The Inefficiency of Markets for Provisioning Interconnected Communication Networks

Pedro Ferreira and Marvin Sirbu

Engineering and Public Policy Department, Carnegie Mellon University

5000 Forbes Avenue, Baker Hall 129, Pittsburgh PA 15213, USA

emails: {pferreira, sirbu}@cmu.edu

Abstract

The Internet is modelled as a collection of Internet Service Providers (ISPs) that exchange IP traffic through transit and peering agreements, using the traditional models for multi-commodity flow networks. Each ISP provisions the cheapest possible network to meet an exogenous fixed level of demand. Transit prices and the price of bandwidth are also fixed and exogenously given. In this context, a non-cooperative game among ISPs arises whose Nash equilibrium characterizes the topology of the resulting network. This network, which exists but is not unique, may be more expensive to provision than the network that ISPs would provision if they cooperated to reduce overall provisioning costs, which we call the optimal network. This extra cost constitutes a welfare loss. We show that if peering agreements are allowed between ISPs but transit agreements are not, there is no welfare loss. We also show that the optimal network with only peering agreements is more expensive to provision than the optimal network when both peering and transit agreements are allowed. Also, we show that the optimal network does not depend on the prices of transit. We then ask if transit prices could be adjusted to make the network at Nash equilibrium look more like the optimal network and thus reduce overall provisioning costs. We show by example that we can find networks at Nash equilibrium that are more expensive to provision than the optimal network, which shows that the market is not a sufficient mechanism to drive down provisioning costs to the minimum simply by adjusting transit prices. We conclude with an analysis of the difficulties that the regulator would face if it were to intervene in this market failure.

I. INTRODUCTION AND CONTEXT

The Internet is a collection of independent networks, called Autonomous Systems (ASes) [1], [2], which, many times, map into Internet Service Providers (ISPs). Smaller ISPs operate the edges of the Internet and provide access to end-users while larger ISPs operate the core of the network and provide access mainly to smaller ISPs [3]. ISPs must interconnect to allow users of different ISPs to exchange traffic. In the Internet, interconnection is accomplished through the Border Gateway Protocol (BGP) [4], which allows ISPs to advertise routes to sets of destinations in the network. Which routes are advertised to which ISPs depends on the interconnection agreements established, which reflect the interconnection strategies of the various ISPs [5], [6].

ISPs interconnect through transit and peering agreements. Under a transit agreement, an ISP pays another ISP to carry traffic on his behalf to and from every destination in the Internet [7], [8], [9], [10], [11]. The former ISP is seen as a customer of the latter ISP. Under a peering agreement, two providers exchange traffic destined to themselves, to their customers and to the customers of their customers, usually without cash settlements [7], [8], [9], [10], [11]. The dramatic reduction in the price of bandwidth [12], as Figure 1 shows for leased lines of 155 Mbps between NYC, US and London, UK, lead smaller ISPs to create a ”peering-donut” around the larger ISPs relying less frequently on transit from these providers [13], [14].

The declining trend in the price of bandwidth put pressure on transit prices, which have also fallen significantly since January 2003, as Figure 2 shows. This trend in the price of transit together with the steady increase in IP traffic worldwide (about 110% between 2003 and 2004 [15]) in most major international backbones) resulted in a substantial increase in

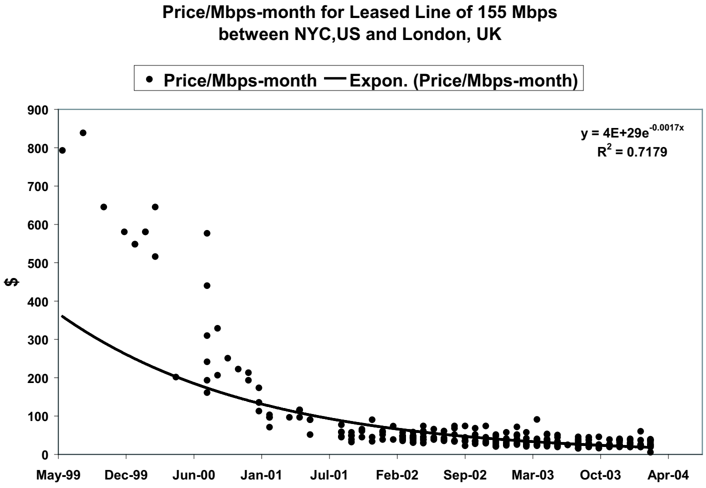

Fig. 1. Trend in the price/Mbps-month for leased lines of 155 Mbps between NYC,US and London, UK between May 1999 and April 2004 (source: Band-X.com).

transit revenues [16]. Therefore, it is clear that both peering and transit agreements are two main constructs of today's communication networks and ISPs do rely heavily on them to interconnect and exchange IP traffic across the globe.

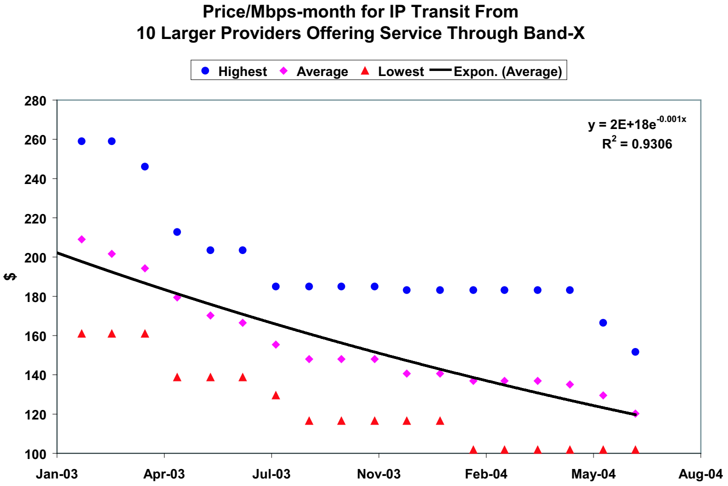

Fig. 2. Trend in the price/Mbps-month for IP transit from the 10 larger providers offering service through Band-X (source: Band-X.com).

