

Estudo da Fila FB em Sistemas Multiprocessados

Alexandre Magno de Sousa, Cristina Duarte Murta

¹Departamento de Computação, CEFET/MG
Belo Horizonte - MG

Resumo. A política de escalonamento FB (*Foreground-Background*) se mostra bastante apropriada para sistemas que experimentam, ao mesmo tempo, cargas com grande variabilidade no tempo de serviço e intensidade de tráfego elevada. Estas características colocam o sistema em um estado crítico, que pode ser tratado pela fila FB, sem exigir conhecimento do tamanho da tarefa. Este artigo apresenta um estudo da fila FB em ambientes distribuídos, para sistemas fortemente e fracamente acoplados. Os resultados são avaliados quanto ao tempo de execução das tarefas e quanto à justiça do escalonamento. Este artigo contribui para o conhecimento mais amplo das características da fila FB em ambientes paralelos e distribuídos.

Abstract. The *Foreground-Background* (FB) scheduling policy seems to be quite appropriate for systems that experiences at the same time load with high variability in service times and high traffic intensity. These characteristics can take the system into a critical condition, which can be handled by the FB queue, without requiring knowledge of the task size. This paper presents a study of the FB queue in strongly and weakly coupled distributed systems. The results are evaluated according to the response time and the scheduling fairness. This paper contributes to improve the knowledge on the characteristics of the FB scheduling policy in parallel and distributed environments.

1. Introdução

O uso crescentemente disseminado dos sistemas computacionais, aliado à evolução de sua arquitetura e à complexidade da carga, torna esses ambientes cada vez mais complexos. Neste contexto, situações críticas, representadas por intensidades de tráfego elevadas e cargas altamente variáveis, são encontradas. Nesses ambientes é desejável que as políticas de escalonamento sejam capazes de tratar estas situações. Este trabalho apresenta um estudo da política de escalonamento FB (*Foreground-Background*), que se mostra bastante apropriada para sistemas que experimentam, ao mesmo tempo, cargas com grande variabilidade no tempo de serviço e intensidade de tráfego elevada.

A política de escalonamento FB trata cada tarefa em função de seu tamanho, sem precisar conhecê-lo. A política FB tem sido estudada recentemente de maneira mais intensa [Nuyens 2004a, Nuyens and Wierman 2008, Wierman 2007b]. No entanto, estes trabalhos abordam o desempenho da política majoritariamente em sistemas com um processador. Em um artigo anterior, apresentamos um estudo da política de escalonamento FB em sistemas mono e multiprocessados [Sousa and Murta 2010]. O presente artigo amplia os conhecimentos alcançados no trabalho anterior, apresentando um estudo de simulação acerca do comportamento da política FB em sistemas multiprocessados fortemente e fracamente acoplados.

Neste trabalho, a política **FB** é comparada com as políticas tradicionais de escalonamento *Processor-Sharing* (**PS**) e *First-Come-First-Served* (**FCFS**) e com uma política ótima de escalonamento, a **SRPT** (*Shortest-Remaining-Processing-Time*). Estas políticas são simuladas em ambientes multiprocessados, submetidos a cargas cujos tempos de serviço apresentam baixa e alta variabilidade, e intensidades de tráfego variáveis. A principal métrica analisada é o tempo médio de resposta. A análise da justiça no tratamento das tarefas quando o sistema se aproxima de uma situação de instabilidade é também apresentada.

A comparação de desempenho da política **FB** nos dois modelos de acoplamento de sistemas, e frente a outras políticas, comparadas nas mesmas condições, revela as situações em que ela pode ser aplicada com vantagens frente às demais opções. Devemos observar que, geralmente, o projeto de sistemas de acoplamento forte, também vistos como sistemas paralelos, pode apresentar custos mais altos do que o de sistemas de acoplamento fraco, isto é, sistemas distribuídos.

Este artigo está organizado em cinco seções. A próxima seção trata da fundamentação teórica e da discussão dos trabalhos relacionados. A seção 3 descreve o planejamento dos experimentos. Na seção 4 são apresentados os resultados dos experimentos realizados neste artigo. Por fim, a seção 5 apresenta a conclusão deste trabalho e algumas direções para trabalhos futuros.

2. Fundamentação Teórica e Trabalhos Relacionados

Sistemas multiprocessados podem ser organizados de duas maneiras quanto ao acoplamento: fortemente acoplados e fracamente acoplados [Casavant and Kuhl 1988, Anger et al. 1990, Tanenbaum and Steen 2007]. Nos sistemas fortemente acoplados há uma única fila para todos os servidores, enquanto nos sistemas fracamente acoplados há uma fila para cada servidor. Estes modelos são apresentados nas Figuras 1 e 2.

O sistema fortemente acoplado pode ser visto como um sistema paralelo, em geral modelado pela fila $M/G/m$. Neste modelo, a utilização do sistema é dada por $\rho = \lambda/(m\mu)$, em que m representa o número de servidores.

Nos sistemas fracamente acoplados há duas decisões de escalonamento: a estratégia de despacho ou roteamento e a política de escalonamento aplicada à fila de cada servidor, ao qual tarefas são assinaladas. A decisão de roteamento é feita no momento da chegada da tarefa. O sistema fracamente acoplado pode ser visto como um sistema distribuído e pode ser modelado por um conjunto de filas $M/G/1$. Esse sistema consiste de m filas paralelas Q_1, \dots, Q_m , uma para cada servidor, conforme apresentado na Figura 2. O sistema é homogêneo, isto é, os servidores têm a mesma capacidade de processamento. Neste sistema, a taxa de chegada é definida como $\lambda = \sum_{i=1}^m \lambda_i$, e λ_i representa a taxa de chegada da i -ésima fila. A utilização de cada recurso do sistema é dada por $\rho_i = \lambda_i/\mu$ e a utilização do sistema é dada por $P = \sum_{i=1}^m \rho_i$.

Nesse trabalho consideramos, para a decisão de roteamento nos sistemas fracamente acoplados, as políticas de despacho **RR** (*Round-Robin*) e **JSQ** (*Join the Shortest Queue*). No roteamento **RR**, a i -ésima tarefa é assinalada ao $(i \bmod m)$ -ésimo servidor, enquanto no **JSQ** as tarefas são assinaladas ao servidor com a menor fila de espera.

