November 2011
Geoff Huston

BGP Growth Revisited

BGP has been toiling away, literally holding the Internet together, for more than two decades, and nothing seems to be falling off the edge of the Internet. As far as we can tell everyone can still see everyone else, and routing appears to be working. So why should we be interested in BGP? One cause for concern is the inexorable growth of the Internet's routing system. Does this constant growth in routing imply that our routing system is growing faster than our capacity to afford ever larger and faster routers, assuming of course that we can keep on building ever larger and faster routers in the first place? Lets take a look at the metrics of growth in BGP.

> I last looked at the state of BGP in March 2009, and I have received a number of queries about what has happened in BGP since then. Here is an update of that article, reflecting data up to late 2011.

Here's some possible reasons why BGP data can be useful for folk in Internet business.

For the ISP network operator, this information may be help in figuring out how big a router should you buy today if you want it to cope with the full BGP routing load in 3 - 5 years time. Perhaps you might want to understand what FIB size is necessary in that time, and what level of TCAM size might be appropriate, in which case you may want to have a conservative estimate of the anticipated number of entries in the routing table over that period.

The same consideration applies to a vendor of routing equipment: How big a router should a vendor build to cope with the BGP load over the next 3 - 5 years? What growth factors for the routing system should be added into the product design phase? What are the Internet's scaling factors at play here?

Underlying these questions are a more basic set of questions about BGP itself. Is BGP scaling or failing? Do we need to develop a new Inter-Domain Routing protocol to take over from BGP? If BGP is failing, when will this failure become critical? How much time do we have before a new approach is needed? And if we are going to head down this path of attempting to slide in an entirely new routing protocol, is the routing problem simply one of routing over an ever larger and more diverse population, or is this an expression of a more fundamental scaling limitation of the Internet's current concepts of names and addresses? In other words, if we are facing a major problem with routing scalability do we need now to examine alternate architectural models of identity and location separation in order to build truly massive and highly diverse networks? Or, is routing scaling an intractable problem within the confines of the current architecture and we need to shift around the basic building blocks of the Internet architecture in order to allow a different routing architecture that has radically different scaling properties? Or, on the other hand, is BGP coping just fine and there is no current expectation of its imminent demise?

These questions were last studied by the Internet Architecture Board at its workshop in October 2006, which was written up as RFC4984, and projections of routing table inflation in the coming years were a source of considerable concern at the time:

> The workshop participants believe that **routing scalability is the most important problem facing the Internet today and must be solved**, although the time frame in which these problems need solutions was not directly specified. The routing scalability problem includes the size of the DFZ RIB and FIB, the implications of the growth of the RIB and FIB on routing convergence times, and the cost, power (and hence, heat dissipation) and ASIC real estate requirements of core router hardware.
>
> It is commonly believed that the IPv4 RIB growth has been constrained by the limited IPv4 address space. However, even under this constraint, the DFZ IPv4 RIB has been growing at what appears to be an accelerating rate [DFZ]. Given that the IPv6 routing architecture is the same as the IPv4 architecture (with substantially larger address space), if/when IPv6 becomes widely deployed, it is natural to predict that routing table growth for IPv6 will only exacerbate the situation.
>
> RFC4984: Report from the IAB Workshop on Routing and Addressing, September 2007

At the time the picture was not looking overly optimistic for the longer term prospects of BGP, and the workshop prompted further studies of routing techniques and architectures that were capable of sustaining a greater level of information aggregation.

> First of all, the workshop participants would like to reiterate the importance of solving the routing scalability problem. They noted that the concern over the scalability and flexibility of the routing and addressing system has been with us for a very long time, and the current growth rate of the DFZ RIB is exceeding our ability to engineer the routing infrastructure in an economically feasible way. We need to start developing a long-term solution that can last for the foreseeable future.
>
> RFC4984: Report from the IAB Workshop on Routing and Addressing, September 2007

But I wasn't going to head in that direction of describing the areas of possible direction for future routing systems in this article. The question I'd like to ask here is somewhat more pragmatic in nature: has anything changed in this perspective on BGP? Are the prospects of the medium term collapse of BGP through scaling overload still a realistic option for the routing environment? Should we still be concerned about routing scaling? Is the BGP sky about to fall on our heads?

> I can't let that pass without at least a passing reference to Asterix and the chief of the Gaulish village, Abraracourcix (or Vitalstatistix in the English version, if you prefer), who had nothing to fear except for the sky falling on his head!
> <http://en.wikipedia.org/wiki/Recurring_characters_in_Asterix#Vitalstatistix>

The BGP Measurement Environment

In trying to analyse long baseline data series the ideal approach is to keep as much of the local data gathering environment as stable as possible. In this way the changes that occur in the collected data reflect changes in the larger environment, as distinct from changes as a result of changes in the local configuration of the data collection equipment.

In this case the measurement point being used is a BGP router configured as AS131072. This AS generates no traffic and originates no routes in BGP. It's a passive measurement point that has been logging all received BGP updates since 1 July 2007, and is the successor to an earlier setup located in AS1221. The router is fed with a default-free eBGP feed from AS 4608, which is the APNIC network located in Australia, and AS 4777, which is the APNIC network located in Japan. For IPv6 routes the measurement system is being fed with complete route sets from AS4608, AS4777, and AS5539 (SpaceNet)[1].

There is also no iBGP in this particular measurement setup. While it has been asserted at various times that iBGP is a major contributor to BGP scalability concerns in BGP, the consideration here in trying to objectively measure this assertion is that there is no "standard" iBGP configuration, and each network has its own rather unique configuration of Route Reflectors and iBGP peers. This makes it hard to generate a "typical" iBGP load profile, let alone analyse the general trends in iBGP update loads over time.

In this study the scope of attention is limited to a simple eBGP configuration that is likely to be found at a "stub" AS at the edge of the Internet. This AS is not an upstream for any third party, it has no transit role, and does not have a large set of BGP peers. It's a simple view of the routing world that I see when I sit at an edge of the Internet.

The measurement system that I'm using takes a snapshot of the BGP RIB every hour, as well as logging all received BGP updates.

The Data

IPv4 Routing Table

The following tables show some of the vital statistics for IPv4 in BGP since the start of 2009.

Figure 1 shows the total number of routes in the routing table over this period. This is a classic "up and to the right" Internet curve, but it should be noted that the days of the strong "J curve" that indicated growth at an exponential rate that doubled every year, or even growth that doubled every few months, are over. The growth elements in the Internet today are more strongly aligned to a far more modest linear growth model.

There are a few discontinuities in the plot, coinciding with efforts by a number of ISPs to aggregate their routing advertisements into a smaller advertised set of routes. Such efforts

[1] My thanks to Gert Doring of SpaceNet for this willingness to feed me an IPv6 eBGP session for this work.

are notable positive steps in alleviating the constant pressure of growth in the routing system, but these aggregation measures are not hugely significant, in that they have pushed back the overall timetable of routing growth by a small number of weeks at best.