Meanwhile, interconnection policy in the US is still defined by the Telecommunications

Act of 1996, which does not require ISPs to provide interconnection for the transport of IP traffic [17] and, usually, smaller ISPs have to meet a number of requirements to obtain peering from a larger provider, such as a minimum number of peering locations and a minimum bandwidth for those connections [18], [19]. If they do not meet these requirements, they usually have to pay for transit.

The Telecommunications Act defined a set of obligations that apply only to telecommunications service providers, particularly to Incumbent Local Exchange Carriers (ILECs), such as the duty to allow Competitive Local Exchange Carriers (CLECs) to "physically co-locate at the ILEC's premises to obtain access to Unbundled Network Elements", and therefore interconnect, "with at least the same quality, rate, terms and conditions as the access provided to subsidiaries and affiliates". However, as long as the transport of IP traffic is classified as an information service, none of these obligations apply to the Internet.

Often, larger ISPs agree to interconnect with smaller ISPs for free only if the latter meet a number of requirements. However, the only legal instruments that ISPs can use to dispute the reasonability of these requirements are anti-trust laws and tort law [20]. The absence of a specific law for interconnection applicable in the context of the Internet is particularly tied to the issue of market power in the Internet backbone. Concerns regarding dominance in this market were indeed the central topic of the review of the proposed WorldCom-Sprint merger of June 2000. The Department of Justice (DoJ) filed a suit to enjoin WorldCom from acquiring Sprint, arguing that the proposed merger "would substantially lessen competition in violation of Section 7 of the Clayton Act" [21]. As argued at that time by the provider Level 3, the harmful effects of such a concentration could be further exacerbated in the absence of an open interconnection policy framework for IP networks. Such a framework should, it was argued, encompass measurable and published peering criteria by all providers. Smaller providers would then know what would be needed in order to obtain interconnection from larger ISPs. Level 3 and Genuity published their interconnection policies [18], [19] as a way to pressure larger ISPs to publish theirs.

The FCC has also looked into the subjects of interconnection and market power and produced a series of reports on this matter. A first study concluded that in the absence of a dominant provider, market forces are enough to encourage interconnection[22]. Still, this same report notes that if a dominant ISP should emerge, regulation may be necessary to prevent monopoly rents, as it was before, for example, for the case of traditional telephony. Another report by the FCC made the case for settlement-free agreements as the default regime for interconnection at central offices in the context of traditional telephony [23]. Still another report by the FCC suggested that carriers should split equally the incremental costs of interconnection [24]. However, for the last three years, the FCC has been silent with respect to interconnection issues.

An objective of this paper is to contribute to this stream of research on interconnection in the context of IP networks, which in our view is still a topic insufficiently addressed in the current literature. In our paper, we model each ISP as a separate entity run by an independent administrator. Therefore, we may expect each ISP to provision its network selfishly and thus we can characterize the topology of the Internet as the Nash equilibrium of the non-cooperative game among ISPs that arises in this context [25], [26], [27], [28], [29]. However, it is well known that the Nash equilibrium carries a welfare loss [30], [31], [25], [32], in the sense that the aggregate cost to provision a network with such configuration may be greater than the cost to provision the network that would emerge if ISPs would cooperate to reduce overall provisioning costs still using peering and transit agreements to interconnect to meet the same level of exogenous demand. We call this latter network, the optimal network.

We show how large provisioning costs for the network at Nash equilibrium can become relative to the provisioning costs for the optimal network and we show how this result is a function of the level of economies of scale in the price of bandwidth. We also show how and why welfare loss is primarily related to transit agreements rather than peering agreements. Finally, we show that it is possible to have networks at Nash equilibrium that are more expensive to provision than the optimal network even when ISPs choose transit prices, which indicates that the market for provisioning communication networks that interconnect through transit and peering agreements set up the way they are today is not necessarily efficient. We conclude with a discussion of the difficulties that the regulator faces if it chooses to intervene to help mitigate this market failure.

II. RELATED WORK

In a network of selfish users that route their own traffic without making use of ISPs, the price of anarchy - defined as the worst-case ratio between the overall latency experienced in a network at Nash equilibrium and that of the best coordinated network [25], has been shown to be no more than the $sup_{\{e \in E\}}[(\lambda_e \mu_e + (1-\lambda_e))^{-1}]$, where λ_e satisfies $l_e^*(\lambda_e r) = l_e(r)$, μ_e is defined by $\mu_e = l_e(\lambda_e r)/l_e(r)$, $l_e(x)$ is the latency function for link e and $l_e^*(x)$ is the marginal latency function for this link [26]. $sup_{\{e \in E\}}[.]$ indicates the supremum of the expression within brackets over all the edges in the network. For the case of affine functions, this expression reduces to 4/3. For polynomial latency functions with non-negative coefficients and degree at most p, the price of anarchy is asymptotic to $[1 - p(p+1)^{-(p+1)/p}]^{-1}$ as $p \rightarrow \infty$. It has also been shown that this bound is tight, in fact it occurs in simple networks of parallel links [26], [30], [31], and independent of the network topology [33]. For the cases of M/M/1 and M/G/1 queues the price of anarchy approaches $+\infty$ as the sum of traffic rates approaches the smallest link capacity [26]. All these results are obtained assuming standard latency functions, for which $x.l_e(x)$ is convex. More recently, it has been shown that these results can be extended to the case of networks with capacities and with discontinuous latency functions that are non-decreasing in the amount of traffic carried [34].

