Dakota Performance Solutions

About Erlang C – Call Center Modeling

The dimensionless unit named the **Erlang** is used in telephony as a statistical measure of the volume of telecommunications traffic. It is named after the Danish telephone engineer A. K. Erlang, the originator of traffic engineering and queueing theory. Traffic of one Erlang refers to a single resource being in continuous use, or two channels being at fifty percent use, and so on, pro rata. For example, if an office had two telephone operators who are both busy all the time, that would represent two Erlangs of traffic.

Alternatively, an Erlang may be regarded as a "use multiplier" per unit time, so 100% use is 1 Erlang, 200% use is 2 Erlangs, and so on. For example, if total cell phone use in a given area per hour is 180 minutes, this represents 180/60 = 3 Erlangs. In general, if the mean arrival rate of new calls is λ per unit time and the mean call holding time is h, then the traffic in Erlangs A is:

$$A = \lambda h$$

This may be used to determine if a system is over-provisioned or under-provisioned (has too many or too few resources allocated). For example, the traffic measured over many busy hours might be used for a T1 or E1 circuit group to determine how many voice lines are likely to be used during the busiest hours. If no more than 12 out of 24 channels are likely to be used at any given time, the other 12 might be made available as data channels.

Traffic measured in Erlangs is used to calculate grade of service (GOS) or quality of service (QoS). There are a range of different Erlang formulae to calculate these, including Erlang B, Erlang C and the related Engset formula. These are discussed below, and may each be derived by means of a special case of continuous-time Markov processes known as a birth-death process.

The Erlang C formula also assumes an infinite population of sources, which jointly offer traffic of A Erlangs to N servers. However, if all the servers are busy when a request arrives from a source, the request is queued. An unlimited number of requests may be held in the queue in this way simultaneously. This formula calculates the probability of queueing offered traffic, assuming that blocked calls stay in the system until they can be handled. This formula is used to determine the number of agents or customer service representatives needed to staff a call center, for a specified desired probability of queueing.

$$P_W = \frac{\frac{A^N}{N!}\frac{N}{N-A}}{\sum_{i=0}^{N-1}\frac{A^i}{i!} + \frac{A^N}{N!}\frac{N}{N-A}}$$

where:

- A is the total traffic offered in units of Erlangs
- N is the number of servers
- P_W is the probability that a customer has to wait for service