Traffic Analysis for a Multiplexer with PER-VC Queuing

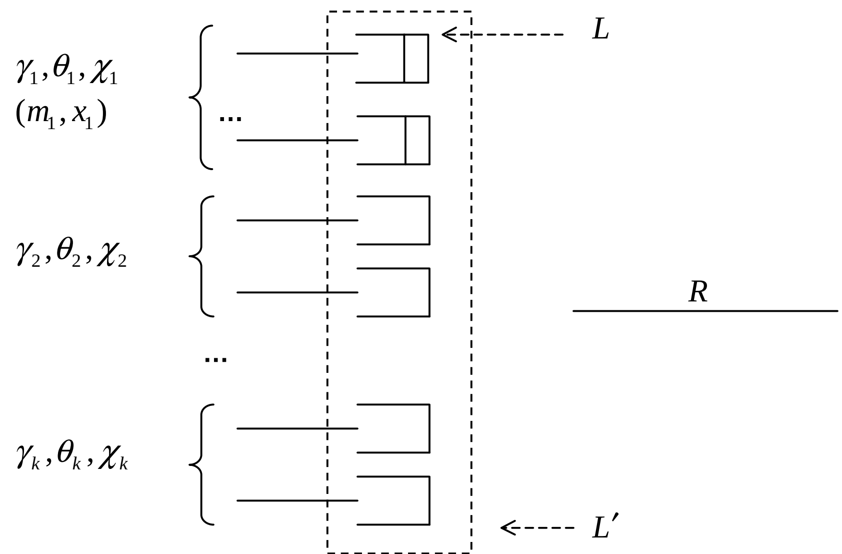
Nicolae Santean
Traffic Engineering Discipline, 0G60

The Mathematical Model

The purpose of this study is to engineer the traffic for UE9000. It answers the question: what is an optimal number of services that can be accommodated in a multiplexer with per-VC queuing. The result of this work is a mathematical model and an engineering tool used to analyze various configurations of the system.

The value added to previous work in this area is a more accurate model of the system obtained by capturing the “per-VC queuing” aspect and by considering a heterogeneous traffic mix. In addition the model can describe both the upstream and the downstream traffic flow and can be used to independently configure various segments of the system.

We consider the following model of a multiplexer with per-VC queuing:



Traffic services are offered to various users connected to the multiplexer via connection lines. Each connection can carry multiple virtual circuits (VC) via virtual paths (VP) – shown in the above figure. Each VC can be seen as a potential traffic stream. The *link* is a high-speed line carrying the multiplexed traffic and R is its bandwidth offer. The traffic per each VC is described by an on/off, fluid-flow model and is characterized by a traffic descriptor $T(\gamma, \theta, \chi)$ where γ is the peak rate of the connection, θ is the activity rate and χ is the mean burst size (bursts are transmitted at the peak rate). We group together all sources with the same traffic descriptor obtaining k groups. By m_i we denote the number of VCs with traffic descriptor T_i and by x_i the number of VCs of type T_i active (carrying a burst) at a given time. Then, $(m_1, ..., m_k)$ represents the configuration of the system and $(x_1, ..., x_k)$ represents the active subset of the configuration at a given time. All per-VC queues have the same size L and L' is the total buffer space limit. Note that L' can have a value significantly lower than the sum of capacities L of all queues; but at any given time the sum of fill-levels of all queues can not exceed L'. Note also that there are two cell-loss causes: when an arriving cell finds the assigned queue completely filled or when upon an arrival the cumulated fill-level of all queues is already equal to the total buffer space limit. Clearly, all queues share the same limited buffer memory. All sources have similar quality of service requirements given by an upper-limit for the probability of joining the queue β - QoS_β - and an upper-limit for the cell loss probability given that the queue is not empty η - QoS_η.