Numerosas estratégias de roteamento podem ser encontradas na literatura. Além

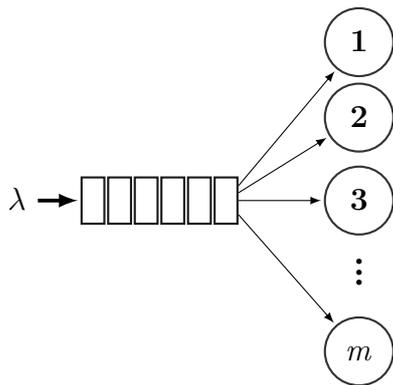


Figura 1. Modelo multiprocessado fortemente acoplado.

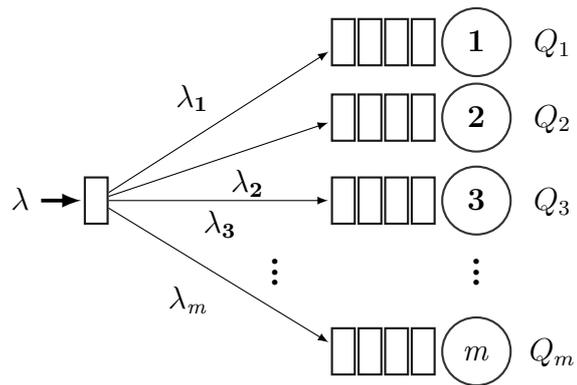


Figura 2. Modelo multiprocessado fracamente acoplado.

das citadas acima, outra estratégia comumente citada é a *Random*, que assinala tarefas aos servidores com uma probabilidade $1/m$, sendo m o número de servidores. *Random* e RR não necessitam de informação dos servidores para fazer a decisão de roteamento. Por sua vez, JSQ requer informação do tamanho da fila de espera de cada servidor.

Há estratégias mais elaboradas, por exemplo, *Size-Based Task Assignment* (SITAE), em que as tarefas são enviadas aos servidores de acordo com suas faixas de tamanhos [Harchol-Balter et al. 1999, Down and Wu 2006]. A escolha da estratégia de roteamento depende de vários fatores. Se o *overhead* de comunicação é uma medida a ser considerada, sugere-se o uso de *Random* ou RR. No entanto, políticas que utilizam mais informação, tal como JSQ, produzem melhores resultados quanto ao desempenho.

Políticas de serviço ou políticas de escalonamento definem a ordem em que as tarefas são atendidas. Essa ordem pode produzir impacto no desempenho de um sistema que implementa filas de espera para suas solicitações. Exemplos de políticas de escalonamento são *First-Come-First-Served* (FCFS), *Processor-Sharing* (PS), *Shortest-Remaining-Processing-Time* (SRPT), *Shortest-Job-First* (SJF) e *Foreground-Background* (FB).

Políticas de escalonamento que dão prioridade a tarefas menores reduzem o tempo médio de resposta [Conway et al. 2003, Wierman 2007b]. Uma das políticas que prioriza tarefas menores, sendo considerada teoricamente ótima, é a política SRPT [Schrage 1968, Bansal and Harchol-Balter 2001]. No entanto essa política depende do conhecimento a priori do tamanho das tarefas, o que pode tornar difícil ou mesmo impossibilitar sua utilização em ambientes nos quais não é possível conhecer antecipadamente o tamanho das tarefas. As dificuldades observadas na implementação da política SRPT em sistemas reais são discutidas em [Bansal and Harchol-Balter 2001].

Na implementação de políticas de escalonamento preemptivas, tais como SRPT e PS, os custos de preempção devem ser observados. Tais custos são oriundos do sistema operacional, uma vez que a cada preempção ou interrupção de uma tarefa em execução ocorre uma troca de contexto [Tanenbaum and Woodhull 2006]. Os custos de preempção da política SRPT são menores do que os da política PS [Bansal and Harchol-Balter 2001]. Na execução da política SRPT, uma tarefa só sofre preempção quando outra tarefa chega ou quando finaliza. Como cada tarefa chega

e termina exatamente uma vez, então, em média, as tarefas são interrompidas duas vezes. Estudos acerca da influência do tamanho do quantum no desempenho de uma fila M/G/1/RR são apresentados em [Gupta 2008]. Nesses estudos, os valores do quantum são relacionados à distribuição de serviço por meio de uma função matemática.

A política SRPT é ótima, isto é, produz os menores tempos médios de resposta dentre todas as políticas [Schrage and Miller 1966, Schrage 1968]. Todavia, em situações reais, o tamanho das tarefas pode não ser conhecido antecipadamente. Assim, é interessante encontrar alternativas que sejam um indicativo para estimar o tamanho da tarefa. Uma indicação importante do tamanho da tarefa é o número de vezes que ela foi escalonada pelo processador para ser servida. Esse número é conhecido na literatura por *age*. A política FB utiliza essa medida e prioriza tarefas de acordo com o menor valor de *age* [Nuyens 2004b]. Trabalhos recentes têm sido motivados por propostas de projetos de sistemas computacionais que façam o uso da política FB, tais como o escalonamento do fluxo de roteadores e o escalonamento de processos em sistemas operacionais [Nuyens and Wierman 2008, Wierman 2007b].

Resultados mostraram uma relação entre a variabilidade da distribuição de serviço e o desempenho da política FB [Nuyens and Wierman 2008]. Uma forma de expressar a relação de desempenho dessa política frente à variabilidade nas características da carga é por meio do coeficiente de variação [Feng and Misra 2004], que é dado por $C[X] = \sigma/E[X]$, em que σ representa o desvio padrão e $E[X]$ o valor esperado da distribuição de serviço X . De acordo com [Feng and Misra 2004] a fila FB apresenta desempenho ruim para cargas com $C[X] < 1$; comportamento semelhante à política PS para cargas com $C[X] = 1$; e apresenta bom desempenho para cargas com $C[X] > 1$. Quanto maior o $C[X]$, melhor o desempenho dessa política [Nuyens 2004b, Wierman 2007b].

A política FB é classificada como parcialmente justa, pois age com justiça para determinadas cargas de trabalho, e age com injustiça para outras cargas de trabalho [Wierman 2007a]. De acordo com Nuyens e Wierman, apesar da política FB priorizar teoricamente tarefas menores, apenas uma pequena quantidade de tarefas grandes é discriminada [Nuyens and Wierman 2008]. Resultados sobre o comportamento dessa política são encontrados na literatura apenas para sistemas monoprocessados, pois a avaliação analítica dessa fila é complexa e um modelo analítico para sistema multiprocessados ainda não foi desenvolvido [Nuyens 2004b, Wierman 2007b, Nuyens and Wierman 2008].

