

Latest revision: 17 August 2001

Author: Geoff Huston (gih@telstra.net)

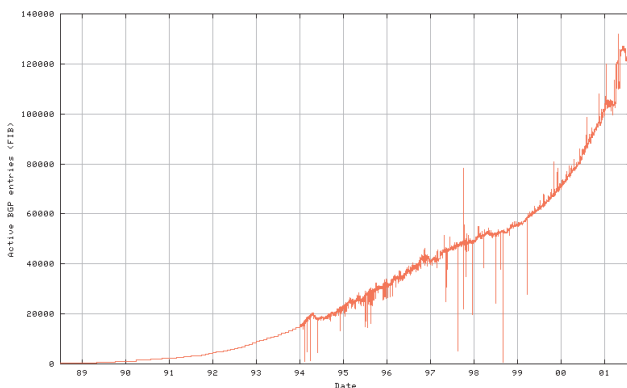
BACKGROUND

The routing table is the complete set of routes that describe the origin of every routed address within the Internet. As new networks connect to the Internet they announce their address prefix into this table. As the Internet grows so does the size of this table. Looking at this table at regular intervals can give us a good idea of what is happening within the routing system.

In routing circles these days you will often hear talk about the Big Question: how will routing scale deal with the demands of tomorrow's Internet? While many aspects of the Internet are prone to scaling pressure, routing appears to be one of the technologies at the pointy end of the scaling problem, and the issues involved are illustrative of the more general issues of technology design within a rapidly expanding domain.

TECHNICAL ISSUES

There's quite a story behind the chart below, and it can tell us a lot about what is likely to happen in the future. The chart appears to have four distinct phases: exponential growth between 1988 and 1994, a correction through 1994, linear growth from 1995 to 1998 and a resumption of exponential growth in the past two year, and some oscillation in the past few months.



Prior to 1994 the Internet used a routing system based on classes of addresses. One half of the address space was termed class A space, and used a routing element of 8 bits (or a /8) and the remaining 24 bits was used to number hosts within the network. One quarter of the

space was termed class B space, with 16 bits of routing address (/16) and 16 bits of host address space, and one eighth was the Class C space, with 24 bits of routing address (/24) and 8 bits of host space. According to the routing system, routed networks came in just three sizes, small (256 hosts), medium (65,535) hosts and large (16,777,215 hosts). Real routed networks however came in different sizes, most commonly one or two thousand hosts. For such networks a Class B routing address was a severe case of oversupply of addresses, and the most common technique was to use a number of Class C networks. As the network expanded so did the number of Class C network routes appearing in the routing table. By 1992 it was becoming evident that if we didn't do something quickly the routing table would expand beyond the capabilities of the routers being used at the time, and by "quickly" we were talking months rather than years.

The solution was termed 'CIDR' or Classless Inter-Domain Routing. The technique was elegant and effective. Instead of dividing the network into just three fixed prefix lengths, you allow each routing advertisement to have an associated prefix length.

CIDR led to the other change in routing policy, that of provider-based addresses and provider route aggregation. Instead of allocating network address blocks to every network, the address registry allocated a larger address block (a /19 prefix) to a provider, who in turn allocated smaller address blocks from this block to each customer. Now a large number of client networks would be encompassed by a single provider routing advertisement. This technique, hierarchical routing, is used in a number of network architectures, and is a powerful mechanism to aggregate routing information.

Through 1995 to 1998 the combination of CIDR and hierarchical provider routing proved very effective. While the Internet continued to double in size each year (or more!), the routing space grew at a linear rate, increasing in size by some 10,000 routes per year. For the routing system this was good news. Vendors were able to construct larger routers at a pace that readily matched the growth of the Internet, and a combination

of Moore's law in hardware and CIDR and hierarchical routing in the routing system proved very effective in coping with dramatic growth in the Internet.

But midway through 1998 something changed. The routing system stopped growing at a linear rate and resumed a pattern of exponential growth again, at a rate of a little under 50% per year. This is a worrying pattern. While the size of the routing table is some 105,000 routes at the middle of 2001, in a years time it could be some 150,000 routes and 225,000 routes the year after, and so on. Within six years the table will be reach some 1,000,000 routes at this rate of growth.

There are many factors driving this pattern of routing table growth, including detailed connectivity policy being expressed on fine-grained prefixes, traffic engineering across a dense mesh of interconnectivity, and even some level of operator inattention to aggregation management. But if there was one driver that appears to be the dominant driver of growth, then it would be, in a word, multi-homing. Multi-homing is when an ISP has a number of external connections to other networks. This may take the form of using a number of upstream ISPs as service providers, or using a combination of upstream providers and peer relationships established either by direct links or via a peering exchange. The way in which multi-homing impacts the global BGP table is that multi-homing entails pushing small address fragments into the global table with a distinct connection policy that differs from any of its upstream neighbours. What we are seeing in this sharp rise in the size of the BGP table is a rapid increase in the number of small address blocks being announced globally.

Connecting to multiple upstream services and connecting to peering exchanges implies the use of more access circuits. While the cost of these circuits was high, the offset in terms of benefit was low enough as to negate most of the potential benefits of the richer connectivity mesh. Over the past few years the increasing level of competition in the largely deregulated activity of provision of communications bearers has bought about reductions in the price of these services. This, coupled with an increasing technical capability in the ISP sector, has resulted in the increasing adoption of multi-homed ISPs. In the quest for ever increasing resiliency we are also starting to see multi-homed customers in addition to ISPs. In the routing system we are seeing the most challenging of environments: a densely interconnected semi-mesh of connectivity with very fine-grained policies being imposed on top of this connectivity. Any topological or other form of logical hierarchical abstrac-

tion is largely lost in such an environment, and the routing system is being faced with increasing overheads in its efforts to converge the distributed algorithm to a stable state.

IMPLICATIONS

If multi-homing becomes a common option of corporate customers, then what is happening is that the function of providing resiliency within a network has shifted from a value-added role within the network to that of a customer responsibility. And if customers are not prepared to pay for highly robust network services from any single ISP then there is little economic incentive for any single ISP to spend the additional money to engineer robustness within their service. From that perspective, what the ISP industry appears to be heading into is a case of a somewhat disturbing self-fulfilling prophesy of minimalist network engineering with no margin for error.

But then, as the economists tell us, such are the characteristics of a strongly competitive open commodity market. That's quite a story that lurks behind a simple chart of the size of the BGP routing table.

Expanded Coverage from ISOC

In-depth articles, papers, links and other resources related to this topic are available from the ISOC site at <http://www.isoc.org/issues/003/>.

For More Information

For more references to the routing issues described in this article, see <http://www.telstra.net/ops/bgp>.

Relevant IETF RFCs

Many IETF RFCs pertain to routing and scalability. Visit the RFC Editor page at <http://www.rfc-editor.org/> for more information

About the Author

Geoff Huston is the Chief Scientist, Internet for Telstra. He is also a Trustee Emeritus of the ISOC Board of Trustees and has served a term as Chair of the Board in 1999. He is an active member of the IETF and a member of the Internet Architecture Board. He has also published a number of books, on Quality of Service in IP networks and the ISP Survival Guide.



About the Background Paper Series

Published by: The Internet Society • 1775 Wiehle Ave., Suite 102 • Reston, Virginia 20190 USA • Tel: +1 703 326 9880 • Fax: +1 703 326 9881 4, rue des Falaises • CH-1205 Geneva, Switzerland • Tel: +1 703 326 9880 • Fax: +1 703 326 9881 • Email: info@isoc.org • Web: <http://www.isoc.org/> • *Series Editor:* Lance Laack (lance@isoc.org).