Real Time Traffic Management by Dynamic Bandwidth Allocation

Åke Arvidsson
Department of Telecommunications and Mathematics*
The University of Karlskrona/Ronneby
S-371 79 Karlskrona
Sweden

Abstract

We introduce a six layer model of traffic variations and a corresponding model of management actions. In particular, we traffic study real time traffic management on the traffic and session levels which correspond to variations on the hour and minute levels. Using virtual paths (VPs) and virtual channels (VCs) as the managed entities on the two levels respectively, we put forward a simple and robust strategy for traffic management in real time: For slow variations, *i.e.* on the traffic level, all nodes simultaneously monitor offered traffics and forward the result to a network management centre (NMC); The NMC computes a new set of optimal VPs compares the result to the existing network; If a change appears profitable, the necessary information is sent back to the nodes and the new design is implemented. For fast variations, *i.e.* on the session level, dynamic routing schemes similar to those of public switched telephone networks are employed.

Applying our strategy to control general networks subject to traffics which in advance are known only as expectations, we develop simple methods for traffic observation, derive the optimal observation interval, briefly review algorithms to compute VP networks, develop a cost optimal test to determine whether to implement a new design or not, and derive the optimal updating interval. Running our strategy in a simulator, we discover a considerable adaptivity and a performance far better than obtained by alternative methods.

1 Introduction

1.1 Traffic Management

Any telecommunications network is subject to traffic variations, some of which are related to human factors, such as office hours, and others are technology dependent, for example compression of video information. The way in which such variations are handled is, in our terms, the essence of traffic management. Typical management issues, *e.g.* statistical multiplexing and advanced routing, have gained considerable attention in the past. Current interest in traffic management is motivated by a number of factors: Services with widely different traffic characteristics and quality of service demands must be jointly managed in the integrated services digital network (ISDN) and in the broadband-ISDN (B-ISDN); Recent technological achievements in optical fibre technology mean that transmission is becoming faster, cheaper and more reliable; Switching technology has taken a new course with the introduction of the synchronous digital hierarchy (SDH) or its equivalent the slotted envelope network (SONET) and the asynchronous transfer mode (ATM); User needs are changing with the continuous deployment of new equipment and applications at the user's end.

Figure 1 shows a simple, hierarchical view of traffic variations in a B-ISDN using ATM. The **first stratum** refers to traffic growth from new users, equipment and applications; the **second stratum** refers to variations in the number of potential users, *i.e.* the number users in a position where they may use their equipment; the **third stratum** refers to variations in the number of active users, *i.e.* the number of potential users actually using their equipment; the **fourth stratum** refers to variations in user behaviour, *i.e.* the number of active users currently generating information; the **fifth stratum** refers to variations in user information delivery, *i.e.* the number of users currently presenting information to the network; and the **sixth stratum** refer to cell formation and presentation to the network. For more details, refer to *e.g.* [4].

Our view of traffic management, more thoroughly presented in [4] and illustrated in figure 2, follows the above model of variations. Similar approaches are proposed also by [14, 15, 28] and others. Management means for the **first stratum** include installing and activating new resources; the **second stratum** is managed by rearranging major parts of transmission network in terms of its virtual paths; the **third stratum** by rearranging minor parts of the transmission network in terms of virtual channels; the **fourth stratum** by

*The majority of this work was carried out while at the Department of Communication Systems, Lund Institute of Technology, Box 118, S-221 00 Lund, Sweden, and at the Teletraffic Research Centre, The University of Adelaide, Adelaide, South Australia 5005, Australia.

Layer	Variation factors	Time scale
1 Network	Traffic growth	Days
2 Traffic	Users preferences	Hours
3 Session	Call placements	Minutes
4 Activity	Applications employed	Seconds
5 Burst	Information generation	Milliseconds
6 Cell	Cell delivery	Microseconds

Figure 1: *Layered model of traffic variations.*

controlling the flow into and out of buffers, *i.e.* by means of end user backward congestion notifications and dynamic alternations of transmission capacity; the **fifth stratum** by potentially long buffers with dynamic buffer allocations and intelligent actions on congestion; and the **sixth stratum** by small buffers, short distance backward congestion notifications and forward reservations. Our main concern in the present paper is strata 2 and 3.

1.2 Virtual Paths and Virtual Networks

A VP is formed by reserving a certain amount of transmission capacity on a series of links and cross connecting the reserved channels through possible, intermediate transit nodes. A network of VPs, a virtual network (VN), forms a higher layer which is logically independent of the underlying physical network. The VPs thus constitute the links of a VN. The path concept in SDH/SONET and path addressing of ATM both lend themselves well to implementation of VNs.

The capacities of VN links are set so that current traffic demands can be handled with an acceptable grade of service under the constraints of the underlying physical network. If all demands cannot be accomodated, capacities are set to maximise some network performance metric, *e.g.* carried traffic revenues minus running costs. Since capacity allocations depend on available physical link capacities and currently offered traffics, changes in these will necessitate re-evaluations of the allocations. We refer to this process as VN management (VNM), other names include bandwidth switching [1], capacity management [8], bandwidth management [23], and bandwidth control [34, 35, 38].

Besides the general appeal of a structured approach to traffic management, some of the motives behind VPs and VNs [2, 5, 6, 7, 8, 9, 10, 11, 16, 27, 28, 33, 34, 35, 38] are: Reduced costs resulting from simplified transit exchanges, faster connection establishment by excluding intermediate node processing at set-up time, improved traffic management capabilities such as possibilities to redirect traffic in congested or faulty networks, and a means for providing customer-dedicated closed VNs. For multi-service networks, we may additionally gain simplified statistical multiplexing and grade of service control on strata 4–6 by grouping services into service classes (SCs) and carrying each class on a separate VN, a virtual service network (VSN) [4]. The members of an SC have similar strata 4–6 traffic characteristics, *e.g.* peak rate and burstiness, and similar strata 4–6

Layer	Management means	Time scale
1 Network	Expanding facilities	Days
2 Traffic	VP rearrangements	Hours
3 Session	VC rearrangements	Minutes
4 Activity	Flow operations	Seconds
5 Burst	Large buffers	Milliseconds
6 Cell	Small buffers	Microseconds

Figure 2: *Layered model of traffic management.*

quality of service demands, *e.g.* cell loss and delay.

2 Preliminaries

2.1 Networks and Traffics

We consider networks with N nodes on which VSNs with full end-to-end connectivity are established. For the sake of simplicity, we restrict ourselves to one SC. However, the results are readily extended to multi-service networks.

Strata 4–6 are visible from stratum 3 as a table of "equivalent circuits" (ECs). For a VP with origin o and destination d, $o, d = 1, \ldots, N$, $o \neq d$, the number of ECs is denoted by $m(o, d)$. This number is computed from the capacity of the VP, the buffer space for burst stratum statistical multiplexing, and service requirements such as loss, delay and jitter for the class in question, see works on equivalent circuits, *e.g.* [22, 26, 32], for further details. The capabilities of the activity level stratum may optionally be taken into account.

On strata 2 and 3, each origin-destination pair (OD-pair) $o, d = 1, \ldots, N$ is offered a pure chance stream of connection requests, the rate of which which changes every Tth time unit. We set the mean session holding time to unity and denote the traffic offered to o, d during the kth T-interval by $A_k(o, d)$. The sequence of traffics is limited to $k = 1, \ldots, K$ and cyclically repeated, so that $A_K(o, d)$ is followed by $A_1(o, d)$. Periods T and traffics $A_k(o, d)$ are known only as forecast expectations, *e.g.* averages from long term measurements, and will in reality deviate randomly from these values, in our paper according to normal distributions.

2.2 Traffic Management Strategy

As mentioned above, management functions on strata 4–6 interface stratum 3 by means of tables of ECs, and stratum 1 is not considered in this work. We will now give the strategies for strata 3 and 2.

2.2.1 Stratum 3: Sessions

The normalisation of time means that the time scale for stratum 3 is unity, and the scope is essentially stochastic variations in the number of arriving connection requests. This is done by establishing and releasing VCs.

Requests from o to d are granted a new VC on the direct o,d-VP from if the number of sessions in progress $l(o, d)$ is less than the VP's number of ECs, *i.e.* if $l(o, d) < m(o, d)$. Should this not be the case, is the request rejected (fixed session routing) or overflowed to alternative VP arrangements (dynamic session routing).

Dynamic session routing is accomplished by means of DAR, Dynamic Alternative Routing [19]. Secondary attempts in DAR are made over OD-pair specific tandem nodes $n(o, d)$, so that a VC may be established over the o,n- and n,d-VPs in tandem. Note that only two-hop alternatives are used. DAR is combined with trunk reservation, which means that a VP accepts overflowing requests only if $l(o, d)$ is less than the trunk reservation level $r(o, d) < m(o, d)$. A successful overflow thus requires that $l(o, n) < r(o, n)$ and $l(n, d) < r(n, d)$. Should the secondary choice fail is the request rejected, and a new $n(o, d)$ is selected for future overflows. Selection takes place at random with equal probabilities for all nodes.

