Internet Regulation:
Monopolistic Bottlenecks in Internet Service Markets?

Margit A. Vanberg*

ZEW (Centre for European Economic Research), Mannheim

August 2003

Abstract

The paper focuses on the cost characteristics of Internet technology and on the question whether there are monopolistic bottlenecks in Internet services which justify regulatory intervention into the market. The analysis is prompted by a discussion which followed the MCI and Worldcom merger in 1998. In this discussion it was often claimed that top-level ISPs have market power, due to their high market shares in backbone traffic and backbone connectivity. The paper proposes a separate analysis of the cost characteristics of the Internet, which could be the source of monopolistic bottlenecks in Internet services markets and the network effects of the Internet, which could lead to strategic actions in Internet interconnection agreements. By separating the analysis of the cost characteristics of the network elements making up the Internet from the issue of universal connectivity in Internet interconnection the paper comes to the conclusion that the claim of market power in the Internet backbone industry is not substantiated by monopolistic bottlenecks in backbone network capacity.

Keywords: Internet, Networks, Interconnection, Regulation, Competition Policy.
JEL Classification: D4, K2, L51; L96

* Research Fellow, Research Group Information and Communication Technologies, ZEW (Centre for European Economic Research); P.O.Box 103443, D-68034 Mannheim, Germany; E-Mail: vanberg@zew.de

1 Introduction

Developments in the market for Internet backbone capacity and especially the merger of two of the four largest firms in this market, Worldcom and MCI, in 1998 and the failed merger of MCI/Worldcom and Sprint Corporation in 2000 spurred a large body of literature on the role of antitrust policy in Internet Interconnection. Unlike most other telecommunications activities, Internet service provision evolved in a largely unregulated environment.[1] The recent discussions surrounding Internet backbone activities suggest, however, that it is by far not certain that this market will remain unregulated.

The concerns raised with respect to the market for Internet service provision focus on two issues. The first is the fact that a high proportion of Internet traffic is transported on the capacity of only a few top-level networks. The second concern addresses the strong network externalities in the Internet industry. End users demand *universal connectivity* from their Internet Service Provider (ISP), that is, the ability to communicate with all other Internet users, irrespective of the network they are connected to. ISPs can achieve this universal connectivity only by interconnecting (directly or indirectly) with at least one top-level backbone network (European Commission, 2000: 13). Both concerns, it is said, show that the competitiveness in the top-level Internet backbone market is essential for the competitiveness of the entire industry.

The present paper argues that the discussion on the potential need for regulation in the Internet is disorganized. Markets closely related to the Internet are often confused with the Internet itself. In particular, the closely related telecommunications market is mixed up with Internet service provision. Since telecommunications markets are still governed by many sector-specific regulations, the insufficient separation of the Internet and telecommunications results in a demand for regulation of the Internet which is not separated clearly from existing regulations in telecommunications markets. In chapter 2, I introduce a separation of Internet service provision and Internet periphery as a basis for the subsequent policy analysis.

Chapter 3 provides a brief overview of the literature on the potential for regulation of the Internet. I argue that the two problem areas predominately being discussed are not differentiated clearly and that this is a reason why the literature remains vague on the proper regulatory remedies for the Internet. However, the two discussed potential sources of market

[1] In the USA, for instance, Internet Services are counted among the “information services“, which are not subject to common carrier regulation (Kende, 2000: 13).

would, if confirmed, call for very different regulatory remedies. I suggest a separate analysis of the potential sources of market power mentioned in the literature.

Chapter 4 presents a framework for analyzing the cost characteristics of network industries. This framework is used to analyze the question whether there are essential facilities in Internet service provision which justify regulatory intervention into the market. The analysis assumes a hypothetical world in which network effects are not an issue. By separating the issue of universal connectivity from the cost characteristics of the market, the analysis arrives at the conclusion that the cost characteristics of Internet service provision, do not justify ex-ante sector-specific regulation in long-distance network capacity or in Internet connectivity services. The analysis can therefore exclude one of the most often cited sources for potential market power in Internet backbones. What remains to be analyzed is the possibility that market power results from the network effects in the Internet connectivity market.

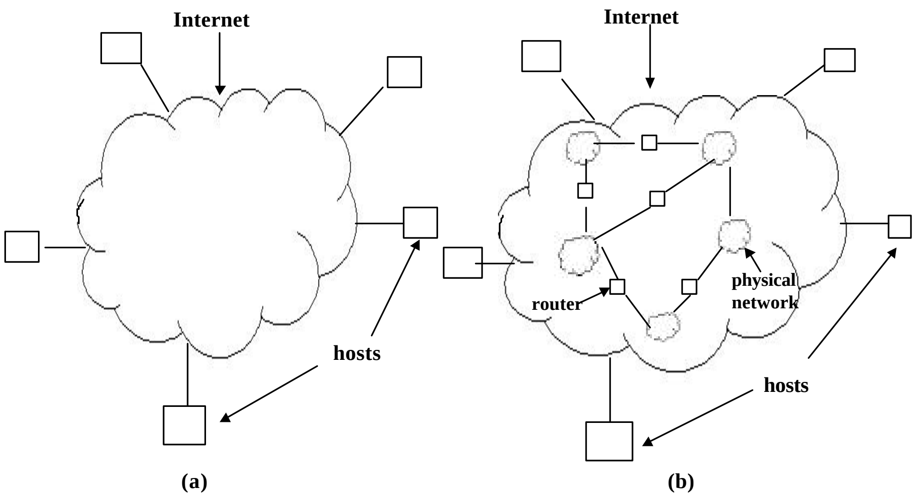
2 Defining the Internet

This chapter will give a short introduction into the technology of the Internet and especially into Internet interconnection. This understanding of the structure of the Internet is the basis for the policy discussion surrounding Internet backbones.

2.1 Differentiating Internet services and telecommunications services

Very often the Internet is loosely defined as “a network of networks”. Comprised in this definition is the characteristic of the Internet, that the many networks which make up the Internet are perceived by the end-user as a single large network (see Fig. 2.1 (a)).

Figure 2.1 The end-user's view of the Internet



(a) The user's view of a TCP/IP internet in which each computer appears to attach to a single large network, and
(b) the structure of physical networks and routers that provide the interconnection.

Source: Comer, 1995: 55

The user is not aware if content she downloads from the Internet is hosted on her home network or on an unrelated network. This is because the individual networks making up the Internet are made compatible by a set of common standards (see section 2.1), which govern network interconnection and data exchange. Comer places an emphasis on these standards, by defining the Internet as "...a method of interconnecting physical networks and a set of conventions for using networks that allow the computers they reach to interact" (Comer, 1995: 17).

An "official" definition of the term Internet was given by the Federal Networking Council (FNC) on October 24, 1995, when it unanimously passed a resolution defining the term Internet as:

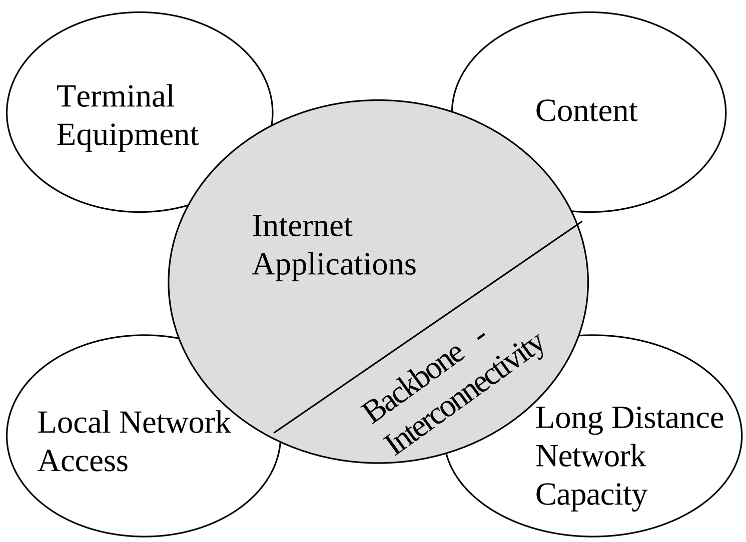
> "Internet" refers to the global information system that -- (i) is logically linked together by a globally unique address space based on the Internet Protocol (IP) or its subsequent extensions/follow-ons; (ii) is able to support communications using the Transmission Control Protocol/Internet Protocol (TCP/IP) suite or its subsequent extensions/follow-ons, and/or other IP-compatible protocols; and (iii) provides, uses or makes accessible, either publicly or privately, high level services layered on the communications and related infrastructure described herein. (Leiner et al., 2000:14).

As in Comer's definition, the FNC also emphasizes the common logic of data transportation in the Internet, the common address space and the standardized protocols.

That the Internet is not defined in reference to its physical network infrastructure is due to the fact, that Internet services are realized on a range of platforms, which have viable functions apart from the Internet. The platforms in use are for instance telecommunications networks, cable TV networks and wireless networks. Knieps (2003) proposes a differentiation of ***Internet Service Provision*** and ***Internet Periphery*** in order to distinguish primary functions of the Internet, as defined by the FNC, from complementary functions to Internet service provision, which are also viable apart from the Internet.

The Internet periphery (non-shaded areas in Figure 2.2) includes the complementary functions to Internet service provision. These are the infrastructure platforms, further distinguished into (1) **Local Network Access** (the infrastructure which joins the end-user's host and the Point of Presence (PoP) of the Internet Service Provider (ISP)) and (2) **Long Distance Network Capacity** (bandwidth for long-distance data transportation in the network of the ISP). The Internet periphery also includes (3) **Terminal Equipment** (Phones, PCs, Notebooks and Work-Stations) and (4) **Content** (music files, videos, news, magazine articles, etc.). Just like network infrastructure elements, terminal equipment and content are necessary complements to Internet services provision. However, they also have viable functions independent from the Internet.

