Journal of Universal Computer Science, vol. 8, no. 7 (2002), 698-721
submitted: 18/11/01, accepted: 17/4/02, appeared: 28/7/02 © J.UCS

Bounded Flooding Routing Algorithm for Provisioning the Globally Optimal Route in a Hierarchical ATM Network

Daniel Won-Kyu Hong
(R&D Group, KT, Korea
wkhong@kt.co.kr)

Choong Seon Hong
(Kyung Hee University, Korea
cshong@khu.ac.kr)

Yoo Jae Hyoung
(R&D Group, KT, Korea
styoo@kt.co.kr)

Dongsik Yun
(R&D Group, KT, Korea
dsyun@kt.co.kr)

Woo-Sung Kim
(R&D Group, KT, Korea
wskim@kt.co.kr)

Abstract: ATM virtual path (VP) contains bundles of virtual channels (VCs). A VP layer network can be used as a server layer network of VC layer networks and each VC layer network can be a client layer. Therefore the effective provision of VC services can be achieved by a better routing scheme of the VP layer network. However, the traditional hierarchical routing scheme of PNNI signaling protocol does not provide the globally optimal route in hierarchical transport network due to its successive network partitioning and topology abstraction. We propose a new VP routing scheme suitable for a nation-wide hierarchical transport network and a network model suitable for the scalable VP network management system. The routing algorithm can provide the globally optimal route in the hierarchical network environment from the perspectives of maximization of network resource utilization and satisfaction of the end user's QoS requirement. In addition, we describe the implementation model of the ATM virtual path network management system (VP-NMS). Lastly, we show the routing performance evaluated in the High Speed Information Network (HSIN) of Korea Telecom.

Categories: C.2.0, C.2.3, C.2.4, G.2.2
Key Words: ATM, Hierarchical Network Management, Hierarchical QoS Routing, Inter-domain Network Management System, ATM VP Optimal Route Provision

1 Introduction

An important concept in ATM networks is the introduction of a virtual path. A virtual path is a logical connection which contains groups of virtual channels that are routed together as a bundled unit. We can obtain many advantages by using the VP concept [ITU731, ATM96a, WONG00, MCM94]. For example, it can reduce the connection set-up delay, simplify the physical topology by making an abstraction of nodes and their related physical links, and can handle smoothly various types of traffic based on their quality of service (QoS) requirements. It can also reduce processing time for routing and the bandwidth control burden in a node [WONG00]. On the other hand, network scalability and survivability are important issues in the telecommunications industry keen on deploying large-scale ATM networks. The scalability is designed with the aim of providing the capabilities of easy network expansion, coping with the constantly changed network environments and achieving the network manageability.

Hierarchical network deployment is one of the popular approaches for providing network scalability, such as PNNI [ATM96b] and OSPF [MOY94]. It is based on network topology abstraction and decomposition of large-scale networks from the perspective of the management domain. ITU-T Rec. G.805 also defines the generic transport network architecture in terms of layering and partitioning [ITU805, SATO90a, SATO90b]. The layering concept can be used to describe the relationship between ATM VC and VP in terms of the client/server layer relationship. Since each VP contains bundles of VCs, a VP layer network can serve as a server layer and each VC layer network can be a client layer. Therefore the provision of effective VC service depends completely on reliability from the perspective of network survivability and optimality from the perspective of routing of the VP server layer [HONG99]. The partitioning concept can be used to deploy a hierarchical transport network based on network topology abstraction and decomposition of the management or administrative domain. It decomposes a specific layer network into smaller management scopes of an subnetwork in order to determine the internal management structure of a layer network, demarcate the administrative boundaries between network operators and routing boundaries within a layer network of a single operator and fault localization domain. Since the partitioning is based intrinsically on the network topology abstraction in order to maximize the network scalability from the perspective of network management, it is naturally used to deploy the hierarchical transport network.

One of the research items involves the study of a dynamic routing algorithm that can provide a globally optimal route in a hierarchical network environment from the perspectives of maximization of network resource utilization and satisfaction of the end user's QoS requirement. We propose a hierarchical ATM transport network model based on the ITU-T Rec. G.805 layering and partitioning concept. In addition, we propose a hierarchical dynamic routing algorithm that can provide the globally optimal route with the propagation of summarized routing information (SRI) in a hierarchical network. This has been thus far impossible because of the successive network topology abstraction in the hierarchical transport network. Lastly we describe the implementation and analyze the performance of the proposed hierarchical routing algorithm followed by concluding remarks.

2 Hierarchical ATM Transport Network

To maximize the network scalability, it is highly necessary to define the common network model that is uniformly applicable to all kinds of networks. We define the generic transport network model with the ITU-T Rec. G.805 layering and partitioning concepts [SATO90a].

The transport network can be decomposed into a number of independent layer networks with a client/server relationship between adjacent layer networks. A layer network describes the generation, transport and termination of a specific characteristic information or protocol. There can be a client/server layer network relationship between the ATM VC and VP because each VP contains bundles of VCs. The client/server layer network relationship is shown in Figure 1.

In order to generally describe the transport network not specific to a vendor's technology and a specific layer network, in this paper, two concepts are defined: network topology and connectivity. The network topology represents the topological structure of a transport network which is composed of layer network (LNW), subnetwork, link, and link termination point (LTP). A layer network describes the generation, transport and termination of particular characteristic information. An subnetwork represents an administrative domain, the management domain or routing and rerouting scope. The subnetwork is further classified in terms of the TMN functional layered architecture: in the element management layer, a subnetwork (emlSNW) can be mapped to the network element, and in the network management layer, a subnetwork (nmlSNW) can be represented by an interconnected group of network elements. An subnetwork can be further partitioned into a number of tinyer subnetworks and links. A link is a logical or physical transmission path interconnecting two network elements or subnetworks and it is supported by a trail in the server layer network. The link termination point represents the near or far end points of a link. There are two kinds of link termination points: one represents the NNI link termination point (nLTP) and the other the UNI link termination point (uLTP).