Trunk reservation levels are set per VP to fulfill a simple profit maximisation criterion: A VP earns a profit of 1 for each primary choice VC established, and $1/2$ for each secondary choice VC. The traffic overflowing from a primary choice is assumed to be randomly distributed over all possible alternatives in proportion to their relative idle capacities, and the sum of all overflowing traffics to any VP is assumed to form a Poisson process.

2.2.2 Stratum 2: Traffic

The time scale of stratum 2 is by our definitions T, and the primary concern is to adjust resource allocations to systematic variations in the number of arriving connection requests. This is done by adjusting the capacities and physical layouts of VPs.

We require that all nodes simultaneously monitor offered traffics during intervals of t_{M} time units, with new intervals commencing every t_{U}th time unit. Traffic estimates are forwarded to the NMC which computes updated VSNs and analyses the result. If implementing the new design appears profitable, the necessary information is sent back to the nodes and the design is implemented. Transmitting traffic information to the NMC, computing and analysing the design, returning results to the nodes, and implementing the design is assumed to take a total of t_{E} time units.

Implementing new designs may result in that sessions in progress must be moved, or that VPs no longer can support all sessions in progress. The former is not considered further here, but for the latter we provide two-hop rerouting over the least loaded alternative, if any, and use premature clearing as a last resort if all alternative two-hop paths are blocked. Rerouting is further combined with limited repacking: At updating intervals are rerouted sessions moved back to direct routes if possible, but without optimisation. Rerouting and repacking may seem complicated, but we point out that this is already done in today's cellular mobile systems and high speed packet switching, [36]. Moreover, in section 4.1 we show that the actual number of sessions subject to such actions in reality is very small.

There is clearly a trade off between the resources spent on management actions such as rerouting and repacking active sessions, and the associated increase in carried traffic. To compare the two and optimise our strategy, we have assigned a profit of 1 for all accepted requests and accounted for the following costs: C_{L} for every connection attempt lost, C_{T} for each updating attempt (transmission of data to the NMC, computing and analysing a design), and for implementing a new design C_{I} for the implementation as such plus C_{P} for processing a session not routed on its direct path, C_{R} for processing a session that cannot be carried on its direct path, and C_{C} for dropping a session that cannot be carried at all.

2.3 Numerical Assumptions

Throughout this work, we have used a set of test networks and a simulator which implements the above strategy. The set of test networks comprises eight different physical networks of $N = 20$ nodes, each for which a series of $K = 8$ distinct sets of offered traffics was provided. For details on network topologies, transmission capacities and offered traffics, refer to other works by the author. Moreover, we have set the traffic interval length $T = 30$ and chosen the following costs: $C_{\mathrm{L}} = 1$, $C_{\mathrm{T}} = 10$, $C_{\mathrm{I}} = 100$, $C_{\mathrm{P}} = 0.5$, $C_{\mathrm{R}} = 0.5$, $C_{\mathrm{C}} = 10$. Numerical values obtained below are to some extent dependent on our particular choices.

3 Strategy Details

3.1 Traffic Observation

The first step in our strategy consists of nodes observing offered traffics and reporting results to the NMC. We therefore need a method according to which traffics are observed and a way to compute t_{M}, the time during which observations takes place.

3.1.1 Method

For our purpose, only short term, on-line measurements, possibly combined with short term forecasts, are relevant. We have considered two basic approaches, arrival counting (AC) and carried traffic measurements (CT).

We first consider AC, which means counting the number of connection attempts received during an observation interval of length t, $N(t)$, from which an estimate $\hat{A}$ of the offered traffic A is obtained as $\hat{A} = N(t)/t$. The analysis is straightforward and we find $E\{\hat{A}\} = A$ and $V\{\hat{A}\} = \frac{A}{t}$.

Proceeding to CT, the number of busy channels at time τ, $a'(\tau)$, is recorded during t time units, $\int_t a'(\tau)\,d\tau$, from which an estimate $\hat{A}'$ of the carried traffic A' is obtained as $\hat{A}' = \frac{1}{t}\int_t a'(\tau)\,d\tau$. This, in turn, gives an estimate of the offered traffic by "backward Erlang computation", *i.e.* by solving for $\hat{A}$ in $\hat{A}' = \hat{A}(1 - E_m(\hat{A}))$, where $E_m(A)$ is the Erlang loss formula. (For estimation of A' only, see also *e.g.* [13, 15].) Assuming a loss of 0, hence $\hat{A} = \hat{A}'$, and negative exponential session holding times, we find $E\{\hat{A}\} = A$ and $V\{\hat{A}\} = \frac{2A}{t}\left(1 - \frac{1-e^{-t}}{t}\right)$. Comparing AC to CT, it is noted that the latter provides better accuracy if $t \lesssim 1.5936$.

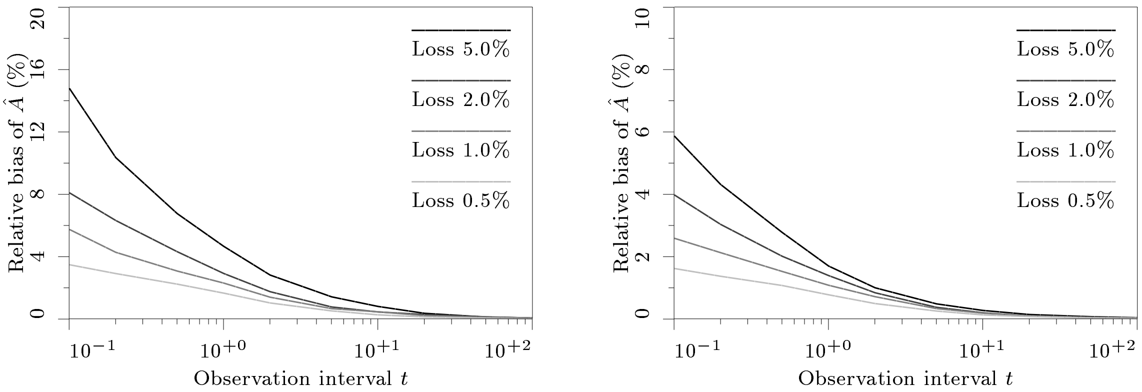

Figure 3: *Bias for the CT measurement technique for a small trunk (left) and a large trunk (right).*

For a loss > 0, however, CT fails as backward Erlang will introduce a positive bias to $\hat{A}$. In short, this is because of the non-linearity of the Erlang loss function. Figure 3 shows the bias recorded by applying CT measurements on a simulated trunk with $m = 100$ (left) and $m = 1000$ (right) ECs and $E_m(A) = 0.5\%$, 1.0%, 2.0% and 5.0% respectively.

In conclusion, CT cannot be used for real-time estimation of offered traffics, and AC is chosen as our estimation method. AC also has the advantage of ease of implementation: Counting the number of arriving service requests is straightforward, and is actually already implemented in many switching systems, *e.g.* as part of the overload control of the system itself.

3.1.2 Observation Interval

We now turn to the problem of selecting a proper observation interval $t = t_{\mathrm{M}}$ for AC. On one hand, the variance of the estimate becomes smaller the longer the observation interval. On the other hand, we consider traffic processes the rate of which varies with time, and the longer the interval, the more old, possibly invalid, information will be contained in an estimate. In other words, the optimal observation interval is a balance between low variance (large t), and exclusion of invalid information (small t).

Consider an arbitrary OD-pair o, d, an arbitrary traffic T-interval k, and let $\hat{A}$ be an estimate of $A_k(o,d)$ used to redesign the network during this interval. We define $t_{\mathrm{M}}^{\mathrm{opt}}(k,o,d)$, the optimal observation interval with respect to k, o, d, as the one for which the expected, squared error of $\hat{A}$ takes its minimum,

$$t_{\mathrm{M}}^{\mathrm{opt}}(k,o,d) = t_{\mathrm{M}} : \min_{t_{\mathrm{M}}} E\{(\hat{A}(t_{\mathrm{M}}) - A_k(o,d))^2\} \tag{1}$$

For convenience we temporarily drop the reference to o, d. Moreover, we assume that A_k confirms exactly with the forecast average, and ignore possible information on expected deviations in our analysis.

