Design of ATM Networks with Multiple Traffic Classes

Byung Han Ryu, Cheol Hye Cho, and Jee Hwan Ahn

CONTENTS

ABSTRACT

In this paper, we propose a new heuristic design algorithm for the virtual path (VP)-based ATM network with multiple traffic classes, in which QoS constraints associated with traffic class are taken into account. The minimum bandwidth of VP required to carry given amount of traffic is obtained by utilizing an equivalent bandwidth concept, and the route of each VP is placed so that the network cost is minimized while the QoS requirement is fulfilled. To evaluate our design algorithm, we consider two kinds of traffic: voice traffic as low speed service and still picture traffic as high speed service. Through numerical examples, we demonstrate that our design method can achieve an efficient use of network resources, which results in providing a cost-effective VP-based ATM network.

I. INTRODUCTION

An asynchronous transfer mode (ATM) technology is capable of supporting a wide variety of connections with different bandwidth requirements and traffic characteristics by means of resource management control. As an effective way to facilitate the coexistence of traffic with diverse traffic characteristics and different quality of service (QoS) requirements in ATM networks, a virtual path (VP) concept has been proposed [1]-[4]. A VP is defined as a direct logical connection between two end nodes, and can accommodate multiple virtual channels (VCs) simultaneously. As its name indicates, VPs are virtual and are multiplexed on the physical transmission link in a cell multiplexing manner. When a VP is established between two nodes, all VCs on that VP are multiplexed and routed via an identical predefined path. In a multimedia network environment, the network is expected to be able to treat various kinds of information, which have different demands of QoS requirements. In attempting to design the ATM network, it requires considering the network topology and traffic pattern generated from users so that the network cost can be minimized while satisfying QoS requirements such as cell/call loss probabilities and delay times.

In this study, we propose a new heuristic design algorithm for the VP-based ATM network which can handle multiple traffic classes. In the proposed algorithm, for given traffic load, routes of VPs are established to minimize the network cost while satisfying the QoS constraints of each traffic class for every pair of end nodes. For this purpose, the physical network topology, which is represented by a connection graph of links between nodes, is assumed to be given. We further assume that each traffic class is assigned to a separate VP even if several traffic classes have identical source and destination nodes. As related works, the heuristic design algorithm for packet switching communication networks, which handles multimedia information with different QoS, has been evaluated in [5]. The design methods of ATM networks based on VP have been presented in [6]-[11]. The planing method of ATM networks based on evaluation of two design items, VC/VP routing, taking account of virtual path identifier (VPI) field constraints have been described in [6]. In [7], the virtual path layout design on ATM networks for minimizing the overall VC setup and switching cost is presented. An efficient algorithm that finds VP routes to minimize the traffic load on individual links is presented in [8]. A mathematical formulation for an optimal strategy of VP connections, considering the trade-off between increased capacity costs and reduced control costs, has been developed in [9]. A network design problem that arises when designing VPs to support client/server applications has been presented in [10]. More recently, heuristic design algorithms for VP-based ATM networks have been proposed in [11]. In these network design methods, however, the link cost has been assumed to be a linear function of the total VP bandwidth [6]-[10], and the route of each VP is established based on the shortest path [6],[8]. However, in our study, the link cost is formulated as a linear function of the number of physical lines instead of the total bandwidths of VPs passing through the link, we try to minimize the network cost by eliminating unnecessary links through our algorithm.

In this paper, for determining the minimum bandwidth required to accommodate a given number of VCs while satisfying QoS requirements in terms of cell loss probability at cell level and call loss probability at call level, an equivalent bandwidth concept [12]-[15] is introduced. More precisely, we first determine the maximum number of VCs, which should be admitted simultaneously, for each VP to satisfy the constraint of call loss probability at call level. This can be accomplished through a classical Erlang-B formula. Then, for the given number of VCs, we determine the minimum bandwidth of VP by utilizing an equivalent bandwidth method under the constraint of cell loss probability at cell level. Since the equivalent bandwidth approach does not take into account the delay times, however, we restrict the number of hops of every VP in determining its route. By this method, the maximum delay for each cell can be bounded for given propagation delays and buffer sizes at the switching nodes.

Our algorithm presented in this paper is based on the one in [11] but extended to establish the ATM networks with multiple traffic classes as will be explained in Section 4. Following [11], our design algorithm is divided into two phases: an initialization phase and an adjustment phase. In the initialization phase, the bandwidth of each VP is first determined, and then all the routes of VPs are temporarily established through shortest paths in similarity in conventional network design methods [6],[8]. In the adjustment phase, the possibility of cost reduction is sought by (1) alternation of VP routes, (2) separation of the single VP into multiple VPs, and (3) introduction of VCX nodes by which several VPs on the link are multiplexed in a statistical cell multiplexing manner. It is desirable to establish a VP-based network by its simplicity of network and traffic management. Therefore, we first establish VPs for all pairs of end nodes. In this case, while the statistical multiplexing gain at cell level within the VP can be achieved, we cannot expect it by VP multiplexing on each link. As a result, the capacity of the physical link, which is given by the sum of the required bandwidth of VPs on that link, tends to be large. Then, the network cost would also become large. Therefore, we will also check the possibility of introducing VCX nodes in our algorithm. Since the establishment of VP has a benefit in call admission control (CAC) because of the simplification of the call setup procedure (although its overhead is not formulated in our algorithm), we treat the introduction of VCX nodes as an option. This is why the introduction of VCX nodes is examined in the last phase of our algorithm.

This paper is organized as follows. Section II briefly summarizes the equivalent bandwidth method, which will be used to determine the VP bandwidth under QoS constraints. Section III describes a formulation of our network design problem. Then, we introduce our design algorithm in detail in Section IV. In Section V, we present some numerical results. Concluding remarks and further studies are described in Section VI.

II. DERIVATION OF EQUIVALENT BANDWIDTH

An equivalent bandwidth of a set of VCs multiplexed on a VP is defined as an amount of bandwidth required to achieve a desired QoS. In order to characterize the equivalent bandwidth or effective bit rate of VCs in terms of known parameters, we need to select an appropriate model. In this paper, we calculate the equivalent bandwidth of VCs on the basis of its statistical characteristics at cell level by means of fluid flow approximation and stationary approximation methods [12]. By assuming that VCs within the same VP have identical traffic characteristics, each VC is modeled by a two-state Markov source whose peak rate, mean idle period, and mean busy period are denoted by R_{peak}, $1/\lambda$, $1/\mu$, respectively. Let the buffer size be K. Further, an allowable cell loss probability is given by ε. Then the equivalent bandwidth at cell level by fluid flow approximation for N multiplexed VCs, $\widehat{C}_{(F)}(N)$, is given by

$$\widehat{C}_{(F)}(N) = \frac{N\{A - \sqrt{A^2 + 4\lambda\alpha R_{peak}/K}\}}{-2\alpha/K} \qquad (1)$$

where $\alpha = \ln(1/\varepsilon)$ and $A \equiv \lambda + \mu - R_{peak}\alpha/K$.

