Análise Temporal sobre a Diferenciação de Tráfego em Redes Ethernet Micro-segmentadas

Rafael Pinto Costa[1] **, João Cesar Netto**[1]

[1]Instituto de Informática – Universidade Federal do Rio Grande do Sul
Caixa Postal 15064 – 90501-970 Porto Alegre, RS

rcosta@inf.ufrgs.br, netto@inf.ufrgs.br

Abstract. *This work aims to evaluate, through a deterministic analysis, the utilization of switched ethernet networks on distributed control systems. The analysed environment considers the use of traffic differentiation mechanisms inserted in IEEE standards for local networks since 1998. On top of this structure it is proposed a calculus analytic model capable to predict the maximum, average and minimum values of end-to-end delay on different traffic classes communications.*

Resumo. *Este trabalho procura avaliar, através de uma análise determinística, a utilização de redes ethernet micro-segmentadas em ambientes de controle distribuído. O ambiente analisado considera o emprego dos mecanismos de diferenciação de tráfego introduzidos nos padrões IEEE para redes locais desde 1998. Sobre esta estrutura propõe-se um modelo analítico de cálculo capaz de prever os valores máximos, médios e mínimos do atraso fim-a-fim observado nas comunicações de diferentes classes de tráfego.*

1. Introdução

Quase todos sistemas críticos e a maioria dos sistemas de computação embarcada são sistemas de tempo real [Stankovic et al., 1996]. O funcionamento correto destes sistemas depende não apenas de resultados lógicos, mas também do tempo no qual estes resultados são produzidos.

Anos de desenvolvimento prático e de pesquisa demonstram que o fornecimento de garantias para o atendimento aos requisitos temporais que emergem destes cenários não é uma tarefa trivial [Stankovic, 1988]. Sob a ótica da computação, o projeto destes sistemas, especialmente os críticos, requer que os conceitos previsibilidade e determinismo sejam considerados em todas camadas da arquitetura empregada.

As redes de comunicação são figuras extremamente relevantes neste contexto, pois muitas das aplicações envolvidas englobam a comunicação entre entidades espacialmente separadas no ambiente local. Este é o caso geral na área de automação, industrial ou predial/residencial por exemplo, com o controle em rede utilizado para coordenação de sensores e atuadores.

Algumas das métricas de desempenho que impactam diretamente na aplicação de tecnologias de rede em sistemas de controle incluem: tempo de acesso, tempo de transmissão, tempo de resposta, tempo de mensagem, percentual de colisões, vazão e tamanho

das mensagens, além do percentual de utilização da rede e determinismo. A atenção a estas métricas pode ser resumida em duas características fundamentais: atraso temporal limitado para as comunicações e transmissão sem perdas [Lian et al., 2001]. Isto significa que as mensagens devem ser corretamente transmitidas dentro de um limite temporal previsível. Transmissões sem sucesso ou com tempo muito elevado entre sensores e atuadores podem degradar o desempenho de um sistema ou levá-lo à instabilidade.

Neste sentido, diversas tecnologias dedicadas de rede foram desenvolvidas para atender a estas exigências, basicamente através do suporte à troca freqüente de pequenas quantidades de dados [Thomesse, 1999]. Genericamente conhecidas como *fieldbuses*, os exemplos mais conhecidos são: TTP/C, CAN, DeviceNet, ProfiBus, WorldFip, P-Net e Foundation FieldBus [ISO, 1999, ISO, 1992] . Embora adequadas em muitos casos, estas tecnologias apresentam algumas desvantagens como: alto custo, dificuldade de integração com outras tecnologias de rede, incompatibilidades entre dispositivos de fabricantes distintos e escassez na banda disponível [Decotignie, 2001].

Devido a estas limitações, alguns protocolos de uso geral passaram a ser cogitados como alternativas para emprego em sistemas com características temporais, como FDDI e ATM [Song, 2001] e, principalmente, a tecnologia mais utilizada no âmbito local: redes tipo *ethernet* [Shimokawa and Shiobara, 1985, Court, 1992, Chiueh and Venkatramani, 1994].

Esta tendência é fortemente presenciada no controle de processos industriais, com o emprego massivo de redes *ethernet* observado nos últimos anos [Kaplan, 2001, Caro, 2001] . Entretanto, sua efetiva utilização requer a solução para o problema do indeterminismo, característico da tecnologia, originado no mecanismo de arbitragem (colisões) e no processo de enfileiramento [Wheelis, 1993, Khanna and Singh, 1994]. Entre os trabalhos desenvolvidos na análise e na adequação desta estrutura estão: [Lee and Lee, 2002], [Hoang et al., 2002], [Hoang and Jonsson, 2003], [Jasperneite and Neumann, 2001], [Jasperneite et al., 2002], [Bello and Mirabella, 2001],[Steffen et al., 2003],[Moldovansky, 2002] e [Song, 2001].

O objetivo principal deste trabalho é a avaliação de um modelo que contemple soluções para estes dois problemas básicos da tecnologia *ethernet* em aplicações desta natureza:

- ausência de colisões com o uso de *switched ethernet*; e
- previsibilidade no enfileiramento a partir do emprego dos mecanismos de qualidade de serviço previstos na padronização de redes locais IEEE 802.1D/Q. As bases destes mecanismos são a diferenciação de tráfego prioritário e o escalonamento de prioridade absoluta entre filas distintas.

Este artigo está organizado em 6 seções. Na Seção 2 são apresentadas as características básicas das normas IEEE 802.1D/Q fundamentais para a análise desenvolvida. O modelo analítico de cálculo para o atraso das comunicações é deduzido na Seção 3, enquanto a Seção 4 ilustra sua aplicação em dois casos de estudo específicos. A validação do modelo é exposta na Seção 5, confrontando os resultados teóricos e aqueles obtidos por simulação. Por último, a Seção 6 exibe as conclusões e considerações finais a respeito da análise apresentada.

2. Mecanismos IEEE 802.1 D/Q

A preocupação com a transmissão de informações críticas em relação a variável temporal foi introduzida na família de protocolos para redes locais IEEE através da especificação 802.1p, incorporada na edição de 1998 da norma IEEE 802.1D [IEEE, 1998].

Entre outras especificações, a norma 802.1D apresentou a idéia de classificação do tráfego e definiu, de modo independente da tecnologia de nível físico empregada, o comportamento esperado das entidades intermediárias [1]. Embora baseada na diferenciação por sinalização explícita, não apresentou novos formatos para os quadros de tecnologias que não possuíssem a capacidade de carregar informações de prioridade, como por exemplo, *ethernet*.

O complemento necessário veio também em 1998 através do padrão IEEE 802.1Q [IEEE, 2003] que estendeu o conceito dos comutadores, adicionando funcionalidades para o suporte e gerenciamento de redes locais virtuais (VLANs). Entre as definições apresentadas estava um novo formato de quadro capaz de carregar informações de identificação de VLAN e de prioridade de usuário. No caso de redes tipo *ethernet*, estas informações foram incluídas em um novo campo de cabeçalho inserido imediatamente após os campos de endereços origem e destino do quadro original (Figura 1).

