THE OVERLOAD PERFORMANCE OF ENGINEERED NETWORKS WITH NONHIERARCHICAL AND HIERARCHICAL ROUTING

Akinpelu, J. M.
Bell Telephone Laboratories, Incorporated
Holmdel, New Jersey, U.S.A.

ABSTRACT

We report the results of a study of the performance of engineered nonhierarchical and hierarchical routing networks under overloads, using analytical and simulation models. We also examine the effect of one control, namely, trunk reservation for first-routed traffic. Based on our results, we recommend that such a control be considered in the introduction of nonhierarchical routing networks in order to ensure maximum useful network throughput under all traffic conditions.

1. INTRODUCTION

Hierarchical network design has been used since the early days of the telephone toll network. However, recent advances in switching and signaling technology, as well as the tremendous growth in toll traffic, have provided new incentives to increase network design efficiency. Since the mid 1970's, efforts have been underway at Bell Laboratories to develop a new method of engineering large-scale nonhierarchical networks. These efforts have culminated in the unified algorithm (1) (2). This algorithm engineers nonhierarchical networks which are less expensive than hierarchical networks and which give service comparable to that of the hierarchical networks under engineered traffic conditions. However, there was a need for a better understanding of the performance of the nonhierarchical networks and the effect of controls under other traffic conditions, specifically loads in excess of engineered loads. Such an understanding is a prerequisite for the deployment of nonhierarchical routing, so that maximum useful network throughput can be maintained at all times.

In this paper, we report the results of a study of the performance of engineered nonhierarchical and hierarchical networks under overloads, using analytical and simulation models. In addition, we examine the effect on these networks of one control, namely, trunk reservation for first-routed traffic.

2. BACKGROUND

Krupp (3) and Nakagome and Mori (4) carried out approximate analyses of nonhierarchical routing applied to small uniformly-loaded networks which are easily analyzed because of their simple symmetric designs. Their analyses revealed in some cases the existence of network instabilities. Thus for certain loads the network can operate in two states: (1) a low network blocking state in which almost all calls use their shorter first-choice path, and (2) a congested state in which a large proportion of calls use longer alternate paths and many calls are blocked. Furthermore, the value of the load at which congestion disappears is lower than the value of the load at which congestion first appears. This instability implies that a temporary increase in offered load may cause the network to enter a state of congestion where it remains for some period after the load has returned to normal.

For example, Krupp and Nakagome and Mori showed that instabilities occur for a ten-node symmetric network with 100 trunks per link for point-to-point offered loads between 71 and 88 erlangs. We sought to verify this behavior using a simulation model developed by Krupp and found that the predicted behavior did in fact occur. Fig. 1 shows a simulation run for this network with a point-to-point offered load of 80 erlangs, starting with an empty network. (Time is measured in call holding times.) In this run few calls are blocked and carried load is fairly constant at around 3600 calls. To determine the behavior with this same load when the network is initially congested, we simulated the network with an offered load of 90 erlangs for approximately 14 holding times. Then the load was dropped to 80 erlangs. This simulation is shown in Fig. 2. Note that the initial load causes the network to become
congested with carried load only at about 3 000 calls. When the load is dropped to 80 erlangs, congestion persists for another 16 holding times before the carried load increases to the level seen in Fig. 1.

When the load is dropped to 80 erlangs, congestion persists for another 16 holding times before the carried load increases to the level seen in Fig. 1.

These results spurred interest in the performance of nonhierarchical routing networks engineered using the unified algorithm. To address this question, we developed an analytical model that extends the models in (3) and (4) to more general nonhierarchical networks. We also extended a simulation model for nonhierarchical networks developed by Krupp. In these models we consider link performance only - no switching system effects are included. We applied our analytical and simulation models to engineered nonhierarchical networks based on three representative network models developed at Bell Laboratories - a 30-node model, a 25-node model, and a 140-node model. For comparison, we also analyzed versions of the 30-node and 25-node network models engineered for hierarchical routing. For the 30- and 25-node network models, the nonhierarchical network has about nine percent fewer trunks than its corresponding hierarchical network.

