On Topological Design of Service Overlay Networks

Arunabha Sen[1], Ling Zhou[1], Bin Hao[1], Bao Hong Shen[1], and Samrat Ganguly[2]

[1] Dept. of Computer Science, Arizona State University, Tempe, AZ 85287, USA
{asen,ling.zhou,binhao,bao}@asu.edu
[2] Dept. of Broadband & Mobile Networks, NEC Laboratories, USA
samrat@nec-lab.com

Abstract. The notion of *service overlay network* (SON) was proposed recently to alleviate difficulties encountered in providing end-to-end *quality of service* (QoS) guarantees in the current Internet architecture. The SONs are able to provide QoS guarantees by purchasing bandwidth from individual network domains and building a logical end-to-end data delivery infrastructure on top of existing Internet. In this paper, we consider a generalized framework for SON, which is categorized based on three different characteristics: a) single-homed/multi-homed end-system b) usage-based/leased cost model and c) capacitated/uncapacitated network. We focus on the algorithmic analysis of the topology design problem for the above generalized SON. We prove that for certain case, polynomial-time optimal algorithm exists, while for other cases, the topology design problem is NP-complete. For the NP-complete cases, we provide approximation algorithms and experimental results.

1 Introduction

The Internet today comprises multiple independently operated networks (autonomous systems or domains) joined at the peering points. The independently operated networks (often Internet Service Providers, ISPs) may have an interest in providing QoS guarantees within their own network, but they do not have any incentive to provide service guarantees to customers of other remote ISPs. The notion of *service overlay network* (SON) was proposed in [3] to overcome this problem, so that end-to-end guarantees can be provided to the customers of different ISPs. Service overlay network is an outcome of the recent studies on overlay networks such as Detour [10], Resilient Overlay Network [1] and Internet Indirection Infrastructure [11].

The SONs are able to provide end-to-end QoS guarantees by building a logical delivery infrastructure on top of the existing transport network by purchasing bandwidth from individual network domains. The SONs provide various flexibilities in deploying and supporting new services by allowing the creation of service-specific overlay network without incorporating changes in the underlying network infrastructure. This mechanism can be utilized to support applications

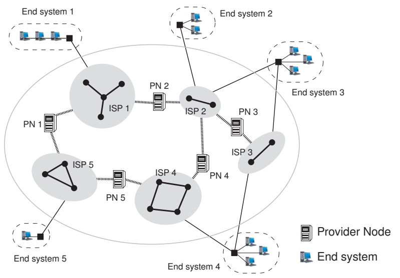

Fig. 1. Service Overlay Network Model

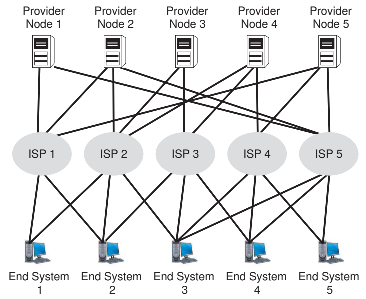

Fig. 2. Relationship between Provider Nodes, End-systems and ISPs

for fault-tolerance, multi-cast communication, security, file sharing and QoS [1, 2, 7].

We consider the SON model described in [8], where it is constructed on top of an infrastructure of ISPs and is capable of providing QoS guarantees to a set of customers. Because of this capability, the network is referred to as a QoS Provider Network or *Provider Network*. The provider network comprises a collection of *provider nodes*, and a set of customers referred to as the *end-systems or enterprises*. The provider nodes and the end-systems gain access to the Internet through ISPs. An illustration of the SON is given in Figure 1. The provider nodes are connected to each other through ISPs and the end-systems are also connected to the provider nodes through ISPs. Two provider nodes are said to be connected to each other, if they are connected to the same ISP. Similarly, an end-system is said to be connected to a provider node if they are connected to the same ISP. Figure 2 illustrates the relationships between provider nodes, ISPs and end-systems. The provider node buys services (guaranteed bandwidth) from ISPs and sells them to the end-systems with end-to-end service guarantees. Currently, there exists at least one commercial service overlay network (Internap [6]) that closely resembles the model used in this paper as well as in [8].

The topology design problem of a SON can be described as follows: Given a set of end-systems, provider nodes, access cost of traffic from an end-system to a provider node, transport cost of traffic among provider nodes, traffic demand for each pair of end-systems, find the least cost design that satisfies the traffic bandwidth demand between each pair of end-systems. Our work is motivated by the recent study done by Vieira *et.al.* [8] on topology design problem for a specific SON model. In this paper, we introduce a generalized framework for SON, which provides a comprehensive view of the overall topology design space. We categorize the generalized SON model based on the following scenarios:

– **Single-homed vs multi-homed:** The term *multihoming* is generally used to indicate that an end-system is connected to multiple ISPs [12]. In the context of SON, we extend this notion and let *multihoming* refer to the scenario where one end-system can be connected to multiple *provider nodes*

