# A Passive System for Server Selection within Mirrored Resource Environments Using AS Path Length Heuristics

*AppliedTheory Communications, Inc.*
*Patrick R. McManus <mcmanus@AppliedTheory.com>*
*Lead Developer, Software Engineering Group*

**Abstract: This work presents the results of an experiment using a heuristic algorithm that is utilized by clients to select the closest available web server from a group of mirrors. The heuristic is based on BGP AS path lengths and can be determined without the introduction of any additional measurement traffic into the network. The results indicate this is a promising strategy for passive determination of good, though not always optimal, servers.**

## Background

Many traditional load balancing solutions distribute load based solely on server side criteria such as current processor load, the number of outstanding queued requests, and network saturation. These schemes fail to consider the client's relationship to the members of the server pool when making the decision of which host to use to process a request. They consider only the state of other potential servers and their environments without consideration of the proximity of the client, which is at a fixed point, to the location of each member of the pool of prospective servers.

This project attempts to separate the concept of proximate server selection from distributed processing of requests within a locally connected server farm. When applied in the context of a network of strongly interconnected servers, the latter is a well understood problem with a number of quality commercial solutions available. However, when attempting to coordinate servers across a wide area network, or even across diverse providers the complexity of the problem increases.

First, the state sharing process between remote hosts must be done within a timed interval that is short enough to make the relevant decisions with. The ability of existing protocols to do this has been asserted but not clearly demonstrated. Additionally, the characteristics of the network link between the client and potential servers are not factored into the decision. The dominant factor impacting end-to-end client perception of performance on a wide are network still likely remains the bandwidth and delay properties of the path between client and server and not the performance characteristics of the server itself.

A system of motor vehicle registries would not be considered efficient if people were traveling to offices 300 miles from their home because another city had queues that on average were 20 minutes shorter than those in their home cities. Only in a system that does not consider the customer transit time does such an algorithm makes sense.

We suggest using traditional load balancing schemes to schedule traffic within distributed server clusters on LANs or on other strongly connected infrastructures. However, when working within environments that are mirrored on diverse networks the decision of which cluster the client should use should be made independently of server selection within that cluster.

Optimal cluster selection would involve that cluster with an ingress path exhibiting sufficient bandwidth, low packet latency, low variation in packet latencies, and low packet loss rates. In cases of short lived sessions, so-called mice flows that are exhibited by HTTP traffic, factors that impact the retransmission timeout value have a greater importance than normal because of the reduced probability of using the fast retransmit algorithm. Both the mean latency and latency variation variables fit this criteria.

Obviously a myriad of factors contribute to these measurements. Physical link distance, link medium, other contention for the medium, hardware speed and reliability, etc… Active measurement of these properties is inappropriate for a high transaction volume context. In high volume circumstances (which correlate strongly to diversely mirrored environments) active probing introduces an unacceptably high amount of traffic into the network.

Recall that the measurement must be from cluster ingress to the fixed client point so results

are not able to be cached for application to more than one client. Even the addition to the connection overhead of at least a single round trip time to determine the server with the lowest latency is not sufficient for approximating the values of the link characteristics such as drop rates, and latency variations. To compile a measurement of all these values a multiple packet test must be conducted for each server. This not only multiples new traffic introduced into the network but it introduces a very high delay at the beginning of a connection while the measurements are taken. A metric that can approximate these measurements either by passive monitoring of already present network conditions, or by the introduction of a fixed amount of traffic that is not linearly related to the number of flows being measured is needed.

As a hypothesis, the number of Autonomous Systems in between a client and server was used as a measurement of the Internet Distance. This hypothesis assumes that the dominant delay and risk incurred in packet processing occurs at exchange points in between service providers. There is intuitive acceptance of that argument based on the fact that exchange points inherently require co-operation between multiple parties with different sets of priorities. This can lead to under-provisioning of bandwidth by some or all of the providers, or under-provisioning of the router or switch itself. Note that even an insufficient allocation of bandwidth by a single provider not directly involved in a particular transaction can have a transitive impact on the transaction if it causes a backlog on the switch or router.