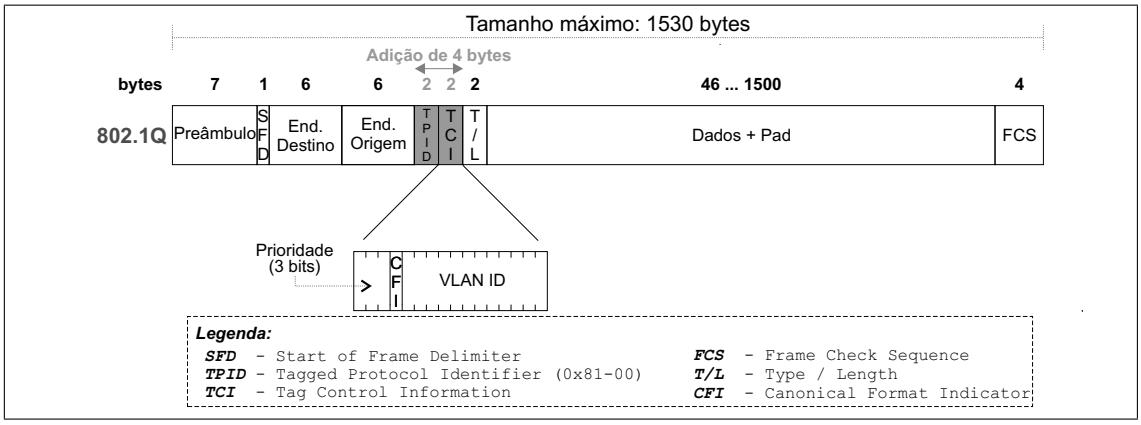

Figura 1: Novo formato de quadro introduzido

Nesta nova arquitetura, três características são fundamentais para a análise temporal das comunicações:

Enfileiramento dos quadros : quadros são classificados de acordo com o valor sinalizado em seu campo de prioridade. Cada classe definida é associada a uma fila independente.

Escalonamento de prioridade absoluta : quadros de determinada fila são selecionados para transmissão apenas quando as filas de prioridade superior estão vazias.

Algoritmo *Spanning Tree* : o emprego deste protocolo entre os diversos comutadores da rede faz com que topologias arbitrárias sejam percebidas, do ponto de vista lógico funcional, sob a forma de árvores[2]. Sua utilização garante a existência de

[1] entidades de repasse de quadros – *switches* ou comutadores.

[2] Pela teoria dos grafos *spanning tree* é uma árvore de arcos que se estende por um grafo, mantendo sua conectividade sem conter caminhos fechados.

no máximo uma rota entre duas estações quaisquer, preservando-a na ausência de falhas físicas.

3. Modelo Analítico para Cálculo do Atraso das Comunicações em Redes *Ethernet*

O modelo analítico apresentado nesta seção considera um sistema distribuído que agrega tanto tráfego de controle (NCS - *Network-based Control Systems*) [Walsh and Hong, 2001, Chow and Tipsuwan, 2001], quanto tráfego sem requisitos temporais (rede de dados). Seu objetivo é a obtenção do perfil de temporal das comunicações, através da previsão de valores mínimos, médios e máximos para o atraso fim-a-fim (*"One-Way Delay – OWD"* [Almes et al., 1999]) observado por diferentes classes de quadros.

3.1. Premissas

Topologia: em sincronia com a exigência cada vez maior por desempenho e sua ampla utilização, assume-se que as redes estão totalmente estruturadas ao redor *switches* (comutadores). A interconexão das n estações presentes na rede se dá no maior grau de segmentação possível, ou seja, cada porta de um comutador possui apenas uma estação conectada ou outro *switch* encadeado.

Switches: tipo *store-and-forward* e compatíveis com as normas IEEE 802.1D/Q, isto é: [1] utilização do formato de quadro introduzido pela norma IEEE 802.1Q (VLANs), admitindo a sinalização explícita de prioridade (3 bits do cabeçalho - Figura 1); [2] enfileiramento nas portas de saída; [3] filas independentes para classes de tráfego distintas (máximo de oito níveis)[3]; [4] escalonamento de prioridade absoluta entre as filas; [5] compatibilidade com o protocolo *Spanning Tree*;[6] .

Modelo de tráfego: o padrão de geração de quadros adotado para a rede de controle é periódico e com comunicações baseadas no modelo mestre-escravo. A adoção de um modelo mestre-escravo implica na presença na rede de pelo menos uma estação controladora ou mestra, responsável por receber os quadros com relatórios de estado enviados pelas estações escravas. Este formato de comunicação é coerente com o perfil de redes de automação, sendo utilizado freqüentemente na coordenação de sensores e atuadores [Chow and Tipsuwan, 2001].

Já em relação à rede de dados, nenhuma restrição é imposta sobre o perfil de tráfego utilizado. Contudo os quadros desta classe são invariavelmente considerados como de prioridade mais baixa, recebendo tratamento tradicional de melhor esforço (*best effort*).

3.2. Cálculo do Atraso Observado nas Comunicações

A operação estável de uma rede *ethernet* baseada em *switches* está atrelada a algumas condições de contorno, conforme apresentado em [Lee and Lee, 2002], fundamentais para validação da análise apresentada.

Em primeiro lugar há de se observar que o tráfego total presente na rede deve ser compatível com a capacidade dos comutadores. Isto é expresso pela relação 1, onde:

[3] o mapeamento admitido entre prioridades e classes é um para um, com cada valor de prioridade associado a uma classe distinta. Com isto os termos classes e prioridades passam a representar sinônimos.

- $Capa_s(m)$ = capacidade do *switch* m em termos de quadros que ele é capaz de processar por unidade de tempo;
- n = o número de estações diretamente conectadas ao comutador m; e
- $Quadros_i$ = número máximo de quadros produzidos pela estação i por unidade de tempo.

$$Capa_s(m) \geq \sum_{i=0}^{n} Quadros_i, \qquad \text{para todo m} \tag{1}$$

A principal conseqüência desta relação é que cada comutador deve ser capaz de processar o número máximo de quadros gerados pelas entidades vizinhas em uma unidade de tempo. Considera-se que a capacidade de *backplane* dos *switches* é suficiente para suportar o tráfego gerado, não prejudicando o tempo de repasse dos quadros.

Uma segunda condição importante diz respeito a capacidade das linhas de transmissão. A soma do tráfego endereçado a determinada estação, a cada período, deve ser menor que a capacidade da linha de recepção desta mesma estação, conforme apresentado na relação 2.

$$Capa_e(j) \geq \sum_{i=0}^{n} Quadros_i(j), \qquad \text{para todo j} \tag{2}$$

em que:

- $Capa_e(j)$ = taxa da linha de transmissão entre o *switch* e a estação j;
- $Quadros_i(j)$ = número de bits a serem transmitidos, numa unidade de tempo, da estação i para estação j.

Caso esta relação (vazão do canal) não seja satisfeita, a fila formada no comutador para o canal de saída de uma estação específica pode crescer indefinidamente, mesmo que a capacidade do *switch* seja suficiente para processar todo o tráfego.

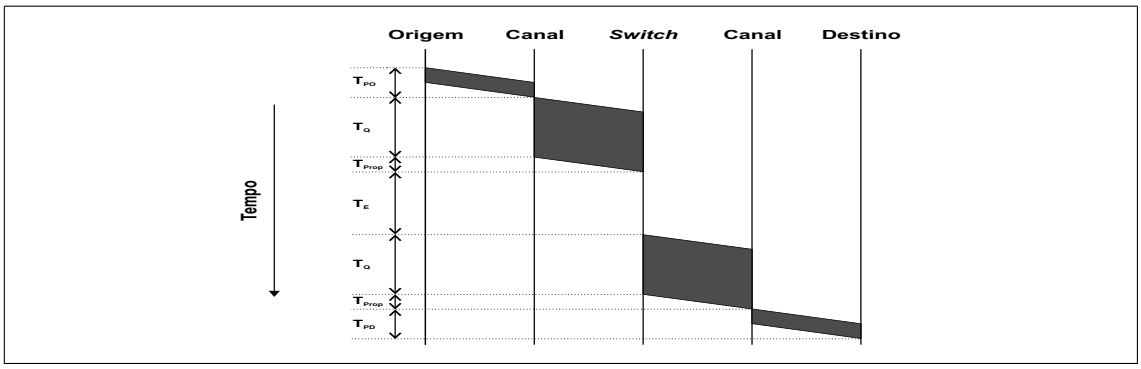

Figura 2: Diagrama temporal – componentes de A_T

Com a rede em operação estável, satisfeitas a relações 1 e 2, o atraso total fim-a-fim (A_T) experimentado por quadros de prioridade p ao transitarem entre duas estações distintas quaisquer, pode ser definido como a soma de várias parcelas (Figura 2) – equação 3.

$$A_T(p) = T_{PO} + T_Q + T_{Prop} + T_E(p) + T_{PD} \quad (3)$$

na qual:

- T_{PO} e T_{PD} = tempos de processamento consumidos nas estações origem e destino, respectivamente;
- T_Q = tempo de quadro, representa o tempo necessário para injeção dos quadros em cada linha de transmissão percorrida;
- T_{Prop} = tempo de propagação dos quadros através das linhas. Está diretamente relacionado com a velocidade de propagação e com a distância entre as estações e *switches*;
- T_E = Tempo de enfileiramento que cada quadro aguarda nas filas possivelmente formadas ao longo do caminho (*switches*). Este tempo é função da prioridade/classe de cada quadro, pois esta informação determina sua fila de espera.