Due to transmission and processing time, designs implemented during the first t_{E} time units of the interval, will be based on measurements made during $k-1$. Similarly, designs completed during the following t_{M} time units will be based on measurements covering both $k-1$ and k. For the rest of the interval, however, all designs completed will be based on information relating to k. Assuming that designs are completed with the same probability throughout the interval and summing up the squared errors for the three cases gives

$$\begin{aligned} E\{(\hat{A} - A_k(o,d))^2\} \quad &= \\ &\left((A_{k-1} - A_k)^2 + \frac{A_{k-1}}{t_{\mathrm{M}}}\right)\frac{t_{\mathrm{E}}}{T} + \left(\frac{(A_{k-1} - A_k)^2}{3} + \frac{A_{k-1} + A_k}{2t_{\mathrm{M}}}\right)\frac{t_{\mathrm{M}}}{T} + \left(\frac{A_k}{t_{\mathrm{M}}}\right)\frac{T - t_{\mathrm{E}} - t_{\mathrm{M}}}{T} \end{aligned} \tag{2}$$

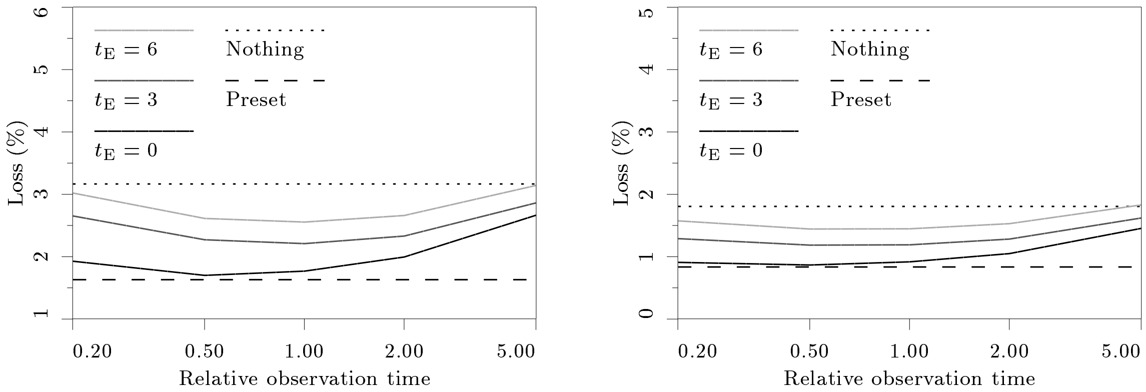

Figure 4: *Network loss versus observation interval with fixed session routing (left) and dynamic session routing (right).*

Differentiating (2) with respect to t_M and solving for a positive, real valued root yields

$$t_M^{\text{opt}} = \frac{\sqrt{3A_kT + 3(A_{k-1} - A_k)t_E}}{|A_{k-1} - A_k|} \tag{3}$$

If the difference between A_{k-1} and A_k is small, (3) may give $t_M^{\text{opt}} > T - t_E$, a result for which our model is not valid, but a sequence of more than two traffics must be considered. We refrain from this and simply set t_M^{opt} as the minimum of $T - t_E$ and (3).

Extending the result to networks, we compute an overall t_M^{opt} by weighting $t_M^{\text{opt}}(k, o, d)$ for each k, o, d by the expected differences respectively

$$t_M^{\text{opt}} = \sum_{k,o,d} \frac{\sqrt{E\{(\hat{A} - A_k(o,d))^2\}}}{\sum_{k,o,d}\sqrt{E\{(\hat{A} - A_k(o,d))^2\}}} t_M^{\text{opt}}(k, o, d) \tag{4}$$

where $E\{(\hat{A} - A_k(o,d))^2\}$ is computed from (2) with $t_M = t_M^{\text{opt}}$ according to (3).

Applying (4) to our set of test networks, we typically obtain $t_M^{\text{opt}} \approx 6$, 5, and 4 for $t_E = 0$, 3, and 6 respectively. Figure 4 shows network loss with fixed session routing (left) and dynamic session routing (right) versus t_M/t_M^{opt} as observed in our simulator, with different values of t_E indicated by solid lines. Dotted and dashed lines give upper and lower bounds respectively: The former refers to a fixed network which is permanently dimensioned for an average traffic over all k, and the latter to "ideal" management were exact information about all traffics is available in advance and the network is updated exactly when traffics change.

The figure basically confirms the optimum and we observe an overall insensitivity to t_M^{opt}: Setting t_M to $t_M^{\text{opt}}/2$ or $2t_M^{\text{opt}}$ has very little impact on the loss. This means that (4) provides a robust estimate in the sense that even poor forecasts of $A_k(o,d)$ will yield a t_M^{opt} that performs close to optimal.

There is also a tendency to reward t_M smaller than optimum, especially for small t_E. This is explained by exploitation of stratum 3 variations, the stochastic nature of which demand short observation intervals and short design and implementation times. These are visible here because of the unrealistically high frequency of network updates used to produce these curves, $t_U = 1$.

Finally we note the general increase in loss caused by longer execution times. This leads to the conclusion that, because of traffic dynamics, a fast algorithm may be preferred to a slower one even if the latter is known to produce in some meaning better network designs.

3.2 Network Design

We have used the design algorithm of [3], though a possibly better alternative now is available in [37]. Other more or less related proposals are given in Evans [12], Gopal *et al.* [20, 21], Herzberg [23, 24, 25], Gersht *et al.* [17, 18, 30] and Mase *et al.* [31]. These were not suitable for our purpose, however. For reasons of space, we refrain from elaborating on the details, but refer the interested reader to the appropriate references.

3.3 Network Updating

Upon receiving traffic estimates, the NMC computes the optimum VSN design, in our case by means of the algorithm in [3], after which it must decide whether to implement the result or not. Basing this decision on a comparison of costs, we need to evaluate the expected costs for implementing the new design and for keeping the old one respectively. We also need to determine the interval t_{U} at which the cycle of measurements and updating should be repeated.

3.3.1 Evaluation

Updating a network means that we can expect higher revenues as resources are reallocated according to current demands. However, updating also means costs, for example by processing VPs and VCs, and should hence be made only if increasing revenues exceed the costs. As costs are paid when an update is made while increasing revenues follow later, a decision on whether to implement a design or not must be based on predictions.

Consider a network with channel allocations $m(o,d)$, which are based on traffic estimates $A(o,d)$, and suppose that a new design $m^*(o,d)$, which is based on the most recent traffic estimates $A^*(o,d)$, is considered for implementation.

The costs C of implementing the design, consists of a fixed cost C_{I} for the implementation itself, and a variable cost depending on the state of the network. The latter is composed from costs for packing, rerouting and clearing, which amount to C_{P}, C_{R}, and C_{C} respectively per transaction. The number of transactions can be worked out exactly by performing an imaginary update. The total cost can thus be expressed as

$$C = C_{\mathrm{I}} + \sum_{\substack{\text{Packed}\\\text{sessions}}} C_{\mathrm{P}} + \sum_{\substack{\text{Rerouted}\\\text{sessions}}} C_{\mathrm{R}} + \sum_{\substack{\text{Cleared}\\\text{sessions}}} C_{\mathrm{C}} \tag{5}$$

The gain G from implementing the design is the additional traffic carried in the future if the design is implemented, but that otherwise would have been rejected. We assume that such a gain can only be achieved if the new design is based on real changes in arrival rates, but not if the difference $\Delta A(o,d)$ between $A(o,d)$ and $A^*(o,d)$ merely reflects random fluctuations. The expected gain is therefore

$$G = \sum_{o,d} P(o,d)\Delta A^{*\prime}(o,d)\tau \tag{6}$$

where $P(o,d)$ is the probability that $\Delta A(o,d)$ is not a result of stochastic variations, $\Delta A^{*\prime}(o,d)$ is the difference in carried traffic if $A^*(o,d)$ is offered to $m^*(o,d)$ rather than $m(o,d)$ ECs, and τ is the time during which present conditions are expected to last.

To compute $P(o,d)$, we assume that traffic estimates are normally distributed, with the mean and variance given in section 3.1.1. If $A^*(o,d)$ is an independent sample of the same traffic as $A(o,d)$, both have the same mean and variance, and we may define a normalised difference $Z(o,d)$

$$Z(o,d) = \frac{\Delta A(o,d) - E\{\Delta A(o,d)\}}{\sqrt{V\{\Delta A(o,d)\}}} = \frac{\Delta A(o,d)}{\sqrt{2\,\frac{A(o,d)}{t_{\mathrm{M}}}}} \in \mathrm{N}(0,1) \tag{7}$$

which allows us to approximate $P(o,d)$ as

$$P(o,d) \approx 2\Phi(|Z(o,d)|) - 1 \tag{8}$$

where Φ denotes the distribution function for the standard normal distribution, and we have used a two-sided test.