3. ANALYTICAL MODELS

3.1 ANALYTICAL MODELS WITHOUT CONTROLS

The model for the nonhierarchical networks without controls requires that a network, together with link sizes, point-to-point offered loads, and a fixed route for each point-to-point pair, be specified. It is assumed that (1) the mixture of traffic offered to a link is Poisson, (2) link blocking probabilities are independent, and (3) the time needed to make connections (setup time) is small enough, relative to the average holding time of calls, to be ignored. We also assume that a call blocked on a link of a path can always be cranked back to the originating switching system so that the call can access the next path in its route.

Let \( U \) be the offered load for point-to-point pair \( j \) and let \( L \) be the total offered load for all point-to-point pairs. Let \( p_i \), \( n_i \), and \( a_i \) denote respectively the blocking probability, link size, and offered load for link \( i \) (in erlangs), and let \( q_i = 1 - p_i \). We denote a path by \( r \), a route by \( R \), the route for point-to-point pair \( j \) by \( R_j \), and the route formed by the first \( k \) paths of \( R \) by

\[
R_k^j = \{r_1^j, \ldots, r_k^j\}.
\]

Finally, we define \( D(R) \) to be the probability that route \( R \) is blocked, i.e., each path in \( R \) has at least one blocked link.

The basic idea of our analysis is to determine the offered load \( a_i \) for link \( i \) as a function of the link blockings. For each route containing link \( i \), we determine the contribution that the route makes to the total link offered load. Suppose link \( i \) is in path \( r_{i}^{j} \), the \( k^{th} \) path in the route for point-to-point pair \( j \). The load carried by \( r_{i}^{j} \), \( c_{i}^{j} \), is given by
\[ c_k^j = L^j \left[ D \left( R_{k-1}^j \right) - D \left( R_k^j \right) \right] , \]  

which is the point-to-point load for pair \( j \) that overflows the first \( k - 1 \) paths but not the \( k \)th path. The load \( c_k^j \) contributes to the carried load for each link \( i \in \mathcal{I}_k^j \). The total carried load for link \( i \), \( K_i \), is obtained by considering all paths containing link \( i \) and is given by

\[ K_i = \sum_{j,k} c_k^j . \]

From the relation

\[ K_i = a_i q_i \]

for Poisson traffic, we immediately obtain

\[ a_i = \sum_{j,k} \frac{c_k^j}{q_i} . \]  

(2)

In addition to the relations given by equation (2), based on our assumption of Poisson link-offered loads, we relate the link offered loads to the link blockings by the Erlang-B formula:

\[ p_j = B(n_j, a_j) . \]  

(3)

Equations (2) and (3) can be solved iteratively, starting with an initial estimate of the link blockings, to determine the link offered loads and blocking probabilities in equilibrium. (The calculation of \( D(R_k^j) \) is discussed below.) Once a solution has been obtained, several quantities of interest can be calculated. In particular, network blocking, \( z \), is given by

\[ z = \sum_j L^j D(R_k^j) , \]

and network carried load, \( C \), by

\[ C = L(1 - z) . \]

If the paths of \( R_k^j \) are disjoint, then, by our link independence assumption, the blocking probabilities for the paths are independent, so that

\[ D(R_k^j) = \prod_{t=1}^{k} \left[ 1 - \frac{\pi_{i_t}^{j}}{\pi_{i_t}^{j}} q_i \right] . \]  

(4)

For the nonhierarchical networks studied here, the paths have all been restricted to contain either one or two links. This implies that all paths in a route are disjoint, and \( D(R_k^j) \) can be calculated using the formula above.

The model used for the hierarchical networks without controls was taken from (5). We used only the link portion of the model, omitting the parts dealing with switching system queuing dynamics, retrials and DABY (don’t answer, busy).

