QoS support in a switch

Analysis of Policed Traffic Through a Switch with Shared Buffer Space

by

Yuval Shavitt

Dept. of Electrical Engineering – Systems
Tel Aviv University

Dagstuhl Oct/2002

Yuval Shavitt


**Background**

- ISPs sign SLAs with customers
- Typically an SLA contains:
  - Priority or QoS class
  - Committed rate: rate the ISP guarantee to the customer
  - Promise for ‘best-effort’ for excess rate.
- The ISP police and mark the incoming traffic at the network edge, e.g., using leaky bucket.
Background (2)

- Committed rates are used for routing, traffic engineering, etc.
- Excess traffic brings additional revenue
- To maximize profit, ISPs may use statistical multiplexing to overbook committed traffic.
- In case of congestion, committed traffic have precedence over excess traffic even with higher QoS class.
Implications

- Every switching element must be able to differentiate between QoS classes and the commit bit.
- We assume no PFQ, hardware must be kept economically operational. Thresholds are the way to go.
Input port model

- VOQs
- Queue per VOQ and per class.
- Classes share buffer space. protection vs. utilization
- Committed and excess traffic must be kept in the same Q (or OOO)

Typically 4-8 QoS classes.
1 or 2 has strict priority; rest are served with WRR.
Analysis Model

Priorities

- Priority queue 2 will simulate all the higher priority traffic in the system and will have strict priority over the other queues.
- Priority queue 1 has two traffic types, committed and excess.
- Priority queue 0 has two traffic types, committed and excess.
Thresholds

Thresholds limit the portion of buffer space traffic type can capture.

- Priority queue 2 rate decreases the service rate.
- Priority queue 1 committed and excess traffic have thresholds, $\alpha_{1C}$ and $\alpha_{1E}$.
- Priority queue 0 committed and excess traffic have thresholds, $\alpha_{0C}$ and $\alpha_{0E}$.
The analyzed system

Two queues that share a buffer space of $n$ packets (or cells).
The high priority queue has committed traffic and excess traffic with rates $\lambda_1$ and $\lambda_2$. The low priority queue has committed traffic with rate $\lambda_3$. Service rate is $\mu$.
Threshold is $n_{th} = \alpha_{th} n$, above threshold occupancy of the buffer excess high priority traffic is not accepted.
Strict priority.
The above system can be modeled by a continuous-time Markov chain with \((n + 1)(n + 2)/2\) states.

Each state is represented by the ordered pair \((t, s)\), where \(t\) is the number of high priority packets in the buffer and \(s\) the number of low priority packets.
The infinitesimal transition rates from state \((t, s)\) to state \((t', s')\), \(q_{t,s,t',s'}\) are:

\[
\begin{align*}
q_{t,s,t-1,s} &= \mu \\
q_{0,s,0,s-1} &= \mu \\
q_{t,s,t,s+1} &= \lambda_3 \\
q_{t,s,t+1,s} &= \begin{cases} 
\lambda_1 + \lambda_2 & \text{if } t + s \leq n_{th} \\
\lambda_1 & \text{if } t + s > n_{th}
\end{cases} \\
-q_{t,s,t,s} &= \begin{cases} 
\lambda_1 + \lambda_2 + \lambda_3 & \text{if } t + s = 0 \\
\lambda_1 + \lambda_2 + \lambda_3 + \mu & \text{if } 0 < t + s \leq n_{th} \\
\lambda_1 + \lambda_3 + \mu & \text{if } t + s > n_{th}
\end{cases}
\end{align*}
\] (1)

Note that \(-q_{t,s,t,s}\) is the transition rate out of state \((t, s)\)
We wish to find the steady state probabilities, $\pi_{t,s}$.

$\bar{\pi}Q = 0$

$\sum_{(t,s)} \pi_{t,s} = 1$

This numerical solution requires $O(n^{2(2+\alpha)}) \simeq O(n^5)$.

In the following, we shall describe methods to make the problem more tractable, by presenting a recursive solution that requires only $O(n^3)$ operations.
Calculating Drop probabilities

Let $\lambda_1 + \lambda_3 < \mu$ or else our solution is meaningless.

\[ \eta_1 = \sum_{i=0}^{n} \pi_{i,n-i} \]  \hspace{1cm} (2)

\[ \eta_2 = \sum_{i+j > n_{th}} \pi_{i,j} \]  \hspace{1cm} (3)

\[ \eta_3 = \sum_{i=0}^{n} \pi_{i,n-i} \]  \hspace{1cm} (4)

By definition $\eta_1 = \eta_3$ which shows that there is no preference between the two priority classes in the acceptance probability.
Average delay for the lower class:

Let $\bar{N}_i$ be the average number of cells of type $i$ in the system.

$$\bar{N}_3 = \sum_{i,j} \pi_{i,j} j$$

Using Little’s Law we know that the average delay, $T_3$, is given by

$$T_3 = \frac{\bar{N}_3}{(1 - \eta_3)\lambda_3} = \frac{\sum_{i,j} \pi_{i,j} j}{(1 - \sum_{i=0}^{n} \pi_{i,n-i})\lambda_3} \quad (5)$$
Reducing the Analysis Complexity Using Recurrence

Grand plan

Write the steady state probabilities of all the system states, \( \pi_{t,s} \), as functions of \( \pi_{0,s} \), \( 0 \leq s \leq n \).

