ATM FORUM: Technical Committee Working Group
ATM Forum/94-0077 (TM SWG)
****
SOURCE: Bellcore
Ljiljana Trajković
445 South St.
Morristown, NJ 07962-1910
Phone: 201 829-4531
FAX: 201 829-2504
Email: ljilja@thumper.bellcore.com

****
TITLE: Buffer Requirements in ATM Networks with Leaky Buckets, FIFOs, and Weighted Fair Queuing Scheduling Mechanisms
****
DATE: January 17-20, 1994
****
DISTRIBUTION TO: ATM Forum Technical Committee Working Group Members
****

ABSTRACT

In this contribution we present two studies on buffer allocation requirements in ATM networks. We first illustrate, by constructing a worst case deterministic scenario, that the buffer allocation requirements at various hops through the network may grow unexpectedly high if no cell loss is required even if the traffic patterns entering an ATM network are well regulated and conform to the agreed rates. We then present simulation examples comparing per VC buffer overflows in a node with FIFO and a node with a Weighted Fair Queueing (WFQ) scheme. Our simulation results illustrate that the WFQ scheme distributes losses (due to finite buffers) among traffic streams of various burstiness more fairly than the FIFO scheme.
****

****

1 Introduction

High speed packet networks will support diverse traffic such as voice, data, image, and real time video. In such a heterogeneous environment, different services will demand a range of Quality of Services (QoS) [1]. Some applications require no packet loss, while others are sensitive to packet delay or to packet delay-jitter. One of the goals of a good admission control scheme for ATM networks is to guarantee contracted performance for heterogeneous services while optimally using network resources (bandwidth) [2], [3], [4], [5], [6], [7], [8], [9], [10].

An ATM network with guaranteed QoS may support Virtual Channels of various bandwidth allocations. We assume that the sources on the outskirts of the network have pre-shaped traffic patterns that conform to the Connection Traffic Descriptor specified by the Generic Cell Rate Algorithm (GCRA). One ATM layer QoS that a network will provide is guaranteed service. This service will guaranty agreed upon low losses for customers who adhere to agreed Peak Cell Rates (PCR) defined by the GCRA [11], [12]. We show, by constructing a possible traffic scenario, that the buffer allocation requirements at various hops through the network may grow unexpectedly high for a zero loss QoS even if the traffic entering the network is well regulated and conforms to the agreed PCR.

2 Buffer Requirements in an ATM Network

An ATM network may support Virtual Channels (VCs) with various bandwidth allocations. Various VCs may be allocated a portion of the total ATM bandwidth (149.76 Mbps) and may have several hops where MUX/DMUX of ATM cells is performed. Network Elements such as MUXs and DMUXs may, therefore, processes VCs of various speeds. To illustrate required buffer space per VC at each hop we consider a scenario with a MUX/DMUX scheme and multiple VCs with distinct bit rates.

Consider two types of traffic streams shown in Figure 1(a): a fast stream A arriving to a MUX point at rate 74.88 Mbps, and five slow streams B arriving at rate 14.976 Mbps each. We assume a round robin service at the input ports. We also assume that the cells belonging to the slow VCs arrive first and that the cells belonging to the fast VC arrive last. The output stream L, shown in Figure 1(b), will then have the pattern shown in Figure 2.

Figure 3(a) shows stream C consisting of cells from stream A, after being DMUXed from stream L. Streams C and D have the bandwidth allocations of the original fast streams A and the total of streams B, respectively. Since we deal with a peak allocation service, the outgoing link is reserved for 149.76 Mbps in each case, as was done at the beginning of the MUX/DMUX process. Note that due to the MUXing and DMUXing of VCs with different bandwidth allocations the cells of stream C need to be buffered. In the example shown in Figure 3(b), up to three cells of each stream need to be stored in the MUX/DMUX point before being sent to the output link.

From the given example we can conclude that the buffer allocation requirements depend on the bandwidth ratio of MUXed and DMUXed VCs. The ATM network will support video, image, voice, and data traffic that will require vastly different bandwidths, and will, therefore, require substantial buffering for a zero loss QoS [13], [4], [5]. For example, in the case of MUXing a video application requiring 74.88 Mbps and a voice application requiring 64 Kbps bandwidth, we can calculate that the buffer allocation required for zero loss service. The calculation is performed by following the same argument as illustrated in the above example. When two VCs (voice and data) were MUXed and then DMUXed, the buffer allocation requirement is equal to $74.88 - (74.88)^2/(74.88 + 0.064)$ Mbits $= 63.95$ Kbits for the VC supporting the video application.

Figure 4 shows the buffer allocation requirement calculation for a general case of two streams involved in the previously described MUX/DMUX scenario. The two streams have rates i_1 and i_2 respectively. After being MUXed together into a single stream of rate $i_1 + i_2$, they are DEMUXed and the first stream of cells is sent out with the rate o_1. For a given time interval T, the buffer size is calculated as the difference between the rate of the incoming traffic $(i_1 + i_2)$ and the rate of outgoing traffic $o_1(t - s)$. (The service time s that might be needed to serve a cell in an ATM switch was taken into account for the sake of generality,) The buffer size required to assure no loss is equal to $i_1T - o_1(m-s)$, where $m = \frac{i_1 T}{i_1+i_2}$. Note that the buffer size allocation requirement is proportional to the ratio of the fast and slow VCs.

One may envision a worst case scenario of the case described above. Consider a network where at each physical multiplexing point an ATM cell arrives simultaneously with other cells into a FIFO queue. If all other cells on all lines happen to be served first, the buffer allocation requirement at all multiplexing points will be considerable and of magnitude similar to the example we presented.