O tempo gasto com o processamento das mensagens, T_{PO} e T_{PD}, são tipicamente constantes e dependem exclusivamente da capacidade de processamento das estações. Não apresentam relação direta com as condições físicas da rede ou com peculiaridades do protocolo empregado e, portanto, são ignorados na discussão apresentada. Estes dois termos são suprimidos na equação 4.

$$A_T(p) = T_Q + T_{Prop} + T_E(p) \quad (4)$$

O cálculo de T_Q compreende a soma do tempo gasto com a inserção do quadro em cada um dos k canais percorridos em determinada comunicação, como exposto pela equação 5.

$$T_Q = \sum_{c=1}^{k} \frac{Tam_q}{Taxa(c)} \quad (5)$$

onde Tam_q é o tamanho em bits do quadro considerado, k o número de canais percorridos e $Taxa(c)$ a taxa de transferência característica do canal c, expressa em bits/segundo (bps).

Da mesma forma é preciso considerar o tempo gasto por cada bit durante seu trânsito (T_{Prop}) através de toda extensão dos canais – equação (6).

$$T_{Prop} = \sum_{c=1}^{k} \frac{Comp(c)}{C} \quad (6)$$

em que:

- $Comp(c)$ = comprimento em metros de cada canal;
- C = velocidade de propagação do sinal no meio (algo em torno de $2 \times 10^8 m/s$)[4].

[4]Aproximadamente 2/3 do velocidade de propagação das ondas eletromagnéticas no vácuo ($3 \times 10^8 m/s$)[Tanenbaum, 1996]

Resta portanto a definição da parcela correspondente ao tempo gasto com o enfileiramento nos *switches*. Esta parcela pode agregar grande variação ao valor final de A_T, fato que ressalta sua importância em sistemas tempo dependentes.

Em um ambiente com filas isoladas para cada uma das classes de tráfego, o tempo de contenção de um quadro com prioridade p, em cada comutador, depende do número de quadros (N_Q) na sua própria fila e também nas filas de maior prioridade, como definido na equação 7.

$$T_E(p) = \sum_{i=0}^{p} (N_Q(i) \times T_{trans}) \tag{7}$$

O índice i representa a inspeção da fila p e de todas prioritárias a esta ($0 \leq i \leq p$), considerando o valor zero como a maior prioridade no sistema ($p0$). O tempo de transmissão gasto com cada quadro presente nestas filas (T_{trans}) é dado pela equação 8.

$$T_{trans} = \frac{Tam_q}{Taxa(c)} + T_{EQ} \tag{8}$$

T_{EQ} representa o tempo mínimo que deve ser mantido sem transmissão entre quadros consecutivos. Este tempo é definido no padrão como o tempo necessário para injeção de 96 bits no canal e constitui o intervalo necessário para recuperação dos níveis inferiores da arquitetura, além de impedir, no caso de acesso múltiplo, o monopólio sobre o meio de transmissão. Com uma taxa de operação de 10 Mbps $T_{EQ} = 9,6\mu s$, e para 100 Mbps $T_{EQ} = 0,96\mu s$.

A melhor hipótese possível em busca do menor valor de T_E é que não ocorra enfileiramento ao longo do caminho. Para um quadro de prioridade p isto significa que o tamanho da própria fila de prioridade p e das filas de maior prioridade seja nulo, isto é:

$$N_Q(i) = 0, \qquad \text{para todo i} \qquad 0 \leq i \leq p \tag{9}$$

Entretanto, é importante observar que ainda assim, sem enfileiramento, dependendo da carga na rede e da topologia adotada, o tempo de contenção nos *switches* pode não tender a zero. Isto ocorre porque existe a possibilidade de que um quadro, mesmo de prioridade inferior, já esteja em processo de transmissão quando uma mensagem prioritária alcança o comutador. Como não existe previsão de preempção deste processo, um quadro qualquer pode ser obrigado a aguardar, a cada *switch*, o final da transmissão de uma outra mensagem de tamanho arbitrário. Em uma rede padrão IEEE 802.3 o menor quadro possível é de 576 bits (72 bytes) e o maior 12240 bits (1530 bytes), incluindo os bits de preâmbulo e sinalizador de início de quadro.

Trata-se de um caso possível de inversão de prioridades cujos efeitos sobre o valor final de A_T não são desprezíveis. Esta situação exige que a equação 7 seja redefinida de modo a considerar este tempo residual (T_{Res}), possivelmente gasto com uma transmissão já em andamento.

$$T_E(p) = \sum_{i=0}^{p} \left(N_Q(i) \times T_{trans} \right) + T_{Res} \tag{10}$$

À medida que a carga na rede aumenta, mais relevante pode se tornar o valor de T_{Res}. Havendo disputa por canais entre classes distintas (concentração em linhas de saída), maior será a probabilidade de que um quadro de prioridade inferior já esteja em processo de transmissão. Redes com alta carga de tráfego são o foco da análise aqui apresentada, pois estudos prévios [Wheelis, 1993][Lian et al., 2001] já demonstraram a possibilidade do emprego de redes *ethernet* em aplicações de controle, quando a carga total presente é mantida em um valor bem baixo.

4. Aplicação - Cenários de Estudo

Para exemplificar o emprego do modelo de cálculo apresentado, é necessário definir alguns parâmetros principais do ambiente. Com este objetivo são considerados dois cenários hipotéticos nos quais se considera, de acordo com as premissas apresentadas anteriormente, o compartilhamento da estrutura de rede entre aplicações de controle e de dados (sem requisitos temporais). O emprego do modelo é direcionado para temporização dos quadros da rede de controle no sentido estações → controlador.

4.1. Caso 1: N Estações x 1 *Switch* - 2 Prioridades

Neste primeiro caso a rede está estruturada ao redor de apenas um comutador, com todas entidades diretamente conectadas a ele (Figura 3).

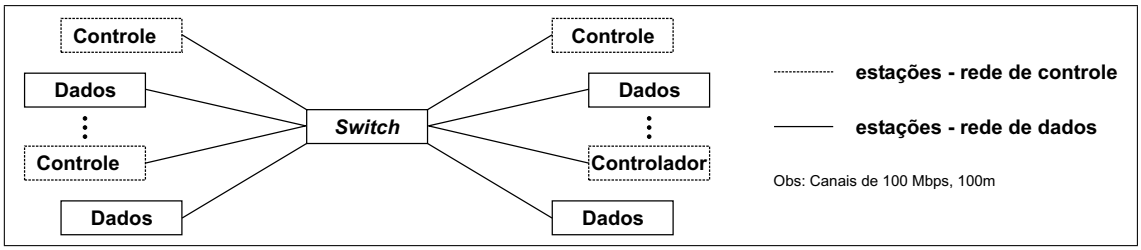

Figura 3: Caso 1 – Topologia

A tabela 1 define os principais parâmetros adotados para a rede de controle.