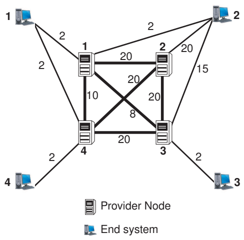

Fig. 3. Service Overlay Network Model

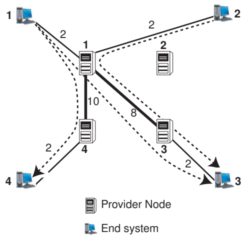

Fig. 4. Single-homed Solution for SON Design

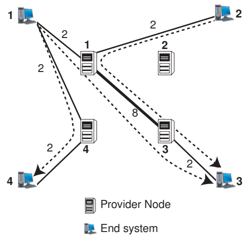

Fig. 5. Multi-homed Solution for SON Design

instead of multiple ISPs. In a multi-homed environment, an end-system has more flexibility in connecting to a set of provider nodes. This flexibility enables the designer to find a lower cost solution. Figures 4 and 5 show the solution of the same problem in single-homed and multi-homed scenarios, where the cost of the single-homed design is 26 and that of the multi-homed is 18.

– **Usage-based vs leased(fixed) cost model:** In the usage-based cost model, the cost of the link is proportional to the volume of data sent through the link. In a leased or fixed cost model, we assume that each link has an associated cost that is independent of the traffic sent through it. Such fixed cost scenario is often applicable to enterprises who buy leased lines from ISPs at a flat rate.
– **Capacitated vs uncapacitated network:** In case of a capacitated network, we assume that any link in the SON has a capacity bound that cannot be exceeded. While in an uncapacitated case, there exist no such constraints.

It may be noted that the authors in [8] provide solution only for the single-homed, uncapacitated network with usage-based cost model. In this paper, we provide results of our comprehensive study of the SON design problem. The key contributions of this paper are as follows:

– We prove that the SON topology design problem with (a) multi-homed enterprise, (b) usage-based cost model and (c) uncapacitated network can be solved in *polynomial* time.
– We prove that the SON topology design problem with (a) single-homed enterprise, (b) usage-based cost model and (c) uncapacitated network is NP-Complete.
– We prove that the SON topology design problem with (a) single-homed/multi-homed enterprise and (b) fixed cost model is NP-Complete in both capacitated and uncapacitated network scenarios. We present approximation algorithms for the solution of uncapacitated version of these problems.
– We show that all the four problems in the capacitated version of the SON design problem are NP-Complete.

A summary of the complexities involved in the topology design problem for the various cases is shown in Table 1.

Table 1. Complexity results for different versions of SON design problem

Cost Model	Uncapacitated Network		Capacitated Network	
	Single-homed	Multi-homed	Single-homed	Multi-homed
Usage-based cost	NPC	Poly. Solution	NPC	NPC
Fixed cost	NPC	NPC	NPC	NPC

Table 2. Basic Notations

ES_i	end-system i
PN_j	Provider node i
M	Number of end-systems
N	Number of provider nodes
α_{ij}	Access cost (per unit of reserved bandwidth) for traffic from ES_i to PN_j
α	Access cost matrix for traffic from all ES_i to all PN_j
l_{ij}	Transport cost (per unit of reserved bandwidth) for traffic on the transport link from PN_i to PN_j
L	Transport cost matrix for traffic on the transport link from all PN_i to all PN_j
b_{ij}	Cost of least-cost route (per unit of reserved bandwidth) for traffic between PN_i to PN_j
B	Cost of least-cost route matrix for traffic between all PN_i to all PN_j
ω_{ij}	Reserved bandwidth for traffic from ES_i to ES_j
Ω	Reserved bandwidth matrix for traffic from all ES_i to all ES_j

2 Problem Formulation

The optimal topology design problem of a SON is described in the previous section. We consider different versions of the problem based on different application environments: (i) *single-homed* or *multi-homed* end-system, (ii) *usage-based* or *fixed* cost model [9], and (iii) *finite* or *infinite* capacity links. The notations used in this paper are same as in [8] and are given in Table 2.

The *access cost* α_{ij} of an *access link* connecting an end-system ES_i to a provider node PN_j refers to the cost of transmitting one unit of data over that link in usage-based cost model and the cost of transmitting any number of units of data in fixed cost model. In case ES_i can be connected to PN_j through more than one ISP, α_{ij} represents the cheapest way of connecting ES_i to PN_j. If ES_i cannot reach PN_j through any ISP, access cost $\alpha_{ij} = \infty$. The *transport cost* l_{ij} of a *transport link* connecting PN_i to PN_j refers to the cost of transmitting one unit of data over that link in usage-based cost model and the cost of transmitting any number of units of data in fixed cost model. In case PN_i can be connected to PN_j through more than one ISP, l_{ij} represents the cheapest way of connecting PN_i to PN_j. If PN_i cannot reach PN_j through any ISP, transport cost $l_{ij} = \infty$.