3 How to Remedy the High Buffer Allocation Requirements

Our example shows how the buffer allocation requirements may grow in an ATM network with guaranteed service where the network supports VCs of various bandwidth allocations. The sources on the outskirts of the network have shaped traffic patterns with parameters that conform to a leaky bucket [14] traffic control scheme. We show that the buffer allocation requirements at various hops through the network may grow unexpectedly high for a zero loss QoS even if the traffic entering the network is well regulated and conforms to the agreed rates.

Thus, this example implies that leaky bucket mechanism is not sufficient to provide guaranteed zero loss QoS without extensive buffers provided at all MUX/DMUX points in the network. Conforming and regular traffic patterns at the outskirts of an ATM network will not be sufficient in reducing these buffer allocation requirements. Therefore, leaky bucket policing, without additional congestion control mechanisms, will not be able to provide guaranteed QoS. Furthermore, other admission control schemes (e.g., certain round robin schemes that would not take into account the speed of various VCs) might not be able to prevent similar high buffer allocation requirements.

A possible remedy to high buffer allocation requirements would be a scheduling algorithm for MUXing VCs with various speeds. Interleaving fast and slow cells in the MUXing points would prevent clustering of fast cells. This, in turn, will alleviate the need for buffering the fast cells in the DMUX network pints. One way to implement such interleaving would be a Weighted Fair Queueing (WFQ) [15], [16], [17], [18], [19], [20] of cells belonging to VCs of different bandwidths.

4 Simulation of FIFO and WFQ Schemes in a Network with Bursty Traffic

We simulated a network node with ten sources shown in Figure 5. Each source generates bursty traffic, with a geometric distribution of active and idle periods. The traffic is switched at the switching node (shown in Figure 5) that contains a finite size buffer (256 ATM cells). Traffic generated by each source is regulated per VC with leaky buckets that are used as a policing mechanism. Each leaky bucket has two associated parameters: leaking rate and bucket size. The internal structure of the example is shown in Figures 5 and 6.

We implemented two scheduling schemes: FIFO and WFQ. The WFQ scheme is based on the predicted packet finishing time that is used as the priority criterion for serving the packet. The finishing time t_{finish} is calculated based on the packet arrival time $t_{arrival}$ as: $t_{finish} = t_{arrival} + \frac{packet\ size}{leaky\ bucket\ rate}$. Note that the calculation that has to be performed per packet is simple and fast. Each packet leaving the source node is stamped with t_{finish} and is inserted into the buffer with the priority based on the calculated finishing time. (A variation of the WFQ scheme could include a QoS parameter that can be used to favor streams of packets coming form a particular node.)

The statistics of per VC overflows were collected at the FIFO and the WFQ buffers. Source f_0 is more bursty and has a larger leaky bucket size than the remaining nine sources. Figure 7 shows the policing action of the leaky bucket and the buffer overflows at the switch with FIFO scheduling scheme, for two sources f_0 and f_1. A comparison between FIFO and WFQ schemes is shown in Figures 8 and 9 for four sources. In the FIFO case the more bursty source takes advantage of the scheme and has smaller losses due to overflows than the sources with less bursty traffic. Note that this "unfair" service can be remedied by a WFQ scheme that distributes the losses due to finite buffer sizes more fairly.

5 Conclusions

We demonstrated how the buffer allocation requirements may grow in ATM networks with guaranteed service where the network supports VCs of various bandwidth allocations. This occurs even if the traffic entering the network is well regulated and conforms to the agreed rates. We suggested that a scheduling scheme ensuring interleaving fast and slow cells in the MUXing points would prevent clustering of fast cells. A possible implementation for such interleaving would be a Weighted Fair Queueing of cells belonging to VCs of distinct bandwidths. Furthermore, our simulations illustrate that such a scheme also distributes losses (due to finite buffers) among traffic streams of various burstiness more fairly than the FIFO scheme.

References

[1] J. Kurose, "Open Issues and Challenges in Providing Quality of Service Guarantees in High-Speed Networks," *CCR*, Jan. 1993, pp. 6–15.

[2] A. Demers, S. Keshav, and S. Shenker, "Analysis and Simulation of a Fair Queueing Algorithm," *Proc. ACM SIGCOMM 89,* Sept. 1989, *Computer Communications Review,* vol. 19, no. 4, 1989, pp. 1–12, and *Journal of Internetworking Research and Experience*, Sept. 1990, pp. 3–26.

[3] D. Ferrari and D. C. Verma, "A Scheme for Real-Time Channel Establishment in Wide-Area Networks," *IEEE Journal on Selected Areas in Communications,* April 1990, pp 368–379.

[4] S. J. Golestani, "A Stop-and-Go Queueing Framework for Congestion Management," *Proc. ACM SIGCOMM 90,* Sept. 1990, Philadelphia, PA, pp. 8-18.

[5] S. J. Golestani, "A Framing Strategy for Congestion Management," *IEEE JSAC,* vol. 9, no. 7, pp. 1064-1077, Sept. 1991.

[6] D. C. Verma, H. Zhang, and D. Ferrari, "Delay Jitter Control for Real-Time Communications in a Packet Switching Network," *Proc. TriComm '91,* Chapel Hill, NC, April 1991, pp. 35–43.

[7] L. Zhang, "VirtualClock: A New Traffic Control Algorithm for Packet Switching Networks," *Proc. SIGCOMM 90,* Philadelphia, PA, Sept. 24–27, 1990, pp. 19–29.

[8] C. R. Kalmanek and H. Kanakia, "Rate Controlled Servers for Very High-Speed Networks," *Proc. GLOBECOM,* San Diego, CA, Dec. 1990, pp. 12-20.

[9] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks—The Single Node Case," *Proc. IEEE INFOCOM '92*, Florence, Italy, May 1992, pp. 915–924.

