SRR: An O(1) Time-Complexity Packet Scheduler for Flows in Multiservice Packet Network

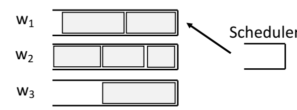
Ghuanxiong Guo

IEEE/ACM Transaction on Networking, Vol.12, No.6, December 2004

Presenters: Taeju Park, Yibo Pi

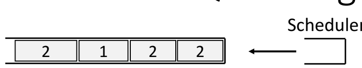
Introduction

- Different types of services on the Internet:
  - Delay insensitive: email
  - Delay sensitive: video and audio conferencing
- ***Resource isolation*** is needed to provide quality of service (QoS)
  - Flows are served based on their requirements
- Packet Scheduler
  - Decide which packet to be transmitted when the output link is idle

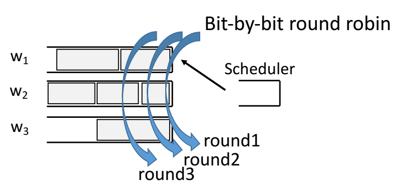


Fairness: the number of bits served for each flow is proportional to their weights.

Intro to Fair Queueing



First in first out (FIFO): no isolation among different flows



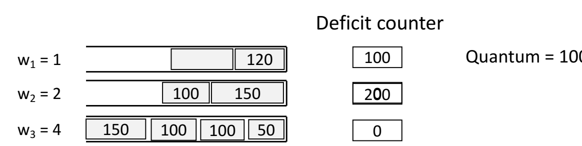
Generalized Processor Sharing (GPS)

In each round,
Flow 1: w_1 bits
Flow 2: w_2 bits
Flow 3: w_3 bits

GPS: ideal fairness, but not practical to use

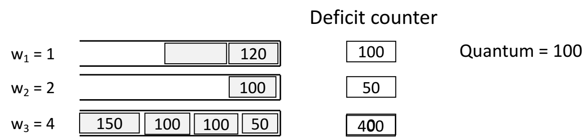
Packetized Queueing Schemes

- Weighted fair queueing (WFQ)
- Deficit round robin (DRR)



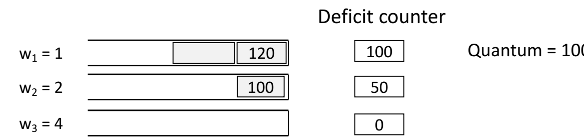
Packetized Queueing Schemes

- Weighted fair queueing (WFQ)
- Deficit round robin (DRR)



Packetized Queueing Schemes

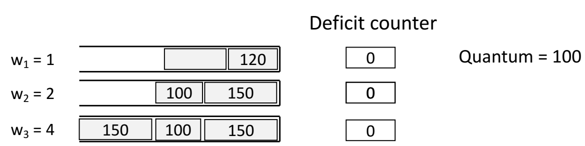
- Weighted fair queueing (WFQ)
- Deficit round robin (DRR)



Sequence of service: flow1, flow2, $\underbrace{\text{flow3, flow3, flow3, flow3}}_{\text{4 packets}}$

Problems of DRR: 1) bursty output and 2) short-term unfairness

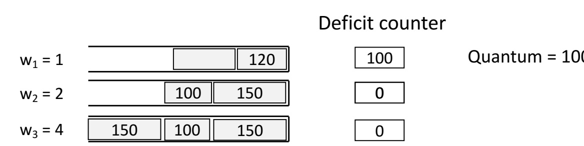
How to improve DRR?



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

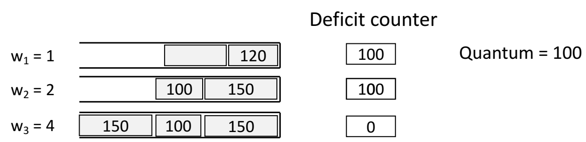
How to improve DRR?



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

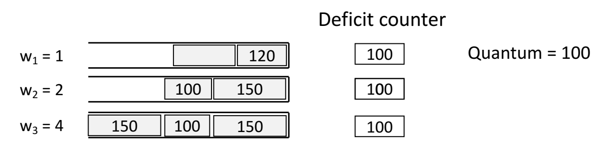
How to improve DRR?