These results have been criticized because the worst-case welfare loss occurs very seldom [35]. Let $C(r)$ represent the criticality factor of a network as a function of the traffic demand r. $C(r)$ is defined as the ratio between the aggregate running cost at Nash equilibria to meet r and that to meet a demand of $r/2$ [35]. The criticality factor of a network serves as an upper bound for its price of anarchy. Let A_φ represent the set of rates that imply a price of anarchy greater than φ and represent by $\nu(A_\varphi)$ the measure of this set. It has been shown that $\nu(A_\varphi) = O(log(C(r))/\varphi)$ and therefore A_φ has a small Lebesgue measure [35]. Additionally, it is well known that most traffic in the Internet uses TCP/IP [36], which implements flow control. Hence, the rates at which users send traffic to the network depend on the congestion in the network. It has been shown that the cost function that users minimize over each TCP connection is approximately linear [28]. Thus, in this case, we have $\varphi \leq 4/3$ and, in practice, one can expect generic losses to be small, at least when compared to the bounds provided before for the cases of M/M/1 and M/G/1 queues.

The price of anarchy is just an upper bound on the welfare loss in the network at Nash equilibrium and therefore it may convey little information about the actual loss experienced. The relevance of the price of anarchy depends on how often we are at the worst-case welfare loss. This issue has been addressed by simulation over realistic Internet-like topologies and traffic demands [37] considering only intra-domain traffic. These simulations have shown that selfish routing achieves close to optimal average latency in such environments, however such performance benefit comes at the expense of significant increased congestion on certain links.

Note, however, that all these results were generated without taking into consideration the existence of ISPs and assuming that end-users route their own traffic, a situation that is not true whatsoever in today's communication networks. As such, the models currently used in the literature do not address the issue of interconnection, which is key for the development of the Internet. The current paper aims at filling this gap by studying how the Nash configuration of interconnected networks differs from the best coordinated network of ISPs and how these differences result in increased provisioning costs. Such understanding is paramount for determining whether the market is a sufficient mechanism to minimize welfare loss.

III. Previous Work

This section introduces the model we used to describe a network of ISPs that interconnect through transit and peering agreements to exchange IP traffic. This model has been extensively studied in [38]. This section summarizes the main results obtained in the context of this model, namely the price of anarchy and the relationship between welfare loss and interconnection agreements.

A. Overview of the Model

In a previous paper, we modelled the Internet as a collection of ISPs that exchange IP traffic through transit and peering agreements [38]. In this model, ISPs meet an exogenous fixed level of demand. We have also assumed a price of transit and a price of bandwidth exogenously given. It is easy to see that all the results obtained in that paper do not change if instead of assuming a single price of transit we consider a different price of transit for each possible transit agreement in the network. Therefore, in this paper we assume that each transit agreement has a different price, still exogenously given, and thus we can think of each and every provider behaving as a (transit) price taker. Finally, both the transit prices and the price of bandwidth exhibit economies of scale, that is, the more traffic they carry the less the price per unit of traffic. Transit prices and the price of bandwidth exhibit different levels of economies of scale.

In this earlier paper, we assumed that each ISP maximizes is own profit regardless of what happens to other ISPs [38]. That is, each ISP behaves selfishly to provision its own network and to agree to interconnect to other ISPs. Therefore, the interaction among ISPs can be modelled as a non-cooperative game. The strategy of an ISP is to build up network between the cities it serves and to interconnect to other ISPs using peering and transit agreements. Interconnection agreements cannot be forced unilaterally, that is, they can only take place if both ISPs agree to do so.

The problem faced by each ISP is to maximize its own profit subject to a number of constraints. The profit function of each ISP is given by the difference between profits from selling and buying transit and provisioning costs, which reflect bandwidth costs. We have assumed that the cost of a peering link between two ISPs is split equally between them and a transit link is paid entirely by the ISP purchasing transit. The constraints of the ISP's optimization problem reflect the following aspects: i) traffic flows are positive amounts; ii) all exogenous traffic is delivered; iii) peering agreements only carry traffic to the parties involved in peering and their customers; iv) an ISP can only use a peering link to another ISP if the latter ISP agrees, also an ISP can only sell transit to another ISP if the latter ISP agrees to buy, and finally; v) an ISP cannot send traffic destined to himself over a transit agreement in which he is the selling party.

In the context described above, a non-cooperative game arises among the ISPs whose Nash equilibrium characterizes the topology of the resulting network [27]. We used the concept of Wardrop equilibrium [28] to characterize the set of flows in this network. A flow of any type of traffic in this network exists if this flow cannot be switched to another path in the network that allows the ISP responsible for delivering this flow of traffic to enjoy a greater profit per unit of traffic flow. We have shown that a network at equilibrium exists, using the facts that the set of strategies for each ISP is non-empty, compact and convex and the profit function for each ISP is C-concave [39]. Furthermore, we have shown, by example, that the network at Nash equilibrium is not unique. Therefore, when we refer to the cost of the network at Nash equilibrium, we are actually referring to the cost of the most expensive network at Nash equilibrium.

