

October 2020
Geoff Huston

Going Postal

When a service is constructed using diverse components, then the way in which service revenues are distributed to the various suppliers of the components of the service can follow a number of quite distinct models. There are various forms of revenue redistribution models where the revenue per transaction is distributed to the various suppliers according to their inputs to support each transaction. And there are various “bill and keep” models where a set of service providers use each other’s services essentially donate their service to each other. This allows each service provider to bill their customers for the entire service, and rely on a “knock for knock” model where the foregone revenue for the donated services matches the opportunity revenue gained in the efficiencies not having to redistribute the fees paid by the customer for the compound service.

At one point in the telephone saga one large telco acknowledged that it cost the company some 21c to bill their customers 25c per local phone call. The shift to flat fees without usage tariffs is not only attractive to many customers, but also attractive to many operators as a means of stripping out huge costs from their billing systems as well as eliminating one of the major points of friction with their customer base.

The Internet is a good example of this compound service, where each Internet carriage provider provides services to both their customers and to the customers of other carriage providers. If you look at the Internet as a packet transit network, then many individual providers play a role in passing an IP packet from a sender to a receiver.

Contrary to the utopian mantra of the late 80’s the Internet is certainly not free, and ultimately there is only one source of revenue: users like you and me. We each pay our local access providers some form of service tariff and this revenue is then notionally distributed across the network in a manner that is intended, to varying levels of accuracy, to compensate each contributing service provider according to the costs that they incur in providing the service. We each pay our access service providers, and they in turn pay various transit and exchange providers, or carriage providers, for their services. But this raises a perennial topic for the Internet. How do the various carriage service providers pay each other? Who sets these inter-provider financial settlement rates? On what basis are these rates set? How fair are these arrangements? Much has been written on this topic of settlements and peering in the Internet, and doubtless much more will be written. As we continue to explore this topic it’s useful to look at other communications systems and look at how they coped with much the same problem.

For many years it felt that International Telecommunications Union (ITU) had been cast as the bad boy by the United States, due largely to the structure of international call accounting financial settlements and the perception that these arrangements placed US carriers, particularly AT&T, at a massive disadvantage. For much of the latter part of the twentieth century US enterprises saw in the ITU an institution that was hostile to US interests in that it was apparently determined to extract money from US telephone operators and divide up the spoils in a completely unfair and potentially corrupt manner. Even though the US economy is significant, the voting model of the ITU with one country one vote meant that the extent of

US both in corporate and government terms could effect some redress of their grievances with this situation was very limited. For this reason, among others of course, the US took a firm position to in the 1990's withhold the Internet from the same international treaty organisational regulatory structure, and this position was staunchly maintained by the US for many years. This is still largely the case today, although aspects of direct US oversight of the coordinating infrastructure of the Internet are no longer as evident as they were in the past. There is no UN agency lurking behind ICANN, or the Regional Internet Registries. There are no inter-provider arrangements described in some regulatory recommendations from the ITU or any other such recognised international body. And when we come to the issue of interconnection between Internet carriage providers there is no imposed common framework that governs any such interconnections. Peering and settlements in the Internet are a market, much like any other, and it's up to each service provider to negotiate a position with other providers of their choice that they can live with. Yes, there are winners and there are losers, as with any such framework, and the winners are generally happier with the current situation than those who perceive that they are net losers here. However, after some thirty years of working within this system we seem to have become accustomed to this system, complete with its perceived benefits and inequities. Or, at a minimum, the US interests who felt so aggrieved by the telco international call accountment settlement frameworks now feel that they are not a disadvantaged party in Internet's arrangements.

But it's useful to remind ourselves that it could've been so very different. There were a number of alternate models that were examined along the way, and the ITU model of half circuits, call accounting and call accounting financial settlements was certainly one that was supported by many, particularly those with a background in the telco sector. The fact that we couldn't find a "call" or even a "caller", and the fact that the retail tariffs were based on access rather than individual service transactions did not seem to bother these proponents. The market-based system we're using seems to reward the early and the largest adopters and everyone else has to pay. Considering the care that the telco sector had taken to try and balance the inter-provider interconnection space to ensure that no single national provider could leverage their position to the detriment of others, including half circuit funding frameworks and originating call minute accounting with agreed transfer rates, this seems like a regressive step on the part of the Internet.

However, the telco interconnection model was not the only show in town. There was another body in the international treaty space that did not appear to engender the opprobrium of the United States, or at least not until quite recently! And that body is the Universal Postal Union, or UPU. Here I'd like to look at the UPU's track record of trying to cope with a similar settlement and peering problem and contrast their various responses to those used within the Internet.

