March 2007

Geoff Huston

Infrastructure ENUM

After much initial fanfare a couple of years ago ENUM has matured to a state where it is currently yet another under-achiever in the technology deployment stakes. ENUM initially presented itself as a very provocative response to the legacy telco position of monopolising public voice services through their exclusive control over the Public Switched Telephone Network (PSTN) and the associated controlling position over the telephone number space (the so-called "E.164" number space, after the ITU-T recommendation E.164 which recommends country code assignments for switched telephony services). The perception was that ENUM was going to dismantle these levers of control and open up the voice market to a new wave of competitive carriers. If the address plan was the key to the PSTN, then ENUM was intended unlock this network and position the new wave of Voice Over IP (VOIP) carriers to take over any residual treasures of the traditional voice market.

Events have not played out according to these expectations, or at least not yet, and not via ENUM as it was originally conceived. Yes, there is a lot of VOIP happening, from user services such as Skype and Google Talk through to VOIP systems that have pretty much replaced the historic PABX configurations. But ENUM itself is not taking over the voice services world. We have yet to see a major change in the major regulatory regimes to head beyond limited scope "ENUM trials" and to open up their entire national number space to ENUM registration. We've yet to see end users in a position to exercise complete and total control over the manner and type of service delivery associated with their phone number.

ENUM, it would appear, is languishing quietly in an increasingly dim and dusty regulatory corner.

But that does not mean that the VOIP industry has been passively waiting for interest in ENUM to pick up. Indeed, a widely held view is that the original concept of ENUM completely missed the target and that the real leverage of ENUM is not as a service for end users to express their preferences, but as a means of inter-carrier signalling, or "Infrastructure" ENUM as distinct from the original concept of "User" ENUM. In this article I'd like to look at the current state of Infrastructure ENUM and look at the tensions that this topic has raised between the open and closed service models.

A Recap of User ENUM

If you are attached to a VOIP network that is connected to the wider PSTN network via some kind of VOIP gateway and you make an outgoing call, then how should your local gateway route the call? The initial operation is to see if you are calling someone else that is behind the same VOIP gateway, in which case the call can be completed locally via VOIP. If not then the only thing the gateway can do is to pass the call request out into the PSTN for call completion, and pay the PSTN the standard call termination charges to complete the call. But if your local VOIP gateway had a means of access to "complete" knowledge of the ways to reach every dialled number, then if the number you called was in some other carrier's VOIP world, then, ideally, your gateway could avoid using the PSTN to complete the call, and simply route the call via IP to the corresponding called party's VOIP gateway.

From this perspective, the core User ENUM problem is how to support a PSTN bypass function. In other words, how can a VOIP gateway dynamically discover that a called number maps to a VOIP service that is directly reachable via the Internet, without resorting to the PSTN.

The approach taken by User ENUM is to use the public DNS as the discovery mechanism. The starting point is where clients publish their call preferences

using the DNS. If I wish to be reachable via SIP, for example, I use the DNS to associate my conventional E.164 number with a SIP address. I create a DNS zone using a transform of my phone number and then insert a Resource Record in this DNS zonefile that references my SIP URI. The calling party (or its agent that is attempting call completion) uses a DNS lookup to establish my call preferences. It applies a simple transform to the E.164 called number to generate a DNS query, and the DNS responses describe the set of services that are associated with that number, the called party's relative preferences and the description of the rendezvous point to actually complete the requested transaction. My SIP URI is returned and the call can be completed using SIP.

So this is "User" ENUM. A zone of the public DNS (**e164.arpa**) is populated with E.164 PSTN numbers Each complete E.164 number is a terminal DNS zone, and is populated with a collection of service rendezvous URIs. This is an opt-in multi-service where each E.164 number may be added into the DNS, and the contents of the zone file reflect multiple services that may be provided by multiple different service providers.

For all its virtues, "User" ENUM really hasn't happened yet. Yes, there are trials, but mass deployment in the scale of hundreds of millions of ENUM entries in the DNS has proved challenging so far.