[10] A. K. Parekh and R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks—The Multiple Node Case," *Proc. IEEE INFOCOM '93*, San Francisco, May 1993, pp. 521–530.

[11] CCITT, "Recommendation I.371: Traffic Control and Congestion Control in B-ISDN," Geneva, June 1992.

[12] The ATM Forum, "ATM User-Network Interface Specification," Version 2.2, June 21, 1993.

[13] A. I. Elwalid and D. Mitra, "Analysis and Design of Rate-Based Congestion Control of High Speed Networks, I: Stochastic Fluid Models, Access Regulation," *Queueing Systems*, vol. 9, 1991, pp. 29–64.

[14] J. S. Turner, "New Directions in Communications (or Which Way to the Information Age?)," *IEEE Comm. Magazine*, Oct. 1986, Vol. 24, No. 10, pp.8–15.

[15] A. Fraser, "Designing a Public Data Network," *IEEE Comm. Magazine*, Oct. 1991, pp. 31–35.

[16] E. L. Hahne, C. R. Kalmanek, and S. P. Morgan, "Fairness and Congestion Control on a Large ATM Data Network With Dynamically Adjustable Windows," *Proceedings Telletraffic and Datatraffic, in a Period of Change, ITC-13*, A. Jensen and V. B. Iversen (Eds.), North-Holland, 1991, pp. 867–872.

[17] S. Floyd, "Congestion with Multiple Congested Gateways in Packet-Switched Networks, Part I: One-Way Traffic," *CCR*, Oct. 1991, pp. 30–47.

[18] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance," available by anonymous ftp from ftp.ee.lbl.gov.

[19] T. E. Anderson, S. S. Owicki, J. B. Saxe, and C. P. Thacker, "High Speed Switch Scheduling for Local Area Networks," *Proceedings ASPLOS 92*, 1992, pp. 98–110.

[20] D. Clark, S. Shenker, and L. Zhang, "Supporting Real-Time Applications in an Integrated Services Packet Network: Architecture and Mechanisms," *ACM Proc. SIGCOMM 92*, Baltimore, MA, 1992, pp. 14–26.

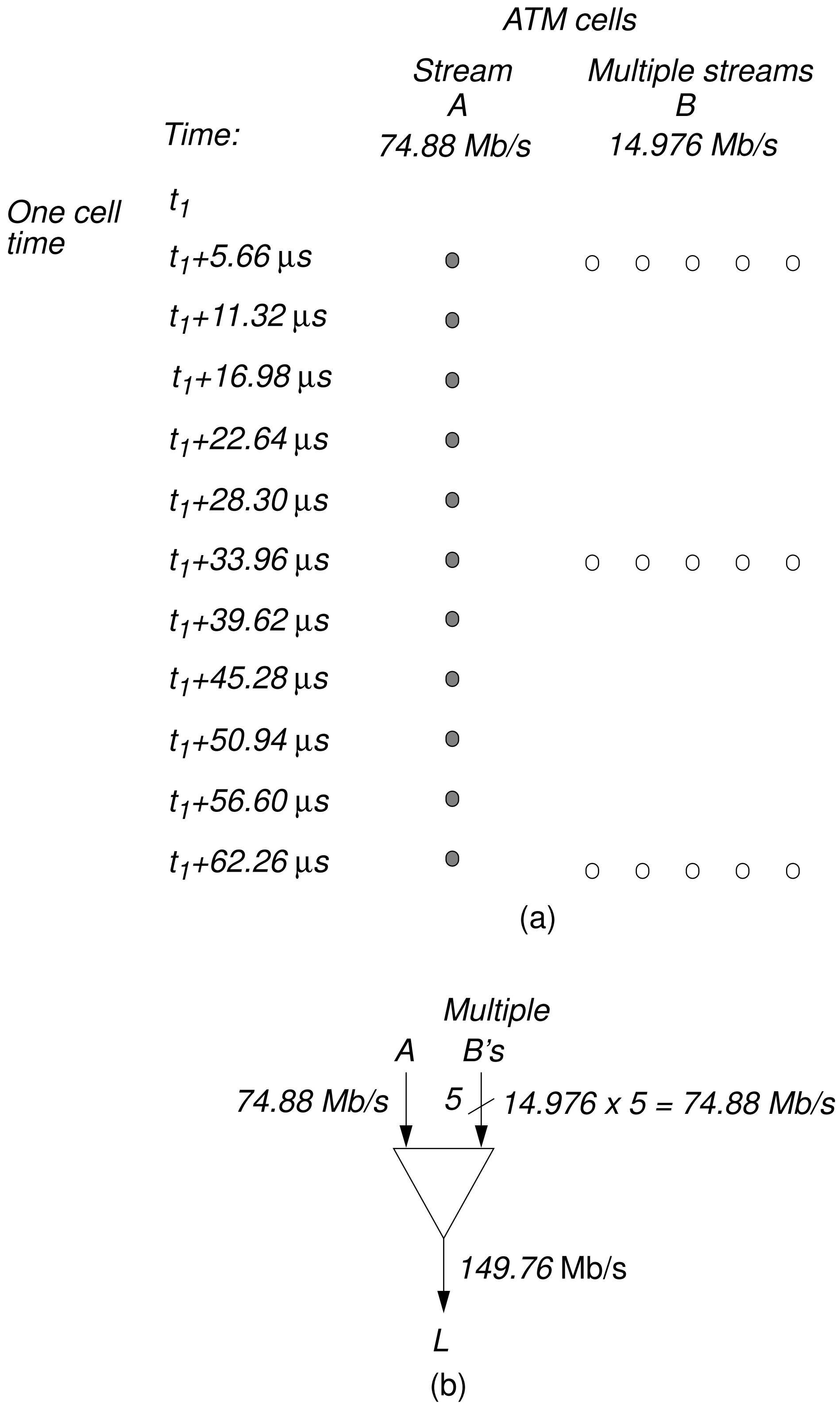


Figure 1: (a) Fast and slow ATM streams arriving to a MUX point. (b) MUX point with one A stream and multiple (five) B streams. Time scale arbitrarily starts at t_1 in increments of one cell time for stream A (5.66 μsec).