Fig. 2.2 Internet Periphery vs. Internet Service Provision



Source: Based on Knieps (2003: 219)

Internet service provision (gray-shaded areas in Figure 2.2) includes (1) **Internet applications** (such as E-mail services, file-transfer, web-browsing etc.) and (2) **Backbone Interconnectivity** services (such as routing, network management etc.). Internet application services and Internet connectivity services have meaningful functions only in connection to the Internet. While Internet users are familiar with E-mail, file transfer etc, as these are the end-products supplied by ISPs, they usually do not come in contact with the intermediate services of the backbone interconnectivity market. Section 2.2 will introduce the reader unfamiliar with Internet technology to the basics of Internet connectivity.

Because the elements of the Internet Periphery are complementary services to Internet Service Provision, existing regulations in Internet periphery markets necessarily have an impact also on Internet service provision. For example, equal access regulations which govern local network access in telecommunications networks also eliminate a potential entry barrier in the market for Internet application services. The interdependency of the elements of the Internet periphery and the Internet core has even resulted in new regulations of Internet periphery elements, due to the functions that these elements have in their role as inputs in the markets for Internet service provision. For example, line sharing regulation was introduced in local network access in order to enable ISPs to offer Internet access services to customers of independent telecommunications carriers.

The differentiation of Internet service provision on one hand and Internet periphery on the other hand will be continued throughout the paper. Chapter 3 will show that this differentiation is not common in the literature discussing the potential need for regulation in the Internet. Because of this lack of differentiation, claims for regulation in the Internet have been rather diffuse. A separation of Internet periphery and Internet service provision helps in differentiating clearly between existing regulations in the Internet periphery and the demand for new regulations of Internet service provision.

2.2 Internet Technology

David (2001: 27) states correctly that economists wanting to discuss policy for the Internet need to understand the principles of its technology. This section will therefore give a very brief overview of the most central technological characteristics which play an important role in the market for Internet connectivity, before the discussion on the potential for market power in Internet services is reviewed in chapter 3.[2] This section can be omitted by readers familiar with Internet technology.

Circuit-switching vs. Packet-switching

Perhaps the most notable difference between traditional telephone technology and the communication technology of the Internet is that traditional telephone services are realized on the basis of *circuit-switching*, whereas the Internet is based on *packet-switching technology*. With circuit switching, a dedicated line between caller A and receiver B is set up for the entire duration of a call. This end-to-end link is pieced together by switching through several point-to-point segments between network nodes (i.e. switches) along the way. If only one cable segment or node on this link is lost, then the transmission is interrupted and would need to be set up anew.

Packet switching technology does not rely on an end-to-end link between the communicating parties. It decomposes the data stream into smaller segments, called “packets” or “datagrams”, which can be sent on a variety of different possible routes between A and B, before being reassembled in the correct order at their destination. Because every packet can potentially travel a different route, a defect cable segment or node would not disrupt the transmission. A further advantage of packet switching technology is its efficient use of network capacity. Circuit switching blocks the entire capacity of the switched-through link for the duration of the connection, even though the capacity is idle for much of the time, as during speech pauses. Packet switching instead uses capacity only when data is transmitted, such that the network can handle more than one communication link at a time by placing packets in the pauses of packets from an unrelated data transmission.

Internet Protocol

The rules for data transmission in the Internet are laid down in the software for data transport, the *Internet protocols*. The standard protocols used and supported by all networks in the Internet are part of the TCP/IP protocol family. The protocol is responsible for dividing the data stream into packets and for providing each package with information on the originating IP address (see next paragraph), the destination IP addresses, an indication which packages are part of one message, whether there are more packages to come and the order of the

[2] When no source is explicitly stated, the general understanding of the Internet technology is taken from Comer (1995), Siegmund (1999), Halsall (1996), the Internetworking Technology Handbook from Cisco Systems and on the Online Dictionary for Computer and Internet Terms „Webopedia“ at http://www.webopedia.com/.

packages belonging to one data stream. The packages are recompiled in the order of the original message at their destination.

IP Address

Each computer connected to the Internet is identified by at least one IP address. The Internet Assigned Number Authority (IANA) grants IP addresses to Internet registries, which in turn assign blocks of IP addresses to Internet Service Providers. ISPs finally assign IP-addresses to their customers. [3]

IP addresses identify both the home network of a host as well as the host itself (Comer: 1995:61). This is important because it allows routing to be based only on the information concerning the home network until the home network itself is reached. Within the home network, the information on the individual host will be evaluated. As a consequence of the fact that the address contains both information on the home network as well as an identifier for the host, a host connected to multiple networks will also have multiple IP addresses.

The format of an Ipv4 address is a 32-bit numeric address. It contains four numbers in the range from zero to 255, separated by periods. For example, 193.196.11.208 is the IP address of the site www.zew.de. A new generation of IP addresses, Ipv6, was deployed beginning in 1999. However, Ipv4 addresses continue to remain the most common.

Routing

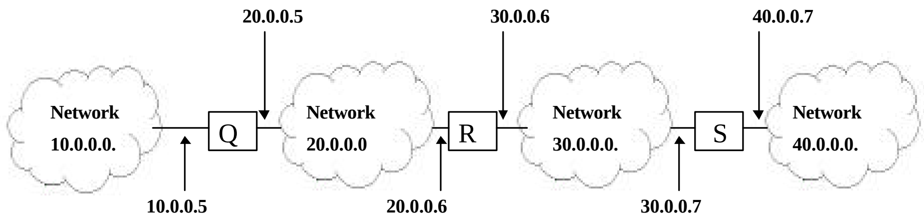
As mentioned above, data transmitted over the Internet is divided into small packages, each of which contains the IP address of the destination host. The transmission of the data packages to their destination is accomplished by so called *routers*. Routers are computers situated at network nodes and interconnection points where private lines and networks making up the Internet interconnect. The routers process incoming data packages according to rules specified in routing protocols. When more than one possible path to the final destination is available, ***routing*** refers to the process of deciding on a specific route.

The routing of data packages involves a router evaluating the IP address in the header of an incoming data package and determining the optimal routing path based on information given in its routing table. A ***routing table*** pairs IP addresses (destinations in the network) and routing paths (next router in line to the destination). Since the routing path is given as the IP address of the next router along the path, a router never knows the complete path to the final destination, but only the next hop the packet has to take in order to travel the optimal route.

This is called "next-hop routing". To keep the information which is stored in routing tables to a minimum, routing tables do not actually list all IP addresses, but only the network prefixes which are part of the IP address. Routing tables are furthermore kept to a minimum by the possibility of defining a default case. Most routers have a specified default route for all network prefixes which are not explicitly listed in the routing table.

Figure 2.3 illustrates the logic of next-hop routing. The routing table for router R specifies that all incoming packages with network prefixes 20.0.0.0 or 30.0.0.0 can be delivered directly because router R has a direct connection to both of these networks. For packages destined to networks 10.0.0.0 or 40.0.0.0 the routing table lists the IP address of the next-hop router, to which R has a direct connection.

Figure 2.3 Next-hop routing



(a)

TO REACH HOSTS ON NETWORK	ROUTE TO THIS ADRESS
20.0.0.0	DELIVER DIRECTLY
30.0.0.0	DELIVER DERECTLY
10.0.0.0	20.0.0.5
40.0.0.0	30.0.0.7

(b)

(a) An example internet with 4 networks and 3 routers, and
(b) the routing table in R.

Source: Comer, 1995: 114

Routing tables can be filled manually by network administrators. Routers, however, also "learn" new routing paths from other routers. Because the Internet is dynamic with many hosts being connected and disconnected at any given time, automated systems are clearly more reliable and stable compared to manual systems. The automated systems function with the help of protocols such as *the Routing Information Protocol* or the *Open Shortest Path First* (*OSPF) Protocol*. Using these protocols, routers advertise their current network capacity and the connections they can establish to neighboring routers and receive analogous

[3] See www.iana.org for details.

information. By this exchange of information routers acquire a complete picture of network topology within their region and fill their routing tables accordingly. The choice of the path a datagram is to take to its destination is based on the fixed assumptions of shortest path. When routers advertise which destinations can be reached via their network, they also advertise the number of hops (the number of routers passed) to the final destination via this particular route. A router will add that route to its routing table which minimizing the number of hops to the destination.[4]

Not all routers have the same functions and abilities. Rather, there is a hierarchy within the routing architecture. Routers at the bottom of the hierarchy have relatively small routing tables, as they are connected to only a few networks. For instance, routers of local area networks quite possibly have only one default route which is resorted to for all datagrams not destined for a host to which it can route directly. This default router is also called an interior gateway. Routers in the middle of the hierarchy will have routing information for several local networks reachable via interior gateways. This set of networks can be either a Wide Area Network or a set of independent Local Area Networks. These networks typically have a common default route for network addresses not within their group of networks. The default route leads to an exterior gateway.

