

IPV6 Supply and Demand

Jan 7, 2013

Charles Prince

Computer Science Dept.

Naval Postgraduate School

Why is IPV6 important

- We are, or have already run out of IPV4 addresses to assign so we need an alternative
- An alternative for supply of web sites is for web sites to use one host with many sites (available for HTTP1.1)
- An alternative for web site demand, that is how the users connect to the internet can be by using NAT
- An alternative for both is to use IPV6

Where is IPV6 headed

- Will demand of IPV6 increase greatly now that available addresses of IPV4 have virtually disappeared
- This is important for planning purposes for internet infrastructure
- Perhaps we can look at IPV4 statistics and IPV6 statistics and gauge what demand is of IPV6 compared to IPV4

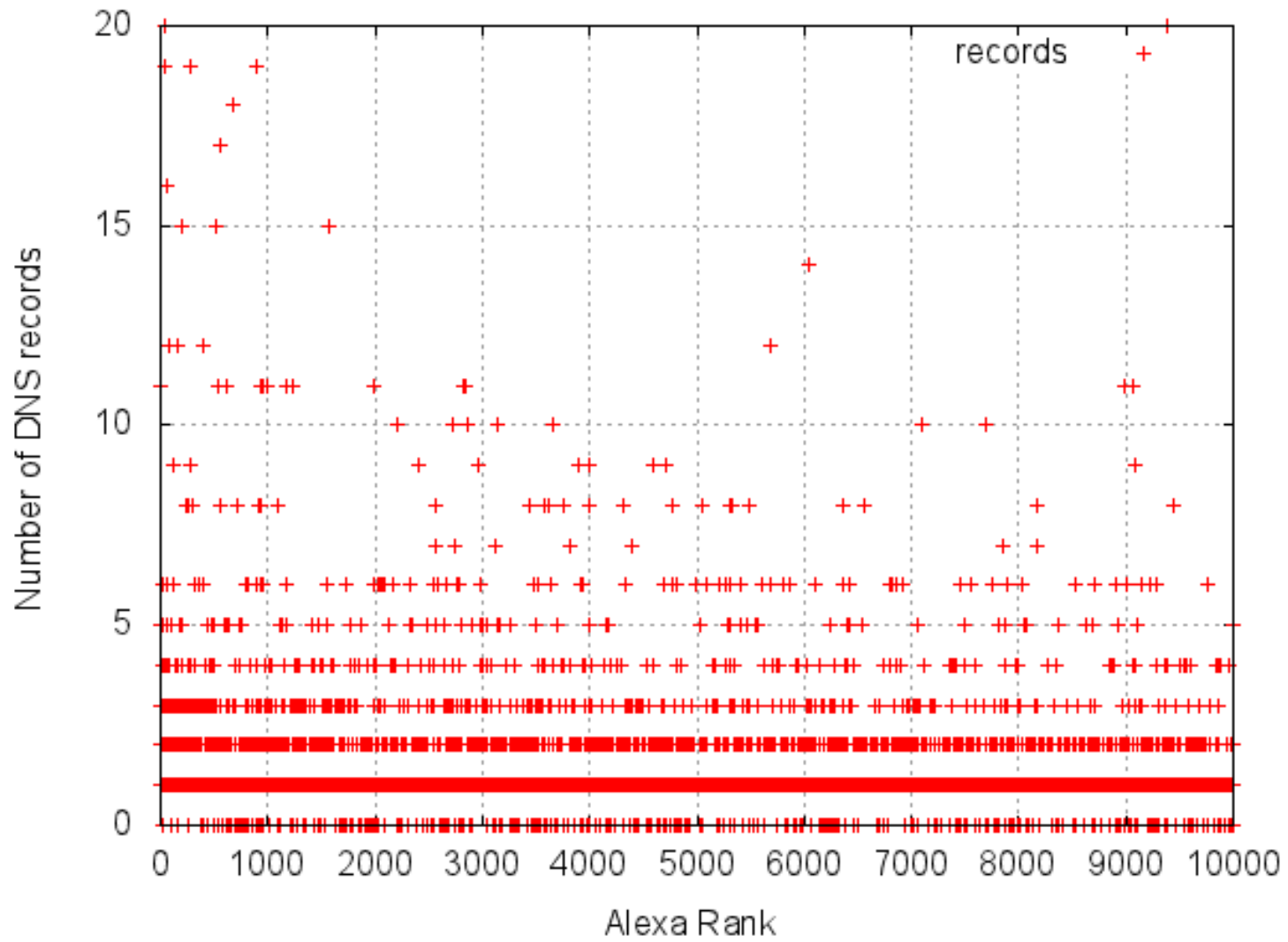
IPV4 Demand

- Do highly requested web sites correlate to having more DNS records?
- Do highly requested web sites correlate to having better connectivity? That is, is latency less for the most popular web sites? Is hop count less for the most popular web sites?
- If true, will it be similar for IPV6?