Engineering Aspects

The fact that the multiplexer does not offer a queue common to all traffic sources makes the use of Gibbens-Hunt formula difficult. Nevertheless, the equivalent bandwidth as computed by the EGH algorithm is an important traffic estimation since it captures the statistical multiplexing behavior in two areas. First, various traffic streams - of stochastic nature - share the same link (with a given bandwidth offer) and second, cells owned by different traffic sources are temporarily stored in the same buffer pool (with a given capacity offer).

We first analyze a simplified system by considering a queue of size L' common to all connections, ignoring therefore the queuing per-VC. Secondly, we analyze another simplified system by considering the queuing per-VC but ignoring the total buffer space limit this time. Finally, we make estimations about the initial system by combining the obtained results.

Let us first ignore the "per-VC queues" and consider a single common buffer of size L'. We apply the EGH method to this simplified system considering the same quality of service (QoS_β, QoS_η) and find the equivalent bandwidth for each type of source. Let E_i be the equivalent bandwidth corresponding to a source of type T_i,

$$E_i := EGH(\gamma_i, \theta_i, R, \chi_i, L', QoS_\beta, QoS_\eta) \ , \ 1 \le i \le k \quad .$$

Then the maximum number of sources of type T_i that can be accommodated in this simplified system has an upper-bound $\frac{R}{E_i}$. One can observe that the conditional probability of cell drop caused by reaching the total buffer space limit in the initial system is less than or equal to the conditional cell loss probability in this simplified system (considering the same number of VCs).

Therefore, by searching for an optimal configuration (a maximal number of VCs) for the initial system in the range given by these upper-bounds we ensure that the probability of cell loss caused by reaching the total buffer space limit is less than QoS_η. Let therefore assume that the solution $(m_1, ..., m_k)$ verifies the system of inequalities:

$$\begin{cases} 0 \le m_i \le \left\lfloor \dfrac{R}{E_i} \right\rfloor, \ 1 \le i \le k \\ \sum_{i=1}^{k} mi \cdot E_i \le R \end{cases}$$

and that $\eta_{1,t}$ is an estimation for the cell loss rate (conditional cell loss probability) of a VC of type t in this configuration.

Let us analyze another reduced system derived from the initial one by considering an infinite buffer space this time. In this case the cell loss is due only to the per-VC queue limit. Let ξ_t be a random variable denoting the bandwidth offered to an active source of type T_t, $1 \le t \le k$. Assuming that all VCs are in session and the system is in a steady state we can express

$$\Pr\{\xi_t > u\} = \sum_{\substack{(x_1,\ x_2,\dots,x_k) \\ 0 \le xj \le m_t,\ 1 \le j \le k \\ \sum_{i=1}^{k} x_i \gamma_i < \frac{\gamma_t}{u} R}} \left(\prod_{r=1}^{k} C_{m_r}^{x_r} \theta_r^{x_r} (1-\theta_r)^{m_r - x_r} \right),\ 0 < u \le R \quad .$$

Then, the CDF (cumulative distribution function) of ξ_t is given by

$$F_{\xi_t}(u) = \Pr\{\xi_t \le u\} = 1 - \Pr\{\xi_t > u\},\ 0 < u \le R \quad ,$$

with a probability density function

$$f_{\xi_t}(u) = \Pr\{\xi_t = u\},\ 0 < u \le R \quad .$$

Let abw_{ξ_t} be the average bandwidth available to the considered VC. Then

$$abw_{\xi_t} = E_{\xi_t} = < L > \int_{0+}^{R} v f(v) dv \quad .$$

For practical purpose the integral can be approximated by a smaller value

$$abw_{\xi_t} \ge \gamma_t \cdot \Pr\{\xi_t \ge \gamma_t\} \ ,$$

where

$$\Pr\{\xi_t \ge \gamma_t\} = \sum_{\substack{(x_1,\ x_2,\dots,x_k) \\ 0 \le xj \le m_t,\ 1 \le j \le k \\ \sum_{i=1}^{k} x_i \gamma_i \le R}} \left(\prod_{r=1}^{k} C_{m_r}^{x_r} \theta_r^{x_r} (1-\theta_r)^{m_r - x_r} \right) \quad .$$