Only routers at the top of the hierarchy have *complete* routing tables, containing information on all possible destinations within the Internet. The routing information of these routers suffices to direct datagrams to the correct gateway, which will send the datagram on to the router which can deliver the datagram to the destination host. Routers at the top of the hierarchy by definition have no default router. They need to know every single Internet destination in order to ensure consistent routing. These routers of the highest level in the router hierarchy are also called *core routers* (Comer, 1995: 235). What distinguishes core routers from all other routers within the Internet is that they have the responsibility to drop any datagram to which they cannot match a route and sent an error message to the originating device (Milgrom et al., 2000: 179-180 and Halsall, 1996: 506). This is important in order that inconsistencies in routing do not paralyze the system (as would for instance be the case if *routing loops* develop, by routers of connected networks, pointing their default routes at each other).

[4] More sophisticated software would use additional characteristics such as the network load, the quality requirements of the services specified in the datagram etc. (Comer, 1995: 110). While the actual hop-count does not necessarily reflect the fastest route to a destination, many routers will set hop counts artificially high for slower connections (Comer, 1995: 242).

2.3 Internet Interconnection

Internet interconnections agreements between Internet Service Providers take the form of either "peering" or "transit". The most important difference between peering and transit is that peering partners exchange traffic on a settlement-free basis, whereas transit involves one interconnection party paying the other for the traffic it sends via the interconnection link. Peering partners therefore also avoid having to measure traffic for billing purposes and do not incur any costs for billing and collection, while this is an important cost element of transit agreements.

Whether interconnecting ISPs exchange data traffic via peering or transit also determines how the costs of realizing the physical network interconnection between the networks are shared. Generally, ISPs bear the cost of bringing their network to the Internet exchange point, irrespective of whether they are transit customers or peering partners. The costs of the link connecting the networks at the interconnection site are however shared only in the case of peering. In the case of transit, the transit-customer usually covers these costs (European Commission, 2000a: 9).

Important implications result for the **reachability** of an interconnection agreement depending on whether the interconnecting operators have agreed on peering or transit. Peering agreements allow the interconnecting parties the exchange of traffic originating and terminating with customers of one or the other of the peering parties. "Customers", as used in peering agreements, are retail and wholesale customers. The retail customers are the direct end-users of the peering partners. Wholesale customers are ISPs which have purchased transit from one of the peering partners (European Commission, 1998: 7-8). Peering excludes traffic originating and/or terminating in third party networks, that are peering partners of one of the interconnecting parties. Transit, as opposed to peering, allows the transit customer to pass all traffic on to the transit provider, regardless of the originating and terminating address of the traffic exchanged. The characteristics of peering and transit are summarized in Table 2.1.

Table 2.1. Characteristics of Peering and Transit Agreements

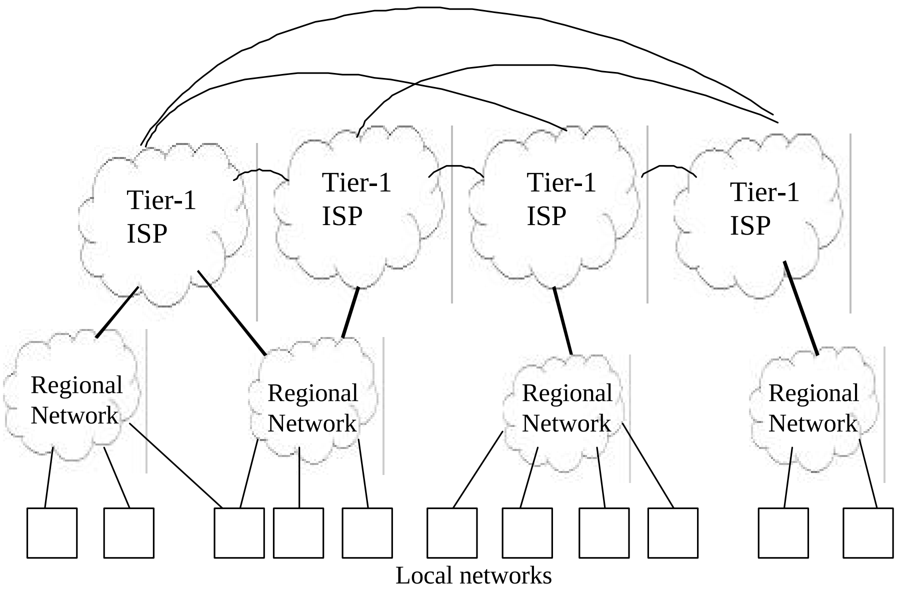
	Peering	Transit
Charges for traffic exchanged	Settlement free traffic exchange	Transit customer pays for interconnection
Cost of the interconnection link	Shared according to commercial negotiations	Generally financed by transit customer
Type of traffic accepted	Traffic originating with customers of one peering party and terminating with customers of the other peering party	Agreement extends also to traffic originating in third party networks and terminating in third party networks

2.4 The market for Backbone Interconnectivity

With the above background on Internet interconnection and Internet routing technology we can now define the market for Internet connectivity, as introduced in section 2.1. Ceteris paribus, interconnection agreements are the more valuable, the greater the Internet-coverage that they offer. The essential value traded on the Internet connectivity market is therefore the reachability of specific networks which make up the Internet.

ISPs wanting to offer universal connectivity to their customers have to combine a set of interconnection agreements which guarantee full coverage of the Internet. The easiest solution to this problem is to have at least one transit agreement which can be used as a default route for all traffic not covered by possible peering agreements. However, since transit is more costly than peering, ISPs will mix peering and transit agreements in order to attain the coverage they want. The logic of Internet interconnection has resulted in a hierarchical structure of the networks making up the Internet. ISPs often peer with ISPs of the same hierarchy level and arrange a transit agreement with an ISPs further up in the Internet hierarchy to have full coverage of the Internet. Only the carriers on top of the Internet hierarchy do not require a transit agreement in order to have full reachability of the Internet. They have peering agreements with all other ISPs on the highest Internet level and through these peering agreements they can reach any network in the Internet, as all other networks are either direct or indirect customers of the highest level ISPs. These ISPs on top of the Internet hierarchy are often referred to as **tier-1 ISPs** or **top-level ISPs**. The literature generally assumes the number of top-level ISPs to be below 10 carriers.

Fig. 2.4 Stylized depiction of the Internet hierarchy



Top-level ISPs are the only ISPs which can offer access to all Internet users. Any ISPs wanting full coverage of the Internet will therefore need to purchase transit services from at least one top-level ISP (directly of indirectly through intermediate ISPs). Furthermore, any ISP wanting to become part of the highest Internet level will have to have peering with all top-level ISPs. Otherwise he cannot offer universal connectivity solely on the basis of his peering agreements. It is because of this dependence of lower-level ISPs on the transit and peering services of the tier-1 ISPs that there is a discussion on the potential market power which results from the network effects of the Internet. Since universal connectivity is a quality characteristic which is important to most Internet users, ISPs are highly dependent on the interconnection services of the ISPs further up in the Internet hierarchy.

3 Market Power in Internet Service Provision?

The following overview of the literature on competition and market structure in the Internet will show that one can identify two issues which are being discussed in connection with the claim that top-level ISPs have market power. The first is the high market share tier-1 operators are said to have in backbone traffic and backbone capacity. The second is the need of lower-level ISPs to interconnect with at least one of the top-level backbone networks in order to achieve universal connectivity. This literature overview shoes that the discussion does not differentiate clearly between these issues, as though both can be traced back to the same source of market power. However, differentiating the markets for Internet service provision from the complementary markets of the Internet periphery, as was done in chapter 2, shows that the central issues are located on two separate markets. Backbone capacity is traded on the market for ***long-distance network capacity,*** whereas Internet interconnection agreements are negotiated on the ***backbone interconnectivity*** market. This paper argues that these issues need to be analyzed separately in order to arrive at non-ambiguous results regarding the potential for market power in the Internet.

3.1 Summing up the literature

Already in 1998 the OECD (1998) analyzed “Internet traffic exchange” on the grounds that smaller ISPs had voiced the concern that operators with a large share of “IP backbone traffic and connectivity” might be able to leverage higher rates for interconnection with their backbone network (p. 18). From the reproduced wording it is not clear, whether the smaller ISPs were concerned with a lack of competition for long-distance network capacity or with a lack of competition in backbone interconnectivity (or with both). In its analysis, the OECD defines backbone networks by their characteristic that they do “... not have to buy Internet access from any other company” (p. 17). It would therefore seem that the focus of the analysis is on the special features of tier-1 operators active on the backbone interconnectivity market. However, in its analysis of the potential for market power in Internet backbones, the OECD goes on to consider only the market for long-distance network capacity. It comes to the conclusion that “... the United States does not yet have cause for concern, because the earlier introduction of infrastructure competition has encouraged numerous backbone networks...” (p. 18). The mentioned infrastructure competition obviously has an affect only on the market for long-distance network capacity, but not on Internet backbone connectivity. It seems that the OECD (at least at that time) did not realize that the tier-1 networks active in the Internet connectivity market are not necessarily identical to the suppliers of long-distance network capacity.

