Interconnection and Peering among Internet Service Providers

A Historical Perspective

An Interisle White Paper

Prepared by:
Interisle Consulting Group, LLC
39 S Russell Street
Boston, MA 02114
http://www.interisle.net
Lyman Chapin
+1 617 686 2527
lyman@interisle.net
Chris Owens
+1 617 413 3734
chris@interisle.net

About this Report

Studies of Internet Service Provider (ISP) interconnection arrangements have been performed from many different perspectives, including the technical architecture of exchange points, the business and economic models that underlie peering and transit agreements, and the interaction between market-driven interconnection arrangements and public policy (at both the national and international levels). These studies have been extensively reported and analyzed (see *Sources and References*). This report is intended to provide an historical context for, and concise summary of, the evolution of ISP interconnection—how it originated, how it developed, and how it is practiced today—without exhaustively reiterating information that is available from other sources.

The goal of this report is to describe the way in which the self-organized and self-regulating structures that govern today's global Internet—including the arrangements that enable ISPs to connect their networks to each other—have evolved naturally, over a period of roughly 35 years, according to principles that are deeply embedded in the Internet architecture. These structures are self-organized and self-regulating not because the Internet is an anachronistic "untamed and lawless wild west" environment, but because years of experience have shown that self-management is the most effective and efficient way to preserve and extend the uniquely valuable properties of the Internet.

Following an introduction to ISP interconnection in section 1, section 2 of this report describes the way in which today's model of ISP interconnection has evolved over the past 35 years in parallel with the evolution of the Internet architecture. Section 3 describes the economics and management structures that have emerged from that process to govern today's Internet. Section 4 identifies new challenges that have emerged since the turn of the century to the traditional arguments for and against the regulation of ISP interconnection. Section 5 summarizes the conclusions of the report. Section 6 contains a list of sources and references.

About the Author

Lyman Chapin has actively contributed to the development and evolution of the technologies and self-governance structures of what is now the Internet since 1977. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), and during almost three decades of international technical and diplomatic leadership has played a key role in the development of the network routing and interconnection architecture, protocols, and policy framework that support today's globally pervasive Internet. Mr. Chapin has served as the chairman of the Internet Architecture Board (IAB), the Special Interest Group on Data Communication of the Association for Computing Machinery (ACM SIGCOMM), and the U.S. and international standards committees responsible for network and transport layer architecture, service, and protocol standards. He was a principal architect of the Open Systems Interconnection (OSI) reference model and protocols, and is the co-author of *Open Systems Networking—TCP/IP and OSI*. Mr. Chapin has also served as a Director of the Internet Corporation for Assigned Names and Numbers (ICANN); standards area director for the Internet Engineering Steering Group (IESG); co-founder and trustee of the Internet Society (ISOC); U.S. representative to the networking panel of the NATO Science Committee; and U.S. representative to the computer communications technical committee of the International Federation for Information Processing (IFIP TC6).

Before co-founding Interisle Consulting Group, Mr. Chapin was Chief Scientist at BBN Technologies, which (as Bolt, Beranek and Newman) developed the hardware and software for the first ARPAnet nodes installed in 1969. Beginning in 1989, BBN operated the NSFnet regional network NEARnet, which became one of the first commercial Internet Service Providers (BBN Planet) in 1993, acquired two other regional networks (BARRnet and SURAnet), and was separately incorporated in 2000 as Genuity.

1 Introduction

Internet Service Providers (ISPs) connect their networks to each other in order to exchange traffic between their customers and the customers of other ISPs. *ISP Interconnection*[1] allows traffic originating at a source connected to one ISP's network to reach a destination connected to another ISP's network, around the block or around the world. End users see the seamless, global, ubiquitous communication medium known as the Internet; behind the scenes lie many individual networks, owned and operated by many different corporate, institutional, and governmental entities, joined to each other by interconnection arrangements. Interconnection is the glue that holds the Internet together.

Interconnection enables the Internet as a whole to be ubiquitously fully-connected, despite the fact that no single network operator could possibly provide Internet access in every part of the world. The unregulated market-driven model on which today's global interconnection arrangements are based has developed over the past three decades in parallel with the development of the Internet itself, and studies by a wide variety of public and private organizations[2] have repeatedly concluded that it represents the most effective and efficient way to provide ubiquitous public Internet connectivity without being either anti-competitive or inequitable.

Internet interconnection is fundamentally different from interconnection in the traditional, circuit-switched telephony world, for reasons that are intrinsic to the architecture of the Internet and how it has evolved. As a result, the nature of Internet interconnection agreements, the range of choices that are available to participants, the economics of interconnection, and the number and variety of

[1] In *The Evolution of the U.S. Internet Peering Ecosystem* (see *Sources and References*), Bill Norton coins the roughly equivalent term Internet Peering Ecosystem: "a community of loosely affiliated network operators that interact and interconnect their networks in various business relationships."

[2] An excellent summary of the case that these studies collectively make for the unnecessity of ISP interconnection regulation is contained in FCC Office of Plans and Policy (now Office of Strategic Planning and Policy Analysis) Working Paper 32, *The Digital Handshake: Connecting Internet Backbones* (see *Sources and References*).

participants in the market are different from their counterparts in the telephony world.

