Economics of Network Pricing with Multiple ISPs

Srinivas Shakkottai R. Srikant

***Abstract*—In this paper we examine how transit and customer prices are set in a network consisting of multiple ISPs. Some ISPs may be geographically co-located so that they compete for the same set of end users. We examine the existence of equilibrium price strategies in this situation and show how positive profit can be achieved using threat strategies. It is shown that if the number of ISPs competing for the same customers is large then it can lead to price wars. ISPs that are not geographically co-located may not directly compete for users, but are nevertheless involved in a non-cooperative game of setting access and transit prices for each other. We study how such ISPs are linked economically through transit ISPs by considering a multi-stage game. We also consider the economics of private exchange points and show that they could become far more wide spread then they currently are.**

I. INTRODUCTION

***Index Terms*— Internet economics, pricing, repeated games, Nash equilibrium.**

The Internet is a heterogeneous body of privately owned infrastructure. Roughly speaking, it consists of two types of networks: (i) densely meshed networks in geographically localized regions which specialize in providing consumers with connection points to the network and (ii) networks traversing large geographical distances which provide connectivity between the local networks [1]. All the networks are connected by means of an identical (or at least inter-operable) protocol stack agreed upon by the Internet Engineering Task Force (IETF). Figure 1 illustrates the Internet as it looks like today. There are local Internet Service Providers (ISPs) providing services in small regions and transit ISPs which transfer data between local groups. The groups exchange data with each other at Network Access Points (NAPs). Each transit ISP would have a point of presence in each of the regions that it is interested in providing transit service. A NAP may be provided as a public resource or may be privately funded, with each constituent paying some charge to the NAP for traffic exchange [2]. In the figure, the small "clouds" show local ISPs as well as the points of presence of transit ISPs. The large transit clouds represent multiple transit (either in parallel or in series) ISPs linking the regions. The NAPs are shown as routers linked by a common media.

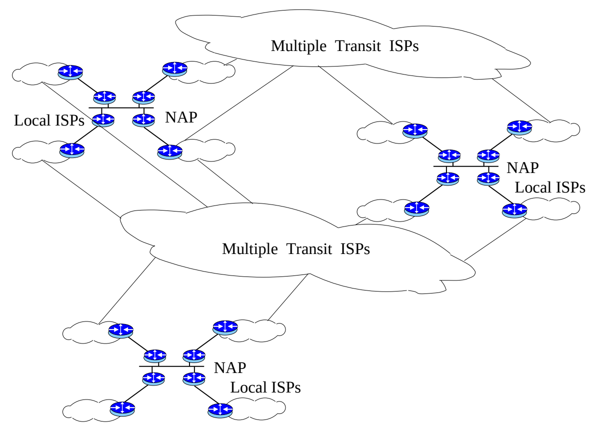

Fig. 1. The structure of the Internet today consisting of local and transit ISPs.

A topic of current interest is the development of good pricing models for the Internet. Most ISPs currently employ flat rate pricing, i.e., end-users pay a fixed sum of money every month for unlimited access (most broadband providers and dial-up services use this scheme). Others, following the telephony model, price the time spent connected to the Internet (some dial-up services are priced this way). They charge an hourly rate to end-users. Still others charge based on actual bytes transferred (many Australia/New Zealand ISPs use this scheme [3]). Pricing models for Internet economics have been discussed in [1], [4], [5]. The authors discuss how pricing may be used to provide different qualities of service and argue that differentiated pricing for different types of data transfers might be a good approach.

While some aspects of the traditional telephony model [6] may provide a starting point for economic analysis of the Internet, one must keep in mind that flows on the Internet are usually biased very heavily in one direction, whereas telephone calls impose equal loads in both directions. The transfer of data in the Internet may be assumed to be unidirectional in most cases. We assume that traffic originates at websites and terminates at end-users. In [7] this model is used to find a Nash equilibrium solution for prices charged to websites and end-users.

Research supported in part by DARPA grant F30602-00-2-0542 and NSF grant ANI-0312976

Department of Electrical and Computer Engineering and Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Emails: (sshakkot,rsrikant)@uiuc.edu

There is also another difference in economic interaction between players on the Internet as compared to the telephony model. In the telephony model a single party, usually the origin, funds the transport of traffic on the complete path from sender to receiver. Thus, each receiving ISP in the path of the flow charges each transmitting ISP a termination charge. This is called *bilateral settlement*. A bilateral settlement in which two ISPs reciprocally set their termination charges to zero is called a *peering* arrangement.

In the Internet, there is also a hierarchy of providers [1]. The sender does not fund the end-to-end transport of traffic, only a part of the path is sender funded. Once the traffic passes beyond the sending ISP's service funded domain, the receiver implicitly assumes funding responsibility for the traffic and the second part of the complete transport path is funded by the receiver. Thus, the set of connectivity paths within the Internet can be seen as a collection of path pairs, where the sender funds the initial path component and the receiver funds the second terminating path component. The hierarchical model of economic interaction is illustrated in Figure 2 for a particular direction of traffic. It is seen that higher tiers charge regardless of traffic direction. The higher level ISP is said to provide a *transit* service to the lower levels. However, the hierarchy is not rigid. Traffic could be exchanged within local groups within a region. ISPs in different regions could also exchange traffic at *private exchanges*. Companies such as Equinix [8] provide the infrastructure for the establishment of such exchanges. These are indicated in Figure 2 by a circle containing a 'P'. Thus, traffic need not traverse the whole of the hierarchy. It is for this reason that only the Tier-I ISPs (which do not pay transit charges to any other ISPs) and the local groups (which potentially have to pay for all traffic to and from their infrastructure and compete for the same customers) are well defined. Understanding economic interactions between the different groups is the main focus of the paper.

Several models have been proposed to study peering versus transit relations [9], [10], [11], [12], [2]. A problem with some of these models is that they do not take into account the fact that the Internet is separated into local and transit ISPs with different economic interests. In this paper we divide ISPs into *local* ISPs which are co-located in a small geographical region (for example Region 1 in Figure 2) and compete for the same customers and *transit* ISPs (shown as Tier-I and Tier-II in the economic hierarchy) which transfer traffic between local ISPs. We then study the economic interaction of different groups assuming a variable demand for traffic by the websites and end-users. We illustrate our ideas using a byte-wise pricing model. However, our results apply to flat rate pricing as well.

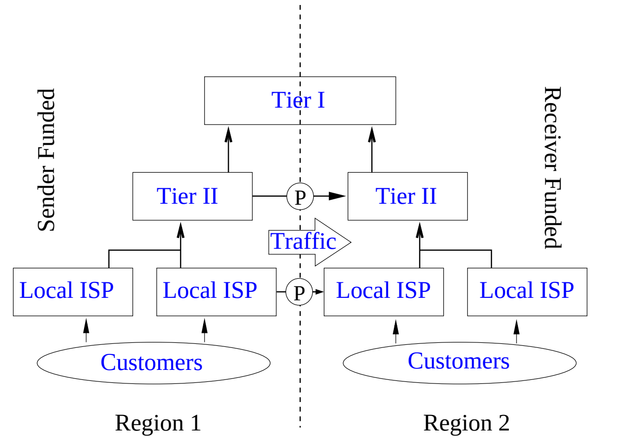

Fig. 2. Illustrating the economic hierarchy in the current Internet for traffic from Region 1 to Region 2. The solid arrows indicate monetary transfers. The arrows pointing upwards indicate the fact that the higher tiers charge regardless of traffic direction. Bilateral settlement may be accomplished at private exchanges as shown.