Let now consider the following notations for the upstream traffic:

$$b_t := 1 - \Pr\{\xi_t \ge \gamma_t\}, \quad \psi_t := \frac{abw_{\xi_t}}{\gamma_t}, \quad \rho_t := \frac{\theta_t}{\psi_t}, \quad r_t := \theta_t \gamma_t \quad .$$

One can observe that if r_t (the average rate of source) is strictly greater than abw_{ξ_t} then $\rho_t > 1$, the system becomes unstable and congestion appears (i.e. the condition of ergodicity is not met).

If $\rho_t \le 1$ then the conditional probability of cell loss is given by a derivation of the G-H formula:

$$\ln(\eta_{2,t}) = \frac{L}{\chi_t} \cdot \frac{1}{\psi_t} \left(\frac{\theta_t}{1-\theta_t} - \frac{\psi_t}{1-\psi_t} \right)$$

and the probability of joining the queue is given by

$$\beta_{2,t} = \frac{b_t}{1-\rho_t + \rho_t b_t} \quad .$$

One can also observe that – back to the first simplified system – an estimation for $\eta_{1,t}$ is given by

$$\ln(\eta_{1,t}) = \frac{L'\alpha_t}{\chi_t} \cdot \frac{1}{\psi_t} \left(\frac{\theta_t}{1-\theta_t} - \frac{\psi_t}{1-\psi_t} \right), \quad where \quad \alpha_t := \frac{\beta_{2,t}}{\sum_{j=1}^{k} m_j \beta_{2,j}} \quad .$$

Assuming that a solution $(m_1,...,m_k)$ verifies the system of inequalities defined in the first part, the following upper-estimations for the initial system hold:

$$\begin{cases} \eta \le \max_{1\le t\le k}(\eta_{1,t}) + \max_{1\le t\le k}(\eta_{2,t}) \\ \beta \le \max_{1\le t\le k}(\beta_{2,t}) \end{cases}$$

Concluding, a vector $(m_1,...,m_k)$ that verifies the following system of inequalities:

$$\begin{cases} 0 \le m_i \le \left\lfloor \dfrac{R}{E_i} \right\rfloor \ , \ 1 \le i \le k \\ \sum_{i=1}^{k} mi \cdot E_i \le R \\ \max_{1\le t\le k}(\eta_{1,t}) + \max_{1\le t\le k}(\eta_{2,t}) \le QoS_\eta \\ \max_{1\le t\le k}(\beta_{2,t}) \le QoS_\beta \end{cases}$$

is a valid configuration of the initial system. One can observe that the role of the first two inequalities is to limit the solution space to finite boundaries whereas the last two inequalities ensure the meeting of the QoS requirements. Furthermore, the average cell transit delay for this system can be estimated by:

$$\frac{\sum_{i=1}^{k} m_i \cdot \left(\frac{1}{r_i \cdot \beta_{2,i}} \cdot \frac{\rho_i'}{(1-\rho_i')^2} + cell\,size \cdot \frac{1}{abw_{\xi_i}} \right)}{\sum_{i=1}^{k} m_i} \quad , \quad where\ \rho_i' = \frac{r_i \cdot \beta_{2,i}}{abw_{\xi_i}} \quad .$$

The system utilization for a given configuration can be estimated by the following formula:

$$U_{(m_1,\ldots m_t)} = \frac{1}{1 + \frac{R}{\sum_{t=1}^{k} m_t \cdot abw_{\xi_t}}}$$

Experiment

The following chart plots η and β as estimated by the above formulas. A system with two types T_1 and T_2 of traffic has been considered. The number of VCs of type T_1 (m_1) varies on the *x-axis* and the number of VCs of type T_2 (m_2) remains fix. The rate of cell loss is plotted against the primary *y-axis*, whereas the rate of joining the queue corresponds to the secondary *y-axis*.