Tabela 1: Parâmetros principais da rede de controle – Caso 1

Descrição	Valor
Número de entidades	15 estações e um controlador
Taxa de amostragem	1000 quadros/s
Tamanho dos quadros	576 bits

O primeiro ponto a ser verificado é a operação estável da rede - relações 1 e 2. A transmissão periódica de 1000 quadros a cada segundo, efetivada por cada uma das 15 estações da rede de controle, corresponde a um período de amostragem de $1\ ms$ e

produz um total de 15000 quadros por unidade de tempo. Esses quadros com tamanho de 576 bits (Tabela 1) geram o equivalente a uma taxa de $8.64\ Mbps$ ($576\ bits/quadro \times 1000\ quadros/s \times 15$ estações) destinada a estação de controle; um valor bem abaixo da taxa do canal ($100\ Mbps$) e, portanto, plenamente compatível com a relação 2. O período de $1ms$ adotado para amostragem na rede de controle satisfaz a grande maioria das aplicações [Lee and Lee, 2002].

O valor de 15000 quadros por segundo é muito inferior também a capacidade dos *switches* atuais que é da ordem de milhões de mensagens por segundo [3Com, 2003]. Portanto, na verdade não há problemas para suportar o tráfego gerado pelas outras estações pertencentes a rede de dados, satisfazendo facilmente a relação 1. Por exemplo, considerando uma taxa de geração de 5000 quadros/s nas estações da rede de dados, um valor cinco vezes maior do que na rede de controle, ainda seriam necessárias 197 estações ativas para transpor a barreira de $1.000.000$ quadros/s.

Como existe apenas um *switch* na rede, toda mensagem ao trafegar entre duas estações distintas quaisquer irá transpor duas linhas de transmissão idênticas: origem $\rightarrow$ *switch* e *switch* $\rightarrow$ destino. Por aplicação direta das equações 5 e 6, para os quadros da rede de controle:

$$T_Q = 2 \times \left(\frac{576\ bits}{100 \times 10^6\ bits/s} \right) = 11,52\mu s \qquad T_{Prop} = 2 \times \left(\frac{100\ m}{2 \times 10^8\ m/s} \right) = 1\mu s$$

Para o cálculo da parcela variável que compõe o atraso (T_E), é importante observar que a cada período de transmissão, o pior cenário para uma estação específica ocorre quando o seu quadro é o último a atingir o *switch*. Nesta situação e considerando a transmissão praticamente simultânea de 15 estações de mesma prioridade, o tamanho máximo da fila possivelmente encontrado por um quadro é de 14 mensagens ($N_{Q_{max}}(p0) = 14$).

O emprego da topologia apresentada na figura 3, onde o controlador não participa da rede de dados, impossibilita que um quadro com prioridade superior seja contido por uma transmissão já iniciada de um quadro com menor prioridade, independentemente da carga de tráfego presente. A inexistência de canais compartilhados entre as redes de dados e de controle, somada ao enfileiramento independente em cada porta do *switch*, faz com que estas duas classes de tráfego fiquem completamente isoladas, levando o valor do tempo de contenção residual para zero ($T_{Res} = 0$).

Com estes dados, através das equações 8 e 10:

$$T_{trans} = \frac{576\ bits}{100 \times 10^6\ bits/s} + 0,96\ \mu s$$
$$T_{trans} = 6,72\ \mu s$$

$$T_{Emax}(p0) = (14 \times T_{trans}) + T_{Res}$$
$$= 14 \times 6,72\mu s + 0$$
$$T_{Emax}(p0) = 94,08\mu s$$

Já o menor tempo de enfileiramento possível ocorre quando a fila está vazia ($N_{Q_{min}}(p0) = 0$), ou seja, o quadro de determinada estação é o primeiro a atingir o *switch* no período considerado. Nesta situação todas parcelas da equação 10 se anulam ($T_{E_{min}}(p0) = 0$).

Com estas considerações acerca da parcela variável que compõe o cálculo de A_T, através da aplicação da equação 4, são definidos os limites mínimo e máximo para o atraso total dos quadros da rede de controle.

$$
\begin{aligned}
A_{T_{max}}(p0) &= T_Q + T_{Prop} + T_{E_{max}}(p0) \\
&= 11,52\mu s + 1\mu s + 94,08\mu s \\
A_{Tmax}(p0) &= 0,106ms
\end{aligned}
$$

$$
\begin{aligned}
A_{Tmin}(p0) &= T_Q + T_{Prop} + T_{Emin}(p0) \\
&= 11,52\ \mu s + 1\ \mu s \\
A_{Tmin}(p0) &= 12,52\mu s
\end{aligned}
$$

De forma semelhante a aplicada nos limites extremos de A_T, é coerente supor que em média o quadro transmitido por determinada estação encontrará a fila no comutador com a metade de seu tamanho máximo $(N_{Q_{med}}(p0) = 14/2 = 7)$.

$$
\begin{aligned}
T_{E_{med}}(p0) &= (7 \times T_{trans}) + T_{Res} \\
&= 7 \times 6,72\mu s + 0 \\
T_{E_{med}}(p0) &= 47,04\mu s
\end{aligned}
$$

$$
\begin{aligned}
A_{T_{med}}(p0) &= T_Q + T_{Prop} + T_{E_{med}}(p0) \\
&= 11,52\mu s + 1\mu s + 47,04\mu s \\
A_{T_{med}}(p0) &= 59,56\mu s
\end{aligned}
$$

O gráfico apresentado na figura 4 mostra o aumento linear do atraso enquanto um número maior de estações é inserido na rede de controle e os demais parâmetros são mantidos constantes. O ponto de validade das previsões do modelo é atingido com aproximadamente 148 estações, quando o valor de $A_{T_{max}}$ atinge o limiar de $1ms$ definido como período de geração de quadros prioritários.

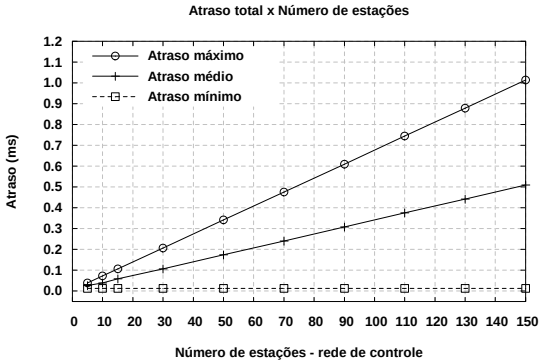

Figura 4: Número de estações

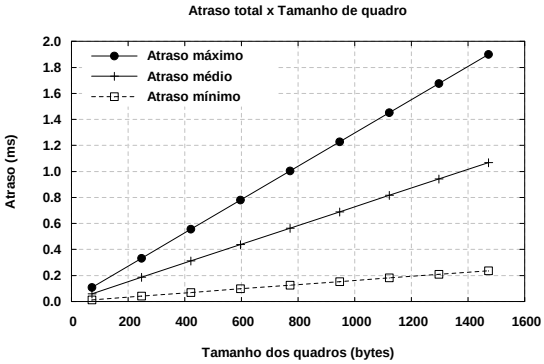

Figura 5: Tamanho dos quadros

Já a figura 5 mostra o comportamento de A_T quando o número de estações é mantido constante (15), mas o tamanho dos quadros na rede de controle é gradativamente incrementado de 72 até 1530 *bytes*. Neste caso, o número de quadros em disputa por posições na fila é mantido sempre o mesmo, mas o tempo necessário para transmissão de cada um deles (T_{trans}) é proporcionalmente elevado. Novamente o limite de 1ms é atingido, significando que nesta configuração adotada o maior quadro recomendável para a rede de controle possui cerca de 770 *bytes*.

A operação com valores de A_T superiores ao período da rede de controle significa o comprometimento das garantias temporais e da estabilidade do sistema.

4.2. Caso 2: N estações x 2 *Switches* - 3 Prioridades

No caso anterior, a ausência de disputa pelos recursos da rede, originada no isolamento criado entre as distintas classes de tráfego, impede a observação dos efeitos da interferência entre os fluxos de prioridades distintas. Com este objetivo, a figura 6 apresenta uma nova topologia em que se sugere a existência de um canal de concentração para o trânsito dos quadros na rede.