For N multiplexed VCs, let the mean aggregate bit rate and the standard deviation of aggregate bit rate be $m_N (= Nm)$ and $\sigma_N (= \sqrt{N\sigma^2})$, respectively. Then the equivalent bandwidth by stationary bit rate approximation, $\widehat{C}_{(S)}(N)$, is derived as

$$\widehat{C}_{(S)}(N) = m_N + \sqrt{-2\ln(\varepsilon) - \ln(2\pi)}\,\sigma_N. \qquad (2)$$

Note that while the value of the equivalent bandwidth calculated by both approximation methods is overestimated when compared with a practical bandwidth, it can be utilized as an upper bound value which can satisfy QoS requirements. Finally, we obtain the equivalent bandwidth, $\widehat{C}_{\text{cell}}(N)$, as a function of the number of multiplexed VCs by the following expression:

$$\widehat{C}_{\text{cell}}(N) = \min\{\widehat{C}_{(F)}(N), \widehat{C}_{(S)}(N)\}. \qquad (3)$$

Next, assuming that the amount of traffic load for a single VP is given by a in Erlang, we use the following well-known Erlang loss formula in order to obtain the minimum number of acceptable VCs, k^*, at call level which keeps the call loss probability below given η.

$$k^* = \min\{k : E(k,\ a) \le \eta\}, \qquad (4)$$

where $E(k,\ a) = \dfrac{a^k/k!}{\sum\limits_{i=0}^{k} a^i/i!}$.

We first determine k^* using (4) for the given a. Then, the equivalent bandwidth C^* of a VP, which is required to transfer a given amount of traffic while satisfying QoS requirements at both cell and call levels, is obtained by

$$C^* = \overset{\frown}{C}_{\text{cell}}\,(k^*). \tag{5}$$

III. NETWORK COST FUNCTION

In this section, we first introduce notation and modeling assumptions followed by the introduction of the network cost function. For the initial network topology, it is assumed that we only know whether the establishment of the direct link between two nodes is allowed or not. A link consists of one or more physical lines of, e.g., 150 Mbps. Then, in what follows, we use the term link as a logical connection between two nodes, and its capacity is given by the multiplicity of the physical lines which can accommodate one or more VPs on that line.

To formulate the network construction cost, we use the following notations.

h_k : maximum allowable number of hops for traffic class $k(1,\ \ldots,\ s)$

$P_{l;k}$: cell loss (call blocking) probability of traffic class k at level $l = 1 (l = 2)$ $(l = 1$; cell level, $l = 2$; call level)

$P_{l;k}^{\max}$: maximum allowable cell loss (call blocking) probability for traffic class $k$at level $l = 1 (l = 2)$

$\omega_{pq;k}$: VP of an end node pair p, q of traffic class

l_{ij} : the link connecting adjacent nodes i and j

τ_{ij} : the number of physical lines on linkl_{ij}

l_{ij}^m : m-th physical line on link l_{ij}

Ω_j : a set of all nodes directly connected to node j

L_{P_j} : a set of all lines connected to VPX node j

L_{C_j} : a set of all lines connected to VCX node j

L_j : a set of all lines connected to node j $(= L_{P_j} \cup L_{C_j})$

$C_{\omega_{pq;k}}(l_{ij}^m)$: an equivalent bandwidth of VP $\omega_{pq;k}$ on line l_{ij}^m

$\{\omega_{pq;k}$ on $l_{ij}^m\}$: a set of all VPs $\omega_{pq;k}$'s passing through the physical line l_{ij}^m

In the case of a single traffic class, the subscript k which represents traffic class is negligible, i.e., $\omega_{pq;k} = \omega_{pq}$, $h_k = h$, $P_{l;k} = P_l$ and $P_{l;k}^{\max} = P_l^{\max}$.

A network cost is assumed to consist of a link cost and a node cost. The link cost is assumed to be a linear function of the number of lines instead of a total bandwidth of VPs passing through the link. Since we explicitly model the discrete nature of link capacity (i.e., 150 Mbps per a single line), two or more VPs with the same route are established if the required bandwidth between a source-destination pair is larger than the line capacity. We do not explicitly model the physical length of links. Instead, the number of hops between end nodes is considered for delay constraint. Noting that the total bandwidth of VPs passing through a single line is not larger than the line capacity S, we use the variable to represent the total bandwidth of VPs on the line l_{ij}^m, i.e., $X_{ij}^m \equiv \sum_{k=1}^{s} \sum_{\{\omega_{pq;k} \text{ on } l_{ij}^m\}} C_{\omega_{pq;k}}(l_{ij}^m)$ for $i \in \Omega_j$. Further, we define the operator $I(X)$ to model the discrete nature of physical line cost as:

$$I(X) = \begin{cases} 0, & \text{if } X = 0, \\ 1, & \text{if } X > 0. \end{cases}$$

Then, the cost of the link between adjacent nodes i and j, $G_{\text{link},\ l_{ij}}$, is formulated as follows:

$$G_{\text{link},\ l_{ij}} = \alpha \sum_{m=1}^{\tau_{ij}} I(X_{ij}^m), \qquad (6)$$

where α is the cost per physical line.

The node cost is assumed to be proportional to the sum of the equivalent bandwidths of VPs passing through the node. Further, the node cost is classified into the cost of VPX node with the cross-connect function and the cost of VCX node with the switching function. Thus, by gathering the equivalent bandwidths for all VPs on all lines $l_{ij}^m \in L_{P_j}$, we can express the VPX cost at node j, $G_{\text{VPX},\ j}$, as:

$$G_{\text{VPX},\ j} = \beta \sum_{\{\text{all } i \text{ for } i \in \Omega_j\}} \sum_{m=1}^{\tau_{ij}} y_{ij}^m, \qquad (7)$$

where β is VPX cost per equivalent bandwidth, and $y_{ij}^m = \sum_{k=1}^{s} \sum_{\{\omega_{pq;k} \text{ on } l_{ij}^m\}} C_{\omega_{pq;k}}(l_{ij}^m)$. On the other hand, the VCX cost at node j, $G_{\text{VCX},\ j}$, should be considered only if VC switching is required at node j. The VCX cost at node j for all $\omega_{pq;k}$'s associated with all lines can be expressed as:

$$G_{\text{VCX},\ j} = \gamma\beta \sum_{\{\text{all } i \text{ for } i \in \Omega_j\}} \sum_{m=1}^{\tau_{ij}} z_{ij}^m, \qquad (8)$$

where γ is a relative value of VCX cost to VPX cost per unit equivalent bandwidth, and $z_{ij}^m = \sum_{k=1}^{s} \sum_{\{\omega_{pq;k} \text{ on } l_{ij}^m\}} C_{\omega_{pq;k}}(l_{ij}^m)$. Then, we can express the total network cost, G_{nc}, as:

$$G_{nc} = \sum_{\text{for all } i,\ j} G_{\text{link},\ l_{ij}} + \sum_{\text{for all modes}} G_{\text{node},\ j}, \qquad (9)$$

where

$$G_{\text{node},\ j} = G_{\text{VPX},\ j} + G_{\text{VCX},\ j}.$$

For the delay time requirement, we use the number of hops of every VP in determining its route, which is related to the number of switching nodes which each VP passes by. That is, the maximum call setup delay is bounded by multiplying call processing times in a switching node by the number of intermediate nodes when the VP passes by VCX. Further, the cell delay times can be bounded by the number of hops as well. Therefore, all the routes of VPs should be established so as to satisfy the following equation.

