SRR: An O(1) time complexity packet scheduler for flows in multi-service packet networks[1]

Reviewed by
Ryan Marcotte and Xintong Wang

Summary

The author proposes Smoothed Round Robin (SRR), a fair queueing scheme, to overcome the burstiness issue that many ordinary round-robin schedulers have. SRR scheduler uses a Weight Matrix, which codes the weights of flows into binary vectors, and a Weight Spread Sequence (WSS), which distributes the output traffic of each flow evenly to achieve short-term fairness, certain scheduling delay bound and O(1) time complexity. The author further shows that the space needed to store the 32nd WSS is practical and is enough for current packet networks. Simulations are run to compare the mean and worst-case end-to-end delay of SRR with the delay of WFQ and DRR in three different settings, where the author varies the weights of the CBR flows. Experimental results show that the end-to-end delay property of SRR is similar to that of WFQ and should be a qualified scheduler for services that do not have strict end-to-end delay requirements.

Novelty and Contributions

1. Four desirable properties of a scheduler

From a narrative point-of-view, we commend the author's explanation of the four desirable properties (low time complexity, fairness, delay bounds, and simplicity) of schedulers and why they are hard to achieve simultaneously. Framing the previous state-of-the-art in the context of these properties ("[the] simplicity and time complexity properties always collide with the fairness and delay bound properties") immediately primes the reader to appreciate a primary contribution of this work: the simultaneous achievement of *all* of these properties in a scheduling algorithm.

[1] G. Chuanxiong, "SRR: An O(1) time complexity packet scheduler for flows in multi-service packet networks," *ACM SIGCOMM Computer Communication Review*, vol. 31, no. 4, pp. 211–222, Oct. 2001.

2. The scalability of SRR

The author makes several notes on the scalability of SRR that are worthy of comment. First, it seems highly useful that SRR can accommodate a wide range of rates simply by adjusting the rate granularity. This allows the same SRR implementation (specifically the same WSS) to operate in different environments under different traffic conditions. For example, the author points out that a similar WSS could be used in both core routers and edge routers. In addition, it is exciting that the functioning of SRR is independent of the performance of the underlying link, which allows SRR to maintain fairness even as the bandwidth of the output link may change. Next, one of the most important aspects of SRR is its constant time complexity, which positions it well for use in high-speed networks. Finally, the very useful property of WSSs that they can be constructed easily from smaller versions is interesting. This allows for much more scalability than their exponential space complexity would otherwise allow. The wide applicability of SRR through these aspects of scalability makes it an attractive scheduling algorithm.

3. Double link data structure

The double-link data structure, which is used to link the non-zero terms in a column of the Weight Matrix, is a key and interesting design that helps SRR maintain a low time complexity. Based on whether the end of a double-link of a column is reached or not, the algorithm will decide whether to enter the loop portion of the algorithm, which should be visited at most twice when transmitting a packet. Therefore, double-link acts as a signal to update the pointers. Though this data structure is not necessarily at the core of the SRR algorithm, it is a useful implementation component that aids in efficient functioning of the algorithm.

Possible Improvements and Extensions

1. Inadequacy of O(log N) *time complexity for schedulers*

The author's argument for why the *O(log N)* schedulers are not good enough for high-speed links is not very compelling. He states that it takes approximately 80 ns to transmit a

100-byte packet on a 10 Gbps link, so the scheduler must finish packet selection within the same time period regardless of the number of flows. This argument starts off promising to motivate the inadequacy of *O(log N)* schedulers (which of course do take longer with larger N), but the author fails to complete this line of reasoning. With time complexities this small (e.g. *O(log N)* or *O(1)*), constant factors in a particular algorithm (or even a particular implementation) may dominate until N grows very large. Therefore, to better argue that existing *O(log N)* schedulers are inadequate, the author could have shown how the runtimes of these algorithms grow beyond the necessary bounds (e.g. 80 ns) for numbers of flows that are of interest. This questionable justification is continued much later in the paper when discussing the complexity of SRR (Section 4.4). The author states that in cases in which a flow is not always backlogged, the $O(K_{max})$ complexity of adding and deleting flows is comparable to the *O(log N)* complexity of competing methods. However, there is no further justification that such a statement is true or applicable for practical scenarios.

2. Experimental results

For the simulation and experiment part, while mainly presenting the end-to-end delay metric, the author fails to show any experimental comparisons on fairness and local scheduling delay, which are equally important properties as shown in the theoretical part. Though the author directs the readers to his PhD dissertation for more experimental results, we failed to find his dissertation and any supplementary materials with respect to experiments on fairness and local scheduling delay online. It would be interesting if we could see comparisons of fairness between SRR and WFQ, especially under the case that weights of the CBR flows are in the diagonal. Since as the author states "the end-to-end delay of SRR is similar to that of WFQ under various conditions", the author could possibly avoid repetition by combining some of the results on end-to-end delay and save more space for experimental results on fairness and local scheduling delay. This will help confirm the theoretical section and make the paper more well-rounded. Another possible improvement is to consider the case of adding or deleting a flow in the simulation. By varying on the frequency of adding or deleting a flow, we can understand its impact on the overall performance of SRR. Further comparisons could also be made among SRR, WFQ and revised DRR.

3. Presentation of results

As we discussed in the previous section, we would have liked to see much more thorough experimental validation of the presented ideas, especially in comparison to competing algorithms. Another weakness we see is in the actual presentation of the results that do appear in the paper. First, the service curves in Figure 1 are not adequately explained as the author seems to assume that readers would already have seen such plots before and would be able to interpret them on their own. This is a bad assumption. Figures and experimental results are critical components of conveying an academic idea and should be able to be understandable on their own. A second weakness occurs in Figure 4 (a and b), which serve as the primary evaluation of the idea presented in the paper. The author argues about the relative merits of SRR to WFQ and DRR based on the shapes of the particular curves in the figures. However, the plots are shown without any error bars to indicate the statistical significance of these findings. Furthermore, there are no mentions of such considerations in the text. Overall, these factors combine to make for a weak experimental evaluation of an otherwise promising idea.

4. Intuition and motivation of proofs

From an expository standpoint, it would be helpful to the reader if the author motivated the various mathematical formalisms he presents before diving into their details. Mathematical propositions and proofs, especially those that are highly symbolic, are much more easily digestible once their purpose is explained sufficiently. It may not be possible to fully appreciate or understand a particular mathematical statement without grokking its details, but that is not an excuse for failing to provide initial context and motivation to the reader. As one example, it would be nice to receive more intuition on Weight Spread Sequences before describing them in a formal manner. The similarity between the Hanoi tower and the definition of WSS, as the author briefly mentioned but failed to elaborate, would be a good starting point to provide the readers some intuition. These WSSs and their interesting properties are arguably the most important insights in the whole paper, yet they are presented initially without any indication of why they have this particular structure. In some cases, the author later explains the utility of WSS properties, but other times the reader must fill in the gaps for himself.

5. Minor errors

Besides the expository aspect, the paper also contains several typos, including misspellings on page 1147 and the wrong title for Figure 4 (should be *The weights of the CBR flows are in the diagonal*), and should be proofread. These types of issues are often excused in writing as being minor or unimportant, but they can have a significant effect on the overall readability and understandability of a document. Furthermore, in this type of academic writing, such errors can detract from the perceived credibility of the authors and their claims.