Computer Networks 56 (2012) 1286–1302

Contents lists available at SciVerse ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Path-vector contracting: Profit maximization and risk management

Praveen K. Muthuswamy [a,*], Aparna Gupta [a], Murat Yuksel [b], Koushik Kar [a]

[a] Rensselaer Polytechnic Institute, Troy, NY 12180, United States
[b] University of Nevada, Reno (UNR), NV 89557-0042, United States

ARTICLE INFO

Article history:
Received 4 November 2010
Received in revised form 18 September 2011
Accepted 10 November 2011
Available online 13 December 2011

Keywords:
Internet
Non-linear pricing
End-to-end contracts
Modeling
Optimization
Risk management

ABSTRACT

We consider an Internet Service Provider's (ISP's) problem of providing end-to-end (e2e) services with bandwidth guarantees, using a *path-vector* based approach. In this approach, an ISP uses its edge-to-edge (g2g) single-domain contracts and vector of contracts purchased from neighboring ISPs as the building blocks to construct, or participate in constructing, an end-to-end "contract path". We develop a *spot-pricing* framework for the e2e bandwidth guaranteed services utilizing this path contracting strategy, by formulating it as a stochastic optimization problem with the objective of maximizing *expected profit* subject to *risk* constraints. In particular, we present time-invariant path contracting strategies that offer high expected profit at low risks, and can be implemented in a fully distributed manner. Simulation analysis is employed to evaluate the contracting and pricing framework under different network and market conditions. An admission control policy based on the path contracting strategy is developed and its performance is analyzed using simulations.

1. Introduction

The Internet consists of several network domains owned and administered by independently operated ISPs. Today's end-to-end (e2e) Internet services are mostly best-effort without any explicit QoS guarantee mainly due to the fact that "all bets are off" once the traffic crosses into another ISP's domain. Significant amount of research has been carried out to provide QoS-assured Internet services, most of them focusing on intra-domain QoS guarantees. Though e2e QoS contracts are currently possible via virtual private networks, such possibilities only work with static and long-term contracts. Customers often demand services that require more dynamism and crossing of multiple ISPs, which cannot be realistically accommodated via rigid and static inter-ISP service level agreements (SLAs) in the current Internet architecture. The problem of end-to-end (e2e) QoS provisioning is further complicated due to the size and heterogeneity of users and providers in the Internet. Towards this end, we believe that using path-vector routing principles to construct end-to-end QoS-enabled services hold promise, as it provides a scalable and flexible mechanism for solving the e2e QoS provisioning problem.

In this work, we consider a path-vector based bandwidth contracting framework as a possible future architecture for Internet routing, and attempt to answer if such an inter-domain routing architecture is economically viable from an ISP's perspective. We develop the necessary tools and techniques that can help ISPs maximize their profitability while meeting their contractual obligations, and thereby satisfying the e2e demands. In particular, we develop a spot-contracting and non-linear pricing framework for e2e services over the Internet for time durations finer than the existing SLAs. An ISP constructs or participates in constructing a set of "contract paths" to destination networks (ISPs) in order to serve its customer base. A customer (which can be an end user or an upstream ISP) can then enter into spot-contracts in order to get a flexible-duration of QoS guaranteed services from itself to the destination networks (ISPs). The QoS metric that we consider

* Corresponding author.
E-mail addresses: muthup@rpi.edu (P.K. Muthuswamy), guptaa@rpi.edu (A. Gupta), yuksem@cse.unr.edu (M. Yuksel), kark@rpi.edu (K. Kar).

1389-1286/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2011.11.004

in this paper is data rate that can be supported between different source–destination pairs by the ISP. Most of the SLAs involve some form of delay or even loss guarantees. However, such guarantees are typically defined as additional parameters on top of a baseline average and/or maximum traffic rate or bandwidth. The average and maximum rates are more realistic at a more aggregate level, such as peering SLAs. Because of these reasons, we have focused our efforts on formulating the optimization problem just on the bandwidth. It is, however, possible to extend our formulation with other metrics. We could, for example, define utility functions for the provider that depends on the experienced delay. With some changes to our formulation, these can then be incorporated in the objective function to provide delay guarantees. However, the problem of reserving enough bandwidth along the paths would still be the major bottom line. Therefore, we consider only bandwidth as the QoS metric in this paper. Our framework allows the ISP both temporal (by tuning the duration of the contracts it advertises) and spatial (by advertising different prices for contracts crossing its different ingress and egress points) flexibility in contracting. Note that in general, an ISP's strategy for constructing the contract paths determines its profitability and customer satisfaction.

1.1. Path-vector contracting framework

We consider a path-vector contracting framework in which an ISP can announce different contract paths and prices for different destinations. This can be realized using a contract-switched Internet architecture proposed in our earlier work [1], where each ISP is abstracted as a set of edge-to-edge (g2g) contracts (which we call intra-domain "contract links") as shown in Fig. 1. In such an architecture, an ISP can compose e2e contract paths in a path-vector style by using its g2g single-domain contracts and the vector of service contracts purchased from the neighboring ISPs. This framework is scalable since it is based on path-vector style of contracting and is directly compatible with the existing inter-domain routing protocols such as BGP [2]. In our framework, an ISP would receive advertisements for *contract paths* to a destination from its neighboring ISPs through its egress points. These *contract paths* (hereafter referred as *extra-domain paths*) ensure guaranteed service for traffic starting from a neighboring ISP to the destination. Therefore, the ISP can construct end-to-end contract paths to a destination by choosing some of these *extra-domain paths* and prepending them with its edge-to-edge contract links.

A key component in our inter-ISP architecture is the g2g contract links which provide guaranteed service between two end points of an ISP's domain. We illustrate this using the contract-switching Internet abstraction shown in Fig. 1. Traditionally, for inter-domain routing, ISPs in the Internet are abstracted as nodes as shown in Fig. 1(b). Now, consider the path-vector contract-switched model [1], shown in Fig. 1(a) and let us suppose that the ISP A wants to setup end-to-end contracts between point (router/network) 1 located within its domain and point (router/network) 9 located in ISP D. The links 1-2, 1-3, 3-5 are g2g contract links belonging to ISP A. Similarly 3-6 and 2-8 are g2g contract links belonging to ISP B. Now, the ISP A would look for advertisements for extra-domain paths to point 9 from its neighboring ISPs B and C. Examples of extra-domain paths from the neighbor ISP B include 2-8-9 (involves the g2g links 2-8, 8-9), and 3-6-10-9 (involves links 3-6, 6-10, and 10-9). ISP A then provides one or more of its available intra-domain contract links, and enters into contract with its neighboring ISPs (B and C) to carry the traffic forward from 1 up to 9. The idea of path formation using negotiation and contracting has been proposed earlier in [3–5], although these approaches still use the node level ISP abstraction. The focus of this paper, however, is not on the protocol design and policy issues of how a path-vector (i.e. extra-domain contract path) gets constructed, but *on developing a profit maximizing risk management strategy to help ISPs select which contract paths to participate.*

Since dynamic, end-to-end provisioning for bandwidth and other QoS metrics is largely absent from the current Internet, limited-term services that require end-to-end QoS guarantees are not readily enabled. On-demand high-quality video-conferencing/streaming is one such example – even though such applications are run over the current Internet, they are not associated with any QoS guarantees, and often do not scale well to high quality/resolution. The path-vector contracting framework that

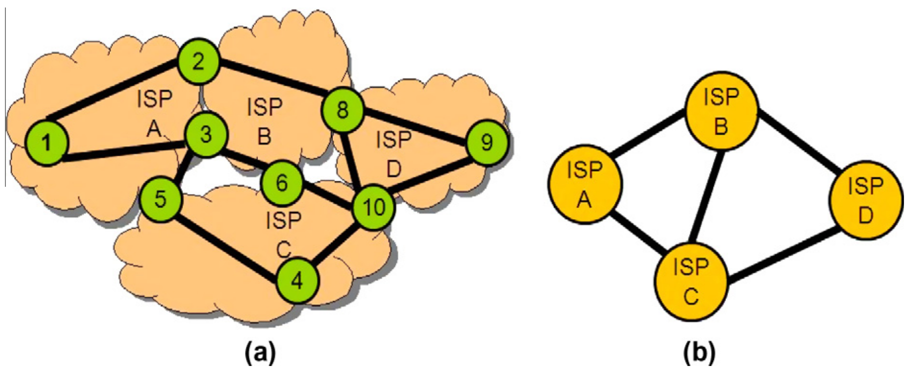

Fig. 1. (a) Contract-switching framework: An ISP is abstracted as a collection of several edge-to-edge (g2g) contract links. For ISP A, 1-2, 1-3, 3-5 are intra-domain contract links, while 2-8-9, 3-6-10-9 and 3-5-4-10-9 are extra-domain paths available for an end-to-end contract path establishment to destination 9. (b) Network abstraction traditionally used in inter-domain routing, where each ISP is abstracted as a single node.

we consider is well suited for this purpose, as it can concatenate several guaranteed but single-domain contracts into end-to-end reliable and QoS-guaranteed services.

1.2. E2E contracting considerations: an ISP's perspective

Viewed from the perspective of an ISP participating in path-vector style contracting as described above, the end-to-end (e2e) contracting problem boils down to establishing contract paths between the ISP's customers (end users or other ISPs) and specific destinations that the customers want to send traffic to. In the following, our reference to the e2e contracting problem, e2e contract paths or e2e traffic (service) demands must thus be viewed from a perspective of a single ISP. Note that the e2e contract paths are not set up dynamically for each arriving incoming request, which may happen on the seconds to minutes time-scale. Instead, we obtain a relatively longer-term e2e contracting strategy covering hours to days, wherein the contract paths are set up based on past data on the demand for intra-domain and extra-domain (inter-domain) services.

Typically, the demand for e2e services and network conditions can be dynamic and stochastic. Therefore, provisioning of e2e contract paths becomes risky due to uncertainties caused by competing traffic in the Internet. In such situations, an ISP cannot determine the exact amount of capacity to contract along each of the links/paths, since traffic is inherently stochastic. A time-invariant contracting strategy developed for the entire planning period can lead to capacity deficit or under-utilization of resources at certain times. Moreover, time-varying demands or router failures within any other Internet domains can potentially cause changes in the prices of the contract paths offered by the neighboring ISPs, premature termination of contract by neighboring ISPs, etc. Therefore, the *ISP's cost* of providing or participating in creating the e2e service can increase or decrease, due to which the provider may or may not deliver the e2e service at the promised quality level. In this paper, the marginal cost of extra-domain service is modeled as *mean-reverting random walk* process. We consider two models for the demand process, namely, the mean-reverting process and the time-of-day process, described in Section 3.2.1.

For pricing links and paths, we use the *Ramsey pricing model* with a reasonable choice of demand profile for the ISP's customer base. Due to uncertainty in the customer demand as well as the cost of the e2e service, any contracting and pricing strategy would lead to fluctuations in the ISP's profit. Therefore, *risk management* becomes a critical issue while providing e2e services. In this paper, we assume that the average e2e user demand and cost can be determined using historical data. We show that the ISP can achieve significantly higher expected profit, at a given risk level, by reserving capacities along the links and paths that target supporting the average e2e demands *inflated* by a certain factor.

1.3. Contribution

The major contributions of this paper can be outlined as follows.

1. We introduce an architectural framework for scalable construction of e2e contracts by using the edge-to-edge single-domain contracts and the vector of contract paths obtained from other ISPs using a path-vector approach.
2. The e2e path contracting and pricing strategy for a participating ISP is formulated as a stochastic optimization problem with the objective of maximizing expected profit subject to risk constraints.
3. We obtain a solution to the problem in the space of time-invariant contracting strategies. The path contracting solution achieves maximum expected profit, for a given constraint on the risk (standard deviation) of profit.
4. The optimal time-invariant contracting strategy is assessed for the effect of input volatility, path failures, and path correlation properties.
5. We use the Rocketfuel dataset [6] and the GTITM models [7] to evaluate the performance of the proposed path contracting solution.
6. Finally, we develop and evaluate an admission control policy that can be combined with the contracting strategy to obtain implementable solutions for e2e service provisioning.