As a contributing factor, interior routing within an AS using a system such as OSPF is much more capable of adapting to an end-to-end concept of routing because of its smaller size and single source administration. Because of this greater agility, backlogs and other shortcomings can be more effectively addressed than at the inter-AS level.

While it was clear from the outset that this scheme would not provide 100% reliable decisions, its non-intrusive nature and positive scaling attributes made it an attractive base system. Because BGP already propagates the information necessary to build this service it can be constructed in a completely passive mode with respect to wide area traffic.

Optimal deployment of this technique is as close to the client as possible. However, it is unrealistic to expect personal workstations to dedicate the resources necessary for holding complete BGP feeds. Additionally, the load on the router supplying that feed would be extraordinary if it needed to maintain a peering session for each client. A more pragmatic approach is to have one or more proxies on the same network as the clients and have all the end user machines share those proxies. This results in just one BGP feed per proxy and meshes well with a highly scalable ISP cascading caching proxy infrastructure.

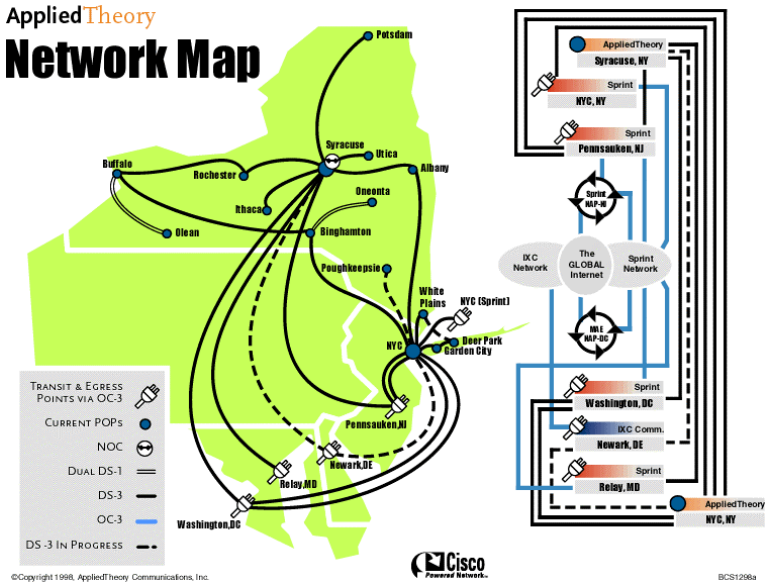
A less attractive, but not infeasible, alternative is to deploy the proximate cluster selection on the sever side. This scheme has the drawback of requiring at least one round trip interaction with a random cluster in order to have the routing decision made. Subsequent to that all subsequent traffic would pass through only the designated proximate cluster.

The server responsible for making the redirection decision needs to implement a protocol to get an Internet Distance measurement to the client from each of the prospective clusters. Design of such a system was out of the scope of this project. The main advantage of this scheme is one of deployment. No cooperation is required from the client side so it may be effectively used as a transition scheme.

## Implementation Procedure

A path length server was created that established itself as a version 4 BGP peer with a cisco router in the same Autonomous System (AS) as the router (i.e. an internal peering session). The router contained the full Internet routing table and the path length server (PLS) built an internal database based on that feed in real time according to RFC 1771.

No routes were ever announced to the router by the PLS. The router (SYR-F1-R2-C7K1-0E0-0.appliedtheory.com) was located on Applied Theory's network in our Syracuse New York Network Operations Center. Egresses from the AppliedTheory network were provisioned in cooperation Sprintlink during this experiment. See figure on next page for network configuration.




The PLS implemented a simple UDP based protocol to provide comparative path length information to clients. This protocol is used to service proximate server resolution requests from clients who present a set of IP addresses to the PLS. The clients could package up to 255 IP addresses in the request datagram and the PLS would return a path length score for each one and also identify which address has the best score which is represented by the lowest path length metric. The PLS resolution protocol is defined separately from this report.

