

Bandwidth Utilization in Sorted-Priority Schedulers

Tae Joon Kim

Kongju National University, 275 Budaе-Dong, Cheonan, Chungnam, 330-240, Korea,
tjkim@kongju.ac.kr

Abstract. This paper first introduces bandwidth utilization metric and then analyzes sorted-priority schedulers in the terms of the metric. The results show that the utilization is directly proportional to both the number of delay bound classes and the dependency of delay bound on rate but inversely proportional to packet size.

1 Introduction

Packet scheduling algorithm has been extensively studied in the last decade due to its importance in the provision of Quality of Service (QoS) guarantees. Numerous sorted-priority scheduling algorithms have been developed to emulate the ideal algorithm called General Processor Sharing (GPS) [1]: Weighted Fair Queuing (WFQ) [2] has an ideal latency property with the complexity of $O(V)$. The extreme complexity has been significantly reduced in Self-Clocked Fair Queuing (SCFQ) [3] with sacrificing the latency. The theory of Rate Proportional Schedulers (RPS) [4] was formulated to reduce the complexity without deteriorating the latency, and then applied to various schedulers [5][6]. Now, sorted-priority schedulers become to achieve both the latency of WFQ and the complexity of $O(\log V)$.

For each flow i with the desired rate of r_i and the maximum packet size of M_i , its latencies in RPS based and SCFQ schedulers, denoted by Q_i^{RPS} and Q_i^{SCFQ} , respectively, are expressed as follows: $Q_i^{RPS} = M_i/r_i + M/G$ [4] and $Q_i^{SCFQ} = M_i/r_i + \sum_{k=1, k \neq i}^V M_k/G$ [3], where M is the maximum packet size in the scheduler, G is the capacity of outgoing link termed scheduler bandwidth and V is the maximum number of flows that the scheduler can admit. When the latency violates required delay bound, the scheduler should reduce it with even raising the rate reserved for the flow and consequently the bandwidth corresponding to the raised rate will be lost. This loss can not be, unfortunately, evaluated by the three legacy metrics of latency, complexity and unfairness used in previous works [2-6]. In this paper, we first introduce bandwidth utilization metric and then analyze sorted-priority schedulers in the terms of the metric.

2 Utilization metric

From the latency equation of Q_i^{RPS} we can see that scheduler can improve the latency of each flow as much as it needs with raising its reserved rate. Thus,

bandwidth utilization may be more useful than the latency used as a key metric in previous works [2-6], and, in addition, it enables for us to exactly compute the effective capacity of the scheduler. We define the bandwidth utilization ρ in a scheduler as the ratio of the amount of bandwidth practically used in servicing traffic flows requiring QoS guarantees to the amount of scheduler bandwidth reserved for them.

Now, let us derive a general expression of the bandwidth utilization. For each flow i with the desired rate of r_i , critical rate r_i^{crt} and reservation rate r_i^{res} are introduced: r_i^{crt} is the minimum rate needed to satisfy the delay bound that the flow requires and r_i^{res} is the rate that the scheduler should reserve for the flow in order to simultaneously guarantee both desired rate and required delay bound. r_i^{res} is equal to $\max(r_i, r_i^{crt})$. Bandwidth loss due to the excess reservation rate of $(r_i^{res} - r_i)$ is termed reservation loss. Let us define the desired rate of each flow as a random variable R distributed within $[a, b]$ and $r^{req} \equiv E[R]$. Then

$$\rho = \frac{1}{G} \sum_{i=1}^V r_i = \frac{r^{req}}{G}, \text{ where } V = \{k | \sum_{i=1}^k r_i^{res} \leq G \ \& \ \sum_{i=1}^{k+1} r_i^{res} > G\} \quad (1)$$

3 Utilization analysis

Let us first consider a scheduler supporting only single delay bound of B second. For simplicity, it is assumed that flows use all the same maximum packet size of M . Then flows have all the same critical rate r^{crt} written as $r^{crt} = GM/(GB - M)$ from the equation of Q_i^{RPS} . The reservation rate of each flow also becomes a random variable of $\max(R, r^{crt})$. Since V is recast as $\lfloor G/r^{res} \rfloor$, where $r^{res} \equiv E[\max(R, r^{crt})]$, $\rho = (r^{req}/G) \lfloor G/r^{res} \rfloor$.

