4   Traffic Modelling for ASON/GMPLS networks dimensioning

Automatic Switched Optical Networks represent a very promising technology in offering flexibility for bandwidth allocation in future telecommunication networks. However, they open also new perspectives in providing high bandwidth connection services to a wide range of clients. For example, in this context, it can be highlighted that a great variety of transport services have to be carried by the same transport network. In particular, as we stated in Chapter 2, the way to provide connections ranges from the permanent (by means the NMS) to the switched (by means the CP).

Concerning the dynamic traffic (that related with the switched transport services), it could probably show various kinds of statistics for the Holding Time (HT) and the Inter-Arrival Time (IAT), depending on the clients.

Recent research activities have been related basically to ASON architectural issues. However, it has to be considered that all the new functionalities introduced by the ASON/GMPLS paradigm need a solid methodological background to be properly adopted.

The evolution towards intelligent optical networks is driven by the growing importance of traffic that requires on-demand switched connections for its efficient transportation.

Planning such a novel type of transport network hinges on investigating traffic models suited to represent (at the connection level) the dynamic component of the traffic. In the past some teletraffic models were developed that describe the user request patterns offered to a circuit-switched network, namely Plain Old Telephone Service (POTS)/Integrated Services Digital Network (ISDN). The models available in literature are mainly oriented to the classic telephone networks.

There are some analogies but also some important differences between ASONs and traditional circuit-switched networks. Specifically, an ASON allows the user to dynamically establish connections from one UNI to another in a way similar to that of a telephone conversation. The process to set up, hold and release connections is very similar. But, contrary to traditional switched networks, in ASON networks both permanent and switched connections have to coexist simultaneously, and while the permanent connections are provided as leased lines and established by the NMS, the switched connections require signalling between the customer premises and the network, and they require the support of an intelligent control plane.

This part of the Ph.D. Thesis is devoted to present the simulation case study we carried out about the applicability of the classical teletraffic models to dimension the switched part of the ASON networks. It requires the characterization of the switched connections demands statistical distribution (i.e., traffic arrivals/intensity process statistics). The traffic arrival process is generated by applying the triggering demands procedure described in Chapter 3. Specifically, for this case study, we considered the procedure based on the monitoring the average buffer occupancy (ABO).

Our case study consisted in characterising the traffic arrivals process through its two first statistical moments, namely the mean and the variance. Then, we evaluated by simulation the suitability of the classical teletraffic models for ASON switched part dimensioning purposes.

The organization of this Chapter is as follows: firstly we introduce the most important classical teletraffic models, namely the *Poisson*, the *Engset* and the *Fredericks* models, discussing their characteristics and their applicability to ASON networks environment and then we present in detail the simulation case study we carried out to evaluate their suitability to be used as analytical models to calculate the blocking probability for ASON network dimensioning.

4.1   Introductory notations

Arrival traffic process is characterized by the busy circuits distribution induced by it in an infinite pool of resources [81], [82]. In fact, the average traffic intensity, offered by a set of users generating call accommodation requests, is the average number of the observed busy resources, evaluated during a given period of time. Its measure unit is *Erlang*.

When a pool of resources is finite, there is the possibility of resource exhaustion; this occurrence is called congestion. There are some ways to measure congestion in a system. Two

important congestion indicators are the blocking congestion (the fraction of call attempts which observes the system busy) and the traffic congestion (the fraction of offered traffic that is not carried and is consequently lost).

The difference between the two indices lies in the fact that the former is based on users call attempts and the latter on the actual offered traffic.

To obtain the traffic arrival process distribution the so-called *moment-matching technique*, commonly used in teletraffic theory to study overflow streams in telephone networks, can be used [83], [84]. This technique consists of choosing an equivalent process that yields the same moments and whose distribution is easily obtained, and using that equivalent process to obtain the busy circuit distribution. In general, the first two statistical moments are considered, although it represents an approximation of the traffic arrival process [85].