3. Planejamento dos Experimentos

O objetivo do trabalho é avaliar o desempenho da política FB em ambientes multiprocessados, organizados com acoplamento fraco ou forte, conforme descrito anteriormente. Para avaliarmos a fila FB de maneira comparativa, também foram feitos experimentos com as políticas de escalonamento FCFS e PS, que são comumente utilizadas em sistemas computacionais, e com a política teoricamente ótima, SRPT. As políticas são comparadas quanto ao tempo médio de resposta para a execução das tarefas.

A carga de trabalho foi gerada sinteticamente. Os resultados de desempenho são avaliados em função da variabilidade da carga, dada por $C[X]$. Para os testes foram escolhidas duas cargas de trabalho pouco variáveis (cauda leve), que são exponencial (M) e hiperequencial (H); e uma carga com variabilidade elevada (cauda pesada), representada pela distribuição de Pareto (P), com parâmetro de variabilidade (*shape*) $\alpha = 1,5$.

As tarefas chegam ao sistema de acordo com processo de Poisson, a taxa de chegada é definida como λ e a taxa de serviço é definida como μ . A utilização do sistema ou a intensidade de tráfego é dada por $\rho = \lambda/(m\mu)$.

Os resultados foram obtidos por simulação baseada em eventos discretos. O simulador utilizado para realização dos experimentos foi implementado em linguagem de programação C. Os resultados foram verificados e validados em um intervalo de confiança de 95%. O número de observações da amostra foi baseada em replicações e é dada por $n = k \times m$, em que m representa o número de execuções, também chamado de tamanho da execução, e k o número de replicações. Para realização dos experimentos, o número de execuções foi definido como $k = 10$ e o número de replicações como $m = 1 \times 10^6$, então $n = 10 \times 10^6$. Conforme MacDougall, o número de replicações deve ser relativamente pequeno, entre 1 e 10, e o tamanho das execuções deve ser relativamente grande [MacDougall 1987].

Em simulações que utilizam distribuição de serviço de cauda pesada, o número de amostras para que o sistema alcance o estado estável pode ser muito grande. Isso acontece porque a distribuição apresenta grande variabilidade e a convergência é lenta. Para a distribuição de Pareto, o número de amostras necessárias para alcançar uma precisão de dois dígitos quando o parâmetro de variabilidade é 1,5 é de um milhão [Crovella and Lipsky 1997]. Assim, isso justifica o grande número de execuções e foi definido um valor padrão de $m = 1 \times 10^6$ para todos os tipos de carga nos experimentos. Além disso, um número de execuções suficientemente grande também soluciona o problema de estado transiente (*warmup*) [MacDougall 1987, Jain 1991].

O simulador foi verificado realizando-se uma comparação entre os resultados do simulador e os resultados dos modelos analíticos disponíveis na literatura. Ressalta-se que a grande maioria dos modelos analíticos disponíveis são para sistemas com apenas um servidor. Para sistemas múltiplos servidores, não existem modelos analíticos para as políticas SRPT e FB, portanto, apenas o modelo da fila M/M/m/FCFS foi utilizado [Wierman 2007b, Nuyens and Wierman 2008].

As políticas de escalonamento no simulador funcionam conforme descrito a seguir. Para FCFS, as tarefas são alocadas aos servidores de acordo com a ordem de chegada, se não houver servidor livre, a tarefa é encaminhada à fila. As tarefas são atendidas concorrentemente de acordo com sua ordem de chegada para PS, cada tarefa é executada por um período de tempo (quantum), ao final desse período, se a tarefa não terminou sua execução, ela é encaminhada para fila. Para SRPT, quando uma tarefa chega ao sistema e há um servidor livre, este é alocado à mesma. Caso não haja servidor disponível, o tempo restante de execução das tarefas já alocadas aos servidores é avaliado, se houver tarefas com tempo restante de execução maior do que a nova tarefa, é feito um sorteio entre os servidores dessas tarefas. O servidor sorteado é alocado à nova tarefa e a tarefa que estava alocada a ele vai para fila. Por fim, para FB, existem contadores para verificar quantas vezes a tarefa foi servida. Quando uma tarefa solicita serviço, a prioridade é dada à tarefa com o menor valor desses contadores, a fila é ordenada por ordem crescente dos contadores.

No simulador, a política de despacho ou roteamento RR funciona da seguinte forma: a i -ésima tarefa é assinalada ao $(i \bmod m)$ -ésimo servidor. A política de rotea-

mento JSQ executa como descrito a seguir: cada tarefa é assinalada ao servidor com o menor número de tarefas na fila. Neste modelo, os custos de comunicação para realizar o roteamento das tarefas não são levados em consideração.

O planejamento de experimentos foi definido como do tipo fatorial e considera os fatores e níveis apresentados na Tabela 1 para a avaliação dos sistemas multiprocessados distribuídos e paralelos.

Fatores	Níveis
Política	FB; FCFS; PS; SRPT
Tempo de Serviço	M; H; P
No. de servidores	8
Acoplamento	fraco; forte
Roteamento	RR; JSQ
Utilização do Sistema	0,2 ; 0,4 ; 0,6 ; 0,8 ; 0,9 ; 0,95

Tabela 1. Principais fatores e níveis planejados para os experimentos.

Além dos resultados para o tempo médio de resposta, foram realizados mais dois experimentos para sistemas paralelos. O primeiro experimento teve como meta avaliar a justiça das políticas de escalonamento em relação ao atendimento das tarefas, à medida que a utilização do sistema se aproxima de uma situação de instabilidade. Foram feitos dois experimentos, um para $\rho \approx 0,90$ e outro para $\rho \approx 0,99$. Considerando que as filas FB e SRPT discriminam as tarefas de acordo com seu tamanho, é necessário avaliar a questão da justiça.

O segundo experimento consistiu em analisar o comportamento das políticas FB e PS com relação aos diferentes valores de quantum. Uma vez que essas políticas têm como parâmetro o quantum, é necessário avaliar o efeito de sua variação. A Tabela 2 apresenta os principais fatores e níveis para os experimentos realizados com o quantum.