Frieden (2001), in his own words, wants to “explore the nature of the Internet interconnection dispute”, which he characterizes as the allegation by some ISPs and policy makers, that transit payments to Tier-1 ISPs “violate international trade, antitrust, and economic development policies” (p. 159). From this description it would seem that the author is focusing on the backbone connectivity market. However, in the subsequent analysis, the long-distance network capacity becomes the central theme. Frieden weighs up the investments into long-haul/international bandwidth necessary to operate a backbone network, and maintains that "...a major backbone Internet operator does not appear overnight. These Tier-1 operators must have the financial and operational wherewithal to construct or lease and manage a nationwide network of high capacity lines. Few enterprises can amass the needed investment and skills" (p. 165). The focus on a "nationwide network" makes it clear, that Frieden is analyzing the services offered on the long-distance capacity market. What distinguishes tier-1 operators from other operators with extensive networks is not the reach of their physical network but rather the reach in terms of customers accessible via the network. From his observations on the long-distance capacity market, Frieden, unlike the OECD, comes to the conclusion that there is a potential for anti-competitive practices by Tier-1 operators. He urges that "Internet service provision requires collaboration, e.g. network interconnection, and cooperation, e.g. access by small ISPs to the backbone trunks of Tier-1 ISPs on fair, cost-based terms and conditions" (p. 166-167). Given the initial focus on Internet interconnection and the subsequent analysis of long-distance network capacity, it is not clear to the reader which market is finally being referred to in Frieden's claim for regulatory oversight of terms and conditions. Since Frieden does not differentiate between the motivation to interconnect with a top-level network in order to reach the customers attached to that network and the motivation to interconnect with a particular network in order to share its long-distance network capacity, it could be either one.

In its decision on the MCI and Worldcom merger of 1998 the European Commission (1998) focused especially on Internet interconnection issues, belonging to the market for backbone interconnectivity. The Commission found a separate market for "top-level or 'universal' Internet connectivity" (para. 70) and defined it by the characteristic that only the top-level ISPs active on this market can offer full coverage of all users connected to the Internet. This definition coincides with the above description of the Internet backbone connectivity market. However, in its assessment of the potential for market power among top-level networks offering universal connectivity, the Commission does not spend a lot of effort on an analysis of the strategies open to ISPs in order to attain universal connectivity. Rather, the commission focused on static data of revenues and traffic flow, without a motivation why these figures should reflect the market structure in the Internet connectivity market. Based on this data, the Commission determined the combined market share of the merging parties to be above 50% (para. 114). It is not clear in which relevant market the merging parties would have had 50%. Nevertheless, the Commission feared that the merged entity would be in a position to act independently of competitors and customers (para. 117) and therefore imposed additional requirements before authorization of the merger.

The Commission applied the same market definition in its assessment of the proposed merger between MCI Worldcom and Sprint in 2000 (European Commission, 2000a). Again the Commission concluded, on the basis of an assessment of market shares, that the proposed merger would either create or strengthen a position of dominance in the market for top-level Internet connectivity (para. 196). The merger was declared incompatible with the common market and the functioning of the EEA agreement.

Kesan and Shah (2001) consider Internet connectivity the central theme in the analysis of the potential for market power in the Internet. They see a shortcoming in the National Science Foundation's (NSF) privatization scheme for the Internet in the fact that the NSF did not impose any interconnection requirements on commercial ISPs. The authors hold this neglect accountable for a lack of competition in the Internet Backbone since smaller ISPs depend on interconnections with the top-level networks at conditions dictated by the latter (p. 148). That lower level ISPs are dependent on top-level ISPs is argued first with "the substantial role of nationwide backbone providers... the only ones who could carry traffic throughout the Internet" (p. 142). This seems to refer to long-distance network capacity, however, further into the analysis, the authors also state that "the enormous power of a few large backbone providers can be explained by 'network effects' " (p. 151). Overall, the analysis remains very vague on the regulations demanded for the Internet, such that this is also no indication as to where they see the source of the market power of tier-1 ISPs. They only generally find fault with the "lack of interconnection policy".

Crémer et al. (2000) in their own words, examine "the backbone market in the Internet" (p. 433). They identify connectivity as the central issue for an analysis of the competitiveness of the backbone market and consider the strategies available to ISPs wanting to circumvent an increase in the price for backbone service, or a decrease in the quality of that service. According to the authors, a strategy of secondary peering among lower-level ISPs can be of only limited influence on the backbone market, as the "setting up of new connectivity structure would involve a lag, since one must build bilateral interconnections and test new routing procedures" (p. 446). This argumentation misses the point that secondary peering can never be a complete substitute for transit with top-level ISPs because the customers connected to tier-1 networks will not be reached. From their discussion it becomes clear, that the authors focus the connectivity problem not on the universal connectivity service provided by tier-1 ISPs but have in mind a wider definition of the Internet backbone which also includes long-distance network capacity. Their market for "backbone" services therefore does not coincide with the market for backbone connectivity and the localization of the source of market power is blurred by the combined analysis of the two separable markets.

3.2 Preliminary Conclusions

The overview of the existing literature shows, that the markets for long-distance network capacity and for backbone interconnectivity have not been differentiated clearly in the discussion on the potential for market power in top-level Internet networks. Rather, the analyses often confuse the activities of tier-1 ISPs and of providers of long-distance network capacity. Because of this limited differentiation, the calls for regulation are necessarily vague. Should market power exist in either the market for long-distance network capacity or in the market for Internet backbone connectivity, very different regulatory remedies would need to be applied, depending on where the market power is located.

Market power in the market for long-distance network capacity can result if there are entry barriers into this market which keep potential competitors from entering and offering long-distance capacity at prices below the rates alleged to be above the competitive level. An analysis of the cost characteristics of the markets for Internet services would disclose if there are any entry barriers which support market power in the market for long-distance network capacity.

Market power in the market for backbone interconnectivity could be based on the network effects of the Internet, which, ceteris paribus, make those networks more valuable to customers which can offer the highest reachability of potential communication partners. An analysis of the network effects in the Internet and of the substitutes in Internet interconnection which can offer universal connectivity would therefore disclose if market power exists in the Internet backbone connectivity market.

Only a separate analysis of the cost characteristics of the Internet and the network effects in the Internet can come to unambiguous conclusions on the question of regulation of the Internet and appropriate regulatory remedies. Chapter 4 presents a framework for locating market power in the Internet based on the cost characteristics of the network. This chapter assumes that network effects do not matter. An analysis of the network effects in the Internet is not undertaken in this paper. It is the subject of related research by the author.

4 A framework for analyzing market power in Internet infrastructure services

4.1 Prologue on sector-specific regulation

It seems that at least some contributions to the discussion on the potential need to regulate the Interne are calling for government intervention without consideration to the costs of this alternative. In his seminal paper on information and efficiency Demsetz (1969) urges that public policy considerations should not be based on a nirvana approach, which entails comparing the existing institutional arrangement to an ideal norm, but rather on a comparative institution approach, which entails comparing the existing arrangement to other feasible institutional arrangements. Applied to the present analysis, a consideration of regulation in the market for Internet infrastructure services needs to take into account the incentive structure of a regulatory agency, the incentives which follow from regulatory instruments, the workings of the political process etc. before a decision to favor regulation above the market outcome can be justified.

A correct appraisal of the efficiency of government intervention is naturally difficult. Limited knowledge with respect to the consequences of government action often stands in the way of a detailed justification of regulatory activities. Economic theory, especially research in public choice theory does, however, allow the conclusion that government intervention is not likely to lead to an ideal outcome. Rather, the deviation from the ideal norm is likely to be significant in regulated markets. As a consequence, not every market failure can be drawn upon to justify government intervention into free markets. The non-intervention market outcome may well be closer to the ideal norm, than the intervention outcome. This is especially true for dynamic markets, in which conditions can change rapidly.

4.2 The disaggregated regulatory framework

The Internet is such a dynamic market, in which regulatory intervention could do a lot of harm, if imposed without sufficient justification. This section therefore introduces a normative framework for defining regulatory policies in network industries, which acknowledges the arguments of the positive theory of regulation and recommends regulation only in a limited class of cases.

The ***disaggregated regulatory approach*** (Knieps, 2000b) is specialized in the localization of market power in network industries. Network industries, are typical examples for industries characterized by large economies of scale and scope, and few active firms.

Because of deregulation in these industries, the localization of the remaining market power, which requires sector-specific regulation, is a topic of ongoing political debate. The disaggregated approach comprehends network industries as the sum of smaller vertically and horizontally integrated markets of which many are potentially competitive and only some may require regulation.

From the point of view of the disaggregated regulatory approach ex-ante regulation of a market demands more precise criteria for market power than those used in the ex-post analyses carried out in the context of general competition policy (Knieps, 2000b). The conjectural criteria of market power used in general competition policy are not considered sufficient for establishing that a firm has stable market power, which justifies ex-ante control over the market. This is so, because market characteristics, such as high market shares and great financial strength, are reconcilable also with situations in which other structural characteristics of a market, such as low barriers to entry, impede any attempt of large firms to act in a welfare reducing manner.[5]

The disaggregated regulatory framework introduces the concept of ***network-specific market power*** as a sufficient criterion which justifies ex-ante regulation,. This concept takes into account not only observable market characteristics, such as active competition, but also market characteristics which determine the potential for market entry by firms not yet active but willing to enter the market, should incumbent firms realize supra-normal profits over an extended time-period. Thus, the disaggregated regulatory approach places great emphasis on market entry conditions.

