Internet Interconnection Agreements

Burton Filstrup
212 N. Fifth Ave.
Ann Arbor, MI 48104
bfilstru@umich.edu

Final Project Report
SI646: Information Economics
Marshall VanAlstyne
May 2, 2001

1. Introduction

The Internet is a network of networks. In order to provide universal service, all the networks must interconnect to exchange traffic. Contracts for governing traffic exchange come in two very specific forms – peering agreements and transit agreements. The chosen contract depends primarily on the relative geographic scopes of the interconnecting networks. The limited range of feasible contracts between competing networks can be attributed to high transaction costs and information asymmetries. The cornerstone of this paper is a peering decision model that explains the observed structure of interconnection agreements on the Internet.

The networks on the Internet can be broken into a number of tiers [HUS99]. Lower-tier networks connect directly to businesses and consumers, but rely on higher-tier networks to deliver packets to their destinations. For simplicity, we consider only two tiers: Internet Service Providers (ISPs) and national backbone networks. Backbones span the entire U. S., delivering packets for other networks at the edge of the Internet (see Figure 1). ISP's sell Internet connectivity to consumers and businesses, relying on backbones to deliver packets to their destinations. Most backbone network operators also operate their own vertically integrated ISP.

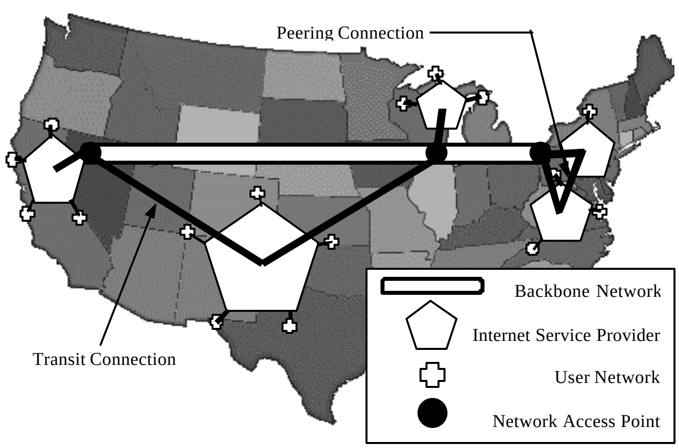

Figure 1 – Backbones and ISPs

Both the backbone and ISP industry are highly competitive. Figure 2 shows that there are over 7,000 ISP's in the market [ERI01]. The backbone industry is more concentrated, with WorldCom and Sprint cornering 43% of the market (see Figure 3). Based on the Boardwatch market share estimates, the Herfindahl-Hirschman index of market concentration for the backbone industry is 1300. This is a moderately concentrated market, according to the FTC merger guidelines [FTC97]. Due to low barriers to entry, this level of market concentration is unlikely to result in significant exertion of market power. Entrants can easily lease existing fiber, and we will see that interconnection agreements are available for any reasonably sized entrant. Therefore, in this paper we assume that the backbones and ISP's are engaged in perfect competition. This view is echoed by Michael Kende, Director of Internet Policy Analysis at the FCC [KEN00].

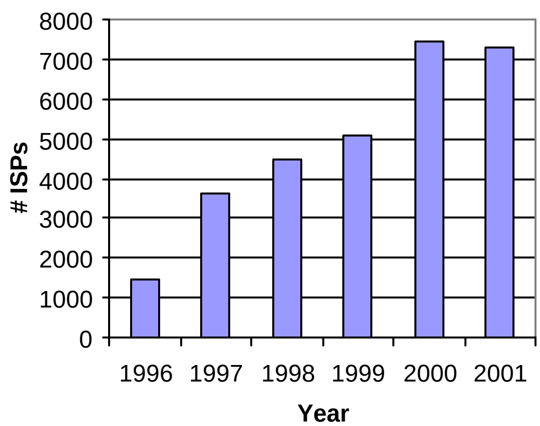

Figure 2 - Number of ISP's by Year

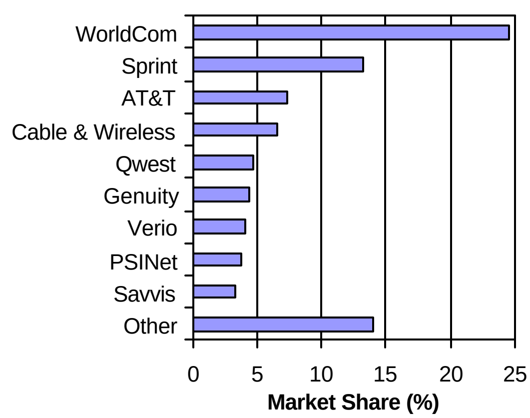

Figure 3 - Backbone Market Share

Networks exchange traffic at private or public exchange points. Historically, a small number of public exchanges, also termed Network Access Points (NAP's) were the primary exchange points on the Internet [HUS99]. At MAE-East, a NAP located in New York City, all of the national backbones plus all of the regional ISP's in the New England area exchanged traffic over a shared switching fabric. These NAP's became congested due to a classic tragedy of the commons – abusers reaped the benefits of overusing the exchange, while all of the member networks shared the cost of running the exchange. Part of the problem was a lack of monitoring tools to enforce usage-based pricing. Also, the volume of traffic exchanged at these bottleneck points exceeded technological limits. Currently, over 80% of traffic is exchanged at private exchange points, where the exchange point is owned entirely by the two connecting networks [KEN00].