3.2 ANALYTICAL MODELS WITH TRUNK RESERVATION FOR FIRST-ROUTED TRAFFIC

Network performance was also modeled with trunk reservation for first-routed traffic. Under this control, a threshold is specified for each link, and alternate-routed calls attempting to seize a trunk on the link are refused if the number of busy trunks on the link has reached the threshold. For the nonhierarchical networks, this control was implemented by subjecting a call to trunk reservation on all legs of an alternate path. In the hierarchical networks, the classification of a call as first-routed or alternate-routed was made at each switching system the call traversed as follows: On the first-choice link out of a switching system the call is considered first-routed, whereas on any other link it was considered alternate-routed. All alternate-routed calls offered to a link were subjected to this control. A call overflowing a link because of trunk reservation was offered to the next path in its route. When the first path of a route is shorter (in terms of the number of links) than the alternate paths, then under large loads, this control has the effect of decreasing the average number of links per call and thus increasing network carried load.

For the trunk reservation model, let \( m \) be the trunk reservation threshold on the link, \( \hat{a} \) the probability that no more than \( m - 1 \) trunks in the link are busy, \( \hat{a} \) the link offered load subject to trunk reservation on the link, \( \hat{R} \) the link carried load subject to trunk reservation on the link, and \( r = \hat{a}/a \) (a is the total link offered load). Using a birth-death model for the behavior of \( n \) servers with offered load \( a \) when less than \( m \) servers are busy and offered load \( a(1 - r) \) when at least \( m \) servers are busy, we obtain the probabilities \( P_j \) that exactly \( j \) trunks on the link are busy:

\[ P_j = \binom{n}{j} \hat{a}_j (1 - \hat{a})^{n-j} , \]

where \( \hat{a}_j = \frac{1}{n!} \hat{a}^j (1 - \hat{a})^{n-j} \) and \( \hat{a} = \frac{1}{m} \).
\[ P_j = \frac{a_j}{j!} P_0, \quad j = 0, \ldots, m - 1, \]
\[ = \frac{a_j}{j!} (1 - r)^j m P_0, \quad j = m, \ldots, n. \]

Here
\[ P_0 = \left[ \sum_{k=0}^{m} \frac{a^k}{k!} + \sum_{k=m+1}^{n} \frac{a^k}{k!} (1 - r)^k m \right]^{-1}. \]

It follows that
\[ p = \frac{a^n}{n!} (1 - r)^n m P_0 \quad (5) \]
\[ \bar{q} = \sum_{j=0}^{m-1} \frac{a_j}{j!} P_0. \quad (6) \]

The quantities \( p \) and \( \bar{q} \) are easily calculated using recursive formulas.

We also need formulas for the calculation of \( a \) and \( \bar{q} \), which we obtain by relating these quantities to the corresponding carried loads. The total carried load \( K \) is given by
\[ K = \sum_{j=0}^{n} jP_j \]
\[ = ar\bar{q} + a(1 - r)q, \]
so that the offered load that is not subject to trunk reservation is given by
\[ a - \bar{q} = \frac{K(1 - r)}{a\bar{q} + (1 - r)q}. \quad (7) \]

We relate the offered load \( a \) to the carried load \( K \) by
\[ a = \frac{K}{a\bar{q}}. \quad (8) \]

Carried load \( K \) is still calculated as in the previous section with \( \bar{K} \) being the portion of \( K \) that was subjected to trunk reservation.

The nonhierarchical model with trunk reservation for first-routed traffic is obtained by replacing equation (2) with equations (7) and (8) and equation (3) with equations (5) and (6). The calculation of \( D(R_{ik}^j) \) in equation (4) is now given by
\[ D \left( R_{ik}^j \right) = \left[ 1 - \prod_{i=1}^{\pi} q_i \right] \prod_{t=2}^{k} \left[ 1 - \prod_{i=1}^{\pi} \hat{a}_i \right]. \]

The hierarchical model with trunk reservation for first-routed traffic is obtained by combining equations (5)-(8) with the Franks and Rishel formulas for link carried load.