Minimize

$$G_{nc} = \sum_{i,j} G_{\text{link},l_{ij}} + \sum_{\text{nodes}} G_{\text{node},j}$$

Subject to:

$$P_{l;k} \le P_{l;k}^{\max},\ l = 1,\ 2,\ \ldots,\ s.$$

hops for traffic class $k \le h_k$

Design variable:

bandwidth of each VP; $C_{\omega_{pq;k}}(l_{ij}^{m})$

routing of each VP

IV. DESIGN ALGORITHM

The design algorithm proposed in this chapter is divided into two phases for route establishment of each VP: an initialization phase and an adjustment phase. In the initialization phase of our algorithm, we first determine the required bandwidth of each VP by considering QoS of traffic sources. We then determine the routes of VPs on the links using the shortest paths on the given network topology. In the adjustment phase, we increase or decrease the capacity of the links by taking account of (1) QoS constraints, (2) a discrete nature of the link (i.e., the number of physical lines on each link is taken into account), and (3) a statistical multiplexing gain of VCs, and possibly VPs by allowing to introduce the VCX nodes. As a result, we may have lines or links with zero capacity, which results in eliminating those lines/links.

1. Initialization Phase

In the initialization phase, we temporarily establish the VP between every source-destination node pair on its shortest path in which the number of hops of each VP is considered as a path length. The bandwidth of each VP is determined by taking account of cell loss and call blocking probabilities as QoS constraints by means of the equivalent bandwidth method. As described earlier, the maximum allowable number of hops is used for the delay time requirement. However, it is likely to happen that the delay time constraint is not satisfied even if the shortest path is used. Since, in this study, the initial network topology is assumed to be given, we have no means for resolving such a problem.

We assign each VP using the shortest path for the two end nodes. That is, in this phase, we do not take into account a discrete nature of the link in principle. Our assumption that the link cost is proportional to the number of physical lines (in the unit of, say, 150 Mbps) is ignored in that sense. Therefore, the selection order of VPs is irrelevant. However, when we have two or more shortest paths (i.e., the paths with a same number of hops), we can use the following selection rule. When the examined VP is put on some links, it may require the additional line due to the lack of remaining capacity on that link. Since it apparently increases the link cost, such a link should be avoided for cost reduction. More specifically, when we have multiple shortest paths with same distance for the examined VP, we choose the path which requires the smallest number of additional physical lines. If the above condition is fulfilled and if we still have multiple shortest paths, the path is arbitrarily chosen for the designated VP. The discrete nature of links is explicitly treated in the following adjustment phase.

2. Adjustment Phase

Once we have the VP-based network by the initialization phase, the adjustment phase is performed next. Recall that we assume the link cost is proportional to the number of physical lines by taking into account the discrete nature of the link capacity, the adjustment phase tries to eliminate some physical lines to reduce the link cost. First, we check the possibility that the previously assigned VPs can be moved to the other links. We also examine the possibility that a single VP can be split into several VPs such that each of the separated VPs can be fit onto other links. Finally, the possibility of introducing VCX nodes is examined. It is true that the VP-based network is desirable as mentioned before. However, if we introduce the VCX node, a gain obtained by multiplexing several VPs in a statistical manner may lead to eliminate the unnecessary lines and even links at the expense of the increased node cost.

Multimedia traffic can be classified on the basis of its traffic characteristics and QoS requirements. In this study, we need to determine the selection order of VPs by which the possibility of cost reduction is examined. For this purpose, we introduce the traffic class ordering by considering the statistical multiplexing gain of traffic sources. Suppose that a single VP of n multiplexed VCs (denoted by $\omega_{pq;k}(n)$) is separated into VPs, $\omega_{pq;k}(n_1)$, $\omega_{pq;k}(n_2)$, $\ldots$, $\omega_{pq;k}(n_m)$, for class k. By separating the single VP into several VPs, the total bandwidth for accommodating these n VCs is increased due to the loss of statistical multiplexing gain. By allowing the associated bandwidth of as $\omega_{pq;k}(i)$ as $C^{\omega_{pq;k}(i)}$, the total bandwidth is increased by

$$\Delta_k = \sum_{i=1}^{m} C^{\omega_{pq;k}(n_i)} - C^{\omega_{pq;k}(n)}.$$

In what follows, we sort the traffic classes according to Δ_k, and we will call the traffic with small Δ_k as low class traffic and the one with large Δ_k as high class traffic.

A. Step (1): VP Movement Phase

By moving the previously assigned route to a new route, we try to reduce the network cost. For this purpose, we first choose the link which has a largest remaining capacity in the network. This is because the VPs on that link, which has a smaller capacity, can be accommodated by other links. After we select the link which may have multiple VPs, the VP with largest bandwidth is chosen. Then, the next shortest path is sought and checked whether the VP can be fit into each link of that path. The rational behind this ordering of VPs is as follows. When the VP with larger bandwidth is first moved to

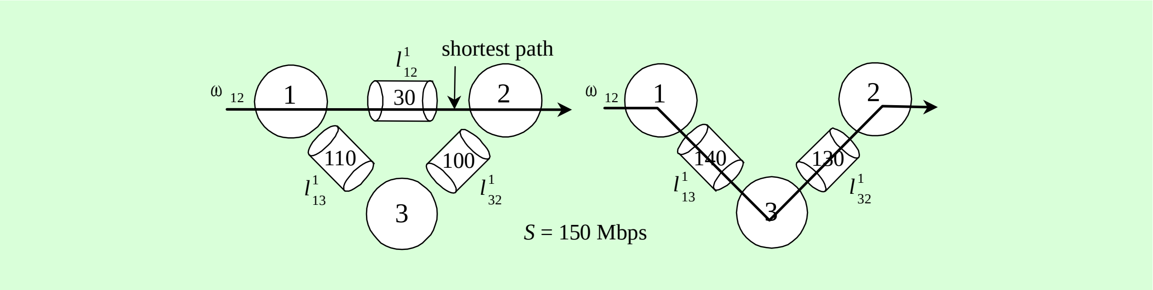


Fig. 1. An example of VP movement.

the other links, the remaining capacities on those links become small. However, even in that case, there still remains a possibility that the VP with smaller bandwidth can be fit into those links. Regardless of traffic class, this phase is performed in a decreasing order of the equivalent bandwidths of VPs on that link. For the chosen VP, we find the next shortest path so that it can be fit into the links on the path. If a suitable path cannot be found for that VP, then the next link with largest remaining capacity is checked. Further, if at least one line cannot be eliminated for the chosen link, the movement of VPs is canceled since it only increases the network cost by selecting the longer paths of VPs. We repeat this way until all VPs and links in the network are checked.

An illustrative example for this phase is shown below. In this example, we consider that the capacity of the single line is 150 Mbps. In Fig. 1, suppose that VP ω_{12} with 30 Mbps has been assigned to line l^1_{12} at the initialization phase, and the current total bandwidths of VPs on lines l^1_{13} and l^1_{32} are 110 Mbps and 100 Mbps, respectively. We can eliminate the unnecessary link if the hop count of each VP does not exceed the maximum allowable number of hops on the new route. That is, if the VP ω_{12} is moved such that it passes through lines, l^1_{13} and l^1_{32}, then the unnecessary line l^1_{12} can be eliminated.

