

### **Escalonador WF<sup>2</sup>Q**

O escalonador WF<sup>2</sup>Q [3] é uma aproximação baseada em pacotes do GPS, que tem por objetivo emular este escalonador fluido o mais próximo possível. De acordo com Giroux e Ganti [1], o escalonador WF<sup>2</sup>Q é classificado como um escalonador com divisão justa e conservativa (*work-conserving fair-share scheduler*). Portanto, sempre que houver células aguardando para serem servidas, no mínimo uma célula será servida a cada *slot* de serviço do escalonador. Em outras palavras, o escalonador nunca é “desligado” enquanto existirem células aguardando para serem servidas.

Consideremos um conjunto de conexões  $B_{t_k}$  sendo servidas simultaneamente em um escalonador GPS no instante  $t_k$ . Para cada conexão  $i$  é atribuído um “peso”  $\phi_i$ , que determina a fração de serviço que a conexão  $i$  receberá. Como no GPS cada conexão tem exatamente um pacote sendo servido, a taxa de serviço instantânea para a conexão  $i$  é dada por

$$r_i = \frac{\phi_i}{\sum_{j \in B_{t_k}} \phi_j} \cdot r, \quad (1)$$

onde  $B_{t_k}$  é o número de conexões sendo servidas simultaneamente no escalonador no instante  $t_k$ ,  $\phi_i$  é o peso da conexão  $i$  no escalonador GPS,  $\phi_j$  é o peso da conexão  $j$  no escalonador GPS e  $r$  é a taxa total disponível no escalonador.

O funcionamento do escalonador WF<sup>2</sup>Q é bastante semelhante ao do escalonador PGPS (ou WFQ), porém com uma diferença na forma em que os pacotes a serem servidos são selecionados. Quando um escalonador PGPS escolhe um pacote para ser servido no instante  $t_s$ , ele seleciona, entre todos os pacotes presentes no servidor, o pacote que terminaria serviço primeiro no GPS. Já o escalonador WF<sup>2</sup>Q, ao invés de escolher o pacote que terminaria serviço primeiro no GPS dentre todos os pacotes presentes no servidor, escolhe somente entre os pacotes que já teriam iniciado serviço no GPS (pacotes elegíveis). Ou seja, no escalonador WF<sup>2</sup>Q, o

pacote escolhido para ser servido no instante  $t_s$  é o pacote que terminaria serviço primeiro no GPS e que já teria começado a ser servido pelo GPS no instante  $t_s$ .

Tanto o PGPS quanto o WF<sup>2</sup>Q utilizam o conceito de tempos virtuais [2] para emular o funcionamento do GPS e a partir daí escolher a ordem de serviço dos pacotes presentes no escalonador. Consideremos a chegada de um pacote da conexão  $i$  no instante  $t_k$  com tamanho igual a  $L_k^i$  bits. No momento da chegada deste pacote, um tempo virtual geral  $V_{t_k}$  é incrementado de acordo com a expressão

$$V_{t_k} = V_{t_{k-1}} + \frac{(t_k - t_{k-1})}{\sum_{j \in B_{t_k}} \phi_j}, \quad (2)$$

onde  $V_{t_{k-1}}$  é o tempo virtual geral no instante  $t_{k-1}$  e  $B_{t_k}$  é o número de conexões sendo servidas simultaneamente no escalonador no instante  $t_k$ .

Uma vez atualizado o tempo virtual geral, é necessário calcular então os tempos virtuais de início de serviço ( $S_{t_k}^i$ ) e de final de serviço ( $F_{t_k}^i$ ) deste pacote no GPS emulado. O tempo virtual de início de serviço  $S_{t_k}^i$  é calculado de acordo com a expressão

$$S_{t_k}^i = \max\{F_{t_{k-1}}^i, V_{t_k}\}, \quad (3)$$

onde  $F_{t_{k-1}}^i$  é o tempo virtual de final de serviço anterior.