2 The Origins of Interconnection

As we observe it today, ISP interconnection is not an intrinsic technical feature of the Internet; it is a management feature necessitated by the fact that the ownership and administration of the physical components of the Internet infrastructure are distributed among many different commercial, non-commercial, and governmental organizations. Thus, there is an important distinction between *internetworking*, which enables networks based on different telecommunication technologies and protocols to exchange data, and *interconnection*, which enables the owners and operators of different networks to collaborate as business entities in the provision of seamless end-to-end Internet connectivity to all of their individual customers. Today internetworking, using the standard Internet Protocol[3], is the common operating mode throughout the Internet; interconnection takes place at specific public and private *exchange points*, at which two or more ISPs make technical and administrative arrangements to exchange traffic.

2.1 Networking

Neither internetworking nor interconnection were features of the Internet's most distant precursors. In the 1950s and 1960s, before LANs and PCs, "computer communication" meant connecting I/O and storage peripherals (such as card readers, terminals, and printers) to resolutely self-contained mainframe computers. Early efforts to connect computers to each other led to "networks" based on a variety of different proprietary communications technology and protocols. The Information Processing Techniques Office (IPTO) of the Advanced Research Projects Agency (ARPA) of the U.S. Department of Defense funded several projects to build homogeneous networks before Bob Taylor, who took

[3] The standard Internet Protocol is the "IP" in the familiar acronym "TCP/IP."

over as IPTO head in 1966, recruited Larry Roberts to design a "distributed communications network" which laid the foundation for the ARPAnet.

When there were just a few of these homogeneous networks, it was possible to exchange information between them by building a translator; but as the number of networks grew, the n-squared scaling inefficiency of pair-wise translation led to the idea of "internetworking"—creating a network of networks.

2.2 Internetworking

It is remarkable to realize that the very earliest thinking[4] about what a "network of networks"—an "internet"—should be embraced the three key concepts that underlie the architecture of today's global Internet:

1) The concept of *packet switching*, which originated in at least three distinct places during 1961-1965: in Paul Baran's work at the RAND corporation in Santa Monica, CA; in Leonard Kleinrock's work at UCLA in Los Angeles, CA; and in Donald Davies's work at the National Physical Laboratory in Teddington, UK. All three concluded that the strongest communication system would be a distributed network of computers with (a) redundant links; (b) no central control; (c) messages broken into equal-size packets; (d) variable routing of packets depending on the availability of links and nodes; and (e) automatic reconfiguration of routing tables after the loss of a link or node.

2) The concept of *best-effort service*, which originated in the multi-access channels of ALOHAnet at the University of Hawaii (by Abramson, Kuo, and Binder, through 1970)[5].

---

[4] In the early to mid-1960s, culminating in the July 1968 ARPA request for proposals for the interconnection of four ARPA research sites into what would be called the ARPAnet.

[5]ALOHAnet was a radio network, and the idea of contention for channels was widely familiar in the radio context; it was Bob Metcalfe's brilliant leap from ALOHA to Ethernet (at PARC in 1973) that brought the concept of stochastic (non-deterministic) channel access into the networking mainstream.

3) The concept of *application independence*—that the network should be adaptable to any purpose, whether foreseen or unforeseen, rather than tailored specifically for a single application (as the public switched telephone network had been purpose-built for the single application of analog voice communication).

At the outset, in 1969, the ARPAnet was not an "internet"—each of its four computer hosts was connected to an Interface Message Processor (IMP) by a proprietary serial link and protocol[6], and the IMPs communicated with each other over 56Kb/sec. lines leased from the telephone company, using an ARPAnet-specific "host-to-host protocol" that was referred to as the Network Control Program (NCP). Other packet networks, based on other protocols, were being developed at the same time[7]. The first papers describing "packet network interconnection" were published by Vint Cerf and Bob Kahn in 1973; the ARPAnet began using IP in 1977.

From the beginning the ARPAnet was managed by an informal and mostly self-selected group of engineers and managers who began meeting as the Network Working Group (NWG) in the summer of 1968. The tradition of self-management by the people designing, installing, and operating the network was established at the very first NWG meeting, and has carried through to the governance structures that oversee the Internet today—particularly the Internet Engineering Task Force (IETF).

2.3 Interconnection

The clearly evident usefulness of the ARPAnet to the U.S. Defense Department contractors who were permitted to use it led other U.S. Government agencies to

[6] Dubbed "1822," after the serial number of the BBN report that described it.

[7] Initially the Packet Radio Network (Bob Kahn) and Packet Satellite Network (Larry Roberts); later Cyclades (Louis Pouzin), and the X.25-based networks that became Telenet, Datapac, PSS, and Transpac.

develop similar networks[8]. Eventually, disgruntled computer scientists[9] who could not connect to one of the government-controlled networks established CSNET for the (academic and industrial) computer science community. AT&T's wide dissemination of the Unix operating system encouraged the creation of USENET, based on the Unix UUCP communication protocols, and in 1981 Ira Fuchs and Greydon Freeman developed BITNET, which linked academic mainframe computers.

With the exception of BITNET and USENET, these early networks were restricted to closed communities defined by an "acceptable use policy" (AUP) that specified the uses to which the networks could legitimately be put (e.g., to conduct research funded by a particular government agency). The prevalence of highly restrictive AUPs provided little incentive for the networks to interconnect, and initially they did not.

2.3.1 Federal Internet Exchanges

The practical awkwardness of operating multiple non-communicating networks eventually led to the establishment of two exchange points for federally-funded networks operated by NASA, DoE, ARPA, and NSF: the Federal Internet Exchanges at the University of Maryland (FIX-East) and NASA's Ames Research Center in Mountain View, CA (FIX-West). These interconnection points were managed by two informal groups of engineers and managers, the Federal Networking Council (for administrative matters) and the Federal Engineering Planning Group (for technical matters). The interconnection regime was designed primarily to isolate the regions within the emerging technologically uniform IP "internet" that were subject to different acceptable use policies.