Local ISPs

In our first model, we study the economic interactions of local ISPs of a region with each other, which we model as a repeated game with discount factor $\alpha \in (0, 1)$. The discount factor is the number by which a payoff at the next step must be multiplied in order to obtain its value at the current step. For instance, a discount factor of $\frac{1}{2}$ would mean that the same payoff would be worth half as much in the next step, one quarter as much two steps later and so on.

We show that for local ISP interaction that there exists an optimal scenario where all ISPs peer with each other and jointly maximize their profits. In a non-cooperative game situation, we show that this optimal price can be enforced by the threat of lowering prices to a minimal value. We show that the discount factor at which the threat fails to work scales with the number of ISPs N, as $\frac{N-1}{N}$ and converges to unity as N gets large. This means that ISPs could co-exist in two possible ways:

1) There could either be a small number of ISPs interested in a short term profit or a large number which are all interested in long term profits (discount factor close to unity). They would be content to peer with each other and split the customer revenues amongst themselves.
2) Some or all of a large number of ISPs could be interested in short term gains and try to set prices below others to make quick profits by getting a larger customer share. We show that the natural outcome of

the low discount factor scenario is that prices fall to the bare minimum and all profits to go to zero, the equivalent of a price war.

In the current Internet, the number of data carriers in a local region is small, usually limited to a couple of DSL, cable based providers and some Tier-II web hosts, which means that $\frac{N-1}{N}$ is not too large. This corresponds to the first case above and is what is happening today [13] in this segment. However, the Voice over IP (VoIP) market is very different. Customers are very mobile as they can choose any long-distance provider available with access through 1-800 numbers. This implies that there are a very large number of providers, most of whom are interested in quick gains (low discount factor). The model predicts a price-war in such a scenario, and indeed this is exactly what is happening [14].

Local and Transit ISPs

We next expand the model to include interactions between local groups and transit providers, which we study as a multi-stage game. Our model now takes into account the hierarchy of providers. The main result of this section is that in economic terms, Tier-II providers can be considered to be price transfer agents, transferring prices between two levels in the hierarchy. Any charges of the Tier-I ISPs would get transferred to the ISP originating/terminating the traffic along with an additional charge which is the Tier-II's profit. This means that Tier-II ISPs are "middlemen" between the local groups and Tier-I providers. This is the reason why no ISP wants to be seen as Tier-II [13]. The objectives of the two groups (local and transit ISPs) are entirely different and the division of ISPs into local and transit is thus the basis for peering. Any sort of peering or mutual-benefit traffic arrangement can be made only between entities of the same class, i.e., local ISP with local ISP or transit with transit. The model is extended to the case when there exist multiple transit providers between two geographical regions. We show that the game here is very similar to that of local ISP interaction and has the same potential for price wars.

Private exchanges

As indicated in Figure 2, ISPs have the option of exchanging traffic at private exchange points. This provides a new dimension to the problem. There has been some interest in the question of when it is economically viable to establish private exchange points. The marginal cost of traffic transfer for a Tier-II ISP is fairly low due to the large volumes of traffic generated by an aggregate of local ISPs. If local ISPs were to establish a private exchange, the marginal cost of transfer would depend on the volume of traffic exchanged. A study of the cost tradeoffs involved in such exchange is present in [8], where it is assumed that such ISPs at such exchange points would have similar amounts of traffic destined for each other and have a peering agreement.

We show that the above scenario is a special case of a more general multi-stage game where the players could have any bilateral settlement, rather than having to peer. Our results show that in the general case, the decision of whether or not to go in for private exchange is decided by fundamental limitations on the cost function and can be taken *unilaterally* by any ISP. There is no requirement for ISPs to sign a cooperative contract and reach a peering agreement. So asymmetric traffic cases (where there is a difference in amounts of traffic exchanged) are possible, as well as cases where one ISP uses a private exchange and the other uses a transit provider, with both making higher profits than using the transit provider alone. This suggests that private exchanges could become far more widespread than they currently are.

Organization of the paper

In Section II we discuss the game theoretic concepts used in the paper. Then in Section III, we discuss the theory of repeated games and how it applies to the case of interaction of local ISPs. The main idea used here is that of Nash-reversion being used as a threat strategy. We show how this could be used to enforce the optimal prices for high enough discount factors. In Section IV, we consider an extended model with a transit provider and show how there is a natural Stackelberg solution. We also study the case of multiple transit providers. The next logical extension of private exchange is then studied in Section V, with the conditions for private exchange being derived.

The objective of this paper is to understand and formalize some of the "Folk-Theorem" like results on Internet economics prevalent today. Our contribution is to show how these results naturally arise using surprisingly simple models.

II. Basic Notions

We first introduce the game theoretic concepts[1] that are used in this paper using a simple matrix game, which we call G_s. Consider Figure 3. There are two players who have identical *strategy spaces* $\mathcal{S}_1 = \mathcal{S}_2$, with elements *High* and *Low* called pure strategies. These can be thought of as the actions that they could possibly take. A *strategy profile* is an element of the product-space of strategy

[1] A good reference for this is [15].

Player 1 \ Player 2	High	Low
High	5 , 5	0 , 10
Low	10 , 0	0 , 0

Fig. 3. An example game illustrating Nash-reversion.

spaces of each player and is denoted by $\mathcal{S}$. An example of a strategy profile would be for Player 1 to play $s_1 = \textit{High}$ and for Player 2 to play $s_2 = \textit{Low}$, denoted by $s = (s_1, s_2) = (H, L)$. The payoffs that each strategy profile yields to each player are the elements of the matrix. The objective of each player is to maximize his individual payoff.

Definition A pure strategy Nash equilibrium is a strategy profile from which no player has a unilateral incentive to change his strategy. ■

The only pure strategy Nash equilibrium for this game (easily verified) is the strategy profile (L, L), which yields a payoff of zero to both players.

Suppose now that the game G_s above is repeated infinitely many times. We then obtain a new game which we denote by G_r. The payoff that the players receive in this game is the sum of (discounted) payoffs in each step. Formally, if the payoff in step k to player i be denoted by $\pi_i(k)$, then each player now has to maximize

$$\Pi_i(0, \infty) \triangleq \sum_{k=0}^{\infty} \alpha^k \, \pi_i(k), \tag{1}$$

where $\alpha \in (0, 1)$. Also, each player is aware of all the actions taken by both players up to the previous step. Let us denote the set of all possible histories of the game from step 0 to $k-1$ by H^k. A particular history $h^k \in H^k$ constitutes the information available to each player in step k. The concept of a pure strategy has to be extended a little in this setting. A pure strategy is now a contingent plan on how to play in each step k for a possible history h^k, making it a map from H^k to the set of possible actions. For example, in G_s a strategy could be "Play *Low*", whereas in G_r a strategy could be "Play *High* until the other player plays *Low*, and then play *Low* for ever". The concept of a Nash equilibrium remains unchanged.

We next consider the concept of *subgame perfection*. Since the players both know h^k, we can view the game from step k onwards with history h^k, as a game in its own right, which we denote as $G_r(h^k)$. This new game is called a *subgame* of the game G_r. Suppose there were some Nash equilibrium strategy profile s for the game G_r. Then, given some history h^k, s would recommend playing an action suggested by $s|h^k$ at step k. If the same strategy profile $s|h^k$ were a Nash equilibrium for the game $G_r(h^k)$, the strategy profile s is called subgame perfect. One way of thinking about this is to suppose that at step k all the players payoffs upto that step vanish. Then if $s|h^k$ is still the best thing to do, s is subgame perfect.