There are a number of factors here that may explain this situation. It can be quite a complex operation to correctly populate an ENUM DNS zone, and there is the question of who should be maintaining the user's zone. Given that a user may want to use multiple service providers for different services, then how do multiple service providers each maintain a subsection of a single DNS zone without treading on each other's toes? And whose E.164 number is it anyway? If I cancel my PSTN service can I still use the same number exclusively for SIP calls? But within all these potential factors, perhaps the most significant one is that most, if not all, of the traditional PSTN operators see ENUM as a threat rather than an opportunity. ENUM allows users to exercise fine-grained control over service selection, with the potential of constructing a highly volatile and very price sensitive environment voice market where the user has complete control over service selection on a scale of hours or even minutes. For incumbent telephone operators this is undoubtedly a nightmare of the worst order, and ENUM represents a future that simply has to be resisted in any and every possible way.

Even the oft-touted dreams of E.164 phone numbers taking up the role of a universal identity schema so that one number could lead to your web page, your SIP service, your mailbox, and so on, were perhaps just another instance of incredibly wishful thinking on the part of an insatiably voracious DNS name registration industry. The issue here appears to be that ENUM is simply pushing too hard against a well-entrenched incumbent PSTN voice industry, for whom ENUM represents a future that simply is not aligned to their perception of self-interest.

So if User ENUM is having such a hard time, why has Infrastructure ENUM managed to spark some interest? One possible reason is that Infrastructure ENUM services a slightly different client base. These clients, the competitive voice carriers, have similar cost minimization motives to end users, but also have a far greater ability to translate such motivations into outcomes.

What do the Competitive Carriers want from ENUM?

ENUM takes on a different guise from the perspective of a competitive VOIP-based carrier. From this perspective, ENUM is an efficient way to dynamically discover the ultimate identity of the terminating carrier, their preferences in the way in which the call should be terminated, and the termination rendezvous point. It's not the preferences of the end-user here, but the end-user's VOIP service provider's preferences that are important. This is what is being referred to as "Infrastructure ENUM".

In theory, infrastructure ENUM could be considered to be somewhat unnecessary, as it appears to duplicate the role of the SS7/C7 signalling mechanism in the PSTN framework. So why is infrastructure ENUM of interest to VOIP carriers? As is usual in things ENUM, the answer turns inevitably to money. Not only does SS7 equipment represent a considerable financial and technical investment for a competitive carrier, but SS7 is, in effect, no more than a PSTN traffic director. If the primary objective of the signalling system is to discover the cheapest possible call termination path for any call across all possible transports, then SS7 and the PSTN are generally a last resort, and the most valuable information for competitive carriers relates to alternative IP-based call termination mechanisms that would be used in preference to the PSTN.

From the competitive VOIP carrier's perspective infrastructure ENUM is not really about the end users' fine-grained preferences relating to local service preferences. It's actually about call termination options for number blocks and leverage of the Internet as an inter-VOIP transit network. Infrastructure ENUM is about dynamic discovery of the terminating carrier's preferred call termination paths that bypass the imposed inter-carrier SS7-defined paths across the PSTN, and, consequently. It's about creating sufficient impetus to re-define inter-carrier call accounting settlement regimes and related call termination fees. The essential characteristic of infrastructure ENUM is that its all about describing entire E.164 carrier number blocks, rather than opted-in individual end user's numbers, and it's motivated by redefining the financial aspects of inter-provider arrangements, rather than publishing the end user's individual preferences.

Infrastructure ENUM is ideal for competitive voice carriers to announce to other carriers a collection of rendezvous points for termination of services that relate to calls made to numbers within their number blocks. Infrastructure ENUM is logically populated by number blocks rather than individual numbers, for which the carrier is the nominated carrier of record. Infrastructure ENUM allows carrier to announce the services that the carrier is willing to terminate and the associated rendezvous points that will perform service termination. It allows the carriers to undertake call service termination bilaterally, removing any third party intermediaries from the picture, and also removing any associated forms of call termination fees and call accounting settlements that these intermediaries would normally impose on these competitive carriers. This is why Infrastructure ENUM, or I-ENUM, is a topic of persistent interest in this industry.