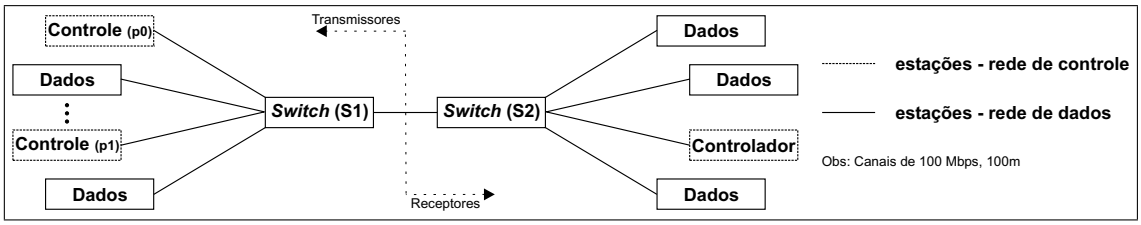

Figura 6: Caso 2 – Topologia

Outra diferença significativa é que os quadros da rede de controle são transmitidos sob duas prioridades distintas (0 e 1), conforme tabela 2. Novamente o tráfego de menor prioridade pertence às comunicações de dados ($p > 1$).

Tabela 2: Parâmetros principais da rede de controle – Caso 2

	Prioridade 0 (p0)	***Prioridade 1 (p1)***
Número de estações	10	20
Tamanho dos quadros	768 bits	576 bits
Período	1ms	1ms

Em relação às comunicações, o tráfego de controle produz neste caso um total de 30.000 quadros a cada segundo. A taxa gerada por estas comunicações é de 19,20Mbps, e pode ser facilmente obtida pela soma das transmissões efetuadas sob cada umas das prioridades da rede de controle (0 e 1) a cada período.

$$\begin{aligned} Taxa_{controle} &= Taxa_{p0} + Taxa_{p1}; \\ &= [768\ bits/quadro \times 1000\ quadros/seg. \times 10 \text{ estações}] + \\ &\quad [576\ bits/quadro \times 1000\ quadros/seg. \times 20\ estacoes] \\ &= 7,68Mpbs + 11,52Mbps = 19,20Mbps \end{aligned}$$

De maneira semelhante ao exemplo anterior, os valores de 30.000 quadros por segundo e taxa de transmissão de 19,20Mbps satisfazem plenamente as relações 1 e 2.

A presença de dois comutadores encadeados na rede (Figura 6) faz com que o caminho percorrido por determinado quadro, quando em trânsito entre uma estação qualquer e o controlador, seja composto por três linhas de transmissão: origem→*switch*(S1), *switch*(S1)→*switch*(S2) e *switch*(S2)→controlador.

As equações 5 e 6 fornecem:

$$T_Q(p0) = 3 \times \left(\frac{768\ bits}{100 \times 10^6\ bits/s}\right) = 23,04\ \mu s$$

$$T_Q(p1) = 3 \times \left(\frac{576\ bits}{100 \times 10^6\ bits/s}\right) = 17,28\ \mu s$$

$$T_{Prop}(p0) = T_{Prop}(p1) = 3 \times \left(\frac{100\ m}{2 \times 10^8\ m/s}\right) = 1,5\ \mu s$$

A respeito do tempo de enfileiramento, o primeiro ponto a destacar é que o fato de todos quadros alcançarem o *switch* S2 através de um único canal (S1→S2), com taxa

de transmissão igual a da linha de saída (S2→controlador), impede a formação de filas. Além disso, o tempo residual de contenção em S2 também não é significativo, uma vez que apenas tráfego de controle, originado de uma mesma porta de entrada e naturalmente enfileirado, é direcionado ao controlador ($T_{Res_{max}}(S2, p0) = T_{Res_{max}}(S2, p1) = 0$).

Como T_E representa uma parcela variável de A_T, seu estudo em S1 pode ser derivado de previsões em torno de limites extremos e valores médios, conforme abaixo:

T_E **mínimo:** quadros de ambas prioridades podem encontrar o melhor cenário possível em S1, ou seja, filas vazias e tempo residual de contenção nulo. Para isto, basta atingir o comutador antes de qualquer outra mensagem no período considerado.

T_E **máximo:** limite oposto, representa a pior configuração possível. O tamanho máximo da fila para determinado quadro de prioridade 0 é de 9 mensagens ($N_{Qmax}(p0, S1) = 9$). Já para um quadro de prioridade 1 deve ser considerada também a fila de maior prioridade absoluta, somando um total de 29 mensagens no pior caso ($N_{Qmax}(p1, S1) = 10\ quadros_{p0} + 19\ quadros_{p1}$). Além disso, quadros de ambas classes estão sujeitos ao maior tempo de contenção residual: $T_{Res_{max}}(S1, p0) = T_{Res_{max}}(S1, p1) = 12240bits/100 \times 10^6 bps = 122,40\ \mu s$.

T_E **médio:** supõe-se que quadros de prioridade 0 encontrem em média a fila com metade de seu tamanho máximo, ou seja, $N_{Q_{med}}(p0, S1) = \frac{N_{Q_{max}}(p0,S1)}{2} = 4$. Por outro lado, quadros de prioridade 1 serão naturalmente contidos, caso não encontrem a fila prioritária vazia, durante a transmissão e chegada de novos quadros de prioridade 0. Assim, em um ambiente de transmissões periódicas e praticamente simultâneas entre as diversas estações, determinado quadro de prioridade 1 será obrigado a aguardar em média o tempo necessário para transmissão de todos quadros com prioridade superior e de metade dos quadros de mesma prioridade ($N_{Qmed}(p1, S1) = 10\ quadros_{p0} + 9\ quadros_{p1}$). Para o tempo residual de contenção, deve-se considerar um tamanho médio de quadros: $T_{Res_{med}} = \frac{T_{Res_{max}}}{2} = 61,2\ \mu s$.

A aplicação destas considerações nas equações 8 e 10 para ambas classes prioritárias fornece:

$$T_{trans}(p0) = \frac{768\ bits}{100 \times 10^6\ bits/s} + 0,96\ \mu s = 8,64\ \mu s$$

$$\begin{aligned} T_{E_{max}}(p0) &= T_{E_{max}}(p0, S1) + T_{E_{max}}(p0, S2) \\ T_{E_{max}}(p0) &= [(9 \times T_{trans}(p0)) + T_{Res_{max}}(S1)] + 0 \\ &= 9 \times 8,64\ \mu s + 122,40\ \mu s \\ &= 200,16\ \mu s \end{aligned}$$

$$T_{trans}(p1) = \frac{576\ bits}{100 \times 10^6\ bits/s} + 0,96\ \mu s = 6,72\ \mu s$$

$$\begin{aligned} T_{E_{max}}(p1) &= T_{E_{max}}(p1, S1) + T_{E_{max}}(p1, S2) \\ T_{E_{max}}(p1) &= [(19 \times T_{trans}(p1)) + (10 \times T_{trans}(p0)) + \\ &\qquad T_{Res_{max}}(S1)] + 0 \\ &= 19 \times 6,72\ \mu s + 10 \times 8,64\ \mu s + 122,40\ \mu s \\ &= 336,48\ \mu s \end{aligned}$$

$$T_{E_{med}}(p0) = T_{E_{med}}(p0, S1) + T_{E_{med}}(p0, S2)$$
$$T_{E_{med}}(p0) = \left[(4 \times T_{trans}) + T_{Res_{med}}(S1)\right] + 0$$
$$= 4 \times 8,64\ \mu s + \frac{122,40}{2}\ \mu s$$
$$= 95,76\ \mu s$$

$$T_{E_{min}}(p0) = 0\ \mu s$$

$$T_{E_{med}}(p1) = T_{E_{med}}(p1, S1) + T_{E_{med}}(p1, S2)$$
$$= [(9 \times T_{trans}(p1)) + (10 \times T_{trans}(p0)) + T_{Res_{med}}(S1)] + 0$$
$$= 9 \times 6,72\ \mu s + 10 \times 8,64\ \mu s + \frac{122,40}{2}\ \mu s$$
$$= 208,08\ \mu s$$

$$T_{E_{min}}(p1) = 0\ \mu s$$

Com os valores máximos, médios e mínimos estabelecidos para o tempo de enfileiramento é possível calcular, através da relação 4, os limites para a variação do atraso total observado (Tabela 3).