1.4. Organization

The rest of the paper is organized as follows. Section 2 provides a brief review of state-of-the-art for QoS support and pricing. In Section 3, we develop the models and formulate the e2e pricing and risk management problem as a stochastic optimization problem, and present an approximation solution to the problem. Section 4 presents a detailed discussion of simulation results including the study of the effects of input volatility, failures, and correlation on the contracting strategy. Section 5 addresses the challenges involved while implementing the long-term strategy in practice. We conclude the paper and provide prospects for future research in Section 6.

2. Related work

Technology to Support QoS: In the Internet, QoS deployment in multi-domain, IP-based inter-networks has been an elusive goal, partly due to complex deployment issues [8]. From an architectural standpoint, contemporary QoS research has recognized the need to *simplify* and *de-couple* building blocks to promote implementation and inter-network deployment. RSVP [9] de-coupled inter-network signaling from routing. The IntServ [10] de-coupled e2e support from network support for QoS. IntServ is not scalable because of the complexity and overheads caused by per flow control and data-plane functions in the entire network. The DiffServ services and stateless or soft-state packet processing techniques (e.g., CSFQ [11]) further simplified core architecture and moved data-plane complexity to the edges, and allowed a range of control-plane options [12,13,9].

More recently, overlay networks [14,15] even further de-coupled QoS delivery goals from network-level implementation and have become an attractive alternative for e2e QoS delivery, possibly overcoming the problems at peering points [16]. Technologies and concepts are being developed to address the challenge of provisioning QoS assurances at various levels with pilot studies [17].

Economic Tools for QoS Provisioning: Though the technological constraints have been an important part of e2e QoS provisioning problem, economic constraints have been recognized to be a major factor as well. Construction of economic tools such as pricing for improving the Internet QoS has been an active research area in the past decade. Various pricing schemes have been extensively studied in the literature. Pricing schemes can be classified as being static or dynamic based on whether prices change with the state of the network. Static pricing, including some class dependent pricing schemes [18] as well as the traditional flat rate or time-of-the-day pricing [19], does not react to the congestion state of the network. On the other hand, dynamic pricing schemes such as Smart Market [20], proportional fair pricing schemes [21], Priority Pricing [22], take varying network conditions into account. Dynamic pricing schemes have been shown to be useful in formulating the provider's pricing decisions when the provider and customers act to maximize their own benefits [21], or identifying the value of services to customers [20]. Savagaonkar et al. [23] studied the dynamic pricing problem of bandwidth services with different service classes, where the customer's demand are driven by an underlying traffic-state process. A comparative study of dynamic pricing schemes can be found in [24]. Dynamic pricing schemes proposed thus far are in general computationally expensive and may raise scalability concerns.

A common approach to handling QoS issues in pricing is to use the concept of "customer class", where each class is associated with certain QoS level [18,25–28]. Prices are usually determined based on the definitions of "class". Analysis is performed on how prices may affect resource allocation and the actual QoS experienced by the customers due to traffic intensities. However, a precise QoS specification itself is often missing. QoS delivery in the packet-switching Internet has an inherently stochastic, or risky, nature [29–31]. It was argued that lack of mechanisms for managing the risks in QoS delivery has contributed to the failure of QoS assured services to thrive, despite active research and development of standards [30].

In pricing e2e services, the interaction between the providers involved in delivering the service needs to considered, in addition to the profit or welfare maximization objectives. Pricing is tied with the e2e service delivery architectures. Li et al. [27] propose a hierarchical pricing scheme in DiffServ with e2e admission control, which uses congestion pricing within each domain. However, such congestion prices are more geared to support network management or traffic regulation tasks and are not necessarily the monetary prices a provider will charge the customers.

As a key difference, we consider implementable pricing of e2e contracts of shorter durations ranging from several minutes to hours. Our work builds on a nonlinear pricing model proposed in [32] for cost recovery, and develops it for e2e services by using path-vector based approach for finding the service paths. Our earlier work [33] also focussed on pricing e2e bandwidth guaranteed services, however, it does not address the risk involved in providing assured services and assumes existence of a multi-ISP overlay provider. In this paper, we propose a path-vector based approach to construct e2e contract paths without additional infra-structural requirements. Unlike previous works, we focus on minimizing the risks in ISP's profit arising due to changing network conditions and e2e demands.

3. Problem formulation and solution approach

Consider any ISP that is participating in path-vector style contracting strategy. For this ISP, traffic from the customer (end-user or upstream ISP) enters at an *ingress* edge router located inside the ISP's domain. If the destination of the traffic is outside the ISP's domain, then this traffic exits the ISP at one of its *egress* edge routers, beyond which it traverses an *extra-domain* path (through one or more other ISPs) to the destination.

Each contract link is defined between two edge points: ingress and egress. As the baseline QoS parameter, a contract link has an edge-to-edge (g2g) effective available capacity between the ingress and egress points. The g2g effective capacity of the contract link will depend on the loads and capacities of the individual links being traversed by it. The ISP that owns the contract link will estimate the g2g effective capacity using historical data on maximum and average utilization levels of the links through which the g2g contract link may be routed. Since e2e paths are comprised of g2g contract links and extra-domain contract paths, there are three economic characteristics of the e2e service that the ISP needs to consider: *g2g contract link maintenance and opportunity cost, extra-domain path contracting cost, and e2e contract path price.* An ISP is responsible for the contracts made on its own g2g contract links. The maintenance and opportunity cost of each g2g contract link used for e2e service creation should be an input to the owner ISP's profit maximization problem. Extra-domain path contracting cost depends on the offer of neighboring ISPs. The contracts are typically short-term, where each provider acquires bandwidth from other providers dynamically as a 'customer'. Therefore, these are customer-provider relationships and not peering relationships. As mentioned earlier, the cost of the extra-domain paths are dynamic and stochastic. Considering the composition of e2e paths, the price of e2e service is a function of the intra-domain g2g contract link cost and the extra-domain path contracting cost. In the following sections, we describe our models for each of these components and address the end-to-end path contracting problem.

3.1. Topology and contract link/path models

Let the planning horizon be divided into N epochs of fixed duration (an hour or day) each. In our formulation, it suffices to model the specific ISP under consideration by a set of ingress and egress nodes, with logical "intra-domain" contract links connecting the ingress nodes with the egress nodes.

Let $\mathcal{I}$ and $\mathcal{E}$ represent the set of ingress and egress nodes, respectively, and let $L = (\mathcal{I} \times \mathcal{E})$ represent the set of intra-domain contract links connecting them, as shown in Fig. 2. The extra-domain contract paths are abstracted as paths connecting the egress node (nodes in $\mathcal{E}$) with the set of destination nodes, represented by $\mathcal{D}$. Let the set of extra-domain contract paths be denoted by $P = (\mathcal{E} \times \mathcal{D})$. (Refer Table 1).

The set of intra-domain contract links is associated with a capacity matrix, $\mathcal{U} = (U_{ie},\ i \in \mathcal{I},\ e \in \mathcal{E})$, where $U_{ie} > 0$ indicates that ingress node i and egress node e are linked with total capacity U_{ie}, and $U_{ie} = 0$ implies no connectivity between i and e. Similarly, the set of extra-domain contract paths is associated with a capacity matrix, $\mathcal{W} = (W_{ed},\ e \in \mathcal{E},\ d \in \mathcal{D})$, where $W_{ed} > 0$ indicates that egress node e and destination node d are linked with total capacity W_{ed}, and $W_{ed} = 0$ implies no connectivity between e and d.

Let $\mathcal{M}(t) = (\mu_{id}(t),\ i \in \mathcal{I},\ d \in \mathcal{D})$ represent the demand matrix between the ingress-destination pairs at time t; let μ_{id} be the time-averaged demand. The demand $\{\mu_{id}(t)\}_{t=1}^{N}$ is modeled as a random process with known time-average value of μ_{id}. Recall that the ISP's customer can be an end-user or an upstream ISP. Therefore, the traffic entering an ingress router of the ISP could originate from an end-user within the ISP's domain, or from another ISP that chooses relay traffic to the ISP under consideration through that ingress router. The demand matrix captures the sum of both these types of traffic (of "local" as well of "distant" origin) arriving at the ingress node, categorized in terms of their final destinations.

To meet these demands, the ISP reserves bandwidth on the available "intra-domain" contract links and "extra-domain" contract paths. Let the decision variables $y_{ie}(t)$, $x_{ed}(t)$ represent the amount of bandwidth reserved on the intra-domain contract link between ingress–egress pair (i,e) and the extra-domain path between the egress-destination pair (e,d), respectively, for time t. The ISP must determine $y_{ie}(t)$, $x_{ed}(t)$ for the entire planning horizon.

3.2. Profit maximization

We formulate the e2e contracting problem as a stochastic optimization problem with the objective of maximizing the mean profit subject to risk constraints. In the following sections, we describe the components that determine the ISP's profit, namely, the demand, the revenue, and the cost.

Table 1
Table of symbols.

$\mathcal{I},\ \mathcal{E},\ \mathcal{D}$	Set of ingress, egress, and destination nodes
π	Price schedule
c_I, c_S	Marginal costs of Intra-domain link and e2e service
$c_E(t)$	Marginal cost of extra-domain path
α	Ramsey number
U_{ie}, W_{ed}	Capacity of link and path
$\mu_{id}(t)$	End-to-end demand
ϵ	Inflation factor

3.2.1. Demand modeling

The demand for services between a source i and a destination d is typically stochastic, denoted for the tth epoch by $\mu_{id}(t)$. The demand process, $\mu_{id}(t)$, is modeled as a mean-reverting random walk process with μ_{id} as its long-term mean. The initial value of the process is its long-term mean, μ_{id}. The size of increment during each epoch is fixed and denoted by δ_μ. Initially, the probability of positive and negative increments are both 0.5. However, if there was a positive increment during the previous time step, then the probability of positive increment for the current time-step decrease by p. Thus, the parameter p controls the rate of mean-reversion of the process. For $p = 0.5$, the process alternates between the values $\mu_{id} + \delta_\mu$ and $\mu_{id} - \delta_\mu$. For $p = 0$, the demand process is a symmetric random walk process. For $t > 1$, the mean-reverting random walk process can be represented by the following equations:

$$\begin{aligned}
&\text{If}\quad \mu_{id}(t) > \mu_{id}(t-1), \quad \mu_{id}(t+1) = \begin{cases} \mu_{id}(t) + \delta_\mu, & \text{w.p.}\quad 0.5 - p, \\ \mu_{id}(t) - \delta_\mu, & \text{w.p.}\quad 0.5 + p. \end{cases} \\
&\text{If}\quad \mu_{id}(t) < \mu_{id}(t-1), \quad \mu_{id}(t+1) = \begin{cases} \mu_{id}(t) + \delta_\mu, & \text{w.p.}\quad 0.5 + p, \\ \mu_{id}(t) - \delta_\mu, & \text{w.p.}\quad 0.5 - p. \end{cases}
\end{aligned} \tag{1}$$

It can be shown that the expected value of the demand during any epoch t is, $E[\mu_{id}(t)] = \mu_{id}$. The demand processes, $\mu_{id}(t)$, for $i \in \mathcal{I}, e \in \mathcal{E}$ are assumed independent of one

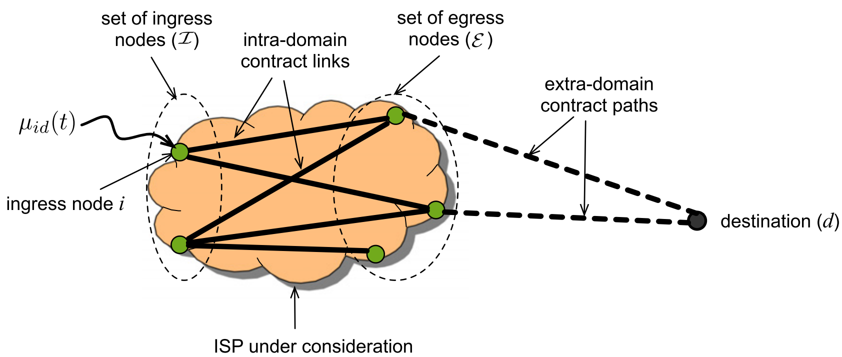

Fig. 2. Network model abstraction: Demand $\mu_{id}(t)$ between ingress node i and destination d must be supported using intra-domain contract links between the ingress i and egress nodes $e \in \mathcal{E}$, and extra-domain contract paths between the egresses $e \in \mathcal{E}$ and destination d. Note that contract links may not exist between all ingress–egress pairs; similarly, extra-domain paths may not exist from all egress nodes to a destination.

another, each having different mean and step size values. A sample realization of the mean-reverting process is shown in Fig. 3(a).