I-ENUM

The promise of I-ENUM is best seen with an example. Imagine that you are a PSTN carrier that supports IP-based services internally, and you are using E.164 numbers for called party identification when undertaking service completion operations. So VOIP, MMS, SMS, and so on, are all terminated to a a service termination point using E.164 numbers as the lookup token into some service database. When presented with a call request then the only available database to consult is the E.164 number database, and to achieve this you need to launch an SS7 request, and you'll get back an E.164 record, a carrier of record for that number and a path to complete the call. But what if you don't like the answer? What if you with to use IP-based carriage services to bypass any intermediate PSTN carriers, and perform the call completion function directly with the terminating carrier, avoiding forced intermediate carrier-imposed settlement changes?

Here I-ENUM comes into play. I-ENUM is an ideal way for VOIP carriers to selectively announce to other VOIP carriers a set of rendezvous points for service termination. Its an implementation of an inter-carrier PSTN bypass, if you want to look at it that way.

Mechanically, the approach is to announce in some I-ENUM DNS domain the E.164 number block for which this carrier is the carrier-of-record, and populate this DNS zone with the description of the services that the carrier is willing to terminate, and nominate the IP rendezvous URI that performs service termination in the terminating carrier's network.

So it's the same ENUM technology, but now it's the carriers attempting to undertake the discovery and termination operation relative to the terminating carrier rather than relative to the end user. The operations are similar to ENUM, but translated into a carrier context:
   o   Identify the service being requested
   o   Lookup the called number in the I-ENUM DNS domain
   o   Select the terminating carrier's URI for a compatible terminating service entry that is published against an enclosing number block entry.

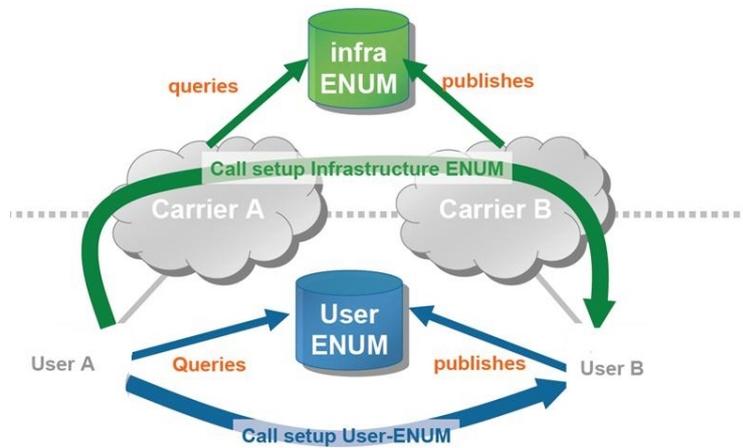o   Complete the call request using a service rendezvous operation

Figure 1 – Infrastructure-ENUM and User-ENUM

It appears that what carriers want from ENUM is the ability to selectively disclose E.164 called numbers to nominated terminating service bindings that are accessible to certain other carriers at specified rendezvous points. They would like this disclosure to be under the full control of the terminating carrier, and that the information disclosed by the terminating carrier to be relative to a nominated carrier's query, rather than being a unilateral act of publication for any party to query. In other words they would like the answer to be variable, depending on the party making the query. They would like number blocks to be described via this mechanism, and the publication framework to be maintained with a minimal provisioning overhead and minimal operating expenditure. What the carriers want is for the originating and terminating carriers to be jointly in full control of their service interconnection policies, and for the products of these policies not to be unnecessarily exposed to a wider audience. So neither the number blocks, the services, the rendezvous points or the bilateral interconnection arrangements are necessarily to be exposed to a wider audience here.

Implementing I-ENUM

It appears that the industry knows, roughly, what it wants, but the means of achieving it are not so clear.

The approaches to I-ENUM appear to be along the lines of:
- o   splitting the existing ENUM DNS delegation hierarchy into multiple trees close to the root of the hierarchy and operate ENUM and I-ENUM as parallel and distinct hierarchies, or
- o   work at the ENUM terminal point and enrich ENUM with further DNS Resource Records and query sequences, or
- o   use 'private' ENUM contexts.