To find $\Delta A^{*\prime}(o,d)$ we formally write

$$\Delta A'(o,d) = A^{**\prime}(o,d) - A^{*\prime}(o,d) \tag{9}$$

where $A^{**\prime}(o,d)$ and $A^{*\prime}(o,d)$ refer to the carried traffic on the proposed and existing networks respectively. If fixed routing is employed stratum 3, we simply get

$$A^{**\prime}(o,d) = A^*(o,d)\left(1 - E_{m*(o,d)}(A^*(o,d))\right) \tag{10}$$

$$A^{*\prime}(o,d) = A^*(o,d)\left(1 - E_{m(o,d)}(A^*(o,d))\right) \tag{11}$$

where $E_m(A)$ again denotes the Erlang loss formula. If dynamic routing is used on stratum 3 we instead get

$$\begin{aligned} A^{*\prime}(o,d) &= A^*(o,d)\left(1 - E^A_{m(o,d),r(o,d)}(A^*(o,d),B^*(o,d))\right) + \\ &\quad + \frac{1}{2}B^*(o,d)\left(1 - E^B_{m(o,d),r(o,d)}(A^*(o,d),B^*(o,d))\right) \end{aligned} \tag{12}$$

$$\begin{aligned} A^{**\prime}(o,d) &= A^*(o,d)\left(1 - E^A_{m^*(o,d),r^*(o,d)}(A^*(o,d),B^{**}(o,d))\right) + \\ &\quad + \frac{1}{2}B^{**}(o,d)\left(1 - E^B_{m^*(o,d),r^*(o,d)}(A^*(o,d),B^{**}(o,d))\right) \end{aligned} \tag{13}$$

where the two terms of each equation refer to the contributions from primary and secondary choice traffics respectively. $E^A_{m,s}(A,B)$ and $E^B_{m,r}(A,B)$ are modified Erlang loss functions for traffic A and B respectively, when two traffics A and B are offered to a common group of m circuits and traffic B is blocked if $r < m$ circuits are busy

$$E^A_{m,r}(A,B) = \frac{A^m}{m!} \Big/ \left(\sum_{\mu=0}^{r} \frac{(A+B)^\mu}{\mu!} + \sum_{\mu=r+1}^{m} \frac{A^\mu}{\mu!}\right) \tag{14}$$

$$E^B_{m,r}(A,B) = \left(\frac{(A+B)^r}{r!} + \sum_{\mu=r+1}^{m} \frac{A^\mu}{\mu!}\right) \Big/ \left(\sum_{\mu=0}^{r} \frac{(A+B)^\mu}{\mu!} + \sum_{\mu=r+1}^{m} \frac{A^\mu}{\mu!}\right) \tag{15}$$

and $B^*(o,d)$ and $B^{**}(o,d)$ is the overflowing traffic offered to the o,d-link with the existing and proposed designs respectively

$$B^*(o,d) = \sum_{\substack{n=1 \\ n \neq o,d}}^{N} A^*(o,d)E_{m(o,d)}(A^*(o,d))\, I(o,d) \Big/ \sum_{\substack{n=1 \\ n \neq o,d}}^{N} I(o,d) \tag{16}$$

$$B^{**}(o,d) = \sum_{\substack{n=1 \\ n \neq o,d}}^{N} A^*(o,d)E_{m^*(o,d)}(A^*(o,d))\, I^*(o,d) \Big/ \sum_{\substack{n=1 \\ n \neq o,d}}^{N} I^*(o,d) \tag{17}$$

Here $I(o,d)$ and $I^*(o,d)$ denote the effective idle capacity of a two hop path with the existing and proposed designs respectively,

$$I(o,d) = \min(i(o,n), i(n,d)) \tag{18}$$

$$I^*(o,d) = \min(i^*(o,n), i^*(n,d)) \tag{19}$$

where $i(o,d)$ and $i^*(o,d)$ denote the effective idle capacity of a single link with the existing and proposed designs respectively

$$i(o,d) = m(o,d) - A^*(o,d)\left(1 - E_{m(o,d)}(A^*(o,d))\right) \tag{20}$$

$$i^*(o,d) = m^*(o,d) - A^*(o,d)\left(1 - E_{m^*(o,d)}(A^*(o,d))\right) \tag{21}$$

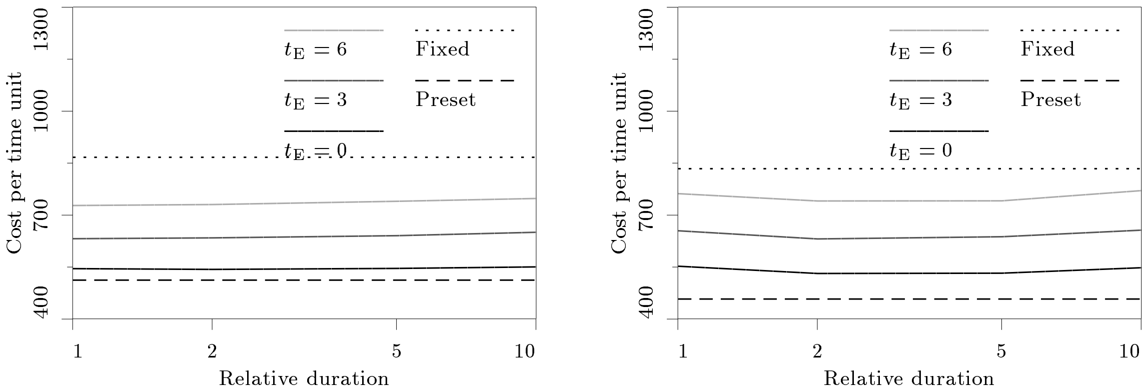

Figure 5: *Network cost versus assumed duration of an update with fixed session routing (left) and dynamic session routing (right).*

Clearly (12) – (21) should be cyclically repeated until the solution converges, *i.e.* until the Erlang fixed point is found. In light of all approximations above and the feeling that this is too time consuming for the intended real time purpose, we refrain from this for now.

Figure 5 shows network cost versus τ/t_{U} as observed in our simulator, with routing strategies, different values of t_{E}, and upper and lower bounds as in figure 4. Combining the results for fixed and dynamic VC routing, it appears that network cost takes its smallest for $\tau/t_{\mathrm{U}} \approx 2$, hence we choose this value.

3.3.2 Interval

Selecting a proper interval t_{U} between two attempts to update a network, is again a question of minimising costs: The more immediate the response to changing demands, the more resources must be spent on processing, but the less traffic will be lost. We define the optimum updating interval as the one for which the expected, total costs takes its minimum

$$t_{\mathrm{U}}^{\mathrm{opt}} = t_{\mathrm{U}} : \min_{t_{\mathrm{U}}} \sum_{k=1}^{K} C_k(t_{\mathrm{U}}) \tag{22}$$

where $C_k(t_{\mathrm{U}})$, the costs paid during period k with an updating interval of t_{U}, is a sum of management costs and costs for lost traffic.

As for management, results on traffic measurements are transmitted to the NMC which computes a new design and analyses it. This process is repeated at intervals of t_{U}, at a cost of C_{T} each.

Proceeding to losses, costs are associated to traffic lost due to the delay in updating the network. This traffic is denoted by $\Delta A'_k(o,d)$. For each T-interval k, the first traffic estimate $A^*(o,d)$ that contains information about the new traffic will, on the average, be completed after $t_{\mathrm{U}}/2$ time units, after which computation, evaluation and possible implementation will take another t_{E} time units. We call this time the initial updating delay. If a redesign is not made at this point, an additional updating delay, a multiple of updating intervals t_{U}, will be added until a redesign is implemented. We denote the average number of such attempts by $H_k(t_{\mathrm{U}})$.

Summing up, we can expect the following costs over a T-interval k of

$$C_k(t_{\mathrm{U}}) \quad = \quad \frac{T}{t_{\mathrm{U}}} C_{\mathrm{T}} + \left(\frac{t_{\mathrm{U}}}{2} + t_{\mathrm{E}}\right) \sum_{o,d} \Delta A'_k(o,d) + H_k(t_{\mathrm{U}})\, t_{\mathrm{U}} \sum_{o,d} \Delta A'_k(o,d) \tag{23}$$

where the first term refer to processing, the second one to traffic lost during the initial delay, and the third one to traffic lost during the additional delay.

$\Delta A'_k(o,d)$ is estimated from equation (9) with the new sample $A^*(o,d)$ equal to the offered traffic $A_k(o,d)$, $m^*(o,d)$ is taken from the design for forecasts of $A_k(o,d)$, and $m(o,d)$ from the one for $A_{k-1}(o,d)$.

To find $H_k(t_{\mathrm{U}})$, consider the probability $h_k(t_{\mathrm{U}})$ that the first traffic estimate containing information about a new traffic k will lead to a network redesign. According to section 3.3.1, a redesign will take place only if the cost C is less than the expected gain G, and we may write

$$h_k(t_{\mathrm{U}}) = \mathrm{Prob}(c_k < g_k) \tag{24}$$

where c_k and g_k is the cost and gain respectively, as predicted by the first sample containing k-traffic after a change.

Starting with the cost c_k, we simplify (5) by not going into details with the processing of individual sessions, which would require approximations of computed designs and of the states of the network at updating instants, and only consider the cost for implementing the design, hence $c_k \approx C_{\mathrm{I}}$.

Proceeding to the gain g_k, let $a_{k-1}(o,d)$ denote the estimates of $A_{k-1}(o,d)$ according to which the network is dimensioned prior to traffic period k, let $a^*_k(o,d)$ be the estimates which contain the first information on $A_k(o,d)$, and define $\delta a_k(o,d)$ as the difference between the two estimates. Similar to equation (6) we get

$$g_k \approx \sum_{o,d} p_k(o,d)\delta a'_k(o,d)\tau \tag{25}$$

where $p_k(o,d)$ approximates $P(o,d)$, the significance of the detected difference, and $\delta a'_k(o,d)$ approximates $\Delta A'(o,d)$, the expected increase in carried traffic following the implementation of a new design.