---

[8] The U.S. Department of Energy (DoE) built MFENet for its researchers in Magnetic Fusion Energy; DoE's High Energy Physicists responded by building HEPNet. NASA Space Physicists followed with SPAN.

[9] Led by Rick Adrion, David Farber, and Larry Landweber.

2.3.2 CSNet and NSFnet

The USENET, BITNET, and commercial X.25 networks could not be connected to the ARPAnet (or to the other federal networks that were interconnected at the FIXes) because of the government policy limiting ARPAnet to government agencies and their contractors. The turning point that eventually brought them all together was the CSNET project, which was created in 1981[10] under a grant from the National Science Foundation . The purpose of CSNET was to link all of the computer science departments and industry labs engaged in computing research. It provided TCP/IP interfaces with USENET, BITNET, and the X.25 networks, and established nameserver databases to enable any computing researcher to locate any other.

The development of CSNet highlighted the disconnect between the “haves” and the “have nots” in the computing research community—between those who could find a government agency or contractor to sponsor their connection to the ARPAnet, and those who could not (connecting instead to CSNet). In modern terms, we would say that the customers of one ISP (ARPAnet) could not communicate with the customers of another ISP (CSNet), because no mechanism existed to reconcile the different Acceptable Use Policies of the two networks. This disconnect persisted as both sides assumed that any agreement to exchange traffic would necessarily involve the settlement of administrative, financial, contractual, and a host of other issues, the bureaucratic complexity of which daunted even the most fervent advocates of interconnection—until the CSNet managers came up with the idea that we now call “peering,” or interconnection without explicit accounting or settlement. A landmark agreement between NSF and ARPA allowed NSF grantees and affiliated industry research labs access to ARPAnet, as long as no commercial traffic flowed through ARPAnet. This agreement was the turning point at which the evolution of commercial network interconnection began.

[10] by Larry Landweber (University of Wisconsin), David Farber (University of Delaware), Anthony Hearn (Rand Corporation), and Peter Denning (Purdue University).

NSF went on to sponsor NSFnet, a high-speed backbone connecting its supercomputing research centers. NSF also commissioned the development[11] of a deliberate architecture of backbones and regional networks that introduced the idea of hierarchy into the Internet topology. By 1990, the NSFnet had become the backbone of the modern Internet, and in 1996, NSF handed over its management to commercial Internet Service Providers (ISPs).

2.3.3 Network Access Points

In 1993, as the National Science Foundation began the transition to private ownership and management of the NSFnet infrastructure, it established four geographically distributed, privately owned and operated Network Access Points (NAPs), operated by Sprint, Pacific Bell, Ameritech, and MFS. Under the terms established by the NFS, a NAP operator was required to provide and operate an interconnection facility on a nondiscriminatory basis, using published pricing and established technical operating specifications.

These NAPs were the first commercial Internet exchange points,where any interested party could co-locate equipment and connect its network to the NFSnet backbone or to other networks. <<CIX in San Diego was the first to engineer the interconnection at the IP layer, using routers.>>

As the original NAPs (also referred to as Metropolitan Area Exchanges, or MAEs) became increasingly congested, many network providers began creating their own private NAPs, which extended the commercial Internet exchange model yet further.

2.3.4 Commercial Internet Exchange Points

As the number and diversity of NAPs increased, the potential complexity of hundreds or thousands of ad-hoc bilateral arrangements pointed to the need for

[11] By Peter Ford, Bob Aiken, Hans-Werner Braun, and Steve Wolff.

an overarching, neutral policy framework within which providers could implement mutually beneficial cost-sharing interconnection agreements.

It was at this juncture that there began to emerge a large number of privately operated NAPs, a.k.a. "exchange points," which provided a uniform set of technical and administrative services (*e.g.* interconnection, traffic routing based on sophisticated criteria, operational support of routing equipment, traffic metering, billing, and clearing and settlement of charges between parties). These exchanges provided a framework that allowed multiple providers of different sizes, scopes, and operating philosophies, serving the same or different markets, to interconnect in ways appropriate to each.

2.3.5 Internet Service Providers

If exchange points, NAPs, and backbone providers were the wholesalers of the emerging commercial Internet, Internet Service Providers, or ISPs, were the retailers. ISPs served end users by providing connectivity between them and the rest of the Internet. End users connected to ISPs by placing calls over the public telephone network to modem banks operated by the ISPs, or via leased circuits of higher capacity. ISPs, in turn, connected to regional or backbone networks at NAPs or exchange points.

This interconnection hierarchy did not, however, correspond to a strict hierarchy of ISPs and backbone providers as business entities. Some providers were vertically integrated, operating in every business from high-capacity backbone traffic down to dial-up lines. Others specialized in providing one form or another of connectivity to one or more specific markets.

