópias e Supressão de NAKs

O atraso para o protocolo **RMS** será obtido de forma similar a que foi elaborada anteriormente para os dois protocolos (**TM** e **RM**). Particularmente, boa parte de seu funcionamento é idêntica ao protocolo **RM** já que todos dois são orientados pela recepção.

Inicialmente, calcula-se os tempos de espera dos pacotes nos *buffers* de recepção e transmissão. Para o transmissor, os fluxos existentes são similares aos que existem para o protocolo **RM**. Os NAKs podem ser subdivididos em um conjunto que gera retransmissões e outro que não as gera. O número total de NAKs recebido pelo transmissor será bem menor por causa da utilização do mecanismo de supressão de NAKs. Esta quantidade corresponderá a uma variável aleatória denominada Q. Vale lembrar que este número de NAKs será também transmitido por um receptor a todos os outros receptores, pois os NAKs neste protocolo são enviados em multiponto. Logo, a taxa de chegada de NAKs ao transmissor bem como aos receptores é $\lambda E[Q]$. Em Yamamoto, Kurose e Towsley [15], determina-se uma expressão para E[Q]. Seguindo o mesmo raciocínio realizado para o protocolo **RM**, tem-se $\lambda_t^T = \lambda$, $\lambda_r^T = \lambda(E[N]-1)$ e $\lambda_n^T = \lambda(E[Q] - E[N] + 1)$

Os tempos de serviço são respectivamente X, $X + Y$ e Y. A carga (ρ_{RMS}^T)e o tempo médio de espera em *buffer* ($E[W_{RMS}^T]$) serão obtidos pela mesma expressão (5), utilizando-se estas taxas de chegada.

Para o receptor, a chegada de dados ocorre de forma similar a que acontece nos sistemas que operam com os protocolos **TM** e **RM**. Logo, a taxa de chegada de pacotes de dados é mais uma vez $\lambda_t^R = m\lambda E[N](1 - p_r)$. O tempo de processamento de cada pacote é X. Para os NAKs, a taxa de chegada já é conhecida ($\lambda E[Q]$), pois é a mesma para todos os participantes do sistema de comunicação multiponto. Como o tratamento destes NAKs possui um tempo de processamento Y e o mesmo parâmetro para pacotes de dados tem valor X, a carga imposta ao receptor e o tempo médio de espera em *buffer* são, respectivamente,

$$\rho_{RMS}^R = m\lambda E[N](1-p_r)E[X] + \lambda E[Q] \text{ e } E[W_{RMS}^R] = \frac{\lambda_t E[X^2] + \lambda E[Q]E[Y^2]}{2(1-\rho_{RMS}^R)} \ . \qquad (9)$$

A operação executada pelo protocolo é bastante similar a que acontece para o protocolo **RM**. Também são identificadas duas fases correspondentes a deteção e a recuperação do pacote. As expressões para os tempos médios relativos a estas duas fases são quase idênticas as que foram determinadas anteriormente, alterando-se apenas os tempos de espera em *buffer* para $E[W_{RMS}^T]$ e $E[W_{RMS}^R]$. A principal diferença que há na expressão para o atraso médio total está no tempo aleatório introduzido pelo mecanismo de supressão de NAKs. Este tempo será identificado por uma variável aleatória denominada V. Este tempo aleatório está localizado entre as duas fases identificadas, pois irá justamente retardar o envio do NAK que dá partida a fase de recuperação do pacote perdido. A expressão para $E[V]$ é encontrada em [15].

Finalmente, a expressão correspondente ao atraso médio total, compreendido no intervalo entre a chegada do pacote de dados ao *buffer* de transmissão e a recepção completa pelo receptor sob operação do protocolo **RMS**, será:

$$E[D_{RMS}]$$
$$= \quad E[W_{RMS}^T] + E[W_{RMS}^R] + 2E[X] + \tau + pE[V] + \frac{p^2(T_R + E[W_{RMS}^R] + E[Y])}{1-p} +$$

$$p\left(\sum_{k=0}^{R-1}\left(\begin{array}{c} R-1 \\ k \end{array}\right)p^k(1-p)^{R-k-1}\frac{p^{k+1}/\lambda}{1-p^{k+1}}+E[W_{RMS}^T]+E[W_{RMS}^R]+2(E[X]+\tau)\right).$$