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

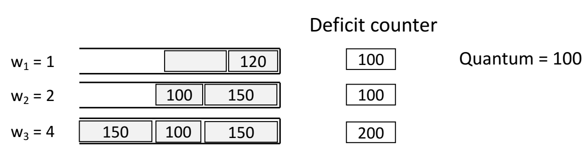
How to improve DRR?



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

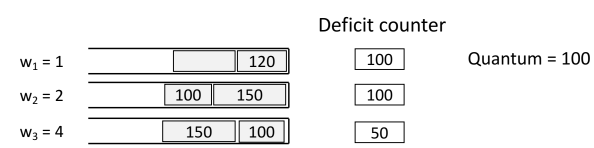
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

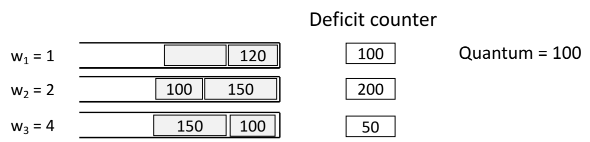
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

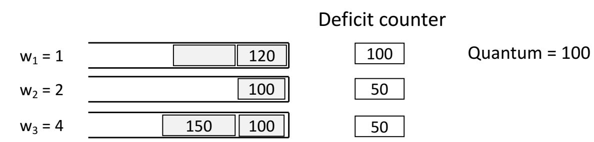
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

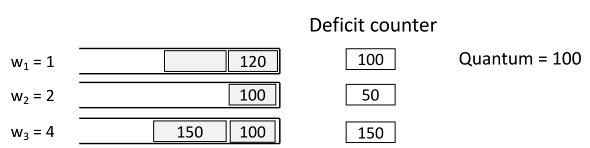
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

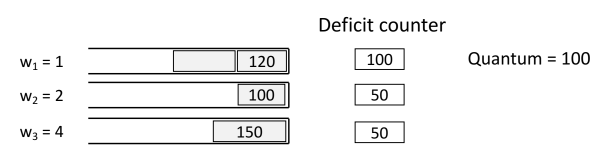
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

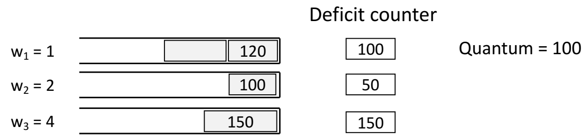
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

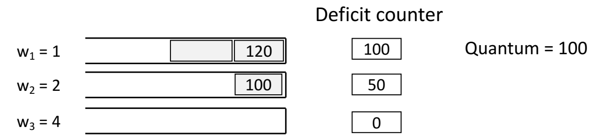
How to improve DRR



Sequence of service: flow1, flow2, flow3, flow3, flow3, flow3

sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

How to improve DRR?



sequence of service: flow1, flow2, flow3, flow3, flow2, flow3, flow3

2 packets 1 packet 2 packets

service of flow 1: 1
service of flow 2: 2
service of flow 4: 4

4 consecutive packets from flow 3 → 2 consecutive packets from flow 3

The design goal of Smoothed Round Robin

Weighted Fair Queueing
Pro: Short-term fairness
Con: high complexity O(# of active flow)

Round robin
Pro: low complexity O(1)
Con: short-term unfairness

Combine the pros of WFQ and round robin →

Smoothed Round robin
Short-term fairness +
low complexity O(1)

Weight Spread Sequence (WSS)

- WSS is a specially designed sequence that distributes the output traffic of each flow evenly.
- A set of WSSs is defined recursively as follows:
  - $S^1 = 1$
  - $S^k = \{a_i^k\} = S^{k-1}, k, S^{k-1}$
- Total number of terms in k^{th} WSS is $len_k = 2^k - 1$
- WSS Example
  - $S^3 = \{\underbrace{1,2,1}_{S^2},\underbrace{3}_{k},\underbrace{1,2,1}_{S^2}\}$
  - $S^5 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1,5,1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$
  - $len_5 = 2^5 - 1 = 31$

Weight Matrix

- Each flow is assigned a weight in proportion to its reserved rate.
  - (r_1 =64 kb/s, r_2 =256 kb/s, r_3 =512 kb/s, r_4 =192 kb/s) => (w_1 =1, w_2 =4, w_3 =8, w_4 =3)
- Weight of $flow_f$ is encoded as binary number $(4 = 100_2)$ in weight matrix

$$WM = \begin{bmatrix} WV_1 \\ \vdots \\ WV_N \end{bmatrix} = \begin{bmatrix} a_{1,(k-1)} & \cdots & a_{1,0} \\ \vdots & \ddots & \vdots \\ a_{N,\boxed{(k-1)}} & \cdots & a_{N,0} \end{bmatrix}$$