B. Step (2): VP Separation Phase

In this phase, we separate a single route of VP into several routes i) if it is possible to decrease the link cost (and node cost), and also ii) if there exists the remaining capacity on other links larger than the bandwidth which each separated VP requires. The statistical multiplexing gain implies that the total required bandwidth is increased when the single VP is separated into several VPs. Nevertheless, it is meaningful because the network cost can be reduced by eliminating the unnecessary lines. Of course, we should consider the re-calculation of equivalent bandwidth, the (re)assignment of VPs to other links, and the effect of the separated VPs on other links. However, it would be difficult if we perform the above procedure in parallel. Similar to the previous phase, we order links and VPs for investigation as follows. We first choose the link which has a largest remaining capacity in the network. It implies that VPs on that link, which has the smallest capacity, can be fit into other links if those are separated into multiple VPs adequately. In actual, however, the selected link may have a single VP while another link (with larger capacity) has multiple VPs, each of which has a smaller bandwidth. In this case, it is not clear which is better for reducing the total cost since it depends on the traffic characteristics. From a viewpoint of a statistical multiplexing gain, to separate the VP with larger bandwidth may not be more desirable than to separate the multiple VPs with smaller bandwidths even when the total bandwidths of the latter VPs is greater than the former single VP. On the other hand, it is desirable to try to fit the VP with a larger bandwidth into other links first. Anyway, noticing that both links are checked at the end of this phase, we adopt the approach explained in this subsection

This phase is performed in such an order that the link with a largest remaining capacity is chosen first, and then VPs for high class traffic are examined within the selected link. The sum of equivalent bandwidths of separated VPs is not greatly increased in the case of low class traffic by its definition. Therefore, the possibility that VPs for the low class traffic can be separated is high even after the VP for the high class traffic is first selected. In other words, it would become impossible to establish separated routes of high class traffic by lack of remaining capacities if VPs of low class traffic are first examined. Among VPs within the same traffic class, the VP with the largest bandwidth is first examined. We should note here that in this phase, it is possible that the VP movement can be allowed for some VPs when movement of the VP was canceled at the VP movement phase because of failing line elimination. Therefore, for the selected VP, we first try to move the route of that VP under satisfying the delay time constraint. If it is not possible by the lack of the remaining capacity of the other links, we perform the following three steps for VP separation.

(i) Find the next shortest path. If there are multiple shortest paths, then check the remaining capacities of the links on the next shortest paths, and select the one which has the largest available capacity.
(ii) On the condition that the original VP cannot move to a new path, the VP separation is performed as follows. We establish a new VP on the selected path by assigning the part of traffic. It is easy to calculate the amount of carried traffic (in Erlang) for given bandwidth under the QoS constraints in terms of the cell loss probability through the equivalent bandwidth approach. Then, we need another VP in addition to the newly established VP for the remaining traffic. For this purpose, the next shortest path is examined.

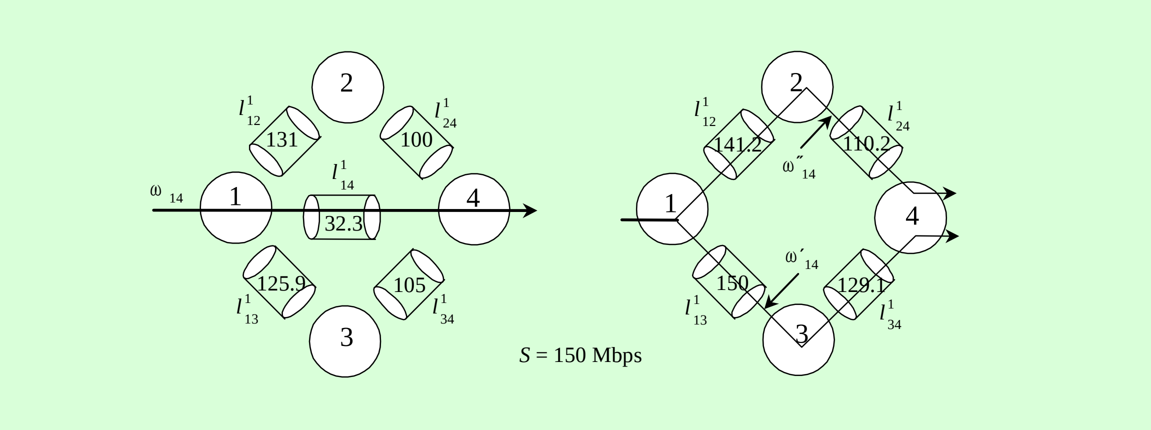


Fig. 2. An example of VP separation.

(iii) Repeat the above procedure until all of the traffic of the VP are assigned to the separated VPs. In carrying out this procedure, the constraint regarding the maximum number of hops may not be fulfilled. In this case, the separation of the original VP should be canceled.

The above procedure is iterated until all the VPs and links are checked.

Figure 2 shows an example of the VP separation phase. In the figure, each link is assumed to have a single line, and therefore link l_{ij} is identical to line l^1_{ij}. Suppose that the VPω_{14} from node 1 to node 4 requires the bandwidth of 32.3 Mbps. Since the remaining capacity of link l_{14} is largest, the VPω_{14} on that link is first examined for eliminating the link between nodes 1 and 4. Since the remaining capacities of links, l_{12}, l_{24}, l_{13}, and l_{34}, are currently 19 Mbps, 50 Mbps, 24.1 Mbps, 45 Mbps, respectively, it is impossible to establish a single VP from node 1 to node 4 unless the additional line is added between node 1 and node 2 or between node 1 and node 3. Therefore, we try to establish separated VPs. We have two next shortest paths; the one with links l_{12} and l_{24}, and the other with links l_{13} and l_{34}. Since the latter path has a largest available capacity (24.1 Mbps) on link l_{13}, we first choose it to establish a new VP. To fulfill the remaining capacity on link l_{13} of the path, VP ω'_{14} with the equivalent bandwidth 24.1 Mbps is established on the path. Then, the remaining traffic is carried on the path going through links l_{12} and l_{24}. Its required equivalent bandwidth is 10.2 Mbps, and the VP with 10.2 Mbps (ω''_{14}) is added on that link. Note that when the route of VP is separated, the equivalent bandwidth of each VP is calculated again because of the decrease of statistical multiplexing gain.

C. Step (3): VCX Introduction Phase

In the VCX introduction phase, we put VCX on the node i) if the separation of route of VPs is impossible by a lack of remaining capacity on any link, and ii) if there exists the remaining capacity which can accommodate the single VP after some VPs on the lines are multiplexed onto a single VP. Here, we only consider the case that the VPs of same traffic class on two or more lines between two adjacent nodes are multiplexed for simplicity. In this phase, a set of lines between two adjacent nodes on the line with the largest remaining capacity is chosen first. Then, we first examine VPs for the highest class traffic because the gain by multiplexing such VPs is large. In other words, cost reduction is large because the required bandwidth by multiplexing several VPs with high class traffic can be largely decreased. Among VPs within the same traffic class, VPs with the largest bandwidth is first examined due to the expectation of a large decrease of the equivalent bandwidth. We perform the following steps for selected VPs.

(i) Multiplex selected VPs in a statistical manner.

(ii) If a line can be deleted by the decrease of the equivalent bandwidth, the VCX is put at the associated node. Otherwise, select the VP with the next larger bandwidth.

(iii) Repeat the above steps until all VPs on the selected lines are checked.

The above procedure is repeated until all links in the network are examined.