Tabela 3: Previsões para A_T

Previsão	Prioridade 0	Prioridade 1
$A_{T_{max}}$	$0,224\ ms$	$0,355\ ms$
$A_{T_{med}}$	$0,120\ ms$	$0,226\ ms$
$A_{T_{min}}$	$24,54\ \mu s$	$18,78\ \mu s$

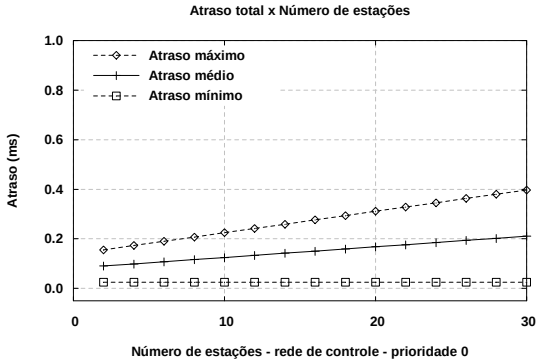

Figura 7: Previsões - Prioridade 0

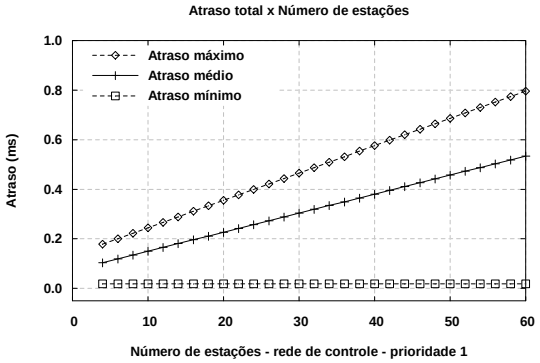

Figura 8: Previsões - Prioridade 1

Os valores apresentados nos gráficos das figuras 7 e 8 mostram a evolução linear de A_T para ambas classes prioritárias enquanto um número maior de estações é inserido na rede de controle. Durante a montagem dos gráficos foi mantida a mesma proporção apresentada na tabela 2, na qual o número de estações transmitindo sob prioridade 1 é mantido igual ao dobro de estações de prioridade 0.

Através da inspeção das linhas representativas de $A_{T_{max}}$ é possível estabelecer o limite extremo para operação estável do sistema. A soma do tempo máximo de atraso de ambas prioridades ultrapassa o limiar de 1ms aproximadamente quando a rede de controle é composta por 24 ($A_{T_{max}}(p0) = 0,3457ms$) e 48 ($A_{T_{max}}(p1) = 0,6644ms$) estações transmitindo, respectivamente, sob prioridades 0 e 1. A partir deste ponto não é possível garantir que todos quadros gerados em um período de transmissão sejam efetivamente entregues ao controlador. A taxa de entrada nas filas pode ultrapassar a taxa de saída, possibilitando inclusive o descarte por escassez de recursos nos comutadores.

5. Resultados de Validação do Modelo

A abordagem escolhida para validação do modelo de cálculo proposto é a comparação dos resultados teóricos com os obtidos através de simulação dos cenários escolhidos como estudos de caso. O simulador empregado é o *NS* ou *Network Simulator* [NS2, 2003, Bajaj et al., 1999].

Como bases para modelagem destas configurações de interesse estão a ligação ponto a ponto entre as diferentes entidades e a utilização do módulo de serviços diferenciados disponível no simulador [Pieda et al., 2000]. Embora os dispositivos previstos neste módulo lidem com o nível de rede ao invés da camada de enlace da arquitetura, sua utilização para estudo dos mecanismos previstos nas normas IEEE 802.1 é válida uma vez que deriva do mesmo conceito teórico baseado na marcação de tráfego e na utilização de prioridades relativas.

O ambiente simulado dispõe então das características principais necessárias como enfileiramento independente e escalonamento de prioridade absoluta. O fato de estarem implementadas em outro nível da arquitetura não influi nos resultados pois o modelo abstrai as especificidades físicas. O tempo de processamento dos quadros é desconsiderado no modelo deduzido e inexiste nas simulações, evitando discrepâncias características de abordagens em diferentes camadas.

Para confronto com os resultados teóricos obtidos com a aplicação do modelo analítico o atraso das comunicações no ambiente simulado é extraído pela diferença temporal entre o momento de partida do pacote de sua estação origem e o instante de chegada ao destino. Este valor engloba todos tempos intermediários de injeção, transmissão e enfileiramento.

Os resultados obtidos para a rede de controle passam ao final por um processo de consolidação, agregando valores de todos micro-fluxos existentes. Neste processo, os valores médios são obtidos para cada experimento através da média aritmética de todas replicações e de todos fluxos de mesma prioridade na rede. Já os limites mínimos e máximos são reflexos de valores pontuais observados em algum ponto da simulação e selecionados por comparação direta com todos demais.

O modelo de tráfego adotado para a rede de dados durante as simulações considera taxas constantes de geração do maior quadro possível (1530 bytes). O objetivo é manter a utilização da rede muito próxima de 100%, ocupando a banda não utilizada pelas transmissões de controle.

5.1. Caso 1 - Validação

A figura 9 mostra o confronto entre os resultados teóricos e aqueles extraídos por simulação do cenário de interesse analisado na seção 4.1. Para demonstrar o isolamento e independência entre as duas classes de tráfego, os experimentos foram repetidos, excluindo a comunicação de dados e mantendo ativas apenas as estações da rede de controle. Os resultados obtidos foram praticamente os mesmos, conforme figuras 10 e 11.

5.2. Caso 2 - Validação

Os gráficos expostos nas figuras 12 e 13 apresentam, para o segundo caso de interesse (Seção 4.2), o confronto entre resultados experimentais e teóricos. Novamente observa-se

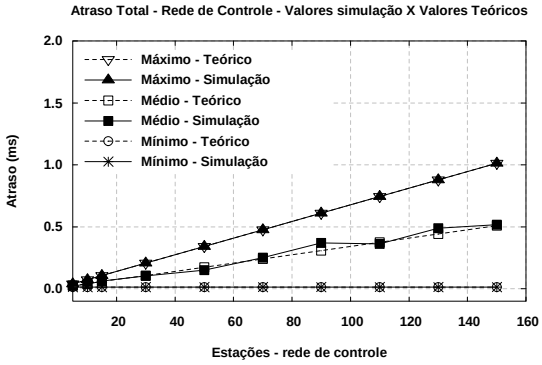

Figura 9: **Teóricos x Experimentais**

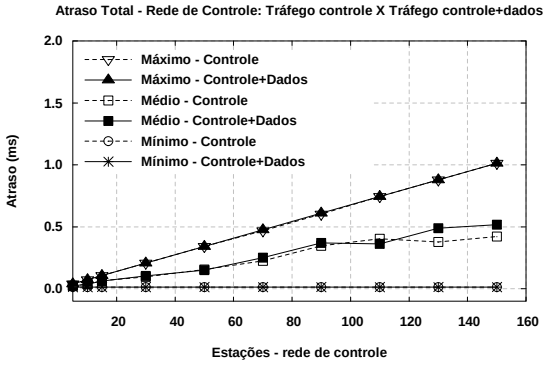

Figura 10: **Controle x Controle+Dados**

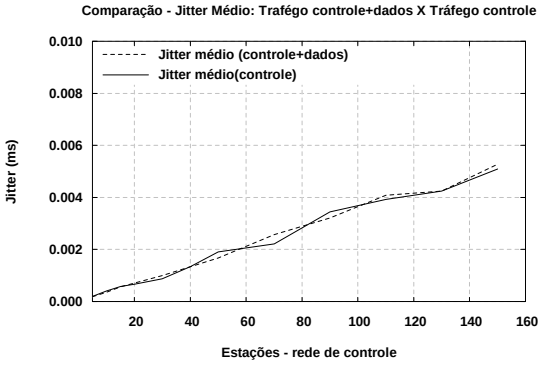

Figura 11: ***Jitter*** **(Controle x Controle+Dados) - Caso 1**

extrema compatibilidade entre os resultados, reforçando a validade das previsões obtidas com o modelo analítico.