Compliant clients were configured to use a PLS via this protocol. The PLS and client seeking proximate resolution services should be within the same AS, as that is where the measurements are in relationship too. Because the granularity of measurement is based on intervening autonomous systems there is no accuracy advantage to forcing the client and its PLS to be on the same subnet as long as they reside within the same AS. Location of the PLS on an AS directly up or down stream from the client with no alternative path in between would be acceptable as well because this would introduce a constant error into all calculated metrics and would not impact the cluster selection algorithm. This only makes sense to do when the PLS is used on the client side of the configuration.

Identifying what addresses comprise a set of mirrored servers was out of the scope of this project. Any such scheme must be designed to not conflict with the multiple A record DNS round robin scheme currently in use for load balancing. Implementations using LDAP or with a reverse DNS tree such as was done with in-addr.arpa may be good solutions to this problem. Overloading of the key to this service as a legitimate member of the server pool also provides a good fallback mechanism if the directory is unable to resolve the request.

As a final caveat, even a simple metric such as AS Path length becomes complicated with the introduction of CIDR aggregations [RFC1519] [RFC1965]. CIDR breaks an AS Path into a linear path sequence and an unordered set of additional networks aggregated into that announcement that are reachable (but not necessarily directly reachable) by the last AS in the sequence. This scheme has succeeded in its goal of reducing the number of routes that need to be stored in a BGP table, but correspondingly that results in a reduction of useful information for tasks other than routing.

This experiment approximates the length introduced by the unordered set as 1 + ln(set length) AS hops. The derivation of that evaluation is a wholly arbitrary one and demands further study if the project proves useful. This length value is then multiplied by 1000 and any fraction dropped in order to preserve integer arithmetic within the implementation.

## Empirical Data

### Path Length Metric Distribution

Before reaching the stage of being able to evaluate any correlation between our Internet Distance metric and the performance exhibited by the corresponding link an unanticipated trait surfaced in the data obtained from the BGP peering session. 90% to 92% of all route announcements on January 20, 1999 were represented by just 3 distinct path length values.

Lack of diversity in observed values potentially mitigates its value for use as criteria in cluster selection. Full tables detailing the path length value distributions are included at the end of this paper. Further research indicates CAIDA has independently obtained data that confirms this observed distribution [CAIDA98].

The probability of 2 hosts having the same path length measurement is between 30% and 34%. Two values are given because the calculation depends if the probability is based on the distribution of network announcements, or on the size of the blocks represented by those announcements when weighting the results. While being able to differentiate between networks just 2/3 of the time does not render the scheme completely infeasible, it was a discouraging beginning.

The actual cause of this tight grouping has not yet been determined. Potential causes include the lack of differentiation between long haul transit networks containing transcontinental links and regional second or third tier providers. Indeed, when doing some non-scientific browsing of the accumulated data it was observed a German university and a Boston university both received a path length value of 3000 when measured from the New York state network!

However, while an exhaustive survey of locations real life deployments of mirrored servers has not yet been done the anecdotal data collected indicates that there is sufficient diversity in path lengths to use this as a differentiator. This is largely because organizations that deploy mirrored server use several more than just two. See the section titled *Metric Value Considered for Real Mirrors* for the results of path length metric distribution of the current servers of six large organizations.

Internet routing has come to bear a striking resemblance to the theory of *six degrees of separation*. That must mean the current topology of the Internet is of a more multidimensional mesh configuration than a flat puzzle layout that relies heavily on cooperative transit to move long distances. Increased reliance and deployment of private peering points and the move away from traditional first, second, and third tier providers is a likely contributing factor.

### Metric Value by Individual Anecdotes

Individual servers can be found that both substantiate and refute the modeling of Internet Distance using AS path lengths. The important consideration is whether in general the system is capable of selecting a good server from its possible choices even if it does not always choose the optimal one.

Also consider that physical geography is not relevant to the decision making process. For example, in one non-intuitive case path length values are indeed indicative of network response even though this doesn't correspond to physical arrangement. The Microsoft web site in Redmond Washington registers a path length metric of 3000, while a central Massachusetts university, which is an order of magnitude closer as the crow flies, yields a 4693.