Networks enter into two types of interconnection agreements: peering agreements and transit agreements [HUS99]. Two peering networks agree to exchange traffic that is destined for their networks (see Figure 4). Peering agreements are always settlement-free, implying that both networks benefit equally from the traffic exchange. Transit agreements involve a payment from the customer network to the provider network. The transit provider will deliver traffic from the customer network to any destination on the Internet (see Figure 4). This is in contrast to peering agreements, where a network accepts a packet only if it is destined for that network. Note that there are no peering-with-settlement contracts.

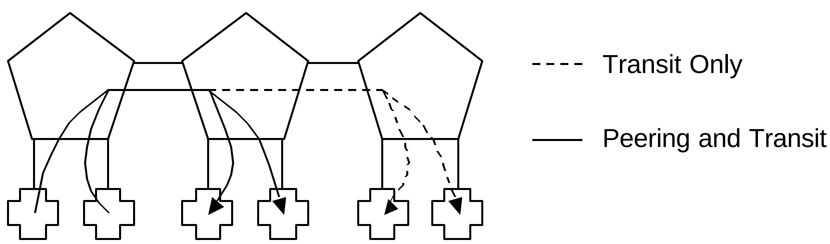

Figure 4 – Peering and Transit Traffic Flows

Because interconnection agreements are bilateral contracts between two businesses, the terms of interconnection are usually obscured by non-disclosure agreements [CER01]. Recently, a number of backbone networks, including WorldCom, have publicly released their criteria for selecting peering partners. WorldCom requires that the peering network, or requestor, operate a network that meets the following criteria [WOR01]:

- *Geographic scope* - The requestor must have a geographic scope at least 50% of the size of WorldCom. Specifically, the requestor must be able to exchange traffic at no less than four peering points across the nation.
- *Traffic exchange ratio* – The traffic across the peering points in each direction should be roughly balanced, and should not exceed a 1 : 1.5 ratio.

- *Backbone capacity* – The majority of the requestor's backbone should have a capacity of 622 Mbps.
- *Traffic volume* – The volume of traffic exchanged across the peering points should exceed 150 Mbps.

[KEN00] includes examples of instances where peering was discontinued because of the geographic scope, traffic exchange ratio or traffic volume criteria. Transit agreements remain secret, but the assumption is that any requestor that does not meet WorldCom's criteria must either buy transit from WorldCom or buy transit from a network that peers with WorldCom. Settlements for transit service are typically based on the capacity of the interconnection link [HUS99] [KEL99].

Peering becomes more common as the size of the network increases. At the top tier, national backbones of all sizes peer with each other in at many as 9 different peering points. Regional networks will usually peer with nearby regional networks, but they must always supplement peering with transit service from a top-tier network. Local ISPs tend to send all of their traffic upstream via a transit provider [HUS99].

This paper attempts to explain the observed structure of interconnection agreements. Section 2 briefly explains some technical details that come into play when designing interconnection agreements. In Section 3, a decision model for interconnection is presented. This model is used to explain the features of interconnection agreements in Sections 4 and 5. Section 4 explains the conditions under which two networks will peer with each other, and Section 5 proves that the only alternatives to peering agreements are transit agreements. The issue of free-riding is considered in Section 6. This project is compared to previous work in Section 7, and Section 8 concludes.

2 Network Properties

Three properties of the Internet are featured in our interconnection decision model. First, the quality of a connection between two users on the Internet is determined by speed of the slowest network along the connection's path. Second, the costs of transmitting a message are borne primarily by the destination network. Third, it is technologically impossible for a network to determine the precise load placed upon it by a connected network.

When two users communicate on the Internet, they generate a flow of traffic. This flow will usually consist of traffic traveling in both directions. In order to get from the source to the destination, a packet in the flow must travel through several networks. For example, a download from a Microsoft web server to a student at the University of Michigan would be passed from Microsoft's private network to Cable & Wireless's backbone network. Cable & Wireless operates a peering point with Michigan's regional network in Chicago, which routes the flow into the University of Michigan's network. The length of the path taken from Seattle to Michigan does not significantly affect the speed of the download. What really matters is how much bandwidth is available on the most congested network along the path. For example, if MichNet is heavily congested and can deliver only 1 kilobyte per second, my 10 kilobyte file will take ten seconds to download, regardless of how much faster the other networks are, or whether the traffic is exchanged in Chicago or Alaska. The takeaway point is that the quality perceived by the end user is determined by the network along that path with the least available bandwidth.

