

The Lord of the Numbers

Geoff Huston
May 2002

I don't know about you, but I've been reading a lot about a thing called "ENUM" lately. It appears to be a topic of considerable interest to all kinds of people, from ISPs to domain name registrars to national regulators.

What's ENUM all about, and why all the fuss?

Well, if I were to say that ENUM is the most essential part of the technology that allows the public telephone network to be seamlessly integrated into the Internet, then perhaps you can see why so many parties are interested. The interesting questions are how the telephone system is mapped into the Internet world, and, more particularly, which party will undertake which role.

The basics of the telephony world are very simple indeed. Telephone handsets are little more than a speaker and a microphone and when a call is made, the network connects the microphone of one party to the speaker of the other. So, one of the primary objectives of the telephone network is to support this carriage of a voice signal. Of course you don't need a specialized telephone network to support the carriage of voice. As any user of a desktop computer would confirm, there are now a plethora of applications that can deliver a voice signal across the network. For an application to support a voice conversation, a conventional approach is to use a network base of the UDP transport protocol, with a Real Time Protocol (RTP) overlay, and the RTP payload is an encoded version the original analogue voice signal. Carrying voice signals in real time across an Internet is a well understood network service, with an accompanying set of existing protocols and associated applications.

So if carrying a real time voice signal across the Internet is not the major technical challenge, what's the problem? Looking once more at the handset, in addition to the microphone and speaker there is course the number keypad. It's this keypad, or more exactly the call signaling invoked by this keypad, that is the true value of the telephone network. There is a wealth of associated design to support the notion that if I press enough of these numbered buttons, in the right order, then somewhere else in the world a handset will commence ringing. The implication of this is that each and every telephone handset has a unique number, or "address" and that there is adequate structure within the address space to allow the telephone network's call setup functional entities to know the relative location of each and every telephone address at all times. This address space is often referred to as the "E.164" address space because this is the document number assigned by the International Telecommunications Union to the International Public Telecommunication Numbering Plan (it must be important because it has all those capital letters!).

To make Internet telephony truly useful as an extension of the telephone network, the Internet telephony world has to be able to interface to the telephone network, allowing Internet-connected telephones to make and receive calls in precisely the same fashion as conventional telephone handsets. For this to work, each and every Internet device that supports telephone operation needs to also have an alias in the form of a telephone address. But there's a bit more to it than that. Don't forget that each Internet telephone is also an IP device, and that, for the Internet part at least, the voice traffic will be carried by IP packets which include the IP address of the Internet telephone device. So each Internet telephone needs to maintain both an Internet address and a matching telephone address. It's the mapping from a telephone number to an IP address which is the crucial part.

So when Alice, on a normal telephone, wants to call Bob, on an Internet phone, all Alice needs to do is simply dial Bob's telephone number, or E.164 address. Of course Bob's phone is connected to the Internet, and can't directly receive Alice's call request, so a gateway is necessary. The telephone system will be able to map Alice's call request to an Internet telephony gateway, and the gateway will then need to translate Bob's phone number into an IP address. Here is where ENUM comes in.

Of course the gateway could have a configured list of E.164 phone numbers and IP addresses, and that they way many of them are configured today. But what happens when the IP device is numbered dynamically using DHCP, or if it's mobile, and moves from one service provider's IP network to another? In other words how can this mapping be dynamic rather than static?

The way a dynamic domain name to IP address mapping is maintained on the Internet is through the Internet's Domain Name System (DNS). ENUM is a way of using the DNS to support a distributed mapping from an E.164 address into an IP address. Using ENUM the telephony gateway can request the DNS to return the IP address that is mapped by Bob's E.164 address. The gateway can then use the Session Initiation Protocol (SIP) to send to Bob's Internet phone a call request, which the starts Bob's phone ringing. When Bob accepts the call the gateway can then pass all data originating from Alice to Bob's IP address, and all data received from Bob's IP address across to the telephone connection to Alice for the duration of the call. Alice need never know that Bob is using an Internet device. Alice dialed a phone number, heard it ring and then heard Bill answer the call. For Alice, nothing has changed.

Obviously these days the telephone network is more than simple voice conversations, and any serious attempt to bridge the telephone network and the Internet should also be able to also handle various forms of text messaging as well as faxes, and should be able to seamlessly redirect the telephone service to the appropriate Internet device. In other words we are seeing a requirement that different services associated with the same E.164 address should be able to be mapped to different IP addresses. The approach used in ENUM is to avoid a direct mapping from an E.164 address into an IP address in one pass, but first map an E.164 number onto a collection of service-specific Uniform Resource Identifiers (URIs). A gateway will receive the complete collection of service-specific URIs in response to a request to translate an E.164 address to a URI. Depending on the type of service being requested, the gateway can then select the most appropriate URI and use the DNS a second time to translate the domain name part of the URI to an IP address using the URI-specific DNS resource record as a query term. The gateway can then use the full URI specification to open up an IP session with the selected service port and complete the service transaction.