Also, in the context of this model, we defined the concept of optimal network. This is the network that would be achieved if ISPs would cooperate to reduce overall provisioning costs, still using transit and peering agreements to interconnect and exchange IP traffic to meet the exogenous demand. That is, the optimal network is the solution of a cost minimization problem over the whole network, but still taking into account the constraints in terms of traffic flows listed before. We have shown that this network also exists and that, in this case, but it is not unique. Also, we noticed that this network depends only of the exogenous demand and on the exogenous price of bandwidth. In other words, the optimal network is independent of the transit prices because transit payments cancel out when evaluated over the entire network.

B. The Price of Anarchy

By definition, the network at Nash equilibrium is no less costly to provision than the optimal network. We define welfare loss in ratio terms as the ratio between the costs to provision the network at Nash equilibrium and those to provision the optimal network. Therefore, welfare loss is a function of the exogenous parameters in the model and it is always greater than or equal to 1. If welfare loss is 1, the network at Nash equilibrium costs as much as the optimal network to provision. Next, we defined price of anarchy as the worst-case welfare loss. That is, the largest welfare loss possibly experienced in the network, due to the selfishness of each and every provider, evaluated over all the possible assignments of the exogenous parameters of the model. The price of anarchy, defined as a ratio of provisioning costs, is a function of the price of bandwidth and is independent of the prices of transit.

In our previous paper, we computed the price of anarchy and we introduced a lower bound for it which is a function of the level of economies of scale in the price of bandwidth [38]. Subsequently, we have shown that this lower bound for the price of anarchy is tight, since it occurs in networks with only peering agreements. Figure 3 summarizes the results obtained. The y-axis represents the lower bound for the price of anarchy and the x-axis represents the level of economies of scale in the price of bandwidth, decreasing from left to right. When the level of economies of scale is significant, ISPs observe the benefits from aggregating traffic, both within their own networks and over interconnection agreements, and seize these benefits, which results in a low anarchy value. On the other hand, when economies of scale are not so significant, the cost per unit of bandwidth deployed is about the same everywhere in the network and therefore there is no immediate benefit from aggregating traffic and the cheapest network is not much better than the network at Nash equilibrium and again anarchy is low. For the current level of economies of scale in the price of bandwidth the price of anarchy is at least 25%. See [38] for more details on this figure.

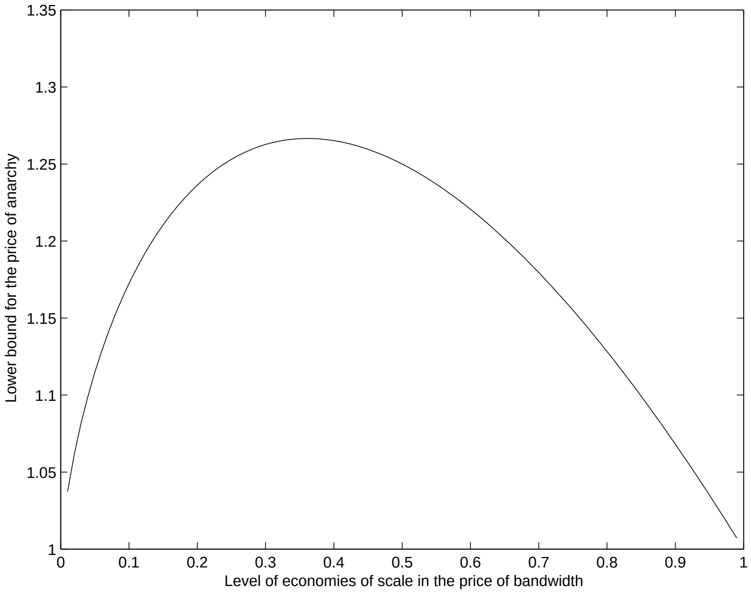

Fig. 3. Lower bound for the price of anarchy as a function of the level of economies of scale in the price of bandwidth.

The welfare loss between the network at Nash equilibrium and the optimal network is the additional cost that society incurs to provision the network of ISPs due to the selfishness of each and every ISP in the network. No welfare loss would exist if a common administrator would oversee the choices of every ISP in the network and dictate which interconnection agreements to put in place. Thus, we can also look at the price of anarchy as the extra provisioning cost that society incurs to allow different providers in the network, which in turn promotes competition and reduces, for example, the prices of the services provided over the network. So, we can look at the welfare loss as the ”price of competition”. In other words, we can imagine that a common administrator for all ISPs would behave as a monopolist and regulation could be needed to prevent him from abusing his market power by collecting (unfair) monopoly rents. But if the cost of setting up the appropriate regulation mechanisms exceeds the price of anarchy, then it is better not to have this central scheduling administrator and allow different ISPs to compete, because the price of anarchy is the most society pays to allow different providers in the network, which, per se, should be enough to foster competition. Therefore, the price of anarchy can also be understood as a benchmark for how much we might be willing to pay for regulation that lead to nearly optimal networks.

C. The Structure of the Welfare Loss

It is now natural to ask how does the welfare loss embedded in the network at Nash equilibrium relate to the fact that ISPs interconnect through transit and peering agreements. In our previous paper, we have assumed that profit functions are exponential in nature, which is usually the case given the economies of scale in both the price of transit and the price of bandwidth. With this assumption in place, we have shown that if all paths to deliver traffic between a pair of points in the network exhibit the same profit function then there is no welfare loss because in such case the network at Nash equilibrium is the same as the optimal network.