Fatores	Níveis
Política	FB; PS
Quantum	1; 2; 4; 6; 8; 16
Tempo de Serviço	M; H; P
Utilização do Sistema	0,2 ; 0,4 ; 0,6 ; 0,8 ; 0,9 ; 0,95

Tabela 2. Principais fatores e níveis para os testes realizados com o quantum.

4. Resultados

Nesta seção são apresentados os resultados da simulação para os sistemas em teste. Para a geração dos tempos de serviço da carga de trabalho foram escolhidas distribuições de probabilidade que apresentam diferentes coeficientes de variação ($C[X]$). Os valores de $C[X]$ são os seguintes: para a distribuição exponencial, representada por M, $C[X] = 1$; para a distribuição hiperequencial, representada por H, $C[X] = 4$; e para a distribuição de Pareto com $\alpha = 1,5$, representada por P, $C[X]$ apresentou valores entre 4 e 210 durante a realização dos experimentos.

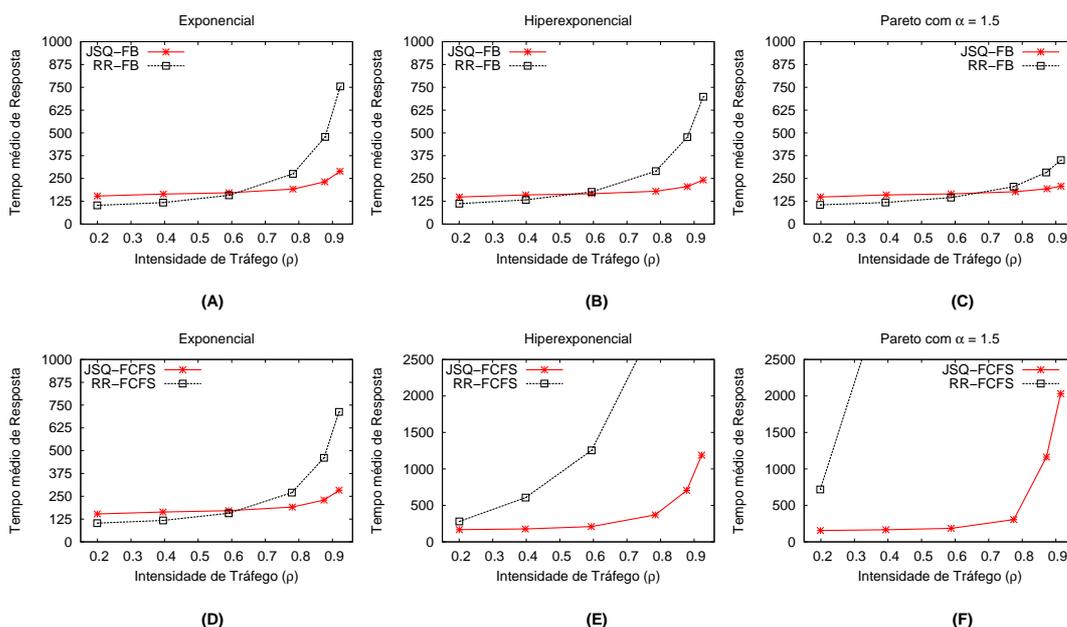


Figura 3. Resultados do sistema distribuído para as filas FB e FCFS com as políticas de roteamento RR e JSQ.

4.1. Sistema fracamente acoplado

Os gráficos de (A) a (F) da Figura 3 apresentam os resultados para as políticas FB e FCFS, frente às políticas de roteamento JSQ e RR, para as cargas M, H e P. Os gráficos (A), (B) e (C) tratam da fila FB, e os demais apresentam os resultados para a fila FCFS.

Para o caso da fila FB, observamos, para as três cargas, que o roteamento RR produz tempos de resposta ligeiramente menores do que JSQ para cargas leves. No entanto, para cargas elevadas, a política JSQ produz os melhores tempos de resposta. Além disso, a diferença entre as duas estratégias de roteamento diminui à medida que a variabilidade da carga aumenta.

No caso da fila FCFS frente às políticas de roteamento JSQ e RR, observamos que quanto maior a variabilidade da carga, maior a diferença entre as duas estratégias de roteamento. Em resumo, a estratégia de roteamento JSQ, em conjunto com a política de escalonamento FB, produz os melhores resultados em condições de tráfego mediano a intenso e cargas muito variáveis.

Na Figura 4 são apresentados os resultados para as políticas PS e SRPT frente às políticas de roteamento em teste, para as cargas M, H e P. Os resultados indicam comportamentos qualitativamente similares para RR e JSQ, sendo que RR apresenta os menores tempos de resposta para intensidades de tráfego baixas a moderadas, enquanto JSQ produz os melhores resultados para tráfego mais intenso. Lembramos que a política SRPT é ótima, isto é, produz os menores tempos de resposta. A combinação de JSQ com SRPT torna o tempo de resposta praticamente insensível à intensidade de tráfego, para todas as cargas.

Na Figura 5 são apresentados os resultados dos tempos médios de resposta do sistema distribuído, agrupados por distribuição de tempo de serviço. Para cada distribuição