Row: weight vector of a flow
If weight is 10, then [1 0 1 0] where k = 4

Column number

The number of columns = order of WSS

Smoothed Round Robin Scheduler

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3$
  - Corresponding WSS,$S^4 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$

scan

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \end{bmatrix} = \begin{array}{c} \text{scan} \\ \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \\ c_3 \quad c_2 \quad c_1 \quad c_0 \end{array}$$

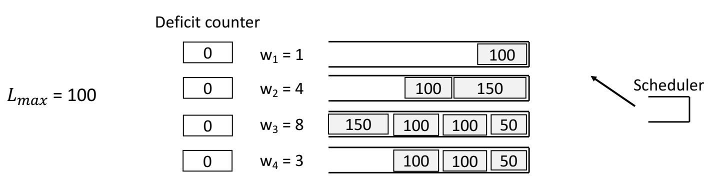
- Basic Idea of Smoothed Round Robin (SRR) Scheduler
  1. scan WSS sequence term by term
  2. When the value of the term is i, $column_{k-i}$ of the WM is chosen.
  3. In the column, the scheduler scan the terms from top to bottom.
  4. If the term is 1, the scheduler serve the corresponding flow.

Smoothed Round Robin

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3$
  - Corresponding WSS,$S^4 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

- Three asynchronous action
  - Schedule, Del_flow, Add_flow

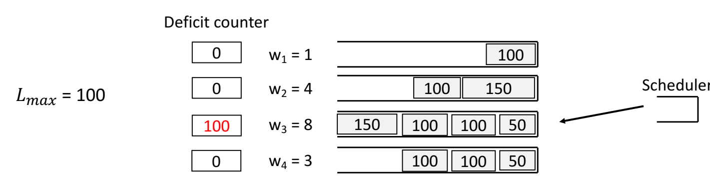


Smoothed Round Robin

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3$
  - Corresponding WSS,$S^4 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

- Three asynchronous action
  - Schedule, Del_flow, Add_flow

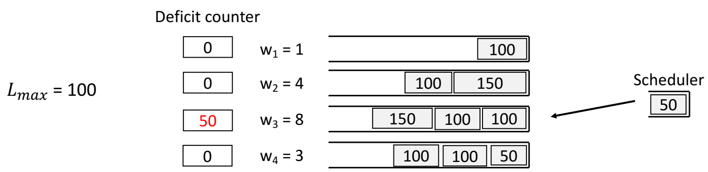


Smoothed Round Robin

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3$
  - Corresponding WSS, $S^4 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

- Three asynchronous action
  - Schedule, Del_flow, Add_flow

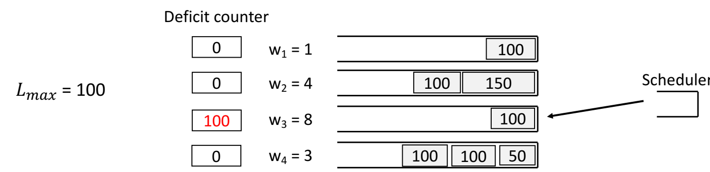


Smoothed Round Robin

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3$
  - Corresponding WSS, $S^4 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

- Three asynchronous action
  - Schedule, Del_flow, Add_flow

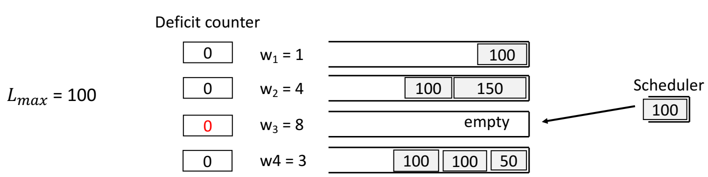


Smoothed Round Robin

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3$
  - Corresponding WSS, $S^4 = \{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