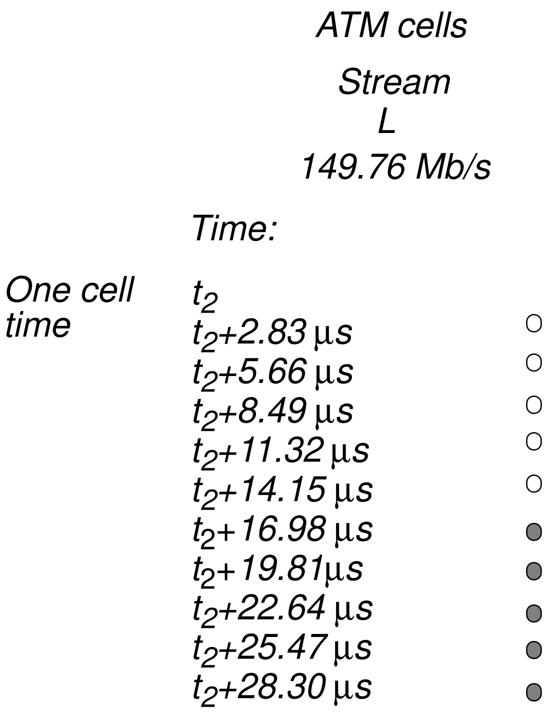
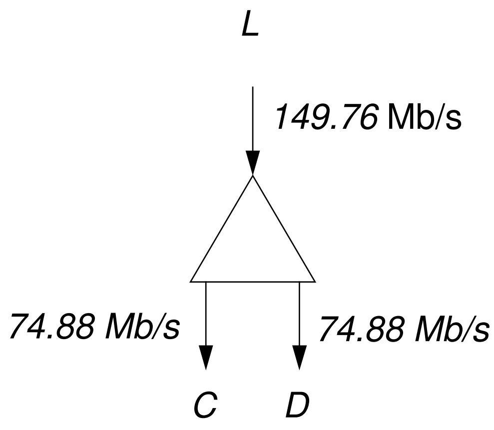
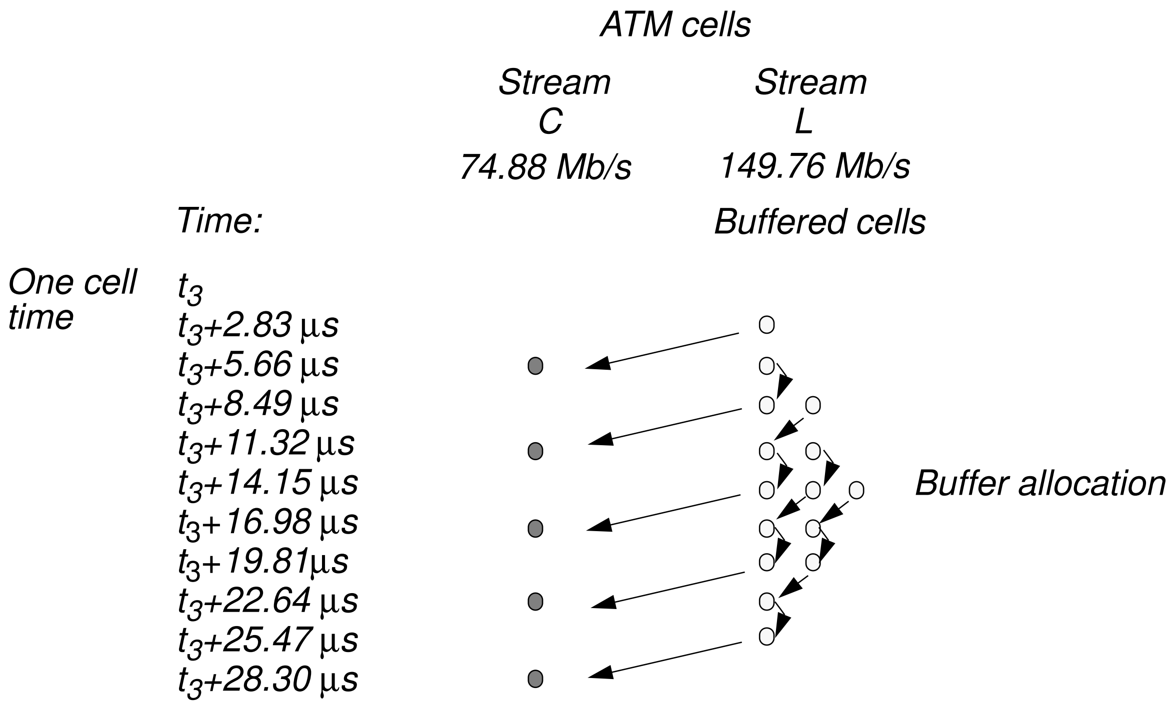


Figure 2: Traffic stream L after the MUX point. Time scale starts at t_2 in increments of one cell time for stream L (2.83 μsec).



(a)



(b)

Figure 3: (a) DMUXing the streams. (b) The buffer requirements at the DMUXing point. The arrows show the buffering process of arriving cells before they can be transmitted. Time scale starts at t_3 in increments of one cell time for stream L (2.83 μsec).