The connectivity describes the logical or physical connection stream traversing the objects of the network topology which is composed of trail, trail termination point (TTP), subnetwork connection (SNC), link connection (LC) and

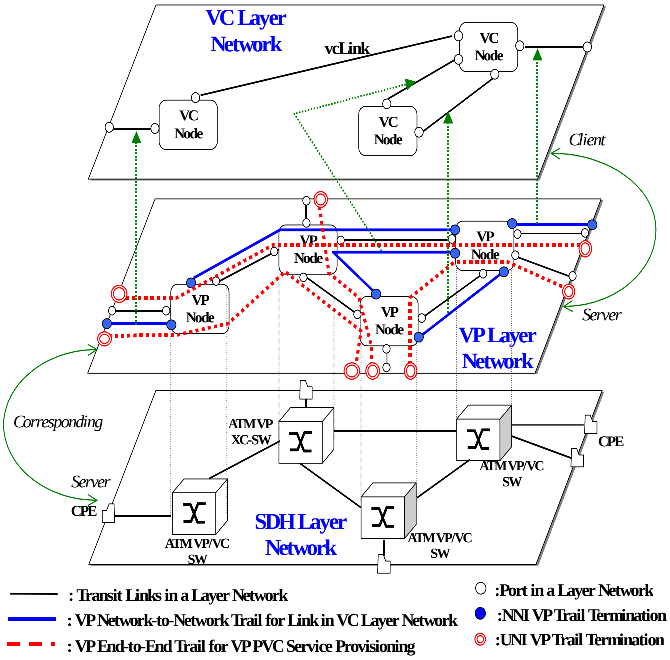

Figure 1: Layered network mode

connection termination point (CTP). A trail represents the logical end-to-end connection within a layer network. The trail termination point describes the near or far ending point of a trail. A link connection is capable of transferring information transparently across a link. It is delimited by link termination points and represents the fixed relationship between the endpoints of the link. An subnetwork connection is capable of transferring information transparently across an subnetwork. It is delimited by connection points at the boundary of the subnetwork and represents the association between these connection points. A connection termination point represents the end point of subnetwork connections or link connections. A trail is composed of a number of subnetwork connection and link connection. The subnetwork connection contained in a trail is composed of a number of smaller subnetwork connections and link connections. In addition, the subnetwork connection that is a part of a trail is composed of a number of smaller subnetwork connections and link connections, as shown in Figure 2.

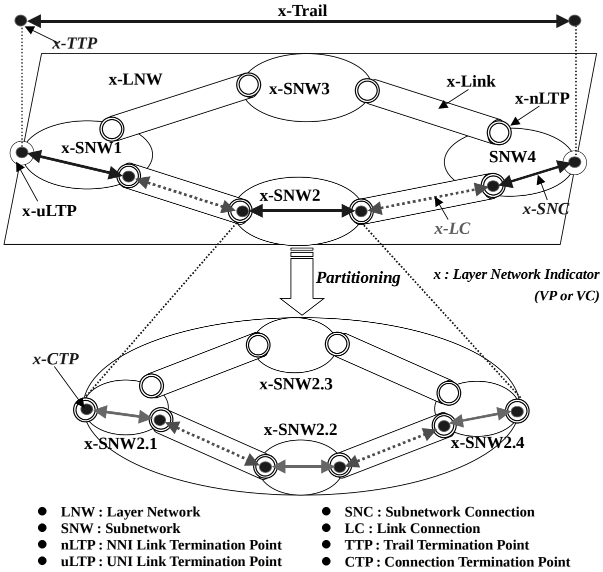

Figure 2: Hierarchical transport network

3 Hierarchical Routing Model

Hierarchical network construction is one of the best approaches for providing network scalability. However, it has a critical disadvantage in that it cannot provide a globally optimal route because of network topology abstraction according to the partitioning concept, as shown in Figure 3.

ATM VP layer network can be partitioned into three subnetworks A, B and C as an example. The partitioned subnetworks have their own management scope and routing policy applied to $A_{PARTITION}$, $B_{PARTITION}$ and $C_{PARTITION}$. Because the route selection is done from a layer network to a partitioned subnetwork, the direct link of L4 between partitioned subnetworks A and B is selected as an optimal one in a layer network which traverses six nodes in the second partitioned level. However, there is a more optimal route than the selected route which traverses five nodes from the perspective of the second partitioned level. In order to select the globally optimal route traversing five nodes at the second partitioned level, another route traversing subnetworks A, C and B should be selected instead of the direct link between A and B at the first partitioned level.

Thus, it is impossible to provide a globally optimal route in a hierarchical transport network. Therefore, we propose a hierarchical routing model that can provide a globally optimal route in a hierarchical VP transport network.

The model is composed of the following five steps, collecting network topol-

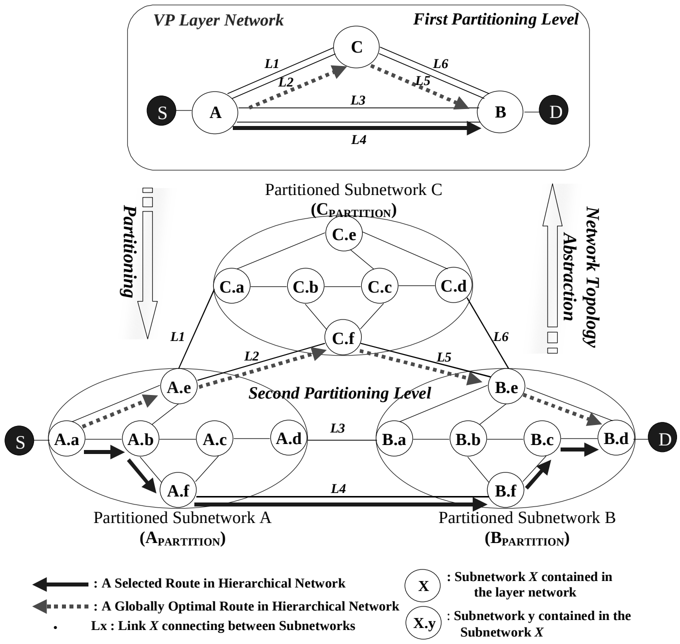

Figure 3: Route selection in the hierarchical transport network

ogy and state parameters, constructing routing table, propagating summarized routing information (SRI) to superior partitioned level, forming a relationship between superior routing information and the propagated SRI from subordinate partitioned level, and selecting an optimal route with connection admission control (CAC). The hierarchical routing information model can be described with Rumbaugh's object model notation [RUMB91], as shown in Figure 4. The RIB maintains the network topology information for constructing the routing table (RT). The RT is composed of a number of *Routes* that have network status information traversing the constructed routes. The *RouteEntry* is the primary information to construct the *Route* which is composed of source/sink link termination point, subnetwork and link.