Initially this makes little sense, but the Microsoft site is actually more responsive (~64ms RTT compared to ~92ms) due to the well provisioned links in its route. This is consistent with the hypothesis of this experiment that network exchanges cause the most inefficient link traversals.

Short path length values do appear directly correlated to variance in round trip times even when using path lengths that don't correspond to intuitive distances. A good case for comparison would be that of McMaster University located in Hamilton, Ontario (a few hundred miles from the point of measurement) and www.oracle.com which is 3000 miles from the point of measurement.

The Ontario server generates a path length value of 5000, contrasted with Oracle's 3000. But, average RTT (taken from 200 samples at just one point in time) was faster to the Ontario site at 71ms compared to 81ms.

That is what we intuitively expect but it doesn't correspond to the path length values. However, the variance in the RTT measurements has a stronger correlation for this sample. The standard deviation on the Ontario measurements was 16.34, while the standard deviation of the California measurements was just 2.25 Not included in these measurements, but also of interest, was the 3% packet loss to Ontario compared to .5% packet loss to California.

A few other sites were selected for measurement, and they also exhibited this property. Therefore, while raw delay was better for the site with a higher path length metric the recovery properties of the preferred link were much more stable.

This seems to give some weight to the argument that the AS path length of a route is a decent indicator of its reliability and stability, if not its proximity. Given that the RTT variance is a major factor in the calculation of the TCP Retransmission Timeout (RTO) this indicates that path length variances may indeed correspond (at least weakly) to end-to-end performance of paths experiencing even minimal packet loss.

**Metric Value Considered for Large Samples**

The next step is to calculate real link performance data for a large set of Internet servers and to calculate their path length values to test for an average case correlation.

A set of 20,000 servers was used. We would have liked to base server selection on traces from a very busy proxy as that would have reflected a real life usage distribution. However, no suitable log could be located. Instead host names were harvested from a set of major non-juried web crawling search engines. Arbitrary search terms were systematically pulled from a dictionary to try and obtain a uniform distribution of host names.

Of the 20,000 servers only about 17,500 produced data that was usable. Some simply no longer existed, some contained links that filtered ICMP or other measurement traffic that we used to actively probe the network, and some were just unreachable for measurement at the time of the experiment.

For each server the path length metric was looked up and a sequence of 10 ICMP_ECHO_REQUEST packets were sent to it at 1 second intervals. The round trip time of each packet (or the fact of its loss) was recorded. 12 different hosts were being measured in parallel at any given time. The experimental host resided on a quiet 10Mb/sec Ethernet in a hosting facility that presented no transfer bottlenecks with respect to the local loop allocation. All active probes were done using 64 byte packets.

Because the proportion of route announcements with path length metrics that are not divisible by 1000 is relatively small, they are represented by the nearest whole metric in the following data for the sake of clarity.

| Path Length Metric (in thousands) | Mean RTT (in milliseconds) |
|---|---|
| 1 | 116.9 |
| 2 | 137.2 |
| 3 | 164.1 |
| 4 | 209.3 |
| 5 | 296.1 |
| 6 | 349.2 |
| 7 | 479.5 |

This data shows an even stronger correlation between our approximation of Internet Distance and raw packet latency than was expected. Perhaps of the greatest importance are the significant differences between the values for path lengths of 3, 4, and 5. That is because those path length values represent over 90% of the real data space. The fact that the RTT for each of them differs by such a large amount quantifies a tangible benefit to the hypothesis that this is a good heuristic in the average case for real world data with respect to raw latency.

| Path Length Metric (in thousands) | Standard Deviation of packet RTT (ms) |
|---|---|
| 1 | 70.1 |
| 2 | 74.6 |
| 3 | 88.5 |
| 4 | 101.0 |
| 5 | 110.0 |
| 6 | 123.6 |
| 7 | 179.2 |

Variance of round trip times, as represented here by standard deviation, is important to end to end throughput of a connection primarily because of its impact on calculation of retransmission timeouts. It is also important to real time stream

based applications that may not be based on TCP.