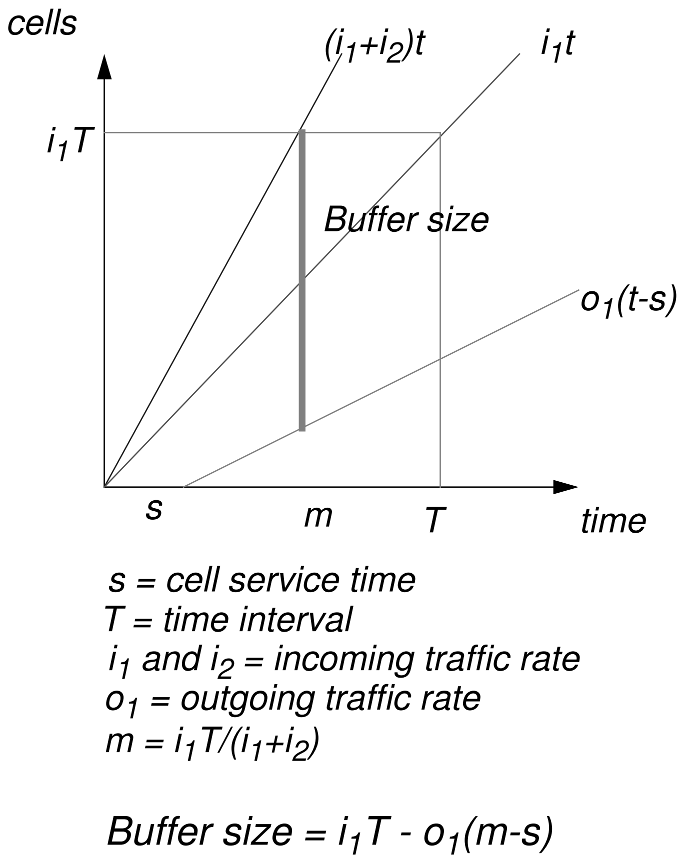
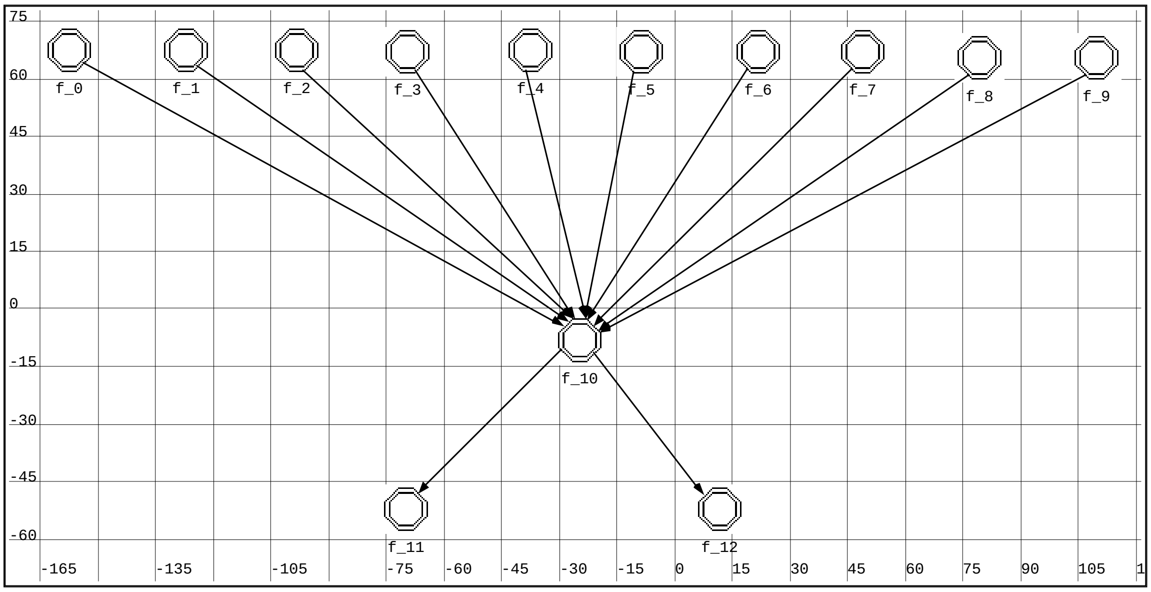


Figure 4: Buffer size estimate.

Example 1: network topology and the switch.