Proceeding analogous to section 3.3.1 we write

$$p_k(o,d) \approx 2\Phi(z_k(o,d)) - 1 \tag{26}$$

with

$$z_k(o,d) = \frac{\delta a_k(o,d)}{\sqrt{2\,\frac{a_{k-1}(o,d)}{t_{\mathrm{M}}}}} \tag{27}$$

The expected gain $\delta a'(o,d)$ used in the evaluation process is approximated as the relative difference detected times the nominal gain

$$\delta a'_k(o,d) \approx \frac{\delta a_k(o,d)}{\Delta A_k(o,d)}\Delta A'_k(o,d) \tag{28}$$

Inserting the approximations (26)–(28) in (25) we get

$$g_k \approx \sum_{o,d}\left(2\Phi\Big(\frac{\delta a_k(o,d)}{\sqrt{2\,\frac{a_{k-1}(o,d)}{t_{\mathrm{M}}}}}\Big) - 1\right)\frac{\delta a_k(o,d)}{\Delta A_k(o,d)}\Delta A'_k(o,d)\tau \tag{29}$$

from which it is observed that each term in g_k is a function of the sampled variables $\delta a_k(o,d)$ and $a_{k-1}(o,d)$. The approximate mean and the variance of g_k may therefore be obtained from from the moments of $\delta a_k(o,d)$ and $a_{k-1}(o,d)$ by standard formulae for approximate moments of functions of stochastic variables [29]

$$\begin{aligned} E\{g_k\} &= \sum_{o,d} E\{g_k(o,d)\} \qquad (30) \\ &= \sum_{o,d}\left(2\Phi\Big(\frac{E\{\delta a_k\}}{\sqrt{2\,\frac{E\{a_{k-1}\}}{t_{\mathrm{M}}}}}\Big) - 1\right)\frac{E\{\delta a_k\}}{\Delta A_k}\Delta A'_k\tau \end{aligned}$$

and

$$V\{g_k\} = \sum_{o,d} V\{g_k(o,d)\} \tag{31}$$

$$= \sum_{o,d} V\{\delta a_k\} \left[\frac{2\phi(\frac{E\{\delta a_k\}}{\sqrt{2\frac{E\{a_{k-1}\}}{t_\mathrm{M}}}})}{\sqrt{2\frac{E\{a_{k-1}\}}{t_\mathrm{M}}}} \frac{E\{\delta a_k\}}{\Delta A_k} \Delta A'_k + \left(2\Phi(\frac{E\{\delta a_k\}}{\sqrt{2\frac{E\{a_{k-1}\}}{t_\mathrm{M}}}}) - 1\right) \frac{\Delta A'_k}{\Delta A_k} \right]^2 \tau^2 +$$

$$\sum_{o,d} V\{a_{k-1}\} \left[2\phi(\frac{E\{\delta a_k\}}{\sqrt{2\frac{E\{a_{k-1}\}}{t_\mathrm{M}}}}) \frac{E\{\delta a_k\}}{\sqrt{2\frac{E\{a_{k-1}\}}{t_\mathrm{M}}}} \frac{E\{\delta a_k\}}{\Delta A_k} \Delta A'_k \right]^2 \tau^2$$

where we for readability have omitted references to (o,d) for δa_k, a_{k-1}, ΔA_k, and $\Delta A'_k$. The mean and the variance of the sample difference and the old sample respectively are obtained as

$$E\{\delta a_k(o,d)\} = \begin{cases} \frac{\Delta A_k(o,d)}{2} \frac{t_\mathrm{U}}{t_\mathrm{M}} & t_\mathrm{M} > t_\mathrm{U} \\ \frac{\Delta A_k(o,d)}{2} \left(1 - \frac{t_\mathrm{M}}{t_\mathrm{U}}\right) & t_\mathrm{M} < t_\mathrm{U} \end{cases} \tag{32}$$

$$V\{\delta a_k(o,d)\} = \begin{cases} \frac{2A_{k-1}(o,d)}{t_\mathrm{M}} + \frac{\Delta A_k(o,d)}{2} \frac{t_\mathrm{U}}{t_\mathrm{M}^2} + \frac{\Delta A_k(o,d)^2}{12} \frac{t_\mathrm{U}^2}{t_\mathrm{M}^2} & t_\mathrm{M} > t_\mathrm{U} \\ \frac{A_k(o,d)+A_{k-1}(o,d)}{t_\mathrm{M}} - \frac{\Delta A_k(o,d)}{2t_\mathrm{U}} - \frac{\Delta A_k(o,d)^2}{3} \frac{t_\mathrm{M}}{t_\mathrm{U}} + \frac{\Delta A_k(o,d)^2}{4} \frac{t_\mathrm{M}^2}{t_\mathrm{U}^2} & t_\mathrm{M} < t_\mathrm{U} \end{cases} \tag{33}$$

$$E\{a_{k-1}(o,d)\} = A_{k-1}(o,d) \tag{34}$$

$$V\{a_{k-1}(o,d)\} = A_{k-1}(o,d)/t_\mathrm{M} \tag{35}$$

Subtracting the cost from the gain and assuming that g_k, which is a sum of independent stochastic variables, is normally distributed, we finally obtain

$$h_k(t_\mathrm{U}) \approx \Phi(\frac{E\{g_k\} - c_k}{\sqrt{V\{g_k\}}}) \tag{36}$$

The following samples will contain successively more and more information about the new k-traffic, hence an update will have a greater chance of being implemented the second time and so on. However, we assume that the probability of implementing a new design independent of the number of previously failed attempts, and finally obtain

$$H(t_\mathrm{U}) = \sum_{k=0}^{\infty} k\,(1 - h(t_\mathrm{U}))^k h(t_\mathrm{U}) = \frac{1 - h(t_\mathrm{U})}{h(t_\mathrm{U})} \tag{37}$$

Applying the result to our series of networks typically gives $t_\mathrm{U}^\mathrm{opt} \approx 3.00$, 2.75, and 2.50 for $t_\mathrm{E} = 0$, 3, and 6 respectively and fixed session routing. With dynamic session routing, updating is less frequent and we get $t_\mathrm{U}^\mathrm{opt} \approx 3.25$, 3.00, and 2.75 respectively. Of the proposed updates, typically about 60% are found profitable for fixed VC routing and 20% for dynamic VC routing. We thus get effective updating intervals which are much longer than the frequency of updating attempts suggests, the difference is more accentuated for dynamic session routing.

Figure 6 shows network cost versus $t_\mathrm{U}/t_\mathrm{U}^\mathrm{opt}$ as observed in our simulator, with routing strategies, different values of t_E, and upper and lower bounds as before. Similar to measurements, the figure confirms the optimum and we also observe a similar insensitivity: Setting t_U to $t_\mathrm{U}^\mathrm{opt}/2$ or $2t_\mathrm{U}^\mathrm{opt}$ only marginally increases the cost.

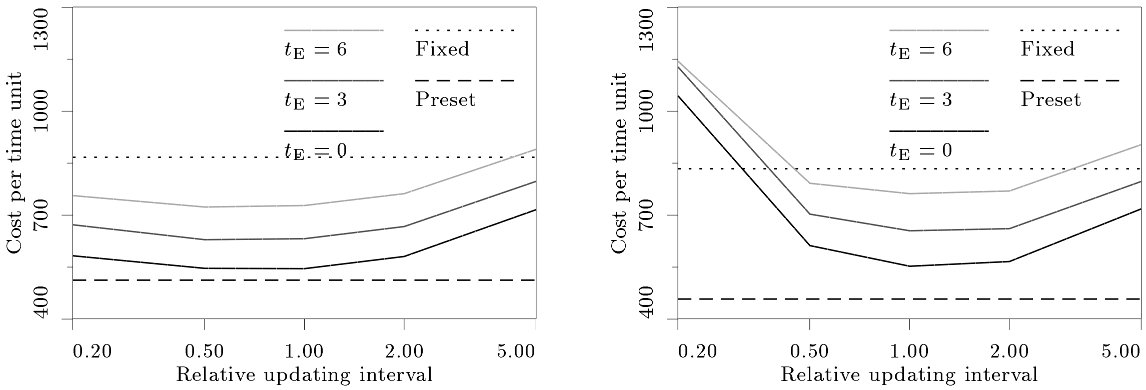

Figure 6: *Network cost versus updating interval with fixed session routing (left) and dynamic session routing (right).*

4 Results

4.1 Performance

So far, analyses and simulations have presumed that offered traffics fully agree with forecasts. In real networks, however, T are A_k are neither known, nor do they actually exist, but traffics change continuously *e.g.* throughout the day and from one day to another. To model this, we introduce slow changes by letting offered traffics change gradually, in 6 steps over 10 time units, rather than instantaneously. We also let $\tilde{T}_k$, the actual value of T-interval k, and $\tilde{A}_k(o,d)$, the actual traffic from o to d during period k, vary randomly around the expected (forecast) values according to normal distributions with a coefficient of variation K_V

$$\tilde{T}_k \in \mathrm{N}(T, K_V T) \quad (38)$$

$$\tilde{A}_k(o,d) \in \mathrm{N}(A_k(o,d), K_V A_k(o,d)) \quad (39)$$

with the modification that $\tilde{T}_k$ over a cycle $k = 1, \ldots, K$ are scaled to sum up to KT. This is done to keep the cycle length, typically a day or a week, fixed. The long term stability thus imposed is important to the alternatives we will discuss below, but not our strategy. Figure 7 gives an idea of resulting differences between a forecast traffic profile, our approximations, and actual realisation for $K_V = 0.05$ (left) and $K_V = 0.20$ (right).