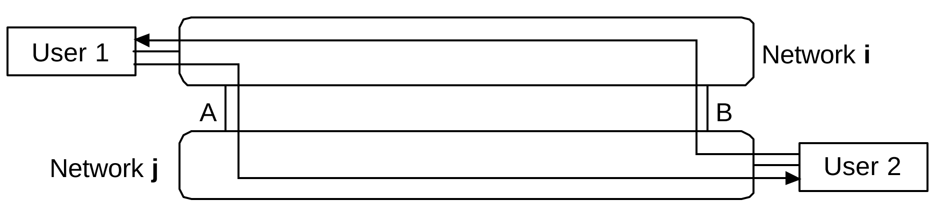

Figure 5 – Hot-Potato Routing

Two networks **i** and **j** often peer at multiple locations. Consider a one-way packet flow from network **i** to network **j**. Network **i** prefers to exchange traffic at the peering point A closest to the sender, because this will decrease the load on network **i** and increase the load on network **j** (see Figure 5). Because network **i** handles the packet first, it may unilaterally decide to exchange traffic at peering point A. Flows in the reverse direction will always be exchanged at peering point B. If an equal amount of traffic is flowing in each direction, the two networks will share the cost of the internetwork traffic equally. This phenomenon is known as hot-potato routing.

On the Internet, packets are routed towards their destinations on a hop-by-hop basis through a series of routers. If we were to sample a packet at a router, the only information available about that packet would be its source, its destination, and its next hop. Given this information, it is extremely difficult to figure out what links in the network will be traversed by the packet. For example, when Cable & Wireless receives a packet from Microsoft, it can't easily tell whether that packet is destined for Japan, in which case it would be expensive to deliver, or whether it will exit Cable & Wireless's network immediately. The upshot is that while measuring how much traffic is being exchanged between two networks is easy, calculating how much load the aggregated traffic places on the network is currently impossible.

This limitation may soon be eliminated. In a recent series of papers [FEL00a] [FEL00b] [FEL01], researchers at AT&T Labs have developed a method of measuring traffic flows. Although the process is complex, it could become a standard tool for measuring traffic flows in large backbone networks. Although the tool is geared towards traffic engineering, it also has implications for interconnection agreements. These thoughts are left until the conclusion.

3 Peering Decision Model and Assumptions

This section sets up the model that will be used to dissect interconnection agreements. The model consists of two cost functions, one for transit interconnection and one for peering interconnection. A number of minor assumptions are made about the costs of delivering traffic on the Internet.

The primary assumption in this model is that the backbone networks are perfectly competitive. This is in sharp contrast to the model in [BAA98], where there are a small number of networks engaged in Cournot competition. At discussed in the Introduction, it appears that the former assumption is a more accurate description of the U. S. market.

Because the market is perfectly competitive, networks cannot affect the market price by adjusting output. This does not rule out the possibility that a network may increase the price of its service by delivering superior or inferior quality service to each destination network. With these assumptions, the inverse demand function would resemble Equation 1.

$$\mathbf{p}_i = \sum \mathbf{u}_j \, (\mathbf{a} + \mathbf{b} \min \{\mathbf{q}_i, \mathbf{q}_{i+1}, \ldots, \mathbf{q}_j\}) \qquad \text{(Equation 1)}$$

The symbols are defined in Table 1. The users in network **i** pay for a bundle of connectivity to every other network **j** on the Internet. The inverse demand function implies a network effect; the value of the Internet to users increases as the number of users increases. As discussed in Section 2, the quality of connections traversing networks **i**, **i**+1, ... , **j** is the minimum available bandwidth among those networks. A typical path through the Internet traverses at least four networks, so the quality seen by a customer will rarely be constrained by the available bandwidth in network **i**'s backbone. Therefore, network **i** cannot control $\mathbf{p}_i$, and networks are participating in a competitive commodity market.

Symbol	Description	Dimension
$\mathbf{p_i}$	price charged per user by network **i**	$
$\mathbf{u}_j$	number of users subscribed to network **j**	users
a	value of connecting	$ per user
b	value of connecting	$ per user per Kbps
$\mathbf{q}_i$	the available bandwidth in network **i**	Kbps

Table 1 – Symbol Definitions for the Inverse Demand Function

The number of flows exchanged between two networks is a function of the product of their sizes.

$$\mathbf{t}_{ij} = \mathbf{d} \, \mathbf{u}_i \, \mathbf{u}_j$$

This is a reasonable assumption, since the total possible number of connections crossing the peering point is $\mathbf{u}_i \, \mathbf{u}_j$. The number of flows multiplied by the bandwidth per flow, min $\{\mathbf{q}_i, \mathbf{q}_j\} \, \mathbf{t}_{ij}$, yields the traffic volume exchanged. Traffic may be asymmetric, and the total traffic flow is the sum of the traffic flows in each direction.

$$\mathbf{t}_{ij} = \mathbf{t}_{i\text{-}j} + \mathbf{t}_{j\text{-}i}$$

The cost to network **i** of servicing flows to network **j** depends on whether the two networks decide to peer. If the two networks decide not to peer, then the networks will have to exchange traffic via their transit providers, and the costs of handling internetwork traffic is given by Equation 2, where the symbols are defined in Table 2.