The Advent of the Penny Post



While sending letters to each other is about as old as civilization itself, it remained a pastime of the ruling class and the privileged elites for millennia. The writing and reading of written messages required education, and the movement of the messages required the services of a specialised courier service. The industrial revolution of the nineteenth century saw many changes in society, including the dislocation of large numbers of people moving from the country to the new industrial cities looking for better working conditions. The mass movement was not only a movement within countries, but emigration was also a major feature of this time. Schooling became more commonplace and literacy rose. There was a greater level of demand for an available and affordable mail delivery service that would replace the rather haphazard process of mail couriers who were paid upon delivery of a letter.

Although a postal service had been available in Britain for hundreds of years, the Royal Mail took two revolutionary steps in the nineteenth century that would make it the leader in postal innovation at the time. In 1840, Roland Hill, an enthusiastic campaigner for postal reform, led the Royal Mail's introduction

of the Penny Post, a cheap single rate for domestic delivery of letters, replacing a complex distance and letter volume tariff. This step made the mail service accessible to a much larger population. At the same time, it was intended to result in a large-scale uptake of the use of private correspondence, allowing the mail delivery service to realise economics of scale and process. The second revolutionary step was a switch from the recipient pays to a sender pays tariff model. This step removed a troubling imposition on the mail service, caused by mail refusal. Under the recipient-funded service model, the refusal of a mail item meant that the postal service incurred the cost of delivery, as well as the cost of return, assuming that the refused mail was ever returned.

These reforms had a profound impact. They generated huge volumes of postal traffic, facilitated the development of commerce and underpinned significant changes in social communication. This measure coincided with the Industrial Revolution, when people were moving from rural areas to the industrial cities and from one country to another, so these postal reforms eased to, some extent, the social pressures that arose from this large-scale transformation in nineteenth century society.



The pressures for postal reform rapidly spread, and there was a debate at the time for what was called “Ocean Penny Postage”. This was largely driven in the United States, where there was a significant immigrant population who wanted to maintain ties with their family back home. The postage rates for trans-oceanic postage were extremely high and recipient-pays refusal rate was high enough to be a significant imposition for the postal carriers. Complicating this was the issue of a complex overlay of courier services to carry a mail item. Each stage of the delivery of the letter would incur a

further change by that courier, sometimes using a different tariff model with a different weight classification.

As the immigration levels increased this system became unworkable and the first international reform in the 1850’s was to shift away from the individual handling of each letter and to use standard weight mail bags as the element of exchange between national mail couriers. This created a predictable tariff for the delivery of a letter, which in turn allowed postal operators to adopt a pre-paid international letter delivery service. The postal operators used intergovernmental agreements to allow the destination postal services to accept the mail bag and not open up and weigh every letter.

Before the UPU international mail was supported by a patchwork set of agreements between governments in bilateral agreements. Individual countries had to negotiate, where they could, bilateral arrangements with other countries for the forwarding of mail. This might’ve involved private arrangements with transit couriers to deliver the mail to the destination country. Once at the destination country the destination postal system would deliver the mail to the address. This did not scale easily to all countries as the bilateral arrangements were piecemeal. This was complex because different countries imposed different conditions to accept mail and used different fees. And for many countries around the world such arrangements were simply not possible to make those bilateral arrangements with state organisations and private carriers were often very unreliable. By the late 1860’s early 1870’s it was evident to all around the world that some form of uniform arrangement was highly desirable. There was the universal expectation that the British Royal Mail would provide leadership here, working from the earlier success of the domestic Penny Postage reforms. But this was a case where the British did not take a lead, largely for domestic reasons. The two national proponents in the 1870’s were America and Germany, while the Swiss agreed to chair the arrangements.

The General Postal Union

An 1863 international meeting in Paris failed to come up with an agreement between the largely European collection of national postal services. Eleven years later, the International Postal Congress in Berne in 1874 managed to obtain agreement, and the “General Postal Union” was established. The new framework which encompasses the European nations, plus Russia, Egypt and the US, normalised the various classes of mail and the auxiliary postal services. Their work on the classification of “Printed Matter” resulted in a definition that was essentially non correspondence, which allowed many other objects other than letters to be handled by the postal services.

One of the big questions was “who should pay for the delivery of mail?” While the sender paid the local postage agency for the feisty part of the onward route of the letter, much of the expense was incurred on the delivery of the item in the addressed country. Who should pay the destination postal agency for the delivery of internal post items?