Seguindo o mesmo procedimento adotado anteriormente, os experimentos foram repetidos mantendo a rede de dados inativa (Figuras 14,15,16 e 17). O confronto dos resultados demonstra claramente a interferência sobre a temporização dos quadros da rede de controle. A elevada taxa de quadros de tamanho máximo ("CBR") adotada para as transmissões de dados constitui um ambiente extremamente desfavorável, estimulando a possibilidade de inversão de prioridades na disputa pelo canal compartilhado S1→S2 (T_{Res}).

Para ilustrar esta característica de pior caso adotada na modelagem da rede de dados, novos experimentos foram realizados substituindo seu tráfego por *traces* reais capturados durante a operação normal de uma rede *ethernet*[5]("Trace").

Os deslocamentos entre as linhas de "controle" e de "controle+dados – CBR" (pior caso) apresentados nestas figuras são extremamente coerentes: aproximadamente $122\mu s$ para $A_{T_{max}}$ e $61\mu s$ para $A_{T_{med}}$. Estes são, respectivamente, os tempos gastos com a inserção no canal de um quadro completo ($T_{Res_{max}}$) e de metade de um quadro da rede de dados ($T_{Res_{med}}$). Como esperado, a linha representativa do tráfego ethernet real ("controle+dados – Trace") encontra-se entre os limites extremos definidos.

[5]Disponibilizados pelo *Bellcore Morristown Research and Engineering Facility* – http://ita.ee.lbl.gov

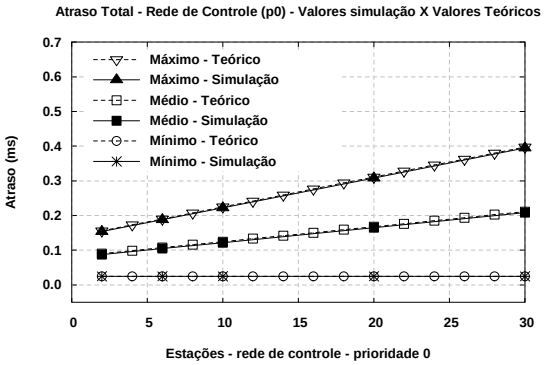

Figura 12: **Confronto - prioridade 0**

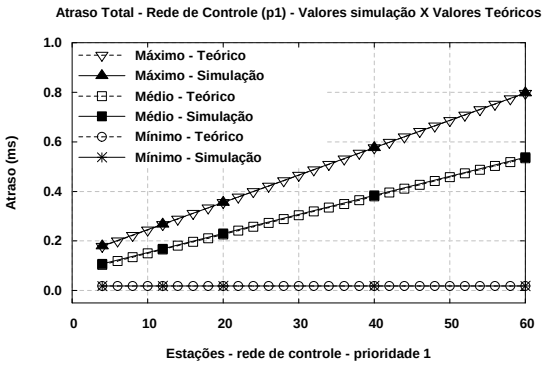

Figura 13: **Confronto - prioridade 1**

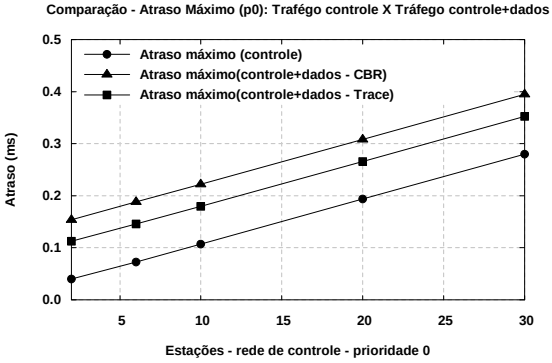

Figura 14: **Interferência -** $A_{T_{max}}(p0)$

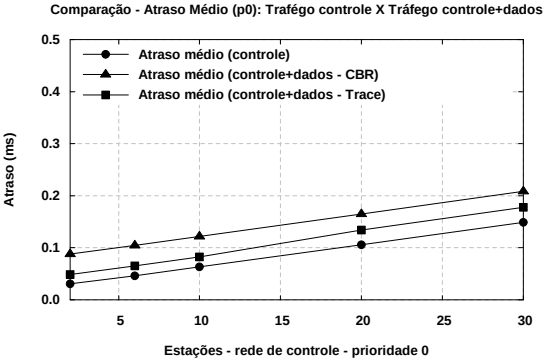

Figura 15: **Interferência -** $A_{T_{med}}(p0)$

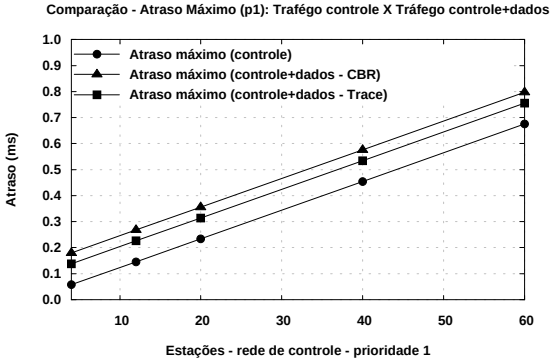

Figura 16: **Interferência -** $A_{T_{max}}(p1)$

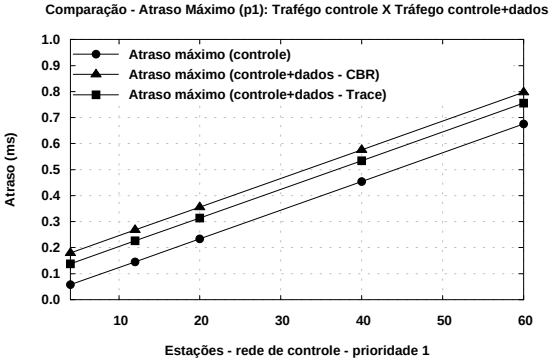

Figura 17: **Interferência -** $A_{T_{med}}(p1)$

6. Conclusões

O modelo analítico de cálculo proposto neste trabalho é capaz de prever os limites de atraso fim-a-fim observado por diferentes classes de tráfego nas comunicações em redes *ethernet* micro-segmentadas. O perfil temporal é traçado unicamente com base nos efeitos provocados pela estrutura de rede, mais especificamente, pelo nível de enlace da arquitetura.

A análise teve seu foco direcionado para ambientes de automação através de um cenário de estudo inspirado em sistemas de controle distribuídos, transmissões cíclicas e no modelo de comunicação mestre-escravo. Além do tráfego de controle, admitiu-se a convivência com fluxos que não possuem requisitos temporais – rede de dados.

A dedução do modelo foi baseada em uma análise determinística através da previsão do tráfego prioritário presente na rede a cada ciclo de transmissão. Esta abordagem é alternativa à técnica de aproximação do tráfego por distribuições conhecidas e análise do problema sob a ótica da teoria de filas, como em [Song et al., 2002] e [Iida et al., 2001].

Os resultados permitem concluir que o uso de redes *ethernet* baseadas em *switches*, agregado à utilização dos mecanismos de diferenciação de tráfego disponíveis, representa uma tecnologia bastante promissora para sistemas de tempo real. No ambiente industrial por exemplo, a efetiva integração desta estrutura por parte dos equipamentos, muitos dos quais já incluem interfaces ethernet, representaria um passo importante para a convergência de todos níveis da arquitetura de rede comumente empregada: planta, controle e campo.