The two moments considered are the average number (A) and the variance (V) of busy circuits. With the first two statistical moments, it can be calculated the peakedness factor (Z) of traffic arrivals, since it is defined as the ratio of the variance to the average number of busy circuits ($Z = V/A$) in an infinite-circuits system.

The peakedness factor is a measure of the deviation of the traffic from Poissonian nature. If Z is greater than 1 the traffic is said to be *peaked* while if it is lower than 1 it is said to be *smoothed*. While peaked traffic is characterized by irregular arrivals (i.e. arrivals in "groups"), *smoothed* traffic is characterised by regular arrivals [86].

Figure 42 presents the simulation results obtained to depict the traffic arrival process characteristics according to the peakedness factor Z. It shows that increasing Z the connection establishment requests arrives in groups.
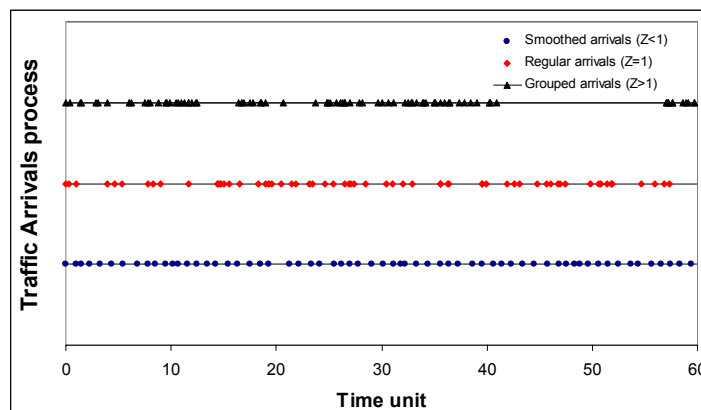

Figure 42: Traffic arrivals process according to the peakedness factor Z

4.2   Classical Teletraffic Models

In the past, some teletraffic models that describe user request patterns offered to a circuit-switched network (typically POTS/ISDN) were developed. The three most significant teletraffic models we considered are the *Poisson*, *Fredericks* and *Engset* models.

In this Subsection, we summarise these basic traffic models that could be used in the ASON context:

- *Poisson* model is characterized by exponentially distributed inter-arrival times of connection requests and exponentially distributed holding times. The variance of traffic intensity (V) is equal to the offered traffic (A = HT/IAT) and therefore the peakedness factor is $Z = 1$. Blocking probability experimented by a Poisson process on a set of n resources (circuits) is the well-known Erlang-B formula [86].

$$B^{Erlang}(n, A) = \frac{A^n/n!}{\sum_{k=0}^{n} A^k/k!}$$

  It is possible to demonstrate that the Erlang-B formula does not require exponentially distributed holding time but it is valid for arbitrary holding time distributions.

- *Fredericks* model gives an approximate formula for evaluating the congestion probability in a system of n circuits when the offered traffic is not poissonian and it is characterized by mean A and peakedness factor Z.

  The principle is to create equivalence between a system characterized by the tern of parameters (n, A, Z) and a poissonian equivalent system (n/Z, A/Z, 1).

  This assumption can be interpreted as an ideal model composed of a set of Z equivalent processes, each one fed by one of the Z simultaneous requests arriving in a group. Request groups follow a Poisson process. This model transformation is illustrated in Figure 43. Under this hypothesis the expression for the congestion can thus derived from the Erlang-B formula in the following way [87]:

$$B^{Fredericks}(n, A, Z) = B^{Erlang}(n/Z, A/Z)$$

  where n is the number of servers, A is the offered traffic in Erlang and Z is the peakedness factor. The Fredericks model was developed to capture the characteristics of the overflow telephone traffic, when an alternative routing strategy is used in the network. The offered