The 1874 conference played with the idea of performing inter-agency financial settlements, recording every instance of a bilateral item of postal delivery, debiting the sending country for every delivered item and crediting the delivering country the same amount, based on a universal postal delivery charge. However, the additional record keeping and associated arrangements for financial settlement per item of delivered mail were considered to pose an almost insurmountable barrier to the scheme. In a desire to get things going quickly, they come up with a pragmatic settlement of “sender keep all” where the sending country would retain all the postage revenues for international services. This was far simpler to manage, of course, and on the assumption that the volumes of mail items between two countries was roughly similar and the delivery costs were also roughly similar, then this arrangement would have a similar net outcome as compared to detailed per transactions financial settlement with significantly lower overheads. This interim measure remained the situation in the ensuing decades and proved surpris[ingly] effective.

The next question was: How much? Should each country be able to set its own rates? Or should there be a single international rate, akin to the British domestic Penny Postage service. The 1874 General Postal Union (GPU) conference came up with a schedule of permitted postage rates. The higher rates were intended to cover the delivery costs where the population was widely dispersed and access was expensive, as was the case for the Russian and Ottoman empires.

A number of countries kept its previous rate. The UK took the decision for a single international rate of 2½d rate for all destination countries, including the US. The US was taken by surprise by this move as this created postage rate asymmetries and as these were public agencies such asymmetries had political ramifications. More than a few countries matched the British 2½d international postage rates.

The initial experience with this flat rate for all countries was further refined in the 1876 GPU conference. Four countries, Germany, France the US and Great Britain came up with a “base rate” that would apply as a minimum rate in all cases to all countries, where the international postage rate was double the domestic postage rate. The theory was that no country would lose out in this scheme, particularly if the volumes of mail in and out of each country were roughly equivalent. The “overspend” on the part of the sender for outgoing mail would fund both the cost of the transit delivery to the addressed country for outgoing mail and the domestic delivery of incoming mail. Noone would be any better or worse off under this sender keep all scheme with a double domestic rate postage tariff. This became the standard rate for all GPU members. In 1878 the GPU changed its name to the Universal Postal Union, or UPU.

Refining the Balanced Traffic Model

Subsequent experience proved that these assumptions about roughly balanced mail traffic volumes were not universal. The original assumptions of a smaller group of countries with an exchange of personal letters that balanced in overall volume was being eroded by the changing patterns of use of the

international postal system. In 1900 Italy complained to the UPU Congress of a large a volume of incoming periodicals being posted into Italy, but it was sending none to other countries. Italy requested an amendment to the constitution of the UPU to require payment for the delivery of the extra incoming mail that exceeded the volumes of outgoing mail. The UPU did nothing on this request at the time. And it the ponderous manner to which many international bodies aspired to, they did nothing for some sixty years!



The twentieth century saw the challenges of air mail, the challenges of maintaining a postal system during times of war, and the challenges of rampant domestic inflation in some countries. The latter factor caused a set of distortions in the postal system which still exist today, where it can be cheaper to privately ship mail bags to a country where the outgoing international postal rate is relatively low, and pass the mail items into the postage system in that country as international mail. This sustains anomalous behaviour where it may be cheaper for a postage agency to ship mail to

another country than to put it into the domestic postal system. Businesses in the Netherlands noted at one point that it was 90% cheaper to use Hungary to send mail to Netherland addressees than to use the domestic postal system in the Netherlands. The Netherlands threatened to block incoming mail from Hungary to prevent this anomaly from wreaking havoc on their domestic postal agency.

Air mail was considered an extraordinary service for some time, but unlike land-based transit, the air mail system passed over land borders and landed the mail directly in the addressed country, bypassing all land-based transit fees. The British shifted to an all-air mail delivery system as it was a cheaper method for all outgoing mail.

Public and Private Mail Services



After 1945 the UPU gained many new members as former colonies became countries, and with the inauguration of the United Nations the UPU affiliated itself with the UN. For a while the model of the postal system being essentially a union of public postal services operated by national entities. The 1950's saw the introduction of private mail couriers, who wanted to have functional interfaces between their services and the public services. One of the early players in the space was the Australian courier company TNT, who, along with the US enterprises Fedex and UPS, led the charge of the private sector into this up to now refined public space previously exclusively occupied by national monopolies. The private carriers were early adopters of automated systems and re-introduced per-item inter-provider settlement tariffs, but at the same time introduced delivery time assurances, such as same-day services between certain destinations.

By the 1990's this was more than a complication, and it created a major schism in the postal world. With the progressive privatisation of former public enterprises in the 1980's and onward the distinction between private and public postal service enterprise was largely lost. TNT, for example, took over the Dutch postal service, KPN, forming the TNT Post Group in 1998 that provided the public postal service for the Netherlands. At that point the UPU had to incorporate private sectors post operators as well as national entities into its structure.