The economic incentives and tradeoffs that are so richly diverse in today's Internet (see section 3.2) began to develop as soon as commercial ISPs recognized that their interconnection arrangements could be a source of competitive advantage. Any ISP could connect to one of the public Internet exchange points, but the opportunity to achieve better performance, particularly for destinations that would be several "hops" away using a public exchange, led many ISPs to explore direct interconnection of their networks with those of other ISPs. The growing number of ISPs, and the variety of different ways in which the rapidly expanding Internet services market drove the development of creative combinations of public and private ISP interconnection, ensured that the Internet as a whole would always be fully interconnected; the customers of every ISP could communicate with the customers of every other ISP, whether or not any particular pair of ISPs installed an explicit public or private interconnection.[12]

2.3.6 Internet Exchange Points outside of North America

Because the Internet developed earlier, and more rapidly, in North America than in other parts of the world[13], the interconnection arrangements between North American ISPs and networks in other countries initially were biased strongly in favor of the North American ISPs. Until relatively recently, it was common for Internet users in Taiwan, for example, to communicate with other Internet users in Japan or Singapore over a path that led through an exchange point in California (MAE-West), with the Asian network operators paying the full cost of the trans-Pacific links. This imbalance arose both from the early absence of an exchange infrastructure in other parts of the world, and from the much more favorable (largely unregulated) economics of Internet telecommunications in North America than in most other countries, which meant that even where a link existed between, for example, Germany and France, the cost of connecting through an exchange point on the east coast of the U.S. could be an order of magnitude lower than the cost of a direct connection . Because North American Internet users were overwhelmingly the sources, rather than the consumers, of

[12] The topology of Internet interconnection has emerged over the past decade as an important factor in studies of Internet resilience and survivability; see, for example, Edward J. Malecki's "The Economic Geography of the Internet's Infrastructure" (*Sources and References*). A corollary to many of these studies is the observation that the self-healing properties of the Internet architecture guarantee that the Internet as a whole will remain fully interconnected even if most of the direct connections between individual ISPs were removed. The fear of Internet "balkanization" as a result of large ISPs refusing to interconnect with smaller ISPs is, in today's Internet, completely unfounded.

[13] With the notable exception of the UK, which was connected to the ARPAnet much earlier than any other non-North American country.

Internet content, they had very little incentive to defray the cost of connections to other countries.

As recently as five years ago, ISPs in non-North American countries were determined to correct this imbalance by forcing North American ISPs to subsidize the cost of inter-regional links. However, as dozens of viable regional Internet exchanges have emerged outside of North America[14], the pressure to regulate international ISP interconnection in favor of non-North American ISPs has substantially evaporated. Market forces now drive ISP interconnection decisions in many other countries as effectively as they do in North America.

3 Interconnection in Today's Internet

The fabric of today's Internet is stitched together from a huge variety of links and individual networks, ranging from individual home users' dial-up connections to globe-spanning networks of massive capacity, owned and operated by a literally uncountable array of providers: private and public, large and small, local and global, special-purpose and generalist. Interconnection is goverened by a wide variety of bilateral and multi-party arrangements.

ISP interconnection, from a technical, operational, administrative, financial, and legal perspective, involves a number of issues that go beyond simply splicing together traffic streams, all of which require cooperation and collaboration among multiple ISPs:

1) secure exchange of interdomain routing information;
2) the provision of services, particularly "quality of service" (QoS) dependent services, that semantically span multiple ISPs;
3) detection of and response to denial of service attacks (and possibly other forms of distributed, multi-ISP attack that have yet to be seen);

[14] A current list of Internet exchange points is maintained at https://www.peeringdb.com/private/exchange_list.php; at the time of this report, 60 of the 92 listed exchanges are located outside of North America.

4) control of spam, phishing, and other intrinsically multi-ISP exploits; and

5) enforcement of national public policy mandates (universal service, emergency warning (cf. the recent IETF proposal), wiretap, etc.).

ISP interconnection operates very differently in the Internet than its counterpart does, for example, in the more familiar public switched telephone network (PSTN). The differences are observable both in the basic architecture of interconnection—the decentralized and self-organizing "Internet approach" to packet switching vs. the centralized and heavily-managed PSTN circuit switching—and in the policies and economics that govern interconnection arrangements.

3.1 Interconnection Architecture

3.1.1 The Internet Approach

What can loosely be termed "the Internet approach" has become the dominant paradigm in networking. The Internet approach has displaced:

- other *technologies* (e.g. x.25, SNA);
- other *architectures* (e.g., point-to-point leased circuits; frame relay);
- other *business models* (e.g. monolithic, end-to-end ownership of the transport infrastructure by a single provider); and
- other *forms of governance* (e.g., regulated monopolies or government-owned PTTs).

What has given vitality to the Internet approach isn't simply that the current approach meets the needs of the current environment; it is that the current approach relies on underlying *processes* that are flexible, adaptive, and compelling. The Internet approach is driven by self-organizing, self-regulating groups that have proved, time and again, their ability to create and maintain technical, architectural, business, and governance policies and practices that encourage high-quality engineering, broad interoperability, and continued creation of value, while truly representing global consensus and thereby keeping participants on board.

Almost every aspect of Internet technical development, operation, and governance is managed by a self-organized, self-regulating structure, and this fact has often been cited as the key to the Internet's phenomenal success. Self-regulation has allowed the Internet to adapt quickly and efficiently to the rapid pace of change and innovation in telecommunications technology, operations, and public policy.

3.1.1.1 Technical Standards

Internet technical standards are developed through the activities of the Internet Engineering Task Force (IETF), coordinated by the Internet Architecture Board (IAB) and housed, administratively, within the Internet Society. The IETF is:

> "…a loosely self-organized group of people who contribute to the engineering and evolution of Internet technologies. It is the principal body engaged in the development of new Internet standard specifications. The IETF is unusual in that it exists as a collection of happenings, but is not a corporation and has no board of directors, no members, and no dues."[15]

The "loosely self-organized" IETF and related organizations have proven, over a 20 year history, to be effective at establishing workable standards and highly adaptive to the rapid growth and change that have occurred within the Internet.