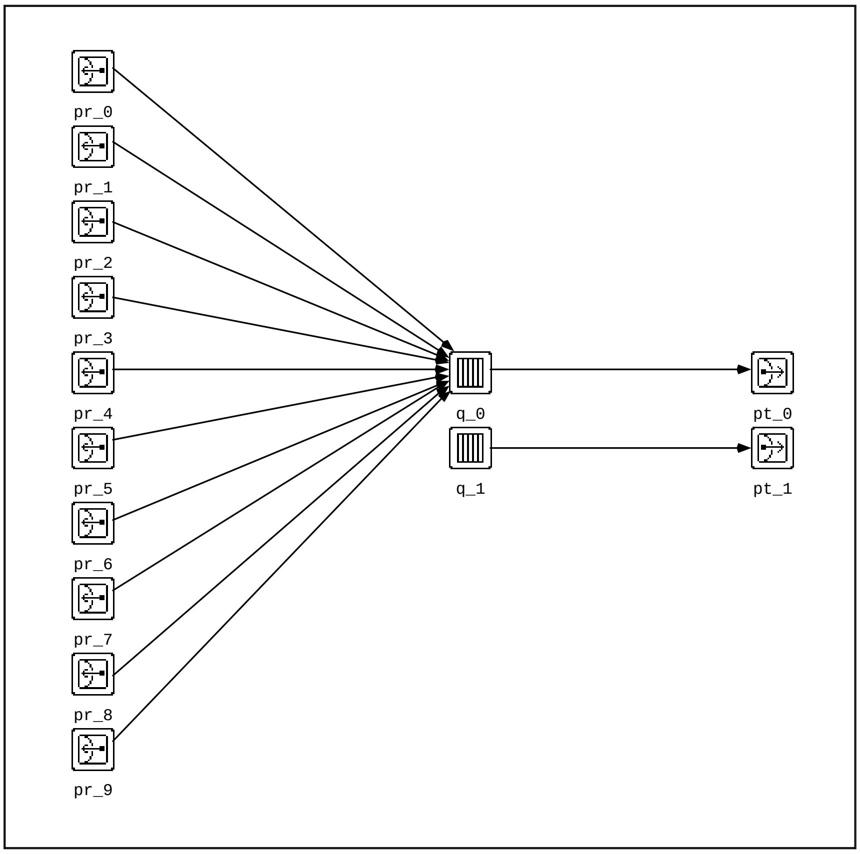


Figure 5: Simulation example: network topology.

Example 1: FIFO and WFQ internal nodes.

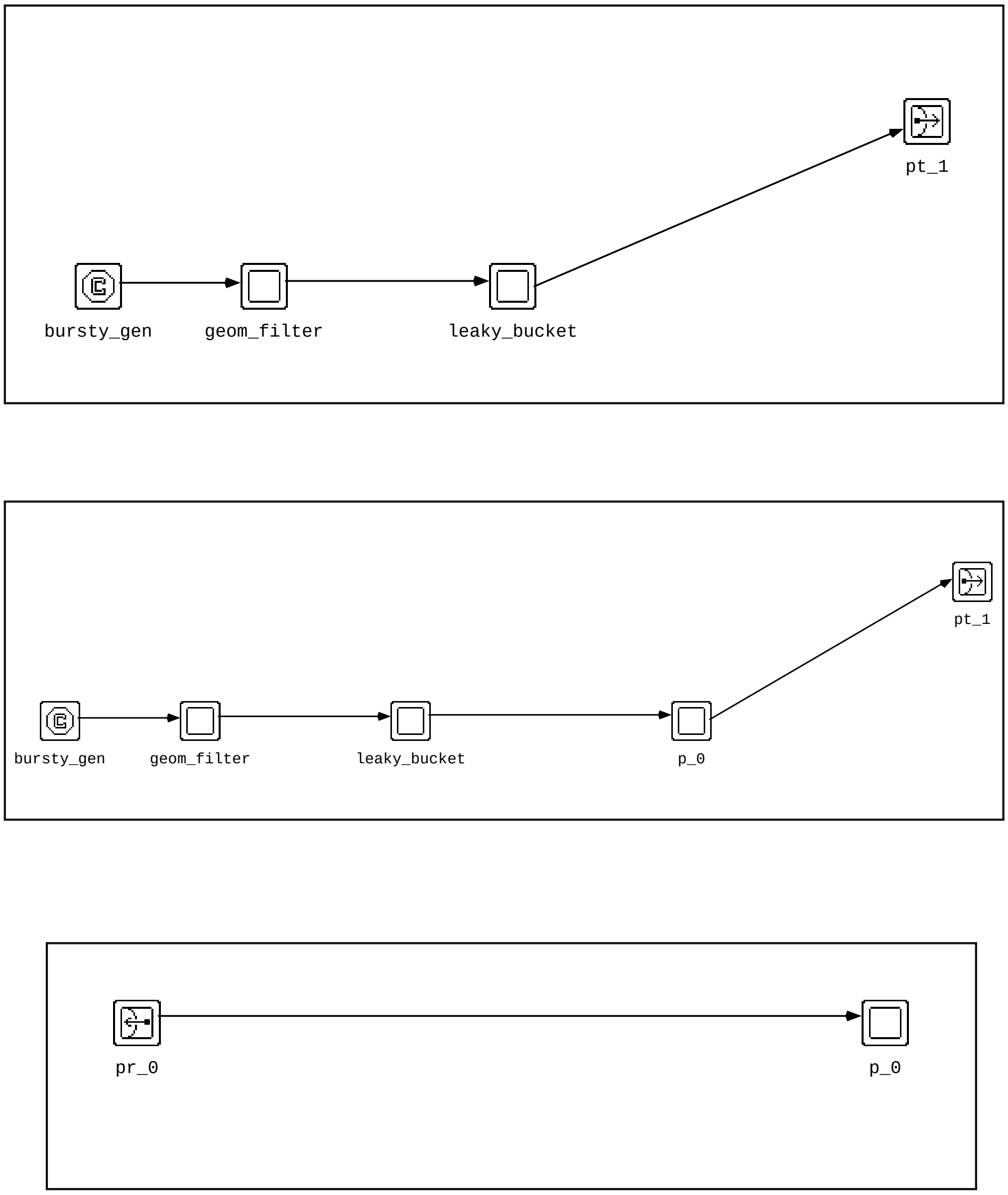


Figure 6: Simulation example: FIFO and WFQ internal nodes.

Example 1: leaky buckets and FIFO, f_0 = 500, f_1 = 200.