Notice that once the VCX is introduced at intermediate node, it cannot be the end-to-end call blocking probability for each connection passing through the VCX node. That is, when a VCX node is introduced, the end-to-end call blocking probability turns into the call blocking probability between the VP termination nodes.

For example, suppose link l_{kj} having two lines, l^1_{kj} and l^2_{kj}, between nodes k and j. We check whether it can be accommodated in a single line $l^{'}_{kj}$ by introducing a VCX node or not. If (10) below is satisfied, the VCX node is introduced at node j as $l^{'}_{kj}$ instead of l^1_{kj} and l^2_{kj}.

$$\alpha I(x^{\prime}_{kj}) + \gamma\beta z^{\prime}_{kj} \leq \alpha I(x^1_{kj}) + \alpha I(x^2_{kj}) + \beta y^1_{kj} + \beta y^2_{kj}. \tag{10}$$

Two terms in the left hand side of the above equation represent the link cost and the VCX node cost, respectively, when node j is changed to a VCX node. The terms in the right hand side are the link cost on l^1_{kj} and l^2_{kj}, and the VPX cost for VPs on l^1_{kj} and l^2_{kj}, respectively, when not considering a VCX node at node j associated with the VPs on two lines between nodes k and j. Figure 3 shows an example of putting a VCX at node j in Step (3). Suppose that the required bandwidths of VPs on lines l^1_{ij} and l^2_{ij} are 142 Mbps and 9.2 Mbps, respectively. If (10) is satisfied by putting a VCX at node j for the VPs from node i, it can establish a single line l^1_{ij}. Of course, the decision should be made by taking into account the increasing cost of a VCX node.

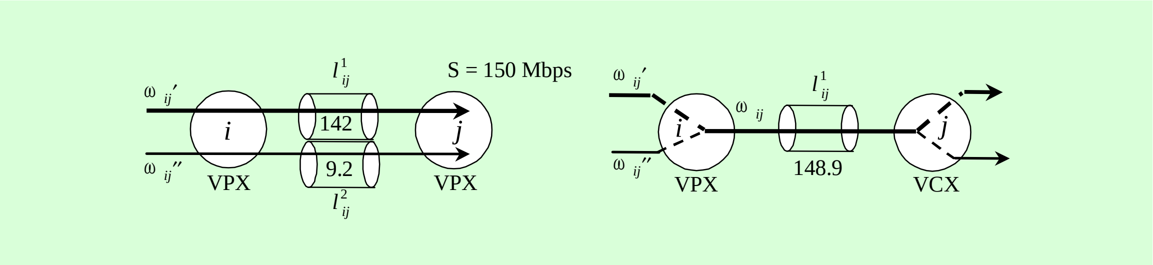


Fig. 3. An example of VCX introduction.

The above example shows a rather simple illustration in which cost reduction can be achieved by simply multiplexing VPs on two lines between two adjacent nodes. However, a more cost-saving network can be obtained if we go back to Steps (1) and (2) as described below. First, we check whether some VPs on two or more lines can be moved or separated to pass through the other links or not. Secondly, if movement or separation of VPs is possible, move or separate those VPs first, and for the other VPs on the lines, we try to introduce the VCX node. If the network cost is reduced, a VCX node may be introduced. That is, when the network cost cannot be reduced by introducing VCX node at Step (3), we examine whether Step (1) or (2) can be applied to the VPs on the concerning link or not. If it is the case, we check the possibility of cost saving by introducing the VCX node. The above procedure is iterated for all links in the networks, and if the cost saving cannot be expected any more, the algorithm is completed.

V. NUMERICAL EXAMPLES AND DISCUSSIONS

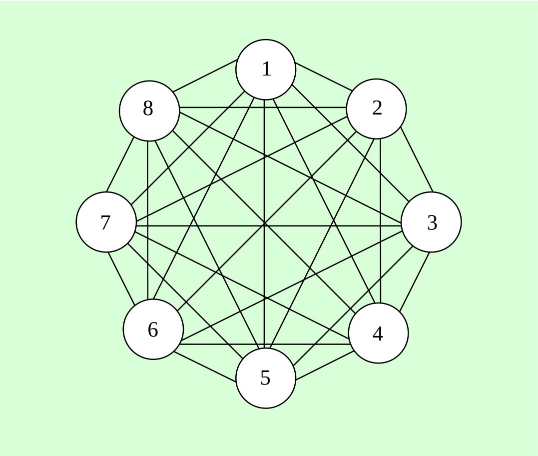


Fig. 4. Network model (metropolitan area network).

In this section, we present numerical examples through employing our proposed design algorithm. In this example, we consider two classes of traffic ($s = 2$); voice ($k = 1$) as low class traffic and still picture transmission ($k = 2$) as high class traffic. For evaluating our heuristic design algorithm, the following network model with 8 nodes is used as shown in Fig. 4. The offered traffic load for voice service is shown

Table 1. Offered traffic load matrix (voice).

s d	1	2	3	4	5	6	7	8
1		4,255	3,791	831	662	1,039	933	1,327
2	3,954		5,625	1,137	931	1,032	1,091	986
3	3,757	6,097		5,044	3,789	5,408	5,006	3,157
4	908	1,132	5,032		2,553	2,643	1,243	642
5	782	1,027	3,825	2,766		2,471	1,034	568
6	1,045	1,521	5,597	3,093	2,441		3,024	1,049
7	960	1,116	5,525	1,147	866	2,595		2,055
8	1,300	1,020	3,720	674	585	1,028	2,561	