Stigler's concept of entry barriers is used to evaluate the possibility of market entry. Stigler defines: "A barrier to entry [...] as a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry" (Stigler, 1968: 67). This definition interprets only real asymmetric cost advantages of incumbent firms as barriers to entry, but not, for instance, the capital requirements which entrants face in any industry. From the point of view of the disaggregated regulatory approach, wider definitions of entry-barriers are not suited to localizing network-specific market power. Wider definitions treat mere hindrances to immediate market entry as entry barriers, while they do not deter market entry in the long run, when supra-competitive profits continue to be made in the industry over a substantial time period. For instance, the classical work on entry barriers by Bain includes product differentiation advantages of

[5] Empirical studies on the correlation between concentration and profitability have produced mixed results and in sum also do not support the hypothesis of a positive concentration-profitability relationship (Schmalensee, 1989: 973-977).

incumbents, such as advertising expenditures, among the barriers to entry (Schmalensee, 1989: 969). It can be argued, however, that as long as the costs of differentiating one's service or product are the same for both incumbents and newcomers, product differentiation activities can be undertaken by either firm, such that incumbents have no strategic advantage.[6] Along the same lines, even economies of scale and scope cannot be considered barriers to entry in the long run, as long as both the incumbent and the entrant have access to the same production technologies and therefore to the same cost functions (Knieps, 2002, 60). A natural monopolist enjoying economies of scale in the relevant output region still has to fear being replaced by potential entrants when the entrants can expect to produce on the same long-run cost function as the incumbent. The disaggregated regulatory approach therefore does not see a need to regulate a natural monopoly as long as the incumbent has no asymmetric cost advantage which secures lower production costs than those faced by potential entrants.[7]

That natural monopoly alone does not lend market power to an incumbent is also the central proposition of the Contestable Markets Theory by Baumol et. al (1988). The authors of the contestable markets theory present a theoretical case of a perfectly contestable market in order to expand the traditional reference scenario of perfect competition to an alternative and more general reference scenario, compatible also with natural monopolies (Mantzavinos, 1994: 56). For a market to be perfectly contestable, Baumol et al. put forward three conditions which must be fulfilled:

- Potential rivals have access to the same productive technology and input prices as the incumbent
- Market entry and exit are free, because entrants are not required to sink substantial investments
- The Bertrand-Nash expectations hold: potential entrants take the prices of the incumbent as given and believe that the incumbent will not instantaneously change prices in response to entry. Entrants evaluate potential profits of entry at a price below the incumbents price and expect customers to switch to the lowest priced product (market transparency is high).

While the ambition of the contestable markets theory was to formulate conditions for markets which guarantee the same stable efficiency properties as perfectly competitive markets, the disaggregated regulatory approach focuses on the opposite viewpoint. It seeks to

6 This theory is backed up also by empirical studies which do not support the hypothesis that advertising expenditures are a barrier to entry (Schmalensee, 1989: 981).

7 Von Weizsäcker points out that when economies of scale are present, very often the incumbent firm will own plant and equipment dedicated to the particular industry, such that the theoretical and empirical work on economies of scale as a barrier to entry may have misjudged the actual cause of the entry barrier which has to be seen in the combined existence of economies of scale and irreversible investments (v. Weizsäcker, 1980: 401).

define market characteristics which lend stable market power to incumbent firms. So while the contestable markets theory focuses on the formulation of precise market conditions which guarantee market entry into markets with above competitive profits, the disaggregated regulatory framework focuses on the formulation of market conditions which discourage market entry, even when above competitive profits are being made.

The disaggregated regulatory approach shares many of the assumptions of the contestable markets theory, but with a different focus. In the contestable markets theory, the assumption of Bertrand/Nash expectations plays the role of a restrictive condition which needs to be fulfilled in order for markets to be perfectly contestable. In the context of defining network specific market power, the assumption of Bertrand/Nash expectations, which ease market entry, support the conclusion that market power is stable. The argument goes, that if Bertrand/Nash expectations are given, but market entry still does not take place, then incumbent firms must have asymmetric cost advantages over potential rivals. The first two assumptions of the contestable markets theory list the strategic advantages which can hinder market entry even when Bertrand/Nash expectations hold: an incumbent has access to more productive technology or lower input prices and an incumbent has invested sunk costs in the industry.

Irreversible investments into production give the incumbent an asymmetric cost advantage which provides the potential for strategic action. They are no longer decision-relevant to the incumbent, but are considered by entrants contemplating market entry. The threat that the incumbent will lower its prices to the level of variable costs becomes credible because sunk costs no longer play a role.

From the viewpoint of the disaggregated regulatory approach, the relevant question to be answered when considering ex-ante regulation is therefore whether a natural monopoly (due to economies of scale and scope in the relevant range of output) is combined with substantial sunk costs. These characteristics are sufficient to hinder potential competition from effectively disciplining active firms. Network areas with these market characteristics are called “monopolistic bottlenecks . They lend network specific market power to the incumbent and therefore warrant ex-ante regulation.

Figure 4.1. The localization of monopolistic bottlenecks

Network area	**w/ sunk costs**	**w/o sunk costs**
natural monopoly	monopolistic bottleneck (1)	contestable market (2)
no natural monopoly	active competition (3)	active competition (4)

Source: Knieps, 2002: 4

The "theory of monopolistic bottlenecks", defines "a class of cases in which the localization of market power is based on the same reasons." (Knieps, 2002:1) Figure 4.1 illustrates this theory. Network areas are differentiated with respect to whether they exhibit natural monopoly characteristics (economies of scale and scope in the relevant range of output) and whether entry into the market would involve substantial sunk investment costs. Because economic theory predicts, that all network areas not exhibiting natural monopoly characteristics will be characterized by active competition, there is no need for market intervention here (quadrants 3 and 4). Those network areas showing natural monopoly characteristics, but not associated with substantial sunk investments, are considered contestable (quadrant 2). Ex-ante regulation is to be considered only in natural monopolies characterized by substantial sunk costs (quadrant 1). The theory of monopolistic bottlenecks does not suggest that contestable markets (quadrant 2) in reality show the efficiency properties of the textbook model by Baumol et al. (Knieps, 2000b: 97). Strategic action by active firms, such as product differentiation, price discrimination, advertising etc. will cause market frictions. However, these frictions are not considered sufficient to warrant ex-ante regulatory intervention into network industries as they are also found on competitive markets and do not cause stable market power over a substantial time period.

4.3 Disaggregated view of the market for Internet service provision

Applying the disaggregated regulatory framework to the question of market power in the market for Internet service provision requires analyzing the cost characteristics and entry conditions in the relevant network areas. This section therefore examines the network areas introduced in section 2.1 with regard to natural monopoly characteristics and investment requirements for entrants. The analysis is extended beyond the network areas of the Internet core to include the network infrastructure belonging to the Internet periphery markets. This is done especially with a view to the market for long-distance network capacity which has been one focus of the ongoing discussion. Table 4.1 summarizes the most important network elements and communications functions which are the principle drivers of costs in the network level and in the identified submarkets of Internet service provision.

Table 4.1. Disaggregated view of the market for Internet Service Provision

Submarkets of communications systems	Network Elements
Submarket 1 **The Network Level** **(= Local network access and long-distance network capacity)**	- building and maintaining transmission lines (requires access to public rights of way) - building and maintaining collocation space - upgrading of network capacity
Submarket 2 **Internet connectivity**	- the operation of Internet Exchange Points - the operation of routers - maintaining routing tables - the operation of DNS servers - the allocation of IP addresses to networks - software for: ▪ establishing and managing logical connections ▪ the fragmentation and reassembly of data messages ▪ monitoring network performance ▪ transmission error control and flow control
Submarket 3 **Internet Applications and Customer Relations**	- Internet access services - E-mail services - portal services - web hosting - customer support - billing and customer care

The **network level** (submarket 1) encompasses the physical infrastructure of the local network access and the long-distance network capacity. The infrastructure includes telecommunications transmission lines, coaxial cable, wireless media such as satellites and/or radio transmission facilities, collocation facilities, multiplexers etc. The costs associated with the network level are the costs of building and maintaining these physical assets as well as the costs of upgrading the network capacity according to network demand.

Level 2, **the Internet connectivity level**, comprises the network elements specialized in data transmission, such as routers, which are installed at network nodes and interconnection points between networks, and software for establishing and managing logical connections, for the fragmentation and reassembly of data messages, for transmission error control and for flow control. A further critical element of data transmission is a common address space. The hardware and software elements of layer 2 are therefore complemented by institutions such as the Internet Assigned Number Authority and its regional Internet registries, which administer the allocation of IP addresses and domain names. Carriers active on level 2 are required to maintain continuously updated routing tables to account for growth in Internet hosts and for changes associated with hosts moving to a new network or being taken off the Internet. They are further required to have a constant connection to Domain Name Servers where they can retrieve the translation of domain names into IP addresses.

The **Internet applications level** (Level 3) comprises the different applications ISPs offer to end-users of the Internet, such as E-mail services, file transfer, web-browsing, content hosting, portal services etc. Carriers offering these applications are responsible for product development, product design, and marketing, etc. To offer E-mail services and web-hosting and content, ISPs operate servers on which data can be stored. Furthermore, the direct relationship with the customer also encompasses functions in the area of customer support and billing activities.

4.3.1 Analyzing the cost-characteristics of the Internet applications level

Given access to the network elements of the network level and the transportation level, entry into Internet applications is similar to entry into competitive product markets of other industries. The costs of providing Internet application services such as E-mail services, portal services, web hosting etc. are driven primarily by comparable cost pools of traditional markets such as product development costs, marketing costs, and costs for customer care. These costs cannot be the source of network-specific market power, as they are no different in the market for Internet services than in any traditional product market.

Included among the Internet-specific costs of Internet application provision are the costs for operating servers required for web-hosting services, for E-mail services and for caching of remote web-sites.[8] Since these servers are scaleable, they cannot be the source of economies of scale in the relevant range of output and therefore also do not qualify as monopolistic bottlenecks. Further Internet-specific cost of Internet application provision are associated with the **billing** of electronically generated data records. The billing of such data records requires powerful rating engines. The fixed investments for these rating engines could be the source of economies of scale in the relevant output region, since smaller ISPs are not likely to generate enough traffic to exhaust the economies of scale of the most cost efficient billing technologies. However, there are companies specialized in offering billing solutions, to ISPs, telecommunications carriers, cable providers etc..[9] Smaller Internet application providers therefore do not need to invest into expensive billing equipment which is not reversible upon exiting the market. Rather, it is possible to outsource all billing functions such as the rating of call records, printing and shipping of invoices etc.. This decreases the investment risk for Internet application providers. In summary, because there are no network elements on the application level which show economies of scale in the relevant output region paired with irreversible investments, there are no monopolistic bottlenecks in the market for Internet application services.