UPU Terminal Dues

The problem of mail imbalances became a larger issue in the UPU post-1945. Like many international bodies of the time the UPU operated their decision-making structure along the lines of "one country one vote", and the introduction of many post-colonial countries into the UPU changed the UPU quite radically. Developing countries complained that the incoming mail volumes were far higher than the

outgoing mail volume. The notion that a double domestic postage rate for outgoing mail essentially ‘prepays’ for the consequent incoming reply relies on a balanced mail volume falls apart when the incoming mail volumes rise far above outgoing mail. This was largely due to the small parcel and packet volumes, and these smaller countries made the case that they wanted delivery payment for these small packets and countries.

The UPU responded by making a distinction between the so-called “developing” countries and the others. Developing countries received delivery payments from developed countries to compensate them for this imbalance in mail volumes. These payments were made irrespective of the actual mail volumes, so in the case of some Asian countries, including China, they were classified as developing so that they received delivery payments for incoming parcels despite the fact that they had a net outgoing mail volume, shipping packets and parcels to Europe and North America.

The UPU introduced “Terminal Dues” in 1969 as their response to mail imbalances. These terminal dues used a process of aggregate volumes of incoming and outgoing mail between any two countries, paid by the net sender country to the net receiver. The advent of electronic commerce in recent times saw countries such as China become high volume net senders and other countries become high volume net receivers, completely upending the established pattern of UPU-determined bilateral financial settlements arising from these terminal dues.

The terminal dues are set at a common level for all countries. That level is below the level of comparable domestic costs for mail delivery in many if not all developed economies. What this means is that the national postal service is forced to cross-subsidise the costs of delivery of incoming postal items with the revenues gathered from the domestic postal service, placing further stress on the domestic postal tariffs, and directly benefiting the sending country. Furthermore, the terminal dues rate paid by developing countries is lower than that paid by developed countries.

This has led to some anomalies in the international environment. Clearly, those countries who are net senders of international mail items are recipients of net terminal dues payments, and China is a good example of a country that has derived net benefit from these arrangements. The British, German, French and Dutch providers responded by expanding their previously domestic postal service into the international delivery market, leveraging this international business to offset their domestic terminal dues net payments. Other national postal services who have remained within their national borders are the big losers here, as they face increasing losses due to terminal dues. Their only available revenue raising measure is to raise domestic tariffs which causes a downward pressure on domestic mail volumes. More individuals and businesses turn to electronic services in response to the increasing prices. The consequent erosion of the domestic mail service further drives domestic commerce to e-commerce services and private courier services, exacerbating the national postal service problem and hastening the decline of these national enterprises.

E-commerce has also increased the pressures on this system in other ways. It has allowed consumers to order products directly from international sellers and manufacturers and this has, in turn, shifted the burden of delivery costs from a small set of import agents to the domestic postal system. For example, to appreciate the scale of these changes, it is noted that out of 100M postal parcels entering the EU in 2017, two thirds were posted in China.

Erosion of Terminal Dues

Membership of the UPU is a burden in Australia, New Zealand, Canada, Italy, Ireland and many other countries where the imbalance of internal postal volumes works against the national postal carrier. Obviously in the UPU there is a huge split between the net importers and the net exporters. The net importers are advocates of higher terminal dues rates. The United States had previously supported low

terminal dues rates, but as the postal content shifted from documents to packages, the US postal service became a net importer and the terminal dues revenue was not covering the costs of delivery of incoming packages and parcels, predominantly from China. The Trump administration saw itself as being on the short end of the stick and saw the terminal dues framework as one that imposed further costs on US domestic commerce. The administration resigned from the UPU. The US subsequently withdrew its resignation only when the UPU voted to permit the US to set its own terminal rate for the delivery of inbound packages.

The shifting patterns of use of the postal service, and its transformation from letters through periodicals and documents to today's package and parcel business have proved to be the major challenges for the UPU. This has been compounded by the presence of private courier services that cherry-pick the postal services on otherwise loss-making services, while retaining higher margin services within their own operational sphere.

The UPU is facing some major existential challenges. The original benefits of facilitated transit and uniform inter-provider rates has all but vaporised. The residual asset was the mail bag system and the UPU measure was that the screening of incoming mail was the responsibility of the receiver, performed at the receiver's cost. But this too is falling apart. Many countries are now requiring the sender to provide the information used for package screening, pushing most of the burden relating to describing the contents of these packages and the grounds for customs clearance back to the sender.