- Three asynchronous action
  - Schedule, Del_flow, Add_flow

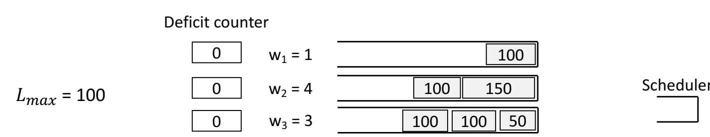


Smoothed Round Robin

- Four flows with fixed packet size (f_1, f_2, f_3, f_4) with corresponding weights (w_1, w_2, w_3, w_4)
  - $w_1 = 1, w_2 = 4, w_4 = 3$
  - Corresponding WSS, $S^3 = \{1,2,1,3,1,2,1\}$

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

- Three asynchronous action
  - Schedule, Del_flow, Add_flow

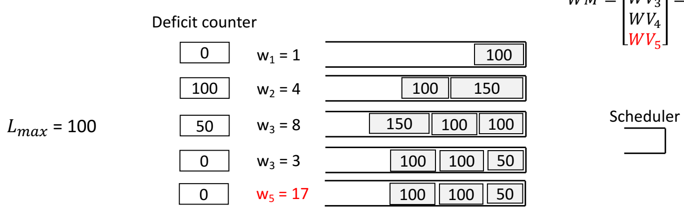


Smoothed Round Robin

- Five flows with fixed packet size $(f_1, f_2, f_3, f_4, f_5)$ with corresponding weights $(w_1, w_2, w_3, w_4, w_5)$
  - $w_1 = 1, w_2 = 4, w_3 = 8, w_4 = 3, w_5 = 17$ => Corresponding WSS,S^5
- Three asynchronous action
  - Schedule, Del_flow, Add_flow

added

$$WM = \begin{bmatrix} WV_1 \\ WV_2 \\ WV_3 \\ WV_4 \\ WV_5 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$



Properties of SRR

- Work-conserving
  - If there are active flows, the SRR always forward it.
- Theorem1: $flow_f$ is visited w_f times by SRR in a round
  - The number of received service by scheduler of each flow is proportional to its weight
  - The number of the occurrences of element i in $S^k (1 \le i \le k)$ is 2^{k-i}
    - The number of element 3 is in S^5 is $2^{5-3} = 4$

$$\{1,2,1,3,1,2,1,4,1,2,1,3,1,2,1,5,1,2,1,3,1,2,1,4,1,2,1,3,1,2,1\}$$

Properties of SRR: Fairness

- Lemma2 (Long-term fairness): For any pair of backlogged flows f and g, at the end of a round in SRR, then

End of a round

$$\left| \frac{V_f(0,\tau)}{w_f} - \frac{V_g(0,\tau)}{w_g} \right| = 0$$

$V_f(0,t)$ is the number of times that $flow_f$ is visited by SRR from time 0 to t

- Corollary1(Short-term fairness): For any pair of backlogged flows f and g in SRR, we have

$$\left| \frac{S_f(0,t)}{w_f} - \frac{S_g(0,t)}{w_g} \right| < \frac{(k+2)L_{max}}{2\min(w_f, w_g)}$$

$S_f(0,t)$ is service received by flows f from time 0 to t

Properties of SRR: Scheduling Delay Bound

- Scheduling Delay Bound (D_f)
  - Scheduling delay: time between queuing packet and transmitting the packet.
- Theorem3: The scheduling delay bound of $flow_f$ is

$$D_f < \frac{2L_{max}}{w_f} + (N-1)\frac{2L_{max}}{C}$$

N: the number of active flows

- Inverse proportional to the weight, proportional to total number of active flows
  - Cannot provide a strictly rate-proportional delay bound.
  - They claim that the delay bound is much better than that of DRR.

Properties of SRR: Scalability

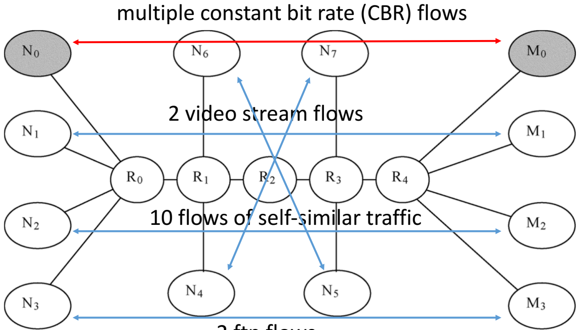
- Different rate ranges can be accommodates by WSS of the same order by adjusting the rate granularity
  - 1 kb/s rate granularity
  - 1 Mb/s rate granularity
- SRR can be used for variable bandwidth capacity
- SRR works well regardless of the number of flows.
  - Time complexity is O(1)