$$\mathbf{c}_{ij} = \mathbf{f}_i\, \mathbf{q}_{ij}\, \mathbf{t}_{ij} + \mathbf{g}\, \mathbf{t}_{j\text{-}i}\, \mathbf{q}_i\, \mathbf{l}_i \qquad \text{(Equation 2)}$$

The first product represents the payment made to the transit provider for handling a given traffic volume $\mathbf{q}_{ij}\,\mathbf{t}_{ij}$. The coefficient $\mathbf{f}_i$ is then the price per Kbps of transit service for network **i**. Recall from Section 1 that transit service is usually based on the connection capacity instead of being based on the actual bandwidth usage. We assume that the capacity of the interconnection link is a good approximation of the bandwidth usage. The second product in Equation 2 represents the cost of transporting the flows through the internal network. Again, this is an increasing function of the number of flows $\mathbf{t}_{j\text{-}i}$ destined for network **i** and the throughput of the internal network $\mathbf{q}_i$. The cost of passing the data through the internal network also depends of the diameter of the network $\mathbf{l}_i$. The larger the diameter, the more routers must be traversed, the higher the cost of servicing a flow.

If the two networks decide to peer, their costs can be calculated using Equation 3.

$$\mathbf{c}_{ij} = \mathbf{m} + \mathbf{h}\, \mathbf{q}_{ij}\, \mathbf{t}_{ij} + \mathbf{g}\, \mathbf{t}_{j\text{-}i}\, \mathbf{q}_i\, \mathbf{l}_i \qquad \text{(Equation 3)}$$

One important assumption is that the internal network cost is independent of the peering decision. In other words, the interconnection points will be the same regardless of whether they are used for peering or transit. The first difference between peering and transit is that to peer the networks must invest in an interconnection point, which typically consists of collocated routers. The coefficient **m** also describes the other administrative work that goes on in order to set up an interconnection, such as signing agreements, researching each other's networks in order to remove any uncertainties, monitoring the link, establishing communications between network operators, etc. The only other difference between peering and transit is that traffic through the interconnection point costs **h** per packet instead of $\mathbf{f}_i$. Note that **m** and **h** are not dependent on the peering networks because we assume that the two networks are physically adjacent to each other. Understanding peering decisions between geographically separated networks would require an extension to this model.

Symbol	Description	Dimension
$\mathbf{c}_{ij}$	cost of servicing flows from network **i** to network **j**	$
$\mathbf{t}_{ij}$	total traffic between network **i** and network **j**	flows
$\mathbf{t}_{i\text{-}j}$	traffic from network **i** to network **j**	flows
d	proportion of possible flows that are active	ratio
$\mathbf{f}_i$	cost of transit service	$ per Kbps
h	variable cost of peering	$ per Kbps
m	fixed cost of peering	$
g	variable internal link service cost	$ per Kbps per link
$\mathbf{l}_i$	network diameter	links per flow
$\mathbf{q}_{ij}$	the available bandwidth across the peering point between networks **i** and **j**	Kbps

Table 2 – Symbols Definitions for the Cost Functions

As a simplification, we argue that both networks will want to set the interconnection quality $\mathbf{q}_{ij} = \min\{\mathbf{q}_i, \mathbf{q}_j\}$. Setting $\mathbf{q}_{ij}$ higher will be a waste of money, since the overall connection quality is constrained by $\mathbf{q}_i$ and $\mathbf{q}_j$. Setting $\mathbf{q}_{ij}$ lower than $\mathbf{q}_i$ is also inadvisable, since it will lower the overall quality of the connection between network **i** and network **j** without decreasing costs significantly. Assume that $\mathbf{h} \approx \mathbf{g}$, then

$$\mathbf{q}_{ij} << \mathbf{q}_i \, \mathbf{l}_i$$

$$\mathbf{h} \, \mathbf{q}_{ij} \, \mathbf{t}_{ij} << \mathbf{g} \, \mathbf{t}_{j\text{-}i} \, \mathbf{q}_i \, \mathbf{l}_i$$

In other words, passing traffic through the internal network is far more expensive than peering, so it would make no sense to cut costs and quality just at the peering points. The fact that backbones have made a strong move towards private exchanges indicates that they are willing to establish high-performance peering points. With these assumptions, our new model is described by Equations 4 and 5.

$$\mathbf{c}_{ij} = \mathbf{f}_i \min\{\mathbf{q}_i, \mathbf{q}_j\} \, \mathbf{t}_{ij} + \mathbf{g} \, \mathbf{t}_{j\text{-}i} \, \mathbf{q}_i \, \mathbf{l}_i \qquad \text{(Equation 4)}$$

$$\mathbf{c}_{ij} = \mathbf{m} + \mathbf{h} \min\{\mathbf{q}_i, \mathbf{q}_j\} \, \mathbf{t}_{ij} + \mathbf{g} \, \mathbf{t}_{j\text{-}i} \, \mathbf{q}_i \, \mathbf{l}_i \qquad \text{(Equation 5)}$$

4 The Peering Decision

The profit maximization problem for each network involves choosing $\mathbf{q}_i$ and deciding whether to peer with each network or use a transit provider. Section 6 addresses the choice of $\mathbf{q}_i$. For now, we assume that a network has chosen $\mathbf{q}_i$, and need only decide whether to peer with a network **j**. In a perfectly competitive environment there are no price effects, so the decision is purely based on cost. If a network can decrease its marginal costs below the industry average, that network can turn an economic profit. There are four conditions that must be true for peering to occur.