3.1.1.2 Operating Principles and Practices

Since the earliest days of the Internet, the operators of interconnected networks have met both informally and formally to share technical information and coordinate operating principles and practices. In the 1990s, members of the former NFSNET "Regional-techs" meeting formed an expanded group, called the "North American Network Operators Group" (NANOG), with a charter to promote and coordinate the interconnection of networks within North America

[15] *Tao of the IETF—A Novice's Guide to the Internet Engineering Task Force.* (see *Sources and References*).

and to other continents, serving as an operational forum for the coordination and dissemination of technical information related to backbone and enterprise networking technologies and operational practices.

NANOG has been highly effective in allowing ISPs and backbone providers to coordinate their activities to efficiently provide seamless service to a broad market. The fact that North American Internet users enjoy transparent access to the entire Internet, regardless of the ISP to which they happen to be locally connected, testifies to the success of the self-regulating NANOG model.

Another measure of the effectiveness of NANOG is that other regions of the world have replicated the approach and have developed or are developing similar groups, including:

- AfNOG—the African Network Operators Group
- SwiNOG—the Swiss Network Operators Group
- JANOG—the JApan Network Operators Group
- FRnOG—the FRench Network Operators Group
- NZNOG—the New Zealand Network Operators Group
- SANOG—the South Asian Network Operators Group
- PACNOG—the Pacific Network Operators Group

3.1.1.3 Resource Allocation

One of the most important governance functions in any domain is promoting an efficient exchange of value and allocation of resources. In the Internet, there are two key types of resources in play: Physical, tangible infrastructure such as communications links and switching facilities, and virtual resources.

Domain names, such as "coca-cola.com", or "lightbulbs.com", constitute one highly visible class of valuable virtual resource in the Internet. Domain names combine aspects of traditional intellectual property (i.e. trademarks and service marks), with the technical infrastructure required to cause the names to perform their intended function.

Another important virtual resource is the ***IP Address***, the numerical address by which each computer connected to the Internet is uniquely addressable. Under any addressing scheme, there are only a fixed number of addresses available; the IETF can establish an addressing scheme (and has done so); the operators groups can establish a plan for deploying it; but there still needs to be a mechanism for allocating the addresses.

ICANN, the Internet Corporation for Assigned Names and Numbers, is an international, broadly participatory organization responsible for overseeing:

- The allocation of domain names, through a highly decentralized, market-driven process
- The allocation of IP addresses
- The operation of the mechanism (also highly decentralized) whereby names are resolved to addresses, an essential function for the proper operation of most Internet services.

From its own description:

> "As a private-public partnership, ICANN is dedicated to preserving the operational stability of the Internet; to promoting competition; to achieving broad representation of global Internet communities; and to developing policy appropriate to its mission through bottom-up, consensus-based processes."

3.1.2 Interconnection Arrangements

From a purely technical standpoint—that is, unencumbered by policy or economics—ISP interconnection is no more complicated (or controversial) than simple internetworking, in which routers connected by communication links of various kinds compute routes through the Internet based on information they have received from hosts (end users) on any networks to which they are directly connected and from other routers. In its simplest form, an Internet exchange point is a physical place (typically a room in a building) in which Internet routers are installed. ISPs that want to use the exchange point to connect to other ISPs run one or more links from their own routers to the exchange point, and connect them to the exchange point routers. The ISP routers and the exchange point routers exchange information about where different groups of Internet hosts—identified by their IP addresses—are located, using routing protocols such as the Border Gateway Protocol (BGP). ISP A might learn, for example, that a group of Internet users who are customers of ISP B can be reached through an exchange point to which both A and B are connected, and decide to use the exchange point to reach those users. Traffic from users on A's network to users on B's network would flow over A's network as far as the exchange point, and then over B's network.[16] A similar arrangement obtains when two ISPs decide to connect their networks directly to each other, rather than at a third-party exchange point.

The most important difference between this model of Internet interconnection and the circuit-switching model of the PSTN is that the Internet dynamically self-organizes to find paths from one point to another without explicit pre-configuration or setup. In the Internet, if an ISP's link to one exchange point (or the exchange point itself) fails, it can quickly re-route traffic through some other exchange point, or to a direct connection to another ISP, without loss of data or manual re-configuration. When multiple carriers are involved, this process is much less dynamic (and much less robust) in the PSTN, where call re-routing depends on the prior negotiation and provisioning not only of alternative circuits but also of switch ports and switching fabric capacity.

In today's richly-interconnected Internet, the possibility that an ISP could find itself unable to connect its customers to some part of the Internet because one or even many other ISPs refused to interconnect with it[17] is vanishingly small; there

---

[16] In practice, of course, the way in which traffic flows are managed at exchange points is much more complicated than in this example.

[17] Some ISPs will refuse to carry traffic that originated with another ISP that has been "blacklisted" for sponsoring spam or phishing attacks, but this is not the classic "holdup" scenario that can arise from simple refusal to interconnect in the PSTN world.

are simply too many available connection points, public and private, and the architecture of the Internet ensures that traffic will flow end-to-end regardless of where an ISP is connected.

3.2 Interconnection Policy and Economics

Interconnection policy refers to the way in which the technical and contractual arrangements that ISPs negotiate with each other to interconnect directly or at public or private peering points (Internet exchanges) are influenced by (a) the business objectives and policies of each of the parties, and (b) external mandates arising from laws, regulations, and other public policy instruments that apply to the jurisdiction in which the interconnection takes place. Interconnection economics refers to the way in which interconnecting ISPs assess and manipulate the economic variables that determine the viability of interconnection as a business proposition.

While it was once the case that networks were quite private about their peering policies, increasingly the market has become one in which networks publish their policies openly[18]. Aside from the publicly available policies providing a useful look into the economics of peering, the fact that they have become increasingly public speaks to an increasingly transparent, participatory market.