4 Resultados Numéricos

Com a modelagem dos sistemas envolvidos em cada protocolo considerado, realiza-se uma análise quantitativa do atraso gerado sob várias condições em que se utilizam os protocolos genéricos propostos. Para efeito de comparação e percepção do efeito acarretado pela introdução de transmissão por múltiplas cópias, são considerados, paralelamente, aos protocolos **TM**, **RM** e **RMS** os que são apresentados em [13, 15]: **A**, orientado pelo transmissor; **N1**, orientado pelo receptor simplesmente; e **N2**, orientado pelo receptor com mecanismo de supressão de NAKs. A fim de que se possa ter noção da escalabilidade obtida, foram medidos, ainda, os efeitos gerados pelo número de participantes no sistema, variando-se o número de receptores.

Nos cálculos efetuados, considerou-se tempos de processamento constantes. Kasera, Kurose e Towsley [7] mediram para o tempo de processamento de pacote o valor 500 μs e para processamento de ACK/NAK o tempo de 100 μs em ambiente com rede Ethernet, protocolos IP/UDP e estação com processador Pentium 133MHz com sistema operacional Linux. Normalizando-se estes tempos, tem-se $E[X] = 1.0$ e $E[Y] = 0.2$. O atraso de propagação considerado foi de 10ms, assumindo, portanto, valor normalizado $\tau = 20.0$.

O gráfico da figura 5 apresenta o atraso médio gerado nos sistemas com os protocolos considerados em relação ao número de receptores para probabilidade de perda 0.1 e taxa de chegada de dados $\lambda = 0.1$. Pode-se verificar, comparando-se as curvas dos protocolos **A** e **TM**, um atraso levemente menor com a operação do protocolo **TM**. Contudo, em termos de escalabilidade, não há ganho, visto que operam de forma estável em uma faixa estreita onde há pequeno número de receptores. Entretanto, ao observar as curvas dos protocolos **N1** e **N2**, que são orientados pelos receptores, percebe-se que a faixa de operação estável é bem mais ampla do que a faixa obtido na curva dos protocolos **A** e **TM**. O protocolo **N2**, por usar supressão de NAKs, abrange uma faixa maior com estabilidade, sofrendo menos os efeitos do problema de implosão de NAKs. Portanto, a opção de orientação pelos receptores permite maior escalabilidade aos sistemas em questão. Ao observar-se as curvas obtidas para os protocolos **RM** e **RMS**, constata-se que operam estavelmente na faixa analisada. Até 10000 receptores, o atraso médio observado tem leve variação. Além disso, enquanto o protocolo **N1** tem limitações já a partir de valores pouco superiores a 1000 receptores, o protocolo **RM** mantém atraso praticamente constante na faixa inteira observada. Apenas ao aproximar-se de 10000 receptores, o atraso médio inicia uma variação positiva. Em relação aos protocolos **N2** e **RMS**, as mesmas observações podem ser colocadas, com a exceção de que, mesmo próximo a 10000 receptores, o atraso não sofre aumento considerável.

Outra observação interessante a ser observada no gráfico da figura 5 é o declínio

da curva do atraso médio nos protocolos **N2** e, menos acentuadamente, **RMS** em faixas inferiores de número de receptores. Tal fato decorre do funcionamento do mecanismo de supressão de NAKs. Ao introduzir um intervalo de tempo aleatório, para um número menor de receptores, relativamente, o tempo a esperar até que o primeiro NAKs seja retornado será maior do que quando há quantidade maior de receptores. Isto determina a inclinação negativa para as curvas analisadas neste caso.