A - Splitting the DNS ENUM hierarchy

This approach to Infrastructure-ENUM uses the same NAPTR DNS Resource Record entries as User-ENUM, and uses the same general lookup mechanism to resolve a called number to a collection of URIs. Here the regular expression substitution capabilities of NAPTRs can be used to take a general NAPTR resource record form and generate a called-number specified rendezvous URI to perform call completion.

This approach has no change to ENUM Resource Records, no change to NAPTR capabilities, and no change to the operation of the DNS. It just relies on a split ENUM DNS hierarchy, as indicated in Figure 2, where a new DNS tree, named something like "ie164.arpa" would hold I-enum records, while leaving "e164.arpa" to continue to contain User ENUM records.
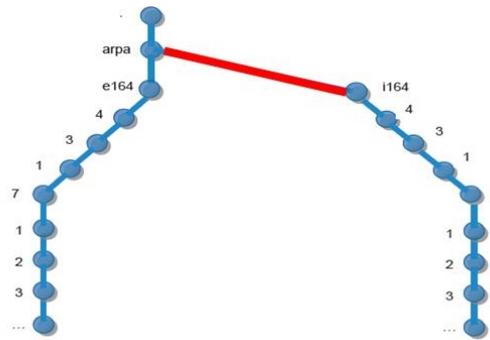
Figure 2 – Split ENUM hierarchy for I-ENUM

And here's where the problems start!

The e164.arpa ENUM hierarchy was a significant challenge to all concerned. Representing one realm's identity space within another quite distinct identity realm is invariably ill-advised, and more so when the political framework that governs to two realms differ markedly. The E.164 number space has its challenges, where, oddly enough, "countries" are sometimes difficult to define and difficult to deal with.

> The classic example of the "country problem" in the E.164 context is Taiwan. Yes, you can place a phone call to a Taiwanese number in most parts of the world, but you'll not find the country "Taiwan" listed in the E.164 number plan under any alias whatsoever. Its just not there! And its not in ENUM as a consequence of the ITU-T and IETF arrangements over ENUM governance that allow the ITU-T the right of veto over the inclusion of country code entries in the e164.arpa DNS space. So how can country dial code for Taiwan exist within the PSTN, yet be completely excluded from ENUM? The answer is that the nature of the inter-governmental treaty arrangements that lie behind the ITU allow some member states to exercise a veto on the recognition of others as peer member states, and Taiwan is the subject of such a veto. In the PSTN world this is not exactly an insurmountable problem, in that PSTN operators can add additional configuration entries into their international switches that in essence 'synthesise' Taiwan as if it had its own country code. And by some coincidence almost every country has decided to use the same E.164 country code. The DNS is not so flexible in this respect and a similar approach to mapping +886 into 6.8.8.e164.arpa using informal local arrangements simply does not work that way!

So if ENUM has presented some particularly tough challenges, why would I-ENUM be any easier to construct? Why is there a need to create a second DNS hierarchy that maps from E.164 to DNS resource records? Why does this require the involvement of member states of the ITU-T? Why would any service provider ask for higher levels of government intervention and regulatory control in competitive signalling infrastructure? The question is relevant because the use of "ie164.arpa" for I-ENUM will require government approval and involvement, and in deregulated regimes this is precisely the opposite of the current regulatory direction.

Another aspect of the emerging service industry is that there is no longer a uniform scenario of monolithic all-inclusive service providers, and instead we have a variety of actors, some of whom specialize in offering only a small selection of services. If we want to preserve user choice here then we probably need to get back to addressing the basic ENUM problem of many hands wanting to edit the same DNS zone file. If we don't wish to bias I-ENUM for this monolithic single vertically integrated service provider world then you have the problem of multiple service entities wanting to maintain their service record within the same DNS zone file. So the

basic question here how do you know that 2 DNS "trees', namely e164.apra and i164.arpa, are enough? And if not, then how many trees really is enough? Who gets to make that decision?

We have been here before, by the way, and in that case ENUM was seen as being the solution, not the problem!