3.2.2. Cost modeling

The cost of e2e service has two components, namely the *intra-domain cost* and the *extra-domain cost*. We assume a non-linear pricing scheme to determine the cost of g2g (i.e., ingress–egress) contract links cost and the extra-domain paths. A non-linear pricing scheme refers to one where the tariff is not proportional to the quantity purchased and the marginal price for successive purchases decreases. In this paper, we explore the Ramsey pricing model [34], which is widely popular in telecommunications and power sectors, and produces an efficient tariff design in situations where due to either regulation or competition, revenues sufficient to only recover the provider's total cost may be achievable. Nonlinear pricing models have also been considered in our earlier work, [32], for pricing Internet services. However, in [32], the focus is on studying the price of intra-domain (g2g) services for different demand profiles of the intra-domain customer base.

In non-linear pricing schemes, the demand characteristics of a customer-base are often described by a demand profile, $N(\pi(q),q)$, defined as the number or fraction of the customer-base that will buy at least q units at the marginal price $\pi(q)$. $\pi(q)$ is called the price schedule. In the Ramsey pricing model, the price schedule maximizes a commonly used measure for the aggregate customers' benefits, namely the total consumer surplus given by,

$$CS(q) = \int_{\pi(q)} N(\pi, q)dq. \quad (2)$$

We choose a sample demand profile $N(\pi(z),z) = 1 - \pi(z) - z$, where $\pi(z)$ is the price schedule. The optimal price schedule that maximizes the consumer surplus is given by $\pi^*(z) = \frac{c+(1-z)\alpha}{1+\alpha}$, where the parameter c is the marginal cost and α is a Ramsey number. Ramsey number is an indicator of the market structure the provider operates in. A larger value of α indicates a higher revenue requirement or a greater monopoly power. For instance, a profit-maximizing monopolist has $\alpha = 1$, while a regulated firm with no binding revenue requirement has an $\alpha = 0$. In the case of a budget-constrained welfare maximization and an oligopolistic competition, $0 < \alpha < 1$. Since Ramsey number is to a large extent dependent on the competitive structure of the market, the provider has but limited flexibility in choosing its value at a given time, and needs to take it as given when deciding the prices. In addition, a regulatory agency may require it to be within a certain range in order to control the pricing adopted by the providers. For a budget-constrained welfare maximization and oligopolistic competition, which is the general setting for our work, $0 < \alpha < 1$.

The optimal price schedule of an intra-domain contract link, which is the opportunity cost of intra-domain bandwidth, is determined by the ISP, given by, $\pi_I^* = \frac{c_I+(1-z)\alpha_I}{1+\alpha_I}$. (The suffix I refers to intra-domain). The total intra-domain cost obtained by summing over all intra-domain contract link capacity utilized is given by,

$$C_I(t) = \sum_{i\in\mathcal{I},e\in\mathcal{E}} \int_0^{\frac{y_{ie}(t)}{U_{ie}}} \pi_I^*(q)dq. \quad (3)$$

The cost of the extra-domain paths to the ISP represents the price charged by the neighboring ISPs for providing service from an egress node to a destination node. Due to dynamic network conditions, the cost of the extra-domain paths would change with time. To incorporate this in our framework, we consider stochastic marginal costs while determining the cost of the extra-domain paths using Ramsey model. In other words, the parameter c of the Ramsey pricing equation is now time-varying and different for each extra-domain path, denoted by $c_E^{ed}(t)$ for the path between egress e and destination d. The marginal cost processes, $c_E^{ed}(t)$, are modeled as mean-reverting process, with known time-averaged value c_E^{ed}. The price schedule for the extra-domain paths is,

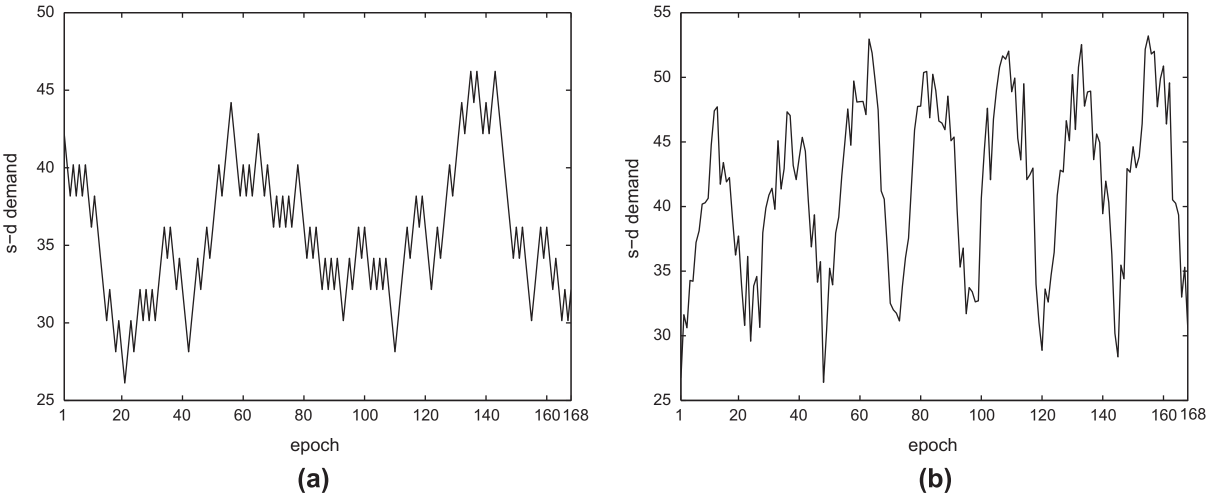

Fig. 3. (a) Mean-reverting process, (b) Time-of-day model. The duration of each epoch is an hour. Thus, the total duration shown in the figures is the duration of a week (168 h). In the Time-of-day model (b), the demand for first 48 epochs appears lower than the rest of the epochs. This is due to the fact that these epochs correspond to week end hours, where the traffic is assumed to be slightly lower than the traffic during week day hours.

$$\pi_E(z) = \frac{c_E + (1-z)\alpha_E}{1+\alpha_E}, \tag{4}$$

where c_E and α_E are the marginal cost and ramsey number of the pricing model. Note that each path has a different marginal cost, $c_E^{ed}(t)$, for epoch t. The total extra-domain path cost at time t is a function of the current marginal costs and demand, given by:

$$C_E(t) = \sum_{e\in\mathcal{E},d\in\mathcal{D}} \int_0^{\frac{x_{ed}(t)}{W_{ed}}} \pi_E(q)dq. \tag{5}$$

The ISP's total cost is the sum of intra-domain and extra-domain cost, $C(t) = C_I(t) + C_E(t)$.

We consider two models for the cost processes, $c_E^{ed}(t)$. In the first model, we assume that the marginal costs of different paths are independent of each other and model them in the same way as the demand process. That is, each marginal cost process, $c_E^{ed}(t)$, is a mean-reverting random walk process (as in Eq. 1) with parameters $c_E(0)^{ed}$, δ_c, and p. In general, the cost processes cannot be assumed to be independent of one another. The marginal cost of a path, $c_E^{ed}(t)$, represents an aggregate value decided by all the ISPs in the path. Hence, the cost processes for different paths will naturally be correlated due to common ISPs residing on the paths. For ease of generation of correlated random process in our simulation analysis, we consider a second model for the cost process (Eq. 1), in which, $c_E^{ed}(t)$ is modeled as a continuous mean-reverting random walk process instead of the discrete model. The cost process $c_E^{ed}(t)$, is defined by

$$c_E^{ed}(t+1) = c_E^{ed}(t) + dc_E^{ed}(t),$$
$$dc_E^{ed}(t) = \gamma^{ij}\left(c_E^{ed} - c_E^{ed}(t)\right)dt + b_1^{ed} \times c_E^{ed}(t) \times dW_t^{1ed}. \tag{6}$$

Here, c_E^{ed} is the mean and γ^{ij} is the rate of convergence to mean. The increments dW_t^{1ed} are correlated normal random variables, because, we consider continuous mean-reverting random walk for correlated marginal costs. We have assumed mean-reverting models for the marginal costs due to the following reason. The cost of the extra-domain paths changes due to dynamic network conditions, variability in the path congestion being one of the main factors behind it. Thus one of the primary contributors of the dynamism in the network state is the demand from customers in other ISPs. Hence, we have modeled the cost also as mean-reverting processes similar to the demand.

3.2.3. Revenue modeling

Next, we develop the ISP's revenue model for the e2e service, which is also assumed to be based on the non-linear pricing approach. The optimal price schedule using the Ramsey rule is, $\pi_S^*(z) = \frac{c_S+(1-z)\alpha_S}{1+\alpha_S}$, where c_S and α_S are the marginal cost and the Ramsey number for the e2e service. c_S must be set based on the ISP's total cost, C. The revenue generated from e2e service depends on the pricing parameters, the level of e2e demand, and the links/paths chosen to route the traffic. We define a maximum demand matrix, $\mu^{\max}$, whose entries ($\mu_{id}^{\max}$) denotes the maximum demand that can be supported between source (ingress)-destination pair (i,d). $\mu_{id}^{\max}$ can be defined in terms of the capacity elements U_{ie} and W_{ed} as follows: $\mu_{id}^{\max} = \min\{\sum_{e\in\mathcal{E}} U_{ie}, \sum_{e\in\mathcal{E}} W_{ed}\}$. The total revenue collected by the ISP per unit time is

$$R^d(t) = \sum_{i\in\mathcal{I}} \sum_{d\in\mathcal{D}} \int_0^{\frac{\mu_{id}(t)}{\mu_{id}^{max}}} \pi_S^*(q)dq. \tag{7}$$

The demand, $\mu_{id}(t)$, is stochastic and therefore, the total demand at an ingress node during some epochs can become so high as to not be feasibly allocated to reserved capacity (i.e. $y'_{ie}s, x'_{ed}s$). This excess traffic cannot contribute to the ISP's revenue. Since the ISP is planning its long-term pricing and contracting strategy, it is difficult to choose in advance, a particular ingress-destination pair, whose traffic should be cut in order to make the total demand at an ingress node feasible. Hence, we associate a cost to the total excess traffic, $Q(t)$, namely the Cost of QoS Loss, $CQ(t)$. $Q(t)$ is defined as,

$$\begin{aligned} Q(t) &= \sum_{i=1}^{s}\left\{\sum_{d=1}^{k}\mu_{id}(t) - \sum_{e=1}^{m} y_{ie}(t)\right\}^+ \\ &\quad + \sum_{d=1}^{k}\left\{\sum_{i=1}^{s}\mu_{id}(t) - \sum_{e=1}^{m} x_{ed}(t)\right\}^+ \\ &= \sum_{i=1}^{s} Q_I(i,t) + \sum_{d=1}^{k} Q_E(d,t) \end{aligned} \tag{8}$$

The terms $Q_I(i,t)$ and $Q_E(d,t)$ are the total excess traffic at an ingress node i and destination node d respectively, at time t. We associate a penalty (or cost) to the total excess traffic rather than the individual source–destination demands that exceeds the average levels. The reason being the following: Even if some of the source–destination demands exceed their average level, the overall demand can still be supported if some other demand decreases by the same amount. Moreover, a particular i-d pair cannot be made responsible for creating excess traffic.