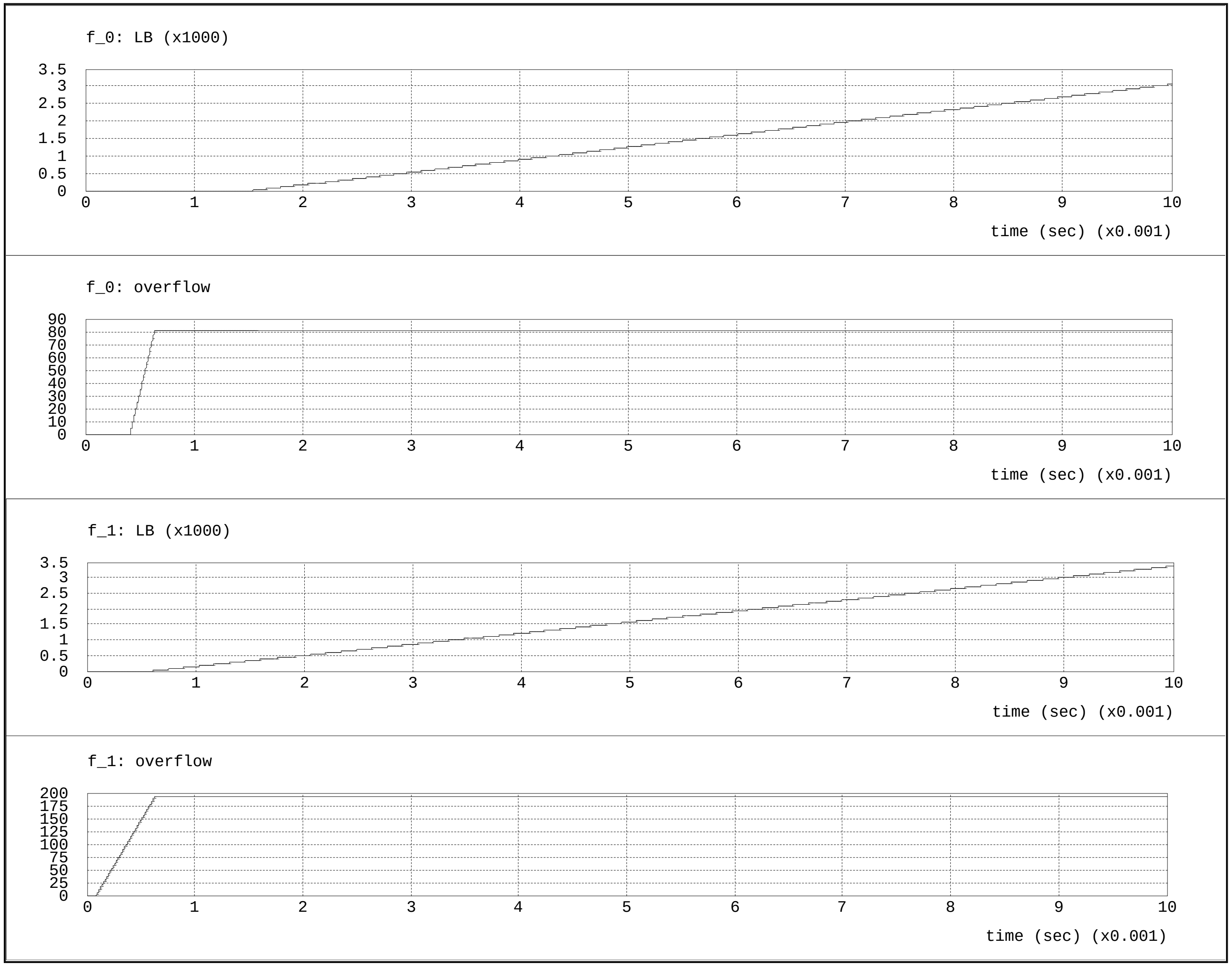


Figure 7: Simulation example: policing and buffer overflows.

Example 1: FIFO vs. WFQ for streams f_0 and f_1.