There is a major difference between the proposed model and the existing hierarchical routing model: the SRI propagation scheme which will be described in sections 3.4 and 3.5. In short, the existing hierarchical routing scheme does not propagate the SRI information to the superior partitioned level, but the proposed hierarchical routing scheme does.

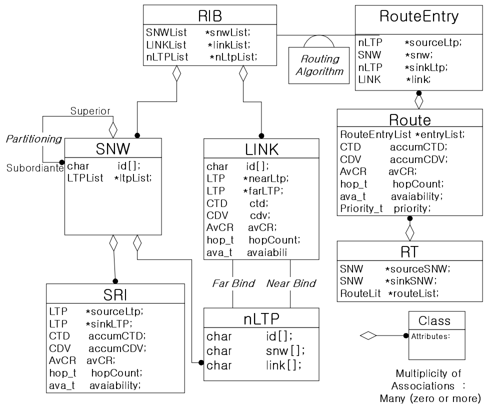

Figure 4: Hierarchical routing information object diagram

3.1 Collecting Network Topology and State Parameters

In this step, we construct the network topology and necessary routing metrics. Network topology is described as a routing information base (RIB) that is composed of subnetwork, link, nLTP and summarized routing information (SRI), as shown in Figure 4. An ATM VP network can be represented as an undirected graph of $RIB(V, E)$, where $Vertex(V)$ represents a set of subnetworks and $Edge(E)$ represents a set of links. $Link_{i,j}$ represents the combination of $Link_{i,j}$ and $Link_{j,i}$. As ATM VP network is partitioned, V of $RIB(V, E)$ at the first partitioned level in Figure 3 has its own partitioned network topology of $RIB'_{Pn}(V'_{Pn}, E'_{Pn})$.

Topology state parameter is a generic term that refers to either a link state parameter or a nodal state parameter. Topology state parameters fall into two categories: topology metric and topology attribute. The topology metric requires the values of the state parameters of all links and nodes (subnetworks) along a given path to be combined to determine whether the path is acceptable and/or desirable for carrying a given connection that satisfies the bounds of cell delay variance (CDV), maximum cell transfer delay (maxCTD) and hop count (HC). The topology attributes are individually considered to determine whether a given link or node (subnetwork) is acceptable and/or desirable for requirements such as CLR, maxCR, avCR and CRM.

There are three performance constraints: link constraints, node constraints, and path constraints. For link constraints, non-additive link-state parameters are used. For node constraints, non-additive nodal state parameters are used. These two constraints correspond to topology attributes. For path constraints, additive link-state parameters are required which belong to topology metrics. Link and node constraints are used to prune the network graph during path selection. For instance, we will prune the links and the nodes based on the link and node constraints, such as CLR and generic CAC parameters (i.e., avCR, and CRM). Once the network topology is trimmed, we will choose a path that meets the path constraints, such as accumulated end-to-end delay (accumulated maxCTD and propagation delay), and end-to-end delay jitter (accumulated CDV).

From the perspective of network topology determination, we consider only the topology attributes, the node and link constraints. Each link has the link-state parameters of CDV, maxCTD, the available bandwidth and the reachability. The reachability refers to the possibility of cell transfer through any node, link or route. If one or more nodes or links are abnormal, the reachability of any route traversing any of these abnormal nodes or links is not allowable.

3.2 Constructing the Routing Table

We construct the routing table (RT) with network topology of $RIB(V, E)$ using the bounded flooding tree construction algorithm (BFTCA) that will be proposed in this section. The BFTCA is used to construct the pre-computed route of bounded flooding tree (BFT) with the moderation of the NP-complexity of the existing flooding algorithm. BFRA can be differentiated from the existing flooding algorithm in two aspects. First, it constructs pre-computed routes without CAC. Second, it prunes the routes exceeding the named bounded depth (Bd) corresponding to hop count.

Because each layer network and the partitioned subnetworks have their own RIBs, each of them constructs its own routing table (RT) with the $RIB(V, E)$ and $RIB'_{Pn}(V'_{Pn}, E'_{Pn})$. There are four routing tables, as shown in Figure 1. One is for the layer network at the first partitioned level and the others are for the partitioned subnetworks A, B and C at the second partitioned level. Though each of them has its own routing policy and network management policy, we apply the BFRA to all of them from the perspective of routing table construction.

In order to construct the bounded flooding tree (BFT) with the $RIB(V, E)$ and $RIB'_{Pn}(V'_{Pn}, E'_{Pn})$, we first decide the source and destination (SD) pairs. There are three SD pairs of (A, B), (A, C) and (B, C) at the first partitioned level in Figure 3. If the number of V at $RIB(V, E)$ and the number of V'_{Pn} at the partitioned $RIB'_{Pn}(V'_{Pn}, E'_{Pn})$ is N, there are $N(N-1)/2$ source and destination pairs.

The BFRA is a very simple algorithm, as shown in Figure 5. It traverses the network topology in a depth-first search manner with the constraint of Bd and stores the traversing traces into RT. The *ComposeRouteEntry()* is used to add a route entry that is composed of an subnetwork and link traversing the RIB. Once the *ComposeRouteEntry()* is called, one route entry is added. It is recursively called until the destination subnetwork is reached or the designated Bd. The

```
class ENTRY {
puclic:
        SNW     *V;     // Vertex (Sub-Network)
        LINK    *E;     // Edge (Link)
        ENTRY *next;
        ENTRY(SNW*,LINK*,ENTRY*);
};

VS   : Source Vertex
VD   : Destination Vertex
Bd   : Bounded Depth

ComposeRouteEntry(RIB, VS, VD, Db, &Entry) {
        VTMP = NULL;
        if(VS == VD ) {
                Entry = new ENTRY(VD, NULL, Entry);
                return;
        }
        find Source Vertex VS from RIB;
        for all Edge E∈ VS {
                VTMP = BranchEntry(Entry, VS);
                if(ValidateCycle(Entry, VTMP) == FALSE) {
                        Entry = new ENTRY(VTMP, E, Entry)
                        ComposeRouteEntry(RIB, VTMP, VD, Bd, Entry);
                } else {
                        VTMP = BranchEntry(Entry, VTMP);
                        if(ValidateBoundedDepth(Entry, Bd) == TRUE)
                                return;
                }
        }
} // end of ComposeRouteEntry()

BranchEntry(&Entry, VS) {
        branch = 0;
        EntryTMP = Entry;
        while(VTMP != NULL) [
                if(VTMP ->V == VS)
                        branch++;    break;
        }
        if(branch != 0) {
                while(VTMP != NULL) [
                        if(Entry->V != VS) {
                                VBEFORE = Entry;
                                Entry = Entry->next;
                                EntryBEFORE = NULL;
                        } else {
                                EntryBEFORE = Entry;
                                Entry = Entry->enxt;
                                delete EntryBEFORE;
                                return;
                        }
                } // end of while
        } // end of if
} // end of BranchEntry()
```

Figure 5: Algorithm for construing the bounded fooding tree

BranchEntry() routine is used to determine the next traversing link in RIB. The *validateCycle()* determines whether there are any previously visited subnetworks or links in the route entry. If there are any previously visited subnetworks or links in the route entry, it traverses the next branch in RIB.