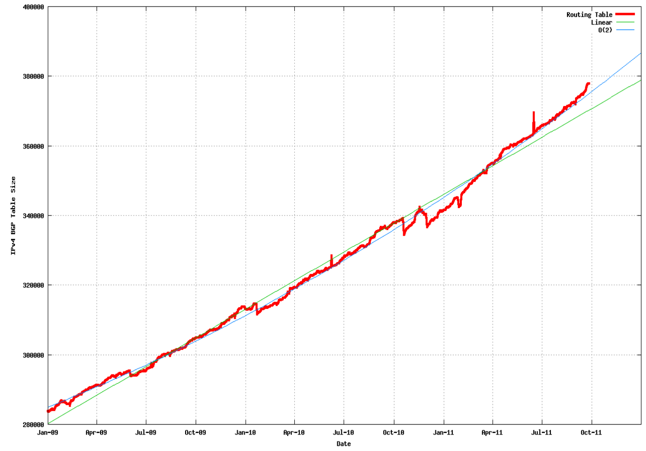

Figure 1 - IPv4 BGP Routing Table Size (RIB)

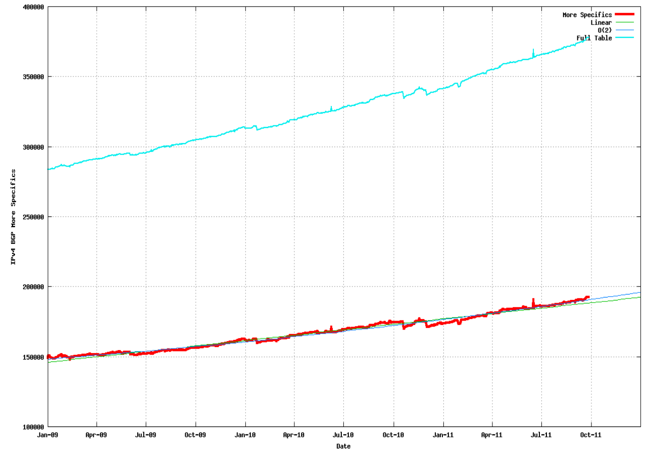

Figure 2 - IPv4 More Specific Entries

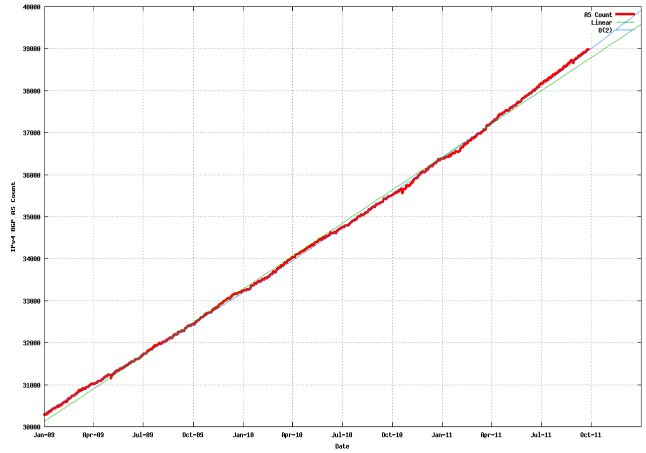

Figure 3 - IPv4 AS Count

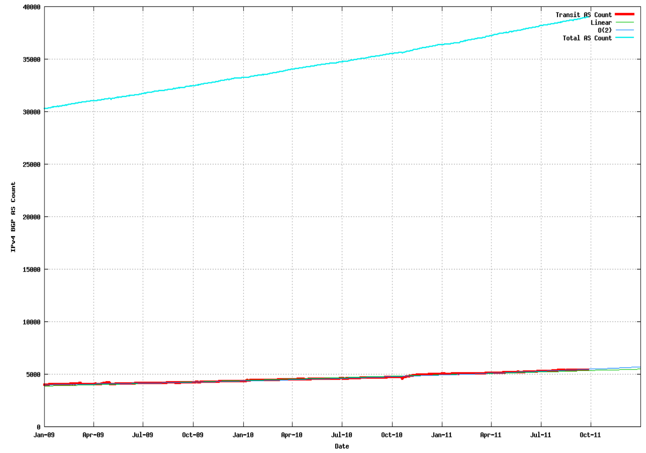

Figure 4 - IPv4 Transit AS Count

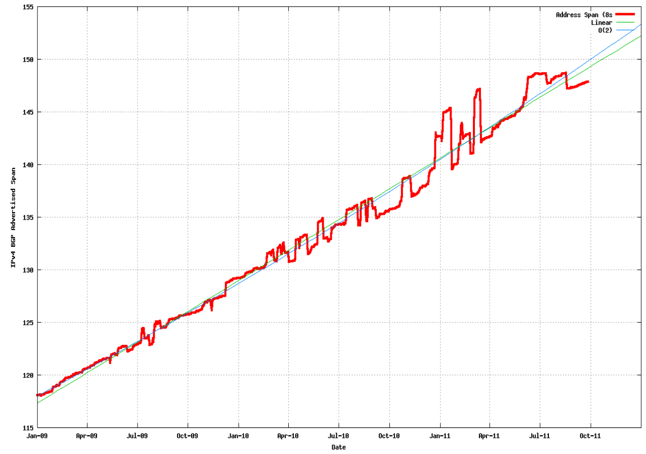

Figure 5 - IPv4 Advertised Address Space

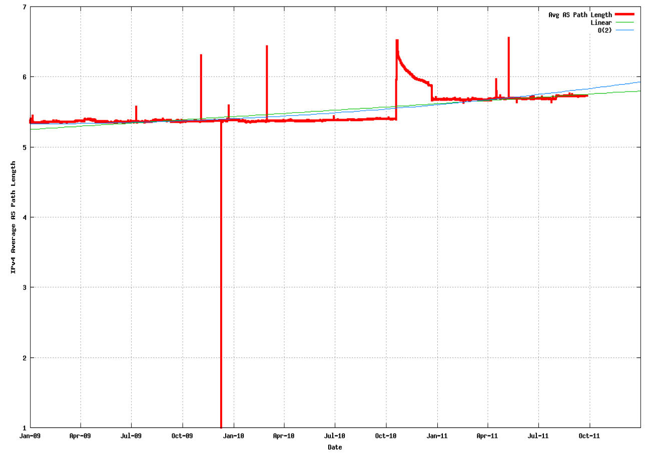

Figure 6 - IPv4 Average AS Path Length

However, buried in Figure 1 is one significant event, namely the exhaustion of the IPv4 address space pools in APNIC, serving the Asia Pacific region, in the last week of April 2011. This event triggered an immediate change in the growth trends of the routing table. However the reaction to exhaustion was not an immediate upward spike in the number of routing table entries, as some folk have been expecting. The reaction has been a shift to a suppressed growth model (see Figure 7). Its likely that the interruption to the steady inflow of new address allocations in the Asian region has been partly responsible for a dropping of the rate of growth of the routing system in the remainder of 2011. It is also appropriate to note that in the same period we have witnessed the second round of a major financial recession in many parts of the world, and the continued suppression of routing growth in the latter part of 2011 may have as much to do with widespread adverse economic conditions in many parts of the world as it has to do with the change of the IPv4 Internet growth model in the Asia Pacific region due to the IPv4 address run out.

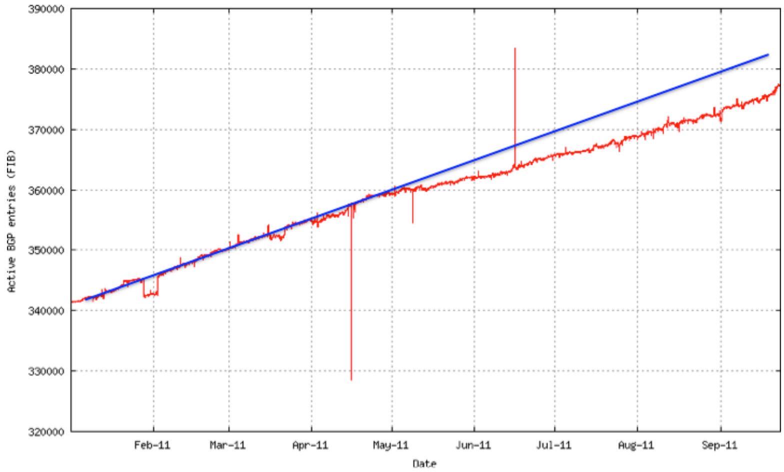

Figure 7 – IPv4 BGP Table in 2011

The summary of the IPv4 BGP network over the 2009-2011 period is shown in Table 1.