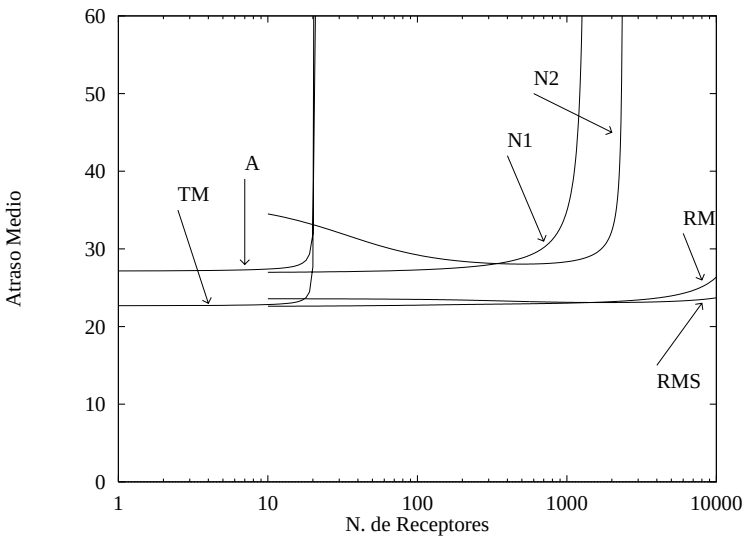

Figura 5: Atraso médio versus número de receptores com $p = 0.1$, $\lambda = 0.1$ e $m = 2$.

O gráfico da figura 6(a) apresenta o atraso médio em relação ao número de receptores no sistema multiponto. As curvas traçadas permitem observar as mesmas comparações que foram discutidas para o gráfico da figura 5. O que se observa de novo é a maior limitação existente com a utilização de uma probabilidade mais elevada. Todas as curvas apresentam faixas de operação estáveis menores do que as obtidas para $p = 0.1$. Particularmente, para os protocolos **RM** e **RMS**, a limitação em atraso médio ocorre, na faixa considerada, próximo a 2000 e 3000 receptores. O efeito de declínio do atraso em valores inferiores de número de receptores para os protocolos **N2** e **RMS**, em decorrência da supressão de NAKs, torna-se mais acentuado neste caso. Como há uma chance maior de perda de pacote neste caso, a parcela relativa ao atraso aleatório tem maior influência no atraso médio gerado. Também é interessante notar a razão entre os atrasos médios gerados em **A** e **TM** e em **N1** e **RM**. Com um valor aproximadamente 1.48, esta razão mostra a diminuição de atraso proporcionada pela transmissão em avanço.

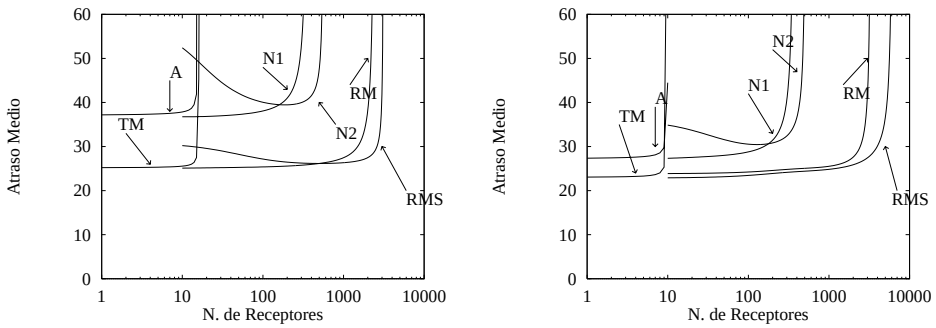

Figura 6: Atraso médio versus número de receptores: (a)$p = 0.25$, $\lambda = 0.1$ e $m = 2$, (b)$p = 0.1$, $\lambda = 0.2$ e $m = 2$.

O gráfico da figura 6(b) mostra o atraso médio gerado nos sistemas com os mesmos protocolos abordados em probabilidade de perda $p = 0.1$, taxa de geração de pacotes $\lambda = 0.2$ e $m = 2$. As comparações entre os protocolos permitem repetir as conclusões observadas anteriormente nos dois gráficos discutidos. De fato, o aumento da taxa de geração de pacotes irá acarretar em restrições maiores devido ao problema de implosão de reconhecimentos, uma vez que a taxa de perda não alterou em relação ao gráfico da figura 5, mas o aumento na quantidade de pacotes provoca o efeito analisado. É interessante observar que a razão entre os atrasos médios em faixas inferiores, discutida no gráfico da figura 6(a) correspondem aproximadamente aos mesmos valores obtidos para o gráfico da figura 5, como consequência do número de cópias enviadas em avanço e da probabilidade de perda, já que o número de receptores não influi de forma crítica nesta situação.