From the set of input data, we construct a graph $G_{ESPN} = (V_{ESPN}, E_{ESPN})$, where the vertex set V_{ESPN} consists of two different types of nodes, V_{PN} and V_{ES}, representing the provider nodes and the end-systems respectively. Similarly, the edge set E_{ESPN} consists of two different types of edges, $E_{PN,PN}$ and $E_{ES,PN}$. For any $v_i, v_j \in V_{PN}$, there is an edge in $E_{PN,PN}$ connecting them with an associated weight l_{ij}. l_{ij} values for all pairs of provider nodes are denoted by matrix L. The length of the shortest path between v_i and v_j is denoted by b_{ij}. b_{ij} values for all pairs of provider nodes are denoted by matrix B. For any $v_i \in V_{ES}$ and $v_j \in V_{PN}$, there is an edge in $E_{ES,PN}$ connecting them with an associated weight α_{ij}. α_{ij} values for all end-system to provider node pairs are denoted by matrix α. For any $v_i, v_j \in V_{ES}$, there is a traffic demand ω_{ij} associated with this ordered pair of nodes (ω_{ij} may be zero). Traffic demands for all pairs of end-systems are denoted by matrix Ω. An illustration of such a graph is shown in Figure 3. In this example, there is a non-zero traffic demand for the pairs (ES_1, ES_3), (ES_1, ES_4) and (ES_2, ES_3). In the fixed cost model, the actual bandwidth request by each pair is not relevant. The optimal solutions for the single-homed and multi-homed versions of the SON design problem are shown in Figures 4 and 5 respectively. The optimal cost of the single-homed version is 26, whereas the multi-homed version is 18.

The main difference between usage-based model and fixed cost model is how the access and transport costs are calculated, especially when the same edge appears on more than one path between end-system pairs. For example, if an edge $e_{pq} \in E_{ES,PN}$ is used for transferring $\omega_{i,j}$ amount of data from ES_i to ES_j and also used for transferring $\omega_{i,k}$ amount of data from ES_i to ES_k, then the cost of using this link will be $\alpha_{pq}(\omega_{i,j} + \omega_{i,k})$ in usage-based cost model and only α_{pq} in the fixed cost model. Similarly, If an edge $e_{pq} \in E_{PN,PN}$ is used for transferring $\omega_{i,j}$ amount of data from ES_i to ES_j and $\omega_{r,s}$ amount of data from ES_r to ES_s, then the cost of using this link will be $l_{pq}(\omega_{i,j} + \omega_{r,s})$ in usage-based cost model and only l_{pq} in the fixed cost model.

3 SON Topology Design - Algorithms and Complexities

We consider eight different versions of the SON topology design problem. We show that only one of the four different versions with uncapaciated network model is polynomial-time solvable, and the other three are NP-complete. Since uncapacitated version of the problem is just a special case of the capacitated version, the NP-Completeness of the capacitated version will follow from the NP-Completeness of the uncapacitated version. The complexity results of various versions are summarized in Table 1.

3.1 SON Design Problem with Multi-homed Enterprise, Uncapacitated Network and Usage-based Cost Model (SONDP-MHE/UN/UBC)

Instance: Given a graph $G_{ESPN} = (V_{ESPN}, E_{ESPN})$ with matrices L, α, Ω, and a positive integer K.
Question: Is it possible to construct a SON topology with total cost less than or equal to K so that all traffic demands given in matrix Ω are satisfied?

Theorem 1. *SON design problem with MHE/UN/UBC can be solved in polynomial time.*

Proof: From the special properties of this problem, it is no hard to see that the cost of establishing a path to transmit w_{ij} units of data from ES_i to ES_j is independent of the cost of establishing paths for other pairs. Therefore, we can minimize the total cost by establishing a shortest path for each pair of end-systems separately in G_{ESPN}, and thus obtain the optimal solution for this problem.

The computation complexity of this algorithm is $O(k(|V_{ESPN}|\ log\ |V_{ESPN}|+|E_{ESPN}|))$, where k is the number of end-systems pairs that need to transfer data between each other.

3.2 SON Design Problem with Single-homed Enterprise, Uncapacitated Network and Usage-based Cost Model (SONDP-SHE/UN/UBC)

From the transport cost matrix L (in Table 2), we compute the least-cost route matrix B. The problem instance is described in terms of matrix B.

Instance: Given a graph $G_{ESPN} = (V_{ESPN}, E_{ESPN})$ with matrices B, α, Ω, and a positive integer K.
Question: Is it possible to construct a SON topology with total cost less than or equal to K so that all traffic demands given in matrix Ω are satisfied, and meanwhile each end-system is connected to only one provider node?

Theorem 2. *SON design problem with SHE/UN/UBC is NP-complete.*

Proof: We can restate the question more formally in the following way:
Question: Is there a function $g : \{1, 2, ..., M\} \rightarrow \{1, 2, ..., N\}$, such that

$$\sum_{i=1}^{M}\sum_{j=1}^{M} w_{ij}(\alpha_{ig(i)} + b_{g(i)g(j)} + \alpha_{jg(j)}) \leq K? \qquad (1)$$

Clearly SONDP-SHE/UN/UBC is in NP. We prove its NP-Completeness by reduction from the Matrix Cover problem [4]. The reduction maps a Matrix Cover instance (an $n \times n$ matrix $A = \{a_{ij}\}$, K) to a SONDP-SHE/UN/UBC instance $(\Omega, B, \alpha, K^{'})$, so that there is a function $f : \{1, 2, ..., n\} \rightarrow \{-1, +1\}$ such that $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}f(i)f(j) \leq K$ if and only if there is a function g : $\{1, 2, ..., M\} \rightarrow \{1, 2, ..., N\}$ such that $\sum_{i=1}^{M}\sum_{j=1}^{M} w_{ij}(\alpha_{ig(i)}+b_{g(i)g(j)}+\alpha_{jg(j)}) \leq K^{'}$. Given any instance of Matrix Cover: An $n \times n$ matrix $A = \{a_{ij}\}$ with nonnegative integer entries, and an integer K, we construct the instance for SONDP-SHE/UN/UBC problem as follows:

1. Let $M = n$. For the $M \times M$ bandwidth reservation matrix $\Omega = \{w_{ij}\}$, $\forall 1 \leq i \leq M, 1 \leq j \leq M$, let $w_{ij} = 1$.

