

October 2016

Geoff Huston

## The Death of Transit?

I was struck at a recent NANOG meeting just how few presentations looked at the ISP space and the issues relating to ISP operations and how many were looking at the data centre environment. If the topics that we use to talk to each other are any guide, then this is certainly an environment which appears to be dominated today by data centre design and the operation of content distribution networks. And it seems that the ISP function, and particularly the transit ISP function is waning. It's no longer a case of getting users to content, but getting content to users. Does this mean that the role of transit for users is now over? Let's look into this a little further..

The Internet is composed of many tens of thousands of separate networks, each playing a particular role. If the Internet's routing system is a good indicator, there are some 55,400 discrete networks (or "Autonomous Systems" to use the parlance of the routing system). Most of these networks sit at the periphery of the network, and don't visibly provide a packet transit service to any other network - there are some 47,700 such networks at the present time. The remaining 7,700 networks operate in a slightly different capacity. As well as announcing a set of addresses into the internet they also carry the announcements of other networks - in other words they appear to operate in the role of a "transit" provider.

Why has this distinction been important in the past? What was so special about "transit"?

A large part of the issue was tied up in inter-provider financial arrangements. In many activities where multiple providers are involved in the provision of a service to a customer it is common to see the provider who received the customer payment to then compensate the other providers for their inputs to this service provided to the customer. As logical as that sounds, the Internet did not readily admit such arrangements. There is no clear model of what a component "service" may be, nor was there any common agreement on how to perform the service accounting that would underpin any form of inter-provider financial settlement. On the whole, the Internet dispensed with such arrangements and adopted a simpler model based on just two forms of interconnection. In interconnection model, one provider positioned itself as a provider to the other, who became the customer, and the customer paid the provider for the service. In the other model, the two providers positioned themselves in a roughly equivalent position, in which case they could find a mutually acceptable outcome in being able to exchange traffic at no cost to each other (or "peering") (Figure 1).

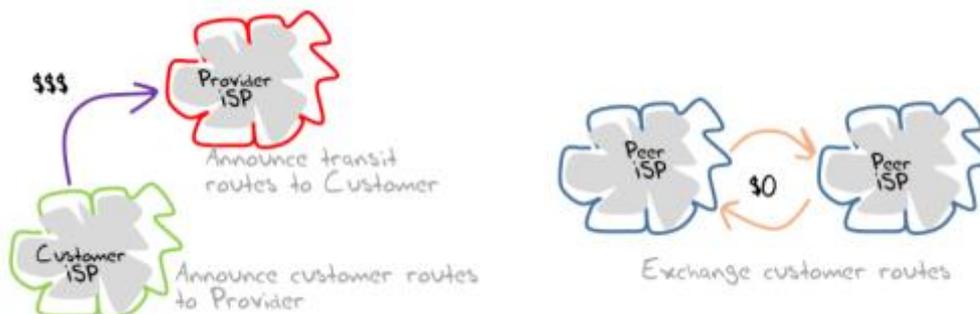


Figure 1 – Provider/Customer and Peering in the ISP world

In this environment sorting out who was the provider and who was the customer could be a difficult question in some cases. Equally, working out if peering with another provider was in your interest, or contrary to your interest was sometimes a challenging question. But there was one yardstick that appeared to be pretty much self-evident: Access providers paid transit providers to carry their traffic. The overall result of these arrangements was a set of “tiers” where players who co-resided in a tier normally found that any form of interconnection resulted in a peering arrangement, while if the interconnection spanned tiers, the lower tier network became the customer, and the higher tier entity was the provider. These tiers were not formally defined, nor was “admission” into any tier a pre-determined outcome. In many ways this layered model described the outcome of a negotiation process where each ISP was essentially defined into a tier by the ISPs with whom they peered, those who were their customers and those to whom they themselves were a customer.