We concluded that welfare loss in terms of provisioning costs might arise from the fact that the market for the transport of traffic is not homogenous and there are different levels of economies of scale across cities. Links where the level of economies of scale is higher tend to attract more traffic, because ISPs realize the benefits from transporting traffic over these links as opposed to splitting such traffic over separate links. While these advantages may be fully explored by a central authority aiming at minimizing overall provisioning costs, this will hardly be the case with multiple ISPs. Each ISP seizes the benefits of strong economies of scale within its own network (and over interconnection agreements to other ISPs whenever possible), but this optimization is done locally at each network. As a consequence, traffic might be shifted differently across paths and might be handed-off between ISPs at sub-optimal interconnection points, which, on aggregate, results in higher provisioning costs.

A peering agreement is, in essence, no more than establishing a direct link between two ISPs and therefore its cost reflects solely the cost of this bandwidth link. Thus, the level of economies of scale in the price of peering is just the same as the level of economies of scale in the price of bandwidth. However, transit agreements exhibit a different level of economies of scale because their price does not reflect directly the true cost of the infrastructure. The price of a transit agreement includes a service fee to carry traffic on the behalf of the customer ISP to and from every destination in the Internet. When prices reflect costs, we expect welfare loss to be low. This is not the case with the price of transit agreements, and the distortion between prices and true costs introduced by this service charge results in some welfare loss.

In sum, we conclude that peering agreements do not introduce welfare loss, as long as they reflect the true cost of the infrastructure. However, transit agreements do introduce welfare loss. But, at the same time, transit agreements might contribute to reduce overall provisioning costs because they allow for aggregating more traffic than peering agreements do, which in turn allows for benefiting from economies of scale to a larger extent. Therefore, overall, transit agreements can be seen as a welfare enhancing mechanism, yet not a fully efficient one. But, recall that all the results obtained so far were derived assuming that transit prices are exogenous. Could prices of transit be somehow adjusted to make the network at Nash equilibrium become as inexpensive as the optimal network? The next section looks at this question in detail.

IV. The Inefficiency of the Market

In this section, we show by example that situations occur in which a network at Nash equilibrium is more expensive to provision than the optimal network even when ISPs can choose transit prices and, therefore, we conclude that the market with transit and peering agreements set up the way they are today is not efficient. We start by highlighting a few important characteristics of the Nash network and of the optimal network and we finish with a discussion of the difficulties that the regulator would face if it were to intervene to help mitigate the market inefficiency.

A. The Effect of Transit Agreements and Transit Prices

In the previous section, we showed that when only peering agreements are allowed no welfare loss arises. This situation can be modelled by adding a new constraint to the model introduced before. To rule out transit agreements, we can simply say that traffic in any transit link in the network must equal zero. Therefore, the overall cost minimization problem that defines the optimal network is, in the case of only peering, a minimization problem over a decision set that is a subset of the decision space of the overall cost minimization problem when both peering and transit agreements are allowed. Therefore, it is trivial to show that the costs to provision the optimal network cannot increase when transit becomes available.

In reality, transit agreements contribute significantly to reduce provisioning costs because they allow for better aggregating traffic that is destined to the same places in the network, which might in turn result in networks with fewer links. For example, a network with n nodes, and only peering, in the absence of Network Access Points (NAPs), requires on the order of $n(n-1)/2$ links; whereas a network with only transit agreements will most likely result in a tree structure that only needs in the order of $n.log(n)$ links. NAPs have emerged as a way to reduce the number of links in the network. An ISP can run a fat single link to the NAP, which it uses to carry all the traffic to all the ISPs with which it interconnects at the NAP, instead of running several separate links, one to each of these ISPs. Note that NAPs can be easily modelled with the model we introduced before, when we consider a city where several ISPs meet but none of them serves demand from end-users at that city.

Moreover, we must note that the network at Nash equilibrium is not unique and thus Nash networks can exhibit different provisioning costs. The optimal network is not unique as well. However, all optimal networks will have the same overall provisioning cost. More importantly, an optimal network does not depend on the prices of transit. This is because the optimal network is a cost minimization problem over the entire network and whatever payments ISPs make to each do not really matter. The fact that sometimes the optimal network uses transit links depends only on how using these links allows for aggregating more traffic thus reducing costs.

Therefore, the question we aim at answering in this section is "is it possible that a network at Nash equilibrium is more expensive to provision than the optimal network when ISPs can choose transit prices?". If we can show that this is the case, we conclude that the market can end up in a network configuration that is more expensive than the optimal network. This does not mean, however, that the market is always inefficient, because the Nash network is not unique and therefore there might still exist other Nash networks that are no more expensive to provision than the optimal network. But showing, by example, that there exists a Nash network that is more expensive than the optimal network suffices to prove that the market is not always efficient. The next subsection provides such an example.

B. Expensive Nash Network

The following is a very simple example that shows that a Nash network can be more expensive than the optimal network. Consider ISP i serving city u and city v and ISP j serving city r. There is an exogenous demand of 10 Mbps from users of ISP j at city r to users of city ISP i at city v (we can assume that there is also a demand of traffic less than 10 Mbps from the latter users to the former users, but this changes nothing in the what come next). All other demand in the network is set to zero. Furthermore, assume that the length of shortest links between city u and city r, between city v and city r and between city u and city v are, respectively, 2105, 2125 and 332 miles, which is clearly a realizable network. Assume that the price of bandwidth links is given by the function $PB(M, L) = 248.29M^{0.5269}L^{0.3774}$, where M represents the bandwidth of the link and L represents its length. This price function for bandwidth has been estimated in a previous paper [38]. Figure 4 illustrates this network.