4.3.2 Analyzing the cost-characteristics of the Internet connectivity level

Of the network elements on the Internet connectivity level, large investments may be associated particularly with the operation of Internet Exchanges, with the acquisition and operation of routers and investments into software. It is therefore worthwhile to analyze the cost characteristics of these network elements in more detail.[10]

Operation of Internet Exchanges[11]

Internet Exchange Points (IXPs) are shared interconnection points, where carriers meet for the purpose of exchanging Internet traffic.[12] The primary service provided by an Internet Exchange (IX) is a central switching platform and a pre-provisioned fiber mesh (often comparable to a local area network), to which carriers can connect their own equipment in order to realize interconnections with other carriers. Secondary services of IXs include help in establishing contacts between carriers, for instance by establishing mailing lists of carriers

[8] Caching refers to the temporary storage of remote content closer to the user. ISPs use the caching of web-sites to lower interconnection costs and to enhance the performance of their network (download speed).

[9] See, for instance, www.billcom.de or www.intec-telecom-systems.com/.

[10] The allocation of IP addresses from a common address pool is an important coordinating function within the Internet. While this function is not associated with substantial sunk costs, it is nevertheless sensible that only one central institution will be charged with this task.

[11] I am very grateful to Arnold Nipper, the technical manager of the DE-CIX, in Frankfurt Germany, for an interview which provided extensive background knowledge for this section on IXPs. The DE-CIX is an IX organized by the Association of the German Internet Economy (www.eco.de).

present at the IX. Generally, IXPs are located at commercial collocation sites. The costs of renting collocation are passed on to carriers as part of the fee for the IX service.

Carriers can gain efficiency from interconnecting at IXs, because interconnection at IXs minimizes the cost for building fiber to the location of peering points. Furthermore, an IX offers the opportunity for direct interconnection between networks when the amount of traffic between the two networks would not have warranted a bilateral interconnection. Smaller carriers have also used IXs to bundle traffic in order to negotiate a joined agreement for interconnection with larger carriers and thereby better the commercial terms of the contract (Hussain, 2002).

There are economies of scale in providing Internet-exchange services. Firstly, the cost of the shared infrastructure at the IX falls with the number of carriers supplied by it. Secondly, a smaller number of Internet Exchanges in a given area is more efficient because the benefit of membership at an IXP rises with the number of carriers interconnected there. It is difficult to assess when these economies of scale are exhausted. Easier to determine is the fact that providing Internet Exchange services does not require making substantial sunk investments. The collocation space is rented and therefore is not a fixed investment into IX service provision. The capital outlay on equipment used at IXs can be counted among the reversible investments, because a carrier can move equipment to a new IX location when leaving a particular IXP. Furthermore, the equipment installed at such central network nodes is generally high-end equipment, which can be reconfigured to fulfil other functions within the network should the carrier no longer need the services of an IXP. The only sunk costs of joining an IXP consist of the costs which arise from installing the equipment. Since IXs are generally organized by a cooperative of several ISPs, these sunk costs are shared among several carriers and are therefore relatively low compared to the investment into equipment. These easy entry condition into providing Internet Exchange services ensure that there is not market power associated with an IXP.

Routers

Routers perform different tasks depending on their place in the network hierarchy (see section on the technology of the Internet above). Investments into routers utilized in lower network levels are small compared to investments into the high-end equipment which is used to assume the role of a core router. The investments into the routers utilized on lower network levels must be considered part of the variable costs of production, as routers are replaced

[12] The European Internet Exchange Association has an informative website on Internet Exchange Points

regularly, according to advances in the technology. Furthermore, these routers are subject to capacity limits such that the economies of scale associated with a router are exhausted with a large enough customer base (Srinagesh, 1997: 124ff).

The core routers of Tier-1 networks necessarily have far greater capacity than the routers of lower network levels. Milgrom et al. (2000: 183) argue that the costs of core routers are likely to be subadditive and that from a cost-efficiency point of view a centrally managed backbone would be beneficial. They make the case that demand could be served by only one core router because each core router has to have a full routing table, such that one is always a perfect substitute for the other. Increasing the number of core routers would increase the cost of core routing without extra benefit. Furthermore, costs would rise not only because of the investments into the routers but also because the number of circuits connecting the routers would increase and the number of routing information exchanges would increase with the number of core routers.

What this argumentation does not consider is that with respect to technical items of such central importance as core routers networks are always redundant such that in the case of a malfunctioning of one element, another can take over. Therefore, it can be expected that even a single top-tier backbone operator would own and manage at least two if not more core routers. From a risk aversion perspective it is more efficient to have the redundant system be owned and operated by two independent backbone operators, thereby duplicating not only the hardware, but also the personnel knowledgeable with respect to core routing facilities.

Furthermore, the number of efficient core routers depends not only on the investment costs of core routing and the security aspects discussed above, but also on the quality of service which the routing system offers to connected customers. If a higher number of core routers can offer a better quality of service then a smaller number of these routers, then this may countervail the considerations which suggest that minimizing investment costs is efficient. Rather, the opportunity costs of a smaller number of routers in terms of bygone quality of service may tip the scale in favor of a higher number of core routers.

Quality of service in data transmission is generally measured by the parameters jitter (undesired visual or audio signals in data transmission), latency (the time it takes for a datagram to get from source to destination), throughput (the amount of datagrams a router can process in a given timer period) and packet loss (incomplete data transfer). Applications particularly sensitive to Quality of Service (QoS) characteristics, are real-time applications such as VOIP, video-gaming etc.. The QoS parameters are a function of the quality of the

(www.euro-ix.net).

circuits connecting network nodes (optical fiber, wireless, etc.) but also of the number of routers passed between the source and the destination of a message (because each router requires time to reevaluate the packet and possibly change entries in the header such as the hop count). With a single core router, whenever two autonomous systems that are not connected on a lower hierarchy level want to exchange datagrams, the traffic would be routed through the single core router in order to retrieve the necessary routing information stored in the full routing table of the core router. When the core router is geographically distant from the source and the destination network, then many hops will be required for the transmission, which will have a negative impact on the quality of the connection. The quality of the interconnection would be increased if a core router were closer to the origination and destination points.[13]

To sum up, there is a trade-off between the cost of core routing and the quality of service. Furthermore, there is a risk associated with only few core routers and few carriers capable of managing core routing systems. This seems to suggest that core routing is not a natural monopoly. It is more likely that the tradeoff will lead to an equilibrium with several core routers owned by several different carriers (as is currently the case).[14]

Whether core routers have the potential to sustain market power, depends not only on the number of core routers in the market, but also on the extent to which investments into core routing are sunk upon entering the market. According to the Yankee Group, routing platforms are currently being replaced every 18-24 months (Yankee Group, 2002: 4). The capital costs of routers as well as the installation and configuration costs can therefore be counted among the variable costs of Internet service provision. Variable costs factor in the calculation of both incumbent carriers and carriers contemplating to enter the market, such that there is no value to incumbency arising from previous investments into routing capabilities.

Furthermore, there exits a second-hand market for used routers, such that routers do not loose their worth upon exiting the market. Cisco, the leading manufacturer of core routers, for instance offers to buy back used routers for its Cisco Authorized Refurbished Equipment program.[15] Through this program new entrants to the market have the opportunity of buying certified and supported equipment without having to invest into the most expensive technology. In summary, there is no market power associated with routing facilities.

13 See also Huston (1999:6)

14 An aspect not considered here is the organizational advantage when core routers are owned by only one company. This issue will be considered in the section on firm subadditivity.

15 http://www.cisco.com/en/US/ordering/or6/or17/order_refurbished_equipment_high_level_listing.html

Software[16]

Router software is generally purchased in combination with a router and upgraded as regularly as the hardware. Furthermore, software is scaled to individual demand, such that a small ISP with a simple network has far lower software costs than a Tier-1 ISP which has core-routing functions in his network. Investments into software can therefore be counted among the variable costs of Internet service provision, such that they cannot be the source of substantial sunk costs. There are therefore also no large fixed investments into software which could be the source of economies of scale in the relevant output region.

Network management

Network management costs are likely to show increasing returns to scale as the cost of managing a network with twice the capacity will not double the costs of network management. Management costs do not count among the irreversible investments, however, because they are made up largely of personnel costs, which are included among variable costs of network operation.

Summary Internet connectivity level

None of the network elements of the transportation level show the characteristics of a monopolistic bottleneck. While there are economies of scale in network management, in Internet exchange services and in core routing capabilities, there are no substantial sunk costs associated with these network elements. The conclusion from the perspective of the disaggregated regulatory approach with respect to the transportation level is therefore that these elements do not sustain market power in Internet service provision.