The first three conditions are derived from Equation 6, which states that network **i** will want to peer with network **j** only if they share the cost of servicing the internetwork flows equally. If network **i** bears more of the cost, then its average marginal cost will exceed the average marginal cost of network **j**. Because this is a perfectly competitive market, network **i** would exit the market and network **j** would turn a profit.

$$\mathbf{h} \min \{\mathbf{q}_i, \mathbf{q}_j\} \mathbf{t}_{ij} + \mathbf{g}\, \mathbf{t}_{j\text{-}i}\, \mathbf{q}_i\, \mathbf{l}_i = \mathbf{h} \min \{\mathbf{q}_i, \mathbf{q}_j\} \mathbf{t}_{ij} + \mathbf{g}\, \mathbf{t}_{i\text{-}j}\, \mathbf{q}_j\, \mathbf{l}_j \qquad \text{(Equation 6)}$$

Equation 6 is true if the following three conditions hold:

$$\mathbf{t}_{j\text{-}i} = \mathbf{t}_{i\text{-}j} \qquad \text{(Condition 1)}$$

$$\mathbf{q}_i = \mathbf{q}_j \qquad \text{(Condition 2)}$$

$$\mathbf{l}_i = \mathbf{l}_j \qquad \text{(Condition 3)}$$

Condition 1 requires that the traffic flows in each direction be balanced. Condition 2 specifies that the networks must be capable of providing the same bandwidth to the flows. Condition 3 implies that traffic should traverse the same number of links in each network. These three conditions roughly correspond to WorldCom's Traffic Exchange Ratio, Backbone Capacity, and Geographic Scope criteria.

The fourth condition is derived from Equations 7 and 8, which state that network **i** will want to peer with network **j** only if peering is a better option than using a transit provider, and vice versa.

$$(\mathbf{f}_i - \mathbf{h}) \min \{\mathbf{q}_i, \mathbf{q}_j\} \mathbf{t}_{ij} - \mathbf{m} > 0 \qquad \text{(Equation 7)}$$

$$(\mathbf{f}_j - \mathbf{h}) \min \{\mathbf{q}_i, \mathbf{q}_j\} \mathbf{t}_{ij} - \mathbf{m} > 0 \qquad \text{(Equation 8)}$$

Because $\mathbf{t}_{j-i} = \mathbf{t}_{i-j}$, $\mathbf{q}_i = \mathbf{q}_j$, and $\mathbf{l}_i = \mathbf{l}_j$, the networks must be roughly symmetrical, and $\mathbf{f}_i = \mathbf{f}_j$. Therefore, Equations 7 and 8 are equivalent, and can be simplified to Condition 4.

$$\mathbf{q}_i \mathbf{t}_{ij} > \mathbf{m} (\mathbf{f}_i - \mathbf{h})^{-1} \qquad \text{(Condition 4)}$$

Condition 4 states that the traffic flow between the two networks $\mathbf{q}_i \mathbf{t}_{ij}$ must be larger than some constant. This condition corresponds to WorldCom's Traffic Volume peering criterion.

The first condition guards against asymmetric traffic flows. If more traffic is flowing from network **j** to network **i** than in the reverse direction, then a greater load will be placed on network **i**, and network **i** will have higher marginal costs. This is due to hot-potato routing, where packets are delivered primarily by the destination network (see Section 2). Asymmetric traffic flows are common on the Internet because WWW traffic flows from the web server to the web browser. In the aggregate, most networks have the same proportion of web servers and web browsers, so this is not a problem. However, networks such as Exodus that specialize in web hosting services can generate asymmetric traffic flows. WorldCom's Traffic Exchange Ratio policy specifies that $\mathbf{t}_{j-i} / \mathbf{t}_{i-j} < 1.5$.

The second condition ensures that the networks deliver the same bandwidth to internetwork flows. If network **i** overprovisions its infrastructure relative to network **j**, then network **i** will have higher marginal costs. The throughput of flows is $\min \{\mathbf{q}_i, \mathbf{q}_j\}$, so the increase in $\mathbf{q}_i$ is not matched by an increase in service to network **i**'s customers, nor does it increase network **i**'s revenue. WorldCom's Backbone Capacity criterion requires requestors to meet a minimum backbone capacity. The aggregate backbone capacity is a very poor approximation for $\mathbf{q}_i$, since $\mathbf{q}_i = \mathbf{capacity} / \mathbf{t}_{j-i}$, i.e. their metric does not take into account the load on the requestor's backbone. This indicates that WorldCom is not serious about enforcing Condition 2. The reason for this is covered in Section 6.