Figure 6 shows the constructed BFT with the partitioned network topology of $RIB'_{Pc}(V'_{Pc}, E'_{Pc})$ in Figure 2. All of the leaf vertexes in BFT have the additive link-state parameters of CDV and maxCTD, the concave metric of bandwidth,

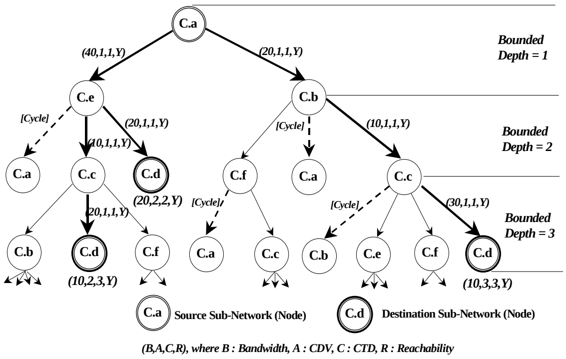

Figure 6: Bounded flooding tree (BFT)

reachability and priority which are used to determine route priority among routes in BFT and to select a path at CAC. Let $td(i,j)$be an accumulated maxCTD and propagation delay and $cdv(i,j)$ be an accumulated CDV from root to leaf vertex. For any route $R = (i,j,k,...,l,m)$, link-state parameters of ctd and cdv and concave bandwidth metric b are like eq. (1).

$$
\begin{aligned}
additive,\ if\ ctd(R) &= ctd_{(i,j)} + ctd_{(j,k)} + ... + ctd_{(k,m)} \\
additive,\ if\ cdv(R) &= cdv_{(i,j)} + cdv_{(j,k)} + ... + cdv_{(k,m)} \\
concave,\ if\ b(R) &= min\{b_{(i,j)}, b_{(j,k)}, ..., b_{(k,m)}\}
\end{aligned} \quad (1)
$$

The reachability refers to the possibility of cell transfer from root to leaf vertex. If one or more nodes or links are abnormal, the reachability of any route traversing one or more of these abnormal nodes or links is not allowable. If all the parameters of CTD, CDV, bandwidth and reachability have been set to the leaf vertex, the BFT is sorted by the following criteria:

- If one wants to provide delay-sensitive VPC, the route being the minimum CTD and CDV will be assigned a higher priority than the other metrics. If there are two or more routes having equal CTD and CDV, the route having the most residual bandwidth will be assigned a higher priority than the others.

- If one wants to concentrate on the maximization of network resource utilization and load balancing, the route with the most bandwidth will be assigned a higher priority than the others. If there are two or more routes having equal bandwidths, the route having the minimum CTD and CDV will be assigned a higher priority than the other metrics.

– If the reachability of the leaf vertex is not allowable, then the least priority is assigned to this route that is excluded in CAC.

3.3 Propagating SRI to Superior Partitioning Level

All of the subnetworks construct their BFT with their network topology, as shown in Figure 6. The first priority route in BFT being determined by its own routing policy is defined as the SRI. The SRI is composed of CTD, CDV, bandwidth, reachability and hop count of the leaf vertex of the highest priority in a BFT. For example, the first priority route between $C.a$ and $C.d$ in Figure 6 traverses $C.a$, $C.e$, $C.c$ and $C.d$. It is selected as the route to represent the SRI between $C.a$ and $C.d$ from the perspective of $C_{PARTITION}$.

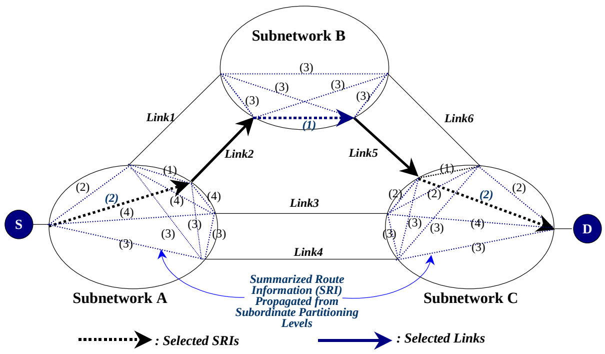

Figure 7: Relationship between superior RIB and subordiante SRI

The SRI is defined to solve the difficulty of finding a globally optimal route in an existing hierarchical transport network. The SRI represents the highly summarized network topology of the subordinate partitioned level from the perspective of the superior partitioned level. There is one SRI per BFT. For example, there are six BFTs at the first PL and 28 BFTs of partitioned subnetwork3 at the second PL in Figure 3. All of the subnetworks except the layer network construct the RIB based on the BFT. The subordinate partitioned levels propagate SRI of the subnetwork to a superior partitioned level of the layer network. The superior partitioned level uses the SRI to determine the internal topology parameters of partitioned subnetworks, as shown in Figure 7.

The network is recursively partitioned until the last partitioned subnetworks such as $A.a$, $B.a$, $C.a$, etc., in Figure 3 are equivalent to the switching elements. The SRI of the last partitioned subnetwork has only the default nodal constraints as follows: the reachability is allowable and hop count is one.

Because the network states are dynamically changed by the transient load fluctuation, the establishment and release of the connections, and the link state change, the BFT priority is adjusted whenever the topology parameters are changed. Whenever the first priority of BFT at the subordinate partitioned level is changed, the SRI is changed in accordance with the change of the first priority route and the changed SRI is propagated to a superior partitioned level of layer network in order to reflect the changed network status of the subordinate partitioned level to a superior partitioned level.

The growing network size makes it increasingly difficult to gather up-to-date information in a dynamic environment. Because the network state gathering is localized within a partitioned subnetwork in a hierarchical network, the overhead for gathering the global network status information can be minimized. For example, if the first partitioned level of the layer network has overall network topology that includes the second partitioned levels in Figure 3, the advantage of scalability should be abandoned. Thus, we propose a mechanism to take advantage of promoting the scalability and narrowing the management domain of a hierarchical network and to provide a globally optimal route with suitable complexity and performance using the SRI propagation.