This data shows a correlation between path lengths and link instability. We hypothesize that this is because of an associated increase in the average number of routers and switched in the path that have their queueing delays aggregated. Links exhibiting less stable transmission times will take longer to timeout in the case of lost, discarded, and damaged packets that cannot use the fast-retransmit and recovery algorithms.

| Path Length Metric (in thousands) | Approx. Timeout (RTT + 4 SD) in ms |
|---|---|
| 1 | 397.3 |
| 2 | 435.6 |
| 3 | 518.1 |
| 4 | 613.3 |
| 5 | 736.1 |
| 6 | 843.6 |
| 7 | 1196.3 |

Given the significant performance gaps for timeout periods across different path lengths it is reasonable to inspect their respective packet loss rates.

| Path Length Metric (in thousands) | Mean Losses per ten packets |
|---|---|
| 1 | 0.2 |
| 2 | 0.2 |
| 3 | 0.3 |
| 4 | 0.4 |
| 5 | 0.5 |
| 6 | 0.8 |
| 7 | 0.0[Δ] |

Indeed, those links exhibiting longer path length metrics also experience greater packet loss rates. That makes use of hosts with a lower metric even more attractive.

**Metric Value Considered by Interval**

Reliance on mean statistics can produce misleading results in the face of data with large variances. It is useful to study data based on statistical intervals as well.

[Δ] The number of samples with path length of 7000 is just 73. This outlier is probably too small of a sample to be significant for events that happen less than 10% of the time.

The traces used for this experiment have a path length distribution that is similar to the one exhibited by all the announced BGP routes.

| Path Length Metric (in thousands) | Percentage of Sample |
|---|---|
| 1 | 0.41 |
| 2 | 1.26 |
| 3 | 40.73 |
| 4 | 41.67 |
| 5 | 13.07 |
| 6 | 2.48 |
| 7 | 0.41 |

Using those distributions we can calculate the theoretical RTT ranges for our sample as distributed by path length metric.

| Path Length Metric (in thousands) | Percentage of samples with RTT within 2 standard deviations for corresponding PLM mean |
|---|---|
| 1 | 89.2 |
| 2 | 93.8 |
| 3 | 93.6 |
| 4 | 91.6 |
| 5 | 78.3 |
| 6 | 79.0 |
| 7 | 68.5 |

In order for the mean and variance measures to be statistically useful indicators, about 95% of data points should lie within 2 standard deviations of the mean. Our data comes close to that but falls a little short. This likely indicates our model is a reasonable approximation but is not completely appropriate for use with simulation studies.

**Metric Value Considered for Real Mirrors**

As a final test of the appropriateness of this heuristic, the active measurement test was conducted against six organizations currently utilizing mirrored web servers. All of these organizations use a manual system that requires the end user to select what they perceive to be the closest mirror.

Servers are listed grouped by metric. The grouping which contained the optimal metric is denoted with a + in the instance column. The grouping containing the least attractive server is marked with a -.

| Yahoo! | | | | |
|---|---|---|---|---|
| Metric | Instances | Mean RTT | Std Deviation | Loss per 10 |
| 3000 | 8 + | 104.8 | 69.4 | 0.4 |
| 5000 | 3 | 299.7 | 18.0 | 0.0 |
| 6000 | 7 - | 298.4 | 240.9 | 2.0 |

| Safesurf | | | | |
|---|---|---|---|---|
| Metric | Instances | Mean RTT | Std Deviation | Loss per 10 |
| 3000 | 11 + | 92.4 | 15.0 | 0.5 |
| 4000 | 11 - | 139.2 | 93.3 | 0.3 |
| 5000 | 1 | 113.0 | 12.0 | 4.0 |