The third condition says that the number of links traversed on each network should be the same. If network **i** carries flows over more links on average than network **j**, then network **i** will have higher marginal costs. This condition is met if the two networks peer at uniformly distributed points across their respective networks. Figure 6 shows a scenario where $\mathbf{l}_i \neq \mathbf{l}_j$. National backbone networks peer at a number of sites across the country, so they all have symmetric link counts. However, a regional network in Virginia may only peer at one site on the East Coast. These regional networks are not allowed to peer with national backbone networks. WorldCom's Geographic Scope clause requires peering at a number of geographically distributed sites so that $\mathbf{l}_i = \mathbf{l}_j$.

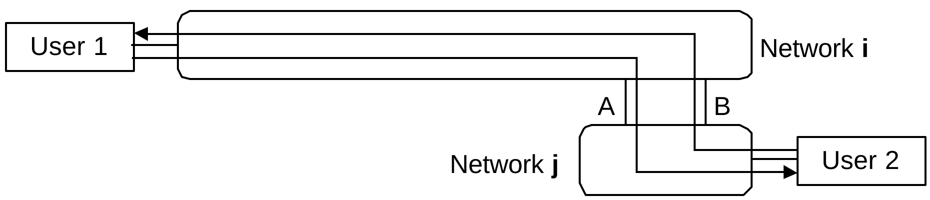

Figure 6 – Asymmetric Networks

According to the fourth condition, peering is only better than transit if the amount of traffic exceeds a certain amount. If the exchanged traffic volume is small, then it may not be profitable to set up a private peering point to route traffic. Because the exchanged traffic is proportional to $\mathbf{u}_i \mathbf{u}_j$, the peering decision depends on the product of the number of users in each network. This hypothesis is consistent with the Internet, where local ISPs never peer, regional ISPs peer with other regional ISPs and with national ISPs, and national ISPs peer with every other national ISP. It also explains the rationale behind WorldCom's Traffic Volume peering criteria, which specifies a minimum amount of traffic to be exchanged at peering points.

In order for this peering decision model to be accurate, the conditions must be observable. Network **i** must be able to measure $\mathbf{t}_{j\text{-}i}$, $\mathbf{t}_{i\text{-}j}$, $\mathbf{q}_i$, $\mathbf{q}_j$, $\mathbf{l}_i$, $\mathbf{l}_j$, $\mathbf{m}$, $\mathbf{f}_i$, and $\mathbf{h}$. The routers at peering points automatically monitor the traffic volumes in each direction, min $\{\mathbf{q}_i, \mathbf{q}_j\}\, \mathbf{t}_{j\text{-}i}$ and min $\{\mathbf{q}_i, \mathbf{q}_j\}\, \mathbf{t}_{i\text{-}j}$. These can be used directly to calculate the traffic flow ratio $\mathbf{t}_{j\text{-}i} / \mathbf{t}_{i\text{-}j}$, which can be used to test

Condition 1. The values $\mathbf{q}_i$ and $\mathbf{q}_j$ are very difficult to measure, since they represent per-flow available bandwidth, so Condition 2 cannot be enforced. The equivalence of $\mathbf{l}_i$ and $\mathbf{l}_j$, required by Condition 3, is easy to ascertain by looking at the distribution of peering points. A network knows its own $\mathbf{f}_i$, $\mathbf{m}$, $\mathbf{h}$, and min $\{\mathbf{q}_i, \mathbf{q}_j\}$ $\mathbf{t}_{ij}$, so it can evaluate Condition 4.

In summary, two networks will peer only if they ascertain that the costs of carrying traffic are symmetric and that peering is more cost efficient than using a transit provider. These requirements can be expressed as four simple conditions based on predicted cost functions. The conditions explain the rationale behind WorldCom's peering policy. The two networks can verify all but the second condition.

5 Transit Agreements

The model predicts that two networks will not peer without settlement if they are asymmetric, i.e. if the first three conditions are not met. If the fourth condition is met, peering is still a profitable alternative. It is natural to expect two asymmetric networks to use a settlement to remedy any cost imbalance. However, high transaction costs make peering agreements with settlement unattractive, so networks usually resort to a single transit agreement instead. Also, transit providers can use risk-sharing to reduce the risks of interconnection.

If network $\mathbf{i}$ decides that there is a significant cost imbalance, it can demand settlement from network $\mathbf{j}$. The settlement should be calculated using Equation 9.

$$\mathbf{g}\,(\mathbf{t}_{j\text{-}i}\,\mathbf{q}_i\,\mathbf{l}_i - \mathbf{t}_{i\text{-}j}\,\mathbf{q}_j\,\mathbf{l}_j) \qquad \text{(Equation 9)}$$

Unfortunately, none of these variables is easily ascertainable. The result is a great deal of uncertainty over the magnitude of the settlement. In order to cope with this uncertainty, interconnecting networks may have to perform costly research and engage in costly negotiation. The increased transaction costs $\mathbf{m}$ dictate that it is more efficient to establish a single transit connection than a large number of peering-with-settlement relationships. Furthermore, the transit providers have expertise in evaluating networks, so even the single negotiation can be streamlined. The following paragraphs describe the degree of uncertainty of each variable in Equation 9 and explain how transit providers can reduce this uncertainty.