Lessons

What can the experience of the UPU tell us about the Internet's arrangements for settlement and peering? The postal system, the telephone system, the airline system and the Internet have all had to address the issue of multi-provider service delivery, and each of these activities have addressed the issue in various ways.

The UPU came the closest to make a Sender Keep All interconnection mechanism viable. To maintain the stability of such an arrangement does not necessarily imply that all providers are the same size. But what is implied is that all bilateral arrangements are made between roughly equally sized transaction volumes. The correspondence model was based on an exchange of letters, where volumes in and out were commensurate.

Where the exchanged volumes are not symmetric then the Sender Keep All arrangements are not stable, in that one party is leveraging the other. In the postal situation when the predominate transaction shifted from correspondence to parcels asymmetries emerged and the Sender Keep All arrangements became unstable.

The Internet inter-provider space has also made extensive use of Sender Keep All arrangements, but they are only used by mutual consent. This is based on a perception of equal value in the arrangement. If one party has a greater dependence on the arrangement than the other party would be advantaged by converting the arrangement into a customer/provider arrangement. By making the interconnection into a market rather than a common imposition then parties are free to enter into Sender Keep All arrangements on the basis of mutual agreement, and free to set out the terms when the agreement is rescinded. Parties are not forced into an arrangement where they are disadvantaged for the sake of imposing a common arrangement on all actors.

The UPU transformed these agreements into a form of sender pays agreements with the introduction of terminal dues. The issue with the terminal dues system was that it was not necessarily a cost-based schedule of tariffs, and a form of structural cross subsidisation in the form of discounted rates to so-called developing countries entered into the structure. From an economic perspective there is much to

be said for removing hidden cross-subsidies is such exchange systems. Aid to developing countries should be a deliberate decision with an overt payment structure. The approach of deeply embedding discounts into exchange fee schedules has the potential to create anomalies that expose the entire system to leverage and potential abuse. The US experience with terminal dues is a good example of the problems that can arise here. Similar issues were evident in the telco sector with the ITU structure of international call accounting settlements, where the call termination rate included a cross-subsidy to developing nations.

The Internet has not gone down this path of balanced settlement payments. The critical observation as to why this has not happened is that stable inter-provider arrangements need to be based on stable uniform retail tariff structures.

A good example of what happens when this is not the case can be found in the telco regime when the retail tariff for local calls was an untimed call for a flat fee, while the inter-provider tariff was based on a per minute transfer charge. In the dial-up modem market of the early 90's some retail telcos saw advantage in hosting banks of modems for dual-up ISPs so that they could charge termination per minute charges to other retail telcos who were only getting a flat fee from their customers for the modem call. The timed interconnection fee revenues could be used to both subsidise the modem banks and provide a net financial return to the telco.

The retail tariff structure for the Internet is not based on "calls" or any other form of discrete customer transaction. There have been efforts to base a retail tariff on received volumes in the so-called "capped" volume-based tariffs, but these tariffs do not extend into the inter-provider space, as they are generally seen as a measure to mitigate over-consumption in the access network of mobile networks, and have no clear counterpart in a wired network. In general, customers pay their access ISP for a "whole of Internet" service. It's then no surprise that the access ISP pays its providers for a "whole of Internet" service as well.

This leads to the rather sparse set of choices in the Internet:

- We can agree that one of us is the provider and the other is the customer, and the customer pays the provider, and after that it's just a matter of determining the price.
- We can agree that we can't figure out who should be the provider and who should be the customer, but we both have reached the conclusion that we'd be better off interconnecting than not. Then we can agree to exchange traffic on a Sender Keep All basis and feel comfortable with this.
- We can't agree on either of the above, in which case there is no interconnection.

While the future of the telephone world and the institutional foundation of the ITU-T sounds today like a contradiction in terms, there is still a future in the parcel delivery world and the logistics behind that. Increasingly however there is little to distinguish between the remnants of the national postal systems and the private courier services. Within the world of these courier enterprises the same forces of volume-based economies of scale sustains of a small set of dominant actors across the international space. This does not bode well for the outlook for the UPU and its residual national constituency. And as for the Internet, the almost unconscious decision to leave the entire system of interconnection to market forces appears in retrospect to have been one of the most robust decisions we've made. It's allowed the provider

framework to adapt to the changing picture of the Internet from edge-based service models to core-based content delivery seamlessly, and I'm confident that when the time is right it will allow us to adapt this environment to future service models with equal ease. If the history of the UPU tells us anything, it should warn us of the perils of memorialising in an institution the arrangements that work today, because it inevitably adds to the burden of the future in resisting change and imposing a dead weight on further innovation.

Disclaimer

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

Author

Geoff Huston B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region.

www.potaroo.net