The H Ratio for Address Assignment Efficiency

Status of this Memo

   This memo provides information for the Internet community. This memo
   does not specify an Internet standard of any kind. Distribution of
   this memo is unlimited.

Abstract

   This document was submitted to the IETF IPng area in response to RFC
   1550. Publication of this document does not imply acceptance by the
   IPng area of any ideas expressed within. Comments should be
   submitted to the author and/or the sipp@sunroof.eng.sun.com mailing
   list.

Table of Contents

1. Efficiency of address assignment

   A substantial part of the "IPng" debate was devoted to the choice of
   an address size. A recurring concept was that of "assignment
   efficiency", which most people involved in the discussion expressed
   as a the ratio of the effective number of systems in the network over
   the theoretical maximum. For example, the 32 bits IP addressing plan
   could in theory number over 7 billions of systems; as of today, we
   have about 3.5 millions of addresses reported in the DNS, which would
   translate in an efficiency of 0.05%.

But this classic evaluation is misleading, as it does not take into
account the number of hierarchical elements. IP addresses, for
example, have at least three degrees of hierarchy: network, subnet
and host. In order to remove these dependencies, I propose to use a
logarithmic scale for the efficiency ratio:

$$H = \frac{\log (number\ of\ objects)}{available\ bits}$$

The ratio H is not too dependent of the number of hierarchical
levels. Suppose for example that we have the choice between two
levels, encoded on 8 bits each, and one single level, encoded in 16
bits. We will obtain the same efficiency if we allocate in average
100 elements at each 8 bits level, or simply 10000 elements in the
single 16 bits level.

Note that I use base 10 logs in what follows, because they are easier
to compute mentally. When it comes to large numbers, people tend to
use "powers of 10", as in "IPng should be capable of numbering 1 E+15
systems". It follows from this choice of units that H varies between
0 and a theoretical maximum of 0.30103 (log base 10 of 2).

2. Estimating reasonable values for the ratio H:

Indeed, we don't expect to achieve a ratio of 0.3 in practice, and
the interesting question is to assert the values which can be
reasonably expected. We can try to evaluate them from existing
numbering plans. What is especially interesting is to consider the
moment where the plans broke, i.e. when people were forced to add
digits to phone number, or to add bits to computer addresses. I have
a number of such figures handy, e.g.:

* Adding one digit to all French telephone numbers, moving from 8
  digits to 9, when the number of phones reached a threshold of 1.0
  E+7. The log value is 7, the number of bits was about 27 (1 decimal
  digit is about 3.3 bits). The ratio is thus 0.26

* Expending the number of areas in the US telephone system, making it
  effectively 10 digits long, for about 1.0 E+8 subscribers. The log
  value is 8, the number of bits is 33, the ratio is about 0.24

* Expending the size of the Internet addresses, from 32 bits to
  something else. There are currently about 3 million hosts on the
  net, for 32 bits. The log of 3.E6 is about 6.5; this gives a ratio
  of 0.20. Indeed, we believe that 32 bits will still be enough for
  some years, e.g. to multiply the number of hosts by 10, in which
  case the ratio would climb to 0.23

* Expending the size of the SITA 7 characters address. According to
  their documentation, they have about 64000 addressed points in
  their network, scattered in 1200 cities, 180 countries. An upper
  case character provides about 5 bits of addressing, which results
  in an efficiency of 0.14. This is an extreme case, as SITA uses
  fixed length tokens in its hierarchy.

* The globally-connected physics/space science DECnet (Phase IV)
  stopped growing at about 15K nodes (i.e. new nodes were hidden)
  which in a 16 bit space gives a ratio of 0.26

* There are about 200 million IEEE 802 nodes in a 46 bit space, which
  gives a ratio of 0.18. That number space, however, is not
  saturated.

From these examples, we can assert that the efficiency ratio usually
lies between 0.14 and 0.26.

3. Evaluating proposed address plans

Using a reverse computation, we get the following population counts
in the network:

	Pessimistic (0.14)	Optimistic (0.26)
32 bits	3 E+4 (!)	2 E+8
64 bits	9 E+8	4 E+16
80 bits	1.6 E+11	2.6 E+27
128 bits	8 E+17	2 E+33

I guess that the figure explains well why some feel that 64 bits is
"not enough" while other feel it is "sufficient by a large margin":
depending of the assignment efficiency, we are either well below the
target or well above. But there is no question, in my view, that 128
bits is "more than enough". Even if we presume the lowest efficiency,
we are still way above the hyperbolic estimate of 1.E+15 Internet
hosts.

It is also interesting to note that if we devote 80 bits to the
"network" and use 48 bits for "server less autoconfiguration", we can
number more that E.11 networks in the pessimistic case - it would
only take an efficiency of 0.15 to reach the E+12 networks hyperbole.

I guess this explains well why I feel that 128 bits is entirely safe
for the next 30 year. The level of constraints that we will have to
incorporate in the address assignment appears very much in line with
what we know how to do, today.

4. Security Considerations

    Security issues are not discussed in this memo.

5. Author's Address

    Christian Huitema
    INRIA, Sophia-Antipolis
    2004 Route des Lucioles
    BP 109
    F-06561 Valbonne Cedex
    France

    Phone: +33 93 65 77 15
    EMail: Christian.Huitema@MIRSA.INRIA.FR