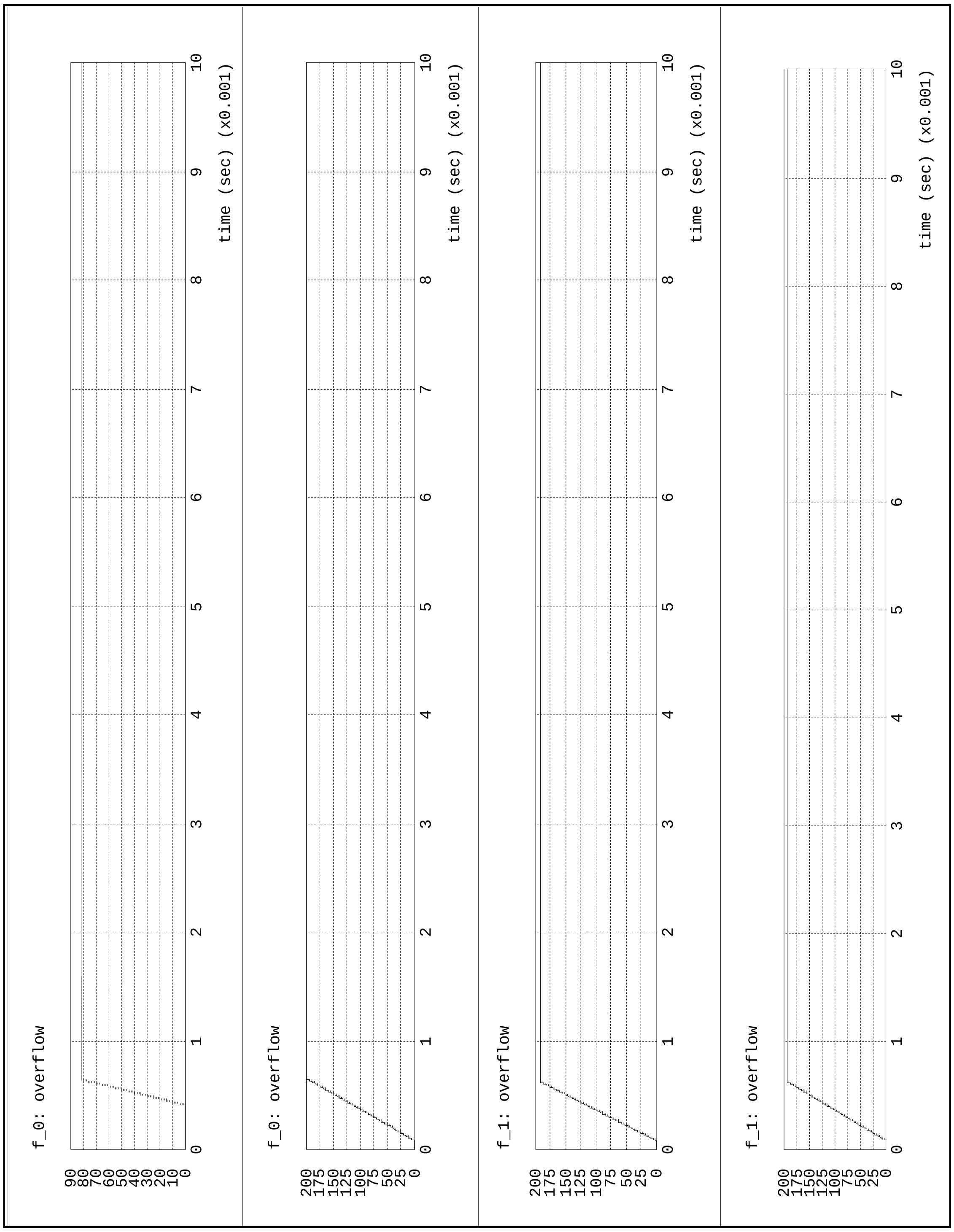


Figure 8: Simulation example: FIFO vs. WFQ buffer overflows in streams with different bursts.

Example 1: FIFO vs. WFQ for streams f_2 and f_3.

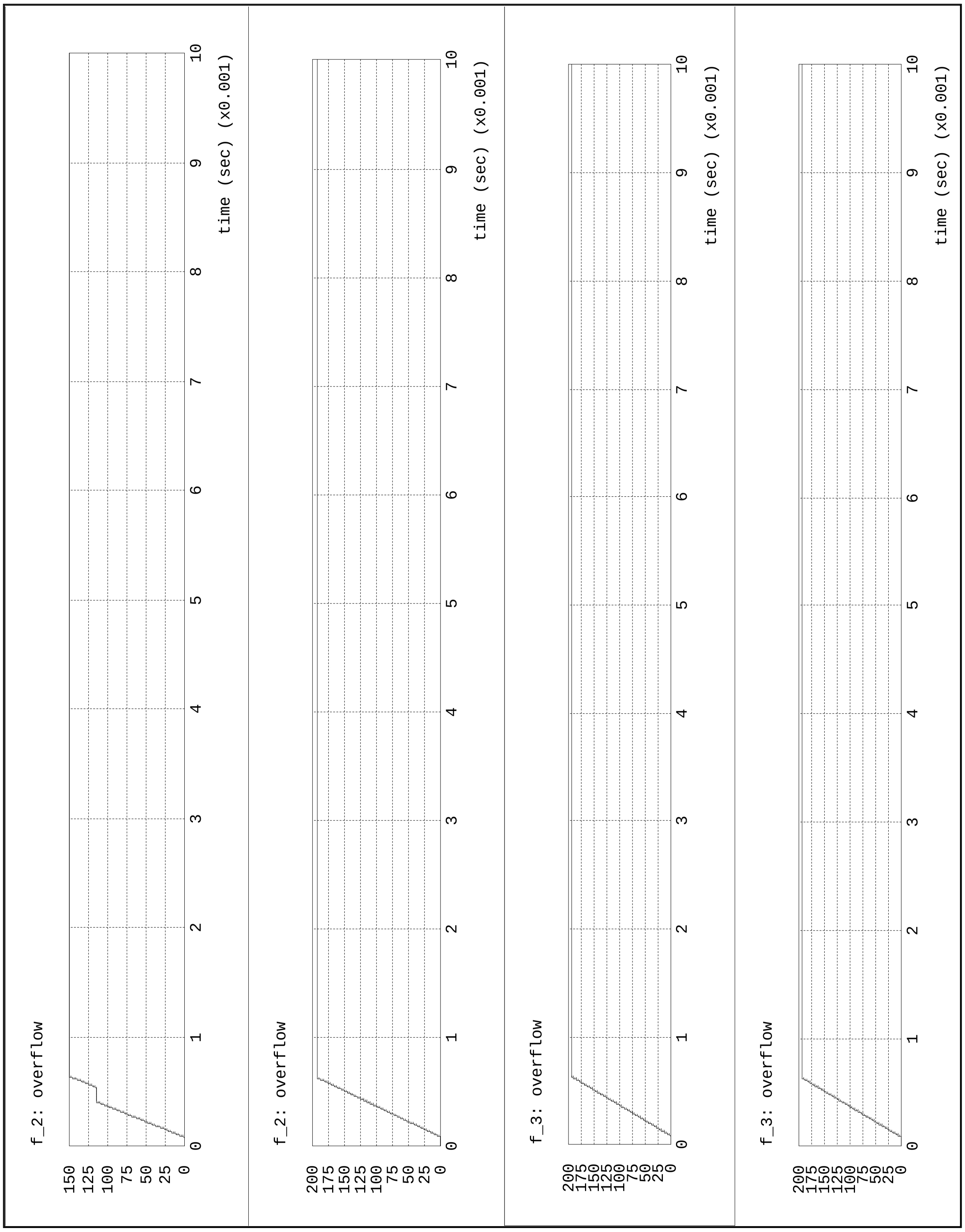


Figure 9: Simulation example: FIFO vs. WFQ buffer overflows in streams with equal bursts.