traffic (that is supposed to be Poisson) tries to get the first routing choice, and if rejected for congestion reasons, is offered to the next choice. The Fredericks approach allows modelling the carried as well as the overflow traffic by its mean and variance. The overflow traffic is shown to be characterized by a Z factor greater than 1, while the opposite (Z <1) for the carried traffic.
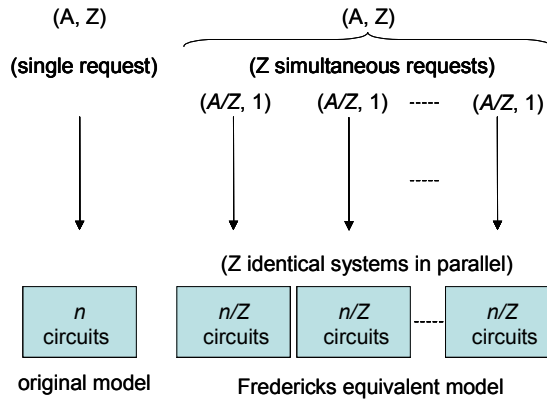

Figure 43: Fredericks model

In general Fredericks model require the evaluation of the Erlang-B formula with a non-integer number of channels and thus it needs for numeric computations an extended version of the Erlang-B formula for a continuous number of channels.

- *Engset* model assumes a finite number of sources S, each requiring one connection at a time. A source may be idle or active (when it has a connection in place): the sojourn time in both states is exponentially distributed. When a source requires a connection, it switches from idle to active state if it finds a free resource otherwise it falls in the idle state again. The blocking congestion for the Engset model is given by the following formula [87]:

$$B^{Engset}(n,S) = \frac{\binom{S-1}{n}\beta^n}{\sum_{k=0}^{n}\binom{S-1}{k}\beta^k}$$

where $\beta = \alpha/(1-\alpha)$ is the so-called offered traffic per idle source and α is the offered traffic per source (the carried traffic if the sources were never blocked); $A = S\cdot\alpha$ is the total offered traffic. The peakedness factor is given by $Z = 1-\alpha$.

4.3   Applicability of the teletraffic models to the ASON networks

The models above summarized can be applied in an ASON context with the following assumptions:

- Each user can request just one optical channel at a time.
- $\alpha = A/S$ is the offered traffic per user and it indicates the ratio of the time in which the user is using an optical channel.
- n is the number of circuits required on a specific trunk.
- Z is the degree of simultaneity of user requests.

As a matter of fact, in an ASON network, switched connections are likely to present dynamics substantially different from those observed on a traditional telephone network. Due to their huge bandwidth (typical bandwidth for optical channels is 2.5 or 10 Gb/s), these connections are likely to be required by very demanding services like network-to-network interconnections, big business customers, etc.

The variability associated to these kinds of services in terms of IATs and HTs is far from being known and we cannot therefore test the conformity of the models with the real behaviour. The behaviour of people calling on a phone network usually fits very well with a model of IAT and HT exponentially distributed, while potential customers that require connections to an ASON could show a very heterogeneous behaviour in terms of IAT and HT. This is because traffic sources in ASON networks could be residential users that require huge bandwidth connections for broadband applications but more realistically, in a medium term perspective, they could be ISPs, enterprises, branch offices of banks or some other kind of business customers that need bandwidth on demand. In addition, the behaviour could be very different depending on the type of customer. As a consequence, the offered traffic could not be poissonian in many cases. Thus, studies on statistical behaviour of customers of ASON will be necessary when such networks and related customer applications are available.

Our intention here was to make a survey of the available, analytically tractable models, and verify what kind of behaviour they are able to catch and taken into account for sizing the network resources. Then, by applying these models for a single link dimensioning, we investigate, by simulation, their applicability for ASON dimensioning purposes.