Com a utilização de probabilidades de perda baixas, caso da figura 7(a) onde $p = 0.01$, os efeitos são pouco notados pois as curvas estão muito próximas, à exceção dos protocolos orientados pelo transmissor **TM** e **A** que continuam sendo limitados para valores bem inferiores de número de receptores. Os protocolos restantes apresentam um comportamento constante na faixa considerada. Apenas próximo a 10000 receptores a curva de **N1** começa a sofrer variação positiva mais acentuada.

Por fim, um último levantamento interessante foi realizado com a utilização de 3 cópias enviadas em avanço. Foram obtidos os resultados númericos para todos os protocolos abordados. As curvas obtidas para os protocolos **RM** e **RMS** estão mostradas na figura 7(b), juntamente com as mesmas curvas apontadas anteriormente para o caso em que $m = 2$. A faixa de número de receptores considerada no gráfico figura entre 1000 e 10000 receptores e, como se pode notar, apresentam valores muito próximos.

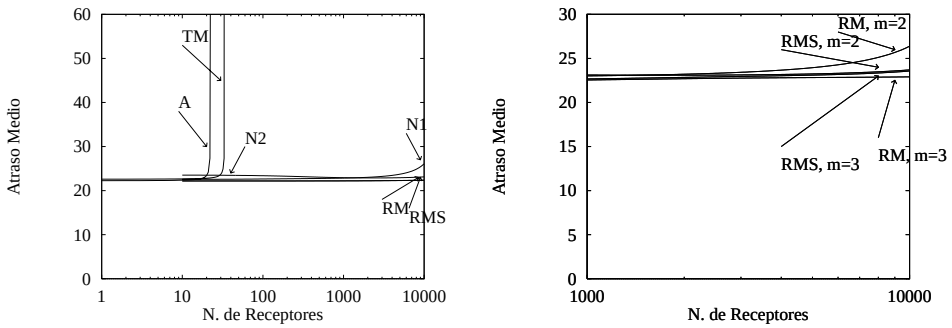

Figura 7: Atraso médio versus número de receptores: (a)$p = 0.01$ e $\lambda = 0.1$, (b)$p = 0.1$, $\lambda = 0.1$, $m = 2$ e $m = 3$.

5 Conclusões

Este trabalho analisa o desempenho de protocolos multiponto confiáveis que utilizam estratégia de transmissão por múltiplas cópias. O objetivo das múltiplas cópias é obter escalabilidade para os sistemas que operem sob tais protocolos com alto desempenho, por meio de garantias de qualidade de serviço. Esta estratégia permitiu a definição de protocolos genéricos que utilizam abordagens distintas, servindo-se de conceitos como orientação por transmissor e receptor e supressão de NAKs.

Foi realizada a avaliação de desempenho para os três modelos associados aos protocolos. A análise realizada objetivou o parâmetro atraso médio gerado. A distribuição do tempo de processamento de pacotes e de reconhecimentos não possui restrição. Estes modelos consideram que as chegadas de pacotes aos *buffers* do transmissores dos sistemas ocorrem segundo distribuição de Poisson.

Os resultados númericos obtidos apresentam desempenho superior dos protocolos que adotam a replicação de pacotes em relação aos que não a adotam. Este efeito foi notado consideravelmente com probabilidades de perda mais altas. Com probabilidades de perda mais baixas os protocolos possuem atraso médio bastante próximo em faixas de receptores que variam até 10000 receptores. O aumento da probabilidade de perda pode acarretar em limitação na escalabilidade. Um possível aumento na taxa de geração de pacotes gera o mesmo efeito, de forma que protocolos que operem em faixas maiores permanecendo estáveis são altamente desejáveis.

Os protocolos definidos com a adoção de transmissões de cópias permitiram maior faixa de número de receptores com estabilidade assegurada. Por isso, a estratégia proposta neste trabalho adequou-se bem ao objetivo de maior alcance de escalabilidade. Ainda, além de reduzir os efeitos do problema de implosão de reconhecimentos, também há uma diminuição do atraso permitida pelo envio de cópias em avanço dadas condições em que a taxa de perda é relativamente alta.