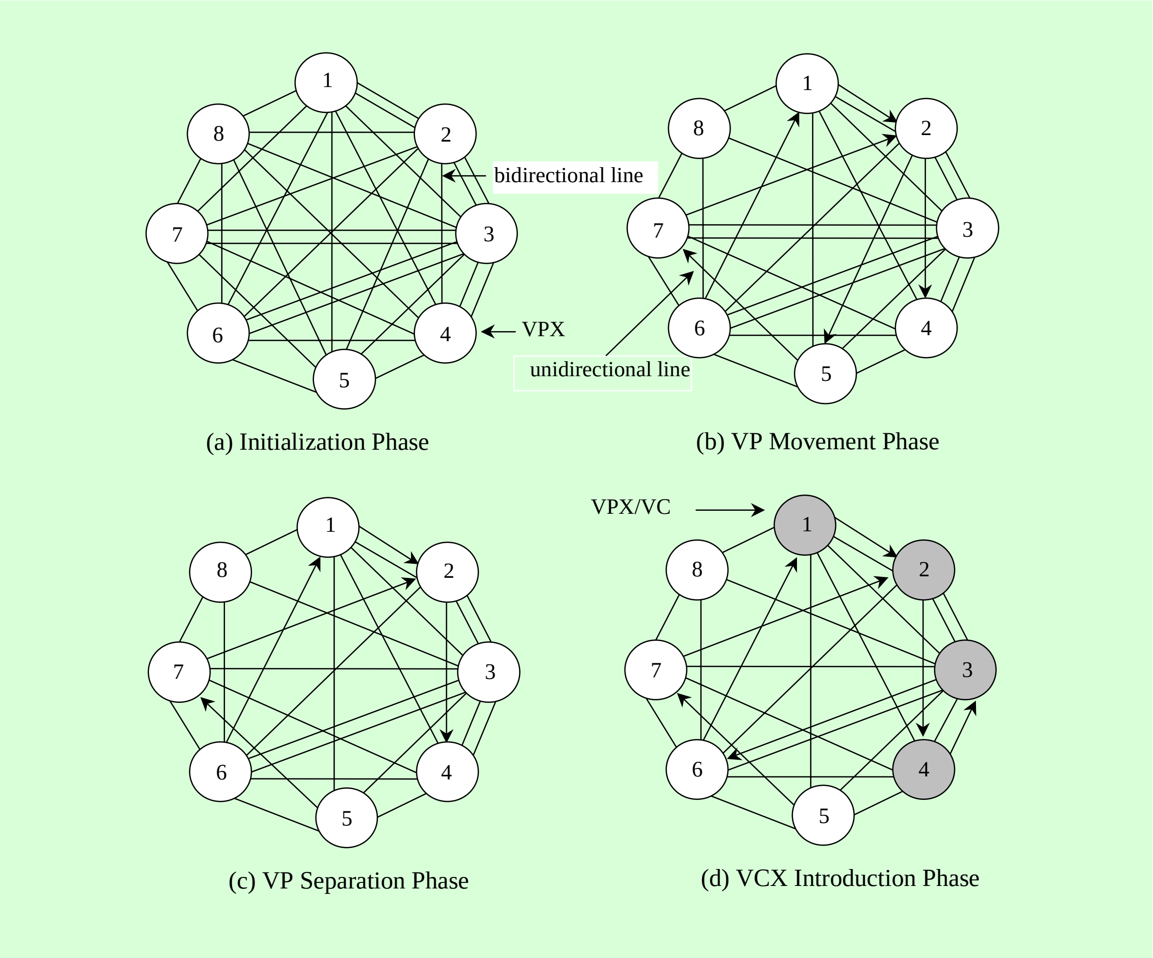
s: source node, d: destination node

in Table 1 in Erlang. In voice service, the allowable cell loss probability ($P_{1;1}^{\max}$) and call blocking probability ($P_{2;1}^{\max}$) are assumed to be 10^{-3} and 10^{-2}, respectively. The maximum allowable number of hops (h_1) is assumed to be 3 and the buffer size (K) is 150 cells. Peak rate, meanburst and idle periods of all sources are identically set to be 64 Kbps, 352 msec, 650 msec, respectively. For the still picture service, the allowable cell loss probability ($P_{1;2}^{\max}$) and call blocking probability ($P_{2;2}^{\max}$) are set to 10^{-9}, 10^{-2}, respectively, and the maximum allowable number of hops (h_2) and the buffer size are set to 3 and 1,000 cells, respectively. The traffic load of the still picture is assumed to have one-twentieth of voice traffic load as presented in [1]. Peak rate, mean burst, and idle periods of all sources are assumed to be 2.0 Mbps, 0.5 sec, 11 sec, respectively.

The capacity of a single physical line and the line cost, α, are assumed to be 150 Mbps and 150 per physical line, respectively. The unit cost, β and γ, is set to be relative to the line cost as described in Sect. III. The VCX node needs call processing while the VPX node does not. Therefore, the ratio of VCX node cost to VPX node cost per equivalent bandwidth of VP, i.e., γ, is expected to be larger than 1 as presented in [6]. Here, it is assumed that its value, β, is 1 per Mbps. In addition, for the given initial network topology, all nodes are assumed to be VPX nodes with cross-connect functions, and then source and destination nodes are also considered as the VPX nodes.

The results are shown in Fig. 5 and Table 2. The equivalent bandwidth (in Mbps) and route of each VP are partly shown in Table 2. In the table, the phases in which the route is finally fixed are also listed.

The total network cost after performing the initialization phase is obtained as 20,753



(a) Initialization Phase

(b) VP Movement Phase

(c) VP Separation Phase

(d) VCX Introduction Phase

Fig. 5. Network topology at each phase (voice and still picture).

(Fig. 5(a)). It consists of 10,853 for the VPX cost and 9,900 for the link cost.

By performing the VP movement phase, we obtain cost-saving considerably compared with the cost obtained in the initialization phase. In applying the movement of routes, the maximum number of hops is assumed to be 3. Therefore, while the node cost is increased to 11,593 from 10,853, the link cost is reduced to 7,800 from 9,900 by eliminating 14 unnecessary lines. Consequently, the total network cost is reduced to 19,393 from 20,753 (Fig. 5(b)).

When a single VP of a low class traffic is separated into multiple VPs, the increase of total equivalent bandwidth of separated VPs is not large due to the narrow bandwidth of low class traffic. Therefore, several VPs for voice traffic can be divided into multiple routes. It is true that if a single route is divided into several routes, the

Table 2. Equivalent bandwidths and route of VP (voice and still picture).

o	t	bandwidth	route	∗	o	t	bandwidth	route	∗
3	2	145.382	3-2(v)	1	2	1	95.943	2-1(v)	1
		75.168	3-2(s)	1			57.300	2-3-6-1(v)	2
2	3	134.536	2-3(v)	1	5	3	92.957	5-3(v)	1
		71.351	2-3(s)	1			55.809	5-3(s)	1
6	3	133.887	6-3(v)	1	1	3	46.695	1-3(v)	4
		71.300	6-1-3(s)	4			42.961	1-4-3(v)	4
7	3	79.449	7-3(v)	3			4.694	1-2-3(v)	4
		40.702	7-4-3(v)	3			55.634	1-3(s)	1
		19.636	7-2-3(v)	3	.	.	.	.	.
		70.547	7-3(s)	1	.	.	.	.	.
3	6	129.526	3-6(v)	1	.	.	.	.	.
		69.578	3-6(s)	1	7	5	17.310	7-6-5(v)	3
3	4	121.146	3-4(v)	1			7.731	7-2-4-5(v)	3
		66.480	3-2-4(s)	4			24.079	7-6-5(s)	2
.	.	.	.	.	4	8	17.801	4-1-8(v)	2
.	.	.	.	.			20.853	4-1-8(s)	2
3	7	83.824	3-7(v)	3	8	5	16.400	8-1-5(v)	2
		28.840	3-4-7(v)	3			20.001	8-1-5(s)	2
		14.474	3-4-7(v)	3	5	8	15.957	5-1-8(v)	2
		66.151	3-7(s)	1			19.713	5-1-8(s)	2

o: source node, t: destination node, v: voice traffic, s: still picture traffic
∗ indicates the phase number that route is determined;
1:initialization, 2:movement, 3:separation, 4:VCX introduction

network cost would increase due to the decrease of the statistical multiplexing gain. However, we can expect to eliminate unnecessary lines by dividing a single VP to several VPs. By performing VP separation phase, each of four single routes of VP is separated into two or three routes as shown in Fig. 5(c) and Table 2 (all not shown in this table). It results in decrease of the statistical multiplexing gain, and the node cost is increased to 11,791 from 11,593. However, by separating a single route, unnecessary 3 lines are eliminated and the total network cost becomes 19,141 and cost-saving is about 7.77% at the end of this phase compared with the cost obtained in the initialization phase. On the other hand, for VPs of still picture traffic, i.e., for high class traffic in this example, the possibility that a single route is divided into several routes was very rare.