Definition A game is *continuous at infinity* if for each player i the payoff π_i satisfies

$$\sup_{h, \tilde{h} \ s.t. \ h^k = \tilde{h}^k} |\pi_i(h) - \pi_i(\tilde{h})| \to 0 \text{ as } k \to \infty$$

■

The condition says that events in the distant future are relatively unimportant. It is obvious that the discounted payoff function in our example is continuous at infinity since the payoff in any step is uniformly bounded by 10.

One-step deviation principle: In an infinite-horizon multi-step game which is continuous at infinity, profile s is subgame perfect iff there is no player i and no strategy $\hat{s}_i$ that agrees with s_i except at a single step k and h^k, and such that a unilateral deviation by player i to $\hat{s}_i$ gives him a better payoff conditional on history h^k being reached.

In our example game, using the one-step deviation principle, it is straightforward to show that if the discount factor $\alpha > \frac{1}{2}$, then the strategy profile "Play *High* until the other player plays *Low*, and then play *Low* for ever" (for both players) is a sub-game-perfect Nash equilibrium. We make the following observations about the strategy profile:

- The strategy profile being used when everything is "going fine" is (H, H), the best possible outcome in G_s if both players were to cooperate.
- The strategy used when any player deviates is L, the Nash equilibrium of the single step game G_s.

Such a threat strategy is called *Nash-reversion* [16]. It may sometimes be possible that a sub-game perfect threat exists which is "harsher" than Nash-reversion. Such a threat could work for lower discount factors. Abreu [17] calls such a strategy as an *optimal penal code*, which is optimal in the sense that it causes a deviating player to suffer the worst possible penalty. Of course, this penalty cannot be greater than the minimum guaranteed payoff that a player receives, no matter what the other players do. In the example given, the minimum guaranteed payoff is zero and hence Nash-reversion is also the optimal penal code.

III. Interaction between local ISPs

In this section, we consider the dynamics of pricing among local ISPs considered as a Bertrand game. In our model, the number of customers (end-users and websites) varies with prices set by ISPs. We first find the optimum price, which would maximize the profit of individually rational ISPs. We then show that this price can be enforced by threat in a repeated Bertrand game for a suitably high discount factor.

Consider the scenario illustrated in Figure 4. This shows a small section of the Internet. ISPs $1, 2$ and 3 are localized in a small geographical location, while the Tier-II ISP is involved in transit services. They are interconnected at a NAP where routing of traffic between ISPs is accomplished. We will consider bilateral interactions of ISPs $1, 2$ and 3 in this section. As in [7] consumers are divided into websites, which originate traffic and end-users who request the traffic. ISP i charges prices $\tilde{p}_i$ and p_i per unit traffic to websites and end-users respectively. Consumers are free to choose between the different local ISPs due to the geographical localization. Both websites' and end-users' demand (in terms of the number of consumers) vary with respect to the prices charged. Since the consumers are free to move, the demand depends on the lowest of the prices charged by the ISPs. We denote the demand functions of websites and end-users by $\tilde{n}(\tilde{p})$ and $n(p)$ respectively, where $\tilde{p} = min_i(\tilde{p}_i)$ and $p = min_i(p_i)$. We assume that each website and each end-user have one unit of traffic between them. Thus the total traffic in the region is $n(p)\tilde{n}(\tilde{p})$. This is just a restatement of Metcalfe's law, which states that the utility of a network is proportional to the square of the number of users present. An ISP incurs a marginal cost c_o per unit of traffic originating in a website subscribing to its services. Similarly, a marginal cost c_t is incurred by an ISP for terminating traffic at an end-user. The sum total of these costs is denoted by $c_o + c_t = c$. The costs for all ISPs in the region is assumed to be identical. All ISPs have identical fixed costs, which (since it makes no difference to any of the results) we take to be equal to zero.

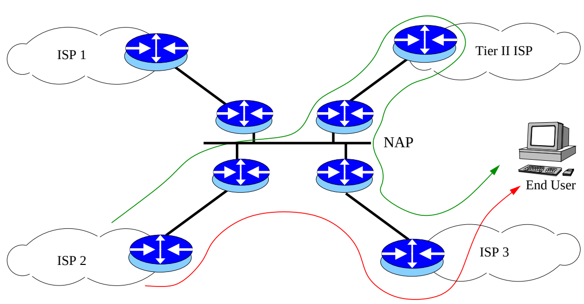

Fig. 4. A local ISP structure which shows how policy decisions can be implemented at NAPs.

Each ISP must provide a guarantee of connectivity to all other users. This means that ISPs have to exchange traffic. At the local level being considered, they may do this in one of two ways as shown in Figure 4. ISP i could charge ISP j an access charge a_{ji} for termination of traffic, i.e., they could have a bilateral settlement. Note that the access charge may be different for different ISPs. The other option is for both ISPs to pay a transit charge to a higher tier ISP (such as the Tier-II ISP in Figure 4), which would provide connectivity. Let the charge that the transit ISP asks for be κ. Then the ceiling on the access charge which a terminating ISP can ask for in a bilateral settlement is also κ as otherwise no ISP would want to directly interact with it. We study the game that such local ISPs play amongst themselves. At each step of the game ISPs must declare their prices and access charges. Customers then choose their service providers. Each step models temporal changes of prices, for example on a monthly basis. We assume that customers are charged based on usage, i.e., the number of bytes transferred. The profit function of ISP i in step k is given by $(i, j \in \{1, 2 \cdots N\})$

$$\begin{aligned}\pi_i(k) &= n_i(p_i)\tilde{n}_i(\tilde{p}_i)(p_i + \tilde{p}_i - c) \\ &\quad + \sum_{j \neq i} n_i(p_i)\tilde{n}_j(\tilde{p}_j)(p_i - (c_t - a_{ji})) \\ &\quad + \sum_{j \neq i} n_j(p_j)\tilde{n}_i(\tilde{p}_i)(\tilde{p}_i - (c_o + a_{ij})). \quad (2)\end{aligned}$$

Here, the first term is for traffic which originates and terminates in ISP i. The second is for traffic which originates from another ISP and terminates in ISP i. The third term is for traffic which originates in ISP i and terminates in some other ISP. The objective of the ISPs is to maximize the overall profit function (1) by setting appropriate prices and access charges. Using the same notation as in the example game in Section II, we will refer to the single step game by G_s and the infinitely repeated game by G_r.

The above expression assumes usage based pricing. If customers were charged a flat rate regardless of usage, the profit function would look like

$$\begin{aligned}\pi_i(k) &= n_i(p_i)p_i + \tilde{n}_i(\tilde{p}_i)\tilde{p}_i - n_i(p_i)\tilde{n}_i(\tilde{p}_i)c \\ &\quad + \sum_{j \neq i} n_i(p_i)\tilde{n}_j(\tilde{p}_j)(-c_t + a_{ji}) \\ &\quad + \sum_{j \neq i} n_j(p_j)\tilde{n}_i(\tilde{p}_i)(-c_o - a_{ij}). \quad (3)\end{aligned}$$

The analysis and conclusions are very similar for both pricing schemes with some differences in the expression for the equilibrium prices. Therefore, we only present results for usage-based pricing (2).

A. Properties of the demand functions

We assume that the number of consumers is infinitesimally divisible. This a reasonable assumption since the number of customers is very large in practice. The demand functions $f(x) : \mathbb{R}^+ \to \mathbb{R}^+$, are assumed to belong to a class of functions $\mathcal{S}$, which satisfy the following conditions:

1) They are continuous, twice differentiable and strictly decreasing ($f'(x) < 0$).
2) $f(0) = 1$ and $lim_{x\to\infty} f(x) = 0$. The first of these implies that the population is finite and normalized to be equal to one. The second implies that we have no customers at infinite price.
3) $\lim_{x\to\infty} x f(x) = 0$. This means that revenue obtained at infinite price is zero.