The cost **g** of passing a flow over an internal link is difficult to calculate. In order for a backbone to be profitable, the peering network should pay not only a share of the link installation and operating costs, but also a portion of the administrative overhead. If a network has peers and customers of different sizes, calculating **g** may be tricky. Transit providers enter into a large number of interconnection agreements, so they should have the expertise to find the correct value of **g**.

Determining the values of $\mathbf{t}_{j-i}$ and $\mathbf{t}_{i-j}$ is a straightforward procedure ex post. The number of flows is measured by standard hardware in high-end routers. Ex ante, the parties may need to set up a test circuit in order to determine these values. Although this could potentially be costly, it should be a routine procedure for transit providers.

Perhaps the most uncertain variable is $\mathbf{l}_i$, the average number of links that network **j**'s packets will traverse over network **i**. The value of $\mathbf{l}_i$ depends first on the location of peering points, which are directly observable to both parties. However, it also depends on the number of peering relationships that network **j** establishes with other nearby networks. If network **j** dumps all of its inexpensive traffic through peering relationships, network **i** will end up delivering only expensive traffic. Network **j**'s peering relationships are secret, so network **i** cannot calculate $\mathbf{l}_i$. Even worse, $\mathbf{l}_i$ cannot be measured with existing tools. As discussed in Section 2, tools for measuring $\mathbf{l}_i$ are just now coming online [FEL01]. However, these tools could be too expensive for all but the largest transit providers.

For each of the variables, imperfect measurement of relative costs presents opportunities for moral hazard. The customer has an incentive to step up its auxiliary peering agreements after the transit contract is signed, leaving the transit provider with the expensive traffic. Likewise, the customer has an incentive to reduce its available bandwidth and enlist web browsers instead of web servers.

Even an expert transit provider cannot resolve all of the uncertainties involved with interconnection agreements. Because transit providers negotiate a large number of contracts, they can use risk-sharing to reduce the risks of interconnection. This makes transit providers more efficient than a standard network, which would presumably need to buy insurance.

6 Free Riding

In Section 4, we said that Condition 2, which requires peers to allocate the same bandwidth to flows, is difficult to enforce. This has dire consequences for infrastructure investment on the Internet. Imagine that the all networks on the Internet provide the same $\mathbf{q}_i$. Then no network will ever upgrade its infrastructure, because increasing $\mathbf{q}_i$ unilaterally would have no impact on the performance seen by network **i**'s customers. In fact, every network **i** has an incentive to decrease its $\mathbf{q}_i$. The reason is that the cost savings would accrue to network **i**, while the decrease in performance would be shared by all of the other networks. The problem is termed free-riding, and the result is under-investment in the Internet [HUS99].

Previously, we argued that WorldCom's Bandwidth Capacity requirement was a weak instrument for ensuring that requestors had the same $\mathbf{q}_i$. The difficulty of measuring $\mathbf{q}_i$ offers a partial explanation for this. A better explanation is that WorldCom is trying to pressure other networks into upgrading their backbones. If WorldCom refuses to peer with a requestor, they could be forced to increase costs by purchasing transit to WorldCom's network. This theory is further supported by WorldCom's request that other major backbones publish similar peering guidelines.

7 Related Work

Interconnection agreements are an interesting topic because the parties involved are in direct competition with each other. The economics of these agreements has been discussed in a number of papers, starting with Bailey [BAI96] and Srinagesh [SRI96] in 1995 when the Internet was first privatized and interconnection agreements had not yet matured. Since then, two peering decision models have emerged. The ISPs in Europe have been described as being engaged in Cournot competition with each other [BAA98] [REI99], and Baake and Wichmann developed a model to explain their behavior. Peering decisions in the competitive U. S. market have been loosely characterized by Huston [HUS99].

This paper was inspired by "On the Economics of Internet Peering," by Pio Baake and Thorsten Wichmann [BAA98]. They develop a peering decision model where two ISPs are in Cournot competition with each other. That is, they are the only two firms in the German market and can therefore make positive profits by restricting output. The two networks can both reduce cost by peering, but the larger network may decide not to peer in order to preserve its competitive advantage. The decision is not whether to peer, but rather how good the interconnection between the two networks should be. The model allows for a settlement between the two networks. While this model may provide insight about the European industry, it does not describe the U. S. market, which is highly competitive. American ISPs do not care about the quality of the interconnection, only about whether or not peering occurs. The Baake-Wichmann model does not explain any of the WorldCom peering criteria, and it sheds no light on why peering with settlement is not evident on the Internet.

Geoff Huston, an industrial insider, has been writing about interconnection since before the Internet was privatized in 1995. His article, "Peering and Settlements", is a comprehensive and accessible overview of interconnection [HUS99]. Huston's description of peering decisions hints at the model presented in this paper. He postulates that the peering decision depends on the relative "values" provided by the two networks, where a network's value is measured by its geographic spread, backbone capacity, total traffic volume and number of customers. Our model concurs on the first property (Condition 3), but endorses a different view on the last three. The total backbone capacity matters only in that it may impact the available bandwidth on a network (Condition 2). The total traffic volume is not a direct concern, but the traffic volume exchanged across the peering points does (Condition 4). The number of customers determines the traffic volume, but otherwise does not factor into the peering decision.