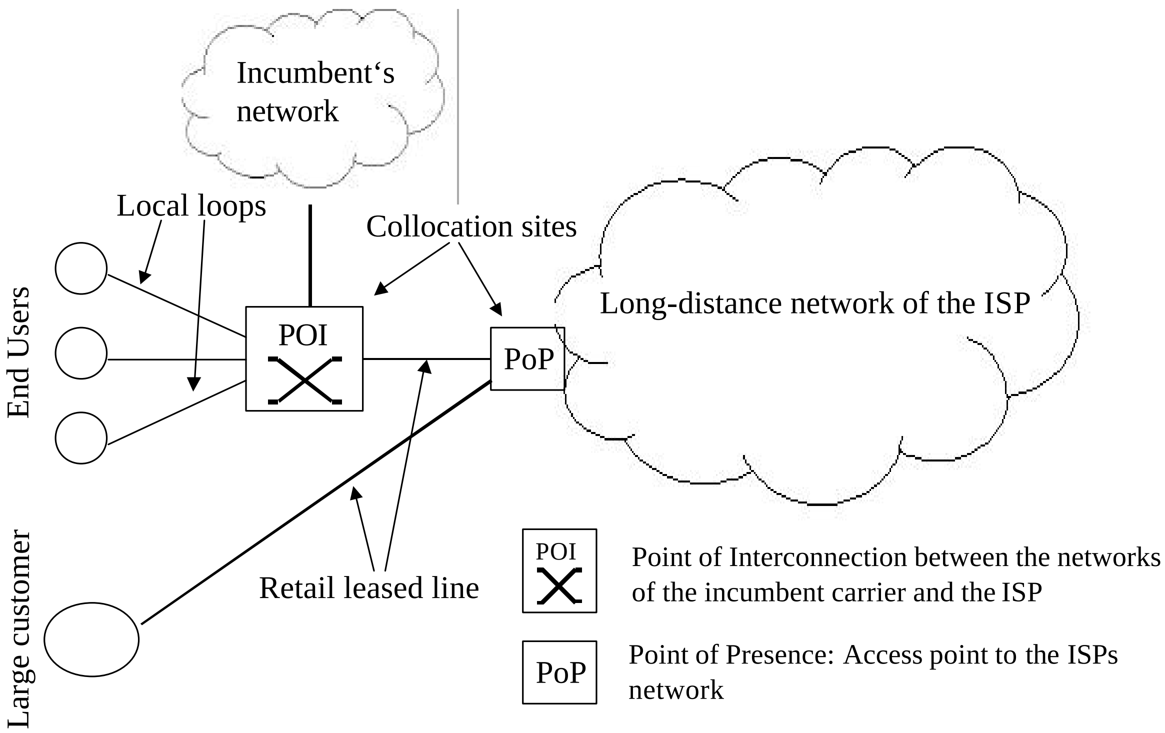
4.3.3 Analyzing the cost-characteristics of the network level

The elements of the network level which are used in Internet services belong to the Internet periphery. They are not Internet specific, because depending on the equipment installed at network nodes, the circuits of the network level can be used either for the transmission of voice signals or for the transmission of data packets. Therefore, the physical network lines and the collocation spaces at network nodes are an input to both voice telecommunications and Internet service provision. Although the demand for Internet services today represents a high portion of the demand for transmission capacity in the network level, there is no separate market for transmission lines for Internet services.

[16] See web pages of Cisco (www.cisco.com) for an example of Software for IP services.

The network level is made up of the local network access, the lines connecting end-user sites with the nearest Point of Presence (PoP) of a carriers network, and the long-distance network capacity, the links making up a transmission network and collocation spaces, in which transmission equipment is installed. The network level by far encompasses the largest investments necessary for Internet service provision. This is not because of high costs for optical fiber and copper lines, but rather because of the substantial expenditures for the first-time installation of these lines. And because these installation costs, such as digging trenches and laying cable, are sunk upon entering the market, the network layer also has a high potential of sustaining market power. Figure 3.3. illustrates the network level.

Figure 4.2 Stylized illustration of the network level



Local network access

When the end user utilizes the ***local loop*** of the incumbent telecommunications carrier to access the Internet, the actual usage of the ISPs network begins at a point of interconnection between the incumbent carrier and the ISP. The ISP will choose a POI located near a PoP of its own long-distance network and bridge the distance between the POI and the PoP, by employing a retail leased line. The ISP may connect larger end-users with a higher traffic volume directly to the nearest PoP of its long-distance network.

Due to the substantial installation costs which are divided among only few end-users, local lines are characterized by large economies of scale and scope and also by sunk investment costs. Traditionally, the local loop and local retail leased lines were therefore considered monopolistic bottlenecks for the provision of voice communication and Internet services. Recently, however, the view that the local lines are monopolistic bottlenecks has been put into new perspective by the trend of convergence between media platforms. While duplicating the existing local telecommunications infrastructure is indeed prohibitively expensive, when economies of scale are not exhausted, there already exist alternative local infrastructures, connecting a large number of households, which can potentially be upgraded to serve as substitutes for local telecommunications lines. These are for instance Cable TV lines (CATV) and electric powerlines. The redeployment of these alternative infrastructures has been especially successful for the purpose of providing Internet access services. A study by the OECD on the development of broadband Internet access in OECD countries, for instance, reveals, that among broadband alternatives, CATV access is often even more successful than broadband DSL services provided on the copper local loops of telecommunications incumbents (OECD, 2001).

In the face of new alternative technologies for local Internet access, the view that local loops are generally monopolistic bottlenecks can no longer be adhered to. However, it is also wrong to assume, that the market for local access is everywhere competitive. Instead, economists generally agree, that active or potential competition in the local loop is very selective (Woroch, 2002). The institutional framework (regulatory environment) and the historical circumstances in which alternative technologies have evolved, as well as factors such as the population density all influence the chances of intermodal competition in a particular area.

Consider, for example, the success of voice telephony and Internet access via cable modem in Great Britain. The positive development of the cable-modem technology in Great Britain is due to the fact that the British cable infrastructure was only deployed in the 1990’s. At this time legal entry barriers to competition in local telecommunication services had already been removed, such that cable companies had an incentive to build a technology capable of supporting two-way communication services (Höckels, 2001: 36). In Germany, in contrast, cable modem services have not been successful. To a large extent this is a result of the fragmented ownership structure of the German cable infrastructure (Vanberg, 2002: 13).

The installation of new circuit lines as an alternative to the incumbents retail leased lines has proved economical only in the market for higher bandwidth circuits (above 2 Mbit/s) and in metropolitan areas where traffic volumes are sufficiently high, such that more than one operator can operate at optimal scale (OFTEL, 2000:16). Therefore retail leased lines for the purpose of connecting individual customers also remain monopolistic bottlenecks whenever alternative access technologies are not viable.

For the purposes of the present analysis, it can be concluded that when economies of scale and scope are not exhausted, and when the upgrading of alternative technologies or the building of new access lines involves substantial sunk investments, the incumbent telecommunications carrier continues to have market power which justifies sector-specific regulation. Because technological advances are likely to further expand the realm in which competition is possible, the regulatory assessment of the local loop needs to be reevaluated regularly.

Long-distance network capacity

The installment of long-distance transmission lines involves large fixed investments, many of which are sunk, such as acquiring rights of way and labor costs of burying cable. Once a network is in place, the variable costs of operation are relatively small compared to the fixed costs of building a network. This constellation leads to substantial economies of scale. Because the communications lines can be used for more than one kind of service (for instance voice telephony and data transmission), long-haul transportation capacity is also characterized by economies of scope.

Through growing demand for capacity, triggered especially by the large volume of Internet traffic, the economies of scale and scope of long-haul transmission lines have, however, largely been exhausted. At the capacity limit, marginal cost of network use rises sharply as capacity expansion becomes necessary. Particularly on main routes connecting large city pairs, demand is large enough to support multiple lines.

Elixmann (2001) gives a comprehensive overview of the international market for long-haul transmission capacity. The study comprises data on the length of transmission lines, investment outlays for network capacity, stock exchange capitalization, turnover and locations of the most important international carriers for the year 2000. For the US market Elixmann identifies considerable market entry in long-distance lines after the divestiture of AT&T in 1984. Furthermore, extensive market entry occurred in the 1990's when carriers extrapolated the demand growth for capacity which resulted from the exponential growth rates in Internet usage. Counting the large telecommunications carriers AT&T, MCI, Worldcom and Sprint among the "incumbents", Elixmann considers additional 7 new carriers important players in the US market for long-haul data transmission: Broadwing, Global Crossing, Qwest, Teleglobe, Williams, Winstar and 360networks (p. 6).

For the European market, Elixmann identifies 22 important players in the market for long-distance capacity. The most significant of these being Cable & Wireless, Colt, Global Crossing, Interoute, KPNQwest, Level3 and MCIWorldcom (p. 20). Although demand for capacity continues to grow, the rate has decreased substantially compared to the peak rates in the mid 1990's.[17] Due to higher expected rates of growth, the market is now characterized by a substantial supply surplus (p. 49). The current situation in the market for long-haul capacity is characterized by price drops for rental prices in leased lines of varying bandwidth and varying length in almost all European countries (European Commission, 2000b: ANNEX 3 and European Commission, 2001: 44). These price developments can be taken as a sign of active competition in line segments connecting important commercial centers and highly frequented network segments. There is consequently no need for regulatory intervention in this market segment.

Entry into leased line provision in **rural areas**, with relatively little traffic volume, has however been rare (Elixmann, 2001: 54). These line segments are comparable to the retail leased lines in local access. Whenever traffic volume is low, economies of scale are not exhausted. And because building leased lines requires making investments which are not reversible upon exiting the market, these line segments can be considered monopolistic bottlenecks. As with the local access lines, technological advances can result in competition from alternative technologies. From the point of view of the disaggregated regulatory approach, regulation of these leased lines is justified whenever alternative technologies do not discipline the incumbent telecommunications carrier.