Now, let us define the cooperative profit function as $R(x, y) \triangleq f(x)g(y)(x + y - c)$ where $f, g \in \mathcal{S}$ over the set $\mathcal{P} = \{(x, y) : x + y \geq c$ and $x > 0, y > 0\}$. We will see later that when ISPs cooperate and set identical prices and access charges the profit function (2) looks similar to $R(x, y)$. We note the following properties of the above function:

- $\lim_{x+y\to c} R(x, y) = 0$. Follows immediately from properties 1 and 2.
- $\lim_{(x,y)\to\infty} R(x, y) = 0$. This follows from property 3.
- The function is positive everywhere except for $x + y = c$ and $x = \infty$ or $y = \infty$.
- There exists $(x^*, y^*) \in \mathcal{P}$ (not necessarily unique) which maximizes $R(x, y)$, yielding a non-zero maximum. This is seen as follows. Since $lim_{x,y\to\infty} f(x)g(y)(x + y - c) = 0$, we have that given any $\varepsilon > 0$ there exist finite p, q such that $f(x)g(y)(x + y - c) \leq \varepsilon$ for all $x > p$ or $y > q$. Thus, we can approximate the set $\mathcal{P}$ over which we maximize, infinitesimally closely by the set $\{(x, y) : x + y \geq c$ and $0 < x \leq p, 0 < y \leq q\}$. Now, since this is a compact set and the function $R(x, y)$ is continuous over this set, there must exist a (finite) maximum.

Example

An example of a function which satisfies the above conditions is $f(x) = \frac{1}{1+x^2}$, $x \in \mathbb{R}^+$. Let us take $c = 0$ and also assume that $f(x) = g(x)$ in the definition of $R(x, y)$. In this particular case, the cooperative profit function $R(x, y) = f(x)f(y)(x + y)$. This is plotted in Figure 5. We see that there exists a maximizing value of (x^*, y^*) as expected.

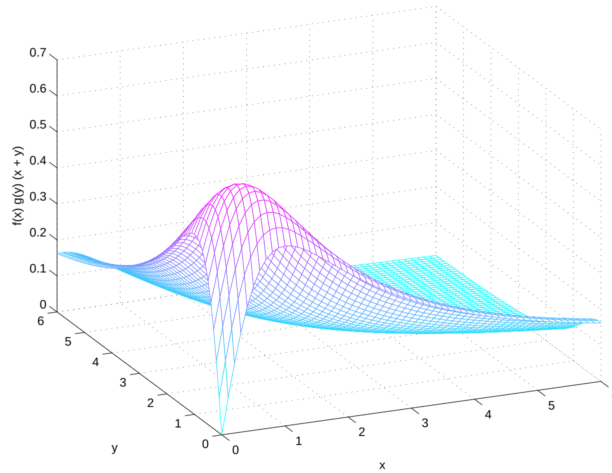

Fig. 5. Cooperative profit function for the demand function $\frac{1}{1+x^2}$.

B. Cooperative maxima between local ISPs for the single step game G_s

Consider the single step game G_s. We would like to know if there exists any cooperative maxima, i.e., is there a set of prices and access charges whereby the all the ISPs could maximize their joint profits? We will then show how threat strategies may be used to enforce cooperation.

Theorem 1: Suppose there are N ISPs in the region. Assuming that the ISPs cooperate and set equal access charges and prices and that the demand functions $\tilde{n}(\tilde{p})$ and $n(p)$ belong to the class $\mathcal{S}$, there exists a set of prices $\mathcal{P}_{opt}$ (consisting of pairs of form $(p_{opt}, \tilde{p}_{opt})$) which maximize the profits.

Proof: The profit function of ISP i is given by (2), repeated here for convenience.

$$\begin{aligned}\pi_i(k) &= n_i(p_i)\tilde{n}_i(\tilde{p}_i)(p_i + \tilde{p}_i - c) \\ &+ \sum_{j\neq i} n_i(p_i)\tilde{n}_j(\tilde{p}_j)(p_i - (c_t - a_{ji})) \\ &+ \sum_{j\neq i} n_j(p_j)\tilde{n}_i(\tilde{p}_i)(\tilde{p}_i - (c_o + a_{ij})), \quad (4)\end{aligned}$$

where $i, j \in \{1, 2 \cdots N\}$. We may simplify the above expression in the manner of [7] as follows:

$$\begin{aligned}\pi_i(k) &= n_i(p_i)\tilde{n}_i(\tilde{p}_i)(p_i + \tilde{p}_i - c_o - c_t - a_{ii} + a_{ii}) \\ &+ \sum_{j\neq i} n_i(p_i)\tilde{n}_j(\tilde{p}_j)(p_i - (c_t - a_{ji})) \\ &+ \sum_{j\neq i} n_j(p_j)\tilde{n}_i(\tilde{p}_i)(\tilde{p}_i - (c_o + a_{ij})). \quad (5)\end{aligned}$$

Simplifying we obtain

$$\pi_i(k) = n_i(p_i)\left(\sum_j \tilde{n}_j(\tilde{p}_j)(p_i - (c_t - a_{ji})\right) + \tilde{n}_i(\tilde{p}_i)\left(\sum_j n_j(p_j)(\tilde{p}_i - (c_o + a_{ij})\right) \quad (6)$$

By assumption since the ISPs cooperate (we assume that if all ISPs charge identical prices, they have an equal number of customers), we have that

$$\begin{aligned} p_j = p\,\forall j \quad &\text{and} \quad \tilde{p}_j = \tilde{p}\,\forall j \\ \sum_j n_j(p_j) = n(p) \quad &\text{and} \quad \sum_j \tilde{n}_j(\tilde{p}_j) = \tilde{n}(\tilde{p}) \\ n_j(p) = \frac{1}{N}n(p) \quad &\text{and} \quad \tilde{n}_j(\tilde{p}) = \frac{1}{N}\tilde{n}(\tilde{p}) \\ a_{ij} \quad &= \quad a_{ji}. \end{aligned}$$

Also, $p_j + \tilde{p}_j \geq c$ as the ISPs are individually rational. The individual profit functions are now

$$\pi_i(k) = \frac{1}{N}\, n(p)\, \tilde{n}(\tilde{p})\, (p + \tilde{p} - c). \quad (7)$$

The proof follows from the results on demand functions. ■

We will use $s_{cooperate}$ to denote the strategy profile in which all the ISPs play cooperative prices, and all set the access charges to zero (peering). In general, however, ISPs have no incentive to cooperate. Therefore a relevant question is the following: is the cooperative maximum found above enforceable by some means? We show the existence a threat strategy Nash equilibrium which would achieve the cooperative maximum. Our result is an example of Friedman [16] type punishment, wherein players revert to the single step Bertrand-Nash equilibrium if a deviation occurs.