2. Let $N = 2M$. For the $N \times N$ transport matrix $B = \{b_{ij}\}$, $\forall 1 \le k \le M, 1 \le l \le M$, let

$$\begin{aligned} b_{2k,2l} &= \frac{a_{kl}+a_{lk}}{2} + max; \qquad & b_{2k-1,2l-1} &= \frac{a_{kl}+a_{lk}}{2} + max; \\ b_{2k,2l-1} &= -\frac{a_{kl}+a_{lk}}{2} + max; \qquad & b_{2k-1,2l} &= -\frac{a_{kl}+a_{lk}}{2} + max; \end{aligned} \tag{2}$$

where max is the maximum element of matrix A in the instance of Matrix Cover. It is added to make sure that b_{ij} is nonnegative.

3. For the $M \times N$ Access Cost matrix $\alpha = \{\alpha_{ij}\}$, $\forall 1 \le i \le M$

$$\alpha_{ij} = \begin{cases} 0 & \text{if } j = 2i-1, \text{ or } 2i \\ \infty & \text{otherwise} \end{cases} \tag{3}$$

4. Let $K' = K + M^2 \cdot max = K + n^2 \cdot max$.

The construction can be done in polynomial time. To complete the proof, we show that this transformation is a reduction. Suppose for the instance of matrix cover problem, there is a function $f : \{1, 2, .., n\} \to \{-1, +1\}$ such that $\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} f(i) f(j) \le K$, then we can build the corresponding $g : \{1, 2, .., M\} \to \{1, 2, ..., N\}$ for the instance of SONDP-SHE/UN/UBC as follows:

$$\begin{aligned} &\forall 1 \le i \le M (M = n, N = 2M) \\ &g(i) = \begin{cases} 2i & \text{if } f(i) = +1 \\ 2i-1 & \text{if } f(i) = -1 \end{cases} \end{aligned} \tag{4}$$

Due to the process of construction, there exists a relationship between the objective functions of the two problems, as shown in the following table. Therefore,

Table 3. Relationship between two objective functions

$f(i)$	$f(j)$	$a_{ij}f(i)f(j)$	$a_{ji}f(j)f(i)$	$g(i)$	$g(j)$	$b_{g(i)g(j)}$	$b_{g(j)g(i)}$
$+1$	$+1$	a_{ij}	a_{ji}	$2i$	$2j$	$\frac{a_{ij}+a_{ji}}{2} + max$	$\frac{a_{ij}+a_{ji}}{2} + max$
-1	-1	a_{ij}	a_{ji}	$2i-1$	$2j-1$	$\frac{a_{ij}+a_{ji}}{2} + max$	$\frac{a_{ij}+a_{ji}}{2} + max$
$+1$	-1	$-a_{ij}$	$-a_{ji}$	$2i$	$2j-1$	$-\frac{a_{ij}+a_{ji}}{2} + max$	$-\frac{a_{ij}+a_{ji}}{2} + max$
-1	$+1$	$-a_{ij}$	$-a_{ji}$	$2i-1$	$2j$	$-\frac{a_{ij}+a_{ji}}{2} + max$	$-\frac{a_{ij}+a_{ji}}{2} + max$

given the g function we have build, it is true that

$$\begin{aligned} &\sum_{i=1}^{M}\sum_{j=1}^{M} w_{ij}(\alpha_{ig(i)} + b_{g(i)g(j)} + \alpha_{jg(j)}) \\ =&\sum_{i=1}^{M}\sum_{j=1}^{M} b_{g(i)g(j)} \qquad (w_{ij}=1, \alpha_{ig(i)} = \alpha_{jg(j)} = 0) \\ =&\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} f(i)f(j) + \sum_{i=1}^{n}\sum_{j=1}^{n} max \\ =&\leq K + n^2 \cdot max \\ =&K^{'} \end{aligned} \tag{5}$$

Conversely, suppose that for the instance we have built for the SONDP-SHE/UN/UBC problem, there is a function $g : \{1, 2, .., M\} \rightarrow \{1, 2, ..., N\}$ such that $\sum_{i=1}^{M}\sum_{j=1}^{M} w_{ij}(\alpha_{ig(i)} + b_{g(i)g(j)} + \alpha_{jg(j)}) \leq K^{'}$. Then $\forall 1 \leq i \leq M$, $g(i)$ must be equal to $2i$ or $2i-1$, otherwise $\alpha_{ig(i)}$ will be equal to ∞. Then we can build the corresponding $f : \{1, 2, .., n\} \rightarrow \{-1, +1\}$ for the instance of Matrix Cover as follows:

$$\begin{aligned} &\forall 1 \leq i \leq n (n = M, N = 2M); \\ &f(i) = \begin{cases} +1 \text{ if } g(i) = 2i \\ -1 \text{ if } g(i) = 2i - 1 \end{cases} \end{aligned} \tag{6}$$