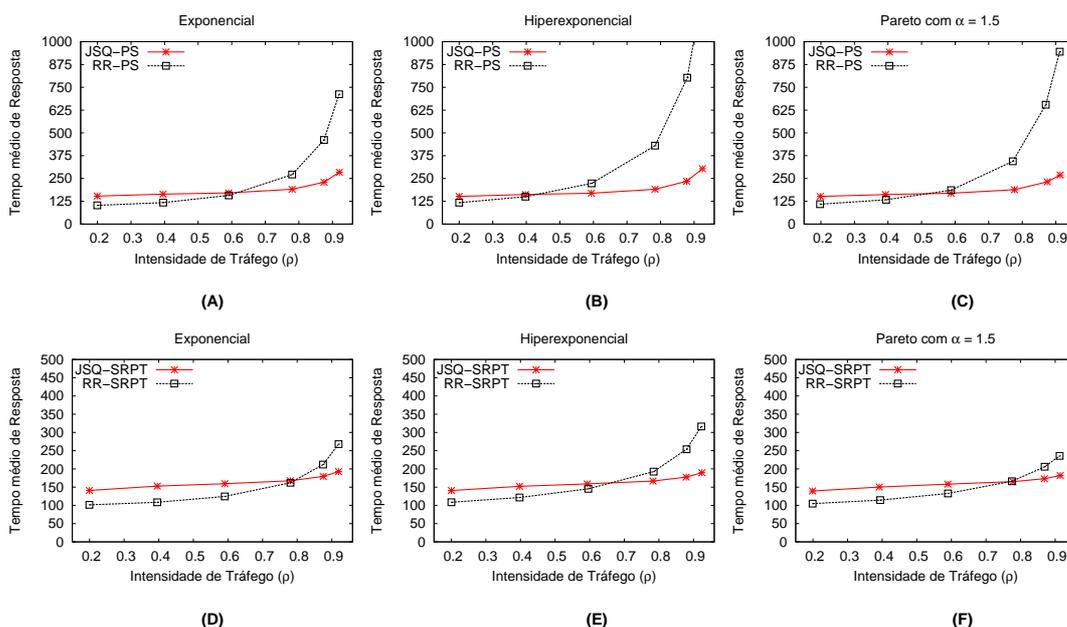


Figura 4. Resultados do sistema distribuído para as filas PS e SRPT com as políticas de roteamento RR e JSQ.

e cada estratégia de roteamento podemos comparar as quatro políticas de escalonamento em teste.

Para a distribuição exponencial (gráficos **(A)** e **(D)**) observa-se que as políticas de escalonamento, exceto SRPT, apresentam resultados iguais sob a mesma estratégia de roteamento. Sob RR, os resultados são melhores para pouco tráfego, e sob JSQ os resultados são melhores para tráfego intenso.

A política SRPT apresenta o melhor desempenho em todos os casos, conforme esperado, independente da estratégia de roteamento e da distribuição do tempo de serviço. Para os casos de distribuição hiperexponencial e Pareto, observa-se que a política FB apresenta resultados melhores do que PS e FCFS, para as duas estratégias de roteamento. O resultado de FB é mais próximo de SRPT para carga de Pareto, que é a mais variável.

A fila FCFS obtém os piores resultados para sistemas distribuídos que utilizam RR, resultados que não são apresentados no gráfico **(F)** devido a seus valores extremamente elevados.

4.2. Sistema fortemente acoplado

Os resultados da simulação para o sistema fortemente acoplado são apresentados nas Figuras 6 e 7. Pode-se observar que os tempos de resposta para todas as políticas são similares para intensidades de tráfego pequenas a moderadas. A política SRPT apresenta novamente o melhor desempenho. Para cargas intensas, FB apresenta o segundo melhor resultado, que é tanto mais próximo de SRPT quanto mais variável for a carga. Os resultados de FB são, novamente, melhores do que PS, que é uma das políticas mais utilizadas em sistemas computacionais.

A Figura 7 apresenta os resultados agrupados por política de escalonamento. Esses gráficos permitem avaliar o efeito da distribuição do tempo de serviço em cada tipo de

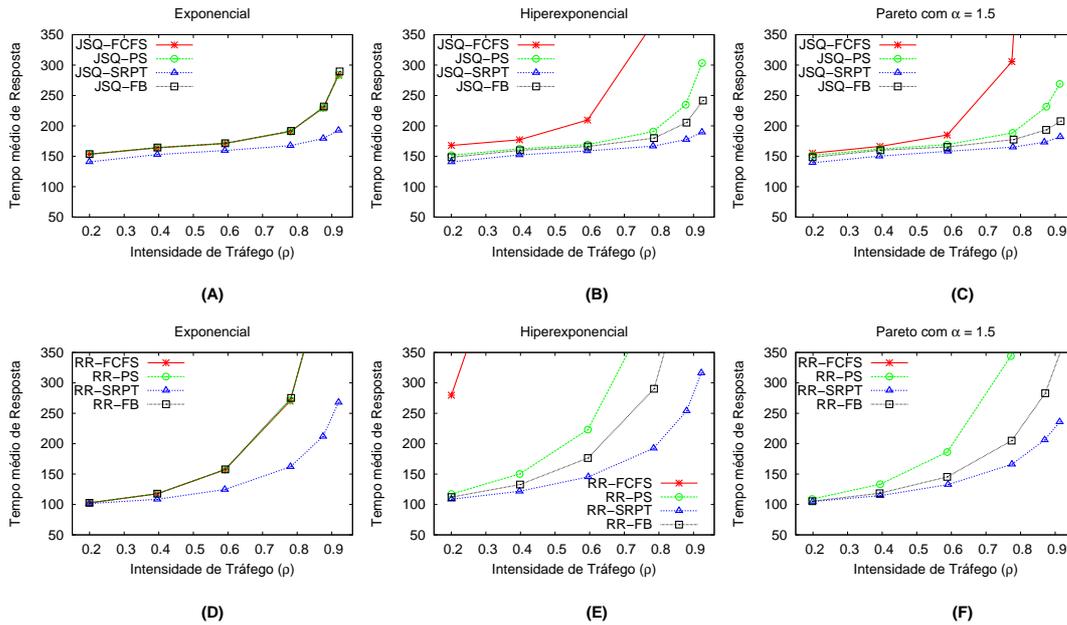


Figura 5. Resultados do sistemas fracamente acoplados agrupados por distribuição de serviço.

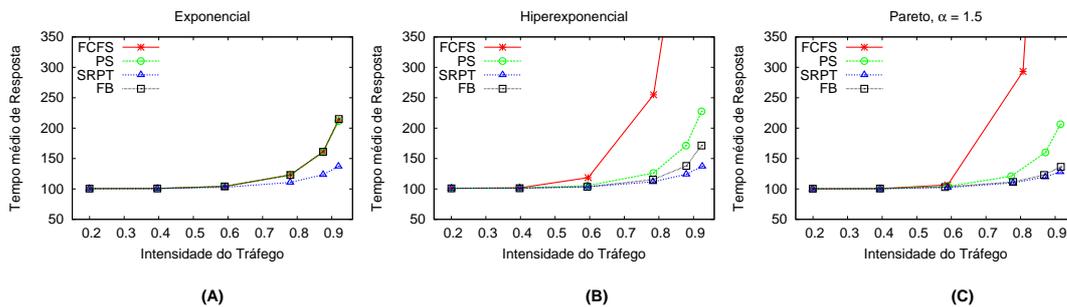


Figura 6. Resultados do sistemas fortemente acoplados de oito servidores para as cargas testadas.

escalonamento. Observa-se que a fila PS é praticamente insensível ao tipo de carga. A fila FB produz tempos menores para cargas mais variáveis.