	Jan-09	Jan-10	Jan-11	Oct-11	2010 growth	*2011 growth(*)*
Prefix Count	284,000	313,000	341,000	379,000	9%	*14%*
Root Prefixes	135,000	152,000	168,000	185,000	10%	*13%*
More Specifics	150,000	162,000	173,000	192,000	7%	*15%*
Address Span (/8s)	118.0	128.9	140.4	147.9	9%	*7%*
AS Count	30,200	33,200	36,400	39,000	10%	*10%*
Transit AS Count	4,000	4,350	5,050	5,390	16%	*9%*
Stub AS Count	26,200	28,850	31,350	33,610	9%	*10%*

- The January-September data has been adjusted to an annual rate using pro rata projection

Table 1 – IPv4 BGP Table Growth Profile

What this table indicates is that the IPv4 network continues to grow steadily at around 10%, and has done since 2009.

Other observations about the trends of the IPv4 BGP routing system include noting that the use (or abuse!) of more specifics in the routing system has not improved, nor has it become significantly worse. The average size of advertisements is slowly getting smaller in terms of address span per routing table entry, as is the span of originating addresses per origin AS. The average AS path length is constant at around 5 AS hops (which would translate to 4 AS hops if the measurement setup overhead was removed). The number of AS's is increasing at a very regular pace, and the interconnection degree of AS's is getting higher. The implication of this collection of observations is that the granularity of the IPv4 inter-domain routing system continues to get finer, albeit more slowly than in the past, and the density of interconnection continues to increase. The growth of the Internet is not "growth at the edge" as the network is not getting any larger in terms of average AS path change. Instead, the growth is happening by increasing the density of the network by attaching new networks into the existing transit structure and peering at established exchange points. This makes for a network whose diameter, measured in AS hops, is essentially static, yet whose density, measured in terms of prefix count, AS interconnectivity and AS Path diversity, continues to increase. This denser mesh of interconnectivity could be potentially problematical in terms of convergence times if the BGP routing system used a dense mesh of peer connectivity, but the topology of the network continues along a clustered hub and spoke model, where a small number of transit ASs directly service a large number of stub edge networks.

IPv6 BGP Table Data

A similar exercise has been undertaken for IPv6 routing data, and the comparable figures are shown below.

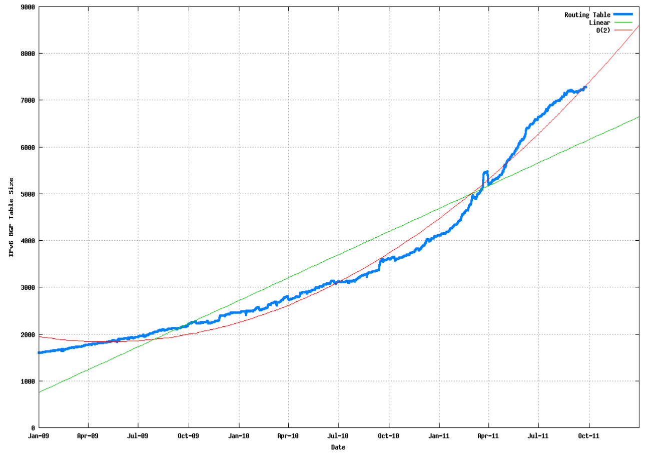

Figure 8 - IPv6 BGP Routing Table Size (RIB)