Similarly, due to the relationship shown in Table 3, it is true that

$$\begin{aligned} \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} f(i)f(j) &= \sum_{i=1}^{M}\sum_{j=1}^{M} b_{g(i)g(j)} - \sum_{i=1}^{n}\sum_{j=1}^{n} max \\ &= \sum_{i=1}^{M}\sum_{j=1}^{M} w_{ij}(\alpha_{ig(i)} + b_{g(i)g(j)} + \alpha_{jg(j)}) - \sum_{i=1}^{n}\sum_{j=1}^{n} max \\ &\leq K^{'} - n^2 \cdot max \\ &= K \end{aligned} \tag{7}$$

This proves the theorem.

3.3 SON Design Problem with Multi-homed/single-homed Enterprise, Uncapacitated Network and Fixed Cost Model (SONDP-MHE/UN/FC)

In this section we consider both the multi-homed and single-homed versions of the uncapacitated network with fixed cost model, which are described as follows:

Instance: Given a graph $G_{ESPN} = (V_{ESPN}, E_{ESPN})$ with matrices L, α, Ω, and a positive integer K.

Question: Is it possible to construct a SON topology with total cost (under fixed cost model) less than or equal to K so that all traffic demands given in matrix Ω are satisfied?

Instance: Given a graph $G_{ESPN} = (V_{ESPN}, E_{ESPN})$ with matrices L, α, Ω, and a positive integer K.
Question: Is it possible to construct a SON topology with total cost (under fixed cost model) less than or equal to K so that all traffic demands given in matrix Ω are satisfied, and meanwhile each end-system is connected to only one provider node?

Theorem 3. *SON design problems with MHE/UN/FC and SHE/UN/FC are NP-complete.*

Proof. Clearly, the SONDP-SHE/UN/FC problem belongs to NP. We prove its NP-Completeness by reduction from the Steiner Tree Problem. Given any instance of Steiner Tree Problem: undirected graph $G(V, E)$, weights $c : E(G) \rightarrow \mathbb{R}_+$, the set of terminals $S \subseteq V(G)$, and a positive integer K, we construct an instance $(G', L, \alpha, \Omega, K')$ for SONDP-SHE/UN/FC problem as follows: $G' = (V \cup U, E \cup E')$, where V is the set of provider nodes; $U = \{u | u$ is new added node adjacent to $v, \forall v \in S\}$ is the set of end-systems; $E' = \{(u, v) | \forall v \in S\}$; $\forall e \in E, l(e) = c(e)$; $\forall e' \in E', \alpha(e) = 0$; $\forall u_i, u_j \in U, \omega(u_i, u_j) = 1$; and $K' = K$. The construction can be done in polynomial time.

Now we show that this transformation is a reduction. Suppose for the instance of Steiner Tree problem, there is a Steiner tree $T = (V_{ST}, E_{ST})$ for S in G with total cost $c(E_{ST})$ less than or equal to K, then we can construct a solution T' for SONDP-SHE/UN/FC problem as follows: $T' = (V_{ST} \cup U, E_{ST} \cup E')$, where T' connects all the end-systems, which means the bandwidth requirement for each pair of end-systems is satisfied. In addition, each end-system is connected to only one provider node, and the cost of T' is less than or equal to K'. Similarly, given the solution $T' = (V' \cup U, E'' \cup E')$ for the instance of the SONDP-SHE/UN/FC problem, we can construct a corresponding solution T for Steiner Tree problem as $T = (V', E'')$, by removing all the end-systems and the associated edges. T is a solution for the Steiner Tree Problem, since all the terminals are connected and $c(E(T))$ is no greater than K. Therefore, the transformation is a reduction and the SONDP-SHE/UN/FC problem is NP-Complete.

It's true that the instance of SONDP-SHE/UN/FC problem we constructed can also be seen as a special instance for SONDP-MHE/UN/FC problem, since in the problem description, there is no constraint on the number of provider nodes each end-system can connect to. Therefore, a similar proof can show that SONDP-MHE/UN/FC problem is also NP-Complete.