Referências

[1] H. Eriksson, "Mbone: The multicast backbone", *Communications of the ACM*, vol. 37, pp. 54–60, agosto 1994.

[2] J. Nonnenmacher e E. W. Biersack, "Performance modelling of reliable multicast transmission", in *Proc IEEE INFOCOM'97*, Kobe, abril 1997.

[3] B. N. Levine e J. J. Garcia-Luna-Aceves, "A comparison of known classes of reliable multicast protocols", in *Proceedings of International Conference on Network Protocols*, Columbus, outubro 1996.

[4] J. C. Lin e S. Paul, "RMTP: A reliable multicast transport protocol", in *Proc. IEEE INFOCOM'96*, pp. 1414–1424, março 1996.

[5] J. Liebeherr e B. S. Sethi, "A scalable control topology for multicast communications", in *IEEE INFOCOM'98*, abril 1998.

[6] M. Hofmann, "Adding scalability to transport level multicast", in *3rd COST 237 Workshop - Multimedia Telecommunications and Applications*, Barcelona, novembro 1996.

[7] S. Kasera, J. Kurose e D. Towsley, "Scalable reliable multicast using multiple multicast groups", in *Proceedings of 1997 ACM Sigmetrics Conference*, 1997.

[8] J. Nonnenmacher e E. W. Biersack, "Parity-based loss recovery for reliable multicast transmission", in *Proc. SIGCOMM'97*, Cannes, setembro 1997.

[9] J. Nonnenmacher, E. W. Biersack e D. Towsley, "Reliable multicast: Where to use forward error correction", in *Proc. 5th Workshop on Protocols for High Speed Networks*, Sophia Antipolis, France, outubro 1996.

[10] H. M. de Lima e O. C. M. B. Duarte, "An effective selective repeat ARQ strategy for high speed point-to-multipoint communications", in *IEEE Global Telecommunications Conference GLOBECOM'96*, pp. 1059–1063, novembro 1996.

[11] S. Floyd, V. Jacobson, S. McCanne, C. G. Liu e L. Zhang, "A reliable multicast framework for light-weight sessions and application level framing", *IEEE/ACM Transactions on Networking*, pp. 784–803, dezembro 1997.

[12] M. D. Nunes e O. C. M. B. Duarte, "MCF: Incorporando escalabilidade, mobilidade e novas estratégias de implementação de alto desempenho", in *II Seminário Franco-Brasileiro de Sistemas Informáticos Distribuídos*, novembro 1997.

[13] S. Pingali, D. Towsley e J. Kurose, "A comparison of sender-initiated and receiver initiated reliable multicast protocols", *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 3, abril 1997.

[14] S. Ramakrishnan e B. N. Jain, "A negative acknowledgement with periodic polling protocol for multicast over LAN", in *Proc. IEEE INFOCOM '87*, pp. 502–511, março 1987.

[15] M. Yamamoto, J. F. Kurose, D. F. Towsley e H. Ikeda, "A delay analysis of sender-initiated and receiver-initiated reliable multicast protocols", in *Proc. of IEEE INFOCOM'97*, 1997.

[16] D. A. M. Villela e O. C. M. B. Duarte, "Mecanismos com prioridades para escalabilidade e desempenho em comunicações multiponto confiáveis", in *XVI Simpósio Brasileiro de Redes de Computadores*, pp. 301–318, maio 1998.

[17] D. A. M. Villela e O. C. M. B. Duarte, "A priority scheme applied to scalable reliable multicast communications", tech. rep., Grupo de Teleinformática e Automação/UFRJ, 1998.

[18] A. R. K. Sastry, "Improving repeat-request (ARQ) performance on satellite channels under high error rate conditions", *IEEE Transactions on Communications*, vol. COM-23, no. 4, pp. 436–439, abril 1975.

[19] N. D. Birrel, "Preemptive retransmission for communication over noisy channels", *IEEE Proceedings*, vol. 128, pp. 393–400, novembro 1981.

[20] L. Kleinrock, *Queueing Systems Vol. 1: Theory.* Wiley Interscience, 1974.