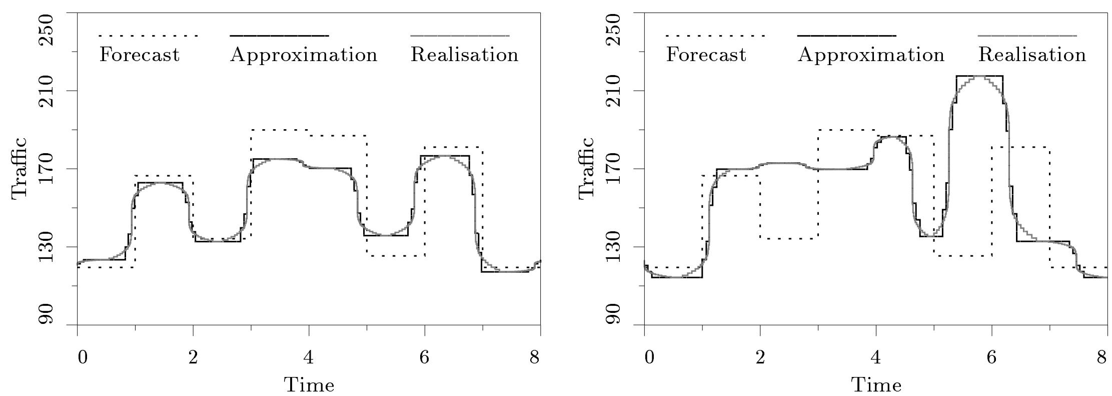

Figure 7: *Examples of possible realisations of traffic sequences, our approximations and associated forecasts with low uncertainty (left) and high uncertainty (right).*

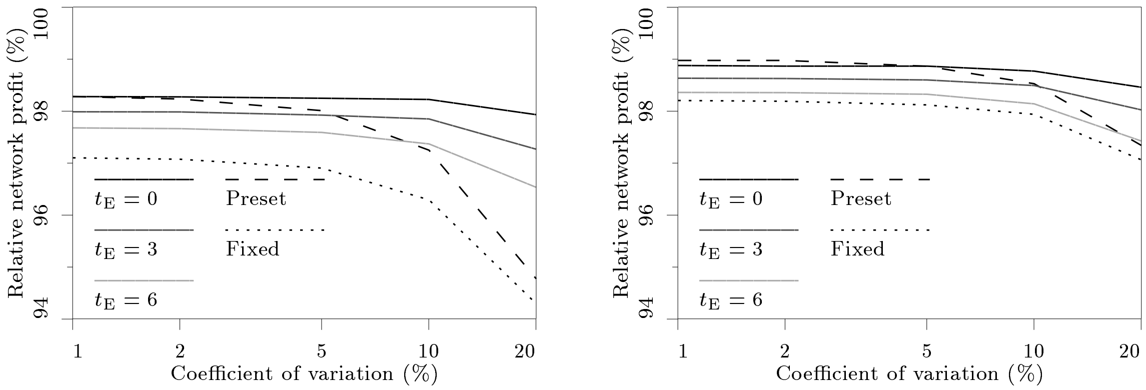

Figure 8: *Relative network profit versus forecast uncertainty with fixed session routing (left) and dynamic session routing (right).*

To get a compact representation of network performance, *i.e.* both carried traffic, lost traffic and processing costs, we define the performance metric relative network profit (r.n.p.) as carried traffic minus management costs over offered traffic. That is, an r.n.p. of 100% corresponds to a case where all offered traffic is accepted at no additional management costs. Figure 8 shows the r.n.p. observed when simulating the networks for different values of K_V, with t_M and t_U computed from the forecast traffic profiles. Routing strategies, different values of t_E, and bounds are indicated as above.

For fixed routing, it is noted that very small deviations, on the order of 1%, have little impact. However, for larger variations, our strategy adapts very well and significantly outperforms the "lower bound" of prescheduled rearranging. This observation holds for all values of t_E and irrespective of whether session routing is fixed or dynamic, but the degree of uncertainty required to motivate our strategy depends on both t_E and the session routing strategy. Small t_E and fixed session routing means lower degrees of uncertainty and *vice versa*.

To give an idea of the background to the curves, table 1 gives examples of the amount of processing behind the curves. The table provides the number of sessions carried, overflowed, redirected, repacked, and cleared for dynamic, preset and fixed VP networks, and fixed and dynamic VC routing. The numbers refer to an uncertainty K_V of 0.05, and $t_E = 0$ for dynamic VP routing.

It is seen that out of approximately 15 million sessions carried, no more than 30,000 (0.2%) are subject to redirection or repacking. We also observe that the dynamic strategy means less redirection and no clearing but more repacking than the preset strategy, and that the fixed strategy is the one that exploits alterative VC routing the most.

Finally we demonstrate the impact of our cost assumptions by showing the same plots as in figure 8. The upper and lower diagrams in figure 9 show r.n.p. when processing costs are divided by three and multiplied by

VP routing	VC routing	Carried	Overflowed	Redirected	Repacked	Cleared
Dynamic	Fixed	14232310	0	1665	99	0
Dynamic	Dynamic	14352452	225238	2753	31634	0
Preset	Fixed	14194258	0	6430	0	281
Preset	Dynamic	14365672	229665	7506	9846	618
Fixed	Fixed	14026139	0	0	0	0
Fixed	Dynamic	14252193	385185	0	0	0

Table 1: *Examples of actual amounts of processing.*

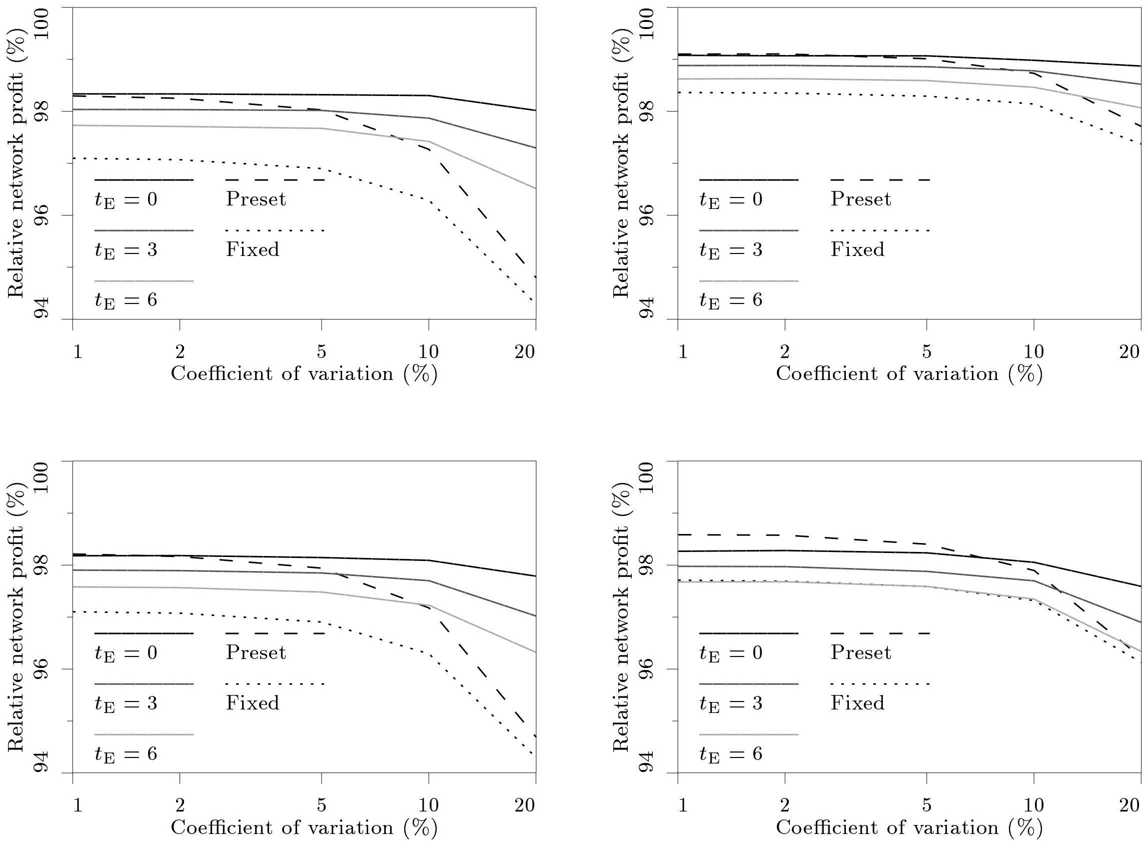

Figure 9: *Relative network profit versus forecast uncertainty with fixed session routing (left) and dynamic session routing (right). The upper row refers to lower processing costs, the lower row to higher processing costs.*

three respectively. The differences are quite small, but in particular for dynamic VC routing, it is noted that our strategy performs better the smaller the processing costs.