3.2.1 Interconnection Agreements

At its most basic, an interconnection agreement says “You carry some traffic for me, in return for which I’ll do something—either carry traffic for you, or pay you, or some combination of the two.” Interconnection agreements are often tailored very carefully and minutely to the specific circumstances of the parties involved, particularly when those parties are large ISPs.

---

[18] Representative examples of large, medium, and small networks’ peering policies are the MCI UUNet policy (at <http://global.mci.com/uunet/peering/>), the Speakeasy policy (at <http://www.speakeasy.net/network/peeringpolicy.php>), or (third example),

The current environment is one in which a heterogeneous mix of network providers—large and small; local, national, and global; public and private—connect to each others' networks under a variety of arrangements, which adhere to one of four basic models:

1) **Bilateral settlements.** Two operators interconnect. Each accepts traffic destined for its own customers and originating within the other's network. Neither network delivers traffic to third parties on behalf of the other. Each charges for the volume of traffic it accepts from the other. (It follows that if the value of traffic in both directions is equal, the net settlement amount would be zero)

2) **Sender Keep All.** As with bilateral settlements, two operators each accept traffic from the other, for delivery to the accepting network's customers. But no charge is made.

3) **Transit.** One operator, the provider, accepts traffic originating within the other's network, destined not only for its own customers but for third party networks with whom the provider in turn connects. The provider charges a fee for carrying the other network's traffic.

4) **Multilateral exchanges.** An operator connects to an *exchange*, a (usually commercial) facility carrying connections from multiple operators. There, traffic is routed to other operators' networks via equipment provided by the exchange and according to rules administered by the exchange; the operator settles through the exchange for traffic that others carry on its behalf and that it carries on behalf of others.

When all the different network interconnection arrangements are considered, it is possible to consider them all as variations on a common theme (see Table 1) :

- Networks "A" and "B" connect to each other, possibly through a third-party exchange point or other intermediary.
- Each accepts traffic destined for its own customers (peering), and/or for the customers of other networks to which it is in turn connected (transit).

- The arrangement either includes a cash payment made by one network to the other (again, possibly through a third party intermediary), or it doesn't.
- The arrangement is either purely bilateral, or it is a multi-party agreement.

<table>
<tr><th>"A" accepts traffic for:</th><th>"B" accepts traffic for:</th><th>Financial settlement</th><th>Networks connect:</th><th>Nature of Agreement</th></tr>
<tr><td rowspan="2">Its own customers only</td><td rowspan="2">Its own customers only</td><td>None</td><td rowspan="2">Directly</td><td rowspan="2">Bilateral</td></tr>
<tr><td rowspan="2">Cash</td></tr>
<tr><td rowspan="2">Other networks to whom it connects</td><td rowspan="2">Other networks to whom it connects</td><td rowspan="2">Through an Exchange</td><td rowspan="2">Multi-Party</td></tr>
<tr><td>Other</td></tr>
</table>

Table 1: Any given interconnection arrangement can be characterized by choosing one value from each of the columns above.

3.2.2 Micro-economics of Interconnection

Many economic and business-policy factors affect an individual ISP's decision to peer or not to peer with another ISP, and in the case of a paid arrangement (bilateral settlement or transit), the pricing[19] External, publicly-advertised factors include:

- Geographic coverage of the two networks: either overlapping, such that a peering relationship would be symmetrical; or non-overlapping, such that a peering relationship would extend each network's geographic reach.
- Technical factors: networks may require certain technical standards, or preferentially choose interconnection partners where the peering relationship gives access to a desired technology.

[19] Two recent studies of peering economics are reported in *Economics of Peering* and *A Business Case for Peering in 2004* (see *Sources and References*).

- Operational: networks may require a certain level of operational support.
- Routing: networks may require specific routing policies and practices.
- Size: networks may choose to peer only with similarly sized networks.
- Anticipated traffic volumes.

Additional, idiosyncratic factors apply. For example, if one network's specific geography or customer mix or traffic mix dovetails with an important element of the other network's strategy, it would lead to a higher perceived value and price than otherwise.

In any given case, the arrangement is made on the basis of a *perceived equitable exchange of value* between the two interconnecting parties, where the value of the arrangement to each of the parties is determined by a number of factors, some obvious (direct cash payment, cost-effective transit, or access to a large user community, for example); others entirely idiosyncratic .

Because so many idiosyncratic factors affect each interconnection decision, it is extremely difficult to analyze the economics of any particular interconnection arrangement using external, objective criteria in order to determine whether or not the market is distorted and the agreement gives either party undue advantage.

The argument has been made in the past, for example, that certain bilateral relationships between overseas and US-based networks are "unfair" on the grounds that the cost of the transatlantic or transpacific link was borne entirely by the overseas network, where as the origination of traffic was split more evenly between the two. In some cases, European ISPs in one country were connecting to US backbones in order to send traffic back to a neighboring European country, bearing the cost in effect of two transatlantic hops.

Two arguments against intervention apply here: one addresses the argument itself and the other examines historical outcomes. At a theoretical level, the implicit assumption that the cost of a link, in a perfectly fair market, should be borne by the two connected parties in proportion to the volume of traffic they originate, and that anything else is perforce distored; is flawed due to the additional factors other than traffic volume, discussed above, that influence the value of interconnection to either party. On a more pragmatic level, it has been observed that the European networks now connect to each other at multiple exchange points within Europe. It was the rational, financial desire to avoid paying transatlantic round-trips to connect to one's neighbor, and not regulatory intervention, that led to the emergence of this more effective network topology.