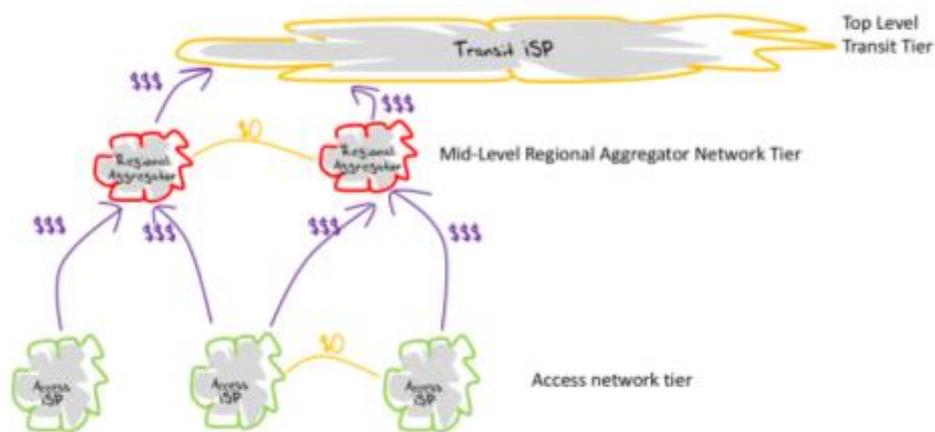


Figure 2 – Tiering in the ISP environment

It’s pretty clear to see where this headed, because at the top of this hierarchy sat the dominant providers of transit service. At the ‘top’ level were the so-called “Tier One” internet service providers. Collectively, these providers formed an informal oligarchy of the Internet.

To gain an insight as to why transit was such an important role, you only had to look at the more distant parts of the world where transit costs completely dominated the costs of Internet provision. In Australia a decade or so ago some 75% of all data that was passed to end users was hauled across (and back) the Pacific Ocean, with the Australian ISP carrying the cost of this transit function. The outcome was high retail costs for Internet access, the use of data caps and all kinds of efforts to perform local data caching.

Even in markets where the transit costs were not so dominant there was still a clear delineation of tiers where the combination of infrastructure assets, total direct and indirect customer numbers (aggregate market share, in other words) and negotiation ability determined one’s position relative to other ISPs in the same market space. But even in such markets transit mattered, as the consideration of infrastructure assets included the breadth of the assets in terms of span of coverage. A regional or local ISP would find it challenging to peer with an ISP that operated across a national footprint, for example.

But that was then and this is now, and it appears that much has changed in this model of the Internet. The original model of the Internet was carriage function, and its role was to carry client traffic to and from wherever servers might be located. To quote Senator Ted Stevens out of context, this carriage-oriented model Internet was indeed a “series of tubes.” Services populated the edge of the network and the network was uniformly capable of connecting users to services, wherever they may be located. But a distinction needs to be drawn between capability and performance. While the model allowed any user to connect to any service, the carriage path within the Internet differed markedly between users when accessing the same service. This often lead to anomalous performance where the further the user was

located from the content the slower the service. For example, the following slide comes from a Facebook presentation to NANOG 68 (Figure 3).