3.4 Forming a Relationship between Superior Routing Information and Subordinate Routing Information

On receiving SRI from the subordinate partitioned level, we form a relationship between the RIB at the superior partitioned level and the propagated SRIs. The SRI is used to determine the internal structure of partitioned subnetworks at the superior partitioned level from the perspective of routing.

The propagated SRIs are the first priority routes in the BFTs of the subordinate partitioned level. The source or destination of the propagated SRI is a border node. The border node represents the node that terminates the links between subnetworks and the uLTPs. For example, there are four border nodes in the second partitioned level of partitioned subnetwork A in Figure 3. $A.a$ terminates uLTPs. $A.d$, $A.e$, and $A.f$ terminate links between subnetworks. There are six SRIs that are propagated to the first partitioned level and that are used to determine the internal structure of subnetwork A at the first partitioned level. The relationship between a partitioned subnetwork at the first partitioned level and the SRI is shown in Figure 7.

3.5 Selecting Optimal Route

After making forming a relationship between the partitioned subnetwork at the first partitioned level and the SRI propagated from the second partitioned level, the BFTs of the superior partitioned level automatically adjust the CTD, CDV, bandwidth, and reachability for all leaf vertices at their BFTs considering the propagated SRIs such as those in Figure 6. From now on, the optimal routing service can be provided with the BFTs. In Figure 6, we assume that the bandwidth, CTD, and CDV are equal at every node and link and the reachability of all BFTs are allowable for simplicity. If the source is any uLTP contained in $A.a$ and the destination is any uLTP contained in $B.d$, we can select the indirect route traversing subnetworks A, C and B as an optimal route in the first partitioned level. As reflecting the SRI in the process of route selection, the directed link between A and B is not selected which is not a globally optimal route from the perspective of the second partitioned level, as shown in Figure 7. Thus, we can provide a globally optimal route with the SRI propagation mechanism in a hierarchical transport network.

3.6 Differences between Existing and Proposed Hierarchical Routing Schemes

The existing hierarchical routing schemes [ATM96b, TINA97] cannot provide the end-to-end optimal route in the hierarchical transport network because of successive subnetwork partitioning and abstraction. It is composed of three steps: collecting network topology and network status parameters, constructing the routing table based on the gathered network topology and status information, and selecting a route including CAC in line with the requested traffic parameters. However, the proposed hierarchical routing scheme can provide the end-to-end optimal route in the hierarchical transport network with the proposed SRI concept.

4 Implementation and Performance Analysis

We implement the ATM VP-NMS that provides the optimal routing service in a hierarchical transport network with the proposed routing scheme.

4.1 Implementation Model

ATM VP-NMS is composed of four management components of Configuration (*ConfMgr*), Performance (*PerfMgr*), Fault (*FaultMgr*) and Route Managers (*RouteMgr*). ATM VP-NMS is allocated to the layer network and to every partitioned subnetwork except the last partitioned subnetworks that correspond to switching elements. The detailed implementation model of the ATM VP-NMS is depicted in Figure 8. Configuration Manager (*ConfMgr*) maintains the partitioned topology information, periodically collects nodal and link-state parameters from switches and stores them into RIB of *RouteMgr*. Fault Manager (em FaultMgr) receives various alarms from switches, identifies the root cause

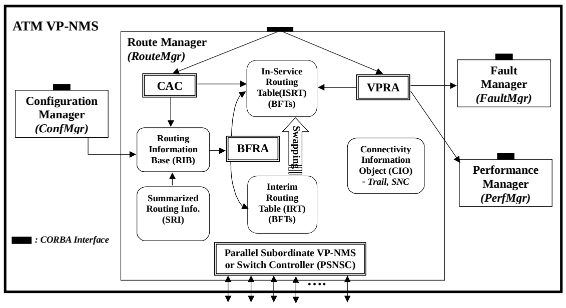

Figure 8: ATM VP-NMS management system components

of alarms by alarm correlation and notifies it to *RouteMgr* to initiate VPC reconfiguration. It also activates or deactivates continuity check (CC) to validate the integrity of VPC connectivity periodically. CC activation is also done at the request of *RouteMgr* when the IP-NMS requests VPC reconfiguration. Performance Manager (*PerfMgr*) periodically collects performance data from switches and stores them in a performance log to trace the performance trends. It also activates or deactivates the threshold crossing alert (TCA) on the ATM VPC to detect the congestion phenomena and to notify it to *RouteMgr* to reconfigure ATM VPC reconfiguration.

Route Manager (*RouteMgr*) plays a key role in provisioning of the globally optimal routing and rerouting paths in the hierarchical transport network. For rapid computation, we store the partitioned network topology or its nodal and link-state parameters in RIB. In the initializing step of the network, we construct the In-Service Routing Table (ISRT) that corresponds to BFT with RIB network topology. If the physical network topology of RIB is changed in the midst of network operation, we should reconstruct the BFT and the VPC provisioning or reconfiguration should be blocked until the BFT reconstruction is completed. For unceasing VPC management service, we construct the Interim Routing Table (IRT) with changed RIB. Of course the RIB maintains the changed topology information. While constructing IRT, ATM VP-NMS can provide continuous VPC management services with ISRT. Immediately we swap ISRT for IRT for reflecting a newly constructed routing table to VPC management service from this point. The ISRT also plays another important role of routing table backup. When the ISRT crashes, it automatically selects the IRT for the unceasing routing service. Therefore ATM VP-NMS provides robust VPC management services.

ATM VP-NMS determines the route priority of BFT with the mixture information of BFT and SRI. After priority determination, it provides VPC provisioning and reconfiguration services with the proposed optimal route selection using GCAC and the reconfiguration algorithm. If an optimal route is selected, ATM VP-NMS creates connectivity information objects of trail and subnetwork connection and saves it into persistent database storage. Parallel subordinate VP-NMS or switch controller (PSNSC) requests the optimal route and reroute provisioning to subordinate VP-NMSs or switches in parallel according to the restoration flags. With the parallel control of subordinate ATM VP-NMSs or switches, ATM VP-NMS maximizes the VPC manipulation performance.