Feito isso calcula-se o tempo virtual de final de serviço  $F_{t_k}^i$  do pacote

$$F_{t_k}^i = S_{t_k}^i + \frac{L_k^i}{\phi_i}. \quad (4)$$

Tanto no escalonador PGPS quanto no escalonador WF<sup>2</sup>Q, a ordem de serviço dos pacotes é determinada em função dos tempos virtuais  $S_{t_k}^i$  e  $F_{t_k}^i$ . Suponhamos que ambos os escalonadores queiram escolher no instante  $t_s$  um pacote para ser servido. Como vimos, no

PGPS é servido primeiro o pacote que terminaria serviço primeiro no GPS, ou seja o pacote que tiver o menor tempo virtual  $F_{t_k}^i$ , independente do instante de tempo  $t_s$ . Já no escalonador WF<sup>2</sup>Q, é servido primeiro o pacote elegível que terminaria serviço primeiro no GPS. Ou seja, é servido primeiro o pacote com menor tempo virtual  $F_{t_k}^i$ , cujo  $S_{t_k}^i$  seja menor que o instante de tempo  $t_s$ .

Portanto, o escalonador WF<sup>2</sup>Q, ao invés de utilizar uma disciplina de seleção de pacotes “menor tempo virtual de final de serviço primeiro” (SFF – *Smallest virtual Finish time First*), utiliza uma disciplina “menor tempo virtual de final de serviço elegível primeiro” (SEFF – *Smallest Eligible virtual Finish time First*). Ou seja, o escalonador WF<sup>2</sup>Q escolhe entre todos os pacotes elegíveis, o pacote que tem o menor tempo virtual de final de serviço.

Como vimos, o WF<sup>2</sup>Q serve primeiro o pacote com menor tempo virtual  $F_{t_k}^i$  cujo  $S_{t_k}^i$  seja menor que o instante de tempo de serviço real  $t_s$ . Para implementarmos esta disciplina de seleção de pacotes, utilizamos estruturas de dados do tipo filas com prioridades [4]. Tais filas possuem desempenho ótimo de inserção e de retirada de objetos, e foram implementadas utilizando árvores binárias balanceadas [4]. Duas filas com prioridades foram utilizadas: fila com prioridades para o  $F_{t_k}^i$  e fila com prioridades para o  $F_{t_k}^i$  temporária. Uma vez que nosso modelo de escalonador WF<sup>2</sup>Q foi desenvolvido voltado para a simulação de redes ATM, ambas as filas com prioridades são utilizadas para ordenar células ATM conforme os seus tempos virtuais de final de serviço  $F_{t_k}^i$ . A Figura 1 mostra a estrutura do modelo desenvolvido.

As células são armazenadas na fila com prioridades por  $F_{t_k}^i$  onde aguardam para serem servidas. No momento em que uma célula é servida, é verificado se a célula no topo da fila é elegível. Se a célula escolhida for elegível, ela é servida. Caso contrário, ela é armazenada na fila com prioridades por  $F_{t_k}^i$  temporária, e uma nova célula é verificada. O processo prossegue até que uma célula elegível seja escolhida ou até que a fila com prioridades fique vazia. Após uma célula ter sido escolhida, as células não elegíveis que se encontram na fila temporária são retornadas à fila principal.

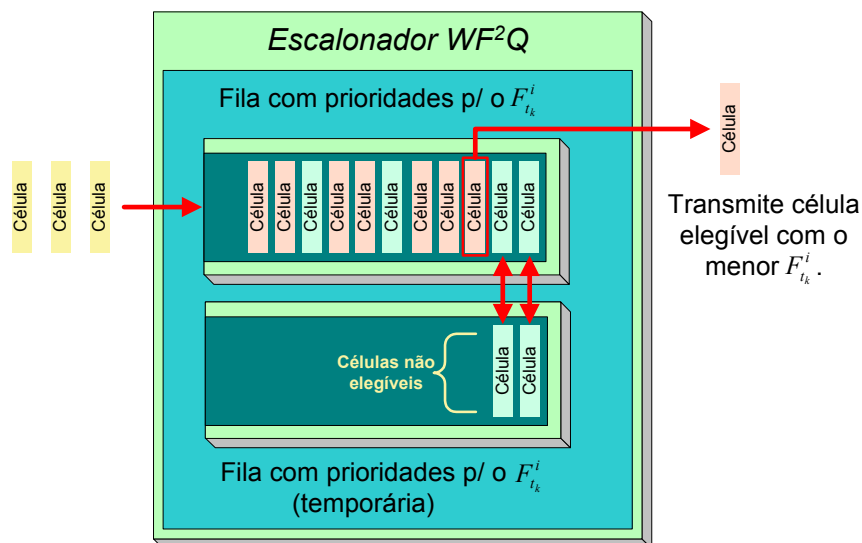


Figura 1 – Estrutura do modelo do escalonador WF<sup>2</sup>Q.

O modelo desenvolvido para o escalonador WF<sup>2</sup>Q possui dois algoritmos principais: algoritmo de armazenamento de células ATM e algoritmo de serviço de células ATM.

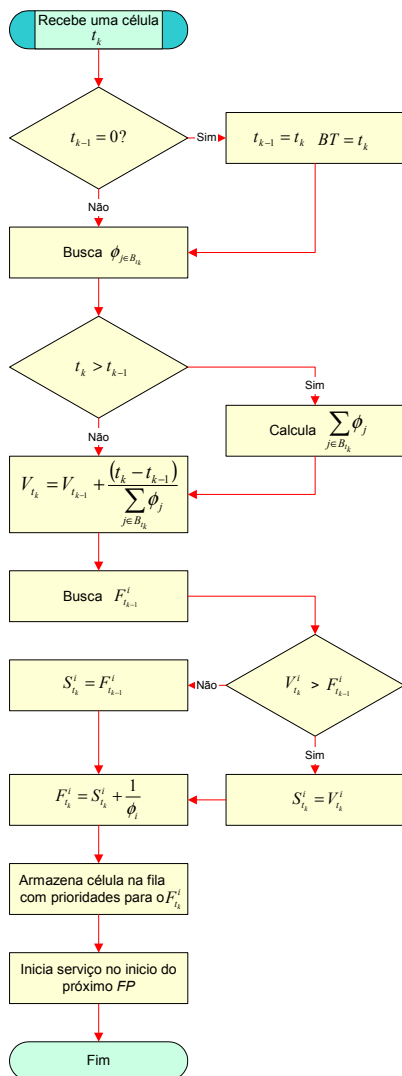
### Algoritmo de Armazenamento de Células ATM

A Figura 2 mostra o fluxograma do algoritmo de armazenamento de células ATM. Quando uma célula ATM é recebida (instante de tempo  $t_k$ <sup>1</sup>), o modelo do escalonador verifica se o tempo de chegada da célula anterior ( $t_{k-1}$ ) é igual a zero. Se esta variável for igual a zero, ela é configurada com o valor de  $t_k$ . Uma variável chamada *BT* (*Begin Time*) é utilizada para marcar o instante de tempo em que pelo menos uma conexão está presente no escalonador. Em outras palavras, o *BT* marca o início de um período de atividade. Se  $t_{k-1}$  for igual a zero, a variável *BT* também é configurada com o valor de  $t_k$ .

<sup>1</sup> Quando a primeira célula ATM é recebida  $k=1$ ,  $F_{t_0}^i=0$ ,  $S_{t_0}^i=0$  e  $V_{t_0}=0$ .

Ainda na Figura 2, o modelo do escalonador WF<sup>2</sup>Q busca o valor dos pesos para as conexões do conjunto  $B_{t_k}$ . Feito isso, é verificado se a variável  $t_k$  é maior que a variável  $t_{k-1}$ . Se esta condição for verdadeira, o somatório dos pesos de todas as conexões ativas é recalculado. Caso contrário, o mesmo valor obtido anteriormente é utilizado. Em seguida é atualizado o valor do tempo virtual do sistema ( $V_{t_k}$ ). Também é lido o valor do tempo virtual de final de serviço ( $F_{t_{k-1}}^i$ ) para a conexão  $i$  no instante  $t_{k-1}$ .

Se o tempo virtual  $V_{t_k}$  for maior que o  $F_{t_{k-1}}^i$ , o valor do tempo virtual de início de serviço ( $S_{t_k}^i$ ) para a conexão  $i$  é tomado como sendo igual ao  $V_{t_k}$ . Caso contrário, o valor do tempo virtual de início de serviço para a conexão  $i$  é tomado como sendo igual ao  $F_{t_{k-1}}^i$ . Então, o  $F_{t_k}^i$  é atualizado para o valor do  $S_{t_k}^i$  acrescido de  $1/\phi_i$ . A célula é então armazenada na fila com prioridades para o  $F_{t_k}^i$ , onde aguarda pelo serviço.