Let the maximum excess traffic at an ingress i and destination d be $Q_I^{max}(i)$ and $Q_E^{max}(d)$. The cost of QoS loss is described as price of the unmet demand, $Q(t)$, as follows,

$$CQ(t) = \sum_{i=1}^{s} \int_0^{\frac{Q_I(i,t)}{Q_I^{max}(i)}} \pi_S^*(z)dz + \sum_{d=1}^{k} \int_0^{\frac{Q_E(d,t)}{Q_E^{max}(d)}} \pi_S^*(z)dz, \tag{9}$$

where

$$Q_I^{max}(i) = \max_t Q_I(i,t) = \max_t \sum_{e=1}^{m}\{U_{ie} - y_{ie}(t)\}, \quad \text{and}$$

$$Q_E^{max}(d) = \max_t Q_E(d,t) = \max_t \sum_{e=1}^{m}\{W_{ed} - x_{ed}(t)\}.$$

The above definitions for $Q_I^{max}(i)$ and $Q_E^{max}(d)$ follows from the fact that capacity values W_{ed} and U_{ed} are such that they can support the maximum possible demands. The cost of QoS loss is deducted from the revenue, thus the contracting solution penalizes the ISP for not reserving enough bandwidth to support the entire demand.

The effective revenue for the t epoch after adjusting for the cost of QoS loss, is given by

$$R(t) = R^d(t) - CQ(t). \tag{10}$$

3.2.4. End-to-end path contracting problem

We formulate the ISP's e2e path contracting problem as a profit maximization problem subject to constraints on the risk of profit. We introduce a new decision variable, ϵ, termed as the *inflation factor*, whose significance is discussed shortly.

$$\max \quad E[Profit] = \sum_{t=1}^{N} E\{R(t,\epsilon) - C(t,\epsilon)\}, \tag{11}$$

$$\text{s.t.} \quad E\{R(t,\epsilon)\} \geqslant E\{C(t,\epsilon)\}, \tag{12}$$

$$\sum_{e\in\mathcal{E}} y_{ie}(t) = \sum_{d\in\mathcal{D}} \mu_{id} \times (1+\epsilon), \quad \forall i \in \mathcal{I}, \tag{13}$$

$$\sum_{i\in\mathcal{I}} y_{ie}(t) = \sum_{d\in\mathcal{D}} x_{ed}(t), \quad \forall e \in \mathcal{E}, \tag{14}$$

$$\sum_{e\in\mathcal{E}} x_{ed}(t) = \sum_{i\in\mathcal{I}} \mu_{id} \times (1+\epsilon), \quad \forall d \in \mathcal{D}, \tag{15}$$

$$0 \leqslant y_{ie}(t) \leqslant U_{ie}, \quad 0 \leqslant x_{ed}(t) \leqslant W_{ed}. \tag{16}$$

$$StdDev\{Profit\} \leqslant V_{th}. \tag{17}$$

The decision variables in this problem are $y_{ie}(t)$, $x_{ed}(t)$ and the inflation factor ϵ. Eqs. (13)–(15) are the flow conservation equations at the ingress, egress and destination nodes respectively. (Eq. 17) specifies the maximum amount of risk (Vth) the ISP is willing to take. We use the standard deviation of profit as the measure of risk. In our problem formulation, we are bounding the variability in the total profit realized over the entire planning period. The planning period N could be a week or a month or even a year. The variability in profit is caused both by the demand as well as the cost of the extra-domain paths. Thus, the ISPs have the choice to contract along different extra-domain paths based on their risk preferences.

From the flow conservation Eqs. (13 to 15), we find that the ISP reserves just enough bandwidth to meet the average demand inflated by a factor, ϵ. This is due to the fact that the only information known about the demand is its time-average value and volatility.

$R(t,\epsilon)$ and $C(t,\epsilon)$ represent the net revenue and cost respectively, for a fixed ϵ, given by

$$R(t,\epsilon) = R^d(t) - CQ(t,\epsilon), \tag{18}$$

$$C(t,\epsilon) = C_I(t,\epsilon) + C_E(t,\epsilon). \tag{19}$$

The terms $CQ(t,\epsilon)$, $C_I(t,\epsilon)$ and $C_E(t,\epsilon)$ denotes cost of QoS loss, intra-domain cost, and extra-domain cost, respectively, when the reservation is made for ϵ-inflated demands. The revenue from demand, $R^d(t)$, depends only on the demand and not the reservation levels. Therefore, $R^d(t)$ is independent of ϵ. When the inflation factor $\epsilon = 0$, the reservation is such that only the average demands can be met. But such a reservation could increase the cost of QoS loss due to frequent demand violation, and therefore, reduce the ISP's revenue and profit. On the contrary, a high ϵ, would significantly increase the cost of the service, thereby reducing the profit. Hence, it is important to determine an optimal value for the inflation factor ϵ as well as the corresponding reservation levels, ($y_{ie}(t)$ and $x_{ed}(t)$).

Clearly, the solution to the problem depends on the specifications of the demand and cost processes. We present a solution approach in the following section by considering time-invariant strategies as opposed to time-dependent solutions.

3.3. Solving the path contracting problem and risk management

Our objective is to find simple, and practical contracting strategies that can offer high expected profits. Hence, we restrict ourselves to time-invariant contracting solutions. With this simplification, the problem can be reformulated as,

$$\max \quad E[Profit] = \sum_{t=1}^{N} E\{R(t,\epsilon) - C(t,\epsilon)\}, \tag{20}$$

$$\text{s.t.} \quad \sum_{e\in\mathcal{E}} y_{ie} = \sum_{d\in\mathcal{D}} \mu_{id} \times (1+\epsilon), \quad \forall i \in \mathcal{I}, \tag{21}$$

$$\sum_{i\in\mathcal{I}} y_{ie} = \sum_{d\in\mathcal{D}} x_{ed}, \quad \forall e \in \mathcal{E}, \tag{22}$$

$$\sum_{e\in\mathcal{E}} x_{ed} = \sum_{i\in\mathcal{I}} \mu_{id} \times (1+\epsilon), \quad \forall d \in \mathcal{D}, \tag{23}$$

$$0 \leqslant y_{ie} \leqslant U_{ie}, \quad 0 \leqslant x_{ed} \leqslant W_{ed}, \tag{24}$$

$$StdDev\{Profit\} \leqslant V_{th}. \tag{25}$$

Based on the models for demand and cost discussed in the previous section, several inferences can be made about the revenue and the cost function. They are,

1. Expected revenue $E[R(t,\epsilon)]$ increases with ϵ but saturates eventually. Higher ϵ increases the chance of meeting the demand completely and thereby generating more revenue. But, the revenue cannot continue to increase, as the demand is bounded.
2. Expected cost $E[C(t,\epsilon)]$ increases with ϵ.
3. Standard deviation of revenue and cost increases with ϵ.

From 1 and 2, we conclude that the expected profit is a concave decreasing function of ϵ and from 3, we conclude the standard deviation of profit is a convex increasing function of ϵ. We use these properties to devise the following simulation-based approximation Algorithm (OPT) to determine the optimal ϵ and the corresponding time-invariant contracting strategy:

Algorithm OPT: Determine ϵ^* and optimal contracting

$\epsilon \leftarrow 0$
while $\sigma_{Profit}(\epsilon)$ > = Vth **do**
  $\epsilon \leftarrow \epsilon + \delta$
end while
$\mu^*_{Profit} \leftarrow \mu_{Profit}(\epsilon)$
$\epsilon^* \leftarrow \epsilon$
while $\sigma_{Profit}(\epsilon)$ < = Vth **do**
  $\epsilon \leftarrow \epsilon + \delta$
  if $\mu_{Profit}(\epsilon) > \mu^*_{Profit}$ **then**
    $\mu^*_{Profit} \leftarrow \mu_{Profit}(\epsilon)$
    $\epsilon^* = \epsilon$
  end if
end while

$\mu_{Profit}(\epsilon)$ and $\sigma_{Profit}(\epsilon)$ are the mean and standard deviation of the profit for a given ϵ. They are obtained by evaluating the solution of the deterministic optimization problem described below in a dynamic environment through simulation analysis.

$$\max N \times (R - C) \tag{26}$$

$$\text{s.t.} \sum_{e \in \mathcal{E}} y_{ie} = \sum_{d \in \mathcal{D}} \mu_{id} \times (1 + \epsilon), \quad \forall i \in \mathcal{I}, \tag{27}$$

$$\sum_{i \in \mathcal{I}} y_{ie} = \sum_{d \in \mathcal{D}} x_{ed}, \quad \forall e \in \mathcal{E}, \tag{28}$$

$$\sum_{e \in \mathcal{E}} x_{ed} = \sum_{i \in \mathcal{I}} \mu_{id} \times (1 + \epsilon), \quad \forall d \in \mathcal{D}, \tag{29}$$

$$0 \leqslant y_{ie} \leqslant U_{ie}, \quad 0 \leqslant x_{ed} \leqslant W_{ed}. \tag{30}$$

The above problem finds the cheapest time-invariant contracting strategy for supporting ϵ-inflated demand levels for the entire planning period, assuming that the marginal costs remain fixed at their time-averaged values. The term R denotes the revenue generated by the inflated demand levels $\mu_{id} \times (1 + \epsilon)$, and C denotes the total cost for supporting these demand levels. Since each epoch generates the same profit, $(R - C)$, the profit generated for the entire planning period is simply $N \times (R - C)$, which is maximized by choosing appropriate y_{ie}, $x'_{ed}s$. There is no constraint on the risk of profit, since there is no uncertainty in demand or the cost in this formulation. The demand level for each epoch is assumed to be $\mu_{id} \times (1 + \epsilon)$, while the marginal costs are equal to their average values. Note that the optimization problem defined above (Eq. 26) is quadratic since the objective function is quadratic in y_{ie}, $x'_{ed}s$, while the constraints are linear. Therefore, Algorithm OPT solves a sequence of quadratic optimization problems to determine the optimal solution for the original stochastic optimization problem (Eq. 20). Quadratic optimization problems of large sizes can be solved fast. Moreover, the total number of quadratic optimization problems solved is typically a hundred or two, since it suffices to choose a value of δ that is reasonably small (say 0.01). The solution provided by Algorithm OPT (ϵ^* and the corresponding contracting strategy) can be used for the next N epochs. It also serves as a guideline for admissions policy developed in Section 5.

In order to obtain the mean and standard deviation of the profit $(\mu_{Profit}(\epsilon), \sigma_{Profit}(\epsilon))$, the solution to the deterministic problem (Eq. 26) for a given ϵ, is evaluated in presence of stochastic demand and marginal costs. The following section describes the steps involved.

3.3.1. Obtaining the mean and standard deviation of profit

We consider the solution to deterministic problem (Eq. 26) for a fixed ϵ. We generate several realizations of the demand and cost processes using the mean-reverting random walk models defined earlier. For each realization, Eqs. 18 and 19 are used to determine the revenue and total cost. The statistics $\mu_{Profit}(\epsilon)$ and $\sigma_{Profit}(\epsilon)$ are obtained by averaging the profit over several realizations. In our simulations, around 20,000 instances of demand and cost processes were generated. The simulation was performed on the topology described in Section 4.2. After generating a sequence of time-invariant solutions $(y'_{ie}s, x'_{ed}s)$, one for each value of ϵ, the algorithm described in the previous section picks the solution that behaves optimally (in terms of mean–variance of profit) in a stochastic environment.

4. Simulation results and evaluation

We first solve the deterministic problem (Eq. 26) on a realistic network topology in order to gain insights on the time-invariant contracting solutions. The network topology and the contracting solution are discussed in Sections 4.1 and 4.2, respectively. Later in Section 4.3, we present the time-invariant contracting strategy obtained using our approximation algorithm. The contracting strategy achieves higher mean profit in presence of stochastic demands and costs.