4.4 Suitability of classical teletraffic theory for ASON network dimensioning: simulation case study

As a comparison among the above-discussed models, in Figure 44, a graph depicting the number of circuits required using these models is reported, assuming a blocking target probability of 0.1% and a population of 100 users. All the curves are drawn as a function of the traffic specific for each user. The first two curves (starting from below) refer respectively to the Engset model and to the Erlang (Poisson) model while the rest refers to the Fredericks model calculated for various values of peakedness factor.
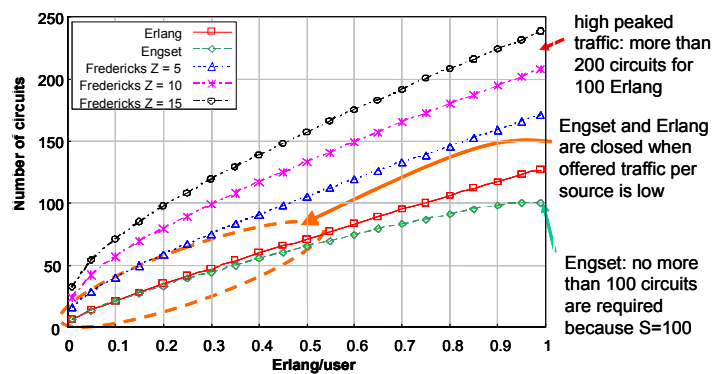

Figure 44: Number of circuits required for a population of 100 users and a blocking probability target of 0.1% as a function of the traffic intensity per user

It can be seen that the differences between Erlang and the Engset models are quite small. In particular, Engset model gives values always less than 100 because it is accounting for the finite number of users while the Erlang model starts to exhibit values greater than 100 for traffic of 76 Erlang or more, which corresponds to a 0.76 Erlang/user. For values lower than 0.2 Erlang/user the two models give the same results, while for higher values the Erlang model gives greater values than the Engset model in such a way that the dimensioning approach that uses the Erlang model is always conservative.

Regarding the Fredericks model, the case Z = 1 is equivalent to the Erlang case while for higher values of Z, when increasing the simultaneity of the users requests, the number of required circuits increases in order to maintain the assigned level of performance.

The intersection of these curves with the horizontal line (corresponding to a value of 100 circuits) is emphasized in the graph. Under these types of traffic, the points of intersection represent the break-even points between a full resource allocation and a statistical allocation planning strategy. For traffic values lower than the identified thresholds, some statistical gain can be

achieved, while for higher values the reduction of just one circuit with respect to the allocation of one circuit per user would degrade the performance to unacceptable values.

The presented models have the advantage to be simple general purposes models. Especially in the case of systems loaded by high number of sources the Fredericks model could be flexibly used to represent the traffic by only two parameters, the offered traffic and the peakedness factor. This is an appreciable advantage because no additional information on statistical distributions is required. The drawback concerns the accuracy of the model that cannot be very high as it is an approximated model that does not take into account distributions but only the first two moments of the traffic intensity.

Assuming the above summarized as reference models, we investigated whether such models can be effectively used as "*black box*" models with the aim to estimate the congestion probability in a simulated system reproducing connection requests to an ASON network. Such models could be taken into consideration as a tool for dimensioning ASON network resources and employed in network planning processes.

In order to obtain numerical results on the applicability of the three above-mentioned teletraffic models to ASONs, we have evaluated a simulation case study in a simple IP-over-ION-scenario. It consists of an IP router collecting the traffic from different IP networks (e.g., ISPs) on top of an optical switch (OXC) provided with a control plane.
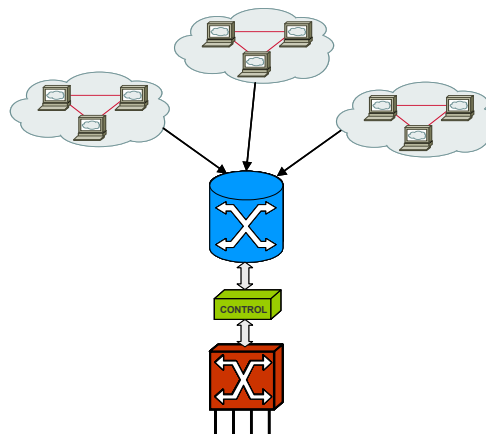

Figure 45: Simulated scenario, IP router on top of an OXC

The control function in Figure 45 implements the procedure based on the monitoring the average buffer occupancy (ABO) of the IP router interfaces (described in Chapter 3), which triggers the requests for switched connections on the basis of the ABO monitoring.