The probability that the VCX node is introduced is not too high for VPs of voice traffic. The reason why the possibility of the VCX introduction is not high can be explained as follows; when several VPs are accommodated on a single VP by multiplexing several VPs, the decrease of total required bandwidth is not large due to the narrow bandwidth of voice traffic. This is because, in the case of voice traffic, the statistical multiplexing gain can be fully obtained even in the small bandwidth of VP. Suppose that we have two VPs for accommodating voice traffic. If these two VPs accommodate 200 VCs and 300 VCs under the QoS constraint in terms of 10-3 cell loss probability and 10-2 call blocking probability, the required bandwidths become 5.99 Mbps and 8.58 Mbps as seen in Fig. 6. If these two VPs are multiplexed onto a single VP, its required bandwidth becomes 13.16 Mbps. Since the sum of two VPs is 14.57 Mbps, the statistical multiplexing gain is only 1.41 Mbps. Therefore, the probability of introducing the VCX node is not too high in the case of voice traffic.

On the other hand, in still picture traffic, if several VPs are multiplexed onto a single VP and a VCX is put at the associated node, we can utilize the network resource cost-effectively. Consequently, in this example, VCXs are put at nodes 2 and 4 for still picture traffics on lines l^1_{32}, l^1_{24} and at nodes 1 and 3 for still picture traffics on lines l^1_{61}, l^1_{13}, respectively. Of course, it depends on the relative value of VCX cost, γ, and in our case, it can be applied when γ is below 1.38 (Fig.5(d)).

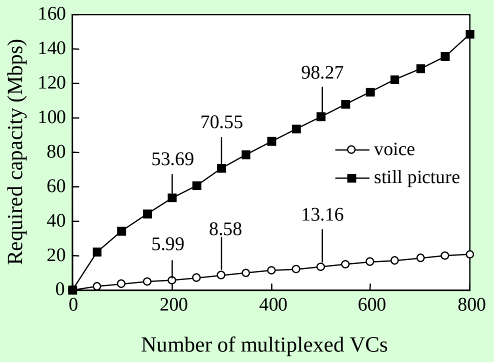


Fig. 6. Statistical multiplexing effect.

For example, we explain how the VCXs at nodes 2 and 4 are put for still picture traffic on lines l^2_{32}, l^1_{24}. At the end of VP separation phase, the assigned capacities on lines l^1_{32}, l^2_{32}, l^1_{24}, are 145.382 Mbps, 132.701 Mbps, 148.095 Mbps, respectively (Fig. 7(a)). Suppose that the VCX is put at node 2 for traffics on both lines l^1_{32}, l^2_{32}. In this case, we cannot expect a cost-saving because no lines can be eliminated. The network cost can be reduced by the following manner. First, VP $\omega_{27;1}$ with equivalent bandwidth of 28.72 Mbps passing through lines l^1_{24}, l^1_{47} is divided into two routes of VP $\omega'_{27;1}$ (19.318 Mbps), $\omega''_{27;1}$ (11.409 Mbps), each of which passes through lines $\{l^1_{24}, l^1_{47}\}$, and lines $\{l^1_{26},$

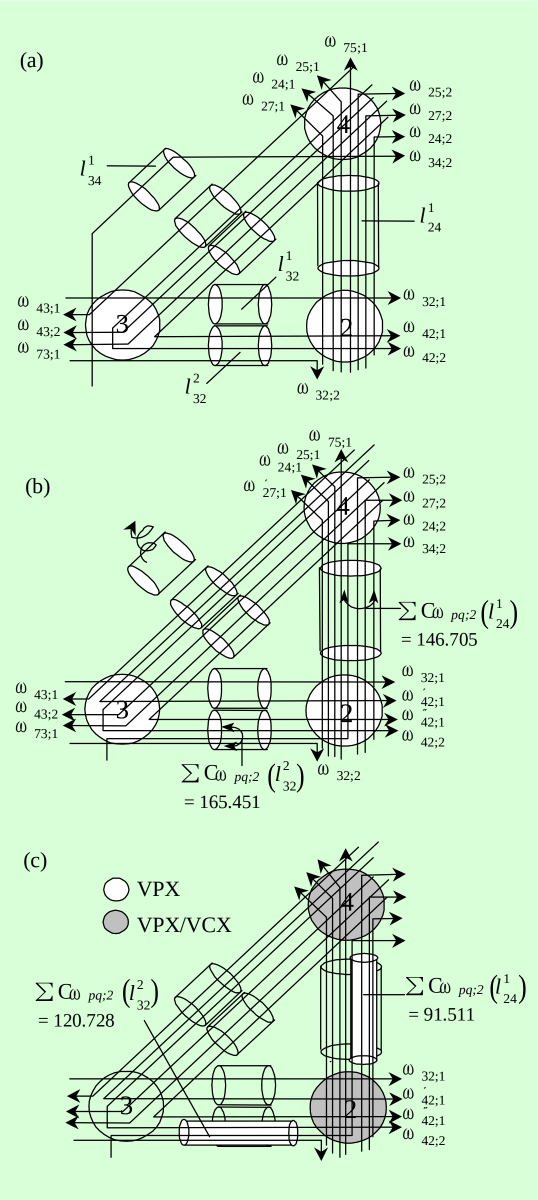


Fig. 7. Example of VCX introduction (voice and still picture).

$l^1_{67}\}$, respectively. Secondly, VP $\omega_{42;1}$ with 29.73 Mbps passing through lines $\{l^1_{43},\ l^2_{32}\}$ is divided into two routes of VP, $\omega'_{42;1}$ (4.613 Mbps), $\omega''_{42;1}$ (26.599 Mbps), each of which passes through lines $\{l^1_{43},\ l^1_{32}\}$ and other lines $\{l^1_{43},\ l^2_{32}\}$, respectively. Next, VP $\omega_{34;2}$ with 66.48 Mbps passing through line l^1_{34} is moved to pass through lines $\{l^2_{32},\ l^1_{24}\}$ and then line l^1_{34} is eliminated while the total capacities for both traffic classes on lines l^2_{32}, l^1_{24} turn into 196.050 Mbps, 205.173 Mbps, respectively. This means that it requires to construct two lines for traffics on lines l^2_{32} and l^1_{24} (Fig. 7(b)). However, by multiplexing the still picture traffic on lines l^2_{32} and l^1_{24}, respectively, and then introducing the VCX, the equivalent bandwidths for still picture traffic are reduced to 120.728 Mbps, 91.511 Mbps from 169.451 Mbps, 146.705 Mbps, respectively. Finally, by summing the equivalent bandwidths for both traffic classes, the total capacities on lines l^2_{32}, l^1_{24} result in 147.327 Mbps, 149.979 Mbps, respectively (Fig. 7(c)).

Finally, we show the influence of the VCX node cost-ratio γ in Fig. 8 which shows the network cost dependent on γ. For example, when γ is 1.38, the network cost obtained in the initialization, the VP movement, the VP separation, and the VCX introduction phases are 20,753, 19,393, 19,141, and 19,045, respectively. That is, our adjustment phase can perform cost reduction in about 8.23% compared to the initialization phase. In addition, if the value of γ is below 1.83, the network cost can be reduced than the one obtained in the VP separation phase.