4.1. Network setup

We devise a realistic network topology using Rocketfuel data [6] and the GTITM technique [7], which are considered good abstractions of the real internetwork structure. GTITM models have been extensively used for Internet topology generation [35]. We model the connectivity between the ISPs in the Internet by a GTITM random network. Each node in this network represents an ISP and each edge represents the link between the two ISPs. We choose a node (denote by x) in the graph and replace it with the topology of a real ISP obtained using Rocketfuel technology [6]. This graph would show all the internal links of an ISP and the connectivity of this ISP to other ISPs.

Rocketfuel data repository [6] provides router-level topology data for six ISPs: Abovenet, Exodus, Ebone, Sprintlink, Telstra and Tiscali. We choose the topology of Abovenet ISP, since it has been widely used for intra-domain analysis. The total number of routers and links in the Abovenet are 150 and 922 respectively. The edge routers of the Abovenet are identified as those which have low degree and higher distance from the core of the ISP. The degree and distance threshold was set such that 21 routers out of 150 are edge routers. The edge routers are classified into 12 ingress and 9 egress routers. We then identify the node in the GTITM network with the maximum degree (denoted by x), for replacing with the Abovenet topology. The egress routers of Abovenet are grouped (depending on the degree of x) based on their geographical locations and each group is connected to a neighboring ISP of x. Egress routers located in the same city are combined into a group, if needed. An intra-domain contract link between an ingress–egress pair of Abovenet is the shortest path between the two routers. Before we set up the extra-domain contract paths, each destination router is assigned to a randomly chosen ISP in the GTITM network. Now, an extra-domain contract path between an egress of Abovenet and the destination is defined as the shortest path between the two ISPs in the GTITM network.

An intra-domain contract link exists between every ingress/egress pair of Abovenet. To assign capacity values to them, we use a technique based on the Breadth-First Search (BFS) algorithm. We select the maximum-degree router in the Abovenet topology as the center node for the BFS. After running the BFS from the max-degree router, each router is assigned a BFS distance value with respect to the center node. The BFS distance value of the center node is 0. Given the BFS distances, we apply a simple strategy to assign link capacities. Let the BFS distances for routers i and j be d_i and d_j respectively. The estimated capacity of the link connecting i and j is, $C_{ij} = k[max(d_i; d_j)]$, where k is a decreasing vector of conventional link capacities. In this paper, we used: $k[1] = 40$, $k[2] = 10$, $k[3] = 2.5$, $k[4] = 0.62$, $k[5] = 0.15$, $k[6] = 0.045$, and $k[7] = 0.01$ Gbps. So, for e.g., a link between the center router and a router with BFS distance 5 will be assigned 155 Mbps as its estimated link capacity. The intuition behind this BFS-based method is that an ISP's network would have higher capacity and higher degree links towards center of its topology. The extra-domain path capacity depends on the capacities of intra-domain contract links that make this path. We define path capacity as follows: If there are d ISPs on the extra-domain path between an egress router of Abovenet and the destination router, we sample d values randomly from the pool of intra-domain capacities of Abovenet and take the minimum of those values to be the path capacity.

We use gravity models to construct a feasible demand matrix composed of e2e flows. Here, the traffic between two routers is proportional to the product of the populations of the two cities where the routers are located. The ingress and egress routers belong to the Abovenet, whose locations are given by Rocketfuel. For the destination router, we pick a location randomly from a data set of available cities and populations. The proportionality factor of the gravity model is finally adjusted to make the flow feasible.

4.2. Deterministic optimal solution

We solve the profit maximization problem (26) on the network topology described in Section 4.1. The number of ingress, egress and destination nodes are $S = 12$, $M = 9$, $K = 10$, respectively. Thus, there are 108 intra-domain contract links and 90 extra-domain paths. The demand matrix μ, capacity matrices $\mathcal{U}$ and *cal W* are obtained using Gravity models and BFS based estimations [36], respectively.

The Ramsey number is set at, $\alpha_I = \alpha_E = \alpha_S = 0.5$, assuming a competitive oligopoly market-structure. The marginal cost for the intra-domain links is set at, $c_I = 1$ unit, for extra-domain paths at, $c_E = 2$ units, and for e2e service at, $c_S = 25$ units. The value of c_E is usually higher than c_I, since an extra-domain path is made of several intra-domain contract links. Hence, we set a value of 2 for c_E, since there are at least two intra-domain links in a path. The marginal price of the e2e service, c_S, is set such that the ISP can recover the total cost. Also, we currently assume uniform c_I and c_E values for all the links and paths for ease of analysis.

We solve the optimization problem for this choice of parameters. A key set of variables of the problem are, x_{ed}, displaying the extent of extra-domain contracting needed to fulfill the e2e demand while maximizing the profit. In Fig. 5, we plot the contracted extra-domain capacities (x_{ed}), corresponding to the inflation factor, $\epsilon = 0.04$, for all the 90 extra-domain paths. The maximum allowable capacity, W_{ed}, of each extra-domain path is arranged in decreasing order. The paths with higher capacity are utilized more than the ones with lower capacity. This makes sense, since our non-linear pricing scheme for the links and paths favors low-congestion zones, as the quantity of bandwidth purchased is normalized by the maximum available capacity for price determination. Thus, paths with higher capacities tend to be cheaper and are contracted to minimize the cost.

As mentioned earlier, the value of c_S must be set such that the profit is non-negative. However, c_S depends on the parameters of the problem being solved. For the above problem instance, a c_S value of 25 was found to be sufficient. In general, this may not be true. Therefore, we obtain the c_S for which the revenue breaks even with the cost for different choices of the key parameters, c_I and c_E. The value of ϵ was kept fixed at 0.04. We will show in the next section that the optimal value of the inflation factor, ϵ, for the topology we consider turns out to be 0.04. From the solutions, it was observed that the relationship between c_S and c_E, c_I is linear, modeled approximately as $c_S = 3 \times c_I + 6 \times c_E + 5$. This relation assumes that c_E and c_I values are same for all the links and paths. In practice, the marginal costs are different for each path/link and are stochastic. Nevertheless, this relation could still act as a guideline for choosing c_S if the range of fluctuations in c_E are known.

4.3. Mean and standard deviation of profit

Fig. 4(a) and (b) shows the mean and standard deviation of profit, respectively, for a range of ϵ, when both demand and extra-domain marginal costs are stochastic. The duration of each simulation, N, is a week or $24 \times 7 = 168$ hours. The step size for the demand process, $\mu_{id}(t)$, is 5% of its time-averaged value, μ_{id}, while the step size for the marginal cost is 1% of its initial value. The time-average value of marginal cost, c_E^{ed}, was set to 2 for all the paths. The convergence parameter p for all the mean-reverting processes was set at 1/6.

The smooth curve in Fig. 4(a) is a 3rd-degree polynomial approximation of the mean profit curve. The mean profit initially increases with ϵ, but beyond a point decreases sharply. This behavior can be explained using the equation for total return (Eq. 18). The term CQ decreases sharply with ϵ when ϵ is small. But, the cost of intra-domain and extra-domain resources does not increase significantly for small values of ϵ. This is due to the fact that the extra demand can be accommodated by purchasing little extra capacity on currently used links/paths, without buying additional links/paths. Therefore, we see an increase in the mean profit initially. However, for higher values of ϵ, the cost term, $C_I + C_E$, increases sharply outweighing the decrease in CQ. Hence, the profit curve starts decreasing after a certain ϵ value. The curve attains maximum at $\epsilon = 0.04$.

Next, consider standard deviation curve shown in Fig. 4(b). The standard deviation of profit decreases as ϵ increases, initially, but increases beyond a certain ϵ value. The minimum is achieved for $\epsilon = 0.02$. The region a in Fig. 4(a) and (b) represent the trade-off region for the ISP. The mean profit can be made higher than that for $\epsilon = 0$ by choosing an ϵ in region a. But, as it can be observed from Fig. 4(b), the standard deviation of the profit could also increase. Thus, a higher mean profit is associated with a higher risk. An appropriate choice for ϵ, and the corresponding contracting strategy, would need to be made

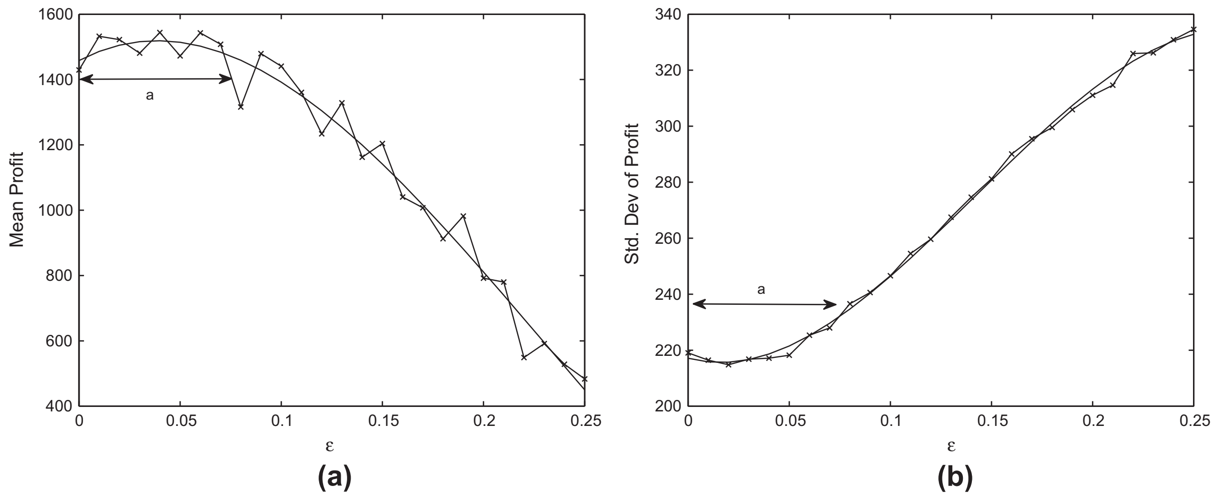

Fig. 4. (a) Mean Profit vs Inflation factor (ϵ), (b) Standard deviation of profit vs Inflation factor (ϵ). The demand and marginal cost process are modeled as mean-reverting random walk processes.

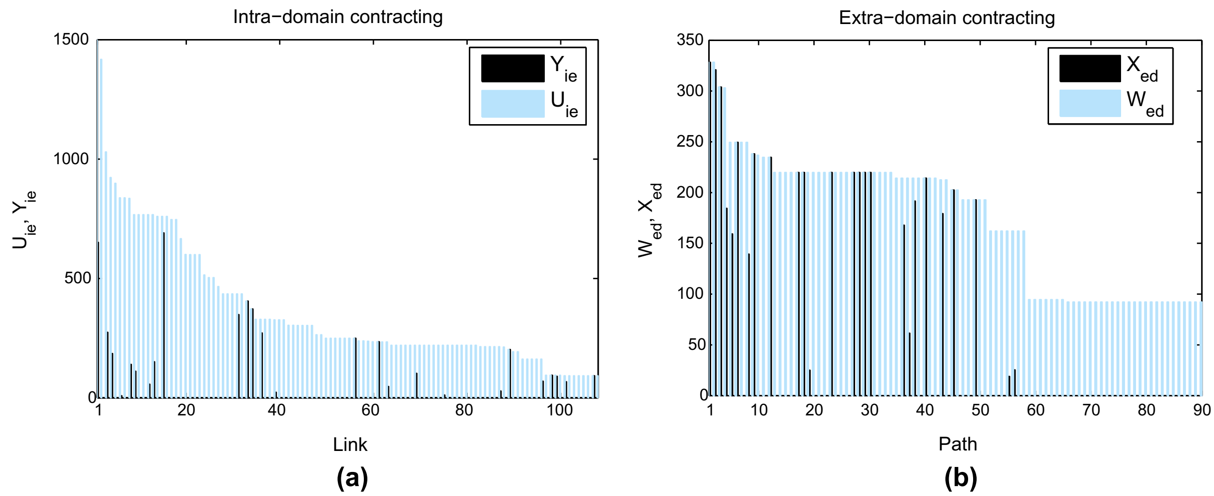

Fig. 5. owing the optimal contracting strategy, i.e., the contracted capacities for $\epsilon = \epsilon^*$. (a) Contracted intra-domain capacities, (b) Contracted extra-domain capacities.

based on the risk-aversion of the ISP. Choosing an ϵ beyond region *a* would be completely inefficient.