The aim of this simulation case study was to study the suitability of the classical teletraffic models presented in the previous Section to be used as analytical models for dimensioning the OXC outgoing trunk. The simulation conditions were the same of those described in Section 3.2.

Specifically, three steps composed our study:

Step 1: Running simulations for different values of the OW and triggering thresholds in order to characterize the switched channels demand statistics. Both the case of a single threshold and of different thresholds for set up and tear down triggering were considered. The number of available channels on the OXC trunks was assumed big enough to avoid any connection request blocking. As a result, we obtained the statistical distribution for IAT and HT, which allowed us to compute the total offered traffic intensity A as $\sum_i$ (HT_i/IAT_i) for any channel *i* that was switched on and off during the simulation time. Moreover, we obtained the peakedness factor Z as V/A, where V is the variance of the number of channels that were switched on and off during the simulation time (Table 9).

OW (s)	TH_{high} (KB)	TH_{low} (KB)	Traffic Intensity (A)	Peakedness Factor (Z)
60	4	4	2.483	0.645
60	50	50	2.653	0.721
60	50	4	1.894	0.792
60	450	50	2.234	0.680
10	4	4	3.49	0.735
10	50	4	2.31	0.837

Table 9: A and Z obtained for different configurations of the triggering mechanism

Step 2: Calculating the analytical blocking congestion probability applying the teletraffic models (Erlang, Fredericks and Engset) as a function of the offered traffic A and peakedness factor Z obtained by simulation in the first step, assuming a limited number of available switched channels.

Step 3: Comparison of the figures obtained in step 2 with the actual blocking probability obtained by simulation, in which the number of available channels is assigned in such a way that the blocking congestion assumed significant values. Specifically, the applicability of the Poisson, Fredericks and Engset models was then evaluated by comparing their blocking probability (calculated on the basis of the IAT and HT results obtained in the first set of simulations) to the blocking probability obtained trough the second set of simulations.