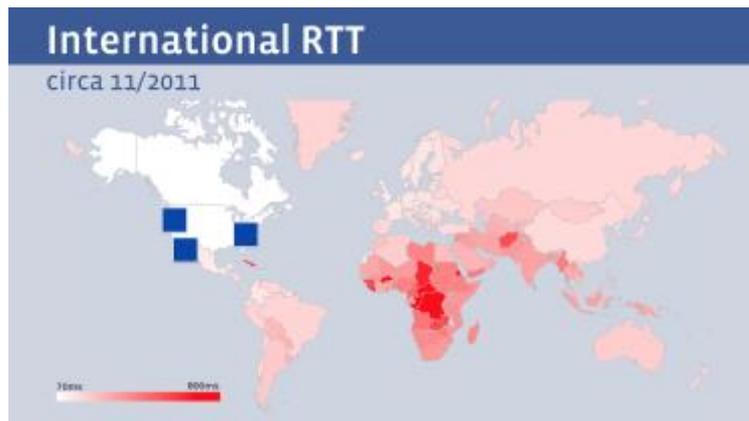


Figure 3 – Service Performance across the Internet c. 2011

Those areas served with geostationary satellite services and those with extremely long paths in terms of circuit miles obviously saw much larger round trip times (RTT) for data to be exchanged between the user and the service provision point. Of course this RTT delay is not a one-off hit. When you consider that there may be an extended DNS transaction, then a TCP handshake, potentially followed by a TLS handshake, then a request transaction and then the delivery of the served content with TCP slow start, then the average user experiences a transaction delay of a minimum of 3 RTTs, and more likely double that. A transaction that may take a user located in the same city as the server under a tenth of a second would take a distant user up to 6 seconds for precisely the same transaction. These delays are not always due to poor path selection by the Internet's routing system, or sparse underlying connectivity, though these considerations are a factor. Part of the problem is the simple physical realities of the size the Earth and the speed of light. It takes a third of a second to send a packet to a geostationary satellite and back to Earth again. The factor for satellite delay is the combination of the speed of light and the distance of the satellite above the earth to achieve a geostationary orbit. For undersea cable its better, but not by much. The speed of light in a fibre optic cable is two thirds of the speed of light, so a return trip across the Pacific Ocean can be as much as 160 milliseconds. So in some ways there is only so much you can achieve by trying to organise an optimal transit topology for the Internet, and after that there is no more to be achieved unless you are willing to wait upon the results of continental drift to pull us all together again! If you really want better performance for your service then clearly the next step is to bypass transit completely and move a copy of the content closer to the user.

And this is what is happening today. What we are seeing these days is that the undersea cable projects are now being constructed to link the data centres of the largest providers, rather than carry user traffic towards remote data centres. Google started taking ownership positions in undersea cables in 2008 and now has positions in six of these cables (as illustrated in a snippet of a recent cable announcement: Figure 4).

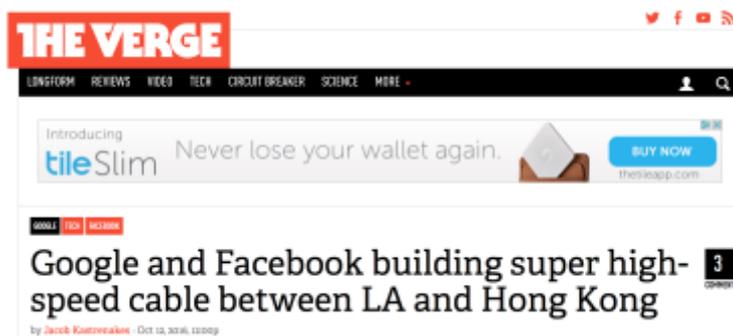


Figure 4 – CDN Cable Announcement

Tim Stronge, the vice president of TeleGeography, was reported in WIRED to claim that the new cable is the continuation of a current trend. "Large content providers have huge and often unpredictable traffic requirements, especially among their own data centers," he said. "Their capacity needs are at such a scale that it makes sense for them, on their biggest routes, to build rather than to buy. Owning subsea fibre pairs also gives them the flexibility to upgrade when they see fit, rather than being beholden to a third-party submarine cable operator." (<http://www.wired.co.uk/article/google-facebook-plcn-internet-cable>)

This shift in the manner of content access also changes the role of the various Internet exchanges. The original motivation of the exchange was to group together a set of comparable local access providers to allow them to directly peer with each other. The exchange circumvented paying the transit providers to perform this cross-connect, and as long as there was a sizeable proportion of local traffic exchange, then exchanges filled a real need in replacing a proportion of the transit provider's role with a local switching function. However, exchange operators quickly realised that if they also included data centre services at the exchange then the local exchange participants saw an increased volume of user traffic that could be provided at the exchange, and this increased the relative importance of the exchange.

There is another factor driving content and service into these dedicated content distribution systems, and that's the threat of malicious attack. Individual entities, particularly those entities whose core business is not the provision of online services, find it challenging to make the considerable investment in expertise and infrastructure to withstand such denial of service attacks. This desire to increase service resilience in the face of attack, coupled with the demands of users for rapid access, implies a demand to use some form of anycast service that distributes the service to many access points close to the various user populations. So it's little wonder that content distribution enterprises, such as Akamai and Cloudflare to name just two, are seeing such high levels of demand for their services.