A Tabela 3 apresenta resultados que permitem a comparação do sistema fortemente acoplado com o sistema fracamente acoplado para a política FB. Os valores apresentados correspondem à razão entre os tempos médios de resposta do sistema fracamente acoplado, representados por RR-FB ou JSQ-FB, pelo resultado da fila FB.

Os resultados indicam que, para todos os tipos de carga, o sistema fortemente acoplado apresenta melhor desempenho do que o sistema fracamente acoplado independente da política de roteamento. Ao comparar as duas estratégias de roteamento, observamos que RR apresenta os melhores resultados do tempo médio de resposta para intensidades de carga baixa e média. No entanto, quando há elevada utilização do sistema, a política de roteamento JSQ apresenta melhor desempenho para os três tipos de carga. A estratégia JSQ provê melhores resultados para cargas mais variáveis.

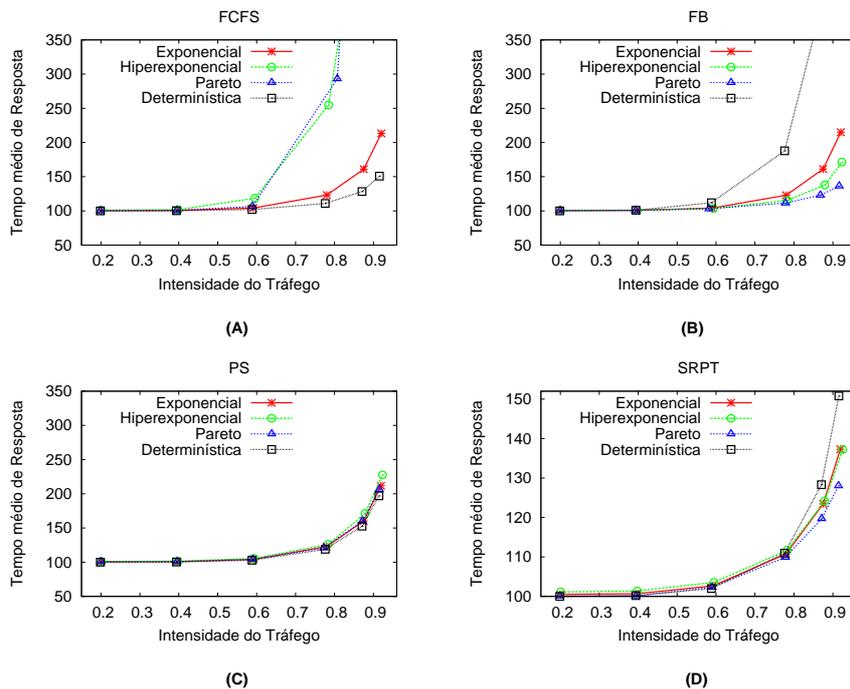


Figura 7. Resultados dos sistemas fortemente acoplados de oito servidores agrupados por política.

ρ	Exponencial		Hiperexponencial		Pareto	
	RR-FB/FB	JSQ-FB/FB	RR-FB/FB	JSQ-FB/FB	RR-FB/FB	JSQ-FB/FB
0,20	1,0206	1,5273	1,1106	1,4710	1,0491	1,4777
0,40	1,1682	1,6323	1,3135	1,5806	1,1771	1,5895
0,60	1,5152	1,6467	1,7052	1,6072	1,4098	1,6046
0,80	2,2397	1,5607	2,5174	1,5596	1,8378	1,5906
0,90	2,9667	1,4389	3,4563	1,4891	2,2997	1,5721
0,95	3,5111	1,3466	4,0737	1,4103	2,5698	1,5224

Tabela 3. Razão dos tempos médios de resposta dos sistemas com acoplamento fraco em relação ao sistema com acoplamento forte para a fila FB.

4.3. Avaliação da Justiça quando $\rho \rightarrow 1$

Esta seção apresenta a avaliação da justiça no atendimento das tarefas quando $\rho \rightarrow 1$, isto é, quando o sistema se aproxima de situação de instabilidade. Para essa avaliação, as tarefas foram agrupadas em cinco faixas de tamanhos: Faixa 1: $x = 10^0$; Faixa 2: $10^0 < x \leq 10^1$; Faixa 3: $10^1 < x \leq 10^2$; Faixa 4: $10^2 < x \leq 10^3$; Faixa 5: $10^3 < x \leq 10^4$; Faixa 6: $10^4 < x \leq 10^5$; Faixa 7: $10^5 < x$. As distribuições de serviço simuladas geraram tarefas de diferentes tamanhos, em função de suas características. Para a carga M, as tarefas se limitaram às faixas 1 a 5; para a carga H, as tarefas ficaram entre as faixas 1 e 6; e para distribuição de serviço P, as tarefas geradas ficaram entre as faixas 3 e 7.

O objetivo deste teste é avaliar se tarefas de todas as faixas de tamanho são tratadas de forma equivalente pelas políticas de escalonamento. A justiça é avaliada pela métrica *slowdown*, definida pela equação $S = E[T(x)]/x$, que é a razão entre o tempo de resposta