Then, we can write \( n \) equilibrium equations and together with the probability conservation equation we obtain \( n + 1 \) linear equations that can be solved with complexity of \( O(n^3) \).
we can write the following $n(n+1)/2$ equilibrium equations

\[
-q_{t,s,t,s} \pi_{t,s} = q_{t,s-1,t,s} \pi_{t,s-1} \\
+q_{t+1,s,t,s} \pi_{t+1,s} + q_{t-1,s,t,s} \pi_{t-1,s} \\
-q_{t,0,t,0} \pi_{t,0} = q_{t+1,0,t,0} \pi_{t+1,0} + q_{t-1,0,t,0} \pi_{t-1,0} \\
1 \leq t \leq n - 1 \\
-q_{0,s,0,s} \pi_{0,s} = q_{0,s-1,0,s} \pi_{0,s-1} + q_{1,s,0,s} \pi_{1,s} \\
+q_{0,s+1,0,s} \pi_{0,s+1} \quad 1 \leq s \leq n - 1 \\
-q_{0,0,0,0} \pi_{0,0} = q_{1,0,0,0} \pi_{1,0} + q_{0,1,0,0} \pi_{0,1}
\]
\[
\begin{align*}
-(\lambda_1 + \lambda_2 + \lambda_3 + \mu) \pi_{t,s} &= \lambda_3 \pi_{t,s-1} \\
&+ \mu \pi_{t+1,s} + (\lambda_1 + \lambda_2) \pi_{t-1,s}, \quad t > 0, t + s \leq n_{th} \\
-(\lambda_1 + \lambda_3 + \mu) \pi_{t,s} &= \lambda_3 \pi_{t,s-1} \\
&+ \mu \pi_{t+1,s} + (\lambda_1 + \lambda_2) \pi_{t-1,s}, \quad t > 0, t + s = n_{th} + 1 \\
-(\lambda_1 + \lambda_3 + \mu) \pi_{t,s} &= \lambda_3 \pi_{t,s-1} \\
&+ \mu \pi_{t+1,s} + \lambda_1 \pi_{t-1,s}, \quad t > 0, t + s > n_{th} + 1 \\
-(\lambda_1 + \lambda_2 + \lambda_3 + \mu) \pi_{t,0} &= \mu \pi_{t+1,0} + (\lambda_1 + \lambda_2) \pi_{t-1,0}, \quad 1 < t \leq n_{th} \\
-(\lambda_1 + \lambda_3 + \mu) \pi_{n_{th}+1,0} &= \mu \pi_{n_{th}+2,0} + (\lambda_1 + \lambda_2) \pi_{n_{th},0} \\
-(\lambda_1 + \lambda_3 + \mu) \pi_{t,0} &= \mu \pi_{t+1,0} + \lambda_1 \pi_{t-1,0}, \quad n_{th} + 1 < t \leq n - 1 \\
-(\lambda_1 + \lambda_2 + \lambda_3 + \mu) \pi_{0,s} &= \lambda_3 \pi_{0,s-1} + \mu \pi_{1,s} + \mu \pi_{0,s+1}, \quad s \leq n_{th} \\
-(\lambda_1 + \lambda_3 + \mu) \pi_{0,s} &= \lambda_3 \pi_{0,s-1} + \mu \pi_{1,s} + \mu \pi_{0,s+1}, \quad s > n_{th} \\
-(\lambda_1 + \lambda_2 + \lambda_3) \pi_{0,0} &= \mu \pi_{1,0} + \mu \pi_{0,1}
\end{align*}
\]
Now, we can write the following recursion relations for $\pi_{t,s}$, $t > 0$:

$$\pi_{1,0} = \frac{(-q_{0,0,0,0}\pi_{0,0} - q_{0,1,0,0}\pi_{0,1})}{q_{1,0,0,0}}$$

$$\pi_{1,s} = \frac{(-q_{0,s,0,s}\pi_{0,s} - q_{0,s-1,0,s}\pi_{0,s-1} - q_{0,s+1,0,s}\pi_{0,s+1})}{q_{1,s,0,s}} \quad s = 1, 2, \ldots, n - 1$$

$$\pi_{t,0} = \frac{(-q_{t-1,0,t-1,0}\pi_{t-1,0} - q_{t-2,0,t-1,0}\pi_{t-2,0})}{q_{t,s,t-1,s}} \quad 2 \leq t \leq n$$