4.2 Network Configuration for Performance Analysis

We partitioned the ATM VP layer network into seven subnetworks A, B, C, E, D, F and G. There are 43 ATM switches and 106 links connecting ATM switches, as shown in Figure 9. We allocate ATM VP-NMS to layer network and all of the partitioned subnetworks. We summarized the RIB of the network in terms of the number of subnetworks, links, source and destination (SD) pair, border subnetworks and SRIs in Table 1.

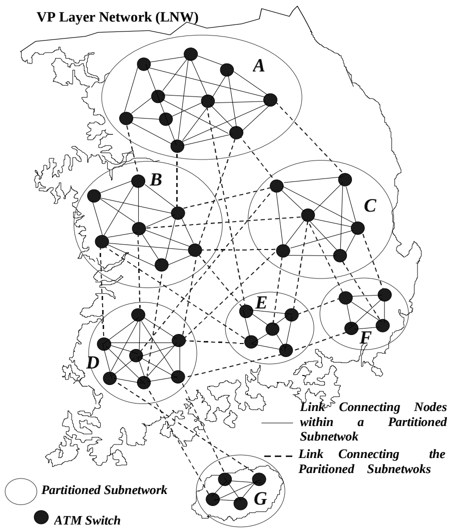

Figure 9: ATM VP-NMS management system components

4.3 Performance Analysis

We evaluate the performance of routing table construction, SIR propagation, adjustment of superior RIB with the propagated SRI, and path selection from the perspective of globally optimal route provisioning. From the perspective of ATM VP PVC provisioning, the most important thing is to find the optimal end-to-end route in the hierarchical transport network. Because the existing hierarchical routing schemes [ATM96b, TINA97] cannot provide the end-to-end optimal route, it subsequently shows poor network resource utilization. However the proposed hierarchical routing scheme can provide the end-to-end optimal route in the hierarchal transport network and it can promote the network resource utilization.

We aim at verifying that the proposed hierarchical routing scheme can provide a globally optimal route without significant performance degradation in comparison with the existing hierarchical routing schemes [ATM96b, TINA97]. There is no doubt that the route selection performance of the existing routing schemes that do not propagate SRI to a superior partitioned level is faster than that of the proposed routing scheme that propagates the SRI to a superior partitioned level. Because this paper focuses on the maximization of network resource utilization and the optimal end-to-end route provisioning in a hierarchical transport network, the route selection performance is not a critical factor. We implement all eight VP-NMSs using the SUN E6500 Base System that is equipped with four 333MHz CPUs with 8MB cache and 4GB main memory. We use the Oracle database to support the information persistency of the RIB and routing table.

First we compare the performance of the two routing schemes. One is the proposed routing scheme that propagates SRI to a superior partitioned level and the other is the routing scheme that does not. The former can provide the globally optimal route but the latter cannot. We assign the Bd to four which trims the route exceeding four hop counts.

Figure 10 shows the performances of BFT construction and SRI propagation under the network topology of Figure 8. The BFT construction is involved in the existing and proposed hierarchical routing schemes simultaneously. However the SRI propagation and reflection of the propagated SRI into the internal network

Table 1: RIB and VP-NMS allocation

RIB	LNW	SNW A	SNW B	SNW C	SNW D	SNW E	SNW F	SNW G
SNW	7	10	7	6	7	5	4	4
Link	29	24	16	12	15	8	6	6
SD	21	45	21	15	21	10	6	6
Border SNW	-	7	6	4	6	5	4	4
SRI	-	21	15	6	15	10	6	6
VP-NMS	NMS 1	NMS 2	NMS 3	NMS 4	NMS 5	NMS 6	NMS 7	NMS 8

topology of the partitioned subnetwork are only parts of the proposed routing scheme. Therefore the performance of the routing table construction of the existing routing schemes is high. On the other hand, the proposed routing scheme

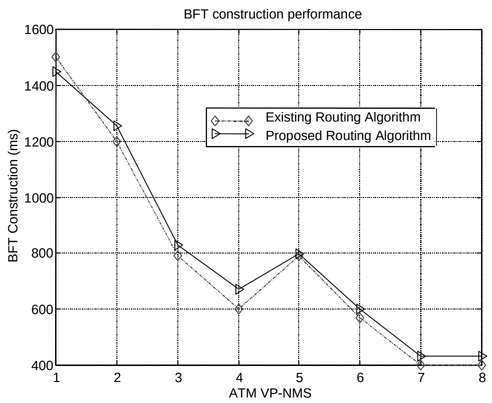

Figure 10: Performance for constructing the bounded flooding tree

is somewhat inferior to the existing schemes. However, there was some performance degradation when the VP-NMS associated the partitioned subnetwork internal routing structure with the propagated SRI from subordinate VP-NMSs. It took 1,450ms to construct BFT, propagate SRI to superior partitioned level and make an association between the propagated SRI and the internal structure of the layer network. However the primary goal of the proposed routing algorithm lies in providing the globally optimal route in the hierarchical transport network while maximizing network resource utilization. Therefore the proposed hierarchical routing scheme can compensate the performance degradation of the routing table construction with the maximization of network resource utilization with globally optimal route provision. Because every partitioned subnetwork in Figure 8 manage nearly the same number of nodes and links in terms of network topology complexity, the performances of BFT construction and SRI propagation of other ATM-NMSs except for ATM-NMS1 are nearly the same. Figure 11 shows the performance of BFT construction and SRI propagation from the perspective of variable Bd. It took 2,340ms when the Bd was 3, 2,500ms when Bd was 4, 2,878ms when Bd was 5, and 3,507ms when Bd was 6. The Bd can be adjustable according to the routing policies of each partitioned subnetwork taking into account the end-to-end transit delay. As a result of the performance analysis of BFT construction and SRI propagation with the Bd constraint, the routing table construction performance depends entirely on the number of Bd's.

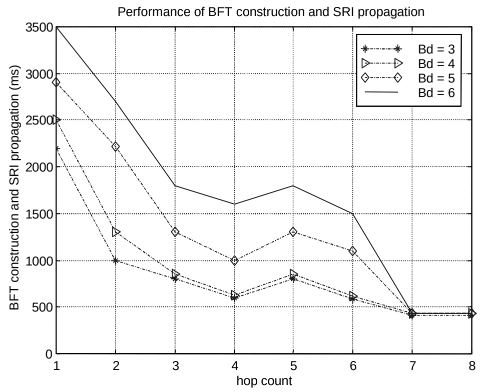

Figure 11: Performance for BFT construction and SRI propagation with the bounded depth constraint

When the routing table construction time is increased as the Bd is increased, the possible routing paths are also increased. The number of possible routing paths is important from the perspective of rerouting because it directly affects the alternative route selection at the restoration. As a result of our performance analysis of the proposed routing scheme under the High Speed Information Network (HSIN), the reasonable Bd can be four or five. If the target network is changed, the Bd can be adjusted according to the complexity of network topology and routing policy.

