Chapter 8

Circuit traffic

The Erlang model was developed a century ago in order to dimension the first telephone networks. Today, it is still a reference in the field of telecommunications. The model and its extensions are instrumental in the performance analysis of so-called "loss" systems whose resources are reserved and incoming clients are rejected in case of congestion.

8.1. Erlang's model

Consider a link consisting of m circuits. Each telephone call requires one circuit. Calls arrive according to a Poisson process of intensity λ and have exponential durations with parameter μ (this assumption is in fact not essential due to the insensitivity property, see exercise 4 in section 8.9). When all circuits are occupied, incoming calls are blocked and lost. We are interested in the *blocking rate*, that is the probability that an incoming call is blocked. A crucial parameter is the traffic intensity, defined as the product of the call arrival rate and the mean call duration:

$$\alpha = \frac{\lambda}{\mu}. \tag{8.1}$$

This dimensionless quantity, generally expressed in *erlangs* (symbol E), corresponds to the mean number of calls in the absence of blocking, that is when $m = \infty$. The system is then equivalent to an $M/M/\infty$ queue whose steady state has a Poisson distribution with mean α (see section 6.4). The randomness of traffic is illustrated by figure 8.1 for an arrival rate of 2.5 calls per minute and a mean call duration of 4 minutes, corresponding to a traffic intensity of $\alpha = 10\,\text{E}$.

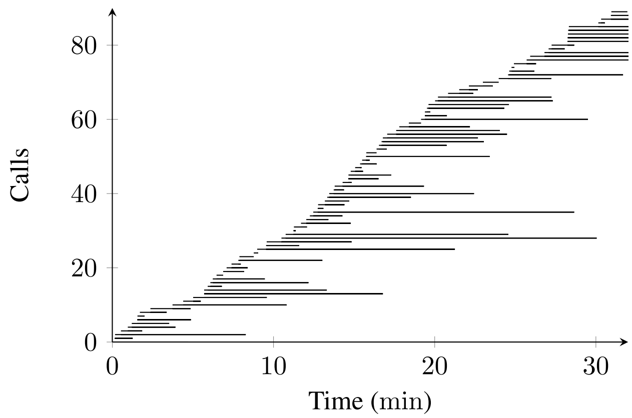

Figure 8.1: Randomness of telephone traffic.

The link load is the ratio of traffic intensity to link capacity:

$$\rho = \frac{\alpha}{m} = \frac{\lambda}{m\mu}.$$

8.2. Erlang's formula

The Erlang model is nothing more than an $M/M/m/m$ queue. In particular, the stationary distribution of the number of on-going calls is given by:

$$\forall x = 0, 1, \ldots, m, \quad \pi(x) = \frac{\frac{\alpha^x}{x!}}{1 + \alpha + \frac{\alpha^2}{2} + \ldots + \frac{\alpha^m}{m!}}. \tag{8.2}$$

This is a truncated Poisson distribution, as illustrated by figure 8.2. According to the PASTA property, each call sees the system in steady state at its arrival and thus is blocked with probability $\pi(m)$. We deduce the blocking rate:

$$B = \frac{\frac{\alpha^m}{m!}}{1 + \alpha + \frac{\alpha^2}{2} + \ldots + \frac{\alpha^m}{m!}}. \tag{8.3}$$

This is the Erlang formula, published in 1917.

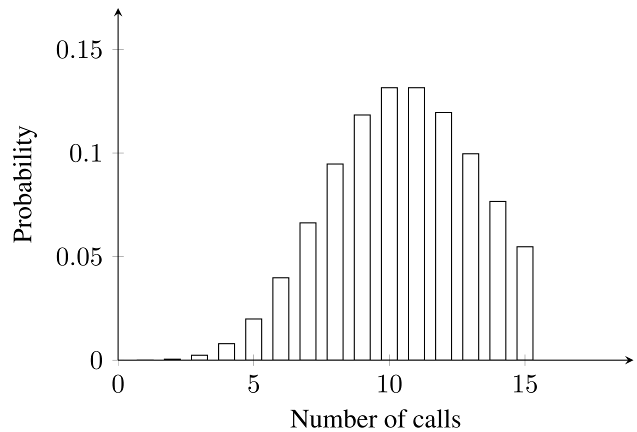

Figure 8.2: Distribution of the number of calls for $\alpha = 10\,\mathrm{E}$ and $m = 15$ circuits.