That was probably a bit indigestible, so you can either read the last paragraph a few more times, or follow this example:

Let's say Bob's Internet telephone services are mapped to the E.164 address **+61-2-629599**. When Alice tries to call Bob, the Internet gateway takes the call setup request with Bob's number and firstly reverses the digits. It then inserts a '.' between each digit, and appends **"e164.arpa"**. The resultant DNS string is the fully qualified domain name **"9.9.5.9.2.6.2.1.6.e164.arpa"**. This name is then passed as a query to the DNS, to retrieve all associated **"NAPTR"** DNS resource records.

Let's say that Bob has both a SIP internet telephone and he wants his SMS messages sent to him via email. In the DNS BOB would've placed two NAPTR records for **"9.9.5.9.2.6.2.1.6.e164.arpa"**. One would be a URI for a SIP service, such as **"sip:bob@tele.example.com"**, and the second is a URI to a mail delivery service, such as **"mailto:bob@po.example.com"**.

For a phone call, the gateway picks the **"sip:"** URI and then uses the DNS a second time to translate the domain **"tele.example.com"** into an IP address using a DNS **A** record, and

opens up a session with this SIP server to complete the call setup, requesting a voice session with the user "**bob**" on this server.

If, on the other hand, Alice is sending Bob a short text message, the gateway would use the "**mailto:**" URI and use the DNS to find an **MX** record for "**po.example.com**", use the DNS a second time to find an **A** record for the **MX** target, and then open up a SMTP session with this IP address in order to deliver the message to the mailbox "**bob**".

So if this is all there is to ENUM, why all the fuss? It may be somewhat indirect, but on the face of it there is nothing controversial in this - is there?

Maybe looks can be deceptive. After all, the DNS didn't look all that controversial when it started, and look at where we've got to with ICANN. ENUM appears to have a similar, or possibly even more powerful, set of attractors for regulatory and social controversy.

One issue is: who gets to populate the E164.arpa domain with all these URIs? It could be that this is a responsibility of existing telephone service providers, since after all these entities operate the E.164 address space in each country. It could be that this is a responsibility of Internet Service Providers (ISPs), since the data in the resource records can also describe Internet-based services. Or maybe the end subscriber gets to populate the DNS with their own entries. It is quite conceivable that we could see ISPs that have no direct role in carrying voice traffic wanting access to a country's E.164 number plant in order to provide various forms of ENUM services. Given that each element of an ENUM service collection can use URIs that refer to different ISP services, it is possible that the one ENUM record can be populated by URIs referring to a number of different service providers. This model of multi-agent access to such infrastructure resource records is a novel concept to many regulatory and operating regimes, where a single operator manages the entire associated infrastructure elements that are needed to deliver a service.

But most of the discussion about ENUM has been on more subtle aspects of this mapping. Firstly there's the choice of "arpa" as the common DNS root for ENUM DNS entries. At an international level there's a perception that "arpa" is too American and that a name root of "int" appears to be more neutral. It's been reported that China and France are keen for this change to the ENUM specification.

But there's something else lurking here, which has surfaced within the regulatory debate in the United States. North America has the IDDD code of "+1", implying that under ENUM there is a single DNS domain for ENUM, namely "1.e164.arpa". Single domains imply single operators, and single operators imply a non-competitive monopoly service regime. There has a call for multiple e164 DNS root locations for North America, allowing for two or more competing service operators using different DNS hierarchies to locate their ENUM services. On the one side there is the view that such attempts to create multiple partially populated ENUM name hierarchies to support create competitive service provision in ENUM-based services are no more than an incitement to chaos. This chaos would, in turn, seriously hamper the uptake of ENUM services. On the other hand the competitive provision proponents of multiple DNS root domains for the "+1" IDDD code argue that a government-sanction monopoly is still a monopoly, and this monopoly situation will likely lead to high service prices for ENUM services. This escalated pricing structure would, in turn, seriously hamper the uptake of ENUM services.

The proponents of ENUM envisage a single telephone number as being an alias not only for your internet phone service, but also instant messaging, email, your web page and any other service location point. One identifier is all that would be required to reach you, using a service point of your choice. No more business cards cluttered with phone numbers, fax numbers, mobile numbers, email addresses, web addresses, and instant messaging handles. All we need is just one ENUM address, or just one number, for all these services.

"One number to rule them all, one number to find them, one number to bring them all and in the darkness bind them", is the ENUM version of Tolkien's classic saga.

But one person's ease of use is often another's opportunity to exploit. ENUM can be seen as yet another erosion of personal privacy on the Internet. Not only is it one more step towards a single individual digital identity that could be used to track individuals within the Internet, on a more immediate and practical level it allows spammers a wealth of new ways to drive you to complete distraction.

It appears that the technical components of ENUM are by and large the most straightforward part. The regulatory and social implications of ENUM are more of a concern, and its here that we are entering into *"the Land of Mordor where the Shadows lie"*.

Further reading:

<http://www.enum.org> has a good overview of ENUM and its potential application

RFC 2916 (<http://rfc2916.x42.com>) contains a more formal description of how ENUM uses the DNS