Co-location surfaces

Co-location surfaces are sites located at network nodes, where carriers install hardware equipment, such as switches and routers. These are also the sites where networks are interconnected. Because of the sensitivity of the communications equipment with regards to environmental factors such as temperature and humidity as well as the high security demands for such a site, collocation facilities afford large investments into fire protection, water protection, power supply, air conditioning, and security.

Generally, the large investment costs for a collocation site will cause substantial economies of scale, which only the largest carriers can be expected to exhaust fully. Furthermore, a large amount of the investment cost are sunk. Smaller carriers that require far less space for their equipment will therefore always be at an economical disadvantage. For smaller carriers it is therefore efficient to share collocation sites and distribute the overhead costs. Security solutions, such as lockable racks, can ensure that carrier equipment is protected in shared rooms as well as in shared cabinets.

The amount of space a carrier will demand at a given site is a function of the number of users and the amount of traffic volume which this site serves. The geographical location of the site is therefore important for the demand a carrier can expect to serve there. In cosmopolitan areas, where many carriers are in need of collocation space, there are often independent companies specialized in offering shared collocation space to carriers.[18] Some carriers have also specialized in offering collocation services for the purpose of realizing network interconnections.[19] Wherever such offers exist, a carrier can rent collocation space scaled to demand. Compared to the investment costs for building own collocation space, the set up costs involved with using rented collocation space are negligible.

However, the more rural the collocation site is, the lower will be the number of carriers active in the area and the more unlikely that an independent organization will offer scalable collocation space. In the lower network levels, it is likely that only the incumbent telecommunications carrier will have demand enough to justify investments into collocation facilities. Co-location space in lower network levels can therefore be a potential monopolistic bottleneck. In these instances sector-specific regulation is justified which requires the incumbent telecommunications carrier to offer collocation space on non-discriminatory terms.

Summary Network Level

In summary, the disaggregated analysis of the network level suggests that those network elements which are associated with providing long-haul data transportation services, the long-distance network capacity and the collocation facilities in higher network levels, are competitive. The network elements associated with offering access services, such as local loops and collocation space in lower network levels show different degrees of competition,

[17] The European Commission estimates that the market for data and leased lines grew at 3.8% in 2001 (European Commission, 2001: 3).

[18] Companies such as Interxion (www.interxion.com) and Equinix (www.equinix.com) for example build and operate Internet exchange centers in large cities worldwide.

[19] These are for instance Level(3) and MCI/Worldcom.

depending on the population density in a particular region and the possibility of upgrading alternative infrastructures for Internet service provision. It is argued that there remain monopolistic bottlenecks in local services, especially in rural areas. Sector-specific regulation is justified in these instances, in order to enable competition in horizontally and vertically related markets. Because regulation of incumbent telecommunications carriers already encompasses the network elements in the local network access which were isolated by this analysis, there is no need for a further strengthening of regulation from the point of view of Internet service provision.[20]

4.4 Summary of the disaggregated analysis of the Internet Services Market

Looking at the market for Internet service provision from the viewpoint of the disaggregated regulatory approach, this chapter comes to the conclusion, that potential monopolistic bottlenecks can be found only in local network access. Local access lines, local leased lines, and collocation on lower network levels are all associated with economies of scale in the relevant output region and substantial sunk investment costs. These network elements belong to the Internet periphery. Regulation of traditional telecommunications markets already pays regard to the functions that elements of the local access network take on in Internet service provision. Any demand for regulation of the Internet, that is based on the market power incumbent telecommunications carriers have in local network access, will therefore have to analyze existing regulations in the local access market and explain why these regulations are not considered appropriate or sufficient.

The disaggregated analysis of the market for Internet services did not disclose any regulatory requirements for long-distance network capacity or for core elements of Internet service provision. These network elements, which are at the center of the discussion on Internet regulation, do not show characteristics of monopolistic bottlenecks. From the analysis of the cost characteristics of Internet service provision, there is no need for ex-ante sector-specific regulation in long-distance network capacity or in Internet connectivity services. What remains to be analyzed with respect to the potential for market power in Internet backbones is the possibility of market power resulting from the network effects in the Internet connectivity market.

[20] Indeed, the analysis rather suggests that there is too much regulation in telecommunications markets, as regulation often still encompasses network elements of higher network levels.

References

Baumol, W., J. Panzar and R. Willig (1988), *Contestable Markets and the Theory of Industry Structure*, 2nd ed., New York, NY.

Comer, D. (1995), *Internetworking with TCP/IP Principles, Protocols and Architecture,* Englewood Cliffs, NJ.

Crémer, J., P. Rey and J. Tirole (1999), Connectivity in the Commercial Internet Market, mimeo.

David, P. (2001) The Beginnings and Prospective Ending of "End-to-End": An Evolutionary Perspective on the Internet's Architecture, mimeo.

Demsetz, H. (1969), Information and Efficiency: Another Viewpoint, *Journal of Law and Economics* 12, 1-22.

Elixmann, D. (2001), Der Markt für Übertragungskapazität in Nordamerika und Europa, *WIK Diskussionsbeiträge* 224, Bad Honnef.

European Commission (1998), Commission Decision of 8 July 1998 declaring a concentration to be compatible with the common market and the functioning of the EEA Agreement (Case IV/M.1069 – WorldCom/MCI), *Official Journal of the European Commission* L116, 1-35.

European Commission (2000a), Commission Decision of 28 June 2000 declaring a Concentration incompatible with the common market and the EEA Agreement (Case COMP/M.1741 – MCI WorldCom/Sprint).

European Commission (2000b), *Working Document on the Initial Results of the Leased Lines Sector Inquiry*, Brussels.

European Commission (2001), *Seventh Report on the Implementation of the Telecommunications Regulatory Package*, Brussels.

Frieden, R. (2001), The Potential for Scrutiny of Internet Peering Policies in Multilateral Forums, in: Compaine, B. and S. Greenstein (Eds.), *Communications Policy in Transition: The Internet and Beyond*, London.

Halsall F. (1996), *Data Communications, Computer Networks and Open Systems*, Harlow.

Höckels, A. (2001), Internationaler Vergleich der Wettbewerbsentwicklung im Local Loop, *WIK Diskussionsbeiträge* 228, Bad Honnef.

Hussain, F. (2002), *Network Conceptions*, Peering Business News, 1/1, 1-14.

Huston, G. (1999), *Interconnection, Peering and Settlements,* URL: http://www.isoc.org/inet99/proceedings/1e/1e_1.htm

Katz, M. and C. Shapiro (1985), Network Externalities, Competition, and Compatibility, *American Economic Review* 75, 424-440.

Kende, M. (2000), The Digital Handshake: Connecting Internet Backbones, *OPP Working Paper* 32, Washington D.C.

Kesan, J. and R. Shah (2001), Fool Us Once Shame on You - Fool Us Twice Shame on Us: What We Can Learn from the Privatizations of the Internet Backbone Network and the Domain Name System, *Washington University Law Quarterly* 79, 89-220

Knieps, G. (2000a), Der disaggregierte Regulierungsansatz der Netzökonomie. in: Knieps, Brunekreeft (Eds.): *Zwischen Regulierung und Wettbewerb: Netzsektoren in Deutschland*, Heidelberg.

Knieps, G. (2000b), Interconnection and Network Access, *Fordham International Law Journal* 23, 90-115.

Knieps, G. (2002), Does the System of Letter Conveyance Constitute a Bottleneck Resource?, *Institute of Transport Economics and Regional Policy Discussion Paper No.88*, University of Freiburg.

Knieps, G. (2003), Competition in Telecommunications and Internet Services: A Dynamic Perspective, in: Barfield, C. et al. (Eds.): *Internet, Economic Growth and Globalization: Perspectives on the New Economy in Europe, Japan and the USA*, Berlin.

Leiner, B., et al. (2000), *A Brief History of the Internet*, URL: http://www.isoc.org/internet/history/brief.shtml.

Mantzavinos, C. (1994), *Wettbewerbstheorie: Eine kritische Auseinandersetzung*, Berlin.

Milgrom, P., B. Mitchell and P. Srinagesh (2000), Competitive Effects of Internet Peering Policies, in: Vogelsang, I. and B. Compaine (Eds.), *The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy*, London.

OECD (1998), *Internet Traffic Exchange: Developments and Policy*, Paris.

OECD (2001), *The Development of Broadband Access in OECD Countries*, Paris.

OFTEL (2000), *National Leased Lines: Effective Competition Review and Policy Options*, London.

Schmalensee. R. (1989), Inter-Industry Studies of Structure and Performance, in: Schmalensee, R. and R. Willig (Eds.), *Handbook of Industrial Organization, Vol. II*, London.

Srinagesh, P. (1997), Internet Cost Structures and Interconnection, in: McKnight, L. and J. Bailey (Eds.), *Internet Economics*, Cambridge, MA.

Siegmund G., (1999), *Technik der Netze*, Heidelberg.

Stigler, G. (1968), *The Organization of Industry*, Homewood, Ill.

Vanberg, M. (2002), Competition in the German Broadband Access Market, *ZEW Discussion Paper* 02-80, Mannheim.

Von Weizsäcker, C. (1980), A Welfare Analysis of Barriers to Entry, *Bell Journal of Economics* 11, 399-420.

Woroch, G. (2002), Local Network Competition, in: Cave, M, S. Majumdar and I. Vogelsang (Eds.), *Handbook of Telecommunications Economics*, New York, NY.

Yankee Group (2002), Core Competence: New Requirements for Core Routing URL: http://www.juniper.net/products/features/core/core_competence.pdf.