Test Setup

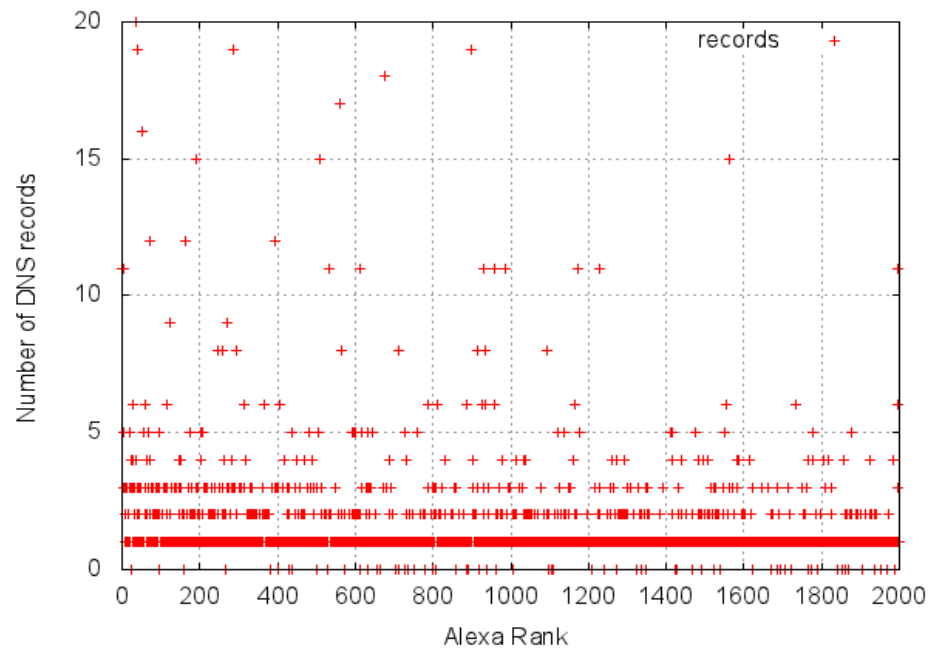
- We used the Alexa 1 million as our starting point to represent the most popular sites
- Using python scripts we collected data on the number of DNS records, hop count, and latency between one vantage point and the web sites for both IPV4 and IPV6
- The data was put into a table and then Pearson correlations were made between the data

Results



Results

- The majority of the data appears to be rather random, but a closer look reveals some interesting insights up to Alexa 2000



Test Results

- The results so far show good correlation about 60% between DNS records of IPV4, or IPV6 sites to Alexa rank for the first 2000 web sites with very solid P-values
- Pearson correlation values for Alexa 1 to 2000 with bin size of 20 and number of values at 100

Alexa 2k-4k	Pearson Correlation	P-Values	# of Values bin=20
IPV4 DNS records	-0.595	$6.6(10)^{-11}$	100
IPV6 DNS records	-0.542	$5.9(10)^{-9}$	100
IPV4 to IPV6	0.626	$3.3(10)^{-12}$	100

Test Results

- Some unexpected results deserve further investigation
 - Alexa from 2k to 10k yields no correlation, while Alexa 10k to 12k shows some correlation

Alexa 2k-4k	Pearson Correlation	P-Values	# of Values bin=20
IPV4 DNS records	0.002	0.97	100
IPV6 DNS records	-0.099	0.32	100
IPV4 to IPV6	0.242	0.015	100

Alexa 10k-12k	Pearson Correlation	P-Values	# of Values bin=20
IPV4 DNS records	-0.200	0.048	100
IPV6 DNS records	0.003	0.96	100
IPV4 to IPV6	0.226	0.024	100

Why These Results

- One possibility for these results is this may be due to economic reasons that basic infrastructure costs are used when web site traffic levels reach a certain threshold seen at around Alexa 2000, above which the DNS records depend more on the volume of traffic
- Another possibility may be due to inaccuracies in Alexa
- An intriguing possibility may be due to geo-location of web site origin
- The results for hop count and latency show little correlation of around 20% we discuss this further in future work

Test Results -- Continued

- From Alexa sites 2000 to 4000 we see no correlation between DNS records for IPV4/6 to Alexa rank
- From Alexa sites 10000 to 12000 we see 20% correlation for IPV4 – this is actively being investigated
- There seems to be very little correlation in IPV4 latency to Alexa rank for the first 2000 web sites (bin size of 40 sites), but a much higher correlation between IPV6 latency to rank at 17%

Future Work

- The results so far are from just one vantage point and it would be very interesting to get many vantage points. We may find a lot more DNS records out there.
- In addition to the above testing we can try and track the number of IPs that host multiple web sites. This would let us know how much volume the alternative to IPV6 was and if tracked over time, how IPV6 was catching on.
- We also want to determine where in terms of DNS records IPV6 is compared to IPV4 and to see if we can infer a volume of IPV6 relative to IPV4

Future Work

- We would like to determine if we can find out how many single IP multi hosted web sites there are for both IPV4 and IPV6 and track this statistic
- We would also like to follow web site link chains to look up additional URLs to see if we can gather additional DNS records per Alexa site
- Latency and hop count correlation may increase with more vantage points – there could be a correlation between geo-location of users of a site to location of the Alexa web site and site rank

Comments

- We feel that we are probably grossly under representing DNS records at this time