In his FCC report, Kende provides a thorough review of the backbone industry, including an endorsement of Huston's peering analysis. He cites WorldCom's peering policy and provides examples of these policies in action [KEN00]. In 1997, UUNet discontinued peering relationships with regional ISPs because they did not have the requisite geographical scope. In 1998, PSINet refused to peer with Exodus because of an unfair traffic exchange ratio. Also in 1998, MCI stopped peering with Level-3 because it could not exchange enough traffic volume.

8 Conclusions

Based on some simple assumptions about the cost of servicing internetwork flows, a peering decision model has been developed that explains peering policies on the Internet. The three enforceable conditions for peering are balanced traffic flow, comparable geographic scope and a minimum exchanged traffic volume. The relative costs of peering also depend on differences in available bandwidth, but this quantity is difficult to measure. Because the quality of network infrastructures cannot be enforced, free-riding is a serious problem.

If two networks are asymmetric, the smaller network must pay for interconnection. Because of the high transaction costs of settlement, the smaller network will enter into a single transit agreement instead of multiple peering with settlement agreements. The existence of transit providers that handle a large number of transit agreements helps mitigate the risks of interconnection.

The lack of peering agreements with settlement means that some efficient peering arrangements will never be established. The major obstacles to these relationships are the high transaction costs of settlement and the high risks of interconnection. Automated settlement protocols are needed to reduce transaction costs. Improved network traffic measurement tools would reduce risks. However, these new systems would require a massive overhaul of the Internet. Until the Internet industry matures, this type of massive coordinated expenditure is unlikely.

The peering decision model has implications on antitrust regulation. Understanding the peering decision process will help the authorities determine whether a network refuses to peer for anticompetitive reasons, or because of network asymmetries. The FCC's study on the possible anticompetitive nature of Internet interconnection agreements was on the right track [KEN00], but an analytical peering decision model would strengthen its conclusions.

9 References

[AWD98] D. Awduche, J. Agogbua and J. McManus. "An Approach to Optimal Peering Between Autonomous Systems in the Internet." Proceedings of IC3N98, Lafayette, Louisiana. October, 1998.

[BAA98] P. Baake and T. Wichmann. "On the Economics of Internet Peering." Netnomics, vol. 1, no. 1. 1999.

[BAI95] J. Bailey. "Economics and Internet Interconnection Agreements." Journal of Electronic Publishing. May, 1996.

[CER01] V. Cerf. Interview on Private Peering. Available at http://www.zdnet.com/intweek. January, 2001.

[ECO96] N. Economides, G. Lopomo and G. Woroch. "Strategic Commitments and the Principle of Reciprocity in Interconnection Pricing." Discussion Paper no. EC-96-13, Stern School of Business, NYU. 1996.

[ECO92] N. Economides and G. Woroch. "Benefits and Pitfalls of Network Interconnection." Discussion Paper no. EC-92-31, Stern School of Business, NYU. 1992.

[ERI01] T. Erickson. "Introduction to the Directory of Internet Service Providers, 13th Edition." Available at http://www.ispworld.com/. March, 2001.

[FEL00b] A. Feldmann, A. Greenberg, C. Lund, N. Reingold, J. Rexford and F. True. "Deriving Traffic Demands for Operational IP Networks: Methodology and Experience." Proceedings of ACM SIGCOMM. August 2000.

[FEL01] A. Feldmann, C. Lund, N. Reingold and J. Rexford. "NetScope: Traffic Engineering for IP Networks." IEEE Network Magazine. 2000.

[FEL00a] A. Feldmann and J. Rexford. "IP Network Configuration for Traffic Engineering." Technical Report, AT&T Labs. May 2000.

[FTC97] "Horizontal Merger Guidelines." from the Federal Trade Commission and the Department of Justice. Available at http://www.ftc.gov/. April, 1997.

[HUS99] G. Huston. "Peering and Settlements Parts 1 and 2." Internet Protocol Journal, vol. 2, no. 1 and no. 2. March 1999.

[KEL99] T. Kelly. "Pricing Internet Services: Trends and Tactics." Presentation at International Telecommunication Union Workshop. November, 1999.

[KEN00] M. Kende. "The Digital Handshake: Connecting the Internet Backbones." OPP Working Paper, Federal Communication Commission. September, 2000.

[REI99] P. Reichl and B. Stiller. "Pricing Models for Internet Services." CATI Deliverable CATI-TIKSWI-DN-P-008-2.2, ETH Zürich. March 1999.

[SRI95] P. Srinagesh. "Internet Cost Structures and Interconnection Agreements." Journal of Electronic Publishing. May, 1996.

[WOR01] "WorldCom Policy for Settlement-Free Interconnection with Internet Networks." WorldCom. Available at http://www.uu.net. January, 2001.