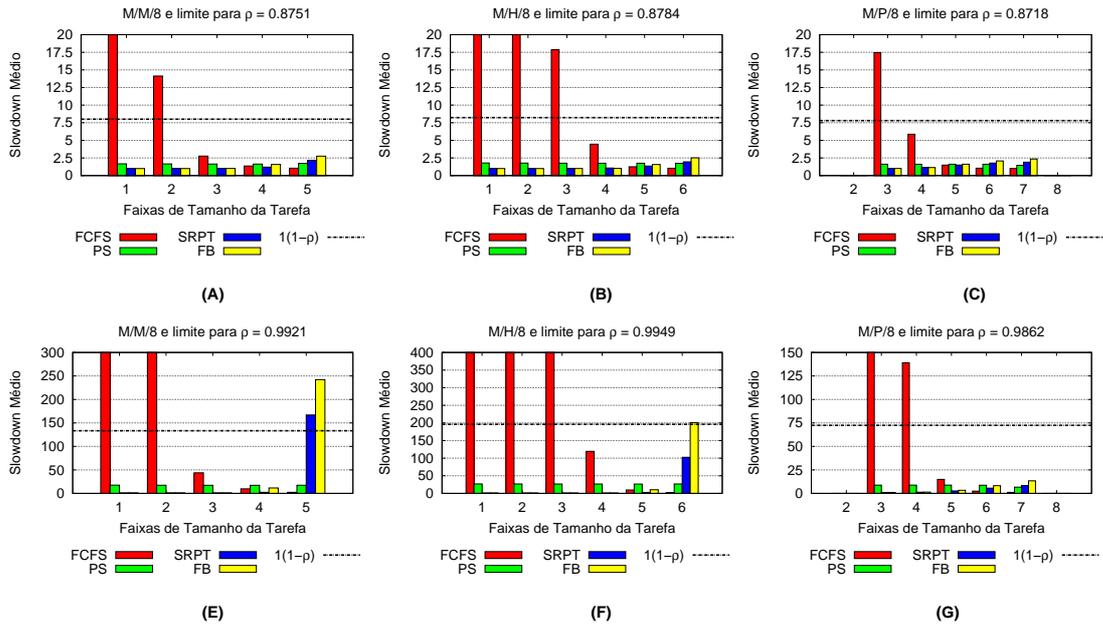


Figura 8. Slowdown médio no sistema fortemente acoplado.

e o tamanho da tarefa. Uma tarefa é tratada com justiça se a condição definida por $S \leq 1/(1 - \rho)$ for satisfeita [Nuyens and Wierman 2008]. Esse limite é representado por uma reta paralela ao eixo x nos gráficos da Figura 8.

A Figura 8 apresenta os gráficos¹ dos resultados do *slowdown* médio para o sistema fortemente acoplado, as três distribuições do tempo de serviço, e duas intensidades de tráfego (ρ) por tipo de carga. Os gráficos na primeira linha são para $\rho \approx 0,90$ e os da segunda linha são para $\rho \approx 0,99$. Os gráficos (A) e (E) apresentam o resultado para a distribuição exponencial (M/M/8) para duas intensidades de carga. Observamos que a fila FCFS trata com injustiça tarefas das classes 1 e 2. As políticas FB e SRPT discriminam apenas as tarefas que se encontram na faixa 5, para a carga mais elevada (gráfico (E)). No entanto, para as duas políticas, menos de 0,005% das tarefas são discriminadas. PS trata com justiça tarefas de todas as faixas de tamanho.

Os gráficos (B) e (F) apresentam o resultado para a distribuição hiperexponencial (M/H/8) para duas intensidades de carga. Neste caso, a fila FCFS trata com justiça apenas as tarefas que se encontram nas faixas 4, 5 e 6. As outras políticas tratam com justiça tarefas em todas as faixas de tamanho. FB se aproxima do limite estabelecido para tarefas da faixa 6. A política SRPT apresenta resultados para *slowdown* melhores do que a política FB para tarefas que se encontram nas faixas 5 e 6.

Finalizando, os gráficos (C) e (G) apresentam o resultado para a distribuição de Pareto (M/P/8). A fila FCFS continua penalizando tarefas de tamanhos pequenos. As políticas PS, FB e SRPT tratam com justiça tarefas de todos os tamanhos.

Em resumo, observamos que a fila FB se comportou muito bem quanto à justiça para tarefas cujo tamanho apresenta maior variabilidade (hiperexponencial e Pareto) e para intensidades de carga consideráveis, próximas do limite máximo ($\rho \approx 0,99$).

¹Os gráficos são melhor visualizados na tela devido ao esquema de cores.

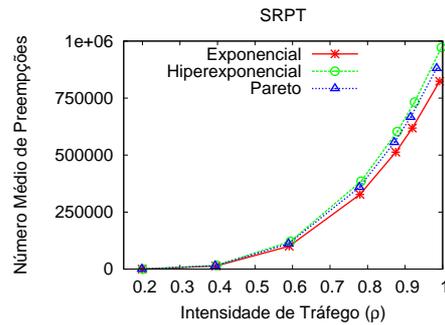


Figura 9. Número de preempções das políticas FB, PS e SRPT.

4.4. Número de Preempções

Para políticas preemptivas, é importante analisar o número de preempções que ocorrem no sistema, uma vez que a preempção gera custos adicionais de tempo para as tarefas de troca de contexto. Foram feitos experimentos para avaliar o número de preempções das políticas FB, PS e SRPT para o sistema fortemente acoplado. Observamos que o número de preempções gerado na execução das políticas FB e PS é semelhante para todos os tipos de carga. O total de tarefas escalonadas em cada caso foi de dez milhões. Para essas tarefas, o número total de preempções para essas duas políticas é de 1×10^8 . Esse número não depende da intensidade de tráfego. Assim, conclui-se que o custo de preempção para a política FB é igual ao da política PS.

O gráfico da Figura 9 mostra que o número de preempções da política SRPT. Nesse caso, o número de preempções varia de acordo com a intensidade de tráfego e com os tipos de carga de trabalho. Quanto maior a intensidade de tráfego, maior o número de preempções. O número de preempções foi menor para carga exponencial e maior para carga hipereexponencial. Para Pareto, o número de preempções foi intermediário. Pode-se concluir que, em um problema de escalonamento, o custo de preempções da política SRPT será sempre menor do que o das políticas FB e PS.

4.5. Alterações no Valor do Quantum

O efeito das alterações no quantum é estudado somente para as políticas FB e PS, que definem esse parâmetro. Apesar da política FB priorizar tarefas de acordo com o menor valor de *age*, no momento em que existe um conjunto de tarefas com a mesma prioridade, essa política executa à maneira PS e utiliza um quantum de tempo para execução das tarefas. Para tanto, foram realizados experimentos com diferentes valores de quantum para analisar o impacto no comportamento das políticas FB e PS.