3.4 Optimal Solution for SON Topology Design using Integer Linear Programming

In this section, we provide a 0-1 integer linear programming formulations for both SONDP-SHE/UN/FC and SONDP-MHE/UN/FC problems. The formulation

$$\text{Minimize} \sum_{i=1}^{M}\sum_{j=1}^{N} \alpha_{i,j} y_{i,j} + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} l_{i,j} z_{i,j} \tag{8}$$

Subject to

$$\sum_{j=1}^{N} q_{i,j}^{k,l} >= 1, \qquad \text{for } 1 \le i \le M,\ w_{k,l} > 0; \tag{9}$$

$$\sum_{l=1}^{N} x_{i,k}^{j,l} + q_{k,j}^{i,k} - \sum_{l=1}^{N} x_{i,k}^{l,j} - q_{i,j}^{i,k} = 0, \qquad \text{for } 1 \le j \le N,\ w_{i,k} > 0; \tag{10}$$

$$\sum_{k=1}^{M}\sum_{l=1}^{M} q_{i,j}^{k,l} \le M^2 \times y_{i,j}, \qquad \text{for } 1 \le i \le M,\ 1 \le j \le N; \tag{11}$$

$$\sum_{i=1}^{N}\sum_{k=1}^{N} \left(x_{i,k}^{j,l} + x_{i,k}^{l,j}\right) \le 2N^2 \times z_{j,l}, \qquad \text{for } 1 \le j < l \le N; \tag{12}$$

$$y_{i,j} = 0/1, \qquad \text{for } 1 \le i \le M,\ 1 \le j \le N; \tag{13}$$

$$z_{j,l} = 0/1, \qquad \text{for } 1 \le j < l \le N; \tag{14}$$

$$q_{i,j}^{k,l} = 0/1, \qquad \text{for } 1 \le i,k,l \le M,\ 1 \le j \le N; \tag{15}$$

$$x_{i,k}^{j,l} = 0/1, \qquad \text{for } 1 \le i,k \le M,\ 1 \le j,l \le N; \tag{16}$$

Fig. 6. ILP for SONDP-MHE/UN/FC Problem

for multi-homed problem is shown in Figure 6. For single-homed problem, we only need to add one more set of constraints for the ILP to ensure that exactly one access link is used for each end-system, i.e. $\sum_{j=1}^{N} y_{i,j} = 1$ for $1 \le i \le M$.

The variable $y_{i,j} = 1$ indicates that bandwidth is reserved on the access link from end-system i to provider node j. The variable $z_{j,l} = 1$ indicates that bandwidth is reserved on the transport link between provider node j and provider node l. The variable $q_{i,j}^{k,l} = 1$ indicates that the traffic from end-system k to end-system l is using the access link between end-system i and provider node j,where i is equal to k or l. The variable $x_{i,k}^{j,l} = 1$ indicates that traffic from end-system i to end-system k is using transport link between provider node j and provider node l.

The objective function in Figure 6 is the sum of the costs of access links and transport links. Constraint (9) ensures that at least one access link is used to connect an end-system to the overlay network. Constraint (10) is for flow conservation at each provider node. No traffic is initiated or terminated at a provider node. Constraints (11) and (12) determine the access links and the transport links used by the solution.

4 Approximate Algorithms for SON Topology Design

In this section we present approximate algorithms for the solution of SHE/UN/FC and MHE/UN/FC problems. It may be noted that we have shown that the

MHE/UN/UBC problem is polynomial-time solvable, and approximate solution for the SHE/UN/UBC problem has been presented in [8]. Since in the fixed cost model, the cost of each link is independent of the amount of data transmitted on it, the amount of reserve bandwidth ω_{ij} between end-systems ES_i and ES_j can be ignored. If $\omega_{ij} > 0$, then ES_i and ES_j should be connected in the resulting topology, otherwise, they don't need to be connected. Therefore, from the reserve bandwidth matrix Ω, we construct a connectivity requirement set $R = \{(s_1, t_1), \ldots, (s_k, t_k)\}$, where each (s_i, t_i) is an ordered pair of end-systems which has positive bandwidth demand.

We provide three different heuristics for the solution of SHE/UN/FC problem: (i) Randomized Heuristic, (ii) Gain-based Heuristic and (iii) Spanning Tree based heuristic. It may be noted that by shortest path between any two nodes (end-systems or provider nodes), we implies the shortest path that only uses provider nodes as intermediate nodes. In analyzing the computational complexity of each heuristic, M is the number of end-systems, N is the number of provider nodes and k is the number of connections to be established.

Heuristic 1: Randomized Approach

Step 1: Intialize $C_{RH} = \infty$ and $D_{RH} = 0$.
Step 2: Repeat steps 3 to 13 W times (the parameter W is set by the user to determine the number of times the random process is repeated).
Step 3: Set $R = \{(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)\}$.
Step 4: Randomly choose a pair (s_i, t_i) from R, and remove it.
Step 5: Compute the shortest path from s_i to t_i. Suppose in the computed shortest path, s_i is connected to provider node P_j, and t_i is connected to P_k. Call these provider nodes *gateways* for s_i and t_i, and denote them $G(s_i)$ and $G(t_i)$ respectively.
Step 6: Set $D_{RH} = D_{RH} + \{$weight of the shortest path computed in step 5$\}$.
Step 7: Set the weights of all the links on the computed shortest path zero.
Step 8: Repeat steps 9-12 till R is empty.
Step 9: Randomly choose a pair (s_i, t_i) from R, and remove it.
Step 10: If $G(s_i)$ and $G(t_i)$ are known, compute the shortest path between $G(s_i)$ and $G(t_i)$; else if $G(s_i)$ is known while $G(t_i)$ is not known, compute the shortest path between $G(s_i)$ and t_i; else if $G(s_i)$ is not known while $G(t_i)$ is known, compute the shortest path between s_i and $G(t_i)$; else if neither $G(s_i)$ nor $G(t_i)$ is known, compute the shortest path between s_i and t_i.
Step 11: Set $D_{RH} = D_{RH} + \{$weight of the shortest path computed in step 10$\}$.
Step 12: Set weights of all the links on the computed shortest path zero.
Step 13: Set $C_{RH} = min(C_{RH}, D_{RH})$.
Step 14: Output C_{RH}. This is the cost of the solution.