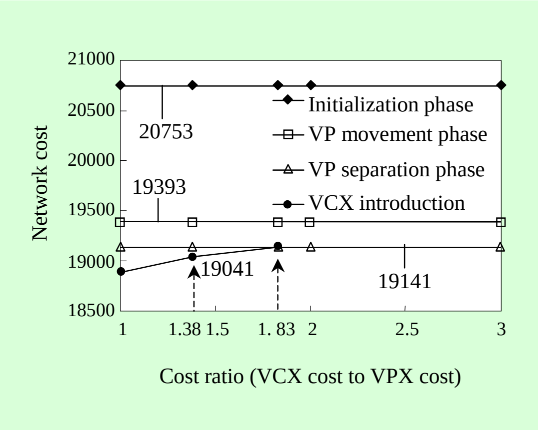


Fig. 8. Network cost dependent on γ (voice and still picture).

VI. CONCLUSIONS

In this paper, we have proposed a heuristic algorithm for designing a VP-based ATM network with multiple traffic classes, which makes possible to achieve an efficient use of network resources under QoS constraints. We tried to minimize the network cost through the alternation of VP route and the separation of single VP into several VPs, and optionally the introduction of VCX nodes after several VPs within some physical lines are statistically multiplexed. In this algorithm, the QoS requirements such as cell loss probability and call blocking probability for various traffic are taken into account as design constraints. For this purpose, we have utilized an equivalent bandwidth concept to determine required bandwidths of VPs. Through numerical examples, we have demonstrated that our design method can facilitate an efficient use of network resources, which results in providing a cost-effective VP-based ATM network.

Our design algorithm proposed in this paper may not provide an optimal solution, but can be a suboptimal one. We should investigate the degree of optimality of our design algorithm and describe the dependency of the proposed algorithm on the network topology as further research topics.

ACKNOWLEDGMENTS

We thank Associate Professor Masayuki Murata and Professor Hideo Miyahara of Osaka University for their invaluable comments in improving this study.

REFERENCES

[1] S. Shioda and H. Uose, "Virtual Path Bandwidth Control Method for ATM Networks: Successive Modification Method," *IEICE Trans. Commun.*, Vol. E74-B, No. 12, Dec. 1991, pp. 4061-4068.

[2] S. Ohta and K. Sato, "Dynamic Bandwidth Control of the Virtual Path in an Asynchronous Transfer Mode Network," *IEEE Trans. Commun.*, Vol. 40, No. 7, July 1992, pp. 1239-1247.

[3] T. Aoyama, I. Tokizawa, and K. Sato, "Introduction Strategy and Technologies for ATM VP-Based Broadband Networks," *IEEE J. Select. Areas Commun.*, Vol. 10, No. 9, Dec. 1992, pp. 1434-1447.

[4] K. Sato, "Transport Network Technology Evaluation: SDH, ATM, and Beyond," *2nd Asia-Pacific Conference on Communications* (*Tutorial*), June 1995.

[5] M. Sugano, M. Murata, and H. Miyahara, "Capacity Assignment Algorithm for Packet Switching Network with Different Grades of Service," *Trans. IEICE*, Vol. J76-B-I, No. 8, Aug. 1993, pp. 581-589, (in Japanese).

[6] M. Yoshida and H. Okazaki, "A Study on ATM Network Planning Based on Evaluation of Design Items," *IEICE Trans. Commun.*, Vol. E76-B, No.11, Nov. 1993, pp. 1333-1340.

[7] S.H. Ahn, R.P. Tsang, S.R, Tong, and D.H.C. Du, "Virtual Path Layout Design on ATM Networks," *Proceedings of IEEE INFOCOM'94*, 1994, pp. 192-200.

[8] I. Chlamtac, A. Farago, and T. Zhang, "Optimizing the System of Virtual Paths," *IEEE/ACM Trans. Networking*, Vol. 2, No. 6, Dec. 1994, pp. 581-586.

[9] S.B. Kim, "An Optimal Establishment of Virtual Path Connections for ATM Networks," *Proceedings of IEEE INFOCOM'95*, 1995, pp. 72-79.

[10] O. Gerstel, I. Cidon, and S. Zaks, "Efficient Support for the Client/Server Paradigm over Heterogeneous ATM Networks," *Proceedings of IEEE INFOCOM'95*, 1996, pp. 1294-1301.

[11] B.H. Ryu, H. Ohsaki, M. Murata, and H. Miyahara, "Design Algorithm for Virtual Path Based ATM Networks," *IEICE Trans. Commun.*, Vol. E79-B, No. 2, Feb. 1996, pp. 97-107.

[12] R. Guerin, H. Ahmadi, and M. Naghshineh, "Equivalent Capacity and its Application to Bandwidth Allocation in High-Speed Networks," *IEEE J. Select. Areas. Commun.*, Vol. 9, No. 7, Sept. 1991, pp. 968-981.

[13] H.Y. Hui, M.B. Gursoy, N. Moayeri, and R.D. Yates, "A Layered Broadband Switching Architecture with Physical or Virtual Path Configurations," *IEEE J. Select. Areas. Commun.*, Vol. 9, No. 9, Dec. 1991, pp. 1416-1426.

[14] G. Kesidis, J. Walrand, and C.S. Chang, "Effective Bandwidths for Multiclass Markov Fluids and Other ATM Sources," *IEEE/ACM Trans. Networking*, Vol. 1, No. 4, Aug. 1993, pp. 424-428.

[15] G.L. Choudhury, D.M. Lucantoni, and W. Whitt, "Squeezing the Most out of ATM," *IEEE Trans. Commun.*, Vol. 44, No. 2, Feb, pp. 203-217. 1996.

Byung Han Ryu received the B.S. and M.S. degrees in industrial engineering from Hanyang University and Seoul National University, Seoul, Korea, in 1985 and 1988, respectively, and the Ph.D degree in information and computer sciences from Osaka University, Osaka, Japan, in 1997. He joined the Switching Technology Division, ETRI, in March 1988, where he was involved in evaluating the performance of the TDX-10 Switching System. He is currently with the Mobile Traffic Management Section, ETRI. Since 1997, he has participated in developing the IMT-2000 (FPLMTS) system. From 1992, he has been a senior member of research staff in ETRI. His research interests include communication network design, mobile traffic engineering in mobile communication network, and performance modeling and evaluation.

Cheol Hye Cho received the B.E. and M.S. degrees in computer science from Hongik University and Taejon University, Korea, in 1982 and 1997, respectively. He joined ETRI in March 1982, where he was involved in developing Software Development Environments for Telecommunication System for TDX-1 and TDX-10 Switching System from 1982 to 1990. He also engaged in developing the CDMA Digital Cellular System from 1991 to 1995. Since 1996, he has participated in developing the IMT-2000 (FPLMTS). Currently he is Head of the Operation Technology Section in ETRI. His special interests include mobile telecommunication service, network management, software engineering and object-oriented methodologies.

Jee Hwan Ahn received the B.S. and M.S. degrees in electronic engineering from Chungnam National University and Seoul National University, Korea, in 1979 and 1981, respectively. He joined ETRI in July 1982, where he was involved in developing the call processing software of the TDX-1 and TDX-10. He participated in the CMS development project since 1992, and developed call processing software of the mobile exchange (CMS-MX) and then call control software of the base station (CMS-BS). Since 1996, he has participated in developing the IMT-2000 system. He is Head of the Mobile Traffic Management Section at ETRI. His research interests include mobile traffic engineering in mobile network and software development of the base station for the IMT-2000.