Next, consider a scheduler in which L delay bounds of B_1, B_2, \dots , and B_L seconds are supported, where $B_1 > B_2 > \dots > B_L$. The set of flows requiring B_n is termed class n traffic and so the designated class number of each flow requiring B_n becomes n . The scheduler can be decomposed into L sub-schedulers as shown in Fig. 1, in which each Sub-Scheduler n (SS n) services only traffic belonging to the corresponding class, i.e., class n traffic. Every flow arrived at the scheduler goes to the sub-scheduler servicing its designated class traffic. Thus the bandwidth utilization can be computed as the weighted sum of the utilizations in L sub-schedulers by the amounts of scheduler bandwidth allocated to them. The utilization in each sub-scheduler can be obtained by the same way as that in the scheduler with single delay bound if both the amount of scheduler bandwidth allocated to the sub-scheduler and the desired rate distribution of flow being arrived at the sub-scheduler are known.

We first develop a framework to obtain the desired rate distribution of flow being arrived at each sub-scheduler, and then analyze the bandwidth utilization.

3.1 Desired rate distribution

Every flow being arrived at a scheduler can be characterized by two variables of desired rate and designated class number. We define the designated class

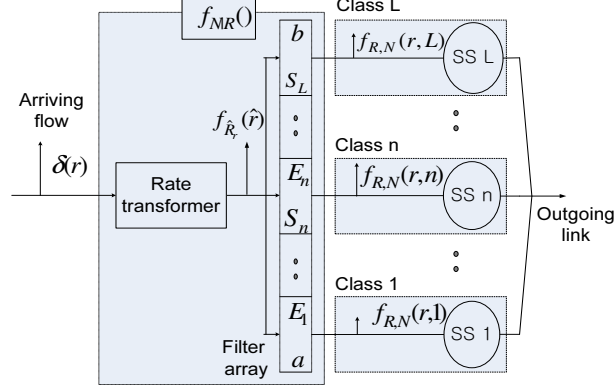


Fig. 1. Internal model of scheduler with L delay bounds

number of an arriving flow as a random variable on the sample space $S(L)$, where $S(L)$ is the set of all integers within $[1, L]$. Then, the arriving flow can be represented as a two-dimensional random vector whose components R and N have the joint-probability $f_{R,N}$. In general, f_R is known from the traffic load condition. f_N is, however, not easy to formalize because the designated class of each arriving flow may be relying on various factors such as its desired rate, the number of nodes along its end-to-end path, and the end-to-end delay bound of the service application to which it belongs. For simplicity, only the relation between designated class and desired rate is considered in this paper.

We first formulate a methodology to obtain $f_{N|R}$ meaning the dependency of designated class number on desired rate, in which a rate transformer and filter array shown in Fig. 1 are used. The desired rate of each arriving flow is transformed to an intermediate rate such that the designated class number of the flow becomes a function proportional to the intermediate rate. Then the designated class number n can be expressed as an increasing stepwise function of intermediate rate \hat{r} , i.e., $n = ku(\hat{r} - S_k)u(E_k - \hat{r})$, where $k \in S(L)$, $u(\hat{r})$ is a unit step function, $S_1 = a$ and $E_L = b$. Since the range of intermediate rate to be mapped to each class n is $[S_n, E_n)$, a class n filtering function $F^n(\hat{r})$ that extracts the class n traffic can be defined as $F^n(\hat{r}) \equiv u(\hat{r} - S_n)u(E_n - \hat{r})$. We define the intermediate rate of the desired rate r as a random variable \hat{R}_r on the same sample space as that of the random variable R . Then the flow with the desired rate of r can be represented as $f_{\hat{R}_r}(\hat{r})$ in the terms of the intermediate rate \hat{r} as shown in Fig. 1, where $f_{\hat{R}_r}(\hat{r})$ is the probability density function of the random variable \hat{R}_r . Note that Fig.1 illustrates an example of how the designated class of an arriving flow with the desired rate of r , represented by $\delta(r)$, is determined. Therefore $f_{N|R}$ can be obtained as $f_{N|R}(n|r) = \int_a^b F^n(\hat{r})f_{\hat{R}_r}(\hat{r})d\hat{r}$.