Earlier work in this space by Marshall Rose in 1992 used the same approach of placing rendezvous information in the DNS, using a transformed E.164 number as the lookup key into the DNS hierarchy rooted at "tpc.int". Here the exercise was not call termination, but fax termination, where the A record in the DNS mapped to an internet termination point was capable of completing the 'last mile" as a true fax. A messaging paging service was later added using a new hierarchy of E.164 numbers, in a tree rooted at "pager.tpc.int". The implication was that each additional service implied a need to create and populate a new DNS tree of the form <service>.tpc.int.

ENUM was a way of combining all these service-specific DNS hierarchies into a single number hierarchy, where the services are placed in resource records at the leaves of the tree, rather than creating a multiplicity of service-specific trees.

It seems to me that ie164.arpa, and its variants that also rely on a split of the DNS tree, is not really an answer to the carrier's needs in this space. It manages to bring back into focus the relatively ad-hoc arrangements between the IETF and the ITU-T that did resolve the situation in terms of control of the number space to any party's satisfaction, without any new insight that might make this task more straightforward the second time around, and contains no promises that this won't occur a third time, or even more!

It appears, sensibly, that we are somewhat reluctant to simply repeat the mistakes of the past, and the concept of multiple instances of the entire E.164 number space in the DNS is not one that looks like it is the most appropriate way of supporting I-ENUM. But we are nothing if not endlessly creative, and the next proposal is that each national regime start its own infrastructure enum hierarchy just below the country code. So if you happen to look at Australia (E.164 country code of 61), then the enum DNS delegation tree is a 1.6.e164.arpa, and the infrastructure enum tree for Australia would be rooted at "infra.1.6.e164.arpa." Of course not every country code in the E.164 recommendations is a two digit code. The North American numbering plan uses the E.164 code is 1, while many other countries use a three digit code. Aside from loading static tables (something we are trying to avoid in this entire dynamic information discovery domain), how do you know which country code is of which length? And would every national regime use the English abbreviation "infra" to label its infrastructure ENUM hierarchy?

B – More games with NAPTRS

It appears that the "ie164.arpa" is incapable of meeting all the various service-specific requirements that may arise from I-ENUM, and that a more general approach is a collection of DNS hierarchies of the form <service>164.arpa. But this is also a relatively unattractive solution, in that it tends to create more and more DNS hierarchies.

The original ENUM concept was to compress these multiple trees into a single instance and use a set of NAPTR resource resources to describe the set of services associated with an E.164 number. Could this be extended to include I-ENUM requirements? The idea here is that the service records are not placed in a single DNS zone file, but are themselves further points of delegation, allowing the user to delegate DNS zones for individual services.

There is a proposal intended to achieve precisely that by introducing a number of modifications to the syntax and semantics of the NAPTR DNS resource records to allow for NAPTR records that point to further delegation zones rather than terminal entries. This is intended to allow a leaf ENUM zone file to contain service-specific