Properties of SRR: Complexity

- Space complexity
  - $len_k = 2^k - 1$ becomes very large if k is large number.
  - They claim that $K_{max} = 32$ is enough
    - It can provide 4 Tb/s rate with granularity of 1 kb/s
  - Since $(2k)^{th}$ WSS can be constructed by using k^{th} WSS and $(k+1)^{th}$ WSS, the space complexity of SRR is $2^{17} + \underbrace{O(N \times K_{max})}_{\text{To store } K_{max} \text{ double links.}}$
- Time complexity
  - $O(1)$ time to choose a packet for transmission
  - $O(k)$ time to add or delete a flow, where k is the order of WSS currently used by SRR.

Evaluation

Simulation tool: NS2



Weights of CBR flows are powers of 2

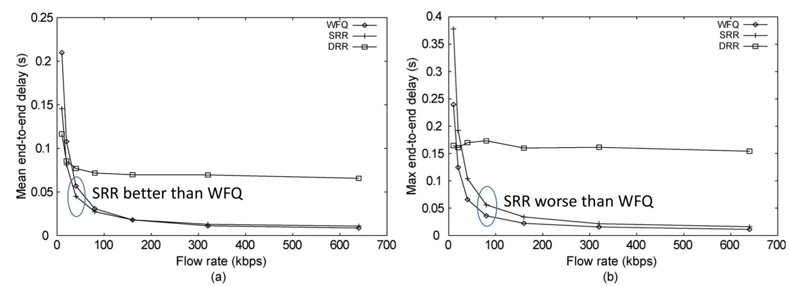


Fig. 4. (a) Mean delays of the CBR flows. (b) Maximum delays of the CBR flows. The weights of the CBR flows are randomly chosen.

Weights of CBR flows are randomly chosen

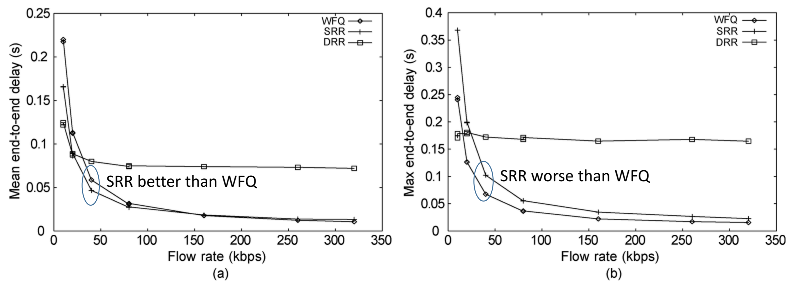


Fig. 5. (a) Mean delays of the CBR flows. (b) Maximum delays of the CBR flows. The weights of the CBR flows are randomly chosen.

Weights of CBR flows are equal

TABLE III
MAXIMUM AND MEAN DELAYS OF THE 10 CBR FLOWS WITH RATE 100 kb/s

Flow number	WFQ		SRR (DRR)	
	Max (ms)	Mean (ms)	Max (ms)	Mean (ms)
1	27.2	23.7	31.2	22.2
2	27.8	24.3	31.6	22.7
3	28.6	24.9	31.9	23.2
4	29.2	25.5	33.0	24.1
5	29.7	26.1	32.7	24.8
6	30.4	26.7	33.1	25.3
7	30.7	27.3	33.3	25.8
8	31.4	28.0	33.8	26.2
9	31.8	28.7	34.3	26.7
10	32.7	29.3	34.8	27.3

Discussion

- Weakness
  - The paper is not well written
  - Bad worst-case fairness
  - Ignore time overhead to construct high order WSS (32^{th} WSS) using low order (16^{th} & $17^{th}WSS$)

$$\left|\frac{S_f(0,t)}{w_f} - \frac{S_g(0,t)}{w_g}\right| < \frac{(k+2)L_{max}}{2\min(w_f, w_g)}$$

```
for each term a_i^j of S^j
  if (a_i^j = 1)
    replace a_i^j with S^{k+1};
  else
    a_i^j = a_i^j + k;
```

$O(Len_j)$

- Extension
  - Single scheduler to multi-scheduler fairness?
  - Singe resource (bandwidth) to multi-resource fairness?
  - Queue-independent fairness to queue-dependent fairness?