Computational complexity: The computational complexity of the Randomized Heuristic is $O(kW(M + N)log(M + N))$, where W is the number of times the random process is repeated.

Heuristic 2: Gain Based Approach

Step 1: Initailize $C_{GBH} = 0$.
Step 2: Set $R = \{(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)\}$.

Step 3: Compute shortest paths for all pairs of end-systems in R.
Step 4: Identify the source-destination pair (s_i, t_i) from R that has the longest path length. Remove this pair from R. Suppose in the computed shortest path, s_i is connected to provider node P_j, and t_i is connected to P_k. Call these provider nodes *gateways* for s_i and t_i, and denote them $G(s_i)$ and $G(t_i)$ respectively.
Step 5: Set $C_{GBH} = C_{GBH} + \{$weight of the path chosen in step 4$\}$.
Step 6: Set the weights of all the links on the path chosen in step 4 zero.
Step 7: Repeat steps 8-12 till R is empty.
Step 8: Compute shortest paths for all the pairs of end-systems in R. If either $G(s_i)$ or $G(t_i)$ is identified in one of the earlier iterations, in the shortest path computation, $G(s_i)$ and $G(t_i)$ should replace s_i and t_i respectively.
Step 9: Note the *gain*, i.e. the change in path length, for all the pairs in the set R.
Step 10: Identify the end-system pair (s_i, t_i) with largest gain. Remove it from R.
Step 11: Set $C_{GBH} = C_{GBH} + \{$weight of the path chosen in step 10$\}$.
Step 12: Set the weights of all the links on the path chosen in step 10 zero.
Step 13: Output C_{GBH}. This is the cost of the solution.

Computational complexity: The computational complexity of the Gain-based Heuristic is $O(k(M+N)^3)$. A different implementation can realize this in $O(k^2(M+N)^2)$. The implementation should be chosen based on the values of M, N and k.

Heuristic 3: Spanning Tree Based Approach

Step 1: Initialize $C_{STH} = 0$.
Step 2: Set $R = \{(s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)\}$.
Step 3: Compute the minimum spanning tree MST_{PN} of the subgraph induced by the Provider Nodes. Set $C_{STH} =$ Cost of MST_{PN}.
Step 4: Connect each end-system to its nearest provider node. Update C_{STH} with the additional cost of connecting all the end-systems.
Step 5: Remove those provider nodes from MST_{PN} that are not used to connect any end-systems pair, and also remove the cost used to connecting them from C_{STH}.
Step 6: Output C_{STH}. This is the cost of the solution.

Computational complexity: The computational complexity of the Spanning Tree based Heuristic is $O((M+N)^2 log(M+N))$.

The approximate algorithms for the multi-homed version are similar to the ones for the single-homed version, except that end-system is no longer required to connect to only one provider node. So the shortest path for each end-system pair should be computed directly, and the *gateway* information is not needed.

5 Experimental Results

In this section, we compare the performance of our three heuristics for the SONDP-SHE/UN/FC problem against the optimal solution obtained by solving ILP. Simulation experiments are carried out using randomly generated input sets. We develop a random graph generator, which takes as input the number of nodes and average degree, and generates connected undirected graphs. It also

randomly generates the weights on the links from a uniform distribution over a specified range (we use the ranges of 3 to 8, and 3 to 80 for our experiments). The graphs produced by the generator are used as the network for the provider nodes. The random weights on the edges are the transport cost among provider nodes. Once the network for provider nodes is generated, a specified number of end-systems are connected to the provider nodes in the following way:
Step 1: The degree of an end-system is randomly generated from a uniform distribution over the range of 1 to 10.
Step 2: The provider node neighbors of an end-system are randomly generated with uniform distribution.
Step 3: The access cost from the end-system to the provider node is randomly generated with a uniform distribution over the range of 3 to 8 (small access cost variation) or 3 to 80 (large access cost variation).
Step 4: Communication requests between end-systems are also randomly generated.

In our simulation experiments, we compute the optimal cost of SON design and the costs obtained by three different heuristics. These results are presented in Table 4. The time taken by the optimal solution as well as the heuristic solutions are also presented. In Table 4, M and N represent the number of end-systems and provider nodes respectively. There could potentially be $M(M-1)/2$ possible requests between M end-systems. The term *Req%* represents the percentage of $M(M-1)/2$ possible requests that is considered for the instance. The term *Cost-variation ratio* is defined to be the ratio of the cost difference between the heuristic solution(s) and the optimal solution to the cost of the optimal solution. The values of cost-variation for three different heuristics are presented. It may be observed that ILP fails to find a solution within a reasonable amount of time when the problem instance increases in size. The heuristics however are able to produce solutions for these instances. As noted earlier, the link cost distribution is taken to be 3 to 8 for some of the instances and 3 to 80 for the rest. We did not notice any perceptible impact of the variation of the link weights on results. From the experiment results, it may be concluded that Heuristic 2 produces the best solution for most of the instances, whereas Heuristic 3 produces the solution in the least amount of time. Clearly, a tradeoff between quality of solution and the time taken to find it exists in these two heuristics. It may be noted that all three heuristics produce a reasonable quality solution in a fraction of time needed to find the optimal solution.