| Redhat Distribution | | | | |
|---|---|---|---|---|
| Metric | Instances | Mean RTT | Std Deviation | Loss per 10 |
| 1000 | 1 | 58.0 | 71.0 | 0.0 |
| 2000 | 2 | 262.5 | 180.0 | 0.0 |
| 3000 | 8 | 202.4 | 39.1 | 0.5 |
| 4000 | 37 + | 311.9 | 85.3 | 0.6 |
| 5000 | 22 | 923.0 | 280.6 | 1.8 |
| 6000 | 8 - | 1133.4 | 278.1 | 3.5 |

| Microsoft | | | | |
|---|---|---|---|---|
| Metric | Instances | Mean RTT | Std Deviation | Loss per 10 |
| 3000 | 8 + | 133.9 | 21.1 | 0.8 |
| 4000 | 3 - | 283.3 | 53.0 | 0.0 |
| 5000 | 1 | 161.0 | 21.0 | 2.0 |
| 6000 | 1 | 362.0 | 8.0 | 0.0 |

| Linux Kernel Distribution | | | | |
|---|---|---|---|---|
| Metric | Instances | Mean RTT | Std Deviation | Loss per 10 |
| 3000 | 2 | 411.5 | 195.5 | 0.5 |
| 4000 | 29 + | 1259.8 | 114.5 | 3.5 |
| 5000 | 71 - | 2634.9 | 975.9 | 1.3 |
| 6000 | 4 | 1825.2 | 621.2 | 1.5 |

| NASA Jet Propulsion Laboratory | | | | |
|---|---|---|---|---|
| Metric | Instances | Mean RTT | Std Deviation | Loss per 10 |
| 4000 | 4 - | 357.2 | 43.0 | 1.2 |
| 5000 | 4 + | 261.0 | 124.0 | 0.0 |
| 6000 | 1 | 105.0 | 34.0 | 0.0 |

These results are very good. For all but one organization the servers comprising the lowest metric exhibit the best latency characteristics. For half of the cases the best server is also in the grouping with the best metric. For only one case would a bad decision be made, and the worst server would be selected 25% of the time.

## Other Considerations

Even if AS distance does not provide a good general purpose mechanism for all values of the heuristic distance measurement, it may be useful for distances of 1000 or less. These hosts are either located on the current network itself or on a directly peered network. Because long haul networks tend to have their own AS numbers, and contain very few hosts, any route involving these would need a path value of at least 2000 and would not be considered. While this scheme would not be able to provide a “closest mirrored server” service under such conditions, it would be able to keep all traffic that could be served locally from passing through needless exchange points.

Supplementary materials created during the course of this study are available to interested parties. This includes the source code to the Path Length Server and full logs of the data collected during the active measurement experiments. Contact the author for information.

## References


[CAIDA98] Claffy, K.C. "CAIDA Annual Report 1998", University of California, San Diego / San Diego Supercomputer Center, http://www.caida.org/Caida/annual98.html, October 1998.

[GUPTA99] S. Gupta, A.L.N. Reddy. "A Client Oriented, IP Level Redirection Mechanism". Department of Electrical Engineering, Texas A&M University, College Station, TX. Infocomm 1999.

[IDMAPS98] P. Francis, et al. "Internet Distance Maps (DIMaps)". Notes from IEPG 03/29./98. http://idmaps.eecs.umich.edu/

[IDMAPS99] P. Francis, et al. "An Architecture for a Global Internet Host Distance Estimation Service". MIT Software Laboratories, Tokyo. Infocomm 1999.

[PAXSON97] V. Paxson. "End-to-End Internet Packet Dynamics". Network Research Group, Lawrence Berkeley National Laboratory. University of California Berkeley. 1997.