entries that in turn point to service-specific DNS zone files. This allows a distinction to be drawn between the queries of what services are associated with a particular E.164 number and what URIs are the rendezvous points for these services. An example of this approach is shown in Figure 3.

```
DNS information

$ORIGIN 3.8.0.0.6.9.2.3.6.1.4.4.e164.arpa.

        NAPTR 10 100 "u" "E2U+sip" "!^.*$!sip:info@example.com!"   .
        NAPTR 10 102 "u" "E2U+msg" "!^.*$!mailto:info@example.com!"   .

        NAPTR 10 100 "d" "E2U+sip" "" 3.8.0.0.6.9.2.3.6.1.4.4.e164.arpa.
        NAPTR  10 102 "d" "E2U+msg" "" 3.8.0.0.6.9.2.3.6.1.4.4.e164.arpa.

$ORIGIN _e2u.3.8.0.0.6.9.2.3.6.1.4.4.e164.arpa.
        _sip   NS  sipservice.example.com
        _msg   NS  mailservice.example.com

$ORIGIN _sip._e2u.3.8.0.0.6.9.2.3.6.1.4.4.e164.arpa.

        . URI 10 10 "sip:info@example.com"
        . URI 10 10 "sip:info@example2.net"

$ORIGIN _msg._e2u.3.8.0.0.6.9.2.3.6.1.4.4.e164.arpa.

        . URI 10 10 mailto:info@example.com
```
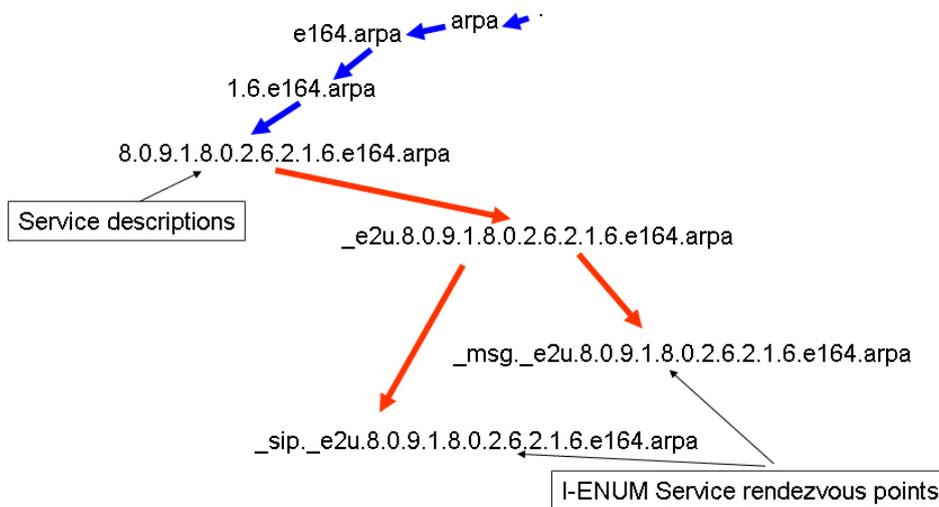
Figure 3 – Delegated service rendezvous descriptions in the DNS

What this ends up is still a rather unsatisfying compromise. On the positive side this approach avoids endlessly replicating ENUM delegation trees for each service type, and it avoids the administrative problem of multiple service providers attempting to edit the same single DNS zone file, which, in general is Good. On the other hand this does represent more load on the DNS, with yet another Resource Record type (the URI type) and another layer of indirection when attempting to query the DNS, which is Not So Good. However, it still exposes the inter-carrier service rendezvous points to public view and exposes inter-carrier signalling to public view, which is probably Bad. It requires the inter-carrier information to be placed at the leaf of the ENUM delegation tree, and not at the higher point in the delegation tree that corresponds precisely to the carrier's number block, and this is just plain Ugly!

As far as I can see, both these approaches fail in critical ways. If the desire from I-ENUM is the ability to perform selective disclosure of information then these approaches simply fail to meet that objective. From the carrier's perspective there is an evident aversion to make this inter-carrier service termination information generally available to the public because of the ever-present risks of exposing the service to denial of service attacks and attempts to subvert the service. And the desire to disclose responses that may vary according to the carrier performing the query is to implement bilateral interconnection policies that may entail different services being exposed to different carriers.

So is I–ENUM looking for answers in all the wrong places? Is DNS just the wrong tool to implement these requirements?

Perhaps not, but here we need to consider the differences between "public" and "private" DNS capabilities.

C – Private I-ENUM

Another approach to the I-ENUM requirement is to use ENUM technology, but base this upon a private DNS structure that is not part of the public DNS hierarchy. There is no doubt that as far as distributed databases go DNS is a remarkably slick, efficient and cheap solution, and possibly it really the best in its class. So if the problem is just the "public" part of the public DNS, then another approach is to continue to use the DNS, but use a private DNS hierarchy.

Private DNS allows each carrier to craft its DNS advertisement to suit individual interconnection arrangements, and through DNS filters and DNS selective response generation only divulge information to the correct querying party and only expose service rendezvous points according to the associated interconnection policy with the carrier in question.