C. The threat: Nash Equilibrium for the single step game

In the illustrative example considered in Section II, we saw that the single step Nash equilibrium could be used to construct a sub-game perfect Nash equilibrium for the infinitely repeated game. We find the single step Nash equilibrium below. Suppose that the ISPs set prices and access charges as follows:

$$\begin{aligned} p = c_t - \kappa \quad &\text{and} \quad \tilde{p} = c_o + \kappa \\ a \quad &= \quad \kappa \end{aligned}$$

We refer to the strategy profile in which all players play the above as s_{threat}. From (7), this means that all profits are zero. We will show that the above is actually the Nash equilibrium for the single step game G_s. First of all note that any unilateral increase in p or $\tilde{p}$ would have the effect of losing customers leading to a loss of revenue (a, of course, cannot be raised beyond κ, as it is the maximum possible). Also note that reducing the termination charge a has absolutely no effect on the number of customers. Any reduction in a only causes loss of revenue. So the only unilateral decision which has potential to increase profits is to decrease one or both of p and $\tilde{p}$. We then have the following three cases:

Case I: Set $p = c_t - \kappa - \varepsilon$. An ISP doing this will get all the end-users as customers. Its profit for each unit of traffic can easily be shown to be $-\varepsilon$ by considering two cases:

- Traffic entirely within the ISP's infrastructure yields a profit of $c_o + \kappa + c_t - \kappa - \varepsilon - c = -\varepsilon$.
- Traffic originating in another ISP and terminating within the ISP which charges the reduced price yields a profit of $c_t - \kappa - \varepsilon + \kappa - c_t = -\varepsilon$.

Case II: Set $\tilde{p} = c_o + \kappa - \varepsilon$. An ISP doing this will get all websites as customers. Again, we can show that the profit per unit traffic is $-\varepsilon$.

Case III: Set $p = c_t - \kappa - \varepsilon$ and $\tilde{p} = c_o + \kappa - \delta$. An ISP doing this would get all consumers as customers. However, its profit would be $-(\varepsilon + \delta)$.

We see that a unilateral shift by any ISP from the above prices would not yield a higher profit. Thus, the above set of parameters, which gives zero profits, defines the Nash equilibrium solution for the game. Noting that any ISP can guarantee itself a profit of at least zero, we see that this set of prices also yields the guaranteed minimum payoff to any player. Thus, playing this threat is also the *Optimal Penal code*. No ISP would play this threat unless its profits are hurt by some other ISP's choice of parameters. Thus, the threat of playing this set of parameters can be used to enforce that set of parameters which jointly imply a maximization of profit.

D. Subgame-perfect Nash equilibrium for the repeated game

We are now ready to address the question of whether there exists a sub-game perfect Nash equilibrium which ensures that the ISPs set prices according to $\mathcal{P}_{opt}$. We first require the following lemma.

Lemma 2: (proof is trivial) The repeated game G_r is continuous at infinity.

The lemma ensures that we can use the one-step deviation principle to prove sub-game perfection.

Theorem 3: The strategy profile "Play $s_{cooperate}$ until any player deviates and then play s_{threat} for ever" (for all players) is a sub-game perfect Nash equilibrium for the repeated game G_r under the condition that the discount factor

$$\alpha > \frac{\pi_{opt}\frac{N-1}{N} + \delta}{\pi_{opt} + \delta},$$

where π_{opt} denotes the optimal profit under cooperation and $\delta = \max(0, \kappa - \tilde{p}_{opt})$.

Proof: Consider the implications of a one-step deviation, i.e., suppose an ISP i sets the prices lower than the others at some step k. Since the ISP must maximize its profit it would do so by a reduction of p by ε for end-users, thereby obtaining all end-users as customers. Also, depending on the value of κ it would decide on whether it would be more profitable decrease $\tilde{p}$ (to get all the websites as customers) or increasing $\tilde{p}$ (thus ensuring that it has no websites and making a profit κ by terminating all the traffic). Thus, its profit for that one step is bounded in the following manner:

$$\begin{aligned} \pi_{i(dev)}(k) &\leq \alpha^k (n(p_{opt})\tilde{n}(\tilde{p}_{opt}) \\ &\quad \times (p_{opt} + \max(\tilde{p}_{opt}, \kappa) - c)) \\ &= \alpha^k N\,(\pi_{opt} + \delta). \end{aligned} \tag{8}$$

The other ISPs would employ the threat strategy from the next step onwards, causing all profits to go to zero. So the above value is also the total profit $\Pi_{i(dev)}(k, \infty)$ from step k to ∞.

The profit from step k to ∞ if the ISP chooses not deviate from the cooperative values is

$$\begin{aligned} \Pi_{i(ndev)}(k, \infty) &= \alpha^k \left(\pi_{opt} + \alpha\pi_{opt} + \alpha^2\pi_{opt}\ + ...\right) \\ &= \alpha^k\ \frac{\pi_{opt}}{1-\alpha} \end{aligned} \tag{9}$$

By assumption, since

$$\alpha > \frac{\pi_{opt}\frac{N-1}{N} + \delta}{\pi_{opt} + \delta},$$

we obtain from (8) and (9) that

$$\Pi_{i(dev)}(k, \infty)\ <\ \Pi_{i(ndev)}(k, \infty). \tag{10}$$

Thus, by the one-step deviation principle the strategy profile is a sub-game perfect Nash equilibrium. ■

The fact that the discount factor varies with the number of ISPs (N) as $\frac{N-1}{N}$ for the threat to be credible is worth noting. For large N the required discount factor tends to unity. This has the implication that ISPs have to be interested in long-term profits rather than quick gains if cooperative prices are to be viable. In market terminology the Bertrand-Nash equilibrium is the same as a price-war. This means that if there are a large number of ISPs in the market, price-wars are bound to follow since some of them would be interested only in short term gains resulting in a downward spiral of prices. The only solution to this would be for larger players to buy up the smaller ones. Then the threat strategy would be credible and prices would be stable.

IV. Interaction between Local and Transit Service Providers

In this section we consider the role of the transit provider. In general, transit providers are not concerned with providing services to websites or end-users. Instead, they obtain revenue from access charges that the local ISPs pay them. We have already seen in Figure 2 how transit providers between two locations connect two groups of local ISPs. Consumers cannot move between the two locations. This is a generalization of the scenario presented in [18]. There is a pair of economic transfer segments associated with a data transfer. In each economic segment each receiving ISP in the path of the flow charges each transmitting ISP a termination charge.

We extend the game in this scenario to include multiple regions connected by a hierarchy of ISPs. In Internet pricing, consumers are charged the same price regardless of how many transit hops that their traffic takes. However, for ease of exposition, we assume that consumers can be charged different prices for intra- and inter-regional traffic. It can be easily shown that similar results also apply for the case of constant price (irrespective of the number of hops), but the resulting expressions for the equilibrium prices are more complicated.

We use notation similar to the previous sections, with an additional superscript to denote the region. There are three types of players involved:

1) Local ISPs, which must set prices for intra and inter-regional traffic. As before, let the intra-regional parameters be denoted by $p_i^r, \tilde{p}_i^r$ and a_{ij}^r, where ISPs i and j both belong to region r. Also, let the prices charged by local ISP i to end-users and websites in region r for inter-regional traffic be denoted by $(q_i^r, \tilde{q}_i^r)$. The costs of origination and termination are taken to be c_o^r and c_t^r as before.
2) Tier-II providers, which charge the local ISPs for transit. Let the transit charges for outgoing (forward) and incoming (reverse) traffic in region r be t_f^r and t_r^r. The cost incurred by the Tier-II provider is denoted by t_c^r.
3) The Tier-I provider, which charges the Tier II provider (in region r) b_f^r and b_r^r for traffic from and to region r. Its cost for the transfer is t_c.