So if content is shifting into data centres, and we are distributing these centres closer to users, where does that leave the ISP transit providers?

The outlook is not looking all that rosy, and while it may be an early call, it's likely that we've seen the last infrastructure projects of long distance cable being financed by the transit ISP carriers, for the moment at any rate. In many developed Internet consumer markets there is just no call for end users to access remote services at such massive volumes any more. Instead the content providers are moving their mainstream content far closer to the user, and it's the content distribution systems that are taking the leading role in funding further expansion in the Internet's long distance infrastructure. For the user that means higher performance and the potential for lower data costs, particularly when considering the carriage cost component. The demands for long distance transit infrastructure are still there, but now it's a case of the content providers driving the bulk of that demand, and the residual user-driven transit component is becoming an esoteric niche activity that is well outside the mainstream data service.

The overall architecture of the internet is also changing as a result. From the peer-to-peer any-to-any flat mesh of the Internet of the early 90's we've rebuilt the Internet into a largely client/server architecture, where much of the traffic is driven between clients and servers, and client-to-client traffic has largely disappeared. NATs helped that along by effectively concealing clients unless they initiated a connection to a server. It is probably solely due to NATs that the exhaustion of IPv4 addresses was not the disastrous calamity we all expected, By the time the supplies of new addresses were dwindling we had largely walked away from a peer-to-peer model of network connectivity, and the model of increasing reliance of NATs in the access network rewrote on the whole fully compatible with this client/server access model. By the time we turned to broad scale use of NATs the client/server network service model was firmly in place, and the two approaches complemented each other.

It's possible now that we are further breaking the client/server model into smaller pieces, by effectively corralling each client into a service "cone" defined by a collection of local data centres (Figure 5).

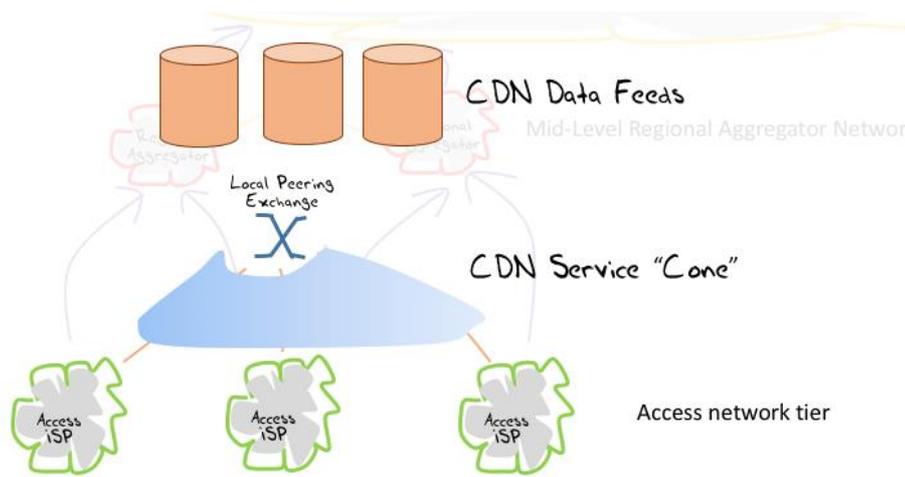


Figure 5 – CDN “Service Cones”

If the Internet service world truly is defined by a small set of CDN providers, including Google, Facebook, Amazon, Akamai, Microsoft, Apple, Netflix, Cloudflare and perhaps just a couple of others, and all other service providers essentially push their services into these large content clouds, then why exactly do we need to push client traffic to remote data centres? What exactly is the residual need for long distance transit? Do transit service providers even have a future?