4.2 Alternatives

Having proposed a dynamic strategy for VNM and demonstrated its efficiency, one can ask what alternative solutions that can be found. We can see at least three: applying a deterministic strategy for rerouting VPs rather than our dynamic one; skipping routing of VPs altogether by relying on dynamic routing of sessions (VCs) only; or supplying so much extra transmission capacity that dynamic tuning of the resources is not meaningful.

The first alternative of predefined rearrangements is identical to the lower bound of preset VPs in figures 4–9. From figures 8 and 9, it is obvious that our strategy is superior except for very accurate forecasts, even when dynamic routing is applied on stratum 3.

The second alternative of no rearrangements is identical to the upper bound of fixed VPs in the same figures. From figures 8 and 9 it is seen that dynamic VP routing always outperforms fixed VP routing irrespective of whether VC routing is fixed or dynamic, *i.e.* dynamic routing of VCs is a compliment rather than an alternative to dynamic routing of VPs. Similar results have also been reported by [20, 21, 38, 35] and others. It is also seen that if one was restricted to keeping either VP or VC routing fixed, figures 8 and 9 indicate that there is more gain from applying the dynamics to VPs than to VCs if the VP design computation and implementation time t_E is not too large.

The third alternative of supplying more capacity rather than rearranging existing resources is rather costly as illustrated in figure 10. The curve shows the average number of OD-pairs serviced per unit of capacity for

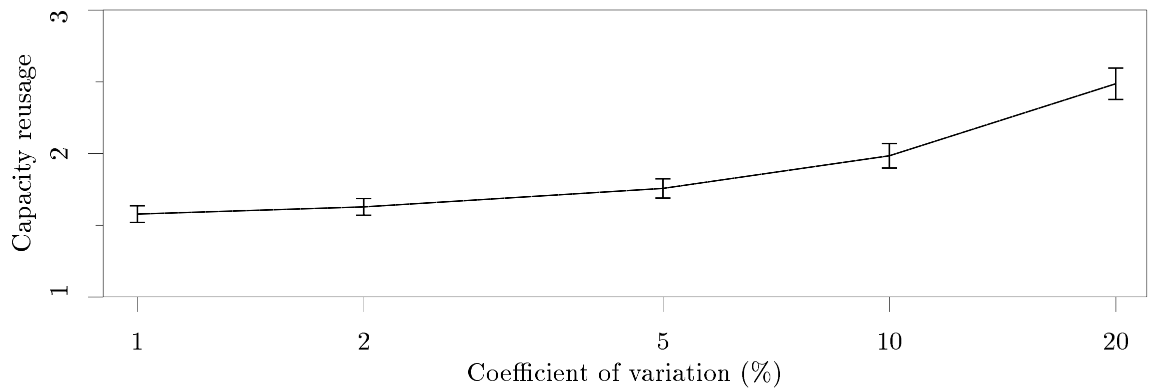

Figure 10: *Capacity reusage versus traffic profile coefficient of variation.*

various values of K_V. This number gives an idea of the additional capacity required to maintain service under traffic variations without VNM: about 50% for $K_V = 0.01$ up to 150% for $K_V = 0.2$. Moreover, VNM would still be required for situations like faulty links or nodes. The figure is obtained analytically for the current series of networks by averaging designs for 50 realisations of each traffic pattern and each network.

4.3 Bounds

Finally we turn to the question of how to identify a suitable time scale T for "slow" rate variations. We will show that T is not so much a matter of the general properties of traffic variations, but rather of technical and economical constraints.

As for traffic variation properties, it might be suspected that fast rate variations are averaged by a network and therefore not noticeable in terms of network performance. This is illustrated in figure 11, which shows network loss *vs.* the duration of a "stable state" T for networks which are permanently dimensioned for average traffics. Bounds are indicated by dotted lines: At lower bound at $T = 0$ means that traffics are equal to average at any time, and an upper bound at $T = \infty$ means that average applies only over an infinitely long period. The diagrams to the left and right refer to networks with fixed and dynamic VC routing respectively.

It is seen that the longer the time scale over which the average applies, the worse does a network permanently set to average traffics perform. In particular, we observe that not even rate variations which occur on a time scale faster than a session holding time are fast enough to be perceived as average by the network, a fact

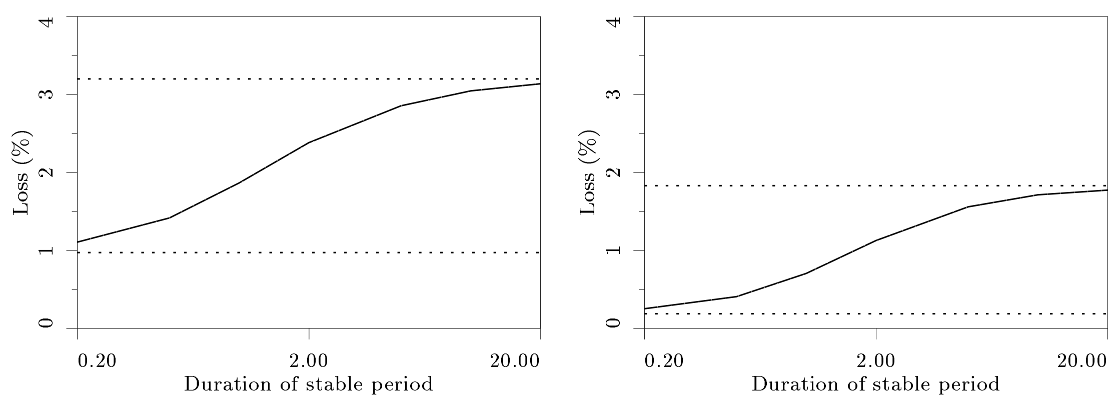

Figure 11: *Influence of rate variations to network performance.*

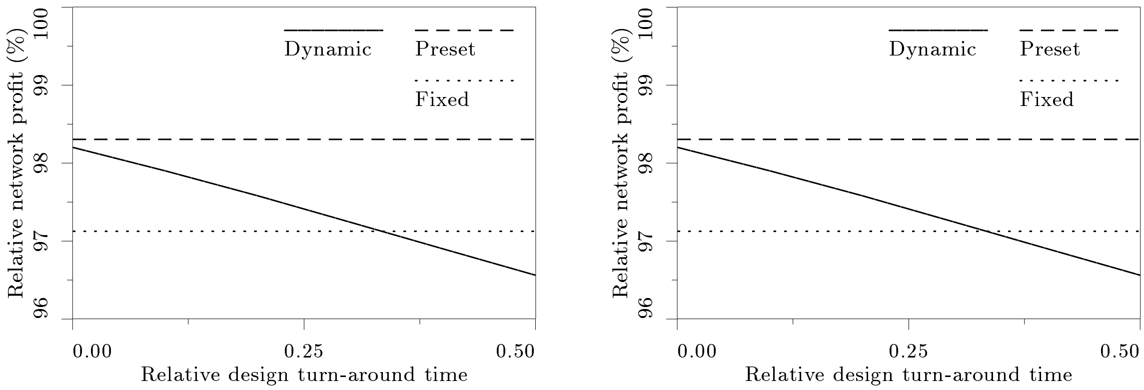

Figure 12: *Technical and economical constraints imposed by relative design turn-around time with fixed session routing (left) and dynamic session routing (right).*

which is not affected by employing advanced VC routing. Consequently, there is no distinct lower limit to the durability T of rate variations below which network performance is not affected, and T cannot be determined from general properties of variable traffics.

Figure 12 attempts to illustrate the technical and economical constraints. The diagrams show the impact of the relative design turn-around time t_{E}/T to network profit. As before, our strategy is indicated by solid lines, dotted and dashed lines refer to upper and lower bounds, and fixed VC routing is shown to the left, dynamic VC routing to the right. We note that the profit decreases almost linearly with t_{E}/T, from close to the upper bound for $t_{\mathrm{E}} = 0$ down to and below the lower bound marked by a fixed network. The points at which the these are crossed indicate absolute minima of T, below which our dynamic reallocation scheme falls short of permanent allocations based on traffics which are averages over k.

It is seen that T is bounded by t_{E}, and that the smaller the ratio t_{E}/T, the better will the strategy perform. Selecting T is therefore a compromise between good performance of the strategy, which is obtained for a long interval, and accurate control of the network, which requires a short interval. Further investigations are required to determine the optimal trade-off between the two aspects, but it is clear that the shorter the design turn-around time t_{E}, the smaller T can be used (or, conversely, the better the performance for any given T).