The profit function of the local ISPs in region r connected by transit providers to region s is now given by

$$\begin{aligned}\pi_i(k) &= n_r(p_i^r)\tilde{n}_r(\tilde{p}_i^r)(p_i^r + \tilde{p}_i^r - c^r) \\ &+ \sum_{j \neq i, j \in r} n_r(p_i^r)\tilde{n}_j^r(\tilde{p}_j^r)(p_i^r - (c_t^r - a_{ji}^r)) \\ &+ \sum_{j \neq i, j \in r} n_r(p_j^r)\tilde{n}_r(\tilde{p}_i^r)(\tilde{p}_i^r - (c_o^r + a_{ij}^r)) \\ &+ \sum_{l \in s} n_s(q_l^s)\tilde{n}_r(\tilde{q}_i^r)(\tilde{q}^r - c_o^r - t_f^r) \\ &+ \sum_{l \in s} \tilde{n}_s(\tilde{q}_l^s) n_r(q_i^r)(q^r - c_t^r - t_r^r) \qquad (11)\end{aligned}$$

The first three terms above are for intra-regional traffic and the last two are for inter-regional traffic. The objective is the maximization of the discounted sum of profits given by (1). We make the following observations:

- Intra-regional and inter-regional traffic pricing can be solved separately, as they do not share any terms.
- For inter-regional traffic, ISPs play two games: (i) amongst themselves to obtain consumers, and (ii) with the transit provider to maximize their profits.

Following the notation of the previous section we call the single step game G_s and the repeated game G_r. We first assume that the local ISPs in a region cooperate and study their interaction with transit providers. We then construct a threat strategy by which the cooperative prices may be enforced in a group of local ISPs.

A. *The single step Stackelberg game*

We model the interaction of ISPs as a Stackelberg game. In such a game players play in a definite sequence. Each player plays knowing that the next player would optimize his play based on what he does currently. The equilibrium in such a case is called a Stackelberg equilibrium. We assume for now that the local ISPs in a region cooperate and all them charge the same prices (in the next subsection, we will show how cooperation can be enforced among the local ISPs). This means that the local ISPs in a region can be grouped together as one player. The only problem now is that of setting prices for inter-regional traffic.

The Stackelberg solution, which we denote by $s_{stackelberg}$, for the single step game G_s is found as follows. Consider traffic which originates in Region$_1$, travels through the transit infrastructure and terminates in Region$_2$ (for traffic in the reverse direction, the roles of the two regions are reversed). We refer to the collection of local ISPs in a region as a *group*. Then the two economic segments are Group$_1$ $\rightarrow$ Tier-II (Region$_1$) $\rightarrow$ Tier-I and Group$_2$ $\rightarrow$ Tier-II (Region$_2$) $\rightarrow$ Tier-I. Consider Group$_1$'s problem:

$$\max_{\tilde{q}^1} \tilde{n}_1(\tilde{q}^1) n_2(q^2)(\tilde{q}^1 - c_o^1 - t_f^1), \qquad (12)$$

where n_1 and n_2 are the demand functions for inter-regional traffic in regions 1 and 2 respectively and we require that $\tilde{q}^1 \geq c_o^1 + t_f^1$ (individual rationality). Note that we have dropped the subscript identifying a particular local ISP since the whole group of local ISPs in a region is taken as a single player.

Lemma 4: The solution $\tilde{q}_{opt}^1$ to the maximization above is a strictly increasing function of t_f^1.

Proof: Differentiating (12) and equating to zero we have that $\tilde{q}_{opt}^1$ satisfies

$$\tilde{n}_1'(\tilde{q}_{opt}^1)\left(\tilde{q}_{opt}^1 - c_o^1 - t_f^1\right) + \tilde{n}_1(\tilde{q}_{opt}^1) = 0 \qquad (13)$$

where $n_1'(\tilde{q}^1) \stackrel{\triangle}{=} \frac{d\,\tilde{n}_1(\tilde{q}^1)}{d\,\tilde{q}^1}$ if $\tilde{q}_{opt}^1$ lies in the interior of $[c_o^1 + t_f^1, \infty)$. Since $\tilde{q}_{opt}^1$ is a maximum, we also have

$$2\tilde{n}_1'(\tilde{q}_{opt}^1) + \tilde{n}_1''(\tilde{q}_{opt}^1)\left(\tilde{q}_{opt}^1 - c_o^1 - t_f^1\right) < 0 \qquad (14)$$

Differentiating (13) with respect to t_f^1 and rearranging, we obtain

$$\begin{aligned}\frac{d\,\tilde{q}_{opt}^1}{d\,t_f^1} &= \frac{\tilde{n}_1'(\tilde{q}_{opt}^1)}{2\tilde{n}_1'(\tilde{q}_{opt}^1) + \tilde{n}_1''(\tilde{q}_{opt}^1)\left(\tilde{q}_{opt}^1 - c_o^1 - t_f^1\right)} \\ &> 0.\end{aligned}$$

The last inequality follows from the properties of the demand function $\tilde{n}_1(\tilde{q}^1)$ and (14). Suppose $\tilde{q}^1 = c_o^1 + t_f^1$ (boundary condition), then it is automatically increasing in t_f^1. Hence the proof. ■

Corollary $\tilde{n}_1\left(\tilde{q}_{opt}^1(t_f^1)\right)$ is a strictly decreasing function of t_f^1. The proof follows trivially from the lemma and the properties of $\tilde{n}(\tilde{q}^1)$.

Now consider the maximization problem faced by the Tier-II (Region$_1$) provider:

$$\max_{t_f^1} \tilde{n}_1\left(\tilde{q}_{opt}^1(t_f^1)\right) n_2(q^2)(t_f^1 - t_c^1 - b_f^1), \qquad (15)$$

where $t_f^1 \geq t_c^1 + b_f^1$.

Lemma 5: The solution t_{fopt}^1 to the maximization above is a strictly increasing function of b_f^1.

Proof: From the corollary of lemma 4 we have that $\tilde{n}_1\left(\tilde{q}_{opt}^1(t_f^1)\right)$ is strictly decreasing in t_f^1. Then proceeding as in the proof of Lemma 4, we obtain the result. ■

Theorem 6: $\tilde{q}_{opt}^1$ is a strictly increasing function of b_f^1.

Proof: The proof is a direct consequence of the Lemmas 4 and 5. ■

The theorem defines the role of the Tier-II provider in the Internet economy. It says that it acts as a middleman who transfers the access charge b_f^1 from Tier-I to Group$_1$ and makes a profit in the bargain. If there are multiple Tier-II providers in sequence then the entire sequence can be condensed into a block whose input is the access charge b_f^1 and whose output is the transit charge $t_f^1 > b_f^1$. Thus, the role of the Tier-II provider in the network is that of a price transfer agent between levels of the hierarchy.

Now, the theorems proved above hold for the other segment of the transfer as well. In this case price b_r^2 is transferred to Group$_2$. Thus, the problem faced by Tier-I is as follows:

$$\max_{b_f^1,\, b_r^2} \tilde{n}_1(\tilde{q}_{opt}^1(b_f^1)) n_2(q_{opt}^2(b_r^2)(b_f^2 + b_r^2 - t_c) \tag{16}$$

We define $\hat{n}_1(b_f^1) \triangleq \tilde{n}_1(\tilde{q}_{opt}^1(b_f^1))$ and $\hat{n}_2(b_r^2) \triangleq n_2(q_{opt}^2(b_r^2)$. It is easy to show that both of these functions belong to the class $\mathcal{S}$. Then we have that the Tier-I provider must solve

$$\max_{b_f^1,\, b_r^2} \hat{n}_1(b_f^1)\hat{n}_2(q^2)(b_f^1 + b_r^2 - t_c). \tag{17}$$

This looks identical in form to (7) and there exist values of $(b_f^1,\ b_r^2)$ which maximize it.

Example

We illustrate the Stackelberg game (for traffic from Region$_1$ to Region$_2$) by taking the demand functions $\tilde{n}_1(\tilde{q}^1) = e^{-\tilde{q}^1}$ and $n_2(q^2) = e^{-q^2}$. We then have the optimization problems as follows.