$$\pi_{t,s} = \frac{(-q_{t-1,s,t-1,s}\pi_{t-1,s} - q_{t-2,s,t-1,s}\pi_{t-2,s} - q_{t-1,s-1,t-1,s}\pi_{t-1,s-1})}{q_{t,s,t-1,s}} \quad t = 2, 3, \ldots, n \quad s = 1, 2, \ldots, n - t$$
The above recurrence suggests that all $\pi_{t,s}$ can be written as functions of $\pi_{0,s}$, i.e.,

$$\pi_{t,s} = \sum_{l=0}^{n} C_{t,s}(l)\pi_{0,l}, \quad (9)$$

It is easier to calculate the recurrence for the coefficients, $C_{t,s}(l)$, rather than directly for $\pi_{t,s}$. 
First, we calculate the coefficients of $\pi_{1,s}$ by

\[
C_{1,s}(s) = -q_{0,s,0,s}/q_{1,s,0,s} \quad s = 0, 1, 2, \ldots, n - 1 \quad (10)
\]
\[
C_{1,s}(s - 1) = -q_{0,s-1,0,s}/q_{1,s,0,s} \quad s = 1, 2, \ldots, n - 1
\]
\[
C_{1,s}(s + 1) = -q_{0,s+1,0,s}/q_{1,s,0,s} \quad s = 0, 1, 2, \ldots, n - 1
\]
\[
C_{1,s}(l) = 0 \quad |l - s| > 1
\]

Next, we calculate the coefficients of $\pi_{t,s}$ for $t = 2, 3, \ldots, n - 1$:

\[
C_{t,s}(m) = (q_{t-1,s,t-1,s}C_{t-1,s}(m) + q_{t-1,s-1,t-1,s}C_{t-1,s-1}(m) + q_{t-2,s,t-1,s}C_{t-2,s}(m))/q_{t,s,t-1,s}
\]

The recurrence calculation requires $O(n^3)$ operations.
The following $n + 1$ linear equation system, is made from $n$ (out of $n + 1$) unused equilibrium equations plus the probability conservation equation:

\[
-q_{t,n-t,t,n-t\pi t,n-t} = q_{t,n-(t+1),t,n-t\pi t,n-(t+1)} + q_{t-1,n-t,t,n-t\pi t-1,n-t} \\
1 \leq t \leq n - 1
\]  

\[
-q_{0,n,0,n\pi 0,n} = q_{0,n-1,0,n\pi 0,n-1} \\
\sum_{(t,s)} \pi_{t,s} = 1
\]

Solution complexity is lower than $O(n^3)$. 
Using the recurrence on the coefficients we can write Eq. 12 as

\[- \sum_{m=0}^{n} q_{t,n-t,t,n-t} C_{t,n-t}(m) \pi_{0,m} =\]

\[\sum_{m=0}^{n} \left( q_{t,n-(t+1),t,n-(t+1)} C_{t,n-(t+1)}(m) + q_{t-1,n-t,t,n-t} C_{t-1,n-t}(m) \right) \pi_{0,m}\]

\[1 \leq t \leq n - 1\]

and rewrite the probability conservation equation as

\[\sum_{t=0}^{n-1} \sum_{s=0}^{n-t} \sum_{m=0}^{n} C_{t,s}(m) \pi_{0,m} = 1\]
The analysis as a function of the buffer size, $n$.

$\lambda_1 = \lambda_2 = \lambda_3 = 0.4$ and $\alpha_{th} = 0.8$
Simulation Results

Line rate is 1Gbps
The buffer size is set to 100 packets
The packet arrival process is Poisson with combined rate 1.2
($\lambda_1 = \lambda_2 = \lambda_3 = 0.4$).
The packet length is Exponentially distributed, but bounded to be between 40 and 1500 bytes.
No thresholds

No admission control to the buffer
Threshold admission control to the buffer, where
\[ \lambda_1 = \lambda_2 = \lambda_3 = 0.4. \]
Threshold admission control to the buffer, where \( \lambda_1 = \lambda_2 = \lambda_3 = 0.34 \).
Low threshold for excess class 1 traffic ⇒ good protection and lower delay for committed class 0 traffic.

High threshold for excess class 1 traffic ⇒ good capability to pass bursts of excess traffic.

We want to have both!
Cross traffic threshold

We ALSO drop excess high priority traffic when the queue of class 0 committed traffic crosses a threshold.

This enables us to set the previous threshold quite high.

We can use the similar analysis to drive numerical results.
Threshold admission control to the buffer, where \( \lambda_1 = \lambda_2 = \lambda_3 = 0.4 \).
Cross-threshold admission control to the buffer, $\alpha_{1E0} = 0.4$, where $\lambda_1 = \lambda_2 = \lambda_3 = 0.4$. 
Conclusions

- Initial results show potential
- We need to study the system numerically and by simulations to get better understanding of the capabilities.
- Need to simulate bursty arrivals.