[PAXSON98] V. Paxson. "On Calibrating Measurements of Packet Transit Times". Network Research Group, Lawrence Berkeley National Laboratory. University of California Berkeley. 1998

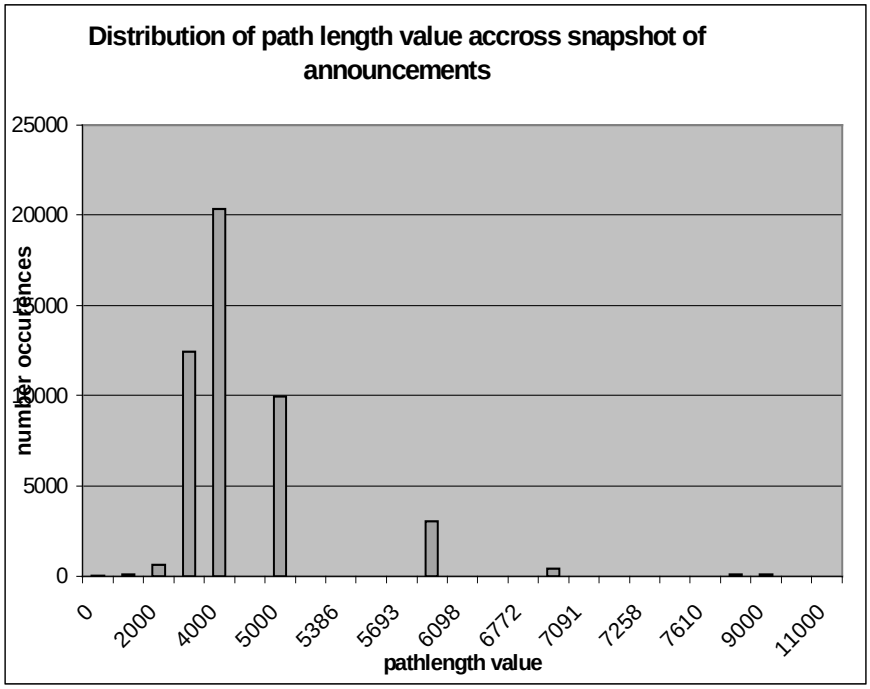
[RFC1519] V. Fuller, T. Li, J. Yu, K. Varadhan, "Classless Inter Domain Routing (CIDR): an Address Assignment and Aggregation Strategy", RFC 1519, BARRNet, cisco, MERIT, OARnet, September 1993.

[RFC1771] Y.Rekheter, T. Li, "A Border Gateway Protocol 4 (BGP-4)", RFC 1771, IBM Corp, cisco Systems, March 1995.

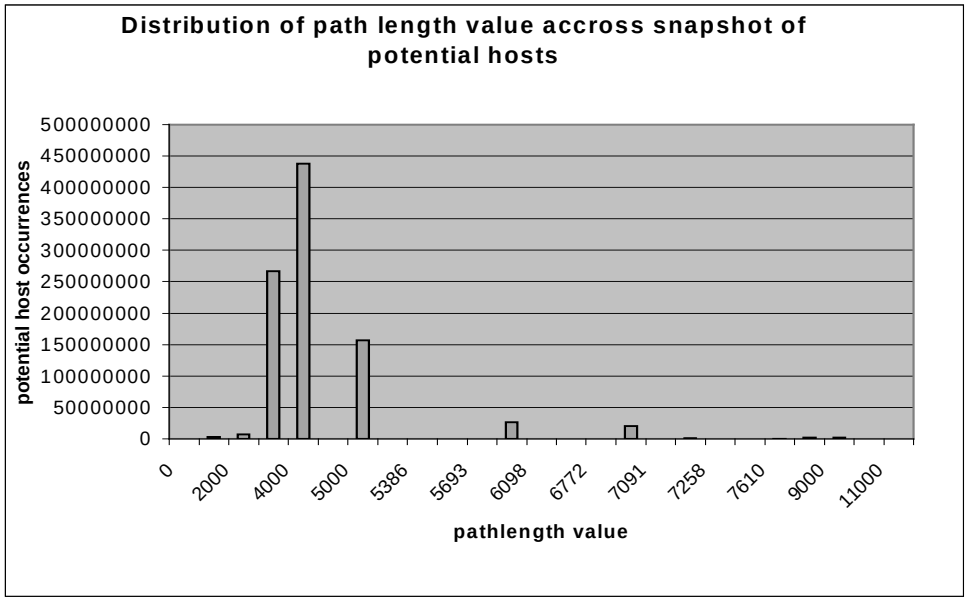
[RFC1965] P. Traina, "Autonomous System Confederations for BGP". RFC 1965, cisco Systems, June 1996.