Group$_1$'s problem is

$$\begin{aligned} &\max_{\tilde{q}^1} \quad n_2(q^2)e^{-\tilde{q}^1}(\tilde{q}^1 - c_o^1 - t_f^1) \\ &\Rightarrow \quad \tilde{q}_{opt}^1 = 1 + c_o^1 + t_f^1, \end{aligned} \tag{18}$$

The Tier-II (Region$_1$) ISP's problem is

$$\begin{aligned} &\max_{t_f^1} \quad n_2(q^2)e^{-\tilde{q}_{opt}^1(t_f^1)}(t_f^1 - t_c^1 - b_f^1) \\ \equiv\ &\max_{t_f^1} \quad n_2(q^2)e^{-(1+c_o^1+t_f^1)}(t_f^1 - t_c^1 - b_f^1) \\ &\Rightarrow \quad t_{fopt}^1 = 1 + t_c^1 + b_f^1, \end{aligned} \tag{19}$$

where the second line follows from (18). We see that the transit provider has transferred the access charge b_f^1 to Group$_1$ along with its cost t_c^1 and a profit of one unit.

Group$_2$ and Tier-II (Region$_2$) have similar problems to the above. Their Stackelberg solutions are:

$$q_{opt}^2 = 1 + c_t^2 + t_r^2 \text{ and} \tag{20}$$

$$t_{ropt}^2 = 1 + t_c^2 + b_r^2 \tag{21}$$

Finally, the Tier-I provider's problem (after substituting the above parameters) is:

$$\begin{aligned} &\max_{b_f^1, b_r^2} \quad e^{-(2+c_o^1+t_c^1+b_f^1)}e^{-(2+c_t^2+t_c^2+b_r^2)} \\ &\qquad\quad \times (b_f^1 + b_r^2 - t_c) \\ &\Rightarrow \quad b_{fopt}^1 + b_{ropt}^2 = 1 + t_c \end{aligned} \tag{22}$$

The Stackelberg solution here is non unique. One would expect that in this case, the Tier-I provider would charge both sides equally, but this is not required.

B. A local threat to enforce the Stackelberg equilibrium

The Stackelberg solution found above applies to each local ISP group as a whole. What would prevent a member of a group from charging customers less than the optimal, thus getting all the customers? The situation here is similar to that of the local ISP game studied earlier in Section III, and we construct a Nash reversion strategy for the ISPs in a local group, which would ensure cooperation for a sufficiently high discount factor. Suppose that the local ISPs in a region r sets prices as follows:

$$\begin{aligned} p^r = c_t^r - \kappa_f^r \quad &\text{and} \quad \tilde{p}^r = c_o^r + \kappa_f^r \\ a^r \quad &= \quad \kappa_f^r \\ q^r = c_t^r + t_r^r \quad &\text{and} \quad \tilde{q}^r = c_o^r + t_f^r, \end{aligned}$$

Where, as in Section III, κ_f^r is charged by the Tier-II ISP to the originator of traffic for traffic originating and terminating in the same region, and would be determined by using an expression similar to (16). It is easy to verify the above is a Nash equilibrium for the single step game played within the group and that no ISP in the group can make a non-zero profit in such a regime. We note that the threat strategy of Section III (s_{threat}) is a special case of this Nash equilibrium. We denote this more general threat strategy using the same notation s_{threat}. The threat of playing this set of parameters can be used to enforce that set of parameters which jointly imply a maximization of profit.

Theorem 7: (Proof is identical to that of Section III) The strategy profile "Play the $s_{cooperate}$ for intra-regional traffic and $s_{stackelberg}$ for inter-regional traffic until a group member deviates and then play the single step Nash equilibrium prices s_{threat} for ever" (for all players) is a sub-game perfect Nash equilibrium for the infinitely repeated game if the discount factor

$$\alpha > \frac{\pi_{opt}\frac{N-1}{N} + \delta}{\pi_{opt} + \delta},$$

where π_{opt} denotes the optimal profit under cooperation, N is the number of local ISPs in a group, $s_{cooperate}$ is the cooperative strategy for intra-regional traffic (found in Section III), and $\delta = \max(0, \kappa_f^r - \tilde{p}_{opt}^r)$. ■

Thus, the scaling law under which local ISPs can make positive profits holds for the extended model as well.

C. Multiple Transit Providers

If there were multiple transit providers (either Tier-II or Tier-I) spanning the same levels in Figure 2 , then how would this affect the transit costs? This question is easily answered in the framework of maxima enforced by threat which we have developed.

Theorem 8: (Proof is obvious) If there exist N transit providers providing the same service, then they all must charge identical transit charges (given by the relevant Stackelberg solution). This can be enforced by Nash-reversion threat strategy for $\alpha > \frac{N-1}{N}$. ■

In the case of VoIP providers, the case is that of transit providers directly offering long-distance services to customers, who naturally pick the cheapest option. It has already been discussed how this leads to price wars.

V. The Case for Private Internet Exchanges

Suppose that an ISP in each local group in Figure 2 wishes to establish a private exchange point, bypassing the transit infrastructure. Norton [8] discusses the economic tradeoffs involved in this. According to [8], the cost per

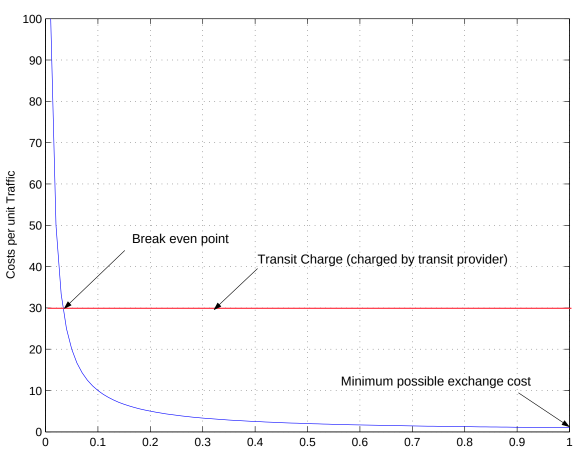

Fig. 6. Illustrating the cost tradeoffs in private exchanges.

unit traffic passing through a private exchange point is inversely proportional to the amount of traffic carried on it, as shown in Figure 6. By assumption, the traffic between two ISPs is equal to the product of the websites and the end-users each of them have. If the two ISPs have an identical number of websites and end-users, it makes sense for them to peer. The notation used in this section is similar to that of the previous section, except that we study the interaction of single ISPs in each region. Consider two ISPs 1 and 2 in regions 1 and 2 respectively. ISP_1's profit is given by,

$$\begin{aligned}\pi_1 &= \tilde{n}_1(\tilde{q}_1)n_2(q_2)\left(\tilde{q}_1 - c_{o1} - \frac{k}{\tilde{n}_1(\tilde{q}_1)n_2(q_2)}\right) \\ &\quad + \tilde{n}_2(\tilde{q}_2)n_1(q_1)(q_1 - c_{t1}),\end{aligned} \tag{23}$$

where, by assumption, $\tilde{n}_1(\tilde{q}_1)n_2(q_2) = \tilde{n}_2(\tilde{q}_2)n_1(q_1)$ and k is a proportionality constant. ISP_2's profit function is the dual of the above. Then comparing with (12), it is obvious that for

$$\frac{k}{\tilde{n}_1(\tilde{q}_1)n_2(q_2)} < t_{fopt}^1 \text{ and } \frac{k}{\tilde{n}_2(\tilde{q}_2)n_1(q_1)} < t_{fopt}^2, \tag{24}$$

it increases profits if a private peering point were established. This is exactly what Norton requires in [8] and the break-even condition is shown in Figure 6.