From Fig. 4(a), we see that, for a standard deviation threshold (*V*th) large enough, the mean profit attains maximum at 0.04, i.e. $\epsilon^* = 0.04$. Considerable increase in mean profit (around 10%) can be achieved by contracting for $\epsilon = 0.04$, rather than $\epsilon = 0$. The corresponding optimal contracting can be obtained by the solving the optimization problem in Eqn. (26). Fig. 5 shows the intra-domain and the extra-domain contracting for $\epsilon = \epsilon^*$, respectively. Note that the extra-domain paths with lower capacity remain largely unused, while the lower capacity intra-domain links tend to be used. This is due to the fact that the extra-domain paths are costlier than the intra-domain links. Therefore, the intra-domain links are used in such a way that the extra-domain cost is reduced, instead of using only the higher capacity links to reduce the intra-domain cost. Thus, in a dynamic environment, it is inefficient to reserve bandwidth just to meet the mean demands. Significant increase in mean profit can be achieved for the same level of risk by contracting for inflated mean demands (4% for the above case). Next, we consider a different demand model to study the effect of demand model on the optimal contracting strategy. The new model, called as time-of-day model is developed based on the Internet traffic data in [37]. In this model, each s-d demand has a daily pattern, as shown in the Fig. 3(b). The demand shows fluctuations over short time-scales (hourly), while following a daily pattern. To model this behaviour in the s-d traffic, we consider the continuous-time mean reverting process given by (Eq. 6), with time-varying means, $\bar{\mu}_{id}(t)$, instead of constant mean. The mean function $\bar{\mu}_{id}(t)$ is periodic, but weekends have lower mean than the weekdays. The model for marginal cost process is kept unchanged (i.e. mean-reverting process).

Fig. 6(a) and (b) show the mean and standard deviation of profit for different values of the inflation factor, ϵ. We find that the mean profit increases with ϵ till a certain value, and then begins to drop, similar to Fig. 4(a). However, the standard deviation increases steadily with ϵ. Thus, considerable increase in mean profit can be achieved by contracting for higher ϵ. But, in this case, there is no possibility of achieving a risk that is lesser than the risk resulting from contracting for mean-demands ($\epsilon = 0$). Also, we find that the optimal epsilon (=0.16) for the time-of-day model is higher than that for the simple mean-reverting model shown in 4(a).

In the next section, we study the effect of increased demand and cost volatility (or standard deviation) on the contracting strategy.

4.4. Effects of demand and cost volatility

The standard deviation of the demand and cost process is increased by reducing the value of parameter p of the mean-reverting random walk (Eq. 1) that models demand and cost. From (Eq. 1), we see that, for $p = 0.5$, the process $\mu_{id}(t)$ takes the values $\mu_{id} + \delta_m$ and $\mu_{id} - \delta_m$ alternately during successive time instants. Hence, the standard deviation is δ_m. But, for $p = 0$, the process $\mu_{id}(t)$ is a symmetric random walk function with standard deviation, $\sigma_m(t) = \delta_m \times \sqrt{t}$. Thus, decreasing p from 0.5 to 0 increases the standard deviation.

Fig. 7(a) and (b) shows the effect of increased volatility on the curves for mean and standard deviation of the profit. As the volatility is increased (i.e. p is reduced) for a fixed ϵ, we find that the mean profit reduces, while the standard deviation increases. Fig. 7(a) also shows the optimal inflation, ϵ^*, for different volatilities (i.e. different values of p). The constraint on the standard deviation (or risk) of profit is shown in Fig. 7(b). As expected, we see that the optimal inflation factor ϵ^* increases as the demand and cost becomes more volatile. If the reservation (or the inflation factor) is not increased for increased demand volatility, the expected QoS loss would increase, leading to reduced profit levels. The optimal contracting (for fixed p) can be obtained by solving the problem in Eq. (26) for $\epsilon = \epsilon^*$.

4.5. Correlation and failure in the paths

In this section, we extend our previous models for revenue and marginal cost in order to account for the correlation among the extra-domain paths, and the possibility of path failures. Correlation between paths and path failures do not make the optimization problem any different, however, the optimal contracting solution could be different. The motivation for correlated marginal costs stems from the fact that the extra-domain paths could pass through common ISPs. Any change in the cost of service of this ISP would lead to an equivalent change in the marginal cost of all the extra-domain paths that requires the services of this ISP. Therefore, we now consider correlated marginal cost processes, $c_E^{ed}(t)$, instead of independent processes. For ease of generation of correlated random processes, we consider the continuous mean-reverting process defined by (Eq. 6). The degree of correlation between two extra-domain paths is defined by

$$\rho_{ed,e'd'} = \frac{\text{Number of ISPs common to paths } ed \text{ and } e'd'}{\text{Diameter}}. \quad (31)$$

The above definition leads to a symmetric correlation matrix ρ. $\rho_{ed,ed}$ is defined to be 1.

The extra-domain contract paths are usually available for service throughout the entire planning period. However, occasionally, a path could become unavailable due to link failures or violation of contract terms by some participating ISP. During those epochs, the ISP will experience a sudden increase in the net QoS loss due to capacity deficit along some of the extra-domain paths. Thus, there will be a drop in the revenue due to increased Cost of QoS loss. But, the marginal cost would continue to change according to

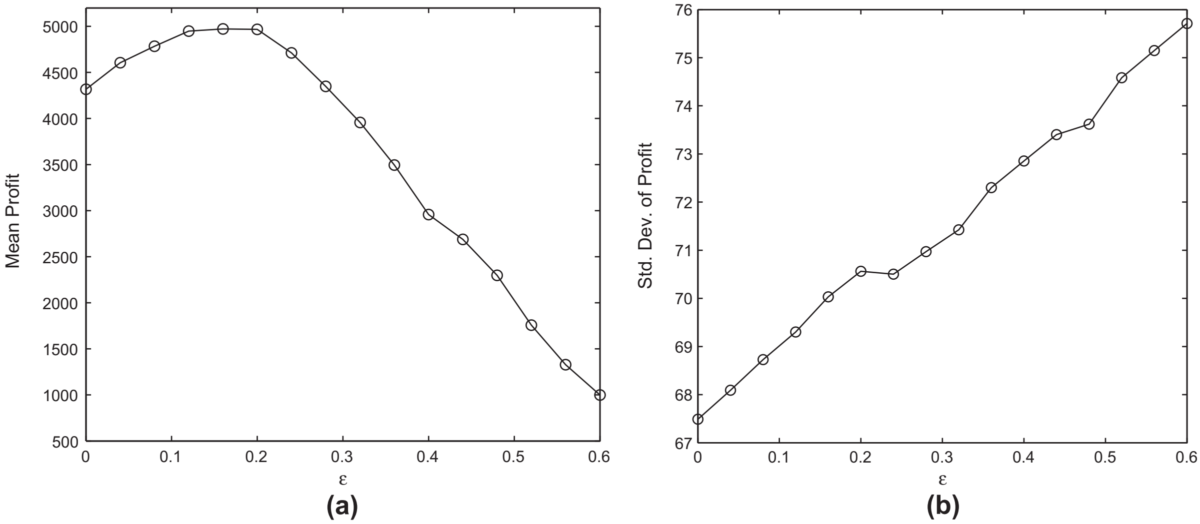

Fig. 6. (a) Mean Profit vs Inflation factor (ϵ), (b) Standard deviation of profit vs Inflation factor (ϵ). The demand process evolves using the time-of-the-day model, while the marginal cost follows mean-reverting random walk process.

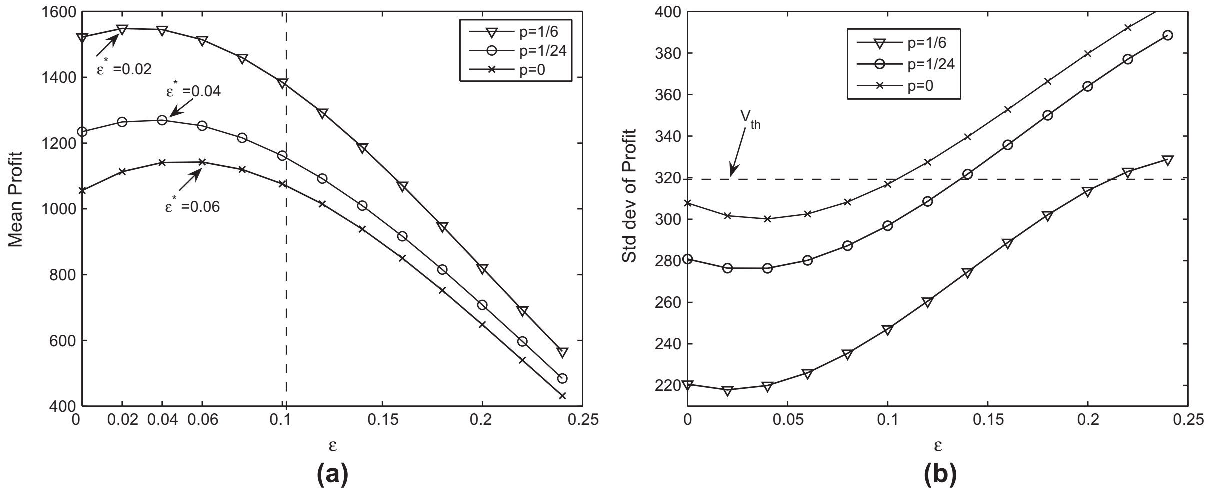

Fig. 7. (a) Mean Profit vs ϵ, (b) Standard deviation of profit vs ϵ. The standard deviation of the demand and cost process decreases with the parameter p.

the mean-reverting process, even when an extra-domain path fails temporarily. This assumption is reasonable, since the extra-domain path contracts are set-up at the beginning of the planning period. In our failure model, ISPs fail independently of each other randomly and for random durations of time during the planning period. Moreover, an ISP can fail and recover several times during the planning period by a failure and recovery rate. At any epoch t, all the extra-domain paths that pass through a failed ISP, are made unavailable. The paths recover to their initial state after failure. They do not operate in degraded mode.

As before, we determine the solution for different levels of inflation factor, ϵ, and evaluate their performance in the presence of stochastic demand and costs. The difference now is that the increments in marginal costs are normal (Eq. 6), and correlated. Moreover, in the new scenario, the paths can fail at random instants, leading to decrease in the net revenue. We use (Eq. 6) and the correlation matrix ρ to generate correlated processes, $c_E^{ed}(t)$. The net data loss for time t, $Q(t)$ (Eq. 8), is calculated after setting the contracted capacity term, $x_{ed}(t)$, to 0 for all those paths that are currently unavailable.

Fig. 8 shows the mean and standard deviation of profit for several ϵ values using independent and correlated marginal costs. Observe that the mean profit decreases with the introduction of path failures and correlation. This behavior can be explained using Figs. 5 and 9 that show the optimal extra-domain contracting and the failure of the paths, respectively. We see that several contracted extra-domain paths experienced non-zero failure periods during the planning period. The revenue generated by these contracted paths drops to zero during failures, while the ISP still incurs cost of contracting. Thus, the expected profit reduces. On the other hand, the standard deviation of profit increases in presence of path failures and correlated marginal costs. This can be explained using the equation for the total cost. Since the total extra-domain cost is an aggregation of individual path costs, the presence of non-zero covariance terms increases the total variance (and the standard deviation) of the profit.