Embora definidos em 1998 e implementados na maioria dos equipamentos utilizados atualmente, os dispositivos de priorização especificados nas normas IEEE não foram amplamente difundidos. Um passo fundamental seria sua disponibilidade para o nível de aplicação com a alteração de protocolos ou interfaces intermediárias (e.g. *API* de *sockets*).

Um ponto importante revelado é a possibilidade de inversão de prioridades quando o tráfego de diferentes classes é concentrado em canais de saída. A alta carga de tráfego adotada na modelagem da rede de dados durante as simulações, expõe uma influência negativa sobre a temporização dos quadros prioritários, mas demonstra que as oscilações inseridas são limitadas e previsíveis. A interferência verificada está atrelada aos tamanhos máximos (Figuras 14 e 16) e médios (Figuras 15 e 17) dos quadros envolvidos na disputa. Por outro lado, a ausência de canais compartilhados entre fluxos de prioridade distinta permite verificar a independência entre classes com tratamento diferenciado, conforme demonstrado pela figura 11.

Embora a validação do modelo através de simulação demonstre extrema coerência, a observação de alguns fatores adicionais, como o impacto dos tempos de processamento envolvidos nos processos de classificação e repasse nos *switches*, indica a necessidade de montagem de um cenário real e extração de medidas práticas. Para tal, devem ser respeitadas as premissas do modelo deduzido, especialmente em relação aos comutadores empregados.

Referências

3Com (2003). SuperStack Switches – Data Sheets. http://www.3com.com.

Almes, G., Kalidindi, S., and Zekauskas, M. (1999). A One-way Delay Metric for IPPM. Request for Comments 2679, IETF - Internet Engineering Task Force.

Bajaj, S., Breslau, L., Estrin, D., Kevin Fall, S. F., Haldar, P., Handley, M., Helmy, A., Heidemann, J., Huang, P., Kumar, S., McCanne, S., Rejaie, R., Sharma, P., Varadhan, K., Xu, Y., Yu, H., and Zappala., D. (1999). Improving Simulation For Network Research. USC Technical Report 99-702, USC - University of Southern California.

Bello, L. L. and Mirabella, O. (2001). Design Issues for Ethernet in Automation. In *Proc. of the 8th IEEE International Conference on Emerging Technologies and Factory Automation – ETFA2001*, Antibes, France. IEEE.

Caro, R. H. (2001). Ethernet Wins Over Industrial Automation. *IEEE Spectrum*, (38):114–115.

Chiueh, T. and Venkatramani, C. (1994). Supporting Real-time Traffic on Ethernet. In *Proc. of IEEE Real-time Systems Symposium – 1994*, San Juan, Puerto Rico.

Chow, M. Y. and Tipsuwan, Y. (2001). Network-based Control Systems: A Tutorial. In *Proc. of the 27th Annual Conference of the IEEE Industrial Eletronics Society*, pages 1593–1602, Denver, United States. IEEE.

Court, R. (1992). Real-Time Ethenet. *Computer Communications*, 15(3):198–201.

Decotignie, J.-D. (2001). A Perspective on Ethernet as a Fieldbus. In *Proceedings of the 4th International Conference on Fieldbus Systems and their Applications – FET'01*, Nancy, France.

Hoang, H. and Jonsson, M. (2003). Switched Real-Time Ethernet in Industrial Applications - Deadline Partitioning Scheme. In *Proc. of the 9th Asian Pacific Conference on Communication*, Penang, Malaysia.

Hoang, H., Jonsson, M., Hagström, U., and Kallerdahl, A. (2002). Switched Real-Time Ethernet and Earliest Deadline First Scheduling - Protocols and Traffic Handling. In *Proc. of International Parallel and Distributed Processing Symposium – IPDPS 2002*, Fort Lauderdale, Florida,USA.

IEEE (1998). 802.1D - Part 3: Media Access Control Bridges. Std, Piscataway, NJ - USA.

IEEE (2003). 802.1Q - Virtual Bridged Local Area Networks. Std, Piscataway, NJ - USA.

Iida, K., Takine, T., Sunahara, H., and Oie, Y. (2001). Delay Analysis for CBR Traffic under Static-Priority Scheduling. *IEEE/ACM Trans. Netw.*, 9(2):177–185.

ISO (1992). Road vehicles – interchange of digital information – controller area network (can) for high speed communication. DIS 11898, ISO – International Organization for Standardization.

ISO (1999). Digital data communications for measurement and control – fieldbus for use in industrial control systems – part 4: Data link protocol specification. IEC 61158-4, ISO – International Organization for Standardization.

Jasperneite, J. and Neumann, P. (2001). Switched Ethernet for Factory Communication. In *Proc. of the 8th IEEE International Conference on Emerging Technologies and Factory Automation – ETFA2001*, Antibes, France. IEEE.

Jasperneite, J., Neumann, P., Theis, M., and Watson, K. (2002). Deterministic Real-Time Communication with Switched Ethernet. In *Proc. of the 4th IEEE Workshop on Factory Communication Systems – WFCS'02*, pages 11–18, Vasteras, Sweden. IEEE.

Kaplan, G. (2001). Ethernet's Winning Ways. *IEEE Spectrum*, (38):113–114.

Khanna, V. K. and Singh, S. (1994). An Improved Piggyback Ethernet Protocol and Its Analysis. *Computer Networks ISDN Systems*, 26(11):1437–1446.

Lee, K. C. and Lee, S. (2002). Performance Evaluation of Switched Ethernet for Real-Time Industrial Communications. *Computer Standards & Interfaces*, 24(5):411–423.

Lian, F., Moyne, J., and Tilbury, D. (2001). Performance Evaluation of Control Networks: Ethernet, ControlNet, and DeviceNet. *IEEE Control Systems Magazine*, 21(1):66–83.

Moldovansky, A. (2002). Utilization of Modern Switching Technology in Ethernet/IP Networks. In *Proc. of the 1st Int. Workshop on Real-Time LANS in the Internet Age – RTLIA'02*, Porto, Portugal. Politema.

NS2 (2003). The Network Simulator – NS2. http://www.isi.edu/nsnam/ns/index.html.

Pieda, P., Ethridge, J., Baines, M., and Shallwani, F. (2000). A Network Simulator Differentiated Services Implementation. Open IP, Nortel Networks.

Shimokawa, Y. and Shiobara, Y. (1985). Real-Time Ethernet for Industrial Applications. In *Proc. of International Conference on Industrial Eletronics – IECON'85*, San Francisco, CA, United States. IEEE.

Song, Y. (2001). Time Constrained Communication over Switched Ethernet. In *Proceedings of the 4th International Conference on Fieldbus Systems and their Applications – FET'01*, pages 138–143, Nancy, France.

Song, Y., Koubaa, A., and Simonot, F. (2002). Switched Ethernet for Real-Time Industrial Communication: Modelling and Message Buffering Delay Evaluation. In *Proc. of the 4th IEEE International Workshop on Factory Communication Systems – WFCS'02*, Vasteras, Suécia. IEEE.

Stankovic, J. A. (1988). Misconceptions About Real-Time Computing: A Serious Problem for Next-Generation Systems. *Computer*, 21(10):10–19.

Stankovic, J. A. et al. (1996). Strategic Directions in Real-Time and Embedded Systems. *ACM Computing Surveys*, 28(4):751–763.

Steffen, R., Zeller, M., and Knorr, R. (2003). Real-Time Communication over Shared Media LANs. In *Proc. of the 2nd Int. Workshop on Real-Time LANs in the Internet Age – RTLIA'03*, Porto, Portugal. Politema.

Tanenbaum, A. S. (1996). *Computer Networks*. Prentice-Hall PTR, New Jersey, 3 edition.

Thomesse, J. (1999). Fieldbus and Interoperability. *Control Engineering Practice*, 7(1):81–94.

Walsh, G. C. and Hong, Y. (2001). Scheduling of Networked Control Systems. *IEEE Control Systems Magazine*, 21(1):57–65.

Wheelis, J. D. (1993). Process Control Communications: Token Bus, CSMA/CD or Token Ring? *ISA Trans*, 32(2):193–198.