Figure 46 shows, for the various configurations simulated, that the Poisson model is not suited to characterize the optical switched connection demand generated by the IP traffic filtered by applying the designed procedure to trigger requests for switched connections. In fact, it experiences a significantly higher blocking probability compared to the results obtained through the simulation. In contrast, the Engset and Fredericks model provide only a slightly higher blocking probability than the simulated case, i.e., slightly overestimate the number of optical channels that have to be provisioned in order to guarantee the required Grade of Service (GoS). The Engset model is the one that best approximates the simulated traffic pattern, and could thus be chosen as candidate to dimension an ASON that has to cope with the traffic offered by an IP client network.
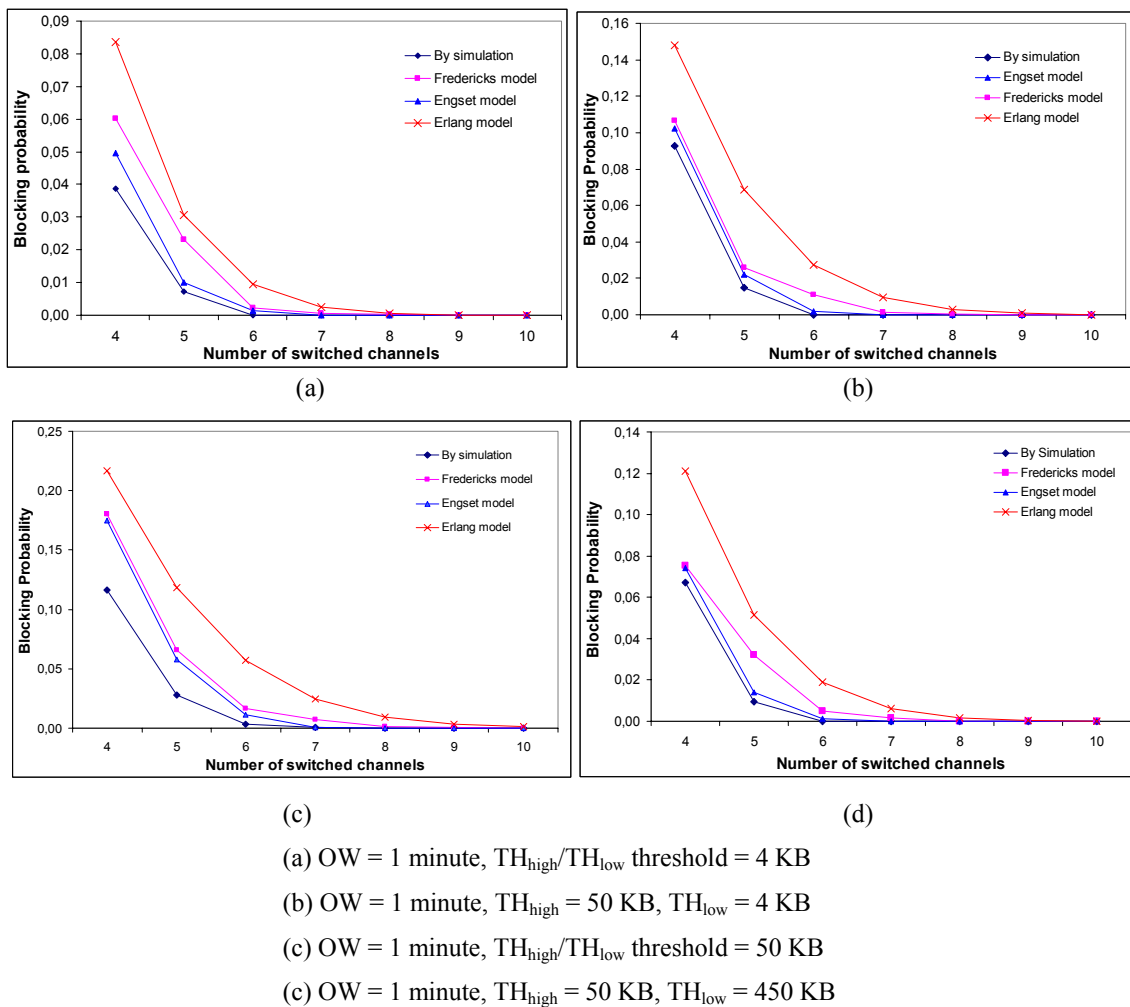

(a)　　　　　　　　　　　　　　　　　(b)

(c)　　　　　　　　　　　　　　　　　(d)

(a) OW = 1 minute, TH_{high}/TH_{low} threshold = 4 KB

(b) OW = 1 minute, TH_{high} = 50 KB, TH_{low} = 4 KB

(c) OW = 1 minute, TH_{high}/TH_{low} threshold = 50 KB

(c) OW = 1 minute, TH_{high} = 50 KB, TH_{low} = 450 KB

Figure 46: Blocking probability as a function of the number of switched channels available

However, it has to be underlined that this is a simple case study and that the simulation results are strongly influenced, on one hand, by the offered IP traffic and, on the other, by the parameters of the triggering mechanism (e.g., OW, TH_{high} and TH_{low} thresholds).

Moreover, it is worth mentioning that the suitability of classical teletraffic models in the analysed scenario is considered under particular conditions. In fact, the model identification has been done for a small range of model parameter values (offered traffic about 2 Erlang and peakedness between 0.65 and 0.84), while in telephone networks, where such models have been successfully applied since the beginning of the past century, it is of the order of tens or hundreds of Erlangs. Anyway, it has to be taken into account that probably just few channels will be requested automatically in an ASON environment.

Summarizing, as a minor contribution of this Thesis, we suggest an approximated method for modelling and dimensioning the switched part of ASON networks, which consists of:

1. Characterizing the connection demand statistics generated by applying the triggering request mechanism, by means of the average traffic intensity (A) and the peakedness factor (Z).

2. Applying the resulting values of A and *Z* to compute the well know Engset model to obtain the number of switched channels required to cope with a given GoS.