5 Further Work

The results indicate a number of possible improvements. Further work therefore includes better prediction methods optimised for slow changes over longer periods of time, and a more strict optimisation of the duration of a particular design. Another important issue is to develop suitable criteria for selecting an optimal T from a general traffic profile.

A more fundamental issue is to develop strategies that do not involve a central NMC but let the nodes communicate with each other and adjust the VPs as necessary.

Acknowledgement

The generous support of the Department of Communications Systems, Lund Institute of Technology, Sweden, and of the Teletraffic Research Centre, The University of Adelaide, S.A., Australia, which has made this work possible is gratefully acknowledged.

References

[1] Addie, R. & Warfield, R.: *Bandwidth Switching and New Network Architectures*, paper no. 2.3iiA.1 in Proc. 12th Int. Teletraffic Cong., Torino, Italy (1988).

[2] Addie, R., Burgin, J. & Sutherland, S.: *Information Transfer Protocols for the Broadband ISDN*, paper no. 22.6 in Proc. IEEE Globecom '88, Hollywood, Florida, U.S.A. (1988).

[3] Arvidsson, Å.: *High Level B-ISDN/ATM Traffic Management in Real Time*, paper no. 2.6 in Proc. 2nd Workshop on Perf. Modelling and Eval. of ATM Netw., Bradford, U.K. (1994).

[4] Arvidsson, Å.: *On the Usage of Virtual Paths, Virtual Channels, and Buffers in ATM Traffic Management*, paper no. 10.3 in Proc. 12th Nordic Teletraffic Sem., Helsinki, Finland (1995).

[5] Atkinson, D., Anido, G. & Bradlow, H: *Designing ATM Virtual Networks*, pp. 183–192 in Proc. 6th Australian Teletraffic Sem., Wollongong, New South Wales, Australia (1991).

[6] Boda, M., Faragó, A., Blaabjerg, S., Henk, T., Gordos, G. & Haraszti, Zs.: *Virtual Networking and Real-Time dimensioning — A Paradigm Shift in Network Management*, Telecom'95, Geneva, Switzerland (1995, to appear).

[7] Brungard, D., Grotjohann, H. & Kallenberg, P.: *Impact of New Transmission Technologies on the Network Architecture, Network Management and Economics*, paper no. A1.2 in Proc. XIIIth Int. Switching Symp., Stockholm, Sweden (1990).

[8] Burgin, J.: *Management of Capacity and Control in Broadband ISDN*, Int. J. of Dig. and Ana. Cabled Sys., vol. 2, pp. 155–165 (1989).

[9] Burgin, J.: *Broadband ISDN Resource Management*, Computer Netw. and ISDN Syst., vol. 20, no. 1–5, pp. 323-331 (1990).

[10] Burgin, J.: *Traffic Control and Resource Management for the Broadband ISDN*, pp. 93–100 in Proc. Australian Broadband Switching and Services Symposium '92, Melbourne, Victoria, Australia (1992).

[11] Dutkiewicz, E. & Anido, G: *Traffic Management and Control in Broadband Networks*, pp. 105–112 in Proc. 5th Australian Fast Packet Switching Workshop, Melbourne, Victoria, Australia (1990).

[12] Evans, S.: *A Mathematical Model and Related Problems of Optimal Management and Design in a Broadband Integrated Services Network*, J. Australian Math. Soc., ser. B, vol. 31, part 2, pp. 150–175 (1989).

[13] Filipiak, J. & Chemouil, P.: *Modeling and Prediction of Traffic Fluctuations in Telephone Networks*, IEEE Trans. on Commun., vol. 35, no. 9, pp. 931–941 (1987).

[14] Filipiak, J.: *M-Architecture: A Structural Model of Traffic Management and Control in Broadband ISDNs*, IEEE Comm. Mag., vol. 25, no. 5, pp. 25–31 (1989).

[15] Filipiak, J.: *Automatic Network Management Controls*, IEEE Trans. on Commun., vol. 39, no. 12, pp. 1776–1786 (1991).

[16] Gerla, M., Monteiro, J. & Pazos, R: *Topology Design and Bandwidth Allocation in ATM Nets*, IEEE J. Sel. Areas in Commun., vol. 7, no. 8, pp. 1253–1262 (1989).

[17] Gersht, A. & Shulman, A.: *Optimal Routing in Circuit Switched Networks*, IEEE Trans. on Commun., vol. 37, no. 11, pp. 1203–1211 (1989).

[18] Gersht, A. & Kheradpir, S.: *Integrated Traffic Management in SONET-Based Multi-Service Networks*, pp. 67–72 in Proc. 13th Int. Teletraffic Cong., Copenhagen, Denmark (1991).

[19] Gibbens, R., Kelly, F. & Key, P.: *Dynamic Alternative Routing — Modelling and Behaviour*, paper no. 3.4A.3 in Proc. 12th Int. Teletraffic Cong., Torino, Italy (1988).

[20] Gopal, G., Kim, C.-K. & Weinrib, A.: *Dynamic Network Configuration Management*, paper no. 302.2 in IEEE Int. Conf. on Commun., Atlanta, Georgia, U.S.A. (1991).

[21] Gopal, G., Kim, C.-K. & Weinrib, A.: *Algorithms for Reconfigurable Networks*, pp. 341–347 in Proc. 13th Int. Teletraffic Cong., Copenhagen, Denmark (1991).

[22] Guérin, R., Ahmadi, H. & Naghshineh, M.: *Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks*, IEEE J. Sel. Areas in Commun., vol. 9, no. 7, pp. 968–981 (1991).

[23] Herzberg, M.: *Network Bandwidth Management — A New Direction in Network Management*, pp. 218–225 in Proc. 6th Australian Teletraffic Sem., Wollongong, New South Wales, Australia (1991).

[24] Herzberg, M. & Byes, S.: *Bandwidth Management for Reconfigurable Multi-Service Networks*, Australian Teletraffic Research, vol. 25, no. 2, pp. 1–13 (1993).

[25] Herzberg, M. & Byes, S.: *Bandwidth Management in Reconfigurable Networks*, pp. 280–290 in Proc. 8th Australian Teletraffic Sem., Melbourne, Victoria, Australia (1993).

[26] Hong, D. & Suda, T.: *Congestion Control and Prevention in ATM Networks*, IEEE Netw. Mag., vol. 5, no. 4, pp. 10–16 (1991).

[27] Hughes, D., Bradlow, H. & Anido, G: *Congestion Control in an ATM Network*, pp. 835–840 in Proc. 13th Int. Teletraffic Cong., Copenhagen, Denmark (1991).

[28] Hughes, D. & Wajda, K: *Comparison of Virtual Bandwidth Assignment and Routing Methods*, Annales des Telecommunication, to appear.

[29] Kendall, M. & Stuart, A.: *The Advanced Theory of Statistics*, Charles Griffin & Co. Ltd., London (1962).

[30] Kheradpir, S., Gersht, A. & Stinson, W.: *Performance Management in SONET-Based Multi-Service Networks*, paper no. 39.3 in Proc. IEEE Globecom '91, Phoenix, Arizona, U.S.A. (1991).

[31] Mase, K. & Imase, M.: *An Adaptive Capacity Allocation Scheme in Telephone Networks*, IEEE Trans. on Commun., vol. 30, no. 2, pp. 354–359 (1982).

[32] Mase, K. & Shioda, S.: *Real-Time Network Management for ATM Networks*, pp. 133–140 in Proc. 13th Int. Teletraffic Cong., Copenhagen, Denmark (1991).

[33] Moondra, S.: *Impact of Emerging Switching-Transmission Cost Tradeoffs on Future Telecommunications Network Architectures*, IEEE J. Sel. Areas in Commun., vol. 7, no. 8, pp. 1207–1218 (1989).

[34] Ohta, S. & Sato, K.-I.: *Dynamic Bandwidth Control of the Virtual Path in an Asynchronous Transfer Mode Network*, IEEE Trans. on Commun., vol. 40, no. 7, pp. 1239–1247 (1992).

[35] Papanikos, I., Logothetis, M. & Kokkinakis, G.: *Virtual Path Bandwidth Control Versus Dynamic Routing Control*, paper no. 81 in Proc. 3rd Workshop on Perf. Modelling and Eval. of ATM Netw., Ilkley, U.K. (1995).

[36] Parris, C., Zhang, H. & Ferrari, D.: *A Dynamic Management Scheme for Real-Time Connections*, paper no. 6a.1 in Proc. IEEE Infocom '94, Toronto, Ontario, Canada (1994).

[37] Pióro, M. & Gajowniczek, P.: *Stochastic Allocation of Virtual Paths to ATM Links*, paper no. 2.1 in Proc. 2nd Workshop on Perf. Modelling and Eval. of ATM Netw., Bradford, U.K. (1994).

[38] Shioda, S.: *Evaluating the Performance of Virtual Bandwidth Control in ATM Networks*, IEICE Trans. on Commun., vol. E77-B, no. 10, pp. 1175–1187 (1994).