Now it may be that all of this is very premature - to adapt Mark Twain’s quote, you could say that the reports of the death of the ISP transit provider are greatly exaggerated, and you would probably be right today. But one should bear in mind that there is no regulatory fiat that preserves a fully connected Internet, and no regulatory impost that mandates any particular service model in the Internet. In the world of private investment and private activity it’s not unusual to see providers solving the problems they chose to solve, and to solve them in ways that work to their benefit and at the same time ignore what is a liability in their eyes, and given that time and energy is finite, to ignore what is of no commercial benefit to them. So while “universal service” and “universal connectivity” may have meaning in the regulated old world of the telephone system (and in the explicit social contract that was the underpinning of that system), such concepts of social obligations are not imperatives in the Internet. If users don’t value it to the extent that they are willing to pay for it, then service providers have limited incentive to provide it!

If transit is a little desired and little used service, then are we able to envisage an Internet architecture that eschews transit altogether? It’s certainly feasible to see the internet evolve in to a set of customer “cones” that hang off local data distribution points, and we may choose to use entirely different mechanisms to coordinate the movement or synchronisation of data between the centres. Rather than a single Internet we would see a distributed framework that would look largely like the client / service environment we have today, but with some level of implicit segmentation between each of these service “clouds”. Admittedly that’s taking the current situation and running a pretty extensive degree of extrapolation to get to this picture, but even so, it’s not beyond the realms of possibility.

In this case, then in such a collection of service cones is there any residual value in a globally consistent address space? If all the traffic flows are constrained to sit within each service cone, then the residual addressing requirement is at best a requirement to uniquely distinguish end points just within the scope of the cone without any particular requirement for network-wide uniqueness (whatever “network-wide” might mean in such a system). If anyone still remembers the IETF work on RSIP (Realm-Specific IP, RFC 3102) then this is probably what I’m referring to with such a thought.

But that is still a future that may be some years hence, if at all.

Today, what we are seeing are further changes in the continuing tension between carriage and content. We are seeing the waning use of a model that invests predominately in carriage, such that the user is “transported” to the door of the content bunker. Instead, we are using a model that pushes a copy of the content towards the user, bypassing much of the previous carriage function. In other words, right now it appears that content is in the ascendant and the carriage folk, particularly the area of long haul transit carriage, are seeing a marked decline in the perceived value of their services.

But this model also raises some interesting questions about the coherence of the Internet. Not all services are the same and not all content is delivered in the same way. A CDN may well make assumptions about the default language to use to present its servers, or honour regional licensing of content. For example, Netflix uses different catalogues in different parts of the world. It’s just not quite the same service everywhere. At a more generic level we are seeing some degree of segmentation, or fragmentation, in the architecture of the Internet as a result of service delivery specialization. In such an environment that is segmented into almost self-contained service cones what is the market driver that would see all possible services loaded into every service distribution point? Who would pay for such over-provisioning? And if no one is prepared to underwrite this model of “universal” content, then to what extent will “the Internet” differ depending on where you are? Will all domain names be consistently available and resolve consistently in this segmented environment? Obviously, such considerations of customising the local service that is delivered in each local region heads down a path of Internet Fragmentation pretty quickly, so perhaps I’ll stop at this point and leave these as open questions that may be a side-effect of the possible demise of transit and the rise of content networking.

How long this model will last is anyone’s guess. It may well be that this is just a phase we’re going through, and the current data centre-centric view is just where we are today, rather than where we might end up. So maybe it’s not quite time to completely give up on the traditional concepts of universal connectivity and encompassing transit services just yet, even though it seems a bit like today’s Internet is taking a rather large step in a different direction!

---

## Author

*Geoff Huston* B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. He has been closely involved with the development of the Internet for many years, particularly within Australia, where he was responsible for building the Internet within the Australian academic and research sector in the early 1990's. He is author of a number of Internet-related books, and was a member of the Internet Architecture Board from 1999 until 2005, and served on the Board of Trustees of the Internet Society from 1992 until 2001 and chaired a number of IETF Working Groups. He has worked as an Internet researcher, as an ISP systems architect and a network operator at various times.

*[www.potaroo.net](http://www.potaroo.net)*

---

## Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre.