How Internet Traffic Moves Between Regions

- or -

Why IXPs Cannot Interconnect With Each Other

Version 1.0
November 2006

Bill Woodcock
Philip Smith
Barry Greene
Pindar Wong
Steve Feldman
Kurtis Lindqvist
Frederico Neves

Purpose

The purpose of this document is to explicate the method by which Internet traffic is passed between regions, and to address misconceptions occasionally arrived at about the "interconnection of IXPs."

Background: Transit and Peering

Internet exchange points, IXPs, are the points at which Internet service providers interconnect their networks, in order to produce the Internet traffic carriage service which they sell to customers. IXPs are thus the "factories" which produce the value of the Internet. IXPs are trivial to operate, but can only exist as a collaborative project between ISPs since, by definition, they exist at the point of adjacency between two or more ISP networks. Thus ISPs jointly operate IXPs, but their principle business is the carriage of traffic, or "backhauling" of the Internet traffic from the IXPs to their customers. Since the scope of any one ISP's operations are geographically limited, ISPs also buy wholesale "transit" service from other ISPs in order to receive traffic from regions outside of their geographic scope. Transit is the same service which they're selling retail to their own customers. There are only two types of relationship between parties in the Internet: peering, which is the symmetric cost-neutral connection between two ISPs at an IXP, and transit, which is the direction-specific paid connection between an ISP and a customer, whether that customer is another ISP or an end-user.

Peering occurs at an IXP or, said the opposite way and equally true, an IXP is created where peering occurs. Because an IXP is by definition a point, rather than a span or an area, peering does not involve carriage (transport) service. Transit, on the other hand, does involve the carriage of traffic between the point at which the ISP receives it and the customer's location. Because that carriage is the major contributor to Internet service costs, transit is a paid service. Because no external costs are involved in peering, and because both parties to a peering session create value as a consequence of it, peering is not a paid service, and could not be, because it doesn't have a "direction." To think about it a different way, in transit, an ISP provides a service to a customer, in the direction from ISP to customer, and the customer pays the ISP for that service, in the direction from customer to ISP. In peering, two ISPs are involved, neither provides a service to the other, and neither pays the other. These relationships are not mutually exclusive: peering and transit in both directions sometimes exist between just two parties.

If we diagram the Internet, we see that it always follows a hierarchy of one of two very similar forms:

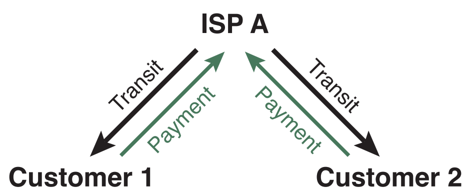

Figure 1 - Two customers interconnect through a single ISP

In the simplest but rarest case, diagrammed in Figure 1, customers 1 and 2 happen to coincidentally purchase transit from the same ISP A, rather than different ISPs, and their traffic is simply passed from one to the other through ISP A. The infrequency of this case is due to the fact that there are nearly four thousand ISPs operating at this level of the Internet topology, and more than a billion customers, so it's much more frequently the case that the customer of one ISP wishes to exchange traffic with the customers of the other 3,999 than with another customer of the same ISP. This is, in effect, a restatement of the "network effect."

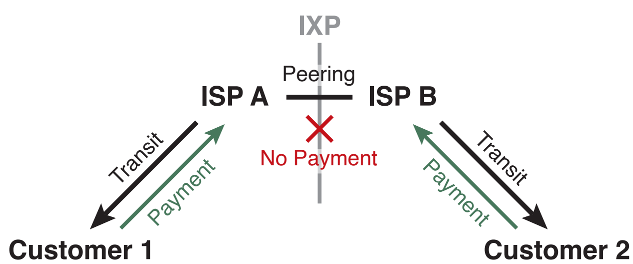

Figure 2 - Two customers interconnect through two ISPs

In the simplest expression of the usual case, diagrammed in Figure 2, customer 1 purchases transit from ISP A, while customer 2 purchases transit from ISP B. ISPs A and B peer with each other at an Internet exchange point to pass customer traffic back and forth.

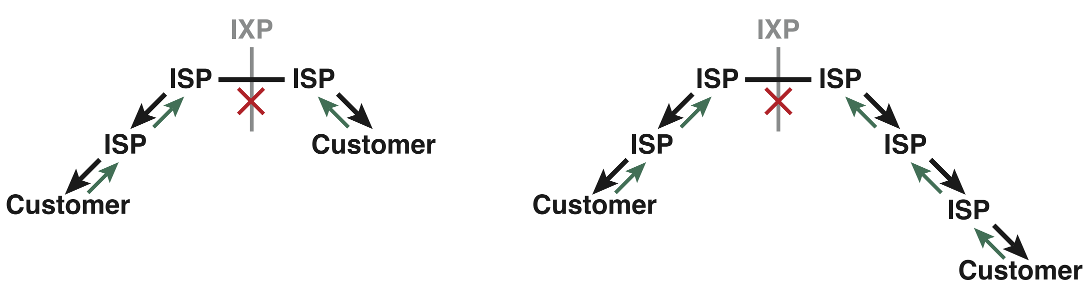

Figure 3 - All paths through the Internet follow the same topological model

In every case, paths through the Internet follow one of these two models: either "up" from one customer through one or more ISPs to an ISP at the top of the hierarchy, and then back "down" through a similar chain to a customer on the other side, or much more commonly, up from one customer through one or more ISPs, then through peering at an IXP to another ISP, and down to the customer on the other side. Figure 3 diagrams these more typical cases, which are embellishments of, but conceptually identical to, Figure 2.

It's important to note that regardless of the number of ISPs in the path, the inverse correlation between direction of transit (always "down," and never "up" or "across") and direction of payments (always "up," and never "down" or "across") remains necessarily uniform, since the service must be paid for, and the customer is the only party with an incentive to do so. Likewise, it is necessarily the case that money never flows across an IXP, because if money flowed, the connection would be, by definition, transit rather than peering, and the connection would be a directional customer-provider one (“up/down”), rather than an IXP (“across”).

Background: Quality Characteristics of the Internet

The quality characteristics of an Internet connection are loss, latency, jitter, and out-of-order delivery.

Loss is the portion of data packets which are lost in transmission; sent, but not received. Loss increases as a function of the number and reliability of components in the path, and the amount of contention for capacity.

Latency is the amount of time elapsed between when a packet is sent, and when it arrives at the recipient. Because of the limited speed of light, latency increases as a function of distance, but also as a function of the degree of utilization of transmission buffers by competing traffic sources.

Jitter is the degree of variability in loss and latency, which negatively affects the efficacy and efficiency of the encoding schemes which mitigate their effects. Jitter increases relative to the ratio of traffic burstiness to number of sources.

Out-of-order delivery is the portion of packets which arrive later than other, subsequently-transmitted packets. It increases as a function of the difference in queueing delay on parallel paths.

All of these properties become worse with time and distance, which is a reasonable definition of a perishable commodity. For this reason, keeping the connection between the sender and receiver of a packet as short as possible is very desirable. This can be generalized by saying that the speed, quality, and value of an Internet connection are inversely correlated with the distance from the IXP through which traffic passes. The quality of Internet service is increased when a customer is brought near an IXP (as in the case of a server colocation facility), or when an IXP is brought near a customer (as when a new local IXP is created). For real-world traffic, which is between two customers, quality is increased when the *average* distance between the customers and the IXP is decreased; that is, when the overall distance from the first customer to the IXP to the second customer decreases.

Background: Hot-Potato Routing

The Internet is a network of networks, the vast majority of which are commercial in nature, and which operate using standard economic rules; each network seeks to optimize its behavior for maximal self-interest. This is of particular interest because, as we’ve seen above, the work of an Internet service provider is to carry traffic between its customers and the IXP or IXPs to which it connects. The ISP’s self-interest is maximized by making the mean distance between each of its customers and the nearest IXP as short as possible, since distance multiplied by capacity defines the ISP’s cost, and the ISP has only indirect influence over the ultimate sources and destinations of its customers’ traffic. Fortunately, this aligns perfectly with the customer’s interest in keeping the distance between themselves and the IXP as short as possible, in order to maximize performance and minimize price.

All customers, however, are not close to the same IXP, since customers exist the world over, and each is near their own local IXP, of which more than four hundred exist, globally.[1]

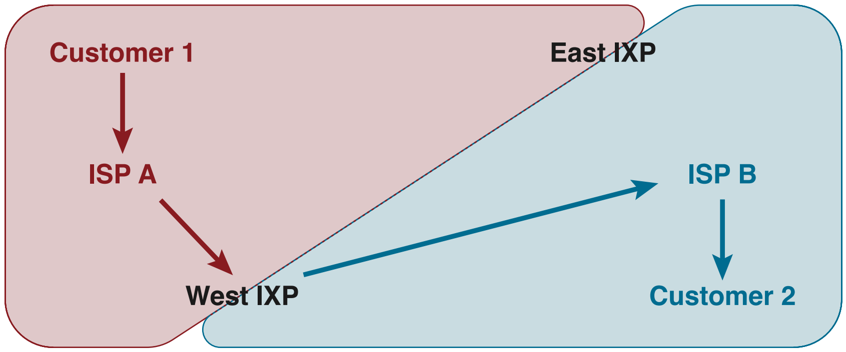

Figure 4 - Hot potato routing of traffic from Customer 1 to Customer 2

In the diagram above, Customer 1 exists within the service region of ISP A, and is a customer of ISP A (shown here in red). ISP A peers at West IXP and East IXP. The IXPs exist, by definition, at the adjacency between ISP A's network and its neighbors. In this case, Customer 1 sends a packet, perhaps a request for a web page, to Customer 2, who is within the region of, and a customer of, ISP B (shown in blue). ISPs A and B share common adjacencies at West IXP and East IXP, the former being nearer to Customer 1, and the latter being nearer to Customer 2.

Customer 1's packet is transmitted from Customer 1's network to ISP A. ISP A examines its routing table, determines that Customer 2 is a customer of ISP B and that the nearest adjacency with ISP B is at West IXP, and delivers the packet to ISP B there. Upon receiving the packet at West IXP, ISP B must haul the packet the long distance back to Customer 2, incurring substantial backhaul costs, which are recovered from Customer 2.

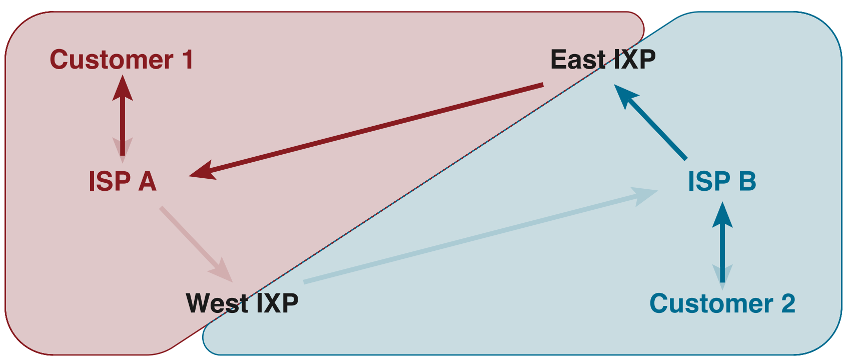

Figure 5 - Hot potato routing of reply demonstrates fair symmetric sharing of long-haul costs

Since Internet communications are normally bidirectional, Customer 2 will then reply to Customer 1, in this case with the content of the requested web page. They send the data to ISP B, which examines its routing tables, and determines that the destination is a customer of ISP A, and its nearest adjacency with ISP A is at East IXP, where it proceeds to deliver the packet to ISP A. ISP A must now undertake the cost of the long-haul of the packet back from East IXP to Customer 1.

What's fundamentally interesting about this transaction is that it's exactly symmetric. By each acting completely in its own self-interest, the two ISPs share the costs of providing service to their two customers exactly and fairly. This is, again, why there are no settlements in the modern communications network: each provider is paid by their own customer, and neither owes the other anything, because they have already divided the costs in a manner which they have both already determined to be fair. This beneficial economic consequence of hot potato routing is often referred to as "bill-and-keep" after the manner of handling money received from customers, and it represents a vast improvement over the prior century's very expensive and inefficient negotiated-settlement-based interconnection, which in any case did not ensure fairness. The prices which customers of the modern era pay for network services could no longer accommodate the waste-costs which were implicit in settlements.

A Summary of Beneficial Properties in the Network

Since IXPs are the means of Internet production, a larger total aggregate volume of IXP capacity means more Internet product, which is good.

Since a shorter distance between an IXP and a customer means that an ISP encounters lower transport costs and is able to deliver a higher-quality service at a lower price, having a large number of IXPs, widely-distributed, so as to exist in close proximity to the largest possible number of customers, is good.

The network effect applies in IXPs, as in the network as a whole, such that the greater the number of participating ISPs in any IXP, the greater the likelihood that an ISP will be able to find peering partners at the IXP with whom to exchange traffic, thereby increasing the efficiency with which they utilize the IXP.

The Misconception

If more, larger, and more populous exchanges are good, a naïve misconception is to believe that an even greater good could be created by somehow interconnecting geographically disparate exchanges to form a kind of hypothetical meta-exchange or "virtual exchange." This theory would have one believe it possible to simultaneously receive the benefits of widely-distributed exchanges and the benefits of highly-aggregated exchanges by interconnecting a number of widely-distributed exchanges together into one very large exchange which combined the populations of all of its constituent exchanges, yielding a greater network effect, while damaging neither the aggregate capacity nor the mean-distance-to-customers. Much akin to the attractive myth of perpetual motion machines, this fallacy has occurred independently in different regions during early phases of Internet development.[234]

The Reality

In fact, there is no free lunch. All long-distance backhaul has a cost, and that cost must be paid by someone. The economics of the modern network are mature, in as much as we have gradually arrived at the simplest, most efficient means of allocating that cost: it is equally divided between the two ISPs involved in the transaction, in a way which they have, by interconnecting at IXPs, already agreed is fair prior to the start of the transaction. Any step backwards from this endpoint of economic evolution would necessarily come at a severe cost in efficiency, trust, and goodwill, and literally vast monetary cost. Evolution goes in one direction for a reason, and throwbacks are eliminated through out-competition, just as their forebears were. Any more complex, inefficient, and expensive way of allocating those costs is not tolerated in the competitive marketplace. Interconnecting two IXPs, and incurring a backhaul cost which would then have to be allocated between parties via some less efficient method, is a self-correcting throwback. ISPs subjected to such a pricing regime depart, and form new, functional IXPs to replace those which have been destroyed by the attempted interconnection.

The Greater Cost

Unfortunately, the friction of the vast one-time expenses involved in building new facilities and moving many operating networks to them from the destroyed facilities is not the only detriment. A far greater long-term cost exists, and that is the ceiling imposed upon the growth of ISPs in any region where an IXP interconnection is attempted. In normal market conditions, ISPs compete with each other to deliver traffic in a balanced environment: each is assured that their major cost, that of backhauling traffic between IXPs and their customers, is approximately the same as that incurred by their competitors, and that the price customers will pay for service is approximately the same as well. This leaves the ISPs competing on quality of service, value-added features, and efficiency of operation, all worthy investments. If the basic engine of value-production, the long-distance backhaul of traffic from remote IXPs to one's customers, is removed by the imposition of an IXP interconnection which at once is not a direct competitor, but removes the ISP's ability to generate profit by providing this service, the ISP no longer has an economic reason to undertake the cost of connecting to remote IXPs. They can still operate at a small scale and reduced profitability, within the context of a single local market, but if extending their network to additional markets incurs backhaul costs without the ability to profit from them, they become less profitable than their smaller competitors in each individual market, and fail. This creates a regional ceiling in the vicinity of the inter-connected IXPs, in which no ISP can grow beyond a single locality, and all excess profits are extracted by larger out-of-region transit providers.

Conclusion

By definition, there is no possible model under which geographically-diverse Internet exchange points can directly interconnect, nor any reason for doing so. Internet exchange points are naturally interconnected by multiple Internet service providers, who compete to provide that service as a core function of their business. This is the only mechanism by which Internet service providers can grow from metro-area to regional or international-scale carriers, and if attempts are made to interconnect IXPs, they will not only directly cause the failure of the value-proposition of the involved IXPs, but also place a ceiling upon the growth of all of the ISPs in the region, preventing them from growing into the regional or international role which the successful among them would eventually fill in a competitive ecosystem. Any attempt to inter-connect IXPs would, by definition, take the form of a new incumbent monopoly, and would have to be permanently funded at public expense, since it precludes the possibility of competitive commercial ISPs or carriers becoming capable of assuming their natural role as profitable, competitive long-haul carriers. This precludes technological progress, Internet growth, and decrease in consumer prices, and would come at huge public expense. It offers no conceivable benefits, and the liabilities are legion. This is not speculation, this is established fact: there are more than 400 IXPs, and nearly 4,000 transit-providing ISPs in the world today. None have ever found a successful model for interconnecting IXPs to each other while all, the world over, participate in the existing proven model of ISPs interconnecting IXPs.

Authors' Contact Information

Bill Woodcock
Research Director
Packet Clearing House
woody@pch.net

Philip Smith
Consulting Engineer
Cisco Systems
pfs@cisco.com

Barry Greene
Consulting Engineer
Cisco Systems
bgreene@cisco.com

Pindar Wong
Technical Advisory Board
Packet Clearing House
pindar@hk.super.net

Steve Feldman

Network Architect
MAE-East & MAE-West
feldman@twincreeks.net

Kurtis Lindqvist
Managing Director
NetNod
kurtis@kurtis.pp.se

Frederico Neves
Director of Services and Technology
NIC.BR
fneves@registro.br

[1] More than four hundred IXPs exist globally: http://www.pch.net/ixpdir/.

[2] The proposal for a "ASEAN Regional Internet Exchange," was documented in 1999 at http://www.e-aseantf.org/pilot_prj/arix.html, now defunct, as are the links on the ASEAN web site, though some old press releases can be found using Google.

[3] The notion of a "Pan-African Virtual Internet Exchange" was described in 2002 at http://www.afrispa.org/Initiatives.htm.

[4] Interconnection of exchanges was mentioned as recently as June, 2005, as point 1.1 of eLAC2007, a Latin-American regional development plan: http://www.eclac.org/socinfo/noticias/documentosdetrabajo/5/21685/eLAC%202007%20English.pdf.