References

1. D. G. Anderson, H. Balakrishnan, M.F. Kaashoek and R. Morris, "Resilient Overlay Network", *Proc. 18th ACM SOSP 2001,* Banff, October 2001.
2. A. Keromytis, V. Mishra and D. Rubenstein, "Secure Overlay Networks", *Proc. of SIGCOMM'02*, pp. 61-72, 2002.
3. Z. Duan, Z. L. Zhang and Y.T. Hou, "Service Overlay Networks:SLAs, QoS and Bandwidth Provisioning", *Proc. 10th IEEE International Conference on Network*

Table 4. Simulation Results for SONDP-SHE/UN/FC Problem

	Instance			**Cost**				**Cost-var. Ratio (%)**			**Running Time (s)**			
#	**M**	**N**	**Req%**	**Opt**	**H1**	**H2**	**H3**	**H1**	**H2**	**H3**	**Opt**	**H1**	**H2**	**H3**
1*	10	10	44	124	134	134	142	8.1^+	8.1^+	14.5	1	<1	<1	<1
2*	12	10	36	93	131	128	141	40.9	37.6^+	51.6	<1	<1	<1	<1
3*	15	10	29	103	141	165	145	36.9^+	60.2	40.8	<1	1	<1	<1
4	20	15	53	94	130	124	132	38.3	31.9^+	40.4	3	<1	<1	<1
5	25	15	42	102	150	140	144	47.1	37.3^+	41.2	2	1	1	<1
6	30	25	34	143	186	191	202	30.1^+	33.6	41.3	40	1	1	<1
7	35	25	29	129	195	185	205	51.2	43.4^+	58.9	14	3	1	<1
8*	40	30	26	704	1083	1247	1073	53.8	77.1	52.4^+	43	3	1	<1
9	45	35	23	169	301	285	309	78.1	68.6^+	82.8	11	3	2	<1
10*	50	40	20	904	1646	1475	1441	82.1	63.2	59.4^+	68	5	3	<1
11	55	45	19	237	393	364	371	65.8	53.6^+	56.5	693	6	4	<1
12*	60	50	17	1285	2245	2208	2004	74.7	71.8	56.0^+	72	6	6	<1
13	65	55	16	242	408	392	445	68.6	62.0^+	83.9	325	8	8	<1
14*	70	60	14	1464	2038	1962	2248	39.2	34.0^+	53.6	578	9	8	<1
15	75	65	14	295	496	482	548	68.1	63.4^+	85.8	422	11	11	<1
16	85	75	12	313	566	552	594	80.8	76.4^+	89.8	176	14	17	<1
17	95	85	11	N/A	608	592	638	N/A	N/A	N/A	N/A	19	25	<1
18*	80	70	13	N/A	2721	2851	2726	N/A	N/A	N/A	N/A	13	14	<1
19	105	95	10	N/A	673	674	729	N/A	N/A	N/A	N/A	23	34	<1

+: Result with best cost variation ratio
*: Link cost is between 3 and 80; otherwise, it is between 3 and 8.
N/A: ILP failed to find an optimal solution.

Protocols", Paris, France, November 2002. Also in *IEEE/ACM Trans. on Networking*, vol. 11, no. 6, pp. 870-883, 2003.
4. M. Garey and D. Johnson, "Computers and Intractability: A guide to the theory of NP-COmpleteness", W.H. Freeman, 1979.
5. X. Gu, K. Nahrstedt, R. H. Chang and C. Ward, "QoS-Assured Service Composition in Managed Service Overlay Networks", *Proc. of 23rd IEEE International Conference on Distributed Computing Systems*, Providence, May 2003.
6. Internap Network Services Corporation, *http://www.internap.com*
7. L. Subramanian, I. Stoica, H. Balakrishnan and R. H. Katz, "OverQoS: Offereing Internet QoS using Overlays", *Proc. HotNET-I Workshop*, October 2002.
8. S. L. Vieira and J. Liebeherr, "An algorithmic approach to topological design of Service Overlay Networks", *Proc. of IWQoS'04*, 2004, Montreal, Canada.
9. J. Walrand and P. Varaiya, "High Performance Communication Networks", Morgan Kaufman Publishers, 2000.
10. S. Savage, T. Anderson, and et al. "Detour: a case for informed internet routing and transport." *IEEE Micro,* 19(1):5059, January 1999.
11. Ion Stoica, Daniel Adkins, Shelley Zhuang, Scott Shenker and Sonesh Surana, "Internet Indirection Infrastructure," *Proc. of ACM SIGCOMM,* August, 2002
12. A. Akella, J. Pang, A. Shaikh, B. Maggs and S. Seshan, " A Comparison of Overlay Routing and Multihoming Route Control," *In Proc. of ACM SIGCOMM,* 2004