Finding $f_{\hat{R}_r}$ is beyond this paper. Instead we introduce an intermediate rate with the following $f_{\hat{R}_r}(\hat{r})$, termed ϕ intermediate rate, to analyze schedulers

with L delay bounds: $f_{\hat{R}_r}(\hat{r}) = \frac{1}{\phi(b-a)}$ for $r - \phi(r - a) \leq \hat{r} \leq r + \phi(b - r)$ and $f_{\hat{R}_r}(\hat{r}) = 0$ otherwise. ϕ is an independency factor indicating the degree that designated class gets free of desired rate: If $\phi = 0$, $f_{\hat{R}_r}(\hat{r})$ is equal to $\delta(r)$ which means that the designated class number of an arriving flow with the desired rate of r becomes a deterministic one proportional to the desired rate r itself. As ϕ increases, the number becomes a random one distributed within more various class numbers because the intermediate rate of the desired rate is more widely distributed. If $\phi = 1$, then it becomes a random number distributed within because the intermediate rate is uniformly distributed within $[a, b]$. For each arriving flow with a desired rate, as its designated class number distributes more widely within $S(L)$, the desired rate of flow being arrived at each sub-scheduler also distributes more widely within $[a, b]$, and then the scheduler will suffer from higher excess reservation rate. Thus the amount of reservation loss in each sub-scheduler will increase with raising ϕ .

Under the ϕ intermediate rate $f_{N|R}(n|r)$ can be developed as $f_{N|R}(n|r) = \int_a^b F^n(\hat{r})d\hat{r} = \int_a^b u(\hat{r}-B_n)u(\hat{r}-E_n)f_{\hat{R}_r}d\hat{r}$. The desired rate distribution $f_{R|N}(r|n)$ of flow being arrived at each sub-scheduler n is finally obtained as follows: $f_{R|N}(r|n) = f_{N|R}(n|r)f_R(r) / \int_a^b f_{N|R}(n|r)f_R(r)dr$

3.2 Bandwidth Utilization

Before analyzing the utilization, let us consider how to distribute the scheduler bandwidth to L sub-schedulers. Two policies are possible: explicit allocation in which the amount of bandwidth allocated to each sub-scheduler is previously determined, and implicit allocation in which the amount is implicitly determined by the property of traffic load.

Let us investigate the expected desired and reservation rates, denoted by r_n^{req} and r_n^{res} , respectively, of flow being arrived at each sub-scheduler n . Since r_n^{req} is the conditional expectation of the random variable R given $N = n$, $r_n^{req} = E[R|N = n]$. Because of the same maximum packet size of M , flows within each class n also have all the same critical rate r_n^{crit} written as $r_n^{crit} = GM/(GB_n - M)$ from the equation of Q_i^{RPS} . Then the reservation rate also becomes a random variable of $max(R, r_n^{crit})$. Thus $r_n^{res} = E[max(R, r_n^{crit})|N = n]$.

Now, let us obtain the bandwidth utilization ρ_E in a scheduler with the explicit allocation. Let G_n^E denote the amount of scheduler bandwidth allocated to each sub-scheduler n . Then the sub-scheduler can use the bandwidth of G_n^E regardless of other sub-schedulers and so it can be regarded as a scheduler with single delay bound. Thus from (1) ρ_E can be written as $\rho_E = \frac{1}{G} \sum_{n=1}^L G_n^E \rho_E^n$, where $\rho_E^n \equiv (r_n^{req}/G_n^E)[G_n^E/r_n^{res}]$.

Meanwhile, sub-schedulers under the implicit allocation share the scheduler bandwidth without any regulation until it will be exhausted. In other words, the scheduler bandwidth becomes a kind of common resource for them. In a consequence, the expected reservation rate r_I^{res} of an arriving flow at the scheduler becomes equal to the sum of the expected reservation rates of flows going to

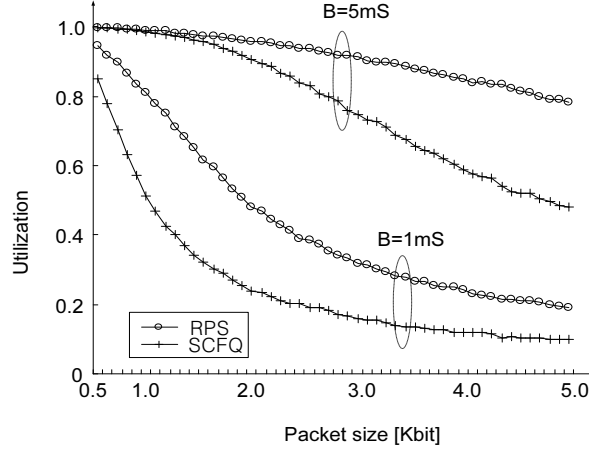


Fig. 2. Bandwidth utilization for single delay bound

L sub-schedulers. Since the expected reservation rate of flow going to each sub-scheduler n means the weighted expected reservation rate of flow being arrived at the sub-scheduler n by $f_N(n)$, $r_I^{res} = \sum_{k=1}^L r_k^{res} f_N(k)$. Thus the bandwidth utilization ρ_I under the implicit allocation can be written as $\rho_I = (r_{req}/G)[G/r_I^{res}]$.