- $t_k$  - Tempo de chegada da célula atual
- $t_{k-1}$  - Tempo de chegada da célula anterior
- $\phi_i$  - Peso para a conexão  $i$
- $V_{t_{k-1}}$  - Tempo virtual para o instante  $t_{k-1}$
- $V_{t_k}$  - Tempo virtual para o instante  $t_k$
- $B_k$  - Conjunto de conexões ativas em  $t_k$
- $F_{t_{k-1}}^i$  - Tempo virtual de final de serviço para a conexão  $i$  em  $t_{k-1}$
- $F_{t_k}^i$  - Tempo virtual de final de serviço para a conexão  $i$  em  $t_k$
- $S_{t_k}^i$  - Tempo virtual de início de serviço para a conexão  $i$  em  $t_k$
- $FP$  - Período de um *frame* de serviço
- $BT$  - Início de um período de atividade no escalonador

**Figura 2 – Algoritmo de armazenamento de células ATM.**

## Algoritmo de Serviço de Células ATM

A Figura 3 mostra o fluxograma do algoritmo de serviço de células ATM. Quando uma primeira célula ATM é armazenada na fila com prioridades para o  $F_{t_k}^i$ , automaticamente o modelo do escalonador é “ligado” e o serviço de células começa no início do próximo *slot* de serviço. Uma única célula é servida no período de um *slot* de serviço ( $FP$  – *Frame Period*). Consideremos o caso em que uma célula ATM será servida no instante  $t_s$ . Inicialmente, é executado um laço na fila com prioridades para o  $F_{t_k}^i$ . O objetivo deste laço é encontrar a célula elegível com menor  $F_{t_k}^i$ . Para verificar a elegibilidade de uma célula, foi necessário desenvolver uma expressão que permitisse comparar o seu tempo virtual de início de serviço  $S_{t_k}^i$  no sistema GPS com o tempo real de serviço  $t_s$ . A seguinte expressão foi desenvolvida para possibilitar tal comparação

$$S_{t_k}^i \cdot FP + BT \leq t_s. \quad (5)$$

Uma vez que o equacionamento do escalonador WF<sup>2</sup>Q foi desenvolvido considerando-se uma taxa de serviço normalizada de 1 célula por segundo, a multiplicação do tempo virtual de início de serviço  $S_{t_k}^i$  pelo período de um *slot* de serviço ATM torna este tempo virtual compatível com a taxa de serviço  $FP$  do modelo do escalonador WF<sup>2</sup>Q. O acréscimo do tempo  $BT$  é necessário para reajustar o escalonador após um período de inatividade.

Portanto, a cada iteração do laço, o tempo virtual  $S_{t_k}^i$  da célula que se encontra no topo da fila com prioridades para o  $F_{t_k}^i$  é lido e a expressão (5) é utilizada para verificar a elegibilidade desta célula. Se a expressão (5) for verdadeira, a célula é considerada elegível, sendo, portanto, removida da fila  $F_{t_k}^i$  e servida. Neste caso a execução do laço acaba. Entretanto, se a expressão (5) não for verdadeira, a célula é considerada inelegível, sendo portanto removida da fila  $F_{t_k}^i$  e

armazenada na fila para o  $F_{t_k}^i$  temporária, onde aguarda para ser re-inserida na fila  $F_{t_k}^i$  principal. Neste caso, o laço prossegue verificando a elegibilidade das demais células armazenadas na fila  $F_{t_k}^i$  principal até que uma célula elegível seja encontrada ou até que esta fila principal fique vazia.

Como vimos, o escalonador WF<sup>2</sup>Q é um escalonador com divisão justa e conservativa. Portanto, se nenhuma célula da fila com prioridades  $F_{t_k}^i$  principal for elegível para ser servida no instante  $t_s$ , o escalonador deverá escolher igual uma célula para ser servida. Neste caso será servida a célula com menor  $F_{t_k}^i$  que se encontra na fila com prioridade temporária.