The above is a specific case of a more general problem. Suppose that the two ISPs are not constrained to peer (i.e., they can charge each other termination costs b_1 and b_2 respectively). We then have a new Stackelberg game, very similar in form to the one studied earlier. There are only two players now, since there is no transit provider involved. Each member of the group would establish a private exchange point (if it were profitable to do so), since all of them have an incentive to make higher profits. We first consider the asymmetric case where it is profitable for one group to use a private exchange, while the other uses the transit provider. As before, consider traffic from Region_1 to Region_2. Now, ISP_1's problem is

$$\begin{aligned}&\max_{\tilde{q}_1} \tilde{n}_1(\tilde{q}_1)n_2(q_2)\left(\tilde{q}_1 - c_{o1} - \frac{k}{\tilde{n}_1(\tilde{q}_1)n_2(q_2)} - b_2\right) \\ &\qquad = \max_{\tilde{q}_1} \tilde{n}_1(\tilde{q}_1)n_2(q_2)\left(\tilde{q}_1 - c_{o1} - b_2\right) - k\end{aligned} \tag{25}$$

and ISP_2's problem is

$$\max_{q_2,\, b_2} \tilde{n}_1(\tilde{q}_{1opt}(b_2))n_2(q_2)(q_2 - c_{t2} + b_2). \tag{26}$$

It is easy to show that $\tilde{n}_1(\tilde{q}_{1opt}(b_2))$ is monotone decreasing in b_2 and we have a very similar set of optimization problems as before. Then, if it turns out that there exists $(q_{1opt}, b_{2opt}, q_{2opt})$ such that

$$\frac{k}{\tilde{n}_1(\tilde{q}_{1opt})n_2(q_{2opt})} + b_{2opt} < t_{f1opt} \tag{27}$$

it would make sense for ISP_1 to ask for private exchange of traffic with ISP_2. For ISP_2 since from (15) we know that $t_{ropt}^2 > 0$, it is simple to see that the profit when the transit provider is eliminated is always higher. For the reverse problem of traffic from Region_2 to Region_1 the required condition is similar to (27).

Example

Similar to the example considered before, let the demand functions be exponentially decreasing i.e., $\tilde{n}_1(\tilde{q}_1) = e^{-\tilde{q}_1}$ and $n_2(q_2) = e^{-q_2}$. Then from (25) we have that ISP_1 maximizes

$$\max_{\tilde{q}_1} \quad e^{-q_2}e^{-\tilde{q}_1}(\tilde{q}_1 - c_{o1} - b_2) - k \qquad (28)$$
$$\Rightarrow \quad \tilde{q}_{1opt} = 1 + c_{o1} + b_2. \qquad (29)$$

From (26) we have that ISP_2's problem is

$$\max_{q_2, b_2} \quad e^{-q_2}e^{-(1+c_{o1}+b_2)}(q_2 - c_{t2} + b_2) \quad (30)$$
$$\Rightarrow \quad q_{2opt} + b_{2opt} = 1 + c_{t2} \qquad (31)$$

Then we have from (27), (29), (31) and (22) (assuming that the Tier-I provider charges both halves equally) that the condition for which private exchange is better for ISP_1 is

$$ke^{2+c_{o1}+c_{t2}} + b_{2opt} \leq \frac{3}{2} + c_{t1} + \frac{c_t}{2}, \qquad (32)$$

where b_{2opt} satisfies (31). The case where both providers use private exchange can be looked upon as two optimization problems similar to that considered above, only the providers maximize over the sum of revenues obtained from origination and termination of traffic. Thus, the decision as whether or not to establish a private exchange point depends on fundamental limitations on demand functions and costs. Note that the ISP in which the traffic is terminated always makes a higher profit by private exchange. The decision can be taken unilaterally by any of the ISPs in a group and all of the members of the group would take the same decision. Asymmetric cases where one group could go in for private point, while the other could stay with the transit provider are also possible. The conclusion is that when the requirement that the two groups must *peer* is relaxed, it increases the space where private exchange is possible.

VI. CONCLUSION

In this paper we have studied many facets of interaction between ISPs: both local service providers as well as large scale transit providers. We first studied interactions of ISPs in a geographically localized region and showed that as the number of ISPs increases, price wars would be the natural outcome due to Nash reversion. We predict that the outcome of this would be coalescing of ISPs into a small number of players interested in long term profits, who would then set prices which the market would naturally support.

We then tried to understand the role of transit ISPs and by finding the Stackelberg solution, showed how they act as price transfer agents between economic levels. We showed how interactions between different transit providers could be modeled as a repeated game. We also studied the effect of introducing the further option of private exchange and showed when it would be viable for ISPs to do so.

ACKNOWLEDGMENT

The first author would like to acknowledge Prof. Nevil Brownlee (University of Auckland) for the many interesting discussions he had with him on aspects of Internet architecture and pricing schemes.

REFERENCES

[1] G. Huston, *ISP Survival Guide: Stratagies for Running a Competitive ISP*, John Wiley and Sons, New York, 1999.
[2] W. B. Norton, "Internet Service Providers and Peering," 2001, White Paper from Equinix: http://www.equinix.com.
[3] "The Australian ISP Directory," 2003, http://www.australianispdirectory.com.
[4] A. M. Odlyzko, "Paris metro pricing for the Internet," in *Proceedings of the ACM Conference on Electronic Commerce (EC'99)*, 1999, pp. 140–147.
[5] A. M. Odlyzko, "Pricing and architecture of the Internet: Historical perspectives from telecommunications and transportation," March 2004, Available at: http://www.dtc.umn.edu/∼odlyzko/doc/networks.html.
[6] J-J. Laffont and J. Tirole, *Competition in Telecommunication*, MIT Press, Cambridge, Massachusetts, 2000.
[7] J-J. Laffont, S. Marcus, P. Rey, and J. Tirole, "Internet Interconnection and the Off-Net Cost Pricing Principle," 2001, Mimeo, Institut d'Economic Industrielle (IEDI), France.
[8] W. B. Norton, "A Business Case for ISP Peering," 2002, White Paper from Equinix: http://www.equinix.com.
[9] R. Dewan, M. Freimer, and P. Gundepudi, "Interconnection Agreements between Competing Internet Service Providers," in *Proccedings of the 33rd Hawaii International Conference on System Sciences*, January 2000.
[10] J-J. Laffont, S. Marcus, P. Rey, and J. Tirole, "Internet peering," *American Economic Review*, vol. 91, pp. 287–291, 2001.
[11] N. Badasyan and S. Chakrabarti, "Private Peering Among Internet Backbone Providers," January 2003, Economics Working Paper Archive at WUSTL, Series on Industrial Organization.
[12] P. Baake and T. Wichmann, "On the Economics of Internet Peering," 1998, Technical report, Berlecon Research.
[13] W. B. Norton, "The Evolution of the U.S. Internet Peering Ecosystem," 2003, White Paper from Equinix: http://www.equinix.com.
[14] B. Charny, "VoIP Providers Face Price War," November 3 2003, Article in news.com http://www.news.com.
[15] D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, MA, 1991 (7th printing, 2000).
[16] J. W. Friedman, "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies*, vol. 38, no. 113, pp. 1–12, January 1971.
[17] D. Abreu, "On the Theory of Infinitely Repeated Games with discounting," *Econometrica*, vol. 56, no. 2, pp. 383–396, March 1988.
[18] B. Hajek and G. Gopal, "Do greedy autonomous systems make for a sensible Internet?," in *Conference on Stochastic Networks*, 2002.