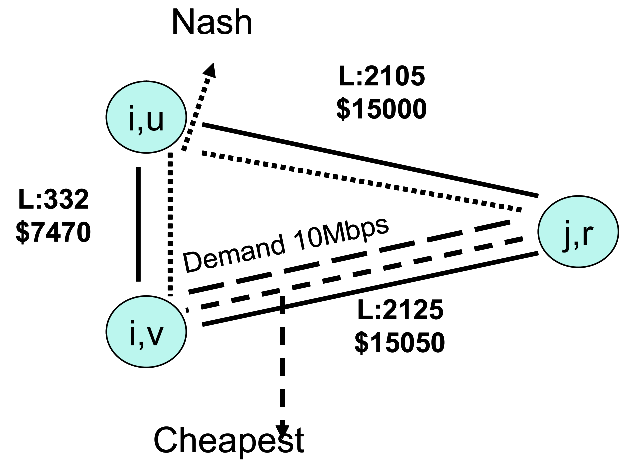

Fig. 4. Example of two ISPs serving three cities in which there is a Nash network strictly more expensive than an optimal network.

In this case, the optimal network is to send 10 Mbps from ISP j at city r to ISP i at city v for a total provisioning cost of \$15050. Now, consider the network in which ISP i at city u and ISP j at city r peer. In this latter case, ISP i and ISP j pay a total cost of \$14970 and \$7500, respectively, for a total provisioning cost of \$22740. We can show that this network, which is more expensive than the optimal network, is at Nash equilibrium. For that, we will show that ISP i and ISP j cannot agree to change it.

For example, an alternative network would be to have ISP i at city v and ISP j at city r peer. In this case, both ISP i and ISP j would have to pay \$7525. Therefore, ISP i is not willing to proceed with the change. Other alternative networks include transit agreements. Consider the network in which ISP i at city v buys transit from ISP j at city r. In this case, ISP i would have to pay \$15050+T and ISP j would have receive T, where T represents the transit payment. Therefore, ISP i would always be worse off and would not agree to change.

A similar reasoning applies when ISP i at city v sells transit to ISP j at city r.

The remaining two network configurations involve transit between ISP i at city u and ISP j at city r. Consider the network in which ISP i at city u buys transit from ISP j at city r. In this case, ISP i would have to pay $22470+T and ISP j would have receive T, where again T represents the transit payment. Therefore, ISP i would always be worse off and would not agree to change. Finally, consider the case in which ISP i at city u sells transit to ISP j at city r. In this case, ISP i would incur in a cost of $7470-T and ISP j would incur in a cost of $15000+T. Therefore, ISP j would always become worse off and would not agree to change. In sum, ISP i and ISP j cannot find an alternative network to which they would both agree to change to and the network with the peering link between ISP i at city u and ISP j at city r (and the internal link in ISP i's network between city u and city v) is a Nash equilibrium, which is more expensive to provision than the optimal network.

Finally, note that not all possible networks in the previous example are at Nash equilibrium. For example, the network in which ISP i at city u buys transit from ISP j at city r is not a Nash equilibrium. To see this, note that under that network ISP i pays \$22470+T_1 and ISP j receives T_1, where T_1 is the transit payment. But, consider the network in which ISP i at city v buys transit from ISP j at city r. In this case, ISP i pays \$15050+T_2 and ISP j receives T_2, where T_2 is the transit payment in this case. Therefore, any positive payment T_2 between T_1 and $T_1 - 7470$ will make both ISPs agree to change to the latter network and hence the former one is never a Nash equilibrium.

The example above proves that the market is not efficient to provision interconnected communication networks because even when ISPs choose transit prices there exist Nash networks that are (strictly) more expensive than the optimal networks. Hence, there is scope for the regulator to intervene to improve total welfare.

C. A Difficult Role for the Regulator

The previous subsection showed, by example, that the market can end up in a Nash network configuration that is more expensive to provision than the optimal network. This indicates that there is scope for the regulator to intervene in order to help mitigate this source of welfare loss. However, the regulator faces a series of difficulties to do so that lead us to believe that it is better to allow the market to run and to give us a good allocation of bandwidth and interconnection agreements yet possibly not the optimal one.

First, we do not know how often situations as the one discussed in the previous subsection occur and it would be clearly unwise to have the regulator focus on welfare loss cases that occur seldom. Second, the regulator is at a severe loss in terms of accurate knowledge about the topology of the network and actual traffic flows. In face of such lack of information, it is impossible for the regulator to create and to use a good enough model of the network that could allow him to identify the appropriate changes to reduce provisioning costs. Moreover, the changes in the network that the regulator could promote to reduce provisioning costs depend on the demand for the transport of traffic, which in our model was considered exogenous. Therefore, the model the regulator would have to use to study what network changes should occur would have to be more complex than the one we present in this paper.

Third, and assuming that the regulator could identify these changes, he would have to devise ways to intervene in the market to change the behavior of ISP, which could include designing the proper incentive mechanisms to induce ISPs to route traffic flows and/or to interconnect differently. However, the process of incentive design and implementation takes a significant amount of time and it could be likely that the regulator's policies would end up being applied to an already different state of the world, which could render them useless, if not harmful. Forth, the regulator should take into account that ISPs will certainly behave differently once they factor in the fact that the regulator will act upon the market to mitigate welfare losses. That is, again, the model that the regulator should use to predict the behavior of ISPs would have to be more complex than the model we presented in this paper, which does not account for the role of a regulator.