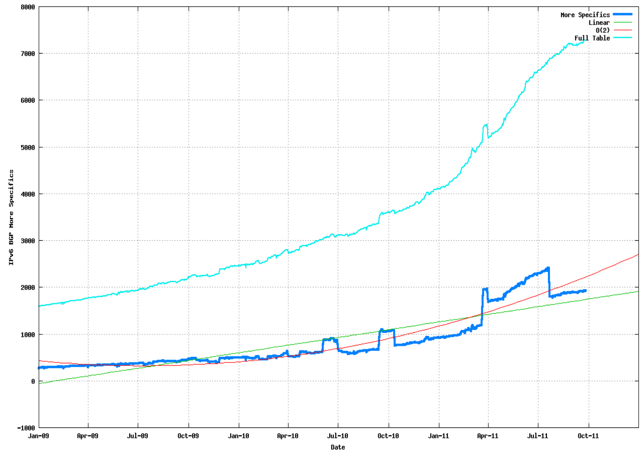

Figure 9 - IPv6 More Specific Entries

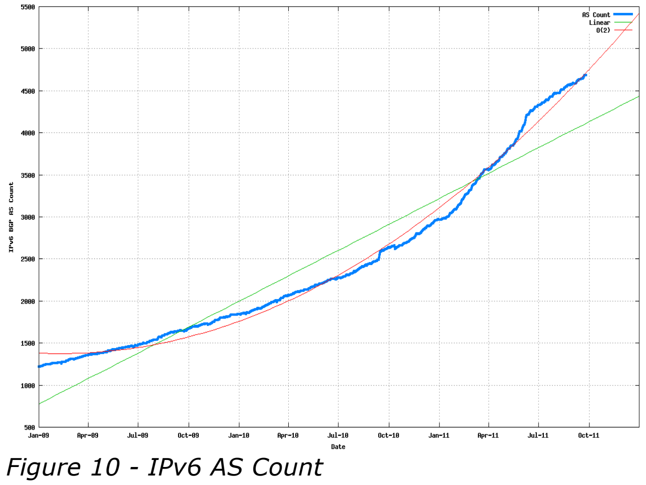

Figure 10 - IPv6 AS Count

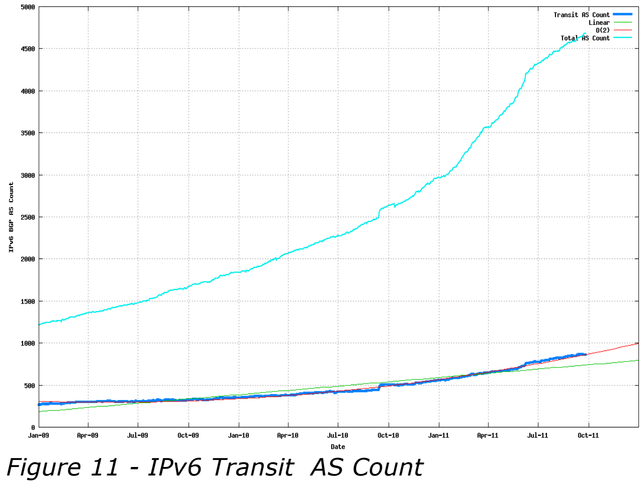

Figure 11 - IPv6 Transit AS Count

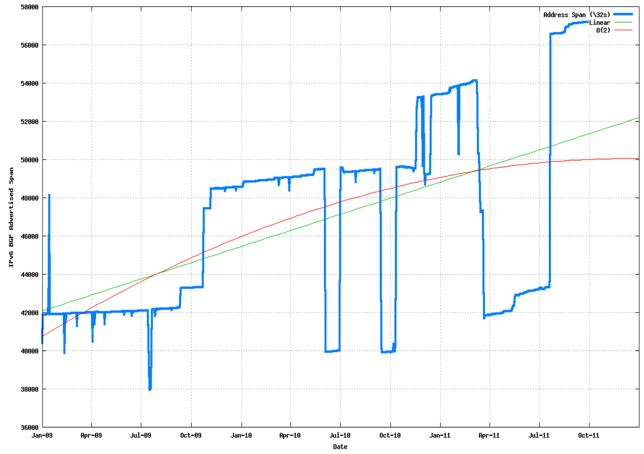

Figure 12 - IPv6 Advertised Address Space

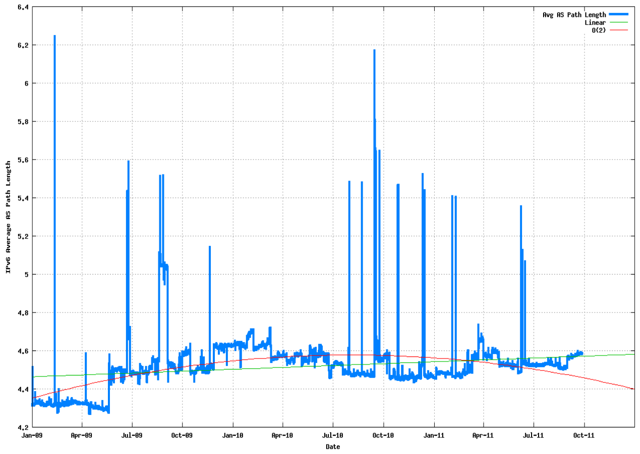

Figure 13 - IPv6 Average AS Path Length

The summary of the IPv6 Internet for the period 2009-2011 is as follows:

	Jan-09	Jan-10	Jan-11	Oct-11	2010 growth	*2011 growth(*)*
Prefix Count	1,600	2,400	4,100	7,300	70%	*104%*
Root Prefixes	1,300	2,000	3,100	5,400	55%	*99%*
More Specifics	300	400	1,000	1,900	250%	*120%*
Address Span (/32s)	40,400	48,500	53,400	57,200	10%	*12%*
AS Count	1,200	1,800	3,000	4,700	66%	*74%*
Transit AS Count	260	350	560	860	60%	*70%*
Stub AS Count	940	1,450	2,440	3,840	68%	*75%*

- The January-September data has been adjusted to an annual rate using pro rata projection

Table 2 – IPv6 BGP Table Growth Profile

Here are those "J-curves" that have disappeared from the IPv4 network, particularly when looking at the growth of the number of entries in the IPv6 routing table and the number of Autonomous System numbers. The rate of growth in the IPv6 network is currently exponential growth with a doubling interval of every 12 – 14 months. There are a number of interesting aspects to this growth where the characteristics of IPv4 look to be appearing in IPv6. The number of more specific advertisements of existing aggregate announcements is rising faster than the number of aggregate announcements. More specifics are now 25% of the total IPv6 routing table. Also the number of transit AS networks is falling, and its now 18% of all AS's in IPv6, as compared to 13% in IPv4. However there is one major aspect of 2011 that is the major event in the IPv6 routing year so far.

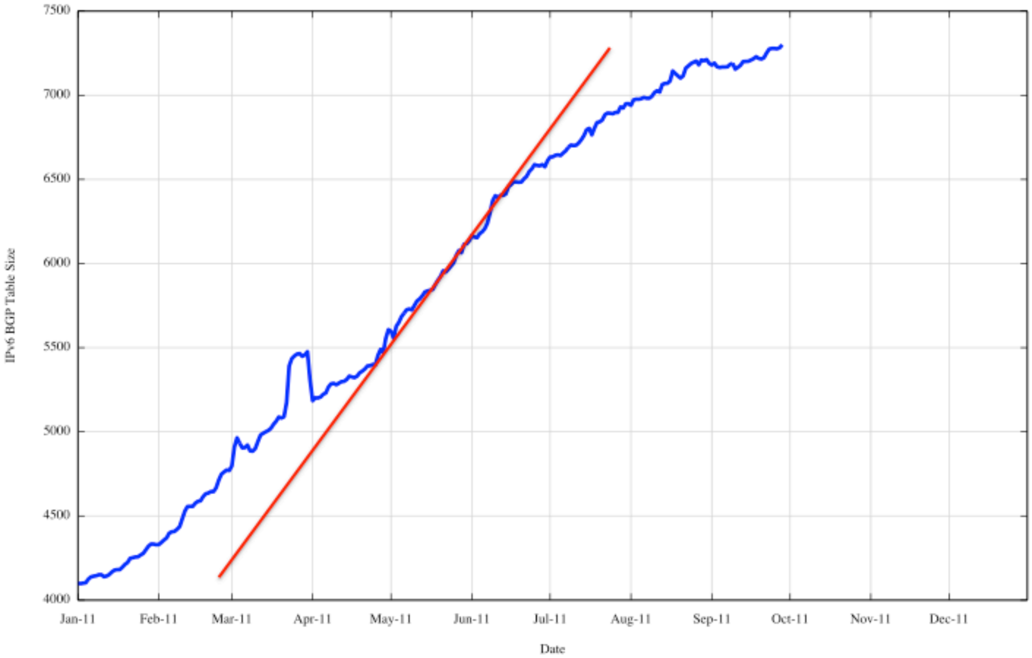

Figure 14 – IPv6 Table Growth in 2011

There is one 6 week period where the growth rates were substantially higher than the rest of the year, namely the period from the end of April 2011 until the first week of June 2011. This period starts with the exhaustion of the remaining IPv4 address pool in APNIC, and ends with World IPv6 day (Figures 14 and 15). These events, particularly World IPv6 data provided a focal point for IPv6 efforts in the year, and these figures show that these events had a visible impact in the routing metrics of IPv6 in terms of IPv6 takeup.

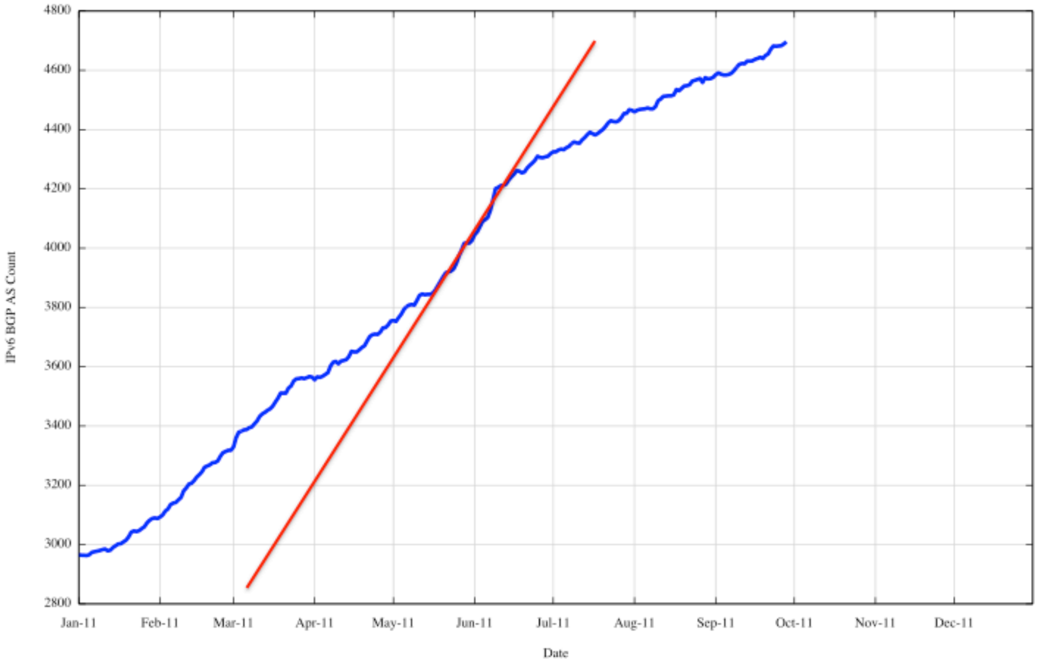

Figure 15 – IPv6 AS Growth in 2011

The Predictions

What can this data tell us in terms of projections of the future of BGP in terms of BGP table size?