Os testes foram realizados para as cargas exponencial, hipereexponencial e Pareto, com $\alpha = 1, 5$, e foram definidos cinco valores de quantum, a saber, 1, 2, 4, 8 e 16 unidades de tempo de acordo com a Tabela 2 do planejamento de experimentos. A alteração nos valores do quantum não causa impacto no tempo médio de resposta, independente do tipo de carga escalonada pela política FB. Da mesma forma, para PS, os diferentes valores de quantum não influenciam o tempo médio de resposta da política PS. Diante dos resultados apresentados, pode-se perceber que os diferentes valores de quantum não influenciam o comportamento das políticas PS e SRPT e não causam nenhum impacto no tempo médio

de resposta. Uma explicação para este resultado é o fato de que, em média, o tamanho das tarefas é pequeno, sendo que a maioria das tarefas é menor do que os valores testados para o quantum, o que torna o quantum sem efeito, uma vez que as tarefas são finalizadas em um quantum.

5. Conclusão

Este artigo apresentou uma análise de desempenho da política de escalonamento FB em ambientes distribuídos com acoplamento fraco e forte. Os resultados de desempenho desta política foram comparados aos resultados das políticas PS, FCFS e SRPT, executadas em condições iguais. As políticas PS e FCFS foram escolhidas por serem bastante conhecidas e estudadas. Assim, elas provêem uma base segura para comparação. Por outro lado, a política SRPT foi escolhida por ser uma política ótima, que provê os melhores resultados em termos da métrica alvo, tempo médio de resposta. Dessa forma, ela oferece um limite inferior, igualmente importante para mostrar o desempenho da fila FB frente ao melhor desempenho possível. No modelo de acoplamento fraco as requisições foram roteadas, no momento da chegada, pelas estratégias de roteamento JSQ e RR.

Os resultados indicam que a fila FB produz melhores tempos médios de resposta do que as filas PS e FCFS e compatíveis com SRPT para cargas com maior variabilidade e em situações de intensidade de tráfego moderada a elevada. Devemos observar que quanto maior a variabilidade da carga e maior a intensidade de tráfego, mais crítica é a situação do sistema, uma vez que o aumento destas condições é prejudicial ao desempenho do sistema. Portanto, FB produz bons resultados em sistemas em estado crítico. Por outro lado, FB não requer conhecimento do tamanho da tarefa, enquanto SRPT, que produz os melhores resultados, requer esse conhecimento. Assim, FB demonstra ser uma política bastante interessante para sistemas em estado crítico que não têm conhecimento a priori do tamanho da tarefa, ou que preferem não utilizar essa informação no escalonamento.

Os trabalhos futuros estão direcionados para a realização de experimentos com sistemas heterogêneos, compostos por recursos de capacidades diferentes. Além disso, experimentos com cargas altamente variáveis, como Pareto com $\alpha = 1,2$, são necessários para simular cargas ainda mais variáveis. Tendo em vista o contexto de discriminação de tarefas por classes de prioridades, experimentos com diferentes classes de prioridades de tarefas devem ser realizados.

6. Agradecimentos

Os autores agradecem ao UNIFOR/MG, à FAPEMIG e ao INCT Web pelo apoio.

Referências

- Anger, F. D., Hwang, J.-J., and Chow, Y.-C. (1990). Scheduling with Sufficient Loosely Coupled Processors. *Journal of Parallel and Distributed Computing*, 9(1):87–92.
- Bansal, N. and Harchol-Balter, M. (2001). Analysis of SRPT Scheduling: Investigating Unfairness. In *Proceedings of ACM SIGMETRICS 2001*.
- Casavant, T. and Kuhl, J. (1988). A Taxonomy of Scheduling in General-purpose Distributed Computing Systems. *Software Engineering, IEEE Transactions on*, 14(2).

- Conway, R. W., Maxwell, W. L., and Miller, L. W. (2003). *Theory of Scheduling*. Dover Publications, NY, USA.
- Crovella, M. E. and Lipsky, L. (1997). Long-lasting Transient Conditions in Simulations With Heavy-tailed Workloads. In *WSC '97: Proceedings of the 29th conference on Winter simulation*, pages 1005–1012, Washington, DC, USA. IEEE Computer Society.
- Down, D. G. and Wu, R. (2006). Multi-layered Round robin Routing for Parallel Servers. *Queueing Syst. Theory Appl.*, 53:177–188.
- Feng, H. and Misra, V. (2004). On the Relationship between Coefficient of Variation and the Performance of M/G/1 - FB queues. *SIGMETRICS Perform. Eval. Rev.*, 32(2).
- Gupta, V. (2008). Finding the Optimal Quantum Size: Sensitivity Analysis of the M/G/1 Round-robin Queue. *ACM SIGMETRICS Performance Evaluation Review*, 36(2).
- Harchol-Balter, M., Crovella, M., and Murta, C. (1999). On Choosing a Task Assignment Policy for a Distributed Server System. *IEEE J. of Parallel and Distrib. Computing*.
- Jain, R. (1991). *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation and Modeling*. John Wiley and Sons, New York, NY, USA.
- MacDougall, M. H. (1987). *Simulating Computer Systems: Techniques and Tools*. MIT Press, Cambridge, MA, USA.
- Nuyens, M. (2004a). *The Foreground-Background Queue*. PhD thesis, Universit van Amsterdam - Met samenvatting in het Nederlands.
- Nuyens, M. (2004b). *The Foreground-Background Queue*. PhD thesis, Universit van Amsterdam - Met samenvatting in het Nederlands.
- Nuyens, M. and Wierman, A. (2008). The Foreground-Background Queue: A Survey. *Perform. Eval.*, 65(3-4):286–307.
- Schrage, L. E. (1968). A Proof of the Optimality of the Shortest Processing Remaining Time Discipline. *Operations Research*, 16:678–690.
- Schrage, L. E. and Miller, L. W. (1966). The Queue M/G/1 with the Shortest Remaining Processing Time Discipline. *Operations Research*, 14(4):670–684.
- Sousa, A. M. and Murta, C. D. (2010). Simulation of the Foreground-Background Queue in Parallel Systems. *XI WSCAD–SSC 2010*, pages 17–24.
- Tanenbaum, A. S. and Steen, M. V. (2007). *Distributed Systems Principles and Paradigms*. Prentice Hall, USA.
- Tanenbaum, A. S. and Woodhull, A. S. (2006). *Operating Systems Design and Implementation*. Prentice Hall, USA.
- Wierman, A. (2007a). Fairness and Classifications. *SIGMETRICS Performance Evaluation Review*, 34(4):4–12.
- Wierman, A. (2007b). *Scheduling for Today's Computer Systems: Bridging Theory and Practice*. PhD thesis, School of Comp. Science, Carnegie Mellon Univ., Pittsburgh.