Though the mean profit reduces in presence of correlation and path failures, the curve still peaks at the same value of ϵ. Similarly, the standard deviation curve attains a minimum at the same value of ϵ. These results suggest that the contracting solution developed for independent cost model is robust in the presence of path failures and correlation.

5. Implementing the end-to-end contracting and pricing strategy

5.1. Resolving path-vector contracts to admission control policies

So far, we presented a long-term contracting strategy for providing e2e services over the Internet. The proposed time-invariant strategy minimizes the risk of profit by contracting enough bandwidth to support e2e demands exceeding their time-averaged levels. The exact level of contracting on the paths and links can be obtained by solving a static optimization problem, given the models for e2e demand and cost. The strategy is easy to implement in practice, as the ISP can simply reserve enough capacity on its own intra-domain links and enter into long-term path contracts with the neighboring ISPs offering fixed level bandwidth. However, it should be noted that the solution does not specify any admission control policy for the ISP to determine whether or not an incoming demand (and what fraction of it) must be accepted while maintaining satisfactory levels of profit and quality of service delivery. In this section, we focus on devising such an admission control policy.

While developing the revenue models and the contracting strategy, it was assumed that the incoming e2e demand is always accepted. This approach could lead to very high levels of QoS loss during some epochs, which is undesirable in practice. It might be better to reject an incoming demand completely or partially, in order to keep the QoS loss at a satisfactory level. Moreover, at times the incoming demand may not bring enough revenue to the

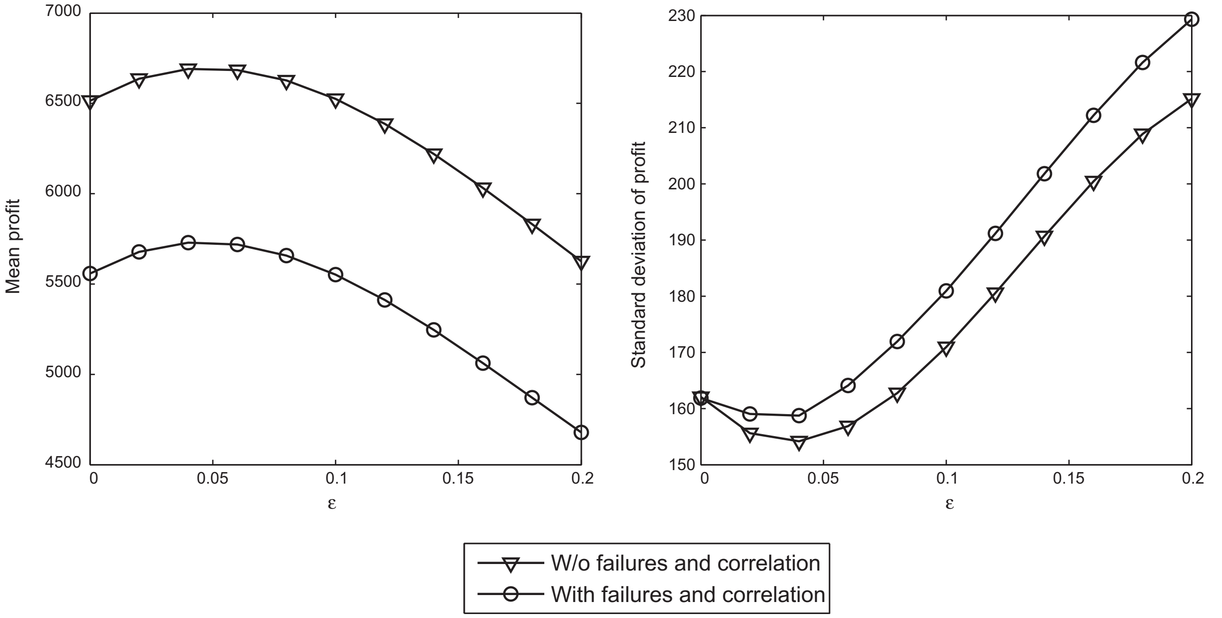

Fig. 8. Mean Profit and Standard deviation of profit vs ϵ with failures and correlation

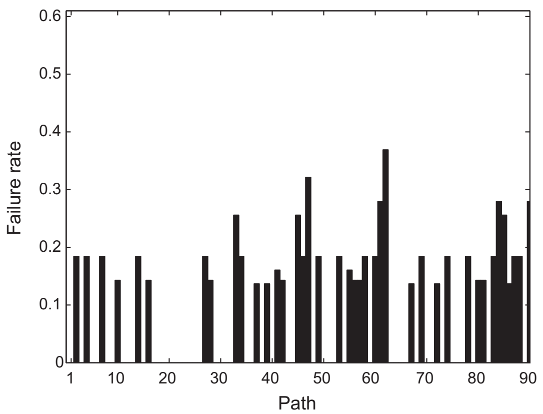

Fig. 9. Figure showing the fraction of time for which an extra-domain path failed during the planning period.

ISP while also contributing to some (though not severe) degradation in the QoS loss. Again, it would be better to reject (partially or completely) such requests for service. Thus, the decision to accept or reject an incoming demand is critical to maintain the congestion and economic state of the network at a satisfactory level. We develop an admission control policy based on the revenue and QoS loss terms defined earlier.

Let us assume that the e2e demand for all source–destination pairs remain at their average levels, i.e. μ_{id}, until time t, so that the net QoS loss $Q(t') = 0$ for $t' \leqslant t$. At time $t+1$, let there be an additional demand W for a particular source–destination pair $p-q$. The net incoming traffic at time $t+1$ for the pair $p-q$ is $\mu_{pq}(t+1) = \mu_{pq} + W$. Let the demand for the other ingress-destination pairs remain the same, i.e. $\mu_{id}(t+1) = \mu_{id}$, for $i \neq p$.

In order to decide whether the additional demand W be accepted or not, the ingress station p considers the impact of the decision on the congestion state (i.e. QoS) of network and the revenue brought about by the additional demand. Naturally, the demand must be accepted if the addition of this demand does not increase the QoS loss significantly. Assume that, all the other ingress stations (i.e. excluding p) have already made their decisions before the ingress station p. Let the accepted demands for time $t+1$ for these i-d pairs be $\mu'_{id}(t+1)$ (where $i \neq p, d \neq q$). Now, suppose the ingress node p decides to accept this demand, the net accepted demand for p-q becomes $\mu'_{pq}(t+1) = \mu_{pq} + W$. The QoS loss between the points p and q, due to the addition of this demand is given by,

$Q^{pq} = max(Q_I(p,t+1), Q_E(q,t+1))$ where,

$$Q_I(p,t+1) = \left\{\sum_{d=1}^{k}\mu'_{pd}(t+1) - \sum_{e=1}^{m}y_{pe}\right\}^+, \text{ and } Q_E(q,t+1) = \left\{\sum_{i=1}^{s}\mu'_{iq}(t+1) - \sum_{e=1}^{m}x_{eq}\right\}^+.$$

Note that Q^{pq} considers the increase in QoS loss caused by the additional demand, both within the ISP's domain as well as outside the domain. If the resulting QoS loss, Q^{pq}, is higher than the threshold (Qth), it implies that this traffic would seriously affect the existing traffic flowing in the network. In this case, the ingress p should reject the additional demand.

On the contrary, if Q^{pq} is less Qth, the demand could be accepted without affecting the current traffic. But, in order to accept the demand, it should also generate sufficient revenue to the ISP. The net revenue for tth epoch, $R'(t)$, is a function of the accepted demands $\mu'_{id}(t)$. $R'(t)$ can be calculated using the (Eq. 18) after replacing $\mu_{id}(t)$ by $\mu'_{id}(t)$. Similarly, $R'(t+1)$ can be calculated if the accepted demands $\mu'_{id}(t+1)$ are known for all i, d. Assuming that $\mu'_{id}(t+1)$ is available for all $i \neq p$, the ingress p sets $\mu'_{pq}(t+1) = \mu_{pq} + W$ and computes the increase in the revenue brought by the new demand, i.e. $R'(t+1) - R'(t)$. If the increase in revenue is not considerable (higher than Rth), the ingress p can reject demand.

One issue with the above admission control algorithm is that it assumes that only one i-d pair $(p - q)$ experiences an increase in demand, while the others remain the same. However, in the demand model, each user demand, $\mu_{id}(t)$, was assumed to evolve independently according to a mean-reverting random walk process. Therefore, at any given time $t + 1$, the demand levels at several ingress stations can increase or decrease compared to their levels at the tth epoch. The value of W for these ingress stations is accordingly $+ve$ or $-ve$. We solve the problem of multiple incoming demands as follows: Clearly, if W is $-ve$, the ingress must accept the decrease in demand. Therefore, consider only the case where $W > 0$. If $W > 0$ at several ingress stations, the ingress stations cannot make decisions simultaneously, since it was assumed that the decisions of other ingress stations are available before a particular ingress makes a decision. We propose a simple heuristic solution that the ingress stations make their decisions one after the other based on some priority. The priority can be randomly changed later to make it fair for all the ingresses.

The proposed admission control algorithm was run on the topology described in Section 4.2. Fig. 10 shows an instance of the incoming and accepted demand evolutions for the ingress-destination pair (11, 1). The threshold Rth was set to 0, and two values were considered for Qth, namely 0 and 2. We observe that, during some epochs, the incoming demands gets accepted (e.g. $t = 60$), while during some other epochs (e.g. $t = 100$), the demand is rejected. Therefore, there is a gap between the incoming and accepted demand waveforms for both the values of Qth. Moreover, the gap increases, when QoS threshold, Qth, is reduced. As shown in Fig. 10, the average and inflated demand levels for the ingress-destination pair (11, 1) are 206 and 215 respectively. The incoming demand gets accepted always when its value is below the average level. On the other hand, if the demand is above the inflated average level, it gets rejected most of the times. However, if the demand is in between these two levels, it gets accepted sometimes and rejected other times.

Table 2 summarizes the performance of two admission policies, P_1 and P_2. For both P_1 and P_2, the revenue threshold Rth was set to zero. The QoS threshold for P_1 was set to zero for all ingress-destination pairs. However, for P_2, the QoS threshold for an ingress-destination pair was set to 1% of the average demand between the pair. Table 2 shows that the expected profit reduces significantly for P_1, where the QoS threshold is too stringent. This is due to frequent rejection of incoming demand and the revenue coming from it. However, for P_2, the expected profit is higher than the profit that we can achieve without any admission control. As shown in Table 2, the QoS loss in the network is very high (54000), without admission control. Therefore, the net revenue is reduced due to high cost of QoS loss. P_2 performs the best among the three, achieving high expected profit at satisfactory level of QoS loss, but only at the cost of increased risk of profit.

The proposed admission control policy acts as a tool for the ISP to control the economic and congestion state of the network. The actual profit the ISP can realize by implementing the contracting solution along with the admission

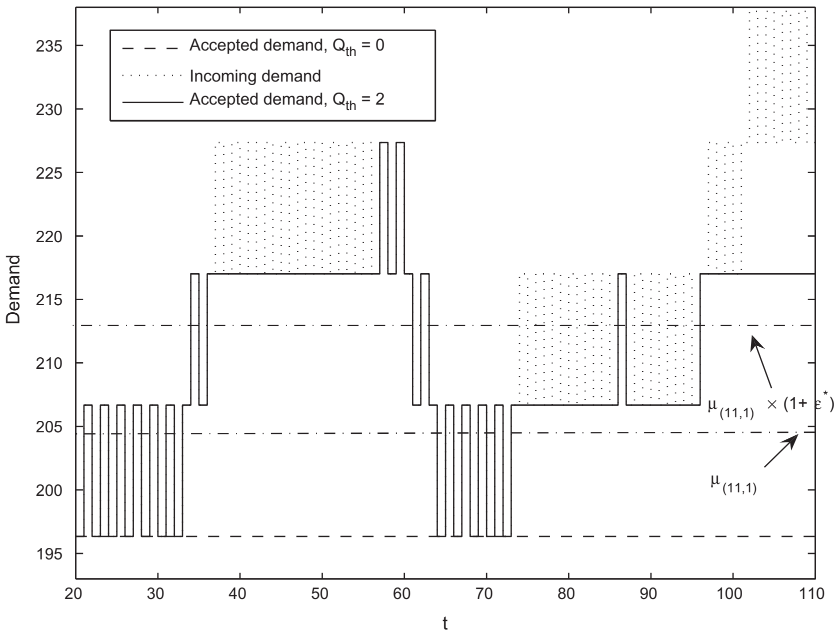

Fig. 10. Incoming and Accepted demand for the ingress-destination pair (11, 1). The revenue threshold, Rth = 0.