4. RESULTS WITHOUT CONTROLS

In this section we describe the results obtained by applying the analytical and simulation models without controls to the 30-, 25-, and 140-node network models. All three network models were studied under uniform overloads of up to 200 percent, which were obtained by multiplying the engineered loads by the appropriate factor.

The results for the 30-node networks are presented in Figs. 3 and 4. These results show a striking difference in the performance of the nonhierarchical and the hierarchical networks. The two networks show similar performance up to about a ten percent overload, with carried load increasing with increasing offered load. However, at that point the number of calls carried in the nonhierarchical network falls sharply, due to an increase in the number of multi-link calls. The drop continues until around 100 percent overload, where a gradual increase in carried load begins. This increase results because the network has become congested to the point that the probability of finding available trunks for a multi-link call is very small, so that now one-link calls begin to be favored over multi-link calls. Such throughput degradation is not seen in the hierarchical network which, because of the presence of final links, the presence of primary high usage links that do not carry alternate-routed traffic, the absence of crankback, and a larger number of trunks than in the nonhierarchical network, is intrinsically better able to limit the number of multi-link calls under overloads.

In fact, the number of calls carried in the hierarchical network increases steadily as offered load increases over the entire range of overloads considered.
Another view of the difference in the behavior of the two networks is shown in Fig. 5. Here we have plotted the ratio of the number of multi-link calls to the number of one-link calls as determined from the analytical model for various overloads. This ratio grows sharply for the nonhierarchical network for overloads of up to 50 percent, then levels off and begins to decline. On the other hand, for the hierarchical network, this ratio rises slowly before leveling off at around 30 percent overload. This shows that the hierarchical network makes more efficient use of its trunks under overloads than does the nonhierarchical network.

The results for the 25-node network are presented in Fig. 6. As with the 30-node network, the 25-node nonhierarchical network shows a drop in carried load: at about ten percent overload in the analytical model and at about 15 percent overload in the simulations. The hierarchical network exhibits a continuous increase in carried load with increasing offered load.

The performance of the 140-node nonhierarchical network is displayed in Fig. 7. These results were obtained...
from the analytical model only. The results agree qualitatively with those derived for the 30- and 25-node nonhierarchical networks. In Fig. 7 we also show the number of one-link calls in the network as derived from the analytical model. We see that the direction of change in the number of one-link calls is almost always the same as the direction of change in the total number of calls. Again we conclude that the degree to which network capacity is efficiently used is related to the network's ability to favor one-link calls over multi-link calls.

5. COMPARISON OF THE NONHIERARCHICAL NETWORKS

We have shown that the nonhierarchical networks studied here demonstrate a drop in carried load under overloads without controls. The most severe drop occurs in the 30-node network, where carried load drops 8.37 percent below the level at ten percent overload before leveling off (based on the analytical model). The 25- and 140-node networks demonstrate respective drops of 0.58 and 1.13 percent. Krupp (3) has shown that, for symmetric nonhierarchical networks, all other things being constant, performance under heavy loads worsens as the number of paths per route increases. In fact, the 30-node network provides on average 6.91 alternate paths per route, the 25-node network provides 2.12, and the 140-node network provides 4.41. Thus we see a strong relationship between route size and poor overload performance. This further substantiates the conclusion that the reason for the poor overload performance of the nonhierarchical networks is the large amount of alternate-routing under overloads made possible by the large number of alternate paths.

6. RESULTS WITH TRUNK RESERVATION FOR FIRST-ROUTED TRAFFIC

We now consider the effect of trunk reservation for first-routed traffic on the performance of these networks. On each link, five percent of the trunks, with a minimum of one trunk, were reserved for first-routed traffic.