[TCPCONG] M.Allman, V. Paxson, W. Stevens, "TCP Congestion Control". IETF Internet Draft draft-ietf-tcpimp-cong-control-03.txt, NASA Lewis, LBNL, Consultant, December 1998.

[ZHANG97] L. Zhang, S. Floyd, V. Jacobson. "Adaptive Web Caching". UCLA and LBNL. February 1997.




| value | occurences | P(value) | P(value repeated) |
|---|---|---|---|
| 0 | 32 | 0.0679 | 0.0000% |
| 1000 | 85 | 0.1804 | 0.0003% |
| 2000 | 643 | 1.3646 | 0.0186% |
| 3000 | 12426 | 26.3715 | 6.9546% |
| 4000 | 20329 | 43.1440 | 18.6140% |
| 4693 | 3 | 0.0064 | 0.0000% |
| 5000 | 9950 | 21.1167 | 4.4592% |
| 5098 | 1 | 0.0021 | 0.0000% |
| 5386 | 2 | 0.0042 | 0.0000% |
| 5609 | 1 | 0.0021 | 0.0000% |
| 5693 | 2 | 0.0042 | 0.0000% |
| 6000 | 3039 | 6.4496 | 0.4160% |
| 6098 | 1 | 0.0021 | 0.0000% |
| 6484 | 1 | 0.0021 | 0.0000% |
| 6772 | 1 | 0.0021 | 0.0000% |
| 7000 | 390 | 0.8277 | 0.0069% |
| 7091 | 2 | 0.0042 | 0.0000% |
| 7218 | 2 | 0.0042 | 0.0000% |
| 7258 | 1 | 0.0021 | 0.0000% |
| 7295 | 1 | 0.0021 | 0.0000% |
| 7610 | 1 | 0.0021 | 0.0000% |
| 8000 | 106 | 0.2250 | 0.0005% |
| 9000 | 92 | 0.1953 | 0.0004% |
| 10000 | 4 | 0.0085 | 0.0000% |
| 11000 | 4 | 0.0085 | 0.0000% |
| total announcements | 47119 | 100.0000 | 30.4705% |




| value | occurences | P(value) | P(value repeated) |
|---|---|---|---|
| 0 | 80696 | 0.0087 | 0.0000% |
| 1000 | 2902436 | 0.3137 | 0.0010% |
| 2000 | 6816630 | 0.7368 | 0.0054% |
| 3000 | 267033660 | 28.8622 | 8.3303% |
| 4000 | 437699136 | 47.3085 | 22.3810% |
| 4693 | 50176 | 0.0054 | 0.0000% |
| 5000 | 157106432 | 16.9808 | 2.8835% |
| 5098 | 32768 | 0.0035 | 0.0000% |
| 5386 | 17408 | 0.0019 | 0.0000% |
| 5609 | 65536 | 0.0071 | 0.0000% |
| 5693 | 131072 | 0.0142 | 0.0000% |
| 6000 | 26534960 | 2.8680 | 0.0823% |
| 6098 | 65536 | 0.0071 | 0.0000% |
| 6484 | 4096 | 0.0004 | 0.0000% |
| 6772 | 65536 | 0.0071 | 0.0000% |
| 7000 | 20474624 | 2.2130 | 0.0490% |
| 7091 | 131072 | 0.0142 | 0.0000% |
| 7218 | 1048576 | 0.1133 | 0.0001% |
| 7258 | 65536 | 0.0071 | 0.0000% |
| 7295 | 65536 | 0.0071 | 0.0000% |
| 7610 | 262144 | 0.0283 | 0.0000% |
| 8000 | 2424064 | 0.2620 | 0.0007% |
| 9000 | 1925888 | 0.2082 | 0.0004% |
| 10000 | 131584 | 0.0142 | 0.0000% |
| 11000 | 66304 | 0.0072 | 0.0000% |
| total potential hosts | 925201406 | 100.0000 | 33.7336% |