There exist other performance measurements of route selection and connection admission control (CAC) performed after completion of the BFT construction and SRI propagation. The route selection and CAC performance methods are shown in Figure 11. The route selection is composed of two major procedures: (1) sorting the BFT with the route selection criteria and (2) performing connection admission control along the selected route. Because the number of branches in BFT depends fully on the network topology complexity and the designated Bd, most of the time is spent in sorting the BFT in the processes of route selection and CAC. Figure 12 shows that as the Bd increased, more time was spent in selecting a route than CAC. Because Bd directly affects the end-to-end transit delay, we normally set Bd in the VP network to three or four. When Bd was four, it took 68ms to select a route and perform CAC which is the most reasonable performance in ATM VP PVC provisioning by the network management system. Until now, we have only discussed the performance of the BFRA under the 2-level hierarchy as shown in Figure 11. In order to show its performance from the perspective of scalability, BFRA performance under a 4-level hierarchy

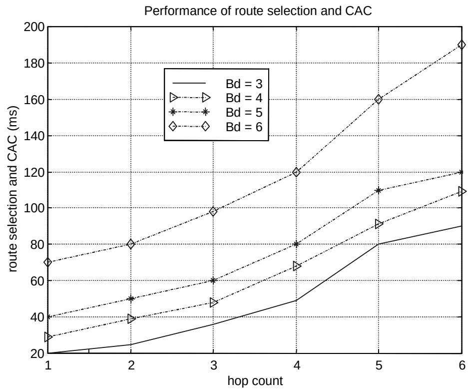

Figure 12: Performance for route selection and CAC with the bounded depth constraint

must also be evaluated. We assume that there are four partitioning levels: the top, first, second, and third. There are twelve third-partitioned subnetworks. Each of them manages 100 ATM switches, which makes a total of 1,200 ATM switches that are connected in full mesh.

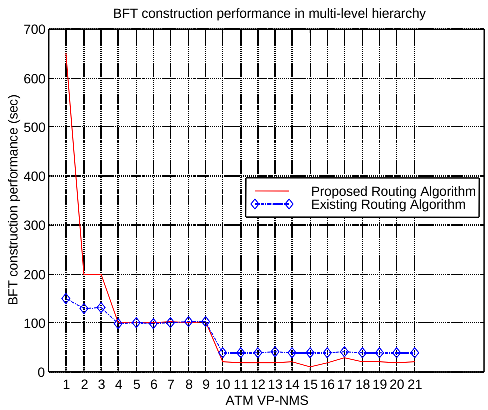

Figure 13: Performance in constructing the bounded flooding tree under a 4-level hierarchy

There are six second-partitioned subnetworks containing two third-partitioned subnetworks, and one top layer network containing six second-partitioned subnetworks. Logically therefore, we would need twenty-one ATM-NMSs for each subnetwork. Under this complex network topology and 4-level hierarchy, we compare the performance of the BFT construction with the existing routing schemes [ATM96b, MOY94], as shown in Figure 12. In addition, we evaluate the performance of the BFT construction and SRI propagation with the Bd constraint, as shown in Figure 14. Figure 13 shows the performance of the BFT construc-

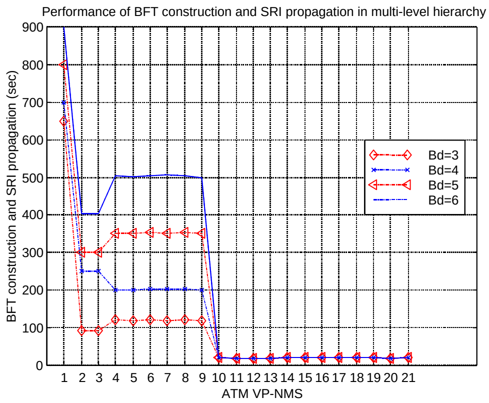

Figure 14: Performance of the BFT construction and SRI propagation with the bounded depth constraint under a 4-level hierarchy

tion under a 4-level hierarchy when the Bd is four. Our findings reveal that the algorithm under a 4-level hierarchy performs higher than the existing algorithms under a hierarchy level of one. On the other hand, its performance nearly equals the existing algorithms under a hierarchy level of two. However, its performance degrades significantly than the existing algorithms under a hierarchy level greater than three. This performance degradation mainly caused by the SRI propagation and the association between the propagated SRI from the subordinate partitioned level and the internal structure of the superior partitioned level.

Under a 3-level hierarchy, the top network management system, ATM VP-NMS1, took approximately 650 seconds. However, the existing algorithm was only at 160 seconds. The result of this performance analysis shows our algorithm with an approximate 30 percent and 75 percent performance degradation compared with that of the existing algorithm under a hierarchy level of three and four, respectively. Therefore, our algorithm exhibits the most reasonable performance when the hierarchy level is two and the most significant performance degradation at a hierarchy level larger than two.

Figure 14 shows the performance of the BFT construction and SRI propagation with the variable Bd constraint under 4-level hierarchy and complex network topology. From the perspective of the top partitioned level, ATM VP-NMS1, it took 650 seconds when the Bd is three, 700 seconds when it is four, 800 seconds when five and 890 seconds when six. Unfortunately, it shows significant performance degradation in direct proportion to the complexity of network topology and hierarchy level. Fortunately because Bd directly affects the end-to-end transit delay of VP PVC, we normally set Bd in the VP network to three or four. If we set Bd to three under a complex 4-level hierarchy, the performance of 650 seconds could be endurable from the perspective of the VP PVC management, not the VP SVC management.