At the outset it's appropriate to observe that this is a time of extreme uncertainty in the BGP prediction business! The run out of the Ipv4 address pool in the Asia Pacific region will be followed by a similar run down in the European and the Middle East anticipated to occur in the first half of 2012, and the run down in North America will follow some time thereafter. What was observed in the Asia Pacific region was an initial dampening on the growth in the address table, but there is some expectation that once an after-market picks up momentum there will be movement of address fragments into the routing table, with the potential to cause significant growth in the number of entries in the routing table. Quantifying this is extremely difficult, as this is a case where the past is a poor indicator of the future.

So with the caveat that we are now heading deep into highly speculative areas, and the associated warning that the predictions being made here come with a very high level of uncertainty, lets look at the predictions for the Internet's routing system for the coming few years.

Forecasting the IPv4 BGP Table

The technique used here is to use the least squares curve fitting approach to the historical data (http://en.wikipedia.org/wiki/Least_squares). The data used here is the sequence of daily averages of the 24 individual data measurements that have been made every hour. Figure 16 shows the data set for BGP from the 1st January 2004 until October 2011.

The first order differential, or the rate of growth, of the BGP routing table is shown in Figure 17. This figure is perhaps the most obvious illustration of the correlation between the growth in the Internet, as measured by the growth in the Internet's routing table, and the broader economic conditions in the world. The tend of growth from 2004 showed a sharp decline in late 2008, and the subsequent recovery in early 2009 established a new trend of growth which has still not yet recovered to the levels seen in mid 2007 and 2008.

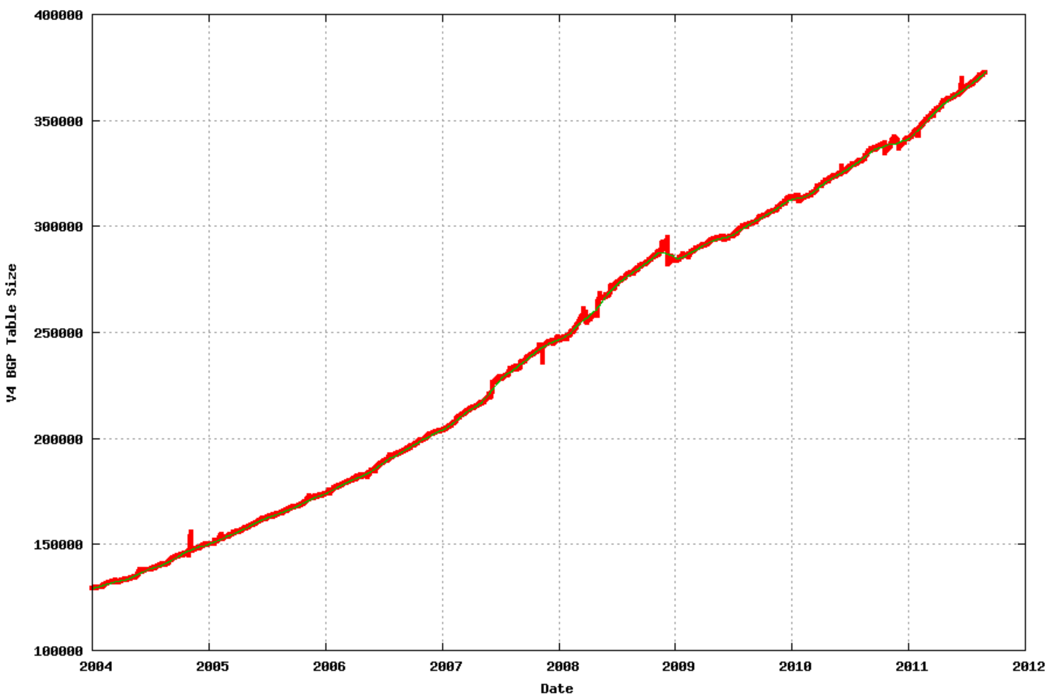

Figure 16 - IPv4 BGP Table Size from 1 January 2004

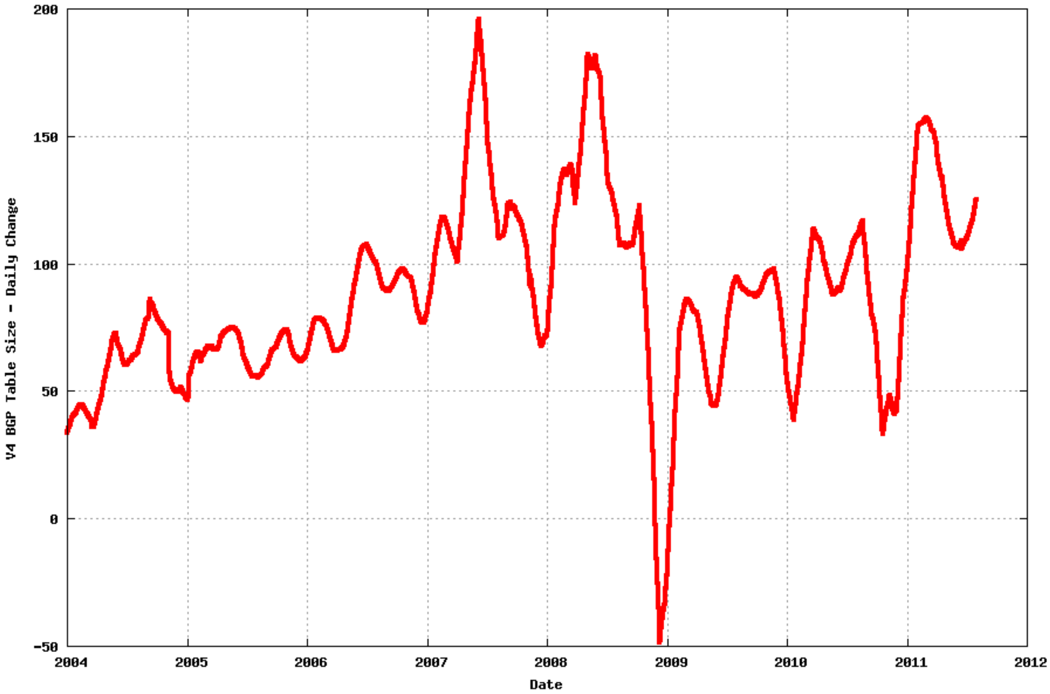

Figure 17 -First Order Differential of Smoothed IPv4 BGP Table Size

The longer term trend of the routing table growth is modelled by a least squares polynomial fit to the data, which generates the quadratic equation

$$y = 496\, x^2 - 1959364\, x + 1934188362$$

Figure 18 shows the fit of this model against the BGP table size data.

The predictions of BGP table size using this model are shown in Table 3.