Finally, consider the case of international networks, for which no regulator holds overall jurisdiction. Networks that span different countries operate under the regulatory regime defined by different regulators who would then have to cooperate to reduce overall costs, which is a task that might become quite hard to implement. Moreover, different regulators might have different incentives and each of them might behave selfishly to meet his goals, which might not be aligned with the overarching goal of reducing overall provisioning costs. Taking into account all the issues listed above, it is reasonable to say that it is extremely difficult for the regulator to act upon the market inefficiency identified in this paper. It is therefore fair to acknowledge that it might be better to allow the market to operate, which will result in a good yet possibly not optimal allocation of bandwidth and of interconnection agreements.

V. CONCLUSIONS

This paper addresses the issue of economic inefficiencies in terms of provisioning costs for communication networks. A number of ISPs, assumed to behave rationally, engage in a non-cooperative game to build their networks and to establish transit and peering agreements to interconnect and thus meet an exogenous fixed level of deterministic demand for the transport of IP traffic between a number of cities. Each ISP maximizes its own profit, which is given by the difference between the profits from selling and buying transit and provisioning costs. The network that results from this interaction, which we call network at Nash equilibrium, is, by definition, no less expensive to provision than the network that ISPs would provision to meet the same level of exogenous demand if they were to cooperate to reduce overall provisioning costs. The latter network, we call the optimal network. We call welfare loss the ratio of provisioning costs between the network at Nash equilibrium and the optimal network. In this paper, we are interested in determining to what extent the market is a sufficient mechanism to cope with the selfishness of each and every provider in the network and to reduce both the welfare loss and provisioning costs to a minimum.

A model for interconnected networks based on the traditional models of multi-commodity flows is provided to study this question. We have discussed how a lower and tight bound for the price of anarchy, which in the context of our model is an upper bound on the welfare loss, changes with the level of economies of scale in the price of bandwidth. This lower bound is low when the level of economies of scale is either very significant and very insignificant. When economies of scale are significant, ISPs seize the benefits from aggregating traffic and anarchy is low. When economies of scale are not so significant, the cheapest network of networks is not much better than the network that results from the non-cooperative interaction among ISPs, because there are no immediate benefits from aggregating traffic to realize, and, again, anarchy is low. The current level of economies of scale in the market for bandwidth yields a lower bound for the price of anarchy of 25%.

In addition, we have discussed how welfare loss relates to the type of interconnection agreement. Peering agreements reflect directly true provisioning costs and therefore no major welfare loss arises from this type of interconnection agreement. Transit agreements exhibit a different level of economies of scale due to a service charge collected by the transit seller. This service charge is mainly determined by market conditions and does not reflect directly the true costs to transport traffic. This distortion between prices and costs results in economic welfare losses. In sum, transit agreements result in welfare loss while peering agreements per se do not. However, at the same time, transit agreements are welfare enhancing, in the sense that they help reduce overall provisioning costs, because they allow for better aggregating traffic destined to the same places in the network even when in their presence it is possible for ISPs acting selfishly not to realize all the provisioning cost reductions which transit makes possible.

In our model, transit prices are fixed and exogenously given and therefore we can think of ISPs as (transit) price takers. For a particular instance of these prices, the network at Nash equilibrium is at least as expensive as the optimal network. We then ask if these prices can be adjusted so that the former network seems more like the latter network and thus provisioning costs for the Nash network decline. We have shown that by example that there are situations in which there exists a Nash network that is more expensive to provision than the optimal network even when transit prices may vary. This means that the market is inefficient, because even when ISPs can choose transit prices they can end up in a Nash network that costs more than the optimal network. Still, we must keep in mind that the Nash network is not unique and the fact that there exist a Nash network that is more expensive than the optimal network does not prevent other Nash networks from being no more expensive to provision than the optimal network.

Nevertheless, this market failure indicates that there is scope for a regulator to intervene to help mitigate the sources of welfare loss. However, we also discuss a number of difficulties that the regulator would have to face for it to act properly upon the market. Some of these difficulties are as follows: i) these inefficient outcomes may occur seldom, in which case it would be unwise for the regulator to focus on them; ii) there is lack of information about actual traffic flows, which might prevent the regulator from identifying properly the sources of welfare loss; iii) the regulator would have to act, in a timely fashion, but designing and implementing the appropriate incentive mechanisms is a time consuming process; iv) the regulator would have to take into account that ISPs would behave differently if they knew that the regulator were to intervene; v) if we are looking at international networks we have to take into account that such networks are under the jurisdiction of several regulators who then would have to cooperate to reduce overall provisioning costs. In the face of all these issues, it is fair to say that it might be better to let the market work to provide us with a good allocation of bandwidth and interconnection agreements, yet possibly not the optimal one.

REFERENCES

[1] T. Bates, E. Gerich, L. Joncheray, J. Jouanigot, D. Karrengerg, M. Terpstra, and J. Yu, “Representation of IP Routing Policies in a Routing Registry (ripe-81++),” Request for Comments 1786, 1995.

[2] J. Hawkison and T. Bates, “Guidelines for Creation, Selection and Registration of an Autonomous System (AS),” Request for Comments 1930, 1996.

[3] W. Fang and L. Peterson, “Inter-AS traffic patterns and their implications,” in *IEEE Global Internet Symposium*, December 1999, Rio de Janeiro, Brazil.