The effects of trunk reservation on the 30-, 25-, and 140-node networks are shown in Figs. 8-10. By comparing these results with the results without trunk reservation, we see that trunk reservation has a beneficial effect with nonhierarchical routing under large overloads. The drop in carried load with increasing offered load seen without controls disappears with the use of trunk reservation. Trunk reservation limits the number of multi-link calls, allowing more efficient use of the trunks. Similarly, the performance of the hierarchical networks under large overloads is improved, although the improvement is not as dramatic as for the nonhierarchical networks. However, in both the nonhierarchical and hierarchical networks, the use of trunk reservation results in a small decrease in network carried load at the lighter overloads. Based on these characteristics, a control strategy for nonhierarchical routing networks can be formulated which provides maximum useful network throughput (6).
7. SUMMARY

We have examined the performance of three network models under uniform overloads using analytical and simulation models. Results obtained demonstrate that, without controls, the performance of nonhierarchical networks is inferior to that of comparable hierarchical networks under overloads. No instabilities of the type described in (3) and (4) are seen in the nonhierarchical networks, but a drop in carried load consistently occurs at about ten percent overload. This results because of the tendency of the nonhierarchical networks to alternate-route calls when the networks become overloaded. The severity of this drop appears to be correlated with the number of paths per route; nonhierarchical networks with more paths per route exhibit greater throughput degradation under overloads. Such behavior is not seen in the hierarchical networks, which are intrinsically better able to limit the amount of alternate-routing under overloads.

We also have applied a control, namely, trunk reservation for first-routed traffic, to the network models. This control has an extremely beneficial effect on the performance of the nonhierarchical and hierarchical networks under overloads. By diminishing the amount of alternate-routing in the networks, it allows for more efficient use of the trunks by one-link calls. For the nonhierarchical networks, this results in a continuous increase in carried load with increasing offered load over the entire range of overloads considered. We recommend that such a control be considered in the introduction of nonhierarchical routing networks as exemplified in the proposed control strategy described in (6).

REFERENCES


Date: 13 June 1983
Session: 3.2
Paper: 4

Q.1 (Halgreen)

Have you studied the effect of trunk reservation on network performance in case of focused overload?

A.1 (J.M. Akinpelu)

No, I have not looked at this with my models, although the models are applicable to focused overloads. However, focused overloads have been studied using the SPC-CCIS simulator discussed in paper 3.2-5. In the simulator, selective trunk reservation and selective dynamic overload control for switching systems are applied, which utilize the hard-to-reach transmission capability over the CCIS network. The inherent robustness of the SPC-CCIS network prevents the overload from spreading through the network. The selective controls serve to keep calls coming into and going out of the focus.

Q.2 (W. Robinson)

The analytical model agrees very well with simulation results for both FHR and DNHR networks in global average blocking.

How good is the agreement between the analytical model and simulation for individual node pair blocking?

A.2 (J.M. Akinpelu)

Node pair blocking comparison between the analytical and simulation models have been made only casually; no specific results are available.

Q.3 (W.H. Cameron)

DNHR has engineered paths for call routing, and real-time back-up paths which can be used as needed and as available and on which some reservation is made for first-route traffic. Would trunk reservation on the engineered paths improve the overload performance of DNHR, letting it come closer to carrying one call per trunk? Would it affect network performance under normal conditions? How much?
A.3 (J.M. Akinpelu)

Trunk reservation on engineered paths does improve overload performance under large overloads, even when real-time paths are used with trunk reservation. In fact, under large general overloads, the real-time paths essentially become unavailable for use, and trunk reservation of the engineered paths allows more efficient use of the network by one-link calls. At engineered loads, trunk reservation can degrade performance if it is too severe. We have experimented with trunk reservation variants that do improve overall network performance at engineered loads, but which result in increases in some parcel blockings. There has been some concern that these increases would adversely affect customer service, Trunk reservation on engineered paths does improve overload performance under large overloads, even when real-time paths are used with trunk reservation. In fact, under large general overloads, the real-time paths essentially become unavailable for use, and trunk reservation on the engineered paths allows more efficient use of the network by one-link calls. At engineered loads, trunk reservation can degrade performance if it is too severe. We have experimented with trunk reservation variants that do improve overall network performance at engineered loads, but which result in increases in some parcel blockings. There has been concern that these increases would adversely affect customer service, and so to avoid these effects, an automatic triggering mechanism will be used as discussed in paper 3.2-5.