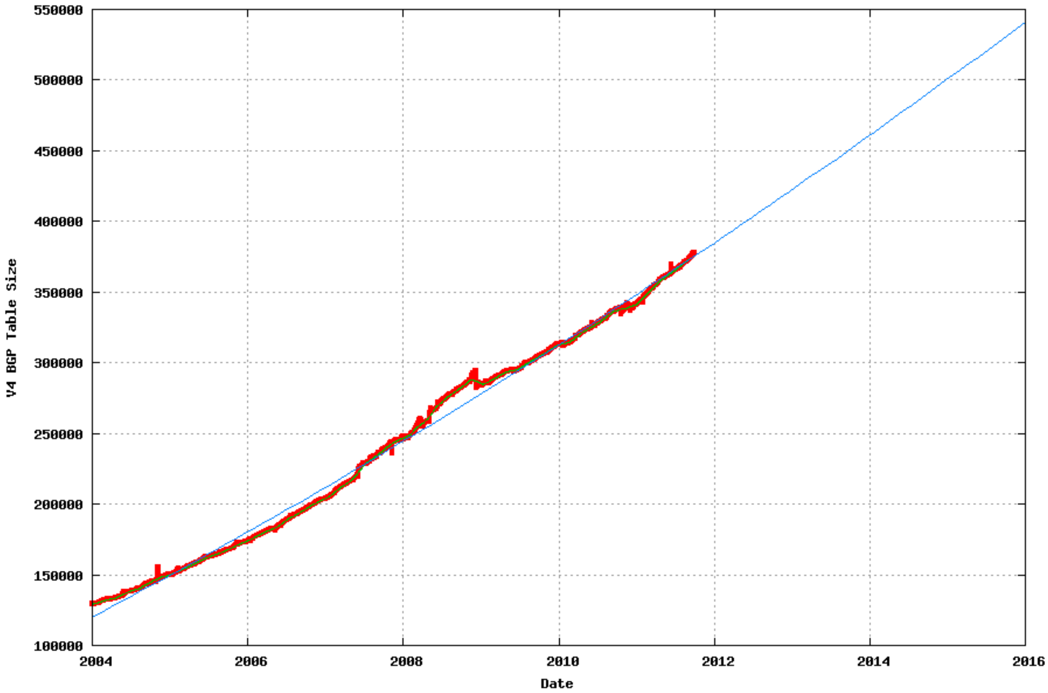

Figure 18 -Prediction of IPv4 BGP Table Size

	IPv4 Table	IPv4 Prediction
Jan 2009	284,194	
Jan 2010	312,989	
Jan 2011	341,483	
Jan 2012		384,857
Jan 2013		422,438
Jan 2014		461,020
Jan 2015		500,595
Jan 2016		541,161

Table 3 – IPv4 BGP Table Size Prediction

With the caveat that this prediction is based on the assumption that tomorrow is a lot like today and that the influences that shape tomorrow have already shaped today, then its reasonable to predict that the routing table in a little over three years time, at the start of 2015, will contain an additional 120,000 entries, making a total for IPv4 of some 500,000 entries in the BGP routing table at that time.

However I'm not confident in the predictions generated by this model. It is simply not possible to use the current models of BGP growth to peer into this post-exhaustion IPv4 routing environment, so the numbers given in Table 3 are extremely uncertain.

Perhaps there is another way of looking at this. If one assumes that the major objective here is to ensure that the "unit cost" of routing continues to decline over time, or at least remain constant, what benchmark could be used to compare the BGP prediction against in terms of a constant unit cost curve?

One possible model that could be used as a benchmark of a prediction of constant unit cost in terms of this form of routing and packet forwarding hardware in packet networks is Moore's Law <http://en.wikipedia.org/wiki/Moore%27s_law>. Here the general assumption is that as long as the growth parameters of the routing table sit within the parameters of Moore's law then the expectation is that the unit cost of routing and switching hardware should not escalate to any appreciable extent. The following figure compares the quadratic projection model of the size of the BGP default free zone with an exponential model of doubling every two years, as used in Moore's Law. As can be seen in the figure below there is no real cause for alarm at this stage, and the BGP table size appears to fit comfortably within these parameters within the current projection model.

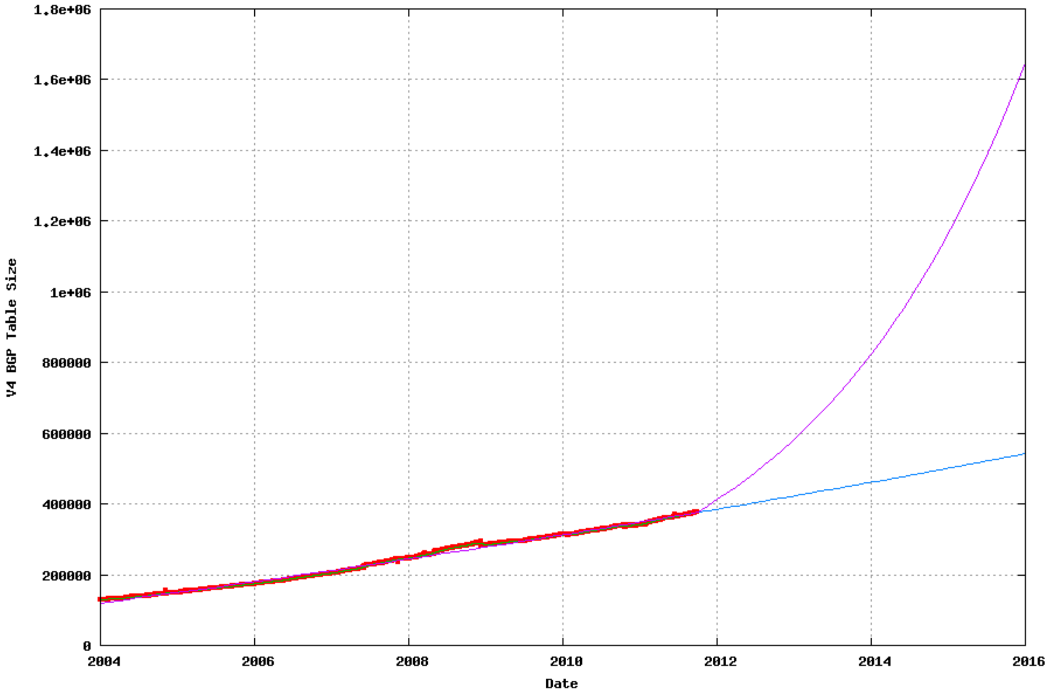

Figure 19 -Comparison of BGP RIB prediction to Moore's Law Growth

Of course, if address exhaustion causes a rapid doubling on the routing table across 2012, inflating the routing table by early 2014 to a size of 2 million entries or more, then this would represent a somewhat different scenario. What could potentially drive this rapid inflation scenario is some form of IPv4 address redistribution function that is focussed solely on the public addressing requirements in IPv4 NAT scenarios, probably increasing the prevalence of global routes at the /24 level, or potentially at even smaller prefix sizes. Such scenarios are related to the levels of speculation concerning the industry reaction to the exhaustion of the existing mechanism of IPv4 address distribution, and at this point in time the predictions about the nature of the redistribution function and the pressures placed on the routing space in consequence are highly speculative.

Forecasting the IPv6 BGP Table

The same technique can be used for the Ipv6 routing table. Figure 20 shows the data set for BGP from the 1st January 2004 until October 2011.

The first order differential, or the rate of growth of the IPv6 BGP routing table is shown in Figure 21. The growth levels in IPv6 have been very low for the period 2004 though to the start of 2009, with average routing table growth levels of no more than 1 entry per day for much of that period. The picture for IPv6 started to change in 2008, when average daily growth levels in the size of the IPv6 Internet started to grow by an average of up to 2 entries per day. This has rapidly increased since then, with the visible characteristics of an exponential growth curve. In 2011 there were two notable periods when the IPv6 table grew from an average of 8 new entries per day to in excess of 14 new entries per day, with the first being in February 2011, coinciding with the widespread reporting of the exhaustion of the remaining IPv4 IANA address pools, and the second being the period across May 2011, starting with the exhaustion of the APNIC IPv4 address pool in late April and ending with World IPv6 Day on the 8th June 2011.