In addition, the size of the resulting routing tables is of interest. The size of resulting routing tables is gradually increased in proportion to the complexity of network topology and the assigned bounded depth (Bd). As we propose the bounded depth for moderating the NP-complexity problem of flooding algorithm, the size of resulting routing tables can be also moderated in proportion to the assigned number of bounded depth. What we can gain from the proposed

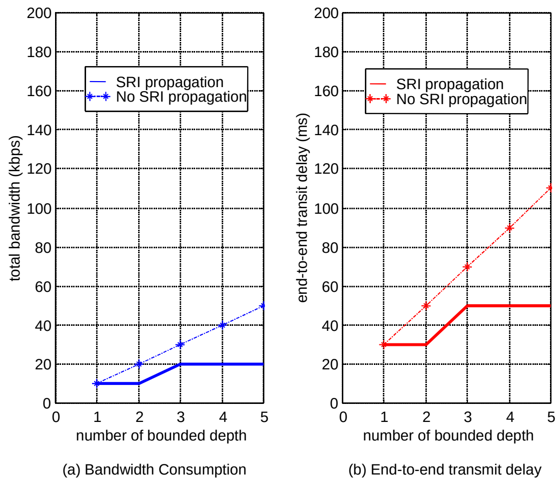

Figure 15: Gains from the bounded flooding routing algorithm in terms of bandwidth and end-to-end transit delay

routing algorithm is to provide the globally optimal route in hierarchical network. Though the optimal route simply implies the shortest path, it is most important taking into account the overall network resource utilization and end-to-end transit delay. Let's assume that the transit delay of each node and link is 10ms and the requested bandwidth is 100kbps. For this, if we select the route traversing $A.a$-$A.b$-$A.f$-$B.f$-$B.c$-$B.d$ in Figure 2, the total bandwidth for accommodate the route is 500kbps and the end-to-end transit delay is 110ms. If we select the globally optimal route traversing $A.a$-$A.e$-$C.f$-$B.e$-$B.d$ with our algorithm in Figure 2, the total bandwidth for accommodate the route is 400kbps and the end-to-end transit delay is 900ms.

Under the complex network topology composed of 43 ATM nodes and 106 links in Figure 9, we measured the gains in terms of required bandwidth and end-to-end transit delay comparing our algorithm propagating SRI with the algorithm does not propagating SRI, as shown in Figure 15. The gains from the proposed routing algorithm in terms of required bandwidth and end-to-end transit delay are increased in proportion to the number of bounded depth (Bd).

5 Conclusions

After designing a hierarchical ATM transport network model based on the ITU-T G.805 layering and partitioning concept, we proposed a hierarchical dynamic routing algorithm that can find the globally optimal route with the propagation of SRI in the hierarchical network. This has been thus far impossible due to the network topology abstraction and aggregation. In addition, we described the implementation procedure and analyzed the performance of the proposed hierarchical routing and rerouting algorithms. We implemented the ATM VP-NMS, adopting the proposed dynamic routing and rerouting schemes in a hierarchical VP network. We also evaluated the routing and rerouting performance in the High Speed Information Network (HSIN) of Korea Telecom.

The performance evaluation in the real network environment of the High Speed Information Network (HSIN) of Korea Telecom showed that there was no problem in applying the proposed routing scheme to a large-scale ATM virtual path network management. Even though the proposed routing scheme shows slight performance degradation compared with that of existing routing schemes, the major aim of the proposed routing scheme lies in finding the globally optimal route in hierarchical transport network while maintaining highly efficient network resource utilization.

References

[ADAM88] J.L. Adams, "The virtual path identifier and its applications for routing and priority of connectionless and connection-oriented services," International Journal Digital and Analog Cabled Systems 1, 257-262, 1988.

[ATM96a] ATM Forum Technical Committee, "Traffic Management Specification Versions 4.0," at-nm-0056.000, April 1996.

[ATM96b] ATM Forum Technical Committee, "Private Network-Network Interface Specification Version 1.0," af-pnni-0055.000, March 1996.

[HONG99] W.K. HONG, D.S. YUN, "Distributed Connection Management Architecture for Optimal VPC Provisioning on Hierarchical ATM Transport Network," DSOM'99, 63-75, October 1999.

[ITU731] ITU-T Recommendation I.371, "Traffic Control and Congestion Control in B-ISDN," March 1993.

[ITU805] ITU-T Recommendation G.805, "Generic Function Architecture Of Transport Networks," November 1995.

[MCM94] J. C. McDonald, "Public network integrity - avoiding a crisis in trust," IEEE Journal on Selected Areas in Communications, vol. 12, no. 1, 5-12, January 1994.

[MOY94] J. Moy, OSPF Version 2, Ascend Communications, Inc., April 1998.

[RUMB91] James Rumbaugh, Michael Blaha, William Premerlani, Frederick Eddy and William Lorensen, "Object-Oriendted Modeling and Design," Prentice-Hall. Inc., 1991.

[SATO90a] K.I. Sato and I. Tokizawa, "Flexible asynchronous transfer mode networks utilizing virtual paths," GLOBCOMM'90, 831-838, 1990.

[SATO90b] K.I. Sato and I. Tokizawa, "Broad-band ATM network architecture based on virtual paths," IEEE Trans. Communications., vol.38, 1212-1222, 1990

[TINA97] TINA-C Deliverable, "Network Resource Architecture Version 3.0," February 1997.

[WONG00] Eric W.M. WONG, Andy K.M. CHAN, Sammy CHAN, and King-Tim KO, "Bandwidth Allocation for Virtual Paths in ATM Networks with Dynamic Routing," IEICE Trans.Communications, Vol.E83-B, No.3, March 2000.

List of Abbreviations

ATM	Asynchronous Transfer Mode
Bd	Bounded Depth
BFT	Bounded Flooding Tree
BFTCA	BFT Construction Algorithm
CAC	Connection Admission Control
CDV	Cell Delay Variance
CLR	Cell Loss Ratio
CRM	Cell Rate Margin
CTD	Cell Transfer Delay
CTP	Connection Termination Point
EML	Element Management Layer
GCAC	Generic Connection Admission Control
HC	Hop Count
HSIN	High Speed Information Network
IRT	Interim Routing Table
ISRT	In-Service Routing Table
LC	Link Connection
LNW	Layer Network
LTP	Link Termination Point
nLTP	NNI Link Termination Point
NML	Network Management Layer
NNI	Network-Network Interface
OSPF	Open Shortest Path First
RIB	Routing Information Base
PL	Partitioning Level
PNNI	Private Network-Network Interface
RT	Routing Table
QoS	Quality of Service
SD	Source and Destination
SNC	Subnetwork Connection
SNW	Subnetwork
SRI	Summarized Routing Information
TCA	Threshold Crossing Alert
TTP	Trail Termination Point
VC	Virtual Channel
VP	Virtual Path
VPC	Virtual Path Connection
uLTP	UNI Link Termination Point
UNI	User-Network Interface