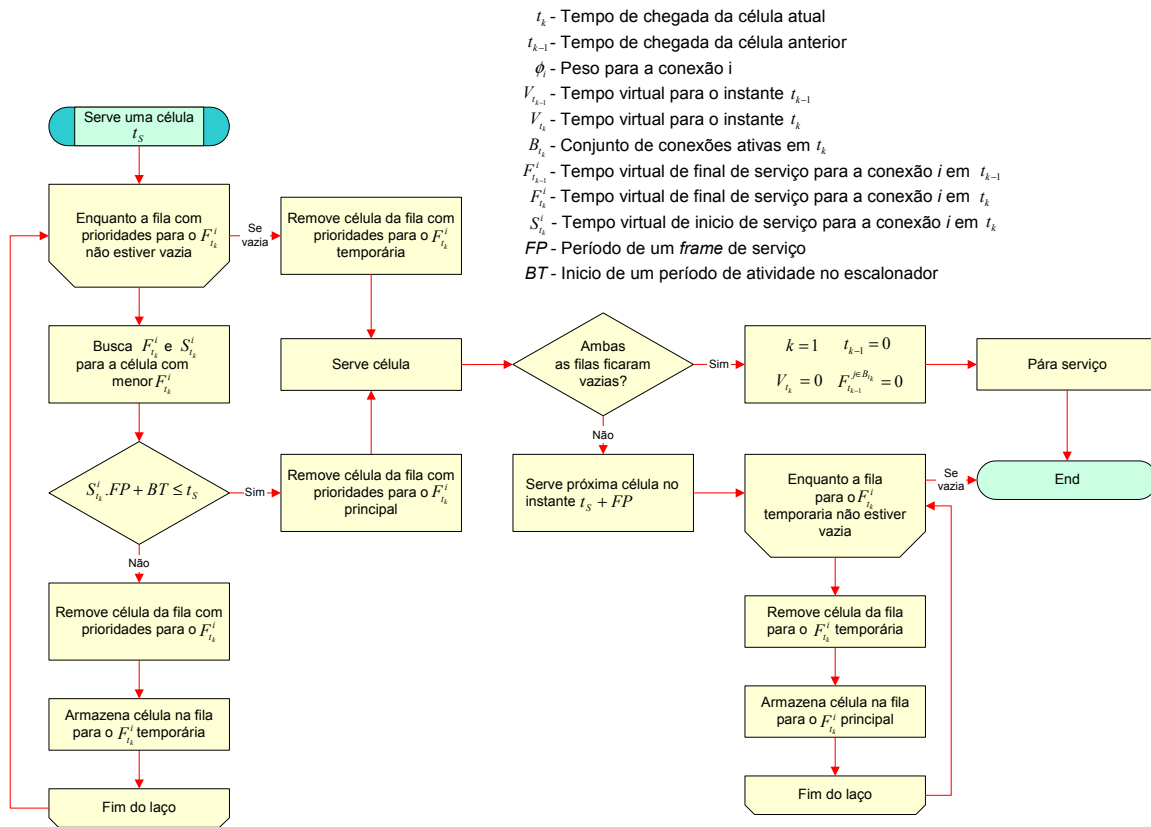


Figura 3 – Algoritmo de serviço de células ATM.



Depois de escolhida a célula a ser servida, é verificado se restaram células nas filas com prioridades principal e temporária. Se ambas as filas ficaram vazias, as variáveis  $k$ ,  $t_{k-1}$ ,  $V_{t_k}$  e  $F_{t_{k-1}}^{j \in B_{t_k}}$  são reinicializadas e o escalonador é “desligado”. Caso contrário, é agendando o serviço de uma próxima célula ATM no instante  $t_s + FP$ . Neste ponto outro laço é executado. O objetivo é retornar as células inelegíveis não servidas de volta para a fila principal.

- [1] GIROUX, N., GANTI, S., “Quality of Service in ATM Networks: State-of-Art Traffic Management”, Prentice Hall, 1998.
- [2] PAREKH, A. K., “A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case”, IEEE/ACM Transactions on Networking, Vol. 1. No 3. Junho de 1993.
- [3] BENNETT, J., ZHANG, H., “WF2Q: Worst-case Fair Weighted Fair Queueing”, IEEE Infocom 1996.
- [4] BUDD, T., “*Classic Data Structures in C++*”, Addison Wesley, 1994.
- [5] ALBERTI, A. M. “Desenvolvimento de Modelos de Simulação para a Análise de Qualidade de Serviço em Redes ATM”, Tese de Doutorado, 2003. Disponível em <http://libdigi.unicamp.br/document/?code=vtls000305133>.