However to put this into some relative perspective with the IPv4 routing table, at the start of 2011 when the IPv6 routing table growth rate peaked at 14 new routing entries per day, the IPv4 routing table was experiencing a peak of some 150 new routing table entries per day.

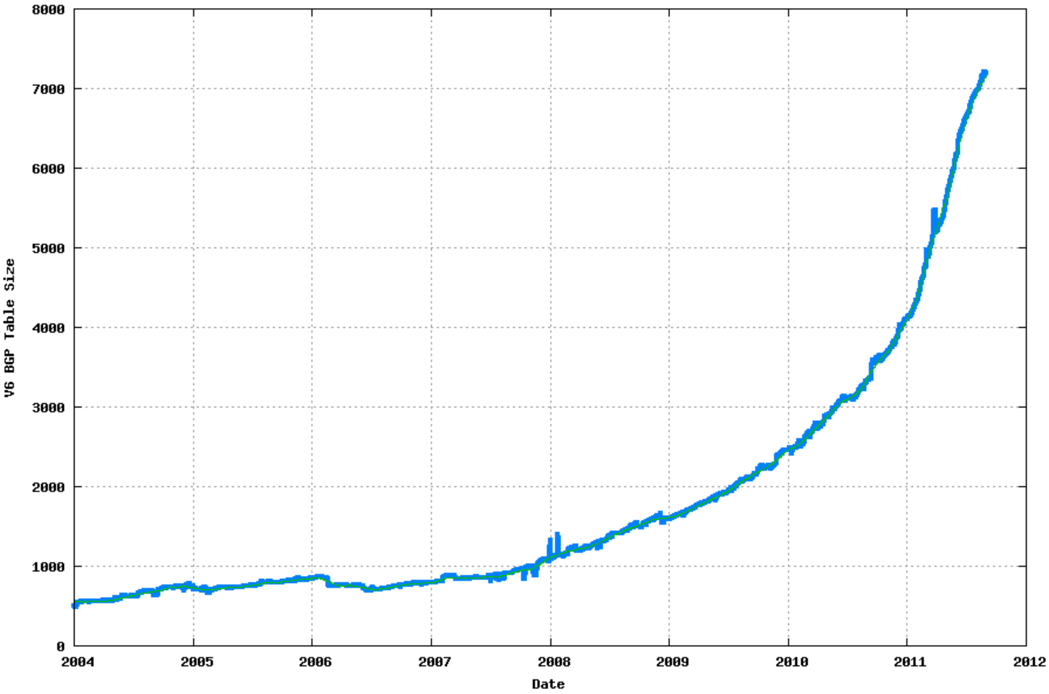

Figure 20 – IPv6 BGP Table Size from 1 January 2004

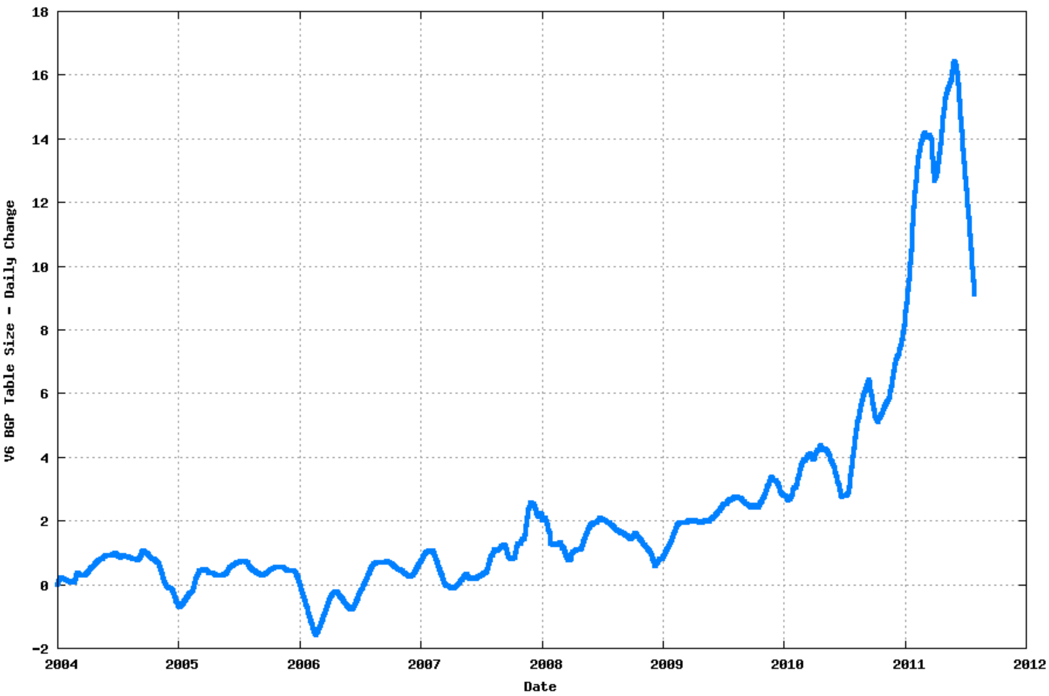

Figure 21 -First Order Differential of Smoothed IPv6 BGP Table Size

The curve-fitting exercise for IPv6 uses the data set commencing 1 January 2009, as this period shows a consistent growth characteristic. Figure 22 shows two curve fitting approaches, one using a quadratic best fit (which assumes a constantly increasing first order differential), and the other using an exponential curve as a best fit (which assumes an exponentially increasing first order differential).

The equations are:

$$y = 1415\,x^2 - 5688923\,x + 5716440317$$

$$y = e^{0.562\,x - 1120.903}$$

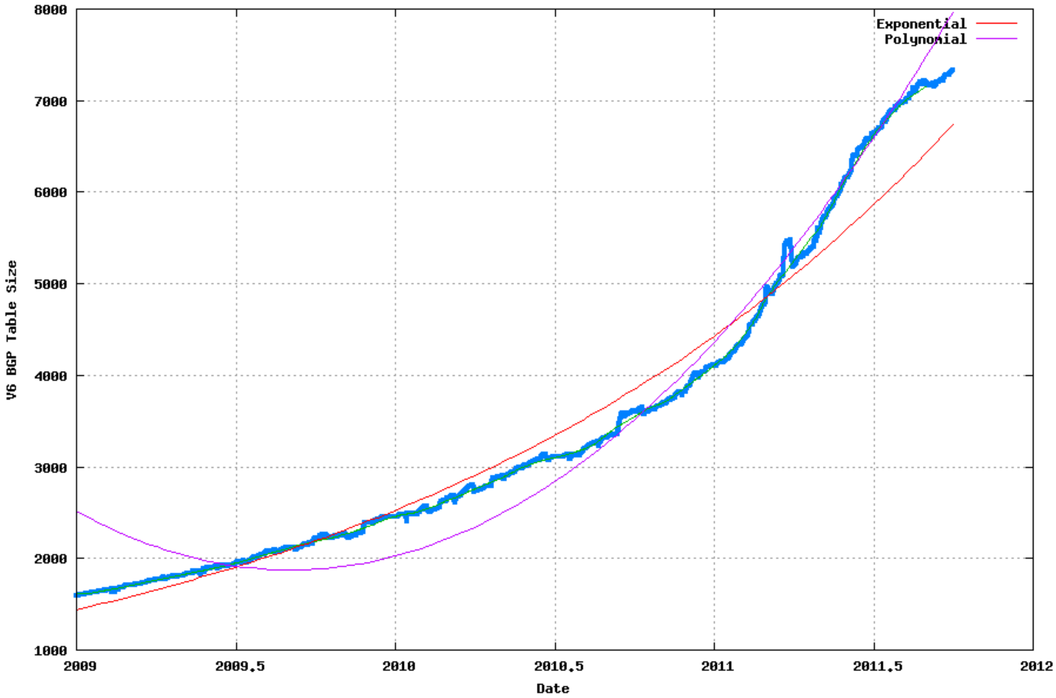

Figure 22 – Curve fit to IPv6 BGP Table Growth

The polynomial provides a tolerably good fit across the period from August 2010 to May 2011, while the exponential curve provides a better fit across the entire period from January 2009, with the caveat that it is underestimating the growth period in the April 2011 – May 2011 period.