3.2.3 Macro-Economics of Interconnection

In addition to examining the factors influencing a single interconnection arrangement, it is worth examining the overall characteristics of the market in which these arrangements happen.

An essential characterization of a market is its *liquidity*. Does a buyer or a seller have a choice of many parties to deal with? Or is there a monopoly or oligopoly limiting choice?

In some sense, choice is intrinsic to the Internet's routed, connectionless architecture, as contrasted with the circuit-switched, connection-oriented public telephone networks. Just because ISP "X" can't strike a satisfactory bargain with ISP "Y", does not mean that X's customers will be unable to reach Y's customers: X always has the option of buying transit from some third party who is in turn connected to Y. (Y has a strong incentive not to make the terms of interconnection too onerous for at least some well-connected peers, at risk of cutting its customers off from regions of the Internet)

In addition to the intrinsic choice, the overall economic environment in which interconnection agreements are negotiated can be characterized as a free market with a large number of players. Models of interconnection economics have been developed by <<citations, including "Internet interconnection and the off-net-cost pricing principle": "The purpose of this article is to develop a framework for modeling the competition among interconnected Internet "backbone operators" or "networks.">>

It has been suggested that the free market operation of ISP interconnection would be threatened by consolidation and the emergence of a small number of dominant players, who would be able to form in effect a cartel and disadvantage their competitors, resulting in the usual effects of reduced competition: higher prices overall, a slower pace of innovation, fewer choices, and damage to the end consumer. There has, in fact, been considerable consolidation in the ISP market. Is it hurting the market?

In assessing this question, it is important to understand what might be the symptoms, or *signatures* of a distorted, uncompetitive, oligopolistic or monopolistic market. Given the idiosyncratic nature of peering decisions, it is not clear that just because the cash economics of given a peering arrangement do not track the data transport volumes, geographic footprints, or other obvious criteria, implies market distortion.

On the other hand, tracking the number of backbone operators and the entry barriers to the backbone business is likely to provide insight into the dynamics of the market. Recently, it has been claimed that the barriers to entry in the backbone business have been lowered, for example:

> "Trends in transport pricing over the past six months have created a disruptive change by lowering the barriers for small and regional networks to develop robust national backbones for application delivery, peering, network performance and business expansion. This presentation will review pricing trends and the opportunities that are being created for small and regional networks. It will draw upon specific examples and case studies of ISPs that have leveraged this trend, as well as a review of specific products and their price points. The presentation will be technical and geared toward an engineering audience."[20]

It is worth following these predictions to determine their accuracy and applicability.

[20] Jay Adelson, founder and CTO, Equinix, Session announcement, 2005 ISPCON conference.

4 New Challenges

The Internet approach to interconnection faces several new challenges, which will force it to adapt, as it has to other challenges and changes over the past decades. The first three challenges described below are well within the scope of the "Internet approach": the existing policy mechanisms are well equipped to adapt to these changes, as they have to equally disruptive challenges and changes in the past. The fourth, relating to external attempts to bypass the self-organizing aspects of the Internet approach and impose poicy, are in many senses orthogonal to the operation of the Internet itself, and represent a significant and potentially distorting force.

4.1 Multi-Layer Interconnection Arrangements

Today: interoperability and interconnection at the packet level (architecture—the Internet hourglass model). This was the focus of NRIC V Focus Group 4, for example. More recently, attention has been focused on interoperability at the application layer—VoIP in particular. This comes about in part because the PSTN, which has traditionally been the vehicle for global voice telecommunications, does not observe the same functional layering as do IP networks. In the PSTN, voice is both the application and the driver for the architecture of every other part of the system. In the Internet, voice is just another application, and at the IP layer a VoIP packet is indistinguishable from any other data packet. Because the architecture of the Internet is application-insensitive, and the architecture of the PSTN is highly application-sensitive, interoperability between the Internet and the PSTN is freighted with serious difficulties.

In 2000 the final report of NRIC V Focus Group 4 (see *Sources and References*), dealt with ISP interconnection at the level of packet exchange—how IP packets are conveyed across the boundary (physical and administrative) between different ISPs. In November 2003, the final report of NRIC VI Focus Group 3 dealt with VoIP: "The recommendations and best practices included in this report address the interoperability of Voice over Internet Protocol (VoIP) and the Public Switched Telephone Network (PSTN)."

With the rise of VoIP and QoS-dependent applications, interconnection arrangements are likely to involve multiple layers of the Internet architecture. This will affect technical standards, operating practices and policies, the economic decisions surrounding a provider's decision to interconnect, the terms of interconnection agreements, and the overall market.

4.2 Balkanization

Potential for balkanization of the Internet as backbone ISPs try to differentiate themselves (competitively) by offering services only to their own customers, resulting in a network infrastructure that does not provide a uniform, universal standard of coverage. (see OPP WP 32 pg. 26). This was anticipated by studies conducted in the late 1990s, which concluded at the time that balkanization was not likely to occur because of other forces.

4.3 Traffic-Load Sensitive Peering Agreements

Today: peering agreements are almost uniformly traffic-load insensitive. Possible emergence of a traffic-sensitive settlement system as ISPs in different situations try to deal economically with the asymmetry inherent in WWW.

4.4 Government Intervention

Government attempts to control various aspects of the Internet (ITU, national governments). As the Internet becomes more of a core enabler of human activities, it may start to look like a tool for the achievement of public policy objectives (for example, addressing the "Digital Divide" at a national or global level, or controlling trans-border data flows).