[4] Y. Rekhter and T. Li, “A Border Gateway Protocol (BGP-4),” Request for Comments 1654, 1994.

[5] J. Stewart, *BGP-4: Inter-Domain Routing in the Internet*, Addison-Wesley Publishers Company, 1998.

[6] S. Halabi and D. McPherson, *Internet Routing Architectures*, Cisco Press, 2000.

[7] J. Marcus, “Global Traffic Exchange Among Internet Service Providers (ISPs),” in *OECD - Internet Traffic Exchange*, June 2001, Berlin, Germany.

[8] G. Huston, *ISP Survival Guide: Strategies for Running a Competitive ISP*, John Wiley & Sons, 1999.

[9] T. McGarty, “Peering, Transit and Interconnection: Internet Access in Central Europe,” MIT Internet Telephony Consortium, Working Paper, January.

[10] W. Norton, “Interconnection Strategies for ISPs,” White Paper, Equinix, 1999.

[11] W. Norton, “Internet Service Providers and Peering,” White Paper, Equinix, 1999.

[12] P. Ferreira, J. Mindel, and L. McKnight, “Why Bandwidth Trading Markets Have not Matured? Analysis of Technological and Market Issues,” *International Journal of Technology, Management and Policy*, vol. 2, to appear.

[13] G. Cook, “Peering, Transit and IXs: Economics and Policy,” summary available at http://www.cookreport.com/11.08.shtml, November.

[14] M. O’Kelly and T. Grubesic, “Backbone Topology, Access and the Commercial Internet: 1997-2000,” *Environmental and Planning B*, vol. 29.

[15] Telegeography, “Annual Traffic Report,” 2004.

[16] Telegeography, “,” Light Reading, August.

[17] U.S. Congress, “Telecommunications Act of 1996,” Pub. L. No. 104-104, 110 Stat. 56, 1996.

[18] Level 3, “Level 3 Policy for Settlement Free Interconnection,” available from Level 3’ website at http://www.level3.com/1511.html, 2001.

[19] Genuity, “Interconnection Guidelines for Genuity,” available from Genuity’s website at http://www.genuity.com/infrastructure/interconnection.htm, 2001.

[20] T. McGarty, “Competition in the Local Exchange Market: an Economic and Antitrust Perspective,” MIT Internet Telephony Consortium, September 1996.

[21] Department of Justice, “United States v. WorldCom, Inc., and Sprint Corporation,” Complain filed in the United States District Court for the District of Columbia (Case IV/M.11741 - MCI WorldCom/Sprint), June 26.

[22] M. Kende, “The Digital Handshake: Connecting Internet Backbones,” Working Paper 32, Office of Plans and Policy, Federal Communications Commission, FCC, September.

[23] P. DeGraba, “Bill and Keep at the Central Office as the Efficient Interconnection Regime,” Working Paper 33, Office of Plans and Policy, Federal Communications Commission, FCC, December.

[24] J. Atkinson and C. Barnekov, “A Competitive Neutral Approach to Network Interconnection,” Working Paper 34, Office of Plans and Policy, Federal Communications

Commission, FCC, December.

[25] E. Koutsoupias and C. Papadimitriou, “Worst-case Equilibria,” in *16th Annual Symposium on Theoretical Aspects of Computer Science*, pp. 404–413.

[26] T. Roughgarden, “Selfish Routing,” PhD Dissertation, 2002.

[27] J. Nash, “Non-cooperative Games,” *Annals of Mathematics*, vol. 54, no. 2, pp. 286–295.

[28] J. Wardrop, “Some Theoretical Aspects of Road Traffic Research,” in *Institute of Civil Engineers*, vol. 1, Part II.

[29] M. Beckmann, C. McGuire, and C. Winsten, *Studies in the Economics of Transportation*, Yale University Press, 1956.

[30] C. Pigou, *The Economics of Welfare*, Macmillan, 1920.

[31] D. Braess, “Uber ein Paradoxon der Verkehrsplanung,” *Unternehmensforschung*, available from http://homepage.ruhr-uni-bochum.de/Dietrich.Braess/.

[32] T. Roughgarden, “Designing Networks for Selfish Users is Hard,” *42nd Annual Symposium on Foundation of Computer Science*, pp. 472–481, 2001.

[33] T. Roughgarden, “The Price of Anarchy is Independent of the Network Topology,” in *34th Annual ACM Symposium on the Theory of Computing*, 2002.

[34] J. Correa, A. Schulz, and N. Stier, “Selfish Routing in Capacitated Networks,” in *INFORMS Annual Meeting*, Atlanta, Georgia.

[35] E. Friedman, “A Generic Analysis of Selfish Routing,” Working Paper, available at http://www.orie.cornell.edu/ friedman/papers.htm, 2001.

[36] J. Postel, “The Transport Control Protocol(TCP),” Request for Comments 793, 1985.

[37] L. Qiu, Y. Yang, Y. Zang, and S. Shenker, “On Selfish Routing in Internet Like-Environments,” in *SIGCOMM’03*, August 25-29 2003, Karlsruhe, Germany.

[38] P. Ferreira and M. Sirbu, “The Price of Anarchy for Interconnected Communication

Networks Provisioned Selfishly," in *Conference on Competition in Networks: Wired and Wireless*, London Business School, London, UK.

[39] W. Kim and K. Lee, "The Existence of Nash Equilibrium in n-Person Games with C-Concavity," *International Journal of Computers & Mathematics with Applications*, vol. 44.