These two curves are projected through to January 2016 in Figure 23. The two curves show different outcomes, but not only in fine detail. Both curves indicate a current growth momentum for IPv6 such that in 4 years time the IPv6 routing table will still be under 100,000 entries.

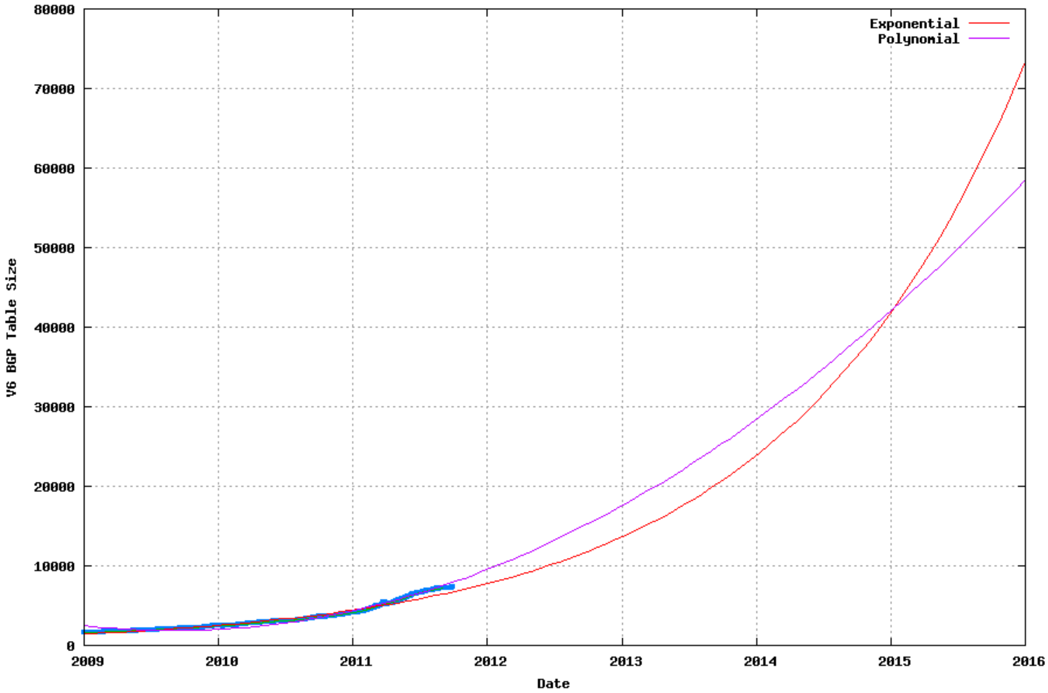

Figure 23 –IPv6 BGP Table Growth projections

The projections for the IPv6 table size are shown in Table 4.

	IPv6 Table	IPv6 Prediction Exponential	IPv6 Prediction Polynomial
Jan 2009	1,596		
Jan 2010	2,458		
Jan 2011	4,100		
Jan 2012		7,665	9,536
Jan 2013		13,614	17,538
Jan 2014		23,871	28,370
Jan 2015		41,857	42,034
Jan 2016		73,392	58,528

Table 4 – IPv4 BGP Table Size Prediction

The comparison of these curves to Moore's law is show in Figure 24.

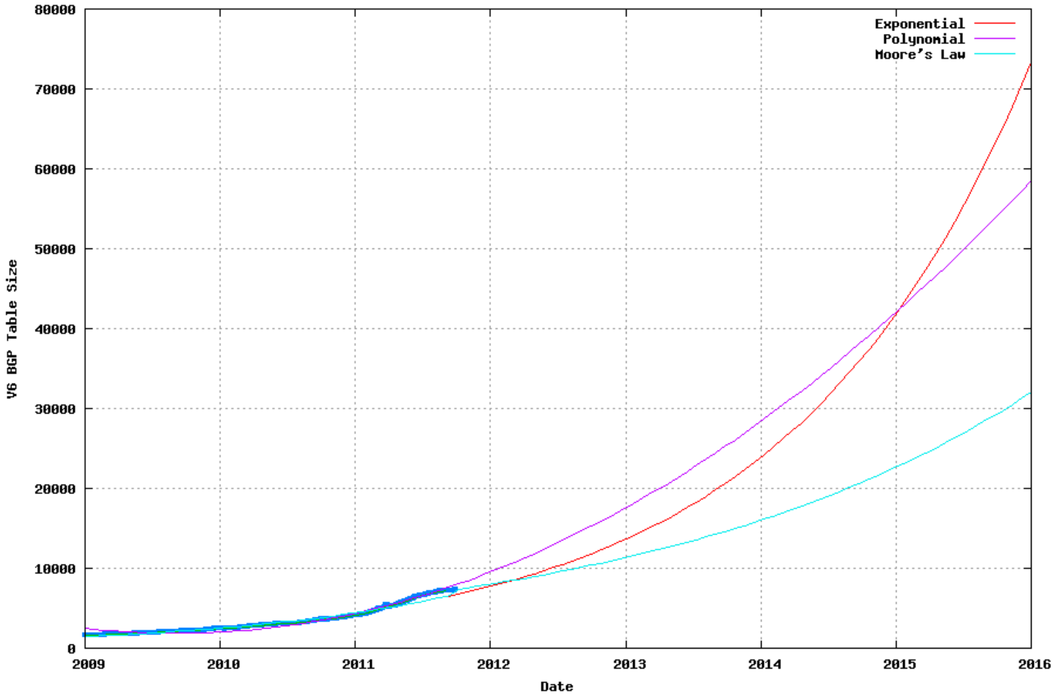

Figure 24 –IPv6 BGP Table Growth projections and Moree's Law Projection

At this stage it is evident that the recent growth in the IPv6 routing table is one that doubles the size of the table every 13 months or so, and this rate is significantly faster than Moore's Law, which doubles every 24 months. However at this stage the numbers are still small. Today's IPv6 routing table is 1/50 of the size of the IPv4 routing table, and over the next four years these figures indicate that it will get to 1/5 of the current size of the IPv4 routing table.

Conclusions

These predictions are highly uncertain. Despite this uncertainty, nothing in these figures indicates any serious cause for alarm at present. There is no clear signal of the need to invent a new inter-domain routing protocol in the near future, assuming that the motivation to do so was based on a reasonable prediction of the imminent collapse of BGP.

None of the metrics indicate that we are seeing such an explosive level of growth in the routing system that it will fundamentally alter the viability of carrying a complete eBGP routing table in the near future. In terms of the table size in the IPv4 and IPv6 networks, the BGP sky is not about to fall on our heads any time soon.

But this is only half the story of BGP growth. The other half of the story is concerned with the characteristics of the dynamics of the BGP protocol, and the rate of growth of updates in BGP. We'll look at that in next month's column.

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. 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 was a member of the Internet Architecture Board from 1999 until 2005, and served on the Board of Trustees of the Internet Society from 1992 until 2001.

www.potaroo.net