5 Conclusions

Today's Internet is the way it is because of the way it developed. In every arena: technical standards, operating practices, resource allocation, and others, policy is established by self-organized, inclusive organizations, operating with a high degree of transparency, and representing a broad constituency.

This approach is nearly inevitable, given the inherently decentralized native architecture of the Internet and the heterogeneous, global market in which the Internet operates. The incentives are well aligned: due to the network effect, continued growth of the Internet is a rising tide that lifts all boats, which creates a strong bias toward policies that facilitate growth and efficiency. If the policy making organizations didn't respond to that imperative, the participants wouldn't follow, and the policy makers would lose their mandate.

On the other hand, top-down attempts to regulate, either in the service of "improving" the Internet itself, to redress perceived inequalities in access or pricing, or in furtherance of orthogonal policy objectives (solving the “digital divide” problem, for example), no matter how well intentioned or carefully crafted, are contrary to the fundamental, decentralized nature of the Internet, which is an important source of the Internet's vitality, and run the risk of being destabilizing and harmful.

At present, the self-organized, self-regulating aspects of the Internet are thriving. Regulatory policy-makers should remain attuned to the possibility that future developments would lead to a less competitive environment, and watch for the signatures of a distorted market, but until such problems present themselves, should refrain from action.

6 Sources and References

The FCC and the Unregulation of the Internet. Jason Oxman, July 1999. FCC Office of Plans and Policy Working Paper No. 31. http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp31.pdf

The Digital Handshake: Connecting Internet Backbones. Michael Kende, September 2000. FCC Office of Plans and Policy Working Paper No. 32. http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp32.pdf

Bill and Keep at the Central Office As the Efficient Interconnection Regime. Patrick DeGraba, December 2000. FCC Office of Plans and Policy Working Paper No. 33. http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp33.pdf

A Competitively Neutral Approach to Network Interconnection. Jay M. Atkinson and Christopher C. Barnekov, December 2000. FCC Office of Plans and Policy Working Paper No. 34. http://www.fcc.gov/Bureaus/OPP/working_papers/oppwp34.pdf

Evolution of Internet Infrastructure in the Twenty-First Century: The Role of Private Interconnection Agreements. Rajiv Dewan, Marshall Friemer, and Pavan Gundepudi, University of Rochester, October 1999.

NRIC V Focus Group 4 Final Report, Appendix B: Service Provider Interconnection for Internet Protocol Best Effort Service. FCC Network Reliability and Interoperability Council, September 2000. <http://www.nric.org/pubs/nric5/2B4appendixb.doc>

NRIC VI Focus Group 3 Final Report. FCC Network Reliability and Interoperability Council, November 2003. <http://www.nric.org/pubs/nric5/2B3finalreport.pdf>

A Business Case for Peering in 2004. William B. Norton, September 2004. Presentation to "Gigabit Peering Forum IX," New York City, 14 September 2004. <https://ecc.equinix.com/peering/downloads/A Business Case for Peering in 2004.ppt>

Competitive Effects of Internet Peering Policies. Paul Milgrom, Bridger Mitchell, and Padmanabhan Srinagesh, 2000. Reprinted from The Internet Upheaval, Ingo Vogelsang and Benjamin Compaine (eds), Cambridge: MIT Press (2000): 175-195. <http://www.stanford.edu/~milgrom/publishedarticles/TPRC 1999.internet peering.pdf>

Economic Trends in Internet Exchanges. Bill Woodcock, February 2005. Presentation to NZNOG 2005 conference, 3 February 2005. <http://www.pch.net/resources/papers/asia-pac-ix-update/asia-pac-ix-update-v11.pdf>

Economics of Peering. Steve Gibbard, April 2005. <http://www.pch.net/documents/papers/Gibbard-peering-economics.pdf>

The Art of Peering—The Peering Playbook. William B. Norton, May 2002. <http://www.xchangepoint.net/info/wp20020625.pdf>

The ISP Survival Guide: Strategies for Running a Competitive ISP. Geoff Huston, Wiley Computer Publishing, 1998.

Internet interconnection and the off-net-cost pricing principle. Jean-Jacques Laffont, Scott Marcus, Patrick Rey, and Jean Tirole, RAND Journal of Economics Vol. 34, No. 2, Summer 2003 pp. 370–390.

Interconnection, Peering and Financial Settlements in the Internet. Geoff Huston, June 1999. Presentation at the 1999 Internet Society Conference (INET 99). <http://www.isoc.org/inet99/proceedings/1e/1e_1.htm >

Without Public Peer: The Potential Regulatory and Universal Service Consequences of Internet Balkanization. Rob Frieden, 1998. Virginia Journal of Law and Technology, Fall 1998. <http://vjolt.student.virginia.edu/graphics/vol3/vol3_art8.html>

The Evolution of the U.S. Internet Peering Ecosystem. William B. Norton, November 2003. < http://www.equinix.com/pdf/whitepapers/PeeringEcosystem.pdf>

Internet Economics. Lee W. McKnight and Joseph P. Bailey (eds.), MIT Press, 1997.

Netheads vs. Bellheads: Research into Emerging Policy Issues in the Development and Deployment of Internet Protocols. Timothy Denton, May 1999. <http://www.tmdenton.com/pub/bellheads.pdf>

Tao of the IETF—A Novice's Guide to the Internet Engineering Task Force. Susan Harris and Paul Hoffman, October 2004. <http://edu.ietf.org/tao>

The Economic Geography of the Internet's Infrastructure. Edward J. Malecki, October 2002. < http://www.clarku.edu/econgeography/2002_04.html>