4 Numerical evaluation

Uniformly distributed desired rate within $[2, 2048) Kbps$ and scheduler bandwidth of $10Gbps$ are considered. The bandwidth utilizations of RPS based and SCFQ schedulers for two typical delay bounds under single delay bound are compared in Fig.2. It shows that the utilization decreases with increasing the packet size. This is because longer packet size brings about higher critical rate resulting in larger reservation loss. We observe that the RPS based scheduler has better utilization by up to 50 % than that of the SCFQ one.

Next, RPS based scheduler with L delay bounds is evaluated under the following additional considerations: the delay bound class of each arriving flow is determined both the ϕ intermediate rate and the filter array such that the rate range of each filter n , i.e., $[S_n, E_n)$ is set as $[(r_{n-1}^{crt} + r_n^{crt})/2, (r_n^{crt} + r_{n+1}^{crt})/2] Kbps$, where $r_n^{crt} = 2 + 2048n/(L+1)$, $r_0^{crt} = 2a - r_1^{crt}$ and $r_{L+1}^{crt} = 2b - r_L^{crt}$. Each class has equal bandwidth share under the explicit allocation, i.e., $G_n^E = G/L$ for all $n \in S(L)$. The results are plotted in Fig. 3. We can observe that the utilization becomes better with increasing L or decreasing ϕ , which is due to reducing the reservation loss, and the explicit allocation policy yields some better utilization.

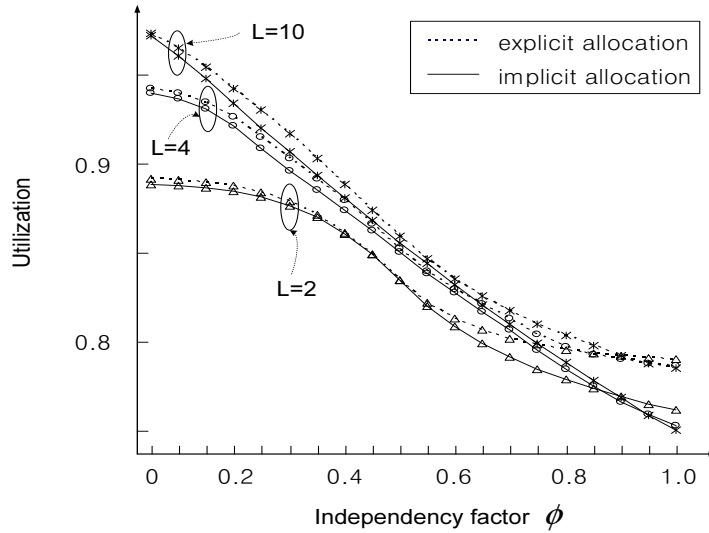


Fig. 3. Bandwidth utilization for L delay bounds

5 Conclusions

In this paper sorted-priority schedulers were analyzed in the terms of bandwidth utilization. A methodology to obtain the delay bound class of an arriving flow from the dependency of delay bound on desired rate was formulated and then used in evaluating the performance of scheduler with multiple delay bounds. The numerical evaluation showed that the bandwidth utilization is directly proportional to both the number of delay bound classes and the dependency of delay bound on rate but inversely proportional to packet size. In particular, schedulers with the latency property of WFQ had much better bandwidth utilization by up to 50 % than that in the SCFQ one.

References

- [1] Parekh, A.K.: A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks. PhD dissertation MIT (1992)
- [2] Demers, A., Keshav, S., Shenker, S.: Design and analysis of a fair queuing algorithm. Proc. ACM SIGCOMM (1989) 1-12
- [3] Golestani, S.J.: A Self-Clocked Fair Queuing Scheme for Broadband Applications. Proc. IEEE INFOCOM (1994) 636-646.
- [4] Stiliadis, D., Varma, A.: Rate Proportional Schedulers: A Design Methodology for Fair Queuing Algorithms. IEEE/ACM Trans. Net. **6** (1998) 164-174
- [5] Stiliadis, D., Varma, A.: Efficient Fair Queuing Algorithms for Packet-Switched Networks. IEEE/ACM Trans. Net. **6** (1998) 175-185
- [6] Kwak, D., Ko, N., Park, H.: Medium Starting Potential Fair Queuing for High-Speed Networks. IEICE Trans. Comm. **E87-B** (2004) 188-198