This approach does not require any additional DNS Resource Records, nor does it create a multiplicity of service-specific ENUM DNS hierarchies, nor does it create the situation of multiple distinct service providers all attempting to edit parts of a single DNS zone file. Just as critically, private DNS approaches can be accompanied by various forms of restricted access to the DNS service. The implication is that service rendezvous points are not unilaterally published in a relatively hostile world. The service rendezvous points, and the number blocks under the control of each carrier are not implicitly passed into the public domain through use of private ENUM.

There is no compelling need to have inter-carrier service termination information made public through a public DNS I-ENUM model, and there are strong arguments why the opposite approach of private bilateral information exchange is a better fit to the actual inter-carrier industry dynamics here. It should come as no surprise to learn that in this space I-ENUM as a private DNS solution is already replete with vendors, products, customers and carrier users.

I-ENUM and The Standards Perspective

The IETF's ENUM Working Group has been considering I-ENUM for some years, taking the general approach of adapting the User ENUM structure in e164.arpa as the starting position. This activity has served to highlight the observation that the requirements for end user preference settings in U-ENUM are intrinsically quite different for carrier interconnection policy expression in I-ENUM. The general directions of the ENUM WG's efforts to resolve this difference between the ENUM requirements are described above. Interestingly, the ie164.araa-styled approach using a Branch Location Record remains a Working Group item, and appears to have reached a consensus position in the Working Group. The next step is that of publication as a Proposed Standard in the IETF.

But if the industry's preferred approach truly lies in the direction of private DNS approaches then this calls into question the potential role of the IETF in defining technology to support the information exchange between various forms of service providers. Should the IETF embark of a rather difficult path to standardize a second E.164 hierarchy within the DNS and once more embark on a rather complex conversation with the ITU-T over relative roles and responsibilities for this administration of this second ENUM domain? Should the IETF embark

of standardization of an information model for carrier inter-connection that has some distinct implications on the types of inter-connection arrangements that can be supported within this standards framework? Who should be setting business roles for carrier-to-carrier interconnection? Is this logically or practically within the purview of the IETF?

This appears to be a case where standards cannot provide a complete framework in addressing the various requirements and managerial perspectives where the DNS and the E.164 identity spaces are made to intersect. While technology folk can reach a consensus position that a particular technology is viable and can produce running code, the ultimate acceptance of that standard by industry players relies on the additional constraint that the standard serves a real need, and does so in a manner that at the very least does not constrain the actions of these industry players.

Perhaps the problem is actually with the industry itself, and the underlying extent of the changes that are happening within the industry as VOIP-based competitive carriers grow into being significant players in the voice market.

The Emerging Realities of the Global PSTN Industry

If the competitive VOIP-based voice carriers can leverage VOIP and Infrastructure ENUM to dismantle the dominance of the conventional regulatory-inspired call termination paths and the associated inter-PSTN operator call accounting regimes then it is possible that we will witness a "perfect storm" for the traditional PSTN providers.

VOIP providers are completely uninterested in "call minutes". The revenue stream from the customer comes in the form of a fixed monthly fee, and in this environment variable costs are unacceptable. Consequently the VOIP carriers are demanding "bill and keep" arrangements as opposed to call minute-based inter-carrier compensation. The additional motivation here is that not only does call accounting introduce variable costs into a fixed revenue business sector, but the costs imposed on the industry to perform this form of financial settlement range from the sublime to the outrageous!

So it is possible that the combination of VOIP, Infrastructure ENUM, end-to-end IP carrier–to-carrier services without forced PSTN intermediation, bill and keep services, the avoidance of SS7 and PSTN transit paths all combine to make a mutually supportive set of factors that just might force the residual PSTN carriers completely out of the voice market. This indeed would be a revolutionary outcome for the Infrastructure ENUM effort!

Disclaimer

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

About the Author

GEOFF HUSTON holds a B.Sc. and a M.Sc. from the Australian National University. He has been closely involved with the development of the Internet for many years, particularly within Australia, where he was responsible for the initial build of the Internet within the Australian academic and research sector. He is author of a number of Internet-related books, and is currently the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region. He was a member of the Internet Architecture Board from 1999 until 2005, and served on the Board of the Internet Society from 1992 until 2001.

http://www.potaroo.net