Figure 8.3 gives the blocking rate with respect to the link load $\rho = \alpha/m$ for different values of m. Note that, at constant load, the blocking rate decreases with capacity. These economies of scale can be explained by the lower traffic fluctuations (see exercise 1 in section 8.9) and can be proved using the integral form of the Erlang formula (see exercise 2). In the limit, we get the loss rate of a "fluid" model, without any traffic fluctuations: null if $\rho < 1$ and equal to the fraction of traffic in excess $(\rho - 1)/\rho$ otherwise; this corresponds to the case $m = \infty$ in figure 8.3.

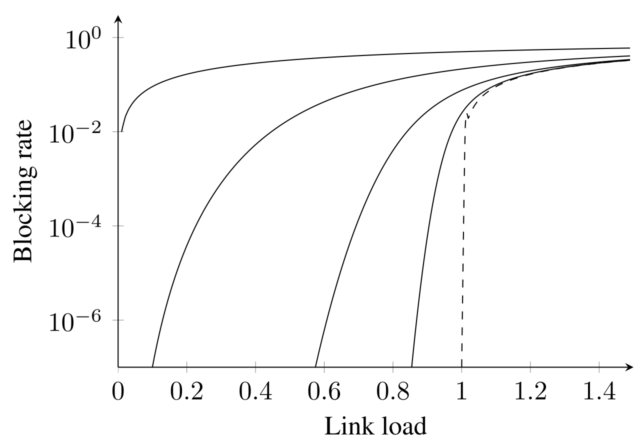

Figure 8.3: The Erlang formula ($m = 1, 10, 100, 1000, \infty$, from top to bottom).

The success of the Erlang formula is mainly due to its simplicity and robustness: it depends on the traffic characteristics through the traffic intensity α only. In particular, it holds for any distribution of call durations (see exercise 4). Erlang's model is said to be "insensitive". The only critical assumption is the Poisson process of call arrivals, which is satisfied when calls are generated by a large number of users (refer to section 3.7); for a small number of users, the Engset model described below applies.

Recursive formula

The numerical computation of Erlang's formula can be troublesome [1] for large values of m. In this case, it is preferable to compute the inverse of the blocking rate $I(m)$ for m circuits. Using (8.3), we get the following recursive formula:

$$I(m) = 1 + \frac{m}{\alpha} I(m-1), \quad \text{with} \quad I(0) = 1. \tag{8.4}$$

The limiting value I of the inverse of the blocking rate when the capacity m tends to infinity and the link load $\rho = \alpha/m$ is kept constant can be obtained by solving the corresponding equation: $I = 1 + I/\rho$. We get $I = \infty$ for $\rho < 1$, which corresponds to a null blocking rate, and $I = \rho/(\rho - 1)$ otherwise, which corresponds to the blocking rate $(\rho - 1)/\rho$.

8.3. Engset's formula

The Engset model is very similar to the Erlang model; the only difference is that calls are not generated according to a Poisson process but by the activity of a fixed number of users, denoted by K. Specifically, each user generates a continuous sequence of calls separated by random idle periods, as illustrated by figure 8.5.

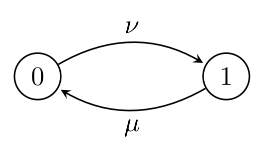

Figure 8.4: Transition graph of the Markov process describing the state of a user.

Call durations are exponential with parameter μ and idle periods are exponential with parameter ν. Denote by β the ratio ν/μ. In the absence of blocking, the traffic

1. For example, the number 1000! is of the same order as 10^{250}.