Table 2
Performance of the admission control policy.

Policy	Mean profit	Std.dev	Mean QoS loss
P_1: Qth = 0	31	237.83	0
P_2: Qth = 1%	1670	230.84	42
W/o admn ctrl	1540	216	54000

control policy depends on the choice of the revenue and QoS thresholds. Note that implementing the admission control system only needs exchange of information among the ingress/egress nodes in an ISP and later verifying whether the thresholds are satisfied. Therefore, it is not costly. However, repeating this process for each incoming flow might be computationally expensive. Therefore, the admission control is performed only once for each epoch for each ingress–egress pair.

5.2. Exposure of ISP Internals

The ISPs only need to share the information about the contract links, so that end-to-end paths can be constructed. This does not mean that they will have to cooperate by exposing their internal information. End-to-end contract paths can be established as long as some very high level information is available about the edge-to-edge (g2g) contract links: ingress point, egress point, quality (i.e., effective capacity), and price (i.e., dollars per Mb). Except the "egress point", all the other parameters are being used extensively in SLA establishments in the current Internet. However, the time-scale of the current SLAs is too long. We argue that finer time-scale SLAs will help the inter-ISP economics and establish a healthy market for bandwidth exchange and management.

Notice that day-to-day information about ISP networks is publicly posted even today even though there is no contracted requirement for ISPs to share that information. Please see [38] as an example. We suggest that ISPs are already motivated to share such high level performance of their networks and will be even more motivated to provide finer granularity (i.e. between more specific points of their networks at more specific times) performances of their networks as long as there is significant enough revenue potential.

6. Conclusion

We developed a bandwidth (QoS) contracting and pricing strategy for e2e services using the path-vector based contracting approach and the Ramsey pricing model. The proposed time-invariant contracting strategy achieves high expected profit in a dynamic environment for a given level of risk, by contracting for an inflated average demand levels rather the average demand levels. The contracting strategy utilizes the high capacity intra-domain links and extra-domain paths, where some low-capacity intra-domain links get used as a cost trade-off since the extra-domain paths are costlier than the intra-domain links.

Examination of the contracting problem under different demand and cost volatility parameters shows that our solution framework is robust for degree of demand and cost volatility. Our detailed simulation-based study using a realistic topology suggested that the inflation factor and hence the contracted capacities should be increased as the network becomes more volatile. Furthermore, counter to intuition, we find that the solution obtained from our stochastic optimization framework under no failures continues to be optimal under correlation and failure of extra-domain links. Finally, the admission control policy implementing the contracting strategy is seen to often admit demand levels much higher than the targeted inflated demand used to determine the contracting strategy, and performs remarkably better than when there is no admissions control implemented.

There are several interesting issues to be studied in the future, such as, constructing better models for cost and demand fluctuations and extending the optimization framework to study non-stationary changes in demand and cost uncertainty. The end-to-end contracting solution that we develop does not directly provide guarantees on the experienced delay or QoS metrics other than the bandwidth. In future work, we plan to extend our formulation and analysis to incorporate additional QoS metrics such as the delay.

Acknowledgement:

This work was supported by the National Science Foundation through the NeTS-FIND program (awards CNS-0721609, CNS-0721600, CNS-0831830 and CNS-0831957).

References

[1] M. Yuksel, A. Gupta, S. Kalyanaraman, Contract-switching paradigm for internet value flows and risk management, in Proceedings of IEEE Global Internet Symposium, 2008.
[2] Rfc1771- a border gateway protocol 4 (bgp-4). 1995.
[3] Scott Shenker P. Brighten Godfrey, Igor Ganichev and Ion Stoica. Pathlet routing, in: Proceedings of SIGCOMM, 2009.
[4] Ramesh Johari Nick Feamster, Vijay Vazirani, Agora: A market for internet connectivity, in: Workshop on Programmable Routers for Extensible Services of Tomorrow (PRESTO), Princeton, NJ, 2007.
[5] David Wetherall Ratul Mahajan, Thomas Anderson, Negotiation-based routing between neighboring isps, in: Proc. USENIX Symposium on Networked Systems Design and Implementation, 2005.
[6] N. Spring, R. Mahajan, D. Wetherall, Measuring isp topologies with rocketfuel, in: ACM SIGCOMM, 2002.
[7] <ttp://www.cc.gatech.edu/projects/gtitm/>.
[8] G. Huston, Internet Performance Survival Guide: QoS Strategies for Multiservice Networks, John Wiley, Hoboken, NJ, 2000.
[9] R. Braden et al, Resource Reservation Protocol (RSVP)-V1 functional Spec. IETF Internet RFC 2205, September 1997.
[10] R. Braden et al, Integrated services in the Internet architecture: An overview. IETF Internet RFC 1633, June 1994.
[11] I. Stoica, S. Shenker, H. Zhang, Core-stateless fair queueing: A scalable architecture to approximate fair bandwidth allocations in high speed networks, IEEE/ACM TON 11 (1) (2003) 33–46.
[12] E. Rosen et al., Multiprotocol label switching architecture, IETF Internet RFC (2001) 3031.
[13] D. Awduche, A. Chiu, A. Elwalid, I. Widjaja, X. Xiao, Overview and principles of internet traffic engineering, RFC (2002) 3272.
[14] Z. Duan, Z.L. Zhang, Y.T. Hou, Service overlay networks: SLAs, QoS, and bandwidth provisioning, ACM/IEEE Transactions on Networking (TON) 11 (6) (2003) 870–883.
[15] L. Subramanian, I. Stoica, H. Balakrishnan, R.H. Katz, OverQoS: offering internet QoS using overlays, CCR 33 (1) (2003) 1–16.
[16] Akamai Inc. <http://www.akamai.com>, 2004.

[17] GÉANT. <http://www.geant.net/server/show/conwebdoc.500>, 2005.
[18] I. Paschalidis, Y. Liu, Pricing in multiservice loss networks: static pricing, asymptotic optimality, and demand substitution effects, IEEE/ACM Transactions on Networking 10 (3) (2002) 425–438.
[19] A.M. Odlyzko, Internet pricing and history of communications, AT& T labs, 2000.
[20] J.K. MacKie-Mason, H.R. Varian, Public Access to the Internet, chapter Pricing the Internet, MIT Press, Boston, MA, 1995, pp. 89–100.
[21] F.P. Kelly, A.K. Maulloo, D.K.H. Tan, Rate control in communication networks: shadow prices, proportional fairness and stability, Journal of the Operational Research Society 49 (1998) 237–252.
[22] A. Gupta, D.O. Stahl, A.B. Whinston, Internet Economics, chapter Priority pricing of Integrated Services networks, MIT Press, Boston, MA, 1997, pp. 323–352.
[23] U. Savagaonkar, E.K.P. Chong, R.L. Givan, Online pricing for bandwidth provisioning in multi-class networks, Computer Networks 44 (6) (2002) 835–853.
[24] X. Wang, H. Schulzrinne, Comparative study of two congestion pricing schemes: auction and tâtonnement, Computer Networks 46 (1) (2004) 111–131.
[25] C. Courcoubetis, V.A. Siris. Managing and pricing services level agreements for differentiated services, in: IEEE IWQoS, 1999.
[26] E.W. Fulp, D.S. Reeves, Bandwidth provisioning and pricing for networks with multiple classes of service, Computer Networks 46 (1) (2004) 41–52.
[27] T. Li, Y. Iraqi, R. Boutafa, Pricing and admission control for QoS-enabled Internet, Computer Networks 46 (1) (2004) 87–110.
[28] A.J. O'Donnell, H. Sethu, Congestion control, differentiated services, and efficient capacity management through a novel pricing strategy, Computer Communications 26 (13) (2003) 57–1469.
[29] V. Firoiu, J.-Y. Le Boudec, D. Towsley, Z.-L. Zhang, Theories and models for internet quality of service, Proceedings of the IEEE 90 (9) (2002) 1565–1591.
[30] B. Teitelbaum, S. Shalunov, What QoS research hasn't understood about risk, in: Proc. of ACM SIGCOMM Workshops, 2003, pp. 148–150.
[31] A. Gupta, S. Kalyanaraman, L. Zhang, Pricing of risk for loss guaranteed intra-domain internet service contracts, Computer Networks 50 (2006) 2787–2804.
[32] A. Gupta, S. Kalyanaraman, L. Zhang, A spot pricing framework for pricing intra-domain assured bandwidth services, Int. J. Inf. Tech. Decision Mak. 4 (1) (2005) 35–58.
[33] A. Gupta, L. Zhang, Pricing for end-to-end assured bandwidth services, Int. J. Inf. Technol. Decis. Mak. 7 (2) (2008) 361–389.
[34] R.B. Wilson, Nonlinear Pricing, Oxford University Press Inc.,, New York, NY, 1993.
[35] Ye Wang, Hao Wang, Ajay Mahimkar, Richard Alimi, Yin Zhang, Lili Qiu, Yang Richard Yang. R3: resilient routing reconfiguration, in: Proceedings of the ACM SIGCOMM, 2010.
[36] W. Liu, Hasan T. Karaoglu, A. Gupta, M. Yuksel, K. Kar. Edge-to-edge bailout forward contracts for single domain internet services, in: Proceedings of IWQoS 08, 2008.
[37] Gregory J. Miller Kevin Thompson, Rick Wilder, Widearea internet traffic patterns and characteristcs, in: IEEE Network, 1997.
[38] <http://ipnetwork.bgtmo.ip.att.net/pws/index.html>.

Praveen Kumar M is a Ph.D. student with the Department of Electrical, Computer and Systems Engineering at Rensselaer Polytechnic Institute (RPI) since Fall 2008. His research interests include wireless sensor networks, network economics, and IT service management.

Aparna Gupta is a faculty member in the Lally School of Management and Technology at Rensselaer Polytechnic Institute. Her research interest is in financial decision support, risk management, and financial engineering, as well as applying mathematical modeling and financial engineering techniques for risk management in technology-enabled services, such as, Internet service provisioning, future Internet design, wireless spectrum markets, and technology enabled service contracts.

Murat Yuksel is currently an Assistant Professor at the CSE Department of The University of Nevada – Reno (UNR). He received a B.S. degree from Computer Engineering Department of Ege University, Izmir, Turkey in 1996. He received M.S. and Ph.D. degrees from Computer Science Department of RPI in 1999 and 2002. He has been a Postdoctoral Research Associate and a member of Adjunct Faculty at RPI during 2002–2006. Dr. Yuksel's research interests are in the area of computer communication networks with a focus on protocol design, network economics, wireless routing, freespace-optical mobile ad hoc networks (FSO-MANETs), and peer-to-peer networks.

Koushik Kar has been a faculty member at the Department of Electrical, Computer and Systems Engineering at Rensselaer Polytechnic Institute (RPI) since 2002. He obtained the Ph.D. degree in Electrical and Computer Engineering from the University of Maryland, College Park, and his B.Tech. degree in Electrical Engineering from the Indian Institute of Technology, Kanpur, India, in 1997. His research is in modeling, analysis and optimization questions for the Internet as well as wireless networks, including study of medium access control, scheduling and routing issues in these networks. Dr. Kar received the CAREER award from the National Science Foundation in 2005.